Skript zur Vorlesung

Lineare Algebra II
Basiskurs

Sommersemester 2006
(dreistündig)

Prof. Dr. Annette Werner

Inhaltsverzeichnis

8 Die Jordan'sche Normalform

Wir haben im letzten Paragraphen gelernt, dass ein Endomorphismus $f : V \to V$ genau dann diagonalisierbar ist, wenn erstens das charakteristische Polynom von der Form

$$\chi_f(X) = (X - \alpha_1)^{m_1} \dots (X - \alpha_r)^{m_r}$$

mit $\alpha_i \in K$ ist und wenn zweitens

$$m_i = \dim V_{\alpha_i}$$

gilt.
Den Exponenten m_i nennt man auch die algebraische Vielfachheit von α_i, die Zahl $\dim V_{\alpha_i}$ (die Dimension des Eigenraums zu α_i) heißt auch geometrische Vielfachheit von α_i.
Ist $f : V \to V$ ein beliebiger Endomorphismus eines endlich-dimensionalen Vektorraums, so gilt

$$\dim V_{\alpha_i} \leq n_i,$$

d.h. es ist geometrische Vielfachheit $\leq$ algebraische Vielfachheit. (Beweisen Sie das !)
Wir wollen nun auf die Bedingung geometrische Vielfachheit = algebraische Vielfachheit verzichten und trotzdem noch eine möglichst einfache Koordinatenmatrix für f angeben.

Satz 8.1 Sei $f : V \to V$ ein Endomorphismus des n-dimensionalen K-Vektorraums V. Wir nehmen an, dass das charakteristische Polynom χ_f vollständig in Linearfaktoren über K zerfällt, d.h.

$$\chi_f(X) = (X - \alpha_1)^{m_1} \cdot \ldots \cdot (X - \alpha_r)^{m_1}$$

mit $\alpha_i \in K$. Dann gibt es eine Basis B von V, so dass die Koordinatenmatrix $A_{f,B,B}$ von f bezüglich B eine obere Dreiecksmatrix ist. Auf der Diagonalen von $A_{f,B,B}$ stehen die Eigenwerte $\alpha_1, \dots, \alpha_r$ mit ihren Vielfachheiten $m_1, \dots, m_r$ gezählt, d.h.

jedes α_i kommt m_i-mal vor:

$$A_{f,B,B} = \begin{pmatrix} \alpha_1 & & & & & & \\ & \ddots & & & & * & \\ & & \alpha_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_r & & \\ & 0 & & & & \ddots & \\ & & & & & & \alpha_r \\ & \underbrace{\hphantom{\alpha_1 \quad \ddots \quad \alpha_1}}_{m_1} & & & & \underbrace{\hphantom{\alpha_r \quad \ddots \quad \alpha_r}}_{m_r} & \end{pmatrix}$$

Beweis : mit Induktion nach $n = \dim V$. Für $n = 1$ ist nichts zu zeigen. Wir nehmen jetzt an, $n \geq 2$ und die Behauptung gelte für alle Vektorräume der Dimension $\leq n-1$. Es sei $b_1 \neq 0$ ein Eigenvektor zum Eigenwert α_1. Wir ergänzen b_1 zu einer Basis $C = (b_1, c_2, \ldots, c_n)$ von V. Die Koordinatenmatrix von f bezüglich C hat dann die Gestalt

$$A_{f,C,C} = \begin{pmatrix} \alpha_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

mit einer Matrix $A' \in K^{(n-1)\times(n-1)}$. Es sei $W := \langle c_2, \ldots, c_n \rangle \subset V$ der von $c_2, \ldots, c_n$ erzeugte Unterraum von V. Offenbar ist $\dim W = n-1$.
Es sei $g : W \to W$ die lineare Abbildung, deren Koordinatenmatrix bezüglich der Basis $(b_2, \ldots, b_n)$ gerade die Matrix A' ist. Dann gilt für alle $w \in W$

$$f(w) = \gamma(w) b_1 + g(w)$$

mit einem $\gamma(w) \in K$.
Nun gilt:

$$\begin{aligned} \chi_f(X) &= \det(XE_n - A_{f,C,C}) \\ &= (X - \alpha_1) \det(XE_n - A') \\ &= (X - \alpha_1)\chi_g(X). \end{aligned}$$

Also ist $\chi_g(X) = (X - \alpha_1)^{m_1 - 1}(X - \alpha_2)^{m_2} \cdot \ldots \cdot (X - \alpha_r)^{m_r}$, d.h. χ_g zerfällt vollständig in Linearfaktoren über K. Nach Induktionsvoraussetzung gibt es also eine Basis $B' = (b_2, \ldots, b_n)$ von W, so dass die Koordinatenmatrix von g bezüglich B' eine obere Dreiecksmatrix der Form

$$A_{g,B',B'} = \begin{pmatrix} \alpha_1 & & & & & & \\ & \ddots & & & & * & \\ & & \alpha_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_r & & \\ & 0 & & & & \ddots & \\ & & & & & & \alpha_r \\ & \underbrace{\qquad\qquad}_{m_1 - 1} & & & & \underbrace{\qquad\qquad}_{m_r} & \end{pmatrix}$$

ist. Wir setzen $B = (b_1, b_2, \dots, b_n)$. Dann gilt

$$A_{f,B,B} = \begin{pmatrix} \alpha_1 & * & \dots & * \\ 0 & & & \\ \vdots & A_{g,B',B'} & & \\ 0 & & & \end{pmatrix},$$

woraus die Behauptung folgt. ☐

Einen Endomorphismus, der bezüglich einer geeigneten Basis eine Koordinatenmatrix besitzt, die eine obere Dreiecksmatrix ist, nennt man auch trigonalisierbar. Satz 8.1 sagt also, dass f trigonalisierbar ist, falls χ_f vollständig in Linearfaktoren über K zerfällt.

Satz 8.2 Für eine Matrix $A \in K^{n \times n}$ sind äquivalent:

i) χ_A zerfällt vollständig in Linearfaktoren über K.

ii) A ist ähnlich zu einer oberen Dreiecksmatrix.

Beweis : i) ⇒ ii) folgt aus Satz 8.1, angewandt auf die lineare Abbildung $\lambda_A : K^n \to K^n$. (Führen Sie das aus!)
ii) ⇒ i). Nach Lemma 7.14 haben ähnliche Matrizen dasselbe charakteristische Polynom. Ist $D = \begin{pmatrix} d_1 & & * \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ eine obere Dreiecksmatrix, so gilt offenbar

$$\chi_D(X) = \det(XE_n - D) = \prod_{i=1}^{n} (X - d_i),$$

d.h. χ_D zerfällt vollständig in Linearfaktoren über K. ☐

Der Beweis von Satz 8.1 gibt uns auch ein Verfahren an die Hand, zu einer gegebenen Matrix eine ähnliche Matrix in oberer Dreiecksgestalt zu finden.
Beispiel:

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 2 & 2 & 3 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$$

Es ist

$$\begin{aligned}
\chi_A(X) &= \det \begin{pmatrix} X-2 & 1 & -1 \\ 1 & X-2 & 1 \\ -2 & -2 & X-3 \end{pmatrix} \\
&= (X-2)\det \begin{pmatrix} X-2 & 1 \\ -2 & X-3 \end{pmatrix} - \det \begin{pmatrix} 1 & 1 \\ -2 & X-3 \end{pmatrix} - \det \begin{pmatrix} 1 & X-2 \\ -2 & -2 \end{pmatrix} \\
&= (X-2)[(X-2)(X-3)+2] - (X-3) - 2 + 2 - 2(X-2) \\
&= (X-2)^2(X-3) - (X-3) \\
&= X^3 - 7X^2 + 15X - 9 \\
&= (X-1)(X-3)^2
\end{aligned}$$

Also zerfällt χ_A vollständig in Linearfaktoren über $\mathbb{Q}$, die Eigenwerte von A sind 1 und 3. Wir suchen zunächst einen Eigenvektor b_1 zum Eigenwert $\alpha_1 = 1$. Den finden wir, indem wir das lineare Gleichungssystem $(E_3 - A)x = 0$ lösen. Wir wählen $b_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Offenbar ist (b_1, e_2, e_3) eine Basis von $\mathbb{Q}^3$. Bezüglich dieser Basis hat $\lambda_A : \mathbb{Q}^3 \to \mathbb{Q}^3$ die Koordinatenmatrix

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 1 & 4 \end{pmatrix},$$

diese Matrix ist also konjugiert zu A. Wir müssen nun die durch $A' = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$ gegebene lineare Abbildung $\langle e_2, e_3 \rangle \to \langle e_2, e_3 \rangle$ untersuchen.
Die Matrix A' hat als einzigen Eigenwert 3. Lösen des linearen Gleichungssystems $(3E_2 - A')x = 0$ liefert den Eigenvektor $\binom{1}{-1}$ zum Eigenwert 3 der Matrix A'. Also ist $b_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^3$ ein Eigenvektor von $\lambda_{A'} : \langle e_2, e_3 \rangle \to \langle e_2, e_3 \rangle$ zum Eigenwert 3. So-

mit hat die Koordinatenmatrix von λ_A bezüglich der Basis (b_1, b_2, e_3) obere Dreiecksgestalt. Indem wir diese Koordinatenmatrix berechnen, erkennen wir, dass A ähnlich zu der Matrix $\begin{pmatrix} 1 & -2 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 3 \end{pmatrix}$ ist.

Definition 8.3 Ein Endomorphismus $f : V \to V$ eines n-dimensionalen K-Vektorraums heißt **nilpotent**, wenn es ein $k \in \mathbb{N}$ mit $f^k = \underbrace{f \circ \ldots \circ f}_{k-\text{mal}} = 0$ gibt.

Mit Hilfe von Satz 8.1 können wir nilpotente Endomorphismen folgendermaßen charakterisieren:

Proposition 8.4 Sei $f : V \to V$ ein Endomorphismus und $n = \dim V$. Dann sind äquivalent

i) f ist nilpotent

ii) $\chi_f(X) = X^n$

iii) Es gibt eine Basis B von V, so dass $A_{f,B,B}$ eine obere Dreiecksmatrix mit Nullen auf der Hauptdiagonalen ist, d.h. es gilt

$$A_{f,B,B} = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

iv) $f^n = 0$

Beweis : i) $\Rightarrow$ ii): Angenommen $f^k = 0$. Wir zeigen $\chi_f(X) = X^n$ mit Induktion nach n, indem wir ein ähnliches Argument verwenden wie im Beweis von Satz 8.1.

Ist $n = 1$, so muss $f = 0$ gelten, also gilt die Behauptung.

Für den Induktionsschluss nehmen wir an, die Behauptung gilt für alle nilpotenten Endomorphismen $f : W \to W$ mit dim $W < n$. Für jede Basis B von V ist wegen $f^k = 0$

$$A^k_{f,B,B} = 0,$$

also folgt $\det A_{f,B,B} = 0$.

Nach Lemma 7.9 ist 0 also ein Eigenwert von f. Es sei $b_1 \in V$ ein Eigenvektor zum Eigenwert 0. Wir ergänzen b_1 zu einer Basis $C = (b_1, c_2, \ldots, c_n)$ von V. Dann ist

$$A_{f,C,C} = \begin{pmatrix} 0 & * & \ldots & * \\ & & A' & \\ 0 & & & \end{pmatrix}.$$

Wie im Beweis von Satz 8.1 sei $W = \langle c_2, \ldots, c_n \rangle$ und $g : W \to W$ der Endomorphismus, dessen Koordinatenmatrix bezüglich der Basis $(c_2, \ldots, c_n)$ die Matrix A' ist.

Aus $f^k = 0$ folgt $A^k_{f,C,C} = 0$. Nun ist $A^k_{f,C,C} = \begin{pmatrix} 0 & * & \ldots & * \\ \vdots & & & \\ & & A'^k & \\ 0 & & & \end{pmatrix}$, also ist $A'^k = 0$.

Daraus folgt $g^k = 0$, d.h. g ist nilpotent. Nach Induktionsvoraussetzung ist also $\chi_g(X) = X^{n-1}$. Wie im Beweis von Satz 8.1 sieht man

$$\chi_f(X) = X\chi_g(X).$$

Also ist $\chi_f(X) = X^n$, damit gilt ii).
ii) $\Rightarrow$ iii). Da $\chi_f(X) = X^n$ vollständig in Linearfaktoren zerfällt und 0 die einzige Nullstelle von $\chi_f(X)$, also der einzige Eigenwert von f ist, folgt die Behauptung aus Satz 8.1.
iii) $\Rightarrow$ iv): Ist $A_{f,B,B} = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$, so rechnet man nach, dass $(A_{f,B,B})^n = 0$ gilt. (Führen Sie das aus!) Somit ist $f^n = 0$.
iv) $\Rightarrow$ i) ist klar. □

Wir wollen jetzt die obere Dreiecksgestalt der Koordinatenmatrix aus Satz 8.1 noch weiter vereinfachen. Dazu brauchen wir folgendes Lemma:

Lemma 8.5 Sei $f : V \to V, \dim V = n$, und α ein Eigenwert von f.
Wir setzen $g = \alpha\mathrm{id} - f : V \to V$. Wie immer bezeichne $g^j : V \to V$ die Verknüpfung $\underbrace{g \circ \ldots \circ g}_{j-\mathrm{mal}}$. Wir setzen

$$d := \min\{j : \mathrm{Kern}(g^j) = \mathrm{Kern}(g^{j+1})\}.$$

Für $U = \mathrm{Kern}(g^d)$ und $W = \mathrm{Bild}(g^d)$ gilt dann:

i) $g(U) \subset U$ und $g(W) \subset W$, d.h. U und W sind g-invariante Unterräume.

ii) Für die Einschränkungen $g|_U : U \to U$ und $g|_W : W \to W$ gilt:

$$(g|_U)^d = 0 \text{ und } \mathrm{Kern}(g|_W) = 0$$

iii) $V = U \oplus W$.

Beweis : Kern(g) $\subset$ Kern(g^2) $\subset$ Kern$(g^3)\dots$ ist eine aufsteigende Kette von Unterräumen von V. (Prüfen Sie das!) Da V endlich-dimensional ist, muss es ein j mit Kern(g^j) = Kern(g^{j+1}) geben. Die Zahl d ist also wohldefiniert.

i) Ist $u \in U =$ Kern(g^d), so gilt

$$g^d(g(u)) = g^{d+1}(u) = g(g^d(u)) = g(0) = 0$$

Also liegt $g(u)$ in Kern$(g^d) = U$, d.h. es ist $g(U) \subset U$. Jetzt wollen wir $g(W) \subset W$ zeigen. Nach der Dimensionsformel für lineare Abbildung gilt für alle $j \geq 1$:

$$\dim \text{ Kern}(g^j) + \dim \text{ Bild}(g^j) = \dim V = \dim \text{ Kern}(g^{j+1}) + \dim \text{ Bild}(g^{j+1}).$$

Also folgt aus $\dim$ Kern(g^d) $=$ $\dim$ Kern(g^{d+1}) auch $\dim$ Bild(g^d) $=$ $\dim$ Bild(g^{d+1}). Mit Bild (g^{d+1}) $\subset$ Bild(g^d) ergibt sich Bild(g^d) = Bild(g^{d+1}). Also folgt für $w = g^d(v) \in$ Bild$(g^d) = W$:

$$g(w) = g(g^d(v)) = g^{d+1}(v) \in \text{ Bild}g^{d+1} = \text{ Bild}g^d = W,$$

d.h. es gilt $g(W) \subset W$.

ii) Definitionsgemäß ist

$$(g_{|U})^d = (g^d)_{|U} = 0.$$

Ist $w \in W$ mit $g(w) = 0$, so gibt es ein $v \in V$ mit $w = g^d(v)$. Also ist $g^{d+1}(v) = 0$. Wegen Kern(g^d) = Kern(g^{d+1}) folgt $w = g^d(v) = 0$. Also ist Kern $(g|_W) = 0$.

iii) Wir zeigen zunächst mit Induktion nach k, dass Kern (g^d) = Kern(g^{d+k}) für alle $k \geq 1$ gilt. Für $k = 1$ ist die Behauptung klar nach Definition von d. Angenommen, es gilt Kern (g^d) = Kern(g^{d+k}). Wir haben die triviale Inklusion Kern(g^{d+k}) $\subset$ Kern(g^{d+k+1}) und müssen Kern(g^{d+k+1}) $\subset$ Kern(g^{d+k}) zeigen. Ist $v \in$ Kern(g^{d+k+1}), so folgt $0 = g^{d+k+1}(v) = g^{d+1}(g^k(v))$, d.h. $g^k(v) \in$ Kern(g^{d+1}) = Kern(g^d). Also ist $g^{d+k}(v) = g^d(g^k v) = 0$, was den Beweis abschließt.

Nun zeigen wir $U \oplus W = V$. Sei $u \in U \cap W$, d.h. $g^d(u) = 0$ und $u = g^d(v)$ für ein $v \in V$. Es folgt $g^{2d}(v) = 0$, d.h. $v \in$ Kern(g^{2d}) = Kern(g^d). Also ist $u = g^d(v) = 0$.

Die Räume U und V sind also unabhängig. Nach der Dimensionsformel für lineare Abbildungen ist

$$\dim V = \dim \text{ Kern}(g^d) + \dim \text{ Bild}(g^d) = \dim U + \dim W = \dim(U \oplus W),$$

also gilt $V = U \oplus W$.

□

Wir haben gezeigt, dass für einen diagonalisierbaren Endomorphismus $f : V \to V$ eines endlich-dimensionalen Vektorraums V gilt, dass V direkte Summe aller Eigenräume ist. Das können wir auf trigonalisierbare Endomorphismen verallgemeinern, wenn wir statt der Eigenräume die sogenannten Haupträume betrachten.

Definition 8.6 Sei $f : V \to V, \dim V = n$, und α ein Eigenwert von f. Dann heißt

$$H_\alpha := \{v \in V : (f - \alpha \mathrm{id})^k = 0 \text{ für ein } k \in \mathbb{N}\}$$

der **Hauptraum** von f zum Eigenwert α.

Satz 8.7 Sei $f : V \to V$, dim $V = n$. Das charakteristische Polynom von f zerfalle vollständig in Linearfaktoren über k, d.h. es ist

$$\chi_f(X) = (X - \alpha_1)^{m_1} \cdot \ldots \cdot (X - \alpha_r)^{m_r}$$

mit paarweise verschiedenen $\alpha_i \in K$.

Dann gilt:

i) $H_{\alpha_i} =$ Kern$(f - \alpha_i \mathrm{id})^{m_i}$ für alle i

ii) $f(H_{\alpha_i} \subset H_{\alpha_i}$ und $\dim H_{\alpha_i} = m_i$ für alle i.

iii) $V = \bigoplus\limits_{i=1}^{r} H_{\alpha_i}$.

Beweis : Wir setzen $g = f - \alpha_1 \mathrm{id} : V \to V$ und wenden Lemma 8.5 auf g an. Es gibt also ein $d \geq 1$ mit

$$\text{Kern}(g) \subsetneq \text{Kern}(g^2) \subsetneq \cdots \subsetneq \text{Kern}(g^d) = \text{Kern}(g^{d+1}) = \text{Kern}(g^{d+2}) = \ldots$$

Definitionsgemäß ist

$$H_{\alpha_1} = \bigcup_{k=1}^{\infty} \text{Kern}(g^k) = \text{Kern}(g^d).$$

Es sei $W =$ Bild(g^d). Nach Lemma 8.5 sind H_{α_1} und W jeweils g-invariante Unterräume mit $V = H_{\alpha_1} \oplus W$, so dass

$$(g_{|H_{\alpha_1}})^d = 0$$

gilt. Also ist $g_{|H_{\alpha_1}}$ nilpotent. Proposition 8.4 liefert $(g_{|H_{\alpha_1}})^{\dim H_{\alpha_1}} = 0$. Also ist $H_{\alpha_1} =$ Kern$(g^{\dim H_{\alpha_1}})$. Also folgt Kern$(g^d) =$ Kern$(g^{\dim H_{\alpha_1}})$. Ferner sind H_{α_1}

und W als g-invariante Unterräume auch f-invariant. Aus $V = H_{\alpha_1} \oplus W$ folgt also $\chi_f(X) = \chi_{f_{|H_{\alpha_1}}}(X)\chi_{f_{|W}}(X)$. Ist A eine Koordinatenmatrix von f, so ist $A - \alpha_1 E_n$ eine Koordinatenmatrix von g. Also folgt

$$\begin{aligned}\chi_f(X) &= \det(XE_n - A)\\ &= \det[(X-\alpha_1)E_n - (A - \alpha_1 E_n)]\\ &= \chi_g(X - \alpha_1).\end{aligned}$$

Eine analoge Gleichung gilt für $f_{|H_{\alpha_1}}$ und $f_{|W}$. Also folgt

$$\begin{aligned}\chi_f(X) &= \chi_{f_{|H_{\alpha_1}}}(X)\chi_{f_{|W}}(X)\\ &= \chi_{g_{|H_{\alpha_1}}}(X-\alpha_1)\chi_{f_{|W}}(X)\\ &\overset{8.4}{=} (X-\alpha_1)^{\dim H_{\alpha_1}}\chi_{f_{|W}}(W).\end{aligned}$$

Nach Lemma 8.5 ist Kern$(g_{|W}) = 0$, also ist $\chi_{g_{|W}}(0) \neq 0$, woraus $\chi_{f_{|W}}(\alpha_1) \neq 0$ folgt. Also muss dim$H_{\alpha_1} = m_1$ sein und $\chi_{f_{|W}}(X) = (X-\alpha_2)^{m_2} \cdot \ldots \cdot (X-\alpha_r)^{m_r}$ gelten. Damit sind i) und ii) für α_1 bewiesen. Genauso zeigt man diese Aussagen für die anderen α_i.

iii) zeigen wir mit Induktion nach $n = \dim V$. Die Behauptung ist klar für $n = 1$. Wir nehmen also an, für jeden Endomorphismus $h : W \to W$ eines Vektorraums W der Dimension $< n$ ist W die direkte Summe der Haupträume, falls $\chi_h(X)$ vollständig in Linearfaktoren zerfällt.

