Skript zur Vorlesung

Lineare Algebra für Informatiker und Ingenieure

Wintersemester 2014/ 15

Prof. Dr. Helmut Maier
Dr. Hans- Peter Reck

Institut für Zahlentheorie und Wahrscheinlichkeitstheorie
Universität Ulm

Inhaltsverzeichnis

Kapitel 1

Einleitung

Vorbemerkung

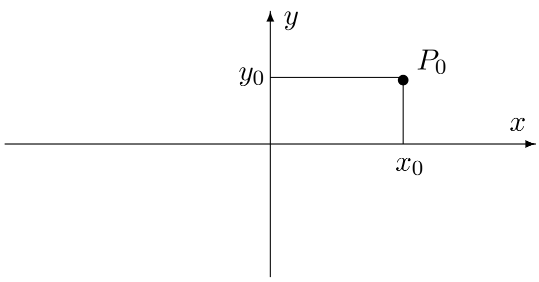
Zentral in der linearen Algebra ist der Begriff des Vektorraums: Dieser Begriff hat sich aus der analytischen Geometrie der Ebene und des Raums entwickelt, umfasst jedoch wesentlich allgemeinere Strukturen. Jedoch wird auch das Verständnis der allgemeineren Strukturen durch die Konzepte, die auf der geometrischen Anschauung in der Ebene und im Raum beruhen, erleichtert.

1.1 Die Ebene

Wir setzen die Menge $\mathbb{R}$ der reellen Zahlen mit den Rechenoperationen Addition, Subtraktion, Multiplikation und Division als bekannt voraus. Unter $\mathbb{R}^2$ verstehen wir die Menge aller Paare reeller Zahlen:

$$\mathbb{R}^2 = \{(x, y) \colon x, y \in \mathbb{R}\}.$$

Die Ebene E, die wir uns zum Beispiel als Zeichenebene vorstellen, kann durch Einführung eines kartesischen Koordinatensystems mit dem $\mathbb{R}^2$ identifiziert werden. Ein kartesisches Koordinatensystem entsteht durch Vorgabe eines Punktes 0 und einer Zahlengeraden, die wir x- Achse nennen, mit dem Nullpunkt 0. Die y- Achse entsteht durch eine positive Drehung gegen den Uhrzeigersinn um $90°$ um den Punkt 0 aus der x- Achse. Fällt man für einen beliebigen Punkt $P_0 \in E$ die Lote auf die Achsen, so bestimmen die beiden Fußpunkte die x- bzw. y- Koordinate x_0 bzw. y_0 von P_0, und man schreibt $P_0 = (x_0, y_0)$.



Der Punkt $0 = (0, 0)$ heißt Nullpunkt oder Ursprung des Koordinatensystems. Nach der Festlegung eines kartesischen Koordinatensystems gibt es zu jedem Zahlenpaar $(x, y) \in \mathbb{R}^2$ genau einen Punkt $X \in E$ mit $X = (x, y)$, und umgekehrt. Zu je zwei Punkten P und Q der Ebene gibt es genau eine Parallelverschiebung (der Ebene), die P nach Q bringt. Diese Verschiebung wird mit $\overrightarrow{PQ}$ bezeichnet, und heißt "Vektor von P nach Q". Der Vektor $\vec{v} = \overrightarrow{PQ}$ wird durch einen Pfeil, der von P nach Q zeigt, dargestellt.

Wird unter $\overrightarrow{PQ}$ ein anderer Punkt R nach S verschoben, dann hat offenbar $\overrightarrow{RS}$ die gleiche Wirkung wie $\overrightarrow{PQ}$, das heißt $\overrightarrow{PQ} = \overrightarrow{RS}$. Zwei gleich lange und gleichgerichtete Pfeile im Raum stellen somit den selben Vektor dar.
Offenbar gibt es zu einem Vektor $\overrightarrow{PQ}$ genau einen Punkt S, so dass $\overrightarrow{PQ} = \overrightarrow{0S}$ ist. Wir können den Vektor $\overrightarrow{PQ}$ dann mit dem Punkt S der Ebene identifizieren. Jeder Punkt der Ebene kann also ein Vektor gedeutet werden und umgekehrt. Insbesondere kann jeder Vektor der Ebene durch ein Zahlenpaar beschrieben werden. Wir schreiben dieses Paar als Spalte

$$\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Zahlen $x, y \in \mathbb{R}$ heißen die Komponenten von $\vec{v}$.

Die Addition von Vektoren

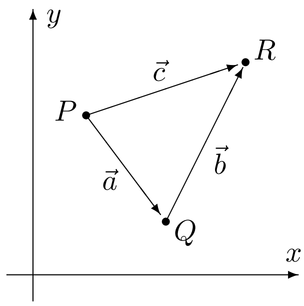
Den zu $\vec{v}$ gleichgerichteten, aber entgegengesetzten Vektor bezeichnen wir mit

$$-\vec{v} = \begin{pmatrix} -x \\ -y \end{pmatrix}.$$

Insbesondere ist $-\overrightarrow{PQ} = \overrightarrow{QP}$. Der Nullvektor ist

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

mit $\vec{0} = \overrightarrow{PP}$ für alle Punkte P. Führt man zwei Parallelverschiebungen hintereinander aus, etwa zuFerst $\vec{a} = \overrightarrow{PQ}$ und dann $\vec{b} = \overrightarrow{QR}$, so ergibt sich wieder eine Parallelverschiebung, nämlich $\vec{c} = \overrightarrow{PR}$. Wir nennen $\vec{c}$ die Summe von $\vec{a}$ und $\vec{b}$ und schreiben $\vec{c} = \vec{a} + \vec{b}$.



Sind $\vec{a}$ und $\vec{b}$ durch ihre Komponenten gegeben, so kann die Summe durch Addition der Komponenten erhalten werden:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Rightarrow \vec{a} + \vec{b} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}.$$

Offenbar gelten für beliebige Vektoren $\vec{a}, \vec{b}, \vec{c}$ die folgenden Rechenregeln:

$$\begin{aligned} \vec{a} + \vec{0} &= \vec{a}, \quad \vec{a} + (-\vec{a}) = \vec{0} \\ \vec{a} + \vec{b} &= \vec{b} + \vec{a} \qquad \text{(Kommutativgesetz)} \qquad\qquad (VA) \\ \vec{a} + (\vec{b} + \vec{c}) &= (\vec{a} + \vec{b}) + \vec{c} \qquad \text{(Assoziativgesetz).} \end{aligned}$$

Die skalaren Vielfachen eines Vektors

Zu einer reellen Zahl $\lambda \geq 0$ und einem Vektor $\vec{a}$ bezeichne $\lambda\vec{a}$ denjenigen Vektor, der dieselbe Richtung wie $\vec{a}$ besitzt, aber die λ-fache Länge. Im Fall $\lambda < 0$ setzt man $\lambda\vec{a} := -(|\lambda|\vec{a})$.
Sonderfälle dieser Definition sind $0\vec{a} = \vec{0}$ und $\lambda\vec{0} = \vec{0}$ für jede Zahl $\lambda \in \mathbb{R}$ und jeden Vektor $\vec{a}$.
Für diese Multiplikation von Vektoren mit Zahlen (Skalarmultiplikation) gelten mit mit $\lambda, \mu \in \mathbb{R}$ und Vektoren $\vec{a}, \vec{b}$ folgende Rechenregeln:

$$\begin{aligned} \lambda(\mu\vec{a}) &= (\lambda\mu)\vec{a} \\ \lambda(\vec{a}+\vec{b}) &= \lambda\vec{a}+\lambda\vec{b} \qquad\qquad (VM) \\ (\lambda+\mu)\vec{a} &= \lambda\vec{a}+\mu\vec{a}. \end{aligned}$$

Geraden

Ein Punkt X liegt genau dann auf der Geraden g durch A in Richtung $\vec{c}$, sofern $\vec{c} \neq \vec{0}$ gilt, wenn $\overrightarrow{AX}$ zu $\vec{c}$ parallel ist, falls es eine also Zahl $t \in \mathbb{R}$ mit $\overrightarrow{AX} = t\vec{c}$ gibt. Man sagt, dass g die Punkt- Richtungsgleichung

$$\overrightarrow{AX} = t\vec{c} \tag{$*$}$$

mit $t \in \mathbb{R}$ besitzt. Die in $(*)$ auftretende Variable t nennt man einen Parameter.
Zu jedem Parameter $t = t_0$ gehört genau ein Punkt X_0 auf der Geraden g mit $\overrightarrow{AX_0} = t_0\vec{c}$ und umgekehrt. Wegen $\overrightarrow{AX} = \overrightarrow{PX} - \overrightarrow{PA}$ lässt sich g bezüglich eines beliebigen Punktes P durch

$$\overrightarrow{PX} = \overrightarrow{PA} + t\vec{c}$$

mit $t \in \mathbb{R}$ darstellen. Ist $A = (a_1, a_2)$, $B = (b_1, b_2)$ und $C = (c_1, c_2)$, so ergibt ein Komponentenvergleich für die Geradenpunkte $X = (x, y)$ die zwei Gleichungen

$$\begin{cases} x = a_1 + tc_1 \\ y = a_2 + tc_2 \end{cases} \quad \text{(Punkt- Richtungs- Gleichung)}$$

$$\begin{cases} x = a_1 + t \cdot (b_1 - a_1) \\ y = a_2 + t \cdot (b_2 - a_2) \end{cases} \quad \text{(Zwei- Punkte- Gleichung)}$$

Löst man die zwei unteren Gleichungen nach t auf, und setzt die Ausdrücke gleich, so erhält man mit $A = (a_1, a_2)$ und $B = (b_1, b_2)$ eine parameterfreie Darstellung.

Koordinatengleichung der Geraden durch A und B

Es ist

- $\dfrac{x - a_1}{b_1 - a_1} = \dfrac{y - a_2}{b_2 - a_2}$, falls $a_i \neq b_i$ für $i = 1, 2$
- $x = a_1$, falls $a_1 = b_1$
- $y = a_2$, falls $a_2 = b_2$.

Aus der (parameterfreien) Zwei- Punkte- Form für g findet man über

$$t = \frac{x - a_1}{b_1 - a_1} \quad \text{bzw.} \quad t = \frac{y - a_2}{b_2 - a_2}$$

zur Parameterform zurück.

Beispiel 1.1.1. Die durch die Parametergleichung

$$g: \begin{cases} x = 3 - 2t \\ y = 4 + 5t \end{cases}$$

bestimmte Gerade in der Ebene hat die Koordinatengleichung

$$\frac{3-x}{2} = \frac{y-4}{5}.$$

Beispiel 1.1.2. Man finde die Parameterdarstellung der durch die Gleichung $2x+3y = 5$ gegebenen Geraden. Die Rechnung ist

$$\begin{array}{rcl} 2x + 3y & = & 5 \\ 2x - 5 & = & -3y \\ \dfrac{x - \frac{5}{2}}{\frac{1}{2}} & = & \dfrac{y}{-\frac{1}{3}} \\ x - \frac{5}{2} & = & \frac{1}{2}t \\ y & = & -\frac{1}{3}t \quad \Rightarrow \quad \begin{cases} x = \frac{5}{2} + \frac{1}{2}t \\ y = -\frac{1}{3}t \end{cases} \end{array}$$

Schnittpunkt zweier Geraden

Die Bestimmung des Schnittpunkts zweier Geraden führt auf die Lösung eines linearen Gleichungssystems, und zwar in jedem Fall: ob die Gerade nun durch die Punkt- Richtungs- Gleichung oder durch die Zwei- Punkte- Gleichung gegeben ist. Auch lineare Gleichungssysteme stehen im Zentrum der Linearen Algebra. Wir werden sie in später systematisch behandeln.

Beispiel 1.1.3. Es seien zwei Geraden durch Punkt- Richtungs- Gleichungen gegeben:

$$g_1: \begin{cases} x = 3 + 5t \\ y = 2 - t \end{cases} \qquad \text{und} \quad g_2: \begin{cases} x = 4 - 2u \\ y = 1 + 3u \end{cases}$$

Es ist wichtig, zwei verschiedene Variablen für die Parameter zu benützen. Gleichsetzen liefert

$$\begin{array}{rcl} 3 + 5t & = & 4 - 2u \\ 2 - t & = & 1 + 3u \end{array}$$

oder

$$\begin{array}{rcl} 5t + 2u & = & 1 \\ -t - 3u & = & -1. \end{array}$$

Addition des 5-fachen der 2. Zeile zur 1. Zeile ergibt

$$\begin{array}{rcl} -13u & = & -4 \\ -t - 3u & = & -1 \\ u & = & \dfrac{4}{13}. \end{array}$$

Einsetzen in das System von g_2 ergibt $x = \frac{44}{13}$ und $y = \frac{25}{13}$, also erhalten wir den Schnittpunkt $\left(\frac{44}{13}, \frac{25}{13}\right)$.

Sind die beiden Geraden durch Koordinatengleichungen gegeben, so erhält man das Gleichungssystem unmittelbar:

Beispiel 1.1.4. Es seien $g_1 : x + 3y = 5$ und $g_2 : 3x - 2y = 7$ gegeben, dann ist

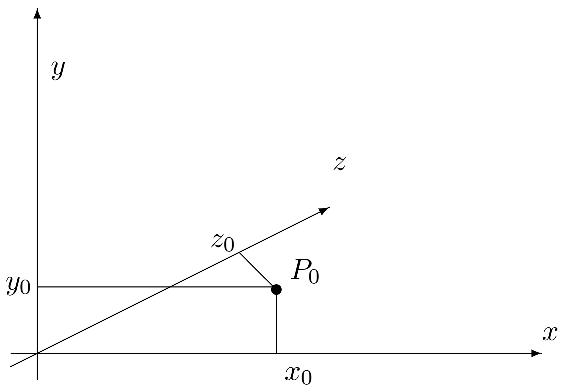
$$\begin{cases} x + 3y = 5 \\ 3x - 2y = 7 \end{cases} \Rightarrow \begin{cases} x + 3y = 5 \\ -11y = -8 \end{cases}$$

nach Subtraktion des 3-fachen der ersten Zeile von der zweiten Zeile. Dies führt zu $y = \frac{8}{11}$ und $x = \frac{31}{11}$, also zu dem Schnittpunkt $P_0 = \left(\frac{31}{11}, \frac{8}{11}\right)$.

Die Gleichungssysteme haben keine bzw. unendlich viele Lösungen, falls es sich um parallele bzw. identische Geraden handelt.

1.2 Der Raum

Es sei nun $\mathbb{R}^3 = \{(x, y, z) \colon x, y, z \in \mathbb{R}\}$.
Ähnlich wie die Ebene mit dem $\mathbb{R}^2$ kann der Raum nun mit dem $\mathbb{R}^3$ identifiziert werden. Ein kartesisches Koordinatensystem im Raum besteht aus dem Nullpunkt 0 und drei sich in 0 schneidenden Zahlengeraden gleicher Längeneinheit. Man bezeichnet sie als x, y und z-Achse derart, dass diese drei Achsen ein Rechtssystem bilden, das heißt, die Drehung der positiven x-Achse um 90° in die positive y-Achse, zusammen mit einer Verschiebung in Richtung der positiven z-Achse, muss eine Rechtsschraube darstellen. Diese drei durch je zwei Achsen bestimmten Ebenen heißen Koordinatenebenen, bzw. (x, y)-Ebene, (y, z)-Ebene und (z, x)-Ebene. Die Koordinaten x_0, y_0 und z_0 eines Punktes P_0 gewinnt man aus den Schnittpunkten der entsprechenden Achsen mit dem zu den Koordinatenebenen parallelen Ebenen durch P_0. Man schreibt $P_0 = (x_0, y_0, z_0)$.



Völlig analog zum Fall der Ebene werden nun auch im Raum Vektoren als Parallelverschiebungen des Raumes definiert. Auch die Pfeildarstellung sowie die Operationen der Addition und der Skalarmultiplikation verlaufen völlig analog zum Fall der Ebene. Der einzige Unterschied liegt in der Tatsache, dass Vektoren im Raum drei Komponenten besitzen. Wie in der Ebene besitzt eine Gerade im Raum eine Parameterdarstellung

$$\overrightarrow{OX} = \overrightarrow{OA} + t\vec{x}$$

mit $t \in \mathbb{R}$. Falls $A = (a_1, a_2, a_3)$ und

$$\vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

gilt,

so ergibt ein Komponentenvergleich die drei Gleichungen

$$\begin{cases} x = a_1 + tc_1 \\ y = a_2 + tc_2 \\ z = a_3 + tc_3 \end{cases} \qquad \text{(Punkt- Richtungs- Gleichung).}$$

Für eine Gerade durch die zwei verschiedenen Punkte $A = (a_1, a_2, a_3)$ und $B = (b_1, b_2, b_3)$ erhält man die Gleichung

$$\begin{cases} x = a_1 + t \cdot (b_1 - a_1) \\ y = a_2 + t \cdot (b_2 - a_2) \\ z = a_3 + t \cdot (b_3 - a_3) \end{cases} \qquad \text{(Zwei- Punkte- Gleichung)}$$

Die parameterfreien Koordinatengleichungen der Geraden durch A und B sind

- $\dfrac{x - a_1}{b_1 - a_1} = \dfrac{y - a_2}{b_2 - a_2} = \dfrac{z - a_3}{b_3 - a_3}$, falls $a_i \neq b_i$ für $i = 1, 2, 3$
- $\dfrac{x - a_1}{b_1 - a_1} = \dfrac{y - a_2}{b_2 - a_2}$ und $z = a_3$, falls $a_i \neq b_i$ für $i = 1, 2$ und $a_3 = b_3$
- $x = a_1$ und $y = a_2$, falls $a_1 = b_1$, $a_2 = b_2$ und $a_3 \neq b_3$.

Neben den Geraden sind die Ebenen wichtige Teilmengen des Raumes.
Wir betrachten nun die Ebene E mit dem "Aufpunkt" A und den beiden (von $\vec{0}$ verschiedenen und nicht parallelen) Richtungsvektoren $\vec{u}$ und $\vec{v}$. Ein Raumpunkt X liegt genau dann auf E, wenn sich der Vektor $\overrightarrow{AX}$ in der Form $\overrightarrow{AX} = t\vec{u} + s\vec{v}$ mit Zahlen $s, t \in \mathbb{R}$ darstellen lässt, das heißt, man hat mit den zwei Parametern $t, s \in \mathbb{R}$ die Parameterdarstellung von E:

$$\overrightarrow{AX} = t\vec{u} + s\vec{v}. \tag{1}$$

Wird ein Kartesisches Koordinatensystem festgelegt, so dass $A = (a_1, a_2, a_3)$ und

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \quad \text{bzw.} \quad \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

ist, so die Parameterdarstellung (1) äquivalent zu den drei Koordinatengleichungen

$$\begin{cases} x = a_1 + tu_1 + sv_1 \\ y = a_2 + tu_2 + sv_2 \\ z = a_3 + tu_3 + sv_3. \end{cases} \tag{2}$$

Werden $\vec{u}$ und $\vec{v}$ durch die verschiedenen Punkte $A = (a_1, a_2, a_3)$, $B = (b_1, b_2, b_3)$ und $C = (c_1, c_2, c_3)$ mit $\vec{u} = \overrightarrow{AB}$, $\vec{v} = \overrightarrow{AC}$, also $u_i = b_i - a_i$ und $v_i = c_i - a_i$, bestimmt, dann geht (2) in die Drei-Punkte-Gleichung für die Ebene

$$\begin{cases} x = a_1 + t \cdot (b_1 - a_1) + s \cdot (c_1 - a_1) \\ y = a_2 + t \cdot (b_2 - a_2) + s \cdot (c_2 - a_2) \\ z = a_3 + t \cdot (b_3 - a_3) + s \cdot (c_3 - a_3) \end{cases}$$

über.

Kapitel 2

Grundlegende Strukturen

2.1 Mengen, Abbildungen und Verknüpfungen

Georg Cantor begründete die aus der Schule bekannte Mengenlehre mit folgender Definition:

Eine Menge ist die Zusammenfassung von Objekten,
die Gegenstand unseres Denkens sein können, zu einer Gesamtheit.

Von zentraler Bedeutung sind die Mengen

- $\mathbb{N} := \{1, 2, 3, \ldots\}$ (Menge der natürlichen Zahlen),
- $\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\}$,
- $\mathbb{Z} := \{0, 1, -1, 2, -2, \ldots\}$ (Menge der ganzen Zahlen),
- $\mathbb{Q}$ als die Menge der rationalen Zahlen, d.h. der Brüche ganzer Zahlen,
- $\mathbb{R}$ als die Menge der reellen Zahlen.

Wir nehmen in dieser Vorlesung an, dass wir wissen, was unter natürlichen, ganzen, rationalen und reellen Zahlen zu verstehen ist.

Definition 2.1.1. Es seien M und N Mengen.
Eine Menge M heißt Teilmenge von N (bzw. heißt N Obermenge von M), falls jedes Element von M auch ein Element von N ist (Schreibweise: $M \subseteq N$ bzw. $N \supseteq M$). Beispielsweise gilt $\mathbb{N} \subseteq \mathbb{Z}$, $\mathbb{Z} \subseteq \mathbb{Q}$ und $\mathbb{Q} \subseteq \mathbb{R}$. Wir schreiben zusammenfassend $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$.
Die Verneinung dieser Relation schreibt man $M \not\subseteq N$, d.h. M ist keine Teilmenge von N, es gibt also ein Element in M, das kein Element von N ist. Die Vereinigung ist $M \cup N = \{x \colon x \in M \text{ oder } x \in N\}$. Es seien beispielsweise $M = \{1, 2, 3\}$ und $N = \{2, 3, 5\}$. Dann ist $M \cup N = \{1, 2, 3, 5\}$.
Der Durchschnitt ist $M \cap N = \{x \colon x \in M \text{ und } x \in N\}$. Für $M = \{1, 2, 3\}$ und $N = \{2, 3, 5\}$ gilt dann $M \cap N = \{2, 3\}$.

Definition 2.1.2. Sind M und N Mengen, so definieren wir deren Differenz durch

$$M \backslash N := \{x \in M \colon x \notin N\}.$$

Ist $\mathcal{S}$ eine Menge von Mengen oder ein Mengensystem, so definieren wir die Vereinigung aller Mengen aus $\mathcal{S}$ als

$$\bigcup_{M \in \mathcal{S}} M := \{x\colon \text{es gibt ein } M \in \mathcal{S} \text{ mit } x \in M\}.$$

Sind n Mengen $M_1, M_2, \ldots, M_n$ gegeben, so schreiben wir $M_1 \cup M_2 \cup \cdots \cup M_n$ auch als

$$\bigcup_{k=1}^{n} M_k.$$

Ist für jedes $k \in \mathbb{N}$ eine Menge M_k gegeben, so schreiben wir statt

$$\bigcup_{k \in \mathbb{N}} M_k \quad \text{auch} \quad \bigcup_{k=1}^{\infty} M_k.$$

Eine entsprechende Schreibweise verwenden wir auch für den Durchschnitt:

$$\begin{aligned} \bigcap_{M \in \mathcal{S}} M &:= \{x\colon x \text{ liegt in jedem } M \in \mathcal{S}\} \\ \bigcap_{k=1}^{n} M_k &:= M_1 \cap M_2 \cap \cdots \cap M_n \\ \bigcap_{k=1}^{\infty} M_k &:= \bigcap_{k \in \mathbb{N}} M_k. \end{aligned}$$

Definition 2.1.3. Es seien alle Mengen $M \in \mathcal{S}$ eines Mengensystems $\mathcal{S}$ Teilmengen einer "Universalmenge" U. Wir bezeichnen die Komplementmenge $U \backslash M$ einer Menge $M \in \mathcal{S}$ mit M'. Mit der Schreibweise $U - M$ statt $U \backslash M$ deuten wir an, dass M eine Teilmenge von U ist.

Beispiel 2.1.4. Es sei $U = \mathbb{N} = \{1, 2, 3, \ldots\}$ sowie $M = \{2, 4, 6, 8, \ldots\}$.
Dann ist $U - M = M' = \{1, 3, 5, 7, \ldots\}$.

Es gelten die Komplementierungsregeln von de Morgan:

$$\left(\bigcup_{M \in \mathcal{S}} M \right)' = \bigcap_{M \in \mathcal{S}} (M') \quad \text{und} \quad \left(\bigcap_{M \in \mathcal{S}} M \right)' = \bigcup_{M \in \mathcal{S}} (M'),$$

d.h. das Komplement der Vereinigung ist gleich dem Schnitt der Komplemente, und das Komplement des Schnitts ist gleich der Vereinigung der Komplemente.

Definition 2.1.5. Es seien X und Y zwei nichtleere Mengen.
Unter einer Funktion oder Abbildung f von X nach Y versteht man eine Vorschrift, die jedem $x \in X$ genau ein $y \in Y$ zuordnet. Dieses dem Element x zugeordnete Element y bezeichnen wir mit $f(x)$ und nennen es den Wert der Funktion f an der Stelle x, oder das Bild von x unter f, während x das Urbild von $y = f(x)$ heißt.
Die Menge X wird die Definitionsmenge oder auch der Definitionsbereich, Y die Zielmenge von f genannt.

Bemerkung 2.1.6. Die detaillierteste Darstellung einer Funktion geschieht in der Form

$$f\colon \begin{cases} X & \to & Y \\ x & \mapsto & f(x) \end{cases}.$$

Es werden also zunächst die Mengen angegeben, zwischen denen die Funktion abbildet, dann die Zuordnung der einzelnen Elemente.

Beispiel 2.1.7. Es sei

$$f\colon \begin{cases} [1,3] & \to & \mathbb{R} \\ x & \mapsto & 2x \end{cases}$$

diejenige Funktion, die jeder Zahl aus dem Intervall $[1,3]$ ihr Doppeltes zuordnet. Dann ist beispielsweise $f(\frac{4}{3}) = \frac{8}{3}$.

Bemerkung 2.1.8. Anstelle von f kann natürlich jedes andere beliebige Symbol benutzt werden. Neben der Schreibweise des Arguments in Klammern sind auch andere Notationen üblich:

i) Das Argument im Index: a_n für $a(n)$, gebräuchlich für Folgen,

ii) Darstellung ohne Klammern: Fg für $F(g)$, gebräuchlich für höhere Operatoren,

iii) Funktionssymbol hinter dem Argument: A' für das Komplement von A.

Definition 2.1.9. Die Bildmenge oder schlichtweg das Bild einer Abbildung $f\colon A \to B$ ist die Menge $\mathrm{Bild}(f) = f(A) = \{f(a)\colon a \in A\}$.

Definition 2.1.10. Eine Abbildung $f\colon A \to B$ heißt

i) injektiv, falls aus $a \neq a'$ stets $f(a) \neq f(a')$ folgt, d.h. falls verschiedene Elemente aus A stets verschiedene Bilder in B besitzen,

ii) surjektiv, falls $f(A) = B$ ist, d.h. falls es zu jedem $b \in B$ ein Urbild $a \in A$ mit $f(a) = b$ gibt,

iii) bijektiv, falls f injektiv und surjektiv ist.

Definition 2.1.11. Es seien $k \in \mathbb{N}$ und $M_1, \dots, M_k$ beliebige Mengen. Das Kartesische Produkt der Mengen $M_1, \dots, M_k$ ist als

$$M_1 \times M_2 \times \cdots \times M_k := \{(m_1, \dots, m_k)\colon m_j \in M_j\}$$

definiert, d.h. als die Menge aller k- Tupel $(m_1, \dots, m_k)$, deren j-te Komponenten jeweils in M_j liegen. Ist $M_1 = M_2 = \cdots = M_k$, so schreiben wir kurz M^k.

Definition 2.1.12. Es sei X eine nichtleere Menge. Eine Abbildung $X^n \to X$ heißt eine n-stellige Verknüpfung auf X. Unter einer Verknüpfung schlechthin verstehen wir stets eine zweistellige Verknüpfung $\circ\colon X \times X \to X$.

Beispiel 2.1.13. Es sei $\mathbb{R}$ die Menge der reellen Zahlen. Wir haben auf $\mathbb{R}$ unter anderem die Verknüpfungen der Addition und der Multiplikation.

Bemerkung 2.1.14. Im Falle einer endlichen Menge $X = \{x_1, x_2, \dots, x_n\}$ kann man $\circ$ durch eine Verknüpfungstafel darstellen:

$\circ$	x_1	x_2	$\cdots$	x_n
x_1	$x_1 \circ x_1$	$x_1 \circ x_2$	$\cdots$	$x_1 \circ x_n$
x_2	$x_2 \circ x_1$	$x_2 \circ x_2$	$\cdots$	$x_2 \circ x_n$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_n	$x_n \circ x_1$	$x_n \circ x_2$	$\cdots$	$x_n \circ x_n$

Statt $a \circ b$ schreiben wir $a \cdot b$, ab, $a + b$ usw. für die verschiedenen Verknüpfungen. Beispielsweise besitzt $X = \{-1, 0, 1\}$ mit der üblichen Multiplikation die Verknüpfungstafel

$\cdot$	-1	0	1
-1	1	0	-1
0	0	0	0
1	-1	0	1

2.2 Gruppen

Definition 2.2.1. Es sei $\circ$ eine Verknüpfung auf einer Menge $G \neq \emptyset$. Dann heißt $(G, \circ)$ eine Gruppe, wenn die folgenden Axiome gelten:

(G1) Es gilt $a \circ (b \circ c) = (a \circ b) \circ c$ für alle $a, b, c \in G$ (Assoziativgesetz).

(G2) Es gibt mindestens ein neutrales Element $e \in G$ mit $e \circ a = a \circ e = a$ für alle $a \in G$.

(G3) Ist ein neutrales Element $e \in G$ gegeben, so gibt es zu jedem $a \in G$ ein inverses Element $a' \in G$ mit $a \circ a' = a' \circ a = e$.

Gilt zusätzlich das Axiom

(G4) Es gilt $a \circ b = b \circ a$ für alle $a, b \in G$ (Kommutativgesetz),

so heißt G eine abelsche oder auch kommutative Gruppe.

Beispiel 2.2.2. Es sei $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ die Menge der ganzen Zahlen mit der Verknüpfung der Addition. Das Assoziativgesetz (G1) ist offenbar erfüllt, denn für alle $a, b, c \in \mathbb{Z}$ gilt $(a + b) + c = a + (b + c)$. Die Zahl 0 ist das einzige neutrale Element wegen $0 + a = a + 0 = a$ für alle $a \in \mathbb{Z}$. Zu jeder Zahl $a \in \mathbb{Z}$ existiert ein eindeutiges Inverses, nämlich $-a$ mit $(-a) + a = 0$. Außerdem ist (G4) über $a + b = b + a$ für alle $a, b \in \mathbb{Z}$ erfüllt. Somit ist $\mathbb{Z}$ bzgl. der Addition eine abelsche Gruppe.

Beispiel 2.2.3. Ebenso bilden die Mengen $\mathbb{Q}$ der rationalen Zahlen und die Menge $\mathbb{R}$ der reellen Zahlen abelsche Gruppen bzgl. der Addition. Die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$ bilden keine Gruppe bzgl. der Addition, da kein neutrales Element existiert.

Beispiel 2.2.4. Die Menge V_2 aller Vektoren in der Ebene bildet eine abelsche Gruppe bzgl. der Vektoraddition: Der Nullvektor $\vec{0}$ ist das einzige neutrale Element, denn es gilt $\vec{0} + \vec{a} = \vec{a} + \vec{0} = \vec{a}$ für alle $a \in V_2$. Zu jedem Vektor $\vec{a}$ existiert ein eindeutig bestimmtes Inverses, nämlich $-\vec{a}$ mit $(-\vec{a}) + \vec{a} = \vec{0}$ für alle $\vec{a} \in V_2$. Ebenso bildet die Menge V_3 der Vektoren im Raum eine abelsche Gruppe.

Beispiel 2.2.5. Die Menge $\mathbb{Z}$ der ganzen Zahlen bildet keine Gruppe bzgl. der Verknüpfung Multiplikation. Es gibt zwar ein neutrales Element, nämlich die Zahl 1, aber es gibt für die Zahlen $a \neq \pm 1$ in $\mathbb{Z}$ kein Inverses.

Beispiel 2.2.6. Die Mengen $\mathbb{Q}\backslash\{0\}$ und $\mathbb{R}\backslash\{0\}$ bilden je abelsche Gruppen bzgl. der Multiplikation. In beiden Fällen ist die Zahl 1 das einzige neutrale Element. Die Zahl $a \neq 0$ hat $\frac{1}{a} = a^{-1}$ als Inverses.

Beispiel 2.2.7. (Eine nicht-abelsche Gruppe)
Es sei γ_3 die Menge der Permutationen von $\{1, 2, 3\}$, d.h. die Menge der bijektiven Abbildungen von $\{1, 2, 3\}$ auf sich selbst. Die Verknüpfung auf γ_3 ist die Komposition der Permutationen. Jede Permutation τ werde in der Form

$$\tau = \begin{pmatrix} 1 & 2 & 3 \\ \tau(1) & \tau(2) & \tau(3) \end{pmatrix}$$

notiert. Dann bildet γ_3 eine Gruppe mit neutralem Element

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}.$$

Jedes $\tau \in \gamma_3$ besitzt als Inverses seine Umkehrabbildung τ^{-1}. Die Gruppe γ_3 ist nicht abelsch, denn mit etwa

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \quad \text{und} \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

folgt

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \quad \text{aber} \quad \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Wir zeigen nun unseren ersten

Satz 2.2.8. *In einer beliebigen Gruppe $(G, \cdot)$ gelten folgende Eigenschaften:*

i) Es gibt genau ein neutrales Element.

ii) Jedes $a \in G$ hat genau ein Inverses, das wir mit a^{-1} bezeichnen.

Beweis. i) Wir vollziehen einen Beweis durch Widerspruch.
Angenommen es gebe verschiedene neutrale Elemente $e_1, e_2 \in G$. Da e_1 neutral ist gilt $e_1 \cdot a = a$ für alle $a \in G$, wir können also auch $a = e_2$ einsetzen und erhalten $e_1 \cdot e_2 = e_2$. Da auch e_2 neutral ist gilt $a \cdot e_2 = a$ für alle $a \in G$. Einsetzen von $a = e_1$ ergibt $e_1 \cdot e_2 = e_1$. Zusammensetzen der beiden Gleichungen ergibt

$$e_1 = e_1 \cdot e_2 = e_2,$$

ein Widerspruch zur Annahme, dass $e_1 \neq e_2$ ist.

ii) Es sei $a \in G$ beliebig und $a_1, a_2 \in G$ zwei Inversen, also $a \cdot a_1 = a_1 \cdot a = e$ bzw. $a \cdot a_2 = a_2 \cdot a = e$ mit dem nach i) eindeutig bestimmten neutralen Element e von G. Wir multiplizieren die zweite Gleichung von links mit a_1 und erhalten $a_1 \cdot (a \cdot a_2) = a_1 \cdot e$. Anwendung des Axioms (G1) ergibt die Gleichungen $(a_1 \cdot a) \cdot a_2 = a_1 \cdot e$. Anwendung von (G3) auf der linken Seite ergibt $e \cdot a_2 = a_1 \cdot e$, und nach (G2) folgt $a_2 = a_1$, also waren die Inversen gleich.

□

Bemerkung 2.2.9. Für eine multiplikativ geschriebene Gruppe nennt man das neutrale Element auch <u>Einselement</u> und schreibt 1 statt e. Der Multiplikationspunkt wird oft weggelassen, man schreibt also ab statt $a \cdot b$.

Wir zeigen nun ein paar Rechenregeln für Gruppen:

Satz 2.2.10. *In jeder Gruppe $(G, \cdot)$ gilt:*

i) Zu beliebigen $a, b \in G$ gibt es eindeutig bestimmte $x, y \in G$ mit $ax = b$ und $ya = b$, nämlich $x = a^{-1}b$ und $y = ba^{-1}$.

ii) Es gilt $(a^{-1})^{-1} = a$ für alle $a \in G$.

Beweis. i) Wir müssen zeigen, dass es eine Lösung der Gleichungen $ax = b$ bzw. $ya = b$ gibt, und dass diese eindeutig bestimmt ist.

- Existenz:
  Einsetzen von $x = a^{-1}b$ und $y = ba^{-1}$ ergibt

$$ax = a(a^{-1})b \underset{(G1)}{=} (aa^{-1})b \underset{(G2)}{=} eb = b \quad \text{bzw.} \quad ya = (ba^{-1})a \underset{(G1)}{=} b(a^{-1}a) \underset{(G2)}{=} be = b.$$

- Eindeutigkeit:
  Es seien $x_1, x_2 \in G$ zwei Lösungen von $ax = b$, so gilt $ax_1 = b = ax_2$. Nach Multiplikation von links mit a^{-1} erhalten wir $a^{-1}ax_1 = a^{-1}ax_2$, woraus mit (G3) und (G2) die Aussage $x_1 = x_2$ folgt. Die Lösungen waren also gleich.
  Die Rechnung für die Lösungen von $ya = b$ verläuft analog.

ii) Nach (G3) ist $\tilde{a} = (a^{-1})^{-1}$ ein Element aus G mit der Eigenschaft $\tilde{a} \cdot a^{-1} = a^{-1} \cdot \tilde{a} = e$. Auch a erfüllt diese Eigenschaft wegen (G2). Nach Satz 2.2.8 ii) ist $a = \tilde{a} = (a^{-1})^{-1}$.

□

Satz 2.2.11. *Für endlich viele Elemente $a_1, a_2, \ldots, a_n \in G$ gilt $(a_1 \cdot a_2 \cdots a_n)^{-1} = a_n^{-1} \cdot a_{n-1}^{-1} \cdots a_1^{-1}$.*

Beweis. Diese Aussage beweisen wir durch vollständige Induktion nach n.
Der Induktionsanfang ist $n = 1$: für nur ein einziges Element ist die Aussage $a_1^{-1} = a_1^{-1}$ richtig.
Die Induktionsannahme ist, dass für ein beliebiges $n \geq 1$ die Aussage $(a_1 \cdots a_n)^{-1} = a_n^{-1} \cdots a_1^{-1}$ gilt.
Der Induktionsschritt besteht darin, dass wir die Aussage für $n+1$ zeigen, indem wir die Induktionsannahme für n verwenden. Die Aussage für n lautet, dass die rechte Seite des Satzes das Axiom (G3) erfüllt, also

$$(a_1 \cdots a_n) \cdot (a_n^{-1} \cdots a_1^{-1}) = e.$$

Nach (G2) dürfen wir e in der Mitte einfügen ohne den Wert der linken Seite zu ändern:

$$(a_1 \cdots a_n) \cdot e \cdot (a_n^{-1} \cdots a_1^{-1}) = e.$$

Nach (G3) können wir e durch $a_{n+1}a_{n+1}^{-1}$ ersetzen und erhalten

$$(a_1 \cdots a_n) \cdot (a_{n+1}a_{n+1}^{-1}) \cdot (a_n^{-1} \cdots a_1^{-1}) = e.$$

Wegen (G1) dürfen wir die Klammen zu

$$(a_1 \cdots a_n \cdot a_{n+1}) \cdot (a_{n+1}^{-1} \cdot a_n^{-1} \cdots a_1^{-1}) = e$$

umsetzen. Damit ist die rechte Klammer das Inverse der linken Klammer nach (G3). Das ist die gewünschte Aussage für $n+1$. Damit haben wir die Induktion abgeschlossen, und der Satz gilt für alle $n \in \mathbb{N}$. □

Insbesondere ist die aus der Schule bekannte Regel $(a_1 \cdots a_n)^{-1} = a_1^{-1} \cdots a_n^{-1}$ nur in abelschen Gruppen richtig!

Bemerkung 2.2.12. Für eine additive geschriebene Gruppe $(G, +)$ ändern sich die Bezeichnungen und die Rechenregeln folgendermaßen:

i) Man schreibt "0" für das neutrale Element, und nennt es das Nullelement von $(G, +)$. Es gilt $a + 0 = 0 + a = a$ für alle $a \in G$.

ii) Man schreibt $-a$ für das Inverse statt a^{-1} und $a - b$ als Abkürzung für $a + (-b)$. Es gilt somit $a - a = -a + a = a + (-a) = 0$ für alle $a \in G$.

iii) Die Gleichungen $a + x = b$ und $y + a = b$ sind mit $x = -a + b$ und $y = b - a$ eindeutig lösbar.

iv) Es gilt $-(-a) = a$ und

$$-(a_1 + a_2 + \cdots + a_n) = -a_n - a_{n-1} - \cdots - a_1 = (-a_n) + (-a_{n-1}) + \cdots + (-a_1).$$

Üblicherweise benützt man die additive Schreibweise nur für abelsche Gruppen.

2.3 Ringe und Körper

Definition 2.3.1. Es seien $+$ und $\cdot$ zwei Verknüpfungen auf einer Menge $R \neq \emptyset$. Das Tripel $(R, +, \cdot)$ heißt Ring, wenn die folgenden Axiome erfüllt sind:

(R1) $(R, +)$ ist eine abelsche Gruppe mit dem Nullelement 0.

(R2) Für alle $a, b, c \in R$ ist $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativgesetz bzgl. $\cdot$).

(R3) Es gelten die Distributivgesetze

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \quad \text{und} \quad (b + c) \cdot a = (b \cdot a) + (c \cdot a).$$

Gilt außerdem

(R4) Für alle $a, b \in R$ ist $a \cdot b = b \cdot a$ (Kommutativgesetz),

so heißt R ein kommutativer Ring.
Gibt es ein $e \in R$ mit $e \cdot a = a \cdot e = a$ für alle $a \in R$, so heißt R ein Ring mit Eins, und wir schreiben 1 statt e.

Bemerkung 2.3.2. Wir vereinbaren die "Punkt-vor-Strich -Regel", d.h. der Ausdruck $a+bc$ ist eine Abkürzung für $a + (b \cdot c)$.

Satz 2.3.3. *In einem Ring R gilt für alle $a, b \in \mathbb{R}$*

$$a \cdot 0 = 0 \cdot a = 0, \quad a \cdot (-b) = (-a) \cdot b = -(a \cdot b) \quad \textit{und} \quad (-a) \cdot (-b) = a \cdot b.$$

Es gelten also die von den gewöhnlichen Zahlensystemen her bekannten Rechenregeln.

Beweis. Es gilt

$$a \cdot 0 = a \cdot (0 + 0) \underset{\text{(R3)}}{=} a \cdot 0 + a \cdot 0.$$

Subtraktion von $a \cdot 0$ auf beiden Seiten ergibt $0 = a \cdot 0$.
Weiter gilt

$$ab + a(-b) \underset{\text{(R3)}}{=} a \cdot (b + (-b)) = a \cdot 0 = 0$$

nach dem ersten Teil. Also gilt $a(-b) = -ab$. Der Beweis für $(-a)b = -ab$ verläuft analog.
Schließlich ist damit

$$(-a)(-b) = -(-a)b = -(-ab) = ab.$$

$\square$

Definition 2.3.4. Ein Ring $(K, +, \cdot)$ heißt Körper, wenn $(K \backslash \{0\}, \cdot)$ eine abelsche Gruppe ist.

In einem Körper schreibt man oft

$$ab^{-1} = b^{-1}a = \frac{a}{b}$$

mit $a, b \in K$ und $b \neq 0$.

Beispiel 2.3.5. Die Mengen $\mathbb{Q}$ der rationalen Zahlen und $\mathbb{R}$ der reellen Zahlen sind Körper bzgl. der vertrauten Addition und Multiplikation.

Beispiel 2.3.6. Die Menge $K = \{0, 1\}$ bildet mit den Verknüpfungen

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \text{und} \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

einen Körper. In ihm gilt $1 + 1 = 0$.

Beispiel 2.3.7. Die Menge $K = \{0, 1, a\}$ mit den Verknüpfungen

$$\begin{array}{c|ccc} + & 0 & 1 & a \\ \hline 0 & 0 & 1 & a \\ 1 & 1 & a & 0 \\ a & a & 0 & 1 \end{array} \qquad \text{und} \qquad \begin{array}{c|ccc} \cdot & 0 & 1 & a \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & a \\ a & 0 & a & 1 \end{array}$$

bildet einen Körper. Hier gilt $1 + 1 + 1 = 0$.

Beispiel 2.3.8. Die Menge $\mathbb{Z}$ der ganzen Zahlen bildet bzgl. der Addition und Multiplikation einen Ring, jedoch keinen Körper.

