Ralf Gerkmann

Mathematisches Institut
Ludwig-Maximilians-Universität München

Lineare Algebra

(Version vom 16. Juli 2015)

Inhaltsverzeichnis

§ 1. Matrizen und Lineare Gleichungssysteme

Im gesamten Text bezeichnet K stets einen beliebigen Körper, solange nichts genaueres festgelegt wird. Mit $K^\times$ bezeichnen wir die Menge der Elemente in K ungleich 0_K. Für jedes $n \in \mathbb{N}$ ist 0_{K^n} das Tupel $(0_K, \ldots, 0_K) \in K^n$, dessen sämtliche Einträge gleich 0_K sind. Ist $x \in K^n$, dann bezeichnen wir mit $x_1, \ldots, x_n$ die Komponenten von x. Es gilt also jeweils $x = (x_1, \ldots, x_n)$.

Die Elemente in K^n werden auch ***Vektoren*** genannt. Sie lassen sich komponentenweise addieren oder mit einem „Skalar" $\lambda \in K$ multiplizieren. Für beliebige $v, w \in K^n$ definieren wir

$$v + w = (v_1 + w_1, \ldots, v_n + w_n) \qquad \text{und} \qquad \lambda v = (\lambda v_1, \ldots, \lambda v_n).$$

Definition 1.1 Sei $n \in \mathbb{N}$. Eine ***Linearform*** auf dem K^n ist eine Abbildung $\phi : K^n \to K$, die in der Form $\phi(x) = \sum_{j=1}^n a_j x_j$ mit geeigneten Koeffizienten $a_1, \ldots, a_n \in K$ dargestellt werden kann.

Definition 1.2 Seien $m, n \in \mathbb{N}$. Ein ***lineares Gleichungssystem*** (LGS) bestehend aus m Gleichungen in n Unbekannten ist ein Paar

$$((\phi_1, \ldots, \phi_m), b)$$

bestehend aus einem Tupel von m Linearformen ϕ_i und einem Vektor $b \in K^m$. Ist $b = 0_{K^m}$, dann spricht man von einem ***homogenen***, sonst von einem ***inhomogenen*** LGS.

Häufig werden lineare Gleichungssysteme in der folgenden ausgeschriebenen Form dargestellt: Sind die m Linearformen gegeben durch $\phi_i(x) = \sum_{j=1}^n a_{ij} x_j$ für $1 \le i \le m$ und ist $b = (b_1, \ldots, b_m)$, dann schreibt man das LGS $((\phi_1, \ldots, \phi_m), b)$ auch in der Form

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array} \tag{1}$$

Definition 1.3 Ein Element $v \in K^n$ bezeichnet man als ***Lösung*** des LGS $((\phi_1, \ldots, \phi_m), b)$, wenn die Gleichungen $\phi_i(v) = b_i$ für $1 \le i \le m$ erfüllt sind. Die Teilmenge

$$\mathscr{L} \quad = \quad \{v \in K^n \mid \phi_i(v) = b_i \text{ für } 1 \le i \le m\} \quad \subseteq \quad K^n$$

wird die ***Lösungsmenge*** des LGS genannt.

Ein Element $v = (v_1, ..., v_n) \in K^n$ liegt nach Definition genau dann in der $\mathscr{L}$, wenn alle m Gleichungen in (1) erfüllt sind, nachdem man x_j für $1 \le j \le n$ durch v_j ersetzt hat. Wir notieren nun zunächst einige einfachen Beobachtungen zur Gestalt solcher Lösungsmengen.

Lemma 1.4 Ist $\phi : K^n \to K$ eine Linearform und sind $v, w \in K^n$ und $\lambda \in K$, dann gilt $\phi(v + w) = \phi(v) + \phi(w)$ und $\phi(\lambda v) = \lambda\phi(v)$.

Beweis: Beide Gleichungen erhält man durch einfaches Nachrechnen. Nach Definition der Linearformen hat die Abbildung ϕ die Form $\phi(x) = \sum_{j=1}^{n} a_j x_j$ mit $a_1, ..., a_n \in K$. Man erhält nun

$$\phi(v + w) \quad = \quad \sum_{j=1}^{n} a_j(v_j + w_j) \quad = \quad \sum_{j=1}^{n} a_j v_j + \sum_{j=1}^{n} a_j w_j \quad = \quad \phi(v) + \phi(w)$$

und ebenso $\phi(\lambda v) = \sum_{j=1}^{n} a_j(\lambda v_j) = \lambda \sum_{j=1}^{n} a_j v_j = \lambda\phi(v)$. □

Proposition 1.5 Sei $((\phi_1, ..., \phi_m), 0_{K^m})$ ein homogenes LGS mit Lösungsmenge $\mathscr{L} \subseteq K^n$. Ist $\lambda \in K$ und sind $v, w \in \mathscr{L}$, dann gilt auch $v + w \in \mathscr{L}$ und $\lambda v \in \mathscr{L}$.

Beweis: Aus $v, w \in \mathscr{L}$ folgt nach Definition $\phi_i(v) = \phi_i(w) = 0_K$ für $1 \le i \le m$. Nach Lemma (1.4) ist damit auch $\phi_i(v + w) = \phi_i(v) + \phi_i(w) = 0_K + 0_K = 0_K$ für $1 \le i \le m$ und somit $v + w \in \mathscr{L}$. Ebenso ist $\phi_i(\lambda v) = \lambda\phi_i(v) = \lambda \cdot 0_K = 0_K$ und damit $\lambda v \in \mathscr{L}$. □

Proposition 1.6 Sei $((\phi_1, ..., \phi_m), b)$ ein beliebiges LGS, $\mathscr{L}$ die Lösungsmenge dieses Systems und $\mathscr{L}^{\mathrm{h}}$ die Lösungsmenge des zugehörigen homogenen LGS gegeben durch $((\phi_1, ..., \phi_m), 0_{K^m})$. Ist $v \in \mathscr{L}$ ein beliebiges Element, dann gilt $\mathscr{L} = v + \mathscr{L}^{\mathrm{h}}$, also

$$\mathscr{L} \quad = \quad \{v + w \mid w \in \mathscr{L}^{\mathrm{h}}\}.$$

Beweis: Aus $v \in \mathscr{L}$ folgt $\phi_i(v) = b_i$ für $1 \le i \le m$. Beweisen wir nun zunächst die Inklusion „$\supseteq$“. Ist $w \in \mathscr{L}^{\mathrm{h}}$, dann gilt $\phi_i(w) = 0_K$ für $1 \le i \le m$. Wir erhalten $\phi_i(v + w) = \phi_i(v) + \phi_i(w) = b_i + 0_K = b_i$ für $1 \le i \le m$ und somit $v + w \in \mathscr{L}$.

Zum Nachweis von „$\subseteq$“ sei $u \in \mathscr{L}$ vorgegeben. Dann gilt $\phi_i(u) = b_i$ für $1 \le i \le m$. Setzen wir $w = u - v$, dann folgt $\phi_i(w) = \phi_i(u) + \phi_i(-v) = \phi_i(u) - \phi_i(v) = b_i - b_i = 0_K$ für alle i. Dies zeigt, dass w in $\mathscr{L}^{\mathrm{h}}$ enthalten ist. Also ist u in der Form $v + w$ mit $w \in \mathscr{L}^{\mathrm{h}}$ darstellbar. □

Um die Lösungsmenge eines LGS konkret auszurechnen, benötigt man Operationen, die zur Vereinfachung des Systems genutzt werden können, ohne die Lösungsmenge zu ändern. Zunächst bemerken wir

Lemma 1.7 Seien ϕ, ψ Linearformen auf dem K^n und $\lambda \in K$. Dann sind auch die beiden Abbildungen $\phi + \psi$ und $\lambda\phi$ gegeben durch $(\phi+\psi)(x) = \phi(x) + \psi(x)$ und $(\lambda\phi)(x) = \lambda\phi(x)$ Linearformen.

Beweis: Dies sieht man unmittelbar durch Einsetzen. Seien ϕ und ψ gegeben durch $\phi(x) = \sum_{j=1}^n a_j x_j$ und $\psi(x) = \sum_{j=1}^n b_j x_j$ mit $a_j, b_j \in K$ für $1 \le j \le n$. Dann gilt für jedes $x \in K^n$ jeweils

$$(\phi+\psi)(x) = \sum_{j=1}^n (a_j + b_j) x_j \qquad \text{und} \qquad (\lambda\phi)(x) = \sum_{j=1}^n (\lambda a_j) x_j .$$

Also sind auch $\phi + \psi$ und $\lambda\phi$ Linearformen. □

Definition 1.8 Unter einer ***elementaren Umformung*** eines LGS $((\phi_1, ..., \phi_m), b)$ versteht man eine der folgenden Operationen.

($\mathrm{M}_{k,\lambda}$) Ersetzung von ϕ_k durch $\lambda\phi_k$ und von b_k durch λb_k, für ein $k \in \{1, ..., m\}$ und ein $\lambda \in K^\times$

($\mathrm{A}_{k,\ell,\lambda}$) Ersetzung von ϕ_ℓ durch $\lambda\phi_k + \phi_\ell$ und von b_ℓ durch $\lambda b_k + b_\ell$, für $k, \ell \in \{1, ..., n\}$ mit $k \ne \ell$ und beliebiges $\lambda \in K$

Als weiteren Umformungstyp betrachtet man häufig noch

($\mathrm{V}_{k,\ell}$) Vertauschung von ϕ_k und ϕ_ℓ sowie von b_k und b_ℓ, für $k, \ell \in \{1, ..., n\}$ mit $k \ne \ell$

Diese Umformungen betrachten wir aber nicht als elementar, weil sie aus den Umformungen vom Typ ($\mathrm{M}_{k,\lambda}$) und ($\mathrm{A}_{k,\ell,\lambda}$) zusammengesetzt werden kann, wie das folgende Schema zeigt.

$$\begin{pmatrix} \phi_k, b_k \\ \phi_\ell, b_\ell \end{pmatrix} \xrightarrow{\mathrm{A}_{k,\ell,1}} \begin{pmatrix} \phi_k, b_k \\ \phi_k + \phi_\ell, b_k + b_\ell \end{pmatrix} \xrightarrow{\mathrm{M}_{k,-1}} \begin{pmatrix} -\phi_k, -b_k \\ \phi_k + \phi_\ell, b_k + b_\ell \end{pmatrix} \xrightarrow{\mathrm{A}_{\ell,k,1}} \begin{pmatrix} \phi_\ell, b_\ell \\ \phi_k + \phi_\ell, b_k + b_\ell \end{pmatrix} \xrightarrow{\mathrm{A}_{k,\ell,-1}} \begin{pmatrix} \phi_\ell, b_\ell \\ \phi_k, b_k \end{pmatrix}$$

Proposition 1.9 Sei $((\phi_1, ..., \phi_m), b)$ ein LGS und $((\phi'_1, ..., \phi'_m), b')$ ein weiteres, dass aus dem ersten durch Anwendung einer elementaren Umformung entsteht. Bezeichnen $\mathscr{L}, \mathscr{L}' \subseteq K^n$ die Lösungsmengen der beiden Systeme, dann gilt $\mathscr{L} = \mathscr{L}'$.

Beweis: Zunächst betrachten wir den Fall, dass $\mathscr{L}'$ durch Anwendung einer elementaren Umformung vom Typ ($\mathrm{M}_{k,\lambda}$) entsteht. Es gilt dann $\phi'_k = \lambda\phi_k$ und $b'_k = \lambda b_k$ sowie $\phi'_i = \phi_i$ und $b'_i = b_i$ für alle $i \ne k$. Für alle $v \in K^n$ und alle $i \ne k$ gilt dann offenbar $\phi_i(v) = b_i \Leftrightarrow \phi'_i(v) = b'_i$, und ebenso

$$\phi_k(v) = b_k \quad \Leftrightarrow \quad \lambda\phi_k(v) = \lambda b_k \quad \Leftrightarrow \quad \phi'_k(v) = \lambda b'_k .$$

Insgesamt erhalten wir damit

$$v \in \mathscr{L} \quad \Leftrightarrow \quad \phi_i(v) = b_i \text{ für } 1 \le i \le m \quad \Leftrightarrow \quad \phi'_i(v) = b'_i \text{ für } 1 \le i \le m \quad \Leftrightarrow \quad v \in \mathscr{L}'$$

und somit $\mathscr{L} = \mathscr{L}'$. Betrachten wir nun den Fall einer Umformung vom Typ $(\mathrm{A}_{k,\ell,\lambda})$. Hier gilt $\phi_i = \phi'_i$ und $b_i = b'_i$ für alle $i \neq \ell$, außerdem $\phi'_\ell = \lambda\phi_k + \phi_\ell$ und $b'_\ell = \lambda b_k + b_\ell$. Für jedes $v \in K^n$ und $i \neq k, \ell$ ist die Äquivalenz $\phi_i(v) = b_i \Leftrightarrow \phi'_i(v) = b'_i$ offensichtlich. Für die Indizes k, ℓ gilt

$$\phi_k(v) = b_k \wedge \phi_\ell(v) = b_\ell \quad \Leftrightarrow \quad \phi_k(v) = b_k \wedge \lambda\phi_k(v) + \phi_\ell(v) = \lambda b_k + b_\ell \quad \Leftrightarrow \quad \phi'_k(v) = b'_k \wedge \phi'_\ell(v) = b'_\ell.$$

Also ist $\mathscr{L} = \mathscr{L}'$ auch für diesen Umformungstyp erfüllt. □

Eine Umformungen, die die Lösungsmenge eines LGS nicht ändert, wird auch ***Äquivalenzumformung*** genannt. Wir werden nun sehen, wie man ein LGS noch kompakter darstellen kann.

Definition 1.10 Seien $m, n \in \mathbb{N}$. Eine ***$m \times n$ - Matrix*** über K ist eine Abbildung

$$A : \{1, ..., m\} \times \{1, ..., n\} \longrightarrow K.$$

Dabei nennt man $A(i, j)$ den ***Eintrag*** von A an der Stelle (i, j). Die Menge aller $m \times n$-Matrizen über K wird mit $\mathscr{M}_{m\times n,K}$ bezeichnet. An Stelle von $\mathscr{M}_{n\times n,K}$ schreibt man auch kürzer $\mathscr{M}_{n,K}$. Die Elemente dieser Menge werden als ***quadratische*** Matrizen bezeichnet.

Durch die Gleichung $A = (a_{ij})$ ordnet man dem Eintrag $A(i, j)$ der Matrix A die Bezeichnung a_{ij} zu. Allgemein legen wir folgende Konvention fest: Bezeichnet ein Großbuchstabe wie zum Beispiel A, B, C eine Matrix, dann bezeichnen die indizierten Kleinbuchstaben a_{ij}, b_{ij}, c_{ij} immer automatisch die Einträge dieser Matrix. Man kann eine Matrix auch auf übersichtliche Weise als rechteckiges Schema der Form

$$A \quad = \quad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \qquad \text{dargestellen.}$$

Allgemein nennt man $a_{i\bullet} = (a_{i1}, ..., a_{in}) \in K^n$ die ***i-te Zeile*** und $a_{\bullet j} = (a_{1j}, ..., a_{mj}) \in K^m$ die ***j-te Spalte*** von A. Im weiteren Verlauf wird es sich als praktisch erweisen, auch die Elemente aus K^n als spezielle Matrizen mit nur einer Spalte anzusehen. Wir legen deshalb fest, dass von nun an $K^n = \mathscr{M}_{n\times 1,K}$ für alle $n \in \mathbb{N}$ gilt.

Zur Beschreibung der Einträge verwendet man häufig als hilfreiche Abkürzung das sogenannte ***Kronecker-Delta***. Für jeden Körper K und beliebige $m, n \in \mathbb{N}$ definiert man

$$\delta_{mn,K} \quad = \quad \begin{cases} 1_K & \text{falls} \quad m = n \\ 0_K & \text{falls} \quad m \neq n. \end{cases}$$

Falls aus dem Kontext geschlossen werden kann, welcher Körper gemeint ist, wird der Index K auch oft weggelassen. Die folgenden konkreten Beispiele für Matrizen werden uns in den Anwendungen immer wieder begegnen.

(i) die ***Nullmatrix*** $\mathrm{O} = \mathrm{O}^{(m\times n)}$ in $\mathcal{M}_{m\times n,K}$, deren sämtliche Einträge gleich 0_K sind (Wir bezeichnen mit $\mathrm{O}^{(n)} = \mathrm{O}^{(n\times n)}$ die quadratische Nullmatrix.)

(ii) die ***Einheitsmatrix*** $\mathrm{I} = \mathrm{I}^{(n)}$ in $\mathcal{M}_{n,K}$ mit den Einträgen δ_{ij} für $1 \le i \le m$ und $1 \le j \le n$ (Die Einheitsmatrix ist also immer quadratisch.)

(iii) die ***Basismatrizen*** $\mathrm{B}_{k\ell} = \mathrm{B}_{k\ell}^{(m\times n)}$ mit den Einträgen $b_{ij} = \delta_{ik}\delta_{j\ell}$ für $1 \le i \le m$ und $1 \le j \le n$

Definition 1.11 Seien $A \in \mathcal{M}_{m\times n,K}$ und $v \in K^n$. Dann bezeichnet man den Vektor $w \in K^m$ mit den Komponenten w_i gegeben durch

$$w_i \quad = \quad \sum_{j=1}^{n} a_{ij} v_j \qquad \text{für} \quad 1 \le i \le m$$

als das ***Matrix-Vektor-Produkt*** Av von A und v.

Proposition 1.12 Das Matrix-Vektorprodukt erfüllt die Rechenregeln $A(v+w) = Av + Aw$ und $A(\lambda v) = \lambda(Av)$ für beliebige Matrixen $A \in \mathcal{M}_{m\times n,K}$ und Vektoren $v, w \in K^n$.

Beweis: Wir rechnen beide Gleichungen komponentenweise nach. Sei $A = (a_{ij})$, $v' = Av$, $w' = Aw$, $u = A(v+w)$ und $z = A(\lambda v)$. Zu zeigen ist $u = v' + w'$ und $z = \lambda v'$. Nach Definition des Matrix-Vektor-Produkts gelten die Gleichungen

$$v'_i = \sum_{j=1}^{n} a_{ij} v_j \quad , \quad w'_i = \sum_{j=1}^{n} a_{ij} w_j \quad , \quad u_i = \sum_{j=1}^{n} a_{ij}(v_j + w_j) \quad , \quad z_i = \sum_{j=1}^{n} a_{ij}(\lambda v_j).$$

für $1 \le i \le m$. Daraus folgt

$$u_i \quad = \quad \sum_{j=1}^{n} a_{ij}(v_j + w_j) \quad = \quad \sum_{j=1}^{n} a_{ij} v_j + \sum_{j=1}^{n} a_{ij} w_j \quad = \quad v'_i + w'_i$$

und ebenso $z_i = \sum_{j=1}^{n} a_{ij}(\lambda v_j) = \lambda \sum_{j=1}^{n} a_{ij} v_j = \lambda v'_i$ für $1 \le i \le m$, also $u = v' + w'$ und $z = \lambda v'$.

Jedem Paar (A, b) bestehend aus einer Matrix $A \in \mathcal{M}_{m\times n,K}$ und einem Vektor $b \in K^m$ kann ein LGS der Form $((\phi_1, \ldots, \phi_m), b)$ zugeordnet werden, indem man die Linearformen ϕ_i durch

$$\phi_i(x) \quad = \quad \sum_{j=1}^{n} a_{ij} x_j \qquad \text{für} \quad 1 \le i \le m \quad \text{definiert.}$$

Die Lösungsmenge des LGS ist dann durch $\mathscr{L} = \{ v \in K^n \mid Av = b \}$ gegeben. Man bezeichnet A als die ***Koeffizientenmatrix*** des LGS. Die Matrix $\tilde{A} = (A\ b) \in \mathscr{M}_{m\times(n+1),K}$, deren erste n Spalten mit denen von A und deren letzte Spalte mit b übereinstimmt, nennt man die ***erweiterte*** Koeffizientenmatrix des Systems.

Die Aussagen in Prop. (1.5) und Prop. (1.6) ergeben sich nun auch unmittelbar aus den Rechenregeln für das Matrix-Vektor-Produkt. Jede elementare Umformung des LGS entspricht einer Umformung der Matrix A bei gleichzeitiger Umformung des Vektors b. Beispielsweise wird in $(\mathrm{M}_{k,\lambda})$ die k-te Zeile von A und die k-te Komponente von b mit dem Wert λ multipliziert. Bei $(\mathrm{A}_{k,\ell,\lambda})$ addiert man das λ-fache der k-ten Zeile von A zur ℓ-ten Zeile von A und ebenso das λ-fache der Komponente b_k des Vektors b zur Komponente b_ℓ.

§ 2. Das Gaußsche Eliminationsverfahren

In diesem Abschnitt sehen wir uns an, wie die Lösungsmenge eines LGS gegeben durch ein Paar (A, b) bestehend aus einer Matrix $A \in \mathscr{M}_{m\times n,K}$ und einem Vektor $b \in K^m$ ausgerechnet werden kann. Befindet sich die Matrix in einer besonders einfachen Form, dann kann die Lösungsmenge direkt abgelesen werden.

Definition 2.1 Eine Matrix $A \in \mathscr{M}_{m\times n,K}$ befindet sich in ***Zeilenstufenform*** (kurz ZSF), wenn $A = \mathrm{O}^{(m\times n)}$ gilt oder folgende Bedingung erfüllt ist: Es gibt ein $r \in \{1,...,m\}$ und $j_1,..., j_r \in \{1,...,n\}$ mit $j_1 < j_2 < ... < j_r$, so dass

(i) $a_{i j_i} \neq 0_K$ für $1 \le i \le r$ und
(ii) $a_{i j} = 0_K$ für $j < j_i$ oder $i > r$

erfüllt ist. Man nennt r den ***Zeilenrang*** einer solchen Matrix. Das Tupel $(r, j_1,..., j_r)$ bezeichnen wir insgesamt als die ***Kennzahlen*** der ZSF.

Die Positionen (i, j_i) mit $1 \le i \le r$ in der Matrix werden ***Zeilenköpfe*** genannt. Die Bedingung (i) besagt, dass die Einträge in den Zeilenköpfen ungleich Null sind. Nach Bedingung (ii) befinden sich links von den Zeilenköpfen nur Nulleinträge; in den „kopflosen" Zeilen sind alle Einträge gleich Null. Der Zeilenrang kann offenbar nie größer als $\min\{m, n\}$ werden.

Definition 2.2 Eine Matrix $A \in \mathscr{M}_{m\times n,K}$ befindet sich in ***normierter*** ZSF, wenn $A = \mathrm{O}^{(m\times n)}$ gilt oder wenn sie in ZSF mit den Kennzahlen $(r, j_1,..., j_r)$ vorliegt uns außerdem die folgenden Bedingungen erfüllt sind: Es gilt $a_{i j_i} = 1_K$ für $1 \le i \le r$ und $a_{k j_i} = 0_K$ für $1 \le i \le r$ und $1 \le k < i$.

Bei der normierten ZSF kommen also folgende Bedingungen hinzu: Die Einträge in den Zeilenköpfen sind gleich 1_K, und oberhalb der Zeilenköpfe befinden sich nur Nulleinträge. Bei einer Matrix A in normierter ZSF gilt also insgesamt $a_{i j} = 0$ in jedem der drei Fälle

(1) $i > r$ (2) $i \le r$ und $j < j_i$ (3) $i \le r$ und $j = j_k$ für ein $k \in \{1,...,r\} \setminus \{i\}$

wobei aber durchaus noch weitere Einträge von A gleich Null sein können. Wir bemerken noch, dass eine Matrix $A \in \mathscr{M}_{m\times n,K}$ in normierter ZSF mit Zeilenrang $r = n$ in den oberen n Zeilen mit der Einheitsmatrix $\mathrm{I}^{(n)}$ übereinstimmt, denn in diesem Fall muss $j_k = k$ für $1 \le k \le n$ gelten.

Definition 2.3 Für $1 \le \ell \le n$ bezeichnet $e_\ell \in K^n$ jeweils den ***ℓ-te Einheitsvektor*** mit den Einträgen $e_{\ell j} = \delta_{\ell j}$.

In K^3 sind die drei Einheitsvektoren beispielsweise gegeben durch

$$e_1 = (1_K, 0_K, 0_K) \quad , \quad e_2 = (0_K, 1_K, 0_K) \quad \text{und} \quad e_3 = (0_K, 0_K, 1_K).$$

Sei nun A eine Matrix in normierter ZSF mit Kennzahlen $(r, j_1, ..., j_r)$. Unser erstes Ziel in diesem Abschnitt besteht darin, die Lösungsmenge $\mathscr{L}^{\mathrm{h}}$ des homogenen LGS mit Koeffizientenmatrix A explizit anzugeben. Dazu definieren wir

$$S = \{1, ..., n\} \setminus \{j_1, ..., j_r\}$$

und definieren für jede Zahl $\ell \in S$ einen Vektor $b_\ell \in K^m$ durch

$$b_\ell \quad = \quad e_\ell - \sum_{k=1}^{r} a_{k\ell} e_{j_k}.$$

Der Vektor b_ℓ entsteht also aus dem Nullvektor dadurch, dass man die ℓ-te Komponente auf 1_K setzt und die Einträge $-a_{1\ell}, ..., -a_{r\ell}$ der ℓ-ten Spalte auf die Positionen $j_1, ..., j_r$ des Vektors verteilt. Mit Hilfe dieser Vektoren lässt sich nun die Lösungsmenge $\mathscr{L}^{\mathrm{h}}$ folgendermaßen darstellen.

Satz 2.4 Sei $\mathscr{L}^{\mathrm{h}} \subseteq K^n$ die Lösungsmenge eines homogenen LGS mit Koeffizientenmatrix A, und seien S und die Vektoren b_ℓ für $\ell \in S$ definiert wie oben.

(i) Im Fall $S = \emptyset$ gilt $\mathscr{L}^{\mathrm{h}} = \{ 0_{K^n} \}$.

(ii) Ist S nichtleer, dann ist die Lösungsmenge gegeben durch

$$\mathscr{L}^{\mathrm{h}} \quad = \quad \left\{ \sum_{\ell \in S} \lambda_\ell b_\ell \;\middle|\; \lambda_\ell \in K \;\forall\; \ell \in S \right\}.$$

Beweis: zu (i) Unter dieser Vorausssetzung gilt $\{j_1, ..., j_r\} = \{1, ..., n\}$, woraus wiederum $r = n$ und somit $m \ge n$ folgt. Wie oben ausgeführt, stimmen die ersten n Zeilen von A mit der Einheitsmatrix $\mathrm{I}^{(n)}$ überein. Es gilt also $a_{ij} = \delta_{ij}$ für $1 \le i, j \le n$ und $a_{ij} = 0_K$ falls $i > n$. Für jeden Vektor $v \in K^n$ gilt nun die Äquivalenz

$$v \in \mathscr{L}^{\mathrm{h}} \quad \Leftrightarrow \quad \sum_{j=1}^{n} a_{kj} v_j = 0_K \ \text{ für } \ 1 \le k \le m \quad \Leftrightarrow \quad \sum_{j=1}^{n} \delta_{kj} v_j = 0_K \ \text{ für } \ 1 \le k \le n \quad \Leftrightarrow$$
$$v_k = 0_K \ \text{ für } \ 1 \le k \le n \quad \Leftrightarrow \quad v = 0_{K^n}.$$

zu (ii) Auf Grund der Eigenschaften (1) und (3) der normierten ZSF gilt für alle $v \in K^n$ jeweils

$$v \in \mathscr{L}^{\mathrm{h}} \quad \Leftrightarrow \quad \sum_{j=1}^{n} a_{kj} v_j = 0_K \;\text{ für }\; 1 \le k \le m \quad \overset{(1)}{\Leftrightarrow} \quad \sum_{j=1}^{n} a_{kj} v_j = 0_K \;\text{ für }\; 1 \le k \le r$$

$$\overset{(3)}{\Leftrightarrow} \quad \sum_{j \in S \cup \{j_k\}} a_{kj} v_j = 0_K \quad \Leftrightarrow \quad v_{j_k} = -\sum_{j \in S} a_{kj} v_j \;\text{ für }\; 1 \le k \le r.$$

Für die oben konstruierten Vektoren b_ℓ mit $\ell \in S$ muss also die folgende Gleichung überprüft werden.

$$\left\{ v \in K^n \;\middle|\; v_{j_k} = -\sum_{\ell \in S} a_{k\ell} v_\ell \text{ für } 1 \le k \le r \right\} \quad = \quad \left\{ \sum_{\ell \in S} \lambda_\ell b_\ell \;\middle|\; \lambda_\ell \in K \;\forall\; \ell \in S \right\}$$

„$\subseteq$“ Sei $v \in K^n$ ein Vektor mit $v_{j_k} = -\sum_{\ell \in S} a_{k\ell} v_\ell$ für $1 \le k \le r$. Wir definieren $\lambda_\ell = v_\ell$ für alle $\ell \in S$ und setzen $w = \sum_{\ell \in S} \lambda_\ell b_\ell$. Auf Grund der Formel für die Komponenten der Vektoren b_ℓ gilt für $1 \le k \le r$ gilt dann jeweils

$$w_{j_k} \quad = \quad \sum_{\ell \in S} \lambda_\ell b_{\ell j_k} \quad = \quad \sum_{\ell \in S} \lambda_\ell (-a_{k\ell}) \quad = \quad -\sum_{\ell \in S} a_{k\ell} v_\ell \quad = \quad v_{j_k}.$$

Weil für jedes $\ell \in S$ der Vektor b_ℓ an der ℓ-ten Komponente gleich 1_K und an den Stellen $j \in S \setminus \{\ell\}$ gleich 0_K ist, gilt außerdem

$$w_j \quad = \quad \sum_{\ell \in S} \lambda_\ell b_{\ell j} \quad = \quad \sum_{\ell \in S} \lambda_\ell \delta_{\ell j} \quad = \quad \lambda_j \quad = \quad v_j$$

für alle $j \in S$. Insgesamt gilt also $v = w$, und damit ist v in der Menge auf der rechten Seite enthalten.

„$\supseteq$“ Sei $v = \sum_{\ell \in S} \lambda_\ell b_\ell$ mit $\lambda_\ell \in K$ für alle $\ell \in S$. Dann sind die Komponenten von v gegeben durch $v_{j_k} = \sum_{\ell \in S} (-\lambda_\ell) a_{k\ell}$ für $1 \le k \le r$ und $v_j = \sum_{\ell \in S} \lambda_j \delta_{\ell j} = \lambda_j$ für $j \in S$. Wir erhalten für $1 \le k \le r$ also

$$v_{j_k} \quad = \quad \sum_{\ell \in S} (-\lambda_\ell) a_{k\ell} \quad = \quad -\sum_{\ell \in S} a_{k\ell} v_\ell \quad ,$$

und damit ist v in der Menge auf der linken Seite enthalten. □

Als nächstes sehen wir uns nun an, wie man eine Lösung für ein beliebiges, möglicherweise inhomogenes LGS findet. Nach Prop. (1.6) und dem soeben bewiesenen Satz genügt dies, um die gesamte Lösungsmenge des LGS zu bestimmen.

Satz 2.5 Sei also $\tilde{A} = (A\ b) \in \mathscr{M}_{m \times (n+1),K}$ die erweiterte Koeffizientenmatrix eines LGS und $\mathscr{L} \subseteq K^n$ dessen Lösungsmenge. Wir setzen voraus, dass $\tilde{A}$ in normierte ZSF vorliegt, mit den Kennzahlen $(r, j_1, ..., j_r)$.

(i) Ist $j_r = n + 1$, dann gilt $\mathscr{L} = \emptyset$.

(ii) Sei nun $j_r \le n$. Wir definieren einen Vektor $w \in K^n$ durch $\sum_{k=1}^{r} b_k e_{j_k}$, mit den Komponenten $w_{j_k} = b_k$ für $1 \le k \le r$ und $w_j = 0_K$ für $j \in S$. Dann gilt $w \in \mathscr{L}$.

Der spezielle Lösungsvektor w entsteht also einfach dadurch, dass man die Werte $b_1, ..., b_r$ auf die Positionen $j_1, ..., j_r$ verteilt und die übrigen Komponenten auf Null setzt.

Beweis: zu (i) Nehmen wir an, dass $\mathscr{L}$ nichtleer und w ein Element aus $\mathscr{L}$ ist. Dann gilt insbesondere $\sum_{j=1}^n a_{rj} w_j = b_r$. Wegen $j_r = n+1$ gilt aber $a_{rj} = 0_K$ für $1 \le j \le n$ und $b_r = 1_K$. Setzen wir dies in die Gleichung ein, so erhalten wir $\sum_{j=1}^n 0_K w_j = 1_K$. Der Widerspruch $0_K = 1_K$ zeigt, dass unsere Annahme falsch war.

zu (ii) Zu zeigen ist $\sum_{j=1}^n a_{kj} w_j = b_k$ für $1 \le k \le m$. Für $1 \le k \le r$ gilt nach Definition des Vektors w und auf Grund der Eigenschaft (3) der normierten ZSF jeweils

$$\sum_{j=1}^n a_{kj} w_j \quad = \quad \sum_{i=1}^r a_{kj_i} w_{j_i} \quad \overset{(3)}{=} \quad a_{kj_k} b_k \quad = \quad b_k$$

Für $r < k \le m$ gilt nach Eigenschaft (2) der normierten ZSF sowohl $a_{kj} = 0$ für $1 \le j \le n$ als auch $b_k = 0$. Also sind die Gleichungen auch für diese k erfüllt. □

Um ein Lösungsverfahren für beliebige LGS zu bekommen, brauchen wir also nur noch ein Verfahren, mit dem wir beliebige Matrizen in normierte Zeilenstufenform überführen können. Um dieses Problem auf systematische Weise lösen zu können, beschäftigen wir uns zunächst allgemein mit Rechenoperationen für Matrizen.

Definition 2.6 Sei K ein Körper.

(i) Seien $m, n \in \mathbb{N}$ und $A, B \in \mathscr{M}_{m\times n,K}$ mit Einträgen $A = (a_{ij})$ und $B = (b_{ij})$. Dann nennt man die Matrix $C = (c_{ij})$ mit $c_{ij} = a_{ij} + b_{ij}$ für $1 \le i \le m$ und $1 \le j \le n$ die ***Summe*** von A und B. Wir bezeichnen diese Matrix mit $A + B$.

(ii) Sei $A \in \mathscr{M}_{m\times n,K}$ mit $A = (a_{ij})$ und $\lambda \in K$. Dann ist die Matrix $C = (c_{ij})$ mit $c_{ij} = \lambda a_{ij}$ ein ***skalares Vielfaches*** von A, das wir mit λA bezeichnen.

(iii) Seien nun $m, n, r \in \mathbb{N}$ und $A \in \mathscr{M}_{m\times n,K}$, $B \in \mathscr{M}_{n\times r,K}$. Dann heißt die Matrix $C \in \mathscr{M}_{m\times r,K}$ mit den Einträgen

$$c_{ij} \quad = \quad \sum_{k=1}^n a_{ik} b_{kj}$$

Produkt der Matrizen A und B und wird mit AB bezeichnet.

(iv) Sei $A \in \mathscr{M}_{m\times n,K}$. Die Matrix $B \in \mathscr{M}_{n\times m,K}$ mit den Einträgen $b_{ij} = a_{ji}$ für $1 \le i \le n$ und $1 \le j \le m$ wird die zu A ***transponierte*** Matrix ${}^{\mathrm{t}}A$ genannt.

Die Berechnung des Produkts ähnelt dem Matrix-Vektor-Produkt aus dem vorletzten Abschnitt. Um den Eintrag c_{ij} der Matrix $C = AB$ zu erhalten, muss die i-te Zeile von A mit der j-ten Spalte von B multipliziert werden. Man beachte, dass das Produkt AB nur gebildet werden kann, wenn die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt. Die Summe $A + B$ ist nur dann definiert, wenn A und B dieselbe Anzahl Zeilen ***u***nd Spalten besitzen.

Proposition 2.7 Seien $A, A' \in \mathscr{M}_{m\times n,K}$, $B, B' \in \mathscr{M}_{n\times r,K}$ und $C \in \mathscr{M}_{r\times s,K}$. Dann gilt

(i) $A(B+B') = AB + AB'$ und $(A+A')B = AB + A'B$

(ii) $A(\lambda B) = (\lambda A)B = \lambda(AB)$

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

(iv) ${}^{\mathrm{t}}(AB) = {}^{\mathrm{t}}B\,{}^{\mathrm{t}}A$

Beweis: zu (i) Den Beweis der ersten Formel überlassen wir dem Leser als Übungsaufgabe und beweisen statt dessen direkt die zweite. Es sei $C = A + A'$, $D = CB$, $F = AB$, $G = A'B$ und $H = F + G$. Dann ist $D = H$ zu zeigen. Es gilt $c_{ij} = a_{ij} + a'_{ij}$ und

$$d_{ij} \quad = \quad \sum_{k=1}^{n} c_{ik} b_{kj} \quad = \quad \sum_{k=1}^{n} (a_{ik} + a'_{ik}) b_{kj} \quad = \quad \sum_{k=1}^{n} a_{ik} b_{kj} + \sum_{k=1}^{n} a'_{ik} b_{kj}.$$

Für die Einträge der Matrizen F und G erhalten wir

$$f_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad \text{und} \qquad g_{ij} = \sum_{k=1}^{n} a'_{ik} b_{kj}.$$

Es folgt

$$h_{ij} \quad = \quad f_{ij} + g_{ij} \quad = \quad \sum_{k=1}^{n} a_{ik} b_{kj} + \sum_{k=1}^{n} a'_{ik} b_{kj} \quad = \quad d_{ij}$$

also insgesamt $D = H$.

zu (ii) Wir definieren $C = \lambda B$, $D = AC$, $F = \lambda A$, $G = FB$, $H = AB$ und $U = \lambda H$. Zu zeigen ist dann $D = G = U$. Nach Definition gilt $c_{ij} = \lambda b_{ij}$ und

$$d_{ij} \quad = \quad \sum_{k=1}^{n} a_{ik} c_{kj} \quad = \quad \sum_{k=1}^{n} \lambda a_{ik} b_{kj}.$$

Andererseits gilt auch $f_{ij} = \lambda a_{ij}$ und $g_{ij} = \sum_{k=1}^{n} f_{ik} b_{kj} = \sum_{k=1}^{n} \lambda a_{ik} b_{kj} = d_{ij}$, womit die Gleichung $D = G$ bewiesen ist. Nun gilt außerdem $h_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ und

$$u_{ij} \quad = \quad \lambda h_{ij} \quad = \quad \sum_{k=1}^{n} \lambda a_{ik} b_{kj} \quad = \quad d_{ij} \quad ,$$

wodurch auch die Gleichung $U = D$ bewiesen ist.

zu (iii) Wir definieren $D = AB$, $F = DC$, $G = BC$ und $H = AG$. Dann gilt $d_{k\ell} = \sum_{i=1}^{n} a_{ki} b_{i\ell}$, und für die Matrix F erhalten wir

$$f_{k\ell} \quad = \quad \sum_{i=1}^{r} d_{ki} c_{i\ell} \quad = \quad \sum_{i=1}^{r} \sum_{j=1}^{n} a_{kj} b_{ji} c_{i\ell}.$$

Andererseits hat G die Einträge $g_{k\ell} = \sum_{i=1}^{r} b_{ki} c_{i\ell}$, und für die Matrix H gilt

$$h_{k\ell} \quad = \quad \sum_{i=1}^{n} a_{ki} g_{i\ell} \quad = \quad \sum_{i=1}^{n} \sum_{j=1}^{r} a_{ki} b_{ij} c_{j\ell} \quad , \quad \text{also insgesamt} \quad F = H.$$

zu (iv) Hier definieren wir die Hilfsmatrizen $C = AB$, $D = {}^{\mathrm{t}}C$, $F = {}^{\mathrm{t}}A$, $G = {}^{\mathrm{t}}B$ und $H = GF$. Dann müssen wir $D = H$ nachrechnen. Es gilt

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad \text{und} \qquad d_{ij} = c_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

Wegen $f_{ij} = a_{ji}$ und $g_{ij} = b_{ji}$ gilt außerdem

$$h_{ij} \quad = \quad \sum_{k=1}^{n} g_{ik} f_{kj} \quad = \quad \sum_{k=1}^{n} b_{ki} a_{jk} \quad = \quad d_{ij}$$

also $H = D$ wie gewünscht. □

Proposition 2.8 Sei $A \in \mathscr{M}_{m \times n,K}$. Dann gilt $\mathrm{I}^{(m)} A = A\mathrm{I}^{(n)} = A$.

Beweis: Sei $B = \mathrm{I}^{(m)} A$. Für $1 \le i \le m$ und $1 \le j \le n$ gilt dann

$$b_{ij} \quad = \quad \sum_{k=1}^{m} \delta_{ik} a_{kj} \quad = \quad a_{ij}.$$

Damit ist $B = A$ bewiesen. Sei nun $C = A\mathrm{I}^{(n)}$. Für $1 \le i \le m$ und $1 \le j \le n$ gilt dann ebenfalls

$$c_{ij} \quad = \quad \sum_{k=1}^{n} a_{ik} \delta_{kj} \quad = \quad a_{ij} \quad , \qquad \text{also} \quad C = A. \qquad \square$$

Bei Rechnungen mit Matrizen ist es oft günstig, diese in mehrere Bereiche aufzuteilen. Sei $A \in \mathscr{M}_{m \times n,K}$, seien k_1, k_2, ℓ_1, ℓ_2 natürliche Zahlen mit $1 \le k_1 \le k_2 \le m$, $1 \le \ell_1 \le \ell_2 \le n$, und außerdem $r = k_2 - k_1 + 1$, $s = \ell_2 - \ell_1 + 1$. Dann nennt man die Matrix $B \in \mathscr{M}_{r \times s,K}$ mit den Einträgen $b_{ij} = a_{k_1+i-1,\ell_1+j-1}$ für $1 \le i \le r$, $1 \le j \le s$ eine ***Teilmatrix*** von A; es handelt sich um einen „rechteckigen Ausschnitt“ der Matrix A.

Häufig verwendet man die sogenannte ***Blockschreibweise***, um Matrizen darzustellen, die aus bestimmten Teilmatrizen aufgebaut sind. So steht beispielsweise der Ausdruck

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

für die Matrix, deren linker oberer Teil aus den Einträgen von A und entsprechend in den übrigen drei Bereichen aus den Einträgen von B, C und D besteht. Dabei wird voraussetzt, dass untereinander stehende Matrizen (hier: A, C bzw. B, D) stets dieselbe Spaltenzahl und nebeneinander stehende Matrizen (A, B bzw. C, D) dieselbe Zeilenzahl haben. Das Rechnen mit Matrizen in Blockschreibweise wird durch eine Reihe von Rechenregeln vereinfacht.

Proposition 2.9 Seien A, B, C, D Matrizen über K.

(i) Stimmt die Spaltenzahl von A und B mit der Zeilenzahl von C überein, dann gilt

$$\begin{pmatrix} A \\ B \end{pmatrix} C \quad = \quad \begin{pmatrix} AC \\ BC \end{pmatrix}.$$

(ii) Stimmt die Spaltenzahl von A mit der Zeilenzahl von B und C überein, dann gilt

$$A\begin{pmatrix} B & C \end{pmatrix} \quad = \quad \begin{pmatrix} AB & AC \end{pmatrix}.$$

(iii) Stimmt die Spaltenzahl von A mit der Zeilenzahl von C und die Spaltenzahl von B mit der Zeilenzahl von D überein, dann gilt

$$\begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} \quad = \quad AC + BD.$$

Beweis: Wir beschränken uns auf den Beweis von (iii). Nach Voraussetzung gilt $A \in \mathscr{M}_{m\times n_1,K}$, $B \in \mathscr{M}_{m\times n_2,K}$, $C \in \mathscr{M}_{n_1\times r,K}$ und $D \in \mathscr{M}_{n_2\times r,K}$ für geeignete $m, n_1, n_2, r \in \mathbb{N}$. Die Matrix $AC + BD$ auf der rechten Seite ist in $\mathscr{M}_{m\times r,K}$ enthalten. Seien nun k, ℓ mit $1 \le k \le m$ und $1 \le \ell \le r$ vorgegeben. Zu zeigen ist, dass der Eintrag des Matrixprodukts links an der Position (k, ℓ) mit dem Eintrag der Matrix $AC + BD$ an derselben Stelle übereinstimmt. Um den Eintrag auf der linken Seite auszurechnen, muss die k-te Zeile des Faktors $(A \ \ B)$ mit der ℓ-ten Spalte des zweiten Faktors multipliziert werden. Dies liefert den Wert

$$\sum_{j=1}^{n_1} a_{kj} c_{j\ell} + \sum_{j=1}^{n_2} b_{kj} d_{j\ell}.$$

Die erste Summe entspricht dem Eintrag von AC an der Stelle (k, ℓ), die zweite Summe dem Eintrag von BD an derselben Position. Insgesamt erhalten wir also den Eintrag von $AC + BD$ an der Stelle (k, ℓ). □

Aus Prop. (2.9) kann durch vollständige Induktion abgeleitet werden, dass Matrizen mit beliebiger Aufteilung „blockweise“ multipliziert werden können, wobei lediglich vorausgesetzt werden muss, dass die Teilmatrizen, die dabei multipliziert werden sollen, „zusammenpassen“. Genauer lässt sich dies folgendermaßen formulieren: Seien $m, n, r \in \mathbb{N}$, seien A_{ij} für $1 \le i \le m$, $1 \le j \le n$ und B_{jk} für $1 \le j \le n$ und $1 \le k \le r$ Matrizen mit der Eigenschaft, dass die Spaltenzahl von A_{ij} jeweils mit der Zeilenzahl von B_{jk} übereinstimmt, für alle i, j, k. Dann gilt

$$\begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \vdots \\ B_{n1} & \cdots & B_{nr} \end{pmatrix} \quad = \quad \begin{pmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & & \vdots \\ C_{m1} & \cdots & C_{mr} \end{pmatrix}$$

mit $C_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk}$ für $1 \le i \le m$ und $1 \le k \le r$.

Definition 2.10 Eine Matrix in $\mathscr{M}_{m,K}$ der Form $\mathrm{M}_{k,\lambda} = \mathrm{I}^{(m)} + (\lambda - 1)\mathrm{B}_{kk}^{(m\times m)}$ für $k \in \{1,\ldots,m\}$ und $\lambda \in K^\times$ und $\mathrm{A}_{k,\ell,\lambda} = \mathrm{I}^{(m)} + \lambda \mathrm{B}_{\ell k}^{(m\times m)}$ für $k,\ell \in \{1,\ldots,m\}$ und $\lambda \in K$ werden ***Elementarmatrizen*** genannt.

