Lineare Algebra

Skriptum zur Vorlesung von Prof. Dr. Victor Bangert
WS 2002/03 und SS 2003

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:

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

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:

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

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:

$$\mathrm{I} \quad \begin{array}{rcll} 2x+4y & = & 8 & \mid \cdot\left(-\frac{3}{2}\right) \\ 3x-6y & = & 4 & \hookleftarrow\! \mid \end{array}$$

$$\mathrm{II} \quad \begin{array}{rcll} 2x+\ \ 4y & = & 8 & \\ 0x-12y & = & -8 & \mid :(-12) \end{array}$$

$$\mathrm{III} \quad \begin{array}{rcl} 2x+4y & = & 8 \\ y & = & \frac{2}{3} \end{array} \qquad 2x+\frac{8}{3}=8 \qquad 2x=\frac{16}{3} \qquad x=\frac{8}{3}$$

$$\mathrm{IV} \quad \begin{array}{rcl} x & = & \frac{8}{3} \\ y & = & \frac{2}{3} \end{array}$$

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

4) 2 Gleichungen, 2 Unbekannte:

$$\text{I}\qquad \begin{array}{rcl} 2x+4y &=& 8 \quad \Big| \cdot\left(-\frac{3}{2}\right) \\ 3x+6y &=& 4 \quad \hookleftarrow\Big| \end{array}$$

$$\text{II}\qquad \begin{array}{rcl} 2x+4y &=& 8 \\ 0x+0y &=& -8 \end{array}$$

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).

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

$$\text{II}\qquad \begin{array}{rcl} 2x_1+4x_2+4x_3+6x_4 &=& 2 \\ x_2 - \ x_3 \qquad &=& 1 \quad \Big| \ (-1) \\ x_2 - 3x_3 + 2x_4 &=& 0 \quad \hookleftarrow\Big| \end{array}$$

$$\text{III}\qquad \begin{array}{rcrl} 2x_1+4x_2+4x_3+6x_4 &=& 2 & \text{III}_1 \\ x_2 - \ x_3 \qquad &=& 1 & \text{III}_2 \\ -2x_3 + 2x_4 &=& -1 & \text{III}_3 \end{array}$$

Gleichungssystem in Stufenform
("rekursiv auflösbar")

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

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

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

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 $M = \{x \in \mathbb{N} \mid x \text{ ist ungerade natürliche Zahl kleiner als } 9\}$.

Beispiele von Mengen: $\mathbb{N} = \{x \in \mathbb{R} \mid x \text{ natürliche Zahl}\} = \{0, 1, 2, 3, \ldots\}$
$\mathbb{Z} = \{x \in \mathbb{R} \mid x \text{ ganze Zahl}\} = \{0, 1, -1, 2, -2, \ldots\}$

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 <u>Teilmenge einer Menge</u> 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: $\{x \mid x \text{ ungerade natürliche Zahl kleiner } 9\} = \{1, 3, 5, 7\}$

Konstruktionen mit Mengen A und B:

(1.10) <u>Schnittmenge</u> (Durchschnitt) $A \cap B$ <u>von</u> A <u>und</u> B:

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

<u>Vereinigung</u> $A \cup B$ <u>von</u> A <u>und</u> B:

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

<u>Komplement</u> $A \setminus B$ <u>von</u> B <u>in</u> 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:

$$
\begin{array}{llcll}
\text{(a)} & A \cup B & = & B \cup A,\ A \cap B = B \cap A & \text{“Kommutativgesetze”} \\[1ex]
\text{(b)} & (A \cup B) \cup C & = & A \cup (B \cup C) & \text{“Assoziativgesetze”} \\
 & (A \cap B) \cap C & = & A \cap (B \cap C) & \\[1ex]
\text{(c)} & A \cap (B \cup C) & = & (A \cap B) \cup (A \cap C) & \text{“Distributivgesetze”} \\
 & A \cup (B \cap C) & = & (A \cup B) \cap (A \cup C) & \\[1ex]
\text{(d)} & A \setminus (B \cup C) & = & (A \setminus B) \cap (A \setminus C) & \text{“De Morgansche Regeln”} \\
 & A \setminus (B \cap C) & = & (A \setminus B) \cup (A \setminus C) &
\end{array}
$$

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 $a_1 = a'_1, \ldots, a_n = a'_n$ 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:

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

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:*

Ist $x = (x_1, \ldots, x_n) \in L_I$ *und* $y = (y_1, \ldots, y_n) \in L_I$, *so gilt* $x + y \in L_I$.
Ist $x = (x_1, \ldots, x_n) \in L_I$ *und* $r \in \mathbb{R}$, *so gilt* $rx \in L_I$.

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* $\overline{x} = (\overline{x}_1, \dots, \overline{x}_n)$ *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 $y - \overline{x} := y + (-1)\overline{x}$).

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

Subtraktion dieser Gleichungen liefert:

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

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 $i \in \{1, \dots, m\}$:

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

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

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

Addition dieser Gleichungen liefert:

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

also $y + x \in L_I$.

Zurück zur Mengenlehre:

Abbildungen zwischen Mengen

(1.17) Def.: Es seien M, N Mengen. Eine <u>Abbildung</u> $f : M \to N$ <u>von</u> M <u>nach</u> 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):

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

Beispiel zu (b):

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

(1.21) Def.: Ist $f : M \to N$ Abbildung und $A \subseteq M$, so ist die <u>Einschränkung</u> (Restriktion) $f|A (= f|_A) : A \to N$ <u>von</u> f <u>auf</u> 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 $\begin{aligned} f^{-1}(y) &= f^{-1}(f(x)) = (f^{-1} \circ f)(x) \\ &= \mathrm{id}_M(x) = x \text{ gelten.} \end{aligned}$

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

(1.23) Def.: Seien M, N Mengen. Eine <u>Relation auf</u> (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 <u>Diagonale</u> in $M \times M$.
2) Sei $f : M \to N$ Abbildung. Dann ist der <u>Graph von</u> 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

$$
\begin{array}{rcl}
[x]_R & := & \{y \in M \mid (x,y) \in R\} \text{ die Äquivalenzklasse von } x \text{ und} \\
M/R & := & \{[x]_R \mid x \in M\} \text{ die Menge der Äquivalenzklassen, die auch "}M\text{ modulo }R\text{"} \\
& & \text{genannt wird.}
\end{array}
$$

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

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

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

Für diese Äquivalenzrelation ist folgende Bezeichnung gebräuchlich:

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

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 (mod 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)(i) $\Rightarrow -(-a) = a$
(2.3)(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 $1 \in K$ 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: $0 \cdot a + a = 0 \cdot a + 1 \cdot a \overset{(2.4)(\text{iii})}{=} (0+1)a = 1 \cdot a = a \quad | -a \Rightarrow 0 \cdot a = 0$
(ii) $(-a)b + ab \overset{(2.4)(\text{iii})}{=} ((-a)+a)b = 0 \cdot b \overset{(\text{i})}{=} 0$
(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:

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

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

a der Realteil von $z, a =: \mathrm{Re}(z)$
b der Imaginärteil von $z, b =: \mathrm{Im}(z)$

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.

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) $1v = v \quad$ (, 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 homogenes 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: $v \in U \overset{\text{(ii)}}{\Rightarrow} (-1)v \in U \overset{(3.2)\text{(iii)}}{\Rightarrow} -v \in U$

$U \neq \emptyset \Rightarrow \exists v \in U \Rightarrow v$ 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: $\alpha_1 v_1 + \alpha_s v_2 + \ldots + \alpha_k v_k = \sum_{i=1}^{k} \alpha_i v_i$

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 $\alpha \in K, v = \sum_{i=1}^{k} \alpha_i v_i \in U_0$ 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: $\alpha \sum_{i=1}^{k} \alpha_i v_i = \sum_{i=1}^{k} \alpha\alpha_i v_i$).
Also ist U_0 ein M enthaltender Unterraum. Nach Definition (3.6) folgt $\mathrm{span}(M) \subseteq U_0$.
2) Zeige $U_0 \subseteq \mathrm{span}(M)$. Seien $v_1, \ldots, v_k \in M$, $\alpha_1, \ldots, \alpha_k \in K$. Da $\mathrm{span}(M)$ Unterraum ist, folgt nach (3.3)(iii): $\alpha_1 v_1 \in \mathrm{span}(M), \ldots, \alpha_k v_k \in \mathrm{span}(M)$, 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 $e_1 := (1, 0, \ldots, 0) \in K^n$, $e_2 := (0, 1, 0, \ldots, 0) \in K^n$, $\ldots$, $e_n = (0, \ldots, 0, 1) \in K^n$. 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: $\alpha v = \underline{0}, v \neq \underline{0} \overset{(3.2)(\mathrm{ii})}{\Rightarrow} \alpha = 0$.
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.

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

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

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

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

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

Bsp.: $\{e_1, \ldots, e_n\} \subseteq K^n$ 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

$$\begin{array}{rcl} r \quad\; + t & = & 0 \quad | \; - \\ s + t & = & 0 \quad | \\ r + s - t & = & 0 \; \leftarrow| \end{array}$$

Wir lösen es nach dem Gaußschen Eliminationsverfahren:

$$\begin{array}{rcl} r \quad\; + \; t & = & 0 \\ s + \; t & = & 0 \quad | \; - \\ s - 2t & = & 0 \; \leftarrow| \end{array}$$

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

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 $k \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_k \in K$, $v_1, \ldots, v_k \in M$: $v \overset{(*)}{=} \sum_{i=1}^{k} \alpha_i v_i$.

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 $v \in M' \setminus M$. Wegen $\operatorname{span}(M) = V$ (M Basis!) gilt $v \in \operatorname{span}(M) \setminus M$. 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, $v \in M \setminus M''$. 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 $\text{span}(B) = V$ 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 für jedes solche 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 $v_j \in \text{span}(B')$, 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

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

d.h. B' ist Erzeugendensystem von V. Um die lineare Unabhängigkeit von B' zu zeigen, nehmen wir an, daß für $\beta \in K$, $\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_n \in K$ 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 $B = \{v_1, \dots, v_n\}$ 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 $B = \{v_1, \dots, v_n\}$ eine Basis von V mit n Elementen und seien $w_1, \dots, w_m$ m linear unabhängige Vektoren in V. Dann gilt*

(i) $m \leq n$
(ii) *Es gibt $(n-m)$ Vektoren in B, bei geeigneter Numerierung $v_{m+1}, \dots, v_n$, so daß $\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$ 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 $w_1, \dots, w_m$ linear unabhängig gegeben. Die Induktionsvoraussetzung impliziert: $(m-1) \leq n$ und es existieren $n - m + 1$ Vektoren in B, o.E. $v_m \dots, v_n$, so daß $\{w_1, \dots, w_{m-1}, v_m \dots, v_n\}$ eine Basis von V ist.

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

Zeige (ii): Wir wollen in der Basis $\{w_1, \dots, w_{m-1}, v_m, \dots, v_n\}$ eines der v_i, $m \leq i \leq n$, gegen w_m "austauschen" und zwar so, daß wieder eine Basis entsteht.

Da $\mathrm{span}\{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\} = V$ 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. $U = \mathrm{span}(B') = V$.

(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 <u>komplementärer Unterraum</u>)
(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 $U_1 = U_{12} \oplus W_1$, $U_2 = U_{12} \oplus W_2$:

$$\dim U_1 = \dim U_{12} + \dim W_1$$
$$\dim U_2 = \dim U_{12} + \dim W_2$$

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

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

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 $\underline{k\text{-dimensionaler}}$ $\underline{\text{affiner}}$ $\underline{\text{Unterraum}}$, 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 $\xrightarrow{\text{Kodierung}}$ endl. Folge der Länge n von $0, 1$ $\xrightarrow{\text{Übermittlung}}$ fehlerhafter Folge von $0, 1$ $\xrightarrow{\text{automat. Fehlerkorrektur}}$ ursprüngl. Folge von $0, 1$ $\xrightarrow{\text{Dekodierung}}$ 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 $F : \{0, 1\}^n \to \{0, 1\}^p$ 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 $\approx n^2 p$ $(= 31^2 \cdot 5 \approx 5 \cdot 10^3) \ll 2^n \approx 2 \cdot 10^9)$

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 demselben Körper K.

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

(i)	$L(v_1 + v_2)$	$=$	$L(v_1) + L(v_2)$	(L ist "additiv")
(ii)	$L(\alpha v)$	$=$	$\alpha L(v)$	(L ist "homogen")

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

Ein Homomorphismus $L : V \to W$ heißt Isomorphismus, falls L bijektiv ist. Zwei K-Vektorräume V und W heißen isomorph, 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, $\mathrm{im}\, L$ 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 $L \in \mathrm{Hom}(V, W)$. 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 $L \in \mathrm{Hom}(K^n, K^k)$:

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

Das Problem, für eine gegebene rechte Seite $b = (b_1, \ldots, b_k) \in K^k$ 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 $b = (b_1, \ldots, b_k)$ 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 $\stackrel{(4.3)}{\Rightarrow} L(B)$ ist linear unabhängig.
$\mathrm{span}\, L(B) \stackrel{(4.4)}{=} L(\mathrm{span}\, B) = L(V) \stackrel{L \text{ surjektiv}}{=} W \Rightarrow L(B)$ 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 $L(v_1), \dots, L(v_k)$ 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.
$L(V) = L \,\mathrm{span}\, B) \stackrel{(4.4)}{=} \mathrm{span}\, L(B) = W \Rightarrow L$ ist surjektiv.

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

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

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

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

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 $w \in \operatorname{im} L$ 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 $L|U : U \to \operatorname{im} L$ 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: $L_{I^{\text{hom}}} = \ker L$ und

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

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

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

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 $(b_1, \ldots, b_n)$ lösbar $\Leftrightarrow L_{I^{\text{hom}}} = \{0\} \Leftrightarrow I$ besitzt für jede rechte Seite $(b_1, \ldots, b_n)$ genau eine Lösung.

Ist $L \in \operatorname{Hom}(V, W)$, $\alpha \in K$, 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: $\alpha L \in \operatorname{Hom}(V, W)$, 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 $L_1, L_2 \in \operatorname{Hom}(V, W)$, 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: $L_1 + L_2 \in \operatorname{Hom}(V, W)$, 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 $\operatorname{Hom}(V, W)$ ein K-Vektorraum.*

Bew.: Es muß nachgeprüft werden, daß $(\operatorname{Hom}(V, W), +)$ 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 $\operatorname{Hom}(V, W)$ ist die Abbildung $L \in \operatorname{Hom}(V, W)$, die jedes $v \in V$ auf $\underline{0} \in W$ abbildet, genannt die 0-Abbildung. Das additive Inverse zu $L \in \operatorname{Hom}(V, W)$ ist $-L \in \operatorname{Hom}(V, W)$, definiert durch $(-L)(v) := -L(v)$.

Um (3.1)(iii) zu zeigen, seien $\alpha \in K$, $L_1, L_2 \in \operatorname{Hom}(V, W)$. 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 <u>Matrix von L bezüglich $\mathcal{G}$ und $\mathcal{G}'$</u>,

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

(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 $j \in \{1, \dots, n\}$. Aus (4.6) folgt dann: $L_1 = L_2$.

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

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

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.

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

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

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

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

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

(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* $c_{hj} = \sum\limits_{i=1}^{m} a_{hi}b_{ij}$ *für* $1 \leq h \leq k$, $1 \leq j \leq n$.

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.: $b_1(a_{11}v_1 + a_{21}v_2) + b_2(a_{12}v_1 + a_{22}v_2) = (b_1a_{11} + b_2a_{12})v_1 + (b_1a_{21} + b_2a_{22})v_2$.

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

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

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: $A = \mathrm{Mat}_{\mathcal{G}_3}^{\mathcal{G}_4}(L_3), B = \mathrm{Mat}_{\mathcal{G}_2}^{\mathcal{G}_3}(L_2), C = \mathrm{Mat}_{\mathcal{G}_1}^{\mathcal{G}_2}(L_1)$
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:

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

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

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

1. Spezialfall: $V' = V$

Ein Homomorphismus $L : V \to V$ heißt Endomorphismus, $\mathrm{End}(V) := \mathrm{Hom}(V, V)$. 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, man schreibt $\mathrm{Aut}(V) := \{L \mid L \in \mathrm{End}(V) \text{ bijektiv}\}$. $(\mathrm{Aut}(V), \circ)$ 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 $P := \mathrm{Mat}^{\mathcal{G}}_{\overline{\mathcal{G}}}(\mathrm{id}_V)$. Dann gilt $P \in \mathrm{GL}_n(K)$, $P^{-1} = \mathrm{Mat}^{\overline{\mathcal{G}}}_{\mathcal{G}}(\mathrm{id}_V)$ und

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

Bew.: $\mathrm{Mat}^{\overline{\mathcal{G}}}_{\mathcal{G}}(\mathrm{id}_V) \, \mathrm{Mat}^{\mathcal{G}}_{\overline{\mathcal{G}}}(\mathrm{id}_V) = \mathrm{Mat}^{\overline{\mathcal{G}}}_{\overline{\mathcal{G}}}(\mathrm{id}_V) = E_n \Rightarrow P \in \mathrm{GL}_n(K)$ und $P^{-1} = \mathrm{Mat}^{\overline{\mathcal{G}}}_{\mathcal{G}}(\mathrm{id}_V)$. 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: $\begin{array}{rcl} V & = & K^n \text{ mit Basis } \mathcal{G} = (e_1, \ldots, e_n) \\ V' & = & K^m \text{ mit Basis } \mathcal{G}' = (e_1, \ldots, e_m) \end{array}$

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 $1 \le j \le n$.

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

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

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*

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

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

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

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.

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

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):

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

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

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 $1 \leq i \leq n$. 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 $(\mathrm{En} \mid \overline{\overline{B}})$, 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):

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

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

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

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

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

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

Explizite Lösung:

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

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

Zu (4.26)

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

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

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

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

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

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

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

5. Die Determinante

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

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

das von $v_1, \dots, v_n$ 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 $(v_1, \dots, v_n)$ von V eine reelle Zahl $(\neq 0)$ zuordnet, deren Betrag wir als Definition für das Volumen von $P(v_1, \dots, v_n)$ nehmen wollen. (Es stellt sich heraus, daß die Funktion D, die positive und negative Werte annimmt, das mathematisch natürliche Objekt ist. Das Vorzeichen von $D(v_1, \dots, v_n)$ sagt etwas über die "Orientierung" der geordneten Basis $(v_1, \dots, v_n)$ 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) $(v_1, \dots, v_n)$ Basis $\Rightarrow D(v_1, \dots, v_n) \neq 0$.

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

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

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

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

Dabei hat Forderung (V_2) die anschauliche Interpretation, daß das Volumen des Parallotops $P(v_1, \dots, v_{i-1}, rv_i, v_{i+1}, \dots, v_n)$ das $|r|$-fache des Volumens von $P(v_1, \dots, v_n)$ 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 $i \in \{1, \dots, k\}$ und alle $v_1, \dots, v_{i-1}, v, w, v_{i+1}, \dots, v_k \in V$, $\alpha, \beta \in K$ gilt:

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

(ii) alternierend (oder schiefsymmetrisch), falls für alle $1 \leq i < j \leq k$ 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) $\forall 1 \leq i < j \leq k$: $f(\ldots, \overset{i}{v}, \ldots, \overset{j}{w}, \ldots) = -f(\ldots, \overset{i}{w}, \ldots, \overset{j}{v}, \ldots)$

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

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

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 $v_i = \sum\limits_{\substack{j=1 \\ j \neq i}}^{k} \alpha_j v_j$. Dann:

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

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

Sind $a_1, \ldots, a_k$ 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, $f(v_1, \ldots, v_n) = \alpha$

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

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

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

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

(iii) Existenz: Definiere f durch $(*)$. Man sieht leicht, daß f n-linear ist. Außerdem gilt wegen $v_j = \sum_{i=1}^n \delta_{ij} v_i$ (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 $w_k = \sum_{l=1}^n a_{lk} v_l$, $1 \leq k \leq n$, so folgt aus $w_i = w_j$:

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

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) $\Leftrightarrow \{f \mid f$ alternierende n-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 $(v_1, \dots, v_n)$ Basis von V. Dann folgt aus (5.10) ("Speziell:") und $D \neq 0$, daß $D(v_1, \dots, v_n) \neq 0$ gilt. Ist $D(v_1, \dots, v_n) \neq 0$, so folgt aus (5.3)(ii), daß die $v_1, \dots, v_n$ linear unabhängig sind, also - wegen $\dim V = n$ - eine Basis von V.

(ii) Sei $(v_1, \dots, v_n)$ eine Basis von V und $\alpha := \frac{f(v_1,\dots,v_n)}{D(v_1,\dots,v_n)}$. Dann sind f und αD alternierende n-Linearformen, die auf der Basis $(v_1, \dots, v_n)$ 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 $(e_1, \dots, e_n)$, die - nach (5.10) - eine "Standarddeterminantenform" $D_0 : K^n \times \dots \times K^n \to K$ durch

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

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

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

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

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

Speziell: $n = 1: \quad A = (a_{11}) \to \quad \det A = a_{11}$
$\qquad\qquad n = 2: \quad A = \begin{smallmatrix} a_{11} & a_{12} \\ a_{21} a_{22} & \end{smallmatrix} \to \quad \det A = a_{11}a_{22} - a_{12}a_{21}$

$n = 3$: 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) $\det E_n = D_0(e_1, \dots, e_n) = 1$.
(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):

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

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.

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

(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.

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

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) $(A^T = (\tilde{a}_{ji})_{\substack{1\leq j\leq n\\ 1\leq i\leq n}}$ 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.: $a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(n)n} \neq 0 \Rightarrow \forall i \in \{1,\ldots,n\} : \sigma(i) \leq i \Rightarrow \sigma = \mathrm{id} \Rightarrow$ Beh.

Bsp.:

$$\begin{vmatrix} 1 & 2 & 3\\ 4 & 8 & 6\\ 7 & 8 & 10\end{vmatrix}\begin{array}{r|r|} -4 & -7\\ \leftarrow & \\ & \leftarrow\end{array} = \begin{vmatrix} 1 & 2 & 3\\ 0 & 0 & -6\\ 0 & -6 & -11\end{vmatrix} = -\begin{vmatrix} 1 & 2 & 3\\ 0 & -6 & -11\\ 0 & 0 & -6\end{vmatrix} = -36$$

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) $\operatorname{vol}_D(P)$ eines Parallelotops $P = P(v_1,\ldots,v_n) \subseteq V$ durch $\operatorname{vol}_D(P) = |D(v_1,\ldots,v_n)|$. 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:

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

(5.19) Satz und Def. *Sei V n-dimensionaler K-Vektorraum und $L \in \operatorname{End}(V)$. 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:

$$\begin{aligned} \det(L_2 \circ L_1) &= \frac{D(L_2(L_1(v_1)), \ldots, L_2(L_1(v_n)))}{D(v_1, \ldots, v_n)} \\ &= \frac{D(L_2(L_1 v_1), \ldots, L_2(L_1(v_n)))}{D(L_1(v_1), \ldots, L_1(v_n))} \cdot \det L_1 \\ &= \det L_2 \cdot \det L_1 \end{aligned}$$

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 $L \in \mathrm{End}(K^n)$ und $\mathrm{Mat}(L) = A$, d.h. "$L(x) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$".

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 $L \in \mathrm{End}(V)$.

Dann gilt: $\det L = \det(\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L))$.

Bew.:

(i) $\mathrm{Mat}(L) = A \Leftrightarrow L(e_j) = A_j \Rightarrow$

$$\det L = \frac{D_0(L(e_1), \ldots, L(e_n))}{D_0(e_1, \ldots, e_n)} = D_0(A_1, \ldots, A_n) = \det A.$$

(ii) Sei $J \in \mathrm{Hom}(K^n, V)$ der Isomorphismus mit $J(e_j) = v_j$ für $j \in \{1, \ldots, n\}$, vgl. (4.5). Dann gilt $\mathrm{Mat}(J^{-1} \circ L \circ J) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$, da $J^{-1} \circ L \circ J(e_j) = \sum_{i=1}^{n} a_{ij} \underbrace{e_i}_{=J^{-1}(v_i)} \Leftrightarrow$

$L(v_j) = \sum_{i=1}^{n} a_{ij} v_i$. 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

$$\begin{aligned} \det L &= D(L(v_1), \ldots, L(v_n)) = D_0(J^{-1} \circ L(v_1), \ldots, J^{-1} \circ L(v_1)) \\ &= D_0(J^{-1} \circ L \circ J(e_1), \ldots, J^{-1} \circ L \circ J(e_n)) \\ &= \det(J^{-1} \circ L \circ J) \overset{\text{(i)}}{=} \det(\mathrm{Mat}(J^{-1} \circ L \circ J)) = \det(\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)). \end{aligned}$$

Bem.: Es gilt i.a. <u>nicht</u> $\det L = \det(\mathrm{Mat}_{\mathcal{G}}^{\overline{\mathcal{G}}}(L))$, 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) $A \in \mathrm{GL}_n(K) \Rightarrow \det(A^{-1}) = (\det A)^{-1}$

Speziell besagt (5.22): Die Abbildung $A \to \det A$ ist ein Gruppenhomomorphismus von $(\mathrm{GL}_n(K), \cdot)$ auf $(K \setminus \{0\}, \cdot)$. $\mathrm{SL}_n(K) = \{A \in \mathrm{GL}_n(K) | \det A = 1\}$ ist Untergruppe von $\mathrm{GL}_n(K)$, die <u>spezielle lineare Gruppe</u>.

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.: $\det(L-\lambda\,\mathrm{id}_V)=0 \overset{(5.20)(ii)}{\Leftrightarrow} (L-\lambda\,\mathrm{id}_V)\notin \mathrm{Aut}(V) \overset{(4.8)}{\Leftrightarrow}$
$\ker(L-\lambda\,\mathrm{id}_V)\neq\{0\} \Leftrightarrow \exists v\in V\setminus\{0\}:(L-\lambda\,\mathrm{id}_V)(v)=0$
$\Leftrightarrow \exists v\in V\setminus\{0\}: L(v)=\lambda v.$

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 $\lambda)$.

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 $L \in \mathrm{Aut}(\mathbb{R}^n)$, $\det L > 0$, 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

$$\mathrm{I} \qquad \begin{array}{llll} a_{11}x_1 & + \ldots + & a_{1n}x_n = & b_1 \\ \vdots & & & \vdots \\ a_{n1}x_1 & + \ldots + & a_{nn}x_n = & b_1 \end{array} \qquad \text{oder kurz:} \quad A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = b$$

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.: $\det A \neq 0 \overset{(5.20)(\mathrm{ii})}{\Leftrightarrow} x \in \mathbb{R}^n \to Ax \in \mathbb{R}^n$ 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

$$x_1A_1 + x_2A_2 + \ldots + x_nA_n = b \quad (\text{vgl. I!})$$

Also gilt:

$$\begin{aligned} D_0(A_1, \ldots, A_{k-1}, b, A_{k+1}, \ldots, A_n) &= D_0(A_1, \ldots, A_{k-1}, \sum_{i=1}^n x_iA_i, A_{k+1}, \ldots, A_n) \\ &\overset{D_0\ n\text{-linear}}{=} \sum_{i=1}^n x_i D_0(A_1, \ldots, A_{k-1}, A_i, A_{k+1}, \ldots, A_n) \\ &\overset{D_0 \text{ alternierend}}{=} x_k \det A. \end{aligned}$$

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 $A := \begin{pmatrix} 1 & 0\ldots 0 \\ 0 & \\ \vdots & \tilde{A} \\ 0 & \end{pmatrix} \in K^{n\times n}$, 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* $A \in K^{n\times n}$, $j \in \{1, \ldots, n\}$ *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 $i \in \{1, \ldots, n\}$ *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.: $A_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, A_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$ seien die Spaltenvektoren von A.

$$\begin{aligned} \det A &= D_0(A_1, \dots, A_n) = D_0(A_1, \dots, A_j = \sum a_{ij} e_i, \dots, A_n) \\ &= \sum_{i=1}^{n} a_{ij} D_0(A_1, \dots, A_{j-1}, e_i, A_{j+1} \dots, A_n) \end{aligned}$$

Es bleibt zu zeigen: $D_0(A_1, \dots, A_{j-1}, e_i, A_{j+1} \dots, A_n) = (-1)^{i+j} \det A_{ij}$

$D_0(A_1, \dots, A_{j-1}, e_i, A_{j+1} \dots, A_n) = (-1)^{j-1} D_0(e_i, A_1, \dots, A_{j-1}, A_{j+1}, \dots, A_n)$

$$= \quad (-1)^{j-1} \begin{vmatrix} 0 & a_{11} & \dots & \dots & a_{1n} \\ \vdots & & & & \vdots \\ 1 & a_{i1} & \dots & \dots & a_{in} \\ \vdots & & & & \vdots \\ 0 & a_{n1} & \dots & \dots & a_{nn} \end{vmatrix} = (-1)^{i+j-2} \begin{vmatrix} 1 & a_{i1} & \dots & \dots & a_{in} \\ 0 & a_{11} & \dots & \dots & a_{1n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{i-1,1} & \dots & \dots & a_{i-1,n} \\ 0 & a_{i+1,1} & \dots & \dots & a_{i+1,n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{n1} & \dots & \dots & a_{nn} \end{vmatrix}$$

$$= \quad (-1)^{i+j} \begin{vmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & a_{11} & \dots & \dots & a_{1n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{i-1,1} & \dots & \dots & a_{i-1,n} \\ 0 & a_{i+1,1} & \dots & \dots & a_{i+1,n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{n1} & \dots & \dots & a_{nn} \end{vmatrix}$$

$$\overset{(5.27)}{=} \quad (-1)^{i+j} \quad \det A_{ij}$$

(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 <u>euklidischer Vektorraum</u>.

Bsp.: 1) $V = \mathbb{R}^n$,

$$\begin{aligned} b(x,y) &:= \langle x, y\rangle := \sum_{i=1}^n x_i y_i \\ b(x,x) &= \sum_{i=1}^n x_i^2 \end{aligned}$$

2)

$$\begin{aligned} V = C^0([0,2\pi],\mathbb{R}) &= \{f \mid f : [0,2\pi] \to \mathbb{R} \text{ stetig}\} \\ b(f_1,f_2) := \langle f_1, f_2\rangle_{L^2} &= \int_0^{2\pi} f_1(t) f_2(t)\, dt \end{aligned}$$

(6.2) Def.: Sei $V, \langle , \rangle$ euklidischer Vektorraum. Dann heißt $\|v\| := \sqrt{\langle v, v\rangle} \geq 0$ die <u>Norm</u> (oder <u>Länge</u>) von $v \in V$ (bzgl. $\langle , \rangle$) und $\|v - w\|$ der <u>Abstand</u> 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 orthogonal (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: $v_1, \ldots, v_n \in S$ verschieden, $r_1, \ldots, r_n \in \mathbb{R}$ 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) $V = \mathbb{R}^n$, $\langle x,y\rangle := \sum\limits_{i=1}^{n} x_i y_i \Rightarrow \{e_1, \ldots, e_n\}$ 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 $B = (v_1, \ldots, v_n)$* 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 <u>vollständiges</u> ONS ist), aber das können wir hier nicht beweisen.

Bew.: (i) $0 \leq \| v - \sum\limits_{e \in E} \langle v, e\rangle e \| = \|v\|^2 - 2 \sum\limits_{e \in E} \langle v, e\rangle^2 + \sum\limits_{e \in E} \langle v, e\rangle^2$

(ii) $(v_1, \ldots, v_n)$ Basis $\Rightarrow \exists r_1, \ldots, r_n \in \mathbb{R} : v = \sum\limits_{j=1}^{n} r_j v_j \Rightarrow \langle v, v_i\rangle = \sum\limits_{j=1}^{n} r_j \langle v_j, v_i\rangle = r_i$

(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*

$$\operatorname{span}\{v_1, \ldots, v_i\} = \operatorname{span}\{\tilde{v}_1, \ldots, \tilde{v}_i\}$$

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

$$\operatorname{span}\{v_1, \ldots, v_i\} = \operatorname{span}\{\tilde{v}_1, \ldots, \tilde{v}_i\} \text{ für alle } 1 \leq i \leq k-1.$$

Setze $w_k = v_k - \sum\limits_{j=1}^{k-1} \langle v_k, \tilde{v}_j\rangle \tilde{v}_j$ $\quad(\Rightarrow \langle w_k, \tilde{v}_i\rangle = 0$ 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 $\operatorname{span}\{\tilde{v}_1, \ldots, \tilde{v}_k\} = \operatorname{span}\{v_1, \ldots, v_k\}$.

(6.8) Def. Seien $V, \langle,\rangle$ und $\tilde{V}, \tilde{\langle,\rangle}$ euklidische Vektorräume. Ein $L \in \operatorname{Hom}(V, \tilde{V})$ heißt orthogonal(bzgl. $\langle,\rangle$ und $\tilde{\langle,\rangle}$), falls für alle $v, w \in V$ gilt: $\langle v, w\rangle = \widetilde{\langle L(v), L(w)\rangle}$. 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 $f_1, \dots, f_n$ ONB von V. Definiere L durch $L(f_i) = e_i \in \mathbb{R}^n$. Ist $v = \sum r_i f_i$ und $w = \sum s_j f_s$, so folgt $\langle v, w\rangle_V = \sum r_i s_i$. Daraus folgt $L(v) = \sum r_i e_i$, $L(w) = \sum s_j e_j$ und $\langle L(v), L(w)\rangle = \sum r_i s_i = \langle v, w\rangle_V$.

(6.10) Satz. *Sei* $L \in \mathrm{End}(V)$ *und* $\mathcal{G} = (e_1, \dots, e_n)$ 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.:
$$\begin{aligned} \delta_{ij} = \langle e_i, e_j\rangle &= \langle L(e_i), L(e_j)\rangle = \left\langle \sum_{k=1}^n a_{ki} e_k, \sum_{l=1}^n a_{lj} e_l \right\rangle \\ &= \sum_{k,l=1}^n a_{ki} a_{lj} \delta_{kl} = \sum_{k=1}^n a_{ki} a_{kj} \end{aligned}$$
$\Leftrightarrow A^T A = E_n$.

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 $x = \sum\limits_{i=1}^n x_i e_i$, $y = \sum\limits_{j=1}^n y_j e_j$ 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 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 $x = \sum\limits_{i=1}^n x_i e_i$, $y = \sum\limits_{j=1}^n y_j e_j$:

$$b(x,y) = \sum_{i=1}^n \sum_{j=1}^n x_i b_{ij} y_j = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

wobei der letzte Term als Produkt von Matrizen zu interpretieren ist.

Besonders einfach zu verstehen ist eine Bilinearform b, deren Matrix (bzgl. $e_1, \ldots, e_n$) 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 $i = 1, \ldots, n$) positiv sind.

