Lineare Algebra SS2018

Dozent: Prof. Dr. Arno Fehm

19. Juli 2018

Inhaltsverzeichnis

Kapitel V

Endomorphismen

In diesem Kapitel seien K ein Körper, $n \in \mathbb{N}$ eine natürliche Zahl, V ein n-dimensionaler K-VR und $f \in \mathrm{End}_K(V)$ ein Endomorphismus.

Das Ziel dieses Kapitels ist, die Geometrie von f besser zu verstehen und Basen zu finden, für die $M_B(f)$ eine besonders einfache oder kanonische Form hat.

1. Eigenwerte

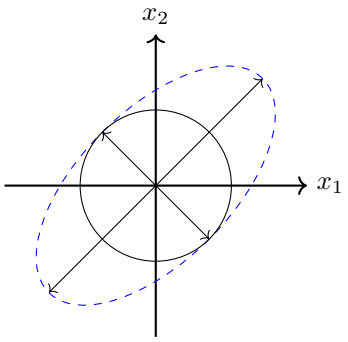
▶ **Bemerkung 1.1**
Wir erinnern uns daran, dass $\mathrm{End}_K(V) = \mathrm{Hom}_K(V,V)$ sowohl einen K-VR als auch einen Ring bildet. Bei der Wahl einer Basis B von V wird $f \in \mathrm{End}_K(V)$ durch die Matrix $M_B(f) = M_B^B(f)$ beschrieben.

■ **Beispiel 1.2**
$K = \mathbb{R}, A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}), f = f_A \in \mathrm{End}_K(K^2)$

$$A \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \; A \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$\Rightarrow$ mit $B = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$ ist $M_B(f) = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$.

Der Endomorphismus $f = f_A$ streckt also entlang der Achse $\mathbb{R} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ um den Faktor 3 und spiegelt entlang der Achse $\mathbb{R} \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}$



Definition 1.3 (Eigenwert, Eigenvektor, Eigenraum)
Sind $0 \neq x \in V$ und $\lambda \in K$ mit $f(x) = \lambda x$ so nennt man λ einen Eigenwert von f und x einen Eigenvektor von f zum Eigenwert λ. Der Eigenraum zu $\lambda \in K$ ist $\mathrm{Eig}(f, \lambda) = \{x \in V \mid f(x) = \lambda x\}$.

▶ **Bemerkung 1.4**
Für jedes $\lambda \in K$ ist $\mathrm{Eig}(f, \lambda)$ ein UVR von V, da

$$\begin{aligned}\mathrm{Eig}(f, \lambda) &= \{x \in V \mid f(x) = \lambda x\} \\ &= \{x \in V \mid f(x) - \lambda \cdot \mathrm{id}_V(x) = 0\} \\ &= \{x \in V \mid (f - \lambda \cdot \mathrm{id}_V)(x) = 0\} \\ &= \mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V)\end{aligned}$$

und $f - \lambda \cdot \mathrm{id}_V \in \mathrm{End}_K(V)$.

▶ **Bemerkung 1.5**
Achtung! Der Nullvektor ist nach Definition kein Eigenvektor, aber $\lambda = 0$ kann ein Eigenwert sein, nämlich genau dann, wenn $f \notin \mathrm{Aut}_K(V)$, siehe Übung. Die Menge der Eigenvektoren zu λ ist also $\mathrm{Eig}(f, \lambda) \backslash \{0\}$ und λ ist genau dann ein Eigenwert von f, wenn $\mathrm{Eig}(f, \lambda) \neq \{0\}$.

■ **Beispiel 1.6**
Ist $A = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ und $f = f_A \in \mathrm{End}_K(K^n)$, so sind $\lambda_1, ..., \lambda_n$ EW von f und jedes e_i ist ein EV zum EW λ_i.

Satz 1.7
Sei B eine Basis von V. Genau dann ist $M_B(f)$ eine Diagonalmatrix, wenn B aus EV von f besteht.

Beweis. Ist $B = (x_1, ...x_n)$ eine Basis aus EV zu EW $\lambda_1, ..., \lambda_n$, so ist $M_B(f) = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ und umgekehrt. □

■ **Beispiel 1.8**
Sei $K = \mathbb{R}$, $V = \mathbb{R}^2$ und $f_\alpha \in \mathrm{End}_K(\mathbb{R}^2)$ die Drehung um den Winkel $\alpha \in [0, 2\pi)$

$$\Rightarrow M_{\mathcal{E}}(f_\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Für $\alpha = 0$ hat $f_\alpha = \mathrm{id}_{\mathbb{R}^2}$ nur den EW 1.
Für $\alpha = \pi$ hat $f_\alpha = -\mathrm{id}_{\mathbb{R}^2}$ nur den EW -1.
Für $\alpha \neq 0, \pi$ hat f_α keine EW.

Lemma 1.9
Sind $\lambda_1, ..., \lambda_n$ paarweise verschiedene EW von f und ist x_i ein EV zu λ_i für $i = 1, ..., m$, so ist $(x_1, ..., x_m)$ linear unabhängig.

Beweis. Induktion nach m
$m = 1$: klar, denn $x_1 \neq 0$

$\underline{m-1 \to m}$: Sei $\sum_{i=1}^{m} \mu_i x_i = 0$ mit $\mu_1, ..., \mu_m \in K$.

$$\begin{aligned} 0 &= (f - \lambda \cdot \mathrm{id}_V) \left(\sum_{i=1}^{m} \mu_i x_i \right) \\ &= \sum_{i=1}^{m} \mu_i (f(x_i) - \lambda_m \cdot x_i) \\ &= \sum_{i=1}^{m-1} \mu_i (\lambda_i - \lambda_m) \cdot x_i \end{aligned}$$

Nach IB ist $\mu_i(\lambda_i - \lambda_m) = 0$ für $i = 1, ..., m-1$, da $\lambda_i \neq \lambda_m$ für $i \neq m$ also $\mu_i = 0$ für $i = 1, ..., m-1$. Damit ist auch $\mu_m = 0$. Folglich ist $(x_1, ..., x_m)$ linear unabhängig. □

Satz 1.10
Sind $\lambda_1, ..., \lambda_m \in K$ paarweise verschieden, so ist

$$\sum_{i=1}^{m} \mathrm{Eig}(f, \lambda_i) = \bigoplus_{i=0}^{m} \mathrm{Eig}(f, \lambda_i).$$

Beweis. Seien $x_i, y_i \in \mathrm{Eig}(f, \lambda_i)$ für $i = 1, ..., m$. Ist $\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} y_i$, so ist $\sum_{i=1}^{m} \underbrace{x_i - y_i}_{z_i} = 0$.

o. E. seien $z_i \neq 0$ für $i = 1, ..., r$ und $z_i = 0$ für $i = r+1, ..., m$. Wäre $r > 0$, so wären $(z_1, ..., z_r)$ linear abhängig, aber $z_i = x_i - y_i \in \mathrm{Eig}(f, \lambda_i) \backslash \{0\}$, im Widerspruch zu Lemma 1.9. Somit ist $x_i = y_i$ für alle i und folglich ist die Summe $\sum \mathrm{Eig}(f, \lambda_i)$ direkt. □

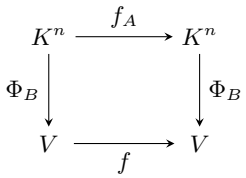
Definition 1.11 (EW und EV für Matrizen)
Sei $A \in \mathrm{Mat}_n(K)$. Man definiert Eigenwerte, Eigenvektoren, etc von A als Eigenwerte, Eigenvektoren von $f_A \in \mathrm{End}_K(K^n)$.

Mathematica/WolframAlpha-Befehle (Eigenwerte und Eigenvektoren)
Um die Eigenwerte und Eigenvektoren einer Matrix A zu berechnen, gibt es in Mathematica bzw. WolframAlpha verschiedene Möglichkeiten:

- `Eigenvalues[A]`: liefert eine Liste der Eigenwerte
- `Eigenvectors[A]`: liefert eine Liste der Eigenvektoren
- `Eigensystem[A]`: liefert zu jeden Eigenwert den Eigenvektor

Satz 1.12
Sei B eine Basis von V und $\lambda \in K$. Genau dann ist λ ein EW von f, wenn λ ein EW von $A = M_B(f)$ ist. Insbesondere haben ähnliche Matrizen die selben EW.

Beweis. Dies folgt aus dem kommutativen Diagramm



denn $f_A(x) = \lambda x \iff (\Phi_B \circ f_A)(x) = \Phi_B(\lambda x) \iff f(\Phi_B(x)) = \lambda \Phi_B(x)$.
Ähnliche Matrizen beschreiben den selben Endomorphismus bezüglich verschiedener Basen, vgl. IV.4.1 □

2. Das charakteristische Polynom

Satz 2.1
Sei $\lambda \in K$. Genau dann ist λ ein EW von f, wenn $\det(\lambda \cdot \mathrm{id}_V - f) = 0$.

Beweis. Da $\mathrm{Eig}(f, \lambda) = \mathrm{Ker}(\lambda \cdot \mathrm{id}_V - f)$ ist λ genau dann ein EW von f, wenn $\dim_K(\mathrm{Ker}(\lambda \cdot \mathrm{id}_V - f)) > 0$, also wenn $\lambda \cdot \mathrm{id}_V - f \notin \mathrm{Aut}_K(V)$. Nach IV.4.6 bedeutet dies, dass $\det(\lambda \cdot \mathrm{id}_V - f) = 0$ □

Definition 2.2 (charakteristisches Polynom)
Das charakteristische Polynom einer Matrix $A \in \mathrm{Mat}_n(K)$ ist die Determinante der Matrix $t \cdot \mathbb{1}_n - A \in \mathrm{Mat}_n(K[t])$.

$$\chi_A(t) = \det(t \cdot \mathbb{1}_n - A) \in K[t]$$

Das charakteristische Polynom eines Endomorphismus $f \in \mathrm{End}_K(V)$ ist $\chi_f(t) = \chi_{M_B(f)}(t)$, wobei B eine Basis von V ist.

Mathematica/WolframAlpha-Befehle (charakteristisches Polynom)
Die folgende Funktion liefert das charakteristische Polynom einer Matrix A mit der Variable x

```
CharacteristicPolynomial[A,x]
```

Satz 2.3
Sind $A, B \in \mathrm{Mat}_n(K)$ mit $A \sim B$, so ist $\chi_A = \chi_B$. Insbesondere ist χ_f wohldefiniert.

Beweis. Ist $B = SAS^{-1}$ mit $S \in \mathrm{GL}_n(K)$, so ist $t \cdot \mathbb{1}_n - B = S(t \cdot \mathbb{1}_n - A)S^{-1}$, also $t \cdot \mathbb{1}_n - B \sim t \cdot \mathbb{1}_n - A$ und ähnliche Matrizen haben die selben Determinante (IV.4.4).
Sind B, B' Basen von V, so sind $M_B(f) \sim M_{B'}(f)$, also $\chi_{M_B(f)} = \chi_{M_{B'}(f)}$ □

Lemma 2.4
Für $\lambda \in K$ ist $\chi_f(\lambda) = \det(\lambda \cdot \mathrm{id}_V - f)$.

Beweis. Sei B eine Basis von V und $A = M_B(f) = (a_{ij})_{i,j}$. Dann ist $M_B(\lambda \cdot \mathrm{id}_V - f) = \lambda \cdot \mathbb{1}_n - A$. Aus IV.2.8 und I.6.8 folgt $\det(t \cdot \mathbb{1}_n - A)(\lambda) = \det(\lambda \cdot \mathbb{1}_n - A)$. Folglich ist

$$\begin{aligned}
\chi_f(\lambda) &= \chi_A(\lambda) \\
&= \det(t \cdot \mathbb{1}_n - A)(\lambda) \\
&= \det(\lambda \cdot \mathbb{1}_n - A) \\
&= \det(\lambda \cdot \mathrm{id}_V - f)
\end{aligned}$$

□

Satz 2.5
Sei $\dim_K(V) = n$ und $f \in \mathrm{End}_K(V)$. Dann ist $\chi_f(t) = \sum_{i=0}^{n} \alpha_i t^i$ ein Polynom vom Grad n mit

$$\begin{aligned} \alpha_n &= 1 \\ \alpha_{n-1} &= -\operatorname{tr}(f) \\ \alpha_0 &= (-1)^n \cdot \det(f) \end{aligned}$$

Die Nullstellen von χ_f sind genau die EW von f.

Beweis. Sei B eine Basis von V und $A = M_B(f) = (a_{ij})_{i,j}$. Wir erinnern uns daran, dass $\operatorname{tr}(f) = \operatorname{tr}(A = \sum_{i=1}^{n} a_{ii}$. Es ist $\chi_f(t) = \det(t - \cdot 1_n - A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} (t\delta_{i,\sigma(i)} - a_{i,\sigma(i)})$.
Der Summand für $\underline{\sigma = \mathrm{id}}$ ist $\prod_{i=1}^{n}(t - a_{ii}) = t^n + \sum_{i=1}^{n}(-a_{ii})t^{n-1} + \ldots + \prod_{i=1}^{n}(-a_{ii})$
Für $\underline{\sigma \neq \mathrm{id}}$ ist $\sigma(i) \neq i$ für mindestens zwei i, der entsprechende Summand hat also Grad höchstens $n-2$. Somit haben α_n und α_{n-1} die oben behauptete Form, und $\alpha_0 = \chi_A(0) = \det(-A) = (-1)^n \cdot \det(f)$.
Die Aussage über die Nullstellen von χ_f folgt aus Satz 2.1 und Lemma 2.4. □

Folgerung 2.6
Ist $\dim_K(V) = n$, so hat f höchstens n Eigenwerte.

Beweis. Satz 2.5 und I.6.10 □

Definition 2.7 (normiertes Polynom)
Ein Polynom $0 \neq P \in K[t]$ mit Leitkoeffizient 1 heißt $\underline{\text{normiert}}$.

Beispiel 2.8

1. Ist $A = (a_{ij})_{i,j}$ eine obere Dreiecksmatrix, so ist $\chi_A(t) = \prod_{i=1}^{n}(t - a_{ii})$, vgl. IV.2.9.c
   Insbesondere ist $\chi_{1_n}(t) = (t-1)^n$, $\chi_0(t) = t^n$
2. Für eine Blockmatrix $A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$ mit quadratischen Matrizen A_1, A_2 ist $\chi_A = \chi_{A_1} \cdot \chi_{A_2}$
   vgl. IV.2.9.e
3. Für

$$\begin{pmatrix} 0 & \ldots & \ldots & \ldots & 0 & -c_0 \\ 1 & \ddots & & & \vdots & \vdots \\ 0 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & 1 & 0 & -c_{n-1} \end{pmatrix} \quad c_0, \ldots, c_{n-1} \in K$$

   ist $\chi_A(t) = t^n + \sum_{i=0}^{n-1} c_i t^i$
   Man nennt diese Matrix die Begleitmatrix zum normierten Polynom $P = t^n + \sum_{i=0}^{n-1} c_i t^i$ und schreibt $M_P := A$

3. Diagonalisierbarkeit

Definition 3.1 (diagonalisierbar)
Man nennt f diagonalisierbar , wenn V eine Basis B besitzt, für die $M_B(f)$ eine Diagonalmatrix ist.

Lemma 3.2
Genau dann ist f diagonalisierbar, wenn

$$V = \sum_{\lambda \in K} \mathrm{Eig}(f, \lambda)$$

.

Beweis. ($\Rightarrow$): Ist B eine Basis aus EV von f (vgl. Satz 1.7), so ist $B \leq \bigcup_{\lambda \in K} \mathrm{Eig}(f, \lambda)$, also $V = \mathrm{span}_K(\bigcup_{\lambda \in K} \mathrm{Eig}(f, \lambda)) = \sum_{\lambda \in K} \mathrm{Eig}(f, \lambda)$.
($\Leftarrow$): Ist $V = \sum_{\lambda \in K} \mathrm{Eig}(f, \lambda)$, so gibt es $\lambda_1, ..., \lambda_n \in K$ mit $V = \sum_{i=1}^{r} \mathrm{Eig}(f, \lambda_i)$. Wir wählen Basen B_i von $\mathrm{Eig}(f, \lambda_i)$. Dann ist $\bigcup_{i=1}^{r} B_i$ ein endliches Erzeugendensystem von V, enthält also eine Basis von V (II.3.6). Diese besteht aus EV von f. □

Satz 3.3
Ist $\dim_K(V) = n$, so hat f höchstens n Eigenwerte. Hat f genau n Eigenwerte, so ist f diagonalisierbar.

Beweis. Ist λ ein EW von f, so ist $\dim_K(\mathrm{Eig}(f, \lambda)) \geq 1$. Sind also $\lambda_1, ..., \lambda_n$ paarweise verschiedene EW von f, so ist

$$\begin{aligned} n = \dim_K(V) &\geq \dim_K \left(\sum_{i=1}^{m} \mathrm{Eig}(f, \lambda_i) \right) \\ &\overset{\text{Satz 1.10}}{=} \dim_K \left(\bigoplus_{i=0}^{m} \mathrm{Eig}(f, \lambda_i) \right) \\ &= \sum_{i=1}^{m} \dim_K(\mathrm{Eig}(f, \lambda_i)) \\ &\geq m \end{aligned}$$

Ist zudem $m = n$, so muss

$$\dim_K(V) = \dim_K(\sum_{i=1}^{m} \mathrm{Eig}(f, \lambda_i)) \text{ sein, also}$$
$$V = \sum_{i=1}^{m} \mathrm{Eig}(f, \lambda_i)$$

Nach Lemma 3.2 ist f genau dann diagonalisierbar. □

Definition 3.4 (a teilt b)
Sei R ein kommutativer Ring mit seien $a, b \in R$. Man sagt, a teilt b (in Zeichen $a|b$), wenn es $x \in R$ mit $b = ax$ gibt.

Definition 3.5 (Vielfachheit)
Für $0 \neq P \in K[t]$ und $\lambda \in K$ nennt man $\mu(P,\lambda) = \max\{r \in \mathbb{N}_{>0} \mid (t-r)^r | P\}$ die Vielfachheit der Nullstelle λ von P.

Lemma 3.6
Genau dann ist $\mu(P,\lambda) \geq 1$, wenn λ eine Nullstelle von P ist.

Beweis. ($\Rightarrow$): $t-\lambda | P \Rightarrow P(t) = (t-\lambda)\cdot Q(t)$ mit $Q(t) \in K[t] \Rightarrow P(\lambda) = 0 \cdot Q(\lambda) = 0$.
($\Leftarrow$): $P(\lambda) = 0 \overset{I.6.9}{=} t-\lambda | P(t) \Rightarrow \mu(P,\lambda) \geq 1$. □

Lemma 3.7
Ist $P(t) = (t-\lambda)^r \cdot Q(t)$ mit $Q(t) \in K[t]$ und $Q(\lambda) \neq 0$, so ist $\mu(P,\lambda) = r$

Beweis. Offensichtlich ist $\mu(P,\lambda) \geq r$. Wäre $\mu(P,\lambda) \geq r+l$, so $(t-\lambda)^{r+l} | P(t)$ also $(t-\lambda)^r \cdot Q(t) = (t-\lambda)^{r+l} \cdot R(t)$ mit $R(t) \in K[t]$, folglich $t-\lambda | Q(t)$, insbesondere $Q(\lambda) = 0$.
(Denn wir dürfen kürzen: R ist nullteilerfrei, genau so wie $K[t]$).
$(t-\lambda)^r(Q(t) - (t-\lambda)R(t)) = 0 \Rightarrow Q(t) = (t-\lambda)R(t)$. □

Lemma 3.8
Sind $P, Q, R \in K[t]$ mit $PQ = PR$, und ist $P \neq 0$, so ist $Q = R$.

Beweis. $PQ = PR \Rightarrow P(Q-R) = 0 \overset{K[t] \text{ nullteilerfrei}}{\Rightarrow} Q - R = 0$, d.h. $Q = R$. □

Lemma 3.9
Es ist $\sum_{\lambda \in K} \mu(P,\lambda) \leq \deg(P)$, mit Gleichheit genau dann, wenn P in Linearfaktoren zerfällt.

Beweis. Schreibe $P(t) = \prod_{\lambda \in K} (t-\lambda)^{r_\lambda} \cdot Q(t)$, wobei $Q(t) \in K[t]$ keine Nullstellen mehr besitzt. Nach Lemma 3.7 ist $\mu(P,\lambda) = r_\lambda$ für alle λ und somit $\deg(P) = \sum_{\lambda \in K} r_\lambda + \deg(Q) \geq \sum_{\lambda \in K} \mu(P,\lambda)$ mit Gleichheit genau dann,wenn $\deg(Q) = 0$, also $Q = c \in K$, d.h. genau dann, wenn $P(t) = c \cdot \prod_{\lambda \in K} (t-\lambda)^{r_\lambda}$. □

Lemma 3.10
Für $\lambda \in K$ ist

$$\dim_K(\mathrm{Eig}(f,\lambda)) \geq \mu(x_f,\lambda)$$

Beweis. Ergänze eine Basis B von $\mathrm{Eig}(f,\lambda)$ zu einer Basis B von V. Dann ist

$$A = M_B(f) = \begin{pmatrix} \lambda \mathbb{1}_s & * \\ 0 & A' \end{pmatrix}$$

mit einer Matrix $A' \in \mathrm{Mat}_{n-s}(K)$, also $\chi_f(t) = \chi_A(t) \overset{\text{Beispiel 2.8}}{=} \chi_{\lambda \mathbb{1}} \cdot \chi_{A'}(t) = (t-\lambda)^s \cdot \chi_{A'}(t)$ und somit $\dim_K(\mathrm{Eig}(f,\lambda)) = s \leq \mu(x_f,\lambda)$. □

Satz 3.11
Genau dann ist f diagonalisierbar, wenn χ_f in Linearfaktoren zerfällt und $\dim_K(\mathrm{Eig}(f,\lambda)) = \mu(x_f,\lambda)$ für alle $\lambda \in K$.

Beweis. Es gilt

$$\begin{aligned}
\dim_K(\sum_{\lambda\in K} \mathrm{Eig}(f,\lambda)) &\overset{\text{Satz 1.10}}{=} \dim_K(\bigoplus_{\lambda\in K} \mathrm{Eig}(f,\lambda)) \\
&\overset{\text{II.4.12}}{=} \sum_{\lambda\in K} \dim_K(\mathrm{Eig}(f,\lambda)) \\
&\overset{\text{Lemma 3.10}}{\leq} \sum_{\lambda\in K} \mu(\chi_f,\lambda) \qquad (1) \\
&\leq \deg(\chi_f) \qquad (2) \\
&= n
\end{aligned}$$

Nach Lemma 3.2 ist f genau dann diagonalisierbar, wenn $\dim_K(\sum_{\lambda\in K} \mathrm{Eig}(f,\lambda)) = n$, also wenn bei (1) und (2) Gleichheit herrscht. Gleichheit bei (1) bedeutet $\dim_K(\mathrm{Eig}(f,\lambda)) = \mu(\chi_f,\lambda)$ für alle $\lambda \in K$, und Gleichheit bei (2) bedeutet nach Lemma 3.9, dass χ_f in Linearfaktoren zerfällt. □

Definition 3.12 (algebraische und geometrische Vielfachheit)
Man nennt $\mu_a(f,\lambda) = \mu(\chi_f,\lambda)$ die algebraische Vielfachheit und $\mu_g(f,\lambda) = \dim_K(\mathrm{Eig}(f,\lambda))$ die geometrische Vielfachheit des Eigenwertes λ von f.

▶ **Bemerkung 3.13**
Wieder nennt man $A \in \mathrm{Mat}_n(K)$ diagonalisierbar, wenn $f_A \in \mathrm{End}_K(K^n)$ diagonalisierbar ist, also wenn $A \sim D$ für eine Diagonalmatrix D.

4. Trigonalisierbarkeit

Definition 4.1
Man nennt f trigonalisierbar , wenn V eine Basis B besitzt, für die $M_B(f)$ eine obere Dreiecksmatrix ist.

■ **Beispiel 4.2**
Ist f diagonalisierbar, so ist f auch trigonalisierbar.

Lemma 4.3
Ist f trigonalisierbar, so zerfällt χ_f in Linearfaktoren.

Beweis. Klar aus Beispiel 2.8 und Satz 2.3. □

Definition 4.4 (invariant)
Ein Untervektorraum $W \leq V$ ist f-invariant , wenn $f(W) \leq W$.

▶ **Bemerkung 4.5**
Ist W ein f-invarianter UVR von V, so ist $f|_W \in \mathrm{End}_K(W)$.

■ **Beispiel 4.6**

1. V hat stets die f-invarianten UVR $W = \{0\}$ und $W = V$.
2. Jeder UVR $W \leq \mathrm{Eig}(f, \lambda)$ ist f-invariant.
3. Ist $B = (x_1, ..., x_n)$ eine Basis von V, für die $M_B(f)$ eine obere Dreiecksmatrix ist, so sind alle UVR $W_i = \mathrm{span}_K(x_1, ..., x_i)$ f-invariant.
4. Sei $V = W \oplus U$, $B_1 = (x_1, ..., x_r)$ Basis von W, $B_2(x_{r+1}, ..., x_n)$ Basis von U und $B = (x_1, ..., x_n)$. Ist W f-invariant, so ist

$$M_B(f) = \begin{pmatrix} M_{B_1}(f|_W) & * \\ 0 & * \end{pmatrix}$$

Sind W und U f-invariant, so ist

$$M_B(f) = \begin{pmatrix} M_{B_1}(f|_W) & 0 \\ 0 & M_{B_2}(f|_U) \end{pmatrix}$$

Lemma 4.7
Ist $W \subset V$ ein f-invarianter UVR, so gilt $\chi_{f|_W} | \chi_f$. Hat W ein lineares Komplement U, dass auch f-invariant ist, so $\chi_f = \chi_{f|_W} \cdot \chi_{f|_U}$.

Beweis. Ergänze eine Basis $B_0 = (x_1, ..., x_r)$ von W zu einer Basis $B = (x_1, ..., x_n)$ von V. Sei $A = M_B(f)$,

$A_0 = M_{B_0}(f|_W)$. Dann ist

$$A = \begin{pmatrix} A_0 & * \\ 0 & C \end{pmatrix} \quad C \in \mathrm{Mat}_{n-r}(K)$$

folglich $\chi_f = \chi_A = \chi_{A_0} \cdot \chi_C$, insbesondere $\chi_{f|_W} | \chi_f$.
Ist auch $U = \mathrm{span}_K(x_{r+1}, ..., x_n)$ f-invariant, so ist

$$A = \begin{pmatrix} A_0 & 0 \\ 0 & C \end{pmatrix}$$

und folglich $\chi_f = \chi_A = \chi_{A_0} \cdot \chi_C = \chi_{f|_W} \cdot \chi_{f|_U}$. □

Theorem 4.8
Genau dann ist f trigonalisierbar, wenn χ_f in Linearfaktoren zerfällt.

Beweis. ($\Rightarrow$): Lemma 4.3
($\Leftarrow$): Induktion nach $n = \dim_K(V)$.
$n = 1$: trivial
$n - 1 \to n$: Nach Annahme ist $\chi_f(t) = \prod_{i=1}^n (t - \lambda_i)$ mit $\lambda_1, ..., \lambda_n \in K$. Sei x_1 ein EV zum EW λ_1. Dann ist $V_1 = K \cdot x_1$ ein f-invarianter UVR. Ergänze $B_1 = (x_1)$ zu einer Basis $B = (x_1, ..., x_n)$ von V und setze $B_2 = (x_2, ..., x_n)$, $V_2 = \mathrm{span}_K(B_2)$. $n - 1 \to n$: Nach Annahme ist $\chi_f(t) = \prod_{i=1}^n (t - \lambda_i)$ mit $\lambda_1, ..., \lambda_n \in K$. Sei x_1 ein EV zum EW λ_1. Dann ist $V_1 = K \cdot x_1$ ein f-invarianter UVR. Ergänze $B_1 = (x_1)$ zu einer Basis $B = (x_1, ..., x_n)$ von V und setze $B_2 = (x_2, ..., x_n)$, $V_2 = \mathrm{span}_K(B_2)$.

$$\Rightarrow M_B(f) = \begin{pmatrix} \lambda_1 & * \\ 0 & A_2 \end{pmatrix} \quad A_2 \in \mathrm{Mat}_{n-1}(K)$$

$$\chi_f(t) = \chi_{\lambda_1 \mathbb{1}_1} \cdot \chi_{A_2} = (t - \lambda_1) \cdot \chi_{A_2}(t)$$

$$\overset{\text{Lemma 3.7}}{\Rightarrow} \chi_{A_2}(t) = \prod_{i=2}^n (t - \lambda_i)$$

Seien $\pi_1, \pi_2 \in \mathrm{End}_K(V)$ gegeben durch $M_B(\pi_1) = \mathrm{diag}(1, 0, ..., 0)$ und $M_B(\pi_2) = \mathrm{diag}(0, 1, ..., 1)$. Dann ist $\pi_1 + \pi_2 = \mathrm{id}_V$ und $f_i = \pi_1 \circ f$ ist $f = \mathrm{id}_V \circ f = f_1 + f_2$ und $f_2|_{V_2} \in \mathrm{End}_K(V_2)$. Nach Induktionshypothese ist $f_2|_{V_2}$ trigonalisierbar, da $M_B(f_2|_{V_2}) = A_2$, also $\chi_{f_2|_{V_2}} = \chi_{A_2}$. Dies bedeutet, es gibt also eine Basis $B_2' = (x_2', ..., x_n')$ von V_2, für die $M_{B_2'}(f_2|_{V_2})$ eine obere Dreiecksmatrix ist. Somit ist für $B' = (x_1, x_2', ..., x_n')$ auch

$$M_{B'}(f) = M_{B'}(f_1) + M_{B'}(f_2)$$
$$= \begin{pmatrix} \lambda_1 & * \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & M_{B_2'}(f_2|_{V_2}) \end{pmatrix}$$

eine obere Dreiecksmatrix. □

Folgerung 4.9
Ist K algebraisch abgeschlossen, so ist jedes $f \in \mathrm{End}_K(V)$ trigonalisierbar.

