Lineare Algebra II

Eine Vorlesung von Prof. Dr. Klaus Hulek
hulek@math.uni-hannover.de

Bei dieser Vorlesung handelt es sich um die Fortsetzung
der Vorlesung *Lineare Algebra I* vom WS 2001/02.

Institut für Mathematik
Universität Hannover
D – 30060 Hannover
Germany
E-Mail : hulek@math.uni-hannover.de

Inhaltsverzeichnis

§ 10 Der Satz von Sylvester und das Hurwitzkriterium 4

Der Trägheitssatz von Sylvester – Das Hurwitz Kriterium – Diagonalisierung symmetrischer Matrizen

§ 11 Eigenwerttheorie II 13

Polynome – Ideale in $K[x]$ – Nilpotente Matrizen – Idempotente Matrizen – Zerfallende Matrizen – Das Minimalpolynom – Eigenwerte und Minimalpolynom – Spektralzerlegung – Die Jordan-Chevalley Zerlegung – Anwendungen – Der Quotientenraum – Kanonische Faktorisierung – Die Joradansche Normalform – Nilzyklische Matrizen – Zur Berechnung der Jordanform

§ 12 Projektive Räume 83

Definition der projektiven Räume – Homogene Koordinaten – Abbildungen – Koordinatensysteme – Affinitäten

§ 13 Quadriken in projektiven Räumen 100

Projektive Quadriken – Projektive Klassifikation von Quadriken in $\mathbb{P}^{\mathbf{n}}(\mathbb{R})$

§ 14 Affine Quadriken 110

Affine Quadriken – Klassifikation der reellen projektiven Quadriken bezüglich affiner Äquivalenz – Klassifikation der reellen affinen Quadriken

§ 15 Das Tensorprodukt 126

Definition und Konstruktion des Tensorprodukts – Tensorprodukte von Unterräumen und Quotienten – Tensorprodukte und lineare Abbildungen – Komposition von Abbildungen – Bild und Kern – Das Kroneckerprodukt von Matrizen – Mehrfache Tensorprodukte – Die Tensoralgebra – Induzierte Abbildungen – Die gemischte Tensoralgebra – Darstellung mittels einer Basis – Transformationsverhalten – Tensorverjüngung

§ 16 Äußere Algebra 154

Das p-fache äußere Produkt – Vorbereitungen zur Konstruktion des äußeren Produkts – Interpretation der Determinante – Induzierte lineare Abbildungen – Äußere Algebra – Multiplikation

§ 17 Symmetrische Algebra 169

Fas p-fache symmetrische Produkt – Symmetrische Tensoren – Die symmetrische Algebra

Literaturverzeichnis 176

§ 10 Der Satz von Sylvester und das Hurwitzkriterium

Der Trägheitssatz von Sylvester

Es sei V ein endlich-dimensionaler reeller oder komplexer Vektorraum mit einer symmetrischen Bilinearform (hermiteschen Sesquilinearform)

$$\langle\ ,\ \rangle : V \times V \longrightarrow K.$$

Bezüglich einer Basis $B = (v_1, \ldots, v_n)$ wird $\langle\ ,\ \rangle$ dargestellt durch die Matrix

$$A = (a_{ij}), \quad a_{ij} = \langle v_i, v_j \rangle.$$

Nach der Transformationsformel ändert sich bei einem Basiswechsel mit Übergangsmatrix $S \in \mathrm{GL}(n; K)$ die darstellende Matrix auf die folgende Weise:

$$A \longmapsto {}^t\overline{S}AS.$$

Wir wissen bereits, daß es ein $S \in O(n)$ (bzw. $U(n)$) gibt mit

$${}^t\overline{S}AS = S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad (\lambda_1, \ldots, \lambda_n \in \mathbb{R}).$$

Wir fragen nun nach der *Normalform*, wenn wir allgemeiner $S \in \mathrm{GL}(n; K)$ zulassen. Eine erste Antwort hierauf liefert der folgende Satz.

Satz 10.1 (Trägheitssatz von Sylvester)
Es sei $\langle , \rangle : V \times V \to K$ eine symmetrische Bilinearform (hermitesche Sesquilinearform). Ferner seien B_1, B_2 Basen von V und A_1, A_2 die zugehörigen darstellenden Matrizen. Sei

$$\begin{aligned} k_i &:= \text{Anzahl positiver Eigenwerte von } A_i \\ l_i &:= \text{Anzahl negativer Eigenwerte von } A_i, \end{aligned}$$

wobei die Eigenwerte mit Vielfachheit gezählt werden. Dann gilt

(i) $k_1 = k_2$,

(ii) $l_1 = l_2$,

(iii) $k_1 + l_1 = \operatorname{Rang} A_1 = \operatorname{Rang} A_2 = k_2 + l_2$.

Beweis. Es sei

$$\begin{aligned} V_i^+ &:= \text{Unterraum aufgespannt von den Eigenvektoren} \\ &\ \text{zu positiven Eigenwerten von } A_i \\ V_i^- &:= \text{Unterraum aufgespannt von den Eigenvektoren} \\ &\ \text{zu negativen Eigenwerten von } A_i \\ V_i^0 &:= \text{Unterraum aufgespannt von den Eigenvektoren} \\ &\ \text{zum Eigenwert } 0 = \operatorname{Ker} A_i. \end{aligned}$$

(Hierbei heißt ein Vektor $v \in V$ ein Eigenvektor von A_i zum Eigenwert λ_i wenn $A_i q_{B_i}(v) = \lambda_i q_{B_i}(v)$ gilt.)

Behauptung. Es gilt

$$V_i^0 = \{v \in V; \langle v, w\rangle = 0 \text{ für alle } w \in V\}.$$

Dies sieht man etwa wie folgt ein. Es gilt:

$$\langle v, w\rangle = \overline{\langle w, v\rangle} = {}^t q_{B_i}(w)\overline{A}_i\overline{q_{B_i}(v)}.$$

Also ist $\langle v, w\rangle = 0$ für alle $w \in V$ genau dann wenn $A_i q_{B_i}(v) = 0$ ist, d. h. wenn $q_{B_i}(v) \in \mathrm{Ker} A_i = \mathrm{Eig}(A_i, 0)$ ist. Diese Darstellung zeigt aber auch, daß

$$V_1^0 = V_2^0 =: V^0.$$

Ferner gilt nach Korollar (I.9.28):

$$V = V_i^+ \perp V_i^- \perp V^0$$

und damit

$$k_1 + l_1 = \mathrm{Rang}\, A_1 = \mathrm{Rang}\, A_2 = k_2 + l_2,$$

bzw.

$$k_i + l_i + \dim V^0 = \dim V.$$

Es gilt ferner:

$$\begin{gathered} 0 \neq v \in V_i^+ \Rightarrow \langle v, v\rangle > 0 \\ 0 \neq v \in V_i^- \Rightarrow \langle v, v\rangle < 0. \end{gathered}$$

Dies zeigt man wie folgt. Es sei $v_1^{(i)}, \dots, v_{k_i}^{(i)}$ eine ONBasis von V_i^+ bestehend aus Eigenvektoren von A_i. Dann ist die Einschränkung von $\langle\ ,\ \rangle$ auf V_i^+ bezüglich dieser Basis durch eine Matrix

$$\begin{pmatrix} \lambda_1^i & & 0 \\ & \ddots & \\ 0 & & \lambda_{k_i}^i \end{pmatrix}; \quad \lambda_1^i, \dots, \lambda_{k_i}^i > 0$$

gegeben, wobei die λ_j die positiven Eigenwerte von A_i sind. Damit gilt:

$$0 \neq v \in V_i^+ \Rightarrow v = \sum_{j=1}^{k_i} x_j^{(i)} v_j^{(i)} \Rightarrow \langle v, v\rangle = \sum_{j=1}^{k_i} |x_j^{(i)}|^2 \underbrace{\langle v_j^{(i)}, v_j^{(i)}\rangle}_{=\lambda_j^i > 0} > 0.$$

Damit folgt

$$V_2^+ \cap (V_1^- \perp V^0) = \{0\}$$

und daher gilt

$$k_2 + l_1 + \dim V^0 \leq \dim V.$$

Wegen

$$k_1 + l_1 + \dim V^0 = \dim V$$

erhalten wir $k_2 \leq k_1$. Analog zeigt man $k_1 \leq k_2$. Damit gilt $k_1 = k_2$, also auch $l_1 = l_2$. $\square$

Korollar 10.2
Sei $A \in \mathrm{Mat}(n; K)$ hermitesch (symmetrisch). Sei $S \in \mathrm{GL}(n; K)$. Dann haben A und ${}^t\overline{S}AS$ dieselbe Anzahl von positiven bzw. negativen Eigenwerten sowie denselben Rang.

Bemerkung. Sei $(V, \langle\ ,\ \rangle)$ gegeben. Dann hat man eine (nicht notwendig eindeutig bestimmte) Zerlegung

$$V = V^+ \perp V^- \perp V^0$$

mit

$$\begin{aligned}
&0 \neq v \in V^+ \Rightarrow \langle v, v\rangle > 0\\
&0 \neq v \in V^- \Rightarrow \langle v, v\rangle < 0\\
&V^0 = \{v \in V;\ \langle v, w\rangle = 0 \text{ für alle } w \in V\}.
\end{aligned}$$

V^0 heißt der *Ausartungsraum* (oder *Nullraum*) der Form $\langle\ ,\ \rangle$.

Definition.

(i) Der *Index* der Form $\langle\ ,\ \rangle$ ist die Anzahl der (mit Vielfachheit gezählten) positiven Eigenwerte einer (und damit jeder) darstellenden Matrix von $\langle\ ,\ \rangle$.

(ii) Der *Rang* von $\langle\ ,\ \rangle$ ist der Rang einer (und damit jeder) darstellenden Matrix.

(iii) Die *Signatur* von $\langle\ ,\ \rangle$ ist die Differenz zwischen der Anzahl der positiven und der Anzahl der negativen Eigenwerte einer (und damit jeder) darstellenden Matrix.

Bemerkung. Für eine Zerlegung

$$V = V^+ \perp V^- \perp V^0$$

gilt also

$$\begin{aligned}
\text{Index} &= \dim V^+,\\
\text{Rang} &= \dim V^+ + \dim V^-,\\
\text{Signatur} &= \dim V^+ - \dim V^-.
\end{aligned}$$

Satz 10.3
Es gibt eine Basis B bezüglich der $\langle\ ,\ \rangle$ dargestellt wird durch

$$\left.\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & 0 & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & 0 & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \quad \begin{matrix} \}k \\ \}l \end{matrix}\right.$$

wobei k der Index der Form $\langle\ ,\ \rangle$ ist, und $l = \text{Index} - \text{Signatur}$ ist.

Beweis. Es gibt zunächst eine Basis

$$B' = (w_1, \dots, w_n)$$

bezüglich der die darstellende Matrix wie folgt lautet:

$$\begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & 0 & \\ & & \lambda_k & & & & & \\ & & & \lambda_{k+1} & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_{k+l} & & \\ & & & & & & 0 & \\ & 0 & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}, \quad \lambda_i \text{ ist } \begin{cases} > 0 & \text{für } i \leq k \\ < 0 & \text{für } k < i \leq k+l. \end{cases}$$

Wir setzen

$$v_i := \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} w_i & 1 \leq i \leq k+l \\ w_i & i > k+l. \end{cases}$$

Dann ist für $1 \leq i \leq k+l$:

$$\langle v_i, v_i \rangle = \frac{1}{|\lambda_i|}\lambda_i = \begin{cases} +1 & 1 \leq i \leq k \\ -1 & k < i \leq k+l. \end{cases}$$

Also können wir

$$B = (v_1, \dots, v_n)$$

wählen um das Gewünschte zu erreichen. $\square$

Folgerung 10.4

(i) Ist A eine symmetrische (hermitesche) Matrix, so gibt es ein $S \in \mathrm{GL}(n, K)$ mit

$$ {}^t\overline{S}AS = \begin{pmatrix} {}^1 \ddots {}_1 & & 0 \\ & {}^{-1} \ddots {}_{-1} & \\ 0 & & {}^0 \ddots {}_0 \end{pmatrix} $$

(ii) Die hermiteschen Sesquilinearformen (symmetrischen Bilinearformen) werden durch *Index* und *Signatur* klassifiziert.

Das Hurwitz-Kriterium

Es sei $A = {}^t\overline{A} \in \mathrm{Mat}(n; K)$ eine symmetrische (hermitesche) Matrix.

Definition. Die Matrix A heißt *positiv definit*, falls die zugehörige Form $\langle\ ,\ \rangle_A$ positiv definit ist, d. h.

$$ \langle x, x \rangle_A = {}^t\overline{x}Ax > 0 \text{ für } x \neq 0. $$

A heißt *negativ definit* wenn $-A$ positiv definit ist.

Schreibweise. Ist A positiv definit, so schreibt man $A > 0$.

Es ist wichtig, ein Kriterium dafür zu haben, daß eine Matrix positiv definit ist.

Satz 10.5

A ist genau dann positiv definit wenn alle Eigenwerte positiv sind.

Beweis. Die Eigenschaft positiv definit zu sein, hängt nur von der Form $\langle\ ,\ \rangle_A$ ab, nicht von der gewählten Basis. Nach Basiswechsel gilt:

$$ {}^t\overline{S}AS = \begin{pmatrix} {}^1 \ddots {}_1 & & 0 \\ & {}^{-1} \ddots {}_{-1} & \\ 0 & & {}^0 \ddots {}_0 \end{pmatrix} \begin{matrix} \}k \\ \\ \\ \end{matrix} . $$

Diese Form ist genau dann positiv definit, wenn $k = n$ ist. $\square$

Sei nun $K = \mathbb{R}$ und $A = {}^tA$. Wir schreiben

$$A = \left(\begin{array}{ccc|cc} a_{11} & \dots & a_{1k} & \dots & a_{1n} \\ \vdots & & \vdots & & \\ a_{k1} & \dots & a_{kk} & & \vdots \\ \hline \vdots & & & & \\ a_{n1} & & \dots & & a_{nn} \end{array}\right)$$

und definieren

$$A_k := (a_{ij})_{1 \leq i,j \leq k}.$$

Satz 10.6 (Hurwitz-Kriterium)
Es sei $A \in \mathrm{Mat}(n; \mathbb{R})$ eine reelle symmetrische Matrix. Dann sind äquivalent:

(i) A ist positiv definit.

(ii) $\det A_k > 0$ für $1 \leq k \leq n$.

Beweis. (i)$\Rightarrow$(ii): Wir zeigen zunächst

$$\det A = \det A_n > 0.$$

Da A positiv definit ist, gibt es ein $S \in \mathrm{GL}(n; \mathbb{R})$ mit

$${}^tSAS = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Also folgt

$$1 = \det({}^tSAS) = \det A (\det S)^2$$

und damit

$$\det A > 0.$$

Um nun $\det A_k > 0$ für $1 \leq k \leq n$ zu zeigen, betrachten wir

$$\mathbb{R}^k = \{x \in \mathbb{R}^n;\ x_{k+1} = \ldots = x_n = 0\} \subset \mathbb{R}^n.$$

Die Form $\langle\ ,\ \rangle$ induziert durch Einschränkung eine Form

$$\langle\ ,\ \rangle_k : \mathbb{R}^k \times \mathbb{R}^k \longrightarrow \mathbb{R}$$

mit darstellender Matrix A_k. Da auch $\langle\ ,\ \rangle_k$ positiv definit ist, folgt

$$\det A_k > 0 \qquad (1 \leq k \leq n).$$

(ii)⇒(i): Wir machen Induktion nach n.

$n = 1$: Klar.

$n-1 \mapsto n$: A_{n-1} ist nach Induktionsvoraussetzung positiv definit. Also gibt es ein $S' \in \mathrm{GL}(n-1;\mathbb{R})$ mit

$$ {}^tS' A_{n-1} S' = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = E_{n-1}. $$

Sei

$$ S := \left(\begin{array}{c|c} & 0 \\ S' & \vdots \\ & 0 \\ \hline 0 \cdots 0 & 1 \end{array}\right) \in \mathrm{GL}(n;\mathbb{R}). $$

Es ist

$$ \begin{aligned} {}^tSAS &= \left(\begin{array}{c|c} & 0 \\ {}^tS' & \vdots \\ & 0 \\ \hline 0 \cdots 0 & 1 \end{array}\right) \left(\begin{array}{c|c} & a_{1,n} \\ A_{n-1} & \vdots \\ & a_{n-1,n} \\ \hline a_{1,n} \cdots a_{n-1,n} & a_{n,n} \end{array}\right) \left(\begin{array}{c|c} & 0 \\ S' & \vdots \\ & 0 \\ \hline 0 \cdots 0 & 1 \end{array}\right) \\ &= \left(\begin{array}{ccc|c} 1 & & & b_1 \\ & \ddots & & \vdots \\ & & 1 & b_{n-1} \\ \hline b_1 & \cdots & b_{n-1} & b_n \end{array}\right) =: B, \end{aligned} $$

wobei $b_i = a_{i,n}$. Dann gilt

$$ (*) \qquad \det B = (\det S)^2 \det A > 0. $$

Es genügt zu zeigen, daß B positiv definit ist. Wir setzen

$$ T := \left(\begin{array}{ccc|c} 1 & & & -b_1 \\ & \ddots & & \vdots \\ & & 1 & -b_{n-1} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \in \mathrm{GL}(n;\mathbb{R}). $$

Dann ist

$$
\begin{aligned}
{}^tTBT &= \left(\begin{array}{ccc|c} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ \hline -b_1 & \cdots & -b_{n-1} & 1 \end{array}\right) \left(\begin{array}{ccc|c} 1 & & & b_1 \\ & \ddots & & \vdots \\ & & 1 & b_{n-1} \\ \hline b_1 & \cdots & b_{n-1} & b_n \end{array}\right) \left(\begin{array}{ccc|c} 1 & & & -b_1 \\ & \ddots & & \vdots \\ & & 1 & -b_{n-1} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \\
&= \left(\begin{array}{ccc|c} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ \hline -b_1 & \cdots & -b_{n-1} & 1 \end{array}\right) \left(\begin{array}{ccc|c} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ \hline b_1 & \cdots & b_{n-1} & b_n - b_1^2 - \ldots - b_{n-1}^2 \end{array}\right) \\
&= \left(\begin{array}{ccc|c} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ \hline 0 & \cdots & 0 & b_n - b_1^2 - \ldots - b_{n-1}^2 \end{array}\right) =: C.
\end{aligned}
$$

Nun ist

$$\det C = (\det T)^2 \det B = \det B \overset{(*)}{>} 0.$$

Also ist

$$b_n - b_1^2 - \ldots - b_{n-1}^2 > 0.$$

Damit ist C positiv definit und damit auch B und A. $\square$

Diagonalisierung symmetrischer Matrizen

Wir hatten bereits früher die Elementarmatrizen

$$F_k(\alpha) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \alpha & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix},$$

wobei α an der Stelle (k,k) steht und

$$F_{kl} = \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \ldots & 1 & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow k \\ \\ \\ \\ \\ \end{array} \\ \uparrow l \end{array}$$

betrachtet. Multipliziert man eine Matrix A mit $F_k(\alpha)$ von links (rechts), so bedeutet dies Multiplikation der k-ten Zeile (Spalte) mit α. Die Multiplikation von A mit F_{kl} von links (rechts) bedeutet die Addition der l-ten Zeile von A zur k-ten Zeile (der k-ten Spalte von A zur l-ten Spalte). Es gilt

$$ {}^tF_k(\alpha) = F_k(\alpha), \quad {}^tF_{kl} = F_{lk}. $$

Satz 10.7
Es sei $A \in \mathrm{Mat}(n; \mathbb{R})$ eine reelle symmetrische Matrix. Dann gibt es Elementarmatrizen $C_1, \ldots, C_s$, so daß

$$ C_s \cdots C_1 A {}^tC_1 \ldots {}^tC_s = D, $$

wobei D eine Diagonalmatrix ist.

Beweis. Wir führen einen Induktionsbeweis nach n. Für $n = 1$ ist nichts zu zeigen.

1. Fall $a_{11} \neq 0$. Man addiere das $-a_{k1}a_{11}^{-1}$-fache der ersten Zeile zur k-ten Zeile und ebenso das $-a_{1k}a_{11}^{-1} = -a_{k1}a_{11}^{-1}$-fache der ersten Spalte zur k-ten Spalte, wobei $k = 2, \ldots, n$. Dies entspricht endlich vielen Zeilen- und Spaltenumformungen. Sind $C_1, \ldots, C_N$ die zu den Zeilenumformungen gehörigen Elementarmatrizen, so gilt

$$ C_N \ldots C_1 A \, {}^tC_1 \ldots {}^t C_N = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix} $$

wobei $A' = {}^tA'$ eine symmetrische $(n-1) \times (n-1)$ Matrix ist. Man kann dann die Induktionsvoraussetzung auf A' anwenden.

2. Fall $a_{11} = 0$, aber $a_{kk} \neq 0$ für ein $k \geq 2$. Die Matrix C sei ein Produkt von Elementarmatrizen, das dem Vertauschen der ersten und der k-ten Zeile entspricht. Dann ist $A' = CA{}^tC$ eine symmetrische Matrix mit $a'_{11} \neq 0$ und wir sind im 1. Fall.

3. Fall $a_{kk} = 0$ für alle k. Ist $A = 0$, so ist nichts zu zeigen. Ansonsten gibt es einen Eintrag $a_{ij} = a_{ji} \neq 0$. Die Elementarmatrix C entspreche der Addition der j-ten Zeile zur i-ten Zeile. Für $A' = CA{}^tC$ gilt dann $a'_{ii} = 2a_{ij} \neq 0$ und wir sind im zweiten Fall. $\square$

Bemerkung. Dieses Verfahren kann man für alle symmetrischen Matrizen $A \in \mathrm{Mat}(n; K)$ anwenden, so lange die Charakteristik des Körpers ungleich 2 ist, d.h. $1 + 1 \neq 0$.

Aus diesem Satz kann man aucheinen Algorithmus zur Diagonalisierung ableiten. Man schreibe die Matrix A und die Einheitsmatrix $E = E_n$ übereinander. Dann führe man A durch das ebene beschriebene Verfahren in eine Diagonalmatrix D über. Auf E wende man nur die entsprechenden Spaltenumformungen an. Dies ergibt

$$ \begin{pmatrix} A \\ E \end{pmatrix} \curvearrowright \begin{pmatrix} D \\ {}^tC_1 \ldots {}^tC_s \end{pmatrix}. $$

§ 11 Eigenwerttheorie II

Wir hatten bereits früher den Polynomring in einer Variablen über einem Körper K betrachtet:

$$K[x] = \mathrm{Abb}[\mathbb{N}, K] = \{P;\ P = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0;\ a_\nu \in K\}.$$

Definition. Eine Funktion $f : K \to K$ heißt eine *Polynomfunktion*, falls es ein Polynom $P \in K[x]$ gibt mit $f(a) = P(a)$ für alle $a \in K$.

Definiert man

$$\mathrm{Pol}(K) := \{f;\ f : K \to K \text{ ist Polynomfunktion}\},$$

so hat man eine surjektive Abbildung von K-Vektorräumen:

$$K[x] \longrightarrow \mathrm{Pol}(K).$$

Satz 11.1
Hat K unendlich viele Elemente, so ist die obige Abbildung ein Isomorphismus.

Bemerkungen.

(i) Für $K = \mathbb{R}$ folgt dies z. B. daraus, daß für jedes nicht-konstante Polynom $P \neq \text{const.}$ gilt: $\lim_{x\to\infty} P(x) = \pm\infty$.

(ii) Im folgenden werden wir meist annehmen, daß K unendlich viele Elemente besitzt und dann nicht mehr zwischen Polynomen und Polynomfunktionen unterscheiden.

(iii) Falls K endlich ist, ist die Aussage von Satz (11.1) stets falsch. Z. B. ist für $K = \mathbb{F}_2$ das Polynom $P(x) = x^2 + x \neq 0$, als Funktion ist jedoch $f(x) = x(x+1) \equiv 0$.

Lemma 11.2 (Vandermonde-Determinante)
Für $a_1, \ldots, a_n \in K$ gilt:

$$\det \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ \vdots & & & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix} = \prod_{1\le i<j\le n} (a_j - a_i).$$

Beweis. Sei

$$\Delta := \det \begin{pmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ \vdots & & & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{pmatrix}.$$

Induktion nach n:

$n = 1$: Klar.

$n-1 \mapsto n$: Wir formen um: Für $k = n, n-1, \ldots, 3, 2$ subtrahiere man das a_1-fache der $(k-1)$-ten Spalte von der k-ten Spalte:

$$\begin{pmatrix} 1 & a_1 & \cdots & a_1^{n-2} & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-2} & a_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} & a_n^{n-1} \end{pmatrix} \curvearrowright \begin{pmatrix} 1 & a_1 & \cdots & a_1^{n-2} & 0 \\ 1 & a_2 & \cdots & a_2^{n-2} & a_2^{n-2}(a_2 - a_1) \\ \vdots & & & & \vdots \\ 1 & a_n & \cdots & a_n^{n-2} & a_n^{n-2}(a_n - a_1) \end{pmatrix}$$

$$\cdots \curvearrowright \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & \cdots & (a_2 - a_1)a_2^{n-2} \\ \vdots & & & \vdots \\ 1 & a_n - a_1 & \cdots & (a_n - a_1)a_n^{n-2} \end{pmatrix}$$

Also erhält man

$$\begin{aligned} \Delta(a_1, \ldots, a_n) &= \det \begin{pmatrix} a_2 - a_1 & \cdots & (a_2 - a_1)a_2^{n-2} \\ \vdots & & \vdots \\ a_n - a_1 & \cdots & (a_n - a_1)a_n^{n-2} \end{pmatrix} \\ &= (a_2 - a_1) \cdots (a_n - a_1) \det \begin{pmatrix} 1 & \cdots & a_2^{n-2} \\ \vdots & & \vdots \\ 1 & \cdots & a_n^{n-2} \end{pmatrix} \\ &= \prod_{1 < i \leq n} (a_i - a_1) \Delta(a_2, \ldots, a_{n-2}) \\ &\overset{\text{(IV)}}{=} \prod_{1 < i \leq n} (a_i - a_1) \prod_{2 \leq i < j \leq n} (a_j - a_i) = \prod_{1 \leq i < j \leq n} (a_j - a_i). \end{aligned}$$

□

Beweis von Satz (11.1). Die Abbildung ist natürlich surjektiv. Es bleibt zu zeigen, daß die Funktionen $\{1, x, x^2, \ldots\}$ in dem Vektorraum $\mathrm{Abb}(K, K)$ linear unabhängig sind. Ansonsten gibt es $a_0, \ldots, a_n$ (nicht alle zugleich 0) mit

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n \equiv 0 \quad \text{(als Funktion)}.$$

Es seien $x_1, \ldots, x_{n+1} \in K$ verschieden. Dann gilt

(L) $$P(x_i) = a_0 + a_1 x_i + \cdots + a_n x_i^n = 0. \quad (i = 1, \ldots, n+1)$$

Man fasse dies als Gleichungssystem in den Koeffizienten a_i auf. Die Koeffizientenmatrix ist

$$A = \begin{pmatrix} 1 & x_1 & \cdots & x_1^n \\ \vdots & & & \vdots \\ 1 & x_{n+1} & \cdots & x_{n+1}^n \end{pmatrix}.$$

Also ist

$$\det A = \Delta(x_1, \ldots, x_{n+1}) = \prod_{1 \leq i < j \leq n+1} (x_j - x_i) \neq 0.$$

Damit ist das quadratische homogene Gleichungssystem L eindeutig lösbar, und es gilt:

$$a_0 = \cdots = a_n = 0.$$

$\square$

Für den Rest des Paragraphen gelte: $\#K = \infty$.

Jedes Polynom $P \neq 0$ (bzw. falls $\#K = \infty$ auch jede Polynomfunktion) besitzt also eine eindeutige Darstellung

$$P(x) = a_0 + \cdots + a_n x^n \quad \text{mit } a_n \neq 0.$$

Definition.

(i) n heißt der *Grad* von P ($n = \deg P$).

(ii) P heißt *normiert*, falls $a_n = 1$.

(iii) $\deg(0) := -\infty$.

Bemerkung. $\deg(P) = 0 \Leftrightarrow P = a_0 \neq 0$.

Es seien

$$\begin{aligned} P(x) &= a_0 + \cdots + a_n x^n \in K[x] \\ Q(x) &= b_0 + \cdots + b_m x^m \in K[x]. \end{aligned}$$

Dann gilt für die zugehörigen Funktionen

$$P(x) \cdot Q(x) = c_0 + \cdots + c_{n+m} x^{n+m}$$

mit

$$c_k = \sum_{i+j=k} a_i b_j.$$

Insbesondere ist

$$c_0 = a_0 b_0, \quad c_{n+m} = a_n b_m.$$

Man definiert also auch für die Polynome

$$(P \cdot Q)(x) := c_0 + \cdots + c_{n+m} x^{n+m}$$

und nennt dies das *Produkt* von P und Q.

Bemerkung.

$$\deg(P \cdot Q) = \deg P + \deg Q \text{ (Gradformel).}$$

Bemerkung. Mit dieser Multiplikation wird $K[x]$ zu einem kommutativen Ring mit Einselement (*Polynomring*), bzw. zu einer K-Algebra.

Definition. Es sei V ein K-Vektorraum versehen mit einer Multiplikation

$$\begin{array}{rcl} V \times V & \longrightarrow & V, \\ (x, y) & \longmapsto & x \cdot y. \end{array}$$

Dann heißt V eine *Algebra* über K (*K-Algebra*), falls gilt

$$\begin{array}{rcl} (\alpha x + \beta y) \cdot z & = & \alpha(x \cdot z) + \beta(y \cdot z) \\ x \cdot (\alpha y + \beta z) & = & \alpha(x \cdot y) + \beta(x \cdot z). \qquad (x, y, z \in V;\ \alpha, \beta \in K) \end{array}$$

Die Algebra heißt *assoziativ*, falls stets

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z, \qquad (x, y, z \in V)$$

und *kommutativ*, falls stets

$$x \cdot y = y \cdot x. \qquad (x, y \in V)$$

Man sagt V hat ein *Einselement*, falls es ein Element $1 \in V$ gibt mit

$$1 \cdot x = x \cdot 1 = x \qquad (x \in V).$$

Beispiele.
(1) Der Polynomring $K[x]$ ist eine assoziative, kommutative K-Algebra mit Einselement.

(2) Der Vektorraum $\mathrm{Mat}(n; K)$ der $(n \times n)$-Matrizen ist zusammen mit der Matrizenmultiplikation eine assoziative K-Algebra mit Einselement E, die sogenannte *Matrizenalgebra*. Für $n \geq 2$ ist diese Algebra nicht kommutativ.

Homomorphismen zwischen *K-Algebren* werden in der üblichen Weise definiert.

Satz 11.3 (Division mit Rest)
Es seien $0 \neq P, Q \in K[x]$. Dann gibt es eindeutig bestimmte Polynome q, r mit

(i) $P = Qq + r$,

(ii) $\deg r < \deg Q$.

Beweis. *Eindeutigkeit:* Es sei

$$Qq + r = Qq' + r' \quad (\deg r, \deg r' < \deg Q).$$

Also gilt

$$Q(q - q') = r' - r.$$

1. Fall: $q = q' \Rightarrow r = r'$.

2. Fall: $q \neq q'$. Dann ist

$$\underbrace{\deg(r'-r)}_{<\deg Q} = \deg Q + \underbrace{\deg(q-q')}_{\geq 0} \geq \deg Q.$$

Dies ist ein Widerspruch.

Existenz: Wir betrachten

$$M := \{\deg(P - Qp);\ p \in K[x]\}.$$

Die Menge M besitzt ein Minimum in $\mathbb{N} \cup \{-\infty\}$. Es sei $q \in K[x]$ so, daß

$$(*) \qquad \deg(P - Qq) \leq \deg(P - Qp) \text{ für alle } p \in K[x].$$

Es sei ferner

$$r := P - Qq,$$

d. h.

$$P = Qq + r.$$

Es genügt zu zeigen, daß $\deg r < \deg Q$.

Annahme: $\deg r \geq \deg Q$.

Dann haben wir

$$\begin{array}{rcl} Q & = & b_0 + b_1 x + \cdots + b_m x^m \quad (b_m \neq 0) \\ r & = & c_0 + c_1 x + \cdots + c_k x^k \quad (c_k \neq 0, k \geq m). \end{array}$$

Es sei

$$p := q + \frac{c_k}{b_m} x^{k-m} \ \in K[x].$$

Dann ist

$$P - Qp = \underbrace{P - Qq}_{=r} - Q\frac{c_k}{b_m}x^{k-m} = r - Q\frac{c_k}{b_m}x^{k-m}.$$

Es ist also

$$P - Qp = \underbrace{c_k x^k - b_m \frac{c_k}{b_m} x^k}_{=0} + \text{Terme niedrigerer Ordnung.}$$

Daher gilt

$$\deg(P - Qp) < k = \deg r = \deg(P - Qq).$$

Dies steht im Widerspruch zur Wahl von q. □

Aus dem eben gegebenen Beweis kann ein Algorithmus für die Polynomdivision abgeleitet werden, der hier kurz an einem Beispiel beschrieben werden soll.

Beispiel. Gegeben seien

$$P(x) = x^4 + 4x^2 - x + 3, \quad Q(x) = x^2 + x + 1.$$

Man erhält

$$\begin{array}{l}
(x^4 \qquad + 4x^2 - x + 3) : (x^2 + x + 1) = x^2 - x + 4 (= q(x)) \\
\underline{x^4 + x^3 + x^2} \\
\quad -x^3 + 3x^2 - x + 3 \\
\quad \underline{-x^3 - x^2 - x \qquad} \\
\qquad\quad 4x^2 \qquad + 3 \\
\qquad\quad \underline{4x^2 + 4x + 4} \\
\qquad\qquad -4x - 1 (= r(x))
\end{array}$$

Damit ergibt sich

$$(x^4 + 4x^2 - x + 3) = (x^2 + x + 1)(x^2 - x + 4) + (-4x - 1).$$

Definition. Ist $0 \neq P \in K[x]$ und $\lambda \in K$, so daß $P(\lambda) = 0$ gilt, so heißt λ eine *Nullstelle* von P.

Korollar 11.4
Es sei $0 \neq P \in K[x]$ und λ eine Nullstelle von P. Dann gibt es genau ein Polynom $Q \in K[x]$ mit

$$P = Q(x - \lambda).$$

Es ist $\deg Q = \deg P - 1$.

Beweis. Der Divisionssatz (11.3) angewandt auf P und $(x - \lambda)$ gibt

$$P = (x - \lambda)Q + r$$

mit

$$\deg r < \deg(x - \lambda) = 1.$$

Also ist $r = \text{const}$. Da

$$0 = P(\lambda) = r(\lambda),$$

folgt $r = 0$. Die Eindeutigkeit folgt ebenfalls aus dem Divisionssatz. $\square$

Korollar 11.5
Es sei $P \neq 0$. Dann ist die Anzahl der verschiedenen Nullstellen von P höchstens gleich $\deg P$.

Beweis. Es sei $k = \deg P$. Gäbe es verschiedene Nullstellen, $\lambda_1, \dots, \lambda_{k+1}$, so wäre

$$P = (x - \lambda_1) \cdots (x - \lambda_k) Q$$

mit $\deg Q = 0$. Nun ist

$$P(\lambda_{k+1}) = \underbrace{(\lambda_{k+1} - \lambda_1)}_{\neq 0} \cdots \underbrace{(\lambda_{k+1} - \lambda_k)}_{\neq 0} Q(\lambda_{k+1}),$$

also ist $Q(\lambda_{k+1}) = 0$. D. h. $Q = 0$, also $P = 0$. □

Satz 11.6
Jedes Polynom $0 \neq P \in K[x]$ besitzt eine Darstellung

$$P = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r} \cdot Q$$

wobei die $\lambda_1, \dots, \lambda_r$ paarweise verschieden sind, und Q ein Polynom ohne Nullstellen ist. Diese Darstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

Beweis. Die *Existenz* der Darstellung folgt aus obigem.

Eindeutigkeit:

(i) Die λ_i sind genau die Nullstellen von P, liegen also eindeutig fest.

(ii) Die Eindeutigkeit der m_i folgt aus dem folgenden Lemma.

Lemma 11.7
Es seien $P, Q \in K[x]$ Polynome mit $P(\lambda) \neq 0 \neq Q(\lambda)$. Falls

$$(x - \lambda)^m P(x) = (x - \lambda)^n Q(x)$$

für alle $x \in K$, so ist $m = n$.

Beweis. Es sei $m \geq n$. Dann gilt

$$(x - \lambda)^{m-n} P(x) - Q(x) = 0 \quad (\text{für } x \neq \lambda).$$

Da K unendlich viele Elemente enthält, gilt für die Polynome

$$(x - \lambda)^{m-n} P - Q = 0.$$

Falls $m > n$, wäre $Q(\lambda) = 0$, ein Widerspruch. □

Beweis von Satz (11.6) (Fortsetzung).

(iii): Eindeutigkeit von Q:

Es sei

$$(x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r} Q = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r} Q'.$$

Für alle $x \neq \lambda_1, \ldots, \lambda_r$ gilt

$$Q(x) = Q'(x).$$

Also hat $Q - Q'$ unendlich viele Nulstellen, und es folgt $Q = Q'$. □

Definition. Man sagt das Polynom $0 \neq P \in K[x]$ *zerfällt* über K, falls es eine Darstellung

$$P(x) = a(x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r} \quad (a \in K)$$

gibt.

Bemerkungen.

(i) a ist der Koeffizient von x^n.

(ii) Diese Darstellung ist (bis auf die Reihenfolge der Faktoren) eindeutig.

(iii) Es gibt viele Polynome, die *nicht* zerfallen, z. B.

$$P(x) = 1 + x^2 \in \mathbb{R}[x].$$

Aber

$$P(x) = (x - i)(x + i) \in \mathbb{C}[x].$$

Satz 11.8 (Fundamentalsatz der Algebra)
Jedes nicht-konstante Polynom $P \in \mathbb{C}[x]$ besitzt eine Nullstelle.

Beweis. Siehe die Vorlesungen Algebra I, Funktionentheorie I.

Korollar 11.9
Jedes Polynom $P \in \mathbb{C}[x]$ zerfällt.

Ideale in $\mathbf{K}[\mathbf{x}]$

Definition. Eine Teilmenge I eines Ringes R heißt ein *Ideal*, falls gilt:

(I1) $I \subset R$ ist additive Untergruppe.

(I1) Für $P \in I, Q \in R$ gilt $Q \cdot P \in I$ (d. h. $R \cdot I \subset I$).

Beispiel. $I := \langle P_1, \ldots, P_n \rangle := \{Q_1 P_1 + \cdots + Q_n P_n;\ Q_i \in R\}$.

Definition. $I = \langle P_1, \ldots, P_n \rangle$ heißt das von $P_1, \ldots, P_n$ erzeugte Ideal.

Definition. I heißt *Hauptideal*, falls es von einem Element erzeugt wird, d. h. $I = \langle P \rangle$ für ein $P \in R$.

Satz 11.10

(i) $K[x]$ ist ein Hauptidealring, d. h. jedes Ideal ist Hauptideal.

(ii) Zu jedem Ideal $I \neq \{0\}$ in $K[x]$ gibt es genau ein normiertes Polynom P mit $I = \langle P \rangle$.

Beweis. (i): Es sei $I \neq \{0\}$ (sonst ist $I = \langle 0 \rangle$). Dann ist

$$M = \{\deg P;\ 0 \neq P \in I\} \neq \emptyset.$$

Es sei $m := \operatorname{Min} M$ und $P \in I$ mit $\deg P = m$.

Behauptung. $I = \langle P \rangle = K[x] \cdot P$. Es ist $K[x] \cdot P \subset I$ nach (I2). Es bleibt zu zeigen, daß $I \subset K[x] \cdot P$.

Es sei $Q \in I$ beliebig. Dann gibt es eine Darstellung

$$Q = qP + r, \qquad \deg r < \deg P.$$

Ist $r = 0$, so ist $Q \in \langle P \rangle$. Ist $r \neq 0$, so folgt mit (I1) und (I2)

$$r = Q - qP \in I.$$

Wegen

$$0 \leq \deg r < \deg P = m$$

ist dies ein Widerspruch zur Wahl von P.

(ii): Mit P liegt auch aP in I. Also kann man P normiert annehmen. Sei

$$I = \langle P \rangle = \langle P' \rangle$$

mit P, P' normiert. Dann gilt

$$P' = Q'P, \quad P = QP'$$

für geeignete $Q, Q' \in K[x]$. Also

$$(*) \qquad P = QQ'P.$$

Daraus folgt

$$(**) \qquad P(1 - QQ') = 0.$$

Nach dem Gradsatz folgt aus $(*)$

$$\deg QQ' = 0, \text{ also } QQ' = 1 \text{ nach } (**).$$

Wieder nach dem Gradsatz sind Q, Q' konstant. Da P, P' normiert sind, folgt $Q = Q' = 1$, d. h. $P = P'$. □

Definition.

(i) Zwei Polynome $P, P' \in K[x]$; $0 \neq P, P'$ heißen *teilerfremd*, falls aus $P = QR$, $P' = Q'R$ folgt $R = \text{const}$.

(ii) Man sagt, ein Polynom $P \in K[x]$ teilt ein Polynom $Q \in K[x]$ (Schreibweise $P|Q$), falls es ein Polynom $R \in K[x]$ gibt, mit $Q = P \cdot R$.

Korollar 11.11
P, P' seien teilerfremd. Dann gibt es $Q, Q' \in K[x]$ mit $PQ + P'Q' = 1$.

Beweis. Die Menge

$$I := \langle P, P' \rangle = \{QP + Q'P';\ Q, Q' \in K[x]\}$$

ist ein Ideal. Also gibt es ein normiertes Polynom $R \in K[x]$ mit

$$I = \langle R \rangle.$$

Damit gilt

$$P = QR, \quad P' = Q'R$$

für geeignete $Q, Q' \in K[x]$. Da P, P' teilerfremd sind, ist $R = \text{const}$. Aus der Normierung von R folgt $R = 1$. Also ist $1 \in I$, d. h. es gibt $Q, Q' \in K[x]$ mit

$$QP + Q'P' = 1.$$

□

Nilpotente Matrizen

Definition. Eine Matrix $A \in \mathrm{Mat}(n; K)$ heißt *nilpotent*, falls es ein $m \in \mathbb{N}$ mit $A^m = 0$ gibt.

Bemerkung. A nilpotent $\Rightarrow A \notin \mathrm{GL}(n, K)$.

Definition.

(i) A heißt *obere Dreiecksmatrix*, falls für $A = (\alpha_{ij})$ gilt $\alpha_{ij} = 0$ für $i > j$.

(ii) A heißt *echte obere Dreiecksmatrix*, falls $\alpha_{ij} = 0$ für $i \geq j$.

Bemerkung. Obere (bzw. echte obere) Dreiecksmatrizen haben also die Gestalt

$$A = \begin{pmatrix} * & \cdots & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}, \quad \text{bzw.} \quad \begin{pmatrix} 0 & * & \cdots & * \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & * \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Lemma 11.12

(i) A sei nilpotent und $B \sim A$. Dann ist auch B nilpotent.

(ii) Ist A nilpotent, so ist 0 Eigenwert, und es gibt keine weiteren Eigenwerte.

Beweis. (i): $B = W^{-1}AW$. Sei $A^m = 0$. Es ist

$$B^m = (W^{-1}AW)(W^{-1}AW)\cdots(W^{-1}AW) = W^{-1}A^mW = 0.$$

(ii): Ist A nilpotent, so ist $A \notin \mathrm{GL}(n,K)$. Also ist 0 Eigenwert von A. Sei $\lambda \neq 0$ ein weiterer Eigenwert mit Eigenwert $v \neq 0$. Dann gilt

$$\begin{aligned} Av &= \lambda v \quad (v \neq 0) \\ &\Rightarrow \; A^m v = \lambda^m v \neq 0 \text{ für alle } m \\ &\Rightarrow \; A^m \neq 0 \text{ stets.} \end{aligned}$$

□

Satz 11.13
Es sind äquivalent für $A \in \mathrm{Mat}(n;K)$:

(i) A ist nilpotent.

(ii) A ist ähnlich zu einer echten oberen Dreiecksmatrix. In diesem Fall ist $A^n = 0$ und $\chi_A(x) = x^n$.

Beweis. (i)⇒(ii): Induktion nach n:

$n = 1 : A = (0)$.

$n-1 \mapsto n$: Ist A nilpotent, dann ist 0 ein Eigenwert von A, und es gibt einen Vektor $0 \neq v \in \operatorname{Ker} A = \operatorname{Eig}(A,0)$. Ergänzt man v zu einer Basis von K^n, so folgt, daß

$$A \sim \begin{pmatrix} 0 & * \cdots * \\ \vdots & B \\ 0 & \end{pmatrix} =: A'.$$

Es ist

$$A^m = \begin{pmatrix} 0 & * \cdots * \\ \vdots & B^m \\ 0 & \end{pmatrix} = 0.$$

Also ist mit A auch B nilpotent. Nach der Induktionsvoraussetzung gibt es

$$W \in \mathrm{GL}(n-1,K)$$

mit

$$W^{-1}BW = B' = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

Setze

$$\tilde{W} := \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & W \end{array}\right).$$

Dann ist
$$\tilde{W}^{-1}A'\tilde{W} = \left(\begin{array}{c|c} 0 & * \\ \hline 0 & \\ \vdots & W^{-1}BW \\ 0 & \end{array}\right) = \left(\begin{array}{c|ccc} 0 & & * & \\ \hline 0 & 0 & & * \\ \vdots & & \ddots & \\ 0 & 0 & & 0 \end{array}\right).$$
Damit ist auch A zu einer echten oberen Dreiecksmatrix ähnlich.

(ii)⇒(i):
$$A \sim A' = (\alpha'_{ij}) \text{ mit } \alpha'_{ij} = 0 \text{ für } i \geq j.$$
Nach Lemma (11.12) genügt $(A')^n = 0$ zu zeigen.

Behauptung. $(A')^s = (\alpha'^{(s)}_{ij})$ mit $\alpha'^{(s)}_{ij} = 0$ für $i \geq j + 1 - s$.

Induktion nach s:

$s = 1$: Dies ist die Voraussetzung, daß A' echte obere Dreiecksmatrix ist.

$s - 1 \mapsto s$:
$$\alpha'^{(s)}_{ij} = \sum_{k=1}^{n} \alpha'_{ik}\alpha'^{(s-1)}_{kj}.$$
Sei

(1) $$i \geq j + 1 - s.$$

Dann ist
$$k \geq (j+1) - (s-1) \Rightarrow \alpha'^{(s-1)}_{kj} = 0 \text{ nach Induktionsvoraussetzung}$$
oder
$$k \leq (j+1) - s \overset{(1)}{\leq} i \Rightarrow \alpha'_{ik} = 0.$$
Also gilt
$$\alpha'^{(s)}_{ij} = 0.$$
Sei nun
$$A \sim \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix} = A'.$$
Dann ist $(A')^n = 0$ also nach Lemma (11.12) auch $A^n = 0$. Ferner gilt, da $A \sim A'$:
$$\chi_A(x) = \chi_{A'}(x) = \det \begin{pmatrix} x & \cdots & \cdots & * \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & x \end{pmatrix} = x^n.$$

□

Korollar 11.14
Ist A nilpotent, so ist $\operatorname{Spur} A = 0$.