(6.11) Satz (Hauptachsentransformation). *Zu jeder symmetrischen Bilinearform* $b : V \times V \to \mathbb{R}$ *existiert eine* ONB $f_1, \ldots, f_n$ *von* V *und* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, *so daß für alle* $x = \sum\limits_{i=1}^{n} x_i f_i$, $y = \sum\limits_{j=1}^{n} y_j f_j$ *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)

$$f(x_0 + h) = f(x_0) + Df(x_0)(h) + \frac{1}{2} D^2 f(x_0)(h,h) + R_{f,x_0}(h),$$

wobei $Df(x_0) \colon \mathbb{R}^n \to \mathbb{R}$ linear (1. Ableitung von f an der Stelle x_0), $D^2 f(x_0) \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ 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 $Df(x_0) = 0$ gilt. Aus (6.11) folgt, daß man eine ONB des $\mathbb{R}^n$ finden kann, bzgl. derer $D^2 f(x_0)(h,h) = \sum\limits_{i=1}^{n} \lambda_i h_i^2$ gilt. Daraus folgt für einen kritischen Punkt x_0 von f:

$$\begin{array}{rcl} \text{alle } \lambda_i > 0 \ (\Leftrightarrow D^2f(x_0) \text{ positiv definiert}) & \Rightarrow & x_0 \text{ lokales Minimum.} \\ \text{alle } \lambda_i < 0 \ (\Leftrightarrow D^2f(x_0) \text{ negativ definiert}) & \Rightarrow & x_0 \text{ lokales Maximum.} \end{array}$$

Im Gegensatz zum (aus der Schule bekannten) Fall $n = 1$ gibt es für $n > 1$ noch andere wichtige Möglichkeiten für $D^2f(x_0)$: 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. $\tilde{B}(z_1, \ldots, z_n) =$

$\left(\sum\limits_{i=1}^{n} b_{1i}z_i, \sum\limits_{i=1}^{n} b_{2i}z_i, \ldots, \sum\limits_{i=1}^{n} b_{ni}z_i\right)$. 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: $b(x_1e_1, x_1e_1) = x_1^2 \underbrace{b(e_1, e_1)}_{=:\lambda_1}$.

$n > 1$: 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.

Orthogonales Komplement und Orthogonalprojektion

Wir betrachten weiterhin einen euklidischen Vektorraum $V, \langle , \rangle$.

(6.13) Def.: Ist $M \subseteq V$, so heißt

$$M^\perp := \{v \in V \mid \forall w \in M : \langle v, w\rangle = 0\}$$

das orthogonale Komplement von M.

(6.14) Fakt.

(i) $M^\perp$ ist Untervektorraum von V.
(ii) $M_1 \subseteq M_2 \subseteq V \Rightarrow M_2^\perp \subseteq M_1^\perp$
(iii) $M^\perp = (\mathrm{span}(M))^\perp$
(iv) $M \subseteq (M^\perp)^\perp$

Bew. von (iii): Wegen (ii) genügt es, $M^\perp \subseteq (\mathrm{span}(M))^\perp$ zu beweisen. Sei also $v \in M^\perp$. Wir müssen zeigen, daß für alle $w \in \mathrm{span}(M)$ gilt: $\langle v, w\rangle = 0$. Zu jedem $w \in \mathrm{span}(M)$ existieren $k \in \mathbb{N}$, $r_1, \ldots, r_k \in \mathbb{R}$ und $w_1, \ldots, w_k \in M$:

$$w = \sum_{i=1}^{k} r_i w_i \qquad \text{(siehe (3.7))}.$$

Wegen $v \in M^\perp$ und $w_i \in M$ gilt $\langle v, w_i\rangle = 0$ für $1 \leq i \leq k$. Also:

$$\langle v, w\rangle = \langle v, \sum_{i=1}^{k} r_i w_i\rangle = \sum_{i=1}^{k} r_i \langle v, w_i\rangle = 0.$$

Bsp.: $V = \mathbb{R}^3$ mit dem Standardskalarprodukt

$$\begin{aligned} M = \{e_1, e_2\} \quad &\Rightarrow \quad \mathrm{span}(M) = \mathbb{R}^2 \times \{0\} \\ M^\perp = (\mathrm{span}(M))^\perp \quad &= \quad \{(0,0)\} \times \mathbb{R} \end{aligned}$$

(6.15) Satz. *Sei* $\dim V < \infty$. *Dann gilt für jeden Untervektorraum* U *von* V*:*

$$U \oplus U^\perp = V \quad \textit{und} \quad (U^\perp)^\perp = U.$$

Speziell: $\dim U + \dim U^\perp = \dim V$.

Bew.: Sei $\dim U =: k$. Ergänze eine Basis $(v_1, \ldots, v_k)$ von U zu einer Basis $(v_1, \ldots, v_n)$ von V, vgl. (3.16). Nach (6.7) (Gram-Schmidt) existiert eine ONB $(w_1, \ldots, w_n)$ von V, so daß

$$\mathrm{span}\{w_1, \ldots, w_k\} = \mathrm{span}\{v_1, \ldots, v_k\} = U$$

gilt. Wir zeigen, daß $U^\perp = \mathrm{span}\{w_{k+1}, \ldots, w_n\}$ gilt. Denn:

(i) Ist $i \leq k < j$, so gilt $\langle w_i, w_j\rangle = 0$, also $w_j \in \{w_1, \ldots, w_k\}^\perp = U^\perp$. Wegen (6.14)(i) folgt $\mathrm{span}\{w_{k+1}, \ldots, w_n\} \subseteq U^\perp$.

(ii) Ist $v = \sum_{j=1}^{n} r_j w_j \in U^\perp$, so gilt für $1 \leq i \leq k$:

$$0 = \langle v, w_i\rangle = \sum_{j=1}^{n} r_j \langle w_j, w_i\rangle = r_i.$$

Also $v = \sum\limits_{j=k+1}^{n} r_j w_j \in \operatorname{span}\{w_{k+1}, \dots, w_n\}^{\perp}$.

Aus $U = \operatorname{span}\{w_1, \dots, w_k\}$ und $U^{\perp} = \operatorname{span}\{w_{k+1}, \dots, w_n\}$ folgt $V = U \oplus U^{\perp}$ und $(U^{\perp})^{\perp} = U$.

Bem.: 1) Der Beweis von (6.15) liefert ein Rechenverfahren zur Bestimmung einer ONB $\{w_1, \dots, w_n\}$ von V, so daß $\{w_1, \dots, w_k\}$ eine ONB von U und $\{w_{k+1}, \dots, w_n\}$ eine ONB von $U^{\perp}$ ist.
2) (6.15) gilt nicht ohne die Voraussetzung $\dim V < \infty$!

Ist U Untervektorraum eines endlich-dimensionalen euklidischen Vektorraums V, so ist $U^{\perp}$ ein zu U komplementärer Vektorraum, vgl. (3.21). Unter den vielen zu U komplementären Untervektorräumen ist $U^{\perp}$ durch die Eigenschaft ausgezeichnet, daß jedes $v \in U^{\perp}$ zu allen $u \in U$ orthogonal ist (d.h. der Name "orthogonales Komplement"). Ist $\dim V = \infty$, so kann $U \oplus U^{\perp} \subsetneqq V$ gelten. In diesem Fall ist das orthogonale Komplement $U^{\perp}$ also <u>kein</u> zu U komplementärer Unterraum. (Ein "orthogonales Komplement" ist also nicht notwendig ein "Komplement". So etwas tritt in der mathematischen Sprache öfters auf, wie auch in der Umgangssprache, in der mit einem "tollen Hecht" oft kein Hecht gemeint ist.)

(6.16) Def.: Sei U Untervektorraum von V und $V = U \oplus U^{\perp}$. Dann existiert für jedes $v \in V$ genau ein Paar $(u_0, u_1) \in U \times U^{\perp}$, so daß $v = u_0 + u_1$ gilt. Die Abbildung $P_U : V \to U$, $P_U(v) := u_0$, heißt <u>Orthogonalprojektion von</u> V <u>auf</u> U.

Bem.:

0) $P_U(v)$ ist durch die Eigenschaften $P_U(v) \in U$ und $v - P_U(v) \in U^{\perp}$ eindeutig bestimmt.
1) $P_U \in \operatorname{Hom}(V, U)$.
2) Für alle $u \in U$ gilt $P_U(u) = u$ (d.h. $P_U|U = \mathrm{id}_U$) und für alle $v \in U^{\perp}$ gilt $P_U(v) = 0$.
3) Ist $(v_1, \dots, v_k)$ ONB von U, so gilt für alle $v \in V$:

$$P_U(v) = \sum_{i=1}^{k} \langle v, v_i \rangle v_i.$$

4) Ist $(v_{k+1}, \dots, v_n)$ ONB von $U^{\perp}$, so gilt für alle $v \in V$:

$$P_U(v) = v - \sum_{i=k+1}^{n} \langle v, v_i \rangle v_i.$$

Bez.: Ist $\emptyset \neq M \subseteq V$ und $v \in V$, so heißt

$$d(v, M) := \inf\{\|v - u\| \mid u \in M\}$$

der Abstand von v zu M.

(6.17) Satz. *Sei $U \subseteq V$ Untervektorraum und $V = U \oplus U^\perp$. Dann gilt für jedes $v \in V$: Es existiert genau ein $\overline{u} \in U$ mit $d(v, U) = \|v - \overline{u}\|$, nämlich $\overline{u} = P_U(v)$.*
Speziell gilt: $d(v, U) = \|v - P_U(v)\|$.

Bew.: Sei $u \in U$ beliebig. Zerlege $v = P_U(v) + (v - P_U(v))$, wobei $P_U(v) \in U$, $(v - P_U(v)) \in U^\perp$. Dann gilt

$$\|v-u\|^2 = \|(P_U(v)-u) + (v - P_U(v))\|^2 = \|(P_U(v)-u)\|^2 + \|v - P_U(v)\|^2.$$

Also $\|v - u\| \geq \|v - P_U(v)\|$, wobei Gleichheit genau für $u = P_U(v)$ eintritt.

Bsp.: Approximation komplizierter Funktionen durch einfache: die Fourierentwicklung. Wir betrachten den $\mathbb{R}$-Vektorraum

$$V = C^0([0, 2\pi], \mathbb{R}) = \{f \mid f : [0, 2\pi] \to \text{stetig}\}$$

mit dem sogenannten L^2-Skalarprodukt

$$\langle f, g\rangle := \int\limits_0^{2\pi} f(x)g(x)dx$$

und dem ONS $f_0(x) = \frac{1}{\sqrt{2\pi}}$, $f_k(x) = \frac{1}{\sqrt{\pi}}\cos(kx)$, $g_k(x) = \frac{1}{\sqrt{\pi}}\sin(kx)$, vgl. Bsp. 1 nach (6.1) und Bsp. nach (6.5). Es sei

$$U_n := \operatorname{span}\{f_0, \ldots, f_n, g_1, \ldots, g_n\}.$$

Wegen Bem. 3 nach (6.16) gilt für alle $f \in V$:

$$(*) \qquad P_{U_n}(f) = \sum_{k=0}^{n}\langle f, f_k\rangle f_k + \sum_{k=1}^{n}\langle f, g_k\rangle g_k.$$

Im Gegensatz zu dem nach (6.6) geschriebenen, definiert man üblicherweise die Fourierkoeffizienten $a_k = a_k(f)$ und $b_k = b_k(f)$ durch

$$\begin{aligned} a_k = \tfrac{1}{\pi}\int\limits_0^{2\pi} f(x)\cos(kx)dx &= \begin{cases} \sqrt{\frac{2}{\pi}}\langle f, f_0\rangle & \text{für } k = 0 \\ \frac{1}{\sqrt{\pi}}\langle f, f_k\rangle & \text{für } k > 0 \end{cases} \\ b_k = \tfrac{1}{\pi}\int\limits_0^{2\pi} f(x)\sin(kx)dx &= \tfrac{1}{\sqrt{\pi}}\langle f, g_k\rangle \text{ für } k > 0. \end{aligned}$$

Damit schreibt sich $(*)$ als

$$P_{U_n}(f)(x) = \frac{a_0}{2} + \sum_{k=1}^{n}\left(a_k\cos(kx) + b_k\sin(kx)\right).$$

Satz (6.17) besagt, daß $P_{U_n}(f)$ das eindeutig bestimmte Element von U_n mit minimalem L^2-Abstand von f ist, d.h. die beste (L^2-)Approximation von f durch ein Element von U_n.

Die (hier nicht bewiesene) Tatsache, daß $\{f_0\} \cup \{f_k, g_k \mid k > 0\}$ ein vollständiges ONS bilden, heißt gerade, daß für jedes $f \in V$ gilt

$$\left(\lim_{n\to\infty} d(f, U_n) =\right) \lim_{n\to\infty} \|f - P_{U_n}(f)\| = 0.$$

Die Gramsche Determinante (J.P. Gram, dän. Math. 1850-1916)

(6.18) Satz. *Sei V euklidischer Vektorraum, $0 < \dim V = n < \infty$. Dann existieren genau zwei Determinantenformen $\pm D$ auf V, so daß gilt:*

$$\textit{Ist } v_1, \ldots, v_n \text{ ONB } \textit{von } V, \textit{so gilt } | \pm D(v_1, \ldots, v_n)| = 1.$$

Bez.: Ein solches D heiße <u>normierte</u> Determinantenform des euklidischen Vektorraums V.

Bew.: Existenz: Sei $(\overline{v}_1, \ldots, \overline{v}_n)$ eine (feste) ONB vn V. Nach (5.10) existiert genau eine Determinantenform D von V mit

$$D(\overline{v}_1, \ldots, \overline{v}_n) = 1.$$

Wir müssen zeigen, daß dann für jede ONB $(v_1, \ldots, v_n)$ von V $|D(v_1, \ldots, v_n)| = 1$ gilt. Sei $L \in \operatorname{End} V$ durch $L(\overline{v}_i) = v_i$ für $1 \leq i \leq n$ definiert. Nach Definition (5.19) gilt:

$$D(v_1, \ldots, v_n) = \det L \;\; D(\overline{v}_1, \ldots, \overline{v}_n) = \det L.$$

Nach (6.10) gilt für die Matrix A von L bzgl. $(v_1, \ldots, v_n)$:

$$A^T A = E_n.$$

Also $1 = \det E_n = \det(A^T A) \overset{(5.20)}{=} \det A^T \det A \overset{(5.17)}{=} (\det A)^2 \overset{(5.21)}{=} (\det L)^2$ und damit

$$|D(v_1, \ldots, v_n| = |\det L| = 1.$$

Daß $\pm D$ die einzigen solchen Determinantenformen sind, folgt direkt aus (5.10).

Aus Satz (6.18) folgt, daß in endlich-dimensionalen euklidischen Vektorräumen $(V, \langle , \rangle)$ ein natürlicher Volumenbegriff für Parallelotope existiert. Ist D eine normierte Determinantenform und sind $w_1, \ldots, w_n \in V$, so definieren wir

$$\operatorname{vol}_n^{\langle , \rangle}(P(w_1, \ldots, w_n)) := |D(w_1, \ldots, w_n)|.$$

Ist $(v_1, \ldots, v_n)$ eine ONB von V, so nennt man $P(v_1, \ldots, v_n)$ einen Würfel der Kantenlänge 1, und es folgt:

$$\operatorname{vol}_n^{\langle , \rangle}(P(v_1, \ldots, v_n)) = 1.$$

Die Gramsche Determinante berechnet $\operatorname{vol}_n^{\langle , \rangle}(P(w_1, \ldots, w_n))$ mittels der Skalarprodukte $\langle w_i, w_k \rangle$ für $1 \leq i, k \leq n$:

(6.19) Satz (Gramsche Determinante). *Sei $(V, \langle , \rangle)$ n-dimensionaler euklidischer Vektorraum und $w_1, \ldots, w_n \in V$. Dann gilt*

$$\operatorname{vol}_n^{\langle , \rangle}(P(w_1, \ldots, w_n)) = \sqrt{\det((\langle w_i, w_k \rangle)_{1 \leq i,k \leq n})}.$$

Bsp.: Wir betrachten $\mathbb{R}^2$ mit dem Standardskalarprodukt. Dann ist D_0 eine normierte Determinantenform, da $D_0(e_1, e_2) = 1$. In diesem Fall sagt (6.19):

$$\begin{aligned}
\mathrm{vol}_2^{\langle,\rangle}(P(w_1, w_2)) &= \sqrt{\det \begin{pmatrix} \langle w_1, w_1\rangle & \langle w_1, w_2\rangle \\ \langle w_2, w_1\rangle & \langle w_2, w_2\rangle \end{pmatrix}} = \sqrt{\|w_1\|^2\|w_2\|^2 - \langle w_1, w_2\rangle^2} \\
&= \|w_1\|\|w_2\|\sqrt{1-\cos^2\varphi} = \|w_1\|\|w_2\|\sin\varphi,
\end{aligned}$$

wobei φ den Winkel zwischen w_1 und w_2 bezeichnet. Das ist die Formel für die Fläche des Parallelogramms $P(w_1, w_2)$ mit den Seitenlängen $\|w_1\|$ und $\|w_2\|$ und dem eingeschlossenen Winkel φ.

Beweis von (6.19): Sei $\mathcal{G} = (v_1, \ldots, v_n)$ eine ONB von V und $L \in \mathrm{End}(V)$ durch $L(v_j) = w_j$ für $1 \leq j \leq n$ definiert. Dann ist $A = (a_{ij})_{1\leq i,j\leq n} = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$ durch

$$w_j = \sum_{i=1}^n a_{ij} v_i$$

definiert, und es gilt mit einer normierten Determinantenform D:

$(*)$ $\qquad \mathrm{vol}_n^{\langle,\rangle}(P(w_1, \ldots, w_n)) = |D(w_1, \ldots, w_n)| = |\det A||D(v_1, \ldots, v_n)| = |\det A|.$

Andererseits:

$$\begin{aligned}
\langle w_i, w_k\rangle &= \left\langle \sum_{j=1}^n a_{ji}v_j, \sum_{l=1}^n a_{lk}v_l \right\rangle = \sum_{j,l=1}^n a_{ji}a_{lk} \underbrace{\langle v_j, v_l\rangle}_{=\delta_{jl}} = \sum_{j=1}^n a_{ji}a_{jk} \\
&= (A^T \cdot A)_{ik}
\end{aligned}$$

Also: $(\langle w_i, w_k\rangle)_{1\leq i,k\leq n} = A^T A$.

Daraus folgt: $(**)$ $\det(\langle w_i, w_k\rangle)_{1\leq i,k\leq n} = \det(A^T A) = (\det A)^2$.

Aus $(*)$ und $(**)$ folgt die Behauptung.

Jeder Unterraum U eines euklidischen Vektorraums $(V, \langle,\rangle)$ "erbt" das Skalarprodukt von V, d.h. $\langle,\rangle \mid U \times U$ ist ein Skalarprodukt auf U. Damit ist auch das Volumen von k $(\leq n)$-dimensionalen Parallelotopen in V definiert:

Sind $v_1, \ldots, v_k \in V$ linear unabhängig, so heißt

$$P(v_1, \ldots, v_k) = \left\{ \sum_{i=1}^k s_i v_i \mid 0 \leq s_i \leq 1 \right\} \subseteq \mathrm{span}\{v_1, \ldots, v_k\}$$

das von $v_1, \ldots, v_k$ aufgespannte (k-dimensionale) Parallelotop.

Ist D^U normierte Determinantenform auf $U := \mathrm{span}\{v_1, \ldots, v_k\}$, so besagt (6.19)

$$|D^U(v_1, \ldots, v_k)| = \sqrt{\det((\langle v_i, v_j\rangle)_{1\leq i,j\leq k})}$$

und wir definieren diesen Wert als das k-dimensionale Volumen $\mathrm{vol}_k^{\langle,\rangle}(P(v_1,\ldots,v_k))$ von $P(v_1,\ldots,v_k)$. Sind $v_1,\ldots,v_k$ linear abhängig, so definieren wir $\mathrm{vol}_k^{\langle,\rangle}(P(v_1,\ldots,v_k))=0$.

Orthogonale Abbildungen

Wir wollen die orthogonale Endomorphismen $L\in O(V)$ eines euklidischen Vektorraums $(V,\langle,\rangle)$ untersuchen.

Erinnerung (an Def. (6.8)):

$$L\in O(V)\Leftrightarrow L\in \mathrm{End}(V) \text{ und für alle } v,w\in V \text{ gilt } \langle L(v),L(w)\rangle=\langle v,w\rangle.$$

Solche $L\in O(V)$ sind stets injektiv und damit – falls $\dim V<\infty$ – auch surjektiv. Wir setzen voraus, daß $\dim V=n<\infty$ gilt. Dann ist $O(V)$ eine Untergruppe von $\mathrm{Aut}(V)$ (mit der Hintereinanderausführung von Abbildungen als Gruppenoperation). Nach Satz (6.10) gilt: Ist $\mathcal{G}=(v_1,\ldots,v_n)$ ONB von V, ist $L\in\mathrm{End}(V)$ und $A:=\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$, so gilt:

$$L\in O(V)\Leftrightarrow A^T\cdot A=E_n.$$

Speziell folgt daraus (wie im Beweis von (6.18) schon benutzt):

$$L\in O(V)\Rightarrow |\det L|=|\det A|=1.$$

(6.20) Def.: $A\in\mathbb{R}^{n\times n}$ heißt orthogonal $\Leftrightarrow A^TA=E_n$
$O(n):=\{A\mid A\in\mathbb{R}^{n\times n} \text{ orthogonal}\}$
$SO(n):=\{A\mid A\in O(n),\ \det A=1\}=O(n)\cap \mathrm{SL}_n(\mathbb{R})$
$SO(V):=\{L\mid L\in O(V),\ \det L=1\}$

Da $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}:(\mathrm{Aut}(V),\circ)\to(\mathrm{GL}_n(\mathbb{R}),\cdot)$ ein Gruppenisomorphismus ist (vgl. (4.18)), ist $O(n)=\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(O(V))$ eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$. Ebenso ist $SO(n)$ (bzw. $SO(V)$) Untergruppe von $O(n)$ (bzw. von $O(V)$).

Wir beschäftigen uns zunächst mit dem Spezialfall $V=\mathbb{R}^2$, $\langle,\rangle=$ Standardskalarprodukt, d.h.

$$\langle(x_1,x_2),(y_1,y_2)\rangle:=x_1y_1+x_2y_2.$$

Beispiele von Elementen in $O(\mathbb{R}^2)$:

1) "Drehungen": Sei $\varphi\in\mathbb{R}$ und $D_\varphi\in\mathrm{End}(\mathbb{R}^2)$ mit

$$\mathrm{Mat}(D_\varphi)=\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$

Dann gilt

$$\begin{aligned}\mathrm{Mat}(D_\varphi)^T\cdot\mathrm{Mat}(D_\varphi)&=\begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix}\cdot\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}\\ &=\begin{pmatrix}\cos^2\varphi+\sin^2\varphi & 0\\ 0 & \cos^2\varphi+\sin^2\varphi\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix},\end{aligned}$$

d.h. $\mathrm{Mat}(D_\varphi) \in O(2)$ und damit $D_\varphi \in O(\mathbb{R}^2)$. Wegen

$$\begin{vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{vmatrix} = 1 \text{ gilt sogar } D_\varphi \in SO(\mathbb{R}^2), \mathrm{Mat}(D_\varphi) \in SO(2).$$

Offenbar gilt $D_{\varphi+2\pi} = D_\varphi$ und $D_\varphi \neq D_\psi$, falls $0 \leq \varphi < \psi < 2\pi$.

2) "Spiegelungen an Geraden durch $0 \in \mathbb{R}^2$": Sei $U \subseteq \mathbb{R}^2$ 1-dimensionaler Untervektorraum. Wähle eine ONB $\mathcal{G} = (v_1, v_2)$ des $\mathbb{R}^2$ mit $\mathrm{span}\{v_1\} = U$. Definiere "die Spiegelung $S_U \in \mathrm{End}(\mathbb{R}^2)$ an U" durch

$$S_U(v_1) = v_1, S_U(v_2) = -v_2.$$

Dann gilt $S_U \in O(\mathbb{R}^2) \setminus SO(\mathbb{R}^2)$, da $\mathrm{Mat}^{\mathcal{G}}_{\mathcal{G}}(S_U) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in O(2) \setminus SO(2)$.

S_U ist durch U eindeutig bestimmt!

(6.21) Fakt. Es gilt:

(i) $SO(\mathbb{R}^2) = \{D_\varphi | \varphi \in [0, 2\pi)\}$.

(ii) $O(\mathbb{R}^2) = \{S_U | U$ 1-dimensionaler Unterraum des $\mathbb{R}^2\}$.

(iii) $\forall \varphi, \psi \in \mathbb{R} : D_{\varphi+\psi} = D_\varphi \circ D_\psi (= D_\psi \circ D_\varphi), D_0 = \mathrm{id}_{\mathbb{R}^2}$ (d.h. $\varphi \in (\mathbb{R}, +) \to D_\varphi \in (SO(\mathbb{R}^2), \circ)$ ist surjektiver Gruppenhomomorphismus).

$$D_\varphi = \mathrm{id}_{\mathbb{R}^2} \Leftrightarrow \varphi \in \{2\pi k | k \in \mathbb{Z}\}$$

(iv) $S_{U_2} \circ S_{U_1} = D_{2\varphi}$, wobei φ der Winkel ist, um den man U_1 drehen muß, um U_2 zu erhalten, d.h. $D_\varphi(U_1) = U_2$. (φ ist bis auf additive Vielfache $\varphi + k\pi, k \in \mathbb{Z}$, von π bestimmt.)

(v) $D_{2\varphi} \circ S_U = S_{D_\varphi(U)}, S_U \circ D_{2\varphi} = S_{D_{-\varphi}(U)}$.

Insbesondere folgt aus (iii), (iv) bzw. (v), daß $SO(\mathbb{R}^2)$ abelsch ist, während $O(\mathbb{R}^2)$ nicht abelsch ist.

Bew.:

(i) Sei $L \in SO(\mathbb{R}^2)$, $L(e_1) =: (x, y) \in \mathbb{R}^2$. Dann gilt $x^2 + y^2 = 1$ und mit den Kenntnissen aus der Analysis I kann man einsehen, daß es genau ein $\varphi \in [0, 2\pi)$ gibt mit $(x, y) = (\cos\varphi, \sin\varphi)$, d.h. $L(e_1) = D_\varphi(e_1)$. Dann sind $L(e_2)$ und $D_\varphi(e_2)$ beides Einheitsvektoren, die zu $L(e_1) = D_\varphi(e_1)$ orthogonal sind und zusammen mit $L(e_1) = D_\varphi(e_1)$ eine positiv orientierte Basis bilden. Da es nur einen solchen Vektor gibt (das ist anschaulich klar, muß aber im Prinzip durch eine Rechnung begründet werden), gilt auch $L(e_2) = D_\varphi(e_2)$, also $L = D_\varphi$.

(ii) Sei $L \in O(2) \setminus SO(2)$. Um zu zeigen, daß L eine Spiegelung ist, suchen wir einen 1-dimensionalen Unterraum U, der punktweise von L festgelassen wird, d.h. einen Vektor $v \in \mathbb{R}^2 \setminus \{0\}$ mit $L(v) = v$ ($\Leftrightarrow v$ EV zum EW 1 von L). Ist $\mathrm{Mat}(L) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so gilt wegen $\det L = -1$: $\det(L - \lambda\,\mathrm{id}_{\mathbb{R}^2}) = \lambda^2 - (a + d)\lambda - 1 =: p(\lambda)$.

Wegen $p(0) = -1$ und $\lim\limits_{\lambda\to\pm\infty} p(\lambda) = \infty$, hat $p(\lambda)$ zwei Nullstellen (d.h. L zwei Eigenwerte) $\lambda_2 < 0 < \lambda_1$. Da L orthogonal ist, hat jeder Eigenwert von L den Betrag 1, also $\lambda_2 = -1, \lambda_1 = 1$. Sei v_1 EV von L zum EW $\lambda_1 = 1$ und (v_1, v_2) ONB von $\mathbb{R}^2$. Aus $L \in O(\mathbb{R}^2)$, $L \neq \mathrm{id}_{\mathbb{R}^2}$ und $L(v_1) = v_1$, folgt $L(v_2) = -v_2$, d.h. $L = S_U$ für $U := \mathrm{span}\{v_1\}$.

(iii) folgt aus den "Additionstheoremen" für sin und cos.

(iv) Es gelte $D_\varphi(U_1) = U_2$. Da $\det(S_{U_2} \circ S_{U_1}) = \det(S_{U_2}) \det(S_{U_1}) = (-1)^2 = 1$ ist, ist $S_{U_2} \circ S_{U_1} \in SO(\mathbb{R}^2)$, d.h. $S_{U_2} \circ S_{U_1}$ ist eine Drehung.

Wie in der Vorlesung mit einem einfachen elementargeometrischen Argument (und ähnlich für (v)) gezeigt wurde, ist $S_{U_2} \circ S_{U_1}$ in der Tat eine Drehung um den Winkel 2φ. Dieses Argument muß aber in der hier aufgebauten "Analytischen Geometrie" durch eine Rechnung bewiesen werden. Es ist nun so, daß solche Rechnungen statt mit (2×2)-Matrizen sehr viel weniger aufwendig mit komplexen Zahlen ausgeführt werden können. Deshalb zunächst der

Exkurs: Beschreibung von Drehungen und Spiegelungen des $\mathbb{R}^2$ mit Hilfe der komplexen Zahlen.

Identifizieren wir wie üblich $(x, y) \in \mathbb{R}^2$ mit $x + iy \in \mathbb{C}$ und definieren (!) wir für $\varphi \in \mathbb{R}$

$$e^{i\varphi} := \cos\varphi + i \sin\varphi$$

(für diese Definition gibt es einen mathematischen Hintergrund, der uns jetzt nicht zu interessieren braucht), so berechnet man:

$$D_\varphi(z) = e^{i\varphi} \cdot z.$$

Die Additionstheoreme für sin und cos sind äquivalent zur Gleichung $e^{i(\varphi+\psi)} = e^{i\varphi} \cdot e^{i\psi}$ für alle $\varphi, \psi \in \mathbb{R}$. Außerdem gilt $\overline{e^{i\varphi}} = e^{-i\varphi}$ und $|e^{i\varphi}| = \sqrt{\cos^2 + \sin^2 \varphi} = 1$. Ist $U = \mathrm{span}_{\mathbb{R}}\{e^{i\psi}\} := \{se^{i\psi} \mid s \in \mathbb{R}\}$, so zeigen wir, daß S_U durch

$$S_U(z) = e^{2i\psi}\overline{z}$$

gegeben ist: Die Abbildung $z \in \mathbb{C} \simeq \mathbb{R}^2 \to \overline{z} \in \mathbb{C} \simeq \mathbb{R}^2$ ist gerade die Spiegelung an der x-Achse, so daß die Abbildung

$$z \to e^{2i\psi}\overline{z} = D_{2\psi}(\overline{z})$$

in $O(\mathbb{R}^2)\backslash SO(\mathbb{R}^2)$ liegt. Wendet man sie auf $z = e^{i\psi}$ an, so erhält man $e^{i\psi} \to e^{2i\psi} \cdot e^{-i\psi} = e^{i\psi}$. Da auch $S_U(e^{i\psi}) = e^{i\psi}$ gilt, folgt

$$S_U(z) = e^{2i\psi}\overline{z}$$

für alle $z \in \mathbb{C}$, denn beide Abbildungen sind Spiegelungen, die die Gerade $\mathrm{span}_{\mathbb{R}}\{e^{2i\psi}\}$ fest lassen.

Wir kommen nun zu einem "analytischen" Beweis von (6.21)(iv) und (v):

(iv): Es sei $U_1 = \mathrm{span}_{\mathbb{R}}\{e^{i\psi_1}\} = D_{\psi_1}(\mathbb{R} \times \{0\})$ und $U_2 = \mathrm{span}_{\mathbb{R}}\{e^{i\psi_2}\} = D_{\psi_2}(\mathbb{R} \times \{0\})$.

Dann gilt $D_{\psi_2-\psi_1}(U_1) = D_{\psi_2-\psi_1}(D_{\psi_1}(\mathbb{R}\times\{0\}) \overset{\text{(iii)}}{=} D_{\psi_2}(\mathbb{R}\times\{0\}) = U_2$, d.h. für $\varphi := \psi_2 - \psi_1$ gilt $D_\varphi(U_1) = U_2$. Andererseits berechnen wir für alle $z \in \mathbb{C}$:

$$S_{U_2} \circ S_{U_1}(z) = S_{U_2}(e^{2i\psi_1}\overline{z}) = e^{2i\psi_2}\overline{(e^{2i\psi_1}\overline{z})} = e^{i2(\psi_2-\psi_1)}z = D_{2\varphi}(z)$$

(v): Ist $U = \operatorname{span}_{\mathbb{R}}\{e^{i\psi}\}$, so gilt

$$D_{2\varphi} \circ S_U(z) = e^{2i\varphi}e^{2i\psi}\overline{z} = e^{2i(\varphi+\psi)}\overline{z} = S_{D_\varphi(U)}(z),$$

da $D_\varphi(U) = \operatorname{span}_{\mathbb{R}}\{D_\varphi(e^{i\psi})\} = \operatorname{span}_{\mathbb{R}}\{e^{i(\varphi+\psi)}\}$.

Ebenso erhalten wir:

$$S_U \circ D_{2\varphi}(z) = e^{2i\psi}(e^{-2i\varphi}\overline{z}) = e^{2i(\psi-\varphi)}\overline{z} = S_{D_{-\varphi}(U)}(z).$$

Normalform von orthogonalen Abbildungen

(6.22) Def.: Sei $L \in \operatorname{End}(V)$. Ein Untervektorraum U von V heißt L-invariant, falls $L(U) \subseteq U$ gilt.

Bem.: $\{0\}$ und V sind L-invariant für jedes $L \in \operatorname{End}(V)$.

Ein großes Ziel bei der Untersuchung eines $L \in \operatorname{End}(V)$ ist es, L-invariante Unterräume $U_1 \neq \{0\}, U_2 \neq \{0\}$ zu finden, so daß V die direkte Summe von U_1 und U_2 ist, d.h. $V = U_1 \oplus U_2$. Dann reduziert sich die Untersuchung von L auf die Untersuchung von $L|U_1 \in \operatorname{End}(U_1)$ und $L|U_2 \in \operatorname{End}(U_2)$. (Leider ist das nicht immer möglich, z.B. nicht für die "Scherung" $L \in \operatorname{End}(\mathbb{R}^2)$ mit $\operatorname{Mat}(L) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.)

