§9 Determinanten

Definition.

Für $n \in \mathbb{N}$ sei $\mathcal{S}_n := \mathbf{S}(\{1,...,n\})$, die Gruppe aller Permutationen der Menge $\{1,...,n\}$.

Man nennt $\mathcal{S}_n$ die *symmetrische Gruppe* (vom Grad n).

Schreibweise für $\sigma \in \mathcal{S}_n$: $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$.

Beispiele ($\mathcal{S}_2$, $\mathcal{S}_3$).

1. $\mathcal{S}_2 = \{\mathrm{id}, \sigma\}$ mit $\sigma = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, d.h. $\sigma(1) = 2$ und $\sigma(2) = 1$. Wegen $\sigma^2 = \mathrm{id}$ ist $\sigma = \sigma^{-1}$.

2. Sei $\sigma := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \in \mathcal{S}_3$ und, $\rho := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \in \mathcal{S}_3$.

Dann $\sigma^2 = \mathrm{id}$ und $\rho^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ $[\rho^2(1) = \rho(3) = 2,\ \rho^2(2) = \rho(1) = 3,\ \rho^2(3) = \rho(2) = 1]$.

Weiter gilt $\rho^3(1) = \rho(2) = 1$, $\rho^3(2) = \rho(3) = 1$, $\rho^3(3) = \rho(1) = 3$, woraus $\rho^3 = \mathrm{id}$ und damit $\rho^{-1} = \rho^2$ folgt.

Definition. Für $1 \le i \ne j \le n$ sei $\tau_{i,j} \in \mathcal{S}_n$ definiert durch $\tau_{i,j}(k) := \begin{cases} k & \text{falls } k \ne i,j \\ j & \text{falls } k = i \\ i & \text{falls } k = j \end{cases}$.

$\tau \in \mathcal{S}_n$ heißt *Transposition*, wenn es $i \ne j$ mit $\tau = \tau_{i,j}$ gibt.

Für jede Transposition τ gilt: $\tau^2 = \mathrm{id}$ bzw. $\tau^{-1} = \tau$.

Satz 9.1. Sei $n \ge 1$.

(a) $|\mathcal{S}_n| = n!$.

(b) Zu jedem $\sigma \in \mathcal{S}_n$ gibt es $m \ge 0$ und Transpositionen $\tau_1, \dots, \tau_m$ mit $\sigma = \tau_1\tau_2 \dots \tau_m$ ($\sigma = \mathrm{id}$, falls $m = 0$).

Beweis:

(a) *Definition.* Für jede Menge M sei $S_n(M)$ die Menge aller injektiven Abbildungen von $\{1, \dots, n\}$ in M.
Für $\sigma : \{1, \dots, n-1\} \to M$ und $a \in M$ sei $\sigma_a^* : \{1, \dots, n\} \to M$ definiert durch $\sigma_a^*(i) := \begin{cases} a & \text{falls } i = n \\ \sigma(i) & \text{sonst} \end{cases}$.

Durch Induktion nach n zeigen wir: $|M| = n \Rightarrow |S_n(M)| = n!$.

1. $n = 1$: trivial. 2. $n > 1$: $S_n(M) = \bigcup_{a \in M} \{\sigma_a^* : \sigma \in S_{n-1}(M \setminus \{a\})\} \Rightarrow$

$\Rightarrow |S_n(M)| = \sum_{a \in M} |S_{n-1}(M \setminus \{a\})| \overset{\text{I.V.}}{=} \sum_{a \in M} (n-1)! = (n-1)! \cdot n = n!$.

(b) Induktion nach $k := \min\{l \le n : \sigma(i) = i \text{ für } l < i \le n\}$.

1. $k = 0$: Dann $\forall i \in \{1, ..., n\}(\sigma(i) = i)$ und somit $\sigma = \mathrm{id}$.

2. $k > 0$: Da σ injektiv ist, gilt dann $\sigma(k) < k$. Mit $\tau := \tau_{k,\sigma(k)}$ gilt außerdem $\tau\sigma(k) = k$ und $\tau\sigma(i) = \tau(i) = i$ für $i = k{+}1, \dots, n$. Nach I.V. haben wir deshalb $\tau\sigma = \tau_1 \dots \tau_m$, und somit $\sigma = \tau\tau_1 \dots \tau_m$.

Definition. Eine Transposition $\tau_{i,j} \in \mathcal{S}_n$ heißt *Nachbarnvertauschung*, falls $j = i{+}1$ oder $i = j{+}1$ ist.

Lemma 9.2. Jede Transposition τ ist das Produkt einer ungeraden Zahl von Nachbarnvertauschungen.

Beweis:

Sei $\tau = \tau_{i,j}$ mit $i < j$. Induktion nach $j - i$:

1. $j - i = 1$: Dann ist τ selbst eine Nachbarnvertauschung.

2. $j - i > 1$: Dann gilt $(*)\ \tau = \tau_{j-1,j} \circ \tau_{i,j-1} \circ \tau_{j-1,j}$ $(i < j{-}1)$, und nach I.V. ist $\tau_{i,j-1}$ Produkt einer ungeraden Zahl von Nachbarnvertauschungen. Daraus folgt die Behauptung für τ.

Definition.

Ein Paar (i, j) heißt *Fehlstand* der Permutation $\pi \in \mathcal{S}_n$, wenn $1 \le i < j \le n$ und $\pi(j) < \pi(i)$ ist.

Für jede Permutation $\pi \in \mathcal{S}_n$ sei $a(\pi)$ die Anzahl der Fehlstände von π, d.h.

$a(\pi) := |\{(i, j) : 1 \le i < j \le n \wedge \pi(j) < \pi(i)\}|$.

Ferner sei $\operatorname{sign}(\pi) := (-1)^{a(\pi)}$ *(Vorzeichen von π).*

π heißt *gerade* falls $a(\pi)$ gerade, d.h. $\operatorname{sign}(\pi) = 1$ ist. Andernfalls heißt π ungerade.

Lemma 9.3

Ist $\tau \in \mathcal{S}_n$ eine Nachbarnvertauschung, so gilt $a(\tau \circ \pi) \in \{a(\pi)+1, a(\pi)-1\}$ für jedes $\pi \in \mathcal{S}_n$.

Beweis:

Sei $\tau = \tau_{k,k+1}$ und $P := \{(i, j) : 1 \le i < j \le n\}$. Offenbar gibt es genau ein Paar $(i, j) \in P$ mit $\pi(i), \pi(j) \in \{k, k+1\}$; und für dieses Paar gilt: $\pi(j) < \pi(i) \Leftrightarrow (\tau \circ \pi)(i) < (\tau \circ \pi)(j)$. Für alle anderen Paare $(i, j) \in P$ gilt dagegen $\pi(j) < \pi(i) \Leftrightarrow (\tau \circ \pi)(j) < (\tau \circ \pi)(i)$.

Satz 9.4

(a) $\operatorname{sign}(\mathrm{id}) = 1$.

(b) $\operatorname{sign}(\pi \circ \sigma) = \operatorname{sign}(\pi) \cdot \operatorname{sign}(\sigma)$.

(c) $\operatorname{sign}(\tau) = -1$ für jede Transposition τ.

Beweis:

(b) HS. Ist τ eine Nachbarnvertauschung, so $\operatorname{sign}(\tau \circ \sigma) = -\operatorname{sign}(\sigma) = \operatorname{sign}(\tau) \cdot \operatorname{sign}(\sigma)$.

Beweis: $\operatorname{sign}(\tau \circ \sigma) = (-1)^{a(\tau\circ\sigma)} = (-1)^{a(\sigma)\pm 1} = -(-1)^{a(\sigma)} = -\operatorname{sign}(\sigma)$ und $\operatorname{sign}(\tau) = (-1)^1 = -1$.

Nach 9.1b und 9.2 gibt es Nachbarnvertauschungen $\tau_1, ..., \tau_k$ mit $\pi = \tau_1 \circ \ldots \circ \tau_k$.

Mit HS folgt also $\operatorname{sign}(\pi \circ \sigma) = \operatorname{sign}(\tau_1) \cdots \operatorname{sign}(\tau_k) \cdot \operatorname{sign}(\sigma) = \operatorname{sign}(\pi) \cdot \operatorname{sign}(\sigma)$.

(c) Nach 9.2 ist $\tau = \tau_1 \circ \ldots \circ \tau_k$ mit Nachbarnvertauschungen $\tau_1, ..., \tau_k$ und k ungerade.

Folglich $\operatorname{sign}(\tau) = \operatorname{sign}(\tau_1) \cdots \operatorname{sign}(\tau_k) = (-1)^k = -1$.

Folgerung.

Eine Permutation ist genau dann (un)gerade, wenn sie Produkt einer (un)geraden Zahl von Transpositionen ist.

Die geraden Permutationen $\pi \in \mathcal{S}_n$ bilden eine Untergruppe $\mathcal{A}_n$ von $\mathcal{S}_n$ *(alternierende Gruppe).*

Definition.

Sei R ein kommutativer Ring mit Eins, $n \ge 1$ und $d : R^{n\times n} \to R$.

Hat $A \in R^{n\times n}$ die Zeilen $a_1, ..., a_n \in R^{1\times n}$ so schreiben wir auch $d(a_1, ..., a_n)$ für $d(A)$.

1. d heißt *multilinear*, falls für $k = 1, ..., n$ und alle $a_i, x, y \in R^{1\times n}$, $\lambda \in R$ gilt:
   $d(a_1, ..., a_{k-1}, x + \lambda y, a_{k+1}, ..., a_n) = d(a_1, ..., a_{k-1}, x, a_{k+1}, ..., a_n) + \lambda \cdot d(a_1, ..., a_{k-1}, y, a_{k+1}, ..., a_n)$.
2. d heißt *alternierend* $:\Leftrightarrow$ $\forall a_1, ..., a_n \in R^{1\times n} \big[\exists i, j (1 \le i < j \le n \;\&\; a_i = a_j) \Rightarrow d(a_1, ..., a_n) = 0\big]$.
3. d heißt *normiert* $:\Leftrightarrow$ $d(E_n) = 1$.

Lemma 9.5. Ist $d : R^{n\times n} \to R$ multilinear und alternierend, so gilt für alle $A, A' \in R^{n\times n}$:

(a) Entsteht A' aus A durch Vertauschen zweier Zeilen, so ist $d(A') = -d(A)$.

(b) Entsteht A' aus A durch elementare Zeilenumformungen vom Typ II, so ist $d(A') = d(A)$.

(c) $d(A) = d(E_n) \cdot \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi)\cdot a_{1\pi(1)}\cdot a_{2\pi(2)} \cdots a_{n\pi(n)}$.

Beweis:

(a) $0 = d(a_1 + a_2, a_1 + a_2, a_3, \ldots, a_n) = d(a_1, a_1, a_3, \ldots, a_n) + d(a_2, a_2, a_3, \ldots, a_n) +$
$+ d(a_1, a_2, a_3, \ldots, a_n) + d(a_2, a_1, a_3, \ldots, a_n) = d(a_1, a_2, a_3, \ldots, a_n) + d(a_2, a_1, a_3, \ldots, a_n)$.

(b) $d(a_1 + \lambda a_j, a_2, \ldots, a_n) = d(a_1, a_2, \ldots, a_n) + \lambda \cdot d(a_j, a_2, \ldots, a_n) = d(a_1, a_2, \ldots, a_n) \quad (j \in \{2, ..., n\})$.

(c) $d(a_1, ..., a_n) = d(\sum_{i_1=1}^n a_{1i_1} e_{i_1}, ..., \sum_{i_n=1}^n a_{ni_n} e_{i_n}) =$
$\sum_{i_1=1}^n \cdots \sum_{i_n=1}^n a_{1i_1} \cdot \ldots \cdot a_{ni_n} \cdot d(e_{i_1}, ..., e_{i_n}) =$
$\sum_{\pi\in F_n} a_{1\pi(1)} \cdot \ldots \cdot a_{n\pi(n)} \cdot d(e_{\pi(1)}, ..., e_{\pi(n)}) \overset{(1)}{=}$
$\sum_{\pi\in\mathcal{S}_n} a_{1\pi(1)} \cdot \ldots \cdot a_{n\pi(n)} \cdot d(e_{\pi(1)}, ..., e_{\pi(n)}) \overset{(2)}{=}$
$\sum_{\pi\in\mathcal{S}_n} a_{1\pi(1)} \cdot \ldots \cdot a_{n\pi(n)} \cdot \operatorname{sign}(\pi) \cdot d(e_1, ..., e_n)$.

(1) Es sei F_n die Menge *aller* Abbildungen $\pi : \{1, ..., n\} \to \{1, ..., n\}$.
Die fragliche Gleichung folgt aus $\mathcal{S}_n = \{\pi \in F_n : \pi \text{ injektiv}\}$ und der Voraussetzung "d alternierend".

(2) Aus (a) folgt durch Induktion nach k: Ist $\pi = \tau_1 \cdots \tau_k$ mit Transpositionen $\tau_1, ..., \tau_k$, so
$d(e_{\pi(1)}, ..., e_{\pi(n)}) = (-1)^k\cdot d(e_1, ..., e_n) \overset{9.4}{=} \operatorname{sign}(\pi)\cdot d(e_1, ..., e_n)$.

Definition.

Für $A = (a_{ij}) \in R^{n\times n}$ sei $\det(A) := \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi)\cdot a_{1\pi(1)}\cdot a_{2\pi(2)} \cdots a_{n\pi(n)}$. (Leibnizsche Formel)

Satz 9.6.

(a) Die soeben definierte Funktion $\det : R^{n\times n} \to R$ ist multilinear, alternierend und normiert.
Man nennt sie *die Determinante*.

(b) Ist $d : R^{n\times n} \to R$ multilinear, alternierend und normiert, so gilt $d(A) = \det(A)$ für alle $A \in R^{n\times n}$.

Beweis:

(a) 1. $\det(a_1, \ldots, a_{k-1}, a_k{+}\lambda b_k, a_{k+1}, \ldots, a_n) = \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots (a_{k\pi(k)} + \lambda b_{k\pi(k)}) \cdots a_{n\pi(n)} =$
$\sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots a_{k\pi(k)} \cdots a_{n\pi(n)} + \lambda \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdots b_{k\pi(k)} \cdots a_{n\pi(n)} =$
$\det(a_1, \ldots, a_{k-1}, a_k, a_{k+1}, \ldots, a_n) + \lambda\det(a_1, \ldots, a_{k-1}, b_k, a_{k+1}, \ldots, a_n)$.

2. Sei $a_{1j} = a_{2j}$ für $j = 1, ..., n$. Für $\pi \in \mathcal{S}_n$ sei $\pi' := \pi \circ \tau_{1,2}$. Ferner $\mathcal{S}_n^* := \{\pi \in \mathcal{S}_n : \pi(1) < \pi(2)\}$.
Offenbar gilt $(*)$ $\mathcal{S}_n = \mathcal{S}_n^* \overset{\text{disj}}{\cup} \{\pi' : \pi \in \mathcal{S}_n^*\}$ und folglich
$\det(A) = \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)} \cdot a_{1\pi(2)} \cdots a_{n\pi(n)} \overset{(*)}{=}$
$\sum_{\pi\in\mathcal{S}_n^*} \operatorname{sign}(\pi) a_{1\pi(1)} \cdot a_{1\pi(2)} \cdots a_{n\pi(n)} + \sum_{\pi\in\mathcal{S}_n^*} \operatorname{sign}(\pi') a_{1\pi'(1)} \cdot a_{1\pi'(2)} \cdots a_{n\pi'(n)} =$
$\sum_{\pi\in\mathcal{S}_n^*} \operatorname{sign}(\pi) a_{1\pi(1)} \cdot a_{1\pi(2)} \cdots a_{n\pi(n)} - \sum_{\pi\in\mathcal{S}_n^*} \operatorname{sign}(\pi) a_{1\pi(2)} \cdot a_{1\pi(1)} \cdots a_{n\pi(n)} = 0$.

3. $\det(E) = \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi) \delta_{1\pi(1)} \cdots \delta_{n\pi(n)} = 1$.

(b) folgt aus 9.5c.

Satz 9.7.

Ist $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \in R^{n\times n}$, wobei A und B quadratisch, so gilt $\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det(A)\cdot\det(B)$.

Beweis:

1. Seien A, C fest. Wir definieren $d : R^{l\times l} \to R$, $d(B) := \det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$.
d ist offenbar multilinear und alternierend.
Nach 9.5c gilt deshalb $d(B) = d(E)\cdot\det(B) = \det \begin{pmatrix} A & C \\ 0 & E \end{pmatrix}\cdot\det(B)$.

2. Sei $\begin{pmatrix} A & C \\ 0 & E \end{pmatrix} = (a_{ij})_{i,j}$ und $k := n-l$. Dann gilt $\det \begin{pmatrix} A & C \\ 0 & E \end{pmatrix} = \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi) a_{1\pi(1)}\cdot a_{2\pi(2)}\cdots a_{n\pi(n)}$
mit $a_{ij} = \delta_{ij}$ für $i \in \{k{+}1,\dots,n\}$, $j \in \{1,\dots,n\}$.
Es folgt $\det \begin{pmatrix} A & C \\ 0 & E \end{pmatrix} = \sum_{\sigma\in\mathcal{S}_k} \operatorname{sign}(\sigma) a_{1\sigma(1)}\cdot a_{2\sigma(2)}\cdots a_{k\sigma(k)} = \det(A)$.

Korollar

Ist $A = (a_{ij}) \in R^{n\times n}$ eine obere [bzw. untere] Dreiecksmatrix, d.h. $a_{ij} = 0$ für $i > j$ [bzw. $i < j$],
so gilt $\det(A) = a_{11}\cdot\ldots\cdot a_{nn}$.

Beweis durch Induktion nach n mittels Satz 9.7

Satz 9.8. $\det(A^{\mathtt{t}}) = \det(A)$

Beweis:

Sei $A = (a_{ij})$. Dann $A^{\mathtt{t}} = (a'_{ij})$ mit $a'_{ij} := a_{ji}$. Es folgt $\det(A^{\mathtt{t}}) = \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi)\cdot a'_{1\pi(1)}\cdots a'_{n\pi(n)} =$
$= \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi)\cdot a_{\pi(1)1}\cdots a_{\pi(n)n} = \sum_{\pi\in\mathcal{S}_n} \operatorname{sign}(\pi)\cdot a_{1\pi^{-1}(1)}\cdots a_{n\pi^{-1}(n)} \overset{(*)}{=} \det(A)$.
$(*)$ Wegen $\operatorname{sign}(\pi)\cdot\operatorname{sign}(\pi^{-1}) = \operatorname{sign}(\pi\circ\pi^{-1}) = 1$ ist $\operatorname{sign}(\pi^{-1}) = \operatorname{sign}(\pi)$.

Korollar. Die Abbildung $R^{n\times n} \ni A \mapsto \det(A)$ ist auch bzgl. der *Spalten* von A multilinear und alternierend.
Folglich gilt für alle $A, A' \in R^{n\times n}$:
(a) Entsteht A' aus A durch Vertauschen zweier Spalten, so ist $\det(A') = -\det(A)$.
(b) Entsteht A' aus A durch elementare Spaltenumformungen vom Typ II, so ist $\det(A') = \det(A)$.

Satz 9.9. Für $A, B \in R^{n\times n}$ gilt: $\det(AB) = \det(A)\cdot\det(B)$.

Beweis:

Wie man leicht nachrechnet (siehe unten) ist die Abbildung $d_B : R^{n\times n} \to R$, $d_B(A) := \det(AB)$ multilinear und alternierend. Nach 9.5c gilt also $\det(AB) = d_B(A) \overset{9.5c}{=} d_B(E)\cdot\det(A) = \det(B)\cdot\det(A)$.
1. $d_B(a_1 + \lambda a'_1, a_2, \dots, a_n) = \det((a_1+\lambda a'_1)B, a_2B, \dots, a_nB) = \det(a_1B + \lambda a'_1B, a_2B, \dots, a_nB) =$
$\det(a_1B, a_2B, \dots, a_nB) + \lambda{\cdot}\det(a'_1B, a_2B, \dots, a_nB) = d_B(a_1, a_2, \dots, a_n) + \lambda\cdot d_B(a'_1, a_2, ..., a_n)$.
2. Sind zwei Zeilen von A gleich, so auch die entsprechenden Zeilen von AB, also $d_B(A) = \det(AB) = 0$.

Definition. Sei $A = \big(a_{kl}\big)_{k,l} \in R^{n\times n}$ und $i, j \in \{1, ..., n\}$.

1. A_{ij} sei die $(n{-}1){\times}(n{-}1)$-Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.
2. A'_{ij} sei die $n{\times}n$-Matrix, die aus A entsteht, wenn man a_{ij} durch 1 und alle anderen Komponenten, die in der i-ten Zeile oder der j-ten Spalte stehen, durch 0 ersetzt.
3. A^*_{ij} sei die $n{\times}n$-Matrix, die aus A entsteht, wenn man die i-te Zeile durch $e_j := (\delta_{j1}\ \dots\ \delta_{jn})$ ersetzt.
4. $\widetilde{A} := \big(\widetilde{a}_{ij}\big)_{i,j} \in R^{n\times n}$ mit $\widetilde{a}_{ij} = \det(A'_{ji})$ *(Die zu A komplementäre Matrix)*

Lemma 9.10. Für jede Matrix $A \in R^{n\times n}$ gilt:
(a) $\det(A'_{ij}) = (-1)^{i+j}\det(A_{ij})$.
(b) $\det(A'_{ij}) = \det(A^*_{ij})$.
(c) $\widetilde{A}\cdot A = A\cdot\widetilde{A} = \det(A)\cdot E_n$.

Beweis:

(a) Durch $i-1$ Vertauschungen benachbarter Zeilen und $j-1$ Vertauschungen benachbarter Spalten kann man A'_{ij} auf die Form $B := \begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix}$ bringen. Also $\det(A'_{ij}) = (-1)^{i-1}\cdot(-1)^{j-1}\det(B) = (-1)^{i+j}\det(A_{ij})$.

(b) Durch Zeilenumformungen vom Typ II kann A^*_{ij} in A'_{ij} überführt werden. Also gilt $\det(A'_{ij}) = \det(A^*_{ij})$.

(c) Seien $a_1, ..., a_n$ die Zeilen von A. Dann gilt für alle $i, k \in \{1, ..., n\}$: $\sum_{j=1}^n a_{ij}\widetilde{a}_{jk} = \sum_{j=1}^n a_{ij}\det(A'_{kj}) \overset{(b)}{=}$
$= \sum_{j=1}^n a_{ij}\det(a_1, ..., a_{k-1}, e_j, a_{k+1}, ..., a_n) = \det(a_1, ..., a_{k-1}, \sum_{j=1}^n a_{ij}e_j, a_{k+1}, ..., a_n) =$
$= \det(a_1, ..., a_{k-1}, a_i, a_{k+1}, ..., a_n) = \det(A)\delta_{ik}$. Also $A\widetilde{A} = \det(A)E_n$. Ebenso zeigt man $\widetilde{A}A = \det(A)E_n$.

Satz 9.11 (Entwicklungssatz von Laplace).
Für jede Matrix $A \in R^{n\times n}$ gilt:
(a) $\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij}\det(A_{ij})$ *(Entwicklung nach der i-ten Zeile)*
(b) $\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij}\det(A_{ij})$ *(Entwicklung nach der j-ten Spalte)*

Beweis von (a):
Aus $\det(A)\cdot E_n = A\cdot\widetilde{A}$ folgt: $\det(A) = \sum_{j=1}^n a_{ij}\widetilde{a}_{ji} = \sum_{j=1}^n a_{ij}\det(A'_{ij}) \overset{\text{L.9.10a}}{=} \sum_{j=1}^n (-1)^{i+j} a_{ij}\det(A_{ij})$.

Satz 9.12 (Die Vandermondsche Determinante).

Für alle $\alpha_0, ..., \alpha_n \in R$ gilt: $\det\begin{pmatrix} 1 & \alpha_0 & \alpha_0^2 & \dots & \alpha_0^n \\ 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^n \\ \vdots & & & & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \dots & \alpha_n^n \end{pmatrix} = \prod_{i<j\le n}(\alpha_j - \alpha_i)$.

Beweis durch Induktion nach n:

Sei D_n die linke Seite der behaupteten Gleichung.

1. $n = 1$: trivial. 2. $n > 1$: Durch Spaltenumformungen vom Typ II erhält man

$$D_n = \det\begin{pmatrix} 1 & \alpha_0 & \alpha_0^2 & \dots & \alpha_0^{n-1} & 0 \\ 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{n-1} & \alpha_1^{n-1}(\alpha_1-\alpha_0) \\ \vdots & & & & \vdots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \dots & \alpha_n^{n-1} & \alpha_n^{n-1}(\alpha_n-\alpha_0) \end{pmatrix} = \det\begin{pmatrix} \dots & 0 & 0 \\ \dots & \alpha_1^{n-2}(\alpha_1-\alpha_0) & \alpha_1^{n-1}(\alpha_1-\alpha_0) \\ \vdots & \vdots & \vdots \\ \dots & \alpha_n^{n-2}(\alpha_n-\alpha_0) & \alpha_n^{n-1}(\alpha_n-\alpha_0) \end{pmatrix} =$$

$$= \det\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & \alpha_1-\alpha_0 & \alpha_1(\alpha_1-\alpha_0) & \dots & \alpha_1^{n-2}(\alpha_1-\alpha_0) & \alpha_1^{n-1}(\alpha_1-\alpha_0) \\ \vdots & & & & \vdots & \vdots \\ 1 & \alpha_n-\alpha_0 & \alpha_n(\alpha_n-\alpha_0) & \dots & \alpha_n^{n-2}(\alpha_n-\alpha_0) & \alpha_n^{n-1}(\alpha_n-\alpha_0) \end{pmatrix}.$$

Durch Entwicklung nach der ersten Zeile und Herausziehen der Faktoren $(\alpha_i - \alpha_0)$ folgt daraus

$$D_n = \prod_{j=1}^n(\alpha_j-\alpha_0)\cdot\det\begin{pmatrix} 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{n-1} \\ \vdots & & & & \\ 1 & \alpha_n & \alpha_n^2 & \dots & \alpha_n^{n-1} \end{pmatrix} \overset{\text{I.V.}}{=} \prod_{j=1}^n(\alpha_j-\alpha_0)\cdot\prod_{i<j\le n-1}(\alpha_{j+1}-\alpha_{i+1}) = \prod_{i<j\le n}(\alpha_j-\alpha_i).$$

Satz 9.13 (Cramersche Regel).

Sei $A \in R^{n \times n}$ mit den Spalten $a_1, \ldots, a_n \in R^n$. Seien ferner $b \in R^n$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in R^n$, und gelte $A \cdot x = b$.

Dann gilt für $i = 1, \ldots, n$: $x_i \cdot \det(A) = \det(a_1 \ \ldots \ a_{i-1} \ b \ a_{i+1} \ \ldots \ a_n)$.