Beweis. Es ist
$$\chi_A(x) = x^n - \operatorname{Spur} A x^{n-1} + \cdots$$
Da $\chi_A(x) = x^n$ gilt, folgt $\operatorname{Spur} A = 0$. □

Definition.

(i) Für $A, B \in \operatorname{Mat}(n; K)$ heißt
$$[A, B] := AB - BA$$
der *Kommutator* von A und B.

(ii) A, B heißen *vertauschbar*, falls $[A, B] = 0$, d. h. falls $AB = BA$.

Satz 11.15
Es sei $[A, B] = 0$. Dann gilt

(i) $(A + B)^m = \sum_{k=0}^{m} \binom{m}{k} A^k B^{m-k}$.

(ii) Ist B nilpotent, so ist auch AB nilpotent.

(iii) Sind A, B nilpotent, so ist auch $A + B$ nilpotent.

Beweis. (i): Dies beweist man wie beim binomischen Lehrsatz.

(ii): $(AB)^m = (AB) \cdots (AB) = A^m B^m = 0$, falls $B^m = 0$.

(iii): $A^m = 0$, $B^p = 0$. Sei $k := m + p$.
$$(A + B)^k \overset{\text{(i)}}{=} \sum_{l=0}^{k} \binom{k}{l} A^l B^{k-l}.$$
Es gibt zwei Fälle:

$l < m \Rightarrow k - l > k - m = (m + p) - m = p \Rightarrow B^{k-l} = 0$,

$l \geq m \Rightarrow A^l = 0$. □

Satz 11.16
Seien $A, N \in \operatorname{Mat}(n; K)$. Es sei N nilpotent und $[A, N] = 0$. Dann gilt
$$\chi_A = \chi_{A+N}.$$
Insbesondere ist $\det(A) = \det(A + N)$.

Beweis. Die letzte Behauptung folgt aus der Gleichheit der charakteristischen Polynome, da für jede Matrix $B \in \operatorname{Mat}(N; K)$ gilt:
$$\det B = (-1)^n \chi_B(0).$$

Wir betrachten nun

$$\det(xE - A) = \chi_A(x) = x^n - \operatorname{Spur} A x^{n-1} + \cdots$$

Da K unendlich viele Elemente enthält, gibt es x mit

$$\det(xE - A) \neq 0.$$

Es sei

$$(*) \qquad M := (xE - A)^{-1} N.$$

(i) M ist nilpotent: Da N nilpotent ist, genügt es nach Satz (**??**) (ii), zu zeigen daß $[(xE - A)^{-1}, N] = 0$. Es ist

$$(xE - A)N = xEN - AN \overset{[A,N]=0}{=} xNE - NA = N(xE - A).$$

Multiplikation mit $(xE - A)^{-1}$ von beiden Seiten ergibt

$$N(xE - A)^{-1} = (xE - A)^{-1} N,$$

d. h.

$$[(xE - A)^{-1}, N] = 0.$$

(ii) $\det(E - M) = 1$: Da M nilpotent ist, gibt es ein $W \in \mathrm{GL}(n, K)$ mit

$$W^{-1} M W = \text{echte obere Dreiecksmatrix.}$$

Also

$$\begin{aligned}
\det(E - M) &= \det(W^{-1}(E - M)W) \\
&= \det(E - W^{-1} M W) \\
&= \det \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = 1.
\end{aligned}$$

(iii) Beweis der Aussage: Es sei zunächst $x \in K$ mit $\chi_A(x) \neq 0$. Dann gilt

$$\begin{aligned}
\chi_{A+N}(x) &= \det(xE - (A + N)) \\
&\overset{(*)}{=} \det((xE - A) - (xE - A)M) \\
&= \det((xE - A)(E - M)) \\
&\overset{\text{(ii)}}{=} \det(xE - A) = \chi_A(x)
\end{aligned}$$

Da K unendlich viele Elemente hat, folgt $\chi_{A+N} = \chi_A$.

Idempotente Matrizen

Definition. $A \in \mathrm{Mat}(n;k)$ heißt *idempotent*, falls $A^2 = A$.

Lemma 11.17
Die Matrix A sei idempotent. Dann ist

$$K^n = \mathrm{Im}\, A \oplus \mathrm{Ker}\, A.$$

Die beiden einzig möglichen Eigenwerte sind $0, 1$.

Beweis. Nach der Dimensionsformel gilt

$$\dim \mathrm{Im}\, A + \dim \mathrm{Ker}\, A = n.$$

Also genügt es zu zeigen, daß

$$\mathrm{Im}\, A + \mathrm{Ker}\, A = K^n.$$

Es sei $x \in K^n$. Wir setzen

$$x_1 := Ax \in \mathrm{Im}\, A;\ x_0 := x - x_1 = x - Ax.$$

Dann ist

$$Ax_0 = Ax - A^2x = Ax - Ax = 0$$

also ist $x_0 \in \mathrm{Ker}\, A$ und

$$x = x_1 + x_0.$$

Wir haben also für jedes $x \in K^n$ eine eindeutige Darstellung $x = x_1 + x_0$ mit x_0, x_1 wie oben beschrieben. Dann ist

$$Ax = A(x_0 + x_1) = x_1,$$

also ist

$$Ax = \lambda x \Leftrightarrow \begin{cases} x_0 = 0, & \lambda = 1 \\ x_1 = 0, & \lambda = 0. \end{cases}$$

□

Die geometrische Deutung der zu einer idempotenten Matrix gehörigen Abbildung $h_A : K^n \to K^n, x \mapsto Ax$ ist, daß es eine *Projektion* auf den Unterraum $\mathrm{Im}\, A$ ist.

Satz 11.18
Es sei $0 \neq A \in \mathrm{Mat}(n;K)$ und $r = \mathrm{Rang}\, A > 0$. Dann sind äquivalent:

(i) A ist idempotent.

(ii) $A \sim \left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right)$ (wobei $E^{(r)}$ die Einheitsmatrix der Größe r bezeichnet).

In diesem Fall ist $\chi_A(x) = (x-1)^r x^{n-r}$ und $\operatorname{Spur} A = r \cdot 1$.

Bemerkung. Bezüglich der Basis, die zur Matrix $\left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right)$ gehört, ist der zugehörige Endomorphismus $K^n \to K^n$ gegeben durch

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0).$$

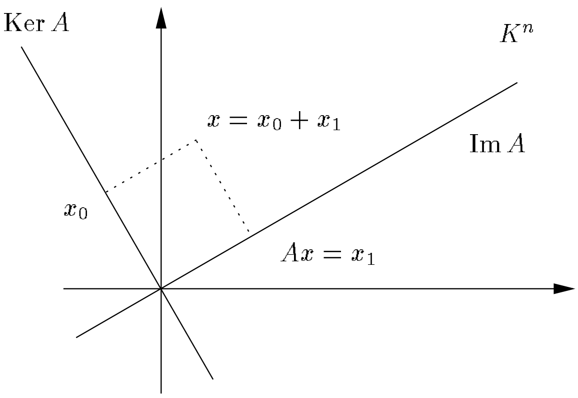


Abb. 1: Projektion auf den Unterraum Im A

Bemerkung. $r \cdot 1 = \underbrace{1 + \cdots + 1}_{r\text{-mal}}$.

Beweis. (ii)⇒(i): Es ist

$$A = W^{-1} \left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right) W \qquad (\text{mit } W \in \mathrm{GL}(n, K)).$$

Also

$$\begin{aligned} A^2 &= W^{-1} \left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right) W W^{-1} \left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right) W \\ &= W^{-1} \left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right) W = A. \end{aligned}$$

(i)⇒(ii): Wir wählen Basen

$$\begin{aligned} b_1, \dots, b_r &\in \operatorname{Im} A \\ b_{r+1}, \dots, b_n &\in \operatorname{Ker} A. \end{aligned}$$

Dann ist $b_1, \dots, b_n$ Basis von K^n. Es gilt

$$Ab_i = 0 \qquad \text{für } i = r+1, \dots, n,$$

also sind $b_{r+1}, \ldots, b_n$ Eigenvektoren zum Eigenwert 0. Für $b_1, \ldots, b_r \in \operatorname{Im} A$ gilt:

$$b_i = Ab_i' \qquad (b_i' \in K^n).$$

Also

$$Ab_i = A^2 b_i' = Ab_i' = b_i. \qquad (i = 1, \ldots, r).$$

D. h. $b_1, \ldots, b_r$ sind Eigenvektoren zum Eigenwert 1. Nach Satz (I.8.1) ist A diagonalisierbar, und es gilt

$$A \sim \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} =: B,$$

wobei die Anzahl der Einträge 1 gleich r ist.

Es gelte nun (i), (ii). Dann ist

$$\chi_A(x) \;=\; \chi_B(x) \text{ mit } B = \left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right).$$

Also folgt

$$\begin{aligned} \chi_A(x) \;&=\; \det(xE - B) = \begin{pmatrix} x-1 & & & & & \\ & \ddots & & & & \\ & & x-1 & & & \\ & & & x & & \\ & & & & \ddots & \\ & & & & & x \end{pmatrix} \\ &=\; (x-1)^r x^{n-r}. \end{aligned}$$

Daraus folgt

$$\operatorname{Spur} A = \operatorname{Spur} B = r \cdot 1.$$

□

Korollar 11.19
A sei idempotent vom Rang r. Dann gilt

$$\det(E - xA) = (1 - x)^r.$$

Beweis. Nach Satz (11.18) gibt es $W \in \mathrm{GL}(n, K)$ mit

$$W^{-1}AW = \left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right).$$

Also ist

$$\begin{aligned} \det(E - xA) &= \det(W^{-1}(E - xA)W) \\ &= \det(E - x\left(\begin{array}{c|c} E^{(r)} & 0 \\ \hline 0 & 0 \end{array}\right)) \\ &= \det \begin{pmatrix} 1-x & & & & & \\ & \ddots & & & & \\ & & 1-x & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} = (1-x)^r. \end{aligned}$$

$\square$

Zerfallende Matrizen

Definition. Man sagt, eine Matrix $A \in \mathrm{Mat}(n; K)$ *zerfällt*, falls das charakteristische Polynom von A zerfällt, d. h.

$$\chi_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}.$$

(In einer solchen Darstellung nehmen wir stets an, daß $\lambda_1, \ldots \lambda_r$ paarweise verschieden sind.)

Bemerkung. Die λ_i sind genau die *Eigenwerte* von A.

Definition. Die m_i heißen die *algebraischen Vielfachheiten* der Eigenwerte λ_i.

Beispiele.
(1) Alle Matrizen $A \in \mathrm{Mat}(n; \mathbb{C})$ zerfallen wegen des Fundamentalsatzes der Algebra.
(2) Ist A nilpotent so zerfällt A, da $\chi_A(x) = x^n$.
(3) Ist A idempotent, so zerfällt A, da $\chi_A(x) = (x-1)^r x^{n-r}$.

Definition. Eine *verallgemeinerte Jordanmatrix (Jordanblock)* zum Eigenwert $\lambda \in K$ ist eine Matrix der Form

$$J = \lambda E + N$$

wobei N echte obere Dreiecksmatrix ist. Damit hat J folgende Gestalt:

$$J = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}.$$

Definition. Die *Jordanmatrix* (der Größe n) zum Eigenwert λ ist die Matrix

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

Bemerkung. In jedem Fall gilt $\chi_J(x) = (x - \lambda)^n$.

Satz 11.20 (Struktursatz, 1. Form der Jordanschen Normalform)
Es sei A eine zerfallende Matrix mit

$$\chi_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k},$$

wobei $\lambda_1, \ldots, \lambda_k$ paarweise verschieden sind. Dann ist A ähnlich zu einer Blockmatrix

$$\begin{pmatrix} \boxed{A_1} & & & 0 \\ & \boxed{A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_k} \end{pmatrix}$$

wobei $A_i \in \operatorname{Mat}(m_i; K)$ verallgemeinerte Jordanmatrizen zum Eigenwert λ_i sind.

Bemerkung. Wir werden später eine Normalform herleiten wobei nur noch (echte) Jordanmatrizen vorkommen.

Korollar 11.21
Für $A \in \operatorname{Mat}(n; K)$ sind äquivalent

(i) A zerfällt.

(ii) A ist zu einer oberen Dreiecksmatrix ähnlich.

Beweis. (ii)⇒(i):

$$A \sim A' = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Dann ist

$$\chi_A = \chi_{A'} = \det \begin{pmatrix} x - \lambda_1 & & * \\ & \ddots & \\ 0 & & x - \lambda_n \end{pmatrix} = (x - \lambda_1) \cdots (x - \lambda_n).$$

(i)⇒(ii): Aus dem Struktursatz. □

Korollar 11.22
Sei $A \in \text{Mat}(n; K)$. Falls A zerfällt, ist $\text{Spur}\, A$ gleich der Summe der Eigenwerte (mit Vielfachheit).

Beweis. A zerfalle. Dann gilt nach dem Struktursatz (11.20)

$$A \sim A' = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_1 \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \lambda_k & & * \\ & \ddots & \\ & & \lambda_k \end{matrix}} \end{pmatrix}$$

Hieraus folgt dann

$$\text{Spur}\, A = \text{Spur}\, A' = m_1\lambda_1 + \cdots + m_k\lambda_k.$$

□

Korollar 11.23
Zerfällt $A \in \text{Mat}(n; K)$ in n verschiedene Linearformen (d. h. hat A genau n verschiedene Eigenwerte), so ist A diagonalisierbar.

Beweis. Sofort aus dem Struktursatz (11.20). □

Bevor wir den Struktursatz beweisen können, benötigen wir das

Lemma 11.24
Es seien $M \in \text{Mat}(p; K)$; $N \in \text{Mat}(q; K)$ nilpotente Matrizen. Ferner sei $C \in \text{Mat}(p, q; K)$. Dann gibt es genau ein $X \in \text{Mat}(p, q; K)$ mit

$$(*) \qquad X = MX - XN + C.$$

Beweis. Wir betrachten $(*)$ als lineares Gleichungssystem. Die Unbekannten sind x_{ij}.

$$\begin{aligned} \#\text{Gleichungen} &= pq \\ \#\text{Unbekannten} &= pq. \end{aligned}$$

Also ist $(*)$ genau dann eindeutig lösbar, wenn das zugehörige homogene Gleichungssystem

$$(**) \qquad X = MX - XN$$

nur $X = 0$ als Lösung hat.

Behauptung. X sei Lösung von $(**)$. Dann gilt für alle $s \in \mathbb{N}$:

$$(***) \qquad X = \sum_{\sigma=0}^{s} (-1)^{\sigma} \binom{s}{\sigma} M^{s-\sigma} X N^{\sigma}.$$

Daraus folgt das Lemma: Da M, N nilpotent gibt es $m, n \in \mathbb{N}$ mit $M^m = 0$ und $N^n = 0$. Sei $s = m + n$. Ist $s = (s - \sigma) + \sigma$, so ist $(s - \sigma) \geq m$ oder $\sigma \geq n$. Also folgt $X = 0$.

Beweis der Behauptung. Wir machen Induktion nach s.

$s = 1$: Für $s = 1$ ist $(***)$ gerade die Gleichung $(**)$.

$s \mapsto s + 1:$ Nach Induktionsvoraussetzung gilt

$$X = \sum_{\sigma=0}^{s} (-1)^{\sigma} \binom{s}{\sigma} M^{s-\sigma} X N^{\sigma}.$$

Da X $(**)$ erfüllt:

$$\begin{aligned} X &\overset{(**)}{=} MX - XN \\ &\overset{(\mathrm{IV})}{=} \sum_{\sigma=0}^{s} (-1)^{\sigma} \binom{s}{\sigma} M^{(s+1)-\sigma} X N^{\sigma} - \sum_{\sigma=0}^{s} (-1)^{\sigma} \binom{s}{\sigma} M^{s-\sigma} X N^{\sigma+1} \\ &= \sum_{\sigma=0}^{s+1} (-1)^{\sigma} \left(\binom{s}{\sigma} + \binom{s}{\sigma - 1} \right) M^{(s+1)-\sigma} X N^{\sigma} \\ &= \sum_{\sigma=0}^{s+1} (-1)^{\sigma} \binom{s+1}{\sigma} M^{(s+1)-\sigma} X N^{\sigma}. \end{aligned}$$

□

Beweis des Struktursatzes (11.20) Wir machen nun Induktion nach n.

$n = 1$: $A = (\lambda_1)$. Trivial.

$n - 1 \mapsto n$: Es sei A gegeben mit

$$(*) \qquad \chi_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}.$$

Nach Satz (I.8.5) gilt:

$$A \sim \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & B \\ 0 & \end{pmatrix} =: A'.$$

Es ist

$$\chi_A(x) = (x - \lambda_1) \chi_B(x).$$

Wegen $(*)$ folgt

$$\chi_B(x) = (x-\lambda_1)^{m_1-1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}.$$

(Dies gilt zunächst für $x \neq \lambda_i$, da $\#K = \infty$, aber dann für alle x und damit für die Polynome.) Also zerfällt auch B. Nach Induktionsvoraussetzung gilt

$$B \sim \begin{pmatrix} A_1^* & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix} =: B'$$

wobei

$$\begin{array}{rcl} A_i & = & \lambda_i E_{m_i} + N_i \in \mathrm{Mat}(m_i;K); \quad i = 2,\ldots,k \\ A_1^* & = & \lambda_1 E_{m_1-1} + N_1 \in \mathrm{Mat}(m_1-1;K) \end{array}$$

verallgemeinerte Jordanblöcke sind. (Falls $m_1 = n$ gilt, ist man an dieser Stelle fertig.)

Behauptung.

$$A \sim \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & B' \\ 0 & \end{pmatrix}$$

Denn: Es sei $W' \in \mathrm{GL}(n-1,K)$ mit

$$W'^{-1}BW' = B'.$$

Dann definiere man

$$W := \begin{pmatrix} 1 & 0\cdots 0 \\ 0 & \\ \vdots & W' \\ 0 & \end{pmatrix} \in \mathrm{GL}(n,K).$$

Dann gilt

$$W^{-1} = \begin{pmatrix} 1 & 0\cdots 0 \\ 0 & \\ \vdots & (W')^{-1} \\ 0 & \end{pmatrix}$$

und wir erhalten

$$W^{-1}A'W = \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & (W')^{-1}BW' \\ 0 & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \cdots * \\ 0 & \\ \vdots & B' \\ 0 & \end{pmatrix}.$$

Also hat man bisher gefunden

$$A \sim \begin{array}{c} \begin{pmatrix} A_1 & C_2 & \cdots & \cdots & C_k \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_k \end{pmatrix} \\ \underbrace{\quad}_{m_1} \ \underbrace{\quad}_{m_2} \ \cdots \ \cdots \ \underbrace{\quad}_{m_k} \end{array} \begin{array}{l} \}m_1 \\ \}m_2 \\ \vdots \\ \vdots \\ \}m_k \end{array} =: C$$

wobei

$$A_i = \lambda_i E_{m_i} + N_i \in \mathrm{Mat}(m_i; K) \qquad (i = 1, \dots, k)$$

verallgemeinerte Jordanblöcke sind, sowie

$$C_i = \begin{pmatrix} * \cdots * \\ 0 \end{pmatrix} \in K(m_1, m_i; K) \qquad (i = 2, \dots, k).$$

Wir müssen nun die Kästchen C_i wegtransformieren. Dazu wählen wir den Ansatz:

$$\tilde{W} = \begin{pmatrix} E_{m_1} & \tilde{W}_2 & \cdots & \cdots & \tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix}$$

mit

$$\tilde{W}_i \in \mathrm{Mat}(m_1, m_i; K).$$

Behauptung.

$$\tilde{W}^{-1} = \begin{pmatrix} E_{m_1} & -\tilde{W}_2 & \cdots & \cdots & -\tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix}$$

Denn: Es gilt

$$\begin{pmatrix} E_{m_1} & -\tilde{W}_2 & \cdots & \cdots & -\tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix} \begin{pmatrix} E_{m_1} & \tilde{W}_2 & \cdots & \cdots & \tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix}$$

$$= \begin{pmatrix} E_{m_1} & \tilde{W}_2 - \tilde{W}_2 & \cdots & \cdots & \tilde{W}_k - \tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix} = E^{(n)}.$$

Also hat man

$$\begin{aligned} \tilde{W}^{-1}C\tilde{W} &= \begin{pmatrix} E_{m_1} & -\tilde{W}_2 & \cdots & \cdots & -\tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix} \begin{pmatrix} A_1 & C_2 & \cdots & \cdots & C_k \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_k \end{pmatrix} \cdot \\ &\quad \cdot \begin{pmatrix} E_{m_1} & \tilde{W}_2 & \cdots & \cdots & \tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix} \\ &= \begin{pmatrix} A_1 & C_2 - \tilde{W}_2 A_2 & \cdots & \cdots & C_k - \tilde{W}_k A_k \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_k \end{pmatrix} \begin{pmatrix} E_{m_1} & \tilde{W}_2 & \cdots & \cdots & \tilde{W}_k \\ 0 & E_{m_2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & E_{m_k} \end{pmatrix} \\ &= \begin{pmatrix} A_1 & A_1\tilde{W}_2 + C_2 - \tilde{W}_2 A_2 & \cdots & \cdots & A_1\tilde{W}_k + C_k - \tilde{W}_k A_k \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & A_k \end{pmatrix}. \end{aligned}$$

Es sei

$$D_r := C_r + A_1\tilde{W}_r - \tilde{W}_r A_r \quad (r = 2, \dots, k).$$

Es gilt

$$A_r = \lambda_r E_{m_r} + N_r \quad (N_r \text{ echte obere Dreiecksmatrix}).$$

Damit ist

$$\begin{aligned} D_r &= C_r + (\lambda_1 E_{m_1} + N_1)\tilde{W}_r - \tilde{W}_r(\lambda_r E_{m_r} + N_r) \\ &= C_r + \lambda_1\tilde{W}_r - \lambda_r\tilde{W}_r + N_1\tilde{W}_r - \tilde{W}_r N_r. \end{aligned}$$

Also

$$D_r = (\lambda_1 - \lambda_r)\tilde{W}_r + N_1\tilde{W}_r - \tilde{W}_r N_r + C_r.$$

Da $\lambda_1 \neq \lambda_r$, gilt $D_r = 0$ genau wenn:

$$\tilde{W}_r = \underbrace{\frac{1}{\lambda_r - \lambda_1}N_1}_{=:M}\tilde{W}_r - \tilde{W}_r \underbrace{\frac{1}{\lambda_r - \lambda_1}N_r}_{=:N} + \underbrace{\frac{1}{\lambda_r - \lambda_1}C_r}_{=:C'}.$$

Da N_1, N_r nilpotent, sind auch M, N nilpotent und die Gleichung

$$\tilde{W}_r = M\tilde{W}_r - \tilde{W}_r N + C'$$

hat nach Lemma (11.24) genau eine Lösung. Diese $\tilde{W}_r$ tun es. □

Das Minimalpolynom

Es sei daran erinnert, daß $\mathrm{Mat}(n;K)$ ein Vektorraum der Dimension n^2 ist. Wie üblich setzen wir für $A \in \mathrm{Mat}(n;K)$:

$$A^n = A \cdots A \quad (n\text{-mal}), \quad A^0 = E.$$

Dann ordnen wir der Matrix A wie folgt einen Unterraum in $\mathrm{Mat}(n;K)$ zu:

Definition. $K[A] := \mathrm{Span}\{A^n;\ n \in \mathbb{N} \cup \{0\}\}$.

Bemerkung. Mit dem Matrizenprodukt ist $K[A]$ ein (kommutativer) Ring und sogar eine kommutative K-Algebra.

Man definiert

$$\begin{array}{rrcl} \varphi_A : & K[x] & \longrightarrow & K[A] \\ & P = \sum_{\nu=0}^n a_\nu x^\nu & \longmapsto & P(A) := \sum_{\nu=0}^n a_\nu A^\nu. \end{array}$$

Bemerkung.

(i) φ_A ist surjektiv.

(ii) φ_A ist sowohl ein Homomorphismus von K-Vektorräumen, als auch ein Ringhomomorphismus:

$$\begin{aligned} (P+Q)(A) &= P(A) + Q(A) \\ (\lambda P)(A) &= \lambda P(A) \\ (P \cdot Q)(A) &= P(A) \cdot Q(A). \end{aligned}$$

Genauer gesagt ist φ_A ein K-Algebrenhomomorphismus.

Wir betrachten nun

$$I_A := \{P \in K[x];\ P(A) = 0\} = \operatorname{Ker} \varphi_A.$$

Nach obigem folgt, daß I_A ein *Ideal* ist. Da $K[x]$ ein Hauptidealring ist, gibt es (falls $I_A \neq \{0\}$ ist) genau ein normiertes Polynom $\mu_A \in K[x]$ mit

$$I_A = \langle \mu_A \rangle.$$

Satz 11.25
Es gibt ein normiertes Polynom $0 \neq \mu_A \in K[x]$ mit folgenden Eigenschaften

(i) $\mu_A(A) = 0$.

(ii) Ist $P \in K[x]$ ein Polynom mit $P(A) = 0$, so ist $P = Q \cdot \mu_A$; d. h. μ_A teilt P.

Ferner gilt:

(iii) Unter allen normierten Polynomen mit $P(A) = 0$ hat μ_A minimalen Grad. Dadurch ist μ_A eindeutig bestimmt.

(iv) $\deg \mu_A = \dim K[A]$.

Definition. μ_A heißt das *Minimalpolynom* von A.

Beweis. Wir zeigen zunächst $I_A \neq \{0\}$. Es ist

$$K[A] \subset \operatorname{Mat}(n; K),$$

also gilt $\dim K[A] \leq n^2 =: N$. Daher sind die Matrizen

$$E, A, A^2, \ldots, A^N$$

linear abhängig, d. h. es gibt $a_0, \ldots, a_N \in K$, nicht alle 0, mit

$$a_0 E + a_1 A + \cdots + a_N A^N = 0.$$

Also ist

$$0 \neq P(x) = a_0 + \cdots + a_N x^N \in I_A.$$

Damit ist nach Satz (11.10)

$$I_A = \langle \mu_A \rangle$$

für ein eindeutig bestimmtes normiertes Polynom $\mu_A \neq 0$. Dies zeigt (i)–(iii).

(iv): Es sei $m := \dim K[A]$. Dann sind

$$E, A, \ldots, A^m$$

linear abhängig. Wie oben zeigt man, daß es ein (normiertes) Polynom $P \in K[x]$, $\deg P \leq m$ gibt mit $P(A) = 0$. Also ist

$$\deg \mu_A \leq m = \dim K[A].$$

Es sei umgekehrt

$$m' := \deg \mu_A.$$

Behauptung. $\dim K[A] \leq m'$. Wir betrachten

$$V := \mathrm{Span}(E, A, \dots, A^{m'-1}).$$

Nach Definition von m' ist $\dim V = m'$. Es genügt zu zeigen, daß

$$K[A] \subset V.$$

Dazu ist zu zeigen, daß

$$A^s \in V \qquad \text{für } s \geq m'.$$

Es sei

$$\mu_A = x^{m'} + a_{m'-1}x^{m'-1} + \dots + a_0.$$

Da $\mu_A(A) = 0$ folgt

$$A^{m'} = -a_{m'-1}A^{m'-1} - \dots - a_0 E \in V.$$

Durch Multiplikation mit A auf beiden Seiten folgt dann aber auch

$$A^{m'+1} = -a_{m'-1}A^{m} - \dots - a_0 A \in V.$$

Induktion liefert die Behauptung. □

Korollar 11.26
Ist $m = \deg \mu_A$, so ist $E, A, \dots, A^{m-1}$ eine Basis von $K[A]$.

Beweis. $E, \dots, A^{m-1}$ sind wegen $m = \deg \mu_A$ linear unabhängig. Die Behauptung folgt dann aus $m = \dim K[A]$. □

Korollar 11.27
(i) A ist invertierbar $\Leftrightarrow \mu_A(0) \neq 0$.

(ii) Ist A invertiertbar, so ist $A^{-1} \in K[A]$.

Beweis. Es sei

$$\mu(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0.$$

Also ist

$$A^m + a_{m-1}A^{m-1} + \cdots + a_1 A = -a_0 E$$

und damit

$$A(\underbrace{A^{m-1} + a_{m-1}A^{m-2} + \cdots + a_1 E}_{=:B}) = -a_0 E.$$

Ist $a_0 = \mu(0) \neq 0$, so ist

$$AB = -a_0 E$$

und damit

$$A(-\frac{1}{a_0}B) = E,$$

d. h. $A \in \mathrm{GL}(n, K)$.

Ist $A \in \mathrm{GL}(n, K)$, so folgt

$$B = -a_0 A^{-1}.$$

Nun ist $B \neq 0$ (sonst wäre μ_A nicht das Minimalpolynom). Also ist $a_0 = \mu(0) \neq 0$.

(ii): Nach obigem ist $A^{-1} = -\frac{1}{a_0}B \in K[A]$. □

Korollar 11.28
Ist $A \in \mathrm{GL}(n, K)$, so gibt es ein Polynom $P \in K[x]$ mit $P(0) = 0$, so daß $P(A) = E$.

Beweis. Ist $A \in \mathrm{GL}(n, K)$, so ist nach Korollar (11.27) (ii) auch $A^{-1} \in K[A]$. Also gibt es ein $Q(x) \in K[x]$ mit

$$A^{-1} = Q(A).$$

Daraus folgt, daß

$$E = A \cdot Q(A).$$

D. h. für

$$P(x) := x \cdot Q(x)$$

gilt $P(A) = E$. □

Lemma 11.29 (Invarianz)

$$A \sim B \Rightarrow \mu_A = \mu_B.$$

Beweis. Für $P \in K[x]$ und $W \in \mathrm{GL}(n, K)$ gilt

$$P(A) = 0 \Leftrightarrow P(W^{-1}AW) = 0.$$

Dies folgt, da

$$\begin{aligned} P(W^{-1}AW) &= a_0E + a_1W^{-1}AW + \cdots + a_m(W^{-1}AW)^m \\ &= a_0W^{-1}EW + a_1W^{-1}AW + \cdots + a_m(W^{-1}A^mW) \\ &= W^{-1}(P(A))W. \end{aligned}$$

□

Bemerkung. $\mu_a = \mu_b \overset{\text{i. a.}}{\not\Rightarrow} A \sim B$.

Beispiel.

$$A = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix}.$$

Dann ist

$$\begin{aligned} (A-E)(A-2E) &= \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} -1 & & \\ & 0 & \\ & & 0 \end{pmatrix} = 0, \\ (B-E)(B-2E) &= \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix} \begin{pmatrix} -1 & & \\ & -1 & \\ & & 0 \end{pmatrix} = 0. \end{aligned}$$

Also ist, da $A - E, A - 2E \neq 0$ und $B - E, B - 2E \neq 0$

$$\mu_A(x) = \mu_B(x) = (x-1)(x-2).$$

Andererseits gilt

$$\begin{aligned} \chi_A(x) &= \det \begin{pmatrix} x-1 & & \\ & x-2 & \\ & & x-2 \end{pmatrix} = (x-1)(x-2)^2, \\ \chi_B(x) &= \det \begin{pmatrix} x-1 & & \\ & x-1 & \\ & & x-2 \end{pmatrix} = (x-1)^2(x-2). \end{aligned}$$

D. h. $\chi_A \neq \chi_B$, und damit $A \not\sim B$.

Eigenwerte und Minimalpolynom

Satz 11.30
Für $A \in \mathrm{Mat}(n; K)$ sind äquivalent

(i) λ ist Eigenwert von A.

(ii) $\mu_A(\lambda) = 0$.

Beweis. (i)$\Rightarrow$(ii): Es sei $v \neq 0$ ein Vektor mit

$$Av = \lambda v \qquad (v \neq 0).$$

Dann gilt $A^k v = \lambda^k v$ und also für jedes $P \in K[x]$:

$$\begin{aligned} P(A)v &= (a_0 E + \cdots + a_n A^n)v = a_0 v + a_1 \lambda v + \cdots + a_n \lambda^n v \\ &= P(\lambda)v. \end{aligned}$$

Ist $P = \mu_A$, so ist $\mu_A(A) = 0$, also $\mu_A(\lambda)v = 0$. Da $v \neq 0$ ist, folgt $\mu_A(\lambda) = 0$.

(ii)$\Rightarrow$(i): Es sei $\mu_A(\lambda) = 0$. Dann gilt

$$\mu_A(x) = (x - \lambda)P(x) \quad \text{für ein geeignetes } P \in K[x].$$

Nun ist, da μ das Minimalpolynom von A ist

$$P(A) \neq 0,$$

d. h. es gibt einen Vektor v mit

$$0 \neq v \in \operatorname{Im} P(A).$$

Also gibt es ein $w \in K^n$ mit

$$0 \neq v = P(A)w.$$

Um zu zeigen, daß λ Eigenwert von A ist, genügt es zu zeigen, daß

$$Av = \lambda v.$$

Dies gilt, da

$$Av - \lambda v = (A - \lambda E)v = (A - \lambda E)P(A)w = \mu_A(A)w = 0.$$

$\square$

Folgerung. χ_A und μ_A haben dieselben Nullstellen.

Wir werden später sehen, daß μ_A das charakteristische Polynom χ_A teilt, d. h.

$$\chi_A = \mu_A \cdot R$$

für ein Polynom $R \in K[x]$.

Wir erinnern daran, daß eine verallgemeinerte Jordanmatrix zum Eigenwert λ eine Matrix der folgenden Form ist:

$$J = \lambda E + N = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda \end{pmatrix},$$

wobei N echte obere Dreiecksmatrix ist.

Satz 11.31
Ist J eine verallgemeinerte Jordanmatrix zum Eigenwert λ, so ist für ein Polynom $P \in K[x]$ die Matrix $P(J)$ eine verallgemeinerte Jordanmatrix zu $P(\lambda)$.

Beweis. Es gilt

$$P(x) = a_0 + \cdots + a_n x^n.$$

Daraus folgt

$$\begin{aligned} P(J) &= a_0 E + a_1 J + \cdots + a_n J^n = a_0 E + a_1(\lambda E + N) + \cdots + a_n(\lambda E + N)^n \\ &= P(\lambda) \cdot E + N'. \end{aligned}$$

Dabei ist N' eine Linearkombination von $N, \ldots, N^n$, also auch wieder eine echte obere Dreiecksmatrix.

Satz 11.32
Es sei $A \in \mathrm{Mat}(n; K)$ eine zerfallende Matrix mit charakteristischem Polynom (die λ_i seien paarweise verschieden):

$$\chi_A = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}.$$

Es sei

$$w_A := (x - \lambda_1) \cdots (x - \lambda_r).$$

Dann ist $w_A(A)$ nilpotent.

Beweis. Es ist

$$A \sim \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix} =: A' \qquad A_i = \lambda_i E + N_i \in \mathrm{Mat}(m_i; K),$$

wobei die N_i echte obere Dreiecksmatrizen sind. (Wir werden hier und im folgenden oft E statt E_{m_i} schreiben, um die Notation zu vereinfachen.)

Da $A \sim A'$, ist auch $w_A(A) \sim w_A(A')$. Also genügt es nach Satz (11.12) (i) zu zeigen, daß $w_A(A')$ nilpotent ist. Nun ist

$$w_A(A') = \begin{pmatrix} w_A(A_1) & & \\ & \ddots & \\ & & w_A(A_r) \end{pmatrix}.$$

Da A_i verallgemeinerte Jordanmatrix zum Eigenwert λ_i ist, ist nach (11.31) $w_A(A_i)$ verallgemeinerte Jordanmatrix zum Eigenwert $w_A(\lambda_i) = 0$. Also gilt

$$w_A(A_i) = \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix}.$$

Damit ist $w_A(A_i)$ eine echte obere Dreicksmatrix, damit auch $w_A(A')$, also ist die Matrix $w_A(A')$ nilpotent. $\square$

Theorem 11.33 (Satz von Cayley)
Es sei $K \subset \mathbb{C}$. Dann gilt für jede Matrix $A \in \mathrm{Mat}(n; K)$:

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

Folgerung 11.34
Das Minimalpolynom μ_A teilt χ_A, d. h.

$$\chi_A = R \cdot \mu_A$$

für ein geeignetes Polynom $R \in K[x]$.

Beweis. Da $K \subset \mathbb{C}$ ist, kann man $A \in \mathrm{Mat}(n; \mathbb{C})$ annehmen. Über $\mathbb{C}$ zerfällt χ_A. Also gilt

$$A \sim A' = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$$

wobei die A_i verallgemeinerte Jordanblöcke zum Eigenwert λ_i sind. Es genügt nun zu zeigen, daß $\chi_A(A') = \chi_{A'}(A') = 0$. Es ist

$$\chi_{A'}(A') = \begin{pmatrix} \chi_{A'}(A_1) & & \\ & \ddots & \\ & & \chi_{A'}(A_r) \end{pmatrix}.$$

Es genügt nun, daß

$$\chi_{A'}(A_i) = (A_i - \lambda_1 E)^{m_1} \cdots (A_i - \lambda_i E)^{m_i} \cdots (A_i - \lambda_r E)^{m_r}$$

Null wird, wobei

$$\chi_{A'} = \chi_A = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}.$$

Nun ist

$$A_i - \lambda_i E = N_i$$

eine echte obere Dreiecksmatrix der Größe m_i. Daraus folgt

$$(A_i - \lambda_i E)^{m_i} = N_i^{m_i} = 0.$$

Also ist $\chi_{A'}(A_i) = 0$ für $i = 1, \dots, r$ und damit auch $\chi_{A'}(A') = 0$. □

Bemerkung. Der Satz gilt für alle Körper K. Man braucht dazu die folgende Aussage der Algebra: Es sei $P \in K[x]$. Dann gibt es einen Körper $K' \supset K$ (*Zerfällungskörper*), so daß $P \in K'[x]$ zerfällt.

Satz 11.35
Es sei $A \in \mathrm{Mat}(n; K)$ zerfallend mit

$$\chi_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}$$

mit paarweise verschiedenen Nullstellen λ_i.

Dann sind äquivalent:

(i) A ist diagonalisierbar.

(ii) Jedes Element von $K[A]$ ist diagonalisierbar.

(iii) Ist $B \in K[A]$ nilpotent, so folgt $B = 0$.

(iv) $\mu_A(x) = (x - \lambda_1) \cdots (x - \lambda_r)$.

(v) μ_A hat nur einfache Nullstellen.

(vi) Es gibt in $K[A]$ linear unabhängige idempotente Matrizen $P_1, \dots, P_r$ mit

  (a) $P_i P_j = 0$ für $i \neq j$.

  (b) $P_1 + \cdots + P_r = E$.

  (c) $A = \lambda_1 P_1 + \cdots + \lambda_r P_r$.

Bemerkung.

(i) $P_1, \dots, P_r$ heißt eine *Spektralzerlegung* von A.

(ii) Ist A diagonalisierbar, so enthält die Algebra $K[A]$ keine von 0 verschiedenen nilpotenten Matrizen. Dann nennt man $K[A]$ auch *halbeinfach*. Man bezeichnet mitunter diagonalisierbare Matrizen auch als *halbeinfache* Matrizen.

Beweis. (i)⇒(ii): Es sei D eine Diagonalmatrix und

$$W^{-1}AW = D.$$

Es sei nun $B \in K[A]$, d. h.

$$B = P(A) \quad \text{für ein } P \in K[x].$$

Damit ist

$$W^{-1}BW = W^{-1}P(A)W = P(W^{-1}AW) = P(D)$$

ebenfalls eine Diagonalmatrix.
(ii)⇒(iii): Es sei $B \in K[A]$. Nach (ii) ist B diagonalisierbar, d. h.

$$B \sim \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} =: C.$$

Nun ist mit B auch C nilpotent, d. h. $\mu_1 = \cdots = \mu_n = 0$. D. h.

$$W^{-1}BW = 0 \Rightarrow B = W0W^{-1} = 0.$$

(iii)⇒(iv): Es sei

$$w_A := (x - \lambda_1) \cdots (x - \lambda_r).$$

Nach Satz (11.32) ist $w_A(A)$ nilpotent. Also ist

$$w_A(A) = 0.$$

Da μ_A nach Satz (11.30) dieselben Nullstellen wie χ_A hat, folgt

$$\mu_A = w_A.$$

(iv)⇒(v): Klar.
(v)⇒(i): Da nach Satz (11.30) μ_A die Nullstellen $\lambda_1, \ldots \lambda_r$ haben muß, folgt

$$\mu_A = (x - \lambda_1) \cdots (x - \lambda_r).$$

Für A kann man nach Satz (11.20) annehmen, daß

$$A \sim B = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix},$$

wobei

$$A_i = \lambda_i E + N_i, \quad N_i \text{ echte obere Dreiecksmatrix.}$$

Unser Ziel ist es zu zeigen, daß $N_i = 0$.

Es gilt nun $\mu_A = \mu_B$. Damit gilt

$$\mu_A(B) = \begin{pmatrix} \mu_A(A_1) & & \\ & \ddots & \\ & & \mu_A(A_r) \end{pmatrix} = 0.$$

Also gilt

$$\begin{aligned} 0 = \mu_A(A_i) &= (A_i - \lambda_1 E) \cdots \underbrace{(A_i - \lambda_i E)}_{N_i} \cdots (A_i - \lambda_r E) \\ &= \prod_{\substack{j=1 \\ j \neq i}}^{r} (A_i - \lambda_j E) \cdot N_i. \end{aligned}$$

Hierbei haben wir verwendet, daß $K[A_i]$ kommutativ ist. Nun ist

$$J_i := \prod_{\substack{j=1 \\ j \neq i}}^{r} (A_i - \lambda_j E) = \varphi_i(A_i)$$

mit

$$\varphi_i(x) := \prod_{\substack{j=1 \\ j \neq i}}^{r} (x - \lambda_j).$$

Also ist J_i nach Satz (11.31) verallgemeinerte Jordanmatrix zum Eigenwert

$$\varphi_i(\lambda_i) = \prod_{\substack{j=1 \\ j \neq i}}^{r} (\lambda_i - \lambda_j) =: c_i \neq 0.$$

D. h.

$$0 = J_i \cdot N_i = \begin{pmatrix} c_i & & * \\ \vdots & \ddots & \\ 0 & \cdots & c_i \end{pmatrix} \begin{pmatrix} 0 & & * \\ \vdots & \ddots & \\ 0 & \cdots & 0 \end{pmatrix} \qquad (c_i \neq 0).$$

Behauptung. $N_i = 0$ ($\Rightarrow A_i$ ist Diagonalmatrix).

Denn: Es sei $N_i \neq 0$. Man wähle das minimale $s \geq 1$ mit $\alpha^{(i)}_{k,k+s} \neq 0$ für ein k (Es sei $A_i = (\alpha^{(i)}_{kl})$). Nun ist

$$0 = J_i N_i = (c_i E + N') \cdot N_i = c_i N_i + N' N_i,$$

wobei N' eine echte obere Dreiecksmatrix ist. Es genügt zu zeigen, daß

$$(N' N_i)_{k,k+s} = 0,$$

denn dann erhalten wir einen Widerspruch zu $\alpha^{(i)}_{k,k+s} \neq 0$. Nun ist

$$\begin{aligned} (N'N_i)_{k,k+s} &= \sum_{l=1}^{n} N'_{kl}(N_i)_{l,k+s} \\ &= \sum_{l=1}^{k} \underbrace{N'_{kl}}_{=0}(N_i)_{l,k+s} + \sum_{l=k+1}^{n} N'_{kl} \underbrace{(N_i)_{l,k+s}}_{=0,\text{da } k+s-l<s} = 0. \end{aligned}$$

(iv)$\Rightarrow$(vi): Wir wissen, daß

$$\mu_A(x) = (x - \lambda_1) \cdots (x - \lambda_r).$$

Es sei

$$\psi_i(x) := \frac{\varphi_i(x)}{\varphi_i(\lambda_i)} \in K[x].$$

Dann ist

$$\psi_i(\lambda_j) = \delta_{ij}.$$

Es sei wie früher

$$I_A = \{P \in K[x];\ P(A) = 0\} = \langle \mu_A \rangle.$$

Behauptung 1. $\psi_i \psi_j - \delta_{ij} \psi_i \in I_A$.

Denn: Es genügt zu zeigen, daß für $\varphi_{ij} := \psi_i \psi_j - \delta_{ij} \psi_i$ gilt

$$\varphi_{ij}(\lambda_k) = 0 \qquad k = 1, \dots, r,$$

denn dann ist μ_A ein Teiler von φ_{ij}.

Nun ist

$$\begin{aligned} \varphi_{ij}(\lambda_k) &= \psi_i(\lambda_k)\psi_j(\lambda_k) - \delta_{ij}\psi_i(\lambda_k) \\ &= \delta_{ik}\delta_{jk} - \delta_{ij}\delta_{ik} = 0. \end{aligned}$$

Behauptung 2. $\psi_1 + \cdots + \psi_r - 1 \in I_A$.

Denn: Dies folgt wie oben wegen

$$\psi_1(\lambda_k) + \cdots + \psi_r(\lambda_k) - 1 = 0 \quad (k = 1, \ldots, r).$$

Behauptung 3. $\lambda_1\psi_1 + \cdots + \lambda_r\psi_r - x \in I_A$.

