10 Normalformen

Wir wenden uns in diesem Kapitel noch einmal der Untersuchung linearer Abbildungen auf endlichdimensionalen Vektorräumen und deren Darstellung mittels Matrizen zu. Speziell geht es um die Frage, welche Normalformen solche Darstellungsmatrizen unter einer zugelassenen Klasse von Transformationen haben.

Primärdekomposition

In diesem Abschnitt betrachten wir häufig eine lineare Abbildung $f : V \to V$ auf einem Vektorraum V über dem Körper K und ein Polynom $p(x) = p_0 + p_1 x + \cdots + p_n x^n \in K[x]$ mit Koeffizienten in K. Der Ausdruck $p(f)$ bezeichnet die Einsetzung von f in p, also die lineare Funktion

$$p(f) = p_0 \cdot \mathrm{id}_V + p_1 \cdot f + \cdots p_n \cdot \underbrace{f \circ \cdots \circ f}_{n \text{ mal}}$$

bzw.

$$\begin{array}{cccc} p(f): & V & \longrightarrow & V \\ & x & \mapsto & p_0 x + p_1 f(x) + \cdots + p_n f^n(x) \end{array} .$$

Oft schreiben wir die Hintereinanderausführung (Komposition) von Funktionen ohne das Operationssymbol "$\circ$", also statt $f \circ p(f)$ schreiben wir etwa einfach $fp(f)$.

Definition 10.1: *Ist $f : V \to V$ eine lineare Abbildung auf dem Vektorraum V, so heisst der Teilraum W von V* **f-invariant** *oder* **f-stabil**, *wenn gilt $f(W) \subseteq W$.* □

Beispiel 10.2: (a) Für jede lineare Abbildung $f : V \to V$ sind die Teilräume $\mathrm{im}(f)$ und $\mathrm{kern}(f)$ f-invariant.

(b) Sei $D : \mathrm{Pol}(\mathbb{R}) \to \mathrm{Pol}(\mathbb{R})$ die Differentiationsabbildung auf dem Raum der Polynome mit reellen Koeffizienten. Für jedes n ist der Teilraum $\mathrm{Pol}_n(\mathbb{R})$, also der Raum der Polynome vom Grad kleiner oder gleich n, D-invariant.

(c) Der eindimensionale Teilraum, welcher von einem nicht-trivialen Vektor $x \neq 0$ aufgespannt wird, ist f invariant genau dann, wenn x ein Eigenvektor von f ist.

(d) Ist V ein Vektorraum über dem Körper K, $f : V \to V$ linear, und $p \in \mathrm{Pol}(K)$ ein Polynom mit Koeffizienten in K, dann ist der Teilraum $\mathrm{kern}(p(f))$ ein f-invarianter Teilraum von V. Dabei ist für ein Polynom $p(x) = p_0 + p_1 x + \cdots p_n x^n$ die Anwendung auf f erklärt als die lineare Abbildung

$$p(f) = p_0 \cdot \mathrm{id}_V + p_1 \cdot f + \cdots p_n \cdot f^n .$$

Ist etwa V ein reeller Vektorraum und $p(x) = 2 + x^2$, dann ist $p(f) = 2 \cdot id_V + f \circ f$. $y \in \mathrm{kern}(p(f))$ bedeutet also $2y + f(f(y)) = 0$. Ist also $x \in \mathrm{kern}(p(f))$, dann gilt wegen der Linearität von f

$$0 = f(\underbrace{2x + f(f(x))}_{=0}) = 2 \cdot f(x) + f(f(f(x))) = p(f)(f(x)) ,$$

also $f(x) \in \mathrm{kern}(p(f))$. □

Satz 10.3: *Sei V ein Vektorraum über dem Körper K, sei $f : V \to V$ linear, und sei $p(x) \in K[x]$ ein Polynom mit Koeffizienten in K. Dann sind die Unterräume* $\mathrm{im}(p(f))$ *und* $\mathrm{kern}(p(f))$ *beide f-invariant.*

Ist nun etwa V ein n-dimensionaler Vektorraum über K, $f : V \to V$ linear, und W ein f-invarianter Teilraum von V der Dimension r, dann induziert f eine lineare Abbildung

$$f_W : \begin{array}{ccc} W & \longrightarrow & W \\ w & \mapsto & f(w) \end{array} .$$

Sei $B_W = \{w_1, \ldots, w_r\}$ eine Basis von W und sei

$$B = \{\underbrace{w_1, \ldots, w_r}_{\in W}, \underbrace{v_1, \ldots, v_{n-r}}_{\in V}\}$$

eine Basis von V. Die Darstellungsmatrix von f bzgl. der Basis B hat also die Form einer sogenannten Blockmatrix

$$\mathcal{A}(f, B, B) \;=\; \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} ,$$

wobei

X die $r \times r$ Matrix $\mathcal{A}(f_W, B_W, B_W)$ ist,
$Y \in \mathrm{Mat}_{r\times(n-r)}(K)$,
$Z \in \mathrm{Mat}_{(n-r)\times(n-r)}(K)$.

Nehmen wir nun weiter an, dass $V = V_1 \oplus V_2$ ist, wobei sowohl V_1 als auch V_2 f-invariante Teilräume von V sind. Ist nun B_1 eine Basis für V_1 und B_2 eine Basis für V_2, dann ist $B = B_1 \cup B_2$ eine Basis für V. Man sieht, dass die Darstellungsmatrix von f bzgl. B von der Form

$$\mathcal{A}(f, B, B) \;=\; \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

ist, wobei

$A_1 = \mathcal{A}(f_{V_1}, B_1, B_1)$,
$A_2 = \mathcal{A}(f_{V_2}, B_2, B_2)$.

Allgemein, für $V = \bigoplus_{i=1}^k V_i$, wobei V_i jeweils f-invariant sind und B_i eine Basis von V_i ist, hat die Darstellungsmatrix von f bzgl. der Basis $B = \bigcup_{i=1}^k B_i$ die **Blockdiagonalform**

$$\mathcal{A}(f, B, B) \;=\; \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_k \end{pmatrix} ,$$

wobei

$$A_i = \mathcal{A}(f_{V_i}, B_i, B_i) \qquad \text{für alle } 1 \leq i \leq k.$$

Unser Ziel besteht nun darin, für den Vektorraum V eine spezielle Basis B zu finden, bzgl. welcher die Darstellungsmatrix von f eine besonders günstige Form hat.

