Skript Lineare Algebra & Geometrie 2, Hertrich-Jeromin

Studierendenmitschrift

7. April 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

$$\begin{aligned} p &: K \to K,\ x \mapsto p(x) := 1 + x + x^2 \\ q &: K \to K,\ x \mapsto q(x) := 1 \end{aligned}$$

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

$$\begin{aligned} &\forall x \in K : x(x+1) = 0 \\ \Rightarrow\ &\forall x \in K : p(x) = q(x) \end{aligned}$$

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 $\mathrm{End}(V)$ 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: $v^k := \underbrace{v \odot \cdots \odot v}_{k-\text{mal}}$ 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

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

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

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

vermittels des Isomorphismus

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

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

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:

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

[1] sonst wäre K^K unendlich dimensional.

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

$$p(t) = \begin{cases} d(t)q(t) + r(t) \\ d(t)\tilde{q}(t) + \tilde{r}(t) \end{cases}$$

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

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

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

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* $d(t) = \operatorname{ggT}\big(p(t), q(t)\big)$ 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*
>
> $$\begin{gathered} d(t) \mid p(t) \wedge d(t) \mid q(t) \text{ und} \\ d'(t) \mid p(t) \wedge d'(t) \mid q(t) \Rightarrow d'(t) \mid d(t). \end{gathered}$$

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 $x_1, \dots, x_n \in \mathbb{C}$ 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,

$$\ker(f - \mathrm{id}_V) = [\{b_1 + b_2\}]$$

hat

$$\dim \ker(f - \mathrm{id}_V) < \dim V.$$

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

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

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 $\xi_B^B(f) = X = (x_{ij})_{i,j\in\{1,\dots,n\}}$, 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 $\deg(\chi_f(t)) = n$,

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

wobei die *Spur* $\mathrm{tr}\, f$ („tr “$\widehat{=}$ trace) von f durch diese Gleichung (wohl-)defininiert ist.

Ist $(x_{ij})_{i,j\in\{1,\dots,n\}} = X = \xi_B^B(f)$ Darstellungsmatrix von f, so gilt

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

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;

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

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 $B = (b_1, \dots, b_n)$ von V mit $\ker(\mathrm{id}_v\, x - f) = [(b_1, \dots, b_g)]$ 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

$$k = 2 \text{ und } g = 1$$

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 $x_1, \dots, x_n \in K$ 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 $v_1, \dots, v_m \in V$ zu paarweise verschiedenen Eigenwerten $x_1, \dots, x_m$ 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 $y_1, \dots, y_m \in K$, so folgt $y_1 = \dots = y_m = 0$.

Seien $y_1, \dots, y_m \in K$ 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 $f(w_i) = w_i x_i$

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

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

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

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

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

also

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

4.4.10 Satz

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

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

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 $C_i, i = 1, \dots, m$. 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 $f \in \operatorname{End}(V)$ 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 $x_1, \ldots, x_n$, 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

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

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

$$\chi_f(t) = t^n + t^{n-1} a_{n-1} + \dots, +t a_1 + a_0.$$

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

Beweis Da $B = \big(b, f(b), \dots, f^{n-1}(b)\big)$ 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 $Y = (y_1, y_2)$

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

mit $X = \begin{pmatrix} 0 & 3 \\ 1 & -2 \end{pmatrix}$, 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 $\frac{d}{ds} \in \operatorname{End}(C^\infty(\mathbb{R}))$ Darstellungsmatrix X hat.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

so folgt

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

und damit

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

4.5.5 Beweis vom Satz von Cayley-Hamilton

Zu zeigen: für $f \in \mathrm{End}(V)$ gilt $\chi_f(f) = 0$:

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

Sei also $v \in V^\times$ und

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

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

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

Damit folgt

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

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

Index