Kurze Wiederholung von den Themen letzter Woche, die in dieser Woche verallgemeinert werden

- Diagonalisierung von (Matrizen von) Endomorphismen
- Polynome von Endomorphismen

Diagonalisierung von (Matrizen von) Endomorphismen

- Den Satz 32 LA I, welchen ich in der Woche 1 wieder bewiesen habe, werde ich nicht noch einmal beweisen oder wiederholen.
- Aber die Philosophie schon: wir haben die Basis gesucht, sodass in der Basis die Matrix möglich einfach ist. Wir haben gesehen, dass in der Basis aus Eigenwerten die Matrix diagonal ist.
- "Diagonalform" ist in diesem Fall nicht wesentlich. Auch wenn wir die Matrix zu einer anderen Form bringen wollen, müssen wir eine Basis mit bestimmten Eigenschaften suchen. Ich mache ein Bsp.

Einfaches Bsp.

- Was machen wir, wenn wir (in Dimension 2) eine Basis suchen, sodass in dieser Basis ein gegebener Endomorphismus f durch die Matrix $A' := \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ gegeben ist?
- Theoretische Antwort, welche aus Definitionen folgt, ist wie folgt: wir sollen eine Basis $(b_1 = \begin{pmatrix} b_1^1 \\ b_1^2 \end{pmatrix}, b_2 = \begin{pmatrix} b_2^1 \\ b_2^1 \end{pmatrix})$ suchen sodass $f(b_1) = 2b_1$ und $f(b_2) = b_1 + 2b_2$ sein. Ich gebe eine rechnerische Erklärung auf der nächsten Seite.

Suche $\left(b_1 = \begin{pmatrix} b_1^1 \\ b_1^2 \end{pmatrix}, b_2 = \begin{pmatrix} b_2^1 \\ b_2^2 \end{pmatrix}\right)$ sodass $f(b_1) = 2b_1$ und $f(b_2) = b_1 + 2b_2$

- Wenn z.B. der Endomorphismus f in der Standardbasis (e_1, e_2) durch der Matrix A gegeben ist, dann gilt für die Matrix $B = \begin{pmatrix} b_1^1 & b_2^1 \\ b_1^2 & b_2^2 \end{pmatrix}$:

$$B^{-1}AB = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},$$

denn:

$$B^{-1}ABe_1 = B^{-1}Ab_1 = 2B^{-1}b_1 = 2e_1,$$

also die erste Spalte von $B^{-1}AB$ ist $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$. Analog gilt:

$$B^{-1}ABe_2 = B^{-1}Ab_2 = B^{-1}(2b_2 + b_1) = e_1 + 2e_2,$$

also die zweite Spalte von $B^{-1}AB$ ist $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ wie wir es wollen.

- Selbstverständlich ist die Existenz einer solchen Basis nicht immer gegeben (unsere Matrix A' hat Eingenwert 2 und Spur 4, deswegen können nur die Matrizen mit Eigenwert 2 und Spur 4 auf A' mit einer Ähnlichkeittransformation $A \mapsto B^{-1}AB$ überfuhrt werden). Auch die Methode, wie man die Basis findet, ist jetzt für Sie wahrscheinlich nicht klar, sie kommt noch.

Polynome von Matrizen

- Ich habe ohne Beweis benutzt, dass für die Polynome $P(A) = \sum_{k=0}^{n} a_k A^k$ und $Q(A) = \sum_{k=0}^{m} b_k A^k$ die folgende Regel gilt: $P(A)Q(A) = PQ(A)$ (wobei PQ das Produkt von Polynomen ist).
  Daraus folgte, dass $P(A)Q(A) = Q(A)P(A)$.
- Heute werde ich Produkt von Polynomen definieren.
  Allgemein werden wir heute viel mit Polynomen und Polynomen von Matrizen arbeiten.

Hauptziel dieser Woche ist Jordansche Normalformsatz

- Das ist die wichtigste Aussage der Matrizenrechnung
- Sie werden mehrere Anwendungen davon sehen: zuerst bei mir, später in Analysis, Differentialgleichungen, Numerik, Algebra 1,2, Lie-Theorie, ...

Exkurs: "Elementare Zahlentheorie" für Polynome

Def. Ein Polynom f über einem Körper $\mathbb{K}$ mit der Unbestimmten x ist eine formale Summe

$$f(x) = \sum_{i \geq 0} a_i x^i,$$

wobei nur endlich viele der Koeffizienten $a_i \in \mathbb{K}$ von Null verschieden sind.
(Koeffizienten, die Null sind, werden normalerweise beim Schreiben weggelassen; $f(x) = \sum_{i=0}^{k} a_i x^i$.)

Bsp. $2x^2 + 1$ ist ein Polynom über $\mathbb{R}$.
In diesem Bsp. habe ich $0 \cdot x^1$ und x^0 "vergessen" (sollte eigentlich $2x^2 + 0 \cdot x^1 + 1 \cdot x^0$ schreiben). Dass ist eine standarte kosmetische Regel; sie wird auch über anderen Körper eingesetz und ist mit allen späteren Regeln kompatibel.

Zwei Polynome heißen gleich, wenn alle ihre Koeffizienten übereinstimmen.
Wir werden später sehen, dass (über einigen Körper) Gleicheit von zwei Polynomen f_1 und f_2 ist nicht dasselbe wie die Gleichung von entsprechenden Funktionen von $\mathbb{K}$ nach $\mathbb{K}$ (sie werden später bespochen und $\tilde{f_1}$, $\tilde{f_2}$ bezeichnet).

Der Grad eines Polynoms ist der größte Index i, für den $a_i \neq 0$ ist. Dieser Koeffizient a_i wird auch Anfangskoeffizient, Leitkoeffizient bzw. höchster Koeffizient genannt. Für den Grad des Nullpolynoms setzt man (symbolisch) $-\infty$.

Auf der Menge $\mathbb{K}[x]$ definiert man die Summe und das Produkt von Polynomen: Die Summe $\sum_i c_i x^i$ der Polynome $\sum_i a_i x^i$ und $\sum_i b_i x^i$ ist definiert durch $c_i = a_i + b_i$;
das Produkt $\sum d_i x^i$ durch $d_i = \sum_{k+\ell=i} a_k \cdot b_\ell$.

Die Summe $\sum_i c_i x^i$ der Polynome $\sum_i a_i x^i$ und $\sum_i b_i x^i$ ist definiert durch $c_i = a_i + b_i$;
das Produkt $\sum d_i x^i$ durch $d_i = \sum_{k+\ell=i} a_k \cdot b_\ell$.

Bemerkung. Die Addition und Multiplikation von Polynomen sind assoziative und kommutative Operationen und erfüllen Distributivgesetz.

Beweis. Distributivität und Kommutativität von "+" und "·" folgen sofort aus Definition und aus Körpereigenschaften: z.B., ist i-ter Koeffizient des Polynoms $(A+B)C$ gleich

$$\sum_{k+\ell=i} (a_k + b_k) c_\ell \qquad (*),$$

i-ter Koeffizient des Polynoms $AC + BC$ ist

$$\sum_{k+\ell=i} a_k c_\ell + \sum_{k+\ell=i} b_k c_\ell \qquad (**).$$

Wir vergleichen $(*)$ und $(**)$ und sehen, dass $(A+B)C = AC + BC$.

