# Lineare Algebra für Physik-Studiengänge

## Ein Kompendium zur Vorlesung von L. Recke

### Inhalt



# 1 Vektorräume

**Axiomensystem des "abstrakten" Vektorraums** Eine Menge $V$ mit einer Abbildung $(\boldsymbol{u},\boldsymbol{v}) \in V \times V \mapsto \boldsymbol{u} + \boldsymbol{v} \in V$ (genannt Addition) und einer Abbildung $(\lambda,\boldsymbol{v}) \in \mathbb{R} \times V \mapsto \lambda\boldsymbol{v} \in V$ (genannt Multiplikation mit Skalaren) heißt Vektorraum, wenn folgende Eigenschaften erfüllt sind:

(I) Für alle $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in V$ gilt $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$.
(II) Für alle $\boldsymbol{u},\boldsymbol{v} \in V$ gilt $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$.
(III) Es existiert ein $\boldsymbol{0} \in V$, so daß für alle $\boldsymbol{v} \in V$ gilt $\boldsymbol{0}+\boldsymbol{v}=\boldsymbol{v}$.
(IV) Für alle $\boldsymbol{v} \in V$ existiert ein $-\boldsymbol{v} \in V$, so daß gilt $-\boldsymbol{v}+\boldsymbol{v}=\boldsymbol{0}$.
(V) Für alle $\lambda, \mu \in \mathbb{R}$ und $\boldsymbol{v} \in V$ gilt $(\lambda\mu)\boldsymbol{v}= \lambda(\mu\boldsymbol{v})$.
(VI) Für alle $\lambda, \mu \in \mathbb{R}$ und $\boldsymbol{v}\in V$ gilt $(\lambda+\mu)\boldsymbol{v}= \lambda\boldsymbol{v}+\mu\boldsymbol{v}$.
(VII) Für alle $\lambda \in \mathbb{R}$ und $\boldsymbol{u},\boldsymbol{v} \in V$ gilt $\lambda(\boldsymbol{u}+\boldsymbol{v}) = \lambda\boldsymbol{u}+\lambda\boldsymbol{v}$.
(VIII) Für alle $\boldsymbol{v} \in V$ gilt $1\boldsymbol{v}=\boldsymbol{v}$.

Die Elemente von $V$ heißen Vektoren. Im Gegensatz dazu werden Zahlen auch Skalare genannt.

**Elementare Folgerungen aus den Axiomen** (i) Wenn für ein $\boldsymbol{u} \in V$ und für alle $\boldsymbol{v} \in V$ gilt, daß $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}$ ist, so folgt $\boldsymbol{u}=\boldsymbol{0}$. Mit anderen Worten: Der besondere Vektor $\boldsymbol{0} \in V$, dessen Existenz in (III) postuliert ist, ist eindeutig bestimmt. Dieser Vektor heißt Nullvektor.
(ii) Wenn für $\boldsymbol{u},\boldsymbol{v} \in V$ gilt daß $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{0}$ ist, so folgt $\boldsymbol{u}= -\boldsymbol{v}$. Mit anderen Worten: Der Vektor $-\boldsymbol{v} \in V$, dessen Existenz in (IV) postuliert ist, ist, bei gegebenem $\boldsymbol{v}$, eindeutig bestimmt. Dieser Vektor heißt zu $\boldsymbol{v}$ entgegengesetzter Vektor. Anstelle von $\boldsymbol{u}+(-\boldsymbol{v})$ schreibt man einfacher $\boldsymbol{u}-\boldsymbol{v}$.
(iii) Für beliebige $\boldsymbol{u},\boldsymbol{v} \in V$ und $\lambda \in \mathbb{R}$ mit $\lambda \neq 0$ ist die Gleichung $\lambda\boldsymbol{x}+\boldsymbol{u}=\boldsymbol{v}$ nach der "Unbekannten" $\boldsymbol{x} \in V$ auflösbar: $\boldsymbol{x}= \frac{1}{\lambda}(\boldsymbol{v}-\boldsymbol{u})$.
(iv) Für beliebige $\boldsymbol{v} \in V$ und $\lambda \in \mathbb{R}$ gilt $0\boldsymbol{v}=\boldsymbol{0}$ und $\lambda\boldsymbol{0}=\boldsymbol{0}$ .

**Bemerkung zur Bezeichnungsweise** Das Symbol + wird in zweierlei Sinn benutzt, nämlich als Addition zwischen Zahlen und zwischen Vektoren. Zum Beispiel steht im Axiom (VI) links vom Gleichheitszeichen in der Klammer eine Summe von Zahlen, dagegen steht rechts vom Gleichheitszeichen eine Summe von Vektoren.

**Beispiel: Der Vektorraum $\mathbb{R}^n$** Die Menge $\mathbb{R}^n$ der $n$-Tupel reeller Zahlen ist ein Vektorraum mit den "komponentenweise" definierten Operationen

$$\begin{aligned} (u_1, \ldots, u_n) + (v_1, \ldots, v_n) &= (u_1 + v_1, \ldots, u_n + v_n), \\ \lambda(v_1, \ldots, v_n) &= (\lambda v_1, \ldots, v_n). \end{aligned}$$

Hier stehen links von den Gleichheitszeichen die Operationen "Addition von Zahlentupeln" und "Multiplikation von Zahlen mit Zahlentupeln", die zu definieren sind, und rechts stehen die bekannten Summen und Produkte von Zahlen. Der Nullvektor in diesem Vektorraum ist das Zahlentupel $(0, 0, \ldots, 0)$. Das obige Axiomensystem und die darauf aufbauende Theorie der Vektorräume ist weitgehend durch Abstraktion aus dem Vektorraum $\mathbb{R}^n$ entstanden.
Die Bezeichnung "Vektor" wird sowohl für Elemente eines abstrakten Vektorraumes als auch für Elemente des $\mathbb{R}^n$, also für Zahlentupel, verwendet. Dabei werden z.B. Zahlentripel $(v_1, v_2, v_3)$ entsprechend der geometrischen Anschauung mit Punkten im 3-dimensionalen Raum oder mit Pfeilen (im Nullpunkt fest verankert) oder mit Äquivalenzklassen von Pfeilen (zwei Pfeile sind äquivalent, wenn sie durch Parallelverschiebung ineinander überführt werden können) identifiziert.

**Kartesisches Produkt zweier Vektorräume** Wenn $V$ und $W$ zwei Vektorräume sind, so ist auch ihr sogenanntes kartesisches Produkt

$$V \times W = \{(\boldsymbol{v}, \boldsymbol{w}) : \; \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$$

ein Vektorraum mit den "komponentenweise" definierten Operationen

$$\begin{aligned} (\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2) &= (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2), \\ \lambda(\boldsymbol{v}, \boldsymbol{w}) &= (\lambda \boldsymbol{v}, \lambda \boldsymbol{w}). \end{aligned}$$

Analog macht man mit Hilfe der algebraischen Operationen in $V$ das $n$-fache kartesische Produkt $V \times \cdots \times V$ ($n$ Faktoren) zu einem Vektorraum, und diesen bezeichnet man mit $V^n$. Offenbar ist die Bezeichnung $\mathbb{R}^n$ ein Spezialfall dieser Bezeichnungsweise.

**Beispiel: Ein Funktionenraum** Die Menge aller Funktionen $f : \mathbb{R} \to \mathbb{R}$ ist ein Vektorraum mit den "punktweiseweise" definierten Operationen

$$\left.\begin{aligned} (f + g)(x) &= f(x) + g(x), \\ (\lambda f)(x) &= \lambda f(x), \end{aligned}\right\} \text{ für alle } x \in \mathbb{R}.$$

Hier stehen wieder links von den Gleichheitszeichen die Operationen "Addition von Funktionen" und "Multiplikation von Zahlen mit Funktionen", die zu definieren sind, und rechts stehen die bekannten Summen und Produkte von Zahlen. Der Nullvektor in diesem Vektorraum ist die Funktion, die identisch gleich Null ist.

**Beispiel: Der Vektorraum der Polynome** Die Menge $\mathbb{R}[x]$ aller Polynome mit reellen Koeffizienten (und mit der sogenannten unabhängigen Variablen $x$) ist ein Vektorraum mit den Operationen

$$\begin{aligned}(\alpha_0 + \ldots + \alpha_m x^m) + (\beta_0 + \ldots + \beta_n x^n) &= \alpha_0 + \beta_0 + \ldots + (\alpha_m + \beta_m)x^m + \beta_{m+1}x^{m+1} + \ldots + \beta_n x^n, \\ \lambda(\alpha_0 + \alpha_1 x + \ldots + \alpha_m x^m) &= \lambda\alpha_0 + \lambda\alpha_1 x + \ldots + \lambda\alpha_m x^m.\end{aligned}$$

Dabei ist in der Definitionsgleichung für die Addition ohne Beschränkung der Allgemeinheit nur der Fall $m \leq n$ betrachtet. Der Nullvektor in diesem Vektorraum ist das Polynom, dessen Koeffizienten alle Null sind.

**Unterräume** Eine Untermenge $U$ eines Vektorraumes $V$ heißt Unterraum von $V$, wenn für alle $\boldsymbol{u}, \boldsymbol{v} \in U$ und $\lambda \in \mathbb{R}$ gilt $\boldsymbol{u} + \boldsymbol{v} \in U$ und $\lambda \boldsymbol{u} \in U$. Man kommt also nicht aus $U$ heraus, indem man mit Elementen von $U$ die Operationen "Addition" und "Multiplikation mit Skalaren" ausführt. Insbesondere ist $U$ mit den von $V$ "geerbten" Operationen wieder ein Vektorraum.

**Affine Unterräume** Eine Untermenge $A$ eines Vektorraumes $V$ heißt affiner Unterraum von $V$, wenn ein $\boldsymbol{v} \in V$ und ein Unterraum $U$ von $V$ existieren, so daß gilt $A = \{\boldsymbol{v} + \boldsymbol{u} \colon \boldsymbol{u} \in U\}$. Dann schreibt man auch

$$A = \boldsymbol{v} + U. \tag{1.1}$$

**Bemerkungen zur Bezeichnungsweise** (i) Ein Unterraum ist auch ein affiner Unterraum, aber ein affiner Unterraum ist im allgemeinen kein Unterraum. Die in der obigen Definition eingeführte und allgemein übliche Sprachregelung ist deshalb etwas unlogisch.
(ii) In der Darstellung (1.1) eines affinen Unterraues $A$ mit Hilfe eines Vektors $\boldsymbol{v}$ und eines Unterraumes $U$ ist $U$ eindeutig bestimmt, nicht aber $\boldsymbol{v}$. Es gilt nämlich $\boldsymbol{v}_1 + U_1 = \boldsymbol{v}_2 + U_2$ genau dann, wenn die Unterräume $U_1$ und $U_2$ gleich sind und wenn $\boldsymbol{v}_1 - \boldsymbol{v}_2$ in diesem Unterraum liegt. In diesem Sinne sagt man, daß $U$ "der" Unterraum zu dem affinen Unterraum $A = \boldsymbol{v} + U$ ist, während $\boldsymbol{v}$ "eine" Translation ist, durch die $A$ aus $U$ hervorgeht.

**Beispiel: Unterräume und affine Unterräume in $\mathbb{R}^2$** In $\mathbb{R}^2$ existieren genau folgende Unterräume: Die trivialen Unterräume $\{(0,0)\}$ und $\mathbb{R}^2$ sowie die Geraden durch den Nullpunkt. Eine Untermenge des $\mathbb{R}^2$ ist genau dann affiner Unterraum, wenn sie mit einem dieser Unterräume zusammenfällt oder wenn sie eine Gerade ist, die nicht den Nullpunkt enthält.

**Beispiel: Unterräume und affine Unterräume im Funktionenraum** Die stetigen Funktionen sind ein Unterraum im Vektorraum aller Funktionen $f : \mathbb{R} \to \mathbb{R}$. Dieser Unterraum wird mit $C(\mathbb{R})$ bezeichnet. In $C(\mathbb{R})$ bilden die stetig differenzierbaren Funktionen wiederum ein Unterraum. Dieser Unterraum wird mit $C^1(\mathbb{R})$ bezeichnet. Faßt man die Polynome als Funktionen auf, so ist $\mathbb{R}[x]$ ein Unterraum in $C^1(\mathbb{R})$. Schließlich bilden die Polynome, deren Grad nicht größer als eine fixierte natürliche Zahl $n$ ist, wiederum einen Unterraum in $\mathbb{R}[x]$. Dieser Unterraum wird mit $\mathbb{R}_n[x]$ bezeichnet.
Im Vektorraum aller Funktionen $f : \mathbb{R} \to \mathbb{R}$ ist ferner die Untermenge $\{f : \ f(0) = 0\}$ ein Unterraum. Die Untermenge $\{f : \ f(0) = 1\}$ ist kein Unterraum, aber ein affiner Unterraum. Wenn man mit dem Symbol 1 nicht nur die Zahl Eins, sondern auch die Funktion, die identisch Eins ist, bezeichnet, so gilt, entsprechend der Bezeichnungsweise (1.1),

$$\{f : \ f(0) = 1\} = 1 + \{f : \ f(x) = 0\}.$$

**Linearkombinationen und lineare Hülle** Es seien $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ Elemente eines Vektorraumes $V$ und $\lambda_1, \cdots, \lambda_n$ Zahlen. Dann nennt man die Summe $\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n$ Linearkombination der Vektoren $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$. Die Menge aller solcher Linearkombinationen ist ein Unterraum von $V$, heißt lineare Hülle der Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ und wird mit $\mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ bezeichnet, d.h.

$$\mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} = \{\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n : \ \lambda_j \in \mathbb{R}\}.$$

**Durchschnitt und Summe von Unterräumen** Wenn $U_1$ und $U_2$ zwei Unterräume eines Vektorraumes $V$ sind, so ist auch ihr Durchschnitt $U_1 \cap U_2$ ein Unterraum von $V$. Ferner ist auch ihre sogenannte Summe

$$U_1 + U_2 = \{\boldsymbol{u}_1 + \boldsymbol{u}_2 : \ \boldsymbol{u}_1 \in U_1, \boldsymbol{u}_2 \in U_2\}$$

ein Unterraum von $V$. Dagegen ist die Vereinigung $U_1 \cup U_2$ im allgemeinen kein Unterraum.

**Komplement, direkte Summe und Projektor** Es seien $U_1$ und $U_2$ zwei Unterräume eines Vektorraumes $V$. Dann gilt $U_1 \cap U_2 = \{0\}$ genau dann, wenn für alle $\boldsymbol{v} \in U_1 + U_2$ genau ein $\boldsymbol{u}_1 \in U_1$ und genau ein $\boldsymbol{u}_2 \in U_2$ existieren, so daß

$$\boldsymbol{v} = \boldsymbol{u}_1 + \boldsymbol{u}_2 \tag{1.2}$$

ist. In diesem Fall heißt $U_1 + U_2$ direkte Summe von $U_1$ und $U_2$, man schreibt $U_1 \oplus U_2$ anstelle von $U_1 + U_2$, und die Abbildungen $\boldsymbol{v} \in V \mapsto \boldsymbol{u}_1 \in U_1$ bzw. $\boldsymbol{v} \in V \mapsto \boldsymbol{u}_2 \in U_2$, die durch (1.2) definiert sind, heißen Projektoren auf $U_1$ entlang $U_2$ bzw. auf $U_2$ entlang $U_1$.
Wenn $U_1 \oplus U_2 = V$ ist, so heißen $U_1$ bzw. $U_2$ Komplement zu $U_2$ bzw. $U_1$ in $V$.

**Beispiel** Wenn $U_1$ und $U_2$ zwei Ebenen in $\mathbb{R}^3$ sind, die den Nullpunkt enthalten und die nicht gleich sind, so gilt $U_1 + U_2 = \mathbb{R}^3$, aber $\mathbb{R}^3$ ist nicht die direkte Summe von $U_1$ und $U_2$.

**Beispiel** Wenn $U_1$ eine Gerade in $\mathbb{R}^2$ durch den Nullpunkt ist, so ist $U_2$ ein Komplement zu $U_1$ in $\mathbb{R}^2$ genau dann, wenn $U_2$ eine andere Gerade durch den Nullpunkt ist.

**Beispiel** Die Menge aller Polynome der Art $\alpha_{n+1}x^{n+1} + \alpha_{n+2}x^{n+2} + \cdots + \alpha_m x^m$ ($n$ fixiert, $m \geq n$ beliebig) ist ein Komplement zu $\mathbb{R}_n[x]$ in $\mathbb{R}[x]$.

**Lineare Abhängigkeit** Elemente $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ eines Vektoraumes $V$ heißen linear abhängig (bzw. linear unabhängig), wenn Zahlen $\lambda_1, \ldots, \lambda_n$ existieren, so daß $\lambda_1^2 + \cdots + \lambda_n^2 > 0$ und $\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n = \boldsymbol{0}$ gilt (bzw. wenn solche Zahlen nicht existieren). Dies ist genau dann der Fall, wenn ein $j \in \{1, \ldots, n\}$ existiert, so daß gilt

$$\boldsymbol{v}_j \in \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1}, \boldsymbol{v}_{j+1}, \ldots, \boldsymbol{v}_n\}$$

(bzw. wenn ein solches $j$ nicht existiert).

**Beispiel** Zwei Vektoren in $\mathbb{R}^2$ sind linear abhängig genau dann, wenn sie auf einer gemeinsamen Geraden, die den Nullpunkt enthält, liegen.

**Beispiel** Die Funktionen $f(x) = \sin^2 x$ und $g(x) = \cos^2 x$ sind linear unabhängig, dagegen sind sie gemeinsam mit der Funktion $h(x) = 1$ linear abhängig.

**Beispiel** Die Funktionen $f_j(x) = \frac{1}{1+jx^2}$, $(j = 0, 1, \ldots, n)$ sind linear unabhängig.

**Beispiel** Die Funktionen $1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots, \sin nx, \cos nx$ sind linear unabhängig.

**Basen und Koordinaten** Für Elemente $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ eines Vektorraumes $V$ sind folgende Bedingungen äquivalent:
(i) Die Vektoren $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ sind linear unabhängig, und gemeinsam mit jedem weiteren Vektor sind sie linear abhängig.
(ii) Für jeden Vektor $\boldsymbol{v} \in V$ existiert genau ein Zahlentupel $(v_1, \ldots, v_n) \in \mathbb{R}^n$, so daß $\boldsymbol{v} = v_1\boldsymbol{b}_1 + \ldots + v_n\boldsymbol{b}_n$ gilt.
In diesem Fall sagt man, daß die Vektoren $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ eine Basis in $V$ bilden, und die Zahlen $v_1, \ldots, v_n$ heißen Koordinaten des Vektors $\boldsymbol{v}$ bezüglich dieser Basis.

**Beispiel: Die kanonische Basis in $\mathbb{R}^n$** Die Vektoren

$$\boldsymbol{e}_1 = (1, 0, \ldots, 0),\ \boldsymbol{e}_2 = (0, 1, \ldots, 0),\ \ldots,\ \boldsymbol{e}_n = (0, 0, \ldots, 1) \tag{1.3}$$

bilden eine Basis in $\mathbb{R}^n$, die sogenannte kanonische Basis. Die Koordinaten eines Tupels $(v_1, \ldots, v_n) \in \mathbb{R}^n$ bezüglich dieser Basis sind die Zahlen $v_1, \ldots, v_n$, denn es gilt

$$(v_1, \ldots, v_n) = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + \ldots + v_n\boldsymbol{e}_n.$$

**Beispiel** Die Polynome $1, x, x^2, \ldots, x^n$ bilden eine Basis in $\mathbb{R}_n[x]$. Die Koordinaten eines Polynoms $\alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$ bezüglich dieser Basis sind die Zahlen $\alpha_0, \ldots, \alpha_n$. Eine andere Basis in $\mathbb{R}_n[x]$ bilden z.B. die Polynome $(x+1)^n, (x+2)^n, \ldots, (x+n+1)^n$.

**Dimensionen** Wenn $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ und $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ zwei Basen in einem Vektorraum $V$ sind, so gilt $m = n$. Diese Zahl heißt Dimension von $V$ und wird mit $\dim V$ bezeichnet, und $V$ heißt dann endlich-dimensional. Wenn in $V$ beliebig viele linear unabhängige Vektoren existieren, so heißt $V$ unendlich-dimensional.
Unter der Dimension eines affinen Unterraums versteht man die Dimension des entsprechenden Unterraums, d.h.

$$\dim(\boldsymbol{v} + U) = \dim U. \tag{1.4}$$

**Beispiel** $\dim \mathbb{R}^n = n$

**Beispiel** $\mathbb{R}_n[x]$ besitzt die Dimension $n + 1$. Dagegen ist $\mathbb{R}[x]$ unendlich-dimensional, weil für beliebiges $n$ die Polynome $1, x, x^2, \ldots, x^n$ linear unabhängig sind.

**Beispiel** Es gilt $\dim \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} = n$ genau dann, wenn die Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ linear unabhängig sind. Andernfalls ist $\dim \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} < n$.

**Dimension des kartesischen Produkts** Wenn $V$ und $W$ zwei endlich-dimensionale Vektorräume sind, so gilt

$$\dim(V \times W) = \dim V + \dim W. \tag{1.5}$$

**Dimensionsformel für Unterräume** Für zwei Unterräume $U_1$ und $U_2$ eines Vektorraumes gilt stets

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

Insbesondere gilt $\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2$, denn die Benutzung der Bezeichnung $U_1 \oplus U_2$ impliziert, daß $U_1 \cap U_2 = \{0\}$ ist.

**Abschätzung der Dimension des Durchschnitts von Unterräumen** Wenn $U_1$ und $U_2$ zwei Unterräume in einem Vektorraum $V$ sind, so gilt $\dim(U_1 \cap U_2) = \dim U_1 + \dim U_2 - \dim(U_1 + U_2) \geq \dim V - (\dim V - \dim U_1) - (\dim V - \dim U_2)$. Analog gilt

$$\dim(U_1 \cap \ldots \cap U_m) \geq \dim V - n_1 - \ldots - n_m, \quad \text{falls} \quad \dim U_j = \dim V - n_j \quad \text{für} \quad j = 1, \ldots, m. \tag{1.6}$$

**Reelle und komplexe Vektorräume** Alle bisherigen Resultate gelten unverändert, wenn überall $\mathbb{R}$ durch $\mathbb{C}$ ersetzt wird. Man spricht dann von komplexen Vektorräumen (oder Vektorräumen über $\mathbb{C}$) bzw., bei Vektorräumen wie bisher, von reellen Vektorräumen (oder Vektorräumen über $\mathbb{R}$).

**Beispiele** Mit wie oben eingeführten Operationen sind die Menge $\mathbb{C}^n$ der $n$-Tupel komplexer Zahlen, die Menge der Funktionen $f : \mathbb{C} \to \mathbb{C}$, die Menge $\mathbb{C}[x]$ der Polynome mit komplexen Koeffizienten sowie die Menge $\mathbb{C}_n[x]$ der Polynome mit komplexen Koeffizienten von Grad kleiner oder gleich $n$ komplexe Vektorräume.

**Reellifizierung komplexer Verktorräume** Es sei $V$ ein komplexer Vektorraum. Dann erfüllt die Einschränkung auf $\mathbb{R} \times V$ der Operation "Multiplikation von Vektoren mit komplexen Zahlen" (gemeinsam mit der in $V$ gegebenen Addition) alle Axiome eines reellen Vektorraumes. Der so entstehende reelle Vektorraum wird mit $V_{\mathbb{R}}$ bezeichnet, und es gelten die folgenden Eigenschaften:
(i) Es sei $\boldsymbol{v} \in V$ mit $\boldsymbol{v} \neq \boldsymbol{0}$, und $i$ sei die imaginäre Einheit (d.h. $i^2 = -1$). Dann sind die Vektoren $\boldsymbol{v}$ und $i\boldsymbol{v}$ linear abhängig in $V$, aber linear unabhängig in $V_{\mathbb{R}}$.
(ii) $\dim V_{\mathbb{R}} = 2 \dim V$.
(iii) Es sei $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ eine Basis in $V$. Dann ist z.B. $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n, i\boldsymbol{b}_1, \ldots, i\boldsymbol{b}_n$ eine Basis in $V_{\mathbb{R}}$.
(iv) Die lineare Hülle von Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in V$ in $V$ ist Menge aller Linearkombinationen $\lambda_1\boldsymbol{v}_1 + \cdots + \lambda_n\boldsymbol{v}_n$ mit $\lambda_j \in \mathbb{C}$. Die lineare Hülle dieser Vektoren in $V_{\mathbb{R}}$ dagegen ist Menge aller Linearkombinationen $\lambda_1\boldsymbol{v}_1 + \cdots + \lambda_n\boldsymbol{v}_n$ mit $\lambda_j \in \mathbb{R}$ und folglich kein Unterraum in $V$.

**Komplexifizierung reeller Vektorräume** Es sei $V$ ein reeller Vektorraum. Dann wird die Menge aller Paare $(\boldsymbol{v}, \boldsymbol{w}) \in V^2$ zu einem komplexen Vektorraum, wenn man die Addition bzw. die Multiplikation mit komplexen Zahlen definiert durch

$$\begin{aligned} (\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2) &= (\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2), \\ (\lambda + i\mu)(\boldsymbol{v}, \boldsymbol{w}) &= (\lambda\boldsymbol{v} - \mu\boldsymbol{w}, \lambda\boldsymbol{w} + \mu\boldsymbol{v}). \end{aligned}$$

Der so entstehende komplexe Vektorraum heißt Komplexifizierung von $V$ und wird mit $V_{\mathbb{C}}$ bezeichnet. Anstelle von $(\boldsymbol{v}, \boldsymbol{0})$ bzw. $(\boldsymbol{0}, \boldsymbol{w})$ schreibt man einfach $\boldsymbol{v}$ bzw. $i\boldsymbol{w}$ und folglich $\boldsymbol{v} + i\boldsymbol{w}$ anstelle von $(\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v}, 0) + (0, \boldsymbol{w})$. Mit diesen Schreibweisen gehen die oben definierten Operationen über in

$$\begin{aligned} (\boldsymbol{v}_1 + i\boldsymbol{w}_1) + (\boldsymbol{v}_2 + i\boldsymbol{w}_2) &= \boldsymbol{v}_1 + \boldsymbol{v}_2 + i(\boldsymbol{w}_1 + \boldsymbol{w}_2), \\ (\lambda + i\mu)(\boldsymbol{v} + i\boldsymbol{w}) &= \lambda\boldsymbol{v} - \mu\boldsymbol{w} + i(\lambda\boldsymbol{w} + \mu\boldsymbol{v}), \end{aligned}$$

und es gelten die folgenden Eigenschaften:
(i) $\dim V_{\mathbb{C}} = \dim V$.
(ii) Es sei $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ eine Basis in $V$. Dann sind z.B. $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ oder $i\boldsymbol{b}_1, \ldots, i\boldsymbol{b}_n$ Basen in $V_{\mathbb{C}}$.
(iii) Wenn $V$ ein reeller (bzw. komplexer) Vektorraum ist, so gilt $\dim(V_{\mathbb{C}})_{\mathbb{R}} = 2 \dim V$ (bzw. $\dim(V_{\mathbb{R}})_{\mathbb{C}} = 2 \dim V$). Folglich sind die Vektorräume $V$ und $(V_{\mathbb{C}})_{\mathbb{R}}$ (bzw. $(V_{\mathbb{R}})_{\mathbb{C}}$) wesentlich verschieden.