Für das gegebene f mit Eigenwert α_1 haben wir oben $V = H_{\alpha_1} \oplus W$ mit einem f-invarianten Vektorraum W gezeigt, für den $\chi_{f_{|W}}(X) = \prod\limits_{i=2}^{r}(X-\alpha_i)^{m_i}$ gilt. Wenden wir die Induktionsvoraussetzung auf $f_{|W} : W \to W$ an, so erhalten wir $W = \bigoplus\limits_{i=2}^{r} \text{Kern}(f_{|W} - \alpha_i \text{id})^{m_i}$.

Nun ist Kern $(f_{|W} - \alpha_i \text{id}_W)^{m_i} = \text{Kern}(f - \alpha_i\text{id})^{m_i} = V\alpha_i$ für $i \geq 2$, denn offenbar ist Kern $(f_{|W} - \alpha_i \text{id}_{|W})^{m_i} \subset \text{Kern}(f - \alpha_i \text{id})^{m_i}$, und die Dimensionen beider Räume stimmen nach ii) überein. Also folgt tatsächlich $V = \bigoplus\limits_{i=1}^{r} V\alpha_i$. □

Definition 8.8 Für alle $m \in \mathbb{N}$ definieren wir die $(m \times m)$-**Jordanmatrix** J_m als

$$J_m = \begin{pmatrix} 0 & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & 0\end{pmatrix} \in K^{m\times m}$$

Offenbar ist $J_m^m = 0$.

Satz 8.9 Es sei $g : V \to V$ nilpotent, $\dim V = n$. Wir setzen $d = \min\{j : g^j = 0\}$. Dann gibt es $s_1, \dots, s_d \in \mathbb{N}_0$ mit

$$ds_d + (d-1)s_{d-1} + \cdots + 2s_2 + s_1 = n$$

und eine Basis B von V mit

$$A_{g,B,B} = \begin{pmatrix} J_d & & & & & & 0 \\ & \ddots & & & & & \\ & & J_d & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s_d\text{-mal} \\ \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s_1\text{-mal} \end{matrix}$$

Die Zahlen $s_1, \dots, s_d$ sind durch diese Bedingungen eindeutig bestimmt.

Beweis : Wir setzen $U_j =$ Kern(g^j). Dann ist $\{0\} = U_0 \subset U_1 \subset U_2 \subset \ldots \subset U_d = V$ eine aufsteigende Kette von Unterräumen von V. Wir zeigen zunächst $g(U_j) \subset U_{j-1}$. Sei $u \in U_j =$ Kern(g^j). Dann ist $g^{j-1}(g(u)) = g^j(u) = 0$, also $g(U) \subset$ Kern$g^{j-1} = U_{j-1}$.
Für jeden Unterraum W von V mit $W \cap U_j = 0$ für ein $j \geq 1$ folgt wegen Kern$g = U_1 \subset U_j$ auch $W \cap$ Kern$g = 0$, d.h. Kern$g_{|W} = 0$. Wir wählen nun ein Komplement W_d von U_{d-1} in V, d.h. $V = U_d = U_{d-1} \oplus W_d$. Dann ist $g(W_d) \subset g(U_d) \subset U_{d-1}, g(U_{d-1}) \subset U_{d-2}$ und Kern$(g_{|W_d}) = 0$.
Ferner gilt $U_{d-2} \cap g(W_d) = 0$, denn ist $v \in U_{d-2} \cap g(W_d)$, so ist $g^{d-2}(v) = 0$ und $v = g(w)$ für ein $w \in W_d$. Dann ist $g^{d-1}(w) = 0$, also $w \in U_{d-1} \cap W_d = 0$, woraus $v = g(w) = 0$ folgt.
Also finden wir einen Unterraum W_{d-1} von U_{d-1} mit

$$g(W_d) \subset W_{d-1} \text{ und } U_{d-1} = U_{d-2} \oplus W_{d-1}.$$

Wie oben folgt $g(U_{d-2}) \subset U_{d-3}$, $g(W_{d-1}) \subset U_{d-2}$ sowie $U_{d-3} \cap g(W_{d-1}) = 0$.
Dieses Verfahren wenden wir solange an, bis alle U_j verbraucht sind. Wir erhalten eine Zerlegung

$$V = W_1 \oplus \cdots \oplus W_d$$

mit $g(W_j) \subset W_{j-1}$ und Kern$(g_{|W_j}) = 0$ für alle j.
Nun wählen wir eine Basis $w_1^{(d)}, \dots, w_{s_d}^{(d)}$ von W_d. Wegen $g(W_d) \subset W_{d-1}$ und Kern$(g_{|W_d}) = 0$ sind $g(w_1^{(d)}), \dots, g(w_{s_d}^{(d)})$ linear unabhängig in W_{d-1}. Wir ergänzen

sie zu einer Basis

$$(g(w_1^{(d)}), \dots, g(w_{s_d}^{(d)}), w_1^{(d-1)}, \dots, w_{s_{d-1}}^{(d-1)})$$

von W_{d-1}. (Hier ist $s_d = \dim W_d$ und $s_{d-1} = \dim W_{d-1} - \dim W_d$). Analog können wir die Bilder unter g dieser Basisvektoren, die in W_{d-2} liegen, zu einer Basis

$$(g^2(w_1^{(d)}), \dots, g^2(w_{s_d}^{(d)}), g(w_1^{(d-1)}), \dots, g(w_{s_{d-1}}^{(d-1)}), w_1^{(d-2)}, \dots, w_{s_{d-2}}^{(d-2)})$$

von W_{d-2} ergänzen.

Dieses Verfahren führen wir so lange durch, bis wir Basen für alle W_j gefunden haben. Insgesamt erhalten wir eine Basis von $V = W_1 \oplus \cdots \oplus W_d$ der Form

$$\begin{aligned} B \;=\; & (g^{d-1}(w_1^{(d)}), g^{d-2}(w_1^{(d)}), \dots, g(w_1^{(d)}), w_1^{(d)}, \dots \\ & g^{d-1}(w_{s_d}^{(d)}), g^{d-2}(w_{s_d}^{(d)}), \dots, g(w_{s_d}^{(d)}), w_{s_d}^{(d)}, \\ & g^{d-2}(w_1^{(d-1)}), g^{d-1}(w_1^{(d-1)}), \dots, g(w_1^{(d-1)}), w_1^{(d-1)}), \dots \\ & g^{d-2}(w_{s_{d-1}}^{(d-1)}), g^{d-1}(w_{s_{d-1}}^{(d-1)}), \dots, g(w_{s_{d-1}}^{(d-1)}), w_{s_{d-1}}^{(d-1)}, \\ & \dots, \\ & w_1^{(1)}, \dots, w_{s_1}^{(1)}) \end{aligned}$$

Die Koordinatenmatrix $A_{g,B,B}$ hat dann die gewünschte Form

$$A_{g,B,B} = \left(\begin{array}{cccccccccc} J_d & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & J_d & & & & & & & \\ & & & J_{d-1} & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & J_{d-1} & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & J_1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & J_1 \\ \multicolumn{3}{c}{\underbrace{\hspace{3cm}}_{s_d}} & \multicolumn{3}{c}{\underbrace{\hspace{3cm}}_{s_{d-1}}} & & \multicolumn{3}{c}{\underbrace{\hspace{3cm}}_{s_1}} \end{array}\right)$$

Da $A_{g,B,B}$ eine $(n \times n)$-Matrix ist, folgt $n = ds_d + (d-1)s_{d-1} + \cdots + s_1$.

Jetzt wollen wir noch die Eindeutigkeit von $s_1, \dots, s_d$ zeigen. Es ist

$$J_k^2 = \begin{pmatrix} 0 & & \\ \vdots & J_{k-1} & \\ 0 & \dots & 0 \end{pmatrix} \text{ und } \dim \mathrm{Kern} J_k = 1.$$

Also folgt

$$\begin{aligned}
\dim \mathrm{Kern}(g) &= \dim\ \mathrm{Kern}\ A_{g,B,B} = s_1 + \cdots + s_d \\
\dim \mathrm{Kern}(g^2) &= \dim\ \mathrm{Kern}\ A^2_{g,B,B} = s_1 + 2s_2 + 2s_3 + \cdots + 2s_d \\
\dim \mathrm{Kern}(g^3) &= \dim\ \mathrm{Kern} A^3_{g,B,B} = s_1 + 2s_2 + 3s_3 + \cdots + 3s_d \\
\ldots \\
\dim \mathrm{Kern}(g^{d-1}) &= \dim \mathrm{Kern} A^{d-1}_{g,B,B} = s_1 + 2s_2 + \cdots + (d-1)s_{d-1} + (d-1)s_d \\
\dim \mathrm{Kern}(g^d) &= n.
\end{aligned}$$

Also folgt für $k = 2, \ldots, d$:

$$\dim\ \mathrm{Kern}\ g^k - \dim \mathrm{Kern} g^{k-1} = s_k + \cdots + s_d.$$

Durch diese Formeln sind offenbar $s_2, \ldots, s_d$ eindeutig durch g bestimmt. Wegen $s_1 + 2s_2 + \cdots + ds_d = n$ ist auch s_1 eindeutig durch g bestimmt. ☐

Satz 8.9 besagt, dass der nilpotente Endomorphismus bis auf Äquivalenz eindeutig durch die Zahlen $s_1, \ldots, s_d$ festgelegt („klassifiziert“) wird. Aus dem Existenzbeweis in Satz 8.9 folgt insbesondere, dass $s_d = \dim W_d = n - \dim U_{d-1} \geq 1$ ist, da $U_{d-1} \subsetneq V$ gilt.

Für kleine Dimensionen können wir aus Satz 8.9 folgende Schlüsse ziehen: Ist $g : V \to V$ nilpotent mit $\dim V = 2$, so gilt entweder $g = 0$ oder ($g^2 = 0$ und $g \neq 0$). Ist $g^2 = 0$ und $g \neq 0$, so ist $d = 2$. Wegen $s_2 \geq 1$ und $2s_2 + s_1 = 2$ folgt schon $s_2 = 1$ und $s_1 = 0$. Die Matrix aus Satz 8.9 hat dann also die Gestalt

$$J_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Ist $g : V \to V$ nilpotent mit $\dim V = 3$, $g \neq 0$, so ist entweder $g^2 = 0$ oder ($g^3 = 0$ und $g^2 \neq 0$). Im ersten Fall $g^2 = 0$ ist $d = 2$. Aus $s_2 \geq 1$ und $2s_2 + s_1 = 3$ folgt schon $s_2 = 1, s_1 = 1$. Die Matrix aus Satz 8.9 hat also die Gestalt

$$J_3 = \begin{pmatrix} J_2 & \\ & J_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Im zweiten Fall $g^3 = 0, g^2 \neq 0$ ist $d = 3$. Aus $s_3 \geq 1$ und $3s_3 + 2s_2 + s_1 = 3$ folgt schon $s_3 = 1, s_1 = s_2 = 0$. Die Matrix aus Satz 8.9 hat also die Gestalt

$$J_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Satz 8.10 (Jordansche Normalform) Sei $f : V \to V$ ein Endomorphismus des n-dimensionalen K-Vektorraums V, so dass $\chi_f(X) = (X-\alpha_1)^{m_1} \cdot \ldots \cdot (X-\alpha_r)^{m_r}$ über K vollständig in Linearfaktoren zerfällt. Hier seien die $\alpha_1, \ldots, \alpha_r$ paarweise verschieden. Dann existiert eine Basis B von V, so dass $A_{f,B,B}$ folgende Gestalt hat:

$$A_{f,B,B} = \begin{pmatrix} \alpha_1 E_{m_1} + N_1 & & \\ & \alpha_2 E_{m_2} + N_2 & 0 \\ & & \ddots \\ 0 & & \alpha_r E_{m_r} + N_r \end{pmatrix} \tag{1}$$

mit nilpotenten Matrizen $N_i \in K^{m_i \times m_i}$ der Form

$$N_i = \begin{pmatrix} J_{d_i} & & & & & & 0 \\ & \ddots & & & & & \\ & & J_{d_i} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s^{(i)}_{d_i}\text{-mal} \\ \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s^{(i)}_1\text{-mal,} \end{array}$$

wobei $s^{(i)}_1, \ldots, s^{(i)}_{d_i} \in \mathbb{N}_0$, $s^{(i)}_{d_i} \geq 1$ Zahlen mit $d_i s^{(i)}_{d_i} + (d_i - 1)s^{(i)}_{d_i-1} + \cdots + 2s^{(i)}_2 + s^{(i)}_1 = m_i$ sind.

Die $s^{(i)}_j, j = 1, \ldots, d_i$ sind durch f eindeutig bestimmt. Die eindeutig bestimmte Koordinatenmatrix $A_{f,B,B}$ der Form (1) heißt Jordansche Normalform von f.

Beweis : Nach Satz 8.7 ist $V = \bigoplus\limits_{i=1}^{r} H_{\alpha_i}$, wobei H_{α_i} den f-invarianten Hauptraum zu α_i bezeichnet. Auf H_{α_i} ist die lineare Abbildung $g_i = f - \alpha_i \mathrm{id}$ nilpotent. Wir wenden Satz 8.9 auf $g_{i|H_{\alpha_i}}$ an. Für

$$d_i = \min\{k : (f - \alpha_i \mathrm{id})^k_{|H_{\alpha_i}} = 0\}$$

gilt also: Es gibt eine Basis B_i von H_{α_i} mit

$$A_{g_{i|H_{\alpha_i}} B_i, B_i} = \begin{pmatrix} J_{d_i} & & & & & & 0 \\ & \ddots & & & & & \\ & & J_{d_i} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & & & & \ddots & \\ 0 & & & & & & J_1 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s^{(i)}_{d_i} \\ \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s^{(i)}_1 \end{array}$$

für eindeutig bestimmte Zahlen $s_1^{(i)}, \dots, s_{d_i}^{(i)} \in \mathbb{N}_0$ mit $d_i s_{d_i}^{(i)} + \dots + 2s_2^{(i)} + s_1^{(i)} = \dim H_{\alpha_i} = m_i$. Nun ist

$$A_{f|H_{\alpha_i}, B_i B_i} = \alpha_i E_{m_i} + A_{g_i|H_{\alpha_i}}.$$

Also liefert die Basis $B = (B_1, \dots, B_r)$ eine Koordinatenmatrix der gewünschten Gestalt. □

Für eine $n \times n$-Matrix A, so dass $\chi_A(X)$ vollständig in Linearfaktoren über K zerfällt, gibt es nach Satz 8.10 eine Matrix der Form

$$\tilde{A} = \begin{pmatrix} \alpha_1 E_{d_1} + J_{d_1} & & & & & & \\ & \ddots & & & & 0 & \\ & & \alpha_1 E_1 + J_1 & & & & \\ & & & \ddots & & & \\ & & & & \alpha_r E_{d_r} + J_{d_r} & & \\ & 0 & & & & \ddots & \\ & & & & & & \alpha_r E_1 + J_1 \end{pmatrix} \begin{array}{l} \} \, s_{d_1}^{(1)}\text{-mal} \\ \\ \} \, s_1^{(1)}\text{-mal} \\ \\ \} \, s_{d_r}^{(r)}\text{-mal} \\ \\ \} \, s_1^{(r)}\text{-mal} \end{array}$$

konjugiert zu A. Die Matrix $\tilde{A}$ ist eindeutig bestimmt bis auf die Reihenfolge der $\alpha_1, \dots, \alpha_r$. $\tilde{A}$ heißt **Jordansche Normalform** zu A. Die Matrizen der Form $\alpha_i E_k + J_k$ auf der Hauptdiagonalen von $\tilde{A}$ heißen **Jordankästchen** zu α_i.

Beispiel: Es sei $A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$.

Dann ist

$$\chi_A(X) = \det \begin{pmatrix} X-3 & -4 & -3 \\ 1 & X & 1 \\ -1 & -2 & X-3 \end{pmatrix}$$

Also ist $\alpha_1 = 2$ der einzige Eigenwert von A. Wir betrachten

$$A - 2E_3 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$$

und berechnen

$$A - 2E_3 \neq 0, (A - 2E_3)^2 \neq 0, (A - 2E_3)^3 = 0.$$

Also ist $d_1 = 3$, woraus $s_3^{(1)} = 1, s_2^{(1)} = s_1^{(1)} = 0$ folgt. Somit ist

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

die Jordan'sche Normalform zu A. Sie besteht aus genau einem Jordankästchen.
Wir wollen nun noch die Jordan'sche Normalform benutzen, um das Minimalpolynom eines Endomorphismus zu definieren. Ist $p(X) = a_m X^m + \cdots + a_1 X + a_0$ ein beliebiges Polynom mit Koeffizienten $a_i \in K$, so können wir einen Endomorphismus $f : V \to V$ des K-Vektorraums V in p einsetzen: Es sei $p(f) : V \to V$ der Endomorphismus

$$p(f) = a_m f^m + \cdots + a_1 f + a_0 \cdot \mathrm{id},$$

wobei natürlich wieder $f^k = \underbrace{f \circ \cdots \circ f}_{k-\text{mal}}$ gilt.

Ist $A \in K^{n \times n}$, so definieren wir

$$p(A) = a_m A^m + \cdots + a_1 A + a_0 E_n \in K^{n \times n}.$$

Ist B eine Basis von V und $A = A_{f,B,B}$, dann gilt $p(A) = A_{p(f),B,B}$. (Prüfen Sie das!)
Für ein Polynom $p(X) = a_m X^m + \cdots + a_1 X + a_0$ mit $a_m \neq 0$ heißt $m = \mathrm{grad}(p)$ der Grad von p. Ist $a_m = 1$, so nennt man p normiert. Sind alle $a_i = 0$, d.h. gilt $p(X) = 0$, so setzen wir $\mathrm{grad}(p) = -\infty$.

Satz 8.11 Es sei $f : V \to V$ ein Endomorphismus V des n-dimensionalen K-Vektorraums V, so dass $\chi_f(X) = (X - \alpha_1)^{m_1} \cdot \ldots \cdot (X - \alpha_r)^{m_r}$ über K vollständig in Linearfaktoren zerfällt, wobei $\alpha_1, \ldots, \alpha_r$ paarweise verschieden sind. Für jedes α_i sei d_i die Größe des größten Jordankästchens zu α_i in der Jordanschen Normalform. Das Polynom

$$p_f(X) = (X - \alpha_1)^{d_1} \cdot \ldots \cdot (X - \alpha_r)^{d_r}$$

hat dann folgende Eigenschaften:

i) $p_f(f) = 0$

ii) Ist $u(X) \neq 0$ ein Polynom mit $u(f) = 0$, so ist u ein Vielfaches von p_f, d.h. es gilt $u(X) = p_f(X)v(X)$ für ein Polynom $v(X)$. Das Polynom p_f heißt **Minimalpolynom** von f.

Beweis : Wir zerlegen wie im Beweis von Satz 8.10 $V = \bigoplus\limits_{i=1}^{r} H_{\alpha_i}$. Dann gibt es eine Basis B_i von H_{α_i} mit

$$A_{f-\alpha_i \mathrm{id}_{|H_{\alpha_i}},B_i,B_i} = \begin{pmatrix} \underbrace{\begin{matrix} J_{d_i} & & \\ & \ddots & \\ & & J_{d_i} \end{matrix}}_{s^{(i)}_{d_i}} & & \\ & \ddots & \\ 0 & & \underbrace{\begin{matrix} J_1 & & \\ & \ddots & \\ & & J_1 \end{matrix}}_{s^{(i)}_{1}} \end{pmatrix}$$

Für alle Vektoren $w \in H_{\alpha_i}$ gilt somit $(f - \alpha_i \mathrm{id})^{d_i}(w) = 0$, da $J_i^{d_i} = 0$ ist. Somit ist für alle Vektoren $w \in H_{\alpha_i}$ auch

$$p(f)(w) = (f - \alpha_1 \mathrm{id})^{d_1} \circ \cdots \circ (f - \alpha_r \mathrm{id})^{d_r}(w) = 0$$

Da der Endomorphismus $p(f)$ somit auf allen H_{α_i} verschwindet, gilt $p(f) = 0$.
Nun sei $u(X) \neq 0$ ein Polynom mit $v(f) = 0$. Wir wollen zeigen, dass $u(X)$ ein Vielfaches von $p_f(X)$ ist. Dazu betrachten wir zunächst den Spezialfall, dass f nilpotent auf V ist. Wir setzen $d = \min\{j : f^j = 0\}$. Dann gilt für das Polynom $w(X) = X^d$ ebenfalls $w(f) = 0$. Wir zeigen nun, dass $\mathrm{grad}(u) \geq d = \mathrm{grad}(w)$ gilt. Angenommen, $\mathrm{grad}(u) < d$. Dann wählen wir unter allen Polynomen $\tilde{u} \neq 0$ mit $\tilde{u}(f) = 0$ eines mit minimalem Grad aus. Also ist auch $\mathrm{grad}(\tilde{u}) < d$. Nun benutzen wir, dass es für Polynome eine Division mit Rest gibt. (Das wurde im Aufbaukurs LA I, § 6 bewiesen). Division von $w(X) = X^d$ durch $\tilde{u}$ mit Rest liefert Polynome $\tilde{v}(X), \tilde{r}(X)$ mit $\mathrm{grad}(\tilde{r}) < \mathrm{grad}(\tilde{u})$ und $w(X) = X^d = \tilde{u}(X)\tilde{v}(X) + \tilde{r}(X)$.
Aus $w(f) = 0$ und $\tilde{u}(f) = 0$ folgt $\tilde{r}(f) = 0$. Da $\tilde{u}$ minimal gewählt wurde, muss $\tilde{r} = 0$ sein. Also ist $\tilde{u}(X) = X^{\tilde{d}}$ für ein $\tilde{d} \leq d$, d.h. es gilt $f^{\tilde{d}} = \tilde{u}(f) = 0$. Nach Definition von d folgt daraus $d \leq \tilde{d}$, was $\tilde{d} = \mathrm{grad}(u) < d$ widerspricht. Also gilt in der Tat $\mathrm{grad}(u) \geq d$. Daher können wir $u(X)$ mit Rest durch $w(X) = X^d$ dividieren und erhalten

$$u(X) = X^d v(X) + r(X)$$

mit Polynomen $v(X), r(X)$, so dass $\mathrm{grad}\ r(X) < d$ gilt.
Aus $u(f) = 0$ und $f^d = 0$ folgt durch Einsetzen von f wieder $r(f) = 0$. Da es kein Polynom $\neq 0$ vom Grad $< d$ mit dieser Eigenschaft gibt, ist $r = 0$. Somit folgt $u(X) = X^d v(X)$, d.h. $u(X)$ ist ein Vielfaches von X^d.