In Blockschreibweise hat die Elementarmatrix $\mathrm{M}_{k,\lambda}$ die Form

$$\mathrm{M}_{k,\lambda} \quad = \quad \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \lambda & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(m-k)} \end{pmatrix}$$

wobei die Einträge O jeweils für Nullmatrizen der passenden Größe stehen. Die Elementarmatrix $\mathrm{A}_{k,\ell,\lambda}$ hat im Fall $k < \ell$ bzw. $k > \ell$ die Form

$$\mathrm{A}_{k,\ell,\lambda} \quad = \quad \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & 1 & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(\ell-k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \lambda & \mathrm{O} & 1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{I}^{(m-\ell)} \end{pmatrix}$$

beziehungsweise

$$\mathrm{A}_{k,\ell,\lambda} \quad = \quad \begin{pmatrix} \mathrm{I}^{(\ell-1)} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & 1 & \mathrm{O} & \lambda & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(k-\ell-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & 1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{I}^{(m-k)} \end{pmatrix}$$

Proposition 2.11 Sei $A \in \mathscr{M}_{m\times n,K}$.

(i) Sei $\lambda \in K^\times$ und $k \in \{1,\ldots,m\}$. Multipliziert man die Matrix A von links mit der Elementarmatrix $\mathrm{M}_{k,\lambda}$, so bewirkt dies eine Multiplikation der k-ten Zeile mit dem Wert λ.

(ii) Seien $k,\ell \in \{1,\ldots,m\}$ mit $k \neq \ell$ und $\lambda \in K$. Multipliziert man die Matrix A mit der Elementarmatrix $\mathrm{A}_{k,\ell,\lambda}$, dann wird das λ-fache der k-ten Zeile zur ℓ-ten Zeile von A addiert.

Beweis: Beide Aussagen lassen sich durch die blockweise Multiplikation von Matrizen unmittelbar nachrechnen.

zu (i) Sei $B \in \mathscr{M}_{(k-1)\times n,K}$ die Teilmatrix bestehend aus den oberen $k-1$ und $C \in \mathscr{M}_{(m-k)\times n,K}$ die Teilmatrix bestehend aus den unteren $m-k$ Zeilen von A. Ferner sei $z \in \mathscr{M}_{1\times n,K}$ die k-te Zeile von A. Dann gilt

$$\mathrm{M}_{k,\lambda} A \quad = \quad \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \lambda & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(m-k)} \end{pmatrix} \begin{pmatrix} B \\ z \\ C \end{pmatrix} \quad = \quad \begin{pmatrix} \mathrm{I}^{(k-1)} B + \mathrm{O}z + \mathrm{O}C \\ \mathrm{O}B + \lambda z + \mathrm{O}C \\ \mathrm{O}B + \mathrm{O}z + \mathrm{I}^{(m-k)} C \end{pmatrix} \quad = \quad \begin{pmatrix} B \\ \lambda z \\ C \end{pmatrix}$$

zu (ii) Hier beschränken wir uns auf den Fall $k < \ell$ und teilen die Matrix A auf in die Matrix $B \in \mathscr{M}_{(k-1)\times n,K}$ bestehend aus den ersten $k-1$ Zeilen, der Matrix $C \in \mathscr{M}_{(\ell-k-1)\times n,K}$ bestehend aus der $(k+1)$-ten bis zur $(\ell-1)$-ten Zeile und der Matrix $D \in \mathscr{M}_{(m-\ell)\times n,K}$ bestehend aus den unteren $m-\ell$ Zeilen. Ferner seien $z_k, z_\ell \in \mathscr{M}_{1\times n,K}$ die k-te und ℓ-te Zeile von A. Dann erhalten wir

$$\mathrm{A}_{k,\ell,\lambda} A \quad = \quad \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & 1 & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(\ell-k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \lambda & \mathrm{O} & 1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{I}^{(m-\ell)} \end{pmatrix} \begin{pmatrix} B \\ z_k \\ C \\ z_\ell \\ D \end{pmatrix} \quad = \quad \begin{pmatrix} B \\ z_k \\ C \\ \lambda z_k + z_\ell \\ D \end{pmatrix} \qquad \square$$

Jede Zeilenumformung einer Matrix A lässt sich also durch Multiplikation mit einer Elementarmatrix von links realisieren. Dementsprechend führt die Multiplikation von A mit einem Produkt $E_m \cdot E_{m-1} \cdot ... \cdot E_1$ von Elementarmatrix dazu, dass A einer Folge von m Zeilenumformungen unterworfen wird. Wir bezeichnen die Menge aller Matrizen in $\mathscr{M}_{m,K}$, die sich als Produkt von Elementarmatrizen schreiben lassen, mit $\mathscr{E}_m(K)$.

Mit Hilfe des Matrixkalküls werden wir nun zeigen, dass sich jede Matrix durch eine endliche Anzahl von Zeilenumformungen auf normierte Zeilenstufenform bringen lässt.

Lemma 2.12 Sei $A \in \mathscr{M}_{m\times 1,K}$ eine Matrix, die aus einer einzigen Spalte besteht, also eine Matrix der Form $A = {}^{\mathrm{t}}(a_1\ a_2\ ...\ a_m)$. Sind nicht alle Einträge von A gleich Null, dann gibt es ein Produkt $E \in \mathscr{E}_m(K)$ von Elementarmatrizen mit $EA = {}^{\mathrm{t}}(1_K\ 0_K\ 0_K\ ...\ 0_K)$.

Beweis: Auf Grund unserer Vorbemerkung genügt es zu zeigen, dass A durch eine endliche Abfolge von elementaren Zeilenumformungen auf die Gestalt ${}^{\mathrm{t}}(1\ 0\ ...\ 0)$ gebracht werden kann. Auch Vertauschungen von Zeilen sind zulässig, weil diese (wie oben gesehen) durch endlich viele elementare Umformungen realsierbar sind. Nach Voraussetzung gibt es ein $k \in \{1,...,m\}$ mit $a_k \neq 0_K$. Nach Multiplikation der k-ten Zeile mit a_k^{-1} und Vertauschung der k-ten mit der ersten Zeile gilt $a_1 = 1_K$. Nun addieren wir für $\ell = 2,...,m$ jeweils das $(-a_\ell)$-fache der ersten Zeile zur ℓ-ten. Dies führt dazu, dass sämtliche Einträge der Matrix mit Ausnahme des ersten zu Null werden. □

Satz 2.13 Jede Matrix $A \in \mathscr{M}_{m\times n,K}$ kann durch endlich viele elementare Zeilenumformungen auf normierte ZSF gebracht werden. Eine äquivalente Formulierung dieser Aussage lautet: Es gibt eine Matrix $E \in \mathscr{E}_m(K)$, so dass EA in normierter ZSF vorliegt.

Beweis: Wir zeigen zunächst, dass A auf ZSF gebracht werden kann und führen den Beweis durch vollständige Induktion über die Anzahl n der Spalten. Der Fall $n = 1$ ist mit Lemma (2.12) bereits erledigt, denn nach Definition ist ${}^{\mathrm{t}}(1_K\ 0_K\ \dots\ 0_K)$ eine Matrix in ZSF (mit den Kennzahlen $r = j_1 = 1$). Sei nun $n \in \mathbb{N}$, und setzen wir die Aussage für dieses n voraus. Sei außerdem $A \in \mathscr{M}_{m\times(n+1),K}$ eine beliebige Matrix. Wir müssen zeigen, dass A auf ZSF gebracht werden kann und unterscheiden dafür zwei Fälle.

1. Fall: Die erste Spalte von A hat nur Nulleinträge.
Dann hat A die Form $(\mathrm{O}^{(m\times 1)}\ B)$ mit einer Matrix $B \in \mathscr{M}_{m\times n,K}$. Nach Induktionsvoraussetzung gibt es eine Matrix $E \in \mathscr{E}_m(K)$, so dass $B' = EB$ in ZSF vorliegt, mit gewissen Kennzahlen $r, j_1, \dots, j_r$. Es gilt

$$EA \quad = \quad E(\mathrm{O}^{(m\times 1)}\ B) \quad = \quad (\mathrm{O}^{(m\times 1)}\ EB) \quad = \quad (\mathrm{O}^{(m\times 1)}\ B').$$

Wie man leicht überprüft, liegt auch $(\mathrm{O}^{(m\times 1)}\ B')$ die Matrix in ZSF vor, mit den Kennzahlen $r, j_1+1, \dots, j_r+1$.

2. Fall: Die erste Spalte von A hat Einträge ungleich Null.
In diesem Fall kann A in der Blockgestalt

$$A \quad = \quad \begin{pmatrix} a_{11} & z \\ s & C \end{pmatrix}$$

dargestellt werden, mit $a_{11} \in K$, $z \in \mathscr{M}_{1\times n,K}$, $s \in \mathscr{M}_{(m-1)\times 1,K}$ und $C \in \mathscr{M}_{(m-1)\times n,K}$, wobei die Teilmatrix ${}^{\mathrm{t}}(a_{11}\, s)$ nicht nur Nulleinträge enthält. Nach Lemma (2.12) gibt es eine Matrix $E \in \mathscr{E}_m(K)$ mit

$$E\begin{pmatrix} a_{11} \\ s \end{pmatrix} \quad = \quad \begin{pmatrix} 1 \\ \mathrm{O} \end{pmatrix}$$

und wir erhalten

$$EA \quad = \quad \begin{pmatrix} 1 & z' \\ \mathrm{O} & C' \end{pmatrix}$$

mit geeigneten Matrizen $z' \in \mathscr{M}_{1\times n,K}$ und $C' \in \mathscr{M}_{(m-1)\times n,K}$. Nach Induktionsvoraussetzung existiert nun eine Matrix $E' \in \mathscr{E}_{m-1}(K)$, so dass $E'C'$ in ZSF vorliegt, mit gewissen Kennzahlen $r, j_1, \dots, j_r$. Außerdem gilt

$$\begin{pmatrix} 1 & \mathrm{O} \\ \mathrm{O} & E' \end{pmatrix}\begin{pmatrix} 1 & z' \\ \mathrm{O} & C' \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & z' \\ \mathrm{O} & E'C' \end{pmatrix}$$

Wieder überprüft man, dass sich die Matrix rechts in ZSF befindet, mit Kennzahlen $r+1, 1, j_1+1, \dots, j_r+1$. Anhand der Gleichung

$$\begin{pmatrix} 1 & \mathrm{O} \\ \mathrm{O} & U \end{pmatrix}\begin{pmatrix} 1 & \mathrm{O} \\ \mathrm{O} & V \end{pmatrix} \quad = \quad \begin{pmatrix} 1 & \mathrm{O} \\ \mathrm{O} & UV \end{pmatrix}$$

für Blockmatrizen sieht man, dass mit E' auch die Matrix

$$\begin{pmatrix} 1 & \mathrm{O} \\ \mathrm{O} & E' \end{pmatrix}$$

als Produkt von Elementarmatrizen darstellbar ist.

Zu zeigen bleibt, dass jede Matrix in ZSF durch elementare Zeilenumformungen auf ***normierte*** ZSF gebracht werden kann. Dazu setzen wir voraus, dass A bereits in ZSF mit den Kennzahlen $r, j_1, \ldots, j_r$ vorliegt. Um $a_{i\,j_i} = 1$ für $1 \le i \le r$ zu erreichen, dividiert man einfach für jedes i die i-te Zeile durch $a_{i\,j_i}$. Die ZSF der Matrix wird durch diese Operation nicht zerstört, da die Eigenschaft eines Eintrages, gleich Null oder ungleich Null zu sein, dadurch nicht verändert wird.

Die Bedingung $a_{k\,j_i} = 0$ für $k < i$ kann dadurch erfüllt werden, dass man nacheinander für die Zeilennummern

$$i \quad = \quad r, r-1, r-2, \ldots, 1$$

jeweils das $a_{k\,j_i}$-fache der i-ten Zeile von der k-ten Zeile subtrahiert, für $1 \le k < i$. Dabei ist darauf zu achten, dass in keinem Schritt die ZSF beeinträchtigt wird und die erreichte Form für die Spalten j_ℓ mit $\ell > i$ erhalten bleibt. Die ZSF bleibt erhalten, da die i-te Zeile ihren ersten Eintrag $\neq 0$ erst in der Spalte j_i hat und $j_i > j_k$ für $1 \le k < i$ gilt. Somit werden weder die Zeilenköpfe der darüberliegenden Zeilen noch die Einträge links davon verändert. Die Zeilen unterhalb der i-ten bleiben völlig unverändert. An der Bedingung $a_{k\,j_\ell} = 0$ für $\ell > i$ und $k < \ell$ bleibt erhalten, da der einzige Eintrag ungleich Null in der j_ℓ-ten Spalte der Eintrag $a_{\ell\,j_\ell} = 1$ ist, und dieser spielt wegen $\ell > i$ bei der Zeilenumformung keine Rolle. □

Dieses „induktive“ Beweisschema kann problemlos verwendet werden, um eine beliebige, konkret vorgegebene Matrix $A \in \mathscr{M}_{m\times n,K}$ zunächst auf Zeilenstufenform und dann auf normierte Zeilenstufenform zu bringen. Man bezeichnet diese Vorgehensweise als das ***Gaußsche Eliminationsverfahren***.

§ 3. Die allgemeine lineare Gruppe

In diesem Abschnitt wiederholen wir zunächst die Definition einiger grundlegender algebraischer Strukturen, die teilweise aus dem ersten Semester bekannt sind, auf die wir im weiteren Verlauf der Vorlesung aber immer wieder zurückgreifen werden.

Definition 3.1 Eine ***Verknüpfung*** auf einer Menge A ist eine Abbildung $A \times A \to A$.

Als Bezeichnungen für eine Verknüpfung sind die Symbole $\cdot$, $\odot$, $*$, $+$, $\oplus$ und einige Varianten üblich. Wird eines der Symbole $\cdot$, $\odot$, $*$ verwendet, dann spricht man von einer ***multiplikativen*** Verknüpfung, bei $+$ oder $\oplus$ nennt man sie ***additiv***. Die beiden Typen unterscheiden sich aber außschließlich durch das verwendete Symbol, mathematisch gesehen besteht zwischen einer additiven und einer multiplikativen Verknüpfung keinerlei Unterschied.

Multiplikative Verknüpfungssymbole werden zur Vereinfachung der Notation häufig auch weggelassen, d.h. an Stelle von $a \cdot b$ schreibt man einfach ab. Sollen mehrere Elemente miteinander verknüpft werden, so ist die Verwendung von ***Klammern*** üblich, um die Reihenfolge der angewendeten Verknüpfungen anzuzeigen. So bedeutet zum Beispiel der Ausdruck $a(b(cd))$, dass zunächst das Element $x_1 = cd$ gebildet wird, anschließend $x_2 = bx_1$ und schließlich $x_3 = ax_2$.

Definition 3.2 Eine Verknüpfung $\cdot$ auf einer Menge A bezeichnet man als

(i) ***kommutativ***, wenn $ab = ba$ für alle $a, b \in G$
(ii) ***assoziativ***, wenn $a(bc) = (ab)c$ für alle $a, b, c \in G$ erfüllt ist.

Bei assoziativen Verknüpfungen können die Klammern auch weggelassen werden. Für beliebige Elemente $a, b, c, d \in A$ ist dann zum Beispiel $abcd$ eine Kurzschreibweise für das Element $a(b(cd))$, welches auf Grund der Assoziativität mit jedem anders geklammerten Ausdruck, etwa $(ab)(cd)$ oder $a((bc)d)$, übereinstimmt.

Definition 3.3 Eine ***Monoid*** ist ein Paar $(G, \cdot)$ bestehend aus einer Menge $G \neq \emptyset$ und einer Verknüpfung $\cdot : G \times G \to G$ mit folgenden Eigenschaften.

(i) Die Verknüpfung ist assoziativ.
(ii) Es gibt ein Element $e \in G$ mit $ae = ea = a$ für alle $a \in G$.

Ist die Verknüpfung $\cdot$ auch kommutativ, dann wird $(G, \cdot)$ ein ***kommutatives*** oder ***abelsches*** Monoid genannt.

Ein Element e mit der unter (ii) genannten Eigenschaft bezeichnet man als ***Neutralelement*** des Monoids. Jeder Monoid $(G,\cdot)$ besitzt nur ein Neutralelement. Ist nämlich e' ein weiteres, dann gilt $e = ee' = e'$. Aus diesem Grund kann von ***d***em Neutralelement des Monoids gesprochen werden. Man verwendet die feststehende Bezeichung e_G für dieses ausgezeichnete Element. Bei einer additiven Verknüpfung wird im allgemeinen an Stelle von e_G das Symbol 0_G verwendet.

Definition 3.4 Sei $(G,\cdot)$ ein Monoid. Wenn für jedes $a \in G$ ein Element $b \in G$ existiert, so dass die Gleichungen

$$ab = ba = e_G$$

erfüllt sind, so bezeichnet man $(G,\cdot)$ als eine ***Gruppe***. Wie bei den Monoiden spricht man von einer kommutativen oder abelschen Gruppe, wenn die Verknüpfung kommutativ ist.

Für jedes $a \in G$ bezeichnet man ein Element b mit der angegebenen Eigenschaft als ein zu a ***inverses*** Element. Jedes Element a in einem Monoid $(G,\cdot)$ besitzt höchstens ein Inverses. Sind nämlich $b, c \in G$ beide zu a invers, dann gilt $ba = e_G$ und $ac = e_G$, insgesamt also

$$b = be_G = b(ac) = (ba)c = e_Gc = c.$$

Auf Grund dieser Eindeutigkeit ist es zulässig, für das Inverse von a die feststehende Bezeichnung a^{-1} zu verwenden. Bei einer additiven Verknüpfung ist die Bezeichnung $-a$ für das Inverse üblich. Diejenigen Elemente in einem Monoid, die ein Inverses besitzen, bezeichnet man als ***invertierbare*** Elemente. Eine Gruppe ist also ein Monoid, in dem jedes Element invertierbar ist.

Für die invertierbaren Elemente eines Monoids gelten eine Reihe von wichtigen Rechenregeln.

Lemma 3.5 Sei $(G,\cdot)$ ein Monoid, und seien $a, b \in G$ invertierbare Elemente.

(i) Ist $c \in G$ ein Element mit $ac = e_G$, dann gilt $c = a^{-1}$. Ebenso folgt aus $ca = e_G$ bereits die Gleichheit $c = a^{-1}$.

(ii) Das Neutralelement e_G ist invertierbar, und es gilt $e_G^{-1} = e_G$.

(iii) Das Element ab ist invertierbar, und es gilt $(ab)^{-1} = b^{-1}a^{-1}$.

(iv) Auch das Inverse a^{-1} von a ist invertierbar, und es gilt $(a^{-1})^{-1} = a$.

Beweis: zu (i) Aus $ac = e_G$ und $a^{-1}a = e_G$ folgt insgesamt $c = e_Gc = (a^{-1}a)c = a^{-1}(ac) = a^{-1}e_G = a^{-1}$. Setzen wir $ca = e_G$ voraus, so erhalten wir die Gleichheit entsprechend durch $aa^{-1} = e_G$ und die Rechnung $c = ce_G = c(aa^{-1}) = (ca)a^{-1} = c(aa^{-1}) = (ca)a^{-1} = e_Ga^{-1} = a^{-1}$.

zu (ii) Ein Element $c \in G$ ist ein Inverses von e_G, wenn die Gleichungen $e_Gc = ce_G = e_G$ erfüllt sind. Aus der Gleichung $e_Ge_G = e_G$ folgt also, dass das Neutralelement e_G sein eigenes Inverses ist.

zu (iii) Dies folgt aus den beiden Gleichungen $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = ae_G a^{-1} = aa^{-1} = e_G$ und $(b^{-1}a^{-1})(ab) = b^{-1}(aa^{-1})b = b^{-1}e_G b = b^{-1}b = e_G$.

zu (iv) Ein Element $c \in G$ ist ein Inverses von a^{-1}, wenn die Gleichungen $a^{-1}c = e_G$ und $ca^{-1} = e_G$ gelten. Also folgt die Aussage aus den Gleichungen $a^{-1}a = e_G$ und $aa^{-1} = e_G$. □

Sei $*$ eine Verknüpfung auf einer Menge A. Eine Teilmenge $U \subseteq A$ bezeichnet man als ***abgeschlossen*** unter $*$, wenn für alle $a, b \in U$ auch das Element $a * b$ in U enthalten ist. Dies bedeutet, dass die Einschränkung der Verknüpfungsabbildung auf die Teilmenge $U \times U \subseteq A \times A$ eine Abbildung mit dem Bildbereich U liefert. Man erhält also eine neue Verknüpfung $*_U$, die nun auf der Menge U definiert ist. Um die Notation aber nicht zu aufwändig werden zu lassen, behält man meistens auch auf U das alte Verknüpfungssymbol bei. Gelegentlich ist die Schreibweise $*_U$ aber hilfreich, wenn betont werden soll, dass die Verknüpfungsabbildung auf U gemeint ist.

Beispielsweise ist nach Lemma (3.5) (iii) ist in jedem Mononoid $(G, \cdot)$ die Menge der invertierbaren Elemente unter der Verknüpfung $\cdot$ abgeschlossen. Also ist auch auf der Menge dieser Elemente, die man üblicherweise mit $G^\times$ bezeichnet, eine Verknüpfung definiert. Dasselbe Lemma zeigt auch, dass das Neutralelement e_G des Monoids in $G^\times$ enthalten ist. Mit jedem Element a liegt auch das Inverse a^{-1} in $G^\times$. Daraus folgt

Satz 3.6 Die Menge $G^\times$ der invertierbaren Elemente eines Monoids $(G, \cdot)$ bildet zusammen mit der auf $G^\times$ eingeschränkten Verknüpfung eine Gruppe.

Die aus dem ersten Semester bekannte Definitionen des ***Rings*** und des ***Körpers*** lassen sich mit dem Begriff des Monoids und der Gruppe nun kürzer formulieren.

Definition 3.7 Ein ***Ring*** ist ein Tripel $(R, +, \cdot)$ bestehend aus einer Menge R und zwei Verknüpfungen auf $+$ und $\cdot$ auf R mit folgenden Eigenschaften.

(i) Das Paar $(R, +)$ ist eine abelsche Gruppe.
(ii) Das Paar $(R, \cdot)$ ist ein abelsches Monoid.
(iii) Es gilt das Distributivgesetz $a(bc) = ab + ac$ für alle $a, b, c \in R$.

Wir erinnern daran, dass das Neutralelement der Gruppe $(R, +)$ das ***N**ullelement* 0_R und das Neutralelement des Monoids $(R, \cdot)$ das ***Einselement*** 1_R des Rings genant wird. Ist darüber hinaus die Gruppe $R^\times$ der invertierbaren Elemete gleich $R \setminus \{0_R\}$, dann nennt man den Ring $(R, +, \cdot)$ einen ***Körper***.

Wir verwenden die neu eingeführten Begriffe nun, um unser bisherigen Wissen über die Regeln zur Matrizenrechnung besser zu strukturieren. Wie immer bezeichnen wir mit K einen beliebigen Körper.

Satz 3.8 Sei K ein Körper, und seien $m, n \in \mathbb{N}$.

(i) Die Menge $\mathcal{M}_{m\times n,K}$ bildet mit der Addition von Matrizen eine abelsche Gruppe, deren Neutralelement die Nullmatrix O ist.

(ii) Die Menge $\mathcal{M}_{n,K}$ bildet mit der Multiplikation von Matrizen ein Monoid, mit der Einheitsmatrix I als Neutralelement. Dieses ist kommutativ für $n = 1$ und nicht-kommutativ für $n \geq 2$.

Beweis: zu (i) Zu zeigen ist, dass für beliebige Matrizen $A, B, C \in \mathcal{M}_{m\times n,K}$ das Assoziativgesetz $A + (B + C) = (A + B) + C$ und das Kommutativgesetz $A + B = B + A$ sowie die Gleichung $A + \mathrm{O} = A$ gültig ist. Das Nachrechnen dieser Rechenregeln ist einfacher als in Prop. (2.7) und wird deshalb hier nicht ausgeführt. Sie zeigen, dass $(\mathcal{M}_{m\times n,K}, +)$ ein abelsches Monoid ist. Ebenso leicht weist man nach, dass für jedes $A \in \mathcal{M}_{m\times n,K}$ mit der Matrix $-A = (-1_K)A$ die Gleichung $A + (-A) = \mathrm{O}$ erfüllt ist. Also ist unser Monoid sogar eine abelsche Gruppe.

zu (ii) Seien $A, B, C \in \mathcal{M}_{n,K}$ vorgegeben. Bereits in Prop. (2.7) wurde das Assoziativgesetz $A(BC) = (AB)C$ wurde nachgewiesen, und nach Prop. (2.8) gilt jeweils $A\mathrm{I} = \mathrm{I}A = A$. Also ist $(\mathcal{M}_{n,K}, \cdot)$ ein Monoid. Im Fall $n = 1$ gilt $(a)(b) = (ab) = (ba) = (b)(a)$ für alle $a, b \in K$, d.h. die Kommutativität von $(\mathcal{M}_{1,K}, \cdot)$ folgt direkt aus der Kommutativität von $(K, \cdot)$. Aber schon das Monoid $(\mathcal{M}_{2,K}, \cdot)$ ist nicht mehr kommutativ, denn beispielsweise gilt

$$\begin{pmatrix} 0_K & 1_K \\ 1_K & 0_K \end{pmatrix}\begin{pmatrix} 1_K & 1_K \\ 0_K & 1_K \end{pmatrix} = \begin{pmatrix} 0_K & 1_K \\ 1_K & 1_K \end{pmatrix} \neq \begin{pmatrix} 1_K & 1_K \\ 1_K & 0_K \end{pmatrix} = \begin{pmatrix} 1_K & 1_K \\ 0_K & 1_K \end{pmatrix}\begin{pmatrix} 0_K & 1_K \\ 1_K & 0_K \end{pmatrix}$$

Im Fall $n > 2$ kann die Nicht-Kommutativität durch Rechnen mit Matrizen in Blockdarstellung auf den Fall $n = 2$ zurückgeführt werden. Dies überlassen wir dem Leser als Übungsaufgabe. □

Eine Matrix $A \in \mathcal{M}_{n,K}$ wird ***invertierbar*** genannt, wenn eine Matrix $B \in \mathcal{M}_{n,K}$ mit $AB = BA = \mathrm{I}^{(n)}$ existiert. Beispielsweise sind die Elementarmatrizen $\mathrm{M}_{k,\lambda}$ und $\mathrm{A}_{k,\ell,\lambda}$ invertierbar. Für alle $\lambda \in K^\times$ und $k \in \{1, ..., n\}$ gilt

$$\mathrm{M}_{k,\lambda}\mathrm{M}_{k,\lambda^{-1}} = \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \lambda & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(n-k)} \end{pmatrix}\begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \lambda^{-1} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(n-k)} \end{pmatrix} =$$

$$= \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & 1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(n-k)} \end{pmatrix} = \mathrm{I}^{(n)}$$

und ebenso $\mathrm{M}_{k,\lambda^{-1}}\mathrm{M}_{k,\lambda} = \mathrm{I}^{(n)}$.

Die Invertierbarkeit von $\mathrm{A}_{k,\ell,\lambda}$ erhält man im Fall $k < \ell$ durch die Gleichungen

$$\mathrm{A}_{k,\ell,\lambda}\mathrm{A}_{k,\ell,-\lambda} \quad = $$

$$\begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & 1 & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(\ell-k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \lambda & \mathrm{O} & 1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{I}^{(n-\ell)} \end{pmatrix} \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & 1 & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(\ell-k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & -\lambda & \mathrm{O} & 1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{I}^{(n-\ell)} \end{pmatrix}$$

$$= \begin{pmatrix} \mathrm{I}^{(k-1)} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & 1 & \mathrm{O} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{I}^{(\ell-k-1)} & \mathrm{O} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & 1 & \mathrm{O} \\ \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{O} & \mathrm{I}^{(n-\ell)} \end{pmatrix} = \mathrm{I}^{(n)}$$

und $\mathrm{A}_{k,\ell,-\lambda}\mathrm{A}_{k,\ell,\lambda} = \mathrm{I}^{(n)}$, und im Fall $k > \ell$ durch eine analoge Rechnung. Aus Satz 3.6 folgt

Folgerung 3.9 Die Menge der invertierbaren $n \times n$-Matrizen über einem Körper K bildet eine Gruppe. Man nennt sie die ***allgemeine lineare Gruppe*** und bezeichnet sie mit $\mathrm{GL}_n(K)$.

Weil $\mathrm{GL}_n(K)$ unter Matrizenmultiplikation abgeschlossen ist und alle Elementarmatrizen in $\mathrm{GL}_n(K)$ liegen, ist auch die Menge $\mathscr{E}_n(K)$ aller Matrizen, die als Produkte von Elementarmatrizen dargestellt werden können, in $\mathrm{GL}_n(K)$ enthalten. Ist allgemein $B \in \mathrm{GL}_n(K)$, dann gilt für jedes $A \in \mathscr{M}_{n,K}$ offenbar die Äquivalenz

$$A \in \mathrm{GL}_n(K) \quad \Leftrightarrow \quad AB \in \mathrm{GL}_n(K) \quad \Leftrightarrow \quad BA \in \mathrm{GL}_n(K).$$

Setzen wir nämlich $A \in \mathrm{GL}_n(K)$ voraus, dann ist AB invertierbar, weil die Menge der invertierbaren Matrizen unter Multiplikation abgeschlossen ist. Setzen wir umgekehrt die Invertierbarkeit von AB voraus, dann folgt mit demselben Argument wegen $(AB)B^{-1} = A(BB^{-1}) = A\mathrm{I}^{(n)} = A$ die Invertierbarkeit von A. Der Beweis der Äquivalenz $A \in \mathrm{GL}_n(K) \Leftrightarrow BA \in \mathrm{GL}_n(K)$ läuft völlig analog.

Wir beschäftigen uns nun mit der Frage, wie man die Invertierbarkeit von Matrizen nachweist, und wie gegebenenfalls die inverse Matrix berechnet werden kann.

Satz 3.10 Lässt sich eine Matrix $A \in \mathscr{M}_{n,K}$ durch endliche viele elementare Zeilenumformungen in eine Matrix A' in normierter ZSF mit Zeilenrang $r = n$ umwandeln, so ist A invertierbar.

Beweis: Zunächst bemerken wir, dass eine Matrix A' in normierter ZSF mit Zeilenrang $r = n$ zwangsläufig mit der Einheitsmatrix $\mathrm{I}^{(n)}$ übereinstimmt. Sind nämlich neben n die Kennzahlen der ZSF gegeben durch $1 \le j_1 < j_2 < ... < j_n \le n$, dann muss $j_k = k$ für $1 \le k \le n$ gelten. Nach Definition der normierten ZSF ist die j_k-te Spalte von A' gleich dem k-ten Einheitsvektor $e_k \in K^n$.

Weil A durch elementare Zeilenumformungen in $A' = \mathrm{I}^{(n)}$ überführt werden kann, gibt es eine Matrix $T \in \mathscr{E}_n(K) \subseteq \mathrm{GL}_n(K)$ mit $TA = \mathrm{I}^{(n)}$. Es folgt $A = \mathrm{I}^{(n)} A = (T^{-1}T)A = T^{-1}(TA) = T^{-1}\mathrm{I}^{(n)} = T^{-1}$. Damit ist die Invertierbarkeit von A bewiesen. □

Die Beweisidee in Satz (3.10) kann genutzt werden, um die zu A inverse Matrix auszurechnen. Wendet man die Zeilenumformungen im Beweis statt auf A auf die Blockmatrix $(A \;\; \mathrm{I}^{(n)})$ an, so erhält man die Matrix

$$T(A \;\; \mathrm{I}^{(n)}) \quad = \quad (TA \;\; T\mathrm{I}^{(n)}) \quad = \quad (\mathrm{I}^{(n)} \;\; T).$$

Aus der rechten Hälfte der umgeformten Matrix kann die Inverse von A abgelesen werden, denn es gilt $A = T^{-1} \Leftrightarrow A^{-1} = T$.

Offen bleibt zunächst noch die Frage, wie es zu interpretieren ist, wenn die Matrix A zwar auf normierte ZSF, aber mit Zeilenrang $r < n$, gebracht werden kann. Dafür ist es notwendig, dass wir uns neben den Zeilen- auch mit ***Spaltenumformungen*** einer Matrix befassen. Unter einer ***elementaren*** Spaltenumformungen verstehen wir, dass die zu den Zeilenumformungen analogen Operationen auf die Spalten einer Matrix $A \in \mathscr{M}_{m\times n,K}$ angewendet werden, also im einzelnen

(i) die Multiplikation der k-ten Spalten einer Matrix mit einem Wert λ, wobei $\lambda \in K^\times$ und $k \in \{1,...,n\}$ ist
(ii) die Addition des λ-fachen der k-ten Spalte zur ℓ-ten, mit $\lambda \in K$ und $k, \ell \in \{1,...,n\}$, $k \neq \ell$.

Lemma 3.11 Die Multiplikationen einer Matrix $A \in \mathscr{M}_{m\times n,K}$ mit den Transponierten von Elementarmatrizen von rechts bewirken elementare Spaltenumformungen. Genauer gilt:

(i) Die Matrix $A\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda}$ entsteht aus der Matrix A durch Multiplikation der k-ten Spalte mit λ.

(ii) Die Matrix $A\,{}^{\mathrm{t}}\mathrm{A}_{k,\ell,\lambda}$ entsteht aus der Matrix A durch Addition des λ-fachen der k-ten Spalte zur ℓ-ten Spalte.

Beweis: Wir beschränken uns auf den Beweis der Aussage (i). Nach der Rechenregel (iv) in Prop. (2.7) gilt

$$A\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda} \quad = \quad {}^{\mathrm{t}}({}^{\mathrm{t}}A)\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda} \quad = \quad {}^{\mathrm{t}}(\mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A).$$

Der Übergang ${}^{\mathrm{t}}A \mapsto \mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A$ bewirkt nach Prop. (2.11) die Multiplikation der k-ten Zeile von ${}^{\mathrm{t}}A$ mit dem Wert λ. Für jedes ℓ ist die ℓ-te Spalte von A gleich der ℓ-ten Zeile von ${}^{\mathrm{t}}A$, und entsprechend ist die ℓ-te Spalte von $A\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda} = {}^{\mathrm{t}}(\mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A)$ gleich der ℓ-ten Zeile von $\mathrm{M}_{k,\lambda}\,{}^{\mathrm{t}}A$. Also stimmt die ℓ-te Spalte von A mit der ℓ-ten Spalte von $A\,{}^{\mathrm{t}}\mathrm{M}_{k,\lambda}$ für $\ell \neq k$ überein, und für $\ell = k$ unterscheiden sie sich um den Faktor λ. □

Wir bemerken noch, dass mit jeder Matrix $A \in \mathrm{GL}_n(K)$ auch die Transponierte ${}^{\mathrm{t}}A$ invertierbar ist, mit $({}^{\mathrm{t}}A)^{-1} = {}^{\mathrm{t}}(A^{-1})$. Dies folgt direkt aus der Rechnung

$$ {}^{\mathrm{t}}A\,{}^{\mathrm{t}}(A^{-1}) \quad = \quad {}^{\mathrm{t}}(A^{-1}A) \quad = \quad {}^{\mathrm{t}}\mathrm{I}^{(n)} \quad = \quad \mathrm{I}^{(n)} $$

und einer analogen Rechnung, die ${}^{\mathrm{t}}(A^{-1})\,{}^{\mathrm{t}}A = \mathrm{I}^{(n)}$ liefert.

Satz 3.12 Für jede Matrix $A \in \mathscr{M}_{m\times n,K}$ gibt es invertierbare Matrizen $T \in \mathrm{GL}_m(K)$ und $U \in \mathrm{GL}_n(K)$ und ein $r \in \{1,...,n\}$, so dass die Matrix TAU die Blockgestalt

$$ \begin{pmatrix} \mathrm{I}^{(r)} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} \end{pmatrix} \qquad \text{besitzt.} $$

Beweis: Wir wissen bereits, dass eine Matrix $T \in \mathrm{GL}_m(K)$ existiert, so dass $B = TA$ in normierter ZSF vorliegt, mit gewissen Kennzahlen $r, j_1,..., j_r$. Nach Lemma (3.11) genügt es nun zu zeigen, dass B durch elementare Spaltenumformungen auf die angegebene Blockgestalt gebracht werden kann. Nach Definition der normierten ZSF befindet sich für $1 \le k \le r$ in der j_k-ten Spalten von B jeweils der k-te Einheitsvektor $e_k \in K^m$. Nun führt man nacheinander für $1 \le k \le r$ die folgende Operation aus:

$$ \text{Addition des } (-a_{k\ell})\text{-fachen der } j_k\text{-ten Spalte zur } \ell\text{-ten, für } j_k < \ell \le n $$

Durch diese Operation werden die Einträge rechts von der Position (k, j_k) zu Null, während alle übrigen Einträge der Matrix unverändert bleiben.

Nach Durchführung dieser Schritte enthält die modifizierte Matrix B' in den Spalten $j_1,..., j_r$ die Einheitsvektoren $e_1,..., e_r$, alle übrigen Spalten sind Null. Nun vertauscht man die Spalten noch so, dass sich die Einheitsvektoren in den ersten r Spalten befinden. Dann hat die Matrix die gewünschte Form. □

Auch hier lassen sich die Matrizen T und U, die die angegebene Blockgestalt erzeugen, explizit berechnen. Zunächst wendet man die erforderlichen Zeilenumformungen statt auf A auf die Blockmatrix $(A \;\; \mathrm{I}^{(m)})$ an und erhält so eine Matrix der Form $(B \;\; T)$ mit $T \in \mathrm{GL}_m(K)$, wobei $B = TA$ sich in normierter ZSF befindet. Anschließend wendet man auf die linke Teilmatrix von $(B \;\; \mathrm{I}^{(n)})$ Spaltenumformungen an, die B auf die Blockgestalt bringen, und ***dieselben*** Spaltenumformungen auch auf die rechte Teilmatrix. Man erhält damit eine Matrix der Form $(C \;\; U)$ mit $U \in \mathrm{GL}_m(K)$, wobei $C = BU$ die angegebene Blockgestalt hat. Die Matrizen T und U haben die gewünschte Umformungeigenschaft.

Durch dieses Rechenverfahren wird auch eine Möglichkeit aufgezeigt, eine Matrix auf Invertierbarkeit zu testen.

Satz 3.13 Sei $A \in \mathcal{M}_{n,K}$ eine Matrix, die durch elementare Zeilenumformungen auf normierte ZSF mit Zeilenrang $r < n$ gebracht werden kann. Dann ist A *nicht* invertierbar.

Beweis: Nehmen wir an, dass die Voraussetzung erfüllt, die Matrix A aber dennoch in $\mathcal{M}_{n,K}$ liegt. Nach Satz (3.12) gibt es Matrizen $T, U \in \mathrm{GL}_n(K)$ mit

$$TAU = \begin{pmatrix} \mathrm{I}^{(r)} & \mathrm{O} \\ \mathrm{O} & \mathrm{O} \end{pmatrix}.$$

Mit A wäre dann auch TAU invertierbar. Aber eine Matrix mit Nullzeilen kann nicht invertierbar sein, denn für beliebiges $B \in \mathcal{M}_{r\times n,K}$ und $V \in \mathrm{GL}_n(K)$ gilt

$$\begin{pmatrix} B \\ \mathrm{O} \end{pmatrix} V = \begin{pmatrix} BV \\ \mathrm{O}V \end{pmatrix} = \begin{pmatrix} BV \\ \mathrm{O} \end{pmatrix} ,$$

wobei die letzte Matrix offensichtlich *nicht* mit der Einheitsmatrix übereinstimmt. Der Widerspruch zeigt, dass die Annahme falsch war. □

Unser Rechenverfahren zur Bestimmung der Inversen einer Matrix A liefert also zugleich ein Entscheidungskriterium zur Invertierbarkeit: Kommt bei der Rechnung eine normierte ZSF mit Zeilenrang $r < n$ heraus, dann existiert die Inverse von A nicht.

§ 4. *Vektorräume und lineare Abbildungen*

Definition 4.1 Sei K ein Körper. Ein K-***Vektorraum*** ist ein Tripel $(V,+,\cdot)$ bestehend aus einer nichtleeren Menge V und Abbildungen $+ : V \times V \to V$ und $\cdot : K \times V \to V$ genannt ***Vektoraddition*** und ***skalare Multplikation***, so dass folgende Bedingungen erfüllt sind.

(i) Das Paar $(V,+)$ ist eine abelsche Gruppe.

(ii) Für alle $v, w \in V$ und $\lambda, \mu \in K$ gelten die Rechenregeln

   (a) $(\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v)$

   (b) $\lambda \cdot (v + w) = (\lambda \cdot v) + (\lambda \cdot w)$

   (c) $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$

   (d) $1_K \cdot v = v$

Die Elemente der Menge V werden ***Vektoren*** genannt.

Bei der skalaren Multiplikation wird häufig auf das Abbildungssymbol $\cdot$ verzichtet. Das Neutralelement der Gruppe $(V,+)$ bezeichnet man als den ***Nullvektor*** 0_V des Vektorraums. Das Inverse eines Vektors $v \in V$ bezüglich der Vektoraddition bezeichnet man mit $-v$ und verwendet $v - w$ als abkürzende Schreibweise für $v + (-w)$. Per Konvention bindet die skalare Multiplikation stärker als die Vektoraddition, d.h. der Ausdruck $\lambda v + w$ ist gleichbedeutend mit $(\lambda v) + w$ für $\lambda \in K$, $v, w \in V$.

Proposition 4.2 Sei $(V,+,\cdot)$ ein K-Vektorraum. Dann gilt für alle $\lambda \in K$ und $v \in V$ die Äquivalenz

$$\lambda v = 0_V \quad \Leftrightarrow \quad \lambda = 0_K \text{ oder } v = 0_V \quad ,$$

außerdem $(-1_K)v = -v$ für alle $v \in V$.

Beweis: Zunächst beweisen wir die Äquivalenz. „$\Leftarrow$“ Ist $\lambda = 0_K$, dann gilt $\lambda v = 0_K v = (0_K + 0_K)v = 0_K v + 0_K v = \lambda v + \lambda v$. Addition von $-\lambda v$ auf beiden Seiten dieser Gleichung liefert

$$\lambda v + (-\lambda v) = \lambda v + \lambda v + (-\lambda v) \quad \Leftrightarrow \quad 0_V = \lambda v + 0_V \quad \Leftrightarrow \quad 0_V = \lambda v.$$

Setzen wir nun $v = 0_V$ voraus, dann erhalten wir $\lambda v = \lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V + \lambda 0_V = \lambda v + \lambda v$. Wieder führt die Addition von $-\lambda v$ auf beiden Seiten zum gewünschten Ergebnis.

„$\Rightarrow$“ Setzen wir $\lambda v = 0_V$ voraus, und nehmen wir an, es ist $\lambda \neq 0_K$. Dann gilt

$$v \quad = \quad 1_K v \quad = \quad (\lambda^{-1}\lambda)v \quad = \quad \lambda^{-1}(\lambda v) \quad = \quad \lambda^{-1}0_V \quad = \quad 0_V \quad ,$$

wobei im letzten Schritt die bereits bewiesene Rechenregel $\mu 0_V = 0_V$ für alle $\mu \in K$ verwendet wurde. Beweisen wir nun noch die Gleichung $(-1_K)v = -v$. Es gilt $v + (-1_K)v = 1_K v + (-1_K)v = (1_K + (-1_K))v = 0_K v = 0_V$.

Addition von $-v$ auf beiden Seiten liefert

$$v + (-1_K)v + (-v) = 0_V + (-v) \quad \Leftrightarrow \quad v + (-v) + (-1_K)v = -v \quad \Leftrightarrow$$
$$0_V + (-1_K)v = -v \quad \Leftrightarrow \quad (-1_K)v = -v \qquad \square$$

Proposition 4.3 Sei K ein Körper.
Die folgenden Strukturen sind Beispiele für K-Vektorräume.

(i) für jedes $n \in \mathbb{N}$ das Tripel $(K^n, +, \cdot)$ mit den Abbildungen

$$+ : K^n \times K^n \to K^n \quad , \quad ((a_1, ..., a_n), (b_1, ..., b_n)) \mapsto (a_1 + b_1, ..., a_n + b_n)$$

und

$$\cdot : K \times K^n \to K^n \quad , \quad (\lambda, (a_1, ..., a_n)) \mapsto (\lambda a_1, ..., \lambda a_n)$$

(ii) für alle $m, n \in \mathbb{N}$ das Tripel $(\mathcal{M}_{m\times n,K}, +, \cdot)$, wobei $+ : \mathcal{M}_{m\times n,K} \times \mathcal{M}_{m\times n,K} \to \mathcal{M}_{m\times n,K}$ für die Addition und $\cdot : K \times \mathcal{M}_{m\times n,K} \to \mathcal{M}_{m\times n,K}$ für die skalare Multiplikation von Matrizen steht

(iii) das Tripel $(K, +, \cdot)$, wobei $+$ die Addition und $\cdot$ die Multiplikation von K bezeichnet

(iv) das Tripel $(\{0_V\}, +, \cdot)$, wobei $+ : \{0_V\} \times \{0_V\} \to \{0_V\}$ durch $(0_V, 0_V) \mapsto 0_V$ und $\cdot : K \times \{0_V\} \to \{0_V\}$ durch $(\lambda, 0_V) \mapsto 0_V$ für alle $\lambda \in K$ gegeben ist

Beweis: Punkt (i) kann als Spezialfall von (ii) angesehen werden, da wir in §1 festgelegt haben, dass die Elemente aus K^n Matrizen mit einer einzigen Spalte sind. Beweisen wir nun die Vektorraum-Eigenschaft von Beispiel (ii). Nach Satz (3.8) (i) ist $(\mathcal{M}_{n,K}, +)$ eine abelsche Gruppe, mit der Nullmatrix als Neutralelement. Zu überprüfen sind noch die Gleichungen $(\lambda+\mu)A = \lambda A+\mu A$, $\lambda(A+B) = \lambda A+\lambda B$, $(\lambda\mu)A = \lambda(\mu A)$ und $1_K A = A$ für beliebige Körperelemente $\lambda, \mu \in K$ und beliebige Matrizen $A, B \in \mathcal{M}_{m\times n,K}$. Der Beweis dieser Rechenregeln erfolgt ähnlich wie in Prop. (2.7) und wird hier nicht ausgeführt.

