f) Hauptvektoren und und die Jordan-Zerlegung

Falls die lineare Abbildung $\varphi: V \to V$ Eigenwerte hat, deren geometrische Vielfachheit kleiner als die algebraische ist, haben wir keine Chance auf eine Basis, bezüglich derer die Abbildungsmatrix von φ Diagonalgestalt hat: Die Elemente einer solchen Basis wären allesamt Eigenvektoren, und bei zu kleiner geometrischer Vielfachheit gibt es nicht genügend linear unabhängige Eigenvektoren. Außerdem gibt es offensichtlich keine Chance auf eine Diagonalgestalt, wenn das charakteristische Polynom von φ nicht in Linearfaktoren zerfällt, denn dann ist schon die Summe der *algebraischen* Vielfachheiten der Eigenwerte kleiner als die Dimension von V.

Das zweite dieser Probleme konnten wir zumindest beim Beispiel der Matrix

$$\begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}$$

dadurch lösen, daß wir zu einem größeren Körper übergegangen sind, nämlich von den reellen zu den komplexen Zahlen.

Tatsächlich läßt es sich *immer* dadurch lösen, daß man zu einem größeren Körper übergeht: Nach dem zuerst von GAUSS bewiesenen *Fundamentalsatz der Algebra* zerfällt jedes Polynom mit komplexen (also insbeondere auch mit reellen) Koeffizienten über dem Körper der komplexen Zahlen in Linearfaktoren. Trotz seines Namens kann dieser Satz nicht rein algebraisch bewiesen werden; ein Beweis, der möglichst weitgehend mit Sätzen aus dieser Vorlesung auskommt, könnte beispielsweise die Integralsätze aus Kaptitel II in Sätze der komplexen Analysis übersetzen und den Satz dann relativ schnell daraus herleiten. Da der Gesamtaufwand allerdings bei ein bis zwei Vorlesungen läge, sei auf Einzelheiten verzichtet.

Für andere Körper als die reellen oder komplexen Zahlen zeigt die Algebra, daß es zu jedem Polynom über einem Körper stets einen Erweiterungskörper gibt, der als Vektorraum über dem Ausgangskörper endliche Dimension hat, so daß das gegebene Polynom dort in Linearfaktoren zerfällt. Mit Methoden, die im allgemeinen nicht konstruktiv sind, folgt sogar, daß es stets einen (im allgemeinen unendlichdimensionalen)

Erweiterungskörper gibt, über dem *jedes* Polynom in Linearfaktoren zerfällt, den sogenannte *algebraischen Abschluß* des Ausgangskörpers. Einzelheiten findet man in jedem Lehrbuch der Algebra.

Somit können wir das Problem, daß das charakteristische Polynom eventuell nicht genügend viele Nullstellen hat, im wesentlichen ignorieren. Ernster ist das Problem mit Eigenwerten, deren geometrische Vielfachheit kleiner ist als die algebraische. Damit wollen wir uns in diesem Abschnitt beschäftigen.

Die Lösung wird darin bestehen, daß wir solchen Eigenwerten Räume zuordnen, die größer sind als die Eigenräume, aber immer noch eine gut an die Abbildung angepaßte Basis haben. Insbesondere sollen sie, genau wie die Eigenräume, *invariant* sein unter der betrachteten Abbildung:

Definition: $\varphi: V \to V$ sei eine lineare Abbildung. Ein Untervektorraum $U \le V$ heißt invariant unter φ oder kurz φ-invariant, wenn $\varphi(U) \le U$ ist.

Die φ-Invarianz der Eigenräume im Sinne dieser Definition ist klar, denn auf einem Eigenraum ist φ einfach die Multiplikation mit dem zugehörigen Eigenwert.

Für das folgende wollen wir der Einfachheit halber annehmen, daß V endliche Dimension habe. Dann ist erst recht jeder φ-invariante Unterraum U endlichdimensional, wir können also eine endliche Basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ von U finden und diese ergänzen zu einer Basis $\{\vec{b}_1, \dots, \vec{b}_n\}$ von V. Da $\varphi(U) \le U$ ist, liegen die Bilder der ersten r Basisvektoren wieder in U, d.h. die Abbildungsmatrix bezüglich dieser Basis hat die Form

$$\begin{pmatrix} \boxed{A} & \boxed{C} \\ \mathbf{0} & \boxed{B} \end{pmatrix}$$

mit einer $r \times r$-Matrix A, der Abbildungsmatrix von $\varphi_{|U}: U \to U$, einer $(n-r)\times(n-r)$-Matrix B und einer $(n-r)\times r$-Matrix C. Die fette Null soll hier, wie auch in den noch folgenden Matrizen, stets eine Nullmatrix der jeweils korrekten Größe bezeichnen.