Damit haben wir die Behauptung im Spezialfall f nilpotent bewiesen. Jetzt können wir den allgemeinen Fall mit Induktion nach n zeigen. Für $n = 0$ ist sie klar. Wir nehmen an, sie gilt für alle Vektorräume der Dimension $< n$. Nach Lemma 8.5 ist $V = H_{\alpha_1} \oplus W$ mit einem f-invarianten Unterraum W von f, für den Kern $(f_{|W} - \alpha_1 \mathrm{id}_{|W}) = 0$ gilt. Also ist Bild$(f_{|W} - \alpha_1 \mathrm{id}_{|W}) = W$. Aus $u(f) = 0$ folgt $u(f_{|H_{\alpha_1}}) = 0$ und $u(f_{|W}) = 0$. Wir setzen $\hat{u}(X) = u(X + \alpha_1)$. Dann ist $\hat{u}(f_{|H_{\alpha_1}} - \alpha_1 \mathrm{id}_{|H_{\alpha_1}}) = u(f_{|H_{\alpha_1}}) = 0$. Da $f_{|H_{\alpha_1}} - \alpha_1 \mathrm{id}_{|H_{\alpha_1}}$ nilpotent ist, folgt nach dem schon behandelten Fall der nilpotenten Endomorphismen

$$\hat{u}(X) = X^d v_0(X)$$

für ein Polynom v_0 und die Zahl $d = \min\{j : (f_{|H_{\alpha_1}} - \alpha_1 \mathrm{id}_{|H\alpha_1})^j = 0\}$. Also ist $d = d_1$ die Größe des größten Jordankästchens zu α_1 aus der Jordan'schen Normalform. Es folgt

$$u(X) = \hat{u}(X - \alpha_1) = (X - \alpha_1)^{d_1} v_0(X - \alpha_1) = v_1(X)(X - \alpha_1)^{d_1}$$

für das Polynom $v_1(X) = v_0(X - \alpha_1)$.
Nun ist Bild$(f_{|W} - \alpha_1 \mathrm{id}_{|W}) = W$, also auch Bild$(f_{|W} - \alpha_1 \mathrm{id}_{|W})^{d_1} = W$. Aus $u(f_{|W}) = 0$ folgt $v_1(f_{|W}) \circ (f_{|W} - \alpha_1 \mathrm{id}_{|W})^{d_1} = 0$. Nun können wir jedes $w \in W$ als $w = (f - \alpha_1 \mathrm{id})^{d_1}(w')$ für ein $w' \in W$ schreiben, woraus

$$v_1(f_{|W})(w) = 0$$

folgt. Somit gilt $v_1(f_{|W}) = 0$. Wir können daher die Induktionsvoraussetzung auf $f_{|W} : W \to W$ anwenden und erhalten

$$v_1(X) = (X - \alpha_2)^{d_2} \cdot \ldots \cdot (X - \alpha_r)^{d_r} v(X)$$

für ein Polynom $v(X)$. Daraus folgt die Behauptung. □

Beispiel: Hat $f : \mathbb{R}^5 \to \mathbb{R}^5$ die Jordan'sche Normalform

$$A_{f,B,B} = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix},$$

so ist $p_f(X) = (X - 3)^3(X + 1)$.

Definitionsgemäß ist das Minimalpolynom eines Endomorphismus $f : V \to V$, für den $\chi_f(X)$ vollständig in Linearfaktoren zerfällt, ein Teiler des charakteristischen Polynoms $\chi_f(X)$. Also folgt

Korollar 8.12 (Satz von Cayley-Hamilton) Sei $f : V \to V$ ein Endomorphismus, so dass $\chi_f(X)$ vollständig in Linearfaktoren zerfällt. Dann gilt

$$\chi_f(f) = 0$$

Beweis : Es ist $\chi_f(X) = v(X) \cdot p_f(X)$ für ein Polynom $v(X)$. Also ist $\chi_f(f) = v(f) \circ p_f(f) = 0$ nach Satz 8.11. □

Der Satz von Cayley-Hamilton gilt auch ohne die Voraussetzung, dass $\chi_f(X)$ vollständig in Linearfaktoren zerfällt (Siehe Aufbaukurs, LA I, § 6.) Zum Abschluss dieses Kapitels wollen wir noch auf die Voraussetzung, dass $\chi_f(X)$ vollständig in Linearfaktoren zerfällt, eingehen.

Definition 8.13 Ein Körper K heißt algebraisch abgeschlossen, falls jedes Polynom mit Koeffizienten in K eine Nullstelle in K hat.

Fundamentalsatz der Algebra: $\mathbb{C}$ ist algebraisch abgeschlossen.

Lemma 8.14 Ist K algebraisch abgeschlossen und f ein Polynom mit Koeffizienten in K, so gilt $f(X) = c(X - \alpha_1) \cdot \ldots \cdot (X - \alpha_n)$ mit $c, \alpha_1, \ldots, \alpha_n \in K$.

Beweis : mit Induktion nach grad(f). Ist grad(f) ≤ 1, so ist nichts zu zeigen. Für beliebiges f von Grad ≥ 2 sei α_1 eine Nullstelle von f. Wir dividieren $f(X)$ mit Rest durch $X - \alpha_1$ und erhalten

$$f(X) = (X - \alpha_1)g(X) + r(X),$$

wobei grad(r) < 1 ist, d.h. $r(X) = c \in K$. Durch Einsetzen von α_1 folgt $c = 0$, d.h. $f(X) = (X - \alpha_1)g(X)$. Durch Anwenden der Induktionsvoraussetzung auf g folgt die Behauptung. □

Ist der Grundkörper K also algebraisch abgeschlossen (etwa $K = \mathbb{C}$), so gelten die Ergebnisse auf § 8 für beliebige Endomorphismen endlich-dimensionaler K-Vektorräume. In der Algebra lernt man, dass es für jeden Körper K einen algebraisch abgeschlossenen Erweiterungskörper $\overline{K}$ gibt.

9 Bilinearformen

Wir beginnen gleich mit der Definition:

Definition 9.1 Es seien U, V, W $\quad K$-Vektorräume. Eine Abbildung

$$f : V \times W \to U$$

heißt **bilinear**, falls folgende Bedingungen erfüllt sind:

i) $f(v_1 + v_2, w) = f(v_1, w) + f(v_2, w)$ für alle $v_1, v_2 \in V$ und $w \in W$

ii) $f(v, w_1 + w_2) = f(v, w_1) + f(v, w_2)$ für alle $v \in V$ und $w_1, w_2 \in W$.

iii) $f(\alpha v, w) = \alpha f(v, w) = f(v, \alpha w)$ für alle $\alpha \in K, v \in V$ und $w \in W$.

Beispiel:

i) Die Produktabbildung

$$\begin{array}{rcl} K^{l\times m} \times K^{m\times n} & \to & K^{l\times n} \\ (A, B) & \mapsto & AB \end{array}$$

ist bilinear.

ii) Es sei $A \in K^{m\times n}$ eine beliebige $(m \times n)$-Matrix. Dann definiert A eine bilineare Abbildung

$$\beta_A : K^m \times K^n \longrightarrow K$$

durch $\beta_A(v, w) = v^t A w$.

Ist etwa $A = E_n \in K^{n\times n}$, so ist die zugehörige bilineare Abbildung gerade

$$\begin{array}{rcl} \beta_{E_n} : K^n \times K^n & \longrightarrow & K \\ \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) & \longmapsto & x_1 y_1 + x_2 y_2 + \cdots + x_n y_n. \end{array}$$

Für $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{2\times 3}$ ergibt sich die bilineare Abbildung

$$\begin{array}{rcl} \beta_A : \mathbb{Q}^2 \times \mathbb{Q}^3 & \longrightarrow & \mathbb{Q} \\ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) & \mapsto & (x_1, x_2) \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ & = & (x_1 + x_2 \;\; 2x_1 \;\; 3x_1 + x_2) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ & = & x_1 y_1 + x_2 y_1 + 2x_1 y_2 + 3x_1 y_3 + x_2 y_3 \end{array}$$

Lemma 9.2 Seien U, V, W K-Vektorräume und $f : V \times W \to U$ eine Abbildung. Dann ist f genau dann bilinear, falls für alle $v \in V$ die Abbildung

$$\begin{aligned} f(v, -) : W &\to U \\ w &\mapsto f(v, w) \end{aligned}$$

linear ist und falls für alle $w \in W$ die Abbildung

$$\begin{aligned} f(-, w) : V &\to U \\ v &\mapsto f(v, w) \end{aligned}$$

linear ist.

Beweis : Das folgt sofort aus Definition 9.1. ☐

Ab jetzt betrachten wir immer den Spezialfall $U = K$. Eine bilineare Abbildung

$$f : V \times W \to K$$

lässt sich folgendermaßen mit einer Matrix beschreiben.

Definition 9.3 Es sei $f : V \times W \to K$ bilinear mit endlich-dimensionalen Vektorräumen V und W. Ferner sei $B = (b_1, \ldots, b_m)$ eine Basis von V und $C = (c_1, \ldots, c_n)$ eine Basis von W. Dann heißt die $(m \times n)$-Matrix

$$M_{f,B,C} = (f(b_i, c_j))_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

Koordinatenmatrix von f bezüglich B und C. Manchmal wird $M_{f,B,C}$ auch **Gram'sche Matrix** genannt. Der Eintrag von $M_{f,B,C}$ an der Stelle (i, j) ist also gerade $f(b_i, c_j)$.
Die Basis B definiert einen Isomorphismus

$$\begin{aligned} \kappa_B : V &\to K^m \\ v &\mapsto \kappa_b(v), \end{aligned}$$

wobei die Einträge β_i von $\kappa_B(v) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$ durch

$$v = \beta_1 b_1 + \cdots + \beta_m b_m$$

definiert sind. Analog definiert die Basis C einen Isomorphismus $\kappa_C : W \to K^n$.

Lemma 9.4 Es sei $f : V \times W \to K$ bilinear und $B = (b_1, \dots, b_m)$ eine Basis von V sowie $C = (c_1, \dots, c_n)$ eine Basis von W. Dann gilt für alle $v \in V$ und $w \in W$:

$$f(v, w) = \kappa_b(v)^t M_{f,B,C} \kappa_c(w).$$

Beweis : Seien b_i und c_j Vektoren aus den Basen B und C. Dann gilt $\kappa_B(b_i) = e_i$ und $\kappa_C(c_j) = e_i$, wobei $e_i \in K^m$ und $e_j \in K^n$ die entsprechenden Einheitsvektoren sind. Also ist

$$\kappa_B(b_i)^t M_{f,B,C} \kappa_C(c_j) = e_i^t M_{f,B,C} e_j$$

gerade der Eintrag von $M_{f,B,C}$ an der Stelle (i, j), d.h. gleich $f(b_i, c_j)$. Sind $v \in V$ und $w \in W$ beliebig, so schreiben wir $v = \sum\limits_{i=1}^{m} \beta_i b_i$ und $w = \sum\limits_{j=1}^{n} \gamma_j c_j$ mit $\beta_i, \gamma_j \in K$. Also ist $\kappa_B(v) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ und $\kappa_C(w) = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}$. Da f bilinear ist, folgt

$$\begin{aligned}
f(v, w) &= f\left(\sum_{i=1}^{m} \beta_i b_i, \sum_{j=1}^{n} \gamma_j c_j\right) \\
&= \sum_{i=1}^{m} \beta_i f(b_i, \sum_{j=1}^{n} \gamma_j c_j) \\
&= \sum_{i=1}^{m} \beta_i \sum_{j=1}^{n} \gamma_j f(b_i, c_j) \\
&= \sum_{i=1}^{m} \beta_i \sum_{j=1}^{n} \gamma_j (e_i^t M_{f,B,C} e_j) \\
&= \sum_{i=1}^{m} \beta_i (e_i^t M_{f,B,C} \sum_{j=1}^{n} \gamma_j e_j) \\
&= \left(\sum_{i=1}^{m} \beta_i e_i^t\right) M_{f,B,C} \left(\sum_{j=1}^{n} \gamma_j e_j\right) \\
&= \kappa_B(v)^t M_{f,B,C} \kappa_c(w).
\end{aligned}$$

□

Dieses Resultat sagt, dass jede bilineare Abbildung $f : V \times W \to K$ nach Wahl von Basen B von V und C von W durch die Koordinatenmatrix beschrieben werden kann.

Wir können Lemma 9.4 auch so ausdrücken: Das Diagramm

$$\begin{tikzcd}
V \times W \arrow[dr, "f"] \arrow[dd, "\kappa_B \times \kappa_C"'] & \\
& K \\
K^m \times K^n \arrow[ur, "\beta_{M_{f,B,B}}"] &
\end{tikzcd}$$

ist kommutativ. Hier ist $\beta_{M_{f,B,C}}$ die durch $M_{f,B,C}$ induzierte bilineare Abbildung aus dem obigen Beispiel.

Jetzt wollen wir untersuchen, wie sich die Koordinatenmatrix einer bilinearen Abbildung ändert, wenn man zu anderen Basen übergeht.

Proposition 9.5 Es sei $f : V \times W \to K$ bilinear und dim$V = n$, dim $W = m$. Ferner seien Basen B und B' von V sowie C und C' von W gegeben.
Wir bezeichnen (wie immer) mit $P^B_{B'}$ bzw. $P^C_{C'}$ die Übergangsmatrix von B nach B' bzw. von C nach C'. Dann gilt

$$M_{f,B',C'} = (P^{B'}_B)^t M_{f,B,C} P^{C'}_C$$

Beweis : Es sei $B' = (b'_1, \dots, b'_m)$ und $C' = (c'_1, \dots, c'_n)$. Dann hat $P^{B'}_B$ die Spalten $(\kappa_B(b'_1), \dots, \kappa_B(b'_m))$, die Matrix $P^{C'}_C$ hat die Spalten $(\kappa_C(c'_1), \dots, \kappa_C(c'_n))$. Definitionsgemäß ist der Eintrag von $M_{f,B',C'}$ an der Stelle (i, j) gerade

$$\begin{aligned}
f(b'_i, c'_j) &\overset{9.4}{=} \kappa_B(b'_i)^t \; M_{f,B,C} \; \kappa_C(c'_j) \\
&= (i\text{-te Zeile von } (P^{B'}_B)^t) M_{f,B,C} (j\text{-te-Spalte von } P^{C'}_C) \\
&= \text{ Eintrag an der Stelle } (i, j) \text{ in der Matrix } (P^{B'}_B)^t M_{f,B,C} P^{C'}_C .
\end{aligned}$$

Also ist in der Tat $M_{f,B',C'} = (P^{B'}_B)^t M_{f,B,C} P^{C'}_C$. ☐

Nach ... ist jede invertierbare $(m \times m)$-Matrix die Übergangsmatrix von B zu einer anderen Basis von V und jede invertierbare $(n \times n)$-Matrix die Übergangsmatrix von C zu einer anderen Basis von W. Wir können Proposition 9.5 also auch so formulieren: Zwei Matrizen $M, N \in K^{m \times n}$ beschreiben genau dann dieselbe bilineare Abbildung (bezüglich verschiedener Basen), wenn Matrizen $P \in \mathrm{GL}(m, K)$ und $Q \in \mathrm{GL}(n, K)$ existieren mit

$$P^t M Q = N$$

Definition 9.6

i) Eine bilineare Abbildung $f : V \times W \to K$ heißt **nicht ausgeartet in der ersten Variablen**, falls aus
$$f(v, w) = 0 \text{ für alle } w \in W$$
schon $v = 0$ folgt.

ii) $f : V \times W \to K$ heißt **nicht ausgeartet in der zweiten Variablen**, falls aus
$$f(v, w) = 0 \text{ für alle } v \in V$$
schon $w = 0$ folgt.

Beispiel: Ist $A = (1 \;\; 2) \in \mathbb{R}^{1\times 2}$, so ist

$$\begin{array}{rcl} \beta_A : \mathbb{R} \times \mathbb{R}^2 & \to & \mathbb{R} \\ (x, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) & \mapsto & xy_1 + 2xy_2 \end{array}$$

nicht ausgeartet in der ersten Variablen, aber ausgeartet in der zweiten Variablen, denn $\beta_A(x, \begin{pmatrix} -2 \\ 1 \end{pmatrix}) = 0$ für alle $x \in \mathbb{R}$.

Ab jetzt betrachten wir nur noch den Fall $V = W$. Eine bilineare Abbildung
$$f : V \times V \to K$$
nennt man auch **Bilinearform auf *V***.

Definition 9.7 Sei $f : V \times V \to K$ eine Bilinearform.

i) f heißt **symmetrisch**, falls für alle $v, w \in V$ gilt $f(v, w) = f(w, v)$.

ii) f heißt **schiefsymmetrisch** oder alternierend, falls für alle $v \in V$ gilt $f(v, v) = 0$.

Im Fall char$(K) \neq 2$ lassen sich schiefsymmetrische Bilinearformen auch anders charakterisieren.

Lemma 9.8 Sei char$K \neq 2$. Eine Bilinearform $f : V \times V \to K$ ist genau dann schiefsymmetrisch, wenn für alle $v, w \in V$
$$f(v, w) = -f(w, v)$$
gilt.

Beweis : Ist f schiefsymmetrisch, so gilt für alle $v, w \in V$:

$$0 = f(v+w, v+w) = f(v,v) + f(v,w) + f(w,v) + f(w,w) = f(v,w) + f(w,v),$$

also $f(v,w) = -f(w,v)$. (Dieses Argument funktioniert auch für $\mathrm{char}(K) = 2$.)
Ist umgekehrt $f(v,w) = -f(w,v)$ für alle $v, w \in V$, so folgt $f(v,v) = -f(v,v)$, also $2f(v,v) = 0$. Wegen $\mathrm{char}(K) \neq 2$ folgt $f(v,v) = 0$. □

Definition 9.9 Sei $A = (a_{ij})_{i,j=1,\dots,n} \in K^{n\times n}$.

i) A heißt **symmetrisch**, falls $a_{ij} = a_{ji}$ für alle i, j gilt.

ii) A heißt **schiefsymmetrisch**, falls $a_{ii} = 0$ für alle i und $a_{ij} = -a_{ji}$ für alle $i \neq j$ gilt.

Proposition 9.10 Sei $f : V \times V \to K$ bilinear und $B = (b_1, \dots, b_n)$ eine beliebige Basis von V. Dann ist f genau dann symmetrisch bzw. schiefsymmetrisch, wenn die Koordinatenmatrix $M_{f,B,B}$ symmetrisch bzw. schiefsymmetrisch ist.

Beweis : Ist f symmetrisch, so gilt

$$(M_{f,B,B})_{i,j} = f(b_i, b_j) = f(b_j, b_i) = (M_{f,B,B})_{j,i},$$

also ist $M_{f,B,B}$ symmetrisch. Ist umgekehrt $M_{f,B,B}$ symmetrisch, so folgt analog $f(b_i, b_j) = f(b_j, b_i)$ für alle i, j. Für $b = \sum\limits_{i=1}^{n} \beta_i b_i$ und $c = \sum\limits_{j=1}^{n} \gamma_j b_j$ in V folgt

$$\begin{aligned} f(b,c) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i \gamma_j f(b_i, b_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_i \gamma_j f(b_j, b_i) = f(c, b). \end{aligned}$$

Für schiefsymmetrische Formen bzw. Matrizen argumentiert man analog. □

Definition 9.11 Es sei $f : V \times V \to K$ eine symmetrische Bilinearform auf V.

i) Dann heißen $v, w \in V$ **orthogonal** bezüglich f (wir schreiben $v \perp w$), falls $f(v,w) = 0$ ist.

ii) Ist $x \subset V$ eine beliebige Teilmenge, so heißt die Menge

$$X^{\perp} = \{v \in V : f(x,v) = 0 \text{ für alle } x \in X\}$$

orthogonales Komplement von X bezüglich f. $X^{\perp}$ ist ein Unterraum von V.

iii) Ist $v \neq 0$ orthogonal zu sich selbst bezüglich f, d.h. gilt $f(v,v) = 0$, so heißt v **isotrop** bezüglich f.

Beispiel: Es sei

$$\begin{array}{rcl} f : \mathbb{R}^4 \times \mathbb{R}^4 & \to & \mathbb{R} \\ \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right) & \mapsto & x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4 \end{array}$$

die sogenannte Lorentzform, d.h. $f = \beta_A$ für $A = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$. Dann ist f nicht ausgeartet in beiden Variablen. Der Vektor $v = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ erfüllt $f(v,v) = 0$, d.h. v ist isotrop.

Besitzt f einen isotropen Vektor v, so ist das orthogonale Komplement $\langle V \rangle^\perp$ kein Komplement im Sinne der Unterräume von V, da

$$v \in \langle v \rangle \cap \langle v \rangle^\perp \neq 0$$

gilt.