**Bemerkung zur Bezeichnungsweise** Üblicherweise wird, soweit möglich, die Theorie für reelle und komplexe Vektorräume gleichzeitig behandelt. Man spricht dann von Vektorräumen über $\mathbb{K}$, wobei sowohl $\mathbb{K} = \mathbb{R}$ als auch $\mathbb{K} = \mathbb{C}$ sein kann. Dabei ist die Bezeichnung $\mathbb{K}$ deshalb gewählt, weil das Gemeinsame von $\mathbb{R}$ und $\mathbb{C}$, das benutzt wird, ihre sogenannten Körper-Eigenschaften sind.

# 2 Skalarprodukte und Orthogonalität

Im folgenden wird für $z \in \mathbb{C}$ mit $\overline{z}$ wird die konjugiert-komplexe Zahl zu $z$, mit $\operatorname{Re} z$ der Realteil, mit $\operatorname{Im} z$ der Imaginärteil und mit $|z|$ der Betrag von $z$ bezeichnet, d.h. $\overline{\lambda + i\mu} = \lambda - i\mu$, $\operatorname{Re}(\lambda + i\mu) = \lambda$, $\operatorname{Im}(\lambda + i\mu) = \mu$ und $|\lambda + i\mu| = \sqrt{\lambda^2 + \mu^2}$ für $\lambda, \mu \in \mathbb{R}$.
Analog werden (in Kapitel 8) die Bezeichnungen $\overline{\boldsymbol{v}}$, $\operatorname{Re} \boldsymbol{v}$ und $\operatorname{Im} \boldsymbol{v}$ für $\boldsymbol{v} \in \mathbb{C}^n$ benutzt.

Ein Skalarprodukt in einem Vektorraum $V$ ist eine Abbildung $(\boldsymbol{u}, \boldsymbol{v}) \in V^2 \mapsto \boldsymbol{u} \cdot \boldsymbol{v} \in \mathbb{K}$, so daß für alle $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ und $\lambda, \mu \in \mathbb{K}$ gilt

$$\begin{aligned} (\lambda \boldsymbol{u} + \mu \boldsymbol{v}) \cdot \boldsymbol{w} &= \lambda(\boldsymbol{u} \cdot \boldsymbol{w}) + \mu(\boldsymbol{v} \cdot \boldsymbol{w}), \\ \boldsymbol{u} \cdot \boldsymbol{v} &= \overline{\boldsymbol{v} \cdot \boldsymbol{u}}, \\ \boldsymbol{u} \cdot \boldsymbol{u} &> 0, \text{ falls } \boldsymbol{u} \neq \boldsymbol{0}. \end{aligned}$$

Ein Vektorraum über $\mathbb{R}$ (bzw. $\mathbb{C}$) mit Skalarprodukt heißt Euklidischer (bzw. unitärer) Vektorraum. Die nichtnegative Zahl

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

heißt Norm oder Länge des Vektors $\boldsymbol{v}$. Zwei Vektoren $\boldsymbol{u}, \boldsymbol{v} \in V$ heißen orthogonal, wenn $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ gilt.

**Elementare Folgerungen aus den Axiomen** Für beliebige $\boldsymbol{u} \in V$ und $\lambda \in \mathbb{K}$ ist $\boldsymbol{u} \cdot \boldsymbol{0} = \boldsymbol{0} \cdot \boldsymbol{u} = 0$ und $\|\lambda \boldsymbol{u}\| = |\lambda| \|\boldsymbol{u}\|$. Ferner gilt $\|\boldsymbol{u}\| = 0$ genau dann, wenn $\boldsymbol{u} = \boldsymbol{0}$.

**Beispiel: Das Euklidische Skalarprodukt** Das sogenannte Euklidische Skalarprodukt in $\mathbb{R}^n$ ist definiert durch

$$(u_1, \ldots, u_n) \cdot (v_1, \ldots, v_n) = \sum_{j=1}^{n} u_j v_j.$$

Die diesem Skalarprodukt entsprechenden Begriffe "Länge" und "orthogonal" sind die der Elementargeometrie, zum Beispiel ist

$$\|(v_1, v_2)\| = \sqrt{v_1^2 + v_2^2}$$

die Länge des Vektors $(v_1, v_2) \in \mathbb{R}^2$ (Satz des Pythagoras). Wenn im folgenden ein Skalarprodukt bzw. eine Norm in $\mathbb{R}^n$ benutzt wird, so ist stets das Euklidische Skalarprodukt bzw. die diesem Skalarprodukt entsprechende Norm gemeint. Es existieren aber auch andere Skalarprodukte in $\mathbb{R}^n$.

**Beispiel: Das Hermitesche Skalarprodukt** Das sogenannte Hermitesche Skalarprodukt in $\mathbb{C}^n$ ist definiert durch

$$(u_1, \ldots, u_n) \cdot (v_1, \ldots, v_n) = \sum_{j=1}^{n} u_j \overline{v_j}.$$

**Beispiel: Skalarprodukte im reellifizierten bzw. komplexifizierten Vektorraum** Wenn $\mathbb{K} = \mathbb{C}$ bzw. $\mathbb{K} = \mathbb{R}$ ist, so kann man in $V_{\mathbb{R}}$ bzw. $V_{\mathbb{C}}$ ein Skalarprodukt einführen durch

$$(\boldsymbol{v}, \boldsymbol{w}) \in V_{\mathbb{R}}^2 \mapsto \operatorname{Re}(\boldsymbol{v} \cdot \boldsymbol{w}) \in \mathbb{R}$$

bzw. durch

$$(\boldsymbol{v}_1 + i\boldsymbol{w}_1, \boldsymbol{v}_2 + i\boldsymbol{w}_2) \in V_{\mathbb{C}}^2 \mapsto \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + \boldsymbol{w}_1 \cdot \boldsymbol{w}_2 + i\,(\boldsymbol{w}_1 \cdot \boldsymbol{v}_2 - \boldsymbol{v}_1 \cdot \boldsymbol{w}_2) \in \mathbb{C}.$$

**Beispiel: Ein Skalarprodukt in einem Funktionenraum** Im Vektorraum der stetigen Funktionen $f : [-1, 1] \to \mathbb{R}$ (mit den "punktweise" definierten algebraischen Operationen) ist

$$f \cdot g = \int_{-1}^{1} f(x)g(x)dx \tag{2.1}$$

ein Skalarprodukt.

**Binomische Formel, Parallelogramm-Identität und Satz des Pythagoras** Für beliebige $\boldsymbol{u}$,$\boldsymbol{v} \in V$ gilt

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + 2\operatorname{Re}(\boldsymbol{u} \cdot \boldsymbol{v}) + \|\boldsymbol{v}\|^2$$

und folglich

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 + \|\boldsymbol{u} - \boldsymbol{v}\|^2 = 2(\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2)$$

und

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u} - \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2, \text{ genau dann, wenn } \operatorname{Re}(\boldsymbol{u} \cdot \boldsymbol{v}) = 0.$$

**Cauchy-Schwarz-Ungleichung und Dreiecksungleichung** Für beliebige $\boldsymbol{u}$,$\boldsymbol{v} \in V$ gilt

$$|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\|\,\|\boldsymbol{v}\| \text{ und } \|\boldsymbol{u}\| - \|\boldsymbol{v}\| \leq \|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\| \text{ für alle } \boldsymbol{u}, \boldsymbol{v} \in V.$$

Im Fall des Euklidischen Skalarprodukts in $\mathbb{R}^n$ ergibt sich

$$\Big(\sum_{j=1}^{n} u_j v_j\Big)^2 \leq \sum_{j=1}^{n} u_j^2 \sum_{j=1}^{n} v_j^2$$

und

$$\Big(\sum_{j=1}^{n} u_j^2\Big)^{1/2} - \Big(\sum_{j=1}^{n} v_j^2\Big)^{1/2} \leq \Big(\sum_{j=1}^{n} (u_j + v_j)^2\Big)^{1/2} \leq \Big(\sum_{j=1}^{n} u_j^2\Big)^{1/2} + \Big(\sum_{j=1}^{n} v_j^2\Big)^{1/2}.$$

**Winkel** Es sei $\mathbb{K} = \mathbb{R}$. Für gegebene $\boldsymbol{u}$,$\boldsymbol{v} \in V$ existiert dann (wegen der Cauchy-Schwarz-Ungleichung) genau ein $\varphi \in [0, \pi]$ mit

$$\cos\varphi = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|u\|\,\|v\|},$$

der sogenannte Winkel zwischen den Vektoren $\boldsymbol{u}$ und $\boldsymbol{v}$. Dabei gilt

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 + 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| \cos\varphi.$$

**Orthonormalbasen** Es sei $\dim V = n$, und für gegebene Vektoren $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n \in V$ gelte $\|\boldsymbol{e}_j\| = 1$ und $\boldsymbol{e}_j \cdot \boldsymbol{e}_k = 0$ für alle $j$ und alle $k \neq j$. Dann ist $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ eine Basis in $V$. Solche Basen werden Orthonormalbasen genannt.

**Beispiel** Die kanonische Basis in $\mathbb{R}^n$ ist eine Orthonormalbasis bzgl. dem Euklidischen Skalarprodukt.

**Beispiel** Die Polynome $p_1(x) = \sqrt{\frac{1}{2}}$, $p_2(x) = \sqrt{\frac{3}{2}}x$ und $p_3(x) = \sqrt{\frac{5}{8}}(3x^2 - 1)$ (die ersten drei sogenannten Legendre-Polynome) bilden eine Orthonormalbasis in $\mathbb{R}_2[x]$ bzgl. dem Skalarprodukt (2.1).

**Hilbert-Schmidt-Orthonormalisierungsverfahren** Es sei $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ eine Basis in $V$. Dann existiert genau (bis auf den skalare Faktoren vom Betrag Eins) eine Orthonormalbasis $\boldsymbol{e}_1, \dots, \boldsymbol{e}_j$ in $V$, so daß gilt

$$\operatorname{span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_j\} = \operatorname{span}\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_j\} \text{ für } j = 1, \dots, n.$$

Dabei ist

$$\boldsymbol{e}_1 = c_1 \frac{\boldsymbol{b}_1}{\|\boldsymbol{e}_1\|} \text{ und } \boldsymbol{e}_j = c_j \frac{\boldsymbol{b}_j - \sum_{k=1}^{j-1} (\boldsymbol{b}_j \cdot \boldsymbol{e}_k)\, \boldsymbol{e}_k}{\|\boldsymbol{b}_j - \sum_{k=1}^{j-1} (\boldsymbol{b}_j \cdot \boldsymbol{e}_k)\, \boldsymbol{e}_k\|} \text{ für } j = 2, \dots, n,$$

wobei $c_1, \dots, c_n$ skalare Faktoren vom Betrag Eins sind.

**Koordinaten bzgl. Orthonormalbasen** Wenn $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ eine Orthonormalbasis in $V$ ist, so gilt für jedes $\boldsymbol{v} \in V$

$$\boldsymbol{v} = \sum_{j=1}^{n} (\boldsymbol{v} \cdot \boldsymbol{e}_j)\, \boldsymbol{e}_j \text{ und } \|\boldsymbol{v}\|^2 = \sum_{j=1}^{n} (\boldsymbol{v} \cdot \boldsymbol{e}_j)^2.$$

Mit anderen Worten: Die Zahlen $\boldsymbol{v}\cdot\boldsymbol{e}_1, \dots, \boldsymbol{v}\cdot\boldsymbol{e}_1$ sind die Koordinaten von $\boldsymbol{v}$ bzgl. dieser Orthonormalbasis, und die Länge von $\boldsymbol{v}$ ist gleich der Wurzel aus der Quadratsumme dieser Koordinaten.

**Orthogonales Komplement, orthogonale Summe und orthogonaler Projektor** Es sei $U$ ein Unterraum in $V$, und $V$ sei endlich-dimensional.
(i) Die Menge $U^\perp = \{\boldsymbol{v} \in V : \boldsymbol{u} \cdot \boldsymbol{v} = 0 \text{ für alle } \boldsymbol{u} \in U\}$ ist ein Unterraum in $V$, das sogenannte orthogonale Komplement zu $U$, und es gilt $V = U \oplus U^\perp$. Eine solche direkte Summe, wobei der eine Summand das orthogonale Komplement des anderen Summanden ist, nennt man auch orthogonale Summe.
(ii) Für jedes $\boldsymbol{v} \in V$ existiert genau ein $\boldsymbol{u} \in U$, so daß gilt

$$\|\boldsymbol{v} - \boldsymbol{u}\| = \min\{\|\boldsymbol{v} - \boldsymbol{w}\| : \boldsymbol{w} \in U\}.$$

Dabei ist $\boldsymbol{v} - \boldsymbol{u} \in U^\perp$, und die Abbildung $\boldsymbol{v} \mapsto \boldsymbol{u}$ ist der Projektor auf $U$ entlang $U^\perp$ (der sogenannte orthogonale Projektor auf $U$). Die Vektoren $\boldsymbol{u}$ bzw. $\boldsymbol{v} - \boldsymbol{u}$ nennt man auch orthogonale Projektion bzw. senkrechtes Lot von $\boldsymbol{v}$ auf $U$, und $\|\boldsymbol{v} - \boldsymbol{u}\|$ heißt Abstand von $\boldsymbol{v}$ zu $U$, und es gilt

$$\|\boldsymbol{v} - \boldsymbol{u}\| = \max\left\{\frac{|\langle \boldsymbol{v}, \boldsymbol{w}\rangle|}{\|\boldsymbol{w}\|} : \boldsymbol{w} \in U^\perp\right\}.$$

Ferner gilt

$$\boldsymbol{u} = \sum_{j=1}^{m} (\boldsymbol{v} \cdot \boldsymbol{e}_j)\, \boldsymbol{e}_j,$$

falls $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m$ eine Orthonormalbasis in $U$ ist.
(iii) Wenn $U_1$ und $U_2$ zwei Unterräume in $V$ sind, so gilt $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$.

**Beispiel** Es sei $U$ eine Ebene in $\mathbb{R}^3$, die den Nullpunkt enthält, also ein zweidimensionaler Unterraum in $\mathbb{R}^3$. Dann ist $U^\perp$ die zu $U$ senkrechte Gerade durch den Nullpunkt, und die kürzeste Verbindung zwischen einem Punkt $\boldsymbol{x} \in \mathbb{R}^3$ und $U$ ist das Lot von $\boldsymbol{x}$ auf $U$, d.h. die Strecke von $\boldsymbol{x}$ zum Schnittpunkt von $U$ und $\boldsymbol{x}+U^\perp$. Wenn dabei $U$ die Menge aller Lösungen $(u_1, u_2, u_3) \in \mathbb{R}^3$ der Gleichung $\lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3 = 0$ (mit $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 > 0$) ist, so gilt $U = (\operatorname{span}\{(\lambda_1, \lambda_2, \lambda_3)\})^\perp$, und folglich ist der Abstand von $(x_1, x_2, x_3)$ zu $U$ gleich

$$\frac{|\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3|}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}.$$

**Beispiel** Es sei $U$ eine Gerade in $\mathbb{R}^3$, die den Nullpunkt enthält, also ein eindimensionaler Unterraum in $\mathbb{R}^3$. Dann ist $U^\perp$ die zu $U$ senkrechte Ebene durch den Nullpunkt, und die kürzeste Verbindung zwischen einem Punkt $\boldsymbol{x} \in \mathbb{R}^3$ und $U$ ist das Lot von $\boldsymbol{x}$ auf $U$, d.h. die Strecke von $\boldsymbol{x}$ zum Schnittpunkt von $U$ und $\boldsymbol{x}+U^\perp$. Wenn dabei $U = \operatorname{span}\{\boldsymbol{u}_0\}$ ist, d.h. wenn die Gerade $U$ durch den Vektor $\boldsymbol{u}_0 \neq 0$ aufgespannt wird, so ist der Abstand von $\boldsymbol{x}$ zu $U$ gleich

$$\left\| \boldsymbol{x} - \left( \boldsymbol{x} \cdot \frac{\boldsymbol{u}_0}{\|\boldsymbol{u}_0\|} \right) \frac{\boldsymbol{u}_0}{\|\boldsymbol{u}_0\|} \right\| = \sqrt{\|\boldsymbol{x}\|^2 - \left( \frac{\boldsymbol{x} \cdot \boldsymbol{u}_0}{\|\boldsymbol{u}_0\|} \right)^2}.$$

**Einheitsnormalen zu Hyperebenen** Es sei $U$ ein Unterraum in $V$ mit $\dim U = \dim V - 1$. Dann heißt $U$ Hyperebene in $V$, und es gilt $\dim U^\perp = 1$, d.h. $U^\perp$ ist eine Gerade durch den Nullpunkt. Folglich existiert genau (bis auf skalare Faktoren vom Betrag Eins) ein Vektor in $U^\perp$ mit der Länge Eins. Diese Vektoren heißen Einheitsnormalen zu $U$. Mit anderen Worten: Die Einheitsnormalen zu einer Hyperebene $U$ sind die Vektoren $\boldsymbol{v} \in V$ mit $\|\boldsymbol{v}\| = 1$ und $\boldsymbol{v} \cdot \boldsymbol{u} = 0$ für alle $\boldsymbol{u} \in U$.

# 3 Lineare Abbildungen

Es seien $V$ und $W$ zwei Vektorräume über $\mathbb{K}$ (bei $V$ und $W$ gleiches $\mathbb{K}$). Eine Abbildung $\boldsymbol{F}: V \to W$ heißt linear, wenn

$$\boldsymbol{F}(\lambda \boldsymbol{u} + \mu \boldsymbol{v}) = \lambda \boldsymbol{F}(\boldsymbol{u}) + \mu \boldsymbol{F}(\boldsymbol{v}) \text{ für alle } \lambda, \mu \in \mathbb{K} \text{ und } \boldsymbol{u}, \boldsymbol{v} \in V$$

gilt. Die Menge aller linearen Abbildungen von $V$ nach $W$ wird mit $\mathcal{L}(V; W)$ bezeichnet. Anstelle von $\mathcal{L}(V; V)$ schreibt man auch $\mathcal{L}(V)$. Mit $\boldsymbol{I} \in \mathcal{L}(V)$ bezeichnet man die identische Abbildung in $V$.

**Beispiel** $F : \mathbb{R} \to \mathbb{R}^n$ ist linear genau dann, wenn ein $\boldsymbol{v}_0 \in \mathbb{R}^n$ existiert, so daß für alle $\lambda \in \mathbb{R}$ gilt $F(\lambda) = \lambda \boldsymbol{v}_0$.

**Beispiel** $F : \mathbb{R}^n \to \mathbb{R}$ ist linear genau dann, wenn ein $\boldsymbol{v}_0 \in \mathbb{R}^n$ existiert, so daß für alle $\boldsymbol{v} \in \mathbb{R}^n$ gilt $F(\boldsymbol{v}) = \boldsymbol{v} \cdot \boldsymbol{v}_0$.

**Beispiel** $F:\mathbb{C}\to\mathbb{C}$ ist linear genau dann, wenn ein $c\in\mathbb{C}$ existiert, so daß für alle $v\in\mathbb{C}$ gilt $F(v)=cv$. Wenn $c=re^{i\varphi}$ ist mit $r,\varphi\in\mathbb{R}$, so ist also $F$ die Superposition der "Drehung" $v\mapsto e^{i\varphi}v$ und der "Streckung" $v\mapsto rv$.

**Beispiel** $\boldsymbol{F}\colon\mathbb{R}^2\to\mathbb{R}^2$ ist linear genau dann, wenn relle Zahlen $a_{11},a_{12},a_{21},a_{22}$ existieren, so daß für alle $(v_1,v_2)\in\mathbb{R}^2$ gilt $\boldsymbol{F}(v_1,v_2)=(a_{11}v_1+v_{12}v_2,a_{21}v_1+a_{22}v_2)$.

**Komplexifizierung linearer Abbildungen** Wenn $\boldsymbol{F}\colon V\to W$ linear ist und $\mathbb{K}=\mathbb{R}$, dann ist

$$\boldsymbol{v}+i\boldsymbol{w}\mapsto\boldsymbol{F}(\boldsymbol{v})+i\boldsymbol{F}(\boldsymbol{w})$$

eine lineare Abbildung von $V_{\mathbb{C}}$ nach $W_{\mathbb{C}}$ und wird Komplexifizierung von $F$ genannt.

**Beispiel** Es seien $g\in C(\mathbb{R})$ und $x_0\in\mathbb{R}$ fixiert. Dann sind die Abbildungen $f\in C(\mathbb{R})\mapsto fg\in C(\mathbb{R})$ mit (fg)(x)=f(x)g(x), $f\in C(\mathbb{R})\mapsto f(x_0)\in\mathbb{R}$ und $f\in C(\mathbb{R})\mapsto F\in C(\mathbb{R})$ mit $F(x)=\int_{x_0}^{x}f(y)dy$ linear. Ferner ist die Abbildung $f\in C^1(\mathbb{R})\mapsto f'\in C(\mathbb{R})$ (die Ableitungsabbildung) linear.

**Eigenschaften** (i) $\mathcal{L}(V;W)$ ist ein Vektorraum über $\mathbb{K}$ mit den Operationen

$$\left.\begin{array}{rcl}(\boldsymbol{F}+\boldsymbol{G})(\boldsymbol{v}) & = & \boldsymbol{F}(\boldsymbol{v})+\boldsymbol{G}(\boldsymbol{v}),\\ (\lambda\boldsymbol{F})(\boldsymbol{v}) & = & \lambda\boldsymbol{F}(\boldsymbol{v}).\end{array}\right\}\ \text{für alle } \boldsymbol{v}\in V.$$

Das Nullelement in $\mathcal{L}(V;W)$ ist die Abbildung, die alle $\boldsymbol{v}\in V$ auf den Nullvektor in $W$ abbildet. Wenn $V$ und $W$ endlich-dimensional sind, so gilt

$$\dim\mathcal{L}(V;W)=\dim V\dim W. \tag{3.1}$$

(ii) Wenn $\boldsymbol{F}:U\to V$ und $\boldsymbol{G}:V\to W$ zwei lineare Abbildungen sind, so ist auch ihre Superposition $\boldsymbol{G}\circ\boldsymbol{F}:U\to W$ linear.
(iii) Wenn $\boldsymbol{F}:V\to W$ linear und bijektiv ist, so ist auch seine Umkehrung $\boldsymbol{F}^{-1}:W\to V$ linear.
(iv) Für fixiertes $\boldsymbol{G}\in\mathcal{L}(V)$ sind die Abbildungen $\boldsymbol{F}\in\mathcal{L}(V)\mapsto\boldsymbol{F}\circ\boldsymbol{G}\in\mathcal{L}(V)$ und $\boldsymbol{F}\in\mathcal{L}(V)\mapsto\boldsymbol{G}\circ\boldsymbol{F}\in\mathcal{L}(V)$ linear. Dagegen sind die Abbildungen $\boldsymbol{F}\in\mathcal{L}(V)\mapsto\boldsymbol{F}\circ\boldsymbol{F}\in\mathcal{L}(V)$ oder $(\boldsymbol{F},\boldsymbol{G})\in\mathcal{L}(V)^2\mapsto\boldsymbol{F}\circ\boldsymbol{G}\in\mathcal{L}(V)$ nicht linear.

**Kern und Bild von linearen Abbildungen** Es sei $\boldsymbol{F}\in\mathcal{L}(V;W)$. Dann heißen die Unterräume

$$\ker\boldsymbol{F}=\{\boldsymbol{v}\in V:\ \boldsymbol{F}(\boldsymbol{v})=\boldsymbol{0}\}\ \text{bzw.}\ \operatorname{im}\boldsymbol{F}=\{\boldsymbol{F}(\boldsymbol{v})\in W:\ \boldsymbol{v}\in V\}$$

Kern bzw. Bild von $\boldsymbol{F}$.

**Dimensionsformel für lineare Abbildungen und Fredholmsche Alternative** Es sei $\boldsymbol{F}\in\mathcal{L}(V;W)$ und $\dim V<\infty$. Dann ist

$$\dim\ker\boldsymbol{F}+\dim\operatorname{im}\boldsymbol{F}=\dim V.$$

Wenn $\dim V=\dim W$ ist, gilt insbesondere, daß $\boldsymbol{F}$ genau dann injektiv (d.h. $\ker\boldsymbol{F}=\{0\}$) ist, wenn es surjektiv (d.h. $\operatorname{im}\boldsymbol{F}=W$) ist. Mit anderen Worten, wenn $\dim V=\dim W$ ist, so gilt die folgende Alternative: Entweder die sogenannte homogene Gleichung $\boldsymbol{F}(\boldsymbol{v})=\boldsymbol{0}$ besitzt eine Lösung $\boldsymbol{v}\neq\boldsymbol{0}$ oder die sogenannte inhomogene Gleichung $\boldsymbol{F}(\boldsymbol{v})=\boldsymbol{w}$ besitzt für jedes $\boldsymbol{w}\in W$ genau eine Lösung $\boldsymbol{v}\in V$.

**Lösungsmengen inhomogener linearer Gleichungen** Wenn $\boldsymbol{F} \in \mathcal{L}(V;W)$ und $\boldsymbol{w} \in W$ fixiert sind, so ist die Menge aller Lösungen $\boldsymbol{v} \in V$ der Gleichung $\boldsymbol{F}(\boldsymbol{v}) = \boldsymbol{w}$ ein affiner Unterraum. Es gilt nämlich

$$\{\boldsymbol{v} :\ \boldsymbol{F}(\boldsymbol{v}) = \boldsymbol{w}\} = \boldsymbol{v_0} + \ker \boldsymbol{F} \ \text{ genau dann, wenn } \ \boldsymbol{F}(\boldsymbol{v_0}) = \boldsymbol{w}. \tag{3.2}$$

Diesen Sachverhalt formuliert man bisweilen auch folgendermaßen: Die allgemeine Lösung der inhomogenen Gleichung, nämlich $\{\boldsymbol{v} :\ \boldsymbol{F}(\boldsymbol{v}) = \boldsymbol{w}\}$, ist gleich der Summe aus einer partikulären Lösung der inhomogenen Gleichung, nämlich $\boldsymbol{v_0}$, und der allgemeinen Lösung der homogenen Gleichung, nämlich $\ker \boldsymbol{F}$.