Ist f linear auf V, so sind alle Darstellungsmatrizen $\mathcal{A}(f, B, B)$ von f für eine beliebige Basis B ähnlich. Ähnliche Matrizen haben dasselbe charakteristische Polynom (vgl. Korollar zu Satz 8.2).

Ausserdem haben ähnliche Matrizen dasselbe Minimalpolynom. Ist nämlich $B = PAP^{-1}$ und $a(A) = \sum_{i=0}^{n} a_i A^i = 0$, dann ist auch

$$a(B) = \sum_{i=0}^{n} a_i (PAP^{-1})^i = \sum_{i=0}^{n} a_i PA^i P^{-1} = P\big(\sum_{i=0}^{n} a_i A^i\big) P^{-1} = 0 \ .$$

Es ist also sinnvoll, vom **charakteristischen Polynom** bzw. **Minimalpolynom** einer linearen Abbildung f zu sprechen, ohne auf eine bestimmte Basis Bezug zu nehmen.

Satz 10.4 (Primärdekomposition): *Sei V ein vom Nullraum verschiedener endlichdimensionaler Vektorraum über K und sei $f : V \to V$ linear. Seien das charakteristische Polynom bzw. das Minimalpolynom von f (bzw. seiner Darstellungsmatrix) von der Gestalt*

$$c_f = p_1^{d_1} p_2^{d_2} \cdots p_k^{d_k} \quad \text{bzw.} \quad m_f = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k} \ ,$$

wobei $p_1, \ldots, p_k \in K[x]$ irreduzible und paarweise relativ prime Polynome sind (also jeweils keinen gemeinsamen nicht-konstanten Faktor haben).

(i) *Dann ist jeder der Teilräume $V_i = \mathrm{kern}(p_i^{e_i}(f))$ f-invariant und es gilt $V = \bigoplus_{i=1}^{k} V_i$.*

(ii) *Sei $f_i : V_i \to V_i$ jeweils die von f auf V_i induzierte Abbildung. Dann ist $p_i^{d_i}$ das charakteristische Polynom von f_i und $p_i^{e_i}$ das Minimalpolynom von f_i.*

Korollar 1: *Für alle $1 \leq i \leq k$ gilt:* $\dim(V_i) = d_i \cdot \mathrm{grad}(p_i)$.

Korollar 2: *Liegen alle Eigenwerte von f im Körper K, ist also*

$$c_f = (x - \lambda_1)^{d_1} \cdots (x - \lambda_k)^{d_k} \quad \text{und} \quad m_f = (x - \lambda_1)^{e_1} \cdots (x - \lambda_k)^{e_k}$$

für $\lambda_1, \ldots, \lambda_k \in K$, dann ist $V_i = \mathrm{kern}(f - \lambda_i \mathrm{id})^{e_i}$ ein f-invarianter Teilraum der Dimension d_i und es gilt $V = \bigoplus_{i=1}^{k} V_i$.

Beispiel 10.5: Wir betrachten die Abbildung

$$\begin{array}{cccc} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \mapsto & (-z, x+z, y+z) \ . \end{array}$$

Bzgl. der kanonischen Basis ist die Darstellungsmatrix von f von der Gestalt

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \ .$$

Daraus sehen wir durch einfache Rechnung, dass

$$c_f = m_f = (x+1)(x-1)^2 \ .$$

Wegen Korollar 2 ist also $\mathbb{R}^3$ darstellbar als die direkte Summe

$$\mathbb{R}^3 = \mathrm{kern}(f + \mathrm{id}) \oplus \mathrm{kern}(f - \mathrm{id})^2 \ ,$$

wobei der Summand $\mathrm{kern}(f + \mathrm{id})$ die Dimension 1 und der Summand $\mathrm{kern}(f - \mathrm{id})^2$ die Dimension 2 hat.
Nun gilt

$$(f + \mathrm{id})(x,y,z) = (x - z, x + y + z, y + 2z) \ ,$$

$\{(1,-2,1)\}$ ist also eine Basis für $\mathrm{kern}(f+\mathrm{id})$.
Weiter gilt

$$(f-\mathrm{id})^2(x,y,z) \;=\; (x-y+z,-2x+2y-2z,x-y+z)\;,$$

$\{(0,1,1),(1,1,0)\}$ ist also eine Basis für $\mathrm{kern}(f-\mathrm{id})^2$.
Eine Basis von $\mathbb{R}^3$, bzgl. welcher die Darstellungsmatrix von f eine Blockdiagonalform besitzt, ist also

$$B \;=\; \{(1,-2,1),(0,1,1),(1,1,0\}\;.$$

Die Transformationsmatrix von B zur kanonischen Basis ist

$$P \;=\; \begin{pmatrix} 1 & 0 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}\;,$$

und die Blockdiagonalmatrix, welche f bzgl. der Basis B darstellt, ist somit

$$P^{-1}AP \;=\; \begin{pmatrix} -1 & & \\ & 2 & 1 \\ & -1 & 0 \end{pmatrix}\;. \qquad \square$$