Beweis. Ist K algebraisch abgeschlossen, so zerfällt nach I.6.14 jedes Polynom über K in Linearfaktoren, ins-

besondere also χ_f. □

Folgerung 4.10
Ist V ein endlichdimensionaler $\mathbb{C}$-VR, so ist jedes $f \in \mathrm{End}_{\mathbb{C}}(V)$ trigonalisierbar.

Beweis. Nach dem Fundamentalsatz der Algebra I.6.16 ist $\mathbb{C}$ algebraisch abgeschlossen. □

5. Das Minimalpolynom

Definition 5.1
Für ein Polynom $P(t) = \sum_{i=0}^{n} c_i t^i \in K[t]$ definieren wir $P(f) = \sum_{i=0}^{m} c_i f^i \in \mathrm{End}_K(V)$, wobei $f^0 = \mathrm{id}_V$, $f^1 = f$, $f^2 = f \circ f$, ...

Analog definiert man $P(A)$ für $A \in \mathrm{Mat}_n(K)$.

▶ **Bemerkung 5.2**
Die Abbildung $\begin{cases} K[t] \to \mathrm{End}_K(V) \\ P \mapsto P(f) \end{cases}$ ist ein Homomorphismus von K-VR und Ringen. Sein Kern ist das Ideal

$$\mathcal{I}_f := \{P \in K[t] \mid P(f) = 0\}$$

und sein Bild ist der kommutative Unterring

$$\begin{aligned} K[f] :&= \{P(f) \mid P \in K[t]\} \\ &= \mathrm{span}_K(f^0, f^1, f^2, ...) \end{aligned}$$

des (im Allgemeinen nicht kommutativen) Rings $\mathrm{End}_K(V)$.

Analog definiert man $\mathcal{I}_A$ und $K[A] \leq \mathrm{Mat}_n(K)$.

Lemma 5.3
$\mathcal{I}_f \neq \{0\}$

Beweis. Wäre $\mathcal{I}_f = \{0\}$, so wäre $K[t] \to \mathrm{End}_K(V)$ injektiv, aber $\dim_K(K[t]) = \infty > n^2 = \dim_K(\mathrm{End}_K(V))$, ein Widerspruch. □

Satz 5.4
Es gibt ein eindeutig bestimmtes normiertes Polynom $0 \neq P \in K[t]$ kleinsten Grades mit $P(f) = 0$. Dieses teilt jedes $Q \in K[t]$ mit $Q(f) = 0$.

Beweis. Nach Lemma 5.3 gibt es $0 \neq P \in K[t]$ mit $P(f) = 0$ von minimalem Grad d. Indem wir durch den Leitkoeffizienten von P teilen, können wir annehmen, dass P normiert ist.
Sei $Q \in \mathcal{I}_f$. Polynomdivision liefert $R, H \in K[t]$ mit $Q = P \cdot H + R$ und $\deg(R) < \deg(P) = d$. Es folgt $R(f) = \underbrace{Q(f)}_{=0} - \underbrace{P(f)}_{=0} \cdot H(f) = 0$. Aus der Minimalität von d folgt $R = 0$ und somit $P|Q$.
Ist Q zudem normiert vom Grad d, so ist $H = 1$, also $Q = P$, was die Eindeutigkeit zeigt. □

Definition 5.5 (Minimalpolynom)
Das eindeutig bestimmte normierte Polynom $0 \neq P \in K[t]$ kleinsten Grades mit $P(f) = 0$ nennt man das Minimalpolynom P_f von f.

Analog definiert man das Minimalpolynom $P_A \in K[t]$ einer Matrix $A \in \mathrm{Mat}_n(K)$.

Mathematica/WolframAlpha-Befehle (Minimalpolynom)
Die Funktion für das Minimalpolynom p mit der Variable t in Mathematica bzw. WolframAlpha lautet:

```
MinimalPolynomial[p,x]
```

■ **Beispiel 5.6**

1. $A = \mathbb{1}_n$, $\chi_A(t) = (t-1)^n$, $P_A(t) = t - 1$
2. $A = 0$, $\chi_A(t) = t^n$, $P_A(t) = t$
3. Ist $A = \operatorname{diag}(a_1, ..., a_n)$ mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_r$, so ist $\chi_A(t) = \prod_{i=1}^{n}(t-a_i) = \prod_{i=1}^{n}(t-\lambda_i)^{\mu_a(f_A,\lambda_i)}$, $P_A(t) = \prod_{i=1}^{r}(t-\lambda_i)$ und es folgt $\deg(P_A) \geq |\{a_1, ..., a_n\}| = r$.

Definition 5.7 (f-zyklisch)
Ein f-invarianter UVR $W \leq V$ heißt f-zyklisch , wenn es ein $x \in W$ mit $W = \operatorname{span}_K(x, f(x), f^2(x), ...)$ gibt.

Lemma 5.8
Sei $x \in V$ und $x_i = f(x)$. Es gibt ein kleinstes k mit $x_k \in \operatorname{span}_K(x_0, x_1, ..., x_{k-1})$, und $W = \operatorname{span}_K(x_0, ..., x_{k-1})$ ein f-zyklischer UVR von V mit Basis $B = (x_0, ..., x_{k-1})$ und $M_B(f|_W) = M_{\chi_{f|_W}}$.

Beweis. Da $\dim_K(V) = n$ ist $(x_0, ..., x_n)$ linear abhängig, es gibt also ein kleinstes k mit $(x_0, ..., x_{k-1})$ linear unabhängig, aber $(x_0, ..., x_k)$ linear abhängig, folglich $x_k \in \operatorname{span}_K(x_0, ..., x_{k-1})$. Mit $x_k = f(x_{k-1}) = \sum_{i=0}^{k-1} -c_i x_i$ ist dann Da $\dim_K(V) = n$ ist $(x_0, ..., x_n)$ linear abhängig, es gibt also ein kleinstes k mit $(x_0, ..., x_{k-1})$ linear unabhängig, aber $(x_0, ..., x_k)$ linear abhängig, folglich $x_k \in \operatorname{span}_K(x_0, ..., x_{k-1})$. Mit $x_k = f(x_{k-1}) = \sum_{i=0}^{k-1} -c_i x_i$ ist dann

$$M_B(f|_W) = \begin{pmatrix} 0 & ... & ... & ... & 0 & -c_0 \\ 1 & \ddots & & & \vdots & \vdots \\ 0 & \ddots & \ddots & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & ... & 0 & 1 & 0 & -c_{k-1} \end{pmatrix}$$

somit $\chi_{f|_W} = t^k + \sum_{i=0}^{k-1} c_i t^i$, also $M_B(f|_W) = M_{\chi_{f|_W}}$. □

Theorem 5.9 (Satz von Cayley-Hamiltion)
Für $f \in \operatorname{End}_K(V)$ ist $\chi_f(f) = 0$.

Beweis. Sei $x \in V$. Definiere $x_i = f^i(x)$ und $W = \operatorname{span}_K(x_0, ..., x_{k-1})$ wie in Lemma 5.8. Sei $\chi_{f|_W} = t^k +$

$\sum_{i=0}^{k-1} c_i t^i$, also $f(x_{k-1}) = \sum_{i=0}^{k-1} -c_i x_i$. Wenden wir $\chi_{f|_W}(f) \in \mathrm{End}_K(V)$ auf x an, so erhalten wir

$$\begin{aligned}\chi_{f|_W}(f)(x) &= \left(f^k + \sum_{i=1}^{k-1} c_i f^i\right)(x) \\ &= \sum_{i=1}^{k-1} -c_i x_i + \sum_{i=1}^{k-1} c_i x_i \\ &= 0\end{aligned}$$

Aus $\chi_{f|_W} | \chi_f$ (Beispiel 4.6) folgt somit $\chi_f(f)(x) = 0$, denn ist $\chi_f = Q \cdot \chi_{f|_W}$ mit $Q \in K[t]$, so ist $\chi_f(f) = Q(f) \circ \chi_{f|_W}(f)$, also $\chi_f(f)(x) = Q(f)(\underbrace{\chi_{f|_W}(f)(x)}_{=0}) = 0$. Da $x \in V$ beliebig war, folgt $\chi_f(f) = 0 \in \mathrm{End}_K(V)$. □

Folgerung 5.10
Es gilt $P_f | \chi_f$. Insbesondere ist $\deg(P_f) \leq n$.

Beweis. Theorem 5.9 + Satz 5.4 □

▶ **Bemerkung 5.11**
Ist B eine Basis von V und $A = M_B(f)$, so ist $P_A = P_f$. Insbesondere ist $P_A = P_B$ für $A \sim B$. Als Spezialfall von Theorem 5.9 erhält man $\chi_A(A) = 0$ und $P_A | \chi_A$.

▶ **Bemerkung 5.12**
Der naheliegende "Beweis" $\underbrace{\chi_A}_{\in \mathrm{Mat}_n(K)} = \det(t\mathbb{1}_n - A)(A) = \det(A\mathbb{1}_n - A) = \det(0) = \underbrace{0}_{\in K}$ ist falsch!

6. Nilpotente Endomorphismen

▶ **Bemerkung 6.1**
Für $f \in \mathrm{End}_K(V)$ sind

- $f\{0\} = \mathrm{Ker}(f^0) \subseteq \mathrm{Ker}(f^1) \subseteq \mathrm{Ker}(f^2) \subseteq ...$
- $V = \mathrm{Im}(f^0) \supseteq \mathrm{Im}(f^1) \supseteq \mathrm{Im}(f^2) \supseteq ...$

Folgen von UVR von V. Nach der Kern-Bild-Formel III.7.13 ist

$$\dim_K(\mathrm{Ker}(f^i)) + \dim_K(\mathrm{Im}(f^i)) = \dim_K(V) \quad \forall i$$

Da $\dim_K(V) = n < \infty$ gibt es ein d mit $\mathrm{Ker}(f^d) = \mathrm{Ker}(f^{d+i})$ und $\mathrm{Im}(f^d) = \mathrm{Im}(f^{d+i})$ für jedes $i \geq 0$.

■ **Beispiel 6.2**
$f = f_A$, $A \in \mathrm{Mat}_2(K)$.

- $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$: $\{0\} = \mathrm{Ker}(f^0) = \mathrm{Ker}(f^1) = ...$
- $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$: $\{0\} = \mathrm{Ker}(f^0) \subset \mathrm{Ker}(f^1) = \mathrm{Ker}(f^2) = ... = \mathrm{span}_K(e_2)$
- $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$: $\{0\} = \mathrm{Ker}(f^0) \subset \underbrace{\mathrm{Ker}(f^1)}_{=\mathrm{span}_K(e_1)} \subset \mathrm{Ker}(f^2) = ... = K^2$
- $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$: $\{0\} = \mathrm{Ker}(f^0) \subset \mathrm{Ker}(f^1) = \mathrm{Ker}(f^2) = ... = K^2$

Lemma 6.3
Seien $f, g \in \mathrm{End}_K(V)$. Wenn f und g kommutieren, d.h. $f \circ g = g \circ f$, so sind die UVR $\mathrm{Ker}(g)$ und $\mathrm{Im}(g)$ f invariant.

Beweis. Ist $x \in \mathrm{Ker}(f)$, so ist $g(f(x)) = f(g(x)) = f(0) = 0$, also $f(x) \in \mathrm{Ker}(g)$. Für $g(x) \in \mathrm{Im}(g)$ ist $f(g(x)) = g(f(x)) \in \mathrm{Im}(g)$. □

Satz 6.4 (Lemma von Fitting)
Seien $V_i = \mathrm{Ker}(f^i)$, $W_i = \mathrm{Im}(f^i)$, $d = \min\{i : V_i = V_{i+1}\}$. Dann sind

$$\{0\} = V_0 \subsetneq V_1 \subsetneq ... \subsetneq V_d = V_{d+1} = ...$$
$$V = W_0 \supsetneq W_1 \supsetneq ... \supsetneq W_d = W_{d+1} = ...$$

Folgen f-invarianter UVR und $V = V_d \oplus W_d$.

Beweis. Da f^i und f^j für beliebige i, j kommutieren, sind V_i und V_j nach Lemma 6.3 f-invariant für jedes i. Aus $\dim_K(V_i)+\dim_K(W_i) = n$ folgt $d = \min\{i : W_i = W_{i+1}\}$, insbesondere ist $\mathrm{Im}(f^d) = \mathrm{Im}(f^{d+1}) = f(\mathrm{Im}(f^d))$, somit $W_{d+i} = \mathrm{Im}(f^{d+i}) = W_d$ für $i \geq 0$, also auch $V_d = V_{d+i}$ für alle $i \geq 0$.
Insbesondere ist $f^d|_{W_d} : W_d \to W_{2d} = W_d$ surjektiv, also auch injektiv, also $V_d \cap W_d = \{0\}$. Aus der Dimensionsformel II.4.12 folgt dann $\dim_K(V_d + W_d) = \dim_K(V_d) + \dim_K(W_d) = \dim_K(V)$. Folglich ist $V_d + W_d = V$ und $V_d \cap W_d = \{0\}$, also $V = V_d \oplus W_d$. □

Definition 6.5 (nilpotent)
Ein $f \in \mathrm{End}_K(V)$ heißt nilpotent , wenn $f^k = 0$ für ein $k \in \mathbb{N}$. Analog heißt $A \in \mathrm{Mat}_n(K)$ nilpotent, wenn $A^k = 0$ für $k \in \mathbb{N}$. Das kleinste k mit $f^k = 0$ bzw. A^k heißt die Nilpotenzklasse von f bzw. A.

Lemma 6.6
Ist f nilpotent, so gibt es eine Basis B von V, für die $M_B(f)$ eine strikte obere Dreiecksmatrix ist.

Beweis. Induktion nach $n = \dim_K(V)$.
$n = 1$: $f^k = 0 \Rightarrow f = 0$
$n > 1$: Sei k die Nilpotenzklasse von f und $U = \mathrm{Ker}(f^{k-1})$. Dann ist $U \subset V$. Da $f^k = f^{k-1} \circ f$ ist $f(V) \subset U$, insbesondere $f|_U \in \mathrm{End}_K(U)$. Da $f|_U$ nilpotent ist, gibt es nach I.H. eine Basis B_0 von U, für die $M_B(f|_U)$ eine strikte obere Dreiecksmatrix ist. Ergänze B_0 zu einer Basis B von V. Da $f(V) \subset U$ ist dann auch

$$M_B(f) = \begin{pmatrix} M_{B_0}(f|_U) & * \\ 0 & 0 \end{pmatrix}$$

eine strikte obere Dreiecksmatrix. □

Satz 6.7
Für $f \in \mathrm{End}_K(V)$ sind äquivalent:

1) f ist nilpotent
2) $f^n = 0$ für $n \in \mathbb{N}$
3) $P_f(t) = t^r$ für ein $r \leq n$
4) $\chi_f(t) = t^n$
5) Es gibt eine Basis B von V, mit

$$M_B(f) = \begin{pmatrix} 0 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ & & & 0 \end{pmatrix}$$

eine strikte obere Dreiecksmatrix ist.

Beweis.

- 1) $\Rightarrow$ 5): Lemma 6.6
- 5) $\Rightarrow$ 4): Beispiel 2.8
- 4) $\Rightarrow$ 3): Nach Folgerung 5.10 ist $P_f|\chi_f = t^n$, also $t^n = P_f(t)Q(t)$ mit $Q \in K[t]$. Schreibe $P_f(t) = t^a \cdot P_1(t), Q(t) = t^b \cdot Q_1(t)$ mit $a, b \in \mathbb{N}$, $P_1, Q_1 \in K[t]$, $P_1(0) \neq 0$, $Q_1(0) \neq 0$
  $\overset{3.8}{\Rightarrow} t^{n-(a+b)} = P_1(t)Q_1(t)$ und $(P_1Q_1)(0) \neq 0$
  $\Rightarrow n - (a+b) = 0 \Rightarrow P_1 = 1$, somit $P_f(t) = t^a$
- 3) $\Rightarrow$ 2): $t^r = 0$, $r \leq n \Rightarrow f^n = 0$
- 2) $\Rightarrow$ 1): nach Definition □

Folgerung 6.8
Die Nilpotenzklasse eines nilpotenten Endomorphismus $f \in \mathrm{End}_K(V)$ ist höchstens $\dim_K(V)$.

Folgerung 6.9
Ist $d := \min\{i \mid \mathrm{Ker}(f^i) = \mathrm{Ker}(f^{i+1})\}$, so ist $d \leq \dim_K(\mathrm{Ker}(f)) = \mu_a(f, 0)$.

Beweis. Sei $V_d = \mathrm{Ker}(f^d)$, $W_d = \mathrm{Im}(f^d)$, $k = \dim_K(V_d)$. Da $V = V_d \oplus W_d$ ist $\chi_f = \chi_{f|_{V_d}} \cdot \chi_{f|_{W_d}}$. Da $f|_{V_d}$ nilpotent ist, ist $\chi_{f|_{V_d}} = t$ nach Satz 6.7. Da $f|_{W_d}$ injektiv ist, ist $\chi_{f|_{W_d}}(0) \neq 0$. Somit ist $\mu_a(f, 0) = \mu(\chi_f, 0) \overset{3.6}{=} k$. Da $\dim_K(\mathrm{Ker}(f^d)) > ... > \dim_K(\mathrm{Ker}(f)) > 0$ ist $k = \dim_K(\mathrm{Ker}(f^d)) \geq d$, falls $d > 0$, sonst klar. □

▶ **Bemerkung 6.10**
Die Bedeutung nilpotenter Endomorphismen beim Finden geeigneter Basen ergibt sich aus der folgenden Beobachtung:
Ist A eine obere Dreiecksmatrix, so ist $A = D + N$, wobei D eine Diagonalmatrix ist und N eine strikte obere Dreiecksmatrix ist. Anders gesagt: Jeder trigonalisierbare Endomorphismus ist Summe aus einem diagonalisierbaren und einem nilpotenten Endomorphismus.

Definition 6.11 (Jordan-Matrix)
Für $k \in \mathbb{N}$ definieren wir die JORDAN-Matrix

$$J_k = \begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & ... & ... & ... & 0 \end{pmatrix} \in \mathrm{Mat}_k(K)$$

weiter setzen wir für $\lambda \in K$ $J_k(\lambda) := \lambda \mathbb{1} + J_k$.

Lemma 6.12
Die JORDAN-Matrix J_k ist nilpotent von Nilpotenzklasse k.

Beweis. Es ist $(J_k)^r = (\delta_{i+r,j})_{i,j}$ für $r \geq 1$. □

Satz 6.13
Ist f nilpotent von Nilpotenzklasse k, so gibt es eindeutig bestimmte $r_1, .., r_k \in \mathbb{N}_{>0}$ mit $\sum_{d=1}^{k} dr_d = n$ und eine Basis B von V mit

$$M_B(f) = \text{diag}(\underbrace{J_k, ..., J_k}_{r_k \text{ viele}}, ..., \underbrace{J_1, ..., J_1}_{r_1 \text{ viele}})$$

Beweis. Sei $U_i = \text{Ker}(f^i)$. Nach Satz 6.4 haben wir eine Folge $\{0\} = U_0 \subset U_1 \subset ... \subset U_k = V$ mit $f(U_i) \subseteq U_{i-1}$ für alle $i > 0$.
Wir konstruieren eine Zerlegung $V = \bigoplus_{d=1}^{k} W_d$ mit $U_i = U_{i-1} \oplus W_i$, $f(W_i) \subseteq W_{i-1}$, $f|_{W_d}$ injektiv für $i > 1$.

$$\begin{aligned} V &= U_k \\ V &= U_{k-1} \oplus W_k \\ V &= U_{k-2} \oplus W_{k-1} \oplus W_k \\ &\vdots \\ V &= U_0 \oplus W_1 \oplus ... \oplus W_k \end{aligned}$$

Wähle W_k mit $V = U_k = U_{k-1} \oplus W_k$. Ist $k > 1$, so ist $W_k \cap \text{Ker}(f) \subseteq W_k \cap U_{k-1} = \{0\}$, also $f|_{W_k}$ ist injektiv. Des weiteren ist $f(W_k) \subseteq U_{k-1}$ und aus $W_k \cap U_{k-1} = \{0\}$ folgt $f(W_k) \cap U_{k-2} = \{0\}$. Wir können deshalb W_{k-1} mit $U_{k-1} = U_{k-2} \oplus W_{k-1}$ und $f(W_k) \subseteq W_{k-1}$ wählen. Somit ist $V = U_{k-1} \oplus W_k = U_{k-2} \oplus W_{k-1} \oplus W_k$. Wir setzen dies fort und erhalten $V = U_0 \oplus W_1 \oplus ... \oplus W_k$ mit $f(W_i) \subseteq W_{i-1}$ und $f|_{W_i}$ injektiv für $i > 1$, wobei $U_0 = \{0\}$ und $W_1 = \text{Ker}(f)$.
Sie $r_d = \dim_K(W_d) - \dim_K(W_{d+1})$, wobei wir $W_{k+1} = \{0\}$. Wähle nun eine Basis $(x_{k,1}, ..., x_{k,r_k})$ von W_k. Ist $k > 1$, so ist $f|_{W_k}$ injektiv und wir können $(f(x_{k,1}), ..., f(x_{k,r_k}))$ durch Elemente $x_{k-1,1}, ..., x_{k-1,r_{k-1}}$ zu einer Basis von W_{k-1} ergänzen, und so weiter.
Da $V = \bigoplus_{d=1}^{k} W_d$ ist

$$B = \{f^i(x_{d,j}) \mid d = 1, ..., k, j = 1, ..., r_d, i = 0, ..., d-1\}$$

eine Basis von V, die bei geeigneter Anordnung das Gewünschte leistet.
Es bleibt zu zeigen, dass $r_1, ..., r_k$ eindeutig bestimmt sind. Ist B_0 eine Basis, für die $M_{B_0}(f)$ in der gewünschten Form ist, so ist

$$\begin{aligned} \dim_K(U_1) &= \sum_{d=1}^{k} r_d \\ \dim_K(U_2) &= \sum_{d=2}^{k} r_d + \sum_{d=1}^{k} r_d \\ &\vdots \\ \dim_K(U_k) &= \sum_{d=k}^{k} r_d + ... + \sum_{d=1}^{k} r_d \end{aligned}$$

woraus man sieht, dass $r_1, ..., r_k$ durch $U_1, ..., U_k$, also durch f eindeutig bestimmt. □

■ **Beispiel 6.14**
Sei $f = f_A$ mit $A = \begin{pmatrix} 0 & 1 & 3 \\ & 0 & 2 \\ & & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$

$$A^2 = \begin{pmatrix} 0 & 0 & 2 \\ & 0 & 0 \\ & & 0 \end{pmatrix}, A^3 = 0$$

$\Rightarrow k = 3, U_0 = \{0\}, U_1 = \mathbb{R}e_1, U_2 = \mathbb{R}e_1 + \mathbb{R}e_2, U_3 = V$.
Wähle W_3 mit $V = U_3 = U_2 \oplus W_3$, z.B. $W_3 = \mathbb{R}e_3$.
Wähle W_2 mit $U_2 = U_1 \oplus W_2$ und $f(W_3) \subseteq W_2$, also

$$W_2 = \mathbb{R} \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$$

Setze $W_1 = U_1 = \mathrm{Ker}(f) = \mathbb{R}e_1 \Rightarrow$ Basis $B = (f^2(e_3), f(e_3), e_3)$

$$M_B(f) = \begin{pmatrix} 0 & 1 & 0 \\ & 0 & 1 \\ & & 0 \end{pmatrix}$$

7. Die Jordan-Normalform

Definition 7.1 (Hauptraum)
Der Hauptraum von f zum EW λ der Vielfachheit $r = \mu_a(f, \lambda)$ ist

$$\mathrm{Hau}(f, \lambda) = \mathrm{Ker}\left((f - \lambda \,\mathrm{id}_V)^r\right)$$

Lemma 7.2
$\mathrm{Hau}(f, \lambda)$ ist ein f-invarianter UVR der Dimension $\dim_K(\mathrm{Hau}(f, \lambda)) = \mu_a(f, \lambda)$, auf dem $f - \lambda \,\mathrm{id}_V$ nilpotent ist und $\chi_{f|_{\mathrm{Hau}(f,\lambda)}} = (t - \lambda)^{\mu_a(f,\lambda)}$

Beweis. f kommutiert sowohl mit f als auch mit id_V, somit auch mit $(f - \lambda \,\mathrm{id}_V)^r$. Die f-Invarianz von $U = \mathrm{Hau}(f, \lambda)$ folgt aus Lemma 6.3. Nach Folgerung 6.9 ist $\dim_K(U) = \mu_a(f - \lambda \,\mathrm{id}_V, 0)$ und da $\chi_f(t) = \chi_{f-\lambda \,\mathrm{id}_V}(t - \lambda)$ ist $\mu_a(f, \lambda) = \mu(\chi_f, \lambda) = \mu_a(f - \lambda \,\mathrm{id}_V, 0)$. Da $f - \lambda \,\mathrm{id}_V \,|_U$ nilpotent ist $\chi_{f-\lambda \,\mathrm{id}_V \,|_U}(t) = t^r$, somit $\chi_{f|_U}(t) = (t - \lambda)^r$. □

Satz 7.3 (Hauptraumzerlegung)
Ist $\chi_f(t) = \prod_{i=1}^m (t - \lambda_i)^{r_i}$ mit $\lambda_1, ..., \lambda_m \in K$ paarweise verschieden und $r_1, ..., r_m \in \mathbb{N}$, so ist $V = \bigoplus_{i=1}^m V_i$ mit $V_i = \mathrm{Hau}(f, \lambda_i)$ eine Zerlegung in f-invariante UVR und für jedes i ist $\chi_{f|_{V_i}}(t) = (t - \lambda_i)^{r_i}$.

Beweis. Induktion nach m.
$m = 1$: $r_1 = n \overset{7.2}{\Rightarrow} V = V_1$.
$m - 1 \to m$: Nach Satz 6.4 ist $V = V_1 \oplus W_1$ mit $W_1 = \mathrm{Im}((f - \lambda_i \,\mathrm{id}_V)^r)$ eine Zerlegung in f-invariante UVR mit $\dim_K(V_1) = r_1$, $\dim_K(W_1) = n - r_1$. Somit ist $\chi_f = \chi_{f|_{V_1}} \cdot \chi_{f|_{W_1}}$ und $\chi_{f|_{V_1}} \overset{7.2}{=} (t - \lambda_1)^{r_1}$ also $\chi_{f|_{W_1}} = \prod_{i=2}^m (t - \lambda_i)^{r_i}$. Nach I.H. ist also $W_1 = \bigoplus_{i=2}^m \mathrm{Hau}(f|_{W_1}, \lambda_i)$. Es ist für $i \geq 2$ $\mathrm{Hau}(f|_{W_1}, \lambda_i) \subseteq \mathrm{Hau}(f, \lambda_i) = V_i$ und da $\dim_K(\mathrm{Hau}(f|_{W_1}, \lambda_i)) = r_i = \dim_K(\mathrm{Hau}(f, \lambda_i))$ gilt Gleichheit. Damit ist

$$\begin{aligned} V &= V_1 \oplus W_1 \\ &= V_1 \oplus \bigoplus_{i=2}^m \mathrm{Hau}(f|_{W_1}, \lambda_i) \\ &= V_1 \oplus \bigoplus_{i=2}^m V_i \\ &= \bigoplus_{i=1}^m V_i \end{aligned}$$

□

■ **Beispiel 7.4**
$f = f_A$

$$A = \begin{pmatrix} 1 & 3 & \\ & 1 & 4 \\ & & 2 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

$\chi_A(t) = (t-1)^2(t-2) \Rightarrow \mathbb{R}^3 = \underbrace{\mathrm{Hau}(f,1)}_{\dim=2} \oplus \underbrace{\mathrm{Hau}(f,2)}_{\dim 1}$
$\mathrm{Hau}(f,1) = \mathrm{Ker}((f-\mathrm{id})^2) = L((A-\mathbb{1})^2, 0)$
$\mathrm{Hau}(f,2) = \mathrm{Ker}(f - 2\,\mathrm{id}) = \mathrm{Eig}(f,2) = L(A - 2\mathbb{1}, 0)$

$$A - \mathbb{1} = \begin{pmatrix} 0 & 3 & \\ & -1 & 4 \\ & & 0 \end{pmatrix}, (A-\mathbb{1})^2 = \begin{pmatrix} 0 & & 12 \\ & 0 & 4 \\ & & 1 \end{pmatrix} \qquad \Rightarrow \mathrm{Hau}(f,1) = \mathbb{R}e_1 + \mathbb{R}e_2$$

$$A - 2\mathbb{1} = \begin{pmatrix} -1 & 3 & \\ & -1 & 4 \\ & & 0 \end{pmatrix} \qquad \Rightarrow \mathrm{Hau}(f,2) = \mathbb{R}\begin{pmatrix} 12 \\ 4 \\ 1 \end{pmatrix}$$

Mit $B = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 12 \\ 4 \\ 1 \end{pmatrix} \right)$ ist

$$M_B(f) = \begin{pmatrix} \begin{pmatrix} 1 & 3 \\ & 1 \end{pmatrix} & \\ & 2 \end{pmatrix}$$

Theorem 7.5 (Jordan-Normalform)
Sei $f \in \mathrm{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom χ_f in Linearfaktoren zerfällt. Dann gibt es $r \in \mathbb{N}$, $\mu_1, ..., \mu_r \in K$ und $k_1, ..., k_r \in \mathbb{N}$ mit $\sum_{i=1}^r k_i = \dim_K(V)$ und eine Basis B von V mit

$$M_B(f) = \mathrm{diag}(J_{k_1}(\mu_1), ..., J_{k_r}(\mu_r))$$

Die Paare $(\mu_1, k_1), ..., (\mu_r, k_r)$ heißen die JORDAN-Invarianten von f und sind bis auf Reihenfolge eindeutig bestimmt.