Die Summe $\sum_i c_i x^i$ der Polynome $\sum_i a_i x^i$ und $\sum_i b_i x^i$ ist definiert durch $c_i = a_i + b_i$;
das Produkt $\sum d_i x^i$ durch $d_i = \sum_{k+\ell=i} a_k \cdot b_\ell$.

Das Assoziativgesetz der Multiplikation folgt aus

$$\sum_{k+\ell=i} a_k \left(\sum_{m+n=\ell} b_m c_n \right) = \sum_{k+m+n=i} a_k b_m c_n \qquad (*)$$

$$\sum_{\ell+n=i} \left(\sum_{k+m=\ell} a_k b_m \right) c_n = \sum_{k+m+n=i} a_k b_m c_n \qquad (**).$$

Hier habe ich in $(*)$ den $i-$ten Koeffizient von $A(BC)$, und in $(**)$ den $i-$ten Koeffizient von $(AB)C$ nach Definition ausgerechnet. Wir sehen, dass $(*) = (**)$ ist, also ist Polynommultiplikation assoziativ. □

Bemerkung. Da die Operationen "+" und "·" assoziativ, kommutativ und distributiv sind, können wir die Polynome multiplizieren indem wir die Klammer mit Hilfe von Assoziativität, Distributivität und Kommutativität ausklammern.

Übliche Rechenregeln für Grad des Polynoms

Die Summe $\sum_i c_i x^i$ der Polynome $\sum_i a_i x^i$ und $\sum_i b_i x^i$ ist definiert durch $c_i = a_i + b_i$;
das Produkt $\sum d_i x^i$ durch $d_i = \sum_{k+\ell=i} a_k \cdot b_\ell$.

Es gilt. $Grad(f+g) \leq max(Grad(f), Grad(g))$, falls $Grad(f) \neq Grad(g)$ ist, dann $Grad(f+g) = max(Grad(f), Grad(g))$. Ferner gilt: $Grad(f \cdot g) = Grad(f) + Grad(g)$.

Beweis. Sei $f = \sum_i a_i x^i$ und $g = \sum_i b_i x^i$. Wenn $i > max(Grad(f), Grad(g))$ ist, gilt $a_i = 0$ und $b_i = 0$ und deswegen $a_i + b_i = 0$, also ist der i-te Koeffizient von $f+g$, gleich $a_i + b_i = 0 + 0 = 0$, wie behauptet. Wenn außerdem $Grad(f) \neq Grad(g)$ ist, oBdA $Grad(f) > Grad(g)$ ist, dann $a_i \neq 0$ und $b_i = 0$ für $i = Grad(f)$ und deswegen ist der i-te Koeffizient von $f+g$ ungleich 0. Sei jetzt $i > Grad(f) + Grad(g)$. Dann ist der i-te Koeffizient von $f \cdot g$ gleich $\sum_{k+\ell=i} a_k \cdot b_\ell = 0$, weil in jedem Summand $a_k \cdot b_\ell$ wegen $k + \ell > Grad(f) + Grad(g)$ mind. einer der Faktoren, a_k oder b_ℓ gleich 0 ist, wie behauptet. □

Übliche Rechenregeln für Polynomdivision

Es gilt. Seien f und g Polynome mit $Grad(f) \geq Grad(g) \geq 0$. Dann gibt es Polynome q und r mit $Grad(r) < Grad(g)$, so dass $f = q \cdot g + r$.

Beweis wiederholt den Beweis für Polynome über $\mathbb{R}$, siehe Satz 30 LA I und wird hier nicht gegeben.

Für ein Polynom $f = \sum_i a_i x^i$ kann man die Abbildung $\tilde{f} : \mathbb{K} \to \mathbb{K}$, $\tilde{f}(z) = \sum_i a_i z^i$ definieren (wir setzen also das Element $z \in \mathbb{K}$ statt Variabel x ein).

Def. Ein $\lambda \in K$ ist eine Nullstelle von f, wenn $\tilde{f}(\lambda) = 0$.

Für ein Polynom $f = \sum_i a_i x^i$ definieren wir die Abbildung $\tilde{f} : \mathbb{K} \to \mathbb{K}$ durch $\tilde{f}(z) = \sum_i a_i z^i$

Wenn $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ ist, ist die Abbildung $\tilde{f}$ im Wesentlichen dasselbe wie das Polynom f, da die Abbildung $\tilde{f}$ das Polynom f, das wir benutzt haben um das Polynom $\tilde{f}$ zu konstruieren, eindeutig bestimmt. Das haben wir in LA I gesehen; ich wiederhole kurz den Beweis:

Beweis dass über $\mathbb{R}$ (oder $\mathbb{C}$ zwei Polynomen $f_a := \sum_{k=0}^{n} a_k x^k$ und $f_b := \sum_{k=0}^{m} b_k x^k$ genau dann gleich sind, wenn sie gleich in Sinne von Polynomengleicheit sind (also, wenn $a_i \equiv b_i$ für alle $i \in \mathbb{N} \cup \{0\}$).

O.B.d.A. ist $n = m$ (wir werden im Beweis nicht benuztz, dass Leitkoeffizient nicht 0 ist und können die fehlende Koeffizienten b_i, falls etwa $n > m$ ist, mit 0 ergänzen). Wir nehmen $n+1$ verschiedene Zahlen $z_0, z_1, ..., z_n$ (z.B. $z_0 = 0$,..., $z_n = n$) und betrachten das lineare Gleichungssystem auf $x_0, ..., x_n$.

$$\begin{pmatrix} 1 \cdot x_0 + z_0 \cdot x_1 + z_0^2 \cdot x_2 + ... + z_0^n \cdot x_n &= \tilde{f}_a(z_0) \\ 1 \cdot x_0 + z_1 \cdot x_1 + z_1^2 \cdot x_2 + ... + z_1^n \cdot x_n &= \tilde{f}_a(z_1) \\ &\vdots \\ 1 \cdot x_0 + z_n \cdot x_1 + z_n^2 \cdot x_2 + ... + z_n^n \cdot x_n &= \tilde{f}_a(z_n) \end{pmatrix}$$

$$\begin{pmatrix} 1 \cdot x_0 + z_0 \cdot x_1 + z_0^2 \cdot x_2 + \ldots + z_0^n \cdot x_n & = & \tilde{f}_a(z_0) \\ 1 \cdot x_0 + z_1 \cdot x_1 + z_1^2 \cdot x_2 + \ldots + z_1^n \cdot x_n & = & \tilde{f}_a(z_1) \\ & & \vdots \\ 1 \cdot x_0 + z_n \cdot x_1 + z_n^2 \cdot x_2 + \ldots + z_n^n \cdot x_n & = & \tilde{f}_a(z_n) \end{pmatrix}$$

Das ist ein LGS aus $(n+1)$ Gleichungen auf $(n+1)$ Unbekannten $x_0, \ldots, x_n$. Dir Koeffizieten "kommen" von $z_0, \ldots, z_n$. Die rechte Seite ist ein Vektor, die Komponenten dafon sind die Werte von Funktion $\tilde{f}_a$ auf z_i. Man bemerke, dass $\tilde{f}_a(z_i) = \tilde{f}_b(z_i)$ und sind die Werten von $\tilde{f}_b$; deswegen ist das System kostruiert vom f_a und vom f_b gleich.

