Skript Lineare Algebra & Geometrie 2, Hertrich-Jeromin

Studierendenmitschrift

13. Juni 2016

Inhaltsverzeichnis

4 Volumenmessung

4.3 Polynome & Polynomfunktionen

Warum? (Vielleicht eher „Algebra“ – allgemein – als „lineare“ Algebra) Wichtig: das charakteristische Polynom eines Endomorphismus – wichtiges Hilfsmittel im Kontext der Struktursätze.

Beispiel Wir definieren Polynomfunktionen $p, q : K \to K$ eines Körpers K in sich durch

$$p : K \to K,\ x \mapsto p(x) := 1 + x + x^2$$
$$q : K \to K,\ x \mapsto q(x) := 1$$

Falls $K = \mathbb{Z}_2$ so gilt dann

$$\forall x \in K : x(x+1) = 0$$
$$\Rightarrow \forall x \in K : p(x) = q(x)$$

d.h., unterschiedliche „Polynome“ liefern die gleiche Polynomfunktion: Koeffizientenvergleich funktioniert nicht.

Wiederholung Auf dem Folgenraum $K^{\mathbb{N}}$ betrachten wir die Familie $(e_k)_{k\in\mathbb{N}}$ mit

$$e_k : \mathbb{N} \to K,\ j \mapsto e_k(j) := \delta_{jk}$$

Wir wissen: $(e_k)_{k\in\mathbb{N}}$ ist linear unabhängig, aber kein Erzeugendensystem:

$$\forall k \in \mathbb{N} : e_k \notin [(e_j)_{j\neq k}] \text{ und } [(e_j)_{j\in\mathbb{N}}] \neq K^{\mathbb{N}}$$

Insbesondere gilt:

$$\forall x \in [(e_j)_{j\in\mathbb{N}}]\ \exists n \in \mathbb{N}\ \forall k > n : x_k = 0$$

4.3.1 Idee & Definition

Wir fassen ein Polynom als (endliche) Koeffizientenfolge auf,

$$\sum_{k=0}^{n} t^k a_k \cong \sum_{k\in\mathbb{N}} e_k a_k \text{ mit } a_k = 0 \text{ für } k > n$$

und führen darauf das *Cauchyprodukt* (vgl. Analysis) als Multiplikation ein:

$$(a_k)_{k\in\mathbb{N}} \odot (b_k)_{k\in\mathbb{N}} := (c_k)_{k\in\mathbb{N}}$$

wobei

$$c_k := \sum_{j=0}^{k} a_j b_{k-j}.$$

Insbesondere gilt damit

$$\forall j,k \in \mathbb{N} : e_j \odot e_k = e_{j+k} \Rightarrow \forall k \in \mathbb{N} : \begin{cases} e_0 \odot e_k = e_k \\ e_1^k = \underbrace{e_1 \odot \cdots \odot e_1}_{k \text{ mal}} = e_k \end{cases}$$

Mit $1 := e_0,\ t := e_1$ und $t^0 := 1$, wie üblich, liefert dies:

$$\sum_{k=0}^{n} t^k a_k = \sum_{k\in\mathbb{N}} e_k a_k \in [(e_k)_{k\in\mathbb{N}}] \subset K^{\mathbb{N}}$$

4.3.2 Definition

$$K[t] := ([(e_k)_{k\in\mathbb{N}}], \odot),$$

mit dem Cauchyprodukt $\odot$, ist die *Polynomalgebra* über dem Körper K; die Elemente von $K[t]$,

$$p(t) = \sum_{k=0}^{n} t^k a_k = \sum_{k\in\mathbb{N}} e_k a_k,$$

heißen *Polynome in der Variablen* $t := e_1$. Der *Grad* eines Polynoms ist

$$\deg \sum_{k=0}^{n} t^k a_k := \max\{k \in \mathbb{N} \mid a_k \neq 0\} \quad (\text{bzw. } \deg 0 := -\infty)$$

Ist (der „höchste“ Koeffizient) $a_n = 1$ für $\deg p(t) = n$, so heißt das Polynom $p(t)$ *normiert.*

Notation Mit $t^k = e_k$, also $K[t] = [(e_k)_{k\in\mathbb{N}}]$ wird das Cauchyprodukt auf $K[t]$ eine „normale“ Multiplikation, gefolgt von einer Sortierung nach den Potenzen der Variablen t. Wir werden das „$\odot$“ daher oft unterdrücken, und z.B. $p(t)q(t)$ schreiben, anstelle von $p(t) \odot q(t)$.

Bemerkung (Koeffizientenvergleich) Mit dieser Definition von „Polynom“ gilt

$$p(t) = \sum_{k=0}^{n} t^k a_k = 0 \quad \Rightarrow \forall k \in \mathbb{N} : a_k = 0,$$

da $(t^k)_{k\in\mathbb{N}} = (e_k)_{k\in\mathbb{N}}$ linear unabhängig ist. Koeffizientenvergleich funktioniert!

Bemerkung Die Polynomalgebra $K[t]$ über K ist eine assoziative und kommutative K-Algebra, weiters ist $K[t]$ unitär mit Einselement $1 = e_0$.

4.3.3 Definition

Eine K-Algebra ist ein K-VR mit einer *bilinearen Abbildung*,

$$\odot : V \times V \to V,\ (v, w) \mapsto v \odot w,$$

d.h. es gilt

(i) $\forall w \in V :\ V \ni v \mapsto v \odot w \in V$ ist linear;

(ii) $\forall v \in V : V \ni w \mapsto v \odot w \in V$ ist linear.

Eine K-Algebra heißt

- unitär (mit Einselement 1), falls $\exists 1 \in V^\times \forall v \in V : 1 \odot v = v \odot 1 = v;$
- assoziativ, falls $\forall u, v, w \in V : (u \odot v) \odot w = u \odot (v \odot w);$
- kommutativ, falls $\forall v, w, \in V : v \odot w = w \odot v.$

Beispiel Die additive Gruppe $\mathrm{End}(V)$ ist (mit Komposition) eine unitäre assoziative Algebra.

Bemerkung In jeder Algebra $(V, \odot)$ gilt:

$$\forall v \in V : 0 \odot v = v \odot 0 = 0$$

da z.B. für $v \in V$ gilt

$$v \odot 0 = v \odot (0 + 0) = v \odot 0 + v \odot 0 \ \Rightarrow\ 0 = v \odot 0$$

Ist $(V, \odot)$ unitär, so liefert $[1] \subset V$ wegen $1 \odot 1 = 1$ einen Körper:

$$([1], + |_{[1]\times[1]}, \odot |_{[1]\times[1]}) \cong K$$

vermöge $K \ni x \mapsto 1 \cdot x \in [1]$ (siehe Aufgabe 5).

4.3.4 Definition

Ein *Algebra-Homomorphismus* zwischen K-Algebren $(V, \odot)$ und $(W, *)$ ist eine lineare Abbildung $\psi \in \mathrm{Hom}(V, W)$, für die gilt:

$$\forall v, v' \in V : \psi(v \odot v') = \psi(v) * \psi(v').$$

Bemerkung $\mathrm{Hom}(V, W)$ wird oft auch für den (Vektor-)Raum der Algebra-Homomorphismen verwendet. In dieser LVA bedeutet „$\mathrm{Hom}(V, W)$“ immer VR-Homomorphismen, bei allen „anderen“ Homomorphismen wird erwähnt, was gemeint ist.

4.3.5 Einsetzungssatz & Definitionen

Seien $(V, \odot)$ eine unitäre assoziative Algebra und $v \in V$. Dann ist

$$\psi_v : K[t] \to V,\ \sum_{k=0}^{n} t^k a_k = p(t) \mapsto \psi_v(p(t)) := \sum_{k=0}^{n} v^k a_k$$

– wobei $v^0 = 1$ sinnvoll ist, da die Algebra unitär ist – ein Algebra-Homomorphismus; ψ_v heißt *Einsetzungshomomorphismus*.

$$p : V \to V,\ v \mapsto p(v) := \psi_v(p(t))$$

heißt die zu $p(t) \in K[t]$ gehörige *Polynomfunktion* auf V.

Bemerkung Wie üblich: $v^k := \underbrace{v \odot \cdots \odot v}_{k-\mathrm{mal}}$ und $v^0 := 1$.

Beweis

1. ψ_v ist linear:
   - für $p(t) = \sum_{k\in\mathbb{N}} t^k a_k$ und $a \in K$ gilt:
   $$\psi_v(p(t)a) = \psi_v\Big(\sum_{k\in\mathbb{N}} t^k a_k a\Big) = \sum_{k\in\mathbb{N}} v^k a_k a = \psi_v(p(t))a;$$
   - für $p(t) = \sum_{k\in\mathbb{N}} t^k a_k$ und $q(t) = \sum_{k\in\mathbb{N}} t^k b_k$ gilt:
   $$\psi_v(p(t) + q(t)) = \psi_v\Big(\sum_{k\in\mathbb{N}} t^k (a_k + b_k)\Big) = \sum_{k\in\mathbb{N}} v^k (a_k + b_k) = \psi_v(p(t)) + \psi_v(q(t))$$
2. ψ_v ist „multiplikativ“, d.h. verträglich mit der beteiligten Multiplikation:

Für die Vektoren der Basis $(t^k)_{k\in\mathbb{N}}$ von $K[t]$ gilt, da $(V,\odot)$ assoziativ ist,

$$\psi_v(t^m t^n) = \psi_v(t^{m+n}) = v^{m+n} = v^m \odot v^n = \psi_v(t^m) \odot \psi_v(t^n).$$

Da aber ψ_v linear und die Multiplikation in $K[t]$ und in $(V,\odot)$ bilinear sind, folgt die Behauptung.

Bemerkung (Fortsetzungssatz für bilineare Abbildungen) Im Beweis haben wir verwendet: Die Abbildungen

$$K[t] \times K[t] \to V,\ \ (p(t), q(t)) \mapsto \begin{cases} \psi_v(p(t)q(t)) & \text{(Cauchyprodukt)} \\ \psi_v(p(t)) \odot \psi_v(q(t)) & \text{(Produkt in } (V,\odot)\text{)} \end{cases}$$

sind bilinear (da ψ_v linear ist), sind also gleich, sobald sie auf einer Basis übereinstimmen. Dies ist die Eindeutigkeit eines Fortsetzungssatzes für bilineare Abbildungen:

Sind V, W K-VR, $(b_i)_{i\in I}$ eine Basis von V und $(\beta_{ij})_{i,j\in I}$ eine Familie in W, so gibt es eine eindeutige bilineare Abbildung

$$\beta : V \times V \to W$$

mit

$$\forall i,j \in I : \beta(b_i, b_j) = \beta_{ij}$$

Dieser Fortsetzungssatz folgt direkt aus dem Fortsetzungssatz für lineare Abbildungen, da

$$\{\beta : V \times V \to W \text{ bilinear}\} \cong \operatorname{Hom}(V, \operatorname{Hom}(V,W))$$

vermittels des Isomorphismus

$$\beta \mapsto (v \mapsto \underbrace{\beta(v,.)}_{\in \operatorname{Hom}(V,W)}),$$

d.h. durch Nacheinandereinsetzen der Argumente.

Bemerkung Die Abbildung eines Polynoms auf seine Polynomfunktion auf dem Körper,

$$K[t] \ni p(t) \mapsto (x \mapsto p(x)) = \psi_x(p(t)) \in K^K$$

ist für $\operatorname{Char} K \neq 0$ nicht injektiv[1]. Das heißt: Koeffizientenvergleich (für Polynomfunktionen) kann nur funktionieren, wenn $\operatorname{Char} K = 0$.

[1] sonst wäre K^K unendlich dimensional.

Beispiel & Bemerkung Ist V K-VR, so ist $\mathrm{End}(V)$ eine K-Algebra (mit Komposition $\circ$). Man erhält also für $f \in \mathrm{End}(V)$ einen Einsetzungshomomorphismus

$$\psi_f : K[t] \to \mathrm{End}(V),\ p(t) \mapsto \psi_f(p(t)) = p(f);$$

und für jedes Polynom $p(t) \in K[t]$ eine zugehörige Polynomfunktion

$$p : \mathrm{End}(V) \to \mathrm{End}(V),\ f \mapsto \psi_f(p(t)) = p(f).$$

Dieses Beispiel ist der Schlüssel zum Satz von Cayley-Hamilton (im nächsten Abschnitt).

4.3.6 Lemma

Für Polynome $p(t), q(t) \in K[t]$ gilt:

- $\deg p(t) \odot q(t) = \deg p(t) + \deg q(t)$,
- $\deg p(t) + q(t) \leq \max\{\deg p(t), \deg q(t)\}$.

Beweis Für $p(t) = \sum_{k\in\mathbb{N}} t^k a_k$ und $q(t) = \sum_{k\in\mathbb{N}} t^k b_k$ ist

$$p(t) \odot q(t) = \sum_{k\in\mathbb{N}} t^k c_k \text{ mit } c_k = \sum_{j=0}^{k} a_j b_{k-j}$$

Gilt nun $\deg p(t) = n$ und $\deg q(t) = m$, d.h.

$$a_n, b_m \neq 0 \wedge \forall k > n, k' > m : a_k = b_{k'} = 0$$

so folgt

$$\left.\begin{aligned} \forall k > m+n : c_k = 0 \\ c_{m+n} = a_n b_m \end{aligned}\right\} \Rightarrow \deg p(t) \odot q(t) = m+n$$

Gilt andererseits $\deg p(t) = -\infty$ oder $\deg q(t) = -\infty$, also $p(t) = 0 \vee q(t) = 0$, so folgt

$$p(t) \odot q(t) = 0 \Rightarrow \deg p(t) \odot q(t) = -\infty.$$

Die zweite Behauptung ist offensichtlich wahr.

Beispiel Für $p(t), q(t), d(t) \in K[t]$ mit $d(t) \neq 0$ gilt

$$d(t)p(t) = d(t)q(t) \Rightarrow p(t) = q(t).$$

Nämlich: da $\deg d(t) \geq 0$,

$$\begin{aligned}
-\infty &= \deg d(t)\big(p(t) - q(t)\big) \\
&= \deg d(t) + \deg\big(p(t) - q(t)\big) \\
\Rightarrow \deg\big(p(t) - q(t)\big) &= -\infty \\
\Rightarrow p(t) &= q(t)
\end{aligned}$$

4.3.7 Euklidischer Divisionsalgorithmus

Seien $p(t), d(t) \in K[t],\ d(t) \neq 0$. Dann existieren eindeutig $q(t), r(t) \in K[t]$, sodass

$$p(t) = d(t)q(t) + r(t) \text{ und } \deg r(t) < \deg d(t).$$

Bemerkung Ist $\deg p(t) \leq \deg d(t)$, so ist die Aussage trivial.

Beweis Eindeutigkeit folgt wie im Beispiel; mit

$$\begin{aligned}
p(t) &= \begin{cases} d(t)q(t) + r(t) \\ d(t)\tilde{q}(t) + \tilde{r}(t) \end{cases} \\
\Rightarrow d(t)\big(q(t) - \tilde{q}(t)\big) &= \tilde{r}(t) - r(t)
\end{aligned}$$

erhält man

$$\begin{aligned}
\deg d(t) + \deg\big(q(t) - \tilde{q}(t)\big) &= \deg(r(t) - \tilde{r}(t)) \\
&\leq \max\{\deg r(t), \deg \tilde{r}(t)\} < \deg d(t).
\end{aligned}$$

Also folgt

$$\deg\big(q(t) - \tilde{q}(t)\big) = -\infty \Rightarrow \deg(r(t) - \tilde{r}(t)) = \deg d(t) - \infty = -\infty$$

und damit

$$\tilde{q}(t) = q(t) \text{ und } \tilde{r}(t) = r(t).$$

Existenz: Mit $k := \deg d(t) \geq 0$ und

$$K[t]_m := \{q(t) \in K[t] \mid \deg q(t) \leq m\} \text{ für } m \in \mathbb{N}$$

betrachte man die Abbildung

$$K[t]_m \times K[t]_{k-1} \to K[t]_{k+m},\ \big(q(t), r(t)\big) \mapsto d(t)q(t) + r(t).$$

Diese Abbildung ist linear (klar) und injektiv, denn: ist $q(t) \neq 0$, so folgt wegen

$$\deg r(t) < k = \deg d(t) \leq \deg d(t)q(t)$$

dass

$$\begin{aligned} \deg\left(d(t)q(t) + r(t)\right) &= \deg d(t)q(t) \geq k > -\infty \\ &\Rightarrow d(t)q(t) + r(t) \neq 0, \end{aligned}$$

also

$$d(t)q(t) + r(t) = 0 \Rightarrow q(t) = 0 \wedge r(t) = 0.$$

Wegen

$$\dim K[t]_m \times K[t]_[k-1] = (m+1) + k = (k+m) + 1 = \dim K[t]_{k+m}$$

liefert diese Abbildung dann für jedes $m \in \mathbb{N}$ einen Isomorphismus

$$K[t]_m \times K[t]_{k-1} \to K[t]_{k+m}$$

4.3.8 Korollar & Definition

Sei $p(t) \in K[t]$ mit $\deg p(t) \geq 1$. Ist $x \in K$ eine *Nullstelle* von $p(t)$, d.h.

$$p(x) = \psi_x(p(t)) = 0,$$

so folgt

$$\exists! q(t) \in K[t] : p(t) = (t-x)q(t)$$

Beweis Seien $p(t) \in K[t]$ mit $\deg p(t) \geq 1$ und $x \in K$ eine Nullstelle von $p(t)$; dann gibt es eindeutig $q(t), r(t) \in K[t]$ mit

$$p(t) = (t-x)q(t) + r(t) \text{ und } \deg r(t) < \deg(t-x) = 1,$$

also

$$p(t) = (t-x)q(t) + r(t) = (t-x)q(t) + c_0.$$

Einsetzen von $x \in K$ liefert dann

$$0 = p(x) = (x-x)q(x) + c_0 = c_0.$$

Bemerkung und Beispiel Dies liefert eine Methode, um Polynome zu *faktorisieren*: Für jede gefundene Nullstelle kann man einen *Linearfaktor* abspalten.

$$p(t) = t^4 - t^3 + t^2 - t = \begin{cases} t(t-1)(t-i)(t+i) \in \mathbb{C}[t] \\ t(t-1)(t^2+1) \in \mathbb{R}[t]. \end{cases}$$

4.3.9 Mehr zu Polynomen

Dies ist der Anfang einer der Teilbarkeitstheorie der natürlichen Zahlen ähnlichen Theorie für Polynome.

Sind $p(t), d(t) \in K[t]$, so heißt $d(t)$ Teiler *von $p(t), d(t) \mid p(t)$, falls*

$$\exists q(t) \in K[t] : p(t) = d(t)q(t).$$

Primpolynome Nennt man $p(t) \in K[t]$ mit $\deg p(t) > 0$ ein *Primpolynom* (oder *irreduzibel*), falls für $d(t), q(t) \in K[t]$ gilt:

$$p(t) = d(t)q(t) \Rightarrow \Big(\deg q(t) = 0 \vee \deg d(t) = 0\Big),$$

so gilt der Satz über die *Primfaktorzerlegung*:

Jedes Polynom $p(t) \in K[t]$ mit $\deg p(t) > 0$ zerfällt eindeutig in Primpolynome,

$$p(t) = a_n p_1(t) \cdots p_m(t),$$

wobei $a_n \in K$ und $p_1(t), \ldots, p_m(t) \in K[t]$ normierte Primpolynome sind.

Beweis Existenz ist einfach zu zeigen (Induktion über n), die weniger leicht zu zeigende Eindeutigkeit benutzt die Existenz des *größten gemeinsamen Teilers* $d(t) = \mathrm{ggT}\,(p(t), q(t))$ zweier Polynome $p(t)$ und $q(t)$:

Zu $p(t), q(t) \in K[t] \setminus \{0\}$ gibt es genau ein normiertes Polynom $d(t) \in K[t]$ mit

$$\begin{aligned} &d(t) \mid p(t) \wedge d(t) \mid q(t) \text{ und} \\ &d'(t) \mid p(t) \wedge d'(t) \mid q(t) \Rightarrow d'(t) \mid d(t). \end{aligned}$$

Lemma von Bézout Für den ggT gilt auch das Lemma von Bézout:

$$\exists p'(t), q'(t) \in K[t] : d(t) = p(t)p'(t) + q(t)q'(t)$$

Bemerkung Aus der Gradformel,

$$\deg d(t)q(t) = \deg d(t) + \deg q(t)$$

folgt direkt:

Jedes Polynom $p(t) \in K[t]$ mit $\deg p(t) = 1$ ist Primpolynom.