Proposition 9.12 Es sei char$K \neq 2$. Eine symmetrische Bilinearform $f : V \times V \to K$ mit $f \neq 0$ besitzt einen anisotropen Vektor, d.h. es existiert ein $v \neq 0$ in V mit $f(v,v) \neq 0$.

Beweis : Ist $f \neq 0$, so gibt es $v, w \in V$ mit $f(v,w) \neq 0$. Falls $f(v,v) \neq 0$ oder $f(w,w) \neq 0$ sind, so folgt die Behauptung. Im Fall $f(v,v) = 0 = f(w,w)$ gilt

$$f(v+w, v+w) = f(v,v) + f(w,w) + 2f(v,w) = 2f(v,w) \neq 0,$$

da char$(K) \neq 2$. □

Jetzt können wir einen Spezialfall angeben, in dem das orthogonale Komplement ein Komplement im üblichen Sinn ist.

Proposition 9.13 Es sei $f : V \times V \to K$ eine symmetrische Bilinearform und $v \in V$ anisotrop bezüglich f. Dann gilt $V = \langle v \rangle \oplus \langle v \rangle^{\perp}$.

Beweis : Wir zeigen zunächst $\langle v \rangle \cap \langle v \rangle^{\perp} = 0$. Ist für $a \in K$ der Vektor $av \in \langle v \rangle \cap \langle v \rangle^{\perp}$, so folgt $0 = f(v, av) = af(v, v)$. Da v anisotrop ist, gilt $f(v, v) \neq 0$, also $a = 0$, und somit $av = 0$.
Nun zeigen wir $\langle v \rangle + \langle v \rangle^{\perp} = V$. Sei $w \in V$. Wir setzen $a := \frac{f(v,w)}{f(v,v)}$ und schreiben $w = (w - av) + av$. Es ist $f(x - av, v) = f(w, v) - af(v, v) = 0$, woraus $w - av \in \langle v \rangle^{\perp}$ folgt. Also liegt $w = (w - av) + av \in \langle v \rangle^{\perp} + \langle v \rangle$. ☐

Offenbar ist eine symmetrische Bilinearform f auf V genau dann nicht ausgeartet in der ersten Variablen, wenn sie nicht-ausgeartet in der zweiten Variablen ist. Wir nennen sie dann einfach nicht ausgeartet. Definitionsgemäß ist f genau dann nicht ausgeartet, wenn $V^{\perp} = 0$ gilt. Wir können diese Eigenschaft wie folgt an der Koordinatenmatrix von f ablesen:

Satz 9.14 Es sei $f : V \times V \to K$ eine symmetrische Bilinearform auf V und $B = (b_1, \dots, b_n)$ eine Basis von V.

i) Dann gilt
$$V^{\perp} = \{v \in V : M_{f,B,B} \kappa_B(v) = 0\}.$$
Im orthogonalen Komplement von V sind also genau die Vektoren in V, deren Koordinatenmatrix im Kern der durch Multiplikation mit $M_{f,B,B}$ gegebenen linearen Abbildung liegen. Wir können $V^{\perp}$ daher durch Lösen eines linearen Gleichungssystems bestimmen.

ii) f ist genau dann nicht ausgeartet, wenn $M_{f,B,B}$ invertierbar ist.

Beweis :

i) Wir zeigen $V^{\perp} = \{v \in V : M_{f,B,B} \kappa_B(v) = 0\}$.
„$\supset$“: Ist $M_{f,B,B} \kappa_B(v) = 0$, so gilt nach Lemma 9.4 für alle $w \in V$:
$$f(w, v) = \kappa_B(w)^t M_{f,B,B} \kappa_B(v) = 0,$$
also ist $v \in V^{\perp}$.
„$\subset$“: Angenommen, $v \in V^{\perp}$. Dann gilt für alle $w \in V$
$$0 = f(w, v) \overset{9.4}{=} \kappa_B(w)^t M_{f,B,B} \kappa_B(w).$$

Der Vektor $x = M_{f,B,B}\kappa_B(v) = K^n$ hat also die Eigenschaft, dass für alle $y \in K^n$

$$y^t x = 0$$

gilt. Indem wir für y die Einheitsvektoren $e_1, \dots, e_n$ einsetzen, sehen wir, dass alle Koeffizienten von $x = M_{f,B,B}\kappa(v)$ verschwinden. Also folgt $M_{f,B,B}\kappa_B(v) = 0$.

ii) f ist genau dann nicht ausgeartet, wenn $V^\perp = 0$ ist, nach i) also genau dann, wenn aus $M_{f,B,B}\kappa_B(v) = 0$ schon $v = 0$ folgt. Da $\kappa_B : V \to K^n$ ein Isomorphismus ist, ist das äquivalent dazu, dass der Kern der Abbildung

$$\begin{array}{rcl} K^n & \to & K^n \\ x & \mapsto & M_{f,B,B}x \end{array}$$

gleich 0 ist, also zu der Tatsache, dass $M_{f,B,B}$ invertierbar ist.

□

Beispiel:

i) Die Lorentzform $f = \beta_A, A = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix}$ ist nicht-ausgeartet.

ii) Die symmetrische Bilinearform

$$\begin{array}{rcl} f : \mathbb{C}^2 \times \mathbb{C}^2 & \to & \mathbb{C} \\ \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) & \mapsto & x_1 y_1 + i x_2 y_1 + i x_1 y_2 - x_2 y_2 \end{array}$$

ist ausgeartet, da

$$\det \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} = 0$$

gilt.

Definition 9.15 Es sei V ein endlich-dimensionaler K-Vektorraum und $f : V \times V \to K$ eine symmetrische Bilinearform. Eine Basis $B = (b_1, \dots, b_n)$ heißt **Orthogonalbasis von *V*** bezüglich f, falls für alle $i \neq j \quad b_i \perp b_j$ (d.h. $f(b_i, b_j) = 0$) gilt. Also ist B genau dann eine Orthogonalbasis bezüglich f, wenn $M_{f,B,B}$ eine Diagonalmatrix ist.

Ist B eine Orthogonalbasis, so können wir leicht prüfen, ob f nicht ausgeartet ist. Nach Satz 9.14 ist f nämlich genau dann nicht ausgeartet, wenn die Diagonalmatrix

$M_{f,B,B}$ invertierbar ist, d.h. wenn alle Diagonalelemente $f(b_i, b_i)(i = 1, \ldots, n)$ dieser Matrix ungleich Null sind.

Satz 9.16 Jede symmetrische Bilinearform $f : V \times V \to K$ auf einem endlich-dimensionalen Vektorraum V besitzt eine Orthogonalbasis.

Beweis : Ist $f = 0$, so ist jede beliebige Basis von V eine Orthogonalbasis. Wir können also $f \neq 0$ annehmen. Wir zeigen die Behauptung mit Induktion nach dim$V$. Für dim$V = 1$ ist jede Basis von V eine Orthogonalbasis. Also nehmen wir an, die Behauptung gilt für alle symmetrischen Bilinearformen auf Vektorräumen der Dimension $< \dim V$. Nach Proposition 9.12 besitzt f einen anisotropen Vektor v, und nach Proposition 9.13 ist $V = \langle v \rangle \oplus \langle v \rangle^{\perp}$. Nach Induktionsvoraussetzung besitzt der Vektorraum $\langle v \rangle^{\perp}$ zusammen mit der Einschränkung von f auf $\langle v \rangle^{\perp} \times \langle v \rangle^{\perp}$ eine Orthogonalbasis $(b_2, \ldots, b_n)$. Da $f(v, b_i) = 0$ für alle $i = 2, \ldots, n$ gilt, ist $(v, b_2, \ldots, b_n)$ eine Orthogonalbasis von V bezüglich f. □

Korollar 9.17 Sei $A \in K^{n \times n}$ eine symmetrische Matrix. Dann existiert ein $P \in$ GL$_n(K)$, so dass $P^t AP$ eine Diagonalmatrix ist.

Beweis : Ist A symmetrisch, so ist $\beta_A : K^n \times K^n \to K$ eine symmetrische Bilinearform. Nach Satz 9.16 gibt es eine Orthogonalbasis B von K^n bezüglich β_A, d.h. $M_{f,B,B}$ ist eine Diagonalmatrix. Ist C die kanonische Basis von K^n, so ist $M_{f,C,C} = A$. Für die Übergangsmatrix $P = P_C^B \in$ GL$_n(K)$ gilt daher nach Proposition 9.5 $M_{f,B,B} = P^t AP$, woraus die Behauptung folgt. □

10 Symmetrische Bilinearformen über $\mathbb{R}$

In diesem Kapitel werden wir symmetrische Bilinearformen auf endlich-dimensionalen $\mathbb{R}$-Vektorräumen untersuchen. Wir beginnen gleich mit einem Satz, der die Existenz besonders einfacher Orthogonalbasen garantiert.

Satz 10.1 (Trägheitssatz von Sylvester) Es sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und $f : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann existiert eine Basis B von V, für die gilt:

$$M_{f,B,B} = \begin{pmatrix} E_r & & \\ & -E_s & \\ & & 0 \end{pmatrix} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} r \\ \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} s \\ \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} n-r-s \end{matrix}$$

Hier sind $r, s \geq 0$ mit $r + s \leq n$. Das Zahlenpaar (r, s) ist durch f eindeutig bestimmt, es heißt **Signatur** von f.

Beweis : Nach Satz 9.16 existiert eine Orthogonalbasis $c_1, \ldots, c_n$ von V bezüglich f. Wir nummerieren diese so, dass

$$\begin{aligned} f(c_i, c_i) &> 0 && \text{für } 1 \leq i \leq r \\ f(c_i, c_i) &< 0 && \text{für } r+1 \leq i \leq r+s \\ \text{und } f(c_i, c_i) &= 0 && \text{für } r+s+1 \leq i \leq n \end{aligned}$$

gilt. Für $i = 1, \ldots, r$ setzen wir $\alpha_i = \sqrt{f(c_i, c_i)}$ und definieren $b_i = \frac{1}{\alpha_i} c_i$.
Dann ist $f(b_i, b_i) = \frac{1}{\alpha_i^2} f(c_i, c_i) = 1$. Für $i = r+1, \ldots, r+s$ setzen wir $\alpha_i := \sqrt{-f(c_i, c_i)}$ und definieren $b_i = \frac{1}{\alpha_i} c_i$. Dann ist $f(b_i, b_i) = \frac{1}{\alpha_i^2} f(c_i, c_i) = -1$.
Für $i = r + s + 1, \ldots, n$ setzen wir $b_i = c_i$.
Dann ist $B = (b_1, \ldots, b_n)$ eine Orthogonalbasis von V bezüglich f, für die

$$M_{f,B,B} = \begin{pmatrix} E_r & & \\ & -E_s & \\ & & 0 \end{pmatrix}$$

gilt.
Jetzt müssen wir noch zeigen, dass (r, s) durch f eindeutig bestimmt sind. Wir nehmen an, B' ist ebenfalls eine Orthogonalbasis mit

$$M_{f,B',B'} = \begin{pmatrix} E'_r & & \\ & -E'_s & \\ & & 0 \end{pmatrix}.$$

Dann müssen wir $r = r'$ und $s = s'$ zeigen. Offenbar ist $r + s = \mathrm{rang}(M_{f,B,B})$ und $r' + s' = \mathrm{rang}(M_{f,B',B'})$. Ist $P = P^B_{B'} \in \mathrm{GL}_n(\mathbb{R})$ die Übergangsmatrix von B nach B', so gilt nach Proposition 9.5 $P^t M_{f,B',B'} P = M_{f,B,B}$. Da P invertierbar ist, folgt $\mathrm{rang}(M_{f,B',B'}) = \mathrm{rang}(M_{f,B,B})$, also $r + s = r' + s'$.
Als nächsten Schritt zeigen wir, dass die $r + (n - r')$ Vektoren $b_1, \ldots, b_r, b'_{r'+1}, \ldots, b'_n$ linear unabhängig sind. Haben wir das gezeigt, so folgt

$$r + (n - r') \leq \dim V = n,$$

also $r \leq r'$.
Es sei also $\sum\limits_{i=1}^{r} \alpha_i b_i + \sum\limits_{i=r'+1}^{n} \beta_i b'_i = 0$ mit $\alpha_i, \beta_i \in K$ eine Linearkombination der Null.
Wir setzen $v = \sum\limits_{i=1}^{r} \alpha_i b_i = -\sum\limits_{i=r'+1}^{n} \beta_i b'_i$. Dann ist

$$f(v,v) = f\left(\sum_{i=1}^{r} \alpha_i b_i, \sum_{i=1}^{r} \alpha_i b_i\right) = \sum_{i=1}^{r} \alpha_i f(b_i, \sum_{i=1}^{r} \alpha_i b_i) = \sum_{i=1}^{r} \alpha_i^2 f(b_i, b_i) = \sum_{i=1}^{r} \alpha_i^2 \geq 0,$$

da $B = (b_1, \ldots, b_n)$ eine Orthogonalbasis mit $f(b_i, b_i) > 0$ für $1 \leq i \leq r$ ist.
Analog rechnen wir

$$f(v,v) = f(-\sum_{i=r'+1}^{n} \beta_i b'_i, -\sum_{i=r'+1}^{n} \beta_i b'_i) = \sum_{i=r'+1}^{n} \beta_i^2 f(b'_i, b'_i) = -\sum_{i=r'+1}^{r'+s'} \beta_i^2 \leq 0,$$

da $B' = (b'_1, \ldots, b'_n)$ eine Orthogonalbasis mit $f(b'_i, b'_i) < 0$ für $r' + 1 \leq i \leq r' + s'$ und $f(b'_i, b'_i) = 0$ für $r' + s' + 1 \leq i \leq n$.
Also folgt $\sum\limits_{i=1}^{r'} \alpha_i^2 = -\sum\limits_{i=r'+1}^{r'+s'} \beta_i^2 = 0$, d.h. $\alpha_1 = \cdots = \alpha_r = 0$ und $\beta_{r'+1} = \cdots = \beta_{r'+s'} = 0$. Dann ist $\sum\limits_{i=r'+s'+1}^{n} \beta_i b'_i = 0$, also auch $\beta_{r'+s'+1} = \cdots = \beta_n = 0$. Somit sind $b_1, \ldots, b_r, b'_{r'+1}, \ldots, b'_n$ in der Tat linear unabhängig und es folgt $r \leq r'$. Dasselbe Argument mit vertauschten Rollen liefert $r' \leq r$, also $r = r'$. Da $r + s = r' + s'$ gilt, folgt auch $s = s'$. Die Signatur (r, s) ist also tatsächlich eindeutig durch f bestimmt.
□

Wir lernen jetzt eine weitere wichtige Eigenschaft von symmetrischen Bilinearformen über $\mathbb{R}$ kennen.

Definition 10.2

i) Eine symmetrische Bilinearform $f : V \times V \rightarrow \mathbb{R}$ auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum V heißt **positiv definit**, wenn für alle $v \neq 0$ aus V die Ungleichung $f(v,v) > 0$ gilt. Eine positiv definite symmetrische Bilinearform auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum V nennt man auch Skalarprodukt auf V.

ii) Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ heißt **positiv-definit**, falls $\beta_A : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ positiv definit ist, d.h. falls für alle $x \neq 0$ aus $\mathbb{R}^n$ die Ungleichung $x^t Ax > 0$ gilt.

Beispiel: Eine Diagonalmatrix $A = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn alle $d_i > 0$ sind. Für $A = E_n$ heißt die positiv definite symmetrische Bilinearform

$$\begin{aligned} \langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) &\mapsto x_1y_1 + \cdots + x_ny_n \end{aligned}$$

kanonisches Skalarprodukt auf $\mathbb{R}^n$.

Lemma 10.3 Eine symmetrische Bilinearform f auf dem endlich-dimensionalen $\mathbb{R}$-Vektorraum V ist genau dann positiv definit, wenn für eine Basis B von V die Koordinatenmatrix $M_{f,B,B}$ positiv definit ist. Dann ist für jede beliebige Basis B von V die Koordinatenmatrix $M_{f,B,B}$ positiv definit.

Beweis : Es sei B eine beliebige Basis von V. Es ist $f(v,v) = \kappa_B(v)^t M_{f,B,B} \kappa_B(v)$ nach Lemma 9.4. Da $\kappa_B : V \rightarrow \mathbb{R}^n$ ein Isomorphismus ist, ist f genau positiv definit, wenn $M_{f,B,B}$ positiv definit ist. ☐

Definition 10.4 Es sei $f : V \times V \rightarrow \mathbb{R}$ eine symmetrische Bilinearform auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum. Eine Basis $B = (b_1, \ldots, b_n)$ von V heißt **Orthonormalbasis** von V bezüglich f, falls

$$f(b_i, b_j) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

gilt.
Eine Basis B ist also genau dann eine Orthonormalbasis von V bezüglich f, wenn $M_{f,B,B} = E_n$ gilt. In diesem Fall ist die Signatur von f gleich $(n, 0)$. Existiert eine Orthonormalbasis von V bezüglich f, dann ist f positiv definit. Wir zeigen nun, dass auch die Umkehrung gilt.

Satz 10.5 Ist f eine positiv definite symmetrische Bilinearform auf dem endlichdimensionalen $\mathbb{R}$-Vektorraum V, so gibt es eine Orthonormalbasis von V bezüglich f.

Beweis : (mit Hilfe des sogenannten **Gram-Schmidt-Orthonormalisierungsverfahrens**) Wir wählen eine beliebige Basis $C = (c_1, \ldots, c_n)$ von V und konstruieren daraus eine Orthonormalbasis. Da f positiv definit ist, ist $f(c_1, c_1) > 0$, also gilt für $b_1 = \frac{1}{\sqrt{f(c_1,c_1)}} c_1$ die Gleichung $f(b_1, b_1) = 1$.
Als nächsten Schritt suchen wir einen Vektor $b_2' \in \langle b_1, c_2 \rangle$ mit $b_1 \perp b_2'$. Der Ansatz $b_2' = \alpha_1 b_1 + c_2$ liefert $f(b_1, b_2') = \alpha_1 f(b_1, b_1) + f(b_1, c_2) = \alpha_1 + f(b_1, c_2)$. Also gilt für

$$\alpha_1 = -f(b_1, c_2) \in \mathbb{R},$$

dass der Vektor $b_2' = \alpha_1 b_1 + c_2$ orthogonal zu b_1 ist. Offenbar gilt $\langle b_1, b_2' \rangle = \langle b_1, c_2 \rangle = \langle c_1, c_2 \rangle$. Wir setzen wieder $b_2 = \frac{1}{\sqrt{f(b_2' b_2')}} b_2$ und erhalten so ein Paar(b_1, b_2) von Vektoren in V mit $\langle b_1, b_2 \rangle = \langle c_1, c_2 \rangle$ und $f(b_i, b_j) = \delta_{ij}$ für $i = 1, 2$. So fahren wir fort. Sind Vektoren $(b_1, \ldots, b_k)$ mit $\langle b_1, \ldots, b_k \rangle = \langle c_1, \ldots, c_k \rangle$ und $f(b_i, b_j) = \delta_{ij}$ gefunden, so sei

$$\alpha_1 = -f(b_1, c_{k+1}), \ldots, \alpha_k = -f(b_k, c_{k+1}).$$

Der Vektor $b_{k+1}' = \alpha_1 b_1 + \cdots + \alpha_k b_k + c_{k+1}$ erfüllt dann

$$\begin{aligned} f(b_i, b_{k+1}) &= f(b_i, \sum_{j=1}^{k} \alpha_j b_j) + f(b_i, c_{k+1}) \\ &= \alpha_i + f(b_i, c_{k+1}) = 0. \end{aligned}$$

Mit $b_{k+1} = \frac{1}{\sqrt{f(b_{k+1}', b_{k+1}')}} b_{k+1}'$ erhalten wir ein System $(b_1, \ldots, b_{k+1})$ von Vektoren, so dass $\langle b_1, \ldots, b_{k+1} \rangle = \langle c_1, \ldots, c_{k+1} \rangle)$ und $f(b_i, b_j) = \delta_{ij}$ gilt. Sind alle c_i verbraucht, so haben wir eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ gefunden. □

Wir hätten Satz 10.5 auch ganz einfach mit dem Sylvesterschen Trägheitsgesetz beweisen können. Der hier angegebene Beweis liefert allerdings ein effektives Verfahren, um eine Orthonormalbasis zu konstruieren. Dieses Resultat wollen wir noch für Matrizen umformulieren.

Satz 10.6 Sei $A \in \mathbb{R}^{n\times n}$. Dann sind äquivalent.

i) Es gibt eine Basis B von $\mathbb{R}^n$ mit $A = M_{\langle,\rangle,B,B}$, wobei $\langle,\rangle$ das kanonische Skalarprodukt auf dem $\mathbb{R}^n$ ist.

ii) Es gibt ein $P \in \mathrm{GL}(n,\mathbb{R})$ mit $A = P^tP$.

iii) A ist symmetrisch und positiv definit.

Beweis : i) $\Rightarrow$ iii) ist klar mit Proposition 9.10 und Lemma 10.3
iii) $\Rightarrow$ ii) Wir wenden Satz 10.5 auf die positiv definite symmetrische Bilinearform $\beta_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ an. Es gibt also eine Orthonormalbasis B von $\mathbb{R}^n$ bezüglich β_A, d.h. $M_{\beta_A,B,B} = E_n$. Es sei C die kanonische Basis des $\mathbb{R}^n$ und $P = P_B^C \in \mathrm{GL}(n,\mathbb{R})$. Dann gilt nach Proposition 9.5: $A = M_{\beta_A,C,C} = P^tE_nP = P^tP$.
ii) $\Rightarrow$ i): Es sei $C = (e_1,\ldots,e_n)$ die kanonische Basis des $\mathbb{R}^n$. Nach Lemma 6.13 gibt es eine Basis B von $\mathbb{R}^n$ mit $P_C^B = P$. Dann ist

$$\begin{aligned} M_{\langle,\rangle,B,B} &\overset{9.5}{=} P^tM_{\langle,\rangle,C,C}P \\ &= P^tP = A. \end{aligned}$$

□

Wir wollen jetzt noch ein Kriterium zeigen, mit dem man feststellen kann, ob eine Matrix positiv definit ist.