Denn: Es gilt

$$\lambda_1\psi_1(\lambda_k) + \cdots + \lambda_r\psi_r(\lambda_k) - \lambda_k = \lambda_k - \lambda_k = 0 \quad (k = 1, \ldots, r).$$

Man setze nun:

$$P_i := \psi_i(A). \quad (i = 1, \ldots, r)$$

Dann folgen die Behauptungen $P_i^2 = P_i$ sowie (a), (b) und (c) sofort aus den Behauptungen (1), (2) und (3). Da $\deg \psi_i < \deg \mu_A$ ist, ist $P_i \neq 0$ für $i = 1, \ldots, r$. Es bleibt die *lineare Unabhängigkeit* der P_i zu zeigen. Es sei

$$\alpha_1 P_1 + \cdots + \alpha_r P_r = 0.$$

Multiplizieren mit P_i gibt wegen (a):

$$0 = \alpha_i P_i^2 = \alpha_i P_i \overset{P_i \neq 0}{\Rightarrow} \alpha_i = 0 \quad (i = 1, \ldots, k).$$

(vi)⇒(iv): Es gilt

$$(\alpha_1 P_1 + \cdots + \alpha_r P_r)(\beta_1 P_1 + \cdots + \beta_r P_r) \overset{P_iP_j = \delta_{ij}P_i}{=} (\alpha_1\beta_1 P_1 + \cdots + \alpha_r\beta_r P_r).$$

Also folgt

$$P(A) \overset{(c)}{=} P(\lambda_1 P_1 + \cdots + \lambda_r P_r) = P(\lambda_1)P_1 + \cdots + P(\lambda_r)P_r.$$

Also

$$P(A) = 0 \Leftrightarrow P(\lambda_1) = \cdots = P(\lambda_r) = 0.$$

Insbesondere folgt daraus

$$\mu_A = (x - \lambda_1) \cdots (x - \lambda_r).$$

□

Bemerkung. Der Beweis zeigt, wie man die Spektralzerlegung berechnen kann.

$$A := \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$$

Es gilt

$$\chi_A = (x-1)(x-4) + 2 = x^2 - 5x + 6 = (x-2)(x-3) = \mu_A.$$

Dann ist

$$\psi_1 = \frac{x-3}{-1} = 3 - x; \quad \psi_2 = \frac{x-2}{1} = x - 2.$$

Also ist

$$P_1 = \psi_1(A) = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix}, \quad P_2 = \psi_2(A) = \begin{pmatrix} -1 & -1 \\ 2 & 2 \end{pmatrix}.$$

Es ist insbesondere

$$2P_1 + 3P_2 = \begin{pmatrix} 4 & 2 \\ -4 & -2 \end{pmatrix} + \begin{pmatrix} -3 & -3 \\ 6 & 6 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} = A.$$

Es sei hier nochmals an den Begriff des *Eigenraums* zum Eigenwert λ erinnert, der wie folgt definiert war

$$\mathrm{Eig}(A, \lambda) = \{v \in K^n; Av = \lambda v\} = \mathrm{Ker}(A - \lambda E).$$

Der Zusammenhang zwischen Eigenräumen und Spektralzerlegung ist dann durch den folgenden Satz gegeben. Es sei A diagonalisierbar. Dann haben wir eine Spektralzerlegung

(1) $A = \lambda_1 P_1 + \ldots + \lambda_r P_r$.

konstruiert. Wir erinnern an folgende Eigenschaften:

(2) $P_i P_j = 0$ für $i \neq j$, $P_i^2 = P_i$,

(3) $E = P_1 + \ldots + P_r$.

Satz 11.36
Die MAtrix A zerfalle. Für die Eigenräume von A gilt dann:

$$\mathrm{Eig}(A, \lambda_i) = \mathrm{Im}\, P_i.$$

Beweis. $\mathrm{Im}\, P_i \subset \mathrm{Eig}(A, \lambda_i)$: Es sei $x \in \mathrm{Im}\, P_i$, d. h. es gibt ein y mit $x = P_i y$. Dann folgt

$$\begin{aligned} Ax &= AP_i(y) \overset{(1)}{=} (\lambda_1 P_1 + \ldots + \lambda_i P_i + \ldots + \lambda_r P_r) P_i(y) \\ &\overset{(2)}{=} \lambda_i P_i^2(y) \overset{(2)}{=} \lambda_i P_i(y) = \lambda_i x. \end{aligned}$$

Also ist $x \in \mathrm{Eig}(A, \lambda_i)$.

$\mathrm{Eig}(A, \lambda_i) \subset \mathrm{Im}\, P_i$: Es sei $x \in \mathrm{Eig}(A, \lambda_i)$. Also gilt $Ax = \lambda_i x$. Daraus folgt mit (1)

$$(\lambda_1 P_1 + \ldots + \lambda_r P_r)x = \lambda_i x,$$

und damit

$$\lambda_j P_j(x) \overset{(2)}{=} P_j(\lambda_1 P_1 + \ldots + \lambda_r P_r)(x) = P_j(\lambda_i x) = \lambda_i P_j(x).$$

Also ist

$$P_j(x)(\lambda_j - \lambda_i) = 0.$$

Für $j \neq i$ folgt $P_j(x) = 0$. Also gilt

$$x = Ex \overset{(3)}{=} (P_1 + \ldots + P_r)(x) = P_i(x) \;\Rightarrow\; x \in \mathrm{Im}\, P_i.$$

□

Satz 11.37
Sind $\lambda_1, \ldots, \lambda_r$ die Eigenwerte von A, so sind äquivalent:

(i) A ist diagonalisierbar.

(ii) $K^n = \mathrm{Eig}(A, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(A, \lambda_r)$.

Beweis. (ii)⇒(i) Es gibt eine Basis aus Eigenvektoren, also ist A diagonalisierbar.

(i)⇒(ii) Daß die Summe direkt ist, folgt aus Lemma (I.8.6). Wegen Satz (11.36) ist $\mathrm{Eig}(A, \lambda_i) = \mathrm{Im}\, P_i$, und es genügt zu zeigen, daß

$$K^n = \mathrm{Im}\, P_1 + \ldots + \mathrm{Im}\, P_n.$$

Es sei $x \in K^n$. Dann gilt

$$x = Ex = (P_1 + \ldots + P_r)(x) = P_1(x) + \ldots + P_r(x) \in \mathrm{Im}(P_1) + \ldots + \mathrm{Im}(P_r).$$

□

Definition. Die *geometrische Vielfachheit* eines Eigenwertes λ_i ist die Dimension des Eigenraumes $\mathrm{Eig}(A, \lambda_i)$.

Satz 11.38
Die geometrische Vielfachheit eines Eigenwertes λ_i ist höchstens gleich der algebraischen Vielfachheit von λ_i.

Beweis. Es sei r die geometrische und s die algebraische Vielfachheit von λ_i. Ergänzt man r linear unabhängige Eigenvektoren zum Eigenwert λ_i zu einer Basis von K^n, so erhält man

$$A \sim \left(\begin{array}{ccc|c} \lambda_i & & 0 & \\ & \ddots & & * \\ & & \lambda_i & \\ \hline & 0 & & B \end{array}\right) \left.\vphantom{\begin{array}{c} \lambda_i \\ \ddots \\ \lambda_i \end{array}}\right\}\mathrm{r}$$

$$\underbrace{\hphantom{\lambda_i \quad \ddots \quad \lambda_i}}_{r}$$

D. h.

$$\chi_A(x) = (x - \lambda_i)^r \chi_B(x)$$

und es folgt, daß $r \leq s$ ist. □

Beispiel. Für

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

gilt

$$\mathrm{Eig}(A, 2) = Ke_1, \qquad \chi_A(x) = (x-2)^2.$$

Die geometrische Vielfachheit eines Eigenwertes kann also echt kleiner als die algebraische Vielfachheit sein.

Die Jordan-Chevalley-Zerlegung

Theorem 11.39
Es sei $A \in \mathrm{Mat}(n; K)$ eine zerfallende Matrix. Dann gibt es eine diagonalisierbare Matrix H, sowie eine nilpotente Matrix N mit $H, N \in K[A]$, so daß

$$A = H + N.$$

Insbesondere gilt $[H, N] = 0$.

Beweis. Wir machen zunächst den

Reduktionsschritt: Es genügt anzunehmen, daß A in Blockform gegeben ist:

$$(*) \qquad A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}.$$

Hierbei sind $A_i \in \mathrm{Mat}(m_i; K)$ verallgemeinerte Jordanblöcke zum Eigenwert λ_i.

Denn: Es sei $A = W^{-1}A'W$ mit A' in Blockform. Man habe eine Zerlegung

$$A' = H' + N'.$$

Dann gilt

$$A = W^{-1}H'W + W^{-1}N'W.$$

Wir setzen

$$H := W^{-1}H'W, \qquad N := W^{-1}N'W.$$

Dann ist H diagonalisierbar, und N ist nilpotent. Ferner gilt

$$H, N \in W^{-1}K[A']W = K[W^{-1}A'W] = K[A].$$

Existenzbeweis: Es sei also A wie in $(*)$. Wir setzen

$$C_i := \begin{pmatrix} \boxed{0} & & & & \\ & \ddots & & & \\ & & \boxed{E_{m_i}} & & \\ & & & \ddots & \\ & & & & \boxed{0} \end{pmatrix}$$

wobei E_{m_i} die Einheitsmatrix der Größe m_i ist. Dann ist

$$H := \lambda_1 C_1 + \ldots + \lambda_r C_r$$

eine Diagonalmatrix. Wir setzen

$$N := A - H.$$

Dies ist eine echte obere Dreiecksmatrix, also nilpotent.

Es bleibt zu zeigen, daß $H \in K[A]$ ist. Dann ist auch $N \in K[A]$, und damit gilt auch $[H, N] = 0$. Um einzusehen, daß $H \in K[A]$ ist, genügt es zu zeigen, daß $C_1, \ldots, C_r \in K[A]$ gilt.

1. Schritt: Es gibt eine verallgemeinerte Jordanmatrix

$$J_i \in \mathrm{Mat}(m_i; K)$$

zum Eigenwert 1 mit

$$B_i := \begin{pmatrix} \boxed{0} & & & & \\ & \ddots & & & \\ & & \boxed{J_i} & & \\ & & & \ddots & \\ & & & & \boxed{0} \end{pmatrix} \in K[A].$$

Dazu betrachten wir die paarweise verschiedenen Eigenwerte $\lambda_1, \dots, \lambda_r$ und

$$\psi_i(x) := \prod_{j \neq i} \frac{x - \lambda_j}{\lambda_i - \lambda_j} \in K[x]. \quad (\psi_i(\lambda_j) = \delta_{ij}).$$

Es gilt

$$\psi_i(A) = \begin{pmatrix} \psi_i(A_1) & & \\ & \ddots & \\ & & \psi_i(A_r) \end{pmatrix} \in K[A].$$

Nun ist $\psi_i(A_j)$ Jordanmatrix zum Eigenwert $\psi_i(\lambda_j) = \delta_{ij}$. Für $i \neq j$ ist also $\psi_i(A_j)$ echte obere Dreiecksmatrix, d. h. nilpotent. Also hat

$$B_i := \psi_i(A)^n = \begin{pmatrix} \boxed{0} & & & & \\ & \ddots & & & \\ & & \boxed{\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}} & & \\ & & & \ddots & \\ & & & & \boxed{0} \end{pmatrix} \in K[A]$$

die gewünschte Form.

2. Schritt: Es gibt nach Korollar (11.28), angewandt auf die Matrix J_i ein Polynom $Q_i \in K[x]$ mit $Q_i(0) = 0$ und $Q_i(J_i) = E_{m_i}$. Damit gilt

$$Q_i(B_i) = \begin{pmatrix} \boxed{0} & & & & \\ & \ddots & & & \\ & & \boxed{Q_i(J_i)} & & \\ & & & \ddots & \\ & & & & \boxed{0} \end{pmatrix} = \begin{pmatrix} \boxed{0} & & & & \\ & \ddots & & & \\ & & \boxed{E_{m_i}} & & \\ & & & \ddots & \\ & & & & \boxed{0} \end{pmatrix} = C_i.$$

Wegen $B_i \in K[A]$ folgt auch

$$C_i \in K[A].$$

□

Satz 11.40
Die Matrizen $A, B \in \mathrm{Mat}(n; K)$ seien diagonalisierbar. Ist $[A, B] = 0$, so gilt:

(i) A, B sind zugleich diagonalisierbar, d. h. es gibt eine Matrix $W \in \mathrm{GL}(n; K)$, so daß $W^{-1}AW$ und $W^{-1}BW$ Diagonalmatrizen sind.

(ii) $A + B$, AB sind diagonalisierbar.

Beweis. (ii) folgt aus (i), da

$$\begin{aligned} W^{-1}(A+B)W &= W^{-1}AW + W^{-1}BW \\ W^{-1}(AB)W &= (W^{-1}AW)(W^{-1}BW). \end{aligned}$$

(i) Nach Übergang von A zu $U^{-1}AU$; $U \in \mathrm{GL}(n;K)$ kann man annehmen, daß

$$A = \begin{pmatrix} \boxed{\lambda_1 E_{m_1}} & & \\ & \ddots & \\ & & \boxed{\lambda_r E_{m_r}} \end{pmatrix}$$

wobei $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A sind, und E_{m_i} die Einheitsmatrix der Größe m_i ist. Wir schreiben B in Blockform

$$B = \begin{pmatrix} \boxed{B_{11}} & \boxed{B_{12}} & & \\ & \boxed{B_{22}} & & \\ & & \ddots & \\ & & & \boxed{B_{rr}} \end{pmatrix}, \quad B_{ij} \in \mathrm{Mat}(m_i, m_j; K).$$

Nun folgt:

$$[A, B] = 0 \ \Rightarrow \ \lambda_i B_{ij} = \lambda_j B_{ij}.$$

Für $i \neq j$ folgt also $B_{ij} = 0$, d. h.

$$B = \begin{pmatrix} \boxed{B_{11}} & & 0 \\ & \ddots & \\ 0 & & \boxed{B_{rr}} \end{pmatrix}.$$

Behauptung. Die Matrizen B_{ii} sind diagonalisierbar.

Denn: Es sei μ_B das Minimalpolynom von B; $\mu_{B_{ii}}$ das Minimalpolynom von B_{ii}. Da

$$0 = \mu_B(B) = \begin{pmatrix} \mu_B(B_{11}) & & \\ & \ddots & \\ & & \mu_B(B_{rr}) \end{pmatrix}$$

folgt $\mu_B(B_{ii}) = 0$. Also gibt es $R_i \in K[x]$ mit

$$\mu_B = R_i \mu_{B_{ii}}.$$

Da B diagonalisierbar ist, hat μ_B nur einfache Nullstellen. Dasselbe gilt dann für $\mu_{B_{ii}}$ und somit ist auch B_{ii} diagonalisierbar (Satz (8.40)). Es gibt also $W_i \in \mathrm{GL}(m_i; K)$ mit

$$W_i B_{ii} W_i^{-1} = D_{ii},$$

wobei D_{ii} eine Diagonalmatrix ist. Es sei

$$W := \begin{pmatrix} W_{11} & & \\ & \ddots & \\ & & W_{rr} \end{pmatrix}.$$

Damit gilt

$$\begin{aligned}
W^{-1}AW &= \begin{pmatrix} W_{11}^{-1} & & \\ & \ddots & \\ & & W_{rr}^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 E_{m_1} & & \\ & \ddots & \\ & & \lambda_r E_{m_r} \end{pmatrix} \begin{pmatrix} W_{11} & & \\ & \ddots & \\ & & W_{rr} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 E_{m_1} & & \\ & \ddots & \\ & & \lambda_r E_{m_r} \end{pmatrix} = A
\end{aligned}$$

sowie

$$W^{-1}BW = \begin{pmatrix} W_{11}^{-1}B_{11}W_{11} & & \\ & \ddots & \\ & & W_{rr}^{-1}B_{rr}W_{rr} \end{pmatrix} = \begin{pmatrix} D_{11} & & \\ & \ddots & \\ & & D_{rr} \end{pmatrix},$$

also ist $W^{-1}BW$ eine Diagonalmatrix. □

Es sei nun $A \in \mathrm{Mat}(n;K)$ zerfallend und

$$A = H + N$$

eine Jordan-Chevalley-Zerlegung, d. h. H ist diagonalisierbar, N ist nilpotent und es gilt $[H,N] = 0$.

Satz 11.41
H und N sind eindeutig bestimmt. Es gilt stets $H, N \in K[A]$.

Beweis. Wir wählen eine Darstellung nach Theorem (11.39):

(2) $$A = H' + N',$$

H' ist diagonalisierbar, N' ist nilpotent, $[H', N'] = 0$ und $H', N' \in K[A]$.

1. Schritt: $[H,H'] = [N,N'] = 0$:

$$[H,N] = 0 \;\Rightarrow\; [H, \underbrace{H+N}_{=A}] = 0 \;\Rightarrow\; [H,A] = 0.$$

Da $H' \in K[A]$ folgt $[H,H'] = 0$. Analog zeigt man $[N,N'] = 0$.

2. Schritt: $H - H'$ ist diagonalisierbar, $N - N'$ ist nilpotent. Dies folgt für $H - H'$ aus Satz (11.40) (ii) für $N - N'$ aus Satz (11.15) (iii).

3. Schritt: Es gilt

$$H + N = A = H' + N',$$

also

$$H - H' = N' - N.$$

Damit ist $H - H'$ diagonalisierbar und nilpotent. Also ist $H - H' = 0$, d. h. $H = H'$ und damit auch $N = N'$. □

Anwendungen

(1) *Multiplikative Jordan-Chevalley Zerlegung*

Satz 11.42
Sei $A \in \mathrm{GL}(n;K)$ zerfallend. Dann gibt es genau eine diagonalisierbare Matrix H, sowie eine nilpotente Matrix N, so daß $A = H(E+N)$ und $[H,N] = 0$.

Beweis. Wir betrachten die additive Jordan-Chevalley Zerlegung

$$A = H + N', \qquad H, N' \in K[A].$$

Nach Satz (11.16) ist

$$\det H = \det(H + N') = \det A \neq 0.$$

Also ist $H \in \mathrm{GL}(n;K)$. Sei

$$N := H^{-1}N'.$$

Dann ist

$$A = H + HN = H(E+N).$$

Da $[H,N] = [H,N'] = 0$ ist, folgt die Nilpotenz von N, d. h. wir haben eine multiplikative Zerlegung gefunden. Umgekehrt ist

$$A = H + HN$$

eine additive Jordan-Chevalley Zerlegung. Die Eindeutigkeit von H, N' folgt aus der Eindeutigkeit von H und N. □

(2) *Symmetrische Matrizen*

Satz 11.43
Es sei $A \in \mathrm{Mat}(n;K)$ symmetrisch und zerfallend. Dann gilt für die Jordan-Chevalley Zerlegung

$$A = H + N,$$

daß auch H, N symmetrisch sind.

Beweis. Es ist

$$H + N = A = {}^tA = {}^tH + {}^tN.$$

Es ist auch ${}^tH$ diagonalisierbar, ${}^tN$ ist nilpotent und $[{}^tH, {}^tN] = 0$. Damit ist auch $A = {}^tH + {}^tN$ eine Jordan-Chevalley Zerlegung von A. Aus der Eindeutigkeit folgt dann $H = {}^tH$, $N = {}^tN$.

(3) *Reelle Jordan-Chevalley Zerlegung*

Es sei

$$Z = (z_{kl}) \in \mathrm{Mat}(n; \mathbb{C}).$$

Wir schreiben

$$z_{kl} = x_{kl} + iy_{kl} \; ; \qquad x_{kl}, y_{kl} \in \mathbb{R}$$

für alle k, l. Also hat man eine eindeutige Zerlegung

$$Z = X + iY \; ; \qquad X, Y \in \mathrm{Mat}(n; \mathbb{R}).$$

Definition. Die *konjugierte* Matrix ist definiert durch

$$\overline{Z} := X - iY.$$

Bemerkung.

(i) $\overline{Z_1 + Z_2} = \overline{Z_1} + \overline{Z_2}$,

(ii) $\overline{Z_1 \cdot Z_2} = \overline{Z_1} \cdot \overline{Z_2}$.

Definition. $A \in \mathrm{Mat}(n; \mathbb{R})$ heißt *halbeinfach*, falls A, aufgefaßt als Element in $\mathrm{Mat}(n; \mathbb{C})$, diagonalisierbar ist.

Satz 11.44

Es sei $A \in \mathrm{Mat}(n; \mathbb{R})$. Dann gibt es eindeutig bestimmte reelle Matrizen H, N mit H halbeinfach und N nilpotent, so daß

$$A = H + N$$

und $[H, N] = 0$ gilt.

Beweis. $A \in \mathrm{Mat}(n; \mathbb{C})$ ist zerfallend. Also gibt es H diagonalisierbar, N nilpotent so daß

$$A = H + N$$

und $[H, N] = 0$. Es gilt

$$A = \overline{A} = \overline{H} + \overline{N}.$$

Wegen der Eindeutigkeit folgt $H = \overline{H}$, $N = \overline{N}$, d. h. H und N sind reell. Die Eindeutigkeit folgt aus dem komplexen Fall. $\square$

Der Quotientenraum

Es sei X eine Menge.

Definition. Eine *Äquivalenzrelation* auf X ist eine Teilmenge $R \subset X \times X$ mit folgenden Eigenschaften:

(i) $(x,x) \in R$ für alle $x \in X$ (*reflexiv*)

(ii) $(x,y) \in R \Rightarrow (y,x) \in R$ für alle $x,y \in X$ (*symmetrisch*)

(iii) $(x,y) \in R, (y,z) \in R \Rightarrow (x,z) \in R$ für alle $x,y,z \in X$ (*transitiv*).

Schreibweise. $x \sim y :\Leftrightarrow (x,y) \in R$

Beispiele.

(1) $R = \{(x,x);\ x \in X\} \subset X \times X$.

Dann gilt:

$$x \sim y \ \Leftrightarrow \ x = y.$$

(2) Es sei $X = \mathrm{Mat}(n;K)$ und

$$R := \{(A,B) \in X \times X;\ \text{es gibt } W \in \mathrm{GL}(n;K) \text{ mit } A = W^{-1}BW\} \subset X \times X,$$

d. h.

$$A \sim B \Leftrightarrow A \text{ ist ähnlich zu } B.$$

(3) Es sei V ein K-Vektorraum und $U \subset V$ ein Unterraum.

$$R := \{(u,v) \in V \times V;\ u - v \in U\}$$

D. h.

$$u \sim v \Leftrightarrow u - v \in U.$$

Lemma 11.45
Die obige Relation R ist eine Äquivalenzrelation.

Beweis.

(i) $(u,u) \in R$, da $u - u = 0 \in U$.

(ii) $(u,v) \in R \Rightarrow u - v \in U \Rightarrow v - u \in U \Rightarrow (v,u) \in R$.

(iii) $(u,v) \in R, (v,w) \in R \Rightarrow (u-v) \in U, (v-w) \in U \Rightarrow u - w \in U \Rightarrow (u,w) \in R$. □

R sei im folgenden stets eine Äquivalenzrelation.

Definition. Eine Teilmenge $A \subset X$ heißt eine *Äquivalenzklasse*, falls

(i) $A \neq \emptyset$,

(ii) $x, y \in A \Rightarrow x \sim y$,

(iii) $x \in A, y \in X$ mit $x \sim y \Rightarrow y \in A$.

Lemma 11.46

(i) Sind A, A' Äquivalenzklassen, so gilt $A \cap A' = \emptyset$ oder $A = A'$.

(ii) Zu jedem $a \in X$ gibt es genau eine Äquivalenzklasse A_a mit $a \in A_a$.

Definition. a heißt ein *Repräsentant* der Äquivalenzklasse $A = A_a$.

Beweis.

(i) Sei $A \cap A' \neq \emptyset$, etwa $a \in A \cap A'$.

$A \subset A'$: $\quad x \in A, a \in A \cap A' \overset{\text{(ii)}}{\Rightarrow} x \sim a \underset{a \in A'}{\overset{\text{(iii)}}{\Rightarrow}} x \in A'$.

$A' \subset A$: $\quad$ Analog.

(ii) Setze

$$A_a := \{x;\ x \sim a\} \subset X.$$

A_a ist eine Äquivalenzklasse, denn:

(i) $a \in A_a$, also $A_a \neq \emptyset$.

(ii) $x, y \in A_a \Rightarrow x \sim a, y \sim a \Rightarrow x \sim y$.

(iii) $x \in A_a, y \sim x \Rightarrow x \sim a, y \sim x \Rightarrow y \sim a \Rightarrow y \in A_a$.

Die Eindeutigkeit folgt aus Teil (i). $\square$

Folgerung. X ist die disjunkte Vereinigung der Äquivalenzklassen.

Wir führen nun die folgende neue Menge ein:

$$X/R := \text{Menge der Äquvalenzklassen.}$$

Man hat eine *kanonische Projektion*

$$\pi: \quad X \longrightarrow X/R, \quad a \longmapsto A_a.$$

Wir kehren nun zu Beispiel (3) zurück.

Definition. $a + U := \{a + u;\ u \in U\}$.

Lemma 11.47

(i) $A_a = a + U$.

(ii) $a + U = b + U \Leftrightarrow a \sim b$, d. h. $a - b \in U$.

Beweis.

(i) $x \sim a \Leftrightarrow x - a \in U \Leftrightarrow x \in a + U$.

(ii) Aus (i).

Definition. Der *Quotientenraum* V/U ist die Menge

$$V/U := \{a + U;\ a \in V\}.$$

Bemerkung. Die Elemente der Menge V/U sind also Äquivalenzklassen bezüglich der Relation R und damit Teilmengen von V.

Geometrische Deutung

Es sei $U \subset V$ ein Unterraum und W ein Komplement von U in V, d. h.

$$V = U \oplus W.$$

Es sei $U_a = a + U$ eine Äquivalenzklasse. Dann gibt es eine eindeutig bestimmte Darstellung

$$a = a_W + a_U, \quad a_W \in W, a_U \in U.$$

Es gilt

$$U_a = a + U = a_W + U.$$

Das Element a_W hängt nur von der Äquivalenzklasse ab und ist durch diese eindeutig bestimmt:

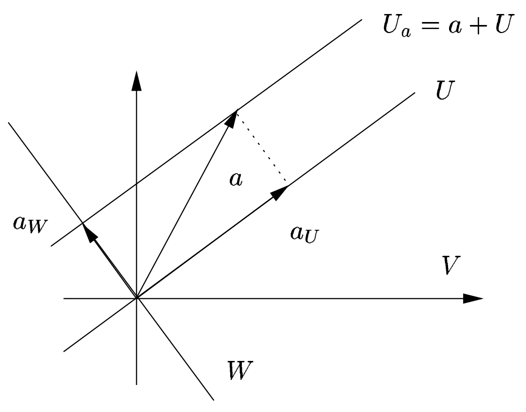


Abb. 2: Geometrische Deutung des Quotientenraumes V/U.

Wir können also die Elemente von V/U mit den Elementen von W identifizieren. Allerdings ist es günstiger mit V/U zu arbeiten, da das Komplement W nicht eindeutig bestimmt ist.

Lemma und Definition 11.48
Die Verknüpfungen

$$\begin{array}{rlr} + : & V/U \times V/U \longrightarrow V/U; & (a+U)+(b+U) := (a+b)+U \\ \cdot : & K \times V/U \longrightarrow V/U; & \alpha(a+U) := \alpha a + U \end{array}$$

sind wohldefiniert (d. h. unabhängig von der Wahl der Repräsentanten a und b) und machen V/U zu einem K-Vektorraum.

Beweis. Wir zeigen zunächst die *Unabhängigkeit* von der Wahl der Repräsentanten.

(i) Es sei $a \sim a', b \sim b'$. Wir müssen zeigen, daß

$$(a+b)+U = (a'+b')+U$$

ist, d. h. daß $a+b \sim a'+b'$. Nach Voraussetzung ist $a-a' \in U$, $b-b' \in U$, also gilt $(a-a')+(b-b') \in U$, d. h. $(a+b)-(a'+b') \in U$, und damit $a+b \sim a'+b'$.

(ii) Die entsprechende Aussage für die Skalarmultiplikation zeigt man analog.

Das Nachrechnen der *Vektorraumaxiome* für V/U ist einfach. Wir führen dies an einem Beispiel vor:

$$\begin{aligned} \alpha(a+U)+\alpha(b+U) &= (\alpha a+U)+(\alpha b+U) \\ &= (\alpha a+\alpha b)+U = \alpha(a+b)+U \\ &= \alpha((a+b)+U) \\ &= \alpha((a+U)+(b+U)). \end{aligned}$$

Das *neutrale* Element von V/U ist $0+U$, das *additive Inverse* von $a+U$ ist $-a+U$. □

Satz 11.49
Die Abbildung

$$\pi : \ V \longrightarrow V/U; \quad a \longmapsto a+U$$

ist ein Epimorphismus mit $\operatorname{Ker}\pi = U$.

Beweis. π ist surjektiv nach Konstruktion. Es gilt

$$\pi(a+b) = (a+b)+U = (a+U)+(b+U) = \pi(a)+\pi(b).$$

Analog zeigt man $\pi(\alpha a) = \alpha\pi(a)$. Ferner ist

$$\pi(a) = 0 \Leftrightarrow a+U = 0+U \Leftrightarrow a \sim 0 \Leftrightarrow a-0 = a \in U.$$

d. h. $\operatorname{Ker}\pi = U$. □

Definition. π heißt der *kanonische Epimorphismus*.

Korollar 11.50
$\dim U + \dim V/U = \dim V$.

Beweis.

$$\begin{aligned} \dim V &= \dim \operatorname{Ker} \pi + \dim \operatorname{Im} \pi \\ &= \dim U + \dim V/U. \end{aligned}$$

□

Kanonische Faktorisierung

Wir betrachten einen Homomorphismus $f : V \to W$.

Satz 11.51
Es gibt genau einen Isomorphismus

$$\overline{f} : V/\operatorname{Ker} f \longrightarrow \operatorname{Im} f,$$

so daß das folgende Diagramm kommutativ ist:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\ V/\operatorname{Ker} f \arrow[r, "\overline{f}"] & \operatorname{Im} f. \arrow[u, "i"'] \end{tikzcd}$$

wobei $i : \operatorname{Im} f \to W$ die Inklusion ist.

Beweis. Falls $\overline{f}$ existiert, muß gelten:

$$\overline{f}(a + \operatorname{Ker} f) = f(a).$$

Dies ist *wohldefiniert*, da:

$$a \sim b \Rightarrow a - b \in \operatorname{Ker} f \Rightarrow f(a-b) = 0 \Rightarrow f(a) = f(b).$$

Wir definieren also $\overline{f}$ wie oben. Dann erhält man einen Homomorphismus, da

$$\begin{aligned} \overline{f}((a + \operatorname{Ker} f) + (b + \operatorname{Ker} f)) &= \overline{f}((a+b) + \operatorname{Ker} f) = f(a+b) \\ &= f(a) + f(b) = \overline{f}(a + \operatorname{Ker} f) + \overline{f}(b + \operatorname{Ker} f). \end{aligned}$$

Analog verfährt man mit der Skalarmultiplikation.

$\overline{f}$ ist ein *Isomorphismus*: Die Surjektivität folgt aus der Konstruktion. Die Injektivität folgt aus

$$\overline{f}(a + \operatorname{Ker} f) = 0 \Leftrightarrow f(a) = 0 \Leftrightarrow a \in \operatorname{Ker} f \Leftrightarrow a + \operatorname{Ker} f = 0 + \operatorname{Ker} f.$$

Kommutativität des Diagramms:

$$(i \circ \overline{f} \circ \pi)(a) = i(\overline{f}(a + \operatorname{Ker} f)) = i(f(a)) = f(a).$$

$\square$

Folgerung. $\overline{f} : V/\operatorname{Ker} f \xrightarrow{\sim} \operatorname{Im} f$.

Beispiel. $V = U \oplus W$. Wir betrachten die Projektion auf W:

$$f : \ V \longrightarrow W, \quad v = u + w \longmapsto w.$$

Dann ist $U = \operatorname{Ker} f$, $W = \operatorname{Im} f$. Man hat

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \arrow[d, equal] \\ V/U \arrow[r, "\overline{f}"] & \operatorname{Im} f = W. \end{tikzcd}$$

also

$$\overline{f} : \ V/U \xrightarrow{\sim} W.$$

Dies ist genau die früher schon von uns beschriebene Deutung des Quotientenraumes.

Die Jordansche Normalform

In diesem Abschnitt leiten wir die Jordansche Normalform her und beschreiben eine Methode, sie zu berechnen.

Im folgenden sei V ein K-Vektorraum von Dimension $n < \infty$ und $f \in \operatorname{End}_K(V)$ ein Endomorphismus. Für jeden Vektor $0 \neq a \in V$ betrachten wir den folgenden Unterraum.

Definition.

$$U(f, a) := \operatorname{Span}\{f^k(a);\ k = 0, 1, 2, \ldots\}.$$

Bemerkung.

(i) $f^0(a) = \operatorname{id}(a) = a$.

(ii) $U(f, a) \subset V$ ist ein Unterraum. Es gilt $f(U(f, a)) \subset U(f, a)$.

Definition. Ein Unterraum $U \subset V$ heißt *f-zyklisch*, falls es ein $0 \neq a \in V$ gibt mit $U = U(f, a)$.

Der Endomorphismus f sei nun nilpotent, d. h. es gibt ein l mit $f^l = 0$.

Definition. Wir definieren $m(f, a)$ als die kleinste Zahl $m \geq 0$ mit $f^m(a) = 0$.

Bemerkung. $f(a) = 0 \Leftrightarrow m(f, a) = 1$ (für $a \neq 0$).

Lemma 11.52
Es sei $0 \neq a \in V$ und $m = m(f, a)$. Dann ist $a, f(a), \dots, f^{m-1}(a)$ eine Basis von $U(f, a)$. Insbesondere ist $\dim U(f, a) = m$.

Beweis. Die Eigenschaft, ein Erzeugendensystem zu sein, ist klar, da $f^k(a) = 0$ für $k \geq m$.

Für die lineare Unabhängigkeit betrachten wir

$$\alpha_0 a + \alpha_1 f(a) + \ldots + \alpha_{m-1} f^{m-1}(a) = 0.$$

Anwendung von f^{m-1} gibt:

$$\alpha_0 \underbrace{f^{m-1}(a)}_{\neq 0} = 0 \;\Rightarrow \alpha_0 = 0.$$

Sukzessive Anwendung von $f^{m-2}, \dots, f$ ergibt

$$\alpha_1 = \ldots = \alpha_{m-1} = 0.$$

□

Satz 11.53
Der Endomorphismus $f : V \to V$ sei nilpotent. Dann ist V eine direkte Summe

$$V = U_1 \oplus \ldots \oplus U_s$$

von f-zyklischen Unterräumen $U_1, \dots, U_s$.

Beweis. *1. Fall:* $f = 0$. Dann gilt für jedes $0 \neq a \in V$:

$$U(f, a) = \operatorname{Span}\{a\}.$$

Damit ist die Aussage klar.

2. Fall: $f \neq 0$. Wir machen Induktion nach $n = \dim V$.

$n = 1: \quad V = U(f, a) = \operatorname{Span}\{a\}$ für jedes $0 \neq a \in V$.

$n - 1 \mapsto n: \quad$ Da $f \neq 0$ gibt es $m \geq 2$ mit

$$f^{m-1} \neq 0, \quad f^m = 0.$$

Man wähle a so, daß $f^{m-1}(a) \neq 0$. Dann ist

$$m(f, a) = m.$$

Es sei

$$U := U(f, a).$$

Nach Lemma (11.52) gilt $\dim U = m$. Ist $U = V$, so ist man fertig.

$U \neq V$:

Behauptung 1: Es sei $b \in V$ und $f^k(b) \in U$ für ein $k \geq 0$. Dann gibt es ein $b' \in b + U$ mit $f^k(b') = 0$.

Denn: Klar für $k = 0$. Sei $k \geq 1$ und

$$(*) \qquad c := f^k(b) \in U.$$

Also hat man wegen Lemma (11.52) eine Darstellung

$$c = \alpha_0 a + \alpha_1 f(a) + \ldots + \alpha_{m-1} f^{m-1}(a).$$

Nun ist wegen $(*)$

$$f^{m-k}(c) = f^m(b) = 0.$$

Anwenden von $f^{m-1}, \ldots, f^{m-k}$ ergibt

$$\alpha_0 = \ldots = \alpha_{k-1} = 0.$$

D. h. für

$$c = \alpha_k f^k(a) + \ldots + \alpha_{m-1} f^{m-1}(a)$$

gilt

$$c = f^k(\underbrace{\alpha_k a + \ldots + \alpha_{m-1} f^{m-1-k}(a)}_{=:c' \in U}) = f^k(c').$$

Es sei

$$b' := b - c' \in b + U.$$

Dann gilt

$$f^k(b') = f^k(b) - f^k(c') = f^k(b) - c \overset{(*)}{=} 0.$$

Behauptung 2: Es gibt einen K-Vektorraum $\tilde{V}$ sowie ein kommutatives Diagramm

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "g"'] & V \arrow[d, "g"] \\ \tilde{V} \arrow[r, "\tilde{f}"] & \tilde{V} \end{tikzcd}$$

mit $\dim \tilde{V} = n - m$, $\operatorname{Ker} g = U$, $\operatorname{Im} g = \tilde{V}$, $\tilde{f}^m = 0$.

Denn: Wir setzen

$$\tilde{V} := V/U, \quad g : V \longrightarrow V/U = \tilde{V} \quad \text{(kanonische Projektion).}$$

Ferner sei

$$\tilde{f}(a + U) := f(a) + U.$$

Dies ist wohldefiniert, denn

$$\begin{aligned} a \sim a' \quad &\Rightarrow \quad a - a' \in U \Rightarrow f(a) - f(a') \in f(U) \subset U \\ &\Rightarrow \quad f(a) + U = f(a') + U. \end{aligned}$$

Nach Konstruktion ist $\operatorname{Ker} g = U$, $\operatorname{Im} g = \tilde{V}$, $\dim \tilde{V} = n - \dim U = n - m$. Das Diagramm kommutiert, da

$$\begin{aligned} (\tilde{f} \circ g)(a) &= \tilde{f}(a + U) = f(a) + U \\ (g \circ f)(a) &= g(f(a)) = f(a) + U. \end{aligned}$$

Es gilt sogar

$$(**) \qquad \tilde{f}^l \circ g = g \circ f^l \qquad (l \in \mathbb{N}).$$

Dies zeigt man durch Induktion nach l. Für $l = 1$ haben wir die Aussage eben bewiesen. Der Induktionsschritt folgt aus:

$$\tilde{f}^l \circ g = \tilde{f} \circ (\tilde{f}^{l-1} \circ g) \overset{\text{(IV)}}{=} \tilde{f} \circ (g \circ f^{l-1}) = (g \circ f) \circ f^{l-1} = g \circ f^l.$$

Damit ist

$$\tilde{f}^m \circ g = g \circ f^m = 0.$$

Da g surjektiv ist, folgt $\tilde{f}^m = 0$.

Ende des Beweises: Im Falle $U \neq V$ betrachten wir $\tilde{V} := V/U$ sowie $\tilde{f} : \tilde{V} \to \tilde{V}$. Es ist $\tilde{f}^m = 0$. Nach Induktionsvoraussetzung gibt es eine Zerlegung

$$\tilde{V} = \tilde{U}_1 \oplus \ldots \oplus \tilde{U}_s$$

wobei $\tilde{U}_i$ $\tilde{f}$-zyklische Unterräume sind.

Behauptung 3: Es gibt $a_1, \ldots, a_s \in V$ mit

(i) $\tilde{U}_i = U(\tilde{f}, \tilde{a}_i)$ wobei $\tilde{a}_i = g(a_i)$.

(ii) Ist $U_i := U(f, a_i)$, so ist $g : U_i \to \tilde{U}_i$ ein Isomorphismus.

Denn: Wir wählen $\tilde{a}_i \in \tilde{V}$ mit

$$\tilde{U}_i = U(\tilde{f}, \tilde{a}_i), \quad (i = 1, \dots , s).$$

Es ist dann

$$m_i := \dim \tilde{U}_i = m(\tilde{f}, \tilde{a}_i) \quad \text{(nach Lemma (11.52))}.$$

Insbesondere gilt $\tilde{f}^{m_i}(\tilde{a}_i) = 0$. Wir wählen nun ein $a_i \in V$ mit

$$g(a_i) = \tilde{a}_i.$$

Die a_i sind nur bis auf Elemente in U festgelegt. Es gilt

$$0 = \tilde{f}^{m_i}(\tilde{a}_i) = (\tilde{f}^{m_i} \circ g)(a_i) \overset{(**)}{=} g(f^{m_i}(a_i)).$$

Also ist $f^{m_i}(a_i) \in U$. Nach Behauptung 1 kann man annehmen, daß $f^{m_i}(a_i) = 0$ ist. Nun ist

$$0 \neq \tilde{f}^k(\tilde{a}_i) = (\tilde{f}^k \circ g)(a_i) \overset{(**)}{=} g(f^k(a_i)) \quad \text{für } k < m_i.$$

Also folgt $f^k(a_i) \neq 0$ für $k < m$. Damit gilt

$$m(f, a_i) = m_i.$$

Also gilt

$$U_i := U(f, a_i); \quad \dim U_i = m_i.$$

Nach Konstruktion ist $g(U_i) = \tilde{U}_i$. Da $\dim U_i = \dim \tilde{U}_i$ ist

$$g : \ U_i \longrightarrow \tilde{U}_i$$

ein Isomorphismus. Der Beweis des Satzes folgt schließlich aus der

Behauptung 4: $V = U \oplus U_1 \oplus \ldots \oplus U_s$.

Denn: Wir zeigen zunächst, daß die Summe auf der rechten Seite direkt ist. Sei $u \in U$; $u_i \in U_i, i = 1, \dots , s$. Es sei

$$u + u_1 + \ldots + u_s = 0.$$

Anwendung von g ergibt

$$g(u_1) + \ldots + g(u_s) = 0.$$

Da die Summe der $\tilde{U}_i$ direkt ist, folgt hieraus

$$g(u_1) = \ldots = g(u_s) = 0.$$

Mit Behauptung 3 (ii) ergibt sich

$$u_1 = \ldots = u_s = 0$$

und damit

$$u = u_1 = \ldots = u_s = 0.$$

Die Behauptung folgt nun aus

$$\begin{aligned} \dim V &= \dim V/U + \dim U = \dim \tilde{V} + \dim U \\ &= \dim \tilde{U}_1 + \ldots + \dim \tilde{U}_s + \dim U \\ &= \dim U_1 + \ldots + \dim U_s + \dim U. \end{aligned}$$

Hieraus folgt die letzte Gleichung aus Behauptung 3 (ii). □

Nilzyklische Matrizen

Es sei $N_m \in \mathrm{Mat}(m; K)$ definiert durch

$$N_1 = (0), \quad N_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \ldots, \quad N_m = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Dann gilt

$$N_m(e_i) = \begin{cases} e_{i-1} & ; \quad i = 2, \ldots, m \\ 0 & ; \quad i = 1. \end{cases}$$

K^m ist N_m-zyklisch, da $K^m = U(N_m, e_m)$. Die Matrizen N_m heißen die *nilzyklischen* Matrizen.

Lemma 11.54
Es sei $f : V \to V$ nilpotent. U sei ein f-zyklischer Unterraum. Dann gibt es eine Basis von U, so daß $f|U : U \to U$ durch eine nilzyklische Matrix dargestellt wird.

Beweis. Wir wählen $a \in U$ mit $U = U(f, a)$. Es sei $m = \dim U$. Dann ist

$$b_i := f^{m-i}(a); \quad i = 1, \ldots, m$$

nach Lemma (11.52) eine Basis B von U. Es sei $M(f) = M_B(f)$ die darstellende Matrix von $f|U$ mit $M(f) = (\mu_{ij})$. Dann gilt

$$f(b_j) = f(f^{m-j}(a)) = f^{m-(j-1)}(a) = \begin{cases} b_{j-1}, & j > 1 \\ 0, & j = 1. \end{cases}$$

D. h.

$$f(b_j) = \sum_{i=1}^{m} \mu_{ij} b_i = \begin{cases} b_{j-1}, & j > 1 \\ 0, & j = 1. \end{cases}$$

Damit hat man

$$(\mu_{ij}) = (\delta_{i,j-1}) = \begin{pmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

□

Theorem 11.55 (Jordansche Normalform)
Es sei $A \in \mathrm{Mat}(n; K)$. Die Matrix A zerfalle. Dann ist A ähnlich zu einer Matrix der folgenden Form:

$$\begin{pmatrix} \boxed{J_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{J_m} \end{pmatrix}$$

wobei

$$J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & 0 & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}.$$

Bemerkung. J_i heißt ein Jordanblock zum Eigenwert λ_i.

Beweis. Nach dem bereits bewiesenen Struktursatz (8.25) können wir für

$$\chi_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}$$

annehmen, daß

$$A = \begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_1 \end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix} \lambda_r & & * \\ & \ddots & \\ & & \lambda_r \end{matrix}} \end{pmatrix} \begin{matrix} \}m_1 \\ \\ \}m_r \end{matrix}$$

$$\underbrace{\qquad\qquad}_{m_1} \qquad\qquad \underbrace{\qquad\qquad}_{m_r}$$

Wir betrachten zunächst λ_1 (die übrigen Eigenwerte sind analog zu behandeln) und setzen

$$U_{\lambda_1} := \mathrm{Span}(e_1, \dots, e_{m_1}).$$

Dann ist U_{λ_1} ein A invarianter Unterraum, d. h. $A(U_{\lambda_1}) \subset U_{\lambda_1}$. Wir beschränken uns auf die Behandlung dieses Unterraums. Es sei

$$h_{\lambda_1} := A|U_{\lambda_1} : \; U_{\lambda_1} \longrightarrow U_{\lambda_1}.$$

Dann ist

$$h_{\lambda_1} = \lambda_1 \mathrm{id} + f_1,$$

wobei f_1 nilpotent ist. Damit hat man nach Satz (11.53) eine Zerlegung

$$U_{\lambda_1} = U^1_{\lambda_1} \oplus \ldots \oplus U^s_{\lambda_1}$$

in f_1-zyklische Unterräume. Bezüglich jedem $U^i_{\lambda_1}$ wird $f_1|U^i_{\lambda_1}$ nach Lemma (11.54) bei geeigneter Wahl einer Basis durch eine nilzyklische Matrix dargestellt. Also wird f_1 durch eine Matrix der Form

$$\begin{pmatrix} \boxed{N^{(1)}} & & 0 \\ & \ddots & \\ 0 & & \boxed{N^{(s)}} \end{pmatrix}, \quad N^{(i)} \text{ nilzyklisch}$$

dargestellt. Für h_{λ_1} erhält man damit die Matrix

$$\begin{pmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix} \lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_1 \end{matrix}} \end{pmatrix}.$$

□

Bemerkung. Die Räume U_{λ_i} kann man offensichtlich wie folgt beschreiben: Es sei

$$H^k_i := \mathrm{Ker}(A - \lambda_i E)^k.$$

Dann ist

$$H^1_i = \mathrm{Eig}(A, \lambda_i).$$

Es gilt

$$H_i^1 \subset H_i^2 \subset \ldots \subset H_i^{l_i} = H_i^{l_i+1} = \ldots = H_i^{m_i} = \ldots$$

wobei l_i die maximale Größe der Blöcke $N^{(1)}, \ldots, N^{(s)}$ ist. Es gilt

$$U_{\lambda_i} = H_i^{l_i} = H_i^{m_i} = \operatorname{Ker}(A - \lambda_i E)^{m_i}.$$

Definition. U_{λ_i} heißt der *Hauptraum* von A zum Eigenwert λ_i.

Schreibweise. $\operatorname{Hau}(A, \lambda_i) := \operatorname{Ker}(A - \lambda_i E)^{m_i}$.