Fundamentalsatz der Algebra Falls $K = \mathbb{C}$, so sind die Polynome mit Grad 1 die einzigen Primpolynome:

In $\mathbb{C}$ zerfällt jedes Polynom (mit Grad ≥ 1) in Linearfaktoren;

$$\forall p(t) \in \mathbb{C}[t],\ \deg \geq 1 :\ \exists x_1, \ldots, x_n \in \mathbb{C}$$

mit

$$p(t) = a_n \prod_{j=1}^{n} (t - x_j).$$

Ist $K = \mathbb{R}$, so ist dies nicht der Fall; ein Primpolynom vom Grad $\deg p(t) = 2$ ist z.B.

$$p(t) = t^2 + 1 \ \in \mathbb{R}[t],$$

denn

$$t^2 + 1 = (t - x_1)(t - x_2) \Rightarrow \begin{cases} 0 = x_1 + x_2 \\ 1 = x_1 \cdot x_2 \end{cases} \Rightarrow 1 = -x^2$$

Andererseits ist $p(t) \in \mathbb{R}[t] \subset \mathbb{C}[t]$, also existieren $x_1, \ldots, x_n \in \mathbb{C}$ mit

$$a_n \prod_{j=1}^{n} (t - x_j) = p(t) = \overline{p(t)} = \overline{a_n} \prod_{j=1}^{n} (t - \overline{x_j}),$$

d.h. mit der Eindeutigkeit der Primfaktorzerlegung, $a_n \in \mathbb{R}$ und die x_j sind entweder reell oder treten in komplex-konjugierten Paaren auf:

$$p(t) = a_n \prod_{j=1}^{m} (t^2 - t(x_j + \overline{x_j}) + x_j \overline{x_j}) \prod_{j=2m+1}^{n} (t - x_j).$$

Ist also $p(t) \in \mathbb{R}[t]$ Primpolynom, so folgt $\deg p(t) \leq 2$ und

$$\deg p(t) = 2 \Rightarrow \exists x, y \in \mathbb{R} : p(t) = (t - x)^2 + y^2 \text{ mit } y \neq 0$$

In $K = \mathbb{Q}$ gibt es noch „mehr“ Primpolynome, wie z.B.:

$$p(t) = t^2 - 2 \text{ oder } p(t) = t^4 + 1$$

4.4 Das charakteristische Polynom

4.4.1 Definition

Seien V ein K-VR und $f \in \mathrm{End}(V)$. Dann heißen

(i) $x \in K$ ein Eigenwert von f, falls $\exists v \in V^\times : f(v) = vx$;

(ii) $v \in V^\times$ ein Eigenvektor von f, falls $\exists x \in K : f(v) = vx$;

(iii) $\ker(f - \mathrm{id}_V\, x) \subset V$ ein Eigenraum, falls $\ker(f - \mathrm{id}_V\, x) \neq \{0\}$.

Bemerkung Der Skalar $x \in K$ ist genau dann ein Eigenwert von $f \in \mathrm{End}(V)$, wenn $\ker(f - \mathrm{id}_V\, x) \neq \{0\}$, d.h., wenn ein Eigenvektor $v \in V^\times$ zu x existiert.

Beispiel Für $\frac{d}{ds} \in \mathrm{End}(C^\infty(\mathbb{R}))$ ist jedes $x \in \mathbb{R}$ ein Eigenwert, da

$$\Big(\frac{d}{ds} - \mathrm{id}_V\, x\Big)v = 0 \text{ für } v : \mathbb{R} \to \mathbb{R}, s \mapsto v(s) := e^{xs},$$

wobei $v \in C^\infty(\mathbb{R}) \setminus \{0\}$, d.h. $s \mapsto v(s) = e^{xs}$ ist ein Eigenvektor zum Eigenwert $x \in \mathbb{R}$.

Beispiel Ist $\dim V < \infty$, so kann die Determinante zur Bestimmung von Eigenwerten von Endomorphismen $f \in \mathrm{End}(V)$ benutzt werden, da

$$\ker(f - \mathrm{id}_V\, x) \neq \{0\} \Leftrightarrow (f - \mathrm{id}_V\, x) \text{ nicht injektiv} \Leftrightarrow \det(f - \mathrm{id}_V\, x) = 0,$$

d.h. das Auffinden von Eigenwerten $x \in K$ von f ist reduziert auf die Bestimmung der Nullstellen der Funktion

$$K \ni x \mapsto \det(f - \mathrm{id}_V\, x) \in K.$$

Beispiel Ist z.B. (b_1, b_2) Basis von V und $f \in \mathrm{End}(V)$ durch $f(B) = BX$ gegeben, so liefern die Nullstellen der Polynomfunktion

$$\det(f - \mathrm{id}_V\, x) = \det(X - E_2 x) = \det \begin{pmatrix} x_{11} - x & x_{12} \\ x_{21} & x_{22} - x \end{pmatrix}$$
$$= (x_{11} - x)(x_{22} - x) - x_{12}x_{21} = x^2 - x(x_{11} + x_{22}) + (x_{11}x_{22} - x_{12}x_{21})$$

die Eigenwerte von f – beispielsweise erhalten wir für

$$X = \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} : \ \det(f - \mathrm{id}_V\, x) = x^2 - 2x - 3 = (x+1)(x-3),$$

also Eigenwerte $x_1 = -1$ und $x_2 = 3$ mit zugehörigen Eigenvektoren als Lösungen von

$$v_i \in \ker(f - \mathrm{id}_V\, x_i),$$

also durch Lösungen der linearen Gleichungssysteme

$$\begin{pmatrix} 2-(-1) & 3 \\ 1 & -(-1) \end{pmatrix} \begin{pmatrix} v_1^1 \\ v_1^2 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1^1 \\ v_1^2 \end{pmatrix} \text{ und}$$

$$\begin{pmatrix} 2-3 & 3 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} v_2^1 \\ v_2^2 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} v_2^1 \\ v_2^2 \end{pmatrix}$$

sodass

$$v_1 = b_1 - b_2 \text{ und } v_2 = b_1 3 + b_2$$

Eigenvektoren zu den Eigenwerten x_1, x_2 liefert.

Rechenbeispiel 1 Für $X = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$ erhält man

$$\det(f - \mathrm{id}_V\, x) = \det \begin{pmatrix} 2-x & -1 \\ 1 & -x \end{pmatrix} = x^2 - 2x + 1$$

und Eigenvektoren zum Eigenwert $x = 1$ durch Lösung der LGS

$$\begin{pmatrix} 2-1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1^1 \\ v_1^2 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1^1 \\ v_1^2 \end{pmatrix}$$

d.h. der Eigenraum zum Eigenwert x,

$$\ker(f - \mathrm{id}_V) = [\{b_1 + b_2\}] \quad \text{hat} \quad \dim\ker(f - \mathrm{id}_V) < \dim V.$$

Rechenbeispiel 2 Ist $K = \mathbb{R}$ und

$$\det(f - \mathrm{id}_V\, x) = x^2 + 1,$$

so hat f keine Eigenwerte: z.B., wenn $X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

4.4.2 Definition

Sei V ein K-VR, für $f \in \mathrm{End}(V)$ ist das *charakteristische Polynom* von f:

$$\chi_f(t) := \det(\mathrm{id}_V\, t - f) \in K[t].$$

Analog definiert man für $X \in K^{n\times n}$ das charakteristische Polynom

$$\chi_X(t) := \det(E_n t - X) \in K[t].$$

Bemerkung Oft wird auch das andere Vorzeichen in der Determinante verwendet, also $\det(f - \mathrm{id}_V t)$ bzw. $\det(X - E_n t)$.

Bemerkung *Diese Definition ist erklärungsbedürftig!*

Da $t \notin K$ ist $\mathrm{id}_V t - f \notin \mathrm{End}(V)$, sondern $\mathrm{id}_V t - f \in \mathrm{End}(V)[t]$. Zwei Lösungsstrategien bieten sich an:

1. Erweiterung der Determinante auf $\mathrm{End}(V)[t]$.
2. Benutzung von Darstellungsmatrizen.

Beide führen schließlich zur Leibniz-Formel:

Ist B eine Basis von V und $\xi_B^B(f) = X = (x_{ij})_{i,j\in\{1,\ldots,n\}}$, so erhält man

$$\chi_f(t) = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) \prod_{j=1}^{n} \underbrace{\left(\delta_{\sigma(j)j} t - x_{\sigma(j)j}\right)}_{\in K[t]} \in K[t].$$

Die Unabhängigkeit von der Basis B folgt aus der Transformationsformel für Darstellungsmatrizen und dem Determinanten-Multiplikationssatz (wie vorher für $\det f = \det \xi_B^B(f)$).

4.4.3 Bemerkung & Definition

Ist $\dim V = n$, so ist $\chi_f(t)$ ein normiertes Polynom vom Grad $\deg(\chi_f(t)) = n$,

$$\chi_f(t) = t^n - t^{n-1}\,\mathrm{tr}\, f + \cdots + (-1)^n \det f,$$

wobei die *Spur* $\mathrm{tr}\, f$ („tr “$\widehat{=}$ trace) von f durch diese Gleichung (wohl-)defininiert ist. Ist $(x_{ij})_{i,j\in\{1,\ldots,n\}} = X = \xi_B^B(f)$ Darstellungsmatrix von f, so gilt

$$\mathrm{tr}\, f = \sum_{j=1}^{n} x_{jj} = \sum_{j=1}^{n} b_j^* f(b_j).$$

Oft wird $\det(f - \mathrm{id}_V t) = (-1)^n \chi_f(t)$ als charakteristisches Polynom definiert – dieses Polynom ist dann nur für gerade n normiert.

4.4.4 Korollar

Ein $x \in K$ ist genau dann Eigenwert von f, wenn $\chi_f(x) = 0$.

Also: Die Eigenwerte von f sind genau die Nullstellen des charakteristischen Polynoms $\chi_f(t)$.

Beweis Klar – das war die Idee hinter der Definition des charakteristischen Polynoms.

4.4.5 Korollar & Definition

Ist $x \in K$ Eigenwert von $f \in \mathrm{End}(V)$, so ist $(t-x)$ Teiler des charakteristischen Polynoms. Insbesondere gilt:

$$\exists! k \in \mathbb{N}^\times : \begin{cases} (t-x)^k \mid \chi_f(t) \\ (t-x)^{k+1} \nmid \chi_f(t) \end{cases}$$

Diese Zahl k heißt die *algebraische Vielfachheit* von x;

$$g := \mathrm{def}(\mathrm{id}_V\, x - f) \leq k$$

ist die *geometrische Vielfachheit* von x.

Beweis Da x Eigenwert von f ist, ist die Existenz und Eindeutigkeit von k klar. Außerdem gilt analog auch $g \geq 1$. Zu zeigen bleibt: $g \leq k$, d.h. $(t-x)^g \mid \chi_f(t)$:
Für eine Basis $B = (b_1, \ldots, b_n)$ von V mit $\ker(\mathrm{id}_v\, x - f) = [(b_1, \ldots, b_g)]$ hat

$$\xi_B^B(f) = \begin{pmatrix} E_g x & Y \\ 0 & X \end{pmatrix} \text{ mit } Y \in K^{g \times (n-g)}, X \in K^{(n-g)\times(n-g)}$$

Blockgestalt, also ist

$$\chi_f(t) = (t-x)^g \cdot \chi_X(t),$$

d.h. $(t-x)^g \mid \chi_f(t)$, da $(t-x)^{k+1} \nmid \chi_f(t)$, gilt also $g \leq k$.

Beispiel Ist $f \in \mathrm{End}(V)$ wie oben durch $f(B) = BX$ gegeben, so haben die Eigenwerte

$$x_1 = -1 \text{ und } x_2 = 3 \text{ für } X = \begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix}$$

algebraische und geometrische Vielfachheiten

$$1 = g_i = k_i, \text{ da } 1 \leq g_i \leq k_i \text{ und } k_1 + k_2 \leq 2;$$

der Eigenwert

$$x = 1 \text{ für } X = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$$

hat algebraische und geometrische Vielfachheiten $k = 2$ und $g = 1$, da

$$f \neq \mathrm{id}_V\, x = \mathrm{id}_V$$

und $\chi_f(t) = (t-x)^2 \in \mathbb{R}[t]$, da ein quadratisches Polynom zwei (relle oder komplex konjugierte) Nullstellen hat, oder aber eine doppelte reelle.

4.4.6 Definition & Lemma

Das Schlüsselargument im Beweis oben kann man verallgemeinern:
Sei $f \in \mathrm{End}(V)$ und $U \subset V$ ein f*-invarianter Unterraum*, d.h. $f(U) \subset U$. Ist dann $V = U \oplus U'$ eine direkte Zerlegung und $p, p' \in \mathrm{End}(V)$ die zugehörigen Projektionen, so gilt

$$\chi_f(t) = \chi_{f|_U}(t) \cdot \chi_{f'}(t),$$

wobei

$$f' := p' \circ f|_{U'} \in \mathrm{End}(U').$$

Bemerkung Man kann $f|_U$ als Endomorphismus $f|_U \in \mathrm{End}(U)$ auffassen, da $f(U) \subset U$.

Beweis Wie oben: Sei $B = (b_1, \dots, b_n)$ Basis von V, sodass

- $C = (b_1, \dots, b_k)$ Basis von U und
- $C' = (b_{k+1}, \dots, b_n)$ Basis von U' ist.

Die Darstellungsmatrix von f bzgl. B hat dann Blockgestalt,

$$\xi_B^B(f) = \begin{pmatrix} X & Y \\ 0 & X' \end{pmatrix} \text{ mit } X = \xi_C^C(f|_U), X' = \xi_{C'}^{C'}(f')$$

Damit folgt die Behauptung (wie oben) mit der Leibniz-Formel.

Bemerkung Alternativ kann man das Lemma mit der von f induzierten Quotientenabbildung $f' \in \mathrm{End}(V/U)$ formulieren, wobei

$$f' : V/U \to V/U, v + U \mapsto f'(v + U) := f(v) + U.$$

4.4.7 Definition

Ein Endomorphismus $f \in \mathrm{End}(f)$ heißt *diagonalisierbar* bzw. *trigonalisierbar*, falls es eine Basis B von V gibt, sodass $\xi_B^B(f) = (x_{ij})_{i,j \in \{1,\dots,n\}}$ eine Diagonalmatrix

$$i \neq j \Rightarrow x_{ij} = 0$$

bzw. obere Dreiecksmatrix ist,

$$i > j \Rightarrow x_{ij} = 0.$$

Bemerkung Falls $\dim V < \infty$, so ist $f \in \mathrm{End}(V)$ genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von f besitzt. Damit kann man „Diagonalisierbarkeit“ auch im Falle $\dim V = \infty$ definieren.

Bemerkung Ist f trigonalisierbar (oder gar diagonalisierbar), so zerfällt $\chi_f(t)$ in Linearfaktoren: für geeignete $x_1, \ldots, x_n \in K$ ist

$$\chi_f(t) = \prod_{j=1}^{n} (t - x_j).$$

4.4.8 Bemerkung & Definition

Man nennt eine Matrix $X \in K^{n \times n}$ diagonalisierbar (bzw. trigonalisierbar), falls $f_X \in \mathrm{End}(K^n)$ diagonalisierbar (bzw. trigonalisierbar) ist.

Dies ist genau dann der Fall, falls es $P \in Gl(n)$ gibt, sodass PXP^{-1} Diagonalmatrix (bzw. obere Dreiecksmatrix) ist.

4.4.9 Lemma

Frage: Was sind hinreichende Kriterien dafür? Notwendigkeit kennen wir: $\chi_f(t)$ zerfällt in Linearfaktoren.

Eigenvektoren $v_1, \ldots, v_m \in V$ zu paarweise verschiedenen Eigenwerten $x_1, \ldots, x_m$ eines Endomorphismus $f \in \mathrm{End}(V)$ sind linear unabhängig.

Bemerkung Anders gesagt: Die Summe von Eigenräumen zu paarweise verschiedenen Eigenwerten ist direkt.

Beweis Zu zeigen: Ist $\sum_{i=1}^{m} v_i y_i = 0$ für Koeffizienten $y_1, \ldots, y_m \in K$, so folgt $y_1 = \cdots = y_m = 0$.
Seien $y_1, \ldots, y_m \in K$ und $w_i := v_i y_i$ und $w := \sum_{i=1}^{m} w_i = \sum_{i=1}^{m} v_i y_i$. Wiederholte Anwendung von f liefert, wegen $f(w_i) = w_i x_i$

$$(f^{m-1}(w), \ldots, f^2(w), f(w), w) = (w_1, \ldots, w_m) \begin{pmatrix} x_1^{m-1} & \cdots & x_1^2 & x_1 & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ x_m^{m-1} & \cdots & x_m^2 & x_m & 1 \end{pmatrix}$$

mit der Vandermonde-Matrix $X \in Gl(m)$, da

$$\det X = \prod_{i<j} (x_i - x_j) \neq 0$$

weil die Eigenwerte $x_1, \ldots, x_m$ paarweise verschieden sind. Damit folgt aus $w = \sum_{i=1}^m v_i y_i = 0$

$$(w_1, \ldots, w_m) = (f^{m-1}(w), \ldots, f(w), w)X^{-1} = (0, \ldots, 0)$$

also

$$\forall i = 1, \ldots, m : 0 = w_i = v_i y_i \text{ und } v_i \neq 0 \Rightarrow y_i = 0.$$

4.4.10 Satz

Ein Endomorphismus $f \in \mathrm{End}(V)$ ist genau dann diagonalisierbar, wenn $\chi_f(t) \in K[t]$ in Linearfaktoren zerfällt und die algebraischen und geometrischen Vielfachheiten aller Eigenwerte übereinstimmen,

$$\chi_f(t) = \prod_{i=1}^m (t - x_i)^{k_i} \text{ und } \forall i = 1, \ldots, m : k_i = g_i.$$

Beweis Ist f diagonalisierbar, so existiert eine Basis B aus Eigenvektoren von f, also ist dann

$$\xi_B^B(f) = \begin{pmatrix} E_{g_1} x_1 & 0 & \cdots & 0 \\ 0 & E_{g_2} x_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{g_m} x_m \end{pmatrix}$$

Damit ist

$$\chi_f(t) = \prod_{i=1}^m (t - x_i)^{g_i}.$$

Hat andererseits das charakteristische Polynom diese Gestalt, so wähle man in jedem Eigenraum $\ker(\mathrm{id}_V\, x_i - f)$ eine Basis $C_i, i = 1, \ldots, m$. Da Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, und wegen

$$g_1 + \cdots + g_m = k_1 + \cdots + k_m = \dim V$$

liefert $B := \bigcup_{i=1}^m C_i$ eine Basis von V.

4.4.11 Korollar

Ein Endomorphismus $f \in \mathrm{End}(V)$ mit $n = \dim V$ paarweise verschiedenen Eigenwerten ist diagonalisierbar.

Beweis Für die geometrischen und algebraischen Vielfachheiten jedes Eigenwerts gilt

$$1 \leq g_i \leq k_i \text{ und } \sum_{i=1}^n k_i \leq n.$$

Damit folgt

$$\forall i = 1, \dots, n : k_i = 1 \text{ und } \sum_{i=1}^{n} k_i = n,$$

d.h. das charakteristische Polynom zerfällt in Linearfaktoren und $\forall i = 1, \dots, n : k_i = g_i$.

4.4.12 Satz

Ein Endomorpismus $f \in \mathrm{End}(V)$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom in Linearfaktoren zerfällt.

Bemerkung Da Diagonalisierbarkeit bzw. Trigonalisierbarkeit durch die Existenz einer Darstellungsmatrix in spezieller Gestalt definiert wurde, wird in den Charakterisierungen immer (implizit) $\dim V < \infty$ angenommen.

Beweis Wir wissen schon: Ist f trigonalisierbar, so zerfällt $\chi_f(t)$ in Linearfaktoren. Umkehrung: Beweis durch vollständige Induktion über $n = \dim V$.
Für $n = 1$ ist nichts zu zeigen. Sei die Behauptung für $n-1$ bewiesen. Für n folgt dann:
Da $\chi_f(t)$ in Linearfaktoren zerfällt

$$\chi_f(t) = \prod_{i=1}^{n} (t - x_i)$$

für geeignete $x_1, \dots, x_n$, ist x_1 Eigenwert von f. Nun seien

- b_1 ein Eigenvektor zum Eigenwert x_1 und $U := [\{b_1\}]$,
- $U' \subset V$ ein zu U komplementärer Unterraum, und
- $p, p' \in \mathrm{End}(V)$ die zur direkten Zerlegung $V = U \oplus U'$ gehörenden Projektionen,

$$U = p(V) = \ker p' \text{ und } U' = p'(V) = \ker p,$$

- und $f' := p' \circ f|_{U'} \in \mathrm{End}(U')$.

Da $U(\neq \{0\})$ f-invarianter UR von V ist, faktorisiert das charakteristische Polynom

$$\chi_f(t) = \chi_{f|_U}(t) \cdot \chi_{f'}(t) = (t - x_1) \cdot \chi_{f'}(t);$$

also zerfällt $\chi_{f'}(t)$ in Linearfaktoren,

$$\chi_{f'}(t) = \prod_{i=2}^{n} (t - x_i).$$

Nach Induktionsannahme existiert also eine Basis $B' = (b_2, \dots, b_n)$ von U', sodass $\xi_{B'}^{B'}(f)$ obere Dreiecksmatrix ist. Mit $B = (b_1, \dots, b_n)$ als Basis von V gilt dann:

$$\xi_B^B(f) = \begin{pmatrix} x_1 & Y \\ 0 & \xi_{B'}^{B'}(f') \end{pmatrix}$$

ist obere Dreiecksmatrix.

4.5 Der Satz von Cayley-Hamilton

4.5.1 Satz

Für $f \in \operatorname{End}(V)$ gilt $\chi_f(f) = 0$.

