Notizen zur Vorlesung der Linearen Algebra und Analytischen Geometrie I, II

Jürg Kramer

Wintersemester 2001/02
Sommersemester 2002

Inhaltsverzeichnis

0 Gruppen, Ringe, Körper **5**

1 Vektorräume **7**
1.1 Definition . . . 7
1.2 Beispiele . . . 8
1.3 Unterräume . . . 11
1.4 Lineare Abhängigkeit . . . 13
1.5 Basen . . . 14
1.6 Dimension . . . 17
1.7 Koordinaten . . . 21

2 Lineare Abbildungen **23**
2.1 Definitionen . . . 23
2.2 Operationen mit linearen Abbildungen . . . 30
2.3 Matrizen linearer Abbildungen . . . 32
2.4 Operationen mit Matrizen . . . 39
2.5 Basiswechsel (Basistransformationsmatrizen) . . . 45

3 Lineare Gleichungssysteme **47**
3.1 Die Problemstellung . . . 47
3.2 Zum Existenzproblem . . . 48
3.3 Zur Lösungsmannigfaltigkeit . . . 49
3.4 Lösungsverfahren I . . . 51
3.5 Determinanten . . . 54

4 Diagonalisierbarkeit und Normalformen **65**
4.1 Problemstellung . . . 65
4.2 Eigenwerte, Eigenvektoren . . . 66
4.3 Diagonalisierbarkeit . . . 70
4.4 Minimalpolynome . . . 77
4.5 Die Jordansche Normalform . . . 81

5 Euklidische und unitäre Vektorräume **89**
5.1 Skalarprodukte 89
5.1.1 Der Fall $K = \mathbf{R}$ 89
5.1.2 Der Fall $K = \mathbf{C}$ 90
5.2 Winkelmessung und Orthogonalität 94
5.3 Orthogonale und unitäre Abbildungen 96
5.4 Selbstadjungierte Abbildungen 99

6 Affine Geometrie **105**
6.1 Operation einer Gruppe auf einer Menge 105
6.2 Affine Räume 108
6.3 Affine Unterräume 110
6.4 Parallelismus 114
6.5 Affine Basen, affine Koordinaten 115
6.6 Affine Abbildungen 117
6.7 Matrizielle Darstellung affiner Abbildungen 122
6.8 Der Hauptsatz der affinen Geometrie 123

7 Projektive Geometrie **131**
7.1 Einführung 131
7.2 Der projektive Raum 133
7.3 Projektive Unterräume 134
7.4 Projektive Basen, homogene Koordinaten 137
7.5 Projektive Abbildungen 139
7.6 Matrizielle Darstellung projektiver Abbildungen 141

8 Quadriken **145**
8.1 Quadratische Formen 145
8.2 Quadriken 147

Kapitel 0

Gruppen, Ringe, Körper

Dieses Kapitel befindet sich noch in Arbeit.

Bemerkung.

Die hier vorliegenden Notizen zur Vorlesung „Lineare Algebra und Analytische Geometrie I“ sind als Unterstützung beim Nacharbeiten des Vorlesungsstoffes zu sehen. Sie erheben nicht den Anspruch auf Vollständigkeit, d.h. es gibt durchaus Inhalte, die in der Vorlesung behandelt wurden, sich aber in den Notizen nicht wiederfinden. Ebenso kann Fehlerfreiheit nicht garantiert werden.

Kapitel 1

Vektorräume

1.1 Definition

Definition. Es sei K ein Körper. Eine nichtleere Menge V zusammen mit einer Addition

$$v, w \longmapsto v + w \;\; (v, w \in V)$$

und einer Multiplikation mit Skalaren

$$\lambda, v \longmapsto \lambda \cdot v \;\; (\lambda \in K, v \in V)$$

heißt *Vektorraum über K* oder *K-Vektorraum*, falls gilt:

(1) V ist bzgl. der Addition eine abelsche Gruppe.

(2) Für alle $\lambda, \mu \in K$ und $v, w \in V$ bestehen die Axiome:

(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) $1 \cdot v = v$.

Bemerkung. Das Symbol 0 bezeichnet im folgenden sowohl das Nullelement von K als auch das neutrale Element der abelschen Gruppe V.

Lemma. *Es gilt:*

(i) *Das neutrale Element* 0 *ist eindeutig bestimmt.*

(ii) $0 \cdot v = 0 \;\; (v \in V)$.

(iii) *Das Inverse $-v$ zu $v \in V$ ist eindeutig bestimmt.*

(iv) $(-\lambda) \cdot v = -(\lambda \cdot v) \;\; (\lambda \in K,\, v \in V)$.

(v) $\lambda \cdot (-v) = -(\lambda \cdot v) \;\; (\lambda \in K,\, v \in V)$.

(vi) *Aus $\lambda \cdot v = 0$ $(\lambda \in K,\, v \in V)$ folgt $\lambda = 0$ oder $v = 0$.*

Beweis.

(i) Seien $0, 0^*$ zwei neutrale Elemente. Es folgt $0 + 0^* = 0$ und $0^* + 0 = 0^*$, somit $0^* = 0^* + 0 = 0 + 0^* = 0$.

(ii) Setze $w := 0 \cdot v \ \in V$. Es folgt $w = 0 \cdot v = (0+0) \cdot v \overset{(2a)}{=} 0 \cdot v + 0 \cdot v = w + w$, also $w + (-w) = w + w + (-w)$, somit $0 = w + (w + (-w)) = w + 0 = w$ bzw. $0 \cdot v = 0$.

(iii) Seien $-v, v^*$ Inverse zu $v \in V$. Es folgt $v + (-v) = 0 = v + v^*$, also $v^* = v^* + 0 = v^* + (v + (-v)) = (v^* + v) + (-v) = 0 + (-v) = -v$.
Beachte: Die Eindeutigkeit des Inversen hat zur Folge, daß die Gleichung $v + x = w$ ($v, w \in V$gegeben; $x \in V$gesucht) eine eindeutige Lösung hat, nämlich $x = w + (-v)$. Wir bezeichnen den Vektor $w + (-v)$ durch $w - v$ und nennen ihn Differenzvektor.

(iv) Nach Eigenschaft (ii) und mit (2a) folgt $0 = 0 \cdot v = (\lambda + (-\lambda)) \cdot v = \lambda \cdot v + (-\lambda) \cdot v$, somit ist $(-\lambda) \cdot v$ Inverses zu $\lambda \cdot v$; nach Eigenschaft (iii) ist dieses eindeutig bestimmt, also $-(\lambda \cdot v) = (-\lambda) \cdot v$.
Beachte: Speziell für $\lambda = 1$ haben wir $-(1 \cdot v) = (-1) \cdot v$, also $(-1) \cdot v = -v$.

(v) Wie Eigenschaft (ii) beweist man $\lambda \cdot 0 = 0$ $(\lambda \in K)$; damit und mit (2b) folgt: $0 = \lambda \cdot 0 = \lambda \cdot (v + (-v)) = \lambda \cdot v + \lambda \cdot (-v)$. Somit ist $\lambda \cdot (-v)$ Inverses zu $\lambda \cdot v$, nach Eigenschaft (iii) ist dieses eindeutig bestimmt, also $-(\lambda \cdot v) = \lambda \cdot (-v)$.

(vi) Sei $\lambda \neq 0$, also existiert $\lambda^{-1} \in K$. Mit (2c) und (2d) folgt $v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0 = 0$.

1.2 Beispiele

(i) **Der Vektorraum $\mathbf{R}^n$:** Wir betrachten die Menge

$$\mathbf{R}^n = \left\{ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \middle| \ \xi_1, \dots, \xi_n \in \mathbf{R} \right\}$$

der sogenannten *Spaltenvektoren.*

Die Addition der Spaltenvektoren $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \in \mathbf{R}^n$ wird komponentenweise erklärt, d.h.:

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} + \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} := \begin{pmatrix} \xi_1 + \eta_1 \\ \vdots \\ \xi_n + \eta_n \end{pmatrix}.$$

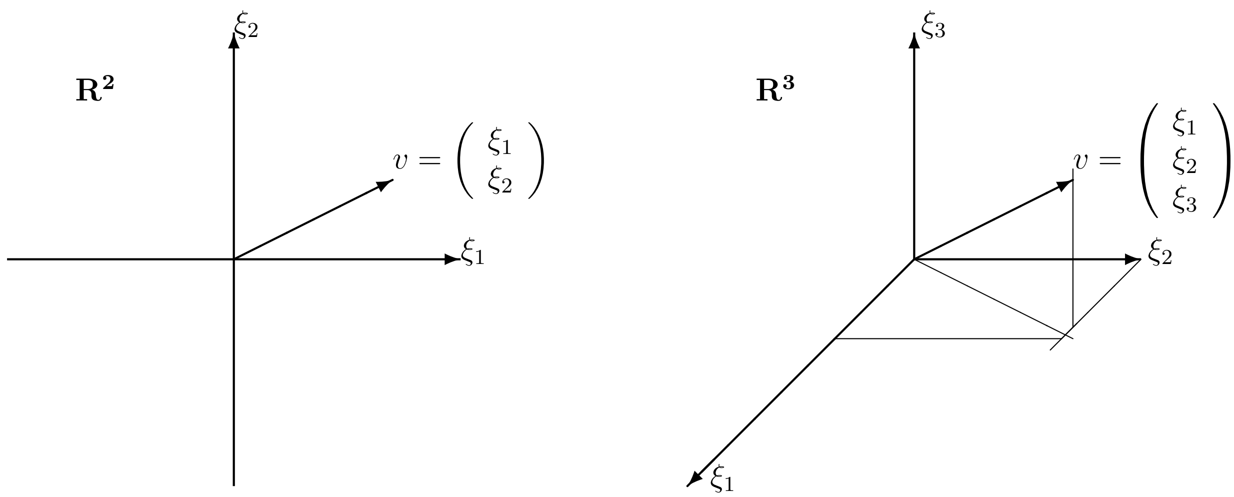
Mit dem Nullvektor $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ und dem zu $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ additiv Inversen $-\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} :=$ $\begin{pmatrix} -\xi_1 \\ \vdots \\ -\xi_n \end{pmatrix}$ wird $\mathbf{R}^n$ zu einer abelschen Gruppe.

Die Multiplikation eines Spaltenvektors $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ mit einem Skalar $\lambda \in \mathbf{R}$ wird ebenfalls komponentenweise erklärt, d.h.

$$\lambda \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} := \begin{pmatrix} \lambda\xi_1 \\ \vdots \\ \lambda\xi_n \end{pmatrix}.$$

Es ist leicht zu überprüfen, daß die Axiome (2a)-(2d) erfüllt sind. Damit wird $\mathbf{R}^n$ zu einem $\mathbf{R}$-Vektorraum, dem *Vektorraum der Spaltenvektoren* (mit n Komponenten) bzw. *Vektorraum der n-Tupel.*

Geometrische Darstellung.



Bemerkung. Anstelle der Spaltenvektoren $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ könnten auch *Zeilenvektoren* $(\xi_1, \ldots, \xi_n)$ betrachtet und analog ein Vektorraum von Zeilenvektoren definiert werden. Unter dem $\mathbf{R}^n$ soll in dieser Vorlesung in der Regel der Vektorraum der Spaltenvektoren betrachtet werden; sollte einmal der Vektorraum der Zeilenvektoren eine Rolle spielen, so werden wir diesen speziell bezeichnen.

(ii) Betrachte m lineare Gleichungen mit n Unbekannten $\xi_1, \ldots, \xi_n$ der Form:

$$\left.\begin{array}{l} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \cdots + \alpha_{1n}\xi_n = 0 \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \cdots + \alpha_{2n}\xi_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \cdots + \alpha_{mn}\xi_n = 0 \end{array}\right\} (S).$$

Hierbei sind die *Koeffizienten* $\alpha_{kj} \in \mathbf{R}$ $(k = 1, \ldots, m;\ j = 1, \ldots, n)$. Das lineare Gleichungssystem (S) wird *homogenes lineares Gleichungssystem* genannt, da rechter Hand nur Nullen vorkommen; ist dies nicht der Fall, so spricht man von einem *inhomogenen linearen Gleichungssystem*. Wir betrachten nun die Lösungsmenge

$$V := \left\{ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \;\middle|\; \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \text{ erfüllt } (S) \right\}.$$

Man verifiziert leicht:

Sind $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \in V$, so sind auch $\begin{pmatrix} \xi_1 + \eta_1 \\ \vdots \\ \xi_n + \eta_n \end{pmatrix}, \begin{pmatrix} \lambda\xi_1 \\ \vdots \\ \lambda\xi_n \end{pmatrix} \in V$ $(\lambda \in \mathbf{R})$.

Damit ist auf V eine Addition und eine Multiplikation mit Skalaren definiert. Es ist nicht schwierig, zu zeigen, daß V die Axiome eines Vektorraumes erfüllt.

(iii) $V := \{f : \mathbf{R} \longrightarrow \mathbf{R} \text{ stetig}\}$. Sind $f, g \in V$ und $\lambda \in \mathbf{R}$, so setzt man

$$\begin{aligned} (f+g)(t) &:= f(t) + g(t)\ (t \in \mathbf{R}), \\ (\lambda \cdot f)(t) &:= \lambda f(t)\ (t \in \mathbf{R}). \end{aligned}$$

Aus der Analysis übernehmen wir: Sind f, g stetige Funktionen auf $\mathbf{R}$, so sind die Funktionen $f+g$ und λf stetig auf $\mathbf{R}$. Damit haben wir auf der Menge V eine Addition und eine Multiplikation mit Skalaren definiert. Man überprüft ohne Mühe, daß V damit zu einem Vektorraum wird, dem *Vektorraum der auf stetigen Funktionen.*

(iv) $V := \{p : \mathbf{R} \longrightarrow \mathbf{R} \mid \exists d \in \mathbf{N} : p(t) = \alpha_d t^d + \ldots + \alpha_1 t + \alpha_0\ ; \alpha_0, \ldots, \alpha_d \in \mathbf{R}\}$. Ein Element $p \in V$ heißt *Polynom vom Grad d*, falls $\alpha_d \neq 0$ ist.
Wie in Beispiel (iii) läßt sich auf V eine Addition und eine Multiplikation mit Skalaren erklären. Sind nämlich $p, q \in V$ und $\lambda \in \mathbf{R}$ und

$$\begin{aligned} p(t) &= \alpha_d t^d + \ldots + \alpha_1 t + \alpha_0 \\ q(t) &= \beta_d t^d + \ldots + \beta_1 t + \beta_0\ , \end{aligned}$$

so gilt

$$\begin{aligned} (p+q)(t) &= (\alpha_d + \beta_d)t^d + \ldots + (\alpha_1 + \beta_1)t + (\alpha_0 + \beta_0) \\ (\lambda \cdot p)(t) &= (\lambda\alpha_d)t^d + \ldots + (\lambda\alpha_1)t + (\lambda\alpha_0). \end{aligned}$$

Wiederum überprüft man mühelos, daß V ein Vektorraum ist, der *Vektorraum der Polynome* (*in einer Variablen über den reellen Zahlen*).

1.3 Unterräume

Für das folgende sei ein K-Vektorraum V festgelegt. Wir definieren:

Definition. Eine nichtleere Teilmenge $U \subseteq V$ heißt ein *linearer Unterraum* oder *Teilraum* von V, falls gilt:

(3) $v, w \in U \;\Rightarrow\; v + w \in U$,

(4) $\lambda \in K,\ v \in U \;\Rightarrow\; \lambda v \in U$.

Lemma. *Jeder Unterraum U eines K-Vektorraumes V ist selbst wieder ein K-Vektorraum.*

Beweis. Mit Hilfe von (3) wird auf U eine additive Struktur definiert; diese ist assoziativ und kommutativ, da sie ja von einer assoziativen und kommutativen Struktur auf V induziert ist.
Ist $v \in U$, so folgt mit (4) (setze $\lambda = -1$) $-v = (-1)v \in U$, d.h. mit $v \in U$ ist auch das Inverse $-v \in U$.
Da weiter U nichtleer ist, existiert mindestens ein $v \in U$ und damit auch $-v \in U$; also folgt mit (3), daß $0 = v + (-v) \in U$ ist. Damit ist U eine abelsche Gruppe.
Die durch (4) festgelegte Multiplikation mit Skalaren erfüllt automatisch die Axiome (2a)-(2d), weil sie es ja sogar auf V tut. Damit ist U ein K-Vektorraum. □

Lemma. *Es seien U_1, U_2 Unterräume von V. Dann ist der Durchschnitt $U_1 \cap U_2$ ebenfalls ein Unterraum von V.*

Beweis: Wir haben für $U_1 \cap U_2$ die Axiome (3),(4) zu bestätigen. Es seien also $v, w \in U_1 \cap U_2$ und $\lambda \in K$. Da insbesondere $v, w \in U_1$ ist, folgt mit (3) bzw. (4), daß $v + w \in U_1$ bzw. $\lambda v \in U_1$. Analog schließt man $v + w \in U_2$ bzw. $\lambda v \in U_2$. Insgesamt ergibt sich $v + w \in U_1 \cap U_2$ bzw. $\lambda v \in U_1 \cap U_2$. □

Beispiele.

(i) Seien $V = \mathbf{R}^n$ der $\mathbf{R}$-Vektorraum aller n-Tupel und U der $\mathbf{R}$-Vektorraum aller n-Tupel, welche ein gegebenes lineares homogenes Gleichungssystem (m Gleichungen, n Unbekannte) lösen.
Dann ist U ein Teilraum von V.

(ii) Sei V der $\mathbf{R}$-Vektorraum der auf ˚stetigen Funktionen und U der $\mathbf{R}$-Vektorraum der Polynome.
Dann ist U ein Unterraum von V.

(iii) Geometrisch:
- Unterräume von $\mathbf{R}^2$: $\{0\}$, Geraden durch 0, $\mathbf{R}^2$.
- Unterräume von $\mathbf{R}^3$: $\{0\}$, Geraden durch 0, Ebenen durch 0, $\mathbf{R}^3$.
Durchschnitt:
- Schnitt zweier Geraden durch $\{0\}$ in $\mathbf{R}^2$.
- Schnitt zweier Ebenen durch $\{0\}$ in $\mathbf{R}^3$.
- Schnitt einer Geraden durch $\{0\}$ mit einer Ebene durch $\{0\}$ in $\mathbf{R}^3$.

Definition. Es seien U_1, U_2 Unterräume von V. Dann heißt die Menge

$$U_1 + U_2 := \{ u_1 + u_2 \,|\, u_1 \in U_1,\ u_2 \in U_2 \}$$

die *Summe der Unterräume* U_1, U_2.

Beispiel. Es sei $V = \mathbf{R}^3$ und

$$U_1 = \left\{ \begin{pmatrix} \xi_1 \\ 0 \\ 0 \end{pmatrix} \middle| \ \xi_1 \in \mathbf{R} \right\},\ U_2 = \left\{ \begin{pmatrix} 0 \\ \xi_2 \\ 0 \end{pmatrix} \middle| \ \xi_2 \in \mathbf{R} \right\}.$$

Dann folgt

$$U_1 + U_2 = \left\{ \begin{pmatrix} \xi_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \xi_2 \\ 0 \end{pmatrix} \middle| \ \xi_1, \xi_2 \in \mathbf{R} \right\} = \left\{ \begin{pmatrix} \xi_1 \\ \xi_2 \\ 0 \end{pmatrix} \middle| \ \xi_1, \xi_2 \in \mathbf{R} \right\}.$$

Lemma. *Es seien U_1, U_2 Unterräume von V. Dann ist ihre Summe $U_1 + U_2$ ebenfalls ein Unterraum von V.*

Beweis. Wir haben für $U_1 + U_2$ die Axiome (3),(4) zu bestätigen. Es seien also $v, w \in U_1 + U_2$, d.h. es existieren $v_1, w_1 \in U_1$ und $v_2, w_2 \in U_2$ mit der Eigenschaft

$$v = v_1 + v_2,\ w = w_1 + w_2.$$

Mit Hilfe der Assoziativität und der Kommutativität der Addition folgt nun

$$v + w = (v_1 + v_2) + (w_1 + w_2) = (v_1 + w_1) + (v_2 + w_2) \ \in U_1 + U_2.$$

Ist weiter $\lambda \in K$, so folgt mit (2b)

$$\lambda v = \lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2 \ \in U_1 + U_2.$$

Damit sind die Axiome (3) und (4) nachgewiesen. □

Definition. Es seien $v_1, \ldots, v_n$ endlich viele Vektoren eines Vektorraumes V. Ein Vektor

$$v \in \{\lambda_1 v_1 + \ldots + \lambda_n v_n \,|\, \lambda_1, \ldots, \lambda_n \in K\}$$

heißt *Linearkombination der Vektoren* $v_1, \ldots, v_n$.
Weiter wird die Menge aller Linearkombinationen von $v_1, \ldots, v_n$ durch $\langle v_1, \ldots, v_n \rangle$ bezeichnet, d.h.

$$\langle v_1, \ldots, v_n \rangle = \{\lambda_1 v_1 + \ldots + \lambda_n v_n \,|\, \lambda_1, \ldots, \lambda_n \in K\}.$$

Lemma. *Mit den vorhergehenden Bezeichnungen gilt: Die Menge $\langle v_1, \ldots, v_n \rangle$ aller Linearkombinationen ist der kleinste Unterraum von V, der die Vektoren $v_1, \ldots, v_n$ enthält.*

Beweis. Übungsaufgabe. □

Definition. Ist $M \subseteq V$ eine beliebige (nicht notwendigerweise) endliche Teilmenge von Vektoren, so heißt $v \in V$ *Linearkombination von* M, falls v eine Linearkombination *endlich* vieler Vektoren aus M ist. Weiter setzen wir

$$\langle M \rangle := \{v \in V \,|\, \exists v_1, \ldots, v_n \in M\,;\, \lambda_1, \ldots, \lambda_n \in K \,:\, v = \lambda_1 v_1 + \ldots + \lambda_n v_n\}.$$

Wiederum erkennt man $\langle M \rangle$ als Unterraum von V. Man bezeichnet $\langle v_1, \ldots, v_n \rangle$ bzw. $\langle M \rangle$ als die *lineare Hülle* oder den *linearen Span* von $v_1, \ldots, v_n$ bzw. M oder auch den durch $v_1, \ldots, v_n$ bzw. M *erzeugten linearen Unterraum* von V.

1.4 Lineare Abhängigkeit

Definition. Endlich viele Vektoren $x_1, \ldots, x_n$ von V heißen *linear unabhängig,* falls nur die triviale Linearkombination den Nullvektor darstellt, d.h. falls aus $\lambda_1 x_1 + \ldots + \lambda_n x_n = 0$ stets $\lambda_1 = \ldots = \lambda_n = 0$ folgt. Andernfalls heißen die Vektoren *linear abhängig.*

Beispiele.

(i) Die beiden Vektoren $\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \end{pmatrix} \in \mathbf{R}^2$ sind linear unabhängig, denn wir haben:

$$\lambda \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 8 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{array}{rcll} 2\lambda + 8\mu & = & 0 & \text{(I)} \\ 3\lambda + 7\mu & = & 0 & \text{(II)} \end{array} \implies$$

$$\begin{array}{rcll} 6\lambda + 24\mu & = & 0 & \text{(I')} = 3 \cdot \text{(I)} \\ 6\lambda + 14\mu & = & 0 & \text{(II')} = 2 \cdot \text{(II)} \end{array} \implies 10\mu = 0 \;\; \text{(I')-(II')}$$

Es ergibt sich $\mu = 0$ und $\lambda = 0$.

(ii) Die drei Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \in \mathbf{R}^3$ sind linear abhängig, denn es gilt:

$$(-1)\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 2 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} + (-1)\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Definition. Eine Teilmenge $M \subseteq V$ heißt *linear unabhängig*, falls je endlich viele Vektoren aus M linear unabhängig sind. Andernfalls heißt die Menge M *linear abhängig.*

Lemma. *Es sei $M \subseteq V$ eine Teilmenge, welche aus mindestens zwei Vektoren besteht. Dann gilt die Äquivalenz:*

$$\begin{array}{lcl} M \subseteq V \textit{ linear abhängig} & \iff & \exists x \in M; x_1, \ldots, x_n \in M \textit{verschieden;} \\ & & x_j \neq x \, (j = 1, \ldots, n) : \\ & & x = \sum\limits_{j=1}^{n} \lambda_j x_j \, (\lambda_1, \ldots, \lambda_n \in K). \end{array}$$

Beweis. $(\Leftarrow)$: Seien also $x \in M$, $x_1, \ldots, x_n \in M$ verschieden und $x_j \neq x\,(j = 1, \ldots, n)$ mit der Eigenschaft $x = \sum_{j=1}^n \lambda_j x_j$ $(\lambda_1, \ldots, \lambda_n \in K)$. Somit ist $1 \cdot x - \lambda_1 x_1 - \ldots - \lambda_n x_n = 0$ eine nichttriviale Darstellung des Nullvektors, d.h. M ist linear abhängig.

$(\Rightarrow)$: Es sei $M \subseteq V$ linear abhängig, also existieren $x_0, x_1, \ldots, x_n \in M$, $\lambda_0, \lambda_1, \ldots, \lambda_n \in K$ (nicht alle gleich Null) mit $\sum_{j=0}^n \lambda_j x_j = 0$.
O.B.d.A. darf $\lambda_0 \neq 0$ angenommen werden. Da die Vektoren $x_0, x_1, \ldots, x_n$ sämtlich verschieden sind, folgt mit $x = x_0$ und $\lambda'_j = \lambda_0^{-1}\lambda_j\,(j = 1, \ldots, n)$:

$$\exists x \in M;\ x_1, \ldots, x_n \in M \text{ verschieden};\ x_j \neq x\ (j = 1, \ldots, n) :$$

$$x = \sum_{j=1}^n \lambda'_j x_j\ (\lambda'_1, \ldots, \lambda'_n \in K).$$

□

1.5 Basen

Definition. Eine Teilmenge $\mathcal{B} \subseteq V$ heißt eine *Basis* von V, falls $\mathcal{B}$ linear unabhängig ist und $V = \langle \mathcal{B} \rangle$ gilt.

Lemma. *Es sei V ein K-Vektorraum und $\mathcal{B} \subseteq V$. Dann gilt:*

(i) *$\mathcal{B}$ ist genau dann eine Basis, wenn $\mathcal{B}$ ein minimales Erzeugendensystem ist.*

(ii) *$\mathcal{B}$ ist genau dann eine Basis, wenn $\mathcal{B}$ eine maximale Menge linear unabhängiger Vektoren ist.*

Beweis. Übungsaufgabe. □

Bemerkung. Als erstes wichtiges Resultat der linearen Algebra wird die Existenz einer Basis eines Vektorraumes gezeigt. Werden keine Voraussetzungen an V gestellt, so muß für den Beweis das sogenannte „Lemma von Zorn“ aus der Mengenlehre herangezogen werden. Dieses soll im folgenden präzise formuliert, aber nicht bewiesen werden.

Definition. Eine *Teilordnung* (oder *partielle Ordnung*) auf einer Menge $\mathfrak{M}$ ist eine Relation $m \leq m'$, die zwischen gewissen Elementen bestehen kann und für entsprechende m, m', $m'' \in \mathfrak{M}$ folgende Axiome erfüllt:

(i) Es gilt $m \leq m$.

(ii) Ist $m \leq m'$ und $m' \leq m''$, so ist $m \leq m''$.

(iii) Ist $m \leq m'$ und $m' \leq m$, so ist $m = m'$.

Eine Teilordnung heißt *Totalordnung* oder *Ordnung* von $\mathfrak{M}$, falls zusätzlich gilt:

(iv) Für alle m, $m' \in \mathfrak{M}$ gilt $m \leq m'$ oder $m' \leq m$.

Definition. Es sei $\mathfrak{M}$ eine teilweise geordnete Menge. Eine Teilmenge $\mathfrak{K} \subseteq \mathfrak{M}$ heißt *Kette*, falls $\mathfrak{K}$ total geordnet ist.

Definition. Es sei $\mathfrak{M}$ eine teilweise geordnete Menge und $\mathfrak{A} \subseteq \mathfrak{M}$ eine Teilmenge. Ein Element $m \in \mathfrak{M}$ heißt *obere Schranke von* $\mathfrak{A}$, wenn für alle $a \in \mathfrak{A}$ die Beziehung $a \leq m$ gilt.

Definition. Es sei $\mathfrak{M}$ eine teilweise geordnete Menge. Ein Element $m_0 \in \mathfrak{M}$ heißt *maximales Element*, wenn $\mathfrak{M}$ kein größeres Element besitzt, d.h. es gibt kein $m \in \mathfrak{M}$ mit $m_0 \leq m$ außer m_0 selbst.

Lemma von Zorn. Es sei $\mathfrak{M}$ eine teilweise geordnete Menge. Falls jede Kette $\mathfrak{K} \subseteq \mathfrak{M}$ eine obere Schranke besitzt, so hat $\mathfrak{M}$ ein maximales Element.

Beweis. Wir wollen den Beweis des Zorn'schen Lemmas hier nicht führen. Wir verweisen auf das Lehrbuch „Algebra“ von B. van der Waerden, Band 1, Kapitel 9, § 69. $\square$

Satz. Es sei V ein K-Vektorraum. Dann besitzt V eine Basis.

Beweis. Wir betrachten die Menge $\mathfrak{M}$, deren Elemente die linear unabhängigen Teilmengen von V sind. Mit Hilfe der Inklusion von Mengen läßt sich auf $\mathfrak{M}$ eine Teilordnung erklären: Sind nämlich M_1, $M_2 \in \mathfrak{M}$, so bedeutet $M_1 \leq M_2$, daß die Inklusion $M_1 \subseteq M_2$ besteht.

Es sei nun $\mathfrak{K} \subseteq \mathfrak{M}$ eine Kette. Wir wollen eine obere Schranke zu $\mathfrak{K}$ konstruieren. Dazu setzen wir

$$M := \bigcup_{K \in \mathfrak{K}} K ,$$

d.h. M ist die Vereinigung aller Elemente von $\mathfrak{K}$. Wir haben zunächst $M \in \mathfrak{M}$ zu zeigen. Dazu gilt es nachzuweisen, daß je endlich viele Vektoren $v_1, \ldots, v_n \in M$ linear unabhängig sind. Wir können annehmen, daß $v_j \in K_j$ mit $K_j \in \mathfrak{K}$ $(j = 1, \ldots, n)$ gilt. Da $\mathfrak{K}$ total geordnet ist, findet sich ein Index $j_0 \in \{1, \ldots, n\}$ derart, daß $v_1, \ldots, v_n \in K_{j_0}$ gilt. Wegen $K_{j_0} \in \mathfrak{K} \subseteq \mathfrak{M}$ ist die Menge K_{j_0} linear unabhängig; somit sind auch die Vektoren $v_1, \ldots, v_n \in K_{j_0}$ linear unabhängig. Damit haben wir $M \in \mathfrak{M}$. Offensichtlich gilt für alle $K \in \mathfrak{K}$ die Beziehung $K \leq M$, d.h. M ist in der Tat eine obere Schranke von $\mathfrak{K}$.

Nach dem Zorn'schen Lemma besitzt $\mathfrak{M}$ ein maximales Element M_0. Konstruktionsgemäß ist M_0 eine maximale, linear unabhängige Teilmenge von V. Nach dem Lemma zu Beginn dieses Abschnitts bildet M_0 somit eine Basis von V. $\square$

Lemma. *Es sei V ein K-Vektorraum mit endlicher Basis $\mathcal{B} = \{b_1, \ldots, b_n\}$, und $b \in V$ besitze die Darstellung*

$$b = \sum_{j=1}^{n} \mu_j b_j .$$

Falls ein $k \in \{1, \ldots, n\}$ mit $\mu_k \neq 0$ existiert, so ist auch die Menge

$$\mathcal{B}' := \{b_1, \ldots, b_{k-1}, b, b_{k+1}, \ldots, b_n\}$$

eine Basis von V.

Beweis. O.B.d.A. kann $k = 1$ angenommen werden; wir haben also $\mathcal{B}' = \{b, b_1, \ldots, b_n\}$ als Basis von V nachzuweisen.
Erzeugtheit: Es ist $V = \langle \mathcal{B} \rangle$ und $\langle \mathcal{B}' \rangle \subseteq \langle \mathcal{B} \rangle$. Da $b = \mu_1 b_1 + \mu_2 b_2 + \ldots + \mu_n b_n$ $(\mu_1 \neq 0)$, d.h. $b_1 = \mu_1^{-1} b - \mu_1^{-1} \mu_2 b_2 - \ldots - \mu_1^{-1} \mu_n b_n$ gilt, folgt $b_1 \in \langle \mathcal{B}' \rangle$, also $\langle \mathcal{B} \rangle \subseteq \langle \mathcal{B}' \rangle$ und somit die Gleichheit $\langle \mathcal{B}' \rangle = \langle \mathcal{B} \rangle = V$.
Lineare Unabhängigkeit: Man gehe aus von der Relation

$$\lambda b + \lambda_2 b_2 + \ldots + \lambda_n b_n = 0$$

mit unbestimmten Skalaren $\lambda, \lambda_2, \ldots, \lambda_n$. Darin ersetzt man $b = \sum_{j=1}^{n} \mu_j b_j$ und erhält

$$\lambda(\mu_1 b_1 + \mu_2 b_2 + \ldots + \mu_n b_n) + \lambda_2 b_2 + \ldots + \lambda_n b_n = 0,$$

d.h.

$$(\lambda \mu_1) b_1 + (\lambda_2 + \lambda \mu_2) b_2 + \ldots + (\lambda_n + \lambda \mu_n) b_n = 0.$$

Da die Vektoren $\{b_1, \ldots, b_n\}$ linear unabhängig sind, folgt aus der letzten Relation

$$\lambda \mu_1 = 0,\ \lambda_2 + \lambda \mu_2 = 0,\ \ldots, \lambda_n + \lambda \mu_n = 0.$$

Wegen $\mu_1 \neq 0$, folgt $\lambda = 0$ und somit auch $\lambda_2 = \ldots = \lambda_n = 0$, d.h. die Menge $\mathcal{B}'$ ist somit auch linear unabhängig. □

Satz (Austauschsatz von Steinitz). *Es sei V ein K-Vektorraum mit endlicher Basis $\mathcal{B} = \{b_1, \ldots, b_n\}$. Weiter seien $c_1, \ldots, c_k$ linear unabhängige Vektoren von V. Dann gilt $k \leq n$, und bei geeigneter Numerierung der Vektoren $b_1, \ldots, b_n$ ist auch*

$$\mathcal{B}' := \{c_1, \ldots, c_k, b_{k+1}, \ldots, b_n\}$$

eine Basis von V.

Beweis. Wir führen eine vollständige Induktion über k durch:
Induktionsverankerung: Ist $k = 1$, so ist $c_1 \in V$ nicht der Nullvektor, da c_1 sonst linear abhängig wäre. Deshalb besteht die nichttriviale Linearkombination $c_1 = \sum_{j=1}^{n} \mu_j b_j$; o.B.d.A. kann $\mu_1 \neq 0$ angenommen werden. Nach dem vorhergehenden Lemma ist dann

$$\mathcal{B}' = \{c_1; b_2, \ldots, b_n\}$$

eine Basis von V; insbesondere gilt $n \geq 1$.
Induktionsvoraussetzung: Ist $k \geq 2$ und sind $c_1, \ldots, c_{k-1} \in V$ linear unabhängig, so ist auch

$$\mathcal{B}'' = \{c_1, \ldots, c_{k-1}; b_k, \ldots, b_n\}$$

eine Basis von V; insbesondere gilt $n \geq k - 1$.
Induktionsschritt: Seien $k \geq 2$ und $c_1, \ldots, c_k \in V$ linear unabhängig. Nach geeigneter Umnumerierung ergibt sich aus der Induktionsvoraussetzung, daß

$$\mathcal{B}'' = \{c_1, \ldots, c_{k-1}; b_k, \ldots, b_n\}$$

eine Basis von V ist. Es folgt insbesondere $n \geq k-1$. Wäre $n = k-1$, so wäre c_k von den Vektoren $c_1, \ldots, c_{k-1}$ linear abhängig, im Widerspruch zur Voraussetzung; also gilt

$$n > k-1 \iff n \geq k.$$

Da nun $\mathcal{B}''$ eine Basis von V ist, gilt speziell

$$c_k = \mu_1 c_1 + \ldots + \mu_{k-1} c_{k-1} + \mu_k b_k + \ldots + \mu_n b_n$$

mit gewissen Skalaren $\mu_1, \ldots, \mu_n$. Da aber c_k nicht als Linearkombination der Vektoren $c_1, \ldots, c_{k-1}$ dargestellt werden kann, existiert mindestens ein $\mu_j \neq 0$ mit Index $j \in \{k, ..., n\}$; o.B.d.A. kann $j = k$ angenommen werden. Mit Hilfe des vorhergehenden Lemmas kann nun aber b_k gegen c_k ausgetauscht werden, und wir erhalten

$$\mathcal{B}' = \{c_1, \ldots, c_k; b_{k+1}, \ldots, b_n\}$$

als Basis von V. □

Nachtrag zum Austauschsatz von Steinitz. Es sei V ein K-Vektorraum mit nicht notwendigerweise endlicher Basis $\mathcal{B}$. Weiter seien $c_1, \ldots, c_k$ linear unabhängige Vektoren von V. Da sich die endlich vielen Vektoren $c_1, \ldots, c_k$ durch endlich viele Basisvektoren aus $\mathcal{B}$ linear kombinieren lassen, ergibt sich aus dem Austauschsatz von Steinitz sofort, daß die endlich vielen Vektoren $c_1, \ldots, c_k$ gegen k Basisvektoren von $\mathcal{B}$ ausgetauscht werden können und damit eine neue Basis $\mathcal{B}'$ von V gewonnen werden kann, welche $c_1, \ldots, c_k$ enthält. Ist insbesondere $\mathcal{B}$ unendlich, so ist auch $\mathcal{B}'$ unendlich.

1.6 Dimension

Definition. Es sei V ein K-Vektorraum und $\mathcal{B}$ eine Basis von V. Dann ist die *Dimension* $\dim_K V$ von V definiert durch

$$\dim_K V := \begin{cases} \#\mathcal{B} & \text{, falls } \mathcal{B} \text{ endliche} \\ \infty & \text{, falls } \mathcal{B} \text{ unendliche} \end{cases}$$

Menge ist.

Bemerkung. Der Dimensionsbegriff hängt gemäß obiger Definition von der Wahl einer Basis $\mathcal{B}$ in V ab. Wir werden im folgenden zeigen, daß der Dimensionsbegriff wohldefiniert, d.h. unabhängig von der Wahl einer Basis ist.

Satz. *Die Dimension* $\dim_K V$ *eines* K*-Vektorraumes* V *ist wohldefiniert, d.h. unabhängig von der Wahl einer Basis in* V*.*

Beweis. Es sei zunächst V ein K-Vektorraum und $\mathcal{B}$ eine unendliche Basis von V. Aus dem Nachtrag zum Austauschsatz von Steinitz folgt dann, daß jede andere Basis von V ebenfalls unendlich sein muß (da wir die Vektoren einer angenommen endlichen Basis $\mathcal{C}$ mit endlich vielen Vektoren von $\mathcal{B}$ austauschen können, womit wir eine neue Basis

$\mathcal{B}'$ erhalten, welche $\mathcal{C}$ als echte Teilmenge enthält, womit $\mathcal{C}$ keine Basis sein kann, im Widerspruch zur Annahme). Damit ist $\dim_K V = \infty$ wohldefiniert.

Es sei nun V ein K-Vektorraum mit endlicher Basis $\mathcal{B}$, $\mathcal{C}$ sei eine weitere endliche Basis von V. Damit setzen wir

$$n := \#\mathcal{B}\ , \quad r := \#\mathcal{C}\ .$$

Nach dem Austauschsatz von Steinitz, angewandt auf die Basis $\mathcal{B}$ und die linear unabhängige Menge $\mathcal{C}$, folgt die Ungleichung

$$r \leq n\ .$$

Ebenso ergibt sich aus dem Austauschsatz, angewandt auf die Basis $\mathcal{C}$ und die linear unabhängige Menge $\mathcal{B}$,

$$n \leq r\ ,$$

also die Gleichheit

$$\#\mathcal{B} = \#\mathcal{C}\ .$$

Somit ist $\dim_K V$ auch in diesem Fall wohldefiniert. $\square$

Beispiel. Sei $V = \mathbf{R}^n$. Dann bilden die Vektoren

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

eine Basis von V. Somit gilt $\dim_{\mathbf{R}} V = n$.

Lemma. *Es sei V ein K-Vektorraum und* $\dim_K V = n$.

(i) *Ist $\mathcal{B}$ eine linear unabhängige Menge von Vektoren aus V, so besteht die Äquivalenz: $\mathcal{B}$ ist Basis von $V \iff \#\mathcal{B} = n$.*

(ii) *Ist $U \subseteq V$ ein Unterraum, so gilt* $\dim_K U \leq n$*; weiter besteht die Äquivalenz:* $\dim_K U = n \iff U = V$.

Beweis.

(i) $(\Rightarrow)$: Da $\dim_K V = n$ vorausgesetzt ist, gilt für jede Basis von V, insbesondere für $\mathcal{B}$, die Gleichheit $\#\mathcal{B} = n$.
$(\Leftarrow)$: Es sei $\mathcal{C}$ eine beliebige Basis von V; es gilt nach Voraussetzung $\#\mathcal{C} = n$. Nach dem Austauschsatz können nun die n Vektoren von $\mathcal{C}$ gegen die n Vektoren von $\mathcal{B}$ ausgetauscht werden, ohne die Basiseigenschaft zu verletzen, d.h. $\mathcal{B}$ ist eine Basis von V.

(ii) Wir setzen $k := \dim_K U$ und nehmen $k > n$ an. Dann existieren insbesondere k linear unabhängige Vektoren $c_1, \dots, c_k \in U \subseteq V$. Es folgt $\dim_K V \geq k > n$, ein Widerspruch.
($\Rightarrow$) : Ist $k = \dim_K U = n$, so bilden die n Basisvektoren $\{c_1, \dots, c_n\}$ von U nach dem Austauschsatz auch eine Basis von V, d.h. $U = \langle c_1, \dots, c_n \rangle = V$.
($\Leftarrow$) : Gilt $U = V$, so ist $\dim_K U = \dim_K V$. □

Satz (Dimensionssatz). *Es seien V ein K-Vektorraum und $S, T \subseteq V$ lineare Unterräume. Dann gilt die Dimensionsformel*

$$\dim_K S + \dim_K T = \dim_K (S \cap T) + \dim_K (S + T).$$

Beweis. Gilt $\dim_K S = \infty$ oder $\dim_K T = \infty$, so gilt die behauptete Gleichheit trivialerweise. Wir nehmen deshalb $\dim_K S < \infty$ und $\dim_K T < \infty$ an. Dann setzen wir $R := S \cap T$ und wählen weiter die Bezeichnungen:

$\mathcal{S}$=Basis von S, $\dim_K S = \#\mathcal{S} = s$,
$\mathcal{T}$=Basis von T, $\dim_K T = \#\mathcal{T} = t$,
$\mathcal{R}$=Basis von R, $\dim_K R = \#\mathcal{R} = r$.

Es sei $\mathcal{R} = \{b_1, \dots, b_r\}$. Es ist $\mathcal{R}$ sowohl eine linear unabhängige Menge von Vektoren in S als auch in T. Nach dem Austauschsatz können wir $\mathcal{S}$ bzw. $\mathcal{T}$ in der folgenden Form wählen:

$$\begin{aligned} \mathcal{S} &= \{b_1, \dots, b_r; c_{r+1}, \dots, c_s\}, \\ \mathcal{T} &= \{b_1, \dots, b_r; d_{r+1}, \dots, d_t\}. \end{aligned}$$

Wir beweisen, daß die Menge

$$\{b_1, \dots, b_r; c_{r+1}, \dots, c_s; d_{r+1}, \dots, d_t\}$$

eine Basis von $S + T$ bildet.
Erzeugtheit: Trivialerweise besteht die Inklusion

$$\langle b_1, \dots, b_r; c_{r+1}, \dots, c_s; d_{r+1}, \dots, d_t \rangle \subseteq S + T.$$

Ist nun umgekehrt $x \in S + T$, so existieren zunächst $y \in S$, $z \in T$ mit $x = y + z$. Für y bzw. z bestehen dann die Darstellungen

$$\begin{aligned} y &= \sum_{j=1}^{r} \alpha_j b_j + \sum_{j=r+1}^{s} \gamma_j c_j, \\ z &= \sum_{j=1}^{r} \beta_j b_j + \sum_{j=r+1}^{t} \delta_j d_j, \end{aligned}$$

also zusammengenommen

$$x = y + z = \sum_{j=1}^{r} (\alpha_j + \beta_j) b_j + \sum_{j=r+1}^{s} \gamma_j c_j + \sum_{j=r+1}^{t} \delta_j d_j.$$

Somit gilt

$$x \in \langle b_1, \ldots, b_r; c_{r+1}, \ldots, c_s; d_{r+1}, \ldots, d_t \rangle,$$

also

$$S + T = \langle b_1, \ldots, b_r; c_{r+1}, \ldots, c_s; d_{r+1}, \ldots, d_t \rangle.$$

Lineare Unabhängigkeit: Mit zunächst unbestimmten Skalaren $\beta_1, \ldots, \beta_r$; $\gamma_{r+1}, \ldots, \gamma_s$; $\delta_{r+1}, \ldots, \delta_t$ werde angesetzt

$$\sum_{j=1}^{r} \beta_j b_j + \sum_{j=r+1}^{s} \gamma_j c_j + \sum_{j=r+1}^{t} \delta_j d_j = 0$$

oder äquivalent dazu

$$\sum_{j=1}^{r} \beta_j b_j + \sum_{j=r+1}^{s} \gamma_j c_j = - \sum_{j=r+1}^{t} \delta_j d_j =: x.$$

Aufgrund der Darstellung linker Hand von x folgt $x \in S$; aufgrund der Darstellung rechter Hand von x ergibt sich $x \in T$. Somit gilt $x \in R = S \cap T$ und es existieren daher Skalare $\lambda_1, \ldots, \lambda_r \in K$ mit der Eigenschaft

$$x = - \sum_{j=r+1}^{t} \delta_j d_j = \sum_{j=1}^{r} \lambda_j b_j \iff \sum_{j=1}^{r} \lambda_j b_j + \sum_{j=r+1}^{t} \delta_j d_j = 0.$$

Aufgrund der linearen Unabhängigkeit der Menge $\mathcal{T} = \{b_1, \ldots, b_r; d_{r+1}, \ldots, d_t\}$ folgt

$$\lambda_1 = \ldots = \lambda_r = \delta_{r+1} = \ldots = \delta_t = 0\,.$$

Somit folgt $x = 0$, also auch

$$\sum_{j=1}^{r} \beta_j b_j + \sum_{j=r+1}^{s} \gamma_j c_j = 0.$$

Aufgrund der linearen Unabhängigkeit der Menge $\mathcal{S} = \{b_1, \ldots, b_r; c_{r+1}, \ldots, c_s\}$ folgt

$$\beta_1 = \ldots = \beta_r = \gamma_{r+1} = \ldots = \gamma_s = 0.$$

Damit ist der Basisnachweis abgeschlossen. Wir finden

$$\dim_K(S + T) = r + (s - r) + (t - r) = s + t - r = \dim_K S + \dim_K T - \dim_K R,$$

d.h. wie behauptet

$$\dim_K S + \dim_K T = \dim_K(S \cap T) + \dim_K(S + T).$$

$\square$

Beispiele.

(i) $V = \mathbf{R}^3$, S sei eine Ebene durch $\{0\}$ und T eine Gerade durch $\{0\}$, wobei $T \nsubseteq S$. Es gilt also $S \cap T = \{0\}$ und somit

$$\dim_{\mathbf{R}}(S + T) = \dim_{\mathbf{R}} S + \dim_{\mathbf{R}} T - \dim_{\mathbf{R}}(S \cap T) = 2 + 1 - 0 = 3,$$

also $S + T = \mathbf{R}^3$.

(ii) Voraussetzungen wie oben, aber jetzt $T \subseteq S$. Also ist $S \cap T = T$; es folgt

$$\dim_{\mathbf{R}}(S + T) = \dim_{\mathbf{R}} S + \dim_{\mathbf{R}} T - \dim_{\mathbf{R}}(S \cap T) = 2 + 1 - 1 = 2\,,$$

und daher $S + T = S$.

1.7 Koordinaten

Es seien V ein n-dimensionaler K-Vektorraum und $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V.

Definition. Die *Koordinaten eines Vektors* $x \in V$ (*bzgl.* $\mathcal{B}$) sind gegeben durch das n-Tupel $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in K^n$, bestimmt durch die Gleichung $x = \sum_{j=1}^n \beta_j b_j$.

Bemerkung. Die Koordinaten von $x \in V$ (bzgl. einer festen Basis $\mathcal{B}$) sind eindeutig bestimmt, denn wären

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \in K^n$$

Koordinaten von x, so hätte man

$$\sum_{j=1}^n \beta_j b_j = \sum_{j=1}^n \gamma_j b_j \Longleftrightarrow \sum_{j=1}^n (\beta_j - \gamma_j) b_j = 0,$$

also

$$\beta_j - \gamma_j = 0 \Longleftrightarrow \beta_j = \gamma_j \ (j = 1, \ldots, n)$$

aufgrund der linearen Unabhängigkeit von $\mathcal{B}$.

Beispiele.

(i) $V = \mathbf{R}^3, \quad \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ (Standardbasis).