**Kern und Bild von Superpositionen** Es gilt

$$\dim \ker \boldsymbol{G} \leq \dim \ker \boldsymbol{F} \circ \boldsymbol{G} \leq \dim \ker \boldsymbol{F} + \dim \ker \boldsymbol{G}$$

und

$$\dim \operatorname{im} \boldsymbol{F} \circ \boldsymbol{G} \leq \min\{\dim \operatorname{im} \boldsymbol{F},\ \dim \operatorname{im} \boldsymbol{G}\}.$$

**Projektoren** Wenn $V = U_1 \oplus U_2$ ist und $\boldsymbol{P}$ der Projektor auf $U_1$ entlang $U_2$, so ist $\boldsymbol{P}$ linear, es gilt $\boldsymbol{P} \circ \boldsymbol{P} = \boldsymbol{P}$, und $\boldsymbol{I} - \boldsymbol{P}$ ist der Projektor auf $U_2$ entlang $U_1$. Ferner gilt: Wenn $\boldsymbol{P} \in \mathcal{L}(V)$ die Gleichung $\boldsymbol{P} \circ \boldsymbol{P} = \boldsymbol{P}$ erfüllt, so ist $\boldsymbol{P}$ der Projektor auf $\operatorname{im} \boldsymbol{P}$ entlang $\ker \boldsymbol{P}$.

**Ein Kriterium für Unendlich-Dimensionalität** Ein Vektorraum $V$ ist genau dann unendlich-dimensional, wenn ein Unterraum $U$ in $V$ und ein $\boldsymbol{F} \in \mathcal{L}(U;V)$ existieren, so daß $U \neq V$ ist und daß $\boldsymbol{F}$ trotzdem bijektiv ist.

# 4 Matrizen

Eine $m \times n$-Matrix über $\mathbb{K}$ ist eine Anordnung

$$\boldsymbol{A} \ = \ [a_{ij}] \ = \ \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

mit $a_{ij} \in \mathbb{K}$. Die Zahlen $a_{ij}$ heißen Koeffizienten von $\boldsymbol{A}$. Die $m$-Tupel bzw. $n$-Tupel

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad \text{bzw.} \quad \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{bmatrix}$$

heißen Spalten bzw. Zeilen von $A$. Wenn $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ die Spalten der Matrix $\boldsymbol{A}$ sind, so schreibt man auch

$$\boldsymbol{A} = [\boldsymbol{a}_1\, \boldsymbol{a}_2\, \dots \boldsymbol{a}_n].$$

Die Menge aller $m \times n$-Matrizen über $\mathbb{K}$ wird mit $\mathbb{M}(m \times n; \mathbb{K})$ bezeichnet. Elemente von $\mathbb{M}(m \times 1; \mathbb{K})$ bzw. $\mathbb{M}(1 \times n; \mathbb{K})$ heißen in Spalten bzw. Zeilen ausgeartete Matrizen und werden mit den entsprechenden Elementen aus $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ identifiziert. Matrizen mit gleicher Spalten- und Zeilenzahl heißen quadratisch. Die quadratische Matrix, auf deren Diagonale Einsen stehen und deren Nichtdiagonalkoeffizienten alle Null sind, heißt Einheitsmatrix und wird mit $\boldsymbol{I}$ bezeichnet. Häufig schreibt man auch $\boldsymbol{I} = [\delta_{ij}]$, dabei ist

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst} \end{cases} \tag{4.1}$$

das sogenannte Kronecker-Symbol.

**Addition und Multiplikation mit Zahlen** $\mathbb{M}(m \times n; \mathbb{K})$ ist ein Vektorraum über $\mathbb{K}$ mit den Operationen

$$\begin{aligned} [a_{ij}] + [b_{ij}] &= [a_{ij} + b_{ij}], \\ \lambda[a_{ij}] &= [\lambda a_{ij}]. \end{aligned}$$

Das Nullelement in $\mathbb{M}(m \times n; \mathbb{K})$ ist die Matrix, deren Koeffizienten alle Null sind, und sie wird auch mit dem Symbol $\mathbf{0}$ bezeichnet. Ferner gilt

$$\dim \mathbb{M}(m \times n; \mathbb{K}) = mn.$$

**Multiplikation von Matrizen bzw. von Matrizen und Vektoren** Es seien $\boldsymbol{A} = [a_{ij}] \in \mathbb{M}(l \times m; \mathbb{K})$, $\boldsymbol{B} = [b_{jk}] \in \mathbb{M}(m \times n; \mathbb{K})$ und $\boldsymbol{C} = [c_{ik}] \in \mathbb{M}(l \times n; \mathbb{K})$ Matrizen, und es gelte

$$c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk} \text{ für alle } i = 1. \cdots, l \text{ und } k = 1, \cdots, n.$$

Dann nennt man die Matrix $\boldsymbol{C}$ Produkt der Matrizen $\boldsymbol{A}$ und $\boldsymbol{B}$, und man schreibt $\boldsymbol{C} = \boldsymbol{AB}$. Wenn insbesondere der zweite Faktor in eine Spalte $\boldsymbol{b} \in \mathbb{K}^m$ ausgeartet ist, so ist auch das Produkt $\boldsymbol{c} \in \mathbb{K}^l$ in eine Spalte ausgeartet, und es gilt

$$c_i = \sum_{j=1}^{m} a_{ij} b_j \quad \text{mit} \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \text{und} \quad \boldsymbol{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_l \end{bmatrix}.$$

In diesem Fall nennt man den Vektor $\boldsymbol{c}$ Produkt der Matrix $\boldsymbol{A}$ mit dem Vektor $\boldsymbol{b}$, und man schreibt $\boldsymbol{c} = \boldsymbol{Ab}$.

**Inverse Matrix** Wenn zu einem gegebenem $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{K})$ ein $\boldsymbol{B} \in \mathbb{M}(n \times n; \mathbb{K})$ existiert, so daß gilt $\boldsymbol{AB} = \boldsymbol{I}$ oder $\boldsymbol{BA} = \boldsymbol{I}$, so gilt auch $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$. Die Matrix $\boldsymbol{A}$ heißt dann invertierbar, $\boldsymbol{B}$ heißt inverse Matrix zu $\boldsymbol{A}$, und man schreibt $\boldsymbol{A}^{-1}$ anstelle von $\boldsymbol{B}$.

**Transponierte und konjugiert-transponierte Matrix** (i) Für $\boldsymbol{A} \in \mathbb{M}(m \times n; \mathbb{K})$ bezeichnet man mit $\boldsymbol{A}^T \in \mathbb{M}(n \times m; \mathbb{K})$ ihre transponierte Matrix, d.h. die Matrix, deren $i$te Zeile (bzw. $j$te

Spalte) gleich der $i$ten Spalte (bzw. $j$ten Zeile) von $\boldsymbol{A}$ ist, also

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nm} \end{bmatrix}.$$

(ii) Für $\boldsymbol{A} \in \mathbb{M}(m \times n; \mathbb{C})$ ist $\boldsymbol{A}^* \in \mathbb{M}(n \times m; \mathbb{C})$ ihre konjugiert-transponierte Matrix, d.h. die Matrix deren Koeffizienten die komplex-konjugierten Zahlen der entsprechenden Koeffizienten von $\boldsymbol{A}^T$ sind, also

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \dots & \overline{a_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \dots & \overline{a_{nm}} \end{bmatrix}.$$

Insbesondere ist $\boldsymbol{A}^T = \boldsymbol{A}^*$, wenn alle Koeffizienten von $\boldsymbol{A}$ reell sind.
(iii) Für beliebige $\boldsymbol{A} \in \mathbb{M}(m \times n; \mathbb{K})$, $\boldsymbol{u} \in \mathbb{K}^n$ und $\boldsymbol{v} \in \mathbb{K}^m$ gilt

$$\boldsymbol{A}\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{A}^*\boldsymbol{v}.$$

**Rechenregeln** Es gilt

$$(\lambda\boldsymbol{A} + \mu\boldsymbol{B})\boldsymbol{C} = \lambda(\boldsymbol{A}\boldsymbol{C}) + \mu(\boldsymbol{B}\boldsymbol{C}), \quad \boldsymbol{C}(\lambda\boldsymbol{A} + \mu\boldsymbol{B}) = \lambda(\boldsymbol{C}\boldsymbol{A}) + \mu(\boldsymbol{C}\boldsymbol{B}),$$
$$\boldsymbol{A}(\boldsymbol{B}\boldsymbol{C}) = (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C}, \quad \boldsymbol{0}\boldsymbol{A} = \boldsymbol{B}\boldsymbol{0} = \boldsymbol{0}, \quad \boldsymbol{I}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I} = \boldsymbol{A},$$
$$(\lambda\boldsymbol{A})^* = \overline{\lambda}\boldsymbol{A}^*, \quad (\lambda\boldsymbol{A})^{-1} = 1/\lambda\, \boldsymbol{A}^{-1}, \quad (\boldsymbol{A}^{-1})^* = (\boldsymbol{A}^*)^{-1}, \quad (\boldsymbol{A}\boldsymbol{B})^* = \boldsymbol{B}^*\boldsymbol{A}^*, \quad (\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}.$$

Dabei sind $\boldsymbol{A}$, $\boldsymbol{B}$ und $\boldsymbol{C}$ bzw. $\lambda$ und $\mu$ beliebige Matrizen bzw. Zahlen derart, daß die Operationen ausgeführt werden können. Ferner gilt

$$\boldsymbol{A}[\boldsymbol{b}_1\ \boldsymbol{b}_2\ \dots\ \boldsymbol{b}_n] = [\boldsymbol{A}\boldsymbol{b}_2\ \dots\ \boldsymbol{A}\boldsymbol{b}_n],$$

d.h. die Spalten eines Produktes $\boldsymbol{A}\boldsymbol{B}$ (mit $\boldsymbol{A} \in \mathbb{M}(l \times m; \mathbb{K})$ und $\boldsymbol{B} \in \mathbb{M}(m \times n; \mathbb{K})$) sind gleich den Produkten von $\boldsymbol{A}$ mit den Spalten von $\boldsymbol{B}$.

**Beispiel** Die Rechnung

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

zeigt, daß in Matrizenprodukten im allgemeinen nicht die Faktoren vertauscht werden dürfen und daß, wenn ein Matrizenprodukt Null ist, im allgemeinen keiner der Faktoren Null sein muß.

**Beispiel** Wegen

$$\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

gilt: Wenn man eine Matrix auf die kanonische Basis anwendet, so erhält man ihre Spalten.

**Ein Analogon zur geometrischen Reihe** Wenn für $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{K})$ gilt $\boldsymbol{A}^p = \boldsymbol{0}$, so ist $\boldsymbol{I} - \boldsymbol{A}$ invertierbar, und es gilt $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2 + \cdots + \boldsymbol{A}^{p-1}$.

**Basiswechsel-Matrix** (i) Es seien $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ und $\boldsymbol{b}'_1, \cdots, \boldsymbol{b}'_n$ zwei Basen in einem Vektorraum $V$ und $\boldsymbol{C} = [c_{ij}] \in \mathbb{M}(n \times n; \mathbb{K})$ eine Matrix. Dann ist die Bedingung

$$\boldsymbol{b}_j = \sum_{i=1}^{n} c_{ij} \boldsymbol{b}'_i \text{ für alle } j \tag{4.2}$$

äquivalent zu der Bedingung

$$\sum_{j=1}^{n} u_j \boldsymbol{b}_j = \sum_{j=1}^{n} u'_j \boldsymbol{b}'_j \text{ genau dann, wenn } u'_i = \sum_{j=1}^{n} c_{ij} u_j \text{ für alle } i. \tag{4.3}$$

Die Matrix $\boldsymbol{C}$ heißt dann Matrix des Basiswechsels von $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n$ zu $\boldsymbol{b}'_1, \cdots, \boldsymbol{b}'_n$. Die Gleichung (4.2) bedeutet, daß die $j$te Spalte von $\boldsymbol{C}$ gleich den Koordinaten des $j$ten alten Basisvektors bzgl. der neuen Basis ist. Die Aussage (4.3) bedeutet, daß $\boldsymbol{C}$ die Koordinaten eines beliebigen Vektors bzgl. der alten Basis in die Koordinaten desselben Vektors bzgl. der neuen Basis überführt.
(ii) Die Matrix $\boldsymbol{C}$ ist invertierbar, und $\boldsymbol{C}^{-1}$ ist die Matrix des Basiswechsels von der Basis $\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_n$ zu der Basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Wenn ferner $\boldsymbol{C}'$ die Matrix des Basiswechsels von der Basis $\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_n$ zu einer Basis $\boldsymbol{b}''_1, \ldots, \boldsymbol{b}''_n$ ist, so ist $\boldsymbol{C}'\boldsymbol{C}$ die Matrix des Basiswechsels von der Basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ zu der Basis $\boldsymbol{b}''_1, \ldots, \boldsymbol{b}''_n$.
(iii) Wenn $\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_n$ eine Orthomormalbasis bzgl. eines Skalarprodukts in $V$ ist, so sind (4.2) und (4.3) äquivalent zu der Bedingung

$$c_{ij} = \boldsymbol{b}_j \cdot \boldsymbol{b}'_i \text{ für alle } i, j. \tag{4.4}$$

Wenn ferner $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ ebenfalls eine Orthomormalbasis ist, so gilt $\boldsymbol{C}^{-1} = \boldsymbol{C}^*$.

**Matrix-Darstellungen linearer Abbildungen** (i) Es seien $V$ und $W$ ein Vektorräume über $\mathbb{K}$, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ und $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ Basen in $V$ und $W$, $\boldsymbol{F} : V \to W$ eine lineare Abbildung und $\boldsymbol{A}_{\boldsymbol{F}} = [a_{ij}] \in \mathbb{M}(m \times n; \mathbb{K})$ eine Matrix. Dann ist die Bedingung

$$\boldsymbol{F}(\boldsymbol{a}_j) = \sum_{i=1}^{m} a_{ij} \boldsymbol{b}_i \text{ für alle } j \tag{4.5}$$

äquivalent zu der Bedingung

$$\boldsymbol{F}(\sum_{j=1}^{n} u_j \boldsymbol{a}_j) = \sum_{j=1}^{m} v_j \boldsymbol{b}_j \text{ genau dann, wenn } v_i = \sum_{j=1}^{n} a_{ij} u_j \text{ für alle } i. \tag{4.6}$$

Die Matrix $\boldsymbol{A}_{\boldsymbol{F}}$ heißt dann Matrixdarstellung von $\boldsymbol{F}$ bzgl. der Basen $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ und $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$. Die Gleichung (4.5) bedeutet, daß die $j$te Spalte von $\boldsymbol{A}_{\boldsymbol{F}}$ gleich den Koordinaten von $\boldsymbol{F}(\boldsymbol{a}_j)$ bzgl. der Basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ ist. Die Aussage (4.6) bedeutet, daß $\boldsymbol{A}_{\boldsymbol{F}}$ die Koordinaten eines beliebigen Vektors

$\boldsymbol{v} \in V$ bzgl. der Basis $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ in die Koordinaten des Bildvektors $\boldsymbol{F}(\boldsymbol{v})$ bzgl. dieser Basis überführt.
(ii) Es gilt für beliebige $\lambda, \mu \in \mathbb{K}$ und $\boldsymbol{F}, \boldsymbol{G} \in \mathcal{L}(V)$

$$\boldsymbol{A}_{\lambda \boldsymbol{F}+\mu \boldsymbol{G}} = \lambda \boldsymbol{A}_{\boldsymbol{F}} + \mu \boldsymbol{A}_{\boldsymbol{G}} \text{ und } \boldsymbol{A}_{\boldsymbol{F}\circ\boldsymbol{G}} = \boldsymbol{A}_{\boldsymbol{F}}\boldsymbol{A}_{\boldsymbol{G}}.$$

(iii) Wenn $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ eine Orthomormalbasis bzgl. eines Skalarprodukts in $W$ ist, so sind (4.5) und (4.6) äquivalent zu der Bedingung

$$a_{ij} = \boldsymbol{F}(\boldsymbol{a}_j) \cdot \boldsymbol{b}_i \text{ für alle } i, j. \tag{4.7}$$

(iv) Insbesondere ist eine Abbildung $\boldsymbol{F} : \mathbb{K}^n \to \mathbb{K}^m$ linear genau dann, wenn eine Matrix $\boldsymbol{A} \in \mathbb{M}(m \times n; \mathbb{K})$ existiert, so daß für alle $\boldsymbol{v} \in \mathbb{K}^n$ gilt $\boldsymbol{F}(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v}$. Die Matrix $\boldsymbol{A}$ ist dann die Matrix-Darstellung der linearen Abbildung $\boldsymbol{F}$ bezüglich der kanonischen Basen in $\mathbb{K}^n$ und $\mathbb{K}^m$. In diesem Sinne identifiziert man lineare Abbildungen aus $\mathcal{L}(\mathbb{K}^n; \mathbb{K}^m)$ und Matrizen aus $\mathbb{M}(m \times n; \mathbb{K})$. Dabei entspricht die Superposition zweier linearer Abbildungen dem Produkt der entsprechenden Matrizen.

**Matrix-Darstellungen unter Basiswechsel** Es seien $\boldsymbol{F} : V \to W$ eine lineare Abbildung, $\boldsymbol{A}$ bzw. $\boldsymbol{A}'$ die Matrix-Darstellungen von $\boldsymbol{F}$ bzgl. der Basen $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ und $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ bzw. $\boldsymbol{a}'_1, \ldots, \boldsymbol{a}'_n$ und $\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_m$ in $V$ und $W$, und $\boldsymbol{C}$ und $\boldsymbol{D}$ seien die Matrizen der Basiswechsel von $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ zu $\boldsymbol{a}'_1, \ldots, \boldsymbol{a}'_n$ und von $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m$ zu $\boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_m$. Dann gilt

$$\boldsymbol{A}' = \boldsymbol{D}\boldsymbol{A}\boldsymbol{C}^{-1}.$$

**Ähnliche Matrizen** Zwei Matrizen $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{M}(n \times n; \mathbb{K})$ heißen ähnlich, wenn eine invertierbare Matrix $\boldsymbol{C} \in \mathbb{M}(n \times n; \mathbb{K})$ existiert, so daß gilt $\boldsymbol{B} = \boldsymbol{C}\boldsymbol{A}\boldsymbol{C}^{-1}$, d.h. wenn eine Basis in $\mathbb{K}^n$ existiert, so daß $\boldsymbol{B}$ die Matrix-Darstellung von $\boldsymbol{A}$ bzgl. dieser Basis ist. Ähnlichkeit ist eine Äquivalenzrelation in $\mathbb{M}(n \times n; \mathbb{K})$.

**Beispiel** Im Vektorraum $\mathbb{R}_3[x]$ der Polynome vom Grad kleiner oder gleich Drei bilden die Polynome $1$, $x$, $x^2$ und $x^3$ eine Basis. Die Ableitungsabbildung $d/dx$ bildet $\mathbb{R}_3[x]$ linear in sich ab, und die Matrix-Darstellung dieser linearen Abbildung bzgl. der obigen Basis ist wegen (4.5),

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Folglich ist die Matrix-Darstellung von $d^2/dx^2$, $d^3/dx^3$ bzw. $d^4/dx^4$ bzgl. derselben Basis gleich

$$\boldsymbol{A}^2 = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ bzw. } \boldsymbol{A}^3 = \begin{bmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ bzw. } \boldsymbol{A}^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

**Beispiel** Es sei $\boldsymbol{u} = (u_1, u_2, u_3) \neq \boldsymbol{0}$ ein fixierter Vektor in $\mathbb{R}^3$, und $\boldsymbol{F} : \mathbb{R}^3 \to \mathbb{R}^3$ sei der Orthoprojektor auf $\operatorname{span}\{\boldsymbol{u}\}$. Dann gilt

$$\boldsymbol{F}(\boldsymbol{v}) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|^2}\boldsymbol{u} \text{ für alle } \boldsymbol{v} \in \mathbb{R}^3,$$

und die Matrix-Darstellung von $\boldsymbol{F}$ bzgl. der kanonischen Basis in $\mathbb{R}^3$ ist wegen (4.7)

$$\frac{1}{\|\boldsymbol{u}\|^2}\begin{bmatrix} u_1^2 & u_1u_2 & u_1u_3 \\ u_1u_2 & u_2^2 & u_2u_3 \\ u_1u_3 & u_2u_3 & u_3^2 \end{bmatrix}.$$

Dagegen ist, wenn $\boldsymbol{v}$, $\boldsymbol{w}$ eine beliebige Basis in $(\operatorname{span}\{\boldsymbol{u}\})^\perp$ ist,

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

die Matrix-Darstellung von $\boldsymbol{F}$ bzgl. der Basis $\frac{\boldsymbol{u}}{\|\boldsymbol{u}\|}$, $\boldsymbol{v}$, $\boldsymbol{w}$ in $\mathbb{R}^3$.

**Beispiel** Es sei $(u_1, u_2) \neq (0,0)$ ein fixierter Vektor in $\mathbb{R}^2$, und $\boldsymbol{F} : \mathbb{R}^2 \to \mathbb{R}^3$ sei die Spiegelung bzgl. $\operatorname{span}\{(u_1, u_2)\}$. Dann bilden die Vektoren

$$\boldsymbol{u} = \frac{(u_1, u_2)}{\sqrt{u_1^2 + u_2^2}} \quad \text{und} \quad \boldsymbol{u}^\perp = \frac{(u_2, -u_1)}{\sqrt{u_1^2 + u_2^2}}$$

eine Orthonormalbasis in $\mathbb{R}^2$, und

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

ist die Matrix-Darstellung von $\boldsymbol{F}$ bzgl. dieser Basis. Wegen (4.4) ist ferner

$$\boldsymbol{C} = \frac{1}{\sqrt{u_1^2 + u_2^2}} \begin{bmatrix} u_1 & u_2 \\ u_2 & -u_1 \end{bmatrix}$$

die Matrix des Wechsels von der Basis $\boldsymbol{u}, \boldsymbol{u}^\perp$ zur kanonischen Basis. Weil beide Basen orthonornal sind, gilt $\boldsymbol{C}^{-1} = \boldsymbol{C}^T$, und folglich ist

$$\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}^T = \frac{1}{u_1^2 + u_2^2} \begin{bmatrix} u_1^2 - u_2^2 & 2u_1u_2 \\ 2u_1u_2 & u_2^2 - u_1^2 \end{bmatrix}$$

die Matrix-Darstellung von $\boldsymbol{F}$ bzgl. der kanonischen Basis. Mit anderen Worten: Für alle $(v_1, v_2) \in \mathbb{R}^2$ gilt

$$\boldsymbol{F}(v_1, v_2) = \frac{1}{u_1^2 + u_2^2} \begin{bmatrix} u_1^2 - u_2^2 & 2u_1u_2 \\ 2u_1u_2 & u_2^2 - u_1^2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \frac{1}{u_1^2 + u_2^2} \begin{bmatrix} (u_1^2 - u_2^2)v_1 + 2u_1u_2v_2 \\ 2u_1u_2v_1 + (u_2^2 - u_1^2)v_2 \end{bmatrix}.$$

**Symmetrische und antisymmetrische Matrizen** (i) Eine Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ heißt symmetrisch (bzw. antisymmetrisch oder schiefsymmetrisch), wenn $\boldsymbol{A}^T = \boldsymbol{A}$ (bzw. $\boldsymbol{A}^T = -\boldsymbol{A}$) gilt.
(ii) In $\mathbb{M}(n \times n; \mathbb{R})$ ist $\boldsymbol{A}\cdot\boldsymbol{B} = \sum_{i,j=1}^n a_{ij}\overline{b_{ij}}$ (bei $\boldsymbol{A} = [a_{ij}]$, $\boldsymbol{B} = [b_{ij}]$) ein Skalarprodukt, und es gilt $\boldsymbol{A}\boldsymbol{B} \cdot \boldsymbol{C} = \boldsymbol{B} \cdot \boldsymbol{A}^T\boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{C}\boldsymbol{B}^T$. Ferner gilt

$$\{\boldsymbol{A} :\, \boldsymbol{A} = \boldsymbol{A}^T\}^\perp = \{\boldsymbol{A} :\, \boldsymbol{A} = -\boldsymbol{A}^T\},$$

d.h. die orthogonale Ergänzung in $\mathbb{M}(n \times n; \mathbb{R})$ (bzgl. des obigen Skalarprodukts) des Unterraums der symmetrischen Matrizen ist der Unterraum der antisymmetrischen Matrizen.

**Quadratische Formen** Eine Abbildung $q : \mathbb{R}^n \to \mathbb{R}$ heißt quadratische Form auf $\mathbb{R}^n$, wenn eine Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ existiert, so daß für alle $\boldsymbol{v}$ gilt

$$q(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{v} \cdot \boldsymbol{v}. \tag{4.8}$$

In diesem Fall existiert auch eine symmetrische Matrix $\boldsymbol{A} = [a_{ij}] \in \mathbb{M}(n \times n; \mathbb{R})$ mit (4.8), und diese ist vermittels

$$a_{ij} = \frac{1}{4}\Big(q(\boldsymbol{e}_i + \boldsymbol{e}_j) - q(\boldsymbol{e}_i - \boldsymbol{e}_j)\Big)$$

durch $q$ eindeutig bestimmt. Dabei ist $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ die kanonische Basis in $\mathbb{R}^n$ (vgl. (1.3)).

# 5 Determinanten

Die Determinante einer quadratischen Matrix $[a_{ij}] \in \mathbb{M}(n \times n; \mathbb{K})$ kann man z.B. definieren durch die sogenannte Leibnitz-Formel

$$\det\,[a_{ij}] = \sum_{j_1,\ldots,j_n=1}^{n} \epsilon_{j_1 \ldots j_n} a_{1j_1} \cdots a_{nj_n}. \tag{5.1}$$

Dabei ist

$$\epsilon_{j_1 \ldots j_n} = \begin{cases} 1 & \text{falls } (j_1 \ldots j_n) \text{ eine gerade Permutation von } (1, \ldots, n) \text{ ist,} \\ -1 & \text{falls } (j_1 \ldots j_n) \text{ eine ungerade Permutation von } (1, \ldots, n) \text{ ist,} \\ 0 & \text{sonst} \end{cases} \tag{5.2}$$

der sogenannte Permutationstensor, und eine Permutation $(j_1 \ldots j_n)$ ist gerade bzw. ungerade, wenn die Anzahl der Zahlenpaare $i < k$ mit $j_i > j_k$ gerade bzw. ungerade ist. In der Summe (5.1) ist also $\epsilon_{j_1 \ldots j_n}$ $n!/2$ mal gleich Eins und $n!/2$ mal gleich minus Eins. Zum Beispiel ist

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

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

**Eigenschaften** (i) Wenn man eine Zeile oder Spalte mit einer Zahl multipliziert und danach zu einer anderen Zeile oder Spalte addiert, ändert sich die Determinante nicht.
(ii) Wenn man zwei Zeilen oder Spalten vertauscht, so ändert sich die Determinante um den Faktor $-1$.
(iii) Wenn man eine Zeile oder Spalte mit einer Zahl $\lambda$ multipliziert, so ändert sich die Determinante

um den Faktor $\lambda$. Insbesondere gilt $\det(\lambda \boldsymbol{A}) = \lambda^n \det \boldsymbol{A}$.
(iv) Wenn die Spalten $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n \in \mathbb{K}^n$ fixiert sind, so ist die Abbildung

$$\boldsymbol{a} \in \mathbb{K}^n \mapsto \det[\boldsymbol{a}_1 \ \ldots \boldsymbol{a}_{j-1}\, \boldsymbol{a}\, \boldsymbol{a}_{j+1} \ldots \boldsymbol{a}_n] \in \mathbb{K}$$

linear. Aber die Abbildung $(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{K}^{n^2} \mapsto \det[\boldsymbol{a}_1 \ldots \boldsymbol{a}_n] \in \mathbb{K}$ ist nicht linear. Eine analoge Aussage gilt für Zeilen.
(v) Die Determinante $\det[a_{ij}]$ hängt glatt von den Koeffizienten $a_{ij}$ ab. Insbesondere gilt: Wenn $\det[a_{ij}] \neq 0$ ist, so existiert ein $\varepsilon > 0$, so daß auch $\det[b_{ij}] \neq 0$ ist, falls nur $|a_{ij} - b_{ij}| < \varepsilon$ für alle $i$ und $j$ ist.
(vi) Es gilt $\det \boldsymbol{A} = \det \boldsymbol{A}^T$ und $\det \boldsymbol{AB} = \det \boldsymbol{A} \det \boldsymbol{B}$ für alle $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{M}(n \times n; \mathbb{K})$.