Beweis. Schreibe $\chi_f(t) = \prod_{i=1}^m (t-\lambda_i)^{r_i}$ mit $\lambda_1, ..., \lambda_m \in K$ paarweise verschieden, $r_i \in \mathbb{N}$. Sei $V_i = \mathrm{Hau}(f, \lambda_i)$. Nach Satz 7.3 ist $V = \bigoplus_{i=1}^m V_i$ eine Zerlegung in f-invariante UVR. Für jedes i wenden wir Satz 6.13 auf $(f - \lambda_i\,\mathrm{id}_V)|_{V_i}$ an und erhalten eine Basis B_i von V_i und $k_{i,1} \geq ... \geq k_{i,s_i}$ mit

$$M_B((f-\lambda_i\,\mathrm{id})|_{V_i}) = \mathrm{diag}(J_{k_{i,1}}, ..., J_{k_{i,s_i}})$$

Es folgt $M_{B_i}(f|_{V_i}) = M_{B_i}(\lambda_i\,\mathrm{id}_{V_i}) + M_{B_i}((f-\lambda_i\,\mathrm{id}_V)|_{V_i})$. Ist nun B die Vereinigung der B_i, so hat $M_B(f)$ die gewünschte Form. Die Eindeutigkeit der JORDAN-Invarianten folgt aus der Eindeutigkeit der $k_{i,j}$ in Lemma 6.3.□

▶ **Bemerkung 7.6**

Ist K algebraisch abgeschlossen, so haben wir nun eine (bis auf Permutationen) eindeutige Normalform für Endomorphismen $f \in \mathrm{End}_K(V)$ gefunden. Aus ihr lassen sich viele Eigenschaften des Endomorphismus leicht ablesen.

Folgerung 7.7

Sei $f \in \mathrm{End}_K(V)$ trigonalisierbar mit $\chi_f(t) = \prod_{i=1}^m (t-\lambda_i)^{\mu_a(f,\lambda_i)}$, $P_f(t) = \prod_{i=1}^m (t-\lambda_i)^{d_i}$ und JORDAN-Invarianten $(\mu_1, k_1), ..., (\mu_r, k_r)$. Mit $J_i = \{j \mid \mu_j = \lambda_i\}$ ist dann

$$\begin{aligned} \mu_g(f, \lambda_i) &= |J_i| \\ \mu_a(f, \lambda_i) &= \sum_{j \in J_i} k_j \\ d_i &= \max\{k_j \mid j \in J_i\} \end{aligned}$$

Beweis.

- μ_a: klar, da $\chi_f(t) = \prod_{j=1}^r (t-\mu_j)^{k_j} = \prod_{i=1}^m (t-\lambda_i)^{\mu_a(f,\lambda_i)}$
- μ_g: lese Basis von $\mathrm{Eig}(f, \lambda_i)$ aus JORDAN-NF: Jeder Block $J_{k_j}(\lambda_i)$ liefert ein Element der Basis.
- d_i: folgt, da J_{k_j} nilpotent von Nilpotenzklasse k_j ist (Lemma 6.12). □

Folgerung 7.8

Genau dann ist f diagonalisierbar, wenn

$$\begin{aligned} \chi_f(t) &= \prod_{i=1}^m (t-\lambda_i)^{r_i} \quad \lambda_1, ..., \lambda_m \in K \text{ paarweise verscheiden und} \\ P_f(t) &= \prod_{i=1}^m m(t-\lambda_i) \end{aligned}$$

Beweis. Genau dann ist f diagonalisierbar, wenn f trigonalisierbar ist und die JORDAN-NF die Diagonalmatrix ist (Eindeutigkeit der JNF), also $k_j = 1$ für alle j. Nach Folgerung 7.7 ist dies äquivalent dazu, dass $d_i = 1$ für alle i, also $P_f = \prod_{i=1}^m (t-\lambda_i)$. □

▶ **Bemerkung 7.9**

Wider definiert man die JORDAN-Invarianten, etc. von einer Matrix $A \in \mathrm{Mat}_n(K)$ als die JORDAN-Invarianten von $f_A \in \mathrm{End}_K(K^n)$.

Folgerung 7.10

Seien $A, B \in \mathrm{Mat}_n(K)$ trigonalisierbar. Genau dann ist $A \sim B$, wenn A und B die gleichen JORDAN-Invarianten haben.

Beweis. Existenz und Eindeutigkeit der JORDAN-Normalform. □

Kapitel VI

Skalarprodukte

In diesem ganzen Kapitel seien

- $K = \mathbb{R}$ oder $K = \mathbb{C}$
- $n \in \mathbb{N}$
- V ein n-dimensionaler K-VR

1. Das Standardskalarprodukt

Sei zunächst $K = \mathbb{R}$.

Definition 1.1 (Standardskalarprodukt in $\mathbb{R}$)
Auf den Standardraum $V = \mathbb{R}^n$ definiert man das Standardskalarprodukt in $\mathbb{R}$ $\langle . \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ durch

$$\langle x, y \rangle = x^t y = \sum_{i=1}^{n} x_i y_i$$

Satz 1.2
Das Standardskalarprodukt erfüllt die folgenden Eigenschaften:

- Für $x, x', y, y' \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ ist:

$$\begin{aligned} \langle x + x', y \rangle &= \langle x, y \rangle + \langle x', y \rangle \\ \langle \lambda x, y \rangle &= = \lambda \langle x, y \rangle \\ \langle x, y + y' \rangle &= \langle x, y \rangle + \langle x, y' \rangle \\ \langle x, \lambda y \rangle &= \lambda \langle x, y \rangle \end{aligned}$$

- Für $x, y \in \mathbb{R}^n$ ist $\langle x, y \rangle = \langle y, x \rangle$
- Für $x \in \mathbb{R}^n$ ist $\langle x, y \rangle \geq 0$ und $\langle x, x \rangle = 0 \iff x = 0$

Beweis.
- klar
- klar
- $\langle x, x \rangle = \sum_{i=1}^{n} x_i^2 \geq x_j^2$ für jedes $j \Rightarrow \langle x, x \rangle \geq 0$ und $\langle x, x \rangle > 0$ falls $x_j \neq 0$ für ein j. □

Definition 1.3 (euklidische Norm in $\mathbb{R}$)
Auf $K = \mathbb{R}^n$ definiert man euklidische Norm in $\mathbb{R}$ $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ durch

$$\|x\| = \sqrt{\langle x, x\rangle}$$

Satz 1.4 (Ungleichung von Cauchy-Schwarz)
Für $x, y \in \mathbb{R}^n$ gilt

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$$

Gleichheit genau dann, wenn x und y linear abhängig sind.

Beweis. siehe Analysis, siehe VI.§3 □

Satz 1.5
Die euklidische Norm erfüllt die folgenden Eigenschaften:

- Für $x \in \mathbb{R}^n$ ist $\|x\| = 0 \iff x = 0$
- Für $x \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ ist $\|\lambda x\| = |\lambda| \cdot \|x\|$
- Für $x, y \in \mathbb{R}^n$ ist $\|x + y\| \leq \|x\| + \|y\|$

Beweis.
- Satz 1.2
- Satz 1.2
- $\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \overset{1.4}{\Rightarrow} \|x+y\| \leq \|x\| + \|y\|$ □

Sei nun $K = \mathbb{C}$.

Definition 1.6 (komplexe Konjugation, Absolutbetrag)
Für $x, y \in \mathbb{R}$ und $z = x + iy \in \mathbb{C}$ definiert man $\overline{z} = x - iy$ heißt komplexe Konjugation .. Man definiert den Absolutbetrag von z als

$$|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2} \in \mathbb{R}_{\geq 0}$$

Für $A = (a_{ij})_{i,j} \in \mathrm{Mat}_{m\times n}(\mathbb{C})$ sehen wir

$$\overline{A} = (\overline{a_{ij}})_{i,j} \in \mathrm{Mat}_{m\times n}(\mathbb{C})$$

Satz 1.7
Komplexe Konjugation ist ein Ringautomorphismus von $\mathbb{C}$ mit Fixkörper

$$\{z \in \mathbb{C} \mid z = \overline{z}\} = \mathbb{R}$$

Beweis. siehe LAAG1 H47 □

Folgerung 1.8
Für $A, B \in \mathrm{Mat}_n(\mathbb{C})$ und $S \in \mathrm{GL}_n(\mathbb{C})$ ist $\overline{A+B} = \overline{A} + \overline{B}, \overline{AB} = \overline{A} \cdot \overline{B}, \overline{A^t} = \overline{A}^t, \overline{S^{-1}} = \overline{S}^{-1}$

Beweis. Satz 1.7, einfache Übung □

Definition 1.9 (Standardskalarprodukt in $\mathbb{C}$)
Auf $K = \mathbb{C}^n$ definiert man das Standardskalarprodukt in $\mathbb{C}$ $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ durch

$$\langle x, y \rangle = x^t \overline{y} = \sum_{i=1}^{n} x_i \overline{y}_i$$

Satz 1.10
Das komplexe Standardskalarprodukt erfüllt die folgenden Eigenschaften:

- Für $x, x', y, y' \in \mathbb{C}^n$ und $\lambda \in \mathbb{C}$ ist:

$$\begin{aligned} \langle x + x', y \rangle &= \langle x, y \rangle + \langle x', y \rangle \\ \langle \lambda x, y \rangle &= = \lambda \langle x, y \rangle \\ \langle x, y + y' \rangle &= \langle x, y \rangle + \langle x, y' \rangle \\ \langle x, \lambda y \rangle &= \overline{\lambda} \langle x, y \rangle \end{aligned}$$

- Für $x, y \in \mathbb{C}^n$ ist $\langle x, y \rangle = \overline{\langle y, x \rangle}$
- Für $x \in \mathbb{C}^n$ ist $\langle x, y \rangle \in \mathbb{R}_{\geq 0}$ und $\langle x, x \rangle = 0 \iff x = 0$

Beweis.
- klar
- klar
- $\langle x, x \rangle = \sum_{i=1}^{n} x_i \overline{x_i} = \sum_{i=1}^{n} |x_i|^2$ □

Definition 1.11 (euklidische Norm in $\mathbb{C}$)
Auf $V = \mathbb{C}$ definiert man die euklidische Norm in $\mathbb{C}$ $\| \cdot \| : \mathbb{C}^n \to \mathbb{R}_{\geq 0}$ durch

$$\|x\| = \sqrt{\langle x, x \rangle}$$

▶ **Bemerkung 1.12**
Schränkt man das komplexe Skalarprodukt auf den $\mathbb{R}^n$ ein, so erhält man das Standardskalarprodukt auf dem $\mathbb{R}^n$. Wir werden ab jetzt die beiden Fälle $K = \mathbb{R}$ und $K = \mathbb{C}$ parallel behandeln. Wenn nicht anders angegeben, werden wir die Begriffe für den komplexen Fall benutzen, aber auch den reellen Fall einschließen.

2. Bilinearformen und Sesquilinearformen

Sei $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Definition 2.1 (Bilinearform, Sesquilinearform)
Eine Bilinearform ($K = \mathbb{R}$) bzw. Sesquilinearform ($K = \mathbb{C}$) ist eine Abbildung $s : V \times V \to K$ für die gilt:

- Für $x, x', y \in V$ ist $s(x + x', y) = s(x, y) + s(x', y)$
- Für $x, y, y' \in V$ ist $s(x, y + y') = s(x, y) + s(x, y')$
- Für $x, y \in V$, $\lambda \in K$ ist $s(\lambda x, y) = \lambda s(x, y)$
- Für $x, y \in V$, $\lambda \in K$ ist $s(x, \lambda y) = \overline{\lambda} s(x, y)$

▶ **Bemerkung 2.2**
Im Fall $K = \mathbb{R}$ ist $\lambda = \overline{\lambda}$. Wir werden der Einfachheit halber auch in diesem Fall von Sesquilinearformen sprechen, vgl. Bemerkung 1.12

■ **Beispiel 2.3**
Für $A = (a_{ij})_{i,j} \in \mathrm{Mat}_n(K)$ ist $s_A : K^n \times K^n \to K^n$ gegeben durch

$$s_A(x, y) = x^t A \overline{y} = x^t \left(\sum_{j=1}^{n} a_{ij} \overline{y}_j \right)_i = \sum_{i,j=1}^{n} a_{ij} x_i \overline{y}_j$$

eine Sesquilinearform auf $V = K^n$.

Definition 2.4
Sei s eine Sesquilinearform auf V und $B = (v_1, ..., v_n)$ eine Basis von V. Die darstellende Matrix von s bzgl. B ist

$$M_B(s) = (s(v_i, v_j))_{i,j} \in \mathrm{Mat}_n(K)$$

■ **Beispiel 2.5**
Die darstellende Matrix des Standardskalarprodukts $s = s_{\mathbb{1}_n}$ auf den Standardraum $V = K^n$ bzgl. der Standardbasis $\mathcal{E}$ ist

$$M_{\mathcal{E}}(s) = \mathbb{1}_n$$

Lemma 2.6
Seien $v, w \in V$. Mit $x = \Phi_B^{-1}(v)$, $y = \Phi_B^{-1}(w)$ und $A = M_B(s)$ ist $s(v, w) = x^t A \overline{y} = s_A(x, y)$.

Beweis. Achtung: v_i beschreibt das i-te Element der Basis B!
$s(v, w) = s(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j) = \sum_{i,j=1}^n x_i \overline{y} s(v, v_j) = x^t A \overline{y}$ □

Satz 2.7
ISei B eine Basis von V. Die Abbildung $s \mapsto M_B(s)$ ist eine Bijektion zwischen den Sesquilinearformen auf V und $\mathrm{Mat}_n(K)$.

Beweis.
- injektiv: Lemma 2.6
- surjektiv: Für $A \in \mathrm{Mat}_n(K)$ wird durch $s(v,w) = \Phi_B^{-1}(v)^t \cdot A \cdot \overline{\Phi_B^{-1}(w)}$ eine Sesquilinearform auf V mit $M_B(s) = (s(v_i, w_j))_{i,j} = (e_i^t A \overline{e_j})_{i,j} = (e_i A e_j)_{i,j} = A$ definiert. □

Satz 2.8 (Transformationsformel)
Seien B und B' Basen von V und s eine Sesquilinearform auf V. Dann gilt:

$$M_{B'}(s) = (T_B^{B'})^t \cdot M_B(s) \cdot \overline{T_B^{B'}}$$

Beweis. Seien $v, w \in V$. Definiere $A = M_B(s)$, $A' = M_{B'}(s)$, $T = T_B^{B'}$ und $x, y, x', y' \in K^n$ mit $v = \Phi_B(x) = \Phi_B(x')$, $w = \Phi_B(y) = \Phi_B(y')$. Dann ist $x = Tx'$, $y = Ty'$ und somit

$$\begin{aligned}(x')^t A' \overline{y'} &\overset{2.6}{=} s(v,w) \\ &\overset{2.6}{=} x^t A \overline{y} \\ &= (Tx')^t A \overline{Ty'} \\ &= (x')^t T^t A \overline{T} \overline{y'}\end{aligned}$$

Da $v, w \in V$ und somit $x', y' \in K$ beliebig waren, folgt $A = T^t A \overline{T}$. □

■ **Beispiel 2.9**
Sei s das Standardskalarprodukt auf dem K^n und $B = (b_1, ..., b_n)$ eine Basis des K^n. Dann ist

$$M_B(s) = (T_{\mathcal{E}}^B)^t \cdot M_{\mathcal{E}}(s) \cdot \overline{T_{\mathcal{E}}^B} = B^t \cdot \mathbb{1}_n \cdot \overline{B} = B^t B$$

wobei $B = (b_1, ..., b_n) \in \mathrm{Mat}_n(K)$.

Satz 2.10
Sei s eine Sesquilinearform auf V. Dann sind äquivalent:

- Es gibt $0 \neq v \in V$ mit $s(v,w) = 0$ für alle $w \in V$.
- Es gibt $0 \neq w \in V$ mit $s(v,w) = 0$ für alle $v \in V$.
- Es gibt eine Basis B von V mit $\det(M_B(s)) = 0$.
- Für jede Basis B von V gilt $\det(M_B(s)) = 0$.

Beweis. Sei B eine Basis von V, $v = \Phi_B(x)$ und $A = M_B(s)$. Genau dann ist die (semilineare) Abbildung $w \mapsto s(v,w)$ die Nullabbildung, wenn $x^t A \overline{y} = 0$ für alle $y \in K^n$, also wenn $0 = x^t A$, d.h. $A^t x = 0$. Somit ist (1) genau dann erfüllt, wenn A^t nicht invertierbar ist, also wenn $0 = \det(A^t) = \det(A)$. Damit (1) $\Rightarrow$ (4) $\Rightarrow$ (3) $\Rightarrow$ (1) gezeigt und (2) $\iff$ (4) zeigt man analog. □

Definition 2.11 (ausgeartet)
Eine Sesquilinearform s auf V heißt ausgeartet , wenn eine der äquivalenten Bedingungen aus Satz 2.10 erfüllt ist, sonst nicht-ausgeartet.

Definition 2.12 (symmetrisch, hermitesch)
Eine Sesquilinearform s auf V heißt symmetrisch , wenn bzw. hermitesch , wenn

$$s(x,y) = \overline{s(y,x)} \quad \text{für alle } x,y \in V$$

Eine Matrix $A \in \mathrm{Mat}_n(K)$ heißt symmetrisch bzw. hermitesch, wenn $A = A^* = \overline{A}^t = \overline{A^t}$.

Mathematica/WolframAlpha-Befehle (symmetrische bzw. hermitesche Matrizen)
Wie für vieles Andere auch, hat Mathematica bzw. WolframAlpha auch dafür eine Funktion:

```
SymmetricMatrixQ[A]
HermitianMatrixQ[A]
```

Satz 2.13
Sei s eine Sesquilinearform auf V und B eine Basis von V. Genau dann ist s hermitesch, wenn $M_B(s)$ dies ist.

Beweis. ($\Rightarrow$): klar aus Definition von $M_B(s)$.
($\Leftarrow$): $x = \Phi_B^{-1}$, $y = \Phi_B^{-1}(w)$, $\overline{s(v,w)} = \overline{s(v,w)^t} = \overline{(x^t A \overline{y})^t} = y^t \overline{A^t} \overline{x} = s(w,v)$ □

Satz 2.14
Für $A, B \in \mathrm{Mat}_n(K)$ und $S \in \mathrm{GL}_n(K)$ ist $(A+B)^* = A^* + B^*$, $(AB)^* = B^* A^*$, $(A^*)^* = A$ und $(S^{-1})^* = (S^*)^{-1}$.

Beweis. Folgerung 1.8, III.1.14, III.1.15 □

3. Euklidische und unitäre Vektorräume

Lemma 3.1
Sei s eine hermitesche Sesquilinearform auf V. Dann ist $s(x,x) \in \mathbb{R}$ für alle $x \in V$.

Beweis. Da s hermitesch ist, ist $s(x,x) = \overline{s(x,x)}$, also $s(x,x) \in \mathbb{R}$. □

Definition 3.2 (quadratische Form)
Sei s eine hermitesche Sesquilinearform auf V. Die quadratische Form zu s ist die Abbildung

$$q_s : \begin{cases} V \to \mathbb{R} \\ x \mapsto s(x,x) \end{cases}$$

▶ **Bemerkung 3.3**
Die quadratische Form q_s erfüllt das $q_s(\lambda x) = |\lambda|^2 \cdot q_s(x)$ für alle $x \in V$, $\lambda \in K$. Im Fall $K = \mathbb{R}$, $V = \mathbb{R}^n$, $x = (x_1, ..., x_n)^t$, $s = s_A$, $A \in \mathrm{Mat}_n(\mathbb{R})$ ist $q_s(x) = s_A(x,x) = x^t A x = \sum_{i,j=1}^n a_{ij} x_i x_j$ ein "quadratisches Polynom in den Variablen $x_1, ..., x_n$".

Satz 3.4 (Polarisierung)
Sei s ein hermitesche Sesquilinearform auf V. Dann gilt für $x, y \in V$:

$$s(x,y) = \frac{1}{2}(q_s(x+y) - q_s(x) - q_s(y)) \quad K = \mathbb{R}$$
$$s(x,y) = \frac{1}{4}(q_s(x+y) - q_s(x-y) + iq_s(x+iy) - iq_s(x-iy)) \quad K = \mathbb{C}$$

Beweis. Im Fall $K = \mathbb{R}$ ist

$$\begin{aligned} q_s(x+y) - q_s(x) - q_s(y) &= s(x+y, x+y) - s(x,x) - s(y,y) \\ &= s(x,x) + s(x,y) + s(y,x) + s(y,y) - s(x,x) - s(y,y) \\ &= s(x,y) + s(y,x) - 2s(x,y) \end{aligned}$$

Im Fall $K = \mathbb{C}$: ÜA □

Definition 3.5 ((semi)definit, euklidischer VR, unitärer VR)
Sei s eine hermitesche Sesquilinearform auf V. Ist $s(x,x) \geq 0$ für alle $x \in V$, so heißt s positiv semidefinit . Ist $s(x,x) > 0$ für alle $0 \neq x \in V$, so heißt s positiv definit (oder ein Skalarprodukt).

Eine hermitesche Matrix $A \in \mathrm{Mat}_n(K)$ heißt positiv (semi)definit, wenn s_A dies ist.

Einen endlichdimensionalen K-VR zusammen mit positiv definiten hermiteschen Sesquilinearformen nennt man einen euklidischen bzw. unitären VR (oder auch Prähilbertraum). Wenn nicht anderes angegeben, notieren wir die Sesquilinearform mit $\langle \cdot, \cdot \rangle$.

■ **Beispiel 3.6**
Der Standardraum $V = K^n$ zusammen mit dem Standardskalarprodukt ist ein euklidischer bzw. unitärer VR.

■ **Beispiel 3.7**
Ist $A = \text{diag}(\lambda_1, ..., \lambda_n)$ mit $\lambda_i \in \mathbb{R}$, so ist s_A genau dann positiv definit, wenn $\lambda_i > 0$ für alle i, und positiv semidefinit, wenn $\lambda_i \geq 0$ für alle i.

Satz 3.8
Ist V ein unitärer VR und $U \subseteq V$ ein UVR, so ist U mit der Einschränkung des Skalarprodukts wieder ein unitärer VR.

Beweis. klar, die Einschränkung ist wieder positiv definit. □

Definition 3.9
Ist V ein unitärer VR, so definiert man die Norm von $x \in V$ als

$$\|x\| = \sqrt{\langle x, x \rangle} \in \mathbb{R}_{\geq 0}$$

Satz 3.10
Die Norm eines unitären VR erfüllt die folgenden Eigenschaften:

- Für $x \in V$ ist $\|x\| = 0 \iff x = 0$
- Für $x \in V$ und $\lambda \in K$ ist $\|\lambda x\| = |\lambda| \cdot \|x\|$
- Für $x, y \in V$ ist $\|x + y\| \leq \|x\| + \|y\|$

Beweis.
- Das Skalarprodukt ist positiv definit.
- klar
- Wie im Fall im $\mathbb{R}^n$ □

Satz 3.11
Ist V ein unitärer VR, so gilt für $x, y \in V$:

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

Dabei gilt Gleichheit genau dann, wenn x und y linear abhängig sind.

Beweis. Für $y = 0$ ist die Aussage klar.
Sei also $y \neq 0$. Für $\lambda, \mu \in K$ ist

$$\begin{aligned} 0 &\leq \langle \lambda x + \mu y, \lambda x + \mu y \rangle \\ &= \lambda\overline{\lambda} \cdot \langle x, x \rangle + \mu\overline{\mu} \cdot \langle y, y \rangle + \lambda\overline{\mu} \cdot \langle x, y \rangle + \mu\overline{\lambda} \cdot \langle y, x \rangle \end{aligned}$$

Setzt man $\lambda = \overline{\lambda} = \langle y, y \rangle > 0$ und $\mu = -\langle x, y \rangle$ ein, so erhält man

$$\begin{aligned} 0 &\leq \lambda \cdot \|x\|^2 \|y\|^2 + \mu\overline{\mu}\lambda - \lambda\mu\overline{\mu} - \langle x, y \rangle \overline{\lambda} \langle y, x \rangle \\ &= \lambda(\|x\|^2 \|y\|^2 - |\langle x, y \rangle|^2) \end{aligned}$$

Teilen durch λ und Wurzelziehen liefert die Ungleichung. Gilt dort Gleichheit, so ist $\|\lambda x + \mu y\| = 0$ folglich (da $\lambda \neq 0$) sind dann x, y linear unabhängig. Ist $x = \alpha y$ mit $\alpha \in K$, so ist $|\langle x, y \rangle| = |\alpha| \cdot \|y\|^2 = \|x\| \cdot \|y\|$ □

4. Orthogonalität

Sei V ein euklidischer bzw. unitärer Vektorraum.

Definition 4.1 (orthogonal, orthogonales Komplement)
Zwei Vektoren $x, y \in V$ heißen orthogonal , in Zeichen $x \perp y$, wenn $\langle x, y\rangle = 0$. Zwei Mengen $X, Y \subseteq V$ sind orthogonal, in Zeichen $X \perp Y$, wenn $x \perp y$ für alle $x \in X$ und $y \in Y$.

Für $U \subseteq V$ bezeichnet

$$U^\perp = \{x \in V \mid x \perp u \text{ für alle } u \in U\}$$

das orthogonale Komplement zu U.

Lemma 4.2
Für $x, y \in V$ ist

- $x \perp y \iff y \perp x$
- $x \perp 0$
- $x \perp x \iff x = 0$

Beweis. klar □

Satz 4.3
Für $U \subseteq V$ ist $U^\perp$ ein Untervektorraum von V mit $U \perp U^\perp$ und $U \cap U^\perp \subseteq \{0\}$.

Beweis. Linearität des Skalarprodukts im ersten Argument liefert, dass $U^\perp$ ein Untervektorraum ist. Die Aussage $U^\perp \perp U$ ist trivial, $U \perp U^\perp$ folgt dann aus Lemma 4.2. Ist $u \in U \cap U^\perp$, so ist insbesondere $u \perp u$, also $u = 0$ nach Lemma 4.2. □

Definition 4.4 (orthonormal)
Eine Familie $(x_i)_{i\in I}$ von Elementen von V ist orthogonal, wenn $x_i \perp x_j$ für alle $i \neq j$, und orthonormal , wenn zusätzlich $\|x_i\| = 1$ für alle i. Eine orthogonale Basis nennt man eine Orthogonalbasis, eine orthonormale Basis nennt man eine Orthonormalbasis.

▶ **Bemerkung 4.5**
Eine Basis B ist genau dann eine Orthonormalbasis, wenn die darstellende Matrix des Skalarprodukts bezüglich B die Einheitsmatrix ist. (Beispiel: Standardbasis des Standardraum bezüglich des Standardskalarprodukts)

Lemma 4.6
Ist die Familie $(x_i)_{i\in I}$ orthogonal und $x_i \neq 0$ für alle $i \in I$, so ist $(x_i)_{i\in I}$ linear unabhängig.

Beweis. Ist $\sum_{i\in I} \lambda_i x_i = 0$, $\lambda_i \in K$, fast alle gleich 0, so ist $0 = \left\langle \sum_{i\in I} \lambda_i x_i, x_j \right\rangle = \sum_{i\in I} \lambda_i \langle x_i, x_j\rangle = \lambda_j \langle x_j, x_j\rangle$
Aus $x_j \neq 0$ folgt $\langle x_j, x_j\rangle > 0$ und somit $\lambda_j = 0$ für jedes $j \in I$. □

Lemma 4.7
Ist $(x_i)_{i\in I}$ orthogonal und $x_i \neq 0$ für alle i, so ist $(y_i)_{i\in I}$ mit

$$y_i = \frac{1}{\|x_i\|} x_i$$

orthonormal.

Beweis. Für alle i ist $\langle y_i, y_i\rangle = \frac{1}{\|x_i\|^2}\langle x_i, x_i\rangle = 1$.
Für alle $i \neq j$ ist $\langle y_i, y_j\rangle = \frac{1}{\|x_i\|\cdot\|x_j\|}\langle x_i, x_j\rangle = 0$. □

Satz 4.8
Sei $U \subseteq V$ ein Untervektorraum und $B = (x_1, ..., x_k)$ eine Orthonormalbasis von U. Es gibt genau einen Epimorphismus $\mathrm{pr}_U : V \to U$ mit $\mathrm{pr}_U|_U = \mathrm{id}_U$ und $\mathrm{Ker}(\mathrm{pr}_U) \perp U$, insbesondere also $x - \mathrm{pr}_U \perp U$ für alle $x \in V$, genannt die orthogonale Projektion auf U, und dieser ist geben durch

$$x \mapsto \sum_{i=1}^{k} \langle x, x_i\rangle x_i \tag{1}$$

Beweis. Sei zunächst pr_U durch (1) gegeben. Die Linearität von pr_U folgt aus (S1) und (S3). Für $u = \sum_{i=1}^k \lambda_i x_i \in U$ ist $\langle u, x_j\rangle = \left\langle \sum_{i=1}^k \lambda_i x_i, x_j\right\rangle = \sum_{i=1}^k \lambda_i \langle x_i, x_j\rangle = \lambda_j$, woraus $\mathrm{pr}_U(u) = u$. Somit ist $\mathrm{pr}_U|_U = \mathrm{id}_U$, und insbesondere ist pr_U surjektiv. Ist $\mathrm{pr}_U(x) = 0$, so ist $\langle x, x_i\rangle = 0$ für alle $i.$, woraus mit (S2) und (S4) sofort $x \perp U$ folgt. Somit ist $\mathrm{Ker}(\mathrm{pr}_U) \perp U$.
Für $x \in V$ ist $\mathrm{pr}_U(x - \mathrm{pr}_U(x)) = \mathrm{pr}_U(x) - \mathrm{pr}_U(\mathrm{pr}_U(x)) = \mathrm{pr}_U(x) - \mathrm{pr}_U(x) = 0$, also $x - \mathrm{pr}_U(x) \in \mathrm{Ker}(\mathrm{pr}_U) \subseteq U^\perp$.
Ist $f : V \to U$ ein weiterer Epimorphismus mit $f|_U = \mathrm{id}_U$ und $\mathrm{Ker}(f) \perp U$, so ist

$$\underbrace{\mathrm{pr}_U(x)}_{\in U} - \underbrace{f(x)}_{\in U} = \underbrace{\mathrm{pr}_U(x) - x}_{\in U^\perp} - \underbrace{f(x) - x}_{\in U^\perp} \in U \cap U^\perp = \{0\}$$

für jedes $x \in V$, somit $f = \mathrm{pr}_U$. □

Theorem 4.9 (Gram-Schmidt-Verfahren)
Ist $(x_1, ..., x_n)$ eine Basis von V und $k \leq n$ mit $(x_1, ..., x_k)$ orthonormal, so gibt es eine Orthonormalbasis $(y_1, ..., y_n)$ von V mit $y_i = x_i$ für $i = 1, ..., k$ und $\mathrm{span}_K(y_1, ..., y_l) = \mathrm{span}_K(x_1, ..., x_l)$ für $l = 1, ..., n$.