Beweis:

Aus 9.10c und $A{\cdot}x = b$ folgt $\det(A){\cdot}x = \widetilde{A}{\cdot}A{\cdot}x = \widetilde{A}{\cdot}b$ und somit $\det(A){\cdot}x_i = \sum_{j=1}^n \widetilde{a}_{ij}{\cdot}b_j \overset{(*)}{=}$
$= \sum_{j=1}^n \det(a_1 \ \ldots \ a_{i-1} \ e_j \ a_{i+1} \ \ldots \ a_n){\cdot}b_j = \det(a_1 \ \ldots \ a_{i-1} \ b \ a_{i+1} \ \ldots \ a_n)$. $(*)$ Die Matrizen A'_{ji} und $(a_1 \ \ldots \ a_{i-1} \ e_j \ a_{i+1} \ \ldots \ a_n)$ lassen sich durch elem. Spaltenumformungen vom Typ II ineinander überführen.

Satz 9.14 Sei K ein Körper und $A \in K^{n \times n}$.

(a) A invertierbar $\iff$ $\det(A) \neq 0$.

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

Beweis:

(a) "$\Leftarrow$": Sei A nicht invertierbar, d.h. (nach 7.1b) $\operatorname{rang}(A) < n$. Dann sind die Zeilen von A linear abhängig, und A kann durch Zeilenumformungen vom Typ II in eine Matrix A' umgeformt werden, bei der eine Zeile Null ist. Dann $\det(A) = \det(A') = 0$.

"$\Rightarrow$": $A \in \mathrm{GL}(K; n) \Rightarrow \det(A)\det(A^{-1}) \overset{9.9}{=} \det(AA^{-1}) = \det(E) = 1 \Rightarrow \det(A) \neq 0$.

(b) folgt aus dem Beweis von (a) "$\Rightarrow$".

Satz 9.15 und Definition (Die Determinante eines Endomorphismus).

Ist V ein n-dimensionaler K-Vektorraum, $f : V \to V$ ein Endomorphismus und sind $\overline{v}$, $\overline{w}$ Basen von V, so gilt $\det(\mathcal{M}_{\overline{v}}(f)) = \det(\mathcal{M}_{\overline{w}}(f))$.

Man kann deshalb definieren: $\det(f) := \det(\mathcal{M}_{\overline{v}}(f))$, wobei $\overline{v}$ irgendeine Basis von V sei.

Beweis:

Nach Korollar zu Satz 6.6 existiert ein $S \in \mathrm{GL}(n; K)$ mit $\mathcal{M}_{\overline{w}}(f) = S^{-1}{\cdot}\mathcal{M}_{\overline{v}}(f){\cdot}S$.

Mit 9.9 und 9.14b folgt daraus $\det(\mathcal{M}_{\overline{w}}(f)) = \det(S^{-1}){\cdot}\det(\mathcal{M}_{\overline{v}}(f)){\cdot}\det(S) = \det(\mathcal{M}_{\overline{v}}(f))$.

Lemma 9.16

Ist V endlichdimensionaler K-Vektorraum, so gilt für alle $f, g \in \mathrm{Hom}(V, V)$:

(a) $\det(f) \neq 0 \iff f$ injektiv $\iff f$ ist Isomorphismus.

(b) $\det(g \circ f) = \det(g) \cdot \det(f)$.

(c) $\det(\mathrm{id}_V) = 1$.

Beweis:

Sei $\overline{v}$ Basis von V, $A := \mathcal{M}_{\overline{v}}(f)$, $B := \mathcal{M}_{\overline{v}}(g)$.

(a) $\det(f) \neq 0 \Leftrightarrow \det(A) \neq 0 \Leftrightarrow A$ invertierbar $\Leftrightarrow f$ bijektiv.

(b) $\det(g \circ f) = \det(\mathcal{M}_{\overline{v}}(g \circ f)) = \det(BA) = \det(B)\det(A) = \det(g)\det(f)$.

(c) $\det(\mathrm{id}) = \det(E) = 1$.

§10 Ringe, Polynome

Sei $(R, +, \cdot)$ ein Ring mit Eins.

Definition.

Für $a \in R$ und $n \in \mathbb{N}$ sei a^n rekursiv definiert durch $a^0 := 1$ und $a^{n+1} := a^n \cdot a$.

Bemerkung. Für $a, b \in R$ und $m, n \in \mathbb{N}$ gilt: $a^{m+n} = a^m \cdot a^n$, $a^{mn} = (a^m)^n$ und, falls R kommutativ ist, $(a{\cdot}b)^n = a^n{\cdot}b^n$.

Definition.

$R^* := \{x \in R : \exists y \in R(xy = yx = 1)\}$. Die Elemente von R^* heißen *Einheiten* von R

Lemma 10.1.

(a) $a, b \in R^* \;\Rightarrow\; ab \in R^*$

(b) R^* mit der induzierten Verknüpfung $(a, b) \mapsto ab$ ist eine Gruppe.

Beweis:

(a) $a, b \in R^* \Rightarrow (\exists c, d \in R)\, ac = ca = 1 \;\&\; bd = db = 1 \Rightarrow (ab)(dc) = ac = 1 \;\&\; (dc)(ab) = db = 1 \Rightarrow ab \in R^*$.

(b) (i) $1 \in R^* \;\&\; \forall a \in R^*(1{\cdot}a = a)$. (ii) $a \in R^* \Rightarrow (\exists b \in R)\, ab = ba = 1 \;\Rightarrow\; b \in R^* \;\&\; ba = 1$.

Definition. Ein Ring $(R, +, \cdot)$ heißt *Integritätsring*, wenn gilt:

(i) R ist kommutativ.

(ii) R ist nullteilerfrei, d.h. $\forall x, y \in R(x \neq 0 \;\&\; y \neq 0 \Rightarrow xy \neq 0)$.

(iii) R besitzt ein Einselement $1 \neq 0$.

Bemerkung. In jedem Integritätsring gilt die *Kürzungsregel*: $a \cdot c = b \cdot c \;\&\; c \neq 0 \Rightarrow a = b$.

Definitionen.

Sei R ein kommutativer Ring mit Eins, und seien $a, b, a_1, ..., a_n \in R$.

$b|a \;:\Leftrightarrow\; \exists c \in R(a = bc)$ (b ist *Teiler* von a)

Ein Element d von R heißt *gemeinsamer Teiler* von $a_1, ..., a_n$, wenn $d|a_i$ für $i = 1, ..., n$.

Die Menge der gemeinsamen Teiler von $a_1, ..., a_n$ bezeichnen wir mit $gT(a_1, ..., a_n)$.

$a_1, ..., a_n$ heißen *teilerfremd*, wenn $gT(a_1, ..., a_n) \subseteq R^*$ gilt.
(Dann ist sogar $gT(a_1, ..., a_n) = R^*$: $e \in R^* \;\Rightarrow\; ee' = 1 \;\Rightarrow\; a_i = e(e'a_i)$.)

$(a_1, ..., a_n) := \{\sum_{i=1}^n x_i a_i : x_1, ..., x_n \in R\}$ heißt das *von $a_1, ..., a_n$ erzeugte Ideal.*

Definition (Euklidischer Ring)

Ein Integritätsring R heißt *euklidischer Ring*, wenn es eine Abbildung $\mathsf{d} : R \setminus \{0\} \to \mathbb{N}$ gibt, so daß gilt:
$\forall a, b \in R \setminus \{0\} \exists q, r \in R\big(a = qb + r \;\&\; (r \neq 0 \Rightarrow \mathsf{d}(r) < \mathsf{d}(b))\big)$.

Satz 10.2.

Ist R ein euklidischer Ring und sind $a_1, ..., a_n \in R$, so existiert ein $b \in R$ mit $(b) = (a_1, ..., a_n)$.

Beweis:

Sei $\mathfrak{a} := (a_1, ..., a_n) \neq \{0\}$. Dann existiert ein $b \in \mathfrak{a} \setminus \{0\}$ mit (1) $\forall x \in \mathfrak{a}(x \neq 0 \Rightarrow \mathsf{d}(b) \leq \mathsf{d}(x))$.

Wir zeigen $(b) = \mathfrak{a}$.

1. $(b) \subseteq \mathfrak{a}$ gilt wegen $b \in \mathfrak{a}$.

2. Sei $a \in \mathfrak{a}$. Dann gibt es $q, r \in R$ mit $a = qb + r$ und (2) $(r \neq 0 \Rightarrow \mathsf{d}(r) < \mathsf{d}(b))$. Es ist $r = a - qb \in \mathfrak{a}$. Wegen (1) gilt deshalb $(r \neq 0 \Rightarrow d(b) \leq d(r))$. Mit (2) folgt daraus $r = 0$. Also $a = qb \in (b)$.

Satz 10.3.
Sei R ein euklidischer Ring und seien $a_1, ..., a_n, b \in R$. Dann gilt:
(a) $a_1, ..., a_n$ teilerfremd $\iff$ $1 \in (a_1, ..., a_n)$.
(b) a_i, b teilerfremd für $i = 1, ..., n \implies a_1 \cdot \ldots \cdot a_n$, b teilerfremd.

Beweis:
(a) "$\Leftarrow$": $c \in gT(a_1, ..., a_n)$ & $1 \in (a_1, ..., a_n) \;\Rightarrow\; c|1 \;\Rightarrow\; c \in R^*$.
"$\Rightarrow$": Nach 10.2 existiert ein $a \in R$ mit $(a) = (a_1, ..., a_n)$.
$a_1, ..., a_n \in (a) \;\Rightarrow\; a \in gT(a_1, ..., a_n) \overset{\text{teilerfremd}}{\Rightarrow} a \in R^* \;\Rightarrow\; 1 \in (a)$.
(b) Nach Voraussetzung und (a) gibt es $x_1, y_1, ..., x_n, y_n \in R$, so daß $1 = x_i a_i + y_i b$ für $i = 1, ..., n$.
Daraus folgt $1 = (x_1 a_1 + y_1 b) \cdot \ldots \cdot (x_n a_n + y_n b)$ und weiter durch Ausmultiplizieren
$\exists x, z \in R(1 = x{\cdot}a_1{\cdot} \ldots {\cdot}a_n + z{\cdot}b)$, also nach (a) $a_1 \cdot \ldots \cdot a_n$, b teilerfremd.

Polynome

Sei K ein Körper.

Definition.
$K^{(\mathbb{N})} := \{(a_i)_{i\in\mathbb{N}} \in K^{\mathbb{N}} : \exists n \forall i > n(a_i = 0)\}$.

Die Elemente von $K^{(\mathbb{N})}$ nennt man *Polynome über K*.

Auf $K^{(\mathbb{N})}$ definiert man *Addition* + und *Multiplikation* $\cdot$ durch

$(a_0, a_1, \ldots) + (b_0, b_1, \ldots) := (a_0 + b_0, a_1 + b_1, \ldots)$,

$(a_0, a_1, \ldots) \cdot (b_0, b_1, \ldots) := (c_0, c_1, \ldots)$, wobei $c_k := \sum\limits_{i=0}^{k} a_i b_{k-i} = \sum\limits_{i+j=k} a_i b_j$.

Satz 10.4.
$K^{(\mathbb{N})}$ mit obiger Addition und Multiplikation ist ein kommutativer Ring
mit Nullelement $(0, 0, \ldots)$ und Einselement $(1, 0, 0, \ldots)$.

Diesen Ring bezeichnet man mit $K[t]$ und nennt ihn den *Polynomring über K in einer Veränderlichen*.

Beweis:
Wir weisen nur die Eigenschaften der Multiplikation nach.
Seien $f = (a_i)$, $g = (b_i)$ und $h = (c_i)$ Elementen von $K^{(\mathbb{N})}$.
1. $f \cdot g = (\sum_{i=0}^{k} a_i b_{k-i})_{k\in\mathbb{N}} = (\sum_{i=0}^{k} b_{k-i} a_i)_{k\in\mathbb{N}} = (\sum_{j=0}^{k} b_j a_{k-j})_{k\in\mathbb{N}} = g \cdot f$.
2. $(f \cdot g) \cdot h = (\sum_{i=0}^{k} \big(\sum_{j=0}^{i} (a_j b_{i-j})\big) c_{k-i})_{k\in\mathbb{N}} = (\sum_{i=0}^{k} \sum_{j=0}^{i} a_j b_{i-j} c_{k-i})_{k\in\mathbb{N}} = (\sum\limits_{j_1+j_2+j_3=k} a_{j_1} b_{j_2} c_{j_3})_{k\in\mathbb{N}} =$
$= (\sum_{i=0}^{k} \sum_{j=0}^{k-i} a_i b_j c_{k-i-j})_{k\in\mathbb{N}} = (\sum_{i=0}^{k} a_i \sum_{j=0}^{k-i} b_j c_{k-i-j})_{k\in\mathbb{N}} = f \cdot (g \cdot h)$
3. $f \cdot (g + h) = (\sum_{i=0}^{k} a_i (b_{k-i} + c_{k-i}))_{k\in\mathbb{N}} = (\sum_{i=0}^{k} a_i b_{k-i} + \sum_{i=0}^{k} a_i c_{k-i})_{k\in\mathbb{N}} = f \cdot g + f \cdot h$.
4. $(1, 0, 0, \ldots) \cdot f = (\delta_{0i})_i \cdot (a_i)_i = (\sum_{i=0}^{k} \delta_{0i} a_{k-i})_k = (a_k)_k = f$.

Bemerkung.

Für $a, b \in K$ gilt: $(a,0,0,...) +_{K[t]} (b,0,0,...) = (a+b,0,0,...)$ und $(a,0,0,...) \cdot_{K[t]} (b,0,0,...) = (ab,0,0,...)$.

Wir können deshalb im folgenden $a \in K$ mit $(a,0,0,...) \in K[t]$ identifizieren.

Dadurch wird K zu einem Unterring von $K[t]$.

Definition. $t := (0,1,0,0,\ldots) \in K^{(\mathbb{N})}$

Lemma 10.5.

(a) Für $f = (a_0, a_1, \ldots) \in K[t]$ mit $\forall i > n(a_i = 0)$ gilt $f = \sum_{i=0}^n a_i t^i$.

(Addition und Multiplikation sind hier in $K[t]$ zu verstehen.)

(b) In $K[t]$ gilt: $\sum_{i=0}^n a_i t^i = \sum_{j=0}^m b_j t^j \ \&\ n \le m \implies \forall k \le n(a_k = b_k) \ \&\ \forall k > n(0 = b_k)$.

Beweis:

(a) 1. Durch Induktion nach i erhalten wir $t^i = (\delta_{ij})_{j \in \mathbb{N}}$:

$t^0 = 1 = (1,0,0,\ldots) = (\delta_{0j})_j$, $t^{i+1} = t^i \cdot t \overset{\text{IH}}{=} (\delta_{ij})_j \cdot (\delta_{1j})_j = (\sum_{j=0}^k \delta_{ij}\delta_{1,k-j})_k = (\delta_{1,k-i})_k = (\delta_{i+1,k})_k$.

Außerdem ist $a \cdot (b_k)_k = (a\delta_{0i})_i \cdot (b_k)_k = (\sum_{i=0}^k a\delta_{0i} b_{k-i})_k = (ab_k)_k$.

Für $f = (a_i)_i$ mit $\forall i > n(a_i = 0)$ gilt deshalb $f = \sum_{i=0}^n (a_i \delta_{ij})_j = \sum_{i=0}^n a_i t^i$.

(b) Sei $\sum_{i=0}^n a_i t^i = \sum_{i=0}^m b_i t^i \ \&\ n \le m$. Wir setzen $a_i := 0$ für $i > n$, und $b_i := 0$ für $i > m$.

Dann $(a_i)_i \overset{\text{(a)}}{=} \sum_{i=0}^n a_i t^i = \sum_{i=0}^m b_i t^i \overset{\text{(a)}}{=} (b_i)_i$ und folglich $a_i = b_i$ für alle $i \in \mathbb{N}$.

Definition (Der Grad eines Polynoms).

Für $f = (a_i)_{i \in \mathbb{N}} \in K[t]$ sei $\deg(f) := \begin{cases} \max\{n \in \mathbb{N} : a_n \neq 0\} & \text{falls } f \neq 0 \\ -\infty & \text{falls } f = 0 \end{cases}$ *(Grad von f)*

Ist $f \neq 0$ und $n = \deg(f)$, so heißt a_n der *Leitkoeffizient* von f, und f heißt normiert, falls $a_n = 1$ ist.

Für $\xi \in \mathbb{N} \cup \{-\infty\}$ sei $\xi + (-\infty) := (-\infty) + \xi := -\infty$.

Lemma 10.6. $f, g \in K[t] \Rightarrow \deg(f+g) \le \max\{\deg(f), \deg(g)\}$ und $\deg(fg) = \deg(f) + \deg(g)$.

Beweis:

Für $f = 0 \vee g = 0$ ist die Behauptung trivial.

Sei jetzt $f = (a_i)_i$, $g = (b_i)_i$ mit $\deg(f) = m \in \mathbb{N}$ und $\deg(g) = n \in \mathbb{N}$.

Dann $\forall k > \max\{m,n\}(a_k + b_k = 0)$ und $c_k = \sum_{i=0}^k a_i b_{k-i} = \begin{cases} a_m b_n \neq 0 & \text{für } k = m+n \\ 0 & \text{für } k > m+n \end{cases}$.

Folgerung.

$K[t]$ ist nullteilerfrei und somit ein Integritätsring.

Satz 10.7.

Zu $f, g \in K[t]$ mit $g \neq 0$ existieren eindeutig bestimmte Polynome $q, r \in K[t]$ derart, daß $f = q \cdot g + r$ und $\deg(r) < \deg(g)$.

Beweis:

Eindeutigkeit: Gelte $q \cdot g + r = q' \cdot g + r'$ und $\deg(r), \deg(r') < \deg(g)$.

Dann $r' - r = (q - q') \cdot g$ und $\deg(r - r') < \deg(g)$. Aus $q \neq q'$ würde deshalb $\deg(r - r') < \deg(g) \overset{!!!}{\le} \deg(q - q') + \deg(g) \overset{\text{L.10.6}}{=} \deg((q - q') \cdot g) = \deg(r - r')$ folgen. Also ist $q = q'$ und damit auch $r = r'$.

Existenz: Durch Rekursion nach $\deg(f)$ definieren wir Polynome $[\frac{f}{g}]$ und r, so daß $f = [\frac{f}{g}]\cdot g + r$ und $\deg(r) < m := \deg(g)$.

1. $\deg(f) < m$: $r := f$, $[\frac{f}{g}] := 0$.

2. $m \le \deg(f)$: $f = \sum_{i=0}^{n} a_i t^i$, $g = \sum_{i=0}^{m} b_i t^i$ mit $m \le n$ und $a_n, b_m \neq 0$. Sei $f' := f - g\cdot\frac{a_n}{b_m}\cdot t^{n-m}$.

Dann $\deg(f') < n = \deg(f)$, und nach I.V. existieren $[\frac{f'}{g}]$ und r mit $f' = [\frac{f'}{g}]\cdot g + r$ und $\deg(r) < m$.

Es folgt $f = f' + g\cdot\frac{a_n}{b_m}\cdot t^{n-m} = (\frac{a_n}{b_m}t^{n-m} + [\frac{f'}{g}])\cdot g + r$. Also $[\frac{f}{g}] = \frac{a_n}{b_m}t^{n-m} + [\frac{f'}{g}]$.

Folgerung. $K[t]$ ist ein euklidischer Ring.

Beispiel zur Polynomdivision

$$
\begin{array}{l}
3t^5 - 2t^4 + t^3 - 5t^2 + 1 : t^3 + t^2 - 2t = 3t^2 - 5t + 12 \text{ Rest } -27t^2 + 24t + 1 \\
3t^5 + 3t^4 - 6t^3 \\
\quad - 5t^4 + 7t^3 \\
\quad - 5t^4 - 5t^3 + 10t^2 \\
\qquad\quad 12t^3 - 15t^2 \\
\qquad\quad 12t^3 + 12t^2 - 24t \\
\qquad\qquad - 27t^2 + 24t + 1
\end{array}
$$

Allgemeines Schema:

$$
\begin{array}{l}
a_n t^n + \quad a_{n-1}t^{n-1} + \ldots + a_{n-m}t^{n-m} + \ldots + a_1 t + a_0 \;:\; b_m t^m + \ldots + b_0 \;=\; \frac{a_n}{b_m}t^{n-m} + \ldots.. \\
\underline{a_n t^n + b_{m-1}\frac{a_n}{b_m}t^{n-1} + \ldots + b_0\frac{a_n}{b_m}t^{n-m} \qquad\qquad} \\
\qquad c_{n-1}t^{n-1} + \ldots + c_{n-m}t^{n-m} + \ldots + a_1 t + a_0 \qquad \text{wobei } c_{n-i} := a_{n-i} - b_{m-i}\frac{a_n}{b_m}
\end{array}
$$

Definitionen.

Sei $f = \sum_{i=0}^{n} a_i t^i \in K[t]$ und $\lambda \in K$.

$f(\lambda) := \sum_{i=0}^{n} a_i \lambda^i \in K$ nennt man das Ergebnis der *Einsetzung* von λ in f.

λ heißt *Nullstelle von* f, wenn $f(\lambda) = 0$ ist.

f heißt *konstant*, wenn $\deg(f) \le 0$ ist.

Bemerkung.

$f \in K[t]$ ist genau dann konstant, wenn $f = a_0 \in K$. In diesem Fall gilt $f(\lambda) = a_0$ für alle $\lambda \in K$.

Lemma 10.8.

Für $f, g \in K[t]$ und $\lambda \in K$ gilt: $(f+g)(\lambda) = f(\lambda) + g(\lambda)$ und $(f\cdot g)(\lambda) = f(\lambda)\cdot g(\lambda)$.

Beweis des zweiten Teils:

Sei $f = \sum_{i=0}^{m} a_i t^i$ und $g = \sum_{i=0}^{n} b_i t^i$. Ferner $c_k := \sum_{i=0}^{k} a_i b_{k-i}$. Dann gilt: $f(\lambda)\cdot g(\lambda) =$
$= (\sum_{i=0}^{m} a_i\lambda^i)\cdot(\sum_{i=0}^{n} b_i\lambda^i) = \sum_{i=0}^{m}\sum_{j=0}^{n} a_i b_j \lambda^{i+j} = \sum_{k=0}^{m+n}(\sum_{i+j=k} a_i b_j)\lambda^k = \sum_{k=0}^{m+n} c_k\lambda^k = (f\cdot g)(\lambda)$.

Lemma 10.9.

Ist $\lambda \in K$ eine Nullstelle von $f \in K[t]$, so existiert ein $q \in K[t]$ mit $f = q\cdot(t-\lambda)$.

Beweis :

Nach 10.7 existieren $q, r \in K[t]$ mit $f = q\cdot(t-\lambda) + r$ und $\deg(r) < \deg(t-\lambda) = 1$.

Dann $0 = f(\lambda) = r(\lambda)$ und $\deg(r) \le 0$, also $r = 0$.

Definition.
Für $f \in K[t] \setminus \{0\}$ und $\lambda \in K$ sei
$\mu(f;\lambda) := \max\{k \in \mathbb{N} : \exists g \in K[t](f = (t-\lambda)^k \cdot g)\}$.
Ist λ eine Nullstelle von f, so heißt $\mu(f;\lambda)$ die *Vielfachheit der Nullstelle λ von f*.

Bemerkung. Nach 10.6 und 10.9 gilt: $f(\lambda) = 0 \;\Rightarrow\; 1 \le \mu(f;\lambda) \le \deg(f)$.

Lemma 10.10.
Für $f \in K[t] \setminus \{0\}$ und $k \in \mathbb{N}$ gilt: $k = \mu(f;\lambda) \;\Leftrightarrow\; \exists g(f = (t-\lambda)^k \cdot g \;\&\; g(\lambda) \neq 0)$.

Beweis :
"$\Rightarrow$" folgt aus der Definition von $\mu(f;\lambda)$ mittels Lemma 10.9.
"$\Leftarrow$": Sei $f = (t-\lambda)^k \cdot g$ mit $g(\lambda) \neq 0$. Dann ist $k \le \mu := \mu(f;\lambda)$ und $f = (t-\lambda)^\mu \cdot q$ für ein $q \in K[t]$.
Aus $(t-\lambda)^k \cdot g = f = (t-\lambda)^\mu \cdot q$ folgt $g = (t-\lambda)^{\mu-k} \cdot q$ und weiter $k = \mu$ wegen $g(\lambda) \neq 0$.

Satz 10.11.
Sind $\lambda_1, ..., \lambda_k$ paarweise verschiedene Nullstellen von $f \in K[t] \setminus \{0\}$ so existiert ein $g \in K[t]$ mit
(i) $f = (t-\lambda_1)^{\mu_1} \cdot \ldots \cdot (t-\lambda_k)^{\mu_k} \cdot g$, wobei $\mu_i := \mu(f;\lambda_i)$ für $i = 1, ..., k$, und
(ii) $g(\lambda_i) \neq 0$ für $i = 1, \ldots, k$.

Ist g konstant, so sagt man *"f zerfällt in Linearfaktoren"*.

Beweis durch Induktion nach k:
Nach 10.10 existiert $h \in K[t] \setminus \{0\}$ mit $f = (t-\lambda_1)^{\mu_1} \cdot h$ und $h(\lambda_1) \neq 0$.
Offenbar sind dann $\lambda_2, ..., \lambda_k$ Nullstellen von h. — Fall 1: $k = 1$. Dann ist nichts weiter zu beweisen.
Fall 2: $k > 1$. Nach I.V. existiert ein g, so daß $h = (t-\lambda_2)^{\nu_2} \cdot \ldots \cdot (t-\lambda_k)^{\nu_k} \cdot g$, wobei $\nu_i := \mu(h;\lambda_i)$ und $g(\lambda_i) \neq 0$ für $i = 2, \ldots, k$. Wegen $h(\lambda_1) \neq 0$ ist auch $g(\lambda_1) \neq 0$. Es folgt $f = (t-\lambda_1)^{\mu_1} \cdot (t-\lambda_2)^{\nu_2} \cdot \ldots \cdot (t-\lambda_k)^{\nu_k} \cdot g$ und daraus mit 10.10 $\nu_i = \mu(f;\lambda_i)$.

Folgerung.
Ist $f \in K[t]$ und $\deg(f) = n \ge 0$, so hat f höchstens n verschiedene Nullstellen.

Lemma 10.12.
Ist K unendlich, so gilt: $f, g \in K[t] \;\&\; \forall \lambda \in K(f(\lambda) = g(\lambda) \Rightarrow f = g)$.

Beweis:
Nach Voraussetzung hat $f - g$ unendlich viele Nullstellen; nach obiger Folgerung ist also $f - g = 0$.

Bemerkung
Ist $K = \{a_0, ..., a_n\}$, so ist $f := (t-a_0) \cdot ... \cdot (t-a_n) \in K[t]$ vom Nullpolynom verschieden,
aber trotzdem $f(\lambda) = 0$ für alle $\lambda \in K$.

Lemma 10.13.
Für $Q = (f_{ij})_{i,j} \in K[t]^{n\times n}$ und $\lambda \in K$ gilt $\det(Q)(\lambda) = \det\Big(f_{ij}(\lambda)\Big)_{i,j}$.