Noch besser wird die Situation, wenn U ein φ-invariantes Komplement hat, wenn es also einen weiteren φ-invarianten Untervektorraum W gibt, so daß $V = U + W$ ist und $U \cap W = \{\vec{0}\}$. (Wir sagen dann, $V = U \oplus W$ sei die *direkte Summe* von U und W.) In diesem Fall können wir für $\vec{b}_{r+1}$ bis $\vec{b}_n$ die Vektoren einer Basis von W nehmen, und da nun auch W auf sich selbst abgebildet wird, haben wir eine Abbildungsmatrix der Form

$$\begin{pmatrix} \boxed{A} & \mathbf{0} \\ \mathbf{0} & \boxed{B} \end{pmatrix}.$$

Allgemein sagen wir für s Untervektorräume $U_1, \dots, U_s$ von V, daß V die direkte Summe

$$V = U_1 \oplus \cdots \oplus U_s = \bigoplus_{i=1}^{s} U_i$$

sei, wenn

$$V = U_1 + \cdots + U_s = \sum_{i=1}^{s} U_i \quad \text{und} \quad U_i \cap \sum_{j \neq i} U_j = \{\vec{0}\}$$

ist. Falls hierbei die U_i allesamt φ-invariant sind, können wir ihre Basen aneinandersetzen und erhalten eine Basis, bezüglich derer die Abbildungsmatrix die Gestalt

$$\begin{pmatrix} \boxed{A_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boxed{A_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boxed{A_s} \end{pmatrix}$$

hat, wobei die A_i die Abbildungsmatrizen der Einschränkungen $\varphi_{|U_i}$ zu Abbildungen von U_i nach U_i sind.

Kandidaten für Untervektorräume U_i liefern die Haupträume:

Definition: *a)* Ein Vektor $\vec{v} \in V$ heißt *Hauptvektor* von φ zum Eigenwert λ, wenn es ein $\ell \in \mathbb{N}_0$ gibt, so daß $(\varphi - \lambda\,\mathrm{id})^\ell(\vec{v}) = \vec{0}$ ist. Falls $(\varphi - \lambda\,\mathrm{id})^{\ell-1} \neq \vec{0}$ ist, bezeichnen wir ℓ als die *Stufe* des Hauptvektors.

b) Die Menge aller Hauptvektoren von φ zum Eigenwert λ heißt *Hauptraum* zu λ und wird mit H_λ bezeichnet.

Insbesondere sind die Hauptvektoren der Stufe eins genau die Eigenvektoren zum Eigenwert λ: Der Nullvektor ist nämlich kein Hauptvektor erster Stufe, da er bereits von $(\varphi - \lambda\,\mathrm{id})^0 = \mathrm{id}$ auf $\vec{0}$ abgebildet wird.

Es ist klar, daß die Hauptvektoren einen Untervektorraum bilden, denn mit $(\varphi - \lambda\,\mathrm{id})$ sind auch dessen Schachtelungen

$$(\varphi - \lambda\,\mathrm{id})^\ell = (\varphi - \lambda\,\mathrm{id}) \circ \cdots \circ (\varphi - \lambda\,\mathrm{id})$$

lineare Abbildungen, und die Hauptvektoren der Stufe höchstens ℓ sind gerade die Elemente des Kerns dieser Abbildung. Da wir von einem endlichdimensionalen Vektorraum V ausgehen, kann die Folge dieser Kerne nicht unbeschränkt wachsen, es gibt also ein maximales ℓ, das als Stufe eines Hauptvektors auftreten kann. Mit diesem ℓ ist der Hauptraum H_λ gerade der Kern von $(\varphi - \lambda\,\mathrm{id})^\ell$.

Der Nutzen der Haupträume ergibt sich aus folgendem

Lemma: H_λ ist ein φ-invarianter Unterraum von V. Bezeichnet ℓ die größte Stufe eines Hauptvektors aus H_λ, so ist $\mathrm{Bild}(\varphi - \lambda\,\mathrm{id})^\ell$ ein φ-invariantes Komplement.