Der Vektor $x = \begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix}$ besitzt somit die Koordinaten $\begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix}$ bzgl. $\mathcal{B}$, da ja gilt

$$\begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 7 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(ii) $V = \mathbf{R}^3, \; \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$ (dies ist eine Basis).

Der Vektor $x = \begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix}$ besitzt somit die Koordinaten $\begin{pmatrix} 4 \\ -8 \\ 7 \end{pmatrix}$ bzgl. $\mathcal{B}$, da ja gilt

$$\begin{pmatrix} 4 \\ 3 \\ 7 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - 8 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 7 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Bemerkung. Seien V ein n-dimensionaler K-Vektorraum und $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V. Die wohldefinierte Zuordnung

$$V \ni x \longmapsto \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in K^n \qquad (x = \sum_{j=1}^{n} \beta_j b_j)$$

definiert eine Abbildung $f : V \longrightarrow K^n$.

Übungsaufgabe. Man zeige, daß die Abbildung f injektiv und surjektiv, d.h. bijektiv, ist.

Kapitel 2

Lineare Abbildungen

2.1 Definitionen

Im folgenden seien K ein Körper und V, W Vektorräume über K.

Definition. Eine Abbildung $f : V \longrightarrow W$, gegeben durch die Zuordnung $x \mapsto f(x)\,(x \in V)$, heißt *lineare Abbildung*, falls gilt:

$$\begin{array}{ll} (1) & f(x+y) = f(x) + f(y) \ \ (x, y \in V), \\ (2) & f(\lambda x) = \lambda f(x) \ \ (\lambda \in K, x \in V). \end{array}$$

Bemerkung. Aus der vorhergehenden Definition folgert man sofort, daß die Linearität einer Abbildung $f : V \longrightarrow W$ äquivalent ist zu

$$(3)\ f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \ \ (\lambda, \mu \in K; x, y \in V)\,,$$

oder allgemeiner zu

$$(4)\ f\left(\sum_{j=1}^{n} \lambda_j x_j\right) = \sum_{j=1}^{n} \lambda_j f(x_j)\ (\lambda_1, \ldots, \lambda_n \in K; x_1, \ldots, x_n \in V).$$

Beispiele.

(i) V sei n-dimensional und $\mathcal{B} = \{b_1, \ldots, b_n\}$ sei eine geordnete Basis von V. Die Koordinatenabbildung $f : V \longrightarrow K^n$, gegeben durch die Zuordnung $x \mapsto \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ mit $x = \sum_{j=1}^{n} \beta_j b_j$, ist eine lineare Abbildung, denn wir haben für $x, y \in V$,

$$x = \sum_{j=1}^{n} \beta_j b_j,\ y = \sum_{j=1}^{n} \gamma_j b_j\,,$$

und $\lambda \in K$:

$$x+y=\sum_{j=1}^{n}(\beta_j+\gamma_j)b_j,\ \lambda x=\sum_{j=1}^{n}(\lambda\beta_j)b_j,$$

also

$$\begin{aligned} f(x+y) &= \begin{pmatrix} \beta_1+\gamma_1 \\ \vdots \\ \beta_n+\gamma_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} = f(x)+f(y), \\ f(\lambda x) &= \begin{pmatrix} \lambda\beta_1 \\ \vdots \\ \lambda\beta_n \end{pmatrix} = \lambda \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \lambda f(x). \end{aligned}$$

(ii) Seien $V=\mathbf{R}^2, W=\mathbf{R}^2$ und $f:\mathbf{R}^2\longrightarrow\mathbf{R}^2$ definiert durch

$$f\left(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\right) := \begin{pmatrix} \pi\xi_1+e\xi_2 \\ 2\xi_1+3\xi_2 \end{pmatrix}.$$

Es ist f eine lineare Abbildung, denn wir haben für $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \in \mathbf{R}^2$ und $\lambda\in\mathbf{R}$:

$$\begin{aligned} & f\left(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}\right) = f\left(\begin{pmatrix} \xi_1+\eta_1 \\ \xi_2+\eta_2 \end{pmatrix}\right) = \begin{pmatrix} \pi(\xi_1+\eta_1)+e(\xi_2+\eta_2) \\ 2(\xi_1+\eta_1)+3(\xi_2+\eta_2) \end{pmatrix} \\ & = \begin{pmatrix} \pi\xi_1+\pi\eta_1+e\xi_2+e\eta_2 \\ 2\xi_1+2\eta_1+3\xi_2+3\eta_2 \end{pmatrix} = \begin{pmatrix} \pi\xi_1+e\xi_2 \\ 2\xi_1+3\xi_2 \end{pmatrix} + \begin{pmatrix} \pi\eta_1+e\eta_2 \\ 2\eta_1+3\eta_2 \end{pmatrix} \\ & = f\left(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\right) + f\left(\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}\right), \end{aligned}$$

$$\begin{aligned} f\left(\lambda\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\right) &= f\left(\begin{pmatrix} \lambda\xi_1 \\ \lambda\xi_2 \end{pmatrix}\right) = \begin{pmatrix} \pi\lambda\xi_1+e\lambda\xi_2 \\ 2\lambda\xi_1+3\lambda\xi_2 \end{pmatrix} = \lambda\begin{pmatrix} \pi\xi_1+e\xi_2 \\ 2\xi_1+3\xi_2 \end{pmatrix} \\ &= \lambda f\left(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\right). \end{aligned}$$

Geometrische Deutung: Die Gerade durch $\{0\}$ mit der Richtung $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$ wird abgebildet auf die Gerade durch $\{0\}$ mit der Richtung $\begin{pmatrix} \pi\xi_1+e\xi_2 \\ 2\xi_1+3\xi_2 \end{pmatrix}$.

(iii) Die identische Abbildung $f:V\longrightarrow V$, gegeben durch die Zuordnung $x\mapsto x$ ($x\in V$), ist linear. Sie wird *Identität* genannt und mit id bezeichnet.

(iv) Die Abbildung $f:V\longrightarrow W$, gegeben durch die Zuordnung $x\mapsto 0_W$ ($x\in V$, 0_W=Nullvektor von W), ist linear. Sie wird *Nullabbildung* genannt und mit 0 bezeichnet.

Lemma. *Es seien $f : V \longrightarrow W$ eine lineare Abbildung, 0_V der Nullvektor von V und 0_W der Nullvektor von W. Dann gilt:*

$$\begin{array}{ll} \text{(i)} & f(0_V) = 0_W. \\ \text{(ii)} & f(-x) = -f(x) \; (x \in V). \end{array}$$

Beweis.

(i) Mit der Regel (2) der Definition einer linearen Abbildung haben wir

$$f(0_V) = f(0 \cdot 0_V) \overset{(2)}{=} 0 \cdot f(0_V) = 0_W.$$

(ii) Wiederum mit Hilfe von (2) erhalten wir

$$f(-x) = f((-1)x) \overset{(2)}{=} (-1)f(x) = -f(x).$$ □

Definition. Es sei $f : V \longrightarrow W$ eine lineare Abbildung. Ist f injektiv bzw. surjektiv, so sprechen wir von einer *injektiven* bzw. *surjektiven, linearen Abbildung*. Ist f bijektiv, so sprechen wir von einem *Isomorphismus*.

Lemma. *Es seien V ein n-dimensionaler K-Vektorraum mit einer Basis $\mathcal{B} = \{b_1, \ldots, b_n\}$, W ein beliebiger K-Vektorraum und $\{w_1, \ldots, w_n\}$ n beliebige Vektoren in W. Dann existiert genau eine lineare Abbildung $f : V \longrightarrow W$ mit der Eigenschaft*

$$f(b_j) = w_j \; (j = 1, \ldots, n).$$

Beweis. Wir haben die Eindeutigkeit und die Existenz von f zu zeigen.
Eindeutigkeit: Ist $f : V \longrightarrow W$ eine lineare Abbildung mit der Eigenschaft $f(b_j) = w_j$ $(j = 1, \ldots, n)$, so haben wir zu zeigen, daß damit die Zuordnung $x \mapsto f(x)$ eindeutig bestimmt ist. Ist nämlich

$$x = \sum_{j=1}^{n} \lambda_j b_j$$

(mit eindeutig bestimmten Skalaren $\lambda_1, \ldots, \lambda_n$), so folgt in der Tat

$$f(x) = f\left(\sum_{j=1}^{n} \lambda_j b_j\right) = \sum_{j=1}^{n} \lambda_j f(b_j) = \sum_{j=1}^{n} \lambda_j w_j,$$

also ist $f(x)$ eindeutig bestimmt.
Existenz: Wir definieren eine Abbidung $f : V \longrightarrow W$ durch die Zuordnung

$$x = \sum_{j=1}^{n} \lambda_j b_j \longmapsto f(x) := \sum_{j=1}^{n} \lambda_j w_j.$$

Offensichtlich erfüllt diese Abbildung die Eigenschaft $f(b_j) = w_j$ $(j = 1, \ldots, n)$. Es bleibt die Linearität von f nachzuweisen. Dazu seien $x = \sum_{j=1}^n \lambda_j b_j$, $y = \sum_{j=1}^n \mu_j b_j \in V$ und $\lambda, \mu \in K$; wegen

$$\lambda x + \mu y = \sum_{j=1}^{n} (\lambda\lambda_j + \mu\mu_j) b_j$$

ergibt sich dann

$$f(\lambda x + \mu y) = \sum_{j=1}^{n} (\lambda\lambda_j + \mu\mu_j) w_j = \lambda \sum_{j=1}^{n} \lambda_j w_j + \mu \sum_{j=1}^{n} \mu_j w_j = \lambda f(x) + \mu f(y),$$

also die Linearität von f. □

Definition. Es seien V, W K-Vektorräume und $f : V \longrightarrow W$ eine lineare Abbildung.

(i) Die Menge
$$\ker(f) := \{v \in V \mid f(v) = 0\}$$
heißt der *Kern der linearen Abbildung* f; anstelle von $\ker(f)$ wird auch $\mathrm{Kern}(f)$ geschrieben.

(ii) Die Menge
$$\mathrm{im}(f) := f(V) = \{w \in W \mid \exists v \in V : \, w = f(v)\}$$
heißt das *Bild der linearen Abbildung* f; anstelle von $\mathrm{im}(f)$ wird auch $\mathrm{Bild}(f)$ geschrieben.

Lemma. *Es seien V, W K-Vektorräume und $f : V \longrightarrow W$ eine lineare Abbildung. Dann gilt:*

(i) *Es ist* $\ker(f) \subseteq V$ *ein Unterraum von* V*.*

(ii) *Es ist* $\mathrm{im}(f) \subseteq W$ *ein Unterraum von* W*.*

Beweis.

(i) Es seien $x, y \in \ker(f)$, d.h. $f(x) = f(y) = 0$, und $\lambda, \mu \in K$. Dann haben wir $\lambda x + \mu y \in \ker(f)$ zu zeigen. Dies ergibt sich aber sofort aufgrund der Linearität von f, nämlich
$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = 0,$$
also
$$\lambda x + \mu y \in \ker(f).$$

(ii) Es seien $x', y' \in \operatorname{im}(f)$, d.h. es existieren $x, y \in V$ mit $x' = f(x), y' = f(y)$, und $\lambda, \mu \in K$. Dann haben wir zu zeigen: $\lambda x' + \mu y' \in \operatorname{im}(f)$. Dazu betrachten wir den Vektor $\lambda x + \mu y \in V$; für dessen Bild unter der linearen Abbildung f berechnen wir

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = \lambda x' + \mu y',$$

d.h. $\lambda x + \mu y$ ist ein Urbild von $\lambda x' + \mu y'$. Somit gilt

$$\lambda x' + \mu y' \in \operatorname{im}(f).$$

□

Definition. Es seien V, W K-Vektorräume und $f : V \longrightarrow W$ eine lineare Abbildung. Der *Rang der linearen Abbildung* f ist gegeben durch

$$\operatorname{rg}(f) := \dim_K(\operatorname{im}(f)).$$

Proposition. *Es seien V, W K-Vektorräume, W endlich dimensional und $f : V \longrightarrow W$ eine lineare Abbildung. Dann bestehen die Aussagen:*

(i) f surjektiv $\iff$ $\operatorname{rg}(f) = \dim_K W$.

(ii) f injektiv $\iff$ $\dim_K(\ker(f)) = 0$.

Beweis.

(i) $(\Longrightarrow)$: Die Abbildung f sei surjektiv, also gilt $\operatorname{im}(f) = f(V) = W$ und somit $\operatorname{rg}(f) = \dim_K(\operatorname{im}(f)) = \dim_K W$.
$(\Longleftarrow)$: Es ist $\operatorname{im}(f) = f(V) \subseteq W$ ein Unterraum. Aufgrund der Gleichheit $\dim_K(\operatorname{im}(f)) = \operatorname{rg}(f) = \dim_K W$ folgt nun $\operatorname{im}(f) = f(V) = W$, also ist f surjektiv.

(ii) $(\Longrightarrow)$: Da f injektiv ist, gilt $\ker(f) = \{v \in V | \, f(v) = 0\} = \{0\}$, also $\dim_K(\ker(f)) = 0$.
$(\Longleftarrow)$: Ist $\dim_K(\ker(f)) = 0$, so gilt $\ker(f) = \{0\}$, d.h. $\{v \in V | \, f(v) = 0\} = \{0\}$. Somit besteht für beliebige $x, y \in V$ die Äquivalenz

$$f(x) = f(y) \Leftrightarrow f(x - y) = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y.$$

Also ist f injektiv. □

Satz. *Es seien V, W endlich dimensionale K-Vektorräume und $f : V \longrightarrow W$ eine lineare Abbildung. Dann gilt die Formel*

$$\dim_K(\ker(f)) + \operatorname{rg}(f) = \dim_K V.$$

Beweis. Wir setzen $\dim_K(\ker(f)) = r$ und $\dim_K V = n$; da $\ker(f) \subseteq V$ ein Unterraum ist, gilt $r \leq n$. Es sei nun $\mathcal{B}' = \{b_1, \ldots, b_r\}$ eine Basis von $\ker(f)$. Mit Hilfe des Austauschsatzes von Steinitz kann $\mathcal{B}'$ zu einer Basis

$$\mathcal{B} = \{b_1, \ldots, b_r; b_{r+1}, \ldots, b_n\}$$

von V ergänzt werden. Wir behaupten jetzt, daß die Menge

$$\mathcal{F} = \{f(b_{r+1}), \ldots, f(b_n)\}$$

eine Basis von $\mathrm{im}(f) = f(V)$ bildet.
Erzeugtheit: Offensichtlich gilt die Inklusion $\langle f(b_{r+1}), \ldots, f(b_n)\rangle \subseteq \mathrm{im}(f)$. Sei andererseits $x' \in \mathrm{im}(f)$, d.h. es existiert $x \in V$ mit der Eigenschaft $x' = f(x)$. Mit der Darstellung

$$x = \sum_{j=1}^{n} \beta_j b_j$$

erhalten wir

$$x' = f(x) = f\left(\sum_{j=1}^{n} \beta_j b_j\right) = \sum_{j=1}^{n} \beta_j f(b_j) = \sum_{j=r+1}^{n} \beta_j f(b_j).$$

Somit haben wir $\mathrm{im}(f) \subseteq \langle f(b_{r+1}), \ldots, f(b_n)\rangle$, also $\mathrm{im}(f) = \langle f(b_{r+1}), \ldots, f(b_n)\rangle$.
Lineare Unabhängigkeit: Mit unbestimmten Skalaren $\lambda_{r+1}, \ldots, \lambda_n$ setzen wir an:

$$\lambda_{r+1} f(b_{r+1}) + \ldots + \lambda_n f(b_n) = 0,$$

also

$$f\left(\sum_{j=r+1}^{n} \lambda_j b_j\right) = \sum_{j=r+1}^{n} \lambda_j f(b_j) = 0.$$

Somit folgt

$$\sum_{j=r+1}^{n} \lambda_j b_j \in \ker(f),$$

also existieren eindeutig bestimmte Skalare $\mu_1, \ldots, \mu_r \in K$ mit der Eigenschaft

$$\sum_{j=r+1}^{n} \lambda_j b_j = \sum_{j=1}^{r} \mu_j b_j \iff \mu_1 b_1 + \ldots + \mu_r b_r - \lambda_{r+1} b_{r+1} - \ldots - \lambda_n b_n = 0.$$

Da $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine Basis von V ist, folgt sofort

$$\mu_1 = \ldots = \mu_r = 0, \; \lambda_{r+1} = \ldots = \lambda_n = 0,$$

also sind die Vektoren $f(b_{r+1}), \ldots, f(b_n)$ linear unabhängig, und somit ist $\mathcal{F}$ eine Basis von $\mathrm{im}(f)$. Es folgt schließlich

$$\dim_K(\ker(f)) + \mathrm{rg}(f) = r + (n - r) = n = \dim_K V. \qquad \square$$

Beispiel. Es seien $V = \mathbf{R}^3, W = \mathbf{R}^3$ und $f : \mathbf{R}^3 \longrightarrow \mathbf{R}^3$ eine lineare Abbildung, gegeben durch die Vorschrift
$$f\left(\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}\right) = \begin{pmatrix} \xi_1 + \xi_2 + \xi_3 \\ \xi_1 + \xi_2 - \xi_3 \\ \xi_3 \end{pmatrix}.$$
Das Bild $\operatorname{im}(f)$ von f wird erzeugt durch die Bilder $f(e_1), f(e_2), f(e_3)$ der drei Standardbasisvektoren e_1, e_2, e_3 von $\mathbf{R}^3$; damit gilt also

$$\operatorname{im}(f) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle,$$

und wir erhalten
$$\dim_{\mathbf{R}}(\operatorname{im}(f)) = 2.$$
Nach dem vorhergehenden Satz ergibt sich daraus

$$\dim_{\mathbf{R}}(\ker(f)) = 1.$$

Insbesondere stellen wir fest, daß der Lösungsraum des linearen homogenen Gleichungssystems

$$\begin{aligned} \xi_1 + \xi_2 + \xi_3 &= 0 \\ \xi_1 + \xi_2 - \xi_3 &= 0 \\ \xi_3 &= 0 \end{aligned}$$

1-dimensional ist.

Corollar. *Es seien V, W endlich dimensionale K-Vektorräume. Dann besteht die Äquivalenz:*
$$\exists\, f : V \longrightarrow W \textit{ Isomorphismus} \Longleftrightarrow \dim_K V = \dim_K W.$$

Beweis. ($\Longrightarrow$): Es sei $f : V \longrightarrow W$ ein Isomorphismus. Damit folgt

$$\dim_K(\ker(f)) = 0\ ,\ \ \dim_K(\operatorname{im}(f)) = \operatorname{rg}(f) = \dim_K W.$$

Aus dem vorhergehenden Satz ergibt sich dann

$$\dim_K W = \dim_K(\ker(f)) + \operatorname{rg}(f) = \dim_K V.$$

($\Longleftarrow$): Es werde jetzt umgekehrt $\dim_K V = \dim_K W = n$ vorausgesetzt. Es seien dann $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine Basis von V und $\mathcal{C} = \{c_1, \ldots, c_n\}$ eine Basis von W. Nach dem Lemma über die Konstruktion linearer Abbildungen existiert (genau) eine lineare Abbildung $f : V \longrightarrow W$ mit der Eigenschaft

$$f(b_j) = c_j\ (j = 1, \ldots, n).$$

Da f die Basis $\mathcal{B}$ von V auf die Basis $\mathcal{C}$ von W abbildet, ist f surjektiv, d.h. es ist

$$\mathrm{rg}(f) = \dim_K(\mathrm{im}(f)) = \dim_K W = n.$$

Aus dem vorhergehenden Satz folgt dann

$$\dim_K(\ker(f)) + \mathrm{rg}(f) = \dim_K(\ker(f)) + n = \dim_K V = n,$$

also

$$\dim_K(\ker(f)) = 0.$$

Somit ist f auch injektiv, also ein Isomorphismus. □

Bemerkung. Das Corollar ist eine Verallgemeinerung der bereits bekannten Tatsache, daß ein n-dimensionaler K-Vektorraum vermöge der Koordinatenabbildung isomorph zum K^n ist.

2.2 Operationen mit linearen Abbildungen

Definition. Für K-Vektorräume V, W definieren wir:

$$\begin{aligned} \mathrm{L}(V,W) &:= \{f \mid f : V \longrightarrow W, \text{ linear } \}, \\ \mathrm{L}(V) &:= \mathrm{L}(V,V). \end{aligned}$$

Proposition. *Mit den vorhergehenden Bezeichnungen gilt:* $\mathrm{L}(V,W)$ *ist ein* K*-Vektorraum.*

Beweis. Wir definieren zuerst eine Addition und Skalarmultiplikation für die Elemente von $\mathrm{L}(V,W)$.
Addition: Sind $f, g \in \mathrm{L}(V,W)$, so ist die Summe $f+g$ definiert als die Abbildung, welche gegeben ist durch die Vorschrift

$$(f+g)(x) := f(x) + g(x) \ (x \in V).$$

Man prüft nach (Übungsaufgabe), daß diese Abbildung linear ist, also in der Tat $f+g \in \mathrm{L}(V,W)$ gilt.
Skalarmultiplikation: Sind $\lambda \in K$ und $f \in \mathrm{L}(V,W)$, so ist die Multiplikation von f mit dem Skalar λ definiert als die Abbildung, welche gegeben ist durch die Vorschrift

$$(\lambda f)(x) := \lambda f(x) \ (x \in V).$$

Wiederum prüft man nach (Übungsaufgabe), daß diese Abbildung linear ist, also in der Tat $\lambda f \in \mathrm{L}(V,W)$ gilt.
Die Nullabbildung $0 : V \longrightarrow W$ spielt die Rolle des neutralen Elements von $\mathrm{L}(V,W)$ bzgl. der Addition. Ist weiter $f \in \mathrm{L}(V,W)$, so ist die additiv-inverse Abbildung $-f$ von f gegeben durch $(-f)(x) = -f(x)\,(x \in V)$. Da man schließlich alle übrigen Vektorraumaxiome für $\mathrm{L}(V,W)$ leicht nachprüft (Übungsaufgabe), folgt die Behauptung. □

Lemma. *Es seien V_1, V_2, V_3 K-Vektorräume und*

$$\mathrm{L}(V_1, V_2) \times \mathrm{L}(V_2, V_3) = \{(f, g) | f \in \mathrm{L}(V_1, V_2), g \in \mathrm{L}(V_2, V_3)\}.$$

Dann besteht die Abbildung

$$\mathrm{L}(V_1, V_2) \times \mathrm{L}(V_2, V_3) \longrightarrow \mathrm{L}(V_1, V_3),$$

gegeben durch die Zuordnung (Komposition von Abbildungen)

$$(f, g) \longmapsto g \circ f.$$

Beweis. Es ist einfach zu zeigen, daß mit $f : V_1 \longrightarrow V_2$ linear und $g : V_2 \longrightarrow V_3$ linear auch die komponierte Abbildung $g \circ f : V_1 \longrightarrow V_3$ linear ist. Dazu seien $x, y \in V_1$ und $\lambda, \mu \in K$; wir erhalten:

$$\begin{aligned}(g \circ f)(\lambda x + \mu y) &= g(f(\lambda x + \mu y)) \\ &= g(\lambda f(x) + \mu f(y)) \\ &= \lambda g(f(x)) + \mu g(f(y)) \\ &= \lambda (g \circ f)(x) + \mu (g \circ f)(y).\end{aligned}$$

□

Definition. Für einen K-Vektorraum V setzen wir

$$\mathrm{GL}(V) := \{f \in \mathrm{L}(V) | \, f \ \text{ Isomorphismus } \}.$$

Proposition. GL(V) *ist eine Gruppe bzgl. der Komposition von Abbildungen, die sogenannte lineare Gruppe von V.*

Beweis. Sind $f, g \in \mathrm{GL}(V)$, so gilt nach dem vorhergehenden Lemma $g \circ f \in \mathrm{L}(V)$; da mit f, g auch $g \circ f$ bijektiv ist, folgt $g \circ f \in \mathrm{GL}(V)$. Damit liefert die Komposition eine assoziative Struktur auf GL(V).
Die identische Abbildung id ist in GL(V) und spielt die Rolle des neutralen Elements von GL(V) bzgl. der Komposition. Ist schließlich $f \in \mathrm{GL}(V)$, so besitzt f aufgrund der Bijektivität eine Umkehrabbildung f^{-1}; für diese gilt $f^{-1} \circ f = f \circ f^{-1} = \mathrm{id}$.
Wir zeigen, daß f^{-1} auch linear ist. Dazu seien $x', y' \in V$ und $\lambda, \mu \in K$; dann existieren eindeutig bestimmte $x, y \in V$ mit der Eigenschaft

$$\begin{aligned}x' = f(x) &\iff x = f^{-1}(x'), \\ y' = f(y) &\iff y = f^{-1}(y').\end{aligned}$$

Damit ergibt sich

$$\begin{aligned}f^{-1}(\lambda x' + \mu y') &= f^{-1}(\lambda f(x) + \mu f(y)) = f^{-1}(f(\lambda x + \mu y)) = (f^{-1} \circ f)(\lambda x + \mu y) \\ &= \mathrm{id}(\lambda x + \mu y) = \lambda x + \mu y = \lambda f^{-1}(x') + \mu f^{-1}(y'),\end{aligned}$$

also gilt $f^{-1} \in \mathrm{GL}(V)$. □

2.3 Matrizen linearer Abbildungen

Es seien V, W endlich dimensionale K-Vektorräume, $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V (d.h. $\dim_K V = n$) und $\mathcal{C} = \{c_1, \ldots, c_m\}$ eine geordnete Basis von W (d.h. $\dim_K W = m$). Ist $f : V \longrightarrow W$ eine lineare Abbildung, so werden durch diese folgendermaßen eindeutig bestimmte Skalare $\alpha_{k,j}$ $(k = 1, \ldots, m; j = 1, \ldots, n)$ definiert:

$$f(b_j) = \sum_{k=1}^{m} \alpha_{k,j} c_k \ (j = 1, \ldots, n).$$

Wir ordnen diese Skalare in einem rechteckigen Schema an, d.h. wir setzen

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} = (\alpha_{k,j})_{\substack{1 \le k \le m \\ 1 \le j \le n}}.$$

Definition. A heißt *die der linearen Abbildung f: $V \longrightarrow W$ zugeordnete $m \times n$-Matrix bzgl. der Basen $\mathcal{B}$, $\mathcal{C}$*. Die Skalare

$$(\alpha_{k,1}, \ldots, \alpha_{k,n}) \ (k = 1, \ldots, m)$$

bilden die *Zeilen* (*Zeilenvektoren*) von A. Die Skalare

$$\begin{pmatrix} \alpha_{1,j} \\ \vdots \\ \alpha_{m,j} \end{pmatrix} (j = 1, \ldots, n)$$

bilden die *Spalten* (*Spaltenvektoren*) von A.

Lemma. *Es seien V, W endlich dimensionale K-Vektorräume, $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V, $\mathcal{C} = \{c_1, \ldots, c_m\}$ eine geordnete Basis von W und $f : V \longrightarrow W$ eine lineare Abbildung mit zugehöriger $m \times n$-Matrix $A = (\alpha_{k,j})$ bzgl. $\mathcal{B}$, $\mathcal{C}$ (d.h. $f(b_j) = \sum_{k=1}^{m} \alpha_{k,j} c_k$).*

Besitzt $x \in V$ bzgl. $\mathcal{B}$ die Koordinaten $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$, so hat $f(x) \in W$ bzgl. $\mathcal{C}$ die Koordinaten

$$\begin{pmatrix} \sum_{j=1}^{n} \alpha_{1,j} \beta_j \\ \vdots \\ \sum_{j=1}^{n} \alpha_{m,j} \beta_j \end{pmatrix}.$$

Beweis. Es seien $\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix}$ die Koordinaten von $f(x)$ bzgl. $\mathcal{C}$, d.h. es gilt

$$f(x) = \sum_{k=1}^{m} \gamma_k c_k.$$

Andererseits berechnen wir

$$f(x) = f\left(\sum_{j=1}^{n}\beta_j b_j\right) = \sum_{j=1}^{n}\beta_j f(b_j) = \sum_{j=1}^{n}\beta_j\sum_{k=1}^{m}\alpha_{k,j}c_k = \sum_{k=1}^{m}\left(\sum_{j=1}^{n}\alpha_{k,j}\beta_j\right)c_k,$$

also

$$\gamma_k = \sum_{j=1}^{n}\alpha_{k,j}\beta_j \ (k = 1, \ldots, m).$$

□

Definition. Die Maximalzahl linear unabhängiger Spaltenvektoren einer Matrix A heißt der *Spaltenrang von A*. Wir bezeichnen diesen mit $\mathrm{rg}_S(A)$.
Die Maximalzahl linear unabhängiger Zeilenvektoren einer Matrix A heißt der *Zeilenrang von A*. Wir bezeichnen diesen mit $\mathrm{rg}_Z(A)$.

Satz 1. *Es seien V, W endlich dimensionale K-Vektorräume, $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V, $\mathcal{C} = \{c_1, \ldots, c_m\}$ eine geordnete Basis von W und $f : V \longrightarrow W$ eine lineare Abbildung mit zugehöriger $m \times n$-Matrix $A = (\alpha_{k,j})$ bzgl. $\mathcal{B}$, $\mathcal{C}$ (d.h. $f(b_j) = \sum_{k=1}^{m}\alpha_{k,j}c_k$). Dann gilt die Gleichheit*

$$\mathrm{rg}(f) = \mathrm{rg}_S(A).$$

Beweis. Die Gleichungen

$$f(b_j) = \sum_{k=1}^{m}\alpha_{k,j}c_k \ (j = 1, \ldots, n)$$

besagen, daß die Spaltenvektoren

$$\begin{pmatrix} \alpha_{1,j} \\ \vdots \\ \alpha_{m,j} \end{pmatrix} (j = 1, \ldots, n)$$

die Koordinaten von $f(b_j)$ bzgl. $\mathcal{C}$ sind. Damit induziert die Koordinatenabbildung $W \xrightarrow{\cong} K^m$ durch Einschränkung auf das Bild $\mathrm{im}(f)$ von f eine lineare, injektive Abbildung

$$\mathrm{im}(f) = \langle f(b_1), \ldots, f(b_n)\rangle \longrightarrow \left\langle \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \ldots, \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix} \right\rangle.$$

Da diese Abbildung konstruktionsgemäß auch surjektiv ist, ergibt sich also der Isomorphismus

$$\mathrm{im}(f) \cong \left\langle \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \ldots, \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix} \right\rangle.$$

Somit ergibt sich

$$\mathrm{rg}(f) = \dim_K(\mathrm{im}(f)) = \dim_K \left\langle \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \ldots, \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix} \right\rangle = \mathrm{rg}_S(A).$$

□

Beispiel. Wir gehen aus von der 3×4-Matrix

$$A = \begin{pmatrix} 1 & 7 & 2 & 6 \\ 3 & 8 & 5 & 6 \\ 4 & 1 & 9 & -4 \end{pmatrix}.$$

Bestimmung von $\mathrm{rg}_S(A)$: Man überprüft leicht, daß die ersten drei Spaltenvektoren

$$\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 9 \end{pmatrix}$$

linear unabhängig sind. Für den vierten Spaltenvektor erkennt man andererseits

$$\begin{pmatrix} 6 \\ 6 \\ -4 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} + 1 \cdot \begin{pmatrix} 7 \\ 8 \\ 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 2 \\ 5 \\ 9 \end{pmatrix}.$$

Somit ergibt sich

$$\mathrm{rg}_S(A) = 3.$$

Bestimmung von $\mathrm{rg}_Z(A)$: Wir überprüfen die drei Zeilenvektoren auf lineare Unabhängigkeit. Dazu setzen wir mit zunächst unbestimmten Skalaren λ, μ, ν an:

$$\lambda(1,7,2,6) + \mu(3,8,5,6) + \nu(4,1,9,-4) = (0,0,0,0).$$

Wir erhalten das Gleichungssystem

$$\begin{array}{rcll} \lambda + 3\mu + 4\nu & = & 0 & \text{(I)} \\ 7\lambda + 8\mu + \ \nu & = & 0 & \text{(II)} \\ 2\lambda + 5\mu + 9\nu & = & 0 & \text{(III)} \\ 6\lambda + 6\mu - 4\nu & = & 0 & \text{(IV)}\ . \end{array}$$

Gleichung (I) liefert

$$\lambda = -3\mu - 4\nu;$$

eingesetzt in (II), (III) folgt

$$\begin{array}{l} 7(-3\mu - 4\nu) + 8\mu + \ \nu \ = 0 \\ 2(-3\mu - 4\nu) + 5\mu + 9\nu = 0 \end{array} \iff \begin{array}{r} 13\mu + 27\nu = 0\ , \\ \mu - \ \nu = 0\ . \end{array}$$

Somit ergibt sich die Lösung

$$\mu = \nu = 0 \quad \text{und} \quad \lambda = 0,$$

also

$$\mathrm{rg}_Z(A) = 3.$$

Elementare Umformungen. Wir gehen aus von der $m \times n$-Matrix

$$A = (\alpha_{k,j})_{\substack{1\le k\le m\\ 1\le j\le n}} = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

Damit haben wir die Spaltenvektoren

$$a_j = \begin{pmatrix} \alpha_{1,j} \\ \vdots \\ \alpha_{m,j} \end{pmatrix} (j = 1, \ldots, n)$$

und die Zeilenvektoren

$$\tilde{a}_k = (\alpha_{k,1}, \ldots, \alpha_{k,n}) \ (k = 1, \ldots, m).$$

Die drei elementaren Spaltenumformungen von A lauten wie folgt:

(S1) Vertauschen zweier verschiedener Spalten $a_j, a_{j'}$ $(j, j' = 1, \ldots, n; j \neq j')$.

(S2) Multiplikation einer Spalte a_j $(j = 1, \ldots, n)$ mit einem Skalar $\lambda \neq 0$.

(S3) Addition einer Spalte a_j zu einer anderen Spalte $a_{j'}$ $(j, j' = 1, \ldots, n; j \neq j')$.

Die drei elementaren Zeilenumformungen von A lauten wie folgt:

(Z1) Vertauschen zweier verschiedener Zeilen $\tilde{a}_k, \tilde{a}_{k'}$ $(k, k' = 1, \ldots, m; k \neq k')$.

(Z2) Multiplikation einer Zeile $\tilde{a}_k$ $(k = 1, \ldots, m)$ mit einem Skalar $\lambda \neq 0$.

(Z3) Addition einer Zeile $\tilde{a}_k$ zu einer anderen Zeile $\tilde{a}_{k'}$ $(k, k' = 1, \ldots, m; k \neq k')$.

Proposition. *Es bezeichne $A = (\alpha_{k,j})_{\substack{1\le k\le m\\ 1\le j\le n}}$ eine $m \times n$-Matrix. Dann ändern sich der Spaltenrang $\mathrm{rg}_S(A)$ und der Zeilenrang $\mathrm{rg}_Z(A)$ von A bei elementaren Umformungen nicht.*

Beweis.

(i) Die Invarianz des Spaltenranges $\mathrm{rg}_S(A)$ bei elementaren Umformungen:
Wir beginnen mit den elementaren Spaltenumformungen. Definitionsgemäß gilt

$$\mathrm{rg}_S(A) = \dim_K \langle a_1, \ldots, a_n \rangle.$$

Aufgrund der drei Gleichungen

$$\begin{aligned}
\langle a_1,\ldots,a_j,\ldots,a_{j'},\ldots,a_n\rangle &= \langle a_1,\ldots,a_{j'},\ldots,a_j,\ldots,a_n\rangle\\
&\quad (j,j'=1,\ldots,n; j\neq j'),\\
\langle a_1,\ldots,a_j,\ldots,a_n\rangle &= \langle a_1,\ldots,\lambda a_j,\ldots,a_n\rangle\\
&\quad (j=1,\ldots,n;\lambda\neq 0),\\
\langle a_1,\ldots,a_j,\ldots,a_{j'},\ldots,a_n\rangle &= \langle a_1,\ldots,a_j,\ldots,a_{j'}+a_j,\ldots,a_n\rangle\\
&\quad (j,j'=1,\ldots,n; j\neq j')
\end{aligned}$$

ergibt sich sofort die Invarianz des Spaltenranges $\mathrm{rg}_S(A)$ bei den elementaren Umformungen (S1), (S2), (S3).
Wir kommen zu den elementaren Zeilenumformungen. Die Koordinaten der Spaltenvektoren $a_1,\ldots,a_n$ bzgl. der Standardbasis $\mathcal{B}=\{e_1,\ldots,e_n\}$ des K^m sind gegeben durch

$$\begin{pmatrix}\alpha_{1,1}\\ \vdots\\ \alpha_{k,1}\\ \vdots\\ \alpha_{k',1}\\ \vdots\\ \alpha_{m,1}\end{pmatrix},\ldots,\begin{pmatrix}\alpha_{1,n}\\ \vdots\\ \alpha_{k,n}\\ \vdots\\ \alpha_{k',n}\\ \vdots\\ \alpha_{m,n}\end{pmatrix}.$$

Durch Vertauschen der beiden Standardbasisvektoren $e_k, e_{k'}\,(k,k'=1,\ldots,m; k\neq k')$ erhalten wir eine neue Basis $\mathcal{B}'$; bzgl. $\mathcal{B}'$ haben die Spaltenvektoren $a_1,\ldots,a_n$ die neuen Koordinaten

$$a_1'=\begin{pmatrix}\alpha_{1,1}\\ \vdots\\ \alpha_{k',1}\\ \vdots\\ \alpha_{k,1}\\ \vdots\\ \alpha_{m,1}\end{pmatrix},\ldots,a_n'=\begin{pmatrix}\alpha_{1,n}\\ \vdots\\ \alpha_{k',n}\\ \vdots\\ \alpha_{k,n}\\ \vdots\\ \alpha_{m,n}\end{pmatrix}.$$

Dieser Prozeß entspricht ersichtlich der elementaren Umformung (Z1); damit folgt

$$\mathrm{rg}_S(A)=\dim_K\langle a_1,\ldots,a_n\rangle=\dim_K\langle a_1',\ldots,a_n'\rangle,$$

d.h. der Spaltenrang $\mathrm{rg}_S(A)$ bleibt invariant bei der elementaren Umformung (Z1). Ersetzen wir den k-ten Standardbasisvektor e_k durch $\lambda^{-1}e_k\,(k=1,\ldots,m;\lambda\neq 0)$, so erhalten wir ebenfalls eine neue Basis $\mathcal{B}'$; bzgl. $\mathcal{B}'$ haben die Spaltenvektoren

$a_1, \ldots, a_n$ die neuen Koordinaten

$$a_1' = \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \lambda\alpha_{k,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \ldots, a_n' = \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \lambda\alpha_{k,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix}.$$

Dieser Prozeß entspricht ersichtlich der elementaren Umformung (Z2); damit folgt

$$\mathrm{rg}_S(A) = \dim_K\langle a_1, \ldots, a_n\rangle = \dim_K\langle a_1', \ldots, a_n'\rangle,$$

d.h. der Spaltenrang $\mathrm{rg}_S(A)$ bleibt invariant bei der elementaren Umformung (Z2). Ersetzen wir schließlich den k-ten Standardbasisvektor e_k durch $e_k - e_{k'}$ $(k, k' = 1, \ldots, m; k \neq k')$, so erhalten wir ebenfalls eine neue Basis $\mathcal{B}'$; bzgl. $\mathcal{B}'$ haben die Spaltenvektoren $a_1, \ldots, a_n$ die neuen Koordinaten

$$a_1' = \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{k,1} \\ \vdots \\ \alpha_{k',1} + \alpha_{k,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \ldots, a_n' = \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{k,n} \\ \vdots \\ \alpha_{k',n} + \alpha_{k,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix}.$$

Dieser Prozeß entspricht ersichtlich der elementaren Umformung (Z3); damit folgt

$$\mathrm{rg}_S(A) = \dim_K\langle a_1, \ldots, a_n\rangle = \dim_K\langle a_1', \ldots, a_n'\rangle,$$

d.h. der Spaltenrang $\mathrm{rg}_S(A)$ bleibt invariant bei der elementaren Umformung (Z3).

(ii) Die Invarianz des Zeilenranges $\mathrm{rg}_Z(A)$ bei elementaren Umformungen:
Wir beginnen hier mit den elementaren Zeilenumformungen. Analog zu (i) erkennt man leicht, daß der Zeilenrang $\mathrm{rg}_Z(A)$ bei den elementaren Umformungen (Z1), (Z2), (Z3) invariant bleibt. Danach betrachten wir die elementaren Spaltenumformungen. Analog zu (i) erkennt man auch hier, daß der Zeilenrang $\mathrm{rg}_Z(A)$ bei den elementaren Umformungen (S1), (S2), (S3) invariant bleibt. $\square$

Satz 2. *Es seien V, W endlich dimensionale K-Vektorräume, $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V, $\mathcal{C} = \{c_1, \ldots, c_m\}$ eine geordnete Basis von W und $f : V \longrightarrow W$ eine lineare Abbildung mit zugehöriger $m \times n$-Matrix $A = (\alpha_{k,j})$ bzgl. $\mathcal{B}$,$\mathcal{C}$ (d.h. $f(b_j) = \sum_{k=1}^{m} \alpha_{k,j} c_k$). Dann gilt die Gleichheit*

$$\mathrm{rg}(f) = \mathrm{rg}_S(A) = \mathrm{rg}_Z(A).$$

Beispiel. Bevor wir zum Beweis von Satz 2 kommen, zeigen wir anhand eines Beispiels, wie die elementaren Umformungen es erleichtern, den Spalten- und den Zeilenrang einer Matrix zu berechnen: Mit Hilfe der elementaren Spaltenumformungen (S1), (S2), (S3) und der elementaren Zeilenumformungen (Z1), (Z2), (Z3) läßt sich die Matrix

$$A = \begin{pmatrix} 0 & 1 & 9 & 2 & 8 \\ 3 & 2 & 3 & 1 & 1 \\ 4 & 7 & 4 & 3 & 2 \\ 5 & 8 & 6 & 7 & 3 \end{pmatrix}$$

in die Form

$$A' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

bringen. Damit erhält man sofort

$$\mathrm{rg}_S(A) = \mathrm{rg}_S(A') = 4, \ \ \mathrm{rg}_Z(A) = \mathrm{rg}_Z(A') = 4.$$

Beweis. Die erstere Gleichheit

$$\mathrm{rg}(f) = \mathrm{rg}_S(A)$$

wurde bereits in Satz 1 bewiesen. Es bleibt also noch die zweite Gleichheit

$$\mathrm{rg}_S(A) = \mathrm{rg}_Z(A)$$

zu beweisen. Wie im vorhergehenden Beispiel läßt sich auch im allgemeinen Fall die Matrix $A = (\alpha_{k,j})_{\substack{1\le k\le m\\ 1\le j\le n}}$ mit Hilfe der elementaren Umformungen (Z1), (Z2), (Z3) und (S1) wie folgt umformen:

$$A' = \begin{pmatrix} 1 & \alpha'_{1,2} & \alpha'_{1,3} & \cdots & \alpha'_{1,r} & \alpha'_{1,r+1} & \cdots & \alpha'_{1,n} \\ 0 & 1 & \alpha'_{2,3} & \cdots & \alpha'_{2,r} & \alpha'_{2,r+1} & \cdots & \alpha'_{2,n} \\ 0 & 0 & 1 & & & & & \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 & \alpha'_{r,r+1} & \cdots & \alpha'_{r,n} \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix};$$

hierbei ist der Index r eine gewisse natürliche Zahl mit der Eigenschaft $1 \le r \le \min(m, n)$. Unter Verwendung der elementaren Spaltenumformungen (S2), (S3) kann nun die Matrix

A' weiter umgeformt werden zu

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

Mit Hilfe der vorhergehenden Proposition folgt somit

$$\mathrm{rg}_S(A) = \mathrm{rg}_S(A') = \mathrm{rg}_S(A'') = r = \mathrm{rg}_Z(A'') = \mathrm{rg}_Z(A') = \mathrm{rg}_Z(A). \quad \square$$

Definition. Der *Rang* $\mathrm{rg}(A)$ *einer Matrix* $A = (\alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}}$ wird definiert durch

$$\mathrm{rg}(A) := \mathrm{rg}_S(A) = \mathrm{rg}_Z(A).$$

2.4 Operationen mit Matrizen

Definition. Für natürliche Zahlen m, n definieren wir:

$$\begin{aligned} \mathrm{M}_{m,n}(K) &:= \{A \mid A = (\alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}},\ \alpha_{k,j} \in K\}, \\ \mathrm{M}_n(K) &:= \mathrm{M}_{n,n}(K). \end{aligned}$$

Proposition. *Mit den vorhergehenden Bezeichnungen gilt:* $\mathrm{M}_{m,n}(K)$ *ist ein* K*-Vektorraum der Dimension* $m \cdot n$.

Beweis. Wir definieren zuerst eine Addition und Skalarmultiplikation für die Elemente von $\mathrm{M}_{m,n}(K)$.
Addition: Sind $A = (\alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}}$, $B = (\beta_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} \in \mathrm{M}_{m,n}(K)$, so ist die Summe $A + B$ definiert durch die Matrix

$$A + B := (\alpha_{k,j} + \beta_{k,j})_{\substack{1\le k\le m\\1\le j\le n}};$$

offensichtlich gilt $A + B \in \mathrm{M}_{m,n}(K)$.
Skalarmultiplikation: Sind $\lambda \in K$ und $A = (\alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}} \in \mathrm{M}_{m,n}(K)$, so ist die Multiplikation von A mit dem Skalar λ definiert durch die Matrix

$$\lambda A := (\lambda \cdot \alpha_{k,j})_{\substack{1\le k\le m\\1\le j\le n}};$$

offensichtlich gilt $\lambda A \in \mathrm{M}_{m,n}(K)$.
Die Nullmatrix $0 = (0)_{\substack{1\le k\le m\\1\le j\le n}}$ spielt die Rolle des neutralen Elements von $\mathrm{M}_{m,n}(K)$ bzgl.

der Addition. Ist weiter $A = (\alpha_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}} \in \mathrm{M}_{m,n}(K)$, so ist die additiv-inverse Matrix $-A$ von A gegeben durch $-A = (-\alpha_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}}$. Da man alle übrigen Vektorraumaxiome leicht nachprüft, stellt man $\mathrm{M}_{m,n}(K)$ als K-Vektorraum fest.
Schließlich zeigt man leicht, daß die Matrizen $E_{k,j} \in \mathrm{M}_{m,n}(K)$ mit einer 1 an der (k,j)-ten Stelle $(1 \leq k \leq m, 1 \leq j \leq n)$ und sonst lauter Nullen eine Basis von $\mathrm{M}_{m,n}(K)$ bilden. Somit folgt

$$\dim_K(\mathrm{M}_{m,n}(K)) = m \cdot n.$$

□

Folgerung. *Es seien V, W endlich dimensionale K-Vektorräume, $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V und $\mathcal{C} = \{c_1, \ldots, c_m\}$ eine geordnete Basis von W. Dann induziert die Abbildung*

$$\varphi : \mathrm{L}(V,W) \longrightarrow \mathrm{M}_{m,n}(K),$$

gegeben durch die Zuordnung

$$f \longmapsto A = (\alpha_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}} \text{ (Matrix von } f \text{ bzgl. } \mathcal{B}, \mathcal{C}),$$

einen Isomorphismus von K-Vektorräumen

$$\mathrm{L}(V,W) \cong \mathrm{M}_{m,n}(K).$$

Beweis. Linearität von φ: Seien $f, g \in \mathrm{L}(V,W)$ und $\lambda, \mu \in K$. Weiter seien

$$\varphi(f) = A = (\alpha_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}} \in \mathrm{M}_{m,n}(K) \quad \text{(d.h. } f(b_j) = \sum_{k=1}^{m} \alpha_{k,j} c_k\text{)},$$

$$\varphi(g) = B = (\beta_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}} \in \mathrm{M}_{m,n}(K) \quad \text{(d.h. } g(b_j) = \sum_{k=1}^{m} \beta_{k,j} c_k\text{)}.$$

Da nun

$$(\lambda f + \mu g)(b_j) = \lambda f(b_j) + \mu g(b_j) = \lambda \sum_{k=1}^{m} \alpha_{k,j} c_k + \mu \sum_{k=1}^{m} \beta_{k,j} c_k = \sum_{k=1}^{m} (\lambda \alpha_{k,j} + \mu \beta_{k,j}) c_k$$

gilt, folgt

$$\varphi(\lambda f + \mu g) = (\lambda \alpha_{k,j} + \mu \beta_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}} = \lambda A + \mu B = \lambda \varphi(f) + \mu \varphi(g).$$

Injektivität von φ: Sei $f \in \mathrm{L}(V,W)$ mit $\varphi(f) = 0$, d.h. $f(b_j) = 0$ für $j = 1, \ldots, n$. Damit folgt $f(x) = 0$ für alle $x \in V$, d.h. f ist die Nullabbildung. Somit ist φ injektiv.
Surjektivität von φ: Es sei $A = (\alpha_{k,j})_{\substack{1\leq k\leq m \\ 1\leq j\leq n}} \in \mathrm{M}_{m,n}(K)$ eine beliebige Matrix. Durch die Zuordung

$$b_j \longmapsto \sum_{k=1}^{m} \alpha_{k,j} c_k \;(j = 1, \ldots, n)$$

wird nach dem Lemma zur Konstruktion linearer Abbildungen eine (eindeutig bestimmte) lineare Abbildung $f \in \mathrm{L}(V, W)$ mit der Eigenschaft

$$f(b_j) = \sum_{k=1}^{m} \alpha_{k,j} c_k \; (j = 1, \ldots, n)$$

definiert. Für diese lineare Abbildung f gilt konstruktionsgemäß $\varphi(f) = A$. Somit ist φ auch surjektiv. □

Matrixmultiplikation. Es seien V_1, V_2, V_3 endlich dimensionale K-Vektorräume mit den geordneten Basen $\mathcal{B}_1 = \{b_1^{(1)}, \ldots, b_n^{(1)}\}, \mathcal{B}_2 = \{b_1^{(2)}, \ldots, b_m^{(2)}\}, \mathcal{B}_3 = \{b_1^{(3)}, \ldots, b_r^{(3)}\}$ (resp.). Nach dem Lemma über die Komposition linearer Abbildungen besteht die Abbildung

$$\mathrm{L}(V_1, V_2) \times \mathrm{L}(V_2, V_3) \longrightarrow \mathrm{L}(V_1, V_3),$$

gegeben durch die Zuordnung

$$(f, g) \longmapsto g \circ f.$$

Sind nun f die Matrix $A = (\alpha_{k,j})_{\substack{1 \le k \le m \\ 1 \le j \le n}}$ bzgl. $\mathcal{B}_1, \mathcal{B}_2$ und g die Matrix $B = (\beta_{l,k})_{\substack{1 \le l \le r \\ 1 \le k \le m}}$ bzgl. $\mathcal{B}_2, \mathcal{B}_3$ zugeordnet, d.h. es bestehen die Formeln

$$\begin{aligned} f(b_j^{(1)}) &= \sum_{k=1}^{m} \alpha_{k,j} b_k^{(2)} \; (j = 1, \ldots, n), \\ g(b_k^{(2)}) &= \sum_{l=1}^{r} \beta_{l,k} b_l^{(3)} \; (k = 1, \ldots, m), \end{aligned}$$

so berechnen wir jetzt die der Komposition $g \circ f$ bzgl. $\mathcal{B}_1, \mathcal{B}_3$ zugeordnete Matrix $C = (\gamma_{l,j})_{\substack{1 \le l \le r \\ 1 \le j \le n}}$. Wir haben

$$\begin{aligned} (g \circ f)(b_j^{(1)}) = g(f(b_j^{(1)})) &= g\left(\sum_{k=1}^{m} \alpha_{k,j} b_k^{(2)}\right) = \sum_{k=1}^{m} \alpha_{k,j} g(b_k^{(2)}) \\ &= \sum_{k=1}^{m} \alpha_{k,j} \sum_{l=1}^{r} \beta_{l,k} b_l^{(3)} = \sum_{l=1}^{r} \left(\sum_{k=1}^{m} \beta_{l,k} \alpha_{k,j}\right) b_l^{(3)}, \end{aligned}$$

also

$$C = (\gamma_{l,j})_{\substack{1 \le l \le r \\ 1 \le j \le n}} \in \mathrm{M}_{r,n}(K), \quad \gamma_{l,j} = \sum_{k=1}^{m} \beta_{l,k} \alpha_{k,j}.$$

Definition. Das *Produkt* $B \cdot A \in \mathrm{M}_{r,n}(K)$ der Matrix $B = (\beta_{l,k})_{\substack{1 \le l \le r \\ 1 \le k \le m}} \in \mathrm{M}_{r,m}(K)$ mit der Matrix $A = (\alpha_{k,j})_{\substack{1 \le k \le m \\ 1 \le j \le n}} \in \mathrm{M}_{m,n}(K)$ ist definiert durch die Matrix $C = (\gamma_{l,j})_{\substack{1 \le l \le r \\ 1 \le j \le n}}$, wobei

$$\gamma_{l,j} = \sum_{k=1}^{m} \beta_{l,k} \alpha_{k,j} \; (l = 1, \ldots, r; j = 1, \ldots, n)$$

gilt. Ausgeschrieben hat man also die folgende Situation:

$$\begin{aligned} B \cdot A &= \begin{pmatrix} \beta_{1,1} & \cdots & \beta_{1,m} \\ \vdots & & \vdots \\ \beta_{r,1} & \cdots & \beta_{r,m} \end{pmatrix} \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \\ &= \begin{pmatrix} (\beta_{1,1}\alpha_{1,1} + \ldots + \beta_{1,m}\alpha_{m,1}) & \cdots & (\beta_{1,1}\alpha_{1,n} + \ldots + \beta_{1,m}\alpha_{m,n}) \\ \vdots & & \vdots \\ (\beta_{r,1}\alpha_{1,1} + \ldots + \beta_{r,m}\alpha_{m,1}) & \cdots & (\beta_{r,1}\alpha_{1,n} + \ldots + \beta_{r,m}\alpha_{m,n}) \end{pmatrix}. \end{aligned}$$

Man beachte hierbei insbesondere die Reihenfolge!