Kommen wir nun zu Beispiel (iii). Dass $(K, +)$ eine abelsche Gruppe mit 0_K als Neutralelement ist, ergibt sich direkt aus Def. (3.7) (i). Seien nun $\lambda, \mu \in K$ und $a, b \in K$ vorgegeben, wobei λ, μ die Rolle der Körperelemente und a, b die Rolle der „Vektoren" übernehmen. Die Rechenregel (ii)(b) in Def. (4.1) gegeben durch $\lambda(a + b) = \lambda a + \lambda b$ entspricht dem Distributivgesetz im Körper K. Die Regel (ii)(a) erhält man durch Kommutativ- und Distributivgesetz, denn es gilt

$$(\lambda + \mu)a \quad = \quad a(\lambda + \mu) \quad = \quad a\lambda + \mu\lambda \quad = \quad \lambda a + \mu a.$$

Nach dem Assoziativgesetz der Multiplikation gilt $(\lambda\mu)a = \lambda(\mu a)$, also ist auch (ii)(c) gültig. Nach Definition des Einselements in einem Körper gilt schließlich noch $1_K a = a$, also Regel (ii)(d).

Betrachten wir nun Beispiel (iv) und überprüfen zunächst, dass $(\{0_V\},+)$ eine abelsche Gruppe ist. Assoziativ- und Kommutativgesetz sind wegen $0_V + 0_V = 0_V + 0_V$ und $(0_V + 0_V) + 0_V = 0_V + 0_V = 0_V + (0_V + 0_V)$ erfüllt. Wegen $a + 0_V = 0_V = a$ und $0_V + a = 0_V = a$ für alle $a \in \{0_V\}$ ist 0_V das Neutralelement der Gruppe. Die Gleichung $0_V + 0_V = 0_V + 0_V = 0_V$ zeigt, dass 0_V sein eigenes Inverses ist. Damit ist der Nachweis der Gruppeneigenschaften abgeschlossen. Zur Überprüfung der Rechenregeln (ii)(a) bis (d) seien $\lambda, \mu \in K$ vorgegeben. Es gilt

$$(\lambda + \mu)0_V \quad = \quad 0_V \quad = \quad 0_V + 0_V \quad = \quad \lambda 0_V + \mu 0_V \quad ,$$

also ist (ii)(a) erfüllt. Die Regel (ii)(b) erhält man durch $\lambda(0_V + 0_V) = \lambda 0_V = 0_V = 0_V + 0_V = \lambda 0_V + \mu 0_V$, und (ii)(c) durch die Rechnung $(\lambda\mu) \cdot 0_V = 0_V = \lambda \cdot 0_V = \lambda \cdot (\mu \cdot 0_V)$. Schließlich ist wegen $1_K 0_V = 0_V$ auch (ii)(d) erfüllt. □

Nach Prop. (4.3) (i) ist $\mathbb{C}^n$ für jedes $n \in \mathbb{N}$ mit der komponentenweisen Addition + und skalaren Multiplikation · ein $\mathbb{C}$-Vektorraum. Wir bemerken noch, dass die skalare Multiplikation $\cdot : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$ zu einer Abbildung $\cdot_{\mathbb{R}} : \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ eingeschränkt werden kann. Das Tripel $(\mathbb{C}^n, +, \cdot_{\mathbb{R}})$ ist dann ein $\mathbb{R}$-Vektorraum, denn wenn die Rechenregeln (ii)(a) bis (d) für alle $v, w \in \mathbb{C}^n$ und $\lambda, \mu \in \mathbb{C}$ gültig sind, dann gelten sie erst recht für alle $v, w \in \mathbb{C}^n$ und $\lambda, \mu \in \mathbb{R}$. Auch jeder andere $\mathbb{C}$-Vektorraum kann durch Einschränkung der skalaren Multiplikation auf diese Weise zu einem $\mathbb{R}$-Vektorraum gemacht werden.

Wir betrachten noch ein weiteres Beispiel für Vektoräume, das auf den ersten Blick etwas exotisch erscheinen mag.

Proposition 4.4 Sei K ein Körper, $(V, +_V, \cdot_V)$ ein K-Vektorraum und X eine beliebige Menge. Mit $A = \mathrm{Abb}(X, V)$ bezeichnen wir die Menge der Abbildungen $X \to V$. Außerdem

(i) definieren wir auf A eine Verknüpfung $\oplus$, indem wir jedem Paar (f, g) von Elementen aus A die Abbildung gegeben durch $(f \oplus g)(x) = f(x) +_V g(x)$ für alle $x \in X$ zuordnen

(ii) definieren wir eine Abbildung $\odot : K \times A \to A$, indem wir jedem Paar (λ, f) die Abbildung gegeben durch $(\lambda f)(x) = \lambda \cdot_V f(x)$ für alle $x \in X$ zuordnen.

Dann ist $(A, \oplus, \odot)$ ein K-Vektorraum.

Beweis: Zunächst zeigen wir, dass $(A, \oplus)$ eine abelsche Gruppe ist. Seien $f, g, h \in A$ vorgegeben. Für ein beliebiges $x \in X$ gilt

$$\begin{aligned} ((f \oplus g) \oplus h)(x) \quad &= \quad (f \oplus g)(x) +_V h(x) \quad = \quad (f(x) +_V g(x)) +_V h(x) \quad = \quad f(x) +_V (g(x) +_V h(x)) \\ &= \quad f(x) +_V (g \oplus h)(x) \quad = \quad (f \oplus (g \oplus h))(x) \quad , \end{aligned}$$

wobei wir im dritten Schritt die Assoziativität der Verknüpfung $+_V$ auf V verwendet haben. Weil $x \in X$ beliebig vorgegeben war, folgt daraus die Gleichheit der Abbildungen, also $(f \oplus g) \oplus h = f \oplus (g \oplus h)$. Ebenso folgt aus $(f \oplus g)(x) = f(x) +_V g(x) = g(x) +_V f(x) = (g \oplus f)(x)$ für $f, g \in V$ und $x \in X$ die Gleichung $f \oplus g = g \oplus f$ für alle $f, g \in A$ und somit die Kommutativität der Verknüpfung $\oplus$ auf A.

Sei nun $0_A \in A$ die Abbildung gegeben durch $0_A(x) = 0_V$ für alle $x \in X$. Für alle $f \in A$ und $x \in X$ gilt dann

$$(f \oplus 0_A)(x) \quad = \quad f(x) +_V 0_A(x) \quad = \quad f(x) +_V 0_V \quad = \quad f(x) \quad ,$$

also $f \oplus 0_V = f$, und auf Grund der bereits bewiesenen Kommutativität auch $0_V \oplus f = f$. Schließlich definieren wir für jedes $f \in A$ die Abbildung $-f \in A$ durch $(-f)(x) = -f(x)$ für alle $x \in X$. Für jedes $x \in X$ gilt dann

$$(f \oplus (-f))(x) \quad = \quad f(x) +_V (-f)(x) \quad = \quad f(x) +_V (-f(x)) \quad = \quad 0_V \quad = \quad 0_A(x)$$

und somit $f \oplus (-f) = 0_A$. Auf Grund der Kommutativität von $\oplus$ gilt damit auch $(-f) \oplus f = 0_A$. Damit ist insgesamt nachgewiesen, dass es sich bei $(A, \oplus)$ um eine abelsche Gruppe handelt.

Nun müssen wir noch die Gleichungen (a) bis (d) in Def. (4.1) (ii) überprüfen. Seien dazu $f, g \in A$ und $\lambda, \mu \in K$ vorgegeben. Für alle $x \in X$ gilt

$$((\lambda + \mu) \odot f)(x) \quad = \quad (\lambda + \mu) \cdot_V f(x) \quad = \quad \lambda \cdot_V f(x) +_V \mu \cdot_V f(x) \quad =$$

$$(\lambda \odot f)(x) +_V (\mu \odot f)(x) \quad = \quad (\lambda \odot f \oplus \mu \odot f)(x)$$

und somit $(\lambda + \mu) \odot f = \lambda \odot f \oplus \mu \odot f$ Ebenso gilt für alle $x \in X$ die Gleichung

$$\begin{aligned} (\lambda \odot (f \oplus g))(x) \quad &= \quad \lambda \cdot_V ((f \oplus g)(x)) \quad = \quad \lambda \cdot_V (f(x) +_V g(x)) \quad = \quad \lambda \cdot_V f(x) +_V \lambda \cdot_V g(x) \\ &= \quad (\lambda \odot f)(x) +_V (\lambda \odot g)(x) \quad = \quad (\lambda \odot f \oplus \lambda \odot g)(x) \quad , \end{aligned}$$

also $\lambda \odot (f \oplus g) = \lambda \odot f \oplus \lambda \odot g$. Zum Nachweis von Bedingung (c) rechnen wir nach, dass für alle $x \in X$ die Gleichung

$$((\lambda\mu) \odot f)(x) \quad = \quad (\lambda\mu) \cdot_V f(x) \quad = \quad \lambda \cdot_V (\mu \cdot_V f(x)) \quad = \quad \lambda \cdot_V (\mu \odot f)(x) \quad = \quad (\lambda \odot (\mu \odot f))(x)$$

und somit $(\lambda\mu) \odot f = \lambda \odot (\mu \odot f)$ gilt; dabei haben wir im zweiten Schritt die Rechenregel (ii)(c) im Vektorraum $(V, +_V, \cdot_V)$ verwendet. Schließlich gilt für alle $f \in A$ und $x \in X$ noch

$$(1_K \odot f)(x) \quad = \quad 1_K \cdot_V f(x) \quad = \quad f(x) \quad ,$$

also $1_K \odot f = f$ für alle $f \in A$. Damit haben wir auch die Regel (ii)(d) auf die entsprechende Regel in $(V, +_V, \cdot_V)$ zurückgeführt. □

Definition 4.5 Seien $(V, +_V, \cdot_V)$ und $(W, +_W, \cdot_W)$ K-Vektorräume. Eine Abbildung $\phi : V \to W$ heißt ***lineare Abbildung*** oder ***Homomorphismus*** von K-Vektorräumen, wenn folgende Bedingungen erfüllt sind.

(i) $\phi(v +_V w) = \phi(v) +_W \phi(w)$ für alle $v, w \in V$

(ii) $\phi(\lambda \cdot_V v) = \lambda \cdot_W \phi(v)$ für alle $v \in V$ und $\lambda \in K$

Lemma 4.6 Ist $\phi : V \to W$ eine lineare Abbidlung wie in der Definition angegeben. Dann gilt $\phi(0_V) = 0_W$, $\phi(-v) = -\phi(v)$ und $\phi(v-w) = \phi(v) - \phi(w)$ für alle $v, w \in V$.

Beweis: Die erste Gleichung erhält man mit Hilfe der Eigenschaft (ii) von linearen Abbildungen durch $\phi(0_V) = \phi(0_K \cdot_V 0_V) = 0_K \cdot_W \phi(0_V) = 0_W$. Die zweite ergibt sich durch die Rechnung

$$\phi(-v) \quad = \quad \phi((-1_K)\cdot_V v) \quad = \quad (-1)_K \cdot_W \phi(v) \quad = \quad -\phi(v).$$

Die dritte Gleichung schließlich erhält man durch

$$\phi(v-w) \quad = \quad \phi(v +_V (-w)) \quad = \quad \phi(v) +_W \phi(-w) \quad = \quad \phi(v) +_W (-\phi(w)) \quad = \quad \phi(v) - \phi(w). \qquad \square$$

Proposition 4.7 Seien $m, n \in \mathbb{N}$ und $A \in \mathcal{M}_{m\times n,K}$. Dann ist durch $\phi_A : K^n \to K^m$, $v \mapsto Av$ eine lineare Abbildung gegeben.

Beweis: Seien $v, w \in K^n$ und $\lambda \in K$ vorgegeben. Auf Grund der Rechenregeln aus Prop. (1.12) für das Matrix-Vektor-Produkt gilt

$$\phi_A(v+w) = A(v+w) = Av + Aw = \phi_A(v) + \phi_A(w)$$

und $\phi_A(\lambda v) = A(\lambda v) = \lambda Av = \lambda\phi_A(v)$. $\square$

Sei V ein K-Vektorraum, $n \in \mathbb{N}$, und seien $v_1, ..., v_n \in V$ beliebige Vektoren. Wie bei Körpern verwenden wir den Ausdruck $\sum_{k=1}^n v_k$ als Kurzschreibweise für die Summe $v_1 + ... + v_n$ in V.

Lemma 4.8 Seien V, W K-Vektorräume, $n \in \mathbb{N}$, außerdem $v_1, ..., v_n \in V$ und $\phi : V \to W$ eine lineare Abbildung. Dann gilt

$$\phi\left(\sum_{k=1}^n v_k\right) \quad = \quad \sum_{k=1}^n \phi(v_k).$$

Beweis: Wir beweisen die Aussage durch vollständige Induktion über n. Im Fall $n = 1$ lautet die Behauptung nur $\phi(v_1) = \phi(v_1)$ für alle $v_1 \in V$ und ist offensichtlich erfüllt. Sei nun $n \in \mathbb{N}$, und setzen wir die Aussage für dieses n voraus. Seien $v_1, ..., v_{n+1} \in V$ beliebige Vektoren. Dann erhalten wir

$$\phi\left(\sum_{k=1}^{n+1} v_k\right) \quad = \quad \phi\left(\sum_{k=1}^{n} v_k + v_{n+1}\right) \quad = \quad \phi\left(\sum_{k=1}^{n} v_k\right) + \phi(v_{n+1}) \quad \overset{(*)}{=}$$

$$\sum_{k=1}^n \phi(v_k) + \phi(v_{n+1}) \quad = \quad \sum_{k=1}^{n+1} \phi(v_k) \quad ,$$

wobei an der Stelle (*) die Induktionsvoraussetzung angewendet wurde. $\square$

Proposition 4.9 Die Linearformen auf dem K-Vektorraum K^n sind genau die linearen Abbildungen $K^n \to K$.

Beweis: Dass jede Linearform ϕ auf K^n eine lineare Abbildung $K^n \to K$ ist, folgt direkt aus den Rechenregeln für Linearformen aus Lemma (1.4). Sei nun $\phi : K^n \to K$ eine lineare Abbildung. Für $1 \le k \le n$ sei $e_k \in K^n$ jeweils der k-te Einheitsvektor und $a_k = \phi(e_k)$. Ist nun $v = (\lambda_1, ..., \lambda_n) \in K^n$ ein beliebiger Vektor, dann gilt $v = \sum_{k=1}^n \lambda_k e_k$. Durch Anwendung von Lemma (4.8) erhalten wir

$$\phi(v) \quad = \quad \phi\left(\sum_{k=1}^n \lambda_k e_k\right) \quad = \quad \sum_{k=1}^n \phi(\lambda_k e_k) \quad = \quad \sum_{k=1}^n \lambda_k \phi(e_k) \quad = \quad \sum_{k=1}^n \lambda_k a_k.$$

Dies zeigt, dass ϕ eine Linearform im Sinne von Def. (1.1) ist. □

Seien X, Y Mengen, sei V ein K-Vektorraum, und seien $A = \mathrm{Abb}(X, V)$ und $B = \mathrm{Abb}(Y, V)$. Nach Prop. (4.4) gibt es Verknüpfungen $\oplus_A$, $\oplus_B$ auf A, B und Abbildungen $\odot_A : K \times A \to A$ und $\odot_B : K \times B \to B$, so dass $(A, \oplus_A, \odot_A)$ und $(B, \oplus_B, \odot_B)$ zu K-Vektorräumen werden.

Proposition 4.10 Sei $u : X \to Y$ eine Abbildung. Dann ist durch $\phi_u : B \to A$, $f \mapsto f \circ u$ eine lineare Abbildung zwischen $(A, \oplus_A, \odot_A)$ und $(B, \oplus_B, \odot_B)$ gegeben.

Beweis: Seien $f, g \in B$ und $\lambda \in K$ vorgegeben. Zu zeigen ist $\phi_u(f \oplus_B g) = \phi_u(f) \oplus_A \phi_u(g)$ und $\phi_u(\lambda \odot_B f) = \lambda \odot_A \phi_u(f)$. Zum Beweis der ersten Gleichung sei $x \in X$ ein beliebiges Element. Dann gilt

$$\begin{aligned} \phi_u(f \oplus_B g)(x) \quad &= \quad ((f \oplus_B g) \circ u)(x) \quad = \quad (f \oplus_B g)(u(x)) \quad = \quad f(u(x)) +_V g(u(x)) \\ &= \quad (f \circ u)(x) +_V (g \circ u)(x) \quad = \quad \phi_u(f)(x) +_V \phi_u(g)(x) \quad = \quad (\phi_u(f) \oplus_A \phi_u(g))(x). \end{aligned}$$

Weil $x \in X$ beliebig vorgegeben war, folgt daraus $\phi_u(f \oplus_B g) = \phi_u(f) \oplus_A \phi_u(g)$. Zum Beweis der zweiten Gleichung betrachten wir wiederum ein beliebiges $x \in X$. Es gilt

$$\begin{aligned} \phi_u(\lambda \odot_B f)(x) \quad &= \quad ((\lambda \odot_B f) \circ u)(x) \quad = \quad (\lambda \odot_B f)(u(x)) \quad = \quad \lambda \cdot_V f(u(x)) \\ &= \quad \lambda \cdot_V (f \circ u)(x) \quad = \quad \lambda \cdot_V \phi_u(f)(x) \quad = \quad (\lambda \odot_A \phi_u(f))(x) \end{aligned}$$

und somit $\phi_u(\lambda \odot_B f) = \lambda \odot_A \phi_u(f)$. □

Proposition 4.11 Seien U, V, W drei K-Vektorräume und $\phi : U \to V$, $\psi : V \to W$ lineare Abbildungen. Dann ist $\psi \circ \phi$ eine lineare Abbildung von U nach W. Ist ϕ bijektiv, dann ist ϕ^{-1} eine lineare Abbidlung von V nach U.

Beweis: Wir überprüfen die Linearität der Abbildung $\psi \circ \phi$. Seien dazu $v, w \in U$ und $\lambda \in K$ vorgegeben. Dann gilt

$$\begin{aligned} (\psi \circ \phi)(v +_U w) \quad &= \quad \psi(\phi(v +_U w)) \quad = \quad \psi(\phi(v) +_V \phi(w)) \quad = \quad \psi(\phi(v)) +_W \psi(\phi(w)) \\ &= \quad (\psi \circ \phi)(v) +_W (\psi \circ \phi)(w) \end{aligned}$$

und $(\psi \circ \phi)(\lambda v) = \psi(\phi(\lambda v)) = \psi(\lambda\phi(v)) = \lambda\psi(\phi(v)) = \lambda(\psi \circ \phi)(v)$. Setzen wir nun voraus, dass ϕ bijektiv ist. Um zu zeigen, dass die Abbildung ϕ^{-1} linear ist, seien $v, w \in V$ und $\lambda \in K$ vorgegeben. Sei $v' = \phi^{-1}(v)$ und $w' = \phi^{-1}(w)$. Unter Verwendung der Linearität von ϕ erhalten wir

$$\begin{aligned} \phi^{-1}(v) +_U \phi^{-1}(w) &= v' +_U w' = \mathrm{id}_U(v' +_U w') = (\phi^{-1} \circ \phi)(v' +_U w') \\ &= \phi^{-1}(\phi(v' +_U w')) = \phi^{-1}(\phi(v') +_V \phi(w')) = \phi^{-1}(v +_V w). \end{aligned}$$

Ebenso gilt $\lambda\phi^{-1}(v) = \lambda v' = \mathrm{id}_U(\lambda v') = (\phi^{-1} \circ \phi)(\lambda v') = \phi^{-1}(\phi(\lambda v')) = \phi^{-1}(\lambda\phi(v')) = \phi^{-1}(\lambda v)$. □

Definition 4.12 Eine lineare Abbildung $\phi : V \to W$ heißt

(i) ***Monomorphismus*** (von K-Vektorräumen), wenn ϕ injektiv ist,
(ii) ***Epimorphismus***, wenn ϕ surjektiv ist,
(iii) ***Isomorphismus***, wenn ϕ bijektiv ist.

Eine lineare Abbildung $\phi : V \to V$ bezeichnet man als ***Endomorphismus*** von V, und ist sie außerdem bijektiv, dann spricht man von einem ***Automorphismus***. Zwei K-Vektorräume V, W werden ***isomorph*** genannt, wenn ein Isomorphismus $\phi : V \to W$ existiert.

Folgerung 4.13 Die Menge der Automorphismen eines K-Vektorraums V ist mit der Komposition $\circ$ von Abbildungen als Verknüpfung eine Gruppe. Man bezeichnet sie mit $\mathrm{GL}(V)$ und nennt sie die ***allgemeine lineare Gruppe*** des Vektorraums V.

Beweis: Die vorhergende Proposition zeigt, dass mit $\phi, \psi : V \to V$ auch die Abbildungen $\psi \circ \phi$ und ϕ^{-1} Automorphismen des Vektorraums V sind. Die Assoziativität ergibt sich aus der allgemeinen Regel $h \circ (g \circ f) = (h \circ g) \circ f$ für beliebige Abbildungen zwischen Mengen. Die identische Abbildung id_V besitzt die definierende Eigenschaft des Neutralelements (es gilt $\phi \circ \mathrm{id}_V = \mathrm{id}_V \circ \phi = \phi$ für jeden Automorphismus ϕ von V), und die Umkehrabbildung ϕ^{-1} von ϕ erfüllt die Bedingung $\phi^{-1} \circ \phi = \phi \circ \phi^{-1} = \mathrm{id}_V$ für das inverse Element. □

§ 5. Untervektorräume

Definition 5.1 Sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ wird ***Unterverktorraum*** von V genannt, wenn folgende Bedingungen erfüllt sind.

(i) $0_V \in U$

(ii) $v + w \in U$ für alle $v, w \in U$

(iii) $\lambda v \in U$ für alle $\lambda \in K$ und $v \in U$

Satz 5.2 Sei $(V, +, \cdot)$ ein Vektorraum und $U \subseteq V$ ein Untervektorraum von V. Definieren wir Abbildungen $+_U : U \times U \to V$ und $\cdot_U : K \times U \to V$ durch

$$v +_U w \quad = \quad v + w \qquad \text{und} \qquad \lambda \cdot_U v \quad = \quad \lambda \cdot v \qquad \text{für} \quad v, w \in U \text{ und } \lambda \in K,$$

dann ist durch $(U, +_U, \cdot_U)$ ein K-Vektorraum gegeben.

Beweis: Weil U ein Untervektorraum ist, gilt $v + w \in U$ für alle $v, w \in U$, also auch $v +_U w = v + w \in U$. Dies zeigt, dass $+_U$ eine Abbildung $U \times U \to U$, also eine Verknüpfung auf U gegeben ist. Ebenso gilt $\lambda \cdot_U v = \lambda \cdot v \in U$ für alle $v \in U$ und $\lambda \in K$. Somit ist $\cdot_U$ eine Abbildung $\cdot_U : K \times U \to U$. Wir müssen nun überprüfen, dass $(U, +_U, \cdot_U)$ die Vektorraum-Bedingungen aus Def. (4.1) erfüllt.

Zunächst überprüfen wir, dass $(U, +_U)$ eine abelsche Gruppe ist. Für alle $u, v, w \in U$ gilt $(u +_U v) +_U w = (u+v)+w = u+(v+w) = u+_U(v+_U w)$, also ist das Assoziativgesetz erfüllt. Nach Voraussetzung liegt 0_V in U, und für alle $v \in U$ gilt $u+_U 0_V = u+0_V = u$ und $0_V +_U u = 0_V + u = u$. Damit besitzt 0_V die Eigenschaften des Neutralelements in $(U, +_U)$. Sei nun $v \in U$ vorgegeben. Nach Voraussetzung liegt der Vektor $-v = (-1)v$ in U. Außerdem gilt $v +_U (-v) = v + (-v) = 0_V$ und $(-v) +_U v = (-v) + v = 0_V$. Also besitzt jedes $v \in U$ in $(U, +_U)$ ein Inverses, nämlich $-v$. Ingesamt bedeutet dies, dass $(U, +_U)$ eine Gruppe ist. Das Kommutativgesetz erhält man durch die Rechnung $v +_U w = v + w = w + v = w +_U v$ für alle $v, w \in U$.

Nun müssen wir noch die Eigenschaften (ii) (a)-(d) aus Def. (4.1) überprüfen. Seien dazu $v, w \in U$ und $\lambda, \mu \in K$ vorgegeben. Es gilt $(\lambda + \mu) \cdot_U v = (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v = \lambda \cdot_U v +_U \mu \cdot_U v$. Ebenso erhält man $\lambda \cdot_U (v +_U w) = \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w = \lambda \cdot_U v +_U \lambda \cdot_U w$. Weiter gilt $(\lambda\mu) \cdot_U v = (\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v) = \lambda \cdot_U (\mu \cdot_U v)$ und schließlich $1_K \cdot_U v = 1_K \cdot v = v$. □

Man sieht, dass das Nachrechnen der Vektorraum-Axiome in $(U, +_U, \cdot_U)$ eine ziemliche Routineangelegenheit war: Überall wurden nur die Symbole $+_U$ und $\cdot_U$ durch $+$ und $\cdot$ ersetzt und anschließend verwendet, dass die Axiome im Vektorraum V gültig sind.

Folgende konkrete Beispiele lassen sich für Untervektorräume angeben.

(i) Ist V ein beliebiger K-Vektorraum, dann sind $\{0_V\}$ und V Untervektorräume von V.

(ii) Für jedes $v \in V$ ist $\mathrm{lin}(v) = \{\lambda v \mid \lambda \in K\}$ ein Untervektorraum. Im Fall $v \neq 0_V$ bezeichnet man ihn als ***lineare Gerade***. Für beliebige v, w ist auch durch

$$\mathrm{lin}(v, w) \quad = \quad \{\lambda v + \mu w \mid \lambda, \mu \in K\}$$

ein Untervektorraum gegeben. Ist $v \neq 0_V$ und w kein skalares Vielfaches von v (also $w \neq \lambda v$ für alle $\lambda \in K$), dann nennt man $\mathrm{lin}(v, w)$ eine ***lineare Ebene***.

(iii) Die Lösungsmenge $\mathscr{L} \subseteq K^n$ eines homogenen linearen Gleichungssystems bestehend aus m Gleichungen in n Unbekannten ist ein Untervektorraum von K^n. Denn der Nullvektor 0_{K^n} ist immer in $\mathscr{L}$ enthalten, und die Bedingung $v, w \in \mathscr{L}, \lambda \in K \Rightarrow v + w, \lambda w \in \mathscr{L}$ haben wir bereits in Prop. (1.5) nachgerechnet.

(iv) Seien $a, b \in \mathbb{R}$, $a < b$, und sei V der Vektorraum der reellwertigen Funktionen auf dem abgeschlossenen Intervall $[a, b]$. Dann bilden die stetigen Funktionen auf $[a, b]$ jeweils einen Untervektorraum von V. Denn aus der Analysis einer Variablen ist bekannt, dass mit f, g auch die Funktionen $f + g$ und λf stetig sind, für jedes $\lambda \in \mathbb{R}$.

Proposition 5.3 Seien V und W zwei K-Vektorräume. Dann bilden die linearen Abbildungen $\phi : V \to W$ einen Untervektorraum von $\mathrm{Abb}(V, W)$.

Beweis: Sei $A = \mathrm{Abb}(V, W)$ und $U \subseteq A$ die Teilmenge der linearen Abbildungen zwischen V und W. Zunächst müssen wir zeigen, dass die Abbildung 0_A, die jeden Vektor aus V auf 0_W abbildet, in U enthalten ist. Seien dazu $v, v' \in V$ und $\lambda \in K$ vorgegeben. Es gilt $0_A(v + v') = 0_A(v) + 0_A(v') = 0_W + 0_W = 0_W$ und $0_A(\lambda v) = \lambda 0_A(v) = \lambda \cdot 0_W = 0_W$. Also ist 0_A in U enthalten.

Seien nun $\phi, \psi \in U$ und $\lambda \in K$ vorgegeben. Zu zeigen ist $\phi + \psi \in U$ und $\lambda\phi \in U$, d.h. wir müssen nachrechnen, dass diese Abbildungen linear sind. Seien dazu $v, v' \in V$ und $\mu \in K$. Weil ϕ und ψ linear sind, gilt

$$(\phi + \psi)(v + v') \quad = \quad \phi(v + v') + \psi(v + v') \quad = \quad \phi(v) + \phi(v') + \psi(v) + \psi(v') \quad =$$

$$\phi(v) + \psi(v) + \phi(v') + \psi(v') \quad = \quad (\phi + \psi)(v) + (\phi + \psi)(v')$$

und ebenso

$$(\phi + \psi)(\mu v) \quad = \quad \phi(\mu v) + \psi(\mu v) \quad = \quad \mu\phi(v) + \mu\psi(v) \quad = \quad \mu(\phi(v) + \psi(v)) \quad = \quad \mu(\phi + \psi)(v).$$

Damit ist $\phi + \psi \in U$ nachgewiesen. Nun zeigen wir noch, dass $\lambda\phi$ in U liegt. Es gilt

$$(\lambda\phi)(v + v') \quad = \quad \lambda\phi(v + v') \quad = \quad \lambda(\phi(v) + \phi(v')) \quad = \quad \lambda\phi(v) + \lambda\phi(v') \quad = \quad (\lambda\phi)(v) + (\lambda\phi)(v')$$

und $(\lambda\phi)(\mu v) = \lambda\phi(\mu v) = \lambda\mu\phi(v) = \mu\lambda\phi(v) = \mu(\lambda\phi)(v)$. Also ist $\lambda\phi$ eine lineare Abbildung. □

Wegen Satz (5.2) zeigt Prop. (5.3) insbesondere, dass die linearen Abbildungen zwischen zwei K-Vektorräumen V, W selbst einen K-Vektorraum bilden. Man bezeichnet diesen üblicherweise mit $\mathrm{Hom}_K(V, W)$.

Proposition 5.4 Sei V ein K-Vektorraum, und seien U, U' Untervektorräume von V. Dann sind auch die Mengen $U \cap U'$ und

$$U \cap U' \qquad \text{und} \qquad U + U' = \{u + u' \mid u \in U, u' \in U'\} \qquad \text{Untervektorräume von } V.$$

Man bezeichnet $U + U'$ als die ***Summe*** von U und U'.

Beweis: Zunächst beweisen wir die Untervektorraum-Eigenschaft von $U \cap U'$. Weil U, U' nach Voraussetzung Untervektorräume sind, gilt $0_V \in U$ und $0_V \in U'$. Es folgt $0_V \in U \cap U'$. Seien nun Elemente $v_1, v_2 \in U \cap U'$ und $\lambda \in K$ beliebig vorgegeben. Dann gilt insbesondere $v_1, v_2 \in U$. Weil U ein Untervektorraum ist, folgt $v_1 + v_2 \in U$ und $\lambda v_1 \in U$, und ebenso gilt $v_1 + v_2 \in U'$ und $\lambda v_1 \in U'$, weil U' ein Untervektorraum ist. Aus $v_1 + v_2 \in U$ und $v_1 + v_2 \in U'$ folgt $v_1 + v_2 \in U \cap U'$, ebenso erhalten wir $\lambda v_1 \in U \cap U'$. Damit sind die Untervektorraum-Eigenschaften für die Menge $U \cap U'$ nachgewiesen.

Nun zeigen wir, dass auch die Menge $U + U'$ ein Untervektorraum von V ist. Wegen $0_V \in U$ und $0_V \in U'$ gilt zunächst $0_V = 0_V + 0_V \in U + U'$. Seien nun $v_1, v_2 \in U + U'$ und $\lambda \in K$ vorgegeben. Dann gibt es $u_1, u_2 \in U$ und $u'_1, u'_2 \in U'$ mit $v_1 = u_1 + u'_1$ und $v_2 = u_2 + u'_2$. Weil U ein Untervektorraum ist, gilt $u_1 + u_2 \in U$, ebenso gilt $u'_1 + u'_2 \in U'$. Es folgt $v_1 + v_2 = (u_1 + u'_1) + (u_2 + u'_2) = (u_1 + u_2) + (u'_1 + u'_2) \in U + U'$. Aus der Untervektorraum-Eigenschaft von U und U' folgt auch, dass $\lambda u_1 \in U$ und $\lambda u'_1 \in U'$ gilt. Wir erhalten $\lambda v_1 = \lambda(u_1 + u'_1) = \lambda u_1 + \lambda u'_1 \in U + U'$. Damit haben wir auch die Untervektorraum-Eigenschaften von $U + U'$ nachgerechnet.
□

Auch aus mehr als zwei Untervektorräumen kann eine Summe gebildet werden. Sei V ein K-Vektorraum, und sei $U_1, U_2, U_3, \ldots$ eine beliebige Anzahl von Untervektorräumen von V. Man definiert

$$\sum_{k=1}^{1} U_k = U_1 \qquad \text{und} \qquad \sum_{k=1}^{r+1} U_k = \left(\sum_{k=1}^{r} U_k\right) + U_{r+1} \quad \text{für} \quad r \geq 1.$$

Der Nachweis, dass es sich bei $\sum_{k=1}^{r} U_k$ für jedes $r \in \mathbb{N}$ um einen Untervektorraum von V handelt, erfolgt durch vollständige Induktion über r, was hier aus Zeitgründen aber nicht ausgeführt wird. Ebenso zeigt man durch vollständige Induktion, dass

$$\sum_{k=1}^{r} U_k \quad = \quad \left\{ \sum_{k=1}^{r} u_k \;\middle|\; u_k \in U_k \text{ für } 1 \leq k \leq r \right\} \qquad \text{gilt.}$$

Definition 5.5 Ein K-Vektorraum V wird ***direkte Summe*** der Untervektorräume $U, U' \subseteq V$ genannt, wenn die Bedingungen

$$V = U + U' \quad \text{und} \quad U \cap U' = \{0_V\} \qquad \text{erfüllt sind.}$$

Die direkte Summe zweier Untervektorräume U, U' wird mit $U \oplus U'$ bezeichnet.

Lemma 5.6 Sei V ein K-Vektorraum mit Untervektorräumen $U, U' \subseteq V$. Dann sind die folgenden Bedingungen äquivalent:

(i) $V = U \oplus U'$

(ii) Für jedes $v \in V$ gibt es eindeutig bestimmte Vektoren $u \in U$ und $u' \in U'$ mit $v = u + u'$.

Beweis: „(i) ⇒ (ii)" Wegen $V = U + U'$ gibt es für jeden Vektor $v \in V$ Elemente $u \in U$ und $u' \in U'$ mit $v = u + u'$. Wir beweisen nun die Eindeutigkeit. Sei $v \in V$, und seien $u_1, u_2 \in U$ und $u'_1, u'_2 \in U'$ mit $v = u_1 + u'_1 = u_2 + u'_2$. Dann gilt $u_1 - u_2 = u'_2 - u'_1 \in U \cap U'$. Wegen $U \cap U' = \{0_V\}$ folgt $u_1 - u_2 = u'_2 - u'_1 = 0_V$, also $u_1 = u_2$ und $u'_1 = u'_2$.

„(ii) ⇒ (i)" Weil jeder Vektor $v \in V$ in der Form $v = u + u'$ mit $u \in U$ und $u \in U'$ geschrieben werden kann, gilt $V = U + U'$. Wir zeigen nun, dass auch $U \cap U' = \{0_V\}$ erfüllt ist. Die Inklusion „$\supseteq$" ist offensichtlich, da U und U' Untervektorräume sind und somit 0_V in U und U' enthalten ist. Sei nun $v \in U \cap U'$ vorgegeben. Nach Voraussetzung gibt es eindeutig bestimmte $u \in U$, $u' \in U'$ mit $v = u + u'$. Aus $v = 0_V + v$ mit $0_V \in U$ und $v \in U'$ folgt auf Grund der Eindeutigkeit $u = 0_V$. Ebenso können wir v auch in der Form $v = v + 0_V$ mit $v \in U$ und $0_V \in U'$ schreiben. Diesmal liefert die Eindeutigkeit die Gleichung $u' = 0_V$. Insgesamt erhalten wir $v = u + u' = 0_V + 0_V = 0_V$. □

Auch die direkte Summe von mehreren Untervektorräumen lässt sich rekursiv definieren. Sei V ein K-Vektorraum, und seien $U_1, U_2, U_3, \ldots$ Untervektorräume von V. Dann setzt man

$$\bigoplus_{k=1}^{1} U_k = U_1 \qquad \text{und} \qquad \bigoplus_{k=1}^{r+1} U_k = \left(\bigoplus_{k=1}^{r} U_k\right) \oplus U_{r+1} \quad \text{für} \quad r \geq 1.$$

Damit die direkte Summe $\bigoplus_{k=1}^{r+1} U_k$ gebildet werden kann, dürfen sich $\bigoplus_{k=1}^{r} U_k$ und U_{r+1} jeweils nur in $\{0_V\}$ schneiden.

Satz 5.7 Sei V ein Vektorraum, $r \in \mathbb{N}$, und seien $U_1, \ldots, U_r$ Untervektorräume von V. Dann sind die folgenden Aussagen äquivalent.

(i) Es gilt $V = \bigoplus_{k=1}^{r} U_k$.

(ii) Jeder Vektor $v \in V$ kann auf eindeutige Weise als Summe $v = \sum_{k=1}^{r} u_k$ dargestellt werden, mit $u_k \in U_k$ für $1 \leq k \leq r$.

(iii) Für $1 \leq k \leq r$ gilt $V = \sum_{k=1}^{r} U_k$ und $U_k \cap \left(\sum_{j \neq k} U_j\right) = \{0_V\}$.

Beweis: Wir beweisen die Äquivalenz der drei Aussagen durch vollständige Induktion über r. Im Fall $r = 1$ besteht (i) nur in der Aussage $V = U_1$. Aussage (ii) besagt, dass für jedes $v \in V$ ein eindeutig bestimmter

Vektor $u_1 \in U_1$ mit $v = u_1$ existiert, was offenbar zu (i) äquivalent ist. Die Aussage (iii) besteht aus den Gleichungen $V = U_1$ und $U_1 \cap \{0_V\} = \{0_V\}$, und wiederum ist „(i) ⇔ (iii)" offensichtlich.

Sei nun $r \in \mathbb{N}$ vorgegeben, und setzen wir die Äquivalenz von (i), (ii) und (iii) für dieses r voraus. Seien $U_1, ..., U_{r+1}$ beliebige Untervektorräume von V. Wir beginnen mit dem Beweis der Implikation „(i) ⇒ (ii)". Hier lautet die Voraussetzung

$$V \quad = \quad \bigoplus_{k=1}^{r+1} U_k \quad = \quad \left(\bigoplus_{k=1}^{r} U_k \right) \oplus U_{r+1}.$$

Insbesondere gilt $V = \sum_{k=1}^{r+1} U_k$; dies bedeutet, dass jedes $v \in V$ jedenfalls als Summe $v = u_1 + ... + u_{r+1}$ dargestellt werden kann, mit $u_k \in U_k$ für $1 \le k \le r+1$. Nehmen wir nun an, dass $v = u'_1 + ... + u'_{r+1}$ eine weitere solche Darstellung ist. Weil V nach Voraussetzung die direkte Summe von $\bigoplus_{k=1}^{r} U_k$ und U_{r+1} ist, folgt $u_1 + ... + u_r = u'_1 + ... + u'_r$ und $u_{r+1} = u'_{r+1}$ nach Lemma (5.6). Nach Induktionsvoraussetzung besitzt jedes Element in $\bigoplus_{k=1}^{r} U_k$ eine eindeutige Darstellung als Summe von Elementen in $U_1, ..., U_k$. Aus $u_1 + ... + u_r = u'_1 + ... + u'_r$ folgt also $u_k = u'_k$ für $1 \le k \le r$.

Beweisen wir nun die Implikation „(ii) ⇒ (iii)" und setzen dazu (ii) voraus. Zunächst zeigen wir die Gleichung $V = \sum_{k=1}^{r+1} U_k$. Die Inklusion „⊇" ist nach Definition der Summe offensichtlich. Andererseits hat auf Grund unserer Voraussetzung jedes Element $v \in V$ eine Darstellung $v = u_1 + ... + u_{r+1}$ mit $u_k \in U_k$ für $1 \le k \le r+1$. Also gilt auch „⊆". Sei nun $k \in \{1, ..., r+1\}$ vorgegeben. Zu zeigen ist die Gleichung

$$U_k \cap \left(\sum_{j \neq k} U_j \right) \quad = \quad \{0_V\}.$$

Hier ist „⊇" offensichtlich erfüllt. Zum Beweis von „⊆" nehmen wir an, dass ein Vektor $v \neq 0_V$ im Durchschnitt existiert. Dann liegt v einerseits in U_k, andererseits gilt $v = \sum_{j \neq k} u_j$ für gewisse Elemente u_j mit $u_j \in U_j$ für $1 \le j \le r+1$ und $j \neq k$. Setzen wir $u_k = -v$, dann gilt $\sum_{j=1}^{r+1} u_j = 0_V$. Weil aber der Nullvektor auch in der Form $0_V + ... + 0_V$ mit $0_V \in U_j$ für $1 \le j \le r+1$ dargestellt werden kann, und weil diese Darstellung nach (ii) eindeutig ist, folgt $u_j = 0_V$ für $1 \le j \le r+1$ mit $j \neq k$ und auch $v = -u_k = 0_V$, im Widerspruch zur Annahme.

Zeigen wir nun noch die Implikation „(iii) ⇒ (i)" und setzen dazu (iii) voraus. Zu zeigen ist

$$V \quad = \quad \bigoplus_{k=1}^{r+1} U_k \quad = \quad \left(\bigoplus_{k=1}^{r} U_k \right) \oplus U_{r+1}.$$

Wir betrachten den Untervektorraum $U = \sum_{k=1}^{r} U_k$. Nach Voraussetzung gilt $U_k \cap (\sum_{j \neq k} U_j) = \{0_V\}$ für $1 \le k \le r+1$. Damit ist für $1 \le k \le r$ jeweils erst recht der Durchschnitt von U_k mit $\sum_{j \neq k, r+1} U_j$ gleich $\{0_V\}$. Also ist die Bedingung (iii) für den K-Vektorraum U und die Untervektorräume $U_1, ..., U_r$ von U erfüllt. Die Induktionsvoraussetzung liefert uns damit

$$U \quad = \quad \bigoplus_{k=1}^{r} U_k.$$

Weiter gilt nach Voraussetzung $V = U + V_{r+1}$, außerdem $U \cap V_{r+1} = \{0_V\}$. Wie gewünscht erhalten wir $V = U \oplus V_{r+1} = \bigoplus_{k=1}^{r+1} U_k$. □

Proposition 5.8 Seien V, W K-Vektorräume und $\phi : V \to W$ eine lineare Abbildung. Ferner seien $V' \subseteq V$ und $W' \subseteq W$ Untervektorräume. Dann sind die Teilmengen

$$\phi(V') = \{\phi(v) \mid v \in V'\} \qquad \text{und} \qquad \phi^{-1}(W') = \{v \in V \mid \phi(v) \in W'\}$$

Untervektorräume von W bzw. von V.

Beweis: Wir rechnen die Untervektorraum-Axiome für beide Teilmengen direkt nach. Seien $w, w' \in \phi(V')$ und $\lambda \in K$. Dann gibt es nach Definition von $\phi(V')$ Vektoren $v, v' \in V'$ mit $w = \phi(v)$ und $w' = \phi(v')$. Da V' ein Untervektorraum ist, gilt $v + v' \in V'$ und damit

$$w + w' = \phi(v) + \phi(v') = \phi(v + v') \quad \in \quad \phi(V').$$

Ebenso gilt $\lambda v \in V'$ auf Grund der Untervektorraum-Eigenschaft und somit $\lambda w = \lambda\phi(v) = \phi(\lambda v) \in \phi(V')$.

Nun zeigen wir, dass auch $\phi^{-1}(W')$ ein Untervektorraum ist. Seien dazu $v, v' \in \phi^{-1}(W')$ und $\lambda \in K$ vorgegeben. Dann gilt $\phi(v), \phi(v') \in W'$ und $\phi(v) + \phi(v') \in W'$, da W' ein Untervektorraum von W ist. Aus $\phi(v + v') = \phi(v) + \phi(v') \in W'$ folgt $v + v' \in \phi^{-1}(W')$. Da auch $\lambda\phi(v)$ in W' liegt, erhalten wir $\phi(\lambda v) = \lambda\phi(v) \in W'$ und somit $\lambda v \in \phi^{-1}(W')$. □

Definition 5.9 Seien V, W zwei K-Vektorräume und $\phi : V \to W$ eine lineare Abbildung. Dann nennt man

(i) $\ker(\phi) = \phi^{-1}(\{0_W\}) = \{ v \in V \mid \phi(v) = 0_W \}$ den ***Kern*** und

(ii) $\mathrm{im}(\phi) = \phi(V) = \{ \phi(v) \mid v \in V \}$ das ***Bild*** von ϕ.

Nach Prop. (5.8) ist $\ker(\phi)$ ein Untervektorraum von V und $\mathrm{im}(\phi)$ ein Untervektorraum von W.

Proposition 5.10 Seien V, W K-Vektorräume und $\phi : V \to W$ eine lineare Abbildung.

(i) Die Abbildung ϕ ist genau dann surjektiv, wenn $\mathrm{im}(\phi) = W$ gilt.

(ii) Sie ist genau dann injektiv, wenn $\ker(\phi) = \{0_V\}$ erfüllt ist.

Beweis: Aussage (i) ist nach Definition der Surjektivität unmittelbar klar. Zum Beweis von (ii) setzen wir zunächst voraus, dass ϕ injektiv ist. Die Inklusion $\{0_V\} \subseteq \ker(\phi)$ ist erfüllt, weil der Kern ein Untervektorraum von V ist. Zum Nachweis von $\ker(\phi) \subseteq \{0_V\}$ sei $v \in \ker(\phi)$ vorgegeben. Dann gilt $\phi(v) = 0_W = \phi(0_V)$, und aus der Injektivität von ϕ folgt $v = 0_V$.

Setzen wir nun umgekehrt die Gleichung $\ker(\phi) = \{0_V\}$ voraus, und beweisen wir die Injektivität von ϕ. Seien dazu $v, v' \in V$ mit $\phi(v) = \phi(v')$ vorgegeben. Dann gilt $\phi(v' - v) = \phi(v') - \phi(v) = 0_W$ und somit $v' - v \in \ker(\phi)$. Aus der Voraussetzung an den Kern folgt $v' - v = 0_V \Leftrightarrow v = v'$. □

Definition 5.11 Eine Teilmenge $A \subseteq V$ wird ***affiner Unterraum*** von V genannt, wenn entweder $A = \emptyset$ gilt oder ein Untervektorraum U und ein Vektor $v \in V$ existieren, so dass

$$A = v + U = \{v + u \mid u \in U\} \qquad \text{erfüllt ist.}$$

Betrachten wir einige konkrete Beispiele für affine Unterräume.