2.4 Der Körper der komplexen Zahlen

Definition 2.4.1. Die Menge $\mathbb{C}$ der <u>komplexen Zahlen</u> ist mit $\mathbb{R}^2 = \{(x, y) \colon x, y \in \mathbb{R}\}$ identisch.
Die <u>Addition</u> komplexer Zahlen wird <u>komponentenweise</u> erklärt:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

für $(x_1, y_1), (x_2, y_2) \in \mathbb{C}$. Die <u>Multiplikation</u> komplexer Zahlen wird durch

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + y_1 x_2)$$

definiert.

Es enthält $\mathbb{C}$ den Körper $\mathbb{R}$ der reellen Zahlen, wenn man $x \in \mathbb{R}$ mit $(x, 0) \in \mathbb{C}$ identifiziert. Diese Definitionen sind transparenter, wenn wir folgende Notation einführen:

Definition 2.4.2. Wir schreiben $(0, 1) = i$ (die <u>imaginäre Einheit</u>).
Jedes $z = (x, y) \in \mathbb{C}$ kann dann als $z = x + iy$ geschrieben werden. Man nennt x den <u>Realteil</u> von z, und y den <u>Imaginärteil</u> und schreibt $x = \Re(z)$ sowie $y = \Im(z)$.

Die Definition der Multiplikation folgt nun einfach aus den beiden Regeln

1. Es gilt $i^2 = -1$ (keine reelle Zahl erfüllt diese Eigenschaft).

2. Distributiv-, Kommutativ- und Assoziativgesetze gelten wie in einem Ring.

Dann gilt

$$(x + iy)(u + iv) \underset{\text{(R3)}}{=} x(u + iv) + iy(u + iv) \underset{\text{(R3)}}{=} xu + ixv + iyu + (iy)(iv) \underset{1.}{=} (xu - yu) + i(xv + yv).$$

Satz 2.4.3. *Es bildet* $(\mathbb{C}, +, \cdot)$ *einen Körper.*

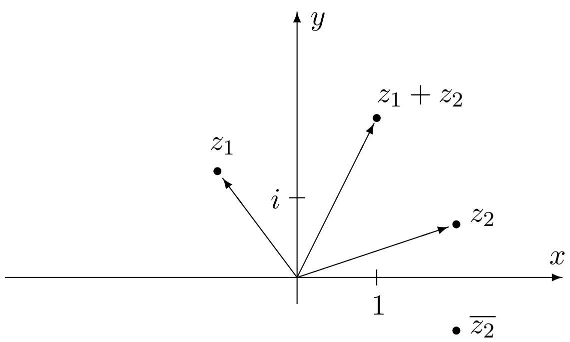
Beweis. Man prüft leicht die Gültigkeit der Körperaxiome. Zum Nachweis der inversen Elemente verwendet man die Regeln

$$-(x+yi) = -x + (-yi) \quad \text{und} \quad (x+yi)^{-1} = \frac{x}{x^2+y^2} - i \cdot \frac{y}{x^2+y^2},$$

falls $x + iy \neq 0$ gilt. □

Da die Menge $\mathbb{C}$ der komplexen Zahlen mit dem $\mathbb{R}^2$ identisch ist, ergibt sich die Möglichkeit der Darstellung in einer mit einem kartesischen Koordinatensystem versehenen Ebene. Die Zahl $x + iy$ identifizieren wir mit dem Punkt (x, y) oder auch mit dem Vektor $\binom{x}{y}$. Die Addition von komplexen Zahlen entspricht dann der Vektoraddition.

Definition 2.4.4. Ist $z = x + iy \in \mathbb{C}$ so nennt man $\overline{z} = x - iy$ die zu z konjugiert komplexe Zahl. Geometrisch gesehen ist $\overline{z}$ das Spiegelbild von z bzgl. der reellen Achse.



Satz 2.4.5. *Es seien $z, w \in \mathbb{C}$ komplexe Zahlen. Dann gilt*

i) Es ist $\overline{z+w} = \overline{z} + \overline{w}$.

ii) Es ist $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$ sowie

$$\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}},$$

falls $w \neq 0$ ist.

iii) Es gilt $\Re(z) = \frac{1}{2} \cdot (z + \overline{z})$ und $\Im(z) = \frac{1}{2i} \cdot (z - \overline{z})$. Insbesondere gilt $\overline{z} = z \Leftrightarrow z \in \mathbb{R}$.

Beweis. Übungsaufgabe □

Definition 2.4.6. Ist $z = x + iy$ mit $x, y \in \mathbb{R}$, so nennt man die nicht-negative reelle Zahl

$$|z| := \sqrt{x^2 + y^2}$$

den Absolutbetrag von z. Geometrisch ist $|z|$ der Abstand von z zum Ursprung bzw. die Länge des Vektors z.

Satz 2.4.7. *Für $z, w \in \mathbb{C}$ gilt*

$$|z|^2 = z \cdot \overline{z}, \quad |z \cdot w| = |z| \cdot |w| \quad \text{und} \quad \left|\frac{z}{w}\right| = \frac{|z|}{|w|}$$

für $w \neq 0$ sowie $|\overline{z}| = |z|$ und die Dreiecksungleichung

$$|w + z| \leq |w| + |z|.$$

Beweis. Durch Nachrechnen. □

Ursprünglich wurden die komplexen Zahlen zur einheitlichen Behandlung von quadratischen Gleichungen

$$az^2 + bz + c = 0 \tag{$*$}$$

mit $a, b, c \in \mathbb{R}$ und $a \neq 0$ eingeführt. Die Gleichung $z^2 + 1 = 0$ hat offenbar in $\mathbb{R}$ keine Lösung, wohl aber $\mathbb{C}$, nämlich $z = \pm\sqrt{-1} = \pm i$, und es gilt $z^2 + 1 = (z + i) \cdot (z - i)$.
Allgemein gilt: Die Gleichung $(*)$ hat die in $\mathbb{C}$ stets die beiden (nicht notwendigerweise verschiedenen) Lösungen

$$z_1 = -\frac{b}{2a} + \frac{1}{2a} \cdot \sqrt{b^2 - 4ac} \quad \text{und} \quad z_2 = -\frac{b}{2a} - \frac{1}{2a} \cdot \sqrt{b^2 - 4ac},$$

wobei $\sqrt{d}$ für $d \geq 0$ die positive Quadratwurzel von $d \in \mathbb{R}$ bezeichne und $\sqrt{d} = i \cdot \sqrt{-d}$ für $d < 0$.
Ferner bestätigt man leicht durch Nachrechnen die Gleichung

$$az^2 + bz + c = a \cdot (z - z_1) \cdot (z - z_2).$$

Kapitel 3

Vektorräume

3.1 Der Begriff des Vektorraums

Es sei V die Menge der Vektoren der Ebene oder des Raums. Diese Menge hat zusammen mit den Operationen der Addition und Skalarmultiplikation die folgenden Eigenschaften:

1. Es bildet $(V, +)$ eine abelsche Gruppe.
2. Es gilt $(\lambda + \mu)\vec{v} = \lambda\vec{v} + \mu\vec{v}$ für alle $\lambda, \mu \in \mathbb{R}$ und $\vec{v} \in V$ sowie $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$ für alle $\lambda \in \mathbb{R}$ und $\vec{u}, \vec{v} \in V$ (Distributivgesetze).
3. Es gilt $\lambda(\mu\vec{v}) = (\lambda\mu)\vec{v}$ (Assoziativgesetz).
4. Es gilt $1 \cdot \vec{v} = \vec{v}$.

Diese Eigenschaften dienen nun zur Definition der allgemeinen Struktur eines Vektorraums.
Ein größerer Grad an Allgemeinheit wird noch erzielt, indem der zur skalaren Multiplikation verwendete Körper $\mathbb{R}$ durch einen beliebigen Körper K ersetzt wird.

Definition 3.1.1. Es sei V eine Menge mit einer Verknüpfung $\oplus$ und $(K, +, \cdot)$ ein Körper mit Null 0 und Eins 1, sowie $\circ\colon K \times V \to V$ eine Abbildung. Das Tripel $(V, K, \circ)$ heißt ein Vektorraum über K (auch kurz VR) oder linearer Raum, wenn die folgenden Axiome für alle $\vec{u}, \vec{v} \in V$ und $\lambda, \mu \in K$ gelten:

(V1) Es bildet $(V, \oplus)$ eine abelsche Gruppe.

(V2) Es gelten die Distributivgesetze: $(\lambda + \mu) \circ \vec{u} = (\lambda \circ \vec{u}) \oplus (\mu \circ \vec{u})$ sowie $\lambda \circ (\vec{u} \oplus \vec{v}) = (\lambda \circ \vec{u}) \oplus (\lambda \circ \vec{v})$.

(V3) Es gilt das Assoziativgesetz: $\lambda \circ (\mu \circ \vec{u}) = (\lambda \cdot \mu) \circ \vec{u}$.

(V4) Es gilt das Unitaritätsgesetz: $1 \circ \vec{u} = \vec{u}$.

Die Elemente von V heißen Vektoren, die von K Skalare. Man nennt K den Skalarkörper des Vektorraums, die Verknüpfung $\circ$ nennt man die äußere oder skalare Multiplikation von Vektoren mit Skalaren, und die Operationen $\oplus$ und $\circ$ nennt man auch lineare Operationen.

Statt $(V, K, \circ)$ nennt man meist kurz V einen Vektorraum, wenn aus dem Zusammenhang Klarheit über K und $\circ$ besteht. In den wichtigen Fällen $K = \mathbb{R}$ und $K = \mathbb{C}$ spricht man auch von einem reellen Vektorraum bzw. einem komplexen Vektorraum.

Bemerkung 3.1.2. In dieser Vorlesung bezeichnen wir Vektoren mit einem Pfeil, etwa $\vec{v}$, und schreiben $\vec{0}$ für das Nullelement der abelschen Gruppe $(V, \oplus)$, welcher der Nullvektor genannt wird. In der Literatur gebräuchlich sind auch die altdeutschen Buchstaben $\mathfrak{u}$, Unterstriche $\underline{u}$ oder Fettdruck $\mathbf{u}$, um Vektoren von Skalaren zu unterscheiden. Sobald wir mit dem Sachverhalt vertraut sind, schreiben wir statt $\oplus$ und $\circ$ einfach $+$ und $\cdot$.

Wir stellen einige einfache Rechenregeln zusammen:

Satz 3.1.3. *In einem Vektorraum $(V, K, \circ)$ gilt für alle $\vec{v}, \vec{w} \in V$ und $\lambda, \mu \in K$:*

i) Es gilt genau dann $\lambda \circ \vec{v} = \vec{0}$, wenn $\lambda = 0$ oder $\vec{v} = \vec{0}$ gilt.

ii) Es gilt $(-\lambda) \circ \vec{v} = \lambda \circ (\ominus \vec{v}) = \ominus(\lambda \circ \vec{v})$, wobei $\ominus$ die Inversion in $(V, \oplus)$ ist.

iii) Es gelten $(\lambda - \mu) \circ \vec{v} = \lambda \circ \vec{v} \ominus \mu \circ \vec{v}$ sowie $\lambda \circ (\vec{v} \ominus \vec{w}) = \lambda \circ \vec{v} \ominus \lambda \circ \vec{w}$.

Beweis. i) Wir zeigen beide Richtungen:
"$\Rightarrow$": Es sei $\lambda \circ \vec{v} = \vec{0}$. Wenn $\lambda = 0$ gilt, so sind wir fertig. Ist $\lambda \neq 0$, so existiert ein λ^{-1} im Körper K. Es folgt

$$\vec{v} \underset{(V4)}{=} 1 \circ \vec{v} = (\lambda^{-1} \cdot \lambda) \circ \vec{v} \underset{(V3)}{=} \lambda^{-1} \circ (\lambda \circ \vec{v}) = \lambda^{-1} \circ \vec{0} = \vec{0}.$$

"$\Leftarrow$":
Es gelte $\lambda = 0$. Dann folgt

$$0 \circ \vec{v} = (0 + 0) \circ \vec{v} \underset{(V2)}{=} 0 \circ v \oplus 0 \circ \vec{v}.$$

Kürzen nach Satz 2.2.10 in der Gruppe $(V, \oplus)$ ergibt $0 \circ \vec{v} = \vec{0}$.
Andererseits gilt für $\vec{v} = \vec{0}$

$$\lambda \circ \vec{0} = \lambda \circ (\vec{0} \oplus \vec{0}) \underset{(V2)}{=} \lambda \circ \vec{0} \oplus \lambda \circ \vec{0}.$$

Kürzen ergibt $\lambda \circ \vec{0} = \vec{0}$.

ii) Es gilt

$$\lambda \circ \vec{v} \oplus (-\lambda) \circ \vec{v} \underset{(V2)}{=} (\lambda - \lambda) \circ \vec{v} = 0 \circ \vec{v} \underset{i)}{=} \vec{0},$$

also $(-\lambda) \circ \vec{v} = \ominus(\lambda \circ \vec{v})$. Ebenso folgt aus

$$\lambda \circ \vec{v} \oplus \lambda \circ (\ominus \vec{v}) \underset{(V2)}{=} \lambda \circ (\vec{v} \ominus \vec{v}) = \lambda \circ \vec{0} \underset{i)}{=} \vec{0}$$

die Gleichung $\lambda \circ (\ominus \vec{v}) = \ominus(\lambda \circ \vec{v})$.

iii) Dies folgt sofort aus (V2) und ii).

□

Beispiel 3.1.4. Es sei $(K, +, \cdot)$ ein Körper und $n \in \mathbb{N}$. Wir setzen

$$V = K^n = \{\vec{x} = (x_1, \ldots, x_n) \colon x_1, \ldots, x_n \in K\}.$$

Für $\vec{x} = (x_1, \ldots, x_n)$ und $\vec{y} = (y_1, \ldots, y_n)$ aus K^n sowie $\lambda \in K$ definieren wir

$$\vec{x} + \vec{y} = (x_1 + y_1, \ldots, x_n + y_n) \quad \text{und} \quad \lambda \cdot \vec{x} = (\lambda x_1, \ldots, \lambda x_n),$$

und nennen K^n den n-dimensionalen Standardraum über dem Körper K.

Satz 3.1.5. *Es ist* $(K^n, K, \cdot)$ *ein Vektorraum.*

Beweis. Wir haben zunächst zu beweisen, dass $(K^n, +)$ eine abelsche Gruppe ist. Die Gesetze einer abelschen Gruppe gelten alle, da sie für jede Komponente gelten. Beispielsweise ist

$$\begin{aligned}(\vec{x}+\vec{y})+\vec{z} &= ((x_1,\ldots,x_n)+(y_1,\ldots,y_n))+(z_1,\ldots,z_n)\\ &= (x_1+y_1,\ldots,x_n+y_n)+(z_1,\ldots,z_n)\\ &= ((x_1+y_1)+z_1,\ldots,(x_n+y_n)+z_n)\\ &= (x_1,\ldots,x_n)+(y_1+z_1,\ldots,y_n+z_n)\\ &= (x_1,\ldots,x_n)+((y_1,\ldots,y_n)+(z_1,\ldots,z_n)) = \vec{x}+(\vec{y}+\vec{z}).\end{aligned}$$

Das Nullelement von K^n ist $\vec{0} = (0,\ldots,0)$, und zu $(x_1,\ldots,x_n)$ ist $(-x_1,\ldots,-x_n)$ das Inverse. Auch die Gesetze (V2) bis (V4) beweist man durch Betrachtung der Komponenten. □

Beispiel 3.1.6. Es sei $F = \{f\colon [0,1] \to \mathbb{R}\}$ die Menge der auf dem Intervall $[0,1]$ definierten reellwertigen Funktionen. Für $f, g \in F$ und $\lambda \in \mathbb{R}$ werden $f+g$ und $\lambda \cdot f$ durch

$$(f+g)(x) = f(x)+g(x) \quad \text{und} \quad (\lambda\cdot f)(x) = \lambda \cdot f(x)$$

für alle $x \in \mathbb{R}$ erklärt. Dann ist, wie man leicht sieht, $(F, \mathbb{R}, \cdot)$ ein Vektorraum.

Beispiel 3.1.7. Es sei der Körper $(K, +, \cdot)$ im Körper $(L, +, \cdot)$ enthalten. Auf L betrachten wir die gewöhnliche Addition, während wir für die skalare Multiplikation nur Produkte $\lambda \cdot x$ für $\lambda \in K$ und $x \in L$ betrachten. Dann ist $(L, K, \cdot)$ ein Vektorraum. Zum Beispiel ist der Körper $\mathbb{C}$ der komplexen Zahlen ein Vektorraum über dem Körper $\mathbb{R}$ der reellen Zahlen.

Beispiel 3.1.8. Es sei wiederum der Körper $(K, +, \cdot)$ im Körper $(L, +, \cdot)$ enthalten. Jeder Vektorraum V über L kann dann auch als Vektorraum über K betrachtet werden. Man braucht lediglich die Skalarmultiplikation auf Skalare aus K zu beschränken. Insbesondere kann jeder komplexe Vektorraum auch als reeller Vektorraum angesehen werden.

3.2 Unterräume

Definition 3.2.1. Eine Teilmenge U eines Vektorraums V über K heißt <u>Unterraum</u> oder <u>Teilraum</u> von V, wenn U bzgl. der in V gegebenen Operationen selbst ein Vektorraum über K ist.

Um nachzuweisen, dass $U \subseteq V$ ein Unterraum von V ist, braucht man nicht alle Vektorraumaxiome nachzuprüfen. Vielmehr gilt

Satz 3.2.2. *Es sei* $U \neq \emptyset$ *eine Teilmenge des Vektorraums* V *über* K*. Diese Teilmenge* U *ist genau dann ein Unterraum von* V*, wenn die folgenden Eigenschaften gelten:*

(U1) Aus $\vec{v}, \vec{w} \in U$ *folgt* $\vec{v}+\vec{w} \in U$ *(Abgeschlossenheit der Addition).*

(U2) Aus $\vec{v} \in U$ *und* $\lambda \in K$ *folgt* $\lambda\vec{v} \in U$ *(Abgeschlossenheit der skalaren Multiplikation).*

Beweis. Ist U ein Unterraum, so gelten (U1) und (U2) offensichtlich. Umgekehrt mögen nun (U1) und (U2) für eine Teilmenge $U \subseteq V$ gelten. Wegen (U1) ist $+$ tatsächlich eine Verknüpfung auf U. Da das Assoziativ- und das Kommutativgesetz in V gelten, gelten sie erst recht in U. Wegen $U \neq \emptyset$ gibt es mindestens ein $\vec{v}_0 \in U$. Mit (U2) folgt $0 \cdot \vec{v}_0 = \vec{0} \in U$. Nach (U2) gilt mit $\vec{v} \in U$ auch $-\vec{v} = (-1)\cdot\vec{v} \in U$. Also ist $(U, +)$ eine abelsche Gruppe. Wegen (U2) ist die Skalarmultiplikation auf U definiert, außerdem gelten die Eigenschaften (V2) bis (V4) in V, also erst recht in U. Damit ist U ein Vektorraum. □

Die Bedingungen (U1) und (U2) lassen sich zu einer Eigenschaft zusammenfassen:

Satz 3.2.3. *(Unterraumkriterium)*
Eine Teilmenge $U \neq \emptyset$ eines Vektorraums V über K ist genau dann ein Unterraum von V, wenn für alle $\vec{v}, \vec{w} \in U$ und $\lambda, \mu \in K$ dann $\lambda\vec{v} + \mu\vec{w} \in U$ gilt.

Beweis. Ist U ein Unterraum, so gilt diese Aussage offensichtlich. Nun gelte $\lambda\vec{v} + \mu\vec{w} \in U$. Dann folgen (U1) und (U2) durch die Wahlen $\lambda = \mu = 1$ bzw. $\mu = 0$. □

Beispiel 3.2.4. Jeder Vektorraum V besitzt die Unterräume $\{\vec{0}\}$, welcher auch der Nullraum genannt wird, und V selbst. Die leere Teilmenge $\emptyset$ ist kein Unterraum.

Beispiel 3.2.5. Es sei V_2 der Vektorraum der Vektoren der Ebene. Unterräume von V_2 sind dann $\{\vec{0}\}$, der V_2 selbst und für jedes $\vec{v} \in V_2 \backslash \{\vec{0}\}$ die Menge $\{\lambda\vec{v} \colon \lambda \in \mathbb{R}\}$, d.h. die Menge der Richtungsvektoren einer Geraden.

Beispiel 3.2.6. Es sei V_3 der Vektorraum der Vektoren des Raums. Unterräume sind analog zum Vektorraum V_2 die Mengen $\{\vec{0}\}$, V_3 und für jedes $\vec{v} \in V_3 \backslash \{\vec{0}\}$ die Menge $\{\lambda\vec{v} \colon \lambda \in \mathbb{R}\}$. Außerdem bildet für jedes Paar nichtparalleler Vektoren $\vec{v}, \vec{w}$ die Menge $\{\lambda\vec{v} + \mu\vec{w} \colon \lambda, \mu \in \mathbb{R}\}$ einen Unterraum, d.h. die Menge der Richtungsvektoren einer Ebene.

Beispiel 3.2.7. Der Standardraum $\mathbb{R}^2$ enthält die Unterräume $\{\vec{0}\}$, den $\mathbb{R}^2$ selbst, sowie für jedes $\vec{v} \in \mathbb{R}^2$ mit $\vec{v} \neq \vec{0}$ die Menge $\{\lambda\vec{v} \colon \lambda \in \mathbb{R}\}$, d.h. die Geraden durch den Nullpunkt.

Beispiel 3.2.8. Der Standardraum $\mathbb{R}^3$ enthält die Unterräume $\{\vec{0}\}$, den $\mathbb{R}^3$ selbst, sowie sämtliche Geraden und Ebenen durch den Nullpunkt.

Beispiel 3.2.9. Es sei $K = \{0, 1\}$ der Körper mit zwei Elementen aus Beispiel 2.3.6. Ein Beispiel eines Unterraums U des Standardraums K^n ist die Menge aller n-Tupel mit einer geraden Anzahl von Einsen. Diese bezeichnet man auch als Tupel mit gerader Parität. Diese kann wegen $1 + 1 = 0$ in K auch über

$$U = \{\vec{x} \in K^n \colon x_1 + x_2 + \cdots + x_n = 0\}$$

definiert werden. Solche Unterräume spielen bei der Nachrichtenübertragung eine große Rolle: Jedes Stück Information wird durch eine Kette von Nullen und Einsen codiert. Damit Fehler bei der Übermittlung nicht unerkannt bleiben, kann man sich entscheiden, nur Ketten gerader Parität zu verwenden. Auch andere Unterräume des K^n spielen in der Codierungstheorie eine große Rolle.

Satz 3.2.10. *Der Durchschnitt beliebig vieler Unterräume eines Vektorraums V ist wieder ein Unterraum von V.*

Beweis. Übungsaufgabe □

Definition 3.2.11. Für eine beliebige Teilmenge M eines Vektorraums V heißt

$$\langle\, M \,\rangle := \bigcap \{U \colon M \subseteq U,\ U \text{ ist Unterraum von } V\}$$

der von M erzeugte oder aufgespannte Unterraum oder die lineare Hülle von M. Man nennt M auch ein Erzeugendensystem und $\langle\, M \,\rangle$ das Erzeugnis. Im Falle einer endlichen Menge $M = \{\vec{v}_1, \ldots, \vec{v}_k\}$ schreibt man auch $\langle\, \vec{v}_1, \ldots, \vec{v}_k \,\rangle$ statt $\langle\, \{\vec{v}_1, \ldots, \vec{v}_k\} \,\rangle$.

Beispiel 3.2.12. Das Erzeugnis der leeren Menge $\emptyset$ ist der Nullraum $\{\vec{0}\}$.

Definition 3.2.13. Es sei V ein Vektorraum über K und $\vec{v}_1, \ldots, v_k \in V$. Jede Summe der Form

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k$$

mit Skalaren $\lambda_1, \ldots, \lambda_k \in K$ heißt eine <u>Linearkombination</u> von $\vec{v}_1, \ldots, \vec{v}_k$.
Man nennt λ_j den <u>Koeffizienten</u> von $\vec{v}_j$ mit $1 \leq j \leq k$ in der Summe.

Nach dem Unterraumkriterium liegt jede Linearkombination der Länge $k = 2$ wieder in V. Per Induktion sieht man, dass auch Linearkombinationen beliebiger Länge aus V wieder in V liegen. Mithilfe des Begriffs der Linearkombination lässt sich nun die lineare Hülle $\langle M \rangle$ einer Menge M viel einfacher beschreiben:

Satz 3.2.14. *Es sei V ein Vektorraum über K und $M \subseteq V$ mit $M \neq \emptyset$. Dann ist $\langle M \rangle$ gleich der Menge aller Linearkombinationen von Elementen aus M mit Koeffizienten aus K:*

$$\langle M \rangle = \{\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k \colon k \in \mathbb{N},\ \vec{v}_j \in M,\ \lambda_j \in K \text{ für } 1 \leq j \leq k\}.$$

Beweis. Bezeichnen wir die Menge der Linearkombinationen mit $L(M)$, so gilt

$$L(M) \subseteq \langle M \rangle, \tag{1}$$

da jeder über M liegende Unterraum von V jede Linearkombination über M enthält.
Andererseits zeigt man mit Satz 3.2.3 leicht, dass $L(M)$ ein Unterraum von V ist. Es gilt nämlich mit $\vec{v} = \lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k \in L(M)$ und $\vec{w} = \mu_1 \vec{w}_1 + \cdots + \mu_m \vec{w}_m \in L(M)$ und $\vec{v}_1, \ldots, \vec{v}_k, \vec{w}_1, \ldots, \vec{w}_m \in M$ auch

$$\lambda \vec{v} + \mu \vec{w} = \lambda \lambda_1 \vec{v}_1 + \cdots + \lambda \lambda_k \vec{v}_k + \mu \mu_1 \vec{w}_1 + \cdots + \mu \mu_m \vec{w}_m \in L(M)$$

für alle $\lambda, \mu \in K$. Da $\langle M \rangle$ nach Definition in jedem Unterraum von V enthalten ist, der M enthält, folgt

$$\langle M \rangle \subseteq L(M). \tag{2}$$

Aus (1) und (2) folgt die Behauptung des Satzes. □

Beispiel 3.2.15. Falls M nur aus ein oder zwei Elementen besteht, und es sich bei V um den $\mathbb{R}^2$ oder den $\mathbb{R}^3$ handelt, so hat $\langle M \rangle$ eine einfache geometrische Bedeutung: Ist $M = \{\vec{v}_1\}$ mit $\vec{v}_1 \neq \vec{0}$, so ist nach Satz 3.2.14 eben $\langle M \rangle = \{\lambda \vec{v}_1 \colon \lambda \in \mathbb{R}\}$, die Gerade durch den Nullpunkt, die $\vec{v}_1$ enthält. Ist $M = \{\vec{v}_1, \vec{v}_2\}$, so ist $\langle M \rangle = \{\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \colon \lambda_1, \lambda_2 \in \mathbb{R}\}$ die Ebene durch den Nullpunkt, die $\vec{v}_1$ und $\vec{v}_2$ enthält, falls $\vec{v}_1, \vec{v}_2$ nicht zueinander parallel sind.

Definition 3.2.16. Es seien $U_1, \ldots, U_k$ Unterräume des Vektorraums V. Dann heißt

$$U_1 + \cdots + U_k = \{\vec{u}_1 + \cdots + \vec{u}_k \colon u_j \in U_j\}$$

die <u>Summe</u> der Vektorräume $U_1, \ldots, U_k$.

Satz 3.2.17. *Die Summe $U_1 + \cdots + U_k$ ist ein Unterraum von V.*

Beweis. Es ist $U_1 + \cdots + U_k \neq \emptyset$ wegen $\vec{0} = \vec{0} + \cdots + \vec{0} \in U_1 + \cdots + U_k$. Es seien $\vec{v} = \vec{v}_1 + \cdots + \vec{v}_k$ und $\vec{w} = \vec{w}_1 + \cdots + \vec{w}_k$ beliebig aus der Menge $U_1 + \cdots + U_k$ sowie $\lambda, \mu \in K$. Dann ist auch

$$\lambda \vec{v} + \mu \vec{w} = \underbrace{(\lambda \vec{v}_1 + \mu \vec{w}_1)}_{\in U_1} + \cdots + \underbrace{(\lambda \vec{v}_k + \mu \vec{w}_k)}_{\in U_k}$$

ein Vektor aus der Summe $U_1 + \cdots + U_k$. Mit Satz 3.2.3 folgt die Behauptung. □

3.3 Lineare Abhängigkeit, Basis, Dimension

Es sei V stets ein Vektorraum über dem Körper K.

Definition 3.3.1. Vektoren $\vec{v}_1, \dots, \vec{v}_k \in V$ heißen linear abhängig (kurz l.a.), wenn es $\lambda_1, \dots, \lambda_k \in K$ gibt, so dass

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k = \vec{0}$$

und $(\lambda_1, \dots, \lambda_k) \neq (0, \dots, 0)$ gilt. Andernfalls, d.h. wenn für $\lambda_j \in K$ stets

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_k \vec{v}_k = \vec{0} \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$$

gilt, heißen die Vektoren $\vec{v}_1, \dots, \vec{v}_k$ linear unabhängig (kurz l.u.).

Beispiel 3.3.2. Für $V = \mathbb{R}^3$, $\vec{v}_1 = (1,2,0)$, $\vec{v}_2 = (0,3,2)$, $\vec{v}_3 = (1,0,0)$ und $\vec{v}_4 = (3,2,0)$ sind $\vec{v}_1, \dots, \vec{v}_4$ linear abhängig, denn es gilt $1 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2 + 2 \cdot \vec{v}_3 + (-1) \cdot \vec{v}_4 = \vec{0}$.
Die Vektoren $\vec{v}_1, \dots, \vec{v}_3$ sind dagegen linear unabhängig, denn aus $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \lambda_3 \vec{v}_3 = \vec{0}$ folgt

$$\begin{cases} \lambda_1 & & & + & \lambda_3 & = & 0 \\ 2\lambda_1 & + & 3\lambda_2 & & & = & 0 \\ & & 2\lambda_2 & & & = & 0 \end{cases} \Rightarrow \lambda_2 = 0,\ \lambda_1 = 0,\ \lambda_3 = 0.$$

Beispiel 3.3.3. Im K^n sind $\vec{e}_1 = (1, 0, \dots, 0)$, $\vec{e}_2 = (0, 1, 0, \dots, 0)$,$\dots$,$\vec{e}_n = (0, \dots, 0, 1)$ stets linear unabhängig.

Beispiel 3.3.4. Im $\mathbb{R}^3$ hat die lineare Abhängigkeit von zwei bzw. drei Vektoren eine anschauliche geometrische Bedeutung. Sind $\vec{v}_1$ und $\vec{v}_2$ nämlich linear abhängig, so gilt $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \vec{0}$ mit $(\lambda_1, \lambda_2) \neq (0, 0)$. Es sei OBdA $\lambda_2 \neq 0$, woraus $\vec{v}_2 = -\lambda_2^{-1} \lambda_1 \vec{v}_1$ folgt, d.h. die beiden Vektoren $\vec{v}_1, \vec{v}_2$ liegen auf einer Geraden durch den Nullpunkt. Ähnlich bedeutet im $\mathbb{R}^3$ die lineare Abhängigkeit zweier Vektoren deren Parallelität. Sind nämlich $\vec{v}_1, \vec{v}_2, \vec{v}_3$ linear abhängig, so gilt $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \lambda_3 \vec{v}_3 = \vec{0}$ mit $(\lambda_1, \lambda_2, \lambda_3) \neq (0, 0, 0)$. Es sei OBdA $\lambda_3 \neq 0$, woraus $\vec{v}_3 = -\lambda_3^{-1} \lambda_1 \vec{v}_1 - \lambda_3^{-1} \lambda_2 \vec{v}_2$ folgt, d.h. die drei Vektoren $\vec{v}_1, \vec{v}_2, \vec{v}_3$ liegen auf einer Ebene durch den Nullpunkt.

Satz 3.3.5. *Es gilt:*

i) Der Nullvektor $\vec{0}$ ist stets linear abhängig, ein einzelner Vektor $\vec{v} \neq \vec{0}$ dagegen ist stets linear unabhängig.

ii) Mit $\vec{v}_1, \dots, \vec{v}_k$ sind auch $\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_l$ mit $l \geq k$ linear abhängig.

iii) Sind $\vec{v}_1, \dots, \vec{v}_k$ linear unabhängig, so auch $\vec{v}_1, \dots, \vec{v}_m$ für $1 \leq m \leq k$.

iv) Ist $\vec{v}$ eine Linearkombination der Vektoren $\vec{v}_1, \dots, \vec{v}_k$, so sind $v_1, \dots, v_k, \vec{v}$ linear abhängig.

v) Sind $k \geq 2$ Vektoren linear abhängig, so ist wenigstens einer von ihnen eine Linearkombination der anderen.

vi) Sind $\vec{v}_1, \dots, \vec{v}_k$ linear unabhängig, aber $\vec{v}_1, \dots, \vec{v}_k, \vec{v}$ linear abhängig, so ist $\vec{v}$ eine Linearkombination der Vektoren $\vec{v}_1, \dots, \vec{v}_k$.

Beweis. i) Es gilt $1 \cdot \vec{0} = \vec{0}$, also ist $\vec{0}$ linear abhängig. Andererseits folgt aus $\lambda \vec{v} = \vec{0}$ mit $\vec{v} \neq \vec{0}$, dass $\lambda = 0$ ist, also ist $\vec{v}$ nach Satz 3.1.3 i) linear unabhängig.

ii) Aus $\lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k = \vec{0}$ und $(\lambda_1, \ldots, \lambda_k) \neq (0, \ldots, 0)$ folgt auch

$$\lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k + \lambda_{k+1}\vec{v}_{k+1} + \cdots + \lambda_l\vec{v}_l = \vec{0}$$

mit $(\lambda_1, \ldots, \lambda_l) \neq (0, \ldots, 0)$, wenn man $\lambda_{k+1} = \cdots = \lambda_l = 0$ einsetzt.

iii) Wären $\vec{v}_1, \ldots, \vec{v}_m$ linear abhängig, so ist nach ii) auch $\vec{v}_1, \ldots, \vec{v}_k$.

iv) Ist $\vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k$, so folgt $\lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k + (-1)\vec{v} = \vec{0}$.

v) Es sei $\lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k = \vec{0}$ und $(\lambda_1, \ldots, \lambda_k) \neq (0, \ldots, 0)$. Nach eventueller Umnummerierung können wir annehmen, dass $\lambda_k \neq 0$ ist. Dann folgt $\vec{v}_k = (-\lambda_k^{-1})\lambda_1\vec{v}_1 + \cdots + (-\lambda_k)^{-1})\lambda_{k-1}\vec{v}_{k-1}$.

vi) Es gelte $\lambda_1\vec{v}_1 + \cdots + \lambda_k\vec{v}_k + \lambda\vec{v} = \vec{0}$ mit $(\lambda_1, \ldots, \lambda_k) \neq (0, \ldots, 0)$. Dann muss $\lambda \neq 0$ sein, da wegen der linearen Unabhängigkeit von $\vec{v}_1, \ldots, \vec{v}_k$ die Gleichung $\lambda_1\vec{v}_1 + \cdots + \lambda\vec{v}_k = \vec{0}$ nur für $\lambda_1 = \cdots = \lambda_k = 0$ möglich ist. Dann ist aber $\vec{v} = (-\lambda^{-1})\lambda_1\vec{v}_1 + \cdots + (-\lambda^{-1})\lambda_k\vec{v}_k$.

□

Der Begriff der linearen (Un-) Abhängigkeit lässt sich auf beliebige (auch unendliche) Mengen von Vektoren erweitern:

Definition 3.3.6. Eine Menge $M \subseteq V$ heißt linear unabhängig, wenn je endlich viele verschiedene Vektoren aus M linear unabhängig sind, andernfalls linear abhängig.

Bemerkung 3.3.7. Wir stellen einige Spezialfälle zusammen:

i) Die leere Menge ist linear uanabhängig.

ii) Ist $\vec{0} \in M$, so ist M linear abhängig.

iii) Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.

iv) Jede in V liegende Obermenge einer linear abhängigen Menge ist linear abhängig.

v) Achtung: Im Falle $\vec{a}_1 = \vec{a}_2 \neq \vec{0}$ sind $\vec{a}_1$ und $\vec{a}_2$ linear abhängig, aber $\{\vec{a}_1, \vec{a}_2\}$ ist linear unabhängig, weil es tatsächlich die Menge $\{\vec{a}_1\}$ ist.

Beispiel 3.3.8. Es sei $G = \{f\colon \mathbb{R} \to \mathbb{R}\}$ der Vektorraum der auf $\mathbb{R}$ definierten reellwertigen Funktionen und $M := \{f_\nu\colon \nu \in \mathbb{Z}\}$ mit

$$f_\nu(x) := \begin{cases} 1, & \text{falls } \nu \leq x < v + 1, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist M linear unabhängig.
Es sei nämlich $\lambda_1 f_{\nu_1} + \cdots + \lambda_k f_{\nu_k} = 0$, wobei 0 die Nullfunktion bezeichne, mit paarweise verschiedenen Indizes $\nu_1, \ldots, \nu_k$. Für x mit $\nu_i \leq x < \nu_i + 1$ erhalten wir

$$0 = (\lambda_1 f_{\nu_1} + \cdots + \lambda_k f_{\nu_k})\,(x) = \lambda_i f_{\nu_i}(x) = \lambda_i,$$

also $\lambda_1 = \cdots = \lambda_k = 0$.

Beispiel 3.3.9. Es sei $\mathbb{R}$ aufgefasst als Vektorraum über dem Unterkörper $\mathbb{Q}$ und $M = \{\pi^\nu\colon \nu \in \mathbb{N}_0\}$ mit der Kreiszahl $\pi = 3,14159\ldots$. Dann ist M linear unabhängig, denn

$$\sum_{\nu=0}^{n} \lambda_\nu \pi^\nu = 0$$

mit $\lambda_\nu \in \mathbb{Q}$ ist nur für $\lambda_0 = \ldots = \lambda_n = 0$ möglich.
Der Beweis dieser Aussage ist schwierig. Man sagt, dass π eine transzendente Zahl ist.

Definition 3.3.10. Ist $M \subseteq V$ linear unabhängig, so schreiben wir $\langle\, M \,\rangle = \langle\langle\, M \,\rangle\rangle$, und nennen M eine Basis des Erzeugnisses $\langle\, M \,\rangle$, kurz:

$$V = \langle\langle\, M \,\rangle\rangle \Leftrightarrow \; M \text{ ist linear unabhängig und } \langle\, M \,\rangle = V.$$

Speziell ist eine Basis eines Vektorraums V also ein linear unabhängiges Erzeugendensystem von V. Im Falle $M = \{\vec{v}_1, \ldots, \vec{v}_k\}$ sagen wir auch, dass die Vektoren $\vec{v}_1, \ldots, \vec{v}_k$ eine Basis von V bilden, und schreiben wieder kurz $\langle\langle\, \vec{v}_1, \ldots, \vec{v}_k \,\rangle\rangle$ statt $\langle\langle\, \{\vec{v}_1, \ldots, \vec{v}_k\} \,\rangle\rangle$.

Beispiel 3.3.11. Es ist $\{\vec{0}\} = \langle\langle\, \emptyset \,\rangle\rangle$, und $\emptyset$ ist die einzige Basis von $\{\vec{0}\}$.

Beispiel 3.3.12. Die Vektoren $\vec{e}_1, \ldots, \vec{e}_n \in K^n$ aus Beispiel 3.3.3 sind linear unabhängig, und wir nennen $\{\vec{e}_1, \ldots, \vec{e}_n\}$ die Standardbasis von K.

Beispiel 3.3.13. Der Körper $\mathbb{C}$ der komplexen Zahlen, aufgefasst als Vektorraum über $K = \mathbb{R}$, hat die Basis $\{1, i\}$.

Man kann zeigen, dass jeder Vektorraum eine Basis besitzt. Wir werden uns jedoch bei der Diskussion der Basis auf relativ einfache Fälle, sogenannte endlichdimensionale Vektorräume beschränken.

Definition 3.3.14. Gibt es eine maximale Zahl n linear unabhängiger Vektoren in V, so heißt n die Dimension von V:

$$\dim(V) = \max\{|M| \colon M \subseteq V \text{ linear unabhängig}\} \in \mathbb{N}_0.$$

Gibt es kein solches n, existiert also zu jedem $k \in \mathbb{N}$ eine linear unabhängige Teilmenge $M \subseteq V$ mit $|M| = k$, so heißt V unendlichdimensional, und wir schreiben $\dim(V) = \infty$.

Beispiel 3.3.15. Es ist $\dim(\{\vec{0}\}) = 0$.

Beispiel 3.3.16. Es sei $V = K$ ein Körper aufgefasst als ein Vektorraum über sich selbst.
Die Menge $\{1\}$ ist linear unabhängig, und sind $\lambda_1, \lambda_2 \in K$ mit $\lambda_1 \neq \lambda_2$, so gilt $\lambda_2\lambda_1 + (-\lambda_1)\lambda_2 = 0$, d.h. die Vektoren λ_1, λ_2 sind stets linear abhängig, also $\dim(K) = 1$.

Beispiel 3.3.17. Es ist $\dim(\mathbb{R}^2) = 2$. Die Vektoren $(1, 0)$ und $(0, 1)$ sind linear unabhängig, also gilt $\dim(\mathbb{R}^2) \geq 2$. Andererseits sind je drei Vektoren $\vec{a} = (a_1, a_2)$, $\vec{b} = (b_1, b_2)$, $\vec{c} = (c_1, c_2)$ linear abhängig, denn es gilt

$$(b_1c_2 - b_2c_1)\vec{a} + (c_1a_2 - a_1c_2)\vec{b} + (a_1b_2 - a_2b_1)\vec{c} = \vec{0}.$$

Ist hier etwa $b_1c_2 - b_2c_1 = 0$, so sind schon $\vec{b}$ und $\vec{c}$ linear abhängig, also $c_1\vec{b} - b_1\vec{c} = \vec{0} = c_2\vec{b} - b_2\vec{c}$. Ist aber $b_1 = b_2 = 0$, so ist schon $\vec{b}$ allein linear abhängig.

Beispiel 3.3.18. Ebenso ist $\dim(\mathbb{C}) = 2$, wenn $\mathbb{C}$ als Vektorraum über $\mathbb{R}$ aufgefasst wird.

Beispiel 3.3.19. Der Vektorraum $G = \{f \colon \mathbb{R} \to \mathbb{R}\}$ besitzt die Dimension $\dim(G) = \infty$.

Satz 3.3.20. *Es sei* $\dim(V) = n < \infty$. *Dann besitzt V eine Basis, genauer bildet jede linear unabhängige Teilmenge mit n Vektoren eine Basis von V.*

Beweis. Es seien $\vec{v}_1, \ldots, \vec{v}_n \in V$ linear unabhängig und $\vec{v} \in V$ beliebig. Nach Definition der Dimension sind $\vec{v}_1, \ldots, \vec{v}_n, \vec{v}$ linear abhängig. Nach Satz 3.3.5 vi) ist $\vec{v}$ eine Linearkombination von $\vec{v}_1, \ldots, \vec{v}_n$, also $\vec{v} \in \langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle$. Da $\vec{v} \in V$ beliebig war, folgt $V = \langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle = \langle\langle\, \vec{v}_1, \ldots, \vec{v}_n \,\rangle\rangle$. □

Wir werden in Kürze zeigen, dass es keine Basis von V mit weniger als $\dim(V)$ Elementen gibt.

Satz 3.3.21. *Für $\vec{v}_1, \dots \vec{v}_n \in V$ sind folgende Aussagen äquivalent:*

i) Es gilt $V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_n \,\rangle\rangle$.

ii) Jedes $\vec{v}$ besitzt eine Darstellung

$$\vec{v} = \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n \tag{$*$}$$

mit eindeutig bestimmten Koeffizienten $(\lambda_1, \dots, \lambda_n) \in K^n$, oder die durch $(*)$ vermittelte Abbildung $K^n \to V$, $(\lambda_1, \dots, \lambda_n) \to \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n$ ist bijektiv.