Unfug-Beweis Durch direktes Einsetzen erhält man

$$\chi_f(f) = \det(\operatorname{id}_V f - f) = \det 0 = 0.$$

Zum Verständnis des Satzes Ist V ein K-VR mit $n = \dim V < \infty$ und $f \in \operatorname{End}(V)$, so ist

$$\chi_f(t) = \sum_{k=0}^{n} t^k a_k \in K[t]$$

ein (abstraktes) Polynom in der Variablen t $(= e_1 \in K^{\mathbb{N}})$ und der Einsetzungshomomorphismus $\psi_f : K[t] \to \operatorname{End}(V)$ (also ein Algebrahomomophismus) liefert

$$\chi_f(f) = \psi_f\left(\chi_f(t)\right) = \sum_{k=0}^{n} f^k a_k.$$

Der Satz sagt, dass $0 = \chi_f(f) \in \operatorname{End}(V)$, d.h.

$$\forall v \in V : \chi_f(f)(v) = 0.$$

4.5.2 Definition & Lemma

Seien $f \in \operatorname{End}(V)$ und B eine *f-zyklische Basis* von V, d.h. eine Basis der Form

$$B = (b_1, \dots, b_n) = \left(b, f(b), \dots, f^{n-1}(b)\right).$$

Dann existieren $a_0, \dots, a_{n-1} \in K$ mit

$$f^n(b) + \sum_{k=0}^{n-1} f^k(b) a_k = 0,$$

mit diesen Koeffizienten ist

$$\chi_f(t) = t^n + t^{n-1}a_{n-1} + \ldots, +ta_1 + a_0.$$

Bemerkung Im Allgemeinen existiert zu $f \in \mathrm{End}(V)$ keine f-zyklische Basis von V, z.B. für $f = \mathrm{id}_V$ und $\dim V \geq 2$.

Beweis Da $B = (b, f(b), \ldots, f^{n-1}(b))$ eine Basis ist, ist $f^n(b) \in [B]$ und damit existieren die a_k mit

$$0 = f^n(b) + \sum_{k=0}^{n-1} f^k(b)a_k.$$

Damit ist die Darstellungsmatrix von f

$$\xi_B^B(f) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \vdots & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -a_{n-1} \end{pmatrix} =: X$$

und Entwicklung von $\chi_f(t) = \det(E_n t - \xi_B^B(f))$ nach der ersten Zeile(nach Laplaceschem Entwicklungssatz – dieser Satz war „nur" eine Methode, die Terme in der Leibniz-Formel zu sortieren) liefert

$$\det(E_n t - X) = \det \begin{pmatrix} t & 0 & \cdots & 0 & a_0 \\ -1 & t & \vdots & \vdots & a_1 \\ 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & t & \vdots \\ 0 & \cdots & 0 & -1 & t + a_{n-1} \end{pmatrix}$$

$$= t \det \begin{pmatrix} t & 0 & \cdots & 0 & a_1 \\ -1 & t & \vdots & \vdots & a_2 \\ 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & t & \vdots \\ 0 & \cdots & 0 & -1 & t + a_{n-1} \end{pmatrix} + (-1)^{n+1} a_0 \det(X_{1n})$$

$$\overset{\text{mit Ind.}}{=} t\{t^n n - 1 + tn - 2a_{n-1} + \cdots + a_1\} + a_0 = t^n + t^{n-1}a_{n-1} + \ldots, +ta_1 + a_0,$$

wie behauptet.

Beispiel Zur Lösung des reellen *Anfangswertproblems*

$$y'' + 2y' - 3y = 0, \quad \begin{cases} y(0) = 4 \\ y'(0) = 0 \end{cases}$$

schreiben wir dieses als System erster Ordnung mit dem Ansatz $y_1 = y$ und $y_2 = y'$:

Daraus erhält man mit $Y = (y_1, y_2)$

$$\begin{aligned} Y' &= (y_1', y_2') = (y', y'') = (y', -2y' + 3y) \\ &= (y, y') \begin{pmatrix} 0 & 3 \\ 1 & -2 \end{pmatrix} = YX \text{ mit } X = \begin{pmatrix} 0 & 3 \\ 1 & -2 \end{pmatrix}, \end{aligned}$$

d.h. wir suchen eine $\frac{d}{ds}$-zyklische Basis $(y, \frac{d}{ds}y) = (y, y')$ eines 2-dim UVR $[(y, y')] \subset C^\infty(\mathbb{R})$ bezüglich derer $\frac{d}{ds} \in \mathrm{End}(C^\infty(\mathbb{R}))$ Darstellungsmatrix X hat.

Der Ansatz $y(s) = e^{xs}(v_0, v_1)$ reduziert das AWP auf ein Eigenwertproblem.

$$0 = \left(Y' - YX\right)(s) = \left(\frac{d}{ds}Y - YX\right)(s) = \underbrace{e^{xs}(v_0, v_1)}_{Y}\{E_2 x - X\}$$

bzw. (vgl. Abschnitt 3.1) mit dem zur transponierten Matrix X^t assoziierten Endomorphismus $f_{X^t} \in \mathrm{End}(\mathbb{R}^2)$

$$f_{X^t}(v) = vx \text{ für } x \in \mathbb{R} \text{ und } v \in \mathbb{R}^2.$$

Nach obigem Lemma sind die Eigenwerte Lösungen der Gleichung

$$0 = \chi_{X^t}(x) = \chi_X(x) \overset{\text{Lemma}}{=} x^2 + 2x - 3 = (x-1)(x+3).$$

Also sind $x_1 = 1$ und $x_2 = -3$ die Eigenwerte; zugehörige Eigenvektoren erhält man als Lösungen der linearen Gleichungssysteme

$$(0,0) = (v_0, v_1)(E_2 x_i - X) = (v_0, v_1)\begin{pmatrix} x_i & -3 \\ -1 & x_i + 2 \end{pmatrix} = \begin{cases} (v_0, v_1)\begin{pmatrix} 1 & -3 \\ -1 & 3 \end{pmatrix} & \text{für } i = 1 \\ (v_0, v_1)\begin{pmatrix} -3 & -3 \\ -1 & -1 \end{pmatrix} & \text{für } i = 2 \end{cases}$$

Damit bekommt man Eigenvektoren $(v_0, v_1) = (1, 1)$ zum Eigenwert $x = 1$ und $(v_0, v_1) = (1, -3)$ zum Eigenwert $x = -3$.

Die allgemeine, durch *Superposition* (Linearkombination) erhaltene Lösung der Differentialgleichung ist also

$$s \mapsto Y(s) = e^s(1,1)c_1 + e^{-3s}(1,-3)c_2$$

mit Koeffizienten $c_1, c_2 \in \mathbb{R}$. Abgleich der „Integrationskonstanten“ c_1 und c_2 mit den Anfangsbedingungen liefert dann die Lösung

$$s \mapsto y(s) = 3e^s + e^{-3s}.$$

Bemerkung Man bemerke: (y, y') ist linear unabhängig für die Lösung, ist also tatsächlich $\frac{d}{ds}$-zyklische Basis eines 2-dim URs $[(y, y')] \subset C^\infty(\mathbb{R})$ – obwohl die den gleichen Raum aufspannenden „Basislösungen“

$$s \mapsto e^s \text{ und } s \mapsto e^{-3s}$$

keine $\frac{d}{ds}$-zyklischen Basen erzeugen, da sie lineare Differentialgleichungen erster Ordnung (mit konstanten Koeffizienten) lösen.

4.5.3 Korollar

Besitzt V eine f-zyklische Basis für $f \in \mathrm{End}(V)$, so gilt $\chi_f(f) = 0$.

Beweis Sei also $B = (b_1, \dots, b_n) = (b, f(b), \dots, f^{n-1}(b))$ f-zyklische Basis von V und $a_0, \dots, a_{n-1} \in K$ so, dass

$$0 = f^n(b) + \sum_{k=0}^{n-1} f^k(b) a_k.$$

Dann gilt

$$\chi_f(f)(b_1) = \chi_f(f)(b) = \left(f^n + \sum_{k=0}^{n-1} f^k a_k \right)(b) = f^n(b) + \sum_{k=0}^{n-1} f^k(b) a_k.$$

Damit folgt für $i = 2, \dots, n$

$$\chi_f(f)(b_i) = \chi_f(f)\left(f^{i-1}(b)\right) \overset{2}{=} f^{i-1}\left(\chi_f(f)(b)\right) = 0.$$

Da also $V = [B] \subset \ker \chi_f(f)$, folgt $\chi_f(f) = 0$.

Bemerkung Damit ist der Satz von Cayley-Hamilton bewiesen, sofern V eine f-zyklische Basis besitzt.

4.5.4 Lemma

Für $f \in \mathrm{End}(V)$ und $v \in V^\times$ sei

$$U := \left[\left(f^k(v) \right)_{k \in \mathbb{N}} \right].$$

[2] Aufgrund der Linearität der Endomorphismen $\mathrm{End}(V)$ als unitäre Algebra.

Damit ist U ein f-invarianter UVR von V. Ist $\dim V < \infty$, so besitzt U eine f-zyklische Basis $(v, f(v), \ldots, f^{r-1}(v))$.

Beweis Offenbar ist U f-invarianter UR:

- U ist (als lineare Hülle einer Familie) ein UVR von V;
- da gilt
$$\forall k \in \mathbb{N} : f\left(f^k(v)\right) = f^{k+1}(v) \in U$$
folgt, dass $f(U) = f\left(\left[\left(f^k(v)\right)_{k\in\mathbb{N}}\right]\right) = \left[\left(f^{k+1}(v)\right)_{k\in\mathbb{N}}\right] \subset U.$

Ist $\dim V < \infty$ und $v \neq 0$, so existiert $r \in \mathbb{N}$, sodass

$$\left(v, \ldots, f^{r-1}(v)\right) \text{ linear unabhängig und } f^r(v) \in \left[\left(v, \ldots, f^{r-1}(v)\right)\right];$$

damit ist $(v, f(v), \ldots, f^{r-1}(v))$ f-zyklische Basis von U:

1. $(v, \ldots, f^{r-1}(v))$ ist linear unabhängig.
2. $f^r(v) \in [(v, \ldots, f^{r-1}(v))]$, damit gilt
$$\forall k \in \mathbb{N} : k \geq r \Rightarrow f^k(v) \in \left[\left(v, \ldots, f^{r-1}(v)\right)\right]$$

wie man z.B. mit Induktion sehen kann: ist

$$f^{k-1}(v) = \sum_{j=0}^{r-1} f^j(v) x_j \in \left[\left(v, \ldots, f^{r-1}(v)\right)\right],$$

so folgt

$$f^k(v) = \sum_{j=1}^{r} f^j(v) x_{j-1} = f^r(v) x_{r-1} + \sum_{j=1}^{r-1} f^j(v) x_{j-1} \in \left[\left(v, \ldots, f^{r-1}(v)\right)\right]$$

und damit
$$U = \left[\left(f^k(v)\right)_{k\in\mathbb{N}}\right] \in \left[\left(v, \ldots, f^{r-1}(v)\right)\right].$$

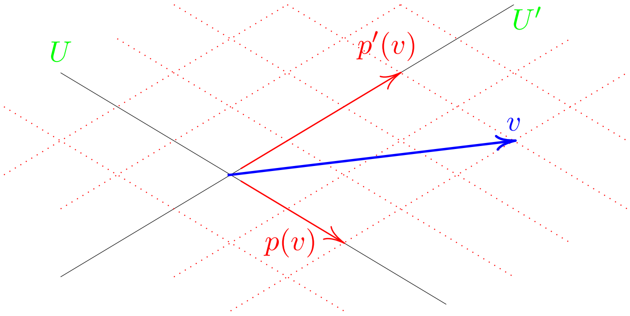
4.5.5 Beweis vom Satz von Cayley-Hamilton

Zu zeigen: für $f \in \operatorname{End}(V)$ gilt $\chi_f(f) = 0$, d.h.

$$\forall v \in V : \chi_f(f)(v) = 0.$$

Seien also $v \in V^\times$ und
$$U := \left[\left(f^k(v)_{k\in\mathbb{N}}\right)\right] \subset V.$$

Mit einem zu U komplementären UVR $U' \subset V$, $V = U \oplus U'$, und den zugehörigen Projektionen



$$p : V \to V, p(V) = U, \ker p = U' \text{ bzw.}$$
$$p' : V \to V, p'(V) = U', \ker p' = U,$$

ist dann $\chi_f(t) = \chi_{f'}(t) \cdot \chi_{f|_U}(t)$ mit $f' := p' \circ f \mid_{U'} \in \mathrm{End}(U')$.
Damit folgt

$$\chi_f(f)(v) = \chi_{f'}(f)\left(\chi_{f|_U}(f)(v)\right) = \chi_{f'}(f)(0) = 0$$

nach Korollar oben, da U eine f-zyklische Basis besitzt und $v \in U$.

4.5.6 Definition

Sei V ein K-VR und $f \in \mathrm{End}(V)$. Dann heißt $p \in K[t]$

- *Annulatorpolynom von f*, falls $p(f) = 0$;
- *Minimalpolynom von f*, falls $p(t)$ normiertes Annulatorpolynom minimalen Grades ist.

Bemerkung Jedes (polynomiale) Vielfache

$$p(t) = q(t)\mu_f(t) \in K[t]$$

eines Minimalpolynoms $\mu_f(t)$ von f ist ein Annulatorpolynom, da

$$\forall v \in V : p(f)(v) = (q(f) \circ \mu_f(f))(v) = q(f)(\mu_f(f)(v)) = q(f)(0) = 0$$

Bemerkung Nach dem Satz von Cayley-Hamilton hat jeder Endomorphismus $f \in \mathrm{End}(f)$ ein Annulatorpolynom, also auch ein Minimalpolynom – wenn $\dim V < \infty$.

4.5.7 Lemma

Ist $p(t) \in K[t]$ Annulatorpolynom von $f \in \mathrm{End}(V)$, so ist jedes Minimalpolynom $\mu_f(t) \in K[t]$ Teiler von $p(t)$.

Beweis Seien $q(t), r(t) \in K[t]$ die (nach dem euklidischen Divisionsalgorithmus) eindeutigen Polynome mit

$$p(t) = q(t)\mu_f(t) + r(t) \text{ und } \deg r(t) < \deg \mu_f(t).$$

Dies liefert

$$r(f) = p(f) - q(f) \circ \mu_f(f) = 0 - q(f)(0) = 0,$$

also $r(t) = 0$, denn andernfalls wäre $\mu_f(t)$ nicht normiertes Annulatorpolynom minimalen Grades.

4.5.8 Korollar

Das Minimalpolynom $\mu_f(t) \in K[t]$ eines Endomorphismus $f \in \mathrm{End}(V)$ ist eindeutig.

Beweis Sind $\mu_f(t), \tilde{\mu}_f(t) \in K[t]$ Minimalpolynome von $f \in \mathrm{End}(V)$, so gilt

$$\exists! q(t) \in K[t] : \tilde{\mu}_f(t) = q(t)\mu_f(t)$$

wobei

- $\deg q(t) = 0$, da $\deg \tilde{\mu}_f(t) \leq \deg \mu_f(t)$,
- $q(t) = 1$, da $\tilde{\mu}_f(t)$ und $\mu_f(t)$ normiert sind.

Daher ist

$$\tilde{\mu}_f(t) = 1 \cdot \mu_f(t) = \mu_f(t).$$

Bemerkung Wie für Endomorphismen kann man Annulatorpolynome, Minimalpolynome, usw. auch für Matrizen $X \in K^{n \times n}$ definieren:

- mithilfe der assoziierten Endomorphismen $f_X \in \mathrm{End}(K^n)$, oder
- mithilfe des Einsetzungshomomorphismus $\psi_X : K[t] \to K^{n \times n}$.

Beide Methoden liefern das gleiche Ergebnis durch den Algebrahomomorphismus zwischen den Endomorphismen und den quadratischen Matrizen.

Bemerkung & Beispiel Zerfällt das charakteristische Polynom in Linearfaktoren, so zerfällt auch das Minimalpolynom in dieselben Linearfaktoren:

$$\chi_f(t) = \prod_{i=1}^{m} (t - x_i)^{k_i} \Rightarrow \mu_f(t) = \prod_{i=1}^{m} (t - x_i)^{m_i},$$

wobei für $i = 1, \ldots, m$ gilt $1 \leq m_i \leq k_i$.

Zum Beispiel:

- $X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$: $\chi_{f_X} = t(t-1) = \mu_{f_X}(t)$

- $X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$: $\chi_{f_X} = (t-1)^2 \Rightarrow \mu_{f_X}(t) = (t-1)$

- $X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$: $\chi_{f_X} = (t-1)^2 = \mu_{f_X}(t)$.

Bemerkung Die Definition des charakteristischen Polynoms ist etwas problematisch:

$$\chi_f(t) := \det(\mathrm{id}_V t - f)$$

ist „gut“ für Polynomfunktionen, aber „nicht korrekt“ für abstrakte Polynome; die Definition

$$\chi_f(t) := \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} \left(\delta_{\sigma(j)j} - x_{\sigma(j)j}\right) \in K[t]$$

mithilfe der Darstellungsmatrix $X = (x_{ij})_{i,j \in \{1,\dots,n\}} = \xi_B^B(f)$ von f bzgl. einer Basis B und der Leibniz-Formel ist nicht sehr übersichtlich. Vergleiche auch [Axler, Kap. 8] zum Thema.

Im Gegensatz dazu: Definitionen von „Annulatorpolynom“ und „Minimalpolynom“ etc. sind einfach (konzeptionell).

Frage: Braucht man das charakteristische Polynom überhaupt? Man kommt auch ohne das charakteristische Polynom „recht weit“:

- Für $\dim V < \infty$ folgt die Existenz eines Annulatorpolynoms, und damit des Minimalpolynoms recht einfach wegen $\dim \mathrm{End}(V) < \infty$.

- Durch Einsetzen: Jeder Eigenwert eines Endomorphismus ist Nullstelle seines Minimalpolynoms.

- Umgekehrt ist auch jede Nullstelle des Minimalpolynoms Eigenwert – ist $\mu_f(x) = 0$, so existiert $q(t) \in K[t]$ mit
$$\mu_f(t) = q(t)(t-x);$$
wäre x kein Eigenwert, also $f - \mathrm{id}_V x \in Gl(V)$, so gälte
$$(f - \mathrm{id}_V x)(V) = V \Rightarrow \{0\} = \mu_f(f)(V) = q(f)(V),$$
d.h. $\mu_f(t)$ wäre nicht Minimal-Polynom.

- Ein Endomorphismus ist diagonalisierbar, wenn sein Minimal-Polynom in paarweise verschiedene Linearfaktoren zerfällt.

Nachteil des Minimal-Polynoms: schwierig berechenbar?

5 Längen- und Winkelmessung

Plan: Längen und Winkel (in „Punkträumen" $\cong$ affinen Räumen) verstehen.
Algebraisch: via Produkte (bilineare – oder fast bilineare – Abbildungen).

5.1 Bilinearformen & Sesquilinearformen

Zur Erinnerung Sind V und W K-VR, so nennt man eine Abbildung

$$\beta : V \times V \to W$$

bilinear oder ein *Produkt*, wenn sie in jedem Argument linear ist:

(i) $\forall w \in V : V \ni v \mapsto \beta(v, w) \in W$ ist linear;

(ii) $\forall v \in V : V \ni w \mapsto \beta(v, w) \in W$ ist linear.

Zu vorgegebenen Werten $\beta_{ij} \in W$ auf einer Basis $(b_i)_{i\in I}$ von V existiert dann eine eindeutige Bilinearform β (Fortsetzungssatz Abschnitt 4.3):

$$\exists! \beta : V \times V \to W \text{ bilinear} : \forall i, j \in I : \beta(b_i, b_j) = \beta_{ij}.$$

Bemerkung Man kann auch bilineare Abbildungen $V \times V' \to W$ betrachten und, zum Beispiel, auch einen Fortsetzungssatz beweisen.

Wir benötigen eine Verallgemeinerung in eine andere Richtung:

5.1.1 Definition

Seien V ein K-VR und $K \ni x \mapsto \overline{x} \in K$ ein (Körper-) Automorphismus, d.h. eine bijektive Abbildung mit

$$\overline{x + y} = \overline{x} + \overline{y} \text{ und } \overline{xy} = \overline{x} \cdot \overline{y}$$

für alle $x, y \in K$. Eine Abbildung $\sigma : V \times V \to K$ heißt dann *Sesquilinearform* (bzgl. $\bar{.}$), falls

(i) $\forall v \in V : V \ni w \mapsto \sigma(v, w) \in K$ ist linear, d.h. $\sigma(v, .) \in V^*$;

(ii) $\forall w \in V : V \ni v \mapsto \sigma(v, w) \in K$ ist *semilinear*, d.h.

   (a) $\forall v, v' \in V : \sigma(v + v', w) = \sigma(v, w) + \sigma(v', w)$ und

   (b) $\forall v \in V \forall x \in K : \sigma(vx, w) = \overline{x}\sigma(v, w)$.

Beispiel Die Identität $K \ni x \mapsto \overline{x} := x \in K$ ist offensichtlich ein Körperautomorphismus für jeden Körper K. *Bilinearformen* sind genau die Sesquilinearformen bezüglich id_K.

Beispiel Für $K = \mathbb{C}$ liefert *komplexe Konjugation* einen Körperautomorphismus (keinen VR-Automorphismus, vgl. Abschnitt 1.4):

$$\mathbb{C} \ni x + iy \mapsto \overline{x + iy} := x - iy \in \mathbb{C}.$$

Dieses Beispiel ist unser Grund für die Einführung des Begriffs der Sesquilinearform.

