Die Jordansche Normalform

VON MARKUS HAASE

§1 Einführung

Das Normalformenproblem kann man so formulieren:

> Gegeben sei eine $n \times n$–Matrix A über dem Körper $\mathbb{C}$. Gibt es in der Ähnlichkeitsklasse von A eine Matrix von einfacher Gestalt, an der man viele Eigenschaften von A direkt ablesen kann?

Das Problem der *Normal*form ist eigentlich noch etwas weiter gefasst. Man möchte ein Konstruktionsverfahren finden, das zu einer gegebenen Matrix A eine zu A ähnliche Matrix $\mathrm{Nf}(A)$ ermittelt, die sehr übersichtlich ist und zwar so, dass zwei Matrizen A und B genau dann zueinander ähnlich sind, wenn $\mathrm{Nf}(A) = \mathrm{Nf}(B)$ gilt.

Man interpretiert dazu A als lineare Selbstabbildung eines n-dimensionalen $\mathbb{C}$-Vektorraumes V (d.h. genauer als Darstellungsmatrix einer solchen Abbildung bezüglich einer Basis). Die zu A ähnlichen Matrizen entsprechen dann genau den Basen von V via Matrixdarstellung. Das Problem ist also, eine Basis zu konstruieren, bezüglich der die Abbildung eine übersichtliche und möglichst charakteristische Darstellung hat.

Ein zentrales Mittel dafür ist, den Vektorraum in kleinere Teilräume zu „zerlegen".

DEFINITION: Sei V ein (sagen wir mal endlichdimensionaler) Vektorraum über einem Körper $\mathbb{F}$, und $U_1, U_2, \ldots, U_k$ Unterräume von V. Dann heißt V die (innere) **direkte Summe** der U_i, in Zeichen: $V = U_1 \oplus \cdots \oplus U_k$, wenn folgende (zueinander äquivalente) Bedingungen erfüllt sind:

(i) Jeder Vektor $v \in V$ hat genau eine Darstellung $v = u_1 + \cdots + u_k$ mit $u_i \in U_i$ $(1 \leq i \leq k)$.

(ii) Ist B_i Basis von U_i $(1 \leq i \leq k)$ so ist $B_1 \cup \cdots \cup B_k$ Basis von V.

(iii) $V = U_1 + \cdots + U_k$ und $U_i \cap \sum_{j \neq i} U_j = 0$ für alle $1 \leq i \leq k$.

(Die behauptete Äquivalenz nachzuweisen ist eine leichte Übungsaufgabe!) Eine Basis von V wie in (ii) heißt eine der Zerlegung **angepasste** Basis.

Natürlich ist nichts gewonnen, wenn man den Vektorraum *irgendwie* zerlegt. Das Ganze muss schon was mit der Abbildung $A : V \to V$ zu tun haben. Darum noch eine

DEFINITION: Ein Teilraum U von V heißt A-**invariant**, wenn $A(U) \subset U$ gilt.

Jeder A-invariante Teilraum U ist auch $p(A)$-invariant für jedes Polynom $p \in \mathbb{C}[x]$ (Beweis?). Für $\lambda \in \mathbb{C}$ ist U genau dann A-invariant, wenn U $(A - \lambda)$-invariant ist.

Wenn man einen A-invarianten Teilraum U hat, dann kann man A auf U einschränken. Gilt $V = U_1 \oplus U_2$ und ist U_1 sogar A-invariant, dann hat A bezüglich einer angepassten Basis die Darstellung

$$\begin{pmatrix} \tilde{A}_1 & * \\ 0 & \tilde{A}_2 \end{pmatrix}$$

wobei $\tilde{A}_1$ die Darstellungsmatrix der Einschränkung von A auf U_1 ist.
Sind beide Teilräume U_1, U_2 A-invariant, so sieht die Darstellungsmatrix bezüglich einer angepassten Basis sogar noch besser aus, nämlich

$$\begin{pmatrix} \tilde{A}_1 & 0 \\ 0 & \tilde{A}_2 \end{pmatrix}.$$

Ist schliesslich $V = U_1 \oplus \cdots \oplus U_k$ eine direkte Zerlegung in A-invariante Teilräume U_i, so führt das zu einer Matrixdarstellung der Form

$$\begin{pmatrix} \tilde{A}_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{A}_k \end{pmatrix}.$$

Umgekehrt kommt jede Matrixdarstellung dieser Form von einer Basis B der Form $B = B_1 \cup \cdots \cup B_k$ derart, dass die B_i paarweise disjunkt sind, die $< B_i >$ jeweils A-invariante Teilräume bilden und $V = < B_1 > \oplus \cdots \oplus < B_k >$ ist.