Beweis: Beginnen wir mit der Invarianz von H_λ unter φ.

Ist $\vec{v}$ ein Hauptvektor der Stufe j, so ist

$$\begin{aligned}(\varphi - \lambda\,\mathrm{id})^j(\vec{v}) &= (\varphi - \lambda\,\mathrm{id})^{j-1}\big((\varphi - \lambda\,\mathrm{id})(\vec{v})\big) \\ &= (\varphi - \lambda\,\mathrm{id})^{j-1}\big(\varphi(\vec{v}) - \lambda\vec{v}\big) = \vec{0}\,,\end{aligned}$$

$\varphi(\vec{v}) - \lambda\vec{v}$ ist also ein Hauptvektor der Stufe höchstens $j-1$ und somit insbesondere ein Element von H_λ. Da mit $\vec{v}$ auch $\lambda\vec{v}$ in H_λ liegt, ist damit auch

$$\varphi(\vec{v}) = \big(\varphi(\vec{v}) - \lambda\vec{v}\big) + \lambda\vec{v} \in H_\lambda$$

ein Hauptvektor.

Für $\vec{v} \in \mathrm{Bild}(\varphi - \lambda\,\mathrm{id})^\ell$ liegt zunächst auch $\varphi(\vec{v}) - \lambda\vec{v}$ im Bild von $(\varphi - \lambda\,\mathrm{id})^\ell$, denn das Bild von $(\varphi - \lambda\,\mathrm{id})^\ell$ enthält das Bild von

$(\varphi - \lambda\,\mathrm{id})^{\ell-1}$. Da mit $\vec{v}$ auch $\lambda\vec{v}$ im Bild liegt, folgt wie oben die Behauptung.

Als nächstes müssen wir zeigen, daß der Durchschnitt der beiden Summanden nur aus dem Nullvektor besteht. Dazu sei $\vec{v}$ ein Vektor aus diesem Durchschnitt. Dann liegt $\vec{v}$ sowohl im Kern als auch im Bild der linearen Abbildung $(\varphi - \lambda\,\mathrm{id})^{\ell}$, es gibt also einen Vektor $\vec{w} \in v$, so daß $\vec{v} = (\varphi - \lambda\,\mathrm{id})^{\ell}(\vec{w})$ ist, und $(\varphi - \lambda\,\mathrm{id})^{\ell}(\vec{v}) = (\varphi - \lambda\,\mathrm{id})^{2\ell}(\vec{w}) = \vec{0}$. Damit liegt $\vec{w}$ aber im Hauptraum zu λ, d.h. $\vec{v} = (\varphi - \lambda\,\mathrm{id})^{\ell}(\vec{w}) = \vec{0}$.

Nach der Dimensionsformel ist

$$\dim \mathrm{Bild}(\varphi - \lambda\,\mathrm{id})^{\ell} = \dim V - \dim \mathrm{Kern}(\varphi - \lambda\,\mathrm{id})^{\ell}\,,$$

also ist

$$\dim \mathrm{Kern}(\varphi - \lambda\,\mathrm{id})^{\ell} + \dim \mathrm{Bild}(\varphi - \lambda\,\mathrm{id})^{\ell} = \dim V\,,$$

die beiden Untervektorräume erzeugen somit ganz V. ∎

Die Zerlegung von V nach diesem Lemma heißt FITTING-Zerlegung.

Der deutsche Mathematiker HANS FITTING (1906–1938) beschäftigte sich vor allem mit der Untersuchung von Operatoren (und Operatorenringen). Trotz seines frühen Todes konnte er damit wesentliche Beiträge zur Algebra leisten, vor allem auch zur Erforschung der Struktur von Gruppen.

Das Schöne an der FITTING-Zerlegung ist, daß sie rekursiv fortgesetzt werden kann: Da $\mathrm{Bild}(\varphi - \lambda\,\mathrm{id})^{\ell}$ auch φ-invariant ist, können wir für die Einschränkung von φ auf diesen Unterraum einen Hauptraum zu einem anderen Eigenwert abspalten *usw.* Bevor wir uns das genauer überlegen, wollen wir uns aber zunächst eine gute Basis für den Hauptraum H_λ verschaffen.