Man wird natürlich versuchen, solche L-invarianten Unterräume U_1 und U_2 in noch kleinere L-invariante Unterräume zu zerlegen, und dazu benötigen wir den Begriff der direkten Summe von endlich vielen Unterräumen (vgl. (3.19)–(3.21) und Blatt 3, Aufgabe 4): Sind $U_1, \ldots, U_k$ Unterräume eines Vektorraums V, so heißt V die direkte Summe von $U_1, \ldots, U_k$ (geschrieben $V = U_1 \oplus \ldots \oplus U_k = \bigoplus\limits_{i=1}^{k} U_i$), falls gilt:

(i) $V = \operatorname{span}(\bigcup\limits_{i=1}^{k} U_i)$ und

(ii) Für alle $i \in \{1, \ldots, k\}$ gilt: $U_i \cap \operatorname{span}(\bigcup\limits_{\substack{j=1\\ j\neq i}}^{k} U_j) = \{0\}$.

Daraus folgt: Ist für $i = 1, \ldots, k$ B_i eine Basis von U_i, so gilt $B_i \cap B_j = \emptyset$ für $i \neq j$, und $B := \bigcup\limits_{i=1}^{k} B_i$ ist Basis von V. Ist $\dim V < \infty$, so folgt:

$$\dim V = \sum_{i=1}^{k} \dim U_i.$$

Im Fall eines euklidischen Vektorraums $(V, \langle , \rangle)$ ist folgender Spezialfall der direkten Summe wichtig:

(6.23) Def.: Seien $U_1, \ldots, U_k$ Unterräume von V Dann heißt V die orthogonale Summe von $U_1, \ldots, U_k$, falls gilt:

(i) $V = \operatorname{span}(\bigcup\limits_{i=1}^{k} U_i)$ und
(ii) Für alle $1 \leq i \neq j \leq k$ gilt $U_i \subseteq U_j^{\perp}$.

Zur Bedingung (ii) sagt man, die U_i, $1 \leq i \leq k$, seien paarweise orthogonal.

Aus (6.23)(i) und (ii) folgt, daß V die direkte Summe von $U_1, \ldots, U_k$ ist.

(6.24) Satz. *Sei $L \in O(V)$. Dann existieren $k \in \mathbb{N}$, 2-dimensionale, L-invariante Unterräume $U_1, \ldots, U_k$ von V und L-invariante Unterräume U_- und U_+ von V, so daß V die orthogonale Summe von $U_1, \ldots, U_k$, U_-,U_+ ist und so daß gilt*

(i) $L|U_i \in SO(U_i) \setminus \{\pm \mathrm{id}_{U_i}\}$ *für* $1 \leq i \leq k$
(ii) $L|U_- = -(\mathrm{id}_{U_-})$
(iii) $L|U_+ = \mathrm{id}_{U_+}$

Bem.: 1) Es kann $k = 0$ oder $U_- = \{0\}$ oder $U_+ = \{0\}$ gelten.
2) Jedes $v \in V$ läßt sich dann eindeutig darstellen als

$$v = u_1 + \ldots + u_k + u_- + u_+$$

mit $u_i \in U_i$ für $1 \leq i \leq k$, $u_- \in U_-$ und $u_+ \in U_+$. Es gilt dann:

$$L(v) = L(u_1) + \ldots + L(u_k) - u_- + u_+.$$

L setzt sich also aus k "Drehungen" in den 2-dimensionalen Unterräumen U_i, $1 \leq i \leq k$, aus der "Punktspiegelung" $-(\mathrm{id}_{U_-})$ im Unterraum U_- und der Identität auf U_+ zusammen.

3) Es gilt $2k + \dim U_- + \dim U_+ = n$. Ist n ungerade, so folgt $\dim U_- + \dim U_+ \neq 0$.

4) $\det(L) = (-1)^{\dim U_-}$, vgl. Blatt 3, Aufgabe 3.

Umformuliert für Matrizen besagt (6.24):

(6.24)' Satz. *Sei $L \in O(V)$. Dann existieren eine ONB $\mathcal{G} = (v_1, \ldots, v_n)$ von V, $k \in \mathbb{N}$ und $\varphi_1, \ldots, \varphi_k \in (0, \pi)$, so daß gilt:*

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} \boxed{\begin{matrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{matrix}} & & & & & & & & \\ & \boxed{\begin{matrix} \cos\varphi_2 & -\sin\varphi_2 \\ \sin\varphi_2 & \cos\varphi_2 \end{matrix}} & & & & & 0 & & \\ & & \ddots & & & & & & \\ & 0 & & \boxed{\begin{matrix} \cos\varphi_k & -\sin\varphi_k \\ \sin\varphi_k & \cos\varphi_k \end{matrix}} & & & & & \\ & & & & -1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & -1 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix}$$

Äquivalent dazu ist: *Zu jedem $A \in O(n)$ existiert ein $B \in O(n)$, so daß $B^T A B$ die obige Form hat.*

(6.25) Lemma. Sei $L \in \mathrm{End}(\mathbb{R}^n)$. Dann existiert ein L-invarianter Untervektorraum U des $\mathbb{R}^n$ mit $0 < \dim U \leq 2$.

Bew.: Das wurde im 1. Schritt des Beweis von Satz (6.11) (Hauptachsentransformation) wie folgt gezeigt: Wir betrachten die $\mathbb{C}$-lineare Fortsetzung $\tilde{L}$ von L auf $\mathbb{C}^n$ und erhalten einen EW $\lambda = \alpha + i\beta \in \mathbb{C}$ mit α, $\beta \in \mathbb{R}$ und einen zugehörigen EV $v = x + iy \in \mathbb{C}^n \setminus \{0\}$ mit x, $y \in \mathbb{R}^n$. Dann gilt:

$$(*) \qquad L(x) = \alpha x - \beta y \text{ und } L(y) = \beta x + \alpha y.$$

Das zeigt, daß $\mathrm{span}\{x, y\}$ L-invariant ist. Wegen $v = x + iy \neq 0$ gilt $\mathrm{span}\{x, y\} \neq \{0\}$.

Beweis von (6.24): Durch vollständige Induktion nach $n = \dim V$.

Induktionsanfang: Ist $\dim V = 1$, so gilt $O(V) = \{\pm \mathrm{id}_V\}$. Die Behauptung ist richtig mit $k = 0$ und entweder $U_+ = V$ oder $U_- = V$.

Induktionsschritt: Sei $L \in O(V)$ und $\dim V = n > 1$. Wegen (6.9) genügt es, den Fall zu betrachten, daß $(V, \langle , \rangle)$ der $\mathbb{R}^n$ mit dem Standardskalarprodukt ist. Aus Lemma (6.25) folgt, daß ein L-invarianter Unterraum $U \subseteq \mathbb{R}^n$ existiert mit $0 < \dim U \leq 2$.

(i) Gilt $n = 2$ und $U = \mathbb{R}^n$, so folgt die Behauptung aus (6.21):
Entweder L ist eine Spiegelung ($\to \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ für eine geeignete ONB des $\mathbb{R}^2$), oder $L = D_\varphi$ für ein $\varphi \in [0, 2\pi)$.
Ist $\varphi = 0$, so $L = \mathrm{id}_{\mathbb{R}^2}$ ($\Rightarrow k = 0$, $U_- = \{0\}$, $U_+ = \mathbb{R}^2$).
Ist $\varphi = \pi$, so $L = -\mathrm{id}_{\mathbb{R}^2}$ ($\Rightarrow k = 0$, $U_- = \mathbb{R}^2$, $U_+ = \{0\}$). Sonst gilt $D_\varphi \in SO(\mathbb{R}^2) \setminus \{\pm \mathrm{id}_{\mathbb{R}^2}\}$.

(ii) Gilt $U \subsetneqq \mathbb{R}^n$, so folgt aus (6.15)

$$\mathbb{R}^n = U \oplus U^\perp$$

Da U L-invariant ist und L orthogonal ist, ist auch $U^\perp$ L-invariant (vgl. Blatt 1, Aufgabe 1). Wegen $\dim U^\perp = n - \dim U < n$ ist auf $L|U^\perp$ die Induktionsvoraussetzung anwendbar. Daraus folgt zusammen mit unseren Kenntnissen über $O(\mathbb{R})$ und $O(\mathbb{R}^2)$ die Behauptung.

Die einzige zusätzliche Information, die man zum Beweis von (6.24)' benötigt, ist folgende: Ist $\dim V = 2$ und $L \in SO(V) \setminus \{\pm \mathrm{id}_V\}$, so existiert eine ONB $\mathcal{G}$ von V und $\varphi \in (0, \pi)$, so daß $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$ gilt. Das sieht man so ein. Ist $\tilde{\mathcal{G}} = (v_1, v_2)$ irgendeine ONB von V, so existiert $\tilde{\varphi} \in (0, \pi) \cup (\pi, 2\pi)$ mit

$$\mathrm{Mat}_{\tilde{\mathcal{G}}}^{\tilde{\mathcal{G}}}(L) = \begin{pmatrix} \cos\tilde{\varphi} & -\sin\tilde{\varphi} \\ \sin\tilde{\varphi} & \cos\tilde{\varphi}. \end{pmatrix}$$

Ist $\tilde{\varphi} \in (\pi, 2\pi)$, so betrachten wir $\mathcal{G} = (v_1, -v_2)$ und erhalten

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} \cos\tilde{\varphi} & \sin\tilde{\varphi} \\ -\sin\tilde{\varphi} & \cos\tilde{\varphi}. \end{pmatrix}$$

Wegen $\cos(-\tilde{\varphi}) = \cos\tilde{\varphi}$, $\sin(-\tilde{\varphi}) = -\sin\tilde{\varphi}$ folgt mit $\varphi := 2\pi - \tilde{\varphi} \in (0, \pi)$:

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi, \end{pmatrix}$$

wie behauptet.

Die anschauliche Begründung für die letzte Überlegung ist wie folgt: Ist $\dim V = 2$ und $L \in SO(V)$, so ist der "Drehwinkel" $\tilde{\varphi}$ von L erst nach Wahl einer Orientierung ($\Rightarrow$ eines "positiven Drehsinns") für V definiert. Ändert man die Orientierung, so geht $\tilde{\varphi}$ in $-\tilde{\varphi}$ über. Ist $\tilde{\varphi} \in (\pi, 2\pi)$, so ist $-\tilde{\varphi} \in (-2\pi, -\pi)$ und statt $-\tilde{\varphi}$ kann man natürlich auch $\varphi := 2\pi - \tilde{\varphi} \in (0, \pi)$ nehmen.

Oft ist es wichtig zu wissen, ob die Unterräume in (6.24) und ob die Normalform in (6.24)' und die zugehörige ONB eindeutig durch L bestimmt sind.

Hierzu kann man folgendes sagen:
Es gilt $U_+ = \ker(L - \mathrm{id}_V)$, $U_- = \ker(L + \mathrm{id}_V)$, so daß U_-, U_+ und damit auch $k = \frac{1}{2}(n - (\dim U_- + \dim U_+))$ eindeutig durch L bestimmt sind. Die $\varphi_1, \ldots, \varphi_k \in (0, \pi)$ sind (bis auf ihre Reihenfolge) eindeutig durch L bestimmt (und damit auch die Normalform in (6.24)'), da $e^{i\varphi_1}, e^{-i\varphi_1}, \ldots e^{i\varphi_k}, e^{-i\varphi_k}$ gerade die Eigenwerte $\neq \pm 1$ von $\tilde{L}$ sind. Aus (6.24)' folgt nämlich sofort:

$$\det(L - \lambda\,\mathrm{id}) = \prod_{j=1}^{k} \underbrace{(\lambda^2 - 2(\cos\varphi_j)\lambda + 1)}_{(\lambda - e^{i\varphi_j})(\lambda - e^{-i\varphi_j})} \cdot (1 + \lambda)^{\dim U_-} \cdot (1 - \lambda)^{\dim U_+}$$

Eine ONB, in der die Matrix von L die Normalform (6.24)' annimmt, ist nicht eindeutig durch L bestimmt. Z.B. ist die Matrix einer Drehung, $D_\varphi \in SO(\mathbb{R}^2)$, $\varphi \in (0, \pi)$, bezüglich jeder positiv orientierten ONB des $\mathbb{R}^2$ in der Normalform $\begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix}$.

Nach so vielen Worten möchte man hoffen, nun alles über orthogonale Abbildungen zu wissen. Weit gefehlt! Über die Gruppenstruktur von $SO(n)$ für $n \geq 3$ wurde etwa noch nichts gesagt. Man möchte auch gern die Elemente von $SO(3)$ durch 3 (warum gerade 3?) reelle "Parameter" beschreiben (so wie wir die Elemente von $SO(2)$ durch den Drehwinkel $\varphi \mod 2\pi$ beschrieben haben). Aber geht das und wie am besten? Da gibt es Fragen und Antworten, genug für ein ganzes Mathematikstudium...

7. Dualität

In diesem Kapitel geht es um die Begriffe "Dualraum" und "dualer Homomorphismus", die sowohl im endlich- wie im unendlich-dimensionalen Fall (hier als algebraischer Hintergrund für die "Funktionalanalysis") wichtig sind. Dennoch ist diesen Begriffen eine gewisse Unanschaulichkeit eigen. Neben der Vorbereitung auf weiterführende Gegenstände dient dieses Kapitel auch einer sachgerechten Behandlung der Transposition von Matrizen (vgl. (5.16) und (4.23)).

Im folgenden sei K ein Körper und V ein K-Vektorraum. Eine <u>Linearform</u> <u>auf</u> V ist eine Abbildung $l : V \to K$, so daß für alle $\alpha, \beta \in K$ und alle $v, w \in V$ gilt:

$$l(\alpha v + \beta w) = \alpha l(v) + \beta l(w).$$

Mit anderen Worten ausgedrückt ist eine Linearform l gerade ein Vektorraumhomomorphismus von V in den 1-dimensionalen K-Vektorraum K, d.h. $l \in \mathrm{Hom}(V, K)$, vgl. (4.1). In (4.9) wird erklärt, daß $\mathrm{Hom}(V, K)$ eine natürliche K-Vektorraumstruktur besitzt, in der etwa die Addition von zwei Linearformen $l_1, l_2 \in \mathrm{Hom}(V, K)$ "punktweise" definiert ist durch:

$$\forall v \in V : \quad (l_1 + l_2)(v) := l_1(v) + l_2(v).$$

(7.1) Definition. Der K-Vektorraum $V^* := \mathrm{Hom}(V, K) = \{l : V \to K \mid l \text{ Linearform}\}$ heißt Dualraum von V.

Bem.: Aus (4.12)–(4.14) folgt: Ist $\dim V < \infty$, so gilt $\dim V^* = \dim V$.

Erinnerung an (4.6): Ist B eine Basis von V und $\tilde{l} : B \to K$ irgendeine Abbildung, so existiert genau ein $l \in V^*$ mit $l \mid B = \tilde{l}$, d.h. mit $l(v) = \tilde{l}(v)$ für alle $v \in B$.

Im Spezialfall $V = K^n$ mit der Standardbasis $(e_1, \ldots, e_n)$ erhalten wir: Ist $l \in (K^n)^*$ und $l(e_i) =: a_i \in K$ für $1 \le i \le n$, so gilt

$$l(x_1, \ldots, x_n) = a_1 x_1 + \ldots + a_n x_n = \sum_{i=1}^{n} a_i x_i.$$

(Denn $l(x_1, \ldots, x_n) = l(\sum_{i=1}^{n} x_i e_i) = \sum_{i=1}^{n} x_i l(e_i) = \sum_{i=1}^{n} a_i x_i$.) In diesem Fall ist eine Linearform $l \in (K^n)^*$ also "nichts anderes" als die linke Seite einer einzigen homogenen linearen Gleichung (mit n Unbekannten in K), und die Vektorraumstruktur in $(K^n)^*$ formalisiert genau das, was wir schon immer (Gaußsches Eliminationsverfahren) mit solchen Gleichungen getan haben: wir haben sie mit Körperelementen multipliziert und zueinander addiert. Bezüglich einer Basis $(v_1, \ldots, v_n)$ eines n-dimensionalen K-Vektorraums V entspricht nach (4.11) einem $l \in V^*$ die $(1 \times n)$-Matrix $(a_1, \ldots, a_n) \in K^{1 \times n}$ mit $a_i = l(v_i)$ für $1 \le i \le n$, die man oft auch als "Zeilenvektor" interpretiert.

Bsp.: Auf dem $\mathbb{R}$-Vektorraum $V = C^0([a,b],\mathbb{R}) = \{f \mid f : [a,b] \to \mathbb{R} \text{ stetig}\}$ ist die Integration eine Linearform, d.h. definiert man für alle $f \in V$

$$l(f) := \int_a^b f(t)dt,$$

so gilt $l \in V^*$.

(7.2) Satz. *Sei* $\dim V = n$ *und* $\mathcal{G} = (v_1, \ldots, v_n)$ *eine Basis von* V. *Definiere* $l_i \in V^*$ *für* $1 \le i \le n$ *durch*

$$l_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \quad \text{für } 1 \le j \le n.$$

Dann ist $\mathcal{G}^* = (l_1, \ldots, l_n)$ *Basis von* V^*, *genannt die* <u>*Dualbasis*</u> <u>*zu*</u> $\mathcal{G}$.
Für alle $l \in V^*$ *gilt:* $l = \sum_{i=1}^n l(v_i) l_i$.

Bew.:

(i) $\mathcal{G}^*$ ist linear unabhängig: Sei $\sum_{i=1}^n \alpha_i l_i = 0$. Dann gilt für $1 \le j \le n$:
$0 = (\sum_{i=1}^n \alpha_i l_i)(v_j) = \sum_{i=1}^n \alpha_i l_i(v_j) = \sum_{i=1}^n \alpha_i \delta_{ij} = \alpha_j$.

(ii) $\mathcal{G}^*$ erzeugt V^*: Wir zeigen, daß für jedes $l \in V^*$ gilt: $l = \sum_{i=1}^n l(v_i) l_i$. Da beiden Seiten dieser Gleichung Linearformen sind, genügt es nachzuweisen, daß beide Seiten auf allen Basiselementen $v_j, 1 \le j \le n$, die gleichen Werte annehmen, vgl. (4.6):

$$\left(\sum_{i=1}^n l(v_i) l_i\right)(v_j) = \sum_{i=1}^n l(v_i) l_i(v_j) = \sum_{i=1}^n l(v_i)\delta_{ij} = l(v_j).$$

<u>Wichtige Bemerkung</u>: Analog kann man im Fall $\dim V = \infty$ aus einer Basis B eine (ebenfalls unendliche) linear unabhängige Menge $B^* \subseteq V^*$ konstruieren ($\Rightarrow \dim V^* = \infty$). Es gilt aber nicht $\operatorname{span} B^* = V^*$, z.B. läßt sich die Linearform $l \in V^*$ mit $l(b) = 1$ für alle $b \in B$ <u>nicht</u> als (endliche!) Linearkombination von Elementen aus B^* darstellen.

(7.3) Def.: Sind V, W K-Vektorräume und ist $L \in \operatorname{Hom}(V,W)$, so heißt die Abbildung $L^* : W^* \to V^*$, die durch

$$L^*(l) = l \circ L \quad (\text{oder explizit:} \forall v \in V \text{ gilt } (L^*(l))(v) = l(L(v)))$$

definiert ist, die zu L duale Abbildung.

(7.4) Fakt. (i) Die Abbildung $L \in \operatorname{Hom}(V,W) \to L^* \in \operatorname{Hom}(W^*, V^*)$ ist ein injektiver K-Vektorraumhomomorphismus von $\operatorname{Hom}(V,W)$ nach $\operatorname{Hom}(W^*, V^*)$.
(ii) Ist Z ein weiterer K-Vektorraum und ist $J \in \operatorname{Hom}(V,W), L \in \operatorname{Hom}(W,Z)$,
so gilt: $\quad (L \circ J)^* = J^* \circ L^* \in \operatorname{Hom}(Z^*, V^*)$.
Außerdem gilt: $\quad (\mathrm{id}_V)^* = \mathrm{id}_{V^*}$.

Bem.:

1) Aus (4.13), (4.14) und (7.2) folgt, im Fall $\dim V = n < \infty$, $\dim W = m < \infty$: $\dim \operatorname{Hom}(V, W) = m \cdot n = \dim \operatorname{Hom}(W^*, V^*)$. In diesem Fall ist also die Abbildung $L \in \operatorname{Hom}(V, W) \to L^* \in \operatorname{Hom}(W^*, V^*)$ ein Isomorphismus, vgl. (4.8).
2) In der Sprache der Kategorien besagt (7.4)(ii), daß * ein kontravarianter Funktor auf der Kategorie der K-Vektorräume ist.

Bew.:

(i) Zeige z.B.: $L_1, L_2 \in \operatorname{Hom}(V, W) \Rightarrow (L_1+L_2)^* = L_1^*+L_2^*$. Denn es gilt für alle $l \in W^*$: $(L_1 + L_2)^*(l) = l \circ (L_1 + L_2) \overset{l \text{ linear}}{=} l \circ L_1 + l \circ L_2 = L_1^*(l) + L_2^*(l) = (L_1^* + L_2^*)(l)$.
Injektivität: Zeige $L \neq 0 \Rightarrow L^* \neq 0$. Ist $L \neq 0$, so existiert ein $v \in V$ mit $L(v) \neq 0$. Ergänze $w := L(v) \neq 0$ zu einer Basis B von W und definiere $l \in W^*$ durch $l(w) = 1$, $l(w') = 0$ für $w' \in B \setminus \{w\}$. Dann gilt $1 = l(L(v)) = (L^*(l))(v)$, also $L^*(l) \neq 0$, also $L^* \neq 0$.

(ii) Für $l \in Z^*$ gilt $(L \circ J)^*(l) = l \circ (L \circ J) = (l \circ L) \circ J = J^*(l \circ L) = J^*(L^*(l)) = (J^* \circ L^*)(l)$.

(7.5) Fakt. Sind $\mathcal{G}_V = (v_1, \ldots, v_n)$ bzw. $\mathcal{G}_W = (w_1, \ldots, w_m)$ Basen von V bzw. W und ist $l \in \operatorname{Hom}(V, W)$, so gilt:

$$\operatorname{Mat}^{\mathcal{G}_V^*}_{\mathcal{G}_W^*}(L^*) = \left(\operatorname{Mat}^{\mathcal{G}_W}_{\mathcal{G}_V}(L)\right)^T$$

(oder mit Worten: Bezüglich dualer Basen ist die Matrix von L^* gerade die transponierte Matrix von L).

Bem.: Aus (7.4)(ii) und (7.5) folgt, daß für alle $A \in K^{m \times n}$, $B \in K^{n \times p}$ gilt:

$$(A \cdot B)^T = B^T \cdot A^T,$$

was man auch leicht direkt nachrechnen kann.

Bew.: Sei $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \operatorname{Mat}^{\mathcal{G}_W}_{\mathcal{G}_V}(L)$, d.h. es gilt

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

Sei $\mathcal{G}_V^* = (l_1, \ldots, l_n)$, $\mathcal{G}_W^* = (f_1, \ldots, f_m)$. Dann ist $B = (b_{kr})_{\substack{1 \le k \le n \\ 1 \le r \le m}} = \operatorname{Mat}^{\mathcal{G}_{V^*}}_{\mathcal{G}_W^*}(L^*)$ durch

$$L^*(f_r) = \sum_{k=1}^{n} b_{kr} l_k \text{ definiert.}$$

Dann gilt für $1 \le j \le n$, $1 \le r \le m$:

$$(L^*(f_r))(v_j) = \sum_{k=1}^{n} b_{kr} l_k(v_j) = \sum_{k=1}^{n} b_{kr} \delta_{kj} = b_{jr}.$$

Also

$$\begin{aligned} b_{jr} &= (L^*(f_r))(v_j) = f_r(L(v_j)) = f_r\left(\sum_{i=1}^m a_{ij} w_i\right) \\ &= \sum_{i=1}^m a_{ij} f_r(w_i) = \sum_{i=1}^m a_{ij}\delta_{ri} = a_{rj}, \end{aligned}$$

d.h. $B = A^T$.

Bez.: $V^{**} := (V^*)^* = \mathrm{Hom}(V^*, K)$ heißt der <u>Bidualraum</u> <u>von</u> V.

(7.6) Fakt. Es existiert ein "natürlicher" injektiver Homomorphismus

$$h = h_V : V \to V^{**}$$

definiert durch: Für alle $v \in V$, $l \in V^*$ gilt

$$(h(v))(l) := l(v).$$

Bem.:

1) Ist $\dim V < \infty$, so ist $h : V \to V^{**}$ ein Isomorphismus, da $\dim V^{**} = \dim V^* = \dim V$.
2) h heißt "natürlich (oder kanonisch)", weil h unabhängig von irgendwelchen Wahlen (z.B. von Basen) definiert ist. In der Sprache der "Kategorien" läßt sich die Natürlichkeit von h mathematisch präzis formulieren.

Bew.: Injektivität von h: Ist $v \in V \setminus \{0\}$, so existiert ein $l \in V^*$ mit $l(v) \neq 0$ (vgl. den Beweis von (7.4)(i)). Dann gilt $(h(v))(l) = l(v) \neq 0$, also $h(v) \in V^{**} \setminus \{0\}$.

Bez.: Zu einem Untervektorraum U von V definieren wir

$$U^s := \{l \in V^* \mid \forall u \in U : l(u) = 0\}.$$

Zu einem Untervektorraum W von V^* definieren wir

$$W_s := \{v \in V \mid \forall l \in W : l(v) = 0\}.$$

Dann ist U^s Untervektorraum von V^* und W_s Untervektorraum von V, und es gilt

(7.7) Satz. Ist $\dim V < \infty$, so

(i) $\dim U + \dim U^s = \dim V$
(ii) $(U^s)^s = h(U) \subseteq V^{**}$
(iii) $\dim W + \dim W_s = \dim V$

Bem.: (7.7)(iii) kann als Präzisierung und Verallgemeinerung von (3.25) angesehen werden: Ist I ein homogenes lineares Gleichungssystem mit k Gleichungen für n Unbekannte (in K), so gilt $\dim L_I \geq n - k$. Jede Gleichung ist gegeben durch eine Linearform $l_1, \ldots, l_k \in (K^n)^*$. Wir setzen $W := \mathrm{span}\{l_1, \ldots, l_k\} \subseteq (K^n)^*$. Dann gilt $L_I = W_s$ und (7.7)(iii) impliziert: $\dim L_I = \dim W_s = n - \dim W \geq n - k$ mit "=" genau dann, wenn die "Gleichungen" $l_1, \ldots, l_k$ linear unabhängig sind. (Ausgedrückt mittels Matrizen folgt Entsprechendes aus (4.21), (4.22)).

Bew.:

(i) Mit der Bemerkung nach (3.13) ("Basisergänzungssatz") können wir eine Basis $\mathcal{G} = (v_1, \dots, v_n)$ von V finden, so daß für $k := \dim U$ gilt:

$$U = \operatorname{span}\{v_1, \dots, v_k\}.$$

Sei $\mathcal{G}^*$ die zu $\mathcal{G}$ duale Basis. Dann gilt mit (7.2):

$$\begin{aligned} l = \sum_{i=1}^{n} l(v_i) l_i \text{ liegt in } U^s &\Leftrightarrow l(v_i) = 0 \text{ für } i = 1, \dots, k \\ &\Leftrightarrow l = \sum_{i=k+1}^{n} l(v_i) l_i \Leftrightarrow l \in \operatorname{span}\{l_{k+1}, \dots, l_n\}. \end{aligned}$$

D.h. $U^s = \operatorname{span}\{l_{k+1}, \dots, l_n\}$ und speziell:

$$\dim U^s = n - k = \dim V - \dim U.$$

(ii) Zeige: $h(U) \subseteq (U^s)^s$. Sei $v \in U$ und $l \in U^s$. Dann gilt

$$(h(v))(l) = l(v) = 0,$$

also $h(v) \in (U^s)^s$. Nach (i) gilt $\dim(U^s)^s = \dim V^* - \dim U^s = \dim V^* - (\dim V - \dim U) = \dim U$. Da h injektiv ist, gilt $\dim h(U) = \dim U (= \dim(U^s)^s)$. Wegen $h(U) \subseteq (U^s)^s$, folgt $h(U) = (U^s)^s$.

(iii) Wir zeigen, daß $W^s = h(W_s)$ gilt, denn daraus folgt mit (i): $\dim W_s = \dim h(W_s) = \dim W^s = \dim V - \dim W$. Ist $w \in W_s$, so gilt für alle $l \in W$: $0 = l(w) = (h(w))(l)$. Daraus folgt $h(w) \in W^s$. Ist $\overline{v} \in W^s \subseteq V^{**}$, so existiert nach (7.6), Bem. 1), ein $v \in V$ mit $h(v) = \overline{v}$. Dann gilt für alle $l \in W$: $0 = \overline{v}(l) = (h(v))(l) = l(v)$. Also $v \in W_s$, und damit $\overline{v} = h(v) \in h(W_s)$.

(7.8) Satz. *Seien V und W K-Vektorräume, $L \in \operatorname{Hom}(V, W)$. Dann gilt:*

(i) $\ker L^* = (\operatorname{im} L)^s$

(ii) $\ker L = (\operatorname{im} L^*)_s$

Speziell: L *surjektiv* $\Leftrightarrow$ L^* *injektiv*

L^* *surjektiv* $\Rightarrow$ L *injektiv. Die Umkehrung gilt, falls* $\dim W < \infty$.

Bew.:

(i) Zeige: $\ker L^* \subseteq (\operatorname{im} L)^s$. Ist $l \in \ker L^* \subseteq W^*$, so gilt für alle $v \in V$:

$$0 = (L^*(l))(v) = l(L(v)), \text{ d.h. } l \in (\operatorname{im} L)^s.$$

Zeige: $(\operatorname{im} L)^s \subseteq \ker L^*$. Ist $l \in (\operatorname{im} L)^s$, so gilt für alle $v \in V$:

$$l(L(v)) = 0, \text{ also } (L^*(l))(v) = 0. \text{ Daraus folgt } l \in \ker L^*.$$

(ii) Analog zu (i).

"Speziell": vgl. Anwesenheitsaufgabe 4 auf Blatt 4a.

(7.9) Folgerung. Seien V und W K-Vektorräume, $\dim W < \infty$ und $L \in \mathrm{Hom}(V, W)$. Dann gilt: $\dim(\mathrm{im}\, L) = \dim(\mathrm{im}\, L^*)$.

Bew.:
$$\begin{aligned} \dim(\mathrm{im}\, L^*) &\overset{(4.8)}{=} \dim W^* - \dim(\ker L^*) \overset{(7.2)}{=} \dim W - \dim(\ker L^*) = \\ &\overset{(7.8)(\mathrm{i})}{=} \dim W - \dim((\mathrm{im}\, L)^s) \overset{(7.7)(\mathrm{i})}{=} \dim W - (\dim W - \dim(\mathrm{im}\, L)) \\ &= \dim(\mathrm{im}\, L) \end{aligned}$$

Eine der berühmtesten mathematischen Erkenntnisse der letzten 50 Jahre – der Atiyah-Singer-Indexsatz – hat damit zu tun, daß (7.9) ohne die Voraussetzung "$\dim W < \infty$" nicht gilt.

Bem.: (7.5) und (7.9) bilden den mathematischen Hintergrund für Satz (4.23) ("Zeilenrang = Spaltenrang"), d.h. $\mathrm{rg}(A) = \mathrm{rg}(A^T)$.

Zusammenhang: Dualraum $\leftrightarrow$ Bilinearformen

Zu einem K-Vektorraum V betrachten wir die Menge
$$B(V) = \{b \mid b : V \times V \to K \text{ bilinear}\}$$
der Bilinearformen auf V. Sie bildet (mit den wie üblich "punktweise"definierten Verknüpfungen) selbst einen K-Vektorraum. Jedes $b \in B(V)$ definiert Abbildungen $\lambda(b) \in \mathrm{Hom}(V, V^*)$ und $\rho(b) \in \mathrm{Hom}(V, V^*)$ durch: Für alle $v \in V$ ist $\lambda(b)(v) \in V^*$ durch
$$\lambda(b)(v) = b(v, \cdot)$$
definiert, d.h. $(\lambda(b)(v))(w) = b(v, w)$ für alle $w \in V$.

Analog ist $\rho(b) \in \mathrm{Hom}(V, V^*)$ definiert durch:

Für alle $v \in V$ gilt $\rho(b)(v) = b(\cdot, v)$, d.h.
$$(\rho(b)(v))(w) = b(w, v) \text{ für alle } w \in V.$$
Man prüft nach, daß in der Tat $\lambda(b), \rho(b) \in \mathrm{Hom}(V, V^*)$.

(7.10) Def.: $b \in B(V)$ heißt nicht ausgeartet, falls folgende Bedingungen (i) und (ii) gelten:

(i) Ist $v \in V$ und gilt $b(v, w) = 0$ für alle $w \in V$, so gilt $v = 0$.
(ii) Ist $v \in V$ und gilt $b(w, v) = 0$ für alle $w \in V$, so gilt $v = 0$.

Bsp.: Jedes Skalarprodukt ist eine nicht ausgeartete Bilinearform. ($\langle v, w\rangle = 0 \ \forall w \in V \Rightarrow \langle v, v\rangle = \|v\|^2 = 0 \Rightarrow v = 0$.)

Bem.: Die Bedingung (7.10)(i) ist äquivalent dazu, daß $\lambda(b) \in \mathrm{Hom}(V, V^*)$ injektiv ist, und (7.10)(ii) ist äquivalent dazu, daß $\rho(b)$ injektiv ist.

(7.11) Satz. *Die Abbildungen*
$$\lambda : B(V) \to \mathrm{Hom}(V, V^*), b \in B(V) \to \lambda(b) \in \mathrm{Hom}(V, V^*)$$

und
$$\rho(b) : B(V) \to \mathrm{Hom}(V, V^*), b \in B(V) \to \rho(b) \in \mathrm{Hom}(V, V^*)$$
sind natürliche K-Vektorraumisomorphismen.

(7.11) besagt, daß ein Homomorphismus $V \to V^*$ "nichts anderes" als eine (etwas anders geschriebene) Bilinearform auf V ist.