Beweis:
$\det(Q) = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) f_{1\pi(1)} \cdots f_{n\pi(n)} \overset{10.8}{\Rightarrow}$
$\Rightarrow \det(Q)(\lambda) = \sum_{\pi \in \mathcal{S}_n} \operatorname{sign}(\pi) f_{1\pi(1)}(\lambda) \cdots f_{n\pi(n)}(\lambda) = \det\left(f_{ij}(\lambda)\right)_{i,j}$.

Fundamentalsatz der Algebra. Jedes Polynom $P \in \mathbb{C}[t] \setminus \{0\}$ zerfällt in Linearfaktoren.

§11 Eigenwerte

Definition

Sei V ein K-Vektorraum, $f \in \mathsf{End}(V)$ und $\lambda \in K$.

1. λ heißt *Eigenwert* von f, wenn es ein $v \neq 0$ aus V mit $f(v) = \lambda v$ gibt.
2. Ist λ Eigenwert von f, so heißt jedes $v \in V \setminus \{0\}$ mit $f(v) = \lambda v$ ein *Eigenvektor von f zum Eigenwert* λ.
3. $\mathrm{Eig}(f;\lambda) := \{v \in V : f(v) = \lambda v\}$ heißt *Eigenraum* von f bzgl. λ.

Bemerkungen.

(i) $\mathrm{Eig}(f;\lambda) \setminus \{0\}$ ist die Menge der zu λ gehörigen Eigenvektoren von f.

(ii) λ ist Eigenwert von $f \iff \mathrm{Eig}(f;\lambda) \neq \{0\}$.

(iii) $\mathrm{Eig}(f;\lambda) = \mathrm{Ker}(f - \lambda{\cdot}\mathrm{id}_V)$, insbesondere $\mathrm{Eig}(\mathsf{f}_A;\lambda) = \mathrm{Lös}(A - \lambda{\cdot}E_n; 0)$ für $A \in K^{n\times n}$.

Beweis von (iii): $v \in \mathrm{Eig}(f;\lambda) \Leftrightarrow f(v) = \lambda v \Leftrightarrow f(v) - \lambda v = 0 \Leftrightarrow (f - \lambda{\cdot}\mathrm{id}_V)(v) = 0$.

Satz 11.1.

Ist $\dim(V) < \infty$, so gilt für $f \in \mathsf{End}(V)$ und $\lambda \in K$: $\quad \lambda$ Eigenwert von $f \iff \det(f - \lambda\,\mathrm{id}_V) = 0$.

Beweis :

λ ist EW $\overset{\text{(ii)+(iii)}}{\Leftrightarrow}$ $\mathrm{Ker}(f - \lambda\,\mathrm{id}) \neq \{0\} \Leftrightarrow f - \lambda\,\mathrm{id}$ ist nicht injektiv $\overset{9.16a}{\Leftrightarrow} \det(f - \lambda\,\mathrm{id}) = 0$.

Beispiele.

1. $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ mit $\alpha \notin \{n\cdot\pi : n \in \mathbb{Z}\}$: Nach §5, Beisp. 2 ist f_A eine Drehung um den Winkel α. Deshalb ist anschaulich klar, daß f_A keinen Eigenwert hat.

2. $B = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$: Sei $v_1 := (\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2})^{\mathsf{t}}$ und $v_2 := (-\sin\frac{\alpha}{2}, \cos\frac{\alpha}{2})^{\mathsf{t}}$. f_B ist die Spiegelung an der Geraden $\mathbb{R}{\cdot}v_1$. Deshalb hat f_B die Eigenvektoren v_1 zum EW $\lambda_1 = 1$ und v_2 zum EW $\lambda_2 = -1$.

$\overline{v} = (v_1, v_2)$ ist eine Basis von $\mathbb{R}^2$ mit $\mathcal{M}_{\overline{v}}(\mathsf{f}_B) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

3. Sei $I \subseteq \mathbb{R}$ ein Intervall und $V = \mathcal{D}(I;\mathbb{R})$ der unendlichdimensionale $\mathbb{R}$-Vektorraum der auf I beliebig oft differenzierbaren Funktionen. Der Endomorphismus $f : V \to V$, $\varphi \mapsto \varphi'$ hat jedes $\lambda \in \mathbb{R}$ als Eigenwert, denn die Funktion $\varphi_\lambda : I \to \mathbb{R}$, $x \mapsto e^{\lambda x}$ ist Eigenvektor zu λ. $[\frac{d}{dx} e^{\lambda x} = \lambda e^{\lambda x}]$

Satz 11.2.

Sind $\lambda_1, ..., , \lambda_k$ paarweise verschiedene Eigenwerte von $f \in \mathsf{End}(V)$, so gilt:

(a) Ist v_i Eigenvektor von λ_i $(i = 1, ..., k)$, so ist das Tupel $(v_1, ..., v_k)$ linear unabhängig.

(b) Die Summe $\mathrm{Eig}(f;\lambda_1) + \ldots + \mathrm{Eig}(f;\lambda_k)$ ist direkt.

Beweis :

(a) Induktion nach k:

1. $k = 1$: v_1 linear unabhängig, denn v_1 ist Eigenvektor und damit ungleich 0.
2. $k > 1$: Nach IV sind $v_1, ..., v_{k-1}$ linear unabhängig. Sei nun $\sum_{i=1}^k \alpha_i v_i = 0$. Durch Anwenden von f bzw. durch Multiplikation mit λ_k erhalten wir daraus $\sum_{i=1}^k \alpha_i \lambda_i v_i = 0$ und $\sum_{i=1}^k \alpha_i \lambda_k v_i = 0$. Durch Subtraktion folgt $\sum_{i=1}^{k-1} \alpha_i(\lambda_i - \lambda_k)v_i = 0$ und daraus mit IV $\alpha_i(\lambda_i - \lambda_k) = 0$ für $i = 1, ..., k{-}1$.

Wegen $\lambda_i \neq \lambda_k$ für $i = 1, ..., k{-}1$ folgt nun $\alpha_1 = \ldots = \alpha_{k-1} = 0$, also auch $\alpha_k v_k = 0$ und damit $\alpha_k = 0$.

(b) Sei $u_1 + ... + u_k = 0$ mit $u_i \in Eig(f;\lambda_i)$ für $i = 1, ..., k$. Z.z.: $u_1 = ... = u_k = 0$.
Annahme: $\exists i \in \{1, ..., k\}(u_i \neq 0)$. Dann o.E.d.A. $u_1, ..., u_r \neq 0$ und $u_{r+1} = ... = u_k = 0$ mit $1 \leq r \leq k$. Also sind $u_1, ..., u_r$ Eigenvektoren zu paarweise verschiedenen Eigenwerten und somit nach (a) $(u_1, ..., u_r)$ linear unabhängig, woraus $u_1 + ... + u_r \neq 0$ und weiter $u_1 + ... + u_k \neq 0$ folgt. *Widerspruch.*

Definition. Sei $A \in K^{n\times n}$.
λ [bzw. v] heißt Eigenwert [bzw. Eigenvektor] von A :⇔ λ [bzw. v] ist Eigenwert [bzw. Eigenvektor] von f_A.
$\mathrm{Eig}(A;\lambda) := \mathrm{Eig}(\mathsf{f}_A;\lambda)$ heißt Eigenraum von A bzgl. λ.

Das charakteristische Polynom

Definition.
Für $A = (a_{ij}) \in K^{n\times n}$ sei $\quad P_A := \det(A - t{\cdot}E) \in K[t]$.

P_A heißt *charakteristisches Polynom* der Matrix A.

Lemma 11.3.
(a) Für $A \in K^{n\times n}$ und $\lambda \in K$ gilt: $\;P_A(\lambda) = \det(A - \lambda{\cdot}E_n)$.
(b) Sind $A, B \in K^{n\times n}$ darstellende Matrizen von $f \in \mathsf{End}(V)$, so ist $P_A = P_B$.

Beweis:
(a) $P_A(\lambda) = \det(A - t{\cdot}E)(\lambda) \overset{10.13}{=} \det\Big((a_{ij} - t{\cdot}\delta_{ij})(\lambda)\Big)_{ij} = \det\Big(a_{ij} - \lambda{\cdot}\delta_{ij}\Big)_{ij} = \det(A - \lambda{\cdot}E_n)$.
(b) Nach dem Korollar zu Satz 6.6 existiert $S \in \mathrm{GL}(n;K)$ mit $B = S^{-1}AS$. Dann $P_B = \det(B - t{\cdot}E) = \det(S^{-1}AS - S^{-1}(tE)S) = \det(S^{-1}(A - tE)S) \overset{9.9}{=} \det(A - tE) = P_A$.

Lemma 11.4.
Sei $A = (a_{ij})_{i,j} \in K^{n\times n}$ und $P_A = \sum_i \alpha_i t^i$. Dann gilt:
(a) $\deg(P_A) = n$.
(b) $\alpha_n = (-1)^n$.
(c) $\alpha_{n-1} = (-1)^{n-1}(a_{11} + \ldots + a_{nn})$.
(d) $\alpha_0 = \det(A)$.
(e) P_A zerfällt $\Longrightarrow \exists \lambda_1, ..., \lambda_n \in K(\,P_A = (\lambda_1 - t)\cdot \ldots \cdot (\lambda_n - t)\,)$.

Beweis:
(a),(b),(c) Sei $\mathcal{S}_n' := \mathcal{S}_n \setminus \{\mathrm{id}\}$. Dann $P_A = \det(A - t{\cdot}E) = (a_{11} - t)\cdot \ldots \cdot (a_{nn} - t) + P'$ mit
$P' := \sum_{\pi\in\mathcal{S}_n'} \mathrm{sign}(\pi){\cdot}q_{1\pi(1)} \cdot \ldots \cdot q_{n\pi(n)}$ und $q_{ij} := a_{ii} - \delta_{ij}t$.

$\deg(P') \leq \max\{\sum_{i=1}^n \deg(q_{i\pi(i)}) : \pi \in \mathcal{S}_n'\} \leq n - 2$ (O.E.d.A. $n \geq 2$).
$\deg((a_{11} - t)\cdot \ldots \cdot (a_{nn} - t)) = \deg(a_{11} - t) + \ldots + \deg(a_{nn} - t) = n$.
Sei $(a_{11} - t)\cdot \ldots \cdot (a_{nn} - t) = \sum_{i=0}^n \beta_i t^i$. Dann $\alpha_n = \beta_n = (-1)^n$, $\alpha_{n-1} = \beta_{n-1} = (-1)^{n-1}(a_{11} + \ldots + a_{nn})$.
(d) $\alpha_0 = P_A(0) = \det(A - 0{\cdot}E) = \det(A)$.
(e) Nach Voraussetzung ist $P_A = (\widetilde{\lambda}_1 - t)^{\mu_1}\cdot \ldots \cdot(\widetilde{\lambda}_k - t)^{\mu_k}\cdot a$ mit $a \in K\setminus\{0\}$ und $\mu_1 + ... + \mu_k = \deg(P_A) \overset{(a)}{=} n$. Daraus folgt $P_A = a{\cdot}\prod_{i=1}^n(\lambda_i - t)$ für passende $\lambda_1, ..., \lambda_n \in K$, sowie $(-1)^n \overset{(b)}{=} \alpha_n = (-1)^n a$, also $a = 1$.

Definition.
Für $f \in \mathsf{End}(V)$ $(\dim(V) < \infty)$ sei $P_f := P_A$, wobei A eine darstellende Matrix von f (siehe L.11.3b).
P_f heißt *charakteristisches Polynom* des Endomorphismus f.

Bemerkung. Für alle $\lambda \in K$ gilt:
(a) $P_f(\lambda) = \det(f - \lambda \cdot \mathrm{id}_V)$.
(b) λ ist Nullstelle von $P_f \iff \lambda$ ist Eigenwert von f.

Beweis:
(a) $P_f(\lambda) = P_A(\lambda) \overset{11.3a}{=} \det(A - \lambda E) \overset{(*)}{=} \det(f - \lambda\,\mathrm{id}_V)$. [$(*)$ $A - \lambda E$ ist darstellende Matrix von $f - \lambda\,\mathrm{id}_V$.]
(b) folgt aus (a) und 11.1.

Beispiele.
1. $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. $P_A = (\cos\alpha - t)^2 + \sin^2\alpha = t^2 - 2t\cos\alpha + 1$.
Nullstellen von P_A: $\lambda_{1/2} = \frac{1}{2}(2\cos\alpha \pm \sqrt{4\cos^2\alpha - 4}) = \cos\alpha \pm \sqrt{\cos^2\alpha - 1}$.
Im Fall $|\cos\alpha| < 1$ (d.h. wenn $\alpha \notin \{n \cdot \pi : n \in \mathbb{Z}\}$) besitzt A also keinen (reellen) Eigenwert.

2. $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$.
$P_A = (\cos\alpha - t)(-\cos\alpha - t) - \sin^2\alpha = -(\cos^2\alpha - t^2) - \sin^2\alpha = t^2 - 1 = (t-1)(t+1)$.
A hat somit die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$. Bestimmung der Eigenräume: Sei $\beta := \frac{\alpha}{2}$.

$$\begin{pmatrix} \cos\alpha - 1 & \sin\alpha \\ \sin\alpha & -\cos\alpha - 1 \end{pmatrix} = \begin{pmatrix} \cos^2\beta - \sin^2\beta - 1 & 2\sin\beta\cos\beta \\ 2\sin\beta\cos\beta & -\cos^2\beta + \sin^2\beta - 1 \end{pmatrix} = \begin{pmatrix} -2\sin^2\beta & 2\sin\beta\cos\beta \\ 2\sin\beta\cos\beta & -2\cos^2\beta \end{pmatrix}$$

$$\begin{pmatrix} \cos\alpha + 1 & \sin\alpha \\ \sin\alpha & -\cos\alpha + 1 \end{pmatrix} = \begin{pmatrix} \cos^2\beta - \sin^2\beta + 1 & 2\sin\beta\cos\beta \\ 2\sin\beta\cos\beta & -\cos^2\beta + \sin^2\beta + 1 \end{pmatrix} = \begin{pmatrix} 2\cos^2\beta & 2\sin\beta\cos\beta \\ 2\sin\beta\cos\beta & 2\sin^2\beta \end{pmatrix}$$

$\Rightarrow$ $\mathrm{Eig}(A;1) = \mathbb{R}\begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}$ und $\mathrm{Eig}(A;-1) = \mathbb{R}\begin{pmatrix} -\sin\beta \\ \cos\beta \end{pmatrix}$.

3. $A = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}$. $P_A = (-1-t)(4-t) + 6 = t^2 - 3t + 2 = (t-1)(t-2)$.
Eigenwerte: $\lambda_1 = 1$ und $\lambda_2 = 2$. Bestimmung der Eigenräume:
$\begin{pmatrix} -2 & 6 \\ -1 & 3 \end{pmatrix} \overset{\text{EZU}}{\longrightarrow} \begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix} \Rightarrow \mathrm{Eig}(A;1) = \mathbb{R}\begin{pmatrix} 3 \\ 1 \end{pmatrix}$. $\begin{pmatrix} -3 & 6 \\ -1 & 2 \end{pmatrix} \overset{\text{EZU}}{\longrightarrow} \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} \Rightarrow \mathrm{Eig}(A;2) = \mathbb{R}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

4. $A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$: $P_A = -t \cdot \det\begin{pmatrix} -2-t & 3 \\ -2 & 3-t \end{pmatrix} + 3 \cdot \det\begin{pmatrix} -1 & 1 \\ -2 & 3-t \end{pmatrix} - 2 \cdot \det\begin{pmatrix} -1 & 1 \\ -2-t & 3 \end{pmatrix} =$
$= -t(t^2 - t) + 3(t-1) - 2(t-1) = -(t-1)(t^2-1) = -(t-1)^2(t+1)$. Eigenwerte: $\lambda_1 = 1$ und $\lambda_2 = -1$.

Eigenräume:
$\lambda_1 = 1$: $\begin{pmatrix} -1 & -1 & 1 \\ -3 & -3 & 3 \\ -2 & -2 & 2 \end{pmatrix} \overset{\text{EZU}}{\longrightarrow} \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $\mathrm{Eig}(A;1) = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 + x_2 - x_3 = 0\} = \mathrm{span}(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix})$.

$\lambda_2 = -1$: $\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ -2 & -2 & 4 \end{pmatrix} \overset{\text{EZU}}{\longrightarrow} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -4 & 6 \\ 0 & -4 & 6 \end{pmatrix} \overset{\text{EZU}}{\longrightarrow} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -3 \\ 0 & 0 & 0 \end{pmatrix}$,
$\mathrm{Eig}(A;-1) = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 - x_2 + x_3 = 0 \;\&\; 2x_2 - 3x_3 = 0\} = \mathbb{R}\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$.

Abkürzung. $Diag(\lambda_1, ..., \lambda_n) := (\lambda_i \delta_{ij})_{i,j} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

Lemma 11.5.

(a) Seien $f \in \mathsf{End}(V)$, $(v_1, ..., v_n)$ eine Basis von V und $\lambda_1, ..., \lambda_n \in K$. Dann gilt:
$\mathcal{M}_{\overline{v}}(f) = Diag(\lambda_1, ..., \lambda_n) \iff \forall j \in \{1, .., n\}(v_j$ ist Eigenvektor von f zum Eigenwert λ_j).

(b) Seien $A, S \in K^{n \times n}$, $v_1, ..., v_n \neq 0$ die Spalten von S und $\lambda_1, ..., \lambda_n \in K$. Dann gilt:
$A \cdot S = S \cdot Diag(\lambda_1, ..., \lambda_n) \iff \forall j \in \{1, .., n\}(v_j$ ist Eigenvektor von A zum Eigenwert λ_j).

Beweis:

(a) Nach Definition von $\mathcal{M}_{\overline{v}}(f)$ gilt: $\mathcal{M}_{\overline{v}}(f) = Diag(\lambda_1, ..., \lambda_n) \Leftrightarrow \forall j \in \{1, .., n\}(f(v_j) = \lambda_j v_j)$.

(b) Die Behauptung folgt aus $A \cdot S = (Av_1 \ \ldots \ Av_n)$ und $S \cdot Diag(\lambda_1, ..., \lambda_n) = (\lambda_1 v_1 \ \cdots \ \lambda_n v_n)$.

Definition.

Ein Endomorphismus $f : V \to V$ heißt *diagonalisierbar*, wenn V eine Basis aus Eigenvektoren von f besitzt. Eine Matrix $A \in K^{n \times n}$ heißt *diagonalisierbar*, falls $\mathsf{f}_A : K^n \to K^n$ diagonalisierbar ist.

Lemma 11.6.

(a) $A \in K^{n \times n}$ ist genau dann diagonalisierbar, wenn es eine invertierbare Matrix $S \in K^{n \times n}$ gibt, so daß $S^{-1}AS$ eine Diagonalmatrix ist.

(b) Hat $f \in \mathsf{End}(V)$ die paarweise verschiedenen Eigenwerte $\lambda_1, ..., \lambda_n$ $(n = \dim V)$, so ist f diagonalisierbar.

Beweis :

(a) "$\Rightarrow$": Sei $(v_1, ..., v_n)$ eine Basis aus Eigenvektoren von f_A und seien $\lambda_1, ..., \lambda_n$ die zugehörigen Eigenwerte. Für $S := (v_1 \ \ldots \ v_n)$ gilt dann: $A \cdot S = S \cdot Diag(\lambda_1, ..., \lambda_n)$ (nach 11.5b) und S invertierbar (nach 5.14).
"$\Leftarrow$": Sei $S = (v_1 \ \ldots \ v_n)$ invertierbar und $S^{-1}AS = Diag(\lambda_1, ..., \lambda_n)$. Dann ist $(v_1, ..., v_n)$ eine Basis von K^n (nach 5.14), und nach 11.5b gilt: v_j ist Eigenvektor zum Eigenwert λ_j $(j = 1, ..., n)$.

(b) Für $j = 1, ..., n$ sei jeweils v_j ein Eigenvektor von f zu λ_j. Dann ist $(v_1, ..., v_n)$ nach 11.2a eine Basis.

Satz 11.7.

Sind $\lambda_1, ..., \lambda_k$ die verschiedenen Eigenwerte von $f \in \mathsf{End}(V)$ und ist $\dim(V) < \infty$, so sind äquivalent:

(i) f diagonalisierbar

(ii) $V = \mathrm{Eig}(f; \lambda_1) \oplus ... \oplus \mathrm{Eig}(f; \lambda_k)$

(iii) $\dim(V) = \dim \mathrm{Eig}(f; \lambda_1) + ... + \dim \mathrm{Eig}(f; \lambda_k)$.

Beweis :

1. (i) $\Leftrightarrow$ V besitzt Erzeugendensystem aus Eigenvektoren $\Leftrightarrow$ $V = \mathrm{Eig}(f; \lambda_1) + \ldots + \mathrm{Eig}(f; \lambda_k) \overset{11.2}{\Leftrightarrow}$ (ii).

2. Sei $U_i := \mathrm{Eig}(f; \lambda_i)$ und $U := U_1 + \ldots + U_k$. Nach 11.2 und 4.12 (Korollar) gilt dann $U = U_1 \oplus \ldots \oplus U_k$ und $\dim(U) = \dim(U_1) + \ldots + \dim(U_k)$. Folglich: (ii) $\Leftrightarrow$ $V = U \overset{4.10b}{\Leftrightarrow} \dim(V) = \dim(U) \Leftrightarrow$ (iii).

Definition. Ist λ Eigenwert des Endomorphismus $f : V \to V$, so definiert man:
Geometrische Vielfachheit von $\lambda := \dim(Eig(f; \lambda))$,
algebraische Vielfachheit von $\lambda := \mu(P_f; \lambda)$ (Vielfachheit von λ als Nullstelle des charakteristischen Polynoms von f)

Satz 11.8.
Ist $\dim(V) = n < \infty$ und $f \in \mathsf{End}(V)$, so gilt:
(a) Für jeden Eigenwert λ von f ist $\dim \mathrm{Eig}(f;\lambda) \leq \mu(P_f;\lambda)$.

(b) f diagonalisierbar $\iff \begin{cases} \text{Das charakteristische Polynom } P_f \text{ zerfällt in Linearfaktoren und} \\ \dim \mathrm{Eig}(f;\lambda) = \mu(P_f;\lambda) \text{ für jeden Eigenwert } \lambda \text{ von } f. \end{cases}$

Beweis :
(a) Wir wählen eine Basis $\overline{v} := (v_1, ..., v_n)$ von V, so daß $(v_1, ..., v_k)$ Basis von $\mathrm{Eig}(f;\lambda)$ ist. Sei $A := \mathcal{M}_{\overline{v}}(f)$.

$$\text{Dann } A = \begin{pmatrix} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ & 0 & & A' \end{pmatrix}, \quad A - t{\cdot}E_n = \begin{pmatrix} \lambda - t & & 0 & \\ & \ddots & & * \\ 0 & & \lambda - t & \\ & 0 & & A' - t{\cdot}E_{n-k} \end{pmatrix} \text{ mit } A' \in K^{n-k \times n-k}.$$

Es folgt $P_f = \det(A - t{\cdot}E_n) = (\lambda - t)^k {\cdot} \det(A' - t{\cdot}E_{n-k}) = (\lambda - t)^k {\cdot} P_{A'}$ und somit $\dim \mathrm{Eig}(f;\lambda) = k \leq \mu(P_f;\lambda)$.

(b) Seien $\lambda_1, ..., \lambda_k$ die verschiedenen Eigenwerte von f (i.e. Nullstellen von P_f), und sei $\mu_i := \mu(P_f;\lambda_i)$.

Nach 10.11 ist $P_f = (t - \lambda_1)^{\mu_1} \cdot \ldots \cdot (t - \lambda_k)^{\mu_k} \cdot g$ und somit $n = \deg(P_f) = \mu_1 + ... + \mu_k + \deg(g)$.

Daraus folgt: (1) $\mu_1 + \ldots + \mu_k \leq n$,
(2) P_f zerfällt in Linearfaktoren $\iff n = \mu_1 + \ldots + \mu_k$.

Nach 11.7 gilt: (3) f diagonalisierbar $\iff n = d_1 + \ldots + d_k$, wobei $d_i := \dim \mathrm{Eig}(f;\lambda_i)$.

Aus (1)-(3) und (a) folgt die Behauptung: "$\Leftarrow$": P_f zerfällt und $d_i = \mu_i \overset{(2),(3)}{\Rightarrow} f$ diagonalisierbar.
"$\Rightarrow$": f diag. $\overset{(3),(a),(1)}{\Rightarrow} n = d_1 + ... + d_k \leq \mu_1 + \ldots + \mu_k \leq n \overset{(a)}{\Rightarrow} n = \mu_1 + \ldots + \mu_k$ und $d_i = \mu_i$.

Bemerkung.
Ist $A \in K^{n \times n}$ eine darstellende Matrix von $f \in \mathsf{End}(V)$, so gilt $\dim \mathrm{Eig}(f;\lambda) = \dim \mathrm{Eig}(A;\lambda)$.

Beweis: Sei $A = \mathcal{M}_{\overline{v}}(f)$. Wie man leicht sieht, gilt dann $\mathrm{Eig}(f;\lambda) = \Phi_{\overline{v}}(\mathrm{Eig}(A;\lambda))$.

Definition.
$A, B \in K^{n \times n}$ heißen *ähnlich*, wenn es eine invertierbare Matrix S mit $B = S^{-1}AS$ gibt.

Lemma 11.9.
(a) "ähnlich" ist eine Äquivalenzrelation auf $K^{n \times n}$.
(b) $A, B \in K^{n \times n}$ sind genau dann ähnlich, wenn es eine Basis $\overline{v}$ von K^n mit $B = \mathcal{M}_{\overline{v}}(\mathsf{f}_A)$ gibt.
(c) Zwei Matrizen sind genau dann ähnlich,
wenn sie darstellende Matrizen desselben Endomorphismus sind.
(d) Sind $A, B \in K^{n \times n}$ ähnlich, so ist $\mathrm{rang}(A) = \mathrm{rang}(B)$, $\det(A) = \det(B)$ und $P_A = P_B$.
(e) Eine Matrix $A \in K^{n \times n}$ ist genau dann diagonalisierbar, wenn sie ähnlich zu einer Diagonalmatrix ist.

Beweis:
(a) klar. (b) Satz 6.4. (c) Korollar zu Satz 6.6. (d) Lemma 7.2, Satz 9.14, Lemma 11.3b.

§12 Euklidische und unitäre Vektorräume

Für $z = x + \mathrm{i}y \in \mathbb{C}$ sei

$\overline{z} := x - \mathrm{i}y$ (die zu z konjugiert komplexe Zahl),

$\mathrm{Re}(z) := x$ (Realteil), $\mathrm{Im}(z) := y$ (Imaginärteil), $|z| := \sqrt{x^2 + y^2}$ (Absolutbetrag).