**Entwicklung nach Zeilen oder Spalten** Für alle $i = 1, \cdots, n$ gilt

$$\det \boldsymbol{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \boldsymbol{A}'_{ij} \quad \text{bei} \quad \boldsymbol{A} = [a_{ij}].$$

Dabei ist $\boldsymbol{A}'_{ij}$ die $(n-1) \times (n-1)$-Matrix, die man erhält, wenn man aus der Ausgangsmatrix $\boldsymbol{A}$ die $i$te Zeile und die $j$te Spalte entfernt. Zum Beispiel ist

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

Analog kann man Matrizen auch nach Spalten entwickeln.

**Determinantenformel für die inverse Matrix** Für ein $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{K})$ sind folgende Eigenschaften äquivalent:
(i) $\boldsymbol{A}$ ist invertierbar.
(ii) $\det \boldsymbol{A} \neq 0$.
(iii) Die Zeilen von $\boldsymbol{A}$ sind linear unabhängig.
(iv) Die Spalten von $\boldsymbol{A}$ sind linear unabhängig.
In diesem Fall gilt

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \begin{bmatrix} \det \boldsymbol{A}'_{11} & -\det \boldsymbol{A}'_{21} & \ldots & (-1)^{n+1} \det \boldsymbol{A}'_{n1} \\ -\det \boldsymbol{A}'_{12} & \det \boldsymbol{A}'_{22} & \ldots & (-1)^{n+2} \det \boldsymbol{A}'_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^{1+n} \det \boldsymbol{A}'_{1n} & (-1)^{2+n} \det \boldsymbol{A}'_{2n} & \ldots & \det \boldsymbol{A}'_{nn} \end{bmatrix}, \tag{5.3}$$

zum Beispiel

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Insbesondere hängen die Koeffizienten von $\boldsymbol{A}^{-1}$ glatt von den Koeffizienten von $\boldsymbol{A}$ ab.

**Beispiel** In $\mathbb{R}_2[x]$ sei $(x+1)^2$, $(x+2)^2$ und $(x+3)^2$ die alte Basis, und 1, $x$ und $x^2$ sei die neue Basis. Wegen $(x+1)^2 = x^2+2x+1$, $(x+2)^2 = x^2+4x+4$, $(x+3)^2 = x^2+6x+9$ und (4.2) ist dann

$$C = \begin{bmatrix} 1 & 4 & 9 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{bmatrix}$$

die Matrix des Basiswechsels von der alten zur neuen Basis. Ihre Determinante ist $-4$. Folglich ergibt die Formel (5.3)

$$C^{-1} = -\frac{1}{4} \begin{bmatrix} \det \begin{bmatrix} 4 & 6 \\ 1 & 1 \end{bmatrix} & -\det \begin{bmatrix} 4 & 9 \\ 1 & 1 \end{bmatrix} & \det \begin{bmatrix} 4 & 9 \\ 4 & 6 \end{bmatrix} \\ -\det \begin{bmatrix} 2 & 6 \\ 1 & 1 \end{bmatrix} & \det \begin{bmatrix} 1 & 9 \\ 1 & 1 \end{bmatrix} & -\det \begin{bmatrix} 1 & 9 \\ 2 & 6 \end{bmatrix} \\ \det \begin{bmatrix} 4 & 9 \\ 4 & 6 \end{bmatrix} & -\det \begin{bmatrix} 1 & 9 \\ 2 & 6 \end{bmatrix} & \det \begin{bmatrix} 1 & 4 \\ 2 & 4 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1/2 & -5/4 & 3 \\ -1 & 2 & -3 \\ 3 & -3 & 1 \end{bmatrix}.$$

Die Matrix-Darstellung der Ableitungsabbildung $d/dx$ bzgl. der neuen Basis ist

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix},$$

bzgl. der alten Basis aber

$$C^{-1}AC = \begin{bmatrix} 1/2 & -5/4 & 3 \\ -1 & 2 & -3 \\ 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 4 & 9 \\ 2 & 4 & 6 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -3/2 & -1/2 & 1/2 \\ 2 & 0 & -2 \\ 0 & 6 & 12 \end{bmatrix}.$$

Das Beispiel zeigt, daß die Matrix-Darstellung einer linearen Abbildung bzgl. einer Basis "gut" sein kann, d.h. viele verschwindende Koeffizienten besitzen kann, während die Matrix-Darstellung derselben linearen Abbildung bzgl. einer anderen Basis "schlecht" ist.

**Rang** Für jede Matrix $\boldsymbol{A} \in \mathbb{M}(m \times n; \mathbb{K})$ sind die maximale Anzahl linear unabhängiger Spalten von $\boldsymbol{A}$, die maximale Anzahl linear unabhängiger Zeilen von $\boldsymbol{A}$ sowie die maximale Ordnung von quadratischen Untermatrizen von $\boldsymbol{A}$ mit nichtverschwindender Determinante gleich. Diese Zahl, die zwischen Null und $\min\{m,n\}$ liegt, heißt Rang von $\boldsymbol{A}$ und wird mit $\operatorname{rang} \boldsymbol{A}$ bezeichnet, und es gilt wegen (3.1)

$$\operatorname{rang} \boldsymbol{A} = \dim \operatorname{im} \boldsymbol{A} = n - \dim \ker \boldsymbol{A}. \tag{5.4}$$

**Orthogonale Matrizen** (i) Für eine Martix $\boldsymbol{Q} \in \mathbb{M}(n \times n; \mathbb{R})$ gilt

$$\boldsymbol{Q}^T = \boldsymbol{Q}^{-1} \text{ genau dann, wenn } \boldsymbol{Qu} \cdot \boldsymbol{Qv} = \boldsymbol{u} \cdot \boldsymbol{v} \text{ für alle } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n,$$

d.h. wenn $\boldsymbol{Q}$ das Skalarprodukt zweier beliebiger Vektoren nicht ändert. In diesem Fall heißt $\boldsymbol{Q}$ orthogonal, und die Menge aller solcher Matrizen wird mit $O(n)$ bezeichnet.

(ii) Es gilt $\boldsymbol{Q} \in O(n)$ genau dann, wenn $\boldsymbol{Q}$ die Matrix des Basiswechsels zwischen zwei Orthonormalbasen in $\mathbb{R}^n$ ist.
(iii) Wenn $\boldsymbol{Q}_1$ und $\boldsymbol{Q}_2$ aus $O(n)$ sind, so sind auch die Produkte $\boldsymbol{Q}_1\boldsymbol{Q}_2$ und $\boldsymbol{Q}_2\boldsymbol{Q}_1$ in $O(n)$. Mit anderem Worten: $O(n)$ ist eine Gruppe bzgl. der Matrizenmultiplikation (die sogenannte orthogonale Gruppe $n$ter Ordnung).
(iv) Für $\boldsymbol{Q} \in O(n)$ gilt $1 = \det(Q^TQ) = (\det Q)^2$, also $\det Q = \pm 1$. Die Menge

$$SO(n) = \{\boldsymbol{Q} \in O(n) : \det \boldsymbol{Q} = 1\}$$

ist ebenfalls eine Gruppe bzgl. der Matrizenmultiplikation (die sogenannte spezielle orthogonale Gruppe $n$ter Ordnung).

**Die orthogonale Gruppe zweiter Ordnung** Die Matrizen

$$\boldsymbol{Q}(\varphi) = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

(Drehungen um den Winkel $\varphi$ entgegen dem Uhrzeigersinn) bilden die Elemente von $SO(2)$, und $O(2) = SO(2) \cup \{JQ(\varphi) : \varphi \in \mathbb{R}\}$. Dabei ist

$$\boldsymbol{J} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

die Spiegelung an der $x_1$-Achse, und es gilt

$$\boldsymbol{J}^2 = \boldsymbol{I},\ \boldsymbol{Q}(\varphi)\boldsymbol{Q}(\psi) = \boldsymbol{Q}(\varphi+\psi) \text{ und } \boldsymbol{J}\boldsymbol{Q}(\varphi) = \boldsymbol{Q}(-\varphi)\boldsymbol{J}.$$

Daraus folgt: Die Elemente von $O(2)$ sind entweder Drehungen oder Spiegelungen, denn

$$\boldsymbol{J}\boldsymbol{Q}(\varphi) = \boldsymbol{Q}(-\varphi/2)\boldsymbol{J}\boldsymbol{Q}(\varphi/2)$$

ist die Spiegelung an der um den Winkel $-\varphi/2$ im Uhrzeigersinn gedrehten $x_1$-Achse. Insbesondere gilt, daß die Nacheinanderausführung einer Drehung und einer Spiegelung oder einer Spiegelung und einer Drehung eine Spiegelung (an einer anderen Achse) ist, während die Nacheinanderausführung zweier Spiegelungen eine Drehung ist.

**Die orthogonale Gruppe dritter Ordnung und Eulersche Winkel** Die Matrizen

$$\boldsymbol{Q}(\psi,\theta,\varphi) = \begin{bmatrix} \cos\psi\cos\varphi - \sin\psi\cos\theta\sin\varphi & -\cos\psi\sin\varphi - \sin\psi\cos\theta\cos\varphi & \sin\psi\sin\theta \\ \sin\psi\cos\varphi + \cos\psi\cos\theta\sin\varphi & -\sin\psi\sin\varphi + \cos\psi\cos\theta\cos\varphi & -\cos\psi\sin\theta \\ \sin\theta\sin\varphi & \sin\theta\cos\varphi & \cos\theta \end{bmatrix}$$

bilden die Elemente von $SO(3)$. Wenn $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ die kanonische Basis in $\mathbb{R}^3$ ist (vgl. (1.3)), $\boldsymbol{e}'_j = \boldsymbol{Q}(\psi,\theta,\varphi)\boldsymbol{e}_j$ $(j = 1, 2, 3)$ die neue Orthonormalbasis nach der Drehung $\boldsymbol{Q}(\psi,\theta,\varphi)$ und $\boldsymbol{v} \in \mathbb{R}^3$ ein Vektor in der Schnittgeraden von $\operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ und $\operatorname{span}\{\boldsymbol{e}'_1, \boldsymbol{e}'_2\}$ mit $\det[\boldsymbol{e}_3\ \boldsymbol{Q}(\psi,\theta,\varphi)\boldsymbol{e}_3\ \boldsymbol{v}] > 0$, so gilt:
(i) Der sogenannte Präzessionswinkel $\psi \in [0, 2\pi)$ ist der Winkel von $\boldsymbol{e}_1$ nach $\boldsymbol{v}$ (Drehung um die

$\boldsymbol{e}_3$-Achse gegen den Uhrzeigersinn der $\boldsymbol{e}_1, \boldsymbol{e}_2$-Ebene).
(ii) Der sogenannte Nutationswinkel $\theta \in [0, \pi)$ ist der Winkel zwischen $\boldsymbol{e}_3$ und $\boldsymbol{e}_3'$ (Drehung um die $\boldsymbol{v}$-Achse).
(iii) Der Winkel $\varphi \in [0, 2\pi)$ ist der Winkel von $\boldsymbol{v}$ nach $\boldsymbol{e}_1'$ (Drehung um die $\boldsymbol{e}_3'$-Achse gegen den Uhrzeigersinn der $\boldsymbol{e}_1', \boldsymbol{e}_2'$-Ebene).
Die Winkel $\psi$, $\theta$ und $\varphi$ heißen Eulersche Winkel der Drehung $\boldsymbol{Q}(\psi, \theta, \varphi)$. Sie beschreiben eine in vielen Anwendungen nützliche Art der Parametrisierung der "3-dimensionalen Fläche" $SO(3)$ im 9-dimensionalen Vektorraum $\mathbb{M}(3 \times 3; \mathbb{R})$.

# 6 Vektorprodukt und Orientierung

In diesem Kapitel wird ausschließlich im Vektorraum $\mathbb{R}^3$ gearbeitet. Am Ende des Kapitels wird gezeigt, wie die Ergebnisse auf den Fall eines beliebigen endlichdimensionalen reellen Vektorraumes mit Skalarprodukt verallgemeinert werden können.

**Positiv orientierte Dreibeine und Regel der rechten Hand** Drei linear unabhängige Vektoren $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \in \mathbb{R}^3$ heißen positiv oder rechts orientiert (bzw. negativ oder links orientiert), wenn die Determinante der Matrix des Basiswechsels von der kanonischen Basis zur Basis $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ positiv (bzw. negativ ist). Dies ist genau dann der Fall, wenn folgendes gilt:
Es sei $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ die kanonische Basis in $\mathbb{R}^3$ (vgl. (1.3)). Ferner sei $\boldsymbol{Q} \in SO(3; \mathbb{R})$ eine Drehung des $\mathbb{R}^3$, so daß $\boldsymbol{Q}\boldsymbol{v}_1 = \lambda \boldsymbol{e}_1$ ist mit $\lambda > 0$, daß $\boldsymbol{Q}\boldsymbol{v}_2 \in \operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ ist und daß $\boldsymbol{Q}\boldsymbol{v}_2 \cdot \boldsymbol{e}_2 > 0$ ist (d.h. $\boldsymbol{Q}\boldsymbol{v}_1$ ist ein skalares Vielfaches von $\boldsymbol{e}_1$, und $\boldsymbol{Q}\boldsymbol{v}_2$ und $\boldsymbol{e}_2$ zeigen in dieselbe der beiden Halbebenen, die durch die $\boldsymbol{e}_1$-Gerade in der $\boldsymbol{e}_1\boldsymbol{e}_2$-Ebene erzeugt werden). Dann muß $\boldsymbol{Q}\boldsymbol{v}_3 \cdot \boldsymbol{e}_3 > 0$ gelten (d.h. $\boldsymbol{Q}\boldsymbol{v}_3$ und $\boldsymbol{e}_3$ zeigen in denselben der beiden Halbräume, die durch die $\boldsymbol{e}_1\boldsymbol{e}_2$-Ebene in $\mathbb{R}^3$ erzeugt werden).

**Vektorprodukt** Das Vektorprodukt $\boldsymbol{u} \times \boldsymbol{v} \in \mathbb{R}^3$ von zwei Vektoren $\boldsymbol{u}$ und $\boldsymbol{v}$ aus $\mathbb{R}^3$ ist folgendermaßen definiert: Falls $\boldsymbol{u}$ und $\boldsymbol{v}$ linear abhängig sind, so gilt $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{0}$. Andernfalls bilden $\boldsymbol{u}$, $\boldsymbol{v}$ und $\boldsymbol{u} \times \boldsymbol{v}$ ein positiv orientiertes Dreibein, $\boldsymbol{u} \times \boldsymbol{v}$ steht auf $\operatorname{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ senkrecht steht und besitzt die Länge

$$\|\boldsymbol{u} \times \boldsymbol{v}\| = \|\boldsymbol{u}\|\|\boldsymbol{v}\| \sin \angle(\boldsymbol{u}, \boldsymbol{v}).$$

Dabei ist $\angle(\boldsymbol{u}, \boldsymbol{v}) \in (0, \pi)$ der Winkel zwischen $\boldsymbol{u}$ und $\boldsymbol{v}$, d.h. $\|\boldsymbol{u}\|\|\boldsymbol{v}\| \cos \angle(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{v}$.
Wenn $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ die kanonische Basis in $\mathbb{R}^3$ ist, so kann man die Koordinaten des Vektorprodukts aus den Koordinaten der Faktoren (Koordinaten jeweils bzgl. der kanonischen Basis) folgendermaßen berechnen:

$$(u_1, u_2, u_3) \times (v_1, v_2, v_3) = \begin{bmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{bmatrix} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).$$

**Rechenregeln** Für alle $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ und $\lambda, \mu \in \mathbb{R}$ gilt $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$, $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u} \cdot \boldsymbol{v})\boldsymbol{w}$, $\|\boldsymbol{u} \times \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 - (\boldsymbol{u} \cdot \boldsymbol{v})^2$ und

$$(\lambda \boldsymbol{u} + \mu \boldsymbol{v}) \times \boldsymbol{w} = \lambda(\boldsymbol{u} \times \boldsymbol{w}) + \mu(\boldsymbol{v} \times \boldsymbol{w}). \tag{6.1}$$

Die Rechenregel (6.1) entspricht den analogen Rechenregeln bei Skalarprodukten und Matrizenprodukten und besitzt Analoga auch bei allen anderen "Produkten", die in der linearen Algebra eingeführt werden (wie Tensorprodukt oder äußeres Produkt). Sie ist eine Verallgemeinerung der Rechenregel "Punktrechnung geht vor Strichrechnung". Sie bedeutet, daß, wenn man einen Faktor fixiert, das Produkt linear von dem anderen Faktor abhängt. Von beiden Faktoren (als Paar betrachtet) hängt das Produkt allerdings nicht linear ab, denn

$$(\boldsymbol{u}_1 + \boldsymbol{v}_1) \times (\boldsymbol{u}_2 + \boldsymbol{v}_2) = \boldsymbol{u}_1 \times \boldsymbol{u}_2 + \boldsymbol{u}_1 \times \boldsymbol{v}_2 + \boldsymbol{v}_1 \times \boldsymbol{u}_2 + \boldsymbol{v}_1 \times \boldsymbol{v}_2.$$

Allen Produkten ist gemeinsam, daß, wenn die Faktoren glatt von einem Parameter abhängen, die Ableitung des Produkts nach dem Parameter nach der bekannten Produkt-Regel berechnet werden kann. Zum Beispiel gilt

$$\frac{d}{dt}\Big(\boldsymbol{u}(t) \times \boldsymbol{v}(t)\Big) = \frac{d\boldsymbol{u}(t)}{dt} \times \boldsymbol{v}(t) + \boldsymbol{u}(t) \times \frac{d\boldsymbol{v}(t)}{dt}.$$

Allerdings muß dabei beachtet werden, ob die Produkte kommutativ sind oder nicht. Zum Beispiel gilt für das Skalarprodukt reeller Vektoren (das kommutativ ist)

$$\frac{d}{dt}\Big(\boldsymbol{u}(t) \cdot \boldsymbol{u}(t)\Big) = 2\,\frac{d\boldsymbol{u}(t)}{dt} \cdot \boldsymbol{u}(t),$$

während für das Produkt von Matrizen (das im allgemeinen nicht kommutativ ist) gilt

$$\frac{d}{dt}\Big(\boldsymbol{A}(t)\boldsymbol{A}(t)\Big) = \frac{d\boldsymbol{A}(t)}{dt}\boldsymbol{A}(t) + \boldsymbol{A}(t)\frac{d\boldsymbol{A}(t)}{dt},$$

wobei die beiden Summanden der rechte Seite im allgemeinen ungleich sind.

**Anwendungsbeispiel: Einheitsnormalen** Wenn $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ linear unabhängig sind, so ist $\operatorname{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ die Ebene mit den Einheitsnormalen

$$\pm\frac{\boldsymbol{u} \times \boldsymbol{v}}{\|\boldsymbol{u} \times \boldsymbol{v}\|}.$$

**Anwendungsbeispiel: Darstellung der Schnittgeraden zweier Ebenen** Wenn $\boldsymbol{u}$, $\boldsymbol{v}$ und $\boldsymbol{u}'$, $\boldsymbol{v}'$ jeweils zwei linear unabhängige Vektoren in $\mathbb{R}^3$ sind, so gilt

$$\operatorname{span}\{\boldsymbol{u}, \boldsymbol{v}\} \cap \operatorname{span}\{\boldsymbol{u}', \boldsymbol{v}'\} = \operatorname{span}\{(\boldsymbol{u} \times \boldsymbol{v}) \times (\boldsymbol{u}' \times \boldsymbol{v}')\}.$$

**Anwendungsbeispiel: Volumen von Parallelepipeds** Für beliebige $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ gilt

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \boldsymbol{v} \cdot (\boldsymbol{w} \times \boldsymbol{u}) = \boldsymbol{w} \cdot (\boldsymbol{u} \times \boldsymbol{v}),$$

und der Betrag dieser Zahl einerseits gleich dem Volumen des von $\boldsymbol{u}$, $\boldsymbol{v}$ und $\boldsymbol{w}$ aufgespannten Paralellepipeds $\{\alpha\boldsymbol{u} + \beta\boldsymbol{v} + \gamma\boldsymbol{w} :\ 0 \le \alpha, \beta, \gamma \le 1\}$ und andererseits gleich

$$|\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})| = \sqrt{\det \begin{bmatrix} \boldsymbol{u} \cdot \boldsymbol{u} & \boldsymbol{u} \cdot \boldsymbol{v} & \boldsymbol{u} \cdot \boldsymbol{w} \\ \boldsymbol{v} \cdot \boldsymbol{u} & \boldsymbol{v} \cdot \boldsymbol{v} & \boldsymbol{v} \cdot \boldsymbol{w} \\ \boldsymbol{w} \cdot \boldsymbol{u} & \boldsymbol{w} \cdot \boldsymbol{v} & \boldsymbol{w} \cdot \boldsymbol{w} \end{bmatrix}}. \tag{6.2}$$

Insbesondere ist die Determinante unter der Wurzel, die sogenannte Gramsche Determinante der Vektoren $\boldsymbol{u}$, $\boldsymbol{v}$ und $\boldsymbol{w}$, positiv, wenn $\boldsymbol{u}$, $\boldsymbol{v}$ und $\boldsymbol{w}$ linear unabhängig sind (und sonst Null).
Wenn $u_1, u_2, u_3$ bzw. $v_1, v_2, v_3$ bzw. $w_1, w_2, w_3$ die Koordinaten von $\boldsymbol{u}$ bzw. $\boldsymbol{v}$ bzw. $\boldsymbol{w}$ bzgl. einer beliebigen positiv orientierten Orthonormalbasis sind, so gilt

$$|\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})| = \left| \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix} \right|$$

**Anwendungsbeispiel: Die Determinante als Maß für Volumenänderung** Es seien eine Matrix $\boldsymbol{A} \in \mathbb{M}(3 \times 3; \mathbb{R})$ und drei Vektoren $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ gegeben. Dann ist das Volumen des von $\boldsymbol{Au}$, $\boldsymbol{Av}$ und $\boldsymbol{Aw}$ aufgespannten Paralellepipeds gleich

$$|\det[\boldsymbol{Au}\ \boldsymbol{Av}\ \boldsymbol{Aw}]| = |\det \boldsymbol{A}| |\det[\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}]|,$$

d.h. $\boldsymbol{A}$ überführt Paralellepipeds in Paralellepipeds und ändert dabei das Volumen um den Faktor $|\det \boldsymbol{A}|$.

**Anwendungsbeispiel: Fläche von Parallelogrammen** Die Fläche des von zwei Vektoren $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ aufgespannten Parallelogramms $\{\alpha\boldsymbol{u} + \beta\boldsymbol{v} :\ 0 \leq \alpha, \beta \leq 1\}$ ist gleich

$$\|\boldsymbol{u} \times \boldsymbol{v}\| = \sqrt{\det \begin{bmatrix} \boldsymbol{u} \cdot \boldsymbol{u} & \boldsymbol{u} \cdot \boldsymbol{v} \\ \boldsymbol{u} \cdot \boldsymbol{v} & \boldsymbol{v} \cdot \boldsymbol{v} \end{bmatrix}}. \tag{6.3}$$

Wenn $u_1, u_2, u_3$ bzw. $v_1, v_2, v_3$ die Koordinaten von $\boldsymbol{u}$ bzw. $\boldsymbol{v}$ bzgl. einer beliebigen positiv orientierten Orthonormalbasis $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ sind, und wenn $\boldsymbol{P}_{jk}$ der Orthoprojektor auf span$\{\boldsymbol{b}_j, \boldsymbol{b}_k\}$ ist, so gilt

$$\|\boldsymbol{u} \times \boldsymbol{v}\|^2 = \sum_{j<k} \|(\boldsymbol{P}_{jk}\boldsymbol{u}) \times (\boldsymbol{P}_{jk}\boldsymbol{v})\|^2 = \sum_{j<k} \det \begin{bmatrix} u_j^2 + u_k^2 & u_j v_j + u_k v_k \\ u_j v_j + u_k v_k & v_j^2 + v_k^2 \end{bmatrix},$$

d.h. das Quadrat der Fläche des Parallelogramms ist gleich der Summe der Quadrate der Flächen der orthogonalen Projektionen des Parallelogramms auf die Koordinatenebenen. Das ist ein Analogon zum Satz von Pythagoras.