Beweis. i) $\Rightarrow$ ii):
Da $\{\vec{v}_1, \dots, \vec{v}_n\}$ ein Erzeugendensystem von V bildet, hat jedes $\vec{v} \in V$ mindestens eine Darstellung der Form $(*)$. Es seien $\vec{v} = \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n$ und $\vec{v} = \mu_1 \vec{v}_1 + \cdots + \mu_n \vec{v}_n$ zwei Darstellungen des gleichen Vektors. Dann folgt aus $(\lambda_1 - \mu_1)\vec{v}_1 + \cdots + (\lambda_n - \mu_n)\vec{v}_n = \vec{0}$ über die lineare Unabhängigkeit der $\vec{v}_j$ sogleich $\lambda_j - \mu_j = 0$ für $j = 1, \dots, n$ und damit auch $\lambda_j = \mu_j$, d.h. die Darstellung $(*)$ ist eindeutig.

ii) $\Rightarrow$ i):
Eine und damit die einzige Darstellung $(*)$ von $\vec{0}$ ist $\vec{0} = 0 \cdot \vec{v}_1 + \cdots + 0 \cdot \vec{v}_n$. Also sind $\vec{v}_1, \dots, \vec{v}_n$ linear unabhängig.

□

Satz 3.3.22. *Es sei $V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_n \,\rangle\rangle$ und $\vec{w} \in V$ mit $\vec{w} \neq \vec{0}$. In der Darstellung*

$$\vec{w} = \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n \tag{$*$}$$

mit $(\lambda_1, \dots, \lambda_n) \in K^n$ sei j ein Index mit $\lambda_j \neq 0$. Dann gilt

$$V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_{j-1}, \vec{w}, \vec{v}_{j+1}, \dots, \vec{v}_n \,\rangle\rangle\,.$$

Beweis. Es sei OBdA $j = 1$, ansonsten nummerieren wir die $\vec{v}_j$ um. Wir müssen nun $V = \langle\langle\, \vec{w}, \vec{v}_2, \dots, \vec{v}_n \,\rangle\rangle$ zeigen.

- $V = \langle\, \vec{w}, \vec{v}_2, \dots, \vec{v}_n \,\rangle$:
  Auflösen von $(*)$ nach $\vec{v}_1$ ergibt
  $$\vec{v}_1 = \mu_1 \vec{w} + \mu_2 \vec{v}_2 + \cdots + \mu_n \vec{v}_n \tag{$**$}$$
  mit $\mu_i = -\frac{\lambda_i}{\lambda_1}$ für $i = 2, \dots, n$. Es sei nun $\vec{v} \in V$ beliebig, etwa $\vec{v} = \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n$ mit $\alpha_1, \dots, \alpha_n \in K$. Einsetzen von $(**)$ ergibt
  $$\begin{aligned} \vec{v} &= \alpha_1(\mu_1 \vec{w} + \mu_2 \vec{v}_2 + \cdots + \mu_n \vec{v}_n) + \alpha_2 \vec{v}_2 + \cdots + \alpha_n \vec{v}_n \\ &= \alpha_1 \mu_1 \vec{w} + (\alpha_1 \mu_2 + \alpha_2)\vec{v}_2 + \cdots + (\alpha_1 \mu_n + \alpha_n)\vec{v}_n \in \langle\, \vec{w}, \vec{v}_2, \dots, \vec{v}_n \,\rangle\,. \end{aligned}$$

- $\{\vec{w}, \vec{v}_2, \dots, \vec{v}_n\}$ ist linear unabhängig:
  Angenommen wir haben
  $$\mu_1 \vec{w} + \mu_2 \vec{v}_2 + \cdots + \mu_n \vec{v}_n = \vec{0}.$$
  Dann folgt mit $(*)$
  $$\vec{0} = \mu_1(\lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n) + \mu_2 \vec{v}_2 + \cdots + \mu_n \vec{v}_n = (\mu_1 \lambda_1)\vec{v}_1 + (\mu_1 \lambda_2 + \mu_2)\vec{v}_2 + \cdots + (\mu_1 \lambda_n + \mu_n)\vec{v}_n$$
  und damit
  $$\mu_1 \lambda_1 = \mu_1 \lambda_2 + \mu_2 = \cdots = \mu_1 \lambda_n + \mu_n = 0,$$
  da $\vec{v}_1, \dots, \vec{v}_n$ linear unabhängig sind, weswegen $\mu_1 = 0$ wegen $\lambda_1 \neq 0$ ist. Daraus folgt dann $\mu_2 = \cdots = \mu_n = 0$, und somit ist $\{\vec{w}, \vec{v}_2, \dots, \vec{v}_n\}$ linear unabhängig.

□

Wir verallgemeinern Satz 3.3.22 zu dem wichtigen

Satz 3.3.23. *(Austauschsatz von Steinitz, Basisergänzungssatz)*
Es sei V ein Vektorraum mit $V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_n \,\rangle\rangle$ und $\vec{w}_1, \dots, \vec{w}_k \in V$ linear unabhängig, womit $k \leq n$ gilt. Ferner lässt sich aus $\{\vec{v}_1, \dots, \vec{v}_n\}$ eine Teilmenge $\{\vec{v}'_{k+1}, \dots, \vec{v}'_n\}$ so auswählen, dass $V = \langle\langle\, \vec{w}_1, \dots, \vec{w}_k, \vec{v}'_{k+1}, \dots, \vec{v}'_n \,\rangle\rangle$ gilt.
Mit anderen Worten: Man kann k geeignet gewählte Vektoren der Basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ gegen die $\vec{w}_i$ austauschen, so dass man wieder eine Basis von V erhält. Insbesondere lässt sich jede linear unabhängige Teilmenge eines endlichdimensionalen Vektorraums zu einer Basis von V ergänzen.

Beweis. Wir führen eine vollständige Induktion nach k für festes n durch.
Induktionsanfang $k = 1$:
Es ist $n \geq 1$. Wegen $\vec{w}_1 \neq \vec{0}$ ist in der Basisdarstellung $\vec{w}_1 = \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n$ mindestens ein $\lambda_j \neq 0$ enthalten. Nach Satz 3.3.22 kann man $\vec{v}_j$ gegen $\vec{w}_1$ austauschen, so dass $\{\vec{v}_1, \dots, \vec{v}_{j-1}, \vec{w}_1, \vec{v}_{j+1}, \dots, \vec{v}_n\}$ wieder eine Basis von V ist.
Induktionsschritt $k \to k + 1$:
Die Behauptung des Satzes gelte für ein $k \in \mathbb{N}$, und es seien $\vec{w}_1, \dots, \vec{w}_k, \vec{w}_{k+1} \in V$ linear unabhängig. Nach Induktionsannahme, angewandt auf $\vec{w}_1, \dots, \vec{w}_k$, ist $k \leq n$, und es gibt eine Teilmenge $\{\vec{v}'_{k+1}, \dots, \vec{v}'_n\} \subseteq \{\vec{v}_1, \dots, \vec{v}_n\}$ derart, dass

$$V = \langle\langle\, \vec{w}_1, \dots, \vec{w}_k, \vec{v}'_{k+1}, \dots, \vec{v}'_n \,\rangle\rangle\,. \qquad (*)$$

Wäre $k = n$, so wäre schon $V = \langle\langle\, \vec{w}_1, \dots, \vec{w}_k \,\rangle\rangle$, also $\vec{w}_{k+1}$ eine Linearkombination von $\vec{w}_1, \dots, \vec{w}_k$. Da $\vec{w}_1, \dots, \vec{w}_k, \vec{w}_{k+1}$ aber linear unabhängig sein sollen, muss folglich $k < n$ gelten, d.h. $k + 1 \leq n$ sein. Wegen $(*)$ und $\vec{w}_{k+1} \neq 0$ gilt

$$\vec{w}_{k+1} = \mu_1 \vec{w}_1 + \cdots + \mu_k \vec{w}_k + \mu_{k+1} \vec{v}'_{k+1} + \cdots + \mu_n \vec{v}'_n$$

mit $\mu_1, \dots, \mu_n \in K$, wobei mindestens ein $\mu_i \neq 0$ ist. Dabei kann nicht $\mu_{k+1} = \cdots = \mu_n = 0$ sein, denn sonst wären $\vec{w}_1, \dots, \vec{w}_k, \vec{w}_{k+1}$ linear abhängig, also ist $\mu_i \neq 0$ für mindestens ein $i \in \{k+1, \dots, n\}$. Austauschen von $\vec{v}'_i$ gegen $\vec{w}_{k+1}$ gemäß Satz 3.3.22 ergibt die Behauptung für $k + 1$. □

Satz 3.3.24. *Es sei $V = \langle\langle\, \vec{v}_1, \dots, \vec{v}_n \,\rangle\rangle$. Dann ist $\dim(V) = n$, und eine Teilmenge $B \subseteq V$ ist genau dann eine Basis von V, wenn B aus n linear unabhängigen Vektoren besteht.*

Beweis. Für jede linear unabhängige Menge $M \subseteq V$ gilt $|M| \leq n$, also $\dim(V) \leq n$ nach Satz 3.3.23. Nach Definition der Dimension ist andererseits $n \leq \dim(V)$, woraus $\dim(V) = n$ folgt.
Nun sei B eine beliebige Basis von V. Dann folgt zunächst $m := |B| \leq \dim(V) = n$ und dann wie oben $\dim(V) = m$, also $m = n$. Umgekehrt ist nach Satz 3.3.20 auch jede linear unabhängige Menge $B \subseteq V$ mit $|B| = \dim(B)$ eine Basis von V. □

Beispiel 3.3.25. Für die Standardvektorräume gilt $\dim(K^n) = n$, denn die Standardbasis $\{\vec{e}_1, \dots, \vec{e}_n\}$ hat n Elemente. Insbesondere ist $\dim(\mathbb{R}^n) = n$ und $\dim(\mathbb{C}^n) = n$ als Vektorraum über $\mathbb{C}$. Allerdings besitzt $\mathbb{C}^n$ als Vektorraum über $\mathbb{R}$ die Dimension $2n$, und eine Basis ist $\{\vec{e}_1, \dots, \vec{e}_n, i\vec{e}_1, \dots, i\vec{e}_n\}$.

Satz 3.3.26. *Es sei $\dim(V) < \infty$ und U ein Unterraum von V Dann gilt*

i) $\dim(U) \leq \dim(V)$.

ii) $\dim(U) = \dim(V) \Leftrightarrow U = V$.

Beweis. i) Dies folgt sofort aus der Definition der Dimension.

ii) Es sei $\dim(U) = \dim(V) = n$. Dann besitzt U nach Satz 3.3.20 eine Basis $\{\vec{v}_1, \dots, \vec{v}_n\}$. Diese bildet dann ebenfalls nach Satz 3.3.20 eine Basis von U, also $U = V$. Die Rückrichtung ist klar.

□

Satz 3.3.27. *(Dimensionssatz für Summenräume)*
Es seien U_1 und U_2 endlichdimensionale Unterräume eines Vektorraums V. Dann gilt

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

Beweis. Da $U_1 \cap U_2$ ein Unterraum von U_1 und ebenso von U_2 ist, gilt nach Satz 3.3.26 die Aussage $d = \dim(U_1 \cap U_2) < \infty$. Es sei also nach Satz 3.3.20 nun $\{\vec{a}_1, \dots, \vec{a}_d\}$ eine Basis von $U_1 \cap U_2$, was im Falle $d = 0$ auch die leere Menge sein kann. Wir ergänzen diese Basis nach Satz 3.3.23 zu je einer Basis von U_1 und U_2:

$$\begin{aligned} B_1 &:= \{\vec{a}_1, \dots, \vec{a}_d, \vec{b}_1, \dots, \vec{b}_r\} \quad \text{eine Basis von } U_1 \\ B_2 &:= \{\vec{a}_1, \dots, \vec{a}_d, \vec{c}_1, \dots, \vec{c}_s\} \quad \text{eine Basis von } U_2. \end{aligned}$$

Wir behaupten nun, dass $B := B_1 \cup B_2 = \{\vec{a}_1, \dots, \vec{a}_d, \vec{b}_1, \dots, \vec{b}_r, \vec{c}_1, \dots, \vec{c}_s\}$ eine Basis von $U_1 + U_2$ ist.
Wir haben zwei Aussagen zu zeigen: $\langle B \rangle = U_1 + U_2$ und die lineare Unabhängigkeit von B:

- $\underline{\langle B \rangle = U_1 + U_2}$:
  Wegen $B \subseteq U_1 \cup U_2 \subseteq U_1 + U_2$, da $U_1 + U_2$ ein Unterraum ist, folgt $\langle B \rangle \subseteq U_1 + U_2$. Andererseits ist
  $$U_1 + U_2 = \langle B_1 \rangle + \langle B_2 \rangle \subseteq \langle B \rangle + \langle B \rangle = \langle B \rangle$$
  wegen $B_1, B_2 \subseteq B$.

- $\underline{B \text{ ist linear unabhängig}}$:
  Es sei
  $$\alpha_1\vec{a}_1 + \cdots + \alpha_d\vec{a}_d + \beta_1\vec{b}_1 + \cdots \beta_r\vec{b}_r + \gamma_1\vec{c}_1 + \cdots + \gamma_s\vec{c}_s = \vec{0}$$
  mit $\alpha_i, \beta_j, \gamma_k \in K$. Wir müssen nun zeigen, dass alle Koeffizienten verschwinden. Sortieren ergibt
  $$\underbrace{\alpha_1\vec{a}_1 + \cdots + \alpha_d\vec{a}_d + \beta_1\vec{b}_1 + \cdots + \beta_r\vec{b}_r}_{\in U_1} = -\underbrace{\gamma_1\vec{c}_1 + \cdots + \gamma_s\vec{c}_s}_{\in U_2} \subseteq U_1 \cap U_2.$$
  Dann gibt es Koeffizienten λ_l mit $-(\gamma_1\vec{c}_1 + \cdots + \gamma_s\vec{c}_s) = \lambda_1\vec{a}_1 + \cdots + \lambda_d\vec{a}_d$, also
  $$\lambda_1\vec{a}_1 + \cdots + \lambda_d\vec{a}_d + \gamma_1\vec{c}_1 + \cdots + \gamma_s\vec{c}_s = \vec{0}.$$
  Da B_2 als Basis linear unabhängig ist, gilt $\lambda_1 = \cdots = \lambda_d = \gamma_1 = \cdots = \gamma_s = 0$. Daraus folgt
  $$\alpha_1\vec{a}_1 + \cdots + \alpha_d\vec{a}_d + \beta_1\vec{b}_1 + \cdots + \beta_r\vec{b}_r = \vec{0}$$
  und somit $\alpha_1 = \cdots = \alpha_d = \beta_1 = \cdots = \beta_r = 0$, da auch B_1 linear unabhängig ist. Also verschwinden sämtliche Koeffizienten $\alpha_i, \beta_j, \gamma_k$, weswegen B linear unabhängig ist.

Da wir jetzt Basen für alle beteiligten Räume haben, können wir die Aussage des Satzes durch Zählen der Basisvektoren zeigen:

$$\dim(U_1 + U_2) = |B| = d + r + s = (d + r) + (d + s) - d = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

□

Beispiel 3.3.28. Es sei $V = \mathbb{R}^3$ und $U_1 = \langle\langle\, (1,0,0), (0,1,0) \,\rangle\rangle$ sowie $U_2 = \langle\langle\, (1,-1,0), (0,0,1) \,\rangle\rangle$. Also gilt $\dim(U_1) = \dim(U_2) = 2$. Dabei sind U_1 und U_2 Ebenen durch $\vec{0}$, und zwar explizit

$$\begin{aligned} U_1 &= \{(x,y,0)\colon x,y \in \mathbb{R}\}, \text{ die } xy\text{-Ebene}, \\ U_2 &= \{(x,-x,z)\colon x,z \in \mathbb{R}\}, \\ U_1 \cap U_2 &= \{(x,-x,0)\colon x \in \mathbb{R}\} = \langle\langle\, (1,-1,0) \,\rangle\rangle\,, \text{ die Gerade durch } y = -x \text{ und } z = 0. \end{aligned}$$

Insbesondere ist $\dim(U_1) \cap \dim(U_2) = 1$. Mit dem Dimensionssatz folgt $\dim(U_1 + U_2) = 2 + 2 - 1 = 3$, also $U_1 + U_2 = \mathbb{R}^3$, d.h. alle $\vec{v} \in \mathbb{R}^3$ lassen sich als $\vec{v} = \vec{u}_1 + \vec{u}_2$ mit $\vec{u}_1 \in U_1$ und $\vec{u}_2 \in U_2$ schreiben. Diese Darstellung ist jedoch nicht eindeutig.

Kapitel 4

Matrizen

4.1 Grundlegende Definitionen

Im folgenden sei K wieder ein Körper.

Definition 4.1.1. Unter einer Matrix vom Typ (m, n) mit $m, n \in \mathbb{N}$ oder einer $m \times n$-Matrix über dem Körper K versteht man ein rechteckiges Schema der Gestalt

$$\mathcal{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

mit $a_{ij} \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$.
Die Einträge a_{ij} heißen Komponenten oder Koeffizienten der Matrix.
Den Vektor $\vec{a}_i = (a_{i1}, \ldots, a_{in}) \in K^n$ für $1 \leq i \leq m$ bezeichnet man als den i-ten Zeilenvektor oder kurz die i-te Zeile von $\mathcal{A}$, den Vektor $\vec{b}_j = (a_{1j}, \ldots, a_{mj}) \in K^m$ für $1 \leq j \leq n$ als den j-ten Spaltenvektor oder kurz die j-te Spalte von $\mathcal{A}$. Wir schreiben $\vec{b}_j$ oft in Spaltenschreibweise

$$\vec{b}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Man schreibt dann auch kurz

$$\mathcal{A} = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_m \end{pmatrix} = (\vec{b}_1, \ldots, \vec{b}_n) = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Die Menge aller Matrizen vom Typ (m, n) bezeichnet man mit $K^{(m,n)}$ oder $K^{m \times n}$. Die Gerade in $\mathcal{A}$, auf der die Elemente $a_{11}, \ldots, a_{rr}$ mit $r = \min(m, n)$ stehen, nennt man die Hauptdiagonale von $\mathcal{A}$. Durch Spiegelung an der Hauptdiagonalen erhält man aus der Matrix $\mathcal{A}$ die Matrix $\mathcal{A}^T$, die als die Transponierte von $\mathcal{A}$ bezeichnet wird. Sie hat die Gestalt

$$\mathcal{A}^T = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ a_{12} & \cdots & a_{m2} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} \in K^{(n,m)}.$$

Die Zeilen von $\mathcal{A}$ werden also die Spalten von $\mathcal{A}^T$, und die Spalten von $\mathcal{A}$ werden die Zeilen von $\mathcal{A}^T$, also

$$\mathcal{A} = (a_{kl})_{\substack{1\leq k\leq m\\ 1\leq l\leq n}} \Leftrightarrow \mathcal{A}^T = (b_{kl})_{\substack{1\leq k\leq n\\ 1\leq l\leq m}}$$

mit $b_{kl} = a_{lk}$.

Matrizen desselben Typs über K können komponentenweise addiert und mit Skalaren aus M multipliziert werden:

Definition 4.1.2. Es seien $\mathcal{A}, \mathcal{B} \in K^{(m,n)}$ mit

$$\mathcal{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \text{und} \quad \mathcal{B} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}.$$

Dann versteht man unter der <u>Summe</u> von $\mathcal{A}$ und $\mathcal{B}$ die Matrix

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

Ist $\lambda \in K$, so setzt man

$$\lambda\mathcal{A} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}.$$

Man schreibt $(-1) \cdot \mathcal{A} = -\mathcal{A}$.

Beispiel 4.1.3. Es sei $K = \mathbb{R}$ und

$$\mathcal{A} = \begin{pmatrix} 3 & -1 & 5 & 7 \\ 2 & 1 & 4 & 3 \\ 0 & -8 & 2 & 6 \end{pmatrix} \quad \text{und} \quad \mathcal{B} = \begin{pmatrix} 2 & 0 & -4 & -3 \\ -2 & -1 & 0 & -5 \\ 1 & 8 & 1 & 4 \end{pmatrix}.$$

Dann ist

$$\mathcal{A} + \mathcal{B} = \begin{pmatrix} 5 & -1 & 1 & 4 \\ 0 & 0 & 4 & -2 \\ 1 & 0 & 3 & 10 \end{pmatrix} \quad \text{und} \quad 3\mathcal{A} = \begin{pmatrix} 9 & -3 & 15 & 21 \\ 6 & 3 & 12 & 9 \\ 0 & -24 & 6 & 18 \end{pmatrix}.$$

Definition 4.1.4. Die Matrix vom Typ (m, n), deren sämtliche Komponenten verschwinden, nennt man die <u>Nullmatrix</u>

$$0 = 0^{(m,n)} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Man zeigt leicht

Satz 4.1.5. *Der $K^{(m,n)}$ bildet bzgl. der Matrizenaddition und Skalarmultiplikation mit der Nullmatrix als Nullelement einen Vektorraum über K mit Dimension* $\dim(K^{(m,n)}) = m \cdot n$.

Beweis. Übungsaufgabe $\square$

Von großer Bedeutung ist auch das Produkt von Matrizen $\mathcal{A}$ und $\mathcal{B}$. Im allgemeinen Fall sind hier jedoch $\mathcal{A}$ und $\mathcal{B}$ von verschiedenem Typ. Die Motivation für die kompliziert anmutende Definition wird erst später im Kapitel über lineare Abbildungen ersichtlich.

Definition 4.1.6. Es seien

$$\mathcal{A} = (a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \in K^{(m,n)} \quad \text{und} \quad \mathcal{B} = (b_{kl})_{\substack{1\leq k\leq n\\ 1\leq l\leq r}} \in K^{(n,r)}$$

gegeben. Dann versteht man unter dem <u>Produkt</u> $\mathcal{C} = \mathcal{A}\mathcal{B}$ die Matrix $\mathcal{C} = (c_{il})_{\substack{1\leq i\leq m\\ 1\leq l\leq r}} \in K^{(m,r)}$ mit

$$c_{il} = \sum_{j=1}^{n} a_{ij} b_{jl}$$

mit $1 \leq i \leq m$ bzw. $1 \leq l \leq r$.

Bemerkung 4.1.7. Das Element in der i-ten Zeile und der l-ten Spalte der Produktmatrix $\mathcal{C}$ wird also erhalten, indem die Elemente der i-ten Zeile von $\mathcal{A}$ und der l-ten Spalte von $\mathcal{B}$ paarweise multipliziert und die Produkte addiert werden:

$$\mathcal{C} = \left(\begin{array}{c|c} & \\ \hline & \end{array}\right) \leftarrow \mathcal{A} = \left(\begin{array}{c} \\ i\text{-te Zeile} \\ \hline \\ \end{array}\right), \quad \mathcal{B} = \begin{array}{c} l\text{-te Spalte} \\ \left(\begin{array}{c|c} \quad & \quad \end{array}\right) \end{array}.$$

Damit das Produkt zweier Matrizen $\mathcal{A}$ und $\mathcal{B}$ definiert ist, muss die Anzahl der Spalten von $\mathcal{A}$ mit der Anzahl der Zeilen $\mathcal{B}$ übereinstimmen. Die Produktmatrix $\mathcal{C} = \mathcal{A}\mathcal{B}$ hat dann dieselbe Anzahl Zeilen wie $\mathcal{A}$ und die diesselbe Anzahl Spalten wie $\mathcal{B}$.

Beispiel 4.1.8. Es sei $K = \mathbb{R}$ und

$$\mathcal{A} = \begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & 1 & 1 \\ 3 & 7 & 4 & 2 \end{pmatrix}, \ \mathcal{B} = \begin{pmatrix} 1 & -1 \\ 3 & 0 \\ 2 & 5 \\ 1 & 6 \end{pmatrix}.$$

Dann ist das Produkt

$$\mathcal{A} \cdot \mathcal{B} = \begin{pmatrix} 14 & 23 \\ -3 & 11 \\ 34 & 29 \end{pmatrix}.$$

Ein neutrales Element bezüglich dieser Multiplikation bildet die sogenannte <u>Einheitsmatrix</u>.

Definition 4.1.9. Es sei $n \in \mathbb{N}$. Die Matrix vom Typ (n, n), $E_n = (\delta_{ij})$, $1 \leq i, j \leq n$ mit dem <u>Kroneckersymbol</u>

$$\delta_{ij} = \begin{cases} 1, & \text{falls } i = j \\ 0 & \text{sonst.} \end{cases}$$

heißt <u>Einheitsmatrix</u> vom Typ (n, n).

Die Matrix E_n hat also Einsen auf der Hauptdiagonale und Nullen außerhalb der Hauptdiagonalen.

Im folgenden Satz werden die wichtigsten Eigenschaften der Matrizenmultiplikation zusammengefasst:

Satz 4.1.10. *Für Matrizen $\mathcal{A}, \mathcal{B}, \mathcal{C}$ über K gilt, falls die Ausdrücke definiert sind, d.h. die Spalten- und Zeilenzahl zueinander passen:*

i) das Assoziativgesetz: $(\mathcal{A}\mathcal{B})\mathcal{C} = \mathcal{A}(\mathcal{B}\mathcal{C})$.

ii) die Distributivgesetze: $\mathcal{A}(\mathcal{B}+\mathcal{C}) = \mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C}$ und $(\mathcal{B}+\mathcal{C})\mathcal{A} = \mathcal{B}\mathcal{A}+\mathcal{C}\mathcal{A}$.

iii) die Neutralität der Einheitsmatrix E_n: $E_n \cdot \mathcal{A} = \mathcal{A} \cdot E_n = \mathcal{A}$.

iv) Im Allgemeinen ist $\mathcal{A}\mathcal{B} \neq \mathcal{B}\mathcal{A}$, d.h. das Kommutativgesetz gilt nicht.

Beweis. i) Es sei

$$\mathcal{A} = (a_{fg})_{\substack{1\le f\le l\\ 1\le g\le m}}, \quad \mathcal{B} = (b_{hi})_{\substack{1\le h\le m\\ 1\le i\le n}} \quad \text{und} \quad \mathcal{C} = (c_{jk})_{\substack{1\le j\le n\\ 1\le k\le r}}.$$

Die in der Behauptung auftretenden Produkte sind offenbar definiert, wir bezeichnen sie mit

$$\mathcal{A}\mathcal{B} = (d_{fi})_{\substack{1\le f\le l\\ 1\le i\le n}} \quad \text{und} \quad \mathcal{B}\mathcal{C} = (e_{hk})_{\substack{1\le h\le m\\ 1\le k\le r}}.$$

Nach Definition ist

$$d_{fi} = \sum_{g=1}^{m} a_{fg}b_{gi} \quad \text{und} \quad e_{hk} = \sum_{i=1}^{n} b_{hi}c_{ik}. \qquad (*)$$

Für die Komponente u_{fk} von $(\mathcal{A}\mathcal{B})\mathcal{C}$ gilt nach Definition des Produkts und $(*)$

$$u_{fk} = \sum_{i=1}^{n} d_{fi}c_{ik} = \sum_{i=1}^{n}\left(\sum_{g=1}^{m} a_{fg}b_{gi}\right)c_{ik}.$$

Andererseits gilt für das Element v_{fk} von $\mathcal{A}(\mathcal{B}\mathcal{C})$

$$v_{fk} = \sum_{g=1}^{m} a_{fg}e_{gk} = \sum_{g=1}^{m} a_{fg}\sum_{i=1}^{n} b_{gi}c_{ik}.$$

Nach dem Distributivgesetz in K gilt $u_{fk} = v_{fk}$, und damit $(\mathcal{A}\mathcal{B})\mathcal{C} = \mathcal{A}(\mathcal{B}\mathcal{C})$.

ii) Nun sei

$$\mathcal{A} = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}, \quad \mathcal{B} = (b_{kl})_{\substack{1\le k\le n\\ 1\le l\le r}} \quad \text{und} \quad \mathcal{C} = (c_{kl})_{\substack{1\le k\le n\\ 1\le l\le r}}.$$

Dann gilt für die Komponente d_{il} von $\mathcal{A}(\mathcal{B}+\mathcal{C})$

$$d_{il} = \sum_{j=1}^{n} a_{ij}(b_{jl}+c_{jl}).$$

Für die Komponente e_{il} bzw. f_{il} von $\mathcal{A}\mathcal{B}$ bzw. $\mathcal{A}\mathcal{C}$ gilt

$$e_{il} = \sum_{j=1}^{n} a_{ij}b_{jl} \quad \text{bzw.} \quad f_{il} = \sum_{j=1}^{n} a_{ij}c_{jl},$$

also $d_{il} = e_{il} + f_{il}$, und somit $\mathcal{A}(\mathcal{B}+\mathcal{C}) = \mathcal{A}\mathcal{B}+\mathcal{A}\mathcal{C}$.
Das zweite Distributivgesetz wird analog bewiesen.

iii) Es sei

$$\mathcal{A} = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$$

und E_n sowie δ_{kl} wie oben definiert. Für das Element b_{il} von $\mathcal{A}\cdot E_n$ gilt dann

$$b_{il} = \sum_{j=1}^{n} a_{ij}\delta_{jl} = a_{il},$$

also $\mathcal{A}\cdot E_n = \mathcal{A}$. Analog wird $E_n \cdot \mathcal{B} = \mathcal{B}$ bewiesen.

iv) Oft ist nur eines der beiden Produkte $\mathcal{AB}$ bzw. $\mathcal{BA}$ überhaupt definiert.
Es ist jedoch auch nicht schwer, Gegenbeispiele zu finden, wenn beide Produkte definiert sind. Ein Beispiel, das für jeden Körper K existiert, ist

$$\mathcal{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad \mathcal{B} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Denn dann gilt

$$\mathcal{AB} = \begin{pmatrix} 0 & 1 \\ 1 & 1+1 \end{pmatrix} \neq \mathcal{BA} = \begin{pmatrix} 1 & 1 \\ 1+1 & 1 \end{pmatrix}$$

unabhängig davon, welches Element $1+1 \in K$ ist.

□

Einen wichtigen Spezialfall von Matrizen bilden die quadratischen Matrizen:

Definition 4.1.11. Eine Matrix vom Typ (m, n) heißt quadratisch, wenn $m = n$ ist.

Satz 4.1.12. *Die Menge $K^{(n,n)}$ der quadratischen Matrizen mit n Zeilen bzw. Spalten bildet bzgl. der Matrizenaddition und -multiplikation einen Ring mit Einselement E_n. Der $K^{(n,n)}$ ist für $n \geq 2$ nicht kommutativ.*

Beweis. Die Ringaxiome folgen aus Satz 4.1.10. Das Beispiel zur Nichtkommutativität lässt sich leicht für alle $n \geq 2$ ausdehnen. □

Satz 4.1.13. *Für Matrizen $\mathcal{A}$ und $\mathcal{B}$ über K gilt, sofern $\mathcal{A}+\mathcal{B}$ bzw. $\mathcal{AB}$ definiert ist:*

i) $(\mathcal{A}+\mathcal{B})^T = \mathcal{A}^T + \mathcal{B}^T$.

ii) $(\mathcal{AB})^T = \mathcal{B}^T\mathcal{A}^T$.

Beweis. i) Dies ist trivial.

ii) Es sei

$$\mathcal{A} = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} \quad \text{und} \quad \mathcal{B} = (b_{kl})_{\substack{1\le k\le n\\ 1\le l\le r}}.$$

Dann gilt $\mathcal{A}^T = (a'_{ij})_{\substack{1\le i\le n\\ 1\le j\le m}}$ mit $a'_{ij} = a_{ji}$ bzw. $\mathcal{B}^T = (b'_{kl})_{\substack{1\le k\le r\\ 1\le l\le n}}$ mit $bkl' = b_{lk}$.
Das Element c_{il} von $\mathcal{AB}$ ist

$$c_{il} = \sum_{j=1}^{n} a_{ij} b_{jl},$$

und das Element d_{il} von $\mathcal{B}^T\mathcal{A}^T$ ist

$$d_{il} = \sum_{j=1}^{n} b'_{ij}a'_{jl} = \sum_{j=1}^{n} a_{lj}b_{ji} = c_{li}.$$

Also ist $\mathcal{B}^T\mathcal{A}^T = (\mathcal{A}\mathcal{B})^T$.

□

4.2 Der Rang einer Matrix und elementare Umformungen

Definition 4.2.1. Die Maximalzahl linear unabhängiger Zeilenvektoren in einer Matrix $\mathcal{A} \in K^{(m,n)}$ heißt der Zeilenrang von $\mathcal{A}$, die Maximalzahl linear unabhängiger Spaltenvektoren von $\mathcal{A}$ heißt der Spaltenrang von $\mathcal{A}$.

Bemerkung 4.2.2. Es seien $\vec{a}_1, \ldots, \vec{a}_m \in K^n$ die Zeilenvektoren und $\vec{b}_1, \ldots, \vec{b}_n \in K^m$ die Spaltenvektoren von $\mathcal{A}$. Nach Satz 3.3.24 ist dann

$$\begin{array}{rcl} \text{Zeilenrang von } \mathcal{A} & = & \dim(\langle\, \vec{a}_1, \ldots, \vec{a}_m \,\rangle) \\ \text{Spaltenrang von } \mathcal{A} & = & \dim(\langle\, \vec{b}_1, \ldots, \vec{b}_n \,\rangle). \end{array}$$

Zur einfachen Berechnung von Zeilen- und Spaltenrang einer Matrix bedient man sich elementarer Umformungen:

Definition 4.2.3. Als elementare Umformungen einer Matrix $\mathcal{A} \in K^{(m,n)}$ bezeichnet man:

- Elementare Zeilenumformungen:
  - (Z1) Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0$ aus K,
  - (Z2) Addition einer mit $\lambda \in K$ multiplizierten Zeile zu einer anderen Zeile,
  - (Z3) Vertauschung zweier Zeilen.
- Elementare Spaltenumformungen:
  - (S1) Multiplikation einer Spalte mit einem Skalar $\lambda \neq 0$ aus K,
  - (S2) Addition einer mit $\lambda \in K$ multiplizierten Spalte zu einer anderen Spalte,
  - (S3) Vertauschung zweier Spalten.

Beispiel 4.2.4. Eine typische Umformungsfolge ist

$$\begin{pmatrix} 2 & 1 & 3 \\ 0 & 3 & 2 \\ 2 & -5 & -1 \end{pmatrix} \xrightarrow[\text{(III)-3·(II)}]{\text{(S2)}} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & -7 \\ 2 & -5 & 14 \end{pmatrix} \xrightarrow[\text{(III)-(I)}]{\text{(Z2)}} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & -7 \\ 0 & -6 & 14 \end{pmatrix}$$

$$\xrightarrow[\text{(III)+2·(II)}]{\text{(Z2)}} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & -7 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[\frac{1}{3}\text{·(II)}]{\text{(Z1)}} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & -\frac{7}{3} \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{(I)-(II)}]{\text{(Z2)}} \begin{pmatrix} 2 & 0 & \frac{7}{3} \\ 0 & 1 & -\frac{7}{3} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{(III)-}\frac{7}{6}\text{·(I)}]{\text{(S2)}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -\frac{7}{3} \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{(III)+}\frac{7}{3}\text{·(II)}]{\text{(S2)}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[\frac{1}{2}\text{·(I)}]{\text{(Z1)}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Beispiel 4.2.5. Eine andere Umformungsfolge für die gleiche Matrix ergibt

$$\begin{pmatrix} 2 & 1 & 3 \\ 0 & 3 & 2 \\ 2 & -5 & -1 \end{pmatrix} \xrightarrow{\text{(Z2)}} \begin{pmatrix} 2 & 1 & 3 \\ 0 & 3 & 2 \\ 0 & -6 & -4 \end{pmatrix} \xrightarrow{\text{(Z2)}} \begin{pmatrix} 2 & 1 & 3 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{(S2)}} \begin{pmatrix} 2 & 0 & 3 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{\text{(S2)}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{(S2)}} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{(Z1)}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{(Z1)}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Satz 4.2.6. *Durch elementare Umformungen lässt sich jede Matrix $\mathcal{A} = (a_{ij}) \in K^{(m,n)}$ in eine Matrix $D_r^{(m,n)}$ der Form*

$$D_r^{(m,n)} = \left(\begin{array}{cccc|ccc} 1 & & & & & & \\ & 1 & & & & 0 & \\ & & \ddots & & & & \\ & & & 1 & & & \\ \hline & & & & 0 & & \\ & 0 & & & & 0 & \\ & & & & & & \ddots \end{array}\right) = (c_{ij}) \quad \textit{mit} \quad c_{ij} = \begin{cases} 1, & \textit{falls } i = j \le r \\ 0 & \textit{sonst} \end{cases}$$

überführen.

Beweis. Ist $\mathcal{A} = 0_{m,n}$ die Nullmatrix, dann sind wir mit $r = 0$ fertig.
Gibt es ein $a_{ij} \neq 0$, so kann dieses durch die Umformungen (Z3) und (S3) an die Stelle von a_{11} getauscht werden. Multiplikation der 1. Zeile (oder der 1. Spalte) mit a_{ij}^{-1} liefert in der linken oberen Ecke eine 1. Die Elemente der neuen Matrix nennen wir wieder a_{ij}, jetzt also mit $a_{11} = 1$. Nun wird das a_{12}-fache der 1. Spalte von der 2. Spalte subtrahiert, dann das a_{13}-fache der 1. Spalte von der 3. Spalte usw. Nach diesen Umformungen erhält die 1. Zeile die Gestalt $(1, 0, \dots, 0)$. Jetzt subtrahiert man das a_{21}-fache der 1. Zeile von der 2. Zeile usw., womit man eine Matrix der Gestalt

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a'_{m2} & \cdots & a_{mn} \end{pmatrix} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \mathcal{A}' & \\ 0 & & & \end{array}\right)$$

erhält.
Mit der Restmatrix $\mathcal{A}' \in K^{(m-1,n-1)}$ wird analog verfahren. Offenbar haben elementare Umformungen von $\mathcal{A}'$ wegen Nullen am Rand keinen Einfluss auf die 1. Zeile sowie die 1. Spalte der ursprünglichen Matrix. Falls nicht schon $\mathcal{A}' = 0_{m-1,n-1}$ ist, geht $\mathcal{A}'$ ihrerseits in eine Matrix der obigen Gestalt über. Wir haben also die Umformungskette

$$\mathcal{A} \longrightarrow \cdots \longrightarrow \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \mathcal{A}' & \\ 0 & & & \end{array}\right) \longrightarrow \cdots \longrightarrow \left(\begin{array}{cc|ccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \mathcal{A}'' & \\ 0 & 0 & & & \end{array}\right).$$

Dieses Verfahren bricht spätestens nach $s = \min(m, n)$ Schritten ab, und der Satz ist bewiesen. □

Beispiel 4.2.7. Es gilt

$$\begin{aligned}
\mathcal{A} \quad &= \begin{pmatrix} 0 & 2 & 1 & 1 \\ 1 & 3 & -2 & 0 \\ 2 & 4 & 6 & 5 \end{pmatrix} \xrightarrow{(Z3)} \begin{pmatrix} 1 & 3 & -2 & 0 \\ 0 & 2 & 1 & 1 \\ 2 & 4 & 6 & 5 \end{pmatrix} \xrightarrow{(Z2)} \begin{pmatrix} 1 & 3 & -2 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & -2 & 10 & 5 \end{pmatrix} \\
&\xrightarrow{(S2)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & -2 & 10 & 5 \end{pmatrix} \xrightarrow{(S1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & 10 & 5 \end{pmatrix} \xrightarrow{(Z2)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 11 & 6 \end{pmatrix} \\
&\xrightarrow{(S2)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 11 & 6 \end{pmatrix} \xrightarrow{(S1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 6 \end{pmatrix} \xrightarrow{(S2)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = D_3^{(3,4)}.
\end{aligned}$$

Satz 4.2.8. *Elementare Umformungen verändern weder den Zeilen- noch den Spaltenrang einer Matrix.*

Beweis. Wegen der Analogie zwischen Zeilen und Spalten genügt es zu zeigen, dass elementare Zeilenumformungen weder den Zeilen- noch den Spaltenrang einer Matrix verändern. Wir beschränken uns auf die Operation (Z2), der Beweis für die anderen Operationen verläuft analog. Wir addieren OBdA ein Vielfaches der 1. auf die 2. Zeile:

$$\mathcal{A} = \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_m \end{pmatrix} \xrightarrow{(Z2)} \mathcal{A}' = \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 + \lambda\vec{a}_1 \\ \vdots \\ \vec{a}_m \end{pmatrix}.$$

Dann gilt $\langle\, \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \,\rangle = \langle\, \vec{a}_1, \vec{a}_2 + \lambda\vec{a}_1, \dots, \vec{a}_m \,\rangle$.
Ist nämlich $\vec{v} = \mu_1\vec{a}_1 + \mu_2\vec{a}_2 + \cdots + \mu_m\vec{a}_m$, so gilt auch

$$\vec{v} = (\mu_1 - \mu_2\lambda)\vec{a}_1 + \mu_2(\vec{a}_2 + \lambda\vec{a}_1) + \mu_3\vec{a}_3 + \cdots + \mu_m\vec{a}_m.$$

Andererseits gilt für $\vec{w} = \nu_1\vec{a}_1 + \nu_2(\vec{a}_2 + \lambda\vec{a}_1) + \cdots + \nu_m\vec{a}_m$ auch

$$\vec{w} = (\nu_1 + \lambda\nu_2)\vec{a}_1 + \nu_2\vec{a}_2 + \cdots + \nu_m\vec{a}_m.$$

Also haben $\mathcal{A}$ und $\mathcal{A}'$ nicht nur gleichen Zeilenrang, ihre Zeilenvektoren spannen sogar denselben Unterraum von K^n auf. Sind

$$\vec{b}_{l_1} = \begin{pmatrix} a_{1,l_1} \\ a_{2,l_1} \\ \vdots \\ a_{m,l_1} \end{pmatrix}, \; \dots \;, \vec{b}_{l,s} = \begin{pmatrix} a_{1,l_s} \\ a_{2,l_s} \\ \vdots \\ a_{m,l_s} \end{pmatrix}$$

beliebige Spaltenvektoren von $\mathcal{A}$, dann sind

$$\vec{b}'_{l_1} = \begin{pmatrix} a_{1,l_1} \\ a_{2,l_1} + \lambda a_{1,l_1} \\ \vdots \\ a_{m,l_1} \end{pmatrix}, \; \dots \;, \vec{b}'_{l,s} = \begin{pmatrix} a_{1,l_s} \\ a_{2,l_s} + \lambda a_{1,l_s} \\ \vdots \\ a_{m,l_s} \end{pmatrix}$$

die entsprechenden Spalten in $\mathcal{A}'$. Jede lineare Relation $\mu_1\vec{b}_{l_1} + \cdots + \mu_s\vec{b}_{l_s} = \vec{0}$ ist zur entsprechenden Relation $\mu_1\vec{b}'_{l_1} + \cdots + \mu_s\vec{b}'_{l_s} = \vec{0}$ äquivalent. Die $\vec{b}_l$ sind also genau dann linear unabhängig, wenn die entsprechenden $\vec{b}'_l$ es sind.
Daher haben $\mathcal{A}$ und $\mathcal{A}'$ auch den gleichen Spaltenrang. □

Satz 4.2.9. *Für jede Matrix $\mathcal{A} \in K^{(m,n)}$ gilt:*

$$\textit{Zeilenrang von } \mathcal{A} = \textit{Spaltenrang von } \mathcal{A}.$$

Diese Zahl heißt der <u>*Rang*</u> *von $\mathcal{A}$, geschrieben* $\mathrm{rg}(\mathcal{A})$.

Es ist $\mathrm{rg}(\mathcal{A}) = r$ die Anzahl der Einsen aus der Matrix $D_r^{(m,n)}$ aus Satz 4.2.6.

Bemerkung 4.2.10. Die Zahl r hängt also nur von $\mathcal{A}$ ab, und nicht davon, mit welcher Serie elementarer Umformungen die Matrix $D_r^{(m,n)}$ gewonnen wurde.

Beweis. (Beweis von Satz 4.2.9)
Nach Satz 4.2.8 hat $D_r^{(m,n)}$ denselben Zeilen- bzw. Spaltenrang wie $\mathcal{A}$. Der Zeilen- sowie der Spaltenrang von $D_r^{(m,n)}$ ist aber offensichtlich r. □

Beispiel 4.2.11. Nach den vorigen Beispielen ist

$$\mathrm{rg}\begin{pmatrix} 0 & 2 & 1 & 1 \\ 1 & 3 & -2 & 0 \\ 2 & 4 & 6 & 5 \end{pmatrix} = 3 \quad \text{und} \quad \mathrm{rg}\begin{pmatrix} 2 & 1 & 3 \\ 0 & 3 & 2 \\ 2 & -5 & -1 \end{pmatrix} = 2.$$

Definition 4.2.12. Eine quadratische Matrix $\mathcal{A} \in K^{(n,n)}$ heißt <u>regulär</u>, falls $\mathrm{rg}(\mathcal{A}) = n$ ist. Andernfalls, wenn also $\mathrm{rg}(\mathcal{A}) < n$ gilt, heißt $\mathcal{A}$ <u>singulär</u>.