Die Koeffizientenmatrix davon ist eine $(n+1) \times (n+1)$-Matrix

$$A = \begin{pmatrix} 1 & z_0 & \cdots & z_0^n \\ 1 & z_1 & \cdots & z_1^n \\ \vdots & & \vdots & \\ 1 & z_n & \cdots & z_n^n \end{pmatrix}$$

Das ist eine Vandermonde-Matrix, wir haben sie in LA I in Hausaufgaben zwei mal behandelt und gezeigt, dass sie nichtausgeartet ist. Die Lösung der Gleichungssystem oben ist deswegen eindeutig. Man bemerke aber, dass nach Konstruktion von LGS die Vektoren $(a_0, \ldots, a_n)^T$ und $(b_0, \ldots, b_n)^T$ Lösungen sind, weil wenn wir etwa in der ersten Gleichung a_i statt x_i Einsetzen bekommen wir links
$a_0 + a_1 z_0 + a_2 z_0^2 + \ldots + a_n z_0^n = \tilde{f}_a(z_i)$. Also, Eindeuitgkeit der Lösung impliziert $(a_0, \ldots, a_n)^T = (b_0, \ldots, b_n)^T$, □

Es gibt aber Körper sodass verschiedene Polynome gleiche Funktionen $\tilde{f}$ haben:

z.B. nehmen wir $\mathbb{K} = \mathbb{Z}_2$ und die Polynome x (also $0 + 1 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 + ...$) und $x + x(x+1) = 0 + 0 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 + ... = x^2$. Die entsprechenden Funktionen sind gleich: Der Definitionsbereich ist zweielementig $\{0, 1\}$ und die beide Polynome sind 0, wenn wir 0 einsetzen, und 1, wenn wir 1 einsetzen. Die Polynome sind aber verschieden, da die a_1-Koeffizienten verschieden sind.

Selbstverständlich kann man den Beweis für alle endliche Körper verallgemeinern: wenn $\mathbb{K}$ endlich ist und aus N Elementen $k_0, ..., k_{N-1}$ bestehet (wir wissen aus LA I dass N eine Primzahlpotenz ist, aber für diese Stelle ist es nicht wesentlich), dann gilt für das Polynom $f(x) = (x - k_0)(x - k_1) \cdots (x - k_{N-1})$ (die Regeln für Muptiplizieren von Polynimen haben wir bereit eingefuhrt (oder wiederholt))
$\tilde{f}(x) \equiv 0$. Aber das Polynom ist kein Null-Polynom, weil Grad davon gleich N ist.

Übliche Rechenregeln für Polynomdivision

Def. Ein $\lambda \in K$ ist eine Nullstelle von f, wenn $\tilde{f}(\lambda) = 0$.

Wiederholung – Lemma 27 LA I. Sei P ein Polynom . Ist $\lambda \in \mathbb{K}$ eine Nullstelle von P, so existiert genau ein Polynom Q mit $P = (x - \lambda)Q$.

In LA I haben wir das Lemma für $\mathbb{K} = \mathbb{R}$ bewiesen. Das Lemma ist für alle Körper gültig, mit im wesentlichen gleichen Beweis.

GGT (größter gemeinsamer Teiler) für Polynome

Def. $g \in \mathbb{K}[x]$ ist ein Teiler von $f \in \mathbb{K}[x]$, falls es ein q mit $f = q \cdot g$ gibt. Falls $0 < \mathit{Grad}(g) < \mathit{Grad}(f)$, dann heißt g ein echter Teiler. Hat f keine echten Teiler, dann heißt f irreduzibel.

Bemerkung. Jedes Polynom hat mehrere nichtechte Teiler: jedes Polynom vom Grad 0, d.h., jedes $k \in \mathbb{K}, k \neq 0$ ist ein Teiler, da $f = \underbrace{k}_{q} \cdot \underbrace{\frac{1}{k} f}_{g}$. Außerdem sind alle Vielfachheiten von f auch Teiler von f; daher ist die Bedingung $0 < \mathit{Grad}(g) < \mathit{Grad}(f)$ sinnvoll.

Def. h heißt gemeinsamer Teiler von f und g, falls h ein Teiler von f und g ist. Ein gemeinsamer Teiler h ist ein größter gemeinsamer Teiler, wenn jeder gemeinsame Teiler ein Teiler von h ist.

Wicht. Beispiel (Analog von Primzahlzehrlegung in Zahlentheorie)

Wir betrachten zwei disjunkte Mengen $\{\lambda_1, ..., \lambda_k\}$ und $\{\mu_1, ..., \mu_m\} \subseteq \mathbb{K}$; $\{\lambda_1, ..., \lambda_k\} \bigcap \{\mu_1, ..., \mu_m\} = \varnothing$; und die Polynome $f = (x - \lambda_1)^{\alpha_1} ... (x - \lambda_k)^{\alpha_k}$ und $g = (x - \mu_1)^{\beta_1} ... (x - \mu_m)^{\beta_m}$, wobei $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_m \in \mathbb{N}$.
Dann gilt: $ggT(f, g) = 1$ (oder beliebiges $k \in \mathbb{K} \setminus \{0\}$) .

Def. h heißt gemeinsamer Teiler von f und g, falls h ein Teiler von f und g ist. Ein gemeinsamer Teiler h ist ein größter gemeinsamer Teiler, wenn jeder gemeinsame Teiler ein Teiler von h ist.

Lemma 27 LA I (Vorl. 12). Sei P ein Polynom. Ist $\lambda \in \mathbb{K}$ eine Nullstelle von P, so existiert genau ein Polynom Q mit $P = (x - \lambda)Q$.

Beweis. Sei h ein Teiler von f und g. Dann existieren $q, p \in \mathbb{K}[x]$ mit $f = qh$ und $g = ph$. Wir setzen λ_1 in f und g ein und bekommen $q(\lambda_1)h(\lambda_1) = f(\lambda_1) = 0$ und $p(\lambda_1)h(\lambda_1) = g(\lambda_1) = (\lambda_1 - \mu_1)^{\beta_1} ... (\lambda_1 - \mu_m)^{\beta_m} \neq 0$.Dann muss $q(\lambda_1) = 0$; daraus folgt (Lemma 27 Vorl. 12 LA I), dass $q = (x - \lambda_1)q_1$; dann gilt $(x - \lambda_1)^{\alpha_1 - 1} ... (x - \lambda_k)^{\alpha_k} = q_1 h$. Wir wiederholen diese Argumentation $(\alpha_1 - 1)$-mal für λ_1 (und bekommen, dass $(x - \lambda_2)^{\alpha_2} ... (x - \lambda_k)^{\alpha_k} = q_{\alpha_1} h$, dann α_2-mal für λ_2 usw.; wir bekommen dann $h \in \mathbb{K} \setminus \{0\}$ was unser Ziel war. □

Satz 5. Zu je zwei Polynomen f und g gibt es einen größten gemeinsamen Teiler h. Außerdem gilt: es gibt Polynome a, b sodass $af + bg = h$. Ferner gilt: Der größte gemeinsame Teiler ist bis auf einen Faktor aus $\mathbb{K} \setminus \{0\}$ eindeutig.