(i) Seien $u, v \in V$. Dann ist $u + \mathrm{lin}(v) = \{u + \lambda v \mid \lambda \in K\}$ ein affiner Unterraum. Im Fall $v \neq 0_V$ bezeichnet man ihn als ***affine Gerade***. Für beliebige $u, v, w \in V$ ist auch durch

$$u + \mathrm{lin}(v, w) = \{u + \lambda v + \mu w \mid \lambda, \mu \in K\}$$

ein affiner Unterraum gegeben. Ist $v \neq 0_V$ und w kein skalares Vielfaches von v (also $w \neq \lambda v$ für alle $\lambda \in K$), dann nennt man $\mathrm{lin}(v, w)$ eine ***affine Ebene***.

(ii) Die Lösungsmenge $\mathscr{L} \subseteq K^n$ eines beliebigen linearen Gleichungssystems bestehend aus m Gleichungen in n Unbekannten ist ein affiner Unterraum von K^n. Dies folgt aus der Tatsache, dass jede Lösungsmenge eines *homogenen* LGS ein Untervektorraum von K^n ist, in Verbindung mit Prop. (1.6)

Proposition 5.12 Sei V ein K-Vektorraum und $\emptyset \neq A \subseteq V$ ein affiner Unterraum.

(i) Es gibt einen eindeutig bestimmten Untervektorraum U, so dass $A = v + U$ für ein $v \in V$ erfüllt ist.

(ii) Für jeden Vektor $w \in A$ gilt $A = w + \mathscr{L}(A)$.

Wir nennen U den zu A gehörenden Untervektorraum und bezeichnen ihn mit $\mathscr{L}(A)$.

Beweis: zu (i) Nehmen wir an, dass $v, v' \in V$ Vektoren und U, U' Untervektorräume von V mit $v + U = A = v' + U'$ sind. Wegen $v \in A$ gilt $v = v' + u_0$ für ein $u_0 \in U'$, und wegen $v' \in A$ gilt $v' = v + u_1$ für ein $u_1 \in U$. Der Differenzvektor $v' - v$ ist also sowohl in U als auch in U' enthalten. Wir beweisen nun die Gleichung $U = U'$.

„$\subseteq$" Ist $u \in U$, dann liegt $v + u$ in A, und folglich gibt es ein $u' \in U'$ mit $v + u = v' + u'$. Es folgt $u = (v' - v) + u' \in U'$. „$\supseteq$" Ist $u' \in U'$ vorgegeben, dann gilt $v' + u'$ in A, es gibt also ein $u \in U$ mit $v' + u' = v + u$. Daraus folgt $u' = (v - v') + u \in U$.

zu (ii) Sei $U = \mathscr{L}(A)$, $v \in V$ ein Vektor mit $A = v + U$ und $w \in A$ ein beliebiges Element. Dann gibt es ein $u \in U$ mit $w = v + u$. Wir beweisen nun die Gleichung $v + U = w + U$. „$\subseteq$" Ist $v_1 \in v + U$, dann gibt es ein $u_1 \in U$ mit $v_1 = v + u_1$, und es folgt $v_1 = (w - u) + u_1 = w + (u_1 - u) \in w + U$. „$\supseteq$" Ist $w_1 \in w + U$, dann existiert ein $u_1 \in U$ mit $w_1 = w + u_1$. Es folgt $w_1 = w + u_1 = (v + u) + u_1 = v + (u + u_1) \in v + U$. □

§ 6. *Erzeugendensysteme und lineare Unabhängigkeit*

Definition 6.1 Sei V ein K-Vektorraum und $S \subseteq V$ eine beliebige Teilmenge. Wir bezeichnen einen Vektor $w \in V$ als ***Linearkombination*** von S, wenn eine endliche Teilmenge $T \subseteq S$ und eine Familie $(\lambda_v)_{v \in T}$ in K existieren, so dass

$$w \quad = \quad \sum_{v \in T} \lambda_v v \qquad \text{erfüllt ist.}$$

Die Menge aller Linearkombinationen von S in V bezeichnen wir mit $\mathrm{lin}(S)$. Gilt $V = \mathrm{lin}(S)$, dann wird S ein ***Erzeugendensystem*** von V genannt. In diesem Fall sagt man auch, dass der Vektorraum V von S ***erzeugt*** oder ***aufgespannt*** wird.

Im Fall, dass T die leere Menge und $(\lambda_v)_{v \in T}$ entsprechend die „leere“ Familie in K ist, setzen wir $\sum_{v \in T} \lambda_v v = 0_V$. Daraus folgt, dass der Nullvektor in $\mathrm{lin}(S)$ enthalten ist; insbesondere gilt $\mathrm{lin}(\varnothing) = \{0_V\}$. Der Definition unmittelbar zu entnehmen ist, dass aus $S \subseteq S'$ stets $\mathrm{lin}(S) \subseteq \mathrm{lin}(S')$ folgt.

So lange die Menge S endlich ist, $S = \{v_1, ..., v_n\}$, lässt sich die Menge $\mathrm{lin}(S)$ der Linearkombinationen noch etwas einfacher beschreiben: Es gilt

$$\mathrm{lin}(S) \quad = \quad \left\{ \sum_{k=1}^{n} \lambda_k v_k \mid \lambda_1, ..., \lambda_n \in K \right\}.$$

Für die Vektorräume K^n und $\mathscr{M}_{m \times n, K}$ existieren endliche Erzeugendensysteme, nämlich im ersten Fall die Menge $\{e_1, ..., e_n\}$ der Einheitsvektoren und im zweiten die Menge $\{B_{k\ell} \mid 1 \le k \le m, 1 \le \ell \le n\}$ der Basismatrizen. Es gibt aber auch Vektorräume, die nicht durch keine endliche Teilmenge ihrer Vektoren aufgespannt werden können.

Satz 6.2 *(ohne Beweis)*

Für jeden Körper K gibt es einen Erweiterungsring $K[x] \supseteq K$ mit einem ausgezeichneten Element $x \notin K$, so dass für jedes $f \in K[x]$ folgende Bedingung erfüllt ist. ***E***ntweder ist $f = 0_K$, ***o***der es gibt ein eindeutig bestimmtes $n \in \mathbb{N}_0$ und eindeutig bestimmte Elemente $a_0, ..., a_n \in K$ mit

$$f \quad = \quad \sum_{k=0}^{n} a_k x^k \quad \text{und} \quad a_n \neq 0_K.$$

Man nennt $K[x]$ den ***Polynomring*** über dem Körper K, seine Elemente bezeichnet man als ***P***olynome. Im Fall $f \neq 0_K$ bezeichnet man n als den ***Grad*** des Polynoms.

Es sei noch einmal hervorgehoben, dass das Element x des Polynomrings $K[x]$ kein Element des Körpers K ist! Man kann $K[x]$ als K-Vektorraum betrachten, in dem man die Vektoraddition mit der gewöhnlichen

Addition im Ring $K[x]$ gleichsetzt und die skalare Multiplikation $K \times K[x] \to K[x]$ durch Einschränkung der Multiplikationsabbildung

$$K[x] \times K[x] \longrightarrow K[x] \quad , \quad (f,g) \mapsto fg$$

definiert. In diesem Fall ist dann $\{x^n \mid n \in \mathbb{N}_0\}$ ein (unendliches) Erzeugendensystem von $K[x]$ als K-Vektorraum, wobei $x^0 = 1_K$ ist. Es gibt in $K[x]$ keine endliche Teilmenge S mit $\mathrm{lin}(S) = K[x]$. Dies werden wir aber erst zu einem späteren Zeitpunkt zeigen können.

Satz 6.3 Sei V ein K-Vektorraum und $S \subseteq V$ eine Teilmenge. Dann gilt

(i) Die Menge $U = \mathrm{lin}(S)$ bildet einen Untervektorraum von V mit $U \supseteq S$.
(ii) Ist U' ein weiterer Untervektorraum von V mit $U' \supseteq S$, dann gilt $U' \supseteq U$.

Somit ist $\mathrm{lin}(S)$ der kleinste Untervektorraum von V, der S als Teilmenge enthält.

Beweis: Zunächst beweisen wir, dass U ein Untervektorraum von V ist. Dass 0_V in $\mathrm{lin}(S)$ liegt, haben wir bereits am Anfang dieses Kapitels festgestellt. Seien nun $v_1, v_2 \in U$ und $\lambda \in K$ vorgegeben. Dann gibt es endliche Teilmengen $T_1, T_2 \subseteq S$ und Familien $(\lambda_v)_{v \in T_1}$ und $(\mu_v)_{v \in T_2}$, so dass

$$v_1 = \sum_{v \in T_1} \lambda_v v \qquad \text{und} \qquad v_2 = \sum_{v \in T_2} \lambda_v v$$

erfüllt ist. Wir setzen nun $T = T_1 \cup T_2$ und erweitern $(\lambda_v)_{v \in T}$ zu einer Familie in K, indem wir $\lambda_v = 0_K$ für alle $v \in T \setminus T_1$ definieren. Ebenso erhalten wir eine Familie $(\mu_v)_{v \in T}$, indem wir $\mu_v = 0_K$ für $v \in T \setminus T_2$ setzen. Es gilt dann $v_1 = \sum_{v \in T} \lambda_v v$ und $v_2 = \sum_{v \in T} \lambda_v v$; daraus folgt

$$v_1 + v_2 \quad = \quad \sum_{v \in T_1} \lambda_v v + \sum_{v \in T_2} \mu_v v \quad = \quad \sum_{v \in T} \lambda_v v + \sum_{v \in T} \mu_v v \quad = \quad \sum_{v \in T} (\lambda_v + \mu_v) v.$$

Dies zeigt, dass $v_1 + v_2$ in U enthalten ist. Ebenso sieht man anhand der Gleichung

$$\lambda v_1 \quad = \quad \lambda \left(\sum_{v \in T_1} \lambda_v v \right) \quad = \quad \lambda \left(\sum_{v \in T} \lambda_v v \right) \quad = \quad \sum_{v \in T} \lambda \lambda_v v \quad ,$$

dass auch $\lambda v_1 \in U$ gilt. Damit haben wir nachgewiesen, dass es sich bei U tatsächlich um einen Untervektorraum von V handelt. Dieser enthält S als Teilmenge. Ist nämlich $u \in S$ vorgegeben, dann können wir die endliche Teilmenge $T = \{u\}$ und die Familie $(\lambda_v)_{v \in T}$ durch $\lambda_u = 1_K$ definieren. Offenbar gilt dann

$$\sum_{v \in T} \lambda_v v \quad = \quad \lambda_u u \quad = \quad 1_K u \quad = \quad u \qquad \text{und somit} \quad u \in U.$$

Sei nun U' ein beliebiger Untervektorraum von V mit $U' \supseteq S$. Wir zeigen durch vollständige Induktion über $r \in \mathbb{N}_0$, dass jede für jede Teilmenge $T \subseteq S$ mit $|T| = r$ und jede Familie $(\lambda_v)_{v \in T}$ in K die Linearkombination $\sum_{v \in T} \lambda_v v$ in U' enthalten ist. Im Fall $r = 0$ ist $T = \varnothing$ und somit $\sum_{v \in T} \lambda_v v = 0_V$. Weil es sich bei U' um einen Untervektorraum von V handelt, ist der Nullvektor auf jeden Fall in U' enthalten.

Sei nun $r \in \mathbb{N}_0$, und setzen wir die Aussage für dieses r voraus. Sei $T \subseteq S$ mit $|T| = r+1$ und $(\lambda_v)_{v\in T}$ eine Familie in K. Zu zeigen ist, dass der Vektor $u = \sum_{v\in T} \lambda_v v$ in U' liegt. Dazu wählen wir ein beliebiges Element $v_1 \in T$ und setzen $T_1 = T \setminus \{v_1\}$. Wegen $|T_1| = r$ liegt $u' = \sum_{v\in T_1} \lambda v$ nach Induktionsvoraussetzung in U'. Der Vektor v_1 ist wegen $T \subseteq S \subseteq U'$ ebenfalls in U' enthalten. Aus der Untervektorraum-Eigenschaft von U' folgt $\lambda_{v_1} v_1 \in U'$ und $u = u' + \lambda_{v_1} v_1 \in U'$. Damit ist der Induktionsschritt abgeschlossen. □

Definition 6.4 Sei V ein K-Vektorraum. Wir bezeichnen eine Teilmenge $S \subseteq V$ als ***linear abhängig***, wenn eine endliche Teilmenge $T \subseteq S$ und eine Familie $(\lambda_v)_{v\in T}$ in K existieren, so dass

$$\sum_{v\in T} \lambda_v v \quad = \quad 0_V$$

erfüllt ist und $\lambda_v \neq 0_K$ für mindestens ein $v \in T$ gilt. Ist dies nicht der Fall, dann wird S ***linear unabhängig*** genannt.

Ist S eine endliche Teilmenge von V, $S = \{v_1, ..., v_n\}$ mit $n = |S|$, dann ist die lineare Unabhängigkeit von S gleichbedeutend damit, dass für alle Tupel $(\lambda_1, ..., \lambda_n)$ in K^n die Implikation

$$\sum_{k=1}^{n} \lambda_k v_k = 0_V \quad \Rightarrow \quad \lambda_1 = ... = \lambda_n = 0_K$$

erfüllt ist. Beispielsweise ist die Menge $\{e_1, ..., e_n\}$ der Einheitsvektoren eine linear unabhängige Teilmenge von K^n, und die Menge $\{\, \mathrm{B}_{k\ell}^{(m\times n)} \mid 1 \le k \le m, 1 \le \ell \le n \,\}$ der Basismatrizen ist im K-Vektorraum $\mathcal{M}_{m\times n,K}$ der $m \times n$-Matrizen linear unabhängig. Die Teilmenge $\{x^n \mid n \in \mathbb{N}_0\}$ des Polynomrings $K[x]$ ist ein Beispiel für eine unendliche linear unabhängige Menge.

Man beachte, dass auch die leere Menge $\emptyset$ in jedem K-Vektorraum V nach Definition linear unabhängig ist. Eine einlementige Teilmenge $\{v\} \subseteq V$ ist genau dann linear abhängig, wenn $v = 0_V$ ist, ansonsten linear unabhängig. Allgemein ist jede Teilmenge, die den Nullvektor 0_V enthält, automatisch linear abhängig. Eine zweielementige Menge $\{v, w\}$ ist genau dann linear abhängig, wenn ein $\lambda \in K$ mit $v = \lambda w$ oder $w = \lambda v$ existiert.

Offenbar ist mit einer Menge S auch jede Teilmenge $S_1 \subseteq S$ linear unabhängig.

Proposition 6.5 Sei V ein K-Vektorraum und $S \subseteq V$ eine beliebige Teilmenge.

(i) Genau dann ist S linear unabhängig, wenn $v \notin \mathrm{lin}(S \setminus \{v\})$ für alle $v \in S$ gilt.

(ii) Ist S linear unabhängig und $v \in V \setminus \mathrm{lin}(S)$, dann ist auch $S \cup \{v\}$ linear unabhängig.

Beweis: zu (i) „⇒“ Angenommen, es gibt ein $v \in S$ mit $v \in \mathrm{lin}(S \setminus \{v\})$. Dann existiert eine endliche Teilmenge $T \subseteq S \setminus \{v\}$ und eine Familie $(\lambda_w)_{w \in T}$ mit $v = \sum_{w \in T} \lambda_w w$. Sei nun $T_1 = T \cup \{v\}$ und $(\mu_w)_{w \in T_1}$ die Familie von Elementen aus K gegeben durch $\mu_w = \lambda_w$ für alle $w \in T$ und $\mu_v = -1_K$. Wegen $\mu_v \neq 0_K$ und

$$\sum_{w \in T_1} \mu_w w \quad = \quad \sum_{w \in T} \lambda_w w + (-1_K) v \quad = \quad 0_V$$

ist S linear abhängig.

„⇐“ Setzen wir voraus, dass S linear abhängig ist. Dann gibt es eine endliche Teilmenge $T \subseteq S$ und eine Familie $(\lambda_v)_{v \in T}$ in K mit $\sum_{v \in T} \lambda_v v = 0_V$, wobei $\lambda_{v_1} \neq 0_K$ für (mindestens) ein $v_1 \in T$ gilt. Setzen wir $T_1 = T \setminus \{v_1\}$, so erhalten wir

$$\sum_{v \in T_1} \lambda_v v + \lambda_{v_1} v_1 = 0_V \quad \Leftrightarrow \quad v_1 = \sum_{v \in T_1} \left(-\frac{\lambda_v}{\lambda_{v_1}} \right) v = 0_V.$$

Wegen $T_1 \subseteq S \setminus \{v_1\}$ folgt daraus $v_1 \in \mathrm{lin}(S \setminus \{v_1\})$.

zu (ii) Nehmen wir an, dass S linear unabhängig ist, $v \notin \mathrm{lin}(S)$ gilt und $S \cup \{v\}$ dennoch linear abhängig ist. Dann gibt es eine endliche Familie $T \subseteq S \cup \{v\}$ und eine Familie $(\lambda_w)_{w \in T}$, so dass $\sum_{w \in T} \lambda_w w = 0_V$ gilt, aber λ_w nicht für alle $w \in T$ gleich Null ist. Wäre $v \notin T$ oder $\lambda_v = 0_K$, dann würde daraus die lineare Abhängigkeit von S folgen, im Widerspruch zur Voraussetzung. Also gilt $v \in T$ und $\lambda_v \neq 0_K$. Wir können somit $T_1 = T \setminus \{v\}$ setzen und

$$\lambda_v v + \sum_{w \in T_1} \lambda_w w = 0_V \quad \Leftrightarrow \quad v = \sum_{w \in T_1} \left(-\frac{\lambda_w}{\lambda_v} \right) w$$

schreiben. Wegen $T_1 \subseteq S$ folgt daraus $v \in \mathrm{lin}(S)$, im Widerspruch zur Voraussetzung. Dies zeigt, dass unsere Annahme falsch war. □

§ 7. Basis und Dimension

Definition 7.1 Sei V ein K-Vektorraum. Eine Teilmenge $B \subseteq V$ heißt ***Basis*** von V, wenn sie linear unabhängig und ein Erzeugendensystem von V ist.

Im letzten Kapitel haben wir mehrere Beispiele für Teilmengen eines Vektorraums V gesehen, die einerseits den Vektorraum V erzeugen und andererseits auch linear unabhängig sind. Beispielsweise ist durch die Menge $\{e_1, ..., e_n\}$ der Einheitsvektoren eine Basis des K^n gegeben. Die Menge $\{\mathrm{B}_{k\ell} \mid 1 \le k \le m, 1 \le \ell \le n\}$ der Basismatrizen bildet eine Basis des Vektorraums $\mathscr{M}_{m\times n,K}$ (daher der Name). Die Menge $\{x^n \mid n \in \mathbb{N}_0\}$ bildet ein (unendliche) Basis des K-Vektorraums $K[x]$.

Satz 7.2 Sei V ein K-Vektorraum. Für eine Teilmenge $B \subseteq V$ sind die folgenden Aussagen äquivalent.

(i) Sie ist eine Basis von V.
(ii) Sie ist ein minimales Erzeugendensystem von V.
(iii) Sie ist eine maximale linear unabhängige Teilmenge von V.

Beweis: „(i) ⇒ (ii)" Nehmen wir an, dass B kein minimales Erzeugendensystem von V ist. Dann gibt es eine Teilmenge $S \subsetneq B$ mit $V = \mathrm{lin}(S)$. Wählen wir $v \in B \setminus S$ beliebig, dann gilt $v \in \mathrm{lin}(S)$. Nach Prop. (6.5) (i) ist $S \cup \{v\}$ also linear abhängig. Wegen $B \supseteq S \cup \{v\}$ ist dann auch B linear abhängig, im Widerspruch zur Voraussetzung.

„(ii) ⇒ (iii)" Gehen wir zunächst davon aus, dass B linear abhängig ist. Dann gibt es nach Prop. (6.5) (i) ein $v \in B$ mit $v \in \mathrm{lin}(B\setminus\{v\})$. Aus $B\setminus\{v\} \subseteq \mathrm{lin}(B\setminus\{v\})$ und $v \in \mathrm{lin}(B\setminus\{v\})$ folgt $B \subseteq \mathrm{lin}(B\setminus\{v\})$ und somit $V = \mathrm{lin}(B) = \mathrm{lin}(B\setminus\{v\})$. Aber dies steht im Widerspruch zu der Vorausetzung, dass B ein *minimales* Erzeugendensystem von V ist.

Nehmen wir nun an, B ist zwar linear unabhängig, aber als linear unabhängige Teilmenge nicht maximal. Dann gibt es eine linear unabhängige Teilmenge S von V mit $S \supsetneq B$. Sei nun v ein beliebiges Element in $S\setminus B$. Wegen $V = \mathrm{lin}(B)$ gilt $v \in \mathrm{lin}(B)$ und wegen $B \subseteq S\setminus\{v\}$ damit erst recht $v \in \mathrm{lin}(S\setminus\{v\})$. Nach Prop. (6.5) (i) bedeutet dies, dass S linear abhängig ist. Unsere Annahme hat also erneut zu einem Widerspruch geführt.

„(iii) ⇒ (i)" Nehmen wir an, dass B kein Erzeugendensystem von V ist. Dann existiert ein $v \in V \setminus \mathrm{lin}(B)$. Nach Prop. (6.5) (ii) ist deshalb mit B auch $B \cup \{v\}$ linear unabhängig. Aber dies widerspricht der Voraussetzung, dass B als maximal als linear unabhängige Teilmenge von V ist. □

Satz 7.3 Sei V ein K-Vektorraum, S eine linear unabhängige Teilmenge und $E \supseteq S$ ein endliches Erzeugendensystem von V. Dann gibt es eine Basis B mit $S \subseteq B \subseteq E$.

Beweis: Sei $\mathscr{S}$ die Menge aller linear unabhängigen Teilmengen T von V mit $S \subseteq T \subseteq E$. Weil E endlich ist, handelt es sich auch bei $\{|T| \mid T \in \mathscr{S}\}$ um eine endliche Menge. Sei $n \in \mathbb{N}_0$ ihr Maximum und $B \in \mathscr{S}$ eine Menge mit $|B| = n$. Dann ist B als Element von $\mathscr{S}$ bezüglich Inklusion maximal. Ist $B = E$, dann ist B linear unabhängig und zugleich ein Erzeugendensystem von V, insgesamt also eine Basis.

Setzen wir nun $B \subsetneq E$ voraus, und sei $v \in E \setminus B$ beliebig vorgegeben. Wäre v nicht in $\mathrm{lin}(B)$ enthalten, dann würde nach Prop. (6.5) (ii) daraus die lineare Unabhängigkeit von $B \cup \{v\}$ folgen, im Widerspruch zur Maximalität von B. So aber gilt $v \in \mathrm{lin}(B)$ für alle $v \in E \setminus B$. Zusammen mit $B \subseteq \mathrm{lin}(B)$ folgt $E \subseteq \mathrm{lin}(B)$ und damit $V = \mathrm{lin}(E) = \mathrm{lin}(B)$. Wieder ist B linear unabhängig und ein Erzeugendensystem von V, also eine Basis. □

Wir bezeichnen einen K-Vektorraum V als ***endlich erzeugt***, wenn V ein endliches Erzeugendensystem besitzt, also eine endliche Teilmenge $E \subseteq V$ mit $V = \mathrm{lin}(E)$ existiert. Aus dem soeben bewiesenen Satz folgt unmittelbar

Folgerung 7.4 Sei V ein endlich erzeugter K-Vektorraum.

(i) *(Basisauswahlsatz)*
Aus jedem endlichen Erzeugendensystem von V kann eine Basis gewählt werden.

(ii) *(Basisergänzungssatz)*
Jede endliche linear unabhängige Teilmenge von V kann zu einer Basis von V ergänzt werden.

Insbesondere besitzt jeder endlich erzeugte K-Vektorraum eine Basis.

Beweis: zu (i) Sei $E \subseteq V$ ein endliches Erzeugendensystem von V. Wenden wir Satz (7.3) auf dieses E und $S = \emptyset$ an, so erhalten wir eine Basis $B \subseteq E$.

zu (ii) Sei S eine endliche linear unabhängige Teilmenge von V und $\tilde{E} \subseteq V$ ein endliches Erzeugendensystem von V. Dann ist auch $E = S \cup \tilde{E}$ ein endliches Erzeugendensystem von V. Durch Anwendung von Satz (7.3) auf S und E erhalten wir eine Basis B von V mit $B \supseteq S$. □

Lemma 7.5 *(Austauschlemma)*
Sei V ein K-Vektorraum, E ein Erzeugendensystem und S eine linear unabhängige Teilmenge von V. Dann gibt es für jedes $v \in S \setminus E$ ein $w \in E \setminus S$, so dass die Menge $(S \setminus \{v\}) \cup \{w\}$ linear unabhängig ist.

Beweis: Sei $v \in S \setminus E$ vorgegeben. Dann ist $S_1 = S \setminus \{v\}$ als linear unabhängige Menge nicht maximal, nach Satz (7.2) also auch kein Erzeugendensystem von V. Nehmen wir nun an, für alle $w \in E$ gilt $w \in \text{lin}(S_1)$ oder $w \in S$. Wegen $v \notin E$ ist $w \in S$ gleichbedeutend mit $w \in S_1$. Dann wäre insgesamt $E \subseteq \text{lin}(S_1)$ erfüllt. Wegen $\text{lin}(E) = V$ wäre damit auch S_1 ein Erzeugendensystem von V. Aber dies widerspricht unserer anfangs getroffenen Feststellung.

Es existiert also ein $w \in E \setminus S$ mit $w \notin \text{lin}(S_1)$. Wenden wir nun Prop. (6.5) (ii) an, so folgt daraus zusammen mit der linearen Unabhängigkeit von S_1 auch die lineare Unabhängigkeit der Menge $S_1 \cup \{w\}$. □

Satz 7.6 *(Austauschsatz)*
Sei V ein K-Vektorraum, E ein Erzeugendensystem und S eine linear unabhängige Teilmenge von V. Dann gibt es für jede endliche Teilmenge $T \subseteq S \setminus E$ eine Teilmenge $F \subseteq E \setminus S$ mit $|F| = |T|$, so dass $(S \setminus T) \cup F$ linear unabhängig ist.

Beweis: Wir führen den Beweis durch vollständige Induktion über $n = |T|$. Im Fall $n = 0$ ist $T = \varnothing$, und man kann offenbar $F = \varnothing$ wählen. Sei nun $T \subseteq S \setminus E$ und $n = |T| > 0$, und setzen wir die Aussage für Teilmengen von $S \setminus E$ mit weniger als n Elemente voraus. Sei $v_1 \in T$ ein beliebiges Element und $T_1 = T \setminus \{v_1\}$. Nach Induktionsvoraussetzung gibt es eine Teilmenge $F_1 \subseteq E \setminus S$ mit $|F_1| = |T_1|$, so dass $S_1 = (S \setminus T_1) \cup F_1$ linear unabhängig ist. Wenden wir Lemma (7.5) auf die linear unabhängige Menge S_1, das Erzeugendensystem E und das Element $v_1 \in S_1$ ein, so erhalten wir ein $w_1 \in E \setminus S_1$ mit der Eigenschaft, dass $(S_1 \setminus \{v_1\}) \cup \{w_1\}$ linear unabhängig ist. Wir definieren $F = F_1 \cup \{w_1\}$. Zu überprüfen sind die Eigenschaften

(i) $F \subseteq E \setminus S$
(ii) $|F| = |T|$
(iii) Die Menge $(S \setminus T) \cup F$ ist linear unabhängig.

zu (i) Es gilt $F_1 \subseteq E \setminus S$, zu zeigen bleibt $w_1 \in E \setminus S$. Jedenfalls ist w_1 nach Definition in E enthalten. Wäre $w_1 \in S$, dann würde wegen $w_1 \notin S_1$ und $S \setminus T_1 \subseteq S_1$ daraus $w_1 \in T_1$ folgen. Aber dies steht wegen $T_1 \subseteq T \subseteq S \setminus E$ im Widerspruch zu $w_1 \in E$.

zu (ii) Wäre $w_1 \in F_1$, dann würde sich wegen $F_1 \subseteq S_1$ ein Widerspruch zu $w_1 \notin S_1$ ergeben. So aber gilt $|F| = |F_1| + 1 = |T_1| + 1 = |T|$.

zu (iii) Wegen $v_1 \in S \setminus E$ und $F_1 \subseteq S_1$ gilt $v_1 \notin F_1$. Deshalb ist es gleichgültig, ob aus der Menge $S \setminus T_1$ zuerst v_1 entfernt und anschließend F_1 hinzugenommen wird oder umgekehrt. Es gilt also

$$(S_1 \setminus \{v_1\}) \cup \{w_1\} \quad = \quad (((S \setminus T_1) \cup F_1) \setminus \{v_1\}) \cup \{w_1\} \quad = \quad ((S \setminus T_1) \setminus \{v_1\}) \cup F_1 \cup \{w_1\} \quad = \quad (S \setminus T) \cup F \quad ,$$

und die Menge zu Beginn der Gleichungskette ist nach Konstruktion linear unabhängig. □

Satz 7.7 Sei V ein endlich erzeugter K-Vektorraum. Dann besitzt V eine endliche Basis, und je zwei Basen sind gleichmächtig.

Beweis: Nach Voraussetzung existiert in V ein endliches Erzeugendensystem, und auf Grund des Basisauswahlsatzes enthält dieses eine Basis als Teilmenge. Insbesondere existiert also eine endliche Basis B von V. Sei nun B' eine beliebige weitere Basis. Nehmen wir an, dass die Menge B' unendlich oder zumindest $|B'| > |B|$ ist. Sei $S \subseteq B'$ eine beliebige endliche Teilmenge mit $|S| > |B|$. Nach dem Austauschsatz können wir sämtliche nicht in B liegenden Elemente von S durch Elemente aus B ersetzen, ohne dass sich an der linearen Unabhängigkeit oder an der Mächtigkeit von S etwas ändert.

Genauer gesagt können wir Satz (7.6) auf S, $E = B$ und $T = S \setminus E$ anwenden. Wir erhalten eine endliche Teilmenge $F \subseteq E \setminus S$ mit $|F| = |T|$ und der Eigenschaft, dass $B_1 = (S \setminus T) \cup F$ linear unabhängig ist. Es gilt $|B_1| = |S| - |T| + |F| = |S| > |B|$, andererseits aber

$$B_1 \quad = \quad (S \setminus T) \cup F \quad = \quad (S \setminus (S \setminus E)) \cup F \quad = \quad (S \cap E) \cup F \quad \subseteq \quad E \quad = \quad B$$

und somit $|B_1| \leq |B|$. Dieser Widerspruch zeigt, dass die Annahme falsch war, es muss also $|B'| \leq |B|$ gelten. Durch Vertauschung der Rollen von B und B' zeigt man genauso $|B| \leq |B'|$. □

Definition 7.8 Sei V ein K-Vektorraum. Dann ist die ***Dimension*** von V definiert durch

$$\dim V \quad = \quad \begin{cases} n & \text{falls } V \text{ endlich erzeugt ist und eine } n\text{-elementige Basis existiert} \\ \infty & \text{falls } V \text{ nicht endlich erzeugt ist} \end{cases}$$

Folgerung 7.9 Sei V ein n-dimensionaler K-Vektorraum. Dann ist jede n-elementige, linear unabhängige Teilmenge in V und jedes n-elementige Erzeugendensystem des Vektorraums V eine Basis von V.

Beweis: Nach Definition der Dimension existiert eine Basis B von V mit $|B| = n$. Sei nun $S \subseteq V$ eine linear unabhängige Teilmenge mit n Elementen. Dann ist S als linear unabhängige Menge maximal. Denn andernfalls könnten wir S zu einer Basis mit mehr als n Elementen ergänzen. Aber dies widerspricht der Tatsache, dass je zwei Basen von V gleichmächtig sind.

Sei nun S ein Erzeugendensystem von V mit $|S| = n$. Dann ist S als Erzeugendensystem minimal. Denn ansonsten könnten wir in S eine Basis $T \subsetneq S$ auswählen. Es wäre dann $|T| < |B| = n$, was erneut im Widerspruch dazu steht, dass je zwei Basen von V gleichmächtig sind. □

Ebenso zeigt man

Folgerung 7.10 Sei V ein endlich-dimensionaler K-Vektorraum, $n = \dim V$. Dann gilt

(i) Ist $S \subseteq V$ linear unabhängig, dann gilt $|S| \le n$. Insbesondere ist S endlich.

(ii) Ist U ein Untervektorraum von V, dann gilt $\dim U \le \dim V$.

(iii) Ist $E \subseteq V$ ein Erzeugendensystem von V, dann gilt $|E| \ge n$.

Beweis: zu (i) Nehmen wir an, dass S eine linear unabhängige Menge mit mehr als n Elementen ist. Dann könnten wir S zu einer Basis B mit $|B| > n$ ergänzen. Aber wegen $n = \dim V$ muss jede Basis von V jeweils n-elementig sein.

zu (ii) Sei S eine Basis von U. Dann ist S insbesondere eine linear unabhängige Teilmenge von V. Nach Teil (i) folgt $\dim U = |S| \le n$.

zu (iii) Nehmen wir an, dass E ein Erzeugendensystem von V mit $|E| < n$ ist. Dann können wir aus E eine Basis B von V mit $|B| < n$ Elementen auswählen. Aber wegen $n = \dim V$ muss jede Basis von V aus n Elementen bestehen. □

Wir geben einige konkrete Beispiele für die Dimension von Vektorräumen an.

(i) Ist K ein Körper und $n \in \mathbb{N}$, dann gilt $\dim K^n = n$. Denn wie wir bereits festgestellt haben, bilden die Einheitsvektoren $e_1, \ldots, e_n$ in K^n eine n-elementige Basis.

(ii) Seien K ein Körper und $m, n \in \mathbb{N}$. Dann gilt $\dim \mathcal{M}(m \times n, K) = mn$, weil die Basismatrizen $\mathrm{B}_{k\ell}$ mit $1 \le k \le m$, $1 \le \ell \le n$, eine mn-elementige Basis von $\mathcal{M}(m \times n, K)$ bilden.

(iii) Wir wissen bereits, dass $\mathbb{C}$ auf natürliche Weise als $\mathbb{R}$-Vektorraum angesehen werden kann. Eine Basis dieses Vektorraums ist durch $\{1, i\}$ gegeben. Denn einerseits kann jedes $z \in \mathbb{C}$ auf Grund der Zerlegung in Real- und Imaginärteil in der Form $z = a \cdot 1 + b \cdot i$ mit $a, b \in \mathbb{R}$ dargestellt werden. Dies zeigt, dass $\{1, i\}$ ein Erzeugendensystem ist. Andererseits ist diese Darstellung auch eindeutig, denn aus $z = a \cdot 1 + b \cdot i$ mit $a, b \in \mathbb{R}$ folgt $a = \mathrm{Re}(z)$ und $b = \mathrm{Im}(z)$. Deshalb ist $\{1, i\}$ auch linear unabhängig. Für $\mathbb{C}$ als $\mathbb{R}$-Vektorraum gilt also $\dim \mathbb{C} = 2$; der Deutlichkeit halber schreibt man $\dim_{\mathbb{R}} \mathbb{C} = 2$.

(iv) Fassen wir $\mathbb{C}$ dagegen als $\mathbb{C}$-Vektorraum auf, dann ist $\{1\}$ eine Basis, und es gilt $\dim \mathbb{C} = 1$.

(v) Ist K ein Körper und V ein K-Vektorraum mit $V = \{0_V\}$, dann gilt $\dim V = 0$, denn in diesem Fall ist die leere Menge $\emptyset$ eine Basis.

(vi) Für jeden Körper K gilt $\dim K[x] = \infty$. Um dies zu sehen, überprüft man zunächst, dass für jedes $n \in \mathbb{N}$ die $(n+1)$-elementige Teilmenge $\{1, x, x^2, \ldots, x^n\}$ von $K[x]$ linear unabhängig ist.

Gäbe es nun in $K[x]$ ein endliches Erzeugendensystem E, dann könnte man darin eine Basis mit $m \leq |E|$ Elemente wählen. Damit wäre aber $\dim K[x] = m$ und somit jede linear unabhängige Teilmenge von $K[x]$ höchstens m-elementig. Dies widerspricht aber unserer zuvor getroffenen Feststellung, dass es in $K[x]$ endliche linear unabhängige Mengen beliebiger Größe gibt.

§ 8. Dimensionssätze

Satz 8.1 *(Schnittdimensionssatz)*

Sei V ein endlich erzeugter K-Vektorraum, und seien W, W' Untervektorräume von V. Dann gilt

$$\dim(W + W') \quad = \quad \dim(W) + \dim(W') - \dim(W \cap W').$$

Beweis: Sei $n = \dim(W \cap W')$ und $\{v_1, ..., v_n\}$ eine Basis von $W \cap W'$. Weil $W \cap W'$ ein Untervektorraum sowohl von W als auch von W' ist, gilt $\dim(W \cap W') \le \dim W$ und $\dim(W \cap W') \le \dim W'$ nach Folgerung (7.10) (ii). Es gibt also $k, \ell \in \mathbb{N}_0$ mit $\dim W = n + k$ und $\dim W' = n + \ell$.

Weil $\{v_1, ..., v_n\}$ eine linear unabhängige Menge in W ist, finden wir nach dem Basisergänzungssatz Vektoren $w_1, ..., w_k$, so dass $B = \{v_1, ..., v_n, w_1, ..., w_k\}$ eine Basis von W ist. Ebenso finden wir Elemente $w'_1, ..., w'_\ell$ mit der Eigenschaft, dass die Familie $B' = \{v_1, ..., v_n, w'_1, ..., w'_\ell\}$ eine Basis von W' ist. Der Satz ist bewiesen, wenn wir zeigen können, dass es sich bei

$$B_0 \quad = \quad B \cup B' \quad = \quad \{v_1, ..., v_n, w_1, ..., w_k, w'_1, ..., w'_\ell\}$$

um eine $n + k + \ell$-elementige Basis von $W + W'$ handelt, denn dann gilt

$$\dim(W + W') \quad = \quad n + k + \ell \quad = \quad (n + k) + (n + \ell) - n \quad = \quad \dim(W) + \dim(W') - \dim(W \cap W').$$

Zunächst zeigen wir, dass B_0 ein Erzeugendensystem von $W + W'$ ist. Jedes $v \in W + W'$ lässt sich in der Form $v = w + w'$ mit $w \in W$ und $w' \in W'$ schreiben. Da $\{v_1, ..., v_n, w_1, ..., w_k\}$ eine Basis von W ist, finden wir Koeffizienten $\mu_i, \lambda_i \in K$ mit $w = \sum_{i=1}^n \mu_i v_i + \sum_{i=1}^k \lambda_i w_i$. Ebenso gibt es $\mu'_i, \lambda'_i \in K$ mit $w' = \sum_{i=1}^n \mu'_i v_i + \sum_{i=1}^\ell \lambda'_i w'_i$. Insgesamt erhalten wir

$$v \quad = \quad w + w' \quad = \quad \sum_{i=1}^n (\mu_i + \mu'_i) v_i + \sum_{i=1}^k \lambda_i w_i + \sum_{i=1}^\ell \lambda'_i w'_i \quad ,$$

also kann jedes $v \in W + W'$ tatsächlich als Linearkombination von B_0 dargestellt werden.

Als nächstes überprüfen wir, dass B_0 tatsächlich aus $n + k + \ell$ verschiedenen Elementen besteht. Besteht die Menge aus weniger Elementen, dann muss $w_i = w'_j$ für gewissen i, j mit $1 \le i \le k$ und $1 \le j \le \ell$ gelten. Dies würde bedeuteten, dass w_i in $W \cap W'$ enthalten ist. Damit wäre w_i also in $\mathrm{lin}(v_1, ..., v_n)$ enthalten und die Menge B damit linear abhängig, im Widerspruch zur Basis-Eigenschaft von B. Also ist $|B_0| = n + k + \ell$ erfüllt.

Nun beweisen wir die lineare Unabhängigkeit. Seien $\mu_i, \lambda_i, \lambda'_i \in K$ Koeffizienten mit

$$\sum_{i=1}^n \mu_i v_i + \sum_{i=1}^k \lambda_i w_i + \sum_{i=1}^\ell \lambda'_i w'_i \quad = \quad 0.$$

Sei $v = \sum_{i=1}^n \mu_i v_i + \sum_{i=1}^k \lambda_i w_i \in W$. Wegen $v = -\sum_{i=1}^\ell \lambda'_i w'_i$ liegt v in $W \cap W'$. Da der Unterraum $W \cap W'$ bereits von $\{v_1, ..., v_n\}$ erzeugt wird, und weil auf Grund der Basiseigenschaft von B die Darstellung von v als Linearkombination von B eindeutig ist, gilt $\lambda_1 = ... = \lambda_k = 0$. Wir erhalten somit $\sum_{i=1}^n \mu_i v_i + \sum_{i=1}^\ell \lambda'_i w'_i = 0$. Wegen der linearen Unabhängigkeit von B' folgt daraus wiederum $\lambda'_i = 0$ für $1 \le i \le \ell$ und $\mu_i = 0$ für $1 \le i \le n$. □

Folgerung 8.2 Sei V ein endlich-dimensionaler K-Vektorraum, und seien W, W' Untervektorräume von V, so dass $V = W \oplus W'$ erfüllt ist. Sei B eine Basis von W und B' eine Basis von W'. Dann gilt

(i) $\dim V = \dim W + \dim W'$

(ii) Die Mengen B und B' sind disjunkt.

(iii) Die Vereinigung $B \cup B'$ ist eine Basis von V.

Beweis: Sei $m = \dim W$ und $m' = \dim W'$. Nach Voraussetzung gilt $W \cap W' = \{0_V\}$, also $\dim(W \cap W') = 0$. Aus dem Schnittdimensionssatz folgt

$$\dim V \quad = \quad \dim W + \dim W' - \dim(W \cap W') \quad = \quad m + m' - 0 \quad = \quad m + m'.$$

Sei $B = \{w_1, ..., w_m\}$ eine Basis von W und $B' = \{w'_1, ..., w'_{m'}\}$ eine Basis von W'. Wir zeigen, dass $E = B \cup B'$ ein Erzeugendensystem von V ist. Sei $v \in V$ vorgegeben. Wegen $V = W + W'$ gibt es $w \in W$ und $w' \in W'$ mit $v = w + w'$. Weil B eine Basis von W und B' eine Basis von W' ist, gibt es $\lambda_1, ..., \lambda_m \in K$ und $\mu_1, ..., \mu_{m'} \in K$ mit

$$w = \sum_{k=1}^m \lambda_k w_k \qquad \text{und} \qquad w' = \sum_{k=1}^{m'} \mu_k w'_k.$$

Es folgt

$$v \quad = \quad w + w' \quad = \quad \sum_{k=1}^m \lambda_k w_k + \sum_{k=1}^{m'} \mu_k w'_k.$$

Dies zeigt, dass E tatsächlich ein Erzeugendensystem von V ist. Wegen $\dim V = m + m'$ besteht jedes Erzeugendensystem von V aus mindestens $m + m'$ Elementen. Die Mengen B und B' sind also disjunkt, da ansonsten $|E| < m + m'$ gelten würde. Als $(m + m')$-elementiges Erzeugendensystem ist E wegen $\dim V = m + m'$ eine Basis von V. □

Durch vollständige Induktion über r erhält man

Folgerung 8.3 Sei V ein K-Vektorraum, und seien $W_1, ..., W_r$ Untervektorräume von V mit $V = \bigoplus_{k=1}^r W_k$. Dann gilt $\dim V = \sum_{k=1}^r \dim W_k$. Ist B_k eine Basis von W_k für $1 \le k \le r$, dann ist $B = \bigcup_{k=1}^r B_k$ eine Basis von V.

Als nächstes untersuchen wir die Vektorraum-Dimension im Zusammenhang mit linearen Abbildungen.

Satz 8.4 Seien V, W endlich-dimensionale Vektorräume über einem Körper K, und sei $\phi : V \to W$ eine lineare Abbildung. Dann gilt

$$\dim V \quad = \quad \dim \ker(\phi) + \dim \operatorname{im}(\phi).$$

Beweis: Sei $\{u_1, ..., u_m\}$ eine Basis von $\ker(\phi)$ und $\{w_1, ..., w_n\}$ eine Basis von $\operatorname{im}(\phi)$. Wir wählen für jedes w_i einen Vektor $v_i \in V$ mit $\phi(v_i) = w_i$ und zeigen, dass durch

$$B \quad = \quad \{u_1, ..., u_m, v_1, ..., v_n\}$$

eine $(m+n)$-elementige Basis von V gegeben ist. Haben wir dies gezeigt, dann ist damit $\dim V = m + n = \dim \ker(\phi) + \dim \operatorname{im}(\phi)$ bewiesen. Dass B aus weniger als $m + n$ Elementen besteht ist nur möglich, wenn $u_i = v_j$ für gewisse i, j mit $1 \le i \le m$ und $1 \le j \le n$ gilt. Aber dann wäre $w_j = \phi(v_j) = \phi(u_i) = 0_W$ im Widerspruch dazu, dass w_j in einer Basis von W liegt und somit ungleich Null sein muss.

Zunächst weisen wir nun nach, dass es sich bei B um ein Erzeugendensystem von V handelt. Sei dazu $v \in V$ vorgegeben. Da $\{w_1, ..., w_n\}$ eine Basis von $\operatorname{im}(\phi)$ ist, finden wir $\lambda_1, ..., \lambda_n \in K$ mit

$$\phi(v) \quad = \quad \sum_{i=1}^{n} \lambda_i w_i.$$

Aus der Linearität der Abbildung ϕ folgt $\phi(v) = \sum_{i=1}^{n} \lambda_i \phi(v_i) = \phi(v')$ mit $v' = \sum_{i=1}^{n} \lambda_i v_i$. Wegen $\phi(v) - \phi(v') = \phi(v - v') = 0_W$ liegt dann der Vektor $v - v'$ in $\ker(\phi)$. Da $\{u_1, ..., u_m\}$ eine Basis dieses Untervektorraums ist, existieren $\mu_1, ..., \mu_m \in K$ mit

$$v - v' = \sum_{j=1}^{m} \mu_j u_j \qquad \Leftrightarrow \qquad v = \sum_{j=1}^{m} \mu_j u_j + v' = \sum_{j=1}^{m} \mu_j u_j + \sum_{i=1}^{n} \lambda_i v_i.$$

Damit haben wir gezeigt, dass B ein Erzeugendensystem von V ist. Nun beweisen wir die lineare Unabhängigkeit. Seien $\mu_i, \lambda_j \in K$ mit

$$\sum_{i=1}^{m} \mu_i u_i + \sum_{j=1}^{n} \lambda_j v_j \quad = \quad 0_V$$

vorgegeben. Wenden wir die lineare Abbildung ϕ auf beide Seiten der Gleichung an, dann folgt

$$0_W \quad = \quad \phi(0_V) \quad = \quad \phi\left(\sum_{i=1}^{m} \mu_i u_i + \sum_{j=1}^{n} \lambda_j v_j\right) \quad = \quad 0_W + \sum_{j=1}^{n} \lambda_j \phi(v_j) \quad = \quad \sum_{j=1}^{n} \lambda_j w_j.$$

Dabei haben wir verwendet, dass die Summe $\sum_{i=1}^{m} \mu_i u_i$ in $\ker(\phi)$ enthalten ist. Weil die Menge $\{w_1, ..., w_n\}$ linear unabhängig ist, bedeutet dies $\lambda_1 = ... = \lambda_n = 0$. Setzen wir dies in die Ausgangsgleichung ein, dann erhält man $\sum_{i=1}^{m} \mu_i u_i = 0_V$. Da $\{u_1, ..., u_m\}$ nach Voraussetzung einer Basis von $\ker(\phi)$ und insbesondere linear unabhängig ist, hat dies wiederum $\mu_1 = ... = \mu_m = 0$ zur Folge. Also B tatsächlich linear unabhängig. □

Folgerung 8.5 Für isomorphe Vektorräume V, W gilt $\dim V = \dim W$.