Satz 4.2.13. *Ist $\mathcal{A} \in K^{(n,n)}$ regulär, so lässt sich $\mathcal{A}$ schon allein durch elementare Zeilenumformungen in die Einheitsmatrix E_n überführen, ebenso auch allein durch Spaltenumformungen.*

Beweis. Es sei $\mathcal{A} = (a_{ij})$ mit $1 \leq i, j \leq n$. Mindestens eines der Elemente a_{k1} der 1. Spalte muss von null verschieden sein. Durch die Umwandlungen (Z3) und (Z1) erhält man in der linken oberen Ecke eine 1.
Anwendung von (Z2) ergibt dann eine Matrix der Gestalt

$$\begin{pmatrix} 1 & a'_{12} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{pmatrix}.$$

Da die 1. und 2. Spalte linear unabhängig sind, muss hier mindestens eine der Komponenten $a'_{22}, \ldots, a'_{n2}$ in der 2. Spalte von 0 verschieden sein. Man erhält dann analog eine Matrix

$$\begin{pmatrix} 1 & 0 & a''_{13} & \cdots & a''_{1n} \\ 0 & 1 & a''_{23} & \cdots & a''_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a''_{n3} & \cdots & a''_{nn} \end{pmatrix}.$$

Nach n Schritten ist die Einheitsmatrix hergestellt. Der Beweis für Spaltenumformungen verläuft analog. □

Beispiel 4.2.14. Es gilt

$$\begin{pmatrix}0 & 2 & 3\\ 1 & 1 & 1\\ 3 & 4 & 2\end{pmatrix} \xrightarrow{(Z3)} \begin{pmatrix}1 & 1 & 1\\ 0 & 2 & 3\\ 3 & 4 & 2\end{pmatrix} \xrightarrow{(Z2)} \begin{pmatrix}1 & 1 & 1\\ 0 & 2 & 3\\ 0 & 1 & -1\end{pmatrix} \xrightarrow{(Z1)} \begin{pmatrix}1 & 1 & 1\\ 0 & 1 & \frac{3}{2}\\ 0 & 1 & -1\end{pmatrix}$$

$$\xrightarrow{(Z2)} \begin{pmatrix}1 & 0 & -\frac{1}{2}\\ 0 & 1 & \frac{3}{2}\\ 0 & 0 & -\frac{5}{2}\end{pmatrix} \xrightarrow{(Z1)} \begin{pmatrix}1 & 0 & -\frac{1}{2}\\ 0 & 1 & \frac{3}{2}\\ 0 & 0 & 1\end{pmatrix} \xrightarrow{(Z2)} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} = E_3.$$

Definition 4.2.15. Eine Matrix $\tilde{E} \in K^{(n,n)}$ heißt <u>Elementarmatrix</u>, wenn sie durch elementare Umformungen aus der Einheitsmatrix E_n hervorgeht. Wir sagen, dass $\tilde{E}_n$ zu dieser Umformung gehört.

Bemerkung 4.2.16. Es gibt demnach drei Typen von Elementarmatrizen:

- <u>Typ 1</u>: Multiplikation der i-ten Zeile mit $\lambda \neq 0$:

$$R_1 = \begin{pmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}.$$

- <u>Typ 2</u>: Addition des λ-fachen der j-Zeile zur i-ten Zeile:

$$R_2 = \begin{pmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & \lambda & & \\ & & \vdots & & \vdots & & \\ & & 0 & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}.$$

- <u>Typ 3</u>: Vertauschung der i-ten und j-ten Zeile:

$$R_3 = \begin{pmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}.$$

Die gleichen Matrizen gehören zu den Spaltenoperationen (S1), (S2) und (S3).

Wie man durch Nachrechnen leicht zeigt, gilt der folgende

Satz 4.2.17. *Entsteht $\tilde{\mathcal{A}}$ aus $\mathcal{A} \in K^{(m,n)}$ durch eine elementare Zeilen- bzw. Spaltenumformungen, dann gilt $\tilde{\mathcal{A}} = \tilde{E}\mathcal{A}$ bzw. $\tilde{\mathcal{A}} = \mathcal{A}\tilde{E}_n$ für die zur Umformung gehörende Matrix $\tilde{E}$.*

4.3 Die Inverse einer Matrix

Definition 4.3.1. Eine quadratische Matrix $\mathcal{A} \in K^{(n,n)}$ heißt invertierbar, wenn es eine Matrix $\mathcal{B} \in K^{(n,n)}$ mit $\mathcal{BA} = E_n$ oder $\mathcal{AB} = E_n$ gibt.

Satz 4.3.2. *Eine Matrix $\mathcal{A} \in K^{(n,n)}$ ist genau dann invertierbar, wenn $\mathcal{A}$ regulär ist. Dann besitzt $\mathcal{A}$ sowohl ein Linksinverses als auch ein Rechtsinverses.*

Beweis. "$\Rightarrow$":
Es sei $\mathcal{A} \in K^{(n,n)}$ invertierbar mit $\mathcal{BA} = E_n$ oder $\mathcal{AB} = E_n$. Wir unterscheiden zwei Fälle:

- Fall 1: $\mathcal{BA} = E_n$.
  Es seien $\vec{b}_1, \ldots, \vec{b}_n \in K^n$ die Spaltenvektoren von $\mathcal{A}$, und $\lambda_1, \ldots, \lambda_n \in K$ beliebig mit
  $$\lambda_1 \vec{b}_1 + \cdots + \lambda_n \vec{b}_n = \vec{0}.$$
  Dann folgt
  $$\vec{0} = \mathcal{A} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \underset{\cdot \mathcal{B} \text{ von links}}{\Longrightarrow} \vec{0} = \mathcal{BA} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = E_n \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix},$$
  also sind alle $\lambda_j = 0$. Damit sind $\vec{b}_1, \ldots, \vec{b}_n$ linear unabhängig und $\mathcal{A}$ regulär.
- Fall 2: $\mathcal{AB} = E_n$.
  Dies folgt aus Fall 1 wegen
  $$\mathcal{AB} = E_n \Rightarrow \mathcal{B}^T \mathcal{A}^T = E_n^T = E_n \Rightarrow \mathcal{A}^T \text{ regulär} \Rightarrow \mathcal{A} \text{ regulär}.$$

"$\Leftarrow$":
Es sei $\mathcal{A}$ regulär. Nach Satz 4.2.13 lässt sich $\mathcal{A}$ durch elementare Zeilenumformungen in die Einheitsmatrix E_n überführen. Nach Satz 4.2.17 gibt es daher eine Folge von Elementarmatrizen $\tilde{E}_1, \ldots, \tilde{E}_m$, so dass $\tilde{E}_m \cdots \tilde{E}_1 \cdot \mathcal{A} = E_n$ ist. Also ist $\mathcal{BA} = E_n$ für $\mathcal{B} = \tilde{E}_m \cdots \tilde{E}_1$. Analog erhält man auch eine Rechtsinverse, indem man die Elementarmatrizen aufmultipliziert, die zu den Spaltenumformungen aus Satz 4.2.13 gehören. □

Satz 4.3.3. *Die regulären Matrizen in $K^{(n,n)}$ bilden bzgl. der Multiplikation eine (für $n \geq 2$ nicht-abelsche) Gruppe. Insbesondere gilt: Zu jeder regulären Matrix $\mathcal{A} \in K^{(n,n)}$ gibt es genau eine inverse Matrix $\mathcal{A}^{-1} \in K^{(n,n)}$ mit $\mathcal{A}^{-1}\mathcal{A} = \mathcal{A}\mathcal{A}^{-1} = E_n$. Es gilt zudem $(\mathcal{A}^T)^{-1} = (\mathcal{A}^{-1})^T$.*

Beweis. Es sei $G = \{\mathcal{A} \in K^{(n,n)} \colon \mathcal{A} \text{ regulär}\}$. Es ist $E_n \in G$. Das Assoziativgesetz gilt in $(G, \cdot)$ nach Satz 4.1.10. Zu $\mathcal{A} \in G$ gibt es nach Satz 4.3.2 ein $\mathcal{B} \in K^{(n,n)}$ mit $\mathcal{AB} = E_n$, und zwar ist dann $\mathcal{B}$ auch invertierbar, also $B \in G$ nach Satz 4.3.2. Zu zeigen bleibt, dass G bzgl $\cdot$ auch abgeschlossen ist. Wähle zu $\mathcal{A}, \mathcal{B} \in G$ dann $\mathcal{A}', \mathcal{B}' \in G$ mit $\mathcal{A}'\mathcal{A} = \mathcal{B}'\mathcal{B} = E_n$. Dann ist $(\mathcal{B}'\mathcal{A}')(\mathcal{AB}) = \mathcal{B}' E_n \mathcal{B} = E_n$, weswegen $\mathcal{AB}$ nach Satz 4.3.2 invertierbar ist und somit $\mathcal{AB} \in G$ gilt. Somit ist $(G, \cdot)$ eine Gruppe, die Eindeutigkeit der Inversen $\mathcal{A}' = \mathcal{A}^{-1}$ folgt mit Satz 2.2.8. Ferner ist
$$(\mathcal{A}^{-1})^T \mathcal{A}^T = (\mathcal{A}\mathcal{A}^{-1})^T = \mathcal{E}_n^T = E_n \Rightarrow (\mathcal{A}^{-1})^T = (\mathcal{A}^T)^{-1}.$$
□

Definition 4.3.4. Mit $\mathrm{GL}(n, K) = \{\mathcal{A} \in K^{(n,n)} \colon \mathcal{A} \text{ regulär}\}$ bezeichnen wir die (multiplikative) Gruppe der invertierbaren (n, n)-Matrizen.

Satz 4.3.5. *Es sei $\mathcal{A} \in \mathrm{GL}(n, K)$.*
Man erhält $\mathcal{A}^{-1}$, indem man an E_n simultan diesselben Zeilenumformungen vornimmt, die man verwendet, um $\mathcal{A}$ gemäß Satz 4.2.13 in E_n zu überführen.

Beweis. Nach Satz 4.2.17 entsprechen diese Zeilenumformungen der Multiplikation mit einem Produkt von gewissen Elementarmatrizen $\tilde{\mathcal{C}} = \tilde{E}_m \tilde{E}_{m-1} \cdots \tilde{E}_1 \in \mathrm{GL}(n, K)$ von links, d.h. $\tilde{\mathcal{C}}\mathcal{A} = E_n$. Das Resultat derselben Umformungen an E_n ist dann $\tilde{\mathcal{C}} E_n = \tilde{\mathcal{C}}$, aber $\tilde{\mathcal{C}} = \mathcal{A}^{-1}$. □

Beispiel 4.3.6. Wir berechnen die Inverse der Matrix

$$\mathcal{A} = \begin{pmatrix} 6 & 2 & 3 \\ 4 & 5 & -2 \\ 7 & 2 & 4 \end{pmatrix}$$

mit der folgenden Umformungskette:

$\mathcal{A}$			E		
6	2	3	1	0	0
4	5	-2	0	1	0
7	2	4	0	0	1
-1	0	-1	1	0	-1
4	5	-2	0	1	0
7	2	4	0	0	1
1	0	1	-1	0	1
0	5	-6	4	1	-4
0	2	-3	7	0	-6
1	0	1	-1	0	1
0	1	0	-10	1	8
0	0	-3	27	-2	-22
3	0	3	-3	0	3
0	1	0	-10	1	8
0	0	-3	27	-2	-22
3	0	0	24	-2	-19
0	1	0	-10	1	8
0	0	-3	27	-2	-22
3	0	0	24	-2	-19
0	3	0	-30	3	24
0	0	3	-27	2	22
1	0	0			
0	1	0		$\mathcal{A}^{-1}$	
0	0	1			

mit der Inversen

$$\mathcal{A}^{-1} = \frac{1}{3} \cdot \begin{pmatrix} 24 & -2 & -19 \\ -30 & 3 & 24 \\ -27 & 2 & 22 \end{pmatrix}.$$

Man rechnet leicht die Probe

$$\frac{1}{3} \cdot \begin{pmatrix} 6 & 2 & 3 \\ 4 & 5 & -2 \\ 7 & 2 & 4 \end{pmatrix} \begin{pmatrix} 24 & -2 & -19 \\ -30 & 3 & 24 \\ -27 & 2 & 22 \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = E_3$$

nach.

Kapitel 5

Lineare Abbildungen

5.1 Definitionen und einfache Eigenschaften

Es seien stets $V, W, V', \ldots$ Vektorräume über demselben Körper K. Dieser Abschnitt befasst sich mit Abbildungen zwischen Vektorräumen, die mit den linearen Operationen verträglich sind:

Definition 5.1.1. Eine Abbildung $\varphi\colon V \to V'$ heißt linear oder ein (Vektorraum-) Homomorphismus, falls folgende Eigenschaften gelten:

(L1) $\varphi(\vec{a} + \vec{b}) = \varphi(\vec{a}) + \varphi(\vec{b})$ für alle $\vec{a}, \vec{b} \in V$.

(L2) $\varphi(\lambda\vec{a}) = \lambda\varphi(\vec{a})$ für alle $\lambda \in K$ und $\vec{a} \in V$.

Äquivalent dazu ist

(L) $\varphi(\lambda\vec{a} + \mu\vec{b}) = \lambda\varphi(\vec{a}) + \mu\varphi(\vec{b})$ für alle $\lambda, \mu \in K$ und $\vec{a}, \vec{b} \in V$.

Die Menge aller linearen Abbildungen $\varphi\colon V \to V'$ wird mit $L(V, V')$ bezeichnet.
Ein $\varphi \in L(V, V')$ heißt ein (Vektorraum)- Isomorphismus, wenn φ bijektiv ist.
Existiert ein Isomorphismus $\varphi\colon V \to V'$, so sagen wir, V ist isomorph zu V', geschrieben als $V \cong V'$.
Ein $\varphi \in L(V, V)$, also mit $V = V'$, heißt ein Endomorphismus, bzw. im Falle der Bijektivität ein Automorphismus von V.

Beispiel 5.1.2. Es gibt stets den trivialen Homomorphismus $\varphi_0(\vec{a}) = \vec{0}' \in V'$ für alle $\vec{a} \in V$.

Beispiel 5.1.3. Der Endomorphismus $\varphi\colon V \to V$, $\varphi(\vec{a}) = \lambda\vec{a}$ für festes $\lambda \in K$ ist trivial für $\lambda = 0$, und ein Automorphismus für $\lambda \neq 0$, da er wegen $\varphi(\vec{a}) = \varphi(\vec{b}) \Leftrightarrow \vec{a} = \vec{b}$ injektiv ist, und surjektiv wegen $\varphi(\lambda^{-1}\vec{a}) = \vec{a}$.

Beispiel 5.1.4. Für $V = \mathbb{R}^2$ ist $\varphi(x, y) = (\lambda x, \mu y)$ mit $\lambda, \mu \in \mathbb{R}\backslash\{0\}$ die sogenannte Eulerabbildung. Die Ebene wird in x-Richtung um den Faktor λ und in y-Richtung um den Faktor μ gestreckt. Offenbar ist φ ein Automorphismus.

Beispiel 5.1.5. Es sei $V = \mathbb{R}^2$. Die Projektion auf die x-Achse $\varphi\colon V \to V$, $\varphi(x, y) = (x, 0)$ ist weder injektiv noch surjektiv, also kein Automorphismus.

Beispiel 5.1.6. Für $V = \mathbb{R}^2$ heißt der Automorphismus $\varphi\colon V \to V$, $\varphi(x, y) = (x + \lambda y, y)$ für festes $\lambda \in \mathbb{R}$ Scherung.

Beispiel 5.1.7. Der Homomorphismus $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ mit

$$\varphi(x_1, \ldots, x_n) = \sum_{j=1}^{n} a_j x_j$$

für feste $a_1, \ldots, a_n \in \mathbb{R}$ ist eine sogenannte <u>Linearform</u>.

Beispiel 5.1.8. Es sei $F_0 = \{p\colon \mathbb{R} \to \mathbb{R},\ p \text{ Polynom}\}$. Für $p \in F_0$ mit

$$p(x) = \sum_{\nu=0}^{n} a_\nu x^\nu$$

sei $\varphi(p) = p'$ die Ableitung mit

$$p'(x) = \sum_{\nu=1}^{n} a_\nu \cdot \nu \cdot x^{\nu-1}.$$

Dann ist φ ein Endomorphismus von F_0. Er ist surjektiv, denn für

$$\tilde{p}(x) = \sum_{\nu=0}^{n} \frac{a_\nu}{\nu+1} x^{\nu+1}$$

ist $\varphi(\tilde{p}) = p$. Er ist nicht injektiv, denn es gilt etwa $\varphi(p_0) = 0$ für jedes konstante Polynom $p_0(x) = a_0$. Die Abbildung φ ist ein Beispiel für einen linearen <u>Differentialoperator</u>.

Satz 5.1.9. *Für Vektorräume V, V' und V'' gilt:*

i) Ist $\varphi \in L(V, V')$ und $\psi \in L(V', V'')$, so ist $\psi \circ \varphi \in L(V, V'')$.

ii) Ist $\varphi\colon V \to V'$ ein Isomorphismus, so auch die inverse Abbildung $\varphi^{-1}\colon V' \to V$.

Beweis. i) Es seien $\vec{a}, \vec{b} \in V$ und $\lambda, \mu \in K$. Dann gilt

$$(\psi \circ \varphi)(\lambda\vec{a} + \mu\vec{b}) = \psi(\lambda\varphi(\vec{a}) + \mu\varphi(\vec{b})) = \lambda(\psi \circ \varphi)(\vec{a}) + \mu(\psi \circ \varphi)(\vec{b}).$$

ii) Da φ^{-1} offenbar bijektiv ist, müssen wir nur (L) zeigen. Es seien $\vec{a}', \vec{b}' \in V'$ und $\lambda, \mu \in K$ mit $\vec{a}' = \varphi(\vec{a})$ und $\vec{b}' = \varphi(\vec{b})$ für $\vec{a}, \vec{b} \in V$. Dann folgt

$$\varphi^{-1}(\lambda\vec{a}' + \mu\vec{b}') = \varphi^{-1}(\lambda\varphi(\vec{a}) + \mu\varphi(\vec{b})) = \varphi^{-1}(\varphi(\lambda\vec{a} + \mu\vec{b})) = \lambda\vec{a} + \mu\vec{b} = \lambda\varphi^{-1}(\vec{a}') + \mu\varphi^{-1}(\vec{b}').$$

□

Satz 5.1.10. *Für $\varphi, \psi \in L(V, V')$ und $\lambda \in K$ seien $\varphi + \psi$ und $\lambda\varphi \in L(V, V')$ werteweise durch*

$$(\varphi + \psi)(\vec{a}) := \varphi(\vec{a}) + \psi(\vec{a}) \quad \text{und} \quad (\lambda\varphi)(\vec{a}) := \lambda\varphi(\vec{a})$$

für alle $\vec{a} \in V$ definiert.
Mit diesen Operationen wird $L(V, V')$ zu einem Vektorraum über K. Der Nullvektor ist der triviale Homomorphismus $\varphi_0(\vec{a}) = \vec{0}' \in V'$ für alle $\vec{a} \in V$.

Beweis. Durch Nachrechnen sieht man leicht die Gültigkeit der Regel (L) für $\varphi + \psi$ sowie für $\lambda\varphi$. Auch die Vektorraumaxiome werden leicht nachgeprüft. □

Satz 5.1.11. *Mit der oben definierten Addition sowie der Komposition von Abbildungen* $\circ$ *als Multiplikation ist* $(L(V,V),+,\circ)$ *ein Ring mit Einselement, der* *Endomorphismenring* *von* V*. Einselement ist die* *Identität* $\mathrm{id}_V : \vec{a} \to \vec{a}$ *für alle* $\vec{a} \in V$*.*

Beweis. Nach Satz 5.1.10 ist $(L(V,V),+)$ eine abelsche Gruppe. Nach Satz 5.1.9 ist $\circ$ eine Verknüpfung auf $L(V,V)$. Das Assoziativgesetz der Multiplikation gilt, da es allgemein für die Komposition von Abbildungen gilt. Die Distributivgesetze gelten ebenfalls, denn für $\psi,\varphi,\chi \in L(V,V)$ und alle $\vec{a} \in V$ gilt

$$(\varphi\circ(\psi+\chi))(\vec{a}) = \varphi(\psi(\vec{a})+\chi(\vec{a})) = \varphi(\psi(\vec{a}))+\varphi(\chi(\vec{a})) = (\varphi\circ\psi)(\vec{a})+(\varphi\circ\chi)(\vec{a}) = (\varphi\circ\psi+\varphi\circ\chi)(\vec{a}).$$

Ebenso ist

$$((\psi+\chi)\circ\varphi)(\vec{a}) = \psi(\varphi(\vec{a}))+\chi(\varphi(\vec{a})) = (\psi\circ\varphi+\chi\circ\varphi)(\vec{a}).$$

Für das Einselement gilt $id_V \circ \varphi = \varphi \circ id_V = \varphi$ für alle $\varphi \in L(V,V)$. □

Satz 5.1.12. *Die Automorphismen von* V *bilden bzgl.* $\circ$ *eine Gruppe, die* *lineare Gruppe* *von* V*, geschrieben* $\mathrm{GL}(V)$*.*

Beweis. Es ist $\circ$ eine Verknüpfung auf $\mathrm{GL}(V)$, denn mit φ,ψ ist auch $\psi\circ\varphi\colon V \to V$ bijektiv, und die Linearität folgt nach Satz 5.1.9. Das Assoziativgesetz folgt nach Satz 5.1.11, das Einselement ist id_V, und das Inverse von φ ist die Umkehrabbildung φ^{-1}. □

5.2 Kern und Bild

Satz 5.2.1. *Für* $\varphi \in L(V,V')$ *gilt:*

i) $\varphi(\vec{0}) = \vec{0}'$.

ii) Sind $\vec{a}_1,\dots,\vec{a}_n \in V$ *linear abhängig, so auch* $\varphi(\vec{a}_1),\dots,\varphi(\vec{a}_n) \in V'$*.*

iii) Sind $\varphi(\vec{a}_1),\dots,\varphi(\vec{a}_n) \in V'$ *linear unabhängig, so auch* $\vec{a}_1,\dots,\vec{a}_n \in V$*.*

Beweis. i) Es gilt $\varphi(\vec{0}) = \varphi(\vec{0}+\vec{0}) = \varphi(\vec{0})+\varphi(\vec{0})$, woraus $\varphi(\vec{0}) = \vec{0}'$ folgt.

ii) Aus $\lambda_1\vec{a}_1+\cdots+\lambda_n\vec{a}_n = \vec{0}$ folgt

$$\lambda_1\varphi(\vec{a}_1)+\cdots+\lambda_n\varphi(\vec{a}_n) = \varphi(\lambda_1\vec{a}_1+\cdots+\lambda_n\vec{a}_n) = \vec{0}'$$

nach i).

iii) Dies folgt direkt aus ii).

□

Satz 5.2.2. *Es sei* $\psi \in L(V,V')$*. Dann gilt*

i) Ist $M \subseteq V$*, so gilt* $\varphi(\langle M \rangle) = \langle \varphi(M) \rangle$*. Insbesondere ist das Bild eines Unterraums* $U \subseteq V$ *ein Unterraum von* V'*.*

ii) Für einen Unterraum $U' \subseteq V'$ *ist* $\varphi^{-1}(U') = \{\vec{a} \in V : \varphi(\vec{a}) \in U'\}$ *ein Unterraum von* V*.*

Beweis. i) Wir unterscheiden zwei Fälle:

Fall 1: Es sei $M = \emptyset$. Dann ist $\langle M \rangle = \{\vec{0}\}$ und $\varphi(\langle M \rangle) = \{\vec{0}'\} = \langle \emptyset \rangle = \langle \varphi(M) \rangle$.

Fall 2: Es sei $M \neq \emptyset$. Wir zeigen zunächst $\varphi(\langle M \rangle) \subseteq \langle \varphi(M) \rangle$:
Ist $\vec{a} = \lambda_1\vec{a}_1 + \cdots + \lambda_n\vec{a}_n \in \langle M \rangle$, so folgt direkt $\varphi(\vec{a}) = \lambda_1\varphi(\vec{a}_1) + \cdots + \lambda_n\varphi(\vec{a}_n) \in \langle \varphi(M) \rangle$.
Es ist nun noch $\langle \varphi(M) \rangle \subseteq \varphi(\langle M \rangle)$ zeigen. Es sei dazu $\vec{b} \in \langle \varphi(M) \rangle$, d.h. $\vec{b} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$ mit $\vec{b}_j \in \varphi(M)$. Dann gibt es $\vec{a}_j \in M$ mit $1 \leq j \leq n$ und $\varphi(\vec{a}_j) = \vec{b}_j$. Es gibt $\lambda_j \in K$ mit $\vec{a} = \lambda_1\vec{a}_1 + \cdots + \lambda_n\vec{a}_n \in \langle M \rangle$ und damit

$$\varphi(a) = \lambda_1\varphi(\vec{a}_1) + \cdots + \lambda_n\varphi(\vec{a}_n) = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n = \vec{b},$$

also $\vec{b} \in \varphi(\langle M \rangle)$. Damit ist insgesamt $\langle \varphi(M) \rangle = \varphi(\langle M \rangle)$ gezeigt.

ii) Es ist $\varphi^{-1}(U') \neq \emptyset$, da $\varphi(\vec{0}) = \vec{0}' \in U'$ ist, also $\vec{0} \in \varphi^{-1}(U')$. Es seien $\vec{a}, \vec{b} \in \varphi^{-1}(U')$ gegeben, also $\varphi(\vec{a}), \varphi(\vec{b}) \in U'$. Ferner seien $\lambda, \mu \in K$. Dann gilt

$$\varphi(\lambda\vec{a} + \mu\vec{b}) = \lambda\varphi(\vec{a}) + \mu\varphi(\vec{b}) \in U',$$

d.h. $\lambda\vec{a} + \mu\vec{b} \in \varphi^{-1}(U')$.

□

Satz 5.2.3. *Es sei $\varphi \in L(V, V')$ und U ein Unterraum von V. Dann ist $\dim(\varphi(U)) \leq \dim(U)$. Insbesondere folgt $\dim(V) = \dim(V')$ aus $V \cong V'$.*

Beweis. Es sei OBdA $\dim(U) < \infty$. Dnn besitzt U nach Satz 3.3.20 eine Basis B, also $U = \langle B \rangle$. Nach Satz 5.2.2 gilt $\varphi(U) = \varphi(\langle B \rangle) = \langle \varphi(B) \rangle$, womit $\varphi(B)$ ein Erzeugendensystem von $\varphi(U)$ ist. Dann ist $\dim(\varphi(U)) \leq |\varphi(B)| \leq |B| = \dim(U)$. □

Definition 5.2.4. Es sei $\varphi \in L(V, V')$. Es heißt $\mathrm{Bild}(\varphi) = \varphi(V) = \{\varphi(\vec{a}) \colon \vec{a} \in V\}$ das Bild von φ.
Die Menge $\mathrm{Kern}(\varphi) = \varphi^{-1}(\{\vec{0}'\}) = \{\vec{a} \in V \colon \varphi(\vec{a}) = \vec{0}'\}$ heißt Kern von φ.
Nach Satz 5.2.2 sind dies Unterräume von V' bzw. V.
Ferner heißt $\mathrm{rg}(\varphi) = \dim(\mathrm{Bild}(\varphi))$ der Rang und $\mathrm{def}(\varphi) = \dim(\mathrm{Kern}(\varphi))$ der Defekt von φ.

Beispiel 5.2.5. Wir betrachten die Projektion aus Beispiel 5.1.5.
Es ist $V = \mathbb{R}^2$ und $\varphi(x, y) = (x, 0)$. Dann ist $\mathrm{Bild}(\varphi) = \{(x, 0) \colon x \in \mathbb{R}\}$, d.h. $\mathrm{rg}(\varphi) = 1$. Andererseits ist $\mathrm{Kern}(\varphi) = \varphi^{-1}(\{(0,0)\}) = \{\vec{a} \in V \colon \varphi(\vec{a}) = (0,0)\} = \{(0, y) \colon y \in \mathbb{R}\}$, und somit $\mathrm{def}(\varphi) = 1$.

Satz 5.2.6. *Für $\varphi \in L(V, V')$ sind folgende Aussagen äquivalent:*

i) Die Abbildung φ ist injektiv.

ii) Es gilt $\mathrm{Kern}(\varphi) = \{\vec{0}\}$, d.h. $\mathrm{def}(\varphi) = 0$.

iii) Sind $\vec{a}_1, \ldots, \vec{a}_n \in V$ linear unabhängig, so auch $\varphi(\vec{a}_1), \ldots, \varphi(\vec{a}_n)$.

Beweis. i) $\Rightarrow$ *ii)*:
Ist φ injektiv, so folgt aus $\varphi(\vec{a}) = \vec{0}' = \varphi(\vec{0})$ schon $\vec{a} = \vec{0}$.

ii) $\Rightarrow$ *iii)*:
Es seien $\mathrm{Kern}(\varphi) = \{\vec{0}\}$ und $\vec{a}_1, \ldots, \vec{a}_n \in V$ linear unabhängig. Aus $\lambda_1\varphi(\vec{a}_1) + \cdots + \lambda_n\varphi(\vec{a}_n) = \vec{0}'$ folgt dann $\varphi(\lambda_1\vec{a}_1 + \cdots + \lambda_n\vec{a}_n) = \vec{0}'$, d.h.

$$\lambda_1\vec{a}_1 + \cdots + \lambda_n\vec{a}_n \in \mathrm{Kern}\, K(\varphi) = \{\vec{0}'\} \Rightarrow \lambda_1 = \cdots = \lambda_n = 0,$$

da $\vec{a}_1, \ldots, \vec{a}_n$ linear unabhängig sind.

iii) $\Rightarrow i)$:
Es seien $\vec{a}_1, \vec{a}_2 \in V$ mit $\vec{a}_1 \neq \vec{a}_2$. Dann ist $\vec{a}_1 - \vec{a}_2 \neq \vec{0}$ linear unabhängig, also ist nach iii) auch $\varphi(\vec{a}_1 - \vec{a}_2) = \varphi(\vec{a}_1) - \varphi(\vec{a}_2)$ linear unabhängig, und damit $\varphi(\vec{a}_1) \neq \varphi(\vec{a}_2)$.

□

Satz 5.2.7. *(Rangformel)*
Es sei $\varphi \in L(V, V')$ *mit* $\dim(V) < \infty$. *Dann gilt* $\mathrm{rg}(\varphi) + \mathrm{def}(\varphi) = \dim(V)$, *oder ausführlich*

$$\dim(\mathrm{Bild}(\varphi)) + \dim(\mathrm{Kern}(\varphi)) = \dim(V).$$

Die beiden Extremfälle dieser Gleichheit sind

$$\begin{array}{rcl} \mathrm{def}(\varphi) = 0 & \Leftrightarrow & \dim(\mathrm{Bild}(\varphi)) = \dim(V) \\ \mathrm{rg}(\varphi) = 0 & \Leftrightarrow & \mathrm{Kern}(\varphi) = V. \end{array}$$

Beweis. Wir wählen eine Basis $\vec{a}_1, \ldots, \vec{a}_r$ von $\mathrm{Kern}(\varphi)$ und ergänzen sie nach dem Austauschsatz von Steinitz (3.3.23) zu einer Basis $\vec{a}_1, \ldots, \vec{a}_r, \vec{b}_1, \ldots, \vec{b}_s$ von V. Wir zeigen im folgenden, dass

$$\varphi(V) = \left\langle\left\langle\ \varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_s)\ \right\rangle\right\rangle$$

ist. Daraus folgt dann $\mathrm{rg}(\varphi) + \mathrm{def}(\varphi) = r + s = \dim(V)$, also die Behauptung.

- $\underline{\varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_s) \text{ sind linear unabhängig:}}$
  Es sei $\vec{0} = \lambda_1\varphi(\vec{b}_1) + \cdots + \lambda_s\varphi(\vec{b}_s) = \varphi(\lambda_1\vec{b}_1 + \cdots + \lambda_s\vec{b}_s)$ mit $\lambda_i \in K$. Es folgt
  $$\lambda_1\vec{b}_1 + \cdots + \lambda_s\vec{b}_s \in \mathrm{Kern}(\varphi) \Rightarrow \lambda_1\vec{b}_1 + \cdots + \lambda_s\vec{b}_s = \mu_1\vec{a}_1 + \cdots + \mu_r\vec{a}_r$$
  mit $\mu_j \in K$. Dann ist $\mu_1\vec{a}_1 + \cdots + \mu_r\vec{a}_r + (-\lambda_1)\vec{b}_1 + \cdots + (-\lambda_s)\vec{b}_s = \vec{0}$, woraus
  $$\lambda_1 = \cdots = \lambda_s = \mu_1 = \cdots = \mu_r = 0$$
  folgt, da $\vec{a}_1, \ldots, \vec{a}_r, \vec{b}_1, \ldots, \vec{b}_s$ linear unabhängig sind.

- $\underline{\varphi(V) = \langle\ \varphi(\vec{b}_1), \ldots, \varphi(b_s)\ \rangle:}$
  Es sei dazu $\vec{v}' \in \varphi(V)$, etwa $\vec{v}' = \varphi(\vec{v})$ für ein $\vec{v} \in V$. Dann ist
  $$\vec{v} = \alpha\vec{a}_1 + \cdots + \alpha_r\vec{a}_r + \beta_1\vec{b}_1 + \cdots + \beta_s\vec{b}_S$$
  mit $\alpha_i, \beta_j \in K$. Es folgt
  $$\begin{array}{rcl} \vec{v}' & = & \varphi(\vec{v}) = \alpha_1\varphi(\vec{a}_1) + \cdots + \alpha_r\varphi(\vec{a}_r) + \beta_1\varphi(\vec{b}_1) + \cdots + \beta_s\varphi(\vec{b}_s) \\ & = & \beta_1\varphi(\vec{b}_1) + \cdots + \beta_s\varphi(\vec{b}_s) \in \left\langle\ \varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_s)\ \right\rangle, \end{array}$$
  denn es ist $\varphi(\vec{a}_i) = \vec{0}$ für die $\vec{a}_i \in \mathrm{Kern}(\varphi)$.

□

Satz 5.2.8. *Es sei* $\varphi \in L(V, V')$. *Dann gilt*

i) Ist $\dim(V) < \infty$, *so gilt* φ *injektiv* $\Leftrightarrow \mathrm{rg}(\varphi) = \dim(V)$.

ii) Ist $\dim(V) = \dim(V')$, *so gilt* φ *injektiv* $\Leftrightarrow \varphi$ *surjektiv.*

Beweis. Es gelten die Äquivalenzen

$$\begin{array}{lcl} \varphi \text{ injektiv} & \underset{\text{Satz 5.2.6}}{\Leftrightarrow} & \mathrm{def}(\varphi) = 0 \underset{\text{Satz 5.2.7}}{\Leftrightarrow} \mathrm{rg}(\varphi) = \dim(V) \quad \text{sowie} \\ \varphi \text{ injektiv} & \underset{\text{i)}}{\Leftrightarrow} & \dim(\mathrm{Bild}(\varphi)) = \dim(V) = \dim(V') \underset{\text{Satz 3.3.26 ii)}}{\Leftrightarrow} \varphi(V) = V'. \end{array}$$

□

5.3 Lineare Fortsetzung

Wir kommen nun zur Frage der Beschreibung linearer Abbildungen:

Satz 5.3.1. *Es sei $V = \langle\langle\, \vec{a}_1, \dots, \vec{a}_n \,\rangle\rangle$, und es seien $\vec{b}_1, \dots, \vec{b}_n \in V'$ beliebige Vektoren. Dann gibt es genau eine lineare Abbildung $\varphi \in L(V, V')$ mit $\varphi(\vec{a}_j) = \vec{b}_j$ für $j = 1 \dots n$, d.h. eine lineare Abbildung ist schon völlig festgelegt, wenn ihre Werte auf einer Basis von V bekannt sind. Andererseits können diese Werte beliebig vorgeschrieben werden.*

Beweis.
- Existenz:
  Zu $\vec{v} \in V$ existieren eindeutig bestimmte $\lambda_1, \dots, \lambda_n \in K$ mit $\vec{v} = \lambda_1\vec{a}_1 + \dots + \lambda_n\vec{a}_n$. Die Abbildung φ wird dann durch $\varphi(\vec{v}) = \lambda_1\vec{b}_1 + \dots + \lambda_n\vec{b}_n$ definiert. Es bleibt, zu zeigen, dass φ linear ist. Dazu seien

  $$\vec{v} = \sum_{j=1}^{n} \lambda_j \vec{a}_j \quad \text{und} \quad \vec{w} = \sum_{j=1}^{n} \mu_j \vec{a}_j$$

  aus V beliebig und $\alpha, \beta \in K$. Dann gilt

  $$\begin{aligned} \varphi(\alpha\vec{v} + \beta\vec{w}) &= \varphi\left(\sum_{j=1}^{n}(\alpha\lambda_j + \beta\mu_j)\vec{a}_j\right) = \sum_{j=1}^{n}(\alpha\lambda_j + \beta\mu_j)\vec{b}_j = \alpha \cdot \sum_{j=1}^{n}\lambda_j\vec{b}_j + \beta \cdot \sum_{j=1}^{n}\mu_j\vec{b}_j \\ &= \alpha\varphi(\vec{v}) + \beta\varphi(\vec{w}). \end{aligned}$$

- Eindeutigkeit:
  Es sei $\psi \in L(V, V')$ mit $\psi(\vec{a}_j) = \vec{b}_j$ für $j = 1 \dots n$. Für jedes $\vec{v} = \lambda_1\vec{a}_1 + \dots + \lambda_n\vec{a}_n$ muss dann

  $$\psi(\vec{v}) = \lambda_1\psi(\vec{a}_1) + \dots + \lambda_n\psi(\vec{a}_n) = \lambda_1\vec{b}_1 + \dots + \lambda_n\vec{b}_n = \varphi(\vec{v})$$

  gelten, also $\psi = \varphi$.

□

Definition 5.3.2. Die durch die Zuordnung $\vec{a}_j \to \vec{b}_j$ nach Satz 5.3.1 eindeutig festgelegte lineare Abbildung $\varphi \in L(V, V')$ heißt lineare Fortsetzung dieser Zuordnung.

Satz 5.3.3. *Es seien V und V' endlichdimensionale Vektorräume über K.*
Dann gilt genau dann $V \cong V'$, wenn $\dim V = \dim V'$ gilt. Insbesondere ist also jeder n-dimensionale Vektorraum über K zum Standardraum K^n isomorph.

Beweis. "$\Rightarrow$":
Dies folgt aus Satz 5.2.3.
"$\Leftarrow$":
Es sei also $\dim V = \dim V' = n \in \mathbb{N}$. Für $n = 0$ ist die Aussage trivial. Andernfalls sei $V = \langle\langle\, \vec{a}_1, \dots, \vec{a}_n \,\rangle\rangle$ und $V' = \left\langle\left\langle\, \vec{b}_1, \dots, \vec{b}_n \,\right\rangle\right\rangle$. Wir definieren $\varphi \in L(V, V')$ als die lineare Fortsetzung der Zuordnung $\vec{a}_j \to \vec{b}_j$ für $j = 1 \dots n$ und müssen nur noch die Bijektivität von φ zeigen. Ist $\vec{b} = \lambda_1\vec{b}_1 + \dots + \lambda_n\vec{b}_n \in V'$ beliebig, so ist $\vec{b} = \varphi(\vec{a})$ für $\vec{a} = \lambda_1\vec{a}_1 + \dots + \lambda_n\vec{a}_n \in V$ nach der Definition von φ. Nach Satz 5.2.8 ii) ist φ auch injektiv und damit ein Isomorphismus. □

Satz 5.3.4. *Es seien $\varphi \in L(V, V')$ und $\psi \in L(V', V'')$, also $\psi \circ \varphi \in L(V, V'')$. Dann gilt*

$$\mathrm{rg}(\varphi) + \mathrm{rg}(\psi) - \dim V' \le \mathrm{rg}(\psi \circ \varphi) \le \min\{\mathrm{rg}(\varphi), \mathrm{rg}(\psi)\}.$$

Beweis. Die rechte Ungleichung folgt aus

$$\operatorname{rg}(\psi \circ \varphi) = \dim(\operatorname{Bild}(\psi \circ \varphi)) = \dim(\psi(\varphi(V))) \leq \begin{cases} \dim \varphi(V) = \operatorname{rg}(\varphi) \\ \dim \psi(V') = \operatorname{rg}(\psi). \end{cases}$$

Für die linke Ungleichung betrachten wir die Abbildung $\psi^* = \psi|_{\varphi(V)}$, die Beschränkung von ψ auf den Unterraum $\varphi(V)$, also $\psi^* \colon \varphi(V) \to V''$. Dann gilt

$$\begin{aligned} \operatorname{rg}(\psi \circ \varphi) &= \dim \psi(\varphi(V)) = \dim \psi^*(V) = \operatorname{rg}(\psi^*) = \dim \varphi(V) - \operatorname{def}(\psi^*) \geq \dim \varphi(V) - \operatorname{def}(\psi) \\ &= \operatorname{rg}(\varphi) - (\dim V' - \operatorname{rg}(\psi)) \end{aligned}$$

nach Satz 5.2.7. □

5.4 Lineare Abbildungen und Matrizen

Alle betrachteten Vektorräume seien endlichdimensional. Es sei $\varphi \in L(V, V')$ eine lineare Abbildung von V nach V' und $\mathcal{B}$ bzw. $\mathcal{B}'$ Basen von V bzw. V'. Nach Satz 5.3.1 kann φ durch die Bilder der Vektoren aus $\mathcal{B}$ eindeutig beschrieben werden. Jedes dieser Bilder kann wiederum eindeutig als Linearkombination der Vektoren von $\mathcal{B}'$ ausgedrückt werden. Die Koeffizienten dieser Linearkombinationen sammeln wir nun in einer Matrix, die der Abbildung φ zugeordnet wird.

Definition 5.4.1. Es seien $\dim V = n < \infty$ sowie $\dim V' = m < \infty$, $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis von V sowie $\mathcal{B}' = \{\vec{b}'_1, \dots, \vec{b}'_m\}$ eine Basis von V'. Ferner sei $\varphi \in L(V, V')$. Wegen $\varphi(\vec{b}_l) \in V'$ gibt es eindeutig bestimmte $\alpha_{kl} \in K$ mit

$$\varphi(\vec{b}_l) = \alpha_{1l}\vec{b}'_1 + \cdots + \alpha_{ml}\vec{b}'_m = \sum_{k=1}^{m} \alpha_{kl}\vec{b}'_k \tag{$*$}$$

mit $l = 1, \dots, n$.
Es sei $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = (\alpha_{kl})$ mit $1 \leq k \leq m$ und $1 \leq l \leq n$. Man nennt $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ die φ bzgl. der Basen $\mathcal{B}$ und $\mathcal{B}'$ <u>zugeordnete Matrix</u> oder auch <u>Darstellungsmatrix</u> bzw. <u>Abbildungsmatrix</u> von φ bzgl. $\mathcal{B}$ und $\mathcal{B}'$.

Satz 5.4.2. *Mit den obigen Bezeichnungen sei*

$$\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n \in V \quad \textit{bzw.} \quad \varphi(\vec{v}) = \lambda'_1\vec{b}'_1 + \cdots + \lambda'_m\vec{b}'_m \in V'.$$

Dann gilt

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \tag{$**$}$$

Definition 5.4.3. Wir nennen $\lambda_1, \dots, \lambda_n$ die <u>Koordinaten</u> von $\vec{v}$ bzgl. $\mathcal{B}$. Entsprechend sind $\lambda'_1, \dots, \lambda'_m$ die Koordinaten von $\varphi(\vec{v})$ bzgl. $\mathcal{B}'$.

Beweis. (Beweis von Satz 5.4.2)
Es ist

$$\varphi(\vec{v}) = \sum_{l=1}^{n} \lambda_l \varphi(\vec{b}_l) \underset{(*)}{=} \sum_{l=1}^{n} \lambda_l \sum_{k=1}^{m} \alpha_{kl}\vec{b}'_k,$$

also

$$\lambda'_k = \sum_{l=1}^{n} \alpha_{kl}\lambda_l$$

für $k = 1, \ldots, m$, d.h.