Kennt man die Jordansche Normalform, so kann man viele Invarianten der Matrix sofort hieraus ablesen. Es sei

$$A \sim \begin{pmatrix} \boxed{J_{\lambda_1}^1} & & & & & \\ & \ddots & & & & \\ & & \boxed{J_{\lambda_1}^{k_1}} & & & \\ & & & \ddots & & \\ & & & & \boxed{J_{\lambda_r}^1} & & \\ & & & & & \ddots & \\ & & & & & & \boxed{J_{\lambda_r}^{k_r}} \end{pmatrix} \quad ; \quad J_{\lambda_i}^j = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & 0 & & \lambda_i \end{pmatrix}$$

$$\begin{aligned} l_i^j &:= \text{Größe von } J_{\lambda_i}^j \quad (i = 1, \ldots, r;\ j = 1, \ldots, k_i), \\ l_i &:= \max\{l_i^j;\ j = 1, \ldots, k_i\}, \\ m_i &:= l_i^1 + \ldots + l_i^{k_i}. \end{aligned}$$

Eigenwerte: $\lambda_1, \ldots, \lambda_r$.

Charakteristisches Polynom: $\chi_A(x) = (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}$.

Minimalpolynom: $\mu_A(x) = (x - \lambda_1)^{l_1} \cdots (x - \lambda_r)^{l_r}$.

Eigenräume: $\dim \operatorname{Eig}(A, \lambda_i) = k_i =$ geometrische Vielfachheit von λ_i.

Haupträume: $\dim \operatorname{Hau}(A, \lambda_i) = m_i =$ algebraische Vielfachheit von λ_i.

Bemerkung. Es sind äquivalent:

$$\begin{aligned} k_i = m_i \text{ für } i = 1, \ldots, r \quad &\Leftrightarrow \quad l_i = 1 \text{ für } i = 1, \ldots, r \\ &\Leftrightarrow \quad \mu_A \text{ hat nur einfache Nullstellen} \\ &\Leftrightarrow \quad A \text{ ist diagonalisierbar.} \end{aligned}$$

Satz 11.56
Die Jordansche Normalform ist bis auf Permutation der Kästchen eindeutig bestimmt.

Beweis. Da die Eigenwerte die gleichen sind, haben wir nur die folgende Situation zu betrachten:

$$J = \begin{pmatrix} \boxed{J^1_{\lambda_1}} & & & & \\ & \ddots & & & \\ & & \boxed{J^{k_1}_{\lambda_1}} & & \\ & & & \ddots & \\ & & & & \boxed{J^1_{\lambda_r}} & & \\ & & & & & \ddots & \\ & & & & & & \boxed{J^{k_r}_{\lambda_r}} \end{pmatrix} \sim \begin{pmatrix} \boxed{J'^1_{\lambda_1}} & & & & \\ & \ddots & & & \\ & & \boxed{J'^{k'_1}_{\lambda_1}} & & \\ & & & \ddots & \\ & & & & \boxed{J'^1_{\lambda_r}} & & \\ & & & & & \ddots & \\ & & & & & & \boxed{J'^{k'_r}_{\lambda_r}} \end{pmatrix} = J'$$

Wir halten nun ein i fest und ordnen die Kästchen zum Eigenwert $\lambda = \lambda_i$ wie folgt der Größe nach an:

$$J: \underbrace{n_1 = \ldots = n_1}_{r_1} > \underbrace{n_2 = \ldots = n_2}_{r_2} > \ldots > \underbrace{n_{j(i)} = \ldots = n_{j(i)}}_{r_{j(i)}}.$$

Es gilt

(1) $$\sum_{l=1}^{j(i)} r_l n_l = m_i \quad \text{(algebraische Vielfachheit von } \lambda_i\text{)}.$$

Analog für J':

$$J': \underbrace{n'_1 = \ldots = n'_1}_{r'_1} > \underbrace{n'_2 = \ldots = n'_2}_{r'_2} > \ldots > \underbrace{n'_{s(i)} = \ldots = n'_{s(i)}}_{r'_{s(i)}}.$$

Dann gilt:

(2) $$\sum_{l=1}^{s(i)} r'_l n'_l = m_i \quad \text{(algebraische Vielfachheit von } \lambda_i\text{)}.$$

Wir können annehmen, daß $j(i) \leq s(i)$ ist. Zu zeigen ist:

$$(n_l, r_l) = (n'_l, r'_l) \quad \text{für } l = 1, \ldots, j(i).$$

Wegen (1), (2) folgt dann sofort $j(i) = s(i)$.

Induktion nach l:

$l = 1$: Es ist $n_1 = n'_1$, da dies gerade den Exponenten des Faktors $(x - \lambda_i)$ im Minimalpolynom ergibt. Ferner ist

$$\left.\begin{array}{l} r_1 = \operatorname{Rang}(J - \lambda_i E)^{n_1 - 1} - \sum\limits_{t \neq i} m_t \\ r'_1 = \operatorname{Rang}(J' - \lambda_i E)^{n_1 - 1} - \sum\limits_{t \neq i} m_t \end{array}\right\} \Rightarrow r_1 = r'_1.$$

$l \mapsto l + 1$: Für $\nu \leq n_l$ gilt

$$\operatorname{Rang}(J' - \lambda_i E)^{\nu} = \sum_{t \neq i} m_t + \begin{cases} r_1(n_1 - \nu) + \ldots + r_l(n_l - \nu) \\ \text{falls } \nu \geq n'_{l+1} \\ r_1(n_1 - \nu) + \ldots + r_l(n_l - \nu) + r'_{l+1} \\ \text{falls } \nu = n'_{l+1} - 1. \end{cases}$$

D. h.

$$n'_{l+1} - 1 = \max\{\nu \leq n_l;\ \operatorname{Rang}(J' - \lambda_i E)^{\nu} > \sum_{t \neq i} m_t + r_1(n_1 - \nu) + \ldots + r_l(n_l - \nu)\}.$$

Die rechte Seite ist für J und J' gleich, also gilt

$$n_{l+1} = n'_{l+1}.$$

Dann ist

$$r'_{l+1} = \operatorname{Rang}(J' - \lambda_i E)^{n_{l+1} - 1} - \sum_{t \neq i} m_t - \sum_{k=1}^{l} r_k(n_k - (n_{l+1} - 1)).$$

Dies zeigt

$$r_{l+1} = r'_{l+1}.$$

$\square$

Zur Berechnung der Jordanform

Sei nun

$$A = (a_{ij}) \in \operatorname{Mat}(n; K)$$

gegeben. Wir nehmen an, daß A zerfällt.

1. Schritt: Berechnung des charakteristischen Polynoms und Zerlegung in Linearformen

$$\begin{aligned} \chi_A(x) &= \det(xE - A) = 0, \\ \chi_A(x) &= (x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r}. \end{aligned}$$

2. Schritt: Wir wissen, daß für $V = K^n$ gilt:

$$V = \mathrm{Hau}(A, \lambda_1) \oplus \ldots \oplus \mathrm{Hau}(A, \lambda_r)$$

mit den Haupträumen

$$\mathrm{Hau}(A, \lambda_i) = \mathrm{Ker}(A - \lambda_i E)^{m_i}.$$

Berechnung: Man bilde sukzessive die Potenzen

$$(A - \lambda_i E)^k \qquad k = 1, \ldots, m_i$$

bis

$$\dim \mathrm{Ker}(A - \lambda_i E)^k = m_i$$

oder äquivalent

$$\mathrm{Rang}(A - \lambda_i E)^k = n - m_i$$

gilt.

3. Schritt: Die Haupträume sind invariant, d. h.

$$A(\mathrm{Hau}(A, \lambda_i)) \subset \mathrm{Hau}(A, \lambda_i).$$

Wir beschränken uns darauf, jetzt die Haupträume einzeln zu behandeln. Sei also nun

$$B : \ K^m \longrightarrow K^m \text{ mit } \chi_B(x) = (x - \lambda)^m.$$

Wir setzen

$$H := B - \lambda E.$$

Es sei p minimal mit $H^p = 0$, d. h.

$$H^{p-1} \neq 0, \quad H^p = 0.$$

Wir setzen

$$\begin{aligned} W_0 &:= \{0\} \\ W_i &:= \mathrm{Ker}(B - \lambda E)^i = \mathrm{Ker}\, H^i, \ (i = 1, \ldots, p). \end{aligned}$$

Dann gilt

$$\begin{array}{ccccccc} W_0 & \subset & W_1 & \subset & W_2 \subset \ldots \subset & W_p & = W = K^m. \\ \| & & \| & & & \| & \\ \{0\} & & \mathrm{Eig}(B, \lambda) & & & \mathrm{Hau}(B, \lambda) & \end{array}$$

Lemma 11.57

(i) $W_{i-1} \subsetneqq W_i$. $(i \leq p)$

(ii) $H^{-1}(W_{i-1}) = W_i$ (hierbei bezeichnet $H^{-1}(W_{i-1})$ das Urbild des Unterraums W_{i-1} unter der Abbildung H).

(iii) Ist $U \subset W$ mit $U \cap W_i = \{0\}$ für ein $i \geq 1$, so ist $H|U$ injektiv.

Beweis. (i): Es sei $W_{i-1} = W_i$. Dann ist

$$(*) \qquad \operatorname{Ker} H^i = \operatorname{Ker} H^{i-1}.$$

Dann gilt für $x \in W$:

$$\begin{aligned} H^p(x) = 0 \;&\Leftrightarrow\; H^i(H^{p-i}(x)) = 0 \overset{(*)}{\Leftrightarrow} H^{i-1}(H^{p-i}(x)) = 0 \\ &\Leftrightarrow\; H^{p-1}(x) = 0. \end{aligned}$$

Also ist $H^{p-1} = 0$ im Widerspruch zur Wahl von p.

(ii): $x \in H^{-1}(W_{i-1}) \Leftrightarrow H(x) \in W_{i-1} \Leftrightarrow H^{i-1}(H(x)) = 0 \Leftrightarrow H^i(x) = 0 \Leftrightarrow x \in W_i$.

(iii): $U \cap W_i = \{0\}$ für ein $i \geq 1 \Rightarrow U \cap W_1 = \{0\} \Rightarrow U \cap \operatorname{Ker} H = \{0\} \Rightarrow H|U$ ist injektiv. □

Lemma 11.58

Es gibt Unterräume $U_1, \ldots, U_p$ von W mit

(i) $W_i = W_{i-1} \oplus U_i$.

(ii) $H(U_i) \subset U_{i-1}$ und $H|U_i$ injektiv $(i \geq 2)$.

(iii) $W = U_1 \oplus \ldots \oplus U_p$.

Beweis. Wir konstruieren die U_i durch absteigende Induktion nach i.

$i = p$: Wir starten mit $W_{p-1} \subsetneqq W_p = W$. Wähle U_p mit

$$W_p = W_{p-1} \oplus U_p.$$

Dann ist $H|U_p$ injektiv.

$i \mapsto i-1$ $(i \geq 3)$: Sei gegeben

$$(*) \qquad W_j = W_{j-1} \oplus U_j \quad (j = i, \ldots, p)$$

mit (i), (ii). Dann gilt

$$U_i \subset W_i \overset{(11.57)\,(ii)}{\Rightarrow} H(U_i) \subset W_{i-1}.$$

Wegen $H^{-1}(W_{i-2}) = W_{i-1}$ folgt aus $(*)$ für $j = i$:

$$H(U_i) \cap W_{i-2} = \{0\}.$$

Ist nämlich $u_i \in U_i$ und $h(u_i) \in H(U_i) \cap W_{i-2}$, dann ist $U_i \in U_i \cap W_{i-1}$, also $u_i = 0$ und damit $h(u_i) = 0$. Wir können also U_{i-1} so wählen, daß

$$U_{i-1} \supset H(U_i), \quad W_{i-1} = W_{i-2} \oplus U_{i-1}.$$

Dann ist auch $H|U_{i-1}$ wieder injektiv.

$i \mapsto (i-1)$ $(i = 2)$: Setze $W_1 =: U_1$. Man hat nun

$$\begin{aligned} W = W_p &= W_{p-1} \oplus U_p = W_{p-2} \oplus U_{p-1} \oplus U_p = \dots \\ &= W_1 \oplus U_2 \oplus \dots \oplus U_p = U_1 \oplus U_2 \oplus \dots \oplus U_p. \end{aligned}$$

Dies zeigt uns, wie die gesuchte Basis zu konstruieren ist. Wir starten mit folgenden Basisvektoren, wobei die Vektoren $u_k^{(l)}$ beliebig gewählt werden können, solange die angegebenen Vektoren eine Basis des jeweiligen Raumes bilden.

$$\begin{array}{cccccc} U_p: & u_1^{(p)}, \dots, & u_{l_p}^{(p)} & & & \\ U_{p-1}: & H(u_1^{(p)}), \dots, & H(u_{l_p}^{(p)}), & u_1^{(p-1)}, \dots, & u_{l_{p-1}}^{(p-1)} & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ U_1: & H^{p-1}(u_1^{(p)}), \dots, & H^{p-1}(u_{l_p}^{(p)}), & H^{p-2}(u_1^{(p-1)}), \dots, & H^{p-2}(u_{l_{p-1}}^{(p-1)}), & \dots, u_1^{(1)}, \dots, u_{l_1}^{(1)} \end{array}$$

Wir konstruieren hieraus wie folgt eine neue Basis:

$$w_1 := H^{p-1}(u_1^{(p)}),\ w_2 := H^{p-2}(u_1^{(p)}),\ \dots,\ w_p =: u_1^{(p)}.$$

Auf diesen Vektoren operiert H wie folgt.

$$w_p \longmapsto w_{p-1} \longmapsto \dots \longmapsto w_1 \longmapsto 0.$$

Wir setzen dieses Verfahren fort:

$$\begin{aligned} w_{p+1} &:= H^{p-1}(u_2^{(p)}), \dots, w_{2p} := u_2^{(p)} \\ &\vdots \\ w_{p(l_p-1)+1} &:= H^{p-1}(u_{l_p}^{(p)}), \dots, w_{pl_p} := u_{l_p}^{(p)} \\ w_{pl_p+1} &:= H^{p-2}(u_1^{(p-1)}), \dots, w_{pl_p+p-1} := u_1^{(p-1)} \\ &\vdots \\ w_{m-l_1+1} &:= u_1^{(1)} \\ &\vdots \\ w_m &:= u_{l_1}^{(1)}. \end{aligned}$$

□

Bemerkung. Dieses Konstruktionsverfahren liefert auch einen unabhängigen Beweis für die Existenz der Jordanschen Normalform.

Definition. Eine Basis, bezüglich derer eine Matrix in Jordanform dargestellt wird, heißt eine *Jordanbasis*. Wir haben also explizit eine Jordanmatrix konstruiert.

Beispiel. Man bestimme die Jordansche Normalform von

$$A = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & 3 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix}.$$

1. Schritt: Berechnung des charakteristischen Polynoms:

$$\begin{aligned} \chi_A(x) &= \det(xE - A) = \det \begin{pmatrix} x-2 & 1 & 0 & -1 \\ 0 & x-3 & 1 & 0 \\ 0 & -1 & x-1 & 0 \\ 0 & 1 & 0 & x-3 \end{pmatrix} \\ &= (x-2)(x-3)[(x-3)(x-1)+1] = (x-2)(x-3)(x^2-4x+4) \\ &= (x-2)^3(x-3). \end{aligned}$$

2. Schritt: Eigenräume, Haupträume.

$\lambda = 3$:

$$(A - 3E)x = 0.$$

$$\begin{pmatrix} -1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\left.\begin{array}{rrrrl} -x_1 & -x_2 & & +x_4 & = 0 \\ & & -x_3 & & = 0 \\ & x_2 & -2x_3 & & = 0 \\ & -x_2 & & & = 0 \end{array}\right\} \Rightarrow x_2 = x_3 = 0,\ x_1 = x_4.$$

Also ist der Eigenraum zu λ_3 gleich

$$\mathrm{Eig}(A, 3) = \mathrm{Span} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

$\lambda = 2:$

$$(A - 2E)x = 0.$$

$$\begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\left.\begin{array}{rrrl} -x_2 & & +x_4 & = 0 \\ x_2 & -x_3 & & = 0 \\ x_2 & -x_3 & & = 0 \\ -x_2 & & +x_4 & = 0 \end{array}\right\} \Rightarrow x_2 = x_3 = x_4.$$

Also gilt

$$\mathrm{Eig}(A, 2) = \mathrm{Span}\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \}.$$

Da $\dim \mathrm{Eig}(A, 2) = 2 < 3$ ist, ist A *nicht* diagonalisierbar.

$$(A - 2E)^2 = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}.$$

Es ist

$$\begin{aligned} (A - 2E)^2 x = 0 \quad &\Leftrightarrow \\ \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad &\Leftrightarrow \quad -2x_2 + x_3 + x_4 = 0. \end{aligned}$$

Also ist

$$\mathrm{Ker}(A - 2E)^2 = \mathrm{Span}\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \}.$$

Damit gilt

$$\dim \mathrm{Ker}(A - 2E)^2 = 3.$$

Da die algebraische Vielfachheit von $\lambda = 3$ gerade 3 ist, folgt

$$\mathrm{Hau}(A - 2E)^2 = \mathrm{Ker}(A - 2E)^2.$$

3. Schritt: Bestimmung einer Basis

$$\begin{array}{ccl} W_0 & = & \{0\} \\ \cap & & \\ W_1 & = & \mathrm{Ker}(A-2E) = \mathrm{Eig}(A,2) \\ \cap & & \\ W_2 & = & \mathrm{Ker}(A-2E)^2 = \mathrm{Hau}(A,2). \end{array}$$

Wähle

$$u_1^{(2)} \in \mathrm{Hau}(A,2) \setminus \mathrm{Eig}(A,2),$$

etwa

$$\begin{array}{rcl} u_1^{(2)} & = & \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}. \\ U_2 & := & \mathrm{Span}(u_1^{(2)}) \end{array}$$

Dann gilt

$$W_2 = W_1 \oplus U_2.$$

Ferner setzen wir

$$U_1 := W_1.$$

Für U_1, U_2 haben wir Basen:

$$\begin{array}{lll} U_2: & u_1^{(2)} & \\ U_1: & H(u_1^{(2)}) = (A-2E)(u_1^{(2)}), & u_1^{(1)} \end{array}$$

mit

$$H(u_1^{(2)}) = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}.$$

Wir können für $u_1^{(1)}$ wählen

$$u_1^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Damit gilt dann

$$\mathrm{Hau}(A,2) = \mathrm{Span}(w_1, w_2, w_3), \text{ mit}$$

$$w_1 = H(u_1^{(2)}),\ w_2 = u_1^{(2)},\ w_3 = u_1^{(1)}.$$

Es gilt

$$\begin{array}{rcl} (A-2E)w_1 & = & 0 \\ (A-2E)w_2 & = & w_1 \\ (A-2E)w_3 & = & 0. \end{array}$$

Also wird bezüglich dieser Basis $A - 2E$ durch die Matrix

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

dargestellt. Damit wird $A|\,\mathrm{Hau}(A,2)$ dargestellt durch

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Schließlich setzen wir

$$B := (w_1, w_2, w_3, w_4) \text{ mit } w_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in \mathrm{Eig}(A,3).$$

Damit erhalten wir als darstellende Matrix

$$M_B(h_A) = \begin{pmatrix} \boxed{\begin{matrix} 2 & 1 \\ 0 & 2 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} \boxed{2} & 0 \\ 0 & \boxed{3} \end{matrix} \end{pmatrix} =: J.$$

4. Schritt: Berechnung der Übergangsmatrix

$$W = (w_1, w_2, w_3, w_4) = \begin{pmatrix} -1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}.$$

Es gilt dann

$$W^{-1}AW = J.$$

Kontrolle:

$$\begin{aligned}
AW &= \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & 3 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 & 2 & 3 \\ -2 & -1 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -2 & -3 & 0 & 3 \end{pmatrix} \\
WJ &= \begin{pmatrix} -1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} -2 & -1 & 2 & 3 \\ -2 & -1 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -2 & -3 & 0 & 3 \end{pmatrix}.
\end{aligned}$$

§ 12 Projektive Räume

V sei ein endlich-dimensionaler Vektorraum. Wir setzen

$$V^* := V \setminus \{0\}.$$

Definition. Für $u, v \in V^*$ sei

$$u \sim v :\Leftrightarrow \text{ Es gibt } \lambda \in K^* \text{ mit } u = \lambda v.$$

Bemerkung.

(i) $\sim$ ist eine Äquivalenzrelation.

(ii) $u \sim v \Leftrightarrow Ku = Kv \Leftrightarrow u, v$ spannen dieselbe Gerade auf.

Definition. Der zu V gehörige *projektive Raum* ist

$$\mathbb{P}(V) := V^*/\sim .$$

Bemerkung.

(i) Als Menge ist $\mathbb{P}(V)$ die Menge der Ursprungsgeraden in V:

$$\mathbb{P}(V) = \{Kv;\ v \in V^*\} \xleftrightarrow{1:1} \{\text{Geraden in } V \text{ durch } 0\}.$$

(ii) Man hat die kanonische Projektion

$$\begin{array}{rccl} \pi: & V^* & \longrightarrow & \mathbb{P}(V) \\ & v & \longmapsto & Kv =: [v]. \end{array}$$

Definition. $\dim \mathbb{P}(V) := \dim V - 1$ heißt die *Dimension* des projektiven Raums $\mathbb{P}(V)$.

Bezeichnung. $\mathbb{P}^n(K) := \mathbb{P}(K^{n+1})$ heißt der *n-dimensionale projektive Raum* über K.

Beispiel.

(i) $\mathbb{P}^1(\mathbb{R}) = \mathbb{P}(\mathbb{R}^2)$

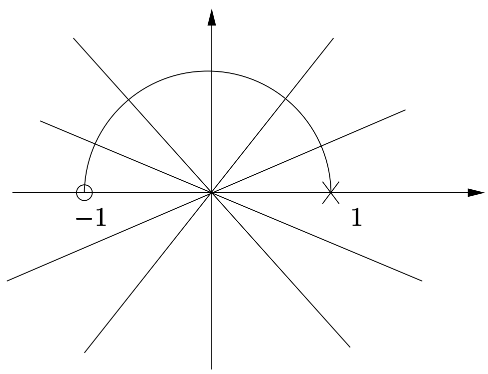


Abb. 3: Die reelle projektive Gerade $\mathbb{P}^1(\mathbb{R})$

Hierbei sind die Punkte -1 und 1 zu identifizieren. Mengentheoretisch gilt

$$\mathbb{P}^1(\mathbb{R}) = S^1 = \mathbb{R} \cup \{\infty\} = \mathbb{R} \cup \mathbb{P}^0(\mathbb{R}).$$

Die Identifikation $S^1 = \mathbb{R} \cup \{\infty\}$ geschieht über die *stereographische Projektion.*

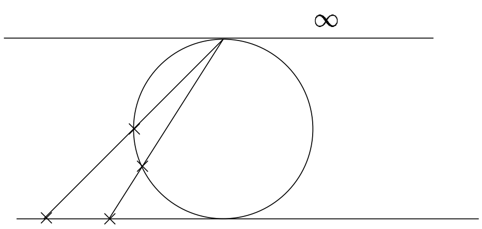


Abb. 4: Stereographische Projektion

(ii) $\mathbb{P}^2(\mathbb{R}) = \mathbb{P}(\mathbb{R}^3)$. Wir haben eine Zerlegung

$$\mathbb{P}^2(\mathbb{R}) = \mathbb{R}^2 \cup \mathbb{P}^1(\mathbb{R}),$$

wobei $\mathbb{P}^1(\mathbb{R})$ gerade den Ursprungsgeraden in der (x, y)-Ebene entspricht.

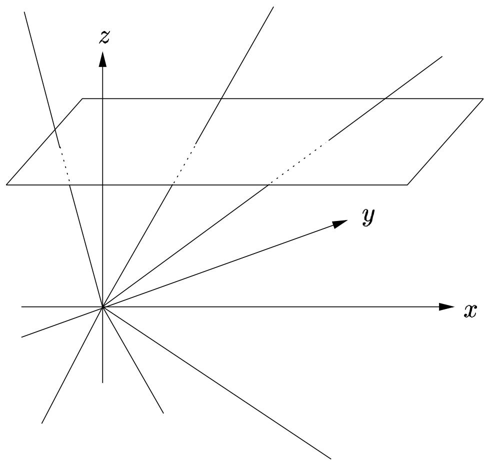


Abb. 5: Zur reellen projektiven Ebene $\mathbb{P}^2(\mathbb{R})$.

Bemerkung. $\mathbb{P}^2(\mathbb{R})$ ist eine nicht-orientierbare differenzierbare Fläche, die nicht in den $\mathbb{R}^3$ eingebettet werden kann.

Es sei nun $U \subset \mathbb{P}(V)$ eine Teilmenge. Wir definieren

$$\tilde{U} := \pi^{-1}(U) \cup \{0\}.$$

Definition. $U \subset \mathbb{P}(V)$ heißt ein *projektiver Unterraum* falls $\tilde{U} \subset V$ ein linearer Unterraum ist.

Bemerkung. Dann ist $U = \mathbb{P}(\tilde{U})$ selbst ein projektiver Raum und hat die Dimension

$$\dim U = \dim_K \tilde{U} - 1.$$

Konvention. $\dim \emptyset := -1$.

Bezeichnungen.

(i) $\dim U = 0$: Punkt,

(ii) $\dim U = 1$: projektive Gerade,

(iii) $\dim U = 2$: projektive Ebene,

(iv) $\dim U = \dim \mathbb{P}(V) - 1$: (projektive) Hyperebene.

Lemma 12.1
Der Durchschnitt von projektiven Unterräumen ist wieder ein projektiver Unterraum.

Beweis. Es sei

$$U_i = \mathbb{P}(\tilde{U}_i); \quad \tilde{U}_i \subset V \quad (i \in I).$$

Dann ist

$$\widetilde{\bigcap_{i\in I} U_i} = \pi^{-1}(\bigcap_{i\in I} U_i) \cup \{0\} = \bigcap_{i\in I} \tilde{U}_i.$$

Dies ist ein linearer Unterraum. □

Definition. Es seien $U_1, \ldots, U_r \subset \mathbb{P}(V)$ projektive Unterräume. Dann ist der *Spann* von $U_1, \ldots, U_r$ definiert durch

$$U_1 \vee \ldots \vee U_r := \bigcap_{\substack{U \supset U_1 \cup \ldots \cup U_r \\ U \text{ ist proj. Unterraum}}} U.$$

Bemerkung.

(i) Der Spann ist der kleinste Unterraum, der $U_1, \ldots, U_r$ enthält.

(ii) $U_1 \vee \ldots \vee U_r = \mathbb{P}(\tilde{U}_1 + \ldots + \tilde{U}_r)$.

Lemma 12.2
$U_1, U_2 \subset \mathbb{P}(V)$ seien projektive Unterräume. Dann gilt

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

Beweis. Es sei $U_i = \mathbb{P}(\tilde{U}_i)$; $i = 1, 2$.

$$\begin{aligned} \dim U_1 + \dim U_2 &= \dim \tilde{U}_1 - 1 + \dim \tilde{U}_2 - 1 \\ &= \dim(\tilde{U}_1 + \tilde{U}_2) - 1 + \dim(\tilde{U}_1 \cap \tilde{U}_2) - 1 \\ &= \dim(U_1 \vee U_2) + \dim(U_1 \cap U_2). \end{aligned}$$

□

Bemerkung. Es sei $\dim \mathbb{P}(V) = 2$ und $L_1, L_2 \subset \mathbb{P}(V)$ seien projektive Geraden. Dann gilt

$$\begin{aligned} \dim(L_1 \cap L_2) &= \dim L_1 + \dim L_2 - \dim(L_1 \vee L_2) \\ &\geq 2 - 2 = 0. \end{aligned}$$

Also ist $L_1 \cap L_2 \neq \emptyset$, d. h. zwei projektive Gerade schneiden sich in einer projektiven Ebene stets.

Homogene Koordinaten

Es sei wieder

$$V = K^{n+1}, \quad \mathbb{P}^n(K) = \mathbb{P}(K^{n+1}).$$

Es sei

$$v = (x_0, \dots, x_n) \neq 0 \qquad \text{(d. h. } x_i \neq 0 \text{ für ein } i).$$

Definition. $(x_0 : x_1 : \dots : x_n) := K(x_0, \dots, x_n) \in \mathbb{P}^n(K)$. $(x_0 : \dots : x_n)$ heißen die *homogenen Koordinaten* von $Kv \in \mathbb{P}^n(K)$.

Bemerkung.

(i) $(x_0 : x_1 : \dots : x_n) = (x'_0 : x'_1 : \dots : x'_n)$ genau dann wenn es ein $\lambda \neq 0$ gibt, mit

$$x_i = \lambda x'_i \quad (i = 0, \dots, n).$$

(ii) Jeder Punkt $P \in \mathbb{P}^n(K)$ bestimmt bis auf einen Skalar $\lambda \neq 0$ ein $(n+1)$-Tupel $(x_0 : \dots : x_n)$. Man schreibt $P = (x_0 : \dots : x_n)$.

Es sei nun $W \subset K^{n+1}$ ein Unterraum. Dann ist W Lösungsmenge eines homogenen Gleichungssystems:

$$W = \left\{ (x_0, \dots, x_n) \in K^{n+1}; \begin{array}{ccc} a_{10}x_0 + \dots + a_{1n}x_n & = & 0 \\ \vdots & \vdots & \vdots \\ a_{m0}x_0 + \dots + a_{mn}x_n & = & 0 \end{array} \right\}.$$

Man kann damit schreiben

$$\mathbb{P}(W) = \{(x_0 : \dots : x_n) \in \mathbb{P}^n(K);\ \sum_{\nu=0}^{n} a_{\mu\nu}x_\nu = 0 \quad \text{für } \mu = 1, \dots, m\}.$$

Dies macht Sinn, da

$$\sum_\nu a_{\mu\nu}x_\nu = 0 \Leftrightarrow \sum_\nu a_{\mu\nu}(\lambda x_\nu) = 0 \qquad (\lambda \in K^*).$$

Wir betrachten nun speziell

$$H_\infty := \{(x_0 : \dots : x_n) \in \mathbb{P}^n(K);\ x_0 = 0\}.$$

Definition. H_∞ heißt die *Hyperebene im Unendlichen*.

Bemerkung. Es sei $x = (x_0 : \ldots : x_n) \notin H_\infty$. Also ist $x_0 \neq 0$, und damit

$$(x_0 : x_1 : \ldots : x_n) = (1 : \frac{x_1}{x_0} : \ldots : \frac{x_n}{x_0}).$$

Sei

$$A := \mathbb{P}^n(K) \setminus H_\infty.$$

Wir betrachten die Abbildung:

$$\begin{array}{rrcl} \iota : & A & \longrightarrow & K^n \\ & (x_0 : x_1 : \ldots : x_n) & \longmapsto & (\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}) =: (y_1, \ldots, y_n). \end{array}$$

Die Abbildung

$$\begin{array}{rrcl} \iota' : & K^n & \longrightarrow & A \\ & (y_1, \ldots, y_n) & \longmapsto & (1 : y_1 : \ldots : y_n) \end{array}$$

ist die Umkehrabbildung von ι. Also ist

$$\iota : \; A \xrightarrow{1:1} K^n$$

eine Bijektion, und mit dieser Identifikation hat man

$$\mathbb{P}^n(K) = A \cup H_\infty = K^n \cup \mathbb{P}^{n-1}(K).$$

Definition. Man nennt $(y_1, \ldots, y_n)$ die *inhomogenen Koordinaten* des Punktes $P = (x_0 : \ldots : x_n) \in A$.

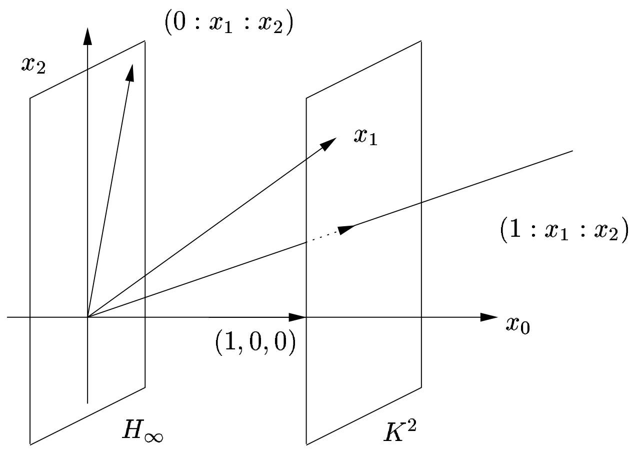


Abb. 6: Geometrische Deutung der Hyperebene im Unendlichen

Sprechweise. Wir nennen $A = \mathbb{P}^n(K) \backslash H_\infty$ den *affinen Teil* von $\mathbb{P}^n(K)$. Mittels $\iota : A \longrightarrow K^n$ identifizieren wir A mit K^n. Man sagt, daß sich der projektive Raum $\mathbb{P}^n(K) = A \cup H_\infty$ aus dem affinen Teil A und der Hyperebene im Unendlichen zusammensetzt.

Bemerkung. Die Auswahl der Hyperebene $H_\infty = \{x_0 = 0\}$ ist willkürlich. Wir hätten im Prinzip auch jede andere Hyperebene als Hyperebene im Unendlichen auswählen können.

Definition.

(i) Eine Teilmenge $W \subset K^n$ heißt ein *affiner Unterraum* des Vektorraums K^n, falls es einen Unterraum W_0 von K^n und ein Element $a \in K^n$ gibt, mit

$$W = a + W_0.$$

(ii) Die *Dimension* von W wird als $\dim W_0$ definiert.

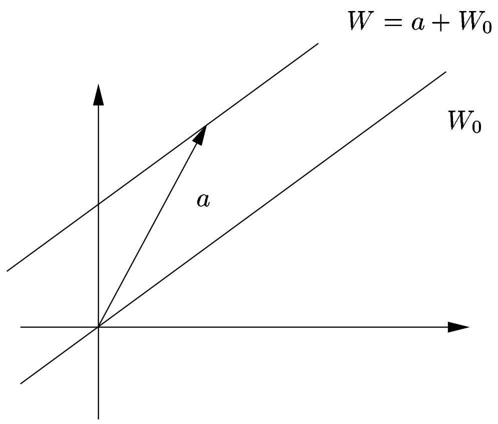


Abb. 7: Affiner Unterraum

Bemerkung. Jeder affine Unterraum $W \subset K^n$ ist die Lösungsmenge eines linearen Gleichungssystems

$$\begin{array}{rcl} a_{11}y_1 + \ldots + a_{1n}y_n & = & b_1 \\ \vdots \qquad\qquad \vdots & & \vdots \\ a_{m1}y_1 + \ldots + a_{mn}y_n & = & b_m. \end{array}$$

Umgekehrt ist die Lösungsmenge eines solchen Gleichungssystems stets ein affiner Unterraum.

Lemma 12.3

(i) Es sei $\overline{W} \subset \mathbb{P}^n(K)$ ein projektiver Unterraum. Dann ist $W := \overline{W} \cap A \subset A = K^n$ ein affiner Unterraum.

(ii) Ist $\emptyset \neq W \subset A = K^n$ ein affiner Unterraum, so gibt es genau einen projektiven Unterraum $\overline{W} \subset \mathbb{P}^n(K)$ mit $\overline{W} \cap A = W$.

Beweis. (i): Die Menge $\widetilde{\overline{W}} := \pi^{-1}(\overline{W}) \cup \{0\}$ ist ein linearer Unterraum von K^{n+1}, d. h. Lösungsmenge eines homogenen linearen Gleichungssystems

$$\widetilde{\overline{W}} = \{(x_0, \dots, x_n);\ a_{10}x_0 + \ldots + a_{1n}x_n = \ldots = a_{m0}x_0 + \ldots + a_{mn}x_n = 0\}.$$

Für einen Punkt

$$x = (x_0 : \ldots : x_n) \notin H_\infty \quad (\text{d. h. } x_0 \neq 0)$$

gilt

$$a_{\mu 0}x_0 + \ldots + a_{\mu n}x_n = 0 \Leftrightarrow a_{\mu 0} + a_{\mu 1}\frac{x_1}{x_0} + \ldots + a_{\mu n}\frac{x_n}{x_0} = 0.$$

Also gilt

$$\overline{W} \cap A = \left\{ (y_1, \dots, y_n);\ \begin{array}{ccc} a_{11}y_1 + \ldots + a_{1n}y_n & = & -a_{10} \\ \vdots \qquad\qquad \vdots & & \vdots \\ a_{m1}y_1 + \ldots + a_{mn}y_n & = & -a_{m0} \end{array} \right\}.$$

D. h. $\overline{W} \cap A$ ist ein affiner Unterraum.

(ii): Sei umgekehrt $W \subset A$ affin. Also

$$W = \left\{ (y_1, \dots, y_n) \in A;\ \begin{array}{ccc} a_{11}y_1 + \ldots + a_{1n}y_n & = & b_1 \\ \vdots \qquad\qquad \vdots & & \vdots \\ a_{m1}y_1 + \ldots + a_{mn}y_n & = & b_m \end{array} \right\}.$$

Setze

$$\overline{W} = \left\{ (x_0 : \ldots : x_n) \in \mathbb{P}^n(K);\ \begin{array}{ccc} -b_1x_0 + a_{11}x_1 + \ldots + a_{1n}x_n & = & 0 \\ \vdots \qquad\qquad \vdots & & \vdots \\ -b_mx_0 + a_{m1}x_1 + \ldots + a_{mn}x_n & = & 0 \end{array} \right\}.$$

Nach Konstruktion ist

$$\overline{W} \cap A = W.$$

Es bleibt die Eindeutigkeit zu zeigen. Es sei also $\overline{W'}$ ein weiterer projektiver Unterraum mit

$$\overline{W'} \cap A = W.$$

Wir setzen

$$\overline{W}_\infty := \overline{W} \cap H_\infty, \quad \overline{W'}_\infty := \overline{W'} \cap H_\infty.$$

Dann sind, da $W \neq \emptyset$ ist,

$$\overline{W}_\infty \subsetneqq \overline{W} \text{ und } \overline{W'}_\infty \subsetneqq \overline{W'}$$

jeweils echte projektive Unterräume. Es gilt

$$\pi^{-1}(W) = \widetilde{\overline{W}} - \widetilde{\overline{W}}_\infty = \widetilde{\overline{W'}} - \widetilde{\overline{W'}}_\infty \neq \emptyset.$$

Das Komplement eines echten Unterraums erzeugt stets den gesamten Vektorraum. Also gilt

$$\widetilde{\overline{W}} = \operatorname{Span}(\widetilde{\overline{W}} - \widetilde{\overline{W}}_\infty) = \operatorname{Span}(\widetilde{\overline{W'}} - \widetilde{\overline{W'}}_\infty) = \widetilde{\overline{W'}}$$

und damit $\overline{W} = \overline{W'}$. □

Beispiel. $\mathbb{P}^2(K)$ (Projektive Ebene über K)

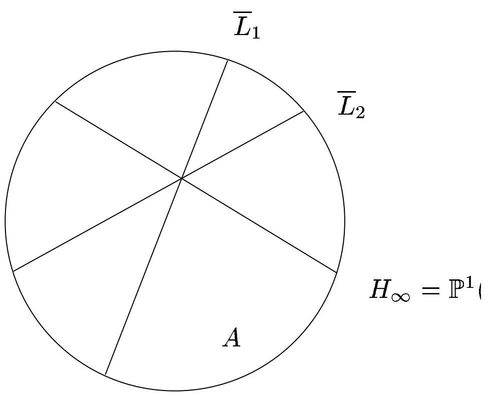


Abb. 8: Parallele Geraden schneiden sich im Unendlichen.

Zwei verschiedene Geraden $\overline{L}_1, \overline{L}_2$ müssen sich in der projektiven Ebene $\mathbb{P}^2(K)$ stets in einem Punkt schneiden. Falls dieser Schnittpunkt in H_∞ liegt (also im Unendlichen) heißen $L_1 = \overline{L}_1 \cap A$ und $L_2 = \overline{L}_2 \cap A$ im affinen Teil *parallel*.

Abbildungen

Es seien V, W endlich-dimensionale Vektorräume und $\varphi : V \to W$ sei ein *injektiver* Homomorphismus. Dann gilt

$$\varphi(Kv) = K\varphi(v) \neq \{0\}$$

und dies induziert eine injektive Abbildung

$$\begin{array}{rccc} \overline{\varphi}: & \mathbb{P}(V) & \longrightarrow & \mathbb{P}(W) \\ & Kv & \longmapsto & K\varphi(v). \end{array}$$

Die Bildmenge

$$\operatorname{Im} \overline{\varphi} = \mathbb{P}(\varphi(V)) \subset \mathbb{P}(W)$$

ist ein projektiver Unterraum.

Beispiel. Es sei $n \geq m$.

$$\begin{array}{rccc} \iota: & K^{m+1} & \hookrightarrow & K^{n+1} \\ & (x_0, \dots, x_m) & \mapsto & (0, \dots, 0, x_0, \dots, x_m) \end{array}$$

ergibt eine Einbettung

$$\overline{\iota}: \quad \mathbb{P}^m(K) \hookrightarrow \mathbb{P}^n(K).$$

Ist $n = m + 1$, so ist das Bild gerade die Hyperebene im Unendlichen.

Ist speziell $\varphi \in \mathrm{Aut}(V)$, so definiert dieser Automorphismus eine bijektive Abbildung

$$\overline{\varphi}: \quad \mathbb{P}(V) \longrightarrow \mathbb{P}(V).$$

Definition. Eine Abbildung $f : \mathbb{P}(V) \to \mathbb{P}(V)$ heißt eine *Projektivität*, falls es einen Automorphismus $\varphi \in \mathrm{Aut}(V)$ gibt mit $f = \overline{\varphi}$.

Bemerkung.

(i) Ist $\overline{\varphi}$ eine Projektivität, so ist

$$\overline{\varphi}^{-1} = \overline{\varphi^{-1}}$$

ebenfalls eine Projektivität.

(ii) Es sei

$$\mathrm{Aut}\,\mathbb{P}(V) = \{\overline{\varphi};\ \overline{\varphi} \text{ ist Projektivität von } \mathbb{P}(V)\}.$$

Dann ist $\mathrm{Aut}\,\mathbb{P}(V)$ mit der üblichen Hintereinanderschaltung eine Gruppe. Es gilt $\overline{\varphi} \circ \overline{\psi} = \overline{\varphi \circ \psi}$.

Lemma 12.4
Es gilt genau dann $\overline{\varphi} = \overline{\psi}$ wenn es ein $\lambda \in K^*$ gibt mit $\varphi = \lambda\psi$.

Beweis. Ist $\varphi = \lambda\psi$, so gilt offensichtlich $\overline{\varphi} = \overline{\psi}$. Es bleibt die Umkehrung zu zeigen:

$$\overline{\varphi} = \overline{\psi} \Rightarrow \overline{\varphi} \circ \overline{\psi}^{-1} = \mathrm{id} \Rightarrow \begin{array}{c} \overline{\varphi} \circ \overline{\psi^{-1}} \\ \| \\ \overline{\varphi \circ \psi^{-1}} \end{array} = \mathrm{id}.$$

Wir setzen also

$$\tau := \varphi \circ \psi^{-1}.$$

Zu zeigen ist $\tau = \lambda \mathrm{id}_V$. Da $\overline{\tau} = \mathrm{id}_{\mathbb{P}(V)}$ ist, gilt

$$\tau(v) = \lambda_v v \text{ für alle } v \in V \text{ mit einem } \lambda_v \in K^*.$$

Zu zeigen ist, daß $\lambda_v = \lambda_w$ für alle v, w gilt.

Annahme: $\lambda_v \neq \lambda_w$. Dann sind v, w notwendig linear unabhängig. Es gilt

$$\left.\begin{aligned} \tau(v+w) &= \tau(v) + \tau(w) = \lambda_v v + \lambda_w w \\ \tau(v+w) &= \lambda_{v+w}(v+w) = \lambda_{v+w} v + \lambda_{v+w} w \end{aligned}\right\} \Rightarrow (\lambda_v - \lambda_{v+w})v + (\lambda_w - \lambda_{v+w})w = 0.$$

Da v, w linear unabhängig sind, folgt hieraus

$$\lambda_v = \lambda_{v+w}, \quad \lambda_w = \lambda_{v+w}$$

und damit $\lambda_v = \lambda_w$ im Widerspruch zur Annahme. □

Folgerung. $\mathrm{Aut}(\mathbb{P}(V)) = \mathrm{GL}(V)/\sim$, wobei

$$\varphi \sim \psi \Leftrightarrow \text{Es gibt ein } \lambda \in K^* \text{ mit } \varphi = \lambda\psi.$$

Ist speziell

$$V = K^{n+1}$$

so erhält man

$$\mathrm{Aut}(\mathbb{P}^n(K)) = \mathrm{GL}(n+1, K)/\sim .$$

Definition. $\mathrm{PGL}(n+1, K) = \mathrm{Aut}(\mathbb{P}^n(K))$ heißt die *projektive lineare Gruppe* (der Dimension n).

Koordinatensysteme

Es sei

$$\mathbb{P} = \mathbb{P}(V), \quad \dim \mathbb{P}(V) = n.$$

Definition. Die Punkte $P_0, \ldots, P_k$ heißen *(linear) unabhängig*, falls $\dim(P_0 \vee \ldots \vee P_k) = k$ gilt.

Bemerkung. Es sei $P_i = Kv_i$; $i = 0, \ldots, k$. Dann gilt

$$P_0, \ldots, P_k \text{ unabhängig} \Leftrightarrow v_0, \ldots, v_k \text{ linear unabhängig.}$$

Beispiele.

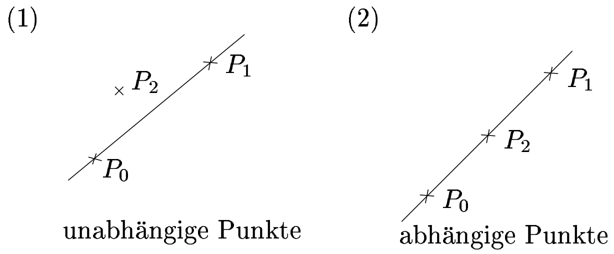


Abb. 9: Linear abhängige/unabhängige Punkte in $\mathbb{P}^2(K)$

Definition. Ein *Koordinatensystem* von $\mathbb{P}$ ist ein geordnetes $(n+2)$-Tupel $(P_0, \dots, P_n, E)$ von denen jeweils $(n+1)$ Punkte linear unabhängig sind. $P_0, \dots, P_n$ heißen die Grundpunkte. E heißt der *Einheitspunkt* des Koordinatensystems.

Lemma 12.5
Es sei $(P_0, \dots, P_n, E)$ ein Koordinatensystem. Dann gibt es eine Basis $v_0, \dots, v_n$ von V mit

(i) $P_i = Kv_i \quad (i = 0, \dots, n)$

(ii) $E = K(v_0 + \ldots + v_n)$.

Die Basis $v_0, \dots, v_n$ ist bis auf einen gemeinsamen Skalar eindeutig bestimmt.