Definition 10.7 Sei $A = (a_{ij})_{i,j=1,\ldots,n} \in \mathbb{R}^{n\times n}$. Es sei

$$A_k = \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{pmatrix} \in \mathbb{R}^{k\times k}$$

die obere linke $k\times k$-Untermatrix von A. Die Werte $\det A_1, \det A_2, \ldots, \det A_n = \det A$ heißen **Hauptminoren von A**.
Beispiel: Für $A = \begin{pmatrix} -1 & 1 \\ 1 & -2 \end{pmatrix}$ sind $\det A_1 = -1, \det A_2 = \det A = 1$ die Hauptminoren.

Satz 10.8 Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ ist genau dann positiv definit, wenn alle Hauptminoren positiv sind, d.h. wenn $\det A_k > 0$ für $k = 1,\ldots,n$ gilt.

Beweis : Angenommen, A ist positiv definit. Dann gibt es nach Satz 10.6 ein $P \in \mathrm{GL}(n,\mathbb{R})$ mit $A = P^tP$. Daher ist $\det A = \det(P^tP) = (\det P)^2 > 0$.

Für alle $k \geq 1$ sei V_k der von den ersten k Einheitsvektoren $e_1, \dots, e_k$ aufgespannte Unterraum von $\mathbb{R}^n$. Die Matrix A definiert eine positiv definite symmetrische Bilinearform β_A auf $\mathbb{R}^n$. Auch ihre Einschränkung $\beta_{A|_{V_k}}$ auf V_k ist positiv definit. Die Koordinatenmatrix von $\beta_{A|_{V_k}}$ bezüglich $e_1, \dots, e_k$ ist gerade A_k. Also gilt $\det A_k > 0$. Wir nehmen umgekehrt an, alle Hauptminoren von A seien positiv und zeigen mit Induktion nach n, dass A dann positiv definit ist. Für $n = 1$ ist das klar. Also können wir annehmen, dass A_{n-1}, deren Hauptminoren gerade $A_1, \dots, A_{n-1}$ sind, positiv definit ist. Nach Satz 10.6 gibt es dann ein $Q \in \mathrm{GL}(n-1, \mathbb{R})$mit $Q\ A_{n-1}Q^t = E_{n-1}$. Es sei $\tilde{P} = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{GL}(n, \mathbb{R})$. Es ist $\tilde{P}\ A\tilde{P}^t = \begin{pmatrix} E_{n-1} & * \\ * & * \end{pmatrix}$. Durch $(n-1)$ elementare Zeilenumformungen, also Multiplikation mit Elementarmatrizen $L_{n-1}, \dots, L_1$ von links, kann man die ersten $(n-1)$-Einträge der letzten Zeile eliminieren. Da $\tilde{P}A\tilde{P}^t$ symmetrisch ist, kann man mit den analogen Spaltenoperationen, also durch Multiplikation mit $L_1^t, \dots, L_{n-1}^t$ von rechts, die ersten $(n-1)$ Einträge der letzte Spalte eliminieren. Für $P = L_{n-1} \cdots L_1 \tilde{P}$ gilt also $PAP^t = \begin{pmatrix} E_{n-1} & 0 \\ 0 & c \end{pmatrix}$ für ein $c \in \mathbb{R}$. Aus $\det A > 0$ folgt $c > 0$. Somit ist PAP^t positiv definit, woraus mit Lemma 10.3 auch A positiv definit folgt. ☐

Definition 10.9 Ein endlich-dimensionaler reeller Vektorraum V mit einer positiv definiten, symmetrischen Bilinearform $\langle , \rangle : V \times V \to \mathbb{R}$ heißt auch **euklidischer Vektorraum**.

Ist $(V, \langle , \rangle)$ ein euklidischer Vektorraum, so können wir für alle $v \in V$ eine Länge $||v|| = \sqrt{\langle v, v \rangle}$ definieren. Es gilt dann

Lemma 10.10

i) $||v|| = 0$ genau dann, wenn $v = 0$.

ii) Für alle $v, w \in V$ ist $||v + w|| \leq ||v|| + ||w||$ (**Dreiecksungleichung**)

iii) Für alle $v, w \in V$ ist $|\langle v, w \rangle| \leq ||v||\ ||w||$ (**Cauchy-Schwarz'sche Ungleichung**).

Beweis :

i) $||v|| = 0 \Leftrightarrow \langle v, v \rangle = 0 \Leftrightarrow v = 0$, da $\langle , \rangle$ positiv definit ist.

ii) folgt aus iii), denn es gilt

$$\begin{aligned} ||v+w||^2 &= \langle v+w, v+w\rangle = \langle v,v\rangle + 2\langle v,w\rangle + \langle w,w\rangle \\ &= ||v||^2 + 2\langle v,w\rangle + ||w||^2 \\ &\leq ||v||^2 + 2|\langle v,w\rangle| + ||w||^2 \\ &\overset{\text{(iii)}}{\leq} ||v||^2 + 2||v||\,||w|| + ||w||^2 \leq (||v|| + ||w||)^2. \end{aligned}$$

iii) Es sei $W = \langle v, w\rangle \subset V$. Die Einschränkung von $\langle,\rangle$ auf W ist ebenfalls eine positiv definite, symmetrische Bilinearform (also ein Skalarprodukt) $\langle,\rangle_W$ auf W.

Ist $\dim W = 1$, so können wir $v = \lambda w$ für ein $\lambda \in K$ annehmen. Dann gilt

$$|\langle v,w\rangle| = |\lambda|\,||w||^2 = ||v||\,||w||.$$

Ist $\dim W = 2$, so ist (v, w) eine Basis von W. Die Koordinatenmatrix der Bilinearform $\langle,\rangle$ bezüglich dieser Basis ist

$$A = \begin{pmatrix} \langle v,v\rangle & \langle v,w\rangle \\ \langle v,w\rangle & \langle w,w\rangle \end{pmatrix}.$$

Nach Lemma 10.3 ist A positiv definit, also ist nach Satz 10.8

$$\det A = \langle v,v\rangle\langle w,w\rangle - \langle v,w\rangle^2 > 0.$$

Es folgt also $||v||^2\,||w||^2 > \langle v,w\rangle^2$, und daher $||v||\,||w|| > |\langle v,w\rangle|$.

□

Ist $(V, \langle,\rangle)$ ein euklidischer Vektorraum und $B = (b_1, \dots, b_n)$ eine Orthonormalbasis von V bezüglich $\langle,\rangle$, so ist $\kappa_B : V \to \mathbb{R}^n$ ein Isomorphismus von $\mathbb{R}$-Vektorräumen, der $\langle,\rangle$ in das kanonische Skalarprodukt des $\mathbb{R}^n$ überführt. Es gilt nämlich für $v = \sum\limits_{i=1}^{n} \alpha_i b_i$ und $w = \sum\limits_{i=1}^{n} \beta_i b_i$:

$$\begin{aligned} \langle v,w\rangle &= \left\langle \sum_{i=1}^{n} \alpha_i b_i, \sum_{i=1}^{n} \beta_i b_i \right\rangle \\ &= \sum_{i=1}^{n} \alpha_i\beta_i = (\alpha_i, \dots, \alpha_n) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \\ &= \kappa_B(v)^t \kappa_B(w). \end{aligned}$$

Auf einem euklidischen Vektorraum $(V, \langle , \rangle)$ kann man nicht-orientierte Winkel definieren.

Definition 10.11 Für $v, w \in V \setminus \{0\}$ sei $\theta \in [-\pi, \pi[$ eine Zahl mit

$$\cos\theta = \frac{\langle v, w\rangle}{||v|| \, ||w||}.$$

Dadurch ist θ bis auf das Vorzeichen ± 1 eindeutig festgelegt. Nach der Cauchy-Schwarz'schen Ungleichung ist nämlich $-1 \le \frac{\langle v,w\rangle}{||v|| \, ||w||} \le 1$. θ heißt **(nicht-orientierter) Winkel** zwischen v und w. Aus dieser Definition folgt sofort $||v+w||^2 = ||v||^2 + ||w||^2 + 2||v|| \, ||w|| \cos\theta$.

In einem euklidischen Vektorraum $(V, \langle , \rangle)$ ist für jeden Unterraum $W \subset V$ das orthogonale Komplement $W^\perp$ ein Komplement im Vektorraumsinne, d.h. es gilt

$$W \oplus W^\perp = V$$

(siehe Übungsaufgaben). Jedes $v \in V$ hat also eine eindeutige Darstellung $v = w + w'$ mit $w \in W, w' \in W^\perp$. Die Abbildung

$$\begin{array}{rcl} p : V & \to & W \\ v & \mapsto & w \end{array}$$

ist linear. Es ist $p(v) = w$ genau dann, wenn $v - w \in W^\perp$ liegt. p heißt **orthogonale Projektion von *V* auf *W***.

Lemma 10.12 Es sei $W \subset V$ ein Unterraum des euklidischen Vektorraums $(V, \langle , \rangle)$ und $(w_1, \dots, w_r)$ eine Orthonormalbasis von W. Dann gilt für die orthogonale Projektion $p : V \to W$: Für alle $v \in V$ ist $p(v) = \langle v, w_1\rangle w_1 + \dots + \langle v, w_r\rangle w_r$.

Beweis : Wir setzen $\tilde{w} = \sum\limits_{i=1}^{r} \langle v, w_i\rangle w_i$. Für alle $j = 1, \dots, r$ gilt

$$\langle v - \tilde{w}, w_j\rangle = \langle v - \sum_{i=1}^{r} \langle v, w_i\rangle w_i, w_j\rangle = \langle v, w_j\rangle - \langle v, w_j\rangle = 0.$$

Also ist $v - \tilde{w} \in W^\perp$, woraus $p(v) = \tilde{w}$ folgt. $\square$

Es sei $(V, \langle , \rangle)$ ein euklidischer Raum.

Definition 10.13

i) Eine lineare Abbildung $f : V \to V$ heißt **Isometrie** (bezüglich $\langle , \rangle$), falls für alle $v, w \in V$
$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$
gilt.

ii) Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt **orthogonal**, falls $A^t A = E_n$ gilt.

Offenbar ist jede Isometrie injektiv, also ein Isomorphismus, falls V endlich-dimensional ist. Eine Matrix A ist definitionsgemäß genau dann orthogonal, wenn $\lambda_A : \mathbb{R}^n \to \mathbb{R}^n$ eine Isometrie bezüglich des kanonischen Skalarproduktes ist. Jede orthogonale Matrix A ist invertierbar mit $A^{-1} = A^t$. Also gilt auch $AA^t = E_n$. Ist V endlich-dimensional, so ist

$$O(V, \langle , \rangle) = \{ f : V \to V : f \text{ Isometrie } \}$$

eine Gruppe bezüglich der Komposition von Abbildungen.
Ebenso bildet die Menge

$$\begin{aligned} O(n, \mathbb{R}) &= \{ A \in \mathbb{R}^{n \times n} : A \text{ orthogonal } \} \\ &= \{ A \in \mathbb{R}^{n \times n} : A^t A = E_n \} \end{aligned}$$

eine Untergruppe von GL$(n, \mathbb{R})$ bezüglich der Multiplikation von Matrizen, die sogenannte **orthogonale Gruppe**. Jedes $A \in O(n, \mathbb{R})$ erfüllt wegen $A^t A = E_n$ die Gleichung $(\det A)^2 = 1$, also ist $\det A \in \{\pm 1\}$. Die Menge SO$(n, \mathbb{R}) = \{ A \in O(n, \mathbb{R}) : \det A = 1 \}$ ist eine Untergruppe von $O(n, \mathbb{R})$, sie heißt **spezielle orthogonale Gruppe**.

11 Hermitesche Formen über $\mathbb{C}$

Ist der Grundkörper $K = \mathbb{C}$, so ist es sinnvoll, statt mit symmetrischen Bilinearformen mit den sogenannnnten hermite'schen Formen zu arbeiten.

Definition 11.1 Sei V ein $\mathbb{C}$-Vektorraum. Eine **hermite'sche Form** auf V ist eine Abbildung
$$f : V \times V \to \mathbb{C}$$
mit folgenden Eigenschaften:

i) f ist linear in der zweiten Variablen:

$$\begin{aligned} f(v, w_1 + w_2) &= f(v, w_1) + f(v, w_2) \text{ und} \\ f(v, \alpha w_1) &= \alpha f(v, w_1) \end{aligned}$$

für alle $v, w_1, w_2 \in V$ und $\alpha \in \mathbb{C}$.

ii) f ist konjugiert-linear in der ersten Variablen:

$$\begin{aligned} f(v_1 + v_2, w) &= f(v_1, w) + f(v_2, w) \text{ und} \\ f(\alpha v_1, w) &= \overline{\alpha} f(v_1, w) \end{aligned}$$

für alle $v_1, v_2, w \in V$ und $\alpha \in \mathbb{C}$.

iii) f ist konjugiert symmetrisch, d.h.

$$f(v, w) = \overline{f(w, v)} \text{ für alle } v, w \in V.$$

Eine Abbildung f, die nur die Eigenschaften i) und ii) besitzt, wird auch **Sesquilinarform** genannt (lat. sesqui bedeutet soviel wie anderthalbfach).

Beispiel: $\langle , \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}, \langle x, y \rangle = \overline{x}^t y$ ist eine hermite'sche Form.

Ist f eine hermite'sche Form auf V, so gilt für alle $v \in V$ nach iii) aus Definition 11.1:

$$f(v, v) = f(v, v), \text{ d.h. } f(v, v) \in \mathbb{R}.$$

Definition 11.2 Eine hermite'sche Form f auf dem komplexen Vektorraum V heißt **positiv definit**, falls für alle $v \neq 0$ aus V

$$f(v, v) > 0$$

gilt. Da $f(v, v)$ eine reelle Zahl ist, ist diese Definition sinnvoll.

Beispiel: Die Form $\langle x, y \rangle = \overline{x}^t y$ ist positiv definit auf $\mathbb{C}^n$.

Definition 11.3 Ein $\mathbb{C}$-Vektorraum V zusammen mit einer positiv definiten hermite'schen Form heißt **unitärer Vektorraum**. In einem unitären Vektorraum $(V, \langle , \rangle)$ ist durch $||v|| = \sqrt{\langle v, v \rangle}$ eine Länge von Vektoren definiert. Diese erfüllt die Eigenschaften i)-iii) aus Lemma 10.10, wobei man in iii) den komplexen Absolutbetrag verwenden muss. (Prüfen Sie das!)

Nun betrachten wir die Koordinatenmatrizen zu hermite'schen Formen.

Definition 11.4 Ist f eine hermite'sche Form auf dem $\mathbb{C}$-Vektorraum V und $B = (b_1, \dots, b_n)$ eine Basis von V, so ist die Koordinatenmatrix $M_{f,B,B}$ von f bezüglich B definiert als die Matrix mit den Einträgen $(f(b_i, b_j))_{i,j}$.

Lemma 11.5 Ist f eine hermite'sche Form auf V und $B = (b_1, \dots, b_n)$ eine Basis von V, so gilt für alle $v, w \in V$ die Gleichung

$$f(v, w) = \overline{\kappa_b(v)}^t M_{f,B,B} \kappa_B(w).$$

Beweis : analog zu Lemma 9.4 □

Proposition 11.6 Ist $M = M_{f,B,B}$ die Koordinatenmatrix einer hermite'schen Form f, so gilt

$$\overline{M}^t = M.$$

Beweis : Der Eintrag von M an der Stelle (i, j) ist definitionsgemäß

$$f(b_i, b_j) = \overline{f(b_j, b_i)},$$

also gleich dem Eintrag von $\overline{M}^t$ an der Stelle (i, j). □

Definition 11.7 Für $A \in \mathbb{C}^{n \times n}$ schreiben wir $A^* = \overline{A}^t$ und nennen A^* die Adjungierte zu A. Eine Matrix A heißt **hermite'sch**, falls $A = A^*$ gilt.

Hier muss man mit der Terminologie aufpassen: Die Adjungierte A^* hat nichts mit der adjungierten Matrix aus den Cramer'schen Regeln zu tun.

Ist $A \in \mathbb{C}^{n \times n}$ hermite'sch, so ist A von der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \overline{a_{12}} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \dots & a_{nn} \end{pmatrix}$$

mit reellen Diagonaleinträgen $a_{11}, \dots, a_{nn}$.

Beispiel: $A = \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix}$ ist hermite'sch. Eine Matrix mit reellen Einträgen ist hermite'sch, wenn sie symmetrisch ist.

Für Koordinatenmatrizen hermite'scher Formen hat man folgendes Basiswechselresultat:

Proposition 11.8 Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und $f : V \times V \to \mathbb{C}$ hermite'sch. Sind B und B' Basen von V und $P = P_B^{B'}$ die Übergangsmatrix, so gilt

$$M_{f,B',B'} = \overline{P}^t M_{f,B,B} P.$$

Beweis : Ist $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$, so ist der Koeffizient an der Stelle (i,j) von $M_{f,B',B'}$ gerade $f(b'_i, b'_j) \overset{11.5}{=} \overline{\kappa_B(b'_i)}^t M_{f,B,B} \kappa_B(b'_j)$, also gleich dem Koeffizienten an der Stelle (i,j) von $\overline{P}^t M_{f,B,B} P$. ☐

Definition 11.9 Ist $(V, \langle, \rangle)$ ein endlich-dimensionaler unitärer Vektorraum, so heißt eine Basis $B = (b_1, \ldots, b_n)$ **Orthonormalbasis von V**, falls $\langle b_i, b_j \rangle = \delta_{ij}$ gilt.

Man kann das Gram-Schmidt-Verfahren auf unitäre Vektorräume übertragen und so die Existenz einer Orthonormalbasis zeigen.

Satz 11.10 Es sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum und $f : V \times V \to \mathbb{C}$ eine positiv definite hermite'sche Form. Dann besitzt V eine Orthonormalbasis bezüglich f.

Beweis : analog zum Beweis von Satz 10.5 ☐

Definition 11.11 Ist $(V, \langle, \rangle)$ ein unitärer Vektorraum, so heißt eine lineare Abbildung $f : V \to V$ eine Isometrie von V (bezüglich $\langle, \rangle$), falls für alle $v, w \in V$

$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$

gilt.

Ist V endlich-dimensional, so ist jede Isometrie ein Isomorphismus und wie im euklidischen Fall bildet die Menge aller Isometrien $f : V \to V$ eine Gruppe, die auch mit $U(V, \langle\rangle)$ bezeichnet wird.

Definition 11.12 Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt **unitär**, falls $A^* A = E_n$ gilt.

Jede unitäre Matrix A ist invertierbar mit $A^{-1} = A^*$. Also gilt auch $AA^* = E_n$. Ist A unitär, so ist $|\det A|^2 = 1$.
Definitionsgemäß ist $A \in \mathbb{C}^{n \times n}$ genau dann unitär, wenn $\lambda_A : \mathbb{C}^n \times \mathbb{C}^n$ eine Isometrie bezüglich der kanonischen hermite'schen Form

$$\begin{array}{rcl} \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ (x,y) & \mapsto & \overline{x}^t y \end{array}$$

ist.
Die reellen unitären Matrizen sind gerade die orthogonalen Matrizen. Die Menge $U(n,\mathbb{C}) = \{A \in \mathbb{C}^{n\times n} : A$ unitär $\}$ ist eine Untergruppe von GL$(n,\mathbb{C})$, die sogenannte **unitäre Gruppe**. Die Menge SU$(n,\mathbb{C}) = \{A \in U(n,\mathbb{C}) : \det A = 1\}$ der unitären Matrizen der Determinante 1 bildet eine Untergruppe von $U(n,\mathbb{C})$, sie heißt **spezielle unitäre Gruppe**.

Ab jetzt sei $(V, \langle,\rangle)$ ein n-dimensionaler unitärer Raum, d.h. $\langle,\rangle$ ist eine positiv definite hermite'sche Form auf V. Wir wollen nun für geeignete Endomorphismen $f : V \to V$ eine Orthonormalbasis von $(V, \langle,\rangle)$ finden, so dass die Koordinatenmatrix von f Diagonalgestalt hat.

Satz 11.13 Es sei $f : V \to V$ linear und es sei B eine Orthonormalbasis von $(V, \langle,\rangle)$. Wir betrachten die Koordinatenmatrix $A = A_{f,B,B}$ der linearen Abbildung f.

i) A ist genau dann hermite'sch, wenn $\langle v, f(w)\rangle = \langle f(v), w\rangle$ für alle $v, w \in V$ gilt. Die lineare Abbildung f heißt dann **selbstadjungiert** bezüglich $\langle,\rangle$.

ii) A ist genau dann unitär, wenn $\langle v, w\rangle = \langle f(v), f(w)\rangle$ für alle $v, w \in V$ gilt, also genau dann, wenn f eine Isometrie von $(V, \langle,\rangle)$ ist.