Beispiele.

(i) $B = \begin{pmatrix} 1 & 4 & 3 \\ 8 & 7 & 1 \end{pmatrix}$, $A = \begin{pmatrix} 3 & 4 & 1 \\ 4 & 8 & 7 \\ 6 & 5 & 1 \end{pmatrix}$.

$$B \cdot A = \begin{pmatrix} 1 & 4 & 3 \\ 8 & 7 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 & 1 \\ 4 & 8 & 7 \\ 6 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 37 & 51 & 32 \\ 58 & 93 & 58 \end{pmatrix}.$$

(ii) Seien $A = (\alpha_{k,j})_{\substack{1\le k\le m \\ 1\le j\le n}} \in \mathrm{M}_{m,n}(K)$ und $x = (\xi_j)_{1\le j\le n} \in \mathrm{M}_{n,1}(K)$. Dann ist das Produkt $A \cdot x \in \mathrm{M}_{m,1}(K)$ gegeben durch

$$A \cdot x = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n \\ \vdots \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n \end{pmatrix}.$$

(iii) Sind $A, B \in \mathrm{M}_n(K)$, so gilt im allgemeinen $A \cdot B \neq B \cdot A$, d.h. die Matrixmultiplikation ist im allgemeinen nicht kommutativ.

(iv) Sind $A \in \mathrm{M}_{m,n}(K), B \in \mathrm{M}_{n,r}(K), C \in \mathrm{M}_{r,s}(K)$, so besteht die Gleichheit

$$(A \cdot B) \cdot C = A \cdot (B \cdot C) \in \mathrm{M}_{m,s}(K),$$

d.h. die Matrixmultiplikation ist assoziativ (Übungsaufgabe).

Bemerkung. Die Matrixmultiplikation läßt sich auch deuten als Abbildung

$$\mathrm{M}_{r,m}(K) \times \mathrm{M}_{m,n}(K) \longrightarrow \mathrm{M}_{r,n}(K),$$

gegeben durch die Zuordnung

$$(B, A) \longmapsto B \cdot A.$$

Berücksichtigen wir (mit den oben eingeführten Bezeichnungen) die Isomorphismen

$$\begin{aligned} \mathrm{L}(V_1, V_2) &\cong \mathrm{M}_{m,n}(K), \\ \mathrm{L}(V_2, V_3) &\cong \mathrm{M}_{r,m}(K), \\ \mathrm{L}(V_1, V_3) &\cong \mathrm{M}_{r,n}(K), \end{aligned}$$

so erhalten wir das folgende Diagramm

$$\begin{array}{ccc} \mathrm{L}(V_2, V_3) \times \mathrm{L}(V_1, V_2) & \longrightarrow & \mathrm{L}(V_1, V_3) \\ \cong\downarrow & & \downarrow\cong \\ \mathrm{M}_{r,m}(K) \times \mathrm{M}_{m,n}(K) & \longrightarrow & \mathrm{M}_{r,n}(K) \end{array} .$$

Unter dem Isomorphismus linker Hand entspricht dem Abbildungspaar (g, f) das Matrixpaar (B, A); unter dem Isomorphismus rechter Hand entspricht die Komposition $g \circ f$ dem Matrixprodukt $B \cdot A$.
Von besonderem Interesse ist der Fall $V = V_1 = V_2 = V_3$ (d.h. $n = m = r$). Dann besteht das folgende Diagramm

$$\begin{array}{ccc} \mathrm{L}(V) \times \mathrm{L}(V) & \longrightarrow & \mathrm{L}(V) \\ \cong\downarrow & & \downarrow\cong \\ \mathrm{M}_n(K) \times \mathrm{M}_n(K) & \longrightarrow & \mathrm{M}_n(K) \end{array} .$$

Definition. Für eine natürliche Zahl n setzen wir

$$\mathrm{GL}_n(K) := \{A \in \mathrm{M}_n(K) | \operatorname{rg}(A) = n\}.$$

Proposition. $\mathrm{GL}_n(K)$ *ist eine Gruppe bzgl. der Matrixmultiplikation, die sogenannte lineare Gruppe vom Grad* n.

Beweis. Es sei V ein endlich dimensionaler K-Vektorraum, $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V und $\varphi : \mathrm{L}(V) \longrightarrow \mathrm{M}_n(K)$ der Isomorphismus, gegeben durch die Zuordnung

$$f \longmapsto A = (\alpha_{k,j})_{\substack{1 \le k \le m \\ 1 \le j \le n}} = \text{Matrix von } f \text{ bzgl. } \mathcal{B}.$$

Nach dem Vorhergehenden wissen wir weiter, daß vermöge φ der Komposition von Abbildungen das Produkt von Matrizen (in der gleichen Reihenfolge) entspricht.
Seien nun $A, B \in \mathrm{GL}_n(K)$. Dann existieren $f, g \in \mathrm{L}(V)$ mit der Eigenschaft $\varphi(f) = A, \varphi(g) = B$. Wegen $\operatorname{rg}(A) = \operatorname{rg}(B) = n$ folgt mit Satz 2 die Gleichheit $\operatorname{rg}(f) = \operatorname{rg}(g) = n$, d.h. f, g sind bijektive lineare Abbildungen. Damit ist auch deren Komposition $f \circ g$ eine bijektive lineare Abbildung, also gilt

$$\operatorname{rg}(A \cdot B) = \operatorname{rg}(f \circ g) = n,$$

d.h. $A \cdot B \in \mathrm{GL}_n(K)$. Damit ist die Menge $\mathrm{GL}_n(K)$ unter der Matrixmultiplikation abgeschlossen; da die Matrixmultiplikation assoziativ ist, definiert diese somit eine assoziative

Struktur auf der Menge $\mathrm{GL}_n(K)$.
Das neutrale Element bzgl. der Matrixmultiplikation ist offensichtlich gegeben durch die Einheitsmatrix

$$E = (\delta_{k,j})_{1\leq k,j\leq n} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix};$$

hierbei bedeutet $\delta_{k,j}$ das Kronecker-Symbol. Sei schließlich $A \in \mathrm{GL}_n(K)$ und $f \in \mathrm{GL}(V)$ mit der Eigenschaft $\varphi(f) = A$. Aus Abschnitt 2.2 wissen wir bereits, daß zu $f \in \mathrm{GL}(V)$ die Umkehrabbildung $f^{-1} \in \mathrm{GL}(V)$ existiert; für diese gilt $f^{-1} \circ f = f \circ f^{-1} = \mathrm{id}$. Damit setzen wir

$$A^{-1} := \varphi(f^{-1}) \in \mathrm{GL}_n(K)$$

und nennen A^{-1} die zu A inverse Matrix; hierbei beachte man, daß $\mathrm{rg}(A^{-1}) = \mathrm{rg}(f^{-1}) = n$ gilt. Wie gewünscht ergibt sich aus der Konstruktion die Gleichheit

$$A^{-1}A = AA^{-1} = E,$$

also spielt in der Tat A^{-1} die Rolle des zu A inversen Elements. Damit folgt die Behauptung. □

Folgerung. *Es sei V ein endlich dimensionaler K-Vektorraum und $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine geordnete Basis von V. Dann induziert die Einschränkung des Isomorphismus $\varphi : \mathrm{L}(V) \xrightarrow{\cong} \mathrm{M}_n(K)$ auf $\mathrm{GL}(V)$ einen Gruppen-Isomorphismus*

$$\varphi' : \mathrm{GL}(V) \longrightarrow \mathrm{GL}_n(K)$$

($\varphi'(f) = A = (\alpha_{k,j})_{1\leq k,j\leq n}$=Matrix von f bzgl. $\mathcal{B}$), d.h.

$$\varphi'(f \circ g) = \varphi'(f) \cdot \varphi'(g)\,.$$

Beweis. Übungsaufgabe. □

Definition. Eine (quadratische) Matrix $A \in \mathrm{M}_n(K)$ mit der Eigenschaft $\mathrm{rg}(A) = n$ wird *regulär* genannt. Eine (quadratische) Matrix $A \in \mathrm{M}_n(K)$ mit der Eigenschaft $\mathrm{rg}(A) < n$ wird *singulär* genannt.

Bemerkung. Die Gruppe $\mathrm{GL}_n(K)$ besteht aus allen regulären $n \times n$-Matrizen. Die Menge $\mathrm{M}_n(K) \setminus \mathrm{GL}_n(K)$ besteht aus allen singulären $n \times n$-Matrizen.

Beispiel. Eine Diagonalmatrix

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

ist genau dann regulär, wenn $\lambda_1 \neq 0, \ldots, \lambda_n \neq 0$ gilt. In diesem Fall berechnet sich die Inverse A^{-1} von A zu

$$A^{-1} = \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix}.$$

Im allgemeinen ist es uns aber an dieser Stelle noch nicht möglich, die Inverse einer regulären Matrix anzugeben.

2.5 Basiswechsel (Basistransformationsmatrizen)

Wir beginnen mit der folgenden Definition:

Definition. Es sei V ein endlich dimensionaler K-Vektorraum, $\mathcal{B} = \{b_1, \ldots, b_n\}$ und $\mathcal{B}' = \{b'_1, \ldots, b'_n\}$ geordnete Basen von V. Die quadratische Matrix

$$S = (\sigma_{k,j})_{1 \leq k,j \leq n} \in \mathrm{M}_n(K),$$

definiert durch

$$b'_j = \sum_{k=1}^{n} \sigma_{k,j} b_k \; (j = 1, \ldots, n),$$

heißt die *Basistransformationsmatrix von $\mathcal{B}$ nach $\mathcal{B}'$*.

Bemerkung. Die Basistransformationsmatrix S von $\mathcal{B}$ nach $\mathcal{B}'$ ist regulär, d.h. $S \in \mathrm{GL}_n(K)$ (insbesondere existiert also die Inverse $S^{-1} \in \mathrm{GL}_n(K)$ mit der Eigenschaft $S^{-1}S = SS^{-1} = E$).

Unter dem Isomorphismus $\varphi : \mathrm{L}(V) \xrightarrow{\cong} \mathrm{M}_n(K)$ entspricht der Matrix S die lineare Abbildung $f : V \longrightarrow V$ mit der Eigenschaft

$$f(b_j) = b'_j \; (j = 1, \ldots, n),$$

d.h. f führt die Basis $\mathcal{B}$ in die Basis $\mathcal{B}'$ über. Somit ist f bijektiv und es gilt

$$\mathrm{rg}(S) = \mathrm{rg}(f) = n.$$

Satz. *Es sei V ein endlich dimensionaler K-Vektorraum, $\mathcal{B} = \{b_1.., b_n\}$, $\mathcal{B}' = \{b'_1, \cdots, b'_n\}$ seien geordnete Basen von V und $S = (\sigma_{k,j})_{1 \leq k,j \leq n} \in \mathrm{GL}_n(K)$ die Basistransformationsmatrix von $\mathcal{B}$ nach $\mathcal{B}'$. Besitzt x bezüglich $\mathcal{B}$ die Koordinaten $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ und bezüglich $\mathcal{B}'$ die Koordinaten $\begin{pmatrix} \beta'_1 \\ \vdots \\ \beta'_n \end{pmatrix}$, so gilt:* $\begin{pmatrix} \beta'_1 \\ \vdots \\ \beta'_n \end{pmatrix} = S^{-1} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$.

Beweis. Wir haben

$$b'_j = \sum_{k=1}^{n} \sigma_{k,j} b_k \ (j = 1, \dots, n).$$

Daraus ergibt sich

$$x = \sum_{j=1}^{n} \beta'_j b'_j = \sum_{j=1}^{n} \beta'_j \left(\sum_{k=1}^{n} \sigma_{k,j} \cdot b_k \right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \sigma_{k,j} \cdot \beta'_j \right) \cdot b_k = \sum_{k=1}^{n} \beta_k b_k \,,$$

also

$$\beta_k = \sum_{j=1}^{n} \sigma_{k,j} \cdot \beta'_j \quad (k = 1, ..., n)\,,$$

und somit

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = S \cdot \begin{pmatrix} \beta'_1 \\ \vdots \\ \beta'_n \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} \beta'_1 \\ \vdots \\ \beta'_n \end{pmatrix} = S^{-1} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

□

Satz. *Es seien V, W endlich dimensionale K-Vektorräume, $\mathcal{B} = \{b_1, \dots, b_n\}$, $\mathcal{B}' = \{b'_1, \dots, b'_n\}$ geordnete Basen von V und $\mathcal{C} = \{c_1, \dots, c_m\}$, $\mathcal{C}' = \{c'_1, \dots, c'_m\}$ geordnete Basen von W. $S = (\sigma_{k,j})_{1 \le k,j \le n} \in \mathrm{GL}_n(K)$ bzw. $T = (\tau_{l,r})_{1 \le l,r \le m} \in \mathrm{GL}_m(K)$ seien die Basistransformationsmatrizen von $\mathcal{B}$ nach $\mathcal{B}'$ bzw. von $\mathcal{C}$ nach $\mathcal{C}'$. Schließlich seien $A = (\alpha_{l,k})_{\substack{1 \le l \le m \\ 1 \le k \le n}} \in \mathrm{M}_{m,n}(K)$ bzw. $A' = (\alpha'_{r,j})_{\substack{1 \le r \le m \\ 1 \le j \le n}} \in \mathrm{M}_{m,n}(K)$, die Matrizen von f bzgl. $\mathcal{B}$,$\mathcal{C}$ bzw. $\mathcal{B}'$,$\mathcal{C}'$. Dann gilt die Formel*

$$A' = T^{-1} A S\,.$$

Beweis. Einerseits gilt für $j = 1, \dots, n$:

$$f(b'_j) = \sum_{r=1}^{m} \alpha'_{r,j} c'_r = \sum_{r=1}^{m} \alpha'_{r,j} \sum_{l=1}^{m} \tau_{l,r} c_l = \sum_{l=1}^{m} \left(\sum_{r=1}^{m} \tau_{l,r} \alpha'_{r,j} \right) c_l\,.$$

Andererseits folgt für $j = 1, \dots, n$:

$$f(b'_j) = f\left(\sum_{k=1}^{n} \sigma_{k,j} b_k \right) = \sum_{k=1}^{n} \sigma_{k,j} f(b_k) = \sum_{k=1}^{n} \sigma_{k,j} \sum_{l=1}^{m} \alpha_{l,k} c_l = \sum_{l=1}^{m} \left(\sum_{k=1}^{n} \alpha_{l,k} \sigma_{k,j} \right) c_l\,.$$

Ein Vergleich liefert

$$TA' = AS \Longleftrightarrow A' = T^{-1} A S.$$

□

Kapitel 3

Lineare Gleichungssysteme

3.1 Die Problemstellung

Ein System von Gleichungen der Form

$$\begin{array}{rcl} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n & = & \beta_1 \\ \vdots & & \vdots \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n & = & \beta_n \end{array} \quad (S),$$

kurz

$$\sum_{j=1}^{n} \alpha_{k,j}\xi_j = \beta_k \, (k = 1, \ldots, m),$$

mit vorgegebenen $\alpha_{k,j} \in K$ $(k = 1, \ldots, m; j = 1, \ldots, n)$, $\beta_k \in K$, $(k = 1, \ldots, m)$ und Unbekannten ξ_j $(j = 1, \ldots, n)$ heißt *lineares Gleichungssystem.* Unsere Aufgabe besteht darin, alle möglichen Werte für $\xi_1, \ldots, \xi_n$ zu finden, die das lineare Gleichungssystem (S) erfüllen.
Indem wir die Matrix $A = (\alpha_{k,j})_{\substack{1 \le k \le m \\ 1 \le j \le n}} \in \mathrm{M}_{m,n}(K)$, den Spaltenvektor $b = (\beta_k)_{1 \le k \le m} \in K^m$ und den Spaltenvektor der Unbekannten $x = (\xi_j)_{1 \le j \le n}$ einführen, schreibt sich (S) matriziell kurz als die Gleichung

$$Ax = b.$$

So betrachtet, besteht unsere Aufgabe darin, alle Vektoren $x \in K^n$ zu finden, so daß die Gleichung $Ax = b$ erfüllt ist. Es stellen sich folgende grundsätzliche Probleme:

(a) *Existenzproblem:* Unter welchen Bedingungen besitzt (S) überhaupt Lösungen?

(b) *Lösungsmannigfaltigkeit:* (S) besitze mindestens eine Lösung: Unter welchen Bedingungen gibt es genau eine Lösung? Wie sieht die Struktur der Lösungsmenge im allgemeinen aus?

(c) *Lösungsverfahren:* Wie kann man die Lösungen von (S) praktisch berechnen?

3.2 Zum Existenzproblem

Wir behalten die Bezeichnungen des vorhergehenden Abschnitts bei. Mit Hilfe der Zuordnung

$$x \longmapsto Ax \ (x \in K^n)$$

wird eine lineare Abbildung

$$f : K^n \longrightarrow K^m$$

definiert. Die Matrix von f bzgl. den Standardbasen des K^n bzw. K^m ist dann gerade gegeben durch die Matrix $A \in M_{m,n}(K)$, denn wir haben für $j = 1, \ldots, n$

$$f(e_j) = A\, e_j = \begin{pmatrix} \alpha_{1,j} \\ \vdots \\ \alpha_{m,j} \end{pmatrix} = \sum_{k=1}^{m} \alpha_{k,j}\, e_k\,.$$

Definition. Die Matrix

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathrm{M}_{m,n}(K)$$

heißt *Koeffizientenmatrix von* (S). Die Matrix

$$A_{\text{erw.}} = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & \beta_m \end{pmatrix} \in \mathrm{M}_{m,n+1}(K)$$

heißt *erweiterte Koeffizientenmatrix von* (S).

Satz. *Das lineare Gleichungssystem (S) ist genau dann lösbar, wenn*

$$\mathrm{rg}(A) = \mathrm{rg}(A_{\text{erw.}})$$

gilt.

Beweis. Wir definieren die Spaltenvektoren

$$a_j := \begin{pmatrix} \alpha_{1,j} \\ \vdots \\ \alpha_{m,j} \end{pmatrix} \in K^m \ (j = 1, \ldots, n), \ b := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in K^m.$$

Mit dieser Definition und den vorhergehenden Vorbereitungen finden wir folgende Reihe von Äquivalenzen

$$\begin{aligned}
(S)\,\text{lösbar} &\iff \exists\, x \in K^n :\ Ax = b \\
&\iff \exists\, x \in K^n :\ f(x) = b \\
&\iff b \in \operatorname{im}(f) \\
&\iff b \in \langle f(e_1), \ldots, f(e_n) \rangle \\
&\iff b \in \langle a_1, \ldots, a_n \rangle \\
&\iff \langle a_1, \ldots, a_n \rangle = \langle a_1, \ldots, a_n, b \rangle \\
&\iff \operatorname{rg}(A) = \operatorname{rg}(A_{\text{erw.}}).
\end{aligned}$$

Dies beweist die Behauptung. □

Beispiel. Ist $m \leq n$ und gilt $\operatorname{rg}(A) = m$, so besitzt (S) eine Lösung: Aufgrund der Voraussetzung spannen nämlich die $n\,(\geq m)$ Spaltenvektoren $a_1, \ldots, a_n$ von A den K^m auf. Damit spannen dann aber auch die $(n+1)$ Spaltenvektoren $a_1, \ldots, a_n, b$ den K^m auf, also gilt

$$\operatorname{rg}(A) = \operatorname{rg}(A_{\text{erw.}}) = m.$$

Die Behauptung ergibt sich somit mit Hilfe des vorhergehenden Satzes.

3.3 Zur Lösungsmannigfaltigkeit

Zur Behandlung dieser Problemstellung betrachten wir zunächst die speziellen linearen Gleichungssysteme mit $\beta_1 = \ldots = \beta_m = 0$, d.h.

$$\begin{matrix} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n = 0 \\ \vdots \qquad\qquad\qquad \vdots \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n = 0 \end{matrix} \qquad (S_0).$$

Definition. Ein lineares Gleichungssystem vom Typ (S_0) heißt *homogenes lineares Gleichungssystem*. Ein lineares Gleichungssystem, das nicht vom Typ (S_0) ist, heißt *inhomogenes lineares Gleichungssystem*.

Bemerkung. Im homogenen Fall gilt offensichtlich immer die Gleichheit $\operatorname{rg}(A) = \operatorname{rg}(A_{\text{erw.}})$, also besitzen homogene lineare Gleichungssysteme immer mindestens eine Lösung; dies ist die triviale Lösung $\xi_1 = \ldots = \xi_n = 0$.

Satz. *Sei*

$$L_0 := \{x \in K^n \,|\, Ax = 0\},$$

d.h. L_0 ist die Menge aller Lösungen des homogenen linearen Gleichungssystems (S_0). Dann ist $L_0 \subseteq K^n$ ein Unterraum der Dimension

$$\dim_K L_0 = n - \operatorname{rg}(A).$$

Beweis. Die Matrix $A \in \mathrm{M}_{m,n}(K)$ gibt, wie bereits ausgeführt, Anlaß zu der linearen Abbildung $f : K^n \longrightarrow K^m$, definiert durch $f(x) = Ax$. Somit haben wir

$$L_0 = \{x \in K^n |\, Ax = 0\} = \{x \in K^n |\, f(x) = 0\} = \ker(f).$$

Damit ist L_0 bereits als Unterraum von K^n nachgewiesen. Für dessen Dimension erhalten wir nach dem Dimensionssatz

$$\dim_K L_0 = \dim_K(\ker(f)) = \dim_K(K^n) - \mathrm{rg}(f) = n - \mathrm{rg}(A).$$

□

Beispiel. Ist $m \geq n$ und gilt $\mathrm{rg}(A) = n$, so besitzt (S_0) nur die triviale Lösung: Dies folgt sofort aus dem vorhergehenden Satz, nämlich der Formel $\dim_K L_0 = n - \mathrm{rg}(A) = 0$.
Besitzt (S_0) umgekehrt nur die triviale Lösung, so folgt wiederum mit Hilfe des vorhergehenden Satzes die Gleichung $\mathrm{rg}(A) = n$ (und damit insbesondere auch $m \geq n$).
Im Spezialfall $m = n$, d.h. $A \in \mathrm{M}_n(K)$, hat also (S_0) genau dann nur die triviale Lösung, wenn A regulär ist. Dies könnten wir auch wie folgt beweisen: Genau dann, wenn A regulär ist, d.h. $A \in \mathrm{GL}_n(K)$ gilt, existiert die zu A inverse Matrix $A^{-1} \in \mathrm{GL}_n(K)$ (diese erfüllt $A^{-1}A = AA^{-1} = E =$ Einheitsmatrix); aus der Gleichung $Ax = 0$ erhalten wir durch Linksmultiplikation mit A^{-1} die Gleichung

$$x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1} \cdot 0 = 0.$$

Wir kommen nun zum allgemeinen Fall zurück. Wir gehen aus vom linearen Gleichungssystem

$$Ax = b \ (S)$$

mit $A \in \mathrm{M}_{m,n}(K)$, $b \in K^m$ und setzen voraus, daß (S) mindestens eine Lösung $x^* \in K^n$ besitzt (d.h. es gilt $\mathrm{rg}(A) = \mathrm{rg}(A_{\text{erw.}})$). Indem wir den Spaltenvektor $b \in K^m$ durch den Nullvektor $0 \in K^m$ ersetzen, erhalten wir aus (S) das homogene lineare Gleichungssystem

$$Ax = 0, \ (S_0)$$

welches wir *das (S) zugeordnete homogene lineare Gleichungssystem* nennen.

Satz. *Sei*

$$L := \{x \in K^n |\, Ax = b\},$$

d.h. L ist die Menge aller Lösungen des linearen Gleichungssystems (S). Weiter sei $L \neq \emptyset$ vorausgesetzt, also existiert $x^* \in L$, welches im folgenden fixiert sei. Dann gilt

$$L = \{x \in K^n |\, x = x^* + y,\, y \in L_0\},$$

d.h. die Menge aller Lösungen von (S) erhält man aus einer (speziellen) Lösung x^* von (S) durch Addition sämtlicher Lösungen von (S_0).

Beweis. Voraussetzungsgemäß gilt $Ax^* = b$, und für $y \in L_0$ gilt $Ay = 0$; somit folgt

$$A(x^* + y) = Ax^* + Ay = b + 0 = b,$$

d.h. $x^* + y \in L$, also gilt

$$\{x \in K^n |\, x = x^* + y,\, y \in L_0\} \subseteq L.$$

Sei nun umgekehrt $x \in L$, d.h. $Ax = b$. Dann setze man: $y := x - x^*$; es folgt

$$Ay = A(x - x^*) = Ax - Ax^* = b - b = 0,$$

d.h. $y \in L_0$ und $x = x^* + y$, also gilt

$$L \subseteq \{x \in K^n |\, x = x^* + y,\, y \in L_0\}.$$

Somit besteht die behauptete Gleichheit. □

Corollar. *Es seien* $m = n, A \in \mathrm{M}_n(K)$ *und* $\mathrm{rg}(A) = n$*, d.h.* $A \in \mathrm{GL}_n(K)$*. Dann besitzt* (S) *genau eine Lösung.*

Beweis. Aus $\mathrm{rg}(A) = n$ folgt unmittelbar auch $\mathrm{rg}(A_{\mathrm{erw.}}) = n$; damit ist (S) lösbar. Aus den beiden vorhergehenden Sätzen folgt, daß wegen $\mathrm{rg}(A) = n$ die Lösung von (S) eindeutig bestimmt ist. □

Bemerkung. Die Aussage des vorhergehenden Corollar kann auch wie folgt bewiesen werden: Zu $A \in \mathrm{GL}_n(K)$ existiert die inverse Matrix $A^{-1} \in \mathrm{GL}_n(K)$. Damit erhält man aus der Gleichung $Ax = b$ durch Linksmultiplikation mit A^{-1} die Gleichung

$$x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}b,$$

also eine eindeutig bestimmte Lösung.

3.4 Lösungsverfahren I

Das Hauptlösungsverfahren linearer Gleichungssysteme ist der im folgenden dargestellte *Gauß'sche Algorithmus*. Wir gehen also aus vom linearen Gleichungssystem

$$Ax = b, \ (S)$$

das vollständig zu lösen ist. Natürlich müßte man sich zunächst überlegen, daß mindestens eine Lösung existiert; dies ergibt sich aber im Zuge der folgenden Überlegungen automatisch. Wir schreiben die erweiterte Koeffizientenmatrix $A_{\mathrm{erw.}}$ in der Form

$$\begin{array}{ccc|c} \alpha_{1,1} & \cdots & \alpha_{1,n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & \beta_m \end{array}$$

auf, wobei wir den Spaltenvektor b durch einen senkrechten Strich abtrennen.

Lemma 1. *Anwendung der elementaren Zeilenumformungen* (Z1) - (Z3) *auf die erweiterte Koeffizientenmatrix* $A_{\mathrm{erw.}}$ *ändert die Lösungsmenge von* (S) *nicht.*

Beweis. Die Aussage des Lemmas ist offensichtlich klar bei den elementaren Zeilenumformungen (Z1) und (Z2). Da bei der elementaren Zeilenumformung (Z3) eine Gleichung lediglich durch eine äquivalente Gleichung ersetzt wird, ist die Behauptung auch in diesem Fall klar. □

Lemma 2. *Anwendung der elementaren Spaltenumformung* (S1) *auf die erweiterte Koeffizientenmatrix* $A_{\text{erw.}}$*, welche den Vektor* b *nicht beinhaltet, ändert die Lösungsmenge von* (S) *nur dahingehend, daß eine Vertauschung der* $\xi_1, \ldots, \xi_n$ *bewirkt wird.*

Beweis. Die Behauptung ist offensichtlich klar. □

Nach dem bereits bekannten Vorgehen können wir durch Anwendung der elementaren Zeilenumformungen (Z1) - (Z3) und elementarer Spaltenumformungen (S1), welche den Vektor b nicht beinhalten, erreichen, daß die erweiterte Koeffizientenmatrix $A_{\text{erw.}}$ zu einer Matrix $A'_{\text{erw.}}$ der folgenden Art umgeformt werden kann:

$$\begin{array}{ccccccc|c}
\alpha'_{1,1} & \cdots & \cdots & \alpha'_{1,r} & \alpha'_{1,r+1} & \cdots & \alpha'_{1,n} & \beta'_1 \\
0 & \alpha'_{2,2} & \cdots & \alpha'_{2,r} & \alpha'_{2,r+1} & \cdots & \alpha'_{2,n} & \beta'_2 \\
\vdots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & \alpha'_{r,r} & \alpha'_{r,r+1} & \cdots & \alpha'_{r,n} & \beta'_r \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \beta'_{r+1} \\
\vdots & & & & & & \vdots & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \beta'_m
\end{array}$$

wobei $\alpha'_{j,j} \neq 0\,(j = 1, \ldots, r)$ gilt. Bei weiterer Anwendung der elementaren Zeilenumformungen (Z2), (Z3) kann die Matrix $A_{\text{erw.}}$ sogar in die folgende, noch einfachere Gestalt gebracht werden:

$$\begin{array}{ccccccc|c}
\alpha'_{1,1} & 0 & \cdots & 0 & \alpha'_{1,r+1} & \cdots & \alpha'_{1,n} & \beta'_1 \\
0 & \alpha'_{2,2} & \ddots & \vdots & \alpha'_{2,r+1} & \cdots & \alpha'_{2,n} & \beta'_2 \\
\vdots & \ddots & \ddots & 0 & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & \alpha'_{r,r} & \alpha'_{r,r+1} & \cdots & \alpha'_{r,n} & \beta'_r \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \beta'_{r+1} \\
\vdots & & & & & & \vdots & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \beta'_m
\end{array}$$

Aus dem linearen Gleichungssystem (S) erhalten wir somit das folgende, im wesentlichen äquivalente, lineare Gleichungssystem (S'); dabei bedeuten im folgenden $\xi'_1, \ldots, \xi'_n$ eine Vertauschung der $\xi_1, \ldots, \xi_n$, die aufgrund der insgesamt vorgenommenen elementaren

Spaltenumformungen (S1) bewirkt wurde:

$$\begin{array}{cccclcl}
\alpha'_{1,1}\xi'_1 & & & & + \alpha'_{1,r+1}\xi'_{r+1} + \ldots + \alpha'_{1,n}\xi'_n & = & \beta'_1 \\
& \alpha'_{2,2}\xi'_2 & & & + \alpha'_{2,r+1}\xi'_{r+1} + \ldots + \alpha'_{2,n}\xi'_n & = & \beta'_2 \\
& & \ddots & & \vdots \qquad\qquad \vdots & & \vdots \\
& & & \alpha'_{r,r}\xi'_r & + \alpha'_{r,r+1}\xi'_{r+1} + \ldots + \alpha'_{r,n}\xi'_n & = & \beta'_r \\
& & & & \qquad\qquad\qquad 0 & = & \beta'_{r+1} \\
& & & & \qquad\qquad\qquad \vdots & & \vdots \\
& & & & \qquad\qquad\qquad 0 & = & \beta'_m
\end{array} \qquad (S')$$

Mit Hilfe des soeben gewonnenen linearen Gleichungssystems (S') erhalten wir folgende Erkenntnisse:

(i) (S) lösbar $\Longleftrightarrow$ (S') lösbar $\Longleftrightarrow$ $\beta'_{r+1} = 0, \ldots, \beta'_m = 0$.

(ii) Falls (i) erfüllt ist, erhalten wir folgende spezielle Lösung $x'^* \in K^n$ von (S'):

$$\xi'^*_1 = \frac{\beta'_1}{\alpha'_{1,1}}, \ldots, \xi'^*_r = \frac{\beta'_r}{\alpha'_{r,r}}, \quad \xi'^*_{r+1} = 0, \ldots, \xi'^*_n = 0.$$

Nach entsprechendem Vertauschen erhalten wir daraus sofort eine spezielle Lösung $x^* \in K^n$ von (S).

(iii) Indem wir den Spaltenvektor $b \in K^m$ durch den Nullvektor $0 \in K^m$ ersetzen, erhalten wir aus dem linearen Gleichungssystem (S) das zugeordnete homogene lineare Gleichungssystem (S_0). Entsprechend erhalten wir aus dem zu (S) äquivalenten linearen Gleichungssystem (S') das homogene lineare Gleichungssystem

$$\begin{array}{cccclcl}
\alpha'_{1,1}\xi'_1 & & & & + \alpha'_{1,r+1}\xi'_{r+1} + \ldots + \alpha'_{1,n}\xi'_n & = & 0 \\
& \alpha'_{2,2}\xi'_2 & & & + \alpha'_{2,r+1}\xi'_{r+1} + \ldots + \alpha'_{2,n}\xi'_n & = & 0 \\
& & \ddots & & \vdots \qquad\qquad \vdots & & \vdots \\
& & & \alpha'_{r,r}\xi'_r & + \alpha'_{r,r+1}\xi'_{r+1} + \ldots + \alpha'_{r,n}\xi'_n & = & 0
\end{array} \qquad (S'_0).$$

Eine Basis des $(n-r)$-dimensionalen Lösungsraumes des homogenen linearen Gleichungssystems (S'_0) ist nun gegeben durch die folgenden $(n-r)$ Spaltenvektoren $y'^{(\rho)} \in K^n$ $(\rho = 1, \ldots, n-r)$:

$$y'^{(\rho)} = \begin{pmatrix} -\alpha'_{1,r+\rho}/\alpha'_{1,1} \\ -\alpha'_{2,r+\rho}/\alpha'_{2,2} \\ \vdots \\ -\alpha'_{r,r+\rho}/\alpha'_{r,r} \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \\ \\ \\ \leftarrow (r+\rho)-\ \text{te Zeile} \\ \\ \\ \end{array}$$

Wir fassen die gewonnenen Resultate im nachfolgenden Satz zusammen.

Satz(Gauß'scher Algorithmus). *Das lineare Gleichungssystem*

$$Ax = b \qquad (S)$$

mit Koeffizientenmatrix $A \in \mathrm{M}_{m,n}(K)$ *und Spaltenvektor* $b \in K^m$ *sei vorgelegt. Dann ist die Menge*

$$L = \{x \in K^n |\, Ax = b\}$$

aller Lösungen von (S) *bis auf eine Vertauschung der Koordinaten (bewirkt durch die Anwendung elementarer Spaltenumformungen (S1), siehe unten) gegeben durch die Menge*

$$\left\{ \begin{pmatrix} \xi'_1 \\ \vdots \\ \xi'_r \\ \xi'_{r+1} \\ \vdots \\ \xi'_n \end{pmatrix} \in K^n \,\middle|\, \begin{pmatrix} \xi'_1 \\ \vdots \\ \xi'_r \\ \xi'_{r+1} \\ \vdots \\ \xi'_n \end{pmatrix} = \begin{pmatrix} \beta'_1/\alpha'_{1,1} \\ \vdots \\ \beta'_r/\alpha'_{r,r} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -\alpha'_{1,r+1}/\alpha'_{1,1} \\ \vdots \\ -\alpha'_{r,r+1}/\alpha'_{r,r} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \lambda_{n-r} \begin{pmatrix} -\alpha'_{1,n}/\alpha'_{1,1} \\ \vdots \\ -\alpha'_{r,n}/\alpha'_{r,r} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}$$

mit $\lambda_1, \ldots, \lambda_{n-r} \in K$. *Hierbei sind* $(\alpha'_{k,j})_{\substack{1\le k\le m \\ 1\le j\le n}}, (\beta'_k)_{1\le k\le m}$ *die Einträge der erweiterten Koeffizientenmatrix* $A'_{\text{erw.}}$, *welche aus der erweiterten Koeffizientenmatrix* $A_{\text{erw.}}$ *durch die Anwendung elementarer Zeilenumformungen (Z1) - (Z3) und elementarer Spaltenumformungen (S1) (welche den Spaltenvektor* $b \in K^m$ *nicht beinhalten) hervorgeht und die spezielle, in der vorhergehenden Diskussion gegebene Form mit* $\beta'_{r+1} = \ldots = \beta'_m = 0$ *besitzt.*

Beweis. Die Behauptung ergibt sich unmittelbar aus den vorhergehenden Überlegungen. □

3.5 Determinanten

Wir erinnern an die folgende Definition.

Definition. Eine bijektive Selbstabbildung π der Menge $\{1, \ldots, n\}$ heißt eine *Permutation* der Elemente $1, \ldots, n$. Eine Permutation π ist bestimmt durch die Angabe des Zahlenschemas

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}.$$

Die Menge aller Permutationen der Elemente $1, \ldots, n$ sei mit S_n bezeichnet.

Lemma. *Die Menge* S_n *ist eine Gruppe bzgl. der Verknüpfung von Abbildungen, die sogenannte symmetrische Gruppe zu* n *Elementen.*

Beweis. Diese wurde bereits in Kapitel 0 bewiesen. □

Definition. Eine Permutation $\tau \in S_n$, durch die zwei Elemente vertauscht werden und alle übrigen festgehalten werden, heißt *Transposition.* Ist τ gegeben durch das Zahlenschema

$$\begin{pmatrix} 1 & \dots & k_1 & \dots & k_2 & \dots & n \\ 1 & \dots & k_2 & \dots & k_1 & \dots & n \end{pmatrix} \quad (1 \leq k_1 < k_2 \leq n),$$

so wird dafür auch kurz (k_1, k_2) geschrieben.

Lemma. *Es gelten die beiden folgenden Aussagen:*

(i) *Jede Permutation $\pi \in S_n$ läßt sich als Produkt von Transpositionen darstellen.*

(ii) *Die Anzahl der Transpositionen, durch welche eine Permutation $\pi \in S_n$ als Produkt dargestellt werden kann, ist entweder immer gerade oder immer ungerade.*

Beweis.

(i) Die Permutation $\pi \in S_n, \pi \neq \mathrm{id}$, sei gegeben durch das Zahlenschema

$$\begin{pmatrix} 1 & 2 & \dots & k & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(k) & \dots & \pi(n) \end{pmatrix};$$

dabei sei die natürliche Zahl k, $1 \leq k \leq n$, dadurch festgelegt, daß $\pi(j) = j$ für $j = 1, \dots, k-1$ und $\pi(k) \neq k$ gilt. Wir multiplizieren nun π von links mit der Transposition $\tau_1 = (k, \pi(k))$. Wir erhalten die Permutation $\pi_1 = \tau_1 \circ \pi \in S_n$ gegeben durch

$$\begin{aligned} (k, \pi(k)) \circ &\begin{pmatrix} 1 & \dots & k-1 & k & k+1 & \dots & n \\ 1 & \dots & k-1 & \pi(k) & \pi(k+1) & \dots & \pi(n) \end{pmatrix} \\ &= \begin{pmatrix} 1 & \dots & k-1 & k & k+1 & \dots & n \\ 1 & \dots & k-1 & k & \pi(k+1) & \dots & \pi(n) \end{pmatrix}. \end{aligned}$$

Durch analoge Behandlung von π_1, erhalten wir nach maximal $l = n - k$ Schritten l Transpositionen, so daß

$$\tau_l \circ \dots \circ \tau_1 \circ \pi = \mathrm{id}$$

gilt. Damit folgt wie behauptet

$$\pi = \tau_1^{-1} \circ \dots \circ \tau_l^{-1} = \tau_1 \circ \dots \circ \tau_l.$$

(ii) Wir beginnen den zweiten Teil des Beweises mit einer Hilfsüberlegung. Mit den Unbestimmten $X_1, \dots, X_n$ definieren wir das Produkt

$$\begin{aligned} P(X_1, \dots, X_n) \;=\; & (X_1 - X_2) \cdot (X_1 - X_3) \cdot \ldots \cdot (X_1 - X_n) \cdot \\ & \qquad\qquad \cdot (X_2 - X_3) \cdot \ldots \cdot (X_2 - X_n) \cdot \\ & \qquad\qquad\qquad \ddots \qquad\qquad \vdots \\ & \qquad\qquad\qquad\qquad \cdot (X_{n-1} - X_n). \end{aligned}$$

Ist $\pi \in S_n$, so ergibt sich durch Einsetzen von $X_1 = 1$, $X_2 = 2$, ... , $X_n = n$, bzw. $X_1 = \pi(1)$, $X_2 = \pi(2)$,..., $X_n = \pi(n)$, die Gleichung

$$P(\pi(1), \pi(2), \ldots, \pi(n)) = c_\pi \cdot P(1, 2, \ldots, n)$$

mit einem eindeutig bestimmten Vorzeichen

$$c_\pi \in \{+1, -1\}.$$

Wir kommen nun zum eigentlichen Beweis: Ist $\tau \in S_n$ eine Transposition, so überlegt man leicht, daß $c_\tau = -1$ gilt. Sind weiter τ, τ' Transpositionen, so zeigt man ebenso leicht, daß $c_{\tau\circ\tau'} = +1$, d.h.

$$c_{\tau\circ\tau'} = c_\tau c_{\tau'}$$

gilt. Wir nehmen nun an, die Permutation $\pi \in S_n$ werde als Produkt von l bzw. l' Transpositionen dargestellt, d.h.

$$\pi = \tau_1 \circ \ldots \circ \tau_l \;\text{ bzw. }\; \pi = \tau_1' \circ \ldots \circ \tau_{l'}'.$$

Nach dem eben Festgestellten folgt

$$(-1)^l = c_\pi = (-1)^{l'},$$

d.h. l, l' sind somit beide zugleich gerade oder ungerade. □

Mit dem obigen Lemma ist die folgende Definition sinnvoll.

Definition. Es sei $\pi \in S_n$ durch l Transpositionen dargestellt. Dann heißt

$$\text{sgn }(\pi) := (-1)^l$$

das *Signum von* π. Gilt $\text{sgn}(\pi) = +1$, so heißt π eine *gerade Permutation*; ist $\text{sgn}(\pi) = -1$, so heißt π eine *ungerade Permutation*.

Beispiel. Die Permutation $\pi \in S_5$ sei gegeben durch das Zahlenschema

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix}.$$

Eine Darstellung von π als Produkt von Transpositionen ist

$$\pi = (1, 4)(1, 5)(1, 2).$$

Somit folgt $\text{sgn}(\pi) = (-1)^3 = -1$, also ist π eine ungerade Permutation. Wir kommen nun zur Definition der Determinante.

Definition. Es sei $A = (\alpha_{k,j})_{1\le k,j\le n} \in \mathrm{M}_n(K)$ eine quadratische Matrix. Die *Determinante von* A ist definiert durch

$$\det(A) = \begin{vmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{vmatrix} := \sum_{\pi\in S_n} \text{sgn}(\pi)\alpha_{1,\pi(1)} \cdot \ldots \cdot \alpha_{n,\pi(n)}.$$

Beispiel. Für $n = 2$ ist $S_2 = \{\mathrm{id}, (1,2)\}$; dabei ist id eine gerade Permutation und $(1,2)$ eine ungerade Permutation. Die Determinante von

$$A = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{pmatrix} \in \mathrm{M}_2(K)$$

ist somit gegeben durch

$$\det(A) = \begin{vmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{vmatrix} = (+1)\alpha_{1,1}\alpha_{2,2} + (-1)\alpha_{1,2}\alpha_{2,1} = \alpha_{1,1}\alpha_{2,2} - \alpha_{1,2}\alpha_{2,1}.$$

Definition. Es sei $A = (\alpha_{k,j})_{1\leq k,j\leq n} \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann heißt die Matrix ${}^tA$, die aus A durch Spiegelung der Einträge an der Diagonalen entsteht, d.h.

$${}^tA = (\alpha_{j,k})_{1\leq j,k\leq n} = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{n,1} \\ \vdots & & \vdots \\ \alpha_{1,n} & \cdots & \alpha_{n,n} \end{pmatrix},$$

die *zu A transponierte Matrix.*

Lemma. *Es sei $A = (\alpha_{k,j})_{1\leq k,j\leq n} \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann gilt*

$$\det(A) = \det({}^tA).$$

Beweis. Aufgrund der Definition der Determinante und der Transponierten einer Matrix folgt sofort

$$\det({}^tA) = \sum_{\pi\in S_n} \mathrm{sgn}(\pi)\alpha_{\pi(1),1}\cdot\ldots\cdot\alpha_{\pi(n),n}.$$

Damit erhalten wir

$$\begin{aligned} \det(A) &= \sum_{\pi\in S_n} \mathrm{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{n,\pi(n)} \\ &= \sum_{\pi\in S_n} \mathrm{sgn}(\pi)\alpha_{\pi^{-1}(1),1}\cdot\ldots\cdot\alpha_{\pi^{-1}(n),n}. \end{aligned}$$

Beachten wir weiter $\mathrm{sgn}(\pi) = \mathrm{sgn}(\pi^{-1})$ und die Tatsache, daß die Zuordnung $\pi \mapsto \pi^{-1}$ eine Bijektion von S_n auf sich bewirkt und somit mit π auch π^{-1} ganz S_n durchläuft, so folgt andererseits

$$\det(A) = \sum_{\pi^{-1}\in S_n} \mathrm{sgn}(\pi^{-1})\alpha_{\pi^{-1}(1),1}\cdot\ldots\cdot\alpha_{\pi^{-1}(n),n}.$$

Indem wir schließlich anstelle von π^{-1} wieder π schreiben, erhalten wir

$$\det(A) = \sum_{\pi\in S_n} \mathrm{sgn}(\pi)\alpha_{\pi(1),1}\cdot\ldots\cdot\alpha_{\pi(n),n} = \det({}^tA).$$

□

Proposition. *Es sei $A = (\alpha_{k,j})_{1\leq k,j\leq n} \in \mathrm{M}_n(K)$ eine quadratische Matrix. Dann gelten die folgenden Aussagen:*

(i) *Geht die Matrix $A' \in \mathrm{M}_n(K)$ aus der Matrix A durch die elementare Zeilenumformung (Z1), bzw. die elementare Spaltenumformung (S1), hervor, so gilt*

$$\det(A') = -\det(A).$$

Besitzt die Matrix A insbesondere zwei gleiche Zeilen, bzw. zwei gleiche Spalten, so gilt

$$\det(A) = 0.$$

(ii) *Geht die Matrix $A' \in \mathrm{M}_n(K)$ aus der Matrix A durch die elementare Zeilenumformung (Z2), bzw. die elementare Spaltenumformung (S2), durch Multiplikation mit einem Skalar $\lambda \neq 0$ hervor, so gilt*

$$\det(A') = \lambda \det(A).$$

Insbesondere gilt damit

$$\det(\lambda A) = \lambda^n \det(A).$$

(iii) *Geht die Matrix $A' \in \mathrm{M}_n(K)$ aus der Matrix A durch die elementare Zeilenumformung (Z3), bzw. die elementare Spaltenumformung (S3), hervor, so gilt*

$$\det(A') = \det(A).$$

Beweis. Aus dem vorhergehenden Lemma folgt, daß es genügt, die Behauptungen (i) - (iii) für die elementaren Zeilenumformungen zu beweisen.