Beweis. Induktion nach $d = n - k$.
$\underline{d = 0}$: nichts zu zeigen
$\underline{d - 1 \to d}$: Für $i \neq k+1$ definiere $y_I = x_i$. Sei $U = \mathrm{span}_K(x_1, ..., x_k)$, $\tilde{x_{k+1}} = x_{k+1} - \mathrm{pr}_U(x_{k-1})$. Dann ist $\tilde{x_{k+1}} \in \mathrm{Ker}(\mathrm{pr}_U) \subseteq U^\perp$ (vgl. Satz 4.8) und $\mathrm{span}_K(x_1, ..., x_k, \tilde{x_{k+1}}) = \mathrm{span}_K(x_1, ..., x_{k+1})$. Setze $y_{k+1} = \frac{1}{\|\tilde{x_{k+1}}\|}\tilde{x_{k+1}}$. Dann ist $(y_1, ..., y_n)$ eine Basis von V mit $(y_1, ..., y_{k+1})$ orthonormal (vgl. Lemma 4.7). Nach Induktionshypothese gibt es eine Orthonormalbasis von V, die das Gewünschte leistet. □

Folgerung 4.10
Jeder endlichdimensionale euklidische bzw. unitäre Vektorraum V besitzt eine Orthonormalbasis.

Beweis. Wähle irgendeine Basis von V und wende Theorem 4.9 mit $k = 0$ an. □

Folgerung 4.11
Ist U ein Untervektorraum von V, so ist $V = U \oplus U^\perp$ und $(U^\perp)^\perp = U$.

Beweis. Wähle eine Orthonormalbasis von U (vgl. Folgerung 4.10), $B = (x_1, ..., x_k)$ und ergänze diese zu einer Orthonormalbasis $(x_1, ..., x_n)$ von V (vgl. Theorem 4.9). Dann sind $x_{k+1}, ..., x_n \in U \perp$, da $U \cap U^\perp = \{0\}$ ist somit $V = U \oplus U^\perp$. Insbesondere ist $\dim_K(U^\perp) = n - \dim_K(U)$, woraus $\dim_K((U^\perp)^\perp) = \dim_K(U)$ folgt. Zusammen mit der trivialen Inklusion $U \subseteq (U^\perp)^\perp$ folgt $U = (U^\perp)^\perp$. □

Folgerung 4.12
Ist s eine positiv definite hermitesche Sesquilinearform auf V und B eine Basis von V, so ist

$$\det(M_B(s)) \in \mathbb{R}_{>0}$$

Beweis. Wähle eine Orthonormalbasis B' von V bezüglich s. Dann ist $M_{B'}(s) = \mathbb{1}_n$, folglich

$$\begin{aligned}
\det(M_B(s)) &= \det\left((T^B_{B'})^t \cdot \mathbb{1}_n \cdot \overline{T^B_{B'}}\right) \\
&= \det\left((T^B_{B'})^t\right) \cdot \det\left(\overline{T^B_{B'}}\right) \\
&= \det\left(T^B_{B'}\right) \cdot \overline{\det\left(T^B_{B'}\right)} \\
&= |\det\left(T^B_{B'}\right)|^2 \\
&> 0
\end{aligned}$$

□

5. Orthogonale und unitäre Endomorphismen

Sei V ein euklidischer bzw. unitärer Vektorraum und $f \in \mathrm{End}_K(V)$.

Definition 5.1 (orthogonale, unitäre Endomorphismen)
f ist orthogonal bzw. unitär , wenn

$$\langle f(x), f(y)\rangle = \langle x, y\rangle \quad \forall x, y \in V$$

Satz 5.2
Ist f unitär, so gelten

- Für $x \in V$ ist $\|f(x)\| = \|x\|$.
- Sind $x, y \in V$ mit $x \perp y$, so ist $f(x) \perp f(y)$.
- Es ist $f \in \mathrm{Aut}_K(V)$ und auch f^{-1} ist unitär.
- Das Bild einer Orthonormalbasis unter f ist eine Orthonormalbasis.
- Ist λ ein Eigenwert von f, so ist $|\lambda| = 1$.

Beweis.

- klar
- klar
- $f(x) = 0 \iff \|f(x)\| = 0 \iff \|x\| = 0 \iff x = 0$, also ist f injektiv, somit $f \in \mathrm{Aut}_K(V)$ und

$$\left\langle f^{-1}(x), f^{-1}(y)\right\rangle \overset{f \text{ unitär}}{=} \left\langle f(f^{-1}(x)), f(f^{-1}(y))\right\rangle = \langle x, y\rangle$$

- Folgt aus 1, 2 und 3
- Ist $f(x) = \lambda x$, $x \neq 0$, so ist

$$\|x\| = \|f(x)\| = \|\lambda x\| = |\lambda| \cdot \|x\| \Rightarrow |\lambda| = 1$$

□

Satz 5.3
Ist $\|f(x)\| = \|x\|$ für alle $x \in V$, so ist f unitär.

Beweis. Aus $\|f(x)\| = \|x\|$ folgt $\langle f(x), f(x)\rangle = \langle x, x\rangle$. Die Polarisierung (Satz 3.4) für $\langle f(x), f(y)\rangle$ und die Linearität von f liefern $\langle f(x), f(y)\rangle = \langle x, y\rangle$. Zum Beispiel im Fall $K = \mathbb{R}$:

$$\begin{aligned}\langle f(x), f(y)\rangle &= \frac{1}{2}\left(\left\langle \underbrace{f(x)+f(y)}_{f(x+y)}, \underbrace{f(x)+f(y)}_{f(x+y)}\right\rangle - \langle f(x), f(x)\rangle - \langle f(y), f(y)\rangle\right)\\ &= \frac{1}{2}\left(\langle x+y, x+y\rangle - \langle x, x\rangle - \langle y, y\rangle\right)\\ &= \langle x, y\rangle\end{aligned}$$

□

Definition 5.4 (orthogonale, unitäre Matrizen)
Eine Matrix $A \in \mathrm{Mat}_n(K)$ heißt orthogonal bzw. unitär , wenn

$$A^*A = \mathbb{1}_n$$

Mathematica/WolframAlpha-Befehle (orthogonale bzw. unitäre Matrizen)
Auch für orthogonale bzw. unitäre Matrizen A gibt es eine Mathematica bzw. WolframAlpha-Funktion

```
OrthogonalMatrixQ[A]
UnitaryMatrixQ[A]
```

▶ **Bemerkung 5.5**
Offenbar ist A genau dann unitär, wenn A^* das Inverse zu A ist. Die folgenden Bedingungen sind daher äquivalent dazu, dass A unitär ist:

$$AA^* = \mathbb{1}_n, \overline{A}A^t = \mathbb{1}_n, A^t\overline{A} = \mathbb{1}_n, A^t = \overline{A^{-1}}$$

Satz 5.6
Sei B eine Orthogonalbasis von V. Genau dann ist f unitär, wenn $M_B(f)$ unitär ist.

Beweis. Sei $A = M_B(f)$, $v = \Phi_B(x)$, $\Phi_B(y)$. Dann ist $\langle v, w\rangle = x^t \underbrace{M_B(\langle\cdot,\cdot\rangle)}_{=\mathbb{1}} \cdot \overline{y} = x^t \cdot \overline{y}$. Somit ist f genau dann unitär, wenn $(Ax)^t\overline{Ay} = x^t\overline{y}$ für alle $x, y \in K^n$, also wenn $A^t\overline{A} = \mathbb{1}$, d.h. A unitär. □

Satz 5.7
Die folgenden Mengen bilden Untergruppen der $\mathrm{GL}_n(K)$.

- $\mathrm{O}_n = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid A \text{ ist orthogonal}\}$ die orthogonale Gruppe
- $\mathrm{SO}_n = \{A \in \mathrm{O}_n \mid \det(A) = 1\}$ die spezielle orthogonale Gruppe
- $\mathrm{U}_n = \{A \in \mathrm{GL}_n(\mathbb{C}) \mid A \text{ ist unitär}\}$ die unitäre Gruppe
- $\mathrm{SU}_n = \{A \in \mathrm{U}_n \mid \det(A) = 1\}$ die spezielle unitäre Gruppe

Beweis. z.B. für U_n: Sind $A^{-1} = A^*$, $B^{-1} = B^*$, so ist $(AB)^{-1} = B^{-1}A^{-1} = B^*A^* = (AB)^*$, $(A^{-1})^{-1} = A = (A^*)^{-1} = (A^{-1})^*$ □

Satz 5.8
Genau dann ist $A \in \mathrm{Mat}_n(K)$ unitär, wenn die Spalten (oder die Zeilen) von A eine Orthonormalbasis des K^n bilden.

Beweis. Sei s das Standardskalarprodukt und $B = (a_1, ..., a_n)$. Nach Bemerkung 4.5 ist B genau dann eine Orthonormalbasis, wenn $M_B(s) = \mathbb{1}_n$, und $M_B(s) = A^t \cdot \mathbb{1}_n \cdot \overline{A}$, vgl. Beispiel 2.9 □

Theorem 5.9
Sei $K = \mathbb{C}$ und $f \in \mathrm{End}_K(V)$. Ist f unitär, so besitzt V eine Orthonormalbasis aus Eigenvektoren von f.

Beweis. Induktion über$n = \dim_K(V)$.
$n = 0$: klar
$n - 1 \to n$: Da K algebraisch abgeschlossen ist, hat χ_f eine Nullstelle λ, es gibt also einen Eigenvektor x_1 von f zum Eigenwert λ. Ohne Einschränkung nehmen wir $\|x\| = 1$ an. Sei $W = K \cdot x_1$. Nach Folgerung 4.11 ist dann $V = W \oplus W^\perp$. Für $v \in W^\perp, w \in W$ ist

$$0 = \langle v, w\rangle = \langle f(v), f(w)\rangle = \overline{\lambda}\,\langle f(v), w\rangle$$

da $\lambda \neq 0$ (f unitär) also $f(W^\perp) \perp W$. Somit ist $f(W^\perp) \subseteq W^\perp$, d.h. $W^\perp$ ist f-invariant. Da auch $f|_{W^\perp}$ unitär ist, gibt es nach Induktionshypothese eine Orthonormalbasis $(x_1, ..., x_n)$ aus Eigenvektoren von $f|_{W^\perp}$. Da $V = W \oplus W^\perp$ und $W \perp W^\perp$ ist $(x_1, ..., x_n)$ eine Orthonormalbasis von V aus Eigenvektoren von f. □

Folgerung 5.10
Jeder unitäre Endomorphismus eines unitären Vektorraums ist diagonalisierbar.

Folgerung 5.11
Zu jeder $A \in \mathrm{U}_n$ gibt es $S \in \mathrm{U}_n$ so, dass

$$S^*AS = S^{-1}AS = \mathrm{diag}(\lambda_1, ..., \lambda_n)$$

mit $|\lambda_i| = 1$ für $i = 1, ..., n$.

Beweis. Da A unitär ist, ist $f_A \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n)$ unitär, nach Theorem 5.9 existiert also eine Orthonormalbasis B des $\mathbb{C}^n$ aus Eigenvektoren von A. Die Transformationsmatrix $S = T_{\mathcal{E}}^{B}$ hat als Spalten die Elemente von B und somit ist S nach Satz 5.8 unitär. Nach Satz 5.2 ist $|\lambda| = 1$ für alle Eigenwerte von f_A. □

▶ **Bemerkung 5.12**
Dies (Theorem 5.9) gilt nicht im Fall $K = \mathbb{R}$. Man kann aber auch orthogonale Endomorphismen immer “fast diagonalisieren“.

6. Selbstadjungierte Endomorphismen

Sei V ein euklidischer bzw. unitärer Vektorraum und $f \in \mathrm{End}_K(V)$.

Definition 6.1 (selbstadjungiert)
f ist selbstadjungiert , wenn

$$\langle f(x), y\rangle = \langle x, f(y)\rangle \quad \forall x, y \in V$$

Satz 6.2
Sei B eine Orthonormalbasis von V. Genau dann ist f selbstadjungiert, wenn $M_B(f)$ hermitesch ist.

Beweis. Seien $A = M_B(f), v = \Phi_B(x), w = \Phi_B(y)$. Es ist

$$\langle f(v), w\rangle = (Ax)^t\overline{y} = x^tA^t\overline{y}$$
$$\langle v, f(w)\rangle = x^t\overline{Ay} = x^t\overline{A}\overline{y}$$

Somit ist $\langle f(v), w\rangle = \langle v, f(w)\rangle$ genau dann, wenn $A^t = \overline{A}$, d.h. $A = A^*$, also A hermitesch. □

Lemma 6.3
Ist f selbstadjungiert und λ ein Eigenwert von f, so ist $\lambda \in \mathbb{R}$.

Beweis. Ist $0 \neq x \in V$ mit $f(x) = \lambda x$, so ist

$$\lambda\langle x, x\rangle = \langle f(x), x\rangle = \langle x, f(x)\rangle = \overline{\lambda}\langle x, x\rangle$$

und mit $\langle x, x\rangle \neq 0$ folgt $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$. □

Satz 6.4
Ist f selbstadjungiert, so ist $\chi_f \in \mathbb{R}[t]$ und χ_f zerfällt über $\mathbb{R}$ in Linearfaktoren.

Beweis. Sei B eine Orthonormalbasis von V. Nach Satz 6.2 ist $A = M_B(f) \in \mathrm{Mat}_n(K) \subseteq \mathrm{Mat}_n(\mathbb{C})$ hermitesch. Da $\mathbb{C}$ algebraisch abgeschlossen ist, ist $\chi_f(t)\prod_{i=1}^n(t-\lambda_i)$ mit $\lambda_1, ..., \lambda_n \in \mathbb{C}$. Nach Lemma 6.3 ist aber schon $\lambda_1, ..., \lambda_n \in \mathbb{R}$. Somit zerfällt $\chi_f\chi_A \in \mathbb{R}[t]$ über $\mathbb{R}$ in Linearfaktoren. □

Theorem 6.5
Ist f selbstadjungiert, so besitzt V eine Orthonormalbasis aus Eigenvektoren von f.

Beweis. Induktion über $n = \dim_K(V)$.
$n = 0$: klar
$n - 1 \to n$: Nach Satz 6.4 hat f einen reellen Eigenwert $\lambda \in \mathbb{R}$. Wähle $x_1 \in V$ mit $f(x_1) = \lambda x_1$ und $\|x_1\| = 1$. Sei $W = K \cdot x_1$. Für $y \in W^\perp$ ist

$$\langle x_1, f(y)\rangle = \langle f(x_1), y\rangle = \lambda\langle x_1, y\rangle = 0$$

und folglich ist $W^\perp$ f-invariant. Nach Folgerung 4.11 ist $V = W \oplus W^\perp$ und $f|_{W^\perp}$ ist wieder selbstadjungiert. Nach Induktionshypothese hat $W^\perp$ eine Orthonormalbasis $(x_1, ..., x_n)$ aus Eigenvektoren von $f|_{W^\perp}$. Da $V = W \oplus W^\perp$ und $W \perp W^\perp$ ist $(x_1, ..., x_n)$ eine Orthonormalbasis von V aus Eigenvektoren von f. □

Folgerung 6.6
Jeder selbstadjungierte Endomorphismus eines euklidischen oder unitären Vektorraums ist diagonalisierbar.

Folgerung 6.7
Ist

- f selbstadjungiert ($K = \mathbb{C}$ oder $\mathbb{R}$)
- f unitär ($K = \mathbb{C}$)

so ist

$$V = \bigoplus_{\lambda \in K} \mathrm{Eig}(f, \lambda)$$

eine Zerlegung von V in paarweise orthogonale Untervektorräume.

Beweis. Nach Theorem 5.9 bzw. Theorem 6.5 existiert eine Orthonormalbasis B aus Eigenvektoren. Insbesondere ist f diagonalisierbar, also

$$V = \bigoplus_{\lambda \in K} \mathrm{Eig}(f, \lambda)$$

Zu jedem λ gibt es eine Teilfamilie von B die eine Basis von $\mathrm{Eig}(f, \lambda)$ bildet. Da B eine Orthonormalbasis ist, folgt, dass die Eigenräume paarweise orthogonal sind. □

▶ **Bemerkung 6.8**
Um eine Orthonormalbasis aus Eigenvektoren wie in Theorem 5.9 oder Theorem 6.5 zu bestimmen, kann man entweder wie im Induktionsbeweis vorgehen, oder man bestimmt zunächst Basen B von $\mathrm{Eig}(f, \lambda_i)$, $i = 1, ..., n$ und orthonormalisiert diese mit Theorem 4.9 zu Basen B'. Nach Folgerung 6.7 ist $\bigcup B'$ dann eine Orthonormalbasis von V aus Eigenvektoren von f.

7. Hauptachsentransformation

Sei V ein euklidischer bzw. unitärer Vektorraum und s eine hermitesche Sesquilinearform auf V.

Satz 7.1
Zu $A \in \mathrm{Mat}_n(K)$ hermitesch gibt es $S \in \mathrm{U}_n(K)$ so, dass

$$S^* AS = S^{-1} AS = \mathrm{diag}(\lambda_1, ..., \lambda_n)$$

mit $\lambda_1, ..., \lambda_n \in \mathbb{R}$.

Beweis. Da A hermitesch ist, ist $f_A \in \mathrm{End}_K(K^n)$ selbstadjungiert, es gibt also nach Theorem 6.5 also eine Orthonormalbasis $B = (x_1, ..., x_n)$ aus Eigenvektoren von f_A. Die Transformationsmatrix $S = T_{\mathcal{E}}^B$ hat $x_1, ..., x_n$ als Spalten und ist somit nach Satz 5.8 unitär. Nach Lemma 6.3 sind die Eigenvektoren $\lambda_1, ..., \lambda_n$ reell. □

Folgerung 7.2
Sei $A \in \mathrm{Mat}_n(K)$ hermitesch. Genau dann ist A positiv definit, wenn alle Eigenwerte positiv sind.

Beweis. Nach Satz 7.1 existiert $S \in \mathrm{U}_n(K)$ mit

$$S^* AS = S^{-1} AS = D = \mathrm{diag}(\lambda_1, ..., \lambda_n) \quad \lambda_1, ..., \lambda_n \in \mathbb{R}$$

Die Eigenwerte von A sind die Eigenwerte von $S^{-1}AS$ (LAAG 1.5.1.11) , also $\lambda_1, ..., \lambda_n$. Sei $T = \overline{S}$. Genau dann ist A positiv definit, wenn $T^t A \overline{T} = S^* AS = D$ positiv definit ist (Satz 2.8), also wenn $\lambda_i > 0$. □

Theorem 7.3 (Hauptachsentransformation)
Zu jeder hermiteschen Sesquilinearform s auf V gibt es eine Orthonormalbasis B von V, für die

$$M_B(s) = \mathrm{diag}(\lambda_1, ..., \lambda_n) \quad \lambda_1, ..., \lambda_n \in \mathbb{R}$$

Beweis. Sei $B_0 = (x_1, ..., x_n)$ eine Orthonormalbasis von V und $A = M_{B_0}(s)$. Da s hermitesch ist, ist auch A hermitesch (Satz 2.13). Nach Satz 7.1 gibt es deshalb $S \in \mathrm{U}_n(K)$ mit $S^* AS = D$ eine reelle Diagonalmatrix. Ist nun $f \in \mathrm{End}_K(V)$ mit $M_{B_0}(f) = \overline{S}$, so ist auch $B = (f(x_1), ..., f(x_n))$ eine Basis von V mit $T_{B_0}^B = \overline{S}$ unitär. Da $M_{B_0}(f)$ unitär ist, ist auch f unitär. Nach Satz 5.2 ist $f(B_0) = B$ somit auch eine Orthonormalbasis. Nach Satz 2.8 ist

$$M_B(s) = (T_{B_0}^B)^t \cdot M_{B_0}(s) \cdot \overline{T_{B_0}^B} = S^* AS = D$$ □

■ **Beispiel 7.4**
$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $s = s_A$, $K = \mathbb{R}$, $V = \mathbb{R}^2$
$\Rightarrow q_s(x) = 2x_1^2 + 2x_1x_2 + 2x_2^2$
Wie verhält sich $q_s : \mathbb{R}^2 \to \mathbb{R}$? Wie sehen die "Höhenlinien"

$$H_c = \{x \in \mathbb{R}^2 \mid q_s(x) = c\} \quad c \in \mathbb{R}$$

aus?

$$\chi_A = (t-2)^2 - 1 = (t-1)(t-3) \Rightarrow \lambda_1 = 3, \lambda_2 = 1$$
$$\Rightarrow B = \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right)$$
$$\Rightarrow M_B(s) = \operatorname{diag}(3,1)$$

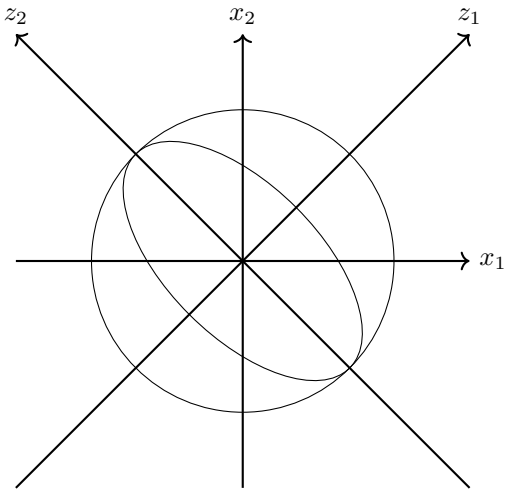
Im neuen Koordinatensystem $z = \Phi_B^{-1}(x)$ ist dann

$$q_s(z) = 3z_1^2 + z_2^2$$

Mit $a_1 = \frac{1}{\sqrt{3}}$, $a_2 = 1$ erhält man "Höhenlinien" der Form

$$\left(\frac{z_1}{a_1}\right)^2 + \left(\frac{z_2}{a_2}\right)^2 = c$$

was für $c > 0$ eine Ellipse beschreibt.



Folgerung 7.5
Zu jeder hermiteschen Sesquilinearform s auf V gibt es eine Basis B von V, für die

$$M_B(s) = \begin{pmatrix} \mathbb{1}_{\mathbb{r}_+(\mathbb{s})} & & \\ & -\mathbb{1}_{\mathbb{r}_-(\mathbb{s})} & \\ & & 0 \end{pmatrix}$$

mit $r_+(s) + r_-(s) \leq n$.

Beweis. Sei $B_0 = (x_1, ..., x_n)$ eine Orthonormalbasis von V mit $A = M_{B_0}(s) = \operatorname{diag}(\lambda_1, ..., \lambda_n)$. Setze

$$\mu_i = \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} & \lambda_i \neq 0 \\ 1 & \lambda_i = 0 \end{cases}$$

Sei $x_i' = \mu_i \cdot x_i$ und $B' = (x_1', ..., x_n')$. Dann ist $M_B(s) = S^t A \overline{S}$ mit $S = T_{B_0}^{B'} = \text{diag}(\mu_1, ..., \mu_n)$ also $M_{B'}(s) = \text{diag}(\lambda_1', ..., \lambda_n')$ mit $\lambda_i' = \mu_i \cdot \lambda_i \cdot \overline{\mu_i} = \mu_i^2 \lambda_i \in \{0, 1, -1\}$. Durch Permutation der Elemente von B' erhält man die gewünschte Basis B. □

Definition 7.6 (Ausartungsraum)
Der Ausartungsraum von s ist

$$V_0 = \{x \in V \mid s(x, y) = 0 \quad \forall y \in V\}$$

Lemma 7.7
V_0 ist ein Untervektorraum von V.

Beweis. Klar aus Linearität im ersten Argument. □

Lemma 7.8
Seien V_+ und V_- Untervektorräume von V mit $V = V_+ \oplus V_- \oplus V_0$ und s positiv definit auf V_+, $-s$ positiv definit auf V_-. Dann ist

$$\dim_K(V_+) = \max\{\dim_K(W) \mid \text{Untervektorraum von } V, s \text{ positiv definit auf } V\}$$
$$\dim_K(V_-) = \max\{\dim_K(W) \mid \text{Untervektorraum von } V, -s \text{ positiv definit auf } V\}$$

Beweis. Beweis nur für V_+, analog für V_-.
$\leq$: klar
$\geq$: Ist $W \leq V$ Untervektorraum mit $s(x, x) > 0 \quad \forall x \in W \backslash \{0\}$, so ist $W \cap (V_- \oplus V_+) = \{0\}$. Ist $x = y + z$ mit $y \in V_-$, $z \in V_0$, so ist $s(x, x) = s(y + z, y + z) = \underbrace{s(y, y)}_{\leq 0} + \underbrace{s(y, z) + s(z, y) + s(z, z)}_{=0} \leq 0 \Rightarrow \dim_K(W) \leq \dim_K(V) - \dim_K(V_-) - \dim_K(V_0) = \dim_K(V_+)$. □

Theorem 7.9 (Trägheitssatz von Sylvester)
Für eine hermitesche Sesquilinearform s auf V sind die Zahlen $r_+(s)$, $r_-(s)$ aus Folgerung 7.5 eindeutig bestimmt.

Beweis. Sei B eine Basis von V wie in Folgerung 7.5, $B = (x_1, ..., x_n)$. Definiere

$$V_+ = \text{span}_K(x_1, ..., x_{r_+(s)})$$
$$V_- = \text{span}_K(x_{r_+(s)+1}, ..., x_{r_+(s)+r_-(s)})$$
$$V_0' = \text{span}_K(x_{r_+(s)+r_-(s)+1}, ..., x_n)$$

Dann ist s positiv definit auf V_+, $-s$ positiv definit auf V_- und $V = V_+ \oplus V_- \oplus V_0'$. Es gilt $V_0' = V_0$
$\subseteq$: klar
$\supseteq$: Ist $x = \sum_{i=1}^n \lambda_i x_i \in V_0$, so ist $0 = s(x, x_i) = \lambda_i \cdot s(x_i, x_i)$ für $i = 1, ..., n$ also $\lambda_i = 0$ für $i = 1, ..., r_+(s) + r_-(s)$, d.h. $x \in V_0'$. Nach Lemma 7.8 ist $r_+(s) = \dim_K(V_+)$ nur von s abhängig, analog für $r_-(s)$. □

Definition 7.10 (Signatur)
Die Signatur von s ist das Tripel

$$(r_+(s), r_-(s), r_0(s))$$

wobei $r_0(s) = \dim_K(V_0)$.

Folgerung 7.11
Ist s eine hermitesche Form auf V und B eine Basis von V, so ist die Zahl der positiven bzw. negativen Eigenwerte von $M_B(s)$ gleich $r_+(s)$ bzw. $r_-(s)$, insbesondere also unabhängig von B.

Beweis. Sei $A = M_B(s)$. Nach Satz 7.1 gibt es $S \in \mathrm{U}_n(K)$ mit S^*AS eine reelle Diagonalmatrix. Da $S^* = S^{-1}$ haben A und S^*AS die selben Eigenwerte. Bringt man S^*AS nun in die Form in Folgerung 7.5, so ändern sich die Vorzeichen der Diagonale nicht mehr. □

8. Quadriken

Sei $n \in \mathbb{N}$.

Definition 8.1 (Quadrik)
Eine Quadrik ist eine Teilmenge von $\mathbb{R}^n$ mit

$$Q = \{x \in \mathbb{R}^n \mid x^t A x + 2b^t x + c = 0\}$$

mit $A \in \mathrm{Mat}_n(\mathbb{R})$ symmetrisch, $b^t \in \mathbb{R}^n$ und $c \in \mathbb{R}$.

▶ **Bemerkung 8.2**

- $Q = \{x \in \mathbb{R}^n \mid \sum_{i,j=1}^n a_{ij} x_i y_j + 2\sum_{i=1}^n b_i x_i + c = 0\}$ also Q ist die Nullstellenmenge eines quadratischen Polynoms in $x_1, ..., x_n$
- Q bestimmt A, b, c nicht eindeutig, da $Q(A, b, c) = Q(\lambda A, \lambda b, \lambda c)$
- Man kann A, b, c so normieren, dass $c = 0$ oder $c = 1$

▶ **Bemerkung 8.3**
Seien A, b, c wie in Definition 8.1, so schreiben wir

$$\tilde{A} = \begin{pmatrix} A & b \\ b^t & c \end{pmatrix}$$
$$\tilde{x} = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

Dann ist $Q = \{x \in \mathbb{R}^n \mid \tilde{x}^t \tilde{A} \tilde{x} = 0\}$. Wir schreiben (A, b) für

$$\begin{pmatrix} A & b \end{pmatrix} \in \mathrm{Mat}_{n,n+1}(\mathbb{R})$$

Es gilt $\mathrm{rk}(A) \leq \mathrm{rk}(A, b) \leq \mathrm{rk}(\tilde{A})$.

▶ **Bemerkung 8.4 (Wiederholung)**
Seien V, W K-Vektorräume. $f : V \to W$ heißt affin, wenn $\exists g \in \mathrm{Hom}_K(V, W)$ mit $f(v) = g(v) + w_0$ $\forall v \in V$. Ist f affin und bijektiv, so ist f^{-1} affin, d.h. $\mathrm{Aff}_K(V) = \{f : V \to V \mid f \text{ affin und bijektiv}\}$.
Im Fall von $V = \mathbb{R}^n$, $K = \mathbb{R}$ ist

$$\mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^n) = \{f = \tau_z \circ f_T \mid T \in \mathrm{GL}_n(\mathbb{R}), z \in \mathbb{R}^n\}$$

mit $f_T(x) = Tx$ und $\tau_z(x) = x + z$.

Lemma 8.5
Ist $Q \subseteq \mathbb{R}^n$ eine Quadrik, so ist $f(Q)$ eine Quadrik, für $f \in \mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^n)$.