**Anwendungsbeispiel: Bewegungen eines starren Körpers um einen fixierten Punkt** Die Teilchen des Körpers mögen zum Zeitpunkt $t = 0$ die Menge $\Omega \subset \mathbb{R}^3$ einnehmen. Ferner möge $\boldsymbol{0} \in \Omega$ gelten, und der Körper möge im Nullpunkt fest verankert sein.
**(i) Bewegung mit konstanter Winkelgeschwindigkeit:**
Es sei $\boldsymbol{e}_3 = (0, 0, 1)$ der dritte Vektor der kanonischen Basis in $\mathbb{R}^3$. Wenn sich der Körper z.B. um die $\boldsymbol{e}_3$-Achse dreht, so daß die Länge des Geschwindigkeitsvektors jedes Teilchens $\boldsymbol{x} \in \Omega$ konstant bzgl. der Zeit $t$ ist, so ist der Ort $\boldsymbol{p}(t, \boldsymbol{x})$ des Teilchens $\boldsymbol{x} = (x_1, x_2, x_3)$ zum Zeitpunkt $t$ gleich

$$\boldsymbol{p}(t, \boldsymbol{x}) = \begin{bmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

und folglich ist dier Geschwindigkeitsvektor $\partial_t \boldsymbol{p}(t,\boldsymbol{x})$ gleich

$$\begin{bmatrix} -\omega\sin\omega t & -\omega\cos\omega t & 0 \\ \omega\cos\omega t & -\omega\sin\omega t & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \omega \boldsymbol{e}_3 \times \boldsymbol{p}(t,\boldsymbol{x}).$$

Analog gilt: Wenn sich der Körper um eine Achse, die von einem Vektor $\boldsymbol{u} \in \mathbb{R}^3$ mit $\|\boldsymbol{u}\| = 1$ aufgespannt wird, dreht, so daß die Länge des Geschwindigkeitsvektors $\partial_t \boldsymbol{p}(t,\boldsymbol{x})$ jedes Teilches $\boldsymbol{x} \in \Omega$ konstant bzgl. der Zeit $t$ ist, so ist $\partial_t \boldsymbol{p}(t,\boldsymbol{x}) = \omega \boldsymbol{u} \times \boldsymbol{p}(t,\boldsymbol{x})$ mit einem $\omega \in \mathbb{R}$. Der Vektor $\omega \boldsymbol{u}$ bzw. die Zahl $|\omega|$ heißen dann vektorielle bzw. skalare Winkelgeschwindigkeit dieser Bewegung. Dabei ist der Winkel, um den sich der Körper in der Zeit $t > 0$ dreht, gleich $|\omega| t$.

**(ii) Bewegung mit variabler Winkelgeschwindigkeit:**
Wenn sich der Körper beliebig um den fixierten Nullpunkt bewegt, so gilt

$$\boldsymbol{p}(t,\boldsymbol{x}) = \boldsymbol{Q}(t)\boldsymbol{x} \quad \text{mit} \quad \boldsymbol{Q}(t) \in SO(3;\mathbb{R}),$$

und zu jedem Zeitpunkt $t$ existiert eine "virtuelle" Winkelgeschwindigkeit $\boldsymbol{\omega}(t) \in \mathbb{R}^3$, so daß für alle $\boldsymbol{x} \in \Omega$ gilt

$$\partial_t \boldsymbol{p}(t,\boldsymbol{x}) = \boldsymbol{\omega}(t) \times \boldsymbol{p}(t,\boldsymbol{x}).$$

Mit anderen Worten: Zu jedem Zeitpunkt existiert eine "virtuelle" Achse durch den Nullpunkt, deren Punkte zu dem gegebenen Zeitpunkt die Geschwindigkeit Null besitzen. Die Geschwindigkeitsvektoren aller anderen Punkte sind zu dem gegebenen Zeitpunkt orthogonal zu dieser Achse und zum Lot aus dem Punkt zur Achse, und die Länge der Geschwindigkeitsvektoren ist proportional zum Abstand der entsprechenden Punkte von der Achse.

**Anwendungsbeispiel: Lorentz-Kraft** Wenn sich ein geladenes Teilchen der Ladung $q$ mit der Geschwindigkeit $\boldsymbol{v}$ durch ein Magnetfeld mit der magnetischen Induktion $\boldsymbol{B}$ bewegt, so wirkt auf das Teilchen die Lorentz-Kraft $q\boldsymbol{v} \times \boldsymbol{B}$. Insbesondere ist die Lorentz- Kraft orthogonal zu der Geschwindigkeit und der Induktion, und Geschwindigkeit, Induktion und Lorentz-Kraft bilden ein positiv orientiertes Dreibein, falls das Teilchen positiv geladen ist.

**Anwendungsbeispiel: Coriolis-Kraft** Eine Punktmasse $m$ befinde sich zum Zeitpunkt $t$ am Ort

$$\boldsymbol{p}(t) = \boldsymbol{Q}(t)\boldsymbol{x}(t) \quad \text{mit} \quad \boldsymbol{Q}(t) = \begin{bmatrix} \cos\omega t & -\sin\omega t & 0 \\ \sin\omega t & \cos\omega t & 0 \\ 0 & 0 & 1 \end{bmatrix} \in SO(3) \quad \text{und} \quad \boldsymbol{x}(t) \in \mathbb{R}^3$$

(z.B. Bewegung auf der Erdoberfläche, überlagert mit der Rotation der Erde). Dann gilt

$$\frac{d^2\boldsymbol{p}(t)}{dt^2} = \boldsymbol{Q}(t)\frac{d^2\boldsymbol{x}(t)}{dt^2} + 2\omega \boldsymbol{e}_3 \times \frac{d\boldsymbol{x}(t)}{dt}.$$

Die Trägheitskraft $-m\frac{d^2\boldsymbol{p}(t)}{dt^2}$ ist also die Summe aus der sogenannten relativen Trägheitskraft $-m\boldsymbol{Q}(t)\frac{d^2\boldsymbol{x}(t)}{dt^2}$ (die Trägheitskraft, die ein Beobachter, der sich mit der Masse $m$ bewegt und nichts von der Überlagerung mit der Rotation $\boldsymbol{Q}(t)$ weiß, erwarten würde) und der sogenannten Coriolis-Kraft

$-2m\omega\boldsymbol{e}_3 \times \frac{d\boldsymbol{x}(t)}{dt}$. Die Coriolis-Kraft ist orthogonal zu der Ebene, die durch $\boldsymbol{e}_3$ und die Relativgeschwindigkeit $\frac{d\boldsymbol{x}(t)}{dt}$ aufgespannt ist, und sie verschwindet (falls $m > 0$ und $\omega > 0$) genau dann, wenn die Relativgeschwindigkeit und $\boldsymbol{e}_3$ parallel sind. Wenn die Relativgeschwindigkeit und $\boldsymbol{e}_3$ nicht parallel sind, bilden Coriolis-Kraft, Relativgeschwindigkeit und $\boldsymbol{e}_3$ ein positiv orientiertes Dreibein.
Im Fall der Bewegung auf der Erdoberfläche bedeutet das folgendes: Wenn die Relativgeschwindigkeit z.B. zum Erdmittelpunkt zeigt, so wirkt die Coriolis-Kraft als nach Osten ablenkende Kraft (maximal am Äquator und verschwindend an den Polen). Wenn die Relativgeschwindigkeit dagegen tangential zur Erdoberfläche ist, so wirkt die Coriolis-Kraft auf der nördlichen (bzw. südlichen) Halbkugel als nach rechts (bzw. links) ablenkende Kraft (maximal an den Polen und verschwindend am Äquator).

**Orientierung eines beliebigen endlichdimensionalen reellen Vektorraumes** Es sei $V$ ein endlichdimensionaler reeller Vektorraum. Zwei Basen in $V$ heißen gleich orientiert, wenn die Determinante der Matrix des Wechsels von der ersten Basis zur zweiten (oder von der zweiten zur ersten) positiv ist. Offenbar zerfällt die Menge aller Basen in $V$ in zwei disjunkte Teilmengen $M_1$ und $M_2$, so daß alle Basen aus $M_1$ (bzw. aus $M_2$) untereinander gleich orientiert sind. Man sagt, daß eine Orientierung in $V$ gewählt ist, wenn eine der beiden Mengen $M_1$ oder $M_2$ ausgezeichnet ist. Die Basen aus dieser ausgezeichneten Menge heißen dann positiv orientiert, die anderen negativ orientiert.
In $\mathbb{R}^n$ z.B. bietet sich eine kanonische Wahl der Orientierung an: Man zeichnet diejenige der Mengen $M_1$ und $M_2$ aus, in der die kanonische Basis liegt, d.h. man nennt alle die Basen positiv orientiert, die zur kanonischen Basis gleich orientiert sind. Wenn $V$ aber z.B. ein echter Unterraum des $\mathbb{R}^n$ ist, so existiert im allgemeinen keine kanonische Wahl einer Orientierung in $V$.
Wenn $\dim V = 1$ (bzw. 2 bzw. 3) ist, so bedeutet die Wahl einer Orientierung in $V$ die Festlegung einer "Richtung" (bzw. einer "Drehrichtung" bzw. einer "Schraubenrichtung") in $V$.
Wenn $V$ eine Hyperebene in einem Vektorraum $W$ ist und wenn in $W$ schon eine Orientierung festgelegt ist, so kann man eine Orientierung in $V$ dadurch wählen, daß man eine Einheitsnormale $\boldsymbol{w}$ an $V$ auszeichnet und dann festlegt, daß eine Basis $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{n-1}$ in $V$ positiv orientiert sein soll, wenn die Basis $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{n-1}, \boldsymbol{w}$ positiv orientiert in $W$ ist. Wenn $W$ z.B. der $\mathbb{R}^3$ mit seiner kanonischen Orientierung ist und wenn $V$ eine Ebene durch den Nullpunkt in $\mathbb{R}^3$ ist, so bedeutet diese Wahl der Orientierung in $V$ folgendes: Wenn man aus dem Halbraum in $\mathbb{R}^3$, in den die ausgezeichnete Normale $\boldsymbol{w}$ zeigt, auf $V$ schaut, so ist die entsprechende "Drehrichtung" gegen den Uhrzeigersinn gerichtet.

**Vektorprodukt in einem beliebigen endlichdimensionalen reellen orientierten Vektorraum mit Skalarprodukt** Es sei $V$ ein reeller Vektorraum mit Skalarprodukt, $\dim V = n$, und in $V$ sei eine Orientierung fixiert. Dann wird das Vektorprodukt $\boldsymbol{v}_1 \times \cdots \times \boldsymbol{v}_{n-1} \in V$ von Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1} \in V$ folgendermaßen definiert: $\boldsymbol{v}_1 \times \cdots \times \boldsymbol{v}_{n-1} = \boldsymbol{0}$, falls die Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1}$ linear abhängig sind. In gegenteiligen Fall ist $\boldsymbol{v}_1 \times \cdots \times \boldsymbol{v}_{n-1}$ eindeutig definiert durch die Festlegungen, daß die Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1}, \boldsymbol{v}_1 \times \cdots \times \boldsymbol{v}_{n-1}$ positiv orientiert sein sollen und daß gelten soll

$$\boldsymbol{v}_1 \times \cdots \times \boldsymbol{v}_{n-1} \in (\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1}\})^{\perp},$$

$$\|\boldsymbol{v}_1 \times \cdots \times \boldsymbol{v}_{n-1}\| = \sqrt{\det \begin{bmatrix} \boldsymbol{v}_1 \cdot \boldsymbol{v}_1 & \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_1 \cdot \boldsymbol{v}_{n-1} \\ \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 & \boldsymbol{v}_2 \cdot \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_2 \cdot \boldsymbol{v}_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{v}_1 \cdot \boldsymbol{v}_{n-1} & \boldsymbol{v}_2 \cdot \boldsymbol{v}_{n-1} & \ldots & \boldsymbol{v}_{n-1} \cdot \boldsymbol{v}_{n-1} \end{bmatrix}}.$$

# 7 Lineare Gleichungssysteme

In diesem Kapitel werden lineare Gleichungssystem

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m \end{array}\right\} \tag{7.1}$$

betrachtet. Anstelle von (7.1) schreibt man auch kompakter $\boldsymbol{Ax} = \boldsymbol{y}$. Dabei sind

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in \mathbb{M}(m \times n; \mathbb{K}), \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{K}^m \quad \text{bzw.} \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n$$

die gegebene Koeffizientenmatrix, der gegebene Vektor der rechten Seiten bzw. der gesuchte Vektor der Unbekannten. Bei der Diskussion des Lösungsverhaltens von (7.1) spielt ferner die sogenannte erweiterte Koeffizientenmatrix

$$[\boldsymbol{A}\ \boldsymbol{y}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & y_1 \\ a_{21} & a_{22} & \dots & a_{2n} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & y_m \end{bmatrix} \in \mathbb{M}(m \times (n+1); \mathbb{K})$$

eine Rolle. Das Gleichungssystem (7.1) heißt homogen, wenn $\boldsymbol{y} = \boldsymbol{0}$ ist.

**Notwendige und hinreichende Lösbarkeitsbedingung** Das Gleichungssystem (7.1) besitzt genau dann mindestens eine Lösung, wenn

$$\operatorname{rang}[\boldsymbol{A}\ \boldsymbol{y}] = \operatorname{rang}\boldsymbol{A}. \tag{7.2}$$

Dies ist wiederum genau dann der Fall, wenn $\boldsymbol{y} \cdot \boldsymbol{v} = 0$ für alle $\boldsymbol{v} \in \mathbb{K}^m$ mit $\boldsymbol{A}^*\boldsymbol{v} = \boldsymbol{0}$. Mit anderen Worten: Es gilt

$$\mathbb{K}^m = \operatorname{im}\boldsymbol{A} \oplus \ker\boldsymbol{A}^* \quad \text{(orthogonale Summe).}$$

Die Menge der $\boldsymbol{y}$, die (7.2) erfüllen, ist $\operatorname{im}\boldsymbol{A}$, also ein Unterraum in $\mathbb{K}^m$ der Dimension $\operatorname{rang}\boldsymbol{A}$. Wenn (7.2) erfüllt ist, so ist $\operatorname{rang}[\boldsymbol{A}\ \boldsymbol{y}] = \operatorname{rang}\boldsymbol{A}$ die maximale Anzahl linear unabhängiger Gleichungen in (7.1).

**Die Lösungsmenge als affiner Unterraum** Es sei $\boldsymbol{x}_0 \in \mathbb{K}^n$ eine beliebige Lösung von (7.1), dann gilt wegen (3.2)

$$\{\boldsymbol{x} \in \mathbb{K}^n : \ \boldsymbol{Ax} = \boldsymbol{y}\} = \boldsymbol{x}_0 + \{\boldsymbol{x} \in \mathbb{K}^n : \ \boldsymbol{Ax} = \boldsymbol{0}\},$$

d.h. die Lösungsmenge von (7.1) ist, wenn sie nicht leer ist, gleich dem affinen Unterraum $\boldsymbol{x}_0 + \ker\boldsymbol{A}$. Insbesondere gilt wegen (1.4) und (5.4): Wenn eine Lösung von (7.1) existiert, so ist die Dimension der Lösungsmenge gleich

$$\dim\{\boldsymbol{x} \in \mathbb{K}^n : \ \boldsymbol{Ax} = \boldsymbol{y}\} = n - \operatorname{rang}\boldsymbol{A},$$

d.h. die maximale Anzahl linear unabhängiger Lösungen von (7.1) ist gleich der Differenz aus der Anzahl der Unbekannten und der maximalen Anzahl linear unabhängiger Gleichungen.

**$m = n$ und $\det \boldsymbol{A} \neq \boldsymbol{0}$** In diesem Fall gilt (7.2) für jedes $y \in \mathbb{K}^n$, und folglich ist (7.1) für jede rechte Seite $\boldsymbol{y}$ lösbar. Die Lösungsmenge besitzt die Dimension $n - \operatorname{rang} \boldsymbol{A} = 0$, d.h. die Lösung ist eindeutig. Mit anderen Worten: Wenn die homogene Gleichung $\boldsymbol{Ax} = \boldsymbol{0}$ nur die Lösung $\boldsymbol{x} = \boldsymbol{0}$ besitzt, so besitzt die inhomogene Gleichung $\boldsymbol{Ax} = \boldsymbol{y}$ für jede rechte Seite genau eine Lösung, nämlich, nach der sogenannten Cramerschen Regel,

$$x_j = \frac{1}{\det \boldsymbol{A}} \det \begin{bmatrix} a_{11} & \dots & a_{1j-1} & y_1 & a_{1j+1} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{nj-1} & y_n & a_{nj+1} & \dots & a_{nn} \end{bmatrix} \quad (j = 1, \dots, n).$$

Die Lösung $\boldsymbol{x}$ hängt glatt von $\boldsymbol{A}$ und $\boldsymbol{y}$ ab.

**$m = n$ und $\det \boldsymbol{A} = \boldsymbol{0}$** In diesem Fall besitzt (7.1) nur für spezielle rechte Seiten $\boldsymbol{y}$ eine Lösung, nämlich nur für solche, die (7.2) erfüllen. Die Dimension des Unterraums dieser $\boldsymbol{y}$ ist $\operatorname{rang} \boldsymbol{A}$, also in diesem Fall kleiner als $n$. Die Dimension der Lösungsmenge ist $n - \operatorname{rang} \boldsymbol{A}$, in diesem Fall also positiv.

**$m < n$ und $\operatorname{rang} \boldsymbol{A} = m$** In diesem Fall ist (7.2) für jedes $\boldsymbol{y} \in \mathbb{K}^m$ erfüllt, und folglich besitzt (7.1) für jede rechte Seite $\boldsymbol{y}$ eine Lösung. Die Lösungsmenge besitzt die Dimension $n - m$.

**$m < n$ und $\operatorname{rang} \boldsymbol{A} < m$** In diesem Fall besitzt (7.1) wiederum nur für einige spezielle rechten Seiten $\boldsymbol{y}$ eine Lösung, und die Dimension der Lösungsmenge ist dann $n - \operatorname{rang} \boldsymbol{A}$, also größer als $n - m$. Die Dimension des Unterraums der $y \in \mathbb{K}^m$, die (7.2) erfüllen, ist wieder $\operatorname{rang} \boldsymbol{A}$, also kleiner als $m$.

**$m > n$ und $\operatorname{rang} \boldsymbol{A} = n$** In diesem Fall besitzt (7.1) ebenfalls nicht für alle rechten Seiten $\boldsymbol{y}$ eine Lösung. Die Lösungsmenge besitzt dann aber die Dimension $n - \operatorname{rang} \boldsymbol{A} = 0$, d.h. die Lösung ist eindeutig. Die Dimension der Menge der $\boldsymbol{y} \in \mathbb{K}^m$, die (7.2) erfüllen, ist $\operatorname{rang} \boldsymbol{A} = n$.

**$m > n$ und $\operatorname{rang} \boldsymbol{A} < n$** Die Dimension des Unterraums der $\boldsymbol{y} \in \mathbb{K}^m$, für die (7.1) lösbar ist, d.h. die (7.2) erfüllen, ist in diesem Fall kleiner als $n$, und die Dimension der Lösungsmenge von (7.1) mit solchen rechten Seiten ist $n - \operatorname{rang} \boldsymbol{A}$, also positiv.

**Störungsbetrachtungen** Die Fälle mit $\operatorname{rang} \boldsymbol{A} = \min\{m, n\}$ und nur diese haben die Eigenschaft, daß sich das Lösungsverhalten von (7.1) ($m = n$: genau eine Lösung; $m < n$: Lösungsmenge der Dimension $n - m$; $m > n$: keine Lösung) nicht ändert, wenn man $\boldsymbol{A}$ und $\boldsymbol{y}$ geringfügig ändert. Alle anderen Fälle haben in gewissem Sinn die gegenteilige Eigenschaft: "Fast jede" kleine Änderung von $\boldsymbol{A}$ und $\boldsymbol{y}$ ändert das Lösungsverhalten (und zwar zu einem der Fälle mit $\operatorname{rang} \boldsymbol{A} = \min\{m, n\}$). In diesem Sinn kann man formulieren: Wenn die Zahl der Gleichungen gleich bzw. kleiner bzw. größer als die Zahl der Unbekannten ist, so besitzt "fast jedes" lineare Gleichungssystem genau eine bzw. unendlich viele bzw. keine Lösung.
Der "seltenste" Fall ist $\operatorname{rang} \boldsymbol{A} = \operatorname{rang} [\boldsymbol{A}\ \boldsymbol{y}] < n < m$: In diesem Fall existieren unendlich viele verschiedene Lösungen, obwohl mehr Gleichungen als Unbekannte auftreten.

**Die Lösungsmenge als Schnittmenge von affinen Hyperebenen** Für jedes $i$ möge ein $j$ existieren, so daß $a_{ij} \neq 0$ ist, d.h. keine Zeile von $\boldsymbol{A}$ möge nur aus Nullen bestehen. Dann sind die

Dimensionen der Unterräume

$$\{(x_1,\dots,x_n)\in\mathbb{K}^n:\ a_{i1}x_1+\cdots a_{in}x_n=0\}=(\operatorname{span}\{(a_{i1},\dots,a_{in})\})^\perp,\ \ i=1,\cdots,m$$

und folglich der affinen Unterräume

$$\{(x_1,\dots,x_n):\ a_{i1}x_1+\cdots a_{in}x_n=y_i\}=\frac{y_i}{a_{i1}^2+\cdots+a_{in}^2}(a_{i1},\dots,a_{in})+(\operatorname{span}\{(a_{i1},\dots,a_{in})\})^\perp \quad (7.3)$$

gleich $n-1$. Solche Unterräume bzw. affinen Unterräume nennt man Hyperebenen bzw. affine Hyperebenen in $\mathbb{K}^n$.
Die Lösungsmenge von (7.1) ist gleich dem Durchschnitt der affinen Unterräume (7.3). Wegen (1.6) ist Dimension der Lösungsmenge von (7.1) folglich nicht kleiner als $n-m$. Durch geeignete Wahl von $\boldsymbol{A}$ und $\boldsymbol{y}$ kann man mit Hilfe von (7.3) jedes Tupel von $m$ affinen Unterräumen in $\mathbb{K}^n$ der Dimension $n-1$ erzeugen. Deshalb kann man die obigen Resultate auch wie folgt geometrisch formulieren:
Es seien Hyperebenen $U_1,\dots,U_m$ in $\mathbb{K}^n$ gegeben, dann gilt: Wenn $m\le n$ und $\dim(U_1\cap\cdots\cap U_m)=n-m$ ist, so folgt

$$\dim\Big((\boldsymbol{v}_1+U_1)\cap\cdots\cap(\boldsymbol{v}_m+U_m)\Big)=n-m \text{ für alle } \boldsymbol{v}_1,\dots,\boldsymbol{v}_m\in\mathbb{R}^n.$$

Wenn dagegen $m>n$ oder $\dim(U_1\cap\cdots\cap U_m)>n-m$ gilt, so ist

$$(\boldsymbol{v}_1+U_1)\cap\cdots\cap(\boldsymbol{v}_m+U_m)=\emptyset \text{ für “fast alle” } \boldsymbol{v}_1,\cdots,\boldsymbol{v}_m\in\mathbb{R}^n.$$

**Übungsaufgabe** Verifizieren Sie die obigen Aussagen in den Fällen $m=n=2$ bzw. $m=2$, $n=3$ bzw. $m=3$, $n=2$.

# 8 Eigenwerte und Eigenvektoren

Es sei $\boldsymbol{A}\in\mathbb{M}(n\times n;\mathbb{C})$, $\lambda\in\mathbb{C}$ und $\boldsymbol{v}\in\mathbb{C}^n$ mit

$$\boldsymbol{A}\boldsymbol{v}=\lambda\boldsymbol{v},\ \ \boldsymbol{v}\neq\boldsymbol{0}.$$

Dann heißen $\lambda$ bzw. $\boldsymbol{v}$ Eigenwert bzw. Eigenvektor von $\boldsymbol{A}$. Die Menge aller Eigenwerte von $\boldsymbol{A}$ wird mit $\operatorname{spec}\boldsymbol{A}$ bezeichnet und heißt Spektrum von $\boldsymbol{A}$. Die gleiche Terminologie kann man für lineare Abbildungen $\boldsymbol{F}: V\to V$ auf einem beliebigen Vektorraum $V$ einführen. Falls $V$ endlich-dimensional ist, gelten dann Analoga zu den fogenden Resultaten.

**Charakteristisches Polynom** Eine Zahl $\lambda$ ist Eigenwert von $\boldsymbol{A}\in\mathbb{M}(n\times n;\mathbb{C})$ genau dann, wenn

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=0$$

ist. Nach dem Hauptsatz der Algebra besteht folglich $\operatorname{spec}\boldsymbol{A}$ aus mindestens einer und höchsten $n$ verschiedenen komplexen Zahlen.

**Geometrische und algebraische Vielfachheit, Eigenraum und verallgemeinerter Eigenraum** Es sei $\lambda$ Eigenwert von $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$.
(i) Der Unterraum $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})$ aller Eigenvektoren zu $\lambda$ heißt Eigenraum zu $\lambda$, und seine Dimension heißt geometrische Vielfachheit von $\lambda$. Mit anderen Worten: Die geometrische Vielfachheit eines Eigenwertes ist die maximale Anzahl linear unabhängiger Eigenvektoren zu diesem Eigenwert.
(ii) Es existiert ein $l \in \{1, \cdots, n\}$, so daß gilt

$$\begin{aligned} \operatorname{rang}(\boldsymbol{A} - \lambda \boldsymbol{I})^{k-1} &> \operatorname{rang}(\boldsymbol{A} - \lambda \boldsymbol{I})^{k} \text{ für } k = 1, \dots, l, \\ \operatorname{rang}(\boldsymbol{A} - \lambda \boldsymbol{I})^{k} &= \operatorname{rang}(\boldsymbol{A} - \lambda \boldsymbol{I})^{k+1} \text{ für } k \geq l. \end{aligned}$$

Der Unterraum $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})^l$ heißt verallgemeinerter Eigenraum von $\lambda$, und seine Dimension heißt algebraische Vielfachheit von $\lambda$. Die algebraische Vielfachheit von $\lambda$ ist gleich der Vielfachheit von $\lambda$ als Nullstelle des charakteristischen Polynoms, d.h. es gilt $a = \dim \ker \ker(\boldsymbol{A} - \lambda \boldsymbol{I})^l$ genau dann, wenn

$$\frac{d}{d\mu}(\boldsymbol{A} - \mu \boldsymbol{I})|_{\mu=\lambda} = \cdots = \left(\frac{d}{d\mu}\right)^{a-1}(\boldsymbol{A} - \mu \boldsymbol{I})|_{\mu=\lambda} = 0 \text{ und } \left(\frac{d}{d\mu}\right)^{a}(\boldsymbol{A} - \mu \boldsymbol{I})|_{\mu=\lambda} \neq 0.$$

**Halbeinfache und einfache Eigenwerte** Es sei $\lambda$ Eigenwert von $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$. Die geometrische Vielfachheit von $\lambda$ ist gleich der algebraischen Vielfachheit (d.h. der Eigenraum ist gleich dem verallgemeinerten Eigenraum) genau dann, wenn

$$\operatorname{rang}(\boldsymbol{A} - \lambda \boldsymbol{I}) = \operatorname{rang}(\boldsymbol{A} - \lambda \boldsymbol{I})^2.$$

In diesem Fall heißt $\lambda$ halbeinfacher Eigenwert. Wenn insbesondere die algebraische (und folglich auch die geometrische) Vielfachheit gleich Eins ist, so heißt $\lambda$ einfach.