Für $z, z_1, z_2 \in \mathbb{C}$ gilt: $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$, $\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z})$, $\mathrm{Im}(z) = \frac{1}{2\mathrm{i}}(z - \overline{z})$, $|z|^2 = z\overline{z}$, und $\quad z \in \mathbb{R} \Leftrightarrow \mathrm{Re}(z) = z \Leftrightarrow \mathrm{Im}(z) = 0 \Leftrightarrow z = \overline{z}$.

Für $A = (a_{ij})_{i,j} \in \mathbb{C}^{m\times n}$ sei $\overline{A} := (\overline{a_{ij}})_{i,j} \in \mathbb{C}^{m\times n}$.

Dann gilt: $\overline{A+B} = \overline{A} + \overline{B}$, $\quad \overline{AB} = \overline{A} \cdot \overline{B}$, $\quad (\overline{A})^{\mathtt{t}} = \overline{A^{\mathtt{t}}}$, $\quad \det(\overline{A}) = \overline{\det(A)}$.

Im Folgenden werden wir ausschließlich Vektorräume über den Körpern $\mathbb{R}$ oder $\mathbb{C}$ betrachten und meist $\mathbb{K}$ anstelle von $\mathbb{R}$ oder $\mathbb{C}$ schreiben. Um Verwechslungen mit dem konjugiert Komplexen zu vermeiden, werden wir Basen von jetzt an mit $\tilde{u}, \tilde{v}, \tilde{w}$ (statt $\overline{u}, \overline{v}, \overline{w}$) bezeichnen.

Definition. Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\sigma : V \times V \to \mathbb{K}$ heißt *Sesquilinearform*, wenn für alle $v, v', w, w' \in V$ und $\lambda \in K$ gilt:

(S1) $\sigma(v + v', w) = \sigma(v, w) + \sigma(v', w)$ und $\sigma(\lambda v, w) = \lambda\sigma(v, w)$,

(S2) $\sigma(v, w + w') = \sigma(v, w) + \sigma(v, w')$ und $\sigma(v, \lambda w) = \overline{\lambda}\sigma(v, w)$.

Im Falle $\mathbb{K} = \mathbb{R}$ gilt stets $\lambda = \overline{\lambda}$, und man spricht dann auch von einer *Bilinearform* σ.
Man bezeichnet σ als *nicht-ausgeartet*, wenn gilt:
(i) $\forall v \in V(\forall w \in V(\sigma(v, w) = 0) \Rightarrow v = 0)$ und (ii) $\forall w \in V(\forall v \in V(\sigma(v, w) = 0) \Rightarrow w = 0)$.

Definition. Eine Sesquilinearform $\sigma : V \times V \to \mathbb{K}$, für die gilt $\forall x, y \in V(\sigma(x, y) = \overline{\sigma(y, x)})$, wird im Falle $\mathbb{K} = \mathbb{R}$ als *symmetrische Bilinearform* (kurz sBF) und im Falle $\mathbb{K} = \mathbb{C}$ als *Hermitesche Form* (kurz HF) bezeichnet. Für solche Formen verwendet man anstelle von $\sigma(x, y)$ häufig auch die Notation $\langle x, y\rangle$. Insbesondere gilt dann $\sigma(x, x) = \langle x, x\rangle \in \mathbb{R}$ für alle $x \in V$.

Beispielsweise definiert $\mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $(x, y) \mapsto x^{\mathtt{t}} \cdot \overline{y} = \sum_{i=1}^n x_i \cdot \overline{y}_i$ eine nicht-ausgeartete sBF bzw. HF.

Eine sBF bzw. HF $\sigma : V \times V \to \mathbb{K}$ heißt *positiv definit*, wenn $\sigma(x, x) > 0$ für alle $x \in V \setminus \{0\}$ gilt. Man spricht dann auch von einem *Skalarprodukt* auf V und nennt das Paar (V, σ) im Falle $\mathbb{K} = \mathbb{R}$ einen *euklidischen* Vektorraum, sowie im Falle $\mathbb{K} = \mathbb{C}$ einen *unitären* Vektorraum. Die oben betrachtete Form $\mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $(x, y) \mapsto x^{\mathtt{t}} \cdot \overline{y}$ ist ein Skalarprodukt auf $\mathbb{K}^n$, das so genannte *kanonische Skalarprodukt.*

Achtung! Es kann $\sigma(v_i, v_i) > 0$ für alle Vektoren v_i einer Basis sein, ohne daß σ positiv definit ist. [Beispiel: $\sigma(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) := x_1y_1 - x_2y_2$]

Beispiele.

Der Hilbertsche Folgenraum $\ell^2 = (V, \sigma)$ mit $V := \{(x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : \sum_{i=0}^\infty x_i^2 < \infty\}$ und $\sigma(x, y) := \sum_{i=0}^\infty x_iy_i$ ist ein euklidischer Vektorraum; ebenso der Raum $C[0, 1]$ aller stetigen Funktionen $f : [0, 1] \to \mathbb{R}$ mit punktweiser Addition und Skalarmultiplikation, und dem Skalarprodukt $\sigma(f, g) := \int_0^1 f(t)g(t)dt$.

Zum kanonischen Skalarprodukt im $\mathbb{R}^2$

Für $x \in \mathbb{R}^n$ $(n = 1, 2, 3)$ ist $\|x\| = \sqrt{\langle x, x\rangle} = \sum_{i=1}^n x_i^2$ die Länge des Vektors x bzw. der Abstand des Punktes x vom Nullpunkt. — Sei jetzt $0 \neq x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ und $0 \neq y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$.

Ist $\alpha_x \in [0, \pi]$ der Winkel zwischen e_1 und x, so gilt: $x_1 = \|x\| \cdot \cos\alpha_x$ und $x_2 = \|x\| \cdot \sin\alpha_x$.

Ist $\varphi \in [0, \pi]$ der Winkel zwischen x und y, so gilt $\varphi = |\alpha_x - \alpha_y|$ und somit

(1) $\cos(\varphi) = \cos(\alpha_x - \alpha_y) = \cos\alpha_x \cos\alpha_y + \sin\alpha_x \sin\alpha_y = \dfrac{x_1y_1 + x_2y_2}{\|x\| \cdot \|y\|} = \dfrac{\langle x, y\rangle}{\|x\| \cdot \|y\|}$,

(2) $\langle x, y\rangle = 0 \Leftrightarrow \cos(\varphi) = 0 \Leftrightarrow \varphi = \dfrac{\pi}{2} \Leftrightarrow x, y$ senkrecht zueinander,

(3) $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\cos(\varphi)\|x\|\|y\|$ (Cosinus-Satz)

[Beweis von (3): $\|x - y\|^2 = \langle x - y, x - y\rangle = \langle x, x\rangle - 2\langle x, y\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2 - 2\cos(\varphi)\|x\|\|y\|$]

Vereinbarung.
Soweit nichts anderes gesagt wird,
sei V im folgenden stets ein euklischer bzw. unitärer Vektorraum mit Skalarprodukt $\langle\ ,\ \rangle$.

Definition.

1. $\|v\| := \sqrt{\langle v, v\rangle}$ heißt die *Norm* (oder *Länge*) von $v \in V$.
2. $v, w \in V$ heißen *orthogonal* (oder *senkrecht*) zueinander (geschrieben $v \perp w$), falls $\langle v, w\rangle = 0$.

Lemma 12.1. (Cauchy-Schwarzsche Ungleichung)
Für alle $v, w \in V$ gilt: (i) $|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$, (ii) $|\langle v, w\rangle| = \|v\| \cdot \|w\| \Leftrightarrow v, w$ linear abhängig.

Beweis:
1. Seien v, w linear unabhängig. Setze $\lambda := \dfrac{\langle v, w\rangle}{\langle w, w\rangle}$. Dann gilt: $(*)$ $\langle v - \lambda w, w\rangle = \langle v, w\rangle - \lambda\langle w, w\rangle = 0$.
Ferner gilt $0 \neq v - \lambda w$ und somit $0 < \langle v - \lambda w, v - \lambda w\rangle \overset{(*)}{=} \langle v - \lambda w, v\rangle = \langle v, v\rangle - \lambda\langle w, v\rangle$.
Daraus folgt $0 < \langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle\langle w, v\rangle = \|v\|^2\|w\|^2 - \langle v, w\rangle\overline{\langle v, w\rangle}$ und weiter $|\langle v, w\rangle| < \|v\| \cdot \|w\|$.
2. Ist $w = \alpha v$, so $|\langle v, w\rangle| = |\alpha|\langle v, v\rangle = |\alpha| \cdot \|v\|^2 \overset{12.2a}{=} \|v\| \cdot \|w\|$.

Definition.
Wegen 12.1(i) kann man im Fall $\mathbb{K} = \mathbb{R}$ für $v, w \in V \setminus \{0\}$ definieren: $\sphericalangle(v, w) := \arccos \dfrac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \in [0, \pi]$.
Dann ist $\langle v, w\rangle = \|v\| \cdot \|w\| \cdot \cos\sphericalangle(v, w)$.

Lemma 12.2.

(a) Die Abbildung $V \to \mathbb{R}_+$, $v \mapsto \|v\|$ ist eine Norm, d.h. es gilt
- (i) $\|v\| = 0 \Leftrightarrow v = 0$,
- (ii) $\|\lambda v\| = |\lambda| \cdot \|v\|$ für alle $\lambda \in \mathbb{K}$,
- (iii) $\|v + w\| \leq \|v\| + \|w\|$.

(b) Die Abbildung $d : V \times V \to \mathbb{R}_+$, $d(v, w) := \|v - w\|$ ist eine Metrik, d.h. es gilt
- (i) $d(v, w) = 0 \Leftrightarrow v = w$,
- (ii) $d(v, w) = d(w, v)$,
- (iii) $d(v, w) \leq d(v, u) + d(u, w)$.

Beweis:
(a) (i) klar. (ii) $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda\tilde{\lambda}\langle v, v\rangle} = |\lambda|\sqrt{\langle v, v\rangle} = |\lambda| \cdot \|v\|$.

(iii) $\|v+w\|^2 = \langle v+w, v+w\rangle = \langle v, v\rangle + \langle w, w\rangle + \langle v, w\rangle + \langle w, v\rangle = \|v\|^2 + \|w\|^2 + \langle v, w\rangle + \overline{\langle v, w\rangle} \overset{(*)}{\leq}$
$\leq \|v\|^2 + \|w\|^2 + 2\|v\|{\cdot}\|w\| = (\|v\| + \|w\|)^2$. $\quad (*)$ $\langle v, w\rangle + \overline{\langle v, w\rangle} = 2\mathrm{Re}(\langle v, w\rangle) \leq 2|\langle v, w\rangle| \leq 2\|v\|{\cdot}\|w\|$.

(b) (ii) $\|v-w\| = \|-(w-v)\| = 1 \cdot \|w-v\|$. (iii) $\|v-w\| = \|(v-u)+(u-w)\| \leq \|v-u\| + \|u-w\|$.

Lemma 12.3.
Für alle $v, w, v_1, ..., v_k \in V$ gilt:
(a) $\|v_1 + \ldots + v_k\|^2 = \|v_1\|^2 + \ldots + \|v_k\|^2$, falls $v_i \perp v_j$ für alle $1 \leq i \neq j \leq k$.
(b) $v \perp w \;\Rightarrow\; \langle v, v \pm w\rangle = \|v\|^2$.
(c) $v \perp (w-v) \;\Rightarrow\; \langle v, w\rangle = \|v\|^2$.

Beweis:
(a) $\|v_1 + \ldots + v_k\|^2 = \langle \sum_{i=1}^k v_i, \sum_{i=1}^k v_i\rangle = \sum\limits_{1\leq i,j\leq k} \langle v_i, v_j\rangle = \sum_{i=1}^k \langle v_i, v_i\rangle$.
(b) $\langle v, v \pm w\rangle = \langle v, v\rangle \pm \langle v, w\rangle$. (c) $\langle v, w\rangle = \langle v, v + (w-v)\rangle$.

Definition.
Seien $v_1, ..., v_k \in V$.
$(v_1, ..., v_k)$ heißt *Orthogonalsystem* (OGS) $:\Longleftrightarrow$ $\begin{cases} v_1, \ldots, v_k \neq 0 \text{ und } \langle v_i, v_j\rangle = 0 \text{ für} \\ \text{alle } i, j \in \{1, ..., k\} \text{ mit } i \neq j \end{cases}$.
$(v_1, ..., v_k)$ heißt *Orthonormalsystem* (ONS) $:\Longleftrightarrow$ $\langle v_i, v_j\rangle = \delta_{ij}$ für alle $i, j \in \{1, ..., k\}$.
$(v_1, ..., v_k)$ heißt *Orthonormalbasis* (ONB) von V $:\Longleftrightarrow$ $(v_1, ..., v_k)$ ist ein ONS und zugleich Basis von V.

Bemerkung. Die Standardbasis des $\mathbb{K}^n$ ist eine ONB.

Lemma 12.4.
(a) Jedes OGS $(v_1, ..., v_k)$ ist linear unabhängig.
(b) Ist $(v_1, ..., v_n)$ eine ONB von V, so gilt $v = \sum\limits_{i=1}^n \langle v, v_i\rangle v_i$ für jedes $v \in V$. *(Entwicklungsformel)*

Beweis:
(a) $\sum_{i=1}^k \alpha_i v_i = 0 \;\Rightarrow\; \langle \sum_{i=1}^k \alpha_i v_i, v_j\rangle = 0 \;\Rightarrow\; \sum_{i=1}^k \alpha_i \langle v_i, v_j\rangle = 0 \;\Rightarrow\; \alpha_j \langle v_j, v_j\rangle = 0 \;\Rightarrow\; \alpha_j = 0$.
(b) Sei $v = \sum_{i=1}^n \alpha_i v_i$. Dann $\langle v, v_j\rangle = \sum_{i=1}^n \alpha_i \langle v_i, v_j\rangle = \alpha_j$.

Folgerung. Ist $\tilde{v} = (v_1, ..., v_n)$ eine ONB von V und $f \in \mathsf{End}(V)$, so gilt $\mathcal{M}_{\tilde{v}}(f) = \big(\langle f(v_j), v_i\rangle\big)_{i,j=1,\ldots,n}$.

Definition.
Für $M \subseteq V$ sei $M^\perp := \{x \in V : \forall y \in M (x \perp y)\}$. Dann gilt: $M \perp M^\perp$.
Ist M ein Untervektorraum, so heißt $M^\perp$ das *orthogonale Komplement* von M.

Lemma 12.5.
Für beliebige Teilmengen M, N von V gilt:
(a) $M^\perp$ ist ein Unterraum von V.
(b) $M \subseteq N \Rightarrow N^\perp \subseteq M^\perp$.
(c) $M^\perp = \mathrm{span}(M)^\perp$.
(d) $M \subseteq (M^\perp)^\perp$.

Beweis:
(a) $v, w \in M^\perp$ & $\lambda \in \mathbb{K} \Rightarrow \forall y \in M(v \perp y \;\&\; w \perp y) \Rightarrow \forall y \in M(v + \lambda w \perp y) \Rightarrow v + \lambda w \in M^\perp$.

(b) $v \in N^\perp \;\Rightarrow\; \forall y \in N(v \perp y) \overset{M \subseteq N}{\Longrightarrow} \forall y \in M(v \perp y) \;\Rightarrow\; v \in M^\perp$.

(c) $v \in M^\perp \;\Leftrightarrow\; M \subseteq \{v\}^\perp \overset{(a)}{\Leftrightarrow} \operatorname{span}(M) \subseteq \{v\}^\perp \;\Leftrightarrow\; v \in \operatorname{span}(M)^\perp$.

(d) $v \in M \;\Rightarrow\; \forall x \in M^\perp(x \perp v) \;\Rightarrow\; v \in (M^\perp)^\perp$.

Definition.
Zwei Unterräume $U, W \subseteq V$ heißen *orthogonal* (in Zeichen $U \perp W$), falls $\langle v, w\rangle = 0$ für alle $u \in U$, $w \in W$.
Eine Summe $U_1 + \ldots + U_k$ von Unterräumen U_i heißt *orthogonal*, falls $U_i \perp U_j$ für $i \neq j$.

Lemma 12.6.
(a) Jede orthogonale Summe ist direkt.
(b) $V = U \oplus W \;\&\; U \perp W \implies W = U^\perp \;\&\; U = (U^\perp)^\perp$.

Beweis :
(a) folgt aus 12.4a (genauso, wie 11.2b aus 11.2a).

(b) 1. $W \subseteq U^\perp$: trivial.
2. $U^\perp \subseteq W$: Sei $v \in U^\perp$. Dann $v = u + w$ mit $u \in U$ und $w \in W$.
Es folgt $0 = \langle u, v\rangle = \langle u, u\rangle + \langle u, w\rangle = \langle u, u\rangle$, also $u = 0$ und somit $v = w \in W$.
3. Aus Symmetriegründen gilt auch $U = W^\perp$ und folglich $U = (U^\perp)^\perp$.

Definition.
Ist $V = U \oplus U^\perp$, so sei pr_U die zu dieser Zerlegung gehörige Projektion von V auf U (vgl. 5.11), d.h. $\mathsf{pr}_U : V \to U$ ist die eindeutig bestimmte lineare Abbildung mit $\forall x \in U(\mathsf{pr}_U(x) = x)$ und $\forall x \in U^\perp(\mathsf{pr}_U(x) = 0)$.
Man nennt $\mathsf{pr}_U(v)$ die *orthogonale Projektion* von v auf U, und $v - \mathsf{pr}_U(v) \in U^\perp$ das *Lot* von v auf U.

Lemma 12.7.
Ist U ein Untervektorraum von V und $(v_1, ..., v_m)$ eine ONB von U, so gilt:
$V = U \oplus U^\perp$ und $\mathsf{pr}_U(v) = \sum_{i=1}^m \langle v, v_i\rangle v_i$ für alle $v \in V$.

Beweis :
Für $v \in V$ sei $f(v) := \sum_{i=1}^m \langle v, v_i\rangle v_i$.
Dann gilt $f(v) \in U$ und $\langle v - f(v), v_j\rangle = \langle v, v_j\rangle - \sum_{i=1}^m \langle v, v_i\rangle\langle v_i, v_j\rangle = \langle v, v_j\rangle - \langle v, v_j\rangle = 0$ für $j = 1, ..., m$, also $v - f(v) \in U^\perp$ und folglich $v \in U + U^\perp$. Damit haben wir bewiesen, daß $V = U \oplus U^\perp$ ist und f, $\mathrm{id}_V - f$ die zu dieser Zerlegung gehörenden Projektionen sind, d.h. insbesondere $f = \mathsf{pr}_U$.

Lemma 12.8.
Sei U ein Untervektorraum von V mit $V = U \oplus U^\perp$. Dann gilt für alle $v \in V$:
(a) $v - \mathsf{pr}_U(v) \perp \mathsf{pr}_U(v)$ und folglich $\|v\|^2 = \|\mathsf{pr}_U(v)\|^2 + \|v - \mathsf{pr}_U(v)\|^2$.
(b) $\mathsf{pr}_U(v) \neq u \in U \;\Rightarrow\; \|v - \mathsf{pr}_U(v)\| < \|v - u\|$.

Beweis von (b): Wegen $0 \neq \mathsf{pr}_U(v) - u \in U$ und $v - \mathsf{pr}_U(v) \in U^\perp$ gilt
$\|v - u\|^2 = \|v - \mathsf{pr}_U(v)\|^2 + \|\mathsf{pr}_U(v) - u\|^2 > \|v - \mathsf{pr}_U(u)\|^2$.

Satz 12.9. (Schmidtsches Orthonormalisierungsverfahren).

(a) Sei $(w_1, ..., w_{k-1})$ ein ONS in V und $v \in V \setminus W$, wobei $W := \text{span}(w_1, ..., w_{k-1})$.
Sei ferner $w := v - \mathsf{pr}_W(v)$ (das Lot von v auf W) und $w_k := \dfrac{w}{\|w\|}$.
Dann ist $(w_1, ..., w_k)$ ein ONS und $\text{span}(w_1, ..., w_k) = \text{span}(w_1, ..., w_{k-1}, v)$.

(b) Sei $(v_1, ..., v_n)$ eine Basis von V, und sei für $k = 1, ..., n$: $w_k := \dfrac{w'_k}{\|w'_k\|}$, wobei $w'_k := v_k - \sum\limits_{i=1}^{k-1} \langle v_k, w_i\rangle w_i$.
Dann ist $(w_1, ..., w_n)$ eine ONB von V und für $k = 1, ..., n$ gilt $\text{span}(v_1, ..., v_k) = \text{span}(w_1, ..., w_k)$.

Beweis:

(a) Nach 12.7 ist $V = W \oplus W^\perp$, also ist pr_W definiert und $\mathsf{pr}_W(v) \in W$, sowie $w = v - \mathsf{pr}_W(v) \in W^\perp$. Wegen $v \notin W$ ist außerdem $w \neq 0$. Es folgt $w_k \in W^\perp$ und $\|w_k\| = 1$. Folglich ist $(w_1, ..., w_k)$ ein ONS.

Aus $w_k = \frac{1}{\|w\|}(v - \mathsf{pr}_W(v))$ und $\mathsf{pr}_W(v) \in W$ folgt $w_k \in \text{span}(w_1, ..., w_{k-1}, v)$ und $v \in \text{span}(w_1, ..., w_k)$ und somit $\text{span}(w_1, ..., w_k) = \text{span}(w_1, ..., w_{k-1}, v)$.

(b) Mittels (a) zeigt man durch Induktion nach k, daß $(w_1, ..., w_k)$ ein ONS mit
$\text{span}(w_1, ..., w_k) = \text{span}(v_1, ..., v_k)$ ist.

Korollar.

(a) Ist $\dim(V) < \infty$, so besitzt V eine ONB, und jedes ONS in V kann zu einer ONB von V ergänzt werden.

(b) Für jeden endlichdimensionalen Unterraum U von V gilt $V = U \oplus U^\perp$.

Beweis: (a) Sei $\tilde{u} = (v_1, \ldots, v_m)$ ein ONS in V. $\tilde{u}$ ist linear unabhängig, kann also zu einer Basis von $(v_1, \ldots, v_n)$ von V ergänzt werden. Aus dieser erhält man mittels 12.9b eine ONB $(w_1, \ldots, w_n)$ von V. Wie man leicht sieht, gilt für $i = 1, \ldots, m$ $w_i = v_i$. (b) folgt aus 12.7 und 12.9.

Im folgenden sei V ein *endlichdimensionaler* euklidischer bzw. unitärer Raum und
U, W seien Untervektorräume von V.

Bemerkung. Nach Lemma 12.6b und Korollar (b) zu 12.9 gilt: $V = U \oplus U^\perp$ und $(U^\perp)^\perp = U$.

Definition.

Für $\emptyset \neq M, M' \subseteq V$ definiert man den *Abstand von* M *und* M' durch
$d(M, M') := \inf\{\|x - y\| : x \in M \;\&\; y \in M'\}$.

Für $q \in V$ sei $d(q, M) := d(M, q) := d(M, \{q\}) = \inf\{\|x - q\| : x \in M\}$.

Lemma 12.10.

Sei $M = p + U$ und sei $q \in V$.

Der Punkt $\ell := p + \mathsf{pr}_U(q - p)$ heißt dann der Fußpunkt des Lotes von q auf M. – Es gilt:

(a) $\ell \in M$ und $q - \ell = \mathsf{pr}_{U^\perp}(q - p)$

(b) $\forall x \in M \setminus \{\ell\}(q - x \notin U^\perp \;\&\; \|q - \ell\| < \|q - x\|)$

(c) $d(q, M) = \|q - \ell\| = (\|q - p\|^2 - \|\mathsf{pr}_U(q - p)\|^2)^{\frac{1}{2}}$

Beweis :

(a) Nach Definition von pr_U gilt $\mathsf{pr}_U(q - p) \in U$ und $q - \ell = (q - p) - \mathsf{pr}_U(q - p) = \mathsf{pr}_{U^\perp}(q - p)$.

(b) 1. Sei $x \in M$ mit $q - x \in U^\perp$. Dann gilt $q - p = (q - x) + (x - p)$ mit $q - x \in U^\perp$ und $x - p \in U$.
Folglich $x = p + (x - p) = p + \mathsf{pr}_U(q - p) = \ell$.

2. Sei $\ell \neq x \in M$. Dann $0 < \|\ell - x\|$ und $\ell - x \in U$, $q - \ell \in U^\perp$. Es folgt $\|q-\ell\|^2 < \|q-\ell\|^2 + \|\ell - x\|^2 = \|q-x\|^2$.
(c) $d(q, M) = \|q - \ell\|$ folgt aus $\ell \in M$ und (b). Die zweite Gleichung folgt aus (a) und 12.3a.

Lemma 12.11 (Abstand zweier affiner Unterräume)
(a) $d(p + U, q + W) = d(p - q, U + W) = \|\mathsf{pr}_{(U+W)^\perp}(p - q)\|$.
(b) Für $a \in p + U$ und $b \in q + W$ sind äquivalent:
(i) $d(p + U, q + W) = \|a - b\|$, (ii) $a - b \in (U + W)^\perp$, (iii) $a - b = \mathsf{pr}_{(U+W)^\perp}(p - q)$.
(c) $a, a' \in p + U$ & $b, b' \in q + W$ & $a - b, a' - b' \in (U + W)^\perp$ & $U \cap W = \{0\} \Rightarrow a = a'$ & $b = b'$.
(d) Es gibt $a \in p + U$ und $b \in q + W$ mit $a - b \in (U + W)^\perp$.

Beweis :
(a) $\{x - y : x \in p + U \ \&\ y \in q + W\} = \{(p - u) - (q + w) : u \in U \ \&\ w \in W\} = \{(p - q) - v : v \in U + W\}$.
Folglich $d(p + U, q + W) = d(p - q, U + W) \overset{12.10}{=} \|\mathsf{pr}_{(U+W)^\perp}(p - q)\|$.
(b) Es gilt $\|\mathsf{pr}_{(U+W)^\perp}(p - q)\| \overset{(a)}{=} d(p + U, q + W) = d(a + U, b + W) \overset{(a)}{=} \|\mathsf{pr}_{(U+W)^\perp}(a - b)\|$.
Folglich: $d(p + U, q + W) = \|a - b\| \Leftrightarrow \|\mathsf{pr}_{(U+W)^\perp}(a - b)\| = \|a - b\| \overset{12.8a}{\Leftrightarrow} a - b \in (U + W)^\perp$.
Ferner gilt $(p - q) - (a - b) \in U + W$ und deshalb $\big(a - b = \mathsf{pr}_{(U+W)^\perp}(p - q) \Leftrightarrow a - b \in (U + W)^\perp\big)$.
(c) Mit (b) folgt $a - b = \mathsf{pr}_{(U+W)^\perp}(q - p) = a' - b'$ und somit $a - a' = b - b' \in U \cap W = \{0\}$

(d) Sei $a = p - u$ und $b = q + w$ mit $u \in U$ & $w \in W$ & $u + w = \mathsf{pr}_{U+W}(p - q)$.
Dann $a - b = p - u - q - w = (p - q) - (u + w) \in (U + W)^\perp$.