Lemma: Der Hauptraum H_λ hat eine Basis, bezüglich derer die Abbildungsmatrix von φ eine obere Dreiecksmatrix ist. Alle Hauptdiagonaleinträge dieser Matrix sind gleich λ.

Beweis: Wir beginnen mit einer Basis $\{\vec{b}_1, \dots, \vec{b}_{r_1}\}$ des Eigenraums zu λ und ergänzen diese zu einer Basis $\{\vec{b}_1, \dots, \vec{b}_{r_2}\}$ des Raums aller

Hauptvektoren der Stufe höchstens zwei und so weiter, bis eine Basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ des gesamten Hauptraums erreicht ist.

Der Vektor $\vec{b}_i$ sei Hauptvektor der Stufe ℓ; dann ist

$$\begin{aligned}(\varphi - \lambda\,\mathrm{id})^{\ell}(\vec{b}_i) &= (\varphi - \lambda\,\mathrm{id})^{\ell-1}\big((\varphi - \lambda\,\mathrm{id})(\vec{b}_i)\big)\\ &= (\varphi - \lambda\,\mathrm{id})^{\ell-1}\big(\varphi(\vec{b}_i) - \lambda\vec{b}_i\big) = \vec{0}\,,\end{aligned}$$

$\varphi(\vec{b}_i) - \lambda\vec{b}_i$ ist also ein Hauptvektor der Stufe höchstens $\ell - 1$. Nach Konstruktion der Basis ist $\varphi(\vec{b}_i) - \lambda\vec{b}_i$ daher eine Linearkombination von Basisvektoren $\vec{b}_j$ mit Indizes echt kleiner i, d.h.

$$\varphi(\vec{b}_i) = \lambda\vec{b}_i + \sum_{j=1}^{i-1} \alpha_{ij}\vec{b}_j\,.$$

Bezüglich der Basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ hat die Abbildungsmatrix daher in der Tat die gewünschte Form. ∎

Abspaltung immer weiterer Haupträume auch vom invarianten Komplement führt schließlich zum

Satz: Zu einer linearen Abbildung $\varphi\colon V \to V$ eines endlichdimensionalen k-Vektorraums V gibt es genau dann eine Basis von V, bezüglich derer die Abbildungsmatrix von φ eine Dreiecksmatrix ist, wenn das charakteristische Polynom von φ über k als Produkt von Linearfaktoren geschrieben werden kann. Alsdann kann die Basis so gewählt werden, daß die Abbildungsmatrix die Form

$$\begin{pmatrix} \boxed{A_1} & \mathbf{0} & \dots & \mathbf{0}\\ \mathbf{0} & \boxed{A_2} & \dots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \dots & \boxed{A_s}\end{pmatrix}$$

hat mit Dreiecksmatrizen A_i, die auf der Hauptdiagonalen den i-ten Eigenwert λ_i stehen haben.

Beweis: Falls es zu φ eine Basis gibt, bezüglich derer die Abbildungsmatrix A von φ eine Dreiecksmatrix ist, ist bezüglich dieser Basis auch $A-\lambda E$ eine Dreiecksmatrix. Da die Determinante einer Dreiecksmatrix gerade das Produkt der Diagonaleinträge ist, bekommen wir als charakteristisches Polynom $\det(A-\lambda E)$ ein Produkt von Linearfaktoren.

Falls umgekehrt das charakteristische Polynom in Linearfaktoren zerfällt, gibt es auf jeden Fall Eigenwerte; λ_1 sei einer davon, und H_1 sei der zugehörige Hauptraum. Dazu gibt es nach dem gerade bewiesenen Lemma eine Basis, bezüglich derer die Abbildungsmatrix von $\varphi_{|H_{\lambda_1}}$ eine obere Dreiecksmatrix ist, deren sämtliche Hauptdiagonaleinträge gleich λ_1 sind.