Bemerkung Ist σ Bilinearform und Sesquilinearform bezüglich $\overline{\cdot}$, so ist σ oder $\overline{\cdot}$ trivial:

$$\forall x \in K \forall v, w \in V : 0 = \sigma(vx, w) - \sigma(vx, w) = (x - \overline{x})\sigma(v, w)$$

$$\Rightarrow \begin{cases} \forall v, w \in V : \sigma(v, w) = 0 \text{ oder} \\ \exists v, w \in V : \sigma(v, w) \neq 0 \wedge \forall x \in K : \overline{x} = x. \end{cases}$$

Bemerkung In $\mathbb{Z}_p, \mathbb{Q}$ und $\mathbb{R}$ gibt es nur *einen* Körperautomorphismus: id_K. Ein Automorphismus $\overline{\cdot}$ von $\mathbb{C}$ mit $\overline{\mathbb{R}} = \mathbb{R}$ ist trivial, $\overline{\cdot} = \mathrm{id}_{\mathbb{C}}$ oder die komplexe Konjugation.

5.1.2 Fortsetzungssatz für Sesquilinearformen

Sind V ein K-VR und $K \ni x \mapsto \overline{x} \in K$ ein Körperautomorphismus, $(b_i)_{i \in I}$ Basis von V und $(s_{ij})_{i,j \in I}$ eine Familie in K, so existiert eine eindeutige Sesquilinearform σ mit

$$\forall i, j \in I : \sigma(b_i, b_j) = s_{ij}.$$

Beweis Wir imitieren den Beweis unseres ersten Fortsetzungssatzes für lineare Abbildungen:

Eindeutigkeit: Sei σ eine Sesquilinearform mit der gewünschten Eigenschaft oben; gilt

$$v = \sum_{i \in I} b_i x_i \text{ und } w = \sum_{i \in I} b_i y_i$$

so folgt

$$\sigma(v, w) = \sum_{i,j \in I} \overline{x_i} \sigma(b_i, b_j) y_j = \sum_{i,j \in I} \overline{x_i} s_{ij} y_j$$

d.h. σ ist durch die Familie $(s_{ij})_{i,j \in I}$ eindeutig bestimmt.

Existenz: Da jeder Vektor $v \in V$ eine eindeutige Basisdarstellung $v = \sum_{i \in I} b_i x_i$ hat, wird durch

$$\sigma : V \times V \to K, (v, w) = \left(\sum_{i \in I} b_i x_i, \sum_{j \in I} b_j y_j \right)$$

$$\mapsto \sigma(v,w) := \sum_{i,j\in I} \overline{x_i} s_{ij} y_j$$

eine Abbildung wohldefiniert. Offenbar (nachrechnen) ist σ dann sesquilinear.

Bemerkung Jede Sesquilinearform $\sigma : V \times V \to K$ liefert eine semi-lineare Abbildung

$$V \ni v \mapsto \sigma(v, .) \in V^*.$$

Mit einem „Fortsetzungssatz für semi-lineare Abbildungen“ (Aufgabe 34) hätte man auch den früher skizzierten Beweis für bilineare Abbildungen imitieren können.

5.1.3 Buchhaltung

Gramsche Matrix Ist $n = \dim V < \infty$ und $B = (b_1, \ldots, b_n)$ Basis von V, so kann man eine Sesquilinearform $\sigma : V \times V \to K$ durch eine Matrix S beschreiben:

$$\begin{array}{c|ccc} \sigma & b_1 & \ldots & b_n \\ \hline b_1 & s_{11} & & s_{1n} \\ \vdots & & \ddots & \\ b_n & s_{n1} & & s_{nn} \end{array}$$

Diese Matrix

$$\Gamma_B(\sigma) = S = (\sigma(b_i, b_j))_{i,j\in\{1,\ldots,n\}}$$

heißt die Darstellungsmatrix oder *Gramsche Matrix* von σ bezüglich B. Für Vektoren

$$v = \sum_{i=1}^{n} b_i x_i = BX \text{ und } w = \sum_{j=1}^{n} b_j y_j = BY$$

ist dann

$$\sigma(v,w) = \sum_{i,j=1}^{n} \overline{x_i} s_{ij} y_j = \overline{X}^t SY$$

$$= (\overline{x_1}, \ldots, \overline{x_n}) \begin{pmatrix} \sum_{i=1}^{n} s_{1j} y_j \\ \vdots \\ \sum_{j=1}^{n} s_{nj} y_j \end{pmatrix} = \sum_{i=1}^{n} \overline{x_i} \sum_{j=1}^{n} s_{ij} y_j.$$

Transformationsformel Ein Basiswechsel $B' = BP$ mit $P = \xi_{B'}^{B} \in Gl(n)$ liefert dann

$$v = BX = (B'P^{-1})X = B' \underbrace{(P^{-1}X)}_{X'} \text{ und } w = B' \underbrace{(P^{-1}Y)}_{Y'}$$

und damit für $X, Y \in K^{n\times 1}$

$$\overline{X}^t SY = \overline{X'}^t \underbrace{(\overline{P}^t SP)}_{S'} Y'$$

woraus die *Transformationsformel für Gramsche Matrizen* folgt

$$S' = \overline{P}^t SP,$$

wobei $\overline{P}^t$ die Transponierte der Matrix mit Einträgen $\overline{p_{ij}}$ ist.

Äquivalenz von Matrizen Dies liefert einen weiteren Äquivalenzbegriff für quadratische Matrizen $S \in K^{n\times n}$:

$$S' \sim S :\Leftrightarrow \exists P \in Gl(n) : S' = \overline{P}^t SP.$$

Die verschiedenen Begriffe der Äquivalenz von Matrizen (vgl. 3.1 & 4.2) spiegeln die verschiedenen Funktionen/Bedeutungen von Matrizen wider.

Bemerkung Die Menge der Sesquilinearformen auf einem K-VR ist selbst ein K-VR. Ist $n = \dim V < \infty$ und B Basis von V, so erhält man (Fortsetzungssatz) einen Isomorphismus

$$K^{V\times V} \supset \{\sigma : V \times V \to K \text{ Sesquilinearform}\} \ni \sigma \mapsto \Gamma_B(\sigma) \in K^{n\times n}.$$

5.1.4 Beispiel & Definition

Sei $\bar{.} : K \to K$ Körperautomorphismus; jedes $S \in K^{n\times n}$ liefert dann eine eindeutige Sesquilinearform

$$\sigma_S : K^n \times K^n \to K \text{ mit } (e_i, e_j) \mapsto \sigma_S(e_i, e_j) := s_{ij},$$

die zu S *assoziierte Sesquilinearform.*

Für $S = E_n$ bezeichnet man σ_S auch als *kanonische Sesquilinearform.*

5.1.5 Definition

Eine Sesquilinearform $\sigma : V \times V \to K$ auf einem K-VR bzgl. eines Automorphismus $\bar{.} : K \to K$ nennen wir

(i) *symmetrisch*, falls $\forall v, w \in V : \sigma(w, v) = \overline{\sigma(v, w)};$

(ii) *schiefsymmetrisch*, falls $\forall v, w \in V : \sigma(w, v) = -\overline{\sigma(v, w)};$

(iii) *alternierend*, falls $\forall v \in V : \sigma(v, v) = 0.$

Falls $K = \mathbb{C}$ und $\bar{.}$ komplexe Konjugation sind, so nennt man eine symmetrische Sesquilinearform auch *Hermitesche Sesquilinearform.*

Bemerkung Ist σ nicht-trivial und (schief-)symmetrisch, so muss $\bar{\cdot}$ eine Involution sein.

Nämlich: Wähle $v, w \in V$ mit $\sigma(v, w) = 1$; dann gilt

$$\forall x \in K : \overline{\overline{x}} = \overline{\sigma(vx, w)} = \pm\sigma(w, vx) = \overline{x}\sigma(v, w) = \pm\sigma(w, v)x = \overline{\sigma(v, w)}x = x.$$

Ist $\mathrm{Char}(K) \neq 2$ und $\bar{\cdot}$ Involution, so kann jede Sesquilinearform in einen symmetrischen und einen schiefsymmetrischen Anteil zerlegt werden:

$$\forall v, w \in V : \sigma(v, w) = \frac{1}{2}\left(\sigma(v, w) + \overline{\sigma(w, v)}\right) + \frac{1}{2}\left(\sigma(v, w) - \overline{\sigma(w, v)}\right).$$

Bemerkung Ist $\mathrm{Char}(K) \neq 2$ und $\bar{\cdot} = \mathrm{id}_K$, so sind „alternierend" und „schiefsymmetrisch" äquivalent für eine Sesquilinearform σ.

Andererseits ist jede alternierende Sesquilinearform bilinear, d.h. $\bar{\cdot} = \mathrm{id}_K$ oder $\sigma = 0$.

Buchhaltung Unter den folgenden Annahmen:

- $\mathrm{Char}(K) \neq 2$ und $\bar{\cdot}$ Involution;
- $n = \dim V < \infty$ und B ist Basis von V;

gilt für die Gramsche Matrix $S = \Gamma_B(\sigma)$ einer Sesquilinearform σ auf V:

- $0 = \overline{S}^t - S \Leftrightarrow \sigma$ symmetrisch;[1]
- $0 = S + \overline{S}^t \Leftrightarrow \sigma$ schiefsymmetrisch.

Nämlich:

$$\overline{S}^t = \begin{pmatrix} \overline{\sigma(b_1, b_1)} & \overline{\sigma(b_1, b_2)} & \cdots & \overline{\sigma(b_1, b_n)} \\ \overline{\sigma(b_2, b_1)} & & & \vdots \\ \vdots & & & \vdots \\ \overline{\sigma(b_n, b_1)} & & \cdots & \overline{\sigma(b_n, b_n)} \end{pmatrix}^t = \begin{pmatrix} \overline{\sigma(b_1, b_1)} & \overline{\sigma(b_2, b_1)} & \cdots & \overline{\sigma(b_n, b_1)} \\ \overline{\sigma(b_1, b_2)} & & & \vdots \\ \vdots & & & \vdots \\ \overline{\sigma(b_1, b_n)} & & \cdots & \overline{\sigma(b_n, b_n)} \end{pmatrix}$$

$$S = \begin{pmatrix} \sigma(b_1, b_1) & \sigma(b_1, b_2) & \cdots & \sigma(b_1, b_n) \\ \sigma(b_2, b_1) & & & \vdots \\ \vdots & & & \vdots \\ \sigma(b_n, b_1) & & \cdots & \sigma(b_n, b_n) \end{pmatrix}$$

[1] bis auf Faktor 2: Gramsche Matrix des schiefsymmetrischen Anteils von σ

5.1.6 Definition

Sei σ symmetrische Sesquilinearform auf einem Vektorraum V. Zwei Vektoren $v, w \in V$ heißen *orthogonal* (bzgl. σ),

$$w \perp v, \text{ falls } \sigma(v, w) = 0.$$

Der *Orthogonalraum* einer Menge $\emptyset \neq S \subset V$ ist der UVR

$$S^{\perp} := \bigcap_{s \in S} \ker \underbrace{\sigma(s, .)}_{\in V^*}.$$

Bemerkung Wegen der Symmetrie von σ ist die *Orthogonalitätsrelation* symmetrisch,

$$w \perp v \Leftrightarrow v \perp w.$$

Bemerkung Da $\forall v \in V : \sigma(v, .) \in V^*$, ist der Orthogonalraum wohldefiniert und (als Schnitt von UVR) ein UVR. Offenbar gilt

$$\tilde{S} \subset S \Rightarrow \tilde{S}^{\perp} \supset S^{\perp}.$$

Damit folgt direkt $S^{\perp} \supset [S]^{\perp}$, sind andererseits $w \in S^{\perp}$ und $v \in [S]$, so gilt

$$v = \sum_{s \in S} s x_s \Rightarrow \sigma(v, w) = \sum_{s \in S} \overline{x_s} \sigma(s, w) = 0, \text{ da } \forall s \in S : w \perp s$$

d.h. $w \in S^{\perp} \Rightarrow w \in [S]^{\perp}$. Insgesamt ist also

$$\forall S \subset V : [S]^{\perp} = S^{\perp}.$$

Ähnlich zeigt man für jede Familie $(U_i)_{i \in I}$ von UVR $U_i \subset V$:

$$\left(\sum_{i \in I} U_i \right)^{\perp} = \bigcap_{i \in I} U_i^{\perp}.$$

Bemerkung & Beispiel Für $S \subset V$ kann man $S^{\perp\perp} = \left(S^{\perp} \right)^{\perp}$ betrachten; im Allgemeinen gilt

$$S \subset S^{\perp\perp} \text{ aber } S \neq S^{\perp\perp}.$$

Ist etwa $\sigma = 0$, so ist $S^{\perp} = V$ für jede Menge $\emptyset \neq S \subsetneq V$; also ist

$$S^{\perp\perp} = V^{\perp} = V \neq S.$$

5.1.7 Definition

$V^{\perp}$ ist der *Radikal(-raum)* eines VR mit symmetrischer Sesquilinearform σ; ist $V^{\perp} = \{0\}$, so heißt σ *radikalfrei* oder *nicht-degeneriert*, andernfalls *degeneriert*.

Beispiel Betrachte $V = \mathbb{R}^2$ mit Standardbasis (e_1, e_2).

Ist für eine symmetrische Sesquilinearform (Bilinearform) σ auf V

$$\sigma(e_1, e_1) = 0, \sigma(e_1, e_2) = 1, \sigma(e_2, e_2) = 0$$

so ist σ nicht-degeneriert, $V^{\perp} = \{0\}$, da

$$v = e_1x_1 + e_2x_2 \perp e_1, e_2 \Rightarrow \begin{cases} 0 = \sigma(e_1, v) = x_2 \\ 0 = \sigma(e_2, v) = x_1 \end{cases}$$

$$\Rightarrow x_1 = x_2 = 0,$$

also $V^{\perp} = \{0\}$, d.h. σ ist nicht-degeneriert.

Ist aber

$$\sigma(e_1, e_1) = 1, \sigma(e_1, e_2) = 1, \sigma(e_2, e_2) = 1,$$

so ist $V^{\perp} = [e_1 - e_2]$, d.h. σ ist degeneriert.

5.1.8 Lemma

Ist $U \subset V$ ein zum Radikal von (V, σ) komplementärer UVR, $V = V^{\perp} \oplus U$, so ist

$$\sigma\big|_{U \times U} : U \times U \to K$$

radikalfrei.

Beweis Sei $u \in U$ im Radikal von $(U, \sigma\big|_{U \times U})$, d.h. es gelte $\forall v \in U : \sigma(v, u) = 0$. Weil

$$\forall v \in V^{\perp} \forall w \in V : v \perp w \Rightarrow \forall v \in V^{\perp} : v \perp u$$

erhalten wir $u \in U \cap V^{\perp} = \{0\}$.

Beispiel Die Einschränkung von σ mit (wie oben)

$$\forall i, j \in \{1, 2\} : \sigma(e_i, e_j) = 1$$

auf jeden UVR $U = [e_1x_1 + e_2x_2]$ mit $x_1 + x_2 \neq 0$ ist radikalfrei, denn

$$\sigma(e_1x_1 + e_2x_2, e_1x_1 + e_2x_2) = (x_1 + x_2)^2 \neq 0.$$

5.2 Der Satz von Sylvester

Beispiel Ist σ symmetrische Sesquilinearform auf $V = \mathbb{Z}_2^2$ mit

$$\sigma(e_1, e_1) = 0, \sigma(e_1, e_2) = 1, \sigma(e_2, e_2) = 0,$$

so ist σ (wie vorher) nicht-degeneriert, $V^\perp = \{0\}$; trotzdem gilt

$$\forall v \in V : \sigma(v, v) = 0.$$

Das folgende Lemma zeigt, dass dies ein degenerierter Fall ist:

5.2.1 Lemma & Definition (Polarisation)

Ist σ symmetrische Bilinearform auf einem K-VR V über einem Körper K mit $\operatorname{Char} K \neq 2$, so gilt

$$\forall v, w \in V : \sigma(v, w) = \frac{1}{2}\left(q(v + w) - q(v) - q(w)\right),$$

wobei

$$q : V \to K, v \mapsto q(v) := \sigma(v, v)$$

die zu σ gehörige *quadratische Form* bezeichnet.

Beweis Ausrechnen: sind $v, w \in V$, so gilt

$$\begin{aligned}
q(v + w) &= \sigma(v + w, v + w) \\
&= \sigma(v, v) + \sigma(v, w) + \sigma(w, v) + \sigma(w, w) \\
&= q(v) + 2\sigma(v, w) + q(w)
\end{aligned}$$

diese Gleichung kann (da $\operatorname{Char} K \neq 2$) nach $\sigma(v, w)$ aufgelöst werden.

Bemerkung Ist $\operatorname{Char} K = 0$ so kann man statt

$$q(v + w) = q(v) + 2\sigma(v, w) + q(w)$$

auch

$$q(v + w) - q(v - w) = 4\sigma(v, w)$$

für die Polarisation verwenden.

5.2.2 Lemma

Ist σ symmetrische Sesquilinearform auf einem K-VR V über einem Körper K mit $\operatorname{Char} K \neq 2$, so gilt

$$\sigma = 0 \Leftrightarrow \forall v \in V : \sigma(v, v) = 0.$$

Bemerkung Im Falle einer Bilinearform folgt dies direkt mit Polarisation.

Im Falle eines nicht-trivialen Körperautomorphismus $\bar{\cdot}$ liefert $v \mapsto \sigma(v, v)$ wegen

$$K \ni x \mapsto \sigma(vx, vx) - x^2\sigma(v, v) = (\overline{x}x - x^2)\sigma(v, v) \neq 0$$

im Allgemeinen *keine* quadratische Form:

$$\exists x \in K : \exists v \in V : \sigma(vx, vx) = \overline{x}\sigma(v, v)x \neq x^2\sigma(v, v).$$

Beweis Ist $\sigma = 0$, so folgt trivialerweise

$$\forall v \in V : \sigma(v, v) = 0.$$

Sei nun $\sigma \neq 0$, d.h.

$$\exists v, w \in V : \sigma(v, w) \neq 0.$$

Wie vorher berechnet man für $v, w \in V$

$$\sigma(v + w, v + w) = \sigma(v, v) + \sigma(v, w) + \overline{\sigma(v, w)} + \sigma(w, w).$$

Wähle nun $v, w \in V$ mit $\sigma(v, w) \neq 0$, o.B.d.A. $\sigma(v, w) = 1$. [2] Ist $\sigma(v, v) \neq 0$ oder $\sigma(w, w) \neq 0$, so sind wir fertig.

Gilt jedoch $\sigma(v, v) = \sigma(w, w) = 0$, so liefert

$$\sigma(v + w, v + w) = 0 + 1 + 1 + 0 \neq 0$$

wieder die Behauptung, da $\operatorname{Char} K \neq 2$.

Vereinbarung Im Folgenden schließen wir $\operatorname{Char} K = 2$ aus.

[2] Ggf. ersetzt man w durch $\frac{w}{\sigma(v,w)}$.

5.2.3 Lemma

Für eine symmetrische Sesquilinearform σ auf V und $b \in V$ mit $\sigma(b,b) \neq 0$ gilt

$$V = [b] \oplus \{b\}^{\perp}.$$

Beweis Es gilt $V = [b] + \{b\}^{\perp}$, da für $v \in V$

$$v = u + b\frac{\sigma(b,v)}{\sigma(b,b)} \text{ mit } u := v - b\frac{\sigma(b,v)}{\sigma(b,b)} \perp b;$$ [3]

ist $v \in [b] \cap \{b\}^{\perp}$, so gilt

$$v = bx \text{ für ein } x \in K \text{ und}$$

$$0 = \sigma(b,v) = \sigma(b,bx) = \sigma(b,b)x$$

$$\Rightarrow x = 0 \wedge v = 0,$$

d.h. $[b] \cap \{b\}^{\perp} = \{0\}$ und damit folgt die Behauptung.

Bemerkung Ist $\sigma(b,b) = 0$ für $b \in V$, so gilt

$$b \in [b] \cap \{b\}^{\perp},$$

d.h. ist $b \neq 0$, so ist $[b] \cap \{b\}^{\perp} \neq \{0\}$. Außerdem ist dann $\sigma|_{U \times U}$ für $U := \{b\}^{\perp}$ degeneriert, da

$$\exists v = b \in U^{\times} \forall u \in U : u \perp b.$$

5.2.4 Diagonalisierungslemma

Zu jeder symmetrischen Sesquilinearform σ auf einem endlichdimensionalen VR V, also $n = \dim V < \infty$, gibt es eine Basis $B = (b_1, \dots, b_n)$ von V, die σ *diagonalisiert*, d.h. für die gilt

$$\sigma(b_i, b_j) = 0, \text{ falls } i \neq j.$$

Beweis Durch Induktion über n.

Für $n = 1$ ist die Behauptung trivial (denn $i \neq j$ existiert nicht).