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = \begin{pmatrix} \alpha_{11}\lambda_1 & + & \cdots & + & \alpha_{1n}\lambda_n \\ & & \vdots & & \\ \alpha_{m1}\lambda_1 & + & \cdots & + & \alpha_{mn}\lambda_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

□

Bemerkung 5.4.4. Die Matrix $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ lässt sich folgendermaßen einfach in Worten beschreiben: Nach $(*)$ besteht die l-te Spalte von $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ aus den Koordinaten des Bildes $\varphi(\vec{b}_l)$ des l-ten Basisvektors.

Wir betrachten jetzt die Matrizen, die einigen der im letzten Paragraphen als Beispiele aufgeführten linearen Abbildungen zugeordnet sind, sowie ein paar andere Beispiele.

Beispiel 5.4.5. Es sei $\dim V = n$ und $\dim V' = m$ sowie $\varphi_0\colon V \to V'$ der triviale Homomorphismus. Dann ist bzgl. beliebiger Basen $\mathcal{B}$ und $\mathcal{B}'$ von V und V' stets $\mathcal{M}(\varphi_0; \mathcal{B}', \mathcal{B}) = 0^{(m,n)}$ die Nullmatrix.

Beispiel 5.4.6. Es sei $\varphi\colon V \to V$ mit $\varphi(\vec{a}) = \lambda\vec{a}$ für festes $\lambda \in K$ und $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine beliebige Basis von V. Die l-te Spalte von $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ besteht dann aus den Koordinaten $\varphi(\vec{b}_l) = \lambda\vec{b}_l$ bzgl. $\mathcal{B}$, also

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{pmatrix} = \lambda \cdot E_n.$$

Beispiel 5.4.7. Es sei $V = \mathbb{R}^2$ und $\varphi\colon V \to V$ mit $\varphi(x, y) = (\lambda x, \mu y)$ die Eulerabbildung. Es sei $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ mit $\vec{e}_1 = (1, 0)^T$ und $\vec{e}_2 = (0, 1)^T$ die Standardbasis. Dann gilt $\varphi(\vec{e}_1) = \lambda\vec{e}_1$ sowie $\varphi(\vec{e}_2) = \mu\vec{e}_2$, also

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Die Darstellungsmatrix hängt im allgemeinen von der Wahl der Basen ab. Ist etwa $\tilde{\mathcal{B}} = \{\vec{b}_1, \vec{b}_2\}$ mit $\vec{b}_1 = (1, 1)^T$ und $\vec{b}_2 = (1, -1)^T$, so ist

$$\begin{aligned} \varphi(\vec{b}_1) &= (\lambda, \mu) = \frac{\lambda + \mu}{2} \cdot \vec{b}_1 + \frac{\lambda - \mu}{2} \cdot \vec{b}_2 \\ \varphi(\vec{b}_2) &= (\lambda, -\mu) = \frac{\lambda - \mu}{2} \cdot \vec{b}_1 + \frac{\lambda + \mu}{2} \cdot \vec{b}_2 \end{aligned}$$

und damit

$$\mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \begin{pmatrix} \frac{\lambda+\mu}{2} & \frac{\lambda-\mu}{2} \\ \frac{\lambda-\mu}{2} & \frac{\lambda+\mu}{2} \end{pmatrix}.$$

Eine der wichtigen Aufgaben der linearen Algebra ist es, zu einer gegebenen Abbildung φ Basen zu finden, bzgl. denen die Darstellungsmatrix besonders einfach wird.

Beispiel 5.4.8. Es sei $V = \mathbb{R}^3$ mit der Standardbasis $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ versehen und $\varphi \in L(V, V)$ die lineare Abbildung mit bzgl. $\mathcal{B}$ zugeordneter Matrix

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix}.$$

Wir betrachten die neue Basis $\tilde{\mathcal{B}} = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Es ist

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ -8 \end{pmatrix} \quad \text{und} \quad \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ -3 \\ 6 \end{pmatrix},$$

also kurz $\varphi(\vec{b}_1) = \vec{b}_1$, $\varphi(\vec{b}_2) = 2\vec{b}_2$ und $\varphi(\vec{b}_3) = 3\vec{b}_3$. Bezüglich der Basis $\tilde{\mathcal{B}}$ hat φ die einfache Diagonalmatrix

$$\mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

als Darstellungsmatrix. Die Abbildung φ streckt daher den $\mathbb{R}^3$ in den Richtungen von $\vec{b}_1$, $\vec{b}_2$ und $\vec{b}_3$ um die Faktoren 1, 2 und 3. Man nennt $\vec{b}_1$, $\vec{b}_2$ und $\vec{b}_3$ die Eigenvektoren von φ mit den zugehörigen Eigenwerten $1, 2, 3$.

Beispiel 5.4.9. Ist $\varphi \in L(V, V')$ ein Isomorphismus, so lassen sich die Basen $\mathcal{B}$ und $\mathcal{B}'$ von V und V' stets so wählen, dass $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = E_n$ für ein $n \in \mathbb{N}_0$ ist.
Nach Satz 5.3.3 ist nämlich $\dim V = \dim V'$, und ist $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V, so sind nach Satz 5.2.6 die Vektoren $\varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n)$ linear unabhängig, d.h. nach Satz 3.3.20 ist dann $\mathcal{B}' = \{\varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n)\}$ eine Basis von V'. Offenbar ist dann $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = E_n$.

Beispiel 5.4.10. Es sei $\varphi \in L(K^n, K)$ mit

$$\varphi(x_1, \ldots, x_n) = \sum_{j=1}^{n} a_j x_j$$

eine Linearform. Weiter seien $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ bzw. $\mathcal{B}' = \{1\}$ die Standardbasen von K^n bzw. K. Dann gilt

$$\varphi(\vec{e}_i) = \varphi(0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1}, 0, \ldots, 0) = a_i,$$

also ist die Darstellungsmatrix die Zeile $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = (a_1, \ldots, a_n)$.

Beispiel 5.4.11. Es sei

$$V = \{p \colon \mathbb{R} \to \mathbb{R},\ p \text{ Polynom vom Grad kleiner gleich } n\}$$

mit der Basis $\mathcal{B} = \{1, x, \ldots, x^n\}$ und $\varphi \colon V \to V$ mit $\varphi(p) = p'$ die Ableitung. Dann gilt

$$\begin{array}{rcccl}
\varphi(1) & = & 0 & = & 0 \cdot 1 + 0 \cdot x + \cdots + 0 \cdot x^{n-1} + 0 \cdot x^n \\
\varphi(x) & = & 1 & = & 1 \cdot 1 + 0 \cdot x + \cdots + 0 \cdot x^{n-1} + 0 \cdot x^n \\
\varphi(x^2) & = & 2x & = & 0 \cdot 1 + 2 \cdot x + \cdots + 0 \cdot x^{n-1} + 0 \cdot x^n \\
\vdots & & \vdots & & \vdots \\
\varphi(x^n) & = & nx^{n-1} & = & 0 \cdot 1 + 0 \cdot x + \cdots + n \cdot x^{n-1} + 0 \cdot x^n
\end{array}$$

und somit

$$\mathcal{M}(\varphi;\mathcal{B},\mathcal{B}) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{(n+1,n+1)}.$$

Der nächste Satz besagt, dass lineare Abbildungen zwischen endlichdimensionalen Vektorräumen praktisch mit Matrizen identifiziert werden können:

Satz 5.4.12. *Es sei* $\dim V = n$ *und* $\dim V' = m$. *Dann gilt* $L(V, V') \cong K^{(m,n)}$.
Genauer: Ist $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ *eine Basis von* V *und* $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_m\}$ *eine Basis von* V', *so ist ein Isomorphismus* $\Phi\colon L(V, V') \to K^{(m,n)}$ *durch* $\Phi(\varphi) := \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ *für* $\varphi \in L(V, V')$ *gegeben. Insbesondere ist* $\dim L(V, V') = \dim K^{(m,n)} = m \cdot n$.
Der inverse Isomorphismus $\Phi^{-1}\colon K^{(m,n)} \to L(V, V')$ *ist durch*

$$\Phi^{-1}(\mathcal{A}) = \varphi_{\mathcal{A}}, \quad \varphi_{\mathcal{A}}(\lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n) = \lambda'_1\vec{b}'_1 + \cdots + \lambda'_m\vec{b}'_m, \quad \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = \mathcal{A} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

gegeben.

Beweis. Man rechnet leicht nach, dass die Abbildung Φ linear ist.
Zur Injektivität von Φ:
Es sei $\varphi \in \text{Kern}(\Phi)$, d.h. $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = 0^{(m,n)}$. Nach Satz 5.4.2 ist dann $\varphi(\vec{v}) = \vec{0}'$ für alle $\vec{v} \in V$, d.h. φ ist der Nullvektor von $L(V, V')$, also $\text{Kern}(\Phi) = \{\vec{0}\}$, womit Φ nach Satz 5.2.6 injektiv ist.
Zur Surjektivität von Φ:
Es sei $\mathcal{A} \in K^{(m,n)}$. Dann wird $\varphi_{\mathcal{A}}$ wie im Satz definiert: für $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n \in V$ sei

$$\varphi_{\mathcal{A}}(\vec{v}) := \lambda'_1\vec{b}'_1 + \cdots + \lambda'_m\vec{b}'_m \quad \text{mit} \quad \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_m \end{pmatrix} = \mathcal{A} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Dann ist $\varphi_{\mathcal{A}} \in L(V, V')$ und $\Phi(\varphi_{\mathcal{A}}) = \mathcal{M}(\varphi_{\mathcal{A}}; \mathcal{B}', \mathcal{B}) = \mathcal{A}$. Gleichzeitig ergibt sich die angegebene Formel für Φ^{-1}. □

Wir geben nun die angekündigte Erklärung für die komplizierte Definition des Matrizenprodukts. Der Komposition von Abbildungen entspricht die Multiplikation der zugehörigen Matrizen.

Satz 5.4.13. *Es seien* V, V' *und* V'' *endlichdimensionale Vektorräume mit Basen* $\mathcal{B}$, $\mathcal{B}'$ *und* $\mathcal{B}''$. *Für* $\varphi \in L(V, V')$ *und* $\psi \in L(V', V'')$ *gilt dann*

$$\mathcal{M}(\psi \circ \varphi; B'', B) = \mathcal{M}(\psi; B'', B') \cdot \mathcal{M}(\varphi; B', B).$$

Beweis. Es sei $\dim V = n$, $\dim V' = m$ und $\dim V'' = p$ sowie $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$, $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_m\}$ und $\mathcal{B}'' = \{\vec{b}''_1, \ldots, \vec{b}''_p\}$. Wir definieren die Matrizen $\mathcal{A}$ und $\mathcal{C}$ durch

$$\begin{aligned} \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) &=: \mathcal{A} = (\alpha_{kl}) \text{ mit } 1 \le k \le m \text{ und } 1 \le l \le n \quad \text{mit} \quad \varphi(\vec{b}_l) = \sum_{k=1}^{m} \alpha_{kl}\vec{b}'_k, \\ \mathcal{M}(\psi; \mathcal{B}'', \mathcal{B}') &=: \mathcal{C} = (\gamma_{jk}) \text{ mit } 1 \le j \le p \text{ und } 1 \le k \le m \quad \text{mit} \quad \psi(\vec{b}'_k) = \sum_{j=1}^{p} \gamma_{jk}\vec{b}''_j. \end{aligned}$$

Dann ist

$$(\psi \circ \varphi)(\vec{b}_l) \quad = \quad \psi(\varphi(\vec{b}_l)) = \psi\left(\sum_{k=1}^{m} \alpha_{kl}\vec{b}'_k\right) = \sum_{k=1}^{m} \alpha_{kl}\psi(\vec{b}'_k) = \sum_{k=1}^{m} \alpha_{kl} \sum_{j=1}^{p} \gamma_{jk}\vec{b}''_j = \sum_{j=1}^{p}\left(\sum_{k=1}^{m} \gamma_{jk}\alpha_{kl}\right)\vec{b}''_j,$$

woraus

$$\mathcal{M}(\psi \circ \varphi; \mathcal{B}'', \mathcal{B}) = \left(\sum_{k=1}^{m} \gamma_{jk}\alpha_{kl}\right)_{\substack{1\le j\le p\\ 1\le l\le n}} = \mathcal{C} \cdot \mathcal{A}$$

folgt. □

Die Rechenregeln für die Matrizenmultiplikation von Satz 4.1.10 ergeben sich aus Satz 5.4.13 unmittelbar, da sie für die Komposition von linearen Abbildungen erfüllt sind.

Bemerkung 5.4.14. Der Endomorphismenring $(L(V,V),+,\circ)$ aus Satz 5.1.11 ist mit dem Matrizenring $K^{(n,n)}$ aus Satz 4.1.12 eng verwandt. Für die in Satz 5.4.12 definierte bijektive Abbildung $\Phi\colon L(V,V) \to K^{(n,n)}$ gilt

$$\Phi(\varphi + \psi) = \Phi(\varphi) + \Phi(\psi) \quad \text{und} \quad \Phi(\varphi \circ \psi) = \Phi(\varphi) \cdot \Phi(\psi).$$

Eine solche Abbildung zwischen Ringen heißt Ringisomorphismus.

Satz 5.4.15. *Es seien V und V' endlichdimensionale Vektorräume mit den Basen $\mathcal{B}$ und $\mathcal{B}'$ sowie $\varphi \in L(V,V')$. Dann gilt $\mathrm{rg}(\varphi) = \mathrm{rg}(\mathcal{M}(\varphi;\mathcal{B}',\mathcal{B}))$. Insbesondere hängt der Rang von $\mathcal{M}(\varphi;\mathcal{B}',\mathcal{B})$ also nicht von der Wahl der Basen $\mathcal{B}$ und $\mathcal{B}'$ ab.*

Beweis. Es seien $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ und $\mathcal{B}' = \{\vec{b}'_1, \ldots, \vec{b}'_m\}$ sowie $\mathcal{M}(\varphi;\mathcal{B}',\mathcal{B}) = \mathcal{A} = (\alpha_{kl})_{\substack{1\le k\le m\\ 1\le l\le n}}$.
Für $l = 1 \ldots n$ ist

$$\varphi(\vec{b}_l) = \sum_{k=1}^{m} \alpha_{kl}\vec{b}'_k,$$

d.h. der l-te Spaltenvektor

$$\vec{a}_l = \begin{pmatrix} \alpha_{1l} \\ \vdots \\ \alpha_{ml} \end{pmatrix}$$

von $\mathcal{A}$ ist der Koordinatenvektor von $\varphi(\vec{b}_l)$ bzgl. der Basis $\mathcal{B}'$. Wir zeigen für $1 \le l_1 < l_2 < \ldots < l_r \le n$ die Äquivalenz

$$\varphi(\vec{b}_{l_1}), \ldots, \varphi(\vec{b}_{l_r}) \text{ sind linear unabhängig} \Leftrightarrow \vec{a}_{l_1}, \ldots, \vec{a}_{l_r} \text{ sind linear unabhängig.} \tag{$*$}$$

Es gilt

$$\begin{aligned} \vec{0} = \sum_{j=1}^{r} \lambda_j \varphi(\vec{b}_{l_j}) = \sum_{j=1}^{r} \lambda_j \left(\sum_{k=1}^{m} \alpha_{kl_j}\vec{b}'_k\right) = \sum_{k=1}^{m}\left(\sum_{j=1}^{r} \lambda_j a_{kl_j}\right)\vec{b}'_k \quad &\Leftrightarrow \quad \sum_{j=1}^{r} \alpha_{kl_j}\lambda_j = 0 \\ &\Leftrightarrow \quad \lambda_1\vec{a}_{l_1} + \cdots + \lambda_r\vec{a}_{l_r} = \vec{0}. \end{aligned}$$

für $k = 1, \ldots, m$, da $\vec{b}'_1, \ldots, \vec{b}'_m$ linear unabhängig sind.
Damit ist $(*)$ bewiesen. Es folgt

$$\begin{aligned} \mathrm{rg}(\varphi) \quad &= \quad \dim \varphi(V) = \dim(\langle\, \varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n)\,\rangle) \\ &= \quad \text{Maximalzahl linear unabhängiger Vektoren unter den } \varphi(\vec{b}_1), \ldots, \varphi(\vec{b}_n) \\ &= \quad \text{Maximalzahl linear unabhängiger Vektoren unter den } \vec{a}_1, \ldots, \vec{a}_n \\ &= \quad \text{Spaltenrang von } \mathcal{A} = \mathrm{rg}(\mathcal{A}). \end{aligned}$$

□

Satz 5.4.16. *Es sei* $\dim V = \dim V' = n$ *und* $\mathcal{B}$ *bzw.* $\mathcal{B}'$ *Basen von* V *bzw.* V' *sowie* $\varphi \in L(V, V')$. *Dann ist* φ *genau dann bijektiv, wenn* $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ *regulär ist.*
In diesem Falle ist $\mathcal{M}(\varphi^{-1}; \mathcal{B}, \mathcal{B}') = \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})^{-1}$.

Beweis. Es gilt

$$\varphi \text{ bijektiv} \underset{\text{Satz 5.2.8 ii)}}{\Leftrightarrow} \varphi \text{ injektiv} \underset{\text{Satz 5.2.8 i)}}{\Leftrightarrow} \operatorname{rg}(\varphi) = n \underset{\text{Satz 5.4.15}}{\Leftrightarrow} \operatorname{rg}(\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})) = n.$$

Und nach Satz 5.4.13 ist

$$\mathcal{M}(\varphi^{-1}; \mathcal{B}, \mathcal{B}') \cdot \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = \mathcal{M}(\varphi^{-1} \circ \varphi; \mathcal{B}, \mathcal{B}) = \mathcal{M}(id_V; \mathcal{B}, \mathcal{B}) = E_n.$$

□

Bemerkung 5.4.17. Die Abbildung $\Phi\colon L(V, V) \to K^{(n,n)}$ mit $\Phi(\varphi) = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ bildet daher die Automorphismengruppe von V bijektiv auf die Gruppe der regulären Matrizen in $K^{(n,n)}$ ab, und es gilt $\Phi(\varphi \circ \psi) = \Phi(\varphi) \cdot \Phi(\psi)$, d.h. Φ ist ein Gruppenisomorphismus.

Satz 5.4.18. *Für Matrizen* $\mathcal{A} \in K^{(m,n)}$ *und* $\mathcal{C} \in K^{(n,p)}$ *gilt*

$$\operatorname{rg}(\mathcal{A}) + \operatorname{rg}(\mathcal{C}) - n \leq \operatorname{rg}(\mathcal{AC}) \leq \min(\operatorname{rg}(\mathcal{A}), \operatorname{rg}(\mathcal{C})).$$

Beweis. Wir wählen Vektorräume V, V', V'' mit den Dimensionen p, n, m und den Basen $\mathcal{B}, \mathcal{B}', \mathcal{B}''$. Nach Satz 5.3.1 gibt es $\varphi \in L(V, V')$, $\psi \in L(V', V'')$ mit $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = \mathcal{C}$ und $\mathcal{M}(\psi; \mathcal{B}'', \mathcal{B}') = \mathcal{A}$, also ist nach Satz 5.4.13 $\mathcal{M}(\psi \circ \varphi; \mathcal{B}'', \mathcal{B}) = \mathcal{A} \cdot \mathcal{C}$. Nach Satz 5.4.15 ist $\operatorname{rg}(\mathcal{A}) = \operatorname{rg}(\psi)$, $\operatorname{rg}(\mathcal{C}) = \operatorname{rg}(\varphi)$ und $\operatorname{rg}(\mathcal{AC}) = \operatorname{rg}(\psi \circ \varphi)$, so dass die Behauptung aus Satz 5.3.4 folgt. □

Satz 5.4.19. *Es sei* $\mathcal{A} \in K^{(m,n)}$ *beliebig, und es seien* $\mathcal{X} \in K^{(m,m)}$ *bzw.* $\mathcal{Y} \in K^{(n,n)}$ *regulär. Dann ist* $\operatorname{rg}(\mathcal{XA}) = \operatorname{rg}(\mathcal{AY}) = \operatorname{rg}(\mathcal{A})$, *d.h. Multiplikationen mit regulären Matrizen ändern den Rang einer Matrix nicht.*

Beweis. Nach Satz 5.4.18 gilt

$$\operatorname{rg}(\mathcal{XA}) \left\{ \begin{array}{lllll} \leq & \min(\operatorname{rg}(\mathcal{X}), \operatorname{rg}(\mathcal{A})) & = & \operatorname{rg}(\mathcal{A}) & \text{wegen } \operatorname{rg}(\mathcal{A}) \leq m = \operatorname{rg}(\mathcal{X}), \\ \geq & \operatorname{rg}(\mathcal{X}) + \operatorname{rg}(\mathcal{A}) - m & = & \operatorname{rg}(\mathcal{A}) & \text{wegen } \operatorname{rg}(\mathcal{X}) = m. \end{array} \right.$$

sowie

$$\operatorname{rg}(\mathcal{AY}) \left\{ \begin{array}{lllll} \leq & \min(\operatorname{rg}(\mathcal{A}), \operatorname{rg}(\mathcal{Y})) & = & \operatorname{rg}(\mathcal{A}) & \text{wegen } \operatorname{rg}(\mathcal{A}) \leq n = \operatorname{rg}(\mathcal{Y}), \\ \geq & \operatorname{rg}(\mathcal{A}) + \operatorname{rg}(\mathcal{Y}) - n & = & \operatorname{rg}(\mathcal{A}) & \text{wegen } \operatorname{rg}(\mathcal{Y}) = n. \end{array} \right.$$

□

5.5 Basiswechsel

Wir untersuchen nun, wie sich die Koordinaten eines Vektors $\vec{v} \in V$ beim Wechsel der Basis des Vektorraums V ändern. Außerdem untersuchen wir die damit eng verwandte Frage, wie sich die zu einer Abbildung $\varphi \in L(V, V')$ bzgl. der Basen $\mathcal{B}$ und $\mathcal{B}'$ gehörende Matrix $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B})$ ändert, wenn die Basen $\mathcal{B}$ und $\mathcal{B}'$ durch andere Basen ersetzt werden. Zunächst beweisen wir einen Satz über die Gesamtheit aller Basen eines Vektorraums.

Satz 5.5.1. *Es sei $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis von V und $\varphi \in L(V, V)$. Die Menge $\tilde{\mathcal{B}} = \{\varphi(\vec{b}_1), \dots, \varphi(\vec{b}_n)\}$ ist genau dann eine Basis von V, wenn φ ein Automorphismus ist.*

Beweis. Es gilt

$\tilde{\mathcal{B}}$ ist eine Basis $\Leftrightarrow \varphi(\vec{b}_1), \dots, \varphi(\vec{b}_n)$ sind linear unabhängig $\Leftrightarrow \operatorname{rg}(\varphi) = n \Leftrightarrow \varphi$ ist ein Automorphismus.

□

Bemerkung 5.5.2. Es besteht also eine umkehrbar eindeutige Entsprechung zwischen

i) den Automorphismen von V,

ii) den Basistransformationen von V,

iii) den regulären $n \times n$- Matrizen über K.

Satz 5.5.3. *Es seien $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ und $\mathcal{B}' = \{\vec{b}'_1, \dots, \vec{b}'_n\}$ Basen von V und $\varphi \in \mathrm{GL}(V)$ ein Automorphismus. Dann gilt $\mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) = \mathcal{M}(id_V; \mathcal{B}', \varphi(\mathcal{B}))$ mit $\varphi(\mathcal{B}) = \{\varphi(\vec{b}_1), \dots, \varphi(\vec{b}_n)\}$.*

Beweis. Die Menge $\varphi(\mathcal{B})$ ist nach Satz 5.5.1 eine Basis von V. Beide Matrizen haben die Gestalt $(\alpha_{kl})_{1 \le k,l \le n}$ mit

$$\varphi(\vec{b}_l) = id_V(\varphi(\vec{b}_l)) = \sum_{k=1}^{n} \alpha_{kl} \vec{b}'_k$$

mit $l = 1, \dots, n$. □

Satz 5.5.4. *Es seien $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ und $\varphi(\mathcal{B}) = \{\varphi(\vec{b}_1), \dots, \varphi(\vec{b}_n)\}$ Basen von V für $\varphi \in \mathrm{GL}(V)$. Sind $\lambda_1, \dots, \lambda_n$ bzw. $\lambda'_1, \dots, \lambda'_n$ die Koordinaten von $\vec{v}$ bzgl. $\mathcal{B}$ bzw. bzgl. $\varphi(\mathcal{B})$, d.h. gilt*

$$\vec{v} = \lambda_1 \vec{b}_1 + \dots + \lambda_n \vec{b}_n \quad \textit{bzw.} \quad \vec{v} = \lambda'_1 \varphi(\vec{b}_1) + \dots + \lambda'_n \varphi(\vec{b}_n),$$

so ist

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})^{-1} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Beweis. Es gilt

$$\begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix} \underset{\text{Satz 5.4.2}}{=} \mathcal{M}(id_V; \varphi(\mathcal{B}), \mathcal{B}) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \underset{\text{Satz 5.4.13}}{=} \mathcal{M}(id_V; \mathcal{B}, \varphi(\mathcal{B}))^{-1} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

wegen

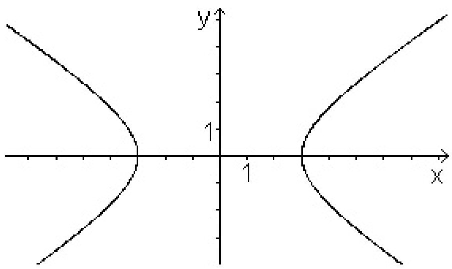
$$E_n = \mathcal{M}(id_V; \mathcal{B}, \mathcal{B}) = \mathcal{M}(id_V; \mathcal{B}, \varphi(\mathcal{B})) \cdot \mathcal{M}(id_V; \varphi(\mathcal{B}), \mathcal{B}),$$

woraus $\mathcal{M}(id_V; \varphi(\mathcal{B}), \mathcal{B}) = \mathcal{M}(id_V; \mathcal{B}, \varphi(\mathcal{B}))^{-1}$ folgt.
Nach Satz 5.5.3 folgt schließlich über

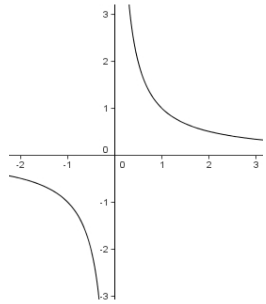
$$\mathcal{M}(id_V; \mathcal{B}, \varphi(\mathcal{B})) = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$$

die Behauptung. □

Beispiel 5.5.5. In der Geometrie oder der Analysis begegnet man der Hyperbel in zwei Formen als Kurven. Die Kurven in der xy-Ebene mit den Gleichungen

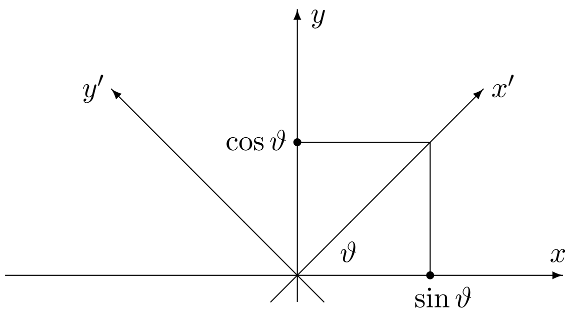


(I) $x^2 - y^2 = b$ $(b > 0)$



(II) $xy = a$ $(a > 0)$

werden jeweils als Hyperbeln bezeichnet.
Woher wissen wir, ob durch die Gleichungstypen (I) und (II) kongruente Kurven beschrieben werden? Die Skizzen legen die Vermutung nahe, dass Kurven vom Typ (II) durch eine Drehung um 45° gegen den Uhrzeigersinn in Kurven des Typs (I) übergeführt werden. Zum Beweis dieser Vermutung drehen wir das xy-Koordinatensystem um 45° in das $x'y'$-Koordinatensystem und untersuchen die Gleichung der Kurve $xy = a$ in $x'y'$-Koordinaten. Zunächst leiten wir die Transformationsgleichungen für eine Drehung des Koordinatensystems um einen Winkel ϑ im Uhrzeigersinn her.



Die x'-Achse wird vom Vektor $\vec{b}_1 = (\cos\vartheta, \sin\vartheta)$ bzw. die y'-Achse von $\vec{b}_2 = (-\sin\vartheta, \cos\vartheta)$ aufgespannt. Während die xy-Koordinaten gerade die Koeffizienten $\vec{v} = (x, y) \in \mathbb{R}^2$ in der Darstellung $\vec{v} = (x, y) = x\vec{e}_1 + y\vec{e}_2$ sind, sind die $x'y'$-Koordinaten von $\vec{v}$ die Koeffizienten x' und y' in der Darstellung $\vec{v} = x'\vec{b}_1 + y'\vec{b}_2$. Nach Satz 5.5.4 ist daher

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})^{-1} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} x \\ y \end{pmatrix} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) \cdot \begin{pmatrix} x' \\ y' \end{pmatrix}$$

Die Matrix der Drehung $\varphi \in L(\mathbb{R}^2, \mathbb{R}^2)$, die die Basis $\mathcal{B} = (\vec{e}_1, \vec{e}_2)$ in die Basis $\mathcal{B}' = (\vec{b}_1, \vec{b}_2)$ überführt, resultiert aus $\varphi\colon \vec{e}_1 \to \vec{b}_1$ und $\varphi\colon \vec{e}_2 \to \vec{b}_2$ sowie $\varphi(\vec{e}_1) = \vec{b}_1 = (\cos\vartheta, \sin\vartheta) = \cos\vartheta \cdot \vec{e}_1 + \sin\vartheta \cdot \vec{e}_2$ bzw. $\varphi(\vec{e}_2) = \vec{b}_2 = (-\sin\vartheta, \cos\vartheta) = -\sin\vartheta \cdot \vec{e}_1 + \cos\vartheta \cdot \vec{e}_2$, was zu

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}$$

führt. Damit ergibt sich

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{oder} \quad \begin{cases} x = (\cos\vartheta)x' - (\sin\vartheta)y' \\ y = (\sin\vartheta)x' + (\cos\vartheta)y'. \end{cases}$$

Für $\vartheta = \frac{\pi}{4}$, was 45° entspricht, erhalten wir $\sin\vartheta = \cos\vartheta = \frac{1}{\sqrt{2}}$, weswegen die Gleichung $xy = a$ zur Gleichung $\frac{1}{2}(x' - y')(x' + y') = a$ oder $x'^2 - y'^2 = 2a$ äquivalent ist. Die Kurve $xy = a$ geht daher aus der Kurve $x^2 - y^2 = 2a$ durch eine Drehung um 45° hervor.
Die Gleichungen (I) und (II) stellen in der Tat kongruente Kurven dar, falls $b = 2a$ ist.

Satz 5.5.6. *Es seien* $\dim V, \dim V' < \infty$ *und* $\mathcal{B}$ *bzw.* $\mathcal{B}'$ *Basen von* V *bzw.* V'. *Zudem seien* $\varrho \in \mathrm{GL}(V)$ *bzw.* $\sigma \in \mathrm{GL}(V')$ *sowie* $\mathcal{X} = \mathcal{M}(\varrho; \mathcal{B}, \mathcal{B})$ *und* $\mathcal{Y} = \mathcal{M}(\sigma; \mathcal{B}', \mathcal{B}')$. *Dann gilt für alle* $\varphi \in L(V, V')$

$$\mathcal{M}(\varphi; \sigma(\mathcal{B}'), \varrho(\mathcal{B})) = \mathcal{Y}^{-1} \cdot \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) \cdot \mathcal{X}.$$

Beweis. Nach Satz 5.4.13 gilt

$$\begin{aligned}
\mathcal{M}(\varphi; \sigma(\mathcal{B}'), \varrho(\mathcal{B})) &= \mathcal{M}(\mathrm{id}_{V'} \circ \varphi \circ \mathrm{id}_V; \sigma(\mathcal{B}'), \varrho(\mathcal{B})) = \mathcal{M}(\mathrm{id}_{V'} \circ \varphi; \sigma(\mathcal{B}'), \mathcal{B}) \cdot \mathcal{M}(\mathrm{id}_V; \mathcal{B}, \varrho(\mathcal{B})) \\
&= \mathcal{M}(\mathrm{id}_{V'}; \sigma(\mathcal{B}'), \mathcal{B}') \cdot \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) \cdot \mathcal{M}(\mathrm{id}_V; \mathcal{B}, \varrho(\mathcal{B})).
\end{aligned}$$

Nach Satz 5.5.3 folgt

$$\mathcal{M}(\mathrm{id}_{V'}; \sigma(\mathcal{B}'), \mathcal{B}') = \mathcal{M}(\mathrm{id}_{V'}; \mathcal{B}', \sigma(\mathcal{B}'))^{-1} = \mathcal{M}(\sigma; \mathcal{B}', \mathcal{B}')^{-1} = \mathcal{Y}^{-1}$$

und ebenso

$$\mathcal{M}(\mathrm{id}_V; \mathcal{B}, \varrho(\mathcal{B})) = \mathcal{M}(\varrho; \mathcal{B}, \mathcal{B}) = \mathcal{X}.$$

Damit ergibt sich die Behauptung. □

Wir erhalten als Spezialfall

Satz 5.5.7. *(Basiswechsel für lineare Abbildungen)*
Es sei $\dim V < \infty$, $\sigma \in \mathrm{GL}(V)$ *und* $\mathcal{B}$ *eine Basis von* V. *Dann gilt für alle* $\varphi \in L(V, V)$

$$\mathcal{M}(\varphi; \sigma(\mathcal{B}), \sigma(\mathcal{B})) = \mathcal{X}^{-1} \cdot \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) \cdot \mathcal{X}$$

mit $\mathcal{X} = \mathcal{M}(\sigma; \mathcal{B}, \mathcal{B})$.

Beispiel 5.5.8. Im Beispiel 5.4.8 betrachteten wir die lineare Abbildung $\varphi \in L(\mathbb{R}^3, \mathbb{R}^3)$ mit der bzgl. der Basis $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ zugeordneten Matrix

$$\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix}.$$

Wir betrachteten die neue Basis $\tilde{\mathcal{B}} = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

Es ist $\tilde{\mathcal{B}} = \sigma(\mathcal{B})$ mit

$$\mathcal{X} = \mathcal{M}(\sigma; \mathcal{B}, \mathcal{B}) = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ -1 & -4 & 2 \end{pmatrix}.$$

Außerdem war

$$\mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Satz 5.5.7 sagt uns daher, dass

$$\mathcal{X}^{-1} \cdot \begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix} \cdot \mathcal{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

ist, was man durch Nachrechnen bestätigt.

Wir haben die Matrix

$$\begin{pmatrix} -3 & -2 & -4 \\ -3 & 0 & -3 \\ 8 & 4 & 9 \end{pmatrix}$$

diagonalisiert.
Im Kapitel über Eigenwerte werden wir das Verfahren zur Diagonalisierung beschreiben.

Ein wichtiger Spezialfall von $L(V, V')$ ist

Definition 5.5.9. Der Dualraum eines K-Vektorraums V ist $V^* = L(V, K)$, die $\varphi \in V^*$ nennt man Linearformen (im Fall $K = \mathbb{R}, \mathbb{C}$ auch lineare Funktionale).

Satz 5.5.10. *Es sei* $\dim V = n < \infty$, *dann ist* $V^* \cong V$. *Ist* $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ *eine Basis von* V, *so gibt es zu jedem* $\varphi \in V^*$ *eindeutig bestimmte* $\alpha_1, \ldots, \alpha_n \in K$ *mit*

$$\varphi(\vec{v}) = \alpha_1\lambda_1 + \cdots + \alpha_n\lambda_n$$

für $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$. *Die Abbildung* $\Psi\colon V^* \to K^n$, $\varphi \to (\alpha_1, \ldots, \alpha_n)$ *ist ein Isomorphismus.*

Beweis. Nach Satz 5.4.12 ist $\dim V^* = n \cdot 1 = \dim V$, womit $V^* \cong V$ nach Satz 5.3.3 gilt. Bzgl. der Basis $\{1\}$ von K ist $\mathcal{M}(\varphi; \{1\}, \mathcal{B}) = (\alpha_1, \ldots, \alpha_n) \in K^{(1\times n)}$, also nach Satz 5.4.2

$$\varphi(\vec{v}) = (\alpha_1, \ldots, \alpha_n) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \cdot 1$$

für $\vec{v} = \lambda_1\vec{b}_1 + \cdots + \lambda_n\vec{b}_n$. Die Linearität und Bijektivität der Abbildung Ψ prüft man leicht durch Nachrechnen. $\square$

Kapitel 6

Lineare Gleichungen

Es sei K stets ein Körper.

6.1 Theorie der Linearen Gleichungen

Definition 6.1.1. Es seien $\alpha_{kl}, \beta_k \in K$ füür $k = 1, \ldots, m$ und $l = 1, \ldots, n$. Man bezeichnet

$$\begin{cases} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \cdots & + & \alpha_{1n}x_n & = & \beta_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + & \cdots & + & \alpha_{2n}x_n & = & \beta_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + & \cdots & + & \alpha_{mn}x_n & = & \beta_m \end{cases} \tag{$*$}$$

als ein lineares Gleichungssystem (LGS) mit m Gleichungen und n Unbekannten $x_1, \ldots, x_n$ und Koeffizienten $\alpha_{kl} \in K$. Man nennt $\mathcal{A} = (\alpha_{kl}) \in K^{(m,n)}$ die Koeffizientenmatrix und

$$\vec{b} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$$

die rechte Seite des Systems $(*)$.
Fügt man zur Matrix $\mathcal{A}$ als $(n+1)$-te Spalte den Vektor $\vec{b}$ hinzu, so erhält man die erweiterte Koeffizientenmatrix $(\mathcal{A}|\vec{b}) \in K^{(m,n+1)}$ des LGS $(*)$.

Fasst man die Unbekannten $x_1, \ldots, x_n$ zu einem Spaltenvektor

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

zusammen, so lässt sich $(*)$ kürzer als Matrixgleichung

$$\mathcal{A}\vec{x} = \vec{b} \tag{$**$}$$

schreiben. Jedes $\vec{x} \in K^n$, so dass $(*)$ bzw. $(**)$ gilt, heißt eine Lösung des LGS. Gibt es ein solches $\vec{x}$, so heißt $(*)$ lösbar, ansonsten unlösbar. Ist $\vec{b} = \vec{0}$, so heißt das LGS homogen, ansonsten inhomogen. In diesem Fall nennt man $\mathcal{A}\vec{x} = \vec{0}$ das zu $(**)$ gehörende homogene LGS.
Ein homogenes LGS besitzt immer die triviale Lösung $\vec{x} = \vec{0}$.

Es erheben sich folgende Fragen:

1. Unter welchen Bedingungen an $\mathcal{A}$ und $\vec{b}$ ist $(*)$ lösbar?
2. Wie sieht die Lösungsmenge von $(*)$ aus?
3. Gibt es ein Verfahren, die Lösungsmenge zu berechnen?

Für eine individuell gegebene rechte Seite $\vec{b}$ hat man das Kriterium

Satz 6.1.2. *Das LGS $\mathcal{A}\vec{x} = \vec{b}$ ist genau dann lösbar, wenn $\mathrm{rg}(\mathcal{A}|\vec{b}) = \mathrm{rg}(\mathcal{A})$ gilt.*

Beweis. Das System $\mathcal{A}\vec{x} = \vec{b}$ ist genau dann lösbar, wenn $\vec{b}$ eine Linearkombination der Spalten $\vec{a}_1, \ldots, \vec{a}_n$ von $\mathcal{A}$ ist, und es gilt

$$\begin{aligned}\vec{b} \in \langle\, \vec{a}_1, \ldots, \vec{a}_n \,\rangle \quad &\Leftrightarrow \quad \langle\, \vec{a}_1, \ldots, \vec{a}_n, \vec{b} \,\rangle = \langle\, \vec{a}_1, \ldots, \vec{a}_n \,\rangle \Leftrightarrow \dim(\langle\, \vec{a}_1, \ldots, \vec{a}_n, \vec{b} \,\rangle) = \dim(\langle\, \vec{a}_1, \ldots, \vec{a}_n \,\rangle) \\ &\Leftrightarrow \quad \text{Spaltenrang von } (\mathcal{A}|\vec{b}) = \text{Spaltenrang von } \mathcal{A} \Leftrightarrow \mathrm{rg}(\mathcal{A}|\vec{b}) = \mathrm{rg}(\mathcal{A}).\end{aligned}$$

□

Betrachtung mehrerer rechter Seiten führt auf folgende Begriffe:

Definition 6.1.3. Ein LGS $\mathcal{A}\vec{x} = \vec{b}$ mit $\mathcal{A} \in K^{(m,n)}$ heißt <u>universell lösbar</u>, wenn $(*)$ für jede rechte Seite $\vec{b} \in K^m$ lösbar ist bzw. <u>eindeutig lösbar</u>, wenn $(*)$ für jede rechte Seite $\vec{b}$ höchstens eine (und eventuell sogar keine) Lösung $\vec{x} \in K^n$ besitzt.

Satz 6.1.4. *Ein LGS $\mathcal{A}\vec{x} = \vec{b}$ mit $\mathcal{A} \in K^{(m,n)}$ ist genau dann*

i) universell lösbar, wenn $\mathrm{rg}(\mathcal{A}) = m$

ii) eindeutig lösbar, wenn $\mathrm{rg}(\mathcal{A}) = n$

gilt.

Beweis. Wir betrachten $\varphi \in L(K^n, K^m)$ mit $\varphi(\vec{x}) = \mathcal{A}\vec{x}$. Dann gilt

$$\mathcal{A}\vec{x} = \vec{b} \text{ universell lösbar} \Leftrightarrow \varphi \text{ surjektiv, d.h. } \varphi(K^n) = K^m \Leftrightarrow \mathrm{rg}(\mathcal{A}) = \mathrm{rg}(\varphi) = \dim(K^m) = m.$$

Ebenso gilt

$$\begin{aligned}\mathcal{A}\vec{x} = \vec{b} \text{ eindeutig lösbar} \quad &\Leftrightarrow \quad \varphi \text{ injektiv} \underset{\text{Satz 5.2.6}}{\Leftrightarrow} \mathrm{Kern}(\varphi) = \{\vec{0}\} \\ &\underset{\text{Satz 5.2.7}}{\Leftrightarrow} \quad 0 = \mathrm{def}(\varphi) = \dim K^n - \mathrm{rg}(\varphi) = n - \mathrm{rg}(\mathcal{A}) \Leftrightarrow \mathrm{rg}(\mathcal{A}) = n.\end{aligned}$$

□

Satz 6.1.5. *Ein LGS $\mathcal{A}\vec{x} = \vec{b}$ mit $\mathcal{A} \in K^{(m,n)}$ ist genau dann universell und eindeutig lösbar, wenn $m = n = \mathrm{rg}(\mathcal{A})$ ist, d.h. wenn $\mathcal{A}$ quadratisch und regulär ist. Für ein LGS der Form $\mathcal{A}\vec{x} = \vec{b}$ mit $\mathcal{A} \in K^{(n,n)}$, also für das die Anzahl der Unbekannten mit der Anzahl der Gleichungen übereinstimmt, gilt*

$$\mathcal{A}\vec{x} = \vec{b} \text{ universell lösbar} \Leftrightarrow \mathcal{A}\vec{x} = \vec{b} \text{ eindeutig lösbar.}$$

Beweis. Das folgt unmittelbar aus Satz 6.1.4. □

Satz 6.1.6. *Die Lösungsmenge* $H = \{\vec{x} \in K^n : \mathcal{A}\vec{x} = \vec{0}\}$ *eines homogenen LGS mit* $\mathcal{A} \in K^{(m,n)}$ *ist ein Unterraum des* K^n *der Dimension* $\dim H = n - \mathrm{rg}(\mathcal{A})$.

Beweis. Es ist $H = \mathrm{Kern}(\varphi)$ mit $\varphi\colon \vec{x} \to \mathcal{A}\vec{x}$, also $\dim H = n - \mathrm{rg}(\varphi) = n - \mathrm{rg}(\mathcal{A})$ nach Satz 5.2.7. □

Um die Lösungsmenge eines inhomogenen LGS beschreiben zu können, führen wir den Begriff der linearen Mannigfaltigkeit ein:

Definition 6.1.7. Eine Teilmenge M eines Vektorraums V heißt eine lineare Mannigfaltigkeit (oder auch affiner Unterraum) der Dimension m, wenn sie als

$$M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U\}$$

für ein festes $\vec{v}_0 \in V$ und einen Untervektorraum U von V mit $\dim U = m$ geschrieben werden kann.