Nach dem Lemma von der FITTING-Zerlegung gibt es zu H_{λ_1} ein φ-invariantes Komplement V_1, so daß $V = H_{\lambda_1} \oplus V_1$ ist. Wir ergänzen die Basis von H_{λ_1} durch eine Basis von V_1 zu einer Basis von V; bezüglich dieser Basis hat φ dann eine Abbildungsmatrix der Form

$$A_1 = \begin{pmatrix} \boxed{D_1} & \mathbf{0} \\ \mathbf{0} & \boxed{B_1} \end{pmatrix}$$

mit einer oberen Dreiecksmatrix

$$D_1 = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_1 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_1 \end{pmatrix} .$$

Entwicklung des charakteristischen Polynoms $\det(A-\lambda E)$ nach der ersten Spalte, gefolgt von der Entwicklung des Rests nach seiner ersten Spalte und so weiter, bis die ersten $r_1 = \dim H_{\lambda_1}$ Spalten aufgebraucht sind, zeigt, daß

$$\det(A_1-\lambda E) = (\lambda_1-\lambda)^{r_1} \cdot \det(B_1-\lambda E)$$

ist, das charakteristische Polynom von $\varphi_{|V_1}$ ist also ein Teiler des charakteristischen Polynoms von φ und zerfällt somit auch in Linearfaktoren.

Insbesondere hat es (mindestens) eine Nullstelle λ_2; wir können deren Hauptraum H_{λ_2} in V_1 betrachten und damit V_1 genau wie oben weiter

zerlegen in H_{λ_2} und dessen invariantes Komplement V_2. Nimmt man nun als Basisvektoren von V zunächst die Basisvektoren von H_{λ_1} wie oben, dann entsprechende Basisvektoren für H_{λ_2} und schließlich noch solche für V_2, hat die Abbildungsmatrix A_2 bezüglich dieser neuen Basis die Form

$$A_2 = \begin{pmatrix} \boxed{D_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boxed{D_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boxed{B_2} \end{pmatrix}$$

mit einer neuen Dreiecksmatrix

$$D_2 = \begin{pmatrix} \lambda_2 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_2 \end{pmatrix} .$$

Auf diese Weise lassen sich sukzessive immer weitere Haupträume abspalten, bis schließlich eine Basis erreicht ist, bezüglich derer die Abbildungsmatrix von φ die Form

$$A = \begin{pmatrix} \boxed{D_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boxed{D_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boxed{D_s} \end{pmatrix}$$

hat mit oberen Dreiecksmatrizen

$$D_i = \begin{pmatrix} \lambda_i & * & \dots & * \\ 0 & \lambda_i & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

zu den Eigenwerten von φ. Ist D_i eine $r_i \times r_i$-Matrix, so ist das charakteristische Polynom von φ

$$\det(A-\lambda E) = (\lambda_1-\lambda)^{r_1}(\lambda_2-\lambda)^{r_2}\cdots(\lambda_1-\lambda)^{r_s} ,$$

die r_i sind also gerade die algebraischen Vielfachheiten der λ_i. ■

Für spätere Anwendungen wollen wir das gerade bewiesene Ergebnis noch etwas umformulieren:

Satz: Falls das charakteristische Polynom von $\varphi: V \to V$ in Linearfaktoren zerfällt, gibt es eine Basis von V, bezüglich derer die Abbildungsmatrix A von φ als $A = D + N$ geschrieben werden kann, wobei D eine Diagonalmatrix ist und N eine obere Dreiecksmatrix mit Nullen in der Hauptdiagonalen. Außerdem ist $DN = ND$.

Beweis: Wir nehmen natürlich die Basis aus dem gerade beendeten Beweis; die Diagonalmatrix D soll genau aus den Diagonalelementen der Abbildungsmatrix A bestehen, also die Eigenwerte entsprechend ihrer algebraischen Vielfachheiten als Diagonalelemente enthalten, und $N = A - D$. Für jede einzelne Dreiecksmatrix D_i aus dem obigen Beweis kommutiert der Diagonalanteil mit dem Rest, da der Diagonalanteil gerade das λ_i-fache der Einheitsmatrix ist. Damit ist auch $DN = ND$, denn bei beiden Multiplikationen treffen, abgesehen von den Nullen, immer nur Einträge aus einem D_i aufeinander. ■

Diese Zerlegung aus diesem Satz bezeichnet man nach dem französischen Mathematiker CAMILLE JORDAN als JORDAN-Zerlegung.