Sei die Behauptung also für $\dim V = n$ bewiesen. Ist σ symmetrische Sesquilinearform auf V mit $\dim V = n + 1$ und o.B.d.A. $\sigma \neq 0$, also

$$\exists b \in V : \sigma(b,b) \neq 0$$

[3] denn $\sigma(b,u) = \sigma(b, v - b\frac{\sigma(b,v)}{\sigma(b,b)}) = \sigma(b,v) - \sigma(b,b)\frac{\sigma(b,v)}{\sigma(b,b)} = 0$

nach obigem Lemma lässt sich also V aufspalten in

$$V = [b] \oplus U \text{ mit } U := \{b\}^\perp$$

und $\dim U = n$. Nach Annahme existiert eine Basis $(b_1, \ldots, b_n)$ von U, die $\sigma|_{U \times U}$ diagonalisiert. Da $b \perp b_1, \ldots, b_n \in U$ liefert $B := (b, b_1, \ldots, b_n)$ eine σ-diagonalisierende Basis von V.

Bemerkung Ist $B = (b_1, \ldots, b_n)$ eine σ-diagonalisierende Basis, also

$$s_{ij} = \sigma(b_i, b_j) = 0 \text{ für } i \neq j$$

so ist

$$\sigma(v, v) = \sum_{i=1}^{n} \overline{x_i} s_{ii} x_i \text{ für } v = \sum_{i=1}^{n} b_i x_i.$$

Sind $a_1, \ldots, a_n \in K^\times$ und $b_i' = b_i a_i$, so zeigt

$$s_{ij}' = \sigma(b_i', b_j') = \overline{a_i} \sigma(b_i, b_j) a_j = \overline{a_i} s_{ij} a_j,$$

dass $B' = (b_1', \ldots, b_n')$ eine weitere σ-diagonalisierende Basis ist. Man kann also die s_{ii} „adjustieren“, sofern man die (unabhängigen) Gleichungen

$$s_{ii}' = \overline{a_i} s_{ii} a_i$$

für gegebene s_{ii}' (nach den a_i) lösen kann. Zum Beispiel:

5.2.5 Korollar

Ist σ symmetrische Bilinearform auf einem $\mathbb{C}$-VR V mit $\dim V < \infty$, so besitzt V eine Basis $B = (b_1, \ldots, b_n)$, sodass

$$\exists r \in \mathbb{N} : s_{ij} = \sigma(b_i, b_j) = \begin{cases} 1 & \text{für } i = j \leq r \\ 0 & \text{sonst.} \end{cases}$$

Bemerkung D.h.

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

Beweis Sei (nach Diagonalisierungslemma) $B' = (b_1', \ldots, b_n')$ eine σ-diagonalisierende Basis von V; durch Umsortierung der Basisvektoren kann man erreichen, dass

$$s_{11}', \ldots, s_{rr}' \neq 0 \text{ und } s_{r+1,r+1}' = \cdots = s_{nn}' = 0$$

für ein $r \in \{0, \ldots, n\}$. Mit einer Wahl der Wurzel bilden die Vektoren

$$b_i := \begin{cases} b'_i \cdot \frac{1}{\sqrt{s'_{ii}}} & \text{für } i = 1, \ldots, r \\ b'_i = 0 & \text{für } i = r+1, \ldots, n \end{cases}$$

dann eine Basis B mit der gewünschten Eigenschaft:

$$\sigma(b_i, b_i) = \underbrace{\sigma(b'_i, b'_i)}_{s'_{ii}} \cdot \left(\frac{1}{\sqrt{s'_{ii}}}\right)^2 = 1 \text{ für } i = 1, \ldots, r$$

$$\sigma(b_i, b_i) = \sigma(b'_i, b'_i) = 0 \text{ für } i = r+1, \ldots, n.$$

5.2.6 Korollar

Ist V ein K-VR mit $\dim V < \infty$ und σ entweder

- symmetrische Bilinearform, wenn $K = \mathbb{R}$, oder
- Hermitesche Sesquilinearform, wenn $K = \mathbb{C}$,

so besitzt V eine Basis $B = (b_1, \ldots, b_n)$, sodass

$$\exists r \in \mathbb{N} : s_{ij} = \sigma(b_i, b_j) = \begin{cases} \pm 1 & \text{für } i = j \leq r \\ 0 & \text{sonst} \end{cases}$$

Beweis Wie oben – aber: In diesen beiden Fällen gilt für eine diagonalisierende Basis $B' = (b'_1, \ldots, b'_n)$ und $b_i = b'_i \cdot \frac{1}{a_i}$ mit $a_i \in K$ für $i = 1, \ldots, n$:

$$s'_{ii} = \sigma(b'_i, b'_i) \in \mathbb{R} \text{ und } \begin{cases} a_i^2 \geq 0 & \text{falls } K = \mathbb{R}, \\ \overline{a_i} a_i \geq 0 & \text{falls } K = \mathbb{C}. \end{cases}$$

Also kann man die s'_{ii} (nur) positiv reskalieren und so $s_{ii} = 0$ oder $s_{ii} = \pm 1$ erreichen.

Notation Im Folgenden bezeichnet $\mathbb{K}$ entweder $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Motivation Für die obige Basis B von V mit den Eigenschaften des Korollars gilt offenbar:

$$v \perp b_1, \ldots, b_r \Rightarrow v \in [\{b_{r+1}, \ldots, b_n\}]$$

und

$$b_{r+1}, \ldots, b_n \perp V,$$

also ist $(b_{r+1}, \dots, b_n)$ Basis des Radikalraums $V^\perp$ von (V, σ),

$$V^\perp = [\{b_{r+1}, \dots, b_n\}] \Rightarrow r = \dim V - \dim V^\perp.$$

Insbesondere ist $\dim V^\perp$ und damit r unabhängig von der Basis B.

5.2.7 Satz von Sylvester

Sei V ein $\mathbb{K}$-VR, $\dim V < \infty$, und σ

- symmetrische Bilinearform, wenn $\mathbb{K} = \mathbb{R}$, oder
- Hermitesche Sesquilinearform, wenn $\mathbb{K} = \mathbb{C}$.

Dann gibt es eine direkte Zerlegung von V mit UVR $V_\pm \subset V$,

$$V = V_+ \oplus_\perp V_- \oplus_\perp V^\perp,$$

wobei

$$V_+ \perp V_- \text{ und } \forall v \in V_\pm^\times : \pm\sigma(v, v) > 0.$$

Die *Signatur* $\operatorname{sgn}(\sigma) := (\dim V_+, \dim V_-, \dim V^\perp)$ von σ ist unabhängig von der direkten Zerlegung von V.

Bemerkung & Definition Ist σ nicht-degeneriert, $V^\perp = \{0\}$, so bezeichnet man auch[4]

- das Paar $\operatorname{sgn}(\sigma) = (\dim V_+, \dim V_-)$ als Signatur von σ, und
- die Differenz $\dim V_+ - \dim V_-$ als *Trägheitsindex* von σ.

Die Dimension $\dim V_\pm$ ist auch der *Positivitäts-* bzw. *Negativitätsindex* von σ. Der Satz von Sylvester wird auch „Trägheitssatz von Sylvester“ genannt.

Beweis Sei $B = (b_1, \dots, b_n)$ eine Basis von V und $p, r \in \mathbb{N}$, sodass (siehe Korollar oben)

$$\sigma(b_i, b_j) = \begin{cases} +1 & \text{für } 0 < i = j \leq p \\ -1 & \text{für } p < i = j \leq r \\ 0 & \text{sonst.} \end{cases}$$

Mit

$$V_+ := [\{b_1, \dots, b_p\}] \text{ und } V_- := [\{b_{p+1}, \dots, b_r\}]$$

[4] Die Reihenfolge kann bei verschiedenen Autoren auch jeweils $-$ vor $+$ sein.

erhält man die gewünschte direkte orthogonale Zerlegung von V,

$$V = V_+ \oplus_\perp V_- \oplus_\perp V^\perp.$$

Zur Eindeutigkeit der Signatur $\operatorname{sgn}(\sigma) = (p, r-p, n-r)$:

Seien

$$V = V_+ \oplus_\perp V_- \oplus_\perp V^\perp = \tilde{V}_+ \oplus_\perp \tilde{V}_- \oplus_\perp \tilde{V}^\perp$$

direkte orthogonale Zerlegungen von V mit

$$\pm\sigma(v,v) > 0 \text{ für } \begin{cases} v \in V_\pm^\times \\ v \in \tilde{V}_\pm^\times. \end{cases}$$

Nun gilt

$$\begin{aligned} &\forall v \in V_-^\times : \sigma(v,v) < 0 \\ &\Rightarrow \forall v \in V_- \oplus V^\perp : \sigma(v,v) \leq 0 \end{aligned}$$

und damit, da $\sigma(v,v) > 0$ für $v \in \tilde{V}_+^\times$,

$$v \in (V_- \oplus V^\perp) \cap \tilde{V}_+ \Rightarrow v = 0.$$

Es folgt, mit dem Dimensionssatz, $\tilde{p} \leq p$, da

$$\tilde{p} + (n-p) = \dim \tilde{V}_+ + \dim(V_- \oplus V^\perp) \leq \dim V = n.$$

Vertauscht man die Rollen der Zerlegungen, so erhält man die Ungleichung $p \leq \tilde{p}$ und damit also

$$p = \tilde{p}.$$

Bemerkung Diese Zerlegung $V = V_+ \oplus_\perp V_- \oplus_\perp V^\perp$ ist im Allgemeinen *nicht* eindeutig!

Beispiel Betrachte eine durch ihre Werte auf der Standardbasis $E = (e_1, e_2)$ gegebene symmetrische Bilinearform $\sigma : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$.

1. $S = (\sigma(e_i, e_j))_{i,j \in \{1,2\}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Mit $P := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in Gl(2)$ liefert ein Basiswechsel $B = EP$

$$(\sigma(b_i, b_j))_{i,j \in \{1,2\}} = P^t S P = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

die Signatur $\operatorname{sgn}(\sigma) = (1,1,0) \cong (1,1)$. Jeder weitere Basiswechsel

$$\tilde{B} = BQ \text{ mit } Q = \begin{pmatrix} \cosh(s) & \sinh(s) \\ \sinh(s) & \cosh(s) \end{pmatrix}, s \in \mathbb{R},$$

liefert eine andere Zerlegung, ohne die Gramsche Matrix zu ändern.

2. $S = (\sigma(e_i, e_j))_{i,j\in\{1,2\}} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Der Basiswechsel $B = EP$ wie oben liefert hier

$$(\sigma(b_i, b_j))_{i,j\in\{1,2\}} = P^t SP = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix},$$

also die Signatur $\operatorname{sgn}(\sigma) = (1,0,1)$ von σ. Hier ist $V^\perp = [\{b_2\}]$ durch σ festgelegt, aber jeder Basiswechsel

$$\tilde{B} = BQ \text{ mit } Q = \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}, s \in \mathbb{R}$$

ändert die der Basis zugeordnete Zerlegung – wieder ohne Änderung der Gramschen Matrix.

5.2.8 Bemerkung & Definition

Zur geometrischen Analyse der Lösungsmengen von quadratischen Gleichungen (Quadriken), ist es hilfreich, eine *Äquivalenz* für symmetrische Bilinearformen/Sesquilinearformen σ und σ' auf einem $\mathbb{K}$-VR V einzuführen:

$$\sigma' \sim \sigma :\Leftrightarrow \exists f \in Gl(V) \forall v, w \in V : \sigma'(v,w) = \sigma(f(v), f(w)).$$

Ist $\dim V < \infty$, so liefert der Satz von Sylvester im Falle

- symmetrische Bilinearform auf $\mathbb{R}$-VR, oder
- Hermitesche Sesquilinearform auf $\mathbb{C}$-VR:

Satz: Zwei symmetrische Sesquilinearformen sind genau dann äquivalent, wenn ihre Signaturen übereinstimmen,

$$\sigma' \sim \sigma \Leftrightarrow \operatorname{sgn}(\sigma') = \operatorname{sgn}(\sigma).$$

5.2.9 Definition

Ein *Skalarprodukt* auf einem K-VR V ist eine nicht-degenerierte symmetrische Sesquilinearform

$$\langle .,. \rangle : V \times V \to K, (v,w) \mapsto \langle v, w \rangle.$$

Eine Familie $(e_i)_{i\in I}$ in einem VR $(V, \langle ., . \rangle)$ mit Skalarprodukt heißt *Orthonormalsystem (ONS)*, falls

$$\forall i, j \in I : \langle e_i, e_j \rangle = \pm\delta_{ij};$$

Orthonormalbasis (ONB), falls $(e_i)_{i\in I}$ zusätzlich Basis ist.

Bemerkung Ein ONS ist linear unabhängig:
Für $v = \sum_{i\in I} e_i x_i$ gilt

$$0 = v \Rightarrow \forall i \in I : 0 = \langle e_i, v\rangle = \sum_{j\in I} \langle e_i, e_j\rangle x_j = \pm x_i$$

Ist $\dim V < \infty$, so hat $(V, \langle ., .\rangle)$ jedenfalls eine ONB, wenn

- symmetrische Bilinearform auf einem $\mathbb{K}$-VR ist, oder
- Hermitesche Sesquilinearform auf einem $\mathbb{C}$-VR ist.

Ist $K \neq \mathbb{K}$, so kann die „Normierung“ problematisch sein.

Beispiel Auf dem $\mathbb{R}$-VR der beschränkten Zahlenfolgen:

$$V = \{(x_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}, \exists c \in \mathbb{R} \forall n \in \mathbb{N} : |x_n| < c\},$$

führen wir ein Skalarprodukt, durch Angabe seiner quadratischen Form (Polarisation!) ein:

$$\langle (x_n)_{n\in\mathbb{N}}, (x_n)_{n\in\mathbb{N}}\rangle := \sum_{n\in\mathbb{N}} \left(\frac{x_n}{2^n}\right)^2.$$

Man erhält ein ONS $(e_n)_{n\in\mathbb{N}}$ aus skalierten Standardvektoren

$$e_m : \mathbb{N} \to \mathbb{R}, n \mapsto e_m(n) := 2^n \delta_{mn}.$$

Dieses ONS kann zu einer Basis ergänzt werden (nach BES), nicht jedoch zu einer ONB (in unserem Sinne):

$$\langle e_m, v\rangle = \frac{x_m}{2^m} \text{ für } v = (x_n)_{n\in\mathbb{N}},$$

also gilt

$$\forall m \in \mathbb{N} : v \perp e_m \Rightarrow v = 0.$$

Bemerkung Später wird der Begriff „Basis“ modifiziert, z.B. in der Funktionalanalysis würde man $(e_m)_{m\in\mathbb{N}}$ aus dem Beispiel als „Orthonormalbasis“ bezeichnen.

5.3 Euklidische & unitäre Vektorräume

Bemerkung Die folgende Definition ist nur für Skalarprodukte $\langle .,. \rangle$ sinnvoll, für die

$$v \mapsto \langle v, v \rangle \in T$$

mit einem angeordneten Teilkörper $T \subset K$ des Körpers K (vgl. Abschnitt 1.2). Ein nicht-triviales Beispiel, mit $T = \mathbb{R} \subset \mathbb{C} = K$, ist ein Hermitesches Skalarprodukt:

$$\forall v \in V : \langle v, v \rangle = \overline{\langle v, v \rangle} \Rightarrow \forall v \in V : \langle v, v \rangle \in \mathbb{R} \subset \mathbb{C}.$$

Vereinbarung Im Folgenden beschränken wir uns bis auf Weiteres auf $\mathbb{K}$-VR mit $\langle .,. \rangle$ Hermitsche Sesquilinearform, falls $\mathbb{K} = \mathbb{C}$ (vgl. Satz von Sylvester).

5.3.1 Definition

Ein Skalarprodukt $\langle .,. \rangle$ auf einem $\mathbb{K}$-VR V heißt *positiv definit*, falls

$$\forall v \in V^\times : \langle v, v \rangle > 0;$$

die *induzierte Norm* eines positiv-definiten Skalarprodukts $\langle .,. \rangle$ ist die Abbildung

$$\|.\| : V \to \mathbb{R}, v \mapsto \|v\| := \sqrt{\langle v, v \rangle} \geq 0.$$

Ein $\mathbb{K}$-VR $(V, \langle .,. \rangle)$ mit positiv-definitem Skalarprodukt ist

- ein *Euklidischer Vektorraum*, falls $\mathbb{K} = \mathbb{R}$, und
- ein *unitärer Vektorraum*, falls $\mathbb{K} = \mathbb{C}$ und $\langle .,. \rangle$ Hermitesche Sesquilinearform ist.

5.3.2 Bemerkung & Definition

Ebenso definiert man ein Skalarprodukt als *negativ definit*, falls

$$\forall v \in V^\times : \langle v, v \rangle < 0;$$

$\langle .,. \rangle$ heißt *indefinit*, falls es weder positiv, noch negativ definit ist. Die Definition der induzierten Norm ist nur im positiv definiten Fall sinnvoll.

Beispiel Der *Betrag* einer komplexen Zahl $z = x + iy \in \mathbb{C} \cong_{\mathbb{R}} \mathbb{R}^2$ ist

$$|z| = \sqrt{\bar{z}z} = \sqrt{x^2 + y^2}$$

die vom Standardskalarprodukt auf $\mathbb{R}^2$ induzierte Norm. Insbesondere gilt

$$\forall z \in \mathbb{C} : |\operatorname{Re} z| \le |z|$$

5.3.3 Bemerkung zum Zusammenhang von reellen und komplexen VR

Da $\mathbb{R} \subset \mathbb{C}$ ein Teilkörper ist, kann jeder $\mathbb{C}$-VR V auch als $\mathbb{R}$-VR aufgefasst werden (Einschränkung der Skalarmultiplikation).

Ist nun $S \subset V$ linear unabhängig über $\mathbb{C}$ (in V als $\mathbb{C}$-VR), so ist

$$S' := S \cup Si = S \cup \{si \mid s \in S\}$$

linear unabhängig über $\mathbb{R}$, denn

$$0 = \sum_{s \in S} s x_s + \sum_{s \in S} s i y_s = \sum_{s \in S} s(x_s + i y_s)$$

$$\Rightarrow \forall s \in S : x_s + i y_s = 0 \Rightarrow \forall s \in S : x_s = y_s = 0,$$

d.h. $S' = S \cup S_i$ ist linear unabhängig über $\mathbb{R}$. Insbesondere folgt

$$\dim_{\mathbb{R}} V = 2 \dim_{\mathbb{C}} V.$$

Weiters definiert für ein Hermitesches Skalarprodukt $\langle .,. \rangle$ auf V (als $\mathbb{C}$-VR)

$$\langle .,. \rangle : V \times V \to \mathbb{R}, (v, w) \mapsto \langle v, w \rangle_{\mathbb{R}} := \operatorname{Re}\langle v, w \rangle$$

ein reelles Skalarprodukt auf V (als $\mathbb{R}$-VR), das genau dann positiv definit ist, wenn $\langle .,. \rangle$ positiv definit ist:

$$\forall v \in V : \langle v, v \rangle_{\mathbb{R}} = \langle v, v \rangle.$$

Damit kann man jeden unitären Vektorraum als Euklidischen Vektorraum auffassen:

- mit verschiedenen Skalarprodukten $\langle .,. \rangle$ bzw. $\langle .,. \rangle_{\mathbb{R}}$, aber
- mit gleichen induzierten Normen.

Komplexifizierung Fasst man einen $\mathbb{C}$-VR V als $\mathbb{R}$-VR auf, so liefert Multiplikation mit $i \in \mathbb{C}$ einen Endomorphismus

$$J : V \to V, v \mapsto J(v) := vi$$

mit

$$J^2 = -\operatorname{id}_V .$$

Insbesondere besitzt J keine reellen Eigenwerte; ist $\dim V < \infty$, so folgt damit

$$\dim V = \deg_{\chi_J}(t) = 0 \mod 2.$$

Umgekehrt: Ist V ein $\mathbb{R}$-VR und $J \in \operatorname{End}(V)$ mit $J^2 = -\operatorname{id}_V$ gegeben, so erhält man eine komplexe Skalarmultiplikation

$$\cdot : \mathbb{C} \times V \to V, (z, v) \mapsto vz := vx + J(v)y,$$

für $z = x + iy$. Ist weiter $\langle ., . \rangle$ ein (reelles) Skalarprodukt auf V, das von J erhalten wird,

$$\forall v, w \in V : \langle Jv, Jw \rangle = \langle v, w \rangle,$$

so definiert

$$\langle v, w \rangle_{\mathbb{C}} := \langle v, w \rangle - i \langle v, Jw \rangle$$

ein Hermitesches Skalarprodukt auf dem so konstruierten $\mathbb{C}$-VR.

Beispiel Ist $\langle ., . \rangle$ das kanonische Skalarprodukt auf $\mathbb{R}^2$, mit der Standardbasis (e_1, e_2) als ONB, so definiert (Fortsetzungssatz)

$$J(e_1) = e_2 \text{ und } J(e_2) = -e_1,$$

einen Endomorphismus $J \in \operatorname{End}(\mathbb{R}^2)$ mit

$$J^2 = -\operatorname{id}_{\mathbb{R}^2} \text{ und } \langle Je_i, Je_j \rangle = \langle e_i, e_j \rangle.$$

Vermöge

$$e_1 i := J(e_1) = e_2 \text{ und } e_2 i := J(e_2) = J^2(e_1) = -e_1 = e_1 i^2$$

wird $\mathbb{R}^2$ zu einem eindimensionalen $\mathbb{C}$-VR, $\mathbb{R}^2 = [\{e_1\}]_{\mathbb{C}}$, da

$$e_1 x + e_2 y = e_1 x + J(e_1) y = e_1(x + iy);$$

und

$$\langle e_1 x + e_2 y, e_1 x' + e_2 y' \rangle = \langle e_1(x + iy), e_1(x' + iy') \rangle_{\mathbb{C}} = \overline{(x - iy)}(x' + iy')$$

liefert das kanonische Skalarprodukt auf $\mathbb{C}$, mit dem kanonischen Euklidischen Skalarprodukt von $\mathbb{R}^2$ als Realteil.