Das Konstruktionsverfahren geht nun in zwei Schritten vor sich. Im ersten Schritt wird der Vektorraum in die sogenannten Primärkomponenten zerlegt. Auf den Primärkomponenten hat die lineare Abbildung dann jeweils nur noch einen Eigenwert und dann sucht man sich sogenannte Jordansche Ketten.

§2 Polynome und das Minimalpolynom

Die Algebra im Polynomring $\mathbb{C}[x]$ ist dadurch recht einfach, dass es eine *Division mit Rest* gibt:

SATZ: Sei $\mathbb{F}$ ein Körper, $f, g \in \mathbb{F}[x]$ zwei Polynome und $g \neq 0$. Dann gibt es Polynome $p, r \in \mathbb{F}[x]$ mit $f = pg + r$ so, dass $r = 0$ oder $\deg r < \deg g$ ist.

Dabei ist $\deg g$ der *Grad* des Polynoms g. (g hat den Grad $n \geq 0$ wenn $g = \sum_{j=0}^{n} a_j x^j$ ist mit $a_n \neq 0$.) Wenn $r = 0$ ist, also $f = pg$ für ein p ist, dann schreibt man dafür $g|f$ (in Worten: „g teilt f).“

[Zu einem Beweis der Aussage des Satzes mache man sich das herkömmliche Polynomdivisionsverfahren klar und kleide das Ganze in einen Beweis durch Induktion nach $\deg f$.]

Man betrachtet nun zu unserer linearen Abbildung $A : V \to V$ die Menge

$$\mathrm{Ann}_A(V) = \{f \in \mathbb{C}[x] : f(A) = 0\}.$$

Klar ist, dass das Nullpolynom in $\mathrm{Ann}_A(V)$ enthalten ist. $\mathrm{Ann}_A(V)$ enthält aber auch nichttriviale Polynome!

[Dies sieht man mit dem Satz von Cayley-Hamilton. Wenn man den nicht mag, kann man auch so argumentieren: Der Vektorraum der linearen Abbildungen des n-dimensionalen Vektorraums V ist n^2-dimensional (da isomorph zur Algebra der $n \times n$-Matrizen). Also ist die Menge $\{1, A, A^2, \ldots, A^{n^2}\}$ linear abhängig. Es gibt daher a_i nicht alle gleich 0, mit $a_0 + a_1 A + \cdots + a_{n^2} A^{n^2} = 0$.]

Man wählt nun in $\mathrm{Ann}_A(V) \setminus \{0\}$ ein normiertes Polynom $m_A \neq 0$ mit kleinstem Grad und nennt es **Minimalpolynom** von A. Dass es nur ein solches Polynom geben kann zeigt der folgende Satz.

SATZ: Sei $f \in \mathrm{Ann}_A(V)$. Dann gibt es ein Polynom $p \in \mathbb{C}[x]$ mit $f = p m_A$.

[Beweis: Nach dem Satz über die Division mit Rest gibt es p und r mit $f = p m_A + r$ und $r = 0$ oder $\deg r < \deg m_A$. Dann ist $0 = f(A) = p(A) m_A(A) + r(A) = r(A)$ Nach Definition von m_A muss also $r = 0$ sein.]

Nach dem *Fundamentalsatz der Algebra* zerfällt jedes Polynom über $\mathbb{C}$ in Linearfaktoren. Wir haben also
$$m_A(x) = \prod_{i=1}^{r} (x - \lambda_i)^{e_i}$$
für gewisse paarweise verschiedene $\lambda_i \in \mathbb{C}$ und $1 \leq e_i \in \mathbb{N}$. Der folgende Satz zeigt, dass die λ_i genau die Eigenwerte von A sind:

SATZ: Sei $\lambda \in \mathbb{C}$. Genau dann gilt $(x - \lambda) \mid m_A$, wenn λ ein Eigenwert von A ist.