Beweis. $f = \tau_z \circ f_T$ mit $T \in \mathrm{GL}_n(\mathbb{R})$ und $z \in \mathbb{R}^n$. Schreibe $S = T^{-1} \in \mathrm{GL}_n(\mathbb{R})$, $\tilde{S} = \begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix}$. Es gilt $\tilde{S}\tilde{x} = \widetilde{Sx}$.

$$\begin{aligned} f_T(Q) &= \{Tx \in \mathbb{R}^n \mid \tilde{x}^t \tilde{A} \tilde{x} = 0\} \\ &= \{y \in \mathbb{R}^n \mid (\tilde{S}\tilde{y})^t \tilde{A} \tilde{S} \tilde{y} = 0\} \\ &= \{y \in \mathbb{R}^n \mid \tilde{y}^t \underbrace{\tilde{S}^t \tilde{A} \tilde{S}}_{\begin{pmatrix} S^tAS & S^tb \\ b^tS & c \end{pmatrix}} \tilde{y} = 0\} \end{aligned}$$

Jetzt für τ_z. Sei $U_z = \begin{pmatrix} \mathbb{1} & z \\ 0 & 1 \end{pmatrix}$. $U_z\tilde{x} = \tilde{\tau}_z(x)$. Man folgert analog, dass

$$\tau_z(Q) = \{y \in \mathbb{R}^n \mid \tilde{y}^t \underbrace{U_z^t \tilde{A} U_z}_{\begin{pmatrix} A & Az+b \\ z^tA+b & z^tAz+b^tz+z^tb+c \end{pmatrix}} \tilde{y} = 0\}$$

□

Definition 8.6 (Typen von Quadriken)
Sei Q gegeben durch (A, b, c) wie in Definition 8.1. Q heißt

- vom kegeligen Typ , wenn $\mathrm{rk}(A) = \mathrm{rk}(A, b) = \mathrm{rk}(\tilde{A})$
- eine Mittelpunktsquadrik , wenn $\mathrm{rk}(A) = \mathrm{rk}(A, b) < \mathrm{rk}(\tilde{A})$
- vom parabolischen Typ , wenn $\mathrm{rk}(A) < \mathrm{rk}(A, b)$
- ausgeartet , wenn $\det(\tilde{A}) = 0$

Lemma 8.7
Ist $Q \subseteq \mathbb{R}^n$ eine Quadrik, $f \in \mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^n)$. Von dem Typ, von dem Q ist, ist auch $f(Q)$.

Beweis. $f = f_{S^{-1}}$, $S \in \mathrm{GL}_n(\mathbb{R})$. Da $\tilde{S}$ invertierbar ist, ist $\mathrm{rk}(\tilde{A}) = \mathrm{rk}(\tilde{S}^t\tilde{A}\tilde{S})$, analog auch $\mathrm{rk}(S^tAS) = \mathrm{rk}(A)$. $(S^tAS, S^tb) = S^t(A, b)\begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \mathrm{rk}(S^tAS, S^tb) = \mathrm{rk}(A, b)$. Für $f = \tau_z$ analog. □

Definition 8.8 (Isometrie)
Eine Isometrie des $\mathbb{R}^n$ ist $f \in \mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^n)$ mit

$$f(x) = Ax + b$$

mit $b \in \mathbb{R}^n$ und $A \in \mathrm{GL}_n(\mathbb{R})$ ist orthogonal.

▶ **Bemerkung 8.9**
$f : \mathbb{R}^n \to \mathbb{R}^n$ ist eine Isometrie genau dann, wenn $\|f(x) - f(y)\| = \|x - y\|$ für alle $x, y \in \mathbb{R}^n$.

Theorem 8.10 (Klassifikation der Quadriken bis auf Isometrien)
Sei Q eine Quadrik. Es gibt eine Isometrie $f \in \mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^n)$ mit $f(Q)$, die eine der folgenden Formen annimmt:

- $f(Q) = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{k} \left(\frac{x_i}{a_i}\right)^2 - \sum_{i=k+1}^{n} \left(\frac{x_i}{a_i}\right)^2 = 0 \right\} \quad k \geq r-k$
- $f(Q) = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{k} \left(\frac{x_i}{a_i}\right)^2 - \sum_{i=k+1}^{n} \left(\frac{x_i}{a_i}\right)^2 = 1 \right\}$
- $f(Q) = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{k} \left(\frac{x_i}{a_i}\right)^2 - \sum_{i=k+1}^{n} \left(\frac{x_i}{a_i}\right)^2 - 2x_{r+1} = 0 \right\} \quad k \geq r-k, r < n$

mit $a_1, ..., a_r \in \mathbb{R}_{>0}$ und $0 \leq k \leq r \leq n$

Beweis. Sei Q gegeben durch (A, b, c). Nach Satz 7.1 gibt es eine orthogonale Matrix $S \in \mathrm{O}_n$ mit $S^t SAS = \mathrm{diag}(\lambda_1, ..., \lambda_n)$. Indem wir Q durch $f_{S^{-1}}(Q)$ ersetzen, können wir also ohne Einschränkung annehmen, dass $A = \mathrm{diag}(\lambda_1, ..., \lambda_n)$. Ohne Einschränkung ist weiter $\lambda_1, ..., \lambda_k > 0$ und $\lambda_{k+1}, ..., \lambda_r < 0$ und $\lambda_{r+1}, ..., \lambda_n = 0$. Dann ist $(e_{r+1}, ..., e_n)$ eine Orthonormalbasis des Ausartungsraums V_0 von s_A.
Wenn wir Q durch $\tau_z(Q)$ ersetzen, wird b durch $Az+b$ ersetzt, wir können deshalb ohne Einschränkung annehmen, dass $b \in V_0$. Ist $n > r$, also $V_0 \neq \{0\}$, so können wir eine Orthonormalbasis $(v_{r+1}, ..., v_n)$ von V_0 mit $b \in \mathrm{span}_{\mathbb{R}}(v_{r+1})$ wählen.
Indem wir Q durch $f_{S^{-1}}(Q)$ mit $S = (e_1, ..., e_r, v_{r+1}, ..., v_n)$ ersetzen, können wir ohne Einschränkung annehmen, dass $b = \mu \cdot e_{r+1}$ mit $\mu \in \mathbb{R}$.
Ist nun $\mathrm{rk}(A) = \mathrm{rk}(A, b)$, so gibt es z mit $Az = -b$, und indem wir Q durch $\tau_z(Q)$ ersetzen, können wir annehmen, dass $b = 0$.

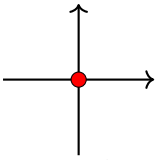
- Im Fall $c = 0$ setzt man $a_i = \frac{1}{\sqrt{|\lambda_i|}}$ und ersetzt gegebenenfalls (A, b, c) mit $(-A, -b, -c)$, um Form 1 zu erhalten.
- Im Fall $c \neq 0$ ersetzt man (A, b, c) durch $(-\frac{1}{c}A, -\frac{1}{c}b, -1)$ und setzt dann $a_i = \frac{1}{\sqrt{|\lambda_i|}}$, um Form 2 zu erhalten.
- Ist $\mathrm{rk}(A) < \mathrm{rk}(A, b)$, so ist insbesondere $r < n$ und $\mu \neq 0$. Nun ersetzten wir Q durch $\tau_z(Q)$ mit $z = -\frac{c}{2\mu} \cdot e_{r+1}$ und können somit auch wieder $c = 0$ annehmen. Ersetzt man $(A, b, 0)$ durch $(-\frac{1}{\mu}A, -1, 0)$ und setzt wieder $a_i = \frac{1}{\sqrt{|\lambda_i|}}$, so erhält man Form 3. (Ist $k < r - k$, so ersetzt man weiter Q durch $f_{-\mathbb{1}_n}(Q)$ und $(A, b, 0)$ durch $(-A, -b, 0)$.) □

Folgerung 8.11
Sei $Q \subseteq \mathbb{R}^n$ eine Quadrik. Es gibt eine invertierbare affine Abbildung $f \in \mathrm{Aff}_{\mathbb{R}}(\mathbb{R}^n)$ für die $f(Q)$ eine der folgenden 3 Formen annimmt:

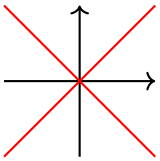
- $f(Q) = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{k} x_i^2 - \sum_{i=k+1}^{r} x_i^2 = 0 \right\} \quad k \geq r-k$
- $f(Q) = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{k} x_i^2 - \sum_{i=k+1}^{r} x_i^2 = 1 \right\}$
- $f(Q) = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^{k} x_i^2 - \sum_{i=k+1}^{r} x_i^2 - 2x_{r+1} = 0 \right\} \quad k \geq r-k, r < n$

Beispiel 8.12
$Q \subseteq \mathbb{R}^2$

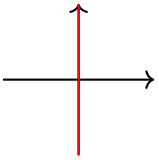
- _ $k = 2, r = 2 : \left\{ x \in \mathbb{R}^2 \mid \left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 = 0 \right\}$



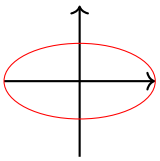
_ $k = 1, r = 2 : \left\{ x \in \mathbb{R}^2 \mid \left(\frac{x_1}{a_1}\right)^2 - \left(\frac{x_2}{a_2}\right)^2 = 0 \right\}$



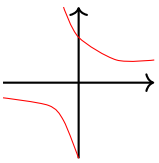
_ $k = 1, r = 1 : \left\{ x \in \mathbb{R}^2 \mid \left(\frac{x_1}{a_1}\right)^2 = 0 \right\}$



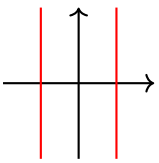
- _ $k = 2, r = 2 : \left\{ x \in \mathbb{R}^2 \mid \left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 = 1 \right\}$



_ $k = 1, r = 2 : \left\{ x \in \mathbb{R}^2 \mid \left(\frac{x_1}{a_1}\right)^2 - \left(\frac{x_2}{a_2}\right)^2 = 1 \right\}$



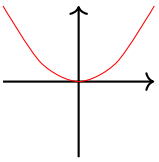
_ $k = 1, r = 1 : \left\{ x \in \mathbb{R}^2 \mid \left(\frac{x_1}{a_1}\right)^2 = 1 \right\}$



_ $k = 0, r = 2 : \left\{ x \in \mathbb{R}^2 \mid -\left(\frac{x_1}{a_1}\right)^2 - \left(\frac{x_2}{a_2}\right)^2 = 1 \right\} = \emptyset$

_ $k = 0, r = 1 : \left\{ x \in \mathbb{R}^2 \mid -\left(\frac{x_1}{a_1}\right)^2 - \left(\frac{x_2}{a_2}\right)^2 = 1 \right\} = \emptyset$

- _ $k = 1, r = 1 : \left\{ x \in \mathbb{R}^2 \mid \left(\frac{x_1}{a_1}\right)^2 - 2x_2 = 0 \right\}$



▶ **Bemerkung 8.13**

- Ist $Q \subseteq \mathbb{R}^2$ eine Quadrik, $U \subseteq V$ affiner Untervektorraum, so ist $Q \cap U$ eine Quadrik in dem Sinne, dass $\exists f$ Isometrie : $f(U) = \mathbb{R}^k$ und $f(Q \cap U)$ ist eine Quadrik.
- Ebene Quadriken sind im wesentlichen Kegelschnitte, $Q' = \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 = x_3^2\}$, außer 2c und 2d in Beispiel 8.12

▶ **Bemerkung 8.14**

Die Situation wird deutlich übersichtlicher, wenn man den affinen Raum $\mathbb{R}^n$ durch Hinzunahme von Punkten im Unendlichen zum projektiven Raum $\mathbb{P}^n(\mathbb{R})$ vervollständigt und den Abschluss der Quadriken darin betrachtet. Es stellt sich dann heraus, dass vom projektiven Standpunkt aus die meisten ebenen Quadriken ähnlich aussehen. (Siehe Vorlesung *Elementare Algebraische Geometrie*)

Kapitel VII

Dualität

1. Das Lemma von Zorn

Sei K ein Körper und U, V, W seien K-Vektorräume. Zudem sei X eine Menge.

Definition 1.1 (Relation)
Eine Relation ist eine Teilmenge $R \subseteq X \times X$. Man schreibt $(x, x') \in R$ als xRx'. R heißt

- reflexiv , wenn $\forall x \in X$: xRx
- transitiv , wenn $\forall x, y, z \in X$: xRy und $yRz \Rightarrow xRz$
- symmetrisch , wenn $\forall x, y \in X$: $xRy \Rightarrow yRx$
- antisymmetrisch , wenn $\forall x, y \in X$: xRy und $yRx \Rightarrow y = x$
- total , wenn $\forall x, y \in X$: $(x, y) \notin R \Rightarrow (y, x) \in R$

■ **Beispiel 1.2 (Äquivalenzrelation)**
Eine Äquivalenzrelation ist eine reflexive, transitive und symmetrische Relation. Wir haben schon verschiedene Äquivalenzrelationen kennengelernt: Isomorphie von K-Vektorräumen und Ähnlichkeit von Matrizen.

Definition 1.3 (Halbordnung)
Eine Halbordnung (oder partielle Ordnung) ist eine reflexive, transitive und antisymmetrische Relation $\leq$. Eine totale Halbordnung heißt Totalordnung oder lineare Ordnung . Man schreibt $x < y$ für $x \leq y \wedge x \neq y$.

■ **Beispiel 1.4**

1. Die natürliche Ordnung $\leq$ auf $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$ und $\mathbb{N}$ ist eine Z Totalordnung.
2. Teilbarkeit $|$ ist eine Halbordnung auf $\mathbb{N}$, aber Teilbarkeit ist keine Halbordnung auf $\mathbb{Z}$, da $1|-1$ und $-1|1$, aber $1 \neq -1$!
3. $\mathcal{P}(X)$ ist die Potenzmenge. "$\subseteq$" ist eine Halbordnung auf $\mathcal{P}$, aber für $|X| > 1$ ist "$\subseteq$" keine Totalordnung.
4. Sei $(X, \leq)$ eine Halbordnung, sei $Y \subseteq X$, so ist $(Y, \subseteq |_Y)$ eine Halbordnung.

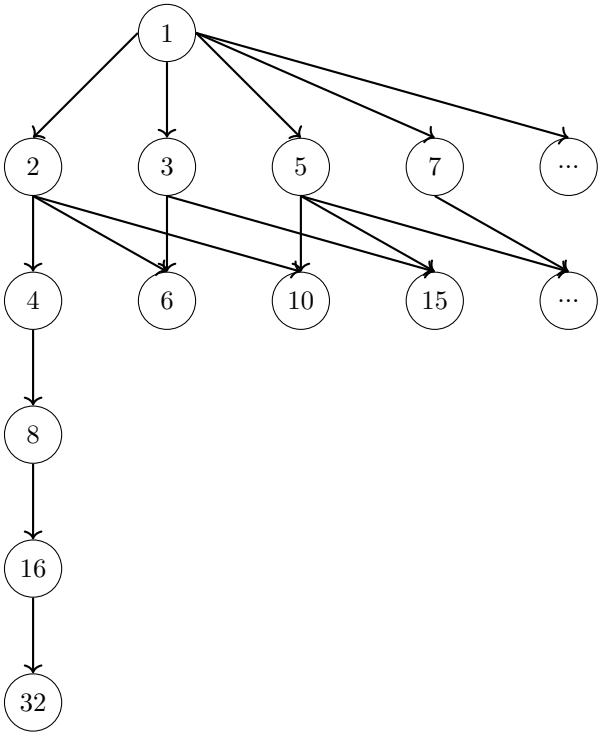
Definition 1.5 (Kette)
Sei $(X, \leq)$ eine Halbordnung, $Y \subseteq X$. Y heißt Kette , wenn $(Y, \leq |_Y)$ total ist.

$x \in Y$ heißt ein minimales Element von Y, wenn $\forall x' \in Y$: $x < x'$.

$x \in Y$ heißt untere Schranke von Y, wenn $\forall y \in Y$: $y \geq x$.

$x \in Y$ heißt kleinstes Element von Y, wenn x untere Schranke von Y ist.

Analog: maximales Element , obere Schranke , größtes Element .



$Y = \{2^n \mid n \in \mathbb{N}\}$ ist eine Kette

▶ **Bemerkung 1.6**

- Hat Y ein kleinstes Element, so ist dies eindeutig bestimmt. Ein kleinstes Element ist minimal.
- Jede endliche Halbordnung hat minimale Elemente. Jede endliche Totalordnung hat ein kleinstes Element. Analog für maximale Elemente und größtes Element.

■ **Beispiel 1.7**
$(\mathbb{N}, \leq)$ hat als kleinstes Element die 1, aber kein größtes Element oder maximale Elemente.

■ **Beispiel 1.8**
$V = \mathbb{R}^3$, $\mathfrak{X}$ die Menge der Untervektorräume des $\mathbb{R}^3$. $(\mathfrak{X}, \leq)$ ist eine Halbordnung auf $Y \subseteq X$ mit $Y = \{U \in \mathfrak{X} \mid \dim_{\mathbb{R}}(U) \leq 2\}$.

- Y hat ein kleinstes Element: $\{0\}$.
- Es gibt unendlich viele maximale Elemente in Y, nämlich die Untervektorräume von V, die die Dimension 2 haben. Es gibt also kein größtes Element.
- V ist die obere Schranke von Y.

Theorem 1.9 (Das Lemma von Zorn)
Sei $(X, \leq)$ eine Halbordnung, die nicht leer ist. Wenn jede Kette eine obere Schranke hat, dann hat X ein maximales Element.

Beweis. Das Lemma von Zorn hat axiomatischen Charakter - es ist äquivalent zum Auswahlaxiom, seine Gültigkeit ist somit abhängig von unseren grundlegenden mengentheoretischen Annahmen. Für einen Beweis des Lemmas von Zorn aus dem Auswahlaxiom siehe die Vorlesung *Mengenlehre*. Wir zeigen hier zumindest die andere Richtung, nämlich dass das Auswahlaxiom aus dem Lemma von Zorn folgt. □

Folgerung 1.10 (Auswahlaxiom)
Zu jeder Familie (x_i), nicht leer, gibt es eine Auswahlfunktion , das heißt eine Abbildung:

$$f : I \to \bigcup_{i \in I} X_i \text{ mit } f(i) \in X_i \quad \forall i$$

Beweis. Sei $\mathcal{F}$ die Menge der Paare (J, f) bestehend aus einer Teilmenge $J \subseteq I$ und einer Abbildung $f : I \to \bigcup_{i \in I} X_i$ mit $f(i) \in X_i \quad \forall i \in J$. Definieren wir $(J, f) \leq (J', f') \iff J \subseteq J'$ und $f'|_J = f$, so ist $\leq$ eine Halbordnung auf $\mathcal{F}$. Da $(\emptyset, \emptyset) \in \mathcal{F}$ ist $\mathcal{F}$ nichtleer. Ist $\mathcal{G} \subseteq \mathcal{F}$ eine nichtleere Kette, so wird auf $J' := \bigcup_{(J,f) \in \mathcal{G}} J$ durch $f'(j) = f(j)$ falls $(J, f) \in \mathcal{G}$ und $j \in J$ eine wohldefinierte Abbildung $f' : J \to \bigcup_{i \in J} X_i$ mit $f'(i) \in X_i \quad \forall i \in J'$ gegeben. Das Paar (J', f') ist eine obere Schranke der Kette $\mathcal{G}$. Nach dem Lemma von Zorn besitzt $\mathcal{F}$ ein maximales Element (J, f). Wir behaupten, dass $J = I$. Andernfalls nehmen wir ein $i' \in I \backslash J$ und ein $x' \in X_{i'}$ und definieren $J' := U \cup \{i'\}$ und $f' : J' \to \bigcup_{i \in J'} X_i$, $j \mapsto \begin{cases} f(j) & j \in J \\ x' & j = i' \end{cases}$. Dann ist $(J', f') \in \mathcal{F}$ und $(J, f) < (J', f')$ im Widerspruch zur Maximalität von (J, f). □

Folgerung 1.11 (Basisergänzungssatz)
Sei V ein K-Vektorraum. Jede linear unabhängige Teilmenge $X_0 \subseteq V$ ist in einer Basis von V enthalten.

Beweis. Sei $\mathfrak{X} = \{X \subseteq V \mid X \text{ ist linear unabhängig, } X_0 \subseteq X\}$ geordnet durch Inklusion. Dann ist $X_0 \in \mathfrak{X}$, also $\mathfrak{X} \neq \emptyset$. Ist $\mathcal{Y}$ eine nichtleere Kette in $\mathfrak{X}$, so ist auch $Y = \bigcup \mathcal{Y} \subseteq V$ linear unabhängig. Sind $y_1, ..., y_n \in Y$ paarweise verschieden, so gibt es $Y_1, ..., Y_n \in \mathcal{Y}$ mit $y_i \in Y_i$ für $i = 1, ..., n$. Da $\mathcal{Y}$ total geordnet ist, besitzt $\{Y_1, ..., Y_n\}$ ein größtes Element, o.E. Y_1. Also sind $y_1, ..., y_n \in Y_1$ und somit linear unabhängig. Folglich ist $Y_1 \in \mathfrak{X}$ eine obere Schranke von $\mathcal{Y}$. Nach dem Lemma von Zorn besitzt $\mathfrak{X}$ ein maximales Element X. Das heißt, X ist eine maximal linear unabhängige Teilmenge von V, nach LAAG1 II.3.5 also eine Basis von V. □

2. Der Dualraum

Sei V ein K-Vektorraum.

Definition 2.1 (Dualraum)
Der Dualraum zu V ist der K-Vektorraum

$$V^* = \mathrm{Hom}_K(V, K) = \{\varphi : V \to K \text{ linear}\}$$

Die Elemente von V^* heißen Linearformen auf V.

■ **Beispiel 2.2**
Ist $V = K^n = \mathrm{Mat}_{n\times 1}(K)$, so wird $V^* = \mathrm{Hom}_K(V, K)$ durch $\mathrm{Mat}_{1\times n}(K) \cong K^n$. Wir können also die Elemente von V als Spaltenvektoren und die Linearformen auf V als Zeilenvektoren auffassen.

Lemma 2.3
Ist $B(x_1)_{i\in I}$ eine Basis von V, so gibt es zu jedem $i \in I$ genau $x_i^* \in V^*$ mit

$$x_i^*(x_j) = \delta_{ij} \quad \forall j \in I$$

Beweis. Siehe LAAG1 III.5.1, angewandt auf die Familie $(y_j)_{j\in I}$, $y_j\delta_{i.j}$ in $W = K$. □

Satz 2.4
Ist $B = (x_1)_{i\in I}$ eine Basis von V, so ist $B^* = (x_i^*)_{i\in I}$ linear unabhängig. Ist I endlich, so ist B^* eine Basis von V^*.

Beweis. Ist $\varphi = \sum_{i\in I} \lambda_i x_i^*$, $\lambda_i \in K$, fast alle gleich 0, so ist $\varphi(x_j) = \sum_{i\in I} \lambda_j x_i^*(x_j) = \lambda_j$ für jedes $j \in I$. Ist also $\varphi = 0$, so ist $\lambda_j = \varphi(x_j) = 0 \quad \forall j \in I$, B^* ist somit linear unabhängig.
Ist zudem I endlich und $\psi \in V^*$, so ist $\psi = \psi' = \sum_{i\in I} \psi(x_i)x_i^*$, denn $\psi'(x_j) = \sum_{i\in I} \psi(x_i)x_i^*(x_j) = \psi(x_i) \quad \forall j \in I$, und somit ist B^* ein Erzeugendensystem von V^*. □

Definition 2.5 (duale Basis)
Ist $B = (x_i)_{i\in I}$ eine endliche Basis von V, so nennt man $B^* = (x_i^*)_{i\in I}$ die zu B duale Basis .

Folgerung 2.6
Zu jeder Basis B von V gibt es einen eindeutig bestimmtem Monomorphismus

$$f_V \to V^* \text{ mit } f(B) = B^*$$

Ist $\dim_K(V) < \infty$, so ist dieser ein Isomorphismus.

Folgerung 2.7
Zu jedem $= 0 \neq x \in V$ gibt es eine Linearform $\varphi \in V$ mit $\varphi(x) = 1$.

Beweis. Ergänze $x_1 = x$ zu einer Basis $(x_i)_{i\in I}$ von V (Folgerung 1.11) und $\varphi = x_1^*$. □

■ **Beispiel 2.8**
Ist $V = K^n$ mit Standardbasis $\mathcal{E} = (e_1, ..., e_n)$, so können wir V^* mit dem Vektorraum der Zeilen-

vektoren identifizieren, und dann ist

$$e_i^* = e_i^t$$

Definition 2.9 (Bidualraum)
Der Bidualraum zu V ist der K-Vektorraum

$$V^{**} = (V^*)^* = \mathrm{Hom}_K(V^*, K)$$

Satz 2.10
Die kanonische Abbildung

$$\iota : \begin{cases} V \to V^{**} \\ x \to \iota_x \end{cases} \quad \text{wobei } \iota_x(\varphi) = \varphi(x)$$

ist ein Monomorphismus. Ist $\dim_K(V) < \infty$, so ist ι ein Isomorphismus.

Beweis.
- $\iota_x \in V^{**}$:
  - $\iota_x(\varphi + \psi) = (\varphi + \psi)(x) = \varphi(x) + \psi(x) = \iota_x(\varphi) + \iota_x(\psi)$
  - $\iota_x(\lambda\varphi) = (\lambda\varphi)(x) = \lambda\varphi(x) = \lambda\iota_x(\varphi)$
- ι linear:
  - $\iota_{x+y}(\varphi) = \varphi(x+y) = \varphi(x) + \varphi(y) = \iota_x(\varphi) + \iota_y(\varphi) = (\iota_x + \iota_y)(\varphi)$
  - $\iota_{\lambda x}(\varphi) = \varphi(\lambda x) = \lambda\iota_x(\iota) = (\lambda\iota_x)(\varphi)$
- ι injektiv: Sei $0 \neq x \in V$. Nach Folgerung 2.7 existiert $\varphi \in V^*$ mit $\iota_x(\varphi) = \varphi(x) = 1 \neq 0$. Somit ist $\iota_x \neq 0$.
- Ist $\dim_K(V) < \infty$, so ist $V \overset{2.6}{\cong} V^* \overset{2.6}{\cong} V^{**}$, insbesondere $\dim_K(V) = \dim_K(V^{**})$. Der Monomorphismus ι ist somit ein Isomorphismus. □

▶ **Bemerkung 2.11**
Sei $\dim_K(V) < \infty$. Im Gegensatz zu den Isomorphismen $V \to V^*$, die von der Wahl der Basis B abhängen, ist der Isomorphismus $\iota : V \to V^{**}$ kanonisch (von der Wahl der Basis B unabhängig).

Die Voraussetzung, dass $\dim_K(V) < \infty$ ist hier essentiell: Für $\dim_K(V) = \infty$ ist ι nicht surjektiv.

Definition 2.12 (Annulator)
Für eine Teilmenge $U \subseteq V$ bezeichne

$$U^0 = \{\varphi \in V^* \mid \varphi(x) = 0 \quad \forall x \in U\}$$

den Annulator von U.

Lemma 2.13
U^0 ist ein Untervektorraum von V^*.

Beweis. Klar. □

Satz 2.14
Ist $\dim_K(V) < \infty$ und $U \subseteq V$ ein Untervektorraum, so ist

$$\dim_K(V) = \dim_K(U) + \dim_K(U^0)$$

Beweis. Ergänze eine Basis $(x_1, ..., x_r)$ von U zu einer Basis $B = (x_1, ..., x_n)$ von V. Dann ist $B^*(x_1^*, ..., x_n^*)$ eine Basis von V^*. Sei $C = (x_{r+1}^*, ..., x_n^*)$. Dann ist C eine Basis von U^0:

- B^* ist Basis $\Rightarrow C$ ist linear unabhängig.
- $C \subseteq U^0$: Für $1 \leq j \leq r < i \leq n$ ist $x_i^*(x_j) = \delta_{ij} = 0$.
- $U^0 \subseteq \text{span}_K(C)$: Ist $\varphi = \sum_{i=1}^n \lambda_i x_i^* \in U^0$, so $0 = \varphi(x_j) = \lambda_j$ für alle $j \leq r$, also $\varphi \in \text{span}_K(x_{r+1}^*, ..., x_n^*)$. □

Folgerung 2.15
Ist $\dim_K(V) < \infty$ und $U \subset V$ ein Untervektorraum, so ist

$$\iota(U) = U^{00}$$

Beweis. Es ist klar, dass $\iota(U) \leq U^{00}$.
Für $\varphi \in U^0$ und $x \in U$ ist $\iota_x(\varphi) = \varphi(x) = 0$. Mit Satz 2.14 ist

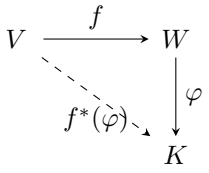
$$\begin{aligned} \dim_K(U^{00}) &= \dim_K(V^*) - \dim_K(U^0) \\ &= \dim_K(V^*) - (\dim_K(V) - \dim_K(U)) \\ &\overset{2.6}{=} \dim_K(U) \end{aligned}$$

und da ι injektiv ist, folgt $\iota(U) = U^{00}$. □

3. Die duale Abbildung

Sei $f \in \mathrm{Hom}_K(V, W)$.

▶ **Bemerkung 3.1**
Ist $\varphi \in W^* = \mathrm{Hom}_K(W, K)$ eine Linearform auf W, so ist $\varphi \circ f \in \mathrm{Hom}_K(V, K) = V^*$ eine Linearform auf V.



Definition 3.2 (duale Abbildung)
Die zu f duale Abbildung ist

$$f^* : \begin{cases} W^* \to V^* \\ \varphi \mapsto \varphi \circ f \end{cases}$$

Lemma 3.3
Es ist $f^* \in \mathrm{Hom}_K(W^*, V^*)$.