Satz 6.1.8. *Der Vektorraum* U *ist durch* M *eindeutig bestimmt, d.h. ist*

$$M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U_0\} = \{\vec{v}_1 + \vec{u} \colon \vec{u} \in U_1\},$$

so ist $U_0 = U_1$. *Für* $\vec{v}_0$ *kann jeder Vektor aus* M *genommen werden.*

Beweis. Es sei $\vec{u}_0 \in U_0$. Dann gilt $\vec{v}_0 + \vec{u}_0 = \vec{v}_1 + \vec{u}_1$ mit gewissen $\vec{v}_1 \in M$ und $\vec{u}_1 \in U_1$.
Wegen $\vec{v}_0 = \vec{v}_0 + \vec{0} \in M$ gilt $\vec{v}_0 = \vec{v}_1 + \vec{w}_1$ mit $\vec{w}_1 \in U_1$, also

$$\vec{v}_0 + \vec{u}_0 = \vec{v}_1 + \vec{w}_1 + \vec{u}_0 = \vec{v}_1 + \vec{u}_1,$$

woraus

$$\vec{u}_0 = \vec{u}_1 - \vec{w}_1 \in U_1$$

und somit $U_0 \subseteq U_1$ folgt. Genauso folgt $U_1 \subseteq U_0$, also $U_0 = U_1$.
Es sei nun $\vec{v}_0' \in M$. Dann ist $\vec{v}_0' = \vec{v}_0 + \vec{u}'$ mit $\vec{u}' \in U$. Daraus folgt

$$M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U\} = \{\vec{v}_0' + \vec{u} \colon \vec{u} \in U\}.$$

□

Beispiel 6.1.9. Den ersten Beispielen sind wir bereits in der Einleitung begegnet. Es sind nämlich die Geraden im $\mathbb{R}^2$ oder im $\mathbb{R}^3$, die als $\{\vec{a} + t\vec{b} \colon t \in \mathbb{R}\}$ mit $\vec{b} \neq 0$ bzw. die Ebenen im $\mathbb{R}^3$, die als $\{\vec{a} + s\vec{b} + t\vec{c} \colon s, t \in \mathbb{R}\}$ mit linear unabhängigen Vektoren $\vec{b}$ und $\vec{c}$ geschrieben werden können, ein- bzw. zweidimensionale lineare Mannigfaltigkeiten.

Satz 6.1.10. *Eine lineare Mannigfaltigkeit eines Vektorraums* V *ist genau dann ein Untervektorraum von* V, *wenn sie den Nullvektor enthält.*

Beweis. Die Richtung, die besagt, dass wenn M ein Untervektorraum ist, diese den Nullvektor enthält, ist klar. Für $\vec{0} \in M$ gilt andererseits $M = \{\vec{0} + \vec{u} \colon \vec{u} \in U\} = U$ nach Satz 6.1.8. □

Satz 6.1.11. *Sind* M *und* N *lineare Mannigfaltigkeiten von* V *mit* $\dim N < \infty$ *und* $M \subseteq N$, *so ist* $\dim M \leq \dim N$. *Es gilt dann genau dann* $\dim M = \dim N$, *wenn* $M = N$ *gilt.*

Beweis. Es sei $\vec{v}_0 \in M$. Dann gilt $M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U_0\}$ und $N = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U_1\}$ mit gewissen Untervektorräumen U_0 und U_1 von V nach Satz 6.1.8. Es folgt $U_0 \subseteq U_1$, und die Restbehauptung folgt aus Satz 3.3.26. □

Satz 6.1.12. *Ist* $\varphi \in L(V, V')$ *mit* $\dim V, \dim V' < \infty$ *und* M *eine lineare Mannigfaltigkeit in* V*, so ist* $\varphi(M) = \{\varphi(\vec{v}) \colon \vec{v} \in M\}$ *eine lineare Mannigfaltigkeit in* V'*.*
Ist φ *ein Isomorphismus, so gilt* $\dim M = \dim \varphi(M)$*.*

Beweis. Es ist $\varphi(M) = \{\varphi(\vec{v}_0) + \varphi(\vec{u}) \colon \vec{u} \in U\} = \{\varphi(\vec{v}_0) + \vec{w} \colon \vec{w} \in \varphi(U)\}$. □

Wir beschreiben nun die Lösungsmenge des inhomogenen LGS:

Satz 6.1.13. *Ist das LGS*

$$\mathcal{A}\vec{x} = \vec{b} \tag{$*$}$$

mit $\mathcal{A} \in K^{(m,n)}$ *und* $\vec{b} \in K^m$ *lösbar, so ist die Lösungsmenge eine lineare Mannigfaltigkeit des* K^n *der Dimension* $n - \mathrm{rg}(\mathcal{A})$*. Man erhält alle Lösungen des LGS in der Form* $\vec{x} = \vec{x}_0 + \vec{x}_h$*, wenn* $\vec{x}_0$ *eine beliebige* <u>*partikuläre*</u> *oder* <u>*spezielle*</u> *Lösung des LGS ist, und* $\vec{x}_h$ *sämtliche Lösungen des zugehörigen homogenen LGS*

$$\mathcal{A}\vec{x} = \vec{0} \tag{$**$}$$

durchläuft. Die Lösungsmenge von $(*)$ *ist genau dann ein Vektorraum, wenn* $\vec{b} = \vec{0}$ *ist.*

Beweis. Es sei $\vec{x}_h$ eine Lösung von $(**)$. Dann ist $\vec{x} = \vec{x}_0 + \vec{x}_h$ eine Lösung von $(*)$, denn es gilt

$$\mathcal{A}\vec{x} = \mathcal{A}(\vec{x}_0 + \vec{x}_h) = \mathcal{A}\vec{x}_0 + \mathcal{A}\vec{x}_h = \vec{b} + \vec{0} = \vec{b}.$$

Jede Lösung $\vec{x}$ von $(*)$ hat die Form $\vec{x} = \vec{x}_0 + \vec{x}_h$, denn aus $\mathcal{A}\vec{x} = \vec{b}$ folgt $\mathcal{A}(\vec{x} - \vec{x}_0) = \vec{0}$, weswegen $\vec{x} - \vec{x}_0 =: \vec{x}_n$ eine Lösung von $(**)$ ist.
Die Lösungsmenge von $(*)$ ist auch eine lineare Mannigfaltigkeit der Dimension $n - \mathrm{rg}(\mathcal{A})$, da nach Satz 6.1.6 die Lösungsmenge von $(**)$ ein Unterraum des K^n mit eben dieser Dimension ist. Die Lösungsmenge von $(*)$ ist nach Satz 6.1.10 genau dann ein Untervektorraum, wenn $\vec{x} = \vec{0}$ eine Lösung von $(*)$ ist. Dies ist jedoch zu $\vec{b} = \mathcal{A}\vec{0} = \vec{0}$ äquivalent. □

Es gilt nun auch die Umkehrung von Satz 6.1.13. Jede lineare Mannigfaltigkeit des K^n lässt sich als Lösungsmenge eines linearen Gleichungssystems darstellen. Wir beweisen den etwas allgemeineren

Satz 6.1.14. *Es sei* M *eine lineare Mannigfaltigkeit des Vektorraums* V *über* K*. Es sei* $\dim M = m$ *und* $\dim V = n$*,* $0 \le m \le n$ *und* $k = n - m$*. Es sei* $L(M) \subseteq V^*$ *die Menge aller Linearformen, die auf* M *konstant sind. Dann ist* $L(M)$ *ein* k*- dimensionaler Unterraum von* V^**. Für jede Basis* $\{\varphi_1, \dots, \varphi_k\}$ *von* $L(M)$ *gibt es* $c_1, \dots, c_k \in K$*, so dass*

$$\vec{v} \in M \Leftrightarrow \begin{cases} \varphi_1(\vec{v}) = c_1 \\ \quad\vdots \\ \varphi_k(\vec{v}) = c_k \end{cases}$$

gilt. Weiter ist M *genau dann ein Untervektorraum von* V*, wenn* $c_1 = \dots = c_k = 0$ *ist.*

Beweis. Es sei zunächst U ein m- dimensionaler Untervektorraum von V, und $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Basis von V. Nach Satz 5.5.10 gibt es zu jedem $\varphi \in V^*$ eindeutig bestimmte $\alpha_1, \dots, \alpha_n \in K$ mit $\varphi(\vec{v}) = \alpha_1\lambda_1 + \dots + \alpha_n\lambda_n$ für $\vec{v} = \lambda_1\vec{b}_1 + \dots + \lambda_n\vec{b}_n \in V$. Die Abbildung $\Psi \colon \varphi \to (\alpha_1, \dots, \alpha_n)$ ist nach Satz 5.5.10 ein Isomorphismus von V^* nach K^n. Die Vektoren

$$\begin{array}{ccccccc} \vec{v}_1 & = & \lambda_{11}\vec{b}_1 & + & \cdots & + & \lambda_{1n}\vec{b}_n \\ \vec{v}_2 & = & \lambda_{21}\vec{b}_1 & + & \cdots & + & \lambda_{2n}\vec{b}_n \\ \vdots & & \vdots & & & & \vdots \\ \vec{v}_m & = & \lambda_{m1}\vec{b}_1 & + & \cdots & + & \lambda_{mn}\vec{b}_n \end{array}$$

mögen eine Basis von U bilden.

Dann hat die Matrix

$$\operatorname{rg}\mathcal{L} = \begin{pmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{m1} & \cdots & \lambda_{mn} \end{pmatrix} = m.$$

Wären nämlich die Zeilen von $\mathcal{L}$ linear abhängig, dann wären auch $\vec{v}_1, \ldots, \vec{v}_m$ linear abhängig, was der Basiseigenschaft widerspricht. Es gilt

$$\forall \vec{v} \in U\colon \varphi(\vec{v}) = 0 \Leftrightarrow \varphi(\vec{v}_1) = \cdots = \varphi(\vec{v}_n) = 0 \Leftrightarrow \begin{cases} \lambda_{11}\alpha_1 + \cdots + \lambda_{1n}\alpha_n = 0 \\ \lambda_{21}\alpha_1 + \cdots + \lambda_{2n}\alpha_n = 0 \\ \vdots \\ \lambda_{m1}\alpha_1 + \cdots + \lambda_{mn}\alpha_n = 0. \end{cases} \tag{$*$}$$

Das System $(*)$ fassen wir nun als LGS mit der Koeffizientenmatrix $\mathcal{L}$ und den Unbekannten $\alpha_1, \ldots, \alpha_n$ auf. Da $\operatorname{rg}(\mathcal{L}) = m$ ist, bildet nach Satz 6.1.6 die Lösungsmenge $\{(\alpha_1, \ldots, \alpha_n)\}$ von $(*)$ einen Unterraum $\tilde{L}$ des K^n der Dimension $k = n - m$. Die Menge $L(U) = \Psi^{-1}(\tilde{L})$ bildet einen k- dimensionalen Unterraum von V^*. Es sei nun M eine m- dimensionale lineare Mannigfaltigkeit, also $M = \{\vec{v}_0 + \vec{u}\colon \vec{u} \in U\}$ mit einem Unterraum U von V mit $\dim U = m$. Dann gilt

$$\varphi(\vec{v}) \text{ konstant auf } M \Leftrightarrow \forall \vec{u} \in U\colon \varphi(\vec{v}_0 + \vec{u}) = \varphi(\vec{v}_0) \Leftrightarrow \forall u \in U\colon \varphi(\vec{u}) = 0.$$

Diese φ bilden somit einen k- dimensionalen Unterraum von V^*. Es sei nun $\{\varphi_1, \ldots, \varphi_k\}$ eine Basis von $L(M)$ und

$$\varphi_i(\vec{v}) = c_i \tag{$**$}$$

für $1 \leq i \leq k$ und für alle $\vec{v} \in M$. Dazu sei $\Psi(\varphi_i) = (\alpha_{i1}, \ldots, \alpha_{in})$ und

$$\mathcal{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{k1} & \cdots & \alpha_{kn} \end{pmatrix}.$$

Dann ist $\operatorname{rg}(\mathcal{A}) = k$. Für $\vec{v} = x_1\vec{b}_1 + \cdots + x_n\vec{b}_n$ ist $(**)$ zu

$$\mathcal{A} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} \tag{$***$}$$

äquivalent. Nach Satz 6.1.13 ist die Lösungsmenge N von $(***)$ eine m- dimensionale Mannigfaltigkeit von V mit $M \subseteq N$. Nach Satz 6.1.11 gilt $M = N$. □

6.2 Der Gaußsche Algorithmus

Der Gaußsche Algorithmus ist ein algorithmisches Verfahren zur Berechnung der Lösungsmenge eines linearen Gleichungssystems, das wir in diesem Abschnitt beschreiben. Gegeben sei ein LGS $\mathcal{A}\vec{x} = \vec{b}$ mit $\mathcal{A} \in K^{(m,n)}$ und

$$(\mathcal{A}|\vec{b}) = \left(\begin{array}{ccc|c} \alpha_{11} & \cdots & \alpha_{1n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} & \beta_m \end{array}\right). \tag{$*$}$$

- <u>1. Schritt:</u>
  Falls in der erste Spalte mindestens ein Element $\alpha_{k1} \neq 0$ vorkommt, bringe man $(\mathcal{A}|\vec{b})$ durch elementare Zeilenumformungen in die Gestalt

$$(\mathcal{A}^{(1)}|\vec{b}^{(1)}) = \left(\begin{array}{cccc|c} 1 & \alpha_{12}^{(1)} & \cdots & \alpha_{1n}^{(1)} & \beta_1^{(1)} \\ 0 & \alpha_{22}^{(1)} & \cdots & \alpha_{2n}^{(1)} & \beta_2^{(1)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \alpha_{m2}^{(1)} & \cdots & \alpha_{mn}^{(1)} & \beta_m^{(1)} \end{array}\right). \tag{1}$$

  Das LGS $\mathcal{A}^{(1)}\vec{x} = \vec{b}^{(1)}$ hat dann dieselbe Lösungsmenge wie $\mathcal{A}\vec{x} = \vec{b}$, denn jede elementare Zeilenumformung lässt sich ja durch passende elementare Zeilenumformungen wieder rückgängig machen. Enthält die erste Spalte von $\mathcal{A}$ nur Nullen, so vertausche man sie zuvor mit einer Spalte von $\mathcal{A}$, wobeidie Spalte $\vec{b}$ nicht in Betracht kommt, welche mindestens ein von 0 verschiedenes Element enthält. Diese Umformung muss registriert werden, denn sie bedeutet eine Umnummerierung der Unbekannten. Die so erhaltene Matrix bringe man dann durch elementare Zeilenumformungen auf die Gestalt (1). Ist $\mathcal{A} = 0^{(m,n)}$, so ist man von vorneherein fertig.

- <u>2. Schritt:</u>
  Falls in der Matrix $\mathcal{A}^{(1)}$ in der zweiten Spalte unter den $\alpha_{k2}^{(1)}$ für $k = 2, \ldots, m$ mindestens einvon 0 verschiedenes Element vorkommt, bringe man $(\mathcal{A}^{(1)}|\vec{b}^{(1)})$ durch elementare Zeilenumformungen in die Gestalt

$$(\mathcal{A}^{(2)}|\vec{b}^{(2)}) = \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_{13}^{(2)} & \cdots & \alpha_{1n}^{(2)} & \beta_1^{(2)} \\ 0 & 1 & \alpha_{23}^{(2)} & \cdots & \alpha_{2n}^{(2)} & \beta_2^{(2)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \alpha_{m3}^{(2)} & \cdots & \alpha_{mn}^{(2)} & \beta_m^{(2)} \end{array}\right). \tag{2}$$

  Sind $\alpha_{22}^{(1)}, \ldots, \alpha_{m2}^{(1)} = 0$, dann vertausche man die zweite Spalte von $\mathcal{A}^{(1)}$ mit einer Spalte der Nummer l_0 mit $3 \leq l_0 \leq n$, welche mindestens ein von 0 verschiedenes $\alpha_{k_0,l_0}^{(1)}$ für $k_0 \geq 2$ enthält. Danach bringe man die Matrix durch elementare Zeilenumformungen in die Form (2). Ist jedoch die gesamte Teilmatrix $(\alpha_{kl}^{(1)})$ für $2 \leq k \leq m$ und $2 \leq l \leq n$ Null, so ist man fertig. So fortfahrend, gelangt man zu einer Matrix der Gestalt

$$(\mathcal{A}^{(p)}|\vec{b}^{(p)}) = \left(\begin{array}{cccccccc|c} 1 & 0 & 0 & \cdots & 0 & \alpha_{1,p+1}^{(p)} & \cdots & \alpha_{1,n}^{(p)} & \beta_1^{(p)} \\ 0 & 1 & 0 & \cdots & 0 & \alpha_{2,p+1}^{(p)} & \cdots & \alpha_{2,n}^{(p)} & \beta_2^{(p)} \\ 0 & 0 & 1 & \cdots & 0 & \alpha_{3,p+1}^{(p)} & \cdots & \alpha_{3,n}^{(p)} & \beta_3^{(p)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \alpha_{p,p+1}^{(p)} & \cdots & \alpha_{p,n}^{(p)} & \beta_p^{(p)} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_{p+1}^{(p)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_m^{(p)} \end{array}\right). \tag{p}$$

  Die Lösungsmenge des LGS $\mathcal{A}^{(p)}\vec{x} = \vec{b}$ stimmt nun bis auf eventuelle Umnummerierung der Unbekannten mit der von $\mathcal{A}\vec{x} = \vec{b}$ überein. An den letzten $m - p$ Zeilen kann die Lösbarkeit des Systems abgelesen werden:

– Fall 1:
Ist mindestens eines der $\beta_{p+1}^{(p)}, \dots, \beta_m^{(p)}$ von 0 verschiedenl, so ist das LGS mit

$$\mathrm{rg}(\mathcal{A}|\vec{b}) = \mathrm{rg}(\mathcal{A}^{(p)}|\vec{b}^{(p)}) = p + 1 = \mathrm{rg}(\mathcal{A}^{(p)}) + 1 = \mathrm{rg}(\mathcal{A}) + 1$$

nicht lösbar.

– Fall 2:
Ist dagegen $\beta_{p+1}^{(p)} = \dots = \beta_m^{(p)} = 0$, so ist das LGS mit

$$\mathrm{rg}(\mathcal{A}|\vec{b}) = \mathrm{rg}(\mathcal{A}^{(p)}|\vec{b}^{(p)}) = p = \mathrm{rg}(\mathcal{A}^{(p)}) = \mathrm{rg}(\mathcal{A})$$

lösbar.

- Explizite Angabe der Lösungsmenge:
Der Einfachheit halber sei angenommen, dass keine Spaltenvertauschungen stattgefunden haben. Das LGS $\mathcal{A}^{(p)}\vec{x} = \vec{b}^{(p)}$ lautet explizit

$$\begin{array}{ccccccccccc}
x_1 & & & + & \alpha_{1,p+1}^{(p)} x_{p+1} & + & \cdots & + & \alpha_{1,n}^{(p)} x_n & = & \beta_1^{(p)} \\
& x_2 & & + & \alpha_{2,p+1}^{(p)} x_{p+1} & + & \cdots & + & \alpha_{2,n}^{(p)} x_n & = & \beta_2^{(p)} \\
& & \ddots & & \vdots & & & & \vdots & & \vdots \\
& & x_p & + & \alpha_{p,p+1}^{(p)} x_{p+1} & + & \cdots & + & \alpha_{p,n}^{(p)} x_n & = & \beta_p^{(p)}
\end{array}$$

und kann zu

$$\begin{array}{ccccccccc}
x_1 & = & \beta_1^{(p)} & - & \alpha_{1,p+1}^{(p)} x_{p+1} & - & \cdots & - & \alpha_{1,n}^{(p)} x_n \\
x_2 & = & \beta_2^{(p)} & - & \alpha_{2,p+1}^{(p)} x_{p+1} & - & \cdots & - & \alpha_{2,n}^{(p)} x_n \\
\vdots & & \vdots & & \vdots & & & & \vdots \\
x_p & = & \beta_p^{(p)} & - & \alpha_{p,p+1}^{(p)} x_{p+1} & - & \cdots & - & \alpha_{p,n}^{(p)} x_n
\end{array}$$

umgeschrieben werden. Dabei sind die $x_{p+1}, \dots, x_n$ frei wählbare Parameter. Der allgemeine Lösungsvektor $\vec{x}$ hat also die Form

$$\begin{aligned}
\vec{x} &= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \beta_1^{(p)} - \alpha_{1,p+1}^{(p)} x_{p+1} - \dots - \alpha_{1,n}^{(p)} x_n \\ \beta_2^{(p)} - \alpha_{2,p+1}^{(p)} x_{p+1} - \dots - \alpha_{2,n}^{(p)} x_n \\ \vdots \\ \beta_p^{(p)} - \alpha_{p,p+1}^{(p)} x_{p+1} - \dots - \alpha_{p,n}^{(p)} x_n \\ x_{p+1} \\ \vdots \\ x_n \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} \beta_1^{(p)} \\ \vdots \\ \beta_p^{(p)} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{\text{partikuläre Lösung von } \mathcal{A}\vec{x} = \vec{b}} \underbrace{+ x_{p+1} \begin{pmatrix} -\alpha_{1,p+1}^{(p)} \\ \vdots \\ -\alpha_{p,p+1}^{(p)} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_{p+2} \begin{pmatrix} -\alpha_{1,p+2}^{(p)} \\ \vdots \\ -\alpha_{p,p+2}^{(p)} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} -\alpha_{1,n}^{(p)} \\ \vdots \\ -\alpha_{p,n}^{(p)} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}}_{\text{Basis des Lösungsraums von } \mathcal{A}\vec{x} = \vec{0}} .
\end{aligned}$$

Bemerkung 6.2.1. Betrachte ein LGS $\mathcal{A}\vec{x} = \vec{b}$ simultan für mehrere rechte Seiten:

$$\mathcal{A}\vec{x}_1 = \vec{b}_1, \quad \mathcal{A}\vec{x}_2 = \vec{b}_2, \quad \dots \quad \mathcal{A}\vec{x}_q = \vec{b}_q$$

oder äquivalent dazu die Matrixgleichung $\mathcal{A}\mathcal{X} = \mathcal{B}$ mit $\mathcal{X} = (\vec{x}_1, \dots, \vec{x}_q)$ und $\mathcal{B} = (\vec{b}_1, \dots, \vec{b}_q)$. Diesen Fall behandelt man analog durch Umformung der Matrix $(\mathcal{A}|\vec{b}_1, \dots, \vec{b}_q) = (\mathcal{A}|\mathcal{B})$. Ein Spezialfall dieses Problems ist uns schon in Satz 4.3.5 bei der Bestimmung der inversen Matrix $\mathcal{A}^{-1}$ begegnet, was ja auf die Lösung der Gleichung $\mathcal{A}\mathcal{X} = E_n$ hinausläuft.

Bemerkung 6.2.2. Es genügt auch, $(\mathcal{A}|\vec{b})$ durch Zeilenumformungen in die folgende allgemeinere Form zu bringen:

$$\left(\begin{array}{cccccccc|c}
\alpha'_{11} & * & * & \cdots & * & * & \cdots & * & \beta'_1 \\
0 & \alpha'_{22} & * & \cdots & * & * & \cdots & * & \beta'_2 \\
0 & 0 & \alpha'_{33} & \cdots & * & * & \cdots & * & \beta'_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \alpha'_{pp} & * & \cdots & * & \beta'_p \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta'_{p+1} \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta'_m
\end{array}\right),$$

wobei $*$ für beliebige Einträge aus K stehe und $\alpha'_{jj} \neq 0$ sei.
Im Falle der Lösbarkeit, d.h. für $\beta'_{p+1} = \cdots = \beta'_m = 0$, erhält man die Lösungen durch sukzessives Auflösen der Gleichungen von unten nach oben.

Beispiel 6.2.3. Wir lösen das System

$$\left\{\begin{array}{rcrcrcrcr}
x & + & 2y & - & 3z & + & 4w & = & 7 \\
2x & + & 4y & - & 6z & + & 10w & = & 20 \\
-x & - & 2y & + & 9z & - & 9w & = & -10
\end{array}\right.$$

Es gilt

$$(\mathcal{A}|\vec{b}) = \left(\begin{array}{rrrr|r}
1 & 2 & -3 & 4 & 7 \\
2 & 4 & -6 & 10 & 20 \\
-1 & -2 & 9 & -9 & \text{-10}
\end{array}\right) \longrightarrow \begin{array}{c}\left(\begin{array}{rrrr|r}
1 & 2 & -3 & 4 & 7 \\
0 & 0 & 0 & 2 & 6 \\
0 & 0 & 6 & -5 & \text{-3}
\end{array}\right) \\ \begin{array}{cccc} x & y & z & w \end{array}\end{array}$$

$$\xrightarrow{\text{Spaltentausch}} \begin{array}{c}\left(\begin{array}{rrrr|r}
1 & 4 & -3 & 2 & 7 \\
0 & 2 & 0 & 0 & 6 \\
0 & -5 & 6 & 0 & \text{-3}
\end{array}\right) \\ \begin{array}{cccc} x & w & z & y \end{array}\end{array} \longrightarrow \left(\begin{array}{rrrr|r}
1 & 0 & -3 & 2 & \text{-5} \\
0 & 1 & 0 & 0 & 3 \\
0 & 0 & 6 & 0 & 12
\end{array}\right)$$

$$\longrightarrow \left(\begin{array}{cccc|c}
1 & 0 & 0 & 2 & 1 \\
0 & 1 & 0 & 0 & 3 \\
0 & 0 & 1 & 0 & 2
\end{array}\right) \Rightarrow \left\{\begin{array}{rcc}
x & = & 1 - 2y \\
w & = & 3 \\
z & = & 2
\end{array}\right.$$

mit einem beliebigen $y \in \mathbb{R}$.
Damit ergeben sich sämtliche Lösungsvektoren in der Form

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 3 \end{pmatrix} + y \cdot \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

mit $y \in \mathbb{R}$.

Beispiel 6.2.4. Bestimme eine Gleichung der Ebene E des $\mathbb{R}^3$, die durch die Punkte $P = (1, -1, 2)$, $Q = (2, 4, -5)$ und $R = (3, 1, 0)$ verläuft.

Lösung:
Da die Ebene E eine zweidimensionale lineare Mannigfaltigkeit des dreidimensionalen $\mathbb{R}^3$ ist, gibt es nach Satz 6.1.14 eine Linearform $\varphi \in L(\mathbb{R}^3, \mathbb{R})$ und ein $d \in \mathbb{R}$ mit

$$\vec{v} \in E \Leftrightarrow \varphi(\vec{v}) = d.$$

Diese hat die Form $\varphi(x, y, z) = ax + by + cz$. Die Bestimmung von φ geschieht nach der im Beweis von Satz 6.1.14 beschriebenen Methode. Die Richtungsvektoren von E sind $\overrightarrow{PQ} = (1, 5, -7)$ und $\overrightarrow{PR} = (2, 2, -2)$, also ist

$$E = \left\{ \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + t \begin{pmatrix} 1 \\ 5 \\ -7 \end{pmatrix} + u \begin{pmatrix} 2 \\ 2 \\ -2 \end{pmatrix} : t, u \in \mathbb{R} \right\}.$$

Die Koeffizienten (a, b, c) von φ bestimmen sich als Lösungen des LGS

$$\begin{cases} a + 5b - 7c = 0 \\ 2a + 2b - 2c = 0 \end{cases} \Leftrightarrow \begin{cases} a + 5b - 7c = 0 \\ \quad - 8b + 12c = 0 \end{cases}$$

Wir wählen $c = 2$, dann ist $b = 3$ und $a = -1$, sowie $\varphi(x, y, z) = -x + 3y + 2z$, denn die Menge der möglichen φ bildet nach Satz 6.1.14 einen eindimensionalen Vektorraum, weswegen φ nur bis auf einen konstanten Faktor bestimmt ist. Den verbleibenden Koeffizienten d in der Gleichung von E erhält man, indem man einen beliebigen Punkt von E in φ einsetzt: beispielsweise ist $\varphi(1, -1, 2) = 0$, also $d = 0$, und die Gleichung von E ist

$$-x + 3y + 2 = 0.$$

Wegen $d = 0$ geht E durch den Nullpunkt und ist sogar ein Untervektorraum von $\mathbb{R}^3$.

Beispiel 6.2.5. Finde ein Gleichungssystem für die dreidimensionale lineare Mannigfaltigkeit M des Vektorraums $\mathbb{R}^5$, die durch die Punkte

$$P_1 = (2, -1, 0, 3, 4), \quad P_2 = (1, 2, 1, 3, 5), \quad P_3 = (3, 1, 1, 4, 5) \quad \text{und} \quad P_4 = (2, 0, 1, 5, 0)$$

verläuft.

Lösung:
Wir schreiben M zunächst wieder als $M = \{\vec{v}_0 + \vec{u} \colon \vec{u} \in U\}$ mit einem Unterraum $U \subseteq \mathbb{R}^5$, welcher die Basis

$$\mathcal{B} = \{\overrightarrow{P_1P_2},\ \overrightarrow{P_1P_3},\ \overrightarrow{P_1P_4}\}$$

besitzt. Nach Satz 6.1.14 gibt es zwei Linearformen $\varphi_1, \varphi_2 \in L(M)$ und $d_1, d_2 \in \mathbb{R}$, so dass

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_5 \end{pmatrix} \in M \Leftrightarrow \varphi_1(\vec{x}) = d_1 \quad \text{und} \quad \varphi_2(\vec{x}) = d_2$$

gilt. Die φ_1, φ_2 können als beliebige Basis des zweidimensionalen Unterraums $L(M)$ von $L(\mathbb{R}^5, \mathbb{R})$ bestimmt werden. Die Koeffizienten $\alpha_1, \ldots, \alpha_5$ der $\varphi \in L(M)$ bestimmen sich als Lösungen des LGS

$$\begin{cases} -\alpha_1 + 3\alpha_2 + \alpha_3 + \alpha_5 = 0 \\ \alpha_1 + 2\alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 = 0 \\ \alpha_2 + \alpha_3 + 2\alpha_4 - 4\alpha_5 = 0 \end{cases}$$

Diese Koeffizienten sind die Komponenten der Vektoren der Basis $\mathcal{B}$. Der Gaußsche Algorithmus liefert

$$\begin{pmatrix} -1 & 3 & 1 & 0 & 1 \\ 1 & 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & -4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -3 & -1 & 0 & -1 \\ 0 & 5 & 2 & 1 & 2 \\ 0 & 1 & 1 & 2 & -4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & -3 & -1 & 0 & -1 \\ 0 & 1 & 1 & 2 & -4 \\ 0 & 0 & -3 & -9 & 22 \end{pmatrix}.$$

Die Wahlen $\alpha_4 = 0$, $\alpha_5 = -3$ bzw. $\alpha_4 = 1$, $\alpha_5 = 0$ liefern zwei linear unabhängige Linearformen φ_1 und φ_2 mit

$$\begin{aligned} \alpha_4 = 0, \alpha_5 = -3 \quad &\Rightarrow \quad \alpha_3 = -22, \alpha_2 = 10, \alpha_1 = 5 \\ \alpha_4 = 1, \alpha_5 = 0 \quad &\Rightarrow \quad \alpha_3 = -3, \alpha_2 = 1, \alpha_1 = 0. \end{aligned}$$

Also ist $\varphi_1(x_1, \ldots, x_5) = 5x_1 + 10x_2 - 22x_3 - 3x_5$ und $\varphi_2(x_1, \ldots, x_5) = x_2 - 3x_3 + x_4$. Die Zahlen $d_1, d_2 \in \mathbb{R}$ ergeben sich wiederum durch Einsetzen eines beliebigen Punktes von M in φ_1 und φ_2, beispielsweise ist

$$\varphi_1(2, -1, 0, 3, 4) = -12 \quad \text{und} \quad \varphi_2(2, -1, 0, 3, 4) = 2.$$

Also ist M durch das LGS

$$\left\{ \begin{array}{rcrcrcrcrcr} 5x_1 & + & 10x_2 & - & 22x_3 & & & - & 3x_5 & = & -12 \\ & & x_2 & - & 3x_3 & + & x_4 & & & = & 2 \end{array} \right.$$

bestimmt. Wegen $d_1, d_2 \neq 0$ ist M kein Untervektorraum.

Kapitel 7

Determinanten

7.1 Permutationen

Definition 7.1.1. Eine Permutation σ der Menge $N = \{1, 2, \ldots, n\}$ ist eine bijektive Abbildung $\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Unter γ_n verstehen wir die Gruppe der Permutationen von N mit der Komposition als Verknüpfung (siehe Beispiel 2.2.7 für den Fall $n = 3$).
Wir schreiben Permutationen in der Form

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

Eine Permutation $\tau \in \gamma_n$, welche nur zwei Zahlen $i \neq j$ vertauscht, also $\tau(i) = j$, $\tau(j) = i$ und $\tau(k) = k$ für $k \notin \{i, j\}$, heißt Transposition. Wir schreiben sie in der Form $\tau = (i\ j)$.

Satz 7.1.2. *Es ist* $|\gamma_n| = n!$

Beweis. Übungsaufgabe. □

Definition 7.1.3. Es sei $\sigma \in \gamma_n$. Ein Paar (i, j) mit $1 \leq i < j \leq n$ heißt Fehlstandspaar (oder Inversion) der Permutation σ, falls $\sigma(i) > \sigma(j)$ ist. Die Anzahl der Fehlstandspaare von σ heißt Fehlstandszahl $\phi(\sigma)$. Die Parität oder das Signum von σ ist durch $\operatorname{sgn}(\sigma) = (-1)^{\phi(\sigma)}$ definiert. Dabei gilt

$$\begin{array}{llclclcl} \sigma & \text{heißt} & \underline{\text{gerade}} & \Leftrightarrow & \operatorname{sgn}(\sigma) = & 1 & \Leftrightarrow & \phi(\sigma) \quad \text{gerade}, \\ \sigma & \text{heißt} & \underline{\text{ungerade}} & \Leftrightarrow & \operatorname{sgn}(\sigma) = & -1 & \Leftrightarrow & \phi(\sigma) \quad \text{ungerade}. \end{array}$$

Beispiel 7.1.4. Es sei $\sigma \in \gamma_5$ durch

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{pmatrix}.$$

gegeben. Dann hat σ die Fehlstandspaare $(1,3)$,$(1,4)$,$(2,3)$,$(2,4)$,$(2,5)$. Es ist also $\phi(\sigma) = 5$ und $\operatorname{sgn}(\sigma) = (-1)^5 = -1$, womit σ eine ungerade Permutation ist.

Satz 7.1.5. *Für jedes* $\sigma \in \gamma_n$ *gilt:*

i) Es kann σ *als Produkt von Transpositionen geschrieben werden:* $\sigma = \tau_1 \circ \cdots \circ \tau_r$.

ii) Es ist $\operatorname{sgn}(\sigma) = (-1)^r$.

iii) Insbesondere gilt $\operatorname{sgn}(\varrho \circ \sigma) = \operatorname{sgn}(\varrho) \cdot \operatorname{sgn}(\sigma)$ *für alle* $\varrho, \sigma \in \gamma_n$.

Bemerkung 7.1.6. Die Anzahl r ist durch σ nicht eindeutig bestimmt. Es ist aufgrund von

$$\operatorname{sgn}(\sigma) = (-1)^r = \begin{cases} 1, & \text{falls } r \text{ gerade ist} \\ -1, & \text{falls } r \text{ ungerade ist} \end{cases}$$

aber bestimmt, ob r gerade oder ungerade ist.

Beweis. (Beweis von Satz 7.1.5)

i) Wir zeigen zunächst:
Ist $\sigma \in \gamma_n$ mit $\phi(\sigma) = m$ und $\tau = (\sigma(h), \sigma(h+1)) \in \gamma_n$ mit $1 \leq h \leq n-1$, dann gilt

$$\phi(\tau \circ \sigma) = \begin{cases} m-1, & \text{falls } (\sigma(h), \sigma(h+1)) \text{ ein Fehlstandspaar von } \sigma \text{ ist,} \\ m+1, & \text{falls } (\sigma(h), \sigma(h+1)) \text{ kein Fehlstandspaar von } \sigma \text{ ist.} \end{cases} \qquad (*)$$

Dazu vergleichen wir die Fehlstandspaare der Permutationen

$$\begin{aligned} \sigma &= \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix} \quad \text{und} \\ \varrho &= \tau \circ \sigma = \begin{pmatrix} 1 & \cdots & h-1 & h & h+1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(h-1) & \sigma(h+1) & \sigma(h) & \cdots & \sigma(n) \end{pmatrix}. \end{aligned}$$

Ist $i \notin \{h, h+1\}$, so ist (i, h) genau dann ein Fehlstandspaar von σ, wenn $(i, h+1)$ ein Fehlstandspaar von ϱ ist. Ebenso ist (i, h) genau dann ein Fehlstandspaar von ϱ, wenn $(i, h+1)$ ein Fehlstandspaar von σ ist. Unter den Paaren $(i, j) \neq (h, h+1)$ haben ϱ und σ daher dieselbe Anzahl von Fehlstandspaaren. Ist $(h, h+1)$ ein Fehlstandspaar von σ, so ist es keines von ϱ. Ist $(h, h+1)$ kein Fehlstandspaar von σ, so ist eines von ϱ. Damit ist $(*)$ gezeigt.
Wir führen nun den Beweis von (i) durch Induktion nach der Zahl m der Fehlstandspaare. Ist $\phi(\sigma) = 0$, so muss $\sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma(n)$ sein, also

$$\sigma = id = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Die Identität kann als leeres Produkt von Transpositionen geschrieben werden, oder auch als $id = (1, 2) \circ (1, 2)$.
Nun sei die Behauptung i) schon für die Fehlstandszahl $m-1$ bewiesen.
Es sei $\sigma \in \gamma_n$ mit $\phi(\sigma) = m$ mit einem Fehlstandspaar $(h, h+1)$ sowie $\tau = (\sigma(h), \sigma(h+1))$. Wegen $(*)$ ist dann $\phi(\sigma \circ \tau) = m-1$. Nach Induktionsannahme ist $\tau \circ \sigma = \tau \circ \tau_1 \circ \ldots \circ \tau_k = \varrho$ für gewisse Transpositionen τ_i und somit wegen $\tau \circ \tau = id$

$$\sigma = \tau \circ \varrho = \tau \circ \tau_1 \circ \ldots \circ \tau_k.$$

ii) Sei $\tau = (i, j) \in \gamma_n$ mit $i < j$. Die Vertauschung von i und j kann folgendermaßen erreicht werden: Durch $(j-i)$ Vertauschungen benachbarter Elemente wird i an die j-te Stelle, und j an die $(j-1)$-te Stelle gebracht. Danach wird j durch $(j-i-1)$ Vertauschungen mit benachbarten Elementen an die i-te Stelle gebracht.

Skizze:

$$
\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & i & \cdots & j & \cdots & n \end{pmatrix}
\overset{(i\ i+1)}{\rightarrow}
\begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i & \cdots & j & \cdots & n \end{pmatrix}
$$

$$
\overset{(i+1\ i+2)}{\rightarrow}
\begin{pmatrix} 1 & 2 & \cdots & i & i+1 & i+2 & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i+2 & i & \cdots & j & \cdots & n \end{pmatrix}
$$

$$
\rightarrow
\begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & j-1 & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i+2 & \cdots & i & j & \cdots & n \end{pmatrix}
$$

$$
\overset{(j-1\ j)}{\rightarrow}
\begin{pmatrix} 1 & 2 & \cdots & i & i+1 & \cdots & j-1 & j & \cdots & n \\ 1 & 2 & \cdots & i+1 & i+2 & \cdots & j & i & \cdots & n \end{pmatrix}
$$

$$
\overset{(j-1\ j-2)}{\rightarrow}
\begin{pmatrix} 1 & 2 & \cdots & j-2 & j-1 & j & \cdots & n \\ 1 & 2 & \cdots & j & j-1 & i & \cdots & n \end{pmatrix}
$$

$$
\rightarrow
\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & j & \cdots & n \\ 1 & 2 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}.
$$

Wir haben also die Produktdarstellung

$$\tau = (i, i+1) \circ \cdots \circ (j-1, j-2) \circ (j-1, j) \circ \ldots \circ (i+1, i+2) \circ (i, i+1),$$

ein Produkt von $2(j-i)-1$ Transpositionen der Form $(\sigma(h), \sigma(h+1))$, also einer ungeraden Anzahl solcher Transpositionen. Sei nun $\sigma = \tau_1 \circ \cdots \circ \tau_r$. Jedes τ_i schreiben wir, wie eben skizziert, als Produkt einer ungeraden Anzahl von Transpositionen der Form $(\sigma(h), \sigma(h+1))$. Dies führt zu einer Produktdarstellung $\sigma = \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_s$, wobei jedes $\tilde{\tau}_j$ von der Form $(\sigma(h), \sigma(h+1))$ ist. s ist gerade falls r gerade ist, bzw. ungerade falls r ungerade ist. Nach $(*)$ unterscheiden sich $\phi(\varrho)$ und $\phi(\tilde{\tau} \circ \varrho)$ um ± 1. Daher ist $\phi(\sigma)$ gerade, falls s und somit r gerade ist, und andernfalls ungerade. Daraus folgt ii).

iii) Sei $\sigma = \tau_1 \circ \cdots \circ \tau_r$ und $\varrho = \nu_1 \circ \cdots \circ \nu_s$ mit Transpositionen τ_i, ν_j. Dann ist

$$\sigma \circ \varrho = \tau_1 \circ \ldots \circ \tau_r \circ \nu_1 \circ \ldots \circ \nu_s.$$

Nach (ii) gilt dann $\mathrm{sgn}(\sigma) = (-1)^r$, $\mathrm{sgn}(\varrho) = (-1)^s$ sowie $\mathrm{sgn}(\sigma \circ \varrho) = (-1)^{r+s}$, woraus die Behauptung $\mathrm{sgn}(\sigma \circ \varrho) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\varrho)$ folgt.

□

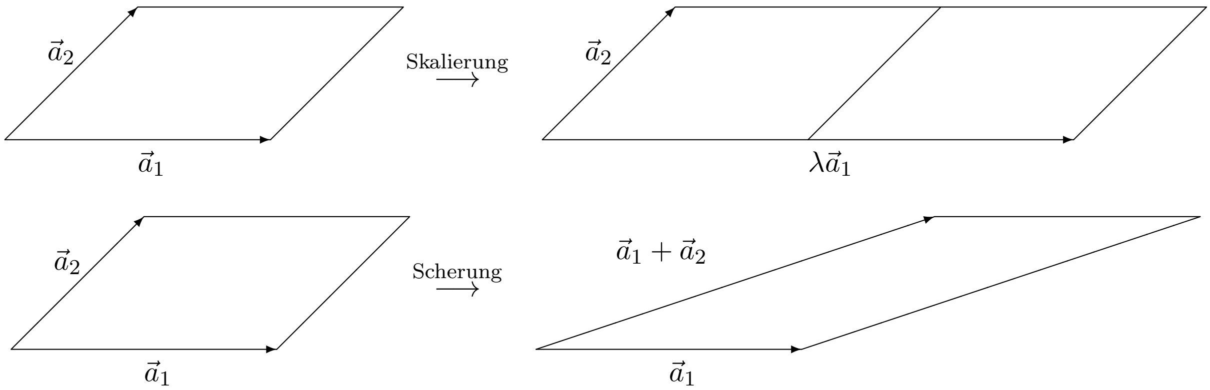
7.2 Determinantenfunktionen

Eine Motivation für die Einführung einer Determinantenfunktion liegt im Wunsch, das Volumen eines von n Vektoren $\vec{a}_1, \ldots, \vec{a}_n$ des $\mathbb{R}^n$ aufgespannten Parallelepipeds zu definieren. Wir beschränken uns bei der Illustration der Einfachheit halber auf den Fall $n = 2$, also den Flächeninhalt eines Parallelogramms. Es sei $|f(\vec{a}_1, \vec{a}_2)|$ die Fläche des von $\vec{a}_1, \vec{a}_2 \in \mathbb{R}^2$ aufgespannten Parallelogramms. Dann gilt offenbar

(H) Homogenität: $f(\lambda\vec{a}_1, \vec{a}_2) = f(\vec{a}_1, \lambda\vec{a}_2) = \lambda f(\vec{a}_1, \vec{a}_2)$ für alle $\lambda \in \mathbb{R}$.

(S) Scherungsinvarianz: $f(\vec{a}_1 + \vec{a}_2, \vec{a}_2) = f(\vec{a}_1, \vec{a}_1 + \vec{a}_2) = f(\vec{a}_1, \vec{a}_2)$.

Skizze:



Wir verallgemeinern diese Eigenschaften nun auf beliebige Abbildungen $f\colon V^n \to K$, wobei V ein n-dimensionaler Vektorraum über einem Körper K und $V^n = V \times V \times \ldots \times V$ ist.