(i) Mit einer geeigneten Transposition $\tau \in S_n$ können wir schreiben:

$$A = (\alpha_{k,j})_{1 \leq k,j \leq n}, \; A' = (\alpha_{\tau(k),j})_{1 \leq k,j \leq n}.$$

Da mit π auch $\pi \circ \tau$ ganz S_n durchläuft, folgt somit

$$\begin{aligned} \det(A') &= \sum_{\pi \in S_n} \mathrm{sgn}(\pi) \alpha_{\tau(1),\pi(1)} \cdot \ldots \cdot \alpha_{\tau(n),\pi(n)} \\ &= \sum_{\tau \circ \pi \in S_n} \mathrm{sgn}(\pi) \alpha_{1,(\pi \circ \tau)(1)} \cdot \ldots \cdot \alpha_{n,(\pi \circ \tau)(n)}. \end{aligned}$$

Beachten wir nun $\mathrm{sgn}(\pi) = -\mathrm{sgn}(\pi \circ \tau)$ und schreiben wir anstelle von $\pi \circ \tau$ wieder π, so folgt

$$\det(A') = -\sum_{\pi \in S_n} \mathrm{sgn}(\pi) \alpha_{1,\pi(1)} \cdot \ldots \cdot \alpha_{n,\pi(n)} = -\det(A).$$

Sind in der Matrix A insbesondere zwei Zeilen gleich und ist A' die Matrix die aus A durch Vertauschen dieser beiden Zeilen entstanden ist, so gilt einerseits

$$\det(A') = -\det(A);$$

andererseits folgt wegen $A' = A$

$$\det(A) = -\det(A),$$

also $\det(A) = 0$, falls die Charakteristik von K ungleich zwei ist. Da die Gleichung $\det(A) = 0$ eine polynomiale Relation darstellt, die bei Vorhandensein zweier gleicher Zeilen erfüllt werden muß, besteht diese Gleichheit auch im Fall der Charakteristik zwei.

(ii) Dies ergibt sich unmittelbar aus der Determinantendefinition. Insbesondere erhalten wir

$$\begin{aligned}\det(\lambda A) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)(\lambda\alpha_{1,\pi(1)})\cdot\ldots\cdot(\lambda\alpha_{n,\pi(n)}) \\ &= \lambda^n \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{n,\pi(n)} \\ &= \lambda^n \det(A).\end{aligned}$$

(iii) Die Matrix $A' = (\alpha'_{k,j})_{1\le k,j\le n} \in \mathrm{M}_n(K)$ gehe aus der Matrix A hervor, indem die k-te Zeile von A durch die Summe der k-ten plus der k'-ten Zeile von $A\,(k, k' \in \{1,\ldots,n\}; k \neq k')$ ersetzt werde. Es ergibt sich

$$\begin{aligned}\det(A') &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\alpha'_{1,\pi(1)}\cdot\ldots\cdot\alpha'_{k,\pi(k)}\cdot\ldots\cdot\alpha'_{n,\pi(n)} \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot(\alpha_{k,\pi(k)}+\alpha_{k',\pi(k)})\cdot\ldots\cdot\alpha_{n,\pi(n)} \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{k,\pi(k)}\cdot\ldots\cdot\alpha_{n,\pi(n)} + \\ &\quad \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{k',\pi(k)}\cdot\ldots\cdot\alpha_{n,\pi(n)}.\end{aligned}$$

Die erstere Summe ist nun gleich $\det(A)$, die letztere Summe ist die Determinante einer Matrix mit gleicher k-ter und k'-ter Zeile, also nach (i) gleich Null. Somit folgt wie behauptet

$$\det(A') = \det(A). \qquad \square$$

Bemerkung. Der Beweis von (iii) zeigt sogar, daß sich die Determinante einer Matrix nicht ändert, wenn eine Zeile bzw. eine Spalte ersetzt wird durch die Summe dieser Zeile bzw. dieser Spalte und einer Linearkombination der übrigen Zeilen bzw. Spalten.

Lemma. *Es sei $A = (\alpha_{k,j})_{1\le k,j\le n} \in \mathrm{M}_n(K)$ eine obere Dreiecksmatrix, d.h.*

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & & \alpha_{1,n} \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \alpha_{n,n} \end{pmatrix}.$$

Dann gilt

$$\det(A) = \alpha_{1,1} \cdot \ldots \cdot \alpha_{n,n}.$$

Beweis. Es ist

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \alpha_{1,\pi(1)} \cdot \ldots \cdot \alpha_{n,\pi(n)}.$$

Ist nun $\pi \in S_n, \pi \neq \mathrm{id}$, so existiert ein $k \in \{1, \ldots, n\}$ mit $k > \pi(k)$; das zugehörige $\alpha_{k,\pi(k)}$ ist dann Null. Somit liefert in obiger Summe höchstens $\pi = \mathrm{id}$ einen von Null verschiedenen Beitrag, d.h.

$$\det(A) = \alpha_{1,1} \cdot \ldots \cdot \alpha_{n,n}.$$

□

Beispiel. Berechnung von $\det(A)$ für die Matrix

$$A = \begin{pmatrix} 1 & 3 & 7 \\ 2 & 8 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

Die vorhergehenden Überlegungen führen zu

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 3 & 7 \\ 2 & 8 & 4 \\ 3 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 3 & 7 \\ 0 & 2 & -10 \\ 0 & -7 & -20 \end{vmatrix} = 2 \cdot \begin{vmatrix} 1 & 3 & 7 \\ 0 & 1 & -5 \\ 0 & -7 & -20 \end{vmatrix} \\ &= 2 \cdot \begin{vmatrix} 1 & 3 & 7 \\ 0 & 1 & -5 \\ 0 & 0 & -55 \end{vmatrix} = 2 \cdot 1 \cdot 1 \cdot (-55) = -110. \end{aligned}$$

Satz. *Es sei* $A = (\alpha_{k,j})_{1 \leq k,j \leq n} \in \mathrm{M}_n(K)$ *eine quadratische Matrix. Dann besteht die Äquivalenz*

$$A \text{ regulär} \Longleftrightarrow \det(A) \neq 0.$$

Beweis. Mit Hilfe elementarer Zeilen- und Spaltenumformungen kann A umgeformt werden zu

$$A' = \begin{pmatrix} \alpha'_{1,1} & \ldots\ldots\ldots & & \alpha'_{1,r} & \alpha'_{1,r+1} & \cdots & \alpha'_{1,n} \\ 0 & \alpha'_{2,2} & \cdots & \alpha'_{2,r} & \alpha'_{2,r+1} & \cdots & \alpha'_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \alpha'_{r,r} & \alpha'_{r,r+1} & \cdots & \alpha'_{r,n} \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix}$$

mit $\alpha'_{1,1}, \ldots, \alpha'_{r,r} \neq 0$. Für die Determinante von A' ergibt sich mit Hilfe der vorhergehenden Proposition die Beziehung

$$\det(A) = C \cdot \det(A'),$$

wobei C ein von Null verschiedener Skalar ist, der sich mit Hilfe der benutzten elementaren Zeilen- und Spaltenumformungen berechnet. Das vorhergehende Lemma liefert dann die Formel

$$\det(A) = C \cdot \alpha'_{1,1} \cdot \ldots \cdot \alpha'_{r,r} \cdot 0 \cdot \ldots \cdot 0.$$

Die bereits bekannte Äquivalenz

$$A\,\text{regulär} \Longleftrightarrow r = n$$

führt somit unmittelbar zur behaupteten Äquivalenz

$$A\,\text{regulär} \Longleftrightarrow \det(A) \neq 0.$$

□

Bemerkung. Die Aussage des vorhergehenden Satzes kann auch in der Form

$$A\,\text{singulär} \Longleftrightarrow \det(A) = 0$$

wiedergegeben werden.

Definition. Es sei $A = (\alpha_{k,j})_{1 \leq k,j \leq n} \in \mathrm{M}_n(K)$ eine quadratische Matrix. Die quadratische Matrix $A_{k,j} \in \mathrm{M}_{n-1}(K)$, die aus A durch Streichen der k-ten Zeile und j-ten Spalte hervorgeht, heißt das *algebraische Komplement des Matrixelements* $\alpha_{k,j}$. Weiter heißt die Größe

$$\mathrm{M}_{k,j} := (-1)^{k+j} \det(A_{k,j})$$

die *Adjunkte* oder der *Minor von* $\alpha_{k,j}$.

Satz (Entwicklungssatz von Laplace). *Es sei* $A = (\alpha_{k,j})_{1 \leq k,j \leq n} \in \mathrm{M}_n(K)$ *eine quadratische Matrix. Dann besteht die Formel „Entwicklung nach der* k*-ten Zeile von* A*“* $(k = 1, \ldots, n)$

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} \alpha_{k,j} (-1)^{k+j} \det(A_{k,j}) \\ &= \sum_{j=1}^{n} \alpha_{k,j} \mathrm{M}_{k,j}. \end{aligned}$$

Analog besteht die Formel „Entwicklung nach der j*-ten Spalte von* A*“* $(j = 1, \ldots, n)$

$$\begin{aligned} \det(A) &= \sum_{k=1}^{n} \alpha_{k,j} (-1)^{k+j} \det(A_{k,j}) \\ &= \sum_{k=1}^{n} \alpha_{k,j} \mathrm{M}_{k,j}. \end{aligned}$$

Beweis. Aufgrund der Invarianz der Determinantenbildung gegenüber Transposition genügt es, die Entwicklung der Determinante von A nach einer Zeile zu betrachten. Für diese Betrachtung genügt es ohne Beschränkung der Allgemeinheit, die Entwicklung der

Determinante von A nach der n-ten Zeile zu untersuchen. Aufgrund der Leibnizschen Determinantendefinition erhalten wir in diesem Fall zunächst

$$\begin{aligned}\det(A) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{n,\pi(n)}\\ &= \alpha_{n,1}\sum_{\substack{\pi\in S_n\\ \pi(n)=1}} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{n-1,\pi(n-1)}+\ldots+\\ &\qquad \alpha_{n,n}\sum_{\substack{\pi\in S_n\\ \pi(n)=n}} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{n-1,\pi(n-1)}.\end{aligned}$$

Wir bezeichnen den j-ten Summanden ($j = 1, \ldots, n$) in obiger Summe mit $\tilde{S}_j$, d.h.

$$\tilde{S}_j := \alpha_{n,j}\sum_{\substack{\pi\in S_n\\ \pi(n)=j}} \operatorname{sgn}(\pi)\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{n-1,\pi(n-1)}.$$

Für $j = n$ wird in der $\tilde{S}_n$ definierenden Summe über alle $\pi \in S_n$ mit $\pi(n) = n$ summiert; diese Permutationen entsprechen aber gerade den Permutationen $\pi' \in S_{n-1}$. Also folgt

$$\begin{aligned}\tilde{S}_n &= \alpha_{n,n}\sum_{\pi'\in S_{n-1}} \operatorname{sgn}(\pi')\alpha_{1,\pi'(1)}\cdot\ldots\cdot\alpha_{n-1,\pi'(n-1)}\\ &= (-1)^{n+n}\alpha_{n,n}\det(A_{n,n}) = \alpha_{n,n}\mathrm{M}_{n,n}.\end{aligned}$$

Wenden wir uns nun dem allgemeinen Fall zu: Hierbei wird für festes $j \in \{1, \ldots, n\}$ über alle Permutationen $\pi \in S_n$ mit $\pi(n) = j$ summiert, d.h. über alle Permutationen der Form

$$\begin{pmatrix} 1 & 2 & \ldots & n-1 & n\\ \pi(1) & \pi(2) & \ldots & \pi(n-1) & j\end{pmatrix}.$$

Komponieren wir die zur Diskussion stehenden π's von links sukzessive mit den $(n-j)$ Transpositionen $\tau_1 = (j, j+1), \ldots, \tau_{n-j} = (n-1, n)$, so erhalten wir Permutationen $\tilde{\pi}$ der Form

$$\begin{aligned}\tilde{\pi} &= \tau_{n-j}\circ\ldots\circ\tau_1\circ\pi\\ &= (n-1,n)\circ\ldots\circ(j,j+1)\circ\begin{pmatrix} 1 & 2 & \ldots & n-1 & n\\ \pi(1) & \pi(2) & \ldots & \pi(n-1) & j\end{pmatrix}\\ &= \begin{pmatrix} 1 & 2 & \ldots & n-1 & n\\ * & * & \ldots & * & n\end{pmatrix},\end{aligned}$$

d.h. es gilt $\tilde{\pi}(n) = n$; diese Permutationen entsprechen nun aber gerade wieder den Permutationen $\pi' \in S_{n-1}$. Somit ergibt sich

$$\begin{aligned}\tilde{S}_j &= \alpha_{n,j}\sum_{\pi'\in S_{n-1}} (-1)^{n-j}\operatorname{sgn}(\pi')\alpha_{1,\pi(1)}\cdot\ldots\cdot\alpha_{n-1,\pi(n-1)}\\ &= (-1)^{n+j}\alpha_{n,j}\det(A_{n,j}) = \alpha_{n,j}\mathrm{M}_{n,j}.\end{aligned}$$

Zusammengenommen folgt

$$\det(A) = \sum_{j=1}^{n} \tilde{S}_j = \sum_{j=1}^{n} \alpha_{n,j} \mathrm{M}_{n,j},$$

wie behauptet. □

Corollar. *Es sei* $A = (\alpha_{k,j})_{1 \le k,j \le n} \in \mathrm{M}_n(K)$ *eine quadratische Matrix. Dann bestehen die Formeln*

$$\begin{aligned} \sum_{j=1}^{n} \alpha_{i,j} \mathrm{M}_{k,j} &= \delta_{i,k} \cdot \det(A) \quad (i,k = 1, \ldots, n), \\ \sum_{k=1}^{n} \alpha_{k,i} \mathrm{M}_{k,j} &= \delta_{i,j} \cdot \det(A) \quad (i,j = 1, \ldots, n). \end{aligned}$$

Beweis. Es genügt wiederum, nur die erste der beiden Formeln zu beweisen. Im Fall $i = k$ folgt die Richtigkeit der Formel unmittelbar aus dem Entwicklungssatz von Laplace. Im Fall $i \neq k$ ist die linke Seite der behaupteten Formel, d.h.

$$\sum_{j=1}^{n} \alpha_{i,j} \mathrm{M}_{k,j},$$

nach dem Entwicklungssatz gleich der Determinante derjenigen Matrix A', die aus A durch Ersetzen der k-ten Zeile durch die i-te Zeile entsteht. Da nun die Matrix A' zwei gleiche Zeilen besitzt, folgt

$$\sum_{j=1}^{n} \alpha_{i,j} \mathrm{M}_{k,j} = \det(A') = 0.$$

Insgesamt erhalten wir wie behauptet

$$\sum_{j=1}^{n} \alpha_{i,j} \mathrm{M}_{k,j} = \delta_{i,k} \cdot \det(A)$$

für $i, k = 1, \ldots, n$. □

Satz (Berechnung der Inversen). *Es sei* $A = (\alpha_{i,j})_{1 \le i,j \le n} \in \mathrm{GL}_n(K)$ *eine reguläre Matrix. Dann ist ihre Inverse* $A^{-1} \in \mathrm{GL}_n(K)$ *gegeben durch*

$$A^{-1} = \det(A)^{-1}\, {}^{t}B,$$

wobei $B = (\mathrm{M}_{j,k})_{1 \le j,k \le n} \in \mathrm{M}_n(K)$ *die Matrix bestehend aus den Minoren von* $\alpha_{j,k}$ $(1 \le j, k \le n)$ *ist.*

Beweis. Mit Hilfe der ersten Formel des vorhergehenden Corollar ergibt sich für den (i,k)-ten Eintrag der Produktmatrix $A \cdot {}^t\!B$

$$\sum_{j=1}^{n} \alpha_{i,j}\mathrm{M}_{k,j} = \delta_{i,k}\det(A),$$

also

$$A \cdot \det(A)^{-1}\, {}^t\!B = E.$$

Mit Hilfe der zweiten Formel des Corollar folgt in analoger Weise

$$\det(A)^{-1}\, {}^t\!BA = E.$$

Somit folgt wie behauptet

$$A^{-1} = \det(A)^{-1}\, {}^t\!B.$$

□

Satz (Cramersche Regel). *Gegeben sei das lineare Gleichungssystem*

$$\begin{array}{ccccc} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n & = & \beta_1 & & \\ \vdots & & \vdots & & (S), \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n & = & \beta_n & & \end{array}$$

mit quadratischer Koeffizientenmatrix $A = (\alpha_{k,j})_{1\le k,j\le n} \in \mathrm{M}_n(K)$*, welche* $\det(A) \neq 0$ *erfüllt. Dann besitzt* (S) *genau eine Lösung* ${}^t(\xi_1,\dots,\xi_n) \in K^n$ *gegeben durch*

$$\xi_j = \frac{1}{\det(A)}\sum_{k=1}^{n}\beta_k M_{k,j} = \frac{1}{\det(A)}\begin{vmatrix} \alpha_{1,1} & \cdots & \beta_1 & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \beta_n & \cdots & \alpha_{n,n} \end{vmatrix} \quad (j = 1,\dots,n);$$

dabei steht rechter Hand die Determinante der Matrix, die aus A *entsteht, indem der* j*-te Spaltenvektor von* A *durch den Spaltenvektor* $(\beta_k)_{1\le k\le n} \in K^n$ *ersetzt wird.*

Beweis. Aufgrund der Voraussetzung $\det(A) \neq 0$ ist A regulär, d.h. $A \in \mathrm{GL}_n(K)$. Damit wissen wir, daß das lineare Gleichungssystem (S) eindeutig lösbar ist und die Inverse $A^{-1} \in \mathrm{GL}_n(K)$ existiert. Die eindeutige Lösung ist nach dem vorhergehenden Satz mit $B = (M_{j,k})_{1\le j,k\le n}$ gegeben durch

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = A^{-1}\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \det(A)^{-1}\, {}^t\!B \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

also

$$\xi_j = \det(A)^{-1}\sum_{k=1}^{n}\mathrm{M}_{k,j}\beta_k = \det(A)^{-1}\sum_{k=1}^{n}\beta_k\mathrm{M}_{k,j}$$

für $j = 1,\dots,n$.

□

Kapitel 4

Diagonalisierbarkeit und Normalformen linearer Abbildungen

4.1 Problemstellung

In diesem Kapitel fixieren wir einen beliebigen Körper K und einen n-dimensionalen Vektorraum V über K. Wir erinnern an die Bezeichnung

$$\mathrm{L}(V) = \{f : V \longrightarrow V, \text{linear}\};$$

eine lineare Abbildung $f \in \mathrm{L}(V)$ nennen wir einen *Endomorphismus von* V.
Ist nun $f \in \mathrm{L}(V)$ gegeben, so läßt sich f bezüglich einer Basis $\mathcal{B} = \{b_1, ..., b_n\}$ von V durch die Matrix

$$A = (\alpha_{k,j})_{1 \leq k,j \leq n} \in \mathrm{M}_n(K)$$

beschreiben; hierbei gilt

$$f(b_j) = \sum_{k=1}^{n} \alpha_{k,j} \cdot b_k.$$

Die Beschreibung von f durch die Matrix A hängt ersichtlich von der Wahl der geordneten Basis $\mathcal{B}$ ab. Bekanntlich steht die Menge der geordneten Basen von V in Bijektion zur Menge $\mathrm{GL}_n(K)$. Es stellt sich somit folgende Frage: Sei $f \in \mathrm{L}(V)$ beliebig vorgegeben. Existiert dann eine Basis $\mathcal{B} = \{b_1, ..., b_n\}$ von V derart, dass die Matrix $A \in \mathrm{M}_n(K)$ von f bezüglich $\mathcal{B}$ eine möglichst einfache Gestalt, z.B. Diagonalform, besitzt, d.h.

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit gewissen $\lambda_1, ..., \lambda_n \in K$?
Wir werden im folgenden sehen, daß die obige Frage in dieser Allgemeinheit nicht bejaht

werden kann. Dennoch werden wir einerseits diejenigen Endomorphismen von V aussortieren, die sich durch Diagonalmatrizen beschreiben lassen und andererseits für alle $f \in \mathrm{L}(V)$ eine Normalform, die Jordansche Normalform, finden.

Geometrische Deutung: Falls $f \in \mathrm{L}(V)$ bezüglich einer Basis $\mathcal{B}$ durch eine Diagonalmatrix $A =$ diag $(\lambda_1, ..., \lambda_n)$ beschrieben werden kann, so heißt dies für den j-ten Basisvektor $b_j \in \mathcal{B}$:

$$f(b_j) = \lambda_j \cdot b_j,$$

d.h. f ist eine Streckung in Richtung b_j mit dem Streckungsfaktor λ_j $(j = 1, ..., n)$.

4.2 Eigenwerte, Eigenvektoren

Wir beginnen mit der folgenden

Definition. Ein Vektor $x \in V, x \neq 0$, heißt *Eigenvektor von* $f \in \mathrm{L}(V)$, falls $\lambda \in K$ mit

$$f(x) = \lambda \cdot x$$

existiert. Der Skalar $\lambda \in K$ heißt der *Eigenwert zum Eigenvektor* $x \in V$.

Analog lassen sich Eigenvektor und Eigenwert einer Matrix $A \in \mathrm{M}_n(K)$ definieren.

Definition. Ein Spaltenvektor $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n, \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$,
heißt *Eigenvektor von* $A \in \mathrm{M}_n(K)$, falls $\lambda \in K$ mit

$$A \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \lambda \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

existiert. Der Skalar $\lambda \in K$ heißt der *Eigenwert zum Eigenvektor* $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n$.

Bemerkung. Die letztere Definition ist natürlich ein Spezialfall der ersteren und hängt mit dieser wie folgt zusammen: wir fixieren eine Basis $\mathcal{B}$ von V und beschreiben $f \in \mathrm{L}(V)$ durch die Matrix $A \in \mathrm{M}_n(K)$ bezüglich $\mathcal{B}$. Mit Hilfe der Koordinatenbildung κ bezüglich $\mathcal{B}$ erhalten wir dann das folgende kommutative Diagramm:

Anhand dieses Diagramms erkennen wir, wie Eigenvektoren von f vermöge der Koordinatenabbildung κ in Eigenvektoren von A übergeführt werden. Wir bemerken allerdings, dass dieser Prozeß nicht kanonisch ist, da er von der Wahl der Basis $\mathcal{B}$ abhängt. Bei der Wahl einer anderen Basis $\mathcal{B}'$, welche vermöge der Basistransformationsmatrix $S \in \mathrm{GL}_n(K)$ aus $\mathcal{B}$ hervorgeht, wird $A \in \mathrm{M}_n(K)$ durch $A' = S^{-1} \cdot A \cdot S \in \mathrm{M}_n(K)$ ersetzt.

Definition. Es sei $A \in \mathrm{M}_n(K)$ und t eine Unbestimmte (Variable). Dann heißt das Polynom n-ten Grades

$$p_A(t) := \det(A - t \cdot E)$$

das *charakteristische Polynom von A*; hierbei ist $E \in \mathrm{M}_n(K)$ die Einheitsmatrix.

Proposition. *Es seien $A \in \mathrm{M}_n(K)$ und $p_A \in K[t]$ das charakteristische Polynom von A. Dann besteht die Äquivalenz*

$$\lambda \textit{ Eigenwert von } A \Longleftrightarrow p_A(\lambda) = 0.$$

Beweis. Wir greifen auf die Definition von Eigenwert und Eigenvektor zurück: Die Gleichung

$$A \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \lambda \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \quad \text{bzw. } (A - \lambda \cdot E) \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

hat genau dann eine nicht-triviale Lösung $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n$, wenn die Matrix $(A - \lambda \cdot E)$ singulär ist. Nach der Determinantentheorie ist dies genau dann der Fall, wenn

$$\det(A - \lambda \cdot E) = 0$$

ist, d.h. wenn λ eine Nullstelle des Polynoms p_A ist. □

Bemerkung. (i) In den Übungen wird der Multiplikationssatz für Determinanten bewiesen, d.h. sind $A, B \in \mathrm{M}_n(K)$, so gilt

$$\det(A \cdot B) = \det(A) \cdot \det(B).$$

Ist $A \in \mathrm{GL}_n(K)$, so schließt man daraus sofort

$$\det(A^{-1}) = \det(A)^{-1}.$$

Damit erkennt man das folgende: Sind $A \in \mathrm{M}_n(K), S \in \mathrm{GL}_n(K)$ und $A' := S^{-1} \cdot A \cdot S$, so gilt

$$\begin{aligned} \det(A') &= \det(S^{-1} \cdot A \cdot S) = \det(S^{-1}) \cdot \det(A) \cdot \det(S) \\ &= \det(S)^{-1} \cdot \det(A) \cdot \det(S) = \det(A). \end{aligned}$$

(ii) Das in (i) Festgestellte findet nun die folgende Anwendung: Seien $f \in \mathrm{L}(V), \mathcal{B}$ eine Basis von V und A die Matrix von f bezüglich $\mathcal{B}$. Ist $\mathcal{B}'$ eine weitere Basis von V und $S \in \mathrm{GL}_n(K)$ die Basistransformationsmatrix von $\mathcal{B}$ nach $\mathcal{B}'$, so gilt für die Matrix A' von f bezüglich $\mathcal{B}'$

$$A' = S^{-1} \cdot A \cdot S.$$

Ist nun p_A das charakteristische Polynom zu A und $p_{A'}$ das charakteristische Polynom zu A', so erkennen wir mit Hilfe von (i)

$$\begin{aligned}
p_{A'}(t) &= \det(A' - t \cdot E) \\
&= \det(S^{-1} \cdot A \cdot S - t \cdot S^{-1} \cdot E \cdot S) \\
&= \det(S^{-1}(A - t \cdot E) \cdot S) \\
&= \det(S)^{-1} \cdot \det(A - t \cdot E) \cdot \det(S) \\
&= \det(A - t \cdot E) \\
&= p_A(t).
\end{aligned}$$

Somit stellen wir fest, daß das charakteristische Polynom unabhängig von der Matrix ist, mit welcher $f \in \mathrm{L}(V)$ beschrieben wird. Wir können mithin künftig auch vom *charakteristischen Polynom p_f von $f \in \mathrm{L}(V)$* sprechen, indem wir

$$p_f(t) := p_A(t)$$

setzen, wobei A die Matrix von f zu irgendeiner Basis von V ist.

Bemerkung. Ist $f \in \mathrm{L}(V)$ und p_f sein charakteristisches Polynom, so kann die Frage nach der Existenz von Eigenwerten bzw. Eigenvektoren gestellt werden. Da V als n-dimensional vorausgesetzt wurde, ist $p_f \in K[t]$ ein Polynom vom Grad n. Somit wissen wir, daß höchstens n Eigenwerte existieren können. Andererseits ist a priori nicht gesichert, daß überhaupt ein Eigenwert von f, d.h. eine Nullstelle von p_f in K existiert. Dies belegen die nachfolgenden Beispiele.
Grundsätzlich ist in diesem Zusammenhang der Begriff des algebraisch abgeschlossenen Körpers von Bedeutung:

Definition. Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes Polynom $p \in K[t]$ von positivem Grad eine Nullstelle in K besitzt.

Mit Hilfe der Division mit Rest zeigt man dann, dass für einen algebraisch abgeschlossenen Körper ein Polynom vom Grad n mit Vielfachheiten gezählt genau n Nullstellen hat. Z.B. ist $K = \mathbb{C}$ ein algebraisch abgeschlossener Körper; dies ist der Inhalt des sogenannten *Fundamentalsatzes der Algebra*, der hier aber nicht bewiesen werden soll.

Ist also K ein algebraisch abgeschlossener Körper, so besitzt das charakteristische Polynom p_f (mit Vielfachheiten gezählt) genau n Nullstellen, und mithin besitzt f (mit Vielfachheiten gezählt) genau n Eigenwerte.

Beispiele.

(1) Sei $A = \begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R})$. Dann erhalten wir dazu das charakteristische Polynom

$$p_A(t) = \det(A - t \cdot E) = \begin{vmatrix} -4-t & 1 \\ -2 & -1-t \end{vmatrix} =$$
$$(4+t)(1+t) + 2 = t^2 + 5t + 6 = (t+2)(t+3).$$

Somit erhalten wir die beiden Eigenwerte

$$\lambda_1 = -2, \lambda_2 = -3.$$

Wir bestimmen nun die Eigenvektoren zu $\lambda_1 = -2$. Wir haben dazu das folgende homogene lineare Gleichungssystem nicht-trivial zu lösen:

$$\begin{aligned}
\begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} &= -2 \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \\
\Longleftrightarrow \quad \left[\begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix} - (-2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\Longleftrightarrow \quad \begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\Longleftrightarrow \quad -2\xi_1 + \xi_2 &= 0 \\
\Longleftrightarrow \quad \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} &= \begin{pmatrix} \alpha \\ 2\alpha \end{pmatrix},
\end{aligned}$$

mit $\alpha \in K, \alpha \neq 0$. Damit ist $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ Eigenvektor der Matrix $\begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix}$ mit Eigenwert -2. Wir bestätigen dies mit Hilfe einer Probe

$$\begin{pmatrix} -4 & 1 \\ -2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \end{pmatrix} = -2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Entsprechend erkennt man $\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$ mit $\alpha \in K, \alpha \neq 0$, als Eigenvektor zum Eigenwert -3.

(2) Sei $A = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ mit $\phi \in (0, 2\pi), \phi \neq \pi$. Wir erhalten für das charakteristische Polynom

$$\begin{aligned}
p_A(t) &= \begin{vmatrix} \cos\phi - t & -\sin\phi \\ \sin\phi & \cos\phi - t \end{vmatrix} \\
&= (\cos\phi - t)^2 + \sin^2\phi = t^2 - 2t\cos\phi + (\cos^2\phi + \sin^2\phi) \\
&= t^2 - 2t\cos\phi + 1.
\end{aligned}$$

Damit erhalten wir für die Eigenwerte

$$\lambda_{1,2} = \cos\phi \pm \sqrt{\cos^2\phi - 1}.$$

Wegen $\cos^2\phi \neq 1$, folgt $\cos^2\phi - 1 < 0$, also folgt $\lambda_{1,2} \notin \mathbb{R}$. Somit besitzt die Drehmatrix A keine Eigenwerte in $\mathbb{R}$, also auch keine Eigenvektoren, was geometrisch natürlich unmittelbar klar ist.

4.3 Diagonalisierbarkeit

Wir erinnern an die Verabredung: K ist ein beliebiger Körper und V ein n-dimensionaler Vektorraum über K. Damit definieren wir

Definition. Es seien $f \in \mathrm{L}(V)$ und $\lambda \in K$ ein Eigenwert von f, d.h. λ ist eine Nullstelle des charakteristischen Polynoms $p_f \in K[t]$; bekanntlich ist dann $(t - \lambda)$ ein Teiler von $p_f(t)$. Der maximale, ganzzahlige Exponent $\alpha_\lambda \in \mathbb{N}$ mit der Eigenschaft

$$(t - \lambda)^{\alpha_\lambda} \mid p_f(t)$$

heißt die *algebraische Vielfachheit des Eigenwerts* λ.

Bemerkung. Sind $f \in \mathrm{L}(V)$ und $x \in V$ ein Eigenvektor von f zum Eigenwert λ, so gilt

$$f(x) = \lambda \cdot x \iff (f - \lambda \cdot \mathrm{id}) \cdot x = 0;$$

dies ist wiederum gleichbedeutend zu

$$x \in \ker(f - \lambda \cdot \mathrm{id}).$$

Definition. Es seien $f \in \mathrm{L}(V)$ und $\lambda \in K$ ein Eigenwert von f, d.h. λ ist eine Nullstelle des charakteristischen Polynoms $p_f \in K[t]$. Dann nennen wir den Unterraum

$$E_\lambda := \ker(f - \lambda \cdot \mathrm{id}) \subseteq V$$

den *Eigenraum von f zum Eigenwert* λ. Wir nennen dessen Dimension

$$\beta_\lambda := \dim_K E_\lambda$$

die geometrische Vielfachheit des Eigenwerts λ.

Bemerkung. Der Eigenraum E_λ besteht aus sämtlichen Eigenvektoren von f zum Eigenwert λ. Nach der Theorie der linearen homogenen Gleichungssysteme erhalten wir für die geometrische Vielfachheit von λ

$$\beta_\lambda = n - \mathrm{rg}(f - \lambda \cdot \mathrm{id}).$$

Beispiel. Es sei $V = \mathbb{R}^3$ und $f \in \mathrm{L}(V)$ gegeben durch die Matrix

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Es folgt

$$p_f(t) = p_A(t) = (t-2)^3,$$

d.h. f besitzt nur den Eigenwert $\lambda = 2$. Für diesen berechnen wir weiter

$$\alpha_\lambda = 3, \beta_\lambda = 3 - \mathrm{rg}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 2.$$

Lemma. *Seien $f \in \mathrm{L}(V)$ und $\lambda \in K$ ein Eigenwert von f. Dann bestehen zwischen algebraischer und geometrischer Vielfachheit die Ungleichungen*

$$1 \le \beta_\lambda \le \alpha_\lambda.$$

Beweis. Die Ungleichung $\beta_\lambda \ge 1$ folgt unmittelbar aus der Definition des Eigenwerts. Sei nun $\{b_1, ..., b_{\beta_\lambda}\}$ eine Basis des Eigenraums E_λ; wir ergänzen diese zu einer Basis $\mathcal{B}$ von V. Die Matrix A von f bezüglich $\mathcal{B}$ hat die Form

$$A = \left(\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & * \\ & & \lambda & \\ \hline & 0 & & A' \end{array}\right)$$

mit $A' \in \mathrm{M}_{n-\beta_\lambda}(K)$. Nach den Rechenregeln für Determinanten ergibt sich dann

$$p_f(t) = p_A(t) = (t-\lambda)^{\beta_\lambda} \cdot p_{A'}(t),$$

d.h. wir haben die Teilbarkeitsbeziehung

$$(t-\lambda)^{\beta_\lambda} \mid p_f(t).$$

Da α_λ maximal mit dieser Eigenschaft definiert wurde, folgt

$$\beta_\lambda \le \alpha_\lambda.$$

□

Definition. Ein Endomorphismus $f \in \mathrm{L}(V)$ heißt *diagonalisierbar*, falls V eine Basis von Eigenvektoren besitzt.

Bemerkung. Seien $f \in \mathrm{L}(V)$ und $A \in \mathrm{M}_n(K)$ die Matrix von f bezüglich irgendeiner Basis $\mathcal{B}$. Ist f diagonalisierbar, so können wir speziall eine Basis $\mathcal{B}'$ bestehend aus Eigenvektoren von f wählen. Bezeichnet $A' \in \mathrm{M}_n(K)$ die Matrix von f bezüglich $\mathcal{B}'$, so folgt mit der Basistransformationsmatrix $S \in \mathrm{GL}_n(K)$ von $\mathcal{B}$ nach $\mathcal{B}'$

$$A' = S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

d.h. A' ist eine Diagonalmatrix; hierbei sind $\lambda_1, ..., \lambda_n \in K$ die Eigenwerte von f.

Definition. Eine Matrix $A \in \mathrm{M}_n(K)$ heißt *diagonalisierbar*, falls eine invertierbare Matrix $S \in \mathrm{GL}_n(K)$ existiert, so daß $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix ist.

Lemma. *Es sei $f \in \mathrm{L}(V)$ ein Endomorphismus. Eigenvektoren zu paarweise verschiedenen Eigenwerten von f sind linear unabhängig.*

Beweis. Es seien $v_1, ..., v_m$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$ von f. Zum Beweis der Behauptung führen wir eine vollständige Induktion über m. Ist $m = 1$, so ist die Behauptung wegen $v_1 \neq 0$ klar. Die Behauptung sei also für $(m-1)$ Eigenvektoren bewiesen. Für den Induktionsschritt betrachten wir die Linearkombination

$$\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + ... + \alpha_m \cdot v_m = 0$$

mit zunächst unbestimmten Skalaren $\alpha_1, ..., \alpha_m$. Nach Anwendung von f auf diese Relation, bzw. Multiplikation dieser Relation mit λ_1 folgt:

$$\alpha_1 \lambda_1 \cdot v_1 + \alpha_2\, \lambda_2 \cdot v_2 + ... + \alpha_m \lambda_m \cdot v_m = 0,$$

bzw.

$$\alpha_1 \lambda_1 \cdot v_1 + \alpha_2\, \lambda_1 \cdot v_2 + ... + \alpha_m \lambda_1 \cdot v_m = 0.$$

Subtraktion dieser beiden Gleichungen ergibt

$$\alpha_2(\lambda_2 - \lambda_1) \cdot v_2 + ... + \alpha_m(\lambda_m - \lambda_1) \cdot v_m = 0.$$

Da $\lambda_2, ..., \lambda_m$ von λ_1 verschieden sind, folgt nach der Induktionsvoraussetzung

$$\alpha_2 = ... = \alpha_m = 0.$$

Damit erhalten wir $\alpha_1 \cdot v_1 = 0$, woraus $\alpha_1 = 0$ folgt, was die lineare Unabhängigkeit von $v_1, ..., v_m$ beweist. $\square$

Satz. *Es sei $f \in \mathrm{L}(V)$. Dann besteht die Äquivalenz:*

$$f \text{ diagonalisierbar} \iff \begin{array}{l} p_f(t) = \pm(t - \lambda_1)^{\alpha_1} \cdot ... \cdot (t - \lambda_r)^{\alpha_r}, \\ \alpha_j = \beta_j \qquad (j = 1, ..., r). \end{array}$$

Hierbei haben wir für $\alpha_{\lambda_j}, \beta_{\lambda_j}$ abkürzungshalber α_j, β_j $(j = 1, ..., r)$ geschrieben.

Beweis. $\Longrightarrow$: Sei also f diagonalisierbar, d.h. es existiert eine Basis $\mathcal{B}$ von V bestehend aus Eigenvektoren; seien $\lambda_1, ..., \lambda_r$ die dazugehörigen, paarweise verschiedenen Eigenwerte. Indem wir die Matrix von f bezüglich $\mathcal{B}$ bilden, erhalten wir für das charakteristische Polynom von f sofort
$$p_f(t) = \pm(t - \lambda_1)^{\alpha_1} \cdot ... \cdot (t - \lambda_r)^{\alpha_r},$$
wobei α_j die algebraische Vielfachheit zu λ_j $(j = 1, ..., r)$ bedeutet. Zum Eigenwert λ_j von f existieren definitonsgemäß β_j linear unabhängige Eigenvektoren. Nach dem vorhergehenden Lemma gibt es somit maximal
$$\beta_1 + ... + \beta_r$$
linear unabhängige Eigenvektoren zu f. Da f diagonalisierbar ist, muß
$$\beta_1 + ... + \beta_r = n$$
gelten. Da andererseits $\beta_j \leq \alpha_j$ gilt, ist obige Gleichheit nur möglich, falls
$$\alpha_j = \beta_j \quad (j = 1, ..., r)$$
gilt.
$\Longleftarrow$: Die Umkehrung ist nach dem Vorhergehenden klar. $\square$

Corollar. *Es sei $f \in \mathrm{L}(V)$ ein Endomorphismus mit $n = \dim_K V$ paarweise verschiedenen Eigenwerten. Dann ist f diagonalisierbar.*

Beweis. Voraussetzungsgemäß ist das charakteristische Polynom von f gegeben durch
$$p_f(t) = \pm(t - \lambda_1) \cdot ... \cdot (t - \lambda_n),$$
wobei $\lambda_1, ..., \lambda_n$ die n paarweise verschiedenen Eigenwerte von f sind, d.h. die algebraische Vielfachheit α_j von λ_j ist Eins. Aufgrund der Ungleichungen
$$1 = \alpha_j \geq \beta_j \geq 1$$
folgt sofort
$$\alpha_j = \beta_j = 1 \quad (j = 1, ..., r).$$
Nach dem vorhergehenden Satz ist f somit diagonalisierbar. $\square$

Bemerkung. Wir erinnern an den Begriff der direkten Summe: Dazu seien $U_1, U_2 \subseteq V$ zwei Unterräume und $U = U_1 + U_2$. Diese Summe heißt *direkt*, falls jeder Vektor $u \in U$ in genau einer Weise als Summe
$$u = u_1 + u_2$$
mit $u_1 \in U_1, u_2 \in U_2$ dargestellt werden kann. Wir schreiben dann
$$U = U_1 \oplus U_2$$

und sagen, daß U *direkte Summe* von U_1, U_2 ist.

Satz. *Es sei $f \in \mathrm{L}(V)$. Dann besteht die Äquivalenz:*

$$f \text{ diagonalisierbar} \iff V = \bigoplus_{j=1}^{r} E_{\lambda_j},$$

wobei $\lambda_1, ..., \lambda_r$ die paarweise verschiedenen Eigenwerte von f sind.

Beweis. $\Longrightarrow$: Da f diagonalisierbar ist, zerfällt das charakteristische Polynom nach dem vorhergehenden Satz in Linearfaktoren

$$p_f(t) = \pm(t - \lambda_1)^{\alpha_1} \cdot ... \cdot (t - \lambda_r)^{\alpha_r},$$

und es gilt

$$\alpha_j = \beta_j \ (j = 1, ..., r).$$

Indem wir Basen von $E_{\lambda_1}, ..., E_{\lambda_r}$ vereinigen, erhalten wir

$$\beta_1 + ... + \beta_r$$

linear unabhängige Vektoren von V. Wegen

$$\beta_1 + ... + \beta_r = \alpha_1 + ... + \alpha_r = n$$

folgt somit

$$V = \sum_{j=1}^{r} E_{\lambda_j}.$$

Ist nun $v \in V$, so gilt also

$$v = v_1 + ... + v_r$$

mit $v_j \in E_{\lambda_j}$ $(j = 1, ..., r)$. Ist

$$v = v_1' + ... + v_r'$$

eine weitere Darstellung von v mit $v_j' \in E_{\lambda_j}$ $(j = 1, ..., r)$, so folgt

$$(v_1 - v_1') + ... + (v_r - v_r') = 0.$$

Aufgrund der linearen Unabhängigkeit von Eigenvektoren zu paarweise verschiedenen Eigenwerten folgt somit

$$v_1 = v_1', ..., v_r = v_r',$$

also

$$V = \bigoplus_{j=1}^{r} E_{\lambda_j}.$$

$\Longleftarrow$: Gilt umgekehrt $V = \bigoplus_{j=1}^{r} E_{\lambda_j}$ mit r paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_r$, so besitzt V insbesondere eine Basis von Eigenvektoren; somit ist f diagonalisierbar. $\square$

Beispiel. Es sei

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \in \mathrm{M}_3(\mathbb{R}).$$

Entscheide, ob A diagonalisierbar ist und bestimme gegebenenfalls $S \in \mathrm{GL}_n(\mathbb{R})$, so daß $S^{-1} \cdot A \cdot S$ diagonal ist. Es ist

$$\begin{aligned} p_A(t) &= \begin{vmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{vmatrix} \\ &= (-t) \begin{vmatrix} -2-t & 3 \\ -2 & 3-t \end{vmatrix} + 3 \begin{vmatrix} -1 & 1 \\ -2 & 3-t \end{vmatrix} - 2 \begin{vmatrix} -1 & 1 \\ -2-t & 3 \end{vmatrix} \\ &= (-t)((t+2)(t-3)+6) + 3(t-3+2) - 2(-3+t+2) \\ &= -t(t^2-t) + 3(t-1) - 2(t-1) \\ &= -t^3 + t^2 + t - 1 = -(t-1)^2(t+1). \end{aligned}$$

Es folgt $\lambda_1 = 1,\ \alpha_1 = 2,\ \lambda_2 = -1,\ \alpha_2 = 1.$

$\lambda_1 = 1$: Betrachte

$$\begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Es folgt

$$\dim E_{\lambda_1} = 3 - \mathrm{rg}(A - \lambda_1 \cdot E) = 3 - 1 = 2,$$

also

$$\beta_1 = 2 = \alpha_1.$$

$\lambda_2 = -1$: Betrachte

$$\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Es folgt

$$\dim E_{\lambda_2} = 3 - \mathrm{rg}(A - \lambda_2 \cdot E) = 3 - 2 = 1,$$

also

$$\beta_2 = 1 = \alpha_2.$$

Damit ist A diagonalisierbar. Zur Bestimmung von S betrachten wir das Gleichungssystem

$$-\xi_1 - \xi_2 + \xi_3 = 0,$$

welches zur Basis

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

von E_{λ_1} führt. Weiter betrachten wir das Gleichungssystem

$$\begin{array}{rcrcrcl} \xi_1 & - & \xi_2 & + & \xi_3 & = & 0 \\ -3\xi_1 & - & \xi_2 & + & 3\xi_3 & = & 0 \\ -2\xi_1 & - & 2\xi_2 & + & 4\xi_3 & = & 0 \end{array} \iff \begin{array}{rcrcrcl} \xi_1 & - & \xi_2 & + & \xi_3 & = & 0 \\ & & -4\xi_2 & + & 6\xi_3 & = & 0\,, \end{array}$$

welches zur Basis

$$\left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right\}$$

von E_{λ_2} führt. Damit setzen wir

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}$$

und berechnen

$$S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Damit folgt:

$$\begin{aligned} A' &= S^{-1} \cdot A \cdot S \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -3 \\ 1 & 1 & -2 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned}$$

4.4 Minimalpolynome

Wir erinnern daran, daß eine Teilmenge $\mathfrak{a}$ eines kommutativen Rings R Ideal genannt wird, falls $\mathfrak{a}$ bezüglich der Addition eine abelsche Gruppe ist und $R \cdot \mathfrak{a} = \mathfrak{a} \cdot R \subseteq \mathfrak{a}$ gilt; speziell können wir von einem Ideal $\mathfrak{a}$ des Polynomrings $K[t]$ sprechen. Ein Ideal $\mathfrak{a} \subseteq R$ heißt Hauptideal, falls $p \in \mathfrak{a}$ existiert, so daß $\mathfrak{a} = (p)$ ist, d.h. alle Elemente von $\mathfrak{a}$ sind Vielfache von p.

Proposition. *Jedes Ideal $\mathfrak{a} \subseteq K[t]$ ist ein Hauptideal.*

Beweis. Ohne Einschränkung können wir $\mathfrak{a} \neq (0)$ annehmen. Wir setzen dann

$$d := \min\{\delta \mid \exists f \in \mathfrak{a}, f \neq 0, \deg f = \delta\}$$

und wählen $p \in \mathfrak{a}$ mit $\deg p = d$. Mit Hilfe des Euklidischen Algorithmus erhalten wir für ein beliebiges $a \in \mathfrak{a}$

$$a = p \cdot q + r$$

mit $q, r \in K[t]$ und $r = 0$ oder $\deg r < \deg p$. Wäre nun $r \neq 0$, so wäre

$$r = a - p \cdot q \in \mathfrak{a}$$

ein von Null verschiedenes Element des Ideals $\mathfrak{a}$, dessen Grad echt kleiner als $\deg p$ wäre. Dies ist aber aufgrund der Wahl von p nicht möglich, also gilt $r = 0$ und somit $a = p \cdot q$. Dies beweist $\mathfrak{a} = (p)$. □

Bemerkung. Ist $\mathfrak{a} \subseteq K[t], \mathfrak{a} \neq 0$, so können wir das erzeugende Polynom p im Beweis der vorhergehenden Proposition mit einem beliebigen Skalar ungleich Null multiplizieren und erhalten wieder ein erzeugendes Element des Ideals. Damit können wir erreichen, daß der höchste Koeffizient von p Eins ist; wir nennen p normiert. Damit ist p eindeutig bestimmt.

Definition. Es sei $\mathfrak{a} \subseteq K[t], \mathfrak{a} \neq 0$, ein Ideal. Dann heißt das normierte Polynom m, welches $\mathfrak{a}$ erzeugt, das *Minimalpolynom von* $\mathfrak{a}$.