Beweis. Sind $\varphi, \psi \in W^*$ und $\lambda \in K$ ist

$$\begin{aligned} f^*(\varphi + \psi) &= (\varphi + \psi) \circ f \\ &= \varphi \circ f + \psi \circ f \\ &= f^*(\varphi) + f^*(\psi) \\ f^*(\lambda\varphi) &= (\lambda\varphi) \circ f \\ &= \lambda \cdot (\varphi \circ f) \\ &= \lambda \cdot f^*(\varphi) \end{aligned}$$

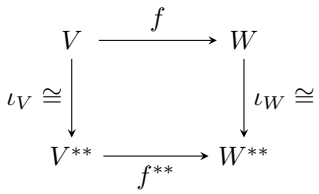
□

Satz 3.4
Sind $B = (x_1, ..., x_n)$ und $C = (y_1, ..., y_m)$ Basen von V bzw. W, so ist

$$M_{B^*}^{C^*}(f^*) = \left(M_C^B(f)\right)^t$$

Beweis. Sei $A = M_C^B(f) = (a_{ij})_{i,j}$ und $B = M_{B^*}^{C^*}(f^*) = (b_{ji})_{j,i}$. Dann ist $f(x_j) = \sum_{i=1}^m a_{ij} y_i$, also $a_{ji} = y_i^*(f(x_j)) = f^*(y_i^*)(x_j)$ und $f^*(y_i^*) = \sum_{j=1}^n b_{ji} x_j^*$, also $b_{ji} = f^*(y_i^*)(x_j) = a_{ij}$. □

Folgerung 3.5
Sind V und W endlichdimensional, und identifizieren wir $V = V^{**}$ und $W = W^{**}$, so ist $f = f^{**}$, das heißt $\iota \circ f = f^{**} \circ \iota$.



Beweis. Seien B und C Basen von V bzw. W. Unter der Identifizierung ist $B^{**} = B$ und $C = C^{**}$, das heißt $\iota(x_i) = x_i^{**}$ bzw. $\iota(y_j) = y_j^{**}$, denn $\iota(x_i)(x_j^*) = x_j^*(x_i) = \delta_{ij} = x_i^{**}(x_j^*) \quad \forall i, j$ und somit

$$M_C^B(f^{**}) \overset{3.4}{=} \left(M_{B^*}^{C^*}(f^*)\right)^t \overset{3.4}{=} \left(M_C^B(f)\right)^{tt} = M_C^B(f)$$

Also $f^{**} = f$. □

Folgerung 3.6
Sind V, W endlichdimensional, so liefert die Abbildung $f \mapsto f^*$ einen Isomorphismus von K-Vektorräumen.

$$\mathrm{Hom}_K(V, W) \to \mathrm{Hom}_K(W^*, V^*)$$

Beweis. Sind $f, g \in \mathrm{Hom}_K(V, W)$ und $\lambda \in K$, $\varphi \in W^*$, so ist

$$\begin{aligned}(f+g)^*(\varphi) &= \varphi \circ (f+g) = \varphi \circ f + \varphi \circ g = f^*(\varphi) + g^*(\varphi) = (f^* + g^*)(\varphi)\\ (\lambda f)^*(\varphi) &= \varphi \circ (\lambda f) = \lambda \cdot (\varphi \circ f) = \lambda \circ f^*(\varphi) = (\lambda f^*)(\varphi)\end{aligned}$$

Die Abbildung ist somit linear. Nach Folgerung 3.5 ist sie injektiv. Da

$$\begin{aligned}\dim_K(V, W) &= \dim_K(V) \cdot \dim_K(W)\\ &= \dim_K(V^*) \cdot \dim_K(W^*)\\ &= \dim_K(\mathrm{Hom}_K(W^*, V^*))\end{aligned}$$

ist sie auch ein Isomorphismus. □

Satz 3.7
Sind V, W endlichdimensional so ist

$$\begin{aligned}\mathrm{Im}(f^*) &= \mathrm{Ker}(f)^0\\ \mathrm{Ker}(f^*) &= \mathrm{Im}(f)^0\end{aligned}$$

Beweis.

- $\mathrm{Im}(f^*) \subseteq \mathrm{Ker}(f)^0$: Ist $\varphi \in W^*$, $x \in \mathrm{Ker}(f)$, so ist

$$f^*(\varphi)(x) = (\varphi \circ f)(x) = \varphi(0) = 0$$

- $\mathrm{Ker}(f)^0 \subseteq \mathrm{Im}(f^*)$: Sei $\varphi \in \mathrm{Ker}(f)^0$. Setze eine Basis $(x_1, ..., x_r)$ von $\mathrm{Ker}(f)$ zu einer Basis $(x_1, ..., x_n)$ von V fort. Dann sind $f(x_{r+1}), ..., f(x_n)$ linear unabhängig nach der Kern-Bild-Formel (LAAG 1 III.7.13), es gibt also $\psi \in W^*$ mit

$$\psi(f(x_i)) = \varphi(x_i) \quad \forall i$$

Es folgt

$$f^*(\psi)(x_i) = \psi(f(x_i)) = \varphi(x_i) \quad \forall i$$

also $\varphi = f^*(\psi)$.

- Mit der Identifizierung $V = V^{**}$ ist

$$\mathrm{Im}(f)^0 \overset{3.5}{=} \mathrm{Im}(f^{**})^0 = \mathrm{Ker}(f^*)^{00} \overset{2.15}{=} \mathrm{Ker}(f^*)$$ □

Folgerung 3.8
Sind V, W endlichdimensional, so ist

$$\mathrm{rk}(f) = \mathrm{rk}(f^*)$$

Beweis.

$$\begin{aligned}\mathrm{rk}(f) &= \dim_K(\mathrm{Im}(f)) \\ &\overset{2.14}{=} \dim_K(W) - \dim_K(\mathrm{Im}(f)^0) \\ &\overset{LAAG1.III.7.13}{=} \dim_K(W^*) - \dim_K(\mathrm{Ker}(f^*)) \\ &= \mathrm{rk}(f^*)\end{aligned}$$ □

Folgerung 3.9
Ist $\dim_K(V) < \infty$ und $U \subseteq V$ ein Untervektorraum, so lässt sich jede Linearform auf U zu einer Linearform auf V fortsetzen.

Beweis. Ist $f : U \to V$ die Inklusionsabbildung, so ist $f^* : V^* \to U^*$, $\varphi \mapsto \varphi|_U$ und

$$\mathrm{rk}(f^*) = \mathrm{rk}(f) = \dim_K(U) = \dim_K(U^*)$$

f^* ist somit surjektiv. □

▶ **Bemerkung 3.10**
Folgerung 3.9 gilt auch ohne die Voraussetzung $\dim_K(V) < \infty$, siehe Übung.

▶ **Bemerkung 3.11**
Ein homogenes lineares Gleichungssystem $Ax = 0$ hat als Lösungsraum $L(A, 0) \subseteq K^n$ ein Untervektorraum des K^n. Unter der Identifizierung $K^n = (K^n)^{**}$ ist $L(A, 0)$ der Annulator der Linearformen beschrieben durch die Zeilen $a_1, ..., a_m \in (K^n)^*$ von A. Wir wollen umgekehrt zu einem Untervektorraum $W \subseteq K^n$ ein $A = (a_1, ..., a_m) \in \mathrm{Mat}_{n\times m}(K)$ mit $W = L(A, 0)$ finden. Ist $W = \mathrm{span}_K(b_1, ..., b_r)$, so ist $W = \mathrm{Im}(f_B)$ mit $B = (b_1, ..., b_r) \in \mathrm{Mat}_{n\times r}(K)$.
$\Rightarrow W \overset{3.7}{=} \mathrm{Ker}(f_B^*)^0$ und $M_{\mathcal{E}^t}(f_B^*) = B^t$. Wenn man also eine Basis $(a_1, ..., a_s)$ von $L(B^t, 0)$ bestimmt und daraus eine Matrix $A = (a_1^t, ..., a_s^t) \in \mathrm{Mat}_{s\times n}(K)$ bildet, so ist $W = L(A, 0)$.

4. Die adjungierte Abbildung

Sei $K = \mathbb{R}$ oder $K = \mathbb{C}$ und V ein endlichdimensionaler unitärer K-Vektorraum.

Definition 4.1 (weitere Skalarmultiplikation)
Wir definieren auf V eine Skalarmultiplikation

$$\lambda * x = \overline{\lambda} \cdot x$$

und schreiben $\overline{V} = (V, +, *)$.

Lemma 4.2
$\overline{V}$ ist ein K-Vektorraum und $\mathrm{End}_K(V) = \mathrm{End}_K(\overline{V})$.

Beweis. Mit LAAG1 VI.1.7 nachprüfen, zum Beispiel:

- $\lambda * (x + y) = \overline{\lambda} \cdot (x + y) = \overline{\lambda}x + \overline{\lambda}y = \lambda * x + \lambda * y$
- $\lambda * (\mu * x) = \overline{\lambda}(\overline{\mu} \cdot x) = \overline{\lambda\mu}x = (\lambda\mu) * x$ □

Weiterhin sei: $f \in \mathrm{End}_K(V)$, $x \in V$, $\lambda \in K$
$\Rightarrow f(\lambda * x) = f(\overline{\lambda}x) = \lambda * f(x)$
$\Rightarrow f \in \mathrm{End}_K(\overline{V})$.
Umgekehrt sei $g \in \mathrm{End}_K(\overline{V})$, $x \in V$, $\lambda \in K$
$\Rightarrow g(\lambda \cdot x) = g(\overline{\lambda} * x) = \lambda \cdot g(x)$
$\Rightarrow g \in \mathrm{End}_K(V)$.

Lemma 4.3
Für $y \in V$ ist

$$\Phi_y : \begin{cases} V \to K \\ x \mapsto \langle x, y \rangle \end{cases}$$

eine Linearform auf V.

Die Abbildung $y \mapsto \Phi_y$ liefert einen Isomorphismus $\Phi : \overline{V} \to V^*$.

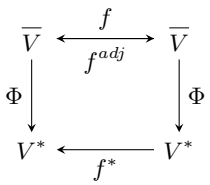
Beweis.
- $\Phi_y \in V^*$: Linearität in ersten Argument.
- $\Phi \in \mathrm{Hom}_K(\overline{V}, V^*)$: Für $y, y' \in V$, $\lambda \in K$, $x \in V$ ist
  - $\Phi_{y+y'}(x) = \langle x, y + y' \rangle = \langle x, y \rangle + \langle x, y' \rangle = \Phi_y(x) + \Phi_{y'}(x)$
  - $\Phi_{\lambda * y}(x) = \langle x, \lambda * x \rangle = \left\langle x, \overline{\lambda} y \right\rangle = \lambda \langle x, y \rangle = \lambda \Phi_y(x)$
- Φ injektiv: Skalarprodukt ist nicht ausgeartet.
- Da $\dim_K(\overline{V}) = \dim_K(V) = \dim_K(V^*)$ ist Φ somit ein Isomorphismus. □

Satz 4.4
Zu $f \in \mathrm{End}_K(V)$ gibt es ein eindeutig bestimmtes $f^{adj} \in \mathrm{End}_K(V)$ mit

$$\langle f(x), y\rangle = \left\langle x, f^{adj}(y)\right\rangle \quad \forall x, y \in V$$

Beweis. Existenz und Eindeutigkeit sind zu zeigen.

- Existenz:



  Für $f^{adj} = \Phi{-}1 \circ f^* \circ \Phi \in \mathrm{End}_K(\overline{V}) = \mathrm{End}_K(V)$ ist

$$\Phi_y \circ = (f^* \circ \Phi)(y) = (\Phi \circ f^{adj\cdot})(y) = \Phi_{f^{adj}(y)}$$

  also

$$\langle f(x), y\rangle = (\Phi_y \circ f)(x) = \Phi_{f^{adj}(y)}(x) = \left\langle x, f^{adj}(y)\right\rangle \quad \forall x, y \in V$$

- Eindeutigkeit: Erfüllen f_1, f_2 für Gleichung

$$\langle f(x), y\rangle = \left\langle x, f^{adj}(y)\right\rangle \quad \forall x, y \in V$$

  so ist

$$0 = \langle x, f_1(y)\rangle - \langle x, f_2(y)\rangle = \langle x, f_1(y), f_2(y)\rangle \quad \forall x, y \in V$$

  da $\langle \cdot, \cdot \rangle$ nicht ausgeartet ist, folgt daraus, dass $f_1 = f_2$. □

Definition 4.5 (adjungierter Endomorphismus)
Die Abbildung f^{adj} heißt der zu f adjungierte Endomorphismus .

Beispiel 4.6

- Ist f selbstadjungiert, so ist $f^{adj} = f$.
- Ist f unitär, so ist $f \in \mathrm{Aut}_K(V)$ und

$$\langle f(x), y\rangle = \left\langle x, f^{-1}(y)\right\rangle \quad \forall x, y \in V$$

  also $f^{adj} = f^{-1}$.

Lemma 4.7
Ist B eine Orthonormalbasis vin V, so ist

$$M_B(f^{adj}) = M_B(f^*)$$

Beweis. Ist $A = M_B(f)$ und $B = M_B(f^{adj})$, $v = \Phi_B(x)$, $w = \Phi_B(y)$, so ist

$$\begin{aligned}(Ax)^t\overline{y} = \langle f(v), w\rangle &= \left\langle v, f^{adj}(w)\right\rangle \\ x^t A^t\overline{y} &= x^t\overline{B}\overline{y} \\ \Rightarrow B = \overline{A^t} &= A^*\end{aligned}$$

□

Lemma 4.8
Für $f, g \in \mathrm{End}_K(V)$ und $\lambda, \mu \in K$ ist

$$\begin{aligned}(\lambda f + \mu g)^{adj} &= \overline{\lambda} f^{adj} + \overline{\mu} g^{adj} \\ (f^{adj})^{adj} &= f\end{aligned}$$

Beweis. Für $x, y \in V$ ist

$$\begin{aligned}\langle(\lambda f + \mu g)(x), y\rangle &= \lambda \langle f(x), y\rangle + \mu \langle g(x), y\rangle \\ &= \lambda \left\langle x, f^{adj}(y)\right\rangle + \mu \left\langle x, g^{adj}\right\rangle \\ &= \left\langle x, (\overline{\lambda} f^{adj} + \overline{\mu} g^{adj})(y)\right\rangle\end{aligned}$$

und

$$\left\langle f^{adj}(x), y\right\rangle = \overline{\langle y, f^{adj}(y)\rangle} = \overline{\langle f(y), x\rangle} = \langle x, f(y)\rangle$$

□

5. Der Spektralsatz

Sei V ein endlichdimensionaler unitärer K-Vektorraum und $f \in \mathrm{End}_K(V)$.

Definition 5.1 (normaler Endomorphismus, normale Matrix)
Der Endomorphismus f heißt normal , wenn

$$f \circ f^{adj} = f^{adj} \circ f$$

Entsprechend heißt $A \in \mathrm{Mat}_n(K)$ normal , wenn

$$AA^* = A^*A$$

Mathematica/WolframAlpha-Befehle (normale Matrix)
Ob eine Matrix A normal ist, beantwortet folgende Funktion für Mathematica bzw. WolframAlpha:

```
NormalMatrixQ[A]
```

■ **Beispiel 5.2**

- Ist f selbstadjungiert, so ist $f^{adj} = f$, insbesondere ist f normal.
- Ist f unitär, so ist $f^{adj} = f^{-1}$, insbesondere ist f normal.

Lemma 5.3
Genau dann ist $f \in \mathrm{End}_K(V)$ normal, wenn

$$\langle f(x), f(y)\rangle = \left\langle f^{adj}(x), f^{adj}(y)\right\rangle \quad \forall x, y \in V$$

Beweis.

- Hinrichtung: Ist f normal, so ist

$$\begin{aligned}\langle f(x), f(y)\rangle &= \left\langle x, (f^{adj} \circ f)(y)\right\rangle \\ &= \left\langle x, (f \circ f^{adj})(y)\right\rangle \\ &= \left\langle f^{adj}(x), f^{adj}(y)\right\rangle \quad \forall x, y \in V\end{aligned}$$

- Rückrichtung: Ist umgekehrt $\left\langle f^{adj}(x), f^{adj}(y)\right\rangle$, so ist

$$\begin{aligned}\left\langle x, (f^{adj} \circ f)(y)\right\rangle &= \left\langle x, (f \circ f^{adj})(y)\right\rangle \\ 0 &= \left\langle x, (f^{adj} \circ f - f \circ f^{adj})(y)\right\rangle \\ f^{adj} \circ f &= f \circ f^{adj}\end{aligned}$$

□

Lemma 5.4
Ist f normal, ist ist

$$\mathrm{Ker}(f) = \mathrm{Ker}(f^{adj})$$

Beweis. Nach Lemma 5.3 ist

$$\|f(x)\| = \|f^{adj}(x)\| \quad \forall x \in V$$

Insbesondere gilt

$$f(x) = 0 \iff f^{adj}(x) = 0$$ □

Lemma 5.5
Ist f normal, so ist

$$\mathrm{Eig}(f, \lambda) = \mathrm{Eig}(f^{adj}, \overline{\lambda}) \quad \forall \lambda \in K$$

Beweis. Da $(\lambda \cdot \mathrm{id} - f)^{adj} \overset{4.8}{=} \overline{\lambda} \cdot \mathrm{id} - f^{adj}$ ist auch $\lambda \cdot \mathrm{id} - f$ normal. Somit ist

$$\begin{aligned}\mathrm{Eig}(f, \lambda) &= \mathrm{Ker}(\lambda\,\mathrm{id} - f) \\ &\overset{5.4}{=} \mathrm{Ker}((\lambda\,\mathrm{id} - f)^{adj}) \\ &= \mathrm{Ker}(\overline{\lambda}\,\mathrm{id} - f^{adj}) \\ &= \mathrm{Eig}(f^{adj}, \overline{\lambda})\end{aligned}$$ □

Theorem 5.6 (Spektralsatz)
Sei $f \in \mathrm{End}_K(V)$ ein Endomorphismus, für den χ_f in Linearfaktoren zerfällt. Genau dann besitzt V eine Orthonormalbasis aus Eigenvektoren von f, wenn f normal ist.

Beweis.

- Hinrichtung: Ist B eine Orthonormalbasis aus Eigenvektoren von f, so ist $A = M_B(f)$ eine Diagonalmatrix. Dann ist auch $M_B(f^{adj}) \overset{4.7}{=} A^*$ eine Diagonalmatrix und $AA^* = A^*A$. Somit ist f normal.
- Rückrichtung: Sei f normal und $\chi_f(t) = \prod_{i=1}^{n}(t - \lambda_i)$. Beweis nach Induktion nach $n = \dim_K(V)$.
  $\underline{n = 0}$: klar
  $\underline{n - 1 \to n}$: Wähle Eigenvektor zum Eigenwert λ_1, o.E. $\|x_1\| = 1$. Sei $U = K \cdot x_1$. Nach Lemma 5.5 ist $f^{adj}(x_1) = \overline{\lambda_1} x_1$, insbesondere ist U f-invariant und f^{adj}-invariant. Für $x \in U^\perp$ ist

  $$\langle f(x), x_1 \rangle = \left\langle x, f^{adj}(x_1) \right\rangle = \left\langle x, \overline{\lambda_1} x_1 \right\rangle = \lambda_1 \langle x, x_1 \rangle = 0$$

  also $f(x) \in U^\perp$ und

  $$\left\langle f^{adj}(x), x_1 \right\rangle = \langle x, f(x_1) \rangle = \langle x, \lambda_1 x_1 \rangle = \overline{\lambda_1} \langle x, x_1 \rangle = 0$$

  also $f^{adj}(x) \in U^\perp$. Somit ist $V = U \oplus U^\perp$ eine Zerlegung in Untervektorräume, die sowohl f-invariant als auch f^{adj}-invariant sind. Insbesondere st $f^{adj}|_{U^\perp} = (f|_{U^\perp})^{adj}$, woraus folgt, dass auch $f|_{U^\perp}$ normal ist:

  $$f|_{U^\perp} \circ (f|_{U^\perp})^{adj} = f \circ f^{adj}|_{U^\perp} = f^{adj} \circ f|_{U^\perp} = f^{adj}|_{U^\perp} \circ f|_{U^\perp} = (f|_{U^\perp})^{adj} \circ f|_{U^\perp}$$

  Außerdem zerfällt auch $\chi_{f|_{U^\perp}} = \prod_{i=2}^{n}(t - \lambda_i)$ in Linearfaktoren. Nach Induktionshypothese existiert eine Orthonormalbasis $(x_2, ..., x_n)$ von $U^\perp$ bestehend aus Eigenvektoren von $f|_{U^\perp}$ und $(x_1, ..., x_n)$ ist dann eine Orthonormalbasis von V aus Eigenvektoren von f. □

Folgerung 5.7
Sei $A \in \mathrm{Mat}_n(\mathbb{C})$. Genau dann gibt es $S \in \mathrm{U}_n$ mit $S^*AS = D$ eine Diagonalmatrix, wenn A normal ist.

▶ **Bemerkung 5.8**
Theorem 5.6 ist eine gemeinsame Verallgemeinerung von Theorem VI.5.9 und Theorem VI.6.5

6. Tensorprodukte

Definition 6.1 (billineare Abbildung)
Eine Abbildung $\xi : V \times W \to U$ ist bilinear , wenn für jedes $v \in V$ die Abbildung

$$\begin{cases} W \to U \\ w \mapsto \xi(v, w) \end{cases}$$

und für jedes $w \in W$ die Abbildung

$$\begin{cases} V \to U \\ v \mapsto \xi(v, w) \end{cases}$$

linear sind.

Wir definieren

$$\mathrm{Bil}_K(V, W, U) = \{\xi \in \mathrm{Abb}(V \times W, U) \mid \xi \text{ bilinear}\}$$

■ **Beispiel 6.2**
Seien $V = W = K[t]_{\leq d}$, $U = K[t]_{\leq 2d}$. Die Abbildung

$$\xi : \begin{cases} V \times W \to U \\ (f.g) \mapsto fg \end{cases} \quad \text{ist bilinear}$$

Wir sehen, dass $\mathrm{Im}(\xi)$ im Allgemeinen kein Untervektorraum von U ist. Ist zum Beispiel $K = \mathbb{Q}$, $d = 1$, so liegen $t^2 = \xi(t, t)$ und $-2 = \xi(-2, 1)$ im $\mathrm{Im}(\xi)$ nicht jedoch $t^2 - 2$, denn wäre $t^2 - 2 = fg$ mit $f, g \in \mathbb{Q}[t]$ linear, so hätte $t^2 - 2$ eine Nullstelle in $\mathbb{Q}$, aber $\sqrt{2} \notin \mathbb{Q}$.

Lemma 6.3
$\mathrm{Bil}_K(V, W, U)$ bildet einen Untervektorraum des K-Vektorraum $\mathrm{Abb}(V \times W, U)$.

Beweis. klar, zum Beispiel

$$(\xi + \xi')(\lambda v, w) = \xi(\lambda v, w) + \xi'(\lambda v, w) = \lambda\xi(v, w) + \lambda\xi(v, w) = \lambda(\xi + \xi')(v, w)$$ □

Lemma 6.4
Ist $\xi \in \mathrm{Bil}_K(V, W, U)$ und $f \in \mathrm{Hom}_K(U, U')$ für einen K-Vektorraum, so ist

$$f \circ \xi \in \mathrm{Bil}_K(V, W, U')$$

Beweis. klar, zum Beispiel

$$(f \circ \xi)(\lambda v, w) = f(\xi(\lambda v, w)) = f(\lambda\xi(v, w)) = \lambda \cdot (f \circ \xi)(v, w)$$ □

Lemma 6.5
Sei $(v_i)_{i\in I}$ eine Basis von V und $(w_j)_{j\in J}$ eine Basis von W. Zu jeder Familie $(u_{ij})_{(i,j)\in I\times J}$ in U gibt es genau ein $\xi \in \mathrm{Bil}_K(V,W,U)$ mit

$$\xi(v_i, w_i) = u_{ij} \quad \forall i \in I, j \in J$$

Beweis.

- Eindeutigkeit: Ist ξ bilinear, $v = \sum_{i\in I} \lambda_i v_i$, $w = \sum_{j\in J} \mu_j w_j$ so ist

$$\begin{aligned}\xi(v,w) &= \xi\left(\sum_{i\in I}\lambda_i v_i, \sum_{j\in J}\mu_j w_j\right)\\ &= \sum_{i\in I}\lambda_i \xi\left(v_i, \sum_{j\in J}\mu_j w_j\right)\\ &= \sum_{i,j}\lambda_i\mu_j u_{ij} \qquad (1)\end{aligned}$$

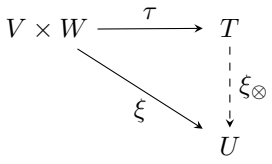
  durch die Familie $(u_{ij})_{i,j}$ bestimmt.

- Existenz: Wird ξ durch (1) definiert, so ist ξ bilinear: Für festes $w = \sum_{j\in J}\mu_j w_j$ ist

$$\begin{cases} V & \to U \\ v = \sum_{i\in I}\lambda_i v_i & \mapsto \xi(v,w) = \sum_{i\in I}\lambda_i\left(\sum_{j\in J}\mu_j u_{ij}\right)\end{cases}$$

  linear (LAAG1 III.5.1), analog für festes v. □

Definition 6.6 (Tensorprodukt)
Ein Tensorprodukt von V und W ist ein Paar (T,τ) bestehend aus einem K-Vektorraum T und einer bilinearen Abbildung $\tau \in \mathrm{Bil}_K(V,W,T)$ welche die folgende universelle Eigenschaft erfüllt: *Ist U ein weiterer K-Vektorraum und $\xi \in \mathrm{Bil}_K(V,W,U)$ so gibt es genau ein $\xi_\otimes \in \mathrm{Hom}_K(T,U)$ mit $\xi = \xi_\otimes \circ \tau$.*



Anmerkung
Sind V und W zwei Vektorräume und K ein gemeinsamer Körper, so kann man das Tensorprodukt $V \otimes W$, was auch ein Vektorraum ist, wie folgt konstruieren: Wenn $B = (b_1, ..., b_n)$ eine Basis von V und $C = (c_1, ..., c_m)$ eine Basis von W ist, dass ist $V \otimes W$ ein Vektorraum, genannt *Tensorproduktraum*, in dem es eine Basis gibt, die auf eindeutige Weise mit den geordneten Paaren des kartesischen Produkts

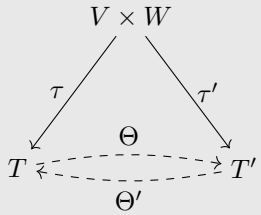
$$B \times C = \{(b_i, c_j)\}$$

der Basen der Ausgangsräume identifiziert werden kann. Die Dimension von $V \otimes W$ ist dann das Produkt der Dimensionen von V und W. Ein Element der Basis von $V \otimes W$, das dem Paar (b_i, c_j) entspricht, wird als $b_i \otimes c_j$ notiert, das $\otimes$ hat also keine tiefere Bedeutung. Ein Element des Tensorproduktes $V \otimes W$ hat dann die Gestalt:

$$\sum_{i,j} \lambda_{ij} \cdot (b_i \otimes c_j)$$

mit $\lambda_{ij} \in K$.

Lemma 6.7
Sind (T, τ) und (T', τ') Tensorprodukte von V und W, so gibt es einen eindeutig bestimmten Isomorphismus $\Theta : T \to T'$ mit $\tau' = \Theta \circ \tau$.



Beweis. Da (T, τ) die universelle Eigenschaft erfüllt, gibt es ein eindeutig bestimmtes $\Theta = (\tau')_{\otimes} \in \mathrm{Hom}_K(T, T')$ mit $\tau' = \Theta \circ \tau$. Analog gibt es $\Theta' \in \mathrm{Hom}_K(T', T)$ mit $\tau = \Theta' \circ \tau'$. Es folgt, dass $\tau = \Theta' \circ \tau' = \Theta' \circ \Theta \circ \tau$. Da auch $\tau = \mathrm{id}_T \circ \tau$ liefert die Eindeutigkeitsaussage in der universellen Eigenschaft von (T, τ), für $U = T$, $\xi = \tau$, dass $\Theta \circ \Theta' = \mathrm{id}_T$. Analog sieht man, dass $\Theta \circ \Theta' = \mathrm{id}_{T'}$. Somit ist Θ ein Isomorphismus. □

Definition 6.8 (Vektorraum mit Basis X)
Sei X eine Menge. Der K-Vektorraum mit Basis X ist der Untervektorraum $V = \mathrm{span}_K((\delta_x)_{x \in X})$ des K-Vektorraum $\mathrm{Abb}(X, K)$ mit $\delta_x(y) = \delta_{x,y} = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$

Lemma 6.9
Sei X eine Menge und V der K-Vektorraum mit Basis X. Dann ist V ein K-Vektorraum und $(\delta_x)_{x \in X}$ ist eine Basis von V.

Beweis. Zu zeigen ist nur, dass $(\delta_x)_{x \in X}$ linear unabhängig ist. Ist $f = \sum_{x \in X} \lambda_x \delta_x$, $\lambda_x \in K$, fast alle gleich 0, und $f = 0$, so ist $\lambda_x = f(x) = 0$ für jedes $x \in X$. □

Lemma 6.10
Sei $(v_i)_{i\in I}$ eine Basis von V und $(w_j)_{j\in J}$ eine Basis von W. Sei T der K-Vektorraum mit der Basis $I \times J$ (im Sinne von Definition 6.8) und $\tau : V \times W \to T$ die bilineare Abbildung gegeben durch $(v_i, w_j) \mapsto \delta_{i,j}$, vergleiche Lemma 6.5. Dann ist (T, τ) ein Tensorprodukt von V und W.

Beweis. Wir schreiben $v_i \otimes w_j$ für $\delta_{i,j}$. Sei U ein weiterer K-Vektorraum und $\xi \in \mathrm{Bil}_K(V, W, U)$. Da $(v_i \otimes w_j)_{(ij)\in I\times J}$ eine Basis von T ist, gibt es genau ein $\xi_\otimes \in \mathrm{Hom}_K(T, U)$ mit $\xi_\otimes(v_i \otimes w_j) = \xi(v_i, w_j)$ für alle i, j, also mit $\xi_\otimes \circ \tau = \xi$ nach Lemma 6.5. Die universelle Eigenschaft ist somit erfüllt. □

Satz 6.11
Es gibt ein bis auf Isomorphie (im Sinne von Lemma 6.7) eindeutig bestimmtes Tensorprodukt

$$(V \otimes_K W, \otimes)$$

von V und W. Sind V und W endlichdimensional, so ist

$$\dim_K(V \otimes_K W) = \dim_K(V) \cdot \dim_K(W)$$

Beweis. Lemma 6.10 und Lemma 6.7 □

■ **Beispiel 6.12**
Durch die Wahl der Standardbasis erhält man einen kanonischen Isomorphismus $K^m \otimes_K K^n \cong \mathrm{Mat}_{m\times n}(K)$.