Definition 7.2.1. Eine Abbildung $f\colon V^n \to K$ heißt homogen, falls für alle $\vec{a}_1, \ldots, \vec{a}_n \in V$ und $i = 1, \ldots, n$

$$f(\vec{a}_1, \ldots, \vec{a}_{i-1}, \lambda\vec{a}_i, \vec{a}_{i+1}, \ldots, \vec{a}_n) = \lambda f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$$

gilt. Die Abbildung f heißt scherungsinvariant, falls für alle i, k mit $i \neq k$ und alle $\vec{a}_1, \ldots, \vec{a}_n \in V$

$$f(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{a}_i + \vec{a}_k, \vec{a}_{i+1}, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$$

gilt.

Im folgenden sei V ein K-Vektorraum.

Satz 7.2.2. *Es sei* $\dim V = n$ *und* $f\colon V^n \to K$ *eine homogene und scherungsinvariante Abbildung. Dann gilt:*

i) $f(\vec{a}_1, \ldots, \vec{a}_{i-1}, \vec{0}, \vec{a}_{i+1}, \ldots, \vec{a}_n) = 0$.

ii) Für alle $i \neq k$ *und alle* $\lambda \in K$ *ist* $f(\vec{a}_1, \ldots, \vec{a}_i + \lambda\vec{a}_k, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$.

iii) Sind $\vec{a}_1, \ldots, \vec{a}_n$ *linear abhängig, so ist* $f(\vec{a}_1, \ldots, \vec{a}_n) = 0$.

iv) Die Abbildung f *ist in jedem Argument additiv, d.h. für alle* i *gilt*

$$f(\vec{a}_1, \ldots, \vec{a}_i + \vec{b}_i, \ldots \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) + f(\vec{a}_1, \ldots, \vec{b}_i, \ldots, \vec{a}_n).$$

v) Die Abbildung f *ist in jedem Argumentpaar alternierend, d.h. für alle* $i \neq j$ *ist*

$$f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n) = -f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n).$$

Beweis. i) Wegen der Homogenität gilt

$$f(\vec{a}_1, \ldots, \vec{0}, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, 0 \cdot \vec{0}, \ldots, \vec{a}_n) \underset{(\mathrm{H})}{=} 0 \cdot f(\vec{a}_1, \ldots, \vec{0}, \ldots, \vec{a}_n) = 0.$$

ii) Es sei ohne Einschränkung $\lambda \neq 0$. Dann gilt

$$\begin{aligned} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_j,\dots,\vec{a}_n) &\underset{(H)}{=} \lambda^{-1} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\lambda\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{(S)}{=} \lambda^{-1} f(\vec{a}_1,\dots,\vec{a}_i+\lambda\vec{a}_j,\dots,\lambda\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{(H)}{=} f(\vec{a}_1,\dots,\vec{a}_i+\lambda\vec{a}_j,\dots,\vec{a}_j,\dots,\vec{a}_n). \end{aligned}$$

iii) Sind $\vec{a}_1,\dots,\vec{a}_n$ linear abhängig, so ist $\lambda_1\vec{a}_1+\cdots+\lambda_n\vec{a}_n=\vec{0}$ für $\lambda_1,\dots,\lambda_n \in K$, und mindestens ein λ_i nicht null; ohne Einschränkung sei dies $\lambda_1 \neq 0$. Dann erhält man

$$\begin{aligned} \lambda_1 f(\vec{a}_1,\dots,\vec{a}_n) &\underset{(H)}{=} f(\lambda_1\vec{a}_1,\dots,\vec{a}_n) \\ &\underset{(ii)}{=} f(\lambda_1\vec{a}_1+\lambda_2\vec{a}_2,\vec{a}_2,\dots,\vec{a}_n) \\ &\vdots \\ &\underset{(ii)}{=} f(\lambda_1\vec{a}_1+\cdots+\lambda_n\vec{a}_n,\vec{a}_2,\dots,\vec{a}_n) \;=\; f(\vec{0},\vec{a}_2,\dots,\vec{a}_n)=0 \end{aligned}$$

und somit $f(\vec{a}_1,\dots,\vec{a}_n)=0$.

iv) Wir unterscheiden hier drei Fälle:

- $\vec{a}_1,\dots,\vec{a}_n$ sind linear unbhängig:
  Dann ist $\{\vec{a}_1,\dots,\vec{a}_n\}$ eine Basis von V, also $\vec{b}_i=\lambda_1\vec{a}_1+\cdots+\lambda_n\vec{a}_n$ mit $\lambda_i \in K$. Es folgt

$$\begin{aligned} f(\vec{a}_1,\dots,\vec{a}_i+\vec{b}_i,\dots,\vec{a}_n) &= f(\vec{a}_1,\dots,\vec{a}_i+\lambda_1\vec{a}_1+\cdots+\lambda_i\vec{a}_i+\cdots+\lambda_n\vec{a}_n,\dots,\vec{a}_n) \\ &\underset{ii)}{=} f(\vec{a}_1,\dots,\vec{a}_i+\lambda_i\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{(H)}{=} (1+\lambda_i) f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &= f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+\lambda_i f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{(H)}{=} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+f(\vec{a}_1,\dots,\lambda_i\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{ii)}{=} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &\quad +f(\vec{a}_1,\dots,\lambda_1\vec{a}_1+\cdots+\lambda_i\vec{a}_i+\cdots+\lambda_n\vec{a}_n,\dots,\vec{a}_n) \\ &= f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+f(\vec{a}_1,\dots,\vec{b}_i,\dots,\vec{a}_n). \end{aligned}$$

- $\vec{a}_1,\dots,\vec{a}_{i-1},\vec{a}_{i+1},\dots,\vec{a}_n$ sind linear unabhängig, aber $\vec{a}_1,\dots,\vec{a}_n$ sind linear abhängig:
  Dann ist $\vec{a}_i$ eine Linearkombination von $\vec{a}_1,\dots,\vec{a}_{i-1},\vec{a}_{i+1},\dots,\vec{a}_n$, also gilt

$$f(\vec{a}_1,\dots,\vec{a}_i+\vec{b}_i,\dots,\vec{a}_n) \underset{ii)}{=} f(\vec{a}_1,\dots,\vec{b}_i,\dots,\vec{a}_n)=f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)+f(\vec{a}_1,\dots,\vec{b}_i,\dots,\vec{a}_n),$$

  da nach (iii) $f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_n)=0$ ist.
- $\vec{a}_1,\dots,\vec{a}_{i-1},\vec{a}_{i+1},\dots,\vec{a}_n$ sind linear abhängig:
  In diesem Fall ist die Aussage iv) richtig, weil alle f-Werte nach iii) verschwinden.

v) Es gilt

$$\begin{aligned} f(\vec{a}_1,\dots,\vec{a}_i,\dots,\vec{a}_j,\dots,\vec{a}_n) &\underset{(H)}{=} -f(\vec{a}_1,\dots,\vec{a}_i,\dots,-\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{ii)}{=} -f(\vec{a}_1,\dots,\vec{a}_i+\vec{a}_j,\dots,-\vec{a}_j,\dots,\vec{a}_n) \\ &\underset{ii)}{=} -f(\vec{a}_1,\dots,\vec{a}_i+\vec{a}_j,\dots,\vec{a}_i,\dots,\vec{a}_n) \\ &\underset{ii)}{=} -f(\vec{a}_1,\dots,\vec{a}_j,\dots,\vec{a}_i,\dots,\vec{a}_n). \end{aligned}$$

$\square$

Definition 7.2.3. Es sei $\dim V = n$ und $f\colon V^n \to K$ eine Abbildung.

i) Dann heißt f eine <u>n-fache Linearform</u>, wenn f in jedem Argument linear ist, d.h. wenn

$$f(\vec{a}_1, \ldots, \lambda\vec{a}_i + \mu\vec{b}_i, \ldots, \vec{a}_n) = \lambda f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) + \mu f(\vec{a}_1, \ldots, \vec{b}_i, \ldots, \vec{a}_n)$$

für alle $\vec{a}_1, \ldots, \vec{a}_n, \vec{b}_i \in V$ und $\lambda, \mu \in K$ gilt.

ii) Weiter heißt f <u>alternierend</u>, falls für alle $i \neq j$ und alle $\vec{a}_1, \ldots, \vec{a}_n \in V$

$$f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n) = -f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n)$$

gilt.

Satz 7.2.4. *Es sei $\dim V = n$ und $f\colon V^n \to K$ eine Abbildung. Dann sind folgende Aussagen äquivalent:*

i) f ist homogen und scherungsinvariant.

ii) f ist eine n-fache alternierende Linearform.

Beweis. $i) \Rightarrow ii)$:
Aus der Homogenität und der Additivität (Satz 7.2.2 iv)) folgt die Linearität. Nach 7.2.2 v) ist f auch alternierend.
$ii) \Rightarrow i)$:
Es sei f eine n-fache alternierende Linearform. Dann ist

$$f(\vec{a}_1, \ldots, \vec{a}_i + \vec{a}_j, \ldots, \vec{a}_n) = f(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n) + f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_j, \ldots, \vec{a}_n).$$

Da f alternierend ist, folgt $f(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_j, \ldots, \vec{a}_n) = 0$ und damit die Scherungsinvarianz von f. Die Homogenität von f ergibt sich unmittelbar aus der Linearität. $\square$

Vertauschung zweier Argumente einer alternierenden Linearform ändert den Wert von f um den Faktor -1. Wir wollen nun untersuchen, wie sich der Wert bei einer beliebigen Umordnung der Argumente verhält:

Satz 7.2.5. *Es sei f eine n-fache alternierende Linearform auf V und $\sigma \in \gamma_n$. Für alle $\vec{a}_1, \ldots, \vec{a}_n \in V$ ist dann*

$$f(\vec{a}_{\sigma(1)}, \ldots, \vec{a}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot f(\vec{a}_1, \ldots, \vec{a}_n).$$

Beispiel 7.2.6. Es sei f eine 3-fache alternierende Linearform und $\dim V = 3$. Für $\vec{a}_1, \vec{a}_2, \vec{a}_3 \in V$ ist

$$f(\vec{a}_2, \vec{a}_3, \vec{a}_1) = f(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \vec{a}_{\sigma(3)}) \quad \text{mit} \quad \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Die Permutation σ hat die Fehlstandspaare $(1,3)$ und $(2,3)$, also $\operatorname{sgn}(\sigma) = 1$. Daher ist

$$f(\vec{a}_2, \vec{a}_3, \vec{a}_1) = \operatorname{sgn}(\sigma) f(\vec{a}_1, \vec{a}_2, \vec{a}_3) \;=\; f(\vec{a}_1, \vec{a}_2, \vec{a}_3).$$

Beweis. (Beweis von Satz 7.2.5)
Wir schreiben σ als Produkt von Transpositionen $\sigma = \tau_1 \circ \cdots \circ \tau_r$, was nach Satz 7.3.2 möglich ist. Für $1 \le j \le r$ setzen wir $\sigma_j = \tau_j \circ \cdots \circ \tau_r$. Nun ist

$$f(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \ldots, \vec{a}_{\sigma(n)}) = f(\vec{a}_{(\tau_1\circ\sigma_2)(1)}, \vec{a}_{(\tau_1\circ\sigma_2)(2)}, \ldots, \vec{a}_{(\tau_1\circ\sigma_2)(n)}) = -f(\vec{a}_{\sigma_2(1)}, \vec{a}_{\sigma_2(2)}, \ldots, \vec{a}_{\sigma_2(n)}),$$

da f alternierend ist, und τ_1 zwei Argumente tauscht. Wiederholtes Abspalten der τ_j ergibt

$$\begin{aligned}
f(\vec{a}_{\sigma(1)}, \vec{a}_{\sigma(2)}, \ldots, \vec{a}_{\sigma(n)}) &= (-1)^1 f(\vec{a}_{\sigma_2(1)}, \vec{a}_{\sigma_2(2)}, \ldots, \vec{a}_{\sigma_2(n)}) \\
&= (-1)^2 f(\vec{a}_{\sigma_3(1)}, \vec{a}_{\sigma_3(2)}, \ldots, \vec{a}_{\sigma_3(n)}) \\
&\vdots \\
&= (-1)^r f(\vec{a}_{\sigma_n(1)}, \ldots, \vec{a}_{\sigma_n(n)}) = (-1)^r f(\vec{a}_1, \ldots, \vec{a}_n).
\end{aligned}$$

Nach Satz 7.1.5 ist $(-1)^r = \operatorname{sgn}(\sigma)$. □

Satz 7.2.7. *Es sei f eine n-fache alternierende Linearform auf V. Dann gilt:*

i) Sind $\vec{a}_1, \ldots, \vec{a}_n, \vec{b}_1, \ldots, \vec{b}_n \in V$ und $\lambda_{ik} \in K$ mit $\vec{a}_i = \sum_{k=1}^{n} \lambda_{ik}\vec{b}_k$, dann ist

$$f(\vec{a}_1, \ldots, \vec{a}_n) = f(\vec{b}_1, \ldots, \vec{b}_n) \cdot \sum_{\sigma\in\gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)}.$$

ii) Ist $f \neq 0$, so gilt

$$f(\vec{b}_1, \ldots, \vec{b}_n) \neq 0 \Leftrightarrow \vec{b}_1, \ldots, \vec{b}_n \textit{ sind linear unabhängig.}$$

Beweis. i) Wir wenden wiederholt die Linearität von f an:

$$\begin{aligned}
f(\vec{a}_1, \ldots, \vec{a}_n) &= f\left(\sum_{k=1}^{n} \lambda_{1,k}\vec{b}_k, \vec{a}_2, \ldots, \vec{a}_n\right) = \sum_{k=1}^{n} \lambda_{1,k} f(\vec{b}_k, \vec{a}_2, \ldots, \vec{a}_n) \\
&= \sum_{k_1=1}^{n} \lambda_{1,k_1} f\left(\vec{b}_{k_1}, \sum_{k_2=1}^{n} \lambda_{2,k_2}\vec{b}_{k_2}, \vec{a}_3, \ldots, \vec{a}_n\right) \\
&= \sum_{k_1=1}^{n}\sum_{k_2=1}^{n} \lambda_{1,k_1}\lambda_{2,k_2} f(\vec{b}_{k_1}, \vec{b}_{k_2}, \vec{a}_3, \ldots, \vec{a}_n) \\
&\vdots \\
&= \sum_{k_1=1}^{n} \cdots \sum_{k_n=1}^{n} \lambda_{1,k_1} \cdots \lambda_{n,k_n} f(\vec{b}_{k_1}, \vec{b}_{k_2}, \ldots, \vec{b}_{k_n}).
\end{aligned}$$

Da f alternierend ist, sind alle Summanden null, für die zwei k_i übereinstimmen. Wir brauchen also nur über die n-Tupel $(k_1, \ldots, k_n)$ zu summieren, für die $(k_1, \ldots, k_n) = (\sigma(1), \ldots, \sigma(n))$ für ein $\sigma \in \gamma_n$ ist. Also ist

$$\begin{aligned}
f(\vec{a}_1, \ldots, \vec{a}_n) &= \sum_{\sigma\in\gamma_n} \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)} f(\vec{b}_{\sigma(1)}, \vec{b}_{\sigma(2)}, \ldots, \vec{b}_{\sigma(n)}) \\
&= \sum_{\sigma\in\gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)} f(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n)
\end{aligned}$$

nach Satz 7.2.5.

ii) Sind $\vec{b}_1, \ldots, \vec{b}_n$ linear unabhängig, so bilden sie eine Basis von V. Für beliebige $\vec{a}_1, \ldots, \vec{a}_n \in V$ gibt es dann $\lambda_{ik} \in K$ für $1 \leq i, k \leq n$ mit $\vec{a}_i = \sum \lambda_{ik}\vec{b}$. Aus i) folgt dann

$$f(\vec{a}_1, \ldots, \vec{a}_n) = f(\vec{b}_1, \ldots, \vec{b}_n) \cdot \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)}.$$

Wäre $f(\vec{b}_1, \ldots, \vec{b}_n) = 0$, so gilt auch $f(\vec{a}_1, \ldots, \vec{a}_n) = 0$. Da die $\vec{a}_1, \ldots, \vec{a}_n \in V$ aber beliebig sind, ist f bereits die Nullabbildung. Sind nun andererseits $\vec{b}_1, \ldots, \vec{b}_n$ linear abhängig, so ist nach Satz 7.2.2 iii) dann $f(\vec{b}_1, \ldots, \vec{b}_n) = 0$.

□

Satz 7.2.8. *Es sei* $\dim V = n$. *Dann bildet die Menge der n-fach alternierenden Linearformen auf V mit werteweiser Addition und Skalarmultiplikation einen K-Vektorraum der Dimension* 1.

Beweis. Es sei $\mathcal{L}$ die Menge aller n-fachen Linearformen auf V. Man rechnet leicht nach, dass mit $f, g \in \mathcal{L}$ auch $\lambda f + \mu g \in \mathcal{L}$ ist. Weiter ist $\mathcal{L}$ auch nicht leer, da die Nullform mit $N(\vec{a}_1, \ldots, \vec{a}_n) = 0$ für alle $\vec{a}_1, \ldots, \vec{a}_n \in V$ stets in $\mathcal{L}$ liegt. Nach dem Unterraumkriterium ist $\mathcal{L}$ damit ein Untervektorraum von $\operatorname{Abb}(V^n, K)$. Die Abbildung $f_0 \colon V^n \to K$ werde folgendermaßen definiert: Es sei $\{\vec{b}_1, \ldots, \vec{b}_n\}$ eine fest gewählte Basis von V und

$$f_0(\vec{a}_1, \ldots, \vec{a}_n) = \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)}, \ \vec{a}_i = \sum_{k=1}^{n} \lambda_{ik}\vec{b}_k.$$

Wir zeigen, dass f_0 ein nichttriviales Element von $\mathcal{L}$ ist:

- f_0 ist n-fache Linearform:
  Es seien $\vec{a}_1, \ldots, \vec{a}_n \in V$ wie oben gegeben, und es sei für $1 \leq i \leq n$

  $$\vec{a}'_i = \sum_{k=1}^{n} \lambda'_{ik}\vec{b}_k.$$

  Dann gilt für $\mu, \nu \in K$

  $$\begin{aligned} f_0(\vec{a}_1, \ldots, \vec{a}_{i-1}, \mu\vec{a}_i + \nu\vec{a}'_i, \vec{a}_{i+1}, \ldots, \vec{a}_n) &= \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,\sigma(1)} \cdots (\mu\lambda_{i,\sigma(i)} + \nu\lambda'_{i,\sigma(i)}) \cdots \lambda_{n,\sigma(n)} \\ &= \mu \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,\sigma(1)} \cdots \lambda_{i,\sigma(i)} \cdots \lambda_{n,\sigma(n)} \\ &\quad +\nu \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,\sigma(1)} \cdots \lambda'_{i,\sigma(i)} \cdots \lambda_{n,\sigma(n)} \\ &= \mu f_0(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_n) + \nu f_0(\vec{a}_1, \ldots, \vec{a}'_i, \ldots, \vec{a}_n). \end{aligned}$$

- f_0 ist alternierend:
  Für $i \neq j$ betrachte man

  $$f_0(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n) = f_0(\vec{a}_{\tau(1)}, \ldots, \vec{a}_{\tau(i)}, \ldots, \vec{a}_{\tau(j)}, \ldots, \vec{a}_{\tau(n)})$$

  mit der Transposition $\tau = (i, j)$. Es ist

  $$\begin{aligned} f_0(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n) &= \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{\tau(1),\sigma(1)} \cdots \lambda_{\tau(n),\sigma(n)} \\ &= \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{\tau(1),(\sigma\circ\tau)(\tau(1))} \cdots \lambda_{\tau(n),(\sigma\circ\tau)(\tau(n))} \\ &= \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma)\lambda_{1,(\sigma\circ\tau)(1)} \cdots \lambda_{n,(\sigma\circ\tau)(n)}. \end{aligned}$$

Wegen der Gruppeneigenschaft von γ_n durchläuft $\tilde{\sigma} = \sigma \circ \tau$ alle Permutationen aus γ_n wenn σ dies tut. Die Abbildung $\sigma \mapsto \tilde{\sigma}$ ist nämlich bijektiv: ist $\sigma \circ \tau = \sigma' \circ \tau$, so folgt nach Multiplikation mit τ von rechts $\sigma = \sigma'$ (Injektivität), und für $\tilde{\sigma} \in \gamma_n$ ist $\sigma = \tilde{\sigma} \circ \tau$ wegen $\tau \circ \tau = id$ stets ein Urbild (Surjektivität). Also kann man die Summe umindizieren, es folgt

$$\begin{aligned}
f_0(\vec{a}_1, \ldots, \vec{a}_j, \ldots, \vec{a}_i, \ldots, \vec{a}_n) &= \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) \lambda_{1,(\sigma\circ\tau)(1)} \cdots \lambda_{n,(\sigma\circ\tau)(n)} \\
&= \sum_{\tilde{\sigma} \in \gamma_n} \operatorname{sgn}(\tilde{\sigma} \circ \tau) \lambda_{1,\tilde{\sigma}(1)} \cdots \lambda_{n,\tilde{\sigma}(n)} \\
&\underset{\text{Satz 7.1.5}}{=} -\sum_{\tilde{\sigma} \in \gamma_n} \operatorname{sgn}(\tilde{\sigma}) \lambda_{1,\tilde{\sigma}(1)} \cdots \lambda_{n,\tilde{\sigma}(n)} \\
&= -f_0(\vec{a}_1, \ldots, \vec{a}_i, \ldots, \vec{a}_j, \ldots, \vec{a}_n).
\end{aligned}$$

Also ist f_0 alternierend.

- f_0 ist nicht die Nullform:
  Für die spezielle Wahl $\vec{a}_i = \vec{b}_i$ für $i = 1, \ldots, n$ ist $\lambda_{ik} = \delta_{ik}$ das Kroneckersymbol, und wir erhalten
  $$f_0(\vec{b}_1, \ldots, \vec{b}_n) = \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) \delta_{1,\sigma(1)} \cdots \delta_{n,\sigma(n)}.$$
  Ist σ nicht die Identität, so gibt es ein i mit $\sigma(i) \neq i$ und der zugehörige Summand verschwindet wegen $\delta_{i,\sigma(i)} = 0$. Nur der Summand für $\sigma = id$ bleibt mit $\delta_{1,1} = \cdots = \delta_{n,n} = 1$ übrig. Also ist $f_0(\vec{b}_1, \ldots, \vec{b}_n) = 1 \neq 0$.

Ist nun $f \in \mathcal{L}$ beliebig, so ist nach Satz 7.2.7 für beliebige Vektoren $\vec{a}_1, \ldots, \vec{a}_n \in V$

$$f(\vec{a}_1, \ldots, \vec{a}_n) \;=\; f(\vec{b}_1, \ldots, \vec{b}_n) \cdot \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)} = f(\vec{b}_1, \ldots, \vec{b}_n) \cdot f_0(\vec{a}_1, \ldots, \vec{a}_n),$$

wobei $\mu = f(\vec{b}_1, \ldots, \vec{b}_n) \in K$ ein von f abhängiges Skalar ist, also ist $f = \mu \cdot f_0$, und es folgt insgesamt $\mathcal{L} = \langle\langle\, f_0 \,\rangle\rangle$ und $\dim \mathcal{L} = 1$. □

Definition 7.2.9. Eine Abbildung $D\colon V^n \to K$ heißt Determinantenfunktion, falls D eine n-fach alternierende Linearform, aber nicht die Nullform ist.

Satz 7.2.10. *Es sei* $\dim V = n$ *und* $\{\vec{b}_1, \ldots, \vec{b}_n\}$ *eine Basis von* V*. Dann gibt es für jedes* $d \in K \backslash \{0\}$ *genau eine Determinantenfunktion* $D\colon V^n \to K$ *mit* $D(\vec{b}_1, \ldots, \vec{b}_n) = d$*.*

Beweis. Es sei $f_0 \in \mathcal{L}$ mit $f_0 \neq 0$ beliebig, was nach Satz 7.2.8 auch existiert. Dann ist nach Satz 7.2.7 ii) speziell $f_0(\vec{b}_1, \ldots, \vec{b}_n) = \eta \neq 0$, da $\vec{b}_1, \ldots, \vec{b}_n$ linear unabhängig sind. Wir setzen $D = d\eta^{-1} f_0 \in \mathcal{L}$. Offenbar ist dann $D(\vec{b}_1, \ldots, \vec{b}_n) = d$. Es sei nun $\tilde{D} \in \mathcal{L}$ mit $\tilde{D}(\vec{b}_1, \ldots, \vec{b}_n) = d$ gegeben. Weiter ist nach Satz 7.2.8 dann $\tilde{D} = \mu f_0$ für ein $\mu \in K$, also $\tilde{D}(\vec{b}_1, \ldots, \vec{b}_n) = \mu f_0(\vec{b}_1, \ldots, b_n) = \mu\eta$. Es folgt $\mu\eta = d$ bzw. $\mu = d\eta^{-1}$, also $\tilde{D} = d\eta^{-1} f_0 = D$. □

Satz 7.2.11. *Ist* $D\colon V^n \to K$ *eine Determinantenfunktion auf* V*, so gilt*

$$\vec{a}_1, \ldots, \vec{a}_n \textit{ linear unabhängig} \Leftrightarrow D(\vec{a}_1, \ldots, \vec{a}_n) \neq 0.$$

Beweis. Dies folgt unmittelbar aus Satz 7.2.7 ii). □

7.3 Die natürliche Determinantenfunktion des K^n

Definition 7.3.1. Es sei $\{\vec{e}_1, \dots, \vec{e}_n\}$ die Standardbasis des $V = K^n$ (geschrieben als Zeilenvektoren). Dann verstehen wir unter der natürlichen Determinantenfunktion des K^n die nach Satz 7.2.10 eindeutig bestimmte Determinantenfunktion $D_n \colon V^n \to K$ mit $D_n(\vec{e}_1, \dots, \vec{e}_n) = 1$.

Satz 7.3.2. *(Leibniz-Formel)*
Es seien $\vec{a}_1, \dots, \vec{a}_n \in K^n$ mit

$$\vec{a}_i = (a_{i1} \cdots a_{in}) = \sum_{k=1}^{n} a_{ik}\vec{e}_k$$

gegeben. Dann ist

$$D_n(\vec{a}_1, \dots, \vec{a}_n) = \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Beweis. Nach Satz 7.2.7 ist

$$D_n(\vec{a}_1, \dots, \vec{a}_n) = D_n(\vec{e}_1, \dots, \vec{e}_n) \cdot \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

mit $D_n(\vec{e}_1, \dots, \vec{e}_n) = 1$. □

Definition 7.3.3. Es sei $\mathcal{A} = (a_{ik}) \in K^{(n,n)}$. Dann ist die Determinante von $\mathcal{A}$ der Wert

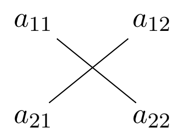
$$\det(A) = |A| = D_n(\vec{a}_1, \dots, \vec{a}_n)$$

mit den Zeilen $\vec{a}_i = (a_{i1} \cdots a_{in})$ von $\mathcal{A}$.

Beispiel 7.3.4. Im Fall $n = 2$ ist die Determinante nach der Leibniz-Formel durch

$$\det(\mathcal{A}) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \sum_{\sigma \in \gamma_2} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} = \underbrace{a_{11} \cdot a_{22}}_{\substack{\sigma = id \\ \text{sgn}=1}} - \underbrace{a_{12} \cdot a_{21}}_{\substack{\sigma = (1,2) \\ \text{sgn}=-1}}$$

gegeben. Dazu gehört das Schema:



Dabei ist das Produkt über die Querdiagonale mit dem Vorfaktor -1 zu versehen. Dann kann $|\det(\mathcal{A})|$ als der Flächeninhalt des von $\vec{a}_1 = (a_{11}\ a_{12})$ und $\vec{a}_2 = (a_{21}\ a_{22})$ aufgespannten Parallelogramms angesehen werden.

Beispiel 7.3.5. Im Fall $n = 3$ ist die Determinante

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{\sigma \in \gamma_3} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)}.$$

Die sechs Elemente $(6 = 3!)$ von γ_3 sind

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}}_{\text{sgn}=1}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}}_{\text{sgn}=-1}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}}_{\text{sgn}=-1}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}}_{\text{sgn}=1}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}}_{\text{sgn}=1}, \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}}_{\text{sgn}=-1}.$$

Also ist

$$\det(\mathcal{A}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Dies kann schematisch durch die Sarrussche Regel dargestellt werden:

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Dabei sind die Produkte über die Querdiagonalen mit dem Vorfaktor -1 zu versehen, und $|\det(\mathcal{A})|$ kann als das Volumen des von $\vec{a}_1 = (a_{11}, a_{12}, a_{13})^T$, $\vec{a}_2 = (a_{21}, a_{22}, a_{23})^T$ und $\vec{a}_3 = (a_{31}, a_{32}, a_{33})^T$ aufgespannten Parallelepipeds angesehen werden.

Satz 7.3.6. *Für $\mathcal{A} \in K^{(n,n)}$ gilt $\det(\mathcal{A}) = \det(\mathcal{A}^T)$.*

Beweis. Es sei $\mathcal{A} = (a_{ik})$ mit $1 \le i, k \le n$. Dann ist

$$\det(\mathcal{A}) = \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \text{ und } \det(\mathcal{A}^T) = \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Es ist $a_{1,\sigma(1)} \cdots a_{n,\sigma_n} = a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n}$, weswegen

$$\det(\mathcal{A}) = \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n} = \sum_{\sigma \in \gamma_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1),1} \cdots a_{\sigma^{-1}(n),n}$$

gilt. Durchläuft σ alle Elemente von γ_n, so auch $\varrho = \sigma^{-1}$, also ist nach Umindizierung der Summe

$$\det(\mathcal{A}) = \sum_{\varrho \in \gamma_n} \operatorname{sgn}(\varrho) a_{\varrho(1),1} \cdots a_{\varrho(n),n} = \det(\mathcal{A}^T).$$

□

Satz 7.3.7. *Für Determinanten von Matrizen gelten folgende Regeln:*

i) Für alle $\lambda \in K$ gilt

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{i1} & \cdots & \lambda a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

sowie

$$\det \begin{pmatrix} a_{11} & \cdots & \lambda a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \lambda a_{ni} & \cdots & a_{nn} \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix}.$$

ii) Es gilt

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+b_{i1} & \cdots & a_{in}+b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} + \det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

sowie

$$\begin{aligned} \det\begin{pmatrix} a_{11} & \cdots & a_{1i}+b_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni}+b_{ni} & \cdots & a_{nn} \end{pmatrix} &= \det\begin{pmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix} \\ &\quad + \det\begin{pmatrix} a_{11} & \cdots & b_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_{ni} & \cdots & a_{nn} \end{pmatrix}. \end{aligned}$$

iii) Besitzt $\mathcal{A}$ eine Zeile oder Spalte, die nur aus Nullen besteht, so ist $\det(\mathcal{A}) = 0$.

iv) Die Determinante einer Matrix ändert sich nicht, wenn man zu einer Zeile (oder Spalte) das Vielfache einer Zeile (oder Spalte) addiert.

v) Die Determinante einer Matrix ändert beim Vertauschen zweier Zeilen (oder zweier Spalten) das Vorzeichen.

vi) Es gilt

$$\begin{aligned} \det(\mathcal{A}) \neq 0 \quad &\Leftrightarrow \quad \textit{die Zeilenvektoren sind linear unabhängig} \\ &\Leftrightarrow \quad \textit{die Spaltenvektoren sind linear unabhängig} \Leftrightarrow \mathcal{A} \textit{ ist regulär.} \end{aligned}$$

Beweis. Per Definition ist die Determinante einer Matrix eine n-fache alternierende Linearform der Zeilenvektoren der Matrix. Nach Satz 7.3.6 ist sie auch eine n-fache alternierende Linearform der Spaltenvektoren. Alle Aussagen folgen daher aus den entsprechenden Aussagen für alternierende Linearformen. □

Satz 7.3.8. *Es ist*

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} & a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \det\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}.$$

Beweis. Es sei $\mathcal{A} = (a_{ik})$. Dann ist nach nach Satz 7.3.7

$$\det(\mathcal{A}) \underset{\text{Satz 7.3.2}}{=} \sum_{\sigma\in\gamma_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}.$$

Ist $\sigma \in \gamma_n$ eine Permutation mit $\sigma(n) \neq n$, so ist $\sigma(j) = n$ mit $j < n$, so verschwindet der zugehörige Summand, weil er den Faktor $a_{j,\sigma(j)} = a_{j,n} = 0$ enthält. Die Determinante von $\mathcal{A}$ ist dann

$$\begin{aligned}
\sum_{\substack{\sigma\in\gamma_n \\ \sigma(n)=n}} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \quad &= \quad \sum_{\varrho\in\gamma_{n-1}} \operatorname{sgn}(\varrho) a_{1,\varrho(1)} \cdots a_{n-1,\varrho(n-1)} \cdot \underbrace{a_{n,n}}_{=1} \\
&\underset{\text{S. 7.3.2}}{=} \det\begin{pmatrix} a_{11} & \cdots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix},
\end{aligned}$$

da jedes $\sigma \in \gamma_n$ mit $\sigma(n) = n$ als Permutation von $\{1, \ldots, n-1\}$ aufgefasst werden kann (ohne Änderung der Fehlstandszahl). $\square$

Definition 7.3.9. Es sei $n \geq 2$ und $\mathcal{A} \in K^{(n,n)}$. Mit $\mathcal{A}_{ij}$ bezeichnen wir die Matrix, die aus $\mathcal{A}$ durch Streichung der i-ten Zeile und der j-ten Spalte entsteht:

$$\mathcal{A}_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Satz 7.3.10. *Es sei $n \geq 2$ und $\mathcal{A} \in K^{(n,n)}$. Dann gilt*

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1j} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix} = (-1)^{i+j} \cdot \det(\mathcal{A}_{ij}).$$

Beweis. Durch $n - j$ Vertauschungen benachbarter Spalten erhalten wir

$$\det(\mathcal{A}) = (-1)^{j-1} \cdot \det\begin{pmatrix} a_{11} & \cdots & a_{1n} & a_{1j} \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & a_{nj} \end{pmatrix}.$$

Weitere $n - i$ Vertauschungen benachbarter Zeilen ergeben

$$\det(\mathcal{A}) = (-1)^{n-j}(-1)^{n-i} \cdot \left(\begin{array}{ccc|c} & & & * \\ & \mathcal{A}_{ij} & & \vdots \\ & & & * \\ \hline 0 & \cdots & 0 & 1 \end{array}\right).$$

Aus $(-1)^{n-i+n-j} = (-1)^{i+j}$ und Satz 7.3.8 folgt die Behauptung. $\square$

Satz 7.3.11. *(Entwicklungssatz von Laplace)*
Es sei $n \geq 2$ und $\mathcal{A} \in K^{(n,n)}$.
Für jedes feste i gilt die Entwicklung nach der i-ten Zeile:

$$det(\mathcal{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathcal{A}_{ij}).$$

Für jedes feste j gilt die Entwicklung nach der j-ten Spalte:

$$\det(\mathcal{A}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\mathcal{A}_{ij}).$$

Beweis. Es ist $\det(\mathcal{A}) = D_n(\vec{a}_1, \ldots, \vec{a}_n)$ mit den Zeilen $\vec{a}_k = (a_{k1}, \ldots, a_{kn})^T$ für $1 \leq k \leq n$. Die Entwicklung nach der i-ten Zeile folgt aus

$$\begin{aligned} D(\vec{a}_1, \ldots, \vec{a}_n) &= D_n\left(\vec{a}_1, \ldots , \sum_{j=1}^{n} a_{ij}\vec{e}_j\ , \ \ldots, \vec{a}_n\right) = \sum_{j=1}^{n} a_{ij} D_n(\vec{a}_1, \ldots, \vec{e}_j, \ldots, \vec{a}_n) \\ &= \sum_{j=1}^{n} a_{ij} \det \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix} \underset{\text{S. 8.1.11}}{=} \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det(\mathcal{A}_{ij}). \end{aligned}$$

Die Entwicklung nach der j-ten Spalte wird durch Übergang zur Transponierten gezeigt. □

Beispiel 7.3.12. Wir bestimmen die Determinante von

$$\mathcal{A} = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 1 & 0 & 2 & 0 \\ 3 & 6 & 4 & 9 \\ 5 & 8 & 7 & 13 \end{pmatrix}.$$

Wegen des Auftretens von zwei Nullen in der zweiten Zeile empfiehlt es sich, die Determinante nach der zweiten Zeile zu entwickeln:

$$\det(\mathcal{A}) = (-1) \cdot 1 \cdot \det \begin{pmatrix} 2 & 1 & 3 \\ 6 & 4 & 9 \\ 8 & 7 & 13 \end{pmatrix} + (-1) \cdot 2 \cdot \det \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \\ 5 & 8 & 13 \end{pmatrix}.$$

Die zweite Determinante ist null, da die ersten beiden Zeilen linear abhängig sind. Also ist

$$\begin{aligned} \det(\mathcal{A}) &= -\det \begin{pmatrix} 2 & 1 & 3 \\ 6 & 4 & 9 \\ 8 & 7 & 13 \end{pmatrix} \underset{\substack{(II)-3\cdot(I) \\ (III)-4\cdot(I)}}{=} -\det \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix} \underset{\text{entwickeln}}{\overset{\text{2. Zeile}}{=}} -(-1)^2 \det \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix} \\ &= -(2 \cdot 1 - 0 \cdot 3) = -2. \end{aligned}$$

Determinanten können auch dazu verwendet werden, explizite Formeln für die Lösungen eindeutig lösbarerer quadratischer Linearer Gleichungssysteme zu erhalten.

Satz 7.3.13. *(Cramersche Regel)*
Es sei $\mathcal{A} \in K^{(n,n)}$ *und* $\vec{b} = (b_1, \ldots, b_n)^T \in K^n$ *(ein Spaltenvektor). Das LGS* $\mathcal{A}\vec{x} = \vec{b}$ *besitzt genau dann eine eindeutige Lösung, wenn* $\det(\mathcal{A}) \neq 0$ *ist. Ist dies der Fall, dann ist die Lösung* $\vec{x} = (x_1, \ldots, x_n)^T$ *durch*

$$x_k = \frac{1}{\det(\mathcal{A})} \cdot \det \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix}$$

gegeben, wobei die k*-te Spalte von* $\mathcal{A}$ *durch* $\vec{b}$ *ersetzt wird.*

Beweis. Nach Satz 6.1.4 ist $\mathcal{A}\vec{x} = \vec{b}$ genau dann eindeutig lösbar, wenn $\mathcal{A}$ regulär ist. Das ist nach Satz 7.3.7 vi) zu $\det(\mathcal{A}) \neq 0$ äquivalent. In diesem Fall gilt für die eindeutig bestimmte Lösung $\vec{x} = (x_1, \ldots, x_n)^T$

$$\vec{b} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n,$$

wobei

$$\vec{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

die j-te Spalte von $\mathcal{A}$ ist. Dann gilt

$$\begin{aligned} \det \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix} &= D_n(\vec{a}_1, \ldots, \vec{a}_{k-1}, \vec{b}, \vec{a}_{k+1}, \ldots, \vec{a}_n) \\ &= D_n\left(\vec{a}_1, \ldots, \vec{a}_{k_1}\, ,\, \sum_{j=1}^{n} x_j\vec{a}_j\, ,\, \vec{a}_{k+1}, \ldots, \vec{a}_n\right) \\ &= \sum_{j=1}^{n} x_j D_n(\vec{a}_1, \ldots, \vec{a}_{k_1}, \vec{a}_j, \vec{a}_{k+1}, \ldots, \vec{a}_n)\,. \end{aligned}$$

Nun ist $D_n(\vec{a}_1, \ldots, \vec{a}_{k-1}, \vec{a}_j, \vec{a}_{k+1}, \ldots, \vec{a}_n) = 0$ für $j \neq k$, da dann zwei gleiche Argumente vorliegen. Der verbleibende Summand ist $x_k \det(\mathcal{A})$, also ist

$$\det \begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix} = x_k \cdot \det(\mathcal{A})\,,$$

woraus die Behauptung folgt, weil man diese Gleichung wegen $\det(\mathcal{A}) \neq 0$ durch $\det(\mathcal{A})$ teilen darf. □

7.4 Der Multiplikationssatz

Satz 7.4.1. *(Determinantenmultiplikationssatz)*
Für $\mathcal{A}, \mathcal{B} \in K^{(n,n)}$ *gilt* $\det(\mathcal{A} \cdot \mathcal{B}) = \det(\mathcal{A}) \cdot \det(\mathcal{B})$*. Ist* $\mathcal{A}$ *regulär, so gilt* $\det(\mathcal{A}^{-1}) = \det(\mathcal{A})^{-1}$*.*

Beweis. Wir zeigen dies in drei Schrirtten:

- Es ist $f_{\mathcal{A}}$ eine n- fach alternierende Linearform.

  Nach Satz 7.2.4 genügt es zu zeigen, dass $f_{\mathcal{A}}$ homogen und schwerungsinvariant ist. Es gilt

$$\begin{aligned} f_{\mathcal{A}}(\vec{b}_1, \ldots, \lambda\vec{b}_i, \ldots \vec{b}_n) &= \det(\mathcal{A}\vec{b}_1, \ldots, \lambda\mathcal{A}\vec{b}_i, \ldots, \mathcal{A}\vec{b}_n) = \lambda \cdot \det(\mathcal{A}\vec{b}_1, \ldots, \mathcal{A}\vec{b}_i, \ldots, \mathcal{A}\vec{b}_n) \\ &= \lambda \cdot f_{\mathcal{A}}(\vec{b}_1, \ldots, \lambda\vec{b}_i, \ldots \vec{b}_n), \end{aligned}$$

  also ist $f_{\mathcal{A}}$ homogen. Weiter ist für $k \neq i$

$$\begin{aligned} f_{\mathcal{A}}(\vec{b}_1, \ldots, \vec{b}_i + \vec{b}_k, \ldots \vec{b}_n) &= \det(\mathcal{A}\vec{b}_1, \ldots, \mathcal{A}(\vec{b}_i + \vec{b}_k), \ldots, \mathcal{A}\vec{b}_n) \\ &= \det(\mathcal{A}\vec{b}_1, \ldots, \mathcal{A}\vec{b}_i + \mathcal{A}\vec{b}_k, \ldots, \mathcal{A}\vec{b}_n) = \det(\mathcal{A}\vec{b}_1, \ldots, \mathcal{A}\vec{b}_i, \ldots, \mathcal{A}\vec{b}_n) \\ &= f_{\mathcal{A}}(\vec{b}_1, \ldots, \vec{b}_i, \ldots \vec{b}_n), \end{aligned}$$

  womit $f_{\mathcal{A}}$ auch scherungsinvariant ist. Daraus folgt die Behauptung.

- Es gilt $f_{\mathcal{A}}(\vec{b}_1, \ldots, \vec{b}_n) = (\det \mathcal{A}) \cdot (\det \mathcal{B})$, wobei $\mathcal{B} = (\vec{b}_1, \ldots, \vec{b}_n)$ die Matrix mit den Spaltenvektoren $\vec{b}_1, \ldots, \vec{b}_n$ ist.