Beweis. Es gibt $v'_0, \dots, v'_n$ mit

$$P_i = Kv'_i \quad (i = 0, \dots, n).$$

Da die P_i unabhängig sind, so sind auch die Vektoren v'_i unabhängig, bilden also eine Basis. Sei

$$E = Ke.$$

Dann gibt es eine Darstellung

$$e = e_0 v'_0 + e_1 v'_1 + \ldots + e_n v'_n.$$

Es sind (auf Grund der Unabhängigkeitsvoraussetzung) alle $e_i \neq 0$. Sei

$$v_i := e_i v'_i.$$

Dies ist eine Basis mit den Eigenschaften (i), (ii). Nun gilt

$$E = K(v_0 + \ldots + v_n) = K(\lambda_1 v_0 + \ldots + \lambda_n v_n)$$

genau wenn $\lambda_0 = \ldots = \lambda_n \in K^*$. Dies liefert die Eindeutigkeit, da die v_i wegen (i) schon bis auf einen Skalar eindeutig festliegen. $\square$

Es sei nun $(P_0, \dots, P_n, E)$ ein Koordinatensystem. Sei $B = (v_0, \dots, v_n)$ eine zugehörige Basis. Dies liefert einen Isomorphismus

$$\begin{array}{rrcl} q_B: & V & \longrightarrow & K^{n+1} \\ & v_i & \longmapsto & e_i. \end{array}$$

Diese Abbildung liegt bis auf einen Skalar fest. Also ist

$$\overline{q}_B: \quad \mathbb{P}(V) \longrightarrow \mathbb{P}^n(K)$$

eindeutig festgelegt. Wir haben damit

$$\begin{array}{rrcl} \overline{q}_B: & \mathbb{P}(V) & \longrightarrow & \mathbb{P}^n(K) \\ & P = Kv & \longmapsto & Kq_B(v) = K(x_0, \dots, x_n). \end{array}$$

Definition. $(x_0 : \ldots : x_n)$ heißen die *homogenen Koordinaten* des Punktes P bezüglich des Koordinatensystems $(P_0, \dots, P_n, E)$.

Schreibweise. $P = (x_0 : \ldots : x_n)$. Die homogenen Koordinaten sind wieder bis auf einen gemeinsamen von Null verschiedenen Skalar festgelegt. Damit gilt:

$$\begin{array}{rcl} P_0 & = & (1:0:0:\ldots:0) \\ P_1 & = & (0:1:0:\ldots:0) \\ \vdots & & \vdots \\ P_n & = & (0:0:\ldots:0:1) \\ E & = & (1:1:\ldots:1:1). \end{array}$$

Es sei nun

$$\overline{\varphi}: \quad \mathbb{P}(V) \longrightarrow \mathbb{P}(V)$$

eine Projektivität. Dann hat man ein kommutatives Diagramm

$$\begin{tikzcd} \mathbb{P}(V) \arrow[r, "\overline{\varphi}"] \arrow[d, "\overline{q}_B"'] & \mathbb{P}(V) \arrow[d, "\overline{q}_B"] \\ \mathbb{P}^n(K) \arrow[r, "\overline{\psi}"] & \mathbb{P}^n(K). \end{tikzcd}$$

$\overline{\psi}$ ist durch $\overline{\varphi}$ eindeutig bestimmt und wieder eine Projektivität. Also gibt es eine Matrix $A \in \mathrm{GL}(n+1; K)$ mit $\overline{\psi} = \overline{A}$. Für

$$P = (x_0 : \ldots : x_n), \quad \overline{\varphi}(P) = (y_0 : \ldots : y_n)$$

gilt

$$\begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix}.$$

Dabei ist A bis auf einen Skalar $\neq 0$ festgelegt.

Damit reduziert sich das Studium der Projektivitäten beliebiger projektiver Räume auf das Studium von Projektivitäten des $\mathbb{P}^n(K)$.

Affinitäten

V sei wiederum ein endlich-dimensionaler K-Vektorraum.

Definition. Eine Abbildung

$$f: \quad V \longrightarrow V$$

heißt eine *Affinität*, falls es einen Automorphismus f_0 von V und einen Vektor $t \in V$ gibt, mit

$$f(v) = f_0(v) + t \quad (v \in V).$$

Bemerkung.

(i) Affine Abbildungen entstehen also aus der Hintereinanderschaltung einer linearen Abbildung mit einer Translation.

(ii) Ist f eine Affinität, so ist f bijektiv und f^{-1} ist wieder eine Affinität. Es gilt:

$$f^{-1}(v) = f_0^{-1}(v) - f_0^{-1}(t).$$

Denn:

$$\begin{aligned}
(f^{-1} \circ f)(v) &= f_0^{-1}(f(v)) - f_0^{-1}(t) \\
&= f_0^{-1}(f_0(v) + t) - f_0^{-1}(t) \\
&= v + f_0^{-1}(t) - f_0^{-1}(t) = v. \\
(f \circ f^{-1})(v) &= f_0(f^{-1}(v)) + t \\
&= f_0(f_0^{-1}(v) - f_0^{-1}(t)) + t \\
&= v - t + t = v.
\end{aligned}$$

(iii) Sind f, g Affinitäten, so auch $f \circ g$:

$$\begin{aligned}
f(v) &= f_0(v) + t \\
g(v) &= g_0(v) + s.
\end{aligned}$$

Dann gilt

$$\begin{aligned}(f \circ g)(v) &= f(g(v)) = f(g_0(v) + s) = f_0(g_0(v) + s) + t \\ &= f_0(g_0(v)) + f_0(s) + t \\ &= (\underbrace{f_0 \circ g_0}_{\in \mathrm{Aut}(V)})(v) + (\underbrace{f_0(s) + t}_{\in V}).\end{aligned}$$

Speziell: $V = K^n$.

$$\begin{aligned} f: \quad K^n &\longrightarrow K^n \\ x &\longmapsto Ax + b \quad (A \in \mathrm{GL}(n; K), b \in K^n).\end{aligned}$$

Definition. $\mathrm{Aff}(V) := \{f;\ f : V \to V \text{ ist Affinität}\}$ heißt die *Gruppe der Affinitäten* auf V.

Wir untersuchen nun den Zusammenhang zwischen Projektivitäten und Affinitäten. Dazu erinnern wir an die Zerlegung

$$\mathbb{P}^n(K) = K^n \cup H_\infty = K^n \cup \mathbb{P}^{n-1}(K).$$

Satz 12.6

(i) Es sei $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ eine Projektivität mit $f(H_\infty) = H_\infty$. Dann ist $f_a = f|_{K^n} : K^n \to K^n$ eine Affinität.

(ii) Ist $f_a : K^n \to K^n$ eine Affinität, so gibt es genau eine Projektivität $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ mit $f|_{K^n} = f_a$.

Beweis. (ii): Es sei

$$f_a(x) = A_0 x + t \quad (A_0 \in \mathrm{GL}(n; K)).$$

Setzt man

$$f_a(x) = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$$

so ergibt sich

(1)
$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = A_0 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}.$$

Dies ist äquivalent zu

$$(2) \qquad \begin{pmatrix} 1 \\ x_1' \\ \vdots \\ x_n' \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & \cdots & 0 \\ t_1 & & & \\ \vdots & & A_0 & \\ t_n & & & \end{pmatrix}}_{A \in \mathrm{GL}(n+1;K)} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Man kann also setzen

$$f = \overline{A}.$$

Um die *Eindeutigkeit* zu sehen, genügt:

Behauptung: Ist $g : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ eine Projektivität mit $g|_{K^n} = \mathrm{id}_{K^n}$, so ist $g = \mathrm{id}_{\mathbb{P}^n(K)}$.

Denn: Man kann wie folgt argumentieren: Zunächst ist $g(H_\infty) = H_\infty$. Sei $g|_{H_\infty} \neq \mathrm{id}_{H_\infty}$. Nach Voraussetzung gibt es also einen Punkt $x \in H_\infty$ mit $g(x) \neq x$. Wir betrachten die Gerade $P \vee x$ mit einem Punkt $P \notin H_\infty$. Da

$$P \vee x = \mathbb{P}(K^2)$$

ist $\#(P \vee x) \geq 3$. Also gibt es $Q \neq P$, $Q \notin H_\infty$ auf $P \vee x$ mit $g(Q) \neq Q$. Dies ist ein Widerspruch dazu, daß $g|_{K^n} = \mathrm{id}_{K^n}$ gilt.

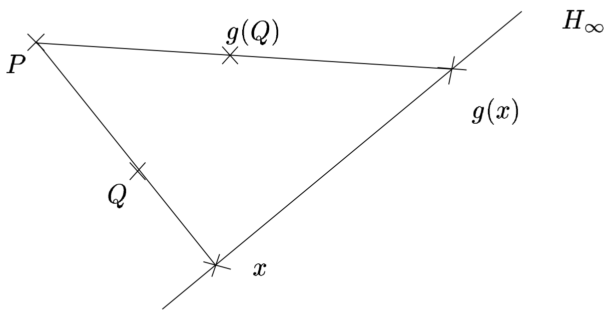


Abb. 10: Zum Beweis von Satz (12.6)

(i): Es sei f gegeben durch

$$\begin{pmatrix} x_0' \\ \vdots \\ x_n' \end{pmatrix} = A \begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix} \qquad A \in \mathrm{GL}(n+1;K).$$

Wegen $f(H_\infty) = H_\infty$ muß gelten

$$A = \begin{pmatrix} a_{00} & 0 & & 0 \\ a_{10} & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n0} & a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Wegen $A \in \mathrm{GL}(n+1; K)$ ist $a_{00} \neq 0$. Man kann also annehmen, daß

$$A = \left(\begin{array}{c|ccc|} 1 & 0 & \cdots & 0 \\ \hline a_{10} & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n0} & a_{n1} & \cdots & a_{nn} \\ \hline \end{array} \right).$$

$$\underbrace{\phantom{a_{11} \cdots a_{1n}}}_{=:A_0 \in \mathrm{GL}(n;K).}$$

Es gilt daher

$$\begin{pmatrix} 1 \\ x_1' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} 1 & 0 \quad \cdots \quad 0 \\ a_{10} & \\ \vdots & A_0 \\ a_{n0} & \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

d. h.

$$f_a \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A_0 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \underbrace{\begin{pmatrix} a_{10} \\ \vdots \\ a_{n0} \end{pmatrix}}_{=:t}.$$

$\square$

§ 13 Quadriken in projektiven Räumen

Es sei

$$\mathbb{P}^n := \mathbb{P}^n(K) = \mathbb{P}(K^{n+1}), \quad \operatorname{char} K \neq 2.$$

Definition. Eine *homogene quadratische Form* in $n+1$ Variablen ist ein Polynom der Form

$$q(x_0, \dots, x_n) = \sum_{\nu,\mu=0}^{n} a_{\nu\mu} x_\nu x_\mu \qquad (a_{\nu\mu} \in K).$$

Bemerkung.

(i) Wir haben bereits gesehen, daß wir $a_{\nu\mu} = a_{\mu\nu}$ annehmen können (hier verwenden wir $\operatorname{char} K \neq 2$).

(ii) Unter der Annahme $a_{\nu\mu} = a_{\mu\nu}$ ist die Matrix $A = (a_{\nu\mu})$ symmetrisch, und es gilt

$$q(x) = q(x_0, \dots, x_n) = {}^t x A x =: q_A(x).$$

(iii) Für jeden Skalar $\lambda \in K$ gilt

$$q(\lambda x_0, \dots, \lambda x_n) = \lambda^2 q(x_0, \dots, x_n).$$

Definition. Eine Teilmenge $Q \subset \mathbb{P}^n(K)$ heißt eine *Quadrik* (*Hyperfläche 2. Ordnung*), falls es eine homogene quadratische Form q gibt, mit

$$Q = \{x = (x_0 : \ldots : x_n) \in \mathbb{P}^n(K);\ q(x) = 0\}.$$

Bemerkung. Dies ist wegen der obigen Bemerkung (iii) wohldefiniert.

Schreibweise. $Q = \{q = 0\}$.

Beispiele. $\mathbb{P}^2 = \mathbb{P}^2(\mathbb{R})$.

(i) $x_0^2 + x_1^2 - x_2^2 = 0$, d. h. $A = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & -1 \end{pmatrix}$.

Wir wollen den *affinen* (eigentlichen) Teil betrachten, wobei wir die Wahl der Hyperebene im Unendlichen variieren.

(α) $H_\infty = \{x_0 = 0\}$. Auf $\mathbb{R}^2 = \mathbb{P}^2 \setminus H_\infty$ erhalten wir

$$1 + \left(\frac{x_1}{x_0}\right)^2 - \left(\frac{x_2}{x_0}\right)^2 = 0$$

bzw.

$$y_2^2 - y_1^2 = 1.$$

Diese Gleichung definiert eine Hyperbel.

(β) $H_\infty = \{x_2 = 0\}$.

$$\left(\frac{x_0}{x_2}\right)^2 + \left(\frac{x_1}{x_2}\right)^2 - 1 = 0$$

bzw.

$$y_0^2 + y_1^2 = 1.$$

Diese Gleichung ergibt einen Kreis.

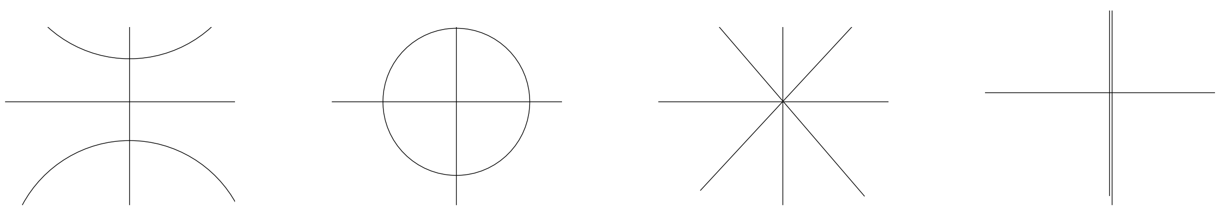
(ii) $x_0^2 - x_1^2 = 0$. Diese Gleichung zerfällt in zwei Linearfaktoren.

$$(x_0 - x_1)(x_0 + x_1) = 0.$$

(iii) $x_0^2 = 0$.

(iv) $x_0^2 + x_1^2 + x_2^2 = 0$. Hier ist $Q = \emptyset$.

(v) $q \equiv 0$. Hier ist $Q = \mathbb{P}^2(\mathbb{R})$.



(i)(α): Hyperbel | (i)(β): Kreis | (ii): Geradenpaar | (iii): Zusammenfallende Geraden

Abb. 11: Darstellungen zu den Beispielen

Definition. Zwei Quadriken $Q, Q' \subset \mathbb{P}^n(K)$ heißen (*projektiv*) *äquivalent*, falls es eine Projektivität $\varphi : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ gibt mit $\varphi(Q) = Q'$.

Bezeichnung. $Q \sim Q'$.

Satz 13.1
Es sei $A = {}^tA$ und $Q = \{q_A = 0\}$. Ferner sei $Q' \subset \mathbb{P}^n(K)$ eine weitere Quadrik. Dann sind äquivalent:

(i) $Q \sim Q'$.

(ii) Es gibt $S \in \mathrm{GL}(n+1, K)$, so daß $Q' = \{q_B = 0\}$, wobei $B = {}^tSAS$.

Beweis. (i)$\Rightarrow$(ii): Es sei $\varphi \in \mathrm{Aut}(\mathbb{P}^n(K))$ eine Projektivität mit $\varphi(Q') = Q$. Sei $\varphi = \overline{S}$ mit $S \in \mathrm{GL}(n+1, K)$. Es gilt

$$\begin{aligned} x \in Q' &\Leftrightarrow \varphi(x) \in Q \Leftrightarrow {}^t(Sx)ASx = 0 \Leftrightarrow {}^tx\,{}^tSASx = 0 \\ &\Leftrightarrow {}^txBx = 0. \end{aligned}$$

Also ist

$$Q' = \{q_B = 0\}.$$

(ii)$\Rightarrow$(i): Wir betrachten $\varphi = \overline{S} \in \mathrm{Aut}(\mathbb{P}^n(K))$. Dann gilt:

$$\begin{aligned} x \in Q' &\Leftrightarrow {}^txBx = 0 \Leftrightarrow {}^tx\,{}^tSASx = 0 \Leftrightarrow {}^t(Sx)A(Sx) = 0 \\ &\Leftrightarrow \varphi(x) \in Q. \end{aligned}$$

Also gilt

$$\varphi(Q') = Q.$$

□

Definition. Für eine Quadrik $Q \subset \mathbb{P}^n(K)$ setzen wir

$$u(Q) := \max\{\dim U;\ U \subset Q \text{ ist ein projektiver Unterraum}\}.$$

Bemerkung. $Q \sim Q' \Rightarrow u(Q) = u(Q')$.

Definition.

(i) Ein Punkt $x \in Q$ heißt *Doppelpunkt* (*singulärer Punkt*) von Q, falls für jede Gerade $L \subset \mathbb{P}^n(K)$ durch x gilt: $L \cap Q = \{x\}$ oder $L \subset Q$.

(ii) $\mathrm{Sing}(Q) := \{x \in Q;\ x \text{ ist ein Doppelpunkt}\}$ heißt der *singuläre Ort* von Q.

(iii) Q heißt *regulär* (*glatt*), falls $\mathrm{Sing}\, Q = \emptyset$ gilt. Ansonsten heißt Q *singulär*.

Beispiele. $\mathbb{P}^2 = \mathbb{P}^2(\mathbb{R})$.

(i) $Q = \{x_0x_1 = 0\}$: Hier ist $u(Q) = 1$.

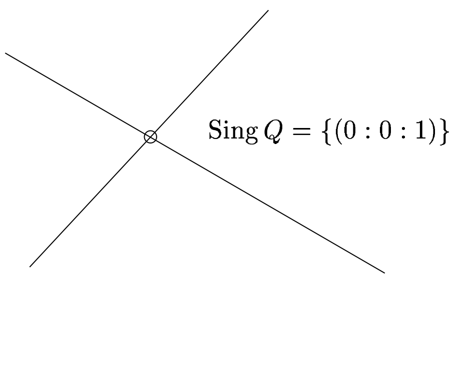


Abb. 12: Singulärer Ort eines Geradenpaares

(ii) $Q = \{x_0^2 + x_1^2 - x_2^2\}$. Für $x_0 = 1$ erhalten wir die Hyperbel $y_2^2 - y_1^2 = 1$ (siehe vorheriges Beispiel). Hier gilt $u(Q) = 0$ und $\mathrm{Sing}(Q) = \emptyset$. (Wir werden dies später streng beweisen.)

(iii) $Q = \{x_0^2 = 0\}$. Hier gilt $u(Q) = 1$ und $\mathrm{Sing}\, Q = Q$.

Satz 13.2
Es sei $Q = \{q_A = 0\} \subset \mathbb{P}^n(K)$ eine Quadrik. Dann gilt

$$\mathrm{Sing}\, Q = \{x \in \mathbb{P}^n(K);\ Ax = 0\} = \mathbb{P}(\mathrm{Ker}\, A).$$

Insbesondere ist $\mathrm{Sing}\, Q \subset \mathbb{P}^n(K)$ ein projektiver Unterraum.

Beweis. Es sei L eine Gerade durch die zwei Punkte $x \neq y \in \mathbb{P}^n(K)$. Dann ist

$$L = \mathbb{P}(\{\lambda x + \mu y;\ \lambda, \mu \in K\})$$

die Verbindungsgerade der Punkte x und y. (Wir verwenden hier dieselben Bezeichnungen für Vektoren in K^{n+1} und die hierdurch definierten Punkte in $\mathbb{P}^n(K)$. Es gilt nun

$$(*)\qquad \begin{aligned} q_A(\lambda x + \mu y) &= {}^t(\lambda x + \mu y)A(\lambda x + \mu y) \\ &= \lambda^2\, {}^t x A x + \mu^2\, {}^t y A y + \lambda\mu\, {}^t y A x + \lambda\mu\, {}^t x A y \\ &= \lambda^2\, {}^t x A x + \mu^2\, {}^t y A y + 2\lambda\mu\, {}^t y A x \\ &= \lambda^2 q_A(x) + \mu^2 q_A(y) + 2\lambda\mu\, {}^t y A x. \end{aligned}$$

„$\supset$“: Es sei x gegeben mit $x \in \mathrm{Ker}\, A$. Ferner sei L eine Gerade durch x, die Q in einem weiteren Punkt y trifft. Zu zeigen ist, daß dann L in Q enthalten ist. Nach $(*)$ gilt:

$$q_A(\lambda x + \mu y) = \lambda^2 q_A(x) + \mu^2 q_A(y) + 2\lambda\mu\, {}^t y A x = 0,$$

da $q_A(x) = q_A(y) = 0$ und $Ax = 0$ gilt.

„$\subset$“: Es sei $x \in \mathrm{Sing}\, Q$. Dann gilt nach $(*)$:

$$(**)\qquad q_A(\lambda x + \mu y) = \mu^2 q_A(y) + 2\lambda\mu\, {}^t y A x.$$

Behauptung: ${}^t y A x = 0$ für alle y. (Hieraus folgt dann, daß $Ax = 0$ gilt.)

1. Fall: $y \in Q$. Ist $y = x$, so ist ${}^t y A y = q_A(y) = 0$. Ansonsten ist nach Voraussetzung $L \subset Q$ und aus $(*)$ folgt dann $2\lambda\mu\, {}^t y A x = 0$ für alle λ und μ, also ${}^t y A x = 0$.

2. Fall: $y \notin Q$. Dann gilt, da $x \in \mathrm{Sing}\, Q$ ist, daß $L \cap Q = \{x\}$ ist, d. h. $q_A(\lambda x + \mu y) \neq 0$ für $\mu \neq 0$. Wir nehmen an, daß ${}^t y A x =: c \neq 0$ ist. Da $y \notin Q$ ist $d := q_A(y) \neq 0$. Es sei nun $\mu = 1$, $\lambda = -\frac{d}{2c}$. Dann gilt

$$\mu^2 q_A(y) + 2\lambda\mu\, {}^t y A x = d - 2\frac{d}{2c}c = 0,$$

ein Widerspruch. □

Definition. $d(Q) := \dim \operatorname{Sing} Q$.

Bemerkung.

(i) Q ist regulär $\Leftrightarrow d(Q) = -1$.

(ii) $Q \sim Q' \Rightarrow d(Q) = d(Q')$.

Theorem 13.3
Es sei $K = \mathbb{R}$. Dann ist jede Quadrik $Q \subset \mathbb{P}^n(\mathbb{R})$ äquivalent zu genau einer der folgenden Quadriken:

$$Q_{t,r}: \quad x_0^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = 0$$

mit $-1 \leq t \leq r \leq n$ und $t + 1 \geq r - t$. Es ist

$$u(Q_{t,r}) = n - t - 1, \quad d(Q_{t,r}) = n - r - 1.$$

Insbesondere gilt

$$Q \sim Q' \Leftrightarrow u(Q) = u(Q') \text{ und } d(Q) = d(Q').$$

Bemerkungen.

(i) r heißt der *Rang* der Quadrik.

(ii) Die obigen Gleichungen nennt man die *Normalformen*.

(iii) $Q_{-1,-1} = \mathbb{P}^n(\mathbb{R})$.

(iv) $Q_{n,n} = \emptyset$.

Beweis. Es sei

$$Q = \{q_A(x) = 0\}, \quad A = {}^tA.$$

Nach Satz (10.3) gibt es $S \in \mathrm{GL}(n+1, \mathbb{R})$ mit

$$ {}^tSAS = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix} =: A_{t,r}.$$

$$\underbrace{\underbrace{\quad\quad}_{t+1}\underbrace{\quad\quad}_{r-t}}_{r+1}$$

Da

$$q_A(x) = 0 \Leftrightarrow q_{-A}(x) = 0$$

können wir annehmen, daß $t + 1 \geq r - t$ ist. Sei

$$Q_{t,r} := \{q_{A_{t,r}} = 0\}.$$

Wir können also jede Quadrik in eine der behaupteten Normalformen bringen. Es bleibt zu zeigen, daß

$$u(Q_{t,r}) = n - t - 1,\ d(Q_{t,r}) = n - r - 1.$$

Die zweite Behauptung folgt mit Satz (13.2), da

$$\mathrm{Sing}(Q_{t,r}) = \{x = (x_0 : \ldots : x_n);\ A_{t,r}x = 0\},$$

d. h.

$$\mathrm{Sing}(Q_{t,r}) = \{x = (x_0 : \ldots : x_n);\ x_0 = \ldots = x_r = 0\}.$$

Also gilt

$$\dim \mathrm{Sing}(Q_{t,r}) = n - r - 1.$$

Es sei nun

$$\tilde{U} := \{x;\ x_0 - x_{t+1} = x_1 - x_{t+2} = \ldots = x_{r-t-1} - x_r = x_{r-t} = \ldots = x_t = 0\}.$$

Sei

$$U := \mathbb{P}(\tilde{U}).$$

Dann ist $U \subset Q_{t,r}$ und $\dim U = n - t - 1$. Also gilt

$$u(Q_{t,r}) \geq n - t - 1.$$

Um zu zeigen, daß $u(Q_{t,r}) \leq n - t - 1$ gilt, betrachten wir

$$\tilde{W} := \{x;\ x_{t+1} = \ldots = x_n = 0\}.$$

Sei

$$W := \mathbb{P}(\tilde{W}).$$

Dann ist

$$\dim W = t, \quad W \cap Q = \emptyset.$$

Annahme: $u(Q_{t,r}) > n - t - 1$.

Dann sei $U' \subset Q$ ein projektiver Unterraum mit $\dim U' > n - t - 1$. Es ist

$$\begin{aligned} \dim(U' \cap W) &= \dim U' + \dim W - \dim(U \vee W') \\ &> (n - t - 1) + t - n = -1. \end{aligned}$$

Also ist

$$U' \cap W \neq \emptyset$$

im Gegensatz zu $U' \subset Q$, $W \cap Q = \emptyset$. □

Projektive Klassifikation von Quadriken in $\mathbb{P}^n(\mathbb{R})$

In den beiden folgenden Tabellen listen wir alle Normalformen von Quadriken in $\mathbb{P}^2(\mathbb{R})$ und $\mathbb{P}^3(\mathbb{R})$ auf.

$\mathbb{P}^2(\mathbb{R})$, $n = 2$: Es gibt sechs verschiedene Äquivalenzklassen.

r	t	d	u	Gleichung (Normalform)	Bezeichnung
-1	-1	2	2	$0 = 0$	projektive Ebene
0	0	1	1	$x_0^2 = 0$	Doppelgerade
1	0	0	1	$x_0^2 - x_1^2 = 0$	Geradenpaar
1	1	0	0	$x_0^2 + x_1^2 = 0$	Doppelpunkt
2	1	-1	0	$x_0^2 + x_1^2 - x_2^2 = 0$	nicht-ausgeartete Kurve
2	2	-1	-1	$x_0^2 + x_1^2 + x_2^2 = 0$	leere Menge

$\mathbb{P}^3(\mathbb{R})$, $n = 3$: Es gibt neun verschiedene Äquivalenzklassen.

r	t	d	u	Gleichung (Normalform)	Bezeichnung
-1	-1	3	3	$0 = 0$	3-dim. projektiver Raum
0	0	2	2	$x_0^2 = 0$	Doppelebene
1	0	1	2	$x_0^2 - x_1^2 = 0$	Ebenenpaar
1	1	1	1	$x_0^2 + x_1^2 = 0$	Doppelgerade
2	1	0	1	$x_0^2 + x_1^2 - x_2^2 = 0$	Kegel
2	2	0	0	$x_0^2 + x_1^2 + x_2^2 = 0$	Doppelpunkt
3	1	-1	1	$x_0^2 + x_1^2 - x_2^2 - x_3^2 = 0$	nicht-ausgeartete Fläche, die Geraden enthält (Ringfläche)
3	2	-1	0	$x_0^2 + x_1^2 + x_2^2 - x_3^2 = 0$	nicht-ausgeartete Fläche, die keine Geraden enthält (Ovalfläche)
3	3	-1	-1	$x_0^2 + x_1^2 + x_2^2 + x_3^2 = 0$	leere Menge

Wir betrachten nun den Fall $K = \mathbb{C}$. Sei also

$$A = {}^tA \in M(n; \mathbb{C}).$$

Dies definiert eine *symmetrische Bilinearform*

$$\begin{array}{rrcl} \langle\, , \,\rangle_A : & \mathbb{C}^n \times \mathbb{C}^n & \longrightarrow & \mathbb{C} \\ & (x, y) & \longmapsto & \langle x, y\rangle_A = {}^txAy. \end{array}$$

Bemerkung. Dies unterscheidet sich von der Behandlung in der Linearen Algebra I, wo wir hermitesche (und nicht symmetrische) Matrizen und damit hermitesche Sesquilinearformen (und nicht symmetrische Bilinearformen) betrachtet haben.

Lemma 13.4
Es gibt eine Basis $v_1, \dots, v_n$ von $\mathbb{C}^n$ mit $\langle v_i, v_j\rangle_A = 0$ für $i \neq j$.

Beweis. Induktion nach n.

$n = 1$: Hier ist nichts zu zeigen.

$n - 1 \mapsto n$: Es gilt zunächst:

$$\langle x, y\rangle_A = \frac{1}{2}(\langle x+y, x+y\rangle_A - \langle x, x\rangle_A - \langle y, y\rangle_A). \tag{1}$$

1. Fall: $\langle x, x\rangle_A = 0$ für alle x.

Dann gilt wegen (1), daß

$$\langle x, y\rangle_A = 0 \qquad (\text{für alle } x, y).$$

Damit kann man jede Basis nehmen.

2. Fall: Es gibt v_1 mit $\langle v_1, v_1\rangle_A \neq 0$. Es sei

$$W := \{y \in \mathbb{C}^n;\ \langle v_1, y\rangle_A = 0\}.$$

Behauptung: $\mathbb{C}^n = \mathbb{C}v_1 \oplus W$.

(i) $\mathbb{C}v_1 \cap W = \{0\}$: Nach Konstruktion.

(ii) $\mathbb{C}^n = \mathbb{C}v_1 + W$: Es sei $v \in \mathbb{C}^n$.

Setzen wir

$$\tilde{v} := \frac{\langle v_1, v\rangle_A}{\langle v_1, v_1\rangle_A} v_1,$$

dann gilt

$$\langle v_1, v - \tilde{v}\rangle_A = \langle v_1, v\rangle_A - \frac{\langle v_1, v\rangle_A}{\langle v_1, v_1\rangle_A}\langle v_1, v_1\rangle_A = 0.$$

Also ist $v - \tilde{v} \in W$, und damit

$$v = \underset{\mathbb{C}v_1}{\underset{\Cap}{\tilde{v}}} + \underset{W}{\underset{\Cap}{(v - \tilde{v})}}.$$

Nach Induktionsvoraussetzung gibt es eine Basis $v_2, \dots, v_n$ von W mit

$$\langle v_i, v_j\rangle_A = 0 \qquad (2 \leq i \neq j \leq n).$$

Wir können daher die Basis

$$B = (v_1, \dots, v_n).$$

wählen. $\square$

Lemma 13.5
Es gibt $S \in \mathrm{GL}(n, \mathbb{C})$ mit

$$ {}^tSAS = \begin{pmatrix} {\scriptstyle 1} & & & & \\ & \ddots & & 0 & \\ & & {\scriptstyle 1} & & \\ & & & {\scriptstyle 0} & \\ & 0 & & & \ddots \\ & & & & & {\scriptstyle 0} \end{pmatrix} \left.\vphantom{\begin{matrix} 1 \\ \ddots \\ 1 \end{matrix}}\right\} r $$

Dabei ist r der Rang von A.

Beweis. Es sei $v_1, \ldots , v_n$ wie oben. Wir können nach Umordnung annehmen, daß

$$ \begin{aligned} \langle v_i, v_i \rangle_A &\neq 0, \qquad && i = 1, \ldots , r \\ \langle v_i, v_i \rangle_A &= 0, \qquad && i \geq r + 1. \end{aligned} $$

Es sei

$$ S_1 := (v_1, \ldots , v_n) \in \mathrm{GL}(n, \mathbb{C}). $$

Dann gilt:

$$ {}^tS_1AS_1 = \begin{pmatrix} {\scriptstyle c_1} & & & & \\ & \ddots & & 0 & \\ & & {\scriptstyle c_r} & & \\ & & & {\scriptstyle 0} & \\ & 0 & & & \ddots \\ & & & & & {\scriptstyle 0} \end{pmatrix} =: A_1 $$

mit

$$ c_i = \langle v_i, v_i \rangle_A \neq 0 \qquad (i = 1, \ldots , r). $$

Wähle nun $\lambda_i \in \mathbb{C}^*$ mit

$$ \lambda_i^2 = \frac{1}{c_i}. $$

Es sei

$$ S_2 := \begin{pmatrix} {\scriptstyle \lambda_1} & & & & \\ & \ddots & & 0 & \\ & & {\scriptstyle \lambda_r} & & \\ & & & {\scriptstyle 1} & \\ & 0 & & & \ddots \\ & & & & & {\scriptstyle 1} \end{pmatrix}. $$

Dann gilt

$$ {}^tS_2A_1S_2 = \begin{pmatrix} {\scriptstyle 1} & & & & \\ & \ddots & & 0 & \\ & & {\scriptstyle 1} & & \\ & & & {\scriptstyle 0} & \\ & 0 & & & \ddots \\ & & & & & {\scriptstyle 0} \end{pmatrix}. $$

Also gilt

$$^t(S_1 \circ S_2)A(S_1 \circ S_2) = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$

Die Behauptung über den Rang ist klar, da

$$\operatorname{Rang} A = \operatorname{Rang}(^tSAS) \qquad (S \in \mathrm{GL}(n; \mathbb{C})).$$

□

Theorem 13.6
Es sei $Q \subset \mathbb{P}^n(\mathbb{C})$ eine Quadrik. Dann ist Q äquivalent zu genau einer der folgenden Quadriken:

$$Q_r: \ x_0^2 + \ldots + x_r^2 = 0 \qquad (-1 \leq r \leq n).$$

Es ist

$$d(Q_r) = n - r - 1.$$

Insbesondere gilt

$$Q \sim Q' \Leftrightarrow d(Q) = d(Q').$$

Bemerkung.

(i) $Q_{-1} = \mathbb{P}^n(\mathbb{C})$.

(ii) $u(Q_r) = n - 1 - [\frac{r}{2}]$.

Beweis. Es sei

$$Q = \{q_A = 0\}.$$

Dann gibt es $S \in \mathrm{GL}(n+1, \mathbb{C})$ mit

$$^tSAS = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$
$$\underbrace{\qquad\qquad}_{r+1}$$

Die Behauptung über $d(Q)$ folgt, da $d(Q) = d(Q_r)$ und

$$\operatorname{Sing} Q_r = \{x_0 = \ldots = x_r = 0\}.$$

□

Zum Abschluß geben wir noch ein Verfahren an, wie man symmtrische Matrizen mit Hilfe von Elementarmatrizen diagonalisieren kann.

§ 14 Affine Quadriken

Es sei

$$V = K^n \qquad (\operatorname{char} K \neq 2).$$

Definition. Eine *(inhomogene) quadratische Form* in n Variablen ist ein Ausdruck der Form

$$q_0(x) = \sum_{\nu,\mu=1}^{n} a_{\nu\mu} x_\nu x_\mu + \sum_{\nu=1}^{n} a_\nu x_\nu + a.$$

Definition. Eine Teilmenge $Q_0 \subset K^n$ heißt *(affine) Quadrik*, falls es eine quadratische Form q_0 gibt mit

$$Q_0 = \{(x_1, \dots, x_n) \in K^n;\ q_0(x_1, \dots, x_n) = 0\}.$$

Schreibweise. $Q_0 = \{q_0 = 0\}$.

Beispiele. Im $\mathbb{R}^2$:

(i) $x_1^2 + x_2^2 - 1 = 0$ (Kreis).

(ii) $x_1^2 - x_2 = 0$ (Parabel).

(iii) $x_1^2 - x_2^2 - 1 = 0$ (Hyperbel).

(iv) $x_1 x_2 = 0$ (sich schneidendes Geradenpaar).

(v) $x_1(x_1 - 1) = 0$ (parallele Gerade).

(vi) $x_1 = 0, x_1^2 = 0$ (Gerade, bzw. „Doppelgerade“).

(vii) $x_1^2 + x_2^2 = 0$ (Ursprung).

(viii) $x_1^2 + x_2^2 + 1 = 0$ (leere Menge).

(ix) $q \equiv 0$ ($\mathbb{R}^2$).

Sei

$$q_0(x) = \sum_{\nu,\mu=1}^{n} a_{\nu\mu} x_\nu x_\mu + \sum_{\nu=1}^{n} a_\nu x_\nu + a_0.$$

Man kann daraus eine *homogene* quadratische Form machen, indem man setzt:

$$q(x) = \sum_{\nu,\mu=1}^{n} a_{\nu\mu} x_\nu x_\mu + \sum_{\nu=1}^{n} a_\nu x_\nu x_0 + a_0 x_0^2.$$

Setzen wir

$$\begin{array}{rcl} a_{\nu 0} & := & a_{0\nu} := \frac{1}{2} a_\nu \\ a_{00} & := & a_0 \end{array}$$

so ist

$$q(x) = \sum_{\nu,\mu=0}^{n} a_{\nu\mu} x_\nu x_\mu \qquad (A = (a_{\nu\mu});\ A = {}^t A).$$

Wir setzen nun

$$\begin{array}{rcl} Q_0 & := & \{q_0 = 0\} \subset K^n \\ Q & := & \{q = 0\} \subset \mathbb{P}^n(K) = K^n \cup \mathbb{P}^{n-1}(K). \end{array}$$

Dann ist

$$Q_0 = Q \cap K^n \qquad (\text{„}\textit{affiner Teil}\text{“ der Quadrik } Q)$$

Definition. Man nennt Q den *projektiven Abschluß* von Q_0.

Definition. Zwei (affine) Quadriken $Q_0, Q'_0 \subset K^n$ heißen *geometrisch äquivalent*, wenn es eine Affinität $K^n \to K^u$ gibt mit $\varphi(Q_0) = Q'_0$.

Schreibweise. $Q_0 \approx Q'_0$.

Definition. Zwei (projektive) Quadriken $Q, Q' \subset \mathbb{P}^n(K)$ heißen *affin-äquivalent*, falls es eine Projektivität $\varphi \in \mathrm{Aut}(\mathbb{P}^n(K))$ gibt, mit $\varphi(H_\infty) = H_\infty$ und $\varphi(Q) = Q'$.

Schreibweise. $Q \sim_a Q'$.

Bemerkung.

(i) Aus $Q \sim_a Q'$ folgt natürlich $Q \sim Q'$.

(ii) Ist $Q \sim_a Q'$ und $Q_0 := Q \cap K^n$, $Q'_0 := Q' \cap K^n$, dann gilt auch $Q_0 \approx Q'_0$.

Es sei nun $Q \subset \mathbb{P}^n(K)$ eine projektive Quadrik:

$$Q = \{(x_0 : \ldots : x_n) \in \mathbb{P}^n(K);\ \sum_{\nu,\mu=0}^{n} a_{\nu\mu} x_\nu x_\mu = 0\}.$$

Dann ist

$$Q^* := Q \cap H_\infty$$

gegeben durch

$$Q^* = \{(x_1 : \ldots : x_n) \in \mathbb{P}^{n-1}(K) = H_\infty;\ \sum_{\nu,\mu=1}^{n} a_{\nu\mu} x_\nu x_\mu = 0\}.$$

Wir definieren nun neue Invarianten (bezüglich der Relation $\sim_a$) für projektive Quadriken.

Definition.

$$d^*(Q) := d(Q^*) \qquad u^*(Q) := u(Q^*).$$

Bemerkung.

$$Q \sim_a Q' \Rightarrow d(Q) = d(Q'),\ u(Q) = u(Q'),\ d^*(Q) = d^*(Q'),\ u^*(Q) = u^*(Q')$$

Theorem 14.1
Jede Quadrik $Q \subset \mathbb{P}^n(\mathbb{R})$ ist affin-äquivalent zu genau einer der folgenden Quadriken:

Typ	Gleichung	Bedingungen	d	u	d^*	u^*
(1)	$x_1^2 + \ldots + x_t^2$	$0 \le t \le r \le n$	$n-r$	$n-t$	$n-r$	$n-t-1$
	$-x_{t+1}^2 - \ldots - x_r^2 = 0$	$r-t \le t$			-1	
(2)	$x_1^2 + \ldots + x_t^2$	$0 \le t \le r \le n$	$n-r$	$n-t$	$n-r$	$n-t-1$
	$-x_{t+1}^2 - \ldots - x_r^2 = x_0^2$	$r-t < t$	-1		-1	
(3)	$x_1^2 + \ldots + x_t^2$	$0 \le t \le r \le n$	$n-r$	$n-(r-t)$	$n-r$	$n-(r-t)$
	$-x_{t+1}^2 - \ldots - x_r^2 = x_0^2$	$r-t \ge t$	-1	-1	-1	-1
(4)	$x_1^2 + \ldots + x_t^2$	$0 \le t \le r \le n-1$	$n-r$	$n-t-1$	$n-r$	$n-t-1$
	$-x_{t+1}^2 - \ldots - x_r^2 = x_0 x_{r+1}$	$r-t \le t$	-2		-1	

Zwei Quadriken $Q, Q' \subset \mathbb{P}^n(\mathbb{R})$ sind also genau dann affin-äquivalent, wenn d, d^*, u, u^* übereinstimmen.

Bemerkung.

(i) Mann kann die Gleichungen auch in der Form

$$x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = \begin{cases} 0 \\ x_0^2 \\ -x_0^2 \\ x_0 x_{r+1} \end{cases}$$

mit $r - t \le t$ schreiben. Dann ändern sich allerdings die Invarianten d, u, d^*, u^* im Fall (3) entsprechend.

(ii) Jede Quadrik $Q_0 \subset \mathbb{R}^n$ ist geometrisch äquivalent zum affinen Teil einer der oben beschriebenen Quadriken.

(iii) Es kann sein, daß $Q \not\sim_a Q'$ aber $Q_0 \approx Q'_0$. Etwa

$$Q = \{x_0 x_1 = 0\}, \qquad Q' = \{x_1^2 = 0\}.$$

Es gilt jedoch

$$Q_0 = \{x_1 = 0\} = \{x_1^2 = 0\} = Q'_0.$$

Wir werden hierauf noch zurückkommen.

Beweis. Wir stellen zunächst fest, daß die angegebenen Quadriken alle verschieden sind. Es gilt

(1) $d^* = d - 1, u^* = u - 1$

(2) $d^* = d, u^* = u - 1$

(3) $d^* = d, u^* = u$

(4) $d^* = d + 1, u^* = u.$

Also sind Quadriken in verschiedenen Typen auf keinen Fall äquivalent. Innerhalb eines Typs bestimmen d, u (bzw. d^*, u^*) schon r, t und umgekehrt.

Bleibt zu zeigen, daß man jede Quadrik auf eine dieser Normalformen bringen kann. Es sei

$$Q: \quad \sum_{\mu,\nu=0}^{n} a_{\mu\nu} x_\mu x_\nu = 0, \qquad A = (a_{\mu\nu});\ A = {}^tA.$$

Es ist

$$Q^* = Q \cap \{x_0 = 0\}.$$

Also ist

$$Q^*: \quad \sum_{\mu,\nu=1}^{n} a_{\mu\nu} x_\mu x_\nu = 0.$$

Es gibt nun nach Theorem (13.3) ein

$$\varphi \in \operatorname{Aut}(H_\infty) = \operatorname{Aut}(\mathbb{P}^{n-1}(\mathbb{R}))$$

mit

$$\varphi(Q^*) = (Q^*)'$$

und

$$(Q^*)': \quad x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = 0 \quad (0 \le t \le r \le n, t \ge r - t).$$

Sei

$$\varphi = \overline{S}$$

mit $S \in \mathrm{GL}(n, \mathbb{R})$. Setze

$$T := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & S & \\ 0 & & & \end{pmatrix} \in \mathrm{GL}(n+1, \mathbb{R}).$$

Sei

$$\psi := \overline{T} \in \mathrm{Aut}(\mathbb{P}^n(\mathbb{R})).$$

Dann ist

$$\psi|_{H_\infty} = \varphi$$

Es sei

$$Q' := \psi(Q).$$

Man hat

$$Q' : \quad x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = bx_0^2 + \sum_{\nu=1}^{n} b_\nu x_0 x_\nu.$$

Wir definieren nun $\tau \in \mathrm{Aut}(\mathbb{P}^n(\mathbb{R}))$ durch

$$\begin{aligned} x_0 &\longmapsto x_0 \\ x_\mu &\longmapsto x_\mu + \frac{b_\mu}{2} x_0 \qquad (1 \le \mu \le t) \\ x_\varrho &\longmapsto x_\varrho - \frac{b_\varrho}{2} x_0 \qquad (t+1 \le \varrho \le r) \\ x_\nu &\longmapsto x_\nu \qquad (r+1 \le \nu \le n). \end{aligned}$$

Dann wird Q' übergeführt in

$$Q'' : \quad x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = cx_0^2 + \sum_{\nu=r+1}^{n} b_\nu x_0 x_\nu.$$

Denn:

$$\begin{aligned} \text{LS:} \quad & (x_\mu + \frac{b_\mu}{2} x_0)^2 = x_\mu^2 + b_\mu x_\mu x_0 + \frac{1}{4} b_\mu^2 x_0^2 \\ \text{RS:} \quad & b_\mu x_0 (x_\mu + \frac{b_\mu}{2} x_0) = b_\mu x_\mu x_0 + \frac{1}{2} b_\mu^2 x_0^2 \end{aligned}$$

bzw.

$$\begin{aligned} \text{LS:} \quad & -(x_\varrho - \frac{b_\varrho}{2} x_0)^2 = -x_\varrho^2 + b_\varrho x_\varrho x_0 - \frac{1}{4} b_\varrho^2 x_0^2 \\ \text{RS:} \quad & b_\varrho x_0 (x_\varrho - \frac{b_\varrho}{2} x_0) = b_\varrho x_0 x_\varrho - \frac{1}{2} b_\varrho^2 x_0^2. \end{aligned}$$

1. Fall: $c = 0$ und $b_\nu = 0$ für $\nu = r+1, \ldots, n$. Dann hat Q'' eine Normalform vom Typ (1). Q'' und $(Q'')^*$ haben dieselbe Gleichung. Nach Theorem (13.3) gilt (beachte daß die Summation bei 1 beginnt):

$$\begin{aligned} u &= n - t, & d &= n - r \\ u^* &= n - t - 1, & d^* &= n - r - 1. \end{aligned}$$

2. Fall: $c > 0$ und $b_\nu = 0$ für $\nu = r+1, \dots, n$. Wir führen folgende Transformation durch:

$$x_0 \longmapsto \frac{1}{\sqrt{c}} x_0, \quad x_\nu \longmapsto x_\nu \qquad (\nu = 1, \dots, n).$$

Damit wird Q'' übergeführt in:

$$Q''' : \quad x_1^2 + \dots + x_t^2 - x_{t+1}^2 - \dots - x_r^2 = x_0^2 \qquad (0 \le t \le r \le n,\, t \ge r - t).$$

(a) $t > r - t$: Dann sind wir im Fall (2). Die Quadrik

$$\underbrace{x_1^2 + \dots + x_t^2}_{t} - \underbrace{x_{t+1}^2 - \dots - x_r^2 - x_0^2}_{r+1-t} = 0$$

hat

$$t \ge r + 1 - t.$$

Dies ist eine Normalform im Sinne von Theorem (13.3). Damit gilt

$$u = n - t, \quad d = n - r - 1.$$

$(Q''')^*$ hat dieselbe Gleichung wie im 1. Fall. Also gilt

$$u^* = n - t - 1, \quad d^* = n - r - 1.$$

(b) $r - t = t$. Dann sind wir im Typ (3) (mit $t = r - t$). Es ist

$$\underbrace{x_0^2 + x_{t+1}^2 + \dots + x_r^2}_{r+1-t} - \underbrace{x_1^2 - \dots - x_t^2}_{t} = 0$$

ebenfalls eine Normalform im Sinne von Theorem (13.3). Es folgt

$$u = n - (r + 1 - t) = n - (r - t) - 1; \quad d = n - r - 1.$$

Im Unendlichen liegt dieselbe Gleichung wie in (a) vor, also gilt:

$$u^* = n - t - 1 \overset{t=r-t}{=} n - (r - t) - 1; \quad d^* = n - r - 1.$$

3. Fall: $c < 0$ und $b_\nu = 0$ für $\nu = r+1, \dots, n$. Wir transformieren

$$\begin{aligned} x_0 &\longmapsto \frac{1}{\sqrt{|c|}} x_0 \\ x_\nu &\longmapsto x_\nu \quad (\nu = 1, \dots, n). \end{aligned}$$

Damit erhalten wir:

$$Q''' : \quad x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = -x_0^2 \quad (0 \le t \le r \le n, t \ge r-t).$$

Bzw.

$$Q''' : \quad \underbrace{x_{t+1}^2 + \ldots + x_r^2}_{r-t} - \underbrace{x_1^2 - \ldots - x_t^2}_{t} = x_0^2.$$

Nach Umbenennen der Variablen:

$$\underbrace{x_1^2 + \ldots + x_{r-t}^2}_{r-t} - \underbrace{x_{r-t+1}^2 - \ldots - x_r^2}_{t} = x_0^2.$$

Mit

$$t' := r - t$$

hat man

$$\underbrace{x_1^2 + \ldots + x_{t'}^2}_{t'} - \underbrace{x_{t'+1}^2 - \ldots - x_r^2}_{r-t'} = x_0^2.$$

Es ist nun

$$0 \le t' \le r \le n; \quad t' \le r - t'.$$

Also sind wir im Typ (3). Die Berechnung von u, d, u^*, d^* erfolgt genau wie zuvor.