■ **Beispiel 6.13**
Ist V ein $\mathbb{R}$-Vektorraum mit Basis $(x_1, ..., x_n)$, so ist $\mathbb{C} \otimes_\mathbb{R} V$ ein $\mathbb{R}$-Vektorraum der Dimension $2n$ mit Basis $(1 \otimes x_1 ..., 1 \otimes x_n, i \otimes x_1, ..., i \otimes x_n)$. Durch $\lambda \cdot z \otimes x = (\lambda z) \otimes x$ für $\lambda, z \in \mathbb{C}$, $x \in V$ wird $\mathbb{C} \otimes_\mathbb{R} V$ zu einem $\mathbb{C}$-Vektorraum der Dimension $1 \otimes x_1, ..., 1 \otimes x_n$, $V_\mathbb{C}$, genannt die Komplexifizierung von V.

Satz 6.14
Sei $V \otimes_K W$ ein Tensorprodukt von V und W. Für jeden weiteren K-Vektorraum U liefert die Abbildung $\xi \to \xi_\otimes$ ein Isomorphismus

$$\mathrm{Bil}_K(V, W, U) \xrightarrow{\cong} \mathrm{Hom}_K(V \otimes_K W, U)$$

Beweis. Diese Abbildung heiße Λ.

- Λ ist linear: klar aus Eindeutigkeitsaussage, z.B.

$$(\xi_\otimes + \xi'_\otimes) \circ \otimes = \xi_\otimes \circ \otimes + \xi'_\otimes \circ \otimes = \xi + \xi' = (\xi + \xi')_\otimes \circ \otimes$$

  und somit $\xi_\otimes + \xi'_\otimes = (\xi + \xi')_\otimes$.
- Λ ist injektiv: Ist $\xi \neq 0$, so wegen $\xi = \xi_\otimes \circ \otimes$ auch $\xi_\otimes \neq 0$.
- Λ ist surjektiv: Ist $f \in \mathrm{Hom}_K(V \otimes_K W, U)$, so ist $\xi = f \circ \otimes$ bilinear, die universelle Eigenschaft liefert somit $f = \xi_\otimes \in \mathrm{Im}(\Lambda)$. □

Folgerung 6.15
Sind V und W endlichdimensional, so ist

$$V \otimes_K W \cong \mathrm{Bil}_K(V, W, K)^*$$

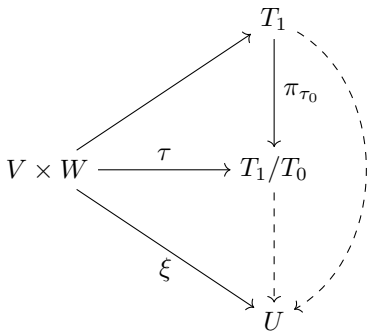
Beweis. Es ist $\dim_K(V \otimes_K W) < \infty$ und deshalb

$$V \otimes_K W \cong (V \otimes_K W)^{**} \overset{6.14}{\cong} \mathrm{Bil}_K(V, W, K) \qquad \square$$

▶ **Bemerkung 6.16**
Während obige Konstruktion des Tensorprodukts von der Wahl (und Existenz) von Basen abhängt, ist die folgende Konstruktion "basisfrei":
Sei T_1 der K-Vektorraum mit Basis $V \times W$ und T_0 der Untervektorraum von T_1 erzeugt von Elementen der Form:

$$\begin{aligned} \delta_{v+v',w} - \delta_{v,w} - \delta_{v',w} \\ \delta_{v,w+w'} - \delta_{v,w} - \delta_{v,w'} \\ \delta_{\lambda v,w} - \lambda \cdot \delta_{v,w} \\ \delta_{v,\lambda w} - \lambda \cdot \delta_{v,w} \end{aligned}$$

mit $v, v' \in V$, $w, w' \in W$ und $\lambda \in K$. Sei weiter $T = T_1/T_0$ und $\tau : V \times W \to T$ gegeben durch $(v, w) \mapsto \delta_{v,w} + T_0$. Dann ist (T, τ) ein Tensorprodukt von V und W.



▶ **Bemerkung 6.17**
Analog kann man für $k \geq 2$ und die K-Vektorräume $V_1, ..., V_k$ k-lineare Abbildungen $V_1 \times ... \times V_k \to U$ definieren und erhält dann Tensorprodukte $V_1 \otimes_K ... \otimes_K V_k$.

Kapitel VIII

Moduln

In diesem ganzen Kapitel sei R ein kommutativer Ring mit Einselement.

1. Moduln

Definition 1.1
Ein R-Modul ist ein Tripel $(M, +, \cdot)$ bestehend aus einer Menge M, einer Verknüpfung $+ : M \times M \to M$ und der Abbildung $\cdot : R \times M \to M$ (Skalarmultiplikation) für die gelten:

- (M1): $(M, +)$ ist eine abelsche Gruppe
- (M2): Addition und Skalarmultiplikation sind verträglich. Für alle $x, y \in M$ und $a, b \in R$ gelten
  1. $a(x + y) = ax + ay$
  2. $(a + b)x = ax + bx$
  3. $a \cdot bx = ab \cdot x$
  4. $1 \cdot x = x$

■ **Beispiel 1.2**

1. Ist $R = K$ ein Körper, so sind die R-Moduln genau die K-Vektorräume.
2. Ist $R = \mathbb{Z}$, so sind die R-Moduln genau die abelschen Gruppen mit der einzig möglichen Skalarmultiplikation

$$\mathbb{Z} \times A \to A, (k, a) \mapsto ka = \underbrace{1 + ... + 1}_{k\text{-mal}} a = \underbrace{a + ... + a}_{k\text{-mal}}$$

   vergleiche Laag 1 III.2.3
3. Jedes Ideal $M \subseteq R$ ist ein R-Modul mit Einschränkung der Multiplikation als Skalarmultiplikation.
4. Ist K ein Körper, V ein K-Vektorraum und $f \in \mathrm{End}_K(V)$, so wird V durch $P(t) \cdot x := P(f)(x)$ zu einem Modul über dem Ring $R = K[t]$, siehe auch V.5.2

▶ **Bemerkung 1.3**
Sei M ein R-Modul. Wie für Vektorräume (LAAG 1 II.1.5) überzeugt man sich leicht, dass $0x = 0$, $a0 = 0$, $(-a)x = a(-x) = -ax$ für alle $a \in R$, $x \in M$.

Im Gegensatz zu Vektorräumen folgt aber aus $ax = 0$ nicht, dass $a = 0$ oder $x = 0$, siehe zum

Beispiel das $\mathbb{Z}$-Modul $M = \mathbb{Z}/n\mathbb{Z}$. Es ist

$$n \cdot \overline{1} = \overline{n} = \overline{0} \in \mathbb{Z}/n\mathbb{Z}$$

aber $0 \neq n \in \mathbb{Z}$.

Definition 1.4 (Homomorphismus von R-Moduln)
Seien M, M' R-Moduln. Eine Abbildung $f : M \to M'$ ein Homomorphismus von R-Moduln (oder R-Homomorphismus oder R-linear), wenn

$$f(x + y) = f(x) + f(y)$$
$$f(ax) = a \cdot f(x)$$

Wir bezeichnen die Menge der R-Homomorphismen $f : M \to M'$ mit $\mathrm{Hom}_R(M, M')$. Wie üblich definiert man den Kern eines R-Homomorphismus, sowie die Begriffe Monomorphismus , Epimorphismus , Isomorphismus , Endomorphismus und Automorphismus von R-Moduln.

■ **Beispiel 1.5**

- Ist $R = K$, so sind die R-Homomorphismen genau die lineare Abbildungen.
- Ist $R = \mathbb{Z}$, so sind die R-Homomorphismen genau die Gruppenhomomorphismen.

■ **Beispiel 1.6**
Für jedes $a \in R$ ist die Abbildung

$$\begin{cases} M \to M \\ x \mapsto ax \end{cases}$$

einen Endomorphismus von M.

Definition 1.7 (Untermodul, Erzeugendensystem)
Ein Untermodul ist eine nichtleere Teilmenge $N \subseteq M$, für die gilt:

- Sind $x, y \in N$, so ist auch $x + y \in N$.
- Ist $a \in R$ und $x \in N$, so ist auch $ax \in N$.

Für eine Familie $(x_i)_{i \in I}$ ist

$$\sum_{i \in I} Rx_i = \{\sum_{i \in I} ax_i \mid a \in R, \text{ fast alle gleich } 0\}$$

der von $(x_i)_{i \in I}$ erzeugte Untermodul von M. Ist $\sum_{i \in I} Rx_i = M$, so ist $(x_i)_{i \in I}$ ein Erzeugendensystem von M. Der R-Modul M ist endlich erzeugt , wenn er ein endliches Erzeugendensystem besitzt.

▶ **Bemerkung 1.8**
Wieder ist der Kern eines R-Homomorphismus $f : M \to M'$ ein Untermodul von M. Leicht sieht man auch hier, dass $\sum_{i \in I} Rx_i$ ein Untermodul von M ist, und zwar der kleinste, der alle x_i enthält.

■ **Beispiel 1.9**

- Ist $R = K$ ein Körper, so sind die Untermoduln von M genau die Untervektorräume.
- Ist $R = \mathbb{Z}$, so sind die Untermoduln von M genau die Untergruppen und der von einer Familie erzeugte Untermodul ist genau gleich der davon erzeugten Untergruppe.
  Ist zum Beispiel $M = \mathbb{Z}$, so sind alle $n\mathbb{Z}$ Untermoduln von M.

Definition 1.10 (freie Familie, Basis)
Eine Familie $(x_i)_{i\in I}$ in M ist frei oder (R-linear unabhängig), wenn es keine Familie $(\lambda_i)_{i\in I}$ von Elementen von R, fast alle gleich 0, aber nicht alle gleich 0, mit $\sum_{i\in I} \lambda_i x_i = 0$ gibt.

Ein freies Erzeugendensystem heißt Basis . Besitzt M eine Basis, so nennt man M frei .

Satz 1.11
Seien M, M' R-Moduln, $(x_i)_{i\in I}$ eine Basis von M und $(y_i)_{i\in I}$ eine Familie in M'. Dann gibt es genau eine R-lineare Abbildung $f : M \to M'$ mit $f(x_i) = y_i$ für alle i.

Beweis. klar, siehe LAAG 1 III.5.1 □

■ **Beispiel 1.12**

- Für $n \in \mathbb{N}$ ist $M = R^n$ mit komponentenweiser Addition und Skalarmultiplikation ein endlich erzeugter freier R-Modul mit der üblichen Standardbasis.
- Allerdings ist zum Beispiel der $\mathbb{Z}$-Modul $\mathbb{Z}/n\mathbb{Z}$ zwar endlich erzeugt aber nicht frei. Für $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ ist $n\overline{a} = \overline{0}$, also $\overline{a}$ linear abhängig.

Definition 1.13 (Summen von Moduln)
Die Summe einer Familie $(N_i)_{i\in I}$ von Untermoduln von M ist

$$\sum_{i\in I} N_i = \left\{ \sum_{i\in I} x_i \mid x_i \in N_i, \text{ fast alle gleich } 0 \right\}$$

Lässt sich jedes $x \in \sum_{i\in I} N_i$ eindeutig als $\sum_{i\in I} x_i$ mit $x_i \in N_i$ schreiben, so nennt man die Summe direkt und schreibt dafür auch $\bigoplus_{i\in I} N_i$.

Ist $(M_i)_{i\in I}$ eine Familie von R-Moduln, so definiert man deren (externe) direkte Summe als das R-Modul

$$\bigoplus_{i\in I} M_i := \left\{ (x_i)_{i\in I} \in \prod_{i\in I} M_i \mid x_i = 0 \text{ für fast alle } i \in I \right\}$$

mit komponentenweiser Addition und Skalarmultiplikation.

▶ **Bemerkung 1.14**
Wie auch für Vektorräume ist eine externe direkte Summe eine direkte Summe der entsprechenden Untermoduln und ist $M = \bigoplus_{i\in I} N_i$, so ist M isomorph zur externen direkten Summe der N_i.

Definition 1.15 (Torsionsmodul)
Für $a \in R$ definiert man den a-Torsionsmodul von M als

$$M[a] := \{x \in M \mid ax = 0\}$$

Die Elemente des Torsionsmoduls

$$M_{tor} := \bigcup_{0 \neq a \in R} M[a] = \{x \in M \mid ax = 0 \text{ für ein } a \in R \backslash \{0\}\}$$

nennt man die Torsionselemente von M.

Satz 1.16
Für $a \in R$ ist $M[a]$ ein Untermodul von M. Ist R nullteilerfrei, so ist auch M_{tor} ein Untermodul von M.

Beweis. $M[a]$ ist der Kern des Endomorphismus $x \mapsto ax$ (Beispiel 1.6), somit ein Untermodul (Bemerkung 1.8). Seien $a, b \in R\backslash\{0\}$ und $x \in M[a]$, $y \in M[b]$. Ist R nullteilerfrei so ist $ab \neq 0$ und

$$(ab) \cdot (x + y) = b \cdot \underbrace{ax}_{=0} + a \cdot \underbrace{by}_{=0} = 0$$

also $x + y \in M[ab] \subseteq M_{tor}$. Somit ist M_{tor} in diesem Fall einer Untermodul von M. □

Beispiel 1.17
Sei $R = \mathbb{Z}$ und $M = \mathbb{Z}/n\mathbb{Z}$, dann ist $M_{tor} = M = M[n]$.

2. Teilbarkeit

Definition 2.1 (Teilbarkeit)
Seien $a, b \in R$.

1. a teilt b (in Zeichen $a \mid b$): Es existiert $x \in R$ mit $b = ax$.
2. a und b sind assoziiert (in Zeichen $a \sim b$): Es existiert $x \in R^\times$ mit $b = ax$.

Mathematica/WolframAlpha-Befehle (Teiler)
Möchte man mit Mathematica bzw. WolframAlpha überprüfen, ob n von m geteilt wird, also $m \mid n$ (!), kann man folge Funktion aufrufen:

```
Divisible[n,m]
```

Eine Liste der Teiler einer Zahl x erhält man mit

```
Divisors[x]
```

Lemma 2.2
Für $a, b, c, d \in R$ gelten

1. $a \mid a$
2. $a \mid b$ und $b \mid c \Rightarrow a \mid c$
3. $a \mid b$ und $a \mid c \Rightarrow a \mid (b + c)$
4. $a \mid b$ und $c \mid d \Rightarrow (ac) \mid (bd)$

Beweis. klar □

Lemma 2.3
Für $a, b, c, d \in R$ gelten

1. $a \sim a$
2. $a \sim b$ und $b \sim c \Rightarrow a \sim c$
3. $a \sim b \Rightarrow b \sim a$
4. $a \sim b$ und $c \sim d \Rightarrow (ac) \sim (bd)$

Beweis. klar, da $(R^\times, \cdot)$ eine Gruppe ist. □

▶ **Bemerkung 2.4**
Teilbarkeit auf R ist insbesondere eine Präordnung , das heißt reflexiv und transitiv, und Assoziiertheit ist eine Äquivalenzrelation.

Lemma 2.5
Sei R nullteilerfrei und seien $a, b \in R$. Genau dann ist $a \sim b$, wenn $a \mid b$ und $b \mid a$.

Beweis.
- Hinrichtung: $b = ax$ mit $x \in R^\times \Rightarrow a = bx^{-1}$.
- Rückrichtung: $b = ax$, $a = by$ mit $x, y \in R^\times$

$$a = by = axy$$
$$a(1 - xy) = 0$$

Also $a = 0$ und damit $b = 0$ oder $xy = 1$, also $x, y \in R^\times$. In beiden Fällen folgt $a \sim b$. □

■ **Beispiel**
Offenbar $2 \mid -2$ und $-2 \mid 2$. Es gilt $2 \sim -2$ und $-2 \sim 2$.

Satz 2.6
Sie R nullteilerfrei. Mit $[a] := \{a' \in R \mid a \sim a'\}$ wird durch $[a][b] \iff a \mid b$ eine wohldefinierte Halbordnung auf $R/\sim := \{[a] \mid a \in R\}$ gegeben.

Beweis.
- wohldefiniert: $a \mid b$, $a \sim a'$, $b \sim b' \Rightarrow a' \mid b'$: $ax = b$, $au = a'$, $bv = b$ mit $x \in R$ und $u, v \in R^\times$

$$b' = bv = axv = a' \underbrace{u^{-1}vx}_{\in R}$$

also $a' \mid b'$.
- reflexiv: klar
- transitiv: aus Transitivität von $\mid$
- antisymmetrisch: Lemma 2.5 □

Definition 2.7 (größter gemeinsamer Teiler, kleinstes gemeinsames Vielfaches)
Seien $a, b \in R$. Ein $c \in R$ ist ein größter gemeinsamer Teiler von a und b in Zeichen $c = \mathrm{ggT}(a, b)$, wenn gilt: $c \mid a$ und $c \mid b$ und ist $d \in R$ mit $d \mid a$ und $d \mid b$, so auch $d \mid c$.

Ein $c \in R$ ist ein kleinstes gemeinsames Vielfaches von a und b, in Zeichen $c = \mathrm{kgV}(a, b)$, wenn gilt: $a \mid c$ und $b \mid c$ und ist $d \in R$ mit $a \mid d$ und $b \mid d$, so ist $c \mid d$.

Mathematica/WolframAlpha-Befehle (ggT und kgV)
Die Funktionen für den größten gemeinsamen Teiler und das kleinste gemeinsame Vielfache in Mathematica bzw. WolframAlpha sind

```
GCD[6,12,4,32]
LCM[6,12,4,32]
```

▶ **Bemerkung 2.8**
Wenn ggT und kgV in einem nullteilerfreien Ring R existieren, sind sie eindeutig bestimmt, aber nur bis auf Assoziiertheit (Lemma 2.5).

Definition 2.9 (Primzahl, irreduzibel)
Sei $x \in R$.

- x ist prim $\iff x \notin R^\times \cup \{0\}$ und $\forall a, b \in R$ gilt $x \mid (ab) \Rightarrow x \mid a \vee x \mid b$.
- x ist irreduzibel $\iff x \notin R^\times \cup \{0\}$ und $\forall a, b \in R$ gilt $x = ab \Rightarrow a \in R^\times \vee b \in R^\times$.

▶ **Bemerkung 2.10**
Leicht sieht man: Ist $p \in R$ prim und $a_1, ..., a_n \in R$ mit $p \mid (a_1 \dots a_n)$, so gilt $p \mid a_i$ für ein i.

■ **Beispiel 2.11**

- In $R = \mathbb{Z}$ gilt: p prim $\iff$ p irreduzibel
- Sei $f \in R = \mathbb{Q}[t]$.
  - $\deg(f) = 1 \Rightarrow f \sim (t - a)$ ist irreduzibel und prim (denn $(t - a) \mid g \iff g(a) = 0$)
  - $\deg(f) = 2$: $f = t^2 - 1$ ist nicht irreduzibel, $t^2 - 2$ ist irreduzibel

Satz 2.12
Sei R nullteilerfrei und $0 \neq p \in R \backslash R^\times$. Ist p prim, so ist es auch irreduzibel.

Beweis. Sei $p = ab$ mit $a, b \in R$. Da insbesondere $p \mid ab$ und p prim ist, folgt $p \mid a$ oder $p \mid b$. Sei ohne Einschränkung $p \mid a$, das heißt $a = pa'$ mit $a' \in R$.

$$\begin{aligned} &\Rightarrow p = ab = pa'b \\ &\Rightarrow p(1 - ab) = 0 \\ &\Rightarrow a'b = 1, \text{ insbesondere } b \in R^\times \end{aligned}$$

Somit ist p irreduzibel. □

▶ **Bemerkung 2.13**
Erinnerung: Ein Ideal von R ist eine Untergruppe $I \subseteq (R, +)$ mit

$$a \in I, r \in R \Rightarrow ra \in I$$

also genau ein Untermodul des R-Moduls R.

Definition 2.14 (erzeugtes Ideal, Hauptideal)
Sei $A \subseteq R$. Das von A erzeugte Ideal mit

$$\langle A \rangle := \left\{ \sum_{i=1}^{n} r_i a_i \mid n \in \mathbb{N}_0, a_1, ..., a_n \in A, r_1, ..., r_n \in R \right\}$$

Ist $A = \{a_1, ..., a_n\}$, so schreibt man auch $(a_1, ..., a_n)$ für $\langle A \rangle$. Ein Ideal der Form $I = (a)$ ist ein Hauptideal .

▶ **Bemerkung 2.15**
Das von A erzeugte Ideal $\langle A \rangle$ ist gleich dem von A erzeugten Untermodul des R-Moduls R, und ist das kleinste Ideal von R, das A enthält.

▶ **Bemerkung 2.16**

Für $a \in R$ ist $(a) = Ra$ und für $a, b \in R$ sind äquivalent:

1. $a \mid b$
2. $b \in (a)$
3. $(b) \subseteq (a)$

Für R nullteilerfrei sind zudem äquivalent:

1. $a \sim b$
2. $(a) = (b)$

■ **Beispiel 2.17**

Jeder Ring hat die Ideale $(0) = \{0\}$ und $(1) = R$. Für jedes $a \in R^\times$ ist $(a) = (1)$, ist R also ein Körper, so hat R keine weiteren Ideale.

■ **Beispiel 2.18**

In $R = \mathbb{Z}$: Für $n \in \mathbb{Z}$ ist $(n) = \mathbb{Z} \cdot n = n\mathbb{Z}$.

3. Hauptidealringe

Sei R nullteilerfrei.

Definition 3.1 (Hauptidealring)
Ein Ring R ist ein Hauptidealring , wenn R nullteilerfrei ist und jedes Ideal von R ein Hauptideal ist.

■ **Beispiel 3.2**
Ist $R = K$ ein Körper, so hat R nur die Ideale (0) und (1), und somit ist R ein Hauptidealring.

Definition 3.3 (euklidische Gradfunktion)
Eine euklidische Gradfunktion auf R ist eine Abbildung $\delta : R\backslash\{0\} \to \mathbb{N}_0$ für die gilt:
Für jedes $a \in R$ und $0 \neq b \in R$ gibt es $q, r \in R$ mit $a = bq + r$, wobei $r = 0$ oder $\delta(r) < \delta(b)$.

Ein nullteilerfreier Ring R ist euklidisch , wenn es eine euklidische Gradfunktion auf R gibt.

■ **Beispiel 3.4**

1. Auf $R = \mathbb{Z}$ ist der Absolutbetrag
$$\delta(x) = |x|$$
eine euklidische Gradfunktion. (LAAG 1 I.4.6)
2. Auf $R = K[t]$, K ein Körper, ist der Grad
$$\delta(f) = \deg(f)$$
eine euklidische Gradfunktion. (LAAG 1 I.6.5)
3. $R = K$ ein Körper ist
$$\delta(x) = 0$$
eine euklidische Gradfunktion, da man in einem Körper jedes Element durch jedes Element (Ausnahme: 0) teilen kann.

Lemma 3.5
Sei $\delta : R\backslash\{0\} \to \mathbb{N}_0$ eine euklidische Gradfunktion und $(0) \neq \trianglelefteq R$ ein Ideal. Ist $0 \neq a \in I$ mit $\delta(a) = \min\{\delta(b) \mid 0 \neq b \in I\}$, so ist $I = (a)$.

Beweis.
- "$\supseteq$": $a \in I \Rightarrow (a) \subset I$
- "$\subseteq$": Sei $0 \neq b \in I$. Schreibe $b = qa + r$ mit $q, r \in R$ und $r = 0$ oder $\delta(r) < \delta(a)$. Da $r = \underbrace{b}_{\in I} - q\underbrace{a}_{\in I} \in I$ folgt wegen der Minimalität von $\delta(a)$, dass $r = 0$, also $b \in (a)$. □

Satz 3.6
Ist R euklidisch, so ist R ein Hauptidealring.

Beweis. Sei $I \trianglelefteq R$ ein Ideal. Ist $I = (0)$, so ist I ein Hauptideal. Andernfalls existiert ein $0 \neq a \in I$ mit $\delta(a)$

minimal. Nach Lemma 3.5 ist $I = (a)$ ein Hauptideal. □

Folgerung 3.7
Die Ringe $\mathbb{Z}$ und $K[t]$, K ein Körper, sind Hauptidealringe.

Lemma 3.8 (Lemma von Bézout)
Sei R ein Hauptidealring und $a, b \in R$. Es existiert ein $c \in R$ mit $c = \mathrm{ggT}(a, b)$ und $(c) = (a, b)$. Insbesondere gibt es $x, y \in R$ mit $c = ax + by$ und $\mathrm{ggT}(x, y) = 1$.

Beweis. R Hauptidealring $\Rightarrow \exists c \in R$ mit $(c) = (a, b)$, insbesondere $c = ax + by$ mit $x, y \in R$.

- $c = \mathrm{ggT}(a, b)$: $a, b \in (c) \Rightarrow c \mid a$ und $c \mid b$. Ist $d \in R$ mit $d \mid a$ und $d \mid b$, so ist $d \mid (ax + by) = c$
- $\mathrm{ggT}(x, y) = 1$: Ist $d \in R$ mit $d \mid x$ und $d \mid y$, so gelten $(cd) \mid (ax)$ und $(cd) \mid (by) \Rightarrow (cd) \mid (ax + by) = c \Rightarrow d \in R^\times$, also $d \sim 1$. □

Satz 3.9
Sei R ein Hauptidealring, $p \in R$. Ist p irreduzibel, so auch prim.

Beweis. Seien $a, b \in R$ mit $p \mid (ab)$. Angenommen $p \nmid a$. DA p irreduzibel ist, ist $\mathrm{ggT}(p, a) = 1$, also $1 = px + ay$ mit $x, y \in R$ nach Lemma 3.8. Also $p \mid (pbx + aby) = b$. □

4. Faktorielle Ringe

Sei R nullteilerfrei.

Definition 4.1 (faktorielle Ringe)
R ist faktoriell $\iff$ jedes $0 \neq x \in R \backslash R^\times$ ist ein Produkt von Primelementen.

Lemma 4.2
Sei R faktoriell und $x \in R$. Ist x irreduzibel, so auch prim.

Beweis. Sei x irreduzibel, insbesondere $0 \neq x \in R\backslash R^\times$. Da R faktoriell, ist $x = p_1, ..., p_n$ mit $p_1, ..., p_n \in R$ prim. Da x irreduzibel ist und $p_i \notin R^\times$ ist $n = 1$ und somit $x = p_1$ prim. □

Lemma 4.3
Sei R ein Hauptidealring und

$$I_1 \subseteq I_2 \subseteq ...$$

eine Kette von Idealen in R. Dann existiert ein $n \in \mathbb{N}$ mit $I_n = I_m$ für alle $m \geq n$.

Beweis. Behauptung: $I = \bigcup_{n=1}^{\infty} I_n$ ist wieder ein Ideal von R.
Beweis: schon in den Übungen zum Teil behandelt, aber hier noch mal kurz bewiesen

- $i \in I$, $r \in R \Rightarrow x \in I_n$ für ein $n \overset{I_n \trianglelefteq}{\Rightarrow} rx \in I_n \subseteq I$
- $x, y \in I \Rightarrow x \in I_n, y \in I_m$ mit $n, m \in \mathbb{N} \overset{\text{Kette}}{\Rightarrow} x + y \in I_k \subseteq$ mit $k = \max\{n, m\}$

Da R Hauptidealring ist, ist somit $I = (x)$ für ein $x \in R$. Mit $I = \bigcup_{n\in\mathbb{N}} I_n$ folgt $x \in I_n$ für ein n, und somit $(x) \subseteq I_n \subseteq I_m \subseteq I = (x)$, für $m \geq n$, also $I_n = I_m$. □

Satz 4.4
Ist R ein Hauptidealring, so ist R faktoriell.

Beweis. Sei $X := \{a \in R \mid a \text{ ist Produkt von Primelementen}\} \cup \{0\} \cup R^\times$. Zu zeigen ist $X = R$. Angenommen, es gebe $a \in R\backslash X$. Da nicht prim ist, insbesondere nicht irreduzibel (Satz 3.9), ist $a = a_1 \cdot a_1'$ mit $a_1, a_1' \in R\backslash R^\times$. Wären a_1 und a_1' in X, so auch a, also ohne Einschränkung $a_1 \notin X$. Fährt man nun mit a_1 so fort, erhält man eine Folge $a_1, a_2, ...$ von Elementen von $R\backslash X$ mit $a_{i+1} \mid a_i$ und $a_{i+1} \nsim a_i$ für alle i. Die entsprechenden Hauptideale bilden eine Kette

$$(a) \subsetneqq (a_1) \subsetneqq (a_2) \subsetneqq ...$$

im Widerspruch zu Lemma 4.3. Somit ist $X = R$, also R faktoriell. □

Anmerkung
Es gilt also euklidisch ⇒ Hauptidealring ⇒ faktoriell.

