Skript zur Vorlesung

Lineare Algebra II

Sommersemester 2005

Universität Konstanz

private Mitschrift

Stand: 14. Juli 2005
www.meidert.net/uni

Achtung:
Dies ist kein offizielles Script, sondern nur eine private Mitschrift. Ich kann daher keine Gewähr für die Richtigkeit und Vollständigkeit übernehmen. Vor allem können die Nummerierungen zum Teil von den in den Vorlesungen verwendeten abweichen. Falls jemand einen Fehler entdeckt, so möge er/sie mir bitte eine eMail schicken - vielen Dank!

Frieder Meidert (uni@meidert.net)

Inhaltsverzeichnis

6. Bilinearformen und Skalarprodukte

a. Linearformen und Dualraum

Sei $\mathbb{K}$ *ein Körper, alle Vektorräume sind* $\mathbb{K}$*-Vektorräume.*

Definition 6.1.
Sei V ein $\mathbb{K}$-Vektorraum. Eine ($\mathbb{K}$-)**Linearform** auf V ist eine lineare Abbildung $\lambda : V \to \mathbb{K}$. Der Vektorraum aller Linearformen auf V heißt **Dualraum** von V, in Zeichen $V^* = Hom_{\mathbb{K}}(V, \mathbb{K}) = \{\lambda : V \to \mathbb{K} | \lambda \text{ ist linear}\}$.

Erinnerung:
$(\lambda + \mu)(v) = \lambda(v) + \mu(v)$, $(a\lambda)(v) = a \cdot \lambda(v)$ für $\lambda, \mu \in V^*$, $a \in \mathbb{K}$ (siehe LA I - 3.36).

Beispiele 6.2.

1. Sei $V = \mathbb{K}^n$. Für jeden Vektor $a = (a_1, \ldots, a_n) \in \mathbb{K}^n$ ist $\lambda_a : \mathbb{K}^n \to \mathbb{K}$, $\lambda_a(x_1, \ldots, x_n) := \sum_{i=1}^{n} a_i x_i$, eine Linearform auf $\mathbb{K}^n$.
   Ist umgekehrt $\lambda \in (\mathbb{K}^n)^*$, so setze $a_i := \lambda(e_i)$, $i = 1, \ldots, n$, dann ist $\lambda = \lambda_a$ mit $a := (a_1, \ldots, a_n)$.
   Denn für jedes $i = 1, \ldots, n$ ist $\lambda_a(e_i) = \sum_{j=1}^{n} a_j \delta_{ij} = a_i = \lambda(e_i)$.
   Die Abbildung $\mathbb{K}^n \to (\mathbb{K}^n)^*$, $a \mapsto \lambda_a$, ist also bijektiv, und sie ist auch linear $\lambda_{a+b} = \lambda_a + \lambda_b$, $\lambda_{c \cdot a} = c \cdot \lambda_a$ für $a, b \in \mathbb{K}^n, c \in \mathbb{K}$. Somit ist sie ein Vektorraum-Isomorphismus.
   Wir werden häufig Elemente von $(\mathbb{K}^n)^*$ als Zeilenvektoren $a = (a_1, \ldots, a_n)$ auffassen (und weiterhin Elemente von $\mathbb{K}^n$ also Spalten). Dann wird die „Auswertungspaarung“ $(\mathbb{K}^n)^* \times \mathbb{K}^n \to \mathbb{K}, (\lambda, v) \mapsto \lambda(v)$ zu Abbildung $((a_1, \ldots, a_n), \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}) \mapsto$
   $(a_1, \ldots, a_n) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} a_i v_i$.
2. Sei V ein Vektorraum, sei $\mathcal{B} = (v_i)_{i \in I}$ eine Basis von V. Für jedes $\lambda \in V^*$ ist $(\lambda(v_i))_{i \in I} \in \mathbb{K}^I = \prod_{i \in I} \mathbb{K}$. Die so definierte Abbildung $V^* \to \mathbb{K}^I = \prod_{i \in I} \mathbb{K}$ ist ein Vektorraum-Isomorphismus. (Übung!)

Sei jetzt $\dim(V) = n < \infty$. Dann ist $\dim(V^*) = n$ (da $\dim(\mathbb{K}) = 1$).
Wir wollen Basen von V^* konstruieren. Sei dazu $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Definiere $v_i^* \in V^*$ $(i = 1, \ldots, n)$ durch $v_i^*\left(\sum_{j=1}^{n} a_j v_j\right) := a_i$. Die v_i^* sind also charakterisiert durch $v_i^*(v_j) = \delta_{ij}$.

Definition und Satz 6.3.
*Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Dann ist $B^* := (v_1^*, \ldots, v_n^*)$ eine Basis von V^*, genannt die zu $\mathcal{B}$ **duale Basis**. Für jedes $\lambda \in V^*$ ist $\lambda = \sum_{i=1}^{n} \lambda(v_i) \cdot v_i^*$.*

Beweis:
$\mathcal{B}^*$ ist linear unabhängig: ist $\sum_{i=1}^{n} a_i v_i^* = 0$, so gibt Einsetzen von v_j: $a_j = 0$ $(j = 1, \ldots, n)$.
Es genügt nun, die letzte Identität zu zeigen:
für $j = 1, \ldots, n$ ist $\left(\sum_{i=1}^{n} \lambda(v_i) \cdot v_i^*\right)(v_j) = \sum_{i=1}^{n} \lambda(v_i) \cdot \underbrace{v_i^*(v_j)}_{\delta_{ij}} = \lambda(v_j) \Rightarrow$ Behauptung.

□

Bemerkungen 6.4.

1. Die Linearform v_i^* hängt nicht nur von v_i, sondern von ganz $\mathcal{B}$ ab.
2. Die VRe V und V^* sind isomorph: z.B. ist $f : V \to V^*$, $f\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i v_i^*$, ein Isomorphismus. Dieser Isomorphismus ist nicht kanonisch, da er von der Basis $\mathcal{B}$ abhängt.
   Man kann (präzisieren und) zeigen, dass es keinen kanonischen Isomorphismus $V \to V^*$ gibt.
3. Ist $\dim(V) = \infty$, so sind V und V^* niemals isomorph. (ohne Beweis)

Lemma 6.5.
*Sei $f : V \to W$ eine lineare Abbildung. Für $\mu \in W^*$ sei $f^*(\mu) = \mu \circ f : V \to \mathbb{K}$. Dann ist $f^*(\mu) \in V^*$ und die Abbildung $f^* : W^* \to V^*$ ist linear. Sie heißt die zu f **duale Abbildung**.*

Beweis:
f^* ist linear: $f^*(a_1\mu_1 + a_2\mu_2)(v) = (a_1\mu_1 + a_2\mu_2)(f(v)) = a_1\mu_1(f(v)) + a_2\mu_2(f(v)) = (a_1 f^*(\mu_1) + a_2 f^*(\mu_2))(v)$.

□

ACHTUNG:
f^* geht in die „umgekehrte“ Richtung!

LEMMA 6.6.
Seien $f, f_1, f_2 : V \to W$ lineare Abbildungen.

(a) $(a_1 f_1 + a_2 f_2)^* = a_1 f_1^* + a_2 f_2^*$ $(a_1, a_2 \in \mathbb{K}^*)$, *mit anderen Worten: die Abbildung $Hom(V,W) \to Hom(V^*, W^*)$, $f \mapsto f^*$, ist linear.*

(b) *Ist $g : U \to V$ weitere lineare Abbildungen, so ist $(f \circ g)^* = g^* \circ f^*$ (Reihenfolge!!).*

(c) $(id_V)^* = id_{V^*}$.

(d) *Ist f bijektiv, so auch f^*, und dann gilt $(f^{-1})^* = (f^*)^{-1}$.*

BEWEIS:

(a) klar

(b) $(f \circ g)^*(\mu) = \mu \circ (f \circ g) = (\mu \circ f) \circ g = (f^*(\mu)) \circ g = g^*(f^*(\mu))$.

(c) klar

(d) aus (b) und (c): $\underbrace{(f \circ f^{-1})^*}_{=id_{W^*}} = (f^{-1})^* \circ f^*$.

□

LEMMA 6.7.
Sei $f : V \to W$ eine lineare Abbildung, sei $\mathcal{B}$ eine Basis von V, sei C eine Basis von W. Dann ist $M_{\mathcal{B}^*}^{C^*}(f^*) = \left(M_C^{\mathcal{B}}(f)\right)^t$.

BEWEIS:
Sei $\mathcal{B} = (v_1, \ldots, v_n)$, $C = (w_1, \ldots, w_n)$, also $\mathcal{B}^* = (v_1^*, \ldots, v_n^*)$, $C^* = (w_1^*, \ldots, w_n^*)$. Sei $M_C^{\mathcal{B}}(f) = (a_{ij}) \in M_{m \times n}(\mathbb{K})$. Es ist also $f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$ $(i = 1, \ldots, n)$.
Nach 6.3 berechnen wir:
$(f^*(w_i^*))(v_j) = w_i^*(f(v_j)) = w_i^*\left(\sum_{k=1}^{m} a_{kj} w_k\right) = a_{ij}$.

Nach 6.3 ist also $f^*(w_i^*) = \sum_{j=1}^{n} a_{ij} v_j^*$ $(i = 1, \ldots, m)$.
Das ist genau die Behauptung.

□

Korollar 6.8.
Sei $f : V \to W$ linear, v und W endlich-dimensional.

(a) $\dim(\ker(f)) + \dim(im(f^*)) = \dim(V)$,
$\dim(\ker(f^*)) + \dim(im(f)) = \dim(W)$.

(b) *f injektiv $\Leftrightarrow$ f^* surjektiv; f surjektiv $\Leftrightarrow$ f^* injektiv.*

(c) *f bijektiv $\Leftrightarrow$ f^* bijektiv.*

Beweis:

(a) Sei $n = \dim(V)$, $m = \dim(W)$, wähle Basen $\mathcal{B}, C$ von V, W wie oben, sei $A := M_C^{\mathcal{B}}(f)$. Es ist $M_{\mathcal{B}^*}^{C^*}(f^*) = A^t$. Es ist $rk(A) = rk(A^t)$, also ist $r := rk(f) = rk(f^*)$.
$\Rightarrow \dim(\ker(f)) = n - r$, $\dim(\ker(f^*)) = m - r$. Daraus folgt direkt (a).

(b) f injektiv $\Leftrightarrow \ker(f) = \{0\} \Leftrightarrow \dim(im(f^*)) = \dim(V) = \dim(V^*) \Leftrightarrow f^*$ surjektiv.
Analog die zweite Aussage.

(c) folgt direkt aus (b).

□

Satz 6.9.
Sei V ein VR, sei $\dim(V) < \infty$, *sei $V^{**} := (V^*)^*$.*
Es gibt einen kanonischen *Isomorphismus $e : V \to V^{**}$.*

Beweis:
Für $v \in V$ sei $e(v) : V^* \to \mathbb{K}$ definiert durch $e(v)(\lambda) := \lambda(v)$ $(\lambda \in V^*)$. Es ist $e(v) \in V^{**}$ und die Abbildung $V \to V^{**}$, $v \mapsto e(v)$, ist linear.
z.B. $e(v + w)(\lambda) = \lambda(v + w) = \lambda(v) + \lambda(w) = e(v)(\lambda) + e(w)(\lambda)$.

Die lineare Abbildung e ist injektiv: denn ist $v \in V$ mit $e(v) = 0$, so ist $\lambda(v) = 0$ für alle $\lambda \in V^*$. Es ist klar, dass hieraus $v = 0$ folgt. Wegen $\dim(V) = \dim(V^{**}) < \infty$ folgt aus e injektiv auch e bijektiv.

□

Folgerung:
Für $\dim(V) < \infty$ sind V und V^{**} **kanonisch isomorph**.

Ist $\dim(V) = \infty$, so zeigt der letzte Beweis: $e : V \to V^{**}$ ist injektiv, aber nicht mehr surjektiv.

b. Bilinearformen, quadratische und alternierende Formen

Definition 6.10.
Seien V_1, V_2, W $\mathbb{K}$-VRe.

(a) Eine Abbildung $b : V_1 \times V_2 \to W$ heißt ($\mathbb{K}$-)**bilinear**, wenn gilt:

(i) $b(v_1 + v_1', v_2) = b(v_1, v_2) + b(v_1', v_2)$

(ii) $b(v_1, v_2 + v_2') = b(v_1, v_2) + b(v_1, v_2')$,

(iii) $b(av_1, v_2) = a \cdot b(v_1, v_2) = b(v_1, av_2)$

für alle $v_1, v_1' \in V_1, v_2, v_2' \in V_2, a \in \mathbb{K}$.

(b) (Fall $W = \mathbb{K}$) Eine bilineare Abbildung $b : V_1 \times V_2 \to \mathbb{K}$ heißt eine **bilineare Paarung** zwischen V_1 und V_2.

(c) (Falls $W = \mathbb{K}$, $V_1 = V_2 = V$) Eine bilineare Abbildung $b : V \times V \to \mathbb{K}$ heißt eine **Bilinearform** auf V.

Bemerkung 6.11.

1. Bilinear bedeutet linear in jeder Komponente. Eine bilineare Abbildung $V_1 \times V_2 \to W$ ist nie linear, außer wenn $b \equiv 0$ ($(v_1, v_2) = (v_1, 0) + (0, v_2)$; es ist $b(v_1, 0) = b(v_1, 0 \cdot 0) = 0$, genauso $b(0, v_2) = 0$)
   Induktiv folgt aus der Definition: ist $b : V_1 \times V_2 \to W$ bilinear, so ist $b\left(\sum_{i=1}^{m} a_i v_i, \sum_{j=1}^{n} a_j' v_j'\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i a_j' \cdot b(v_i, v_j')$.

2. Für jeden VR V ist die Abbildung $b : V^* \times V \to \mathbb{K}$, $b(\lambda, v) := \lambda(v)$ eine bilineare Paarung zwischen V^* und V.

3. Die Abbildung $\mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $(x, y) \mapsto \sum_{i=1}^{n} x_i y_i$, ist eine Bilinearform auf $\mathbb{K}^n$ (die sogenannte **kanonische** Bilinearform auf $\mathbb{K}^n$)

4. Die Abbildung $M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to \mathbb{K}$, $(A, B) \mapsto tr(AB)$, ist bilinear.

5. Für feste V_1, V_2, W bilden die bilinearen Abbildungen $V_1 \times V_2 \to W$ einen $\mathbb{K}$-VR unter den Operationen $(b_1+b_2)(v_1, v_2) := b_1(v_1, v_2)+b_2(v_1, v_2)$, $(ab)(v_1, v_2) := ab(v_1, v_2)$.

Bemerkung 6.12.
Jede bilineare Paarung $b : V_1 \times V_2 \to \mathbb{K}$ definiert zwei **lineare** Abbildungen $l_b V_1 \to V_2^*$, $r_b V_2 \to V_1^*$ durch $l_b(v_1)(v_2) := b(v_1, v_2)$, $r_b(v_2)(v_1) := b(v_1, v_2)$
Umgekehrt erhält man b aus l_b, oder aus r_b, wieder zurück.

Lemma 6.13.
Sei $\dim(V_1) < \infty$, $\dim(V_2) < \infty$, *sei* $b : V_1 \times V_2 \to \mathbb{K}$ *bilinear.*
Man nennt b ***nicht-ausgeartet*** *(oder* ***vollkommen****), falls die folgenden drei äquivalenten Bedingungen gelten:*

(i) $l_b : V_1 \to V_2^*$ *ist ein Isomorphismus;*

(ii) $r_b : V_2 \to V_1^*$ *ist ein Isomorphismus;*

(iii) l_b *und* r_b *sind injektiv.*

Andernfalls heißt b ***ausgeartet****.*

Beweis:
Das Dreieck $V_2 \to (r_b)V_1^* \leftarrow (l_b^*)V_2^{**} \leftarrow (e \sim)V_2$ kommutiert.

Denn für $v_2 \in V_2$ und $v_1 \in V_1$ ist $l_b^*(e(v_2))(v_1) = e(v_2)(l_b(v_1)) = l_b(v_1(v_2) = b(v_1, v_2) = r_b(v_2)(v_1)$, also $l_b^*(e(v_2)) = r_b(v_2)$.
Daraus sehen wir sofort, da e ein Isomorphismus ist (6.8):
(i) $\Leftrightarrow$ (ii). klar damit auch (i) $\Rightarrow$ (iii),
(ii) $\Rightarrow$ (iii).

Umgekehrt gelte $\dim(V_1 \underbrace{\leq}_{l_b \text{ injektiv}} \dim(V_2^*) = \dim(V_2) \underbrace{\leq}_{r_b \text{injektiv}} \dim(V_1^*) = \dim(V_1.$
$\Rightarrow \dim(V_1) = \dim(V_2) \Rightarrow l_b, r_b$ bijektiv.

□

Bemerkung 6.14.

1. Genau dann ist b nicht-ausgeartet, wenn es zu jedem $0 \neq v_1 \in V_1$ ein $v_2 \in V_2$ mit $b(v_1, v_2) \neq 0$ ist.
(analog auch die symmetrische Formulierung)

2. Die bilinearen Paarungen in 6.11 2. bis 4. sind alle nicht-ausgeartet.

Im Folgenden sei stets $\dim(V) < \infty$ *Wir konzentrieren uns auf den Fall* $V_1 = V_2 = V$.

Konstruktion 6.15.
Sei $b \in Bil(V)$:= Vektorraum aller Bilinearformen $V \times V \to \mathbb{K}$.
Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V.

(a) Die Matrix $M_{\mathcal{B}}(b)$ von b bezüglich $\mathcal{B}$ ist die Matrix $M_{\mathcal{B}}(b) := (a_{ij})_{1 \leq i,j \leq n}$ mit $a_{ij} := b(v_i, v_j)$. Durch die Matrix $M_{\mathcal{B}}(b)$ ist die Bilinearform b festgelegt wegen

$$b\left(\sum_{i=1}^{n} a_i v_i, \sum_{j=1}^{n} a'_j v_j\right) = \sum_{i,j} a_i a'_j \cdot \underbrace{b(v_i, v_j)}_{a_{ij}} = (a_1, \ldots, a_n) \cdot M_{\mathcal{B}}(b) \cdot \begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix} \qquad (*)$$

(b) Umgekehrt definiert jede Matrix $A \in M_n(\mathbb{K})$ eine Bilinearform $b_A = b_{A,\mathcal{B}}$ auf V durch $b_A\left(\sum_i x_i v_i, \sum_j y_j v_j\right) = x^t \cdot A \cdot y$
Die so definierte Abbildung $M_{\mathcal{B}} : Bil(V) \to M_n(\mathbb{K})$ ist also ein Isomorphismus der VRe.

Bsp: Zur kanonischen Bilinearform $\mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ $(x, y) \mapsto \sum_{i=1}^{n} x_i y_i = x^t y$ gehört bezüglich der kanonischen Basis auf $\mathbb{K}^n$ die Einheitsmatrix.

Satz 6.16.
Sei $b \in Bil(V)$, *sei* $\mathcal{B}$ *Basis von* V *und* $A := M_{\mathcal{B}}(b)$. *Dann gilt:*
b *nicht-ausgeartet* $\Leftrightarrow$ A *ist regulär.*

Beweis:
Wir bestimmen die Matrix der linearen Abbildung $r_b : V \to V^*$ bezüglich $\mathcal{B}$ und $\mathcal{B}^*$. Sei $\mathcal{B} = (v_1, \ldots, v_n), \mathcal{B}^* = (v_1^*, \ldots, v_n^*)$. Nach 6.3(b) gilt $r_b(v_j) = \sum_i \underbrace{r_b(v_j)(v_i)}_{b(v_i,v_j)=a_{ij}} \cdot v_i^* =$

$\sum_j a_{ij} \cdot v_j^*$
$A =: (a_{ij})$. Das zeigt $A = M^{\mathcal{B}}_{\mathcal{B}^*}(r_b)$.
wegen b nicht-ausgeartet $\Leftrightarrow$ r_b Isomorphismus folgt die Behauptung, dass A regulär ist.

□

Lemma 6.17.
Sei $b \in Bil(V)$, seien $\mathcal{B}, C$ Basen von V. Dann gilt $M_C(b) = S^t \cdot M_{\mathcal{B}}(b) \cdot S$ mit $S := T^C_{\mathcal{B}} = M^C_{\mathcal{B}}(id)$.

Beweis:
Sei $\mathcal{B} = (v_1, \ldots, v_n), C = (w_1, \ldots, w_n)$. Sei $S = (s_{ij})$, dann gilt also $w_j = \sum_{i=1}^{n} s_{ij} v_i$ $(j = 1, \ldots, n) \Rightarrow$
Seien $M_{\mathcal{B}}(b) =: (a_{ij}), M_C(b) = (b_{ij}) \Rightarrow b_{ij} = b(w_i, w_j) = b\left(\sum_k s_{ki} v_k, \sum_l s_{lj} v_l\right) = \sum_{k,l} s_{ki} s_{lj} a_{kl} =$ $(S^t A S)_{ij}$.

□

Nebenbemerkung:
Dieses Transformationsgesetz ist **verschieden** von dem Basiswechsel für die Matrizen linearer Abbildungen (**??**).

Definition 6.18.
Zwei Matrizen $A, B \in M_n(\mathbb{K})$ heißen **kongruent**, in Zeichen $A \simeq B$, wenn es $S \in GL_n(\mathbb{K})$ gibt mit $B = S^t A S$.

Bemerkung 6.19.

0. Hatten schon $\sim$ und $\approx$.

1. $\simeq$ ist eine Äquivalenzrelation auf $M_n(\mathbb{K})$.

2. $A \simeq B \Rightarrow \exists c \in \mathbb{K}$ mit $\det(B) = c^2 \cdot \det(A)$. Denn $B = S^t A S \Rightarrow \det(B) = \det(A) \cdot \det(S)^2$.

3. $A \simeq B \Rightarrow rk(A) = rk(B)$, siehe **??**. Also ist $\simeq$ feiner als $\sim$.

Definition 6.20.
Eine Bilinearform $b : V \times V \to \mathbb{K}$ heißt **symmetrisch**, wenn gilt $b(v_1, v_2) = b(v_2, v_1)$ $\forall v_1, v_2 \in V$. Sei $Bilsym(V)$ der VR der symmetrischen Bilinearformen.

Definition 6.21.
Eine Matrix $A \in M_n(\mathbb{K})$ heißt symmetrisch, wenn $A = A^t$ ist. Man schreibt $Sym_n(\mathbb{K}) := \{A \in M_n(\mathbb{K}) : A = A^t\}$

$b : V \times V \to \mathbb{K}$ $b(v, w) = b(w, v)$ $\mathcal{B}$ von V $\rightsquigarrow$ $M_{\mathcal{B}}(b)$.

Bemerkung 6.22.

1. Ist $b \in Bil(V)$ und $\mathcal{B}$ Basis von V, so gilt: b symmetrisch $\Leftrightarrow$ $M_{\mathcal{B}}(b)$ ist symmetrisch.

2. Ist $f : \mathbb{R}^n \to \mathbb{R}$, 2-mal stetig differenzierbar, so ist für $x \in \mathbb{R}^n$ die zweite Ableitung
   $D^2 f(x) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(u, v) \mapsto \sum_{i,j=1}^{n} \frac{\delta^2 f(x)}{\delta x_i \delta x_j} u_i v_j$ eine symmetrische Bilinearform auf $\mathbb{R}^n$, die sogenannte **Hesseform** von f.

Definition 6.23.
Sei $b \in Bil^{sym}(X)$. Dann nennt man die Abbildung $q_b : V \to \mathbb{K}$, $q_b(v) := b(v, v)$, die zu b gehörende **quadratische Form**.

Lemma 6.24.
Seien b und q_b wie eben. Dann gilt für $v, w \in V$ für $a \in \mathbb{K}$:

(a) $2b(v, w) = q_b(v + w) - q_b(v) - q_b(w)$
(Polarisierungs-Identität);

(b) $q_b(a \cdot v) = a^2 q_b(v)$;

(c) $q_b(v + w) + q_b(v - w) = 2(q_b(v) + q_b(w))$.
(Parallelogramm-Identität).

BEWEIS: Schreibe $q := q_b$.

(a) $q(v+w) = b(v+w, v+w) = q(v) + q(w) + 2b(v,w)$.

(b) $q(av) = b(av, av) = a^2 b(v,v) = a^2 q(v)$.

(c) aus (a) $q(v+w) + q(v-w) = 2q(v) + 2q(w) + \underbrace{b(v,w) + b(v,-w)}_{=0}$.

□

BEMERKUNG 6.25.

1. Die Polarisierungs-Identität sagt für $char(\mathbb{K}) \neq 2$: aus q_b gewinnt man b wieder zurück.

2. 6.24 (c): Betrachte $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^n$,
   $b :=$ kanonische symmetrische Bilinearform auf $\mathbb{R}^n$, dann ist $q(x) = \sum_{i=1}^{n} x_i^2 = \|x\|^2$, also sagt hier 6.24 (c): $\|v+w\|^2 + \|v-w\|^2 = 2(\|v\|^2 + \|w\|^2)$

   Fall $n = 2$: Die Summe der Flächeninhalte der Quadrate über den Diagonalen ist gleich der Summe der Flächeninhalte der Quadrate über den vier Seiten.

DEFINITION 6.26.
Sei V ein $\mathbb{K}$-VR. Eine Abbildung $q : V \to \mathbb{K}$ heißt eine **quadratische Form** auf V, wenn gilt:

(a) $q(av) = a^2 \cdot q(v)$ $\forall a \in \mathbb{K}, v \in V$,

(b) die Abbildung $\beta_q : V \times V \to \mathbb{K}$, $\beta_q(v,w) := q(v+w) - q(v) - q(w)$, ist eine (symmetrische) Bilinearform auf V.

Die quadratische Form q heißt **nicht-ausgeartet**, falls β_q **nicht-ausgeartet** ist.

Für jede Symmetrische Bilinearform b ist also q_b (wie in 6.23) eine quadratische Form im Sinne von Definition 6.26.

Satz 6.27.
Sei char$(\mathbb{K}) \neq 2$. *Sei* V *ein* $\mathbb{K}$*-VR, schreibe* $QF(V)$ *für die Menge aller quadratischen Formen* $q : V \to \mathbb{K}$. *Die Abbildung*
$Bil^{sym}(V) \to QF(V)$, $b \mapsto q_b$,
ist bijektiv. Die Umkehrabbildung ist $q \mapsto \frac{1}{2}\beta_q$.

Beweis:
Für $b \in Bil^{sym}(V)$ und $q_b(v) = b(v,v)$. ist $\frac{1}{2}\beta_q(v,w) = \frac{1}{2}q_b(v+w-q_b(v)-q_b(w)) = b(v,w)$ (Polarisierungs-Identität).
Für $\phi \in QF(V)$ ist $q_{\beta_\phi}(v) = \beta_\phi(v,v) = \phi(v+v) - 2\phi(v) = 2\phi(v)$ also $\phi(v) = q_{\frac{1}{2}\beta_\phi}(v)$.
□

Bemerkung 6.28.

1. Für *char*$(\mathbb{K}) \neq 2$ sind also symmetrische Bilinearformen und quadratische Formen im Wesentlichen dasselbe.

2. Für jede Basis $\mathcal{B}$ von V ($\dim(V) < \infty$) gibt $M_\mathcal{B}$ also auch Bijektionen zwischen $QF(V)$ und $Sym_n(\mathbb{K})$, $n := \dim(V)$.

1. Beispiel:
$V = \mathbb{K}^n$, $\mathcal{B}$:= kanonische Basis:
zur symmetrischen Matrix $A = (a_{ij}) \in Sym_n(\mathbb{K})$ gehört die quadratische Form
$q_A : \mathbb{K}^n \to \mathbb{K}$, $q_A(x) = x^t \cdot A \cdot x = \sum_{i,j=1}^{n} a_{ij}x_ix_j$.

2. Beispiel:
$\det : M_2(\mathbb{K}) \to \mathbb{K}$ ist eine **nicht-ausgeartete** quadratische Form (vgl. Übungsblatt 1, Aufgabe 4).

3.Vorsicht:
Es kann passieren, dass die quadratische Form $q : V \to \mathbb{K}$ nicht-ausgeartet ist, aber es ein $0 \neq v \in V$ gibt mit $q(v) = 0$ (z.B. 2. Beispiel; oder $q : \mathbb{R}^2 \to \mathbb{R}, q(x_1,x_2) := x_1x_2$,

$q(e_1) = q(e_2) = 0$, aber q ist nicht-ausgeartet).

Die einfachsten quadratischen Formen auf $\mathbb{K}^n$ sind die Diagonalformen $q(x) = \sum_{i=1}^{n} a_i x_i^2$.

Theorem 6.29.
Sei char$(\mathbb{K}) \neq 2$. *Sei* $\dim(V) < \infty$, *und sei* $b \in Bil^{sym}(V)$. *Dann gibt es eine Basis* $\mathcal{B} = (v_1, \ldots, v_n)$ *von* V *mit* $b(v_i, v_j) = 0$ *für alle* $i \neq j, i, j = 1, \ldots, n$.

Definition 6.30.
Jede solche Basis $\mathcal{B}$ heißt eine **Orthogonalbasis** (O-Basis) von V bezüglich b.

Korollar 6.31.
Sei char$(\mathbb{K}) \neq 2$. *Jede symmetrische Matrix* $A \in Sym(\mathbb{K})$ *ist kongruent zu einer Diagonalmatrix, d.h.* $\exists S \in GL_n(\mathbb{K})$ *mit* $SAS^t =$ *Diagonalmatrix.*

Beweis:
Ist $A \in Sym_n(\mathbb{K})$, so betrachte die symmetrische Bilinearform $b_A(x, y) = x^t \cdot A \cdot y$ auf $\mathbb{K}^n$ und wende Theorem 6.29 darauf an: es wird also $S = T^{\mathbb{K}}_{\mathcal{B}}$ (wahrscheinlich nicht ganz richtig; $\Leftrightarrow$ nachrechnen ...)

□

Korollar 6.32.
Sei char$(\mathbb{K}) \neq 2$. *Jedes homogene quadratische Polynom* $q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ *über* $\mathbb{K}$ *(die* x_i *seinen „Unbestimmte") lässt sich schreiben in der Form* $q(x) = \sum_{i=1}^{n} c_i \cdot (b_{i1}x_1 + \ldots + b_{in}x_n)^2$ *mit geeigneten* $b_{ij}, c_i \in \mathbb{K}$.

Beweis:
Sei $A = (\widetilde{a_{ij}})$ die symmetrische Matrix mit $q(x) = x^t \cdot A \cdot x$ (also $\widetilde{a_{ij}} = \frac{1}{2}(a_{ij} + a_{ji})$, $i \neq j$, $\widetilde{a_{ii}} = a_{ii}$). Nach Korollar 6.31 existiert $B = (b_{ij}) \in GL_n(\mathbb{K})$ mit $A = B^t \cdot diag(c_1, \ldots, c_n) \cdot B$ für geeignete $c_1, \ldots, c_n \in \mathbb{K}$. Einsetzen ergibt $q(x) = x^t \cdot A \cdot x =$

$$x^t \cdot \left(B^t \cdot \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} \cdot B \right) \cdot x = (Bx)^t \cdot \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix} \cdot (Bx) = \sum_{i=1}^{n} c_i \dot{(}Bx)_i^2.$$

□

Nun zum Beweis von 6.29:

Definition 6.33.
Zwei Vektoren $v, w \in V$ heißen **orthogonal** bezüglich der symmetrischen Bilinearform b, in Zeichen $v \perp w$, wenn $b(v, w) = 0$ ist.

Zwei Teilmengen M, N von V heißen **orthogonal**, in Zeichen $M \perp N$, falls $v \perp w$ für alls $v \in M, w \in N$.

Beachte:
Aus $v \perp w$ folgt $q_b(v + w) = q_b(v) + q_b(w)$.

Definition 6.34.
Zu jedem Unterraum U von V heißt $U^\perp := \{v \in V : v \perp U\}$ der zu U **orthogonale Unterraum** von V.
($U^\perp$ ist ein Unterraum, denn $v \perp w, v' \perp w \Rightarrow av + a'v' \perp w$)

Bemerkung 6.35.

1. Der UR $V^\perp = \{v \in V : \forall w \in V v \perp w\}$ heißt der **Ausartungsraum** von b. Genau dann ist b ausgeartet, wenn $V^\perp \neq \{0\}$ ist.

2. Für jeden UR U von V gilt: $U \cap U^\perp$ ist der Ausartungsraum von $b|_{U\times U}$.
   Also folgt: $b|_{U\times U}$ ist genau dann nicht-ausgeartet, wenn $U \cap U^\perp = \{0\}$.

Lemma 6.36.
Sei $\dim(V) < \infty$. *Für jeden UR U von V gilt: $b|_{U\times U}$ nicht-ausgeartet $\Leftrightarrow V = U \oplus U^\perp$.*
*In dieser Situation nennt man $U^\perp$ das **orthogonale Komplement** von U.*

Beweis: „$\Leftarrow$“ :klar nach vorher, wegen $U \cap U^\perp = \{0\}$.
„$\Rightarrow$“ : Sei $U \cap U^\perp = \{0\}$. Zu zeigen ist $V = U + U^\perp$.
Sei dazu $v \in V$. Die zu $b|_{U\times U}$ gehörende lineare Abbildung $l := l_{b|_{U\times U}} : U \to U^*$ ist bijektiv wegen $b|_{U\times U}$ nicht-ausgeartet und $\dim(U) < \infty$. Also gibt es ein $u' \in U$ mit $b(v, u) = b(u', u)$ für alle $u \in U$. Es folgt $v - u' \perp U \Rightarrow v = \underbrace{(v - u')}_{\in U^\perp} + \underbrace{u'}_{\in U}$.

□

Beweis: (von Theorem 6.29)
Induktion nach $\dim(V)$, der Beginn $\dim(V) = 1$ ist trivial.
Ist $b \equiv 0$ (also $b(v,w) = 0\ \forall v,w$), so fertig. Andernfalls gibt es $v \in V$ mit $b(v,v) \neq 0$.
Betrachte den Unterraum $U := \mathbb{K}v$ von V. Die Einschränkung $b|_{U\times U}$ ist nichtausgeartet. Nach Lemma 6.36 folgt $V = U \oplus U^\perp$.
Wegen $\dim(U^\perp) < \dim(V)$ hat die Einschränkung $b|_{U^\perp\times U^\perp}$ eine O-Basis $(v_2,\dots,v_n)$.
Also ist $(v, v_2, \dots, v_n)$ eine O-Basis von V.

□

Beispiel 6.37.
Rechenverfahren:
Sei $A \in Sym_n(\mathbb{K}), A = (a_{ij})$. Durch elementare Zeilen- und Spaltentransformationen führt man A schrittweise in eine Diagonalmatrix über. Alle diese Transformationen müssen simultan von links und von rechts durchgeführt werden, d.h. mit jeder Zeilentransformation muss auch die entsprechende Spaltentransformation einher gehen und umgekehrt.
Genauer: ist $a_{11} \neq 0$, so erreicht man durch Addition von geeigneten Vielfachen der ersten Zeile und der ersten Spalte Nullen an den Stellen $(i,1)$ und $(1,i)$, $i = 2,\dots,n$.
Ist $a_{11} = 0$, aber $a_{ii} \neq 0$ für ein i, so vertausche zuerst S_1 und S_i, sowie Z_1 und Z_i.
Ist $a_{ii} = 0$ für alle $i = 1,\dots,n$ (aber $A \neq 0$), so wähle (i,j) mit $a_{ij} = 0$. Ersetze Z_i durch $Z_i + Z_j$ und S_i durch $S_i + S_j$; dann wird an der Stelle (i,i) eine Zahl $\neq 0$ erreicht.
Ist $a_{1i} = a_{i1} = 0$, $i = 2,\dots,n$, so mache mit der $(n-1)\times(n-1)$-Matrix rechts unten weiter.

Beispiel:
$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{pmatrix} \in Sym_4(\mathbb{R}).$$
Wollen $S \in GL_4(\mathbb{R})$ finden mit SAS^t = diagonal:
beginne mit 2 Spalten, links mit A, rechts mit I.

$$\begin{array}{cccc|cccc} 0 & 1 & 2 & 3 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}$$
$(q_{12}(1))$
$$\begin{array}{cccc|cccc} 2 & 1 & 3 & 5 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 2 & 0 & 1 & 0 & 0 \\ 3 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 5 & 2 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}$$
$(diag(1,2,2,2))$

$$\begin{matrix} 2 & 2 & 6 & 10 \\ 2 & 0 & 4 & 8 \\ 6 & 4 & 0 & 4 \\ 10 & 8 & 4 & 0 \end{matrix} - \begin{matrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{matrix}$$

$(q_{21}(-1), q_{31}(-3), q_{41}(-5))$

$$\begin{matrix} 2 & 0 & 0 & 0 \\ 0 & -2 & -2 & -2 \\ 0 & -2 & -18 & -26 \\ 0 & -2 & -26 & -50 \end{matrix} - \begin{matrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -3 & -3 & 2 & 0 \\ -5 & -5 & 0 & 2 \end{matrix}$$

$(q_{32}(-1), q_{42}(-1))$

$$\begin{matrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -16 & -24 \\ 0 & 0 & -28 & -48 \end{matrix} - \begin{matrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -2 & -4 & 2 & 0 \\ -4 & -6 & 0 & 2 \end{matrix}$$

$(q_{43}(-\frac{3}{2}))$

$$\begin{matrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -16 & 0 \\ 0 & 0 & 0 & -12 \end{matrix} - \begin{matrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -2 & -4 & 2 & 0 \\ -1 & 0 & -3 & 2 \end{matrix}$$

Es folgt: für $S = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -2 & -4 & 2 & 0 \\ -1 & 0 & -3 & 2 \end{pmatrix}$ ist $SAS^t = diag(2, -2, -16, -12)$.

Für die symmetrische Bilinearform $Ab_A(x, y) = x^t \cdot A \cdot y$ $(x, y \in \mathbb{R}^n)$ bilden also die Zeilen von S eine Orthogonal-Basis $\mathcal{B}$ von b_A mit $M_{\mathcal{B}}(b_A) = diag(2, -2, -16, -12)$.