**Faktorzerlegung des charakteristischen Polynoms** Es seien $\lambda_1, \dots, \lambda_m$ die Eigenwerte einer Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$, und die entsprechenden algebraischen Vielfachheiten seien $a_1, \dots, a_m$. Dann gilt $a_1 + \cdots + a_m = n$ und

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = (\lambda_1 - \lambda)^{a_1}(\lambda_2 - \lambda)^{a_2} \cdots (\lambda_m - \lambda)^{a_m}.$$

**Zerlegung des $\mathbb{C}^n$ mit Hilfe der verallgemeinerten Eigenräume** Es seien $\lambda_1, \dots, \lambda_m$ die Eigenwerte einer Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$, und die entsprechenden algebraischen Vielfachheiten seien $a_1, \dots, a_m$. Dann sind $\ker(\boldsymbol{A} - \lambda_i \boldsymbol{I})^{a_i}$ die entsprechenden verallgemeinerten Eigenräume ($i = 1, \dots, m$), und es gilt

$$\mathbb{C}^n = \ker(\boldsymbol{A} - \lambda_1 \boldsymbol{I})^{a_1} \oplus \ker(\boldsymbol{A} - \lambda_2 \boldsymbol{I})^{a_2} \oplus \cdots \oplus \ker(\boldsymbol{A} - \lambda_m \boldsymbol{I})^{a_m}.$$

Insbesondere sind Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig.

**Matrizen mit ausschließlich einfachen Eigenwerten** Eine Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$ besitzt ausschließlich einfache Eigenwerte genau dann, wenn sie $n$ verschiedene Eigenwerte besitzt. Diese Eigenschaft besitzen in gewissem Sinn "fast alle" Matrizen, es gilt nämlich:
(i) Wenn $\boldsymbol{A}$ ausschließlich einfache Eigenwerte besitzt, so gilt für alle $\boldsymbol{B} \in \mathbb{M}(n \times n; \mathbb{C})$, die "hinreichend nahe" bei $\boldsymbol{A}$ liegen, daß auch $\boldsymbol{B}$ ausschließlich einfache Eigenwerte besitzt.

(ii) Wenn $\boldsymbol{A}$ nicht ausschließlich einfache Eigenwerte besitzt, so existieren Matrizen $\boldsymbol{B} \in \mathbb{M}(n \times n; \mathbb{C})$, die "beliebig nahe" bei $\boldsymbol{A}$ liegen und die ausschließlich einfache Eigenwerte besitzten.

**Störungsverhalten einfacher Eigenwerte** Es sei $\lambda$ ein einfacher Eigenwert von $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$, und $\boldsymbol{B} \in \mathbb{M}(n \times n; \mathbb{C})$ sei beliebig fixiert. Dann existieren positive Zahlen $\epsilon_0$ und $\delta$, so daß für alle $\epsilon \in (-\epsilon_0, \epsilon_0)$ genau ein Eigenwert $\lambda_\epsilon$ von $\boldsymbol{A} + \epsilon \boldsymbol{B}$ mit $|\lambda_\epsilon - \lambda| < \delta$ existiert. Dabei gilt $\lambda_\epsilon \to \lambda$ für $\epsilon \to 0$. Mit anderen Worten: Einfache Eigenwerte hängen stetig von der entsprechenden Matrix ab. Das Störungsverhalten von Eigenwerten, die nicht einfach sind, ist erheblich komplizierter.

**Matrizen mit ausschließlich halbeinfachen Eigenwerten und Hauptachsentransformation** Eine Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$ besitzt ausschließlich halbeinfache Eigenwerte genau dann, wenn in $\mathbb{C}^n$ eine Basis existiert, die aus Eigenvektoren von $\boldsymbol{A}$ besteht.
Genauer gesagt gilt folgendes: Es seien $\lambda_1, \ldots, \lambda_m$ die Eigenwerte von $\boldsymbol{A}$, und alle seien halbeinfach. Ferner sei $g_j$ die geometrische (und algebraische) Vielfachheit von $\lambda_j$, und $\boldsymbol{v}_{j1}, \ldots \boldsymbol{v}_{jg_j}$ sei eine Basis im zu $\lambda_j$ gehörenden Eigenraum $\ker(\boldsymbol{A} - \lambda_j \boldsymbol{I})$ $(j = 1, \cdots, m)$. Dann ist $\boldsymbol{v}_{11}, \ldots, \boldsymbol{v}_{1g_1}, \boldsymbol{v}_{21}, \ldots, \boldsymbol{v}_{mg_m}$ eine Basis in $\mathbb{C}^n$. Insbesondere gilt $n = g_1 + \cdots + g_m$ und

$$\mathbb{C}^n = \ker(\boldsymbol{A} - \lambda_1 \boldsymbol{I}) \oplus \cdots \oplus \ker(\boldsymbol{A} - \lambda_m \boldsymbol{I}). \tag{8.1}$$

Die Matrix-Darstellung von $\boldsymbol{A}$ in dieser Basis ist eine Diagonalmatrix, auf deren Diagonalen die Eigenwerte von $\boldsymbol{A}$ (so oft, wie ihre Vielfachheit ist, und in der entsprechenden Reihenfolge) stehen, d.h. es gilt

$$\boldsymbol{CAC}^{-1} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \lambda_1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \ldots & 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & \lambda_m \end{bmatrix}, \tag{8.2}$$

wobei $\boldsymbol{C}$ die Matrix des Basiswechsels von der kanonischen Basis zur Eigenvektor-Basis ist. Den Übergang von der Matrix $\boldsymbol{A}$ zur Diagonalmatrix $\boldsymbol{CAC^{-1}}$ nennt man auch Hauptachsentransformation von $\boldsymbol{A}$. Falls $\mathbb{K} = \mathbb{R}$ ist und alle $\lambda_j$ reell sind, so besagt die Darstellung (8.2), daß $A$ die Geraden $\operatorname{span}\{\boldsymbol{C^{-1}e}_j\}$ auf sich abbildet und dabei um den Faktor $\lambda_j$ streckt.

**Hermitesche Matrizen** Eine Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$ heißt hermitesch (oder selbstadjungiert), wenn $\boldsymbol{A} = \boldsymbol{A}^*$ gilt. Die Eigenwerte hermitescher Matrizen sind alle halbeinfach und reell, und die direkten Summen in (8.1) sind alles orthogonale Summen. Mit anderen Worten: Es existiert eine Orthonormalbasis in $\mathbb{C}^n$, bestehend aus Eigenvektoren von $\boldsymbol{A}$, so daß die Matrix-Darstellung von $\boldsymbol{A}$ bzgl. dieser Basis eine Diagonalmatrix ist (mit den Eigenwerten von $\boldsymbol{A}$ auf der Diagonalen). In diesem Fall gilt

$$\lambda_- \|\boldsymbol{v}\|^2 \leq \boldsymbol{Av} \cdot \boldsymbol{v} \leq \lambda_+ \|\boldsymbol{v}\|^2 \text{ für alle } \boldsymbol{v} \in \mathbb{C}^n.$$

Dabei sind $\lambda_-$ bzw. $\lambda_+$ der kleinste bzw. der größte Eigenwert von $\boldsymbol{A}$.

**Matrizen mit nicht halbeinfachen Eigenwerten und beigeordnete Eigenvektoren** Es sei $\lambda$ Eigenwert von $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$ mit der geometrischen Vielfachheit $g$ und der algebraischen

Vielfachheit $a$, und $\lambda$ sei nicht halbeinfach, d.h. $g < a$. Ferner sei $\boldsymbol{v}_{11}, \ldots, \boldsymbol{v}_{g1}$ eine Basis im zu $\lambda$ gehörenden Eigenraum $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})$. Dann gilt: Diese Basis kann ergänzt werden zu einer Basis

$$\boldsymbol{v}_{11}, \boldsymbol{v}_{12}, \ldots, \boldsymbol{v}_{1l_1}, \boldsymbol{v}_{21}, \ldots, \boldsymbol{v}_{gl_g}$$

im zu $\lambda$ gehörenden verallgemeinerten Eigenraum $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})^l$. Dabei gilt

$$\boldsymbol{A}\boldsymbol{v}_{j1} = \lambda \boldsymbol{v}_{j1} \text{ und } \boldsymbol{A}\boldsymbol{v}_{jk} = \lambda \boldsymbol{v}_{jk} + \boldsymbol{v}_{jk-1} \text{ für } j = 1, \ldots, g \text{ und } k = 2, \ldots, l_j. \tag{8.3}$$

Die Vektoren $\boldsymbol{v}_{j2}, \ldots, \boldsymbol{v}_{jl_j}$ heißen beigeordnete Eigenvektoren zum Eigenvektor $\boldsymbol{v}_{j1}$, und $\boldsymbol{v}_{j1}, \boldsymbol{v}_{j2}, \ldots, \boldsymbol{v}_{jl_j}$ heißt Jordan-Kette zu $\boldsymbol{v}_{j1}$. Insbesondere gilt $l = \max\{l_1, \ldots, l_g\}$ und $l_1 + \cdots + l_g = a$.

**Jordan-Basis** Es seien $\lambda_i$ die Eigenwerte von $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{C})$ mit den geometrischen Vielfachheiten $g_i$ und den algebraischen Vielfachheiten $a_i$ $(i = 1, \ldots, m)$. Ferner seien $\boldsymbol{v}_{ij1}$ $(j = 1, \ldots, g_i)$ Basen in den Eigenräumen $\ker(\boldsymbol{A} - \lambda_i \boldsymbol{I})$, und $\boldsymbol{v}_{ijk}$ $(k = 1, \ldots, l_{ij})$ seien Jordan-Ketten zu den Eigenvektoren $\boldsymbol{v}_{ij1}$. Dann bilden die Vektoren aller dieser Jordan-Ketten gemeinsam eine Basis (eine sogenannte Jordan-Basis) in $\mathbb{C}^n$, insbesondere gilt

$$\sum_{i=1}^{m} \sum_{j=1}^{g_i} l_{ij} = n.$$

Die Matrix-Darstellung von $\boldsymbol{A}$ bzgl. dieser Basis, angeordnet in der Reihenfolge

$$\boldsymbol{v}_{111}, \boldsymbol{v}_{112}, \ldots, \boldsymbol{v}_{11l_{11}}, \boldsymbol{v}_{121}, \boldsymbol{v}_{122}, \ldots, \boldsymbol{v}_{12l_{12}}, \ldots, \boldsymbol{v}_{mg_m1}, \boldsymbol{v}_{mg_m2}, \ldots, \boldsymbol{v}_{mg_ml_{mg_m}},$$

kann mit Hilfe von (8.3) leicht konstruiert werden. Sie besitzt eine sogenannte Block-Diagonalstruktur, d.h. alle ihre Koeffizienten sind Null außer denen, die in gewissen quadratischen Blöcken entlang der Diagonale von $A$ liegen. Diese Blöcke sind die sogenannten Jordan-Kästchen

$$\begin{bmatrix} \lambda_1 & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_1 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_1 & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_1 \end{bmatrix},$$

Es treten $g_1 + \cdots + g_m$ Jordan-Kästchen auf, nämlich zu jedem Basis-Vektor $v_{ij1}$ eins (mit der Seitenlänge $l_{ij}$). Diese Matrix-Darstellung von $\boldsymbol{A}$ heißt Jordansche Normalform von $\boldsymbol{A}$. Sie ist, bis auf die Reihenfolge der Jordan-Kästchen, eindeutig durch $\boldsymbol{A}$ bestimmt, obwohl bei der Wahl der Jordan-Basis sehr viele Freiheiten existieren: Die Basen in den Eigenräumen können beliebig gewählt werden, und jeder beigeordnete Eigenvektor $\boldsymbol{w}$ zu einem Eigenvektor $\boldsymbol{v}$ kann durch $\boldsymbol{v} + \boldsymbol{w}$ ersetzt werden.

**Reelle Matrizen** Es sei $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ eine Matrix mit reellen Koeffizienten. Dann können trotzdem einige oder sogar alle ihrer Eigenwerte nicht reell sein.
Ein reeller Eigenvektor $\boldsymbol{v} \in \mathbb{R}^n$ zu einem Eigenwert $\lambda$ von $\boldsymbol{A}$ existiert genau dann, wenn $\lambda$ reell ist. Dabei sind die Dimensionen von $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})$ und $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})^l$, aufgefaßt als Unterräume von $\mathbb{R}^n$, gleich den Dimensionen von $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})$ und $\ker(\boldsymbol{A} - \lambda \boldsymbol{I})^l$, aufgefaßt als Unterräume von $\mathbb{C}^n$. Mit

anderen Worten: Die geometrischen und die algebraischen Vielfachheiten reeller Eigenwerte reeller Matrizen können "in $\mathbb{R}^n$" bestimmt werden.

**Konjugiert-komplexe Eigenwertpaare reeller Matrizen** Wenn $\lambda$ Eigenwert von $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ ist, so ist auch $\overline{\lambda}$ Eigenwert von $\boldsymbol{A}$, und die entsprechenden geometrischen und algebraischen Vielfachheiten sind gleich. Wenn ferner $\boldsymbol{v} \in \mathbb{C}^n$ ein Eigenvektor zum Eigenwert $\lambda$ ist, so ist auch $\overline{\boldsymbol{v}}$ ein Eigenvektor zu dem Eigenwert $\overline{\lambda}$. Wenn schließlich $\boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l$ eine Jordan-Kette zu $\boldsymbol{v}$ ist, so ist auch $\overline{\boldsymbol{v}}, \overline{\boldsymbol{v}_1}, \overline{\boldsymbol{v}_2}, \ldots, \overline{\boldsymbol{v}_l}$ eine Jordan-Kette zu $\overline{\boldsymbol{v}}$.

**Beispiel** Das charakteristische Polynom der reellen $2 \times 2$-Matrix

$$\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

ist

$$\det \begin{bmatrix} a - \lambda & b \\ c & d - \lambda \end{bmatrix} = \lambda^2 - (a + d)\lambda + ad - bc.$$

Die beiden Eigenwerte

$$\lambda_{1,2} = \frac{1}{2}(a + d \pm \sqrt{(a - d)^2 + 4bc}) \tag{8.4}$$

bilden ein konjugiert-komplexes Eigenwertpaar, wenn $(a - d)^2 < -4bc$ ist. Andernfalls sind sie reell. Zugehörige Eigenvektoren $\boldsymbol{v}_{1,2}$ müssen der Gleichung

$$(\boldsymbol{A} - \lambda_{1,2}\boldsymbol{I})\boldsymbol{v}_{1,2} = \begin{bmatrix} \frac{1}{2}(a - d \mp \sqrt{(a - d)^2 + 4bc}) & b \\ c & \frac{1}{2}(d - a \mp \sqrt{(a - d)^2 + 4bc}) \end{bmatrix} \boldsymbol{v}_{1,2} = \boldsymbol{0}$$

genügen. Folglich sind z.B.

$$\boldsymbol{v}_{1,2} = (2b, d - a \pm \sqrt{(a - d)^2 + 4bc})$$

zugehörige Eigenvektoren, falls $b \neq 0$ ist. Im Fall $b = 0$ sind z.B.

$$\boldsymbol{v}_1 = \begin{cases} (0, 1) & \text{falls } a < d, \\ (2(a - d), c) & \text{falls } a > d, \end{cases}$$

bzw.

$$\boldsymbol{v}_2 = \begin{cases} (2(a - d), c) & \text{falls } a < d, \\ (0, 1) & \text{falls } a > d, \end{cases}$$

zu $\lambda_1$ bzw. $\lambda_2$ gehörende Eigenvektoren. Wenn $(a - d)^2 \neq -4bc$ ist, so sind die beiden Eigenwerte (8.4) verschieden und folglich einfach.
Es sei nun

$$(a - d)^2 = -4bc. \tag{8.5}$$

Dann besitzt $\boldsymbol{A}$ genau einen Eigenwert

$$\lambda_1 = \lambda_2 = \frac{1}{2}(a + d). \tag{8.6}$$

Er ist genau dann halbeinfach, wenn der Rang von

$$\boldsymbol{A} - \sqrt{-bc}\boldsymbol{I} = \begin{bmatrix} \sqrt{-bc} & b \\ c & -\sqrt{-bc} \end{bmatrix}$$

gleich dem Rang von $(\boldsymbol{A}-\sqrt{-bc}\boldsymbol{I})^2$, also Null, ist. Folglich besitzt $\boldsymbol{A}$ genau dann einen nicht einfachen, aber halbeinfachen Eigenwert, wenn $a = d$ und $b = c = 0$ (und dann sind alle $\boldsymbol{v} \in \mathbb{R}^2$ Eigenvektoren). In diesem Sinn gilt für "fast alle" reellen $2\times 2$-Matrizen, die einen nicht einfachen Eigenwert besitzen, daß dieser auch nicht halbeinfach ist. Ein analoges Resultat existiert auch für $n \times n$-Matrizen.
Es sei schließlich (8.5) erfüllt, aber $b \neq 0$ oder $c \neq 0$, d.h. der Eigenwert (8.6) sei nicht halbeinfach. In diesem Fall müssen Eigenvektoren $\boldsymbol{v}$ bzw. diesen beigeordnete Eigenvektoren $\boldsymbol{w}$ den Gleichungen $(\boldsymbol{A} - \sqrt{-bc}\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}$ bzw. $(\boldsymbol{A} - \sqrt{-bc}\boldsymbol{I})\boldsymbol{w} = \boldsymbol{v}$ genügen. Lösungen sind z.B.

$$\boldsymbol{v} = \begin{cases} (b, -\sqrt{-bc}) & \text{falls } b \neq 0, \\ (0, c) & \text{falls } b = 0, \end{cases}$$

und

$$\boldsymbol{w} = \begin{cases} (0, 1) & \text{falls } b \neq 0, \\ (1, 0) & \text{falls } b = 0. \end{cases}$$

Die Vektoren $\boldsymbol{v}$ und $\boldsymbol{w}$ bilden dann eine Jordan-Basis in $\mathbb{R}^2$. Eine Basis in $\mathbb{R}^2$, die nur aus Eigenvektoren besteht, existiert in diesem Fall nicht.

**Reelle Matrizen mit ausschließlich einfachen Eigenwerten** Es seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$, alle Eigenewerte seien einfach, die nicht rellen Eigenwerte seien $\lambda_1, \ldots, \lambda_{2m}$, und sie seien folgendermaßen zu konjugiert-komplexen Paaren angeordnet: $\lambda_1 = \overline{\lambda_2}, \ldots, \lambda_{2m-1} = \overline{\lambda_{2m}}$. Ferner seien $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ entsprechende Eigenvektoren, und es gelte $\boldsymbol{v}_1 = \overline{\boldsymbol{v}_2}, \ldots, \boldsymbol{v}_{2m-1} = \overline{\boldsymbol{v}_{2m}}$ und $\boldsymbol{v}_{2m+1}, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$. Dann bilden die Vektoren

$$\operatorname{Re}\boldsymbol{v}_1, \operatorname{Im}\boldsymbol{v}_1, \operatorname{Re}\boldsymbol{v}_3, \ldots, \operatorname{Im}\boldsymbol{v}_{2m-1}, \boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2}, \ldots, \boldsymbol{v}_n$$

eine Basis in $\mathbb{R}^n$, und bzgl. dieser Basis besitzt $\boldsymbol{A}$ folgende Matrix-Darstellung:

$$\begin{bmatrix}
\operatorname{Re}\lambda_1 & -\operatorname{Im}\lambda_1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
\operatorname{Im}\lambda_1 & \operatorname{Re}\lambda_1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
0 & 0 & \operatorname{Re}\lambda_3 & -\operatorname{Im}\lambda_3 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
0 & 0 & \operatorname{Im}\lambda_3 & \operatorname{Re}\lambda_3 & \ldots & 0 & 0 & 0 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \ldots & \operatorname{Re}\lambda_{2m-1} & -\operatorname{Im}\lambda_{2m-1} & 0 & 0 & \ldots & 0 \\
0 & 0 & 0 & 0 & \ldots & \operatorname{Im}\lambda_{2m-1} & \operatorname{Re}\lambda_{2m-1} & 0 & 0 & \ldots & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & \lambda_{2m+1} & 0 & \ldots & 0 \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & \lambda_{2m+2} & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 & \ldots & \lambda_n
\end{bmatrix}.$$

Wenn $m > 0$ ist, existiert allerdings keine Basis in $\mathbb{R}^n$, bzgl. derer die Matrix- Darstellung von $\boldsymbol{A}$ Diagonalgestalt hat.

Im Fall reeller Matrizen mit ausschließlich halbeinfachen Eigenwerten existiert ein analoges Resultat. Der Fall reeller Matrizen mit möglicherweise nicht halbeinfachen Eigenwerten ist, wie schon bei komplexen Matrizen, komplizierter.

**Reelle symmetrische Matrizen** Wenn die relle Matrix $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ symmetrisch ist, d.h. wenn $\boldsymbol{A} = \boldsymbol{A}^T$ gilt, so existiert eine Orthonormalbasis in $\mathbb{R}^n$, bestehend aus Eigenvektoren von $\boldsymbol{A}$, so daß die Matrix-Darstellung von $\boldsymbol{A}$ bzgl. dieser Basis eine Diagonalmatrix ist (mit den Eigenwerten von $\boldsymbol{A}$ auf der Diagonalen). Die Matrix $\boldsymbol{C}$ in (8.2) ist dann orthogonal, d.h. $\boldsymbol{C}^{-1} = \boldsymbol{C}^T$, und $\boldsymbol{A}$ ist eine Streckung in die zueinander orthogonalen Richtungen, die durch die Orthonormalbasis aus Eigenvektoren gegeben sind, jeweils um den Faktor, der durch den entsprechenden Eigenwert gegeben ist.

**Anwendungsbeispiel: Quadriken (Kurven und Flächen zweiter Ordnung)** Es seien $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ eine symmetrische Matrix, $\boldsymbol{a} \in \mathbb{R}^n$ ein Vektor und $\alpha \in \mathbb{R}$ eine Zahl. Dann heißt die Menge aller Punkte $\boldsymbol{x} \in \mathbb{R}^n$, die die Gleichung

$$\boldsymbol{Ax} \cdot \boldsymbol{x} + \boldsymbol{a} \cdot \boldsymbol{x} + \alpha = \boldsymbol{0}$$

erfüllen, zu $\boldsymbol{A}$, $\boldsymbol{a}$ und $\alpha$ gehörende Quadrik.
Es sei $\det \boldsymbol{A} \neq 0$. Dann gilt

$$\boldsymbol{A}(\boldsymbol{y} - \frac{1}{2}\boldsymbol{A}^{-1}\boldsymbol{a}) \cdot (\boldsymbol{y} - \frac{1}{2}\boldsymbol{A}^{-1}\boldsymbol{a}) + \boldsymbol{a} \cdot (\boldsymbol{y} - \frac{1}{2}\boldsymbol{A}^{-1}\boldsymbol{a}) + \alpha = \boldsymbol{Ay} \cdot \boldsymbol{y} + \alpha - \frac{1}{4}\boldsymbol{Aa} \cdot \boldsymbol{a}.$$

Mit Hilfe dieser Darstellung läßt sich die Form der Quadrik leicht diskutieren.
Es sei z.B. $n = 2$, und $\lambda_1$, $\lambda_2$ seien die Eigenwerte von $\boldsymbol{A}$. Ferner seien $\boldsymbol{v}_1$ und $\boldsymbol{v}_1$ zugehörige Eigenvektoren mit $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0$ und $\|\boldsymbol{v}_1\| = \|\boldsymbol{v}_2\| = 1$. Wenn nun die Zahlen $\lambda_1$, $\lambda_2$ und $\frac{1}{4}\boldsymbol{Aa} \cdot \boldsymbol{a} - \alpha$ alle positiv oder alle negativ sind, so ist die Quadrik eine Ellipse (im Spezialfall ein Kreis), deren Zentrum im Punkt $-1/2\boldsymbol{A}^{-1}\boldsymbol{a}$ liegt, und deren Hauptachsen $\lambda_1^{-1/2}\boldsymbol{v}_1$ und $\lambda_2^{-1/2}\boldsymbol{v}_2$ sind. Wenn die Zahlen $\lambda_1$, $\lambda_2$ gleiches Vorzeichen haben, aber $\frac{1}{4}\boldsymbol{Aa} \cdot \boldsymbol{a} - \alpha$ das entgegengesetzte Vorzeichen besitzt (bzw. Null ist), so ist die Quadrik leer (bzw. besteht aus einem Punkt). Wenn die Zahlen $\lambda_1$, $\lambda_2$ unterschiedliches Vorzeichen haben, und $\frac{1}{4}\boldsymbol{Aa} \cdot \boldsymbol{a} - \alpha$ nicht Null ist (bzw. Null ist), so ist die Quadrik eine Hyperbel (bzw. besteht aus zwei nichtparallelen Geraden).
Im Fall $n = 3$ sind natürlich mehr verschiedene Quadrikformen zu unterscheiden. Hier treten u.a. Ellipsoide (wenn alle Eigenwerte und $\frac{1}{4}\boldsymbol{Aa} \cdot \boldsymbol{a} - \alpha$ gleiches Vorzeichen besitzen) und einschalige und zweischalige Hyperboloide (wenn die Eigenwerte unterschiedliche Vorzeichen besitzen und $\frac{1}{4}\boldsymbol{Aa} \cdot \boldsymbol{a} - \alpha \neq 0$ ist) auf.
Etwas komplizierter ist der Fall $\det \boldsymbol{A} = 0$, d.h. wenn mindestens ein Eigenwert von $\boldsymbol{A}$ gleich Null ist. Dann treten u.a. Parabeln (im Fall $n = 2$) und elliptische und hyperbolische Paraboloide (im Fall $n = 3$) auf.

**Positiv definite Matrizen** Eine reelle symmetrische Matrix $[a_{ij}] \in \mathbb{M}(n \times n; \mathbb{R})$ heißt positiv definit, wenn alle ihre Eigenwerte positiv sind. Nach der sogenannten Regel von Sylvester ist das genau dann der Fall, wenn die Determinanten der sogenannten Hauptminore

$$\begin{bmatrix} a_{11} & \dots & a_{1j} \\ \vdots & \ddots & \vdots \\ a_{j1} & \dots & a_{jj} \end{bmatrix} \quad (j = 1, \cdots, n)$$

alle positiv sind. Zum Beispiel ist eine reelle symmetrische $2 \times 2$-Matrix $[a_{ij}]$ positiv definit genau dann, wenn $a_{11} > 0$ und $a_{11}a_{22} - a_{12}^2 > 0$ ist.

**Anwendungsbeispiel: Minima von glatten Funktionen mehrerer Veränderlicher** Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine glatte Funktion, deren erste partielle Ableitungen in einem fixierten Punkt $\boldsymbol{x_0} \in \mathbb{R}^n$ alle verschwinden. Ferner sei die sogenannte Hessesche Matrix

$$\left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0)\right]$$

der zweiten partiellen Ableitungen in diesem Punkt positiv definit. Dann ist $\boldsymbol{x_0}$ ein strenges lokales Minimum von $f$, d.h. es existiert ein $\epsilon > 0$, so daß für alle $\boldsymbol{x} \in \mathbb{R}^n$ mit $0 < \|\boldsymbol{x}\| < \epsilon$ gilt $f(\boldsymbol{x_0} + \boldsymbol{x}) > f(\boldsymbol{x_0})$.