Beispiel. Es sei $f \in \mathrm{L}(V)$ ein Endomorphismus; dann sind auch $f^j = f \circ ... \circ f$ (j-mal) Endomorphismen von V. Ist speziell $p \in K[t]$,

$$p(t) = a_d \cdot t^d + ... + a_1 \cdot t + a_0,$$

so können wir anstelle von t den Endomorphismus f einsetzen und erhalten

$$p(f) = a_d \cdot f^d + ... + a_1 \cdot f + a_0 \cdot \mathrm{id} \in \mathrm{L}(V);$$

unter Umständen kann dies der triviale Endomorphismus $0 \in \mathrm{L}(V)$ sein. Wir setzen

$$\mathfrak{a}_f := \{p \in K[t] \mid p(f) = 0 \in \mathrm{L}(V)\}.$$

Man verifiziert sofort, daß $\mathfrak{a}_f$ ein Ideal ist. Wir bezeichnen das nach dem Vorhergehenden existierende Minimalpolynom von $\mathfrak{a}_f$ durch $m_f \in K[t]$. Dies ist das normierte Polynom kleinsten Grades mit $m_f(f) = 0 \in \mathrm{L}(V)$.

Definition. Sei $f \in \mathrm{L}(V)$ und $\mathfrak{a}_f$ wie zuvor. Wir nennen $m_f \in K[t]$ das *Minimalpolynom von f.*

Bemerkung. Ist $A \in \mathrm{M}_n(K)$, so kann man speziell auch vom Ideal $\mathfrak{a}_A$ und seinem Minimalpolynom m_A sprechen. Aufgrund der intrinsischen Konstruktion ist klar, dass mit $S \in \mathrm{GL}_n(K)$ die Gleichheit $\mathfrak{a}_A = \mathfrak{a}_{S^{-1}AS}$ und somit

$$m_A(t) = m_{S^{-1}AS}(t)$$

gilt. Dies läßt sich aber auch direkt nachrechnen.

Satz. (Cayley-Hamilton). *Seien $f \in \mathrm{L}(V)$ und $p_f \in K[t]$ das charakteristische Polynom von f. Dann gilt $p_f(f) = 0 \in \mathrm{L}(V)$.*

Bemerkung. Mit den obigen Bezeichnungen haben wir: $p_f \in \mathfrak{a}_f$, also $m_f \mid p_f$.

Beweis. Nach Wahl einer Basis $\mathcal{B}$ von V wird der Endomorphismus $f \in \mathrm{L}(V)$ durch die Matrix $A \in \mathrm{M}_n(K)$ dargestellt. Die Behauptung des Satzes von Cayley-Hamilton übersetzt sich dann wie folgt in die matrizielle Sprache: Ist $A \in \mathrm{M}_n(K)$ und $p_A \in K[t]$ das charakteristische Polynom von A, so gilt $p_A(A) = 0 \in \mathrm{M}_n(K)$.

Um den Satz zu beweisen, genügt es, die letztere Behauptung zu beweisen. Dazu betrachten wir die Matrix

$$B(t) := {}^t(A - t \cdot E) \in \mathrm{M}_n(K[t]);$$

die Einträge von $B(t)$ außerhalb der Diagonalen sind Elemente von K, in der Diagonalen stehen lineare Polynome. Weiter gilt

$$\det B(t) = \det(A - t \cdot E) = p_A(t).$$

Wir ersetzen nun die Unbestimmte t durch die Matrix A und jeden Eintrag $\alpha_{k,j}$ durch $\alpha_{k,j} \cdot E$; damit erhalten wir

$$B(A) = \begin{pmatrix} \alpha_{1,1} \cdot E - A & & \alpha_{n,1} \cdot E \\ \alpha_{1,2} \cdot E & & \vdots \\ & \dots & \\ \vdots & & \alpha_{n,n-1} \cdot E \\ \alpha_{1,n} \cdot E & & \alpha_{n,n} \cdot E - A \end{pmatrix} \in \mathrm{M}_n(K[A])\,.$$

Mit Hilfe der Standardbasis $\{e_1, ..., e_n\}$ von K^n berechnen wir nun leicht

$$B(A) \cdot \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}e_1 - Ae_1 & + & \alpha_{2,1}e_2 & + & ... & + & \alpha_{n,1}e_n \\ & & & \vdots & & & \\ \alpha_{1,n}e_1 & + & \alpha_{2,n}e_2 & + & ... & + & \alpha_{n,n}e_n - Ae_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Es sei nun $\widetilde{B}(t) \in \mathrm{M}_n(K[t])$ die Matrix der Minoren von $(A - t \cdot E)$. Gemäß Corollar, Seite 61, erhalten wir dann

$${}^t(A - t \cdot E) \cdot \widetilde{B}(t) = \det B(t) \cdot E,$$

d.h.

$$\widetilde{B}(t) \cdot B(t) = p_A(t) \cdot E.$$

Einsetzen von A anstelle von t und Multiplikation mit $\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$ liefert unter Berücksichtigung der vorhergehenden Rechnungen

$$\begin{pmatrix} p_A(A) & & \\ & \ddots & \\ & & p_A(A) \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \widetilde{B}(A) \cdot B(A) \cdot \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

d.h.

$$p_A(A) \cdot e_j = 0$$

für $j = 1, ..., n$. Damit folgt

$$p_A(A) = 0 \in \mathrm{M}_n(K),$$

wie behauptet. □

Beispiel. Sei

$$A = \begin{pmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & 1 & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} \in \mathrm{M}_n(K)$$

mit $d-1$ Einsen oberhalb der Hauptdiagonalen, $1 \leq d \leq n$. Man prüft leicht nach:

$$p_A(t) = \pm\, t^n, \quad m_A(t) = t^d.$$

Extremfälle:

$$\begin{aligned} d = 1 &\implies & p_A(t) &= \pm\, m_A(t)^n, \\ d = n &\implies & p_A(t) &= \pm\, m_A(t). \end{aligned}$$

Satz. *Seien $f \in \mathrm{L}(V)$ und $p_f \in K[t]$, bzw. $m_f \in K[t]$ das charakteristische, bzw. das Minimalpolynom von f. Dann gilt:*
(i) $m_f \,|\, p_f$.
(ii) $p_f \,|\, m_f^n$.

Beweis. (i) Dies folgt unmittelbar aus dem Satz von Cayley-Hamilton.
(ii) Wir nehmen an, daß K algebraisch abgeschlossen ist; dies ist keine Einschränkung, da man andernfalls den Körper K in einen algebraisch abgeschlossenen Körper einbetten kann und dann über diesem Körper argumentiert (Basiswechsel). Somit haben wir

$$p_f(t) = \pm\, (t - \lambda_1)^{\alpha_1} \cdot ... \cdot (t - \lambda_r)^{\alpha_r}$$

mit paarweise verschiedenen $\lambda_1, ..., \lambda_r$. Aufgrund von (i) gilt

$$m_f(t) = (t - \lambda_1)^{\gamma_1} \cdot ... \cdot (t - \lambda_r)^{\gamma_r}$$

mit

$$0 \le \gamma_j \le \alpha_j \quad (j = 1, ..., r).$$

Wir haben $\gamma_j \ge 1 \quad (j = 1, ..., r)$ nachzuweisen. Im Gegenteil dazu nehmen wir z.B. $\gamma_1 = 0$ an. Es sei dann $v_1 \in E_{\lambda_1} \setminus \{0\}$ ein Eigenvektor zu λ_1. Einerseits gilt dann wegen $m_f(f) = 0$ die Beziehung

$$m_f(f) \cdot v_1 = 0.$$

Andererseits gilt für $j = 2, ..., r$

$$(f - \lambda_j \cdot \mathrm{id}) \cdot v_1 = (\lambda_1 - \lambda_j) \cdot v_1 \in E_{\lambda_1} \setminus \{0\},$$

also auch

$$(f - \lambda_j \cdot \mathrm{id})^k \cdot v_1 \in E_{\lambda_1} \setminus \{0\}$$

für $k = 1, 2, 3, \ldots$. Dies beweist sofort

$$m_f(f) \cdot v_1 = (f - \lambda_2 \cdot \mathrm{id})^{\gamma_2} \cdot ... \cdot (f - \lambda_r \cdot \mathrm{id})^{\gamma_r} \cdot v_1 \ne 0,$$

ein Widerspruch. Somit gilt $\gamma_j \ge 1$ für $j = 1, ..., r$, und es folgt daraus

$$p_f \,|\, m_f^n.$$

□

Definition. Ein Endomorphismus $f \in \mathrm{L}(V)$ heißt nilpotent, falls eine Basis $\mathcal{B}$ von V existiert, so daß die Matrix $A \in \mathrm{M}_n(K)$ von f bezüglich $\mathcal{B}$ die Form

$$A = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

hat.

Bemerkung. Man entnimmt dieser Definition sofort:

$$\begin{array}{rcl} p_f(t) & = & \pm\, t^n, \\ m_f(t) & = & t^d \quad (1 \le d \le n). \end{array}$$

Insbesondere stellen wir also fest, dass

$$f^k = 0 \in \mathrm{L}(V)$$

für $k \ge d$ gilt.
Wir bemerken, daß auch die Umkehrung davon gilt, d.h. ist $f \in \mathrm{L}(V)$ und existiert $k \in \mathbb{N}$ mit $f^k = 0$, so ist f nilpotent.

4.5 Die Jordansche Normalform

Es seien K ein Körper und V ein K-Vektorraum mit $n = \dim_K V$.

Definition. Es seien $f \in \mathrm{L}(V)$ und $\lambda \in K$ ein Eigenwert von f mit der algebraischen Vielfachheit α. Dann nennen wir

$$\begin{aligned} H_\lambda &:= \{v \in V | (f - \lambda \cdot \mathrm{id})^\alpha = 0\} \\ &= \ker(f - \lambda \cdot \mathrm{id})^\alpha \end{aligned}$$

den *Hauptraum von f zum Eigenwert λ.*

Satz 1. *Es sei $f \in \mathrm{L}(V)$, und das charakteristische Polynom von f zerfalle in Linearfaktoren*

$$p_f(t) = \pm(t - \lambda_1)^{\alpha_1} \cdot \ldots \cdot (t - \lambda_r)^{\alpha_r}.$$

Dann gilt:

(i) $f(H_{\lambda_j}) \subseteq H_{\lambda_j}$, $\dim_K H_{\lambda_j} = \alpha_j \quad (j = 1, \ldots, r)$.

(ii) $V = \bigoplus_{j=1}^{r} H_{\lambda_j}$.

(iii) $f = f_D + f_N$, *wobei $f_D \in \mathrm{L}(V)$ diagonalisierbar, $f_N \in \mathrm{L}(V)$ nilpotent und $f_D \circ f_N = f_N \circ f_D$ ist.*

Beachte. Ist K algebraisch abgeschlossen, so ist die Voraussetzung des Satzes automatisch erfüllt.

Bemerkung. Der vorhergehende Satz übersetzt sich wie folgt in die Sprache der Matrizen: Ist $A \in \mathrm{M}_n(K)$ und $p_A(t) = \pm(t - \lambda_1)^{\alpha_1} \cdot \ldots \cdot (t - \lambda_r)^{\alpha_r}$, so existiert $S \in \mathrm{GL}_n(K)$ mit

$$A' = S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 E_1 + N_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_r E_r + N_r \end{pmatrix},$$

wobei

$$\lambda_j E_j + N_j = \begin{pmatrix} \lambda_j & & * \\ & \ddots & \\ 0 & & \lambda_j \end{pmatrix} \in M_{\alpha_j}(K)$$

für $j = 1, \ldots, r$. Insbesondere gilt $A' = D + N$, wobei $D \in \mathrm{M}_n(K)$ diagonal und $N \in \mathrm{M}_n(K)$ nilpotent ist.

Dem Beweis des Satzes schicken wir das folgende Lemma voraus.

Lemma. *Seien $f \in \mathrm{L}(V), \lambda \in K$ ein Eigenwert von f mit algebraischer Vielfachheit α und $g := f - \lambda \cdot \mathrm{id} \in \mathrm{L}(V)$. Wir setzen*

$$d := \min\{j | \ker(g^j) = \ker(g^{j+1})\},$$

und $U := \ker(g^d), W := \mathrm{im}(g^d)$. Dann gilt:

(i) $d = \min\{j | \mathrm{im}(g^j) = \mathrm{im}(g^{j+1})\}$.

(ii) $g(U) \subseteq U, g(W) \subseteq W$, *d.h.* U, W *sind* g*-invariante Unterräume von* V.

(iii) $(g|_U)^d = 0$, $(g|_W)$ *ist ein Isomorphismus.*

(iv) $m_{g|_U}(t) = t^d$.

(v) $V = U \oplus W,\ \dim_K U = \alpha \geq d,\ \dim_K W = n - \alpha$.

Beweis. (i) Mit Hilfe der Äquivalenzen

$$\begin{aligned}
&\ker(g^j) = \ker(g^{j+1}) \iff \\
&\dim_K(\ker(g^j)) = \dim_K(\ker(g^{j+1})) \iff \\
&\dim_K(\mathrm{im}(g^j)) = \dim_K(\mathrm{im}(g^{j+1})) \iff \\
&\mathrm{im}(g^j) = \mathrm{im}(g^{j+1})
\end{aligned}$$

ergibt sich sofort die erste Behauptung.
(ii) Sei $u \in U$, d.h. $g^d(u) = 0$. Dann gilt $g^d(g(u)) = g^{d+1}(u) = g(g^d(u)) = 0$, d.h. $g(u) \in U$. Damit folgt $g(U) \subseteq U$.
Sei $w \in W$, d.h. $w = g^d(v)$ für ein $v \in V$. Dann gilt $g(w) = g(g^d(v)) = g^{d+1}(v) \in \mathrm{im}(g^{d+1}) = W$. Damit folgt $g(W) \subseteq W$.
(iii) Wir haben

$$(g|_U)^d = (g|_{\ker(g^d)})^d = g^d|_{\ker(g^d)} = 0.$$

Die lineare Abbildung

$$g|_W : W = \mathrm{im}(g^d) \longrightarrow \mathrm{im}(g|_W) = \mathrm{im}(g^{d+1}) = W$$

ist surjektiv, also auch injektiv. Somit ist $g|_W$ ein Isomorphismus.
(iv) Nach (iii) gilt $(g|_U)^d = 0$. Wir nehmen nun $(g|_U)^{d-1} = 0$ an. Dann folgt

$$\ker(g^d) = U \subseteq \ker(g^{d-1}) \subseteq \ker(g^d),$$

d.h. $\ker(g^{d-1}) = \ker(g^d)$, im Widerspruch zur Definition von d. Damit ist d minimal mit $(g|_U)^d = 0$, also $m_{g|_U}(t) = t^d$.
(v) Wir zeigen zunächst $U \cap W = \{0\}$. Dazu sei $v \in U \cap W$, d.h.

$$g^d(v) = 0,\ v = g^d(w) \quad (w \in V).$$

Zusammengenommen ergibt sich

$$g^{2d}(w) = 0,$$

d.h. $w \in \ker(g^{2d}) = \ker(g^d)$. Damit folgt

$$v = g^d(w) = 0.$$

Damit ergibt sich $U \cap W = \{0\}$. Da nun andererseits

$$\dim_K U + \dim_K W = \dim_K(\ker(g^d)) + \dim_K(\operatorname{im}(g^d)) = n$$

gilt, folgt $V = U \oplus W$.
Es sei $p_g \in K[t]$ das charakteristische Polynom von g. Wegen $p_g(t - \lambda) = p_f(t)$ hat p_g die Nullstelle $t = 0$ mit Vielfachheit α, d.h.

$$p_g(t) = t^\alpha \cdot q(t)$$

mit $q \in K[t], q(0) \neq 0$. Aufgrund der g-Invarianz von U, W und wegen $V = U \oplus W$ gilt andererseits

$$p_g(t) = p_{g|_U}(t) \cdot p_{g|_W}(t);$$

Da $g|_W$ ein Isomorphismus ist, gilt $p_{g|_W}(0) \neq 0$. Damit folgt

$$p_{g|_U}(t) = \pm t^\alpha,$$

woraus $\dim_K U = \alpha$ und somit auch $\dim_K W = n - \alpha$ folgt. Die Ungleichung $\alpha \geq d$ folgt sofort aus (iv). □

Bemerkung. Mit den Bezeichnungen des Lemma, insbesondere (v), folgt $U = H_\lambda$ und $(f - \lambda \cdot \mathrm{id})|_{H_\lambda}$ ist nilpotent.

Beweis von Satz 1. Es sei also $f \in \mathrm{L}(V)$, und das charakteristische Polynom von f zerfalle in Linearfaktoren

$$p_f(t) = \pm(t - \lambda_1)^{\alpha_1} \cdot \ldots \cdot (t - \lambda_r)^{\alpha_r}.$$

Zum Beweis von (i), (ii) führen wir einen Induktionsbeweis. Für $n = 1$ ist die Behauptung offensichtlich richtig; wir nehmen also an, daß (i), (ii) für Vektorräume der Dimension kleiner als n gelten. Mit den Bezeichnungen des vorhergehenden Lemma betrachten wir nun

$$g = f - \lambda_1 \cdot \mathrm{id} \in \mathrm{L}(V)$$

und

$$\begin{aligned} U &= \ker(g^d) = \ker(g^{\alpha_1}) = H_{\lambda_1}, \\ W &= \operatorname{im}(g^d). \end{aligned}$$

Nach dem Lemma gilt

$$V = H_{\lambda_1} \oplus W,$$

wobei sowohl H_{λ_1} als auch W g-invariante und damit auch f-invariante Unterräume von V sind; weiter gilt

$$\dim_K H_{\lambda_1} = \alpha_1.$$

Aufgrund der f-Invarianz von H_{λ_1} und W folgt sofort

$$p_{f|_W}(t) = \pm(t - \lambda_2)^{\alpha_2} \cdot \ldots \cdot (t - \lambda_r)^{\alpha_r}.$$

Damit können wir die Induktionsvoraussetzung auf $f|_W \in \mathrm{L}(W)$ anwenden. Es ergeben sich sofort (i), (ii).
Um (iii) zu beweisen, argumentieren wir in matrizieller Sprache. Aufgrund der Zerlegung

$$V = \bigoplus_{j=1}^{r} H_{\lambda_j}$$

mit $\dim_K H_{\lambda_j} = \alpha_j$ und wegen der f-Invarianz von H_{λ_j} existiert eine Basis $\mathcal{B}$ von V, so daß f bezüglich $\mathcal{B}$ durch die Matrix

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix} \in \mathrm{M}_n(K)$$

mit $A_j \in \mathrm{M}_{\alpha_j}(K) \quad (j = 1, ..., r)$ dargestellt wird. Nach dem vorhergehenden Lemma gilt dabei

$$(A_j - \lambda_j \cdot E_j)^{d_j} = 0 \in \mathrm{M}_{\alpha_j}(K)$$

für genügend großes d_j, d.h.

$$N_j = A_j - \lambda_j \cdot E_j$$

ist nilpotent $(j = 1, ..., r)$. Mit Hilfe eines geeigneten Basiswechsels in H_{λ_j} können wir ohne Einschränkung

$$N_j = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} \in \mathrm{M}_{\alpha_j}(K)$$

erreichen $(j = 1, ..., r)$. Dies zeigt die Existenz von $S \in \mathrm{GL}_n(K)$ mit

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 E_1 + N_1 & & \\ & \ddots & \\ & & \lambda_r E_r + N_r \end{pmatrix},$$

wobei

$$\lambda_j E_j + N_j = \begin{pmatrix} \lambda_j & & * \\ & \ddots & \\ 0 & & \lambda_j \end{pmatrix} \in \mathrm{M}_{\alpha_j}(K)$$

gilt. Daraus entnehmen wir insbesondere die Zerlegung

$$S^{-1} \cdot A \cdot S = D + N,$$

wobei D diagonal und N nilpotent ist. Direktes Nachrechnen zeigt nun $D \cdot N = N \cdot D$. Dies beweist (iii). □

Zur Formulierung der Jordanschen Normalform werden wir abschließend in H_{λ_j} eine geeignete Basis konstruieren, so daß die zugehörige Matrix N_j eine möglichst einfache Gestalt hat.

Definition. Zu $k \in \mathbb{N}$ definieren wir die *Jordanmatrix*

$$J_k = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in \mathrm{M}_k(K).$$

Es gilt $J_k^k = 0$.

Satz 2. *Es seien U ein K-Vektorraum, $g \in \mathrm{L}(U)$ nilpotent und $d := \min\{j | g^j = 0\}$. Dann existieren eindeutig bestimmte $s_1, ..., s_d \in \mathbb{N}$ mit*

$$d \cdot s_d + (d-1) \cdot s_{d-1} + ... + 1 \cdot s_1 = \dim_k U$$

und eine Basis $\mathcal{B}$ von U, so daß die Matrix B von g bezüglich $\mathcal{B}$ gegeben ist durch

$$A = \begin{pmatrix} J_d & & & & & & 0 \\ & \ddots & & & & & \\ & & J_d & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \begin{array}{l} \\ \}s_d - \text{mal} \\ \\ \vdots \\ \\ \}s_1 - \text{mal} \\ \\ \end{array}$$

Beweis. Mit $U_j := \ker(g^j)$ $(j = 0, 1, 2, ...)$ erhalten wir die echt aufsteigende Kette

$$\{0\} = U_0 \subsetneq U_1 \subsetneq ... \subsetneq U_{d-1} \subsetneq U_d = U.$$

Wir stellen nun fest:

(a)
$$\begin{aligned} g^{-1}(U_{j-1}) &= \{u \in U | g(u) \in U_{j-1}\} \\ &= \{u \in U | g^{j-1}(g(u)) = g^j(u) = 0\} \\ &= U_j. \end{aligned}$$

Dies zeigt insbesondere

$$g(U_j) = g(g^{-1}(U_{j-1})) \subseteq U_{j-1};$$

dies ist im allgemeinen eine echte Inklusion.

(b) Ist $W \subseteq U$ ein Unterraum mit $W \cap U_j = \{0\}$ für ein j, so ist $g|_W$ injektiv, denn wegen $\ker(g) = U_1 \subseteq U_j$ gilt $\ker(g|_W) = W \cap \ker(g) \subseteq W \cap U_j = \{0\}$.

Wir bilden nun die Zerlegung

$$V = U_d = U_{d-1} \oplus W_d \quad (\text{d.h.}\, U_{d-1} \cap W_d = \{0\}).$$

Nach (a) gilt

$$g(U_{d-1}) \subseteq U_{d-2},\ g(W_d) \subseteq U_{d-1};$$

weiter folgt mit (a)

$$U_{d-2} \cap g(W_d) = \{0\},$$

denn ist $v \in U_{d-2} \cap g(W_d)$, d.h. $v = g(w)$ mit $w \in W_d$, so folgt wegen $g(w) = v \in U_{d-2}$ die Beziehung $w \in g^{-1}(U_{d-2}) = U_{d-1}$, also $w \in U_{d-1} \cap W_d = \{0\}$.

Indem wir weiter zerlegen

$$U_{d-1} = U_{d-2} \oplus W_{d-1} \quad (\text{d.h. } U_{d-2} \cap W_{d-1} = \{0\}),$$

folgt nach dem Vorhergehenden $g(W_d) \subseteq W_{d-1}$. Iterativ erhalten wir folgende Zerlegungen von V:

$$\begin{array}{ccccccccccc}
U_d & & & & & & & & & & \\
\downarrow & & & & & & & & & & \\
U_{d-1} & \oplus & W_d & & & & & & & & \\
\downarrow & & \downarrow & & & & & & & & \\
U_{d-2} & \oplus & W_{d-1} & \oplus & W_d & & & & & & \\
\downarrow & & \downarrow & & \downarrow & & & & & & \\
\vdots & & \vdots & & \vdots & & & & & & \\
\downarrow & & \downarrow & & \downarrow & & & & & & \\
U_1 & \oplus & W_2 & \oplus & W_3 & \oplus & \ldots & \oplus & W_d & & \\
\downarrow & & \downarrow & & \downarrow & & & & \downarrow & & \\
U_0 & \oplus & W_1 & \oplus & W_2 & \oplus & \ldots & \oplus & W_{d-1} & \oplus & W_d\,,
\end{array}$$

wobei die Pfeile die Wirkung von g zeigen. Also besteht die Zerlegung

$$U = W_1 \oplus \ldots \oplus W_d,$$

wobei die Einschränkungen von g auf

$$W_d \longrightarrow W_{d-1} \longrightarrow \ldots \longrightarrow W_1$$

nach (b) alle injektiv sind. Beginnend mit W_d bauen wir nun wie folgt eine Basis U auf:

$$\{w_1^{(d)}, \ldots, w_{s_d}^{(d)}\}$$

$$\{g(w_1^{(d)}), \ldots, g(w_{s_d}^{(d)}); w_1^{(d-1)}, \ldots, w_{s_{d-1}}^{(d-1)}\}$$

$$\vdots$$

$$\{g^{d-1}(w_1^{(d)}), \ldots, g^{d-1}(w_{s_d}^{(d)}); g^{d-2}(w_1^{(d-1)}), \ldots, g^{d-2}(w_{s_{d-1}}^{(d-1)}); \ldots; w_1^{(1)}, \ldots, w_{s_1}^{(1)}\};$$

hierbei steht in der ersten Zeile eine Basis von W_d, in der zweiten Zeile eine Basis von $W_{d-1}, ...,$ in der letzten Zeile eine Basis von W_1. Die geordnete Basis $\mathcal{B}$ von U erhalten wir nun, indem wir im vorhergehenden Schema in jeder Spalte von unten nach oben laufen und die Spalten von links nach rechts lesen. Die Matrix B von g bezüglich $\mathcal{B}$ hat nun ersichtlich die behauptete Form. $\square$

Jordansche Normalform. *Es sei $f \in \mathrm{L}(V)$, und das charakteristische Polynom von f zerfalle in Linearfaktoren*

$$p_f(t) = \pm(t-\lambda_1)^{\alpha_1} \cdot ... \cdot (t-\lambda_r)^{\alpha_r}.$$

Dann existiert eine Basis $\mathcal{B}$ von V, so daß die Matrix $A \in \mathrm{M}_n(K)$ von f bezüglich $\mathcal{B}$ die folgende Form hat:

$$A = \begin{pmatrix} \lambda_1 E_1 + N_1 & & \\ & \ddots & \\ & & \lambda_r E_r + N_r \end{pmatrix}$$

mit $\lambda_j E_j + N_j \in \mathrm{M}_{\alpha_j}(K)$ und

$$N_j = \begin{pmatrix} J_{d_j} & & & & & & 0 \\ & \ddots & & & & & \\ & & J_{d_j} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \begin{matrix} \\ \}s_{d_j}^{(j)} - \mathrm{mal} \\ \\ \vdots \\ \\ \}s_1^{(j)} - \mathrm{mal}\ , \\ \\ \end{matrix}$$

wobei $d_j \cdot s_{d_j}^{(j)} + (d_j - 1) \cdot s_{d_j-1}^{(j)} + ... + 1 \cdot s_1^{(j)} = \alpha_j$ nach Satz 2 bestimmt werden. Es ist d_j die Vielfachheit von λ_j im Minimalpolynom m_f von f. $\square$

Beispiel. Es sei

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \in \mathrm{M}_3(\mathbb{R}).$$

Man berechnet sofort

$$p_A(t) = -(t-2)^3.$$

Mit $B := A - 2 \cdot E \in \mathrm{M}_3(\mathbb{R})$ berechnet man sofort

$$B \neq 0, B^2 \neq 0, B^3 = 0,$$

d.h.

$$m_A(t) = (t-2)^3.$$

Mit den Bezeichnungen des vorhergehenden Satzes haben wir

$$\begin{aligned} U_0 &= \{0\}, \\ U_1 &= \ker(B) = \left\langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle, \\ U_2 &= \ker(B^2) = \left\langle \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\rangle. \end{aligned}$$

Wir haben die Zerlegungen

$$\begin{aligned} \mathbb{R}^3 &= U_2 \oplus W_3 \\ &= U_1 \oplus W_2 \oplus W_3 \\ &= U_0 \oplus W_1 \oplus W_2 \oplus W_3. \end{aligned}$$

Damit wählen wir

$$\begin{aligned} W_3 &= \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle, \\ W_2 &= B(W_3) = \left\langle \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \right\rangle, \\ W_1 &= B(W_2) = \left\langle \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix} \right\rangle. \end{aligned}$$

Damit setzen wir

$$S = \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix}, \quad S^{-1} = \frac{1}{4} \begin{pmatrix} -1 & -3 & 0 \\ 2 & 2 & 0 \\ 0 & 4 & 4 \end{pmatrix}$$

und erhalten

$$A' = S^{-1} \cdot A \cdot S = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

die Jordansche Normalform von A.

Kapitel 5

Euklidische und unitäre Vektorräume

5.1 Skalarprodukte

In diesem Kapitel werden wir zwischen den Grundkörpern $K = \mathbf{R}$ und $K = \mathbf{C}$ zu unterscheiden haben.

5.1.1 Der Fall $K = \mathbf{R}$

Es sei V ein $\mathbf{R}$-Vektorraum; unter dem kartesischen Produkt $V \times V$ ist die Menge $\{(x,y)|x \in V, y \in V\}$ zu verstehen.

Definition. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbf{R}$, d.h. $(x,y) \mapsto \langle x,y \rangle$ $(x, y \in V)$ heißt ein *Skalarprodukt auf* V, falls die drei folgenden Eigenschaften erfüllt sind:

(i) $\langle \cdot, \cdot \rangle$ ist *bilinear*, also linear in beiden Argumenten, d.h.

$$\begin{aligned} \langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle &= \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle, \\ \langle x, \mu_1 y_1 + \mu_2 y_2 \rangle &= \mu_1 \langle x, y_1 \rangle + \mu_2 \langle x, y_2 \rangle, \end{aligned}$$

für alle $x, x_1, x_2, y, y_1, y_2 \in V$ und alle $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbf{R}$.

(ii) $\langle \cdot, \cdot \rangle$ ist *symmetrisch*, d.h.

$$\langle x, y \rangle = \langle y, x \rangle$$

für alle $x, y \in V$.

(iii) $\langle \cdot, \cdot \rangle$ ist *positiv definit*, d.h.

$$\langle x, x \rangle > 0$$

für alle $x \in V, x \neq 0$.

Bemerkung. Es können auch Skalarprodukte $\langle \cdot, \cdot \rangle$ für Vektorräume V über beliebigen Körpern definiert werden. Dann wird allerdings nur die Bilinearität und die Symmetrie gefordert, weil die positive Definitheit im allgemeinen keinen Sinn macht. Wir wollen uns hier aber nicht mit dieser allgemeineren Situation beschäftigen.

Definition. Ein $\mathbf{R}$-Vektorraum V mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$ heißt ein *euklidischer Vektorraum.*

Beispiel. Es sei $V = \mathbf{R}^n$. Dann wird durch

$$\langle x, y \rangle := \sum_{j=1}^{n} \xi_j \cdot \eta_j \text{ mit } x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

ein Skalarprodukt definiert (Übungsaufgabe). Wir nennen es das *Standardskalarprodukt des* $\mathbf{R}^n$.

5.1.2 Der Fall $K = \mathbf{C}$

Es sei V ein $\mathbf{C}$-Vektorraum; wiederum sei unter $V \times V$ das kartesische Produkt $\{(x, y) | x \in V, y \in V\}$ verstanden.

Definition. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbf{C}$, d.h. $(x, y) \mapsto \langle x, y \rangle$ $(x, y \in V)$ heißt *Skalarprodukt auf* V, falls die drei folgenden Eigenschaften erfüllt sind:

(i) $\langle \cdot, \cdot \rangle$ ist *linear* im ersten und *antilinear* im zweiten Argument, d.h.

$$\begin{aligned} \langle \lambda_1 x_1 + \lambda_2 x_2, y \rangle &= \lambda_1 \langle x_1, y \rangle + \lambda_2 \langle x_2, y \rangle, \\ \langle x, \mu_1 y_1 + \mu_2 y_2 \rangle &= \overline{\mu_1} \langle x, y_1 \rangle + \overline{\mu_2} \langle x, y_2 \rangle, \end{aligned}$$

für alle $x, x_1, x_2, y, y_1, y_2 \in V$ und alle $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbf{C}$.

(ii) $\langle \cdot, \cdot \rangle$ ist *hermitesch*, d.h.

$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$

für alle $x, y \in V$.

(iii) $\langle \cdot, \cdot \rangle$ ist *positiv definit*, d.h.

$$\langle x, x \rangle > 0$$

für alle $x \in V, x \neq 0$.

Bemerkung. Man zeigt leicht, daß die Linearität von $\langle \cdot, \cdot \rangle$ im ersten Argument zusammen mit der Hermitizität die Antilinearität im zweiten Argument impliziert; dementsprechend müßte in (i) nur die Linearität von $\langle \cdot, \cdot \rangle$ gefordert werden.
Aufgrund der Hermitizität folgt für $x \in V$ die Gleichung $\langle x, x \rangle = \overline{\langle x, x \rangle}$, d.h. $\langle x, x \rangle \in \mathbf{R}$; dementsprechend ist die Eigenschaft (iii) überhaupt erst sinnvoll.

Definition. Ein $\mathbf{C}$-Vektorraum V mit einem Skalarprodukt $\langle\cdot,\cdot\rangle$ heißt ein *unitärer Vektorraum.*

Beispiel. Es sei $V = \mathbf{C}^n$. Dann wird durch

$$\langle x, y\rangle := \sum_{j=1}^{n} \xi_j \cdot \overline{\eta_j} \text{ mit } x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$$

ein Skalarprodukt definiert. Wir nennen es das *Standardskalarprodukt des* $\mathbf{C}^n$.

Lemma (Cauchy-Schwarzsche Ungleichung). *Es sei V ein euklidischer bzw. unitärer Vektorraum. Dann besteht für zwei Vektoren $x, y \in V$ die Ungleichung*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle,$$

wobei das Gleichheitszeichen genau dann gilt, wenn x und y linear abhängig sind.

Beweis. Ist $y = 0$, so ist die Behauptung sofort wegen $\langle x, y\rangle = 0$ und $\langle y, y\rangle = 0$ klar. Daher können wir im folgenden $y \neq 0$ und somit auch $\langle y, y\rangle > 0$ annehmen. Für einen beliebigen Skalar $\lambda \in \mathbf{R}$ bzw. $\lambda \in \mathbf{C}$, folgt dann aufgrund der Definition des Skalarprodukts

$$\begin{aligned} 0 &\leq \langle x - \lambda y, x - \lambda y\rangle \\ &= \langle x, x\rangle - \lambda\langle y, x\rangle - \bar{\lambda}\langle x, y\rangle + \lambda\bar{\lambda}\langle y, y\rangle \\ &= \langle x, x\rangle - \lambda\overline{\langle x, y\rangle} - \bar{\lambda}\langle x, y\rangle + \lambda\bar{\lambda}\langle y, y\rangle. \end{aligned}$$

Wählt man nun speziell $\lambda = \langle x, y\rangle / \langle y, y\rangle$, so ergibt sich die Ungleichung

$$\begin{aligned} \langle x, x\rangle - \frac{\langle x, y\rangle^2}{\langle y, y\rangle} - \frac{\langle x, y\rangle^2}{\langle y, y\rangle} + \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle} &\geq 0 \\ \Longleftrightarrow \qquad \langle x, x\rangle\langle y, y\rangle - |\langle x, y\rangle|^2 &\geq 0 \\ \Longleftrightarrow \qquad |\langle x, y\rangle|^2 &\leq \langle x, x\rangle\langle y, y\rangle. \end{aligned}$$

Dies liefert den ersten Teil der Behauptung.
Zum Beweis des zweiten Teils beachtet man, daß in den obigen Ungleichungen genau dann das Gleicheitszeichen steht, wenn

$$\langle x - \lambda y, x - \lambda y\rangle = 0$$

gilt; dies ist aber gleichbedeutend mit

$$x - \lambda y = 0 \Longleftrightarrow x = \lambda y,$$

d.h. x und y sind linear abhängig. $\square$

Definition. Seien V ein euklidischer bzw. unitärer Vektorraum und $x \in V$. Die nichtnegative Zahl

$$\|x\| := \sqrt{\langle x, x\rangle}$$

heißt *die Länge* oder *die Norm* oder *der Betrag des Vektors x.*

Lemma. *Es sei V ein euklidischer bzw. unitärer Vektorraum. Dann besitzt die Norm folgende drei Eigenschaften:*

(i) $\|\cdot\|$ *ist positiv definit, d.h.*

$$\|x\| \geq 0 \; \forall x \in V \;, \; \|x\| = 0 \Longleftrightarrow x = 0.$$

(ii) $\|\cdot\|$ *ist homogen, d.h.*

$$\|\lambda x\| = |\lambda| \cdot \|x\| \; \forall x \in V \;, \; \forall \lambda \in \mathbf{R}, \text{bzw. } \forall \lambda \in \mathbf{C}.$$

(iii) $\|\cdot\|$ *erfüllt die Dreiecksungleichung, d.h.*

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V.$$

Beweis. Die Eigenschaften (i) und (ii) folgen unmittelbar aus der Definition des Skalarprodukts. Eigenschaft (iii) folgt mit Hilfe der Cauchy-Schwarzschen Ungleichung, nämlich

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \overline{\langle x, y\rangle} + \langle y, y\rangle \\
&= \|x\|^2 + 2\mathrm{Re}(\langle x, y\rangle) + \|y\|^2 \\
&\leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2.
\end{aligned}$$

Bemerkung. Seien V ein euklidischer bzw. unitärer Vektorraum und $x, y \in V$. *Der Abstand* oder *die Distanz $d(x, y)$ von x nach y* ist dann definiert durch

$$d(x, y) := \|x - y\|.$$

Der Abstand bzw. die Distanz erfüllt ebenfalls die Eigenschaften der positiven Definitheit und der Homogenität sowie die Dreiecksungleichung.

Definition. Es sei V ein euklidischer bzw. unitärer Vektorraum. Ein Vektor $x \in V$ heißt *normiert*, wenn $\|x\| = 1$ gilt.

Bemerkung. (Darstellung von Skalarprodukten durch Matrizen)
Es sei V ein n-dimensionaler euklidischer Vektorraum mit Basis $\mathcal{B} = \{b_1, \ldots, b_n\}$. Sind $x, y \in V$ und $x = \sum_{j=1}^n \xi_j b_j,\ y = \sum_{k=1}^n \eta_k b_k$, so ergibt sich

$$\langle x, y \rangle = \left\langle \sum_{j=1}^n \xi_j b_j, \sum_{k=1}^n \eta_k b_k \right\rangle = \sum_{j,k=1}^n \xi_j \eta_k \langle b_j, b_k \rangle = (\xi_1, \ldots, \xi_n) A \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix},$$

wobei die Matrix $A \in \mathrm{M}_n(\mathbf{R})$ gegeben ist durch die Skalarprodukte

$$A = (\langle b_j, b_k \rangle)_{1 \le j,k \le n}.$$

Die Matrix A heißt *Gramsche Matrix*; sie ist offensichtlich *symmetrisch*.

Entsprechend folgt für einen unitären Vektorraum V mit Basis $\mathcal{B} = \{b_1, \ldots, b_n\}$

$$\langle x, y \rangle = (\xi_1, \ldots, \xi_n) A \begin{pmatrix} \bar{\eta}_1 \\ \vdots \\ \bar{\eta}_n \end{pmatrix},$$

wobei die Matrix $A \in \mathrm{M}_n(\mathbf{R})$ wiederum gegeben ist durch die Skalarprodukte

$$A = (\langle b_j, b_k \rangle)_{1 \le j,k \le n}.$$

Die Matrix A wird wieder *Gramsche Matrix* genannt. Im Gegensatz zum euklidischen Fall gilt aber

$$(\langle b_j, b_k \rangle)_{1 \le j,k \le n} = (\overline{\langle b_k, b_j \rangle})_{1 \le j,k \le n} \Longleftrightarrow A = {}^t\bar{A};$$

A heißt nun *hermitesch*.

Bemerkung. (Verhalten der Gramschen Matrix unter Basistransformation)
Es sei V ein n-dimensionaler euklidischer Vektorraum mit Basen $\mathcal{B} = \{b_1, \ldots, b_n\}, \mathcal{B}' = \{b'_1, \ldots, b'_n\}$ und Gramschen Matrizen $A = (\langle b_j, b_k \rangle)_{1 \le j,k \le n},\ A' = (\langle b'_l, b'_m \rangle)_{1 \le l,m \le n}$. Mit der Basistransformationsmatrix $S = (\sigma_{j,l})_{1 \le j,l \le n}$ von $\mathcal{B}$ nach $\mathcal{B}'$, d.h.

$$b'_l = \sum_{j=1}^n \sigma_{j,l} b_j \quad (l = 1, \ldots, n)$$

folgt dann

$$\begin{aligned} \langle b'_l, b'_m \rangle &= \left\langle \sum_{j=1}^n \sigma_{j,l} b_j, \sum_{k=1}^n \sigma_{k,m} b_k \right\rangle \\ &= \sum_{j,k=1}^n \sigma_{j,l} \langle b_j, b_k \rangle \sigma_{k,m} \\ &= \sum_{j,k=1}^n {}^t\sigma_{l,j} \langle b_j, b_k \rangle \sigma_{k,m}. \end{aligned}$$

Somit erhalten wir die Transformationsformel

$$A' = {}^tS \cdot A \cdot S.$$

Analog finden wir für einen n-dimensionalen unitären Vektorraum V mit Basen $\mathcal{B} = \{b_1, \dots, b_n\}, \mathcal{B}' = \{b'_1, \dots, b'_n\}$, Gramschen Matrizen A, A' und Basistransformationsmatrix S von $\mathcal{B}$ nach $\mathcal{B}'$ die Transformationsformel

$$A' = {}^tS \cdot A \cdot \bar{S}.$$

5.2 Winkelmessung und Orthogonalität

Es sei zunächst V ein euklidischer Vektorraum. Sind $x, y \in V$, $x, y \neq 0$, so gilt aufgrund der Cauchy-Schwarzschen Ungleichung

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \Longleftrightarrow -1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq +1.$$

Damit wird die folgende Definition sinnvoll.

Definition. Es sei V ein euklidischer Vektorraum. Der *Zwischenwinkel* γ*,* $0 \leq \gamma \leq \pi$*, zwischen* $x, y \in V$, $x, y \neq 0$, ist gegeben durch

$$\cos(\gamma) := \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

Dies motiviert die folgende, allgemeinere Definition.

Definition. Es sei V ein euklidischer bzw. unitärer Vektorraum. Dann heißen zwei Vektoren $x, y \in V$ *senkrecht* oder *orthogonal zueinander*, in Zeichen $x \perp y$, falls $\langle x, y \rangle = 0$ gilt.

Definition. Es sei V ein euklidischer bzw. unitärer Vektorraum. Eine nichtleere Menge $\{x_1, \dots, x_m\}$ von Vektoren $0 \neq x_j \in V$ $(j = 1, \dots, m)$, die paarweise zueinander orthogonal sind, heißt *ein Orthogonalsystem*; sind die Vektoren überdies normiert, so spricht man von einem *Orthonormalsystem.*
Bildet die Basis $\mathcal{B}$ von V ein Orthogonalsystem bzw. ein Orthonormalsystem so spricht man von einer *Orthogonalbasis* bzw. von einer *Orthonormalbasis.*

Lemma. *Seien* V *ein euklidischer bzw. unitärer Vektorraum und* $\{x_1, \dots, x_m\} \subset V$ *ein Orthogonalsystem. Dann sind die Vektoren* $x_1, \dots, x_m$ *linear unabhängig.*

Beweis. Es sei

$$\sum_{j=1}^{m} \lambda_j x_j = 0.$$

Skalarmultiplikation mit dem Vektor x_k $(k = 1, \ldots, m)$, liefert dann

$$0 = \left\langle \sum_{j=1}^{m} \lambda_j x_j, x_k \right\rangle = \sum_{j=1}^{m} \lambda_j \langle x_j, x_k \rangle = \lambda_k \langle x_k, x_k \rangle.$$

Wegen $x_k \neq 0$ ist $\langle x_k, x_k \rangle \neq 0$, also folgt $\lambda_k = 0$ für $k = 1, \ldots, m$. $\square$

Satz (Gram-Schmidtsches Orthogonalisierungsverfahren). *Es sei V ein n-dimensionaler euklidischer bzw. unitärer Vektorraum. Dann besitzt V eine Orthonormalbasis.*

Beweis. Es sei $\mathcal{B} = \{b_1, \ldots, b_n\}$ eine beliebige Basis von V. Zuerst setzen wir

$$e_1 := \frac{1}{\|b_1\|} b_1.$$

Damit gilt $\|e_1\| = 1$, d.h. e_1 ist ein normierter Vektor. Mit einem unbestimmten Skalar λ setzen wir weiter

$$e_2' := b_2 - \lambda e_1.$$

Wir bestimmen nun λ derart, daß $\langle e_2', e_1 \rangle = 0$ ist, d.h. mit Hilfe der Gleichung

$$0 = \langle e_2', e_1 \rangle = \langle b_2 - \lambda e_1, e_1 \rangle = \langle b_2, e_1 \rangle - \lambda \langle e_1, e_1 \rangle = \langle b_2, e_1 \rangle - \lambda,$$

also

$$\lambda = \langle b_2, e_1 \rangle.$$

Damit erhalten wir

$$e_2' = b_2 - \langle b_2, e_1 \rangle e_1;$$

da b_1, b_2 linear unabhängig voneinander sind, ist $e_2' \neq 0$. Indem wir e_2' normieren, d.h.

$$e_2 := \frac{1}{\|e_2'\|} e_2'$$

setzen, erhalten wir einen zu e_1 orthogonalen, normierten Vektor e_2.

Induktiv konstruieren wir auf diese Weise ein Orthonormalsystem $\{e_1, \ldots, e_m\}$, wobei $1 \leq m < n$ gilt. Den nächsten Basisvektor e_{m+1} konstruieren wir wie zuvor durch den Ansatz

$$e_{m+1}' := b_{m+1} - \lambda_1 e_1 - \ldots - \lambda_m e_m;$$

dabei werden die Skalare $\lambda_1, \ldots, \lambda_m$ derart bestimmt, daß e_{m+1}' auf den Vektoren des Orthonormalsystems $\{e_1, \ldots, e_m\}$ senkrecht steht. Dies führt zu

$$\begin{aligned} \langle e_{m+1}', e_1 \rangle = 0 &\implies \lambda_1 = \langle b_{m+1}, e_1 \rangle, \\ \ldots \quad & \qquad \ldots \\ \langle e_{m+1}', e_m \rangle = 0 &\implies \lambda_m = \langle b_{m+1}, e_m \rangle, \end{aligned}$$

also

$$e_{m+1}' = b_{m+1} - \langle b_{m+1}, e_1 \rangle e_1 - \ldots - \langle b_{m+1}, e_m \rangle e_m.$$

Aufgrund der linearen Unabhängigkeit von $b_1, \ldots, b_{m+1}$ ist $e'_{m+1} \neq 0$. Indem wir schließlich e'_{m+1} normieren, d.h.

$$e_{m+1} := \frac{1}{\|e'_{m+1}\|} e'_{m+1}$$

setzen, erhalten wir einen zu $e_1, \ldots, e_m$ orthogonalen, normierten Vektor e_{m+1}.
Nach n Schritten erhält man endlich die gesuchte Orthonormalbasis $\{e_1, \ldots, e_n\}$. □

5.3 Orthogonale und unitäre Abbildungen

Im folgenden sei V ein euklidischer bzw. unitärer Vektorraum.

Definition. Eine lineare Abbildung $f \in \mathrm{L}(V)$ heißt *orthogonal* bzw. *unitär*, falls für alle $x, y \in V$ die Gleichheit

$$\langle f(x), f(y) \rangle = \langle x, y \rangle$$

gilt.

Definition.

(i) Für einen euklidischen Vektorraum V setzen wir

$$\mathrm{O}(V) = \{f \in \mathrm{L}(V) \mid f \text{ orthogonal}\}.$$

(ii) Für einen unitären Vektorraum V setzen wir

$$\mathrm{U}(V) = \{f \in \mathrm{L}(V) \mid f \text{ unitär}\}.$$

Bemerkung. Ist $f \in \mathrm{O}(V)$ bzw. $f \in \mathrm{U}(V)$, so ist f offensichtlich eine längen- und winkeltreue Abbildung. Umgekehrt definiert eine längen- und winkeltreue lineare Abbildung eine orthogonale bzw. unitäre Abbildung. Aufgrund der Relation

$$\langle x, y \rangle = \frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2)$$

genügt es im letzteren Fall, nur die Längentreue zu fordern.

Orthogonale bzw. unitäre Abbildungen sind wegen

$$f(x) = 0 \Rightarrow \|f(x)\| = 0 \Rightarrow \|x\| = 0 \Rightarrow x = 0$$

immer injektiv.

Lemma.

(i) *Ist V ein endlich dimensionaler euklidischer Vektorraum, so gilt* $\mathrm{O}(V) \subseteq \mathrm{GL}(V)$, *d.h. jede orthogonale Abbildung ist insbesondere bijektiv. Die Menge* $\mathrm{O}(V)$ *ist eine Untergruppe von* $\mathrm{GL}(V)$.

(ii) *Ist V ein endlich dimensionaler unitärer Vektorraum, so gilt* $\mathrm{U}(V) \subseteq \mathrm{GL}(V)$, *d.h. jede unitäre Abbildung ist insbesondere bijektiv. Die Menge* $\mathrm{U}(V)$ *ist eine Untergruppe von* $\mathrm{GL}(V)$.

Beweis.