Beweis: Sei $\phi : V \to W$ ein Isomorphismus. Dann ist $\ker(\phi) = \{0_V\}$ und $\mathrm{im}(\phi) = W$. Also folgt die Aussage aus dem Dimensionssatz. □

Wir werden den Dimensionssatz für lineare Abbildungen nun verwenden, um Genaueres über die Struktur der Matrizen herauszufinden.

Proposition 8.6 Sei $A = (a_{ij})$ eine $(m \times n)$-Matrix über K und $\phi_A : K^n \to K^m$ die lineare Abbildung gegeben durch $v \mapsto Av$. Dann gilt

(i) Für $1 \le k \le n$ gilt $\phi_A(e_k) = a_{\bullet k}$. Die Bilder der Einheitsvektoren sind also genau die Spalten der Matrix.

(ii) Es gilt $\mathrm{im}(\phi_A) = \mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\}$.

Beweis: zu (i) Sei $k \in \{1, ..., n\}$ vorgegeben. Nach Definition des Matrix-Vektor-Produkts erhält man für jedes $i \in \{1, ..., m\}$ den i-ten Eintrag von $\phi_A(e_k)$ durch die Formel

$$\sum_{j=1}^{n} a_{ij}\delta_{jk} \quad = \quad a_{ik}\delta_{kk} \quad = \quad a_{ik}.$$

Dies ist genau der i-te Eintrag des k-ten Spaltenvektors $a_{\bullet k}$ der Matrix.

zu (ii) Sei $v \in K^n$ beliebig vorgegeben, $v = (\lambda_1, ..., \lambda_n)$. Da ϕ_A eine lineare Abbildung ist, gilt

$$\phi_A\left(\sum_{k=1}^{n} \lambda_k e_k\right) \quad = \quad \sum_{k=1}^{n} \lambda_k \phi_A(e_k) \quad = \quad \sum_{k=1}^{n} \lambda_k a_{\bullet k}.$$

Damit ist $\mathrm{im}(\phi_A) \subseteq \mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\}$ nachgewiesen. Andererseits ist $\mathrm{im}(\phi_A)$ ein Untervektorraum von K^m, der nach Teil (i) die Menge $\{a_{\bullet 1}, ..., a_{\bullet n}\}$ der Spaltenvektoren als Teilmenge enthält. Nach Satz (6.3) (ii) folgt daraus $\mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\} \subseteq \mathrm{im}(\phi_A)$. □

Definition 8.7 Sei A eine $(m \times n)$-Matrix über K.

(i) Der Untervektorraum $\mathrm{im}(\phi_A) = \mathrm{lin}\{a_{\bullet 1}, ..., a_{\bullet n}\}$ von K^m wird der ***Spaltenraum*** der Matrix A genannt und von uns mit $\mathrm{SR}(A)$ bezeichnet. Die Dimension $\mathrm{sr}(A) = \dim \mathrm{SR}(A)$ nennt man den ***Spaltenrang*** von A.

(ii) Ebenso nennt man den Untervektorraum von K^n gegeben durch $\mathrm{lin}\{a_{1\bullet}, ..., a_{m\bullet}\}$ den ***Zeilenraum*** von A und bezeichnet ihn mit $\mathrm{ZR}(A)$. Die Dimension $\mathrm{zr}(A) = \dim \mathrm{ZR}(A)$ wird ***Zeilenrang*** von A genannt.

Den Kern der linearen Abbildung $\phi_A : K^n \to K^m$, $v \mapsto Av$ nennt man auch den ***Kern der Matrix*** A und bezeichnet ihn mit $\ker(A)$. Direkt aus dem Dimensionssatz für lineare Abbildungen folgt

Folgerung 8.8 Für alle $A \in \mathcal{M}_{m\times n,K}$ gilt $\dim\ker(A) + \mathrm{sr}(A) = n$.

Folgerung 8.9 Sei $A \in \mathcal{M}_{m\times n,K}$ und $\mathcal{L} \subseteq K^n$ der Lösungsraum des homogenen Gleichungssystems mit A als Koeffizientenmatrix. Sei r der Spaltenrang von A. Dann gilt $\dim\mathcal{L} = n - r$.

Beweis: Dies folgt direkt aus Folgerung (8.8), denn die Lösungsmenge $\mathcal{L}$ des linearen Gleichungssystem stimmt mit dem Kern der Matrix A überein: Es gilt $\mathcal{L} = \{x \in K^n \mid Ax = 0_{K^m}\}$. □

Proposition 8.10 Der in Def. (2.1) definierte Zeilenrang einer Matrix in normierter Zeilenstufenform stimmt mit dem hier soeben definierten Zeilenrang einer Matrix überein.

Beweis: Sei $A \in \mathcal{M}_{m\times n,K}$ eine Matrix in normierter Zeilenstufenform, und seien r und $j_1,\ldots,j_r$ die zugehörigen Kennzahlen. Wir zeigen, dass die Zeilenvektoren $a_{1\bullet},\ldots,a_{r\bullet} \in K^n$ ungleich Null linear unabhängig sind und somit eine Basis des Zeilenraums $\mathrm{ZR}(A)$ bilden. Seien $\lambda_1,\ldots,\lambda_r \in K$ vorgegeben, mit

$$\sum_{i=1}^{r} \lambda_i a_{i\bullet} \quad = \quad 0_{K^n} \tag{1}$$

Nach Definition der normierten ZSF ist in der j_k-ten Spalte der Eintrag $a_{kj_k} = 1_K$ der einzige Eintrag ungleich Null. Insgesamt sind die Einträge der j_k-ten Spalte also gegeben durch $a_{ij_k} = \delta_{ik}$ für $1 \le i \le m$. Betrachtet man in der Gleichung (1) also jeweils die j_k-te Komponenten für $k = 1,\ldots,r$, so erhält man die Gleichungen

$$\sum_{i=1}^{r} \lambda_i a_{ij_k} = 0_K \quad \Leftrightarrow \quad \sum_{i=1}^{r} \lambda_i \delta_{ik} = 0_K \quad \Leftrightarrow \quad \lambda_k = 0_K.$$

Damit ist die lineare Unabhängigkeit nachgewiesen. □

Proposition 8.11 Sei $A \in \mathcal{M}_{m\times n,K}$ und $i \in \{1,\ldots,m\}$ eine Zeilennummer mit der Eigenschaft, dass die i-te Zeile von A eine Linearkombination der übrigen $m-1$ Zeilen ist. Entsteht nun die Matrix $\bar{A} \in \mathcal{M}((m-1)\times n, K)$ aus A durch Streichung der i-ten Zeile, dann gilt $\mathrm{ZR}(\bar{A}) = \mathrm{ZR}(A)$ (also insbesondere $\mathrm{zr}(\bar{A}) = \mathrm{zr}(A)$) und ebenso $\mathrm{sr}(\bar{A}) = \mathrm{sr}(A)$.

Beweis: Nach Definition der Untervektorräume $\mathrm{ZR}(A)$ und $\mathrm{ZR}(\bar{A})$ gilt

$$\mathrm{ZR}(A) = \mathrm{lin}\{a_{1\bullet},\ldots,a_{m\bullet}\} \qquad \text{und} \qquad \mathrm{ZR}(\bar{A}) = \mathrm{lin}\{a_{1\bullet},\ldots,a_{i-1\bullet},a_{i+1\bullet},\ldots,a_{m\bullet}\}.$$

Nach Voraussetzung enthält $\mathrm{ZR}(\bar{A})$ neben den Vektoren $a_{k\bullet}$ mit $k \neq i$ auch den i-ten Zeilenvektor $a_{i\bullet}$. Aus $\{a_{1\bullet}, ..., a_{m\bullet}\} \subseteq \mathrm{ZR}(\bar{A})$ und der Untervektorraum-Eigenschaft von $\mathrm{ZR}(\bar{A})$ folgt $\mathrm{ZR}(A) \subseteq \mathrm{ZR}(\bar{A})$. Die umgekehrte Inklusion $\mathrm{ZR}(\bar{A}) \subseteq \mathrm{ZR}(A)$ ist offensichtlich.

Wir betrachten nun die Abbildung $\pi : K^m \to K^{m-1}$, die aus jedem Vektor $c \in K^m$ die i-te Komponente entfernt, also $\pi(c) = (c_1, ..., c_{i-1}, c_{i+1}, ..., c_m)$ für $c = (c_1, ..., c_m) \in K^m$. Man überprüft unmittelbar, dass π eine lineare Abbildung ist. Die Spalten der Matrix A werden von π auf die Spalten von $\bar{A}$ abgebildet. Durch Übergang zur eingeschränkten Abbildung $\phi = \pi|_{SR(A)}$ erhalten wir also eine surjektive lineare Abbildung $\phi : \mathrm{SR}(A) \to \mathrm{SR}(\bar{A})$.

Nun zeigen wir, dass $\ker(\phi) = \{0_{K^m}\}$ gilt. Weil die i-te Zeile von A eine Linearkombination der übrigen Zeilen ist, gibt es Koeffizienten $\mu_j \in K$ mit

$$a_{i\bullet} \quad = \quad \sum_{k=1}^{i-1} \mu_k a_{k\bullet} + \sum_{k=i+1}^{m} \mu_k a_{k\bullet}.$$

Die Einträge der Matrix erfüllen also die Gleichungen $a_{ij} = \sum_{k=1}^{i-1} \mu_k a_{kj} + \sum_{k=i+1}^{m} \mu_k a_{kj}$ für $1 \leq j \leq n$. Die Spalten $w_1, ..., w_n$ von A sind damit im Untervektorraum

$$W \quad = \quad \left\{ c \in K^m \;\middle|\; c_i = \sum_{k=1}^{i-1} \mu_k c_k + \sum_{k=i+1}^{m} \mu_k c_k \right\}$$

von K^m enthalten, es gilt also $\mathrm{SR}(A) \subseteq W$. Sei nun $c \in \ker(\phi)$ vorgegeben. Es gilt $(c_1, ..., c_{i-1}, c_{i+1}, ..., c_m) = \phi(c) = (0, ..., 0)$, also $c_j = 0$ für $j \neq i$. Wegen $c \in W$ ist damit auch

$$c_i \quad = \quad \sum_{k=1}^{i-1} \mu_k c_k + \sum_{k=i+1}^{m} \mu_k c_k \quad = \quad \sum_{k=1}^{i-1} \mu_k \cdot 0 + \sum_{k=i+1}^{m} \mu_k \cdot 0 \quad = \quad 0 \quad , \quad \text{also} \quad c = 0_{K^m}.$$

Damit ist $\ker(\phi) = \{0_{K^m}\}$ bewiesen. Durch Anwendung des Dimensionssatzes für lineare Abbildungen auf ϕ erhalten wir $\mathrm{sr}(A) = \dim \mathrm{SR}(A) = \dim \ker(\phi) + \dim \mathrm{im}(\phi) = 0 + \dim SR(\bar{A}) = \mathrm{sr}(\bar{A})$. □

Satz 8.12 *(Rangsatz)*

Für jede Matrix $A \in \mathcal{M}_{m\times n,K}$ gilt $\mathrm{zr}(A) = \mathrm{sr}(A)$. Wir bezeichnen die Zahl $\mathrm{zr}(A)$ deshalb einfach als den ***Rang*** $\mathrm{rg}(A)$ der Matrix.

Beweis: Sei $r = \mathrm{zr}(A)$. Nach dem Basisauswahlsatz können wir so lange Zeilen aus A streichen, bis die verbleibenden r Zeilen der Restmatrix $A' \in \mathcal{M}(r \times n, K)$ eine Basis von $\mathrm{ZR}(A)$ bilden. Durch wiederholte Anwendung von Prop. (8.11) erhalten wir $\mathrm{zr}(A) = \mathrm{zr}(A') = r$ und $\mathrm{sr}(A) = \mathrm{sr}(A')$. Wegen $\mathrm{SR}(A') \subseteq K^r$ und $\dim K^r = r$ gibt es in $\mathrm{SR}(A')$ keine linear unabhängige Teilmenge mit mehr als r Elementen; es gilt also $\mathrm{sr}(A) = \mathrm{sr}(A') \leq r = \mathrm{zr}(A)$. Anwendung derselben Abschätzung auf die transponierte Matrix ${}^{\mathrm{t}}A$ liefert $\mathrm{zr}(A) = \mathrm{sr}({}^{\mathrm{t}}A) \leq \mathrm{zr}({}^{\mathrm{t}}A) = \mathrm{sr}(A)$, denn die Zeilen von A sind die Spalten von ${}^{\mathrm{t}}A$ und umgekehrt. Insgesamt gilt also $\mathrm{zr}(A) = \mathrm{sr}(A)$. □

Sei $A \in \mathcal{M}_{m\times n,K}$ eine Matrix in normierter ZSF und $\mathscr{L}$ die Lösungsmenge des zugehörigen homogenen linearen Gleichungssystems. Sei r der Zeilenrang von A. In Satz (2.4) haben wir eine $n-r$ elementige Menge b_ℓ von Vektoren im K^n angegeben, die ein Erzeugendensystem von $\mathscr{L}$ bilden. Weil Zeilen- und Spaltenrang einer Matrix übereinstimmen, gilt $\dim \mathscr{L} = n - r$ nach Folgerung (8.9). Also bilden die Vektoren b_ℓ sogar eine Basis von $\mathscr{L}$.

Als letzte Anwendung des Dimensionssatzes zeigen wir noch

Satz 8.13 Sei $n \in \mathbb{N}$, und sei $\phi : V \to W$ eine lineare Abbildung zwischen Vektorräumen V, W derselben Dimension n. Dann sind äquivalent

(i) Die Abbildung ϕ ist injektiv.
(ii) Sie ist surjektiv.
(iii) Sie ist bijektiv.

Beweis: „(i) $\Rightarrow$ (ii)“ Ist ϕ injektiv, dann gilt $\ker(\phi) = \{0_V\}$ nach Prop. (5.10). Es folgt $\dim \ker(\phi) = 0$, und der Dimensionssatz für lineare Abbildungen liefert $\dim \mathrm{im}(\phi) = \dim V - \dim \ker(\phi) = n - 0 = n$. Sei B eine Basis von $\mathrm{im}(\phi)$. Dann ist B eine n-elementige linear unabhängige Teilmenge von W und wegen $\dim W = n$ nach Folgerung (7.9) eine Basis von W. Es folgt $\mathrm{im}(\phi) = \mathrm{lin}(B) = W$ und damit die Surjektivität von ϕ.

„(ii) $\Rightarrow$ (iii)“ Ist ϕ surjektiv, dann gilt $\mathrm{im}(\phi) = W$ und somit $\dim \mathrm{im}(\phi) = \dim W = n$. Der Dimensionssatz für lineare Abbildungen liefert $\dim \ker(\phi) = \dim V - \dim \mathrm{im}(\phi) = n - n = 0$. Es folgt $\ker(\phi) = \{0_V\}$, also ist ϕ injektiv und damit insgesamt bijektiv.

„(iii) $\Rightarrow$ (i)“ Als bijektive Abbildung ist ϕ insbesondere injektiv. □

§ 9. *Koordinatensysteme, Matrizen und lineare Abbildungen*

Sei $n \in \mathbb{N}$ und V ein n-dimensionaler K-Vektorraum. Ein Tupel $\mathscr{B} = (v_1, ..., v_n)$ mit der Eigenschaft, dass $B = \{v_1, ..., v_n\}$ eine Basis von V ist, bezeichnen wir als ***geordnete Basis***. Aus der Definition folgt insbesondere, dass die Menge B aus n verschiedenen Elementen besteht.

Proposition 9.1 Sei V ein n-dimensionaler K-Vektorraum und $\mathscr{B} = (v_1, ..., v_n)$ eine geordnete Basis. Dann gibt es für jeden Vektor $v \in V$ ein eindeutig bestimmtes Tupel $(\lambda_1, ..., \lambda_n) \in K^n$ mit

$$v = \sum_{k=1}^{n} \lambda_k v_k.$$

Beweis: Weil $B = \{v_1, ..., v_n\}$ ein Erzeugendensystem von V ist, gibt es jedenfalls Elemente $\lambda_1, ..., \lambda_n \in K$, so dass $v = \sum_{k=1}^{n} \lambda_k v_k$ erfüllt ist. Nehmen wir nun an, dass $(\lambda_1, ..., \lambda_n)$ und $(\mu_1, ..., \mu_n)$ beides Tupel mit der Eigenschaft

$$\sum_{k=1}^{n} \lambda_k v_k = v = \sum_{k=1}^{n} \mu_k v_k$$

Dann folgt

$$\sum_{k=1}^{n} (\lambda_k - \mu_k) v_k = \sum_{k=1}^{n} \lambda_k v_k - \sum_{k=1}^{n} \mu_k v_k = v - v = 0_V.$$

Weil B auch linear unabhängig ist, folgt $\lambda_k - \mu_k = 0_K$ und somit $\lambda_k = \mu_k$ für $1 \le k \le n$. Damit ist die Eindeutigkeit bewiesen. □

Satz 9.2 Sei $n \in \mathbb{N}$, V ein n-dimensionaler K-Vektorraum und $\mathscr{B} = (v_1, ..., v_n)$ eine geordnete Basis. Dann ist durch die Zuordnung

$$\Phi_{\mathscr{B}} : V \longrightarrow K^n \quad , \quad \sum_{i=1}^{n} \lambda_i v_i \mapsto (\lambda_1, ..., \lambda_n)$$

ein Isomorphismus von K-Vektorräumen definiert. Wir nennen $\Phi_{\mathscr{B}}$ die ***Koordinatenabbildung*** bezüglich der Basis $\mathscr{B}$. Für jeden Vektor $v \in V$ sind $\Phi_{\mathscr{B}}(v) \in K^n$ die ***Koordinaten*** von V bezüglich der geordneten Basis $\mathscr{B}$.

Beweis: Zunächst überprüfen wir, dass eine $\Phi_{\mathscr{B}}$ eine lineare Abbildung ist. Seien $v, w \in V$ und $\lambda \in K$ vorgegeben. Sei $\Phi_{\mathscr{B}}(v) = (\lambda_1, ..., \lambda_n)$ und $\Phi_{\mathscr{B}}(w) = (\mu_1, ..., \mu_n)$. Dann gilt nach Definition der Koordinaten

$$v = \sum_{k=1}^{n} \lambda_k v_k \quad \text{und} \quad w = \sum_{k=1}^{n} \mu_k v_k.$$

Es folgt $v+w = \sum_{k=1}^{n}(\lambda_k+\mu_k)v_k$ und $\lambda v = \sum_{k=1}^{n}(\lambda\lambda_k)v_k$, also gilt $\Phi_{\mathscr{B}}(v+w) = (\lambda_1+\mu_1,...,\lambda_n+\mu_n) = \Phi_{\mathscr{B}}(v)+\Phi_{\mathscr{B}}(w)$ und $\Phi_{\mathscr{B}}(\lambda v) = (\lambda\lambda_1,...,\lambda\lambda_n) = \lambda\Phi_{\mathscr{B}}(v)$. Damit ist die Linearität nachgewiesen.

Nun zeigen wir, dass $\Phi_{\mathscr{B}}$ injektiv ist. Sei $v \in V$ mit $\Phi_{\mathscr{B}}(v) = (0_K,...,0_K)$. Dann folgt $v = \sum_{k=1}^{n} 0_K v_k = 0_V$. Also ist $\ker(\Phi_{\mathscr{B}}) = \{0_V\}$ und $\Phi_{\mathscr{B}}$ damit injektiv. Als injektive lineare Abbildung zwischen Vektorräumen gleicher Dimension ist $\Phi_{\mathscr{B}}$ nach Satz (8.13). □

Wir bemerken noch, dass nach Definition der Koordinatenabbildung das Bild $\Phi_{\mathscr{B}}(v_j)$ des j-ten Basisvektors gerade der j-te Einheitsvektor e_j ist. Es gilt nämlich $v_j = \sum_{k=1}^{n}\delta_{jk}v_k$, und die Koeffizienten δ_{jk} sind gerade die Komponenten des j-ten Einheitsvektors e_j.

Folgerung 9.3 Zwischen zwei beliebigen K-Vektorräumen derselben endlichen Dimension existiert ein Isomorphismus.

Beweis: Seien V und W zwei n-dimensionale K-Vektorräume, und seien $\mathscr{B},\mathscr{C}$ geordnete Basen von V bzw. W. Dann erhält man durch Komposition der Isomorphismen $\Phi_{\mathscr{B}} : V \to K^n$ und $\Phi_{\mathscr{C}}^{-1} : K^n \to W$ insgesamt einen Isomorphismus $\Phi_{\mathscr{C}}^{-1}\circ\Phi_{\mathscr{B}} : V \to W$ zwischen V und W. □

Für jedes $n \in \mathbb{N}$ sei $\mathscr{E}_n = (e_1,...,e_n)$ die Basis des K^n bestehend aus den Einheitsvektoren. Man nennt $\mathscr{E}_n$ auch die ***kanonische Basis*** von K^n. Für jeden Vektor $v = (v_1,...,v_n)$ gilt $\Phi_{\mathscr{E}_n}(v) = v$, also $\Phi_{\mathscr{E}_n} = \mathrm{id}_{K^n}$. Dies folgt unmittelbar aus der Gleichung $v = \sum_{k=1}^{n} v_k e_k$ und der Definition von $\Phi_{\mathscr{E}_n}(v)$.

Als nächstes untersuchen wir nun den Zusammenhang zwischen Matrizen und linearen Abbildungen.

Satz 9.4 *(Existenz und Eindeutigkeit linearer Abbildungen)*

Sei $n \in \mathbb{N}$, V ein n-dimensionaler und W ein beliebiger K-Vektorraum. Sei $(v_1,...,v_n)$ eine geordnete Basis von V und $(w_1,...,w_n)$ ein Tupel bestehend aus Elementen von W. Dann gibt es eine eindeutig bestimmte lineare Abbildung $\phi : V \to W$ mit $\phi(v_i) = w_i$ für $1 \le i \le n$.

Beweis: Zunächst beweisen wir die Existenz einer solchen linearen Abbildung. Jeder Vektor $v \in V$ besitzt eine eindeutige Darstellung $v = \sum_{i=1}^{n}\lambda_i v_i$ als Linearkombination der Basis, mit $\lambda_i \in K$ für $1 \le i \le n$. Wir definieren das Bild $\phi(v)$ durch $\phi(v) = \sum_{i=1}^{n}\lambda_i w_i$. Die so definierte Abbildung ist in der Tat linear. Ist nämlich $v' \in V$ ein weiterer Vektor mit der Darstellung $v' = \sum_{i=1}^{n}\lambda_i' v_i$ als Linearkombination, dann besitzt der Vektor $v+v'$ die Darstellung $v+v' = \sum_{i=1}^{n}(\lambda_i+\lambda_i')v_i$. Nach Definition der Abbildung ϕ erhalten wir

$$\phi(v+v') \quad = \quad \sum_{i=1}^{n}(\lambda_i+\lambda_i')w_i \quad = \quad \sum_{i=1}^{n}\lambda_i w_i + \sum_{i=1}^{n}\lambda_i' w_i \quad = \quad \phi(v)+\phi(v').$$

Ist $\lambda \in K$ beliebig, dann gilt $\lambda v = \sum_{i=1}^n \lambda\lambda_i v_i$ und folglich $\phi(\lambda v) = \sum_{i=1}^n \lambda\lambda_i w_i = \lambda \sum_{i=1}^n \lambda_i w_i = \lambda\phi(v)$. Damit ist die Linearität von ϕ nachgewiesen.

Es bleibt zu zeigen, dass ϕ durch die genannten Eigenschaften eindeutig bestimmt ist. Sei $\phi' : V \to W$ eine weitere lineare Abbildung mit $\phi'(v_i) = w_i$ für $1 \le i \le n$. Ist $v \in V$ dann ein beliebiger Vektor und $v = \sum_{i=1}^n \lambda_i v_i$ eine Darstellung als Linearkombination der Basis, dann folgt aus der Linearität beider Abbildungen

$$\phi(v) \quad = \quad \phi\left(\sum_{i=1}^n \lambda_i v_i\right) \quad = \quad \sum_{i=1}^n \lambda_i \phi(v_i) \quad = \quad \sum_{i=1}^n \lambda_i w_i \quad =$$

$$\sum_{i=1}^n \lambda_i \phi'(v_i) \quad = \quad \phi'\left(\sum_{i=1}^n \lambda_i v_i\right) \quad = \quad \phi'(v)$$

Also stimmen ϕ und ϕ' als Abbildungen überein. □

Aus dieser Beobachtung ergibt sich unmittelbar, dass jede lineare Abbildung zwischen Vektorräumen V, W der Dimensionen n und m durch eine Matrix $A \in \mathscr{M}_{m\times n,K}$ beschrieben werden kann.

Definition 9.5 Seien V, W endlich-dimensionale K-Vektorräume und $\mathscr{A} = (v_1, ..., v_n)$, $\mathscr{B} = (w_1, ..., w_m)$ geordnete Basen von V bzw. W. Ferner sei $A = (a_{ij})$ eine Matrix aus $\mathscr{M}_{m\times n,K)}$. Dann gibt es nach Satz (9.4) eine eindeutig bestimmte lineare Abbildung

$$\phi : V \longrightarrow W \qquad \text{mit} \qquad \phi(v_j) = \sum_{i=1}^m a_{ij} w_i \quad \text{für} \quad 1 \le j \le n.$$

Wir bezeichnen diese Abbildung ϕ mit $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)$ und nennen sie die ***lineare Abbildung zur Matrix*** A bezüglich der Basen $\mathscr{A}$ und $\mathscr{B}$.

Umgekehrt kann jeder linearen Abbildung auf natürliche Weise eine Matrix zugeordnet werden.

Definition 9.6 Seien V, W endlich-dimensionale K-Vektorräume und $\mathscr{A} = (v_1, ..., v_n)$, $\mathscr{B} = (w_1, ..., w_m)$ geordnete Basen von V bzw. W. Sei $\phi : V \to W$ eine lineare Abbildung. Für jedes $j \in \{1, ..., n\}$ stellen wir $\phi(v_j)$ als Linearkombination von $\mathscr{B}$ dar; es gilt

$$\phi(v_j) \quad = \quad \sum_{i=1}^m a_{ij} w_i \qquad 1 \le j \le n$$

mit eindeutig bestimmten Koeffizienten $a_{ij} \in K$. Wir nennen $A = (a_{ij}) \in \mathscr{M}_{m\times n,K}$ die ***Darstellungsmatrix*** von ϕ bezüglich der Basen $\mathscr{A}, \mathscr{B}$ und bezeichnen sie mit $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$.

Der folgende Satz besagt, dass man mit Hilfe der Darstellungsmatrix $\mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)$ und den $\mathscr{A}$-Koordinaten eines Vektors $v \in V$ die $\mathscr{B}$-Koordinaten des Bildvektors $\phi(v) \in W$ ausrechnen kann.

Satz 9.7 Seien die Bezeichnungen wie in der Definition gewählt. Dann gilt

$$\Phi_{\mathscr{B}}(\phi(v)) \quad = \quad \mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)\Phi_{\mathscr{A}}(v) \qquad \text{für alle} \quad v \in V.$$

Beweis: Sei $A = \mathscr{M}_{\mathscr{B}}^{\mathscr{A}}(\phi)$ und $\phi_A : K^n \to K^m$ die Abbildung $v \mapsto Av$ gegeben durch das Matrix-Vektor-Produkt. Zum Beweis der Gleichung $\Phi_{\mathscr{B}} \circ \phi = \phi_A \circ \Phi_{\mathscr{A}}$ genügt es auf Grund des Existenz- und Eindeutigkeitssatzes zu zeigen, dass

$$(\Phi_{\mathscr{B}} \circ \phi)(v_j) \quad = \quad (\phi_A \circ \Phi_{\mathscr{A}})(v_j) \qquad \text{für} \quad 1 \le j \le n$$

erfüllt ist. Für die linke Seite der Gleichung gilt nach Definition

$$(\Phi_{\mathscr{B}} \circ \phi)(v_j) \quad = \quad \Phi_{\mathscr{B}}\left(\sum_{i=1}^{m} a_{ij} w_i\right) \quad = \quad \sum_{i=1}^{m} a_{ij}\Phi_{\mathscr{B}}(w_i) \quad = \quad \sum_{i=1}^{m} a_{ij} e_i \quad = \quad (a_{1j}, \ldots, a_{mj}).$$

Für die rechte Seite erhalten wir

$$(\phi_A \circ \Phi_{\mathscr{A}})(v_j) \quad = \quad \phi_A(e_j) \quad = \quad (a_{1j}, \ldots, a_{mj}) \quad ,$$

denn nach Prop. (8.6) sind die Bilder der Einheitsvektoren unter ϕ_A genau die Spalten der Matrix A. Damit ist gezeigt, dass die beiden Abbildungen übereinstimmen. □

Proposition 9.8 Seien $m, n \in \mathbb{N}$, $A \in \mathscr{M}_{m\times n,K}$ und $\phi_A : K^n \to K^m$ die lineare Abbildung gegeben durch $\phi_A(v) = Av$ für alle $v \in K^n$. Dann gilt

$$\mathscr{M}_{\mathscr{E}_m}^{\mathscr{E}_n}(\phi_A) \quad = \quad A.$$

Beweis: Für $1 \le j \le n$ gilt nach Satz (9.7) jeweils

$$\mathscr{M}_{\mathscr{E}_m}^{\mathscr{E}_n}(\phi_A)e_j \quad = \quad \mathscr{M}_{\mathscr{E}_m}^{\mathscr{E}_n}(\phi_A)\Phi_{\mathscr{E}_n}(e_j) \quad = \quad \Phi_{\mathscr{E}_m}(\phi_A(e_j)) \quad = \quad \phi_A(e_j) \quad = \quad Ae_j.$$

Dies zeigt, dass die j-te Spalte von $\mathscr{M}_{\mathscr{E}_m}^{\mathscr{E}_n}(\phi_A)$ mit der j-ten Spalte von A übereinstimmt, für $1 \le j \le n$. □

Durch den folgenden Satz wird nun der entscheidende Zusammenhang zwischen Matrizen und linearen Abbildungen hergestellt.

Satz 9.9 Seien V, W endlich-dimensionale K-Vektorräume und $\mathscr{A}, \mathscr{B}$ geordnete Basen von V bzw. W. Dann sind durch die beiden Abbildungen

$$\mathscr{M}_{m\times n,K} \to \mathrm{Hom}_K(V,W)\ ,\ A \mapsto \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A) \qquad , \qquad \mathrm{Hom}_K(V,W) \to \mathscr{M}_{m\times n,K}\ ,\ \phi \mapsto \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$$

zueinander inverse Isomorphismen von K-Vektorräumen definiert.

Beweis: Sei $\mathscr{A} = (v_1,...,v_n)$ und $\mathscr{B} = (w_1,...,w_m)$. Zunächst beweisen wir, dass $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ eine lineare Abbildung ist. Seien dazu $A, B \in \mathscr{M}_{m\times n,K}$ und $\lambda \in K$ beliebig vorgegeben, $A = (a_{ij})$, $B = (b_{ij})$. Für $1 \le j \le n$ gilt

$$\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A+B)(v_j) \quad = \quad \sum_{i=1}^m (a_{ij}+b_{ij})w_i \quad = \quad \sum_{i=1}^m a_{ij}w_i + \sum_{i=1}^m b_{ij}w_i \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) + \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(B)(v_j)$$

und

$$\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(\lambda A)(v_j) \quad = \quad \sum_{i=1}^m (\lambda a_{ij})w_i \quad = \quad \lambda \sum_{i=1}^m a_{ij}w_i \quad = \quad \lambda \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) \quad = \quad (\lambda \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A))(v_j).$$

Damit sind die Gleichungen $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A+B) = \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A) + \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(B)$ und $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(\lambda A) = \lambda \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ bewiesen, und $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ ist tatsächlich eine lineare Abbildung. Um zu zeigen, dass die Abbildungen $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ und $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}$ zueinander invers sind, müssen die Gleichungen

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{L}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathscr{M}_{m\times n,K}} \qquad \text{und} \qquad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{M}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathrm{Hom}_K(V,W)}$$

überprüft werden. Zum Beweis der ersten Gleichung sei $A = (a_{ij}) \in \mathscr{M}_{m\times n,K}$ vorgegeben. Es gilt

$$\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) \quad = \quad \sum_{i=1}^m a_{ij}w_i \qquad \text{für} \quad 1 \le j \le n.$$

Diese Gleichungen zeigen, dass A die Darstellungsmatrix von $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)$ ist. Es gilt also

$$(\mathscr{M}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{L}^{\mathscr{A}}_{\mathscr{B}})(A) \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)) \quad = \quad A \quad = \quad \mathrm{id}_{\mathscr{M}_{m\times n,K}}(A).$$

Damit ist die Gleichung $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{L}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathscr{M}_{m\times n,K}}$ nachgewiesen.

Zum Beweis der zweiten Gleichung sei $\phi \in \mathrm{Hom}_K(V,W)$ vorgegeben und $A = \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ mit $A = (a_{ij})$. Dann gilt

$$\phi(v_j) \quad = \quad \sum_{i=1}^m a_{ij}w_i \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)(v_j) \qquad \text{für} \quad 1 \le j \le n.$$

Diese Gleichungen zeigen, dass die linearen Abbildungen ϕ und $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A)$ auf der Basis $\mathscr{A}$ von V übereinstimmen. Nach dem Existenz- und Eindeutigkeitssatz sind sie damit identisch. Es gilt also

$$(\mathscr{L}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{M}^{\mathscr{A}}_{\mathscr{B}})(\phi) \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)) \quad = \quad \mathscr{L}^{\mathscr{A}}_{\mathscr{B}}(A) \quad = \quad \phi \quad = \quad \mathrm{id}_{\mathrm{Hom}_K(V,W)}(\phi) \quad ,$$

also $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}} \circ \mathscr{M}^{\mathscr{A}}_{\mathscr{B}} = \mathrm{id}_{\mathrm{Hom}_K(V,W)}$. Insgesamt haben wir gezeigt, dass die Abbildungen $\mathscr{L}^{\mathscr{A}}_{\mathscr{B}}$ und $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}$ zueinander invers sind, damit insbesondere Umkehrabbildungen besitzen und folglich bijektiv sind. Auf Grund von Prop. (4.11) ist $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}$ als Umkehrabbildung einer linearen Abbildung ebenfalls linear. □

Folgerung 9.10 Seien V, W endlich-dimensionale K-Vektorräume. Dann gilt

$$\dim \mathrm{Hom}_K(V,W) \quad = \quad (\dim V)(\dim W).$$

Beweis: Sei $m = \dim W$ und $n = \dim V$. Nach Satz (9.9) sind die K-Vektorräume $\mathrm{Hom}_K(V,W)$ und $\mathscr{M}_{m\times n,K}$ isomorph. Daraus folgt $\dim \mathrm{Hom}_K(V,W) = \dim \mathscr{M}_{m\times n,K} = mn$. □

Lemma 9.11 Sei V ein endlich-dimensionaler K-Vektorraum und $\mathscr{A}$ eine geordnete Basis von V. Dann gilt

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V) \quad = \quad \mathrm{I}^{(n)} \quad ,$$

d.h. die Darstellungsmatrix der identischen Abbildung ist die Einheitsmatrix.

Beweis: Sei $\mathscr{A} = (v_1, ..., v_n)$. Für $1 \le j \le n$ gilt nach Satz (9.7) dann jeweils

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)e_j \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)\Phi_{\mathscr{A}}(v_j) \quad = \quad \Phi_{\mathscr{A}}(\mathrm{id}_V(v_j)) \quad = \quad \Phi_{\mathscr{A}}(v_j) \quad = \quad e_j.$$

Die Spalten von $\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)$ sind also gerade die Einheitsvektoren. Daraus folgt $\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V) = \mathrm{I}^{(n)}$. □

Satz 9.12 Seien U, V, W endlich-dimensionale K-Vektorräume mit geordneten Basen $\mathscr{A}$, $\mathscr{B}$ und $\mathscr{C}$. Seien $\phi : U \to V$ und $\psi : V \to W$ lineare Abbildungen. Dann gilt

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi) \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi).$$

Das Matrixprodukt entspricht also der Komposition linearer Abbildungen.

Beweis: Sei $n = \dim U$ und $\mathscr{A} = (v_1, ..., v_n)$. Für $1 \le j \le n$ gilt nach Satz (9.9) dann einerseits

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi)e_j \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi)\Phi_{\mathscr{A}}(v_j) \quad = \quad \Phi_{\mathscr{C}}((\psi \circ \phi)(v_j)) \quad ,$$

andererseits aber auch

$$\mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)e_j \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)\Phi_{\mathscr{A}}(v_j) \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\Phi_{\mathscr{B}}(\phi(v_j)) \quad =$$

$$\Phi_{\mathscr{C}}(\psi(\phi(v_j))) \quad = \quad \Phi_{\mathscr{C}}((\psi \circ \phi)(v_j)).$$

Also stimmt die j-te Spalte von $\mathscr{M}^{\mathscr{A}}_{\mathscr{C}}(\psi \circ \phi)$ mit der j-ten Spalte von $\mathscr{M}^{\mathscr{B}}_{\mathscr{C}}(\psi)\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ überein, für $1 \le j \le n$. □

Satz 9.13 Seien V, W beides n-dimensionale K-Vektorräume mit geordneten Basen $\mathscr{A}, \mathscr{B}$. Eine lineare Abbildung $\phi : V \to W$ ist genau dann bijektiv, wenn die Darstellungsmatrix $\mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ invertierbar ist, und in diesem Fall gilt

$$\mathcal{M}^{\mathscr{B}}_{\mathscr{A}}(\phi^{-1}) \quad = \quad \mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)^{-1}.$$

Beweis: Sei $\mathscr{A} = (v_1, ..., v_n)$ und $\mathscr{B} = (w_1, ..., w_n)$. Setzen wir zunächst voraus, dass die Matrix $A = \mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ invertierbar ist. Sei $B = A^{-1}$ die inverse Matrix und $\psi = \mathcal{L}^{\mathscr{B}}_{\mathscr{A}}(B)$. Weil die Abbildungen $\mathcal{L}^{\mathscr{B}}_{\mathscr{A}}$ und $\mathcal{M}^{\mathscr{B}}_{\mathscr{A}}$ nach Satz (9.9) zueinander invers sind, gilt $B = \mathcal{M}^{\mathscr{B}}_{\mathscr{A}}(\psi)$. Für $1 \le j \le n$ gilt dann einerseits

$$\mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\psi \circ \phi) e_j \quad = \quad \mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\psi \circ \phi) \Phi_{\mathscr{A}}(v_j) \quad = \quad \mathcal{M}^{\mathscr{B}}_{\mathscr{A}}(\psi) \mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \Phi_{\mathscr{A}}(v_j) \quad = \quad BA\Phi_{\mathscr{A}}(v_j) \quad = \quad \Phi_{\mathscr{A}}(v_j)$$

und andererseits

$$\mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V) e_j \quad = \quad \mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)(\Phi_{\mathscr{A}}(v_j)) \quad = \quad \Phi_{\mathscr{A}}(\mathrm{id}_V(v_j)) \quad = \quad \Phi_{\mathscr{A}}(v_j).$$

Die j-te Spalte von $\mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\psi \circ \phi)$ stimmt also mit der j-ten Spalte von $\mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V)$ überein. Also sind die beiden Matrizen gleich. Weil die Zuordnung $\mathcal{M}^{\mathscr{A}}_{\mathscr{A}}$ bijektiv ist, folgt $\psi \circ \phi = \mathrm{id}_V$. Nach dem gleichen Schema zeigt man $\phi \circ \psi = \mathrm{id}_W$. Damit ist die Bijektivität von ϕ bewiesen.

Gehen wir nun umgekehrt davon aus, dass ϕ bijektiv ist, und sei ψ die Umkehrabbildung. Mit Hilfe von Lemma (9.11) erhalten wir

$$\mathcal{M}^{\mathscr{B}}_{\mathscr{A}}(\psi) \mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \quad = \quad \mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\psi \circ \phi) \quad = \quad \mathcal{M}^{\mathscr{A}}_{\mathscr{A}}(\mathrm{id}_V) \quad = \quad \mathrm{I}^{(n)}$$

Dies zeigt die Invertierbarkeit von $\mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ und beweist zugleich auch die Identität $\mathcal{M}^{\mathscr{B}}_{\mathscr{A}}(\phi^{-1}) = \mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)^{-1}$.
□

Definition 9.14 Seien V ein endlich-dimensionaler K-Vektorraum, und seien $\mathscr{A}, \mathscr{B}$ zwei geordnete Basen von V. Dann nennt man $\mathcal{T}^{\mathscr{A}}_{\mathscr{B}} = \mathcal{M}^{\mathscr{A}}_{\mathscr{B}}(\mathrm{id}_V)$ die ***Matrix des Basiswechsels*** von $\mathscr{A}$ nach $\mathscr{B}$ oder auch einfach eine ***Transformationsmatrix***.

Die wesentliche Eigenschaft der Transformationsmatrix $\mathcal{T}^{\mathscr{A}}_{\mathscr{B}}$ besteht darin, dass sie die $\mathscr{A}$-Koordinaten eines Vektors in $\mathscr{B}$-Koordinaten umrechnet.

Proposition 9.15 Seien Bezeichnungen wie in der Definition und $n = \dim V$.

(i) Für alle $v \in V$ gilt $\mathcal{T}^{\mathscr{A}}_{\mathscr{B}} \Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{B}}(v)$.

(ii) Es gilt $\mathcal{T}^{\mathscr{A}}_{\mathscr{A}} = \mathrm{I}^{(n)}$ und $\mathcal{T}^{\mathscr{B}}_{\mathscr{A}} = (\mathcal{T}^{\mathscr{A}}_{\mathscr{B}})^{-1}$.

Beweis: Für jeden Vektor $v \in V$ gilt nach Satz (9.7) jeweils

$$\mathscr{T}^{\mathscr{A}}_{\mathscr{B}} \Phi_{\mathscr{A}}(v) \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\mathrm{id}_V)\Phi_{\mathscr{A}}(v) \quad = \quad \Phi_{\mathscr{B}}(\mathrm{id}_V(v)) \quad = \quad \Phi_{\mathscr{B}}(v).$$

Die Gleichung $\mathscr{T}^{\mathscr{A}}_{\mathscr{A}} = \mathrm{I}^{(n)}$ ist eine direkte Folgerung aus der Gleichung $\mathscr{T}^{\mathscr{A}}_{\mathscr{A}} \Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{A}}(v)$, denn mit $v \in V$ durchläuft $\Phi_{\mathscr{A}}(v)$ alle Vektoren aus K^n. Schließlich liefert Satz (9.13) noch

$$(T^{\mathscr{A}}_{\mathscr{B}})^{-1} \quad = \quad (M^{\mathscr{A}}_{\mathscr{B}}(\mathrm{id}_V))^{-1} \quad = \quad M^{\mathscr{B}}_{\mathscr{A}}(\mathrm{id}_V^{-1}) \quad = \quad M^{\mathscr{B}}_{\mathscr{A}}(\mathrm{id}_V) \quad = \quad T^{\mathscr{B}}_{\mathscr{A}}. \qquad \square$$

Satz 9.16 *(Transformationsformel / Satz vom Basiswechsel)*

Seien V, W endlich-dimensionale K-Vektorräume, $\mathscr{A}, \mathscr{A}'$ zwei Basen von V und $\mathscr{B}, \mathscr{B}'$ zwei Basen von W. Für jede lineare Abbildung $\phi : V \to W$ gilt dann

$$\mathscr{M}^{\mathscr{A}'}_{\mathscr{B}'}(\phi) \quad = \quad \mathscr{T}^{\mathscr{B}}_{\mathscr{B}'} \, \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \, \mathscr{T}^{\mathscr{A}'}_{\mathscr{A}}.$$

Beweis: Die Gleichung erhält man direkt durch Anwendung von Satz (9.12). Es gilt

$$\mathscr{T}^{\mathscr{B}}_{\mathscr{B}'} \, \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \, \mathscr{T}^{\mathscr{A}'}_{\mathscr{A}} \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{B}'}(\mathrm{id}_W) \, \mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi) \, \mathscr{M}^{\mathscr{A}'}_{\mathscr{A}}(\mathrm{id}_V) \quad = \quad \mathscr{M}^{\mathscr{A}}_{\mathscr{B}'}(\mathrm{id}_W \circ \phi) \mathscr{M}^{\mathscr{A}'}_{\mathscr{A}}(\mathrm{id}_V) \quad =$$

$$\mathscr{M}^{\mathscr{A}'}_{\mathscr{B}'}(\mathrm{id}_W \circ \phi \circ \mathrm{id}_V) \quad = \quad \mathscr{M}^{\mathscr{A}'}_{\mathscr{B}'}(\phi). \qquad \square$$

§ 10. Determinanten

In diesem Abschnitt erweist es sich an vielen Stellen als praktisch, Matrizen als Tupel bestehend aus ihren Zeilenvektoren darzustellen. Sei K ein Körper und $n \in \mathbb{N}$. Dann schreiben wir $(a_{1\bullet}, ..., a_{n\bullet})$ für die Matrix $A = (a_{ij}) \in \mathscr{M}_{n,K}$ mit den Zeilenvektoren $a_{k\bullet}$ für $1 \le k \le n$. Ist $v_k \in K^n$ ein beliebiger Vektor, dann schreiben wir $(a_{1\bullet}, ..., v_k, ..., a_{n\bullet})$ für die Matrix, die man erhält, wenn man die k-te Zeile $a_{k\bullet}$ durch den Vektor v_k ersetzt.