5.3.4 Komplexifizierungslemma

Ist $(V, \langle ., . \rangle)$ ein Euklidischer Vektorraum, so liefert

$$(v, w)(x + iy) := (vx - wy, wx + vy)$$

eine komplexe Skalarmultiplikation auf $V_{\mathbb{C}} := V \times V$, und

$$\langle\!\langle ((v, w)), (v', w') \rangle\!\rangle_{\mathbb{C}} := (\langle v, v' \rangle + \langle w, w' \rangle) + i\,(\langle v, w' \rangle - \langle w, v' \rangle)$$

ein Hermitesches Skalarprodukt, das $(V_{\mathbb{C}}, \langle\!\langle ., . \rangle\!\rangle_{\mathbb{C}})$ zu einem unitären VR macht.

Beweis Auf dem Euklidischen VR $(V^2, \langle\!\langle ., . \rangle\!\rangle)$, wobei

$$\langle\!\langle ., . \rangle\!\rangle : V^2 \times V^2 \to \mathbb{R}, \left((v, w), (v', w')\right) \mapsto \langle\!\langle (v, w), (v', w') \rangle\!\rangle := \langle v, v' \rangle + \langle w, w' \rangle,$$

definiere $J \in \operatorname{End}(V^2)$ durch

$$J : V^2 \to V^2, (v, w) \mapsto J\left((v, w)\right) := (-w, v).$$

Offenbar gilt $J^2 = -\operatorname{id}_{V^2}$ und

$$\langle\!\langle J(v, w), J(v', w') \rangle\!\rangle = \langle w, w' \rangle + \langle v, v' \rangle = \langle\!\langle (v, w), (v', w') \rangle\!\rangle,$$

sodass

$$(v, w)(x + iy) = (v, w)x + J(v, w)y = (vx - wy, wx + vy)$$

und

$$\begin{aligned} \langle\!\langle (v, w), (v', w') \rangle\!\rangle_{\mathbb{C}} &= \langle\!\langle (v, w), (v', w') \rangle\!\rangle - i \langle\!\langle (v, w), J(v', w') \rangle\!\rangle \\ &= (\langle v, v' \rangle + \langle w, w' \rangle) - i\,(-\langle v, w' \rangle + \langle w, v' \rangle) \end{aligned}$$

$(V^2, \langle\!\langle ., . \rangle\!\rangle)$ zu einem unitären VR machen, wie vorher.

Bemerkung Mit dem „Komplexifizierungslemma“ kann man jeden Euklidischen VR in einen unitären VR gleicher (komplexer) Dimension einbetten:

$$\dim_{\mathbb{C}} V_{\mathbb{C}} = \frac{1}{2} \dim_{\mathbb{R}} V^2 = \dim_{\mathbb{R}} V.$$

Wichtig für den Zusammenhang zwischen unitären und Euklidischen VR: Die induzierte Norm des Hermitschen Skalarprodukts kann als die eines Euklidischen Skalarprodukts aufgefasst werden.

5.3.5 Definition

Eine Abbildung $\|.\| : V \to \mathbb{R}$ auf einem $\mathbb{K}$-VR V heißt *Norm*, falls

(i) $\forall v \in V^\times : \|v\| > 0$, d.h. $\|.\|$ ist *positiv definit*;

(ii) $\forall v \in V \forall x \in \mathbb{K} : \|vx\| = \|v\| \cdot |x|$, d.h. $\|.\|$ *positiv homogen*;

(iii) $\forall v, w \in V : \|v + w\| \leq \|v\| + \|w\|$, d.h. $\|.\|$ erfüllt die *Dreiecksungleichung.*

Ein Vektorraum mit Norm, $(V, \|.\|)$ heißt *normierter Vektorraum.*

Bemerkung Die von einem positiv definiten Skalarprodukt $\langle ., . \rangle$ induzierte Norm $\|.\|$ erfüllt offenbar (i) und (ii); die Dreiecksungleichung zeigen wir unten.

Cauchy-Schwarzsche Ungleichung Ist $(V, \langle ., . \rangle)$ Euklidisch oder unitär, so gilt[5]

$$\forall v, w \in V : |\langle v, w \rangle|^2 \leq \langle v, v \rangle \langle w, w \rangle$$

Beweis Seien $v, w \in V$, o.B.d.A $v \neq 0$. Wir bestimmen das Minimum der Funktion im Euklidischen Fall (unitärer Fall in der Übung)

$$\mathbb{R} \ni s \mapsto g(s) := \langle vs - w, vs - w \rangle.$$

Einsetzen des kritischen Punktes,

$$0 = g'(s) = 2\langle v, v \rangle s - (\langle v, w \rangle + \langle w, v \rangle) = 2(\langle v, v \rangle s - \mathrm{Re}\langle v, w \rangle)$$

$$\Rightarrow s = \frac{\langle v, w \rangle}{\langle v, v \rangle}$$

liefert

$$0 \leq g(s) = \langle v, v \rangle \frac{\langle v, w \rangle^2}{\langle v, v \rangle^2} - 2\langle v, w \rangle \frac{\langle v, w \rangle}{\langle v, v \rangle} + \langle w, w \rangle$$

$$= \frac{1}{\langle v, v \rangle} \Big(-\langle v, w \rangle^2 + \langle v, v \rangle \langle w, w \rangle \Big) \Leftrightarrow 0 \leq -\langle v, w \rangle^2 + \langle v, v \rangle \langle w, w \rangle.$$

5.3.6 Korollar

Die induzierte Norm in $(V, \langle ., . \rangle)$ erfüllt die Dreiecksungleichung.

[5] Im Euklidischen Fall ist der Betrag offenbar überflüssig.

Beweis Ist $(V, \langle ., . \rangle)$ Euklidischer VR, so gilt für $v, w \in V$:

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

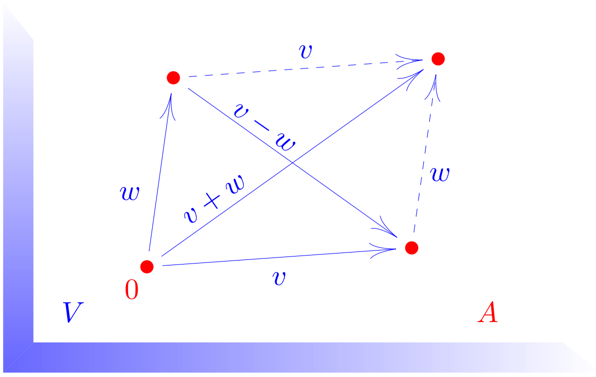
Bemerkung Das Skalarprodukt eines Euklidischen VR $(V, \langle ., . \rangle)$ kann (Polarisation) aus seiner induzierten Norm rekonstruiert werden.

Nicht jede Norm ist jedoch von einem Skalarprodukt induziert (vgl. Aufgabe 56). Hinreichende (Satz von Jordan-von Neumann) und notwendige Bedingung ist die Parallelogrammgleichung:

5.3.7 Parallelogrammgleichung

Für die induzierte Norm $\|.\|$ von $(V, \langle ., . \rangle)$ gilt:

$$\forall v, w \in V : \|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2$$



Beweis Rechnung, wie bei Polarisation.

Beispiel Für die induzierte Norm des kanonischen Skalarprodukts auf $\mathbb{R}^n$ gilt die Parallelogrammgleichung:

$$\sum_{i=1}^{n}(x_i+y_i)^2 + \sum_{i=1}^{n}(x_i-y_i)^2 = 2\sum_{i=1}^{n} x_i^2 + 2\sum_{i=1}^{n} y_i^2.$$

Für die durch

$$\|(x_i)_{i\in\{1,\dots,n\}}\|_1 = \sum_{i=1}^{n}|x_i|$$

auf $\mathbb{R}^n$ definierte Norm $\|.\|_1$ gilt sie nicht; diese Norm ist also nicht induzierte Norm eines Euklidischen Skalarprodukts auf $\mathbb{R}^n$.

Beispiel Auf dem Raum $C^0([0,1])$ der stetigen Funktionen auf $[0,1]$ definiert

$$\|.\|_\infty : C^0([0,1]) \to \mathbb{R}, f \mapsto \|f\|_\infty := \max_{x\in[0,1]} |f(x)|$$

die *Maximumsnorm* (vlg. gleichmäßige Konvergenz).

Für $f, g \in C^0([0,1])$,

$$f(x) := 1 - x \text{ und } g(x) = x$$

ist dann

$$\|f\|_\infty = \|g\|_\infty = \|f+g\|_\infty = \|f-g\|_\infty = 1$$

womit die Parallelogrammgleichung offenbar nicht erfüllt, und die Norm keine induzierte Norm eines Skalarprodukts ist.

5.4 Euklidische Geometrie

5.4.1 Definition

Ein *Euklidischer Raum* ist eine affiner Raum (A, V, τ) über einem Euklidischen Vektorraum $(V, \langle ., . \rangle)$ mit induzierter Norm $\|.\|$.

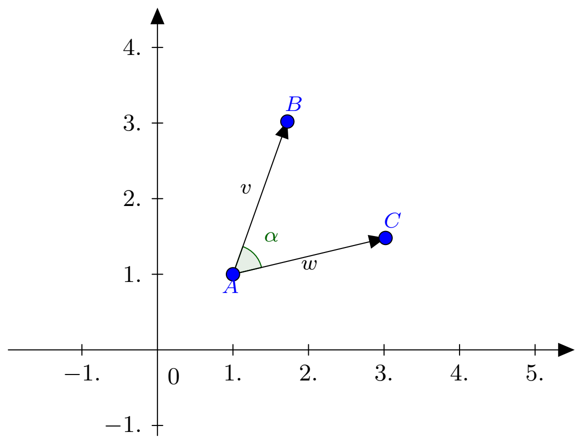
- Die *Länge* eines Vektors $v \in V$ ist seine Norm, der *Abstand* zweier Punkte $a, b \in A$ ist die Länge ihres Verbindungsvektors,

$$d(a,b) := \|b - a\| = \sqrt{\langle b-a, b-a \rangle}.$$

- Der *Winkel* $\alpha \in [0, \pi]$ zweier Vektoren $v, w \in V^\times$ ist durch die Gleichung

$$\langle v, w \rangle = \|v\| \cdot \|w\| \cdot \cos\alpha$$

definiert; der *Winkel* (am Punkt a) in einem nicht-degenerierten Dreieck $\{a, b, c\} \subset A$ ist der Winkel der beiden Seitenvektoren $v = b - a$ und $w = c - a$.



Bemerkung Nach der Cauchy-Schwarzschen Ungleichung ist für $v, w \in V^\times$

$$\frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \in [-1, 1];$$

andererseits ist

$$\cos : [0, \pi] \to [-1, 1] \text{ bijektiv}$$

Damit ist der Winkel von Vektoren bzw. im Dreieck wohldefiniert.

5.4.2 Definition

Eine affine Transformation eines Euklidischen Raumes heißt

- *Kongruenzabbildung* oder *Isometrie*, falls sie Abstandstreu ist,
- *Ähnlichkeitstransformation*, falls sie winkeltreu ist.

Bemerkung Jede Kongruenzabbildung ist Ähnlichkeitstransformation (Polarisation).

Bemerkung Offenbar bilden die Kongruenz- bzw. Ähnlichkeitsabbildungen eines Euklidischen Raumes A auf A operierende (Transformations-)Gruppen.

5.4.3 Definition (Geometrie)

Die auf einem Euklidischen Raum operierende Gruppe der Kongruenzabbildungen bestimmt eine Euklidische Geometrie.

Die Gruppe der Ähnlichkeitstransformationen eines Euklidischen Raumes A bestimmt eine Ähnlichkeitsgeometrie.

Beispiel Jede Translation $\tau_v : A \to A$ ist eine Isometrie: Für $a, b \in A$ gilt

$$\exists! w \in V : b = \tau_w(a)$$

d.h. $w = b - a$; also

$$\tau_v(b) = \tau_v(\tau_w(a)) = \tau_{v+w}(a) = \tau_w(\tau_v(a))$$

d.h. $w = \tau_v(b) - \tau_v(a)$. Damit folgt:

$$\|\tau_v(b) - \tau_v(a)\| = \|w\| = \|b - a\|$$

d.h. τ_v ist abstandstreu, da $a, b \in A$ beliebig waren.

Beispiel Die Streckung mit Zentrum $o \in A$ um den Faktor $s \in \mathbb{R}^\times$,

$$o + v = a \overset{\delta_s}{\mapsto} \delta_s(a) = \delta_s(o + v) := o + vs$$

ist winkeltreu, denn für $a = o + v, b = o + w$ gilt

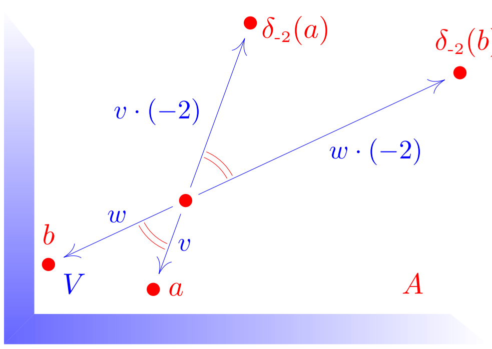
$$\delta_s(b) - \delta_s(a) = (o + ws) - (o + vs) = \cdots = (w - v)s$$

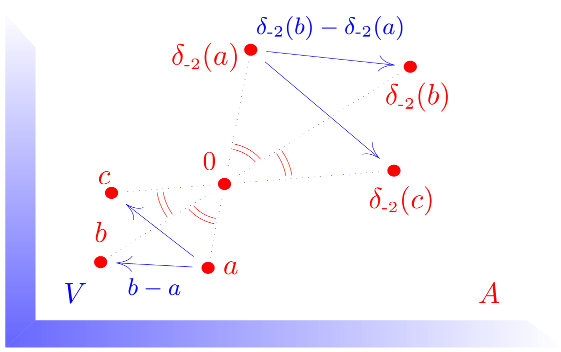
und damit für drei paarweise verschiedene Punkte $a, b, c \in A$

$$\cos\alpha = \frac{\langle \delta_s(b) - \delta_s(a), \delta_s(c) - \delta_s(a)\rangle}{\|\delta_s(b) - \delta_s(a)\|\|\delta_s(c) - \delta_s(a)\|} = \frac{\langle (b-a)s, (c-a)s\rangle}{\|(b-a)s\|\|(c-a)s\|} = \frac{s^2}{|s^2|} \cdot \frac{\langle b-a, c-a\rangle}{\|b-a\|\|c-a\|}$$

d.h. δ_s ist winkeltreu; andererseits ist δ_s für $s \neq \pm 1$ nicht abstandstreu. Ist $a \neq b$, so gilt dann

$$\|\delta_s(b) - \delta_s(a)\| = \|b - a\| \cdot |s| \neq \|b - a\|.$$





Zur Erinnerung Jede affine Abbildung $\alpha : A \to A'$ besitzt einen (eindeutigen) *linearen Anteil* $\lambda : V \to V'$, sodass

$$\forall a \in A \forall v \in V : \alpha(a+v) = \alpha(a) + \lambda(v);$$

ist α eine affine Transformation, so ist $\lambda \in Gl(V)$.

Bemerkung Jede Ähnlichkeitstransformation ist Komposition einer Streckung und einer Kongruenzabbildung.

Nämlich: Ist α Ähnlichkeitstransformation mit linearem Anteil $\lambda \in Gl(V)$, so erhält λ Winkel von Vektoren, insbesondere also Orthogonalität. Nun wähle $w \in V^\times$ und setze

$$s := \frac{\|w\|}{\|\lambda w\|}.$$

Ist dann $v \in V$ mit $\|v\| = \|w\|$, so folgt

$$v + w \perp v - w \Rightarrow \lambda(v+w) \perp \lambda(v-w) \Rightarrow \|\lambda(v)\| = \|\lambda(w)\|,$$

also

$$\forall v \in V^\times : \frac{\|\lambda(v)\|}{\|v\|} = \|\lambda(v\frac{\|w\|}{\|v\|})\|\frac{1}{\|w\|} = \frac{\|\lambda(w)\|}{\|w\|} = \frac{1}{s}.$$

Mit einem beliebigen Streckungszentrum $o \in A$ erhält man also eine Isometrie durch

$$\delta_s \circ \alpha : A \to A.$$

Beispiel Eine *nicht-triviale* Scherung ist *keine* Ähnlichkeitstransformation. Beweis in der Übung.

5.4.4 Lemma & Definition

Eine affine Transformation $\alpha : A \to A$ eines Euklidischen Raumes A ist genau dann eine Kongruenzabbildung, wenn ihr linearer Anteil λ *orthogonal* ist:

$$\lambda \in O(V) := \{f \in Gl(V) \mid \forall v, w \in V : \langle f(v), f(w)\rangle = \langle v, w\rangle\}.$$

$O(V)$ heißt die *orthonogale Gruppe* von $(V, \langle ., .\rangle)$.

Bemerkung $O(V) \subset Gl(V)$ ist eine Gruppe. Beweis in der Übung.

Bemerkung Ist $f \in \mathrm{End}(V)$, so folgt die Injektivität von f aus

$$\forall v, w \in V : \langle f(v), f(w)\rangle = \langle v, w\rangle.$$

Aus $f(v) = 0$ folgt nämlich

$$0 = \|f(v)\| = 0 = \|v\| \Rightarrow v = 0, \text{ da } \langle .,. \rangle \text{ pos. definit.}$$

Ist $\dim V < \infty$, so folgt mit dem Rangsatz, $\dim V = \operatorname{rg} f + \operatorname{def} f = \operatorname{rg}$, dass $f \in Gl(V)$.
Im Fall $\dim V = \infty$ ist f nicht notwendigerweise surjektiv, wie der *Shiftoperator*

$$f \in \operatorname{End}(\mathbb{R}^{\mathbb{N}}), \forall n \in \mathbb{N} : f(e_n) = e_{n+1}$$

zeigt.

Beweis (Lemma) Sei (A, V, τ) Euklidischer Raum über einem Euklidischen VR $(V, \langle .,. \rangle)$ und $\alpha : A \to A$ Affinität mit linearem Anteil $\lambda \in Gl(V)$. Dann ist α genau dann Isometrie, wenn

$$\forall a, b \in A : \|\lambda(b - a)\| = \|\alpha(b) - \alpha(a)\| = \|b - a\|,$$

also (Polarisation), wenn $\lambda \in O(V)$.

5.4.5 Definition

Ist $(V, \langle .,. \rangle)$ unitärer VR, so heißt $f \in Gl(V)$ mit

$$\forall v, w \in V : \langle f(v), f(w) \rangle = \langle v, w \rangle$$

unitär; die *unitäre Gruppe* von $(V, \langle .,. \rangle)$ ist die Gruppe

$$U(V) := \{f \in Gl(V) \mid \forall v, w \in V : \langle f(v), f(w) \rangle = \langle v, w \rangle\}.$$

5.4.6 Schulgeometrie

Betrachte eine Euklidische Ebene A^2 über Euklidischem VR $(\mathbb{R}^2, \langle .,. \rangle)$ mit kanonischem Skalarprodukt $\forall i, j \in \{1, 2\} : \langle e_i, e_j \rangle = \delta_{ij}$.
Weiter (vgl. Abschnitt 5.3) bezeichne $J \in \operatorname{End}(\mathbb{R}^2)$ den durch

$$J(e_1) = e_2 \text{ und } J(e_2) = 1$$

definierten Endomorphismus, also eine „90°-Drehung“, bzw. die $\mathbb{R}^2$ mit $\mathbb{C}$ identifizierende komplexe Multiplikation mit i,

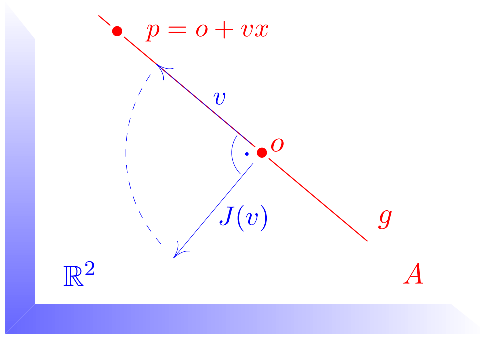
$$v(x + iy) = vx + J(v)y \text{ für } \begin{cases} v \in \mathbb{R}^2 \\ (x + iy) \in \mathbb{C}. \end{cases}$$

Man bemerke: Für $v \in \mathbb{R}^2 \setminus \{0\}$ ist $\{Jv\}^\perp = [v]$ und damit

$$\forall w \in \mathbb{R}^2 : w \perp Jv \Leftrightarrow w \parallel v.$$

So ermöglicht J einen einfachen Wechsel zwischen *parametrischer* und *impliziter Darstellung* (e.g. *Hessesche Normalform*) einer Geraden

$$g = \{p = o + vx \mid x \in \mathbb{R}\} \Leftrightarrow$$
$$g = \{p \in A^2 \mid \langle p - o, Jv\rangle = 0\}$$



5.4.7 Definition

Ein *Kreis* mit *Mittelpunkt* $z \in A^2$ und *Radius* $r \geq 0$ ist die Menge

$$k = \{p \in A^2 \mid \|p - z\| = r\}.$$

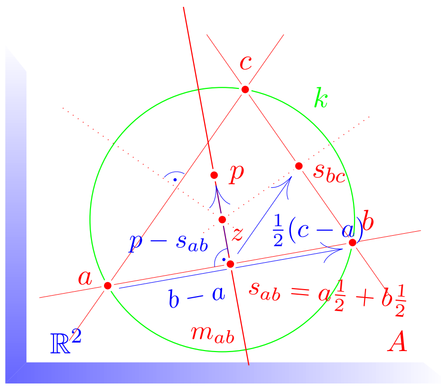
Bemerkung Es ist mitunter sinnvoll, Punkte als Kreise mit Radius $r = 0$ zu betrachten.