(i) Da $f \in \mathrm{O}(V)$ injektiv ist, folgt aufgrund der endlichen Dimension von V auch die Surjektivität von f, also $f \in \mathrm{GL}(V)$.
Offensichtlich ist $\mathrm{id} \in \mathrm{O}(V)$, und mit $f, g \in \mathrm{O}(V)$ ist auch $f \circ g \in \mathrm{O}(V)$. Für die Inverse f^{-1} von $f \in \mathrm{O}(V)$ berechnen wir schließlich

$$\langle x, y\rangle = \langle f(f^{-1}(x)), f(f^{-1}(y))\rangle = \langle f^{-1}(x), f^{-1}(y)\rangle$$

für alle $x, y \in V$, d.h. $f^{-1} \in \mathrm{O}(V)$. Damit ist $\mathrm{O}(V)$ als Untergruppe von $\mathrm{GL}(V)$ nachgewiesen.

(ii) Der Beweis im unitären Fall verläuft analog zu (i). □

Definition.

(i) Ist V ein endlich dimensionaler euklidischer Vektorraum, so nennen wir $\mathrm{O}(V)$ die *orthogonale Gruppe von V*.

(ii) Ist V ein endlich dimensionaler unitärer Vektorraum, so nennen wir $\mathrm{U}(V)$ die *unitäre Gruppe von V*.

Definition.

(i) Eine Matrix $S \in \mathrm{M}_n(\mathbf{R})$ heißt *orthogonal*, falls $S \cdot {}^tS = E$ gilt; äquivalent dazu ist die Bedingung $S^{-1} = {}^tS$. Wir setzen

$$\mathrm{O}_n(\mathbf{R}) = \{S \in \mathrm{M}_n(\mathbf{R}) \mid S \text{ orthogonal}\}.$$

(ii) Eine Matrix $S \in \mathrm{M}_n(\mathbf{C})$ heißt *unitär*, falls $S \cdot {}^t\bar{S} = E$ gilt; äquivalent dazu ist die Bedingung $S^{-1} = {}^t\bar{S}$. Wir setzen

$$\mathrm{U}_n(\mathbf{C}) = \{S \in \mathrm{M}_n(\mathbf{C}) \mid S \text{ unitär}\}.$$

Lemma.

(i) *Die Menge $\mathrm{O}_n(\mathbf{R})$ ist eine Untergruppe von $\mathrm{GL}_n(\mathbf{R})$. Eine Matrix $S \in \mathrm{M}_n(\mathbf{R})$ ist genau dann orthogonal, wenn ihre Zeilen- oder Spaltenvektoren eine Orthonormalbasis des $\mathbf{R}^n$ bezüglich des Standardskalarprodukts bilden.*

(ii) *Die Menge $\mathrm{U}_n(\mathbf{C})$ ist eine Untergruppe von $\mathrm{GL}_n(\mathbf{C})$. Eine Matrix $S \in \mathrm{M}_n(\mathbf{C})$ ist genau dann unitär, wenn ihre Zeilen oder Spaltenvektoren eine Orthonormalbasis des $\mathbf{C}^n$ bezüglich des Standardskalarprodukts bilden.*

Beweis.

(i) Es ist $E \in \mathrm{O}_n(\mathbf{R})$ und mit $S, T \in \mathrm{O}_n(\mathbf{R})$ folgt

$$(ST) \cdot {}^t(ST) = S \cdot T \cdot {}^tT \cdot S = S \cdot E \cdot {}^tS = S \cdot {}^tS = E,$$

also $S \cdot T \in \mathrm{O}_n(\mathbf{R})$. Schließlich folgt für $S \in \mathrm{O}_n(\mathbf{R})$

$$S^{-1} \cdot {}^t(S^{-1}) = {}^tS \cdot S = S \cdot {}^tS = E,$$

d.h. $S^{-1} \in \mathrm{O}_n(\mathbf{R})$. Damit ist $\mathrm{O}_n(\mathbf{R})$ eine Untergruppe von $\mathrm{GL}_n(\mathbf{R})$. Der zweite Teil der Behauptung folgt unmittelbar aus der Orthogonalitätsdefinition einer Matrix.

(ii) Der Beweis im unitären Fall verläuft analog zu (i). □

Definition.

(i) $\mathrm{O_n}(\mathbf{R})$ heißt die *orthogonale Gruppe* (*auf* $\mathbf{R}^n$).

(ii) $\mathrm{U}_n(\mathbf{C})$ heißt die *unitäre Gruppe* (*auf* $\mathbf{C}^n$).

Proposition.

(i) *Es seien V ein n-dimensionaler euklidischer Vektorraum mit Orthonormalbasis $\mathcal{B} = \{b_1, ..., b_n\}$ und $f \in \mathrm{O}(V)$ mit Matrix $S \in (\sigma_{l,j})_{1 \le l,j \le n} \in \mathrm{M}_n(\mathbf{R})$ bezüglich $\mathcal{B}$. Dann gilt $S \in \mathrm{O}_n(\mathbf{R})$. Desweiteren induziert die Zuordnung $f \longmapsto S$ einen Isomorphismus*

$$\mathrm{O}(V) \xrightarrow{\cong} \mathrm{O}_n(\mathbf{R}).$$

(ii) *Es seien V ein n-dimensionaler unitärer Vektorraum mit Orthonormalbasis $\mathcal{B} = \{b_1, ..., b_n\}$ und $f \in \mathrm{U}(V)$ mit Matrix $S = (\sigma_{l,j})_{1 \le l,j \le n} \in \mathrm{M}_n(\mathbf{C})$ bezüglich $\mathcal{B}$. Dann gilt $S \in \mathrm{U}_n(\mathbf{C})$. Desweiteren induziert die Zuordnung $f \longmapsto S$ einen Isomorphismus*

$$\mathrm{U}(V) \xrightarrow{\cong} \mathrm{U}_n(\mathbf{C}).$$

Beweis.

(i) Mit dem Kronecker-Symbol $\delta_{j,k}$ erhalten wir für $j, k = 1, ..., n$

$$\delta_{j,k} = \langle b_j, b_k \rangle = \langle f(b_j), f(b_k) \rangle = \left\langle \sum_{l=1}^{n} \sigma_{l,j} \cdot b_l, \sum_{m=1}^{n} \sigma_{m,k} \cdot b_m \right\rangle = \sum_{l=1}^{n} \sigma_{l,j} \cdot \sigma_{l,k} = \sum_{l=1}^{n} {}^t\sigma_{j,l} \cdot \sigma_{l,k},$$

also

$${}^tS \cdot S = E \quad \Longleftrightarrow \quad S \cdot {}^tS = E.$$

Somit folgt $S \in \mathrm{O}(\mathbf{R})$.

Die durch die Zuordnung $f \longmapsto S$ definierte Abbildung von $\mathrm{O}(V)$ nach $\mathrm{O}_n(\mathbf{R})$ ist nach dem vorhergehenden wohldefiniert und offensichtlich injektiv. Durch die Vorgabe von $S = (\sigma_{l,j})_{1 \leq l,j \leq n} \in \mathrm{O}_n(\mathbf{R})$ wird eine lineare Abbildung f mit

$$f(b_j) = \sum_{l=1}^{n} \sigma_{l,j} \cdot b_l$$

definiert. Die vorhergehende Rechnung zeigt

$$\langle f(b_j), f(b_k)\rangle = \sigma_{j,k} = \langle b_j, b_k\rangle,$$

woraus $f \in \mathrm{O}(V)$ folgt. Dies beweist

$$\mathrm{O}(V) \xrightarrow{\cong} \mathrm{O}_n(\mathbf{R}).$$

(ii) Der Beweis im unitären Fall verläuft analog zu (i). □

5.4 Selbstadjungierte Abbildungen

Im folgenden sei V ein euklidischer bzw. unitärer Vektorraum.

Definition. Eine lineare Abbildung $f \in \mathrm{L}(V)$ heißt *selbstadjungiert*, falls für alle $x, y \in V$ die Gleichheit

$$\langle f(x), y\rangle = \langle x, f(y)\rangle$$

gilt.

Definition. Für einen euklidischen bzw. unitären Vektorraum V setzen wir

$$\mathcal{S}(V) := \{f \in \mathrm{L}(V) \mid f \text{ selbstadjungiert}\}.$$

Definition.

(i) Eine Matrix $A \in \mathrm{M}_n(\mathbf{R})$ heißt *symmetrisch*, falls $A = {}^t A$ gilt. Wir setzen

$$\mathrm{Sym}_n(\mathbf{R}) = \{A \in \mathrm{M}_n(\mathbf{R}) \mid A \text{ symmetrisch}\}.$$

(ii) Eine Matrix $A \in \mathrm{M}_n(\mathbf{C})$ heißt *hermitesch*, falls $A = {}^t \bar{A}$ gilt. Wir setzen

$$\mathrm{Herm}_n(\mathbf{C}) = \{A \in \mathrm{M}_n(\mathbf{C}) \mid A \text{ hermitesch}\}.$$

Proposition.

(i) *Es seien V ein n-dimensionaler euklidischer Vektorraum mit Orthonormalbasis $\mathcal{B} = \{b_1, ..., b_n\}$ und $f \in \mathcal{S}(V)$ mit Matrix $A = (\alpha_{k,j})_{1 \le j,k \le n} \in \mathrm{M}_n(\mathbf{R})$ bezüglich $\mathcal{B}$. Dann ist $A \in \mathrm{Sym}_n(\mathbf{R})$. Desweiteren induziert die Zuordnung $f \longmapsto A$ eine Bijektion*

$$\mathcal{S}(V) \xrightarrow{\text{bijektiv}} \mathrm{Sym}_n(\mathbf{R}).$$

(ii) *Es seien V ein n-dimensionaler unitärer Vektorraum mit Orthonormalbasis $\mathcal{B} = \{b_1, ..., b_n\}$ und $f \in \mathcal{S}(V)$ mit Matrix $A = (\alpha_{k,j})_{1 \le j,k \le n} \in \mathrm{M}_n(\mathbf{C})$ bezüglich $\mathcal{B}$. Dann ist $A \in \mathrm{Herm}_n(\mathbf{C})$. Desweiteren induziert die Zuordnung $f \longmapsto A$ eine Bijektion*

$$\mathcal{S}(V) \xrightarrow{\text{bijektiv}} \mathrm{Herm}_n(\mathbf{C}).$$

Beweis.

(i) Mit

$$f(b_j) = \sum_{k=1}^{n} \alpha_{k,j} b_k \quad (j = 1, ..., n)$$

folgt einerseits

$$\langle f(b_j), b_l \rangle = \left\langle \sum_{k=1}^{n} \alpha_{k,j} b_k, b_l \right\rangle = \sum_{k=1}^{n} \alpha_{k,j} \langle b_k, b_l \rangle = \sum_{k=1}^{n} \alpha_{k,j} \delta_{k,l} = \alpha_{l,j}.$$

Andererseits berechnen wir unter Benutzung der Selbstadjungiertheit von f

$$\langle f(b_j), b_l \rangle = \langle b_j, f(b_l) \rangle = \langle b_j, \sum_{k=1}^{n} \alpha_{k,l} b_k \rangle = \sum_{k=1}^{n} \alpha_{k,l} \langle b_j, b_k \rangle = \sum_{k=1}^{n} \alpha_{k,l} \cdot \delta_{j,k} = \alpha_{j,l}.$$

Ein Vergleich liefert

$$\alpha_{l,j} = \alpha_{j,l}$$

für $j, l = 1, \ldots, n$, also

$$A = {}^{t}A.$$

Damit erhalten wir durch die Zuordnung $f \longmapsto A$ eine wohldefinierte Abbildung von $\mathcal{S}(V)$ nach $\mathrm{Sym}_n(\mathbf{R})$, welche offensichtlich bijektiv ist. Durch die Vorgabe von $A = (\alpha_{k,j})_{1 \le j,k \le n} \in \mathrm{Sym}_n(\mathbf{R})$ wird eine lineare Abbildung mit

$$f(b_j) = \sum_{k=1}^{n} \alpha_{k,j} b_k$$

definiert. Die vorhergehende Rechnung zeigt insbesondere

$$\langle f(b_j), b_l \rangle = \langle b_j, f(b_l) \rangle$$

für $j, l = 1, ..., n$, woraus $f \in \mathcal{S}(V)$ folgt. Dies beweist die behauptete Bijektivität.

(ii) Der Beweis im unitären Fall verläuft analog zu (ii). □

Satz.

(i) *Es seien V ein unitärer Vektorraum und $f \in \mathcal{S}(V)$ eine selbstadjungierte Abbildung. Dann sind alle Eigenwerte von f reell.*

(ii) *Es seien V ein n-dimensionaler euklidischer Vektorraum und $f \in \mathcal{S}(V)$ eine selbstadjungierte Abbildung. Dann besitzt f (mit Vielfachheiten gezählt) n reelle Eigenwerte, d.h. das charakteristische Polynom p_f von f zerfällt über $\mathbf{R}$ in Linearfaktoren.*

Beweis.

(i) Mit einem Eigenwert x von f zum Eigenwert λ berechnen wir sofort

$$\lambda \cdot \langle x, x\rangle = \langle \lambda x, x\rangle = \langle f(x), x\rangle = \langle x, f(x)\rangle = \langle x, \lambda x\rangle = \bar{\lambda} \cdot \langle x, x\rangle.$$

Da $x \neq 0$ ist, ist $\langle x, x\rangle \neq 0$, und somit ergibt sich die Gleichheit

$$\lambda = \bar{\lambda},$$

d.h. $\lambda \in \mathbf{R}$, wie behauptet.

(ii) Wir wollen den Beweis auf (i) zurückführen. Dazu benötigen wir die Konstruktion der Skalarerweiterung für V von $\mathbf{R}$ nach $\mathbf{C}$. Für den n-dimensionalen Vektorraum V fixieren wir folgende Daten:

- $\mathcal{B} = \{b_1, ..., b_n\}$ sei eine Orthonormalbasis von V.
- Sind $x, y \in V$, $x = \sum_{j=1}^n \xi_j b_j$, $y = \sum_{j=1}^n \eta_j b_j$, so ergibt sich für das Skalarprodukt von V
$$\langle x, y\rangle = \sum_{j=1}^n \xi_j \cdot \eta_j,$$
d.h. $\langle \cdot, \cdot \rangle$ ist durch das Standardskalarprodukt gegeben.
- Ist $f \in \mathcal{S}(V)$ eine selbstadjungierte Abbildung von V, so ist dieser die symmetrische Matrix $A = (\alpha_{k,j})_{1 \leq j,k \leq n} \in \mathrm{Sym}_n(\mathbf{R})$ zugeordnet; diese ist charakterisiert durch
$$f(b_j) = \sum_{k=1}^n \alpha_{k,j} b_k \quad (j = 1, ..., n).$$

Wir kommen nun zur angekündigten Skalarerweiterung für V von $\mathbf{R}$ nach $\mathbf{C}$. Dazu definieren wir mit Hilfe der vorhergehenden Daten:

- $V_{\mathbf{C}} := \left\{\sum_{j=1}^n \xi_j b_j \,|\, \xi_j \in \mathbf{C}\right\}$. Dies ist ein n-dimensionaler komplexer Vektorraum mit Basis $\mathcal{B}$. Fasst man $V_{\mathbf{C}}$ als reellen Vektorraum auf, so ist dieser $2n$-dimensional und enthält V als n-dimensionalen Unterraum.

- Für $x, y \in V_{\mathbf{C}}$, $x = \sum_{j=1}^{n} \xi_j b_j, y = \sum_{j=1}^{n} \eta_j b_j$, definieren wir auf $V_{\mathbf{C}}$ durch

$$\langle x, y \rangle_{\mathbf{C}} := \sum_{j=1}^{n} \xi_j \cdot \bar{\eta}_j$$

  ein Skalarprodukt. Schränken wir $\langle \cdot, \cdot \rangle_{\mathbf{C}}$ auf $V \times V \subseteq V_{\mathbf{C}} \times V_{\mathbf{C}}$ ein, so erhalten wir ersichtlich das Skalarprodukt $\langle \cdot, \cdot \rangle$.

- Mit $A = (\alpha_{k,j})_{1 \leq j,k \leq n} \in \mathrm{Sym}_n(\mathbf{R})$ erhalten wir durch

$$f(b_j) = \sum_{k=1}^{n} \alpha_{k,j} b_k$$

  eine lineare Abbildung $f_{\mathbf{C}} : V_{\mathbf{C}} \longrightarrow V_{\mathbf{C}}$. Da nun auch $A \in \mathrm{Herm}_n(\mathbf{C})$ gilt, folgt aufgrund der Bijektion zwischen $\mathcal{S}(V_{\mathbf{C}})$ und $\mathrm{Herm}_n(\mathbf{C})$, daß $f_{\mathbf{C}}$ selbstadjungiert ist.

Wir können nun (ii) beweisen. Nach (i) wissen wir, daß das charakteristische Polynom $p_{f_{\mathbf{C}}}$ von $f_{\mathbf{C}}$ lauter reelle Nullstellen hat. Aufgrund der Gleichheit

$$p_{f_{\mathbf{C}}}(t) = \det(A - t \cdot E) = p_f(t)$$

kann nun auch das charakteristische Polynom p_f von f nur reelle Nullstellen haben. Dies beweist die Behauptung. □

Corollar. *Es sei* $A \in \mathrm{Sym}_n(\mathbf{R})$ *bzw.* $A \in \mathrm{Herm}_n(\mathbf{C})$. *Dann sind sämtliche Nullstellen von* $p_A \in \mathbf{R}[t]$ *bzw.* $p_A \in \mathbf{C}[t]$ *reell.* □

Satz.

(i) *Es sei* $A \in \mathrm{Sym}_n(\mathbf{R})$. *Dann existiert* $S \in \mathrm{O}_n(\mathbf{R})$ *mit*

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ *die* n *(mit Vielfachheiten gezählt) reellen Eigenwerte von* A *sind.*

(ii) *Es sei* $A \in \mathrm{Herm}_n(\mathbf{C})$. *Dann existiert* $S \in \mathrm{U}_n(\mathbf{C})$ *mit*

$$S^{-1} \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ *die* n *(mit Vielfachheiten gezählt) reellen Eigenwerte von* A *sind.*

Beweis.

(i) Wir führen eine vollständige Induktion nach n durch:
Für $n = 1$ ist die Behauptung klar. Für $n \geq 2$ können wir also in der Induktionsvoraussetzung annehmen, daß zu $A' \in \mathrm{Sym}_{n-1}(\mathbf{R})$ eine orthogonale Matrix $S' \in \mathrm{O}_{n-1}(\mathbf{R})$ mit

$$S'^{-1} \cdot A' \cdot S' = \begin{pmatrix} \lambda'_1 & & \\ & \ddots & \\ & & \lambda'_{n-1} \end{pmatrix}$$

existiert, wobei $\lambda'_1, \ldots, \lambda'_{n-1}$ die $(n-1)$ reellen Eigenwerte von A' sind.
Wir kommen nun zum Induktionsschritt. Wir betrachten den euklidischen Vektorraum $V = \mathbf{R}^n$ versehen mit dem Standardskalarprodukt und der Standardorthonormalbasis $\mathcal{E} = \{e_1, \ldots, e_n\}$. Zu $A \in \mathrm{Sym}_n(\mathbf{R})$ betrachten wir die lineare Abbildung $f \in \mathrm{L}(V)$ gegeben durch die Zuordnung

$$x \mapsto A \cdot x \ \ (x \in V = \mathbf{R}^n);$$

konstruktionsgemäß wird f bzgl. $\mathcal{E}$ durch die Matrix $A \in \mathrm{Sym}_n(\mathbf{R})$ dargestellt. Aufgrund der Bijektion zwischen $\mathcal{S}(V)$ und $\mathrm{Sym}_n(\mathbf{R})$ ist f selbstadjungiert. Es sei nun $b_1 \in V = \mathbf{R}^n$ Eigenvektor von f zum Eigenwert λ_1; da f selbstadjungiert ist, ist $\lambda_1 \in \mathbf{R}$. Ohne Einschränkung können wir annehmen, daß b_1 normiert ist. Nach dem Orthogonalisierungsverfahren von Gram-Schmidt können wir den Vektor b_1 zu einer Orthonormalbasis $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ von $V = \mathbf{R}^n$ ergänzen. Damit definieren wir die orthogonale Matrix

$$S_1 := (b_1, \ldots, b_n) \in \mathrm{O}_n(\mathbf{R});$$

S_1 ist die Basistransformation von $\mathcal{E}$ nach $\mathcal{B}$. Bezüglich $\mathcal{B}$ wird die lineare Abbildung f durch die Matrix

$$A_1 := S_1^{-1} \cdot A \cdot S_1$$

dargestellt. Aufgrund der Gleichheit

$${}^t A_1 = {}^t(S_1^{-1} \cdot A \cdot S_1) = {}^t S_1 \cdot {}^t A \cdot {}^t(S_1^{-1}) = S_1^{-1} \cdot A \cdot S_1 = A_1$$

ist A_1 symmetrisch, d.h. $A_1 \in \mathrm{Sym}_n(\mathbf{R})$. Da b_1 Eigenvektor von f mit Eigenwert λ_1 ist, folgt aufgrund der Symmetrie von A_1

$$A_1 = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix},$$

mit $A_2 \in \mathrm{Sym}_{n-1}(\mathbf{R})$. Nach Induktionvoraussetzung existiert $S_2 \in \mathrm{O}_{n-1}(\mathbf{R})$ mit

$$S_2^{-1} \cdot A_2 \cdot S_2 = \begin{pmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

mit den $(n-1)$ reellen Eigenwerten $\lambda_2, \ldots, \lambda_n$ von A_2. Wir betrachten nun die Matrix

$$S := S_1 \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix};$$

da $\mathrm{O}_n(\mathbf{R})$ eine Gruppe ist, folgt sofort $S \in \mathrm{O}_n(\mathbf{R})$. Damit erhalten wir

$$\begin{aligned} S^{-1} \cdot A \cdot S &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2^{-1} & \\ 0 & & & \end{pmatrix} \cdot S_1^{-1} \cdot A \cdot S_1 \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2^{-1} & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_2 & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2 & \\ 0 & & & \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S_2^{-1} A_2 S_2 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \end{aligned}$$

wie behauptet.

(ii) Der Beweis verläuft analog zu (i). $\square$

Kapitel 6

Affine Geometrie

6.1 Operation einer Gruppe auf einer Menge

Wir beginnen dieses Kapitel mit der folgenden, abstrakten

Definition. Es seien G eine Gruppe (mit der Verknüpfung $\circ$ und neutralem Element e) und M eine Menge. Eine Abbildung

$$G \times M \longrightarrow M,$$

gegeben durch die Zuordnung

$$(g, m) \mapsto g \bullet m \quad (g \in G, m \in M)$$

wird *Wirkung* oder *Operation der Gruppe G auf der Menge M* genannt, falls die beiden folgenden Eigenschaften erfüllt sind:

(i) $(g_1 \circ g_2) \bullet m = g_1 \bullet (g_2 \bullet m) \quad (g_1, g_2 \in G; m \in M)$,

(ii) $e \bullet m = m \quad (m \in M)$.

Beispiele.

(a) Sei $G = S_n$ die symmetrische Gruppe und $M = \{1, ..., n\}$. Dann wird durch

$$(\pi, j) \mapsto \pi(j) \quad (\pi \in S_n; j \in \{1, ..., n\})$$

eine Operation von S_n auf $\{1, ..., n\}$ gegeben.

(b) Sei $G = V$ ein K-Vektorraum (bzw. die dem K-Vektorraum zugrunde liegende abelsche Gruppe) und $M = V$. Dann wird durch

$$(x, v) \mapsto v + x \ (x, v \in V)$$

eine Operation von V auf sich selbst gegeben.

(c) Sei $G = \mathrm{SO}_2(\mathbf{R})$ die spezielle orthogonale Gruppe und $M = \mathbf{R}^2$. Dann wird durch

$$\left(\begin{pmatrix}\cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi\end{pmatrix}, \begin{pmatrix}\xi_1 \\ \xi_2\end{pmatrix}\right) \mapsto \begin{pmatrix}\xi_1\cos\varphi & - & \xi_2\sin\varphi \\ \xi_1\sin\varphi & + & \xi_2\cos\varphi\end{pmatrix}$$

$$\left(\begin{pmatrix}\xi_1 \\ \xi_2\end{pmatrix} \in \mathbf{R}^2; \begin{pmatrix}\cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi\end{pmatrix} \in \mathrm{SO}_2(\mathbf{R})\right)$$

eine Operation von $\mathrm{SO}_2(\mathbf{R})$ auf $\mathbf{R}^2$ gegeben.

(d) Sei $G = \mathrm{GL}(V)$ die allgemeine lineare Gruppe und $M = V$ ein K-Vektorraum. Dann wird durch

$$(f, x) \mapsto f(x) \quad (f \in \mathrm{GL}(V), x \in V)$$

eine Operation von $\mathrm{GL}(V)$ auf V gegeben.

(e) Sei $G = (\mathbf{R}, +)$ die additive Gruppe von $\mathbf{R}$ und $M = S^1$ der Einheitskreis, d.h.

$$S^1 = \{z \in \mathbf{C} \mid |\, z\,| = 1\}.$$

Dann wird durch

$$(\alpha, \zeta) \mapsto e^{i\alpha} \cdot \zeta \quad (\alpha \in \mathbf{R}, \zeta \in S^1)$$

eine Operation von $(\mathbf{R}, +)$ auf S^1 gegeben.

(f) Sei $G = \mathbf{C}^*$ die multiplikative Gruppe von $\mathbf{C}$ und $M = \mathbf{C}$. Dann wird durch

$$(z, w) \mapsto z \cdot w \ (z \in \mathbf{C}^*, w \in \mathbf{C})$$

eine Operation von $\mathbf{C}^*$ auf $\mathbf{C}$ gegeben.

Definition. Eine Gruppe G wirke auf der Menge M. Ist $m \in M$, so wird die Teilmenge

$$G \bullet m = \{g \bullet m \mid g \in G\}$$

die *Bahn von m unter G* genannt.

In den vorhergehenden Beispielen (a), (b), (e) überzeugt man sich leicht, daß die Bahn eines einzigen Elements die gesamte zugrunde liegende Menge ausmacht. Dies führt zur folgenden Definition.

Definition. Eine Gruppe G wirke auf der Menge M. Diese Wirkung heißt *transitiv,* wenn für alle $m, m' \in M$ ein $g \in G$ mit der Eigenschaft

$$m' = g \bullet m$$

existiert. Die Wirkung heißt *einfach transitiv*, falls das Gruppenelement g eindeutig bestimmt ist.

Beispiele. Wir überprüfen die vorhergehenden Beispiele auf Transitivität bzw. einfache Transitivität:

(a) S_n wirkt transitiv auf $\{1, ..., n\}$.

(b) V wirkt einfach transitiv auf sich selbst.

(c) $\mathrm{SO}_2(\mathbf{R})$ wirkt nicht transitiv auf $\mathbf{R}^2$.

(d) $\mathrm{GL}(V)$ wirkt nicht transitiv auf V.

(e) $(\mathbf{R}, +)$ wirkt transitiv auf S^1.

(f) $\mathbf{C}^*$ wirkt nicht transitiv auf $\mathbf{C}$.

Lemma. *Eine Gruppe G wirke auf der Menge M.*

(i) *Ist die Wirkung transitiv, so gilt für jedes $m \in M$ die Gleichheit $M = G \bullet m$.*

(ii) *Ist die Wirkung einfach transitiv, so besteht eine Bijektion zwischen M und G.*

Beweis.

(i) Es sei m ein beliebiges Element von M. Für $m' \in M$ existiert aufgrund der transitiven Wirkung von G ein $g \in G$ mit $m' = g \bullet m$, d.h. $m' \in G \bullet m$. Dies zeigt $M = G \bullet m$.

(ii) Wir fixieren ein $m \in M$. Damit definieren wir eine Abbildung

$$\psi = \psi_m : G \longrightarrow M$$

vermöge der Zuordnung $g \mapsto g \bullet m$. Aufgrund der transitiven Wirkung ist ψ surjektiv. Aufgrund der Einfachheit ist die Abbildung auch injektiv, was die Bijektivität von ψ beweist. □

Definition. Eine Gruppe G wirke auf der Menge M, und es sei $m \in M$. Die Menge

$$G_m := \{g \in G \mid g \bullet m = m\}$$

heißt der *Stabilisator von m in G*.

Lemma. *Eine Gruppe G wirke auf der Menge M. Der Stabilisator G_m von $m \in M$ ist eine Untergruppe von G.*

Beweis. Übungsaufgabe □

6.2 Affine Räume

Wir wenden nun das Konzept der einfach transitiven Wirkung einer Gruppe auf Vektorräume an. Dazu sei V ein Vektorraum über K. Wir bezeichnen die V zugrunde liegende abelsche Gruppe ebenfalls mit V.

Definition. Eine Menge M, auf der der K-Vektorraum V einfach transitiv operiert, heißt ein *affiner Raum (mit zugehörigem K-Vektorraum V).*

Bemerkung. Das Beispiel (b) im vorhergehenden Abschnitt zeigt die Existenz eines affinen Raumes mit zugehörigem K-Vektorraum V.

Lemma. *Sind M, M' affine Räume mit zugehörigem K-Vektorraum V, so besteht eine Bijektion zwischen M und M'.*

Beweis. Nach dem ersten Lemma in Abschnitt 6.1 sind sowohl M als auch M' bijektiv zu V, also besteht auch eine Bijektion zwischen M und M'. □

Bezeichnungen. Wir bezeichnen den bis auf Bijektivität eindeutig bestimmten affinen Raum mit zugehörigem K-Vektorraum V durch $\mathbf{A}(V)$. Die Elemente von $\mathbf{A}(V)$ nennen wir *Punkte* und bezeichnen sie durch $O, P, Q, R, ...$ Die einfach transitive Wirkung $\bullet$ von V auf $\mathbf{A}(V)$ bezeichnen wir durch ein $+$ - Zeichen (nicht zu verwechseln mit der Addition von Vektoren), d.h.

$$x \bullet P := P + x \quad (x \in V, P \in \mathbf{A}(V)).$$

Wir verstehen darunter, daß der Punkt $P + x \in \mathbf{A}(V)$ erhalten wird durch „Anheften des Vektors $x \in V$ an den Punkt $P \in \mathbf{A}(V)$". Ist $Q = P + x \in \mathbf{A}(V)$, so nennen wir den eindeutig bestimmten Vektor $x \in V$, mit Hilfe dessen Q aus P hervorgeht, den *Verbindungsvektor von P nach Q* und schreiben dafür

$$x = \overrightarrow{PQ}\ .$$

Indem wir schließlich einen Punkt $O \in \mathbf{A}(V)$ fixieren, stellt sich die Bijektion

$$\psi = \psi_O : V \longrightarrow \mathbf{A}(V)$$

dar durch

$$x \mapsto P = O + x.$$

Die Umkehrabbildung

$$\varphi = \varphi_O : \mathbf{A}(V) \longrightarrow V$$

ist dann gegeben durch die Zuordnung

$$P \mapsto \overrightarrow{OP}\ .$$

Wir beachten, daß die Bijektion $\varphi = \varphi_O$ nicht kanonisch ist, da sie von der Wahl des Punktes O abhängt.

Bemerkung. Wir betrachten jetzt den Spezialfall $V = K^n$. In diesem Falle schreiben wir

$$\mathbf{A}^n(K) = \mathbf{A}(K^n).$$

Die Elemente von K^n, d.h. die Vektoren, bezeichnen wir als n-Tupel

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in K^n$$

und verstehen darunter den entsprechenden Ortsvektor im K^n. Die Elemente von $\mathbf{A}^n(K) = \mathbf{A}(K^n)$, d.h. die Punkte, bezeichnen wir ebenfalls als n-Tupel mit dem Index „aff", d.h.

$$P = \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_n \end{pmatrix}_{\text{aff}} \in \mathbf{A}^n(K)$$

und verstehen darunter den durch den Ortsvektor $x = \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_n \end{pmatrix}$ festgelegten Punkt von $\mathbf{A}^n(K)$. Die einfach transitive Operation von K^n auf $\mathbf{A}^n(K)$ ist damit wie folgt gegeben:

$$\left(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_n \end{pmatrix}_{\text{aff}} \right) \mapsto \begin{pmatrix} \pi_1 + \xi_1 \\ \vdots \\ \pi_n + \xi_n \end{pmatrix}_{\text{aff}}$$

Mit $O = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_{\text{aff}} \in \mathbf{A}^n(K)$ erhalten wir die Bijektion

$$\varphi_O : \mathbf{A}^n(K) \longrightarrow K^n,$$

gegeben durch die Zuordnung

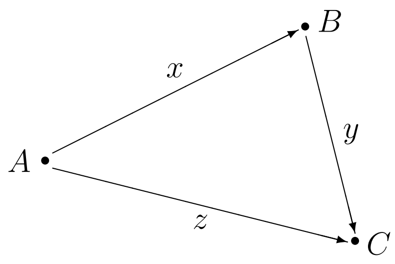
$$\begin{pmatrix} \pi_1 \\ \vdots \\ \pi_n \end{pmatrix}_{\text{aff}} \mapsto \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_n \end{pmatrix}.$$

Lemma. *Es seien $\mathbf{A}(V)$ der affine Raum mit zugehörigem K-Vektorraum V und $A, B, C \in \mathbf{A}(V)$. Dann gilt*

$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}\ .$$

Beweis. Wir haben definitionsgemäß

$$B = A + x, \quad x = \overrightarrow{AB},$$
$$C = B + y, \quad y = \overrightarrow{BC},$$
$$C = A + z, \quad z = \overrightarrow{AC}\ .$$



Aufgrund der Operation von V auf $\mathbf{A}(V)$ folgt damit

$$C = B + y = (A + x) + y = A + (x + y).$$

Aufgrund der einfachen Transitivität der Operation folgt somit

$$\overrightarrow{AC} = z = x + y = \overrightarrow{AB} + \overrightarrow{BC}\ .$$

Dies beweist die Behauptung. $\square$

Definition. Die *Dimension* $\dim_K \mathbf{A}(V)$ eines affinen Raumes $\mathbf{A}(V)$ mit zugehörigem K-Vektorraum V ist definiert durch

$$\dim_K \mathbf{A}(V) := \dim_K V.$$

Beispiel. Für den affinen Raum $\mathbf{A}(V) = \mathbf{A}^n(K)$ erhalten wir

$$\dim_K \mathbf{A}^n(K) = n.$$

6.3 Affine Unterräume

In diesem Abschnitt sei $\mathbf{A}(V)$ ein fixierter affiner Raum mit zugehörigem K-Vektorraum V.

Definition. Eine nicht-leere Teilmenge $N \subseteq \mathbf{A}(V)$ heißt *affiner Unterraum von* $\mathbf{A}(V)$, falls ein Punkt $P \in N$ und ein Unterraum $W \subseteq V$ existieren, so daß

$$N = \{Q \in \mathbf{A}(V) \,|\, Q = P + x, x \in W\}$$

gilt.

Lemma. *Es seien $\mathbf{A}(V)$ ein affiner Raum und $N \subseteq \mathbf{A}(V)$ ein affiner Unterraum. Dann ist N ein affiner Raum.*

Beweis. Definitionsgemäß existieren ein Punkt $P \in N$ und ein Unterraum $W \subseteq V$ mit der Eigenschaft

$$N = \{Q \in \mathbf{A}(V) \,|\, Q = P + x, x \in W\}.$$

Wir behaupten, daß W einfach transitiv auf N operiert: Für einen beliebigen Punkt $Q \in N$, d.h. $Q = P + x$ mit $x \in W$, und einen beliebigen Vektor $y \in W$ setzen wir

$$y \bullet Q := Q + y \in \mathbf{A}(V).$$

Wegen $Q + y = (P + x) + y = P + (x + y)$ und $x + y \in W$ ist $y \bullet Q \in N$. Damit erhalten wir eine Operation von W auf N. Da diese Operation die Einschränkung der Operation von V auf $\mathbf{A}(V)$ ist, ist die Operation $\bullet$ einfach transitiv. Damit erhalten wir

$$N = \mathbf{A}(W),$$

und N ist somit ein affiner Raum. □

Bemerkung. Mit den Bezeichnungen des vorhergehenden Lemma haben wir

$$N = \mathbf{A}(W) = \{Q \in \mathbf{A}(V) \,|\, Q = P + x, x \in W\};$$

wir können dafür auch kurz

$$N = \mathbf{A}(W) = P + W$$

schreiben. Wir stellen fest, daß die Definition des affinen Unterraums N nicht von der Wahl des Punktes $P \in N$ abhängt. Ist nämlich $P' \in N$ ein anderer Punkt, so haben wir

$$N = \mathbf{A}(W) = P + W = (P' + \overrightarrow{P'P}) + W = P' + (\overrightarrow{P'P} + W).$$

Da $P', P \in N = \mathbf{A}(W)$ ist, ist $\overrightarrow{P'P} \in W$, also $\overrightarrow{P'P} + W = W$, und wir erhalten

$$N = \mathbf{A}(W) = P' + W.$$

Definition. Es seien $\mathbf{A}(V)$ ein affiner Raum mit zugehörigem K-Vektorraum V und $\mathbf{A}(W)$ ein affiner Unterraum von $\mathbf{A}(V)$, d.h. W ist ein Unterraum von V; für einen beliebigen Punkt $P \in \mathbf{A}(W)$ gilt dann $\mathbf{A}(W) = P + W$. Wir nennen P *einen Aufpunkt von* $\mathbf{A}(W)$ und den Unterraum W *die Richtung des affinen Unterraums* $\mathbf{A}(W)$.

Definition. Es seien $\mathbf{A}(V)$ ein affiner Raum mit zugehörigem K-Vektorraum V und $\mathbf{A}(W)$ ein affiner Unterraum von $\mathbf{A}(V)$, d.h. W ist ein Unterraum von V.

- Gilt $\dim_K \mathbf{A}(W) = 0$, so nennen wir $\mathbf{A}(W)$ einen *(affinen) Punkt von* $\mathbf{A}(V)$.
- Gilt $\dim_K \mathbf{A}(W) = 1$, so nennen wir $\mathbf{A}(W)$ eine *(affine) Gerade von* $\mathbf{A}(V)$.
- Gilt $\dim_K \mathbf{A}(W) = 2$, so nennen wir $\mathbf{A}(W)$ eine *(affine) Ebene von* $\mathbf{A}(V)$.

- Gilt $\mathrm{codim}_K \mathbf{A}(W) = \dim_K(V/W) = 1$, so nennen wir $\mathbf{A}(W)$ eine *(affine) Hyperebene von* $\mathbf{A}(V)$.

Beispiel. Es seien $P_1, P_2 \in \mathbf{A}(V)$ zwei verschiedene Punkte. Mit dem 1-dimensionalen Unterraum

$$W := \langle \overrightarrow{P_1P_2} \rangle \subseteq V$$

definieren wir die Gerade

$$\mathbf{A}(W) = P_1 + \langle \overrightarrow{P_1P_2} \rangle$$

in $\mathbf{A}(V)$. Diese enthält P_1, P_2; wir nennen sie *die Gerade durch* P_1, P_2.

Definition. Es seien $\mathbf{A}(W_1), \mathbf{A}(W_2)$ affine Unterräume des affinen Raums $\mathbf{A}(V)$. Dann wird der *kleinste affine Raum, der* $\mathbf{A}(W_1)$ *und* $\mathbf{A}(W_2)$ *enthält*, mit $\mathbf{A}(W_1, W_2)$ bezeichnet.

Beispiel. Es seien $P_1, P_2 \in \mathbf{A}(V)$ zwei Punkte; P_1, P_2 sind selbst affine Unterräume von $\mathbf{A}(V)$. Wir bezeichnen (durch leichten Mißbrauch der Bezeichnungen) mit $\mathbf{A}(P_1, P_2)$ den kleinsten affinen Raum, der P_1 und P_2 enthält. Gilt $P_1 = P_2$, so haben wir

$$\mathbf{A}(P_1, P_2) = \{P_1\}.$$

Gilt $P_1 \neq P_2$, so haben wir

$$\mathbf{A}(P_1, P_2) = P_1 + \langle \overrightarrow{P_1P_2} \rangle,$$

d.h. $\mathbf{A}(P_1, P_2)$ ist die Gerade durch P_1, P_2.

Satz. *Es seien* $\mathbf{A}(W_1), \mathbf{A}(W_2)$ *affine Unterräume eines endlich dimensionalen affinen Raums* $\mathbf{A}(V)$*. Dann bestehen die Dimensionsformeln:*

(i) *Gilt* $\mathbf{A}(W_1) \cap \mathbf{A}(W_2) \neq \emptyset$*, so ist*

$$\dim_K \mathbf{A}(W_1, W_2) = \dim_K \mathbf{A}(W_1) + \dim_K \mathbf{A}(W_2) - \dim_K(W_1 \cap W_2).$$

(ii) *Gilt* $\mathbf{A}(W_1) \cap \mathbf{A}(W_2) = \emptyset$*, so ist*

$$\dim_K \mathbf{A}(W_1, W_2) = \dim_K \mathbf{A}(W_1) + \dim_K \mathbf{A}(W_2) + 1 - \dim_K(W_1 \cap W_2).$$

Wir schicken dem Beweis des Satzes zwei Hilfssätze voraus.

Hilfssatz 1. *Es mögen die Voraussetzungen des Satzes gelten. Weiter seien* $P_1 \in \mathbf{A}(W_1)$ *und* $P_2 \in \mathbf{A}(W_2)$*. Dann gilt*

$$\mathbf{A}(W_1, W_2) = P_1 + (\langle \overrightarrow{P_1P_2} \rangle + W_1 + W_2).$$

Beweis. Zunächst haben wir die offensichtlichen Inklusionen

$$\begin{array}{rcccl} \mathbf{A}(W_1, W_2) & \supseteq & \mathbf{A}(W_1) & = & P_1 + W_1, \\ \mathbf{A}(W_1, W_2) & \supseteq & \mathbf{A}(W_2) & = & P_2 + W_2. \end{array}$$

Da $\mathbf{A}(W_1, W_2)$ ein affiner Raum ist, muß mit $P_1, P_2 \in \mathbf{A}(W_1, W_2)$ auch die Gerade durch P_1, P_2, d.h.

$$\mathbf{A}(P_1, P_2) = P_1 + \langle \overrightarrow{P_1P_2} \rangle,$$

in $\mathbf{A}(W_1, W_2)$ enthalten sein (ist $P_1 = P_2$, so degeneriert diese Gerade zu einem Punkt). Zusammengenommen erhalten wir die Inklusion

$$\mathbf{A}(W_1, W_2) \supseteq P_1 + (\langle \overrightarrow{P_1P_2} \rangle + W_1 + W_2).$$

Aufgrund der Minimalität von $\mathbf{A}(W_1, W_2)$ muß sogar die Gleichheit gelten, d.h.

$$\mathbf{A}(W_1, W_2) = P_1 + (\langle \overrightarrow{P_1P_2} \rangle + W_1 + W_2).$$

□

Hilfssatz 2. *Es mögen die Voraussetzungen des Satzes gelten. Weiter seien $P_1 \in \mathbf{A}(W_1)$ und $P_2 \in \mathbf{A}(W_2)$. Dann besteht die Äquivalenz*

$$\mathbf{A}(W_1) \cap \mathbf{A}(W_2) \neq \emptyset \iff \overrightarrow{P_1P_2} \in W_1 + W_2.$$

Beweis. ($\Longrightarrow$): Ist $\mathbf{A}(W_1) \cap \mathbf{A}(W_2) \neq \emptyset$, so existiert $Q \in \mathbf{A}(W_1) \cap \mathbf{A}(W_2)$, d.h. $Q \in \mathbf{A}(W_1)$ und $Q \in \mathbf{A}(W_2)$. Damit folgt

$$\overrightarrow{P_1Q} \in W_1, \ \overrightarrow{QP_2} \in W_2,$$

also

$$\overrightarrow{P_1P_2} = \overrightarrow{P_1Q} + \overrightarrow{QP_2} \in W_1 + W_2.$$

($\Longleftarrow$): Es sei umgekehrt $\overrightarrow{P_1P_2} \in W_1 + W_2$, d.h. es existieren $x_1 \in W_1, x_2 \in W_2$ mit

$$\overrightarrow{P_1P_2} = x_1 + x_2.$$

Damit setzen wir $Q := P_1 + x_1 \in \mathbf{A}(W_1)$. Für den Punkt Q stellen wir andererseits fest

$$\begin{aligned} Q &= P_1 + x_1 = P_1 + (x_1 + (x_2 - x_2)) \\ &= P_1 + ((x_1 + x_2) - x_2) = P_1 + (\overrightarrow{P_1P_2} - x_2) \\ &= P_2 + (-x_2) \in \mathbf{A}(W_2). \end{aligned}$$

Somit haben wir

$$\mathbf{A}(W_1) \cap \mathbf{A}(W_2) \neq \emptyset.$$

□

Wir kommen nun zum Beweis des Satzes.

Beweis.

(i) Aufgrund der beiden vorhergehenden Hilfssätze stellen wir mit $P_1 \in \mathbf{A}(W_1)$ und $P_2 \in \mathbf{A}(W_2)$ fest:

$$\mathbf{A}(W_1, W_2) = P_1 + (\langle \overrightarrow{P_1P_2} \rangle + W_1 + W_2) = P_1 + (W_1 + W_2).$$

Aufgrund der Definition der Dimension affiner Räume folgt mit Hilfe des Dimensionssatzes für Unterräume

$$\begin{aligned} \dim_K \mathbf{A}(W_1, W_2) &= \dim_K (W_1 + W_2) = \dim_K W_1 + \dim_K W_2 - \dim_K (W_1 \cap W_2) \\ &= \dim_K \mathbf{A}(W_1) + \dim_K \mathbf{A}(W_2) - \dim_K (W_1 \cap W_2). \end{aligned}$$

(ii) Wiederum stellen wir mit Hilfe der beiden vorhergehenden Hilfssätze mit $P_1 \in \mathbf{A}(W_1)$ und $P_2 \in \mathbf{A}(W_2)$ fest:

$$\mathbf{A}(W_1, W_2) = P_1 + (\langle \overrightarrow{P_1P_2} \rangle + W_1 + W_2),$$

wobei $\overrightarrow{P_1P_2} \notin W_1 + W_2$ ist, d.h. $\overrightarrow{P_1P_2}$ ist linear unabhängig von $W_1 + W_2$. Wiederum folgt nun mit Hilfe des Dimensionssatzes für Unterräume

$$\begin{aligned} \dim_K \mathbf{A}(W_1, W_2) &= \dim_K (\langle \overrightarrow{P_1P_2} \rangle + W_1 + W_2) = \dim_K (W_1 + W_2) + 1 \\ &= \dim_K W_1 + \dim_K W_2 - \dim_K (W_1 \cap W_2) + 1 \\ &= \dim_K \mathbf{A}(W_1) + \dim_K \mathbf{A}(W_2) + 1 - \dim_K (W_1 \cap W_2). \end{aligned}$$

□

6.4 Parallelismus

Wir fixieren einen affinen Raum $\mathbf{A}(V)$ mit zugehörigem K-Vektorraum V.

Definition.

(i) Zwei affine Unterräume $\mathbf{A}(W_1), \mathbf{A}(W_2) \subseteq \mathbf{A}(V)$ heißen *parallel*, in Zeichen $\mathbf{A}(W_1) \parallel \mathbf{A}(W_2)$, falls sie die gleiche Richtung haben, d.h. falls $W_1 = W_2$ gilt.

(ii) Zwei affine Unterräume $\mathbf{A}(W_1), \mathbf{A}(W_2) \subseteq \mathbf{A}(V)$ heißen *schwach parallel*, in Zeichen $\mathbf{A}(W_1) \lhd \mathbf{A}(W_2)$, falls $W_1 \subset W_2$ und $W_1 \neq W_2$ gilt.

Satz. *Es seien $\mathbf{A}(W_1), \mathbf{A}(W_2)$ affine Unterräume von $\mathbf{A}(V)$. Dann gilt:*

(i) *Sind $\mathbf{A}(W_1)$ und $\mathbf{A}(W_2)$ parallel, so gilt entweder $\mathbf{A}(W_1) = \mathbf{A}(W_2)$ oder $\mathbf{A}(W_1) \cap \mathbf{A}(W_2) = \emptyset$.*

(ii) *Sind $\mathbf{A}(W_1)$ und $\mathbf{A}(W_2)$ schwach parallel, so gilt entweder $\mathbf{A}(W_1) \subset \mathbf{A}(W_2), \mathbf{A}(W_1) \neq \mathbf{A}(W_2)$ oder $\mathbf{A}(W_1) \cap \mathbf{A}(W_2) = \emptyset$.*

Beweis.

(i) Es gelte $\mathbf{A}(W_1) \parallel \mathbf{A}(W_2)$ und $\mathbf{A}(W_1) \cap \mathbf{A}(W_2) \neq \emptyset$. Dann existiert $P \in \mathbf{A}(W_1) \cap \mathbf{A}(W_2)$. Für diesen Punkt gilt

$$\mathbf{A}(W_1) = P + W_1,\ \mathbf{A}(W_2) = P + W_2.$$

Aufgrund der Gleichheit $W_1 = W_2$ folgt jetzt

$$\mathbf{A}(W_1) = \mathbf{A}(W_2).$$

(ii) Übungsaufgabe. □

Satz.