Beweis : Da B eine Orthonormalbasis ist, ist die Koordinatenmatrix von $\langle,\rangle$ bezüglich B gerade die Einheitsmatrix. Nach Lemma 11.5 gilt also für alle $x, y \in V$, dass $\langle x, y\rangle = \overline{\kappa_B(x)^t}\kappa_B(y)$ ist. Ferner gilt nach Proposition 6.11 für alle $x \in V$, dass $\kappa_B(f(x)) = A\kappa_B(x)$ ist.

i) Ist A hermite'sch, d.h. gilt $\overline{A}^t = A$, so folgt für alle $v, w \in V$:

$$\begin{aligned}
\langle v, f(w)\rangle &= \overline{\kappa_B(v)}^t \kappa_B(f(w)) \\
&= \overline{\kappa_B(v)}^t A\kappa_B(w) \\
&= \overline{\kappa_B(v)}^t \overline{A}^t \kappa_B(w) \\
&= \overline{(A\kappa_B(v))}^t \kappa_B(w) \\
&= \langle f(v), w\rangle.
\end{aligned}$$

Ist umgekehrt f selbstadjungiert bezüglich $\langle,\rangle$, so folgt für Elemente der Orthonormalbasis $B = (b_1, \ldots, b_n)$:

$$\langle b_i, f(b_j)\rangle = \langle f(b_i), b_j\rangle,$$

also gilt

$$\begin{aligned} e_iAe_j &= \kappa_B(b_i)^tA\kappa_B(b_j) \\ &= \kappa_B(b_i)^t\kappa_B(f(b_j)) \\ &= \langle b_i, f(b_j)\rangle = \langle f(b_i), b_j\rangle \\ &= \overline{\kappa_B(f(b_i))}^t\kappa_B(b_j) \\ &= \overline{A\kappa_B(b_i))}^t\kappa_B(b_j) \\ &= \overline{\kappa_B(b_i)}^t\overline{A}^t\kappa_B(b_j) \\ &= e_i^t\overline{A}^te_j, \end{aligned}$$

woraus $A = \overline{A}^t$, d.h. A hermite'sch, folgt.

ii) Ist A unitär, d.h. gilt $A^*A = E_n$, so folgt für alle $v, w \in V$:

$$\begin{aligned} \langle f(v), f(w)\rangle &= \overline{\kappa_B(f(v))}^t\kappa_B(f(w)) \\ &= \overline{A\kappa_B(v))}^t(A\kappa_B(w)) \\ &= \overline{\kappa_B(v)}^t\overline{A}^tA\kappa_B(w) \\ &= \overline{\kappa_B(v)}^tA^*A\kappa_B(w) \\ &= \overline{\kappa_B(v)}^t\kappa_B(w) = \langle v, w\rangle. \end{aligned}$$

Ist umgekehrt f eine Isometrie bezüglich $\langle , \rangle$, so gilt für alle $i, j \in \{1, \ldots, n\}$:

$$\begin{aligned} e_i^tA^*Ae_j &= \overline{\kappa_B(b_i)}^t\overline{A}^tA\kappa_B(b_j) \\ &= \overline{A\kappa_B(b_i))}^t(A\kappa_B(b_j)) \\ &= \overline{\kappa_B(f(b_i))}^t\kappa_B(f(b_j)) \\ &= \langle f(b_i), f(b_j)\rangle \\ &= \langle b_i, b_j\rangle \\ &= \delta_{ij}, \end{aligned}$$

d.h. $A^*A = E_n$.

□

Jetzt können wir folgenden wichtigen Satz zeigen:

Satz 11.14 (Spektralsatz, auch Satz von der Hauptachsentransformation genannt)

i) Sei $(V, \langle\rangle)$ ein endlich-dimensionaler unitärer Vektorraum und $f\colon V \to V$ eine selbstadjungierte lineare Abbildung, d.h. es gilt $\langle f(v), w\rangle = \langle v, f(w)\rangle$ für alle $v, w \in V$. Dann gibt es eine Orthonormalbasis von $(V\langle , \rangle)$, die nur aus Eigenvektoren von f besteht. Insbesondere ist f diagonalisierbar.

ii) Sei A eine hermite'sche Matrix in $\mathbb{C}^{n\times n}$. Dann gibt es eine unitäre Matrix $P \in U(n,\mathbb{C})$, so dass

$$PAP^* \text{ eine Diagonalmatrix}$$

mit reellen Einträgen ist.

Beweis : (mit Induktion nach $n = \dim V$)

i) Ist $\dim V = 0$, so ist die Behauptung klar. Wir nehmen also an, $n = \dim V \geq 1$ und die Behauptung gilt für alle Vektorräume der Dimension $< n$. Das charakteristische Polynom $\chi_f(X) \in \mathbb{C}[X]$ hat nach dem Fundamentalsatz der Algebra eine Nullstelle $a \in \mathbb{C}$. Diese ist ein Eigenwert von f, also besitzt f einen Eigenvektor b_1. Wir normieren b_1 so, dass $\langle b_1, b_1\rangle = 1$ ist (indem wir gegebenenfalls b_1 durch $\frac{b_1}{||b_1||}$ ersetzen). Indem wir b_1 zu einer beliebigen Basis von V ergänzen und dann das Gram-Schmidt-Verfahren anwenden, können wir b_1 zu einer Orthonormalbasis $B = (b_1,\ldots,b_n)$ ergänzen. Die Koordinatenmatrix $A_{f,B,B}$ hat dann folgende Gestaalt:

$$A_{f,B,B} = \begin{pmatrix} a & * & \ldots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

mit einer Matrix $A' \in \mathbb{C}^{(n-1)\times(n-1)}$.

Da f selbstadjungiert ist, ist nach Satz 11.13 die Matrix $A_{f,B,B}$ hermite'sch, d.h. es gilt $\overline{A_{f,B,B}}^t = A_{f,B,B}$.

Also folgt $A_{f,B,B} = \begin{pmatrix} a & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$ mit einer hermite'schen Matrix A'.

Somit bildet f den von $B' = (b'_2,\ldots,b'_n)$ erzeugten Unterraum W von V in sich ab, und es gilt $A_{f_{|W},B',B'} = A'$. Da A' hermite'sch ist, ist $f_{|W}$ nach Satz 5.13 selbst-adjungiert. Also gibt es nach Induktionsvoraussetzung eine Orthonormalbasis $C = (c_2,\ldots,c_n)$ von W, so dass $A_{f_{|W},C,C}$ eine Diagonalmatrix ist. Aber dann ist $(b_1, c_2,\ldots,c_n)$ eine Orthonormalbasis von V, so dass die Koordinatenmatrix von f bezüglich dieser Basis eine Diagonalmatrix ist.

ii) Wir betrachten die lineare Abbildung $\lambda_A : \mathbb{C}^n \to \mathbb{C}^n$. Da A hermite'sch ist, ist λ_A nach Satz 11.13 selbstadjungiert bezüglich des kanonischen Skalarprodukts $\langle,\rangle$ auf $\mathbb{C}^n$.

Nach i) existiert also eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ von $\mathbb{C}^n$ bezüglich $\langle , \rangle$, so dass die Koordinatenmatrix von λ_A bezüglich B Diagonalgestalt hat.

Ist P die Übergangsmatrix von der kanonischen Basis zu B, so ist also PAP^{-1} eine Diagonalmatrix. Nun ist P^{-1} die Übergangsmatrix von B zur kanonischen Basis, also die Matrix mit den Spalten $b_1, \ldots, b_n$. Da B eine Orthonormalbasis bezüglich $\langle , \rangle$ ist, folgt also $(P^{-1})^* P^{-1} = E_n$, d.h. P^{-1} ist unitär. Somit ist auch P unitär mit $P^{-1} = P^*$, woraus die Behauptung folgt.

□

Korollar 11.15 Es sei $(V, \langle , \rangle)$ ein n-dimensionaler unitärer Vektorraum und $q : V \times V \to \mathbb{C}$ eine weitere hermite'sche Form auf V. Dann gibt es eine Orthonormalbasis B von $(V, \langle , \rangle)$, die gleichzeitig Orthogonalbasis von q ist. Die Formen $\langle , \rangle$ und q lassen sich also simultan diagonalisieren.

Beweis : Es sei C eine beliebige Orthonormalbasis von $(V, \langle \rangle)$. Nach Proposition 11.6 ist die Koordinatenmatrix $M = M_{q,C,C}$ hermite'sch. Also existiert nach Satz 11.14 ii) ein $P \in U(n, \mathbb{C})$, so dass PMP^* eine Diagonalmatrix ist. Die Matrix P ist nach Lemma 6.13 die Übergangsmatrix von C zu einer Basis $B = (b_1, \ldots, b_n)$ von V. Also ist $P^* = P^{-1} = P_C^B$, so dass aus Proposition 11.8 $PMP^* = M_{f,B,B}$ folgt. Da diese Matrix eine Diagonalmatrix ist, ist B eine Orthogonalbasis bezüglich q. Ferner ist $M_{\langle , \rangle, C, C} = E_n$, also folgt wieder mit Proposition 11.8: $M_{\langle , \rangle, B, B} = PE_nP^* = E_n$. Somit ist B eine Orthonormalbasis bezüglich $\langle , \rangle$. □

Korollar 11.16 Ist $f : V \to V$ eine selbstadjungierte Abbildung eines endlichdimensionalen unitären $\mathbb{C}$-Vektorraums, so sind Eigenvektoren zu verschiedenen Eigenwerten orthogonal zueinander. Insbesondere sind sie linear unabhängig.

Beweis : Nach dem Spektralsatz gibt es eine Orthonormalbasis $B = (b_1, \ldots, b_n)$ von V, die aus Eigenvektoren von f besteht. Wir nummerieren die b_i so, dass für die paarweise verschiedenen Eigenwerte $\alpha_1, \ldots, \alpha_r$ von f gilt:

$$\langle b_1, \ldots, b_{m_1} \rangle = V_{\alpha_1}, \langle b_{m_1+1}, \ldots, b_{m_1+m_2} \rangle = V_{\alpha_2}, \ldots, \langle b_{m_1+\cdots+m_{r-1}+1}, \ldots, b_n \rangle = V_{\alpha_r}.$$

Sind v und w Eigenvektoren zu verschiedenen Eigenwerten, liegt also $v \in V_{\alpha_i}$ und $w \in V_{\alpha_j}$ für $\alpha_i \neq \alpha_j$, so folgt $v \perp w$. □

Ist f also eine selbstadjungierte Abbildung mit paarweise verschiedenen Eigenwerten $\alpha_1, \ldots, \alpha_n$, so ist jede Basis B aus Eigenvektoren von f eine Orthogonalbasis, die man durch Normieren der Basisvektoren in eine Orthonormalbasis überführen kann

Beispiel: Es sei $A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$. Dann ist

$$\chi_A(X) = (X-2)^2 - 1 = X^2 - 4X + 3 = (X-1)(X-3),$$

d.h. A hat die Eigenwerte 1 und 3. Lösen der Gleichungssysteme $(A - E_2)x = 0$ und $(A - 3E_2)x = 0$ liefert Eigenvektoren

$$v_1' = \begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ zu 3 und } v_2' = \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ zu 1.}$$

Sei $v_1 = \frac{1}{||v_1'||} v_1' = \begin{pmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{pmatrix}$ und $v_2 = \frac{1}{||v_2'||} v_2' = \begin{pmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{pmatrix}$. Dann ist (v_1, v_2) eine Orthonormalbasis für $(\mathbb{C}^2, \langle , \rangle)$ aus Eigenvektoren von A. Für

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & i/\sqrt{2} \end{pmatrix}^{-1} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & i/\sqrt{2} \end{pmatrix}^{*} = \begin{pmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & -i/\sqrt{2} \end{pmatrix},$$

gilt also

$$PAP^* = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

Satz 11.17 Es sei $(V, \langle , \rangle)$ ein unitärer Vektorraum. Dann sind alle Eigenwerte eines selbstadjungierten Endomorphismus f von V reell.

Beweis : Die Aussage folgt aus dem Spektralsatz, wir können sie aber auch direkt zeigen. Ist v ein Eigenvektor von f zum Eigenwert α, so gilt nämlich $\langle f(v), v \rangle = \langle v, f(v) \rangle$, da f selbstadjungiert ist. Also folgt

$$\overline{\alpha} \langle v, v \rangle = \langle \alpha v, v \rangle = \langle f(v), v \rangle = \langle v, f(v) \rangle = \langle v, \alpha v \rangle = \alpha \langle v, v \rangle.$$

Da $v \neq 0$ ist, ist $\langle v, v \rangle \neq 0$, also ist $\alpha \in \mathbb{R}$. □

Insbesondere sind also auch alle Eigenwerte einer hermite'schen Matrix reell. Also hat jede symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ nur reelle Eigenwerte, d.h. das charakteristische Polynom $\chi_A(X)$ zerfällt über $\mathbb{R}$ vollständig in Linearfaktoren.

Ganz analog zum Spektralsatz Korollar 11.15 für unitäre Vektorräume können wir einen Spektralsatz für euklidische Vektorräume zeigen. Dazu betrachten wir jetzt für kurze Zeit wieder einen endlich-dimensionalen euklidischen Vektorraum $(V, \langle , \rangle)$. (Zur Erinnerung: V ist also ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit einer positiv definiten symmetrischen Bilinearform $\langle , \rangle$). Dann gilt analog zu Satz 11.13:

Satz 11.18 Es sei $f : V \to V$ eine lineare Abbildung und B eine Orthonormalbasis von $(V, \langle , \rangle)$. Dann gilt für die Koordinatenmatrix $A = A_{f,B,B}$:

i) A ist genau dann symmetrisch, wenn $\langle v, f(w)\rangle = \langle f(v), w\rangle$ für alle $v, w \in V$ gilt. Die lineare Abbildung f heißt dann selbstadjungiert bezüglich $\langle , \rangle$.

ii) A ist genau dann orthogonal, wenn $\langle v, w\rangle = \langle f(v), f(w)\rangle$ für alle $v, w \in V$ gilt, also genau dann, wenn f eine Isometrie von $(V, \langle , \rangle)$ ist.

Beweis : Da B eine Orthonormalbasis ist, ist die Koordinatenmatrix von $\langle , \rangle$ bezüglich B die Einheitsmatrix. Nach Lemma 9.4 gilt also $\langle x, y\rangle = \kappa_B(x)^t \kappa_B(y)$ für alle $x, y \in V$. Ferner gilt nach Proposition 6.11 $\kappa_B(f(x)) = A\kappa_B(x)$ für alle $x \in V$.
Mit diesen Vorüberlegungen können wir dann genauso argumentieren wie in Satz 11.13, wenn wir überall die komplexe Konjugation vergessen. ☐

Satz 11.19 (**Spektralsatz für euklidische Vektorräume**, auch Hauptachsentransformation genannt)

i) Sei $(V, \langle , \rangle)$ ein endlich-dimensionaler euklidischer Vektorraum und $f : V \to V$ eine selbstadjungierte Abbildung, d.h. es gilt $\langle f(v), w\rangle = \langle v, f(w)\rangle$ für alle $v, w \in V$. Dann gibt es eine Orthonormalbasis von $(V, \langle , \rangle)$, die aus Eigenvektoren von f besteht. Insbesondere ist f diagonalisierbar.

ii) Sei A eine symmetrische Matrix in $\mathbb{R}^{n\times n}$. Dann gibt es eine orthogonale Matrix $P \in O(n, \mathbb{R})$, so dass
$$PAP^t$$
eine Diagonalmatrix ist.

Beweis :

i) Wir argumentieren wie im komplexen Fall mit Induktion nach $n = \dim V$. Für $n = 0$ ist nichts zu zeigen. Wir nehmen also an, die Behauptung stimmt für Vektorräume der Dimension $< n$. Nach der Bemerkung zu Satz 11.17 hat f nur reelle Eigenwerte. Sei b_1 ein Eigenvektor von f zu einem Eigenwert α, den wir so normieren, dass $||b_1|| = 1$ gilt. Mit dem Gram-Schmidt-Verfahren ergänzen wir b_1 zu einer Orthonormalbasis $B = (b_1, \dots, b_n)$ von V.

Dann ist $A_{f,B,B} = \begin{pmatrix} a & * & \dots & * \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$ mit einem $A' \in \mathbb{R}^{(n-1)\times(n-1)}$. Da f selbstadjungiert ist, ist nach Satz 11.18 die Matrix $A_{f,B,B}$ symmetrisch. Also gilt

$$A_{f,B,B} = \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{pmatrix}$$

mit einer symmetrischen Matrix $A' \in \mathbb{R}^{(n-1)\times(n-1)}$. Daher ist die Einschränkung von f auf den Unterraum $W = \langle b_2, \dots, b_n \rangle$ nach Satz 11.18 selbstadjungiert. Nach Induktionsvoraussetzung existiert also eine Orthonormalbasis $C = (c_2, \dots, c_n)$ von $\langle , \rangle$ auf W, so dass $A_{f|_W,C,C}$ eine Diagonalmatrix ist. Die Basis $(b_1, c_2, \dots, c_n)$ leistet das Verlangte.

ii) folgt durch Anwenden von i) auf die lineare Abbildung $\lambda_A : \mathbb{R}^n \to \mathbb{R}^n$, die selbstadjungiert bezüglich des kanonischen Skalarproduktes ist.

$\square$

Im euklidischen Fall gilt auch das Analogon zu Korollar 11.15, d.h. das Skalarprodukt $\langle , \rangle$ lässt sich gleichzeitig mit einer symmetrischen Bilinearform diagonalisieren.
Für euklidische Vektorräume gibt der Spektralsatz 11.19 eine erschöpfende Antwort auf die Frage, für welche Endomorphismen es eine Basis aus Eigenvektoren gibt: Es sind genau die selbstadjungierten. Gibt es nämlich für einen Endomorphismus f eine Orthonormalbasis B aus Eigenvektoren, so ist $A_{f,B,B}$ eine Diagonalmatrix, also symmetrisch. Nach Satz 11.18 ist f daher selbstadjungiert. Für unitäre Vektorräume gibt es allerdings eine größere Klasse von Endomorphismen, für die es eine Basis aus Eigenvektoren gibt. Es ist nicht jede komplexe Diagonalmatrix hermite'sch, daher funktioniert das obige Argument nicht.

Proposition 11.20 Es sei $A \in \mathbb{C}^{n\times n}$ eine komplexe Matrix, für die eine unitäre Matrix $P \in U(n, \mathbb{C})$ existiert, so dass

$$PAP^*$$

eine (komplexe) Diagonalmatrix ist. Dann folgt $AA^* = A^*A$.

Beweis : Ist $PAP^* = D$ eine Diagonalmatrix mit komplexen Einträgen, so ist $(PAP^*)^* = PA^*P^* = D^*$ ebenfalls eine. Da die Multiplikation von Diagonalma-

trizen kommutativ ist, gilt $DD^* = D^*D$. Es ist $A = P^{-1}D(P^*)^{-1} = P^*DP$ und $A^* = P^{-1}D^*(P^*)^{-1} = P^*D^*P$.
Also folgt

$$AA^* = P^*DPP^*D^*P = P^*DD^*P = P^*D^*DP = P^*D^*PP^*DP = A^*A.$$

□

Definition 11.21 Eine Matrix $A \in \mathbb{C}^{n\times n}$ mit $AA^* = A^*A$ heißt **normal**.

Wir haben also gesehen: Gilt für eine komplexe Matrix der Spektralsatz, so ist sie normal. Im folgenden wollen wir zeigen, dass auch die Umkehrung gilt, d.h. wir wollen den Spektralsatz auch für normale Matrizen beweisen. Zunächst beweisen wir einige Tatsachen über normale Matrizen.

Lemma 11.22 Sei $A \in \mathbb{C}^{n\times n}$ und $P \in U(n,\mathbb{C})$. Dann ist A genau dann normal, wenn PAP^* normal ist.

Beweis : Das zeigt man genau wie Proposition 11.20. □

Satz 11.23

i) Unitäre Matrizen sind normal.

ii) Hermite'sche Matrizen sind normal.

iii) Schiefhermite'sche Matrizen (d.h. $A^* = -A$) sind normal.

iv) Diagonalmatrizen sind normal.

Beweis :

i) Ist $P \in U(n,\mathbb{C})$, so gilt $PP^* = E = P^*P$.

ii)-iv) ist klar.

□

Lemma 11.24 Sei $\langle , \rangle$ das kanonische Skalarprodukt auf $\mathbb{C}^n$ und A eine **beliebige** Matrix in $\mathbb{C}^{n\times n}$. Dann gilt für alle $v, w \in \mathbb{C}^n$: $\langle Av, w\rangle = \langle v, A^*w\rangle$.

Beweis : Es ist $\langle Av, w\rangle = \overline{(Av)}^t w = \overline{v}^t\overline{A}^t w = \overline{v}^t A^* w = \langle v, A^*w\rangle$. □

Lemma 11.25 Ist $A \in \mathbb{C}^{n\times n}$ normal, so gilt für alle $v \in \mathbb{C}^n$:

$$||A|| = ||A^*v||,$$

wobei $||w|| = \sqrt{\langle w, w\rangle}$ die Länge zum kanonischen Skalarprodukt ist.

Beweis : Es ist $\langle Av, Av\rangle \overset{11.24}{=} \langle v, A^*Av\rangle \overset{A \text{ normal}}{=} \langle v, AA^*v\rangle \overset{11.24}{=} \langle A^*v, A^*v\rangle$, woraus die Behauptung folgt. Wir haben im letzten Schritt die Gleichung $(A^*)^* = A$ benutzt. □

Lemma 11.26 Ist $A \in \mathbb{C}^{n\times n}$ normal und $v \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$, so ist v ein Eigenvektor von A^* zum Eigenwert $\overline{\lambda}$.

Beweis : Ist $Av = \lambda v$ mit $v \neq 0$ aus $\mathbb{C}^n$, so ist $v \in \text{ Kern}(A - \lambda E_n)$. Mit A ist auch $(A - \lambda E_n)$ normal, da $(A - \lambda E_n)^* = A^* - \overline{\lambda} E_n$ gilt. (Prüfen Sie das!) Also ist $||(A^* - \overline{\lambda} E_n)v|| \overset{11.25}{=} ||(A - \lambda E_n)v|| = 0$, woraus $(A^* - \overline{\lambda} E_n)v = 0$ folgt. Somit ist $A^*v = \overline{\lambda} v$. □

Satz 11.27 Sei $A \in \mathbb{C}^{n\times n}$ normal. Dann sind Eigenvektoren zu verschiedenen Eigenwerten von A orthogonal zueinander.