Beispiel 10.6: Wir betrachten die Differentialgleichung

$$(D^n+a_{n-1}D^{n-1}+\cdots+a_1D+a_0)f \;=\; (\frac{d}{dt})^n f+a_{n-1}(\frac{d}{dt})^{n-1}f+\cdots+a_1\frac{d}{dt}f+a_0f \;=\; 0$$

mit Koeffizienten $a_i\in\mathbb{C}$. Der Lösungsraum aller unendlich oft differenzierbaren Funktionen f sei mit V bezeichnet; V ist ein Vektorraum über $\mathbb{C}$. Aus der Theorie der Differentialgleichungen wissen wir, dass die Dimension von V genau n ist.
Aus der Differentialgleichung extrahieren wir nun das Polynom

$$m(x) \;=\; x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\;.$$

Über dem algebraisch abgeschlossenen Körper $\mathbb{C}$ lässt sich $m(x)$ zerlegen in Linearfaktoren, also für gewisse $\alpha_i\in\mathbb{C}$ und natürliche Zahlen e_i haben wir

$$m(x) \;=\; (x-\alpha_1)^{e_1}(x-\alpha_2)^{e_2}\cdots(x-\alpha_k)^{e_k}\;.$$

Also ist $D:V\to V$ eine lineare Abbildung mit Minimalpolynom $m(x)$. Wegen des Korollars 2 ist also der Lösungsraum V die direkte Summe der Lösungsräume V_i der Differentialgleichungen

$$(D-\alpha_i\,\mathrm{id})^{e_i}f \;=\; 0\;.$$

Die Lösungen der Differentialgleichung $(D-\alpha_i\,\mathrm{id})^r f=0$ können aber einfach bestimmt werden. Durch Induktion weist man leicht nach, dass gilt

$$(D-\alpha\,\mathrm{id})^r f \;=\; e^{\alpha t}D^r(e^{-\alpha t}f)\;.$$

Die Funktion f ist also eine Lösung genau dann, wenn $D^r(e^{-\alpha t}f)=0$. Das ist der Fall genau dann, wenn $e^{-\alpha t}f$ ein Polynom vom Grad höchstens $r-1$ ist. Daher ist

$$\{e^{\alpha t},te^{\alpha t},\ldots,t^{r-1}e^{\alpha t}\}$$

eine Basis für den Lösungsraum von $(D - \alpha_i \,\mathrm{id})^r f = 0$. $\square$

Wir betrachten nun noch einen Spezialfall des Primärdekompositionssatzes. Wir nehmen an, dass jeder irreduzible Faktor p_i von m_f linear ist, und jeder Exponent $e_i = 1$; also

$$m_f(x) \;=\; (x - \lambda_1) \cdots (x - \lambda_k)$$

für paarweise verschiedene $\lambda_i \in K$.
Wir wissen schon aus früheren Überlegungen, dass $f : V \to V$ **diagonalisierbar** ist (also eine Diagonalmatrix als Darstellungsmatrix hat für eine geeignete Basis) genau dann, wenn V eine Basis besitzt, welche aus Eigenvektoren von f besteht. Nun können wir die Diagonalisierbarkeit auch mittels der Faktorisierung des Minimalpolynoms charakterisieren.

Satz 10.7: *Sei V ein vom Nullraum verschiedener endlichdimensionaler Vektorraum, und sei $f : V \to V$ linear. Dann sind die folgenden Aussagen äquivalent:*

(i) *das Minimalpolynom m_f von f ist das Produkt verschiedener Linearfaktoren;*

(ii) *f ist diagonalisierbar.*

Beispiel 10.8: Wir betrachten die lineare Abbildung

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \mapsto & (7x - y - 2z, -x + 7y + 2z, -2x + 2y + 10z) \;. \end{array}$$

Bzgl. der kanonischen Basis von $\mathbb{R}^3$ hat f die Darstellungsmatrix

$$A \;=\; \begin{pmatrix} 7 & -1 & -2 \\ -1 & 7 & 2 \\ -2 & 2 & 10 \end{pmatrix} \;.$$

Durch einfaches Nachrechnen sieht man, dass

$$c_f = (x-6)^2(x-12) \quad \text{und} \quad m_f = (x-6)(x-12) \;.$$

Somit folgt aus Satz 10.7, dass f diagonalisierbar ist. $\square$

Beispiel 10.9: Wir betrachten eine 2×2 Matrix A bzw. die dadurch dargestellte lineare Abbildung auf $\mathbb{C}^2$:

$$A = \begin{pmatrix} 2 & -1 \\ 5 & -2 \end{pmatrix}, \qquad \begin{array}{rccc} f_A: & \mathbb{C}^2 & \longrightarrow & \mathbb{C}^2 \\ & \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & A \cdot \begin{pmatrix} x \\ y \end{pmatrix} \;. \end{array}$$

Das charakteristische Polynom von A bzw. der Abbildung f_A ist

$$c_A \;=\; c_{f_A} \;=\; x^2 + 1 \;=\; (x-i)(x+i) \;.$$

Die Eigenwerte sind also $\pm i$.

$$\text{Eigenvektor zu } i: \begin{pmatrix} 2+i \\ 5 \end{pmatrix}, \qquad \text{Eigenvektor zu } -i: \begin{pmatrix} 2-i \\ 5 \end{pmatrix}.$$

$\mathbb{C}^2$ hat also eine Basis B bestehend aus Eigenvektoren von A, und die Transformationsmatrix von B in die kanonische Basis ist

$$P = \begin{pmatrix} 2+i & 2-i \\ 5 & 5 \end{pmatrix}, \quad \text{mit} \quad P^{-1} = \frac{1}{10} \begin{pmatrix} -5i & 1+2i \\ 5i & 1-2i \end{pmatrix}.$$

Mittels dieser Transformationsmatrizen ist A diagonalisierbar:

$$P^{-1}AP \;=\; \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} .$$

Es gibt aber keine Matrix mit rationalen oder reellen Elementen, welche A diagonalisiert.
□

Satz 10.10: *Sei V ein vom Nullraum verschiedener endlichdimensionaler Vektorraum und seien $f, g : V \to V$ diagonalisierbare lineare Abbildungen. Dann sind f und g simultan diagonalisierbar (in dem Sinn, dass es eine Basis von V gibt, welche gleichzeitig aus Eigenvektoren von f und g besteht) genau dann, wenn gilt $f \circ g = g \circ f$.*

Korollar: *Seien $A, B \in \mathrm{Mat}_{n\times n}(K)$ diagonalisierbar. Die Matrizen A und B sind simultan diagonalisierbar (in dem Sinn, dass es eine invertierbare Matrix P gibt, für welche sowohl $P^{-1}AP$ als auch $P^{-1}BP$ Diagonalmatrizen sind) genau dann, wenn $AB = BA$.*