**Die Wurzel aus einer positiv definiten Matrix** Es sei $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ eine positiv definite Matrix. Dann existiert genau eine positiv definite Matrix $\boldsymbol{A}^{1/2} \in \mathbb{M}(n \times n; \mathbb{R})$ (die sogenannte Wurzel aus $\boldsymbol{A}$), so daß $\boldsymbol{A}^{1/2}\boldsymbol{A}^{1/2} = \boldsymbol{A}$ ist. Dabei gilt für jeden Eigenwert $\lambda$ von $\boldsymbol{A}$ und jeden Eigenvektor $\boldsymbol{v}$ zu $\lambda$, daß $\boldsymbol{A}^{1/2}\boldsymbol{v} = \sqrt{\lambda}\boldsymbol{v}$ ist.

**Polarzerlegung** Es sei $\boldsymbol{A} \in \mathbb{M}(n \times n; \mathbb{R})$ eine invertierbare Matrix. Dann sind die Matrizen $\boldsymbol{A}^T\boldsymbol{A}$ bzw. $\boldsymbol{A}\boldsymbol{A}^T$ positiv definit und besitzen folglich Wurzeln $\boldsymbol{U}$ bzw. $\boldsymbol{V}$. Ferner existiert eine orthogonale Matrix $\boldsymbol{Q} \in O(n; \mathbb{R})$, so daß gilt

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{U} = \boldsymbol{V}\boldsymbol{Q}.$$

Wenn $\lambda_i$ die Eigenwerte von $\boldsymbol{U}$ mit Eigenvektoren $\boldsymbol{v_i}$ sind, so sind die $\lambda_i$ auch Eigenwerte von $\boldsymbol{V}$ mit den Eigenvektoren $\boldsymbol{Q}\boldsymbol{v}_i$.
Mit anderen Worten: Wenn man $\boldsymbol{A}$ als "Deformation" des Raumes $\mathbb{R}^n$ auffaßt, so kann man diese Deformation auf zweierlei Art als Nacheinanderausführung zweier "elementarer Deformationen" darstellen. Bei $\boldsymbol{Q}\boldsymbol{U}$ wird zunächst die "reine Streckung" $\boldsymbol{U}$ (Streckung um die Faktoren $\lambda_i$ in die Richtungen $\boldsymbol{v}_i$) und danach die "starre Deformation" $\boldsymbol{Q}$ (Drehung oder Drehung mit Spiegelung) durchgeführt. Bei $\boldsymbol{V}\boldsymbol{Q}$ wird zunächst dieselbe starre Deformation durchgeführt und danach eine reine Streckung mit denselben Streckungsfaktoren, aber anderen Streckungsachsen. Diese Betrachtungsweise wird in der Kontinuumsmechanik benutzt.

# 9 Tensoren

**Multilineare Abbildungen** Es seien $V_1, \ldots, V_n$ und $W$ Vektorräume über $\mathbb{K}$ (überall gleiches $\mathbb{K}$). Eine Abbildung $\boldsymbol{F}$: $V_1 \times \cdots \times V_n \to W$ heißt multilinear, wenn für alle $j = 1, \ldots, n, \boldsymbol{u}_j, \boldsymbol{v}_j \in V_j$ und $\lambda, \mu \in \mathbb{K}$ gilt

$$\begin{aligned}&\boldsymbol{F}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1}, \lambda\boldsymbol{u}_j + \mu\boldsymbol{v}_j, \boldsymbol{v}_{j+1}, \ldots, \boldsymbol{v}_n) = \\ &= \lambda\boldsymbol{F}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1}, \boldsymbol{u}_j, \boldsymbol{v}_{j+1}, \ldots, \boldsymbol{v}_n) + \mu\boldsymbol{F}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1}, \boldsymbol{v}_j, \boldsymbol{v}_{j+1}, \ldots, \boldsymbol{v}_n).\end{aligned}$$

Die Menge aller multilinearen Abbildungen von $V_1 \times \cdots \times V_n$ nach $W$ wird mit $\mathcal{L}(V_1, \ldots, V_n; W)$ bezeichnet. Die Elemente von $\mathcal{L}(V_1, V_2; W)$ heißen bilinear. Die Menge $\mathcal{L}(V_1, \ldots, V_n; W)$ ist ein Vektorraum über $\mathbb{K}$ mit den Operationen

$$\left.\begin{array}{rcl}(\boldsymbol{F}+\boldsymbol{G})(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n) & = & \boldsymbol{F}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)+\boldsymbol{G}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n),\\ (\lambda\boldsymbol{F})(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n) & = & \lambda\boldsymbol{F}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n),\end{array}\right\}\text{ für alle } \boldsymbol{v}_1\in V_1,\dots,\boldsymbol{v}_n\in V_n.$$

Wenn $V_1,\dots,V_n$ und $W$ endlich-dimensional sind, so gilt

$$\dim\mathcal{L}(V_1,\dots,V_n;W)=\dim V_1\cdots\dim V_n\dim W.$$

**Lineare und multilineare Abbildungen** Die im Spezialfall $n=1$ durch die obige Definition erzeugte Menge $\mathcal{L}(V_1;W)$ ist der in Sektion 3 eingeführten Menge $\mathcal{L}(V_1;W)$ der linearen Abbildungen von $V_1$ nach $W$ gleich, d.h. der Begriff der multilinearen Abbildung ist eine natürliche Verallgemeinerung des Begriffs der linearen Abbildung. Im Fall $n>1$ sind lineare und multilineare Abbildungen allerdings sehr verschiedene Objekte: Eine lineare Abbildung von $V_1\times\dots\times V_n$ nach $W$, die nicht die Nullabbildung ist, ist nicht multilinear, und eine multilineare Abbildung von $V_1\times\dots\times V_n$ nach $W$, die nicht die Nullabbildung ist, ist nicht linear. Ferner besitzen die Vektorräume $\mathcal{L}(V_1\times\dots\times V_n;W)$ und $\mathcal{L}(V_1,\dots,V_n;W)$ auch verschiedene Dimensionen, denn es gilt wegen (1.5) und (3.1)

$$\dim\mathcal{L}(V_1\times\dots\times V_n;W)=(\dim V_1+\ldots+\dim V_n)\dim W.$$

Zum Beispiel bildet die bilineare Abbildung $(u,v)\in\mathbb{R}^2\mapsto uv\in\mathbb{R}$ eine Basis in dem eindimensionalen Vektorraum $\mathcal{L}(\mathbb{R},\mathbb{R};\mathbb{R})$, während die linearen Abbildungen $(u,v)\in\mathbb{R}^2\mapsto u\in\mathbb{R}$ und $(u,v)\in\mathbb{R}^2\mapsto v\in\mathbb{R}$ eine Basis in dem zweidimensionalen Vektorraum $\mathcal{L}(\mathbb{R}\times\mathbb{R};\mathbb{R})$ bilden.

**Beispiele** Die Abbildungen

$$\begin{gathered}(\boldsymbol{u},\boldsymbol{v})\in\mathbb{R}^3\times\mathbb{R}^3\mapsto\boldsymbol{u}\times\boldsymbol{v}\in\mathbb{R}^3,\\ (\boldsymbol{A},\boldsymbol{v})\in\mathbb{M}(m\times n,\mathbb{K})\times\mathbb{K}^n\mapsto\boldsymbol{A}\boldsymbol{v}\in\mathbb{K}^m\end{gathered}$$

sind bilinear, die Abbildung

$$(\boldsymbol{A}_1,\dots,\boldsymbol{A}_n)\in\mathbb{M}(m\times m,\mathbb{K})^n\mapsto\boldsymbol{A}_1\cdot\ldots\cdot\boldsymbol{A}_n\in\mathbb{M}(m\times m,\mathbb{K})$$

ist multilinear.

**Beispiele: Höhere Ableitungen vektorwertiger Abbildungen mit vektorwertigen Argumenten**
Es seien $f:\mathbb{R}^m\to\mathbb{R}^n$ eine glatte Abbildung und $x_0\in\mathbb{R}^m$ ein fixierter Punkt. Dann ist die Ableitung $f'(x_0)\in\mathcal{L}(\mathbb{R}^m;\mathbb{R}^n)$ von $f$ in $x_0$ eine lineare Abbildung von $\mathbb{R}^m$ nach $\mathbb{R}^n$, die man z.B. nach der Formel

$$f'(x_0)v=\lim_{t\to 0}\frac{1}{t}\left(f(x_0+tv)-f(x_0)\right)\quad\text{für alle }\ v\in\mathbb{R}^m.$$

berechnen kann. Die zweite Ableitung $f''(x_0)\in\mathcal{L}(\mathbb{R}^m,\mathbb{R}^m;\mathbb{R}^n)$ von $f$ in $x_0$ ist eine bilineare Abbildung von $\mathbb{R}^m\times\mathbb{R}^m$ nach $\mathbb{R}^n$, und es gilt

$$f''(x_0)(v,w)=\lim_{t\to 0}\frac{1}{t}\left(f'(x_0+tw)v-f'(x_0)v\right)\quad\text{für alle }\ v,w\in\mathbb{R}^m.$$

Analog ist die $j$-te Ableitung $f^{(j)}(x_0) \in \mathcal{L}(\mathbb{R}^m, \dots, \mathbb{R}^m; \mathbb{R}^n)$ von $f$ in $x_0$ eine multilineare Abbildung von $\mathbb{R}^m \times \ldots \times \mathbb{R}^m$ ($j$ Faktoren) nach $\mathbb{R}^n$. Alle diese multilinearen Abbildungen spielen z.B. in der Taylor-Formel

$$f(x_0 + v) = \sum_{j=0}^{k} \frac{1}{j!} f^{(j)}(x_0)(v, \dots, v) + \int_0^1 \frac{(1-t)^k}{k!} f^{(k+1)}(x_0 + tv)(v, \dots, v)\, dt$$

eine Rolle.

**Im weiteren sei stets $V$ ein endlich-dimensionaler Vektorraum über $\mathbb{K}$.**

**Lineare Funktionale, dualer Raum und duale Basis** Eine lineare Abbildung $\boldsymbol{\varphi} : V \to \mathbb{K}$ heißt lineares Funktional oder lineare Form oder Linearform oder Kovektor auf $V$. Der Vektorraum $\mathcal{L}(V; \mathbb{K})$ der linearen Funktionale auf $V$ wird mit $V^*$ bezeichnet und heißt dualer Vektorraum zu $V$. Zu jeder Basis $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ in $V$ existiert genau eine Basis $\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^n$ in $V^*$ (die sogenannte zu $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ duale Basis), so daß

$$\boldsymbol{\beta}^j(\boldsymbol{b}_i) = \delta_i^j. \tag{9.1}$$

Dabei ist

$$\delta_i^j = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst} \end{cases}$$

das Kronecker-Symbol, und es gilt

$$\boldsymbol{v} = \sum_{j=1}^{n} \boldsymbol{\beta}^j(\boldsymbol{v}) \boldsymbol{b}_j \text{ für alle } \boldsymbol{v} \in V, \quad \boldsymbol{\varphi} = \sum_{j=1}^{n} \boldsymbol{\varphi}(\boldsymbol{b}_j) \boldsymbol{\beta}^j \text{ für alle } \boldsymbol{\varphi} \in V^*. \tag{9.2}$$

**Einsteinsche Summenkonvention** Daß die Indizes der Basisfunktionale $\boldsymbol{\beta}^j$ oben geschrieben werden und daß in der Tensorrechnung ein Index des Kronecker-Symbols $\delta_i^j$ oben geschrieben wird (im Gegensatz zur Schreibweise (4.1)), liegt an der sogenannten Einsteinschen Summenkonvention. Diese besagt, daß über gleiche, oben und unten auftretende Indizes summiert werden soll. Setzt man diese Konvention voraus, so gehen die Formeln (9.2) über in

$$\boldsymbol{v} = \boldsymbol{\beta}^j(\boldsymbol{v}) \boldsymbol{b}_j \text{ für alle } \boldsymbol{v} \in V, \quad \boldsymbol{\varphi} = \boldsymbol{\varphi}(\boldsymbol{b}_j) \boldsymbol{\beta}^j \text{ für alle } \boldsymbol{\varphi} \in V^*.$$

Diese Konvention erspart nicht nur Schreibarbeit, sondern weist auch häufig auf Fehler in komplizierten Tensorformeln hin: Wenn der gleiche Index zweimal oben oder zweimal unten oder sogar öfter als zweimal auftritt, dann ist irgendwas falsch.

**Der biduale Vektorraum und seine kanonische Identifizierung mit den Grundraum** Der duale Vektorraum zu $V^*$ wird mit $V^{**}$ bezeichnet und bidualer Vektorraum zu $V$ genannt. Dabei ist die Abbildung

$$\boldsymbol{v} \in V \mapsto \boldsymbol{\alpha} \in V^{**} : \quad \boldsymbol{\alpha}(\boldsymbol{\varphi}) = \boldsymbol{\varphi}(\boldsymbol{v}) \text{ für alle } \boldsymbol{\varphi} \in V^*$$

linear und bijektiv, und man identifiziert mit ihrer Hilfe die entsprechenden Elemente von $V$ und $V^{**}$ und damit die Räume $V$ und $V^{**}$.

**Tensoren, Kovarianz und Kontravarianz** Eine multilinearen Abbildung von $V^k \times (V^*)^l$ nach $\mathbb{K}$ heißt Tensor $(k+l)$-ter Stufe auf $V$ oder, genauer, $k$-fach kovarianter und $l$-fach kontravarianter Tensor auf $V$. Den Vektorraum aller $k$-fach kovarianten und $l$-fach kontravarianten Tensoren auf $V$ wird mit $\mathcal{T}_k^l(V)$ bezeichnet. Insbesondere ist dann

$$\mathcal{T}_1^0(V) = V^*,$$

also Funktionale sind einfach kovariante Tensoren, und (wegen der Identifizierung von $V$ und $V^{**}$)

$$\mathcal{T}_0^1(V) = V,$$

also Vektoren sind einfach kontravariante Tensoren.

**Bemerkung zur Bezeichnungsweise** Vektoren $\boldsymbol{v} \in V$ sind einfach kontravariante Tensoren auf $V$. Andererseits sind Tensoren $\boldsymbol{\sigma} \in \mathcal{T}_k^l(V)$ in gewissem Sinne Vektoren, nämlich Elemente des Vektorraumes $\mathcal{T}_k^l(V)$. Das kann aber nicht zur Sprachverwirrung führen, weil in der Tensorrechnung der Begriff "Vektor" ausschließlich für die Elemente des Grundraumes $V$ benutzt wird.

**Tensorprodukt** Für zwei Tensoren $\boldsymbol{\sigma} \in \mathcal{T}_k^l(V)$ und $\boldsymbol{\tau} \in \mathcal{T}_m^n(V)$ ist ihr Tensorprodukt $\boldsymbol{\sigma} \otimes \boldsymbol{\tau} \in \mathcal{T}_{k+m}^{l+n}(V)$ definiert durch

$$\begin{aligned}(\boldsymbol{\sigma} \otimes \boldsymbol{\tau})&(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{k+m}, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{l+n}) \\ &= \boldsymbol{\sigma}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_l)\boldsymbol{\tau}(\boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_{k+m}, \boldsymbol{\varphi}_{l+1}, \dots, \boldsymbol{\varphi}_{l+n})\end{aligned} \tag{9.3}$$

für alle $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{k+m} \in V$ und $\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{l+n} \in V^*$. Das Tensorprodukt besitzt folgende Eigenschaften:
(i) Die Abbildung $(\boldsymbol{\sigma}, \boldsymbol{\tau}) \in \mathcal{T}_k^l(V) \times \mathcal{T}_m^n(V) \mapsto \boldsymbol{\sigma} \otimes \boldsymbol{\tau} \in \mathcal{T}_{k+m}^{l+n}(V)$ ist bilinear.
(ii) Es gilt $\boldsymbol{\rho} \otimes (\boldsymbol{\sigma} \otimes \boldsymbol{\tau}) = (\boldsymbol{\rho} \otimes \boldsymbol{\sigma}) \otimes \boldsymbol{\tau}$, und dieser Tensor wird dann mit $\boldsymbol{\rho} \otimes \boldsymbol{\sigma} \otimes \boldsymbol{\tau}$ bezeichnet.
(iii) Im allgemeinen ist $\boldsymbol{\sigma} \otimes \boldsymbol{\tau} \neq \boldsymbol{\tau} \otimes \boldsymbol{\sigma}$. Für $\boldsymbol{v} \in V$ und $\boldsymbol{\varphi} \in V^*$ z.B. gilt aber stets $\boldsymbol{v} \otimes \boldsymbol{\varphi} = \boldsymbol{\varphi} \otimes \boldsymbol{v}$.

**Basen in Tensorräumen und Koordinaten von Tensoren** Es seien $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ eine Basis in $V$ und $\boldsymbol{\beta}^1, \dots, \boldsymbol{\beta}^n$ die zu dieser Basis duale Basis in $V^*$. Dann ist

$$\{\boldsymbol{b}_{i_1} \otimes \dots \boldsymbol{b}_{i_l} \otimes \boldsymbol{\beta}^{j_1} \otimes \dots \otimes \boldsymbol{\beta}^{j_k} : \, i_1, \dots, i_l, j_1, \dots, j_k = 1, \dots, n\}$$

eine Basis in $\mathcal{T}_k^l(V)$. Insbesondere ist

$$\dim \mathcal{T}_k^l(V) = n^{k+l}.$$

Wenn $\boldsymbol{\sigma} \in \mathcal{T}_k^l(V)$ ein Tensor ist und

$$\boldsymbol{\sigma} = \sigma_{j_1 \dots j_k}^{i_1 \dots i_l} \boldsymbol{b}_{i_1} \otimes \dots \boldsymbol{b}_{i_l} \otimes \boldsymbol{\beta}^{j_1} \otimes \dots \otimes \boldsymbol{\beta}^{j_k}$$

seine Entwicklung nach der obigen Basis (mit Summenkonvention), d.h. wenn die Zahlen $\sigma_{j_1 \dots j_k}^{i_1 \dots i_l}$ die Koordinaten von $\boldsymbol{\sigma}$ bzgl. dieser Basis sind, so nennt man diese Zahlen auch (etwas im Widerspruch zum üblichen Koordinatenbegriff) Koordinaten von $\boldsymbol{\sigma}$ bzgl. der Basis $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ im Grundraum $V$. Dabei gilt

$$\sigma_{j_1 \dots j_k}^{i_1 \dots i_l} = \boldsymbol{\sigma}(\boldsymbol{b}_{j_1}, \dots, \boldsymbol{b}_{j_k}, \boldsymbol{\beta}^{i_1}, \dots, \boldsymbol{\beta}^{i_l}).$$

**Duale Paarung** Die Abbildung $(\boldsymbol{u},\boldsymbol{\varphi}) \in V \times V^* \mapsto \boldsymbol{\varphi}(\boldsymbol{v}) \in \mathbb{K}$ ist bilinear und folglich ein einfach kovarianter, einfach kontravarianter Tensor auf $V$, die sogenannte duale Paarung auf $V$. Die Koordinaten bzgl. einer Basis $\boldsymbol{b}_1,\dots,\boldsymbol{b}_n$ sind $\boldsymbol{\beta^j}(\boldsymbol{b}_i) = \delta_i^j$ (das Kronecker-Symbol, vgl. (9.1)).

**Basiswechsel und entsprechende Koordinatentransformation** Es seien $\boldsymbol{b}_1,\dots,\boldsymbol{b}_n$ und $\boldsymbol{b}'_1,\dots,\boldsymbol{b}'_n$ zwei Basen in $V$, und $\boldsymbol{C} = [C_j^i] \in \mathbb{M}(n \times n;\mathbb{K})$ sei die Matrix des Wechsels von der Basis $\boldsymbol{b}_j$ zur Basis $\boldsymbol{b}'_i$ (vgl. (4.2) und (4.3)), d.h.

$$\boldsymbol{b}_j = C_j^i \boldsymbol{b}'_i. \tag{9.4}$$

Wenn $\boldsymbol{\beta}^1,\dots,\boldsymbol{\beta}^n$ und $\boldsymbol{\beta}'^1,\dots,\boldsymbol{\beta}'^n$ die entsprechenden dualen Basen in $V^*$ sind, so ist $(\boldsymbol{C}^T)^{-1}$ die Matrix des Wechsels von der Basis $\boldsymbol{\beta}_j$ zur Basis $\boldsymbol{\beta}'_i$, d.h. es gilt

$$\boldsymbol{\beta}'^i = C_j^i \boldsymbol{\beta}^j \ \text{ und } \ C_j^i = \langle \boldsymbol{b}_j, \boldsymbol{\beta}'^i \rangle.$$

Es sei schließlich $\boldsymbol{\sigma} \in \mathcal{T}_k^l(V)$ ein Tensor, und $\sigma_{j_1\dots j_k}^{i_1\dots i_l}$ bzw. $\sigma'^{i_1\dots i_l}_{j_1\dots j_k}$ seien die Koordinaten von $\boldsymbol{\sigma}$ bzgl. der Basen $\boldsymbol{b}_j$ bzw. $\boldsymbol{b}'_i$. Dann gilt

$$C_{j_1}^{s_1} \cdots C_{j_k}^{s_k} \sigma'^{r_1\dots r_l}_{s_1\dots s_k} = C_{i_1}^{r_1} \cdots C_{i_l}^{r_l} \sigma_{j_1\dots j_k}^{i_1\dots i_l}.$$

Wenn man diese Formel nach den "neuen" Koordinaten $\sigma'^{r_1\dots r_l}_{s_1\dots s_k}$ umstellen will, muß man für die Komponenten der Matrix $\boldsymbol{C}^{-1}$ Bezeichnungen einführen Es sei also $\boldsymbol{C}^{-1} = [C'^i_j]$, d.h.

$$\boldsymbol{b}'_j = C'^i_j \boldsymbol{b}_i \ \text{ und } \ C'^i_j C_k^j = C_j^i C'^j_k = \delta_k^i. \tag{9.5}$$

Dann folgt

$$\sigma'^{r_1\dots r_l}_{s_1\dots s_k} = C_{i_1}^{r_1} \cdots C_{i_l}^{r_l} C'^{j_1}_{s_1} \cdots C'^{j_k}_{s_k} \sigma_{j_1\dots j_k}^{i_1\dots i_l}. \tag{9.6}$$

Hier sind auch die Bezeichnungen "kovariant" und "kontravariant" begründet: Die unten stehenden, d.h. kovarianten Indizes der Koordinaten eines Tensors transformieren sich unter einem Basiswechsel nach der gleichen Regel wie die Basen (nämlich $\boldsymbol{b}'_j = C'^i_j \boldsymbol{b}_i$), während die oben stehenden, d.h. kontravarianten Indizes sich nach der in gewissem Sinne entgegengesetzten Regel transformieren.

**Im weiteren sei stets $V$ ein endlich-dimensionaler Vektorraum über $\mathbb{R}$ mit einem fixierten Skalarprodukt.**

**Metrischer Tensor** Die Abbildung $(\boldsymbol{u},\boldsymbol{v}) \in V^2 \mapsto \boldsymbol{u} \cdot \boldsymbol{v} \in \mathbb{R}$ ist ein zweifach kovarianter Tensor auf $V$, der sogenannte metrische Tensor (zu dem gegebenen Skalarprodukt). Seine Koordinaten bzgl. einer Basis $\boldsymbol{b}_1,\dots,\boldsymbol{b}_n$ sind

$$g_{ij} = \boldsymbol{b}_i \cdot \boldsymbol{b}_j$$

($g_{ij}$ ist eine Standardbezeichnung in der Tensorrechnung). Wenn $\boldsymbol{b}_1,\dots,\boldsymbol{b}_n$ eine Orthonormalbasis ist, so gilt

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

**Dualitätsabbildung und Transformation von Tensoren beliebigen Typs in rein kovariante Tensoren** Die Abbildung $\boldsymbol{J} : V \to V^*$, die definiert ist durch

$$(\boldsymbol{J}(\boldsymbol{v}))\,(\boldsymbol{u}) = \boldsymbol{u} \cdot \boldsymbol{v} \ \text{ für alle } \ \boldsymbol{u},\boldsymbol{v} \in V,$$

ist linear und bijektiv und heißt Dualitätsabbildung des Euklidischen Vektorraumes $V$. Mit ihrer Hilfe identifiziert man häufig die Elemente von $V$ und $V^*$, d.h. man identifiziert ein lineares Funktional $\boldsymbol{\varphi} \in V^*$ mit dem Vektor $\boldsymbol{J}^{-1}\boldsymbol{\varphi} \in V$.
Analog kann man einem beliebigen Tensor $\boldsymbol{\sigma} \in \mathcal{T}_k^l(V)$ eineindeutig z.B. einen rein kovarianten Tensor $\tilde{\boldsymbol{\sigma}} \in \mathcal{T}_{k+l}^0(V)$ zuordnen nach der Vorschrift

$$\tilde{\boldsymbol{\sigma}}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{k+l}) = \boldsymbol{\sigma}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{J}(\boldsymbol{v}_{k+1}), \dots, \boldsymbol{J}(\boldsymbol{v}_{k+l})) \text{ für alle } \boldsymbol{v}_1, \dots, \boldsymbol{v}_{k+l} \in V. \tag{9.7}$$

Wenn dabei $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ eine Basis in $V$ ist und $\sigma_{j_1 \dots j_k}^{i_1 \dots i_l}$ bzw. $\tilde{\sigma}_{j_1 \dots j_{k+l}}$ die Koordinaten von $\boldsymbol{\sigma}$ bzw. $\tilde{\boldsymbol{\sigma}}$ bzgl. dieser Basis sind, so gilt

$$\tilde{\sigma}_{j_1 \dots j_{k+l}} = g_{i_1 j_1} \dots g_{i_l j_l} \sigma_{j_{1+l} \dots j_{k+l}}^{i_1 \dots i_l} \text{ mit } g_{ij} = \boldsymbol{b}_i \cdot \boldsymbol{b}_j. \tag{9.8}$$

Wenn insbesondere $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ eine Orthonormalbasis ist, so gilt $\tilde{\sigma}_{j_1 \dots j_{k+l}} = \sigma_{j_{1+l} \dots j_{k+l}}^{j_1 \dots j_l}$. Deshalb ist es einfach und sinnvoll, daß man, wenn man ausschließlich mit Orthonormalbasen arbeitet, alle auftretenden Tensoren in rein kovariante transformiert. Folglich werden in einem Kontext, in dem ausschließlich mit Orthonormalbasen gearbeitet wird, die Begriffe "kovariant" und "kontravariant" meist überhaupt nicht benötigt, weil von Anfang an ausschließlich mit einem Typ von Tensoren gearbeitet wird.
Bei nicht-orthonormalen Basen können die Transformationen (9.8) allerdings einen erheblichen Rechenaufwand erfordern.