4. Fall: Es gibt $b_\nu \neq 0$ mit $\nu \in \{r+1, \ldots, n\}$. Damit ist $r \le n-1$. Durch eventuelles Vertauschen der Variablen $x_{r+1}, \ldots, x_n$ kann man erreichen, daß $b_{r+1} \neq 0$. Also erhalten wir

$$Q'' : \quad x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = cx_0^2 + b_{r+1}x_0x_{r+1} + b_{r+2}x_0x_{r+2} + \ldots + b_nx_0x_n.$$

Durch die Transformation

$$\begin{aligned} x_\nu &\longmapsto x_\nu \quad (\nu \neq r+1) \\ x_{r+1} &\longmapsto \frac{1}{b_{r+1}}(x_{r+1} - cx_0 - b_{r+2}x_{r+2} - \ldots - b_nx_n) \end{aligned}$$

wird Q'' übergeführt in

$$Q''' : \quad x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = x_0x_{r+1} \qquad (0 \le t \le r \le n-1,\ r-t \le t).$$

Dies ist Typ (4). Für $x_0 = 0$ stimmt dies mit (1) überein. Dies ergibt d^*, u^*. Um d, u zu bekommen, betrachten wir:

$$\begin{aligned} x_0 &\longmapsto (x_{r+1} + x_0) \\ x_{r+1} &\longmapsto (x_{r+1} - x_0) \\ x_\nu &\longmapsto x_\nu \qquad (\nu \neq 0, r+1). \end{aligned}$$

Dies ist keine affine Projektivität, erhält aber u und d. Wir erhalten dann

$$x_1^2 + \ldots + x_t^2 - x_{t+1}^2 - \ldots - x_r^2 = x_{r+1}^2 - x_0^2,$$

also

$$\underbrace{x_0^2 + \ldots + x_t^2}_{t+1} - \underbrace{x_{t+1}^2 - \ldots - x_{r+1}^2}_{r+1-t} = 0$$

mit $t+1 \geq r+1-t$. Dies ist eine Normalform im Sinne von Theorem (13.3). Daraus folgt

$$u = n - t - 1, \quad d = n - (r+1) - 1 = n - r - 2.$$

□

Beispiel.

$$Q_0 : \quad x_1 + x_1^2 - 2x_2x_3 = 1 \qquad (Q_0 \subset \mathbb{R}^3).$$

Der projektive Abschluß ist dann

$$Q : \quad x_0x_1 + x_1^2 - 2x_2x_3 = x_0^2.$$

Zunächst betrachten wir

$$Q^* = Q \cap H_\infty \qquad (H_\infty = \{x_0 = 0\}).$$

Also

$$Q^* : \quad x_1^2 - 2x_2x_3 = 0.$$

Die darstellende Matrix ist

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

Wir bestimmen nun die Normalform von A und berechnen hierzu zunächst die Eigenwerte von A.

$$\begin{aligned} \chi_A(x) &= \det(xE - A) = \det \begin{pmatrix} x-1 & 0 & 0 \\ 0 & x & 1 \\ 0 & 1 & x \end{pmatrix} \\ &= (x-1)(x^2-1) \\ &= (x-1)^2(x+1). \end{aligned}$$

Also erhalten wir

$$\lambda_{1,2} = 1, \quad \lambda_3 = -1.$$

Damit gilt

$$Q^* \sim (Q^*)' : \qquad x_1^2 + x_2^2 - x_3^2 = 0.$$

Konkret kann man dies so machen:

$$x_1 \longmapsto x_1, \quad x_2 \longmapsto \frac{x_3 - x_2}{\sqrt{2}}, \quad x_3 \longmapsto \frac{x_3 + x_2}{\sqrt{2}}.$$

Damit wird Q übergeführt in

$$Q' : \quad x_0x_1 + x_1^2 + x_2^2 - x_3^2 = x_0^2.$$

Wie im Beweis von Theorem (14.1) kann man dies so transformieren:

$$\begin{array}{ll} x_0 \longmapsto x_0, & x_1 \longmapsto x_1 - \dfrac{x_0}{2} \\ x_2 \longmapsto x_2, & x_3 \longmapsto x_3. \end{array}$$

Damit geht Q' über in

$$x_0(x_1 - \frac{x_0}{2}) + (x_1 - \frac{x_0}{2})^2 + x_2^2 - x_3^2 = x_0^2,$$

d. h.

$$x_0x_1 - \frac{x_0^2}{2} + x_1^2 - x_0x_1 + \frac{x_0^2}{4} + x_2^2 - x_3^2 = x_0^2,$$

bzw.

$$x_1^2 + x_2^2 - x_3^2 = \frac{5}{4}x_0^2.$$

Schließlich ergibt

$$x_0 \longmapsto \frac{2}{\sqrt{5}}x_0, \qquad x_\nu \longmapsto x_\nu \quad (\nu = 1, 2, 3)$$

die Form

$$x_1^2 + x_2^2 - x_3^2 = x_0^2.$$

Die affine Gleichung ist dann

$$Q_0 : \quad x_1^2 + x_2^2 - x_1^3 = 1.$$

Es handelt sich um ein einschaliges Hyperboloid. Dies ist eine Quadrik vom Typ (2). Es gilt

$$d = -1, \quad u = 1, \quad d^* = -1, \quad u^* = 0.$$

Wir kehren nun zur Klassifikation von Quadriken $Q_0 \subset \mathbb{R}^n$ zurück.

Theorem 14.2
Jede affine Quadrik $Q_0 \subset \mathbb{R}^n$ ist (geometrisch) äquivalent zu einer der affinen Quadriken in Theorem (14.1). Zwei Quadriken in dieser Liste sind genau dann (geometrisch) äquivalent, falls sie beide Hyperebenen oder die leere Menge sind.

Bemerkung. Die folgenden Gleichungen führen zu Hyperebenen oder der leeren Menge:

$$\begin{array}{llll} \text{Hyperebenen:} & \text{(i) Typ (1)} & r=1,\, t=1: & x_1^2 = 0. \\ & \text{(ii) Typ (4)} & t=r=0: & x_0 x_1 = 0. \\ \text{Leere Menge:} & \text{Typ (3)} & t=0,\, 0 \le r \le n: & -x_1^2 - \ldots - x_r^2 = x_0^2. \end{array}$$

Beweisskizze Wir haben schon gesehen, daß jede affine Quadrik in obiger Liste vorkommt.

Schritt 1: $\mathbb{P}^n(\mathbb{R})$ ist ein kompakter topologischer Raum. Wir betrachten dazu die Projektion

$$\pi: \quad \mathbb{R}^{n+1} - \{0\} \longrightarrow \mathbb{P}^n(\mathbb{R}).$$

Definition. Eine Menge $U \subset \mathbb{P}^n(\mathbb{R})$ heißt *offen*, falls $\pi^{-1}(U) \subset \mathbb{R}^{n+1} - \{0\}$ offen ist.

Bemerkungen.

(i) Damit ist die Abbildung $\pi : \mathbb{R}^{n+1} - \{0\} \to \mathbb{P}^n(\mathbb{R})$ eine stetige Abbildung.

(ii) Diese Topologie ist so gemacht, daß die Abbildung

$$\begin{array}{rcl} \mathbb{P}^n(\mathbb{R}) \setminus H_\infty & \longrightarrow & \mathbb{R}^n \\ (x_0 : x_1 : \ldots : x_n) & \longmapsto & (\frac{x_1}{x_0}, \frac{x_2}{x_0}, \ldots, \frac{x_n}{x_0}) \end{array}$$

ein Homöomorphismus ist.

(iii) Der topologische Raum $\mathbb{P}^n(\mathbb{R})$ ist deswegen kompakt, weil er das Bild der kompakten Sphäre S^n unter der stetigen Abbildung π ist.

Wir betrachten nun eine affine Quadrik

$$Q_0 = \{q_0 = 0\} \subset \mathbb{R}^n$$

und ihren projektiven Abschluß

$$Q = \{q = 0\} \subset \mathbb{P}^n(\mathbb{R}).$$

Es ist zu beachten, daß Q a priori von der Wahl der Gleichung $q_0 = 0$ von Q_0 abhängt.

Schritt 2: Q ist abgeschlossen in $\mathbb{P}^n(\mathbb{R})$.

Beweis. Es ist zu zeigen, daß

$$\pi^{-1}(Q) = \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} - \{0\};\ q(x_0, \dots, x_n) = 0\}$$

abgeschlossen ist. Dies gilt, da

$$\pi^{-1}(Q) = q^{-1}(0)$$

mit

$$\begin{array}{rccc} q: & \mathbb{R}^{n+1} - \{0\} & \longrightarrow & \mathbb{R} \\ & (x_0, \dots, x_n) & \longmapsto & q(x_0, \dots, x_n). \end{array}$$

D. h. $\pi^{-1}(Q)$ ist das Urbild der abgeschlossenen Menge $\{0\}$ unter der stetigen Abbildung q, also abgeschlossen. Damit hat man in $\mathbb{P}^n(\mathbb{R})$:

$$Q_0 \subset \overline{Q}_0 \subset Q$$

wobei $\overline{Q}_0$ der topologische Abschluß von Q_0 in $\mathbb{P}^n(\mathbb{R})$ ist.

Schritt 3: Ist Q_0 weder leer, noch eine Hyperebene, so gilt $\overline{Q}_0 = Q$. Insbesondere hängt in diesem Fall der projektive Abschluß Q nicht von der gewählten Gleichung q_0 ab.

Beweis. G. Fischer, Analytische Geometrie, p. 66ff.

Schritt 4: Man kann den Beweis nun wie folgt führen. Es seien $Q_0 = \{q_0 = 0\}$ und $Q'_0 = \{q'_0 = 0\}$ Quadriken dieser Liste. Q_0 und Q'_0 seien affin-äquivalent. Es sei $\varphi_0 : \mathbb{R}^n \to \mathbb{R}^n$ also eine Affinität mit $\varphi_0(Q_0) = Q'_0$. Man setze φ_0 fort zu $\varphi : \mathbb{P}^n(\mathbb{R}) \to \mathbb{P}^n(\mathbb{R})$. Dann ist φ ein Homöomorphismus, und es gilt $\varphi(\overline{Q}_0) = \overline{Q}'_0$. Nach Voraussetzung ist Q_0 weder eine Hyperebene noch leer (und damit gilt dies auch für Q'_0). Nach Schritt 3 ist $\overline{Q}_0 = Q$, $\overline{Q}'_0 = Q'$. Also gilt $\varphi(Q) = Q'$. Damit ist $Q = Q'$ nach Theorem (14.1), also $Q_0 = Q'_0$. $\square$

Es sei nun $Q \subset \mathbb{P}^n(\mathbb{R})$ eine Quadrik. Wir fragen, inwieweit die Menge Q eine definierende Gleichung bestimmt.

Beispiel. Die Gleichungen

$$\lambda_0 x_0^2 + \ldots + \lambda_r x_r^2 = 0 \qquad (\lambda_0, \dots, \lambda_r > 0)$$

stellen in $\mathbb{P}^n(\mathbb{R})$ alle dieselbe Teilmenge, nämlich

$$x_0 = x_1 = \ldots = x_r = 0$$

dar. Es wird also nicht in jedem Fall die Menge Q eine (bis auf einen Skalar) eindeutig bestimmte Gleichung q mit $Q = \{q = 0\}$ liefern.

Definition. Wir betrachten nun quadratische Formen

$$\begin{aligned} q_A(x) &:= \langle x, x \rangle_A = {}^t x A x \\ q_{A'}(x) &:= \langle x, x \rangle_{A'} = {}^t x A' x \end{aligned}$$

mit $A = {}^tA$, $A' = {}^tA'$. Für die Form

$$s_A(x,y) := \langle x,y\rangle_A = {}^t xAy$$

hatten wir bereits früher den *Ausartungsraum*

$$V_0^A := \{v \in \mathbb{R}^{n+1};\ \langle v,w\rangle_A = 0 \text{ für alle } w \in \mathbb{R}^{n+1}\} = \operatorname{Ker} A$$

betrachtet. Es gilt natürlich

$$\operatorname{Sing} Q = \mathbb{P}(V_0^A) \subset \{q_A = 0\} = Q.$$

Satz 14.3
Es sei $Q = \{q_A = 0\} = \{q_{A'} = 0\}$. Ferner sei $\operatorname{Sing} Q \subsetneqq Q$. Dann gibt es ein $c \neq 0$ mit $A = cA'$.

Beweis. Es sei

$$C := \pi^{-1}(Q) \cup \{0\}$$

(wobei $\pi : \mathbb{R}^{n+1} \setminus \{0\} \to \mathbb{P}^n(\mathbb{R})$ wieder die Projektion ist) der sogenannte Kegel über Q. D. h.

$$C = \{(x_0,\dots,x_n) \in \mathbb{R}^{n+1};\ q_A(x_0,\dots,x_n) = 0\}.$$

Nach Voraussetzung gibt es einen Vektor

$$v_0 \in C \setminus V_0^A.$$

Wir betrachten nun die Geraden

$$g_w := \{w + \lambda v_0;\ \lambda \in \mathbb{R}\} \qquad (w \in \mathbb{R}^{n+1}).$$

Es gilt

$$g_w \cap C = \{w + \lambda v_0;\ q_A(w + \lambda v_0) = 0\} = \{w + \lambda v;\ q_{A'}(w + \lambda v) = 0\}.$$

Dies führt wegen $q_A(v_0,v_0) = 0 = q_{A'}(v_0,v_0)$ zu:

$$(1) \qquad \begin{aligned} 2\lambda s_A(v_0,w) + q_A(w) &= 0 \\ 2\lambda s_{A'}(v_0,w) + q_{A'}(w) &= 0. \end{aligned}$$

Behauptung.

$$(*) \qquad s_A(v_0,w) = 0 \Leftrightarrow s_{A'}(v_0,w) = 0.$$

Denn: $w \in C$: Dann ist $q_A(w) = q_{A'}(w) = 0$. Es gilt

$$s_A(v_0,w) = 0 \Leftrightarrow g_w \subset C \Leftrightarrow s_{A'}(v_0,w) = 0.$$

$w \notin C$: Dann ist $q_A(w) \neq 0 \neq q_{A'}(w)$. Es gilt

$$s_A(v_0, w) = 0 \Leftrightarrow g_w \cap C = \emptyset \Leftrightarrow s_{A'}(v, w) = 0.$$

Also gilt

$$H := \{w \in \mathbb{R}^{n+1};\ s_A(v_0, w) = 0\} = \{w \in \mathbb{R}^{n+1};\ s_{A'}(v_0, w) = 0\}.$$

Nach Voraussetzung ($v_0 \notin V_0^A$) ist H eine Hyperebene $H \subsetneqq \mathbb{R}^{n+1}$. Eine Hyperebene bestimmt ihre Gleichung bis auf einen Skalar. Also gibt es ein $c \neq 0$ mit:

(2)
$$s_A(v_0, w) = c s_{A'}(v_0, w) \qquad (w \in \mathbb{R}^{n+1}).$$

Da die beiden Gleichungen in (1) äquivalent sind, gilt (man betrachte die Determinante)

(3)
$$s_A(v_0, w) q_{A'}(w) = s_{A'}(v_0, w) q_A(w).$$

Mit (2) ergibt sich

$$c q_{A'}(w) = q_A(w) \qquad (w \in \mathbb{R}^{n+1} \setminus H).$$

Also ist

$$\begin{array}{rcl} \mathbb{R}^{n+1} & \longrightarrow & \mathbb{R} \\ w & \longmapsto & q_A(w) - c q_{A'}(w) \end{array}$$

stetig und verschwindet auf $\mathbb{R}^{n+1} \setminus H$, also auf ganz $\mathbb{R}^{n+1}$. D. h.

$$q_A(w) = c q_{A'}(w) \qquad (w \in \mathbb{R}^{n+1}).$$

Wegen

$$\begin{aligned} s_A(v, w) &= \frac{1}{2}(q_A(v + w, v + w) - q_A(v) - q_A(w)) \\ s_{A'}(v, w) &= \frac{1}{2}(q_{A'}(v + w, v + w) - q_{A'}(v) - q_{A'}(w)) \end{aligned}$$

liefert dies

$$s_A(v, w) = c s_{A'}(v, w) \qquad (v, w \in \mathbb{R}^{n+1}).$$

Also

$$A = cA'.$$

□

Bemerkung. Dieser Beweis funktioniert auch für $K = \mathbb{C}$. Die Aussage selbst gilt für alle Körper K der Charakteristik $\neq 2$.

Die beiden folgenden Tabellen geben die Klassifikation der affinen Kegelschnitte (n=2) und der affinen Quadriken im Raum ($n = 3$).

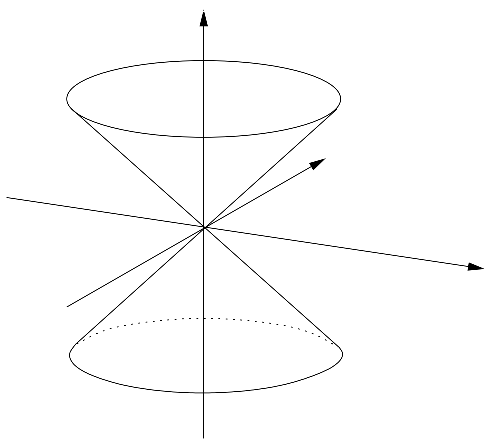
Nr.	Typ	t	r	d	u	d^*	u^*	Gleichung (Normalform)	Bezeichnung (affin)	uneigentlicher Teil
1	(1)	0	0	2	2	1	1	$0 = 0$	Ebene	Gerade
2		1	1	1	1	0	0	$x_1^2 = 0$	eigentl. Gerade	Punkt
3		1	2	0	1	−1	0	$x_1^2 - x_2^2 = 0$	Geradenpaar mit eigentl. Schnittpunkt	Punktepaar
4		2	2	0	0	−1	−1	$x_1^2 + x_2^2 = 0$	eigentl. Punkt	$\emptyset$
5	(2)	1	1	0	1	0	0	$x_1^2 = x_0^2$	Paar paralleler Geraden	Punkt
6		2	2	−1	0	−1	−1	$x_1^2 + x_2^2 = x_0^2$	Ellipse	$\emptyset$
7	(3)	0	0	1	1	1	1	$0 = x_0^2$	uneigentl. Gerade	Gerade
8		0	1	0	0	0	0	$-x_1^2 = x_0^2$	uneigentl. Punkt	Punkt
9		0	2	−1	−1	−1	−1	$-x_1^2 - x_2^2 = x_0^2$	$\emptyset$	$\emptyset$
10		1	2	−1	0	−1	0	$x_1^2 - x_2^2 = x_0^2$	Hyperbel	Punktepaar
11	(4)	0	0	0	1	1	1	$0 = x_0 x_1$	eine eigentl. u. die uneigentl. Gerade	Gerade
12		1	1	−1	0	0	0	$x_1^2 = x_0 x_2$	Parabel	Punkt

Tabelle 1: Affine Klassifikation reeller Quadriken ($n = 2$). Im Affinen fallen die Fälle 2 und 11 (Gerade) und 7,8,9 (leere Menge) zusammen.

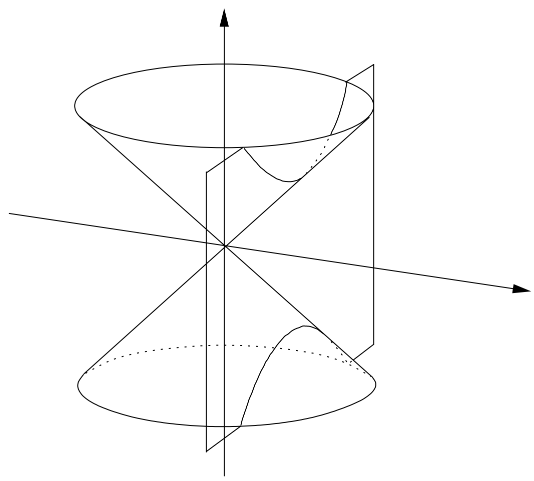
Nr.	Typ	t	r	d	u	d^*	u^*	Gleichung (Normalform)	Bezeichnung (affin)	uneigentlicher Teil
1	(1)	0	0	3	3	2	2	$0 = 0$	3-dim. Raum	Ebene
2		1	1	2	2	1	1	$x_1^2 = 0$	eigentl. Ebene	Gerade
3		1	2	1	2	0	1	$x_1^2 - x_2^2 = 0$	Ebenenpaar mit eigentl. Schnittger.	Geradenpaar
4		2	2	1	1	0	0	$x_1^2 + x_2^2 = 0$	eigentl. Gerade	Punkt
5		2	3	0	1	−1	0	$x_1^2 + x_2^2 - x_3^2 = 0$	Kegel	nicht-ausg. Kurve
6		3	3	0	0	−1	−1	$x_1^2 + x_2^2 + x_3^2 = 0$	eigentl. Punkt	$\emptyset$
7	(2)	1	1	1	2	1	1	$x_1^2 = x_0^2$	Paar paralleler Ebenen	Gerade
8		2	2	0	1	0	0	$x_1^2 + x_2^2 = x_0^2$	elliptischer Zylinder	Punkt
9		2	3	−1	1	−1	0	$x_1^2 + x_2^2 - x_3^2 = x_0^2$	einschaliger Hyperboloid	nicht-ausg. Kurve
10		3	3	−1	0	−1	−1	$x_1^2 + x_2^2 + x_3^2 = x_0^2$	Ellipsoid	$\emptyset$
11	(3)	0	0	2	2	2	2	$0 = x_0^2$	uneigentl. Ebene	Ebene
12		0	1	1	1	1	1	$-x_1^2 = x_0^2$	uneigentl. Gerade	Gerade
13		0	2	0	0	0	0	$-x_1^2 - x_2^2 = x_0^2$	uneigentl. Punkt	Punkt
14		0	3	−1	−1	−1	−1	$-x_1^2 - x_2^2 - x_3^3 = x_0^2$	$\emptyset$	$\emptyset$
15		1	2	0	1	0	1	$x_1^2 - x_2^2 = x_0^2$	hyperbolischer Zylinder	Geradenpaar
16		1	3	−1	0	−1	0	$x_1^2 - x_2^2 - x_3^3 = x_0^2$	zweischal. Hyperboloid	nicht-ausg. Kurve
17	(4)	0	0	1	2	2	2	$0 = x_0 x_1$	eigentl. Ebene und die uneigentl. Ebene	Ebene
18		1	1	0	1	1	1	$x_1^2 = x_0 x_2$	parabol. Zylinder	Gerade
19		1	2	−1	1	0	1	$x_1^2 - x_2^2 = x_0 x_3$	hyperbolischer Paraboloid	Geradenpaar
20		2	2	−1	0	0	0	$x_1^2 + x_2^2 = x_0 x_3$	ellipt. Paraboloid	Punkt

Tabelle 2: Affine Klassifikation reeller Quadriken ($n = 3$). Im Affinen fallen die Fälle 2 und 17 (Ebene), sowie 11,12,13,14 (leere Menge) zusammen.

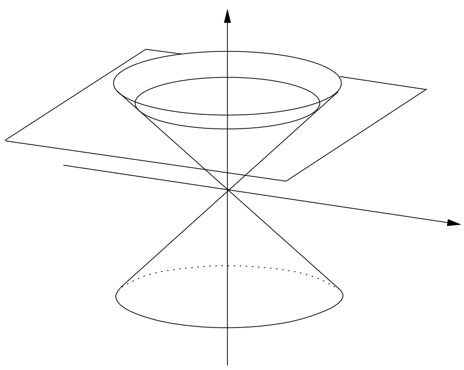
Kegelschnitte



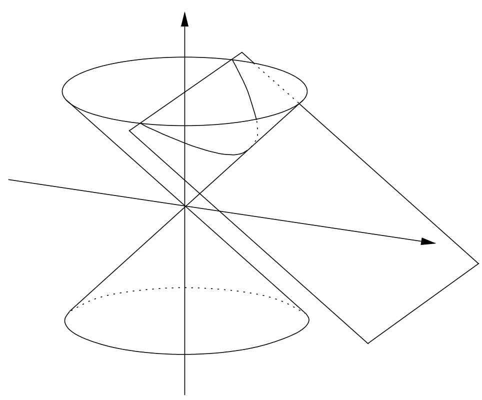
Kegel: $x_0^2 - x_1^2 - x_2^2 = 0.$



Hyperbel

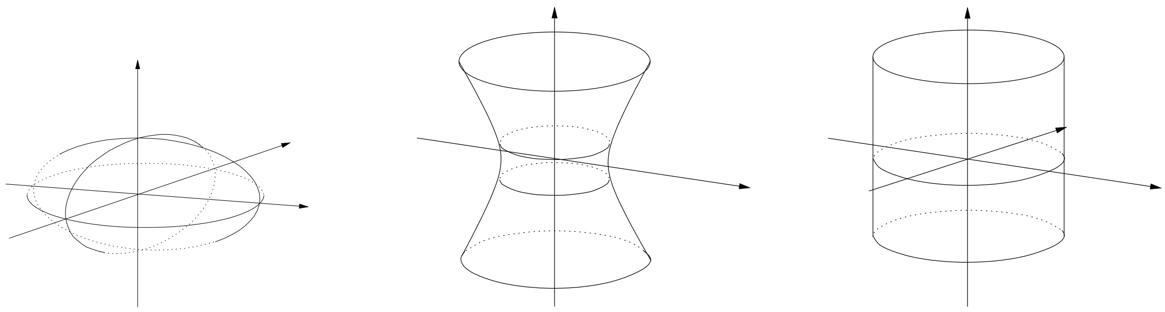


Ellipse



Parabel

Abb. 13: Der quadratische Kegel (Typ (1))



Ellipsoid ($u = 0$)

Einschaliges Hyperboloid ($u = 1, d = -1$)

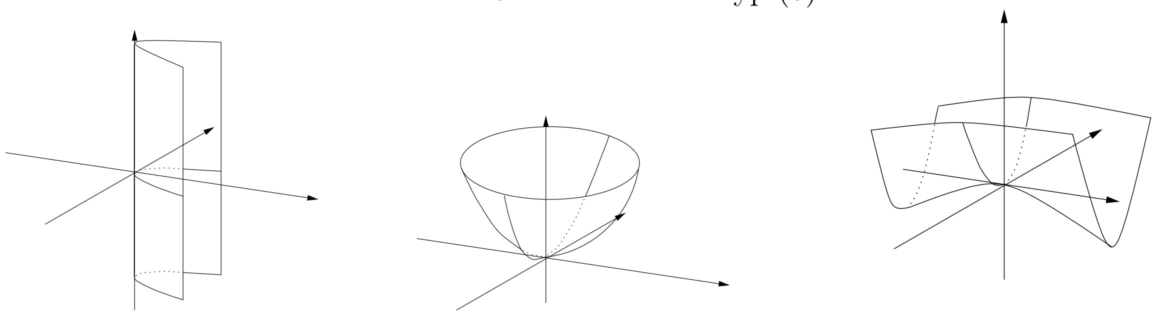
Elliptischer Zylinder ($u = 1, d = d^* = 0$)

Abb. 14: Flächen vom Typ (2)

Zweischaliges Hyperboloid ($u = 0$)

Hyperbolischer Zylinder ($u = 1$)

Abb. 15: Flächen vom Typ (3)



Parabolischer Zylinder ($u = 1, d = 0, d^* = 1$)

Elliptisches Paraboloid ($u = 0, d = -1, d^* = 0$)

Hyperbolischer Paraboloid ($u = 1, d = -1, d^* = 0$)

Abb. 16: Flächen vom Typ (4)

§ 15 Das Tensorprodukt

U, V, W seien K-Vektorräume. Wir definieren zunächst:

Definition. Eine Abbildung $F : U \times V \to W$ heißt *bilinear*, falls gilt

(B1) $$F(\lambda u_1 + \mu u_2, v_1) = \lambda F(u_1, v_1) + \mu F(u_2, v_1)$$

(B2) $$F(u_1, \lambda v_1 + \mu v_2) = \lambda F(u_1, v_1) + \mu F(u_1, v_2)$$

(für $\lambda, \mu \in K$; $u_1, u_2 \in U$; $v_1, v_2 \in V$).

Beispiele.

(1) Jedes *Skalarprodukt* auf einem reellen Vektorraum V:

$$\langle\ ,\ \rangle : V \times V \longrightarrow \mathbb{R}, \quad (u,v) \longmapsto \langle u, v\rangle$$

erfüllt dies, insbesondere das Standardskalarprodukt

$$\langle\ ,\ \rangle : \quad \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad (u,v) \longmapsto {}^t uv = \sum_{i=1}^{n} u_i v_i.$$

(2) $U = V = W = \mathrm{Mat}(n; K)$. Der Kommutator (auch Lie-Produkt genannt):

$$\begin{array}{rrcl} [\quad] : & \mathrm{Mat}(n;K) \times \mathrm{Mat}(n;K) & \longrightarrow & \mathrm{Mat}(n;K) \\ & (A,B) & \longmapsto & [A,B] = AB - BA \end{array}$$

ist bilinear.

Die *Idee* des Tensorprodukts besteht darin, daß man einen Raum konstruiert, so daß *jede* bilineare Abbildung F über das „Tensorprodukt" faktorisiert:

$$\begin{tikzcd} U \times V \arrow[r, "F"] \arrow[d, "\otimes"'] & W \\ U \otimes V \arrow[ur, dotted, "\tilde{F}"'] & \end{tikzcd}$$

Dabei ist $\otimes : U \times V \to U \otimes V$ eine *feste* bilineare Abbildung, die Abbildung $\tilde{F}$ ist linear. Es ist also das Paar $(U \otimes V, \otimes)$ mit dieser „universellen" Eigenschaft zu konstruieren.

Definition. Es sei T ein Vektorraum und

$$\otimes : \quad U \times V \longrightarrow T$$

eine bilineare Abbildung. Man sagt, daß das Paar $(T, \otimes)$ ein *Tensorprodukt* von U und V ist, falls gilt:

(T1) Das Bild von $\otimes$ erzeugt T.

(T2) Ist W ein weiterer Vektorraum und ist $F : U \times V \longrightarrow W$ eine bilineare Abbildung, so gibt es eine lineare Abbildung $\tilde{F} :\ T \to W$, so daß das Diagramm

$$(*) \qquad \begin{array}{ccc} U \times V & \xrightarrow{F} & W \\ \otimes \downarrow & \nearrow \tilde{F} & \\ T & & \end{array}$$

kommutiert.

Theorem 15.1
Zu jedem Paar U, V gibt es ein Tensorprodukt $(T, \otimes)$. Dieses ist bis auf Isomorphie eindeutig bestimmt, d. h. ist $(\tilde{T}, \tilde{\otimes})$ ein weiteres Tensorprodukt, so gibt es einen Isomorphismus $T \xrightarrow{\cong} \tilde{T}$, so daß das Diagramm

$$\begin{array}{ccccc} & & U \times V & & \\ & \otimes \swarrow & & \searrow \tilde{\otimes} & \\ T & & \xrightarrow{\sim} & & \tilde{T} \end{array}$$

kommutiert.

Wir stellen den Beweis zunächst zurück.

Sprechweise. T heißt *das* Tensorprodukt von U und V. Die Elemente von T heißen *Tensoren*.

Schreibweise. $T = U \otimes V$; $u \otimes v := \otimes(u, v)$, $\quad (u \in U, v \in V)$.

Bemerkung. Man beachte, daß die Elemente $u \otimes v$ den Raum T erzeugen, daß aber nicht jeder Tensor von dieser Form ist. Ein allgemeiner Tensor t in T ist eine endliche Summe $t = \sum_{i=1}^{n} u_i \otimes v_i$.

Lemma 15.2
$\otimes : U \times V \to T$ sei bilinear. Dann sind äquivalent:

(i) (T1) und (T2) gelten.

(ii) Es gilt:

> (T) Zu jeder bilinearen Abbildung $F : U \times V \to W$ gibt es *genau eine* lineare Abbildung $\tilde{F} : T \to W$, die $(*)$ kommutativ macht.

Beweis. (i)$\Rightarrow$(ii): Es seien $\tilde{F}_1, \tilde{F}_2 : T \to W$, so daß $(*)$ kommutiert, dann gilt:

$$\tilde{F}_1(u \otimes v) = F(u,v) = \tilde{F}_2(u \otimes v) \qquad (u \in U, v \in V).$$

Da die Elemente $u \otimes v$ den Raum T erzeugen, folgt $\tilde{F}_1 = \tilde{F}_2$.

(ii)$\Rightarrow$(i): Es gelte (T). Dann gilt auch (T2). Es bleibt (T1) zu zeigen. Es sei

$$\begin{gathered} T' := \operatorname{Span}\{u \otimes v,\ u \in U,\ v \in V\} \subset T, \\ i : T' \hookrightarrow T \text{ sei die Inklusion.} \end{gathered}$$

Wir betrachten die bilineare Abbildung

$$\begin{array}{rrcl} F := \otimes : & U \times V & \longrightarrow & T \\ & (u,v) & \longmapsto & u \otimes v. \end{array}$$

Mit $\tilde{F} = \mathrm{id}_T$ gilt $F = \tilde{F} \circ \otimes$. Es sei $F_0 : U \times V \to T'$ durch $F_0(u \times v) = \tilde{F}_0(u \times v)$ definiert. Dann gibt es, wendet man die universelle Eigenschaft auf F_0 an, eine lineare Abbildung $\tilde{F}_0$ mit $F_0 = \tilde{F}_0 \circ \otimes$.

$$\begin{tikzcd} U \times V \arrow[r, "F_0"] \arrow[d, "\otimes"'] \arrow[rr, bend left, "F"] & T' \arrow[r, "i"] & T \\ T \arrow[ur, "\tilde{F}_0"] \arrow[urr, "\tilde{F} = \mathrm{id}_T"'] & & \end{tikzcd}$$

Es gilt

$$(i \circ \tilde{F}_0) \circ \otimes = i \circ (\tilde{F}_0 \circ \otimes) = i \circ F_0 = F.$$

Wegen der Eindeutigkeitsaussage von (T) folgt:

$$\mathrm{id}_T = \tilde{F} = (i \circ \tilde{F}_0).$$

Insbesondere ist i daher surjektiv, d. h. $T' = T$. $\square$

Beispiel. $U = K$. V sei beliebig. Wir setzen

$$\begin{array}{rrrcl} T := V, & \otimes : & K \times V & \longrightarrow & V \\ & & (\lambda, v) & \longmapsto & \lambda v =: \lambda \otimes v. \end{array}$$

Die Abbildung $\otimes$ ist natürlich bilinear.

Behauptung. $(V, \otimes)$ ist Tensorprodukt von K und V. Wir müssen hierzu (T1) und (T2) zeigen.

(T1): Klar.

(T2): $F : K \times V \to W$ sei bilinear.

Wir betrachten das folgende Diagramm:

$$\begin{tikzcd}
K \times V \arrow[r, "F"] \arrow[d, "\otimes"'] & W \\
V \arrow[ur, "\tilde{F}"'] &
\end{tikzcd}$$

Dabei ist $\tilde{F}$ definiert durch

$$\tilde{F}(v) := F(1, v).$$

(α) Das Diagramm kommutiert:

$$\tilde{F}(\otimes(\lambda, v)) = \tilde{F}(\lambda \otimes v) = \tilde{F}(\lambda v) = F(1, \lambda v) = F(\lambda, v).$$

(β) $\tilde{F}$ ist linear:

$$\begin{aligned}
\tilde{F}(\lambda_1 v_1 + \lambda_2 v_2) &= F(1, \lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 F(1, v_1) + \lambda_2 F(1, v_2) \\
&= \lambda_1 \tilde{F}(v_1) + \lambda_2 \tilde{F}(v_2).
\end{aligned}$$

Bevor wir das Theorem (15.1) beweisen, erinnern wir an den Raum $\mathrm{Abb}[X, K]$, den wir für eine Menge X wir folgt definiert hatten:

$$\mathrm{Abb}[X, K] := \{f;\ f : X \longrightarrow K;\ f(x) \neq 0 \text{ für nur endlich viele } x \in X\}.$$

Dies ist ein Vektorraum über K. Wir setzen

$$F(X) := \mathrm{Abb}[X, K]$$

und bezeichnen diesen Vektorraum als den *freien Vektorraum* über der Menge X. Wir erhalten eine *Basis* von $F(X)$ durch die Elemente $f_x, x \in X$ wobei

$$\begin{aligned}
f_x : \quad X &\longrightarrow K \\
f_x(y) &:= \begin{cases} 1 & \text{falls } y = x \\ 0 & \text{falls } y \neq x. \end{cases}
\end{aligned}$$

Jedes Element $f \in F(X)$ besitzt eine eindeutige Darstellung

$$(*) \qquad f = \sum_{x,\text{endlich}} \lambda_x f_x \qquad (\lambda_x \in K).$$

Mittels der Inklusion $X \to F(X), x \mapsto f_x$ kann man X als Teilmenge von $F(X)$ auffassen. Identifiziert man mittels dieser Inklusion x mit f_x, so kann man $(*)$ in der folgenden Form schreiben:

$$(**) \qquad f = \sum_{x,\text{endlich}} \lambda_x x \qquad (\lambda_x \in K).$$

Schließlich benötigen wir noch das

Lemma 15.3
Es sei $f : U \to W$ eine lineare Abbildung mit $V \subset \operatorname{Ker} f$. Dann gibt es genau eine lineare Abbildung $\overline{f} : U/V \to W$, so daß das Diagramm

$$\begin{tikzcd}
U \arrow[r, "f"] \arrow[d, "\pi"'] & W \\
U/V \arrow[ur, "\overline{f}"'] &
\end{tikzcd}$$

kommutiert. (Dabei ist π die kanonische Projektion).

Beweis. Falls das Diagramm kommutiert, muß gelten

$$(*) \qquad \overline{f}(\overline{u}) = f(u).$$

Dies ergibt die Eindeutigkeit. Wir möchten nun $\overline{f}$ durch $(*)$ erklären. Es ist zu zeigen:

(i) $\overline{f}$ *ist wohldefiniert*: Seien $u_1, u_2 \in U$ mit $\overline{u}_1 = \overline{u}_2$. Dann gilt

$$u_1 - u_2 \in V,$$

d. h.

$$u_1 = u_2 + v \qquad (\text{für ein } v \in V).$$

Es gilt

$$\begin{aligned}
\overline{f}(\overline{u}_1) &= f(u_1) = f(u_2 + v) = f(u_2) + f(v) \\
&\overset{V \subset \operatorname{Ker} f}{=} f(u_2) = \overline{f}(\overline{u}_2).
\end{aligned}$$

(ii) $\overline{f}$ *ist linear*:

$$\begin{aligned}\overline{f}(\lambda_1\overline{u}_1+\lambda_2\overline{u}_2) &= \overline{f}(\overline{\lambda_1 u_1+\lambda_2 u_2}) = f(\lambda_1 u_1+\lambda_2 u_2)\\ &= \lambda_1 f(u_1)+\lambda_2 f(u_2) = \lambda_1\overline{f}(\overline{u}_1)+\lambda_2\overline{f}(\overline{u}_2).\end{aligned}$$

Die Kommutativität folgt aus $(*)$. □

Bemerkung. Dieses Lemma ist nur eine geringfügige Umformulierung von Satz (11.51).

Beweis von Theorem (15.1).

Eindeutigkeit: Es seien $(T,\otimes)$, sowie $(\tilde{T},\tilde{\otimes})$ Tensorprodukte von U und V. Dann hat man ein kommutatives Diagramm

$$\begin{tikzcd} U\times V \arrow[rr, "\tilde{\otimes}"] \arrow[d, "\otimes"'] & & \tilde{T} \arrow[dll, shift left, "g"] \\ T \arrow[urr, shift left, "f"] & & \end{tikzcd}$$

wobei $f(u\otimes v)=u\tilde{\otimes}v$ und $g(u\tilde{\otimes}v)=u\otimes v$ ist. Also gilt:

$$\begin{aligned}(g\circ f)(u\otimes v) &= u\otimes v \overset{(T1)}{\Rightarrow} g\circ f=\mathrm{id}_T.\\ (f\circ g)(u\tilde{\otimes}v) &= u\tilde{\otimes}v \overset{(T1)}{\Rightarrow} f\circ g=\mathrm{id}_{\tilde{T}}.\end{aligned}$$

Damit ist f ein Isomorphismus mit $f^{-1}=g$.

Existenz: Wir betrachten den freien Vektorraum $F(U\times V)$. $N(U\times V)$ sei der von folgenden Elementen erzeugte Unterraum:

$$\begin{gathered}(\lambda_1 u_1+\lambda_2 u_2, v_1)-\lambda_1(u_1,v_1)-\lambda_2(u_2,v_1)\\ (u_1,\lambda_1 v_1+\lambda_2 v_2)-\lambda_1(u_1,v_1)-\lambda_2(u_1,v_2).\end{gathered}$$

Dabei sind $u_1,u_2\in U$, $v_1,v_2\in V$. Wir setzen

$$T := F(U\times V)/N(U\times V).$$

Ferner haben wir die natürliche Quotientenabbildung

$$\pi:\quad F(U\times V)\longrightarrow T.$$

Damit definieren wir

$$\begin{array}{cccc}\otimes: & U\times V & \to F(U\times V)\overset{\pi}{\to} & T\\ & (u,v) & \longmapsto & \pi(u,v)=:\overline{(u,v)}=:u\otimes v.\end{array}$$

$\otimes$ *ist bilinear*:

$$(\lambda_1 u_1 + \lambda_2 u_2) \otimes v_1 \quad = \quad \lambda_1 u_1 \otimes v_1 + \lambda_2 u_2 \otimes v_1.$$

Dies gilt, da

$$(\lambda_1 u_1 + \lambda_2 u_2, v_1) - \lambda_1(u_1, v_1) - \lambda_2(u_2, v_1) \in N(U \times V)$$

ist. Analog zeigt man die Behauptung im zweiten Argument.

Nachweis von (T1): Es sei $t \in T$. Da π surjektiv ist, gilt

$$t = \overline{\sum_{\text{endlich}} \alpha_{ij}(u_i, v_j)} = \sum_{\text{endlich}} \alpha_{ij}\overline{(u_i, v_j)} = \sum_{\text{endlich}} \alpha_{ij} u_i \otimes v_j.$$

D. h. die Elemente der Form $u \otimes v$ spannen T auf.

Nachweis von (T2): Es sei

$$F: \quad U \times V \longrightarrow W$$

eine bilineare Abbildung. Wir definieren

$$\hat{F}: \quad F(U \times V) \longrightarrow W$$

durch

(1) $$\hat{F}(u, v) := F(u, v).$$

$\hat{F}$ ist also auf der Basis festgelegt und ergibt so eine eindeutig bestimmte lineare Abbildung $\hat{F}: F(U \times V) \to W$.

Behauptung: $N(U \times V) \subset \operatorname{Ker} \hat{F}$.

Dies folgt, da

$$\begin{aligned} \hat{F}((\lambda_1 u_1 + \lambda_2 u_2, v) - \lambda_1(u_1, v) - \lambda_2(u_2, v)) &= \\ F(\lambda_1 u_1 + \lambda_2 u_2, v) - \lambda_1 F(u_1, v) - \lambda_2 F(u_2, v) &= 0, \end{aligned}$$

wobei wir verwendet haben, daß F bilinear ist. Analog im zweiten Argument. Damit erhalten wir nach Lemma (15.3) ein kommutatives Diagramm

$$\begin{tikzcd} F(U\times V) \arrow[r, "F"] \arrow[d, "\pi"'] & W \\ (U\times V)/N(U\times V) \arrow[ur, "\tilde{F}"'] & \end{tikzcd}$$

Hierbei gilt

$$\tilde{F}(\overline{(u,v)}) = \hat{F}(u,v). \tag{2}$$

Dies liefert eine lineare Abbildung $\tilde{F}$. Es bleibt zu beweisen, daß das folgende Diagramm kommutiert:

$$\begin{tikzcd} U\times V \arrow[r, "F"] \arrow[d, "\otimes"'] & W \\ T \arrow[ur, "\tilde{F}"'] & \end{tikzcd}$$

Dies folgt, da

$$(\tilde{F}\circ\otimes)(u,v) = \tilde{F}(u\otimes v) = \tilde{F}(\overline{(u,v)}) \overset{(2)}{=} \hat{F}(u,v) \overset{(1)}{=} F(u,v).$$

Dies beendet den Beweis. $\square$

Bemerkung. Die Abbildung $u\otimes v \mapsto v\otimes u$ liefert einen Isomorphismus

$$U\otimes V \cong V\otimes U.$$

Wir untersuchen nun *einfache Eigenschaften* des Tensorproduktes.