5.4.8 Umkreissatz

Sei $\{a, b, c\} \subset A^2$ ein nicht-degeneriertes Dreieck. Dann gibt es genau einen Kreis $k \subset A^2$, den *Umkreis* des Dreiecks, der die Eckpunkte a, b und c des Dreiecks enthält. Sein Mittelpunkt ist der Schnittpunkt der drei *Streckensymmetralen/Mittelsenkrechten* m_{ab}, m_{bc} und m_{ca} des Dreiecks, wobei

$$m_{ab} = \{p \in A^2 \mid \langle p - s_{ab}, b - a\rangle = 0\}$$

$$\text{mit } s_{ab} = a\frac{1}{2} + b\frac{1}{2} \text{ etc.}$$



Beweis Definiere

$$g_{ab} : A^2 \to \mathbb{R}, p \mapsto g_{ab}(p) := 2\langle p - s_{ab}, b - a\rangle$$

und analog g_{bc} und g_{ca} (zyklische Vertauschung). Für $p \in A^2$ gilt dann mit

$$g_{ab}(p) \overset{!}{=} \langle (p-a) + (p-b), (p-a) - (p-b)\rangle = \|p-a\|^2 - \|p-b\|^2 \tag{$*$}$$

damit folgt

$$\forall p \in A^2 : (g_{ab} + g_{bc} + g_{ca})(p) = 0,$$

also

$$p \in m_{ab} \cap m_{bc} \Rightarrow p \in m_{ca}.$$

Nun ist

$$m_{ab} = \{p(x) = s_{ab} + J(b-a)x \mid x \in \mathbb{R}\}$$

mit $J(b-a) \not\parallel b-c$, da das Dreieck $\{a, b, c\}$ nicht-degeneriert ist. Dies liefert einen eindeutigen Schnittpunkt $z \in p(x) \in m_{ab} \cap m_{bc}$ als Lösung der linearen Gleichung

$$0 = g_{bc}(p(x)) = 2\langle s_{ab} + J(b-a)x - s_{bc}, c - b\rangle$$

$$= 2\langle J(b-a), c-b\rangle x + \langle a - c, c - b\rangle.$$

Wegen ($*$) gilt nun für diesen Schnittpunkt z

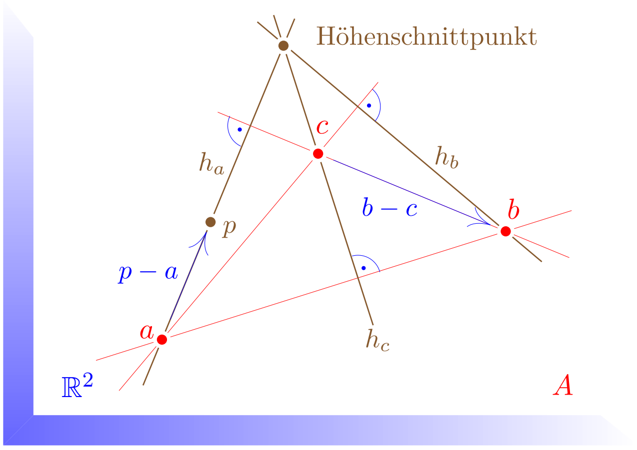
$$\|z-a\| = \|z-b\| = \|z-c\| \tag{$**$}$$

d.h. a, b und c liegen auf einem Kreis mit Mittelpunkt z. Andererseits: Wegen ($*$) impliziert ($**$), dass $z \in m_{ab} \cap m_{bc}$, womit die Eindeutigkeit von z und damit des Umkreises folgt.

5.4.9 Höhensatz

Die *Höhen* h_a, h_b und h_c eines nicht-degenerierten Dreiecks $\{a, b, c\} \subset A^2$ schneiden sich in einem Punkt, dem *Höhenschnittpunkt*, wobei

$$h_a = \{p \in A^2 \mid \langle p - a, b - c\rangle = 0\}, \text{ etc.}$$



Beweis in der Übung, analog zum Umkreissatz.

5.4.10 Euler-Gerade

Seien s, h und z Schwerpunkt, Höhenschnittpunkt und Umkreismittelpunkt eines nicht-degenerierten Dreiecks $a, b, c \subset A^2$. Dann gilt

$$s = z\frac{2}{3} + h\frac{1}{3}.$$

Ist $s \neq z$, so liegen die drei Punkte also auf einer eindeutig bestimmten Geraden, *Euler-Geraden*, mit einem Teilverhältnis $(zs : hs) = -\frac{1}{2}$. Beweis in der Übung.

5.4.11 Satz von Pythagoras

In einem Dreieck $\{a, b, c\} \subset A^2$ mit einem rechten Winkel $\alpha = \frac{\pi}{2}$ bei a gilt stets

$$\|c - a\|^2 + \|a - b\|^2 = \|c - b\|^2.$$

Beweis Offenbar gilt $c - b = (c - a) + (a - b)$, daher

$$\|c - b\|^2 = \|c - a\|^2 + 2\langle c - a, a - b\rangle + \|a - b\|^2 = \|c - a\|^2 + \|a - b\|^2.$$

Bemerkung Für allgemeine Dreiecke liefert die gleiche Rechnung den Cosinussatz:

$$\|b - c\|^2 = \|c - a\|^2 + \|a - b\|^2 - 2\|c - a\|\|a - b\| \cos \alpha.$$

Bemerkung Ist $(o; e_1, e_2)$ ein affines Bezugssystem in A^2 mit

$$e_1 \perp e_2 \text{ und } \|e_1\| = \|e_2\| = 1,$$

so ist jeder Punkt $a \in A^2$ Eckpunkt eines *rechtwinkligen* Dreiecks

$$\{o, i + e_1x_1, o + e_1x_1 + e_2x_2\} \text{ für } a = o + e_1x_1 + e_2x_2;$$

der Abstand vom Ursprung ist also (Pythagoras)

$$\|a - o\| = \sqrt{x_1^2 + x_2^2}.$$

Wegen seiner Translationsinvarianz kann der Abstand zwischen beliebigen Punkten genau so berechnet werden.

5.4.12 Definition

Ein *kartesisches Bezugssystem* (o, E) eines Euklidischen Raumes (A, V, τ) über einem Euklidischen VR $(V, \langle ., . \rangle)$ besteht aus einem Ursprung $o \in A$ und einer ONB E von $(V, \langle ., . \rangle)$.

Bemerkung In jedem endlichdimensionalen Euklidischen Raum gibt es ein kartesisches Bezugssystem, im Allgemeinen ist dies nicht so (vgl. Abschnitt 5.2).

5.4.13 Lemma

Ist $(o; E)$ mit $E = (e_i)_{i \in I}$ kartesisches Bezugssystem eines Euklidischen Raumes (A, V, τ) über $(V, \langle ., . \rangle)$, so ist

$$\forall a \in A : a = o + \sum_{i \in I} e_i \langle e_i, a - o \rangle$$

Beweis Da E Basis ist, existiert zu $a \in A$ eine Familie $(x_i)_{i \in I}$ in $\mathbb{R}$ mit

$$a = o + \sum_{i \in I} e_i x_i,$$

wobei

$$\forall i \in I : \langle e_i, a - o \rangle = \langle e_i, \sum_{j \in I} e_j x_j \rangle = \sum_{j \in I} \delta_{ij} x_j = x_i.$$

5.5 Orthogonalprojektion

5.5.1 Definition

Sei (A, V, τ) ein Euklidischer Raum über einem Euklidischen VR $(V, \langle ., . \rangle)$. Dann heißt

- $p \in \mathrm{End}(V)$ *Orthogonalprojektion*, falls p Projektion ist, $p^2 = p$, mit

$$\ker p \perp p(V)$$

- $\pi : A \to A$ *Orthogonalprojektion*, falls π Parallelprojektion ist, mit einer Orthogonalprojektion $p \in End(V)$ als linearem Anteil.

Bemerkung Ist $p \in \mathrm{End}(V)$ Orthogonalprojektion, so ist auch die komplementäre Projektion $p' = \mathrm{id}_V - p$ Orthogonalprojektion, denn

$$\ker p' = p(V) \perp \ker p = p'(V.)$$

Bemerkung Ist $(o; E)$ mit $E = (e_i)_{i \in I}$ kartesisches Bezugssystem eines Euklidischen Raumes A und $J \subset I$, so liefert

$$\pi : A \to A, a = o + v \mapsto o + p(v) := o + \sum_{i \in J} e_i \langle e_i, v \rangle$$

eine Orthogonalprojektion von A auf

$$\pi(A) = o + p(V) = o + [(e_i)_{i \in J}].$$

5.5.2 Gram-Schmidtsches Orthogonalisierungsverfahren

Sei $(V, \langle ., . \rangle)$ ein Euklidischer VR und $(v_1, \dots, v_n)$ linear unabhängig in V; dann existiert ein ONS $(e_1, \dots, e_n)$ mit

$$[(v_1, \dots, v_k)] = [(e_1, \dots, e_k)] \text{ und } \langle e_k, v_k \rangle > 0 \text{ für } k = 1, \dots, n \tag{$*$}$$

Beweis Induktion über n. Ist $n = 1$, so liefert $e_1 := v_1 \cdot \frac{1}{\|v_1\|}$ das gewünschte Orthonormalsystem.

Ist $(v_1, \dots, v_{n+1})$ linear unabhängig und (nach Induktions-Annahme) $(e_1, \dots, e_n)$ ONS mit

$$V_n := [(v_1, \dots, v_k)] = [(e_1, \dots, e_k)] \text{ und } \langle e_k, v_k \rangle > 0 \text{ für } k = 1, \dots, n,$$

so setzen wir

$$p : V \to V, v \mapsto p(v) := v - \sum_{i=1}^{n} e_i \langle e_i, v \rangle \in \{e_1, \dots, e_n\}^{\perp};$$

da $v_{n+1} \notin V_n$ ist

$$p(v_{n+1}) \neq 0, \text{ und } e_{n+1} := p(v_{n+1}) \frac{1}{\|p(v_{n+1})\|}$$

ergänzt dann $(e_1, \dots, e_n)$ zum gesuchten Orthonormalsystem.

Bemerkung Das ONS $(e_1, \dots, e_n)$ im Gram-Schmidtschen Verfahren ist durch die Bedingungen $(*)$ eindeutig festgelegt.

Bemerkung Der Beweis lässt sich wörtlich auf unitäre VR übertragen.

5.5.3 Korollar & Definition

Ist $U \subset V$ UVR eines Euklidischen VR (oder unitären VR) $(V, \langle ., . \rangle)$ mit $\dim V < \infty$, so gilt

$$V = U \oplus U^{\perp}.$$

Der UVR $U^{\perp}$ heißt dann das *orthogonale Komplement* von U (in $(V, \langle ., . \rangle)$).

Beweis Für $v \in U \cap U^{\perp}$ ist $\langle v, v \rangle = 0$, also $v = 0$, da das Skalarprodukt positiv definit ist. Sei $(e_1, \dots, e_k)$ ONB von U (Gram-Schmidt) und

$$p : V \to V, v \mapsto p(v) := \sum_{i=1}^{k} e_i \langle e_i, v \rangle \in U.$$

Wegen

$$\langle e_j, v - p(v) \rangle = \langle e_j, v \rangle - \sum_{i=1}^{k} \delta_{ij} \langle e_i, v \rangle = 0$$

für $j = 1, \dots, k$ ist dann

$$\forall v \in V : v = p(v) + (v - p(v)) \in U + U^{\perp},$$

also $V = U + U^{\perp}$

Bemerkung Die Einschränkung $\dim V < \infty$ wurde nur benutzt, um die Orthogonalprojektion $p \in \mathrm{End}(V)$ zu definieren/konstruieren. Insbesondere reicht es, $\dim U < \infty$ anzunehmen.

Bemerkung Ist $\dim V < \infty$, so ist $U^{\perp\perp} = U$.

5.5.4 Beispiel & Definition

Ist $p \in \mathrm{End}(V)$ eine Projektion und $p' = \mathrm{id}_V - p$, so erhält man eine Involution

$$s := p - p' \in \mathrm{End}(V)$$

Im Falle einer Orthogonalprojektion p nennt man die zugehörige Transformation

$$\sigma : A \to A, o + v \mapsto \sigma(o + v) := o + s(v)$$

eines Euklidischen Raumes eine *Spiegelung*: σ ist eine Isometrie, da

$$\forall v \in V : \|p(v) \pm p'(v)\|^2 = \begin{cases} \|v\|^2 & \text{für } + \\ \|s(v)\|^2 & \text{für } - \end{cases} = \|p(v)\|^2 + \|p'(v)\|^2 \pm 2 \underbrace{\langle p(v), p'(v) \rangle}_{=0, \text{ da } p(V) \perp p'(V)}$$

Bemerkung Jede Kongruenzabbildung eines endlichdimensionalen Euklidischen Raumes ist Komposition von Spiegelungen.

5.5.5 Beispiel & Definition

Ist A^2 Euklidische Ebene mit kartesischem Bezugssystem $(o; e_1, e_2)$ und $J \in \mathrm{End}(\mathbb{R}^2)$ wie oben, $J(e_1) = e_2$ und $J(e_2) = -e_1$ so liefert

$$\rho_\theta : A^2 \to A^2, o + v \mapsto \rho_\vartheta(o + v) := o + v \cos \vartheta + J(v) \sin \vartheta$$

eine *Drehung* mit *Zentrum* $o \in A^2$ und *Drehwinkel* $\vartheta \in \mathbb{R}$. Die affine Abbildung ρ_ϑ ist dann Komposition zweier Spiegelungen,

$$\rho_\vartheta = \sigma' \circ \sigma,$$

die durch ihre Fixpunktgeraden festgelegt sind:

$$g = o + [e_1] \text{ und } g' = o + [e_1'] \text{ mit } e_1' = e_1 \cos \frac{\vartheta}{2} + e_s \sin \frac{\vartheta}{2}.$$

5.5.6 Lemma

Eine Projektion $p \in \mathrm{End}(V)$ ist genau dann Orthogonalprojektion, wenn

$$\forall v, w \in V : \langle p(v), w \rangle = \langle v, p(w) \rangle$$

Beweis Sei $p \in \mathrm{End}(V)$ Projektion und $p' = \mathrm{id}_V - p$ die komplementäre Projektion mit

$$\ker p = p'(V) \text{ und } p(V) = \ker p'.$$

Ist p Orthogonalprojektion, $p(V) \perp \ker p = p'(V)$, so ist für $v, w \in V$

$$\langle p(v), w \rangle - \langle v, p(w) \rangle = \langle p(v), p(w) \rangle + \underbrace{\langle p(v), p'(w) \rangle}_{\perp} - \langle p(v), p(w) \rangle - \underbrace{\langle p'(v), p(w) \rangle}_{\perp} = 0.$$

Gilt andererseits für $v, w \in V$, also insbesondere für $v \in p(V), w \in \ker p$, stets

$$0 = \langle p(v), w \rangle - \langle v, p(w) \rangle = \langle v, w \rangle,$$

so ist $p(V) \perp \ker p$, also p Orthogonalprojektion.

6 Struktursätze für Endomorphismen

6.1 Adjungierte & duale Abbildungen

Zunächst: Ziel dieses Kapitels ist besseres, strukturelles Verständnis der Bedingungen für orthogonale Transformationen bzw. Orthogonalprojektionen:

$$\forall v, w \in V : \langle f(v), f(w)\rangle = \langle v, w\rangle \text{ bzw. } \langle p(v), w\rangle = \langle v, p(w)\rangle.$$

Motivation Jede Sesquilinearform $\sigma : V \times V \to K$ liefert eine (semi-)lineare Abbildung

$$V \ni v \mapsto \sigma(v, .) \in V^*.$$

Ist die Sesquilinearform σ symmetrisch und nicht-degeneriert, d.h. ein Skalarprodukt auf V, so ist die Abbildung injektiv:

$$\sigma(v, .) = 0 \Rightarrow v = 0 \Rightarrow v \in V^\perp = \{0\}.$$

Ist die Abbildung auch surjektiv, so kann man sie benutzen, um V^* und V zu identifizieren,

$$V^* \simeq V.$$

6.1.1 Rieszsches Darstellungslemma

Sei $(V, \langle ., .\rangle)$ ein K-VR mit Skalarprodukt. Die *kanonische Injektion*

$$\phi : V \to V^*, v \mapsto \phi(v) := \langle v, .\rangle.$$

ist semi-linear und injektiv. Ist $\dim V < \infty$, so ist ϕ auch surjektiv; wir nennen dann

- $\nabla w := \phi^{-1}(w)$ den *Gradienten* von $w \in V^*$, und
- $\phi : V \to V^*$ die *kanonische Identifikation* von $(V, \langle ., .\rangle)$ mit V^*.

Beweis Semi-Linearität und Injektivität folgen sofort aus den Eigenschaften des Skalarprodukts.[1]

[1] Semi-Linearität in der linken Komponente; Nicht-Degeneriertheit impliziert Injektvität.

Ist $\dim V < \infty$, so ist $\phi : V \to V^*$ wegen $\dim V^* = \dim V$ und der Injektivität auch surjektiv.

Bemerkung Dies ist eine „kleine“ Version des Rieszschen Darstellungssatzes

$$\forall \omega \in V^* \exists! w \in V : \omega = \langle w, . \rangle.$$

Der „richtige“ Satz schränkt die Dimension nicht ein, und ist ein wichtiges Hilfsmittel in der Funktional-Analysis.

Bemerkung Ist $E = (e_1, \ldots, e_n)$ ONB von $(V, \langle ., . \rangle)$, so gilt für die Vektoren der dualen Basis $E^* = (e_1^*, \ldots, e_n^*)$

$$\nabla e_i^* = e_i \frac{1}{\langle e_i, e_i \rangle} = \pm e_i,$$

da für $j = 1, \ldots, n$ gilt:

$$\langle e_i \frac{1}{\langle e_i, e_i \rangle}, e_j \rangle = \delta_{ij} = e_i^*(e_j),$$

also

$$\phi(e_i \frac{1}{\langle e_i, e_i \rangle}) = e_i^*.$$

Insbesondere gilt im Falle eines Euklidischen VR

$$\forall i = 1, \ldots, n : \nabla e_i^* = e_i \Leftrightarrow \phi(e_i) = e_i^*,$$

d.h. ϕ realisiert den früher diskutierten (vgl. Abschnitt 1.4) durch duale Basen gegebenen Isomorphismus – im Falle von ONB.

6.1.2 Korollar & Definition

Sind $(W, \langle\langle ., . \rangle\rangle)$ und $(V, \langle ., . \rangle)$ Vektorräume mit Skalarprodukten, $\dim W < \infty$ und $f \in \mathrm{Hom}(W, V)$, so hat f eine eindeutige *Adjungierte* $f^* \in \mathrm{Hom}(V, W)$; dabei ist f^* *adjungiert zu* f, falls

$$\forall v \in V \forall w \in W : \langle\langle f^*(v), w \rangle\rangle = \langle v, f(w) \rangle.$$

Achtung: V und W sind VR über dem gleichen Körper K; die Skalarprodukte sind sesquilinear bzgl. des gleichen Körperautomorphismus!

Beweis Für jedes $v \in V$ definiert

$$\omega_v : W \to K, w \mapsto \omega_v(w) := \langle v, f(w) \rangle$$

eine Linearform $\omega_v \in W^*$; nach Rieszschem Darstellungslemma erhält man daher eine eindeutige Abbildung

$$f^* : V \to W, v \mapsto f^*(v) := \nabla \omega_v.$$

Die Linearität von f^* folgt aus der dualen Abbildung, siehe unten.

Bemerkung Offenbar (Symmetrie) ist $f^{**} = f$, wenn $f^{**} := (f^*)^*$ existiert.

6.1.3 Definition & Lemma

Ist $f \in \mathrm{Hom}(W, V)$, so heißt $f^t \in \mathrm{Hom}(V^*, W^*)$,

$$f^t : V^* \to W^*, \nu \mapsto f^t(\nu) := \nu \circ f$$

zu f *transponiert* oder *dual*. Sind $\langle\langle ., . \rangle\rangle$ und $\langle ., . \rangle$ Skalarprodukte auf W bzw. V und

$$\psi : W \to W^* \text{ und } \phi : V \to V^*$$

die zugehörigen kanonischen Injektionen, und ist $f^* \in \mathrm{Hom}(V, W)$ adjungiert zu f, so gilt

$$\psi \circ f^* = f^t \circ \phi.$$

Beweis Für $v \in V$ und $w \in W$ gilt:

$$((\psi \circ f^*)(v))\,(w) = \langle\langle f^*(v), w \rangle\rangle$$

$$\Big((f^t \circ \phi)(v)\Big)\,(w) = (\phi(v) \circ f)(w) = \langle v, f(w) \rangle$$

und nach Definition der Adjungierten folgt die Gleichheit.