**Beispiel** Wenn $\boldsymbol{\sigma} \in \mathcal{T}_1^1(V)$ die duale Paarung auf $V$ ist und $\tilde{\boldsymbol{\sigma}} \in \mathcal{T}_2^0(V)$ definiert ist entsprechend (9.7), dann ist $\tilde{\boldsymbol{\sigma}}$ der metrische Tensor auf $V$, denn es gilt

$$\tilde{\boldsymbol{\sigma}}(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{\sigma}(\boldsymbol{u}, \boldsymbol{J}(\boldsymbol{v})) = (\boldsymbol{J}(\boldsymbol{v}))(\boldsymbol{u}) = \boldsymbol{u} \cdot \boldsymbol{v} \text{ für alle } \boldsymbol{u}, \boldsymbol{v} \in V.$$

**Heben und Senken von Indizes** Die Abbildung $\boldsymbol{\sigma} \in \mathcal{T}_k^l(V) \mapsto \tilde{\boldsymbol{\sigma}} \in \mathcal{T}_{k+l}^0(V)$, die durch (9.8) "koordinatenweise" bestimmt ist, nennt man auch "Senken aller Indizes". Es werden auch die analogen Abbildungen "Heben oder Senken einiger Indizes" benutzt, die einem gegebenen Tensor eines gegebenen Typs einen anderen Tensor eines anderen Typs zuordnen.
Es sei z.B. ein Tensor $\boldsymbol{\sigma} \in \mathcal{T}_0^2(V)$ mit den Koordinaten $\sigma^{ik}$ bzgl. einer Basis $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ gegeben. Dann definieren die Formeln

$$\tilde{\sigma}_j^i = g_{jk}\sigma^{ik} \text{ und } \tilde{\tilde{\sigma}}_j^i = g_{jk}\sigma^{ki} \text{ mit } g_{ij} = \boldsymbol{b}_i \cdot \boldsymbol{b}_j \tag{9.9}$$

zwei im allgemeinen verschiedene Tensoren $\tilde{\boldsymbol{\sigma}}, \tilde{\tilde{\boldsymbol{\sigma}}} \in \mathcal{T}_1^1(V)$ (genauer deren Koordinaten bzgl. der gegebenen Basis). Oft schreibt man auch ${\sigma^i}_j$ anstelle von $\tilde{\sigma}_j^i$ und ${\sigma_j}^i$ anstelle von $\tilde{\tilde{\sigma}}_j^i$, um Schreibarbeit zu sparen. Dann muß man aber eine Reihenfolge für alle oben und unten stehenden Indizes einführen und beachten. Auf diese Art und Weise können Schreibweisen wie z.B. ${{\sigma^{ij}}_k}^l$ für die Koordinaten eines einfach ko- und dreifach kontravarianten Tensors entstehen.
Übrigens kann man die Tensoren $\tilde{\boldsymbol{\sigma}}$ und $\tilde{\tilde{\boldsymbol{\sigma}}}$ auch basisunabhängig einführen, nämlich durch

$$\tilde{\boldsymbol{\sigma}}(\boldsymbol{v}, \boldsymbol{\varphi}) = \boldsymbol{\sigma}(\boldsymbol{\varphi}, \boldsymbol{J}(\boldsymbol{v})) \text{ und } \tilde{\tilde{\boldsymbol{\sigma}}}(\boldsymbol{v}, \boldsymbol{\varphi}) = \boldsymbol{\sigma}(\boldsymbol{J}(\boldsymbol{v}), \boldsymbol{\varphi}) \text{ für alle } \boldsymbol{v} \in V \text{ und } \boldsymbol{\varphi} \in V^*. \tag{9.10}$$

**Koordinatenweiser Zugang zum Tensorbegriff** Der Begriff "Tensor" wird bisweilen nicht für das "invariante" (d.h. von der Wahl irgendwelcher Basen unabhängige) Objekt "multilineare Abbildung von $V^k \times (V^*)^l$ nach $\mathbb{R}$" benutzt, sondern für die Menge aller seiner Koordinaten bzgl. aller

möglichen Basen. Die entsprechende Sprachregelung ist dann die folgende:
Ein $k$-fach kovarianter und $l$-fach kontravarianter Tensor auf $V$ ist gegeben, wenn für jede Basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ in $V$ Zahlen $\sigma^{i_1 \ldots i_l}_{j_1 \ldots j_k}$ $(i_1 \ldots i_l, j_1 \ldots j_k = 1, \ldots, n)$ gegeben sind, wobei (9.6) gelten soll, wenn die entsprechenden Basen (9.4) gilt. In diesem Sinne ist ein Tensor eine Abbildung "Basis $\mapsto$ Koordinatenmenge" mit einer Transformationsregel.
Wenn allerdings nur eine einzige fixierte Basis benutzt wird, so kann man jeden Tensor mit seinen Koordinaten bzgl. dieser Basis identifizieren. In einem solchen Kontext ist ein $k$-fach kovarianter und $l$-fach kontravarianter Tensor nichts anderes als eine "$k + l$–dimensionale Matrix".

**Invariantes und koordinatenweises Arbeiten** Zwischen den Formeln (9.7) und (9.8) (und analog zwischen (9.10) und (9.9)) besteht ein wesentlicher Unterschied: (9.7) definiert den Tensor $\tilde{\boldsymbol{\sigma}}$ mit Hilfe des Tensors $\boldsymbol{\sigma}$ unabhängig von der Wahl irgendwelcher Basen. (9.8) dagegen definiert für jede Basis $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ Zahlen $\tilde{\sigma}_{j_1 \ldots j_{k+l}}$ (mit Hilfe der Koordinaten $\sigma^{i_1 \ldots i_l}_{j_1 \ldots j_k}$ des Tensors $\boldsymbol{\sigma}$ bzgl. dieser Basis). Ob die so erzeugte Abbildung

$$(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) \mapsto \{\tilde{\sigma}_{j_1 \ldots j_{k+l}} : \; j_1 \ldots j_{k+l} = 1, \ldots, n\}$$

aber tatsächlich einen Tensor erzeugt, d.h. ob die Transformationsregel

$$\tilde{\sigma}'_{s_1 \ldots s_{k+l}} = C'^{r_1}_{s_1} \ldots C'^{r_{k+l}}_{s_{k+l}} \tilde{\sigma}_{r_1 \ldots r_{k+l}} \tag{9.11}$$

erfüllt ist, muß noch nachgeprüft werden. Das ist in diesem Fall nicht schwer, denn (9.11) folgt unmittelbar aus (9.5), (9.6), (9.8) und der (9.8) entsprechenden Formel

$$\tilde{\sigma}'_{j_1 \ldots j_{k+l}} = g'_{i_1 j_1} \ldots g'_{i_l j_l} \sigma'^{i_1 \ldots i_l}_{j_{1+l} \ldots j_{k+l}} \text{ mit } g'_{ij} = \boldsymbol{b}'_i \cdot \boldsymbol{b}'_j.$$

Bei anderen "koordinatenweise" eingeführten Tensoroperationen kann das erheblich komplizierter sein.

**Im weiteren betrachten wir ausschließlich kovariante Tensoren. Anstelle von " $k$-fach kovarianter Tensor" sagt man dann einfacher "Tensor $k$-ter Stufe".**

**Symmetrische und antisymmetrische Tensoren** Ein Tensor $\boldsymbol{\sigma} \in \mathcal{T}^0_k(V)$ heißt symmetrisch bzw. antisymmetrisch, wenn gilt

$$\boldsymbol{\sigma}(\boldsymbol{v}_{j_1}, \ldots, \boldsymbol{v}_{j_k}) = \boldsymbol{\sigma}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) \text{ bzw. } \boldsymbol{\sigma}(\boldsymbol{v}_{j_1}, \cdots, \boldsymbol{v}_{j_k}) = \epsilon_{j_1 \ldots j_k} \boldsymbol{\sigma}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$$

für alle $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in V$ und alle Permutationen $(j_1, \ldots, j_k)$ von $(1, \ldots, k)$. Dabei ist $\epsilon_{j_1 \ldots j_k}$ der in (5.2) eingeführte sogenannte Permutationstensor. Antisymmetrische Tensoren werden auch äußere Formen genannt.
Die Mengen der symmetrischen bzw. antisymmetrischen Tensoren auf $V$ sind jeweils Unterräume in $\mathcal{T}^0_k(V)$. Diese Unterräume werden mit $\mathcal{S}_k(V)$ bzw. $\mathcal{A}_k(V)$ bezeichnet, und es gilt, falls $\dim V = n$ ist,

$$\dim \mathcal{S}_k(V) = \frac{n(n+1) \cdots (n+k-1)}{k!},$$

$$\dim \mathcal{A}_k(V) = \frac{n(n-1) \cdots (n-k+1)}{k!}, \text{ insbesondere } \mathcal{A}_k(V) = \{0\} \text{ für } k > n.$$

Ferner ist

$$\mathcal{T}_2^0(V) = \mathcal{S}_2(V) \oplus \mathcal{A}_2(V),$$

d.h. jeder Tensor zweiter Stufe kann eindeutig als Summe eines symmetrischen und eines antisymmetrischen Tensors zweiter Stufe (seines sogenannten symmetrischen bzw. seines antisymmetrischen Anteils) dargestellt werden, nämlich

$$\boldsymbol{\sigma}(\boldsymbol{u},\boldsymbol{v}) = \frac{1}{2}\Big(\boldsymbol{\sigma}(\boldsymbol{u},\boldsymbol{v}) + \boldsymbol{\sigma}(\boldsymbol{v},\boldsymbol{u})\Big) + \frac{1}{2}\Big(\boldsymbol{\sigma}(\boldsymbol{u},\boldsymbol{v}) - \boldsymbol{\sigma}(\boldsymbol{v},\boldsymbol{u})\Big) \text{ für alle } \boldsymbol{u},\boldsymbol{v} \in V.$$

**Tensoren zweiter Stufe und lineare Abbildungen auf $V$** (i) Jeder linearen Abbildung $\boldsymbol{F} \in \mathcal{L}(V)$ kann man einen Tensor zweiter Stufe $\boldsymbol{\sigma} \in \mathcal{T}_2^0(V)$ zuordnen durch die Vorschrift

$$\boldsymbol{\sigma}(\boldsymbol{u},\boldsymbol{v}) = \boldsymbol{F}(\boldsymbol{u}) \cdot \boldsymbol{v} \text{ für alle } \boldsymbol{u},\boldsymbol{v} \in V. \tag{9.12}$$

Die so definierte Abbildung $\boldsymbol{F} \mapsto \boldsymbol{\sigma}$ ist linear und bijektiv von $\mathcal{L}(V)$ auf $\mathcal{T}_2^0(V)$, und folglich identifiziert man häufig lineare Abbildungen mit den entsprechenden Tensoren zweiter Stufe (und umgekehrt). Dabei gilt

$$\sigma_{ij} = a_{ij} \text{ für alle } i,j = 1.\dots,\dim V,$$

wobei $\sigma_{ij}$ die Koordinaten von $\boldsymbol{\sigma}$ bzgl. einer gegebenen Orthonormalbasis in $V$ sind und $[a_{ij}]$ die Matrixdarstellung von $F$ bzgl. dieser Basis.
(ii) Wenn man Vektoren mit Funktionalen indentifiziert, so ist, für gegebene $\boldsymbol{u},\boldsymbol{v} \in V$, $\boldsymbol{u} \otimes \boldsymbol{v}$ ein Tensor zweiter Stufe. Wenn man die diesem Tensor zweiter Stufe entsprechende lineare Abbildung ebenfalls mit $\boldsymbol{u} \otimes \boldsymbol{v}$ bezeichnet, so gilt wegen (9.3) und (9.12)

$$(\boldsymbol{u} \otimes \boldsymbol{v})(\boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} \text{ für alle } \boldsymbol{w} \in V.$$

Insbesondere ist, wenn $\boldsymbol{b}_1,\dots,\boldsymbol{b}_n$ eine Basis in $V$ ist, $\{\boldsymbol{b}_i \otimes \boldsymbol{b}_j : i,j = 1,\dots,n\}$ eine Basis in $\mathcal{L}(V)$.

**Anwendungsbeispiel: Trägheitstensor** Das Gebiet $\Omega \subset \mathbb{R}^3$ beschreibe einen starren Körper, der sich mit der Winkelgeschwindigkeit $\boldsymbol{\omega} \in \mathbb{R}^3$ drehe, und die Abbildung $m : \Omega \to \mathbb{R}$ sei seine Massendichte. Dann ist

$$\boldsymbol{L} = \int_\Omega m(\boldsymbol{x})\Big(\|\boldsymbol{x}\|^2\boldsymbol{\omega} - (\boldsymbol{\omega}\cdot\boldsymbol{x})\boldsymbol{x}\Big)d\boldsymbol{x}$$

der Drehimpuls des Körpers. Die lineare Abbildung $\boldsymbol{\omega} \mapsto \boldsymbol{L}$ bzw. der ihr entsprechende Tensor

$$\boldsymbol{I} = \int_\Omega m(\boldsymbol{x})\Big(\|\boldsymbol{x}\|^2\boldsymbol{g} - \boldsymbol{x}\otimes\boldsymbol{x}\Big)d\boldsymbol{x} \in \mathcal{S}_2(\mathbb{R}^3)$$

heißen Trägheitstensor des Körpers. Dabei ist $\boldsymbol{g}$ der metrische Tensor in $\mathbb{R}^3$. Ferner ist

$$\frac{1}{2}\boldsymbol{I}(\boldsymbol{\omega},\boldsymbol{\omega}) = \frac{1}{2}\boldsymbol{L}\cdot\boldsymbol{\omega} = \frac{1}{2}\int_\Omega m(\boldsymbol{x})\|\boldsymbol{\omega}\times\boldsymbol{x}\|^2 d\boldsymbol{x}$$

die kinetische Energie des Körpers.
Wenn $x_i$ $(i = 1, 2, 3)$ die Koordinaten bzgl. der kanonischen Basis in $\mathbb{R}^3$ der unabhängigen Ortsvariablen $\boldsymbol{x} \in \Omega$ sind, so sind

$$\int_\Omega m(x_1, x_2, x_3)\Big((x_1^2 + x_2^2 + x_3^2)\delta_{ij} - x_i x_j\Big) dx_1 dx_2 dx_3 \ \ (i, j = 1, 2, 3)$$

die Koordinaten des Trägheitstensors bzgl. dieser Basis.

**Anwendungsbeispiel: Deformationstensor** Das Gebiet $\Omega \subset \mathbb{R}^3$ beschreibe einen deformierbaren Körper und die Abbildung $\boldsymbol{u} : \Omega \to \mathbb{R}^3$ seine Deformation. Ferner sei $\boldsymbol{x}_0 \in \Omega$ fixiert. Dann heißen die lineare Abbildung $\boldsymbol{F} \in \mathcal{L}(\mathbb{R}^3)$, die definiert ist durch

$$\boldsymbol{F}(\boldsymbol{v}) = \lim_{t\to 0} \frac{\boldsymbol{u}(\boldsymbol{x}_0 + t\boldsymbol{v}) - \boldsymbol{u}(\boldsymbol{x}_0)}{t} \ \text{ für alle } \ \boldsymbol{v} \in \mathbb{R}^3,$$

bzw. der ihr entsprechende Tensor zweiter Stufe Deformationstensor im Punkt $\boldsymbol{x}_0$. Es gilt

$$\boldsymbol{u}(\boldsymbol{x}_0 + \boldsymbol{v}) = \boldsymbol{u}(\boldsymbol{x}_0) + \boldsymbol{F}(\boldsymbol{v}) + O(\|\boldsymbol{v}\|^2) \ \text{ für } \ \|\boldsymbol{v}\| \to 0,$$

und folglich beschreibt der Deformationstensor die infinitesimale Deformation jedes infinitesimalen Ortszuwachses im Punkt $\boldsymbol{x}_0$.
Wenn $x_i$ bzw. $u_i(x_1, x_2, x_3)$ $(i = 1, 2, 3)$ die Koordinaten bzgl. der kanonischen Basis in $\mathbb{R}^3$ von der unabhängigen Ortsvariablen $\boldsymbol{x} \in \Omega$ bzw. von $\boldsymbol{u}(\boldsymbol{x})$ sind, so sind die partiellen Ableitungen

$$\frac{\partial u_i}{\partial x_j}(\boldsymbol{x}_0) \ \ (i, j = 1, 2, 3)$$

die Koordinaten des Deformationstensors bzgl. dieser Basis.
Bisweilen wird auch mit anderen Tensoren, die die infinitesimale Deformationen nahe $\boldsymbol{x}_0$ beschreiben, gearbeitet, z.B. mit dem symmetrischen Anteil des Deformationstensors (dem sogenannten Greenschen Deformationstensor). Dessen Koordinaten bzgl. der kanonischen Basis in $\mathbb{R}^3$ sind

$$\frac{1}{2}\Big(\frac{\partial u_i}{\partial x_j}(\boldsymbol{x}_0) + \frac{\partial u_j}{\partial x_i}(\boldsymbol{x}_0)\Big) \ \ (i, j = 1, 2, 3).$$

**Anwendungsbeispiel: Spannungstensor** Es sei $\boldsymbol{x}_0$ ein fixierter Punkt in einem Körper, und $\mathcal{M}$ sei eine Hyperfläche durch $\boldsymbol{x}_0$, die den Körper in zwei Teile $A$ und $B$ teilt. Ferner sei $\boldsymbol{n} \in \mathbb{R}^3$ die Einheitsnormale an $\mathcal{M}$ in $\boldsymbol{x}_0$, die in $B$ zeigt, und $\boldsymbol{t} \in \mathbb{R}^3$ sei die Kraft pro Fläche von $\mathcal{M}$, mit der $A$ auf $B$ in $\boldsymbol{x}_0$ wirkt. Dann gilt, daß $\boldsymbol{t}$ von $A$ und $B$ nur über $\boldsymbol{n}$ abhängt und daß die Abbildung $\boldsymbol{n} \mapsto \boldsymbol{t}$ linear ist. Diese lineare Abbildung bzw. der ihr entsprechende Tensor zweiter Stufe in $\mathbb{R}^3$ heißen Spannungstensor in $\boldsymbol{x}_0$. Wenn sich der Körper in Gleichgewicht befindet, ist der Spannungstensor symmetrisch.

**Tensoren vierter Stufe und lineare Abbildungen auf $\mathcal{L}(V)$** Jeder linearen Abbildung $\boldsymbol{F} \in \mathcal{L}(\mathcal{L}(V))$ kann man einen Tensor vierter Stufe $\boldsymbol{\sigma} \in \mathcal{T}_4^0(V)$ zuordnen durch die Vorschrift

$$\boldsymbol{\sigma}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4) = \Big(\boldsymbol{F}(\boldsymbol{u}_1 \otimes \boldsymbol{u}_2)\Big)(\boldsymbol{u}_3) \cdot \boldsymbol{u}_4 \ \text{ für alle } \ \boldsymbol{u}_1, \dots, \boldsymbol{u}_4 \in V.$$

Die so definierte Abbildung $\boldsymbol{F} \mapsto \boldsymbol{\sigma}$ ist linear und bijektiv von $\mathcal{L}(\mathcal{L}(V))$ auf $\mathcal{T}_4^0(V)$, und folglich identifiziert man häufig lineare Abbildungen auf $\mathcal{L}(V)$ mit den entsprechenden Tensoren vierter Stufe auf $V$ (und umgekehrt).

**Anwendungsbeispiel: Hookescher Tensor** Deformierbare Materialien, bei denen in jedem Punkt und bei hinreichend kleinen Deformationen der Spannungstensor linear vom Deformationstensor abhängt, heißen elastisch. Die lineare Abbildung, die dabei (in einem fixierten Punkt) dem Deformationstensor den Spannungstensor zuordnet, bzw. der ihr entsprechende Tensor vierter Stufe in $\mathbb{R}^3$ heißen Hookescher Tensor.
Der Vektorraum $\mathcal{T}_4^0(\mathbb{R}^3)$ besitzt die Dimension $3^4 = 81$. Allerdings sind bei weitem nicht alle Elemente dieses Vektorraumes auch Hookesche Tensoren, also als elastische Materialgesetze interpretierbar. Im idealen Fall (isotropes Material, spannungsfreier Referenzzustand) ist die Menge der Hookeschen Tensoren ein nur zweidimensionaler Unterraum in $\mathcal{T}_4^0(\mathbb{R}^3)$. Dieser Unterraum wird z.B. durch die beiden sogenannten Lamé-Koeffizienten parametrisiert.

**Tensorfelder** In den meisten Anwendungen hängen die betrachteten Tensoren von Parametern ab. Z. B. ist der Deformationszustand eines deformierbaren Körpers nicht durch einen Deformationstensor beschrieben, sondern duch die Abbildung, die jedem Punkt $\boldsymbol{x}_0 \in \Omega$ den Deformationstensor $\boldsymbol{F} \in \mathcal{L}(\mathbb{R}^3)$ in diesem Punkt zuordnet. Eine solche Abbildung "Punkt $\mapsto$ Tensor" heißt Tensorfeld. Häufig wird dann das Wort "Tensor" auch für den Begriff "Tensorfeld" benutzt.
Die Untersuchung von Tensorfeldern (mit Hilfe der Differential- und Integralrechnung) ist Gegenstand der sogenannten Tensoranalysis.

**Beispiel: Der Volumentensor in $\mathbb{R}^n$** Die Abbildung

$$\left( \begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{bmatrix}, \dots, \begin{bmatrix} v_{1n} \\ v_{2n} \\ \vdots \\ v_{nn} \end{bmatrix} \right) \in (\mathbb{R}^n)^n \mapsto \det \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix} \in \mathbb{R}$$

ist ein antisymmetrischer Tensor $n$-ter Stufe in $\mathbb{R}^n$, der sogenannte Volumentensor (oder die sogenannte Volumenform) auf $\mathbb{R}^n$. Er ordnet einem positiv orientierten $n$-Tupel von Vektoren in $\mathbb{R}^n$ das Volumen des von ihnen aufgespannten Parallelepipeds zu. Seine Koordinaten bzgl. einer positiv orientierten Orthonormalbasis in $\mathbb{R}^n$ sind (vgl. (5.1)) gleich $\epsilon_{j_1 \cdots j_n}$ (der Permutationstensor). Seine Koordinaten bzgl. einer beliebigen positiv orientierten Basis $\boldsymbol{b}_1, \dots, \boldsymbol{b}_n$ sind

$$\sqrt{g}\, \epsilon_{j_1 \cdots j_n} \quad \text{mit} \quad g = \det[\boldsymbol{b}_i \cdot \boldsymbol{b}_j]_{i,j=1}^n$$

($g$ ist eine Standardbezeichnung in der Tensorrechnung).
Wegen $\dim \mathcal{A}_n(\mathbb{R}^n) = 1$ ist jeder antisymmetrische Tensor $n$-ter Stufe auf $\mathbb{R}^n$ zum Volumentensor linear abhängig.

**Beispiel: Der Volumentensor eines Unterraumes des $\mathbb{R}^n$** Es sei $U$ sei ein $k$-dimensionaler Unterraum von $\mathbb{R}^n$, und in $U$ sei eine Orientierung fixiert. Dann existiert genau ein $\boldsymbol{\omega} \in \mathcal{A}_k(\mathbb{R}^n)$, der sogenannte Volumentensor von $U$, so daß gilt:
Wenn ein beliebiges Tupel von Vektoren $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k \in \mathbb{R}^n$ gegeben ist, wenn $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k \in U$ die orthogonalen Projektionen dieser Vektoren auf $U$ sind, und wenn das Tupel $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ positiv orientiert ist,

so ist $\boldsymbol{\omega}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ gleich dem Volumen des von $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ aufgespannten Parallelepipeds, nämlich

$$\boldsymbol{\omega}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k) = \sqrt{\det \begin{bmatrix} \boldsymbol{u}_1 \cdot \boldsymbol{u}_1 & \boldsymbol{u}_1 \cdot \boldsymbol{u}_2 & \ldots & \boldsymbol{u}_1 \cdot \boldsymbol{u}_k \\ \boldsymbol{u}_1 \cdot \boldsymbol{u}_2 & \boldsymbol{u}_2 \cdot \boldsymbol{u}_2 & \ldots & \boldsymbol{v}_2 \cdot \boldsymbol{v}_k \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{u}_1 \cdot \boldsymbol{u}_k & \boldsymbol{u}_2 \cdot \boldsymbol{u}_k & \ldots & \boldsymbol{u}_k \cdot \boldsymbol{u}_k \end{bmatrix}}, \tag{9.13}$$

Dabei ist das Volumen von Parallelepipeds in $\mathbb{R}^n$ induktiv nach dem Prinzip von Cavalieri definiert: Das Volumen des von einem Vektor $\boldsymbol{v} \in \mathbb{R}^n$ aufgespannten Parallelepipeds ist gleich $\|\boldsymbol{v}\|$. Das Volumen des von Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ aufgespannten Parallelepipeds ist gleich dem Produkt aus dem Volumen des von den Vektoren $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1} \in \mathbb{R}^n$ aufgespannten Parallelepipeds und der Länge der orthogonalen Projektion von $\boldsymbol{v}_k$ auf $(\mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}\})^\perp$.
Die Formel (9.13) ist eine Verallgemeinerung von (6.2) ($k = n = 3$) und von (6.3) ($k = 2,\ n = 3$).

**Vektoren und antisymmetrische Tensoren zweiter Stufe in $\mathbb{R}^3$** Jedem Vektor $\boldsymbol{u} \in \mathbb{R}^3$ kann man einen Tensor $\boldsymbol{\omega} \in \mathcal{A}_2(\mathbb{R}^3)$ zuordnen durch die Vorschrift

$$\boldsymbol{\omega}(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) \ \text{ für alle } \ \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3.$$

Die so definierte Abbildung $\boldsymbol{u} \mapsto \boldsymbol{\omega}$ ist linear und bijektiv von $\mathbb{R}^3$ auf $\mathcal{A}_2(\mathbb{R}^3)$, und folglich identifiziert man manchmal Vektoren und antisymmetrische Tensoren zweiter Stufe in $\mathbb{R}^3$. Dabei ist $\boldsymbol{\omega}$ der Volumentensor von $\mathrm{span}\{\boldsymbol{u}\}^\perp$, wenn $\|\boldsymbol{u}\| = 1$ ist und wenn in $\mathrm{span}\{\boldsymbol{u}\}^\perp$ die Orientierung gewählt ist, die durch die Standard-Orientierung in $\mathbb{R}^3$ und durch die Normale $\boldsymbol{u}$ induziert ist. Der Vektor $\boldsymbol{u}$ definiert einen homogenen Fluß in $\mathbb{R}^3$, und $\boldsymbol{\omega}(\boldsymbol{v}, \boldsymbol{w})$ ist dann der Durchsatz dieses Flusses durch das Parallelogramm, das durch $\boldsymbol{v}$ und $\boldsymbol{w}$ aufgespannt wird (positiv, wenn das Tripel $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ positiv orientiert ist).
Analog kann man Vektoren $\boldsymbol{w} \in \mathbb{R}^n$ mit Tensoren $\boldsymbol{\omega} \in \mathcal{A}_{n-1}(\mathbb{R}^n)$ identifizieren.