Beweis : Gilt $Av = \lambda v$ und $Aw = \mu w$ mit $v \neq 0$ und $w \neq 0$ aus $\mathbb{C}^n$, so folgt

$$\langle Av, w\rangle = \langle \lambda v, w\rangle = \overline{\lambda}\langle v, w\rangle.$$

Andererseits gilt

$$\langle Av, w\rangle \overset{11.24}{=} \langle v, A^*w\rangle \overset{11.26}{=} \langle v, \overline{\mu} w\rangle = \overline{\mu}\langle v, w\rangle.$$

Aus $\langle v, w\rangle \neq 0$ folgt also $\overline{\lambda} = \overline{\mu}$, und damit $\lambda = \mu$. □

Satz 11.28 (Spektralsatz für normale Matrizen) Ist $A \in \mathbb{C}^{n\times n}$ eine normale Matrix, so existiert eine unitäre Matrix $P \in U(n, \mathbb{C})$, so dass PAP^* eine Diagonalmatrix ist.

Beweis : mit Induktion nach n. Ist $n = 1$, so ist A bereits eine Diagonalmatrix. Wir nehmen also an, die Behauptung gilt für normale Matrizen in $\mathbb{C}^{(n-1)\times(n-1)}$. Sei $\langle , \rangle$ das kanonische Skalarprodukt auf $\mathbb{C}^n$.
Sei b_1 ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$. Nach Normieren können wir $\langle b_1, b_1\rangle = 1$ annehmen.
Wir ergänzen b_1 zu einer Orthonormalbasis $B = (b_1, \dots, b_n)$ von $(\mathbb{C}^n, \langle , \rangle)$. Es sei P_1 die Übergangsmatrix von der kanonischen Basis zu B. Wie im Beweis von Satz 11.14 ii) folgt $P_1 \in U(n, \mathbb{C})$.

Die Koordinatenmatrix der durch A vermittelten linearen Abbildung bezüglich der Basis $(b_1, \dots, b_n)$ ist also

$$A_1 := P_1 A P_1^{-1} = P_1 A P_1^*.$$

Sie hat die Form

$$A_1 = \begin{pmatrix} \lambda & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}$$

für ein $C \in \mathbb{C}^{(n-1)\times(n-1))}$. Mit A ist nach Lemma 11.22 auch A_1 normal. Es ist

$$A_1^* = \begin{pmatrix} \overline{\lambda} & 0 & \dots & 0 \\ \overline{a_{12}} & & & \\ \vdots & & C^* & \\ \overline{a_{1n}} & & & \end{pmatrix}$$

Wir berechnen die Koeffizienten an der Stelle $(1,1)$ in $A_1 A_1^* = A_1^* A_1$ und erhalten

$$|\lambda|^2 + |a_{12}|^2 + \dots + |a_{1n}|^2 = |\lambda|^2,$$

woraus $|a_{12}|^2 + \dots + |a_{1n}|^2 = 0$, also $a_{12} = \dots = a_{1n} = 0$ folgt.

Also ist $A_1 = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}$ mit einem normalen C. Nach Induktionsvoraussetzung existiert ein $Q \in U(n-1, \mathbb{C})$, so dass QCQ^* diagonal ist. Also ist für $P_2 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{pmatrix} \in U(n, \mathbb{C})$ auch $P_2 A_1 P_2^*$ diagonal, woraus die Behauptung folgt. □

Aus Satz 11.28 kann man einen Spektralsatz für normale Eindomorphismen eines beliebigen Vektorraums folgern, analog zu Satz 11.14 i).

Definition 11.29 Für eine Matrix $A \in K^{n\times n}$ schreiben wir $\sigma(A) = \{\lambda \in K : \lambda \text{ Eigenwert von } A\}$ für die Menge der Eigenwerte von A. Wir nennen $\sigma(A)$ das **Spektrum** von A.

Korollar 11.30 Für eine normale Matrix $A \in \mathbb{C}^{n\times n}$ gilt:

i) A hermite'sch $\Leftrightarrow \lambda \in \mathbb{R}$ für alle $\lambda \in \sigma(A)$.

ii) A positiv definit hermite'sch $\Leftrightarrow \lambda \in \mathbb{R}, \lambda > 0$ für alle $\lambda \in \sigma(A)$.

iii) A unitär $\Leftrightarrow |\lambda| = 1$ für alle $\lambda \in \sigma(A)$.

Beweis : Nach Satz 11.28 gibt es ein $P \in U(n, \mathbb{C})$, so dass PAP^* eine Diagonalmatrix ist. Da P unitär ist, gilt $PAP^* = PAP^{-1}$, woraus $\sigma(A) = \sigma(PAP^*)$ folgt. Auf der Diagonalen in PAP^* stehen also genau die Eigenwerte von A. Somit ist

$$\begin{aligned} PAP^* \text{ hermite'sch} &\Leftrightarrow \lambda \in \mathbb{R} \text{ für alle } \lambda \in \sigma(A), \\ PAP^* \text{ positiv definit hermite'sch} &\Leftrightarrow \lambda \in \mathbb{R}, \lambda > 0 \text{ für alle } \lambda \in \sigma(A) \text{ und} \\ PAP^* \text{ unitär} &\Leftrightarrow |\lambda|^2 = 1 \text{ für alle } \lambda \in \sigma(A). \end{aligned}$$

Da A genau dann hermite'sch/positiv definit hermite'sch/unitär ist, wenn die entsprechende Eigenschaft auf PAP^* zutrifft, folgt die Behauptung. □

12 Die Exponentialabbildung für Matrizen

Aus der Analysis wohlbekannt ist die Exponentialreihe $e^x = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ für $x \in \mathbb{C}$.
Wir wollen jetzt für eine Matrix $A \in \mathbb{C}^{n \times n}$ eine Matrix e^A durch die analoge Reihe

$$e^A = E_n + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots$$

definieren. Die einzelnen Summanden sind Matrizen. Um der Reihe Sinn zu geben, brauchen wir einen Konvergenzbegriff für Matrizen.

Definition 12.1 Sei $A^{(k)} = (A_{ij}^{(k)})_{i,j}$ eine Folge von Matrizen in $\mathbb{C}^{n \times n}$.

i) $A^{(k)}$ konvergiert gegen $A = (A_{ij})_{i,j} \in \mathbb{C}^{n \times n}$, falls für alle $i, j \in \{1, \dots, n\}$ die Folge $A_{ij}^{(k)}$ gegen A_{ij} konvergiert.

ii) Die Reihe $\sum\limits_{k=0}^{\infty} A^{(k)}$ konvergiert, falls für alle $i, j \in \{1, \dots, n\}$ die Reihe $\sum\limits_{k=0}^{\infty} A_{i,j}^{(k)}$ konvergiert.

Definition 12.2 Für $A = (A_{ij}) \in \mathbb{C}^{n \times n}$ definieren wir die Norm von A als

$$||A|| = \max_{i,j} |A_{ij}|$$

Lemma 12.3 Für $A, B \in \mathbb{C}^{n\times n}$ gilt

i) $||AB|| \leq n||A||\,||B||$

ii) $||A^k|| \leq n^{k-1}||A||^k$ für alle $k \geq 0$.

Beweis :

i) Sei $A = (A_{ij})_{i,j}$ und $B = (B_{ij})_{i,j}$. Für den Eintrag $(AB)_{ij}$ von AB an der Stelle (i, j) gilt

$$\begin{aligned} |(AB)_{ij}| &= \left|\sum_{k=1}^{n} A_{ik}B_{kj}\right| \\ &\leq \sum_{k=1}^{n} |A_{ik}|\,|B_{kj}| \\ &\leq n\,||A||\,||B||. \end{aligned}$$

ii) folgt mit Induktion aus i).

□

Satz 12.4 Für jedes $A \in \mathbb{C}^{n\times n}$ konvergiert die Reihe $e^A = E_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots$.

Beweis : Nach Lemma 12.3 gilt für alle $k \geq 1$:

$$\left|\left(\frac{A^k}{k!}\right)_{i,j}\right| \leq \frac{n^{k-1}}{k!}||A||^k = \frac{c^k}{nk!}$$

für $c = n\,||A|| \in \mathbb{R}$. Also hat für alle $i, j \in \{1, \dots, n\}$ die Reihe

$$|(E_n)_{i,j}| + |A_{ij}| + \left|\left(\frac{A^2}{2!}\right)_{ij}\right| + \dots$$

die konvergente Majorante

$$1 + \frac{1}{n}\sum_{k=1}^{\infty}\frac{c^k}{k!}.$$

□

Da die Multiplikation von Matrizen recht kompliziert ist, ist es im allgemeinen nicht so einfach, e^A auszurechnen. Ist A eine Diagonalmatrix, so ist es aber recht einfach.

Lemma 12.5 Ist $A = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{C}^{n\times n}$ diagonal, so ist $e^A = \begin{pmatrix} e^{d_1} & & 0 \\ & \ddots & \\ 0 & & e^{d_n} \end{pmatrix}$.

Beweis : Das folgt sofort aus

$$A^k = \begin{pmatrix} d_1^k & & 0 \\ & \ddots & \\ 0 & & d_n^k \end{pmatrix} \text{ für alle } k \geq 1.$$

□

Proposition 12.6 Ist $A = T^{-1}BT$ für ein $T \in \mathrm{GL}(n, \mathbb{C})$, so ist

$$e^A = T^{-1}e^B T.$$

Beweis : Es ist $A^k = T^{-1}B^kT$ für alle $k \geq 1$, also ist für alle $N \in \mathbb{N}$

$$\begin{aligned} & E_n + A + \frac{A^2}{2!} + \cdots + \frac{A^N}{N!} \\ = \ & E_n + T^{-1}BT + \frac{T^{-1}B^2T}{2!} + \cdots + \frac{T^{-1}B^NT}{N!} \\ = \ & T^{-1}(E_n + B + \frac{B^2}{2!} + \cdots + \frac{B^N}{N!})T. \end{aligned}$$

Die linke Seite konvergiert für $N \to \infty$ gegen e^A, die rechte gegen $T^{-1}e^BT$. □

Für eine diagonaliserbare Matrix A können wir mit Lemma 12.5 und Proposition 12.6 das Exponential e^A also leicht berechnen!
Eine grundlegende Eigenschaft der komplexen Exponentialfunktion e^x ist die Funktionalgleichung

$$e^{x+y} = e^x e^y$$

für alle $x, y \in \mathbb{C}$. Wir wollen jetzt untersuchen, wann die analoge Funktionalgleichung für die Exponentialfunktion auf $\mathbb{C}^{n\times n}$ gilt. Da die Multiplikation von Matrizen nicht kommutativ ist, kann man eine solche Funktionalgleichung nicht für beliebige Matrizen erwarten. Es gilt aber

Satz 12.7 Sind $A, B \in \mathbb{C}^{n\times n}$ mit $AB = BA$, so ist $e^{A+B} = e^A e^B$.
Insbesondere ist für jedes $A \in \mathbb{C}^{n\times n}$ die Matrix e^A invertierbar, und es gilt $(e^A)^{-1} = e^{-A}$.
Um Satz 12.7 zu beweisen, verwenden wir Methoden aus der Analysis. Dazu betrachten wir Matrizen $A(t) = (A(t)_{i,j})_{i,j}$, die von einem reellen Parameter t abhängen. Alle Koeffizienten $A(t)_{i,j}$ sind also Funktionen auf $\mathbb{R}$ mit Werten in $\mathbb{C}$.

Definition 12.8 Eine Funktion

$$\begin{array}{rcl} \mathbb{R} & \to & \mathbb{C}^{n\times n} \\ t & \mapsto & A(t) = (A(t)_{ij})_{i,j} \end{array}$$

heißt **differenzierbar** in $t_0 \in \mathbb{R}$, falls der Grenzwert

$$\frac{d}{dt}A(t)_{|t=t_0} = \lim_{h\to 0} \frac{A(t_0+h) - A(t_0)}{h}$$

existiert.

$A(t)$ ist also differenzierbar in t_0, falls für alle $i,j \in \{1,\ldots,n\}$ der Grenzwert

$$\frac{A(t_0+h)_{i,j} - A(t_0)_{i,j}}{h}$$

existiert, d.h. falls die Funktion

$$\begin{array}{rcl} \mathbb{R} & \to & \mathbb{C} \\ t & \mapsto & A(t)_{i,j} \end{array}$$

in t_0 differenzierbar ist.
Ist $A(t)$ differenzierbar in t_0, so ist also der Eintrag von $\frac{d}{dt}A(t)_{|t=t_0}$ an der Stelle (i,j) gerade die Ableitung $\frac{d}{dt}A(t)_{i,j|t=t_0}$.

Satz 12.9 Es sei $A \in \mathbb{C}^{n\times n}$. Die Funktion $t \mapsto e^{tA}$ ist differentierbar in allen $t \in \mathbb{R}$, und es gilt

$$\frac{d}{dt}e^{tA} = Ae^{tA}.$$

Beweis : Wir müssen zeigen, dass für alle $t \in \mathbb{R}$ gilt

$$\lim_{h\to 0} \frac{e^{(t+h)A} - e^{tA}}{h} = Ae^{tA}.$$

Wir setzen

$$R(t,h) = \left(\frac{e^{(t+h)A} - e^{tA}}{h}\right) - Ae^{tA},$$

und wollen $\lim\limits_{h\to 0} R(t,h) = 0$ zeigen.
Die Reihenentwicklung der Exponentialfunktion liefert (mit $E_n := A^0$):

$$\begin{aligned} R(t,h) &= \frac{1}{h}\left[\sum_{k=0}^{\infty} \frac{1}{k!}[(t+h)^k A^k - t^k A^k]\right] - \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^{k+1} \\ &= \sum_{k=2}^{\infty} \frac{1}{k!}\frac{1}{h}[(t+h)^k - t^k - kht^{k-1}]A^k \end{aligned}$$

Die Taylorentwicklung der Funktion $f(x) = x^k$ um t in zwei Terme mit Restglied liefert

$$\begin{aligned}(t+h)^k &= f(t+h) = f(t) + hf'(t) + h^2 f''(t+\theta_k h)\\ &= t^k + hkt^{k-1} + h^2 k(k-1)(t+\theta_k h)^{k-2}\end{aligned}$$

mit einem reellen $\theta_k \in [0,1]$.
Also ist

$$(t+h)^k - t^k - kht^{k-1} = h^2 k(k-1)(t+\theta_k h)^{k-2}.$$

Das setzen wir in $R(t,h)$ ein und erhalten

$$\begin{aligned}R(t,h) &= \sum_{k=2}^{\infty} \frac{1}{k!}\frac{1}{h}[h^2 k(k-1)(t+\theta_k h)^{k-2}]A^k\\ &= h\sum_{k=2}^{\infty}\frac{1}{(k-2)!}(t+\theta_k h)^{k-2}A^k.\end{aligned}$$

Also folgt für die einzelnen Einträge $R(t,h)_{i,j}$ dieser Matrix

$$\begin{aligned}|R(t,h)_{i,j} \quad | &\leq |h|\sum_{k=0}^{\infty}\frac{1}{k!}(|t|+|h|)^k|A_{i,j}^{k+2}|\\ &\leq |h|\sum_{k=0}^{\infty}\frac{1}{k!}(|t|+|h|)^k\ ||A^{k+2}||\\ &\overset{12.3}{=} |h|\sum_{k=0}^{\infty}\frac{1}{k!}(|t|+|h|)^k n^{k+1}||A||^{k+2}\\ &= |h|\ n\ ||A||^2 e^{(|t|+|h|)n||A||} \overset{h\to 0}{\longrightarrow} 0.\end{aligned}$$

Somit folgt $R(t,h) \overset{h\to 0}{\longrightarrow} 0$ in $\mathbb{C}^{n\times n}$. □

Jetzt können wir die Funktionalgleichung für e^A aus Satz 12.7 zeigen.
Beweis von Satz 12.7 :
Wir zeigen zunächst, dass für jedes $A \in \mathbb{C}^{n\times n}$ die Matrix e^A invertierbar ist mit Inversem e^{-A}. Dazu betrachten wir die Funktion $t \mapsto e^{tA}a^{-tA}$.
Für die Differentiation von Matrizen gilt ein Analogon der Produktregel. (Beweisen Sie das!). Also ist $e^{tA}e^{-tA}$ differenzierbar mit

$$\begin{aligned}\frac{d}{dt}e^{tA}e^{-tA} &= Ae^{tA}e^{-tA} - Ae^{tA}e^{-tA}\\ &= 0.\end{aligned}$$

Da die Einträge von $\frac{d}{dt}e^{tA}e^{-tA}$ die Ableitungen der Einträge von $e^{tA}e^{-tA}$ sind, sind alle Einträge der Matrix $e^{tA}e^{-tA}$ konstant in t. Wir setzen $t = 0$ ein und erhalten $e^{tA}e^{-tA} = E_n$ für alle $t \in \mathbb{R}$. Insbesondere folgt $e^{A}e^{-A} = E_n$.
Jetzt betrachen wir Matrizen $A, B \in \mathbb{C}^{n\times n}$ mit $AB = BA$. Wir wollen $e^{A+B} = e^{A}e^{B}$ zeigen. Dazu genügt es nach den obigen Überlegungen $e^{A+B}e^{-B}e^{-A} = E_n$ zu zeigen.
Wir betrachten die Funktion

$$e^{tA+tB}e^{-tB}e^{-tA}.$$

Mit der Produktregel und Satz 12.9 gilt

$$\begin{aligned}
&\frac{d}{dt}\left[e^{tA+tB}e^{-tB}e^{-tA}\right] \\
= \quad & (A+B)(e^{t(A+B)}e^{-tB}e^{-tA}) + e^{t(A+B)}\frac{d}{dt}(e^{-tB}e^{-tA}) \\
= \quad & (A+B)[e^{t(A+B)}e^{-tB}e^{-tA}] + e^{t(A+B)}[-Be^{-tB}e^{-tA} - e^{-tB}Ae^{-tA}]
\end{aligned}$$

Da A und B kommutieren, kommutieren auch A und e^{-tB}. Also ist

$$\frac{d}{dt}[e^{tA+tB}e^{-tB}e^{-tA}] = (A+B)[e^{t(A+B)}e^{-tB}e^{-tA}] - (A+B)[e^{t(A+B)}e^{-tB}e^{-tA}] = 0,$$

woraus wie oben folgt, dass $e^{tA+tB}e^{-tB}e^{-tA}$ konstant in t ist.
Durch Einsetzen von $t = 0$ erhalten wir

$$e^{tA+tB}e^{-tB}e^{-tA} = E_n.$$

Für $t = 1$ ergibt sich also $e^{A+B}e^{-B}e^{-A} = E_n$. ☐

Diesen Satz wollen wir jetzt anwenden, um für beliebiges $A \in \mathbb{C}^{n\times n}$ die Matrix e^{A} mit Hilfe der Jordanschen Normalform zu berechnen.
Für $A \in \mathbb{C}^{n\times n}$ sei $\sigma(A) = \{\lambda_1, \dots, \lambda_r\}$ und es sei

$$B = \begin{pmatrix} \lambda_1 E_{m_1} + N_1 & & \\ & & 0 \\ & \ddots & \\ 0 & & \\ & & \lambda_r E_{m_r} + N_r \end{pmatrix}$$

die Jordan'sche Normalform von A.

Also ist

$$N_i \quad = \quad \begin{pmatrix} J_{d_i} & & & & & & \\ & \ddots & & & & & \\ & & J_{d_i} & & & & \\ & & & \ddots & & & \\ & & & & J_1 & & \\ & & 0 & & & \ddots & \\ & & & & & & J_1 \end{pmatrix}$$

$$\underbrace{\qquad\qquad}_{s^{(i)}_{d_i}} \qquad\qquad \underbrace{\qquad\qquad}_{s^{(i)}_{1}}$$

Hier ist d_i der Exponent von $(X - \lambda_i)$ im Minimalpolynom von A. Insbesondere ist $N_i^{d_i} = 0$.

Also bricht die Exponentialreihe e^{N_i} ab, und es gilt

$$e^{N_i} = E_n + N_i + \frac{N_i^2}{2!} + \cdots + \frac{N_i^{d_{i-1}}}{(d_i - 1)!}.$$

Das lässt sich also ohne Grenzübergang berechnen. Da $\lambda_i E_{m_i}$ und N_i vertauschbare Matrizen sind, folgt nach Satz 12.7

$$\begin{aligned} e^{\lambda_i E_{m_i} + N_i} \quad &= \quad e^{\lambda_i E_{m_i}} e^{N_i} \\ &= \quad e^{\lambda_i} e^{N_i}. \end{aligned}$$

Für die Matrix $T \in \mathrm{GL}(n, \mathbb{C})$ mit $A = T^{-1}BT$ gilt also

$$e^A = T^{-1} e^B T = T^{-1} \begin{pmatrix} e^{\lambda_1} e^{N_1} & & \\ & \ddots & \\ & & e^{\lambda_r} e^{N_r} \end{pmatrix} T$$

Auf diese Weise lässt sich e^A berechnen.

In der Analysis lernt man, dass die Lösungen der linearen Differentialgleichung

$$\frac{df(t)}{dt} = af(t) \text{ für } a \in \mathbb{R}$$

von der Form $f(t) = ce^{at}$ für ein $c \in \mathbb{R}$ sind. Das wollen wir jetzt verallgemeinern.

Definition 12.10 Ein System homogener linearer Differentialgleichung erster Ordnung mit konstanten Koeffizienten ist eine Matrixgleichung der Form

$$\frac{dX(t)}{dt} = AX(t),$$

wobei $A \in \mathbb{C}^{n \times n}$ und $t \mapsto X(t)$ eine Funktion von $\mathbb{R}$ nach $\mathbb{C}^n$ ist.

Satz 12.11 Sei $A \in \mathbb{C}^{n \times n}$. Dann bilden die Spalten der Matrix e^{tA} eine Basis für den Vektorraum der Lösungen der Differentialgleichung $\frac{dX(t)}{dt} = AX(t)$.