Bemerkung Ist $\dim W < \infty$, so ist ψ bijektiv und das Resultat des Lemmas kann als Definition dienen:

$$f^* := \psi^{-1} \circ f^t \circ \phi.$$

Wegen $f^t \in \mathrm{Hom}(V^*, W^*)$ folgt damit auch $f^* \in \mathrm{Hom}(V, W)$.

Bemerkung $f \in \mathrm{Hom}(W, V)$ hat *immer* eine Transponierte, eine Adjungierte aber nur unter bestimmten Voraussetzungen, z.B. wenn $\dim W < \infty$.

Bemerkung Oft wird die Transponierte/Duale f^t auch mit f^* bezeichnet.

Buchhaltung Sind $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_m)$ Basen von V bzw. W und $f \in \mathrm{Hom}(W, V)$, so gilt:

$$\xi_{B^*}^{C^*}(f^t) = \left(\xi_C^B(f)\right)^t.$$

Sind $(V, \langle ., . \rangle)$ und $(W, \langle\!\langle ., . \rangle\!\rangle)$ unitär (oder Euklidisch) und B, C ONB, so gilt

$$\xi_B^C(f^*) = \left(\xi_C^B(f)\right)^*,$$

wobei

$$X^* := \overline{X}^t \text{ für } X \in K^{n \times m}.$$

In diesem Falle gilt nämlich $b_j^* = \phi(b_j)$ und $c_i^* = \psi(c_i)$ und damit

$$x_{ij}^* = c_i^*(f^*(b_j)) = \langle\!\langle c_i, f^*(b_j) \rangle\!\rangle = \overline{\langle b_j, f(c_i) \rangle} = \overline{x_{ji}}.$$

Bemerkung Sind $f, g \in \mathrm{Hom}(W, V)$ und $x \in K$, so gilt

$$(f + gx)^t = f^t + g^t x \text{ und } (f + gx)^* = f^* + g^* \overline{x}.$$

6.1.4 Lemma

Sind $f \in \mathrm{Hom}(W, V)$ und $g \in \mathrm{Hom}(V, U)$, so gilt

$$(g \circ f)^t = f^t \circ g^t \text{ und } (g \circ f)^* = f^* \circ g^*.$$

Beweis Nachrechnen/-lesen oder über ein kommutatives Diagramm.

$$\begin{tikzcd}
W \arrow[r, bend left, "f"] \arrow[rr, bend left=40, "g \circ f"] \arrow[d, "\psi"] & V \arrow[l, bend left, "f^*"] \arrow[r, bend left, "g"] \arrow[d, "\phi"] & U \arrow[l, bend left, "g^*"] \arrow[d, "\mu"] \arrow[ll, bend left=90, "(g \circ f)^* = f^* \circ g^*"] \\
W^* & V^* \arrow[l, "f^t"'] & U^* \arrow[l, "g^t"'] \arrow[ll, bend left, "(g \circ f)^t = f^t \circ g^t"]
\end{tikzcd}$$

6.1.5 Lemma

Sind $f \in \mathrm{Hom}(W, V)$ und $f^* \in \mathrm{Hom}(V, W)$ adjungiert, so gilt

$$\ker f^* = f(W)^\perp$$

und f^* ist injektiv, wenn f surjektiv ist.

Beweis Da die Skalarprodukte $\langle\!\langle ., . \rangle\!\rangle$ und $\langle ., . \rangle$ auf W bzw. V nicht-degeneriert sind, gilt

$$v \in \ker f^* \Leftrightarrow \forall w \in W : \langle\!\langle f^*(v), w \rangle\!\rangle = 0 \Leftrightarrow \forall w \in W : \langle v, f(w) \rangle = 0 \Leftrightarrow f(W)^\perp,$$

also die erste Behauptung; ist $f(W) = V$, so folgt damit $\ker f^* = f(W)^\perp = V^\perp = \{0\}$.

6.1.6 Korollar

Sind $(W, \langle\!\langle ., . \rangle\!\rangle)$ und $(V, \langle ., . \rangle)$ Euklidisch oder unitär mit $\dim V, \dim W < \infty$ und $f \in \mathrm{Hom}(W, V)$ und $f^* \in \mathrm{Hom}(V, W)$ adjungiert, so gilt

$$\mathrm{rg}\, f^* = \mathrm{rg}\, f.$$

Beweis In der Situation hier (endlich-dimensional, positiv definite Skalarprodukte) sind $f(W)$ und $f(W)^\perp$ komplementäre UR und damit

$$V = f(W) \oplus f(W)^\perp = f(W) \oplus \ker f^*,$$

also

$$\mathrm{rg}\, f = \dim V - \mathrm{def}\, f^* = \mathrm{rg}\, f^*$$

nach Rangsatz für f^*.

Bemerkung Ist $\dim V, \dim W < \infty$, so folgt dann mit der Gleichheit der Abbildungen über die kanonischen Injektionen auch

$$\mathrm{rg}\, f^t = \mathrm{rg}\, f \text{ für } f \in \mathrm{Hom}(V, W).$$

Daher gilt $\mathrm{rg}\, f^* = \mathrm{rg}\, f$ dann auch für allgemeine Skalarprodukte auf $\mathbb{R}$- oder $\mathbb{C}$-VR V und W.

6.1.7 Buchhaltung

Mit den zu $X \in K^{n\times m}$ und $Y \in K^{k\times n}$ assoziierten Homomorphismen $f_X \in \mathrm{Hom}(K^m, K^n)$ und $f_Y \in \mathrm{Hom}(K^n, K^k)$ zeigt man also

$$(YX)^t = X^t Y^t \text{ und } (YX)^* = X^* Y^*.$$

Weiters folgt wegen $\mathrm{rg}\, f_X^* = \mathrm{rg}\, f_X^t = \mathrm{rg}\, f_X$

$$\mathrm{rg}\, X^t = \mathrm{rg}\, X^* = \mathrm{rg}\, X;$$

anders ausgedrückt: der Zeilenrang einer Matrix $X \in K^{n\times m}$ stimmt mit ihrem (Spalten-)Rang überein.

Insbesondere gilt:

$$X \in Gl(n) \Rightarrow X^t \in Gl(n).$$

6.1.8 Lemma

Sei $(V, \langle .,. \rangle)$ Euklidisch (unitär); $f \in Gl(V)$ ist genau dann orthogonal (unitär), wenn f und f^{-1} adjungiert sind, $f^{-1} = f^*$.

Beweis Ist $f^{-1} = f^*$ zu f adjungiert, so ist $f \in O(V)$ (bzw. $f \in U(V)$), da

$$\forall v, w \in V : \langle f(v), f(w) \rangle = \langle (f^* \circ f)(v), w \rangle = \langle v, w \rangle;$$

ist umgekehrt $f \in O(V)$ (bzw. $U(V)$), so ist

$$\forall v, w \in V : \langle v, f(w) \rangle = \langle f(f^{-1}(v)), f(w) \rangle = \langle f^{-1}(v), w \rangle,$$

d.h. f^{-1} ist zu f adjungiert.

Bemerkung $f \in Gl(V)$ ist also genau dann orthogonal/unitär, wenn

$$f^* \circ f = f \circ f^* = \mathrm{id}_V .$$

Bemerkung Das Lemma lässt sich auf Isomorphismen $f \in \mathrm{Iso}(W, V)$ verallgemeinern: f ist genau dann *isometrisch* (längentreu), wenn $f^{-1} = f^*$.

6.1.9 Buchhaltung

Ist $(V, \langle .,. \rangle)$ Euklidischer VR, $\dim V < \infty$, und $X = \xi_B^B(f)$ Darstellungsmatrix von $f \in \mathrm{End}(V)$ bzgl. einer ONB B, so ist

$$f \in O(V) \Leftrightarrow X^* X = E_n,$$

und analog für einen unitären VR V. Daher definiert man die *orthogonale (unitäre) Gruppe in n Variablen*:

$$O(n) := \left\{ X \in \mathbb{R}^{n \times n} \mid X^* X = X^t X = E_n \right\},$$

$$U(n) := \left\{ X \in \mathbb{C}^{n \times n} \mid X^* X = \overline{X}^t X = E_n \right\}.$$

X ist also orthogonal/unitär, wenn die Spalten von X eine ONB von $\mathbb{K}^{n \times 1}$ mit dem kanonischen Skalarprodukt bilden.

6.1.10 Definition

Sei $(V, \langle .,. \rangle)$ Euklidisch oder unitär; $f \in \mathrm{End}(V)$ heißt dann

- *selbstadjungiert* oder *symmetrisch*, falls

$$\forall v, w \in V : \langle f(v), w \rangle = \langle v, f(w) \rangle;$$

- *schiefadjungiert* oder *schiefsymmetrisch*, falls

$$\forall v, w \in V : \langle f(v), w \rangle + \langle v, f(w) \rangle = 0.$$

Bemerkung $f \in \mathrm{End}(V)$ ist also genau dann (schief-)symmetrisch, wenn f eine Adjungierte $f^* \in \mathrm{End}(V)$ besitzt und $f^* = \pm f$. Dies ist genau dann der Fall, wenn

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

eine (schief-)symmetrische Sesquilinearform definiert.

6.1.11 Korollar

Sei $(V, \langle .,. \rangle)$ Euklidischer VR; eine Projektion $p \in \mathrm{End}(V)$ ist genau dann Orthogonalprojektion, wenn sie selbstadjungiert ist, $p^* = p$.

Beweis Lemma Abschnitt 5.5.

6.2 Normale Endomorphismen

Motivation Für orthogonale/unitäre selbst- und schiefadjungierte Endomorphismen $f \in \mathrm{End}(V)$ gilt stets

$$f^* \circ f = f \circ f^*.$$

Dies ist eine „gute“ Eigenschaft: sie liefert viele/wichtige strukturelle Aussagen über Endomorphismen.

Generalvoraussetzung In diesem Abschnitt ist $(V, \langle ., . \rangle)$ Euklidisch oder unitär.

6.2.1 Definition

$f \in \mathrm{End}(V)$ heißt *normal*, wenn f eine Adjungierte $f^* \in \mathrm{End}(V)$ besitzt und

$$f^* \circ f = f \circ f^*.$$

6.2.2 Buchhaltung

Ist $\dim V < \infty$ und $X = \xi_B^B(f)$ Darstellungsmatrix von $f \in \mathrm{End}(V)$ bzgl. einer ONB B von $(V, \langle ., . \rangle)$, so gilt

$$f \text{ normal} \Leftrightarrow X^*X = XX^*,$$

d.h. wenn $X \in \mathbb{K}^{n \times n}$ *normal* ist.

6.2.3 Lemma

Ist $f \in \mathrm{End}(V)$ normal, so gilt:

(i) $\ker f = \ker f^* = f(V)^\perp$;

(ii) $\forall v, w \in V : \langle f^*(v), f^*(w) \rangle = \langle f(v), f(w) \rangle$;

(iii) $\forall x \in \mathbb{K} \forall v \in V : f(v) = vx \Rightarrow f^*(v) = v\overline{x}$.

(iv) Sind $v, w \in V$ Eigenvektoren zu EW $x, y \in \mathbb{K}$ von f, so gilt

$$x = y \text{ oder } v \perp w.$$

Beweis Sei $f \in \mathrm{End}(V)$ normal.

(ii) Wegen $f^{**} = f$ gilt für $v, w \in V$:

$$\langle f^*(v), f^*(w) - \langle f(v), f(w) \rangle \rangle = \langle (f^{**} \circ f^*)(v), w \rangle - \langle (f^* \circ f)(v), w \rangle = \langle (f \circ f^* - f^* \circ f)(v), w \rangle = 0$$

(i) Wegen (ii) gilt für $v \in V$

$$f^*(v) = 0 \Rightarrow 0 = \|f^*(v)\|^2 = \|f(v)\|^2 \Rightarrow f(v) = 0$$

und umgekehrt, und damit

$$\ker f^* = \ker f.$$

Nach früherem Lemma ist

$$\ker f^* = f(V)^\perp$$

(iii) Nach (i) ist für $x \in \mathbb{K}$

$$\ker(f - \mathrm{id}_V\, x) = \ker(f - \mathrm{id}_V\, x)^* = \ker(f^* - \mathrm{id}_v\, \overline{x}),$$

da $f - \mathrm{id}_V\, x$ mit f normal ist.

(iv) Mit (iii) folgt für $v, w \in V$ mit $f(v) = vx$ und $f(w) = wy$

$$(x - y)\langle v, w\rangle = \langle v\overline{x}, w\rangle - \langle v, wy\rangle = \langle f^*(v), w\rangle - \langle v, f(w)\rangle = 0.$$

6.2.4 Lemma

Ist $f \in \mathrm{End}(V)$ normal und $U \subset V$ UVR, so gilt

(i) Ist U f-invariant, so ist $U^\perp$ f^*-invariant;

(ii) Ist U f- und f^*-invariant, so liefert Einschränkng normale Endomorphismen

$$f\big|_U \in \mathrm{End}(U) \text{ und } f\big|_{U^\perp} \in \mathrm{End}(U^\perp).$$

Beweis Für (i) wird nur die Existenz der Adjungierten benutzt.

(i) Sei $v \in U^\perp$, dann gilt:

$$\forall u \in U : \langle f^*(v), u\rangle = \langle v, f(u)\rangle = 0$$

(ii) Da U f- und f^*-invariant ist, ist (nach(i)) $U^\perp$ f^*- und $f^{**} = f$-invariant. Damit ist es sinnvoll

$$f\big|_U \in \mathrm{End}(U), f\big|_{U^\perp} \in \mathrm{End}(U^\perp)$$

$$f^*\big|_U \in \mathrm{End}(U), f^*\big|_{U^\perp} \in \mathrm{End}(U^\perp)$$

zu betrachten. Nun gilt:

$$\forall u, v \in U : \langle f\big|_U^*(u), v\rangle = \langle u, f\big|_U(v)\rangle = \langle u, f(v)\rangle = \langle f^*(u), v\rangle = \langle f^*\big|_U(u), v\rangle$$

und analog für $v, w \in U^\perp$. Damit folgt: $f|_U^* = f^*|_U$ und $f|_{U^\perp}^* = f^*|_{U^\perp}$ und also

$$f|_U^* \circ f|_U = f^* \circ f|_U = f \circ f^*|_U = f|_U \circ f|_U^*.$$

Bemerkung & Beispiel Eine Orthogonalprojektion ist selbstadjungiert, $p \in \mathrm{End}(V)$ mit $p^2 = p, p^* = p$ und damit normal. Ist $p \neq \mathrm{id}_V, 0$, so ist

$$V = U \oplus_\perp U^\perp \text{ mit } \begin{cases} U := p(V), \\ U^\perp = \ker p. \end{cases}$$

Wir definieren:

$$\pi : V \to U, v \mapsto \pi(v) := p(v) \text{ und } \iota : U \to V, u \mapsto \iota(u) := u;$$

man nennt die isometrische Abbildung ι auch die *Inklusion* von U in V. Dann sind π und ι adjungiert:

$$\forall u \in U \forall v \in V : \langle \iota(u), v \rangle = \langle u, \pi(v) \rangle|_U$$

Insbesondere ist die „Projektionsabbildung“ π *nicht* selbstadjungiert.

Achtung: Die Adjungierte hängt von Definitions- und Wertebereich ab!

6.2.5 Spektralsatz (unitärer Fall)

Sei $(V, \langle .,. \rangle)$ unitär, $\dim V < \infty$, und sei $f \in \mathrm{End}(V)$ normal. Dann besitzt V eine ONB aus Eigenvektoren von f.

Bemerkung Es gilt auch die Umkehrung: Ist $(e_1, \ldots, e_n)$ ONB mit

$$f(e_i) = e_i x_i, \text{ also } f^*(e_i) = e_i \overline{x_i},\ i = 1, \ldots, n,$$

so gilt

$$(f^* \circ f)(e_i) = e_i \overline{x_i} x_i = e_i x_i \overline{x_i} = (f \circ f^*)(e_i),\ i = 1, \ldots, n,$$

d.h. f ist normal.

Beweis Induktion über $n = \dim V$.

Für $n = 1$ ist die Aussage trivial. Sei die Aussage für $n \in \mathbb{N}$ wahr. Für $n + 1$ gilt dann:

f hat einen Eigenwert $x \in \mathbb{C}$, da das charakteristische Polynom $\chi_f(t) \in \mathbb{C}[t]$ nach Fundamentalsatz der Algebra in Linearfaktoren zerfällt.

Sei $e \in V^\times$ ein zugehöriger Eigenvektor,

$$f(e) = ex, \text{ o.B.d.A. } \|e\| = 1.$$

Wegen $f^*(e) = e\overline{x}$ ist $[e]$ dann f- und f^*- invariant und damit

$$V = [e] \oplus_\perp [e]^\perp,$$

wobei $f\big|_{[e]} \in \operatorname{End}([e])$ und $f\big|_{[e]^\perp} \in \operatorname{End}([e]^\perp)$ normal sind (Lemma).

Da $\dim[e]^\perp = n$ liefert die Induktions-Annahme eine ONB $(e_1, \ldots, e_n)$ von $[e]^\perp$ aus Eigenvektoren von $f\big|_{[e]^\perp}$. Damit ist $(e, e_1, \ldots, e_n)$ eine ONB aus Eigenvektoren von f.

6.2.6 Buchhaltung

Ein normaler Endomorphismus $f \in \operatorname{End}(V)$ eines unitären VR $(V, \langle .,. \rangle)$ mit $\dim V < \infty$ ist also *orthogonal diagonalisierbar*, d.h. es existiert eine ONB aus Eigenvektoren von f.

Also, bezüglich einer solchen ONB B ist

$$\xi_B^B(f) = \operatorname{diag}(x_1, \ldots, x_n).$$

Da $\xi_B^B(f^*) = (\xi_B^B(f))^*$ gilt:

- ist f selbstadjungiert, so sind alle Eigenwerte reell, $x_i = \overline{x_i}$;
- ist f schiefadjungiert, so sind alle Eigenwerte imaginär, $x_i = -\overline{x_i}$
- ist f unitär, so sind alle Eigenwerte *unitär*, d.h. für $j = 1, \ldots, n$ ist

$$x_j \in S^1 := \{x \in \mathbb{C} : \overline{x}x = 1\} = \{e^{iy} \mid y \in \mathbb{R}\}.$$

6.2.7 Korollar & Definition

Ist $X \in \mathbb{C}^{n\times n}$ normal, $X^*X = XX^*$, so gilt

$$\exists P \in U(n) : P^{-1}XP = \operatorname{diag}(x_1, \ldots, x_n).$$

Dabei gilt:

- ist X *selbstadjungiert*, $X^* = X$, so sind $x_1, \ldots, x_n \in \mathbb{R}$;
- ist X *schiefadjungiert*, $X^* = -X$, so sind $x_1, \ldots, x_n \in i\mathbb{R}$;
- ist X *unitär*, $X \in U(n) = \{Y \in \mathbb{C}^{n\times n} \mid Y^*Y = E_n\}$, so gilt $|x_1|, \ldots, |x_n| = 1$.

Beweis Sei $X \in \mathbb{C}^{n\times n}$, betrachte $\mathbb{C}^n$ als unitären VR mit Standardbasis E als ONB.

$$\Gamma_E(\langle .,. \rangle) = E_n,$$

und den assoziierten Endomorphismus $f_X \in \mathrm{End}(\mathbb{C}^n)$. Orthonormale Basiswechsel $B = EP$ sind dann durch unitäre Matrizen $P \in U(n)$ gegeben:

$$E_n = \Gamma_B(\langle .,. \rangle) = P^* \Gamma_E(\langle .,. \rangle) P = P^* P \Leftrightarrow P \in U(n).$$

Anwendung des Spektralsatzes liefert also die Behauptung.

Beispiel Für $X = \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} \in \mathbb{C}^{2\times 2}$ mit $s \in \mathbb{R}$ ist $X^* = X^t = X^{-1}$, d.h. X ist unitär, also normal, und damit

$$\exists P \in (2) : P^{-1} \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} P = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}$$

wobei x_i die Eigenwerte von f_X sind, d.h. Nullstellen des charakteristischen Polynoms

$$\chi_X(t) = (t - \cos s)^2 + \sin^2 s = t^2 - 2t\cos s + 1 = (t - e^{is})(t - e^{-is}).$$

Bemerke: $|e^{\pm is}| = 1$, d.h. $x_{1,2}$ sind unitär. Eigenvektoren bzw. P:

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$$

6.2.8 Spektralzerlegung (unitärer Fall)

Sei $(V, \langle .,. \rangle)$ unitär, $\dim V < \infty$, und sei $f \in \mathrm{End}(V)$ normal; dann zerfällt V als orthogonale direkte Summe der Eigenräume von f,

$$V = \bigoplus_{x \in \chi_f(\{0\})} \ker(\mathrm{id}_V x - f).$$

Beweis Folgt direkt aus dem Spektralsatz

Bemerkung Mit gewissen Voraussetzungen gilt der Satz auch für $\dim V = \infty$.

Index