Bew.: Linearität von λ, z.B.:
$$\begin{aligned}(\lambda(b_1 + b_2))(v) &= (b_1 + b_2)(v, \cdot) = b_1(v, \cdot) + b_2(v, \cdot) \\ &= (\lambda(b_1))(v) + (\lambda(b_2))(v) = (\lambda(b_1) + \lambda(b_2))(v).\end{aligned}$$
Injektivität von λ: $b \neq 0 \to \exists v, w \in V$: $b(v, w) \neq 0 \Rightarrow$
$$(\lambda(b)(v))(w) \neq 0 \Rightarrow \lambda(b)(v) \neq 0 \Rightarrow \lambda(b) \neq 0.$$
Surjektivität von λ: Sei $L \in \mathrm{Hom}(V, V^*)$. Definiere $b : V \times V \to K$ durch $b(v, w) = (L(v))(w)$. Dann gilt $b \in B(V)$ und für alle $v \in V$:
$$\lambda(b)(v) = b(v, \cdot) = L(v), \text{ d.h. } \lambda(b) = L.$$
Bem.: Ist $b \in B(V)$, so ist $\lambda(b)^* \in \mathrm{Hom}(V^{**}, V^*)$, und es gilt für alle $v, w \in V$:
$$(\lambda(b)^*(h(v)))(w) = h(v)(\lambda(b)(w)) = b(w, v).$$

Ist $\dim V < \infty$, so können wir aus dieser Bemerkung ableiten, daß die Bedingungen (7.10)(i) und (ii) äquivalent sind: $b \in B(V)$ und (7.10)(i) gilt $\Leftrightarrow$ $\lambda(b) \in \mathrm{Hom}(V, V^*)$ ist injektiv $\overset{(7.9)}{\Leftrightarrow}$ $\lambda(b)^* \in \mathrm{Hom}(V^{**}, V^*)$ ist injektiv $\overset{\text{Bem.}}{\Leftrightarrow}$ (7.10)(ii) gilt für b.

(7.12) Folgerung. Sei $\dim V < \infty$ und $b \in B(V)$ nicht ausgeartete Bilinearform. Dann existiert zu jedem $l \in V^*$ genau ein $v \in V$, so daß für alle $w \in V$ gilt:
$$l(w) = b(v, w),$$
nämlich $v = \lambda(b)^{-1}(l)$. Analog existiert genau ein $v' \in V$, so daß für alle $w \in V$ gilt
$$l(w) = b(w, v'),$$
nämlich $v' = \rho(b)^{-1}(l)$.

Bew.: Da b nicht ausgeartet ist, sind $\lambda(b) \in \mathrm{Hom}(V, V^*)$ und $\rho(b) \in \mathrm{Hom}(V, V^*)$ injektiv, vgl. Bem. nach (7.10). Wegen $\dim V = \dim V^* < \infty$ sind $\lambda(b), \rho(b)$ Isomorphismen von V nach V^*. Ist $v := \lambda(b)^{-1}(l)$ und $w \in V$ beliebig, so gilt
$$b(v, w) = (\lambda(b)(v))(w) = (\lambda(b)(\lambda(b)^{-1}(l)))(w) = l(w).$$
Ist umgekehrt $v \in V$ und gilt für alle $w \in V$: $b(v, w) = l(w)$, so folgt
$$b(v, w) = (\lambda(b)(v))(w) = l(w),$$
d.h. $\lambda(b)(v) = l$, und damit $v = \lambda(b)^{-1}(l)$.

Nachtrag: Orientierung von $\mathbb{R}$-Vektorräumen.

Wir wollen nun (7.12) benutzen, um das Kreuzprodukt (auch Vektorprodukt genannt) zu definieren. Dazu benötigen wir den Begriff des "orientierten $\mathbb{R}$-Vektorraums". Das folgende ist eine Verallgemeinerung von (5.25), und zwar auf den Fall beliebiger $\mathbb{R}$-Vektorräume V, mit $1 \leq \dim V = n < \infty$.

(7.13) Def.: Zwei Basen $(v_1, \ldots, v_n)$, $(w_1, \ldots, w_n)$ von V heißen gleich orientiert, falls für den Endomorphismus $L \in \mathrm{End}(V)$, der durch $L(v_i) = w_i$, $1 \leq i \leq n$, definiert ist, gilt: $\det L > 0$.

Bem.: 1) $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_n)$ sind genau dann gleich orientiert, falls für eine ($\Rightarrow$ jede) Determinantenform D auf V gilt:

$$\frac{D(w_1, \ldots, w_n)}{D(v_1, \ldots, v_n)} > 0, \text{ vgl. (5.19).}$$

2) "gleich orientiert" ist eine Äquivalenzrelation auf der Menge der geordneten Basen V mit zwei Äquivalenzklassen.

(7.14) Def.: Eine Orientierung von V ist die Auswahl einer der beiden Äquivalenzklassen von geordneten Basen von V. Die Basen in der ausgewählten Äquivalenzklasse heißen positiv orientiert.

Bsp.: Die "übliche Orientierung" des $\mathbb{R}^n$ besteht in der Auswahl der Äquivalenzklasse, die die Standardbasis $(e_1, \ldots, e_n)$ enthält, vgl. (5.25).

Bez.: Sei $(V, \langle,\rangle)$ orientierter euklidischer Vektorraum, $0 < \dim V = n < \infty$. Dann existiert genau eine Determinantenform D auf V, so daß für eine ($\Rightarrow$ jede) positiv orientierte ONB von V gilt: $D(v_1, \ldots, v_n) = 1$. Dieses D heißt die kanonische Determinantenform von V.

Bem.: Ist $(v_1, \ldots, v_n)$ positiv orientierte ONB von V, so kann man für beliebige $(w_1, \ldots, w_n) \in V^n$ den Wert $D(w_1, \ldots, w_n)$ wie folgt berechnen: Ist $w_j = \sum\limits_{i=1}^{n} a_{ij} v_i$ und $A := (a_{ij})_{1 \leq i,j \leq n}$, so gilt $D(w_1, \ldots, w_n) = \det A$. Das liegt daran, daß diese Formel eine Determinantenform definiert, die auf der gegebenen, positiv orientierten ONB $(v_1, \ldots, v_n)$ den Wert 1 hat (da in diesem Fall $A = E_n$ gilt).

Speziell ist im Fall $V = \mathbb{R}^n$ mit dem Standardskalarprodukt und der üblichen Orientierung, die kanonische Determinantenform gerade die "Standarddeterminantenform" D_0, vgl. (5.13).

Das Kreuzprodukt

(7.15) Def.: Sei $(V, \langle,\rangle)$ orientierter euklidischer Vektorraum, $3 \leq \dim V = n < \infty$, mit kanonischer Determinantenform D. Zu je $n-1$ Vektoren $w_1, \ldots, w_{n-1}$ in V existiert genau ein Vektor $w \in V$, so daß für alle $v \in V$ gilt:

$$(*) \qquad D(w_1, \ldots, w_{n-1}, v) = \langle w, v \rangle.$$

Dieses w heißt das Kreuzprodukt von $w_1, \ldots, w_{n-1}$ und wird durch das Symbol $w_1 \times \ldots \times w_{n-1}$ bezeichnet.
Begründung: Die Abbildung $v \in V \to D(w_1, \ldots, w_{n-1}, v) \in \mathbb{R}$ ist eine Linearform auf V. Da $\langle , \rangle$ eine nichtausgeartete Bilinearform ist, folgt Existenz und Eindeutigkeit eines Vektors w, der $(*)$ für alle $v \in V$ erfüllt, aus Folgerung (7.12).

Bem.: Es ist nur das Kreuzprodukt $w_1 \times \ldots \times w_{n-1}$ von $n-1$ Vektoren ($n = \dim V$) definiert, d.h. für zwei Vektoren $w_1, w_2 \in \mathbb{R}^4$ ist $w_1 \times w_2$ nicht definiert!

Eigenschaften des Kreuzprodukts:

(i) Die Abbildung $(w_1, \ldots, w_{n-1}) \in V \times \ldots \times V \to w_1 \times \ldots \times w_{n-1} \in V$ ist multilinear und alternierend. Speziell gilt: $w_1, \ldots, w_{n-1}$ linear abhängig $\Rightarrow w_1 \times \ldots \times w_{n-1} = 0$.
Z.B. beweist man wie folgt die Additivität im 1. Argument:

$$\begin{aligned}
&\langle (w_1 + w_1') \times w_2 \times \ldots \times w_{n-1}, \cdot \rangle = D(w_1 + w_1', w_2, \ldots, w_{n-1}, \cdot) \\
&\quad = D(w_1, w_2, \ldots, w_{n-1}, \cdot) + D(w_1', w_2, \ldots, w_{n-1}, \cdot) \\
&\quad = \langle w_1 \times w_2 \times \ldots \times w_{n-1}, \cdot \rangle + \langle w_1' \times w_2 \times \ldots \times w_{n-1}, \cdot \rangle \\
&\quad = \langle w_1 \times \ldots \times w_{n-1} + w_1' \times w_2 \times \ldots \times w_{n-1}, \cdot \rangle
\end{aligned}$$

(ii) $L \in O(V) \Rightarrow L(w_1) \times \ldots \times L(w_{n-1}) = \det L \cdot L(w_1 \times \ldots \times w_{n-1})$.
Bew.: $\langle L(w_1) \times \ldots \times L(w_{n-1}), v \rangle = D(L(w_1), \ldots, L(w_n), L(L^{-1}(v)))$

$$\begin{aligned}
&= \det L \, D(w_1, \ldots, w_{n-1}, L^{-1}(v)) = \det L \cdot \langle w_1 \times \ldots \times w_{n-1}, L^{-1}(v)\rangle \\
&\overset{L \in O(V)}{=} \langle \det L \cdot L(w_1 \times \ldots \times w_{n-1}), v \rangle
\end{aligned}$$

(iii) "Geometrische Definition des Kreuzprodukts": Sind $w_1, \ldots, w_{n-1}$ linear unabhängig, so ist $w_1 \times \ldots \times w_{n-1}$ der eineutig bestimmte Vektor mit:
(a) $w_1 \times \ldots \times w_{n-1} \in (\operatorname{span}\{w_1, \ldots, w_{n-1}\})^\perp$,
(b) $(w_1, \ldots, w_{n-1}, w_1 \times \ldots \times w_{n-1})$ ist positiv orientierte Basis von V, und
(c) $\| w_1 \times \ldots \times w_{n-1} \| = \operatorname{vol}_{n-1}^{\langle , \rangle}(P(w_1, \ldots, w_{n-1}))$.

Wir zeigen, daß $w_1 \times \ldots \times w_{n-1}$ die Eigenschaften (a)–(c) hat:
zu (a): $\langle w_i, w_1 \times \ldots \times w_{n-1} \rangle = D(w_1, \ldots, w_{n-1}, w_i) = 0$ für $1 \leq i \leq n-1$.
zu (b):

$$\begin{aligned}
D(w_1, \ldots, w_{n-1}, w_1 \times \ldots \times w_{n-1}) &= \langle w_1 \times \ldots \times w_{n-1}, w_1 \times \ldots \times w_{n-1} \rangle \\
&= \| w_1 \times \ldots \times w_{n-1} \|^2 .
\end{aligned}$$

zu (c): Zeige zunächst: $w_1, \ldots, w_{n-1}$ linear unabhängig $\Rightarrow w := w_1 \times \ldots \times w_{n-1} \neq 0$.

Ergänze $w_1, \ldots, w_{n-1}$ zu einer Basis $(w_1, \ldots, w_{n-1}, v)$ von V.
Dann gilt $0 \neq D(w_1, \ldots, w_{n-1}, v) = \langle w_1 \times \ldots \times w_{n-1}, v \rangle$, also $w = w_1 \times \ldots \times w_{n-1} \neq 0$.

Nun ist die Abbildung, die $v_1, \ldots, v_{n-1} \in \operatorname{span}\{w_1, \ldots, w_{n-1}\}$ die reelle Zahl $D\left(v_1, \ldots, v_{n-1}, \frac{w}{\|w\|}\right)$ zuordnet, eine normierte Determinantenform auf $\operatorname{span}\{w_1, \ldots, w_{n-1}\}$, also

$$\operatorname{vol}_{n-1}^{\langle , \rangle}(P(w_1, \ldots, w_{n-1})) = D\left(w_1, \ldots, w_{n-1}, \frac{w}{\|w\|}\right) \overset{(b)}{=} \| w_1 \times \ldots \times w_{n-1} \| .$$

Umgekehrt sieht man leicht, daß ein Vektor durch die Eigenschaften (a)–(c) eindeutig bestimmt ist.

Schließlich leiten wir eine explizite Formel her, die es erlaubt, $w_1 \times \ldots \times w_{n-1}$ aus den Komponenten von $w_1, \ldots, w_{n-1}$ bezüglich einer positiv orientierten ONB zu berechnen.

(7.16) Fakt. Ist $(v_1, \ldots, v_n)$ positiv orientierte ONB von V und $w_j = \sum\limits_{i=1}^{n} a_{ij} v_i$ für $1 \leq j \leq n-1$, so gilt

$$w_1 \times \ldots \times w_{n-1} = \sum_{i=1}^{n} \big((-1)^{n+i} \det A_i)\big) v_i,$$

wobei $A^i \in \mathbb{R}^{(n-1)\times(n-1)}$ aus $A := (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n-1}} \in \mathbb{R}^{n\times(n-1)}$ durch Streichen der i'ten Zeile entsteht.

Bsp.: Im Fall von $V = \mathbb{R}^3$ mit der üblichen Orientierung und dem Standardskalarprodukt betrachten wir die positiv orientierte ONB (e_1, e_2, e_3) und

$$\begin{aligned} w_1 &=: a = (a_1, a_2, a_3) = \sum_{i=1}^{3} a_i\, e_i \in \mathbb{R}^3 \text{ und} \\ w_2 &=: b = (b_1, b_2, b_3) = \sum_{i=1}^{3} b_i\, e_i \in \mathbb{R}^3. \end{aligned}$$

Dann ist $A \in \mathbb{R}^{3\times 2}$ die Matrix

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_2 \end{pmatrix}$$

und es gilt: $a \times b = \sum_{i=1}^{3} ((-1)^{i+3} \det A_i)\, e_i = (a_2 b_3 - a_3 b_1, -a_1 b_3 + a_3 b_1, a_1 b_2 - a_2 b_1)$.

Bew. von (7.16): Für alle $x = \sum\limits_{i=1}^{n} x_i v_i \in V$ gilt aufgrund der Bem. nach (7.14):

$$\begin{aligned} D(w_1, \ldots, w_{n-1}, x) &= \begin{vmatrix} a_{11} & \ldots & a_{1n-1} & x_1 \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ a_{n1} & \ldots & a_{nn-1} & x_n \end{vmatrix} \overset{\substack{\text{Entwicklung nach} \\ \text{der letzten Spalte}}}{=} \sum_{i=1}^{n} (-1)^{n+i} \det A_i x_i \\ &= \left\langle \sum_{i=1}^{n} ((-1)^{n+i} \det A_i)\, v_i, x \right\rangle. \end{aligned}$$

Also: $w_1 \times \ldots \times w_{n-1} = \sum\limits_{i=1}^{n} ((-1)^{n+i} \det A_i)\, v_i$.

Bsp.: Es sei $V = \mathbb{R}^{n+1}$ mit üblicher Orientierung und Standardskalarprodukt. Es sei $e_1, \ldots, e_n$ die Standardbasis des $\mathbb{R}^n$ und $s_1, \ldots, s_n$ reelle Zahlen. Wir betrachten die Vektoren $w_1 := (e_1, s_1), \ldots, w_n := (e_n, s_n)$ im $\mathbb{R}^{n+1}$. Dann gilt:

$$\mathrm{vol}_n^{\langle,\rangle}(P(w_1, \ldots, w_n)) = \sqrt{1 + \sum_{i=1}^{n} (s_i)^2}$$

(vgl. Blatt 2, Aufgabe 4 für die Fälle $n = 2, 3$).

Bew.: Nach (c) und (7.16) gilt:

$$\mathrm{vol}_n(P(w_1, \ldots, w_n)) \overset{(c)}{=} \| w_1 \times \ldots \times w_n \| \overset{(7.16)}{=} \sqrt{\sum_{i=1}^{n+1} (\det A_i)^2}, \text{ wobei}$$

A_i aus der $((n+1) \times n)$-Matrix

$$A = \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & \ldots & \ldots & 0 \\ \vdots & & \ddots & & \\ \vdots & & & \ddots & \\ 0 & & \ldots & \ldots & 1 \\ s_1 & s_2 & \ldots & \ldots & s_n \end{pmatrix}$$

durch Streichen der i'ten Zeile entsteht. Offensichtlich gilt $\det A_{n+1} = 1$ und $|\det A_i| = |s_i|$ für $1 \leq i \leq n$.

Nachtrag: Quotientenraum (manchmal auch Faktorraum genannt).

Vor der Beschäftigung mit diesem Nachtrag ist es gut, sich (1.23), (1.24) und die darauf folgenden Tatsachen über Äquivalenzrelationen ins Gedächtnis zu rufen.

Sei V ein K-Vektorraum und U Untervektorraum von V. Wir definieren eine Äquivalenzrelation $\sim_U$ auf V durch:

$$v \sim_U w \Leftrightarrow v - w \in U.$$

Die Transitivität von $\sim_U$ sieht man so ein:

$$v \sim_U w \text{ und } w \sim_U z \Rightarrow v - w \in U \text{ und } w - z \in U \overset{U \text{ Untervektorraum}}{\Longrightarrow} (v - w) + (w - z) = v - z \in U \Rightarrow v \sim_U z.$$

Bem.: Die Äquivalenzklasse von $v \in V$ bezüglich $\sim_U$ ist genau der affine Unterraum $v+U = \{v + u \mid u \in U\}$, vgl. (3.26) und die Abschnitte vor (3.26).

Bew.: (i) Ist $w \in v + U$, d.h. $w = v + u$ für ein $u \in U$, so gilt $w - v = u \in U$, also $w \sim_U v$. Deshalb ist $v + U$ in der Äquivalenzklasse von v enthalten.
(ii) Ist $z \in V$ und $z \sim_U v$, so gilt $z - v =: u \in U$, also $z = v + u \in v + U$. Das zeigt, daß die Äquivalenzklasse von v in $v + U$ enthalten ist.

Es ist für das Folgende ganz wichtig, im Gedächtnis zu behalten, daß in der Darstellung einer Äquivalenzklasse in der Form $v+U$ das Element $v \in V$ nicht eindeutig bestimmt ist (es sei denn $U = \{0\}$), denn es gilt:

$$v + U = w + U \Leftrightarrow v \sim_U w \Leftrightarrow v - w \in U.$$

Man nennt jedes Element aus $v+U$ einen Repräsentanten der Äquivalenzklasse $v+U$.

Wir bezeichnen mit

$$V/U := \{v + U \mid v \in V\}$$

die Menge der Äquivalenzklassen von $\sim_U$ und mit

$$\pi : V \to V/U, \pi(v) := v + U$$

die "kanonische Projektion".

(7.17) Fakt: Es gibt genau eine K-Vektorraumstruktur auf der Menge V/U, so daß $\pi : V \to V/U$ ein Homomorphismus ist.

Bew.: Die Eindeutigkeit der K-Vektorraumstruktur folgt aus der Surjektivität von π. Denn sind $\alpha \in K$ und $v+U$, $w+U \in V/U$, so gilt $\pi(v) = v+U$, $\pi(w) = w+U$. Ist nun π ein Homomorphismus, so muß für das (noch zu definierende) Produkt $\alpha(v+U) \in V/U$ gelten:

$$\alpha(v+U) = \alpha\pi(v) \overset{\pi \in \mathrm{Hom}(V,V/U)}{=} \pi(\alpha v) = (\alpha v) + U.$$

Und ebenso:

$$(v+U) + (w+U) = \pi(v) + \pi(w) = \pi(v+w) = (v+w) + U.$$

Wenn π ein Homomorphismus werden soll, müssen wir also die Gleichungen

$$(*) \qquad \alpha(v+U) = (\alpha v) + U$$

und

$$(**) \qquad (v+U) + (w+U) = (v+w) + U$$

zur Definition der in den Gleichungen links stehenden Operationen verwenden. Es ist nicht klar, daß das möglich ist, denn es besteht die Gefahr, daß etwa in $(*)$ $v + U = v' + U$ gilt, aber $(\alpha v)+U \neq (\alpha v')+U$. Wir müssen zeigen, daß das nicht passieren kann. Man sagt dazu, wir müssen zeigen, daß die "Definition" $(*)$ unabhängig von der Wahl des Repräsentanten v von $v+U$ ist, und Analoges für $(**)$: aus $v+U = v'+U$ folgt $v - v' \in U$, und damit $\alpha(v-v') = \alpha v - \alpha v' \in U$, also $\alpha v \sim_U \alpha v'$ oder $(\alpha v)+U = (\alpha v')+U$. Aus $v+U = v'+U$ und $w+U = w'+U$ folgt $v - v' \in U$, $w - w' \in U$, und damit $(v-v') + (w-w') = (v+w) - (v'+w') \in U$, also $(v+w) \sim_U (v'+w')$ oder $(v+w)+U = (v'+w')+U$.

Das zeigt, daß wir $(*)$ und $(**)$ benützen können, um eine Multiplikation mit Skalaren ($\in K$) und eine Addition auf V/U zu definieren. Wir bemerken noch, daß wir beim Beweis der Repräsentantenunabhängigkeit von $(**)$ die Kommutativität der Addition in V benutzt haben. Wir werden später eine analoge Quotientenkonstruktion für Gruppen kennenlernen, und dabei tritt an der entsprechenden Stelle für nichtabelsche Gruppen ein Problem auf.

Nun sind natürlich noch die Vektorraumaxiome für V/U mit den durch $(*)$ und $(**)$ definierten Operationen nachzuprüfen. Das ist länglich, aber leicht. Wir bemerken noch, daß

$$U = \pi(0) = \pi(u) = u + U \qquad (\text{für alle } u \in U)$$

das 0-Element von V/U ist.

Bsp.: $V = \mathbb{R}^3$ und $U = \operatorname{span}\{(1,1,1)\} = \{(x,x,x) \mid x \in \mathbb{R}\}$.

$V/U =$ Menge der affinen Geraden, die parallel zu U sind.

$\pi : \mathbb{R}^3 \to \mathbb{R}^3/U$ ordnet einen Punkt $v \in \mathbb{R}^3$ die v enthaltende Gerade $v + U$ dieser Parallelschar zu. Da jede dieser Geraden $\mathbb{R}^2 \times \{0\}$ in genau einem Punkt trifft, ist $(\pi \mid \mathbb{R}^2 \times \{0\}) : \mathbb{R}^2 \times \{0\} \to \mathbb{R}^3/U$ ein bijektiver Homomorphismus, also ein Isomorphismus, d.h. die (durch $(**)$ definierte) Summe von zwei Geraden kann man wie folgt erhalten: Man bestimmt die Schnittpunkte der zwei Geraden mit $\mathbb{R}^2 \times \{0\}$ und erhält die "Summengerade" als die zu U parallele Gerade durch die Summe (in $\mathbb{R}^2 \times \{0\}$) der Schnittpunkte. Es ist aber wichtiger die Struktur der Quotientenkonstruktion zu verstehen, als dieses explizite Beispiel.

Es ist nicht leicht zu erklären, warum diese Quotientenkonstruktion wichtig ist. Sie macht aus V ein vergröbertes Abbild V/U, in dem alles, was den Unterraum U betrifft, "vergessen" wird. Wir werden etwa in (7.19) sehen, daß zu jedem Homomorphismus $L \in \operatorname{Hom}(V, W)$ ein "natürlicher" Homomorphismus $\overline{L} : V/\ker L \to W$ mit $\overline{L} \circ \pi = L$ gehört, der injektiv ist. Es ist wichtig zu wissen, daß analoge Quotientenkonstruktionen für alle algebraischen Strukturen (Gruppen, Ringe, Moduln) möglich und oft nützlich sind. Hier wird kurz noch der Fall einer (möglicherweise nichtabelschen) Gruppe $(G, \cdot)$ skizziert:

Ist H eine Untergruppe von G, so definiert

$$g_1 \sim_H g_2 \Leftrightarrow g_2^{-1} g_1 \in H$$

eine Äquivalenzrelation $\sim_H$ auf G, deren Äquivalenzklassen gerade die "Linksnebenklassen"

$$gH = \{gh \mid h \in H\}$$

von H sind. Man setzt wieder $G/H = \{gH \mid g \in G\}$ und $\pi : G \to G/H$, $\pi(g) = gH$. Versucht man nun, in Analogie zu (7.17), G/H so mit einer Gruppenstruktur zu versehen, daß π ein Homomorphismus wird, so kann man die (unangenehme) Überraschung erleben, daß das nicht geht: Man würde gern $(g_1H) \circ (g_2H)$ durch $(g_1g_2)H$ definieren, aber es stellt sich heraus, daß (im nichtabelschen Fall) $(g_1g_2)H$ sehr wohl von der Wahl der Repräsentanten g_1 von g_1H und g_2 vn g_2H abhängen kann, d.h. aus $g_1H = g_1'H$ und $g_2H = g_2'H$ folgt nicht notwendig $(g_1g_2)H = (g_1'g_2')H$. Es ist nicht schwer einzusehen, daß das genau dann klappt, wenn für alle $g \in G$

$$gHg^{-1} = H$$

gilt, wobei $gHg^{-1} := \{ghg^{-1} \mid h \in H\}$. Untergruppen H mit dieser Eigenschaft heißen normal, und genau für solche normale Untergruppen H ist es möglich, auf G/H eine

Gruppenstruktur zu definieren, so daß $\pi : G \to G/H$ homomorph ist. Ist G abelsch, so gilt natürlich $ghg^{-1} = gg^{-1}h = h$, und alle Untergruppen von G sind normal. Zu Beginn von Kapitel 2 haben wir (im Prinzip) auf diese Art aus $(G, \cdot) = (\mathbb{Z}, +)$ und einer Zahl $p \in \mathbb{N}$ die endliche zyklische Gruppe $(\mathbb{Z}_p, +)$ konstruiert: Als Untergruppe H nehmen wir

$$p\,\mathbb{Z} = \{n \in \mathbb{Z} \mid p \text{ teilt } n\}$$

und erhalten $\mathbb{Z}_p$ als $\mathbb{Z}/p\,\mathbb{Z}$.

Zurück zu den K-Vektorräumen V:

(7.18) Fakt: (i) Sind U und W komplementäre Untervektorräume von V, d.h. gilt $V = U \oplus W$, so ist $\pi|W : W \to V/U$ ein Isomorphismus.
(ii) Ist $\dim U < \infty$, so gilt $\dim(V/U) = \dim V - \dim U$.

Bew.: (i) Surjektivität von $\pi|W$: Sei $v + U$ ein beliebiges Element von V/U. Dann besitzt v eine (eindeutige) Darstellung als $v = u + w$ mit $u \in U$, $w \in W$. Wegen $\pi(u) = 0 \in V/U$ gilt: $v + U = \pi(v) = \pi(u + w) = \pi(u) + \pi(w) = \pi(w)$.

Injektivität von $\pi|W$: Sei $w \in W$ und $\pi(w) = 0 \in V/U$. Dann gilt $w + U = U$, d.h. $w \in U$. Also $w \in W \cap U$. Wir haben aber vorausgesetzt, daß $U \oplus W$ eine direkte Summe ist, d.h. daß $U \cap W = \{0\}$ gilt, vgl. (3.19). Also folgt aus $w \in W \cap U$, daß $w = 0$ gilt. Das zeigt $\ker(\pi|W) = \{0\}$, d.h. $\pi|W$ ist injektiv.

(ii) Nach (3.21) existiert ein zu U komplementärer Untervektorraum W von V, und es gilt $\dim U + \dim W = \dim V$. Nach (i) gilt $\dim W = \dim(V/U)$, und wegen $\dim U < \infty$ kann man die Gleichung auch als $\dim(V/U) = \dim W = \dim V - \dim U$ schreiben.

(7.19) Isomorphiesatz. *Zu $L \in \operatorname{Hom}(V, W)$ existiert genau ein Homomorphismus $\overline{L} \in \operatorname{Hom}(V/\ker L, W)$, für den $L = \overline{L} \circ \pi$ gilt. $\overline{L}$ ist injektiv und ein Isomorphismus von $V/\ker L$ auf $\operatorname{im} L \subseteq W$.*

Bem.: Ist $\dim(\ker L) < \infty$, so folgt aus der letzten Aussage, daß

$$\dim(\operatorname{im} L) = \dim(V/\ker L) \overset{(7.18)}{=} \dim V - \dim(\ker L)$$

gilt. Das ist ein nicht wirklich neuer Beweis des Dimensionssatzes (4.8).

Bew.: Da $\pi : V \to V/\ker L$ surjektiv ist, existiert höchstens eine Abbildung $\overline{L} : V|\ker L \to W$ mit $\overline{L} \circ \pi = L$. Wenn $\overline{L}$ existiert, muß $\overline{L}$ für alle $v + \ker L \in V/\ker L$ durch

$$(*) \qquad \overline{L}(v + \ker L) = \overline{L}(\pi(v)) = L(v)$$

gegeben sein und wegen $L|\ker L = 0$ ist $(*)$ in der Tat unabhängig von der Wahl des Repräsentanten der Äquivalenzklasse $v + \ker L$. Es gilt dann für alle $\alpha \in K$, $v_1 + \ker L \in$

$V/\ker L$, $v_2 + \ker L \in V/\ker L$:

$$\begin{aligned}
\overline{L}(\alpha(v_1 + \ker L)) &= \overline{L}((\alpha v_1) + \ker L) = L(\alpha v_1) = \alpha L(v_1) \\
&= \alpha \overline{L}(v_1 + \ker L) \text{ und} \\
\overline{L}((v_1 + \ker L) + (v_2 + \ker L)) &= \overline{L}((v_1 + v_2) + \ker L) = L(v_1 + v_2) = L(v_1) + L(v_2) \\
&= \overline{L}(v_1 + \ker L) + \overline{L}(v_2 + \ker L).
\end{aligned}$$

Das zeigt, daß $\overline{L} : V/\ker L \to W$ ein Homomorphismus ist. Ist $v + \ker L \in V/\ker L$ und gilt $\overline{L}(v + \ker L) = 0$, so folgt $L(v) = 0$, d.h. $v \in \ker L$ und damit $v + \ker L = 0 \in V/\ker L$. Das zeigt $\ker \overline{L} = \{0\}$, d.h. $\overline{L}$ ist injektiv. Da $\overline{L}(V/\ker L) = \overline{L}(\pi(V)) = L(V) = \operatorname{im} L$ gilt, ist $\overline{L}$ als Abbildung von $V/\ker L$ nach $\operatorname{im} L$ auch surjektiv, also ein Isomorphismus von $V/\ker L$ auf $\operatorname{im} L$.

8. Polynomringe

Das Umgehen mit Polynomen, d.h. mit Ausdrücken der Form

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

ist aus der Schule vertraut, falls die Koeffizienten $a_0, \ldots, a_n$ ganze oder rationale oder reelle Zahlen sind. In diesen Fällen kann man das Polynom mit einer Abbildung

$$\tilde{p} : \mathbb{R} \to \mathbb{R},\ \tilde{p}(x) := \sum_{i=0}^{n} a_i x^i$$

identifizieren. Es gilt dann $\tilde{p}(0) = a_0, (\tilde{p})'(0) = a_1, (\tilde{p})''(0) = 2a_2$ und die k'te Ableitung von $\tilde{p}$ ausgewertet an der Stelle $x = 0$ ist gerade $k!\,a_k$, d.h. die Funktion $\tilde{p}$ bestimmt die $a_0, \ldots, a_n$ eindeutig. Wenn man nun Polynome mit Koeffizienten a_i aus einem beliebigen Körper K betrachten will, entsteht im Fall von endlichen Körpern folgende Schwierigkeit: es „existieren" unendlich viele verschiedene Polynome, z.B. die „Monome" $1, x, x^2, \ldots, x^n, \ldots$, aber nur endlich viele verschiedene Abbildung $\tilde{p} : K \to K$. Ein Polynom $\sum\limits_{i=0}^{n} a_i x^i$ mit $a_i \in K$ kann also nicht durch die zugehörige Abbildung $\tilde{p} : K \to K$, $\tilde{p}(x) := \sum\limits_{i=0}^{n} a_i x^i$ eindeutig bestimmt sein. Diese Tatsache ruft die Frage hervor, was denn dann ein Polynom mit Koeffizienten $a_i \in K$ wirklich „ist" (oder besser „sein soll"), eine Frage, die wir eigentlich schon bei der Einführung des charakteristischen Polynoms (nach (5.24)) hätten stellen sollen. Das Ziel dieses Kapitels ist eine befriedigende Antwort auf diese Frage. Dann werden wir noch die Polynomdivision (mit Rest) kennenlernen und uns mit dem Zusammenhang zwischen Nullstellen von Polynomen und dem (multiplikativen) „Abspalten" von Linearfaktoren beschäftigen.

Zu einem beliebigen Körper K betrachten wir die Menge

$$K^{\mathbb{N}} := \{f \mid f : \mathbb{N} \to K\},$$

d.h. ein Element $f \in K^{\mathbb{N}}$ ist eine Folge $f = (f_0, f_1, \ldots, f_n, \ldots)$ mit $f_n \in K$ für alle $n \in \mathbb{N}$.

(8.1) Def.: Zu $f, g \in K^{\mathbb{N}}$, $a \in K$ definieren wir

(i) $f + g \in K^{\mathbb{N}}$ durch $f + g := (f_0 + g_0, f_1 + g_1, \ldots, f_n + g_n, \ldots)$
(ii) $af \in K^{\mathbb{N}}$ durch $af := (af_0, af_1, \ldots, af_n, \ldots)$
(iii) $f \cdot g \in K^{\mathbb{N}}$ durch $f \cdot g := ((f \cdot g)_0, (f \cdot g)_1, \ldots, (f \cdot g)_n, \ldots)$,
wobei $(f \cdot g)_0 := f_0 \cdot g_0, (f \cdot g)_1 := f_0 \cdot g_1 + f_1 \cdot g_0, (f \cdot g)_2 := f_0 \cdot g_2 + f_1 \cdot g_1 + f_2 \cdot g_0$ und $(f \cdot g)_n := \sum\limits_{j=0}^{n} f_j \cdot g_{n-j} = \sum\limits_{\substack{(j,k) \in \mathbb{N} \times \mathbb{N} \\ j+k=n}} f_j \cdot g_k$.

Bem.: 1) Die Ausdrücke in (iii) sind den Ausdrücken nachgebildet, die beim „schulmäßigen" Multiplizieren von Polynomen auftreten:

$$(a_0 + a_1x + \ldots + a_nx^n) \cdot (b_0 + b_1x + \ldots + b_mx^m) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \ldots$$

2) Für $a \in K$, $f \in K^{\mathbb{N}}$ gilt $af = (a, 0, \ldots, 0, \ldots) \cdot f$.

(8.2) Fakt.