Bemerkung. $0\otimes v = u\otimes 0 = 0$.

Denn: $0\otimes v = (0+0)\otimes v = 0\otimes v + 0\otimes v$, also ist $0\otimes v = 0$. Analog zeigt man $u\otimes 0 = 0$.

Lemma 15.4
Es seien $u_1,\dots,u_r \in U$ linear unabhängig, sowie $v_1,\dots,v_r\in V$. Ist

$$\sum_{j=1}^{r} u_j\otimes v_j = 0,$$

so ist $v_1 = \ldots = v_r = 0$.

Beweis. Es sei $W := K$. Da $u_1, \dots, u_r$ linear unabhängig sind, gibt es lineare Abbildungen

$$f_i : \quad U \longrightarrow K \quad \text{ mit } f_i(u_j) = \delta_{ij}.$$

Es seien $g_i : V \to K$, $i = 1, \dots, r$ beliebige lineare Abbildungen. Dann ist

$$\begin{aligned} F : \quad U \times V & \longrightarrow K \\ F(u,v) & := \sum_{i=1}^{r} f_i(u) g_i(v) \end{aligned}$$

bilinear. Also kann man F über $U \otimes V$ faktorisieren, d. h. es gibt $\tilde{F} : T \to K$ linear mit

$$\tilde{F}(u \otimes v) = \sum_{i=1}^{r} f_i(u) g_i(v).$$

Damit gilt

$$0 = \tilde{F}(\sum_{j=1}^{r} u_j \otimes v_j) = \sum_{j=1}^{r} \tilde{F}(u_j \otimes v_j) = \sum_{i,j=1}^{r} \underbrace{f_i(u_j)}_{\delta_{ij}} g_i(v_j).$$

Also ist

$$0 = \sum_{i=1}^{r} g_i(v_i),$$

wobei die $g_i : V \to K$ beliebige Homomorphismen waren. Für festes $i_0 \in \{1, \dots, r\}$ wähle man $g_i = 0$ für $i \neq i_0$. Dann gilt

$$g_{i_0}(v_{i_0}) = 0$$

für alle $g_{i_0} : V \to K$. Also ist $v_{i_0} = 0$. $\square$

Korollar 15.5
$u \neq 0, v \neq 0 \Rightarrow u \otimes v \neq 0$.

Lemma 15.6
Es sei $\{e_i\}_{i \in I}$ eine Basis von U. Dann besitzt jeder Vektor $t \neq 0$ eine Darstellung

$$t = \sum_{i,\text{endlich}} e_i \otimes v_i \qquad (v_i \neq 0).$$

Die v_i sind eindeutig bestimmt.

Beweis. Wegen (T1) gibt es eine Darstellung

$$t = \sum_{k,\text{endlich}} u_k \otimes v'_k, \qquad u_k \in U, v'_k \in V.$$

Da $\{e_i\}_{i \in I}$ eine Basis von U ist, hat man

$$u_k = \sum_{j,\text{endlich}} \lambda_k^j e_j.$$

Also ist

$$\begin{aligned} t &= \sum_k (\sum_j \lambda_k^j e_j) \otimes v'_k = \sum_{k,j} e_j \otimes \lambda_k^j v'_k \\ &= \sum_j e_j \otimes \underbrace{(\sum_k \lambda_k^j v'_k)}_{=:v_j} = \sum_j e_j \otimes v_j. \end{aligned}$$

Die Summe ist endlich. Läßt man die v_j mit $v_j = 0$ weg, so erhält man die gewünschte Darstellung.

Um die Eindeutigkeit zu zeigen, sei

$$t = \sum_{j \in J_1} e_j \otimes v_j = \sum_{j \in J_2} e_j \otimes v'_j.$$

Wir setzen

$$\begin{aligned} v_j &= 0 \text{ falls } j \in J_2 \setminus J_1. \\ v'_j &= 0 \text{ falls } j \in J_1 \setminus J_2. \end{aligned}$$

Damit erhält man

$$0 = \sum_{j \in J_1 \cup J_2} e_j \otimes (v_j - v'_j) \overset{(15.4)}{\Rightarrow} v_j - v'_j = 0 \qquad (j \in J_1 \cup J_2).$$

□

Lemma 15.7
Jeder Vektor $0 \neq t \in T$ besitzt eine Darstellung

$$t = \sum_{i=1}^{r} u_i \otimes v_i \qquad (u_i \in U_i, v_i \in V)$$

wobei die $(u_i)_{i=1,\dots,r}$ sowie die $(v_i)_{i=1,\dots,r}$ linear unabhängig sind.

Beweis. Es sei r minimal, so daß eine Darstellung

$$t = \sum_{i=1}^{r} u_i \otimes v_i$$

existiert. Ist $r = 1$, so ist $u_1 \neq 0 \neq v_1$, da $t \neq 0$. Sei also $r \geq 2$.

Annahme: Die u_i sind linear abhängig. Dann kann man annehmen (nach einer eventuellen Umnummerierung), daß

$$u_r = \sum_{i=1}^{r-1} \lambda_i u_i.$$

Also erhält man

$$\begin{aligned} t &= \sum_{i=1}^{r-1} u_i \otimes v_i + \sum_{i=1}^{r-1} \lambda_i u_i \otimes v_r \\ &= \sum_{i=1}^{r-1} u_i \otimes (\underbrace{v_i + \lambda_i v_r}_{=:v_i'}) = \sum_{i=1}^{r-1} u_i \otimes v_i'. \end{aligned}$$

Dies ist ein Widerspruch zur Minimalität von r. Analog zeigt man, daß die v_i linear unabhängig sind. □

Satz 15.8
Es seien $\{e_i\}_{i\in I}$ und $\{f_j\}_{j\in J}$ Basen von U, bzw. V. Die Elemente $\{e_i \otimes f_j\}_{(i,j)\in I\times J}$ bilden eine Basis des Vektorraums $U \otimes V$.

Korollar 15.9
Ist $\dim U = n < \infty$, $\dim U = m < \infty$, so ist

$$\dim(U \otimes V) = nm = \dim U \, \dim V.$$

Beweis von Satz (15.8).

Erzeugendensystem: Es genügt zu sehen, daß jeder Tensor $u \otimes v \in U \otimes V$ durch die Tensoren $e_i \otimes f_j$ dargestellt werden kann. Es sei dazu

$$u = \sum_i \lambda_i e_i, \qquad v = \sum_j \lambda_j f_j.$$

Also gilt

$$u \otimes v = (\sum_i \lambda_i e_i) \otimes (\sum_j \lambda_j f_j) = \sum_{i,j} \lambda_i \lambda_j e_i \otimes f_j.$$

Lineare Unabhängigkeit: Es gilt:

$$\begin{aligned}\sum_{i,j} \lambda_{ij} e_i \otimes f_j = 0 \quad &\Rightarrow \quad \sum_i e_i \otimes (\sum_j \lambda_{ij} f_j) = 0\\ &\overset{(15.4)}{\Rightarrow} \quad \sum_j \lambda_{ij} f_j = 0 \text{ für alle } i\\ &\overset{(f_j) \text{ Basis}}{\Rightarrow} \quad \lambda_{ij} = 0 \text{ für alle } i, j.\end{aligned}$$

□

Beispiel. Es sei $U = K^n, V = K^m$ mit den Standardbasen $e_1, \dots, e_n$, bzw. $e_1, \dots, e_m$. Dann besitzt jeder Tensor $t \in U \otimes V$ eine eindeutige Darstellung der Form

$$t = \sum_{\substack{i=1,\dots,n\\ j=1,\dots,m}} \lambda_{ij} e_i \otimes e_j.$$

Dualraum

Es seien V, W Vektorräume über dem selben Grundkörper K. Dann hatten wir bereits bemerkt, daß

$$\mathrm{Hom}_K(V, W) = \{f;\ f : V \to W \text{ ist Homomorphismus}\}$$

in natürlicher Weise ein K-Vektorraum ist, wenn man Addition und Skalarmultiplikation wie folgt erklärt:

$$\begin{aligned}(f+g)(x) &= f(x) + g(x)\\ (\alpha f)(x) &= \alpha f(x).\end{aligned}$$

Setzt man speziell $W = K$, so erhält man die folgende

Definition. Der Vektorraum

$$V^* := \mathrm{Hom}_K(V, K)$$

heißt der zu V *duale Vektorraum*. Die Elemente aus V^* heißen *Linearformen* auf V (bzw. *lineare Funktionale*).

Beispiele.

(1) $V = K^n$:

$$\begin{aligned}\varphi : \quad V &\longrightarrow K\\ \begin{pmatrix} x_1\\ \vdots\\ x_n \end{pmatrix} &\longmapsto x_1 + \cdots + x_n.\end{aligned}$$

(2) $V = C[a,b]$:

$$\begin{array}{rccc} \varphi: & C[a,b] & \longrightarrow & \mathbb{R} \\ & f & \longmapsto & \int_a^b f(t)\,dt. \end{array}$$

Es sei nun $0 < \dim V = n < \infty$. Ferner sei $v_1, \dots, v_n$ eine Basis von V. Dann gibt es eindeutig bestimmte Linearformen

$$v_i^* : V \longrightarrow K,$$

die durch

$$v_i^*(v_j) := \delta_{ij} := \begin{cases} 1 & \text{falls } j = i \\ 0 & \text{falls } j \neq i \end{cases}$$

definiert sind.

Satz 15.10
$v_1^*, \dots, v_n^*$ sind eine Basis von V^*. Insbesondere gilt $\dim V^* = \dim V = n$.

Definition. $v_1^*, \dots, v_n^*$ heißt die zu $v_1, \dots, v_n$ *duale Basis*.

Beweis. *Lineare Unabhängigkeit:*

$$\sum_{i=1}^n \alpha_i v_i^* = 0 \Rightarrow \sum_{i=1}^n \alpha_i \underbrace{v_i^*(v_j)}_{\delta_{ij}} = 0 \Rightarrow \alpha_j = 0 \quad (j = 1, \dots, n).$$

Erzeugendensystem: Es sei $v^* \in V^*$. Wir setzen

$$\alpha_i := v^*(v_i).$$

Dann gilt $v^* = \sum_{i=1}^n \alpha_i v_i^*$, da

$$(\sum_{i=1}^n \alpha_i v_i^*)(v_k) = \sum_{i=1}^n \alpha_i v_i^*(v_k) = \alpha_k = v^*(v_k) \qquad (k = 1, \dots, n).$$

Also folgt, da eine lineare Abbildung durch die Bilder einer Basis eindeutig festgelegt ist, daß

$$v^* = \sum_{i=1}^n \alpha_i v_i^*.$$

□

Satz 15.11
Es gibt genau eine lineare Abbildung

$$T: \quad U^* \otimes V \longrightarrow \mathrm{Hom}(U, V)$$

mit

$$T(u^* \otimes v)(u) = u^*(u)v.$$

Diese ist injektiv. Sind U, V endlich dimensional, so ist T ein Isomorphismus.

Bemerkung. Für U, V endlich dimensional nennt man T einen „kanonischen" Isomorphismus, da dieser ohne Bezug auf eine Basis definiert werden kann.

Beweis. *1. Schritt*: Wir betrachten die Abbildung

$$T': \quad U^* \times V \longrightarrow \mathrm{Hom}(U, V); \qquad T'(u^*, v)(u) := u^*(u)v.$$

Es ist $T'(u^*, v) \in \mathrm{Hom}(U, V)$, da

$$\begin{aligned} T'(u^*, v)(\lambda_1 u_1 + \lambda_2 u_2) &= u^*(\lambda_1 u_1 + \lambda_2 u_2)v = \lambda_1 u^*(u_1)v + \lambda_2 u^*(u_2)v \\ &= \lambda_1 T'(u^*, v)(u_1) + \lambda_2 T'(u^*, v)(u_2). \end{aligned}$$

Eine ähnliche Rechnung zeigt, daß T' eine bilineare Abbildung ist. Also gibt es genau eine lineare Abbildung

$$T: \quad U^* \otimes V \longrightarrow \mathrm{Hom}(U, V)$$

mit

$$T(u^* \otimes v) = T'(u^*, v).$$

2. Schritt: T ist injektiv. Es sei $0 \neq t \in \mathrm{Ker}\, T$. Wir wählen eine Darstellung nach Lemma (15.7)

$$t = \sum_i u_i^* \otimes v_i$$

mit linear unabhängigen Vektoren v_i. Dann gilt

$$\begin{aligned} T(t) = 0 \quad &\Rightarrow \quad T(t)(u) = 0 \text{ für alle } u \in U \\ &\Rightarrow \quad \sum_i u_i^*(u)v_i = 0 \text{ für alle } u \in U \\ &\overset{v_i \text{ linear unabh.}}{\Rightarrow} \quad u_i^*(u) = 0 \text{ für alle } u \\ &\Rightarrow \quad u_i^* = 0 \text{ für alle } i \\ &\Rightarrow \quad t = 0. \end{aligned}$$

3. Schritt: Ist $\dim U = n, \dim V = m$ mit $n, m < \infty$, so folgt die Behauptung aus

$$\dim(U^* \otimes V) \overset{(15.9)}{=} nm = \dim \mathrm{Hom}(U, V).$$

□

Tensorprodukte von Unterräumen und Quotienten

Es seien $U_1 \subset U$, $V_1 \subset V$ Unterräume.

Satz 15.12
Man hat eine natürliche Inklusion

$$U_1 \otimes V_1 \subset U \otimes V.$$

Beweis. Sind $i_U : U_1 \to U$ und $i_V : V_1 \to V$ die Inklusionen, so induziert die bilineare Abbildung

$$u_1 \times v_1 \longmapsto i_U(u_1) \otimes i_V(v_1), \qquad u_1 \in U_1, v_1 \in V_1$$

eine Inklusion $U_1 \otimes V_1 \subset U \otimes V$. $\square$

Ebenfalls mit Standardmethoden beweist man:

Satz 15.13
Es gibt einen natürlichen Isomorphismus

$$U/U_1 \otimes V/V_1 \cong U \otimes V/(U_1 \otimes V + U \otimes V_1)$$

mit

$$\overline{u} \otimes \overline{v} \longmapsto \overline{u \otimes v}.$$

Tensorprodukte und lineare Abbildungen

Gegeben seien lineare Abbildungen

$$\varphi : \quad U \longrightarrow U', \qquad \psi : \quad V \longrightarrow V'.$$

Dann ist das Produkt

$$\varphi \times \psi : \quad U \times V \longrightarrow U' \times V'$$

eine lineare Abbildung. Diese definiert eine Abbildung

$$\begin{array}{rcccl} \overline{\varphi \times \psi} : & U \times V & \overset{\varphi\times\psi}{\longrightarrow} & U' \times V' \longrightarrow U' \otimes V' \\ & (u,v) & \longmapsto & (\varphi(u), \psi(v)) \longmapsto \varphi(u) \otimes \psi(v). \end{array}$$

Diese Abbildung ist bilinear. Also gibt es ein kommutatives Diagramm

$$\begin{tikzcd} U \times V \arrow[r, "\varphi \times \psi"] \arrow[d, "\otimes"'] & U' \\ U \otimes V \arrow[ur, "T(\varphi, \psi)"'] & \end{tikzcd}$$

mit einer linearen Abbildung $T(\varphi, \psi)$, die durch

$$T(\varphi, \psi)(u \otimes v) = \varphi(u) \otimes \psi(v)$$

bestimmt wird.

Beispiel. $U = U' = V = V' = \mathbb{R}^2$.

$$\begin{array}{rcll} \varphi: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \text{ sei gegeben durch } A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \\ \psi: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \text{ sei gegeben durch } B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \end{array}$$

Als Basis von $\mathbb{R}^2 \otimes \mathbb{R}^2$ wählen wir: $e_1 \otimes e_1$, $e_1 \otimes e_2$, $e_2 \otimes e_1$, $e_2 \otimes e_2$. Dann gilt

$$\begin{array}{rcl} T(\varphi,\psi)(e_1 \otimes e_1) & = & \varphi(e_1) \otimes \psi(e_1) = (e_1 + 2e_2) \otimes e_2 = e_1 \otimes e_2 + 2e_2 \otimes e_2 \\ T(\varphi,\psi)(e_1 \otimes e_2) & = & \varphi(e_1) \otimes \psi(e_2) = (e_1 + 2e_2) \otimes e_1 = e_1 \otimes e_1 + 2e_2 \otimes e_1 \\ T(\varphi,\psi)(e_2 \otimes e_1) & = & \varphi(e_2) \otimes \psi(e_1) = (2e_1 + e_2) \otimes e_2 = 2e_1 \otimes e_2 + e_2 \otimes e_2 \\ T(\varphi,\psi)(e_2 \otimes e_2) & = & \varphi(e_2) \otimes \psi(e_2) = (2e_1 + e_2) \otimes e_1 = 2e_1 \otimes e_1 + e_2 \otimes e_1. \end{array}$$

Also wird $T(\varphi, \psi)$ dargestellt durch die Matrix

$$\begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 2 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1B & 2B \\ 2B & 1B \end{pmatrix}.$$

Insgesamt erhalten wir eine Abbildung

$$\begin{array}{rrcl} T: & \mathrm{Hom}(U, U') \times \mathrm{Hom}(V, V') & \longrightarrow & \mathrm{Hom}(U \otimes V, U' \otimes V') \\ & (\varphi, \psi) & \longmapsto & T(\varphi, \psi). \end{array}$$

Die Abbildung T ist selbst bilinear, induziert also eine lineare Abbildung

$$\tilde{T}: \quad \mathrm{Hom}(U, U') \otimes \mathrm{Hom}(V, V') \longrightarrow \mathrm{Hom}(U \otimes V, U' \otimes V')$$

mit

$$\varphi \otimes \psi \longmapsto T(\varphi, \psi).$$

Satz 15.14
$\tilde{T}$ ist injektiv.

Bemerkung. Dies rechtfertigt die Identifikation $T(\varphi, \psi) = \varphi \otimes \psi$.

Beweis. Es sei

$$\operatorname{Ker} \tilde{T} \ni \ \tau = \sum_{i=1}^{r} \varphi_i \otimes \psi_i \neq 0$$

wobei die φ_i, bzw. die ψ_i linear unabhängig sind. (Eine solche Darstellung existiert nach Lemma (15.7).) Also gilt

$$(*) \qquad 0 = \sum_{i=1}^{r} (\varphi_i \otimes \psi_i)(u \otimes v) = \sum_{i=1}^{r} \varphi_i(u) \otimes \psi_i(v)$$

für alle $u \in U$, $v \in V$. Für festes $u \in U$ sei

$$p := \dim \operatorname{Span}(\varphi_1(u), \ldots, \varphi_r(u)).$$

Wir können annehmen, daß $1 \leq p \leq r-1$ für ein $u \in U$ gilt, da sonst entweder $\varphi_1 = \ldots = \varphi_r = 0$ oder nach Lemma (15.4) $\psi_1 = \ldots = \psi_r = 0$ ist. Wir können ferner nach eventuellem Umnumerieren annehmen, daß $\varphi_1(u), \ldots, \varphi_p(u)$ linear unabhängig sind. Also gibt es Darstellungen

$$\varphi_j(u) = \sum_{i=1}^{p} \lambda_{ij} \varphi_i(u) \qquad (j = p+1, \ldots, r).$$

Aus $(*)$ folgt

$$0 = \sum_{i=1}^{p} \varphi_i(u) \otimes \psi_i(v) + \sum_{j=p+1}^{r} (\sum_{i=1}^{p} \lambda_{ij} \varphi_i(u)) \otimes \psi_j(v).$$

Also gilt

$$0 = \sum_{i=1}^{p} \varphi_i(u) \otimes (\psi_i(v) + \sum_{j=p+1}^{r} \lambda_{ij} \psi_j(v)).$$

Da $\varphi_1(u), \ldots, \varphi_p(u)$ linear unabhängig sind, folgt nach Lemma (15.4):

$$\psi_i(v) + \sum_{j=p+1}^{r} \lambda_{ij} \psi_j(v) = 0 \qquad (1 \leq i \leq p, v \in V).$$

Damit folgt

$$\psi_i = - \sum_{j=p+1}^{r} \lambda_{ij} \psi_j,$$

ein Widerspruch zur linearen Unabhängigkeit der ψ_i. $\square$

Korollar 15.15
Sind U, V, U', V' endlich-dimensional, dann gibt es einen kanonischen Isomorphismus

$$\tilde{T}: \quad \mathrm{Hom}(U,V) \otimes \mathrm{Hom}(U',V') \xrightarrow{\sim} \mathrm{Hom}(U \otimes U', V \otimes V').$$

Beweis. Dies folgt sofort aus Satz (15.14), da $\tilde{T}$ injektiv ist, und die Dimensionen der beiden Räume übereinstimmen. □

Komposition von Abbildungen

Gegeben seien lineare Abbildungen

$$\begin{array}{llll} \varphi: & U & \longrightarrow & U', \qquad \varphi': U' \longrightarrow U'' \\ \psi: & V & \longrightarrow & V', \qquad \psi': V' \longrightarrow V''. \end{array}$$

Bemerkung. Es gilt

$$(\varphi' \otimes \psi') \circ (\varphi \otimes \psi) = (\varphi' \circ \varphi) \otimes (\psi' \circ \psi).$$

Denn:

$$\begin{aligned} (\varphi' \otimes \psi') \circ (\varphi \otimes \psi)(u \otimes v) &= (\varphi' \otimes \psi')(\varphi(u) \otimes \psi(v)) \\ &= \varphi'(\varphi(u)) \otimes \psi'(\psi(v)) = (\varphi' \circ \varphi)(u) \otimes (\psi' \circ \psi)(v) \\ &= ((\varphi' \circ \varphi) \otimes (\psi' \circ \psi))(u \otimes v). \end{aligned}$$

Lemma 15.16
Sind φ, ψ injektiv, dann ist auch $\varphi \otimes \psi$ injektiv.

Beweis. $\varphi : U \to U'$, $\psi : V \to V'$ seien injektiv. Also gibt es $\tilde{\varphi} : U' \to U$, $\tilde{\psi} : V' \to V$ mit:

$$\tilde{\varphi} \circ \varphi = \mathrm{id}_U, \qquad \tilde{\psi} \circ \psi = \mathrm{id}_V.$$

Daraus folgt

$$(\tilde{\varphi} \otimes \tilde{\psi}) \circ (\varphi \otimes \psi) = (\tilde{\varphi} \circ \varphi) \otimes (\tilde{\psi} \circ \psi) = \mathrm{id}_U \otimes \mathrm{id}_V = \mathrm{id}_{U \otimes V}.$$

Damit ist $\varphi \otimes \psi$ injektiv. □

Bild und Kern

Ohne Beweis halten wir die beiden folgenden, relativ leicht zu beweisenden, Aussagen fest:

Satz 15.17
Sind $\varphi : U \to U'$, $\psi : V \to V'$ lineare Abbildungen, so gilt:

(i) $\operatorname{Im}(\varphi \otimes \psi) = \operatorname{Im} \varphi \otimes \operatorname{Im} \psi$

(ii) $\operatorname{Ker}(\varphi \otimes \psi) = \operatorname{Ker} \varphi \otimes V + U \otimes \operatorname{Ker} \psi$.

Bemerkung. Satz (15.17) (ii) impliziert natürlich auch Lemma (15.16).

Das Kroneckerprodukt von Matrizen

Wie stets sei $e_1, \ldots, e_n$, bzw. $e_1, \ldots, e_m$ die Standardbasis von K^n, bzw. K^m. Für $K^n \otimes K^m$ wählen wir $e_1 \otimes e_1, \ldots, e_1 \otimes e_m, e_2 \otimes e_1, \ldots, e_2 \otimes e_m, \ldots, e_n \otimes e_1, \ldots, e_n \otimes e_m$ als Basis. Analog verfahren wir mit K^r, K^s und $K^r \otimes K^s$. Die Abbildungen $\varphi : K^n \to K^r$, $\psi : K^m \to K^s$ werden dargestellt durch die Matrizen $A \in \operatorname{Mat}(r, n; K)$, bzw. $B \in \operatorname{Mat}(s, m; K)$.

Definition. Sei $A \in \operatorname{Mat}(r, n; K)$, $B \in \operatorname{Mat}(s, m; K)$. Das *Kronecker-Produkt* von A und B ist die Matrix

$$C = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \vdots \\ a_{r1}B & \cdots & a_{rn}B \end{pmatrix} \in \operatorname{Mat}(rs, mn; K).$$

Satz 15.18
Werden φ und ψ durch A und B dargestellt, so wird $\varphi \otimes \psi$ bezüglich der oben angegebenen Basen durch das Kroneckerprodukt von A und B dargestellt.

Beweis. Es ist

$$\begin{aligned} (\varphi \otimes \psi)(e_i \otimes e_j) &= \varphi(e_i) \otimes \psi(e_j) \\ &= (\sum_{k=1}^{r} a_{ki} e_k) \otimes (\sum_{l=1}^{s} b_{lj} e_l) \\ &= \sum_{k=1}^{r} \sum_{l=1}^{s} a_{ki} b_{lj} (e_k \otimes e_l). \end{aligned}$$

Andererseits ist die entsprechende Spalte im Kroneckerprodukt

$$\begin{array}{c} a_{1i}b_{1j} \\ \vdots \\ a_{1i}b_{sj} \\ a_{2i}b_{1j} \\ \vdots \\ a_{2i}b_{sj} \\ \vdots \\ a_{ri}b_{1j} \\ \vdots \\ a_{ri}b_{sj}. \end{array}$$

Dies ergibt die Behauptung. □

Mehrfache Tensorprodukte

$U_1, \ldots, U_p, W$ seien K-Vektorräume.

Definition. Eine Abbildung

$$F: \quad U_1 \times \ldots \times U_p \longrightarrow W$$

heißt *p-linear*, falls für jedes $1 \leq i \leq p$ gilt:

$$F(u_1, \ldots, u_{i-1}, \lambda u_i' + \mu u_i'', u_{i+1}, \ldots, u_p) = \\ \lambda F(u_1, \ldots, u_{i-1}, u_i', u_{i+1}, \ldots, u_p) + \mu F(u_1, \ldots, u_{i-1}, u_i'', u_{i+1}, \ldots, u_p).$$

Beispiel. Die Determinantenabbildung

$$\begin{array}{rrcl} \det: & K^n \times \ldots \times K^n & \longrightarrow & K \\ & (a_1, \ldots, a_n) & \longmapsto & \det(a_1, \ldots, a_n) \end{array}$$

ist n-linear.

Definition. T sei ein Vektorraum. Ferner sei

$$\otimes: \quad U_1 \times \ldots \times U_p \longrightarrow T$$

eine p-lineare Abbildung. Dann heißt $(T, \otimes)$ ein *Tensorprodukt* von $U_1, \ldots U_p$, falls gilt:

(T1)' Das Bild von $\otimes$ erzeugt T.

(T2)' Ist $F: U_1 \times \ldots \times U_p \to W$ p-linear, so gibt es eine lineare Abbildung $\tilde{F}: T \to W$, so daß

$$
\begin{tikzcd}
U_1 \times \ldots \times U_p \arrow[r, "F"] \arrow[d, "\otimes"'] & W \\
T \arrow[ur, "\tilde{F}"'] &
\end{tikzcd}
$$

kommutiert.

Bemerkung. Wegen (T1)' ist $\tilde{F}$ eindeutig bestimmt.

Theorem 15.19
Bis auf Isomorphie gibt es genau ein Tensorprodukt von $U_1, \ldots, U_p$.

Beweis. Dies beweist man ganz analog wie wie im Fall $p = 2$. □

Schreibweise.

(i) $T = U_1 \otimes \ldots \otimes U_p$. Man nennt diesen Raum das Tensorprodukt von $U_1, \ldots, U_p$.

(ii) $u_1 \otimes \ldots \otimes u_p := \otimes(u_1, \ldots, u_p)$.

Es sei $\{e^i_\nu\}_{\nu \in J_i}$ eine Basis von U_i $(i = 1, \ldots, p)$.

Satz 15.20
Die Elemente $\{e^1_{\nu_1} \otimes \ldots \otimes e^p_{\nu_p}\}_{(\nu_1, \ldots, \nu_p) \in J_1 \times \ldots \times J_p}$ bilden eine Basis von $U_1 \otimes \ldots \otimes U_p$.

Korollar 15.21
$\dim(U_1 \otimes \ldots \otimes U_p) = \dim U_1 \cdots \dim U_p$.

Beide Aussagen beweist man wiederum genau wie im Fall $p = 2$.

Auch der folgende Satz wird mit Standardmethoden bewiesen.

Satz 15.22
Für $1 \leq p \leq q - 1$ gibt es einen kanonischen Isomorphismus

$$\tilde{\varphi}: \quad (U_1 \otimes \ldots \otimes U_p) \otimes (U_{p+1} \otimes \ldots \otimes U_q) \longrightarrow U_1 \otimes \ldots \otimes U_q$$

mit

$$\tilde{\varphi}((u_1 \otimes \ldots \otimes u_p) \otimes (u_{p+1} \otimes \ldots \otimes u_q)) = u_1 \otimes \ldots \otimes u_q.$$

Die Tensoralgebra

Es sei $(U_i)_{i \in I}$ eine Familie von Vektorräumen.

Definition.

(i) Das *direkte Produkt* der Vektorräume U_i ist definiert durch

$$\prod_{i\in I} U_i := \{(u_i)_{i\in I};\ u_i \in U_i\} := \{u:\ I \longrightarrow \bigcup_{i\in I} U_i\,;\ u:\ i \longmapsto u_i \in U_i\}.$$

(ii) Die *(äußere) direkte Summe* ist definiert durch

$$\bigoplus_{i\in I} U_i := \{(u_i)_{i\in I};\ u_i \in U_i,\ \text{fast alle } u_i = 0\}.$$

Bemerkungen.

(i) Ist I endlich, so stimmen direkte Summe und direktes Produkt überein.

(ii) Das direkte Produkt $\prod_{i\in I} U_i$ wird durch

$$\begin{aligned} (u_i)_{i\in I} + (v_i)_{i\in I} &:= (u_i + v_i)_{i\in I}, \\ \lambda(u_i)_{i\in I} &:= (\lambda u_i)_{i\in I} \end{aligned}$$

in natürlicher Weise zu einem Vektorraum, der $\bigoplus_{i\in I} U_i$ als einen Unterraum enthält.

(iii) Sind $U_1, \dots, U_s \subset V$ Unterräume mit $U_i \cap (\sum_{s\neq i} U_s) = \{0\}$, so gibt es einen natürlichen Isomorphismus der direkten Summe $\bigoplus_{i=1}^{s} U_i \subset V$ mit der oben definierten äußeren direkten Summe.

(iv) Der Zusammenhang mit dem früher eingeführten freien Vektorraum $F(X)$ über einer Menge X ist der, daß $F(X) = \bigoplus_{x\in X} K$ ist.

(v) Sind die $U_i \neq \emptyset$, so folgt die Aussage, daß $\prod U_i \neq \emptyset$ aus dem Auswahlaxiom der Mengenlehre.

Es sei U ein Vektorraum über K. Für $p \geq 2$ setzen wir

$$\otimes^p U := \underbrace{U \otimes \ldots \otimes U}_{p\text{-mal}}.$$

Ferner sei

$$\otimes^0 U := K, \qquad \otimes^1 U := U.$$

Definition.

(i) Die Elemente von $\otimes^p U$ heißen *Tensoren der Stufe p*.

(ii) Ein Tensor $u \in \otimes^p U$ heißt *zerlegbar*, falls es eine Darstellung $u = u_1 \otimes \ldots \otimes u_p$ gibt.

Sei $e_1, \dots, e_n$ eine Basis von U. Dann haben wir bereits gesehen, daß

$$e_{\nu_1} \otimes \ldots \otimes e_{\nu_p}; \qquad 1 \leq \nu_i \leq n$$

eine Basis von $\otimes^p U$ ist. Dies ergibt auch, daß $\dim \otimes^p U = n^p$ ist. Jeder Tensor $u \in \otimes^p U$ hat damit eine Darstellung

$$u = \sum_{\nu_1,\dots,\nu_p} \lambda^{\nu_1\cdots\nu_p} e_{\nu_1} \otimes \ldots \otimes e_{\nu_p}.$$

Schreibweise. $e_{\nu_1\cdots\nu_p} := e_{\nu_1} \otimes \ldots \otimes e_{\nu_p}$.

Damit erhalten wir die Darstellung

$$u = \sum_{\nu_1,\dots,\nu_p} \lambda^{\nu_1,\dots,\nu_p} e_{\nu_1,\dots,\nu_p} =: \lambda^{\nu_1,\dots,\nu_p} e_{\nu_1,\dots,\nu_p}.$$

Summenkonvention: Bei einem Ausdruck der Form $\lambda^{\nu_1,\dots,\nu_p} e_{\nu_1,\dots,\nu_p}$ wird über oben und unten zugleich vorkommende Indizes summiert.

Definition der Tensoralgebra

Wir betrachten die direkte Summe

$$\otimes U := \bigoplus_{p\geq 0} \otimes^p U.$$

Schreibweise. $\otimes U \ni u = (u_0, u_1, \dots) = \sum_{p\geq 0} u_p;\ u_p \in \otimes^p U$ (fast alle $u_p = 0$).

Wir führen nun auf $\otimes U$ eine *Multiplikation* ein. Dazu erinnern wir an folgendes:

Wir hatten gesehen, daß es einen Isomorphismus

$$\tilde{\varphi}: \quad (\otimes^p U) \otimes (\otimes^q U) \xrightarrow{\sim} \otimes^{p+q} U$$

mit

$$\tilde{\varphi}((u_1 \otimes \ldots \otimes u_p) \otimes (u_{p+1} \otimes \ldots \otimes u_{p+q})) = u_1 \otimes \ldots \otimes u_{p+q}$$

gibt. Das heißt: Sind $u \in \otimes^p U$, $v \in \otimes^q U$, so kann man $u \otimes v \in \otimes^{p+q} U$ auffassen.

Sei $u = \sum_{p\geq 0} u_p$, $v = \sum_{p\geq 0} v_p \in \otimes U$. Dann definieren wir die Multiplikation von Tensoren wie folgt:

$$u \cdot v := \sum_{p\geq 0} \Big(\sum_{r+s=p} (u_r \cdot v_s) \Big)$$

wobei

$$u_r \cdot v_s := u_r \otimes v_s \in \otimes^{r+s} U.$$

Bemerkung. $(\otimes U, \cdot)$ ist eine *Algebra*. Sie ist

(i) assoziativ,

(ii) besitzt $1 \in K$ als neutrales multiplikatives Element,

(iii) ist im allgemeinen nicht kommutativ.

Bemerkung. $\otimes U = \oplus_{p\geq 0} \otimes^p U$ ist auch eine *graduierte Algebra*. Ist $u = \sum_{p\geq 0} u_p$, so heißt u_p der *Grad-p Anteil* von u.

Induzierte Abbildungen

Sei

$$\varphi: \quad U \longrightarrow V$$

ein Homomorphismus. Für $p \geq 2$ induziert dies Homomorphismen

$$\otimes^p\varphi := \varphi \otimes \ldots \otimes \varphi : \quad \underbrace{U \otimes \ldots \otimes U}_{p} \longrightarrow \underbrace{V \otimes \ldots \otimes V}_{p}.$$

Ferner sei

$$\varphi^1 := \varphi; \qquad \varphi^0 := \mathrm{id}_K.$$

Dann erhält man den induzierten Homomorphismus

$$\begin{array}{rccl} \otimes\varphi: & \otimes U & \longrightarrow & \otimes V \\ & \displaystyle\sum_{p\geq 0} u_p & \longmapsto & \displaystyle\sum_{p\geq 0} \otimes^p\varphi(u_p). \end{array}$$

Die gemischte Tensoralgebra

Es sei U ein K-Vektorraum. Für $p, q \geq 0$ definieren wir:

Definition. $\otimes^p_q U := \underbrace{U \otimes \ldots \otimes U}_{p} \otimes \underbrace{U^* \otimes \ldots \otimes U^*}_{q}$.

Definition. Die Elemente von $\otimes^p_q$ heißen *Tensoren der Stufe* (p, q).

Klassisch nennt man dies einen p-fach *kontravarianten*, q-fach *kovarianten* Tensor.

Bemerkung.

$$\begin{array}{lll} \otimes^p_0 U = \otimes^p U; & \otimes^1_0 U = U & \text{(„Vektoren“)} \\ \otimes^0_q U = \otimes^q U^*; & \otimes^0_1 U = U^* & \text{(„Linearformen“)} \\ & \otimes^0_0 U = K & \text{(„Skalare“)} \end{array}$$

Darstellung mittels einer Basis

Es sei $\dim U < \infty$ und $e_1, \dots, e_n$ sei eine Basis von U. Wir betrachten hierzu die duale Basis

$$e_1^*, \dots, e_n^*$$

von U^*.

Schreibweise.

(i) $e^i := (e_i)^*$.

(ii) $e_{i_1 \dots i_p}^{j_1 \dots j_q} := e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q}$.

Bemerkung.

(i) $(e_{i_1 \dots i_p}^{j_1 \dots j_q})_{\substack{1 \le i_k \le n \\ 1 \le j_l \le n}}$ ist Basis von $\otimes_q^p U$.

(ii) $\dim \otimes_q^p U = n^{p+q}$.

Jeder Tensor $u \in \otimes_q^p U$ besitzt eine *Darstellung*

$$u = \sum_{\substack{i_1 \dots i_p \\ j_1 \dots j_q}} \lambda_{j_1 \dots j_q}^{i_1 \dots i_p} e_{i_1 \dots i_p}^{j_1 \dots j_q} = \lambda_{j_1 \dots j_q}^{i_1 \dots i_p} e_{i_1 \dots i_p}^{j_1 \dots j_q}.$$

Transformationsverhalten

Sei

$$u = \lambda_{j_1 \dots j_q}^{i_1 \dots i_p} e_{i_1 \dots i_p}^{j_1 \dots j_q}.$$

Ferner sei ein Basiswechsel gegeben:

$$\tilde{e}_i = \sum_{j=1}^n a_i^j e_j = a_i^j e_j$$

$$\tilde{e}^i = \sum_{j=1}^n b_j^i e^j = b_j^i e^j.$$

Dabei ist $\tilde{e}^i$, die zu $\tilde{e}_i$ duale Basis, d. h. $\tilde{e}^i(\tilde{e}_j) = \delta_{ij}$. Für die Matrizen $A = (a_i^j)$ und $B = (b_j^i)$ bedeutet dies, daß $B = ({}^t A)^{-1}$. Dann gilt:

$$\tilde{\lambda}_{l_1 \dots l_q}^{k_1 \dots k_p} = a_{l_1}^{j_1} \dots a_{l_q}^{j_q} b_{i_1}^{k_1} \dots b_{i_p}^{k_p} \lambda_{j_1 \dots j_q}^{i_1 \dots i_p}.$$

Definition. Die direkte Summe

$$\otimes(U, U^*) := \bigoplus_{p,q \geq 0} \otimes^p_q U$$

heißt die *gemischte Tensoralgebra* über U.

Schreibweise.

$$\otimes(U, U^*) \ni u = \sum_{p,q} u^p_q, \; u^p_q \in \otimes^p_q U \qquad \text{(fast alle } u^p_q = 0\text{)}.$$

Um die Multiplikation einzuführen, stellen wir zunächst folgendes fest:
Es gibt einen Isomorphismus

$$(\otimes^r_s U) \otimes (\otimes^p_q U) \xrightarrow{\sim} \otimes^{r+p}_{s+q} U$$

mit

$$\begin{aligned}(u_1 \otimes \ldots \otimes u_r \otimes u^1 \otimes \ldots \otimes u^s) \otimes (v_1 \otimes \ldots \otimes v_p \otimes v^1 \otimes \ldots \otimes v^q) \longmapsto \\ (u_1 \otimes \ldots \otimes u_r \otimes v_1 \otimes \ldots \otimes v_p \otimes u^1 \otimes \ldots \otimes u^s \otimes v^1 \otimes \ldots \otimes v^q).\end{aligned}$$

Für $u \in \otimes^p_q U$, $v \in \otimes^r_s U$ kann man also $u \otimes v \in \otimes^{p+r}_{q+s} U$ auffassen.
Für

$$u = \sum_{p,q} u^p_q, \qquad v = \sum_{r,s} v^r_s$$

setzen wir dann

$$u \cdot v = \sum_{k,l} \Big(\sum_{\substack{p+r=k \\ q+s=l}} u^p_q \otimes v^r_s \Big).$$

Tensorverjüngung

Wir hatten definiert:

$$\otimes^p_q U = \underbrace{U \otimes \ldots \otimes U}_{p} \otimes \underbrace{U^* \otimes \ldots \otimes U^*}_{q}.$$

Wähle nun i, j fest mit $1 \leq i \leq p$, $1 \leq j \leq q$. Die Abbildung

$$\begin{array}{rccl} \gamma^j_i : & U \times \ldots \times U \times U^* \times \ldots \times U^* & \longrightarrow & \otimes^{p-1}_{q-1} U \\ & \begin{array}{c}(u_1, \ldots, u_i, \ldots, u_p, \\ u^1, \ldots, u^j, \ldots, u^q)\end{array} & \longmapsto & \begin{array}{l} u^j(u_i)(u_1 \otimes \ldots \otimes \hat{u}_i \otimes \ldots \otimes u_p \\ \qquad \otimes u^1 \otimes \ldots \otimes \hat{u}^j \otimes \ldots \otimes u^q) \end{array} \end{array}$$

ist multilinear, induziert also eine lineare Abbildung

$$\Gamma_i^j : \quad \otimes_q^p U \quad \longrightarrow \quad \otimes_{q-1}^{p-1} U$$

mit

$$\begin{aligned}\Gamma_i^j(u_1 \otimes \ldots \otimes u_p \otimes u^1 \otimes \ldots \otimes u^q) \quad &= \quad u^j(u_i) u_1 \otimes \ldots \otimes \hat{u}_i \otimes \ldots \otimes u_p \\ &\qquad \otimes u^1 \otimes \ldots \otimes \hat{u}^j \otimes \ldots \otimes u^q.\end{aligned}$$

Hierbei bedeutet die Schreibweise $\hat{u}$, daß das Element u in dem entsprechenden Tensorprodukt weggelassen wird.

Definition. Die Abbildung Γ_i^j heißt *Tensorverjüngung*.

Sei nun

$$u = \sum_{\substack{\nu_1 \ldots \nu_p \\ \mu_1 \ldots \mu_q}} \lambda_{\mu_1 \ldots \mu_q}^{\nu_1 \ldots \nu_p} e_{\nu_1 \ldots \nu_p}^{\mu_1 \ldots \mu_q}.$$

Dann gilt

$$\Gamma_i^j(u) = \sum_{\substack{r_1 \ldots r_{p-1} \\ s_1 \ldots s_{q-1}}} \mu_{s_1 \ldots s_{q-1}}^{r_1 \ldots r_{p-1}} e_{r_1 \ldots r_{p-1}}^{s_1 \ldots s_{q-1}}$$

mit

$$\mu_{s_1 \ldots s_{q-1}}^{r_1 \ldots r_{p-1}} = \sum_k \lambda_{s_1 \ldots s_{j-1} k s_j \ldots s_{p-1}}^{r_1 \ldots r_{i-1} k r_i \ldots r_{p-1}}.$$

Dies rechnet man leicht nach:

$$\begin{aligned}\Gamma_i^j(u) \quad &= \quad \Gamma_i^j\Big(\sum_{\substack{\nu_1 \ldots \nu_p \\ \mu_1 \ldots \mu_q}} \lambda_{\mu_1 \ldots \mu_j \ldots \mu_q}^{\nu_1 \ldots \nu_i \ldots \nu_p} \begin{array}{c} e_{\nu_1} \otimes \ldots \otimes e_{\nu_i} \otimes \ldots \otimes e_{\nu_p} \\ \otimes e^{\mu_1} \otimes \ldots \otimes e^{\mu_j} \otimes \ldots \otimes e^q \end{array} \Big) \\ &= \quad \sum_{\substack{\nu_1 \ldots \nu_p \\ \mu_1 \ldots \mu_q}} \underbrace{e^{\mu_j}(e_{\nu_i})}_{\delta_{\mu_j \nu_i}} \lambda_{\mu_1 \ldots \mu_j \ldots \mu_q}^{\nu_1 \ldots \nu_i \ldots \nu_p} \begin{array}{c} e_{\nu_1} \otimes \ldots \otimes \hat{e}_{\nu_i} \otimes \ldots \otimes e_{\nu_p} \\ \otimes e^{\mu_1} \otimes \ldots \otimes \hat{e}^{\mu_j} \otimes \ldots \otimes e^q \end{array} \\ &= \quad \sum_{\substack{r_1 \ldots r_{p-1} \\ s_1 \ldots s_{q-1}}} \mu_{s_1 \ldots s_{q-1}}^{r_1 \ldots r_{p-1}} e_{r_1 \ldots r_{p-1}}^{s_1 \ldots s_{q-1}}.\end{aligned}$$

Unter Verwendung der Summenkonvention schreibt sich diese Formel als

$$\Gamma_i^j(u) = \mu_{s_1 \ldots s_{q-1}}^{r_1 \ldots r_{p-1}} e_{r_1 \ldots r_{p-1}}^{s_1 \ldots s_{q-1}}.$$

Beispiel. $U = \mathbb{R}^2$. Es sei

$$\begin{aligned}u \quad &= \quad 2e_1 \otimes e^1 - e_2 \otimes e^1 \in \otimes_1^1 U \\ v \quad &= \quad 5e_1 \otimes e_1 \otimes e^1 - 2e_1 \otimes e_1 \otimes e^2 \in \otimes_1^2 U.\end{aligned}$$

Man hat

$$\begin{aligned}
w := u \cdot v &= 10e_1 \otimes e_1 \otimes e_1 \otimes e^1 \otimes e^1 \\
&\quad -4e_1 \otimes e_1 \otimes e_1 \otimes e^1 \otimes e^2 \\
&\quad -5e_2 \otimes e_1 \otimes e_1 \otimes e^1 \otimes e^1 \\
&\quad +2e_2 \otimes e_1 \otimes e_1 \otimes e^1 \otimes e^2 \\
&= 10e^{11}_{111} - 4e^{12}_{111} - 5e^{11}_{211} + 2e^{12}_{211} \in \otimes^3_2 U.
\end{aligned}$$

Ferner ist

$$\begin{aligned}
\Gamma^2_1(w) &= 10e_1 \otimes e_1 \otimes e^1 - 0 - 0 + 2e_1 \otimes e_1 \otimes e^1 \\
&= 12e^1_{11} \in \otimes^2_1 U. \\
\Gamma^1_1(\Gamma^2_1(w)) &= \Gamma^1_1(12e_1 \otimes e_1 \otimes e^1) = 12e_1 \in \otimes^1_0 U = U \\
\Gamma^1_2(\Gamma^2_1(w)) &= \Gamma^2_1(12e_1 \otimes e_1 \otimes e^1) = 12e_1 \in \otimes^1_0 U = U.
\end{aligned}$$

§ 16 Äußere Algebra

Für den Grundkörper K setzen wir in diesem Paragraphen voraus, daß er die Charakteristik 0 hat, d. h. daß

$$n_K := \underbrace{1 + \ldots + 1}_{n\text{-mal}} \neq 0 \text{ für alle } n \geq 1.$$

Wie üblich identifizieren wir dann n_K mit n. U, V seien K-Vektorräume.