Bemerkung 6.38.

1. Bei einer Diagonalmatrix kann man die Diagonale mit Quadraten multiplizieren, ohne die $\simeq$-Klasse zu ändern:
   $diag(a_1, \ldots, a_n) \simeq diag(a_1 c_1^2, \ldots, a_n c_n^2)$ für alle $c_1, \ldots, c_n \in \mathbb{K}^*$. Das entspricht $A \rightsquigarrow SAS^t$ mit $S := diag(c_1, \ldots, c_n)$.

2. Fall $\mathbb{K} = \mathbb{C}$: hier ist jedes $A \in Sym_n(\mathbb{C})$ kongruent zu einer Matrix $diag(\underbrace{1, \ldots, 1}_{r}, 0, \ldots, 0)$; dabei ist $r = rk(A)$. Wir sehen:

   Für $A, B \in Sym_n(\mathbb{C})$ gilt: $A \simeq B \Leftrightarrow rk(A) = rk(B) \Leftrightarrow A \sim B$.

3. Ist $\mathbb{K} = \mathbb{R}$, so ist jedes $A \in Sym_n(\mathbb{R})$ kongruent zu einer Matrix

$diag(\underbrace{1,\dots,1}_{r},\underbrace{-1,\dots,-1}_{s},\underbrace{0,\dots,0}_{n-r-s})$; dabei ist $r+s=rk(A)$. Tatsächlich sind r und s eindeutig bestimmt (siehe unten).

Definition 6.39.
Sei $\mathbb{K}=\mathbb{R}$, sei V ein $\mathbb{R}$-VR.

(a) Eine symmetrische Bilinearform $b: V\times V\to\mathbb{R}$ heißt **positiv definit**, wenn $\forall 0\neq v\in V\ b(v,v)>0$ und **positiv semidefinit**, wenn $\forall v\in V\ b(v,v)\geq 0$.
b heißt **negativ (semi-)definit**, wenn $-b$ positiv (semi-)definit ist.

(b) Eine Matrix $A\in Sym_n(\mathbb{R})$ heißt **positiv definit**, wenn für alle $0\neq x=(x_1,\dots,x_n)\in\mathbb{R}^n$ gilt:
$x^tAx=\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j>0$.
(Analog semidefinit, negativ (semi-)definit).

Bemerkung 6.40.

1. Ist $\dim(V)<\infty$, $\mathcal{B}$ eine Basis von V, so gilt:
b ist positiv definit $\Leftrightarrow M_{\mathcal{B}}(b)$ ist positiv definit.

2. Jede definite symmetrische Bilinearform ist nicht-ausgeartet (Umkehrung ist falsch!).

3. Die (Semi-)Definitheit von $A\in Sym_n(\mathbb{R})$ hängt nur von der Kongruenzklasse von A ab.
Ist also $A=diag(c_1,\dots,c_n)$, so gilt:
A ist positiv definit $\Leftrightarrow c_1>0,\dots,c_n>0$. Also folgt:

Korollar 6.41.
$A\in Sym_n(\mathbb{R})$ ist genau dann positiv definit, wenn $A\simeq I_n$, also genau dann, wenn $A=S\cdot S^t$ mit $S\in GL_n(\mathbb{R})$ ist.

Theorem 6.42. *(Sylvester'scher Trägheitssatz)*
Sei V ein $\mathbb{R}$-VR, $\dim(V)<\infty$, sei $b: V\times V\to\mathbb{R}$ eine symmetrische Bilinearform.

(a) Es gibt eine Zerlegung $V = V_0 \oplus V_+ \oplus V_-$ derart, dass V_0, V_+, V_- paarweise orthogonal bezüglich b sind und gilt: $b|_{V_0 \times V_0} \equiv 0$, $b|_{V_+ \times V_+}$ ist positiv definit, $b|_{V_- \times V_-}$ ist negativ definit.

(b) Dabei ist der UR V_0, sowie $\dim(V_+)$ *und* $\dim(V_-)$*, eindeutig bestimmt.*

Beweis:

(a) Ist $(v_1, \ldots, v_n)$ eine Orthogonal-Basis von V bezüglich b (6.29), und ist nach Umnummerieren etwa $b(v_i, v_i) \begin{cases} = 0 & \text{für } i = 1, \ldots, l \\ > 0 & \text{für } i = l+1, \ldots, m \\ < 0 & \text{für } i = m+1, \ldots, n \end{cases}$
so setze $V_0 := span(v_1, \ldots, v_l)$, $V_+ := span(v_{l+1}, \ldots, v_m)$, $V_- := span(v_{m+1}, \ldots, v_n)$ eine Zerlegung wie in (a).

(b) Seien $V = V_0 \oplus V_+ \oplus V_- = \widetilde{V_0} \oplus \widetilde{V_+} \oplus \widetilde{V_-}$ zwei Zerlegungen wie in (a). Es ist $V_0 = V^\perp = \{v \in V : \forall w \in V \; b(v,w) = 0\}$., insbesondere auch $\widetilde{V_0} = V^\perp = V_0$.
Für alle $v \in V_0 \oplus V_-$ ist $b(v,v) \leq 0$. Also ist $(V_0 \oplus V_-) \cap \widetilde{V_+} = \{0\}$. Es folgt $\dim(\widetilde{V_+}) \leq \dim(V_+)$ wegen $V = (V_0 \oplus V_-) \oplus V_+$. Symmetrisches Argument gibt auch $\dim(V_+) \leq \dim(\widetilde{V_+})$.
Also $\dim(V_+) = \dim(\widetilde{V_+})$, also auch $\dim(V_-) = \dim(\widetilde{V_-})$.

□

Bemerkung 6.43.

1. Die URe V_+ und V_- selbst sind im Allgemeinen *nicht* eindeutig bestimmt: betrachte $V = \mathbb{R}^2$ und $b : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $b(x,y) := x_1y_2 + x_2y_1$: $b(x,x) = 2x_1x_2$. Es ist für $a > 0$:
$V_+ := \mathbb{R}(1,a)$, $V_- := \mathbb{R}(1,-a)$.

Korollar 6.44.
$A \in Sym_n(\mathbb{R})$. Dann ist A kongruent zu genau *einer Matrix der Form* $diag(\underbrace{1, \ldots, 1}_{r}, \underbrace{-1, \ldots, -1}_{s}, \underbrace{0, \ldots, 0}_{n-r-s})$ *mit $r, s \geq 0$ und $r + s \leq n$. Dabei ist $r + s = rk(A)$.*

Beweis:
Betrachte die symmetrische Bilinearform $b_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $b_A(x, y) = x^t \cdot A \cdot y$ und wende das Theorem von Sylvester () an.

□

Definition 6.45.

(a) Für $A \in Sym_n(\mathbb{R})$ heißt $sign(A) := r - s$ (mir r, s wie in 6.44) die **(Silvester-)Signatur** von A.

(b) In der Situation von b heißt $sign(b) := \dim(V_+) - \dim(V_-)$ die **(Sylvester-)Signatur** von b.

Bemerkung 6.46.

(a) Ist $A = diag(a_1, \ldots, a_n)$, so ist $sign(A) = \sum_{i=1}^{n} \operatorname{sgn}(a_i)$

(b) $A, B \in Sym_n(\mathbb{R})$, dann:
$A \simeq B \Leftrightarrow rk(A) = rk(B)$ und $sign(A) = sign(B)$.

(c) Die Matrix A aus Beispiel 6.37 hat $sign(A) = 1 - 3 = -2$.

Definition 6.47.

Sei $A = (a_{ij}) \in M_n(\mathbb{K})$. Dann bezeichnet $d_r(A) := \begin{vmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \ldots & a_{rr} \end{vmatrix}$ $(r = 1, \ldots, n)$ den r-ten **Hauptminor** von A.

Notation:
Ist $x = (x_1, \ldots, x_n)$ eine Folge von reellen Zahlen $x_i \neq 0$, so bezeichne $V(x) = V(x_1, \ldots, x_n)$ die Zahl der Vorzeichenwechsel in x, also der $i \in \{1, \ldots, n-1\}$ mit $x_i x_{i-1} < 0$.

Theorem 6.48.
Sei $A \in M_n(\mathbb{R})$ symmetrisch mit $d_r(A) \neq 0$, $r = 1, \ldots, n$. Dann ist die Anzahl der negativen Eigenwerte in einer Diagonalisierung 6.44 von A gleich $v := V(1, d_1(A), \ldots, d_n(A))$. Es ist also $sign(A) = n - 2v$

Beweis:
Induktion nach n, der Fall $n = 1$ ist klar.
Sei $b(x,y) := b_A(x,y) = x^t \cdot A \cdot y$ $(x,y \in \mathbb{R}^n)$. Sei $U := span(e_1,\dots,e_{n-1}) = \{x \in \mathbb{R}^n : x_n = 0\}$, sei $b' := b|_{U\times U}$. Für $\mathcal{B}' := (e_1,\dots,e_{n-1})$ ist $M'_{\mathcal{B}}(b') = \begin{pmatrix} a_{11} & \dots & a_{1,n-1} \\ \vdots & & \vdots \\ a_{n-1,1} & \dots & a_{n-1,n-1} \end{pmatrix}$
Mit $v' := V(1, d_1(A),\dots,d_{n-1}(A))$ ist nach Induktion $sign(b') = (n-1) - 2v'$.
Wegen $d_{n-1}(A) \neq 0$ ist b' nicht-ausgeartet. Sei $w \in U^\perp$, $w \neq 0$, es ist $w \notin U$ wegen b' nicht ausgeartet. Es ist $\mathbb{R}^n = U \oplus \underbrace{\mathbb{R}w}_{U^\perp}$. Bezüglich $(e_1,\dots,e_{n-1},w)$ hat b die Matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1,n-1} & 0 \\ \vdots & & \vdots & \vdots \\ a_{n-1,1} & \dots & a_{n-1,n-1} & 0 \\ 0 & \dots & 0 & c \end{pmatrix} \text{ mit } c := b(ww) \neq 0.$$

Es folgt: $sign(A) = sign(b) = sign(\widetilde{A}) = n - 1 - 2v' + \operatorname{sgn}(c)$, also

$$sign(A) = \begin{cases} n - 2v' & , c > 0 \\ n - 2v' - 2 & , c < 0 \end{cases} \qquad (*)$$

Andererseits ist $\widetilde{A} \simeq A$, also $\det(\widetilde{A}) = \gamma^2 \cdot \det(A)$ mit $\gamma \neq c$, $\gamma \in \mathbb{R}$, somit wegen $\det(\widetilde{A}) = d_{n-1}(A) \cdot c$: $\operatorname{sgn} d_n(A) = c \cdot \operatorname{sgn} d_{n-1}(A)$.

Also $v = \begin{cases} v' & , c > 0 \\ v' + 1 & , c < 0 \end{cases}$, also $sign(A) = n - 2v$ nach (*).

□

Korollar 6.49.
$A \in Sym_n(\mathbb{R})$ ist genau dann positiv definit, wenn $d_r(A) > 0$ für $r = 1,\dots,n$ ist.

Beweis:
Ist A positiv definit, so ist $A = S^t \cdot S$ mit $S \in GL_n(\mathbb{R})$ (6.41), also $\det(A) = \det(S)^2 > 0$. Da mit b_A auch jede Restriktion von b_A auf einen Unterraum von $\mathbb{R}^n$ positiv definit ist, ist auch $d_r(A) > 0$ für $r = 1,\dots,n$.

Umkehrung sofort aus 6.48

□

Theorem 6.50.
Sei $A \in M_n(\mathbb{R})$ symmetrisch, sei $\det(A) \neq 0$. Für alle $r = 1,\dots,n-1$ sei $d_r(A) \neq 0$ oder $d_{r+1} \neq 0$. Sei $d'(A)$ die Folge, die aus $(1, d_1(A),\dots,d_n(A))$ durch Streichen aller Nullen entsteht, sei $v' := V(d'(A))$. Dann gilt $sign(A) = n - 2v'$.

Beweis:
Wieder Induktion nach n.
Ist $d_{n-1} \neq 0$, so funktioniert der vorige Beweis. Sei also $d_{n-1} = 0$.
Betrachte die Restriktion von b auf $W := span(e_1, \ldots, e_{n-2})$. Wegen $d_{n-2} \neq 0$ ist $\mathbb{R}^n = W \oplus W^\perp$.
$b|_{W^\perp \times W^\perp}$ ist nicht-ausgeartet und muss indefinit sein wegen $d_{n-1} = 0$. Es ist also $sign(b) = sign(b|_{W \times W}$. Andererseits ist $d_n(A) \cdot d_{n-2}(A) < 0$.
$\Rightarrow$ Behauptung (...)

□

Beispiele 6.51.

1. Betrachte noch einmal
   $A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{pmatrix} \in M_4(\mathbb{R})$, hatten gesehen $A \simeq diag(2, -2, -16, -12)$, also $sign(A) = -2$.
   Die Hauptminoren von A sind $d_1 = 0$, $d_2 = -1$, $d_3 = 4$, $d_4 = -12$, also ist $v = V(1, 0, -1, 4, -12) = 3$: OK.

2. Für $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$ ist $sign(A) = 2$, $sign(-A) = -2$, aber A und $-A$ haben dieselben Hauptminoren $0, 0, 0, -1$.

Definition 6.52.

(a) Sei $\mathbb{K}$ ein Körper, sei V ein $\mathbb{K}$-Vektorraum. Eine Bilinearform $b : V \times V \to \mathbb{K}$ heißt **alternierend**, wenn $b(v, v) = 0 \; \forall v \in V$ ist.

(b) Eine Matrix $A \in M_n(\mathbb{K})$ heißt **schiefsymmetrisch**, falls $A^t = -A$ ist.

Lemma 6.53.
Sei $b \in Bil(V)$, also $b : V \times V \to \mathbb{K}$ bilinear.

(a) Ist b alternierend, so gilt $b(w, v) = -b(v, w) \; \forall v, w \in V$.
Ist $char(\mathbb{K}) \neq 2$, so gilt auch die Umkehrung.

(b) Ist $\dim(V) < \infty$, *$\mathcal{B}$ eine Basis von V und char*$(\mathbb{K}) \neq 2$, *so gilt:*
b ist alternierend $\Leftrightarrow$ $M_{\mathcal{B}}(b)$ ist schiefsymmetrisch.

Beweis:

(a) Sei b alternierend. Berechne $0 = b(v+w, v+w) = \underbrace{b(v,v)}_{=0} + b(v,w) + b(w,v) + \underbrace{b(w,w)}_{=0} = b(v,w) + b(w,v)$.
$b(v,v) = -b(v,v)$ heißt $2b(v,v) = 0$, impliziert $b(v,v) = 0$, falls $2 \neq 0$.

□

$\simeq$-Klassen von schiefsymmetrischen Matrizen ($A \simeq B \Leftrightarrow \exists S \in GL_n(\mathbb{K})$: $B = S^t A S$)

Satz 6.54.
Sei char$(\mathbb{K}) \neq 2$, *sei b eine alternierende Form auf V,* $\dim(V) < \infty$.

Dann hat V eine Basis $\mathcal{B}$ mit $M_{\mathcal{B}}(b) = \begin{pmatrix} 0 & -1 & & & & & \\ 1 & 0 & & & & & \\ & & \ddots & & & & \\ & & & 0 & -1 & & \\ & & & 1 & 0 & & \\ & & & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & 0 \end{pmatrix}$ *(*)*

Beweis:
Ähnlich zum Beweis von 6.29: ist $b \equiv 0$, so klar. Andernfalls gibt es $v, w \in V$ mit $b(v,w) \neq 0$. Wegen b alternierend sind v, w linear unabhängig. Sei $c := b(v,w)$, sei $U = \mathbb{K}v + \mathbb{K}w$. Die Restriktion von b auf $U \times U$ hat bezüglich der Basis $(\frac{1}{c}w, v)$ die Matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Habe $U \oplus U^\perp = V$.
Nun betrachte $b|_{U^\perp \times U^\perp}$ und beende mit Induktion.

□

Korollar 6.55.
Sei $A \in M_n(\mathbb{K})$ schiefsymmetrisch, sei char$(\mathbb{K}) \neq 2$. *Dann ist*

$$A \simeq \begin{pmatrix} 0 & -1 & & & & & & 0 \\ 1 & 0 & & & & & & \\ & & \ddots & & & & & \\ & & & 0 & -1 & & & \\ & & & 1 & 0 & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ 0 & & & & & & & 0 \end{pmatrix} \text{ mit } r \text{ Kästchen } \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \text{ Dabei ist } 2r = rk(A).$$

Korollar 6.56.
Schiefsymmetrische Matrizen haben geraden Rang.

c. Skalarprodukte

Bemerkung 6.57.

1. (Längenmessung im $\mathbb{R}^n$) Die euklidische Metrik im $\mathbb{R}^n$ ist $d(x,y) = \|x-y\|$ mit $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2} = \sqrt{\langle x,x\rangle}$ mit $\langle x,y\rangle := x^t y = \sum_{j=1}^{n} x_j y_j$ die kanonische symmetrische Bilinearform auf $\mathbb{R}^n$.

2. (Winkelmessung im $\mathbb{R}^n$) Man hat den Begriff des (unorientierten) Winkels zwischen Vektoren $0 \neq x, y \in \mathbb{R}^n$.
   $\alpha = \sphericalangle(x,y)$.
   $\cos(\alpha) = \frac{\|y'\|}{\|y\|}$.
   Es ist $y' = \frac{\langle x,y\rangle}{\langle x,x\rangle} \cdot x$, denn $\langle y - y', x\rangle = 0$ also $\cos(\alpha) = \frac{|\langle x,y\rangle|}{\langle x,x\rangle} \cdot \frac{\|x\|}{\|y\|} = \frac{|\langle x,y\rangle|}{\|x\|\cdot\|y\|}$.
   $y' = ax$ $\langle y - y', x\rangle = 0$.
   Insbesondere $\langle x,y\rangle = 0 \Leftrightarrow x \perp y$.

3. Auch im $\mathbb{C}^n$ hat man eine kanonische euklidische Metrik, nämlich die aus $\mathbb{C}^n \approx \mathbb{R}^{2n}$.
   Für $w, z \in \mathbb{C}^n$ ist $d(w,z) = \|w-z\|$ mit $\|z\| = \sqrt{|z_1|^2 + \ldots + |z_n|^2} = \sqrt{\langle z,z\rangle}$.
   Dabei wurde $\langle w,z\rangle := \sum_{j=1}^{n} w_j \cdot \overline{z_j}$ gesetzt. Die Form $\langle -,-\rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ ist nicht $\mathbb{C}$-bilinear, denn für $\lambda \in \mathbb{C}$ ist $\langle w, \lambda z\rangle \neq \lambda \cdot \langle w,z\rangle$. Auch ist $\langle z,w\rangle = \overline{\langle w,z\rangle}$.

Definition 6.58. (Sesquilinearform, hermitesche Form)
Sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $h : V \times V \to \mathbb{C}$ heißt **Sesquilinearform**, wenn für alle $v, v_1, v_2 \in V$, $a \in \mathbb{C}$ gilt:

1. $h(v_1 + v_2, v) = h(v_1, v) + h(v_2, v)$;
   $h(v, v_1 + v_2) = h(v, v_1) + h(v, v_2)$;

2. $h(av_1, v_2) = ah(v_1, v_2)$;
   $h(v_1, av_2) = \overline{a}h(v_1, v_2)$;

Gilt zusätzlich

1. $h(v_1, v_2) = \overline{h(v_2, v_1)}$ (bzw. $h(v_1, v_2) = -\overline{h(v_2, v_1)}$), so heißt h **hermitesch** (bzw. **schief-hermitesch**).

Bemerkung 6.59.

1. Man sagt h ist linear im ersten und antilinear im zweiten Argument.

2. Sei $V = \mathbb{C}^n, x, y \in V$.
   $\langle x, y\rangle = \sum_{j=1}^{n} x_j \overline{y_j}$
   heißt die **kanonische hermitesche Form** auf $\mathbb{C}^n$.

   Bemerkung:
   $\mathbb{C}^n \cong \mathbb{R}^{2n}$ als $\mathbb{R}$-Vektorraum.
   $x, y \in \mathbb{C}^n, x_j = a_j + b_j i, y_j = a'_j + b'_j i$
   $\langle x, y\rangle = \sum_{j=1}^{n} (a_j a'_j + b_j b'_j) = \sum_{j=1}^{n} x_j \overline{y_j}$.

3. Ist h eine hermitesche Form, $v \in V$, dann gilt $h(v, v) = \overline{h(v, v)}$. Also $h(v, v) \in \mathbb{R}$.
   Analog: Ist h schief-hermitesch, so folgt $h(v, v) \in i\mathbb{R}$.

4. Sei h eine hermitesche Form, auf V. Dann heißt h **positiv definit**, wenn für alle $v \in V, v \neq 0$ gilt: $h(v, v) > 0$.
   Analog: **positiv semidefinit, negativ (semi-)definit**; vgl. 6.39.

Definition 6.60.
Sei V ein $\mathbb{C}$-VR mit $\dim_{\mathbb{C}}(V) = n < \infty$, sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V, sei h eine Sesquilinearform auf V. Dann heißt $A := (h(v_j, v_k))_{j,k=1,\ldots,n} \in M_n(\mathbb{C})$ die Matrix von h bezüglich $\mathcal{B}$.
Wir schreiben dafür $A = M_{\mathcal{B}}(h)$.

Bemerkung 6.61.

1. Seien $v = \sum_{j=1}^{n} x_j v_j, w = \sum_{k=1}^{n} y_k v_k$, dann ist
   $$h(v, w) = h\left(\sum_{j=1}^{n} x_j v_j, \sum_{k=1}^{n} y_k v_k\right) = \sum_{j,k=1}^{n} x_j \overline{y_k} h(v_j, v_k) = \sum_{j,k=1}^{n} x_j \overline{y_k} a_{jk}. \qquad (*)$$

Umgekehrt ist $A \in M_n(\mathbb{C})$, so definiert (*) eine Sesquilinearform auf V.
Kurz: $v, w \in \mathbb{C}^n : h(v,w) = v^t A\overline{w}$.

2. Dabei gilt: ist h hermitesch, dann gilt: $A = \overline{A^t}$, bzw. falls h schief-hermitesch, so $A = -A^t$.

Definition 6.62.
Eine Matrix $A \in M_n(\mathbb{C})$ heißt hermitesch, falls $A = \overline{A^t}$, bzw. schief-hermitesch, falls $A = -\overline{A^t}$.
Entsprechend: Definitheitseigenschaften.

Satz 6.63.
Sei $\dim_{\mathbb{C}}(V) < \infty$, *sei* h *eine Sesquilinearform auf* V *und seien* $\mathcal{B}, \mathcal{C}$ *Basen von* V. $S = M^{\mathcal{C}}_{\mathcal{B}}(id) = T^{\mathcal{C}}_{\mathcal{B}}$.
Dann gilt:

$$M_C(h) = S^t \cdot M_{\mathcal{B}}(h) \cdot \overline{S}.$$

Beweis: Analog zu Satz 6.17.

□

Von nun an: Bezeichne $\mathbb{K}$ *den Körper* $\mathbb{R}$ *oder* $\mathbb{C}$.

Definition 6.64.
Sei V ein $\mathbb{K}$-Vektorraum.

(a) Ein **Skalarprodukt** auf V ist eine positiv definite symmetrische Bilinearform (falls $\mathbb{K} = \mathbb{R}$) bzw. eine positiv definite hermitesche Form (falls $\mathbb{K} = \mathbb{C}$),

(b) Ein Paar $(V, \langle\ ,\ \rangle)$ aus einem $\mathbb{K}$-Vektorraum V und einem Skalarprodukt $\langle\ ,\ \rangle$ heißt ein **Prähilbertraum** über $\mathbb{K}$.

Üblicher: Für $\mathbb{K} = \mathbb{R}$: euklidischer Vektorraum,
Für $\mathbb{K} = \mathbb{C}$: unitärer Vektorraum.

Definition 6.65.
Sei V ein Prähilbertraum. Dann ist die **Norm** von $x \in V$ definiert als $\|x\| = \sqrt{\langle x, x\rangle}$.

Beispiel 6.66.

1. Das kanonische Skalarprodukt auf $\mathbb{K}^n$ ist $\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j$ (falls $\mathbb{K} = \mathbb{R}$), bzw.
   $\langle x, y\rangle = \sum_{j=1}^{n} x_j \overline{y_j}$, (falls $\mathbb{K} = \mathbb{C}$).
   $\|x\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$ heißt **euklidische Standardnorm**.

2. Seien $a, b \in \mathbb{R}, a < b$, sei $V = C([a, b], \mathbb{K})$ der $\mathbb{K}$-Vektorraum der stetigen Funktionen $[a, b] \to \mathbb{K}$.
   Definiere ein Skalarprodukt $\langle\ ,\ \rangle$ auf B wie folgt:
   Für $f, g \in V$: $\langle f, g\rangle = \int_a^b f(t)g(t)\,dt$ (falls $\mathbb{K} = \mathbb{R}$), bzw.
   $\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\,dt$ (falls $\mathbb{K} = \mathbb{C}$).

   Auf diese Weise wird V zu einem Prähilbertraum. V ist jedoch nicht vollständig (vgl. Analysis).

   Beispiel:
   $\mathbb{K} = \mathbb{R}$:
   Sei $f_n(x) = \sum_{j=1}^{n} \frac{\sin(jx)}{j}; n \in \mathbb{N}, x \in [-\pi, \pi], a = -b = -\pi$.
   $(f_n)_{n\in\mathbb{N}}$ konvergiert punktweise gegen
   $$g(x) := \begin{cases} \frac{\pi - x}{2} & \text{für } x \in (0, \pi] \\ 0 & \text{für } x = 0 \\ -\frac{\pi + x}{2} & \text{für } x \in [\pi, 0). \end{cases}$$
   (Analysis I, Satz 10.6 (b))
   Es folgt:
   Es gibt kein $f \in V$ mit $f_n \to f$ bezüglich $\|\cdot\|$. Dann für ein solches f müsste gelten: $\int_a^b (f(t) - g(t))^2\,dt = 0$.

3. Etwas allgemeiner: Ist $w : [a, b] \to \mathbb{R}$ stetig mit $w(x) > 0\ \forall x \in [a, b]$, dann ist
   $\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}w(t)\,dt$.
   w heißt Gewichtsfunktion.

4. Sei $x = (x_i)_{i\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$. Setze $\|x\|^2 = \sum_{j=1}^{\infty} |x_j|^2 \in \mathbb{R} \cup \{\infty\}$.
   Man definiert:

$l^2_{\mathbb{K}} = \{x \in \mathbb{K}^{\mathbb{N}} : ||x||^2 < \infty\}$ der **Hilbert'sche Folgenraum der quadratsummierbaren Folgen**.
Zeigen: $l^2_{\mathbb{K}} \subset \mathbb{K}^{\mathbb{N}}$ ist Untervektorraum. Seien $x, y \in l^2_{\mathbb{K}}, \lambda \in \mathbb{K}$.

(i) z.z.: $x + y \in l^2_{\mathbb{K}}$.
$\forall a, b \in \mathbb{K}$: $|a + b| \leq |a| + |b|$
$\Rightarrow \forall a, b \in \mathbb{K}$: $|a + b|^2 \leq |a|^2 + |b|^2 + 2|a|\,|b|$.
Außerdem gilt: $2|a|\,|b| \leq |a|^2 + |b|^2$.
Damit $|a + b|^2 \leq 2|a|^2 + 2|b|^2$.
Damit $||x + y||^2 \leq 2||x||^2 + 2||y||^2 \Rightarrow x + y \in l^2_{\mathbb{K}}$.

(ii) z.z.: $\lambda x \in l^2_{\mathbb{K}}$.
$||\lambda x||^2 = |\lambda|^2\,||x||^2$. Damit $\lambda x \in l^2_{\mathbb{K}}$.

Definiere für $x, y \in l^2_{\mathbb{K}}$.
$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k \cdot \overline{y_k} \in \mathbb{K}$.
Dies konvergiert absolut:
$2\sum_{k=1}^{\infty} |x_k\overline{y_k}| = 2\sum_{k=1}^{\infty} |x_k| \cdot |y_k| \leq \sum_{k=1}^{\infty} (|x_k|^2 + |y_k|^2) = ||x||^2 + ||y||^2$.
(wegen $(|x_k| - |y_k|)^2 = |x_k|^2 < 2|x_k|\,|y_k| + |y_k|^2$.

Es gilt insbesondere:
$||x||^2 = \langle x, x\rangle\ \forall x \in l^2_{\mathbb{K}}$)

Satz 6.67.
Der Raum $l^2_{\mathbb{K}}$ ist ein Prähilbertraum über $\mathbb{K}$ mit dem Skalarprodukt aus 6.66 4.

Satz 6.68.
Sei V ein Prähilbertraum über $\mathbb{K}$. Für alle $v, w \in V, \lambda \in \mathbb{K}$ gelten:

(i) $||v|| \geq 0$ *und* $||v|| = 0 \Leftrightarrow v = 0$.

(ii) $||\lambda v|| = |\lambda| \cdot ||v||$ *(Homogenität)*

(iii) $||v + w|| \leq ||v|| + ||w||$ *(Δ-Ungleichung)*

(iv) $||v + w||^2 = ||v||^2 + ||w||^2 + 2Re\,\langle v, w\rangle$.

(v) $||v + w||^2 + ||v - w||^2 = 2||v||^2 + 2||w||^2$ *(Parallelogrammidentität)*

Beweis:

(i) $||v|| = \sqrt{\langle v, v\rangle}$ und $\langle\ ,\ \rangle$ ist positiv definit.

(ii) $||\lambda v||^2 = \lambda\overline{\lambda}||v||^2 = |\lambda|^2||v||^2$. → Wurzelziehen ⇒ Behauptung.

(iii) später.

(iv) $||v+w||^2 = \langle v + w, v + w\rangle = ||v||^2+||w||^2+\langle v, w\rangle+\overline{\langle v, w\rangle} = ||v||^2+||w||^2+2Re\,\langle v, w\rangle$.

(v) $||v + w||^2 + ||v - w||^2 = 2||v||^2 + 2||w||^2 + 2Re\,\langle v, w\rangle + 2Re\,\langle v, -w\rangle = 2||v||^2 + 2||w||^2$.

□

Theorem 6.69. *(Cauchy-Schwarz Ungleichung)*
Sei V ein Prähilbertraum. Es gilt:

$$\forall v, w \in V : |\langle v, w\rangle| \leq ||v||\,||w||$$

Die Gleichheit gilt ⇔ v und w sind linear abhängig.

Beweis:
Fall 1: $w = 0$
$|\langle v, w\rangle| = 0 = ||v||\,||w||$.

Fall 2: $w \neq 0$
$a := \frac{\langle v,w\rangle}{||w||^2}$, $u := v - aw$.

Es ist $u \perp w$, weil
$\langle u, w\rangle = \langle v - aw, w\rangle = \langle v, w\rangle - a\langle w, w\rangle = \langle v, w\rangle - \langle v, w\rangle = 0.$

$||v||^2 = ||u + aw||^2 = ||u||^2 + ||aw||^2 + 2Re\langle u, aw\rangle$ (6.68 (d))
$= ||u||^2 + |a|^2||w||^2 = ||u||^2 + \frac{|\langle v,w\rangle|^2}{||w||^2}$
Nach Multiplikation
$||v||^2||w||^2 = ||u||^2||w||^2 + |\langle v, w\rangle|^2 \geq |\langle v, w\rangle|^2.$

Die Aussage folgt nach Quadratwurzel.

Gilt die Gleichheit $\Leftrightarrow ||u||^2||w||^2 = 0 \Rightarrow ||u|| = 0 \Leftrightarrow u = 0 \Leftrightarrow v = aw$.
Ist $v = aw$, so sind v und w linear abhängig.
Umgekehrt, sind v und w linear abhängig, so existiert $c \in \mathbb{K}$, so dass $V = cw$ (weil $w \neq 0$).
Jetzt ist $\langle v, w\rangle = \langle cw, w\rangle = c\langle w, w\rangle = c||w||^2$, und $c = \frac{\langle v,w\rangle}{||w||^2} = a$.

□

Satz 6.70. *(Dreiecksungleichung)*
Sei V ein Prähilbertraum.

$$\forall v, w \in V \; ||v + w|| \leq ||v|| + ||w||$$

Beweis:
$||v+w||^2 = ||v||^2+||w||^2+2Re\langle v, w\rangle \leq ||v||^2+||w||^2+2|\langle v, w\rangle| \leq ||v||^2+||w||^2+2||v||\,||w|| = (||v|| + ||w||)^2$.

□

Definition 6.71.
Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $||\cdot|| : V \to \mathbb{R}$ heißt eine Norm, falls:

(i) $\forall v \in V$: $||v|| = 0 \Leftrightarrow v = 0$.

(ii) $\forall v \in V, \forall c \in \mathbb{K}$: $||cv|| = |c|\,||v||$.

(iii) $\forall v, w \in V$: $||v + w|| \leq ||v|| + ||w||$.

Ein Paar $(V, ||\cdot||)$ bestehend aus einem $\mathbb{K}$-Vektorraum und einer Norm $||\cdot||$ heißt ein **normierter $\mathbb{K}$-Vektorraum**.

Lemma 6.72.
Sei V ein normierter $\mathbb{K}$-Vektorraum.

1. $d(v,w) := ||v-w||$ *ist eine Metrik auf* V.

2. $\forall v,w \in V \; | \, ||v|| - ||w|| \, | \leq ||v-w||$.

Insbesondere ist die Norm eine gleichmäßig stetige Abbildung.

Beweis:

1.
   - $d(v,w) = ||v-w|| \geq 0$
   - $d(v,w) = 0 \Leftrightarrow ||v-w|| = 0 \Leftrightarrow v-w = 0 \Leftrightarrow v = w$
   - $d(v,w) = ||v-w|| = ||(-1)(w-v)|| = |-1| \, ||w-v|| = ||w-v|| = d(w,v)$
   - $d(v,w) = ||v-w|| = ||(v-u)+(u-w)|| \leq ||v-u|| + ||u-w|| = d(v,u) + d(u,w)$

2. $||v|| = ||(v-w)+w|| \leq ||v-w|| + ||w|| \Rightarrow ||v|| - ||w|| \leq ||v-w||$.
   Ähnlich: $||w|| - ||v|| \leq ||w-v||$.
   Also: $-(||v|| - ||w||) = ||w|| - ||v|| \leq ||w-v|| = ||v-w||$.
   Daraus folgt die Aussage.

   Gleichmäßige Stetigkeit bedeutet:
   $\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall v,w \in V \; (d_V(v,w) < \delta \;\Rightarrow\; d_R(||v||, ||w||) < \varepsilon)$
   Wähle $\delta = \varepsilon$.
   $d_R(||v||, ||w||) = | \, ||v|| - ||w|| \, | \leq ||v-w|| < \delta = \varepsilon$.

□

Bemerkungen 6.73.

1. Ist V ein Prähilbertraum, so lässt sich das Skalarprodukt durch Normen ausdrücken.
   Für $\mathbb{K} = \mathbb{R}$:

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

   Für $\mathbb{K} = \mathbb{C}$: siehe Übungsblatt 3.3.

2. Auf einem $\mathbb{K}$-Vektorraum gibt es verschiedene Normen. Ist $\dim(V) < \infty$, so sind alle diese Normen äquivalent (d.h., sie induzieren dieselbe Topologie).

3. Es existieren auch wichtige nichtdefinite „Metriken".
   Z.B. $q : \mathbb{R}^4 \to \mathbb{R}$ $q(x_1, x_2, x_3, t) = x_1^2 + x_2^2 + x_3^2 - t^2$.

Definition 6.74.
Sei V ein Prähilbertraum.

1. Zwei Vektoren $v, w \in V$ heißen **orthogonal**, falls $\langle v, w\rangle = 0$; in Zeichen: $v \perp w$.
   Sind $M, N \subseteq V$ Teilmengen, dann heißen M und N **orthogonal**, falls $\forall m \in M$, $\forall n \in N$ $(m \perp n)$; in Zeichen $M \perp N$.

2. Eine Familie $(v_i)_{i\in I}$ heißt **orthogonal**, falls $\forall i, j : (i \neq j \;\Rightarrow\; v_i \perp v_j)$.

3. Eine Familie $(v_i)_{i\in I}$ heißt **orthonormal**, falls $\forall i, j : \left\langle v_i, v_j \right\rangle = \delta_{ij}$.

4. Eine Basis von V die eine orthonormale Familie ist, heißt eine **Orthonormal-Basis** (ON-Basis).

Bemerkungen 6.75.

1. $v \perp w \Leftrightarrow w \perp v$ weil $\langle v, w\rangle = \overline{\langle w, v\rangle}$.

2. Jede orthogonale Familie $(v_i)_{i\in I}$ mit $\forall i, v_i \neq 0$ ist linear unabhängig.
   $\sum_{i=1}^{n} a_i v_i = 0 \Rightarrow \forall j : 0 = \left\langle 0, v_j \right\rangle = \left\langle \sum_{i=1}^{n} a_i v_i, v_j \right\rangle = \sum_{i=1}^{n} a_i \left\langle v_i, v_j \right\rangle = a_j \|v_j\|^2 \Rightarrow \forall j :$
   $0 = a_j$.

3. Ist $(v_i)_{i\in I}$ eine orthogonale Familie, die $(\forall i : v_i \neq 0)$ erfüllt, so ist die Familie $(w_i)_{i\in I}$ $w_i = \frac{v_i}{\|v_i\|}$ orthonormal.

$$\left\langle w_i w_j \right\rangle = \tfrac{1}{\|v_i\|}, \|v_j\| \left\langle v_i, v_j \right\rangle = \begin{cases} 0 & , i \neq j \\ \frac{1}{\|v_i\|^2} = 1 & , i = j. \end{cases}$$

4. Geometrische Bedeutung von $\langle\ ,\ \rangle$ für $\mathbb{K} = \mathbb{R}$:
Aus Cauchy-Schwarz folgt:

$$-1 \leq \tfrac{\langle v,w\rangle}{\|v\|\,\|w\|} \leq 1$$

Man definiert

$$\cos(v, w) = \tfrac{\langle v,w\rangle}{\|v\|\,\|w\|}$$

Aus 6.68 (d) folgt: $\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\cos(v, w)\|v\|\,\|w\|$.