Definition 10.1

(i) Eine Abbildung $d : \mathscr{M}_{n,K} \to K$ bezeichnet man als ***multilinear***, wenn für $1 \le k \le n$ und alle $a_{1\bullet}, ..., a_{n\bullet}, a'_{k\bullet} \in K^n$ und alle $\lambda \in K$ jeweils

$$d(a_{1\bullet}, ..., a_{k\bullet} + a'_{k\bullet}, ..., a_{n\bullet}) \quad = \quad d(a_{1\bullet}, ..., a_{k\bullet}, ..., a_{n\bullet}) + d(a_{1\bullet}, ..., a'_{k\bullet}, ..., a_{n\bullet})$$

und $d(a_{1\bullet}, ..., \lambda a_{k\bullet}, ..., a_{n\bullet}) = \lambda d(a_{1\bullet}, ..., a_{k\bullet}, ..., a_{n\bullet})$ erfüllt ist.

(ii) Man bezeichnet eine multilineare Abbildung $d : \mathscr{M}_{n,K} \to K$ als ***alternierend***, wenn $d(A) = 0$ gilt, sobald zwei Zeilen von $A \in \mathscr{M}_{n,K}$ übereinstimmen.

(iii) Eine multilineare, alternierende Abbildung d mit $d(\mathrm{I}^{(n)}) = 1$ bezeichnet man als ***Determinantenfunktion***.

Unser Ziel in diesem Abschnitt besteht darin zu zeigen, dass für jeden Körper K und jedes $n \in \mathbb{N}$ genau eine Determinantenfunktion auf $\mathscr{M}_{n,K}$ existiert. Vor allem für den Nachweis der Existenz benötigen wir als algebraisches Hilfsmittel die symmetrischen Gruppen.

Proposition 10.2 Für jedes $n \in \mathbb{N}$ sei $M_n = \{1, ..., n\}$ die Menge der Zahlen von 1 bis n. Dann bilden die bijektiven Abbildungen $\sigma : M_n \to M_n$ mit der Komposition von Abbildungen als Verknüpfung eine Gruppe. Wir nennen sie die ***symmetrische Gruppe*** in n Elementen und bezeichnen sie mit S_n.

Beweis: Zunächst beweisen wir das Assoziativgesetz. Für alle $\rho, \sigma, \tau \in S_n$ und alle $x \in M_n$ gilt

$$((\rho \circ \sigma) \circ \tau)(x) \quad = \quad (\rho \circ \sigma)(\tau(x)) \quad = \quad \rho(\sigma(\tau(x))) \quad = \quad \rho((\sigma \circ \tau)(x)) \quad = \quad (\rho \circ (\sigma \circ \tau))(x)$$

und somit $(\rho \circ \sigma) \circ \tau = \rho \circ (\sigma \circ \tau)$. Damit ist die Assoziativität nachwiesen. Die identische Abbildung $\mathrm{id} \in S_n$ gegeben durch $\mathrm{id}(x) = x$ für alle $x \in M_n$ ist in $(S_n, \circ)$ das Neutralelement, denn für alle $\sigma \in S_n$ und alle $x \in M_n$ gilt

$$(\sigma \circ \mathrm{id})(x) \quad = \quad \sigma(\mathrm{id}(x)) \quad = \quad \sigma(x) \quad = \quad \mathrm{id}(\sigma(x)) \quad = \quad (\mathrm{id} \circ \sigma)(x)$$

also $\sigma\circ\mathrm{id} = \mathrm{id}\circ\sigma$. Weil jedes $\sigma \in S_n$ bijektiv ist, existiert jeweils die Umkehrabbildung σ^{-1}. Diese ist ebenfalls bijektiv, also ein Element in S_n. Für alle $x \in M_n$ gilt nach Definition der Umkehrabbildung

$$(\sigma^{-1}\circ\sigma)(x) \quad = \quad \sigma^{-1}(\sigma(x)) \quad = \quad x \quad = \quad \mathrm{id}(x)$$

und somit $\sigma^{-1}\circ\sigma = \mathrm{id}$. Ebenso zeigt man $\sigma\circ\sigma^{-1} = \mathrm{id}$. Damit ist nachgewiesen, dass σ^{-1} in $(S_n,\circ)$ das zu σ inverse Element ist. Jedes Element in S_n hat also ein Inverses, damit ist $(S_n,\circ)$ eine Gruppe. □

Elemente in S_n können durch Wertetabellen dargestellt werden. Beispielsweise schreibt man

$$\sigma \quad = \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$$

für das Element $\sigma \in S_4$, das durch $\sigma(1) = 1$, $\sigma(2) = 4$, $\sigma(3) = 2$ und $\sigma(4) = 3$ gegeben ist. Aus der Analysis einer Variablen ist bekannt, dass für jedes $n \in \mathbb{N}$ und beliebige Mengen A, B mit $|A| = |B| = n$ jeweils genau $n!$ bijektive Abbildungen $A \to B$ existieren. Also gilt auch $|S_n| = n!$ für alle $n \in \mathbb{N}$.

Wir werden nun sehen, wie die symmetrische Gruppe S_n mit Determinantenfunktionen auf $\mathscr{M}_{n,K}$ zusammenhängt. Dazu bezeichnen wir mit $\mathrm{Abb}(M_n)$ die Menge aller (nicht notwendig bijektiven) Abbildungen $M_n \to M_n$. Aus dem ersten Semester wissen wir, dass für eine Abbildung $\sigma \in \mathrm{Abb}(M_n)$ die Eigenschaften injektiv, surjektiv und bijektiv äquivalent sind. Ein Element aus $\mathrm{Abb}(M_n)$, das nicht in S_n liegt, ist also weder injektiv noch surjektiv.

Sei nun $d : \mathscr{M}_{n,K} \to K$ eine Determinantenfunktion und $A = (a_{ij}) \in \mathscr{M}_{n,K}$. Mit $e_1, ..., e_n$ bezeichnen wir wie immer die Einheitsvektoren in K^n. Die Darstellung der Zeilenvektoren als Linearkombination der Einheitsvektoren liefert $a_{k\bullet} = \sum_{i=1}^n a_{ki} e_i$ für $1 \le k \le n$. Auf Grund der Multilinearität von d gilt

$$d(A) \quad = \quad d(a_{1\bullet}, ..., a_{n\bullet}) \quad = \quad d\left(\sum_{i_1=1}^n a_{1i_1} e_{i_1}, a_{2\bullet}, ..., a_{n\bullet}\right) \quad =$$

$$\sum_{i_1=1}^n a_{1i_1} d(e_{i_1}, a_{2\bullet}, ..., a_{n\bullet}) \quad = \quad \sum_{i_1=1}^n \sum_{i_2=1}^n a_{1i_1} a_{2i_2} d(e_{i_1}, e_{i_2}, a_{3\bullet}, ..., a_{n\bullet}) \quad = \quad ... \quad =$$

$$\sum_{i_1=1}^n ... \sum_{i_n=1}^n a_{1i_1} ... a_{ni_n} d(e_{i_1}, ..., e_{i_n}) \quad = \quad \sum_{\sigma\in\mathrm{Abb}(M_n)} a_{1\sigma(1)} ... a_{n\sigma(n)} d(e_{\sigma(1)}, ..., e_{\sigma(n)})$$

Jedes Element $\sigma \in \mathrm{Abb}(M_n)$ mit $\sigma \notin S_n$ ist auf Grund unserer Vorbemerkung insbesondere nicht injektiv, es gibt also $i, j \in M_n$ mit $i \neq n$ und $\sigma(i) = \sigma(j)$. Weil die Funktion d alternierend ist, gilt dann $d(e_{\sigma(1)}, ..., e_{\sigma(n)}) = 0$. Somit verschwinden in der Summe sämtliche Summanden, die zu Abbildungen $\sigma \in \mathrm{Abb}(M_n) \setminus S_n$ gehören. Wir erhalten

$$d(A) \quad = \quad \sum_{\sigma\in S_n} a_{1\sigma(1)} ... a_{n\sigma(n)} d(e_{\sigma(1)}, ..., e_{\sigma(n)}).$$

Eine Matrix der Form $P_\sigma = (e_{\sigma(1)}, ..., e_{\sigma(n)})$ mit $\sigma \in S_n$ bezeichnet man als ***Permutationsmatrix***. Insbesondere ist $P_{\mathrm{id}} = \mathrm{I}^{(n)}$ die Einheitsmatrix. Unserer Rechnung hat ergeben

Proposition 10.3 Sei $d : \mathscr{M}_{n,K} \to K$ eine Determinantenfunktion und $A = (a_{ij}) \in \mathscr{M}_{n,K}$. Dann gilt

$$d(A) \quad = \quad \sum_{\sigma \in S_n} a_{1\sigma(1)} \ldots a_{n\sigma(n)} d(P_\sigma).$$

Nun zeigen wir noch, dass auch die Werte $d(P_\sigma)$ durch die Eigenschaften der Determinantenfunktion eindeutig festgelegt sind. Sei $n \in \mathbb{N}$, und seien $k, \ell \in M_n$ zwei verschiedene Zahlen. Dann ist die Abbildung $\sigma : M_n \to M_n$ gegeben durch

$$\sigma(x) \quad = \quad \begin{cases} \ell & \text{falls } x = k \\ k & \text{falls } x = \ell \\ x & \text{sonst} \end{cases}$$

bijektiv, liegt also in S_n. Man verwendet für dieses spezielle Element von S_n die Notation $(k \;\; \ell)$. Allgemein werden Elemente in S_n in dieser Form als ***Transpositionen*** bezeichnet. Jede Transposition τ hat die Eigenschaft $\tau \circ \tau = \mathrm{id}$, denn die Vertauschung von je zwei Elementen wird durch Wiederholungs des Vorgangs wieder rückgängig gemacht. Es gilt also

$$\tau \quad = \quad \tau^{-1} \qquad \text{für jede Transposition } \tau \in S_n.$$

Wir bestimmen nun das Bild $d(P_\sigma)$ unter einer Determinantenfunktion d zunächst für den Fall, dass σ eine Transposition ist. Allgemein gilt

Lemma 10.4 Sei $d : \mathscr{M}_{n,K} \to K$ eine Determinantenfunktion, und seien $A, B \in \mathscr{M}_{n,K}$. Entsteht B aus A durch Vertauschung zweier Zeilen, dann gilt $d(B) = -d(A)$.

Beweis: Seien $k, \ell \in M_n$ die beiden Zeilenindizes mit der Eigenschaft, dass B aus A durch Vertauschung der k-ten mit der ℓ-ten Zeile entsteht, wobei $k < \ell$ ist. Weil die Determinantenfunktion multilinear und alternierend ist, gilt

$$d(A) + d(B) \quad = \quad d(\ldots, a_{k\bullet}, \ldots, a_{\ell\bullet}, \ldots) + d(\ldots, a_{\ell\bullet}, \ldots, a_{k\bullet}, \ldots) \quad =$$

$$d(\ldots, a_{k\bullet}, \ldots, a_{k\bullet}, \ldots) + d(\ldots, a_{k\bullet}, \ldots, a_{\ell\bullet}, \ldots) + d(\ldots, a_{\ell\bullet}, \ldots, a_{k\bullet}, \ldots) + d(\ldots, a_{\ell\bullet}, \ldots, a_{\ell\bullet}, \ldots) \quad =$$

$$d(\ldots, a_{k\bullet}, \ldots, a_{k\bullet} + a_{\ell\bullet}, \ldots) + d(\ldots, a_{\ell\bullet}, \ldots, a_{k\bullet} + a_{\ell\bullet}, \ldots) \quad = \quad d(\ldots, a_{k\bullet} + a_{\ell\bullet}, \ldots, a_{k\bullet} + a_{\ell\bullet}, \ldots) \quad = \quad 0.$$

Daraus folgt $d(B) = -d(A)$. □

Die Permutationsmatrix $P_{(k\;\ell)}$ entsteht aus der Einheitsmatrix $\mathrm{I}^{(n)}$ durch Vertauschung der k-ten und ℓ-ten Zeile. Nach Eigenschaft (iii) der Determinatenfunktionen gilt also $d(P_{(k\;\ell)}) = -d(\mathrm{I}^{(n)}) = -1$. Als nächstes bestimmen wir $d(P_\sigma)$ für beliebige Elemente $\sigma \in S_n$. Dazu bemerken wir zunächst, dass jede Permutation aus Transpositionen zusammengesetzt werden kann.

Proposition 10.5 Jedes Element aus S_n ist darstellbar als Produkt von Transpositionen.

Beweis: Sei $\sigma \in S_n$ vorgegeben. Wir beweisen durch vollständige Induktion über $k \in \{0,...,n\}$: Es gibt ein Produkt τ von Transpositionen, so dass $(\tau \circ \sigma)(i) = i$ für $1 \le i \le k$ erfüllt ist. Die Aussage für $k = n$ liefert dann $\tau \circ \sigma = \mathrm{id} \Leftrightarrow \sigma = \tau^{-1}$. Mit τ ist dann auch das Element $\sigma = \tau^{-1}$ ein Produkt von Transpositionen. Ist nämlich $\tau = \tau_1 \circ ... \circ \tau_\ell$ mit Transpositionen τ_i, dann erhalten wir für τ^{-1} die Produktdarstellung

$$\tau^{-1} \quad = \quad \tau_\ell^{-1} \circ ... \circ \tau_1^{-1} \quad = \quad \tau_\ell \circ ... \circ \tau_1.$$

Kommen wir nun zum Induktionsbeweis. Für $k = 0$ ist nichts zu zeigen. Sei nun $k \in \{0,...,n-1\}$, und setzen wir die Aussage für k voraus. Dann gibt es ein Produkt τ von Transpositionen, so dass die Permutation $\tilde{\sigma} = \tau \circ \sigma$ für $1 \le i \le k$ die Gleichung $\tilde{\sigma}(i) = i$ erfüllt. Gilt nun $\tilde{\sigma}(k+1) = k+1$, dann erfüllt τ die gewünschte Aussage auch für $k+1$. Ansonsten setzen wir $\ell = \tilde{\sigma}(k+1)$; auf Grund der Injektivität von $\tilde{\sigma}$ und wegen $\tilde{\sigma}(i) = i$ für $1 \le i \le k$ und $\tilde{\sigma}(k+1) \ne k+1$ ist $\ell > k+1$. Für das Produkt $\tau' = (k+1 \ \ \ell) \circ \tau$ von Transpositionen gilt dann

$$(\tau' \circ \sigma)(k+1) \quad = \quad ((k+1 \ \ \ell) \circ \tau \circ \sigma)(k+1) \quad = \quad ((k+1 \ \ \ell) \circ \tilde{\sigma})(k+1) \quad = \quad (k+1 \ \ \ell)(\ell) \quad = \quad k+1 \quad ,$$

wodurch der Induktionsschritt abgeschlossen ist. □

Definition 10.6 Sei $\sigma \in S_n$ ein beliebiges Element. Eine zweielementige Teilmenge $\{i, j\}$ von M_n wird ***Fehlstand*** von σ genannt, wenn $i < j$, aber $\sigma(i) > \sigma(j)$ gilt. Ist $k \in \mathbb{N}_0$ die Anzahl der Fehlstände von σ, dann nennt man $\mathrm{sgn}(\sigma) = (-1)^k$ das ***Signum*** oder ***Vorzeichen*** der Permutation.

Lemma 10.7 Für jede Transposition τ gibt es ein Element $\sigma \in S_n$ mit $\tau = \sigma \circ (1 \ \ 2) \circ \sigma^{-1}$.

Beweis: Sei $\tau = (k \ \ \ell)$ mit $k, \ell \in \{1,...,n\}$ und $\sigma \in S_n$ ein beliebiges Element mit $\sigma \in S_n$ mit $\sigma(1) = k$ und $\sigma(2) = \ell$. Setzen wir $\tau' = \sigma \circ (1 \ \ 2) \circ \sigma^{-1}$, dann ist $\tau = \tau'$ zu überprüfen. Zunächst gilt

$$\tau'(k) \quad = \quad (\sigma \circ (1 \ \ 2) \circ \sigma^{-1})(k) \quad = \quad (\sigma \circ (1 \ \ 2))(1) \quad = \quad \sigma(2) \quad = \quad \ell$$

und $\tau'(\ell) = (\sigma \circ (1 \ \ 2) \circ \sigma^{-1})(\ell) = (\sigma \circ (1 \ \ 2))(2) = \sigma(1) = k$. Ist $i \notin \{k, \ell\}$, dann ist $\sigma^{-1}(i) \notin \{1, 2\}$, denn die Elemente 1 und 2 werden in die Menge $\{k, \ell\}$ abgebildet. Wir erhalten $(1 \ \ 2)(\sigma^{-1}(i)) = \sigma^{-1}(i)$ und somit

$$\tau'(i) \quad = \quad (\sigma \circ (1 \ \ 2) \circ \sigma^{-1})(i) \quad = \quad (\sigma \circ (1 \ \ 2))(\sigma^{-1}(i)) \quad = \quad \sigma(\sigma^{-1}(i)) \quad = \quad i.$$

Insgesamt gilt $\tau(i) = \tau'(i)$ für $1 \le i \le n$, also $\tau = \tau'$. □

Lemma 10.8 Für jedes $\sigma \in S_n$ gilt die Produktformel $\mathrm{sgn}(\sigma) = \prod_{i<j} \frac{\sigma(j) - \sigma(i)}{j - i}$.

Beweis: Sei $\sigma \in S_n$ und m die Anzahl der Fehlstände von σ. Dann gilt

$$\prod_{i<j}(\sigma(j)-\sigma(i)) \quad = \quad \prod_{\substack{i<j \\ \sigma(j)>\sigma(i)}}(\sigma(j)-\sigma(i)) \cdot \prod_{\substack{i<j \\ \sigma(j)<\sigma(i)}}(\sigma(j)-\sigma(i)) \quad =$$

$$\prod_{\substack{i<j \\ \sigma(j)>\sigma(i)}}(\sigma(j)-\sigma(i)) \cdot (-1)^m \prod_{\substack{i<j \\ \sigma(j)<\sigma(i)}}|\sigma(j)-\sigma(i)| \quad = \quad (-1)^m \prod_{i<j}|\sigma(j)-\sigma(i)|$$

Sei $\mathscr{T}$ die Menge der zweielementigen Teilmengen von $\{1,...,n\}$. Dann entsprechen die Paare (i,j) mit $i<j$ und $i,j \in \{1,...,n\}$ bijektiv den Mengen in $\mathscr{T}$. Weil σ bijektiv ist, ist mit $\{i,j\} \in \mathscr{T}$ auch $\{\sigma(i),\sigma(j)\}$ eine zweielementige Menge, also in $\mathscr{T}$ enthalten. Die Zuordnung $\mathscr{T} \to \mathscr{T}$, $\{i,j\} \mapsto \{\sigma(i),\sigma(j)\}$ ist bijektiv, da durch $\{i,j\} \mapsto \{\sigma^{-1}(i),\sigma^{-1}(j)\}$ eine Umkehrabbildung gegeben ist; wir bezeichnen diese ebenfalls mit σ. Für jedes $S=\{i,j\} \in \mathscr{T}$ sei außerdem $r_S = |i-j| \in \mathbb{N}$. Wir erhalten nun

$$(-1)^m \prod_{i<j}|\sigma(j)-\sigma(i)| \quad = \quad (-1)^m \prod_{\{i,j\}\in\mathscr{T}}|\sigma(j)-\sigma(i)| \quad = \quad (-1)^m \prod_{S\in\mathscr{T}} r_{\sigma(S)} \quad =$$

$$(-1)^m \prod_{S\in\sigma(\mathscr{T})} r_S \quad = \quad (-1)^m \prod_{S\in\mathscr{T}} r_S \quad = \quad (-1)^m \prod_{i<j}|j-i| \quad = \quad (-1)^m \prod_{i<j}(j-i).$$

Insgesamt haben wir also gezeigt, dass

$$\prod_{i<j}(\sigma(j)-\sigma(i)) \quad = \quad (-1)^m \prod_{i<j}(j-i) \quad = \quad \operatorname{sgn}(\sigma)\prod_{i<j}(j-i) \qquad \text{erfüllt ist.}$$ □

Proposition 10.9 Für beliebige Elemente $\sigma,\tau \in S_n$ gilt $\operatorname{sgn}(\tau\circ\sigma) = \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma)$.

Beweis: Auf Grund des vorhergehenden Lemmas ist $\operatorname{sgn}(\tau\circ\sigma)$ gegeben durch

$$\prod_{i<j}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{j-i} \quad = \quad \prod_{i<j}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \prod_{i<j}\frac{\sigma(j)-\sigma(i)}{j-i}.$$

Das zweite Produkt stimmt mit $\operatorname{sgn}(\sigma)$ überein. Es genügt also zu zeigen, dass das erste Produkt gleich $\operatorname{sgn}(\tau)$ ist. Es gilt

$$\prod_{i<j}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad = \quad \prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad =$$

$$\prod_{\substack{i<j \\ \sigma(i)<\sigma(j)}}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \cdot \prod_{\substack{i>j \\ \sigma(i)<\sigma(j)}}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad = \quad \prod_{\sigma(i)<\sigma(j)}\frac{\tau(\sigma(j))-\tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad ,$$

wobei wir beim zweiten „=“ lediglich die Rollen von i und j im zweiten Faktor vertauscht haben. Um das Produkt weiter zu vereinfachen, beweisen wir die Gleichung

$$\{(k,\ell) \mid 1 \le k < \ell \le n\} \quad = \quad \{(\sigma(i),\sigma(j)) \mid 1 \le i,j \le n\,,\ \sigma(i)<\sigma(j)\}.$$

Der Beweis der Inklusion „$\supseteq$“ ist offensichtlich, denn das Paar (k,ℓ) mit $k = \sigma(i)$, $j = \sigma(\ell)$ erfüllt nach Definition die Bedingungen an die Elemente in der Menge links. Zum Nachweis von „$\subseteq$“ sei ein Paar (k,ℓ) in der Menge links vorgegeben. Sei $i = \sigma^{-1}(k)$ und $j = \sigma^{-1}(\ell)$. Dann gilt $\sigma(i) = k < \ell = \sigma(j)$, also ist $(k,\ell) = (\sigma(i),\sigma(j))$ ein Element der Menge rechts. Es folgt nun

$$\prod_{\sigma(i)<\sigma(j)} \frac{\tau(\sigma(j)) - \tau(\sigma(i))}{\sigma(j)-\sigma(i)} \quad = \quad \prod_{k<\ell} \frac{\tau(\ell)-\tau(k)}{\ell - k} \quad = \quad \operatorname{sgn}(\tau). \qquad \square$$

Folgerung 10.10 Für jede Permutation $\sigma \in S_n$ gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Beweis: Da das Neutralelement id der Gruppe keine Fehlstände besitzt, gilt $\operatorname{sgn}(\mathrm{id}) = 1$. Aus

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

folgt dann die behauptete Gleichung. $\square$

Satz 10.11

(i) Ist τ eine Transposition, dann gilt $\operatorname{sgn}(\tau) = -1$.

(ii) Ist σ als Produkt von r Transpositionen darstellbar, dann gilt $\operatorname{sgn}(\sigma) = (-1)^r$.

Beweis: zu (i) Die Transposition $(1\ \ 2)$ hat offenbar $\{1,2\}$ als einzigen Fehlstand, also gilt $\operatorname{sgn}((1\ \ 2)) = -1$. Für eine beliebige Transposition τ finden wir nach Lemma (10.7) immer ein $\sigma \in S_n$ mit $\tau = \sigma \circ (1\ \ 2) \circ \sigma^{-1}$. Es folgt $\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}((1\ \ 2))\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)^2\operatorname{sgn}((1\ \ 2)) = -1$.

zu (ii) Ist $\sigma = \tau_1 \circ \ldots \circ \tau_r$ bestehend aus Transpositionen $\tau_1,\ldots,\tau_r$, dann folgt aus Teil (i) die Gleichung $\operatorname{sgn}(\sigma) = \prod_{i=1}^r \operatorname{sgn}(\tau_i) = \prod_{i=1}^r (-1) = (-1)^r$. $\square$

Nun sind wir in der Lage, die Determinante der Permutationsmatrizen anzugeben.

Folgerung 10.12 Sei $d : \mathscr{M}_{n,K} \to K$ eine Determinantenfunktion. Dann gilt

$$d(P_\sigma) \quad = \quad \operatorname{sgn}(\sigma) \qquad \text{für alle} \quad \sigma \in S_n.$$

Beweis: Ist $\rho \in S_n$ und $\tau = (k\ \ \ell)$ eine Transposition, dann gilt $d(P_{\rho\circ\tau}) = -d(P_\rho)$ nach Lemma (10.4), denn die Matrix $P_{\rho\circ\tau}$ entsteht aus P_ρ durch Vertauschung k-te und ℓ-ten Zeile: Für $i \neq k,\ell$ ist die i-te Zeile von $P_{\rho\circ\tau}$ gegeben durch $e_{(\rho\circ\tau)(i)} = e_{\rho(i)}$, und die k-te und ℓ-te Zeile sind $e_{(\rho\circ\tau)(k)} = e_{\rho(\ell)}$ bzw. $e_{(\rho\circ\tau)(\ell)} = e_{\rho(k)}$.

Sei nun $\sigma \in S_n$ beliebig vorgegeben. Nach Prop. (10.5) gibt es ein $r \in \mathbb{N}_0$ und Transpositionen $\tau_1, ..., \tau_r \in S_n$, so dass $\sigma = \tau_1 \circ ... \circ \tau_r$ erfüllt ist. Aus Lemma (10.8) folgt daraus $\text{sgn}(\sigma) = (-1)^r$. Wir beweisen die behauptete Gleichung nun durch vollständige Induktion über r. Im Fall $r = 0$ gilt $\sigma = \text{id}$ und $d(P_\sigma) = d(\mathrm{I}^{(n)}) = 1$ auf Grund der Bedingung (iii) für Determinantenfunktionen.

Sei nun $r > 1$, und setzen wir die Gleichung für Werte $< r$ voraus. Definieren wir das Element $\sigma' \in S_n$ durch $\sigma' = \tau_1 \circ ... \circ \tau_{r-1}$, dann gilt $\sigma = \sigma' \circ \tau_r$, außerdem $\text{sgn}(\sigma') = (-1)^{r-1}$ und auf Grund unserer Vorüberlegung

$$d(P_\sigma) \quad = \quad d(P_{\sigma' \circ \tau_r}) \quad = \quad -d(P_{\sigma'}) \quad = \quad -(-1)^{r-1} \quad = \quad (-1)^r \quad = \quad \text{sgn}(\sigma). \qquad \square$$

Setzen wir dieses Ergebnis in Prop. (10.3) ein, so erhalten wir

Folgerung 10.13 Sei $d : \mathcal{M}_{n,K} \to K$ eine Determinantenfunktion und $A = (a_{ij}) \in \mathcal{M}_{n,K}$. Dann gilt

$$d(A) \quad = \quad \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} \ldots a_{n\sigma(n)}.$$

Nach diesen Vorbereitungen können wir nun endlich definieren

Definition 10.14 Sei $A = (a_{ij})_{n \in \mathbb{N}} \in \mathcal{M}_{n,K}$. Dann nennen wir

$$\det(A) \quad = \quad \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{1\sigma(1)} \ldots a_{n\sigma(n)} \qquad \text{die } \textbf{\textit{Determinante}} \text{ der Matrix } A.$$

Der Summenausdruck in der Definition von $\det(A)$ wird auch die ***Leibniz-Formel*** für die Determinante genannt. Wir geben die Formel für die Werte $n = 1, 2, 3$ noch einmal explizit an. Für $n = 1$ ist $S_n = \{\text{id}\}$. In diesem Fall besteht die Summe also nur aus einem einzigen Term. Es gilt $\det(A) = \text{sgn}(\text{id}) a_{1\text{id}(1)} = a_{11}$.

Im Fall $n = 2$ gilt $S_2 = \{\text{id}, (1\ 2)\}$, und die Transposition $\tau = (1\ 2)$ hat nach Satz (10.11) ein negatives Signum. Damit erhalten wir $\det(A) = \text{sgn}(\text{id}) a_{1\text{id}(1)} a_{2\text{id}(2)} + \text{sgn}(\tau) a_{1\tau(1)} a_{2\tau(2)} = a_{11} a_{22} - a_{21} a_{12}$.

Für $n = 3$ besteht S_3 bereits aus $3! = 6$ Elementen, es gilt $S_3 = \{\text{id}, \sigma_1, \sigma_2, \tau_1, \tau_2, \tau_3\}$ mit den Elementen

$$\sigma_1 = (2\ 3) \circ (1\ 3) \ , \ \sigma_2 = (1\ 3) \circ (2\ 3) \ , \ \tau_1 = (1\ 3) \ , \ \tau_2 = (2\ 3) \quad \text{und} \quad \tau_3 = (1\ 2).$$

Zum Beweis genügt es zu überprüfen, dass diese sechs Elemente tatsächlich verschiedene Elemente von S_3 sind. Nach Satz (10.11) gilt $\text{sgn}(\text{id}) = \text{sgn}(\sigma_1) = \text{sgn}(\sigma_2) = 1$ und $\text{sgn}(\tau_1) = \text{sgn}(\tau_2) = \text{sgn}(\tau_3) = -1$.

Mit den sechs Elemente der Gruppe S_3 können wir nun die Formel für die Determinante im Fall $n = 3$ aufstellen. Es gilt

$$\det(A) = a_{1\mathrm{id}(1)}a_{2\mathrm{id}(2)}a_{3\mathrm{id}(3)} + a_{1\sigma_1(1)}a_{2\sigma_1(2)}a_{3\sigma_1(3)} + a_{1\sigma_2(1)}a_{2\sigma_2(2)}a_{3\sigma_2(3)}$$
$$-a_{1\tau_1(1)}a_{2\tau_1(2)}a_{3\tau_1(3)} - a_{1\tau_2(1)}a_{2\tau_2(2)}a_{3\tau_2(3)} - a_{1\tau_3(1)}a_{2\tau_3(2)}a_{3\tau_3(3)} =$$
$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Man bezeichnet diese Formel auch als ***Sarrus-Regel*** zur Berechnung der Determinante. Die drei positiven Summanden entsprechen im Zahlenschema

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}$$

den drei nebeneinanderliegenden Diagonalen von links oben nach rechts unten. Die drei negativen Summanden entsprechen den drei Diagonalen, die von links unten nach rechts oben verlaufen.

Zu beachten ist, dass ein Analogon der Sarrus-Regel für $n = 4$ ***nicht*** gültig ist. Die Leibniz-Formel für eine Matrix $A \in \mathscr{M}_{4,K}$ besteht aus $4! = 24$ Summanden, wärend in der Sarrus-Regel nur acht Terme vorkommen würden.

Nun beweisen wir noch, dass durch die Leibniz-Formel tatsächlich eine Determinantenfunktion definiert ist. Dazu benötigen wir weitere Grundlagen über die symmetrische Gruppe S_n. Für jedes $n \in \mathbb{N}$ bilden die Permutationen mit positivem Signum die sogenannte ***alternierende Gruppe***

$$A_n = \{\sigma \in S_n \mid \mathrm{sgn}(\sigma) = 1\}.$$

Die Gruppe S_n setzt sich zu gleichen Teilen aus Elementen mit positivem und negativem Signum zusammen. Genauer gilt

Proposition 10.15 Sei $A_n = \{\sigma \in S_n \mid \mathrm{sgn}(\sigma) = 1\}$ und $\tau \in S_n$ ein beliebiges Element mit $\mathrm{sgn}(\tau) = -1$. Dann erhält ist durch $S_n = A_n \cup (A_n \circ \tau)$ mit $A_n \circ \tau = \{\sigma \circ \tau \mid \sigma \in A_n\}$ eine Darstellung von S_n als disjunkte Vereinigung gegeben. Zwischen A_n und $A_n \circ \tau$ ist durch $\sigma \mapsto \sigma \circ \tau$ eine Bijektion definiert.

Beweis: Zunächst beweisen wir die Gleichung $S_n = A_n \cup A_n \circ \tau$. Die Inklusion „$\supseteq$“ ist offensichtlich, denn beide Mengen rechts bestehen aus Elementen von S_n. Zum Nachweis von „$\subseteq$“ sei $\sigma \in S_n$ vorgegeben. Liegt σ in A_n, dann ist nichts zu zeigen. Im Fall $\mathrm{sgn}(\sigma) = -1$ liegt $\sigma \circ \tau^{-1}$ in A_n, denn es gilt $\mathrm{sgn}(\sigma \circ \tau^{-1}) = \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)^{-1} = (-1)(-1)^{-1} = 1$. Also ist $\sigma = (\sigma \circ \tau^{-1}) \circ \tau$ in $A_n \circ \tau$ enthalten.

Die Elemente in A_n haben positives Signum. Jedes Element in $A_n \circ \tau$ der Form $\sigma \circ \tau$ mit $\sigma \in A_n$ hat wegen $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) = 1 \cdot (-1) = -1$ negatives Signum. Dies zeigt, dass die Mengen A_n und $A_n \circ \sigma$ disjunkt sind.

Die Abbildung $\sigma \mapsto \sigma \circ \tau$ zwischen A_n und $A_n \circ \sigma$ ist surjektiv, denn jedes Element in $A_n \circ \tau$ kann in der Form $\sigma \circ \tau$ mit $\sigma \in A_n$ geschrieben werden. Sind $\sigma_1, \sigma_2 \in A_n$ mit $\sigma_1 \circ \tau = \sigma_2 \circ \tau$, dann folgt $\sigma_1 \circ \tau \circ \tau^{-1} = \sigma_2 \circ \tau \circ \tau^{-1}$ und damit $\sigma_1 = \sigma_2$. Damit ist auch die Injektivität der Abbildung nachgewiesen. □

Satz 10.16 Für jedes $n \in \mathbb{N}$ ist $\det : \mathcal{M}_{n,K} \to K$ eine Determinantenfunktion.

Beweis: Wir müssen überprüfen, dass die Abbildung det die drei Bedingungen aus Def. (10.1) erfüllt. Sei $k \in \{1, ..., n\}$, und seien $A, B \in \mathcal{M}_{n,K}$ zwei Matrizen, die in allen Zeilen mit Ausnahme der k-ten übereinstimmen. Sei außerdem $C \in \mathcal{M}_{n,K}$ die Matrix mit $c_{k\bullet} = a_{k\bullet} + b_{k\bullet}$ und $c_{\ell\bullet} = a_{\ell\bullet} = b_{\ell\bullet}$ für $\ell \neq k$. Zu zeigen ist $\det(C) = \det(A) + \det(B)$. Tatsächlich gilt

$$\det(C) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} c_{\ell\sigma(\ell)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) c_{k\sigma(k)} \prod_{\ell \neq k} c_{\ell\sigma(\ell)} =$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(a_{k\sigma(k)} + b_{k\sigma(k)}) \prod_{\ell \neq k} c_{\ell\sigma(\ell)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{k\sigma(k)} \prod_{\ell \neq k} c_{\ell\sigma(\ell)} + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{k\sigma(k)} \prod_{\ell \neq k} c_{\ell\sigma(\ell)} =$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{k\sigma(k)} \prod_{\ell \neq k} a_{\ell\sigma(\ell)} + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{k\sigma(k)} \prod_{\ell \neq k} b_{\ell\sigma(\ell)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} a_{\ell\sigma(\ell)} + \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} b_{\ell\sigma(\ell)}$$

$$= \det(A) + \det(B).$$

Sei nun $\lambda \in K$ und $D \in \mathcal{M}_{n,K}$ die Matrix gegeben durch $d_{k\bullet} = \lambda a_{k\bullet}$ und $d_{\ell\bullet} = a_{\ell\bullet}$ für $\ell \neq k$. Dann gilt

$$\det(D) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} d_{\ell\sigma(\ell)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) d_{k\sigma(k)} \prod_{\ell \neq k} d_{\ell\sigma(\ell)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)(\lambda a_{k\sigma(k)}) \prod_{\ell \neq k} a_{\ell\sigma(\ell)}$$

$$= \lambda \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{k\sigma(k)} \prod_{\ell \neq k} a_{\ell\sigma(\ell)} = \lambda \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{\ell=1}^{n} a_{\ell\sigma(\ell)} = \lambda \det(A).$$

Damit ist die Eigenschaft (i) aus Def. (10.1) verifiziert. Zum Nachweis von (ii) sei $A \in \mathcal{M}_{n,K}$ eine Matrix, in der die k-te und ℓ-te Zeile übereinstimmen, wobei $k \neq \ell$ sei. Für die Transposition $\tau = (k\ \ell)$ gilt $S_n = A_n \cup A_n \circ \tau$ nach Prop. (10.15) . Weil die Elemente in A_n positives und die Elemente in $A_n \circ \tau$ negatives Signum haben, erhalten wir für $\det(A)$ den Ausdruck

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma \in A_n} \prod_{i=1}^{n} a_{i\sigma(i)} - \sum_{\sigma \in A_n \circ \tau} \prod_{i=1}^{n} a_{i\sigma(i)} = \sum_{\sigma \in A_n} \prod_{i=1}^{n} a_{i\sigma(i)} - \sum_{\sigma \in A_n} \prod_{i=1}^{n} a_{i\sigma(\tau(i))}$$

Weil die k-te und die ℓ-te Zeile von A übereinstimmen, gilt für die einzelnen Terme der rechten Summe

$$\prod_{i=1}^{n} a_{i\sigma(\tau(i))} = a_{k\sigma(\tau(k))} a_{\ell\sigma(\tau(\ell))} \prod_{i \neq k,\ell} a_{i\sigma(\tau(i))} = a_{k\sigma(\ell)} a_{\ell\sigma(k)} \prod_{i \neq k,\ell} a_{i\sigma(\tau(i))}$$

$$= a_{\ell\sigma(\ell)} a_{k\sigma(k)} \prod_{i \neq k,\ell} a_{i\sigma(\tau(i))} = \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Also heben sich die beiden Summen auf, und es folgt $\det(A) = 0$.

Nun beweisen wir noch die Eigenschaft (iii). Die Determinante der Einheitsmatrix ist nach Definition gegeben durch

$$\det(\mathrm{I}^{(n)}) \quad = \quad \sum_{\sigma \in S_n} s_\sigma \quad , \qquad s_\sigma = \operatorname{sgn}(\sigma)\delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)}.$$

Gilt $\sigma(i) \neq i$ für ein i, dann folgt $\delta_{i\sigma(i)} = 0$ und somit $s_\sigma = 0$. Also ist s_{id} der einzige nicht-verschwindende Summand in der Leibniz-Formel, und dieser ist gleich 1. □

In der folgenden Situation lässt sich die Determinante einer Matrix sehr leicht ausrechnen.

Satz 10.17 Man bezeichnet eine Matrix $A = (a_{ij}) \in \mathcal{M}_{n,K}$ als ***obere Dreiecksmatrix***, wenn $a_{ij} = 0$ für alle $i, j \in \{1, ..., n\}$ mit $i > j$ erfüllt ist. Für jede Matrix dieser Form gilt

$$\det(A) \quad = \quad a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}.$$

Beweis: Wir zeigen, dass in der Leibniz-Formel die Summanden zu den Permutationen $\sigma \neq \mathrm{id}$ alle verschwinden. Das Produkt $a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$ ist dann der verbleibende Summand, der zur Permutation $\sigma = \mathrm{id}$ gehört. Sei also $\sigma \in S_n$ mit $\sigma \neq \mathrm{id}$ vorgegeben; zu zeigen ist, dass der Term $s_\sigma = \operatorname{sgn}(\sigma) \prod_{k=1}^n a_{k\sigma(k)}$ gleich Null ist. Gibt es ein $k \in \{1, ..., n\}$ mit $\sigma(k) < k$, dann gilt auf Grund unserer Voraussetzung $a_{k\sigma(k)} = 0_K$ und damit auch $s_\sigma = 0_K$.

Setzen wir nun $\sigma(k) \geq k$ für $1 \leq k \leq n$ voraus. Wir zeigen durch vollständige Induktion, dass dann $\sigma(n-r) = n-r$ für $0 \leq r \leq n-1$ gilt. Aus $\sigma(n) \geq n$ folgt jedenfalls $\sigma(n) = n$, also ist die Aussage für $r = 0$ richtig. Sei nun $r \geq 1$ vorgegeben, und setzen wir nun $\sigma(n-k) = n-k$ für $0 \leq k < r$ voraus. Wäre $\sigma(n-r) > n-r$, dann wäre die Abbildung nicht injektiv, denn dann würde $\sigma(n-r) = n-k$ für ein $k < r$ gelten, und wir würden auf Grund der Induktionsannahme

$$\sigma(n-r) \quad = \quad n-k \quad = \quad \sigma(n-k) \qquad \text{erhalten.}$$

Somit ist auch hier $\sigma(n-r) = n-r$ die einzige Möglichkeit. Damit ist der Induktionsbeweis abgeschlossen; er zeigt, dass die einzige Permutation mit $\sigma(k) \geq k$ für $1 \leq k \leq n$ die Identität ist. Für jede Permutation $\sigma \in S_n$ mit $\sigma \neq \mathrm{id}$ gibt es also ein $k \in \{1, ..., n\}$ mit $\sigma(k) < k$. Wie oben gezeigt wurde, folgt daraus jeweils $s_\sigma = 0_K$. □

§ 11. Rechenregeln für Determinanten

Zunächst untersuchen wir in diesem Abschnitt, wie sich Zeilenumformungen auf die Determinante einer Matrix auswirken. In Kapitel §2 haben wir die Elementarmatrizen

$$\mathrm{M}_{k,\lambda} = \mathrm{I}^{(m)} + (\lambda - 1)\mathrm{B}_{kk}^{(m\times m)} \qquad \text{und} \qquad \mathrm{A}_{k,\ell,\lambda} = \mathrm{I}^{(m)} + \lambda \mathrm{B}_{\ell k}^{(m\times m)}$$

eingeführt. Die Matrix $\mathrm{M}_{k,\lambda}$ entsteht aus der Einheitsmatrix $\mathrm{I}^{(n)}$ durch Multiplikation der k-ten Zeilen mit dem Wert λ. Auf Grund der Multilinearität der Determinantenfunktion gilt deshalb $\det(\mathrm{M}_{k,\lambda}) = \lambda \det \mathrm{I}^{(n)} = \lambda \cdot 1_K = \lambda$. Ist $A \in \mathscr{M}_{n,K}$ eine beliebige Matrix, dann entsteht $\mathrm{M}_{k,\lambda}A$ aus A durch Multiplikation der k-ten Zeile mit dem Wert λ. Wir erhalten somit die Rechenregel

$$\det(\mathrm{M}_{k,\lambda}A) \quad = \quad \lambda \det A \quad = \quad \det(\mathrm{M}_{k,\lambda})\det(A) \qquad \text{für alle} \quad A \in \mathscr{M}_{n,K}.$$

Allgemein gilt: Entsteht eine Matrix $B = (b_{ij})$ aus $A = (a_{ij}) \in \mathscr{M}_{n,K}$ durch Addition des λ-fachen der k-ten Zeile zur ℓ-ten, dann gilt $\det(A) = \det(A')$, denn aus den Eigenschaften der Determinantenfunktion folgt

$$\det B \quad = \quad \det(..., b_{k\bullet}, ..., b_{\ell\bullet}, ...) \quad = \quad \det(..., a_{k\bullet}, ..., \lambda a_{k\bullet} + a_{\ell\bullet}, ...) \quad =$$

$$\lambda \det(..., a_{k\bullet}, ..., a_{k\bullet}, ...) + \det(..., a_{k\bullet}, ..., a_{\ell\bullet}, ...) \quad = \quad \lambda \cdot 0 + \det A \quad = \quad \det A.$$

Insbesondere entsteht die Matrix $\mathrm{A}_{k,\ell,\lambda}$ aus $\mathrm{I}^{(n)}$ durch Addition des λ-fachen der k-ten Zeile zur ℓ-ten. Somit gilt $\det \mathrm{A}_{k,\ell,\lambda} = \mathrm{I}^{(n)} = 1$ und allgemein

$$\det(\mathrm{A}_{k,\ell,\lambda}A) \quad = \quad \det(\mathrm{A}_{k,\ell,\lambda})\det(A) \qquad \text{für alle} \quad A \in \mathscr{M}_{n,K}.$$

Zusammenfassend kann also formuliert werden

Lemma 11.1 Ist $T \in \mathscr{M}_{n,K}$ eine Elementarmatrix und $A \in \mathscr{M}_{n,K}$ beliebig, dann gilt

$$\det(TA) \quad = \quad \det(T)\det(A).$$

Eine quadratische Matrix in Zeilenstufenform ist insbesondere eine obere Dreiecksmatrix, und deren Determinante erhält man nach Satz (10.17) aus dem Produkt der Diagonalelemente. Wir haben gesehen, wie sich elementare Zeilenumformungen vom Typ $(\mathrm{M}_{k,\lambda})$ und $(\mathrm{A}_{k,\ell,\lambda})$ auf die Determinante auswirken. Außerdem führt die Vertauschung zweier Zeilen nach Lemma (10.4) lediglich zu einem Vorzeichenwechsel bei der Determinante. Dies zusammen liefert uns folgende Strategie für die Berechnung von $\det A$ für eine beliebige Matrix $A \in \mathscr{M}_{n,K}$.

(i) Forme A mit Hilfe des Gaußverfahrens in eine Matrix B in Zeilenstufenform um.
(ii) Bestimme anhand der durchgeführten Zeilenumformungen den Faktor $\mu \in K^\times$ mit $\det(B) = \mu \det(A)$.
(iii) Berechne $\det(B)$ durch Multiplikation der Diagonalelemente $b_{11},...,b_{nn}$ von B.
Das Endergebnis der Rechnung ist dann $\det(A) = \mu^{-1}\det(B)$.

Neben diesem praktischen Rechenverfahren führen unsere Überlegungen auch zu neuen theoretischen Resultaten.

Satz 11.2 *(Charakterisierung invertierbarer Matrizen)*

Für eine Matrix $A \in \mathscr{M}_{n,K}$ sind folgende Aussagen äquivalent.

(i) $A \in \mathrm{GL}_n(K)$ (d.h. die Matrix A ist invertierbar)
(ii) $\mathrm{rg}(A) = n$
(iii) $\det(A) \neq 0$

Beweis: Zunächst beweisen wir die Äquivalenz „(i) ⇔ (ii)". Sei $\phi_A : K^n \to K^n$ gegeben durch $v \mapsto Av$. Nach dem Dimensionssatz für lineare Abbildungen gilt $n = \dim\ker(\phi_A) + \dim\mathrm{im}(\phi_A)$. Aus §8 wissen wir auch, dass $\mathrm{rg}(A) = \dim\mathrm{im}(\phi_A)$ gilt. Also ist $\mathrm{rg}(A) = n$ genau dann erfüllt, wenn

$$\dim\ker(\phi_A) = 0 \quad \text{und} \quad \dim\mathrm{im}(\phi_A) = n$$

erfüllt gilt. Dies wiederum ist genau dann der Fall, wenn ϕ_A bijektiv ist. Nach Satz (9.9), angewendet auf die kanonische Basis $\mathscr{E}$ von K^n, ist die Bijektivität von ϕ_A wiederum äquivalent zur Invertierbarkeit von A.