[Beweis: „$\Rightarrow$“: Wenn $m_A(x) = (x - \lambda)\, p(x)$ ist, so folgt $0 = (A - \lambda)p(A)$. Ausserdem muss p normiert sein mit $\deg p = \deg m_A - 1$. Somit ist $p(A) \neq 0$ und es gibt also $v \in V$ mit $p(A)v \neq 0$. Dann ist $p(A)v$ ein Eigenvektor zu λ.
„$\Leftarrow$“: Sei λ Eigenwert mit Eigenvektor v. Es gibt $p, r \in \mathbb{C}[x]$ mit $m_A = p\,(x - \lambda) + r$ und $r = 0$ oder $\deg r < 1$. Dann ist r jedenfalls ein Skalar und $0 = p(A)(A - \lambda)v + rv = rv$. Wegen $v \neq 0$ folgt $r = 0$.]

Nimmt man die beiden letzten Sätze mit dem Satz von Cayley-Hamilton zusammen, dann ergibt sich
$$\chi_A(x) = \prod_{i=1}^{r} (x - \lambda_i)^{f_i}$$
mit $1 \leq e_i \leq f_i$. Die Zahl f_i heisst die **algebraische Vielfachheit** des Eigenwertes λ_i[1].

§3 Die Primärzerlegung

Wir kommen nun zu der sogenannten Primärzerlegung von V in eine direkte Summe A-invarianter Teilräume. Sie orientiert sich an der Primfaktorzerlegung von m_A.

DEFINITION: Sei λ ein Eigenwert von A und $k \in \mathbb{N}$. Dann heisst der Unterraum $\ker(A - \lambda)^k$ der **verallgemeinerte Eigenraum** zum Exponenten k (bzgl. λ).

Jeder verallgemeinerte Eigenraum ist A-invariant und es gilt $\ker(A - \lambda)^k \subset \ker(A - \lambda)^{k+1}$ für alle k. Man setzt $U_i := \ker(A - \lambda_i)^{e_i}$ für $1 \leq i \leq r$. Dann haben wir den folgende Satz.

SATZ ÜBER DIE PRIMÄRZERLEGUNG: Es ist $V = U_1 \oplus \cdots \oplus U_r$.

[Beweis: Bilde zu i das Polynom $p_i = m_A/(x - \lambda_i)^{e_i} = \prod_{j \neq i}(x - \lambda_j)^{e_j}$. Dann sind die p_i teilerfremd und daher gibt es Polynome q_i mit $p_1 q_1 + \ldots p_r q_r = 1$. (Satz über den grössten gemeinsamen Teiler) Natürlich gilt $p_i(A)V \subset U_i$, denn $(A - \lambda)^{e_i} p(A)V = m_A(A)V = 0$. Also ist $v = \sum p_i(A)q_i(A)v \in \sum U_i$.
Ist $0 = u_1 + \cdots + u_r$ mit $u_i \in U_i$ so gilt für jedes i: $0 = p_i(A)(\sum_j u_j) = \sum_j p_i(A)u_j = p_i(A)u_i$, denn für $j \neq i$ ist $(x - \lambda_j)^{e_j} \mid p_i$ und somit $p_i(A)U_j \subset 0$. Die Polynome p_i und $(x - \lambda_i)^{e_i}$ sind teilerfremd, also gibt es Polynome $a, b \in \mathbb{C}[x]$ mit $1 = ap_i + b(x - \lambda_i)^{e_i}$. Daraus ergibt sich $u_i = a(A)p_i(A)u_i + b(A)(A - \lambda_i)^{e_i}u_i = 0$.]

Für spätere Zwecke notieren wir auch noch den

EINDEUTIGKEITSSATZ: Sei $V = \tilde{U}_1 \oplus \cdots \oplus \tilde{U}_r$. Dabei sei für jedes i der Raum $\tilde{U}_i$ A-invariant, so dass $A|U_i$ nur den Eigenwert λ_i besitzt. Dann ist $U_i = \tilde{U}_i$ für alle i.

[Beweis: Betrachte $u \in \tilde{U}_i$. Dann ist $u = \sum_j u_j$ mit $u_j \in U_j$. Für genügend grosses k ist $(A - \lambda_i)^k u = 0$, also $0 = \sum_j (A - \lambda_i)^k u_j$ und wegen der A-Invarianz der U_j und der Direktheit der Summe ist notwendig $(A - \lambda_i)^k u_j = 0$ für alle j. Wenn $j \neq i$ ist, dann besitzt $A|U_j$ nur den Eigenwert $\lambda_j \neq \lambda_i$; somit ist $u_j = 0$ für alle solche j. Es folgt $u = u_i \in U_i$.
Dies liefert $\tilde{U}_i \subset U_i$. Die Umkehrung verläuft vollständig analog.]