Satz 6.76.
Sei V ein Prähilbertraum, $\dim(V) = n < \infty$. *Sei $(v_i)_{i=1,\dots,n}$ eine ON-Basis.*

1. $\forall v \in V\ v = \sum_{i=1}^{n} \langle v, v_i\rangle\, v_i$
2. $\forall v, w \in V : \langle v, w\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle\, \overline{\langle w, v_i\rangle}$ *(Parseval-Gleichung)*
3. $\forall v \in V : \|v\|^2 = \sum_{i=1}^{n} |\langle v, v_i\rangle|^2$ *(Bessel-Gleichung)*

Die Koeffizienten $\langle v, v_i\rangle$ heißen die Fourier-Koeffizienten von v bezüglich (v_i).

Beweis:
Sei $v = \sum_{i=1}^{n} a_i v_i$, $w = \sum_{i=1}^{n} b_i v_i$.

1.′ $a_i = \langle v, v_i\rangle$

2.′ $\langle v, w\rangle = \sum_{i=1}^{n} a_i \overline{b_i}$

3.′ $\|v\|^2 = \sum_{i=1}^{n} |a_i|^2$

1.′ $\langle v, v_i\rangle = \left\langle \sum_{j=1}^{n} a_j v_j, v_i \right\rangle = \sum_{j=1}^{n} a_j \left\langle v_j, v_i \right\rangle = \sum_{j=1}^{n} a_j \delta_{ij} = a_i.$

2.′ $\langle v, w\rangle = \left\langle \sum_{i=1}^{n} a_i v_i, \sum_{j=1}^{n} b_j v_j \right\rangle = \sum_{i,j=1}^{n} a_i \overline{b_i} \underbrace{\left\langle v_i, v_j \right\rangle}_{\delta_{ij}} = \sum_{i=1}^{n} a_i \overline{b_i}.$

3.′ $\|v\|^2 = \langle v, v\rangle = \sum_{i=1}^{n} a_i \overline{a_i} = \sum_{i=1}^{n} |a_i|^2$

□

Korollar 6.77.
Sei $f : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen Prähilberträumen. Seien $\mathcal{B} = (v_1, \ldots, v_n)$ und $C = (w_1, \ldots, w_m)$ ON-Basen von V und W. Ist $A := M_C^{\mathcal{B}}(f)$, so gilt

$$a_{ij} = \left\langle f v_j, w_i \right\rangle$$

Beweis:
Nach Definition: $f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$
Es folgt: $\left\langle f v_j, w_i \right\rangle = \left\langle \sum_{k=1}^{m} a_{kj} w_k, w_i \right\rangle = \sum_{k=1}^{m} a_{kj} \langle w_k, w_i\rangle = a_{ij}.$

□

Beispiele 6.78.

1. In $\mathbb{K}^n$ mit dem Standardskalarprodukt ist $(e_i)_{i=1,\ldots,n}$ eine ON-Basis.

2. Ist $(v_i)_{i\in I}$ eine ON-Basis, und sind $c_i \in \mathbb{K}$ mit $|c_i| = 1$, so ist $(c_i v_i)_{i\in I}$ auch eine ON-Basis.
   $\left\langle c_i v_i, c_j v_j \right\rangle = c_i \overline{c_j} \left\langle v_i, v_j \right\rangle = \begin{cases} 0 & , i \neq j \\ c_i \overline{c_i} = |c_i|^2 = 1 & , i = j \end{cases}$

3. $V = \{f \in C([0,2\pi],\mathbb{C}) | f(0)f(2\pi)\}$ mit Skalarprodukt.
   $\langle f,g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(x)\overline{g(x)}\,dx$
   Die Familie $f_n(x) = e^{inx}$ $(n \in \mathbb{R})$ ist eine ON-Familie, weil
   $$\langle f_n, f_m\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{i(m-n)x}\,dx = \begin{cases} \frac{1}{2\pi}\int_0^{2\pi} dx = 1 & ,m = n \\ \frac{1}{2\pi}\left[\frac{1}{i(m-n)}e^{i(m-n)x}\right]_0^{2\pi} = 0 & ,m \neq n \end{cases}$$
   $\langle f, f_n\rangle = \frac{i}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\,dx = \widehat{f(n)}$
   $\forall n : \widehat{f(n)} = 0 \Rightarrow f = 0.$

4. In $V = l^2_{\mathbb{K}} = \{(x_n)_{n\in\mathbb{N}} : x_n \in \mathbb{K}, \sum |x_n|^2 < \infty\}$ haben wir das Skalarprodukt $\langle x,y\rangle = \sum_{n=1}^{\infty} x_n\overline{y_n}$. In diesem Prähilbertraum bilden die Elemente $e_n := (0,\ldots,0,\underbrace{1}_{\text{Stelle } n},0,\ldots) = (\delta_{jn})_{j\in\mathbb{N}}$ für $n \in \mathbb{N}$ eine ON-Familie: $\|e_n\|^2 = 1$
   $\langle e_m, e_n\rangle = 0$ für $m \neq n$. Diese ON-Familie ist maximal, denn für alle $x = (x_n)_{n\in\mathbb{N}}$ ist $\langle x, e_n\rangle = x_n$ für $n \in \mathbb{N}$, also ist nur $x = 0$ auf allen e_n orthogonal.
   Andererseits ist $span(e_n : n \in \mathbb{N}) \neq l^2$, denn jede Folge $x = (x_n) \in l^2$ mit $x_n \neq 0$ für unendlich viele n liegt nicht in $span(e_n : n \in \mathbb{N})$.
   z.B. ist $x = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots)$ ein solches $x \in l^2$ (denn $\|x\|^2 = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty$).

KURZE WIEDERHOLUNG:
$\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.
Prähilbertraum V über $\mathbb{K}$: $V, \langle x,y\rangle$ $(x,y \in V)$
$\|x\| := \sqrt{x,x}$ $d(x,y) := \|x-y\|$

BEISPIELE:

1. $V = \mathbb{K}^n, \langle x,y\rangle = \sum_{j=1}^{n} x_j\overline{y_j}$
   $\|x\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2}.$

2. $V = \{C([0,2\pi],\mathbb{C}) \ni f : f(0) = f(2\pi)\}$
   $\langle f,g\rangle := \frac{1}{2\pi}\int_0^{2\pi} f(t)\cdot\overline{g(t)}\,dt$

$\|f\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 \, dt$

ON-Familien: $(v_j)_{j\in I}$ mit $\left\langle v_j, v_k \right\rangle = \delta_{jk}$
$f_n(x) = e^{nix}$, $(n \in \mathbb{Z})$

Ende Wiederholung

Definition 6.79.
Sei V ein Prähilbertraum. Für jede Teilmenge $M \subset V$ sei $M^\perp := \{v \in V : \langle v, w\rangle = 0 \; \forall \; w \in M\}$, der zu M orthogonale Unterraum von V.
Man setzt $M^{\perp\perp} := (M^\perp)^\perp$.
Es gilt: $M^\perp = span(M)^\perp$ und weiter

Lemma 6.80.
Seien $U, U_1, U_2, \ldots$ Unterräume von V.

(a) $U_1 \subset U_2 \Rightarrow U_1^\perp \supset U_2^\perp$,

(b) $\bigcap_j U_j^\perp = \left(\sum_j U_j\right)^\perp$,

(c) $U \cap U^\perp = \{0\}$, *also* $U + U^\perp = U \oplus U^\perp$,

(d) $U \subset U^{\perp\perp}$.

Bemerkung 6.81.
Die ON-Familie aus 6.78 4. erzeugt einen Unterraum U von V mit $U^{\perp\perp} = V = l^2$; aber $U \neq V$.
Im Allgemeinen ist also die Inklusion $U \subset U^{\perp\perp}$ strikt.

Definition 6.82.
Seien $(U_j)_{j\in I}$ Unterräume von V. Man sagt, die Summe (interne Summe) $\sum_{j\in I} U_j$ ist orthogonal, in Zeichen $o \perp U_j = \sum_{j\in I} U_j$ wenn $U_j \perp U_k$ für alle $j \neq k$ ist.

Lemma 6.83.
Jede orthogonale Summe ist direkt.

Beweis:
Sei $0 = \sum_{j\in I} u_j$ mit $u_j \in U_j$, $u_j = 0$ f.f.a. $j \in I$. Für jedes $k \in I$ ist $0 = \left\langle u_k, \underbrace{\sum_{j\in I} u_j}_{=0} \right\rangle =$

$\langle u_k, u_k\rangle = \|u_k\|^2 \Rightarrow u_k = 0$.

□

Satz 6.84. *(Gram-Schmidt Verfahren)*
Sei V ein Prähilbertraum, sei $(v_j)_{j\in I}$ eine linear unabhängige Folge von Vektoren in V, mit $I = \{1, \dots, n\}$ oder $I = \mathbb{N}$.
Dann gibt es eine ON-Familie $(u_j)_{j\in I}$ mit $span(u_1, \dots, u_k) = span(v_1, \dots, v_k)$ für alle $k \in I$.

Beweis:
Sei $U_k := span(v_1, \dots, v_k)$, $k \in I$.
Konstruiere $(u_j)_{j\in I}$ induktiv, beginnend mit $u_1 := \frac{v_1}{\|v_1\|}$. Sei $k > 1$, sei schon eine ON-Familie $(u_1, \dots, u_{k-1})$ mit $span(u_1, \dots, u_{k-1}) = U_{k-1}$ konstruiert. Für $\widetilde{u_k} := v_k - \sum_{j=1}^{k-1} \left\langle v_k, u_j \right\rangle \cdot u_j$ gilt $\left\langle u_k, u_j \right\rangle = 0$, $j = 1, \dots, k-1$, andererseits ist $span(\underbrace{u_1, \dots, u_{k-1}}_{U_{k-1}}, \widetilde{u_k}) =$

U_k.
Wir können also $u_k := \frac{\widetilde{u_k}}{\|\widetilde{u_k}\|}$ setzen.
(NB: $\widetilde{u_k} \neq 0$!).

□

Bemerkung 6.85.
Beweis war konstruktiv, gibt also insbesondere ein Rechenverfahren, das Orthonormalisierungsverfahren von Gram-Schmidt.

Beispiel:
Im $\mathbb{R}^3$ (mit Standardskalarprodukt) seien $v_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix} v_3 = \begin{pmatrix} -5 \\ 4 \\ 3 \end{pmatrix}$.
Aufgabe: Orthonormalisiere (v_1, v_2, v_3) nach Gram-Schmidt.
Beginn:

$u_1 := \frac{v_1}{\|v_1\|} = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$.

zweiter Schritt:

$$\widetilde{u_2} := v_2 - \langle v_2, u_1\rangle \cdot u_1 = \begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix} - \tfrac{3}{9}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

$$\Rightarrow \widetilde{u_2} = \tfrac{1}{3}\begin{pmatrix} -5 \\ 14 \\ -2 \end{pmatrix}, \|\widetilde{u_2}\| = 5, \text{ setze also } u_2 := \tfrac{1}{5}\cdot\widetilde{u_2} = \tfrac{1}{15}\begin{pmatrix} -5 \\ 14 \\ -2 \end{pmatrix}.$$

dritter Schritt:

$$\widetilde{u_3} = v_3 - \langle v_3, u_1\rangle \cdot u_1 - \langle v_3, u_2\rangle \cdot u_2 = \begin{pmatrix} -5 \\ 4 \\ 3 \end{pmatrix} - 0 \cdot u_1 - \tfrac{75}{225} \cdot \begin{pmatrix} -5 \\ 14 \\ -2 \end{pmatrix} = \tfrac{1}{3}\begin{pmatrix} -10 \\ -2 \\ 11 \end{pmatrix}.$$

$\|\widetilde{u_3}\|^2 = \frac{1}{9}\cdot 225 \Rightarrow \|\widetilde{u_3}\| = 5$. Setze $u_3 := \frac{1}{5}\widetilde{u_3} = \frac{1}{15}\begin{pmatrix} -10 \\ -2 \\ 11 \end{pmatrix}$.

Die Familie (u, u_2, u_3) hat die gewünschten Eigenschaften.

Korollar 6.86.
Jeder endlich-dimensionale Prähilbertraum hat eine ON-Basis.

Beweis:
Wende Gram-Schmidt Verfahren auf eine beliebige Basis an.

□

Betrachte den Induktionsschritt im Gram-Schmidt Verfahren. Von U_{k-1} und $v_k (\notin U_{k-1})$ wurde hier der Vektor $\widetilde{u_k} := v_k - u$ mit $u = \sum_{j=1}^{k-1} \left\langle v_k, u_j \right\rangle \cdot u_j$ gebildet.
Untersuche diese Projektion genauer:

Definition und Satz 6.87.
Sei V ein Prähilbertraum, U ein Unterraum von V und $v \in V$. Für $u \in U$ sind äquivalent:

(i) $v - u \in U^\perp$;

(ii) $\|v - u\| = d(v, U)$, *also* $\|v - u\| \leq \|v - u'\|$ *für alle* $u' \in U$.

Genau dann gibt es ein $u \in U$ mit (i) und (ii), wenn $v \in U + U^\perp$ ist. Dann ist u eindeutig bestimmt und heißt die ***orthogonale Projektion*** *von v auf U, in Zeichen $u = \pi_U(v)$.*

Dabei haben wir $d(v, U) := \inf\{\|v - x\| : x \in U\}$ gesetzt. Wegen (ii) nennt man $u = \pi_U(v)$ auch die **beste Approximation** von v auf U.

Beweis:

„⇒“ : Sei $v-u \in U^\perp$ Für $u' \in U$ ist $v-u \perp \underbrace{u-u'}_{\in U}$, also $\|v-u'\|^2 = \|(v-u)+(u-u')\|^2 =$

$\|v-u\|^2 + \|u-u'\|^2 + 2Re\underbrace{\langle v-u, u-u'\rangle}_{=0} \geq \|v-u\|^2$.

„⇐“ : Sei $u \in U$, es gebe $u' \in U$ mit $c := \langle v-u, u'\rangle \neq 0$. Wir können $\|u'\| = 1$ annehmen und außerdem $c \in \mathbb{R}$ (z.B. multipliziere u' mit $\frac{c}{|c|\cdot\|u'\|}$), dann ist
$\|v-(\underbrace{u+cu'}_{\in U})\|^2 = \|v-u\|^2 + \underbrace{|c|^2}_{=1}\|u'\|^2 - \underbrace{2Re\langle v-u, cu'\rangle}_{=2\langle v-u, cu'\rangle} = \|v-u\|^2 + c^2\cdot\underbrace{\|u'\|}_{=1} - 2|c|^2 =$
$\|v-u\|^2 - c^2$.
Widerspruch zur Minimalität (ii).

(i) sagt: $u \in U$ mit $v - u \in U^\perp$, d.h. $v \in U + U^\perp$.
Eindeutigkeit von u ist daraus klar (die Summe ist direkt).

□

Korollar 6.88.
Sei U ein Unterraum von V. Die orthogonale Projektion $\pi_U : U\oplus U^\perp \to U, \pi_U(u+u') = u$ $(u \in U, u' \in U^\perp)$, ist linear, und erfüllt $\|\pi_U(v)\| \leq \|v\|$ für alle $v \in U \oplus U^\perp$. Dabei gilt Gleichheit genau dann, wenn $v \in U$ (also $\pi_U(v) = v$) ist.

Beweis:
Es ist $v = \underbrace{\pi_U(v)}_{\in U} + \underbrace{(v - \pi_U(v))}_{\in U^\perp}$
also $\|v\|^2 = \|\pi_U(v)\|^2 + \|v - \pi_U(v)\|^2$.

□

Korollar 6.89.
Sei $U \subset V$ ein Unterraum mit $\dim(U) < \infty$, dann gilt:

(a) $V = U+U^\perp = U\oplus U^\perp$. Jedes $v \in V$ hat also eine orthogonale Projektion $\pi_U(v)$ auf U.

(b) *Ist* $(u_1, \dots, u_n)$ *eine ON-Basis von* U, *so ist* $\pi_U(v) = \sum_{j=1}^{n} \langle v, u_j \rangle \cdot u_j, \ \forall\, v \in V$.

(c) $U = U^{\perp\perp}$.

Beweis:

(i) $\left\langle v - \sum_{j=1}^{n} \langle v, u_j \rangle \cdot u_j ; u_k \right\rangle = \langle v, u_k \rangle - \sum_{j=1}^{n} \langle v, u_j \rangle \cdot \underbrace{\langle u_j, u_k \rangle}_{\delta_{jk}} = 0$. Also (a) + (b).

(ii)

(iii) folgt formal aus (a), auch ohne $\dim(U) < \infty$:
$U \subset U^{\perp\perp}$ gilt immer.
Umgekehrt sei $v \in U^{\perp\perp}$, schreibe $v = u + w$ mit $u \in U$ und $w \in U^{\perp}$ gemäß (a).
$\Rightarrow w = v - u \in U^{\perp\perp} \cap U^{\perp} = \{0\}$.
$\Rightarrow w = 0$, also $v = u \in U$.

□

Korollar 6.90. *(Bessel'sche Ungleichung)*
Ist $(v_j)_{j \in J}$ *eine ON-Familie in* V, *so gilt für alle* $v \in V$:
$\sum_{j \in J} |\langle v, v_j \rangle|^2 \leq \|v\|^2 < \infty$.

Beweis:
Für jede endliche Teilmenge I von J sei
$U_I := \operatorname{span}(v_i : i \in I)$, dann ist nach 6.89 $\|\pi_{U_{-1}}(v)\|^2 = \|\sum_{j \in J} \langle v, v_j \rangle \cdot v_j\|^2 = \sum_{j \in J} |\langle v, v_j \rangle|^2 \leq$
$\|v\|^2$ (nach 6.88).

□

Bemerkung 6.91.
Ist $U \subset V$ ein Unterraum mit $\dim(V) = \infty$, so ist i.A. $U + U^{\perp} \neq V$. möglich: siehe Beispiel 6.81.
Dort hatten wir Beispiele für $U \neq U^{\perp\perp}$ gesehen (z.B. in $V = l^2_{\mathbb{K}}$; für solches U ist auch $U + U^{\perp} \neq V$, siehe Beweis von 6.89.
Nicht jedes $v \in V$ hat also eine beste Approximation in U.

Vorbereitungen 6.92.
$V, \langle \cdot, \cdot \rangle, \| \cdot \| \rightsquigarrow$ Metrik auf V: $d(x, y) := \|x - y\|$.
Die Abbildungen
$V \times V \to V, (x, y) \mapsto x + y$,
$\mathbb{K} \times V \to V, (a, x) \mapsto ac$
$V \times V \to \mathbb{K}, (x, y) \mapsto \langle x, y \rangle$
sind stetig (Aufgabe 3, Übungsblatt 4).

Daher gilt für jeden Unterraum U von V:
der Abschluss $\overline{U}$ von U ist wieder ein Unterraum von V.
Für jedes $x \in V$ ist die Linearform $V \to \mathbb{K}$, $v \mapsto \langle v, x \rangle$, stetig. Für jedes $x \in V$ ist also der Unterraum $x^\perp$ von V abgeschlossen. Somit ist $M^\perp$ abgeschlossen in V für jede Teilmenge M von V.

Definition 6.93.
Ein Hilbertraum ist ein Prähilbertraum V, welcher unter der von $\| \cdot \|$ induzierten Metrik vollständig ist, d.h. in dem jede Cauchy-Folge konvergiert.

Satz 6.94.
Jeder endlich-dimensionale Prähilbertraum ist vollständig, also ein Hilbertraum.

Beweis:
Sei $v_1, \ldots, v_n$ eine ON-Basis von V. Die Abbildung $\Phi : \mathbb{K}^n \to V$, $\Phi(x_1, \ldots, x_n) := \sum_{j=1}^{n} x_j v_j$, ist eine **Isometrie**, d.h. $\forall\, x, y \in \mathbb{K}^n$ ist $\langle x, y \rangle = \sum_{j=1}^{n} x_j \overline{y_j} = \langle \Phi(x), \Phi(y) \rangle$.
Da $\mathbb{K}^n$ vollständig ist, folgt: V ist vollständig.

□

Bemerkung 6.95.
Ist V ein Prähilbertraum und U ein vollständiger Unterraum von V, so ist U abgeschlossen in V.
Ist umgekehrt V vollständig und U ein abgeschlossener Unterraum von V, so ist auch U vollständig.
(Beides gilt in jedem metrischen Raum)

Insbesondere ist V eine Hilbertraum und U ein Unterraum von V, so gilt: U vollständig $\Leftrightarrow$ U abgeschlossen in V.

Satz 6.96.
Der Folgenraum $V = l^2_{\mathbb{K}}$ ist vollständig, also ein Hilbertraum.

BEWEIS:
Sei $(a^{(m)})_{m\in\mathbb{N}}$ eine Cauchy-Folge in V, mit $a^{(m)} = (a_1^{(m)}, a_2^{(m)}, \ldots) = (a_j^{(m)})_{j\in\mathbb{N}}$.

$a^{(1)} = (a_1^{(1)}, a_2^{(1)}, a_3^{(1)}, \ldots)$
$a^{(2)} = (a_1^{(2)}, a_2^{(2)}, a_3^{(2)}, \ldots)$
$a^{(3)} = (a_1^{(3)}, a_2^{(3)}, a_3^{(3)}, \ldots)$
$\downarrow$
$b := (b_1, b_2, b_3, \ldots)$

Fixiere ein $j \in \mathbb{N}$. Dann ist $(a_j^{(m)})_{m\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{K}$, denn $|a_j^{(m)} - a_j^{(n)}|^2 \leq \sum_k |a_k^{(m)} - a_k^{(n)}|^2 = \|a^{(m)} - a^{(n)}\| := \lim_{n\to\infty} a_j^{(n)} \in \mathbb{K}$.
Sei $b := (b_1, b_2, \ldots) \in \mathbb{K}^{\mathbb{N}}$.
Sei $\varepsilon > 0$, und sei $N = N(\varepsilon) \in \mathbb{N}$ mit $\|a^{(m)} - a^{(n)}\| < \varepsilon \ \forall\ m, n \geq N$.

Für $m, n \geq N$ und alle $k \in \mathbb{N}$ ist also $\sum_{j=1}^{k} |a_j^{(m)} - a_j^{(n)}|^2 \leq \sum_{j=1}^{\infty} |a_j^{(m)} - a_j^{(n)}|^2 = \|a^{(m)} - a^{(n)}\|^2 < \varepsilon^2$.
Halte k und n fest und lasse $m \to \infty$ gehen.
Das ergibt $\sum_{j=1}^{k} |b_j - a_j^{(n)}|^2 \leq \varepsilon^2$.
Also $(k \to \infty)$ $\|b - a^{(n)}\|^2 \leq \varepsilon^2$.
$\Rightarrow b = (b - a^{(n)}) + a^{(n)} \in l^2_{\mathbb{K}}$, und
$(a^{(n)}) \to b, n \to \infty$ in der l^2-Norm.

□

SATZ 6.97.
Sei V ein Hilbertraum, sei $(v_n)_{n\in\mathbb{N}}$ eine ON-Familie in V (abzählbar).
Dann gilt für jede Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$:
die Reihe $\sum_{n=1}^{\infty} a_n v_n$ konvergiert in $V \Leftrightarrow \sum_{n=1}^{\infty} |a_n|^2 < \infty$ (d.h. $(a_n)_{n\in\mathbb{N}} \in l^2_{\mathbb{K}}$).

BEWEIS:
Sei $x_n := \sum_{j=1}^{n} a_j v_j$ $(n \in \mathbb{N})$.

Falls $\lim_{n\to\infty} x_n = v$ existiert, so folgt
$\|v\|^2 = \lim_{n\to\infty} \|x_n\|^2 = \lim_{n\to\infty} \|\sum_{j=1}^{n} a_j v_j\|^2 = \lim_{n\to\infty} \sum_{j=1}^{n} |a_j|^2$.

Umgekehrt sei $\sum |a_n|^2 < \infty$, sei $\varepsilon > 0$ und $N \in \mathbb{N}$ mit $\sum_{j>N} |a_j|^2 < \varepsilon^2$, dann ist für

$m, n > N \; \|x_m - x_n\|^2 \leq \| \sum_{j>N} a_j v_j \|^2 = \sum_{j>N} |a_j|^2 < \varepsilon^2$.
$\Rightarrow (x_n)$ ist Cauchy-Folge in V, hat also einen Limes.

□

Satz 6.98.
Sei V ein Hilbertraum, sei U ein abgeschlossener Unterraum von V. Dann ist $V = U \oplus U^\perp$ und $U^{\perp\perp} = U$. Insbesondere existiert für jedes $v \in V$ die orthogonale Projektion $\pi_U(v)$.

Beweis:
Zu zeigen ist nur $V = U + U^\perp$ ($\Rightarrow U^{\perp\perp} = U$ wie in Beweis von 6.89).

Nach 6.87 ist für jedes $v \in V$ zu zeigen: $\exists\, u \in U$ mit $\|v - u\| \leq \|v - u'\| \; \forall\, u' \in U$ ist. Wähle eine Folge $(u_n)_{n\in\mathbb{N}}$ in U mit $\|v - u_n\| \underbrace{\to}_{n\to\infty} d(v, U) = \inf\{\|v - u'\| : u' \in U\}$.

Behaupte, (u_n) ist Cauchy-Folge. ($\Rightarrow$ fertig: $\exists\, u = \lim_{n\to\infty} u_1 \in U$ und $\|v-u\| = d(v, U)$).

Setze $r := d(v, U)$, sei $\varepsilon > 0$ und $n_0 \in \mathbb{N}$ mit $\|v - u_n\| < r + \varepsilon \; \forall\, n \geq n_0$.

Sei $m, n \geq n_0 \Rightarrow \frac{1}{2}(u_m + u_n) \in U \Rightarrow \|v - \frac{1}{2}(u_m + u_n)\| \geq r \Rightarrow 4r^2 \leq \|(v - u_m) + (v - u_n)\|^2 = \|v - u_m\|^2 + \|v - u_n\|^2 + 2Re\,\langle v - u_m, v - u_n\rangle$
$\Rightarrow \|u_m - u_n\|^2 = \|(v - u_n) - (v - u_m)\|^2 = \|v - u_n\|^2 + \|v - u_m\|^2 - 2Re\,\langle v - u_m, v - u_n\rangle \leq 2\underbrace{\|v - u_m\|^2}_{\leq r+\varepsilon} + 2\underbrace{\|v - u_n\|^2}_{\leq r+\varepsilon} - 4r^2 \leq 4(r + \varepsilon)^2 - 4r^2 = 8r\varepsilon + 4\varepsilon^2$.

□

Korollar 6.99.
Sei V ein Hilbertraum, $U \subset V$ ein Unterraum von V.

(a) $U^\perp = \overline{U}^\perp$,

(b) $U^{\perp\perp} = \overline{U}$,

(c) *U ist dicht in $V \Leftrightarrow U^\perp = \{0\}$,*

(d) $U \oplus U^\perp = V \Leftrightarrow$ *U ist abgeschlossen.*

Beweis:

(a) $U \subset \overline{U} \Rightarrow U^\perp \supset \overline{U}^\perp$
$U^{\perp\perp}$ ist abgeschlossen, $U \subset U^{\perp\perp} \Rightarrow \overline{U} \subset U^{\perp\perp} \Rightarrow \overline{U}\perp \supset (U^\perp)^{\perp\perp} = U^\perp$, nach 6.98, da $U^\perp$ abgeschlossen ist.

(b) $U^{\perp\perp} = \overline{U}^{\perp\perp} = \overline{U}$, nach (a) und 6.98.

(c)

„$\Rightarrow$“ : $\overline{U} = V \Rightarrow U^\perp = \overline{U}^\perp = V^\perp = \{0\}$.

„$\Leftarrow$“ : $U^\perp = \{0\} \Rightarrow \overline{U} = U^{\perp\perp} = \{0\}^\perp = V$.

(d)

„$\Rightarrow$“ : $U = U^{\perp\perp}$, also abgeschlossen.

„$\Leftarrow$“ : 6.98.

□

Korollar 6.100.
Sei V ein Hilbertraum, $(v_n)_{n\in\mathbb{N}}$ eine (abzählbare) ON-Familie in V, sowie $U := \overline{span\ (u_n : n \in \mathbb{N})}$.
$\Rightarrow\ \forall\, v \in V : \pi_U(v) = \sum_{j=1}^{\infty} \left\langle v, v_j \right\rangle \cdot v_j$.

Beweis:
Übung.

□

Bemerkung 6.101.
Sei V ein Hilbertraum. Für $\dim(V) = \infty$ ist der übliche Begriff einer VR-Basis wenig hilfreich. Wichtig ist: eine ON-Familie $\mathcal{B}$ in V, die einen dichten Unter-VR von V erzeugt, heißt eine **Hilbertbasis** von V. Beispiele:

a) Ist $\dim(V) < \infty$, so: Hilbertbasis = ON-Basis.

b) In $V = l^2_{\mathbb{K}}$ ist $(e_n)_{n\in\mathbb{N}}$ eine Hilbertbasis, mit $e_n := (0, \ldots, 0, \underbrace{1}_{n-te\ Stelle}, 0, \ldots)$

Es enthalte V einen dichten Untervektorraum, der eine abzählbar unendliche VR-Basis hat. Nach dem Gram-Schmidt Verfahren hat V dann eine abzählbare Hilbertbasis $(v_n)_{n\in\mathbb{N}}$.
Jedes $v \in V$ schreibt sich als $\sum_{n=1}^{\infty} a_n v_n$ und dabei gilt für $v = \sum_{n=1}^{\infty} a_n v_n, w = \sum_{n=1}^{\infty} b_n v_n$:
$\langle v, w\rangle = \sum_{n=1}^{\infty} a_n \overline{b_n}$.
Insbesondere ist $\sum_{n=1}^{\infty} |a_n|^2 = \|v\|^2 < \infty$, also ist $(a_n) \in l^2_{\mathbb{K}}$.
Somit wird durch
$\Phi : l^2_{\mathbb{K}} \to V, (a_n)_{n\in\mathbb{N}} \mapsto \sum_{n\in\mathbb{N}} a_n v_n$
ein Vektorraum-Isomorphismus $\Phi : l^2_{\mathbb{K}} \to V$ definiert, der die Skalarprodukte erhält. Wir können also V als Hilbertraum mit $l^2_{\mathbb{K}}$ identifizieren.

Beispiel:
Der Prähilbertraum $V_0 = \{C([0, 2\pi], \mathbb{C}) \ni f : f(0) = f(2\pi)\}$ mit Skalarprodukt $\langle f, g\rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) \cdot \overline{g(t)}\, dt$. hat die ON-Familie der $f_n(x) = e^{n\cdot ix}$ $(n \in \mathbb{N})$.
Es gibt kanonische dichte Einbettung von V_0 in einen Hilbertraum $V (= L^2([0, 2\pi]))$.
In V bildet (f_n) eine Hilbertbasis, somit ist $V \cong l^2_{\mathbb{C}}$.

d. Orthogonale und unitäre Abbildungen

Ab jetzt seien alle (Prä-)Hilberträume von endlicher Dimension.

Definition 6.102.
Sei V ein Hilbertraum, $\dim(V) < \infty$. Ein $f \in End(V)$ heißt **orthogonal** (falls $\mathbb{K} = \mathbb{R}$) bzw. **unitär** (falls $\mathbb{K} = \mathbb{C}$), wenn gilt:
$\forall\, v, w \in V : \langle v, w\rangle = \langle f(v), f(w)\rangle$.
*(Andere Name: **Isometrie**)*

Beispiele 6.103.

1. Sei U ein Unterraum von V. Sei $\sigma_U \in End(V)$ definiert durch $\sigma_U(v) := 2\pi_U(v) - v$, $v \in V$ also $\sigma_U(v) = \pi_U(v) - (v - \pi_U(v))$ (σ_U ist Spiegelung an U).
   Dann ist σ_U eine orthogonale bzw. unitäre Abbildung, denn für $v = u + w$ mit $u \in U$, $w \in U^\perp$ ist $\sigma_U(v) = u - w$. Ist auch $v' = u' + w'$ mit $u' \in U, w' \in U^\perp$, so folgt $\langle \sigma_U(v), \sigma_U(v')\rangle = \langle u - w, u' - w'\rangle = \langle u, u'\rangle + \langle w, w'\rangle = \langle u + w, w' + w'\rangle = \langle v, v'\rangle$.

2. Für $c \in \mathbb{K}$ gilt: $f := c \cdot id_V$ ist genau dann orthogonal bzw. unitär, wenn $|c| = 1$ ist. Denn $\langle cv, cw\rangle = c\bar{c}\,\langle v, w\rangle \;\; \forall\, v, w \in V$.
   Dasselbe Argument zeigt: Jeder Eigenwert $c \in \mathbb{K}$ einer orthogonalen bzw. unitären Abbildung hat $|c| = 1$.

3. f orthogonal bzw. unitär $\Rightarrow \;\forall\, v, w \in V\; (v \perp w \Rightarrow f(v) \perp f(w))$. Umkehrung ist falsch, aber es gilt:

Lemma 6.104.
Für $f \in End(V)$ sind äquivalent:

(i) *f ist orthogonal bzw. unitär,*

(ii) *für jede ON-Basis $\mathcal{B}$ von V ist auch $f(\mathcal{B})$ eine ON-Basis von V,*

(iii) *$\forall\, v \in V$ ist $\|f(v)\| = \|v\|$.*

BEWEIS:
(i) $\Rightarrow$ (ii): klar.

(ii) $\Rightarrow$ (iii): ist $v \neq 0$, so ergänze $\frac{v}{\|v\|}$ zu einer ON-Basis von V.

(iii) $\Rightarrow$ (i) gilt, da man $\langle \cdot, \cdot \rangle$ durch Normen ausdrücken kann (Polarisierung).

□

BEISPIEL 6.105.
($\mathbb{K} = \mathbb{R}$) Sei $f \in End_{\mathbb{R}}(\mathbb{R}^2)$ die Drehung um Θ um $(0,0)$. Dann ist f orthogonal: Hinschauen, oder Rechnung:
$\langle f(x), f(y) \rangle = \langle Ax, Ay \rangle = x^t A^t A y = \langle x, y \rangle$
wegen $A^t A = \begin{pmatrix} \cos(\Theta) & \sin(\Theta) \\ -\sin(\Theta) & \cos(\Theta) \end{pmatrix} \begin{pmatrix} \cos(\Theta) & -\sin(\Theta) \\ \sin(\Theta) & \cos(\Theta) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

LEMMA 6.106.
Seien $f, g \in End(V)$ orthogonal bzw. unitär.

(a) f ist bijektiv, und f^{-1} ist orthogonal bzw. unitär.

(b) $f \circ g$ ist orthogonal bzw. unitär..

BEWEIS:
f ist injektiv wegen $\|f(v)\| = \|v\|$ ($\Rightarrow \ker(f) = \{0\}$, also bijektiv wegen $\dim(V) < \infty$.
□

DEFINITION 6.107.
(V Hilbertraum, $\dim(V) < \infty$)
Die Gruppe $\begin{cases} O(V) := \{f \in End_{\mathbb{R}}(V) & : f \text{ ist orthogonal}\} \\ U(V) := \{f \in End_{\mathbb{C}}(V) & : f \text{ ist unitär}\} \end{cases}$ (für $\left\{ \begin{array}{c} \mathbb{K} = \mathbb{R} \\ \mathbb{K} = \mathbb{C} \end{array} \right\}$) heißt die $\left\{ \begin{array}{c} \text{orthogonale} \\ \text{unitäre} \end{array} \right\}$ Gruppe von V.
(Dies ist eine Untergruppe von $GL_{\mathbb{K}}(V)$)

SATZ 6.108.
Sei $\dim(V) = n$, sei $\mathcal{B}$ eine ON-Basis von V. Sei $f \in End_{\mathbb{K}}(V)$, sei $A := M_{\mathcal{B}}^{\mathcal{B}}(f) \in M_n(\mathbb{K})$.

(a) ($\mathbb{K} = \mathbb{R}$) f orthogonal $\Leftrightarrow A \cdot A^t = I$.

(b) *(*$\mathbb{K} = \mathbb{C}$*) f unitär* $\Leftrightarrow A \cdot \overline{A}^t = I$.

Beweis:
Für $\mathcal{B} = (v_1, \ldots, v_n)$ und $v = \sum_{j=1}^{n} a_j v_j, w = \sum_{j=1}^{n} b_j v_j$ ist $\langle v, w \rangle = \sum_{j=1}^{n} a_j \overline{b_j} = a^t \cdot \overline{b}$ ($a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$), andererseits $\langle f(v), f(w) \rangle = (Aa)^t \cdot \overline{(Ab)} = a^t A^t \overline{A} \overline{b}$. Also $\langle f(v), f(w) \rangle = \langle v, w \rangle \ \forall\, v, w \Leftrightarrow A^t \cdot \overline{A} = I$.

□

Definition 6.109.

(a) ($\mathbb{K} = \mathbb{R}$) $O(n) := \{A \in GL_3(\mathbb{R}) : AA^t = I\}$ (**orthogonale Gruppe**)

(b) ($\mathbb{K} = \mathbb{C}$) $U(n) := \{A \in GL_n(\mathbb{C}) : A \cdot \overline{A}^t = I\}$ (**unitäre Gruppe**)

(c) $SO(n) := O(n) \cap SL_n(\mathbb{R})$ **spezielle orthogonale Gruppe**
$SU(n) := U(n) \cap SL_n(\mathbb{C})$ **spezielle unitäre Gruppe**.

Bemerkung 6.110.

1. Für $A \in M_n(\mathbb{K})$ sind äquivalent:

   (i) A ist orthogonal bzw. unitär,

   (ii) die Zeilen von A bilden eine ON-Basis,

   (iii) die Spalten von A bilden eine ON-Basis.