MARIE ENNEMOND CAMILLE JORDAN (1838–1922) arbeitete bei der Herleitung dieser und weiterer Zerlegungen nicht mit komplexen Matrizen, sondern mit Matrizen über endlichen Körpern, motiviert durch Fragen aus der Gruppentheorie und Lösbarkeitsfragen für nichtlineare Gleichungen. Weitere Arbeiten beschäftigen sich mit der Anwendung gruppentheoretischer Methoden auf die Geometrie sowie mit der Topologie, wo er z.B. bewies, daß jede doppeltpunktfreie geschlossene Kurve die Ebene in zwei Gebiete zerlegt. Außerdem entwickelte er neue Methoden zum Nachweis der Konvergenz von FOURIER-Reihen.

Ziel unserer Betrachtungen in diesem Paragraphen war die Berechnung von Potenzen und Exponentialfunktionen einer Matrix. Mit der JORDAN-Zerlegung ist dies im wesentlichen erreicht: Da D und N miteinander

kommutieren, gilt für Potenzen der Summe $D + N$ der „übliche“ binomische Lehrsatz, d.h.

$$(D+N)^m = \sum_{\ell=1}^{m} \binom{m}{\ell} D^{m-\ell} N^{\ell} ,$$

und

$$e^{D+N} = e^D \cdot e^N .$$

Die Potenzen von D sind sehr einfach zu berechnen: D^j ist wieder eine Diagonalmatrix, ihre Diagonalelemente sind die j-ten Potenzen der Diagonalelemente von D; genauso ist e^D einfach die Diagonalmatrix mit den Exponentialfunktionen der Einträge von D als Einträgen.

N ist eine obere Dreiecksmatrix mit Nullen in der Hauptdiagonalen, wir wissen also bereits, daß es einen Exponenten gibt, ab dem alle Potenzen gleich der Nullmatrix sind, so daß die Exponentialreihe zu einer endlichen Summe wird und auch in der binomischen Formel selbst für große m nur relativ wenige Summanden auftreten.

Mit der JORDAN-Zerlegung können wir diese Aussage nun noch etwas präzisieren: Die lineare Abbildung ψ zu N bildet den i-ten Basisvektor $\vec{b}_i$ ab in den von Basisvektoren $\vec{b}_j$ mit $j \le i - 1$ erzeugten Unterraum. Für diese Basisvektoren gilt eine analoge Aussage, $\psi^{(2)}(\vec{b}_i)$ liegt daher im Unterraum, den die $\vec{b}_j$ mit $j \le i - 2$ aufspannen. Induktiv folgt, daß $\psi^{(\ell)}(\vec{b}_i)$ im von den $\vec{b}_j$ mit $j \le i - \ell$ aufgespannten Untervektorraum liegt; falls $i - \ell$ negativ wird, ist das natürlich der Nullraum.

Die Abbildungsmatrix N^ℓ von $\psi^{(\ell)}$ ist daher ebenfalls eine obere Dreiecksmatrix mit Nullen in der Hauptdiagonalen; zusätzlich stehen auch noch in den $\ell - 1$ schrägen Reihen oberhalb und parallel zur Hauptdiagonale lauter Nullen, und spätestens wenn ℓ größer oder gleich der größten Stufe eines Hauptvektors wird, ist N^ℓ gleich der Nullmatrix.

Als Beispiel betrachten wir die Matrix

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

die offensichtlich von der oben betrachteten Form ist; hier ist

$$D = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad N = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

Eine kurze Rechnung zeigt, daß

$$N^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad N^3 = 0$$

ist, also ist beispielsweise

$$A^{10} = D^{10} + 10\,D^9 N + 55\,D^8 N^2 = \begin{pmatrix} 1024 & 5120 & 0 & 0 & 0 \\ 0 & 1024 & 0 & 0 & 0 \\ 0 & 0 & 1 & 20 & 390 \\ 0 & 0 & 0 & 1 & 40 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} .$$

Mehr als Potenzen interessiert uns die Exponentialfunktion einer Matrix; auch diese läßt sich über die JORDAN-Zerlegung berechnen: Da D und N kommutieren, ist $e^A = e^{D+N} = e^D \cdot e^N$ mit

$$e^D = \begin{pmatrix} e^2 & 0 & 0 & 0 & 0 \\ 0 & e^2 & 0 & 0 & 0 \\ 0 & 0 & e & 0 & 0 \\ 0 & 0 & 0 & e & 0 \\ 0 & 0 & 0 & 0 & e \end{pmatrix}$$

und

$$e^N = E + N + \frac{1}{2}N^2 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} ,$$

also ist

$$e^A = e^D \cdot e^N = \begin{pmatrix} e^2 & e^2 & 0 & 0 & 0 \\ 0 & e^2 & 0 & 0 & 0 \\ 0 & 0 & e & 2e & 7e \\ 0 & 0 & 0 & e & 4e \\ 0 & 0 & 0 & 0 & e \end{pmatrix} .$$