Beweis. Es sei $F = \{af + bg \mid a, b \in \mathbb{K}[x]\}$ die Menge aller "Linearkombinationen mit Koeffizienten in $\mathbb{K}[x]$" von f und g. Man wähle ein $h \in F$, so dass es den kleinsten Grad in $F \setminus \{0\}$ hat. Wir zeigen jetzt, dass h ein größter gemeinsamer Teiler ist.
Wir zeigen zuerst, dass h ein Teiler von f ist. Wir dividieren (mit Rest) f durch h und bekommen die Polynome q und r mit $f = qh + r$. Da r einen kleineren Grad als h hat und auch in F liegt, folgt $r = 0$. Also: $f = qh$.
Analog zeigt man, dass h das Polynom g teilt. Also ist h ein gemeinsamer Teiler. Nach Konstruktion ist $h = af + bg$.
Jetzt zeigen wir, dass h der größte gemeinsame Teiler ist. Sei h' ein gemeinsamer Teiler von g und f; wir müssen zeigen, dass h' das Polynom h teilt.
Da h' die Polynome f und g teilt, teilt es jede "Linearkombination" $af + bg$, also jedes Polynom aus F. Da $h \in F$ liegt, teilt h' auch h.

Def. h heißt gemeinsamer Teiler von f und g, falls h ein Teiler von f und g ist. Ein gemeinsamer Teiler h ist ein größter gemeinsamer Teiler, wenn jeder gemeinsame Teiler ein Teiler von h ist.

Satz 5. Zu je zwei Polynomen f und g gibt es einen größten gemeinsamen Teiler h. Außerdem gilt: es gibt Polynome a, b sodass $af + bg = h$. Ferner gilt: Der größte gemeinsame Teiler ist bis auf einen Faktor aus $\mathbb{K} \setminus \{0\}$ eindeutig.

Jetzt zeigen wir, das daß ggT eindeutig (bis auf Multiplizieren mit $k \in \mathbb{K} \setminus \{0\}$) ist. Angenommen, h' ist noch ein ggT von f und g; wir müssen zeigen, dass $h' = kh$ für ein $k \in \mathbb{K} \setminus \{0\}$.
Nach Definition von ggT muss dann h' das Polynom h teilen, und h muss h' teilen. Daraus folgt, dass $Grad(h') = Grad(h)$. Dann ist $h = k \cdot h'$; da $Grad(h) = Grad(h')$ gilt, dass das Polynom k Grad 0 hat, also $k \in \mathbb{K} \setminus \{0\}$. □

Satz 6 Sei $f = (x-\lambda_1)^{\gamma_1}...(x-\lambda_k)^{\gamma_k} \in \mathbb{K}[x]$, wobei λ_i paarweise verschieden sind . Sei $A \in Mat(n,n,\mathbb{K})$. Dann ist
$Kern_{f(A)} = Kern((A-\lambda_1 \cdot Id)^{\gamma_1}) \oplus \cdots \oplus Kern((A-\lambda_k \cdot Id)^{\gamma_k})$.

Beweis. Induktion nach k. IA: Für $k=1$ ist die Aussage offensichtlich: links und rechts stehen gleiche Ausdrücke.
IV: die Aussage gelte für alle $k-1$.
IS: $k-1 \to k$: Setze $f_1 := (x-\lambda_1)^{\gamma_1}$, $f_2 := (x-\lambda_2)^{\gamma_2} \cdots (x-\lambda_k)^{\gamma_k}$.
In wicht. Bsp. haben wir gezeigt, dass $ggT(f_1,f_2)=1$; aus Satz 5 folgt dann, dass $\exists h_1, h_2 \in \mathbb{K}[x]$ so dass $h_1 \cdot f_1 + h_2 \cdot f_2 = 1$.
Wir zeigen: $Kern_{f(A)} = Kern_{f_1(A)} \oplus Kern_{f_2(A)}$. Dann folgt die Aussage nach InduktionsVoraussetzung für $Kern_{f_2(A)}$.

Schema.

Wir zeigen $Kern_{f_i(A)} \overset{(a)}{\subseteq} Kern_{f(A)}$, $Kern_{f_1(A)} + Kern_{f_2(A)} \overset{(b)}{\supseteq} Kern_{f(A)}$.

Daraus folgt, dass $Kern_{f_1(A)} + Kern_{f_2(A)} = Kern_{f(A)}$.

Dann ((c)) zeigen wir , dass $Kern_{f_1(A)} \cap Kern_{f_2(A)} = \{\vec{0}\}$.

Daraus wird folgen, dass die Summe direkt ist.

(a) Behauptung: $Kern_{f_1(A)} \subseteq Kern_{f(A)}$.

Beweis von (a): ist $v \in Kern_{f_1(A)}$, so ist

$f(A)v = (f_1(A)f_2(A))v \overset{\text{da } P(A)Q(A) = Q(A)P(A}{=} (f_2(A)f_1(A))v \overset{\text{Weil } v \in Kern_{f_1(A)}.}{=}$
$f_2(A)(\vec{0}) = \vec{0}. \Longrightarrow Kern_{f_1(A))} \subseteq Kern_{f(A)}$.

Analog: $Kern_{f_2(A)} \subseteq Kern_{f(A)}$.
Es folgt: $Kern_{f_1(A)} + Kern_{f_2(A)} \subseteq Kern_{f(A)}$.

(b): Behauptung: $Kern_{f_1(A)} + Kern_{f_2(A)} \supseteq Kern_{f(A)}$.

Beweis von (b): Sei $v \in Kern_{f(A)}$.
Wie wir oben wiederholt haben, gibt es h_1, $h_2 \in \mathbb{K}[x]$, sodass
$1 = h_1 \cdot f_1 + h_2 \cdot f_2$. Da $1(A) \overset{Def.}{=} Id$, folgt
$v = Id(v) = \underbrace{h_1(A)f_1(A)v}_{:=v_2} + \underbrace{h_2(A)f_2(A)v}_{:=v_1}$.
Wir haben:
$f_2(A)(v_2) = f_2(A)h_1(A)f_1(A)v = h_1(A)f(A)v = h_1(A)(\vec{0}) = \vec{0}$, also
$v_2 \in Kern_{f_2(A)}$.
Analog: $v_1 \in Kern_{f_1(A)}$.
Dann ist jedes v eine Summe von $v_1 \in Kern_{f_1(A)}$ und
$v_2 \in Kern_{f_2(A)}$, folglich ist $Kern_{f_1(A))} + Kern_{f_2(A)} \supseteq Kern_{f(A)}$.