  Da $f_{\mathcal{A}}$ eine Determinantenfunktion und nicht die Nullform ist, gilt

$$f_{\mathcal{A}}(\vec{b}_1, \ldots, \vec{b}_n) = \det(\vec{b}_1, \ldots, \vec{b}_n) \cdot f_{\mathcal{A}}(\vec{e}_1, \ldots, \vec{e}_n)$$

  mit den Standardvektoren $\vec{e}_1, \ldots, \vec{e}_n$. Daraus folgt

$$f_{\mathcal{A}}(\vec{b}_1, \ldots, \vec{b}_n) = \det(\vec{b}_1, \ldots, \vec{b}_n) \cdot f_{\mathcal{A}}(\vec{e}_1, \ldots, \vec{e}_n) = \det \mathcal{B} \cdot \det(\mathcal{A}\vec{e}_1, \ldots, \mathcal{A}\vec{e}_n) = \det \mathcal{B} \cdot \det \mathcal{A}.$$

Mit $\mathcal{C} = \mathcal{A} \cdot \mathcal{B}$ folgt nun $\det(\mathcal{A} \cdot \mathcal{B}) = \det \mathcal{A} \cdot \det \mathcal{B}$. □

Eine wichtige Folgerung aus diesem Satz ist, dass ein Basiswechsel die Determinante einer Darstellungsmatrix nicht ändert: Ist $\varphi \in L(V, V)$ für einen endlichdimensionalen Vektorraum V, und sind $\mathcal{B}_1$ und $\mathcal{B}_2$ zwei Basen von V, so ist $\mathcal{M}(\varphi; \mathcal{B}_2, \mathcal{B}_2) = \mathcal{X}^{-1} \cdot \mathcal{M}(\varphi; \mathcal{B}_1, \mathcal{B}_1) \cdot \mathcal{X}$ für eine reguläre Matrix $\mathcal{X} \in \mathrm{GL}(n, K)$ nach Satz 5.5.7. Nach dem Multiplikationssatz ist

$$\begin{aligned} \det(\mathcal{M}(\varphi; \mathcal{B}_2, \mathcal{B}_2)) &= \det(\mathcal{X}^{-1} \cdot \mathcal{M}(\varphi; \mathcal{B}_1, \mathcal{B}_1) \cdot \mathcal{X}) \\ &= \det(\mathcal{X})^{-1} \cdot \det(\mathcal{M}(\varphi; \mathcal{B}_1, \mathcal{B}_1)) \cdot \det(\mathcal{X}) = \det(\mathcal{M}(\varphi; \mathcal{B}_1, \mathcal{B}_1)). \end{aligned}$$

Definition 7.4.2. Es sei V ein n- dimensionaler K-Vektorraum. Die <u>Determinante</u> von φ ist

$$\det(\varphi) = \det(\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}))$$

für eine Basis $\mathcal{B}$ von V. Die Definition hängt nach der vorigen Rechnung nicht davon ab, welche Basis man wählt.

Der Multiplikationssatz für Endomorphismen lautet

Satz 7.4.3. *Sind* $\varphi, \psi \in L(V, V)$*, so gilt* $\det(\varphi \circ \psi) = \det(\varphi) \cdot \det(\psi)$. *Es gilt*

$$\varphi \textit{ bijektiv} \;\Leftrightarrow\; \det(\varphi) \neq 0.$$

In diesem Fall ist $\det(\varphi^{-1}) = \det(\varphi)^{-1}$.

Kapitel 8

Diagonalisierung

8.1 Eigenwerte und Eigenvektoren

Definition 8.1.1. Es sei $\varphi \in L(V, V)$ ein Endomorphismus. Dann heißt $\lambda \in K$ ein Eigenwert (EW) von φ, falls es einen Vektor $\vec{x} \in V$ mit $\vec{x} \neq \vec{0}$ gibt, so dass $\varphi(\vec{x}) = \lambda\vec{x}$ gilt. Dieser Vektor $\vec{x}$ heißt dann Eigenvektor (EV) von φ zum Eigenwert λ.

Definition 8.1.2. Es sei $\mathcal{A} \in K^{(n,n)}$ eine quadratische Matrix. Dann heißt $\lambda \in K$ ein Eigenwert der Matrix $\mathcal{A}$, falls es einen Vektor $\vec{x} \in K^n$ mit $\vec{x} \neq \vec{0}$ gibt, so dass $\mathcal{A}\vec{x} = \lambda\vec{x}$ gilt. Man nennt $\vec{x}$ dann Eigenvektor von $\mathcal{A}$ zum Eigenwert λ, d.h. λ ist Eigenwert des Endomorphismus $\varphi(\vec{x}) = \mathcal{A}\vec{x}$.

Beispiel 8.1.3. Die Identität $\varphi = id$ hat den Eigenwert 1 und sonst keinen. Die Nullabbildung $N(\vec{x}) = \vec{0}$ hat als einzigen Eigenwert 0. Ein Endomorphismus φ hat genau dann den Eigenwert 0, wenn es ein $\vec{x} \neq 0$ mit $\varphi(\vec{x}) = \vec{0}$ gibt.

Beispiel 8.1.4. Wir betrachten die Projektion $P \colon K^n \to K^n$ mit $P(x_1, \dots, x_n) = (x_1, \dots, x_r, 0, \dots, 0)$ mit $1 < r < n$.
Sie besitzt den Eigenwert 1, denn jeder Vektor der Form $\vec{x} = (x_1, \dots, x_r, 0, \dots, 0)$ erfüllt $P(\vec{x}) = \vec{x}$. Andererseits besitzt P auch den Eigenwert 0, denn jeder Vektor der Form $\vec{x} = (0, \dots, 0, x_{r+1}, \dots, x_n)$ erfüllt $P(\vec{x}) = \vec{0}$.

Satz 8.1.5. *Es sei $\varphi \in L(V, V)$, $\mathcal{B}$ eine Basis von V sowie $\lambda \in K$. Dann gilt*

$$\lambda \textit{ ist Eigenwert von } \varphi \Leftrightarrow \lambda \textit{ ist Eigenwert von } \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}).$$

Bemerkung 8.1.6. Offensichtlich hängt die Aussage auf der linken Seite nicht von der Wahl der Basis $\mathcal{B}$ ab. Damit folgt aus Satz 5.5.7, dass $\mathcal{A} \in K^{(n,n)}$ und $\mathcal{X}^{-1}\mathcal{A}\mathcal{X}$ für alle $\mathcal{X} \in \mathrm{GL}(n, K)$ dieselben Eigenwerte besitzen. Aus Satz 7.4.1 folgt übrigens, dass sie auch dieselbe Determinante besitzen, und sie haben nach Satz 5.4.19 auch den gleichen Rang.

Dadurch wird die folgende Definition motiviert:

Definition 8.1.7. Matrizen $\mathcal{A}, \mathcal{C} \in K^{(n,n)}$ heißen ähnlich (Schreibweise: $\mathcal{A} \approx \mathcal{C}$), wenn es ein reguläres $\mathcal{X} \in \mathrm{GL}(n, K)$ mit $\mathcal{C} = \mathcal{X}^{-1} \cdot \mathcal{A} \cdot \mathcal{X}$ gibt.

Daraus folgt

i) $\mathcal{A} \approx \mathcal{C} \Rightarrow \mathrm{rg}(\mathcal{A}) = \mathrm{rg}(\mathcal{C})$ und $\det \mathcal{A} = \det \mathcal{C}$.

ii) Ist $\varphi \in L(V,V)$ und $\mathcal{B}$ eine Basis von V mit $\mathcal{A} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$, so folgt aus den Sätzen aus Kapitel 5 die Äquivalenz

$$\mathcal{A} \approx \mathcal{C} \Leftrightarrow \text{ Es gibt eine Basis } \tilde{\mathcal{B}} \text{ von } V \text{ mit } \mathcal{C} = \mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}).$$

Beweis. (Beweis von Satz 8.1.5)
Es sei $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V und

$$\vec{x} = \sum_{i=1}^{n} \mu_i \vec{b}_i \quad \text{und} \quad \varphi(x) = \sum_{i=1}^{n} \mu_i' \vec{b}_i$$

mit $\mu_i, \mu_i \in K$. Nach Satz 5.4.2 gilt

$$\begin{pmatrix} \mu_1' \\ \vdots \\ \mu_n' \end{pmatrix} = \mathcal{M}(\varphi; \mathcal{B}', \mathcal{B}) \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}.$$

Nun ist λ genau dann Eigenwert von φ, wenn $\varphi(\vec{x}) = \lambda \vec{x}$ für ein $\vec{x} \neq 0$ ist, also für

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \neq \vec{0} \quad \text{und} \quad \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} \mu_1' \\ \vdots \\ \mu_n' \end{pmatrix} \stackrel{!}{=} \lambda \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix},$$

also genau dann, wenn $(\mu_1, \ldots, \mu_n)$ Eigenvektor von $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B})$ zum Eigenwert λ ist. □

Definition 8.1.8. Es sei $\lambda \in K$ ein Eigenwert von $\varphi \in L(V,V)$. Dann heißt

$$U_\lambda = \{\vec{x} \in V : \varphi(\vec{x}) = \lambda \vec{x}\} = \operatorname{Kern}(\varphi - \lambda id)$$

der Eigenraum von λ. Die Elemente von U_λ sind die Eigenvektoren zum Eigenwert λ und der Nullvektor. Es heißt $\dim U_\lambda = \dim \operatorname{Kern}(\varphi - \lambda id)$ die geometrische Vielfachheit oder geometrische Ordnung des Eigenwerts λ. Entsprechend definiert man für Matrizen $\mathcal{A} \in K^{(n,n)}$

$$U_\lambda = \{\vec{x} \in K^n : (A - \lambda \mathcal{E}_n)\vec{x} = \vec{0}\}$$

sowie $\dim U_\lambda = \dim \operatorname{Kern}(A - \lambda \mathcal{E}_n) = n - \operatorname{rg}(A - \lambda \mathcal{E}_n)$.

Satz 8.1.9. *Es seien $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Eigenwerte von $\varphi \in L(V,V)$ und $\vec{x}_1, \ldots, \vec{x}_r$ zugehörige Eigenvektoren. Dann sind $\vec{x}_1, \ldots, \vec{x}_r$ linear unabhängig.*

Bemerkung 8.1.10. Folglich besitzt φ höchstens $n = \dim V$ verschiedene Eigenwerte.

Beweis. Es ist genau dann $\vec{x}_i$ ein Eigenvektor zum Eigenwert λ, wenn $\vec{x}_i \neq \vec{0}$ und $\varphi(\vec{x}_i) = \lambda_i \vec{x}_i$ für $i = 1, \ldots, r$ ist. Weiter gilt für $i, j = 1, \ldots, r$

$$(\varphi - \lambda_i id)(\vec{x}_j) = \varphi(\vec{x}_j) - \lambda_i \vec{x}_j = \lambda_j \vec{x}_j - \lambda_i \vec{x}_j = (\lambda_j - \lambda_i)\vec{x}_j \tag{$*$}$$

Nun sei $\alpha_1 \vec{x}_1 + \cdots + \alpha_n \vec{x}_r = \vec{0}$ für $\alpha_1, \ldots, \alpha_r \in K$. Wir müssen $\alpha_k = 0$ zeigen. Auf diese Gleichung wenden wir die folgende lineare Abbildung an:

$$\psi_k = (\varphi - \lambda_1 id) \circ \cdots \circ (\varphi - \lambda_{k-1} id) \circ (\varphi - \lambda_{k+1} id) \circ \cdots \circ (\varphi - \lambda_r id).$$

Wir erhalten

$$\vec{0} = \psi_k(\vec{0}) = \psi_k\left(\sum_{j=1}^{n} \alpha_j \vec{x}_j\right) = \sum_{j=1}^{n} \alpha_j \psi_k(\vec{x}_j) = \sum_{j=1}^{n} \alpha_j \prod_{\substack{i=1\\ i\neq k}}^{r} (\varphi - \lambda_i \mathrm{id})(\vec{x}_j) \underset{(*)}{=} \sum_{j=1}^{r} \alpha_j \prod_{\substack{i=1\\ i\neq k}}^{r} (\lambda_j - \lambda_i)\vec{x}_j.$$

Das erste Produkt gehört zur Verknüpfung $\circ$ auf $L(V,V)$, das zweite ist ein gewöhnliches Produkt von Elementen aus K. In der letzten Summe verschwinden alle Summanden bis auf den Summanden für $i = k$, woraus

$$\vec{0} = \alpha_k \cdot \prod_{\substack{i=1\\ i\neq k}}^{r} (\lambda_k - \lambda_i)\vec{x}_k$$

folgt. Da die Eigenwerte paarweise verschieden sind, verschwindet das Produkt nicht. Wegen $\vec{x}_k \neq \vec{0}$ ist dann $\alpha_k = 0$. □

Satz 8.1.11. *Für $\varphi \in L(V,V)$ und $\mathcal{A} \in K^{(n,n)}$ sowie $\lambda \in K$ gilt*

i) λ ist Eigenwert von $\varphi \Leftrightarrow \det(\varphi - \lambda id) = 0$.

ii) λ ist Eigenwert von $\mathcal{A} \Leftrightarrow \det(\mathcal{A} - \lambda\mathcal{E}_n) = 0$.

Beweis. i) Für jedes $\lambda \in K$ gilt

$$\begin{aligned} \lambda \text{ ist Eigenwert von } \varphi \ &\Leftrightarrow\ \mathrm{Kern}(\varphi - \lambda id) \neq \{\vec{0}\} \Leftrightarrow \varphi - \lambda id \text{ ist nicht injektiv} \\ &\Leftrightarrow\ \varphi - \lambda id \text{ ist kein Automorphismus} \underset{\text{Satz 7.4.3}}{\Leftrightarrow} \det(\varphi - \lambda id) = 0. \end{aligned}$$

ii) Es gilt völlig analog

$$\begin{aligned} \lambda \text{ ist Eigenwert von } \mathcal{A} \ &\Leftrightarrow\ \exists \vec{x} \neq 0\colon \mathcal{A}\vec{x} = \lambda\vec{x} \Leftrightarrow \exists \vec{x} \neq 0\colon (\mathcal{A} - \lambda\mathcal{E}_n)\vec{x} = \vec{0} \\ &\Leftrightarrow\ \mathcal{A} - \lambda\mathcal{E}_n\varphi \text{ ist nicht regulär} \Leftrightarrow \det(\mathcal{A} - \lambda\mathcal{E}_n) = 0. \end{aligned}$$

□

Satz 8.1.12. *(Charakteristisches Polynom)*
Für $\mathcal{A} \in K^{(n,n)}$ ist $\det(\mathcal{A} - \lambda\mathcal{E}_n)$ ein Polynom in λ vom Grad n mit Koeffizienten aus K:

$$\det(\mathcal{A} - \lambda\mathcal{E}_n) = P_{\mathcal{A}}(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \cdots + \alpha_n\lambda^n \in K[\lambda].$$

Beweis. Mit $\mathcal{A} = (a_{ij})$ folgt

$$\mathcal{A} - \lambda\mathcal{E}_n = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} = (b_{ij}).$$

Es folgt

$$\det(\mathcal{A} - \lambda\mathcal{E}_n) = (a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda) + \sum_{\substack{\sigma\in\gamma_n\\ \sigma\neq id}} \mathrm{sgn}(\sigma) b_{1,\sigma(1)} b_{2,\sigma(2)} \cdots b_{n,\sigma(n)}.$$

Jeder Summand ist ein Produkt aus einem Skalar aus dem Körper K und einer Potenz von λ höchstens vom Grad $n-2$, also ist die Summe ein Polynom $Q(\lambda) \in K[\lambda]$ ebenfalls höchstens vom Grad $n-2$. Wir multiplizieren das Produkt über die $(a_{ii} - \lambda)$ aus und erhalten

$$\det(\mathcal{A} - \lambda \mathcal{E}_n) = (-1)^n \lambda^n + (-\lambda)^{n-1}(a_{11} + \cdots + a_{nn}) + \tilde{Q}(\lambda)$$

mit einem Polynom $\tilde{Q}(\lambda)$ vom Grad höchstens $n-2$. Damit ist $P_{\mathcal{A}}(\lambda)$ ein Polynom vom Grad n. □

Definition 8.1.13. Das Polynom $P_{\mathcal{A}}(\lambda) = \det(\mathcal{A} - \lambda \mathcal{E}_n)$ heißt charakteristisches Polynom von $\mathcal{A}$.

Es gilt $\alpha_n = (-1)^n$, $\alpha_0 = P_{\mathcal{A}}(0) = \det \mathcal{A}$, sowie $\alpha_{n-1} = (-1)^{n-1} \cdot S(\mathcal{A})$ mit $S(\mathcal{A}) = a_{11} + \cdots + a_{nn}$.

Definition 8.1.14. Die Summe

$$S(\mathcal{A}) = a_{11} + \cdots + a_{nn} \ = \ (-1)^{n-1}\alpha_{n-1}$$

der Diagonalelemente von $\mathcal{A} \in K^{(n,n)}$ nennt man die Spur von $\mathcal{A}$.

Bemerkung 8.1.15. Für $n = 2$ besitzt

$$\mathcal{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

stets das charakteristische Polynom

$$P_{\mathcal{A}}(\lambda) = \lambda^2 - S(\mathcal{A})\lambda + \det \mathcal{A} = \lambda^2 - (a+d)\lambda + (ad - bc).$$

Satz 8.1.12 besagt, dass die Eigenwerte von $\mathcal{A}$ die Nullstellen des charakteristischen Polynoms sind:

$$\lambda_i \text{ ist Eigenwert von } \mathcal{A} \Leftrightarrow P_{\mathcal{A}}(\lambda_i) = 0.$$

Dadurch kann man die Eigenwerte einer Matrix rechnerisch bestimmen:

Beispiel 8.1.16. Bestimme die Eigenwerte von

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

Für die Unbestimmte λ ist

$$\mathcal{A} - \lambda \mathcal{E}_n = \begin{pmatrix} 1-\lambda & 2 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 1 & 1-\lambda \end{pmatrix}.$$

Die Determinante dieser Matrix ist

$$\det \mathcal{A} - \lambda \mathcal{E}_n \underset{\text{2. Zeile}}{=} (1-\lambda) \cdot \det \begin{pmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda) \cdot \left((1-\lambda)^2 - 1 \cdot 1\right) = (1-\lambda) \cdot (-2\lambda + \lambda^2).$$

Dieses Polynom kann man entweder in ausmultiplizierter Form

$$P_{\mathcal{A}}(\lambda) = (-1)\lambda^3 + 3\lambda^2 + (-2)\lambda + 0\lambda^0$$

oder in Produktform

$$P_{\mathcal{A}}(\lambda) = -(\lambda - 2) \cdot (\lambda - 1) \cdot (\lambda - 0)$$

aufschreiben.

An der ersten Form sieht man

$$\det \mathcal{A} = \alpha_0 = 0 \quad \text{und} \quad S(\mathcal{A}) = (-1)^{n-1}\alpha_{n-1} = 3.$$

Aus der Produktform kann man die Eigenwerte $\lambda_{1,2,3} = 0, 1, 2$ ablesen.

Bemerkung 8.1.17. Es ist möglich, dass $\mathcal{A}$ keine Eigenwerte besitzt:

$$\mathcal{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{(2,2)} \Rightarrow P_{\mathcal{A}}(\lambda) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1.$$

Dieses Polynom hat in $\mathbb{R}$ keine Nullstellen. Geometrisch ist

$$\mathcal{A} = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}$$

die Darstellungsmatrix der Drehung im $\mathbb{R}^2$ um 90 Grad.
Nur der Nullvektor wird durch diese Drehung in ein Vielfaches seiner selbst überführt.

Definition 8.1.18. Es sei $\mathcal{A} \in K^{(n,n)}$ und $\lambda_0 \in K$ ein Eigenwert von $\mathcal{A}$. Die Ordnung von λ_0 als Nullstelle von

$$P_{\mathcal{A}}(\lambda) = \det(A - \lambda\mathcal{E}_n)$$

heißt die algebraische Ordnung oder die algebraische Vielfachheit des Eigenwerts λ_0.

Bemerkung 8.1.19. Ist $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n \in K[x]$ ein Polynom mit Grad ≥ 1 und $x_0 \in K$ eine Nullstelle, d.h. $p(x_0) = 0$, so ergibt wiederholte Polynomdivision

$$p(x) = (x - x_0)^{k_0} \cdot q_0(x)$$

für ein maximales $k_0 \in \mathbb{N}$ und ein $q_0(x) \in K[x]$. Die Zahl k_0 heißt Ordnung der Nullstelle x_0. Besitzt $q_0(x)$ eine von x_0 verschiedene Nullstelle x_1, so kann man den Prozess fortsetzen:

$$p(x) = (x - x_0)^{k_0} \cdot q_0(x) = (x - x_0)^{k_0}(x - x_1)^{k_1} \cdot q_1(x) = \cdots = (x - x_0)^{k_0} \cdots (x - x_r)^{k_r} \cdot q_r(x),$$

so dass $q_r(x) \neq 0$ für $x = x_0, \ldots, x_r$ ist. Falls man $q_r(x) = a \in K$ mit einer Konstanten a erreichen kann, sagt man, dass $p(x)$ in K vollständig in Linearfaktoren zerfällt.

Satz 8.1.20. *(Fundamentalsatz der Algebra)*
Über dem Körper $\mathbb{C}$ der komplexen Zahlen zerfällt jedes nichtkonstante Polynom vollständig in Linearfaktoren.

Für den Beweis verweisen wir auf die weiterführenden Vorlesungen der Mathematik.

Beispiel 8.1.21. Das Polynom $p(x) = (x^2 + 1)^2$ hat keine Nullstelle in $\mathbb{R}$, aber in $\mathbb{C}$ hat es die Nullstellen $\pm i$ jeweils zur Ordnung 2. Es zerfällt in der Form

$$(x^2 + 1)^2 = (x - i)^2 \cdot (x + i)^2$$

im Körper $\mathbb{C}$.

Satz 8.1.22. *Ähnliche Matrizen haben dasselbe charakteristische Polynom und dieselben Eigenwerte, und zwar von jeweils derselben geometrischen und derselben algebraischen Vielfachheit.*

Beweis. Ist $\mathcal{A} \approx \mathcal{C}$, so gibt es $\mathcal{X} \in \mathrm{GL}(n, K)$ mit $\mathcal{C} = \mathcal{X}^{-1} \cdot \mathcal{A} \cdot \mathcal{X}$. Dann gilt

$$\begin{aligned} P_{\mathcal{C}}(\lambda) &= \det(\mathcal{X}^{-1}\mathcal{A}\mathcal{X} - \lambda\mathcal{E}_n) = \det(\mathcal{X}^{-1}\mathcal{A}\mathcal{X} - \lambda\mathcal{X}^{-1}\mathcal{X}) \\ &= \det(\mathcal{X}^{-1}(\mathcal{A} - \lambda\mathcal{E}_n)\mathcal{X}) \underset{\text{Satz 7.4.1}}{=} \det(\mathcal{X})^{-1}\det(\mathcal{A} - \lambda\mathcal{E}_n)\det(\mathcal{X}) = P_{\mathcal{A}}(\lambda). \end{aligned}$$

Da die Eigenwerte sowie deren algebraische Ordnung nur vom charakteristischen Polynom abhängen, stimmen Sie für $\mathcal{A}$ und $\mathcal{C}$ überein. Die geometrischen Ordnungen stimmen ebenfalls überein:

$$\begin{aligned} \dim(U_{\mathcal{A},\lambda_i}) &= \dim \mathrm{Kern}(\mathcal{A} - \lambda_i\mathcal{E}) = n - \mathrm{rg}(\mathcal{A} - \lambda_i\mathcal{E}) \\ &= n - \mathrm{rg}(\mathcal{X}^{-1}(\mathcal{A} - \lambda_i\mathcal{E})\mathcal{X}) = n - \mathrm{rg}(\mathcal{X}^{-1}\mathcal{A}\mathcal{X} - \lambda_i) = \dim(U_{\mathcal{B},\lambda_i}), \end{aligned}$$

da die Multiplikation mit regulären Matrizen von links oder rechts den Rang nicht ändert. $\square$

Definition 8.1.23. Das charakteristische Polynom von $\varphi \in L(V, V)$ ist

$$P_\varphi(\lambda) = P_{\mathcal{M}(\varphi;\mathcal{B},\mathcal{B})}(\lambda)$$

für eine beliebige Basis $\mathcal{B}$ von V.

Nach dem vorigen Satz hängt $P_\varphi(\lambda)$ nicht davon ab, welche Basis man nimmt. Damit ist auch klar, was unter $\det(\varphi)$ und $S(\varphi)$ zu verstehen ist.

Satz 8.1.24. *Es sei $\lambda_0 \in K$ ein Eigenwert von $\varphi \in L(V, V)$ mit algebraischer Ordnung k_0 und dem Eigenraum U_{λ_0}. Dann ist $1 \le \dim U_{\lambda_0} \le k_0$, d.h. es gilt stets*

$$\textit{geometrische Ordnung} \le \textit{algebraische Ordnung.}$$

Beweis. Es sei $\dim V = n$ und $l = \dim U_{\lambda_0}$, sowie $\{\vec{b}_1, \ldots, \vec{b}_l\}$ eine Basis von U_{λ_0}. Diese kann nach Satz 3.3.23 (Austauschsatz von Steinitz) zu einer Basis $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_l, \vec{b}_{l+1}, \ldots, \vec{b}_n\}$ von V ergänzt werden. Dann ist

$$\varphi(\vec{b}_j) = \begin{cases} \lambda_0 \vec{b}_j & \text{für } j = 1, \ldots, l \text{ da } \vec{b}_j \text{ Eigenvektor ist} \\ \sum_{i=1}^n \alpha_{ij} \vec{b}_i & \text{für } j = l+1, \ldots, n \text{ mit } \alpha_{ij} \in K. \end{cases}$$

Also besitzt die Darstellungsmatrix bzgl. $\mathcal{B}$ die Form

$$\mathcal{A} = \mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \left(\begin{array}{ccc|c} \lambda_0 & & & \\ & \ddots & & \mathcal{A}_1 \\ & & \lambda_0 & \\ \hline & 0 & & \mathcal{A}_2 \end{array}\right).$$

Entwickelt man diese Matrix l-mal nach der ersten Spalte, so erhält man

$$P_\mathcal{A}(\lambda) = \det \left(\begin{array}{ccc|c} \lambda_0 - \lambda & & & \\ & \ddots & & \mathcal{A}_1 \\ & & \lambda_0 - \lambda & \\ \hline & 0 & & \mathcal{A}_2 - \lambda \mathcal{E}_{n-l} \end{array}\right) = (\lambda_0 - \lambda)^l \cdot \det(\mathcal{A}_2 - \lambda \mathcal{E}_{n-l}),$$

also $P_\mathcal{A}(\lambda) = (\lambda_0 - \lambda)^l \cdot P_{\mathcal{A}_2}(\lambda)$. Also ist die algebraische Ordnung k_0 mindestens l. Sie stimmt aber nicht notwendig mit k_0 überein, da auch $P_{\mathcal{A}_2}(\lambda)$ noch Faktoren $\lambda_0 - \lambda$ enthalten kann. □

Ohne Beweis geben wir noch den

Satz 8.1.25. *(Cayley-Hamilton)*
Jede Matrix $\mathcal{A} \in K^{(n,n)}$ genügt einer charakteristischen Gleichung, d.h. ist $P_\mathcal{A}(\lambda)$ das charakteristische Polynom von $\mathcal{A}$ mit $P_\mathcal{A}(\lambda) = \alpha_0 + \alpha_1 \lambda + \cdots + \alpha_n \lambda^n$, so gilt

$$P_\mathcal{A}(\mathcal{A}) = \alpha_0 \mathcal{E}_n + \alpha_1 \mathcal{A} + \alpha_2 \mathcal{A}^2 + \cdots + \alpha_n \mathcal{A}^n = 0^{(n,n)}.$$

8.2 Diagonalisierung von Endomorphismen

Definition 8.2.1. Ein $\varphi \in L(V,V)$ heißt diagonalisierbar, falls es eine Basis $\tilde{\mathcal{B}}$ von V gibt, so dass $\mathcal{M}(\varphi;\tilde{\mathcal{B}},\tilde{\mathcal{B}})$ eine Diagonalmatrix ist, d.h.

$$\mathcal{M}(\varphi;\tilde{\mathcal{B}},\tilde{\mathcal{B}}) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

mit $\lambda_1,\ldots,\lambda_n \in K$. Eine Matrix $\mathcal{A} \in K^{(n,n)}$ heißt diagonalisierbar, falls $\mathcal{A}$ zu einer Diagonalmatrix ähnlich ist:

$$\mathcal{A} \approx \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Bemerkung 8.2.2. Offenbar ist φ genau dann diagonalisierbar, wenn $\mathcal{M}(\varphi;\mathcal{B},\mathcal{B})$ für eine beliebige Basis $\mathcal{B}$ diagonalisierbar ist (und dann ist $\mathcal{M}(\varphi;\mathcal{B},\mathcal{B})$ auch für jede Basis $\mathcal{B}$ diagonalisierbar).

Satz 8.2.3. *Ein Endomorphismus $\varphi \in L(V,V)$ ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von φ besitzt.*

Beweis. "$\Rightarrow$":
Es sei $\tilde{\mathcal{B}} = \{\vec{b}'_1,\ldots,\vec{b}'_n\}$ eine Basis von V mit

$$\mathcal{M}(\varphi;\tilde{\mathcal{B}},\tilde{\mathcal{B}}) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Offenbar gilt $\varphi(\vec{b}'_j) = \lambda_j \vec{b}'_j$, womit $\vec{b}'_j$ für $1 \le j \le n$ jeweils Eigenvektor von φ zum Eigenwert λ_j ist.

"$\Leftarrow$":
Es seien $\vec{b}'_1,\ldots,\vec{b}'_n$ linear unabhängige Eigenvektoren von φ zu den (nicht notwendigerweise verschiedenen) Eigenwerten $\lambda_1,\ldots,\lambda_n$. Dann ist $\varphi(\vec{b}'_j) = \lambda_j \vec{b}'_j$, also ist die Darstellungsmatrix

$$\mathcal{M}(\varphi;\tilde{\mathcal{B}},\tilde{\mathcal{B}}) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

d.h. φ ist diagonalisierbar. □

Satz 8.2.4. *(Hauptsatz über die Diagonalisierung von Endomorphismen)*
Ein $\varphi \in L(V,V)$ ist genau dann diagonalisierbar, wenn diese beiden Bedingungen erfüllt sind:

i) Das charakteristische Polynom zerfällt über K vollständig in Linearfaktoren.

ii) Für jeden Eigenwert λ_i stimmt die algebraische mit der geometrischen Ordnung von λ_i überein.

Beweis. Es seien $\lambda_1,\ldots,\lambda_r$ die paarweise verschiedenen Eigenwerte von φ, und k_i die algebraische Ordnung von λ_i. Nach Satz 8.1.24 hat man

$$\dim U_{\lambda_i} \le k_i. \qquad (*)$$

Andererseits besitzt $P_{\mathcal{A}}(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_r)^{k_r} \cdot q(\lambda)$ den Grad n, also gilt

$$k_1 + \cdots + k_r \leq n = \dim V = \deg(P_\varphi)(\lambda). \tag{$**$}$$

Nach Satz 8.2.3 ist φ genau dann diagonalisierbar, wenn

$$\dim U_{\lambda_1} + \cdots + \dim U_{\lambda_r} = n \tag{$***$}$$

gilt.
"$\Rightarrow$":
Ist φ diagonalisierbar, d.h. gilt $(***)$, so ist

$$(k_1 - \dim U_{\lambda_1}) + \cdots + (k_r - \dim U_{\lambda_r}) = \underbrace{\sum_{i=1}^{r} k_i}_{\substack{\leq n \\ \text{wegen } (**)}} - \underbrace{(\dim U_{\lambda_1}) + \cdots + \dim U_{\lambda_r})}_{=n \text{ nach } (***)} \leq n - n = 0.$$

Andererseits sind die Summanden $(k_i - \dim U_{\lambda_i}) \geq 0$ wegen $(*)$. Das ist nur möglich, wenn einerseits $k_i = \dim U_{\lambda_i}$ für $i = 1, \ldots, r$ ist, und andererseits $k_1 + \cdots + k_r = n$. Daraus folgt ii), und i) gilt, da in der Gleichung $P_\varphi(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_r)^{k_r} \cdot q(\lambda)$ der Grad n schon durch die Faktoren $(\lambda - \lambda_i)$ ausgeschöpft wird: so ist $q(\lambda)$ eine Konstante, und P_φ zerfällt in Linearfaktoren.

"$\Leftarrow$":
Aus i) folgt $k_1 + \cdots + k_r = n$, woraus mit ii) die Gleichung $(***)$ folgt, weswegen φ diagonalisierbar ist. $\square$

Bemerkung 8.2.5. Für Matrizen $\mathcal{A} \in K^{(n,n)}$ mit paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r$ gilt

$$\mathcal{A} \text{ ist diagonalisierbar} \Leftrightarrow \begin{cases} \text{i)} & P_{\mathcal{A}}(\lambda) = \prod_{j=1}^{r} (\lambda_j - \lambda)^{k_j}, \quad k_1 + \cdots + k_r = n \\ \text{ii)} & \operatorname{rg}(\mathcal{A} - \lambda_j \mathcal{E}_n) = n - k_j, \quad \text{k=1,...,r} \end{cases}$$

Bemerkung 8.2.6. Es sind $\varphi \in L(V, V)$ bzw. $\mathcal{A} \in K^{(n,n)}$ stets diagonalisierbar, wenn das charakteristische Polynom genau n paarweise verschiedene Nullstellen hat, denn nach Satz 8.2.3 ist dann $1 \leq \dim U_{\lambda_i} \leq k_i = 1$.

Zur Überprüfung der Diagonalisierbarkeit eines $\varphi \in L(V, V)$ verwendet man das folgende praktische Vorgehen:

- Eine Darstellungsmatrix $\mathcal{A}$ von φ bestimmen.
- Die Eigenwerte von φ über das charakteristische Polynom $P_{\mathcal{A}}$ bestimmen.
- Zerfällt $P_{\mathcal{A}}$ in K nicht in Linearfaktoren, d.h. hat es unter Berücksichtigung der algebraischen Ordnungen keine n Nullstellen in K, so ist φ nicht diagonalisierbar.
- Eigenräume zu den Eigenwerten bestimmen. Ist $\dim U_{\lambda_i} < k_i$ für ein i, so ist φ nicht diagonalisierbar.

- Ist φ diagonalisierbar, so stehen in der Diagonalmatrix die Eigenwerte. Basen der Eigenräume bestimmen und als Spalten in $\mathcal{X}$ in der Reihenfolge, wie in der Diagonalmatrix $\mathcal{D}$ die Eigenwerte stehen, eintragen:

$$\mathcal{X} = \left(\underbrace{\vec{b}_1^{(1)}, \dots, \vec{b}_{l_1}^{(1)}}_{\text{Basis von } U_{\lambda_1}}, \underbrace{\vec{b}_1^{(2)}, \dots, \vec{b}_{l_2}^{(2)}}_{\text{Basis von } U_{\lambda_2}}, \dots, \underbrace{\vec{b}_1^{(r)}, \dots, \vec{b}_{l_r}^{(r)}}_{\text{Basis von } U_{\lambda_r}} \right).$$

  Dann ist wie gewünscht $\mathcal{X}^{-1} \cdot \mathcal{A} \cdot \mathcal{X} = \mathcal{D}$. Damit ist dann auch $\mathcal{D} = \mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}})$ diagonal für die Basis $\tilde{\mathcal{B}}$ aus den in $\mathcal{X}$ stehenden Basisvektoren.

Wir demonstrieren dieses Vorgehen an einigen Beispielen:

Beispiel 8.2.7. Wir betrachten

$$\mathcal{A} = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} \in \mathbb{R}^{(2,2)}.$$

Wir wollen $\mathcal{A}$ diagonalisieren.

- Eigenwerte bestimmen:
  Das charakteristische Polynom ist

$$P_{\mathcal{A}}(\lambda) = \det(\mathcal{A} - \lambda \mathcal{E}_2) = \det \begin{pmatrix} 1-\lambda & 3 \\ 0 & 2-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda),$$

  die Eigenwerte sind also $\lambda_1 = 1$ und $\lambda_2 = 2$.

- Ordnungen bestimmen:
  Die algebraischen Ordnungen $k_1 = k_2 = 1$ liest man aus dem Polybom $P_{\mathcal{A}}$ ab. Wegen $1 \leq l_i \leq k_i$ nach Satz 8.2.3 ist auch $l_1 = l_2 = 1$. Da die geometrischen und algebraischen Ordnungen übereinstimmen, ist die Matrix $\mathcal{A}$ diagonalisierbar und besitzt die Diagonalgestalt

$$\mathcal{D} = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

- Transformationsmatrix bestimmen:
  Wir müssen die Eigenräume ausrechnen:

$$U_{\lambda_1} = \{\vec{x} \in \mathbb{R}^2 \colon \mathcal{A}\vec{x} = \vec{x}\} = \operatorname{Kern}(\mathcal{A} - \mathcal{E}_n) = \operatorname{Kern} \begin{pmatrix} 0 & 3 \\ 0 & 1 \end{pmatrix}.$$

  Man sieht sofort, dass U_{λ_1} von $\vec{b}_1^{(1)} = (1,0)^T$ aufgespannt wird. Der Eigenraum

$$U_{\lambda_2} = \{\vec{x} \in \mathbb{R}^2 \colon \mathcal{A}\vec{x} = 2\vec{x}\} = \operatorname{Kern}(\mathcal{A} - 2\mathcal{E}_n) = \operatorname{Kern} \begin{pmatrix} -1 & 3 \\ 0 & 0 \end{pmatrix}$$

  wird dagegen von $\vec{b}_1^{(2)} = (3,1)^T$ aufgespannt. Damit haben wir mit

$$\tilde{\mathcal{B}} = \left\{ \vec{b}_1^{(1)}, \vec{b}_1^{(2)} \right\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\}$$

  eine Basis des $\mathbb{R}^2$ aus Eigenvektoren von $\mathcal{A}$. Einsetzen der Vektoren in die Transformationsmatrix ergibt

$$\mathcal{X} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}.$$

  Die gesuchte Diagonalisierung von $\mathcal{A}$ ist somit

$$\mathcal{X}^{-1} \cdot \mathcal{A} \cdot \mathcal{X} = \mathcal{D} = \begin{pmatrix} 1 & \\ & 2 \end{pmatrix}.$$

Beispiel 8.2.8. Es sei $V = \mathbb{R}^3$, $\varphi \in L(V,V)$ mit $\varphi(\vec{x}) = \mathcal{A}\vec{x}$ und der Matrix

$$\mathcal{A} = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}.$$

Wir wollen φ diagonalisieren.

- Startbasis wählen:
  Damit man die geometrischen Ordnungen über den Rang bestimmen kann, muss man eine Darstellungsmatrix von φ wählen. Nimmt man die Standardbasis $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$, so ist $\mathcal{M}(\varphi; \mathcal{B}, \mathcal{B}) = \mathcal{A}$.

- Eigenwerte bestimmen:
  Die Eigenwerte sind die Nullstellen des charakteristischen Polynoms von $\mathcal{A}$. Es empfiehlt sich, das Polynom gleich in Produktform zu bestimmen:

$$\begin{aligned} P_\varphi(\lambda) &= P_{\mathcal{A}}(\lambda) = \det \begin{pmatrix} -\lambda & -1 & 1 \\ -3 & -2-\lambda & 3 \\ -2 & -2 & 3-\lambda \end{pmatrix} = \det \begin{pmatrix} -\lambda & 0 & 1 \\ -3 & 1-\lambda & 3 \\ -2 & 1-\lambda & 3-\lambda \end{pmatrix} \\ &= (1-\lambda) \cdot \det \begin{pmatrix} -\lambda & 0 & 1 \\ -3 & 1 & 3 \\ -2 & 1 & 3-\lambda \end{pmatrix} = (1-\lambda) \cdot \det \begin{pmatrix} -\lambda & 0 & 1 \\ -3 & 1 & 3 \\ 1 & 0 & -\lambda \end{pmatrix} \\ &\underset{\text{entwickeln}}{\overset{\text{2. Spalte}}{=}} (1-\lambda) \cdot \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = (1-\lambda) \cdot (\lambda^2 - 1) = -(\lambda-1)^2 \cdot (\lambda+1). \end{aligned}$$

  Die Eigenwerte sind also $\lambda_1 = 1$ und $\lambda_2 = -1$.

- Ordnungen bestimmen:
  Die algebraischen Ordnungen $k_1 = 2$ und $k_2 = 1$ liest man einfach aus dem Polynom ab. Die geometrische Ordnung von λ_2 muss wegen $k_2 = 1$ und Satz 8.2.3 gleich eins sein. Für die geometrische Ordnung $l_1 = \dim U_{\lambda_1}$ gibt es dagegen die beiden Möglichkeiten $l_1 = 1$ und $l_1 = 2$. Es gilt $l_1 = 3 - \operatorname{rg}(\mathcal{A} - \lambda_1\mathcal{E}_3)$, was ist nach Satz 6.1.13 die Dimension der Lösungsmenge des homogenen LGS $(\mathcal{A} - \lambda_1\mathcal{E}_3)\vec{x} = \vec{0}$ ist. Wir bestimmen die Dimension der Lösungsmenge mit dem Gaußschen Algorithmus:

$$\mathcal{A} - \lambda_1\mathcal{E}_3 = \mathcal{A} - \mathcal{E}_3 = \begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} \underset{\substack{(II)-3\cdot(I) \\ (III)-2\cdot(I)}}{\longrightarrow} \begin{pmatrix} -1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \operatorname{rg}(\mathcal{A} - \lambda_1\mathcal{E}_3) = 1,$$

  also ist $l_1 = 3-1 = 2$. Da die geometrischen und algebraischen Ordnungen somit übereinstimmen, ist auch der zweite Teil des Hauptsatzes erfüllt, womit φ diagonalisierbar ist. Die Diagonalmatrix ist

$$\mathcal{D} = \begin{pmatrix} \lambda_1 & & \\ & \lambda_1 & \\ & & \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

- Transformationsmatrix $\mathcal{X} = \mathcal{M}(id; \mathcal{B}, \tilde{\mathcal{B}})$ bestimmen:
  Wir müssen die Eigenräume zu λ_1 und λ_2 bestimmen. Wegen $l_1 = 2$ gibt es zwei linear unabhängige Lösungsvektoren des Systems $(\mathcal{A} - \lambda_1\mathcal{E}_3)\vec{x} = \vec{0}$. Da bei der Rangbestimmung keine nur Zeilenumformungen verwendet wurden, ist

$$(\mathcal{A} - \lambda_1\mathcal{E}_3)\vec{x} = \vec{0} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \cdot \vec{x} = \vec{0}.$$

Die Lösungsmenge besteht also aus allen Vektoren $(x, y, z)^T \in \mathbb{R}^3$, für die $-x - y + z = 0$ gilt. Man sieht sofort, dass

$$\vec{b}_1^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{b}_2^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

zwei linear unabhängige Lösungsvektoren sind, womit sie eine Basis des Eigenraums U_{λ_1} bilden. Für den Eigenwert λ_2 ist dagegen $l_2 = 1$, weswegen wir also $\mathrm{rg}(\mathcal{A} - \lambda_2 \mathcal{E}_3) = 2$ erwarten. Es ist auch

$$\mathcal{A} - \lambda_2 \mathcal{E}_3 = \mathcal{A} + \mathcal{E} = \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} \underset{\substack{(II)+3\cdot(I) \\ (III)+2\cdot(I)}}{\longrightarrow} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -4 & 6 \\ 0 & -4 & 6 \end{pmatrix} \underset{\substack{(III)-(II) \\ \frac{1}{2}\cdot(II)}}{\longrightarrow} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -2 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ein von dieser Matrix annulierter Vektor ist beispielsweise $\vec{b}_1^{(2)} = (1, 3, 2)^T$, er erzeugt den eindimensionalen Raum U_{λ_2}. Damit ist die Basis

$$\tilde{\mathcal{B}} = \left\{ \vec{b}_1^{(1)}, \vec{b}_2^{(1)}, \vec{b}_1^{(2)} \right\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$$

komplett. Ihr entspricht die Transformationsmatrix

$$\mathcal{X} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}$$

mit

$$\mathcal{X}^{-1} \cdot \mathcal{A} \cdot \mathcal{X} = \mathcal{M}(\varphi; \tilde{\mathcal{B}}, \tilde{\mathcal{B}}) = \mathcal{D} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}.$$

Beispiel 8.2.9. Wir betrachten

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix} = \mathcal{E}_n + E_{ij},$$

wobei E_{ij} die Matrix mit genau einer Eins an der Stelle (i, j) ist. Das charakteristische Polynom ist offensichtlich

$$P_{\mathcal{A}}(\lambda) = \det(\mathcal{A} - \lambda \mathcal{E}_n) = (1 - \lambda)^n,$$

also gibt es nur den Eigenwert $\lambda_1 = 1$, er besitzt die algebraische Ordnung $k_1 = n$. Seine geometrische Ordnung ist dagegen

$$l_1 = n - \mathrm{rg}(\mathcal{A} - \lambda_1 \mathcal{E}_n) = n - \mathrm{rg}(\mathcal{A} - \mathcal{E}_n) = n - \mathrm{rg}(E_{ij}) = n - 1.$$

Wegen $l_1 \neq k_1$ ist $\mathcal{A}$ nicht diagonalisierbar.

Bemerkung 8.2.10. Symmetrische Matrizen $\mathcal{A} = \mathcal{A}^T$ lassen sich stets und besonders einfach auf Diagonalgestalt transformieren. Dies wird als Hauptachsentransformation bezeichnet.