Dreiecksgestalt

Im Primärdekompositionssatz haben wir gesehen, dass die Darstellungsmatrix einer linearen Abbildung auf einem endlichdimensionalen Vektorraum bzgl. einer geeigneten Basis die Form einer Blockdiagonalmatrix hat. Zerfällt das Minimalpolynom über dem Grundkörper K in paarweise verschiedene Linearfaktoren, so ist diese Blockdiagonalmatrix eine Diagonalmatrix. Was aber passiert im Fall, dass das Minimalpolynom zwar in Linearfaktoren zerfällt, welche aber nicht notwendig verschieden sind; also im Fall

$$m_f \;=\; \prod_{i=1}^{k} (x - \lambda_i)^{e_i}, \quad e_i \geq 1 \quad ?$$

Diese Situation tritt immer ein, wenn K ein algebraisch abgeschlossenener Körper ist, etwa $\mathbb{C}$. Wir zeigen, dass in diesem Fall die Darstellungsmatriz bzgl. einer geeigneten Basis obere Dreiecksgestalt hat, also unterhalb der Diagonale nur 0 vorkommt.

Definition 10.11: *Ein quadratische Matrix $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(K)$ heisst* **triangulär** *bzw. in* **(oberer) Dreiecksgestalt**, *wenn $a_{ij} = 0$ für alle $i > j$.* □

Aufgrund von Korollar 2 zu Satz 10.4 (Primärdekompositionssatz) haben wir in unserer Situation

$$V \;=\; \bigoplus_{i=1}^{k} \mathrm{kern}(f - \lambda_i \mathrm{id})^{e_i} \;=\; \bigoplus_{i=1}^{k} V_i \;,$$

wobei die direkten Summanden V_i f-invariant sind. Wie oben betrachten wir die Restriktion von f auf die "Primärkomponente" V_i, also $f_i : V_i \to V_i$. Dann ist die e_i-te Potenz der Abbildung $f_i - \lambda_i\mathrm{id}$ die Nullabbildung auf V_i. Dafür führen wir einen neuen Begriff ein.

Definition 10.12: *Eine lineare Abbildung $f : V \to V$ heisst* **nilpotent** *gdw. es ein $m \in \mathbb{N}$ gibt, sodass $f^m = 0$.*
Ebenso heisst eine quadratische Matrix A **nilpotent** *gdw. es ein $m \in \mathbb{N}$ gibt, sodass $A^m = 0$.* □

Für jede Abbildung $f \in \mathrm{Hom}_K(V,V)$ gilt $f^0 = \mathrm{id}_V$, und für jede Matrix $A \in \mathrm{Mat}_{n\times n}(K)$ gilt $A^0 = I_n$. Ist also V nicht der Nullraum bzw. $n \neq 0$, dann ist $f^0 \neq 0$ und $A^0 \neq 0$. In allen diesen nichttrivialen Fällen bedeutet also nilpotent zu sein, dass eine Potenz f^m bzw. A^m mit $m > 0$ verschwindet.

Beispiel 10.13: (a) Die lineare Abbildung $f(x,y,z) = (0,x,y)$ auf $\mathbb{R}^3$ ist nilpotent. Tatsächlich gilt

$$f^3(x,y,z) = f^2(0,x,y) = f(0,0,x) = (0,0,0)\ ,$$

also $f^3 = 0$.
(b) Sind alle Eigenwerte der linearen Abbildung f auf $\mathbb{C}^n$ gleich 0, dann ist $c_f = x^n$. Der Satz von Cayley-Hamilton besagt dann $f^n = c_f(f) = 0$, f ist also nilpotent.
(c) Die Differentiationsabbildung D auf $\mathrm{Pol}_n(K)$ ist nilpotent. □

Satz 10.14: *Sei V ein vom Nullraum verschiedener n-dimensionaler Vektorraum ($n \in \mathbb{N}$), und sei $f \in \mathrm{Hom}(V,V)$ eine nilpotente lineare Abbildung. Dann gibt es eine geordnete Basis $B = (b_1,\ldots,b_n)$ von V, sodass*

$$\begin{array}{rcl} f(b_1) &=& 0, \\ f(b_2) &\in& \mathrm{span}(\{b_1\}), \\ &\vdots& \\ f(b_n) &\in& \mathrm{span}(\{b_1,\ldots,b_{n-1}\}). \end{array}$$

Korollar: *Ist f nilpotent, dann gibt es eine geordnete Basis von V, bzgl. welcher die Darstellungsmatrix von f triangulär ist, und in welcher ausserdem alle Diagonalelemente 0 sind.*

Satz 10.15 (Dreiecksgestalt): *Sei V ein vom Nullraum verschiedener endlichdimensionaler Vektorraum über dem Körper K. Sei $f \in \mathrm{Hom}_K(V,V)$ eine lineare Abbildung auf V, deren charakteristisches Polynom und Minimalpolynom die Darstellung*

$$c_f = \prod_{i=1}^{k}(x-\lambda_i)^{d_i}, \qquad m_f = \prod_{i=1}^{k}(x-\lambda_i)^{e_i}$$

haben, für paarweise verschiedene $\lambda_1,\ldots,\lambda_k \in K$ und $e_i \leq d_i$.
Dann gibt es eine geordnete Basis von V, bzgl. welcher die Darstellungsmatrix M von f triangulär ist. Noch genauer gesagt ist die Darstellungsmatrix M eine Blockdiagonalmatrix der Form

$$M = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix},$$

wobei jede Teilmatrix A_i eine $(d_i \times d_i)$ obere Dreiecksmatrix ist, in welcher alle Diagonalelemente λ_i sind.