Lemma 12.12.
Für $U, H \subseteq V$ und $n = \dim(V) < \infty$ gilt:
(a) U ist UVR der Dimension $\dim(V) - 1 \iff \exists c \in V\setminus\{0\}(U = \{c\}^\perp)$.
(b) H ist Hyperebene $\iff \exists c \in V\setminus\{0\}, \alpha \in \mathbb{K}(H = \{x \in V : \langle x, c\rangle - \alpha = 0\})$.

Beweis:
(a) Es ist $V = U \oplus U^\perp$, also $n = \dim(U) + \dim(U^\perp)$ und $U = U^{\perp\perp}$.
$\dim(U) = n - 1 \Leftrightarrow \dim(U^\perp) = 1 \Leftrightarrow \exists c \in V\setminus\{0\}(U^\perp = \mathbb{K}c) \Leftrightarrow \exists c \in V\setminus\{0\}(U = (\mathbb{K}c)^\perp)$.

(b) H Hyperebene $\Leftrightarrow H = p + U$ mit $\dim(U) = n - 1 \overset{(a)}{\Leftrightarrow} \exists p, c \in V(c \neq 0 \ \&\ H = p + \{c\}^\perp) \Leftrightarrow$
$\Leftrightarrow \exists p, c \in V(c \neq 0 \ \&\ H = \{x : \langle x - p, c\rangle = 0\}) \Leftrightarrow \exists p, c \in V(c \neq 0 \ \&\ H = \{x : \langle x, c\rangle - \langle p, c\rangle = 0\})$

Bemerkung.
Ist $H = \{x \in V : \langle x, c\rangle - \alpha = 0\}$ mit $c \neq 0$, so $H = \lambda c + \{c\}^\perp$ mit $\lambda := \dfrac{\alpha}{\|c\|^2}$.
Ist $\|c\| = 1$, so nennt man „ $\langle x, c\rangle - \alpha = 0$ " die *Hessesche Normalform* der Hyperebene H.

Lemma 12.13 (Abstand eines Punktes zu einer Hyperebene)
Ist $H = \{x \in V : \langle x, c\rangle - \alpha = 0\}$ mit $\|c\| = 1$, so gilt für jedes $q \in V$: $d(q, H) = |\langle q, c\rangle - \alpha|$.

Beweis:
Es ist $H = \alpha c + U$ mit $U = \{c\}^\perp$, $U^\perp = \mathbb{K}c$.
Nach 12.10 gilt nun $d(q, H) = \|q - \ell\| = \|\mathsf{pr}_{U^\perp}(q - \alpha c)\| = |\langle q - \alpha c, c\rangle| = |\langle q, c\rangle - \alpha|$.

Bemerkung. (Schnittpunkt von Gerade und Hyperebene)

Ist die Gerade $G = p + \mathbb{K}v$ nicht parallel zur Hyperebene $H = \{x : \langle x, c\rangle - \alpha = 0\}$,
so ist $s := p - \dfrac{\langle p, c\rangle - \alpha}{\langle v, c\rangle}\cdot v$ der Schnittpunkt von G und H.

Beweis: G nicht parallel zu $H \Rightarrow v \notin \{c\}^\perp \Rightarrow \langle v, c\rangle \neq 0$ & $\langle s, c\rangle - \alpha = \langle p, c\rangle - \frac{\langle p,c\rangle - \alpha}{\langle v,c\rangle}\langle v, c\rangle - \alpha = 0.$

Definition (Spiegelung an einer Hyperebene).

Für $a \in V$ mit $\|a\| = 1$ sei $\mathsf{s}_a : V \to V$, $\mathsf{s}_a(x) := x - 2\langle x, a\rangle a$.

Offenbar ist s_a die Spiegelung an der zu a senkrechten Hyperebene $\{a\}^\perp$.

Lemma 12.14.

Für $a \in V$ mit $\|a\| = 1$ gilt:

(a) s_a ist lineare Abbildung mit $\mathsf{s}_a(a) = -a$ und $\mathsf{s}_a(x) = x$ für alle $x \in \{a\}^\perp$.

(b) Ist $\overline{v} = (v_1, ..., v_{n-1}, a)$ eine ONB von V, so gilt $\mathcal{M}_{\bar{v}}(\mathsf{s}_a) = \begin{pmatrix} E & 0 \\ 0 & -1 \end{pmatrix}$.

(c) $\mathsf{s}_a \circ \mathsf{s}_a = \mathrm{id}$.

(d) $\det(\mathsf{s}_a) = -1$.

Beweis: (a), (b) klar. (c),(d) folgen aus (b).

Das Vektorprodukt im $\mathbb{R}^3$

Schreibweise: $|A| := \det(A)$.

Definition.

Für $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ und $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3$ sei $x \times y := \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}$ *(Vektorprodukt)*

Merkregel: $x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = e_1 \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - e_2 \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + e_3 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$

Rechenregeln. Für $x, x', y, y', z \in \mathbb{R}^3$ gilt:

($\times$1) $(x + x') \times y = x \times y + x' \times y$,

($\times$2) $x \times (y + y') = x \times y + x \times y'$,

($\times$3) $\lambda x \times y = \lambda(x \times y) = x \times \lambda y$,

($\times$4) $y \times x = -x \times y$, also $x \times x = 0$,

($\times$5) $x \times y = 0 \Leftrightarrow x, y$ linear abhängig,

($\times$6) $\langle x \times y, z\rangle = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$,

($\times$7) $\langle x \times y, x\rangle = \langle x \times y, y\rangle = 0$,

($\times$8) $\|x \times y\|^2 = \|x\|^2\|y\|^2 - \langle x, y\rangle^2$,

($\times$9) $\|x \times y\| = \|x\| \cdot \|y\| \cdot \sin \sphericalangle(x, y)$.

Beweise:

(×5) "⇐": $x \times \lambda x = \lambda(x \times x) = 0$. "⇒": x, y linear unabhängig ⇒ $\operatorname{rang}\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} = 2$ ⇒ $\dim \operatorname{span}(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ y_3 \end{pmatrix}) = 2$ ⇒ o.E.d.A. $(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix})$ lin.unabh. ⇒ $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \neq 0$.

(×6) $\langle x \times y, z\rangle = z_1 \cdot \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - z_2 \cdot \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + z_3 \cdot \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$

(×7) folgt aus (×6)

(×8) $\|x\|^2\|y\|^2 - \langle x, y\rangle^2 = (x_1^2 + x_2^2 + x_3^2) \cdot (y_1^2 + y_2^2 + y_3^2) - x_1^2y_1^2 - x_2^2y_2^2 - x_3^2y_3^2 - 2x_1y_1x_2y_2 - 2x_1y_1x_3y_3 - 2x_2y_2x_3y_3 = x_2^2y_3^2 + x_3^2y_2^2 + x_3^2y_1^2 + x_1^2y_3^2 + x_1^2y_2^2 + x_2^2y_1^2 - 2x_2y_3x_3y_2 - 2x_3y_1x_1y_3 - 2x_1y_2x_2y_1 = \|x \times y\|^2$

(×9) Sei $\vartheta := \sphericalangle(x, y)$. Nach Definition ist $\vartheta \in [0, \pi]$ und somit $\sin\vartheta \geq 0$.

$\|x \times y\|^2 = \|x\|^2\|y\|^2 - \langle x, y\rangle^2 = \|x\|^2\|y\|^2 - \|x\|^2\|y\|^2 \cos^2\vartheta = \|x\|^2\|y\|^2(1 - \cos^2\vartheta) = \|x\|^2\|y\|^2 \cdot \sin^2\vartheta$.

Lemma 12.15.

$u, w \in \mathbb{R}^3$ & $v = u \times w \neq 0$ & $c = \lambda u + \mu w \;\Rightarrow\; \lambda = \dfrac{\langle c \times w, v\rangle}{\|v\|^2}$ und $\mu = \dfrac{\langle u \times c, v\rangle}{\|v\|^2}$.

Beweis: $c \times w = \lambda(u \times w)$ & $u \times c = \mu(u \times w) \;\Rightarrow\; \langle c \times w, v\rangle = \lambda\|v\|^2$ & $\langle u \times c, v\rangle = \mu\|v\|^2$.

Lemma 12.16.

Seien $G = p + \mathbb{R}u$, $G' = q + \mathbb{R}w$ Geraden im $\mathbb{R}^3$. G, G' seien nicht parallel, d.h. $v := u \times w \neq 0$ und $\mathbb{R}u \cap \mathbb{R}w = \{0\}$. Nach Lemma 12.11 gibt es eindeutig bestimmte Punkte $a \in G$, $b \in G'$ mit $d(G, G') = \|a - b\|$.

Behauptung: $a = p + \lambda u$ und $b = q + \mu w$ mit $\lambda = \dfrac{\langle w \times (p-q), v\rangle}{\|v\|^2}$, $\mu = \dfrac{\langle u \times (p-q), v\rangle}{\|v\|^2}$.

Beweis:

Nach 12.11 ist zu zeigen: $\mathsf{pr}_{(\mathbb{R}u+\mathbb{R}w)^\perp}(p - q) = (p + \lambda u) - (q + \mu w)$.

Sei $c := \mathsf{pr}_{\mathbb{R}u+\mathbb{R}w}(p - q)$. Nach 12.15 ist dann $c = (-\lambda')u + \mu' w$ mit $\lambda' = \dfrac{\langle w \times c, v\rangle}{\|v\|^2}$ und $\mu' = \dfrac{\langle u \times c, v\rangle}{\|v\|^2}$.

Wegen $\mathbb{R}v = (\mathbb{R}u + \mathbb{R}w)^\perp$ gilt andererseits $c = (p - q) - \mathsf{pr}_{\mathbb{R}v}(p - q) = p - q + \alpha v$ für ein $\alpha \in \mathbb{R}$, und folglich $\langle w \times c, v\rangle = \langle w \times (p-q), v\rangle + \langle w \times (\alpha v), v\rangle = \langle w \times (p-q), v\rangle$, sowie $\langle u \times c, v\rangle = \langle u \times (p-q), v\rangle$.

Somit $\mathsf{pr}_{(\mathbb{R}u+\mathbb{R}w)^\perp}(p - q) = p - q - c = (p + \lambda' u) - (q + \mu' w) = (p + \lambda u) - (q + \mu w)$.

Definition (Gramsche Determinante)

Für $v_1, ..., v_m \in V$ sei $\quad G(v_1, ..., v_m) := \det\begin{pmatrix} \langle v_1, v_1\rangle & \dots & \langle v_1, v_m\rangle \\ \vdots & & \vdots \\ \langle v_m, v_1\rangle & \dots & \langle v_m, v_m\rangle \end{pmatrix}$.

Ist U ein Untervektorraum von V und $\tilde{u} = (u_1, ..., u_m)$ eine ONB von U, so definiert man

$\det_{\tilde{u}} : U^m \to \mathbb{K}$, $\det_{\tilde{u}}(v_1, ..., v_m) := \det(\Phi_{\tilde{u}}^{-1}(v_1) \;\dots\; \Phi_{\tilde{u}}^{-1}(v_m))$.

Die Abbildung $\det_{\tilde{u}} : U^n \to \mathbb{K}$ ist multilinear (d.h. linear in jedem Argument) und alternierend (d.h. $\det_{\tilde{u}}(v_1, ..., v_m) = 0$, falls $\exists i, j \in \{1, ..., m\}(i \neq j \;\&\; v_i = v_j)$).

Ist $(v_1, \dots, v_m)$ linear abhängig, so $G(v_1, \dots, v_m) = 0$.

Lemma 12.17. Sei $(v_1, ..., v_m)$ linear unabhängig in V.

(a) $G(v_1, ..., v_m) = |\det_{\tilde{u}}(v_1, ..., v_m)|^2$, falls $\tilde{u}$ eine ONB von $\mathrm{span}(v_1, ..., v_m)$.

(b) Ist $(v_1, ..., v_m)$ ein OGS, so $G(v_1, ..., v_m) = \|v_1\|^2 \cdot \ldots \cdot \|v_m\|^2$.

(c) $1 \le i \neq j \le m \;\Rightarrow\; G(v_1, ..., v_{i-1}, v_i + \lambda v_j, v_{i+1}, ..., v_m) = G(v_1, ..., v_m)$.

Beweis:

(a) Sei $A = (a_{ij})$ mit $v_j = \sum_i a_{ij} u_i$. Dann ist $\det_{\tilde{u}}(v_1, ..., v_m) = \det(A)$ und $a_{ij} = \langle v_j, u_i \rangle$ (weil $\tilde{u}$ ONB).

Ferner gilt $\langle v_i, v_j \rangle = \sum_{\mu,\nu} a_{\mu i} \overline{a_{\nu j}} \delta_{\mu\nu} = \sum_\nu a_{\nu i} \overline{a_{\nu j}}$, also $(\langle v_i, v_j \rangle)_{i,j} = A^{\mathsf{t}} \cdot \overline{A}$ und deshalb
$G(v_1, ..., v_m) = \det\big(\langle v_i, v_j \rangle\big)_{i,j} = \det(A^{\mathsf{t}})\det(\overline{A}) = \det(A)\overline{\det(A)} = |\det(A)|^2$.

(b) Sei $u_i := \|v_i\|^{-1} v_i$. Dann ist $\tilde{u} := (u_1, ..., u_m)$ eine ONB von $\mathrm{span}(v_1, ..., v_m)$ und folglich
$\det_{\tilde{u}}(v_1, ..., v_m) = \|v_1\| \cdots \|v_m\| \cdot \det_{\tilde{u}}(u_1, ..., u_m) = \|v_1\| \cdots \|v_m\|$, woraus mit (a) die Behauptung folgt.

(c) Sei $i = 1$. $G(v_1 + \lambda v_j, v_2, ..., v_m) = |\det_{\tilde{u}}(v_1 + \lambda v_j, v_2, ..., v_m)|^2 = |\det_{\tilde{u}}(v_1, ..., v_m)|^2 = G(v_1, ..., v_m)$.

Abkürzung. $\mathit{vol}(v_1, ..., v_m) := \sqrt{G(v_1, ..., v_m)}$

Lemma 12.18.

(a) $\mathit{vol}(v_1) = \sqrt{\det(\langle v_1, v_1 \rangle)} = \|v_1\|$.

(b) $\mathit{vol}(v_1, v_2) = \|v_1\| \cdot \|v_2\| \cdot \sin \sphericalangle(v_1, v_2)$.

(c) $\mathit{vol}(v_1, ..., v_m) = \mathit{vol}(v_1, ..., v_{m-1}) \cdot \|v'_m\|$, wobei $v'_m := v_m - \mathsf{pr}_U(v_m)$ mit $U := \mathrm{span}(v_1, ..., v_{m-1})$

Beweis:

(a) $\mathit{vol}(v_1) = \sqrt{\det(\langle v_1, v_1 \rangle)} = \|v_1\|$.

(b) Mit $\alpha := \sphericalangle(v_1, v_2) \in [0, \pi]$ gilt:
$\mathit{vol}(v_1, v_2) = (\|v_1\|^2 \cdot \|v_2\|^2 - \langle v_1, v_2 \rangle^2)^{1/2} = (\|v_1\|^2 \|v_2\|^2 - \|v_1\|^2 \|v_2\|^2 \cos^2 \alpha)^{1/2} = \|v_1\| \cdot \|v_2\| \cdot \sin \alpha$.

(c) Fall 1: $(v_1, ..., v_{m-1})$ lin. abh.: Dann $\mathit{vol}(v_1, \ldots, v_m) = 0 = \mathit{vol}(v_1, \ldots, v_{m-1})$.
Fall 2: $v_m \in \mathrm{span}(v_1, \ldots, v_{m-1})$: Dann $\mathit{vol}(v_1, \ldots, v_m) = 0$ und $v'_m = 0$.
Fall 3: $(v_1, \ldots, v_m)$ lin. unabh.: Sei $\tilde{u}' = (u_1, ..., u_{m-1})$ eine ONB von U und $u_m := \|v'_m\|^{-1} v'_m$.
Dann ist $\tilde{u} := (u_1, ..., u_m)$ eine ONB von $\mathrm{span}(v_1, \ldots, v_m) = \mathrm{span}(v_1, \ldots, v_{m-1}, v'_m)$.
Ferner $\mathit{vol}(v_1, \ldots, v_{m-1}) = |\det(A)|$, wobei $A = (a_{ij}) \in \mathbb{K}^{m-1 \times m-1}$ mit $v_j = \sum_{i=1}^{m-1} a_{ij} u_i$ $(j = 1, ..., m-1)$, sowie $\mathit{vol}(v_1, \ldots, v_m) \overset{12.17c}{=} \mathit{vol}(v_1, \ldots, v_{m-1}, v'_m) = |\det(A')|$, wobei $A' = \begin{pmatrix} A & 0 \\ 0 & \|v'_m\| \end{pmatrix}$ (wegen $v'_m = \|v'_m\| \cdot u_m$). Folglich $\mathit{vol}(v_1, \ldots, v_m) = |\det(A)| \cdot \|v'_m\| = \mathit{vol}(v_1, \ldots, v_{m-1}) \cdot \|v'_m\|$.

Bemerkung. Ist $(v_1, \ldots, v_m)$ linear unabhängig., so nennt man $\mathit{vol}(v_1, \ldots, v_m)$ das (m-dimensionale) Volumen des von $v_1, \ldots, v_m$ aufgespannten Parallelotops.

§13 Sesquilinearformen und Matrizen; die adjungierte Abbildung

Definition.

Ist V ein $\mathbb{K}$-Vektorraum mit Basis $\tilde{v} = (v_1, ..., v_n)$ und σ eine Sesquilinearform auf V, so nennt man $\mathcal{M}_{\tilde{v}}(\sigma) := \big(\sigma(v_i, v_j)\big)_{i,j} \in \mathbb{K}^{n\times n}$ die *darstellende Matrix von σ bzgl. $\tilde{v}$.*

Bemerkung.

(1) Für $x = (x_1 \, \dots \, x_n)^{\mathtt{t}} \in \mathbb{K}^n$ und $y = (y_1 \, \dots \, y_n)^{\mathtt{t}} \in \mathbb{K}^n$ gilt offenbar

$\sigma(\sum_i x_i v_i, \sum_i y_i v_i) = \sum_{i,j} x_i \sigma(v_i, v_j)\overline{y_j} = x^{\mathtt{t}} \cdot \mathcal{M}_{\tilde{v}}(\sigma) \cdot \overline{y}.$

(2) $A, B \in \mathbb{K}^{n\times n}$ & $\forall x, y \in \mathbb{K}^n (x^{\mathtt{t}} A\overline{y} = x^{\mathtt{t}} B\overline{y}) \Rightarrow A = B.$ [zum Beweis: $e_i^{\mathtt{t}} A\overline{e_j} = a_{ij}$]

Lemma 13.1.

Sei V ein $\mathbb{K}$-Vektorraum mit Basis $\tilde{v} = (v_1, ..., v_n)$, und sei $Sesq(V)$ die Menge aller Sesquilinearformen $\sigma : V \times V \to \mathbb{K}$. Dann gilt:

(a) Die Abbildung $Sesq(V) \to \mathbb{K}^{n\times n}$, $\sigma \mapsto \mathcal{M}_{\tilde{v}}(\sigma)$ ist bijektiv.

(b) $\sigma \in Sesq(V)$ ist genau dann eine sBF bzw. HF, wenn $\mathcal{M}_{\tilde{v}}(\sigma) = \overline{\mathcal{M}_{\tilde{v}}(\sigma)}^{\mathtt{t}}$ ist.

(c) Für $\sigma \in Sesq(V)$ sind äquivalent:

(i) $\forall v \in V (\forall w \in V (\sigma(v, w) = 0) \Rightarrow v = 0)$,

(ii) $\forall w \in V (\forall v \in V (\sigma(v, w) = 0) \Rightarrow w = 0)$,

(iii) $\det \mathcal{M}_{\tilde{v}}(\sigma) \neq 0$.

Beweis :

(a) injektiv: $\mathcal{M}_{\tilde{v}}(\sigma) = \mathcal{M}_{\tilde{v}}(\tau) \Rightarrow \sigma(v_i, v_j) = \tau(v_i, v_j)\ (\forall i, j) \overset{(1)}{\Rightarrow} \sigma(v, w) = \tau(v, w)\ (\forall v, w \in V).$

surjektiv: Sei $A = (a_{ij}) \in \mathbb{K}^{n\times n}$.

Dann wird durch $\sigma(\sum_i x_i v_i, \sum_i y_i v_i) := \sum_{i,j} x_i a_{ij} \overline{y_j}$ eine Sesquilinearform σ mit $\mathcal{M}_{\tilde{v}}(\sigma) = A$ definiert.

(b) Sei $A := \mathcal{M}_{\tilde{v}}(\sigma)$. Dann gilt $\forall x, y \in \mathbb{K}^n (\overline{y^{\mathtt{t}} A\overline{x}} = \overline{y}^{\mathtt{t}} \overline{A} x = x^{\mathtt{t}}\, \overline{A}^{\mathtt{t}} \overline{y})$ und folglich:
$\forall v, w \in V (\sigma(v, w) = \overline{\sigma(w, v)}) \Leftrightarrow \forall x, y \in \mathbb{K}^n (x^{\mathtt{t}} A\overline{y} = \overline{y^{\mathtt{t}} A\overline{x}}) \Leftrightarrow \forall x, y \in \mathbb{K}^n (x^{\mathtt{t}} A\overline{y} = x^{\mathtt{t}}\, \overline{A}^{\mathtt{t}} \overline{y}) \Leftrightarrow A = \overline{A}^{\mathtt{t}}.$

(c) Wir zeigen nur $\neg(i) \Leftrightarrow \neg(iii)$. Sei $A := \mathcal{M}_{\tilde{v}}(\sigma)$. $\exists v \in V \setminus \{0\} \forall w \in V (\sigma(v, w) = 0) \Leftrightarrow \exists x \in \mathbb{K}^n \setminus \{0\}$ $\forall y \in \mathbb{K}^n (x^{\mathtt{t}} A\overline{y} = 0) \Leftrightarrow \exists x \in \mathbb{K}^n \setminus \{0\} (x^{\mathtt{t}} A = 0) \Leftrightarrow \mathsf{f}_{A^t}$ nicht injektiv $\Leftrightarrow \det(A^{\mathtt{t}}) = 0 \Leftrightarrow \det(A) = 0.$

Lemma 13.2 (Transformationsformel).

Sei V ein endlichdim. $\mathbb{K}$-Vektorraum mit Basen $\tilde{v}, \tilde{w}$, und sei $S := \mathcal{M}^{\tilde{w}}_{\tilde{v}}(\mathrm{id}_V)$ die zugehörige Transformationsmatrix. Für jede Sesquilinearform σ auf V gilt dann: $\mathcal{M}_{\tilde{w}}(\sigma) = S^{\mathtt{t}} \mathcal{M}_{\tilde{v}}(\sigma)\overline{S}$.

Beweis :

Es ist $S = (c_{ij})_{ij}$ mit $w_j = \sum_i c_{ij} v_i$. Folglich: $\sigma(w_k, w_l) = \sigma(\sum_i c_{ik} v_i, \sum_i c_{il} v_i) = \sum_i \sum_j c_{ik} \sigma(v_i, v_j)\overline{c_{jl}}$.

Lemma 13.3. Sei V ein $\mathbb{K}$-Vektorraum, $\tilde{v} = (v_1, ..., v_n)$ eine Basis von V und $\sigma \in Sesq(V)$. Dann gilt:

(a) σ Skalarprodukt $\iff$ es existiert ein $S \in \mathrm{GL}(n; \mathbb{K})$ mit $S^{\mathtt{t}} \mathcal{M}_{\tilde{v}}(\sigma)\overline{S} = E_n$.

(b) σ Skalarprodukt $\Longrightarrow \det \mathcal{M}_{\tilde{v}}(\sigma) > 0$.

Beweis :

(a) "$\Leftarrow$": Sei $S = (c_{ij})$ und $w_j := \sum_{i=1}^n c_{ij} v_i$. Dann $\mathcal{M}_{\tilde{w}}(\sigma) = S^{\mathtt{t}} \mathcal{M}_{\tilde{v}}(\sigma)\overline{S} = E_n$ und somit $\sigma(\sum_i x_i w_i, \sum_i y_i w_i) = \sum_{i,j} x_i \delta_{ij} \overline{y_j} = x^{\mathtt{t}} \overline{y}$ für alle $x, y \in \mathbb{K}^n$. Folglich ist σ ein Skalarprodukt.

(a) "$\Rightarrow$" und (b): Sei $\tilde{w}$ eine ONB von (V,σ). Dann $S := \mathcal{M}_{\tilde{v}}^{\tilde{w}}(\mathrm{id}_V) \in \mathrm{GL}(n;\mathbb{K})$ und $E_n = \mathcal{M}_{\tilde{w}}(\sigma) = S^{\mathtt{t}}\mathcal{M}_{\tilde{v}}(\sigma)\overline{S}$. Mit $T := S^{-1}$ folgt daraus $T^{\mathtt{t}}\overline{T} = \mathcal{M}_{\tilde{v}}(\sigma)$; also $\det\mathcal{M}_{\tilde{v}}(\sigma) = \det(T) \cdot \overline{\det(T)} = |\det(T)|^2 > 0$.

Satz 13.4.
Sei V ein $\mathbb{K}$-Vektorraum mit Basis $(v_1, .., v_n)$ und sei $\sigma : V \times V \to \mathbb{K}$ eine sBF bzw. HF.
Für $r = 1, ..., n$ sei $A_r := (\sigma(v_i, v_j))_{i,j=1,...,r}$. — Dann gilt:

σ positiv definit $\iff$ $\det(A_r) > 0$ für $r = 1, ..., n$.

Beweis:

"$\Rightarrow$": Lemma 13.3b.

"$\Leftarrow$": Sei jetzt $\det(A_r) > 0$ für $r = 1, ..., n$. Durch Induktion nach n zeigen wir, daß σ positiv definit ist.

1. $n = 1$: $\sigma(\lambda v_1, \lambda v_1) = |\lambda|^2\sigma(v_1, v_1) = |\lambda|^2\det(A_1) > 0$, falls $\lambda v_1 \neq 0$.

2. $n > 1$: Sei $U := \mathrm{span}(v_1, ..., v_{n-1})$. Nach I.V. ist $\sigma\restriction U$ positiv definit. Nach 12.9 besitzt U eine ONB $(w_1, ..., w_{n-1})$ bzgl. $\sigma\restriction U$. Ferner ist $\tilde{w} := (w_1, ..., w_{n-1}, w_n)$ mit $w_n := v_n - \sum\limits_{i=1}^{n-1}\sigma(v_n, w_i)w_i$ eine Basis von V mit $\mathcal{M}_{\tilde{w}}(\sigma) = \begin{pmatrix} E_{n-1} & 0 \\ 0 & \sigma(w_n, w_n) \end{pmatrix}$. Bleibt zu zeigen $\sigma(w_n, w_n) > 0$. Nach 13.2 existiert $S \in \mathrm{GL}(n;\mathbb{K})$ mit $\mathcal{M}_{\tilde{w}}(\sigma) = S^{\mathtt{t}}\mathcal{M}_{\tilde{v}}(\sigma)\overline{S} = S^{\mathtt{t}}A_n\overline{S}$. Folglich $\sigma(w_n, w_n) = \det\mathcal{M}_{\tilde{w}}(\sigma) = |\det(S)|^2 \cdot \det(A_n) > 0$.