An Stelle der Äquivalenz „(ii) ⇔ (iii)" beweisen wir, dass genau dann $\det A = 0$ ist, wenn $\mathrm{rg}(A) < n$ gilt. Weil sich die Determinante der Matrix bei Zeilenumformungen höchstens um einen Faktor $\lambda \in K^\times$ ändert, können wir voraussetzen, dass sich die Matrix A in Zeilenstufenform befindet. Der Zeilenrang von A ändert sich durch diese Umformungen nicht. Seien $j_1 < ... < j_r$ die Kennzahlen der Zeilenstufenform mit $r = \mathrm{rg} A$. Ist $r < n$, dann ist $a_{nn} = 0$, und somit gilt auch $\det(A) = \prod_{i=1}^n a_{ii} = 0$. Ist dagegen $r = n$, dann muss $j_i = i$ für $1 \leq i \leq n$ gelten, und die Einträge $a_{ij_i} = a_{ii}$ sind nach Definition der Zeilenstufenform ungleich Null. Also ist auch $\det A$ als Produkt der Diagonaleinträge ungleich Null. □

Satz 11.3 *(Multiplikationssatz für Determinanten)*

Seien $A, B \in \mathscr{M}_{n,K}$. Dann gilt $\det(AB) = \det(A)\det(B)$.

Beweis: Wie immer bezeichnen wir mit ϕ_A bzw. ϕ_B die linearen Abbildungen $K^n \to K^n$ gegeben durch $v \mapsto Av$ bzw. $v \mapsto Bv$. Nehmen wir zunächst an, dass $\det(A) = 0$ ist. Dann folgt $\dim\mathrm{im}(\phi_A) = \mathrm{rg}(A) < n$ und somit auch

$$\mathrm{rg}(AB) \quad = \quad \dim\mathrm{im}(\phi_A \circ \phi_B) \quad \leq \quad \dim\phi_A(K^n) \quad < \quad n.$$

Dies wiederum bedeutet $\det(AB) = 0$, d.h. in diesem Fall ist die Gleichung $\det(AB) = \det(A)\det(B)$ erfüllt. Nun betrachten wir den Fall, dass A invertierbar (also $\det A \neq 0$) ist. Für den Fall $A = \mathrm{I}^{(n)}$ ist die Aussage trivial, und im Fall, dass es sich bei A um eine Elementarmatrix handelt, ist sie nach Lemma (11.1) erfüllt. Im allgemeinen Fall ist aus §3 bekannt, dass A als Produkt $T_1 \cdot ... \cdot T_r$ von Elementarmatrizen darstellbar ist. Wir beweisen nun die Gleichung $\det(AB) = \det(A)\det(B)$ durch vollständige Induktion über r. Der Fall $r = 1$ ist bereits erledigt, denn in diesem Fall ist A selbst eine Elementarmatrix. Ist nun $r > 1$ und setzen wir die Gleichung für Werte $< r$ voraus, dann können wir $A' = T_1 \cdot \ldots \cdot T_{r-1}$ setzen und erhalten

$$\det(AB) \quad = \quad \det(A'T_rB) \quad = \quad \det(A')\det(T_rB) \quad =$$
$$\det(A')\det(T_r)\det(B) \quad = \quad \det(A'T_r)(\det B) \quad = \quad \det(A)\det(B). \qquad \square$$

Gelegentlich ist auch die folgende Rechenregel nützlich.

Satz 11.4 Sei $M \in \mathcal{M}_{n,K}$ eine Blockmatrix der Form

$$M \quad = \quad \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

mit $A \in \mathcal{M}_{r,K}$, $B \in \mathcal{M}_{r\times(n-r),K}$ und $C \in \mathcal{M}_{n-r,K}$. Dann gilt $\det M = (\det A)(\det C)$.

Beweis: Aus §3 ist bekannt, dass es Elementarmatrizen $T_1, \ldots, T_k \in \mathrm{GL}_r(K)$ und $U_1, \ldots, U_\ell \in \mathrm{GL}_n(K)$ gibt, so dass $A' = T_k \cdots T_1 A$ und $C' = U_\ell \cdots U_1 C$ in Zeilenstufenform vorliegen, also insbesondere obere Dreiecksmatrizen sind. Für jede Matrix $B' \in \mathcal{M}_{r\times(n-r),K}$ ist dann auch die Blockmatrix

$$M' \quad = \quad \begin{pmatrix} A' & B' \\ 0 & C' \end{pmatrix}$$

eine obere Dreiecksmatrix, und es gilt

$$\det M' \quad = \quad \left(\prod_{i=1}^{r} a'_{ii}\right)\left(\prod_{j=1}^{n} c'_{jj}\right) \quad = \quad \det(A')\det(C').$$

Man überprüft unmittelbar, dass mit T_i und U_j auch

$$\hat{T}_i = \begin{pmatrix} T_i & 0 \\ 0 & \mathrm{I}^{(n-r)} \end{pmatrix} \qquad \text{und} \qquad \hat{U}_j = \begin{pmatrix} \mathrm{I}^{(r)} & 0 \\ 0 & U_j \end{pmatrix}$$

Elementarmatrizen sind, für die $\det T_i = \det \hat{T}_i$ sowie $\det U_j = \det \hat{U}_j$ gilt. Auf Grund der Rechenregel für Produkte von Blockmatrizen (Abschnitt §2, Seite 15) gilt

$$\hat{T}_k \cdots \hat{T}_1 M \quad = \quad \begin{pmatrix} T_k & 0 \\ 0 & \mathrm{I}^{(n-r)} \end{pmatrix} \cdots \begin{pmatrix} T_1 & 0 \\ 0 & \mathrm{I}^{(n-r)} \end{pmatrix}\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} T_k \cdots T_1 A & T_k \cdots T_1 B \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} A' & B' \\ 0 & C \end{pmatrix}$$

mit $B' = T_k \cdots T_1 B$ und ebenso

$$\hat{U}_\ell \cdots \hat{U}_1 \begin{pmatrix} A' & B' \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} \mathrm{I}^{(r)} & 0 \\ 0 & U_\ell \end{pmatrix} \cdots \begin{pmatrix} \mathrm{I}^{(r)} & 0 \\ 0 & U_1 \end{pmatrix}\begin{pmatrix} A' & B' \\ 0 & C \end{pmatrix} \quad = \quad \begin{pmatrix} A' & B' \\ 0 & U_\ell \cdots U_1 C \end{pmatrix} \quad = \quad \begin{pmatrix} A' & B' \\ 0 & C' \end{pmatrix}$$

Damit erhalten wir insgesamt

$$\left(\prod_{i=1}^{k} \det T_i\right)\left(\prod_{j=1}^{\ell} \det U_j\right)\det(A)\det(C) \quad = \quad \det(A')\det(C') \quad = \quad \det M' \quad =$$

$$\left(\prod_{i=1}^{k} \det \hat{T}_i\right)\left(\prod_{j=1}^{\ell} \det \hat{U}_j\right)\det(M) \quad = \quad \left(\prod_{i=1}^{k} \det T_i\right)\left(\prod_{j=1}^{\ell} \det U_j\right)\det(M)$$

und somit $\det(A)\det(C) = \det(M)$. □

Bevor wir als letzte wichtige Rechenregel den Laplaceschen Entwicklungssatz in Angriff nehmen, bemerken wir zunächst, dass sich die Determinate beim Übergang zur transponierten Matrix nicht ändert.

Satz 11.5 Für alle $A \in \mathcal{M}_{n,K}$ gilt $\det A = \det {}^{\mathrm{t}}A$.

Beweis: Hier verwenden wir zum Beweis die Leibnizformel. Setzen wir $B = {}^{\mathrm{t}}A$, dann sind die Einträge b_{ij} von B gegeben durch $b_{ij} = a_{ji}$ für $1 \le i, j \le n$. Es gilt

$$\det({}^{\mathrm{t}}A) \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} b_{i\sigma(i)} \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i} \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)\sigma^{-1}(\sigma(i))}$$

Weil jedes σ bijektiv ist, durchläuft mit i auch $\sigma(i)$ alle Zahlen in $\{1,\ldots,n\}$. Wir können im Produkt also $\sigma(i)$ durch i ersetzen und erhalten

$$\prod_{i=1}^{n} a_{\sigma(i)\sigma^{-1}(\sigma(i))} \quad = \quad \prod_{i=1}^{n} a_{i\sigma^{-1}(i)}.$$

Nach Folgerung (10.10) gilt $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$ für alle $\sigma \in S_n$. Somit gilt insgesamt

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)\sigma^{-1}(\sigma(i))} \quad = \quad \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma^{-1}) \prod_{i=1}^{n} a_{i\sigma^{-1}(i)} \quad =$$

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \quad = \quad \det(A).$$

Beim vorletzten „=“ haben wir verwendet, dass mit σ auch σ^{-1} die gesamte Gruppe S_n durchläuft, so dass wir in jedem Summanden jeweils σ^{-1} durch σ ersetzen können. □

Bevor wird den Laplaceschen Entwicklungssatz formulieren und beweisen können, benötigen wir ein wenig zusätzliche Notation. Sei $A = (a_{ij}) \in \mathcal{M}_{n,K}$. Für beliebige $i, j \in \{1,\ldots,n\}$ sei dann $A'_{ij} = (a_{k\ell})$ die Matrix in $\mathcal{M}_{n,K}$ mit den Einträgen

$$a_{k\ell} \quad = \quad \begin{cases} a_{k\ell} & \text{für } k \neq i, \ell \neq j \\ 0 & \text{für } k = i, \ell \neq j \\ 0 & \text{für } k \neq i, \ell = j \\ 1 & \text{für } k = i, \ell = j \end{cases}$$

Die Matrix A'_{ij} hat also die Form

$$A'_{ij} \quad = \quad \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & 0 & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

Mit $A_{ij} \in \mathscr{M}_{n-1,K}$ bezeichnen wir die Matrix, die aus A durch Streichung der i-ten Zeile und j-ten Spalten zu Stande kommt, also

$$A_{ij} \quad = \quad \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

Lemma 11.6 Es gilt $\det A'_{ij} = (-1)^{i+j} \det A_{ij}$.

Beweis: Durch $i-1$ Zeilenvertauschungen bewegt man die i-te Zeile von A'_{ij} nach oben in die erste Zeile. Anschließend führt man $j-1$ Spaltenvertauschungen durch, um die j-te Spalte ganz nach links zu bewegen. Insgesamt ändert sich das Vorzeichen dadurch um den Faktor $(-1)^{(i-1)+(j-1)} = (-1)^{i+j}$, und man erhält eine Matrix der Form

$$\begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix}$$

Auf Grund der Formel in Satz (11.4) über die Determinante von Blockmatrizen stimmt die Determinante dieser Matrix mit $\det(A_{ij})$ überein. □

Definition 11.7 Sei $A \in \mathscr{M}_{n,K}$. Die Matrix $\tilde{A} \in \mathscr{M}_{n,K}$ mit den Einträgen

$$\tilde{a}_{ij} \quad = \quad \det(A'_{ji}) \quad = \quad (-1)^{i+j} \det(A_{ji})$$

wird die zu A ***komplementäre*** oder ***adjunkte*** Matrix genannt. (Man beachte die Vertauschung von Zeilen- und Spaltenindex, dies ist kein Tippfehler!)

Lemma 11.8 Sei $A \in \mathscr{M}_{n,K}$. Dann gilt für $1 \le i, j \le n$ jeweils

$$\det A'_{ij} = \det(a_{1\bullet}, ..., a_{i-1\bullet}, e_j, a_{i+1\bullet}, ..., a_{n\bullet})$$

Beweis: Sei B die Matrix auf der rechten Seite der Gleichung. Dann kann B in A'_{ij} umgeformt werden, indem man für alle k mit $1 \le k \le n$ und $i \ne k$ jeweils das a_{kj}-fache der i-ten Zeile von der k-ten Zeile von B subtrahiert. Dies zeigt, dass die beiden Determinanten übereinstimmen.

Proposition 11.9 Sei $\tilde{A}$ die zu $A \in \mathscr{M}_{n,K}$ komplementäre Matrix. Dann gilt

$$\tilde{A}A \quad = \quad A\tilde{A} \quad = \quad (\det A)\mathrm{I}^{(n)}.$$

Beweis: Zunächst zeigen wir die Gleichung $A\tilde{A} = (\det A)\mathrm{I}^{(n)}$. Für den Eintrag von $A\tilde{A}$ an der Stelle (i,k) gilt nach Lemma (11.8) und der Rechenregen für die Determiante jeweils

$$\sum_{j=1}^{n} a_{ij}\tilde{a}_{jk} \quad = \quad \sum_{j=1}^{n} a_{ij} \det(A'_{kj}) \quad = \quad \sum_{j=1}^{n} a_{ij} \det(a_{1\bullet}, ..., a_{k-1\bullet}, e_j, a_{k+1\bullet}, ..., a_{n\bullet}) \quad =$$

$$\det\left(a_{1\bullet}, ..., a_{k-1\bullet}, \sum_{j=1}^{n} a_{ij}e_j, a_{k+1\bullet}, ..., a_{n\bullet}\right) \quad = \quad \det(a_{1\bullet}, ..., a_{k-1\bullet}, a_{i\bullet}, a_{k+1\bullet}, ..., a_{n\bullet}) \quad = \quad \delta_{ik} \det(A).$$

Der Eintrag von $A\tilde{A}$ an der Stelle (i,k) stimt also mit dem entsprechenden Eintrag von $(\det A)\mathrm{I}^{(n)}$ überein. Für den Beweis der zweiten Gleichung bemerken wir zunächst, dass ${}^{\mathrm{t}}\tilde{A}$ nach Definition die zu ${}^{\mathrm{t}}A$ komplementäre Matrix ist. Auf Grund der bereits bewiesenen Gleichung gilt also

$$\tilde{A}A \quad = \quad {}^{\mathrm{t}}({}^{\mathrm{t}}A\,{}^{\mathrm{t}}\tilde{A}) \quad = \quad {}^{\mathrm{t}}((\det A)\mathrm{I}^{(n)}) \quad = \quad (\det A)\mathrm{I}^{(n)}. \qquad \square$$

Satz 11.10 *(Laplacescher Entwicklungssatz)*

Sei $A \in \mathscr{M}_{n,K}$.

(i) Für alle $i \in \{1, ..., n\}$ gilt $\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$.

(ii) Für alle $j \in \{1, ..., n\}$ gilt $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$.

Wird die Determinante von A mittels (i) berechnet, spricht man von einer *Entwicklung zur i-ten Zeile*. Die Berechnung mittels (ii) bezeichnet man als *Entwicklung zur j-ten Spalte*.

Beweis: Auf Grund der Proposition ist der Eintrag von $A\tilde{A}$ an der Stelle (i,i) gleich $\det A$. Es gilt also

$$\det(A) \quad = \quad \sum_{j=1}^{n} a_{ij}\tilde{a}_{ji} \quad = \quad \sum_{j=1}^{n} a_{ij}\det(A'_{ij}) \quad = \quad \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}).$$

Auch der Eintrag von $\tilde{A}A$ an der Stelle (j,j) ist gleich $\det(A)$. Folglich gilt

$$\det(A) \quad = \quad \sum_{i=1}^{n} \tilde{a}_{ji}a_{ij} \quad = \quad \sum_{i=1}^{n} a_{ij}\det(A'_{ij}) \quad = \quad \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}). \qquad \square$$

§ 12. Eigenwerte und Eigenvektoren

Im gesamten Text bezeichnet K stets einen beliebigen Körper, solange nichts genaueres festgelegt wird.

Definition 12.1 Sei V ein K-Vektorraum und $\phi : V \to V$ ein Endomorphismus von V.

(i) Ein Element $\lambda \in K$ heißt ***Eigenwert*** von ϕ, wenn es ein $v \in V$ mit $v \neq 0_V$ und $\phi(v) = \lambda v$ gibt.

(ii) Ein Vektor $v \in V$ heißt ***Eigenvektor*** von ϕ, wenn $v \neq 0_V$ ist und ein $\lambda \in K$ mit $\phi(v) = \lambda v$ existiert.

Seien nun $v \in V$ und $\lambda \in K$ vorgegeben. Man nennt v einen Eigenvektor zum Eigenwert λ, wenn $v \neq 0_V$ und die Gleichung $\phi(v) = \lambda v$ erfüllt ist.

Definition 12.2 Sei V ein K-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Für jedes $\lambda \in K$ bezeichnet man die Menge $\mathrm{Eig}(\phi, \lambda) = \{v \in V \mid \phi(v) = \lambda v\}$ als den ***Eigenraum*** von ϕ zum Wert $\lambda \in K$. Er besteht aus dem Nullvektor 0_V und den Eigenvektoren zum Eigenwert λ.

Proposition 12.3 Sei V ein K-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Für jedes $\lambda \in K$ ist der Eigenraum gegeben durch $\mathrm{Eig}(\phi, \lambda) = \ker(\phi - \lambda \mathrm{id}_V)$. Das Element λ ist ein Eigenwert von ϕ genau dann, wenn $\mathrm{Eig}(\phi, \lambda) \neq \{0_V\}$ gilt.

Beweis: Für jeden Vektor $v \in V$ gilt die Äquivalenz

$$v \in \mathrm{Eig}(\phi, \lambda) \quad \Leftrightarrow \quad \phi(v) = \lambda v \quad \Leftrightarrow \quad \phi(v) - \lambda v = 0_V \quad \Leftrightarrow \quad \phi(v) - \lambda \mathrm{id}_V(v) = 0_V$$
$$\Leftrightarrow \quad (\phi - \lambda \mathrm{id}_V)(v) = 0_V \quad \Leftrightarrow \quad v \in \ker(\phi - \lambda \mathrm{id}_V).$$

Daraus folgt $\mathrm{Eig}(\phi, \lambda) = \ker(\phi - \lambda \mathrm{id}_V)$. Ein Element $\lambda \in K$ ist nach Definition Eigenwert genau dann, wenn ein $v \in V$ mit $v \neq 0_V$ und $\phi(v) = \lambda v$ existiert, also genau dann, wenn es ein Element ungleich 0_V in $\mathrm{Eig}(\phi, \lambda)$ gibt. Weil 0_V auf jeden Fall in $\mathrm{Eig}(\phi, \lambda)$ liegt, ist dies wiederum äquivalent zu $\mathrm{Eig}(\phi, \lambda) \neq \{0_V\}$. □

Aus §5 ist bekannt, dass Kerne von linearen Abbildungen Untervektorräume sind. Die Proposition zeigt also, dass $\mathrm{Eig}(\phi, \lambda)$ für jedes $\lambda \in K$ und jedes $\phi \in \mathrm{End}_K(V)$ ein Untervektorraum von V ist. Natürlich kann diese Eigenschaft auch direkt nachgerechnet werden.

Als nächstes sehen wir uns an, wie das Matrixkalkül zur Untersuchung von Eigenwerten und Eigenvektoren eingesetzt werden kann. Sei nun $A \in \mathcal{M}_{n,K}$ eine quadratische Matrix. Wir bezeichnen $v \in K^n$ als einen ***Eigenvektor von*** A, wenn v ein Eigenvektor der Abbildung $\phi_A : K^n \to K^n$, $v \mapsto Av$ ist.

Ebenso sind die ***Eigenwerte von*** A nach Definition die Eigenwerte des Endomorphismus ϕ_A. Für jedes $\lambda \in K$ definieren wir

$$\mathrm{Eig}(A,\lambda) \quad = \quad \mathrm{Eig}(\phi_A,\lambda) \quad = \quad \{v \in K^n \mid Av = \lambda v\}.$$

Wiederum besteht $\mathrm{Eig}(A,\lambda)$ aus den Eigenvektoren von A zum Eigenwert λ und dem Nullvektor 0_{K^n}, und darüber hinaus gilt $\mathrm{Eig}(A,\lambda) = \ker(A - \lambda \mathrm{I}^{(n)})$.

Aus §9 wissen wir, dass jede lineare Abbildung $\phi : V \to W$ zwischen endlich-dimensionalen K-Vektorräumen V,W auf eindeutige Weise durch eine Matrix beschrieben werden kann, sobald man für V und W Basen festgelegt hat. Ist $\mathscr{A}$ eine Basis von V und $\mathscr{B}$ eine Basis von W, dann haben wir die Bezeichung

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$$

für die Darstellungsmatrix von ϕ bezüglich der Basen $\mathscr{A}$ und $\mathscr{B}$ eingeführt. Wir erinnern an den wichtigen Zusammenhang

$$\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)\Phi_{\mathscr{A}}(v) \quad = \quad \Phi_{\mathscr{B}}(\phi(v)).$$

Die Darstellungsmatrix $\mathscr{M}^{\mathscr{A}}_{\mathscr{B}}(\phi)$ ist also dadurch gekennzeichnet, dass sie den Vektor v in $\mathscr{A}$-Koordinaten entgegennimmt und den Vektor $\phi(v)$ in $\mathscr{B}$-Koordinaten als Ergebnis liefert.

Ist nun $V = W$, die Abbildung ϕ also ein ***Endomorphismus*** des Vektorraums V, dann braucht nur noch ***eine*** Basis von V, um ϕ zu beschreiben. Wir setzen $\mathscr{M}_{\mathscr{A}}(\phi) = \mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\phi)$ und nennen diese quadratische Matrix die ***Darstellungsmatrix*** von ϕ bezüglich der Basis $\mathscr{A}$.

Definition 12.4 Zwei Matrizen $A,B \in \mathscr{M}_{n,K}$ werden ***ähnlich*** genannt, wenn eine invertierbare Matrix $T \in \mathrm{GL}_n(K)$ mit $B = TAT^{-1}$ existiert. Zwei Matrizen $A,B \in \mathscr{M}_{m\times n,K}$ bezeichnet man als ***äquivalent*** wenn Elemente $S \in \mathrm{GL}_m(K)$ und $T \in \mathrm{GL}_n(K)$ mit $B = SAT$ existieren.

Ähnliche Matrizen sind also stets äquivalent zueinander. Die Umkehrung ist allerdings im allgemeinen falsch.

Proposition 12.5 Sei V ein n-dimensionaler K-Vektorraum, und sei $\phi \in \mathrm{End}_K(V)$. Sind $A,B \in \mathscr{M}_{n,K}$ Darstellungsmatrizen von ϕ bezüglich unterschiedlicher Basen von V, dann sind A und B ähnlich.

Beweis: Seien $\mathscr{A}$ und $\mathscr{B}$ Basen von V, so dass $A = \mathscr{M}_{\mathscr{A}}(\phi)$ und $B = \mathscr{M}_{\mathscr{B}}(\phi)$ erfüllt ist. Sei außerdem $T = \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$ die Matrix des Basiswechsels von A nach B, also die eindeutig bestimmte Matrix $T \in \mathrm{GL}_n(K)$ mit $T\Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{B}}(v)$ für alle $v \in V$. Auf Grund der Transformationsformel (Satz (9.16)) gilt

$$B \quad = \quad \mathscr{M}_{\mathscr{B}}(\phi) \quad = \quad \mathscr{M}^{\mathscr{B}}_{\mathscr{B}}(\phi) \quad = \quad \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}\mathscr{M}^{\mathscr{A}}_{\mathscr{A}}(\phi)\mathscr{T}^{\mathscr{B}}_{\mathscr{A}} \quad = \quad \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}\mathscr{M}_{\mathscr{A}}(\phi)\left(\mathscr{T}^{\mathscr{A}}_{\mathscr{B}}\right)^{-1} \quad = \quad TAT^{-1}. \qquad \square$$

Proposition 12.6 Sei V ein endlich-dimensionaler K-Vektorraum, $\phi \in \mathrm{End}_K(V)$ und $A \in \mathscr{M}_{n,K}$ die Darstellungsmatrix von ϕ bezüglich einer beliebigen Basis $\mathscr{A}$ von V. Genau dann ist $v \in V$ ein Eigenvektor von ϕ zu einem Eigenwert $\lambda \in K$, wenn der Koordinatenvektor $\Phi_{\mathscr{A}}(v)$ ein Eigenvektor von A zum Eigenwert λ ist.

Beweis: Nach Definition der Darstellungsmatrix gilt $A\Phi_{\mathscr{A}}(v) = \mathscr{M}_{\mathscr{A}}(\phi)\Phi_{\mathscr{A}}(v) = \Phi_{\mathscr{A}}(\phi(v))$ für jedes $v \in V$. Sei nun $\lambda \in K$ vorgegeben. Ein Vektor v ist genau dann Eigenvektor von ϕ zum Eigenwert λ, wenn $v \neq 0_V$ und $\phi(v) = \lambda v$ erfüllt ist. Auf Grund der Bijektivitiät und der Linearität der Koordinatenabbildung $\Phi_{\mathscr{A}}$ ist $v \neq 0_V$ äquivalent zu $\Phi_{\mathscr{A}}(v) \neq 0_{K^n}$, außerdem gilt

$$\phi(v) = \lambda v \quad \Leftrightarrow \quad \Phi_{\mathscr{A}}(\phi(v)) = \Phi_{\mathscr{A}}(\lambda v) \quad \Leftrightarrow \quad \Phi_{\mathscr{A}}(\phi(v)) = \lambda\Phi_{\mathscr{A}}(v) \quad \Leftrightarrow \quad A\Phi_{\mathscr{A}}(v) = \lambda\Phi_{\mathscr{A}}(v). \qquad \square$$

Die Proposition zeigt insbesondere, dass die Eigenwerte von ϕ genau die Eigenwerte der Darstellungsmatrix $\mathscr{M}_{\mathscr{A}}(\phi)$ sind.

Im folgenden beschäftigen wir uns mit der Frage, wie man die Eigenwerte eines Endomorphismus findet. Dafür benötigen wir einige Grundbegriffe und elementare Aussagen über Polynome. Wir nennen ein Polynom $f \in K[x]$ genau dann ***konstant***, wenn $f = 0_K$ oder $\mathrm{grad}(f) = 0$ gilt, wenn f also in K liegt.

Satz 12.7 *(Division mit Rest)*

Seien $f, g \in K[x]$, wobei g nicht-konstant ist. Dann gibt es $q, r \in K[x]$ mit $f = qg + r$, wobei $r = 0_K$ ist oder zumindest $\mathrm{grad}(r) < \mathrm{grad}(g)$ gilt.

Wir verzichten an dieser Stelle auf einen Beweis, weil dieser eher in die Algebra-Vorlesung gehört. Aus dem Schulunterricht ist zumindest für $K = \mathbb{R}$ bekannt, dass die Polynome q und r durch Polynomdivision bestimmt werden können.

Jedem Polynom $f \in K[x]$ kann durch $a \mapsto f(a)$ eine Abbildung $K \to K$ zugeordnet werden, die dadurch zu Stande kommt, dass die Elemente $a \in K$ in die Unbestimmte x eingesetzt werden. Man bezeichnet diese Abbildung auch als die dem Polynom f zugeordnete ***Polynomfunktion***.

Für unendliche Körper gilt allgemein, dass verschiedene Polynome auch verschiedene Polynomfunktionen definieren. Für endliche Körper ist das aber nicht mehr richtig: Beispielsweise definieren die Polynome $f, g \in \mathbb{F}_2[x]$ gegeben durch $f = x$ und $g = x^2$ dieselbe Polynomfunktion, denn es gilt

$$f(\bar{0}) = g(\bar{0}) = \bar{0} \qquad \text{und} \qquad f(\bar{1}) = g(\bar{1}) = \bar{1}.$$

Ein Element $a \in K$ wird ***Nullstelle*** von $f \in K[x]$ genannt, wenn $f(a) = 0_K$ gilt. Man nennt ein Polynom $g \in K[x]$ einen ***Teiler*** von f, wenn ein $h \in K[x]$ mit $f = gh$ existiert. Ist $\mathrm{grad}(g) = 1$, dann nennt man g auch einen ***Linearfaktor*** des Polynoms f. Auch die folgende Aussage ist im Grunde schon aus der Schulmathematik bekannt.

Satz 12.8 Sei $f \in K[x]$ und $a \in K$. Genau dann gilt $f(a) = 0_K$, wenn $x - a$ ein Linearfaktor von f ist.

Beweis: Nach Satz (12.7) gibt es Polynome $g, r \in K[x]$ mit $f = (x-a)g + r$, wobei das Polynom r wegen $r = 0_K$ oder $\mathrm{grad}(r) < \mathrm{grad}(x-a) = 1$ konstant ist. Ist nun a eine Nullstelle von f, dann gilt $r = r(a) = f(a) - (a-a)g(a) = 0_K - 0_K = 0_K$ und somit $f = (x-a)g$. Ist umgekehrt $x - a$ ein Linearfaktor von f, dann gibt es ein $g \in K[x]$ mit $f = (x-a)g$, und es folgt $f(a) = (a-a)g(a) = 0_K$. □

Definition 12.9 Sei $f \in K[x]$ mit $f \neq 0_K$ und $a \in K$ eine Nullstelle von f. Das maximale $r \in \mathbb{N}$ mit der Eigenschaft, dass $(x-a)^r$ ein Teiler von f ist, wird die ***Vielfachheit*** $\mu(f, a)$ der Nullstelle a genannt.

Nach Satz (12.7) gilt also $\mu(f, a) \geq 1$ für jede Nullstelle a von f. Ist $f(a) \neq 0_K$, dann setzen wir $\mu(f, a) = 0$. Das folgende Kriterium ist für die Bestimmung der Vielfachheit einer Nullstelle hilfreich.

Proposition 12.10 Sei $f \in K[x]$ ein Polynom mit einer Zerlegung $f = (x-a)^r g$, wobei $r \in \mathbb{N}_0$ und $g(a) \neq 0_K$ ist. Dann gilt $r = \mu(f, a)$.

Beweis: Die Gleichung $f = (x-a)^r g$ zeigt jedenfalls, dass $\mu(f, a) \geq r$ gilt. Nehmen wir nun an, dass sogar $\mu(f, a) > r$ erfüllt ist. Dann gibt es ein $h \in K[x]$ mit $f = (x-a)^{r+1} h$. Teilt man die Polynomgleichung

$$(x-a)^r g \quad = \quad (x-a)^{r+1} h$$

durch $(x-a)^r$, dann folgt $g = (x-a)h$ und $g(a) = (a-a)h(a) = 0_K$, im Widerspruch zur Voraussetzung $g(a) \neq 0_K$. □

Sei $f \in K[x]$ ein Polynom vom Grad ≥ 1. Man sagt, f ***zerfällt in Linearfaktoren***, wenn es als Produkt von Linearfaktoren geschrieben werden kann. In ausgeschriebener Form bedeutet dies, dass Elemente $c, \lambda_1, ..., \lambda_r \in K$ existieren, so dass

$$f \quad = \quad c \prod_{k=1}^{r} (x - \lambda_k) \qquad \text{gilt.}$$

Ein Körper K wird ***algebraisch abgeschlossen*** genannt, wenn jedes Polynom vom Grad ≥ 1 in $K[x]$ in Linearfaktoren zerfällt. In der Funktionentheorie zeigt man, dass zum Beispiel der Körper $\mathbb{C}$ der komplexen Zahlen diese Eigenschaft besitzt. Dagegen ist $\mathbb{R}$ nicht algebraisch abgeschlossen, denn das Polynom $x^2 + 1$ hat keine Nullstellen in $\mathbb{R}$ und kann deshalb nach Satz (12.8) nicht in Linearfaktoren zerlegt werden. In der Algebra-Vorlesung wird aber gezeigt, dass zu einem Körper K ein algebraisch abgeschlossener Erweiterungskörper existiert. Im Fall $K = \mathbb{R}$ ist dies gerade der Körper $\mathbb{C}$.

Nun werden wir sehen, inwiefern Polynome bei der Bestimmung der Eigenwerte einer Matrix weiterhelfen.

Definition 12.11 Für jede Matrix $A \in \mathcal{M}_{n,K}$ nennt man

$$\chi_A \quad = \quad (-1)^n \det(A - x\mathrm{I}^{(n)}) \quad = \quad \det(x\mathrm{I}^{(n)} - A) \quad \in K[x]$$

das ***charakteristische Polynom*** von A.

Satz 12.12 Die Eigenwerte einer Matrix $A \in \mathcal{M}_{n,K}$ sind genau die Nullstellen des charakteristischen Polynoms χ_A.

Beweis: Für jedes $\lambda \in K$ gilt $\mathrm{Eig}(A,\lambda) = \ker(A - \lambda\mathrm{I}^{(n)})$. Genau dann ist λ ein Eigenwert von A, wenn die Ungleichung $\ker(A - \lambda\mathrm{I}^{(n)}) \neq \{0_V\}$ gilt (vgl. Prop. (12.3)). Nach dem Dimensionssatz für lineare Abbildungen gilt weiter

$$\dim\ker(A - \lambda\mathrm{I}^{(n)}) > 0 \quad \Leftrightarrow \quad \mathrm{rg}(A - \lambda\mathrm{I}^{(n)}) < n \quad \Leftrightarrow \quad \det(A - \lambda\mathrm{I}^{(n)}) = 0_K \quad \Leftrightarrow \quad \chi_A(\lambda) = 0_K \qquad \square$$

Definition 12.13 Ist V ein endlich-dimensionaler K-Vektorraum, $\phi \in \mathrm{End}_K(V)$ und $A \in \mathcal{M}_{n,K}$ die Darstellungsmatrix von V bezüglich einer beliebig gewählen Basis, dann bezeichnen wir $\chi_\phi = \chi_A$ als ***charakteristisches Polynom*** von ϕ.

Proposition 12.14 Das charakteristische Polynom χ_ϕ ist unabhängig von der gewählten Basis des Vektorraums V.

Beweis: Sind $A, B \in \mathcal{M}_{n,K}$ die Darstellungsmatrizen von ϕ bezüglich verschiedener Basen, dann sind A und B nach Prop. (12.5) ähnlich. Es gibt also ein $T \in \mathrm{GL}_n(K)$ mit $B = TAT^{-1}$. Auf Grund der Multiplikativität der Determinantenfunktion folgt

$$\begin{aligned} \chi_B \quad &= \quad \det(x\mathrm{I}^{(n)} - B) \quad = \quad \det(T(x\mathrm{I}^{(n)})T^{-1} - TAT^{-1}) \quad = \quad \det(T(x\mathrm{I}^{(n)} - A)T^{-1}) \\ &= \quad \det(T)\det(x\mathrm{I}^{(n)} - A)\det(T)^{-1} \quad = \quad \det(x\mathrm{I}^{(n)} - A) \quad = \quad \chi_A. \qquad \square \end{aligned}$$

Folgerung 12.15 Auch für jeden Endomorphismus $\phi \in \mathrm{End}_K(V)$ eines endlich-dimensionalen K-Vektorraums V gilt: Die Eigenwerte von ϕ sind genau die Nullstellen des Polynoms χ_ϕ.

Beweis: Sei A die Darstellungsmatrix von ϕ bezüglich einer beliebigen Basis von V. Dann gilt $\chi_\phi = \chi_A$ nach Definition. Auf Grund von Prop. (12.6) sind darüber hinaus die Eigenwerte von ϕ genau die Eigenwerte von A. Also sind die Eigenwerte von ϕ nach Satz (12.12) genau die Nullstellen von χ_A und damit auch genau die Nullstellen von χ_ϕ. □

Definition 12.16 Sind $\lambda_1, ..., \lambda_n \in K$, dann bezeichnen wir mit $\operatorname{diag}(\lambda_1, ..., \lambda_n)$ die Matrix $D = (d_{ij})$ mit den Einträgen

$$d_{ij} \quad = \quad \begin{cases} \lambda_k & \text{falls } i = j = k \\ 0 & \text{falls } i \neq j. \end{cases}$$

Eine Matrix dieser Form wird ***Diagonalmatrix*** genannt. Man bezeichnet eine Matrix $A \in \mathscr{M}_{n,K}$ als ***diagonalisierbar***, wenn sie ähnlich zu einer Diagonalmatrix ist.

Definition 12.17 Einen Endomorphismus ϕ eines endlich-dimensionalen K-Vektorraums V heißt ***diagonalisierbar***, wenn eine Basis von V existiert, so dass die Darstellungsmatrix von ϕ bezüglich dieser Basis eine Diagonalmatrix ist.

Proposition 12.18 Sei V ein endlich-dimensionaler K-Vektorraum, $\phi \in \operatorname{End}_K(V)$ und $A \in \mathscr{M}_{n,K}$ die Darstellungsmatrix von ϕ bezüglich einer beliebigen Basis von V. Genau dann ist A diagonalisierbar, wenn ϕ diagonalisierbar ist.

Beweis: Sei $\mathscr{A}$ eine Basis von V, so dass $A = \mathscr{M}_\mathscr{A}(\phi)$ erfüllt ist.

„$\Leftarrow$“ Weil ϕ nach Voraussetzung diagonalisierbar ist, gibt es eine Basis $\mathscr{B}$ von V, so dass $D = \mathscr{M}_\mathscr{B}(\phi)$ eine Diagonalmatrix ist. Die Matrizen A und D sind also die Darstellungsmatrizen von ϕ bezüglich der Basen $\mathscr{A}, \mathscr{B}$. Nach Prop. (12.5) sind A und D ähnlich, und damit ist A nach Definition diagonalisierbar.

„$\Rightarrow$“ Ist A diagonalisierbar, dann gibt es ein $T \in \mathrm{GL}_n(K)$ mit der Eigenschaft, dass $D = TAT^{-1}$ eine Diagonalmatrix ist. Für jedes $v \in V$ gilt nach Definition der Darstellungsmatrix jeweils

$$A\Phi_\mathscr{A}(v_k) = \Phi_\mathscr{A}(\phi(v)) \quad \Leftrightarrow \quad T^{-1}DT\Phi_\mathscr{A}(v_k) = \Phi_\mathscr{A}(\phi(v)) \quad \Leftrightarrow \quad DT\Phi_\mathscr{A}(v_k) = T\Phi_\mathscr{A}(\phi(v)).$$

Wir wählen nun eine Basis $\mathscr{B}$ mit der Eigenschaft, dass $T = \mathscr{T}^\mathscr{A}_\mathscr{B}$ erfüllt ist (s.u.) und erhalten

$$D\mathscr{T}^\mathscr{A}_\mathscr{B}\Phi_\mathscr{A}(v) = \mathscr{T}^\mathscr{A}_\mathscr{B}\Phi_\mathscr{A}(\phi(v)) \quad \Leftrightarrow \quad D\Phi_\mathscr{B}(v) = \Phi_\mathscr{B}(\phi(v))$$

Weil $\mathscr{M}_\mathscr{B}(\phi)$ die eindeutig bestimmte Matrix mit $\mathscr{M}_\mathscr{B}(\phi)\Phi_\mathscr{B}(v) = \Phi_\mathscr{B}(\phi(v))$ für alle $v \in V$ ist, folgt $\mathscr{M}_\mathscr{B}(\phi) = D$. Dies zeigt, dass ϕ diagonalsierbar ist. □

Beim Beweis von Prop. (12.18) wurde verwendet

Lemma 12.19 Sei V ein endlich-dimensionaler K-Vektorraum, $\mathscr{A}$ eine geordnete Basis von V und $T \in \mathrm{GL}_n(K)$ eine invertierbare Matrix. Dann gibt es eine geordnete Basis $\mathscr{B}$ von V mit $T = \mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$.

Beweis: Sei $\mathscr{A} = (v_1, ..., v_n)$ die vorgegebene geordnete Basis, außerdem $C = (c_{ij})$ die inverse Matrix $C = T^{-1}$ und $\mathscr{B} = (w_1, ..., w_n)$ das Tupel bestehend aus den Vektoren

$$w_\ell \quad = \quad \sum_{k=1}^{n} c_{k\ell} v_k \qquad \text{für} \quad 1 \le \ell \le n.$$

Mit der Matrix $T = (t_{ij})$ gilt dann

$$\sum_{i=1}^{n} t_{i\ell} w_i \quad = \quad \sum_{i=1}^{n} t_{i\ell} \left(\sum_{j=1}^{n} c_{ji} v_j \right) \quad = \quad \sum_{j=1}^{n} \left(\sum_{i=1}^{n} c_{ji} t_{i\ell} \right) v_j \quad = \quad \sum_{j=1}^{n} \delta_{j\ell} v_j \quad = \quad v_\ell$$

wegen $CT = I^{(n)}$. Da die Elemente der Basis $\mathscr{A}$ als Linearkombinationen der Elemente von $\mathscr{B}$ dargestellt werden können, ist auch $\mathscr{B}$ ein Erzeugendensystem, als n-elementiges Tupel wegen $\dim V = n$ sogar eine geordnete Basis von V. Für $1 \le k \le n$ gilt

$$Te_k \quad = \quad \sum_{i=1}^{n} t_{ik} e_i \quad = \quad \sum_{i=1}^{n} t_{ik} \Phi_{\mathscr{B}}(w_i) \quad = \quad \Phi_{\mathscr{B}} \left(\sum_{i=1}^{n} t_{ik} w_i \right) \quad = \quad \Phi_{\mathscr{B}}(v_k).$$

Da die Transformationsmatrix $\mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$ ebenfalls die Gleichungen $\mathscr{T}^{\mathscr{A}}_{\mathscr{B}} e_k = \mathscr{T}^{\mathscr{A}}_{\mathscr{B}} \Phi_{\mathscr{A}}(v_k) = \Phi_{\mathscr{B}}(v_k)$ erfüllt, stimmen die Spalten von T und $\mathscr{T}^{\mathscr{A}}_{\mathscr{B}}$ überein, und es folgt $T = T^{\mathscr{A}}_{\mathscr{B}}$. □

Wir können nun ein neues Kriterium für die Diagonalisierbarkeit herleiten.

Proposition 12.20 Sei $V \neq \{0_V\}$ ein endlich-dimensionaler K-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:

(i) Der Endomorphismus ϕ ist diagonalisierbar.
(ii) Der Vektorraum V besitzt eine Basis bestehend aus Eigenvektoren von ϕ.

Beweis: „(i) ⇒ (ii)" Nach Voraussetzung gibt es eine Basis $\mathscr{A} = (v_1, ..., v_n)$ von V mit der Eigenschaft, dass $D = \mathscr{M}_{\mathscr{A}}(\phi)$ eine Diagonalmatrix ist, $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ mit $\lambda_k \in K$ für $1 \le k \le n$. Der k-te Spaltenvektor von D ist jeweils das λ_k-fache des k-ten Einheitsvektors e_k. Es folgt

$$De_k = \lambda_k e_k \quad \Leftrightarrow \quad \mathscr{M}_{\mathscr{A}}(\phi)\Phi_{\mathscr{A}}(v_k) = \lambda_k \Phi_{\mathscr{A}}(v_k) \quad \Leftrightarrow \quad \Phi_{\mathscr{A}}(\phi(v_k)) = \Phi_{\mathscr{A}}(\lambda_k v_k) \quad \Leftrightarrow \quad \phi(v_k) = \lambda_k v_k$$

für $1 \le k \le n$, wobei im letzten Schritt die Bijektivität von $\Phi_{\mathscr{A}}$ verwendet wurde. Als Element einer Basis ist $v_k \neq 0_V$; zusammen mit der Gleichung $\phi(v_k) = \lambda_k v_k$ zeigt dies, dass $\mathscr{A}$ aus Eigenvektoren von ϕ besteht.

„(ii) ⇒ (i)" Sei $\mathscr{A} = (v_1, ..., v_n)$ eine Basis von V, wobei v_k jeweils ein Eigenvektor von ϕ zum Eigenwert λ_k ist, für $1 \le k \le n$. Außerdem sei $D = \mathscr{M}_{\mathscr{A}}(\phi)$. Dann gilt jeweils $\phi(v_k) = \lambda_k v_k$, und die Rechnung aus dem vorherigen Absatz hat gezeigt, dass dies äquivalent zu $De_k = \lambda_k e_k$ ist. Die k-te Spalte von D ist also gleich $\lambda_k e_k$, für $1 \le k \le n$. Daraus folgt $D = \mathrm{diag}(\lambda_1, ..., \lambda_n)$, also ist D eine Diagonalmatrix und ϕ damit diagonalisierbar. □

Als nächstes zeigen wir, dass der Vektorraum V bezüglich eines diagonalisierbaren Endomorphismus in Eigenräume zerlegt werden kann. Die beiden folgenden Aussagen dienen zur Vorbereitung.

Proposition 12.21 Sei V ein K-Vektorraum, $\phi \in \mathrm{End}_K(V)$, und seien $\lambda_1, ..., \lambda_r$ verschiedene Elemente von K. Für jedes $k \in \{1, ..., r\}$ sei $v_k \in V$ jeweils ein Eigenvektor zum Eigenwert λ_k. Dann ist das Tupel $(v_1, ..., v_r)$ linear unabhängig.

Beweis: Wir führen den Beweis durch vollständige Induktion über r. Für $r = 1$ ist die Aussage wegen $v_1 \neq 0_V$ klar. Sei nun $r \in \mathbb{N}$, und setzen wir nun die Behauptung für dieses r voraus. Es seien $\lambda_1, ..., \lambda_{r+1} \in K$ verschieden, und sei $v_k \in V$ für $1 \le k \le r+1$ jeweils ein Eigenvektor zum Eigenwert λ_k. Zum Nachweis der linearen Unabhängigkeit seien $\alpha_1, ..., \alpha_{r+1} \in K$ mit der Eigenschaft

$$\sum_{k=1}^{r+1} \alpha_k v_k \quad = \quad 0_V. \tag{1}$$

Dann liefert die Multiplikation von (1) mit dem Wert λ_{r+1} einerseits

$$\sum_{k=1}^{r+1} \alpha_k \lambda_{r+1} v_k \quad = \quad 0_V \quad , \tag{2}$$

andererseits erhält man durch Anwendung von ϕ auf (1) aber auch

$$\sum_{k=1}^{r+1} \alpha_k \lambda_k v_k \quad = \quad \sum_{k=1}^{r+1} \alpha_k \phi(v_k) \quad = \quad \phi\left(\sum_{k=1}^{r+1} \alpha_k v_k\right) \quad = \quad \phi(0_V) \quad = \quad 0_V. \tag{3}$$

Subtrahieren wir die Gleichungen (2) und (3) voneinander, so erhalten wir

$$\sum_{k=1}^{r+1} \alpha_k \lambda_{r+1} v_k - \sum_{k=1}^{r+1} \alpha_k \lambda_k v_k \quad = \quad \sum_{k=1}^{r+1} \alpha_k (\lambda_{r+1} - \lambda_k) v_k \quad = \quad 0_V.$$

Da das Tupel $(v_1, ..., v_r)$ nach Induktionsvoraussetzung linear unabhängig sind, folgt $\alpha_k(\lambda_{r+1} - \lambda_k) = 0_K$ für $1 \le k \le r$. Wegen $\lambda_{r+1} - \lambda_k \neq 0_K$ folgt $\alpha_k = 0_K$ für $1 \le k \le r$. Setzen wir dies wiederum in (1) ein, so erhalten wir $\alpha_{r+1} v_{r+1} = 0_V$, und wegen $v_{r+1} \neq 0_V$ folgt $\alpha_{r+1} = 0_K$. Damit ist die linearen Unabhängigkeit von $(v_1, ..., v_{r+1})$ nachgewiesen. □

Proposition 12.22 Sei ϕ ein Endomorphismus eines K-Vektorraums V, und seien $\lambda_1, ..., \lambda_r$ verschiedene Elemente des Körpers K. Dann gilt

$$\mathrm{Eig}(\phi, \lambda_k) \cap \left(\sum_{\ell \neq k} \mathrm{Eig}(\phi, \lambda_\ell)\right) \quad = \quad \{0_V\} \qquad \text{für} \quad 1 \le k \le r.$$

Beweis: Nehmen wir an, dass ein $k \in \{1, ..., r\}$ und ein Vektor $v \neq 0_V$ in der angegebenen Schnittmenge. Dann gibt es Vektoren $v_\ell \in \mathrm{Eig}(\phi, \lambda_\ell)$ für $1 \leq \ell \leq r$ mit

$$v_k = v = \sum_{\ell \neq k} v_\ell \quad \Leftrightarrow \quad \sum_{\ell \neq k} v_\ell + (-1_K) v_k \quad = \quad 0_V.$$

Aber wegen $v_k \neq 0_V$ steht dies steht im Widerspruch zur Prop. (12.21), das Tupel bestehend aus den Vektoren v_ℓ mit $v_\ell \neq 0_V$ linear unabhängig ist. □

Mit diesen Ergebnissen erhalten wir ein neues Kriterium für die Diagonalisierbarkeit eines Endomorphismus.