(c) Behauptung: $Kern_{f_1(A)} \cap Kern_{f_2(A)} = \{\vec{0}\}$.

Beweis von (c). Sei $v \in Kern_{f_1(A)} \cap Kern_{f_2(A)}$. Dann ist

$$v = Id(v) \overset{\text{wie in (b)}}{=} h_1(A) \underbrace{f_1(A)v}_{:=\vec{0}} + h_2(A) \underbrace{f_2(A)v}_{:=\vec{0}} = \vec{0}.$$

Aus (c) folgt, dass die Summe $Kern_{f_1(A)} + Kern_{f_2(A)}$ direkt ist. Nach Definition müssen wir zeigen, dass jedes v eindeutig als Summe $\underbrace{v_1}_{\in Kern_{f_1(A)}} + \underbrace{v_2}_{\in Kern_{f_2(A)}}$ darstellbar ist. Sei

$$v = \underbrace{v_1}_{\in Kern_{f_1(A)}} + \underbrace{v_2}_{\in Kern_{f_2(A)}} = \underbrace{v_1'}_{\in Kern_{f_1(A)}} + \underbrace{v_2'}_{\in Kern_{f_2(A)}}.$$

Dann ist $\underbrace{v_1 - v_1'}_{\in Kern_{f_1(A)}} = \underbrace{v_2' - v_2}_{\in Kern_{f_2(A)}} \in Kern_{f_1(A)} \cap Kern_{f_2(A)}$.

Wegen $Kern_{f_1(A)} \cap Kern_{f_2(A)} = \{\vec{0}\}$ ist dann $v_1 - v_1' = \vec{0}$ und $v_2 - v_2' = \vec{0}$, □

Gleichzeitig haben wir den Induktionsschritt gemacht und damit Satz 6 bewiesen.

Folgerung = Satz 4 vor einer Woche:

$Min_A = (\lambda_1 - x)...(\lambda_k - x)$, wobei λ_i paarweise verschieden sind $\iff$ A ist diagonalisierbar.

Die Richtung „$\Longleftarrow$“ ist einfach: man muss zeigen, dass Minimalpolynom einer Diagonalmatrix die Form $(\lambda_1 - x)...(\lambda_k - x)$ mit paarweise verschieden λ_i hat, siehe Vorl. 1, falls es nicht offensichtlich ist.

Beweis „$\Longrightarrow$“: Da $Min_A(A) = \mathbf{0}$, ist $Kern_{Min_A(A)} = V$. Nach Satz 6 ist $V = \underbrace{Kern_{A-\lambda_1 \cdot Id}}_{=Eig_{\lambda_1}} \oplus \cdots \oplus \underbrace{Kern_{A-\lambda_k \cdot Id}}_{=Eig_{\lambda_k}}$.

Seien $(b_1, ..., b_{Geo_{\lambda_1}})$, $(b_{Geo_{\lambda_1}+1}, ..., b_{Geo_{\lambda_1}+Geo_{\lambda_2}})$, ..., $(b_{Geo_{\lambda_1}+...+Geo_{\lambda_{k-1}}+1}, ..., b_{Geo_{\lambda_1}+...+Geo_{\lambda_k}})$ Basen jeweils in $Eig_{\lambda_1}, Eig_{\lambda_2}, ..., Eig_{\lambda_k}$. Dann ist die Menge $\{b_1, ..., b_{Geo_{\lambda_1}+...+Geo_{\lambda_k}}\}$

(i) erzeugend, weil man nach Satz 6 jedes v als Summe $\underbrace{v_1}_{\in Eig_{\lambda_1}} + ... + \underbrace{v_k}_{\in Eig_{\lambda_k}}$ darstellen kann, und jedes v_i eine Linearkombination von Elementen der $i-$ten Basis sind.

(ii) EINDEUTIG erzeugend ist, weil man jedes v eindeutig als Summe $\underbrace{v_1}_{\in Eig_{\lambda_1}} + ... + \underbrace{v_k}_{\in Eig_{\lambda_k}}$ schreiben kann, und jedes v_i kann man eindeutig als Summe von Elementen von Basis von Eig_{λ_i} schreiben.

Dann ist das Tupel $(b_1, ..., b_{Geo_{\lambda_1}+...+Geo_{\lambda_k}})$ eine Basis. Da alle Elemente Eigenvektoren sind, ist A in dieser Basis nach Satz 32 LA I diagonalisierbar, $\square$

Verallgemeinerte Eigenräume

Sei $A \in Mat(n, n, \mathbb{K})$. Das Minimalpolynom $Min_A = (x - \lambda_1)^{\gamma_1} \cdots (x - \lambda_k)^{\gamma_k}$ zerfalle in Linearfaktoren mit paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$. Dann gilt nach Satz 6:

$$V = \underbrace{Kern((A - \lambda_1 \cdot Id)^{\gamma_1})}_{:=W_{\lambda_1}} \oplus \cdots \oplus \underbrace{Kern((A - \lambda_k \cdot Id)^{\gamma_k})}_{:=W_{\lambda_k}}.$$

Def. Die Räume $W_{\lambda_1}, ..., W_{\lambda_k}$ heißen verallgemeinerte Eigenräume.

Satz 6 in Worten: Das Minimalpolynom von $A : V \to V$ zerfalle in Linearfaktoren. Dann ist V direkte Summe von verallgemeinerten Eigenräumen.

Def. Es sei ϕ ein Endomorphismus des $\mathbb{K}$-Vektorraums V der Dimension $n < \infty$. Ein Untervektorraum $W \subseteq V$ heißt $\phi-$ invariant, falls $Bild_\phi(W) \subseteq W$.
Diesselbe Definition in der Matrizen-Sprache: sei $A \in Mat(n, n, \mathbb{K})$. Ein Untervektorraum $W \subseteq V$ heißt $A-$ invariant, falls $\{Aw : w \in W\} \subseteq W$.

Triv. Bsp. $\{\vec{0}\}$ ist ein $\phi-$ invarianter Untervektorraum; V selbst ist ein $\phi-$ invarianter Untervektorraum.

Wicht. Bsp.

Verallgemeinerte Eigenräume W_i (von ϕ) sind $\phi-$ invariante Untervektorräume.

Beweis. Für jedes $v \in W_\lambda$ ist
$(\phi - \lambda \cdot Id)^\gamma \circ \phi(v) = \phi \circ (\phi - \lambda \cdot Id)^\gamma(v) = \phi(\vec{0}) = \vec{0}$, □

Allgemeine Überlegung

Das Minimalpolynom $Min_\phi = (x-\lambda_1)^{\gamma_1} \cdots (x-\lambda_k)^{\gamma_k}$ zerfalle in Linearfaktoren. Für jedes $i = 1, ..., k$ sei B_i eine Basis in W_{λ_i}. Wir setzen die Basen B_i zu einer Basis $\mathrm{B} = (B_1, ..., B_k)$ von V zusammen (Beweis, dass dies eine Basis ist haben wir in Beweis von Folgerung aus Satz 6 gemacht).