Korollar: *Jede quadratische Matrix über $\mathbb{C}$ ist ähnlich zu einer Dreiecksmatrix.*

Beispiel 10.16: Wie in Beispiel 10.5 betrachten wir $f \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^3, \mathbb{R}^3)$

$$f: \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z) & \mapsto & (-z, x+z, y+z) \ . \end{array}$$

Bzgl. der kanonischen Basis ist die Darstellungsmatrix von f von der Gestalt

$$A \ = \ \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \ .$$

In Beispiel 10.5 haben wir schon gesehen

$$\mathbb{R}^3 \ = \ \mathrm{kern}(f + \mathrm{id}) \oplus \mathrm{kern}(f - \mathrm{id})^2 \ ,$$

wobei der Summand $\mathrm{kern}(f + \mathrm{id})$ die Dimension 1 und der Summand $\mathrm{kern}(f - \mathrm{id})^2$ die Dimension 2 hat.
$V_1 = \mathrm{kern}(f + \mathrm{id}) = \mathrm{span}((1,-2,1))$.
Wir suchen nun eine Basis von $V_2 = \mathrm{kern}(f - \mathrm{id})^2$ im Sinne von Satz 10.14.

$$b_1 = (-1,0,1) \in \mathrm{kern}(f - \mathrm{id}) \setminus \{0\}.$$

Als nächstes suchen wir nun einen Vektor b_2 mit der Eigenschaft

$$(f - \mathrm{id})(b_2) \in \mathrm{span}(b_1), \quad \text{also} \quad (f - \mathrm{id})(b_2) = \alpha b_1 \quad \text{für ein } \alpha \in \mathbb{R}.$$

Der Vektor $b_2 = (0,1,1)$ erfüllt diese Bedingung.
Die Darstellungsmatrix von $f - \mathrm{id}$ bzgl. der geordneten Basis (b_1, b_2) von V_2 ist

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Die Darstellungsmatrix der Abbildung f_2, welche von f induziert wird auf V_2 ist also

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Somit ist also die Darstellungsmatrix von f bzgl. der geordneten Basis

$$B = (\,(1,-2,1), (-1,0,1), (0,1,1)\,)$$

die Dreieckmatrix

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \qquad \square$$

Satz 10.17 (Jordan-Dekomposition): *Sei V ein vom Nullraum verschiedener endlichdimensionaler Vektorraum über dem Körper K, und sei $f \in \mathrm{Hom}_K(V,V)$ eine lineare Abbildung auf V, welche alle ihre Eigenwerte in K hat (also c_f zerfällt über K vollständig in Linearfaktoren).*
Dann gibt es lineare Abbildungen $g, h \in \mathrm{Hom}_K(V,V)$, sodass

$f = g + h$,

g *ist diagonalisierbar,*

h *ist nilpotent, und*

$g \circ h = h \circ g$.

Alternativ könnte man den Beweis zu Satz 10.17 auch über die Darstellungsmatrix von f führen, die ja laut Satz 10.15 Dreiecksgestalt besitzt. Vergleiche dazu Beispiel 10.18.

Man kann auch zeigen, dass die Abbildungen g und h in Satz 10.17 eindeutig bestimmt sind.

Beispiel 10.18: In Beispiel 10.16 haben wir gesehen, dass bzgl. der Basis

$$B = ((1,-2,1),(-1,0,1),(0,1,1))$$

die lineare Abbildung

$$f: \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (x,y,z) & \mapsto & (-z, x+z, y+z) \end{array}.$$

die Darstellungsmatrix

$$T = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

besitzt. Offensichtlich können wir T darstellen als $R + S$ für

$$R = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad S = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Das ist die Jordan-Dekomposition von T.
Mittels der Transformationsmatrix

$$P = \begin{pmatrix} 1 & -1 & 0 \\ -2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

führen wir die Betrachtung zurück auf die kanonische Basis. Wir erhalten die Darstellungsmatrizen

$$PRP^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 2 \\ -1 & 1 & 1 \end{pmatrix} \quad \text{und} \quad PSP^{-1} = \frac{1}{2}\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Der Diagonalanteil von f ist also

$$g(x,y,z) = \frac{1}{2}(x + y - z, 2x + 2z, -x + y + z),$$

und der nilpotente Anteil von f ist

$$h(x, y, z) = \frac{1}{2}(-x - y - z, 0, x + y + z). \qquad \square$$

Jordan-Normalform

Um die Darstellung als Dreiecksmatrix noch zu verbessern, müssen wir uns fragen, ob wir noch bessere Basen finden können für die direkten Summanden im Primärdekompositionssatz.

Definition 10.19: *Für eine nilpotente lineare Abbildung $f \in \mathrm{Hom}_K(V, V)$ ist der* **Index** *die kleinste positive ganze Zahl k mit $f^k = 0$.* $\square$

Lemma 10.20: *Sei $f \in \mathrm{Hom}_K(V, V)$ nilpotent mit Index p. Dann ist*

(i) $\{0\} = \mathrm{kern} f^0 \subset \mathrm{kern} f \subset \cdots \subset \mathrm{kern} f^{p-1} \subset \mathrm{kern} f^p = V$ *eine Kette echt ineinander enthaltener Teilräume von V;*

(ii) $\{x, f(x), \ldots, f^{p-1}(x)\}$ *eine linear unabhängige Teilmenge von V, für jedes $x \in V \setminus \mathrm{kern} f^{p-1}$.*

Satz 10.21: *Sei V ein endlichdimensionaler Vektorraum über K und sei $f \in \mathrm{Hom}_K(V, V)$ so, dass $\mathrm{kern} f^i \subset \mathrm{kern} f^{i+1}$ für ein $i \geq 1$.*
Ist $\{v_1, \ldots, v_s\}$ eine Basis von $\mathrm{kern} f^i$ und ist $\{v_1, \ldots, v_s, w_1, \ldots, w_t\}$ eine Basis von $\mathrm{kern} f^{i+1}$, dann ist $S = \{f(w_1), \ldots, f(w_t)\}$ eine linear unabhängige Teilmenge von $\mathrm{kern} f^i$.