Proposition 12.23 Sei $V \neq \{0_V\}$ endlich-dimensionaler K-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Dann sind die folgenden Aussagen äquivalent:

(i) Der Endomorphismus ϕ ist diagonalisierbar.

(ii) Es gibt verschiedene Elemente $\lambda_1, ..., \lambda_r \in K$, so dass

$$V \quad = \quad \bigoplus_{\ell=1}^{r} \mathrm{Eig}(\phi, \lambda_\ell) \qquad \text{erfüllt ist.}$$

Beweis: „(i) ⇒ (ii)" Nach Voraussetzung existiert eine Basis $\mathscr{A} = \{v_1, ..., v_n\}$ von V bestehend aus Eigenvektoren. Seien $\lambda_1, ..., \lambda_r$ die verschiedenen Eigenwerte von ϕ. Weil alle Elemente der Basis Eigenvektoren sind, gibt es für jedes $k \in \{1, ..., n\}$ ein $\ell \in \{1, ..., r\}$ mit $\phi(v_k) = \lambda_\ell v_k$. Es gilt dann also $v_k \in \mathrm{Eig}(\phi, \lambda_\ell)$. Setzen wir $U = \sum_{\ell=1}^{r} \mathrm{Eig}(\phi, \lambda_\ell)$, dann gilt insgesamt $\mathscr{A} \subseteq U$. Weil $\mathscr{A}$ eine Basis und U ein Untervektorraum von V ist, stimmt V mit der Summe U ein, und nach Prop. (12.22) ist diese Summe direkt.

„(ii) ⇒ (i)" Für jedes $\ell \in \{1, ..., r\}$ sei $\mathscr{A}_\ell$ eine Basis von $\mathrm{Eig}(\phi, \lambda_\ell)$. Auf Grund der direkten Summenzerlegung ist dann $\mathscr{A} = \bigcup_{\ell=1}^{r} \mathscr{A}_\ell$ eine Basis von V. Jedes $\mathscr{A}_\ell$ besteht aus Eigenvektoren von ϕ, somit auch die Basis $\mathscr{A}$. Daraus folgt die Diagonalisierbarkeit von ϕ. □

Definition 12.24 Sei V ein endlich-dimensionaler K-Vektorraum, ϕ ein Endomorphismus von V und $\lambda \in K$ ein Eigenwert von ϕ.

(i) Die Vielfachheit $\mu(\chi_\phi, \lambda)$ von λ als Nullstelle des Polynoms χ_ϕ bezeichnet man als ***algebraische*** Vielfachheit $\mu_a(\phi, \lambda)$ des Eigenwerts λ.

(ii) Die Eigenraum-Dimension $\mu_g(\phi, \lambda) = \dim \mathrm{Eig}(\phi, \lambda)$ nennt man die ***geometrische*** Vielfachheit von λ.

Für eine quadratische Matrix $A \in \mathscr{M}_{n,K}$ definiert man algebraische und geometrische Vielfachheit eines Eigenwerts λ auf analoge Weise: Man setzt $\mu_a(A, \lambda) = \mu(\chi_A, \lambda)$ und $\mu_g(A, \lambda) = \dim \mathrm{Eig}(A, \lambda)$.

In Beweisen ist es oft günstig, wenn die algebraische und geometrische Vielfachheit auch für Nicht-Eigenwerte eines Endomorphismus ϕ definiert sind, weil dadurch einige Fallunterscheidungen unnötig werden. Wenn $\lambda \in K$ kein Eigenwert von ϕ ist, dann setzt man $\mu_a(\phi, \lambda) = \mu_g(\phi, \lambda) = 0$.

Proposition 12.25 Sei V ein n-dimensionaler K-Vektorraum, $\phi : V \to V$ ein Endomorphismus und $\lambda \in K$ ein Eigenwert mit algebraischer Vielfachheit μ_a und geometrischer Vielfachheit μ_g. Dann gilt $1 \le \mu_g \le \mu_a$.

Beweis: Wir haben bereits gesehen, dass λ genau dann ein Eigenwert ist, wenn $\mathrm{Eig}(\phi, \lambda) \neq \{0_V\}$ ist. Deshalb gilt $\mu_g \ge 1$. Sei nun $(v_1, ..., v_r)$ eine Basis von $\mathrm{Eig}(\phi, \lambda)$, die wir durch $v_{r+1}, ..., v_n$ zu einer Basis $\mathscr{A}$ von V ergänzen. Wegen $\phi(v_i) = \lambda v_i$ für $1 \le i \le r$ hat A die Form

$$A \quad = \quad \begin{pmatrix} \lambda \mathrm{I}^{(r)} & B \\ 0 & C \end{pmatrix}$$

mit geeignet gewählten Matrizen $B \in \mathscr{M}_{r \times (n-r), K}$ und $C \in \mathscr{M}_{n-r, K}$. Sei nun $\chi \in K[x]$ das charakteristische Polynom von ϕ und $\tilde{K}$ ein unendlicher Erweiterungskörper von K. Für alle $\alpha \in \tilde{K}$ gilt dann

$$\chi(\alpha) \quad = \quad \det(\alpha \mathrm{I}^{(n)} - A) \quad = \quad \det \begin{pmatrix} (\alpha - \lambda)\mathrm{I}^{(r)} & -B \\ 0 & \alpha \mathrm{I}^{(n-r)} - C \end{pmatrix} \quad = \quad (\alpha - \lambda)^r \det(\alpha \mathrm{I}^{(n-r)} - C) \quad ,$$

wobei wir im letzten Schritt Satz (11.4) für Blockmatrizen angewendet haben. Damit erhalten wir für das charakteristische Polynom $\chi = (x - \lambda)^r g = (x - \lambda)^r g$, wobei $g = \det(x\mathrm{I}^{(n-r)} - C) \in K[x]$ das charakteristische Polynom der Matrix C bezeichnet. Dies beweist $\mu_g = r \le \mu_a$. □

Satz 12.26 Sei $n \in \mathbb{N}$, V ein n-dimensionaler K-Vektorraum, ϕ ein Endomorphismus von V und $\chi \in K[x]$ sein charakteristisches Polynom. Dann sind die folgenden Aussagen äquivalent:

(i) Der Endomorphismus ϕ ist diagonalisierbar

(ii) Der Vektorraum V besitzt eine Basis bestehend aus Eigenvektoren von ϕ.

(iii) Das Polynom χ zerfällt in Linearfaktoren, und für jeden Eigenwert λ von ϕ stimmen algebraische und geometrische Vielfachheit überein.

(iv) Es gibt $\lambda_1, ..., \lambda_r \in K$, so dass $V = \mathrm{Eig}(\phi, \lambda_1) \oplus ... \oplus \mathrm{Eig}(\phi, \lambda_r)$ gilt.

Beweis: Die Äquivalenz von (i), (ii) und (iv) wurde bereits bewiesen.

„(i) ⇒ (iii)“ Sei $\mathscr{A} = (v_1, ..., v_n)$ eine Basis, und seien $\lambda_1, ..., \lambda_n \in K$, so dass $A = M_{\mathscr{A}}(\phi) = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ gilt. Sei außerdem $\tilde{K} \supseteq K$ ein unendlicher Erweiterungskörper. Dann gilt für alle $\alpha \in \tilde{K}$ die Gleichung

$$\chi(\alpha) \quad = \quad \det(\alpha \mathrm{I}^{(n)} - A) \quad = \quad \det\big(\mathrm{diag}(\alpha - \lambda_1, ..., \alpha - \lambda_n)\big) \quad = \quad \prod_{i=1}^{n} (\alpha - \lambda_i)$$

und somit $\chi = \prod_{i=1}^{n}(x-\lambda_i)$. Also zerfällt das charakteristische Polynom in Linearfaktoren. Zu zeigen bleibt, dass für jeden Eigenwert λ_i jeweils algebraische und geometrische Vielfachheit übereinstimmen. Für jeden Eigenwert λ von ϕ sei $S(\lambda) = \{i \mid \lambda_i = \lambda\}$. Dann können wir das charakteristische Polynom in der Form $\chi = gh$ mit

$$g \quad = \quad \prod_{i\in S(\lambda)}(x-\lambda_i) \quad = \quad (x-\lambda)^{|S(\lambda)|} \qquad \text{und} \qquad h \quad = \quad \prod_{i\notin S(\lambda)}(x-\lambda_i)$$

zerlegen. Wegen $h(\lambda) \neq 0$ ist $\mu = |S(\lambda)|$ die Vielfachheit der Nullstelle λ von f, also die algebraische Vielfachheit des Eigenwerts λ. Für jedes $i \in S(\lambda)$ ist aber auch v_i ein Eigenvektor zum Eigenwert λ, es gilt also $\dim \mathrm{Eig}(\phi,\lambda) \geq \mu$. Weil die geometrische Vielfachheit nach Prop. (12.25) immer durch die algebraische Vielfachheit beschränkt ist, folgt $\mathrm{Eig}(\phi,\lambda) = \mu$.

„(iii) ⇒ (iv)“ Sei $\chi = a\prod_{i=1}^{r}(x-\lambda_i)^{\mu_i}$ eine Zerlegung des charakteristischen Polynoms in Linearfaktoren, wobei $a \in K$, die Nullstellen $\lambda_1,\ldots,\lambda_r \in K$ paarweise verschieden und $\mu_i \in \mathbb{N}$ jeweils die Vielfachheiten der Nullstellen λ_i sind. Wegen $\mathrm{grad}(f) = n$ gilt $\sum_{i=1}^{r}\mu_i = n$.

Für jedes $i \in \{1,\ldots,r\}$ definieren wir nun $U_i = \mathrm{Eig}(\phi,\lambda_i)$, außerdem sei $V_r = \sum_{i=1}^{r} U_i$. Da algebraische und geometrische Vielfachheit übereinstimmen, gilt $\mu_i = \dim U_i$ für $1 \leq i \leq r$. Nach Prop. (12.22) ist der Durchschnitt von U_i mit $\sum_{j\neq i} U_j$ jeweils gleich $\{0_V\}$. Dies zeigt, dass V_r die *direkte* Summe der Untervektorräume U_i ist. Nach Folgerung (8.3) erhalten wir

$$\dim V_r \quad = \quad \sum_{i=1}^{r}\dim U_i \quad = \quad \sum_{i=1}^{r}\mu_i \quad = \quad n \quad = \quad \dim V.$$

Aus $V_r \subseteq V$ und $\dim V_r = \dim V$ wiederum folgt $V = V_r$. □

§ 13. Die Jordansche Normalform

Sei V ein endlicher K-Vektorraum und $\phi \in \mathrm{End}_K(V)$. Zerfällt das charakteristische Polynom χ_ϕ in Linearfaktoren und gilt $\mu_a(\phi,\lambda) = \mu_g(\phi,\lambda)$ für jeden Eigenwert, dann lässt sich die Darstellungsmatrix von ϕ auf Diagonalgestalt bringen. Aber auch wenn die zweite Bedingung nicht erfüllt ist, kann für die Darstellungsmatrix noch eine einfache Form gefunden werden, die von der Diagonalgestalt nur geringfügig abweicht.

Definition 13.1 Eine Matrix $J \in \mathscr{M}_{n,K}$ heißt ***Jordanmatrix*** zum Eigenwert $\lambda \in K$, wenn sie die Form

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix} \quad \text{besitzt.}$$

In kleiner Dimension sind die Jordanmatrizen zum Eigenwert λ also gegeben durch

$$\begin{pmatrix} \lambda \end{pmatrix} \quad , \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad , \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad , \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

Eine Matrix $A \in \mathscr{M}_{n,K}$ befindet sich in ***Jordanscher Normalform***, wenn sie als Blockmatrix in der Form

$$A = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix} \quad \text{mit Jordanmatrizen} \quad J_1, \ldots, J_r \quad \text{schreiben lässt.}$$

Man bezeichnet die $J_1, \ldots, J_r$ dann als ***Jordanblöcke*** der Matrix A. Zum Beispiel setzt sich die Matrix

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

aus den Jordanblöcken

$$J_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \quad , \quad J_2 = \begin{pmatrix} 2 \end{pmatrix} \quad \text{und} \quad J_3 = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \quad \text{zusammen.}$$

Ziel dieses Abschnitts ist der

Satz 13.2 Sei $\phi : V \to V$ ein Endomorphismus auf einem K-Vektorraum V mit der Eigenschaft, dass das charakteristische Polynom $\chi_\phi \in K[x]$ in Linearfaktoren zerfällt. Dann gibt es eine geordnete Basis $\mathscr{A}$ von V, so dass sich die Darstellungsmatrix $M_{\mathscr{A}}(\phi)$ in Jordanscher Normalform befindet.

Der erste Schritt zum Beweis besteht in der Zerlegung von V in gewisse Unterräume, die als Verallgemeinerung der Eigenräume angesehen werden können.

Definition 13.3 Sei $\phi : V \to V$ ein Endomorphismus auf einem K-Vektorraum V und $\lambda \in K$ ein beliebiges Element. Dann bezeichnet man

$$\mathrm{Hau}(\phi, \lambda) \quad = \quad \bigcup_{s=1}^{\infty} \ker((\phi - \lambda \mathrm{id}_V)^s) \qquad \text{als } \textbf{\textit{Hauptraum}} \text{ zum Wert } \lambda.$$

Lemma 13.4

(i) Für jedes $\lambda \in K$ ist $\mathrm{Hau}(\phi, \lambda)$ ein Untervektorraum von V, und es gilt jeweils $\phi(\mathrm{Hau}(\phi, \lambda)) \subseteq \mathrm{Hau}(\phi, \lambda)$.

(ii) Es gilt $\mathrm{Hau}(\phi, \lambda) \neq \{0_V\}$ genau dann, wenn λ ein Eigenwert von ϕ ist.

(iii) Der Endomorphismus $\phi|_{\mathrm{Hau}(\phi,\lambda)}$ besitzt außer λ keine weiteren Eigenwerte.

Beweis: zu (i) Sei $\psi = \phi - \lambda \mathrm{id}_V$. Für alle $s \in \mathbb{N}$ gilt offenbar $\ker(\psi^r) \subseteq \ker(\psi^s)$ für $r \leq s$. Seien nun $v, w \in \mathrm{Hau}(\phi, \lambda)$ und $\alpha \in K$ vorgegeben. Dann gibt es nach Definition des Hauptraums $r, s \in \mathbb{N}$ mit $v \in \ker(\psi^r)$ und $w \in \ker(\psi^s)$, wobei wir o.B.d.A. $r \leq s$ annehmen können. Es folgt $v, w \in \ker(\psi^s)$, und weil dies ein Untervektorraum von V ist, folgt

$$v + w, \alpha v \quad \in \quad \ker(\psi^s) \quad \subseteq \quad \mathrm{Hau}(\phi, \lambda).$$

Nun zeigen wir noch, dass ϕ den Hauptraum in sich abbildet. Für vorgegebenes $v \in \mathrm{Hau}(\phi, \lambda)$ gibt es ein $s \in \mathbb{N}$ mit $\psi^s(v) = 0$. Wegen $\phi \circ \psi = \psi \circ \phi$ folgt $\psi^s(\phi(v)) = \phi(\psi^s(v)) = \phi(0_V) = 0_V$, also $\phi(v) \in \mathrm{Hau}(\phi, \lambda)$.

zu (ii) Nach Definition gilt einerseits $\mathrm{Eig}(\phi, \lambda) = \ker(\psi) \subseteq \mathrm{Hau}(\phi, \lambda)$. Sei andererseits $0_V \neq v \in \mathrm{Hau}(\phi, \lambda)$ und $s \in \mathbb{N}$ minimal mit $\psi^s(v) = 0_V$. Setzen wir $w = \psi^{s-1}(v)$, dann ist $w \neq 0_V$ und $\psi(w) = 0_V \Leftrightarrow \phi(w) = \lambda w$, also ist w ein Eigenvektor von λ.

zu (iii) Nehmen wir an, $\mu \neq \lambda$ ist ein weiterer Eigenwert von $\phi|_{\mathrm{Hau}(\phi,\lambda)}$. Dann gibt es einen Vektor $0_V \neq w \in \mathrm{Hau}(\phi, \lambda)$ mit $\phi(w) = \mu w$. Aus $\psi(w) = (\mu - \lambda) w$ folgt $\psi^s(w) = (\mu - \lambda)^s w \neq 0_V$ für alle $s \in \mathbb{N}$, was $w \in \mathrm{Hau}(\phi, \lambda)$ widerspricht. □

Lemma 13.5 Für jedes $\lambda \in K$ existiert jeweils eine kleinste Zahl $f(\lambda)$ mit

$$\mathrm{Hau}(\phi,\lambda) \quad = \quad \ker((\phi - \lambda \mathrm{id}_V)^{f(\lambda)}).$$

Darüber hinaus gilt $V = \mathrm{Hau}(\phi,\lambda) \oplus \mathrm{im}((\phi - \lambda\mathrm{id}_V)^{f(\lambda)})$ wobei neben $\mathrm{Hau}(\phi,\lambda)$ auch der Untervektorraum $\mathrm{im}((\phi-\lambda\mathrm{id}_V)^{f(\lambda)})$ von ϕ in sich abgebildet wird.

Beweis: Zur Vereinfachung der Notation setzen wir wieder $\psi = \phi - \lambda\mathrm{id}_V$. Sei $(v_1,...,v_\ell)$ eine geordnete Basis von $\mathrm{Hau}(\phi,\lambda)$. Dann gibt es für jedes $i \in \{1,...,\ell\}$ ein kleinstes $s_i \in \mathbb{N}$ mit $\psi^{s_i}(v_i) = 0_V$. Sei $f(\lambda) = \max\{s_1,...,s_\ell\}$. Dann gilt $\psi^{f(\lambda)}(v_i) = 0_V$ für $1 \le i \le \ell$, also $\mathrm{Hau}(\phi,\lambda) \subseteq \ker(\psi^{f(\lambda)})$. Nach Definition von $\mathrm{Hau}(\phi,\lambda)$ gilt auch die umgekehrte Inklusion, und offenbar ist $f(\lambda)$ das minimale $s \in \mathbb{N}$ mit der Eigenschaft $\mathrm{Hau}(\phi,\lambda) = \ker(\psi^s)$.

Beweisen wir nun die Gleichung $V = \ker(\psi^{f(\lambda)}) \oplus \mathrm{im}(\psi^{f(\lambda)})$. Auf Grund des Dimensionssatzes für lineare Abbildungen gilt

$$\dim\ker(\psi^{f(\lambda)}) + \dim\mathrm{im}(\psi^{f(\lambda)}) \quad = \quad \dim V.$$

Es genügt deshalb, $\ker(\psi^{f(\lambda)}) \cap \mathrm{im}(\psi^{f(\lambda)}) = \{0_V\}$ zu zeigen. Sei v ein Vektor im Durchschnitt dieser beiden Untervektorräume. Dann gibt es ein $w \in V$ mit $\psi^{f(\lambda)}(w) = v$. Aus $v \in \ker(\psi^{f(\lambda)})$ folgt $\psi^{2f(\lambda)}(w) = 0_V$, also $w \in \mathrm{Hau}(\phi,\lambda)$. Wegen $\mathrm{Hau}(\phi,\lambda) = \ker(\psi^{f(\lambda)})$ erhalten wir $v = \psi^{f(\lambda)}(w) = 0_V$.

Dass der Untervektorraum $\mathrm{im}(\psi^{f(\lambda)})$ von ϕ auf sich abildet wird, folgt schließlich auch hier wieder aus der Gleichung $\phi \circ \psi = \psi \circ \phi$. □

Ohne Beweis sei noch erwähnt, dass die Zahl $f(\lambda)$ immer durch die algebraische Vielfachheit $\mu_a(\lambda)$ des Eigenwerts λ beschränkt ist. Es gilt also

$$\mathrm{Hau}(\phi,\lambda) \quad = \quad \ker((\phi - \lambda\mathrm{id}_V)^{\mu_a(\lambda)}) \tag{1}$$

Lemma 13.6 Sei $\phi : V \to V$ ein Endomorphismus auf einem K-Vektorraum V. Sind $\lambda_1,...,\lambda_s \in K$ verschiedene Eigenwerte von ϕ, dann gilt

$$\mathrm{Hau}(\phi,\lambda_i) \cap \left(\sum_{j\neq i} \mathrm{Hau}(\phi,\lambda_j)\right) \quad = \quad \{0_V\}.$$

Beweis: Wir beweisen die Aussage durch vollständige Induktion über s. Für $s = 1$ ist die Aussage offenbar erfüllt. Setzen wir sie nun für s als gültig voraus; nach Umnummerierung der λ_j können wir auch $i = s+1$ annehmen. Sei $v_j \in \mathrm{Hau}(\phi,\lambda_j)$ für $1 \le j \le s$ und $v \in \mathrm{Hau}(\phi,\lambda_{s+1})$ mit

$$v \quad = \quad \sum_{j=1}^{s} v_j.$$

Zu zeigen ist $v = 0_V$. Nach Lemma (13.5) gilt zunächst

$$\sum_{j=1}^{s} (\phi - \lambda_{s+1}\mathrm{id}_V)^{f(\lambda_{s+1})}(v_j) \quad = \quad (\phi - \lambda_{s+1}\mathrm{id}_V)^{f(\lambda_s)}(v) \quad = \quad 0_V. \tag{2}$$

Jeder der Räume $\mathrm{Hau}(\phi, \lambda_j)$ wird nach Lemma (13.4) (i) von ϕ und damit auch von $\phi - \lambda_{s+1}\mathrm{id}_V$ auf sich abgebildet. Der Vektor $w_j = (\phi - \lambda_{s+1}\mathrm{id}_V)^{f(\lambda_{s+1})}(v_j)$ liegt also in $\mathrm{Hau}(\phi, \lambda_j)$ für $1 \le j \le s$. Die Gleichung (2) kann nun umgestellt werden zu

$$w_s \quad = \quad \sum_{j=1}^{s-1} w_j \quad ,$$

und die Induktionsvoraussetzung liefert $w_j = 0$ für $1 \le j \le s$. Dies wiederum bedeutet $v_j \in \mathrm{Hau}(\phi, \lambda_j) \cap \mathrm{Hau}(\phi, \lambda_{s+1})$ für $1 \le j \le s$.

Wir zeigen nun, dass daraus $v_j = 0_V$ für $1 \le j \le s$ folgt. Ansonsten gäbe es ein minimales $k \in \mathbb{N}$ mit der Eigenschaft $(\phi - \lambda_{s+1}\mathrm{id}_V)^k(v_j) = 0_V$. Setzen wir $u_j = (\phi - \lambda_{s+1}\mathrm{id}_V)^{k-1}(v_j)$, dann gilt $u_j \neq 0_V$ und $(\phi - \lambda_{s+1}\mathrm{id}_V)(u_j) = 0$. Aber dies steht im Widerspruch zu Lemma (13.4) (iii). Also muss $v_j = 0_V$ für $1 \le j \le s$ und damit auch $v = 0_V$ gelten. □

Satz 13.7 Sei V ein endlich-dimensionaler K-Vektorraum und $\phi : V \to V$ ein Endomorphismus, dessen charakteristisches Polynom χ_ϕ in $K[x]$ in Linearfaktoren zerfällt. Sind $\lambda_1, ..., \lambda_s \in K$ die verschiedenen Eigenwerte von ϕ, dann gilt

$$V \quad = \quad \mathrm{Hau}(\phi, \lambda_1) \oplus ... \oplus \mathrm{Hau}(\phi, \lambda_s).$$

Beweis: Wir führen den Beweis durch vollständige Induktion über $n = \dim V$. Im Fall $n = 1$ hat das charakteristische Polynom die Form $\chi_\phi = x - \lambda_1$ für ein $\lambda_1 \in K$. Weil λ_1 als Nullstelle des charakteristischen Polynoms ein Eigenwert und V eindimensional ist, gilt in diesem Fall $V = \mathrm{Eig}(\phi, \lambda_1) = \mathrm{Hau}(\phi, \lambda_1)$.

Sei nun $n > 1$, und setzen wir die Aussage für Endomorphismen auf Vektorräumen kleinerer Dimension voraus. Nach Lemma (13.5) gibt es einen Untervektorraum U von V mit $V = \mathrm{Hau}(\phi, \lambda_1) \oplus U$ und $\phi(U) \subseteq U$. Wegen $\dim \mathrm{Hau}(\phi, \lambda_1) = \dim \mathrm{Eig}(\phi, \lambda_1) \ge 1$ gilt $\dim U < n$. Seien χ_1, χ_U die charakteristischen Polynome von $\phi_1 = \phi|_{\mathrm{Eig}(\phi, \lambda_1)}$ und $\phi_U = \phi|_U$. Damit die Induktionsvoraussetzung auf ϕ_U angewendet werden kann, müssen wir zeigen, dass χ_U in $K[x]$ in Linearfaktoren zerfällt. Dazu beweisen wir die Gleichung

$$\chi \quad = \quad \chi_1 \cdot \chi_U.$$

Ist $\mathscr{A} = (v_1, ..., v_m)$ eine Basis von $\mathrm{Hau}(\phi, \lambda_1)$ und $\mathscr{B} = (w_1, ..., w_{n-m})$ eine Basis von U, dann ist

$$\mathscr{C} \quad = \quad (v_1, ..., v_m, w_1, ..., w_{n-m})$$

wegen $\mathrm{Hau}(\phi, \lambda_1) + U = V$ ein Erzeugendensystem von V und wegen $\dim \mathrm{Hau}(\phi, \lambda_1) + \dim U = \dim V$ auch eine Basis. Die Darstellungsmatrix von ϕ bezüglich V ist dann eine Blockmatrix der Form

$$\mathscr{M}_{\mathscr{C}}(\phi) \quad = \quad \begin{pmatrix} \mathscr{M}_{\mathscr{A}}(\phi) & * \\ 0 & M_{\mathscr{B}}(\phi) \end{pmatrix}.$$

Daraus folgt

$$\chi \quad = \quad \det(\mathscr{M}_{\mathscr{C}}(\phi) - x\mathrm{I}^{(n)}) \quad = \quad \det(\mathscr{M}_{\mathscr{A}}(\phi_1) - x\mathrm{I}^{(m)}) \det(\mathscr{M}_{\mathscr{B}}(\phi_U) - x\mathrm{I}^{(n-m)}) \quad = \quad \chi_1 \cdot \chi_U.$$

Mit χ zerfällt auch χ_U in Linearfaktoren. Nach Lemma (13.4) ist λ_1 der einzige Eigenwert von ϕ_1, und folglich sind $\lambda_2, ..., \lambda_s$ die Eigenwerte von ϕ_U. Die Anwendung der Induktionsvoraussetzung auf ϕ_U liefert eine Zerlegung

$$U \quad = \quad \mathrm{Hau}(\phi, \lambda_2) \oplus ... \oplus \mathrm{Hau}(\phi, \lambda_s) \quad ,$$

so dass wir insgesamt die gewünschte Zerlegung von V erhalten. □

Beispiel: Mit Hilfe der Gleichung (1) können wir die Hauptraumzerlegung eines Endomorphismus ϕ explizit bestimmen. Als Beispiel betrachten wir den Endomorphismus $\phi_A : \mathbb{R}^3 \to \mathbb{R}^3$ zur Matrix

$$A \quad = \quad \begin{pmatrix} 25 & 34 & 18 \\ -14 & -19 & -10 \\ -4 & -6 & -1 \end{pmatrix}$$

Das charakteristische Polynom von A ist gegeben durch

$$\chi_A \quad = \quad -x^3 + 5x^2 - 7x + 3 \quad = \quad -(x-1)^2(x-3).$$

Die Eigenwerte von A sind also 1 und 3, wobei der erste mit algebraischer Vielfachheit 2, der zweite mit algebraischer Vielfachheit 1 auftritt. Nach (1) ist $\mathrm{Hau}(\phi_A, 1)$ der Kern der Abbildung ϕ_B mit

$$B \quad = \quad (A - 1 \cdot \mathrm{I}^{(3)})^2 \quad = \quad \begin{pmatrix} 28 & 28 & 56 \\ -16 & -16 & -32 \\ -4 & -4 & -8 \end{pmatrix}.$$

Durch Anwendung des Gaußalgorithmus bestimmt für $\ker(\phi_B)$ die geordnete Basis $((2,0,-1),(0,2,-1))$. Ebenso gilt $\mathrm{Hau}(\phi_A, 3) = \ker(\phi_C)$ mit

$$C \quad = \quad A - 3 \cdot \mathrm{I}^{(3)} \quad = \quad \begin{pmatrix} 22 & 34 & 18 \\ -14 & -22 & -10 \\ -4 & -6 & -4 \end{pmatrix} \quad ,$$

und der Gaußalgorithmus liefert das einelementige Erzeugendensystem $((-7,4,1))$. Fasst man die Vektoren zur geordneten Basis

$$\mathscr{A} \quad = \quad ((2,0,-1),(0,2,-1),(-7,4,1))$$

zusammen, so gilt

$$\mathscr{T}^{\mathscr{A}}_{\mathscr{E}} \quad = \quad \begin{pmatrix} 2 & 0 & -7 \\ 0 & 2 & 4 \\ -1 & -1 & 1 \end{pmatrix} \quad \text{und} \quad \mathscr{T}^{\mathscr{E}}_{\mathscr{A}} \quad = \quad (\mathscr{T}^{\mathscr{A}}_{\mathscr{E}})^{-1} \quad = \quad \begin{pmatrix} -3 & -\frac{7}{2} & -7 \\ 2 & \frac{5}{2} & 4 \\ -1 & -1 & -2 \end{pmatrix}.$$

Mit der Transformationsformel erhalten wir

$$\mathscr{M}_{\mathscr{A}}(\phi_A) \quad = \quad \mathscr{T}^{\mathscr{E}}_{\mathscr{A}} \cdot A \cdot \mathscr{T}^{\mathscr{A}}_{\mathscr{E}} \quad = \quad \begin{pmatrix} 16 & 25 & 0 \\ -9 & -14 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Um zu sehen, wie man aus der Hauptraumzerlegung die Jordansche Normalform gewinnt, benötigen wir noch ein wenig mehr Theorie.

Definition 13.8 Ein Endomorphismus ϕ eines K-Vektorraums V wird ***nilpotent*** genannt, wenn eine Zahl $p \in \mathbb{N}$ mit $\phi^p = 0$ existiert.

Sei ϕ ein Endomorphismus und λ ein Eigenwert von ϕ mit algebraischer Vielfachheit $\mu_a(\lambda)$. Die Gleichung

$$\mathrm{Hau}(\phi,\lambda) \quad = \quad \ker((\phi - \lambda \mathrm{id}_V)^{\mu_a(\lambda)})$$

zeigt, dass $(\phi - \lambda\mathrm{id}_V)|_{\mathrm{Hau}(\phi,\lambda)}$ ein nilpotenter Endomorphismus ist. Es gilt nun

Satz 13.9 Sei V ein endlich-dimensionaler K-Vektorraum und $\phi : V \to V$ ein nilpotenter Endomorphismus. Dann gibt es eine Basis $(v_1, ..., v_n)$ von V, so dass für $1 \le k \le n$ jeweils

$$\phi(v_k) = v_{k-1} \quad \text{oder} \quad \phi(v_k) = 0_V \qquad \text{gilt.}$$

Eine solche Basis wird ***Jordanbasis*** von V bezüglich ϕ genannt.

Um die Jordansche Normalform eines Endomorphismus zu bestimmen, benötigen wir für jeden Hauptraum eine Basis mit genau dieser Eigenschaft. Leider können wir den Satz aus Zeitgründen hier nicht beweisen.

Statt dessen schauen wir uns an, wie man eine solche Basis in der Praxis findet. Sei $\phi : V \to V$ ein nilpotenter Endomorphismus und $p \in \mathbb{N}$ minimal mit $\phi^p = 0$. Definieren wir $V_k = \ker(\phi^k)$ für $0 \le k \le$, dann gilt

$$\{0_V\} = V_0 \subsetneq V_1 \subsetneq ... \subsetneq V_p = V.$$

Denn wäre $V_k = \ker(\phi^k)$ gleich $V_{k-1} = \ker(\phi^{k-1})$ für ein k, dann würde man durch Komposition mit ϕ^{p-k} die Gleichung $\ker(\phi^{p-1}) = \ker(\phi^p) = V$ und somit $\phi^{p-1} = 0$ erhalten, im Widerspruch zur Minimalität der Zahl p.

Wählt man nun einen Vektor $v_p \in V_p \setminus V_{p-1}$ und definiert $v_k = \phi^{p-k}(v_p)$ für $1 \le k \le p$, dann erhält man ein System $(v_1, ..., v_p)$ mit der im Satz angegebenen Eigenschaft; genauer gilt

$$\phi(v_k) = v_{k-1} \quad \text{für} \quad 2 \le k \le p \qquad \text{und} \qquad \phi(v_1) = 0_V.$$

Ein solches System $(v_1, ..., v_p)$ von Vektoren bezeichnet man als ***Jordankette***.

Im günstigsten Fall spannt eine solche Jordankette bereits den gesamten Vektorraum V auf. Die Darstellungsmatrix von ϕ bezüglich der geordneten Basis $(v_1, \dots, v_p)$ ist dann eine Jordanmatrix.

Beispiel: Wir betrachten den Endomorphismus $\phi : \mathbb{R}^3 \to \mathbb{R}^3$, $v \mapsto Av$ gegeben durch die Matrix

$$A = \begin{pmatrix} 2 & 4 & -3 \\ 1 & 3 & -2 \\ 3 & 7 & -5 \end{pmatrix}$$

Es gilt

$$A^2 = \begin{pmatrix} -1 & -1 & 1 \\ -1 & -1 & 1 \\ -2 & -2 & 2 \end{pmatrix} \quad \text{und} \quad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} ,$$

also ist der Endomorphismus ϕ tatsächlich nilpotent. Mit dem Gauß-Algorithmus berechnet man

$$\begin{aligned} V_1 &= \ker(A) &&= \operatorname{lin}\{(1,1,2)\} \\ V_2 &= \ker(A^2) &&= \operatorname{lin}\{(-1,1,0),(1,0,1)\} \\ V_3 &= \mathbb{R}^3 &&= \operatorname{lin}\{e_1,e_2,e_3\}. \end{aligned}$$

Man überprüft unmittelbar, dass der Vektor $v_3 = e_1$ in $V_3 \setminus V_2$ liegt. Somit ist (v_1, v_2, v_3) gegeben durch $v_2 = \phi(v_3) = (2,1,3)$ und $v_1 = \phi^2(v_3) = \phi(v_2) = (-1,-1,-2)$ eine Jordankette, die wegen $\dim \mathbb{R}^3 = 3$ bereits den gesamten Vektorraum aufspannt. Trägt man die drei Vektoren als Spalten in eine Matrix ein

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

so erhält man die Jordan-Matrix

$$T^{-1}AT = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Im Allgemeinen liefert der folgende Algorithmus eine Jordanbasis. Sei V ein endlich-dimensionaler K-Vektorraum und $\phi \in \operatorname{End}_K(V)$ nilpotent. Sei $p \in \mathbb{N}$ minimal mit $\phi^p = 0$. Für jedes $k \in \{0, \dots, p\}$ sei wieder $V_k = \ker(\phi^k)$. Außerdem setzen wir $U_1 = V_1$ und wählen U_k für $2 \le k \le p$ jeweils so, dass

$$V_k = V_{k-1} \oplus U_k \qquad \text{erfüllt ist.}$$

Danach führt man die folgenden Schritte aus.

(i) Setze $\mathscr{B} = ()$ und $U = \{0_V\}$.

(ii) Sei $m \in \{0, \dots, p\}$ maximal mit der Eigenschaft $U_m \not\subseteq U$. Wähle einen Vektor $v \in U_m \setminus U$.

(iii) Erweitere das Tupel $\mathscr{B}$ um die Kette $(\phi^{m-1}(v), \phi^{m-2}(v), \dots, \phi(v), v)$ und ersetze U durch $\operatorname{lin} \mathscr{B}$.

(iv) Ist $U = V$, dann ist $\mathscr{B}$ bereits eine Jordanbasis von V. Ansonsten gehe zurück zu Schritt (ii).

Beispiel: Wir bestimmen eine Jordan-Basis für $\phi : \mathbb{R}^6 \to \mathbb{R}^6$, $v \mapsto Av$ gegeben durch die Matrix

$$A = \begin{pmatrix} 11 & 29 & -29 & -24 & 5 & 16 \\ -4 & -8 & 9 & 7 & -1 & -5 \\ 7 & 18 & -18 & -15 & 3 & 10 \\ -5 & -12 & 13 & 10 & -2 & -7 \\ 22 & 55 & -56 & -46 & 9 & 31 \\ -2 & -8 & 8 & 6 & -2 & -4 \end{pmatrix}.$$

Es gilt

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 11 & -11 & -9 & 2 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 11 & -11 & -9 & 2 & 6 \\ -4 & -11 & 11 & 9 & -2 & -6 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

und $A^3 = 0$, also ist ϕ in der Tat ein nilpotenter Endomorphismus. Man berechnet

$$\begin{aligned} V_1 &= \ker(A) &&= \operatorname{lin}\{(1,0,0,0,1,-1),(0,1,0,1,-1,0),(0,0,1,1,1,3)\} \\ V_2 &= \ker(A^2) &&= \operatorname{lin}\{(3,0,0,0,0,-2),(0,6,0,0,0,-11),(0,0,6,0,0,11),(0,0,0,2,0,3),(0,0,0,0,3,-1)\} \\ V_3 &= \ker(A^3) &&= \mathbb{R}^6 = \operatorname{lin}\{e_1,e_2,e_3,e_4,e_5,e_6\}. \end{aligned}$$

Setzt man $U_1 = V_1$, $U_2 = \operatorname{lin}\{(3,0,0,0,0,-2),(0,6,0,0,0,-11)\}$ und $U_3 = \operatorname{lin}\{(1,0,0,0,0,0)\}$, dann gilt

$$V_2 = V_1 \oplus U_2 \qquad \text{und} \qquad V_3 = V_2 \oplus U_3 \quad ,$$

wie man leicht nachrechnet (indem man zum Beispiel überprüft, dass die $3+2+1=6$ Vektoren insgesamt den $\mathbb{R}^6$ aufspannen). Führt man nun die Schritte (i) bis (iv) mit dem Vektor $v = e_1 = (1,0,0,0,0,0)$ aus, so erhält man $\mathscr{B} = (v_1, v_2, v_3)$ bestehend aus den Vektoren $v_3 = e_1$, $v_2 = \phi(v_3) = (11,-4,7,-5,22,-2)$ und $v_1 = \phi(v_2) = (0,4,0,4,-4,0)$.

Im nächsten Durchlauf überprüft man, dass der Vektor $v = (3,0,0,0,0,-2)$ nicht in $U = \operatorname{lin} \mathscr{B}$ liegt und erweitert $\mathscr{B}$ durch die Vektoren (v_4, v_5) gegeben durch $v_5 = v$ und $v_4 = \phi(v_5) = (1,-2,1,-1,4,2)$. Im dritten Durchlauf überprüft man $v = (1,0,0,0,1,-1)$ und erweitert $\mathscr{B}$ um den Vektor $v_6 = v$. Insgesamt ist dann durch $\mathscr{B} = (v_1, v_2, v_3, v_4, v_5, v_6)$ eine Jordanbasis gegeben. Trägt man diese Vektoren als Spalten in eine Matrix ein

$$T = \begin{pmatrix} 0 & 11 & 1 & 1 & 3 & 1 \\ 4 & -4 & 0 & -2 & 0 & 0 \\ 0 & 7 & 0 & 1 & 0 & 0 \\ 4 & -5 & 0 & -1 & 0 & 0 \\ -4 & 22 & 0 & 4 & 0 & 1 \\ 0 & -2 & 0 & 2 & -2 & -1 \end{pmatrix}$$

so erhält man die Matrix in Jordanscher Normalform

$$T^{-1}AT \quad = \quad \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

bestehend aus drei Jordanblöcken.

Ohne Beweis bemerken wir noch, dass für einen nilpotenten Endomorphismus $\phi \in \mathrm{End}_K(V)$ die Anzahl a_s der Jordanblöcken einer bestimmten Größe $s \in \mathbb{N}$ auch ermittelt werden kann, ohne dass eine Jordanbasis von V bestimmt werden muss. Setzt man $d_k = \dim\ker(\phi^k)$ für alle $k \in \mathbb{N}_0$, dann erhält man die Anzahl a_s durch die Formel

$$a_s \quad = \quad 2d_s - d_{s-1} - d_{s+1}.$$

In unserem Beispiel ist $d_0 = 0$, $d_1 = 3$, $d_2 = 5$ und $d_s = 6$ für alle $s \geq 3$. Es gilt somit

$$\begin{array}{lllllll} a_1 & = & 2d_1 - d_0 - d_2 & = & 6-0-5 & = & 1 \\ a_2 & = & 2d_2 - d_1 - d_3 & = & 10-3-6 & = & 1 \\ a_3 & = & 2d_3 - d_2 - d_4 & = & 12-5-6 & = & 1 \\ a_s & = & 2d_s - d_{s-1} - d_{s+1} & = & 12-6-6 & = & 0 \quad \text{für alle} \quad s \geq 4 \quad , \end{array}$$

was durch die Jordansche Normalform $T^{-1}AT$ von oben bestätigt wird: Die Matrix besteht aus je einem Jordanblock der Größe 1, 2 bzw. 3.

Nach diesen Vorbereitungen können wir nun ein Verfahren angeben, dass zu einem beliebigen Endormorphismus ϕ auf einem endlich-dimensionalen K-Vektorraum mit zerfallendem charakteristischen Polynom χ_ϕ eine Basis $\mathscr{B}$ liefert, so dass die Darstellungsmatrix $\mathscr{M}_{\mathscr{B}}(\phi)$ in Jordanscher Normalform vorliegt.

(i) Bestimme die Eigenwerte $\lambda_1, ..., \lambda_r$ von ϕ und für jeden Eigenwert den zugehörigen Hauptraum $W_i = \mathrm{Hau}(\phi, \lambda_i)$.

(ii) Die Endomorphismen $\psi_i = \phi|_{W_i} - \lambda_i \mathrm{id}_{W_i}$ sind nilpotent. Bestimme für jedes i eine Jordan-Basis $\mathscr{B}_i$ wie oben beschrieben.

(iii) Dann ist $\mathscr{B} = \mathscr{B}_1 \cup ... \cup \mathscr{B}_r$ eine Basis mit der gewünschten Eigenschaft.

Beispiel: Wir betrachten noch einmal den Endomorphismus $\phi : \mathbb{R}^3 \to \mathbb{R}^3$, $v \mapsto Av$ gegeben durch die Matrix

$$A = \begin{pmatrix} 25 & 34 & 18 \\ -14 & -19 & -10 \\ -4 & -6 & -1 \end{pmatrix}$$

von Seite 99. Wir hatten dort bereits nachgerechnet, dass $\lambda_1 = 1$ und $\lambda_2 = 3$ die einzigen Eigenwerte von ϕ sind, mit zugehörigen Haupträumen

$$W_1 = \mathrm{Hau}(\phi, 1) = \mathrm{lin}\,\{(2,0,-1),(0,2,-1)\} \quad \text{und} \quad W_2 = \mathrm{Hau}(\phi,3) = \mathrm{lin}\,\{(-7,4,1)\}.$$

Für den Raum W_2 haben wir schon eine passende Basis, nämlich $\mathscr{B}_2 = (v_3)$ mit $v_3 = (-7,4,1)$. Auf dem Raum W_1 betrachten wir den nilpotenten Endomorphismus $\psi_1 = \phi|_{W_1} - \mathrm{id}_{W_1}$. Setzen wir $V_k = \ker(\psi_1^k)$ für $0 \le k \le 2$, dann gilt $V_0 = \{0_{W_1}\}$ und $V_2 = W_1$, eine kurze Rechnung ergibt außerdem

$$V_1 = \ker(A - \mathrm{I}^{(3)}) = \mathrm{lin}\,\{(5,-3,-1)\}.$$

Weil der Vektor $v = (2,0,-1)$ in $V_2 \setminus V_1$ liegt, können wir ihn zur Konstruktion einer Jordankette verwenden: Wir definieren $v_2 = v$, $v_1 = \psi_1(v_2) = (30,-18,-6)$ und haben somit durch $\mathscr{B}_1 = (v_1, v_2)$ eine passende Basis für W_1 gefunden. Insgesamt ist damit $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2 = (v_1, v_2, v_3)$ eine Basis des Vektorraums $\mathbb{R}^3$ mit der gewünschten Eigenschaft. Tragen wir die drei Vektoren als Spalten in eine Matrix ein,

$$T = \begin{pmatrix} 30 & 2 & -7 \\ -18 & 0 & 4 \\ -6 & -1 & 1 \end{pmatrix}$$

so erhalten wir

$$T^{-1}AT = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Literatur

[Bo] S. Bosch, *Lineare Algebra*. Springer-Lehrbuch, Berlin 2006.

[Fi] G. Fischer, *Lernbuch Lineare Algebra und Analytische Geometrie*. Vieweg-Teubner, Wiesbaden 2011.

[Jä] K. Jänich, *Lineare Algebra*. Springer-Verlag, Berlin 2001.

[dJ] T. de Jong, *Lineare Algebra*. Pearson-Studium, München 2013.