2. Sei $\mathcal{B}$ eine ON-Basis von V (ein Hilbertraum), sei C eine weitere Basis von V. Dann gilt:
   C ist ON-Basis $\Leftrightarrow T_{\mathcal{B}}^{C}$ ist orthogonal bzw. unitär.
   ($\mathcal{B} = (v_1, \ldots, v_n), C = (w_1, \ldots, w_n), w_k = \sum_j s_{jk} v_j \Rightarrow \langle w_k, w_l \rangle = \sum_j s_{jk} \cdot \overline{s_{jl}} = s^t \cdot \overline{s}$)

3. $\mathbb{K} = \mathbb{R}$: A ist orthogonal $\Leftrightarrow$ A^t ist orthogonal.
   $\mathbb{K} = \mathbb{C}$: A ist unitär $\Leftrightarrow$ A^t ist unitär.

Satz 6.111.

(a) ($\mathbb{K} = \mathbb{R}$)*: Für* $A \in O(n)$ *ist* $\det(A) = \pm 1$.
Der Homomorphismus $\det : O(n) \to \{\pm 1\}$ *ist surjektiv.*

(b) ($\mathbb{K} = \mathbb{C}$)*: Für* $A \in U(n)$ *ist* $|\det(A)| = 1$.
Der Homomorphismus $\det : U(n) \to U(1) = \{a \in \mathbb{C}^* : |a| = 1\}$.
ist surjektiv.

Beweis:

(a) $AA^t = I \Rightarrow \det(A)^2 = 1 \Rightarrow \det(A) \in \{\pm 1\}$

(b) $A\overline{A}^t = I \Rightarrow \det(A) \cdot \overline{\det(A)} = 1 \Rightarrow |\det(A)| = 1$.

□

Beispiele 6.112.

1. $O(1) = \{\pm 1\}, SO(1) = \{1\} U(1) = \{a \in \mathbb{C}^* : |a| = 1\}, SU(1) = \{1\}$.

2. Betrachte die Gruppe $O(2)$. Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$, setze $\alpha := a + ci, \beta := b + di \in \mathbb{C}$, ($i = \sqrt{-1}$). Dann gilt $A \in O(2) \Rightarrow a^2 + c^2 = 1, b^2 + d^2 = 1, ab + cd = 0$
   $|\alpha| = 1$ (1.), $|\beta| = 1$ (2.), $Re(\alpha\overline{\beta}) = 0$ (3.).
   Wegen (1.) und (2.) bedeutet (3.): $\alpha\overline{\beta} = \pm i$, also $\overline{\beta} = \pm i\overline{\alpha}$, also $\beta = \pm i\alpha$. Es gibt also zwei Fälle:

   (a) $A = \begin{pmatrix} \cos(\Theta) & sin(\Theta) \\ \sin(\Theta) & \cos(\Theta) \end{pmatrix}$,

(b) $A = \begin{pmatrix} \cos(\Theta) & \sin(\Theta) \\ \sin(\Theta) & -\cos(\Theta) \end{pmatrix}$

Identifizieren wir $\mathbb{R}^2 \approx \mathbb{C}, (x, y) \leftrightarrow x+iy$, so ist (a) die Abbildung $\mathbb{C} \to \mathbb{C}, z \mapsto \alpha z$ (mit $\alpha := \cos(\Theta) + i\sin(\Theta)$)
Im Fall (b) ist A die Abbildung $1 \mapsto \alpha, i \mapsto -i\alpha$, also $\mathbb{C} \to \mathbb{C}, z \mapsto \alpha\bar{z}$.
Die ist die orthogonale Spiegelung an der Geraden durch den 0 und $\pm\sqrt{\alpha}$.

Wir sehen: $O(2)$ zerfällt in zwei disjunkte Teilmengen. Dies sind die Drehungen ($SO(2)$) und die Spiegelungen (die Elemente von det $= -1$).
Wir können $SO(2)$ mit $U(1) = \{a \in \mathbb{C}^* : |a| = 1\} = S^1$ (1 -Sphäre) identifizieren.

3. Betrachte die Gruppe $SU(2)$.
Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{C})$ gilt:
$A \in SU(2) \Leftrightarrow |a|^2 + |c|^2 = 1$ und $b = -\bar{c}, d = \bar{a}$ (Aufgabe 3, Übungsblatt 6).
Somit ist
$SU(2) =$
können wir identifizieren mit 3-dimensionale Einheitssphäre S^3 im $\mathbb{R}^4$.

4. Betrachte $O(3)$. Ist $A \in O(3)$, so hat $P_A(t)$ eine reelle Nullstelle $\lambda = \pm 1$ (6.103 2). Sei $v \in \mathbb{R}^3$ ein Eigenvektor zu λ, sei $W := (\mathbb{R}v)^\perp$. Dann ist W A-invariant, und $\dim(W) = 2$.
Also wird $A|_W : W \to W$ bezüglich einer ON-Basis von W durch eine Matrix in $O(2)$ beschrieben.
Setze nun $\det(A) = 1$, also $A \in SO(3)$.

Ist $\lambda = +1$, so ist $A|_W \in SO(W)$ (wegen $\det(A) = +1$), also ist $A|_W$ eine Drehung um einen Winkel Θ.
Ist $A \neq I$, so heißt $\mathbb{R}v = Eig(A; 1)$ die **Drehachse** von A und $W = (\mathbb{R}v)^\perp$ die **Drehebene** von A.
Ist (w_1, w_2) eine ON-Basis von W, so hat A bezüglich der Basis (v, w_1, w_2) von $\mathbb{R}^3$ die Form $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\Theta) & -\sin(\Theta) \\ 0 & \sin(\Theta) & \cos(\Theta) \end{pmatrix}$. Man beachte $W = \mathrm{im}(A-I)$ (falls $A \neq I$).

Falls $\lambda = -1$, so ist $\det(A|_W) = -1$; also ist $A|_W$ eine orthogonale Spiegelung.
Auch in diesem Fall hat also A den Eigenwert +1 und ist eine Drehung um den Winkel $\pi = 180°$. Zusammenfassung:

SATZ 6.113.
Ist $A \in SO(3)$, $A \neq I$, so ist A eine Drehung in der Ebene $W = im(A - I) = Eig(A;1)^{\perp}$.

e. Vektorprodukt im $\mathbb{R}^3$

Sei V ein orientierter 3-dimensionaler reeller Hilbertraum. (Erinnerung: siehe Definition 4.69 in LA II).
Standardbeispiel: $V = \mathbb{R}^3$ mit Standardskalarprodukt $\langle x, y\rangle = x_1y_1 + x_2y_2 + x_3y_3$, und der durch Basis (e_1, e_2, e_3) gegebene (kanonische) Orientierung.

Satz 6.114.
Sei (u_1, u_2, u_3) eine positiv orientierte ON-Basis von V. Dann gibt es genau eine bilineare Abbildung $V \times V \to V$, $(v, w) \mapsto v \times w$ mit

(1) $v \times v = 0 \ \ \forall\, v \in V$

(2) $u_1 \times u_2 = u_3,\ u_2 \times u_3 = u_1,\ u_3 \times u_1 = u_2.$

Darüber hinaus hat $\times$ dann die Eigenschaft (2) bezüglich jeder positiv orientierten ON-Basis.

Bemerkung:
Die letzte Aussage bedeutet: für alls $v_1, v_2 \in V$ mit $\|v_1\| = \|v_2\| = 1$ und $\langle v_1, v_2\rangle = 0$ ist $(v_1, v_2, v_1 \times v_2)$ eine positiv orientierte ON-Basis von V.

Beweis:
Es genügt, den Fall $V = \mathbb{R}^3$ (mit kanonischem Skalarprodukt und Orientierung) zu betrachten. Wir müssen definieren $e_i \times e_i = 0$ $(i = 1, 2, 3)$ und $e_1 \times e_2 = e_3, e_2 \times e_3 = e_1, e_3 \times e_1 = e_2$.
Aus (1) folgt $\forall\, v, w \in V$:
$0 = (v + w) \times (v + w) = \underbrace{(v \times v)}_{=0} + (v \times w) + (w \times v) + \underbrace{(w \times w)}_{=0}$, also $(w \times v) = -(v \times w)$.
Also muss auch $e_2 \times e_1 = -e_3, e_3 \times e_2 = -e_1, e_1 \times e_3 = -e_2$ sein.
Wegen Bilinearität folgt allgemein:

$$v \times w = (v_2w_3 - v_3w_2, v_3w_1 - v_1w_3, v_1w_2 - v_2w_1) = \left(\begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix}, \begin{vmatrix} v_3 & w_3 \\ v_1 & w_1 \end{vmatrix}, \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} \right)$$

oder symbolisch: $v \times w = \begin{vmatrix} v_1 & e_3 & e_1 \\ v_2 & e_3 & e_2 \\ v_3 & w_3 & e_3 \end{vmatrix}$
(entwickle formal nach der dritten Spalte).

Umgekehrt sei $\times$ hierdurch definiert. Dann ist $\times$ bilinear, und (1) und (2) gelten. Für den Zusatz ist zu zeigen: ist $A \in SO(3)$ eine Matrix mit Spalten v, w, z, so ist $z = v \times w$. Das folgt so:
$A^t = A^{-1} = \frac{1}{\det(A)} \cdot A^\sharp = A^\sharp$ ist z die dritte Zeile von $A^\sharp$; man sieht: $z = v \times w$.

□

Bemerkung 6.115.

Wir haben gerade gesehen: sind $v, w \in \mathbb{R}^3$ und ist $A = \begin{pmatrix} v_1 & w_1 & * \\ v_2 & w_2 & * \\ v_3 & w_3 & * \end{pmatrix}$ ($*$ beliebig),

so ist $v \times w$ = dritte Zeile in $A^\sharp$.

Definition 6.116.

Sei V ein 3-dimensionaler orientierter $\mathbb{R}$-Hilbertraum. Die durch 6.114 beschriebene Abbildung $V \times V \to V$, $(v, w) \mapsto v \times w$ das **Vektorprodukt** auf V.

Satz 6.117.

V wie eben, $\times$ das Vektorprodukt, $u, v, w \in V$.

(a) $u \times v = -v \times u$,

(b) *Sind $x, y, z \in \mathbb{R}^3$ die Koordinatenvektoren von u, v, w bezüglich einer positiv orientierten ON-Basis, so ist* $\langle u \times v, w \rangle = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$,

(c) *$u \times v$ steht senkrecht auf u und v,*

(d) $\|u \times v\|^2 = \|u\|^2 \|v\|^2 - \langle u, v \rangle^2$,

(e) $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(u, v)|$,

(f) $u \times v = 0 \Leftrightarrow$ *u und v sind linear abhängig,*

(g) $\det(u, v, u \times v) = \|u \times v\|^2$,

(h) $u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0$. *(„Jacobi-Identität“).*

Hierbei ist $\sin(u,v)$ *der Sinus des unorientierten Winkels zwischen* u *und* v*.*

Bemerkung 6.118.

1. (b) sagt: für $u,v,w \in V$ ist $\langle u \times v, w\rangle$ das orientierte Volumen des von u,v,w aufgespannten Spats. *(man nennt* $\langle u \times v, w\rangle$ *auch das* ***Spatprodukt*** *von* u,v,w*.)* Das Spatprodukt ist invariant unter zyklischer Vertauschung von u,v,w.

2. (e) sagt: $\|u \times v\|$ ist genau die Fläche des von u und v aufgespannten Parallelogramms.

3. Sind u und v linear unabhängig, so ist $(u,v,u \times v)$ eine positiv orientierte Basis von V. Aus (f) und (g).

4. $\times$ ist **nicht** assoziativ: z.B. $(e_1 \times e_1) \times e_2 = 0, e_1 \times (e_1 \times e_2) = -e_2$.

Beweis: (von 6.117)

(a) klar.

(b) direkt aus „symbolischer “ Form des Vektorprodukts (siehe 6.114).

(c) sofort aus (b).

(d) später.

(e) besagt $\|u \times v\|^2 = \|u\|^2 \cdot \|v\|^2 \cdot \underbrace{(1 - \cos^2(u,v))}_{=\sin^2(u,v)} \Rightarrow$ Aussage.

(f) aus (e).

(g) aus (b).

(h) Übung.

□

Satz 6.119.
Seien $G, H \subset \mathbb{R}^3$ windschiefe Geraden im $\mathbb{R}^3$, also $G = p + \mathbb{R}u$, $H = q + \mathbb{R}v$ mit $p, q \in \mathbb{R}^3$, $u, v \in \mathbb{R}^3$ linear unabhängig. Dann ist

$$d(G,H) = \frac{|\langle u\times v, p-q\rangle|}{\|u\times v\|}$$

Beweis:
$d(G,H) = \inf\{\|(p+au)-(q+bv)\| : a,b \in \mathbb{R}\} = d(p-q,U)$ mit $U : \mathbb{R}u + \mathbb{R}v$. Nach 6.117 ist $U^\perp = \mathbb{R}(u\times v)$, also ist wegen $d(G,H) = \|(p-q) - \pi_U(p-q)\|$.
Für beliebiges $w \in \mathbb{R}^3$ ist $\pi_U(w) = w + \lambda(u\times v)$, und λ ergibt sich daraus, dass $w + \lambda(u\times v), u, v$ linear abhängig wird. Das bedeutet nach 6.117 (b) $\langle u\times v, w\rangle + \lambda\langle u\times v, u\times v\rangle = 0$, also $\lambda = -\frac{\langle u\times v,w\rangle}{\|u\times v\|^2}$, also $d(w,U) = \|\lambda\cdot(u\times v)\| = |\lambda|\cdot\|u\times v\| = \frac{|\langle u\times v,w\rangle|}{\|u\times v\|}$.
Setze $w = p - q$.

□

6.120.
Die Menge $\mathbb{H} := \left\{\begin{pmatrix} a & b \\ -\overline{b} & \overline{a}\end{pmatrix} : a,b \in \mathbb{C}\right\}$ in $M_2(\mathbb{C})$ ist $\mathbb{R}$-Untervektorraum, ist ein Teilring von $M_2(\mathbb{C})$ und ist ein Schiefkörper.
(Blatt 6, Aufgabe 2 in LA I): für $0 \neq x = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a}\end{pmatrix} \in \mathbb{H}$ ist $\det(x) = |a|^2 + |b|^2 > 0$, und daher $x^{-1} = \frac{1}{\det(x)}\begin{pmatrix} \overline{a} & -b \\ \overline{b} & a\end{pmatrix} = \frac{1}{|a|^2+|b|^2}\begin{pmatrix} \overline{a} & -b \\ \overline{b} & a\end{pmatrix} \in H$.

Man nennt $\mathbb{H}$ den **Schiefkörper der Hamiltonschen Quatermionen**. Die Elemente $E_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}, E_1 = \begin{pmatrix} i & 0 \\ 0 & -i\end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}, E_3 = \begin{pmatrix} 0 & i \\ i & 0\end{pmatrix}$ bilden eine Basis des $\mathbb{R}$-Vektorraumes $\mathbb{H}$, es ist $\dim(\mathbb{H}) = 4$.
Es gelten folgende Rechenregeln:

(1) $E_0 \cdot E_j = E_j \cdot E_0 = E_j$ $(j = 1,2,3)$,

(2) $E_j^2 = -I = -E_0$ $(j = 1,2,3)$,

(3) $E_j \cdot E_k = E_l$, falls (j,k,l) eine zyklische Permutation von $(1,2,3)$ ist,

(4) $E_j \cdot E_k = -E_k \cdot E_j$ für $j,k \in \{1,2,3\}$, $j \neq k$.

Versehe nun $\mathbb{H}$ mit dem Skalarprodukt $\langle -,- \rangle$, welches (E_0, E_1, E_2, E_3) als ON-Basis hat (sogenanntes **kanonisches Skalarprodukt** auf $\mathbb{H}$). Für $x = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} \in \mathbb{H}$ ist $\|x\|^2 = \langle x, x \rangle = |a|^2 + |b|^2 = \det(x)$.
Wegen $\det(x,y) = \det(x) \cdot \det(y)$ gilt also $\|xy\|2 = \|x\|^2 \cdot \|y\|^2 \; \forall \; x, y \in \mathbb{H}$.

Theorem 6.121. *(Euler 1770)*
$(x_0^2 + x_1^2 + x_2^2 + x_3^2) \cdot (y_0^2 + y_1^2 + y_2^2 + y_3^2) = (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)^2 + x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)^2 + (x_0y_2 + x_2y_0 - x_1y_3 + x_3y_1)^2 + (x_0y_3 + x_3y_1 + x_1y_2 - x_2y_1)^2$

Bemerkung 6.122.
Der Beweis (siehe 6.120) hat diese Aussage zunächst für reelle x_j, y_j gezeigt. Aber Nachrechnen zeigt: die Identität gilt für beliebige x_j, y_j in jedem kommutativen Ring (z.B. $\mathbb{Z}$).

Definition 6.123.
Wir setzen $\mathbb{H}_0 := \mathbb{R}E_1 + \mathbb{R}E_2 + \mathbb{R}E_3$ und nennen $\mathbb{H}_0$ den Vektorraum der **reinen Quatermionen**.
Also $\mathbb{H}_0 = \left\{ \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} : a, b \in \mathbb{C}, Re(a) = 0 \right\}$.
Durch die ON-Basis (E_1, E_2, E_3) von $\mathbb{H}_0$ wird $\mathbb{H}_0$ mit $\langle \cdot, \cdot \rangle$ zu einem orientierten 3-dimensionalen $\mathbb{R}$-Hilbertraum.

Sei $\times$ das zugehörige Vektorprodukt auf $\mathbb{H}_0$, siehe Satz 6.114.
Unter Identifizierung von $\mathbb{H}_0$ mit $\mathbb{R}^3$ via $x_1E_1 + x_2E_2 + x_3E_3 \leftrightarrow (x_1, x_2, x_3)$, so identifiziert sich also x auf $\mathbb{H}_0$ mit $\times$ auf $\mathbb{R}^3$.

Satz 6.124.
Für $u, v \in \mathbb{H}_0$ ist

$$uv = -\langle u, v \rangle \cdot E_0 + (u \times v).$$

Also insbesondere: $u \times v$ ist die orthogonale Projektion uv auf den Unterraum $\mathbb{H}_0$ von $\mathbb{H}$.

Beweis:
Ist $u = \sum_{j=1}^{3} x_jE_j$, $v = \sum_{j=1}^{3} y_jE_j$ mit $x_j, y_j \in \mathbb{R}$, so rechnet man aus:

$$uv = -(x_1y_1 + x_2y_2 + x_3y_3)E_0 + (x_2y_3 - x_3y_2)E_1 + (x_3y_1 - x_1y_3)E_2 + (x_1y_2 - x_2y_1)E_3.$$

□

Als Korollar erhalten wir für alle $u, v \in \mathbb{H}_0$:

$$\|uv\|^2 = \|u\|^2 \cdot \|v\|^2 = \langle u, v\rangle^2 + \|u \times v\|^2.$$

Damit ist insbesondere 6.117 (d) bewiesen. Die obige Identität ist eine Präzisierung der Cauchy-Schwarz-Ungleichung $\langle u, v\rangle^2 \leq \|u\|^2 \cdot \|v\|^2$.

6.125.
Für $x = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \in \mathbb{H}$ setze $x^* = \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix} \in \mathbb{H}$. Es ist $(x^*)^* = x$, $(xy)^* = x^*y^*$, und $x^*x = xx^* = \langle x, x\rangle \cdot E_0 = \det(x) \cdot I$ (siehe Aufgabe 7.4).
Setze $T(x) := \frac{1}{2}(x + x^*)$, also $T(x) = \begin{pmatrix} Re(a) & 0 \\ 0 & Re(a) \end{pmatrix} = Re(a) \cdot E_0$. Für $x = \sum_{j=0}^{3} x_jE_j$ ist $T(x) = x_0E_0$.
$T : \mathbb{H} \to \mathbb{R} \cdot E_0$ ist $\mathbb{R}$-linear, $\ker(T) = \mathbb{H}_0$.

Bemerkung:
$(\mathbb{H}^* = \mathbb{H} \setminus \{0\})$
Es ist $\mathbb{H}^* \cap SL_2(\mathbb{C}) = \{x \in \mathbb{H} : \det(x) = 1\} = \{x \in \mathbb{H} : \langle x, x\rangle = 1\} = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} : |a|^2 + |b|^2 = 1 \right\} = SU(2)$.

Lemma 6.126.
Für $u, x \in \mathbb{H}, u \neq 0$, sei $f_u(x) := uxu^{-1} \in \mathbb{H}$.

(a) *Die Abbildung $f_u : \mathbb{H} \to \mathbb{H}$ ist $\mathbb{R}$-linear und orthogonal, und $f_u(\mathbb{H}_0) = \mathbb{H}_0$. Wir setzen $\varphi_u := f_u|_{\mathbb{H}_0} \in End(\mathbb{H}_0)$. Es ist $\det(\varphi_u) = 1$, also $\varphi_{uj} \in SO(\mathbb{H}_0)$.*

(b) *$\varphi : \mathbb{H}^* \to SO(\mathbb{H}_0), u \mapsto \varphi_u$, ist ein Gruppenhomomorphismus mit $\ker(\varphi) = \mathbb{R}^* \cdot E_0$.*

(c) *Für $u \notin \mathbb{R}^* \cdot E_0$ ist die Drehachse von φ_u gleich $\mathbb{R}(u - u^*) \subset \mathbb{H}_0$.*

(d) *Für $0 \neq u \in \mathbb{H}_0$ ist φ_u eine Drehung um $180°$ in $\mathbb{H}_0$ (mit Drehachse $\mathbb{R}u$).*

Beweis:
Die Abbildung f_u ist $\mathbb{R}$-linear. Wegen $\|f_u(x)\| = \|uxu^{-1}\| = \|u\| \cdot \|x\| \cdot \|u\|^{-1} = \|x\|$ ist f_u orthogonal. Es ist $f_u(E_0) = E_0$.
Also lässt f_u auch $(\mathbb{R}E_0)^\perp = \mathbb{H}_0$ invariant.

Es ist $f_u(f_v(x)) = u(vxv^{-1})u^{-1} = (uv)x(uv)^{-1} = f_{uv}(x)$, also $f_{uv} = f_u \circ f_v$, also auch $\varphi_{uv} = \varphi_u \circ \varphi_v$.
Also ist φ ein Gruppenhomomorphismus.

Sei $u \in \mathbb{H}, u \notin \mathbb{R}E_0$. Schreibe $v := u - T(u)$, also $u = \underbrace{T(u)}_{\in \mathbb{R}E_0} + \underbrace{v}_{\in \mathbb{H}_0}$.

Habe $u^2 = (T(u) + v)u = u(T(u) + v) \Rightarrow uv = vu$.
$\Rightarrow uvu^{-1} = v$. Also ist v ein Eigenvektor von φ_u zum Eigenwert 1. Wir zeigen nun $\varphi_u(w) \neq w$ für alle $0 \neq w \in \mathbb{H}_0$ mit $\langle v, w\rangle = 0$.
Daraus folgt dann $\det(\varphi_u) = 1$. *Denn betrachte φ_u eingeschränkt auf das orthogonale Komplement von v in $\mathbb{H}_0$: Dieses ist eine Drehung (und keine Spiegelung), da $+1$ kein Eigenwert ist.*
Es folgt auch $Eig(\varphi_u, 1) = \mathbb{R}v$.
Nach Aufgabe 4, Übungsblatt 7 ist
$vwv = 2\langle w^*, v\rangle \cdot - \langle v, v\rangle \cdot w^*$.
$\Rightarrow$ (wegen $w^* = -w$) $vewv = \langle v, v\rangle \cdot w$.
Andererseits (wegen $v^* = -v$) $v^2w = -(w^*)w = -\langle v, v\rangle \cdot w$.
$\Rightarrow$ Kürzen (links) durch v gibt $vw = -wv$. Damit auch $uw = (T(u) + v)w \neq w(T(u) + v) = wu \Rightarrow \varphi_u(w) = uwu^{-1} \neq w$.

Es folgt insbesondere $\varphi_u \neq id$; also $u \neq \ker(\varphi)$.

Ist sogar $u \in \mathbb{H}_0$, so ist in obiger Diskussion $v = u$, also $uw = -wu \ \forall\ w \in \mathbb{H}_0, \langle w, u\rangle = 0$.
Also $\varphi_u(w) = -w$ für diese w.

□

Theorem 6.127.
Die Einschränkung
$\varphi : SU(2) \to SO(\mathbb{H}_0)$ *von φ ist surjektiv und hat Kern* $\{\pm E_0\}$.

Beweis:
Aussage über Kern ist klar aus Lemma 6.127 (b).

Jede Drehung in $SO(3)$ ist Produkt aus zwei 180°-Drehungen:

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta & -\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Nach (d) des Lemmas liegen alle 180°-Drehungen in $\mathbb{H}_0$ im Bild von $\varphi : SU(2) \to SO(\mathbb{H}_0)$:
ist $\mathbb{R}u$ die Drehachse, so ist $\varphi_{\frac{u}{\|u\|}}$ die 180°-Drehung mit dieser Achse $\Rightarrow \varphi$ ist surjektiv.

□

Korollar 6.128.
φ induziert einen Gruppenisomorphismus

$$SU(2)/\{\pm I\} \to SO(3).$$

Übung: Schreibe ihn explizit auf.

Bemerkung 6.129.
Gruppenisomorphismus $\varphi : SU(2)/\{\pm I\} \to SO(3)$.
Explizit findet man

$$\varphi : \pm\begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} \mapsto \begin{pmatrix} a^2+b^2-c^2-d^2 & 2(bc-ad) & 2(ac-bd) \\ 2(ad+bc) & a^2+c^2-b^2-d^2 & 2(cd-ab) \\ 2(bd-ac) & 2(ab+cd) & a^2+d^2-b^2-c^2 \end{pmatrix},$$

$a^2+b^2+c^2+d^2 = 1$.

f. Adjungierte Abbildung

Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, seien $V = (V, \langle -,- \rangle_V), W = (W, \langle -,- \rangle_W)$ zwei endlichdimensionale $\mathbb{K}$-Hilberträume.

Definition und Satz 6.130.
Sei $f : V \to W$ eine $\mathbb{K}$-lineare Abbildung.
Dann gibt es genau eine lineare Abbildung $g : W \to V$ mit

$$\langle f(v), w \rangle_W = \langle v, g(w) \rangle_V, \ \forall\, v \in V, \ \forall\, w \in W. \tag{*}$$

Man nennt g die zu f adjungierte Abbildung und schreibt $g =: f^{ad}$.
Ist $(v_1, \dots, v_n)$ eine ON-Basis von V, so ist

$$f^{ad}(w) = \sum_{j=1}^{n} \left\langle w, f(v_j) \right\rangle_W \cdot v_j \ (w \in W). \tag{**}$$

Beweis:
Eindeutigkeit:
Seien $g, g' : W \to V$ Abbildungen mit (*). Dann ist $\forall\, v \in V, \ \forall\, w \in W$ $\langle v, g(w) - g'(w) \rangle_V = \langle f(v), w - w \rangle_W = 0$, also ist $g(w) = g'(w)$.

Existenz:
Für die Existenz sei f^{ad} wie in (**) definiert. Dan ist $f^{ad} : W \to V$ $\mathbb{K}$-linear, und (*) gilt:
für $v = \sum_{j=1}^{n} a_j v_j \in V$ und $w \in W$ ist $\left\langle v, f^{ad}(w) \right\rangle_V = \sum_{j=1}^{n} a_j \cdot \overline{\left\langle w, f(v_j)_W \right\rangle} = \sum_{j=1}^{n} a_j \cdot$
$\left\langle f(v_j), w \right\rangle_W = \Big\langle \underbrace{f\left(\sum_{j=1}^{n} a_j v_j \right)}_{=v}, w \Big\rangle_W$. (*)

□

Lemma 6.131.
Sei $f : V \to W$ linear, seien $\mathcal{B}$ bzw. $\mathcal{C}$ ON-Basen in V bzw. W. Dann gilt:

$$M_{\mathcal{B}}^{\mathcal{C}}(f^{ad}) = \overline{M_{\mathcal{C}}^{\mathcal{B}}(f)}^{t}.$$

Beweis:
Seien $\mathcal{B} = (v_1, \ldots, v_n)$, $C = (w_1, \ldots, w_m)$, sei $A := M_{\mathcal{C}}^{\mathcal{B}}(f) =: (a_{jk}), B := M_{\mathcal{B}}^{\mathcal{C}}(f^{ad}) =: (b_{jk})$.

$$a_{jk} = \left\langle f(v_k), w_j \right\rangle_W = \left\langle v_k, f^{ad}(w_j) \right\rangle_V = \overline{\left\langle f^{ad}(w_j), v_k \right\rangle_V} = \overline{b_{kj}}.$$

□

Bemerkung 6.132.
Ist $W = V$, also $f \in End(V)$, so ist $P_{f^{ad}}(t) = \overline{P_f}(t)$ (koeffizientenweises Konjugieren: $p(t) = \sum a_n t^n \Rightarrow \overline{p}(t) = \sum \overline{a_n} t^n$).

Insbesondere: $\det(f^{ad}) = \overline{\det(f)}$ und $\operatorname{tr}(f^{ad}) = \overline{\operatorname{tr}(f)}$.

Beispiel 6.133.
Ist $f \in End(V)$. Dann gilt:

$$f \text{ ist orthogonal bzw. unitär} \Leftrightarrow f \text{ ist invertierbar und } f^{ad} = f^{-1}.$$

Denn: sei f invertierbar. Dann gilt: $f^{ad} = f^{-1} \Leftrightarrow \forall\, v, w \in V \ \langle f(v), w \rangle = \left\langle v, f^{-1}(w) \right\rangle$
$\Leftrightarrow \forall\, v, w \in V \ \langle f(v), f(w) \rangle = \langle v, w \rangle \Leftrightarrow f$ ist orthogonal bzw. unitär.

Lemma 6.134.
Für lineare Abbildungen $f, f_1, f_2 : U \to V$ und $g : V \to W$ gilt:

(a) $(f_1 + f_2)^{ad} = f_1^{ad} + f_2^{ad}$
$(\lambda f)^{ad} = \overline{\lambda} f^{ad}, \lambda \in \mathbb{K}$;

(b) $(g \circ f)^{ad} = f^{ad} \circ g^{ad}$;

(c) $(id_V)^{ad} = id_V$;

(d) $(f^{ad})^{ad} = f$.

Satz 6.135.
Sei $f : V \to W$ linear. Dann ist

$$\ker(f) = im(f^{ad})^{\perp}, \ker(f^{ad}) = im(f)^{\perp}.$$

Beweis:
Für $v \in V$: $v \in \ker(f) \Leftrightarrow \ \forall\ w \in W\ \langle f(v), w\rangle = 0 \Leftrightarrow \ \forall\ w \in W\ \left\langle v, f^{ad}(w)\right\rangle = 0 \Leftrightarrow$ $v \in \mathrm{im}(f^{ad})^{\perp}$.

zweite Gleichung aus erster Gleichung mit f^{ad} statt f.

□

Korollar 6.136.
$f : V \to W$.

(a) f ist injektiv $\Leftrightarrow$ f^{ad} ist surjektiv;

(b) f ist surjektiv $\Leftrightarrow$ f^{ad} ist injektiv;

(c) f ist bijektiv $\Leftrightarrow$ f^{ad} ist bijektiv.

Beweis:

(a) f ist injektiv $\Leftrightarrow 0 = \ker(f) = \mathrm{im}(f^{ad})^{\perp} \Leftrightarrow V = im(f^{ad}) \Leftrightarrow f^{ad}$ ist surjektiv.

(b) analog zu (a)

(c) aus (a) und (b).

□

g. Selbstadjungierte und normale Abbildungen, Spektralsatz

Sei stets V ein Hilbertraum über $\mathbb{K}$, $\dim(V) < \infty$.

Definition 6.137.
Sei $f \in End(V)$.

(a) f heißt **selbstadjungiert**, falls $f^{ad} = f$ ist.

(b) f heißt **antiselbstadjungiert**, falls $f^{ad} = -f$ ist.

(c) f heißt **normal**, falls $f \circ f^{ad} = f^{ad} \circ f$.

Bemerkungen 6.138.

1. Ist $\mathcal{B}$ eine ON-Basis von V, ist $f \in End(V)$, $A := M_{\mathcal{B}}^{\mathcal{B}}(f)$, so gilt:

$$f \text{ ist selbstadjungiert} \Leftrightarrow \overline{A}^t = A;$$
$$f \text{ ist antiselbstadjungiert} \Leftrightarrow \overline{A}^t = A;$$
$$f \text{ ist normal} \Leftrightarrow A \cdot \overline{A}^t = \overline{A}^t \cdot A.$$

2. f ist selbstadjungiert $\Rightarrow \forall\, v \in V$ ist $\langle f(v), v\rangle \in \mathbb{R}$. Denn $\langle f(v), v\rangle = \langle v, f(v)\rangle = \overline{\langle f(v), v\rangle} \in \mathbb{R}$.

3. f, g selbstadjungiert, so auch $f + g$ und λf für $\lambda \in \mathbb{R}$.

4. Für jeden Unterraum U von V sind die orthogonale Projektion $\pi_U : V \to U$ und die orthogonale Spiegelung $\sigma_U : V \to V$ an U selbstadjungiert.

   Denn $V = U \oplus U^{\perp}$. Für jede zu dieser Zerlegung adaptierte ON-Basis von V hat π_U die Matrix $\left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right)$, σ_U die Matrix $\left(\begin{array}{c|c} I & 0 \\ \hline 0 & -I \end{array}\right)$.
   Für beide gilt $A = \overline{A}^t$.

5. Für beliebiges $f \in End(V)$ sind $f \circ f^{ad}$ und $f^{ad} \circ f$ selbstadjungiert, ebenso auch $f + f^{ad}$.

6. Alle selbstadjungierten, antiselbstadjungierten, orthogonalen bzw. unitären Abbildungen sind normal.
   Dies sind die wichtigsten Beispiele normaler Abbildungen.

Satz 6.139.
Sei $f \in End(V)$ normal. Für jedes $\lambda \in \mathbb{K}$ gilt:
$Eig(f, \lambda) = Eig(f^{ad}, \overline{\lambda})$.
Insbesondere ist $\ker(f) = \ker(f^{ad})(= im(f^{ad})^{\perp} = im(f)^{\perp})$.

Beweis:
Für $v \in V$ ist $\|f(v)\|^2 = \langle f(v), f(v)\rangle$
$= \left\langle v, f^{ad} \circ f(v)\right\rangle = \left\langle v, f \circ f^{ad}(v)\right\rangle$ (da f normal)
$= \left\langle f^{ad}(v), f^{ad}(v)\right\rangle = \|f^{ad}(v)\|^2$.

Daraus sehen wir $\ker(f) = \ker(f^{ad})$.
Ist $\lambda \in \mathbb{K}$ beliebig, so ist $Eig(f; \lambda) = \ker(f - \lambda \cdot id_V)$.
Mit f ist auch $f - \lambda \cdot \mathrm{id}_V$ normal, denn $(f - \lambda id)^{ad} = f^{ad} - \overline{\lambda} \cdot id$.
Aus der schon gezeigten Aussage folgt also $Eig(f; \lambda) = \ker(f - \lambda id) = \ker(f^{ad} - \overline{\lambda} id) = Eig(f^{ad}; \overline{\lambda})$.

□

Insbesondere:
$\mathbb{K} = \mathbb{R}$: f und f^{ad} haben dieselben reellen Eigenwerte, falls f normal.

$\mathbb{K} = \mathbb{C}$ f und f^{ad} müssen keinen gemeinsamen komplexen Eigenwert haben.

Theorem 6.140. *(Spektralsatz, Fall $\mathbb{K} = \mathbb{C}$)*
Sei V ein unitärer Vektorraum, $\dim(V) < \infty$. *Für $f \in End(V)$ sind äquivalent:*

(i) *f ist normal;*

(ii) *V hat eine ON-Basis aus Eigenvektoren von f;*

(iii) *V ist die orthogonale Summe der Eigenräume von f.*

Beweis:
(iii) $\Rightarrow$ (ii):
$V = \bigoplus_{j=1}^{k} Eig(f;\lambda_j)$, $Eig(f;\lambda_i) \perp \text{Eig}(f;\lambda_j)$.
Wähle ON-Basis der $Eig(f;\lambda_j)$ und füge diese zusammen.

(ii) $\Rightarrow$ (i):
Sei $\mathcal{B}$ eine ON-Basis von f aus Eigenvektoren von $f \Rightarrow A := M_{\mathcal{B}}^{\mathcal{B}}(f) = \text{diag}\,(\lambda_1,\ldots,\lambda_n)$ ist diagonal $\Rightarrow A \cdot \overline{A}^t = \overline{A}^t \cdot A$.

(i) $\Rightarrow$ (iii):
Sei f normal. Behaupte: $\forall\ \lambda \in \mathbb{C}$ ist $\text{Eig}(f;\lambda)^\perp$ f-invariant.
Denn $\text{Eig}(f;\lambda)^\perp = (\ker(f-\lambda id))^\perp = \text{im}(f^{ad} - \overline{\lambda} id)$ nach 6.135.
Nach Voraussetzung ist $f \circ (f^{ad} - \overline{\lambda} id) = (f^{ad} - \overline{\lambda} id) \circ f$
$\Rightarrow f(\text{Eig}(f;\lambda)^\perp) = \text{im}(f \circ (f^{ad} - \overline{\lambda} id)) = \text{im}((f^{ad} - \overline{\lambda} id) \circ f) \subset \text{im}(f^{ad} - \overline{\lambda} id = \text{Eig}(f;\lambda)^\perp$.

Beweis von (iii) durch Induktion nach $\dim(V)$.
Sei λ ein Eigenwert von f, sei $U := \text{Eig}(f;\lambda)$.
Es ist $f(U^\perp) \subset U^\perp$, und $(f|_{U^\perp})^{ad} = f^{ad}|_{U^\perp}$.
Also ist $f|_{U^\perp} \in End(U^\perp)$ normal.
$\dim(U^\perp) < \dim(V) \Rightarrow$ nach Induktion hat $U^\perp$ eine ON-Basis aus Eigenvektoren von f. Ergänze diese durch eine ON-Basis von U.