Lemma 4.5
Sind $p_1, ..., p_r \in R$ prim, $q_1, ..., q_s \in R$ irreduzibel mit

$$\prod_{i=1}^{r} p_i = \prod_{j=1}^{s} q_j$$

ist $r = s$ und nach Umnummerierung ist

$$p_i \sim q_i \quad \forall i$$

Beweis. Wir zeigen die Behauptung unter der schwächeren Annahme

$$\prod_{i=1}^{r} p_i \sim \prod_{j=1}^{s} q_j$$

durch Induktion nach r.
$\underline{r = 0:}$ $1 \sim \prod_{j=1}^{s} q_j \Rightarrow q_j \in R^\times \; \forall j \overset{q_j \text{ irred.}}{\Rightarrow} s = 0$
$\underline{r - 1 \to r:}$ $p_1 \mid \prod_{i=1}^{r} p_i \sim \prod_{j=1}^{s} q_j \overset{p_1 \text{ prim}}{\Rightarrow} p_1 \mid q_j$ für ein j. Nach Umnummerierung ist $j = 1$. Da q_1 irreduzibel und $p_1 \notin R^\times$ ist $p_1 \sim q_1$, also $q_1 = p_1 \cdot u$ mit $u \in R^\times$. Es folgt

$$p_1 \cdot \left(\prod_{i=2}^{r} p_i - u \cdot \prod_{j=2}^{s} q_j \right) = 0$$

$$\prod_{i=2}^{r} p_i = u \cdot \prod_{j=2}^{s} q_j \sim \prod_{j=2}^{s} q_j$$

Nach Induktionshypothese ist $r - 1 = s - 1$, und nach Umnummerierung ist $p_i \sim q_i$ für $i = 2, ..., r$. □

Satz 4.6
Ist R faktoriell, so lässt sich jedes $0 \neq x \in R \backslash R^\times$ auf eindeutige Weise (bis auf Reihenfolge und Assoziiertheit) als Produkt von Primelementen schreiben.

Beweis. Sei $x = \prod_{i=1}^{r} p_i = \prod_{j=1}^{s} q_j$ mit p_i, q_j prim. Da die q_j nach Satz 2.12 irreduzibel sind, folgt $r = s$ und $p_i \sim q_i$ für alle i aus Lemma 4.5. □

Folgerung 4.7
Sei R faktoriell und enthalte $\mathcal{P} \subseteq R$ für jede Äquivalenzklasse assoziierter Primelemente genau einen Vertreter. Dann lässt sich jedes $0 \neq a \in R$ als

$$a = \varepsilon \cdot \prod_{p \in \mathcal{P}} p^{\mu(p)}$$

mit eindeutig bestimmten $\varepsilon \in R^\times$ und $\mu(p) \in \mathbb{N}_0$, fast alle gleich 0, schreiben.

Beispiel 4.8

1. Jedes $n \in \mathbb{N}$ lässt sich eindeutig als

$$n = \prod_{p \in \mathbb{P}} p^{n_p}$$

schreiben, wobei $\mathbb{P}$ die Menge der Primzahlen ist (Hauptsatz der Arithmetik).

2. Bezeichnet $\mathcal{M}$ die Menge der normierten irreduziblen Polynome in $K[t]$ (K Körper), so lässt sich jedes $0 \neq f \in K[t]$ eindeutig als

$$f = c \cdot \prod_{P \in \mathcal{M}} P^{n_p}$$

mit $c \in K^\times$ und $n_p \in \mathbb{N}_0$, fast alle gleich 0, schreiben.

5. Quotienten von Ringen und Moduln

Seien M und M' zwei R-Moduln und $N \subseteq M$ ein Untermodul.

Definition 5.1 (Quotientenmodul)
Für $x \in M$ schreiben wir

$$x + N := \{x + y \mid y \in N\}$$

Der Quotientenmodul (oder Faktormodul) von M modulo N ist

$$M/N := \{x + N \mid x \in M\}$$

zusammen mit der Addition

$$(x + N) + (y + N) := (x + y) + N \quad (x, y \in M)$$

und der Skalarmultiplikation

$$r \cdot (x + N) := rx + N \quad (x \in M, r \in R)$$

Sei $\pi_N : M \to M/N$ die Abbildung gegeben durch $x \mapsto x + N$.

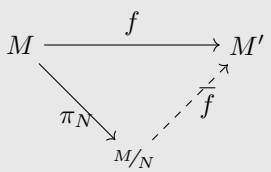
Lemma 5.2
Addition und Skalarmultiplikation sind wohldefiniert und machen M/N zu einem R-Modul. Die Abbildung $\pi_N : M \to M/N$ ist ein R-Epimorphismus mit Kern

$$\mathrm{Ker}(\pi_N) = N$$

Beweis.
- wohldefiniert: wie in LAAG 1 III.7.5
- M/N ist R-Modul: wie in LAAG 1 III.7.7 □

▶ **Bemerkung 5.3**
Durch $x \sim_N x' \iff x - x' \in N$ wird eine Äquivalenzrelation $\sim_N$ auf M definiert, und $x + N$ ist eine $\sim_N$-Äquivalenzklasse $[x]_{\sim_N} = \{y \in M \mid x \sim_N y\}$.

Satz 5.4 (Homomorphiesatz für Moduln)
Sei $f \in \mathrm{Hom}_K(M, M')$ und $N \subseteq M$ ein Untermodul mit $N \subseteq \mathrm{Ker}(f)$. Dann gibt es genau ein $\overline{f} \in \mathrm{Hom}_K(M/N, M')$ mit $f = \overline{f} \circ \pi_N$.



Beweis. Analog zu LAAG 1 III.7.9. Man zeigt, dass jedes $\overline{f} \in \mathrm{Hom}_K({}^M\!/_N, M')$

$$\overline{f}(x+N) = f(x) \quad (x \in M)$$

erfüllen muss, und dass dies wiederum eine wohldefinierte Abbildung liefert. □

Lemma 5.5
Durch $U \mapsto \pi_N(U)$ wird eine Bijektion gegeben zwischen

- den Untermoduln von M, die N enthalten
- den Untermoduln von ${}^M\!/_N$.

Beweis. Sei $\mathcal{U}$ die Menge der Untermoduln von M, die N enthalten, $\overline{\mathcal{U}}$ die Menge der Untermoduln von ${}^M\!/_N$.

- $U \in \mathcal{U} \Rightarrow \pi_N(U) \in \overline{\mathcal{U}}$: klar, da π_N ein Homomorphismus ist
- $\overline{U} \in \overline{\mathcal{U}} \Rightarrow \pi_N^{-1} \in \mathcal{U}$: klar, da π_N ein Homomorphismus ist und $N = \mathrm{Ker}(\pi_N) = \pi_N^{-1}(\{0\}) \subseteq \pi_N^{-1}(\overline{U})$
- $\overline{U} \in \overline{\mathcal{U}} \Rightarrow \pi_N(\pi_N^{-1}(\overline{U})) = \overline{U}$: klar, da π_N surjektiv
- $U \in \mathcal{U} \Rightarrow \pi_N^{-1}(\pi_N(U)) = U$:

$$\begin{aligned}
\pi_N^{-1}(\pi_N(U)) &= \bigcup_{x \in U} \pi_N^{-1}(\pi_N(x)) \\
&= \bigcup_{x \in U} \pi_N^{-1}(x+N) \\
&= \bigcup_{x \in U} (x+N) \\
&= U + N = U
\end{aligned}$$

□

▶ **Bemerkung 5.6**
Das Ideal $I \trianglelefteq R$ ist ein Untermodul des R-Moduls R, somit haben wir ein R-Modul ${}^R\!/_I$ definiert. Man kann ${}^R\!/_I$ mit einer Ringstruktur ausstatten.

Definition 5.7 (Quotientenring)
Sei $I \trianglelefteq R$ ein Ideal. Für $x \in R$ schreiben wir

$$x + I = \{x + a \mid a \in I\}$$

Dann ist

$${}^R\!/_I = \{x + I \mid x \in R\}$$

der Quotientenring von R modulo I mit Addition und Skalarmultiplikation

$$\begin{aligned}
(x+I) + (x'+I) &= (x+x') + I \quad \forall x, x' \in R \\
(x+I) \cdot (x'+I) &= (x \cdot x') + I \quad \forall x, x' \in R
\end{aligned}$$

Und wieder $\pi_I : R \to {}^R\!/_I$ mit $x \mapsto x + I$.

Satz 5.8
Addition und Multiplikation sind wohldefiniert und machen R/I zu einem kommutativen Ring mit Einselement. π_I ist ein Ringhomomorphismus mit Kern

$$\mathrm{Ker}(\pi_I) = I$$

Beweis.

- Addition wohldefiniert: Lemma 5.2
- Multiplikation wohldefiniert: Sind $x, x', y, y' \in R$ mit

$$x + I = x' + I$$
$$y + I = y' + I$$

Dann ist

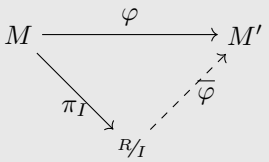
$$x - x' = a \in R \quad \Rightarrow x = x' + a$$
$$y - y' = b \in R \quad \Rightarrow y = y' + a$$

Also

$$xy = (x' + a)(y' + b) = x'y' + \underbrace{ay' + x'b + ab}_{\in I}$$
$$\Rightarrow xy + I = x'y' + I$$

- R/I ist Ring: R1 bis R3 folgen aus den entsprechenden Eigenschaften von R.
- R/I ist kommutativ: folgt auch aus den Eigenschaften von R.
- Einselement: $1 + I$
- π_I ist ein Ringhomomorphismus: folgt nach Definition
- $\mathrm{Ker}(\pi_I)$: klar □

Satz 5.9 (Homomorphiesatz für Ringe)
Sei $\varphi : R \to R'$ ein Ringhomomorphismus, $I \trianglelefteq R$ ein Ideal mit $I \subseteq \mathrm{Ker}(\varphi)$. Dann gibt es genau einen Ringhomomorphismus mit $\overline{\varphi} : R/I \to R'$, sodass $\overline{\varphi} \circ \pi_I = \varphi$.



Beweis. Man sieht, dass

$$\overline{\varphi}(x + I) = \varphi(x) \quad \forall x \in R$$

gelten muss, und das dies auch ein wohldefinierter Ringhomomorphismus ist. □

■ **Beispiel 5.10**

- $R = \mathbb{Z}$, $\forall n \in \mathbb{N}$ ist $n\mathbb{Z}$ ein Ideal.

$$\mathbb{Z}/(n) = \mathbb{Z}\backslash n\mathbb{Z}$$

- Sei K ein Körper und sei $a \in K$. Dann ist $K[t] \to K$, $P \mapsto P(a)$ ist ein Ringepimorphismus. Der Kern $\mathrm{Ker}(\varphi) = (t-a)$, also alle Polynome, die in a eine Nullstelle haben. Es folgt

$$K[t]/(t-a) \cong K$$

☺ $\mathbb{Z}$ ist der Herr der Ringe ☺

■ **Beispiel 5.11**

Sei $0 \neq p \in K[t]$. $K[t]/(p)$ ist ein Ring, aber auch ein $K[t]$-Modul und damit ein K-Vektorraum.

$$\dim_K \left(K[t]/(p)\right) = n = \deg(p)$$

Ist $B = (1, \bar{t}, ..., \overline{t^{n-1}})$ eine Basis wobei $\overline{x} = \pi_{(p)}(x)$ $\forall x \in K[t]$.

6. Der Elementarteilersatz

Sei R Hauptidealring.

Definition 6.1

Seien $a, b, x, y \in R$. Für $i, j \in \{1, ..., n\}$ ist

$$E_{ij} = (\delta_{\sigma,i}, ..., \delta_{\mu,j})_{\sigma,\mu} \in \mathrm{Mat}_n(\mathbb{R})$$

Sei

$$E_{ij}(a, b, x, y) = \mathbb{1}_n - E_{ii} - E_{jj} + aE_{ii} + bE_{ij} + xE_{jj} + yE_{ji}$$

Lemma 6.2

Ist $ax - by \in R^\times$, so ist

$$E_{ij}(a, b, x, y) \in \mathrm{GL}_n(\mathbb{R})$$

Beweis. Folgt aus LAAG1 IV.3.4, da

$$\det(E_{ij}(a, b, x, y)) = ax - by \in R^\times$$

Oder direkt: Das Inverse ist $E_{ij}(xc^{-1}, bc^{-1}, ac^{-1}, -yc^{-1})$, zum Beispiel

$$\begin{pmatrix} a & b \\ y & x \end{pmatrix} \begin{pmatrix} xc^{-1} & -bc^{-1} \\ -yc^{-1} & ac^{-1} \end{pmatrix} = \begin{pmatrix} (ax - by)c^{-1} & 0 \\ 0 & (ax - by)c^{-1} \end{pmatrix}$$

□

▶ **Bemerkung 6.3**

Multiplikation von $E_{ij}(a, b, x, y)$ von links an A führt eine Zeilenumformung durch: Sind $a_1, ..., a_n$ die Zeilen von A, so wird a_i durch $aa_i + ba_j$ ersetzt, und gleichzeitig a_j durch $ya_i + xa_j$ ersetzt. Ist $ax - by = 1$, so sind diese Zeilenumformungen invertierbar.

Spezialfälle: elementare Zeilenumformungen von Typ II und III aus Kapitel III (LAAG 1). Warnung: Im Gegensatz dazu sind über einem Ring R die elementaren Zeilenumformungen vom Typ I (Multiplikation mit einem Skalar) nicht immer invertierbar!

Multiplikation mit $E_{ij}(a, b, x, y)$ von rechts führt entsprechende Spaltenumformungen durch.

Theorem 6.4 (Elementarteilersatz für Matrizen, Smith-Normalform)
Sei $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$. Es gibt $0 \leq r \leq \min\{n, m\}$, $S \in \mathrm{GL}_m(R)$, $T \in \mathrm{GL}_n(R)$ mit

$$SAT = \begin{pmatrix} d_1 & & & \\ & \ddots & & \\ & & d_r & \\ & & & \mathbb{0} \end{pmatrix}$$

$$\mathbb{0} \in \mathrm{Mat}_{m-r\times n-r}$$

wobei $d_i \in R\backslash\{0\}$ mit $d_i \mid d_{i+1}$ für $i = 1, ..., n-1$

Beweis. Induktion nach $\min\{m, n\}$. Für $a \in R$ sei $\delta(a) \in \mathbb{N}_0 \cup \{\infty\}$ die Anzahl der Primelemente in der Primfaktorzerlegung von a, mit $\delta(0) := \infty$, und $\delta(A) := \min_{ij}\{\delta(a_{ij})\}$. Wir können annehmen, dass $\delta(A) \leq \delta(SAT)$ für alle $S \in \mathrm{GL}_m(R)$ und $T \in \mathrm{GL}_n(R)$. Durch Zeilen- und Spaltenvertauschungen erreichen wir, dass $\delta(a_{11}) = \delta(A)$.

- 1. Behauptung: $a_{11} \mid a_{i1}$ für alle i. Gäbe es ein $i \geq 1$ für dass $a_{11} \nmid a_{i1}$, so sei $c = \mathrm{ggT}(a_{11}, a_{i1}) = xa_{11} + ya_{i1}$ mit $\mathrm{ggT}(x, y) = 1$, also $ax - by = 1$ mit $a, b \in R$. Multiplikation mit $E_{1i}(x, y, a, b)$ von links erzeugt an der Position $(1, 1)$ das Element c, und $\delta(c) < \delta(a_{11}) = \delta(A)$, im Widerspruch zur Minimalität von $\delta(A)$. Analog zeigt man, dass $a_{11} \mid a_{1j}$ für alle j. Durch Zeilen- und Spaltenumformungen können wir deshalb nun $a_{i1} = 0$ für alle $i > 1$ und a_{1j} für alle $j > 1$ erreichen.
- 2. Behauptung: $a_{11} \mid a_{ij}$ für alle i, j. Gäbe es $i > 1$ und $j > 1$ mit $a_{11} \nmid a_{ij} := b$, so können wir die j-te Spalte zur ersten Spalte addieren, was a_{11} nicht ändert und $a_{1i} = b$ bewirkt. Wider können wir Behauptung 1 anwenden und erhalten den Widerspruch, dass $a_{11} \mid b$. Damit ist nach diesem Umformungen

$$A = \begin{pmatrix} a_{11} & & \\ & & \\ & a_{11} \cdot A' & \end{pmatrix}$$

mit $A' \in \mathrm{Mat}_{(m-1)\times(n-1)}(R)$. Wir wenden nun die Induktionshypothese auf A' an und sind fertig. □

Mathematica/WolframAlpha-Befehle (Smith-Normalform)
Elementarteiler einer Matrix A lassen sich mit Mathematica mit der Funktion

```
SmithDecomposition[A]
```

die als einziges Argument eine Matrix braucht. Allerdings ist der Output unformatiert, mit folgenden Befehl sieht das deutlich besser aus:

```
MatrixForm/@ ({u,r,v} = SmithDecomposition[A])
```

Der Output sind 3 Matrizen, wobei `u` für S, `v` für T und `r` für das Ergebnis von SAT steht.

▶ **Bemerkung 6.5**

Man kann zeigen, dass die $d_1, ..., d_r$ bis auf Assoziiertheit eindeutig bestimmt sind. Man nennt sie deshalb Elementarteiler der Matrix A.

■ **Beispiel 6.6**

Sei $R = \mathbb{Z}$. Die Elementarteiler von

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

sind

$$\begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix} \to \begin{pmatrix} 4 & 0 \\ 4 & 6 \end{pmatrix} \to \begin{pmatrix} 4 & 0 \\ -2 & 6 \end{pmatrix} \to \begin{pmatrix} 2 & -6 \\ 4 & 0 \end{pmatrix} \to \begin{pmatrix} 2 & 0 \\ 4 & 12 \end{pmatrix} \to \begin{pmatrix} 2 & 0 \\ 0 & 12 \end{pmatrix}$$

2, 2 und 12.

Anmerkung (Teil 1)

Um die Elementarteiler der Matrix A_0 zu ermitteln, muss man geschickt mit Matrizen S und T multiplizieren. Dazu starten wir links oben bei Element $a_{11} \neq 0$ und versuchen nun, auf der ersten Spalte und auf der ersten Zeile nur Nullen zu produzieren, aber $a_{11} \neq 0$ zu erhalten.

Dazu fangen wir mit der ersten Spalte an. Ziel ist es, das letzte Element dieser Spalte durch geschickte Addition der vorletzten Spalte zu 0 werden zu lassen. Wir schauen uns die letzten 2 Elemente, nennen wir sie x und y, dieser ersten Spalte an und bestimmen $\mathrm{ggT}(x, y)$. Weiterhin suchen wir u und v, sodass folgende Gleichung erfüllt ist:

$$\mathrm{ggT}(x, y) = u \cdot x + v \cdot y$$

Da wir eine Zeilenoperation durchführen wollen, brauchen wir eine Matrix S_0, die wir von links an A ranmultiplizieren. Dabei müssen wir auf die richtige Dimension von S_0 aufpassen. Dazu setzen wir S_0 auf $\mathbb{1}_m$ und fügen an der richtigen Stelle die Matrix S_0' ein:

$$S_0' = \begin{pmatrix} u & v \\ -\frac{y}{\mathrm{ggT}(x,y)} & \frac{x}{\mathrm{ggT}(x,y)} \end{pmatrix}$$

Jetzt bestimmen wir $A_1 := A_0 \cdot S_0$. Jetzt haben wir das letzte Element der ersten Spalte zu 0 verwandelt. Wir arbeiten uns jetzt in der ersten Spalte nach oben, versuchen also das vorletzte Element zu 0 zu verwandeln, aber mithilfe der vorvorletzten Zeile. Auch dazu bestimmen wir wieder Matrizen $S_1, S_2, ...$ bis die erste Spalte 0 ist, mit Ausnahme von a_{11}.

Anmerkung (Teil 2)
Jetzt wenden wir uns der ersten Zeile zu: Auch hier versuchen wir das letzte Element zu 0 zu verwandeln, aber eben mit Benutzung der vorletzten Spalte. Die Vorgehensweise ist nahezu identisch, wir bestimmen auch wieder $\mathrm{ggT}(x, y)$ und lösen

$$\mathrm{ggT}(x, y) = u \cdot x + v \cdot y$$

Damit bauen wir uns wieder T_0', die wir an der passenden Stelle in $T_0 = \mathbb{1}_n$ einsetzen

$$T_0' = \begin{pmatrix} u & -\frac{y}{\mathrm{ggT}(x,y)} \\ v & \frac{x}{\mathrm{ggT}(x,y)} \end{pmatrix}$$

Die Matrix T_0 multiplizieren wir aber diesmal von rechts an A_n. So arbeiten wir uns wieder von hinten nach vorne. Es kann passieren, dass wir uns damit leider wieder in der ersten Spalte ein paar Nullen kaputt machen, aber dann bauen wir wieder eine S_n-Matrix mit der wieder Nullen erscheinen. Falls das wieder die Spalten kaputt macht, dann multiplizieren wir wieder mit einer T_n-Matrix. Das Theorem 6.4 garantiert uns, dass wir irgendwann fertig werden.

Anmerkung (Teil 3)
Haben wir nun die erste Zeile und die erste Spalte zu 0 verwandelt, außer a_{11} natürlich, kümmern wir uns um die Untermatrix in Richtung rechts unten. Hier geht der Algorithmus von vorne los; das Schöne ist, dass er uns die erste Zeile/Spalte nicht mehr kaputt machen kann. Irgendwann sind wir rechts unten angekommen und haben nur noch Elemente auf der Hauptdiagonalen stehen. Diese sollten, wie in Theorem 6.4 behauptet eine solche Teilerkette bilden. Tun sie das nicht, kann man wieder mit Matrizen S_n und T_n nachhelfen.

$$S_n' = \begin{pmatrix} u & v \\ -\frac{y}{\mathrm{ggT}(x,y)} & \frac{x}{\mathrm{ggT}(x,y)} \end{pmatrix} \quad T_n' = \begin{pmatrix} 1 & -\frac{vy}{\mathrm{ggT}(x,y)} \\ 1 & \frac{ux}{\mathrm{ggT}(x,y)} \end{pmatrix}$$

$$\text{unter Vorbehalt! } S_n' = \begin{pmatrix} 1 & 1 \\ -\frac{vy}{\mathrm{ggT}(x,y)} & \frac{ux}{\mathrm{ggT}(x,y)} \end{pmatrix}$$

Und dann sind wir endlich fertig! Die Transformationsmatrizen S und T sind dann einfach

$$S = S_1 \cdot S_2 \cdot \ldots$$
$$T = T_1 \cdot T_2 \cdot \ldots$$

Weitere Informationen und Beispiele findet man auf http://www.igt.uni-stuttgart.de/eiserm/lehre/2010/Algebra/Matrizenringe.pdf, ab Abschnitt §7D

Lemma 6.7
Ist M ein endlich erzeugter freier R-Modul und $N \subseteq M$ ein Untermodul, so ist auch N endlich erzeugt.

Beweis. Sei $(x_1, ..., x_m)$ eine Basis von M. Induktion nach m.
$\underline{m = 1:}$ Durch $1 \mapsto x_1$ wird nach Satz 1.11 eine R-lineare Abbildung $f : R \to M$ gegeben, die ein Isomorphismus ist. Der Untermodul $N \subseteq M$ entspricht einem Ideal $I := f^{-1}(N)$ von R. Da R ein Hauptidealring ist, ist $I = (a)$ für ein $a \in R$, somit $N = f(I) = R \cdot f(a)$. Insbesondere ist N endlich erzeugt, sogar von einem Element.
$\underline{m - 1 \to m:}$ Definiere $M' = \sum_{i=1}^{m-1} Rx_i$, $M'' = Rx_m$, $N' = N \cap M'$. Sei unter $\pi : M \to M''$ die R-lineare Abbildung gegeben nach Satz 1.11 durch $\pi(x_i) = \delta_{i,m} x_m$. Nach Induktionshypothese ist N' endlich erzeugt, etwa $N' = \sum_{j=1}^{n} Ry_j$. Aus dem Fall $m = 1$ sehen wir zudem, dass $N'' = \pi(N) = R\pi(y)$ für ein $y \in N$. Sei $\tilde{N} = Ry + \sum_{j=1}^{n} Ry_j \subseteq N$. Da $\mathrm{Ker}(\pi|_N) = M'' \cap N = N' \subseteq \tilde{N}$ und $\pi|_N(\tilde{N}) \supseteq R\pi(y) = N'' = \pi|_N(N)$ ist $\tilde{N} = N$ nach Lemma 5.5 und Satz 5.4. Somit ist N endlich erzeugt. □

Satz 6.8 (Elementarteilersatz für Moduln)
Sei R ein Hauptidealring, $M \cong R^m$ ein endlich erzeugter freier R-Modul, $N \subseteq M$ ein Untermodul. Dann existiert $r \in \mathbb{N}$, eine Basis $B' = (x'_1, ..., x'_m)$ von M und $d_1, ..., d_r \in R \backslash \{0\}$ mit $d_i \mid d_{i+1}$ für $i = 1, ..., r - 1$ für die $(d_1 x'_1, ..., d_r x'_r)$ eine Basis von N ist.

Beweis. Sei $B = (x_1, ..., x_m)$ eine Basis von M. Nach Lemma 6.7 ist N endlich erzeugt, also

$$N = \sum_{j=1}^{n} Ry_j \quad \text{mit} \quad y_j = \sum_{i=1}^{m} a_{ij} x_i \quad a_{ij} \in R$$

Wir betrachten die lineare Abbildung $f : R^n \to M$ gegeben durch $f(e_j) = y_j$. Dann ist $\mathrm{Im}(f) = N$ und

$$M_B^{\mathcal{E}}(f) = A = (a_{ij}) \in \mathrm{Mat}_{m \times n}(R)$$

Nach Theorem 6.4 existieren $S \in \mathrm{GL}_m(R)$, $T \in \mathrm{GL}_n(R)$ mit

$$SAT = D = \mathrm{diag}(d_1, ..., d_r, \mathbb{0})$$

Es gibt somit Basen $\mathcal{E}' = (e'_1, ..., e'_n)$ von R^n, $B' = (x'_1, ..., x'_m)$ von M mit $M_{B'}^{\mathcal{E}'}(f) = D$. Somit ist $N = \mathrm{Im}(f) = \sum_{i=1}^{n} R \cdot f(e'_i) = \sum_{j=1}^{r} Rd_j x'_j$. Da $(x'_1, ..., x'_r)$ frei und R nullteilerfrei ist, ist auch $(d_1 x'_1, ..., d_r x'_r)$ frei, also eine Basis von N. □

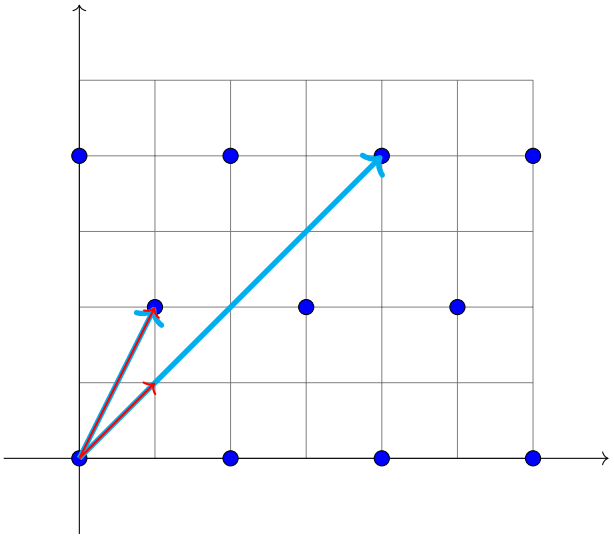
■ **Beispiel**

Sei $R = \mathbb{Z}$, $M = \mathbb{Z}^2$, $N = \mathbb{Z}\begin{pmatrix}2\\0\end{pmatrix} + \mathbb{Z}\begin{pmatrix}1\\2\end{pmatrix}$

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

$$\Rightarrow B = \left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\4\end{pmatrix}\right) \Rightarrow B' = \left(\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right)$$

$$\Rightarrow C = \left(1 \cdot \begin{pmatrix}1\\2\end{pmatrix}, 4 \cdot \begin{pmatrix}1\\1\end{pmatrix}\right) \text{ ist Basis von } N$$



▶ **Bemerkung 6.9**

Wieder kann man zeigen, dass $d_1, ..., d_r$ bis auf Einheiten eindeutig bestimmt sind.

Folgerung 6.10

Ist R ein Hauptidealring, so ist ein Untermodul eines endlich erzeugten freien R-Moduls wieder frei.

▶ **Bemerkung 6.11**

Folgerung 6.10 wird falsch ohne "R Hauptidealring". So ist zum Beispiel $N = (x, y) \trianglelefteq \mathbb{Q}[x, y] = (\mathbb{Q}[x])[y] = R = M$ kein Hauptideal und somit ein nicht freier Untermodul des freien R-Moduls R: Je zwei Elemente von R sind linear abhängig, für $a, b \in R$ ist

$$b \cdot a + (-a) \cdot b = 0$$

Deshalb kann N keine Basis mit mehr als einem Element besitzen.

Die Voraussetzung "endlich erzeugt" ist hingegen nicht notwendig, aber der Beweis wird dadurch einfacher.

Folgerung 6.12
Ist R ein Hauptidealring, so ist ein Untermodul eines endlich erzeugten R-Moduls M wieder endlich erzeugt.

Beweis. Ist $M = \sum_{j=1}^{m} Ry_j$, so betrachte die R-lineare Abbildung $f : R^m \to M$ gegeben durch $f(e_j) = y_j$ für $j = 1, ..., m$. Nach Lemma 6.7 ist $f^{-1}(N) \subseteq R^m$ endlich erzeugt, etwa $f^{-1}(N) = \sum_{i=1}^{n} Rx_i$. Somit ist $N = f(f^{-1}(N)) = \sum_{i=1}^{n} R \cdot f(x_i)$ endlich erzeugt. □

Theorem 6.13 (Hauptsatz über endlich erzeugte Moduln über Hauptidealringen)
Sei R ein Hauptidealring und M ein endlich erzeugter R-Modul. Dann ist

$$M = F \oplus M_{tor}$$

wobei $F \cong R^r$ ein endlich erzeugter freier R-Modul ist und

$$M_{tor} \cong \bigoplus_{i=1}^{n} {}^{R}/_{Rd_i}$$

mit Nichteinheiten $d_1, ..., d_n \in R\backslash\{0\}$, die $d_i \mid d_{i+1}$ für $i = 1, ..., n-1$ erfüllen.

Beweis. Inhalt... □

▶ **Bemerkung 6.14**
Auch gier sind $d_1, ..., d_n$ (bis auf Einheiten) sowie r eindeutig bestimmt. Man nennt r den (freien) Rang von M.

Anhang

Anhang A: Listen

A.1. Liste der Theoreme

A.2. Liste der benannten Sätze

A.3. Liste der Mathematica/WolframAlpha-Befehle

☺ für faule Mathematiker ☺

Index