Definition 10.22: *Sei $\lambda \in K$, und sei n eine positive natürliche Zahl. Eine* **elementare Jordan-Matrix** *zum Eigenwert λ ist eine (quadratische) $n \times n$ Matrix $J = (j_{kl})$, in welcher $j_{kk} = \lambda$ für alle $1 \leq k \leq n$, $j_{k,k+1} = 1$ für alle $1 \leq k < n$, und alle anderen Elemente 0 sind. Also*

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}.$$

Eine **Jordan-Blockmatrix** *zum Eigenwert λ ist eine Blockmatrix der Form*

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix},$$

wobei $k \geq 1$ und jeder Block eine elementare Jordan-Matrix zum Eigenwert λ ist. Alle anderen Elemente sind 0. $\square$

Satz 10.23: *Sei V ein vom Nullraum verschiedener endlichdimensionaler Vektorraum über dem Körper K.*

(i) *Ist $f \in \mathrm{Hom}_K(V,V)$ nilpotent, dann gibt es eine geordnete Basis von V bzgl. welcher die Darstellungsmatrix von f eine Jordan-Blockmatrix zum Eigenwert 0 ist;*

(ii) *ist $f \in \mathrm{Hom}_K(V,V)$ so, dass alle Eigenwerte von f in K sind (also $c_f(x) = (x-\lambda_1)^{d_1} \cdots (x-\lambda_k)^{d_k}$ für $\lambda_1, \ldots, \lambda_k \in K$), dann gibt es eine geordnete Basis von V bzgl. welcher die Darstellungsmatrix von f eine Blockdiagonalmatrix*

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$

ist, in welcher jedes J_i eine Jordan-Blockmatrix zum Eigenwert λ_i ist.

Definition 10.24: *Sei $f \in \mathrm{Hom}_K(V,V)$ so, dass alle Eigenwerte von f in K liegen. Dann heisst eine Matrix von der Form wie in Satz 1.23(ii) eine* **Jordan-Matrix** *von f.*
Sei $A \in \mathrm{Mat}_{n\times n}(K)$ so, dass alle Eigenwerte von A in K liegen. Dann ist die **Jordan-Normalform** *von A die (eine) Jordan-Matrix einer (jeder) linearen Abbildung, welche A als Darstellungsmatrix (bzgl. einer Basis) hat.* □

Bemerkung 10.25: Eine Jordan-Matrix zu einer Abbildung f ist (wenn überhaupt) nicht eindeutig definiert. Insbesondere ist die Ordnung der Jordanblöcke nicht festgelegt. Wohl aber ist die Anzahl der Blöcke, ihre Grösse, und die Anzahl der der elementaren Jordanmatrizen in diesen Blöcken festgelegt. Um einer Eindeutigkeit näher zu kommen, könnte man etwa die Blöcke absteigend nach Grösse anordnen. Über dem Grundkörper $\mathbb{R}$ könnte man auch noch die Eigenwerte der orden, und somit Eindeutigkeit herstellen.

Zusammenfassend lässt sich folgendes feststellen:

Satz 10.26: *Hat die lineare Abbildung $f : V \to V$ bzw. eine Darstellungsmatrix A von f (bzgl. einer Basis) das charakteristische Polynom*

$$c(x) = (x-\lambda_1)^{d_1} \cdots (x-\lambda_k)^{d_k}$$

und das Minimalpolynom

$$m(x) = (x-\lambda_1)^{e_1} \cdots (x-\lambda_k)^{e_k},$$

dann hat die Jordan-Matrix von f bzw. die Jordan-Normalform von A folgende Gestalt:

- *der Eigenwert λ_i kommt d_i mal in der Diagonale vor (also so oft wie die algebraische Vielfachheit von λ_i),*
- *die Anzahl der elementaren Jordan-Matrizen zu λ_i ist $\dim \mathrm{kern}(f_i - \lambda_i \mathrm{id})$ (also die geometrische Vielfachheit von λ_i), und*
- *zumindest eine dieser elementaren Jordan-Matrizen hat die Grösse $e_i \times e_i$.*

Beispiel 10.27: Wir betrachten die nilpotente Abbildung

$$\begin{array}{rccc} f: & \mathbb{R}^4 & \longrightarrow & \mathbb{R}^4 \\ & (x,y,z,w) & \mapsto & (0,x,w,0) \end{array}$$

Der Index von f ist 2. Mit der Notation aus dem Beweis von Satz 10.23 haben wir also

$$\begin{aligned} W_1 &= \mathrm{kern} f = \{(0,y,z,0) | y,z \in \mathbb{R}\}, \\ W_2 &= \mathrm{kern} f^2 = \mathbb{R}^4. \end{aligned}$$

Als Basen für diese Teilräume haben wir etwa

$$\begin{aligned} B_1 = & \quad \{(0,1,0,0),(0,0,1,0)\}, \\ B_2 = B_1 \cup C_2 = & \quad \{(0,1,0,0),(0,0,1,0)\} \cup \{(1,0,0,0),(0,0,0,1)\}. \end{aligned}$$

Das Bild von C_2 unter f ist B_1, und somit erhalten wir (wie im Beweis von Satz 10.23 die Anordnung

$$\begin{matrix} (1,0,0,0), & (0,0,0,1), \\ (0,1,0,0), & (0,0,1,0). \end{matrix}$$

Zählen wir also nun diese Elemente spaltenweise von unten her auf, so erhalten wir die geordnete Basis

$$B = (\,(0,1,0,0),(1,0,0,0),(0,0,1,0),(0,0,0,1)\,).$$

Die Darstellungsmatrix von f bzgl. B ist

$$\begin{pmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & 0 & 1 \\ & & 0 & 0 \end{pmatrix}.$$