□

Korollar 6.141.
Sei V ein unitärer Vektorraum, sei $f \in End(V)$.
f *ist* $\left\{\begin{array}{c} \text{selbstadjungiert} \\ \text{anti-selbstadjungiert} \\ \text{unitär} \end{array}\right\} \Leftrightarrow$ *V hat ON-Basis aus Eigenvektoren von f und für alle Eigenwerte λ von f gilt:* $\left\{\begin{array}{c} \lambda \in \mathbb{R} \\ \lambda \in i\mathbb{R} \\ |\lambda| = 1 \end{array}\right\}$

Beweis:
Sofort aus 6.140: Bezüglich ON-Basis hat f eine Matrix $A = \text{diag}\,(\lambda_1,\ldots,\lambda_n)$.

$$f \text{ ist } \left\{\begin{array}{c} \text{selbstadjungiert} \\ \text{anti-selbstadjungiert} \\ \text{unitär} \end{array}\right\} \Leftrightarrow \left\{\begin{array}{c} A = \overline{A}^t \\ A = -\overline{A}^t \\ A \cdot \overline{A}^t = I \end{array}\right\} \Leftrightarrow \left\{\begin{array}{c} \forall\ j : \lambda_j = \overline{\lambda_j} \\ \forall\ j : \lambda_j = -\overline{\lambda_j} \\ \forall\ j : \lambda_j \overline{\lambda_j} = 1 \end{array}\right\}$$

□

Korollar 6.142. *(Matrixversion)*
Sei $A \in M_n(\mathbb{C})$.

(a) $A\overline{A}^t = \overline{A}^t A \Leftrightarrow \exists\, S \in U(n)$ *mit* $S^{-1}AS = \overline{S}^t = diag\,(\lambda_1, \ldots, \lambda_n)$.

(b) Dabei A hermitesch (schiefhermitesch, unitär) $\Leftrightarrow$ alle $\lambda_j \in \mathbb{R}$ (alle $\lambda_j \in i\mathbb{R}$, alle $|\lambda_j| = 1$).

Bemerkungen 6.143.

1. Theorem 6.140 bleibt (bei identischem Beweis) auch für $\mathbb{K} = \mathbb{R}$ richtig, wenn man zusätzlich voraussetzt, dass $P_f(t)$ über $\mathbb{R}$ in Linearfaktoren zerfällt.

2. Ist $A \in M_n(\mathbb{R})$ symmetrisch, so zerfällt $P_A(t)$ über $\mathbb{R}$ in Linearfaktoren: Denn A ist (als komplexe Matrix aufgefasst) hermitesch, hat also nach 6.142 lauter reelle Eigenwerte. Es gibt also ein $S \in O(n)$ mit $S^{-1}AS = StAS$ = (reelle) Diagonalmatrix.

Notation 6.144.
Für $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ sei $\overline{x} := (\overline{x_1}, \ldots, \overline{x_n}) \in \mathbb{C}^n$, $Re(x) := \frac{1}{2}(x + \overline{x}) \in \mathbb{R}^n$, $Im(x) := \frac{1}{2i}(x - \overline{x}) \in \mathbb{R}^n$, also $x = Re(x) + i\,Im(x)$.
Ist $M \subset \mathbb{C}^n$ eine Teilmenge, so sei $\overline{M} := \{\overline{x} : x \in M\} \subset \mathbb{C}^n$.

Lemma 6.145.
Für jeden $\mathbb{C}$-Vektorraum W von $\mathbb{C}^n$ sind äquivalent:

(a) $W = span_{\mathbb{C}}(W \cap \mathbb{R}^n)$;

(b) W hat eine $\mathbb{C}$-Basis aus Vektoren in $\mathbb{R}^n$;

(c) $\exists$ $\mathbb{R}$- *Untervektorraum U von $\mathbb{R}^n$ mit $W = U + iU$;*

(d) $\overline{W} = W$;

(e) $\forall\, z \in W$: $Re(z) \in W$.

Beweis:
(v) $\Rightarrow$ (i):
$w = Re(w) + i\,Im(w)$ für alle $w \in W$ ist also auch $Im(w) \in W$.
$Re(w), Im(w) \in W \cap \mathbb{R}^n \Rightarrow$ (i).

□

Definition 6.146.
Ein $\mathbb{C}$-Untervektorraum W von $\mathbb{C}^n$ ist **über** $\mathbb{R}$ **definiert**, wenn $(i) - (v)$ aus 6.145 gelten.

Bemerkungen 6.147.

1. Ist $\dim_{\mathbb{C}}(W) = 1$, also $W = \mathbb{C} \cdot w$ mit $w \neq 0$, so gilt W ist über $\mathbb{R}$ definiert $\Leftrightarrow$ $Re(w), Im(w)$ sind linear abhängig (über $\mathbb{R}$).
2. Für $W \subset \mathbb{C}^n$ gilt: $\overline{W}^{\perp} = \overline{W^{\perp}}$, denn $\langle u, \overline{w} \rangle = \overline{\langle \overline{u}, v \rangle}$.
   Daher: W über $\mathbb{R}$ definiert $\Leftrightarrow$ $W^{\perp}$ über $\mathbb{R}$ definiert.
3. $W \subset \mathbb{C}^n$ beliebiger $\mathbb{C}$-UR, dann sind $W + \overline{W}$ und $W \cap \overline{W}$ über $\mathbb{R}$ definiert.
4. Sind W_1, W_2 über $\mathbb{R}$ definiert, so auch $W_1 + W_2$ und $W_1 \cap W_2$.

Lemma 6.148.
Für $A \in M_n(\mathbb{C})$ und $\lambda \in \mathbb{C}$ ist $Eig_{\mathbb{C}}(A; \overline{\lambda}) = \overline{Eig_{\mathbb{C}}(A; \lambda)}$.
*(*Dabei sei $\mathrm{Eig}_{\mathbb{C}}(A; \lambda) := \{x \in \mathbb{C}^n : Ax = \lambda x\}$.*)*

Beweis:
$x \in \mathrm{Eig}_{\mathbb{C}}(\overline{A}; \overline{\lambda}) \Leftrightarrow \overline{A}x = \overline{\lambda}x \Leftrightarrow A\overline{x} = \lambda\overline{x} \Leftrightarrow \overline{x} \in \mathrm{Eig}_{\mathbb{C}}(A; \lambda)$.

□

Korollar 6.149.
Sei $AQ \in M_n(\mathbb{R})$, $\lambda \in \mathbb{C}$:
Ist $\lambda \in \mathbb{R}$, so ist $Eig_{\mathbb{C}}(A; \lambda)$ über $\mathbb{R}$ definiert.
Ist $\lambda \in \mathbb{C} \setminus \mathbb{R}$, so ist $Eig_{\mathbb{C}}(A; \lambda) \oplus Eig_{\mathbb{C}}(A; \overline{\lambda})$ über $\mathbb{R}$ definiert.

(Klar aus 6.148)

Satz 6.150.
Sei $A \in M_n(\mathbb{R})$. Dann sind äquivalent:

(i) A ist über $\mathbb{C}$ diagonalisierbar;
(d.h. $\exists\, S \in GL_n(\mathbb{C})$, $S^{-1}AS$ = diagonal)

(ii) $\exists\, S \in GL_n(\mathbb{R})$ *mit*

$$S^{-1}AS = \begin{pmatrix} \alpha_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \alpha_r & & & & & \\ & & & \beta_1 & -\gamma_1 & & & \\ & & & \gamma_1 & \beta_1 & & & \\ & & & & & \ddots & & \\ & & & & & & \beta_s & -\gamma_s \\ & & & & & & \gamma_s & \beta_s \end{pmatrix} \text{ mit } r, s \geq 0 \text{ und reellen } \alpha_j, \beta_k, \gamma_k$$

(*)

Beweis:
(ii) $\Rightarrow$ (i):
klar, denn $\begin{pmatrix} \beta & -\gamma \\ \gamma & \beta \end{pmatrix}$ mit $\beta, \gamma \in \mathbb{R}$ ist über $\mathbb{C}$ diagonalisierbar. (denn: $\gamma \neq 0$, so hat das charakteristische Polynom zwei verschiedene Nullstellen in $\mathbb{C}$: $P(t) = t^2 - 2\beta \cdot t(\beta^2 + \gamma^2)$)

(i) $\Rightarrow$ (ii):
Seien $\lambda_1, \ldots, \lambda_t$ die komplexen Eigenwerte λ von A mit $Im(\lambda) \geq 0$ (dann sind $\lambda_1, \ldots, \lambda_t, \overline{\lambda_t}, \ldots, \overline{\lambda_t}$ alle komplexen Eigenwerte von A). Seien $\lambda_1, \ldots, \lambda_q \in \mathbb{R}, \lambda_{q+1}, \ldots, \lambda_t \in \mathbb{C} \setminus \mathbb{R}$.
Sei $U_j := \mathrm{Eig}_R(A; \lambda_j)$ für $j = 1, \ldots, q$;
und $U_j := \mathbb{R}^n \cap \left(\mathrm{Eig}_{\mathbb{C}}(A; \lambda_j) \oplus \mathrm{Eig}_C(A; \overline{\lambda})\right)$ für $j = q+1, \ldots, t$.

Dann ist $\mathbb{R}^n = U_1 \oplus \ldots \oplus U_t$.
Denn $\mathbb{C}^n = \bigoplus_{j=1}^{q} \underbrace{\mathrm{Eig}_C(A; \lambda_j)}_{\mathbb{R}-definiert} \oplus \bigoplus_{j=q+1}^{t} \underbrace{\left(\mathrm{Eig}_{\mathbb{C}}(A; \lambda_j) \oplus \mathrm{Eig}_{\mathbb{C}}(A; \overline{\lambda_j})\right)}_{\mathbb{R}-definiert}$

Für $j = 1, \ldots, q$ sei $\mathcal{F}_j$ eine $\mathbb{R}$-Basis von U_j. Sei jetzt $j \in \{q+1, \ldots, t\}$. Sei $(w_1, \ldots, w_p)$ eine $\mathbb{C}$-Basis von $\mathrm{Eig}_{\mathbb{C}}(A; \lambda_j)$. Dann ist $(\overline{w_1}, \ldots, \overline{w_p})$ eine $\mathbb{C}$-Basis von $\mathrm{Eig}_C(A; \overline{\lambda_j})$. Setze $x_k := Re(w_k)$, $y_k := Im(w_k)$, $k = 1, \ldots, p$, und setze $\mathcal{F}_j = (x_1, y_1, \ldots, x_p, y_p)$.
Dann ist $\mathcal{F}_j$ eine $\mathbb{R}$-Basis von U_j. Setze $\beta_j := Re(y_j), \gamma_j := Im(\lambda_j)$, dann ist für $1 \leq k \leq p$:

$$\lambda_j w_k = Aw_k = \underbrace{A(x_k + iy_k)}_{=Ax_k+iAy_k} = (\beta_j + i\gamma_j)(x_k + iy_k) = (\beta_j x_k - \gamma_j y_k) + i(\beta_j y_k + \gamma_j x_k)$$

und es folgt (vergleiche Real- und Imaginärteil)
$Ax_k = \beta_j x_k - \gamma_j y_k, Ay_k = \beta_j y_k + \gamma_j x_k$.

Für die Basis $\mathcal{F} := \mathcal{F} \sqcup \ldots \sqcup \mathcal{F}_t$ von $\mathbb{R}^n$ gilt also $S^{-1}AS$ hat die Form (*) mit $S := T^{\mathcal{F}}_{\mathcal{K}}$.
□

Theorem 6.151. *(Spektralsatz, $\mathbb{K} = \mathbb{R}$)*
Sei V ein euklidischer Vektorraum mit $\dim(V) < \infty$, *und $f \in End_{\mathbb{R}}(V)$.*
Es sind äquivalent:

(i) f ist normal;

(ii) V hat eine ON-Basis, bezüglich der f durch eine Matrix der Form (*) beschrieben ist.

Beweis:
(ii) $\Rightarrow$ (i):
Für die Matrix A aus (*) gilt $A \cdot A^t = A^t \cdot A$

(i) $\Rightarrow$ (ii):
Wir können annehmen $V = \mathbb{R}^n$ (mit Standard-SKP). Sei $A \in M_n(\mathbb{R})$ die Matrix von f bezüglich der kanonischen Basis. Nach dem Spektralsatz über $\mathbb{C}$ (6.140) ist A über $\mathbb{C}$ diagonalisierbar und die $\mathbb{C}$-Eigenräume von A sind paarweise orthogonal.
Gehe in Beweis von 6.150:
Für $\lambda_j \in \mathbb{R}$ wähle $\mathcal{F}_j$ als ON-Basis von U_j; für $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$ wähle $(w_1, \ldots, w_p)$ als ON-Basis von $\mathrm{Eig}_{\mathbb{C}}(A; \lambda_j)$.
Dann ist $(\overline{w}_1, \ldots, \overline{w_p})$ eine ON-Basis von $\mathrm{Eig}_{\mathbb{C}}(A; \overline{\lambda_j})$. Sei wieder $x_k = Re(w_k), y_k = Im(w_k)$. Für $k \neq l$ ist $\{w_k, \overline{w_k}\} \perp \{w_l, \overline{w_l}\}$.
Für $k = l$ ist $0 = \langle w_k, \overline{w_k} \rangle = \langle x_k + iy_k, x_k - iy_k \rangle = \|x_k\|^2 - \|y_k\|^2 + 2i \langle x_k, y_k \rangle$.
Also $\|x_k\| = \|y_k\|$ und $\langle x_k, y_k \rangle = 0$.
Ebenso für $k \neq l$: $\{x_k, y_k\} \perp \{x_l, y_l\}$.
Multipliziere x_k und y_k mit $\frac{1}{\|x_k\|} = \frac{1}{\|y_k\|}$
$\Rightarrow (x_1, y_1, \ldots, x_p, y_p) = \mathcal{F}_j$ ist ON-Basis von U_j.
$\Rightarrow$ Aussage (ii).

□

Theorem 6.152.
Sei V ein euklidischer Vektorraum ($n = \dim(V) < \infty$), sei $f \in End_{\mathbb{R}}(V)$.

(a) *f ist selbstadjungiert $\Leftrightarrow$ V hat eine ON-Basis aus Eigenvektoren von f.*

(b) *f ist antiselbstadjungiert $\Leftrightarrow$ V hat eine ON-Basis, für die zu f die Matrix (*) mit $\alpha_j, \beta_j = 0, \gamma_1, \ldots, \gamma_s \in \mathbb{R}$ gehört.*

(c) *f ist orthogonal $\Leftrightarrow$ V hat eine ON-Basis, bezüglich der f eine Matrix der Form*

$$\begin{pmatrix}
1 & & & & & & & & & \\
& \ddots & & & & & & & & \\
& & 1 & & & & & & & \\
& & & -1 & & & & & & \\
& & & & \ddots & & & & & \\
& & & & & -1 & & & & \\
& & & & & & \cos\theta_1 & -\sin\theta_1 & & \\
& & & & & & \sin\theta_1 & \cos\theta_1 & & \\
& & & & & & & & \ddots & \\
& & & & & & & & \cos\theta_s & -\sin\theta_s \\
& & & & & & & & \sin\theta_s & \cos\theta_s
\end{pmatrix}$$

$(\theta_1, \ldots, \theta_s \in \mathbb{R})$ hat.

Beweis:
klar aus 6.151: die Matrix A aus (*) erfüllt (a) $A = A^t$, (b) $A = -A^t$, (c) $AA^t = I$ genau dann, wenn sie die in (a), (b), (c) angegebene Form hat.

□

Korollar 6.153. *(Matrixversion von 6.152)*
Eine Matrix $A \in M_n(\mathbb{R})$ ist genau dann symmetrisch (bzw. schiefsymmetrisch, bzw. orthogonal), wenn es eine Matrix $S \in O(n)$ gibt, so dass $S^{-1}AS = S^tAS$ eine Diagonalmatrix (bzw. vom Typ 6.152 (b), bzw. 6.152 (c)) ist.

Bemerkung 6.154.
Sei $S \in M_n(\mathbb{R})$ symmetrisch. Wir haben gesehen in 6.152:
Es existiert eine orthogonale Matrix $T \in O(n)$, so dass $T^{-1}ST = T^tST$ = Diagonalmatrix.

Das präzisiert Satz 6.29 (Diagonalisierung von symmetrischen Matrizen): dort wurde gezeigt, dass ein $T \in GL_n(\mathbb{R})$ existiert mit T^tST = Diagonalmatrix. Die jetzt bewiesene Matrix Aussage sagt, dass T dabei orthogonal gewählt werden kann. Das bedeutet, dass T simultan eine $\simeq$-Transformation (Kongruenz: T^tST) und eine $\approx$-Transformation (Ähnlichkeit: $T^{-1}ST$) bewirkt.
Ebenso kann man den schiefsymmetrischen Fall von 6.152 als Präzisierung von Satz 6.54 sehen.

Sei weiter stets V ein $\mathbb{K}$-Hilbertraum, $\dim(V) < \infty$.

Definition 6.155.

(a) Ist $A \in M_n(\mathbb{K})$ symmetrisch ($\mathbb{K} = \mathbb{R}$) bzw. hermitesch ($\mathbb{K} = \mathbb{C}$), so schreibt man $A \geq 0$, falls A positiv semidefinit ist. Die Menge aller solchen A wird mit $P_n(\mathbb{K})$ bezeichnet.

(b) Ein $f \in End(V)$ mit $f = f^{ad}$ heißt **positiv** (bzw. **nichtnegativ**), falls $\forall\ v \in V, v \neq 0$: $\langle f(v), v\rangle > 0$ (bzw. $\langle f(v), v\rangle \geq 0$).
Man setzt $P(V) := \{f \in End(V) : f^{ad} = f, f \text{ nichtnegativ}\}$ und schreibt $f \geq 0$ $\Leftrightarrow$ f nichtnegativ.

Beachte: Auch für $\mathbb{K} = \mathbb{C}$ ist $\langle f(v), v\rangle = \left\langle v, f^{ad}(v)\right\rangle = \langle v, f(v)\rangle = \overline{f(v), v}$ reell.

Bemerkungen 6.156.

1. Ist $\mathcal{B}$ eine ON-Basis von V, $f \in End(V)$ und $f = f^{ad}$, so ist f genau dann positiv (bzw. nichtnegativ), wenn $M_{\mathcal{B}}^{\mathcal{B}}(f)$ positiv definit (bzw. positiv semidefinit) ist.
($A = M_{\mathcal{B}}^{\mathcal{B}}(f), v \in V$, Koordinatenvektor $x \in \mathbb{K}^n$, dann $\langle f(v), v\rangle = x^t A\bar{x}$).

2. Für alle $f \in End(V)$ ist $f^{ad} \circ f \geq 0$, denn $(f^{ad} \circ f)^{ad} = f^{ad} \circ f$, und für $v \in V$ ist $\left\langle (f^{ad} \circ f)(v), v\right\rangle = \langle f(v), f(v)\rangle = \|f(v)\|^2 \geq 0$.

Satz 6.157.
Für $f \in End(V)$, $f^{ad} = f$, gilt:
f positiv (bzw. nichtnegativ) $\Leftrightarrow$ alle (reellen!) Eigenwerte von f sind > 0 (bzw. sind ≥ 0).

Beweis:
Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine ON-Basis von V aus Eigenvektoren von f, etwa $f(v_j) = \lambda_j v_j$ ($\lambda_j \in \mathbb{R}$). Dann ist $M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$. Nach Bemerkung 1. (6.156) folgt: f positiv (nichtnegativ) $\Leftrightarrow$ alle $\lambda_j > 0$ ($\lambda_j \geq 0$). □

Satz 6.158.
$P(V)$ ist ein abgeschlossener und konvexer Kegel in $End(V)$. Das bedeutet: $P(V)$ ist abgeschlossen, und erfüllt: $f, g \in P(V) \Rightarrow f + g \in P(V)$ und $\lambda f \in P(V)\ \forall\ 0 \leq \lambda \in \mathbb{R}$.

Beweis:
Sind $f, g \geq 0$, so ist
$\langle (f+g)(v), v\rangle = \langle f(v), v\rangle + \langle g(v), v\rangle \geq 0$.
$\langle (\lambda f)(v), v\rangle = \langle \lambda \cdot f(v), v\rangle = \lambda \langle f(v), v\rangle \geq 0$.
Also ist $P(V)$ ein konvexer Kegel.

Abgeschlossenheit von $P(V)$:
wir zeigen, dass die Menge $P_n(\mathbb{K}) = \{A = \overline{A}^t : A \geq 0\} \subset M_n(\mathbb{K})$ abgeschlossen ist: ist $(A_\nu)_{\nu\in\mathbb{N}}$ eine Folge in $P_n(\mathbb{K})$, un existiert $A = \lim\limits_{\nu\to\infty} A_\nu$, so ist $\forall\ x \in \mathbb{K}^n$ $x^t A\overline{x} = \lim\limits_{\nu\to\infty} \underbrace{(x^t A_\nu \overline{x})}_{\geq 0} \geq 0$.

□

Satz 6.159.
Sei $f \in P(V)$, sei $r \in \mathbb{N}$. Dann existiert genau ein $g \in P(V)$ mit $g^r = f$.
Man schreibt $\sqrt[r]{f} := g$.

Beweis:
Sei $(v_1, \ldots, v_n)$ eine ON-Basis mit $f(v_j) = \lambda_j v_j$, mit $0 \leq \lambda_j \in \mathbb{R}$: Definiere g durch $g(v_j) := \sqrt[r]{\lambda_j} \cdot v_j$ (dabei sei $\sqrt[r]{\lambda}$ die eindeutige reelle r-te Wurzel ≥ 0).

Dann ist $g = g^{ad}$ und $g^r = f$ (klar).

Eindeutigkeit:
Sei $\widetilde{g} \in P(V)$ mit $\widetilde{g}^r = f$. Ist
$V = \mathrm{Eig}(\widetilde{g}, \mu_1) \oplus \ldots \oplus \mathrm{Eig}(\widetilde{g}, \mu_m)$ (*)
die Eigenraumzerlegung, $\mu_j \geq 0$ reell, so ist $\widetilde{g}^r|\mathrm{Eig}(\widetilde{g}, \mu_j) = \lambda_j^r \cdot id$. Mit den μ_j sind auch die μ_j^r paarweise verschieden.
Also ist (*) auch die Eigenraumzerlegung von $f \Rightarrow \widetilde{g} = g$.

□

Korollar 6.160.
Für alle $f \in End(V)$ gilt:
$f = f^{ad}$ und $f \geq 0 \Leftrightarrow\ \exists\ g \in End(V)$ mit $f = g^{ad} \circ g$.

Beweis:

„$\Leftarrow$“ : ✓

„⇒“ : $f = \sqrt{f} \circ \sqrt{f} = (\sqrt{f})^{ad} \circ \sqrt{f}$.

□

($A = A^t$ eine reelle Matrix: $A \geq 0 \Rightarrow A = B^t \cdot B$ - gesehen in § 2).

Lemma und Definition 6.161.
Für $f, g \in End(V)$ sei $\langle f, g \rangle := tr(f \circ g^{ad})$. Das definiert ein Skalarprodukt auf dem $\mathbb{K}$-Vektorraum der Endomorphismen über V ($End(V)$).

Beweis:
Biadditivität: ✓.

$\langle -, - \rangle$ homogen (bzw. antihomogen) in der ersten (bzw. zweiten) Komponente: ✓.

zu zeigen bleibt: $\langle f, f \rangle > 0$ für $f \neq 0$.
Wir zeigen dies (O.E.) für $V = \mathbb{K}^n$:
$f = A \in M_n(\mathbb{K})$, $\langle f, f \rangle = \text{tr}(A \cdot \overline{A}^t) = \sum_{j=1}^{n} (A\overline{A}^t)_{jj} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jjk} \cdot \overline{a_{jk}} \geq 0$ und $= 0 \Leftrightarrow A = 0$
(hier: $A = (a_{jk})$.

□

Nebenbemerkung:
$A = (a_{jk}), B = (b_{jk})$ ist $\langle A, B \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} \overline{b_{jk}}$.

Die Menge $\{f \in End(V) : f = f^{ad}\}$ ist ein $\mathbb{R}$-Untervektorraum von $End(V)$.
Auf diesem haben wir also das Skalarprodukt $\langle f, g \rangle = \text{tr}(f \circ g)$ definiert.

Theorem 6.162.
Für alle $f, g \in P(V)$ ist $\langle f, g \rangle \geq 0$.
Dabei gilt $\langle f, g \rangle = 0 \Leftrightarrow f \cdot g = 0$.

Das bedeutet also Der Winkel zwischen f und g ist also spitz ($\in [-90°, 90°]$). $P(V)$ ist also ein spitzer Kegel.

Beweis:
$\langle f, g \rangle = \text{tr}(f \circ g) = \text{tr}(\sqrt{f} \circ \sqrt{f} \circ g) = \text{tr}(\sqrt{f} \circ g \circ \sqrt{f})$. Setze $h := \sqrt{f} \circ g \circ \sqrt{f}$.

Es ist $h^{ad} = h$ (da alle Elemente selbstadjungiert und Produkt symmetrisch), und auch $h \geq 0$.
$\langle hv, v\rangle = \left\langle \sqrt{f} \circ g \circ \sqrt{f}(v), v\right\rangle = \left\langle g(\sqrt{f}(v)), \sqrt{f}(v)\right\rangle \geq 0$ wegen $g \geq 0$.
Sind $\mu_1, \ldots, \mu_n \geq 0$ die (reellen) Eigenwerte von h, so ist $\langle f, g\rangle = \operatorname{tr}(h) = \sum_j \mu_j \geq 0$.

Genau dann ist $\langle f, g\rangle = \operatorname{tr}(h) = 0$, wenn alle $\mu_j = 0$ sind, also genau dann, wenn $h = \sqrt{f} \circ g \sqrt{f} = 0$ ist. Aus $h = 0$ folgt
$0 = \langle h(x), x\rangle = \left\langle \gamma \circ \sqrt{f}(x), \sqrt{f}(x)\right\rangle \ \forall\, x \in V$.
Schreibe $\sqrt{f}(x) = \sum_{j=1}^{n} y_j$ mit $g(y_j) = \lambda_j y_j$, $y_j \perp y_k$, $j \neq k$ und $\lambda_j \geq 0$ (Eigenraumzerlegung von g).
Dann folgt $0 = \left\langle g(\sqrt{f}(x)), \sqrt{f}(x)\right\rangle = \left\langle \sum_j l a_j y_j, \sum_j y_j\right\rangle = \sum_j \lambda_j \|y_j\|^2$.
Also ist $\lambda_j = 0$ für alle j mit $y_j \neq 0$, d.h. $\sqrt{f}(x) = \sum_j y_j \in \ker(g)$.
Also: $\langle f, g\rangle = 0 \Rightarrow g \circ \sqrt{f} = 0 \Rightarrow g \circ f = 0$.

Umgekehrt: $\langle f, g\rangle = tr(f \circ g) = 0$, für $f \circ g = 0$.

□

Theorem 6.163.
Sei $f \in End_{\mathbb{K}}(V)$. Dann gibt es $p, g \in End(V)$ mit $p \geq 0$, g orthogonal bzw. unitär und $f = p \circ g$. Dabei ist p eindeutig. Ist f invertierbar, so ist auch g eindeutig.

Beweis:
Angenommen, wir haben solche p und g.
$\Rightarrow f \circ f^{ad} = p \circ \underbrace{(g \circ g^{ad})}_{id} \circ p^{ad} = p \circ p^{ad} = p^2$.
Also ist dann $p = \sqrt{f \circ f^{ad}}$ (und ist damit eindeutig).

Umgekehrt definiere p durch $p = \sqrt{f \circ f^{ad}}$ (6.159).

Betrachte zunächst den Fall, dass f invertierbar ist.
Dann ist p invertierbar (denn $f^{ad} \circ f$ invertierbar), setze $g := p^{-1} \circ f$. Behaupte g ist orthogonal bzw. unitär;
$g \circ g^{ad} = p^{-1} \circ \underbrace{(f \circ f^{ad})}_{p^2} \circ p^{-1} = id$.

Allgemeiner Fall (f nicht notwendig invertierbar); Beweis nur für $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^n$:
Die orthogonale Gruppe $O(n)$ ist kompakt. Das Bild der Produktabbildung $P_n(\mathbb{R}) \times O(n) \to M_n(\mathbb{R})$, $(p, g) \mapsto pg$.
Da $O(n)$ kompakt ist, ist das Bild dieser Abbildung abgeschlossen. Da $GL_n(\mathbb{R})$ in diesem Bild enthalten und dicht in $M_n(\mathbb{R})$ ist, ist die Abbildung surjektiv.

□

Lemma 6.164.
($\mathbb{K} = \mathbb{C}$)
$f \in End(V)$ ist unitär $\Leftrightarrow$ $f = \exp(ih)$ mit einem selbstadjungierten $h \in End(V)$.

Beweis:
f unitär heißt: es existiert eine ON-Basis $(v_1, \ldots, v_n)$ von V mit $f(v_j) = \lambda_j v_j$ und $|\lambda_j| = 1$ (Spektralsatz).

Ist f unitär, so wähle reelle Zahlen $a_j \in \mathbb{R}$ mit $e^{ia_j} = \lambda_j$ $(j = 1, \ldots, n)$, und definiere $h \in End(V)$ durch $h(v_j) := a_j v_j$ $(j = 1, \ldots, n)$.
Dann ist $f = \exp(ih)$, und h ist selbstadjungiert.

□

Korollar 6.165. *(Polarzerlegung)*
($\mathbb{K} = \mathbb{C}$)
Jedes $f \in End(V)$ hat eine Darstellung $f = p \circ \exp(ih)$ mit p, h selbstadjungiert und $p \geq 0$. Dabei ist p eindeutig bestimmt.

Beweis:
6.163 und 6.164

□

Korollar 6.166.
($\mathbb{K} = \mathbb{C}$)
Jedes $A \in M_n(\mathbb{C})$ lässt sich darstellen als $A = P \cdot e^{iH}$ mit P, H hermitesch und $P \geq 0$, P ist eindeutig.

Beispiel:
Fall $n = 1$: $\mathbb{C} \ni a = r \cdot e^{i\varphi}, r, \varphi \in \mathbb{R}, r \geq 0$.
(vergleiche Analysis I)

Die letzten beiden Korollare sind eine Verallgemeinerung von dieser bekannten Zerlegung.

7. Moduln über Hauptidealringen

a. Ringe und Ideale

7.1.
Sei A ein Ring: $(A, +)$ eine abelsche Gruppe mit neutralem 0; Multiplikation $\cdot : A \times A \to A$ assoziativ, distributiv.
Wir fordern hier stets: A ist ein Kommutativer Ring, d.h. $\forall\, a, b \in A$ gilt $ab = ba$.
Und stets: A hat eine Eins, d.h. $\exists\, 1 \in A$ mit $1 \cdot a = a\ \forall\, a \in A$.
(Diese 1 *ist eindeutig bestimmt!)*

Wichtige Beispiele:
Alle Körper $\mathbb{K}$, die Polynomringe über Körpern $\mathbb{K}[t]$, $\mathbb{Z}$, die Ringe $\mathbb{Z}/\mathbb{Z}n (n \in \mathbb{N})$, $\mathbb{Z}[i] := \{a+bi : a, b \in \mathbb{Z}\}$ oder $\mathbb{Z}[t]$, oder allgemeiner Polynomringe über Ringe $A[t]$.

Sei also stets A ein kommutativer Ring mit Eins.

Lemma 7.2.
Ein $a \in A$ heißt eine ***Einheit*** *von A, wenn ein $b \in A$ mit $ab = 1$ existiert.*
Dann ist b eindeutig, und man schreibt $b = \frac{1}{a} = a^{-1}$.
Die Menge A^* der Einheiten von A ist eine Gruppe (abelsch) bezüglich der Multiplikation.

Beweis:
$ab = 1 = ac \Rightarrow b = b \cdot 1 = bac = 1 \cdot c = c$.

$a, b \in A^* \Rightarrow ab \in A^*$: $(ab)(b^{-1}a^{-1}) = 1$.

□

Beispiel 7.3.
A ist ein Körper $\Leftrightarrow A^* = A \setminus \{0\}$.
$A = \mathbb{Z} : A^* = \{\pm 1\}$.
$A = \mathbb{K}[t]$, $\mathbb{K}$ ein Körper: $\mathbb{K}[t]^* = \mathbb{K}^* (= \mathbb{K} \setminus \{0\})$.

7.4.
$A_1, \ldots, A_n$ Ringe, so ist $A_1 \times \ldots \times A_n$ ein Ring bezüglich komponentenweiser Addition und Multiplikation.
Die Eins ist $(1, \ldots, 1)$.

7.5.
Eine Teilmenge $I \subset A$ heißt ein **Ideal** von A, wenn $(I, +)$ eine Untergruppe von $(A, +)$ ist und gilt $\forall\, a \in A,\ \forall\, b \in I$: $ab \in I$.

Für jeden Ringhomomorphismus $\varphi : A \to B$ (stets $\varphi(1) = 1$) ist $\ker(\varphi) = \{a \in A : \varphi(a) = 0\}$ ein Ideal von A.

Sei $I \subset A$ ein Ideal von A. Wir definieren den **Quotientenring** A/I:
Für $a, b \in A$ sei $a \equiv b \pmod{I} :\Leftrightarrow a - b \in I$.
Dies ist eine Äquivalenzrelation auf A, die Äquivalenzklassen sind die $a + I = \overline{a}$ $(a \in A)$.
$A/I := \{a + I : a \in A\}$ als Menge.
Durch die Definitionen
$(a + I) + (b + I) := (a + b) + I$
$(a + I) \cdot (b + I) := ab + I$
wird $(A/I, +, \cdot)$ ein kommutativer Ring mit 1.
Die Wohldefiniertheit von $+, \cdot$ folgt daraus, dass I ein Ideal ist.
Die Abbildung $\pi : A \to A/I, \pi(a) := a + I$, ist ein surjektiver Ringhomomorphismus mit $\ker(\pi) = I$.

Satz 7.6. *(Homomorphiesatz)*
Ist $\varphi : A \to B$ ein surjektiver Ringhomomorphismus, $I := \ker(\varphi)$, so induziert φ einen Ringisomorphismus $\overline{\varphi} : A/I \to B$, nämlich $\overline{\varphi}(a + I) = \varphi(a)$.

(Beweis analog zu LA I)

7.7.
Ist $(I_\lambda)_{\lambda \in \Lambda}$ eine Familie von Idealen in A, so ist auch $\bigcap\limits_{\lambda \in \Lambda} I_\lambda$ ein Ideal in A.

zu jeder Teilmenge $X \subset A$ gibt es also ein eindeutiges kleinstes Ideal I mit $X \subset I$, nämlich $I = \bigcap\{J : J \text{ ist Ideal}, J \supset X\}$.
Man schreibt $I =: (X)$ und nennt (X) das **von** X **erzeugte Ideal von** A. Ist $X = \{a_1, \ldots, a_n\}$ endlich, schreibt man auch $(X) = (a_1, \ldots, a_n)$.

Lemma 7.8.
Sei $X \subset A$. (X) ist die Menge aller endlichen Summen $a_1x_1 + \ldots + a_rx_r$ mit $r \in \mathbb{N}$, $a_i \in A, x_i \in X$.

Beweis:
Jedes Ideal $J \supset X$ enthält auch diese Summen. Umgekehrt bildet die Menge dieser Summen auch ein Ideal, ist also gleich (X).

□

Insbesondere:
$(a_1, \ldots, a_n) = Aa_1 + \ldots + Aa_n$.

Bemerkungen 7.9.

1. Ein Ideal I heißt ein **Hauptideal**, wenn es ein $a \in A$ gibt mit $I = (a) = Aa$. Beispiele $I = (0) = \{0\}, I = (1) = A$.
2. Sind I, J zwei Ideale, so ist $I \cup J$ im Allgemeinen kein Ideal. Das von $I \cup J$ erzeugte Ideal ist vielmehr $I + J := \{a + b : a \in I, b \in J\}$. genannt die **Summe** von I und J.

   Insbesondere: $(a_1, \ldots, a_n) + (b_1, \ldots, b_r) = (a_1, \ldots, a_n, b_1, \ldots, b_r)$.
3. Die Ideale I, J heißen **relativ prim**, falls $I + J = (1)$ ist.

Definition 7.10.
Seien I, J zwei Ideale von A. Das **Idealprodukt** $I \cdot J = IJ$ ist definiert als $IJ := (ab : a \in I, b \in J)$.

Bemerkungen 7.11.

1. $IJ = \{a_1b_1 + \ldots + a_nb_n : n \in \mathbb{N}, a_1, \ldots, a_n \in I, b_1, \ldots, b_n \in J\}$.
2. $I = (a_1, \ldots, a_n)$, $J = (b_1, \ldots, b_r)$, dann ist $IJ = (a_ib_j : i = 1, \ldots n, j = 1, \ldots, r)$.
3. Sei $A = \mathbb{Z}$. In $\mathbb{Z}$ haben wir die Hauptideale $(n) = \{an : a \in \mathbb{Z}\}$ für $n \in \mathbb{Z}$.
   (Tatsächlich sind alle Ideale in $\mathbb{Z}$ - siehe später)
   $m, n > 0$: $(m) + (n) = (m, n) = (ggT(m, n))$
   $(m) \cap (n) = (kgV(m, n))$
   $(m) \cdot (n) = (mn)$.

   Der Quotientenring $\mathbb{Z}/(n)$ ist gerade $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$ aus LA I I.3.

Lemma 7.12.
Ist $I + J = (1)$ (d.h. I, J sind relativ prim), so ist $IJ = I \cap J$.
Im allgemeinen gilt dagegen nur $IJ \subset I \cap J$.

Beweis:
$IJ \subset I \cap J$: klar.
Sei $I + J = (1)$, d.h. habe $1 = a + b$ mit $a \in I, b \in J$. Sei $x \in I \cap J$, dann ist $x = 1 \cdot x = (a + b)x = ax + bx, a, x \in I, b, x \in J. \Rightarrow ax + bx \in IJ$.

□

Beispiel:
$A = \mathbb{Z}$:
$(4) \cdot (6) = (24)$, aber $(4) \cap (6) = (12) \neq (24)$.