[1]Es gibt einen sehr elementaren Beweis des Satzes von Cayley-Hamilton, der mit den Kofaktormatrizen zu A arbeitet. Der gerade bewiesene Satz zeigt, dass charakteristisches und Minimalpolynom dieselben Nullstellen besitzen. Dass die f_i immer mindestens so groß wie die e_i sind, ist dann die Aussage von Cayley-Hamilton. Wir benutzen aber im folgenden dieses Resultat nicht. Vielmehr wird es sich als Konsequenz unserer weiteren Überlegungen ergeben.

§4 Verfahren zur Berechnung des Minimalpolynoms

Wir gehen davon aus, dass wir die verschiedenen Eigenwerte $\{\lambda_1, \ldots, \lambda_r\}$ der Abbildung A kennen. Wie bestimmt man die Exponenten e_i der Linearfaktoren $(x-\lambda_i)$ im Minimalpolynom?

Sei $1 \leq i \leq r$ fest. Wie schon gesehen, werden die verallgemeinerten Unterräume $\ker(A-\lambda_i)^k$ mit wachsendem k immer grösser:

$$\ker(A-\lambda_i)^k \subset \ker(A-\lambda_i)^{k+1}.$$

Gilt beim Übergang von k zu $k+1$ sogar Gleichheit, dann auch beim Übergang von $k+1$ zu $k+2$.

[Wenn $v \in \ker(A-\lambda_i)^{k+2}$ ist, dann ist $(A-\lambda_i)v \in \ker(A-\lambda_i)^{k+1}$. Also auch $(A-\lambda_i)v \in \ker(A-\lambda_i)^k$ nach Voraussetzung und damit $v \in \ker(A-\lambda_i)^{k+1}$.]

Es gilt nun das folgende

LEMMA: Sei $m_A(x) = \prod_{i=1}^r (x-\lambda_i)^{e_i}$ das Minimalpolynom von A mit paarweise verschiedenen λ_i. Sei weiter $1 \leq i \leq r$. Dann gibt es einen Vektor $v \in V$ mit $(A-\lambda_i)^{e_i-1}v \neq 0$ und $(A-\lambda_i)^{e_i}v = 0$. Weiter gilt $\ker(A-\lambda_i)^{e_i+1} = \ker(A-\lambda_i)^{e_i}$.

[Beweis: Sei $u \in \ker(A-\lambda_i)^{e_i+1}$. Man kann schreibt mithilfe der Primärzerlegung $u = u_1 + \cdots + u_r$ mit $u_j \in U_j$. Dann ist $(A-\lambda_i)^{e_i}u = \sum_{j\neq i}(A-\lambda_i)^{e_i}u_j \in \sum_{j\neq i} U_j \cap U_i = 0$. Das ergibt die zweite Behauptung,
Wäre $(A-\lambda_i)^{e_i-1}u = 0$ für alle $u_i \in U_i$, dann wäre $(A-\lambda_i)^{e_i-1}\prod_{j\neq i}(A-\lambda_j)^{e_j}u = 0$ für *alle* $u \in V$ (man schreibe wieder $u = \sum_j u_j$ mit $u_j \in U_j$). Das würde aber bedeuten, dass $p(A) = 0$ ist für $p(x) = (x-\lambda_i)^{e_i-1}\prod_{j\neq i}(x-\lambda_j)^{e_j}$ mit Widerspruch zur Definition von m_A.]

Das Lemma liefert uns sofort eine Methode, um die Zahlen e_i zu bestimmen. Es muss nämlich

$$e_i = \min\{k \in \mathbb{N} : \ker(A-\lambda_i)^k = \ker(A-\lambda_i)^{k+1}\}$$

gelten. Man hat also sukzessive die Dimensionen der Räume $\ker(A-\lambda_i)^k$ zu bestimmen. Die entstehende Folge natürlicher Zahlen wächst streng monoton bis $k = e_i$, dann ist sie konstant.