(i) $(K^{\mathbb{N}}, +)$ ist ein ∞-dimensionaler K-Vektorraum mit dem Nullelement $0 = (0, \ldots, 0, \ldots) \in K^{\mathbb{N}}$.
(ii) $(K^{\mathbb{N}}, +, \cdot)$ ist ein kommutativer Ring mit $1 = (1, 0, \ldots, 0, \ldots) \in K^{\mathbb{N}}$.

Bem.: Die Definition des „kommutativen Rings mit 1" unterscheidet sich von der des Körpers (vgl. (2.4)) nur dadurch, daß auf die Forderung der Existenz von multiplikativen Inversen verzichtet wird. (Man fordert üblicherweise auch nicht, daß $1 \neq 0$ gilt, aber das ist für $K^{\mathbb{N}}$ natürlich erfüllt).

Bez.: $(K^{\mathbb{N}}, +, \cdot)$ heißt der Ring der formalen Potenzreihen über K, meist bezeichnet durch $K[[x]]$.

Der Nachweis der in (8.2) behaupteten Eigenschaften ist leicht. Am längsten dauert der Nachweis der Assoziativität der Multiplikation, den man wie folgt durchführen kann: Seien $f, g, h \in K^{\mathbb{N}}$ und $n \in \mathbb{N}$. Dann gilt:

$$\begin{aligned} ((f \cdot g) \cdot h)_n &= \sum_{j+k=n} (f \cdot g)_j h_k = \sum_{j+k=n} \left(\sum_{l+m=j} f_l \cdot g_m \right) \cdot h_k = \sum_{l+m+k=n} f_l \cdot g_m \cdot h_k \\ (f \cdot (g \cdot h))_n &= \sum_{l+j=n} f_l \cdot (g \cdot h)_j = \sum_{l+j=n} f_l \cdot \left(\sum_{m+k=j} g_m \cdot h_k \right) = \sum_{l+m+k=n} f_l \cdot g_m \cdot h_k \end{aligned}$$

(8.3) Fakt. Die Teilmenge

$$K[x] = \{ f \in K^{\mathbb{N}} \mid f_n \neq 0 \text{ nur für endlich viele } n \in \mathbb{N} \}$$

ist abgeschlossen bezüglich $+$ und $\cdot$ und enthält $1 = (1, 0, \ldots, 0, \ldots)$, und ist ein Unterring (mit 1) von $(K^{\mathbb{N}}, +, \cdot)$.

Die Abgeschlossenheit bzgl. $\cdot$ folgt aus dem Beweis zu (8.5).

(8.4) Def.:

(i) $(K[x], +, \cdot)$ heißt der Polynomring von K.
(ii) Ist $f \in K[x] \setminus \{0\}$, so heißt

$$\operatorname{grad} f := \max\{ i \in \mathbb{N} \mid f_i \neq 0 \} \in \mathbb{N}$$

der Grad von f. Ist $f = 0 \in K[x]$, so setzen wir $\operatorname{grad} f = -\infty$.

(8.5) Fakt. Für alle $f, g \in K[x]$ gilt:

$$\operatorname{grad}(f \cdot g) = \operatorname{grad} f + \operatorname{grad} g.$$

Speziell folgt: $f \cdot g = 0 \Rightarrow f = 0$ oder $g = 0$.

Zu letzterer Eigenschaft sagt man, $(K[x], +, \cdot)$ sei „nullteilerfrei".

Bew.: Ist $f = 0$ oder $g = 0$, so $f \cdot g = 0$, also

$$\operatorname{grad}(f \cdot g) = -\infty = \operatorname{grad} f + \operatorname{grad} g.$$

Ist $\operatorname{grad} f = m \in \mathbb{N}$, $\operatorname{grad} g = n \in \mathbb{N}$, so gilt $f_m \neq 0$, $g_n \neq 0$ und $f_i = 0$ für $i > m$, $g_j = 0$ für $j > n$. Daraus folgt

$$(f \cdot g)_{m+n} = \sum_{\substack{(i,j) \in \mathbb{N} \times \mathbb{N} \\ i+j=m+n}} f_i g_j = f_m \cdot g_n \neq 0$$

und $(f \cdot g)_k = 0$ für $k > m + n$. Also

$$\operatorname{grad}(f \cdot g) = m + n = \operatorname{grad} f + \operatorname{grad} g.$$

Um zur üblichen Darstellung von Polynomen zu kommen, führen wir folgende Bezeichnung ein:

$$x := (0, 1, 0, \ldots, 0, \ldots) \in K[x].$$

Außerdem definieren wir wie üblich $x^0 := (1, 0, \ldots, 0, \ldots)$ und rekursiv $x^{n+1} := x \cdot x^n$ für $n \in \mathbb{N}$.

In diesem Zugang zu den Polynomen ist also x keine „Variable“, sondern ein festes Element des Rings $K[x]$.

(8.6) Fakt. Die Menge $\{x^n \mid n \in \mathbb{N}\}$ ist eine Basis des K-Vektorraums $K[x]$. Es gilt $x^n = (\underbrace{0, \ldots, 0}_{n\text{-mal}}, 1, 0, \ldots) \in K[x]$.

Bew.: Die erste Behauptung folgt aus der zweiten. Die zweite beweisen wir durch Induktion. Die Gleichung $x^0 = (1, 0, \ldots, 0, \ldots)$, die nach Definition von x^0 gilt, ist der Induktionsanfang. Mit der Bezeichnung $\delta_{ij} := 0$ für $i \neq j$ und $\delta_{ii} := 1$ können wir $x = (\delta_{01}, \delta_{11}, \delta_{21}, \ldots)$ und - nach Induktionsvoraussetzung - $x^n = (\delta_{0n}, \delta_{1n}, \ldots, \delta_{nn}, \ldots)$ schreiben. Also gilt für jedes $i \in \mathbb{N}$:

$$(x^{n+1})_i = (x \cdot x^n)_i = \sum_{j+k=i} \delta_{j1} \delta_{kn} = \begin{cases} 1 \text{ falls } i = n+1 \\ 0 \text{ falls } i \neq n+1. \end{cases}$$

(8.7) Folgerung. Jedes Polynom $f \in K[x] \setminus \{0\}$ besitzt genau eine Darstellung

$$f = f_0 + f_1 x + f_2 x^2 + \ldots + f_m x^m \text{ mit } f_m \neq 0.$$

Es gilt dann $\operatorname{grad} f = m$.

Bem.: Bezüglich dieser Darstellung gehen die in (8.1) definierten Operationen in die für Polynome „üblichen“ über.

(8.8) Satz (Division mit Rest). *Seien $f \in K[x]$, $g \in K[x]$, und es gelte* $\operatorname{grad} f \geq \operatorname{grad} g \geq 0$. *Dann existieren $h, r \in K[x]$, so daß*

$$f = gh + r$$

und

$$\operatorname{grad} r < \operatorname{grad} g \qquad \textit{(möglicherweise } r = 0\textit{)}$$

gelten.

Bew.: Durch Induktion nach $n := \operatorname{grad} f$.

Induktionsanfang: Gilt $\operatorname{grad} f = 0$, so auch $\operatorname{grad} g = 0$, also $f = f_0, g = g_0$, und wir können als $h := \frac{g_0}{f_0}$ und $r = 0$ nehmen. Für den Induktionsschritt betrachten wir $f = \sum\limits_{i=0}^{n} f_i x^i$, $g = \sum\limits_{j=0}^{m} g_j x^j$ mit $0 \leq m \leq n$, $f_n \neq 0, g_m \neq 0$. Wir definieren

$$r_1 := f - \frac{f_n}{g_m} x^{n-m} \cdot g.$$

Dann gilt:

$$(*) \qquad f = g \cdot \left(\frac{f_n}{g_m} x^{n-m} \right) + r_1 \text{ und } \operatorname{grad} r_1 < n.$$

1. Fall: $\operatorname{grad} r_1 < \operatorname{grad} g$. Dann erhalten wir die Behauptung, indem wir $h := \frac{f_n}{g_m} x^{n-m}$ und $r := r_1$ setzen.

2. Fall: $\operatorname{grad} r_1 \geq \operatorname{grad} g$. Wegen $n > \operatorname{grad} r_1 \geq \operatorname{grad} g \geq 0$ ist auf r_1 die Induktionsvoraussetzung anwendbar und wir erhalten $h_1, r \in K[x]$, so daß $r_1 = gh_1 + r$ und $\operatorname{grad} r < \operatorname{grad} g$ gelten.

Dann folgt mit $(*)$:

$$f = g \cdot \left(\frac{f_n}{g_m} x^{n-m} + h_1 \right) + r.$$

Setzen wir $h := \frac{f_n}{g_m} x^{n-m} + h_1$, so erhalten wir wegen $\operatorname{grad} r < \operatorname{grad} g$ die Behauptung.

Bem.: Der Beweis von (8.8) besteht in dem üblichen Rechenverfahren zur Polynomdivision, das in die Form eines Beweises gebracht wurde. Umgekehrt liefert der Beweis auch dieses Rechenverfahren, das man beherrschen muß.

Einem Polynom $p = \sum\limits_{i=0}^{n} a_i x^i \in K[x]$ ordnen wir die Abbildung

$$\tilde{p} : K \to K, \ \tilde{p}(b) := \sum_{i=0}^{n} a_i b^i \text{ zu.}$$

Dann gilt für alle $p, q \in K[x]$:

$$\begin{aligned} (p+q)^{\sim} &= \tilde{p} + \tilde{q} \\ (p \cdot q)^{\sim} &= \tilde{p} \cdot \tilde{q}. \end{aligned}$$

Dabei bedeuten $+$ und $\cdot$ auf der linken Seite die Addition und Multiplikation von Polynomen, auf der rechten Seite die Addition und Multiplikation von Abbildungen von K nach

K! Anders ausgedrückt: Die Abbildung $p \to \tilde{p}$ ist ein Ringhomomorphismus von $K[x]$ in den Ring der Selbstabbildungen von K.

Bez.: Ist $p \in K[x]$, $b \in K$, so schreibt man für $\tilde{p}(b)$ meist einfach $p(b)$.

<u>Wichtige Bemerkung</u>: Allgemeiner kann man in $p = \sum\limits_{i=0}^{n} a_i x^i \in K[x]$ auch quadratische Matrizen $A \in K^{m \times m}$ für beliebiges $m \in \mathbb{N}$ einsetzen:

$$p(A) := \sum_{i=0}^{n} a_i A^i \in K^{m \times m}.$$

Es gilt wieder:
$$\begin{aligned}(p+q)(A) &= p(A) + g(A)\\ (p \cdot q)(A) &= p(A) \cdot g(A).\end{aligned}$$

Beweis der zweiten Gleichung: Ist $p = \sum\limits_{i=0}^{m} a_i x^i$, $q = \sum\limits_{j=0}^{n} b_j x^j$, so gilt nach (8.1)(iii): $p \cdot q = \sum\limits_{k=0}^{m+n} \left(\sum\limits_{i+j=k} a_i b_j \right) x^k$, also $(p \cdot q)(A) = \sum\limits_{k=0}^{m+n} \left(\sum\limits_{i+j=k} a_i b_j \right) A^k$. Andererseits berechnen wir in

$$\begin{aligned} K^{m \times m} : p(A) \cdot q(A) &= \left(\sum_{i=0}^{m} a_i A^i \right) \left(\sum_{j=0}^{n} b_j A^j \right) = \sum_{i=0}^{m} \sum_{j=0}^{n} (a_i A^i)(b_j A^j) \\ &= \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i b_j \right) A^k. \end{aligned}$$

(8.9) Def.: Ein $a \in K$ heißt Nullstelle eines Polynoms $p \in K[x]$, falls $p(a) = 0$ gilt.

(8.10) Lemma. $a \in K$ ist genau dann Nullstelle von $p \in K[x] \setminus \{0\}$, wenn ein $q \in K[x]$ existiert, so daß $p = (x-a)q$ gilt.

Bew.: Offenbar folgt aus $p = (x-a)q$, daß

$$\tilde{p}(a) \overset{!}{=} (a-a)\tilde{q}(a) = 0 \text{ gilt.}$$

Gilt $p(a) = 0$, so wenden wir (8.8) auf $f := p$ und $g := x - a$ an und erhalten $q \in K[x]$, $r \in K[x]$ mit $\operatorname{grad} r < \operatorname{grad} g = 1$, so daß

$$p = (x-a)q + r$$

gilt. Aus $0 = p(a) = (a-a)q(a) + r(a)$ folgt $r(a) = 0$ und daraus wegen $\operatorname{grad} r < 1$: $r = 0$.

Bez.: $(x-a)^0 := 1$.

(8.11) Def.: Sei $p \in K[x] \setminus \{0\}$, $a \in K$. Wir bezeichnen mit $k(p,a) \in \mathbb{N}$ das Maximum aller $k \in \mathbb{N}$, für die ein $q \in K[x]$ existiert, so daß $p = (x-a)^k q$ gilt. Ist $k(p,a) > 0$, so gilt $p(a) = 0$ und $k(p,a)$ heißt die Vielfachheit (oder Ordnung) der Nullstelle a von p.

Bem.: 1) $k(p,a) = 0 \Leftrightarrow p(a) \neq 0$.
2) $\exists q \in K[x]$: $p = (x-a)^{k(p,a)}q$ und $q(a) \neq 0$.
3) $k(p,a) \leq \operatorname{grad} p$, denn aus 2) folgt: $\operatorname{grad} p = k(p,a) + \operatorname{grad} q \geq k(p,a)$.

(8.12) Satz. *Seien $p, f, g \in K[x]$, $a \in K$ und es gelte $p = f \cdot g$. Dann gilt*

$$k(p,a) = k(f,a) + k(g,a).$$

Bew.:

(i) Ist $f = (x-a)^{k_1}q_1$, $g = (x-a)^{k_2}q_2$, so gilt

$$p = f \cdot g = (x-a)^{k_1+k_2}q_1q_2.$$

Daraus folgt $k(p,a) \geq k(f,a) + k(g,a)$.

(ii) Ist $p = (x-a)^k q$ mit $k = k(p,a)$ und $f = (x-a)^{k_1}q_1$ mit $q_1(a) \neq 0$ ($\Rightarrow k_1 = k(f,a)$) und $g = (x-a)^{k_2}q_2$ mit $q_2(a) \neq 0$ ($\Rightarrow k_2 = k(g,a)$), so folgt

$$p = (x-a)^k q = f \cdot g = (x-a)^{k_1+k_2}q_1q_2.$$

Wegen (i) gilt $k = k(p,a) \geq k_1 + k_2$, also

$$(x-a)^{k_1+k_2}\left((x-a)^{k-(k_1+k_2)}q - q_1q_2\right) = 0.$$

Da $K[x]$ nulleiterfrei ist, vgl. (8.5), und $(x-a)^{k_1+k_2} \in K[x] \setminus \{0\}$ gilt, folgt $(x-a)^{k-(k_1+k_2)}q - q_1q_2 = 0$. Wenn $k > k_1 + k_2$ gelten würde, so folgte $q_1(a)q_2(a) = 0$, im Widerspruch zu $q_1(a) \neq 0, q_2(a) \neq 0$. Also gilt $k = k(p,a) \leq k_1 + k_2 = k(f,a) + k(g,a)$.

(8.13) Folgerung. Sind $a_1, \ldots, a_s$ verschiedene Nullstellen von $p \in K[x] \setminus \{0\}$ mit den Vielfachheiten $k_1 = k(p,a_1), \ldots, k_s = k(p,a_s)$, so existiert ein $q \in K[x]$ mit:

$$p = (x-a_1)^{k_1} \cdot \ldots \cdot (x-a_s)^{k_s} q.$$

Speziell gilt $\sum\limits_{i=1}^{s} k_i \leq \operatorname{grad} p$.

Bew.: Induktion nach s. Der Induktionsanfang $s = 1$ ist gerade die Definition (8.11). Nach Induktionsvoraussetzung existiert $\overline{q} \in K[x]$ mit

$$p = (x-a_1)^{k_1} \cdot \ldots \cdot (x-a_{s-1})^{k_{s-1}}\overline{q}.$$

Aus (8.12) folgt $k_s = k(p,a_s) = k(\overline{q},a_s)$. Also existiert $q \in K[x]$ mit $\overline{q} = (x-a)^{k_s}q$. Daraus folgt die Behauptung

(8.14) Def.: Ein $p \in K[x] \setminus \{0\}$ zerfällt über K (in Linearfaktoren), falls ein $a \in K$ und $a_1, \ldots, a_s \in K$ und $k_1, \ldots, k_s \in \mathbb{N}$ existieren, so daß gilt

$$p = a(x-a_1)^{k_1} \cdot \ldots \cdot (x-a_s)^{k_s} =: a\prod_{i=1}^{s}(x-a_i)^{k_i}.$$

(8.15) Folgerung (aus (8.13)): $p \in K[x] \setminus \{0\}$ zerfällt genau dann über K, wenn verschiedene $a_1, \ldots, a_s \in K$ existieren, so daß

$$\operatorname{grad} p = \sum_{i=1}^{s} k(p, a_i)$$

gilt.

(8.16) Folgerung. Sind $p, f, g \in K[x] \setminus \{0\}$, gilt $p = f \cdot g$ und zerfällt p über K, so zerfallen auch f und g über K.

Bew.: Nach (8.15) existieren verschiedene $a_1, \ldots, a_s \in K$ mit $\operatorname{grad} p = \sum\limits_{i=1}^{s} k(p, a_i)$.

Nach (8.12) gilt $k(p, a_i) = k(f, a_i) + k(g, a_i)$ für $1 \leq i \leq s$, also

$$\operatorname{grad} f + \operatorname{grad} g = \operatorname{grad} p = \sum_{i=1}^{s} k(p, a_i) = \underbrace{\sum_{i=1}^{s} k(f, a_i)}_{\leq \operatorname{grad} f} + \underbrace{\sum_{i=1}^{s} k(g, a_i)}_{\leq \operatorname{grad} g}.$$

Daraus folgt $\sum\limits_{i=1}^{s} k(f, a_i) = \operatorname{grad} f$, $\sum\limits_{i=1}^{s} k(g, a_i) = \operatorname{grad} g$, so daß f und g nach (8.15) über K zerfallen.

Zum Abschluß wollen wir uns überlegen, wie nun eigentlich das charakteristische Polynom eines Endomorphismus oder einer quadratischen Matrix definiert wird. Das ist nicht unproblematisch, da wir in der Formel $P_L(x) = \det(L - x\,\mathrm{id})$ das x nicht einfach als Körperelement interpretieren können. Der zu diesem Zeitpunkt beste Weg ist zu erkennen, daß man natürlich auch die Determinante einer Matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ definieren kann, deren Einträge a_{ij} Elemente eines festen kommutativen Rings R mit 1 sind, nämlich durch die Leibnizformel $\det A = \sum\limits_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \in R$. Für zwei solche Matrizen A und B ist das Matrizenprodukt AB definiert und es gilt $\det(AB) = \det A \det B$. Das kann man mit einer recht unangenehmen Rechnung direkt nachprüfen. Besser wäre es, wie in Kap. 4 und 5 vorzugehen, aber dabei den Körper K durch einen kommutativen Ring R mit 1 zu ersetzen. Den K-Vektorräumen entsprächen dann „freie R-Moduln“ und den Vektorraumhomomorphismen entsprächen „R-Modulhomomorphismen“, die (im Fall endlich erzeugter R-Moduln) gerade durch Matrizen mit Einträgen aus R beschrieben werden (für Interessierte sei auf Kapitel 9 des Buchs Kowalsky/Michler: Lineare Algebra, de Gruyter 1995, verwiesen). Wir sehen hier ein schönes Beispiel dafür, wie man manchmal von der Mathematik zu allgemeineren Begriffen gezwungen wird.

Jedenfalls kann man dann für eine Matrix $A = (a_{ij})_{1 \leq i,j \leq n} \in K^{n \times n}$ die Matrix $A - xE_n$ als Matrix mit Einträgen im Ring $R = K[x]$ betrachten, und so ist das charakteristische

Polynom

$$P_A = \det(A - xE_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(a_{1\sigma(1)} - x\delta_{1\sigma(1)}) \cdot \ldots \cdot (a_{n\sigma(n)} - x\delta_{n\sigma(n)})$$

von A definiert. Ist $B \in \mathrm{GL}_n(K)$, so gilt

$$\begin{aligned} P_{B^{-1}AB} &= \det(B^{-1}AB - xE_n) = \det(B^{-1}(A - xE_n)B) \\ &= \det B^{-1} P_A \det B = P_A. \end{aligned}$$

Ist nun L Endomorphismus eines n-dimensionalen K-Vektorraums, so wählen wir eine geordnete Basis $\mathcal{G}$ von V, setzen $A := \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$ und definieren das charakteristische Polynom $P_L \in K[x]$ von L durch $P_L := P_A$. Aufgrund der vorangehenden Gleichung ist das unabhängig von der Wahl von $\mathcal{G}$, vgl. (4.19).

9. Die Jordansche Normalform.

Etwas verkürzt ausgedrückt geht es in diesem Kapitel darum, zu einem $L \in \mathrm{End}(V)$, $\dim V = n < \infty$, eine Basis $\mathcal{G}$ zu finden, in der $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$ eine möglichst einfache Form hat.

(9.1) Def.:

(i) $L \in \mathrm{End}(V)$ heißt diagonalisierbar, falls es eine Basis $\mathcal{G} = (v_1, \ldots, v_n)$ von V aus Eigenvektoren v_i von L gibt.
(ii) $A \in K^{n\times n}$ heißt diagonalisierbar, falls es ein $B \in \mathrm{GL}_n(K)$ gibt, so daß $B^{-1}AB$ eine Diagonalmatrix ist (d.h. so daß $(B^{-1}AB)_{ij} = 0$ für $i \neq j$).

Bem.: 1) Ist $\mathcal{G} = (v_1, \ldots, v_n)$ Basis aus Eigenvektoren von L, so gilt

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

wobei λ_i der EW von L zum EV v_i ist.

2) Ist $L \in \mathrm{End}(V), \tilde{\mathcal{G}}$ beliebige Basis von V und $A := \mathrm{Mat}_{\tilde{\mathcal{G}}}^{\tilde{\mathcal{G}}}(L)$, so gilt:

$$L \text{ diagonalisierbar} \Leftrightarrow A \text{ diagonalisierbar}$$

(vgl. (4.19)).

Bsp.: 1) Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, d.h. $A = A^T$, so ist A diagonalisierbar, vgl. den Beweis von (6.11) (Hauptachsentransformation).

2) Ist $V, \langle , \rangle$ euklidischer Vektorraum und ist $L \in \mathrm{End}(V)$ selbstadjungiert ($\Leftrightarrow \forall v, w \in V : \langle L(v), w\rangle = \langle v, L(w)\rangle$), so ist L diagonalisierbar, vgl. (6.11).

Diagonalisierbarkeit einer Matrix $A \in K^{n\times n}$ oder eines $L \in \mathrm{End}(V)$ ist die beste Lösung unseres Problems, die wir uns erhoffen können. Wir werden uns zunächst mit diesem einfachen (aber häufigen) Fall beschäftigen.

(9.2) Fakt. Ist $L \in \mathrm{End}(V)$ und sind $v_1, \ldots, v_k$ EVen zu <u>verschiedenen</u> EWen $\lambda_1, \ldots, \lambda_k$ von L, so sind die $v_1, \ldots, v_k$ linear unabhängig.

Bew.: Durch Induktion nach k.

Induktionsanfang: $k = 1$. Da v_1 EV ist, gilt $v_1 \neq 0$, d.h. v_1 ist linear unabhängig.

Induktionsschritt: Wir können voraussetzen, daß $v_1, \ldots, v_{k-1}$ linear unabhängig sind. Sei $\sum\limits_{i=1}^{k} a_i v_i = 0$. Dann gilt

$$0 = L\left(\sum_{i=1}^{k} a_i v_i\right) = \sum_{i=1}^{k} a_i L(v_i) = \sum_{i=1}^{k} a_i \lambda_i v_i$$

und

$$0 = \lambda_k \left(\sum_{i=1}^{k} a_i v_i\right) = \sum_{i=1}^{k} a_i \lambda_k v_i.$$

Subtraktion der vorangehenden Gleichungen ergibt

$$\sum_{i=1}^{k-1} a_i(\lambda_i - \lambda_k) v_i = 0.$$

Da $v_1, \ldots, v_{k-1}$ linear unabhängig sind, folgt

$$a_i(\lambda_i - \lambda_k) = 0$$

für $1 \leq i \leq k-1$ und wegen $\lambda_i \neq \lambda_k$ auch $a_i = 0$ für $1 \leq i \leq k-1$. Damit reduziert sich $\sum\limits_{i=1}^{k} a_i v_i = 0$ auf $a_k v_k = 0$. Wegen $v_k \neq 0$ folgt auch $a_k = 0$.

(9.3) Folgerung. Seien $\dim V = n$, $L \in \operatorname{End}(V)$. Besitzt L n verschiedene EWe, so ist L diagonalisierbar.

Bem.: L besitzt genau dann n verschiedene EWe, wenn P_L n verschiedene Nullstellen besitzt.

Bew.: Seien $v_1, \ldots, v_n$ EVen zu den verschiedenen EWen von L. Dann zeigt (9.2), daß $\mathcal{G} = (v_1, \ldots, v_n)$ eine Basis aus EVen ist.

(9.4) Def.: Ist λ EW von $L \in \operatorname{End}(V)$, so heißt

$$E(\lambda) = \ker(L - \lambda \operatorname{id}_V) = \{v \in V \mid L(v) = \lambda v\}$$

der Eigenraum von L zum EW λ.

Bem.: $E(\lambda)$ ist Untervektorraum von V und $E(\lambda) \setminus \{0\}$ ist genau die Menge der EVen von L zum EW λ.

(9.5) Satz (Diagonalisierbarkeitskriterium). *Sei $L \in \operatorname{End}(V)$. L ist genau dann diagonalisierbar, wenn*

$$\dim V = \sum_{i=1}^{k} \dim E(\lambda_i)$$

gilt, wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen EWe von L bezeichnen.

Bew.: Setze $d_i := \dim E(\lambda_i)$. Aus (9.2) folgt, daß $E(\lambda_1) \oplus \ldots \oplus E(\lambda_k)$ eine direkte Summe ist, also $\sum\limits_{i=1}^{k} d_i \leq n$.

„$\Rightarrow$" Ist $\mathcal{G} = (v_1, \ldots, v_n)$ Basis aus EVen, so gilt offenbar auch $\sum d_i \geq n$, also $\sum d_i = n$.

„$\Leftarrow$" Gilt $\sum\limits_{i=1}^{k} d_i = n$, so $\dim(E(\lambda_1) \oplus \ldots \oplus E(\lambda_k)) = n$, also $E(\lambda_1) \oplus \ldots \oplus E(\lambda_k) = V$. Also ist die (disjunkte!) Vereinigung von Basen der Eigenräume $E(\lambda_i), 1 \leq i \leq k$, eine Basis von V.

Bem.: Wenn man (9.3) oder (9.5) auf einen konkreten Endomorphismus L anwenden will, muß man seine EWe, d.h. die Nullstellen von P_L bestimmen. Das wird oft nicht explizit, sondern nur approximativ (mit Hilfe der Numerik) möglich sein. Kennt man die EWe, so ist die Bestimmung der Eigenräume leicht. Sie sind die Lösungsmengen von homogenen linearen Gleichungssystemen $\{v \in V \mid L(v) - \lambda v = 0\}$, λ EW von L.

Es ist wichtig zu wissen, daß etwa im Fall des Körpers $K = \mathbb{C}$, „die meisten" $A \in \mathbb{C}^{n\times n}$ (bzw. die meisten Endomorphismen) diagonalisierbar sind. Auf die Frage, was „die meisten" mathematisch bedeuten soll, gibt es (mindestens) zwei (verschiedene) Antworten.

Gegeben eine Eigenschaft E, die Punkte $x \in \mathbb{R}^n$ haben können oder nicht.

1) Definition I: Die meisten $x \in \mathbb{R}^n$ haben die Eigenschaft E, falls die Menge $\{x \in \mathbb{R}^n \mid x \text{ hat Eigenschaft } E\}$ eine offene und dichte Teilmenge des $\mathbb{R}^n$ enthält.
2) Definition II: Die meisten $x \in \mathbb{R}^n$ haben die Eigenschaft E, falls die Menge $\{x \in \mathbb{R}^n \mid x \text{ hat nicht Eigenschaft } E\}$ (Lebesgue-)Maß 0 hat.

Dabei heißt eine Teilmenge A des $\mathbb{R}^n$ vom Maß 0, falls es für jedes $\epsilon > 0$ abzählbar viele Würfel $W_i \subseteq \mathbb{R}^n$ gibt, so daß $A \subseteq \bigcup\limits_{i\in\mathbb{N}} W_i$ und $\sum\limits_{i\in\mathbb{N}} \mathrm{vol}_n(W_i) \leq \epsilon$ gilt. Definition I benutzt topologische, Definition II maßtheoretische Begriffe.

Bsp.: 1) Ist E die Eigenschaft eines $x \in \mathbb{R}$, irrational zu sein, so haben im Sinn von Def. I <u>nicht</u> die meisten $x \in \mathbb{R}$ Eigenschaft E, d.h. es gilt nicht, daß die meisten $x \in \mathbb{R}$ irrational sind: Die Menge $\mathbb{R} \setminus \mathbb{Q}$ enthält überhaupt keine nichtleere offene Teilmenge von $\mathbb{R}$. Im Sinn von Definition II sind dagegen die meisten $x \in \mathbb{R}$ irrational: Ist $\epsilon > 0$ gegeben, so wählen wir eine Folge $(r_i)_{i\in\mathbb{N}}$, die alle rationalen Zahlen durchläuft, und setzen $W_i = [r_i - \frac{\epsilon}{2^{i+2}}, r_i + \frac{\epsilon}{2^{i+2}}]$. Dann gilt $\mathbb{Q} \subseteq \bigcup\limits_{i\in\mathbb{N}} W_i$ und $\sum\limits_{i\in\mathbb{N}} \mathrm{vol}_1(W_i) = \sum\limits_{i\in\mathbb{N}} \frac{\epsilon}{2^{i+1}} = \epsilon$.

2) Es ist nicht schwer, mit einer Konstruktion ähnlich der Konstruktion der Cantormenge eine Teilmenge A von $\mathbb{R}$ zu finden, so daß $\mathbb{R} \setminus A$ offen und dicht ist, aber A nicht Maß 0 hat. Dann sind im Sinn von Def. I die meisten $x \in \mathbb{R}$ nicht Elemente von A, während das im Sinn von Def. II nicht gilt.

Identifiziert man $\mathbb{C}^{n\times n}$ mit $\mathbb{R}^{(2n)\times(2n)} \simeq \mathbb{R}^{4n^2}$, so lassen sich diese Begriffe auch auf Matrizen $A \in \mathbb{C}^{n\times n}$ anwenden. Folgender Satz gilt für beide Definitionen von „die meisten“.

(9.6) Satz (ohne Beweis)

(i) *Die meisten* $p \in \mathbb{C}[x]$ *mit* $\operatorname{grad} p \leq n$ *haben* n *verschiedene Nullstellen.*
(ii) *Die meisten* $A \in \mathbb{C}^{n\times n}$ *sind diagonalisierbar.*

In (i) wird $V := \{p \in \mathbb{C}[x] \mid \operatorname{grad} p \leq n\}$ mit $\mathbb{C}^{n+1} \simeq \mathbb{R}^{2(n+1)}$ identifiziert.

Beweisidee für (i): Wir wollen zeigen, daß die Menge $\{p \in V \mid p$ hat nicht n verschiedene Nullstellen$\}$ Maß 0 hat. Wir betrachten die Abbildung

$$F : \mathbb{C}^{n+1} \to V,\ F(a, \lambda_1, \ldots, \lambda_n) = a \prod_{i=1}^{n} (x - \lambda_i).$$

Der „Fundamentalsatz der Algebra“ impliziert (mit (8.10) und (8.13)), daß jedes $p \in \mathbb{C}[x] \setminus \{0\}$ zerfällt. Das zeigt, daß

$$F(\mathbb{C}^{n+1}) = \{p \in V \mid \operatorname{grad} p = n\}$$

gilt, d.h. $F(\mathbb{C}^{n+1}) = V \setminus H$, wobei $H = \{p \in V \mid \operatorname{grad} p < n\}$ eine Hyperebene ist und Maß 0 hat. Die Menge

$$B := \{(a, \lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n+1} \mid \exists 1 \leq i < j \leq n : \lambda_i = \lambda_j\}$$

ist eine Vereinigung von $\frac{n(n-1)}{2}$ Hyperebenen in $\mathbb{C}^{n+1}$ und hat ebenfalls Maß 0 (in $\mathbb{C}^{n+1} \simeq \mathbb{R}^{2(n+1)}$). Wir benötigen nun den leicht zu beweisenden Satz, daß C^1-Abbildungen $\mathbb{R}^n \to \mathbb{R}^n$ Mengen vom Maß 0 auf Mengen vom Maß 0 abbilden. Wegen $\dim_{\mathbb{R}} \mathbb{C}^{n+1} = 2(n+1) = \dim_{\mathbb{R}} V$ und da F C^1 (sogar polynomial) ist, folgt, daß $F(B) \subseteq V$ Maß 0 hat. Die Mengen aller $p \in V$, die nicht n verschiedene Nullstellen haben, ist aber gerade $H \cup F(B)$, und da sowohl H als auch $F(B)$ Maß 0 haben, folgt das auch für $H \cup F(B)$. Behauptung (ii) läßt sich unter Verwendung von (i) mit ähnlichen Methoden zeigen.

Ziel dieses Exkurses war, zu demonstrieren, wie Ideen und Techniken, die in der Analysis I und II vermittelt werden, hier in der Linearen Algebra nützlich und nötig sind.

Nach (9.6) sind im Fall $K = \mathbb{C}$ Endomorphismen L bzw. quadratische Matrizen A in der Regel diagonalisierbar.

Im Rest dieses Kapitels befassen wir uns mit den nicht diagonalisierbaren Ausnahmefällen. Zwar können wir für solche L keine Basis aus EVen finden, aber doch eine Basis, in der die Matrix von L eine besonders einfache Gestalt – die Jordansche Normalform – annimmt. Die Überlegungen dazu sind weder kurz noch einfach, und Sie können die hier gewählte Darstellung auch im Buch R. Walter: Einführung in die lineare Algebra, Vieweg 1982, nachlesen.

Im folgenden sei V ein K-Vektorraum, $\dim V = n$, und $L \in \operatorname{End}(V)$.

(9.7) Def.:

(i) Ein Untervektorraum U von V heißt L-invariant, falls $L(U) \subseteq U$ gilt.
(ii) Ein L-invarianter Untervektorraum heißt L-reduzibel, falls L-invariante Untervektorräume U_1, U_2 existieren, so daß $U = U_1 \oplus U_2$ gilt.
(iii) Ein L-invarianter Untervektorraum U heißt L-irreduzibel, falls U nicht L-reduzibel ist.