Definition. Es sei

$$F: \quad \underbrace{U \times \ldots \times U}_{p} \longrightarrow V$$

eine Abbildung. Für jedes $\sigma \in S_p$ sei

$$\begin{aligned} \sigma F: \quad U \times \ldots \times U &\longrightarrow V \\ \sigma F(u_1, \ldots, u_p) &:= F(u_{\sigma(1)}, \ldots, u_{\sigma(p)}). \end{aligned}$$

Bemerkungen.

(i) $\mathrm{id} F = F$,

(ii) $\sigma'(\sigma F) = (\sigma'\sigma)F$.

Definition. Eine p-lineare Abbildung

$$F: \quad U \times \ldots \times U \longrightarrow V$$

heißt *alternierend* (*schiefsymmetrisch*), falls gilt

$$\sigma F = \operatorname{sign} \sigma F \qquad \text{für alle } \sigma \in S_p.$$

Beispiel. Wir hatten bereits früher die Determinantenfunktion eingeführt:

$$\begin{aligned} \det: \quad K^n \times \ldots \times K^n &\longrightarrow K \\ (u_1, \ldots, u_n) &\longmapsto \det(u_1, \ldots, u_n). \end{aligned}$$

Diese ist alternierend. Es gilt nämlich:

$$\sigma \det(u_1, \ldots, u_n) = \det(u_{\sigma(1)}, \ldots, u_{\sigma(n)}) = \operatorname{sign} \sigma \det(u_1, \ldots, u_n).$$

Also gilt

$$\sigma \det = \operatorname{sign} \sigma \det.$$

Lemma 16.1
Es sei

$$F: \quad \underbrace{U \times \ldots \times U}_{p} \longrightarrow V$$

eine p-lineare Abbildung. Dann sind die folgenden Aussagen äquivalent.

(i) F ist alternierend.

(ii) $\tau F = -F$ für jede Transposition $\tau \in S_p$,

(iii) Ist $u_i = u_j$ für ein $i \neq j$, so ist $F(u_1, \ldots, u_p) = 0$.

Beweis. (i)⇒(ii): Klar.

(ii)⇒(i): Es sei $\sigma \in S_p$. Dann gibt es eine Darstellung

$$\sigma = \tau_1 \cdots \tau_m$$

mit

$$m = \begin{cases} \text{gerade} & \text{falls } \operatorname{sign} \sigma = 1 \\ \text{ungerade} & \text{falls } \operatorname{sign} \sigma = -1. \end{cases}$$

Damit folgt die Behauptung aus der obigen Bemerkung (ii).

(i)⇒(iii): Es sei $u_i = u_j$ mit $i \neq j$. τ sei die Transposition, die i und j vertauscht. Dann gilt

$$\begin{aligned} F(u_1, \ldots, u_p) &= F(u_{\tau(1)}, \ldots, u_{\tau(p)}) = (\tau F)(u_1, \ldots, u_p) \\ &= -F(u_1, \ldots, u_p). \end{aligned}$$

Also gilt mit $1 + 1 \neq 0$, daß

$$F(u_1, \ldots, u_p) = 0.$$

(iii)⇒(ii): Es sei τ die Transposition, die i und j vertauscht ($i \neq j$). Dann gilt

$$\begin{aligned} 0 &= F(u_1, \ldots, u_i \overset{\overset{i}{\downarrow}}{+} u_j, u_{i+1}, \ldots, u_i \overset{\overset{j}{\downarrow}}{+} u_j, \ldots, u_p) \\ &= F(u_1, \ldots, u_i, \ldots, u_i, \ldots, u_p) + F(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_p) \\ &\quad + F(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_p) + F(u_1, \ldots, u_j, \ldots, u_j, \ldots, u_p). \end{aligned}$$

Also folgt

$$F(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_p) = -F(u_1, \ldots, u_j, \ldots, u_i, \ldots, u_p),$$

d. h.

$$\tau F = -F.$$

□

Definition. Es sei A ein K-Vektorraum, und

$$\bigwedge{}^p : \quad \underbrace{U \times \ldots \times U}_{p} \longrightarrow A$$

sei eine alternierende Abbildung. Dann heißt $(A, \bigwedge{}^p)$ *p-faches äußeres Produkt* von U, falls gilt:

(A1) Die Elemente $u_1 \wedge \ldots \wedge u_p := \bigwedge^p(u_1, \ldots, u_p)$ erzeugen A.

(A2) Ist W ein K-Vektorraum und

$$F : \quad U \times \ldots \times U \longrightarrow W$$

eine alternierende Abbildung, so gibt es eine lineare Abbildung $\hat{F} : A \to W$, so daß das folgende Diagramm kommutiert:

$$\begin{tikzcd} U \times \ldots \times U \arrow[rr, "F"] \arrow[d, "\bigwedge^p"'] & & W \\ A \arrow[urr, "\hat{F}"'] & & \end{tikzcd}$$

Bemerkung. Wegen (A1) ist $\hat{F}$ eindeutig bestimmt.

Theorem 16.2
Es gibt zu jedem Vektorraum ein p-faches äußeres Produkt $(A, \bigwedge^p)$. Dies ist bis auf Isomorphie eindeutig bestimmt, d. h. ist $(\tilde{A}, \widetilde{\bigwedge}^p)$ ein weiteres p-faches äußeres Produkt, so gibt es ein kommutatives Diagramm

$$\begin{tikzcd} & U \times \ldots \times U \arrow[dl, "\bigwedge^p"'] \arrow[dr, "\widetilde{\bigwedge}^p"] & \\ A \arrow[rr, "\sim"] & & \tilde{A}. \end{tikzcd}$$

Der Beweis wird später erfolgen.

Bezeichnungen. $\bigwedge{}^p U := A$.

Lemma 16.3
Es ist $u_1 \wedge \ldots \wedge u_p \neq 0$ genau dann, wenn $u_1, \ldots, u_p$ linear unabhängig sind.

Beweis. „$\Rightarrow$“: $u_1, \ldots, u_p$ seien linear abhängig. Ohne Einschränkung können wir dann annehmen:

$$u_1 = \sum_{i=2}^{p} \lambda_i u_i.$$

Also folgt

$$\begin{aligned} u_1 \wedge \ldots \wedge u_p &= (\sum_{i=2}^{p} \lambda_i u_i) \wedge u_2 \wedge \ldots \wedge u_p \\ &= \sum_{i=2}^{p} \lambda_i (u_i \wedge u_2 \wedge \ldots \wedge u_i \wedge \ldots \wedge u_p) \\ &= \sum_{i=2}^{p} \lambda_i \bigwedge\nolimits^p \underbrace{(u_i, u_2, \ldots, u_i, \ldots, u_p)}_{=0, \text{ da } \bigwedge^p \text{ alternierend}} = 0. \end{aligned}$$

„$\Leftarrow$“: $u_1, \ldots, u_p$ seien linear unabhängig.

$$U_1 := \operatorname{Span}(u_1, \ldots, u_p).$$

Wähle U_2 mit

$$U = U_1 \oplus U_2.$$

Dann besitzt jeder Vektor $u \in U$ eine eindeutige Darstellung $u = u^1 + u^2$ mit $u^1 \in U_1, u^2 \in U_2$. Nach Wahl einer Basis von U_1 können wir die Vektoren in U_1 als Spaltenvektoren auffassen. Wir definieren

$$\begin{array}{rccc} F: & U \times \ldots \times U & \longrightarrow & K \\ & (u_1^1 + u_1^2, \ldots, u_p^1 + u_p^2) & \longmapsto & \det(u_1^1, \ldots, u_p^1). \end{array}$$

Diese Abbildung ist alternierend mit

$$F(u_1, \ldots, u_p) = \det(u_1, \ldots, u_p) \neq 0.$$

Also gibt es ein kommutatives Diagramm

$$\begin{tikzcd} U \times \ldots \times U \arrow[r, "F"] \arrow[d] & K \\ A \arrow[ur, "\hat{F} \text{ linear}"'] & \end{tikzcd}$$

mit

$$\hat{F}(u_1 \wedge \ldots \wedge u_p) = F(u_1, \ldots, u_p) \neq 0,$$

und hieraus folgt sofort

$$u_1 \wedge \ldots \wedge u_p \neq 0.$$

$\square$

Es soll nun das äußere Produkt konstruiert werden.

Vorbereitungen zur Konstruktion des äußeren Produkts

Wir betrachten für festes $\sigma \in S_p$ die Abbildung

$$\begin{array}{rccc} \sigma: & \underbrace{U \times \ldots \times U}_{p\text{-fach}} & \longrightarrow & \otimes^p U \\ & (u_1, \ldots, u_p) & \longrightarrow & u_{\sigma(1)} \otimes \ldots \otimes u_{\sigma(p)}. \end{array}$$

Diese Abbildung ist p-linear, induziert also eine *lineare* Abbildung

$$\sigma: \quad \otimes^p U \longrightarrow \otimes^p(U)$$

mit

$$u_1 \otimes \ldots \otimes u_p \longmapsto u_{\sigma(1)} \otimes \ldots \otimes u_{\sigma(p)}.$$

Definition. Es sei

$$N^p(U) := \operatorname{Span}\{u \in \otimes^p U;\ u = u_1 \otimes \ldots \otimes u_p \text{ mit } u_i = u_j \text{ für ein } i \neq j\}.$$

Bemerkung. Der Unterraum $N^p(U) \subset \otimes^p U$ ist σ-invariant, d. h.

$$\sigma(N^p(U)) = N^p(U).$$

Lemma 16.4
Es sei $u \in \otimes^p U$. Dann gilt

$$u - \operatorname{sign} \sigma\, \sigma(u) \in N^p(U).$$

Beweis. Es genügt, dies für $u = u_1 \otimes \ldots \otimes u_p$ zu zeigen. Sei $\sigma = \tau_1 \cdots \tau_m$ wobei die τ_i Transpositionen sind. Wir machen Induktion nach m.

$m = 1$: τ vertausche i und j $(i < j)$.

$$\begin{aligned} u - \underbrace{\operatorname{sign}\tau}_{-1}\tau(u) &= u_1 \otimes \ldots \otimes u_i \otimes \ldots \otimes u_j \otimes \ldots \otimes u_p \\ &\quad + u_1 \otimes \ldots \otimes u_j \otimes \ldots \otimes u_i \otimes \ldots \otimes u_p \\ &= u_1 \otimes \ldots \otimes (u_i + u_j) \otimes \ldots \otimes (u_i + u_j) \otimes \ldots \otimes u_p \\ &\quad - u_1 \otimes \ldots \otimes u_i \otimes \ldots \otimes u_i \otimes \ldots \otimes u_p \\ &\quad - u_1 \otimes \ldots \otimes u_j \otimes \ldots \otimes u_j \otimes \ldots \otimes u_p, \end{aligned}$$

also

$$u - \operatorname{sign}\tau\,\tau(u) \in N^p(U).$$

$m - 1 \mapsto m$: $\sigma = \tau_1 \underbrace{\tau_2 \cdots \tau_m}_{=:\sigma'} = \tau_1\sigma'$. Es gilt

$$u - \operatorname{sign}\sigma'\,\sigma'(u) \overset{\text{(IV)}}{\in} N^p(U).$$

Da $N^p(U)$ invariant unter τ_1 ist, folgt

$$\tau_1(u) - \operatorname{sign}\sigma'\,\underbrace{\tau_1(\sigma'(u))}_{=\sigma(u)} \in N^p(U).$$

d. h.

$$\tau_1(u) - \operatorname{sign}\sigma'\,\sigma(u) \in N^p(U).$$

Also gilt

$$\operatorname{sign}\tau_1\,\tau_1(u) - \operatorname{sign}\tau_1\operatorname{sign}\sigma'\sigma(u) \in N^p(U),$$

d. h.

$$\operatorname{sign}\tau_1\,\tau_1(u) - \operatorname{sign}\sigma\,\sigma(u) \in N^p(U). \tag{1}$$

Da die Aussage für $m = 1$ gilt, folgt:

$$u - \operatorname{sign}\tau_1\,\tau_1(u) \in N^p(U). \tag{2}$$

Addition von (1) und (2) ergibt

$$u - \operatorname{sign}\sigma\,\sigma(u) \in N^p(U).$$

□

Definition. Ein Tensor $u \in \otimes^p U$ heißt *schiefsymmetrisch (alternierend)*, falls gilt:

$$\sigma(u) = \operatorname{sign}\sigma\, u \qquad (\text{für alle } \sigma \in S_p).$$

Definition. Der Unterraum

$$A^p(U) := \{u \in \otimes^p U;\ u \text{ ist schiefsymmetrisch}\} \subset \otimes^p U$$

heißt der Unterraum der *schiefsymmetrischen (alternierenden)* Tensoren.

Definition. Der *alternierende Operator* ist die lineare Abbildung

$$\begin{array}{rccc} \pi_A : & \otimes^p U & \longrightarrow & \otimes^p U \\ & \pi_A & := & \dfrac{1}{p!} \displaystyle\sum_{\sigma \in S_p} \operatorname{sign}\sigma\, \sigma. \end{array}$$

Beispiel. Sei $U = K^2$.

$$\begin{aligned} \pi_A(e_1 \otimes e_2) &= \frac{1}{2}(e_1 \otimes e_2 - e_2 \otimes e_1) \\ \pi_A(e_1 \otimes e_1) &= \frac{1}{2}(e_1 \otimes e_1 - e_1 \otimes e_1) = 0. \end{aligned}$$

Lemma 16.5
Für $\varrho \in S_p$ gilt:

(i) $\pi_A \varrho = \operatorname{sign}\varrho\, \pi_A$,

(ii) $\varrho \pi_A = \operatorname{sign}\varrho\, \pi_A$.

Beweis. Dies rechnet man direkt nach:

(i)

$$\begin{aligned} \pi_A \circ \varrho &= \Big(\frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sign}\sigma\, \sigma\Big)\varrho \\ &= \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sign}\sigma (\sigma\varrho) = \operatorname{sign}\varrho \frac{1}{p!} \sum_{\sigma \in S_p} \underbrace{(\operatorname{sign}\sigma)(\operatorname{sign}\varrho)}_{=\operatorname{sign}(\sigma\varrho)}(\sigma\varrho) \\ &\overset{\sigma' := \sigma\varrho}{=} \operatorname{sign}\varrho \frac{1}{p!} \sum_{\sigma' \in S_p} \operatorname{sign}\sigma'\, \sigma' = \operatorname{sign}\varrho\, \pi_A. \end{aligned}$$

(ii) Analog. □

Korollar 16.6
Die Abbildung π_A ist eine Projektion, d. h. es gilt $\pi_A^2 = \pi_A$.

Beweis. Auch dies rechnet man direkt nach:

$$\begin{aligned}\pi_A^2 &= \pi_A\Big(\frac{1}{p!}\sum_{\sigma\in S_p}\operatorname{sign}\sigma\,\sigma\Big) = \frac{1}{p!}\sum_{\sigma\in S_p}\operatorname{sign}\sigma(\pi_A\sigma)\\ &\overset{(16.5)(\mathrm{i})}{=} \frac{1}{p!}\sum_{\sigma\in S_p}\underbrace{(\operatorname{sign}\sigma)^2}_{1}\pi_A = \frac{p!}{p!}\pi_A = \pi_A.\end{aligned}$$

□

Lemma 16.7

(i) $\operatorname{Ker}\pi_A = N^p(U)$,

(ii) $\operatorname{Im}\pi_A = A^p(U)$.

Beweis. (i) „⊃“: Es sei

$$u = u_1 \otimes \ldots \otimes u_i \otimes \ldots \otimes u_i \otimes \ldots \otimes u_p.$$

τ sei die Transposition, die i und j vertauscht. Dann ist

$$\tau(u) = u. \tag{1}$$

Also gilt

$$\pi_A(u) \overset{(1)}{=} \pi_A\tau(u) \overset{(16.5)(\mathrm{i})}{=} \operatorname{sign}\tau\pi_A(u) = -\pi_A(u)$$

Daraus folgt

$$\pi_A(u) = -\pi_A(u)$$

und wegen $2 \neq 0$ impliziert dies $\pi_A(u) = 0$.

„⊂“: Es gilt

$$\pi_A(u) - u = \Big(\frac{1}{p!}\sum_\sigma \operatorname{sign}\sigma\,\sigma(u)\Big) - u = \frac{1}{p!}\Big(\sum_\sigma\underbrace{(\operatorname{sign}\sigma\,\sigma(u) - u}_{\in N^p(U)}\Big).$$

Damit gilt

$$\pi_A(u) - u \in N^p(U).$$

Ist also $\pi_A(u) = 0$, so folgt $u \in N^p(U)$.

(ii) „⊂“:

$$\varrho\pi_A(u) \overset{(16.5)(\mathrm{ii})}{=} \operatorname{sign}\varrho\pi_A(u),$$

also $\pi_A(u) \in A^p(U)$.

„⊃“: Sei $u \in A^p(U)$. Dann gilt

$$\pi_A(u) = \frac{1}{p!} \sum_{\sigma \in S_p} \underbrace{\operatorname{sign}\sigma\, \sigma(u)}_{=u} = \frac{1}{p!} \sum_{\sigma \in S_p} u = u.$$

D. h.

$$u = \pi_A(u) \in \operatorname{Im} \pi_A.$$

□

Korollar 16.8
$\otimes^p U = N^p(U) \oplus A^p(U)$.

Beweis. Für jede lineare Abbildung $\pi : V \to V$ mit $\pi^2 = \pi$ gilt

$$V = \operatorname{Im} \pi \oplus \operatorname{Ker} \pi.$$

□

Definition. Ist $u \in \otimes^p(U)$, so heißt $\pi_A(u) \in A^p(U)$ der *alternierende Teil* des Tensors u.

Es sei nun

$$F : \quad U \times \ldots \times U \longrightarrow W$$

p-linear. Dies induziert eine lineare Abbildung $\tilde{F} : \otimes^p U \to W$, so daß gilt:

$$\begin{tikzcd} U \times \ldots \times U \arrow[r, "F"] \arrow[d, "\otimes"'] & W \\ \otimes^p U \arrow[ur, "\tilde{F}"'] & \end{tikzcd}$$

ist kommutativ.

Lemma 16.9
Es sind äquivalent

(i) F ist alternierend.

(ii) $\tilde{F}|N^p(U) \equiv 0$.

Beweis. Es gilt:

$$\begin{aligned} F \text{ ist alternierend} \quad &\Leftrightarrow \quad F(u_1, \ldots, u_i, \ldots, u_j, \ldots, u_p) = 0 \text{ für alle} \\ & \quad u = (u_1, \ldots, u_i, \ldots, u_j, \ldots, u_p) \text{ mit } u_i = u_j \ (i \neq j) \\ &\Leftrightarrow \quad \tilde{F}(u_1 \otimes \ldots \otimes u_i \otimes \ldots \otimes u_j \otimes \ldots \otimes u_p) = 0 \text{ für alle} \\ & \quad \tilde{u} = u_1 \otimes \ldots \otimes u_i \otimes \ldots \otimes u_j \otimes \ldots \otimes u_p \text{ mit } u_i = u_j \ (i \neq j) \\ &\Leftrightarrow \quad \tilde{F}|N^p(U) \equiv 0. \end{aligned}$$

$\square$

Beweis von Theorem (16.2).

Eindeutigkeit: Dies zeigt man genau wie beim Beweis von Theorem (15.1).

Existenz: Wir setzen

$$A := \otimes^p U / N^p(U).$$

Ferner sei

$$\begin{array}{rrcll} \bigwedge^p : & U \times \ldots \times U & \longrightarrow & U \otimes \ldots \otimes U \xrightarrow{\pi} A \\ & (u_1, \ldots, u_p) & \longmapsto & u_1 \otimes \ldots \otimes u_p \longmapsto \pi(u_1 \otimes \ldots \otimes u_p). \end{array}$$

Es gilt

$$\operatorname{Ker} \pi = N^p(U).$$

Damit ist nach Lemma (16.9) die Abbildung $\bigwedge^p$ alternierend.

(A1): Die Elemente

$$u_1 \wedge \ldots \wedge u_p := \bigwedge^p(u_1, \ldots, u_p) = \pi(u_1 \otimes \ldots \otimes u_p)$$

erzeugen A, da die Elemente $u_1 \otimes \ldots \otimes u_p$ das Tensorprodukt $\otimes^p U$ erzeugen und π surjektiv ist.

(A2): Es sei

$$F : \quad U \times \ldots \times U \longrightarrow W$$

alternierend. Dann gilt für die induzierte lineare Abbildung

$$\tilde{F} : \quad \otimes^p U \longrightarrow W$$

nach Lemma (16.9), daß

$$N^p(U) \subset \operatorname{Ker} \tilde{F}.$$

Also gibt es eine lineare Abbildung

$$\hat{F} : \quad A = \quad \otimes^p U / N^p(U) \longrightarrow W$$

mit

$$\hat{F}(u_1 \wedge \ldots \wedge u_p) = \tilde{F}(u_1 \otimes \ldots \otimes u_p) = F(u_1, \ldots, u_p).$$

$\square$

Bemerkung. Die Projektionsabbildung $\pi : \otimes^p U \to A$ definiert einen Isomorphismus $A^p(U) \cong A = \bigwedge^p U$.

Basis von $\bigwedge^p U$

Es sei U ein n-dimensionaler K-Vektorraum mit Basis $e_1, \dots e_n$.

Satz 16.10
Die Tensoren $e_{i_1} \wedge \dots \wedge e_{i_p}$ mit $1 \leq i_1 < \dots < i_p \leq n$ bilden eine Basis von $\bigwedge^p U$.

Korollar 16.11
$\dim \bigwedge^p U = \binom{n}{p}$. Insbesondere ist $\bigwedge^p U = \{0\}$ für $p > n$.

Beweis. *Erzeugendensystem*: Die Tensoren

$$e_{i_1} \otimes \dots \otimes e_{i_p} \in \otimes^p U \qquad (1 \leq i_k \leq n;\ k = 1, \dots, p)$$

erzeugen $\otimes^p U$ und ihre Bilder in $\bigwedge^p U$ erzeugen diesen Raum. Da

$$e_{i_1} \wedge \dots \wedge e_{i_p} = \operatorname{sign} \sigma \, e_{\sigma(i_1)} \wedge \dots \wedge e_{\sigma(i_p)}$$

bzw.

$$e_{i_1} \wedge \dots \wedge e_{i_p} = 0 \quad \text{falls } e_{i_k} = e_{i_l},\ i_k \neq i_l$$

kann man sich auf

$$e_{i_1} \wedge \dots \wedge e_{i_p}; \quad 1 \leq i_1 < i_2 < \dots < i_p \leq n$$

beschränken.

Lineare Unabhängigkeit: Es genügt folgendes zu zeigen:

Behauptung. Für $i_1 < \dots < i_p$ gegeben, gibt es eine alternierende Abbildung

$$F: \quad U \times \dots \times U \longrightarrow K$$

mit

$$\begin{aligned} F(e_{\sigma(i_1)}, \dots, e_{\sigma(i_p)}) &= \operatorname{sign} \sigma, \\ F(e_{j_1}, \dots, e_{j_p}) &= 0 \text{ falls } j_1, \dots, j_p \text{ keine Permutation von } i_1, \dots, i_p \text{ ist.} \end{aligned}$$

Dies beweist man wie im Beweis von Lemma (16.3) mittels einer Determinanten gebildeten Funktion. □

Bemerkung. Jeder Tensor $u \in \bigwedge^p U$ hat damit eine Darstellung

$$u = \sum_{1 \leq i_1 < \dots < i_p \leq n} \lambda^{i_1 \dots i_p} e_{i_1} \wedge \dots \wedge e_{i_p}.$$

Beispiel. Für $U = \mathbb{R}^4$ erhalten wir für $\bigwedge^2 U$ die folgende Basis:

$$e_1 \wedge e_2, e_1 \wedge e_3, e_1 \wedge e_4, e_2 \wedge e_3, e_2 \wedge e_4, e_3 \wedge e_4.$$

Es ist $\dim \bigwedge^2 U = \binom{4}{2} = 6$.

Interpretation der Determinante

Die Determinante ist eine alternierende Abbildung

$$\det : \quad K^n \times \ldots \times K^n \longrightarrow K.$$

Also induziert dies eine lineare Abbildung

$$\widetilde{\det} : \quad \underset{\substack{\cong \\ K}}{\bigwedge^n K^n} \longrightarrow K.$$

Diese Abbildung ist festgelegt durch

(1) $$\widetilde{\det}(e_1 \wedge \ldots \wedge e_n) = 1.$$

Dies zeigt sowohl die Existenz als auch die Eindeutigkeit der Determinantenabbildung. Man kann also die Determinante über das folgende kommutative Diagramm definieren:

$$\begin{tikzcd} K^n \times \ldots \times K^n \arrow[r, "\det"] \arrow[d, "\bigwedge^n"'] & K \\ \bigwedge^n K^n \arrow[ur, "\widetilde{\det}"'] & \end{tikzcd}$$

D. h. man kann die Determinantenfunktion als

$$\det = \widetilde{\det} \circ \bigwedge^p$$

definieren, wobei $\widetilde{\det}$ durch (1) definiert wird.

Induzierte lineare Abbildungen

U, V seien Vektorräume, K sei ein Körper. Es sei $\varphi : U \to V$ eine lineare Abbildung.

Satz 16.12
Es sei $p \geq 2$. Dann gibt es genau eine lineare Abbildung

$$\bigwedge^p \varphi : \quad \bigwedge^p U \longrightarrow \bigwedge^p V$$

mit

$$\bigwedge^p \varphi(u_1 \wedge \ldots \wedge u_p) = (\varphi(u_1) \wedge \ldots \wedge \varphi(u_p)).$$

Bemerkungen.

(i) $\bigwedge^p(\mathrm{id}_U) = \mathrm{id}_{\bigwedge^p U}$.

(ii) $\bigwedge^p(\varphi \circ \psi) = \bigwedge^p(\varphi) \circ \bigwedge^p(\psi)$.

Satz 16.13
Es sei $\varphi : U \to V$ linear. Sei $\dim U = \dim V = n$. Dann sind äquivalent:

(i) φ ist ein Isomorphismus.

(ii) $\bigwedge^n \varphi \not\equiv 0$.

Beweis. Sei $e_1, \dots, e_n$ eine Basis von U. Dann ist

$$\begin{array}{rccc} \bigwedge^n \varphi : & \bigwedge^n U & \longrightarrow & \bigwedge^n V \\ & e_1 \wedge \ldots \wedge e_n & \longmapsto & \varphi(e_1) \wedge \ldots \wedge \varphi(e_n). \end{array}$$

Es gilt

$$\begin{array}{rcl} \varphi \text{ Isomorphismus} & \Leftrightarrow & \varphi \text{ surjektiv} \Leftrightarrow \varphi(e_1), \dots, \varphi(e_n) \text{ linear unabhängig} \\ & \overset{(16.3)}{\Leftrightarrow} & \varphi(e_1) \wedge \ldots \wedge \varphi(e_n) \neq 0 \Leftrightarrow \bigwedge^n \varphi \not\equiv 0. \end{array}$$

□

Bemerkung. Ist $U = V$, so ist $\bigwedge^n \varphi : \bigwedge^n V \to \bigwedge^n V$ als Endomorphismus eines 1-dimensionalen Vektorraums durch einen Skalar gegeben. Dies ist genau die Determinante des Automorphismus φ, d. h.

$$\bigwedge^n \varphi = \det \varphi.$$

Äußere Algebra

U sei ein Vektorraum über K.

Definition.

(i) $\bigwedge^0 U := K$, $\bigwedge^1 U := U$.

(ii)

$$\bigwedge U := \bigoplus_{p \geq 0} \bigwedge^p (U).$$

U heißt die *äußere Algebra* oder (*alternierende Algebra*) über U.

Bemerkung. $\dim U = n \Rightarrow \dim \bigwedge U = 2^n$.

Denn: $\dim \bigwedge^p U = \begin{pmatrix} n \\ p \end{pmatrix}$. Also gilt

$$\dim \bigwedge U = \sum_{p \geq 0} \begin{pmatrix} n \\ p \end{pmatrix} = 2^n.$$

Multiplikation

Wir hatten früher den folgenden Unterraum von $\otimes^p U$ definiert:

$$N^p(U) = \operatorname{Span}\{u;\ u = u_1 \otimes \ldots \otimes u_i \otimes \ldots \otimes u_j \otimes \ldots \otimes u_p \text{ mit } u_i = u_j\}.$$

Ferner lieferte die Quotientenabbildung

$$\pi^{(p)}: \quad \otimes^p U \longrightarrow \otimes^p U / N^p U = \bigwedge{}^p U.$$

einen Epimorphismus

$$\pi: \quad \otimes U \longrightarrow \bigwedge U.$$

Bemerkung.
(i) $N^p(U) \otimes (\otimes^q U)) \subset N^{p+q}(U)$.
(ii) $(\otimes^p U) \otimes (N^q(U)) \subset N^{p+q}(U)$.

Satz 16.14
Es gibt genau eine Multiplikation

$$\cdot: \quad \bigwedge U \times \bigwedge U \longrightarrow \bigwedge U$$

mit

$$\pi(u) \cdot \pi(v) = \pi(u \cdot v).$$

Bemerkung. $u \cdot v$ ist hierbei das Produkt bezüglich der Multiplikation in der Tensoralgebra $\otimes U$.

Beweis. Die Eindeutigkeit ist klar. Zur Wohldefiniertheit betrachten wir

$$u = \sum_{p \geq 0} u_p; \quad v = \sum_{p \geq 0} v_p; \quad u' = \sum_{p \geq 0} u'_p; \quad v' = \sum_{p \geq 0} v'_p.$$

Es gilt

$$\begin{array}{rcl} \pi(u) = \pi(u') & \Leftrightarrow & u_p = u'_p + \tilde{u}_p \text{ mit } \tilde{u}_p \in N^p(U) \\ \pi(v) = \pi(v') & \Leftrightarrow & v_p = v'_p + \tilde{v}_p \text{ mit } \tilde{v}_p \in N^p(U). \end{array}$$

Es ist

$$\begin{array}{rcl} u_p v_q & = & (u'_p + \tilde{u}_p)(v'_q + \tilde{v}_q) \\ & = & u'_p v'_p + \underbrace{\tilde{u}_p v'_q + u'_p \tilde{v}_q + \tilde{u}_p \tilde{v}_q}_{\in N^{p+q}(U)}. \end{array}$$

Also folgt

$$\pi(u \cdot v) = \pi(u' \cdot v').$$

□

Bemerkung. Für $u = u_1 \wedge \ldots \wedge u_p, \quad v = v_1 \wedge \ldots \wedge v_q$ gilt

$$u \cdot v = u_1 \wedge \ldots \wedge u_p \wedge v_1 \wedge \ldots \wedge v_q.$$

Man schreibt daher für die Multiplikation in $\bigwedge U$ meist

$$u \wedge v := u \cdot v.$$

Satz 16.15
Sei $u \in \bigwedge^p U$, $v \in \bigwedge^q U$. Dann gilt

$$u \wedge v = (-1)^{pq} v \wedge u.$$

Beweis. Wir müssen zeigen, daß für

$$\tilde{u} \in \otimes^p U, \quad \tilde{v} \in \otimes^q U$$

gilt

$$\tilde{u} \otimes \tilde{v} - (-1)^{pq} \tilde{v} \otimes \tilde{u} \in N^{p+q}(U).$$

Dazu betrachen wir die Permutation

$$\sigma : \quad (1, \ldots, p, p+1, \ldots, p+q) \longmapsto (p+1, \ldots, p+q, 1, \ldots, p).$$

Es ist

$$\operatorname{sign} \sigma = (-1)^{pq}.$$

Es gilt

$$\sigma(\tilde{u} \otimes \tilde{v}) = \tilde{v} \otimes \tilde{u}.$$

Also folgt

$$\tilde{u} \otimes \tilde{v} - (-1)^{pq} \tilde{v} \otimes \tilde{u} = \tilde{u} \otimes \tilde{v} - \operatorname{sign} \sigma \, \sigma(\tilde{u} \otimes \tilde{v}) \overset{(16.4)}{\in} N^{p+q}(U).$$

□

§ 17 Symmetrische Algebra

Es sei wieder K ein Körper der Charakteristik 0. Ferner seien U, V Vektorräume über dem Körper K.

Definition. Eine p-lineare Abbildung

$$F: \quad \underbrace{U \times \ldots \times U}_{p} \longrightarrow V$$

heißt *symmetrisch*, falls gilt

$$\sigma F = F \quad \text{für alle } \sigma \in S_p.$$

Bemerkung. F ist symmetrisch genau dann wenn $\tau F = F$ für jede Transposition τ.

Definition. Es sei S ein Vektorraum über K. Ferner sei

$$S^p: \quad U \times \ldots \times U \longrightarrow S$$

eine symmetrische Abbildung. Dann heißt (S, S^p) *p-faches symmetrisches Produkt* von U, falls gilt

(S1) Die Elemente $u_1 \cdots u_p := S^p(u_1, \ldots, u_p)$ erzeugen S.

(S2) Ist W ein Vektorraum und

$$F: \quad U \times \ldots \times U \longrightarrow W$$

eine symmetrische Abbildung, so gibt es eine lineare Abbildung $\check{F}: S \to W$, so daß das folgende Diagramm kommutiert

$$\begin{tikzcd}
U \times \ldots \times U \arrow[r, "F"] \arrow[d, "S^p"'] & W \\
S \arrow[ur, "\check{F}"'] &
\end{tikzcd}$$

Bemerkung. Wegen (S1) ist dann $\check{F}$ eindeutig bestimmt.

Theorem 17.1
Sei U ein Vektorraum. Dann gibt es zu U ein p-faches symmetrisches Produkt (S, S^p). Dies ist bis auf Isomorphie eindeutig bestimmt. D. h. ist $(\tilde{S}, \tilde{S}^p)$ ein weiteres p-faches symmetrisches Produkt, so gibt es einen Isomorphismus $\varphi: S \to \tilde{S}$, so daß das folgende Diagramm kommutiert:

$$
\begin{tikzcd}
 & U \times \ldots \times U \arrow[dl, "S^p"'] \arrow[dr, "\tilde{S}^p"] & \\
S \arrow[rr, "\varphi"] & & \tilde{S}.
\end{tikzcd}
$$

Wir werden den Beweis später geben.

Bezeichnungen.

(i) $S^pU := S$.

(ii) $u_1 \cdots u_p := S^p(u_1, \ldots , u_p)$.

Symmetrische Tensoren

Definition. Es sei

$$M^p(U) := \mathrm{Span}\{u - \tau u \in \otimes^p U;\ u \in \otimes^p U, \tau \in S_p \text{ ist Transposition}\}.$$

Lemma 17.2
$M^p(U)$ ist invariant unter S_p.

Beweis. Es genügt zu zeigen, daß $\tau' M^p(U) = M^p(U)$ für jede Transpostion τ'. Sei dazu $v := u - \tau u$. Dann gilt für eine Transposition τ':

$$\tau' v = \tau' u - \tau'\tau u = (\underbrace{\tau' u - u}_{\in M^p(U)}) + (\underbrace{u - \tau u}_{\in M^p(U)}) + (\underbrace{\tau u - \tau'\tau u}_{\in M^p(U)}) \in M^p(U).$$

□

Lemma 17.3
Sei $u \in \otimes^p U$. Dann ist

$$u - \sigma(u) \in M^p(U).$$

Beweis. Es genügt, dies für $u = u_1 \otimes \ldots \otimes u_p$ zu zeigen. Es sei $\sigma = \tau_1 \cdots \tau_m$, wobei die τ_i Transpositionen sind. Wir machen Induktion nach m.

$m = 1$: $u - \tau_1 u \in M^p(U)$ nach Definition von $M^p(U)$.

$m - 1 \mapsto m$: $\sigma = \tau_1 \underbrace{\tau_2 \cdots \tau_m}_{=:\sigma'} = \tau_1\sigma'$. Nun gilt:

$$u - \sigma u = u - \tau_1\sigma' u = \underbrace{u - \sigma' u}_{\in M^p(U)} + \underbrace{\sigma' u - \tau_1\sigma' u}_{\in M^p(U)} \in M^p(U).$$

□

Definition. Ein Tensor $u \in \otimes^p U$ heißt *symmetrisch*, falls

$$\sigma u = u \text{ für alle } \sigma \in S_p.$$

Definition. Der Unterraum

$$\mathrm{Sym}^p(U) := \{u \in \otimes^p U;\ u \text{ ist symmetrisch}\} \subset \otimes^p U$$

heißt der *Unterraum der symmetrischen Tensoren.*

Definition. Der *symmetrische Operator* ist die lineare Abbildung

$$\begin{array}{rccc} \pi_S : & \otimes^p U & \longrightarrow & \otimes^p U \\ & \pi_S & := & \dfrac{1}{p!} \displaystyle\sum_{\sigma \in S_p} \sigma. \end{array}$$

Beispiel. $U = K^2$.

$$\begin{array}{rccc} \pi_S : & \otimes^2 U & \longrightarrow & \otimes^2 U \\ & \pi_S & = & \dfrac{1}{2} \displaystyle\sum_{\sigma \in S_2} \sigma. \end{array}$$

Etwa

$$\begin{aligned} \pi_S(e_1 \otimes e_2) &= \frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1) \\ \pi_S(\frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1)) &= \frac{1}{4}(e_1 \otimes e_2 + e_2 \otimes e_1 + e_2 \otimes e_1 + e_1 \otimes e_2) \\ &= \frac{1}{2}(e_1 \otimes e_2 + e_2 \otimes e_1) \end{aligned}$$

Lemma 17.4
Für $\varrho \in S_p$ gilt

$$\pi_S \varrho = \varrho \pi_S = \pi_S.$$

Beweis.

$$\pi_S \varrho = \frac{1}{p!}(\sum_{\sigma \in S_p} \sigma)\varrho = \frac{1}{p!}\sum_{\sigma \in S_p}(\sigma\varrho) = \pi_S.$$

Analog zeigt man $\varrho \pi_S = \pi_S$. □

Korollar 17.5
Die Abbildung π_S ist eine Projektion, d. h. $\pi_S^2 = \pi_S$.

Beweis.

$$\begin{aligned}\pi_S^2 &= \pi_S\Big(\frac{1}{p!}\sum_{\sigma\in S_p}\sigma\Big) = \frac{1}{p!}\sum_{\sigma\in S_p}\underbrace{\pi_S\sigma}_{=\pi_S}\\ &= \frac{1}{p!}\sum_{\sigma\in S_p}\pi_S = \pi_S.\end{aligned}$$

□

Lemma 17.6
(i) $\operatorname{Ker}\pi_S = M^p(U)$.
(ii) $\operatorname{Im}\pi_S = \operatorname{Sym}^p(U)$.

Beweis. (i) „$\supset$“: $M^p(U)$ wird erzeugt von Elementen der Form $u-\tau u$. Es ist

$$\pi_S(u-\tau u) = \pi_S(u) - \underbrace{\pi_S\circ\tau}_{=\pi_S}(u) = 0.$$

„$\subset$“: Für $u\in\otimes^p U$ gilt mit Lemma (17.3):

$$u - \sigma u \in M^p(U).$$

Also gilt

$$u - \pi_S(u) = \frac{1}{p!}\sum_{\sigma\in S_p}(u-\sigma u)\in M^p(U).$$

Aus $\pi_S(u)=0$ folgt damit, daß $u\in M^p(U)$.
(ii) „$\subset$“: Dies folgt, da

$$\sigma\pi_S(u) \overset{(17.4)}{=} \pi_S(u) \text{ für alle } \sigma\in S_p.$$

„$\supset$“: Es sei $u\in\operatorname{Sym}^p(U)$. Dann gilt

$$\pi_S(u) = \frac{1}{p!}\sum_{\sigma\in S_p}\sigma(u) = \frac{1}{p!}\sum_{\sigma\in S_p}u = u.$$

Also ist $u\in\operatorname{Im}\pi_S$. □

Korollar 17.7
$\otimes^p U = M^p(U)\oplus\operatorname{Sym}^p(U)$.

Definition. Ist $u\in\otimes^p U$, so heißt der Tensor $\pi_S(u)\in\operatorname{Sym}^p(U)$ der *symmetrische Teil* des Tensors u.

Es sei nun wieder

$$F:\quad U\times\ldots\times U\longrightarrow W$$

p-linear. Es gibt damit ein kommutatives Diagramm

$$\begin{tikzcd} U \times \ldots \times U \arrow[r, "F"] \arrow[d] & W \\ \otimes^p U \arrow[ur, "\tilde{F}"'] & \end{tikzcd}$$

Lemma 17.8
Es sind äquivalent:

(i) F ist symmetrisch.

(ii) $\tilde{F}|M^p(U) \equiv 0$.

Beweis. Es gilt

$$\begin{aligned} F \text{ symmetrisch } &\Leftrightarrow F(u) = F(\sigma(u)) \text{ für alle } \sigma \in S_p;\ u = (u_1, \ldots, u_p) \in U \times \ldots \times U \\ &\Leftrightarrow F(u) = F(\tau(u)) \text{ für alle } u, \text{ alle Transpositionen } \tau \\ &\Leftrightarrow \tilde{F}(u) = \tilde{F}(\tau(u)) \text{ für alle } u \in \otimes^p U, \text{ alle } \tau \\ &\Leftrightarrow \tilde{F}(u - \tau u) = 0 \text{ für alle Transpositionen } \tau, \text{ alle } u \\ &\Leftrightarrow \tilde{F}|M^p(U) \equiv 0. \end{aligned}$$

□

Beweis von Theorem (17.1).

Eindeutigkeit: Wie früher.

Existenz: Wir setzen

$$S := \otimes^p U / M^p(U),$$

sowie

$$S^p : \quad U \times \ldots \times U \xrightarrow{\otimes} \otimes^p U \xrightarrow{\pi} S.$$

(S1): Die Elemente $S^p(u_1, \ldots, u_p)$ erzeugen S, da die Elemente $u_1 \otimes \ldots \otimes u_p$ den Raum $\otimes^p U$ erzeugen.

(S2): Es sei

$$F : \quad U \times \ldots \times U \longrightarrow W$$

symmetrisch. Dann gilt für die induzierte Abbildung

$$\tilde{F} : \quad \otimes^p U \longrightarrow W,$$

daß $M^p(U) \subset \operatorname{Ker} \tilde{F}$. Also gibt es eine lineare Abbildung

$$\check{F} : \quad S = \otimes^p U / M^p(U) \longrightarrow W$$

mit

$$\check{F}(S^p(u_1, \dots, u_p)) = \tilde{F}(u_1 \otimes \dots \otimes u_p) = F(u_1, \dots, u_p).$$

Dies leistet das Gewünschte. $\square$

Bemerkung. Die Projektion π definiert einen Isomorphismus $S^pU = S \cong \mathrm{Sym}^p(U)$.

Basis von S^pU

U sei ein endlich-dimensionaler Vektorraum mit Basis $e_1, \dots, e_n$.

Satz 17.9
Die Tensoren $e_{i_1} \cdots e_{i_p}$ mit $1 \leq i_1 \leq i_2 \leq \dots \leq i_p \leq n$ bilden eine Basis von S^pU.

Korollar 17.10
$\dim S^pU = \binom{n+p-1}{p}$.

Die symmetrische Algebra

Es sei U ein K-Vektorraum. Wie in den früheren Fällen definiert man eine Multiplikation auch für symmetrische Tensoren.

Definition.

(i) $S^0U := K$, $S^1U := U$.

(ii) $SU := \oplus_{p \geq 0} S^pU$.

Die Elemente $u \in SU$ sind also von der Form

$$u = \sum_{\text{endlich}} u_p, \quad u_p \in S^pU.$$

Wie im Fall der äußeren Algebra beweist man auch die folgende Aussage.

Satz 17.11

(i) Es gibt genau eine Multiplikationsabbildung

$$\cdot : SU \times SU \longrightarrow SU$$

mit der Eigenschaft

$$\pi(\tilde{u}) \cdot \pi(\tilde{v}) = \pi(\tilde{u} \otimes \tilde{v}) \quad (\tilde{u}, \tilde{v} \in \otimes U).$$

(ii) Die Multiplikation ist kommutativ, d. h. $u \cdot v = v \cdot u$.

Schließlich betrachten wir den Polynomring

$$K[x_1, \dots, x_n] := \{P;\ P = \sum_{\text{endlich}} a_{\nu_1,\dots,\nu_n} x_1^{\nu_1} \cdots x_n^{\nu_n}\}$$

in n Variablen über K. Auch $K[x_1, \dots, x_n]$ ist eine kommutative Algebra mit Eins.

Wir können schreiben

$$K[x_1, \dots, x_n] = \bigoplus_{p \geq 0} K^p[x_1, \dots, x_n]$$

mit

$$\begin{aligned} K^p[x_1, \dots, x_n] &= \{P;\ P \text{ homogen vom Grad } p\} \\ &= \{P;\ P = \sum a_{\nu_1,\dots,\nu_p} x_1^{\nu_1} \cdots x_n^{\nu_n};\ \sum_{i=1}^{n} \nu_i = p\}. \end{aligned}$$

$K^p[x_1, \dots, x_n]$ heißt der *Raum der homogenen Polynome vom Grad* p.

Satz 17.12
Es sei U ein endlich-dimensionaler K-Vektorraum der Dimension n. Dann gibt es einen Isomorphismus von Algebren

$$SU \cong K[x_1, \dots, x_n].$$

Beweis. Es sei $e_1, \dots, e_n$ Basis von U. Dann definiert man

$$SU \longrightarrow K[x_1, \dots, x_n]$$

durch

$$e_{i_1} \dots e_{i_p} \longmapsto x_{i_1} \cdots x_{i_p}.$$

□

Literaturverzeichnis

[Ar] M. Artin, *Algebra*, Birkhäuser Verlag, Basel, 1993.

[Fi1] G. Fischer, *Lineare Algebra*, Vieweg, Braunschweig, 1975.

[Fi2] G. Fischer, *Analytische Geometrie*, Vieweg, Braunschweig, 1983.

[Gr1] W. H. Greub, *Lineare Algebra*, Springer Verlag, Berlin, 1967.

[Gr2] W. H. Greub, *Multilinear Algebra*, Springer Verlag, Berlin, 1967.

[Koe] M. Koecher, *Lineare Algebra und Analytische Geometrie*, Springer Verlag, Berlin, 1992.

[KM] H.–J. Kowalsky, G.–O. Michler, *Lineare Algebra*, W. de Gruyter Verlag, Berlin, 1998 (Neubearbeitung von [Kow]).

[Kow] H.–J. Kowalsky, *Lineare Algebra*, W. de Gruyter Verlag, Berlin, 1970.

[Lo] F. Lorenz, *Lineare Algebra I und II*, BI-Verlag, Mannheim, 1982