Definition 7.13.
Ein Ideal I von A heißt **maximal**, falls $I \neq (1)$ und für alle Ideale J mit $J \supset I$ gilt: $J = I$ oder $J = (1)$.

Satz 7.14.
I ist genau dann maximal, wenn der Ring A/I ein Körper ist.

Beweis:
I maximal $\Leftrightarrow \ \forall\, a \in A, a \notin I$ ist $I + (a) = (1) \Leftrightarrow \ \forall\, a \in A, a \notin I\ \exists\, b \in A$ mit $1 \equiv ab \pmod{I} \Leftrightarrow \ \forall\, 0 \neq x \in A/I\ \exists\, y \in A/I$ mit $xy = 1$ (in A/I) $\Leftrightarrow A/I$ ist ein Körper.

□

Beispiel:
$A = \mathbb{Z}$: sei $n \in \mathbb{N}$. Das Ideal (n) ist genau dann maximal, wenn gilt: n ist Primzahl.

b. Teilbarkeit

7.15.
Der Ring A heißt nullteilerfrei, wenn $\forall\, a, b \in A$ gilt: $ab = 0 \Rightarrow a = 0 \vee b = 0$.
In einem solchen Ring darf man **kürzen**:
Aus $ac = bc$ und $c \neq 0$ folgt $a = b$.

Beispiele 7.16.
Beispiele nullteilerfreier Ringe: Körper $\mathbb{K}$, Polynomringe $\mathbb{K}[t]$, die Ringe $\mathbb{Z}$, $\mathbb{Z}[i]$, ...

Dagegen ist $A_1 \times A_2$ nie nullteilerfrei für $A_1, A_2 \neq \{0\}$: $(0,0) = (1,0) \cdot (0,1)$.

$\mathbb{Z}/(6)$ hat Nullteiler: $6 = 2 \cdot 3$.

Definition 7.17.

(a) Für $a, b \in A$ sagt man $a \mid b$ („a teilt b"), falls ein $c \in A$ existiert mit $b = ac$. Man nennt dann a einen **Teiler** von b.

(b) $a, b \in A$ heißen **assoziiert**, i. Z. $a \sim b$, wenn gilt $a \mid b$ und $b \mid a$.

Lemma 7.18.
Sei A stets nullteilerfrei, seien $a, b, c \in A$.

(a) $a \mid b \Leftrightarrow (b) \subset (a)$;

(b) $a \mid b \wedge b \mid c \Rightarrow a \mid c$; *(Transitivität)*

(c) ist $c \neq 0$, *so* $a \mid b \Leftrightarrow ac \mid bc$;

(d) $a \mid b \wedge a \mid c \Rightarrow a \mid (b + c)$;

(e) $a \sim b \Leftrightarrow (a) = (b) \Leftrightarrow\ \exists\, \varepsilon \in A^*\ b = \varepsilon a$.

Beweis:

(a) $a \mid b \Leftrightarrow\ \exists\, c \in A\ b = ac \Leftrightarrow b \in (a) \Leftrightarrow (b) \subset (a)$.

(e) $a \sim b \Leftrightarrow (a) = (b)$. Ist $a \sim b$, etwa $b = ac$ und $a = bd \Rightarrow a = acd \Rightarrow a(1 - cd) = 0$. Ist $a = 0$, so ist auch $b = 0$. Ist $a \neq 0$, so $c, d \in A^*$.

□

Bemerkung 7.19.
Sei $a|b$. Ist $a \nsim 1, a \nsim b$, so nennt man a einen **echten Teiler** von b.
Es gilt: $a \sim 1 \Leftrightarrow a \in A^*$.

Beispiel:
Sein $0 \neq f, g \in K[t]$ mit $f \mid g$, so ist genau dann f echter Teiler von g, wenn $1 \leq \deg(f) < \deg(g)$ ist.

Definition 7.20.
Ein Element $0 \neq a \in A$ heißt **unzerlegbar**, wenn a keinen echten Teiler hat, also wenn gilt: $a = bc$ mit $b, c \in A \Rightarrow b \in A^* \vee c \in A^*$.

Bemerkungen 7.21.

1. $A = \mathbb{Z}$, dann sind die unzerlegbaren Elemente genau die $\pm p$, p ist Primzahl.

2. Ist $a \sim a'$, so: a unzerlegbar $\Leftrightarrow a'$ unzerlegbar.
   $(a' = b' \cdot c', a = \varepsilon a', \varepsilon \in A^* \Rightarrow a = \varepsilon b'c' \Rightarrow b' \sim 1 \vee c' \sim 1)$

3. $f \in \mathbb{K}[t]$ ($\mathbb{K}$ stets ein Körper) ist genau dann unzerlegbar in $\mathbb{K}[t]$, wenn f **irreduzibel** ist, d.h. wenn $\deg(f) \geq 1$ ist und f keinen Teiler g mit $1 \leq \deg(g) < \deg(f)$ hat.
   Ist $\mathbb{K}$ algebraisch abgeschlossen, so sind die irreduziblen Polynome in $\mathbb{K}[t]$ genau die linearen Polynome. *(Beispiel:* $\mathbb{K} = \mathbb{C}$*)*
   Ist $\mathbb{K}$ nicht algebraisch abgeschlossen, so gibt es irreduzible Polynome in $\mathbb{K}[t]$ von Grad ≥ 2. *(Beispiel:* $\mathbb{K} = \mathbb{R}$*)*
   $f = t^2 + at + b \in \mathbb{R}[t]$ ist genau dann irreduzibel in $\mathbb{R}[t]$, wenn $a^2 - 4b < 0$ ist.

4. $0 \neq a \in A$. Ist a nicht unzerlegbar, so $a = bc$ mit $b, c \nsim 1$. Ist auch b oder c zerlegbar, so kann man a weiter zerlegen, usw.
   Frage: Kommt man nach endlich vielen Schritten an ein Ende?

Definition 7.22.
Eine **unendliche echte Teilerkette** in A ist eine Folge $(a_n)_{n \in \mathbb{N}}$ derart, dass $a_1 \neq 0$ ist und gilt $a_{n+1} \mid a_n$, $a_n \nmid a_{n+1}$ für alle $n = 1, 2, \ldots$.

Satz 7.23.
(A nullteilerfrei)
In A gebe es keine echte unendliche Teilerkette. Dann hat jede Nichteinheit $a \neq 0$ eine Darstellung $a = a_1 \cdot \ldots \cdot a_n$ mit $n \in \mathbb{N}$, $a_i \in A$ unzerlegbar.

Beweis:
Sei $U :=$ Menge aller $a \neq 0$, $a \notin A^*$, die Produkt von endlich vielen unzerlegbaren Elementen sind. Sei $0 \neq a \in A$, $a \notin A^*$:
Annahme: $a \notin U$. Also $a = b \cdot c$ mit $b, c \notin A^*$. Wären $b, c \in U$, so auch $a \in U$. Also O.E. $b \notin U$. Nun induktiv so weiter $b = d \cdot e$ mit $d, e \notin A^*$, und O.E. $d \notin U$ ... $\Rightarrow$ eine unendliche echte Teilerkette.

□

Bemerkung 7.24.

1. In $A = \mathbb{Z}$ gibt es keine echten unendlichen Teilerketten (klar!).
2. Ebenso nicht in $A = \mathbb{K}[t]$, da aus $f|g$, $f \nsim g$ folgt $\deg(f) < \deg(g)$.
   Jedes $0 \neq a \in A \setminus A^*$ ist also endliches Produkt von unzerlegbaren Elementen.

Definition 7.25.
Ein nullteilerfreier Ring A heißt **Hauptidealring**, wenn jedes Ideal von A Hauptideal ist, also von der Form (a) mit $a \in A$ ist.

Beispiel:
Jeder Körper ist aus trivialen Gründen ein Hauptidealring, denn (0) und (1) sind die einzige Ideale.
„Echte“ Beispiele sind $\mathbb{Z}$, $\mathbb{K}[t]$. Beweis später.

Lemma 7.26.
Sei A ein Hauptidealring, sei $p \in A$ unzerlegbar. Dann gilt für alls $a_1, \ldots, a_n \in A$:
$p \mid a_1 \cdot \ldots \cdot a_n \Rightarrow p \mid a_i$ für ein $i \in \{1, \ldots, n\}$.

Beweis:
Es genügt, die für $n = 2$ zu zeigen. Sei also $p \mid ab$. Betrachte das Ideal (a, p). Ist $(a, p) = (1)$, so $\exists\, r, s \in A$ mit $ra + sp = 1$.
Multipliziere dies mit b: $r(ab) + spb = b$. Wegen $p \mid ab \Rightarrow p \mid b$.
Sei jetzt $(a, p) \neq (1)$. Nach Voraussetzung (A ist Hauptidealring) $\exists\, d \in A$ mit $(a, p) = (d)$. Dabei ist $d \notin A^*$. Habe $d \mid p \Rightarrow$ (da d keine Einheit und p unzerlegbar)

$d \sim p$.
$\Rightarrow p \sim d \mid a \Rightarrow p \mid a$.

□

Satz 7.27.
Sei A ein Hauptidealring. Dann hat jede Nichteinheit $0 \neq a \in A$ eine Darstellung $a = p_1 \cdot \ldots \cdot p_n$ mit $n \in \mathbb{N}$, $p_i \in A$ unzerlegbar; dabei sind n und $p_1, \ldots, p_n$ bis auf $\sim$ und Reihenfolge eindeutig bestimmt.

Beweis:
Zeige zunächst: in A gibt es keine unendlichen echten Teilerketten. Sei $\ldots \mid a_3 \mid a_2 \mid a_1$ mit $0 \neq a_n \in A$. Betrachte das Ideal $I = \bigcup_{n \in \mathbb{N}} (a_n)$ von A (dies ist tatsächlich ein Ideal wegen $(a_1) \subset (a_2) \subset \ldots$). Nach Voraussetzung $\exists\, b \in I$ mit $I = (b)$. Also einerseits $b \mid a_n \;\forall\; n$. Andererseits $\exists\, m_0 \in \mathbb{N}$ mit $b \in (a_m) \;\forall\; m \geq m_0$, also $a_m \mid b \;\forall\; m \geq m_0$, also $a_m \sim b \;\forall\; m \geq m_0$.
Also wird die Teilerkette schließlich konstant modulo m_0.

Nach 7.23 folgt: a hat eine Zerlegung wie behauptet. Zur Eindeutigkeit: sei $a = p_1 \cdot \ldots \cdot p_n = q_1 \cdot \ldots \cdot q_n$ mit p_i, q_i unzerlegbar, $m, n \in \mathbb{N}$.
Wegen $p_1 \mid a = q_1 \cdot \ldots \cdot q_m$ folgt aus Lemma 7.26: $p_1 \mid q_j$ für ein $j \in \{1, \ldots, n\}$ Wegen $p_1 \notin A^*$, q_j unzerlegbar also $p_1 \sim q_j$. Kürze durch p_1 und fahre so induktiv fort. es folgt: $\exists\, \varphi : \{1, \ldots, n\} \to \{1, \ldots, m\}$ injektiv mit $p_i \sim q_{\varphi(i)} \;\forall\; i$ Vertausche jetzt die Rollen der p_i und $q_j \Rightarrow$ Behauptung.

□

Bemerkung 7.28.
Sei Π eine Menge von unzerlegbaren Elementen im Hauptidealring A derart, dass jedes unzerlegbare $p \in A$ zu genau einem $\pi \in \Pi$ assoziiert ist (nenne solches Π ein **Vertretersystem** für die unzerlegbaren Elemente in A).
Dann hat jedes $0 \neq a \in A$ eine *eindeutige* Produktdarstellung
$a = \varepsilon \cdot \prod_{\pi \in \Pi} \pi^{v_p(a)}$ mit $\varepsilon \in A^*$, $v_p(a) \in \mathbb{N} \cup \{0\}$, $v_p(a) = 0$ für fast alle $\pi \in \Pi$.

Korollar 7.29.
Sei A ein Hauptidealring. Für jedes $0 \neq p \in A$ sind äquivalent:

(i) *p ist unzerlegbar;*

(ii) *(p) ist ein maximales Ideal von A;*

(iii) *der Ring $A \setminus (p)$ ist ein Körper.*

Beweis:
(ii) $\Leftrightarrow$ (iii):
7.14

(i) $\Rightarrow$ (ii):
$p \notin A^*$, also $(p) \neq (1)$. Wäre (p) nicht maximal, so gäbe es ein $(p) \subsetneq (q) \subsetneq (1)$, also $q \mid p, q \not\sim 1 \Rightarrow$ Widerspruch zu p unzerlegbar.

(ii) $\Rightarrow$ (i):
analog.

□

7.30.
Zeigen nun (u.a.): $\mathbb{Z}, \mathbb{K}[t]$ sind Hauptidealringe.
Verantwortlich dafür: Existenz einer **Division mit Rest**.

Definition 7.31.
Sei A nullteilerfrei. Eine euklidische Wertefunktion auf A ist eine Abbildung $\Phi : A \setminus \{0\} \to \mathbb{N} \cup \{0\}$ mit folgender Eigenschaft („Division mit Rest"):

$$\forall\, a, b \in A \setminus \{0\}\ \exists\, q, r \in A \text{ mit } a = qb + r \text{ und } r = 0 \text{ oder } \Phi(r) < \Phi(b).$$

Existiert ein solches Φ, so heißt A ein **euklidischer Ring**.

Beispiele 7.32.

1. $A = \mathbb{Z}$ mit $\Phi(a) = |a|$ $(0 \neq a \in \mathbb{Z})$.
   Dabei ist in $a = qb + r$ mit $|r| < |b|$ die Darstellung (q, r) nicht eindeutig: $5 = 1 \cdot 3 + 2 = 2 \cdot 3 - 1$.
2. $A = \mathbb{K}[t]$ mit $\Phi(f) = \deg(f)$ $(0 \neq f \in \mathbb{K}[t])$.
   $f = qg + r$ mit $\deg(r) < \deg(g)$ (**??**)
   Hier sind q und r eindeutig.

Satz 7.33.
Jeder euklidische Ring ist ein Hauptidealring.

Beweis:
Sei $I \neq (0)$ ein Ideal in A. Wähle ein $0 \neq a \in I$ mit minimalem $\Phi(a)$, behaupte $I = (a)$:

sei dazu $x \in I$, dividiere x durch a mit Rest:
$x = qa + r$ mit $\Phi(r) < \Phi(a) \vee r = 0$.
Wegen $r \in I$ muss $r = 0$ sein, da sonst Widerspruch zur Wahl von a.

□

Korollar 7.34.
$\mathbb{Z}$ und $\mathbb{K}[t]$ sind Hauptidealringe.

Satz 7.27 gibt uns für $A = \mathbb{Z}$ den **Fundamentalsatz der elementaren Zahlentheorie**:
Jede natürliche Zahl ist Produkt von Primzahlen und die Darstellung ist eindeutig bis auf Reihenfolge der Faktoren.

Korollar 7.35.
Sei $\mathbb{K}$ ein Körper. Jedes $0 \neq f \in \mathbb{K}[t]$ hat eine Produktdarstellung $f = c \cdot p_1 \cdot \ldots \cdot p_m$ mit $0 \neq c \in \mathbb{K}$ und irreduziblen normierten $p_1, \ldots, p_m \in \mathbb{K}[t]$, und diese ist eindeutig bis auf Reihenfolge der p_i.

7.36.
kommt später ... (Frage wie die Zerlegung aussieht ist ziemlich wichtig, aber sehr kompliziert)

Definition 7.37.
Sei A nullteilerfrei, seien $0 \neq a, b \in A$.

(a) Ein $d \in A$ heißt ein **größter gemeinsamer Teiler** von a und b, in Zeichen $d \sim \mathrm{ggT}(a, b)$, wenn $d \mid a$, $d \mid b$ und für alle $d' \in A$ mit $d' \mid a$, $d' \mid b$ gilt: $d' \mid d$.

(b) Ein $e \in A$ heißt **kleinstes gemeinsames Vielfaches** von a, b, in Zeichen $e \sim \mathrm{kgV}(a, b)$ falls $a \mid e$, $b \mid e$ und $\forall\, e' \in A$ mit $a \mid e'$, $b \mid e'$ gilt $e \mid e'$.

Bemerkung 7.38. label7-2-24

1. Falls $\mathrm{ggT}(a, b)$ oder / und $\mathrm{kgV}(a, b)$ existieren, so sind sie eindeutig bis auf $\sim$. Daher die Schreibweise $d \sim \mathrm{ggT}(a, b)$.
2. Im Allgemeinen brauchen $\mathrm{ggT}(a, b)$ oder $\mathrm{kgV}(a, b)$ *nicht* zu existieren.
3. Ist A ein Hauptidealring, so existieren ggT und kgV stets: zu $a, b \in A$ ist $(a, b) = (d)$ mit einem $d \in A$, und für dieses d gilt $d \sim \mathrm{ggT}(a, b)$.
   (denn $d \mid a, d \mid b$; aus $d' \mid a, d' \mid b$ folgt $(a) \subset (d'), (b) \subset (d') \Rightarrow (d) = (a, b) \subset (d') \Rightarrow d' \mid d$)
   Analog: $(a) \cap (b) = (e)$ mit einem $e \in A$, und für dieses e ist $e \sim \mathrm{kgV}(a, b)$.

Satz 7.39.
Sei A ein Hauptidealring, sei Π ein Vertretersystem für die unzerlegbaren Elemente von A.

(a) $a \mid b \Rightarrow \ \forall\ \pi \in \Pi\ v_\pi(a) \leq v_\pi(b)$.

(b) $ggT(a,b) \sim \prod_{\pi\in\Pi} \pi^{\min\{v_\pi(a),v_\pi(b)\}}$
$kgV(a,b) \sim \prod_{\pi\in\Pi} \pi^{\max\{v_\pi(a),v_\pi(b)\}}$

Beweis:

(a)

„$\Leftarrow$“ : klar.

„$\Rightarrow$“ : Sei $\pi \in \Pi$ mit $\pi \mid a \Rightarrow \pi \mid a \mid b$. Setze $a' := \frac{a}{\pi}, b' := \frac{b}{\pi}$, $a' \mid b'$: mache induktiv mit a', b' weiter.

(b) folgt sofort aus (a).

□

Bemerkung 7.40.
$\mathrm{ggT}(a,b) \cdot \mathrm{kgV}(a,b) \sim ab$.

7.41.
Wir zeigen jetzt, dass iterierte Division mit Rest den ggT (und damit auch das kgV) berechnen, ohne dass man die Elemente in Faktoren zerlegen muss (**Euklidischer Algorithmus**):
Sei A ein euklidischer Ring mit Wertefunktion $\Phi : A \setminus \{0\} \to \mathbb{N} \cup \{0\}$. Seien $0 \neq a, b \in A$.

$$\begin{array}{rcl}
a = q_0 b + r_1 & \text{mit} & \Phi(r_1) < \Phi(b), r_1 \neq 0 \\
b = q_1 r_1 + r_2 & \text{mit} & \Phi(r_2) < \Phi(r_1), r_2 \neq 0 \\
\vdots & \vdots & \vdots \\
r_{m-2} = q_{m-1} r_{m-1} + r_m & \text{mit} & \Phi(r_1) < \Phi(r_m), r_{m-1} \neq 0 \\
r_{m-1} = q_m r_m &&
\end{array}$$

Satz 7.42.
$(r_m) = (a,b)$. *Insbesondere* $r_m \sim ggT(a,b)$.

Beweis:
$\subset$:
von oben beginnend: $r_1, r_2, \ldots, r_m \in (a, b)$.
$\supset$:
von unten beginnend: $r_{m-1}, r_{m-2}, \ldots, b, a \in (r_m)$.

□

Bemerkungen und Beispiele 7.43.

(a) Der Algorithmus gibt für $a, b \in A$ auch $r, s \in A$ mit $ra + sb \sim \mathrm{ggT}(a, b)$:
$r_m = r_{m-1} - q_{m-1} r_{m-1}$
$r_{m-1} = r_{m-2} - q_{m-2} r_{m-2}$

(b) $A = \mathbb{Z}, a = 54, b = 35$:
$54 = 2 \cdot 35 - 16$
$35 = 2 \cdot 16 + 3$
$16 = 5 \cdot 3 + 1 \Rightarrow \mathrm{ggT}(35, 56) \sim 1.$

$1 = 16 - 5 \cdot 3 = 16 - 5 \cdot (35 - 2 \cdot 16) = -5 \cdot 35 + 11 \cdot 16 = -5\dot{3}5 + 11 \cdot (2 \cdot 25 - 54) =$
$17 \cdot 35 - 11 \cdot 54.$

(c) Berechnung von $\mathrm{ggT}(f, g)$ in $\mathbb{K}[t]$, $f := t^3 + 2t^2 - t - 1$, $g := t^2 + t - 3$:
$f = t \cdot g + r_1$; $r = t^2 + 2t - 1$
$g = r_1 + r_2$, $r_2 = -t - 2$
$r_1 = -tr_2 - 1$. Also $\mathrm{ggT}(f, g) \sim 1$.
Umgekehrt finde $1 = -r_1 - tr_2 = \ldots = (t - 1) \cdot f - t^2 \cdot g$.

(d) Andere Anwendung: Lösung von simultanen Kongruenzen. Hier das allgemeine Problem:
Sei A ein (beliebiger kommutativer) Ring, seien $I_1, \ldots, I_n \subset A$ Ideale, sowie $a_1, \ldots, a_n \in A$. Wann hat das System

$$\begin{array}{ll} x \equiv a_1 & \mod I_1 \\ \vdots & \vdots \\ x \equiv a_n & \mod I_n \end{array} \qquad (*)$$

von simultanen Kongruenzen eine Lösung $x \in A$?

Anders gesagt: betrachte den Ringhomomorphismus
$\varphi : A \to (A/I_1) \times \ldots \times (A/I_n)$
$\varphi(a) := (a + I_1, \ldots, a + I_n)$ $(a \in A)$
Die Frage ist gerade: was ist das Bild von φ?

Theorem 7.44. *(Chinesischer Restsatz)*
Sind $I_1, \ldots, I_n$ paarweise teilerfremd (d.h., $I_i + I_j = (1)$), so ist φ surjektiv. Für jede Wahl

von $a_1, \ldots, a_n \in A$ hat also (*) eine Lösung in A.

Beweis:
($n = 1 \to \checkmark$). Sei zunächst $n = 2$. Nach Voraussetzung habe $1 = b_1 + b_2$ mit $b_1 \in I_1, b_2 \in I_2$. System $\left| \begin{array}{l} x \equiv a_1 \, (I_1) \\ x \equiv a_2 \, (I_2) \end{array} \right|$: Dann löst $x := a_1 b_2 + a_2 b_1$ das System:
$x \equiv a_1 b_2 \equiv a_1 \, (I_1)$
$x \equiv a_2 b_1 \equiv a_2 \, (I_2)$

Für $n > 2$ zunächst ... □

Hilfssatz 7.1.
Sind $I, J_1, \ldots, J_r$ Ideale mit $I + J_1 = (1)$ für $i = 1, \ldots, r$, so ist $I + (J_1 \cdots J_r) = (1)$. Insbesondere $I + (J_1 \cap \ldots \cap J_r) = (1)$.

Beweis:
Habe $a_i + b_i = 1$ mit $a_i \in I$ und $b_i \in J_i$ $(i = 1, \ldots, r)$. Multipliziere diese:
$1 = (a_1 + b_1) \cdots (a_r + b_r) = \underbrace{b_1 \cdots b_r}_{\in J_1 \cdots J_r} + \underbrace{(\text{übrige Summanden})}_{\in I}$ □

Beweis: (Ende des Beweises von 7.44)
Sei $n \geq 3$. Seien $a_1, \ldots, a_n \in A$.
Hilfssatz $\Rightarrow I_1 + (I_2 \cap \ldots \cap I_n) = (1)$. Nach Induktion gibt es $y \in A$ mit $y \equiv a_i (I_i)$ für $i = 2, \ldots, n$.
Fall $n = 2$ angewandt auf I_1 und $I_2 \cap \ldots \cap I_n$ gibt ein $x \in A$ mit $x \equiv a_1 (I_1)$ und $x \equiv y (I_2 \cap \ldots \cap I_n)$. Also auch $x \equiv y \equiv a_i (I_i)$ für $i = 2, \ldots, n$.
□

Korollar 7.45.
(A beliebig) Seien $I_1, \ldots, I_n$ Ideale in A mit $I_i + I_j = (1)$ für alle $i \neq j$, dann hat man einen Isomorphismus von Ringen
$A/(I_1 \cap \ldots \cap I_n) \to (A/I_1) \times \ldots \times (A/I_n)$
$a + \bigcap I_i \mapsto (a + I_1, \ldots, a + I_n)$.

Beweis:
Für φ wie oben ist $\ker(\varphi) = I_1 \cap \ldots \cap I_n$.
Behauptung folgt aus dem Homomorphiesatz für Ringe (7.6).
□

7.46.
wurde verloren ...

7.47.
Zeige jetzt, wie der Euklidische Algorithmus ein konstruktives Verfahren zur Lösung von simultanen Kongruenzen liefert:
Sei A ein euklidischer Ring, seien $I_i = (b_i)$ $(i = 1, \ldots, n)$ und seien $\mathrm{ggT}(b_i, b_j) \sim 1$ für $i \neq 1$. Wir wollen die simultanen Kongruenzen $x \equiv a_i(b_i), i = 1, \ldots, n$ (*) lösen.

Setze dazu $c_i := \frac{b_1 \cdots b_n}{b_i} = b_1 \cdot \ldots \cdot \widehat{b_i} \cdot \ldots \cdot b_n$ $(i = 1, \ldots, n)$. Dann ist $(b_i, c_i) = (1)$, also gibt es $r_i, s_i \in A$ mit $r_i b_i + s_i c_i = 1$ $(i = 1, \ldots, n)$.
Wegen $s_i c_i \equiv 1(b_i)$ und $s_i c_i \equiv 0(b_j)$ für $j \neq i$.
Also ist $x := \sum_{i=1}^{n} a_i s_i c_i$ eine Lösung von (*).

Die volle Lösungsmenge von (*) ist daher $x + (b_1 \cdot \ldots \cdot b_n)$.

Beispiel 7.48.
Betrachten wir das System
$x \equiv a_1 \mod 54$
$x \equiv a_2 \mod 35$
in $\mathbb{Z}$.
Haben gesehen: $1 = 17 \cdot 35 - 11 \cdot 54 = 595 - 594$.
Somit ist $x := 17 \cdot 35 a_1 - 11 \cdot 54 a_2 = 595 a_1 - 594 a_2$ eine Lösung von (*). Die Lösungsmenge von (*) ist also $\{595 a_1 - 594 a_2 + k \cdot 1890 : k \in \mathbb{Z}\}$.

c. Allgemeines über Moduln

Sei A stets ein kommutativer Ring.

Definition 7.49.
Ein **A-Modul** ist eine abelsche Gruppe $(M, +)$ zusammen mit einer Skalarmultiplikation
$A \times M \to M, (a, x) \mapsto a \cdot x (= ax)$.
so dass gelten:

1. $\forall\, a, b \in A,\ \forall\, x \in M$: $(ab)x = a(bx)$;
2. $\forall\, x \in M$: $1x = x$;
3. $\forall\, a, b \in A,\ \forall\, x \in M$: $(a + b)x = ax + bx$;
4. $\forall\, a \in A,\ \forall\, x, y \in M$: $a(x + y) = ax + ay$.

Beispiele 7.50.

1. Aus 3. mit $b = 0$ folgt $0x = 0$. Aus 3. und 4. mit $b = -a$ und $y = -x$ folgt $(-a)x = -(ax) = a(-x)$.
2. Ist $A = \mathbb{K}$ ein Körper, so sind die A-Moduln dasselbe wie $\mathbb{K}$-Vektorräume.
3. Jede abelsche Gruppe $(M, +)$ besitzt genau eine $\mathbb{Z}$-Modul Struktur. Denn für $n > 0$:
   $nx = (\underbrace{1 + \ldots + 1}_{n-mal})x = \underbrace{1x + \ldots + 1x}_{n-mal} = \underbrace{x + \ldots + x}_{n-mal}$
   $(-n)x = -(nx) = -(\underbrace{x + \ldots + x}_{n-mal})$
   Also sind $\mathbb{Z}$-Moduln genau dasselbe wie abelsche Gruppen.
4. Sei $\mathbb{K}$ ein Körper, $A := \mathbb{K}[t]$ Ist M ein A-Modul, so ist M ein Vektorraum über $\mathbb{K}$.
   Außerdem ist die Abbildung $M \to M, x \mapsto t \cdot x$ $\mathbb{K}$-linear:
   $f(x + y) = t \cdot (x + y) = tx + ty = f(x) + f(y)$
   $f(cx) = t(cx) = (tc)x = (ct)x = c(tx) = c\, f(x)$.
   Also $f \in End_{\mathbb{K}}(M)$.
   Umgekehrt: ist V ein $\mathbb{K}$-Vektorraum und $f \in End_{\mathbb{K}}(V)$, kann man eine Skalarmultiplikation definieren:

$\mathbb{K}[t] \times V \to V, \left(\sum_{i=1}^{n} c_i t^i, x\right) \mapsto \sum_{i=1}^{n} c_i f^i(x)$.
Mit dieser Skalarmultiplikation ist V ein $\mathbb{K}[t]$-Modul (V_f). Also ein $\mathbb{K}[t]$-Modul ist genau dasselbe wie ein Paar (V, f) aus einem $\mathbb{K}$-Vektorraum V, und $f \in End_{\mathbb{K}}(V)$.

Definition 7.51.
Eine Abbildung $f : M \to N$ zwischen A-Moduln heißt **linear**, falls:

(i) $f(x + y) = f(x) + f(y)$;

(ii) $f(ax) = a\, f(x)$.

f heißt ein **Isomorphismus**, falls f bijektiv ist. M, N heißen **isomorph**, falls ein Isomorphismus $f : M \to N$ existiert.

Beispiel 7.52.
Seien V_f, W_g $\mathbb{K}[t]$-Moduln. Ist $\varphi : V_f \to W_g$ $\mathbb{K}[t]$-linear, so gelten:

(i) $\varphi : V \to W$ ist $\mathbb{K}$-linear;

(ii) $\varphi(tx) = t\, \varphi(x)$.

Umgekehrt: gelten (i) und (ii), dann folgt:
$$\varphi\left(\left(\sum_{i=1}^{n} c_i t^i\right) \cdot x\right) = \varphi\left(\sum_{i=1}^{n} c_i \left(t^i x\right)\right)$$
$$= \sum_{i=1}^{n} c_i \varphi(t^i x) = \sum_{i=1}^{n} c_i t^i \varphi(x) = \left(\sum_{i=1}^{n} c_i t^i\right) \varphi(x).$$

Man sieht, dass (ii) bedeutet, dass $\forall\, x \in V$: $\varphi(f(x)) = g(\varphi(x))$.

Lemma 7.53.
Eine Abbildung $\varphi : V_f \to W_g$ zwischen $\mathbb{K}[t]$-Moduln ist $\mathbb{K}[t]$-linear genau dann, wenn sie $\mathbb{K}$-linear ist und $\varphi \circ f = g \circ \varphi$ gilt.

Definition 7.54.
Sei M ein A-Modul. Eine Teilmenge $U \subset M$ heißt ein **(A-)Untermodul**, falls U eine Untergruppe von $(M, +)$ ist und gilt $\forall\, a \in A,\ \forall\, x \in U$: $ax \in U$.

Beispiele 7.55.

1. Eine Teilmenge $U \subset M$ ist genau dann ein Untermodul, wenn

   (i) $0 \in U$;

   (ii) $U + U \subset U$;

   (iii) $AU \in U$.

   Insbesondere sind $\{0\}, M \subset M$ Untermoduln.
   Außerdem ist $Ax := \{ax | a \in A\} \subset M$ ein Untermodul.

2. Jeder Untermodul $U \subset M$ ist ein A-Modul mit der induzierten Skalarmultiplikation
   $$\begin{array}{ccc} A \times U & \to & U \\ \downarrow & & \downarrow \\ A \times M & \to & M \end{array}$$

3. Die Untermoduln von A sind genau die Ideale von A.

4. Ist V_f ein $\mathbb{K}[t]$-Modul, so ist $U \subset V_f$ genau dann ein Untermodul, wenn U ein f-invarianter $\mathbb{K}$-Untervektorraum ist ($f(U) \subseteq U$).

5. Ist M ein $\mathbb{Z}$-Modul, so sind die $\mathbb{Z}$-Untermoduln von M genau die Untergruppen von M.

6. Sei $f : M \to N$ eine lineare Abbildung zwischen A-Moduln. Ist $U \subseteq M$ ein Untermodul, so ist $f(U) = \{f(u) \mid u \in U\} \subseteq N$ ein Untermodul.
   Ist $V \subset N$ ein Untermodul, so ist $f^{-1}(V) = \{x \in M \mid f(x) \in V\} \subset M$ ein Untermodul. Insbesondere

   $$\ker(f) := f^{-1}(\{0\}) = \{x \in M \mid f(x) = 0\}$$

   ist ein Untermodul.

7.56.
Sei M ein A-Modul, $U \subset M$ ein Untermodul. Die Relation

$$x \equiv y \mod U :\Leftrightarrow x - y \in U$$

ist eine Äquivalenzrelation. Wir schreiben $\overline{x}$ für die Äquivalenzklasse von $x \in M$.
Die Menge $M/U := \{\overline{x} \mid x \in M\}$ ist ein A-Modul mit $\overline{x} + \overline{y} := \overline{x+y}$ und $a\overline{x} := \overline{ax}$.
M/U heißt der Quotientmodul.
Die Abbildung $\pi : M \to M/U, x \mapsto \overline{x}$ ist A-linear und surjektiv.

Satz 7.57.
Jede lineare Abbildung $f : M \to N$ zwischen A-Moduln induziert einen Isomorphismus $M/\ker(f) \to im(f), \overline{x} \mapsto f(x)$.

Beweis: Analog zum Isomorphiesatz für Vektorräume.

□

7.58.
Sei M ein A-Modul, $(U_i)_{i\in I}$ eine Familie von Untermoduln: $U_i \subset M$.
Die Menge

$$\sum_{i\in I} U_i := \{\sum_{i\in I} u_i \mid u_i \in U_i, u_i = 0 \text{ f.f.a. } i \in I\} \subset M$$

ist ein Untermodul. $\sum_{i\in I} U_i$ heißt die Summe von $(U_i)_{i\in I}$.
Allgemeiner: ist $X \subset M$ eine beliebige Teilmenge, definiert man $\text{span}_A(X) := \bigcap\{U \subset M \mid U \text{ ist Untermodul}, X \subset U\}$.
span (X) ist der kleinste Untermodul von M, der X enthält.

Es ist span $(X) = \sum_{x\in X} Ax = \{\sum_{x\in X} a_x x \mid a_x \in A, a_x = 0 \text{ f.f.a. } x \in X\}$
und $\sum U_i = \text{span}\left(\bigcup_{i\in I} U_i\right)$.

7.59.
Sei $(M_i)_{i\in I}$ eine Familie von A-Moduln. Das Produkt $\prod_{i\in I} M_i$ ist ein A-Modul:

$$(x+y)_i = x_i + y_i$$
$$(ax)_i := ax_i$$

Die Teilmenge $\bigoplus_{i\in I} M_i := \{x \in \prod_{i\in I} M_i \mid x_i = 0 \text{ f.f.a. } i \in I\}$ ist ein Untermodul. $\oplus M_i$ heißt die (äußere) direkte Summe von $(M_i)_{i\in I}$.
Wenn $I = \{1, \ldots, n\}$ endlich, so ist $\prod_{i=1}^{n} M_i = \bigoplus_{i=1}^{n} M_i (= M_1 \oplus \ldots \oplus M_n)$.

7.60.
Sei M ein A-Modul, $(U_i)_{i\in I}$ eine Familie von Untermoduln von M. Die Abbildung
$f : \bigoplus_{i\in I} U_i \to M, (v_i)_{i\in I} \mapsto \sum_{i\in I} u_i$
ist A-linear. f ist genau dann bijektiv, wenn jedes $x \in M$ eine einzige Darstellung $x = \sum_{i\in I} u_i$ mit $u_i = 0$ *f.f.a.* $i \in I$. In diesem Fall ist M die innere direkte Summe von $(U_i)_{i\in I}$.

Definition 7.61.
Sei M ein A-Modul. Eine Familie $(x_i)_{i\in I}$ von Elementen $x_i \in M$ heißt

- **erzeugend**, falls jedes $x \in M$ eine Darstellung $x = \sum_{i\in I} a_i x_i$ besitzt mit $a_i = 0$ *f.f.a.* $i \in I$. ($\equiv M = \sum_{i\in I} Ax_i$, f in 7.60 ist surjektiv).

- **linear abhängig**, falls eine Gleichung $\sum_{i\in I} a_i x_i = 0$ existiert, mit $a_i = 0$ *f.f.a.* $i \in I$ mit $a_i \neq 0$ für mindestens ein $i \in I$.
  Andernfalls: **linear unabhängig**.

- eine **Basis** von M, wenn sie linear unabhängig und erzeugend ist.
  Der A-Modul M heißt **frei**, wenn er eine Basis hat.

Bemerkung 7.62.

1. Jeder Modul M hat ein Erzeugendensystem (z.B. die Familie aller Elemente von M). Gibt es ein endliches Erzeugendensystem von M, so heißt M **endlich erzeugt**.

2. Aber im Allgemeinen braucht es in M keine nicht-leere linear unabhängige Familie zu geben!
   Bsp: $A = \mathbb{Z}, M = \mathbb{Z}/(n)$ mit $n \in \mathbb{N}, n > 1$. Dann ist $nx = 0$ für alle $x \in M$.

3. Hier sehen wir auch: ein minimales Erzeugendensystem von M muss nicht linear unabhängig sein; eine maximal linear unabhängige Familie muss kein Erzeugendensystem sein.
   M hat im Allgemeinen keine Basis, ist also im Allgemeinen nicht frei.

4. Ist M frei und ist $(x_\lambda)_{\lambda\in\Lambda}$ eine Basis von M, so ist die A-lineare Abbildung

$$\bigoplus_{\lambda\in\Lambda} A \to M, (a_\lambda) \mapsto \sum_{\lambda\in\Lambda} a_\lambda x_\lambda$$

bijektiv, also ein Isomorphismus von A-Moduln.