§5 Die Bedeutung der algebraischen Vielfachheit

Was ist mit der Primärzerlegung gewonnen? Nun, da alle U_i invariant sind, kann man A auf den Unterraum U_i einschränken: $A_i := A|_{U_i} : U_i \to U_i$. Der Vorteil ist, dass die Abbildung A_i nur noch einen Eigenwert hat:

$$m_{A_i}(x) = (x-\lambda_i)^{e_i}$$

[Beweis: Man weiss, dass $(A-\lambda_i)^{e_i}v = 0$ ist für alle $v \in U_i$ (nach Definition von U_i). Also ist $(A_i-\lambda_i)^{e_i} = 0$. Das Minimalpolynom m_{A_i} teilt also das Polynom $(x-\lambda_i)^{e_i}$. Aufgrund der eindeutigen Primfaktorzerlegung im Polynomring muss daher $m_{A_i} = (x-\lambda_i)^k$ für ein $k \leq e_i$ sein. Nach dem Lemma aus dem letzten Abschnitt ist aber $\ker(A-\lambda_i)^{e_i-1} \neq \ker(A-\lambda_i)^{e_i}$, d.h. $(A-\lambda_i)^{e_i-1}U_i \neq 0$. Somit $k = e_i$.]

Da λ_i der einzige Eigenwert von A_i ist, gilt für das charakteristische Polynom von A_i:

$$\chi_{A_i}(x) = (x-\lambda_i)^{\dim U_i}.$$

Sei nun B_i eine Basis von U_i und $\tilde{A}_i$ die Matrixdarstellung von A_i bezgl. B_i. Dann ist $B = B_1 \cup \cdots \cup B_r$ eine der Primärzerlegung angepasste Basis, bezüglich der die Abbildung A die Darstellung

$$\begin{pmatrix} \tilde{A}_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{A}_r \end{pmatrix}$$

hat. Der Determinantenmultiplikationssatz zeigt nun $\chi_A = \chi_{A_1} \cdot \dots \cdot \chi_{A_r}$. Da der Linearfaktor $(x - \lambda_i)$ nur in χ_{A_i} vorkommt, muss schliesslich

$$f_i = \dim U_i = \dim \ker(A - \lambda_i)^{e_i}$$

gelten.

§6 Jordansche Ketten

Ab jetzt betrachten wir eine lineare Abbildung $A : V \to V$, die genau *einen* Eigenwert λ hat; also gilt

$$m_A(x) = (x - \lambda)^e, \quad \chi_A(x) = (x - \lambda)^f.$$

Weiter definieren wir für $1 \leq k \leq e$: $V_k := \ker(A - \lambda)^k$ und $V_0 := 0$. Wie schon gesehen, gilt $V_0 \subset V_1 \subset \dots \subset V_e = V$ und die Inklusionen sind alle echt.

Der Raum V_1 ist offensichtlich der Eigenraum zum Eigenwert λ, seine Dimension $\dim \ker(A - \lambda)$ heisst die **geometrische Vielfachheit** von λ.

Für einen Vektor $v \in V$ nennen wir

$$\operatorname{lev}(v) = \min\{k : v \in V_k\}$$

den **Level** von v. Anwenden von $(A - \lambda)$ auf v vermindert den Level um eins; induktiv ergibt sich:

$$\operatorname{lev}((A - \lambda)^k v) = \operatorname{lev}(v) - k$$

für $0 \leq k \leq \operatorname{lev}(v)$.

DEFINITION: Ein Tupel $(v_1, v_2, \dots, v_m)$ von Vektoren von V heisst **Jordansche Kette**, wenn gilt: $v_1 \neq 0$ und

$$(A - \lambda)v_1 = 0,\ (A - \lambda)v_2 = v_1,\ \dots,\ (A - \lambda)v_m = v_{m-1}.$$

Sei $(v_1, v_2, \dots, v_m)$ ein Jordansche Kette. Dann ist offensichtlich $\operatorname{lev}(v_k) = k$ für alle k. Der von den v_i aufgespannte Unterraum ist A-invariant(!) und die Menge $\{v_1, \dots, v_m\}$ ist linear unabhängig.

[Beweis: (Induktion über m) als Übungsaufgabe.]

Sei für den Moment $U := \operatorname{span}(v_1, \dots, v_m)$. Dann können wir A auf U einschränken (A-Invarianz!) und bezüglich der Basis $\{v_1, \dots, v_m\}$ darstellen. Wir erhalten die Matrix

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}.$$

Eine solche Matrix heisst **Jordankästchen** zum Eigenwert λ.