Da $W_{\lambda i}$ $\phi-$invariant ist, ist für jedes $b \in B_i$ der Vektor $\phi(b)$ eine Linearkommbination der Vektoren aus der Basis B_i, damit ist die Matrix A von ϕ bzgl.der Basis B blockdiagonal:

$$A := \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{pmatrix},$$

wobei A_i die Matrix von $\phi_{|W_{\lambda_i}}$ in der Basis B_i ist.

Wir haben also eine Basis konstruiert, so dass die Matrix von ϕ blockdiagonal ist. Wir werden die Basis noch verbessern.

Dazu werden wir $\phi_{|W_i} - \lambda_i \cdot Id$ untersuchen.

Nilpotente Endomorphismen

Def. Ein Endomorphismus $\phi : W \to W$ heißt nilpotent, falls es ein $\gamma \in \mathbb{N}$ gibt mit $\phi^\gamma \equiv \mathbf{0}$.

Für $w \in W$ heißt $k \in \mathbb{N}$ die $\phi-$Periode von w, falls $\phi^{k-1}(w) \neq \vec{0}$, aber $\phi^k(w) = \vec{0}$.

Bemerkung. Es ist stets $\gamma \geq k$ (falls ϕ wie in Def. oben ist).

Lemma 3. Sei $\phi : W \to W$ nilpotent, sei $w \in W$ mit $\phi-$Periode k. Dann sind w,$\phi(w)$,$\phi^2(w)$, ..., $\phi^{k-1}(w)$ linear unabhängig.

Beweis: Sei $a_0 w + a_1\phi(w) + ... + a_{k-1}\phi^{k-1}(w) = \vec{0}$. $(*)$ Wir wenden ϕ^{k-1} an und bekommen
$a_0\phi^{k-1}(w) + \underbrace{a_1\phi^k(w)}_{\vec{0}} + ... + \underbrace{a_{k-1}\phi^{2k-2}(w)}_{\vec{0}} = \vec{0}$. Da $\phi^{k-1}(w) \neq \vec{0}$, folgt daraus dass $a_0 = 0$.
Dann $(*)$ ist $a_1\phi(w) + ... + a_{k-1}\phi^{k-1}(w) = \vec{0}$.
Anwendung von ϕ^{k-2} liefert uns $a_1 = 0$ u.s.w., $\square$

Def. Ein Endomorphismus $\phi : W \to W$ heißt nilpotent, falls es ein $\gamma \in \mathbb{N}$ gibt mit $\phi^\gamma \equiv \mathbf{0}$.

Für $w \in W$ heißt $k \in \mathbb{N}$ die $\phi-$Periode von w, falls $\phi^{k-1}(w) \neq \vec{0}$, aber $\phi^k(w) = \vec{0}$.

Für $w \in W$ mit $\phi-$ Periode k definieren wir
$Z_w = \mathrm{span}\{w, \phi(w), ..., \phi^{k-1}(w)\}$.
Aus Definition o folgt, dass Z_w ein Untervektorraum ist. Nach Lemma 3 sind die Vektoren $w, \phi(w), ..., \phi^{k-1}(w)$ linear unabhängig. Deswegen ist Z_w ein k-dimensionaler Untervektorraum von W.

Lemma 4 (Zerlegungslemma für nilpotente Endomorphismen) Sei W ein endlich-dimensionaler $\mathbb{K}-$Vektorraum und $\phi : W \to W$ ein nilpotenter Endomorphismus. Sei $w \in W$ ein Element mit maximaler $\phi-$Periode k. Dann existiert ein $\phi-$invarianter Untervektorraum $U \subseteq W$, so dass $W = Z_w \oplus U$.

Beweis. Wähle einen $\phi-$invarianten Untervektorraum $U \subseteq W$ mit maximaler Dimension, so dass $Z_w \cap U = \{\vec{0}\}$. Dann ist die Summe $Z_w + U \subseteq W$ eine direkte Summe: ist $z + u = z' + u'$, so ist $z - z' = u' - u \in Z_w \cap U$, d.h., $z = z', u = u'$.

Z.z.: $Z_w \oplus U = W$. Angenommen das wäre nicht der Fall: Dann gäbe es ein $v \in W$ mit $v \notin Z_w \oplus U$. Wähle j so, dass $\phi^{j-1}(v) \notin Z_w \oplus U$, aber $\phi^j(v) \in Z(w) \oplus U$: Ein solches j existiert, da ϕ nilpotent ist, und deswegen gilt für genügend großes j, dass $\phi^j(v) = \vec{0}$. Setze $x := \phi^{j-1}(v)$. Dann gilt $x \notin Z_w \oplus U$, aber $\phi(x) \in Z_w \oplus U$.
Schreibe $\phi(x) = \underbrace{a_0 w + ... + a_{k-1}\phi^{k-1}(w)}_{\in Z_w} + \underbrace{u}_{\in U}$.

$\vec{0} = \phi^k(x) = \phi^{k-1} \circ \phi(x) \Longrightarrow$
$= \phi^{k-1}(a_0 w + ... + a_{k-1}\phi^{k-1}(w) + u)$ nur $\phi^{k-1}(w)$, $\phi^{k-1}(u)$ dürfen $\neq 0$ sein
$= \underbrace{a_0\phi^{k-1}(w)}_{\in Z_w} + \underbrace{\phi^{k-1}(u)}_{\in U} \overset{\text{Weil die Summe direkt ist}}{\Longrightarrow} a_0 = 0, \phi^{k-1}(u) = 0$

Setze $y := x - (a_1 w + ... + a_{k-1}\phi^{k-2}(w))$.
Nach Konstruktion gilt:
$\phi(y) = \phi(x) - (a_1\phi(w) + ... + a_{k-1}\phi^{k-1}(w)) = u \in U$,

Wir zeigen: $y \notin Z_w \oplus U$. Sonst wäre auch
$x = y + a_1 w + ... + a_{k-1}\phi^{k-2}(w) \in Z_w \oplus U$.

Setze $U' := U \oplus span(y)$. (Die Summe ist direkt, weil $y \notin U$).
$dim(U') = dim(U) + 1$; U' ist $\phi-$invariant. Wir bekommen Widerspruch mit der Annahme, dass U maximale Dimension hat. □

Lemma 4 auf der Sprache von Matrizen. Sei A eine $(n \times n)$- Matrix über $\mathbb{K}$ mit $A^\gamma = 0$. Dann existiert eine Matrix $B \in GL(n, \mathbb{K})$, so dass

$$B^{-1}AB = \begin{pmatrix} \begin{matrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & & 0 & 1 \\ & & & & & 0 \end{matrix} & \\ & \boxed{\mathrm{D}} \end{pmatrix}.$$

Bemerkung. Da $A^\gamma = \mathbf{0}$, ist auch $D^\gamma = \mathbf{0}$. Dann kann man D auch (mit einem geeigneten Basiswechsel) in zwei Blöcke aufsplitten u.s.w.

Satz 7 (Zerlegungssatz für nilpotente Endomorphismen) Sei W ein endlich-dimensionaler $\mathbb{K}-$ Vektorraum, sei $\phi : W \to W$ ein nilpotenter Endomorphismus. Dann existieren $w_1, ...w_\ell \in W$, so dass $W = Z_{w_1} \oplus \cdots \oplus Z_{w_\ell}$ und die Dimensionen der Z_{w_j} sind bis auf die Reihenfolge eindeutig durch ϕ festgelegt.