Beweis : Die Menge aller Lösungen $X(t)$ des Systems $\frac{dX(t)}{dt} = AX(t)$ enthält offenbar die Nullfunktion und ist abgeschlossen unter der Addition. Also ist diese Lösungsmenge ein Unterraum des $\mathbb{C}$-Vektorraums aller Funktionen $\mathbb{R} \to \mathbb{C}^n$.
Mit Satz 12.9 folgt $\frac{d}{dt}e^{tA} = Ae^{tA}$. Lesen wir diese Gleichung spaltenweise, so folgen für die Spalten $(x_1(t), \ldots, x_n(t))$ von e^{tA} die gewünschten Differentialgleichungen $\frac{d}{dt}x_i(t) = Ax_i(t)$.
Sei nun $X(t)$ eine beliebige Lösungvon $\frac{dX(t)}{dt} = AX(t)$. Leiten wir $e^{-tA}X(t)$ mit der Produktregel ab, so folgt

$$\frac{d}{dt}(e^{-tA}X(t)) = (-A)e^{-tA}X(t) + e^{-tA}\frac{dX(t)}{dt} = (-Ae^{-tA} + e^{-tA}A)X(t).$$

Da in der Expontentialreihe von e^{-tA} als Summanden nur skalare Vielfache von Potenzen von A auftreten, ist $Ae^{-tA} = e^{-tA}A$. Also folgt $\frac{d}{dt}(e^{-tA}X(t)) = 0$. Alle Einträge des Spaltenvektors $e^{-tA}X(t)$ sind also konstante Funktion, d.h. $e^{-tA}X(t) = (c_1, \ldots, c_n)^t$. Somit ist $X(t) = e^{tA}(c_1, \ldots, c_n)^t$ eine Linearkombination der Spalten von e^{tA}. □

Möchte man die Lösungen des Systems $\frac{dX(t)}{dt} = AX(t)$ für ein $A \in \mathbb{C}^{n \times n}$ berechnen, so muss man also nach Satz 12.11 nur die Matrix e^{tA} berechnen. Dies lässt sich bewerkstelligen, indem man wie oben die Jordan'sche Normalform von A verwendet.

13 Universelle Konstruktionen

Wir wollen nun noch Summen, Dualräume, Quotientenräume und Tensorprodukte von Vektorräumen kennen lernen.
Ist V ein K-Vektorraum mit Unterräumen W_1 und W_2, so ist $W_1 + W_2$ ein Unterraum von V. Ist $W_1 \cap W_2 = 0$, so gilt $W_1 + W_2 = W_1 \oplus W_2$. Man kann aber auch für beliebige Vektorräume, die nicht als Unterräume desselben V gegeben sind, eine direkte Summe definieren.

Definition 13.1 Es seien zwei K-Vektorräume V_1 und V_2 gegeben. Wir definieren eine Addition und eine skalare Multiplikation auf der Produktmenge $V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$ durch:

i) $(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2)$

ii) $a(v_1, v_2) = (av_1, av_2)$.

Lemma 13.2 Die Abbildungen aus Definition 13.1 machen $V_1 \times V_2$ zu einem K-Vektorraum, den wir mit $V_1 \oplus V_2$ bezeichnen und **direkte Summe** von V_1 und V_2 nennen.

Beweis : Offenbar ist $V_1 \times V_2$ mit der komponentenweisen Addition eine abelsche Gruppe mit neutralem Element $(0,0)$ und Inversem $(-v_1, -v_2)$ zu (v_1, v_2). Es gilt ferner $1(v_1, v_2) = (v_1, v_2)$ und $(ab)(v_1, v_2) = a(b(v_1, v_2))$ sowie die Distributivgesetze. $\square$

Man kann diese Konstruktion auf endlich viele Vektorräume $V_1, \dots, V_n$ übertragen und so die direkte Summe $V_1 \oplus \cdots \oplus V_n$ definieren. Bei unendlich vielen Summanden muss man etwas aufpassen. Ist I eine beliebige Indexmenge und ist für jedes $i \in I$ ein K-Vektorraum V_i gegeben, so ist

$$\bigoplus_{i \in I} V_i = \{(v_i)_{i \in I} : v_i \in V_i, \text{ fast alle } v_i = 0\}$$

mit der komponentenweisen Addition und skalaren Multiplikation.

Lemma 13.3 Es seien K-Vektorräume V und W gegeben. Mit $\mathrm{Hom}_K(V, W)$ bezeichnen wir die Menge der linearen Abbildungen von V nach W. Dann ist $\mathrm{Hom}_K(V, W)$ zusammen mit der Addition

$$(f+g)(v) = f(v) + g(v)$$

und der skalaren Multiplikation

$$(af)(v) = af(v)$$

ein K-Vektorraum.

Beweis : Das folgt sofort aus den Vektorraumaxiomen für V. $\square$

Definition 13.4 Der K-Vektorraum $\mathrm{Hom}_K(V, K)$ heißt **Dualraum von V**, wir bezeichnen ihn mit $V^\vee$. Die Elemente aus $V^\vee$ heißen auch **Linearformen auf V**.

Lemma 13.5 Sind V und W endlich-dimensionale K-Vektorräume, so gilt $\dim_K \mathrm{Hom}(V, W) = \dim_K V \dim_K W$. Insbesondere ist $\dim_K V^\vee = \dim_K V$.

Beweis : Es sei $n = \dim_K V$ und $m = \dim_K W$. Wir wählen Basen $B = (b_1, \dots, b_n)$ von V und $C = (c_1, \dots, c_m)$ von W. Wir betrachten die Abbildung

$$\begin{array}{rcl} \Phi : \mathrm{Hom}_K(V, W) & \to & K^{m \times n} \\ f & \mapsto & A_{f,B,C}, \end{array}$$

die jeder linearen Abbildung ihre Koordinatenmatrix bezüglich B und C zuordnet. Da $A_{(f+g),B,C} = A_{f,B,C}$ und $A_{(\alpha f),B,C} = \alpha A_{f,B,C}$ gilt, ist Φ eine lineare Abbildung. Ist $A_{f,B,C} = 0$, so folgt $f = 0$, d.h. Kern$\Phi = 0$. Ist $A = (a_{ij})$ eine beliebige Matrix in $K^{m\times n}$, so gibt es eine lineare Abbildung $f : V \to W$ mit

$$f(b_j) = \sum_{i=1}^{m} a_{ij} c_i$$

für alle $j = 1, \ldots, n$. Dann ist $A_{f,B,C} = A$. Also ist Φ surjektiv und somit ein Isomorphismus von K-Vektorräumen, woraus $\dim_K \operatorname{Hom}_K(V,W) = \dim_K K^{m\times n} = mn$ folgt. ☐

Nun wollen wir den Quotienten eines Vektorraums nach einem Unterraum definieren.

Es sei V ein K-Vektorraum und $U \subset V$ ein Unterraum. Wir definieren eine Relation auf V durch

$$x \sim y \Leftrightarrow x - y \in U.$$

Lemma 13.6 $\sim$ ist eine Äquivalenzrelation.

Beweis : Es gilt $x \sim x$, also ist $\sim$ reflexiv. Ist $x \sim y$, d.h. $x - y \in U$, so ist auch $y - x = -(x-y) \in U$, also gilt $y \sim x$. Daher ist $\sim$ symmetrisch. Gilt $x \sim y$ und $y \sim z$, so sind $x - y$ und $y - z$ in U. Also ist auch $x - z = (x-y) + (y-z) \in U$, d.h. $x \sim z$. Somit ist $\sim$ auch transitiv. ☐

Beispiel: Ist $V = \mathbb{R}^3$ und $U = \langle e_1, e_2 \rangle = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}$, so ist $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \sim \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \Leftrightarrow$ $x_3 = y_3$.

Jedes $x \in V$ definiert eine Äquivalenzklasse

$$[x] = \{y \in V : x \sim y\} \subset V.$$

Es gilt: $[x] \cap [y] \neq \emptyset \Leftrightarrow [x] = [y] \Leftrightarrow x \sim y \Leftrightarrow x - y \in U$. Mit V/U bezeichnen wir die Menge der Äquivalenzklassen in V bezüglich $\sim$. Die Elemente von V/U sind also Mengen!

Beispiel: In obigem Beispiel ist die Äquivalenzklasse von

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ die Menge } [x] = \left\{ \begin{pmatrix} y_1 \\ y_2 \\ x_3 \end{pmatrix} y_1, y_2 \in \mathbb{R} \right\}.$$

Wir haben hier also eine Bijektion

$$\begin{array}{rcl} \tau:\mathbb{R} & \to & V/U \\ \lambda & \mapsto & \left[\begin{pmatrix}0\\0\\\lambda\end{pmatrix}\right] = \left\{\begin{pmatrix}y_1\\y_2\\\lambda\end{pmatrix} : y_1, y_2 \in \mathbb{R}\right\}. \end{array}$$

Nun definieren wir eine Addition auf der Menge V/U durch

$$[x] + [y] = [x + y]$$

Das ist wohldefiniert, denn aus $[x] = [x']$ und $[y] = [y']$ folgt $x \sim x'$ und $y \sim y'$. Also ist $x - x' \in U$ und $y - y' \in U$, somit auch $(x + y) - (x' + y') \in U$, d.h. $(x + y) \sim (x' + y')$. Daher gilt $[x + y] = [x' + y']$. Die Addition hängt also nicht von der Wahl der Vertreter x und y ab.
Analog definiert man eine skalare Multiplikation auf V/U durch $\lambda[x] = [\lambda x]$. Ist $[x] = [y]$, d.h. gilt $x - y \in U$, so ist auch $\lambda(x - y) \in U$, woraus $\lambda x - \lambda y \in U$, d.h. $[\lambda x] = [\lambda y]$ folgt. Daher ist auch die skalare Multiplikation auf V/U wohldefiniert.

Lemma 13.7 V/U zusammen mit der oben definierten Addition und skalaren Multiplikation ist ein Vektorraum.

Beweis : Übungsaufgabe. ☐

Beispiel: In obigem Beispiel ist

$$\begin{array}{rcl} \left[\begin{pmatrix}0\\0\\\lambda\end{pmatrix}\right] + \left[\begin{pmatrix}0\\0\\\mu\end{pmatrix}\right] & = & \left[\begin{pmatrix}0\\0\\\lambda+\mu\end{pmatrix}\right] \text{ und} \\ a\left[\begin{pmatrix}0\\0\\\lambda\end{pmatrix}\right] & = & \left[\begin{pmatrix}0\\0\\a\lambda\end{pmatrix}\right], \end{array}$$

so dass τ sogar ein Isomorphismus von Vektorräumen ist.

Die Abbildung $p : V \to V/U$, definiert durch $p(x) = [x]$, ist linear und surjektiv. Ferner gilt Kern$(p) = U$, denn es ist $p(x) = 0 \Leftrightarrow [x] = [0] \Leftrightarrow x \in U$. Daraus folgt mit der Dimensionsformel $\dim V/U = \dim V - \dim U$. Wir haben also „$U$ aus V herausgeteilt".

Satz 13.8 (Homomorphiesatz) Es sei $f : V \to W$ eine lineare Abbildung mit $U \subset \operatorname{Kern}(f)$. Dann existiert genau eine lineare Abbildung

$$g : V/U \to W,$$

so dass das Diagramm

$$\begin{tikzcd}
V \arrow[rr, "f"] \arrow[d, "p"'] & & W \\
V/U \arrow[rru, "g"'] & &
\end{tikzcd}$$

kommutiert, d.h. so dass $f = g \circ p$ gilt.

Beweis : Wir wollen eine Abbildung g konstruieren, für die $f = g \circ p$ gilt. Dazu haben wir gar keine Wahl, als $g([x]) = g \circ p(x) = f(x)$ zu setzen. Das ist wohldefiniert, da aus $[x] = [y]$ schon $x - y \in U \subset \operatorname{Kern} f$, d.h. $f(x) = f(y)$ folgt.
Außerdem gilt

$$g([x] + [y]) = g([x + y]) = f(x + y) = f(x) + f(y) = g([x]) + g([y])$$

und

$$g(a[x]) = g([ax]) = f(ax) = af(x) = ag[x].$$

Daher ist g linear. Ferner ist g durch die Gleichung $f = g \circ p$ offenbar eindeutig bestimmt. ☐

Beispiel: Es sei im obigen Beispiel $f : V = \mathbb{R}^3 \to \mathbb{R}^2$ die lineare Abbildung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_3 \\ 2x_3 \end{pmatrix}.$$

Dann ist $U \subset \operatorname{Kern} f$. Die Abbildung

$$\begin{array}{rcl} g : V/U & \to & \mathbb{R}^2 \\ \left[\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right] & \mapsto & \begin{pmatrix} x_3 \\ 2x_3 \end{pmatrix} \end{array}$$

leistet das Verlangte. g ist die Verknüpfung des Isomorphismus $\tau^{-1} : V/U \to \mathbb{R}$ mit der linearen Abbildung $\mathbb{R} \to \mathbb{R}^2$, die durch $\lambda \mapsto \begin{pmatrix} \lambda \\ 2\lambda \end{pmatrix}$ gegeben ist.

Nun wollen wir das Tensorprodukt von zwei Vektorräumen einführen. Es seien V und W zwei K-Vektorräume. Wir betrachten die Menge $I = V \times W$ als Indexmenge und setzen

$$F = \bigoplus_{(v,w)\in V\times W} K$$

F besteht also aus allen Tupeln $(\lambda_{(v,w)})_{v\in V, w\in W}$, für die fast alle Einträge $\lambda_{(v,w)}$ gleich Null sind. Für jedes $v_0 \in V$ und $w_0 \in V$ bezeichnen wir mit $v_0 \tilde{\otimes} w_0$ das Element $v_0 \tilde{\otimes} w_0 = (\lambda_{(v,w)})_{v\in V, w\in W}$, für das

$$\lambda_{(v,w)} = \begin{cases} 0 & \text{falls } v \neq v_0 \text{ oder } w \neq w_0 \\ 1 & \text{falls } v = v_0 \text{ und } w = w_0 \end{cases}$$

gilt. Dann ist $(v \tilde{\otimes} w)_{v\in V, w\in W}$ eine Basis von F.
Wir haben eine Abbildung von Mengen

$$\begin{aligned} t : V \times W &\to F \\ (v,w) &\mapsto v \tilde{\otimes} w. \end{aligned}$$

Auf beiden Seiten stehen Vektorräume, es ist also eine natürliche Frage, ob diese Abbildung bilinear ist.
Es sei $U \subset F$ der Untervektorraum, der von allen Vektoren die Form

$$\begin{array}{rcl} (v+v') \tilde{\otimes} w & - & v \tilde{\otimes} w - v' \tilde{\otimes} w, \\ (\lambda v) \tilde{\otimes} w & - & \lambda(v \tilde{\otimes} w), \\ v \tilde{\otimes} (w+w') & - & v \tilde{\otimes} w - v \tilde{\otimes} w', \\ v \tilde{\otimes} \lambda w & - & \lambda(v \tilde{\otimes} w) \end{array}$$

für beliebige $v, v' \in V$, $w, w' \in W$ sowie $\lambda \in K$ aufgespannt wird. Dann ist t genau dann bilinear, wenn $U = 0$ ist. Das wird im allgemeinen aber nicht der Fall sein. Zum Glück haben wir oben gelernt, wie wir störende Unterräume zunichte machen können.
Wir betrachten den Quotientenraum F/U und setzen

$$V \otimes_K W := F/U.$$

Dieser Vektorraum heißt **Tensorprodukt von *V* und *W*** (über K). Wir haben eine Abbildung

$$T : V \times W \xrightarrow{t} F \xrightarrow{p} F/U = V \otimes_K W,$$

die nach Konstruktion von $V \otimes_K W$ bilinear ist. So gilt etwa

$$\begin{aligned} & T(v+v',w) - T(v,w) - T(v',w) \\ = \;& p((v+v') \tilde{\otimes} w) - p(v \tilde{\otimes} w) - p(v' \tilde{\otimes} w) \\ = \;& p((v+v') \tilde{\otimes} w - v \tilde{\otimes} w - v' \tilde{\otimes} w) \\ = \;& 0, \end{aligned}$$

da $(v+v') \tilde{\otimes} w - v \tilde{\otimes} w - v' \tilde{\otimes} w \in U$ ist. Die anderen Bedingungen rechnet man analog nach.
Wir schreiben $v \otimes w = [v \tilde{\otimes} w]$ für das Bild in $F/U = V \otimes_K W$ des Elements $v \tilde{\otimes} w$. Elemente der Form $v \times w$ nennt man auch **reine Tensoren**. Beliebige Elemente aus $V \otimes_K W$ heißen **Tensoren**. Im allgemeinen ist nicht jeder Tensor ein reiner Tensor, sondern nur eine Summe von reinen Tensoren. Ein beliebiger Tensor ist nämlich von der Form $[f]$ für ein $f \in F$. Es ist

$$\begin{aligned} f &= (\lambda_{(v,w)})_{v \in V, w \in W} \\ &= \sum_{v \in V, w \in W} \lambda_{(v,w)} (v \tilde{\otimes} w). \end{aligned}$$

Dies ist eine endliche Summe, da fast alle $\lambda(v,w) = 0$ sind.
Daher gilt

$$\begin{aligned} [f] = \sum_{v \in V, w \in W} \lambda_{(v,w)} [v \tilde{\otimes} w] &= \sum_{v \in V, w \in W} \lambda_{(v,w)} v \otimes w. \\ &= \sum_{v \in V, w \in W} (\lambda_{(v,w)} v) \otimes w. \end{aligned}$$

In $V \otimes W$ gelten also folgende Rechenregeln:

1) Jedes Element in $V \otimes W$ ist eine endliche Summe $\sum_{i=1}^{n} v_i \otimes w_i$ für $v_i \in V, w_i \in W$

2) $v_1 \otimes w + v_2 \otimes w = (v_1 + v_2) \otimes w$

3) $v \otimes w_1 + v \otimes w_2 = v \otimes (w_1 + w_2)$

4) $\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$.

Beispiel: Ist $W = 0$, so ist $V \otimes_K 0 = 0$. Es gilt nämlich für alle $v \in V$:

$$0 = 0(v \otimes 0) = v \otimes 0,$$

also sind alle reinen Tensoren und damit alle Tensoren Null.

Wenn man mit dem Tensorprodukt arbeitet, so verwendet man statt der Konstruktion meist die folgende universelle Eigenschaft:

Satz 13.9 (Universelle Eigenschaft des Tensorprodukts) Es sei $\beta : V \times W \to Z$ eine beliebige bilineare Abbildung in einem Vektorraum Z. Dann existiert genau eine **lineare** Abbildung $f : V \otimes_K W \to Z$, die das Diagramm

$$\begin{tikzcd}
V \times W \arrow[d, "T"'] \arrow[dr, "\beta"] & \\
V \otimes_K W \arrow[r, "f"'] & Z
\end{tikzcd}$$

kommutativ macht, d.h. für die $\beta = f \circ T$ gilt.

Beweis : Da $(v \overset{\sim}{\otimes} w)_{v \in V, w \in W}$ eine Basis von F ist, gibt es eine lineare Abbildung $h : F \to Z$ mit $h(v \overset{\sim}{\otimes} w) = \beta(v, w)$. Nun ist β bilinear, also verschwindet h auf allen Erzeugern des Vektorraums U. Daher ist $U \subset \mathrm{Kern} h$. Nach dem Homomorphiesatz 13.8 gibt es somit eine lineare Abbildung $f : V \otimes_K W = F/U \to Z$, so dass

$$\begin{tikzcd}
F \arrow[d, "p"'] \arrow[dr, "h"] & \\
V \otimes_K W = F/U \arrow[r, "f"'] & Z
\end{tikzcd}$$

kommutiert. Jetzt schalten wir die Abbildung $V \times W \xrightarrow{t} F$ noch davor. Es gilt $h \circ t(v, w) = h(v \overset{\sim}{\otimes} w) = \beta(v, w)$, also kommutiert wegen $T = p \circ t$ auch das Diagramm

$$\begin{tikzcd}
V \times W \arrow[d, "T"'] \arrow[dr, "\beta"] & \\
V \otimes_K W \arrow[r, "f"'] & Z
\end{tikzcd}$$

Da für einen reinen Tensor $v \otimes w$ die Gleichung $f(v \otimes w) = f \circ T(v, w) = \beta(v, w)$ gilt, ist f durch dieses Diagramm auf reinen Tensoren eindeutig bestimmt. . $\square$

Beispiel: Zu der bilinearen Abbildung

$$\begin{array}{rcl} V \times K & \to & V \\ (v, \lambda) & \mapsto & \lambda v \end{array}$$

gibt es nach Satz 13.9 eine lineare Abbildung

$$f : V \otimes_K K \to V$$

mit $f(v \otimes \lambda) = \lambda v$. Diese ist surjektiv, da $f(v \otimes 1) = v$ ist. Wir zeigen nun, dass f auch injektiv ist. Sei $\sum_{i=1}^{n}(v_i \otimes \lambda_i)$ mit $v_i \in V, \lambda_i \in K$ ein beliebiges Element in $V \otimes_K K$. Es gilt

$$\begin{aligned}\sum_{i=1}^{n}(v_i \otimes \lambda_i) &= \sum_{i=1}^{n}\lambda_i(v_i \otimes 1) = \sum_{i=1}^{n}(\lambda_i v_i) \otimes 1 \\ &= \left(\sum_{i=1}^{n}\lambda_i v_i\right) \otimes 1.\end{aligned}$$

In diesem Beispiel ist also ausnahmsweise jeder Tensor ein reiner Tensor.
Gilt $0 = f\left(\sum_{i=1}^{n} v_i \otimes \lambda_i\right) = \sum_{i=1}^{n} \lambda_i v_i$, so ist auch $\left(\sum_{i=1}^{n} \lambda_i v_i\right) \otimes 1 = \sum_{i=1}^{n}(v_i \otimes \lambda_i)$ gleich 0. Daher ist f injektiv und somit

$$V \otimes_K K \simeq V.$$