Die adjungierte Abbildung

Satz 13.5.
Sei $\sigma : V \times V \to \mathbb{K}$ eine nicht-ausgeartete Sesquilinearform auf dem endlichdim. $\mathbb{K}$-Vektorraum V.

(a) Zu jeder Linearform $\varphi \in V^*$ existiert genau ein $u \in V$ mit $\forall x \in V(\varphi(x) = \sigma(x, u))$.

(b) Zu jedem $f \in \mathsf{End}(V)$ existiert genau eine Abbildung $f^{\mathrm{ad}} : V \to V$ mit $\sigma(f(v), w) = \sigma(v, f^{\mathrm{ad}}(w))$ für alle $v, w \in V$. Diese ist linear und wird *der zu f adjungierte* Endomorphismus genannt.

Beweis :

(a) Für $v \in V$ bezeichne $\sigma(\cdot, v)$ die Linearform $V \to \mathbb{K}$, $x \mapsto \sigma(x, v)$. Ferner bezeichne τ die Abbildung $V \to V^*$, $\tau(v) := \sigma(\cdot, v)$, und $\widehat{V}$ den $\mathbb{K}$-Vektorraum $(V, +, \bullet)$ mit $\bullet : \mathbb{K} \times V \to V$, $\lambda\bullet v := \overline{\lambda}\cdot v$. (Man rechnet leicht nach, daß $\widehat{V}$ tatsächlich ein $\mathbb{K}$-Vektorraum ist.)

Hilfssatz. τ ist ein Isomorphismus von $\widehat{V}$ auf V^*.

Beweis: 1. τ linear, denn $\tau(v + \lambda\bullet w)(x) = \sigma(x, v + \lambda\bullet w) = \sigma(x, v) + \sigma(x, \overline{\lambda}\cdot w) = \sigma(x, v) + \lambda\sigma(x, w) = \tau(v) + \lambda\tau(w)$. 2. τ injektiv, da σ nicht ausgeartet: $\tau(v) = 0 \Rightarrow \forall x \in V(\sigma(x, v) = 0) \Rightarrow v = 0$.
3. Da V endlichdim. und jede Basis von V auch Basis von $\widehat{V}$ ist, gilt $\dim(V^*) = \dim(V) = \dim(\widehat{V})$.

Nach Hilfssatz existiert zu jedem $\varphi \in V^*$ genau ein u mit $\varphi = \tau(u)$, d.h. mit $\forall x \in V(\varphi(x) = \sigma(x, u))$.

(b) Sei $w \in V$ gegeben. Die Abbildung $v \overset{\varphi}{\mapsto} \sigma(f(v), w)$ ist linear, d.h. ein Element von V^*. Nach (a) existiert genau ein $w' \in V$ mit $\forall v \in V(\sigma(f(v), w) = \sigma(v, w'))$; wir setzen $f^{\mathrm{ad}}(w) := w'$.

Linearität von $w \mapsto f^{\mathrm{ad}}(w)$: $\sigma(v, f^{\mathrm{ad}}(w_1 + \lambda w_2)) = \sigma(f(v), w_1 + \lambda w_2) = \sigma(f(v), w_1) + \overline{\lambda}\sigma(f(v), w_2) = \sigma(v, f^{\mathrm{ad}}(w_1)) + \overline{\lambda}\sigma(v, f^{\mathrm{ad}}(w_2)) = \sigma(v, f^{\mathrm{ad}}(w_1) + \lambda f^{\mathrm{ad}}(w_2))$ $(\forall v)$ $\Rightarrow$ $f^{\mathrm{ad}}(w_1 + \lambda w_2) = f^{\mathrm{ad}}(w_1) + \lambda f^{\mathrm{ad}}(w_2)$.

Im folgenden sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$.

Satz 13.6. Die Abbildung $\mathsf{End}(V)\to\mathsf{End}(V)$, $f\mapsto f^{\mathrm{ad}}$ ist semi-linear. Außerdem gilt für alle $f\in\mathsf{End}(V)$:

(a) $(f^{\mathrm{ad}})^{\mathrm{ad}}=f$, d.h. $\langle f^{\mathrm{ad}}(v),w\rangle=\langle v,f(w)\rangle$ für alle $v,w\in V$.

(b) $\mathrm{Ker}(f^{\mathrm{ad}})=(\mathrm{Im}f)^{\perp}$ und $\mathrm{Ker}(f)=(\mathrm{Im}f^{\mathrm{ad}})^{\perp}$.

(c) $\mathrm{rang}(f)=\mathrm{rang}(f^{\mathrm{ad}})$.

(d) Ist $\tilde v$ eine ONB von V (bzgl. $\langle\cdot,\cdot\rangle$), so gilt $\mathcal{M}_{\tilde v}(f^{\mathrm{ad}})=\overline{\mathcal{M}_{\tilde v}(f)}^{\,t}$ für alle $f\in\mathsf{End}(V)$.

Beweis :

$\forall v,w\in V[\langle f(v),w\rangle=\langle v,f^{\mathrm{ad}}(w)\rangle\ \&\ \langle g(v),w\rangle=\langle v,g^{\mathrm{ad}}(w)\rangle]\ \Rightarrow\ \forall v,w\in V[\langle (f+\lambda g)(v),w\rangle=\langle v,f^{\mathrm{ad}}(w)\rangle+\lambda\langle v,g^{\mathrm{ad}}(w)\rangle=\langle v,f^{\mathrm{ad}}(w)+\overline{\lambda}g^{\mathrm{ad}}(w)\rangle]\ \Rightarrow\ \forall w\in V[(f+\lambda g)^{\mathrm{ad}}(w)=(f^{\mathrm{ad}}+\overline{\lambda}g^{\mathrm{ad}})(w)]$.

(a) $\langle f^{\mathrm{ad}}(v),w\rangle=\overline{\langle w,f^{\mathrm{ad}}(v)\rangle}=\overline{\langle f(w),v\rangle}=\langle v,f(w)\rangle$.

(b) $w\in\mathrm{Ker}(f^{\mathrm{ad}})\Leftrightarrow f^{\mathrm{ad}}(w)=0\Leftrightarrow\forall v(\langle v,f^{\mathrm{ad}}(w)\rangle=0)\Leftrightarrow\forall v(\langle f(v),w\rangle=0)\Leftrightarrow w\in\mathrm{Im}(f)^{\perp}$.

(c) $\dim(\mathrm{Im}f^{\mathrm{ad}})=n-\dim\mathrm{Ker}(f^{\mathrm{ad}})=n-\dim((\mathrm{Im}f)^{\perp})=\dim\mathrm{Im}(f)$.

(d) $\mathcal{M}_{\tilde v}(f^{\mathrm{ad}})\overset{12.4\mathrm{b}}{=}\big(\langle f^{\mathrm{ad}}(v_j),v_i\rangle\big)_{i,j}\overset{(a)}{=}\big(\langle v_j,f(v_i)\rangle\big)_{i,j}=\big(\overline{\langle f(v_i),v_j\rangle}\big)_{i,j}$. $\mathcal{M}_{\tilde v}(f)\overset{12.4\mathrm{b}}{=}\big(\langle f(v_i),v_j\rangle\big)_{j,i}$.

Definition. Ein Endomorphismus $f:V\to V$ heißt *normal*, wenn $f\circ f^{\mathrm{ad}}=f^{\mathrm{ad}}\circ f$ gilt.

Satz 13.7. Für jedes $f\in\mathsf{End}(V)$ gilt:

(a) f normal $\iff$ $\langle f(v),f(w)\rangle=\langle f^{\mathrm{ad}}(v),f^{\mathrm{ad}}(w)\rangle$ für alle $v,w\in V$.

(b) f normal $\Longrightarrow$ $\mathrm{Ker}(f)=\mathrm{Ker}(f^{\mathrm{ad}})$ und $\mathrm{Eig}(f;\lambda)=\mathrm{Eig}(f^{\mathrm{ad}},\overline{\lambda})$ für alle $\lambda\in\mathbb{K}$.

Beweis :

(a) "$\Rightarrow$": f normal $\Rightarrow$ $\langle f(v),f(w)\rangle=\langle v,f^{\mathrm{ad}}f(w)\rangle=\langle v,ff^{\mathrm{ad}}(w)\rangle\overset{13.6\mathrm{a}}{=}\langle f^{\mathrm{ad}}(v),f^{\mathrm{ad}}(w)\rangle$.

"$\Leftarrow$": $\langle v,f^{\mathrm{ad}}f(w)\rangle=\langle f(v),f(w)\rangle\overset{\text{Vorauss.}}{=}\langle f^{\mathrm{ad}}(v),f^{\mathrm{ad}}(w)\rangle\overset{13.6\mathrm{a}}{=}\langle v,ff^{\mathrm{ad}}(w)\rangle$.

(b) 1. $f(v)=0\Leftrightarrow\langle f(v),f(v)\rangle=0\overset{(a)}{\Leftrightarrow}\langle f^{\mathrm{ad}}(v),f^{\mathrm{ad}}(v)\rangle=0\Leftrightarrow f^{\mathrm{ad}}(v)=0$.

2. Nach 13.6 ist $(f-\lambda\mathrm{id})^{\mathrm{ad}}=f^{\mathrm{ad}}-\overline{\lambda}\mathrm{id}$. Daraus folgert man, daß auch $f-\lambda\mathrm{id}$ normal ist.

Somit gilt $\mathrm{Eig}(f;\lambda)=\mathrm{Ker}(f-\lambda\cdot\mathrm{id})\overset{1.}{=}\mathrm{Ker}(f^{\mathrm{ad}}-\overline{\lambda}\cdot\mathrm{id})=\mathrm{Eig}(f^{\mathrm{ad}};\overline{\lambda})$.

Satz 13.8 (Spektralsatz für normale Abbildungen).

Es sei $f\in\mathsf{End}(V)$ und das charakteristische Polynom $P_f\in\mathbb{K}[t]$ zerfalle in Linearfaktoren.

Dann gilt: f normal $\iff$ V besitzt eine ONB aus Eigenvektoren von f.

Beweis:

"$\Rightarrow$": Induktion nach $n=\dim(V)$. Sei $n\geq 1$. Nach Voraussetzung besitzt f einen Eigenwert λ. Sei v_1 ein Eigenvektor von λ mit $\|v_1\|=1$. Dann gilt $V=\mathbb{K}v_1\oplus U$ mit $U:=(\mathbb{K}v_1)^{\perp}$. Es gilt $f(U)\subseteq U$ und $f^{\mathrm{ad}}(U)\subseteq U$. [Bew.: $\langle x,v_1\rangle=0\Rightarrow\langle f(x),v_1\rangle=\langle x,f^{\mathrm{ad}}(v_1)\rangle\overset{13.7\mathrm{b}}{=}\langle x,\overline{\lambda}v_1\rangle=0$ und $\langle f^{\mathrm{ad}}(x),v_1\rangle=\langle x,f(v_1)\rangle=\langle x,\lambda v_1\rangle=\overline{\lambda}\langle x,v_1\rangle=0$]. Daraus folgt, daß auch $f|U$ normal ist. Ferner gilt $P_f=(\lambda-t)\cdot P_{f|U}$ (siehe $(*)$). Also zerfällt auch $P_{f|U}$ in Linearfaktoren und wir können die I.V. auf $f|U$ anwenden. Danach besitzt U eine ONB $(v_2,\dots,v_n)$ aus Eigenvektoren von $f|U$. Wegen $v_1\perp U$ ist dann $(v_1,v_2,\dots,v_n)$ eine ONB von V bestehend aus Eigenvektoren von f. $(*)$ Sei $\tilde u=(u_2,\dots,u_n)$ eine Basis von U und $B:=\mathcal{M}_{\tilde u}(f|U)$. Dann ist $\tilde v:=(v_1,u_2,\dots,u_n)$ Basis von V und $\mathcal{M}_{\tilde v}(f)=\begin{pmatrix}\lambda&0\\0&B\end{pmatrix}$. Es folgt:

$P_f=\det(\begin{pmatrix}\lambda&0\\0&B\end{pmatrix}-t\cdot E_n)=\det\begin{pmatrix}\lambda-t&0\\0&B-t\cdot E_{n-1}\end{pmatrix}=(\lambda-t)\cdot\det(B-t\cdot E_{n-1})=(\lambda-t)\cdot P_{f|U}$.

"$\Leftarrow$": Sei $\tilde v$ eine ONB aus Eigenvektoren von f. Dann ist $D:=\mathcal{M}_{\tilde v}(f)$ eine Diagonalmatrix. Mit 13.6d folgt $\mathcal{M}_{\tilde v}(f^{\mathrm{ad}})=\overline{D}$. Da D, $\overline{D}$ Diagonalmatrizen sind, gilt $D\cdot\overline{D}=\overline{D}\cdot D$, woraus $f\circ f^{\mathrm{ad}}=f^{\mathrm{ad}}\circ f$ folgt.

Definitionen.

1. Ein Endomorphismus $f : V \to V$ heißt *selbstadjungiert*, falls $f = f^{\mathrm{ad}}$.
2. Ist $A \in \mathbb{K}^{n\times n}$, so nennt man $\overline{A}^{\mathtt{t}}$ die zu A *adjungierte* Matrix.
3. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *hermitesch* bzw. (im Fall $\mathbb{K} = \mathbb{R}$) *symmetrisch*, falls $A = \overline{A}^{\mathtt{t}}$.
4. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *unitär* bzw. (im Fall $\mathbb{K} = \mathbb{R}$) *orthogonal*, falls $A^{-1} = \overline{A}^{\mathtt{t}}$.

Bemerkungen.

1. Ist $\tilde{v}$ ONB von V und $f \in \mathsf{End}(V)$, so gilt: f selbstadjungiert $\iff$ $\mathcal{M}_{\tilde{v}}(f)$ hermitesch.
2. $A \in \mathbb{K}^{n\times n}$ ist genau dann hermitesch, wenn $\overline{A} = A^{\mathtt{t}}$.
3. Für $A \in \mathbb{K}^{n\times n}$ sind folgende Bedingungen äquivalent:
   - A ist unitär (bzw. orthogonal), d.h. $A \cdot \overline{A}^{\mathtt{t}} = \overline{A}^{\mathtt{t}} \cdot A = E_n$.
   - $A^{\mathtt{t}} \cdot \overline{A} = E_n$
   - Die Spalten von A bilden eine ONB von $\mathbb{K}^n$.
   - Die Zeilen von A bilden eine ONB von $\mathbb{K}^{1\times n}$.
4. Ist $A \in \mathbb{K}^{n\times n}$ unitär bzw. orthogonal, so $|\det(A)| = 1$.

Beispiel

$$A = \begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix}. \qquad P_A = (10-t)\cdot \begin{vmatrix} -14-t & 2 \\ 2 & -11-t \end{vmatrix} - 5\cdot \begin{vmatrix} 5 & 10 \\ 2 & -11-t \end{vmatrix} + 10\cdot \begin{vmatrix} 5 & 10 \\ -14-t & 2 \end{vmatrix} =$$

$(10-t)[t^2 + 25t + 150] - 5[-5t - 75] + 10[150 + 10t] = (15+t)(100 - t^2 + 125) = (15+t)(15+t)(15-t)$.

$$\mathrm{Eig}(A; 15) = \mathbb{R} \cdot \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix}, \text{ denn: } \begin{pmatrix} -5 & 5 & 10 \\ 5 & -29 & 2 \\ 10 & 2 & -26 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -2 \\ 0 & -24 & 12 \\ 0 & 12 & -6 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -2 \\ 0 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathrm{Eig}(A; -15) = \mathrm{Eig}(A; 15)^{\perp} = \mathbb{R} \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}. \quad \left[\begin{pmatrix} 25 & 5 & 10 \\ 5 & 1 & 2 \\ 10 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 5 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right]$$

$$\left(\tfrac{1}{\sqrt{30}} \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix}, \tfrac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}, \tfrac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \right) \text{ ist ONB aus EV. } S = \begin{pmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{-2}{\sqrt{5}} & \frac{-1}{\sqrt{6}} \\ \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \end{pmatrix}, S^{\mathtt{t}}AS = \begin{pmatrix} 15 & 0 & 0 \\ 0 & -15 & 0 \\ 0 & 0 & -15 \end{pmatrix}$$

Lemma 13.9.

Ist $f \in \mathsf{End}(V)$ selbstadjungiert (bzw. $A \in \mathbb{C}^{n\times n}$ hermitesch), so ist P_f (bzw. P_A) von der Form $(\lambda_1 - t)\cdot \ldots \cdot (\lambda_n - t)$ mit $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Folglich sind alle Eigenwerte von f (bzw. A) reell.

Beweis:

Nach L.11.4e und dem Fundamentalsatz der Algebra gilt $P_A = (\lambda_1 - t)\cdot \ldots \cdot (\lambda_n - t)$ mit $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Noch zu zeigen: $\lambda_i \in \mathbb{R}$ $(i = 1, ..., n)$. Da $\lambda := \lambda_i$ Eigenwert von A ist, existiert ein $v \in \mathbb{C} \setminus \{0\}$ mit $Av = \lambda v$. Dann $\lambda\langle v, v\rangle = \langle Av, v\rangle \overset{\text{A herm.}}{=} \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$ und $\langle v, v\rangle \neq 0$, also $\lambda = \overline{\lambda}$, d.h. $\lambda \in \mathbb{R}$.
Die Aussage für selbstadjungiertes $f \in \mathsf{End}(V)$ folgt mit obiger Bemerkung 1.

Satz 13.10. (Hauptachsentransformation oder Spektralsatz für selbstadjungierte Abbildungen)
Ist $f \in \mathsf{End}(V)$ selbstadjungiert, so besitzt V eine ONB aus Eigenvektoren von f.

Beweis: f selbstadj. $\Rightarrow$ f normal $\overset{13.8+13.9}{\Rightarrow}$ Behauptung.

Korollar. Ist $A \in \mathbb{K}^{n\times n}$ hermitesch, so besitzt $\mathbb{K}^n$ eine ONB aus Eigenvektoren von A.

Lemma 13.11.

Ist $A \in \mathbb{K}^{n\times n}$ hermitesch und $(v_1, \dots, v_n)$ eine ONB von $\mathbb{K}^n$ mit $v_j \in \mathrm{Eig}(A; \lambda_j)$ für $j = 1, \dots, n$, so gilt: $S^{-1}AS = Diag(\lambda_1, ..., \lambda_n) \in \mathbb{R}^{n\times n}$, wobei $S := (v_1\ \dots\ v_n) \in \mathbb{K}^{n\times n}$ und $S^{-1} = \overline{S}^{\mathtt{t}}$.

Beweis :

$Av_j = \lambda_j v_j \ \ (j = 1, \dots, n) \overset{(11.5b)}{\Rightarrow} S^{-1}AS = Diag(\lambda_1, \dots, \lambda_n)$. $\qquad \lambda_j$ Eigenwert von $A \overset{13.9}{\Rightarrow} \lambda_j \in \mathbb{R}$.

Satz 13.12. (Hauptachsentransformation für symmetrische Bilinearformen)

Zu jeder sBF bzw. HF σ auf V existiert eine ONB $\tilde{w}$ von $(V, \langle\cdot,\cdot\rangle)$, so daß $\mathcal{M}_{\tilde{w}}(\sigma)$ reelle Diagonalmatrix ist.

Beweis:

Sei $\tilde{v}$ eine ONB von V und $A := \mathcal{M}_{\tilde{v}}(\sigma)$. Dann $A = \overline{A}^{\mathtt{t}}$. Nach 13.11 gibt es eine Diagonalmatrix $D \in \mathbb{R}^{n\times n}$ und eine unitäre Matrix S mit $\overline{S}^{\mathtt{t}}AS = D$, also $C^{\mathtt{t}}A\overline{C} = D$, wobei $C := (c_{ij})_{i,j} := \overline{S}$. Sei $\tilde{w} := (w_1, \dots, w_n)$ mit $w_j := \sum_i c_{ij}v_i$. Da C unitär ist, ist auch $\tilde{w}$ eine ONB. Nach 13.2 gilt $\mathcal{M}_{\tilde{w}}(\sigma) = C^{\mathtt{t}}\mathcal{M}_{\tilde{v}}(\sigma)\overline{C} = D$.

Beispiel

Frage: Welche geometrische Gestalt hat die Menge $Q := \{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : \alpha x_1^2 + 2\beta x_1x_2 + \alpha x_2^2 = 1\}$?

Sei $A := \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$. Dann $Q = \{x \in \mathbb{R}^2 : x^{\mathtt{t}}Ax = 1\}$. Wir bestimmen eine ONB aus Eigenvektoren von A.

$P_A = (\alpha - t)^2 - \beta^2 = (\alpha + \beta - t)(\alpha - b + t)$. Eigenwerte: $\lambda_1 = \alpha + \beta$, $\lambda_2 = \alpha - \beta$.

$A - \lambda_1 E = \begin{pmatrix} -\beta & \beta \\ \beta & -\beta \end{pmatrix}$, $v_1 := \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. $\quad A - \lambda_2 E = \begin{pmatrix} \beta & \beta \\ \beta & \beta \end{pmatrix}$, $v_2 := \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

$S := (v_1\ v_2) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$; $\quad D := S^{\mathtt{t}}AS = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

$Q = \{x \in \mathbb{R}^2 : x^{\mathtt{t}}Ax = 1\} = \{Sy : y \in \mathbb{R}^2\ \&\ (Sy)^{\mathtt{t}}A(Sy) = 1\} = \{Sy : y \in \mathbb{R}^2\ \&\ y^{\mathtt{t}}Dy = 1\} =$

$= \{y_1v_1 + y_2v_2 : \lambda_1 y_1^2 + \lambda_2 y_2^2 = 1\} \overset{(*)}{=} \{y_1v_1 + y_2v_2 : \frac{y_1^2}{a^2} \pm \frac{y_2^2}{b^2} = 1\}$.

$(*)$ Ist $\lambda_1 > 0$ und $\lambda_2 \neq 0$, so setzen wir $a := \frac{1}{\sqrt{\lambda_1}}$ und $b := \frac{1}{\sqrt{|\lambda_2|}}$.

Im Fall $+$ ist Q eine Ellipse, im Fall $-$ eine Hyperbel, a und b sind jeweils die *Hauptachsen*.

§14 Orthogonale und unitäre Endomorphismen

In diesem Abschnitt sei V stets ein endlichdimensionaler euklidischer bzw. unitärer Vektorraum.

Definition.

Ein Endomorphismus $f : V \to V$ heißt *orthogonal* bzw. *unitär*, wenn $\langle f(v), f(w)\rangle = \langle v, w\rangle$ für alle $v, w \in V$.

Lemma 14.1.

Für jeden orthogonalen bzw. unitären Endomorphismus $f : V \to V$ gilt

(a) $\|f(v)\| = \|v\|$ für alle $v \in V$.

(b) $v, w \in V$ & $v \perp w \;\Rightarrow\; f(v) \perp f(w)$.

(c) f ist Isomorphismus und f^{-1} ist ebenfalls orthogonal bzw. unitär.

(d) Ist $\lambda \in \mathbb{K}$ ein Eigenwert von f, so $|\lambda| = 1$.

Beweis:

(c) Wegen (a) ist f injektiv, also ein Isomorphismus.

Ferner $\langle f^{-1}(v), f^{-1}(w)\rangle = \langle f(f^{-1}(v)), f(f^{-1}(w))\rangle = \langle v, w\rangle$ für alle $v, w \in V$.

(d) $\|v\| = \|f(v)\| = \|\lambda v\| = |\lambda| \cdot \|v\| \;\Rightarrow\; |\lambda| = 1$.

Bemerkung.

$f \in \mathsf{End}(V)$ & $\forall v \in V(\,\|f(v)\| = \|v\|\,) \implies f$ orthogonal bzw. unitär. (Beweis folgt später.)

Lemma 14.2.

Für $f \in \mathsf{End}(V)$ gilt: f orthogonal bzw. unitär $\iff f^{\mathrm{ad}} \circ f = \mathrm{id}_V$ (d.h. $f^{\mathrm{ad}} = f^{-1}$).

Beweis:

"$\Rightarrow$": $\forall v, w(\langle f(v), w\rangle = \langle f(v), f(f^{-1}(w))\rangle = \langle v, f^{-1}(w)\rangle) \;\Rightarrow\; f^{\mathrm{ad}} = f^{-1}$.

"$\Leftarrow$": $f^{\mathrm{ad}} \circ f = \mathrm{id}_V \;\Rightarrow\; \langle f(v), f(w)\rangle = \langle v, f^{\mathrm{ad}}(f(w))\rangle = \langle v, w\rangle$.

Korollar. Jeder orthogonale bzw. unitäre Endomorphismus von V ist normal.

Lemma 14.3. Sei $f \in \mathsf{End}(V)$ und $\tilde{v} = (v_1, \ldots, v_n)$ eine ONB von V. Dann sind äquivalent:

(i) f ist orthogonal bzw. unitär.

(ii) $\mathcal{M}_{\tilde{v}}(f)$ ist orthogonal bzw. unitär.

(iii) $(f(v_1), ..., f(v_n))$ ist ONB von V.

Beweis :

Sei $A := \mathcal{M}_{\tilde{v}}(f)$. Nach 13.6d ist dann $\overline{A}^{\mathsf{t}} = \mathcal{M}_{\tilde{v}}(f^{\mathrm{ad}})$.

"(i)$\Leftrightarrow$(ii)": f orthog. bzw. unitär $\overset{14.2}{\Leftrightarrow} f^{\mathrm{ad}} \circ f = \mathrm{id}_V \;\Leftrightarrow\; \overline{A}^{\mathsf{t}} \cdot A = E$.

"(i)$\Rightarrow$(iii)": $\langle f(v_i), f(v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$.

"(iii)$\Rightarrow$(i)": $\langle f(\sum_i x_i v_i), f(\sum_j y_j v_j)\rangle = \sum_{i,j} x_i \overline{y_j} \langle f(v_i), f(v_j)\rangle = \sum_{i,j} x_i \overline{y_j} \langle v_i, v_j\rangle = \langle \sum_i x_i v_i, \sum_j y_j v_j\rangle$.

Korollar.

(a) Ist $f \in \mathsf{End}(V)$ orthogonal bzw. unitär, so $|\det(f)| = 1$.

(b) Ist $\tilde{v} = (v_1, ..., v_n)$ eine ONB von V und $w_j = \sum_{i=1}^n c_{ij} v_i$ für $j = 1, ..., n$, so gilt:
$\tilde{w}$ ONB von $V \iff S := (c_{ij}) \in \mathbb{K}^{n \times n}$ ist orthogonal bzw. unitär.

Definition.