Beweis: Die Existenz (der Zerlegung) folgt induktiv aus dem Zerlegungslemma.
Es gibt also eine Basis, sodass die Matrix von ϕ gleich

$$\begin{pmatrix} \boxed{C_1} & & & \\ & \boxed{C_2} & & \\ & & \ddots & \\ & & & \boxed{C_\ell} \end{pmatrix}, \text{ wobei } C_i = \boxed{\begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & & 0 & 1 \\ & & & & & 0 \end{pmatrix}}.$$

Verschiedene C_i können selbstverständlich verschiedene Dimensionen haben.

Wir müssen jetzt zeigen, dass der Endomorphismus die Anzahl und Dimensionen der Blöcke bestimmt.

Sie werden es zu Hause ausrechnen:Für

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ gilt: $rk(C) = n - 1$ (in unserem Fall $5 - 1$).

Für $C^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ gilt: $rk(C^2) = n - 2$ (in unserem fall 3).

Für $C^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ gilt: $rk(C^3) = n - 3$ (in unserem Fall 2).

$$\vdots$$

$$C^5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ gilt: $rk(C^5) = 0.$ $\quad C^6 = \mathbf{0}$;
$rk(C^6) = rk(C^5) = 0.$

Das ist immer der Fall:
$dim(Kern_{\phi|Z_{w_j}}) = 1$. (Weil
$\phi(a_0 w_j + ... + a_{k-1}\phi^{k-1}(w_j)) = a_0\phi(w_j) + ... + a_{k-2}\phi^{k-1}(w)$ ist genau dann $\vec{0}$,wenn alle $a_0, ..., a_{k-2} = 0$.)

Bezeichnet n_j die ϕ-Periode von w_j, dann gilt für $m \in N$:

$$Kern_{(\phi|Z_{w_j})^m} = \begin{cases} span\left(\phi^{n_j-m}(w_j), ..., \phi^{n_j-1}(w_j)\right), & \text{falls } n_j > m \\ Z_{w_j}, & \text{falls } m \geq n_j \end{cases}$$

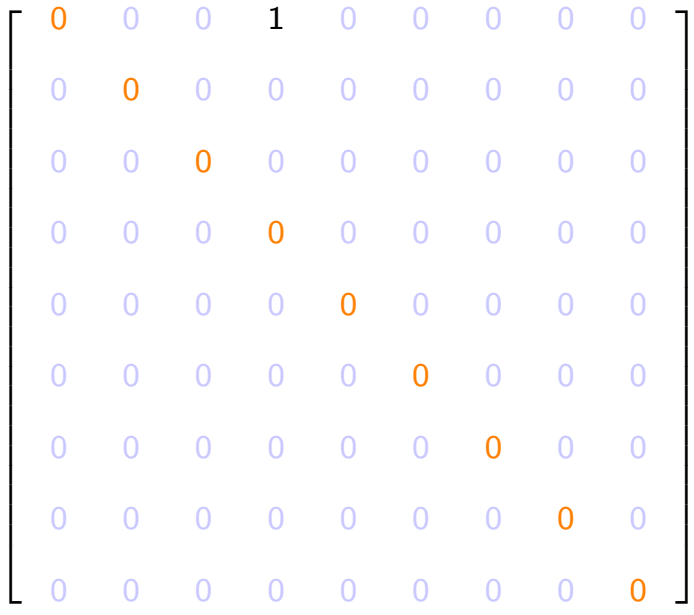
Wir werden diese Beobachtung nutzen um Anzahl und Dimensionen aller Blöcke zu bestimmen; nach dem Beispiel wird es hoffentlich offensichtlich.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

wir sehen: $rk(A) = \underbrace{dim(Bild_\phi)}_{\text{von } \phi \text{ bestimmt}} = n -$ {Anzahl von Kästchen}. Dann ist die Anzahl von Kästchen durch ϕ eindeutig bestimmt.

$$A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

wir sehen: $rk(A^2) = \underbrace{dim(Bild_{\phi^2})}_{\text{von } \phi \text{ bestimmt}} = n -$ {Anzahl von Kästchen}− { Anzahl von Kästchen von Dimension ≥ 2 } . Dann ist die Anzahl von Kästchen der Dimension ≥ 2 durch ϕ eindeutig bestimmt.

$A^3 =$ 

wir sehen: $rk(A^3) = \underbrace{dim(Bild_{\phi^3})}_{\text{von } \phi \text{ bestimmt}} = n -$
$\{\text{Anzahl von Kästchen}\} -$
$\{$ Anzahl von Kästchen
von Dimension ≥ 2 $\}$
$-\{$ Anzahl von Kästchen
von Dimension ≥ 3 $\}$
.

Dann ist Anzahl von Kästchen der Dimension ≥ 3 durch ϕ eindeutig bestimmt.

Analoges gilt für eine beliebige nilpotente Matrix.

Satz 7 in der Matrizen-Sprache: Sei A eine $(n \times n)$- Matrix über $\mathbb{K}$ mit $A^\gamma = \mathbf{0}$. Dann gibt es eine Matrix $B \in GL(n, \mathbb{K})$, so dass die Matrix $B^{-1}AB$ gegeben ist durch $\begin{pmatrix} \boxed{C_1} & & & \\ & \boxed{C_2} & & \\ & & \ddots & \\ & & & \boxed{C_\ell} \end{pmatrix}$, wobei

$$C_i = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

Bemerkung.

$B^{-1}(A - \lambda \cdot Id)B \overset{\text{Linearität}}{=} B^{-1}AB - \lambda B^{-1}IdB = B^{-1}AB - \lambda \cdot Id.$

Deswegen ist (falls $(A - \lambda \cdot Id)^\gamma = \mathbf{0}$ wie in Satz 6) $B^{-1}AB =$

$$\begin{pmatrix} \boxed{C_1} & & & \\ & \boxed{C_2} & & \\ & & \ddots & \\ & & & \boxed{C_\ell} \end{pmatrix} + \lambda \cdot Id.$$

Jordansche Normalform

Def. *Die $k \times k$ Matrix* $J_\lambda^k := \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$ *heißt Jordan-Block.*

Bsp. $J_2^1 = (2)$, $J_1^2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $J_{1+i}^3 = \begin{pmatrix} 1+i & 1 & \\ & 1+i & 1 \\ & & 1+i \end{pmatrix}$

Bemerkung. $alg_{J_\lambda^k}(\lambda) = k$ (da $\chi_{J_\lambda^k} = (\lambda - t)^k$), $geo_{J_\lambda^k} = 1$ (da $J_\lambda^k - \lambda \cdot Id_k = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$ ist; deren Rang ist $k-1$ und deswegen $dim(Kern_{J_\lambda^k - \lambda \cdot Id_k}) = k - (k-1) = 1$. Also, Jordan-Block J_λ^k ist für $k \geq 2$ nicht diagonalisierbar.)