Das ist die Darstellung von f mittels einer Jordan Blockmatrix. □

Beispiel 10.28: Die Matrix

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 0 & 7 & -4 \\ 0 & 9 & -5 \end{pmatrix}$$

hat das charakteristische Polynom bzw. das Minimalpolynom

$$c_A(x) = (x-1)^3 \qquad \text{bzw.} \qquad m_A(x) = (x-1)^2.$$

Aufgrund von Satz 10.26 ist somit

$$\begin{pmatrix} 1 & 1 & \\ & 1 & \\ & & 1 \end{pmatrix}$$

die Jordan-Normalform von A. □

Beispiel 10.29: Die Matrix

$$A = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & -1 & -1 & -1 & 0 \end{pmatrix}$$

hat das charakteristische Polynom

$$c_A(x) = (x-1)^3(x-2)^2.$$

Der Eigenraum zum Eigenwert 1 ist

$$E_1 \;=\; \{(0,0,x,-x,0)^T \mid x \in \mathbb{R}\},$$

also die geomtrische Vielfachheit des Eigenwertes 1 ist $e_1 = 1$.
Der Eigenraum zum Eigenwert 2 ist

$$E_2 \;=\; \{(x,y,-y,0,0)^T \mid x,y \in \mathbb{R}\},$$

also die geomtrische Vielfachheit des Eigenwertes 2 ist $e_2 = 2$.
Somit kommt in der Jordan-Normalform von A der Eigenwert 1 dreimal in der Diagonale vor mit nur einer zugehörigen elementaren Jordan-Matrix; der Eigenwert 2 kommt zweimal in der Diagonale vor mit 2 zugehörigen elementaren Jordan-Matrizen.
Also ist

$$\begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & 1 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}$$

die Jordon-Normalform von A. □

Beispiel 10.30: (i) Wir betrachten die lineare Abbildung $f : \mathbb{R}^7 \to \mathbb{R}^7$ mit

$$c_f(x) = (x-1)^3(x-2)^4 \qquad \text{und} \qquad m_f(x) = (x-1)^2(x-2)^3.$$

Für jede Jordan-Matrix von f gilt also:

- der Eigenwert 1 kommt dreimal in der Diagonale vor, und zumindest einer der zugehörigen Jordan-Blöcke hat die Grösse 2×2,
- der Eigenwert 2 kommt viermal in der Diagonale vor, und zumindest einer der zugehörigen Jordan-Blöcke hat die Grösse 3×3.

Bis auf mögliche Umordnungen der Jordan-Blöcke ist also

$$\begin{pmatrix} 1 & 1 & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 2 & 1 & & \\ & & & & 2 & 1 & \\ & & & & & 2 & \\ & & & & & & 2 \end{pmatrix}$$

die Jordan-Matrix von f.

(ii) Seien f und c_f wie in (i), aber

$$m_f(x) = (x-1)^2(x-2)^2\ .$$

Dann kommt der Eigenwert 2 viermal in der Diagonale der Jordon-Normalform vor, und zumindest einer der zugehörigen Jordan-Blöcke hat die Grösse 2×2.
Abhängig von der konkreten Gestalt von f sind also

$$\begin{pmatrix} 1 & 1 & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 2 & 1 & & \\ & & & & 2 & & \\ & & & & & 2 & 1 \\ & & & & & & 2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 1 & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 2 & 1 & & \\ & & & & 2 & & \\ & & & & & 2 & \\ & & & & & & 2 \end{pmatrix}$$

die möglichen Formen der Jordan-Normalform. □

Definition 10.31: *Sei $f \in \mathrm{Hom}_K(V, V)$ so, dass alle Eigenwerte in K liegen. Eine geordnete Basis von V, bzgl. welcher die Darstellungsmatrix von f eine Jordan-Matrix ist, heisst eine zu f gehörige* **Jordan-Basis**. □

Algorithmus 10.32: (Berechnung einer Jordan-Basis)
(i) Zu einer gegebenen elementaren Jordan-Matrix der Grösse $m \times m$

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

erfüllen die Elemente einer Jordan-Basis $B = \{b_1, \ldots, b_m\}$ folgende Bedingungen:

$$\begin{array}{rcl} f(b_1) & = & \lambda b_1, \\ f(b_2) & = & \lambda b_2 + b_1, \\ f(b_3) & = & \lambda b_3 + b_2, \\ & \vdots & \\ f(b_{m-1}) & = & \lambda b_{m-1} + b_{m-2}, \\ f(b_m) & = & \lambda b_m + b_{m-1}. \end{array}$$

Die Elemente einer Jordan-Basis müssen also natürlich linear unabhängig sein und es muss gelten

(i) $b_1 \in \mathrm{kern}(f - \lambda \mathrm{id})$, und

(ii) $(f - \lambda \mathrm{id})(b_i) = b_{i-1}$ für alle $i = 2, \ldots, m$.

Solche Basiselemente können also bestimmt werden als die Lösungen linearer Gleichungssysteme.
(ii) Für eine allgemeine Jordan-Matrix wendet man obiges Verfahren an auf jede einzelne elementare Jordan-Matrix und fügt schliesslich die so gewonnenen Basen zu einer geordneten Basis des ganzen Raumes zusammen. □

Beispiel 10.33: Wir betrachten die lineare Abbildung

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x, y, z) & \mapsto & (x + y, -x + 3y, -x + y + 2z) \ . \end{array}$$

Bzgl. der kanonischen Basis ist

$$A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 3 & 0 \\ -1 & 1 & 2 \end{pmatrix}$$

die Darstellungsmatrix von f. Das charakteristische Polynom und das Minimalpolynom sind

$$c_f(x) = (x-2)^3 \qquad \text{und} \qquad m_f(x) = (x-2)^2.$$

Also ist die Jordan-Normalform

$$J = \begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 2 \end{pmatrix} .$$

Die relevante Abbildung ist also

$$(f - 2\mathrm{id})(x,y,z) = (-x+y, -x+y, -x+y) .$$

Wir wählen das erste Element der Jordan-Basis so, dass

$$b_1 \in \mathrm{kern}(f - 2\mathrm{id}),$$

also etwa $b_1 = (1,1,1)$.
Wir wählen b_2 so, dass es unabhängig ist von b_1 und

$$(f - 2\mathrm{id})(b_2) = b_1.$$

Für $b_2 = (u,v,w)$ muss also gelten $-u+v = 1$. Der Vektor $b_2 = (1,2,1)$ ist eine mögliche Wahl.
Schliesslich wählen wir b_3 so, dass es unabhängig ist von b_1 und b_2 und $b_3 \in \mathrm{kern}(f - 2\mathrm{id})$.
Offensichtlich erfüllt $b_3 = (1,1,0)$ diese Bedingung.
Somit ist

$$B = (\,(1,1,1),(1,2,1),(1,1,0)\,)$$

eine Jordan-Basis zu f.
Die Basistransformationsmatrix von B zur kanonischen Basis ist