$O(n) := \{A \in \mathrm{GL}(n;\mathbb{R}) : A^{-1} = A^{\mathsf{t}}\}$, *(orthogonale Gruppe)*

$SO(n) := \{A \in O(n) : \det(A) = 1\}$, (*spezielle orthogonale Gruppe)*

$U(n) := \{A \in \mathrm{GL}(n;\mathbb{C}) : A^{-1} = \overline{A}^{\mathsf{t}}\}$, *(unitäre Gruppe)*

Bemerkung.

1. $O(n)$ ist Untergruppe von $\mathrm{GL}(n;\mathbb{R})$, $SO(n)$ Untergruppe von $O(n)$ und $U(n)$ Untergruppe von $\mathrm{GL}(n;\mathbb{C})$.
2. Die Matrizen aus $SO(n)$ nennt man auch *eigentlich orthogonal.*

Lemma 14.4.

Ist $A \in O(2)$, so gibt es ein $\alpha \in [0, 2\pi[$, so daß $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ oder $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$.

Im ersten Fall ist f_A eine *Drehung* um den Winkel α und $\det(A) = 1$.

Im zweiten Fall ist f_A eine *Spiegelung* an der Geraden $\mathbb{R}\cdot \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}$ und $\det(A) = -1$

und es gilt $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Beweis:

Sei $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in O(2)$. Dann $a^2 + b^2 = 1$, $c^2 + d^2 = 1$ und $ac + bd = 0$. Für die komplexen Zahlen $z := a+i\cdot b$ und $z' := d+i\cdot c$ gilt daher $|z| = |z'| = 1$ und $zz' = (ad-bc)+i(ac+bd) = ad-bc = \det(A) = \pm 1$.

Fall 1: $\det(A) = 1$. $zz' = 1 = |z|^2 = z\overline{z} \Rightarrow z' = \overline{z} \Rightarrow d = a \;\&\; c = -b \Rightarrow A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Wegen $a^2 + b^2 = 1$ existiert (genau) ein $\alpha \in [0, 2\pi[$ mit $a = \cos\alpha$ und $b = \sin\alpha$.

Fall 2: $\det(A) = -1$. Analog zu Fall 1 folgt jetzt $d = -a \;\&\; c = b$.

Für den Rest siehe §11, S.58.

Lemma 14.5.

Sei $A \in O(3)$. Dann ist $\det(A) = \pm 1$ ein Eigenwert von A, und es gibt eine ONB $\tilde{v}$ und ein $\alpha \in [0, 2\pi[$, so daß $\mathcal{M}_{\tilde{v}}(\mathsf{f}_A) = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$.

Im Fall $+1$ ist f_A die Drehung mit Drehachse $\mathbb{R}v_1$ um den Winkel α.

Der Unterraum $\mathbb{R}v_2 + \mathbb{R}v_3$ wird in diesem Fall *Drehebene* genannt.

Im Fall -1 gilt $\mathcal{M}_{\tilde{v}}(\mathsf{f}_A) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$.

f_A setzt sich also zusammen aus der Drehung um die Achse $\mathbb{R}v_1$ mit Winkel α und der Spiegelung an der Drehebene $\mathbb{R}v_2 + \mathbb{R}v_3$.

Beweis:

1. Wir zeigen, daß $\det(A)$ Eigenwert von A ist: Das charakteristische Polynom von A hat den Grad 3 und besitzt deshalb mindestens eine reelle Nullstelle λ. Man wähle eine ONB $\tilde{v} = (v_1, v_2, v_3)$ mit $Av_1 = \lambda v_1$. Die darstellende Matrix von f_A bzgl. dieser ONB hat die Gestalt $\begin{pmatrix} \lambda & 0 \\ 0 & C \end{pmatrix}$ mit $C \in O(2)$.

Es gilt also $\det(A) = \lambda\cdot\det(C)$ und somit $\det(A) = \lambda$, falls $\det(C) = 1$. Ist dagegen $\det(C) = -1$, so folgt mit 14.4, daß C und folglich auch A die Eigenwerte $1, -1$ hat.

2. Da wir jetzt schon wissen, daß $\det(A)$ Eigenwert von A ist, können wir in 1. o.E.d.A. $\lambda = \det(A)$ annehmen. Wegen $\det(A) = \lambda \cdot \det(C) \neq 0$ muß dann aber $\det(C) = 1$ und damit $C \in SO(n)$ sein. Nach 14.4 ist C also von der Form $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

Aus dem Fundamentalsatz der Algebra zusammen mit Satz 13.8 und dem Korollar zu 14.2 folgt:

Satz 14.6.

Ist V unitär und $f \in \mathsf{End}(V)$ unitär, so besitzt V eine ONB aus Eigenvektoren von f.

Korollar. Ist $A \in \mathbb{C}^{n\times n}$ unitär, so besitzt $\mathbb{C}^n$ eine ONB aus Eigenvektoren von A.

Lemma 14.7.

Ist V euklidisch, $\dim(V) \geq 1$ und $f \in \mathsf{End}(V)$ orthogonal,
so besitzt V einen Untervektorraum W mit $f(W) = W$ und $\dim(W) \in \{1, 2\}$.

Beweis:

Sei $\tilde{v}$ eine ONB von V und $A := \mathcal{M}_{\tilde{v}}(f)$. Wir betrachten den Endomorphismus $g : \mathbb{C}^n \to \mathbb{C}^n$, $g(z) := Az$.
Da A orthogonal ist, ist g unitär. Sei $0 \neq z \in \mathbb{C}^n$ und $\lambda \in \mathbb{C}$, so daß $Az = \lambda z$.
Es gibt $x, y \in \mathbb{R}^n$ mit $z = x + \mathrm{i}y$. Dann gilt $Ax + \mathrm{i}Ay = Az = \lambda z = (\alpha + \mathrm{i}\beta)(x + \mathrm{i}y) = (\alpha x - \beta y) + \mathrm{i}(\beta x + \alpha y)$.
Wegen $A \in \mathbb{R}^{n\times n}$ folgt daraus $Ax = \alpha x - \beta y$ und $Ay = \beta x + \alpha y$. Wir haben also $\mathsf{f}_A(\mathbb{R}x + \mathbb{R}y) \subseteq \mathbb{R}x + \mathbb{R}y$.
Daraus folgt $f(W) \subseteq W$ für $W := \Phi_{\tilde{v}}(\mathbb{R}x + \mathbb{R}y)$. Da f injektiv ist, gilt sogar $f(W) = W$.

Satz 14.8.

Ist V euklidisch und $f \in \mathsf{End}(V)$ orthogonal, so besitzt V eine ONB $\tilde{v}$ mit

$$\mathcal{M}_{\tilde{v}}(f) = \begin{pmatrix} +1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & +1 & & & \mathbf{0} & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & \mathbf{0} & & & -1 & & & \\ & & & & & & A_1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & A_k \end{pmatrix},$$

wobei für $j = 1, ..., k$

$$A_j = \begin{pmatrix} \cos\vartheta_j & -\sin\vartheta_j \\ \sin\vartheta_j & \cos\vartheta_j \end{pmatrix} \in SO(2)$$

mit $\pi \neq \vartheta_j \in \,]0, 2\pi[$.

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

Nach Lemma 14.7 existiert ein Unterraum $W \subseteq V$ mit $1 \leq \dim(W) \leq 2$ und $f(W) = W$. Dann ist auch $f^{-1}(W) = W$, woraus folgt $\forall v \in W^\perp \forall w \in W(\langle f(v), w\rangle \overset{\text{L.14.2}}{=} \langle v, f^{-1}(w)\rangle = 0)$, d.h. $f(W^\perp) \subseteq W^\perp$.

Damit haben wir f zerlegt in zwei orthogonale Abbildungen $g := f|W \in \mathsf{End}(W)$ und $h := f|W^\perp \in \mathsf{End}(W^\perp)$. Wegen $\dim(W^\perp) < n$ können wir auf h die I.V. anwenden und erhalten eine Basis $\tilde{v}'$ von $W^\perp$ der gewünschten Art. Ist $\dim(W) = 1$, so gibt es einen Eigenvektor $w \in W$ mit $\|w\| = 1$ zu einem Eigenwert ± 1. Ergänzt man $\tilde{v}'$ an passender Stelle durch w zu $\tilde{v}$, so hat diese Basis von V die gewünschten Eigenschaften.

Im Fall $\dim(W) = 2$ gibt es eine ONB (w_1, w_2) von W, so $\mathcal{M}_{(w_1,w_2)}(g)$ eine der folgenden Formen hat
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ oder $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ oder $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ oder $\begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}$ mit $\pi \neq \vartheta \in \,]0, 2\pi[$.
Indem man w_1 und w_2 an passenden Stellen in $\tilde{v}'$ einfügt, erhält man eine Basis $\tilde{v}$ der gewünschten Art.

Satz 14.9.
Sei V ein euklidischer Vektorraum und $f : V \to V$ eine (nicht notwendig lineare) Abbildung.
f heißt *Isometrie* oder *Bewegung*, wenn $\forall v, w \in V(\|f(v) - f(w)\| = \|v - w\|)$.

Es gilt: f ist orthogonaler Endomorphismus $\iff$ f ist Isometrie mit $f(0) = 0$.

Beweis:
"$\Rightarrow$": $\|f(v) - f(w)\| = \|f(v - w)\| = \|v - w\|$.
"$\Leftarrow$":

(1) $\|f(v)\| = \|f(v) - f(0)\| = \|v - 0\| = \|v\|$.

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

Beweis: $\|f(v)\|^2 + \|f(w)\|^2 - 2\langle f(v), f(w)\rangle = \|f(v) - f(w)\|^2 = \|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\langle v, w\rangle \overset{(1)}{\Rightarrow}$ $\langle f(v), f(w)\rangle = \langle v, w\rangle$.

(3) $f(v + w) = f(v) + f(w)$.

Beweis: $\|f(v + w) - f(v) - f(w))\|^2 =$
$\|f(v + w)\|^2 + \|f(v)\|^2 + \|f(w)\|^2 - 2\langle f(v + w), f(v)\rangle - 2\langle f(v + w), f(w)\rangle + 2\langle f(v), f(w)\rangle \overset{(1),(2)}{=}$
$\|v + w\|^2 + \|v\|^2 + \|w\|^2 - 2\langle v + w, v\rangle - 2\langle v + w, w\rangle + 2\langle v, w\rangle = \|(v + w) - v - w\|^2 = 0$.

(4) $f(\lambda v) = \lambda f(v)$. [Beweis wie für (3)]

§15 Die Jordansche Normalform

Im folgenden sei V stets ein endlichdimensionaler K-Vektorraum.

Definition.
Sei U ein Untervektorraum von V und $f \in \mathsf{End}(V)$. U heißt *f-invariant*, falls $f(U) \subseteq U$ gilt.

Lemma 15.1.
Ist $f \in \mathsf{End}(V)$ und U ein f-invarianter Unterraum von V, so ist $P_{f|U}$ ein Teiler von P_f.

Beweis:
Sei $\tilde{v} = (v_1, ..., v_n)$ Basis von V und $1 \leq k \leq n$, so daß $\tilde{u} := (v_1, ..., v_k)$ Basis von U. Sei ferner $m := n - k$.
Dann gilt: $\mathcal{M}_{\tilde{v}}(f) = \begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix}$ mit $A = \mathcal{M}_{\tilde{u}}(f|U)$, $P_{f|U} = \det(A - t{\cdot}E_k)$, $P_f = \det(\begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix} - t{\cdot}E_n) =$
$\det\begin{pmatrix} A - t{\cdot}E_k & * \\ \mathbf{0} & B - t{\cdot}E_m \end{pmatrix} = \det(A - t{\cdot}E_k) \cdot \det(B - t{\cdot}E_m) = P_{f|U} \cdot \det(B - t{\cdot}E_m)$.

Definition.
$f \in \mathsf{End}(V)$ heißt *trigonalisierbar*, wenn es eine Basis $\tilde{v}$ von V gibt, so daß $\mathcal{M}_{\tilde{v}}(f)$ obere Dreiecksmatrix ist.

Eine Matrix $A \in K^{n\times n}$ heißt trigonalisierbar, wenn $\mathsf{f}_A : K^n \to K^n$ trigonalisierbar ist.

Lemma 15.2.
Sei $f \in \mathsf{End}(V)$, $\tilde{v} = (v_1, ..., v_n)$ eine Basis von V, und $V_j := \operatorname{span}(v_1, ..., v_j)$ für $j = 1, ..., n$.

Dann gilt: $f(V_j) \subseteq V_j$ für $j = 1, ..., n \;\Leftrightarrow\; \mathcal{M}_{\tilde{v}}(f)$ ist eine obere Dreiecksmatrix.

Beweis :
Sei $\mathcal{M}_{\tilde{v}}(f) = (a_{ij})_{i,j}$. Dann gilt: $f(V_j) \subseteq V_j \; (j = 1, ..., n) \;\Leftrightarrow$
$\Leftrightarrow\; f(v_1), ..., f(v_j) \in \operatorname{span}(v_1, ..., v_j) \; (j = 1, ..., n) \;\Leftrightarrow\; f(v_j) \in \operatorname{span}(v_1, ..., v_j) \; (j = 1, ..., n) \;\Leftrightarrow$
$\Leftrightarrow\; \sum_{i=1}^n a_{ij} v_i \in \operatorname{span}(v_1, ..., v_j) \; (j = 1, ..., n) \;\Leftrightarrow\; a_{j+1,j} = ... = a_{nj} = 0 \; (j = 1, ..., n)$.

Satz 15.3 (Trigonalisierungssatz).
Gilt $\dim(V) = n$, $f \in \mathsf{End}(V)$ und $\lambda_1, ..., \lambda_n \in K$, so sind äquivalent:

(i) V besitzt eine Basis $\tilde{v}$, so daß $\mathcal{M}_{\tilde{v}}(f)$ eine obere Dreiecksmatrix mit $\lambda_1, ..., \lambda_n$ in der Diagonalen ist.
(ii) $P_f = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t)$.

28.6.2010

Beweis :
(i)$\Rightarrow$(ii): Sei $A := \mathcal{M}_{\tilde{v}}(f)$. Dann $P_f = \det(A - t{\cdot}E) = (\lambda_1 - t) \cdot \ldots \cdot (\lambda_n - t)$.

(ii)$\Rightarrow$(i): Induktion nach n: Sei $n \geq 2$ und v_1 ein Eigenvektor zum Eigenwert λ_1.
Wir ergänzen (v_1) zu einer Basis $\tilde{u} := (v_1, v_2', ..., v_n')$ von V. Dann ist $\mathcal{M}_{\tilde{u}}(f) = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}$.

Wir definieren $h \in \operatorname{Hom}(W, Kv_1)$, $g \in \mathsf{End}(W)$ durch $h(v_j') := a_{1j} v_1$ und $g(v_j') := \sum_{i=2}^n a_{ij} v_i'$ $(j = 2, ..., n)$.

Dann ist $P_f = (\lambda_1 - t) \cdot P_g$, also $P_g = (\lambda_2 - t) \cdot \ldots \cdot (\lambda_n - t)$. Nach I.V. besitzt W deshalb eine Basis $\tilde{w} = (v_2, ..., v_n)$, so daß $\mathcal{M}_{\tilde{w}}(g)$ eine obere Dreiecksmatrix mit Diagonalelementen $\lambda_2, ..., \lambda_n$ ist.
Sei $\tilde{v} := (v_1, ..., v_n)$. Wegen $\forall w \in W(f(w) = h(w) + g(w))$ ist dann $\mathcal{M}_{\tilde{v}}(f) = \begin{pmatrix} \lambda_1 & * \\ 0 & \mathcal{M}_{\tilde{w}}(g) \end{pmatrix}$.

Korollar.

$f \in \mathsf{End}(V)$ ist genau dann trigonalisierbar, wenn P_f in Linearfaktoren zerfällt.

Beispiel.

$V = \mathbb{R}^3$ und $f := \mathsf{f}_A$ mit $A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}$. Dann ist $P_f = (2-t)^3$, also $\lambda_1 = \lambda_2 = \lambda_3 = 2$, und für $\tilde{v} := (e_1 - e_2 + e_3\,,\, e_2 - e_3\,,\, e_3)$ gilt $\mathcal{M}_{\tilde{v}}(f) = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$.

Lemma 15.4.

Sei $f \in \mathsf{End}(V)$ und $\tilde{v} = (v_1, \ldots, v_n)$ eine Basis von V, so daß

$\mathcal{M}_{\tilde{v}}(f)$ eine obere Dreiecksmatrix mit Diagonale $\lambda_1, \ldots, \lambda_n$ ist.

Für $g_i := f - \lambda_i \cdot \mathrm{id}_V$ $(i = 1, \ldots, n)$ gilt dann:

(a) $g_i(V_i) \subseteq V_{i-1}$, wobei $V_i := \mathrm{span}(v_1, \ldots, v_i)$ $\quad (i = 1, \ldots, n)$.

(b) $g_1 \circ \cdots \circ g_n = 0$.

(c) $1 \le k \le n$ & $\lambda_1 = \ldots = \lambda_k \implies v_1, \ldots, v_k \in \mathrm{Ker}(g_1^k)$.

Beweis :

Sei $A = (a_{ij})_{i,j} := \mathcal{M}_{\tilde{v}}(f)$.

(a) $g_i(v_j) = \begin{cases} a_{1j}v_1 + \ldots + a_{jj}v_j - \lambda_i v_j \in V_{i-1} & \text{falls } j < i \\ a_{1i}v_1 + \ldots + a_{ii}v_i - \lambda_i v_i \in V_{i-1} & \text{falls } j = i \text{ (denn } a_{ii} = \lambda_i) \end{cases}$.

(b) Mit (a) folgt $(g_1 \circ \ldots \circ g_n)(V_n) \subseteq (g_1 \circ \ldots \circ g_{n-1})(V_{n-1}) \subseteq (g_1 \circ \ldots \circ g_{n-2})(V_{n-2}) \subseteq \ldots \subseteq g_1(V_1) \subseteq V_0 = \{0\}$.

(c) Induktion nach k: 1. $k = 1$: $g_1(v_1) = f(v_1) - \lambda_1 v_1 = 0$.

2. $k > 1$: Nach I.V. haben wir $v_1, \ldots, v_{k-1} \in \mathrm{Ker}(g_1^{k-1}) \subseteq \mathrm{Ker}(g_1^k)$. Somit gilt:

$g_1^k(v_k) = g_1^{k-1}(g_1(v_k)) = g_1^{k-1}(a_{1k}v_1 + \ldots + a_{kk}v_k - \lambda_1 v_k) = g_1^{k-1}(a_{1k}v_1 + \ldots + a_{k-1,k}v_{k-1}) \overset{\text{I.V.}}{=} 0$.

Definition.

Sei K ein Körper. Eine Menge R zusammen mit Verknüpfungen

$+ : R \times R \to R$, $\otimes : R \times R \to R$, $\odot : K \times R \to R$ heißt K-*Algebra*, wenn gilt:

(i) $(R, +, \otimes)$ ist ein Ring mit Eins,

(ii) $(R, +, \odot)$ ist ein K-Vektorraum,

(iii) Für alle $a, b \in R$, $\lambda \in K$ gilt: $\lambda \odot (a \otimes b) = (\lambda \odot a) \otimes b = a \otimes (\lambda \odot b)$.

Bemerkung.

1. Der Polynomring $K[t]$ zusammen mit seiner auf $K \times K[t]$ eingeschränkten Multiplikation als skalarer Multiplikation ist eine K-Algebra.
2. Für jeden K-Vektorraum V ist der in §5 eingeführte K-Vektorraum $\mathsf{End}(V)$ zusammen mit $\circ$ (der Komposition von Abbildungen) als Ringmultiplikation eine K-Algebra. Das Einselement ist dabei id_V.
3. Für jeden Körper K ist der K-Vektorraum $K^{n \times n}$ zusammen mit der üblichen Multiplikaition von Matrizen eine K-Algebra. Insbesondere ist K selbst eine K-Algebra.

Definition.

Ist $(R, +, \otimes, \odot)$ eine K-Algebra, so definiert man für jedes $\mathbf{r} \in R$ den *Einsetzungshomomorphismus*

$\Phi_{\mathbf{r}} : K[t] \to R$, $\Phi_{\mathbf{r}}(\sum_{i=0}^n a_i t^i) := \sum_{i=0}^n a_i \odot \mathbf{r}^i$, wobei $\mathbf{r}^0 := \mathbf{1}_R$, $\mathbf{r}^{i+1} := \mathbf{r}^i \otimes \mathbf{r}$.

Schreibweise: Für $p \in K[t]$ und $\mathbf{r} \in R$ sei $p(\mathbf{r}) := \Phi_{\mathbf{r}}(p)$.

Lemma 15.5.

$\Phi_{\mathbf{r}}$ ist sowohl Ring- als auch Vektorraumhomomorphimus, d.h. für alle $p, q \in K[t]$ und $\lambda \in K$ gilt:
$(p+q)(\mathbf{r}) = p(\mathbf{r}) + q(\mathbf{r})$, $\quad (p \cdot q)(\mathbf{r}) = p(\mathbf{r}) \otimes q(\mathbf{r})$, $\quad (\lambda \cdot p)(\mathbf{r}) = \lambda \odot p(\mathbf{r})$.

Zum Beweis siehe Lemma 10.8.

Von besonderer Bedeutung sind die Einsetzungshomomorphismen
$\Phi_\lambda : K[t] \to K$, $p \mapsto p(\lambda)$ für $\lambda \in K$ (siehe §10), sowie $\Phi_f : K[t] \to \mathsf{End}(V)$, $p \mapsto p(f)$ für $f \in \mathsf{End}(V)$.

Lemma 15.6.

Für $f \in \mathsf{End}(V)$ und $p, q \in K[t]$ gilt:
(a) $p(f) \circ q(f) = q(f) \circ p(f)$.
(b) Ist U ein f-invarianter Unterraum von V, so $p(f)(x) = p(f|U)(x)$ für alle $x \in U$.
(c) $\mathrm{Ker}(p(f))$ ist f-invariant.

Beweis :
(a) $p(f) \circ q(f) \overset{15.5}{=} (pq)(f) = (qp)(f) \overset{15.5}{=} q(f) \circ p(f)$.

30.6.2010

(b) Sei $g := f|U$. $x \in U \;\Rightarrow\; p(f)(x) = \sum_i a_i \cdot f^i(x) = \sum_i a_i \cdot g^i(x) = (\sum_i a_i \cdot g^i)(x) = p(g)(x)$.

(c) $x \in \mathrm{Ker}(p(f)) \;\Rightarrow\; p(f)(x) = 0 \;\Rightarrow\; p(f)(f(x)) = (p(f) \circ f)(x) \overset{(a)}{=} (f \circ p(f))(x) = f(p(f)(x)) = 0$.

Lemma 15.7.

$f \in \mathsf{End}(V)$ & $p, q \in K[t]$ teilerfremd & $p(f) \circ q(f) = 0 \implies V = \mathrm{Ker}(p(f)) \oplus \mathrm{Ker}(q(f))$.

Beweis :
Nach Satz 10.3a gibt es $p', q' \in K[t]$ mit $1 = p'p + q'q$, also (nach 15.5) $\mathrm{id}_V = p'(f) \circ p(f) + q'(f) \circ q(f)$.

1. $V = \mathrm{Ker}(p(f)) + \mathrm{Ker}(q(f))$:
Sei $x \in V$. Dann $x = \mathrm{id}_V(x) = u + v$ mit $u := (q'(f) \circ q(f))(x)$ und $v := (p'(f) \circ p(f))(x)$.
Es ist $p(f)(u) = (p(f) \circ q'(f) \circ q(f))(x) \overset{15.6a}{=} (p(f) \circ q(f) \circ q'(f))(x) = (0 \circ q'(f))(x) = 0$ und ebenso $q(f)(v) = 0$.
Also $x = u + v \in \mathrm{Ker}(p(f)) + \mathrm{Ker}(q(f))$.

2. Die Summe ist direkt:
$y \in \mathrm{Ker}(p(f)) \cap \mathrm{Ker}(q(f)) \;\Rightarrow\; y = \mathrm{id}_V(y) = (p'(f) \circ p(f))(y) + (q'(f) \circ q(f))(y) = p'(f)(0) + q'(f)(0) = 0$.

Satz 15.8.

Sei $f \in \mathsf{End}(V)$. Sind $p_1, \ldots, p_k \in K[t]$ paarweise teilerfremd, und ist $p_1(f) \circ \ldots \circ p_k(f) = 0$,
so ist $V = \mathrm{Ker}(p_1(f)) \oplus \ldots \oplus \mathrm{Ker}(p_k(f))$.

Beweis durch Induktion nach k:
Sei $p := p_1$ und $q := p_2 \cdot \ldots \cdot p_k$. Nach 10.3b sind p, q teilerfremd. Mit L.15.7 folgt $V = \mathrm{Ker}(p_1(f)) \oplus U$, wobei $U := \mathrm{Ker}(q(f))$ f-invariant ist (nach L.15.6c). Sei $g := f|U$. Dann $p_2(g) \circ \ldots \circ p_k(g) = q(g) \overset{\text{L.15.6b}}{=}$
$= q(f)|U = 0$. Nach I.V. gilt also $U = \mathrm{Ker}(p_2(g)) \oplus \ldots \oplus \mathrm{Ker}(p_k(g))$.
Es folgt $V = \mathrm{Ker}(p_1(f)) \oplus \mathrm{Ker}(p_2(g)) \oplus \ldots \oplus \mathrm{Ker}(p_k(g))$. Somit muß nur noch $\mathrm{Ker}(p_i(g)) = \mathrm{Ker}(p_i(f))$ für $i = 2, \ldots, k$ gezeigt werden: $\mathrm{Ker}(p_i(f)) \subseteq \mathrm{Ker}(q(f)) = U \;\Rightarrow$
$\Rightarrow\; \mathrm{Ker}(p_i(f)) = \{x \in U : p_i(f)(x) = 0\} \overset{\text{L.15.6b}}{=} \{x \in U : p_i(g)(x) = 0\} = \mathrm{Ker}(p_i(g))$.

Satz 15.9 (Zerlegung in Haupträume).

Sei $f \in \mathsf{End}(V)$ und $P_f = (\lambda_1 - t)^{\mu_1} \cdot \ldots \cdot (\lambda_k - t)^{\mu_k}$ mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_k \in K$.

Für $g_i := f - \lambda_i \cdot \mathrm{id}_V$ gilt dann:

(a) $V = \mathrm{Ker}(g_1^{\mu_1}) \oplus \ldots \oplus \mathrm{Ker}(g_k^{\mu_k})$

(b) $\dim \mathrm{Ker}(g_i^{\mu_i}) = \mu_i$ für $i = 1, \ldots, k$.

Beweis:

(1) Für $p_i := (t - \lambda_i)^{\mu_i}$ gilt $p_i(f) = g_i^{\mu_i}$.