Die Matrix der Form $\begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & \\ & \ddots & \\ & & \boxed{J_{\lambda_m}^{k_m}} \end{pmatrix}$, *wobei* $J_{\lambda_j}^{k_j}$ *Jordan-Blöcke sind,*

heißt Jordan-Matrix.

Bsp. Die Matrizen
$\begin{pmatrix} 2 & & \\ & 3 & \\ & & 4 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{pmatrix}$ sind Jordan-Matrizen.

Satz 8 (Jordansche Normalform) Sei $\mathbb{K}$ ein Körper, sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und sei $\phi : V \to V$ ein Endomorphismus, dessen Minimalpolynom in Linearfaktoren zerfällt. Dann existiert eine Basis von V , so dass die Matrix von ϕ eine Jordan-Matrix ist. Diese Jordansche Normalform ist bis auf Reihenfolge der Jordanblöcke eindeutig.

Satz 8 ist eine einfache Folgerung der Sätze 6, 7

Satz 6 – Wiederholung Es gibt eine Basis, so dass in der Basis die Matrix von ϕ die Form $\begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{pmatrix}$ hat, wobei

$(A_i - \lambda_i \cdot Id)^{\gamma_i} = \mathbf{0}$.

Aus Satz 7, siehe auch Bemerkung oben: Für jedes A_i (der Dimension n_i) mit der Eigenschaft $(A_i - \lambda_i \cdot Id)^{\gamma_i} = 0$ gibt es eine

$$B_i \in GL(n_i, \mathbb{K}) \text{ mit } B_i^{-1} A_i B_i = \begin{pmatrix} \boxed{C_1^i} & & & \\ & \boxed{C_2^i} & & \\ & & \ddots & \\ & & & \boxed{C_{l_i}^i} \end{pmatrix} + \lambda_i Id \ .$$

Dann ist

$$\begin{pmatrix} \boxed{B_1} & & & \\ & \boxed{B_2} & & \\ & & \ddots & \\ & & & \boxed{B_k} \end{pmatrix}^{-1} \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{pmatrix} \begin{pmatrix} \boxed{B_1} & & & \\ & \boxed{B_2} & & \\ & & \ddots & \\ & & & \boxed{B_k} \end{pmatrix}$$

$$= \begin{pmatrix} \boxed{B_1^{-1} A_1 B_1} & & & \\ & \boxed{B_2^{-1} A_2 B_2} & & \\ & & \ddots & \\ & & & \boxed{B_k^{-1} A_k B_k} \end{pmatrix} = \text{Wie wir wollen.}$$

Eideutigkeit der Jordan-Normalform folgt aus der Beobachtung, dass ϕ die Untervektorräume W_{λ_i} eindeutig bestimmt, und aus der Eindeutigkeitsaussage im Satz 7. $\square$

Anwendung: Beweis von Hamilton-Cayley für Matrizen über $\mathbb{C}$.

Satz 2  — 

1805 --1856 1821 --1895

$$\chi_A(A) = \mathbf{0}$$

In Worten: Charakteristisches Polynom einer Matrix annihiliert die Matrix.

Beobachtung 1 Für $J = \begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & \\ & \ddots & \\ & & \boxed{J_{\lambda_m}^{k_m}} \end{pmatrix}$, ist

$\chi_J = (\lambda_1 - t)^{k_1} ... (\lambda_m - t)^{k_m}$.
Tatsächlich, da J eine obere Dreiecksmatrix ist (=alle Einträge unterhalb der Hauptdiagonalen sind 0), ist $J - t \cdot Id$ auch eine obere Dreiecksmatrix, und deswegen ist $det(J - t \cdot Id)$ das Produkt von Diagonalementen, die $(\lambda_i - t)$ sind.

Beobachtung 2 Die Matrizen A, A' seien ähnlich: $A = B^{-1}A'B$. Dann gilt: Für jedes $P \in \mathbb{C}[t]$ ist $P(A) = B^{-1}P(A')B$.

Tatsächlich haben wir dies im Wesentlichen in Vorl. 20 LAAG I bewiesen:

$$A^k = \underbrace{A \cdot A \cdot \ldots \cdot A}_{k\text{mal}} = B^{-1}A'\underbrace{BB^{-1}}_{Id}A'B\ldots B^{-1}A'B = B^{-1}A'^kB.$$

Also, $P(A) = P(B^{-1}A'B) =$

$$a_kB^{-1}A'^kB + a_{k-1}B^{-1}A'^{k-1}B + \ldots + a_0B^{-1}B$$

$$\overset{\text{Linearität}}{=} B^{-1}(a_kA'^k + \ldots + a_0Id)B = B^{-1}P(A')B.$$

Beobachtung 3 Für eine Block-diagonale Matrix

$$M = \begin{pmatrix} \boxed{B_1} & & \\ & \ddots & \\ & & \boxed{B_m} \end{pmatrix}, \text{ wobei } B_i \text{ eine } k_i \times k_i\text{-Matrix ist, ist}$$

$$P(M) = \begin{pmatrix} \boxed{P(B_1)} & & \\ & \ddots & \\ & & \boxed{P(B_m)} \end{pmatrix}.$$

Beobachtung 4 $\chi_{J_\lambda^k}(J_\lambda^k) = \mathbf{0}$.
Tatsächlich,

$$\chi_{J_\lambda^k}(J_\lambda^k) = (\lambda \cdot Id - J_\lambda^k)^k = (-1)^k \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}^k \overset{\text{wie in Bsp. vorcher}}{=} \mathbf{0}$$

Alle vier Beobachtungen zusammen:

Für A, die ähnlich ist zu $J = \begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & \\ & \ddots & \\ & & \boxed{J_{\lambda_m}^{k_m}} \end{pmatrix}$, gilt

1. $\chi_J = (\lambda_1 - t)^{k_1} ... (\lambda_m - t)^{k_m} = \chi_{J_{\lambda_1}^{k_1}} \cdots \chi_{J_{\lambda_m}^{k_m}}$.
2. $P(A) \overset{\text{für geignetes } B}{=} B^{-1} P(J) B$, wobei J die Jordan'sche Normalform von A ist
3. $P(J) = \begin{pmatrix} \boxed{P(J_{\lambda_1}^{k_1})} & & \\ & \ddots & \\ & & \boxed{P(J_{\lambda_m}^{k_m})} \end{pmatrix}$,
4. $\chi_J(J_{\lambda_i}^{k_i}) = P_i(J_{\lambda_i}^{k_i}) \cdot \underbrace{\chi_{J_{\lambda_i}^{k_i}}(J_{\lambda_i}^{k_i})}_{\mathbf{0}} = \mathbf{0}$

Dann ist $\chi_A(A) = \chi_J(A) = B^{-1} \chi_J(J) B =$

$$B^{-1} \begin{pmatrix} \boxed{\chi_J(J_{\lambda_1}^{k_1})} & & \\ & \ddots & \\ & & \boxed{\chi_j(J_{\lambda_m}^{k_m})} \end{pmatrix} B = B^{-1} \mathbf{0} B = \mathbf{0},$$

□