Damit gilt insbesondere:

Satz 7.63.
Ein endlich erzeugter A-Modul M ist genau dann frei, wenn es ein $n \in \mathbb{N} \cup \{0\}$ gibt mit $M \cong A^n$ (als A-Moduln).

Definition und Satz 7.64.
Sei A ein Hauptidealring. Sei M ein freier A-Modul. Ist M endlich erzeugt, so ist die Zahl $n \geq 0$ mit $M \cong A^n$ eindeutig durch M bestimmt und heißt der Rang von M, in Zeichen $rk\,(M)$. Ist M nicht endlich erzeugt, so setzt man $rk\,(M) := \infty$.

Beweis:
Sei $M \cong A^n$. Sei $\pi \in A$ unzerlegbar, also ist $k := A/(\pi)$ ein Körper (7.29), und $M/\pi M$ ($\pi M := \{\pi x : x \in M\}$) ist ein Vektorraum über $k = A/(\pi)$, via
$(a + (\pi)) \cdot (x + \pi M) := ax + \pi M$,
und $M/\pi M \cong (A/(\pi))^n = k^n$. Da die k-Dimension des k-Vektorraums $M/\pi M$ eindeutig bestimmt ist, folgt: $n = \dim_k(M/\pi M)$ ist eindeutig durch M bestimmt.
□

Zusatz:
Der Satz gilt auch für alle Ringe A (außer dem Nullring). Beweis wie oben mit maximalem Ideal.

Definition 7.65.
Ein A-Modul heißt **zyklisch**, wenn er von einem Element erzeugt werden kann.

7.66.
Wie sehen zyklische Moduln aus? Sei $M = Ax$ ein zyklischer A-Modul (mit $x \in M$). Betrachte die lineare Abbildung $\varphi_x : A \to M, \varphi_x(a) := ax$.
Sie ist surjektiv, also ist $M \cong A/\ker(\varphi_x)$ (7.57), und dabei ist $\ker(\varphi_x) = \{a \in A : ax = 0\}$ ein A-Untermodul von A, also ein Ideal von A. Somit folgt:

Satz 7.67.
Ein A-Modul M ist genau dann zyklisch, wenn $M \cong A/I$ mit einem Ideal I von A ist. Dabei ist I durch M eindeutig bestimmt.

Beweis:

„⇒“ : ✓

„⇐“ : $M = A/I \Rightarrow M$ ist erzeugt von $1 + I$.

Eindeutigkeit:
Für $M = A/I$ ist $I := \{a \in A : ax = 0 \ \forall \ x \in M\}$ also durch M bestimmt.

□

Lemma 7.68.
Sei M ein zyklischer A-Modul.

(a) Jeder Faktormodul M/U ist wieder zyklisch.

(b) Ist A ein Hauptidealring, so ist auch jeder Untermodul von M zyklisch.

Beweis:

(a) ✓

(b) O.E. $M = A/I$ (I Ideal von A).
Die Untermoduln von $M = A/I$ sind die J/I mit $I \subset J \subset A$ Ideal. Nach Voraussetzung von $J = (b)$ mit $b \in A \Rightarrow J/I$ wird von $b + I$ erzeugt.

□

Satz 7.69.
Sei A ein Hauptidealring, sei M ein endliche erzeugter A-Modul und U ein Untermodul von M.

(a) U ist endlich erzeugt.

(b) Es gibt eine endliche Folge von $\{0\} = U_0 \subset U_1 \subset \ldots \subset U_n = U$ von Untermoduln von U derart, dass alle U_i/U_{i-1} ($i = 1, \ldots, n$) zyklisch sind.

Beweis:
Zunächst (b):
Sei $M = Ax + \ldots + Ax_n$, setze $M_i := Ax_1 + \ldots Ax_i$ ($i = 1, \ldots, n$), $M_0 := \{0\}$ und

$U_i := U \cap M_i$.
Dann ist M_i/M_{i-1} zyklisch, erzeugt von $\overline{x_i} = x_i + M_{i-1}$ $(i = 1, \ldots, n)$.
Der Homomorphismus $U_i \to M_i/M_{i-1}, u \mapsto u + M_{i+1}$ hat den Kern $U_i \cap M_{i-1} = U \cap M_i \cap M_{i-1} = U_{i-1}$.
Somit ist U_i/U_{i-1} isomorph zu einem Untermodul von M_i/M_{i-1}, also selbst zyklisch nach Lemma 7.68.

Aus (b) folgt direkt (a):
Wähle $u_i \in U_i$, so dass U_i/U_{i-1} von $\overline{u_i} := u_i + U_{i-1}$ erzeugt wird. Dann wird U von $u_1, \ldots, u_n$ erzeugt.

□

d. Moduln über Hauptidealringen

7.70.
Sei A ein kommutativer Ring. Wir hatten (in Kapitel 3) $M_{m\times n}(A)$: der (freie!) A-Modul der $m \times n$-Matrizen über A.
$M_n(A) := M_{n\times n}(A)$, ein Ring.
Determinante $\det : M_n(A) \to A$, definiert durch $\det(a_{ij}) := \sum_{\sigma\in S_n} sign(\sigma) \cdot \prod_{j=1}^{n} a_{\sigma(j),j}$.

Es gilt weiterhin $\det(ST) = \det(S) \cdot \det(T)$ ($S, T \in M_n(A)$). Wir haben zu $S \in M_n(A)$ die adjungierte Matrix $S^\sharp$, definiert wie in Kapitel 3.d, und es gilt $S \cdot S^\sharp = S^\sharp \cdot S = \det(S) \cdot I$.
Die Matrix $S \in M_n(A)$ heißt invertierbar, falls $\exists\, T \in M_n(A)$ mit $ST = TS = I$. Dann ist T eindeutig bestimmt, schreibe $S^{-1} := T$.

Lemma 7.71.
$S \in M_n(A)$ ist invertierbar $\Leftrightarrow \det(S) \in A^*$. Dann ist $S^{-1} = \det(S)^{-1} \cdot S^\sharp$.

Beweis:
$ST = I \Rightarrow \det(S) \cdot \det(T) = \det(I) = 1$.
Umkehrung folgt aus $S \cdot S^\sharp = \det(S) \cdot I$.

□

7.72.
Seien M, N freie A-Moduln mit Basen $\mathcal{B} = (x_1, \ldots, x_m)$ von M bzw. $\mathcal{C} = (y_1, \ldots, y_n)$ von N. Lineare Abbildungen $f : N \to M$ werden wie gewohnt durch $m \times n$-Matrizen (über A) beschrieben. Die Matrix $M_{\mathcal{B}}^{\mathcal{C}}(f) = (t_{ij}) \in M_{m\times n}(A)$ ist gegeben durch $f(y_j) = \sum_{i=1}^{m} t_{ij}x_i$ ($j = 1, \ldots, n$).
Insbesondere wird der Untermodul $im(f)$ von M von den aus den Spalten von $M_{\mathcal{B}}^{\mathcal{C}}(f)$ gebildeten Linearkombinationen der x_i erzeugt.
Komposition von linearen Abbildungen entspricht der Matrix-Multiplikation.

7.73.
Über Körpern hatten wir den „Rangsatz“:
$f : V \to W$: kann Basen so wählen, dass $f \sim \left(\begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array}\right)$.
Ziel: Analogon über Hauptidealringen?

Sei also A ein Hauptidealring, sei $T = (t_{ij}) \in M_{m\times n}(A)$ gegeben, entsprechend einer linearen Abbildung $T : A^n \to A^m$, $y \mapsto Ty$. Betrachte die folgenden Operationen auf der Matrix T:

(Z_1): Ersetzen von Z_i durch $Z_i + aZ_j$ ($i \neq j, a \in A$)
(Z_2): Vertauschen von Z_i und Z_j
(S_1): Ersetzen von $_k$ durch $S_k + aS_l$ ($k \neq l, a \in A$)
(S_2): Vertauschen von S_k und S_l

(Z_i := i-te Zeile, S_k := k-te Spalte von A).

Wie in Kapitel 3 entsprechen diese jeweils der Multiplikation der Matrix von links bzw. rechts mit bestimmten „elementaren“ Matrizen ($Q_{ij}(a), \Pi_{ij}$).

Theorem 7.74.
Sei A ein euklidischer Ring, sei $T \in M_{m\times n}(A)$ gegeben. Dann kann man T durch endlich Operationen (S_1), (S_2),(Z_1), (Z_2) in eine Matrix der Gestalt

$$\left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right) \qquad (*)$$

transformieren mit $r \in \mathbb{N} \cup \{0\}$ und mit $d_1 \mid d_2 \mid \ldots \mid d_r \neq 0$.

Beweis:
Sei $T = (a_{ij})$, O.E. $T \neq 0$. Wir setzen $\mathrm{ggT}(T) := \mathrm{ggT}(a_{ij} : 1 \leq i \leq m, 1 \leq j \leq n)$ (d.h. $\mathrm{ggT}(T)$ ist ein Erzeuger des Ideals $(a_{ij} : 1 \leq i \leq m, 1 \leq j \leq n)$).
Für alle $U \in GL_m(A)$, $V \in GL_n(A)$ ist $\mathrm{ggT}(UTV) \sim \mathrm{ggT}(T)$. Denn: $\mathrm{ggT}(T) \mid \mathrm{ggT}(UTV)$.
Umgekehrt: Wegen $T = U^{-1}(UTV)V^{-1}$ gilt auch $\mathrm{ggT}(UTV) \mid \mathrm{ggT}(T)$.
Insbesondere bleibt $\mathrm{ggT}(T)$ unter den elementaren Transformationen Z_1, Z_2, S_1, S_2 unverändert.
Der wesentliche Beweisschritt ist ... □

Hilfssatz 7.2.
Durch Z_1, Z_2, S_1, S_2 kann T in die Form $\begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}$ *mit $d_1 \sim ggT(T)$ gebracht werden, wobei d_1 alle Einträge in $*$ teilt.*

Aus dem Hilfssatz folgt das Theorem 7.74 durch Induktion.
Wir setzen $\Phi(T) := \min\{\Phi(a_{ij}) : 1 \leq i \leq m, 1 \leq j \leq n, a_{ij} \neq 0\}$.

Beweis:
Induktionsbeginn: $\Phi(T) = 0$: durch Vertauschen von Zeilen und Spalten erreiche $\Phi(a_{11}) = 0$. Es gilt $a_{11} \mid a_{ij}\ \forall\ i, j$.
Daher können wir durch Transformationen $(Z_1), (S_1)$ Nullen in der ersten Spalte und der ersten Zeile erzeugen.

O.E. sei $\Phi(T) = \Phi(a_{11})$
1. Schritt: ist $a_{11} \nmid a_{i1}$ für ein $2 \leq i \leq m$, so dividiere mit Rest $a_{i1} = q \cdot a_{11} + r$ mit $\Phi(r) < \Phi(a_{11})$. Ersetze Z_i durch $Z_i - qZ_1$. Die neue Matrix T' hat $a'_{i1} = r$ hat also $\Phi(T') \leq \Phi(r) < \Phi(T)$. Also dann fertig per Induktion.
Daher können wir annehmen $a_{11} \mid a_{i1}\ \forall\ i$. Ersetze Z_i durch $Z_i - \frac{a_{i1}}{a_{11}} Z_1$ $(i = 2, \ldots, n)$,
erreiche so $T' = \begin{pmatrix} a_{11} & a'_{12} & \ldots & a'_{1n} \\ 0 & & & \\ \vdots & & (a'_{ij}) & \\ 0 & & & \end{pmatrix}$.
2. Schritt: ist $a_{11} \mid a_{1j}$ für ein $j \in \{2, \ldots, n\}$, so verkleinere $\Phi(T)$ wie oben. Andernfalls
erzeuge die Form $T'' = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & (a''_{ij}) & \\ 0 & & & \end{pmatrix}$ analog zum 1. Schritt.
3. Schritt: ist jetzt $a_{11} \sim \mathrm{ggT}(T'')$, so fertig.
Andernfalls gibt es $p, q \geq 2$ mit $a_{11} \nmid a''_{pq}$.
Ersetze S_1 durch $S_1 + S_q$, erhalte so die Matrix $T''' = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ a_{2q} & & & \\ \vdots & & (a''_{ij}) & \\ a_{mq} & & & \end{pmatrix}$, und gehe
mit diesem T''' zurück zum 1. Schritt.
Dort tritt dann der erste Fall ein, und wir reduzieren also auf ein T'''' mit $\Phi(T'''') < \Phi(T)$.
Fertig per Induktion.

□

Bemerkung:
Die Größe $\Phi(T)$ war eine Hilfsgröße für den Beweis, sie ist nicht invariant unter den elementaren Transformationen.

Korollar 7.75.
Sei A ein Hauptidealring, sei $T \in M_{m \times n}(A)$. Dann existieren $U \in GL_m(A), V \in GL_n(A)$, so dass UTV die Form $(*)$ mit $d_1 \mid \ldots \mid d_r \neq 0$ hat.

Beweis:
Für A euklidisch: gerade gezeigt (7.74).
(Für die restlichen Fälle (d.h. A nicht euklidisch) siehe z.B. Bröcker LA, Kapitel X.)

□

Definition 7.76.
Die Elemente $d_1, \ldots, d_r$ aus $(*)$ heißen die **Elementarteiler** der Matrix T. (Eindeutigkeit bis auf $\sim$ siehe unten)
Der Beweis oben hat gezeigt: $d_1 \sim \mathrm{ggT}(T)$.

Korollar 7.77.
Sei A ein Hauptidealring, sei $f : N \to M$ eine lineare Abbildung zwischen zwei endlich erzeugten, freien A-Moduln. Dann gibt es Basen $(x_1, ldots, x_m)$ von M und $(y_1, \ldots, y_n)$ von N, sowie $0 \leq r \leq \min\{m, n\}$ und $d_1 \mid \ldots \mid d_r \neq 0$ in A mit $f(y_i) = d_i x_i$, $i = 1, \ldots, r$ und $f(y_i) = 0$, $i = r+1, \ldots, n$.

Beweis:
Starte mit beliebigen Basen von M und N, stelle zu diesen die Matrix T von f auf, finde U, V zu T wie in 7.75, fasse U, V als Basiswechselmatrizen auf, erhalte so die neuen Basen.

□

Korollar 7.78.
(sei A ein Hauptidealring)
Sei M ein freier A-Modul, $rk\,(M) < \infty$, sei U ein Untermodul von M.
Dann existiert eine Basis $(x_1, \ldots, x_m)$ von M, sowie Elemente $d_1 \mid d_2 \mid \ldots \mid d_r \neq 0$ in A $(r \geq 0)$, so dass $(d_1 x_1, \ldots, d_r x_r)$ eine Basis von U ist. Insbesondere ist U selbst frei und $rk\,(U) \leq rk\,(M)$.

Beweis:
Nach 7.69 ist U endlich erzeugt. Es gibt also einen freien A-Modul N, $\mathrm{rk}\,(N) < \infty$ und eine surjektive Abbildung $f : N \to U$.
(sind $x_1, \ldots, x_n$ Erzeuger von U, so nimm $N = A^n$, $f(a_1, \ldots, a_n) := \sum a_i x_i$)
Auf die lineare Abbildung $f : N \to U \subset M$ wende 7.77 an:
es gibt also eine Basis $(x_1, \ldots, x_m)$ von M und $r \geq 0, d_1 \mid \ldots \mid d_r \neq 0$, so dass $U = \mathrm{im}(f)$ von den $d_i x_i$ $(i = 1, \ldots, r)$ erzeugt wird. Die $d_i x_i$ $(i = 1, \ldots, r)$ sind linear unabhängig, denn $0 = \sum_i a_i (d_i x_i) = \sum_i \underbrace{(a_i d_i)}_{=0} x_i \Rightarrow a_i = 0$ wegen $d_i \neq 0$ und A nullteilerfrei.

□

Bemerkung 7.79.
Anders als bei Vektorräumen kann ein echter Untermodul U des freien A-Moduls M denselben Rang wie M haben.
Beispiel: $A = M = \mathbb{Z}$, $U = n\mathbb{Z}$ für $n \in \mathbb{N}, n \geq 2$.

Beispiel 7.80.
Sei $A = \mathbb{Z}$, sei $T = \begin{pmatrix} -10 & 13 & 14 \\ 23 & 3 & 7 \end{pmatrix}$.
Die durch T gegebene lineare Abbildung ist
$T : \mathbb{Z}^3 \to \mathbb{Z}^2, T(y_1, y_2, y_3) = (-10y_1 + 13y_2 + 14y_3, 23y_1 + 3y_2 + 7y_3)$.
ggT$(T) \sim 1$. Führe folgende Umformungen durch:
$S_3 \rightsquigarrow S_3 - S_2$, $X \rightsquigarrow X \cdot Q_{23}(-1)$: $\begin{pmatrix} -10 & 13 & 1 \\ 23 & 3 & 4 \end{pmatrix}$
$S_1 \leftrightsquigarrow S_3$, $X \rightsquigarrow X \cdot \Pi_{13}$: $\begin{pmatrix} 1 & 13 & 23 \\ 4 & 3 & -10 \end{pmatrix}$
$S_2 \rightsquigarrow S_2 - 13S_1$ und $S_3 \rightsquigarrow S_3 + 10S_1$, $x \rightsquigarrow X \cdot Q_{12}(-13) \cdot Q_{13}(10)$: $\begin{pmatrix} 1 & 0 & 0 \\ 4 & -49 & 63 \end{pmatrix}$
$Z_2 \rightsquigarrow Z_2 - 4Z_1$, $X \rightsquigarrow Q_{21}(-4) \cdot X$: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -49 & 63 \end{pmatrix}$
(...), $X \rightsquigarrow X \cdot W$: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \end{pmatrix}$ mit $W := \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 9 \\ 0 & -3 & 7 \end{pmatrix}$. Erklärung von (...):
ggT$(49, 63) \sim 7$, euklidischer Algorithmus gibt $7 = 4 \cdot 49 - 3 \cdot 63$.
$(-49, 63) \cdot \begin{pmatrix} -4 & 9 \\ -3 & 7 \end{pmatrix} = (7, 0)$
Daraus erkenne für das gesuchte $W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 9 \\ 0 & -3 & 7 \end{pmatrix}$ die $-4, -3$. Die letzte Spalte durch Division von (*) durch ggT ~ 7.

Diese Rechnung zeigt unter anderem:

(a) Die Elementarteiler von T sind $1, 7$,

(b) Für $U = Q_{21}(-4) = \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}$ und $V = Q_{23}(-1) \cdots W = \begin{pmatrix} 0 & -3 & 7 \\ -1 & -26 & 56 \\ 1 & 22 & -47 \end{pmatrix}$ ist $UTV = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \end{pmatrix}$, und U, V sind invertierbar.

(c) Sind y_1, y_2, y_3 die Spalten von V (eine Basis von $\mathbb{Z}^3$) und x_1, x_2 die Spalten von U^{-1} (eine Basis von $\mathbb{Z}^2$, $x_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, x_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$), so ist $Ty_1 = d_1x_1, Ty_2 =$

$d_2x_2, Ty_3 = 0$.
Denn: $UTVe = e_1, UTVe_2 = 7e_2, UTVe_3 = 0$, d.h.
$T(Ve_1) = U^{-1}e_1, T(Ve_2) = U^{-1}e_2, T(Ve_3) = 0$.

(d) Die Spalten von T erzeugen die Untergruppe $M := \mathrm{im}(T)$ von $\mathbb{Z}^2$. Sie ist ein freier $\mathbb{Z}$-Modul vom Rang 2, mit Basis $(x_1, 7x_2)$,

Theorem 7.81.
*Sei A ein Hauptidealring, M ein endlich erzeugter A-Modul. Dann gibt es $n \geq 0$, $r \geq 0$ sowie Nichteinheiten $d_1 \mid d_2 \mid \ldots \mid d_r \neq 0$ in A mit $M \cong A^n \oplus A/(d_1) \oplus \ldots \oplus A/(d_r)$. Insbesondere ist M eine direkte Summe von zyklischen Moduln. Die Elemente $d_1, \ldots, d_r$ heißen die **Elementarteiler** von M und n heißt der **torsionsfreie Rang** von M.*

Bemerkung:
Wir werden sehen: $d_1, \ldots, d_r$ (bis auf $\sim$) und n sind eindeutig bestimmt.

Beweis:
Wähle Erzeugendensystem $(x_1, \ldots, x_m)$ von M. Erhalte $M \cong A^m/U$, wobei U der Kern von $A^m \to M$, $(a_1, \ldots, a_m) \mapsto \sum a_i x_i$ ist.

Nach Korollar 7.78 können wir annehmen, dass U von $d_1e_1, \ldots, d_re_r$ erzeugt ist, wobei $0 \leq r \leq m$ und $d_1 \mid \ldots \mid d_r \neq 0$.
Damit ist

$$M \cong A^m/U = \frac{\overbrace{A \oplus \ldots \oplus A}^{r} \oplus A^{m-r}}{Ad_1e_1 \oplus \ldots \oplus Ad_re_r} \cong A/(d_1) \oplus \ldots \oplus A/(d_r) \oplus A^{m-r}.$$

O.E. sind hierbei die $d_i \notin A^*$, da sonst $A/(d_i) = \{0\}$.

□

Definition 7.82.
Sei M ein A-Modul.

(a) $x \in M$ heißt **Torsionselement**, wenn $\exists\, a \in A$ mit $ax = 0$ und $a \neq 0$.
$M_{tors} := \{x \in M : x \text{ ist Torsionselement}\}$.

(b) M heißt **Torsionsmodul**, falls $M = M_{tors}$ und heißt **torsionsfrei**, falls $M_{tors} = \{0\}$.

Bemerkung 7.83.

1. M_{tors} ist ein Untermodul von M.
(Denn: Ist $ax = 0, by = 0$ mit $a, b \neq 0$, so $ab(x+y) = 0$ (und $ab \neq 0$) $\Rightarrow x+y \in M_{tors}$; ebenso $a(cx) = c(ax) = 0 \Rightarrow cx \in M_{tors}$)

2. M/M_{tors} ist torsionsfrei: aus $x \in M$ und $ax \in M_{tors}$ mit $a \neq 0$ folgt: $\exists\, b \neq 0$ $b(ax) = 0 \Rightarrow \underbrace{(ba)}_{\neq 0} x = 0 \Rightarrow x \in M_{tors}$.

Bemerkung 7.84.
Ist M ein A-Modul und $\pi \in A$ ein unzerlegbares Element, so heißt $M(\pi) = \{x \in M : \exists\, n \in \mathbb{N}, \pi^n x = 0\}$ die **π-primäre Komponente** von M_{tors}. Die ist ein Untermodul von M_{tors}, denn $x, y \in M(\pi)$, etwa $n^m x = 0, \pi^n y = 0 \Rightarrow \pi^{m+n}(x+y) = 0 \Rightarrow x + y \in M(\pi)$.
Der Modul M heißt **π-primär**, wenn $M = M(\pi)$ ist.

Beispiele 7.85.

1. Für jedes $n \in \mathbb{N}$ ist $M = A/(\pi^n)$ ein π-primärer Modul.
2. ($A = \mathbb{Z}$) Die endlich erzeugte abelsche Gruppe $M = \mathbb{Z}/(60)$ hat welche Primärkomponenten ($60 = 2^2 \cdot 3 \cdot 5$) $M \cong \mathbb{Z}/(4) \oplus \mathbb{Z}/(3) \oplus \mathbb{Z}/(5)$ (Chinesischer Restsatz).
   $\Rightarrow M(2) \cong \mathbb{Z}/(4), M(3) \cong \mathbb{Z}/(3), M(5) \cong \mathbb{Z}/(5), M(p) = \{0\}\ \forall\ p > 5$ (p Primzahl).

Bemerkungen 7.86.

1. Sind $\pi, \rho \in A$ unzerlegbare Elemente mit $\pi \nsim \rho$, so ist $M(\pi) \cap M(\rho) = \{0\}$.
   Denn: Sei $x \in M(\pi) \cap M(\rho)$, etwa $\pi^m x = \rho^n x = 0$ mit $m, n \in \mathbb{N}$. Dann: $\mathrm{ggT}(\pi, \rho) \sim 1$, also $\exists\, a, b \in A$ mit $1 = a\pi + b\rho$.
   $\Rightarrow x = 1 \cdot x = (a\pi^m + b\rho^n)x = 0$.
2. $(M \oplus N)(\pi) = M(\pi) \oplus N(\pi)$.

Theorem 7.87. *(Variante von 7.81)*
Sei M ein endlich erzeugter A-Modul. Dann ist M die direkte Summe aus einem freien Modul und aus (endlich vielen zyklisch primären Torsionsmoduln.
$M \cong A^n \oplus A/(\pi_1^{n_1}) \oplus \ldots \oplus A/(\pi_s^{n_s})$, *$\pi_i$ unzerlegbar, $n_i \in \mathbb{N}$.*

Beweis:
Folgt aus 7.81: $M \cong A^n \oplus A/(d_1) \oplus \ldots \oplus A/(d_r)$ (mit $d_1 \mid \ldots \mid d_r$). Für $d \in A, d \neq 0, d \notin 0$, $d \notin A^*$, sei $d = \pi^{m_1} \cdots \pi_k^{m_k}$ mit π_i unzerlegbar, $\pi_i \nsim \pi_j$ $(i \neq j)$, $m_i \geq 1$. Chinesischer Restsatz $\Rightarrow$
$A/(d) \cong A/(\pi_1^{m_1}) \oplus \ldots \oplus A/(\pi_k^{m_k})$.

□

Zusatz:
Sind $d_1, \ldots, d_r$ die Elementarteiler von M, so gilt für alle unzerlegbaren Elemente $n \in A$:
$M(\pi) \neq \{0\} \Leftrightarrow \pi$ teilt eines der $d_i \Leftrightarrow \pi \mid d_r$.

7.88.
Sei M endlich erzeugt.
Schreibe $M = F \oplus T$ mit $F \cong A^n$ frei, T Torsion.
Was ist an dieser Zerlegung eindeutig?

1. $T = M_{tors}$, ist also als Untermodul von M eindeutig bestimmt. Damit ist $F \cong M/T$ eindeutig bis auf Isomorphie, und daher ist auch $n = \text{rk}\,(F)$ eindeutig (Lemma in 7.c).
   Ebenso sind die Primärkomponenten $M(\pi)$ eindeutig.

2. Der Untermodul F von M ist im Allgemeinen *nicht* eindeutig bestimmt.
   Beispiel ($A = \mathbb{Z}$): $M = \mathbb{Z} \oplus \mathbb{Z}/(2)$; dabei sind $F_1 = \mathbb{Z}(1,0)$ und $F_2 := \mathbb{Z}(1,1)$ jeweils freie Untermoduln mit $M = F_1 \oplus T = F_2 \oplus T$.

3. Ist $M = M(\pi)$ ein π-primärer Torsionsmodul, etwa $M \cong A/(\pi^{e_1}) \oplus \ldots \oplus A/(\pi^{e_r})$ ($e_i \geq 1$), so sind die zyklischen Untermoduln von M aus dieser Zerlegung im Allgemeinen *nicht* eindeutig.
   Beispiel: $A = \mathbb{Z}, \pi = p$ Primzahl, $M = (\mathbb{Z}/p)^n = \bigoplus_{i=1}^{n} \mathbb{Z}/p$.
   Ist $n \geq 2$, so hat M mehr als eine $\oplus$-Zerlegung in 1-dimensionale $\mathbb{Z}/p$-VRe.

4. Jedoch sind die e_i eindeutig.

Satz 7.89.
Sei $\pi \in A$ unzerlegbar, M ein endlich erzeugter π-primärer Torsionsmodul, etwa $M \cong A/(\pi^{e_1}) \oplus \ldots \oplus A/(\pi^{e_r})$ mit $r \geq 0$, $e_i \in \mathbb{N}$.
Dann sind r und $e_1, \ldots, e_r$ (bis auf Vertauschung) eindeutig durch M bestimmt.

Beweis:
Für $i = 0, 1, 2, \ldots$ sei $\pi^i M = \{\pi^i x : x \in M\}$ ein Untermodul von M: $M = \pi^0 M \supset \pi M \supset \pi^2 M \supset \ldots$
Sei $V_i := V_i(M) := (\pi^{i-1}M)/(\pi^i M)$, $i \geq 1$.
Das sind von π annullierte A-Moduln, also Moduln über $A/(\pi) =: k$, ein Körper (also Vektorräume).
Was ist $\dim_k(V_i)$?
Betrachte dazu einen einzelnen Summanden $N = A/(\pi^e)$ von M. Ist $i \leq e$, so ist $\pi^i N = (\pi^i)/(\pi^e)$; ist $i > e$, so ist $\pi^i N = \{0\}$. Also:

$$V_i(N) \cong \begin{cases} (\pi^{i-1})/(\pi^i) \cong k & i \le e \\ \{0\} & i > e. \end{cases}$$

Daraus sehen wir: für alle $I \in N$ ist $\dim_k V_i(M)$ gleich der Anzahl $j \in \{1, \ldots, r\}$ mit $e_j \ge i$.
Die Zahlen $\dim_k V_i(M)$ $(i = 1, 2, \ldots)$ sind eindeutig bestimmt durch M, also auch die e_i und r.

□

Beispiel 7.90.
$M = A/(\pi) \oplus A/(\pi^2) \oplus A/(\pi^2) \oplus A/(\pi^4) \oplus A/(\pi^5)$.
Für $i = 1, 2, 3, 4, 5$ ist $\dim_k V_i(M) = 5, 4, 2, 2, 1$, für $i \ge 6$: $V_i = \{0\}$.

Folgerung 7.91.
Die Elementarteiler einer Matrix, oder eines endlich erzeugten A-Moduls sind eindeutig bestimmt (bis auf ∼).

Beweis:
Sei $M = M_{tors}$ endlich erzeugt, etwa $M \cong A/(d_1) \oplus \ldots \oplus A/(d_r)$ mit $d_1 \mid \ldots \mid d_r \neq 0$.
Sei Π ein Vertretersystem für die unzerlegbaren Elemente, sei $d_i \sim \prod_{\pi \in \Pi} \pi^{e_i(\pi)}$ mit $e_i(\pi) \ge 0$ und $e_i(\pi) = 0$ *f.f.a.* π, für jedes i.
Es ist $e_1(\pi) \le \ldots \le e_r(\pi)$ für jedes $\pi \in \Pi$, und $M(\pi) \cong A/(\pi^{e_1(\pi)}) \oplus \ldots \oplus A/(\pi^{e_r(\pi)})$.
Nach 7.89 sind die $e_i(\pi)$ eindeutig bestimmt woraus die Behauptung (im Modulfall) folgt.

Ist $T \in M_{m \times n}(A)$, so sind die Elementarteiler von T nach Definition die Elementarteiler des A-Moduls $A^m/\mathrm{im}(T)$. Sie sind also ebenfalls eindeutig bestimmt.

□

Korollar 7.92.
Sei M ein endlich erzeugter Torsionsmodul mit Elementarteilern $d_1 \mid \ldots \mid d_r \neq 0$. Dann gilt:
M zyklisch $\Leftrightarrow r = 1$.

Beweis:
Das folgt aus der Eindeutigkeit der Elementarteiler.

□

Beispiel 7.93.
($A = \mathbb{Z}$) Sei $M = \mathbb{Z}/(30) \oplus \mathbb{Z}/(8) \oplus \mathbb{Z}/(18)$. Zerlege M in $\oplus$ von primärzyklischen

Moduln, und bestimme die Elementarteiler von M.

Lösung:
Chinesischer Restsatz $\Rightarrow M \cong (\mathbb{Z}/2 \oplus \mathbb{Z}/3 \oplus \mathbb{Z}/5) \oplus \mathbb{Z}/2^3 \oplus (\mathbb{Z}/2 \oplus \mathbb{Z}/3^2)$.
Umsortieren: $M \cong (\mathbb{Z}/2 \oplus \mathbb{Z}/2 \oplus \mathbb{Z}/2^3) \oplus (\mathbb{Z}/3 \oplus \mathbb{Z}/3^2) \oplus \mathbb{Z}/5 \cong \mathbb{Z}/2 \oplus \mathbb{Z}/6 \oplus \mathbb{Z}/360$.
Also hat M die Elementarteiler $2, 6, 360$.

Umgekehrt erhält man aus der Elementarteilerform auch wieder die Zerlegung in Primärkomponenten.

Bemerkung 7.94.
Hauptanwendung des Elementarteilersatzes: Lösen von Linearen Gleichunssystemen über Hauptidealringen.
Sei $T \in M_{m\times n}(A)$ (A sei ein Hauptidealring), sei $w \in A^m$, betrachte das Lineare Gleichungssystem $Tx = w$ ($x = (x_1, \ldots, x_n)^t$).
Lösungsmenge über A: $L := \{x \in A^n : Tx = w\}$.

Wie zuvor finde $U \in GL_m(A), V \in GL_n(A)$ mit $UTV = \left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \end{array}\right) =: D$ mit

$d_1 \mid \ldots \mid d_r \neq 0$.
Dann ist $Tx = w$ äquivalent zu $U^{-1}DV^{-1}x = w$, also zu $D\widetilde{x} = \widetilde{w}$ mit $\widetilde{x} := V^{-1}x, \widetilde{w} := Uw$.
Also hat $Tx = w$ genau dann eine Lösung in A^n, wenn gilt $d_i \mid \widetilde{w_i}$ ($i = 1, \ldots, r$) und $\widetilde{w_i} = 0$ ($i = r+1, \ldots, n$).
Sind diese erfüllt, so kann man die Menge aller Lösungen $\widetilde{x}$ sofort hinschreiben und erhält mit $x = V\widetilde{x}$ die Lösungen der Ausgangsgleichung.

Beispiel 7.95.
Betrachte
$-10x_1 + 13x_2 + 14x_3 = w_1$
$23x_1 + 3x_2 + 7x_3 = w_2$
mit gegebenen $w_1, w_2 \in \mathbb{Z}$. Gemäß 7.80 ist für $T = \begin{pmatrix} -10 & 13 & 14 \\ 23 & 3 & 7 \end{pmatrix}$: $UTV = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$ mit $U = \begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}, V = \begin{pmatrix} 0 & -3 & 7 \\ -1 & -26 & 56 \\ 1 & 22 & -47 \end{pmatrix}$.
Schreibe $dtw = UW = \begin{pmatrix} \widetilde{w_1} \\ \widetilde{w_2} \end{pmatrix} = \begin{pmatrix} w_1 \\ -4w_1 + w_2 \end{pmatrix}$. Dann ist (*) lösbar $\Leftrightarrow \widetilde{w_2} = w_2 - 4w_1 \equiv 0 (\mod 7)$.

In diesem Fall hat $D\widetilde{x} = \widetilde{w}$ die Lösungen $\begin{pmatrix} \widetilde{x_1} \\ 7\widetilde{x_2} \\ \widetilde{x_3} \end{pmatrix} = \begin{pmatrix} \widetilde{w_1} \\ \widetilde{w_2} \\ \widetilde{w_3} \end{pmatrix}$, also $\begin{pmatrix} \widetilde{x_1} \\ \widetilde{x_2} \\ \widetilde{x_3} \end{pmatrix} = \begin{pmatrix} w_1 \\ \frac{1}{7}(w_2 - 4w_1) \\ n \end{pmatrix}$
für $n \in \mathbb{Z}$ beliebig.
Die Lösungen von (*) sind also die (setze $k := \frac{1}{7}(w_2 - 4w_1)$) $x = V\widetilde{x}$:
$\begin{pmatrix} -3k \\ -w_1 - 26k \\ w_1 + 22k \end{pmatrix} + n \begin{pmatrix} 7 \\ 56 \\ -47 \end{pmatrix}, n \in \mathbb{Z}.$

e. Zyklische Endomorphismen

7.96.
Sei $\mathbb{K}$ ein Körper, V ein $\mathbb{K}$-Vektorraum, $\dim(V) < \infty$ und $f \in End_{\mathbb{K}}(V)$. Wie in **??** bezeichne V_f den durch $p(t) \cdot v := p(f)(v)$ $(p \in \mathbb{K}[t], v \in V)$ definierten $\mathbb{K}[t]$-Modul. V ist ale $\mathbb{K}[t]$-Modul endlich erzeugt, und ist ein Torsionsmodul: das Minimalpolynom Q_f von f annuliert ganz V: $Q_f(t) \cdot v = Q_f(f)(v) = 0 \; \forall \; v \in V$.
Tatsächlich haben wir gezeigt: Q_f erzeugt das **Annullatorideal** $\text{Ann}\,(V_f) = \{p \in \mathbb{K}[t] : p \cdot v = 0 \; \forall \; v \in V\}$ von V_f.

Satz 7.97.
Sei $V = \mathbb{K}^n$, seien $A, B \in M_n(\mathbb{K})$, seien V_A, V_B die entsprechenden $\mathbb{K}[t]$-Moduln. Dann ist äquivalent:

(i) $A \approx B$,

(ii) $V_A \cong V_B$ *als* $\mathbb{K}[t]$*-Moduln,*

(iii) V_A *und* V_B *haben dieselben Elementarteiler.*

Beweis:
(ii) bedeutet (siehe 7.53): $\exists\, S \in GL_n(A)$ mit $SA = BS$, d.h. mit $B = SAS^{-1}$.
Also (ii) $\Leftrightarrow$ (i).

(ii) $\Leftrightarrow$ (iii) klar.

□

Nach dem Struktursatz 7.81 ist V_f direkte Summe von (endlich vielen) zyklischen Moduln.

Definition 7.98.
$f \in End(V)$ heißt **zyklisch**, wenn der $\mathbb{K}[t]$-Modul V_f zyklisch ist.

Satz 7.99.
Sei $n = \dim(V) > 0$. Es sind äquivalent:

(i) f *ist zyklisch,*

(ii) $\exists\, v \in V$*, so dass* $(v, f(v), f^2(v), \ldots, f^{n-1}(v))$ *eine Basis von* V *ist,*

(iii) bezüglich einer geeigneten Basis von V hat f eine Matrix der Form

$$A = \begin{pmatrix} 0 & & & a_0 \\ 1 & 0 & & a_1 \\ & 1 & \ddots & \vdots \\ & & \ddots & 0 \quad a_{n-2} \\ & & & 1 \quad a_{n-1} \end{pmatrix} \qquad (*)$$

mit $a_0, \ldots, a_{n-1} \in \mathbb{K}$.