Unser Ziel ist, eine Zerlegung von V in L-irreduzible Unterräume U_i zu finden, für die jedes $L|U_i$ von einer einfachen Bauart ist.

(9.8) Def.: Sei λ EW von L. Dann heißt

$$E'(\lambda) = \{v \in V \mid \exists k \in \mathbb{N} : (L - \lambda\, \mathrm{id}_V)^k(v) = 0\}$$

der Hauptraum (oder verallgemeinerte Eigenraum) zum EW λ.

Bem.: Es gilt $E(\lambda) = \ker(L - \lambda\, \mathrm{id}_V) \subseteq E'(\lambda) = \bigcup\limits_{k\in\mathbb{N}} \ker(L - \lambda\, \mathrm{id}_V)^k$.

Fakt: $E'(\lambda)$ ist L-invarianter Untervektorraum.

Wir zeigen: $v \in E'(\lambda) \Rightarrow L(v) \in E'(\lambda)$. Ist $k \in \mathbb{N}$ so, daß $(L - \lambda\, \mathrm{id}_V)^k(v) = 0$ gilt, so folgt

$$(L - \lambda\, \mathrm{id}_V)^k(L(v)) = (L \circ (L - \lambda\, \mathrm{id}_V)^k)(v) = 0,$$

d.h. $L(v) \in E'(\lambda)$.

(9.9) Bem.: λ ist der einzige EW von $L|E'(\lambda)$.

Bew.:

(i) λ ist *EW* von $L|E'(\lambda)$: Da λ EW von L ist, existiert $0 \neq v \in V$ mit $L(v) = \lambda v$, also $(L - \lambda\, \mathrm{id}_V)(v) = 0$. Daraus folgt $v \in E'(\lambda)$, und das zeigt, daß λ EW von $L|E'(\lambda)$ ist.
(ii) Annahme: $\mu \neq \lambda$ ist ein weiterer EW von $L|E'(\lambda)$. Dann existiert $0 \neq w \in E'(\lambda)$ mit $L(w) = \mu w$. Daraus folgt $(L - \lambda\, \mathrm{id}_V)(w) = (\mu - \lambda)v$ und daraus

$$(L - \lambda\, \mathrm{id}_V)^k(w) = (\mu - \lambda)^k w \neq 0$$

für alle $k \in \mathbb{N}$, im Widerspruch zu $w \in E'(\lambda)$.

(9.10) Lemma. Es existiert eine kleinste Zahl $f(\lambda) \in \mathbb{N}$, genannt der Index von λ, so daß $\ker(L - \lambda\, \mathrm{id}_V)^{f(\lambda)} = E'(\lambda)$ gilt. Es gilt $V = E'(\lambda) \oplus \mathrm{im}(L - \lambda\, \mathrm{id}_V)^{f(\lambda)}$.

Bem.: Wegen $L \circ (L - \lambda\, \mathrm{id}_V)^{f(\lambda)} = (L - \lambda\, \mathrm{id}_V)^{f(\lambda)} \circ L$ ist im $(L - \lambda\, \mathrm{id}_V)^{f(\lambda)}$ L-invarianter Unterraum von V.

Bew.: Sei $v_1, \ldots, v_l$ Basis von $E'(\lambda)$. Dann existiert für jedes $i \in \{1, \ldots, l\}$ ein kleinstes $k_i \in \mathbb{N}$ mit $(L - \lambda\, \mathrm{id}_V)^{k_i}(v_i) = 0$. Sei $f(\lambda) := \max\limits_{1\leq i\leq l} k_i$. Dann gilt $(L - \lambda\, \mathrm{id}_V)^{f(\lambda)}(v_i) = 0$ für

alle $i \in \{1,\ldots,l\}$, also
$$E'(\lambda) \subseteq \ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)}.$$
Nach Definition von $E'(\lambda)$ folgt daraus $E'(\lambda) = \ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)}$, und $f(\lambda)$ ist offenbar die kleinste Zahl mit dieser Eigenschaft.

Schlielich zeigen wir, daß
$$\ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)} \oplus \mathrm{im}(L - \lambda\,\mathrm{id}_V)^{f(\lambda)} = V$$
gilt. Wegen des Dimensionssatzes (4.8) gilt
$$\dim\ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)} + \mathrm{im}(L - \lambda\,\mathrm{id}_V)^{f(\lambda)} = \dim V\,,$$
so daß sich die Behauptung auf
$$\ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)} \cap \mathrm{im}(L - \lambda\,\mathrm{id}_V)^{f(\lambda)} = \{0\}$$
reduziert. Sei also $v \in \ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)} \cap \mathrm{im}(L - \lambda\,\mathrm{id}_V)^{f(\lambda)}$. Dann existiert ein $w \in V$, so dass
$$v = (L - \lambda\,\mathrm{id}_V)^{f(\lambda)}(w)$$
gilt. Wegen $v \in \ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)}$, folgt $\ker(L - \lambda\,\mathrm{id}_V)^{2f(\lambda)}(w) = 0$, also $w \in E'(\lambda)$. Aus $E'(\lambda) = \ker(L - \lambda\,\mathrm{id}_V)^{f(\lambda)}$ folgt nun
$$v = (L - \lambda\,\mathrm{id}_V)^{f(\lambda)}(w) = 0\,.$$

(9.11) Lemma. Seien $\lambda_1,\ldots,\lambda_s$ verschiedene EWe von L. Dann ist die Summe der $E'(\lambda_i)$ für $1 \le i \le s$ direkt: $E'(\lambda_1) \oplus \ldots \oplus E'(\lambda_s)$.

Bew.: Wir zeigen durch Induktion nach s, daß für alle $i \in \{1,\ldots,l\}$ gilt:
$$E'(\lambda_i) \cap \mathrm{span}\left(\bigcup_{\substack{j=1\\ j\neq i}}^{s} E'(\lambda_j)\right) = \{0\}.$$
Für $s = 1$ ist das offensichtlich richtig. Im Induktionsschritt können wir nach Umnumerierung der λ_j annehmen, daß $i = s$ gilt. Sei $v \in E'(\lambda_s)$, $v_j \in E'(\lambda_j)$ für $1 \le j < s$, und es gelte
$$v = \sum_{j=1}^{s-1} v_j.$$
Wir müssen beweisen, daß $v = 0$ gilt. Nach Definition von $f(\lambda_s)$ gilt:
$$(*) \qquad 0 = (L - \lambda_s\,\mathrm{id}_V)^{f(\lambda_s)}(v) = \sum_{j=1}^{s-1} (L - \lambda_s\,\mathrm{id}_V)^{f(\lambda_s)}(v_j).$$

Da $E'(\lambda_j)$ L-invariant – und damit auch $(L - \lambda_s\,\mathrm{id}_V)$-invariant – ist, gilt $(L - \lambda_s\,\mathrm{id}_V)^{f(\lambda_s)}(v_j) \in E'(\lambda_j)$ für $1 \le j < s$. Aus der Induktionsvoraussetzung und $(*)$ folgt nun $(L - \lambda_s\,\mathrm{id}_V)^{f(\lambda_s)}(v_j) = 0$ für $1 \le j < s$ und daraus $v_j \in E'(\lambda_s) \cap E'(\lambda_j)$. Wir zeigen, daß daraus $v_j = 0$ folgt: Sonst sei $k \in \mathbb{N}$ minimal mit $(L - \lambda_s \,\mathrm{id}_V)^k(v_j) = 0$. Wäre $v_j \neq 0$, so $k \geq 1$, und wir können $w_j := (L - \lambda_s \,\mathrm{id}_V)^{k-1}(v_j)$ betrachten. Dann gilt $0 \neq w_j \in E'(\lambda_j)$ und $(L - \lambda_s \,\mathrm{id}_V)(w_j) = 0$, im Widerspruch zu (9.9). Das zeigt, daß $v_j = 0$ für $1 \leq j < s$, und daraus folgt $v = 0$, wie behauptet.

(9.12) Satz (1. Zerlegungssatz).*Das charakteristische Polynom $P_L \in K[x]$ von L zerfalle. Es seien $\lambda_1, \ldots, \lambda_s$ die verschiedenen EWe von L. Dann gilt*

$$V = E'(\lambda_1) \oplus \ldots \oplus E'(\lambda_s).$$

Bew.: Nach (9.11) wissen wir schon, daß die Summe direkt ist. Wir beweisen durch Induktion nach $n = \dim V$, daß sie ganz V ist. Ist $n = 1$, so gilt $s = 1$ und $V = E(\lambda_1) = E'(\lambda_1)$.

Im Induktionsschritt verwenden wir, daß nach (9.10) ein L-invarianter Unterraum U (nämlich $\mathrm{im}(L - \lambda_1 \,\mathrm{id}_V)^{f(\lambda_1)}$) existiert, für den

$$(*) \qquad V = E'(\lambda_1) \oplus U$$

gilt. Wegen $\dim E'(\lambda_1) \geq 1$, gilt $\dim U = \dim V - \dim E'(\lambda_1) < n$. Wir wollen die Induktionsvoraussetzung auf $\overline{L} := L|U \in \mathrm{End}(U)$ anwenden und müssen dazu wissen, daß auch $P_{\overline{L}}$ zerfällt. Aus $(*)$ folgt, daß

$$P_L = P_{L|E'(\lambda_1)} \cdot P_{\overline{L}}$$

gilt, vgl. Blatt 3, Aufgabe 3. Nach (8.16) zerfällt auch $P_{\overline{L}}$, und (9.9) impliziert, daß $\lambda_2, \ldots, \lambda_s$ die verschiedenen Nullstellen von $P_{\overline{L}}$ sind. Die Induktionsvoraussetzung ergibt nun, daß

$$U = E'_{\overline{L}}(\lambda_2) + \ldots + E'_{\overline{L}}(\lambda_s)$$

gilt, wobei $E'_{\overline{L}}(\lambda_j) \subseteq E'(\lambda_j)$ den Hauptraum von $\overline{L}$ zum EW λ_j von $\overline{L}$ bezeichnet. Daraus folgt mit $(*)$

$$V = E'(\lambda_1) + E'(\lambda_2) + \ldots + E'(\lambda_s),$$

wie behauptet.

Unser nächstes Ziel ist die Untersuchung von $L_i := L|E'(\lambda_i) \in \mathrm{End}(E'(\lambda_i))$ für $1 \leq i \leq s$. Für $T_i := L_i - \lambda_i \,\mathrm{id}_{E'(\lambda_i)}$ und $f_i := f(\lambda_i)$ gilt (nach (9.10)):

$$(T_i)^{f_i} = 0.$$

(9.13) Def.: Ein $T \in \mathrm{End}(V)$ heißt <u>nilpotent</u>, falls es ein $k \in \mathbb{N}_{>0}$ gibt, so daß $T^k = 0 \in \mathrm{End}(V)$ gilt. Das kleinste solche k heißt dann der Nilpotenzgrad $g = g(T)$ von T.

(9.14) Lemma. Sei $0 \neq T \in \mathrm{End}(V)$ nilpotent. Ist $v \in V \setminus \ker(T^{g-1})$, so sind die Vektoren $v, T(v), \ldots, T^{g-1}(v)$ linear unabhängig und spannen einen g-dimensionalen T-invarianten Unterraum von V auf. Speziell gilt $g \leq \dim V$.

Bew.: Gilt $\sum\limits_{i=0}^{g-1} a_i T^i(v) = 0$, so auch

$$0 = T^{g-1}\left(\sum_{i=0}^{g-1} a_i T^i(v)\right) = \sum_{i=0}^{g-1} a_i T^{g-1+i}(v).$$

Nach Definition von g gilt $T^{g-1+i}(v) = 0$ für $i \geq 1$. Also reduziert sich die vorangehende Gleichung auf $a_0 T^{g-1}(v) = 0$. Nach Voraussetzung gilt $T^{g-1}(v) \neq 0$, also $a_0 = 0$. Wendet man T^{g-2} auf $\sum\limits_{i=0}^{g-1} a_i T^i(v) = 0$ an, so erhält man $a_0 T^{g-2}(v) + a_1 T^{g-1}(v) = 0$. Wegen $a_0 = 0$ folgt wie vorher: $a_1 = 0$. Induktiv erhalten wir durch Anwenden von $T^{g-2}, \ldots, T^0 = \mathrm{id}_V$ auf $\sum\limits_{i=0}^{g-1} a_i T^i(v) = 0$, daß auch $a_2 = 0, \ldots, a_{g-1} = 0$ gilt. Das beweist, daß $v, T(v), \ldots, T^{g-1}(v)$ linear unabhängig sind. Die T-Invarianz von $\mathrm{span}\{v, \ldots, T^{g-1}(v)\}$ folgt aus der Tatsache, daß die Basis $\{v, \ldots, T^{g-1}(v)\}$ von $\mathrm{span}\{v, \ldots, T^{g-1}(v)\}$ durch T auf die Teilmenge $\{T(v), \ldots, T^{g-1}(v), 0\}$ von $\mathrm{span}\{v, \ldots, T^{g-1}(v)\}$ abgebildet wird.

Nilpotente Endomorphismen T mit Nilpotenzgrad $g(T) = \dim V = n$ sind leicht zu verstehen:

(9.15) Lemma. Sei $T \in \mathrm{End}(V)$ nilpotent, $g(T) = n = \dim V$. Dann ist V T-irreduzibel, und es existiert eine Basis $\mathcal{G} = (v, T(v), \ldots, T^{n-1}(v))$ von V, so daß

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(T) = \begin{pmatrix} 0 & \ldots & \ldots & \ldots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 0 \end{pmatrix}$$

gilt.

Bew.: Zum Beweis der Irreduzibilität von V nehmen wir an, daß $V = U_1 \oplus U_2$ mit zwei T-invarianten Unterräumen $U_1 \neq 0$, $U_2 \neq 0$ gilt. Dann sind $T_i := T|U_i \in \mathrm{End}(U_i)$ für $i = 1, 2$ nilpotent und wegen (9.14) gilt $g(T_i) \leq \dim U_i < n$. Daraus folgt

$$g(T) = \max\{g(T_1), g(T_2)\} < n,$$

im Widerspruch zur Voraussetzung $g(T) = n$. Die restliche Behauptung von (9.15) folgt direkt aus (9.14).

Bem.: Haben wir also für einen - jetzt nicht als nilpotent vorausgesetzten - Endomorphismus $L \in \mathrm{End}(V)$ einen EW λ_i, so daß

$$T_i := (L - \lambda_i \, \mathrm{id}_V) \mid E'(\lambda_i)$$

den Nilpotenzgrad $f_i(=g(T_i)) = \dim E'(\lambda_i)$ hat, so können wir mit (9.15) eine Basis $\mathcal{G}_i$ von $E'(\lambda_i)$ finden, so daß

$$\mathrm{Mat}_{\mathcal{G}_i}^{\mathcal{G}_i}(T|E'(\lambda_i)) = \begin{pmatrix} \lambda_i & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \lambda_i \end{pmatrix}$$

gilt.

Nun betrachten wir den Fall eines nilpotenten $T \in \mathrm{End}(V)$ mit $g(T) < n$.

(9.16) Lemma (2. Zerlegungssatz). Sei $T \in \mathrm{End}(V)$ nilpotent und $g = g(T) < n = \dim V$. Dann ist V T-reduzibel.

Bew.: Wegen $T^{g-1} \neq 0$ können wir ein $v \in V \setminus \ker(T^{g-1})$ wählen und dazu ein $l \in V^*$ mit $l(T^{g-1}(v)) \neq 0$. Wir definieren

$$\begin{aligned} U_1 &:= \mathrm{span}\{v, \dots, T^{g-1}(v)\} \quad \text{und} \\ U_2 &:= \ker(l) \cap \ker(l \circ T) \cap \dots \cap \ker(l \circ T^{g-1}). \end{aligned}$$

Dann ist U_1 T-invariant, und es gilt $0 < \dim U_1 = g < n$. Wir haben (9.16) bewiesen, wenn wir zeigen können, daß U_2 T-invariant ist und daß $V = U_1 \oplus U_2$ gilt.

T-Invarianz von U_2: Ist $w \in U_2$, so gilt $l(w) = 0, \dots, l(T^{g-1}(w)) = 0$, also speziell $l(T(w)) = 0, \dots, l(T^{g-2}(T(w)) = 0$ und außerdem wegen $T^g(w) = 0$: $l(T^g(w)) = l(T^{g-1}(T(w)) = 0$. Zusammen zeigen diese Gleichungen, daß $T(w) \in U_2$ gilt. Also ist U_2 T-invariant.

Zeige $U_1 \cap U_2 = \{0\}$: Sei $w \in U_1 \cap U_2$. Da $w \in U_1$ ist, existieren $a_0, \dots, a_{g-1} \in K$ mit $w = \sum\limits_{i=0}^{g-1} a_i T^i(v)$. Da $w \in U_2$ ist, gilt $l(T^{g-1}(w))) \dots = l(w) = 0$.

Wegen $T^{g-1}(w) = a_0 T^{g-1}(v)$ gilt $0 = l(T^{g-1}(w)) = a_0 l(T^{g-1}(v))$, und wegen $l(T^{g-1}(v)) \neq 0$ folgt daraus $a_0 = 0$. Wie im Beweis von (9.14) folgt induktiv $a_1 = 0, \dots, a_{g-1} = 0$, also $w = 0$. Das beweist $U_1 \cap U_2 = \{0\}$. Es liegt also eine direkte Summe $U_1 \oplus U_2$ vor, und wir haben

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

Es gilt $\dim U_1 = g$ und aus (3.23) folgt $\dim U_2 \geq n - g$, also

$$\dim U_1 \oplus U_2 \geq g + (n-g) = n = \dim V.$$

Da $U_1 \oplus U_2$ Untervektorraum von V ist, folgt daraus $V = U_1 \oplus U_2$, wie behauptet.

Zusammen besagen (9.15) und (9.16) gerade, daß für ein nilpotentes $0 \neq T \in \mathrm{End}(V)$ gilt:

$$V \ T\text{-irreduzibel} \Leftrightarrow g(T) = \dim V.$$

Auf jedem T-irreduziblen Unterraum U von V ist $T|U$ also von der einfachen, in (9.15) beschriebenen Bauart. Daß eine Zerlegung in irreduzible Unterräume möglich ist, folgt direkt aus der Definition von irreduzibel:

(9.17) Lemma. Sei $L \in \mathrm{End}(V)$, $V \neq \{0\}$. Dann existieren $k \in \mathbb{N}_{>0}$ und L-irreduzible Unterräume $U_1 \neq \{0\}, \ldots, U_k \neq \{0\}$, so daß

$$V = U_1 \oplus \ldots \oplus U_k$$

gilt.

Bew.: Durch Induktion nach $\dim V = n$. Der Induktionsschritt besteht in der einfachen Bemerkung, daß V entweder L-irreduzibel ist oder eine Zerlegung $V = U_1 \oplus U_2$ mit L-invarianten Unterräumen $U_1 \neq 0$ und $U_2 \neq 0$ gestattet. Wegen $\dim U_1 < n$, $\dim U_2 < n$ ist auf diese die Induktionsvoraussetzung anwendbar.

Wir beweisen nun zunächst den Satz über die Jordansche Normalform in der Formulierung für Endomorphismen und werden das später in eine Aussage über Matrizen übersetzen.

(9.18) Satz (Jordan-Zerlegung). *Sei $L \in \mathrm{End}(V)$ und das charakteristische Polynom von L zerfalle. Sind $\lambda_1, \ldots, \lambda_s$ die verschiedenen EWe von L, so gilt*

$$L = E'(\lambda_1) \oplus \ldots \oplus E'(\lambda_s).$$

Jeder der Haupträume $E'(\lambda_t), 1 \leq t \leq s$, zerfällt in L-irreduzible Unterräume

$$E'(\lambda_t) = U_t^1 \oplus \ldots \oplus U_t^{k_t}$$

und für jedes $i \in \{1, \ldots, k_t\}$ ist $(L - \lambda_t \,\mathrm{id}_V)|U_t^i =: S_t^i \in \mathrm{End}(U_t^i)$ nilpotent vom Nilpotenzgrad $g(S_t^i) = \dim U_t^i$.

Bew.: Die erste Aussage ist gerade der 1. Zerlegungssatz (9.12). Wir wenden dann (9.17) auf $L|E'(\lambda_t)$ an und erhalten eine Zerlegung von $E'(\lambda_t)$ in L-irreduziblen Unterräumen, genannt $U_t^1, \ldots, U_t^{k_t}$. Nach (9.10) wissen wir, daß $(L - \lambda_t \,\mathrm{id}_V)|U_t^i =: S_t^i$ für jedes $t \in \{1, \ldots, s\}$ und jedes $i \in \{1, \ldots, k_t\}$ nilpotent ist. Da U_t^i L-irreduzibel ist, ist U_t^i auch S_t^i-irreduzibel. Aus dem 2. Zerlegungssatz (9.16) folgt nun, daß der Nilpotenzgrad $g(S_t^i)$ von S_t^i gleich der Dimension von U_t^i ist, d.h. daß S_t^i von dem einfachen, durch (9.15) beschriebenen Typ ist.

Nach (9.15) existiert zu jedem der L-irreduziblen Unterräume U_t^i, $1 \leq t \leq s$, $1 \leq i \leq k_t$ eine Basis $\mathcal{G}_t^i$, so daß

$$\mathrm{Mat}_{\mathcal{G}_t^i}^{\mathcal{G}_g^i}(S_t^i) = \begin{pmatrix} 0 & \ldots & \ldots & \ldots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 0 \end{pmatrix}$$

gilt. Wegen $L|U_t^i = S_t^i + \lambda_t \operatorname{id}_{U_t^i}$ folgt

$$(*) \qquad \operatorname{Mat}_{\mathcal{G}_t^i}^{\mathcal{G}_g^i}(L|U_t^i) = \begin{pmatrix} \lambda_t & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \lambda_t \end{pmatrix} =: J_t^i.$$

Damit erhalten wir

(9.19) Satz (Jordansche Normalform). *Sei $L \in \operatorname{End}(V)$ und das charakteristische Polynom P_L von L zerfalle. Die verschiedenen Nullstellen von P_L seien $\lambda_1, \dots, \lambda_s$. Dann existiert eine Basis $\mathcal{G}$ von V, so daß*

$$(**) \qquad \operatorname{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} J_1^1 & & & & 0 \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & J_s^{k_s} \end{pmatrix}$$

gilt, wobei die in der Diagonale stehenden „Jordanblöcke" J_t^i $(\dim U_t^i) \times (\dim U_t^i)$-Matrizen vom Typ $(*)$ sind.

Interpretiert für Matrizen $A \in K^{n \times n}$ besagt (9.19):

(9.20) Folgerung. Sei $A \in K^{n \times n}$ und das charakteristische Polynom von A zerfalle. Dann existiert ein $B \in \operatorname{GL}_n(K)$, so daß $B^{-1}AB$ die Form $(**)$ hat.

Bem.:

1) Ist $K = \mathbb{C}$, so ist die Voraussetzung „das charakteristische Polynom zerfalle" nach dem Fundamentalsatz der Algebra stets erfüllt. Ist K ein Unterkörper von $\mathbb{C}$, z.B. $K = \mathbb{Q}$ oder $K = \mathbb{R}$, und zerfällt das charakteristische Polynom in $K[x]$ nicht, so können wir uns durch „Komplexifizierung" der Situation helfen, d.h. ist $L \in \operatorname{End}(\mathbb{R}^n)$, so betrachten wir die $\mathbb{C}$-lineare Fortsetzung $\tilde{L} \in \operatorname{End}(\mathbb{C}^n)$ von L, vgl. (6.11) und (6.25). In der Algebra konstruiert man zu jedem Körper K einen K enthaltenden Körper $\tilde{K}$, in dem jedes Polynom zerfällt. Also kann man die Voraussetzung „das charakteristische Polynom zerfalle" auch für beliebige Körper durch ein der Komplexifizierung ähnliches Verfahren erzwingen.
2) Wie bereits gesagt ist das Hauptproblem bei der Berechnung der Jordanschen Normalform – die Berechnung der EWe $\lambda_1, \dots, \lambda_s$ – i.a. nicht explizit lösbar. Kennt man $\lambda_1, \dots, \lambda_s$, so ist die Jordansche Normalform und eine zugehörige Basis durch Lösen von linearen Gleichungssystemen berechenbar. Dazu noch einige Hinweise: Zu berechnen sind für jedes $t \in \{1, \dots, s\}$ die Zahl k_t der L-irreduziblen Unterräume

$U_t^1, \ldots, U_t^{k_t}$, in die $E'(\lambda_t)$ zerfällt, und die Dimensionen der U_t^i. Offenbar gilt

$$\sum_{i=1}^{k_t} \dim U_t^i = \dim E'(\lambda_t) \text{ und } \sum_{t=1}^{s} \dim E'(\lambda_t) = \dim V.$$

Berechnet man P_L mit der Jordanschen Normalform, so sieht man, daß $\dim E'(\lambda_t)$ gerade die Vielfachheit $k(P_L, \lambda_t)$ der Nullstelle λ_t ist. $\dim E'(\lambda_t)$kann nach (9.8), (9.10) durch Lösen homogener linearer Gleichungen bestimmt werden. Die Zerlegung von $\dim E'(\lambda_t)$ in L-irreduzible Unterräume kann mit dem im Beweis von (9.16) (2. Zerlegungssatz) verwendeten Verfahren explizit durchgeführt werden.

3) Während eine Basis $\mathcal{G}$, in der $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$ Jordansche Normalform hat, nicht eindeutig durch L bestimmt ist, ist die Jordansche Normalform $(**)$ in (9.19) bis auf die Reihenfolge der Blöcke J_t^i eindeutig durch L bestimmt. Der Beweis dieser Tatsache wurde in der Vorlesung nicht durchgeführt. Er ist nicht schwierig und beruht darauf, daß man für eine nilpotente lineare Abbildung T die Anzahl der irreduziblen Unterräume einer gegebenen Dimension sukzessive aus den Zahlen $\dim(\ker T^j)$ ausrechnen kann, vgl. dazu z.B. R. Walter, Einführung in die lineare Algebra, S. 245, Satz A.

<u>Beispiel zur Berechnung der Jordanschen Normalform.</u>

Wir berechnen die Jordansche Normalform von

$$A = \begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}$$

$$\begin{aligned} \det(A - \lambda E_3) &= (3-\lambda)(\lambda^2 - 3\lambda + 2) + 8(\lambda - 2) - 9(\lambda - 2) \\ &= (3-\lambda)(\lambda-2)(\lambda-1) - (\lambda-2) = -(\lambda-2)(\lambda^2 - 4\lambda + 4) = -(\lambda-2)^3 \end{aligned}$$

Also zerfällt P_A und $\lambda = 2$ ist der einzige EW von A. Nach (9.10) ist $A - 2E_3 =: T$ nilpotent und wegen $A - 2E_3 \neq 0$ gilt für den Nilpotenzgrad g von $A - 2E_3$ nach (9.14)

$$2 \leq g \leq 3.$$

Wir berechnen $(A - 2E_3)^2 = 0$, also $g = 2$. Da der Nilpotenzgrad unabhängig von der Wahl der Basis ist (es gilt $(B^{-1}AB)^j = B^{-1}A^jB$ für jedes $B \in \mathrm{GL}_3(\mathbb{R})$) hat auch die Jordansche Normalform von $A - 2E_3$ Nilpotenzgrad 2. Unter den nilpotenten Jordanschen Normalformen

$$\begin{pmatrix} \boxed{0} & 0 & 0 \\ 0 & \boxed{0} & 0 \\ 0 & 0 & \boxed{0} \end{pmatrix}, \quad \begin{pmatrix} \boxed{\begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \boxed{0} \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

hat aber nur $\left(\begin{smallmatrix}0&0&0\\1&0&0\\0&0&0\end{smallmatrix}\right)$ Nilpotenzgrad 2, also ist

$$\begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix}+2E_3=\begin{pmatrix}2&0&0\\1&2&0\\0&0&2\end{pmatrix}$$

die Jordansche Normalform von A. Eine Matrix $B\in\mathrm{GL}_3(\mathbb{R})$, für die $B^{-1}AB=\left(\begin{smallmatrix}2&0&0\\1&2&0\\0&0&2\end{smallmatrix}\right)$ gilt, kann man wie folgt bestimmen. Wir wählen zunächst einen beliebigen Vektor $v\notin\ker(T)$ für $T:=A-2E_3$, vgl. (9.14). Nach (9.15) ist dann $\mathcal{G}_1:=(v,Tv)$ Basis eines T-irreduziblen, T-invarianten Unterraums, genannt U_1, und $\mathrm{Mat}^{\mathcal{G}_1}_{\mathcal{G}_1}(A|U_1)=\left(\begin{smallmatrix}2&0\\1&2\end{smallmatrix}\right)$. Wir können z.B. $v=\left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$ wählen. Dann ist $Tv=\left(\begin{smallmatrix}1\\4\\3\end{smallmatrix}\right)$. Für Be_3 müssen wir einen von Tv unabhängigen EV von A wählen, z.B. $Be_3=\left(\begin{smallmatrix}3\\0\\1\end{smallmatrix}\right)$. Die Matrix B ist dann durch $Be_1=v=\left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$, $Be_2=Tv=\left(\begin{smallmatrix}1\\4\\3\end{smallmatrix}\right)$ und $Be_3=v=\left(\begin{smallmatrix}3\\0\\1\end{smallmatrix}\right)$ gegeben, d.h. $B=\left(\begin{smallmatrix}1&1&3\\0&4&0\\0&3&1\end{smallmatrix}\right)$. Man rechnet nach, daß $B^{-1}=\begin{pmatrix}1&2&-3\\0&\frac14&0\\0&-\frac34&1\end{pmatrix}$ und $B^{-1}AB=\left(\begin{smallmatrix}2&0&0\\1&2&0\\0&0&2\end{smallmatrix}\right)$ gilt.

Abgesehen von der Einsicht in die möglichen Typen von linearen Abbildungen, die die Sätze (9.18) und (9.19) vermitteln, ist die Jordansche Normalform auch von praktischem Nutzen. Es kommt z.B. oft vor, daß hohe Potenzen A^m einer Matrix $A\in K^{n\times n}$ berechnet werden sollen Für allgemeine Matrizen A ist das mit sehr großem Rechenaufwand verbunden. Kann man jedoch ein $B\in\mathrm{GL}_n(K)$ finden, so daß $A=B^{-1}JB$ gilt und J Jordansche Normalform hat, so gilt $A^m=B^{-1}J^mB$, wobei J^m sehr leicht zu berechnen ist. Eine damit zusammenhängende theoretische Erkenntnis ist der

(9.21) Satz (von Cayley-Hamilton). *Setzt man $L\in\mathrm{End}(V)$ in sein charakteristisches Polynom P_L ein, so erhält man $P_L(L)=0\in\mathrm{End}(V)$.*

Vorbemerkung: Wenn L diagonalisierbar ist, so ist (9.21) leicht einzusehen, vgl. Blatt 7, Aufgabe 3. Nach Satz (9.6)(ii) ist die Menge der diagonalisierbaren $A\in\mathbb{C}^{n\times n}$ dicht in $\mathbb{C}^{n\times n}$. Da die Abbildung $A\in\mathbb{C}^{n\times n}\to P_A(A)\in\mathbb{C}^{n\times n}$ stetig ist und auf der dichten Menge der diagonalisierbaren Matrizen $=0$ ist, folgt, daß $P_A(A)$ für alle $A\in\mathbb{C}^{n\times n}$ (und damit auch für alle $A\in\mathbb{R}^{n\times n}$) gilt.

Bew.: von (9.21): Wir führen den Beweis nur für den Fall durch, daß P_L zerfällt. Wie in Bem. 1 nach (9.20) gesagt wurde, ist das keine wirkliche Einschränkung. Wir verwenden (9.18). Offenbar genügt es zu zeigen, daß für jeden EW λ_t von L gilt:

$$P_L(L)\mid E'(\lambda_t)=0.$$

Wegen (9.14) gilt für $n_t:=\dim E'(\lambda_t)$:

$$(*)\qquad (L-\lambda_t\,\mathrm{id}_V\mid E'(\lambda_t))^{n_t}=0.$$

Andererseits enthält P_L den Faktor $(x-\lambda_t)^{n_t}$. Deshalb existiert ein $Q \in \operatorname{End}(V)$ mit

$$(**) \qquad P_L(L) = Q \circ (L - \lambda_t \operatorname{id}_V)^{n_t}.$$

Aus $(*)$ und $(**)$ folgt $P_L(L) \mid E'(\lambda_t) = 0$, wie behauptet.

Eine weitere wichtige Anwendung der Jordanschen Normalform betrifft das explizite Lösen von linearen Differentialgleichungssystemen mit konstanten Koeffizienten. Gegeben ist dabei eine Matrix $A \in \mathbb{C}^{n\times n}$ und gesucht sind alle n-Tupel $\begin{pmatrix} z_1(t) \\ \vdots \\ z_n(t) \end{pmatrix}$ von differenzierbaren Funktionen $z_i : \mathbb{R} \to \mathbb{C}(\simeq \mathbb{R}^2)$, für die

$$(*) \qquad \begin{pmatrix} z_1'(t) \\ \vdots \\ z_n'(t) \end{pmatrix} = A \begin{pmatrix} z_1(t) \\ \vdots \\ z_n(t) \end{pmatrix}$$

für alle $t \in \mathbb{R}$ gilt. Man führt die Funktion $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} : \mathbb{R} \to \mathbb{C}^n$ ein und schreibt für $(*)$ kürzer:

$$z' = Az.$$

Ist $B \in \mathrm{GL}_n(\mathbb{C})$, gilt $A = B^{-1}JB$ und ist $w : \mathbb{R} \to \mathbb{C}^n$ eine Lösung von $w' = Jw$, so ist $z := B^{-1}w$ eine Lösung von $(*)$:

$$z' = B^{-1}w' = B^{-1}Jw = B^{-1}JBz = Az.$$

Es genügt also, alle Lösungen der Gleichung

$$(**) \qquad w' = Jw$$

zu kennen. Wir betrachten zunächst den (nach (9.6) typischen) Fall, daß

$$J = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist.
Dann „entkoppeln" sich die Differentialgleichungen $(**)$ zu:

$$w_j' = \lambda_j w_j \text{ für } 1 \le j \le n$$

mit $\lambda_j \in \mathbb{C}$. Davon sind die Lösungen explizit angebbar,

$$w_j(t) = w_j(0)e^{\lambda_j t} \text{ für } 1 \le j \le n,$$

wobei die „Anfangswerte" $w_j(0), 1 \le j \le n$, die Lösung eindeutig bestimmen.