Nach 15.3 besitzt V eine Basis $\tilde{v} = (v_1, \ldots, v_n)$, so daß gilt:

(2) $\mathcal{M}_{\tilde{v}}(f)$ ist obere Dreiecksmatrix mit Diagonale $\lambda_1, \ldots, \lambda_1, ..., \lambda_k, ..., \lambda_k$.

Mit 15.4b folgt daraus $g_1^{\mu_1} \circ \ldots \circ g_k^{\mu_k} = 0$ und weiter mit 15.8 und (1) die Behauptung (a).

Wir haben nun $\dim \mathrm{Ker}(g_1^{\mu_1}) + \ldots + \dim \mathrm{Ker}(g_k^{\mu_k}) \overset{(a)}{=} \dim(V) = \deg(P_f) = \mu_1 + \ldots + \mu_k$.

Zum Beweis von (b) genügt es also, $\mu_i \leq \dim \mathrm{Ker}(g_i^{\mu_i})$ für $i = 1, \ldots, k$ zu zeigen.

Wegen der Kommutativität von $K[t]$ reicht es, denn Fall $i = 1$ zu behandeln:

Nach (1) und 15.4c gilt $v_1, \ldots, v_{\mu_1} \in \mathrm{Ker}(g_1^{\mu_1})$. Also ist $\mu_1 \leq \dim \mathrm{Ker}(g_1^{\mu_1})$.

Definition.

Ist λ Eigenwert von $f \in \mathsf{End}(V)$ mit algebraischer Vielfachheit μ, so nennt man $\mathrm{Hau}(f; \lambda) := \mathrm{Ker}(f - \lambda\, \mathrm{id}_V)^{\mu}$ den *Hauptraum* von f zum Eigenwert λ.

Nachtrag: Für $A \in K^{n \times n}$ sei $\mathrm{Hau}(A; \lambda) := \mathrm{Hau}(\mathsf{f}_A, \lambda)$.

Bemerkung.

(a) $\mathrm{Eig}(f; \lambda) = \mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V) \subseteq \mathrm{Hau}(f; \lambda)$. (b) $\mathrm{Hau}(f; \lambda)$ ist f-invariant.

Beweis von (b): Sei $g := f - \lambda\, \mathrm{id}_V$ und $U := \mathrm{Hau}(f; \lambda)$.

$x \in U \Rightarrow \lambda x \in U \;\&\; g^{\mu}(g(x)) = g(g^{\mu}(x)) = 0 \Rightarrow \lambda x, g(x) \in U \Rightarrow f(x) = g(x) + \lambda x \in U$.

5.7.2010

Definition.

Ein Endomorphismus f heißt *nilpotent*, wenn es ein $k \in \mathbb{N}$ mit $f^k = 0$ gibt.

Lemma 15.10.

Sei $f \in \mathsf{End}(V)$, $\dim(V) = n \geq 1$ und $k \geq 1$ mit $f^k = 0$, $f^{k-1} \neq 0$.

Für $U_i := \mathrm{Ker}(f^i)$ gilt dann: (a) $U_i \subseteq U_{i+1} = \overset{-1}{f}(U_i)$. (b) $i < k \;\Rightarrow\; U_i \neq U_{i+1}$.

Beweis:

(a) $f^i(x) = 0 \;\Rightarrow\; f^{i+1}(x) = f(f^i(x)) = 0$. $x \in \overset{-1}{f}(U_i) \Leftrightarrow f(x) \in U_i \Leftrightarrow f^i(f(x)) = 0 \Leftrightarrow x \in U_{i+1}$.

(b) Aus (a) folgt durch Induktion nach l: $U_i = U_{i+1} \;\&\; i \leq l \;\Rightarrow\; U_l = U_{l+1}$.

Mit $f^{k-1} \neq 0 = f^k$ (also $U_{k-1} \neq U_k$) folgt daraus die Behauptung.

Satz 15.11.

Sei $f \in \mathsf{End}(V)$ und $\dim(V) = n \geq 1$. Äquivalent sind:

(i) f ist nilpotent.

(ii) $f^k = 0$ für ein $k \in \{1, \ldots, n\}$.

(iii) $P_f = (-t)^n$.

(iv) Es gibt eine Basis $\tilde{v}$ von V, so daß $\mathcal{M}_{\tilde{v}}(f)$ obere Dreiecksmatrix mit lauter Nullen in der Diagonale.

Beweis:

"(i)⇒(ii),(iv)": Sei k minimal mit $f^k = 0$. Wegen $f^0 = \mathrm{id}_V \neq 0$ ist $k \geq 1$. Sei $U_i = \mathrm{Ker}(f^i)$.

Nach L.15.10 haben wir $\{0\} = U_0 \subsetneqq U_1 \subsetneqq \ldots \subsetneqq U_k = V$, also $k \leq \dim(V) = n$.

Wir wählen eine Basis $\tilde{v} = (v_1, \ldots, v_n)$ von V, so daß gilt:

Es gibt $0 = n_0 < n_1 < \ldots < n_k = n$ mit $U_l = \mathrm{span}(v_1, \ldots, v_{n_l})$ für $l = 1, \ldots, k$.

Behauptung: $\mathcal{M}_{\tilde{v}}(f)$ ist von der gewünschten Form.

Beweis: Sei $1 \leq j \leq n$. Dann existiert ein $l \in \{0, \ldots, k-1\}$, so daß $n_l < j \leq n_{l+1}$. Es folgt $v_j \in U_{l+1}$ und weiter $f(v_j) \in U_l \subseteq \mathrm{span}(v_1, \ldots, v_{j-1})$. Also $f(v_j) = \sum_{i=1}^n a_{ij} v_i$ mit $a_{ij} = 0$ für $i = j, \ldots, n$.

"(iv)⇔(iii)": Satz 15.3. "(ii)⇒(i)": trivial.

"(iv)⇒(ii)": Aus 15.4b folgt $f^n = 0$. Man kann auch direkt $\mathcal{M}_{\tilde{v}}(f)^n = 0$ zeigen.

!!!!!!! Bis hier Stoff der Klausur am 19.7.2010 !!!!!!!

Definition. Sei $g \in \mathsf{End}(V)$.

Ein m-Tupel $\tilde{u}$ von Vektoren aus V heißt *g-zyklisch*, wenn es ein $u \in V$ gibt, so daß

$\tilde{u} = (g^{m-1}(u), g^{m-2}(u), \ldots, g(u), u)$ und $m = \min\{i : g^i(u) = 0\}$.

Ein Unterrraum U von V heißt *g-zyklisch*, wenn es eine g-zyklische Basis von U gibt.

Abkürzung¿: $$J_m := \begin{pmatrix} 0 & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & 0 \end{pmatrix} = (\delta_{i+1,j})_{i,j} \in K^{m \times m}.$$

Bemerkung.

Ist $\tilde{u} = (u_1, \ldots, u_m)$ eine Basis des Unterraums U, so gilt:

$\tilde{u}$ ist g-zyklisch $\iff$ U ist g-invariant und $\mathcal{M}_{\tilde{u}}(g|U) = J_m$.

Satz 15.12.

Sei g ein nilpotenter Endomorphismus von V und $k := \min\{i : g^i = 0\}$.

Dann gibt es $s_1, \ldots, s_k \in \mathbb{N}$ und g-zyklische Unterräume $U_1^{(1)}, \ldots, U_{s_1}^{(1)}, \ldots, U_1^{(k)}, \ldots, U_{s_k}^{(k)}$, so daß

$V = U_1^{(1)} \oplus \ldots \oplus U_{s_1}^{(1)} \oplus \ldots\ldots \oplus U_1^{(k)} \oplus \ldots \oplus U_{s_k}^{(k)}$ und $\dim(U_j^{(i)}) = i$.

Wählt man in jedem $U_j^{(i)}$ eine g-zyklische Basis $\tilde{u}_j^{(i)}$ und setzt dann $\tilde{v} := \tilde{u}_1^{(1)} * \ldots * \tilde{u}_{s_1}^{(1)} * \ldots\ldots * \tilde{u}_1^{(k)} * \ldots * \tilde{u}_{s_k}^{(k)}$, so ist $\tilde{v}$ eine Basis von V und es gilt

$$\mathcal{M}_{\tilde{v}}(g) = \begin{pmatrix} J_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & J_1 & & & & & & \\ & & & J_2 & & & \mathbf{0} & & \\ & & & & \ddots & & & & \\ & & & & & J_2 & & & \\ & & & \mathbf{0} & & & \ddots & & \\ & & & & & & & J_k & \\ & & & & & & & & \ddots \\ & & & & & & & & & J_k \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s_1\text{-mal} \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s_2\text{-mal} \\ \vdots \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} s_k\text{-mal} \end{matrix}$$

Beweis:

Sei $U_i := \mathrm{Ker}(g^i)$. Dann $\{0\} = U_0 \subsetneqq U_1 \subsetneqq \ldots \subsetneqq U_{k-1} \subsetneqq U_k = V$.

Wir definieren rekursiv Unterräume $W_1, \ldots, W_k$ (beginnend mit W_k):

1. W_k sei ein Unterraum von V $(= U_k)$ mit $V = U_{k-1} \oplus W_k$.

2. Sei jetzt $1 < l \leq k$ und $W_l, \ldots, W_k$ schon definiert, so daß $V = U_{l-1} \oplus W_l \oplus \ldots \oplus W_k$ und $U_l = U_{l-1} \oplus W_l$.

Dann $g(W_l) \subseteq g(U_l) \subseteq U_{l-1}$ und $g(W_l) \cap U_{l-2} \overset{(*)}{=} \{0\}$.

$[(*)\ x \in W_l\ \&\ g(x) \in U_{l-2} \Rightarrow x \in W_l \cap \overset{-1}{g}(U_{l-2}) = W_l \cap U_{l-1} = \{0\}]$

Wegen $g(W_l), U_{l-2} \subseteq U_{l-1}$ und $g(W_l) \cap U_{l-2} = \{0\}$ existiert ein Unterraum W_{l-1} von U_{l-1} mit $U_{l-1} = U_{l-2} \oplus W_{l-1}$ und $g(W_l) \subseteq W_{l-1}$.

Schließlich setzen wir noch $W_0 := \{0\}$. Nach Konstruktion gilt dann

(1) $V = W_1 \oplus \ldots \oplus W_k$,

(2) $g(W_l) \subseteq W_{l-1}$ für $l = 1, \ldots, k$, und $g|W_l$ injektiv für $l = 2, \ldots, k$.
[Die Injektivität von $g|W_l$ folgt aus $W_l \cap \mathrm{Ker}(g) \subseteq W_l \cap U_{l-1} = \{0\}$.]

Aus (2) folgt weiter:

(3) Ist $2 \leq l \leq k$ und $\tilde{w} = (w_1, \ldots, w_\nu)$ eine Basis von W_l, so ist $g(\tilde{w}) = (g(w_1), \ldots, g(w_\nu))$ linear unabhängig in W_{l-1} und kann zu einer Basis von W_{l-1} ergänzt werden.

Durch Iteration von (3) kommt man wie folgt zu einer Basis von $V = W_1 \oplus \ldots \oplus W_k$.

$$\begin{array}{rrrrl}
 & & & w_1^{(k)}, \ldots\ldots\ldots, w_{s_k}^{(k)}, & \text{Basis von } W_k \\
 & w_1^{(k-1)}, \ldots\ldots\ldots, w_{s_{k-1}}^{(k-1)}, & & g(w_1^{(k)}), \ldots\ldots, g(w_{s_k}^{(k)}), & \text{Basis von } W_{k-1} \\
 & \vdots \qquad\qquad \vdots & & \vdots \qquad\qquad \vdots & \\
w_1^{(1)}, \ldots, w_{s_1}^{(1)}, \ldots\ldots\ldots, & g^{k-2}(w_1^{(k-1)}), \ldots, g^{k-2}(w_{s_{k-1}}^{(k-1)}), & & g^{k-1}(w_1^{(k)}), \ldots, g^{k-1}(w_{s_k}^{(k)}), & \text{Basis von } W_1
\end{array}$$

Es sei $U_j^{(i)} = \mathrm{span}(g^{i-1}(w_j^{(i)}), \ldots, g(w_j^{(i)}), w_j^{(i)})$.

Wegen $w_j^{(i)} \in W_i \subseteq U_i = \mathrm{Ker}(g^i)$ gilt $g(g^{i-1}(w_j^{(i)})) = 0$, also $U_j^{(i)}$ g-zyklisch.

7.7.2010

Bemerkung. Wegen $U_1 = U_0 \oplus W_1$ und $U_0 = \{0\}$ ist stets $W_1 = U_1 = \mathrm{Ker}(g)$.

Beispiele.

1. $g = \mathsf{f}_A$ mit $A = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, $A^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $A^3 = 0$.

$k = 3$, $U_0 = \{0\}$, $U_1 = \mathrm{span}(e_1)$, $U_2 = \mathrm{span}(e_1, e_2)$, $U_3 = \mathbb{R}^3$.

$U_3 = U_2 \oplus \mathrm{span}(e_3) \rightsquigarrow W_3 := \mathrm{span}(e_3)$. $\qquad g(e_3) = 3e_1 + e_2$

$U_2 = U_1 \oplus \mathrm{span}(g(e_3)) \rightsquigarrow W_2 := \mathrm{span}(g(e_3))$. $\qquad g^2(e_3) = e_1$

$U_1 = U_0 \oplus \mathrm{span}(g^2(e_3)) \rightsquigarrow W_1 := \mathrm{span}(g^2(e_3))$.

Somit $\tilde{v} = (g^2(e_3), g(e_3), e_3) = (2e_1, 3e_1 + 2e_2, e_3)$ und $\mathcal{M}_{\tilde{v}}(g) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

2. $g = \mathsf{f}_A$ mit $A = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $A^2 = 0$.

$k = 2,\ U_0 = \{0\},\ U_1 = \operatorname{span}(e_1, e_2 + e_3),\ U_2 = \mathbb{R}^4$.

$U_2 = U_1 \oplus \operatorname{span}(e_3, e_4) \rightsquigarrow W_2 := \operatorname{span}(e_3, e_4)$. $\quad g(e_3) = -(e_1 + e_2 + e_3),\ g(e_4) = e_2 + e_3$

$U_1 = U_0 \oplus \operatorname{span}(g(e_3), g(e_4)) \rightsquigarrow W_1 := \operatorname{span}(g(e_3), g(e_4))$.

Somit $\tilde{v} = (g(e_3), e_3, g(e_4), e_4)$ und $\mathcal{M}_{\tilde{v}}(g) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Definition.

Unter einem *Jordankästchen* versteht man eine Matrix der Form $\lambda \cdot E_m + J_m = \begin{pmatrix} \lambda & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \mathbf{0} & & & \lambda \end{pmatrix}$.

Eine quadratische Matrix A heißt *Jordanmatrix* (oder *in Jordan Normalform*), wenn sie die Gestalt

$A = \begin{pmatrix} A_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & A_\ell \end{pmatrix}$ hat, wobei $A_1, \ldots, A_\ell$ Jordankästchen sind.

Bemerkung.

Ist $g \in \mathsf{End}(V)$ nilpotent und $f = g + \lambda \cdot \mathrm{id}_V$, so ist die darstellende Matrix von f bzgl. der in 15.12 für g konstruierten Basis $\tilde{v}$ eine Jordanmatrix.

Satz 15.14 (Jordansche Normalform)

Zerfällt das charakteristische Polynom von $f \in \mathsf{End}(V)$ in Linearfaktoren, so gibt es eine Basis $\tilde{v}$ von V derart, daß $\mathcal{M}_{\tilde{v}}(f)$ eine Jordanmatrix ist.

Beweis:

Sei $P_f = (\lambda_1 - t)^{\mu_1} \cdot \ldots \cdot (\lambda_k - t)^{\mu_k}$ mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_k$.

Sei $g_i := f - \lambda_i \cdot \mathrm{id}_V$ und $V_i := \mathrm{Ker}(g_i^{\mu_i}) = \mathrm{Hau}(f; \lambda_i)$. Nach 15.9 gilt dann $V = V_1 \oplus \ldots \oplus V_k$.

Für $i = 1, \ldots, k$ gilt:

(1) V_i ist g_i-invariant und $g_i|V_i$ ist nilpotent. $[(g_i|V_i)^{\mu_i} = g_i^{\mu_i}|V_i = 0]$

(2) V_i ist f-invariant und $f|V_i = g_i|V_i + \lambda_i \cdot \mathrm{id}_{V_i}$

Nach (2) und obiger Bemerkung besitzt V_i eine Basis $\tilde{v}^{(i)}$, so daß $B_i := \mathcal{M}_{\tilde{v}^{(i)}}(f|V_i)$ eine Jordanmatrix ist.

Dann ist $\tilde{v} := \tilde{v}^{(1)} * \ldots * \tilde{v}^{(k)}$ eine Basis von V und es gilt $\mathcal{M}_{\tilde{v}}(f) = \begin{pmatrix} B_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & B_k \end{pmatrix}$.

Somit ist auch $\mathcal{M}_{\tilde{v}}(f)$ eine Jordanmatrix.

§16 Verschiedenes

Satz 16.1 (Cayley-Hamilton).
Sei V ein n-dimensionaler K-Vektorraum, $f \in \mathsf{End}(V)$ und $P_f \in K[t]$ das charakteristische Polynom von f. Dann gilt $P_f(f) = 0$.

Beweis:
Sei A eine darstellende Matrix von f. Wir zeigen $P_A(A) = 0$. Sei $P_A = \sum_{i=0}^n a_i t^i$.

Sei $B := A - t{\cdot}E \in K[t]^{n\times n}$, sowie $\widetilde{B} \in K[t]^{n\times n}$ die in §9 definierte komplementäre Matrix zu B.
Nach Lemma 9.10c gilt dann $\widetilde{B} \cdot B = \det(B) \cdot E$.

Nach Definition von $\widetilde{B}$ (siehe §9) haben alle Komponenten von $\widetilde{B}$ einen Grad $< n$.

Folglich $\widetilde{B} = t^{n-1}{\cdot}C_{n-1} + \ldots + t{\cdot}C_1 + C_0$ mit $C_i \in K^{n\times n}$, und somit
$\widetilde{B}{\cdot}B = (t^{n-1}{\cdot}C_{n-1} + \ldots + t{\cdot}C_1 + C_0){\cdot}(A - t{\cdot}E) =$
$t^n{\cdot}(-C_{n-1}) + t^{n-1}{\cdot}(C_{n-1}A - C_{n-2}) + \ldots + t{\cdot}(C_1A - C_0) + C_0A.$

Andererseits gilt
$\widetilde{B}{\cdot}B = \det(B){\cdot}E = P_A{\cdot}E = (\sum_{k=1}^n a_k t^k){\cdot}E =$
$t^n{\cdot}(a_n{\cdot}E) \ + \ t^{n-1}{\cdot}(a_{n-1}{\cdot}E) \ + \ \ldots \ + \ t{\cdot}(a_1{\cdot}E) + a_0{\cdot}E.$

Komponentenweiser Koeffizientenvergleich (siehe Lemma 10.5b) ergibt:

$a_nE = -C_{n-1}$
$a_{n-1}{\cdot}E = C_{n-1}A - C_{n-2}$
....
$a_1{\cdot}E = C_1A - C_0$
$a_0{\cdot}E = C_0A$

Durch Multiplikation mit A^n bzw. A^{n-1} bzw.... und anschließender Addition erhalten wir

$$\underbrace{a_nA^n + a_{n-1}A^{n-1} + \ldots + a_0E}_{P_A(A)} = -C_{n-1}A^n + (C_{n-1}A - C_{n-2})A^{n-1} + \ldots + (C_1A - C_0)A + C_0A = 0.$$

Bemerkung: Für zerfallendes P_f folgt der Satz schon aus 15.3-15.5.

Der Sylvestersche Trägheitssatz

Sei V ein n-dim. $\mathbb{K}$-Vektorraum und σ eine symmetrische Bilinearform bzw. Hermitesche Form auf V.

Lemma 16.2.
Zu jeder Basis $\tilde{v}$ von V gibt es eine Basis $\tilde{u}$ von V, so daß
$\mathcal{M}_{\tilde{u}}(\sigma)$ reelle Diagonalmatrix und ähnlich zu $\mathcal{M}_{\tilde{v}}(\sigma)$ ist.

Beweis:
Sei $\tau : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $\tau(x,y) := \sigma(\Phi_{\tilde{v}}(x), \Phi_{\tilde{v}}(y))$. Dann ist τ eine sBF bzw. HF auf $\mathbb{K}^n$. Nach 13.12 existiert eine ONB $\tilde{w}$ von $\mathbb{K}^n$, so daß $\mathcal{M}_{\tilde{w}}(\tau)$ eine zu $\mathcal{M}_{\tilde{e}}(\tau)$ ähnliche reelle Diagonalmatrix ist. Dann ist $\tilde{u} := \Phi_{\tilde{v}}(\tilde{w})$ eine Basis von V mit $\sigma(u_i, u_j) = \sigma(\Phi_{\tilde{v}}(w_i), \Phi_{\tilde{v}}(w_j)) = \tau(w_i, w_j)$, also $\mathcal{M}_{\tilde{u}}(\sigma) = \mathcal{M}_{\tilde{w}}(\tau)$. Ferner $\mathcal{M}_{\tilde{v}}(\sigma) = \mathcal{M}_{\tilde{e}}(\tau)$ und deshalb $\mathcal{M}_{\tilde{u}}(\sigma)$ ähnlich zu $\mathcal{M}_{\tilde{v}}(\sigma)$.

Definitionen.
$V_0^\sigma := \{v \in V : \forall w \in V(\sigma(v,w)=0)\}$ (Ausartungsraum von σ)
$r_0^\sigma := \dim(V_0^\sigma)$
$r_+^\sigma := \max\{\dim(U) : U \in \mathcal{U}_+^\sigma\}$ mit $\mathcal{U}_+^\sigma := \{U : U$ Untervektorraum von V & $\forall v \in U \setminus \{0\}(\sigma(v,v) > 0)\}$
$r_-^\sigma := \max\{\dim(U) : U \in \mathcal{U}_-^\sigma\}$ mit $\mathcal{U}_-^\sigma := \{U : U$ Untervektorraum von V & $\forall v \in U \setminus \{0\}(\sigma(v,v) < 0)\}$

Lemma 16.3.
Für jede Zerlegung $V = V_+ \oplus V_- \oplus V_0^\sigma$ mit $V_+ \in \mathcal{U}_+^\sigma$, $V_- \in \mathcal{U}_-^\sigma$ gilt $\dim(V_+) = r_+^\sigma$ und $\dim(V_-) = r_-^\sigma$.

Beweis von $\dim(V_+) = r_+^\sigma$:
Zu zeigen: $\dim(U) \le \dim(V_+)$ für alle $U \in \mathcal{U}_+^\sigma$. – Sei also $U \in \mathcal{U}_+^\sigma$. Dann gilt $U \cap (V_- \oplus V_0^\sigma) \overset{(*)}{=} \{0\}$ und folglich $\dim(U) \le n - (\dim(V_-) + \dim(V_0^\sigma)) = \dim(V_+)$.

[$(*)$ $U \ni u = v_1 + v_0$ & $v_1 \in V_-$ & $v_0 \in V_0^\sigma \;\Rightarrow\; u \in U$ & $\sigma(u,u) = \sigma(v_1,v_1) \le 0 \;\Rightarrow\; u = 0$]

Satz 16.4.
Sei σ eine sBF bzw. HF auf dem n-dimensionalen $\mathbb{K}$-Vektorraum V.

(a) Für jede darstellende Matrix A von σ gilt:
$r_+^\sigma =$ Anzahl der positiven Eigenwerte von A (gezählt mit Vielfachheiten)
$r_-^\sigma =$ Anzahl der negativen Eigenwerte von A (gezählt mit Vielfachheiten)
$r_0^\sigma =$ Vielfachheit des Eigenwerts 0.

(b) σ besitzt eine darstellende Matrix der Gestalt $Diag(1,\dots,1,-1,\dots,-1,0,\dots,0)$.

Beweis:
(a) Sei $\tilde{v}$ eine Basis von V und $A := \mathcal{M}_{\tilde{v}}(\sigma)$. Nach 16.2 existieren $\lambda_1,\dots,\lambda_n \in \mathbb{R}$ und eine Basis $\tilde{u}$ von V, so daß $\mathcal{M}_{\tilde{u}}(\sigma) = Diag(\lambda_1,\dots,\lambda_n)$ und $\mathcal{M}_{\tilde{u}}(\sigma)$ ähnlich zu A. Dabei können wir o.E.d.A. annehmen, daß $\lambda_1,\dots,\lambda_k > 0$, $\lambda_{k+1},\dots,\lambda_m < 0$ und $\lambda_{m+1},\dots,\lambda_n = 0$, wobei $0 \le k \le m \le n$.
Sei $V_+ := \operatorname{span}(u_1,\dots,u_k)$ und $V_- := \operatorname{span}(u_{k+1},\dots,u_m)$.

Dann gilt:
(1) $\sigma(u_i,u_j)\begin{cases} > 0 & \text{für } 1 \le i = j \le k \\ < 0 & \text{für } k < i = j \le m \\ = 0 & \text{sonst} \end{cases}$
(2) $V_0^\sigma = \operatorname{span}(u_{m+1},\dots,u_n)$ und somit $V = V_+ \oplus V_- \oplus V_0^\sigma$
(3) $V_+ \in \mathcal{U}_+^\sigma$ und $V_- \in \mathcal{U}_-^\sigma$
(4) $r_+^\sigma = k$, $r_-^\sigma = m-k$, $r_0^\sigma = n-m$.

Beweis:
(2) "$\supseteq$": $m < i \le n \overset{(1)}{\Rightarrow} \forall j \in \{1,\dots,n\}(\sigma(u_i,u_j)=0) \Rightarrow \forall x \in V(\sigma(u_i,x)=0) \Rightarrow u_i \in V_0^\sigma$.
"$\subseteq$": $v = \sum\limits_{i=1}^{n} x_i u_i \in V_0^\sigma \Rightarrow \forall j \in \{1,\dots,n\}(\,0 = \sigma(v,u_j) = \sum\limits_{i=1}^{n} x_i \sigma(u_i,u_j) = x_j\lambda_j) \Rightarrow x_1 = \ldots = x_m = 0$.
(3) Für $v = \sum_{i=1}^n x_i u_i$ ist $\sigma(v,v) = \sum_{i=1}^n \sigma(u_i,u_i)x_i^2 = \sum_{i=1}^m \lambda_i x_i^2$.
(4) Aus (2) und (3) folgt mit Lemma 16.3: $\dim(V_+) = r_+^\sigma$ und $\dim(V_-) = r_-^\sigma$.
(b) Sei $\tilde{w} := (w_1,\dots,w_n)$ mit $w_i := \begin{cases} |\lambda_i|^{-\frac{1}{2}} \cdot u_i & \text{für } i = 1,\dots,m \\ u_i & \text{für } i = m+1,\dots,n \end{cases}$.
Dann $\mathcal{M}_{\tilde{w}}(\sigma) = Diag(1,\dots,1,-1,\dots,-1,0,\dots,0)$.

Systeme linearer Differentialgleichungen Siehe: Forster, Analysis 2, §16.