Die Idee der Jordanschen Normalform besteht nun darin, eine Basis zu suchen, die sich als disjunkte Vereinigung von Jordanketten schreiben lässt. Dies entspricht einer Zerlegung von V in A-invariante Teilräume, so dass die darstellende Matrix bezüglich einer angepassten Basis die Form

$$\begin{pmatrix} \tilde{J}_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{J}_l \end{pmatrix}$$

hat, wobei die $\tilde{J}_i$ Jordankästchen zum Eigenwert λ sind.

§7 Konstruktion der Jordanschen Normalform

Wir bleiben bei der Ausgangssituation (genau ein Eigenwert λ) und Terminologie des letzten Abschnittes. Um die Konstruktion durchsichtiger zu machen, schicken wir ein Lemma voraus.

LEMMA: Sei U eine Unterraum mit $V_1 \cap U = 0$. Dann ist $(A-\lambda)|_U$ injektiv. Insbesondere gilt: wenn C eine Basis von U ist, so ist $(A-\lambda)C$ eine Basis von $(A-\lambda)U$.

[Beweis: Das ist trivial.]

Man geht induktiv vor. Zuerst sucht man sich (mittels Basisergänzungssatz) einen Unterraum U_e mit Basis C_e

$$V = V_{e-1} \oplus U_e.$$

Dann gilt $V_{e-2} \cap (A-\lambda)U_e = 0$. Wieder mit Basisergänzung findet sich ein Unterraum U_{e-1} mit Basis C_{e-1}, so dass

$$V_{e-1} = V_{e-2} \oplus (A-\lambda)U_e \oplus U_{e-1}$$

gilt.

Angenommen $k \leq 2$ und man hat schon Unterräume $U_e, \dots, U_k$ mit Basen C_i konstruiert, so dass

$$V_k = V_{k-1} \oplus U_k \oplus (A-\lambda)U_{k+1} \oplus \cdots \oplus (A-\lambda)^{e-k}U_e \tag{1}$$

gilt. Wegen $(A-\lambda)V_{k-1} \subset V_{k-2}$ gilt nun:

$$(A-\lambda)V_k \subset V_{k-2} \oplus (A-\lambda)U_k \oplus \cdots \oplus (A-\lambda)^{e-(k-1)}U_e.$$

(Die Pointe liegt in der Aussage, dass die Summe auf der rechten Seite *direkt* ist.)

[Beweis: Angenommen $0 = v + (A-\lambda)u_k + \dots (A-\lambda)^{e-k+1}u_e$ mit $v \in V_{k-2}$ und $u_i \in U_i$. Anwendung von $(A-\lambda)^{k-2}$ ergibt $0 = (A-\lambda)^{k-1}(u_k + \cdots + (A-\lambda)^{e-k}u_e$, also $(u_k + \cdots + (A-\lambda)^{e-k}u_e \in V_{k-1}$. Die Induktionsannahme liefert nun $u_k + \cdots + (A-\lambda)^{e-k}u_e = 0$ was wiederum mit der Induktionsannahme zur Folge hat, dass alle Summanden gleich Null sein müssen.]

Also kann man sich wieder einen Unterraum U_{k-1} mit Basis C_{k-1} suchen, so dass

$$V_{k-1} = V_{k-2} \oplus U_{k-1} \oplus (A-\lambda)U_k \oplus \cdots \oplus (A-\lambda)^{e-k+1}U_e$$

gilt.

Dieses Verfahren kann man fortsetzen bis man bei V_1 angelangt ist. Man setzt nun die konstruierten Räume zusammen und erhält:

$$\begin{array}{ccccccccc} V & = & U_e & \oplus & (A-\lambda)U_e & \oplus & \dots & \oplus & (A-\lambda)^{e-1}U_e \\ & & & \oplus & U_{e-1} & \oplus & \dots & \oplus & (A-\lambda)^{e-2}U_{e-1} \\ & & & & & \oplus & \dots & \oplus & \dots \\ & & & & & & \dots & \oplus & (A-\lambda)U_2 \\ & & & & & & & \oplus & U_1 \end{array} \tag{2}$$

Fasst man die Räume in der ganz rechten Spalte zusammen, so erhält man gerade V_1. Nimmt man die vorletzte Spalte dazu, so erhält man V_2 usw..