(i) *Es seien $\mathbf{A}(W)$ ein affiner Unterraum von $\mathbf{A}(V)$ und $P \in \mathbf{A}(V)$ ein Punkt. Dann existiert ein eindeutig bestimmter affiner Unterraum $N \subseteq \mathbf{A}(V)$ mit den beiden Eigenschaften*

$$N \parallel \mathbf{A}(W),\ P \in N.$$

(ii) *Es seien $\mathbf{A}(W_1) \lhd \mathbf{A}(W_2)$ schwach parallele Unterräume von $\mathbf{A}(V)$. Dann existiert ein affiner Unterraum $N \subset \mathbf{A}(W_2), N \neq \mathbf{A}(W_2)$ mit der Eigenschaft*

$$N \parallel \mathbf{A}(W_1).$$

Beweis.

(i) Der gesuchte affine Unterraum $N \subseteq \mathbf{A}(V)$ ist gegeben durch $N = P + W$; dafür gilt nämlich

$$N \parallel \mathbf{A}(W),\ P \in N.$$

N ist damit eindeutig festgelegt.

(ii) Übungsaufgabe. □

6.5 Affine Basen, affine Koordinaten

Wir fixieren einen affinen Raum $\mathbf{A}(V)$ mit zugehörigem K-Vektorraum V.

Definition.

(i) $(n+1)$ Punkte $P_0, P_1, ..., P_n \in \mathbf{A}(V)$ heißen *affin unabhängig*, falls die n Verbindungsvektoren

$$\overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_n} \in V$$

linear unabhängig sind.

(ii) Bilden die Vektoren $\{\overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_n}\}$ eine (geordnete) Basis von V, so wird die Menge

$$\mathcal{P} = \{P_0, P_1, ..., P_n\}$$

eine *(geordnete) affine Basis von* $\mathbf{A}(V)$ genannt.

Definition. Es sei $\mathcal{P} = \{P_0, P_1, ..., P_n\}$ eine (geordnete) affine Basis von $\mathbf{A}(V)$. Ist $P \in \mathbf{A}(V)$, so heißen die n eindeutig bestimmten Skalare $\xi_1, ..., \xi_n \in K$ mit der Eigenschaft

$$\overrightarrow{P_0P} = \sum_{j=1}^{n} \xi_j \cdot \overrightarrow{P_0P_j}$$

die affinen Koordinaten von P bezüglich der affinen Basis $\mathcal{P}$. Wir notieren diese in der Form

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}_{\text{aff}} \in \mathbf{A}^n(K).$$

Beispiel. Die $(n+1)$ Punkte

$$P_0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_{\text{aff}}, \; P_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{\text{aff}}, \; ..., \; P_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}_{\text{aff}} \in \mathbf{A}^n(K)$$

bilden eine affine Basis von $\mathbf{A}^n(K)$, da die Verbindungsvektoren

$$\overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_n} \in K^n$$

eine Basis des K^n (Standardbasis) bilden. Wir nennen die Menge

$$\mathcal{P} = \{P_0, P_1, ..., P_n\}$$

die affine Standardbasis des $\mathbf{A}^n(K)$.

Beispiel. Die drei Punkte

$$P_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}_{\text{aff}}, \; P_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\text{aff}}, \; P_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_{\text{aff}} \in \mathbf{A}^2(\mathbf{R})$$

bilden eine affine Basis von $\mathbf{A}^2(\mathbf{R})$, da die Verbindungsvektoren

$$\overrightarrow{P_0P_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \overrightarrow{P_0P_2} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in \mathbf{R}^2$$

linear unabhängig sind.
Die affinen Koordinaten des Punktes $P = \begin{pmatrix} 5 \\ 8 \end{pmatrix}_{\text{aff}} \in \mathbf{A}^2(\mathbf{R})$ bezüglich der affinen Basis $P = \{P_0, P_1, P_2\}$ sind gegeben durch $\begin{pmatrix} 2 \\ 3 \end{pmatrix}_{\text{aff}}$, da die Beziehung

$$\overrightarrow{P_0P} = 2 \cdot \overrightarrow{P_0P_1} + 3 \cdot \overrightarrow{P_0P_2} \Longleftrightarrow \begin{pmatrix} 5 \\ 8 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 3 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

gilt.

6.6 Affine Abbildungen

Es seien $\mathbf{A}(V)$ bzw. $\mathbf{A}(W)$ affine Räume mit den zugehörigen K-Vektorräumen V bzw. W.

Vorbemerkung. Es seien

$$f : \mathbf{A}(V) \longrightarrow \mathbf{A}(W)$$

eine beliebige Abbildung (zwischen den $\mathbf{A}(V)$ und $\mathbf{A}(W)$ zugrunde liegenden Mengen) und $P \in \mathbf{A}(V)$ ein beliebiger Punkt. Mit Hilfe der Zuordnung

$$\overrightarrow{PQ} \quad \mapsto \quad \overrightarrow{f(P)f(Q)} \quad (Q \in \mathbf{A}(V))$$

wird durch die Daten f, P eine Abbildung

$$\overrightarrow{f_P} : V \longrightarrow W$$

definiert, d.h. $\overrightarrow{f_P}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}$. Man beachte, daß $\overrightarrow{f_P}$ eine Abbildung zwischen den V und W zugrunde liegenden Mengen ist und somit in der Regel nicht von der Linearität von $\overrightarrow{f_P}$ gesprochen werden kann.
Indem wir an die bijektiven Abbildungen $\varphi_P : \mathbf{A}(V) \longrightarrow V, \varphi_{f(P)} : \mathbf{A}(W) \longrightarrow W$ erinnern, können wir die Definition der Abbildung $\overrightarrow{f_P}$ durch die Kommutativität des folgenden Diagramms ausdrücken:

Nach dieser Vorbemerkung sind wir in der Lage, affine Abbildungen zu definieren.

Definition. Eine Abbildung $f : \mathbf{A}(V) \longrightarrow \mathbf{A}(W)$ zwischen den affinen Räumen $\mathbf{A}(V)$ und $\mathbf{A}(W)$ heißt eine *affine Abbildung*, falls ein Punkt $P \in \mathbf{A}(V)$ existiert, so daß die Abbildung $\overrightarrow{f_P}: V \longrightarrow W$, gegeben durch

$$\overrightarrow{f_P}\,(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)} \qquad (Q \in \mathbf{A}(V)),$$

(siehe Vorbemerkung) linear ist.

Lemma. *Die Definition einer affinen Abbildung $f : \mathbf{A}(V) \longrightarrow \mathbf{A}(W)$ ist unabhängig von dem in der Definition ausgezeichneten Punkt $P \in \mathbf{A}(V)$.*

Beweis. Wir haben zu zeigen, daß für einen beliebigen Vektor $x \in V$ das Bild $\overrightarrow{f_P}\,(x) \in W$ unabhängig von P ist. Dazu sei $Q \in \mathbf{A}(V)$ ein beliebig gewählter Punkt. Indem wir $R := Q + x \in \mathbf{A}(V)$ betrachten, erhalten wir zwei Punkte $Q, R \in \mathbf{A}(V)$ mit Verbindungsvektor $\overrightarrow{QR} = x$. Aufgrund der Beziehung

$$\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR}$$

haben wir

$$x = \overrightarrow{PR} - \overrightarrow{PQ}\,.$$

Mit Hilfe der vorausgesetzten Linearität von $\overrightarrow{f_P}$ erhalten wir somit

$$\overrightarrow{f_P}\,(x) = \overrightarrow{f_P}\,(\overrightarrow{PR} - \overrightarrow{PQ}) = \overrightarrow{f_P}\,(\overrightarrow{PR}) - \overrightarrow{f_P}\,(\overrightarrow{PQ}) = \overrightarrow{f(P)f(R)} - \overrightarrow{f(P)f(Q)}.$$

Wegen

$$\overrightarrow{f(P)f(R)} = \overrightarrow{f(P)f(Q)} + \overrightarrow{f(Q)f(R)}$$

folgt somit

$$\overrightarrow{f_P}\,(x) = \overrightarrow{f(Q)f(R)}.$$

Damit ist die behauptete Unabhängigkeit von $\overrightarrow{f_P}\,(x)$ von P nachgewiesen. □

Bemerkung. Ist $f : \mathbf{A}(V) \longrightarrow \mathbf{A}(W)$ eine affine Abbildung, so ist es nach dem vorhergehenden Lemma gerechtfertigt, die durch f induzierte lineare Abbildung $\overrightarrow{f_P} \in \mathrm{L}(V, W)$ einfach durch $\vec{f} \in \mathrm{L}(V, W)$ zu bezeichnen. Zur Definition von $\vec{f}$ halten wir die beiden äquivalenten, charakterisierenden Gleichungen

$$\vec{f}(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)} \iff f(Q) = f(P) + \vec{f}(\overrightarrow{PQ})$$

fest; hierbei sind $P, Q \in \mathbf{A}(V)$.

Bezeichnungen. Es seien $\mathbf{A}(V)$ bzw. $\mathbf{A}(W)$ affine Räume mit den zugehörigen K-Vektorräumen V bzw. W. Damit setzen wir

$$\mathrm{A}(V, W) := \{f : \mathbf{A}(V) \longrightarrow \mathbf{A}(W) \mid f \text{ affin } \},$$

$$\mathrm{A}(V) := \mathrm{A}(V,V),$$
$$\mathrm{GA}(V) := \{f \in \mathrm{A}(V) \,|\, f \text{ bijektiv }\}.$$

Bemerkung. Die Menge $\mathrm{GA}(V)$ der bijektiven affinen Selbstabbildungen von $\mathbf{A}(V)$ bildet ersichtlich eine Gruppe bezüglich der Komposition von Abbildungen. Wir nennen $\mathrm{GA}(V)$ *die affine Gruppe zum K-Vektorraum V.*

Lemma. *Seien $f \in \mathrm{A}(V,W)$ und $\mathbf{A}(V')$ ein affiner Unterraum von $\mathbf{A}(V)$. Dann ist das Bild $f(\mathbf{A}(V'))$ ein affiner Unterraum von $\mathbf{A}(W)$ mit der Richtung $\vec{f}(V')$.*

Beweis. Definitionsgemäß existiert ein Punkt $P \in \mathbf{A}(V')$ mit der Eigenschaft

$$\mathbf{A}(V') = P + V'.$$

Weiter induziert die affine Abbildung $f \in \mathrm{A}(V,W)$ eine lineare Abbildung $\vec{f} \in \mathrm{L}(V,W)$ derart, daß

$$f(\mathbf{A}(V')) = f(P+V') = f(P) + \vec{f}(V')$$

gilt. Damit ergibt sich die Behauptung. $\square$

Bemerkung. Ist speziell $f \in \mathrm{GA}(V)$, so werden mittels f Geraden, Ebenen etc. von $\mathbf{A}(V)$ in ebensolche überführt.

Wir wollen zunächst zwei wichtige Klassen von Beispielen affiner Abbildungen diskutieren.

Translationen.

Ein Vektor $v \in V$ sei festgehalten. Damit definieren wir durch

$$f_v(P) := P + v \quad (P \in \mathbf{A}(V))$$

eine Abbildung von $\mathbf{A}(V)$ in sich.

Lemma 1. *Mit den vorhergehenden Bezeichnungen gilt $f_v \in \mathrm{GA}(V)$.*

Beweis. Die Abbildung f_v ist offensichtlich eine Bijektion. Es bleibt somit die Affinität von f_v nachzuweisen. Für zwei Punkte $P, Q \in \mathbf{A}(V)$ haben wir definitionsgemäß

$$\begin{aligned} f_v(Q) &= f_v(P) + \overrightarrow{f_v(P)f_v(Q)} \\ &= (P+v) + \overrightarrow{f_v}\,(\overrightarrow{PQ}) \\ &= Q + (\overrightarrow{QP} + v + \overrightarrow{f_v}(\overrightarrow{PQ})). \end{aligned}$$

Da andererseits auch

$$f_v(Q) = Q + v$$

gilt, erhalten wir die Gleichung

$$v = \overrightarrow{QP} + v + \overrightarrow{f_v}(\overrightarrow{PQ}),$$

d.h.

$$\overrightarrow{f_v}(\overrightarrow{PQ}) = -\overrightarrow{QP} = \overrightarrow{PQ}\ .$$

Somit ist die durch $\overrightarrow{f_v}$ induzierte Abbildung $\overrightarrow{f_v} : V \longrightarrow V$ die identische Abbildung, also insbesondere linear. Dies zeigt $f_v \in \mathrm{GA}(V)$. □

Definition. Wir setzen

$$\mathrm{T}(V) := \{f \in \mathrm{GA}(V) \,|\, \exists\ v \in V : f = f_v\}.$$

Die Elemente von $\mathrm{T}(V)$ werden *Translationen um den Vektor* v genannt.

Lemma 2. *Die Menge* $\mathrm{T}(V)$ *der Translationen bildet eine Untergruppe von* $\mathrm{GA}(V)$. *Weiter gilt*

$$\mathrm{T}(V) = \{f \in \mathrm{GA}(V) \,|\, \vec{f} = \mathrm{id}_V\}.$$

Beweis. Übungsaufgabe. □

Bemerkung. Wir haben einen surjektiven Homomorphismus

$$h_1 : \mathrm{GA}(V) \longrightarrow \mathrm{GL}(V)$$

und einen injektiven Homomorphismus

$$h_0 : \mathrm{T}(V) \longrightarrow \mathrm{GA}(V),$$

und wir stellen mit obigem Lemma 2 fest

$$\mathrm{im}(h_0) = \ker(h_1).$$

Diesen Sachverhalt drückt man kurz durch das Diagramm

$$0 \longrightarrow \mathrm{T}(V) \longrightarrow \mathrm{GA}(V) \longrightarrow \mathrm{GL}(V) \longrightarrow 0$$

aus und nennt dies eine *kurze exakte Sequenz von Gruppen*; $\mathrm{GA}(V)$ wird dabei eine *zentrale Gruppenerweiterung von* $\mathrm{GL}(V)$ *durch* $\mathrm{T}(V)$ genannt. Dies besagt, daß die Elemente von $\mathrm{GA}(V)$ als Paare $(\vec{f}, v)$ $(\vec{f} \in \mathrm{GL}(V), v \in \mathrm{T}(V))$ zu betrachten sind; die Komposition zweier solcher Paare erfolgt dabei nicht komponentenweise, sondern in komplizierterer Weise.

Zentrische Streckungen.

In Verallgemeinerung der vorhergehenden Charakterisierung der Translationen definieren wir die Menge

$$\mathrm{D}(V) := \{f \in \mathrm{GA}(V) \,|\, \exists\ \lambda \in K : \vec{f} = \lambda \cdot \mathrm{id}_V\}.$$

Da $\vec{f}$ bijektiv sein muß, gilt $\lambda \neq 0$.

Lemma 1. *Die Menge* $\mathrm{D}(V)$ *bildet eine Untergruppe von* $\mathrm{GA}(V)$, *welche die Translationen* $\mathrm{T}(V)$ *als Untergruppe enthält.*

Beweis. Übungsaufgabe □

Definition. Wir nennen die Elemente von $\mathrm{D}(V)$ die *Dilatationen von* $\mathbf{A}(V)$.

Lemma 2. *Es sei* $f \in \mathrm{D}(V)\backslash\mathrm{T}(V)$, *d.h.* $\vec{f} = \lambda \cdot \mathrm{id}_V$ *mit* $\lambda \neq 1$. *Dann existiert ein Punkt* $Z \in \mathbf{A}(V)$ *mit*

$$f(P) = Z + \lambda \cdot \overrightarrow{ZP}$$

für alle $P \in \mathbf{A}(V)$; Z *ist dabei eindeutig bestimmt.*

Beweis. Zunächst stellen wir fest, daß der gesuchte Punkt $Z \in \mathbf{A}(V)$ ein Fixpunkt der affinen Abbildung f ist, d.h. $f(Z) = Z$ erfüllt. Haben wir umgekehrt einen Fixpunkt Z von f gefunden, so ergibt sich die behauptete Formel für $f(P)$ ($P \in \mathbf{A}(V)$) wie folgt: Aus

$$P = Z + \overrightarrow{ZP}$$

erhalten wir sofort unter Anwendung von f

$$f(P) = f(Z) + \vec{f}(\overrightarrow{ZP}) = Z + \lambda \cdot \overrightarrow{ZP}\,.$$

Wir haben also einen Fixpunkt Z von f zu finden und dessen Eindeutigkeit nachzuweisen. Dazu setzen wir mit beliebigem $O \in \mathbf{A}(V)$

$$Z := O + \frac{1}{1-\lambda} \cdot \overrightarrow{Of(O)}.$$

Damit berechnen wir

$$\begin{aligned} f(Z) &= f(O) + \vec{f}\left(\tfrac{1}{1-\lambda} \cdot \overrightarrow{Of(O)}\right) \\ &= O + \overrightarrow{Of(O)} + \tfrac{\lambda}{1-\lambda} \cdot \overrightarrow{Of(O)} \\ &= O + \tfrac{1}{1-\lambda} \cdot \overrightarrow{Of(O)} = Z, \end{aligned}$$

d.h. Z ist in der Tat Fixpunkt von f. Für einen beliebigen weiteren Fixpunkt Z' von f erhalten wir zunächst

$$\overrightarrow{f(O)Z'} = \overrightarrow{f(O)f(Z')} = \vec{f}(\overrightarrow{OZ'}) = \lambda \cdot \overrightarrow{OZ'}.$$

Aus der Gleichung

$$\overrightarrow{Of(O)} + \overrightarrow{f(O)Z'} = \overrightarrow{OZ'}$$

ergibt sich weiter

$$\overrightarrow{OZ'} - \overrightarrow{Of(O)} = \lambda \cdot \overrightarrow{OZ'},$$

d.h.

$$\overrightarrow{OZ'} = \frac{1}{1-\lambda} \cdot \overrightarrow{Of(O)}.$$

Damit folgt

$$Z' = O + \overrightarrow{OZ'} = O + \frac{1}{1-\lambda} \cdot \overrightarrow{Of(O)} = Z,$$

was die Eindeutigkeit von Z zeigt. □

Definition. Ist $f \in \mathrm{D}(V)\backslash\mathrm{T}(V)$ mit dem Fixpunkt $Z \in \mathbf{A}(V)$ und $\vec{f} = \lambda \cdot \mathrm{id}_V$, so nennen wir f *zentrische Streckung mit dem Zentrum* Z *und dem Streckungsfaktor* λ.

6.7 Matrizielle Darstellung affiner Abbildungen

Es seien $\mathbf{A}(V)$ ein n-dimensionaler affiner Raum mit affiner Basis $\mathcal{P} = \{P_0, P_1, ..., P_n\}$ und $\mathbf{A}(W)$ ein m-dimensionaler affiner Raum mit affiner Basis $\mathcal{Q} = \{Q_0, Q_1, ..., Q_n\}$. Weiter sei $f \in \mathrm{A}(V, W)$, d.h. $f\colon \mathbf{A}(V) \longrightarrow \mathbf{A}(W)$ eine affine Abbildung; wie bisher bezeichne $\vec{f} : V \longrightarrow W$ die durch f induzierte lineare Abbildung. Bezüglich der Basen

$$\mathcal{B} = \{\overrightarrow{P_0P_1}, ..., \overrightarrow{P_0P_n}\} \text{ von } V\,,\ \mathcal{C} = \{\overrightarrow{Q_0Q_1}, ..., \overrightarrow{Q_0Q_n}\} \text{ von } W$$

wird $\vec{f} \in \mathrm{L}(V, W)$ durch die Matrix

$$A = (\alpha_{k,j})_{\substack{k=1,...,m\\ j=1,...,n}} \in \mathrm{M}_{m,n}(K)$$

dargestellt, wobei die Einträge $\alpha_{k,j} \in K$ eindeutig durch

$$\vec{f}(\overrightarrow{P_0P_j}) = \sum_{k=1}^{m} \alpha_{k,j} \cdot \overrightarrow{Q_0, Q_k} \quad (j = 1, ..., n)$$

bestimmt sind.
Für einen beliebigen Punkt $P \in \mathbf{A}(V)$ mit den affinen Koordinaten $\xi_1, ..., \xi_n$ bezüglich $\mathcal{P}$ sollen nun die affinen Koordinaten $\eta_1, ..., \eta_m$ von $f(P) \in \mathbf{A}(W)$ bezüglich $\mathcal{Q}$ berechnet werden. Dazu benötigen wir neben $A \in \mathrm{M}_{m,n}(K)$ noch die affinen Koordinaten $\eta_1^{(0)}, ..., \eta_m^{(0)}$ von $f(P_0)$ bezüglich $\mathcal{Q}$, d.h. es ist

$$\overrightarrow{Q_0 f(P_0)} = \sum_{k=1}^{m} \eta_k^{(0)} \cdot \overrightarrow{Q_0Q_k}.$$

Mit Hilfe von

$$\overrightarrow{P_0P} = \sum_{j=1}^{n} \xi_j \cdot \overrightarrow{P_0P_j}$$

erhalten wir aufgrund der Affinität von f

$$\begin{aligned}
f(P) &= f(P_0) + \overrightarrow{f(P_0)f(P)} \\
&= f(P_0) + \vec{f}\,(\overrightarrow{P_0P}) \\
&= Q_0 + \overrightarrow{Q_0 f(P_0)} + \vec{f}\left(\sum_{j=1}^{n} \xi_j \cdot \overrightarrow{P_0P_j}\right) \\
&= Q_0 + \sum_{k=1}^{m} \eta_k^{(0)} \cdot \overrightarrow{Q_0Q_k} + \sum_{j=1}^{n} \xi_j \cdot \vec{f}\left(\overrightarrow{P_0P_j}\right) \\
&= Q_0 + \sum_{k=1}^{m} \eta_k^{(0)} \cdot \overrightarrow{Q_0Q_k} + \sum_{j=1}^{n} \xi_j \sum_{k=1}^{m} \alpha_{k,j} \cdot \overrightarrow{Q_0Q_k} \\
&= Q_0 + \sum_{k=1}^{m} \left(\eta_k^{(0)} + \sum_{j=1}^{n} \alpha_{k,j} \cdot \xi_j\right) \cdot \overrightarrow{Q_0Q_k}.
\end{aligned}$$

Damit sind die affinen Koordinaten $\eta_1, ..., \eta_m$ von $f(P)$ bezüglich $\mathcal{Q}$ gegeben durch

$$\eta_k = \eta_k^{(0)} + \sum_{j=1}^{n} \alpha_{k,j} \cdot \xi_j \qquad (k = 1, ..., m).$$

Dieses Resultat läßt sich wie folgt in matrizieller Form wiedergeben: Indem wir die Matrix

$$A_{\text{aff}} = \begin{pmatrix} 1 & 0 & ... & 0 \\ \eta_1^{(0)} & \alpha_{1,1} & ... & \alpha_{1,n} \\ \vdots & \vdots & & \vdots \\ \eta_m^{(0)} & \alpha_{m,1} & ... & \alpha_{m,n} \end{pmatrix} \in \mathrm{M}_{m+1,n+1}(K)$$

einführen, lassen sich die affinen Koordinaten $\eta_1, ..., \eta_m$ von $f(P)$ bezüglich $\mathcal{Q}$ in Termen der affinen Koordinaten $\xi_1, ..., \xi_n$ von P bezüglich $\mathcal{P}$ und der Matrix $A_{\text{aff}} \in \mathrm{M}_{m+1,n+1}(K)$ in der folgenden Art berechnen

$$\begin{pmatrix} 1 \\ \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} = \begin{pmatrix} 1 & 0 & ... & 0 \\ \eta_1^{(0)} & \alpha_{1,1} & ... & \alpha_{1,n} \\ \vdots & \vdots & & \vdots \\ \eta_m^{(0)} & \alpha_{m,1} & ... & \alpha_{m,n} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Definition. Unter Berücksichtigung der vorhergehenden Bezeichnungen nennen wir $A_{\text{aff}} \in \mathrm{M}_{m+1,n+1}(K)$ *die Matrix der affinen Abbildung* $f \in \mathrm{A}(V, W)$ *bezüglich der affinen Basen* $\mathcal{P}, \mathcal{Q}$.

6.8 Der Hauptsatz der affinen Geometrie

Zunächst sei $\mathbf{A}(V)$ ein affiner Raum über einem beliebigen Körper K.

Definiton. Drei Punkte $P, Q, R \in \mathbf{A}(V)$ heißen *kollinear*, falls eine affine Gerade $\mathbf{A}(W) \subseteq \mathbf{A}(V)$ mit $P, Q, R \in \mathbf{A}(W)$ existiert.

Lemma. *Ist* $f \in \mathrm{GA}(V)$ *und sind* $P, Q, R \in \mathbf{A}(V)$ *kollinear, so sind auch die Bildpunkte* $f(P), f(Q), f(R)$ *kollinear.*

Beweis. Nach Voraussetzung existiert eine affine Gerade $\mathbf{A}(W) \subseteq \mathbf{A}(V)$ mit $P, Q, R \in \mathbf{A}(W)$. Da nun die bijektive, affine Abbildung f Geraden auf Geraden abbildet, ist das Bild $f(\mathbf{A}(W))$ ebenfalls eine affine Gerade; diese enthält $f(P), f(Q), f(R)$. Somit sind die Bildpunkte $f(P), f(Q), f(R)$ kollinear. □

Bevor wir den Hauptsatz formulieren und beweisen können, betrachten wir den folgenden

Hilfssatz. Es sei $\sigma : \mathbf{R} \longrightarrow \mathbf{R}$ eine bijektive Abbildung, welche additiv und multiplikativ ist, d.h.

$$\sigma(x + y) = \sigma(x) + \sigma(y),$$

$$\sigma(x \cdot y) = \sigma(x) \cdot \sigma(y)$$

für alle $x, y \in \mathbf{R}$ erfüllt. Dann ist σ die identische Abbildung, d.h. $\sigma(x) = x$ für alle $x \in \mathbf{R}$.

Beweis. Aus der Additivität folgt sofort $\sigma(0) = 0$. Wegen $\sigma(0) = 0$, also $\sigma(1) \neq 0$, folgt aus der Multiplikativität sofort $\sigma(1) = 1$. Aus der Additivität folgt somit

$$\sigma(n) = n$$

für alle $n \in \mathbf{N}$. Wegen $\sigma(0) = 0$ folgt aus der Additivität weiter $\sigma(-x) = -\sigma(x)$ für alle $x \in \mathbf{R}$, also haben wir

$$\sigma(n) = n$$

für alle $n \in \mathbf{Z}$. Es sei nun $r \in \mathbf{Q}$, d.h. $r = p/q$ mit $p, q \in \mathbf{Z}, q \neq 0$. Nach dem Vorhergehenden folgt

$$p = \sigma(p) = \sigma(r \cdot q) = \sigma(r) \cdot \sigma(q) = \sigma(r) \cdot q,$$

d.h.

$$\sigma(r) = p/q = r$$

für alle $r \in \mathbf{Q}$.
Wäre σ stetig, so wäre der Beweis zu Ende. Da wir dies aber nicht wissen, gehen wir wie folgt vor. Wir zeigen zunächst, daß σ monoton wachsend ist. Ist nämlich $x \geq 0$, so findet sich $y \in \mathbf{R}$ mit $x = y^2$, also haben wir

$$\sigma(x) = \sigma(y^2) = \sigma(y)^2 \geq 0.$$

Ist also $a \geq b$, d.h. $a - b \geq 0$, so ist auch $\sigma(a - b) \geq 0$, und wir haben

$$0 \leq \sigma(a - b) = \sigma(a) - \sigma(b),$$

d.h.

$$\sigma(a) \geq \sigma(b).$$

Es sei schließlich $x \in \mathbf{R}$ beliebig. Wir können x durch monotone, rationale Zahlenfolgen $\{r_n\}_{n=1}^{\infty}$ von unten bzw. $\{r'_n\}_{n=1}^{\infty}$ von oben approximieren (z.B. durch Folgen von Dezimalbrüchen), d.h. wir haben die Ungleichungen

$$\ldots \leq r_n \leq r_{n+1} \leq \ldots \leq x \leq \ldots \leq r'_{n+1} \leq r'_n \leq \ldots$$

Nach Anwendung von σ erhalten wir

$$\ldots \leq r_n \leq r_{n+1} \leq \ldots \leq \sigma(x) \leq \ldots \leq r'_{n+1} \leq r'_n \leq \ldots,$$

also ergibt sich die Abschätzung

$$|\, \sigma(x) - x \,| \;\leq\; |\, r'_n - r_n \,|$$

für alle $n \in \mathbf{N}$. Lassen wir n gegen unendlich gehen, so finden wir

$$\sigma(x) = x,$$

wie behauptet. □

Hauptsatz. Es seien $K = \mathbf{R}$ und $\mathbf{A}(V)$ ein reeller affiner Raum der Dimension $n \geq 2$. Ist dann $f : \mathbf{A}(V) \longrightarrow \mathbf{A}(V)$ eine bijektive Abbildung, die je drei kollineare Punkte P, Q, R von $\mathbf{A}(V)$ in drei kollineare Punkte $f(P), f(Q), f(R)$ abbildet, so gilt $f \in \mathrm{GA}(V)$.

Beweis. Im nachfolgenden Beweis, den wir in mehrere Schritte unterteilen, fixieren wir einen affinen Punkt O. Die gegebene bijektive Abbildung $f : \mathbf{A}(V) \longrightarrow \mathbf{A}(V)$ induziert eine Abbildung $\vec{f} : V \longrightarrow V$, so daß das folgende kommutative Diagramm besteht:

Mit anderen Worten haben wir für $P \in \mathbf{A}(V)$

$$\vec{f}(\overrightarrow{OP}) = \overrightarrow{f(O)f(P)}.$$

Wir beachten, daß $\vec{f}$ konstruktionsgemäß eine bijektive Abbildung von V auf V ist. Andererseits wissen wir aber nicht, ob $\vec{f}$ linear ist - dies ist ja im folgenden gerade zu beweisen.

Schritt 1. Sind $A, B, C \in \mathbf{A}(V)$ affin unabhängig, so sind auch $f(A), f(B), f(C) \in \mathbf{A}(V)$ affin unabhängig.

Beweis. Wir nehmen das Gegenteil an. Dann sind $f(A), f(B), f(C)$ affin abhängig und liegen somit auf einer affinen Geraden $\mathbf{A}(W)$; wegen der affinen Unabhängigkeit von A, B, C und der Injektivität von f sind die Bilder $f(A), f(B), f(C)$ paarweise verschieden. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, daß $\mathbf{A}(V)$ gleich der durch A, B, C aufgespannten affinen Ebene $\mathbf{A}(A, B, C)$ ist. Es sei dann $P \in \mathbf{A}(V) = \mathbf{A}(A, B, C)$ ein beliebiger Punkt. Ohne Beschränkung der Allgemeinheit können wir annehmen, daß die affine Gerade von P nach A die affine Gerade von B nach C im Punkte P' schneidet.

Da $f(B), f(C) \in \mathbf{A}(W)$ gilt und B, C, P' kollinear sind, folgt $f(P') \in \mathbf{A}(W)$; da nun auch $f(A), f(P') \in \mathbf{A}(W)$ gilt und A, P', P kollinear sind, folgt $f(P) \in \mathbf{A}(W)$. Da $P \in \mathbf{A}(V)$ beliebig gewählt war, folgt $f(\mathbf{A}(V)) \subseteq \mathbf{A}(W)$. Dies ist aber ein Widerspruch, da f die affine Ebene $\mathbf{A}(V)$ bijektiv auf sich selbst abbildet.

Schritt 2. Ist $\mathbf{A}(W)$ eine affine Gerade in $\mathbf{A}(V)$, so ist auch $f(\mathbf{A}(W))$ eine affine Gerade in $\mathbf{A}(V)$.

Beweis. Es seien A, B zwei verschiedene Punkte auf $\mathbf{A}(W)$, d.h. $\mathbf{A}(W) = \mathbf{A}(A, B)$ ist die affine Gerade durch A, B. Wir setzen $N := \mathbf{A}(f(A), f(B))$, die affine Gerade durch $f(A), f(B)$. Da f injektiv ist, sind die beiden Punkte $f(A), f(B)$ verschieden; somit ist N in der Tat eine affine Gerade. Wir behaupten nun

$$f(\mathbf{A}(W)) = N.$$

Aufgrund der Kollinearitätseigenschaft von f folgt sofort die Inklusion $f(\mathbf{A}(W)) \subseteq N$. Sei umgekehrt $C' \in N$. Wegen der Surjektivität von f gibt es einen Punkt $C \in \mathbf{A}(V)$ mit $C' = f(C)$. Wäre $C \notin \mathbf{A}(A, B)$, so wären die drei Punkte A, B, C affin unabhängig. Nach Schritt 1 wären dann aber auch $f(A), f(B), f(C)$ affin unabhängig, was $C' \in \mathbf{A}(f(A), f(B))$ widerspricht.

Schritt 3. Sind $\mathbf{A}(W), \mathbf{A}(W')$ parallele Geraden in $\mathbf{A}(V)$, so sind auch $f(\mathbf{A}(W)), f(\mathbf{A}(W'))$ parallele Geraden in $\mathbf{A}(V)$.

Beweis. Übungsaufgabe. (Hinweis: Man kann ohne Einschränkung $\mathbf{A}(W) \neq \mathbf{A}(W')$ annehmen. Man betrachte dann die durch $\mathbf{A}(W), \mathbf{A}(W')$ erzeugte affine Ebene und deren Bild unter f. Schließlich verwende man die Schritte 1 und 2.)

Schritt 4. $\vec{f} : V \longrightarrow V$ ist additiv, d.h. $\vec{f}(x + y) = \vec{f}(x) + \vec{f}(y)$ für alle $x, y \in V$.

Beweis. Es seien $x, y \in V$ beliebig vorgelegt. Wir nehmen zunächst an, daß x, y linear unabhängig sind; der Fall, daß x, y linear abhängig sind, wird aus Schritt 5 folgen. Aufgrund der zu Beginn des Beweises gemachten Voraussetzungen gibt es eindeutig bestimmte Punkte $A, B \in \mathbf{A}(V)$ mit der Eigenschaft

$$x = \overrightarrow{OA}, y = \overrightarrow{OB}\ .$$

Wir betrachten nun das durch die drei Punkte O, A, B erzeugte Parallelogramm und die durch die vier Seiten $\overrightarrow{OB}, \overrightarrow{AC}, \overrightarrow{OA}, \overrightarrow{BC}$ definierten Geraden $\mathbf{A}(W_1) \parallel \mathbf{A}(W_1'), \mathbf{A}(W_2) \parallel \mathbf{A}(W_2')$, respektive.

Nach Schritt 2 sind die Bilder

$$f(\mathbf{A}(W_1)), f(\mathbf{A}(W_1')), f(\mathbf{A}(W_2)), f(\mathbf{A}(W_2'))$$

wieder Geraden; nach Schritt 3 gilt

$$f(\mathbf{A}(W_1)) \parallel f(\mathbf{A}(W_1')), f(\mathbf{A}(W_2)) \parallel f(\mathbf{A}(W_2')).$$

Somit ergibt sich das folgende Bild:

Mit Hilfe von

$$x + y = \overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OC}$$

folgt schließlich

$$\begin{aligned} \vec{f}(x+y) &= \vec{f}(\overrightarrow{OC}) = \overrightarrow{f(O)f(C)} = \overrightarrow{f(O)f(A)} + \overrightarrow{f(O)f(B)} \\ &= \vec{f}(\overrightarrow{OA}) + \vec{f}(\overrightarrow{OB}) = \vec{f}(x) + \vec{f}(y). \end{aligned}$$

Schritt 5. $\vec{f} : V \longrightarrow V$ ist homogen, d.h. $\vec{f}(\lambda \cdot x) = \lambda \cdot \vec{f}(x)$ für alle $x \in V, \lambda \in \mathbf{R}$.

Beweis. Es sei $x \in V, x \neq 0$. Wir betrachten die affine Gerade $N \subseteq \mathbf{A}(V)$ mit Aufpunkt O und Richtung $\langle x \rangle$, d.h.

$$N = O + \langle x \rangle.$$

Nach Schritt 2 ist das Bild $f(N) \subseteq \mathbf{A}(V)$ von N wieder eine affine Gerade; sie hat Aufpunkt $f(O)$ und Richtung $\vec{f}(\langle x \rangle)$, d.h.

$$f(N) = f(O) + \vec{f}(\langle x \rangle).$$

Es sei nun $P \in N$, d.h. es gibt ein $\lambda \in \mathbf{R}$ mit $P = O + \lambda \cdot x$. Für den Bildpunkt $f(P) \in f(N)$ haben wir dann

$$f(P) = f(O) + \vec{f}(\lambda \cdot x) = f(O) + \lambda' \cdot \vec{f}(x)$$

mit einem $\lambda' \in \mathbf{R}$, das durch die Vorgabe von $\lambda \in \mathbf{R}$ eindeutig bestimmt ist. Mit Hilfe der wohldefinierten Zuordnung $\lambda \mapsto \lambda'$ erhalten wir somit eine Abbildung

$$\sigma : \mathbf{R} \longrightarrow \mathbf{R},$$

die bijektiv ist, da f und mithin auch $\vec{f}$ bijektiv ist.
Wir zeigen nun, daß σ additiv und multiplikativ ist; nach dem vorausgeschickten Hilfssatz ist dann σ die identische Abbildung von $\mathbf{R}$, und wir haben

$$\vec{f}(\lambda \cdot x) = \lambda \cdot \vec{f}(x)$$

für alle $x \in V, \lambda \in \mathbf{R}$.
Da $n = \dim_{\mathbf{R}} \mathbf{A}(V) \geq 2$ ist, existiert ein von x linear unabhängiger Vektor y. Mit y können wir die Summe

$$\lambda \cdot x + \mu \cdot x = (\lambda + \mu) \cdot x \quad (\lambda, \mu \in \mathbf{R})$$

wie folgt mit Hilfe von Parallelen konstruieren:

Wendet man nun f auf diese Skizze an, so ergibt sich wie im Beweis der Addivität mit Hilfe der Schritte 2 und 3 ein analoges Bild, d.h. wir haben:

Damit erhalten wir die Äquivalenzen

$$\vec{f}(\lambda \cdot x) + \vec{f}(\mu \cdot x) = \vec{f}((\lambda + \mu) \cdot x) \Longleftrightarrow$$
$$\lambda' \cdot \vec{f}(x) + \mu' \cdot \vec{f}(x) = (\lambda + \mu)' \cdot \vec{f}(x) \Longleftrightarrow$$
$$\sigma(\lambda) + \sigma(\mu) = \sigma(\lambda + \mu),$$

was die Addivität von σ beweist.

Wiederum mit y können wir das Produkt

$$\lambda \cdot (\mu \cdot x) = (\lambda \cdot \mu) \cdot x \quad (\lambda, \mu \in \mathbf{R})$$

wie folgt mit Hilfe von Parallelen konstruieren:

Wendet man jetzt f auf diese Skizze an, so ergibt sich mit Hilfe der Schritte 2 und 3 ein analoges Bild, d.h. wir haben:

Aus dem Strahlensatz folgern wir

$$(\lambda \cdot \mu)' = \lambda' \cdot \mu' \iff \sigma(\lambda \cdot \mu) = \sigma(\lambda) \cdot \sigma(\mu),$$

was die Multiplikativität von σ zeigt.

Wir tragen abschließend noch den Additivätsbeweis von $\vec{f}$ nach, falls y linear abhängig von x ist, d.h. $y = \lambda \cdot x$ mit $\lambda \in \mathbf{R}$ gilt. Wir haben

$$\begin{aligned}&\vec{f}(x+y) = \vec{f}(x + \lambda \cdot x) = \vec{f}((1+\lambda) \cdot x) = (1+\lambda) \cdot \vec{f}(x) = \\ &\vec{f}(x) + \lambda \cdot \vec{f}(x) = \vec{f}(x) + \vec{f}(\lambda \cdot x) = \vec{f}(x) + \vec{f}(y).\end{aligned}$$

Damit haben wir die Linearität von $\vec{f}$ und somit die Affinität von f vollständig bewiesen. $\square$

Kapitel 7

Projektive Geometrie

7.1 Einführung

Ausgangspunkt der projektiven Geometrie ist der Wunsch nach einer präzisen Definition des „unendlich fernen Punktes“, wie er z.B. in der Kunst im Rahmen der Perspektivenlehre auftritt. Wir suchen nach einem Raum, der einerseits den uns vertrauten affinen Raum umfaßt, der andererseits aber auch die sogenannten „unendlich fernen Punkte“ enthält. Falls ein solcher Raum existiert, so sollen zwei in der eingebetteten affinen Ebene liegende parallele Geraden im umfassenden Raum genau einen Schnittpunkt haben (wie es das entsprechende perspektivische Bild verspricht). Der in der affinen Geometrie diskutierte Begriff des Parallelismus soll also entfallen.

Die Lösung dieser Problemstellung wird durch die projektive Geometrie gegeben, in die wir nun anhand der entsprechenden 2-dimensionalen reellen Problemstellung einführen wollen. Dazu gehen wir aus vom $\mathbf{R}$-Vektorraum $V = \mathbf{R}^3$ und der affinen Ebene

$$\mathbf{E} = \left\{ \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \in V \;\middle|\; \xi_1, \xi_2 \in \mathbf{R},\ \xi_3 = 1 \right\},$$

d.h. wir haben das Bild auf Seite 132. Neben der $\mathbf{E}$ zugrunde liegenden Menge, die wir ebenfalls mit $\mathbf{E}$ bezeichnen, betrachten wir die Menge

$$\begin{aligned} \mathcal{G} &= \{G \subset V \mid G = \text{Gerade durch } 0,\ G \not\subset (\xi_1, \xi_2)\text{-Ebene}\} \\ &= \{G \subset V \mid G = 1\text{-dimensionaler Unterraum},\ G \not\subset (\xi_1, \xi_2)\text{-Ebene}\}. \end{aligned}$$

Wir erkennen damit die Bijektion

$$\mathbf{E} \xrightarrow{\text{bij.}} \mathcal{G},$$

gegeben durch die Zuordnung

$$P \mapsto G := \langle \overrightarrow{OP} \rangle \quad (P \in \mathbf{E});$$

die Umkehrabbildung ist dabei gegeben durch

$$G \mapsto P := \mathbf{E} \cap G \quad (G \in \mathcal{G}).$$

Vom mengentheoretischen Standpunkt aus können wir also die Mengen $\mathbf{E}$ und $\mathcal{G}$ miteinander identifizieren, was wir fortan tun.

Die affine Struktur von $\mathbf{E}$ wird nach dem Hauptsatz der affinen Geometrie im wesentlichen durch die Menge der affinen Geraden in $\mathbf{E}$ charakterisiert. Wir wollen diese Struktur auf $\mathcal{G}$ übertragen. Dazu setzen wir

$$\mathcal{E} = \{E \subset V \mid E = \text{Ebene durch } 0,\ E \neq (\xi_1, \xi_2)\text{-Ebene}\}.$$

Damit erhalten wir die Bijektion

$$\{\mathbf{G} \subset \mathbf{E} \mid \mathbf{G} = \text{affine Gerade}\} \xrightarrow{\text{bij.}} \mathcal{E},$$

gegeben durch die Zuordnung

$$\mathbf{G} \mapsto E := \langle \overrightarrow{OP}, \overrightarrow{OQ} \rangle \quad (P, Q \in \mathbf{G}; P \neq Q);$$

die entsprechende Umkehrabbildung ist gegeben durch

$$E \mapsto \mathbf{G} := \mathbf{E} \cap E \ \ (E \in \mathcal{E}).$$

Damit haben wir die affine Struktur von $\mathbf{E}$ auf die Menge $\mathcal{G}$ übertragen. Bei diesen Betrachtungen ist allerdings die umständliche Bedingung

$$\mathbf{G} \not\subset (\xi_1, \xi_2)\text{-Ebene}$$

an die Elemente $\mathbf{G}$ von $\mathcal{G}$ störend. Indem wir diese Bedingung einfach vergessen, vergrößern wir die Menge $\mathcal{G}$ und erhalten damit den $\mathcal{G} = \mathbf{E}$ umfassenden Raum

$$\overline{\mathcal{G}} = \{G \subset V \mid G = \text{1-dimensionaler Unterraum}\}.$$

Dies ist die gesuchte reelle projektive Ebene, die wir mit $\mathbf{P}(V)$ bezeichnen. Die Struktur von $\mathbf{P}(V)$ wird charakterisiert durch die Menge der „Geraden"

$$\overline{\mathcal{E}} = \{E \subset V \mid E = \text{2-dimensionaler Unterraum}\}.$$

Die Elemente von $\overline{\mathcal{G}}$, welche durch die 1-dimensionalen Unterräume der (ξ_1, ξ_2)-Ebene gegeben sind, nennen wir die „unendlich fernen Punkte" der affinen Ebene $\mathbf{E} = \mathcal{G}$. Damit erkennen wir zum Schluß dieser Einführung das folgende: Sind $\mathbf{G}_1, \mathbf{G}_2 \subset \mathbf{E}$ zwei parallele Geraden, denen die beiden „Geraden" $E_1, E_2 \in \overline{\mathcal{E}}$ entsprechen, so ist der Durchschnitt $E_1 \cap E_2$ eine Gerade der (ξ_1, ξ_2)-Ebene und definiert damit ein Element von $\overline{\mathcal{G}}$, also einen „unendlich fernen Punkt".

7.2 Der projektive Raum

Der einführende Abschnitt legt die folgende Definition nahe.

Definiton. Es sei V ein K-Vektorraum. *Der projektive Raum* $\mathbf{P}(V)$ *zum* K*-Vektorraum* V ist gegeben durch

$$\mathbf{P}(V) = \{G \subset V \mid G = \text{1-dimensionaler Unterraum}\}$$

Ist $\dim_K V = n \geq 1$, so setzen wir

$$\dim_K \mathbf{P}(V) := n - 1$$

und nennen dies die *Dimension von* $\mathbf{P}(V)$. Speziell für $V = K^n, n \geq 1$, setzen wir

$$\mathbf{P}^{n-1}(K) := \mathbf{P}(K^n).$$

Bemerkung. Indem wir

$$\mathbf{P}(V) = \left\{ \, \langle x \rangle \mid x \in V \setminus \{0\} \, \right\}$$

schreiben, erhalten wir die surjektive Abbildung

$$\pi_V : V \setminus \{0\} \longrightarrow \mathbf{P}(V),$$

gegeben durch

$$\pi_V(x) = \langle x \rangle \quad (x \in V \setminus \{0\}).$$

Sind $x, y \in V \setminus \{0\}$, so erkennen wir sofort die Äquivalenz

$$\pi_V(x) = \pi_V(y) \Longleftrightarrow \exists \lambda \in K^* : y = \lambda \cdot x.$$

Aus diesem Grund können wir auch schreiben

$$\mathbf{P}(V) = (V \setminus \{0\})/K^*.$$

Bezeichnungen. Es sei $\mathbf{P}(V)$ der projektive Raum zum K-Vektorraum V. Die Elemente von $\mathbf{P}(V)$ nennen wir *(projektive) Punkte* und bezeichnen sie durch $P, Q, R, \ldots$ Ist $P \in \mathbf{P}(V)$, so gilt $P = \langle x \rangle$, wobei der Vektor $x \in V \setminus \{0\}$ bis auf Streckung (mit einem von Null verschiedenen Skalar) eindeutig bestimmt ist.

7.3 Projektive Unterräume

Es sei $\mathbf{P}(V)$ ein projektiver Raum zum K-Vektorraum V.

Definition. Eine Teilmenge $N \subseteq \mathbf{P}(V)$ heißt ein *projektiver Unterraum von* $\mathbf{P}(V)$, falls ein Unterraum $W \subseteq V$ derart existiert, daß $N = \mathbf{P}(W)$ gilt.

Beispiel. Es sei $V = \mathbf{R}^3$ und $W \subseteq V$ ein 2-dimensionaler Unterraum, d.h. eine Ebene durch $0 \in V$. Dann ist $\mathbf{P}(W) \subseteq \mathbf{P}(V)$ eine (projektive) Gerade der projektiven Ebene $\mathbf{P}^2(\mathbf{R}) = \mathbf{P}(V)$. Man vergleiche hierzu wiederum den einleitenden ersten Abschnitt zu diesem Kapitel.

Definition. Es seien $\mathbf{P}(V)$ ein projektiver Raum zum K-Vektorraum V und $\mathbf{P}(W), W \subseteq V$, ein projektiver Unterraum.

- Gilt $\dim_K \mathbf{P}(W) = -1$ (d.h. $\dim_K W = 0$), so ist $\mathbf{P}(W) = \emptyset$.
- Gilt $\dim_K \mathbf{P}(W) = 0$ (d.h. $\dim_K W = 1$), so nennen wir $\mathbf{P}(W)$ einen *(projektiven) Punkt von* $\mathbf{P}(V)$.
- Gilt $\dim_K \mathbf{P}(W) = 1$ (d.h. $\dim_K W = 2$), so nennen wir $\mathbf{P}(W)$ eine *(projektive) Gerade von* $\mathbf{P}(V)$.
- Gilt $\dim_K \mathbf{P}(W) = 2$ (d.h. $\dim_K W = 3$), so nennen wir $\mathbf{P}(W)$ eine *(projektive) Ebene von* $\mathbf{P}(V)$.
- Gilt $\operatorname{codim}_K \mathbf{P}(W) = \dim_K(V/W) = 1$, so nennen wir $\mathbf{P}(W)$ eine *(projektive) Hyperebene von* $\mathbf{P}(V)$.

Definition. Es seien $\mathbf{P}(W_1), \mathbf{P}(W_2)$ projektive Unterräume eines projektiven Raumes $\mathbf{P}(V)$ zum K-Vektorraum V. Dann wird der kleinste projektive Raum, der $\mathbf{P}(W_1)$ und $\mathbf{P}(W_2)$ enthält, mit $\mathbf{P}(W_1, W_2)$ bezeichnet.