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} .$$

Damit lässt sich die Matrix A in die Jordan-Matrix J ähnlichkeitstransformieren:

$$P^{-1}AP = J . \qquad \square$$

Beispiel 10.34: Bzgl. der kanonischen Basis $\{1, x, x^2, x^3, x^4\}$ von $\mathbb{R}_4[x]$ (Polynome mit reellen Koeffizienten vom Grad ≤ 4) hat die Differentiationsabbildung D die Darstellungsmatrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

Das charakteristische Polynom ist x^5, der einzige Eigenwert ist 0, und der Eigenraum zum Eigenwert 0 wird aufgespannt von $\{1\}$. Daher hat die Jordan-Matrix von D die Gestalt

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

Für eine Jordan-Basis $(p_1, \ldots, p_5)$ muss also gelten

$$Dp_1 = 0 \quad \text{und} \quad Dp_i = p_{i-1} \text{ für } \ i = 2, \ldots, 5.$$

Jordan-Basen zu D sind also

$$(1, x, \frac{1}{2}x^2, \frac{1}{6}x^3, \frac{1}{24}x^4) \quad \text{oder} \quad (24, 24x, 12x^2, 4x^3, x^4).$$ □

Satz 10.35: *Jede quadratische Matrix A über $\mathbb{C}$ ist ähnlich zu ihrer Transponierten A^T.*

Beispiel 10.36: Wir demonstrieren die Nützlichkeit der Jordan-Normalform etwa bei der Lösung von Systemen linearer Differentialgleichungen.
In Beispiel 8.22 haben wir bereits ein **System linearer gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten** betrachtet. Dabei handelt es sich um ein System von Gleichungen der Form

$$\begin{array}{ccccccccc} x_1' & = & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n \\ x_2' & = & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n \\ \vdots & & & & & & & & \\ x_n' & = & a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n \end{array},$$

wobei $a_{ij} \in K$, $K = \mathbb{R}$ oder $\mathbb{C}$, und die gesuchten Lösungen $x_1, \ldots, x_n$ differenzierbare Funktionen in t über dem Körper K sind, also $x_i \in C_K(t)$.
In Matrixnotation können wir dieses Gleichungssystem schreiben als

$$(1) \quad X' = AX,$$

wobei $X = (x_1, \ldots, x_n)^T$ und $A = (a_{ij})$.
Wir nehmen an, dass A alle Eigenwerte in K hat, und ähnlich ist zur Jordan-Normalform J_A. Sei P die zugehörige Transformationsmatrix, also

$$P^{-1}AP = J_A.$$

Setzen wir $Y := P^{-1}X$, so erhalten wir

$$(2) \quad PY' = (PY)' = X' = AX = APY,$$

und somit

$$(3) \quad Y' = P^{-1}X' = P^{-1}APY = J_AY.$$

Die Form von J_A bedeutet im allgemeinen, dass (3) deutlich einfacher zu lösen ist als (1). Die Lösungen Y von (3) können überdies mittels $X = PY$ in Lösungen von (1) transformiert werden.

Wir führen diesen Ansatz für das System

$$\begin{array}{rcrcrcr} x_1' & = & x_1 & + & x_2 & & \\ x_2' & = & -x_1 & + & 3x_2 & & \\ x_3' & = & -x_1 & + & 4x_2 & + & -x_3 \end{array}$$

durch. Also für

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 3 & 0 \\ -1 & 4 & -1 \end{pmatrix}$$

schreiben wir das Gleichungssystem als

$$X' = AX.$$

Durch einfache Rechnung sehen wir

$$c_A(x) = (x+1)(x-2)^2 = m_A(x).$$

Die Jordon-Normalform von A ist also

$$J_A = \begin{pmatrix} -1 & & \\ & 2 & 1 \\ & & 2 \end{pmatrix}.$$

Wir bestimmen nun eine Jordan-Basis (p_1, p_2, p_3) zur von A definierten linearen Abbildung; die Elemente dieser Basis werden später die Spalten der Transformationsmatrix P sein. Die Vektoren p_1, p_2, p_3 müssen linear unabhängig sein und die Bedingungen

$$\begin{array}{rcl} (A + I_3)p_1 & = & 0, \\ (A - 2I_3)p_2 & = & 0, \\ (A - 2I_3)p_3 & = & p_2 \end{array}$$

erfüllen. Geeignete Vektoren sind etwa

$$p_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad p_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}.$$

Also ist

$$P \;=\; \begin{pmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

die Transformationsmatrix von A in die Jordan-Normalform: $P^{-1}AP = J_A$.
Anstatt des ursprünglichen Systems betrachten wir nun das System

$$Y' = J_A Y,$$

also

$$\begin{array}{rcrcrcr} y_1' & = & -y_1, & & & & \\ y_2' & = & & & 2y_2 & + & y_3, \\ y_3' & = & & & & & 2y_3. \end{array}$$

Die erste und die dritte Gleichung ergeben offensichtlich

$$y_1(t) = \alpha_1 e^{-t} \quad \text{und} \quad y_3(t) = \alpha_3 e^{2t}.$$

Die zweite Gleichung lautet somit

$$y_2' = 2y_2 + \alpha_3 e^{2t},$$

und besitzt die Lösung

$$y_2(t) = \alpha_2 e^{2t} + \alpha_3 t e^{2t}.$$

Für den Lösungsvektor Y ergibt sich also

$$Y = \begin{pmatrix} \alpha_1 e^{-t} \\ \alpha_2 e^{2t} + \alpha_3 t e^{2t} \\ \alpha_3 e^{2t} \end{pmatrix}.$$

Schliesslich erhalten wir daraus den allgemeinen Lösungsvektor

$$X = PY = \begin{pmatrix} \alpha_2 e^{2t} + \alpha_3 (t-1) e^{2t} \\ \alpha_2 e^{2t} + \alpha_3 t e^{2t} \\ \alpha_1 e^{-t} + \alpha_2 e^{2t} + \alpha_3 t e^{2t} \end{pmatrix}$$

für das ursprüngliche Gleichungssystem. □