Aufgrund des eingangs zitierten Lemmas gilt nun: Für alle $0 \leq j < k \leq e$ ist $(A-\lambda)^j C_k$ eine Basis von $(A-\lambda)^j U_k$. Man erhält so eine Basis von V die in (disjunkte) Jordansche Ketten zerfällt, und zwar in jeweils $\dim U_k$ Ketten der Länge k.

Insbesondere erhält man mindestens eine Jordankette der Länge e. Das bedeutet $f = \dim V \geq e$, und dies ist genau die Aussage des Satzes von Cayley-Hamilton!

§8 Eindeutigkeit der Jordanschen Normalform

Zunächst einmal eine wichtige Beobachtung: Eine Jordanmatrix ist dadurch vollständig beschrieben, dass wir zu jedem k die Anzahl der Jordankästchen des Rahmens k angeben. Aus unserer Konstruktion ergibt sich, dass dies gerade die Zahlen $\dim U_k$ sind. Diese Dimensionen hängen aber nur von der Folge der Zahlen $\dim V_1, \dim V_2, \ldots, \dim V$ ab. Es ist ja $\dim U_e = \dim V - \dim V_{e-1}$, $\dim U_{e-1} = \dim V_{e-1} - (\dim V_{e-2} + \dim U_e)$, und allgemein:

$$\dim U_k = \dim V_k - (\dim V_{k-1} + \sum_{j=k+1}^{e} \dim U_j).$$

Will man nur die Jordansche Normalform einer Matrix berechnen, ohne die Transformationsmatrix explizit angeben zu müssen, dann reicht es also, die Eigenwerte und die Dimensionen der verallgemeinerten Eigenräume zu kennen!

Es bleibt, die Eindeutigkeit der Jordanform zu zeigen, d.h. wir wollen zeigen, dass zwei ähnliche Matrizen in Jordannormalform schon identisch sind (bis auf Reihenfolge der Jordanblöcke). Wir nehmen also an, wir haben zu einer gegebenen linearen Abbildung $A : V \to V$ Basen C und D so gefunden, dass die zugehörige Darstellungsmatrizen J, L in JNF sind.
An der Matrixdarstellung lassen sich sofort die verschiedenen Eigenwerte $\{\lambda_1, \ldots, \lambda_r\}$ ablesen sowie das charakteristische Polynom. Dieses muss bei beiden Matrizen gleich sein (Ähnlichkeit!), also ist

$$J = \begin{pmatrix} \tilde{A}_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{A}_r \end{pmatrix} \quad \text{und} \quad L = \begin{pmatrix} \tilde{B}_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{B}_r \end{pmatrix}$$

wobei $\tilde{A}_i$ und $\tilde{B}_i$ jeweils aus Jordankästchen zum Eigenwert λ_i bestehen. Ferner müssen $\tilde{A}_i$ und $\tilde{B}_i$ auch jeweils denselben Rahmen haben (das ist ja gerade die algebraische Vielfachheit von λ_i im charakteristischen Polynom).
Für die Basen C und D gilt also $C = C_1 \cup \cdots \cup C_r$ und $D = D_1 \cup \cdots \cup D_r$, wobei $A|_{<C_i>}$ durch $\tilde{A}_i$ dargestellt wird, also nur den Eigenwert λ_i besitzt. Dies gilt analog für $A|_{<D_i>}$ und $\tilde{B}_i$. Der Eindeutigkeitssatz für die Primärzerlegung zeigt nun $U_i =< C_i >=< D_i >$. Insbesondere sind die Matrizen $\tilde{A}_i$ und $\tilde{B}_i$ zueinander ähnlich. Damit reduziert sich das Problem auf den Fall, dass J und L nur einen einzigen Eigenwert λ besitzen.

Gehen wir nun davon aus, dass es nur einen Eigenwert gibt. Dass J in JNF ist, bedeutet, dass die Basis C in disjunkte Jordanketten $K_1, \ldots, K_l$ zerfällt. Sei n_k die Anzahl der Jordanketten der Länge k. Durch Umordnung wird man auf eine Zerlegung des Vektorraums der Form (2) oben geführt. Daraus ersieht man, dass die Folge $(n_k)_k$ nur von der Folge der Dimensionen der verallgemeinerten Eigenräume $\dim\ker(A - \lambda)^j$ abhängt. Diese müssen bei beiden Matrizen J und L gleich sein!

DAS WARS