Satz. *Es seien $\mathbf{P}(W_1), \mathbf{P}(W_2)$ projektive Unterräume eines endlich dimensionalen projektiven Raums $\mathbf{P}(V)$. Dann besteht die Dimensionsformel:*

$$\dim_K \mathbf{P}(W_1, W_2) = \dim_K \mathbf{P}(W_1) + \dim_K \mathbf{P}(W_2) - \dim_K(\mathbf{P}(W_1) \cap \mathbf{P}(W_2)).$$

Dem Beweis schicken wir zwei Hilfssätze voraus.

Hilfssatz 1. *Mit den Voraussetzungen des Satzes gilt die Gleichheit*

$$\mathbf{P}(W_1, W_2) = \mathbf{P}(W_1 + W_2).$$

Beweis. Definitionsgemäß bestehen die Äquivalenzen

$$\begin{aligned}\langle x\rangle \in \mathbf{P}(W_1, W_2) &\iff x \in W_1 \setminus \{0\} \text{ oder } x \in W_2 \setminus \{0\}\\ &\iff x \in (W_1 + W_2) \setminus \{0\}\\ &\iff \langle x\rangle \in \mathbf{P}(W_1 + W_2),\end{aligned}$$

d.h.

$$\mathbf{P}(W_1, W_2) = \mathbf{P}(W_1 + W_2).$$

□

Hilfssatz 2. *Mit den Voraussetzungen des Satzes gilt die Gleichheit*

$$\mathbf{P}(W_1) \cap \mathbf{P}(W_2) = \mathbf{P}(W_1 \cap W_2).$$

Beweis. Definitionsgemäß bestehen die Äquivalenzen

$$\begin{aligned}\langle x\rangle \in \mathbf{P}(W_1) \cap \mathbf{P}(W_2) &\iff x \in W_1 \setminus \{0\} \text{ und } x \in W_2 \setminus \{0\}\\ &\iff x \in W_1 \cap W_2 \setminus \{0\}\\ &\iff \langle x\rangle \in \mathbf{P}(W_1 \cap W_2),\end{aligned}$$

d.h.

$$\mathbf{P}(W_1) \cap \mathbf{P}(W_2) = \mathbf{P}(W_1 \cap W_2).$$

□

Wir kehren nun zum *Beweis des Satzes* zurück: Durch Anwendung der beiden Hilfssätze, des Dimensionssatzes für Unterräume und der Definition der Dimension projektiver Räume ergibt sich

$$\begin{aligned}\dim_K \mathbf{P}(W_1, W_2) &= \dim_K \mathbf{P}(W_1 + W_2) = \dim_K(W_1 + W_2) - 1\\ &= \dim_K W_1 + \dim_K W_2 - \dim_K(W_1 \cap W_2) - 1\\ &= \dim_K W_1 - 1 + \dim_K W_2 - 1 - (\dim_K(W_1 \cap W_2) - 1)\\ &= \dim_K \mathbf{P}(W_1) + \dim_K \mathbf{P}(W_2) - \dim_K \mathbf{P}(W_1 \cap W_2)\\ &= \dim_K \mathbf{P}(W_1) + \dim_K \mathbf{P}(W_2) - \dim_K(\mathbf{P}(W_1) \cap \mathbf{P}(W_2)),\end{aligned}$$

wie behauptet. □

Corollar. *Es sei* $\mathbf{P}(V)$ *ein projektiver Raum zum* K*-Vektorraum* V *mit* $\dim_K \mathbf{P}(V) = n \geq 0$. *Dann bestehen die drei folgenden Aussagen:*

(i) *Sind* $\mathbf{P}(W_1), \mathbf{P}(W_2)$ *zwei projektive Unterräume von* $\mathbf{P}(V)$ *mit der Eigenschaft* $\dim_K \mathbf{P}(W_1) + \dim_K \mathbf{P}(W_2) \geq n$*, so gilt* $\mathbf{P}(W_1) \cap \mathbf{P}(W_2) \neq \emptyset$*.*

(ii) *Es seien* $\mathbf{P}(W) \subset \mathbf{P}(V)$ *eine projektive Hyperebene und* $P \in \mathbf{P}(V)$ *ein projektiver Punkt mit* $P \notin \mathbf{P}(W)$*. Dann schneidet jede projektive Gerade durch* P *die projektive Hyperebene* $\mathbf{P}(W)$ *in genau einem projektiven Punkt.*

(iii) *In* $\mathbf{P}(V)$ *haben* n *Hyperebenen mindestens einen projektiven Punkt gemeinsam.*

Beweis.

(i) Aufgrund der Ungleichungen

$$\dim_K \mathbf{P}(W_1) + \dim_K \mathbf{P}(W_2) \geq n\,,\ \dim_K \mathbf{P}(W_1, W_2) \leq \dim_K \mathbf{P}(V) = n$$

folgt mit Hilfe des vorhergehenden Satzes

$$\begin{aligned} \dim_K(\mathbf{P}(W_1) \cap \mathbf{P}(W_2)) &= \dim_K \mathbf{P}(W_1) + \dim_K \mathbf{P}(W_2) - \dim_K \mathbf{P}(W_1, W_2) \\ &\geq n - \dim_K \mathbf{P}(W_1, W_2) \geq 0, \end{aligned}$$

d.h.

$$\mathbf{P}(W_1) \cap \mathbf{P}(W_2) \neq \emptyset.$$

(ii) Es sei $\mathbf{P}(E)$, $E \subseteq V$, eine projektive Gerade durch P. Wegen $P \notin \mathbf{P}(W)$ gilt $\mathbf{P}(E) \not\subseteq \mathbf{P}(W)$, d.h. $E \not\subseteq W$; aus Dimensionsgründen gilt dann $W + E = V$. Somit folgt

$$\mathbf{P}(W, E) = \mathbf{P}(W + E) = \mathbf{P}(V).$$

Mit Hilfe des vorhergehenden Satzes ergibt sich schließlich

$$\begin{aligned} \dim_K(\mathbf{P}(W) \cap \mathbf{P}(E)) &= \dim_K \mathbf{P}(W) + \dim_K \mathbf{P}(E) - \dim_K \mathbf{P}(W, E) \\ &= (n-1) + 1 - n = 0, \end{aligned}$$

d.h. der Durchschnitt $\mathbf{P}(W) \cap \mathbf{P}(E)$ besteht aus genau einem Punkt.

(iii) Es seien $\mathbf{P}(W_1), \ldots, \mathbf{P}(W_n) \subset \mathbf{P}(V)$ n projektive Hyperebenen; abkürzungshalber setzen wir

$$\mathbf{P}(W) := \mathbf{P}(W_1) \cap \ldots \cap \mathbf{P}(W_{n-1}).$$

Mit Hilfe des vorhergehenden Satzes erhalten wir

$$\begin{aligned} &\dim_K(\mathbf{P}(W_1) \cap \ldots \cap \mathbf{P}(W_n)) \ = \ \dim_K(\mathbf{P}(W) \cap \mathbf{P}(W_n)) = \\ &\dim_K \mathbf{P}(W) + \dim_K \mathbf{P}(W_n) - \dim_K \mathbf{P}(W, W_n) \geq \\ &\dim_K \mathbf{P}(W) + (n-1) - n \ = \ \dim_K(\mathbf{P}(W_1) \cap \ldots \cap \mathbf{P}(W_{n-1})) - 1. \end{aligned}$$

Nach $(n-2)$-maliger Iteration ergibt sich damit

$$\begin{aligned}&\dim_K(\mathbf{P}(W_1)\cap\ldots\cap\mathbf{P}(W_n)) \;\geq\; \dim_K(\mathbf{P}(W_1)\cap\mathbf{P}(W_2))-(n-2)=\\&\dim_K\mathbf{P}(W_1)+\dim_K\mathbf{P}(W_2)-\dim_K\mathbf{P}(W_1,W_2)-(n-2)\geq\\&(n-1)+(n-1)-n-(n-2) \;=\; 0,\end{aligned}$$

d.h.

$$\mathbf{P}(W_1)\cap\ldots\cap\mathbf{P}(W_n)\neq\emptyset.$$

□

7.4 Projektive Basen, homogene Koordinaten

Es sei $\mathbf{P}(V)$ ein projektiver Raum zum K-Vektorraum V.

Definition. $(n+1)$ Punkte $P_0=\langle x_0\rangle,\ldots,P_n=\langle x_n\rangle\in\mathbf{P}(V)$ heißen *projektiv unabhängig*, falls die $(n+1)$ Vektoren

$$x_0,\ldots,x_n\in V$$

linear unabhängig sind. Gilt $\dim_K\mathbf{P}(V)=n$, d.h. $\dim_K V=n+1$, so werden $(n+2)$ Punkte $\mathcal{P}=\{P_0,\ldots,P_{n+1}\}\subseteq\mathbf{P}(V)$ eine *projektive Basis von* $\mathbf{P}(V)$ genannt, falls je $(n+1)$ dieser Punkte projektiv unabhängig sind.

Lemma. *Es sei $\mathbf{P}(V)$ ein projektiver Raum zum K-Vektorraum V mit projektiver Basis $\mathcal{P}=\{P_0,\ldots,P_{n+1}\}$. Dann gibt es $(n+2)$ Vektoren $v_0,\ldots,v_{n+1}\in V$ mit den Eigenschaften*

(a) $P_j=\langle v_j\rangle\quad(j=0,\ldots,n+1)$,

(b) $\{v_0,\ldots,v_n\}$ *ist Basis von* V,

(c) $v_0+\ldots+v_n=v_{n+1}$,

(d) *Erfüllen die Vektoren $w_0,\ldots,w_{n+1}\in V$ ebenfalls* (a) - (c), *so existiert $\lambda\in K^*$ mit* $w_j=\lambda\cdot v_j$ $(j=0,\ldots,n+1)$.

Beweis. Nach Voraussetzung gilt $\dim_K\mathbf{P}(V)=n$ und somit $\dim_K V=n+1$. Es sei $P_j=\langle x_j\rangle$ mit $x_j\in V\setminus\{0\}$ $(j=0,\ldots,n+1)$. Da insbesondere die Vektoren $x_0,\ldots,x_n$ linear unabhängig sind, bilden diese eine Basis von V. Damit finden sich eindeutig bestimmte Skalare $\mu_0,\ldots,\mu_n\in K$ mit

$$x_{n+1}=\sum_{j=0}^{n}\mu_j\cdot x_j.$$

Nun gilt aber $\mu_j\neq0$ für $j=0,\ldots,n$, da andernfalls die Bedingung, daß je $(n+1)$ der $(n+2)$ Vektoren $x_0,\ldots,x_{n+1}\in V$ linear unabhängig sind, verletzt wäre. Damit liegen die Vektoren

$$v_0:=\mu_0x_0,\quad\ldots,\quad v_n:=\mu_nx_n,\quad v_{n+1}:=x_{n+1}$$

alle in $V\setminus\{0\}$ und wir stellen dazu fest:

(a) $P_j = \langle v_j \rangle \quad (j = 0, \dots, n+1)$,

(b) $\{v_0, \dots, v_n\}$ ist Basis von V,

(c) $v_0 + \dots + v_n = v_{n+1}$.

Es bleibt also noch die Eigenschaft (d) nachzuweisen. Dazu seien $w_0, \dots, w_{n+1} \in V$ Vektoren, die ebenfalls (a) - (c) erfüllen. Aus (a) folgt dann wegen $P_j = \langle v_j \rangle = \langle w_j \rangle$ die Gleichung $w_j = \lambda_j \cdot v_j$ mit $\lambda_j \in K^*$ $(j = 0, \dots, n+1)$. Aus (c) ergibt sich weiter

$$\begin{aligned} w_0 + \dots + w_n = w_{n+1} &\iff \lambda_0 v_0 + \dots + \lambda_n v_n = \lambda_{n+1} v_{n+1} \\ &\iff (\lambda_0 - \lambda_{n+1}) v_0 + \dots + (\lambda_n - \lambda_{n+1}) v_n = 0. \end{aligned}$$

Aus (b) folgt schließlich

$$\lambda_j = \lambda_{n+1} =: \lambda \ \ (j = 0, \dots, n)$$

und somit

$$w_j = \lambda v_j \ \ (j = 0, \dots, n+1).$$

□

Bemerkung. Ist $\mathbf{P}(V)$ ein projektiver Raum zum K-Vektorraum V mit projektiver Basis $\mathcal{P} = \{P_0, \dots, P_{n+1}\}$, so existieren nach dem vorhergehenden Lemma Vektoren $v_0, \dots, v_{n+1} \in V$, welche insbesondere den folgenden Eigenschaften

(a) $P_j = \langle v_j \rangle \quad (j = 0, \dots, n+1)$,

(b) $\{v_0, \dots, v_n\}$ ist Basis von V,

(c) $v_0 + \dots + v_n = v_{n+1}$

genügen; dabei sind diese Vektoren bis auf Multiplikation mit einem Skalar $\lambda \in K^*$ eindeutig festgelegt.

Ist nun $P \in \mathbf{P}(V)$ ein beliebiger Punkt mit $P = \langle x \rangle, x \in V \setminus \{0\}$ (dabei ist x nur bis auf Multiplikation mit einem Skalar ungleich Null eindeutig festgelegt), so finden sich Skalare $\xi_0, ..., \xi_n \in K$ mit

$$x = \sum_{j=0}^{n} \xi_j \cdot v_j.$$

Da sowohl x als auch die Vektoren $v_0, ..., v_n$ nur bis auf Multiplikation mit einem Skalar ungleich Null eindeutig gegeben sind, sind auch die Größen $\xi_0, ..., \xi_n \in K$ nur bis auf Multiplikation mit einem Skalar ungleich Null eindeutig bestimmt.

Definition. Mit den vorhergehenden Bezeichnungen nennen wir die (bis auf Multiplikation mit einem Skalar ungleich Null eindeutig bestimmten) Skalare $\xi_0, ..., \xi_n$ *die homogenen Koordinaten von $P \in \mathbf{P}(V)$ bezüglich der projektiven Basis $\mathcal{P} = \{P_0, ..., P_{n+1}\}$*. Wir notieren die homogenen Koordinaten von $P \in \mathbf{P}(V)$ in der Form $[\xi_0 : ... : \xi_n]$ und beachten, daß damit ein Punkt in $\mathbf{P}^n(K)$ definiert wird. Mit Hilfe der homogenen Koordinaten erhalten wir also eine bijektive Abbildung

$$\mathbf{P}(V) \longrightarrow \mathbf{P}^n(K)$$

gegeben durch die Zuordnung

$$P \mapsto [\xi_0 : ... : \xi_n].$$

7.5 Projektive Abbildungen

Es sei $\mathbf{P}(V)$ bzw. $\mathbf{P}(W)$ ein projektiver Raum zum K-Vektorraum V bzw. W und $\pi_V : V \setminus \{0\} \longrightarrow \mathbf{P}(V)$ bzw. $\pi_W : W \setminus \{0\} \longrightarrow \mathbf{P}(W)$ die entsprechende Projektionsabbildung.

Ist $\vec{f} \in \mathrm{L}(V, W)$ eine lineare Abbildung, so könnte man sich durch die Zuordnung

$$\langle x \rangle \mapsto \langle \vec{f}(x) \rangle \quad (\langle x \rangle \in \mathbf{P}(V), \text{ d.h. } x \in V \setminus \{0\})$$

eine Abbildung $f : \mathbf{P}(V) \longrightarrow \mathbf{P}(W)$ erklärt denken. Dabei ist allerdings zu bedenken, daß $\langle \vec{f}(x) \rangle$ nur dann ein Element von $\mathbf{P}(W)$ definiert, wenn $\vec{f}(x) \in W \setminus \{0\}$, d.h. $x \notin \ker(\vec{f})$ ist. Diese Einschränkung an $x \in V \setminus \{0\}$ können wir ausschließen, wenn wir annehmen, daß $\vec{f} \in \mathrm{L}(V, W)$ *injektiv* ist; dann gilt für alle $x \in V \setminus \{0\}$ jeweils auch $\vec{f}(x) \in W \setminus \{0\}$.

Definition. Eine Abbildung $f : \mathbf{P}(V) \longrightarrow \mathbf{P}(W)$ heißt eine *projektive Abbildung*, falls eine injektive lineare Abbildung $\vec{f} : V \longrightarrow W$ existiert, so daß für alle $P = \langle x \rangle \in \mathbf{P}(V)$ die Gleichung

$$f(P) = \langle \vec{f}(x) \rangle$$

besteht. Um auf die Abhängigkeit der projektiven Abbildung f von $\vec{f}$ hinzuweisen, schreiben wir kurz $f = \mathbf{P}(\vec{f})$.

Bemerkung. Die vorhergehende Definition besitzt die folgende, äquivalente Fassung: Eine Abbildung $f : \mathbf{P}(V) \longrightarrow \mathbf{P}(W)$ heißt eine projektive Abbildung, falls eine injektive lineare Abbildung $\vec{f} : V \longrightarrow W$ existiert, so daß das folgende Diagramm kommutiert:

Bemerkung. Wird die Injektivitätsvoraussetzung an $\vec{f}$ fallengelassen, so erhält man mit $f = \mathbf{P}(\vec{f})$ die Abbildung

$$f : \mathbf{P}(V) \setminus \mathbf{P}(\ker(\vec{f})) \longrightarrow \mathbf{P}(W).$$

Der Definitionsbereich von f muß also von $\mathbf{P}(V)$ zu $\mathbf{P}(V) \setminus \mathbf{P}(\ker(\vec{f}))$ eingeschränkt werden; die Menge der Punkte in $\mathbf{P}(\ker(\vec{f}))$ wird *die Menge der Basispunkte von f* genannt.

Bezeichnungen. Es seien $\mathbf{P}(V)$ und $\mathbf{P}(W)$ projektive Räume. Damit setzen wir

$$\begin{aligned} \mathrm{P}(V,W) &:= \{f : \mathbf{P}(V) \longrightarrow \mathbf{P}(W) \mid f \text{ projektiv}\}, \\ \mathrm{P}(V) &:= \mathrm{P}(V,V), \\ \mathrm{PGL}(V) &:= \{f \in \mathrm{P}(V) \mid f \text{ bijektiv}\} \\ &= \{f \in \mathrm{P}(V) \mid \vec{f} \in \mathrm{GL}(V)\}. \end{aligned}$$

Lemma. Die Menge $\mathrm{PGL}(V)$ bildet eine Gruppe bezüglich der Komposition von Abbildungen.

Beweis. Übungsaufgabe. □

Definition. Die Gruppe $\mathrm{PGL}(V)$ wird *die projektive lineare Gruppe zum K-Vektorraum* V genannt. Die Elemente von $\mathrm{PGL}(V)$ werden auch *Homographien* genannt.

Satz. *Es seien $\mathcal{P} = \{P_0, \ldots, P_{n+1}\}$ und $\mathcal{P}' = \{P'_0, \ldots, P'_{n+1}\}$ projektive Basen des n-dimensionalen projektiven Raumes $\mathbf{P}(V)$. Dann existiert genau eine Homographie $f \in \mathrm{PGL}(V)$ mit der Eigenschaft*

$$P'_j = f(P_j) \quad (j = 0, \ldots, n+1).$$

Beweis. Wir beweisen zunächst die Existenz der gesuchten Homographie. Dazu wählen wir Vektoren

$$v_0, ..., v_{n+1} \in V; \quad v'_0, ..., v'_{n+1} \in V$$

mit den Eigenschaften

(a) $P_j = \langle v_j \rangle$ $(j = 0, \ldots, n+1)$, $P'_j = \langle v'_j \rangle$ $(j = 0, \ldots, n+1)$,

(b) $\{v_0, \ldots, v_n\}$ Basis von V, $\{v'_0, \ldots, v'_n\}$ Basis von V,

(c) $v_0 + \ldots + v_n = v_{n+1}$; $v'_0 + \ldots + v'_n = v'_{n+1}$.

Diese Vektoren existieren nach dem Lemma über projektive Basen. Nach der linearen Algebra existiert eine lineare Abbildung $\vec{f} \in \mathrm{GL}(V)$ mit der Eigenschaft

$$\vec{f}(v_j) = v'_j \quad (j = 0, \ldots, n).$$

Die Linearität von $\vec{f}$ zeigt sofort, daß auch

$$\vec{f}(v_{n+1}) = v'_{n+1}$$

gilt. Mit $f = \mathbf{P}(\vec{f})$ erhalten wir jetzt die gesuchte Homographie, welche

$$f(P_j) = \left\langle \vec{f}(v_j) \right\rangle = \langle v'_j \rangle = P'_j \quad (j = 0, \ldots, n+1)$$

erfüllt.
Wir kommen nun zum Beweis der Eindeutigkeit. Dazu gehen wir aus von einer beliebigen Homographie $g = \mathbf{P}(\vec{g})$, welche $P_j = g(P_j)$ $(j = 0, ..., n+1)$ erfüllt, und vergleichen diese mit der im Existenzbeweis konstruierten Homographie $f = \mathbf{P}(\vec{f})$. Wegen $g(P_j) = P_j'$, d.h. $g(\langle v_j \rangle) = \langle v_j' \rangle$, existieren Skalare $\lambda_j \in K^*$ mit der Eigenschaft

$$\vec{g}(v_j) = \lambda_j \cdot v_j' \quad (j = 0, ..., n+1).$$

Durch Einsetzen von (c) in die Relation $\vec{g}(v_{n+1}) = \lambda_{n+1} \cdot v_{n+1}$ ergibt sich dann

$$\vec{g}\left(\sum_{j=0}^{n} v_j\right) = \lambda_{n+1} \cdot \sum_{j=0}^{n} v_j' \iff \sum_{j=0}^{n} \lambda_j \cdot v_j' = \sum_{j=0}^{n} \lambda_{n+1} \cdot v_j'.$$

Aus der linearen Unabhängigkeit der Vektoren $v_0', \ldots, v_n'$ ergibt sich dann sofort

$$\lambda_0 = \lambda_1 = \ldots = \lambda_n = \lambda_{n+1}.$$

Damit folgt $\vec{g} = \lambda_0 \cdot \vec{f}$, woraus wir

$$g = \mathbf{P}(\vec{g}) = \mathbf{P}(\vec{f}\,) = f$$

erhalten. □

7.6 Matrizielle Darstellung projektiver Abbildungen

Wir beginnen mit der folgenden

Proposition. *Es seien* $\mathbf{P}(V), \mathbf{P}(W)$ *projektive Räume zu den* K*-Vektorräumen* V, W *und* $\vec{f}, \vec{g} \in \mathrm{L}(V, W)$ *injektive lineare Abbildungen mit der Eigenschaft* $\mathbf{P}(\vec{f}\,) = \mathbf{P}(\vec{g}) \in \mathrm{P}(V, W)$. *Dann existiert* $\lambda \in K^*$ *mit* $\vec{g} = \lambda \cdot \vec{f}$.

Beweis. Für $\langle x \rangle \in \mathbf{P}(V)$, d.h. $x \in V \setminus \{0\}$, bestehen die Gleichheiten

$$\begin{aligned} \mathbf{P}(\vec{f})\big(\langle x \rangle\big) = \mathbf{P}(\vec{g})\big(\langle x \rangle\big) &\iff \big\langle \vec{f}(x) \big\rangle = \big\langle \vec{g}(x) \big\rangle \\ &\iff \exists \lambda = \lambda(x) \in K^* : \vec{g}(x) = \lambda(x) \cdot \vec{f}(x). \end{aligned}$$

Wir werden zeigen, daß die Funktion $\lambda(x)$ konstant ist. Dazu fixieren wir $x_0 \in V \setminus \{0\}$ und setzen $\lambda_0 := \lambda(x_0)$. Aufgrund der Linearität von $\vec{f}, \vec{g}$ folgt zunächst

$$\lambda(\mu \cdot x_0) = \lambda(x_0) = \lambda_0$$

für alle $\mu \in K^*$. Ist $\dim_K V = 1$, so sind wir fertig. Andernfalls sei $x \in V \setminus \{0\}$ ein beliebiger, von x_0 linear unabhängiger Vektor. Dann betrachten wir die äquivalenten Gleichungen

$$\begin{aligned}
\vec{g}(x_0 + x) = \lambda(x_0 + x) \cdot \vec{f}(x_0 + x) \iff & \vec{g}(x_0) + \vec{g}(x) \\
& = \lambda(x_0 + x) \cdot \vec{f}(x_0) + \lambda(x_0 + x) \cdot \vec{f}(x) \\
\iff & \lambda_0 \cdot \vec{f}(x_0) + \lambda(x) \cdot \vec{f}(x) \\
& = \lambda(x_0 + x) \cdot \vec{f}(x_0) + \lambda(x_0 + x) \cdot \vec{f}(x) \\
\iff & \big(\lambda_0 - \lambda(x_0 + x)\big) \cdot \vec{f}(x_0) \\
& + (\lambda(x) - \lambda(x_0 + x)) \cdot \vec{f}(x) = 0.
\end{aligned}$$

Da nun aufgrund der Injektivität von $\vec{f}$ mit x_0, x auch $\vec{f}(x_0), \vec{f}(x)$ linear unabhängig sind, folgt

$$\lambda(x) = \lambda(x_0 + x) = \lambda_0.$$

Damit ist die Proposition bewiesen. □

Wir kommen nun zur matriziellen Beschreibung projektiver Abbildungen. Dazu seien $\mathbf{P}(V)$ ein n-dimensionaler projektiver Raum mit projektiver Basis $\mathcal{P} = \{P_0, \ldots, P_{n+1}\}$ und $\mathbf{P}(W)$ein m-dimensionaler projektiver Raum mit projektiver Basis $\mathcal{Q} = \{Q_0, \ldots, Q_{m+1}\}$. Damit existieren Vektoren

$$\begin{aligned}
& v_0, \ldots, v_{n+1} \in V \text{ mit } P_j = \langle v_j \rangle \quad (j = 0, \ldots, n+1), \\
& \mathcal{B} = \{v_0, \ldots, v_n\} \text{ Basis von } V, \\
& v_{n+1} = \sum_{j=0}^{n} v_j \, ; \\
& \\
& w_0, \ldots, w_{m+1} \in W \text{ mit } Q_j = \langle w_j \rangle \quad (j = 0, \ldots, m+1), \\
& \mathcal{C} = \{w_0, \ldots, w_m\} \text{ Basis von } W, \\
& w_{m+1} = \sum_{j=0}^{m} w_j .
\end{aligned}$$

Ist nun $f \in \mathrm{P}(V, W)$ eine projektive Abbildung von $\mathbf{P}(V)$ nach $\mathbf{P}(W)$, so gilt $f = \mathbf{P}(\vec{f})$ mit einer injektiven linearen Abbildung $\vec{f} \in \mathrm{L}(V, W)$. Nach der vorhergehenden Proposition ist $\vec{f}$ durch f bis auf Multiplikation mit einem von Null verschiedenen Skalar eindeutig bestimmt. Es sei nun

$$A = (\alpha_{k,j})_{\substack{k=1,\ldots,m \\ j=1,\ldots,n}} \in M_{m+1,n+1}(K)$$

die Matrix von $\vec{f}$ bezüglich $\mathcal{B}, \mathcal{C}$; aufgrund der getroffenen Wahlen ist die Matrix A nur bis auf Multiplikation mit einem von Null verschiedenen Skalar bestimmt.

Definition. Unter Berücksichtigung der vorhergehenden Bezeichnungen nennen wir

$$A_{\mathrm{proj}} := A \text{ modulo Multiplikation mit } K^*$$

die Matrix der projektiven Abbildung $f : \mathbf{P}(V) \longrightarrow \mathbf{P}(W)$ *bezüglich der projektiven Basen* $\mathcal{P}, \mathcal{Q}$.

Bemerkung. Sind $[\xi_0 : \ldots : \xi_n]$ die homogenen Koordinaten von $P \in \mathbf{P}(V)$ bezüglich $\mathcal{P}$, so erhält man die homogenen Koordinaten $[\eta_0 : \ldots : \eta_m]$ von $f(P) \in \mathbf{P}(W)$ bezüglich $\mathcal{Q}$ durch

$$\begin{pmatrix} \eta_0 \\ \vdots \\ \eta_m \end{pmatrix} = A \cdot \begin{pmatrix} \xi_0 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Wir schließen dieses Kapitel mit der Formulierung des Hauptsatzes der projektiven Geometrie, den wir aber nicht beweisen wollen.

Hauptsatz der projektiven Geometrie. *Es seien* $K = \mathbf{R}$ *und* $\mathbf{P}(V)$ *ein projektiver Raum der Dimension* $n \geq 2$. *Ist dann* $f : \mathbf{P}(V) \longrightarrow \mathbf{P}(V)$ *eine bijektive Abbildung, die je drei kollineare Punkte* P, Q, R *von* $\mathbf{P}(V)$ *in drei kollineare Punkte* $f(P), f(Q), f(R)$ *abbildet, so gilt* $f \in \mathrm{PGL}(V)$.

Kapitel 8

Quadriken

8.1 Quadratische Formen

Im gesamten Kapitel arbeiten wir über dem Körper $K = \mathbf{R}$.

Definition. Es sei $A = (\alpha_{j,k})_{j,k=1\ldots n} \in \mathrm{M}_n(\mathbf{R})$ eine symmetrische Matrix. Der Ausdruck

$$\begin{aligned} Q_A(\xi_1, \ldots, \xi_n) &= (\xi_1, \ldots, \xi_n) \cdot A \cdot {}^t(\xi_1, \ldots, \xi_n) \\ &= \sum_{j,k=1}^{n} \alpha_{j,k} \cdot \xi_j \cdot \xi_k \end{aligned}$$

heißt *die quadratische Form zur symmetrischen Matrix* A.

Proposition. *Es sei* $Q_A(\xi_1, \ldots, \xi_n)$ *die quadratische Form zur symmetrischen Matrix* $A \in \mathrm{M}_n(\mathbf{R})$. *Dann existieren Koordinaten* $\eta_1, \ldots, \eta_n$ *derart, daß*

$$Q_A(\xi_1, \ldots, \xi_n) = \lambda_1 \eta_1^2 + \ldots + \lambda_n \eta_n^2$$

gilt, wobei $\lambda_1, \ldots, \lambda_n$ *die Eigenwerte von* A *sind.*

Beweis. Da $A \in \mathrm{M}_n(\mathbf{R})$ symmetrisch ist, existiert eine orthogonale Matrix $S \in \mathrm{O}_n(\mathbf{R})$ (siehe Kapitel 5) mit der Eigenschaft

$${}^tS \cdot A \cdot S = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A sind. Mit den neuen Koordinaten

$$\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} := {}^tS \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = {}^tS \cdot {}^t(\xi_1, \ldots, \xi_n)$$

folgt nun

$$\begin{aligned} Q_A(\xi_1,\ldots,\xi_n) &= (\xi_1,\ldots,\xi_n)\cdot A\cdot {}^t(\xi_1,\ldots,\xi_n) \\ &= (\xi_1,\ldots,\xi_n)\cdot S\cdot {}^tS\cdot A\cdot S\cdot {}^tS\cdot {}^t(\xi_1,\ldots,\xi_n) \\ &= (\eta_1,\ldots,\eta_n)\cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \cdot {}^t(\eta_1,\ldots,\eta_n) \\ &= \lambda_1\eta_1^2+\ldots+\lambda_n\eta_n^2\,, \end{aligned}$$

wie behauptet. □

Bemerkung. Die vorhergehende Proposition besagt, daß bei einer quadratischen Form durch geeignete Koordinatentransformation die gemischten Glieder immer zum Verschwinden gebracht werden können. Die Proposition zeigt überdies, daß dies sogar durch eine orthogonale Koordinatentransformation erreicht werden kann. Natürlich ist es auch möglich, dies durch andere reguläre, nicht notwendigerweise orthogonale Koordinatentransformationen zu erreichen. Der *Trägheitssatz von Sylvester*, der hier nicht bewiesen werden soll, besagt dann, wie man auch immer durch eine reguläre Koordinatentransformation die gemischten Glieder einer quadratischen Form zum Verschwinden bringt, so sind in der neu entstehenden quadratischen Form

$$\lambda_1'\,\xi_1'^2+\ldots+\lambda_n'\,\xi_n'^2$$

die Anzahl der verschwindenden, der positiven und der negativen Koeffizienten immer gleich.

Bezeichnung. Ist

$$Q_A(\xi_1,\ldots,\xi_n)=(\xi_1,\ldots,\xi_n)\cdot A\cdot {}^t(\xi_1,\ldots,\xi_n)$$

die vorgelegte quadratische Form zur symmetrischen Matrix A und gilt für $S'\in \mathrm{GL}_n(\mathbf{R})$

$${}^tS'\cdot A\cdot S'=\begin{pmatrix} \lambda_1' & & \\ & \ddots & \\ & & \lambda_n' \end{pmatrix},$$

d.h. mit ${}^t(\xi_1',\ldots,\xi_n')={}^tS'\cdot {}^t(\xi_1,\ldots,\xi_n)$ besteht die Gleichung

$$Q_A(\xi_1,\ldots,\xi_n)=\lambda_1'\,\xi_1'^2+\ldots+\lambda_n'\,\xi_n'^2,$$

so ordnen wir ab jetzt die Koeffizienten wie folgt an:

$$\lambda_1',\ldots,\lambda_p'>0\,,\quad \lambda_{p+1}',\ldots,\lambda_r'<0\,,\quad \lambda_{r+1}',\ldots,\lambda_n'=0\,;$$

dabei gilt $r=\mathrm{rg(A)}$.

Definition. Mit den vorhergehenden Bezeichnungen wird das Paar $(p,r-p)$ *die Signatur von* A genannt.

Bemerkung. Nach dem Trägheitssatz von Sylvester ist die Signatur wohldefiniert. Indem wir

$$T := S' \cdot \begin{pmatrix} \lambda_1'^{-\frac{1}{2}} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_p'^{-\frac{1}{2}} & & & & & & \\ & & & |\lambda'_{p+1}|^{-\frac{1}{2}} & & & & & \\ & & & & \ddots & & & & \\ & & & & & |\lambda'_r|^{-\frac{1}{2}} & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix}$$

setzen, erhalten wir

$${}^tT \cdot A \cdot T = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix},$$

wobei p-mal die $1, (r-p)$-mal die -1 und $(n-r)$-mal die 0 auftritt.

8.2 Quadriken

Wir beginnen mit der folgenden

Definition. Es seien $A = (\alpha_{j,k})_{j,k=1,\ldots,n} \in \mathrm{M}_n(\mathbf{R})$ eine symmetrische Matrix, $b = {}^t(\beta_1, \ldots, \beta_n) \in \mathbf{R}^n$ und $\gamma \in \mathbf{R}$. Dann heißt die Menge $C \subseteq \mathbf{R}^n$, gegeben durch

$$\begin{aligned} C &:= \{x = {}^t(\xi_1, \ldots, \xi_n) \in \mathbf{R}^n \mid {}^tx \cdot A \cdot x + 2 \cdot {}^tx \cdot b + \gamma = 0\} \\ &= \{x = {}^t(\xi_1, \ldots, \xi_n) \in \mathbf{R}^n \mid \sum_{j,k=1}^{n} \alpha_{j,k}\xi_j\xi_k + 2\sum_{j=1}^{n} \beta_j\xi_j + \gamma = 0\}, \end{aligned}$$

eine *Quadrik* oder *Hyperfläche zweiten Grades des* $\mathbf{R}^n$.

Bemerkung. Indem wir

$$x' := \begin{pmatrix} 1 \\ \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbf{R}^{n+1}, \; A' := \begin{pmatrix} \gamma & \beta_1 & \cdots & \beta_n \\ \beta_1 & \alpha_{1,1} & \ldots & \alpha_{1,n} \\ \vdots & \vdots & & \vdots \\ \beta_n & \alpha_{n,1} & \ldots & \alpha_{n,n} \end{pmatrix} \in \mathrm{M}_{n+1}(\mathbf{R})$$

setzen, können wir die Quadrik C in der vorhergehenden Definition auch wie folgt beschreiben:

$$C = \{x = {}^t(\xi_1, \ldots, \xi_n) \in \mathbf{R}^n \mid {}^t x' \cdot A' \cdot x' = 0\}.$$

Satz (Affine Klassifikation von Quadriken). Mit den vorhergehenden Bezeichnungen sei

$$C = \{x = {}^t(\xi_1, \ldots, \xi_n) \in \mathbf{R}^n \mid {}^t x' \cdot A' \cdot x' = 0\}$$

eine Quadrik. Weiter sei

$$r = \operatorname{rg}(A), \quad r' = \operatorname{rg}(A').$$

Dann existiert eine affine Koordinatentransformation

$$(\xi_1, \ldots, \xi_n) \mapsto (\eta_1, \ldots, \eta_n),$$

gegeben durch eine affine Abbildung $f \in \mathrm{GA}(\mathbf{R}^n)$, so daß C durch eine der drei folgenden Normalformen beschrieben wird:

(a) $\eta_1^2 + \ldots + \eta_p^2 - \eta_{p+1}^2 - \ldots - \eta_r^2 = 0$, falls $r' = r$ $(0 \leq p \leq r)$,

(b) $\eta_1^2 + \ldots + \eta_p^2 - \eta_{p+1}^2 - \ldots - \eta_r^2 = 1$, falls $r' = r + 1$ $(0 \leq p \leq r)$,

(c) $\eta_1^2 + \ldots + \eta_p^2 - \eta_{p+1}^2 - \ldots - \eta_r^2 + 2\eta_{r+1} = 0$, falls $r' = r + 2$ $(0 \leq p \leq r)$.

Beweis. Gemäß dem Abschnitt 8.1 existiert eine reguläre Matrix $T \in \mathrm{GL}_n(\mathbf{R})$ mit der Eigenschaft

$${}^tT \cdot A \cdot T = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& 0 && \\ &&&&&&& \ddots & \\ &&&&&&&& 0 \end{pmatrix},$$

wobei p-mal die 1, $(r - p)$-mal die -1 und $(n - r)$-mal die 0 auftritt. Mit Hilfe von

$$T_1' := \begin{pmatrix} 1 & 0 \\ 0 & T \end{pmatrix} \in \mathrm{GL}_{n+1}(\mathbf{R})$$

berechnet man leicht

$$A'_1 := {}^tT'_1 \cdot A' \cdot T'_1 = \begin{pmatrix} \gamma & \beta'_1 & \cdots & \beta'_p & \beta'_{p+1} & \cdots & \beta'_r & \beta'_{r+1} & \cdots & \beta'_n \\ \beta'_1 & 1 & & & & & & & & \\ \vdots & & \ddots & & & & & & & \\ \beta'_p & & & 1 & & & & & & \\ \beta'_{p+1} & & & & -1 & & & & & \\ \vdots & & & & & \ddots & & & & \\ \beta'_r & & & & & & -1 & & & \\ \beta'_{r+1} & & & & & & & 0 & & \\ \vdots & & & & & & & & \ddots & \\ \beta'_n & & & & & & & & & 0 \end{pmatrix}$$

mit gewissen $\beta'_1, \ldots, \beta'_n \in \mathbf{R}$. Indem wir nun die neuen Koordinaten

$$\begin{aligned} & w' := T_1'^{-1} \cdot x', \text{ d.h.} \\ & {}^t(1, \omega_1, \ldots, \omega_n) = T_1'^{-1} \cdot {}^t(1, \xi_1, \ldots, \xi_n) \end{aligned}$$

einführen, erhalten wir die folgende Vereinfachung der definierenden Gleichung von C

$$\begin{aligned} {}^tx' \cdot A' \cdot x' &= {}^tx' \cdot {}^tT_1'^{-1} \cdot {}^tT'_1 \cdot A' \cdot T'_1 \cdot T_1'^{-1} \cdot x' \;=\; {}^tw' \cdot A'_1 \cdot w' \\ &= \omega_1^2 + \ldots + \omega_p^2 - \omega_{p+1}^2 - \ldots - \omega_r^2 + 2\beta'_1\omega_1 + \ldots + 2\beta'_n\omega_n + \gamma \;=\; 0. \end{aligned}$$

Mit der regulären Matrix

$$T'_2 := \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ -\beta'_1 & 1 & & & & & & & & \\ \vdots & & \ddots & & & & & & & \\ -\beta'_p & & & 1 & & & & & & \\ \beta'_{p+1} & & & & 1 & & & & & \\ \vdots & & & & & \ddots & & & & \\ \beta'_r & & & & & & 1 & & & \\ 0 & & & & & & & 1 & & \\ \vdots & & & & & & & & \ddots & \\ 0 & & & & & & & & & 1 \end{pmatrix}$$

berechnet man leicht

$$A_2' := {}^tT_2' \cdot A_1' \cdot T_2' = \begin{pmatrix} \gamma' & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta_{r+1}' & \cdots & \beta_n' \\ 0 & 1 & & & & & & & & \\ \vdots & & \ddots & & & & & & & \\ 0 & & & 1 & & & & & & \\ 0 & & & & -1 & & & & & \\ \vdots & & & & & \ddots & & & & \\ 0 & & & & & & -1 & & & \\ \beta_{r+1}' & & & & & & & 0 & & \\ \vdots & & & & & & & & \ddots & \\ \beta_n' & & & & & & & & & 0 \end{pmatrix}$$

mit einem gewissen $\gamma' \in \mathbf{R}$. Indem wir die neuen Koordinaten

$$z' := T_2'^{-1} \cdot w', \text{ d.h.}$$
$${}^t(1, \zeta_1, \ldots, \zeta_n) = T_2'^{-1} \cdot {}^t(1, \omega_1, \ldots, \omega_n)$$

einführen, erhalten wir die weitere Vereinfachung der definierenden Gleichung von C

$$\begin{aligned} {}^tx' \cdot A' \cdot x' &= {}^tw' \cdot A_1' \cdot w' = {}^tw' \cdot {}^tT_2'^{-1} \cdot {}^tT_2' \cdot A_1' \cdot T_2' \cdot T_2'^{-1} \cdot w' = {}^tz' \cdot A_2' \cdot z' \\ &= \zeta_1^2 + \ldots + \zeta_p^2 - \zeta_{p+1}^2 - \ldots - \zeta_r^2 + 2\beta_{r+1}' \zeta_{r+1} + \ldots + 2\beta_n' \zeta_n + \gamma' = 0. \end{aligned}$$

Wir unterscheiden nun drei Fälle:

(a') $\gamma' = 0$; $\beta_{r+1}' = \ldots = \beta_n' = 0$.

(b') $\gamma' \neq 0$; $\beta_{r+1}' = \ldots = \beta_n' = 0$.

(c') $\exists\, j \in \{r+1, \ldots, n\} : \beta_j' \neq 0$.

Ad (a'): Indem wir $T' := T_1' \cdot T_2' \in \mathrm{GL}_{n+1}(\mathbf{R})$ setzen und die Koordinaten

$$y' := z' = T'^{-1} \cdot x', \text{ d.h.}$$
$${}^t(1, \eta_1, \ldots, \eta_n) = T'^{-1} \cdot {}^t(1, \xi_1, \ldots, \xi_n)$$

einführen, erhalten wir die behauptete Normalform (a) von C, d.h.

$$\eta_1^2 + \ldots + \eta_p^2 - \eta_{p+1}^2 - \ldots - \eta_r^2 = 0.$$

Dabei stellen wir leicht $r = r'$ fest.

Ad (b'): Indem wir $T' := T_1' \cdot T_2' \in \mathrm{GL}_{n+1}(\mathbf{R})$ setzen, erhalten wir mit den Koordinaten

$$z' = T'^{-1} \cdot x', \text{ d.h.}$$
$${}^t(1, \zeta_1, \ldots, \zeta_n) = T'^{-1} \cdot {}^t(1, \xi_1, \ldots, \xi_n)$$

nach den vorhergehenden Rechnungen die folgende C definierende Gleichung

$$\zeta_1^2 + \ldots + \zeta_p^2 - \zeta_{p+1}^2 - \ldots - \zeta_r^2 = -\gamma'. \qquad (*)$$

Ist $\gamma' < 0$, d.h. $-\gamma' > 0$, so erhalten wir mit den neuen Koordinaten

$$\eta_j := \zeta_j / \sqrt{-\gamma'} \quad (j = 1, \ldots, n)$$

die behauptete Normalform (b) von C, d.h.

$$\eta_1^2 + \ldots + \eta_p^2 - \eta_{p+1}^2 - \ldots - \eta_r^2 = 1.$$

Ist $\gamma' > 0$, so erhalten wir die behauptete Normalform (b) nach Multiplikation von $(*)$ mit -1 und einer analogen Koordinatentransformation. Wir stellen leicht $r' = r + 1$ fest.

Ad (c'): Indem wir die Koordinaten $\xi_1, \ldots, \xi_n$ gegebenenfalls permutieren, können wir ohne Einschränkung $\beta'_{r+1} \neq 0$ annehmen. Mit Hilfe elementarer Umformungen läßt sich erreichen, daß in der Matrix A'_2 der Eintrag β'_{r+1} zu $1, \gamma'$ zu 0 und auch alle übrigen β'_j zu 0 gemacht werden können. Mit anderen Worten findet sich $T'_3 \in \mathrm{GL}_{n+1}(\mathbf{R})$ mit der Eigenschaft

$$A'_3 := {}^tT'_3 \cdot A'_2 \cdot T'_3 = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ 0 & 1 & & & & & & & & \\ \vdots & & \ddots & & & & & & & \\ 0 & & & 1 & & & & & & \\ 0 & & & & -1 & & & & & \\ \vdots & & & & & \ddots & & & & \\ 0 & & & & & & -1 & & & \\ 1 & & & & & & & 0 & & \\ \vdots & & & & & & & & \ddots & \\ 0 & & & & & & & & & 0 \end{pmatrix}.$$

Indem wir $T' := T'_1 \cdot T'_2 \cdot T'_3 \in \mathrm{GL}_{n+1}(\mathbf{R})$ setzen, erhalten wir mit den Koordinaten

$$y' := T'^{-1} \cdot x', \text{ d.h.}$$
$${}^t(1, \eta_1, \ldots, \eta_n) = T'^{-1} \cdot {}^t(1, \xi_1, \ldots, \xi_n)$$

nach den vorhergehenden Rechnungen die folgende C definierende Gleichung

$$\eta_1^2 + \ldots + \eta_p^2 - \eta_{p+1}^2 - \ldots - \eta_r^2 + 2\eta_{r+1} = 0.$$

Wir stellen leicht $r' = r + 2$ fest.

In allen drei Fällen entspricht die Matrix $T' \in \mathrm{GL}_{n+1}(\mathbf{R})$ der gesuchten affinen Transformation $f \in \mathrm{GA}(\mathbf{R}^n)$. $\square$

Klassifikation von nicht-leeren Quadriken im $\mathbf{A}^2(\mathbf{R})$.

Fall (a) :	$r=0$,	$p=0$:	$0=0$,	d.h. affine Ebene $\mathbf{A}^2(\mathbf{R})$.
	$r=1$,	$p=1$:	$\eta_1^2=0$,	d.h. Doppelgerade.
	$r=2$,	$p=1$:	$\eta_1^2-\eta_2^2=0$,	d.h. Geradenpaar mit Schnittpunkt.
	$r=2$,	$p=2$:	$\eta_1^2+\eta_2^2=0$,	d.h. Punkt.
Fall (b) :	$r=1$,	$p=1$:	$\eta_1^2=1$,	d.h. paralleles Geradenpaar.
	$r=2$,	$p=1$:	$\eta_1^2-\eta_2^2=1$,	d.h. Hyperbel.
	$r=2$,	$p=2$:	$\eta_1^2+\eta_2^2=1$,	d.h. Kreis (Ellipse).
Fall (c) :	$r=1$,	$p=1$:	$\eta_1^2+2\eta_2=0$,	d.h. Parabel.

Klassifikation von nicht-leeren Quadriken im $\mathbf{A}^3(\mathbf{R})$.

Fall (a) :	$r=0$,	$p=0$:	$0=0$,	d.h. affiner Raum $\mathbf{A}^3(\mathbf{R})$.
	$r=1$,	$p=1$:	$\eta_1^2=0$,	d.h. Doppelebene.
	$r=2$,	$p=1$:	$\eta_1^2-\eta_2^2=0$,	d.h. Ebenenpaar mit Schnittgerade.
	$r=2$,	$p=2$:	$\eta_1^2+\eta_2^2=0$,	d.h. Gerade.
	$r=3$,	$p=2$:	$\eta_1^2+\eta_2^2-\eta_3^2=0$,	d.h. Kreiskegel.
	$r=3$,	$p=3$:	$\eta_1^2+\eta_2^2+\eta_3^2=0$,	d.h. Punkt.
Fall (b) :	$r=1$,	$p=1$:	$\eta_1^2=1$,	d.h. paralleles Ebenenpaar.
	$r=2$,	$p=1$:	$\eta_1^2-\eta_2^2=1$,	d.h. hyperbolischer Zylinder.
	$r=2$,	$p=2$:	$\eta_1^2+\eta_2^2=1$,	d.h. Kreiszylinder.
	$r=3$,	$p=1$:	$\eta_1^2-\eta_2^2-\eta_3^2=1$,	d.h. zweischaliges Hyperboloid.
	$r=3$,	$p=2$:	$\eta_1^2+\eta_2^2-\eta_3^2=1$,	d.h. einschaliges Hyperboloid.
	$r=3$,	$p=3$:	$\eta_1^2+\eta_2^2+\eta_3^2=1$,	d.h. Kugel (Ellipsoid).
Fall (c) :	$r=1$,	$p=1$:	$\eta_1^2+2\eta_2=0$,	d.h. parabolischer Zylinder.
	$r=2$,	$p=1$:	$\eta_1^2-\eta_2^2+2\eta_3=0$,	hyperbolisches Paraboloid.
	$r=2$,	$p=2$:	$\eta_1^2+\eta_2^2+2\eta_3=0$,	elliptisches Paraboloid.