Folgerung: Sei $A \in \mathbb{C}^{n\times n}$ diagonalisierbar, $A = B^{-1}\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} B$, und $a \in \mathbb{C}^n$. Dann existiert genau eine Lösung $z : \mathbb{R} \to \mathbb{C}^n$ von $(*)$ mit $z(0) = a$, nämlich $z(t) = B^{-1}w(t)$, wobei $w_j(t) = c_j e^{\lambda_j t}$ für $1 \le j \le n$ und $c = Ba$.

Bem.:

1) Offenbar konvergieren die (Diagonal-)Matrizen $\sum\limits_{k=0}^{m} \frac{(tJ)^k}{k!}$ für $m \to \infty$ gegen die Matrix $\begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} =: \exp(tJ)$.

   Entsprechend bezeichnet man $\lim\limits_{m\to\infty} \sum\limits_{k=0}^{m} \frac{(tA)^k}{k!} =: \exp(tA)$ und wegen $A^k = B^{-1}J^kB$ gilt
   $$\exp(tA) = B^{-1}\exp(tJ)B.$$
   Die Lösungen $w(t)$ von $(**)$ lassen sich damit als
   $$w(t) = \exp(tJ)w(0)$$
   und die Lösungen $z(t) = B^{-1}w(t)$ von $(*)$ als
   $$z(t) = \exp(tA)z(0)$$
   schreiben.
2) Ist $A \in \mathbb{R}^{n\times n}$ reell diagonalisierbar, d.h. $A = B^{-1}JB$ mit $B \in \mathrm{GL}_n(\mathbb{R})$, $J = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ mit $\lambda_j \in \mathbb{R}$, so erhalten wir mit demselben Verfahren und denselben Formeln reelle Lösungen $z : \mathbb{R} \to \mathbb{R}^n$, $z = B^{-1}w$ mit $w_j(t) = e^{\lambda_j t} \in \mathbb{R}$ von $z' = Az$. In vielen Fällen wird $A \in \mathbb{R}^{n\times n}$ jedoch nur komplex diagonalisierbar sein, d.h. $A = B^{-1}JB$ mit $B \in \mathrm{GL}_n(\mathbb{C})$, $J = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ mit $\lambda_j \in \mathbb{C}$. Wir erhalten komplexe Lösungen $z : \mathbb{R} \to \mathbb{C}^n$ von $z' = Az$. Schreiben wir $z(t) = x(t) + iy(t)$ mit $x, y : \mathbb{R} \to \mathbb{R}^n$, so gilt auch $x' = Ax$, $y' = Ay$, d.h. $x = \mathrm{Re} z$ und $y = \mathrm{Im} z$ sind reelle Lösungen von $(*)$. Das kann man z.B. daran sehen, daß wegen $A = \overline{A}$ mit $z : \mathbb{R} \to \mathbb{C}^n$ auch $\overline{z} : \mathbb{R} \to \mathbb{C}^n$ Lösung von $(*)$ ist:
   $$\overline{z}' = \overline{Az} = \overline{A}\overline{z} = A\overline{z}.$$
   Da $(*)$ linear ist, sind dann auch $x = \frac{1}{2}(z + \overline{z})$ und $y = \frac{1}{2}(z - \overline{z})$ Lösungen von $(*)$.

Als nächstes betrachten wir den Fall, daß $A \in \mathbb{C}^{n\times n}$ eine Jordansche Normalform hat, die nur ein einziges „Jordankästchen“ hat, d.h. $A = B^{-1}JB$ mit $B \in \mathrm{GL}_n(\mathbb{C})$ und

$$J = T + \lambda E_n = \begin{pmatrix} \lambda & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \lambda \end{pmatrix}, \lambda \in \mathbb{C}.$$

Mit etwas probieren stellt man fest, daß man Lösungen $w : \mathbb{R} \to \mathbb{C}^n$ von $w' = Jw$ wie folgt erzeugen kann: Man wählt ein Polynom $p(t) =: \sum\limits_{k=0}^{n-1} \frac{c_{n-k}}{k!} t^k$, $c_{n-k} = p^{(k)}(0)$, vom Grad $n-1$ und setzt

$$w(t) := e^{\lambda t} \left(p^{(n-1)}(t), p^{(n-2)}(t), \dots, p(t) \right),$$

wobei $p^{(k)}(t)$ die k'te Ableitung von $p(t)$ bezeichnet. Es gilt dann nämlich:

$$w'(t) = \lambda w(t) + e^{\lambda t} \left(0, p^{(n-1)}(t), \dots, p'(t) \right) = Jw(t).$$

Man kann auch leicht sehen, daß jede Lösung w von $w' = Jw$ von diesem Typ ist: Setzt man $q(t) := e^{-\lambda t} w(t)$, so folgt $q' = Tq$. Differenziert man die Gleichung $q' = Tq$ weiter, so erhält man $q^{(k)} = T^k q$ und wegen $T^n = 0$: $q^{(n)} = 0$. Daraus folgt: q ist ein Polynom vom Grad $\leq n-1$. Man erhält also auf diese Art explizit alle Lösungen $w : \mathbb{R} \to \mathbb{C}^n$ von $w' = Jw$ und mit $z := B^{-1}w$ alle Lösungen von $z' = Az$. Die „Anfangswerte“ $w(0) = \left(p^{(n-1)}(0), \dots, p(0)\right) = (c_1, \dots, c_n) \in \mathbb{C}^n$ sind beliebig vorgebbar und bestimmen die Lösung $w(t)$ eindeutig, und analog für $z(t) := B^{-1}w(t)$.

Im allgemeinen hat die Jordansche Normalform J von A mehrere „Jordankästchen“ J_1^1, $\dots$, $J_s^{k_s}$ wie in (9.19)($**$). Die Lösungen von $w' = Jw$ setzen sich dann aus den Lösungen zu den einzelnen Jordankästchen $\left(w_t^{k_t}\right)' = J_t^{k_t} w_t^{k_t}$, die wir nach dem vorangehenden explizit kennen, zusammen.

Bsp.: Wir betrachten das lineare Differentialgleichungssystem

$$(*) \qquad x' = Ax$$

für Vektorfunktionen $x : \mathbb{R} \to \mathbb{R}^3$ und die Matrix $A = \left(\begin{smallmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{smallmatrix}\right)$ aus dem vorangehenden Beispiel. Explizit bedeutet $(*)$ dann:

$$\begin{aligned} x_1'(t) &= 3x_1(t) + 2x_2(t) - 3x_3(t) \\ x_2'(t) &= 4x_1(t) + 10x_2(t) - 12x_3(t) \\ x_3'(t) &= 3x_1(t) + 6x_2(t) - 7x_3(t) \end{aligned}$$

Wir haben die Jordansche Normalform J von A berechnet:

$$J = \left(\begin{smallmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{smallmatrix}\right)$$

Zu dem Jordankästchen $\left(\begin{smallmatrix} 2 & 0 \\ 1 & 2 \end{smallmatrix}\right)$ gehören die Lösungen

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = e^{2t} \left(\begin{smallmatrix} c_1 \\ c_2 + c_1 t \end{smallmatrix}\right)$$

und zu dem EV e_3 von J: $y_3(t) = c_3 e^{2t}$.

Die Lösungen von $y' = Jy$ sind dann $y(t) = e^{2t} \begin{pmatrix} c_1 \\ c_2 + c_1 t \\ c_3 \end{pmatrix}$ mit beliebigen $c_1, c_2, c_3 \in \mathbb{R}$.

Die Lösungen x von $(*)$ sind $x = B^{-1} y$ mit $B^{-1} = \begin{pmatrix} 1 & 2 & -3 \\ 0 & \frac{1}{4} & 0 \\ 0 & -\frac{3}{4} & 1 \end{pmatrix}$.

Bem.:

1) Auch hier kann man zeigen, daß für alle $t \in \mathbb{R}$ die Limites $\lim\limits_{m \to \infty} \sum\limits_{k=0}^{m} \frac{(tJ)^k}{k!} =: \exp(tJ)$ und $\lim\limits_{m \to \infty} \sum\limits_{k=0}^{m} \frac{(tA)^k}{k!} =: \exp(tA)$ existieren, daß $\exp(tA) = B^{-1} \exp(tJ) B$ gilt und daß die Lösungen w von $w' = Jw$ gerade durch $w(t) = \exp(tJ) w(0)$ gegeben sind. Dann folgt, daß die Lösungen von $z' = Az$ als $z(t) = \exp(tA) z(0)$ geschrieben werden können.
2) Ist $A \in \mathbb{R}^{n \times n}$, so zerfällt P_A über $\mathbb{C}$, so daß A eine komplexe Jordansche Normalform $J \in \mathbb{C}^{n \times n}$ besitzt. Ist $A = B^{-1} J B$ mit $B \in \mathrm{GL}_n(\mathbb{C})$ und sind $z = B^{-1} w$ die komplexen Lösungen von $z' = Az$, so erhält man die reellen Lösungen von $x' = Ax$ als die Realteile (oder – gleichwertig – als die Imaginärteile) der komplexen Lösungen.

10. Affine und euklidische Geometrie.

In der analytischen Geometrie beschreibt man nach Wahl eines Koordinatensystems Punkte durch n-Tupel von Zahlen ($n = 2$ für die Ebene, $n = 3$ für den 3-dimensionalen Raum), die man als Vektoren eines Vektorraums betrachten kann. Was diese Punkte eigentlich selbst sind, wird dabei mathematisch nicht gesagt und der Anschauung überlassen. Das Einführen von Koordinaten zerstört die Homogenität des Raumes durch die Auszeichnung eines Punktes als 0-Punkt des Koordinatensystems und durch Auszeichnung der Richtungen der Koordinatenachsen. In diesem Kapitel wird der „affine Punktraum“ eingeführt, mit dessen Hilfe der Unterschied zwischen Punkten und Vektoren und das Einführen von Koordinaten präzisiert werden kann.

(10.1) Def.: Ein affiner Raum $\mathcal{A}$ über einem K-Vektorraum V ist eine Menge $\mathcal{A} \neq \emptyset$ zusammen mit einer Abbildung $+ : \mathcal{A} \times V \to \mathcal{A}$, für die gilt:

(A_1) Für alle $a \in \mathcal{A}$, für alle $v, w \in V$ gilt: $(a + v) + w = a + (v + w)$.
(A_2) Für alle $a_1, a_2 \in \mathcal{A}$ existiert genau ein $v \in V$, so daß $a_2 = a_1 + v$ gilt.
Wir bezeichnen dieses $v \in V$ als $v =: \overrightarrow{a_1 a_2}$.

Bez.: Die Dimension des affinen Raums $\mathcal{A}$ ist die Dimension des zugehörigen Vektorraums V.

Bem.:

1) Aus (A_1) und (A_2) folgt $\overrightarrow{aa} = 0$, $\overrightarrow{a_1 a_2} = -\overrightarrow{a_2 a_1}$ und $\overrightarrow{a_1 a_2} + \overrightarrow{a_2 a_3} = \overrightarrow{a_1 a_3}$.
2) Die Vorstellung hinter der Abbildung $+ : \mathcal{A} \times V \to \mathcal{A}$, $(a, v) \to a + v$ ist, daß der Vektor v am Punkt a angetragen einen neuen Punkt, genannt $a + v$, bestimmt. Entsprechend nennt man $v = \overrightarrow{a_1 a_2}$ den Verbindungsvektor von a_1 nach a_2. Die Bezeichnung $+$ für die Abbildung $\mathcal{A} \times V \to \mathcal{A}$ ist zwar oft praktisch, aber auch irreführend, da sie von der – gleich bezeichneten – Addition in V unterschieden werden muß. In der Gleichung in (A_1) beziehen sich etwa die ersten drei Zeichen $+$ auf die Abbildung $\mathcal{A} \times V \to \mathcal{A}$, das letzte dagegen auf die Addition in V.

Im folgenden Standardbeispiel eines affinen Raums verschwindet dieser Unterschied, und das ist auch der Grund für diese Bezeichnung.

(10.2) Beispiel: Sei V K-Vektorraum. Wir setzen $\mathcal{A} := V$ und nehmen als Abbildung $\mathcal{A} \times V \to V$ die übliche Addition $V \times V \to V$. Dann ist (A_1) gerade das Assoziativgesetz bzgl. $+$, während (A_2) aus der (eindeutigen) Existenz des additiven Inversen folgt.

Hier ein Beispiel, in dem der affine Raum nicht direkt ein Vektorraum „ist“:

(10.3) Beispiel: Sei $A \in K^{n \times n}$ eine Matrix, $w_0 \in K^n$ und

$$\begin{aligned} \mathcal{A} &:= \{v \in K^n \mid Av = w_0\} \neq \emptyset, \\ V &:= \{u \in K^n \mid Au = 0\} \end{aligned}$$

Wir definieren $\mathcal{A} \times V \to \mathcal{A}$ durch die übliche Addition in K^n. Dann gilt in der Tat: $v \in \mathcal{A}$, $u \in V \Rightarrow v + u \in \mathcal{A}$. Denn $A(v + a) = Av + Au = Av = w_0$. Sind $v_1, v_2 \in \mathcal{A}$, so ist der Verbindungsvektor $\overrightarrow{v_1 v_2}$ gerade $v_2 - v_1$, da $v_1 + (v_1 - v_1) = v_1$ gilt, und es gilt in der Tat $\overrightarrow{v_1 v_2} = v_1 - v_1 \in V$, denn $A(v_2 - v_1) = Av_2 - Av_1 = w_0 - w_0 = 0$. In diesem Fall ist

$$\dim \mathcal{A} := \dim V = n - \operatorname{rg}(A).$$

Mengen $\mathcal{A}$ wie in (10.3) kamen schon nach (3.25) vor und wurden damals „affine Unterräume des Vektorraums K^n“ genannt. Wenn man K^n wie in (10.2) als affinen Raum interpretiert, so ist dies ein Spezialfall von

(10.4) Def.: Eine Teilmenge $\mathcal{U}$ eines affinen Raums $\mathcal{A}$ heißt affiner Unterraum von $\mathcal{A}$, falls ein Untervektorraum $V_{\mathcal{U}}$ von V und ein $a_0 \in \mathcal{U}$ existiert, so daß

$$\mathcal{U} = \{a_0 + v \mid v \in V_{\mathcal{U}}\}$$

gilt.

(10.5) Lemma. Sei $\mathcal{U} \subseteq \mathcal{A}$ affiner Unterraum und seien $a_0, V_{\mathcal{U}}$ wie in (10.4). Dann gilt für jedes $a_1 \in \mathcal{U}$:

$$\mathcal{U} = \{a_1 + v \mid v \in V_{\mathcal{U}}\}.$$

Bem.: Das zeigt, daß im Fall eines affinen Unterraums $\mathcal{U}$ die Menge der Verbindungsvektoren von einem festen Punkt in $\mathcal{U}$ zu beliebigen Punkten in $\mathcal{U}$ ein Untervektorraum von V ist, der unabhängig von der Wahl des festen Punktes in $\mathcal{U}$ ist und nur von $\mathcal{U}$ abhängt. $V_{\mathcal{U}}$ heißt auch der Richtungsraum von $\mathcal{U}$. Wir definieren: $\dim \mathcal{U} := \dim V_{\mathcal{U}}$.

Bew.:

(i) $\mathcal{U} \subseteq \{a_1 + v \mid v \in V_{\mathcal{U}}\}$: Sei $b \in \mathcal{U}$. Wir müssen zeigen, daß $\overrightarrow{a_1 b} \in V_{\mathcal{U}}$ gilt (denn das impliziert $b = a_1 + v$ mit $v = \overrightarrow{a_1 b} \in V_{\mathcal{U}}$). Wegen $b \in \mathcal{U}$ und $a_1 \in \mathcal{U}$ gelten $\overrightarrow{a_0 b} \in V_{\mathcal{U}}$ und $\overrightarrow{a_0 a_1} = -\overrightarrow{a_1 a_0} \in V_{\mathcal{U}}$. Da $V_{\mathcal{U}}$ Untervektorraum ist, folgt $\overrightarrow{a_1 a_0} + \overrightarrow{a_0 b} = \overrightarrow{a_1 b} \in V_{\mathcal{U}}$, wie behauptet.

(ii) $\{a_1 + v \mid v \in V_{\mathcal{U}}\} \subseteq \mathcal{U}$: Sei $v \in V_{\mathcal{U}}$. Dann gilt

$$a_1 + v = (a_0 + \overrightarrow{a_0 a_1}) + v \overset{(A_2)}{=} a_0 + (\overrightarrow{a_0 a_1} + v)\,.$$

Wegen $a_1 \in \mathcal{U}$ folgt $\overrightarrow{a_0 a_1} \in V_{\mathcal{U}}$, und da $V_{\mathcal{U}}$ Untervektorraum ist, folgt $\overrightarrow{a_0 a_1} + v \in V_{\mathcal{U}}$. Also gilt $a_1 + v = a_0 + (\overrightarrow{a_0 a_1} + v) \in \mathcal{U}$.

Bez.: Ein affiner Raum (oder Unterraum) der Dimension 1 heißt (affine) Gerade, ein affiner Raum (oder Unterraum) der Dimension 2 heißt (affine) Ebene. Ist $\mathcal{A}$ n-dimensionaler affiner Raum und $\mathcal{U} \subseteq \mathcal{A}$ $(n-1)$-dimensionaler affiner Unterraum, so heißt $\mathcal{U}$ affine Hyperebene in $\mathcal{A}$.

Bem.: Ist $\mathcal{U} \subseteq \mathcal{A}$ affiner Unterraum, so folgt aus (10.5), daß für alle $a \in \mathcal{U}$ gilt:

$$\mathcal{U} = a + V_{\mathcal{U}} := \{a + v \mid v \in V_{\mathcal{U}}\}.$$

Ein 0-dimensionaler affiner Unterraum besteht aus genau einem Punkt.

(10.6) Fakt: Sind $\mathcal{U}_1, \mathcal{U}_2$ affine Unterräume von $\mathcal{A}$ und gilt $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \emptyset$, so ist $\mathcal{U}_1 \cap \mathcal{U}_2$ affiner Unterraum von $\mathcal{A}$, und es gilt

$$V_{\mathcal{U}_1 \cap \mathcal{U}_2} = V_{\mathcal{U}_1} \cap V_{\mathcal{U}_2}.$$

Bew.: Wähle $a \in \mathcal{U}_1 \cap \mathcal{U}_2$. Dann folgt aus $\mathcal{U}_1 = a + V_{\mathcal{U}_1}$ und $\mathcal{U}_2 = a + V_{\mathcal{U}_2}$, daß $\mathcal{U}_1 \cap \mathcal{U}_2 = a + (V_{\mathcal{U}_1} \cap V_{\mathcal{U}_2})$ gilt.

(10.7) Fakt: Sind $\mathcal{U}_1, \mathcal{U}_2$ affine Unterräume von $\mathcal{A}$ und gilt $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \emptyset$, so gilt $\dim(\mathcal{U}_1 \cap \mathcal{U}_2) = \dim(V_{\mathcal{U}_1} \cap V_{\mathcal{U}_2})$.

(10.8) Def.: Zwei affine Unterräume $\mathcal{U}_1, \mathcal{U}_2$ von $\mathcal{A}$ heißen parallel, falls $V_{\mathcal{U}_1} \subseteq V_{\mathcal{U}_2}$ oder $V_{\mathcal{U}_2} \subseteq V_{\mathcal{U}_1}$ gilt.

(10.9) Fakt: Sei $\mathcal{A}$ endlichdimensionaler affiner Raum, $\mathcal{U} \subseteq \mathcal{A}$ affiner Unterraum und $\mathcal{H} \subseteq \mathcal{A}$ affine Hyperebene. Gilt $\mathcal{U} \cap \mathcal{H} = \emptyset$, so sind $\mathcal{U}$ und $\mathcal{H}$ parallel.

Bem.:

1) Gilt $\mathcal{U} \cap \mathcal{H} \neq \emptyset$, so folgt aus (10.7), daß $\mathcal{U} \cap \mathcal{H}$ ein affiner Unterraum mit $\dim(\mathcal{U} \cap \mathcal{H}) = \dim \mathcal{U} - 1$ ist.
2) Ist $\dim \mathcal{A} = 2$, so besagt (10.9), daß zwei Geraden in der affinen Ebene $\mathcal{A}$ entweder parallel sind oder genau einen Schnittpunkt haben.

Bew.: Wir nehmen an, $\mathcal{U}$ und $\mathcal{H}$ seinen nicht parallel. Dann folgt $V_{\mathcal{U}} + V_{\mathcal{H}} = V_{\mathcal{A}}$. Ist $a \in \mathcal{U}$, $b \in \mathcal{H}$, so läßt sich der Vektor $\overrightarrow{ab} \in V_{\mathcal{A}}$ darstellen als

$$\overrightarrow{ab} = v_1 + v_2$$

mit $v_1 \in V_{\mathcal{U}}$, $v_2 \in V_{\mathcal{H}}$. Daraus folgt

$$b = a + (v_1 + v_2) \overset{(A_1)}{=} (a + v_1) + v_2$$

und

$$b + (-v_2) \overset{(A_1)}{=} (a + v_1) + (v_2 - v_2) \overset{(A_1)}{=} a + v_1.$$

Wegen $v_1 \in V_{\mathcal{U}}$, $-v_2 \in V_{\mathcal{H}}$ folgt $a + v_1 = b + (-v_2) \in \mathcal{U} \cap \mathcal{H}$, d.h. $\mathcal{U} \cap \mathcal{H} \neq \emptyset$.

Def.: Sei $\mathcal{A}$ affiner Raum und $(a, b, c) \in \mathcal{A}^3$ ein (geordnetes) Tripel von Punkten in $\mathcal{A}$.

(a) Die Punkte $a, b, c \in \mathcal{A}$ heißen kollinear, falls es eine affine Gerade $\mathcal{G} \subseteq \mathcal{A}$ gibt mit $\{a, b, c\} \subseteq \mathcal{G}$.
(b) Ist $(a, b, c) \in \mathcal{A}^3$ ein kollineares Punktetripel und gilt $a \neq b$, so heißt der durch die Gleichung $\overrightarrow{ac} = \lambda \overrightarrow{ab}$ eindeutig bestimmte Skalar $\lambda \in K$ das Teilverhältnis des kollinearen Punktetripels (a, b, c), bezeichnet durch $\lambda =: TV(a, b, c)$.

Bsp.:

1) c heißt der Mittelpunkt der Strecke von a nach b, wenn $TV(a, b, c) = \frac{1}{2}$ gilt. Das ist äquivalent zu $c = a + \frac{1}{2}\overrightarrow{ab}$.

2) Betrachten wir $\mathbb{R}^2$ nach (10.2) als affinen Raum, so sind $a = (0,1)$, $c = (1,1)$ und $b = (3,1)$ kollinear, da a, b, c auf der affinen Geraden $\mathcal{G} = a + (\mathbb{R} \times \{0\})$ liegen. Wegen

$$\overrightarrow{ac} = (1,0) = \frac{1}{3}\overrightarrow{ab}$$

gilt $TV(a,b,c) = \frac{1}{3}$.

Bem.: Im Fall $K = \mathbb{R}$ hat $TV(a,b,c)$ folgende anschauliche Bedeutung, die für den Namen „Teilverhältnis" verantwortlich ist: Wählt man auf V ein Skalarprodukt (oder – allgemeiner – eine Norm), so definiert man für $a, b \in \mathcal{A}$ die <u>Streckenlänge</u> von a nach b durch $\|\overrightarrow{ab}\|$. Dann gilt

$$|TV(a,b,c)| = \frac{\|\overrightarrow{ac}\|}{\|\overrightarrow{ab}\|},$$

d.h. $|TV(a,b,c)|$ ist das Verhältnis der Streckenlänge von a nach c zur Streckenlänge von a nach b. Dieses Verhältnis ist unabhängig von der Wahl des Skalarprodukts (oder der Norm). Das Vorzeichen von $TV(a,b,c)$ hängt davon ab, ob a auf $\mathcal{G}_{a,b}$ zwischen b und c liegt ($\Rightarrow TV(a,b,c) < 0$) oder nicht.

(10.10) Strahlensatz. *Seien $(a, b_1, c_1) \in \mathcal{A}^3$ und $(a, b_2, c_2) \in \mathcal{A}^3$ kollineare Punktetripel auf verschiedenen Geraden durch $a \notin \{b_1, b_2, c_1, c_2\}$. Dann sind $\mathcal{G}_{b_1,b_2}$ und $\mathcal{G}_{c_1,c_2}$ genau dann parallel, wenn $TV(a,b_1,c_1) = TV(a,b_2,c_2)$ gilt.*

Aufgabe: Fertigen Sie eine Skizze der in (10.10) beschriebenen Situation an.

Bew.: Nach Definition gilt

$$TV(a,b_1,c_1) = \lambda \Leftrightarrow \overrightarrow{ac_1} = \lambda\overrightarrow{ab_1}$$

und

$$TV(a,b_2,c_2) = \mu \Leftrightarrow \overrightarrow{ac_2} = \mu\overrightarrow{ab_2}.$$

Die zwei Geraden $\mathcal{G}_{b_1,b_2}$ und $\mathcal{G}_{c_1,c_2}$ sind genau dann parallel, wenn die Richtungsvektoren $\overrightarrow{b_1b_2}$ und $\overrightarrow{c_1c_2}$ linear abhängig sind. Aus $\overrightarrow{ac_1} + \overrightarrow{c_1c_2} = \overrightarrow{ac_2}$ folgt

$$\overrightarrow{c_1c_2} = \overrightarrow{ac_2} - \overrightarrow{ac_1} = \mu\overrightarrow{ab_2} - \lambda\overrightarrow{ab_1}$$

und ebenso

$$\overrightarrow{b_1b_2} = \overrightarrow{ab_2} - \overrightarrow{ab_1}.$$

Da nach Voraussetzung $\mathcal{G}_{a,b_1} \neq \mathcal{G}_{a,b_2}$ gilt, sind $\overrightarrow{ab_1}$ und $\overrightarrow{ab_2}$ linear unabhängig. Die vorangehenden Gleichungen stellen $\overrightarrow{b_1b_2}$ und $\overrightarrow{c_1c_2}$ als Linearkombinationen dieser linear unabhängigen Vektoren $\overrightarrow{ab_1}$ und $\overrightarrow{ab_2}$ dar. Also sind $\overrightarrow{b_1b_2}$ und $\overrightarrow{c_1c_2}$ genau dann linear abhängig ($\Leftrightarrow \mathcal{G}_{b_1,b_2}$ und $\mathcal{G}_{c_1,c_2}$ sind parallel), wenn $\det\begin{pmatrix} \mu & 1 \\ -\lambda & -1 \end{pmatrix} = 0$ gilt, d.h. wenn $\lambda = TV(a,b_1,c_1) = TV(a,b_2,c_2) = \mu$ gilt.

In der Mathematik sind stets neben den Objekten einer Theorie (hier: neben den affinen Räumen) auch die „strukturerhaltenden“ Abbildungen zwischen diesen Objekten wichtig (hier: die in (10.11) definierten „affinen Abbildungen“).

(10.11) Def.: Für $i = 0, 1$ seien $\mathcal{A}_i$ affine Räume über den K-Vektorräumen V_i. Eine Abbildung $A : \mathcal{A}_1 \to \mathcal{A}_2$ heißt affin, wenn es einen K-Vektorraumhomomorphismus $L_A \in \mathrm{Hom}(V_0, V_1)$ gibt, so daß für alle $a, b \in \mathcal{A}_0$ gilt:

$$L_A(\overrightarrow{ab}) = \overrightarrow{A(a)A(b)}.$$

Eine bijektive affine Abbildung heißt Affinität.

(10.12) Bsp.: Seien $\mathcal{A}$, $\mathcal{A}_0$, $\mathcal{A}_1$ affine Räume über den K-Vektorräumen V, V_0, V_1.

1) $A := \mathrm{id}_{\mathcal{A}}$ ist affin mit $L_A = \mathrm{id}_V$.
2) Sei $v_0 \in V$ und $A : \mathcal{A} \to \mathcal{A}$ definiert durch
$$A(a) = a + v_0.$$
Dann ist A affin mit $L_A = \mathrm{id}_V$. Eine solche Abbildung heißt Translation.
3) Seien $a_0 \in \mathcal{A}_0$, $a_1 \in \mathcal{A}_1$, $L \in \mathrm{Hom}(V_0, V_1)$. Sei $A : \mathcal{A}_0 \to \mathcal{A}_1$ definiert durch
$$A(a) = a_1 + L(\overrightarrow{a_0a}).$$
Dann ist A affin und $L_A = L$.

Beweis zu 3): Wir müssen zeigen, daß für alle $a, b \in \mathcal{A}_0$ gilt:

$$\overrightarrow{A(a)A(b)} = L(\overrightarrow{ab}).$$

Der Vektor $v = \overrightarrow{A(a)A(b)} \in V_1$ ist durch die Gleichung

$$a_1 + L(\overrightarrow{a_0a}) + v = a_1 + L(\overrightarrow{a_0b})$$

definiert. Daraus folgt mit (A_2):

$$L(\overrightarrow{a_0a}) + v = L(\overrightarrow{a_0b}),$$

also

$$v = L(-\overrightarrow{a_0a}) + L(\overrightarrow{a_0b}) = L(\overrightarrow{aa_0} + \overrightarrow{a_0b}) = L(\overrightarrow{ab}),$$

wie behauptet.

Bem.: Ist $A : \mathcal{A}_0 \to \mathcal{A}_1$ affin, so gilt für jedes $a_0 \in \mathcal{A}_0$

$$(*) \qquad A(a) = A(a_0) + L_A(\overrightarrow{a_0a}),$$

d.h. A ist wie in (10.12), 3) darstellbar, wobei $a_0 \in \mathcal{A}_0$ beliebig ist, $a_1 := A(a_0)$ und $L = L_A$.
Begründung von $(*)$: $A(a) = A(a_0) + v$, wobei $v = \overrightarrow{A(a_0)A(a)} \overset{(10.11)}{=} L_A(\overrightarrow{a_0a})$.

Folgende Aussagen sind leicht einzusehen:

(10.13) Eine affine Abbildung $A : \mathcal{A} \to \mathcal{A}$ ist genau dann eine Translation, wenn $L_A = \mathrm{id}_V$ gilt. Eine affine Abbildung $A : \mathcal{A}_0 \to \mathcal{A}_1$ ist genau dann eine Affinität (d.h. A ist bijektiv), wenn $L_A \in \mathrm{Hom}(V_0, V_1)$ ein Isomorphismus ist. Es gilt dann $L_{A^{-1}} = (L_A)^{-1}$. Eine affine Abbildung $A : \mathcal{A} \to \mathcal{A}$ heißt Streckung mit Zentrum $a_0 \in \mathcal{A}$ und Streckfaktor $\lambda \in K$, falls $A(a_0 + v) = a_0 + \lambda v$ für alle $v \in V$ gilt.

(10.14) Ist $A : \mathcal{A}_0 \to \mathcal{A}_1$ affin und $\mathcal{U} \subseteq \mathcal{A}_0$ affiner Unterraum, so ist $A(\mathcal{U}) \subseteq \mathcal{A}_1$ affiner Unterraum. Es gilt: $\mathcal{U} = a_0 + V_{\mathcal{U}} \Rightarrow A(\mathcal{U}) = A(a_0) + L_A(V_{\mathcal{U}})$. Ist $\mathcal{U} \subseteq \mathcal{A}_1$ affiner Unterraum und ist die Urbildmenge

$$A^{-1}(\mathcal{U}) = \{a \in \mathcal{A} \mid A(a) \in \mathcal{U}\}$$

nicht leer, so ist $A^{-1}(\mathcal{U})$ affiner Unterraum von $\mathcal{A}_0$. Ist $a_0 \in A^{-1}(\mathcal{U})$, so gilt

$$A^{-1}(\mathcal{U}) = a_0 + L_A^{-1}(V_{\mathcal{U}}).$$

(10.15) Ist $A : \mathcal{A}_0 \to \mathcal{A}_1$ affin, sind $a, b, c \in \mathcal{A}_0$ und gilt $A(a) \neq A(b)$, so folgt:

$$TV(a, b, c) = TV(A(a), A(b), A(c)).$$

Eine Affinität zwischen zwei affinen Räumen $\mathcal{A}_0$ und $\mathcal{A}_1$ erlaubt es, alle Aussagen, Argumente etc. der affinen Geometrie in $\mathcal{A}_0$ in entsprechende Aussagen, Argumente etc. in $\mathcal{A}_1$ zu verwandeln. Nach (10.13) existiert genau dann eine Affinität von $\mathcal{A}_0$ nach $\mathcal{A}_1$, wenn $\mathcal{A}_0$ und $\mathcal{A}_1$ zu Vektorräumen über dem gleichen Körper gehören und $\dim \mathcal{A}_0 = \dim \mathcal{A}_1$ gilt. Es ist wichtig, zu erkennen, ob eine geometrische Aussage zur affinen Geometrie gehört. Die Aussage, daß sich die Seitenhalbierenden eines Dreiecks in einem Punkt schneiden, der diese Seitenhalbierenden im Verhältnis 2:1 teilt, gehört etwa zur affinen Geometrie, da sie sich nur auf Geraden und Teilverhältnisse bezieht. Man kann die Aussage direkt im affinen Raum $\mathbb{R}^2$ beweisen, aber folgender Weg ist wohl einfacher: Da je zwei Dreiecke durch eine Affinität aufeinander abgebildet werden können (siehe (10.17)), genügt es, die Aussage für gleichseitige Dreiecke zu beweisen. Das ist mit (euklidischen) Symmetrieargumenten leicht möglich.

(10.16) Def. Sei $\mathcal{A}$ affiner Raum über dem K-Vektorraum V, $\dim V = n > 0$. Ein $(n+1)$-Tupel $(a_0, \dots, a_n) \in \mathcal{A}^{n+1}$ von Punkten $a_0, \dots, a_n \in \mathcal{A}$ heißt Koordinatensystem für $\mathcal{A}$, falls die Vektoren $\overrightarrow{a_0 a_1}, \dots, \overrightarrow{a_0 a_n} \in V$ linear unabhängig sind (und damit eine Basis von V bilden). Ist $x \in V$ ein beliebiger Punkt, so gilt

$$\overrightarrow{a_0 x} = \sum_{i=1}^{n} x_i \overrightarrow{a_0 a_i}$$

mit durch x eindeutig bestimmten $x_i \in K$, $1 \leq i \leq n$. Diese x_i, $1 \leq i \leq n$, heißen die Koordinaten von x bezüglich des Koordinatensystems $(a_0, \dots, a_n)$ und $(x_1, \dots, x_n) \in K^n$ heißt der Koordinatenvektor von x (bzgl. $(a_0, \dots, a_n)$).