Beweis:
(i) $\Rightarrow$ (ii): f ist zyklisch, bedeutet: $\exists\, v \in V$, so dass $V = \text{span}\,(f^j(v) : j = 0, 1, \ldots)$. Behaupte, dann ist $\mathcal{B} := (v, f(v), \ldots, f^{n-1}(v))$ eine Basis von V. Es genügt zu zeigen, dass $\mathcal{B}$ linear unabhängig ist.

Sei $\mathcal{B}$ linear abhängig. $\Rightarrow f^k(v) = \sum_{j=0}^{k-1} a_j f^j(v)$ mit $1 \le k \le n-1$ und $a_0, \ldots, a_k \in \mathbb{K}$.
Daraus folgt dann induktiv: $f^m(v) \in \text{span}\,(v, f(v), \ldots, f^{k-1}(v)) \neq V \;\forall\; m \ge k$. $\Rightarrow$ Widerspruch.

(ii) $\Rightarrow$ (iii): Bezüglich der Basis $\mathcal{B}$ hat f die Matrix der Form (*).

(iii) $\Rightarrow$ (i): Anwenden von f auf den ersten Basisvektor liefert nacheinander die gesamte Basis. □

Lemma 7.100.
Sei $V_f \cong \mathbb{K}[t]/(p)$, mit einem normierten Polynom $o = p(t)$. Dann ist $p = P_f$, das charakteristische Polynom von f. Ist $p = t^n - \sum_{i=0}^{n-1} a_i t^i$, so hat f bezüglich einer geeigneten Basis genau die Matrix der Form (*) aus 7.99.

Beweis:
Bezüglich der Basis $\bar{1}, \bar{t}, \ldots, \overline{t^{n-1}}$ von $\mathbb{K}[t]/(p)$ (NB: $n = \dim(V) = \deg(p)$) hat die Multiplikation mit $\bar{t}$ in $\mathbb{K}[t]/(p)$ die Matrix A =(*) aus 7.99.
Direkte Rechnung zeigt:
$P_A(t) = p(t) = tn - \sum_{j=0}^{n-1} a_i t^i$:

$$\begin{vmatrix} t & & & -a_0 \\ -1 & t & & -a_1 \\ & -1 & \ddots & \vdots \\ & & \ddots & t \quad -a_{n-2} \\ & & & -1 \quad t - a_{n-1} \end{vmatrix} = \begin{vmatrix} 0 & & & t^n - a_{n-1}t^{n-1} - \ldots - a_0 \\ -1 & 0 & & \vdots \\ & -1 & \ddots & \vdots \\ & & \ddots & 0 \quad \vdots \\ & & & -1 \quad t - a_{n-1} \end{vmatrix}$$

□

Korollar 7.101.
Sei $f \in End(V)$, nicht notwendig zyklisch. Seien $p_1 \mid \ldots \mid p_r \in \mathbb{K}[t]$ die normierten Elementarteiler der $\mathbb{K}[t]$-Moduls V_f. Dann ist $p_f = p_1 \cdots p_r$ und $Q_f = p_r$.

Beweis:
Nach Voraussetzung ist $V_f \cong \bigoplus_{i=1}^{r} \mathbb{K}[t]/(p_i)$.
Sei $V = V_1 \oplus \ldots \oplus V_r$ die zugehörige Zerlegung von V in f-invariante Unterräume. Dann hat $f|_{V_i} \in End(V_i)$ das charakteristische Polynom $p_i(t)$ $(i = 1, \ldots, r)$. Also $P_f(t) = \prod_{i=1}^{r} p_i(t)$.
Für jedes $i = 1, \ldots, r$ erzeugt $p_i(t)$ das Ideal Ann $(\mathbb{K}[t]/(p_i))$ = Ann $((V_i)_{f|_{V_i}})$. Wegen $p_i \mid p_r$ folgt: (p_r) = Ann (V_f), also $p_r = Q_f$.

□

Bemerkung 7.102.

1. Dies gibt uns einen neuen Beweis des Satzes von Cayley-Hamilton: $Q_f \mid P_f$.
2. Wir sehen insbesondere: P_f und Q_f haben dieselben irreduziblen Teiler in $\mathbb{K}[t]$.

Korollar 7.103.
Für $f \in End(V)$ sind äquivalent:

(i) *f ist zyklisch,*

(ii) *$P_f = Q_f$.*

Falls P_f/t) über $\mathbb{K}$ in Linearfaktoren zerfällt, sind auch äquivalent:

(iii) *Alle Eigenräume von f sind 1-dimensional,*

(iv) *zu jedem Eigenwert von f gibt es nur einen Jordanblock.*

Beweis:
$P_f = p_1, \ldots, p_r$, $Q_f = p_r$.
Also $P_f = Q_f \Leftrightarrow r = 1 \Leftrightarrow f$ ist zyklisch.

□

Bemerkung 7.104.

1. Wir erhalten auch die Jordansche Normalform aus dem Elementarteilersatz: Es zerfalle also $P_f(t)$ in Linearfaktoren. Ist f zyklisch mit $P_f(t) = (t-\lambda)^k$, $\lambda \in \mathbb{K}$, und ist $v \in V$ mit $V = \mathbb{K}[t]\cdot v$, so bilden die $v_j := (f-\lambda)^j(v)$ $(j = 0, 1, \ldots, k-1)$ eine Basis von V.
   Denn 1,
   $ovt-\lambda, (\bar{t}-\lambda)^2, \ldots, (\bar{t}-\lambda)^{k-1}$ bilden eine $\mathbb{K}$-Vektorraumbasis von $\mathbb{K}[t]/(t-\lambda)^k$.
   Bezüglich der Basis $(v_{k-1}, \ldots, v_1, v_0 = v)$ von V hat f die Matrix $\begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$

2. Wir erhalten auch in Fällen, wo $P_f(t)$ nicht in Linearfaktoren zerfällt, verallgemeinerte Jordansche Normalformen:
   ist f zyklisch mit $P_f(t) = p(t)^k$ mit $\deg(p) = d \geq 1$ (und p irreduzibel), und ist $V_f = \mathbb{K}[t] \cdot v$, so kann man die $\mathbb{K}$-Basis $(f^j \circ p(f)^i)(v)$ $(0 \leq i < d, 0 \leq j < k)$ verwenden, um eine verallgemeinerte Jordan-Normalform zu erhalten.
   Über $\mathbb{K} = \mathbb{R}$ etwa ist jedes irreduzible Polynom vom Grad ≤ 2. Für $p = t^2 + at + b$ erhält man so den „verallgemeinerten Jordanblock“
   $$\begin{pmatrix} -a & 1 & & & & & & \\ -b & 0 & 1 & & & & & \\ & & -a & 1 & & & & \\ & & -b & 0 & 1 & 0 & & \\ & & & & \ddots & \ddots & & \\ & & & & & & 1 & 0 \\ & & & & & & -a & 1 \\ & & & & & & -b & 0 \end{pmatrix}$$

8. Affinitäten und Hauptachsentransformation

a. Affine Räume und Affinitäten

Sei V ein $\mathbb{K}$-Vektorraum.

8.1.
Eine Teilmenge A von V heißt ein **affiner Unterraum** von V, falls $A = \emptyset$ ist oder $A = u + W$ mit einem $u \in V$ und einem Untervektorraum W von V ist. Dabei ist $T(A) := W$, der **Translationsraum** von A eindeutig durch A bestimmt. Man setzt $\dim(A) := \dim T(A) = \dim(W)$, $\dim(\emptyset) := -1$.
Jeder Schnitt ($\cap$) von affinen Unterräumen von V ist wieder ein affiner Unterraum von V. (Wiederholung aus Kapitel 3.3).

Definition 8.2.
Seien $v_0, \ldots, v_m \in V$. Die Vektoren $\sum_{i=0}^{m} a_i v_i$ mit $a_i \in \mathbb{K}$ und $\sum_{i=0}^{m} a_i = 1$ heißen **Affinkombinationen** von $v_0, \ldots, v_m$.

Bemerkungen 8.3.

1. Ist $A = u + W$ ein affiner Unterraum von V, so ist jede Affinkombination von Elementen aus A wieder in A.
   Umgekehrt ist eine Teilmenge A von V genau dann ein affiner Unterraum von V, wenn gilt:
   $\forall\, v, w \in A\ \forall\, a \in \mathbb{K}\ (1-a)v + aw \in A$.
   (also, wenn A mit je zwei verschiedenen Punkten auch ihre Verbindungsgerade enthält)
   Beweis: Nach Verschieben können wir annehmen, $O \in A$. Zu zeigen jetzt: A ist Untervektorraum von V.
   Sind $x, y \in A$, so ist $x + y = 2 \cdot (\underbrace{\frac{1}{2}x + \frac{1}{2}y}_{=:z}) = 2z - 1 \cdot 0$

2. Für jede Teilmenge M von V ist die Menge aller Affinkombinationen von Elementen aus M ein affiner Unterraum von V, und damit der kleinste M enthaltende affine Unterraum von V.
   Beweis: Sind $x = \sum_{i=0}^{m} s_i x_i$, $y = \sum_{i=0}^{m} t_i x_i$ Affinkombinationen von $x_0, \ldots, x_m \in M$ (d.h. $\sum s_i = \sum t_i = 1$), und $a \in \mathbb{K}$, so ist $(1-a)\sum s_i x_i + a \sum t_i x_i$ eine Affinkombination von $x_0, \ldots, x_m$, denn $(1-a)\sum s_i + a \sum t_i = 1 - a + a = 1$.

Definition 8.4.
Sind $A_1, \ldots, A_m$ affine Unterräume von V gegeben, so heißt $A_1 \vee \ldots \vee A_m := \bigcap\{A : A \text{ ist affiner Unterraum}, A_i \subset A \text{ für } i = 1, \ldots, m\}$ der **Verbindungsraum** von $A_1, \ldots, A_m$. (LA I Blatt 8 Aufgabe 3).

Bemerkung 8.5.
$A_1 \vee \ldots \vee A_m$ besteht aus allen Affinkombinationen von Elementen in $A_1 \cup \ldots \cup A_m$. Ist speziell $A_i = \{v_i\}$ $(i = 1, \ldots, m)$, so ist $A_1 \vee \ldots \vee A_m = \{v_i\} \vee \ldots \vee \{v_m\} = \{\sum_{i=1}^{m} a_i v_i : \sum a_i = 1\}$.

Definition 8.6.
Seien $v_0, \ldots, v_m \in V$.

(a) $v_0, \ldots, v_m$ heißen **affin unabhängig**, wenn gilt: $a_0, \ldots, a_m \in \mathbb{K}$, $\sum_{i=0}^{m} a_i v_i = 0, \sum_{i=0}^{m} a_i = 0 \Rightarrow a_0 = \ldots = a_m = 0$.
Andernfalls heißen $v_0, \ldots, v_m$ **affin abhängig**.

(b) Ist A ein affiner Unterraum von V, sind $v_0, \ldots, v_m \in A$, so heißt $(v_0, \ldots, v_m)$ eine **affine Basis** von A, wenn $\{v_0\} \vee \ldots \vee \{v_m\} = A$ ist und $v_0, \ldots, v_m$ affin unabhängig sind.

Beachte: Genau dann ist $(v_0, \ldots, v_m)$ eine affine Basis von A, wenn jedes $v \in A$ eine *eindeutige* Affinkombination von $v_0, \ldots, v_m$ ist.

Lemma 8.7.
$(v_0, \ldots, v_m)$ ist genau dann affin unabhängig, wenn $(v_1 - v_0, \ldots, v_m - v_0)$ linear unabhängig ist.

Beweis:
Klar aus Definition.
$$\sum_{i=1}^{m} a_i(v_i - v_0) = \left(-\sum_{i=1}^{m} a_i\right) v_0 + \sum_{i=1}^{m} a_i v_i.$$
□

Bemerkung 8.8.

1. Sei A ein affiner Unterraum. Eine Familie $(v_0, \ldots, v_m)$ von Elementen aus A ist genau dann eine affine Basis von A, wenn $(v_1 - v_0, \ldots, v_m - v_0)$ eine

Vektorraum-Basis von $T(A)$ ist.
Insbesondere: je zwei affine Basen von A haben dieselbe Länge, nämlich $1 + \dim(A)$.

2. Standardbeispiel: für jedes $u \in \mathbb{K}^n$ ist $(u, u + e_1, \ldots, u + e_n)$ eine affine Basis von $\mathbb{K}^n$.

Lemma und Definition 8.9.
Seien V, W $\mathbb{K}$-Vektorräume, seien $A \subset V$, $B \subset W$ affine Unterräume, $A, B \neq \emptyset$, sei $\alpha : A \to B$ eine Abbildung. Für $v_0 \in A$ betrachte die Abbildung $f_{v_0} : T(A) \to T(B)$, $f_{v_0}(v) := \alpha(v_0 + v) - \alpha(v_0) \in T(B)$. Es sind äquivalent:

(i) $f_{v_0} : T(A) \to T(B)$ *ist linear für ein* $v_0 \in A$*;*

(ii) $f_{v_0} : T(A) \to T(B)$ *ist linear* $\forall\, v_0 \in A$.

*Sind (i), (ii) erfüllt, heißt α eine **affine Abbildung**.*
Die lineare Abbildung f_{v_0} hängt dann nicht von v_0 ab und wird mit $T(\alpha)$ bezeichnet.
*Eine **Affinität** ist eine bijektive affine Abbildung.*

Beweis:
Sei f_{v_0} linear, sei auch $v_1 \in A$. Dann ist für $v \in T(A)$: $f_{v_1}(v) = \alpha(v_1 + v) - \alpha(v_1) = \alpha(v_0 + (v_1 + v - v_0)) - \alpha(v_0 + (v_1 - v_0))$
$= f_{v_0}(\underbrace{v_1 + v - v_0}_{(v_1 - v_0) + v}) - f_{v_0}(\underbrace{v_1 - v_0}_{\in T(A)}) = f_{v_0}(v_1 - v_0) + f_{v_0}(v) - f_{v_0}(v - v_0) = f_{v_0}(v)$.

□

Bemerkungen 8.10.

1. Die affinen Abbildungen von A nach B sind also genau die Abbildungen der Form $\alpha : v \mapsto f(v - v_0) + w_0$ mit $v_0 \in A$, $w_0 \in B$ und $f : T(A) \to T(B)$ eine lineare Abbildung.
   Genau dann ist α eine Affinität, wenn f bijektiv, also ein Vektorraum-Isomorphismus, ist.
   Nebenbemerkung: v_0 kann dabei beliebig in A gewählt werden, w_0 ergibt sich dann als $w_0 = \alpha(v_0)$.

2. Die Komposition von affinen Abbildungen ist wieder eine affine Abbildung. Die Umkehrabbildung einer Affinität ist wieder eine Affinität.

Lemma 8.11.
Seien $A \subset V, B \subset W$ affine Unterräume von V bzw. W.
Sei $(v_0, \ldots, v_m)$ eine affine Basis von A.

(a) Zu jeder Familie $(w_0, \ldots, w_m)$ in B gibt es eine eindeutige affine Abbildung $\alpha : A \to B$ mit $\alpha(v_i) = w_i, i = 0, \ldots, m$.

(b) Dabei ist α genau dann eine Affinität, wenn $(w_0, \ldots, w_m)$ eine affine Basis von B ist.

Beweis:
In (a) muss α definiert werden durch $\alpha(v) := w_0 + f(v - v_0)$, wobei $f : T(A) \to T(B)$ die eindeutige lineare Abbildung ist mit $f(v_i - v_0) = w_i - w_0$ $(i = 1, \ldots, m)$ ist.
α ist bijektiv $\Leftrightarrow$ f bijektiv $\Leftrightarrow$ $(w_i - w_0)_{i=1,\ldots,m}$ ist Vektorraum-Basis von $T(B)$.
□

Die Affinitäten $\mathbb{K}^n \to \mathbb{K}^n$ sind also genau die Abbildungen $\alpha_{A,u} : x \mapsto Ax + u$ mit $u \in \mathbb{K}^n, A \in GL_n(\mathbb{K})$. Wir können $\alpha_{A,u}(x) =: x'$ auch ausdrücken durch

$$\begin{pmatrix} 1 \\ x'_1 \\ \vdots \\ x'_n \end{pmatrix} = \left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline u_1 & & & \\ \vdots & & A & \\ u_n & & & \end{array}\right) \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

So sehen wir ...

Satz 8.12.
Die Gruppe Aff $(\mathbb{K}^n)$ $(= GA(\mathbb{K}^n))$ der Affinitäten von $\mathbb{K}^n$ ist isomorph zur Gruppe aller Matrizen in $GL_{n+1}(\mathbb{K})$ mit ersten Zeilen $(1, 0, \ldots, 0)$

Sei weiter jeweils $\mathbb{K} = \mathbb{R}$.

Definition 8.13.
Sei V ein euklidischer Vektorraum $\dim(V) < \infty$. Eine Abbildung $f : V \to V$ heißt eine (euklidische) **Bewegung**, falls gilt $\forall\ v, w \in V\ \|f(v) - f(w)\| = \|v - w\|$ („**abstandstreu**“)

Satz 8.14.
$f : V \to V$ ist genau dann eine Bewegung, wenn f eine Affinität und $T(f)$ eine orthogonale Abbildung ist.
Also genau dann, wenn $f(v) = g(v) + w$ ist mit einem $w \in V$, einem $g \in O(V)$.

Beweis:
Sei f eine Bewegung, sei $u := f(0)$. Wir müssen zeigen, dass die Abbildung $g(v) := f(v) - u$ linear, sogar orthogonal ist.
$\forall\ v \in V$ ist $\|g(v)\| = \|v\|$. Also folgt für $v, w \in V$: $\|g(v) - g(w)\|^2 = \|g(v)\|^2 - 2\langle g(v), g(w)\rangle + \|g(w)\|^2 = \|v\|^2 - 2\langle g(v), g(w)\rangle + \|w\|^2$.
$\|g(v) - g(w)\|^2 = \|f(v) - f(w)\|^2 = \|v - w\|^2 = \|v\|^2 - 2\langle v, w\rangle + \|w\|^2$.
$\Rightarrow \langle g(v), g(w)\rangle = \langle v, w\rangle$.

Sei $(v_1, \ldots, v_n)$ eine ON-Basis von V. Dann ist auch $(g(v_1), \ldots, g(v_n))$ eine ON-Basis von V.
Für beliebige $v, w \in V, a, b \in \mathbb{K}$ ist
$\langle g(av + bw) - ag(v) - bg(w), g(v_i)\rangle = \langle g(av + bw), g(v_i)\rangle - a\langle g(v), g(v_i)\rangle - b\langle g(w), g(v_i)\rangle =$
$\langle av + bw, v_i\rangle - a\langle v, v_i\rangle - b\langle w, v_i\rangle = \ldots = 0$. $(i = 1, \ldots, n)$.
Also ist g linear und orthogonal.

□

Korollar 8.15.
Die Gruppe der euklidischen Bewegungen des $\mathbb{R}^n$ ist isomorph zur Untergruppe von $GL_{n+1}(\mathbb{R})$ aus allen $\left(\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline * & & S & \end{array}\right)$ *mit $S \in O(n)$.*

b. Affine Quadriken, Hauptachsentransformation

Sei $\mathbb{K}$ ein Körper, $char(\mathbb{K}) \neq 2$.

Definition 8.16.
Sei $n \in \mathbb{N}$. Ein **quadratisches Polynom** in den Unbestimmten $x = (x_1, \ldots, x_n)$ ist ein Ausdruck der Form $Q = Q(x_1, \ldots, x_n) = \sum_{j,k=1}^{n} a_{jk} x_j x_k + 2 \sum_{j=1}^{n} u_j x_j + a$ mit $a_{jk} = a_{kj}, u_j, a \in \mathbb{K}$.
Die Teilmenge $Y_Q = Y = \{x \in \mathbb{K}^n : Q(x) = 0\}$ von $\mathbb{K}^n$ heißt die durch Q definierte **affine Quadrik**. Im Fall $n = 2$ nennt man Y_Q auch einen **Kegelschnitt**.

Beispiel 8.17.
$\mathbb{K} = \mathbb{R}, n = 2$:

(a) $Q = x_1^2 + 2x_2^2 - 1$; (Ellipse)

(b) $Q = x_1 x_2 - 1$; (Hyperbel)

(c) $Q = x_1^2 - x_2$; (Parabel)

$n \geq 3$:

(a) $Q = \left(\frac{x_1}{a_1}\right)^2 + \ldots + \left(\frac{x_n}{a_n}\right)^2 - 1$. (n-dimensionales Ellipsoid)

8.18.

Wir können Q schreiben als $Q = x^t \cdot A \cdot x + 2u^t \cdot x + a$ mit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, A = (a_{jk})_{1 \leq j,k \leq n} \in \mathrm{Sym}_n(\mathbb{K}), u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{K}^n$. Man kann Q auch durch eine einzige Matrix ausdrücken:

$$Q = (1, x_1, \ldots, x_n) \begin{pmatrix} a & u_1 & \ldots & u_n \\ u_1 & a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ u_n & a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = (x')^t \cdot A' \cdot x' \text{ mit } x' := \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}, A' := \left(\begin{array}{c|c} a & u^t \\ \hline u & A \end{array}\right) \in \mathrm{Sym}_{n+1}(\mathbb{K}).$$

Man nennt A' die **erweiterte Matrix des quadratischen Polynoms** Q.

Setze ab jetzt $\mathbb{K} = \mathbb{R}$ voraus.
Ziel: Klassifikation der Quadriken bis auf Affinitäten oder bis auf euklidische Bewegungen.

Definition 8.19.
Seien Q, Q' quadratische Polynome (über $\mathbb{R}$), seien $Y = Y_Q, Y' = Y_{Q'}$ die zugehörigen Quadriken.

(a) Die Polynome Q, Q' heißen **affin (algebraisch) äquivalent**, wenn es $S \in GL_n(\mathbb{R})$ und $w \in \mathbb{R}^n$ gibt mit $Q'(x) = Q(Sx + w)$.
Sie heißen **kartesisch (algebraisch) äquivalent**, wenn dabei $S \in O(n)$ gewählt werden kann.

(b) Die Quadriken Y, Y' heißen **affin (geometrisch) äquivalent**, wenn es eine Affinität f von $\mathbb{R}^n$ gibt mit $f(Y) = Y'$. Kann dabei f als Bewegung gewählt werden, so nenne Y, Y' **kartesisch (geometrisch) äquivalent**.

Beispiele 8.20.

1. Sind Q und $\lambda Q'$, mit $0 \neq \lambda \in \mathbb{R}$, affin äquivalent, so sind auch die Quadriken $Y_Q, Y_{Q'}$ affin äquivalent.
Die Umkehrung ist meist auch richtig, siehe später.

2. Sei $Q = Q(x_1, x_2) = x_1^2 + 4x_2^2 + 2x_1 - 8x_2 + 1$.
Der Kegelschnitt $Y_Q = \{x \in \mathbb{R}^2 : Q(x) = 0\}$ wird durch quadratische Ergänzung vereinfacht:
$Q = (x_1 + 1)^2 + 4(x_2 - 1)^2 - 4$.
Durch die Translation $(x_1, x_2) \mapsto (x_1 + 1, x_2 - 1) =: (y_1, y_2)$ wird also Y_Q zur Ellipse $\{y_1^2 + 4y_2^2 = 4\}$.
Also sind $Y = Y_Q$ und $Y' := \{x_1^2 + 4x_2^2 - 4 = 0\}$ zueinander kartesisch äquivalent.
Durch die Affinität $(x_1, x_2) \mapsto (2(x_1 + 1), x_2 - 1)$ wird $Y = Y_Q$ in den Kreis $\{x_1^2 + x_2^2 = 1\}$ überführt.
Dieser ist offensichtlich zu Y nicht *kartesisch* äquivalent.

Konstruktion 8.21. (algebraische Äquivalenz)
Sei $Q = x^t \cdot A \cdot x + 2u^t \cdot x + a$ ein quadratisches Polynom über $\mathbb{R}$ ($a \in \mathbb{R}, u \in \mathbb{R}^n, A \in \mathrm{Sym}_n(\mathbb{R}), x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$). Wollen durch kartesischen Koordinatenwechsel Q in möglichst einfache Gestalt bringen.
Spektralsatz (6.153) $\Rightarrow \exists\, T \in O(n)$, so dass $TAT^t = TAT^{-1} = D = \mathrm{diag}\,(d_1, \ldots, d_n)$

ist.
Dabei erreiche $d_1, \ldots, d_r \neq 0$, $d_{r+1} = \ldots = d_n = 0$ (mit $0 \leq r \leq n, r = \text{rk}\,(A)$). Dann ist also $(A = T^t D T)$ $Q = (Tx)^t D(Tx) + 2(Tu)^t(Tx) + a$.

In den neuen kartesischen Koordinaten $\xi = Tx = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ wird also, mit $v := Tu$,

$$Q = \xi^t \cdot D \cdot \xi + 2v^t \cdot \xi + a = \sum_{j=1}^{r} d_j \xi_j^2 + 2 \sum_{j=1}^{n} v_j \xi_j + a.$$

Quadratisch ergänzen $\Rightarrow = \sum_{j=1}^{r} (\xi_j + \frac{v_j}{d_j})^2 + 2 \underbrace{\sum_{j=r+1}^{n} v_j \xi_j}_{=\widetilde{v}\xi} + \widetilde{a}$ mit $\widetilde{a} \in \mathbb{R}$. Schreibe $\widetilde{v} :=$

$(0, \ldots, 0, v_{r+1}, \ldots, v_n)^t \in \mathbb{R}^n$.
Ist $\widetilde{v} = 0$, so sind wir damit zufrieden.
Ist $\widetilde{v} \neq 0$, so wähle eine orthogonale Matrix $V \in O(n-r)$ mit erster Zeile $\frac{1}{\|\widetilde{v}\|}(v_{r+1}, \ldots, v_n)$. Setze $U := \begin{pmatrix} I_r & 0 \\ 0 & V \end{pmatrix} \in O(n)$. Für die neuen (kartesischen) Koordinaten $\eta = U\xi = UTx$ gilt dann $\eta_1 = \xi_1, \ldots, \eta_r = \xi_r$ und $\eta_{r+1} = \frac{1}{\|\widetilde{v}\|} \cdot \widetilde{v} \cdot \xi = \frac{1}{\|\widetilde{v}\|} \cdot \sum_{j=r+1}^{n} r_j \xi_j$.

Also $Q = \sum_{j=1}^{r} d_j (\eta_j + \frac{v_j}{d_j})^2 + 2\|\widetilde{v}\| \cdot (\eta_{r+1} + \frac{\widetilde{a}}{2\|\widetilde{v}\|})$.

Wir haben also bewiesen:

Theorem 8.22. *(Hauptachsentransformation, kartesische Form)*
Sei $Q = x^t A x + 2u^t x + a = (x')^t \cdot A' \cdot x'$ ein quadratisches Polynom in $x = (x_1, \ldots, x_n)^t$ über $\mathbb{R}$.
Dann ist Q zu einer der folgenden Normalformen kartesisch äquivalent:

I. $Q = d_1 x_1^2 + \ldots + d_r x_r^2$;

II. $Q = d_1 x_1^2 + \ldots + d_r x_r^2 + c$ *mit* $c \neq 0$;

III. $Q = d_1 x_1^2 + \ldots + d_r x_r^2 + c x_{r+1}$ *mit* $c > 0$.

Jeweils mit $0 \leq r \leq n$ und $d_1, \ldots, d_r \neq 0$.
Dabei ist $r = rk\,(A)$, und $d_1, \ldots, d_r$ sind die von 0 verschiedenen Eigenwerte von A.

Genauere Informationen:

Zusatz 8.23.
Zwei der Normalformen aus 8.22 sind genau dann (algebraisch) äquivalent, wenn sie zum selben Fall (I., II., III.) gehören und die $d_1, \ldots, d_r$ bis auf Permutation dieselben sind (und die c übereinstimmen). Dabei gilt:
Fall I $\Leftrightarrow$ rk (A') = rk (A)
Fall II $\Leftrightarrow$ rk (A') = rk $(A) + 1$
Fall III $\Leftrightarrow$ rk (A') = rk $(A) + 2$.

8.24. Beweis:
Sei weiter $Q = x^tAx + 2u^tx + a = (x')^t \cdot A' \cdot x'$ mit $A' = \begin{pmatrix} c & u^t \\ u & A \end{pmatrix}$. Ist y ein anderes affines Koordinatensystem, etwa $x = Sy + w$ mit $S \in GL_n(\mathbb{R})$ und $w \in \mathbb{R}^n$, so ist $x' = Cy'$ mit $C := \begin{pmatrix} 1 & 0^t \\ w & S \end{pmatrix}$
Also ist $Q = (Cy')^t \cdot A' \cdot (Cy') = (y')^t \cdot B'\ddot{y}'$ mit $B' = C^tA'C = \ldots =: \begin{pmatrix} b & v^t \\ v & B \end{pmatrix}$ mit $B = S^tAS, v = S^t(Aw + u), b = a + 2w^tu + w^tAw$.
Wir sehen daraus insbesondere $B \simeq A$ und $B' \simeq A'$.
Insbesondere ändern sich Rand und Sylvester-Signatur von A, A' *nicht*.

Die erweiterte Matrix zur Normalform aus 8.22 ist

I. $\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & D \end{array}\right)$, (rk $(A) = r$, rk $(A') = r$)

II. $\left(\begin{array}{c|c} c & 0 \\ \hline 0 & D \end{array}\right)$, (rk $(A) = r$, rk $(A') = r + 1$)

III. $\left(\begin{array}{c|c} 0 & \frac{c}{2}e_{r+1} \\ \hline \frac{c}{2}e_{r+1} & D \end{array}\right)$ (rk $(A) = r$, rk $(A') = r + 2$)

mit $D = \text{diag}\,(d_1, \ldots, d_r, 0, \ldots, 0)$.

Damit ist Charakteristik von I - III bewiesen (sogar für affine Äquivalenz).

□

Korollar 8.25. *(Hauptachsentransformation, affine Form)*
Jedes reelle quadratische Polynom Q in $x = (x_1, \ldots, x_n)$ ist zu genau einer der drei Normalformen affin (algebraisch) äquivalent:

I. $Q = x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_r^2$ $(0 \le k \le r \le n)$;

II. $Q = x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_r^2 + c$ $(0 \le k \le r \le n, c \ne 0)$;

III. $Q = x_1^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_r^2 - x_{r+1}$ $(0 \le k \le r \le n)$.

Dabei sind k und r bestimmt durch $r = rk\,(A)$ *und* $k - (r-k) = 2k - r = sign(A)$.
Die Fälle I. - III. sind durch $rk\,(A') - rk\,(A)$ *charakterisiert wie in 8.23.*
(Folgt sofort aus 8.23 und Sylvester'scher Trägheitssatz (b))

Bemerkung 8.26.
Sei $Q = x^tAx + 2u^tx + a$.
Genau dann ist rk $(A') \le 1 + \text{rk}\,(A)$ (d.h., liegt Fall I. oder II. vor), wenn wir $v = 0$ in 8.24 erreichen können, also wenn $\exists\, w \in \mathbb{R}^n$ mit $Aw + u = 0$, also wenn $u \in \text{im}(A)$ ist.
Ist dies der Fall, kann man jedes w mit $Aw + u = 0$ als Translationsvektor verwenden.
Ist $S \in O(n)$ mit $S^tAS = D$ (= Diagonalmatrix), so wird Q unter $x = Sy + w$ zu $Q = y^tDy + b$ mit $b = a + 2w^tu + w^tAw = a + w^tu$.
Ist insbesondere $\det(A) \neq 0$, so gibt es genau ein w mit $Aw + u = 0$, nämlich $w = -A^{-1}u$, und mit $x = Sy + w$ erhalte $Q = y^tDy + b$ mit $b = a - u^tA^{-1}u$.

Beispiel 8.27.
Sei $n = 2$, sei $Q = 41x_1^2 - 24x_1x_2 + 34x_2^2 - 58x_1 - 44x_2 + 26 = x^tAx + 2u^tx + a = (x')A'x'$
mit $A = \begin{pmatrix} 41 & -12 \\ -12 & 34 \end{pmatrix}, u = \begin{pmatrix} -29 \\ -22 \end{pmatrix}, a = 26, A' = \begin{pmatrix} 26 & -29 & -22 \\ -29 & 41 & -12 \\ -22 & -12 & 34 \end{pmatrix}$.
Um Q in kartesische Normalform zu bringen, müssen wir A diagonalisieren.
$P_A(t) = t^2 - 75 \cdot t + 1250 = (t - 50)(t - 25)$.
Man findet $\text{Eig}(A; 25) = \mathbb{R}\begin{pmatrix} 3 \\ 4 \end{pmatrix}, \text{Eig}(A; 50) = \mathbb{R}4{-}3$
Für $S := \frac{1}{5}\begin{pmatrix} 4 & 3 \\ -3 & 4 \end{pmatrix}$ gilt also $S \in SO(2)$ (Drehung, da $\det(S) = 1$), und $S^tAS = \begin{pmatrix} 50 & 0 \\ 0 & 25 \end{pmatrix} =: D$.
Wegen $\det(A) \neq 0$, zeigt 8.26: in den neuen kartesischen Koordinaten $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$
mit $x = Sy - A^{-1}u$, also $y = A^t(x - w)$ mit $w = -A^{-1}u = -\frac{1}{1250}\begin{pmatrix} 34 & 12 \\ 12 & 41 \end{pmatrix} = \ldots = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
hat Q die Form $Q = 50y_1^2 + 25y_2^2 + b$ mit $b = a - u^tA^{-1}u = 26 - 51 = -25$.
Also $Q = 50y_1^2 + 25y_2 2 - 25 = 25(2y_1^2 + y_2 2 - 1)$
mit $y_1 = \frac{1}{5}(4x_1 - 3x_2 - 1), y_2 = \frac{1}{5}(3x_1 + 4x_2 - 7)$.
Y_Q ist eine Ellipse.

S ist eine Drehung um $\theta = -\arccos(\frac{4}{5}) \approx -36,9°$.

Bemerkung 8.28.
Sei $Q = x^t Ax + 2u^t x + a$ sei $x = Sy + w$ ein kartesischer Koordinatenwechsel, welcher Q diagonalisiert (also in Normalform bringt). Die Achsen des Koordinatensystems y heißen die **Hauptachsen** der Quadrik $Y_Q = \{Q = 0\}$.
Es sind dies also die um w verschobenen 1-dimensionalen Eigenräume von A. Sie stehen paarweise aufeinander senkrecht.
Hat A lauter verschiedene Eigenwerte, so sind die Hauptachsen eindeutig bestimmt, andernfalls hängen sie von y ab.

8.29.
Also Illustration geben wir die vollständige Klassifikation der Quadriken für $n = 2$ und $n = 3$. Die Bestimmung des Typs von Y_Q bedeutet die Berechnung der Ränge und Signaturen von A und A'. Dabei muss man $sign(A')$ nur im Fall II. berechnen, da sonst stets $sign(A') = sign(A)$ gilt.
Für $\lambda \neq 0$ definieren Q und λQ dieselbe Quadrik, wir führen sie daher nur einmal auf.
Schreibe $r := \text{rk}\,(A)$. $r' := \text{rk}\,(A'), s := sign(A), s' := sign(A')$.

Fall $n = 2$:

r	r'	s	s'	Fall	Gleichung Q	Quadrik Y_Q
2	3	2	1	II.	$x^2 + y^2 - 1 = 0$	Ellipse
2	3	0	1	II.	$x^2 - y^2 + 1 = 0$	Hyperbel
2	3	2	3	II.	$x^2 + y^2 + 1 = 0$	$\emptyset$
2	2	2		I.	$x^2 + y^2 = 0$	Punkt
2	2	0		I.	$x^2 - y^2 = 0$	2 sich schneidende Geraden
1	3	1		III.	$x^2 - y = 0$	Parabel
				II.	$x^2 - 1 = 0$	2 parallele Geraden
				II.	$x^2 + 1 = 0$	$\emptyset$
				I.	$x^2 = 0$	Doppelte Gerade
				III.	$x = 0$	Gerade
				II.	$1 = 0$	$\emptyset$
				I.	$0 = 0$	$\mathbb{R}^2$

Fall $n = 3$:

r	r'	s	s'	Fall	Gleichung Q	Quadrik Y_Q
3	4	3	2	II.	$x^2 + y^2 + z^2 - 1 = 0$	Ellipsoid
3	4	1	0	II.	$x^2 + y^2 - z^2 - 1 = 0$	1-schaliges Hyperboloid
3	4	1	2	II.	$x^2 + y^2 - z^2 + 1 = 0$	2-schaliges Hyperboloid
3	4	3	4	II.	$x^2 + y^2 + z^2 + 1 = 0$	$\emptyset$
3	3	3		I.	$x^2 + y^2 + z^2 = 0$	Punkt
3	3	1		I.	$x^2 + y^2 - z^2 = 0$	Kegel
				III.	$x^2 + y^2 - z = 0$	elliptisches Paraboloid
				III.	$x^2 - y^2 - z = 0$	hyperbolisches Paraboloid
				II.	$x^2 + y^2 - 1 = 0$	elliptischer Zylinder
				II.	$x^2 - y^2 + 1 = 0$	hyperbolischer Zylinder
				II.	$x^2 + y^2 + 1 = 0$	$\emptyset$
				I.	$x^2 + y^2 = 0$	Gerade
				I.	$x^2 - y^2 = 0$	2 sich schneidende Ebenen
				I.	$x^2 - y = 0$	parabolischer Zylinder
				I.	$x^2 - 1 = 0$	2 parallele Ebenen
				I.	$x^2 + 1 = 0$	$\emptyset$
				I.	$x^2 = 0$	Doppelte Ebene
				I.	$x = 0$	Ebene
				I.	$1 = 0$	$\emptyset$
				I.	$0 = 0$	$\mathbb{R}^3$