Wichtige Bem.: Die Abbildung $A : x \in \mathcal{A} \to (x_1, \dots, x_n) \in K^n$ ist eine Affinität, die $\mathcal{A}$ mit dem (nach (10.2)) als affinen Raum betrachteten K^n identifiziert. Dabei gilt $A(a_0) = 0 \in K^n$ und $A(a_i) = e_i \in K^n$ für $1 \leq i \leq n$. Das erlaubt, alle Probleme der affinen Geometrie in einem K^n zu betrachten. Das hat zwei Vorteile: Wir können im üblichen K^n rechnen (ohne die etwas umständlichen Axiome (A_1) und (A_2) zu bemühen) und wir können unser Koordinatensystem dem gegebenen Problem anpassen. Wir haben dadurch natürlich den „Vorteil“ des affinen Raums, keinen ausgezeichneten Punkt und keine ausgezeichneten Richtungen zu besitzen, wieder verloren.

(10.17) Fakt. Seien $\mathcal{A}$ und $\mathcal{B}$ affine Räume (über dem gleichen Körper) und $(a_0, \ldots, a_n) \in \mathcal{A}^{n+1}$ ein Koordinatensystem für $\mathcal{A}$ und $(b_0, \ldots, b_n) \in \mathcal{B}^{n+1}$. Dann gibt es genau eine affine Abbildung $A : \mathcal{A} \to \mathcal{B}$ mit $A(a_i) = b_i$ für $0 \leq i \leq n$. A ist genau dann eine Affinität, wenn $(b_0, \ldots, b_n)$ ein Koordinatensystem für $\mathcal{B}$ ist.

Bew.: Es seien V und W die K-Vektorräume zu $\mathcal{A}$ und $\mathcal{B}$, und es sei $L \in \mathrm{Hom}(V, W)$ definiert durch $L(\overrightarrow{a_0a_i}) = \overrightarrow{b_0b_i}$ für $1 \leq i \leq n$, vgl. (4.6). Wir definieren $A : \mathcal{A} \to \mathcal{B}$ durch

$$A(a) = b_0 + L(\overrightarrow{a_0a}).$$

Nach (10.12) ist A affin, und es gilt $A(a_0) = b_0$ und $A(a_i) = b_0 + L(\overrightarrow{a_0a_i}) = b_0 + \overrightarrow{b_0b_i} = b_i$ für $1 \leq i \leq n$. Die restlichen Aussagen lassen sich ähnlich einsehen.

Die Begriffe „affiner Unterraum“ und „affine Abbildung“ wurden in (10.4) und (10.11) durch Rückgriff auf die Begriffe „Untervektorraum“ und „Vektorraumhomomorphismus“ der linearen Algebra definiert. Es soll nun versucht werden, diese Begriffe allein mit Hilfe des affin-geometrischen Grundbegriffs „Gerade“ zu charakterisieren.

(10.18) Def.: Eine Teilmenge M eines affinen Raums $\mathcal{A}$ heißt affin abgeschlossen, wenn M mit je zwei Punkten $a \neq b$ auch die Gerade

$$\mathcal{G}_{a,b} = \{a + \lambda\overrightarrow{ab} \mid \lambda \in K\}$$

durch a und b enthält (d.h. kürzer: $\{a, b\} \subseteq M$ und $a \neq b \Rightarrow \mathcal{G}_{a,b} \subseteq M$).

(10.19) Satz. *Im Körper K gelte $1 + 1 \neq 0$. Dann gilt: Eine Teilmenge $M \neq \emptyset$ von $\mathcal{A}$ ist genau dann affin abgeschlossen, wenn M ein affiner Unterraum von $\mathcal{A}$ ist.*

Bew.: Offensichtlich ist jeder affine Unterraum affin abgeschlossen. Sei umgekehrt M affin abgeschlossen und $a_0 \in M$. Wir zeigen, daß

$$U := \{\overrightarrow{a_0a} \mid a \in M\}$$

Untervektorraum von V ist. Dann gilt $M = a_0 + U$ und das zeigt, daß M affiner Unterraum ist.

(i) Es gilt $0 = \overrightarrow{a_0a_0} \in U$.
(ii) Ist $v \in U \setminus \{0\}$, $v = \overrightarrow{a_0a}$ mit $a \in M$, so gilt nach Voraussetzung $\mathcal{G}_{a_0,a} \subseteq M$, d.h. $a_0 + (\lambda v) =: a_\lambda \in M$ für alle $\lambda \in K$. Daraus folgt für alle $\lambda \in K$: $\overrightarrow{a_0a_\lambda} = \lambda v \in U$.
(iii) Seien $v = \overrightarrow{a_0a} \in U$, $w = \overrightarrow{a_0b} \in U$ mit $a, b \in M$. Gilt $v = w$, so folgt nach (ii): $v + w = (1+1)v \in U$. Wir können also annehmen, daß $v \neq w$ und damit $a \neq b$ gilt. Nach Voraussetzung gilt dann $\mathcal{G}_{a,b} \subseteq M$. Wegen $\overrightarrow{ab} = \overrightarrow{aa_0} + \overrightarrow{a_0b} = -\overrightarrow{a_0a} + \overrightarrow{a_0b} = w - v$ gilt dann

$$\mathcal{G}_{a,b} = \{a + \lambda(w - v) \mid \lambda \in K\} = \{a_0 + (1 - \lambda)v + \lambda w \mid \lambda \in K\} \subseteq M.$$

Mit $\frac{1}{2} \in K$ sei das multiplikative Inverse von $1+1 \neq 0$ bezeichnet. Für $\lambda = \frac{1}{2}$ erhalten wir $a_0 + \left(1-\frac{1}{2}\right) v + \frac{1}{2} w = a_0 + \frac{1}{2}(v+w) \in M$, also $\frac{1}{2}(v+w) \in U$. Nach (ii) folgt $2 \cdot \left(\frac{1}{2}(v+w)\right) = v+w \in U$.

(10.20) Fundamentalsatz der affinen Geometrie (über $\mathbb{R}$). *Seien $\mathcal{A}, \mathcal{A}'$ affine Räume über den reellen Vektorräumen V, V'. Sei $\dim \mathcal{A} \geq 2$ und $F : \mathcal{A} \to \mathcal{A}'$ eine Bijektion, die Geraden auf Geraden abbildet. Dann ist F eine Affinität.*

Der nicht ganz einfache Beweis wird durch eine Reihe von Hilfssätzen erledigt.

(10.21) Lemma. Sei $\mathcal{E} \subseteq \mathcal{A}$ affine Ebene. Dann ist $F(\mathcal{E}) \subseteq \mathcal{A}'$ eine affine Ebene.

Bew.: Sei (a_0, a_1, a_2) ein Koordinatensystem für $\mathcal{E}$. Dann gilt

$$\mathcal{G}_{a_0,a_1} \cap \mathcal{G}_{a_0,a_2} = \{a_0\}.$$

Da F injektiv ist und $F(\mathcal{G}_{a_0,a_1}) = \mathcal{G}_{F(a_0),F(a_1)}$, $F(\mathcal{G}_{a_0,a_2}) = \mathcal{G}_{F(a_0),F(a_2)}$ gilt, folgt

$$\mathcal{G}_{F(a_0),F(a_1)} \cap \mathcal{G}_{F(a_0),F(a_2)} = \{F(a_0)\}.$$

Deshalb existiert genau eine Ebene $\mathcal{E}' \subseteq \mathcal{A}'$, die $\mathcal{G}_{F(a_0),F(a_1)}$ und $\mathcal{G}_{F(a_0),F(a_2)}$ enthält. Ist $a \in \mathcal{E} \setminus (\mathcal{G}_{a_0,a_1} \cup \mathcal{G}_{a_0,a_2})$, so existiert eine Gerade $\mathcal{G} \subseteq \mathcal{E}$ mit $a \in \mathcal{G}$, die $\mathcal{G}_{a_0,a_1}$ und $\mathcal{G}_{a_0,a_2}$ schneidet. Da F Geraden auf Geraden abbildet, folgt $F(a) \in F(\mathcal{G}) \subseteq \mathcal{E}'$. Damit haben wir $F(\mathcal{E}) \subseteq \mathcal{E}'$ bewiesen. Ähnlich zeigt man $\mathcal{E}' \subseteq F(\mathcal{E})$, aber wir werden das nicht benötigen.

(10.22) Lemma. F bildet parallele Geraden auf parallele Geraden ab.

Bew.: Seien $\mathcal{G} \neq \mathcal{H}$ parallele Geraden in $\mathcal{A}$. Dann existiert eine Ebene $\mathcal{E} \subseteq \mathcal{A}$ mit $\mathcal{G} \cup \mathcal{H} \subseteq \mathcal{E}$. Nach (10.21) existiert eine Ebene $\mathcal{E}' \subseteq \mathcal{A}'$, so daß

$$F(\mathcal{G}) \cup F(\mathcal{H}) \subseteq F(\mathcal{E}) \subseteq \mathcal{E}'$$

gilt. Da F injektiv ist, gilt $F(\mathcal{G}) \cap F(\mathcal{H}) = F(\mathcal{G} \cap \mathcal{H})$.

Bem. 2 nach (10.9) impliziert, daß $\mathcal{G} \cap \mathcal{H} = \emptyset$ gilt. Also gilt $F(\mathcal{G}) \cap F(\mathcal{H}) = \emptyset$ und die gleiche Bem. 2 zeigt dann, daß die Geraden $F(\mathcal{G}) \subseteq \mathcal{E}'$ und $F(\mathcal{H}) \subseteq \mathcal{E}'$ parallel sind.

Um den Fundamentalsatz (10.20) zu beweisen, wählen wir einen festen Punkt $a_0 \in \mathcal{A}$, setzen $a_0' := F(a_0) \in \mathcal{A}'$ und definieren $G : V \to V'$ durch: Für alle $v \in V$ gelte

$$F(a_0 + v) = a_0' + G(v)$$

oder äquivalent

$$G(v) = \overrightarrow{F(a_0)F(a_0+v)}.$$

Dann gilt offensichtlich $G(0) = 0$. Der Rest des Beweises besteht darin, zu zeigen, daß $G \in \mathrm{Hom}(V, V')$ ($\Rightarrow$ (10.20)) gilt. Der Beweis der wichtigen Lemmata (10.23) und (10.25) beruht auf geometrischen Ideen, die am besten durch eine Skizze (!) klar werden.

(10.23) Lemma. Für alle linear unabhängigen v, w in V gilt

$$G(v+w) = G(v) + G(w).$$

Bew.: Wir betrachten die Geraden

$$\mathcal{G}'_v = \{a_0 + w + \lambda v \mid \lambda \in \mathbb{R}\} \text{ und } \mathcal{G}'_w = \{a_0 + v + \mu w \mid \mu \in \mathbb{R}\},$$

die einander im Punkt $a_0 + v + w$ schneiden. Da $\mathcal{G}_{v'}$ parallel zu $\mathcal{G}_v := \{a_0 + \lambda v \mid \lambda \in \mathbb{R}\}$ ist, ist nach (10.22) auch $F(\mathcal{G}'_v)$ parallel zu $F(\mathcal{G}_v)$. Wegen $F(a_0 + w) = a'_0 + G(w) \in F(\mathcal{G}'_v)$ und $a'_0 \in F(\mathcal{G}_v)$, $F(a_0 + v) = a'_0 + G(v) \in F(\mathcal{G}_v)$ gilt

$$F(\mathcal{G}'_v) = \{a'_0 + G(w) + \lambda G(v) \mid \lambda \in \mathbb{R}\}$$

und analog

$$F(\mathcal{G}'_w) = \{a'_0 + G(v) + \mu G(w) \mid \mu \in \mathbb{R}\}.$$

Da F injektiv ist und $\mathcal{G}_v$ und $\mathcal{G}_w := \{a_0 + \mu w \mid \mu \in \mathbb{R}\}$ einander genau in a_0 schneiden, gilt $F(\mathcal{G}_v) \cap F(\mathcal{G}_w) = \{a'_0\}$. Deshalb sind $G(v) = \overrightarrow{F(a_0)F(a_0 + v)}$ und $G(w) = \overrightarrow{F(a_0)F(a_0 + w)}$ linear unabhängig. Da $G(v)$ bzw. $G(w)$ Richtungsvektoren von $F(\mathcal{G}'_v)$ bzw. $F(\mathcal{G}'_w)$ sind, schneiden $F(\mathcal{G}'_v)$ und $F(\mathcal{G}'_w)$ einander genau im Punkt $F(a_0 + v + w) = a'_0 + G(v + w)$, dem Bild des Schnittpunkts von $\mathcal{G}'_v$ und $\mathcal{G}'_w$. Andererseits zeigen die obigen Formeln für $F(\mathcal{G}'_v)$ und $F(\mathcal{G}'_w)$, daß auch $a'_0 + G(v) + G(w) \in F(\mathcal{G}'_v) \cap F(\mathcal{G}'_w)$ gilt. Daraus folgt

$$G(v + w) = G(v) + G(w).$$

(10.24) Lemma. Für alle $v \in V$ und alle $\lambda, \mu \in \mathbb{R}$ gilt:

$$G(\lambda v) + G(\mu v) = G((\lambda + \mu)v).$$

Bem.: Aus (10.23) und (10.24) folgt $G(v + w) = G(v) + G(w)$ für alle $v, w \in V$.

Bew.: Wegen $G(0) = 0$ genügt es, die Fälle zu betrachten, in denen $v \neq 0$, $\lambda \neq 0$ und $\mu \neq 0$ gilt. Aufgrund unserer Voraussetzung $\dim \mathcal{A} \geq 2$ existiert ein zu v linear unabhängiger Vektor $w \in V$.

1. Fall: $\lambda + \mu \neq 0$. Dann sind w und $(\lambda + \mu)v$ linear unabhängig und (10.23) impliziert

$$G(w + (\lambda + \mu)v) = G(w) + G((\lambda + \mu)v).$$

Da $w + \lambda v$ und $w + \lambda v + \mu v$ linear unabhängig sind, folgt aus (10.23)

$$G(w + (\lambda + \mu)v) = G(w + \lambda v) + G(\mu v)$$

und analog

$$G(w + \lambda v) = G(w) + G(\lambda v).$$

Die letzten drei Gleichungen implizieren

$$G((\lambda + \mu)v) = G(\lambda v) + G(\mu v).$$

2. Fall: $\lambda = -\mu$. Da $\lambda v + w$ und $-\lambda v + w = \mu v + w$ linear unabhängig sind, folgt aus (10.23)

$$\begin{aligned} G(2w) = G((\lambda v + w) + (-\lambda v + w)) &= G(\lambda v + w) + G(-\lambda v + w) \\ &= G(\lambda v) + G(w) + G(-\lambda v) + G(w) \\ &= 2G(w) + G(\lambda v) + G(-\lambda v). \end{aligned}$$

Da nach dem 1. Fall $G(2w) = 2G(w)$ gilt ($\lambda = \mu := 1$), folgt daraus die Behauptung: $G(\lambda v) + G(-\lambda v) = 0 = G(\lambda v - \lambda v)$.

(10.25) Lemma. Es existiert eine Funktion $f : \mathbb{R} \to \mathbb{R}$, so daß für alle $v \in V$ und alle $\lambda \in \mathbb{R}$ gilt:

$$G(\lambda v) = f(\lambda)G(v).$$

Bew.: Wegen $G(0) = 0$ ist die Gleichung für $v = 0$ stets erfüllt und sie ist für $\lambda = 0$ und alle $v \in V$ erfüllt, wenn wir $f(0) = 0$ setzen. Wir setzen von jetzt ab $v \neq 0$ und $\lambda \neq 0$ voraus. Da F injektiv ist, gilt $F(a_0 + v) = a_0' + G(v) \neq a_0'$. Nach Voraussetzung liegt $F(a_0 + \lambda v)$ für alle $\lambda \in \mathbb{R}$ auf der Geraden $\mathcal{G}_{a_0', a_0' + G(v)}$. Also existiert genau ein $f(\lambda, v) \in \mathbb{R}$, so daß

$$F(a_0 + \lambda v) = a_0' + f(\lambda, v)G(v)$$

gilt. Wir bemerken, daß $TV(a_0', F(a_0 + \lambda v), F(a_0 + v)) = \frac{f(\lambda,v)}{1-f(\lambda,v)}$ gilt. Wir wollen zeigen, daß $f(\lambda, v)$ unabhängig von v ist, d.h. daß für alle $\lambda \neq 0$, $v \neq 0$, $w \neq 0$ gilt:

$$f(\lambda, v) = f(\lambda, w).$$

1. Fall: v und w sind linear unabhängig. Dann betrachten wir die Geraden $\mathcal{H} = \{a_0 + \mu v \mid \mu \in \mathbb{R}\}$ und $\mathcal{J} = \{a_0 + \mu w \mid \mu \in \mathbb{R}\}$ durch a_0 und die Geraden $\mathcal{G}_1 = \mathcal{G}_{a_0+v, a_0+w}, \mathcal{G}_2 = \mathcal{G}_{a_0+\lambda v, a_0+\lambda w}$, die nach dem Strahlensatz (10.10) parallel sind. Nach (10.21) sind dann auch $F(\mathcal{G}_1)$ und $F(\mathcal{G}_2)$ parallel und die andere Richtung des Strahlensatzes (10.10) impliziert nun, daß

$$TV(a_0', F(a_0 + \lambda v), F(a_0 + v)) = TV(a_0', F(a_0 + \lambda w), F(a_0 + w))$$

gilt. Daraus folgt $f(\lambda, v) = f(\lambda, w)$.

2. Fall: Ist $w = \mu v$, so wählen wir ($\dim \mathcal{A} \geq 2$!) ein von v linear unabhängiges $z \in V$. Dann gilt mit zweifacher Anwendung des 1. Falls:

$$f(\lambda, v) = f(\lambda, z) = f(\lambda, \mu v) = f(\lambda, w)$$

(10.26) Folgerung. $f : \mathbb{R} \to \mathbb{R}$ ist Körperautomorphismus, d.h. es gilt $f(1) = 1$ und für alle $\lambda, \mu \in \mathbb{R}$:

$$\begin{aligned} f(\lambda + \mu) &= f(\lambda) + f(\mu) \\ f(\lambda\mu) &= f(\lambda)f(\mu). \end{aligned}$$

Bew.: Wähle $v \in V \setminus \{0\}$. Da F injektiv ist, folgt $G(v) = \overrightarrow{F(a_0)F(a_0 + v)} \neq 0$.

(i) $G(v) = G(1 \cdot v) = f(1)G(v) \Rightarrow (f(1) - 1)G(v) = 0 \Rightarrow f(1) = 1$

(ii) $G((\lambda + \mu)v) = f(\lambda + \mu)G(v)$
$G((\lambda + \mu)v) = G(\lambda v + \mu v) \overset{(10.24)}{=} G(\lambda v) + G(\mu v) = (f(\lambda) + f(\mu))G(v)$
Diese Gleichungen beweisen $f(\lambda + \mu) = f(\lambda) + f(\mu)$.

(iii) $G((\lambda\mu)v) = f(\lambda\mu)G(v)$
$G((\lambda\mu)v) = G(\lambda(\mu v)) = f(\lambda)G(\mu v) = f(\lambda)f(\mu)G(v)$.
Diese Gleichungen beweisen $f(\lambda\mu) = f(\lambda)f(\mu)$.

Bisher wurde noch nicht benutzt, daß $\mathcal{A}, \mathcal{A}'$ reelle affine Räume sind. Das folgende Lemma aber wäre etwa für den Körper $\mathbb{C}$ (statt $\mathbb{R}$) nicht wahr.

(10.27) Lemma. Die Identität $f = \mathrm{id}_{\mathbb{R}}$ ist der einzige Körperautomorphismus von $\mathbb{R}$.

Bew.: Sei $f : \mathbb{R} \to \mathbb{R}$ Körperautomorphismus. Die Additivität von f und $f(1) = 1$ zeigen, daß $f(2) = f(1+1) = f(1) + f(1) = 1 + 1 = 2$ und – induktiv – daß $f(n) = n$ für alle $n \in \mathbb{N}_{>0}$ gilt. Aus $f(0) + f(0) = f(0+0) = f(0)$ folgt $f(0) = 0$ und aus $f(n + (-n)) = 0 = f(n) + f(-n) = n + f(-n)$ für $n \in \mathbb{N}$ folgt $f(n) = n$ für alle $n \in \mathbb{Z}$. Ist $n \in \mathbb{Z} \setminus \{0\}$, so gilt

$$1 = f(1) = f\left(\frac{1}{n} \cdot n\right) = f\left(\frac{1}{n}\right) f(n) = f\left(\frac{1}{n}\right) \cdot n,$$

also $f\left(\frac{1}{n}\right) = \frac{1}{n}$. Verwenden wir nochmals die Multiplikativität von f, so erhalten wir $f\left(\frac{p}{q}\right) = \frac{p}{q}$ für alle $\frac{p}{q} \in \mathbb{Q}$. Weiter gilt $f(\mathbb{R}_{\geq 0}) \subseteq \mathbb{R}_{\geq 0}$, denn aus $r \geq 0$ folgt

$$f(r) = f\left(\sqrt{r})^2\right) = \left(f(\sqrt{r})\right)^2 \geq 0.$$

Zusammen mit der Additivität von f ergibt das

$$r \geq s \Rightarrow f(r) \geq f(s),$$

d.h. f ist monoton wachsend. Ist schließlich $r \in \mathbb{R}$ beliebig, so wählen wir $x_i, y_i \in \mathbb{Q}$ mit $x_i \leq r \leq y_i$ und $\lim x_i = r = \lim y_i$. Dann folgt $x_i = f(x_i) \leq f(r) \leq f(y_i) = y_i$ und daraus

$$f(r) = \lim x_i = \lim y_i = r.$$

Aus (10.23)–(10.27) folgt, daß die durch

$$F(a_0 + v) = a_0' + G(v)$$

definierte Abbildung $G : V \to V'$ ein $\mathbb{R}$-Vektorraumhomomorphismus ist. Das zeigt, daß F eine affine Abbildung ist, vgl. (10.12), 3). Da F als bijektiv vorausgesetzt wurde, ist F eine Affinität. Das beendet den Beweis des Fundamentalsatzes (10.20).

Weiteres zur affinen Geometrie findet man z.B. in den Büchern G. Fischer: Analytische Geometrie, Vieweg 1978, W. Klingenberg: Lineare Algebra und Geometrie, Springer 1990.

Nach dieser kurzen Einführung in die affine Geometrie wenden wir uns noch kürzer der euklidischen Geometrie zu. Die euklidische Geometrie ist für uns besonders anschaulich, weil sie (in Dimension 3) gerade die mathematische Präzisierung des uns umgebenden „Anschauungsraums“ ist.

(10.28) Def.: Ein euklidischer Raum $\mathcal{E} = \mathcal{E}(V, <,>)$ ist ein affiner Raum $\mathcal{E}$ über einem euklidischen (reellen!) Vektorraum $V, <,>$. Eine Isometrie von $\mathcal{E} = \mathcal{E}(V, <,>)$ auf $\mathcal{E}' = \mathcal{E}'(V', <,>')$ ist eine Affinität $F : \mathcal{E} \to \mathcal{E}'$, deren zugehörige lineare Abbildung $L_F : V \to V'$ orthogonal bzgl. $<,>$ und $<,>'$ ist, vgl. (6.8).

Bez.: $\mathrm{Iso}(\mathcal{E}) = \{F : \mathcal{E} \to \mathcal{E} \mid F \text{ Isometrie}\}$ ist bzgl. $\circ$ eine Gruppe, die „Isometriegruppe von $\mathcal{E}$“. Ein $F \in \mathrm{Iso}(\mathcal{E})$ heißt orientierungserhaltende Isometrie (oder eigentliche Bewegung), falls $L_F \in \mathrm{SO}(V)$ gilt.

Bem.: Zwei endlich-dimensionale euklidische Räume sind genau dann isometrisch, wenn sie die gleiche Dimension haben, vgl. (6.9). Das Standardbeispiel eines euklidischen Raums ist der $\mathbb{R}^n$ (nach (10.2) als affiner Raum aufgefaßt) mit dem Standardskalarprodukt $\langle v, w\rangle = \sum_{i=1}^{n} v_i w_i$.

Zwei Punkten a, b eines euklidischen Raums $\mathcal{E}$ ordnet man den Abstand $d(a, b)$ von a und b durch

$$d(a,b) := \|\overrightarrow{ab}\| = \langle \overrightarrow{ab}, \overrightarrow{ab}\rangle^{\frac{1}{2}}$$

zu. Aus Satz (6.3) folgen leicht die folgenden Eigenschaften des Abstands d:

(i) $d(a,b) \geq 0$ und $d(a,b) = 0 \Leftrightarrow a = b$.
(ii) $d(a,b) = d(b,a)$.
(iii) $d(a,c) \leq d(a,b) + d(b,c)$ „Dreiecksgleichung“.

Isometrien $F : \mathcal{E} \to \mathcal{E}'$ lassen Abstände unverändert: Für alle $a, b \in \mathcal{E}$ gilt:

$$d'(F(a), F(b)) = \|\overrightarrow{F(a)F(b)}\|' = \|L_F(\overrightarrow{ab})\|' = \|\overrightarrow{ab}\| = d(a,b).$$

Sind $\mathcal{G}$ und $\mathcal{G}'$ orientierte Geraden in $\mathcal{E}$, die einander schneiden, so definiert man den Schnittwinkel $\varphi \in [0, \pi]$ von $\mathcal{G}$ und $\mathcal{G}'$ wie folgt: Man wählt positiv orientierte Richtungsvektoren v von $\mathcal{G}$ und v' von $\mathcal{G}'$ und setzt $\varphi = \sphericalangle(v, v') = \arccos \frac{\langle v, v'\rangle}{\|v\|'}$.

Ein affines Koordinatensystem $(a_1, \ldots, a_n)$ eines n-dimensionalen euklidischen Raums $\mathcal{E}$ heißt <u>kartesisch</u>, falls $\overrightarrow{a_0a_1}, \ldots, \overrightarrow{a_0a_n}$ eine ONB von V ist. Die Abbildung, die jedem Punkt $x \in \mathcal{E}$ seinen Koordinatenvektor $(x_1, \ldots, x_n) \in \mathbb{R}$ bzgl. eines festen kartesischen Koordinatensystems zuordnet, ist eine Isometrie von $\mathcal{E}$ auf das Standardbeispiel $\mathbb{R}^n$ eines euklidischen Raums. Speziell gilt dann

$$d(x,y) = \sum_{i=1}^{n} (y_i - x_i)^2.$$

Man kann nun die Aussagen der euklidischen Geometrie, z.B. die der Schulgeometrie, mittels Vektorrechnung im euklidischen $\mathbb{R}^2$ oder $\mathbb{R}^3$ beweisen. Dazu ein Beispiel:

(10.29) Satz. Die Höhen eines Dreiecks in einer euklidischen Ebene schneiden sich in einem Punkt.

Bew.: Seien A, B, C die Eckpunkte des Dreiecks und S der Schnittpunkt der Höhen durch die Eckpunkte A und B. Wir wollen zeigen, daß die Gerade $\mathcal{G}_{C,S}$ die Höhe durch die Ecke C ist, d.h. daß $\mathcal{G}_{C,S}$ die Gerade $\mathcal{G}_{A,B}$ senkrecht schneidet, d.h. daß

$$\langle \overrightarrow{AB}, \overrightarrow{CS}\rangle = 0$$

gilt.
Nun gilt, da $\mathcal{G}_{A,B}$ senkrecht auf $\mathcal{G}_{C,B}$ steht,

$$0 = \langle \overrightarrow{AS}, \overrightarrow{CB}\rangle = \langle \overrightarrow{AS}, \overrightarrow{CS} + \overrightarrow{SB}\rangle = \langle \overrightarrow{AS}, \overrightarrow{CS}\rangle + \langle \overrightarrow{AS}, \overrightarrow{SB}\rangle$$

und analog
$$0 = \langle \overrightarrow{BS}, \overrightarrow{CA} \rangle = \langle \overrightarrow{BS}, \overrightarrow{CS} + \overrightarrow{SA} \rangle = \langle \overrightarrow{BS}, \overrightarrow{CS} \rangle + \langle \overrightarrow{BS}, \overrightarrow{SA} \rangle.$$
Bildet man die Differenz dieser Gleichungen, so folgt
$$0 = \langle \overrightarrow{AS}, \overrightarrow{CS} \rangle - \langle \overrightarrow{BS}, \overrightarrow{CS} \rangle = \langle \overrightarrow{AS} + \overrightarrow{SB}, \overrightarrow{CS} \rangle = \langle \overrightarrow{AB}, \overrightarrow{CS} \rangle$$
wie behauptet.
Das folgende euklidische Analogon zum Fundamentalsatz der affinen Geometrie ist leichter zu beweisen.

(10.30) Satz. *Seien $\mathcal{E} = \mathcal{E}(V, \langle , \rangle)$ und $\mathcal{E}' = \mathcal{E}'(V', \langle , \rangle')$ euklidische Räume und $F : \mathcal{E} \to \mathcal{E}'$ eine surjektive, abstandserhaltende Abbildung. Dann ist F eine Isometrie von $\mathcal{E}$ auf $\mathcal{E}'$.*

Bem.: Daß F abstandserhaltend ist, bedeutet, daß für alle $a, b \in \mathcal{E}$ gilt:
$$d'(F(a), F(b)) = d(a, b).$$
Bew.: Da F abstandserhaltend ist, ist F injektiv. Da F nach Voraussetzung surjektiv ist, ist F sogar bijektiv. Wir wählen $a_0 \in \mathcal{E}$, setzen $a_0' := F(a_0)$ und definieren $G : V \to V'$ durch
$$F(a_0 + v) = a_0' + G(v).$$
Wir müssen zeigen, daß $G \in \mathrm{Hom}(V, V')$ist ($\Rightarrow F$ ist Affinität und $G = L_F$), und daß für alle $v, w \in V$ gilt:

$$(*) \qquad \langle v, w \rangle = \langle G(v), G(w) \rangle'.$$

Zunächst gilt:
$$\begin{aligned} d(a_0 + v, a_0 + w) = \|w - v\| &= d'(F(a_0 + v), F(a_0 + w)) \\ &= d'(a_0' + G(v), a_0' + G(w)) = \|G(w) - G(v)\|'. \end{aligned}$$
Wegen $G(0) = 0$ folgt daraus speziell
$$\|G(v)\|' = \|v\| \quad \text{für alle } v \in V.$$
Damit folgt aus
$$\begin{array}{rcllc} \|w - v\|^2 & = & \|w\|^2 & -2\langle w, v \rangle & +\|v\|^2 = \\ \|G(w) - G(v)\|'^2 & = & \|G(w)\|'^2 & -2\langle G(w), G(v) \rangle' & +\|G(v)\|'^2 \end{array}$$
daß $(*)$ gilt. Es bleibt $G \in \mathrm{Hom}(V, V')$ zu zeigen. Nun gilt für alle $v, w \in V$:

$$(1) \qquad \begin{aligned} \|G(v + w)\|'^2 &= \|v + w\|^2 = \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\ &= \|G(v)\|'^2 + 2\langle G(v), G(w) \rangle' + \|G(w)\|'^2 = \|G(v) + G(w)\|'^2 \end{aligned}$$

und

$$(2) \qquad \begin{aligned} \langle G(v + w), G(v) + G(w) \rangle' = \langle v + w, v \rangle + \langle v + w, w \rangle &= \|v + w\|^2 \\ &= \|G(v + w)\|'^2. \end{aligned}$$

Mit Hilfe von (1) und (2) folgt

$$\begin{aligned}&\|G(v+w)-(G(v)+G(w))\|'^2\\&\qquad=\|G(v,w)\|'^2-2\langle G(v+w),G(v)+G(w)\rangle'+\|G(v)+G(w)\|'^2=0,\end{aligned}$$

also $G(v+w)=G(v)+G(w)$.

Schließlich gilt für alle $\alpha\in\mathbb{R}$, $v\in V$:

$$\|G(\alpha v)\|'^2=\|\alpha v\|^2=\alpha^2\|G(v)\|'^2$$

und

$$\langle G(\alpha v),\alpha G(v)\rangle'=\alpha\langle G(\alpha v),G(v)\rangle'=\alpha^2\|v\|^2=\alpha^2\|G(v)\|'^2.$$

Die vorangehenden Gleichungen implizieren

$$\|G(\alpha v)-\alpha G(v)\|'^2=\|G(\alpha v)\|'^2-2\alpha\langle G(\alpha v),G(v)\rangle'+\alpha^2\|G(v)\|'^2=0,$$

also $G(\alpha v)=\alpha G(v)$.

Jede Isometrie F eines euklidischen Raums $\mathcal{E}=\mathcal{E}(V,\langle,\rangle)$ läßt sich bezüglich eines beliebigen Punktes $a_0\in\mathcal{E}$ in der Form

$$F(a_0+v)=F(a_0)+L_F(v)$$

mit $L_F\in O(V)$ darstellen. Untersucht man F nach Fixpunkten, invarianten Geraden u.s.w., so kann man Darstellungen von F finden, die mehr über F aussagen. Wir begnügen uns mit dem einfachsten Fall.

(10.31) Satz. Sei $\mathcal{E}=\mathcal{E}(V,\langle,\rangle)$ eine euklidische Ebene und $F\in\mathrm{Iso}(\mathcal{E})$ orientierungserhaltend. Dann ist F entweder eine Translation ($\Leftrightarrow L_F=\mathrm{id}_V$) oder eine Drehung um einen Punkt $a\in\mathcal{E}$ ($\Leftrightarrow F(a+v)=a+D(v)$) mit $D\in\mathrm{SO}(V)\setminus\{\mathrm{id}_V\}$).

Bew.: Ist F keine Translation, so gilt $L_F=:D\in\mathrm{SO}(V)\setminus\{\mathrm{id}_V\}$. Wir suchen einen Fixpunkt $a\in\mathcal{E}$ von F, denn dann gilt $F(a+v)=a+D(v)$ für alle $v\in V$. Sei $a_0\in\mathcal{E}$ beliebig. Wir suchen $w\in V$, so daß

$$(*)\qquad F(a_0+w)=F(a_0)+D(w)=a_0+w$$

gilt. Das ist äquivalent zu

$$w-D(w)=\overrightarrow{a_0F(a_0)}=:z.$$

Da $\det(\mathrm{id}_V-D)=\begin{vmatrix}1-\cos\varphi & \sin\varphi\\ -\sin\varphi & 1-\cos\varphi\end{vmatrix}\neq 0$ gilt (vgl. den Beweis von (6.24)'), ist $\mathrm{id}_V-D:V\to V$ surjektiv. Also existiert $w\in V$ mit $w-D(w)=z$, d.h. $(*)$ gilt für dieses w. Damit ist $a:=a_0+w\in\mathcal{E}$ Fixpunkt von F.