Entsprechend läßt sich auch e^{At} berechnen:

$$e^{Nt} = E + Nt + \frac{1}{2}N^2t^2 = \begin{pmatrix} 1 & t & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2t & 3t + 4t^2 \\ 0 & 0 & 0 & 1 & 4t \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} ,$$

und da auch Dt und Nt kommutieren, ist

$$e^{At} = e^{Dt} \cdot e^{Nt} = \begin{pmatrix} e^{2t} & te^{2t} & 0 & 0 & 0 \\ 0 & e^{2t} & 0 & 0 & 0 \\ 0 & 0 & e^t & 2te^t & 3te^t + 4t^2e^t \\ 0 & 0 & 0 & e^t & 4te^t \\ 0 & 0 & 0 & 0 & e^t \end{pmatrix} .$$

Zur Vorsicht sei noch einmal ausdrücklich darauf hingewiesen, daß es für diese Rechnungen sehr wesentlich war, daß D und N miteinander kommutieren; es reicht nicht, wenn wir die Matrix A nur auf *irgendeine* Dreiecksgestalt bringen und dann als Summe einer Diagonalmatrix und einer oberen Dreiecksmatrix mit Nullen in der Hauptdiagonalen schreiben. Für

$$A = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \underset{\text{def}}{=} D + N$$

beispielsweise ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 6 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} ,$$

und in der Tat ist

$$D^2 + 2DN + N^2 = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 6 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 6 \\ 0 & 4 \end{pmatrix}$$

verschieden von
$$A^2 = \begin{pmatrix} 1 & 9 \\ 0 & 4 \end{pmatrix},$$
und genauso ist
$$e^A = \begin{pmatrix} e & 3e^2 - 3e \\ 0 & e^2 \end{pmatrix} \neq e^D \cdot e^N = \begin{pmatrix} e & 3e \\ 0 & e^2 \end{pmatrix}.$$
Der Grund für die Verschiedenheit der Ergebnisse beim Quadrat liegt natürlich darin, daß wir im allgemeinen nur sagen können, daß
$$(D+N)^2 = D(D+N) + N(D+N) = D^2 + DN + ND + N^2$$
ist, aber wir können $DN + ND$ nicht zusammenfassen zu $2DN$. Mit wachsendem Exponenten verschlimmert sich die Situation drastisch; schon
$$(D+N)^3 = D^3 + D^2N + DND + ND^2 + DN^2 + NDN + N^2D + N^3$$
hat acht Summanden; die m-te Potenz hat 2^m, und von denen überleben viele auch dann, wenn N^r schon für relativ kleine r verschwindet. Für die Matrixexponentialfunktion, in die alle Potenzen eingehen, ist also ziemlich klar, daß es für nichtkommutierende Matrizen D und N keinen vernünftigen Zusammenhang zwischen e^{D+N} und $e^D \cdot e^N$ geben kann.

g) Ein Beispiel

Wir haben Eigenvektoren und Hauptvektoren in erster Linie eingeführt, um Differentialgleichungen zu lösen; daher soll das etwas ausführlichere Beispiel in diesem Abschnitt ebenfalls mit einer Differentialgleichung beginnen: Gesucht sind die Lösungen des Differentialgleichungssystems
$$\begin{aligned} \dot{x}(t) &= 2x(t) - y(t) - z(t) \\ \dot{y}(t) &= x(t) + 5y(t) + 2z(t) \\ \dot{z}(t) &= -x(t) - 2y(t) + z(t). \end{aligned}$$
Hier ist
$$A = \begin{pmatrix} 2 & -1 & -1 \\ 1 & 5 & 2 \\ -1 & -2 & 1 \end{pmatrix},$$

und das charakteristische Polynom von A ist
$$\det(A - \lambda E) = -\lambda^3 + 8\lambda^2 - 21\lambda + 18 = -(\lambda - 2)(\lambda - 3)^2.$$
Wir haben also den Eigenwert zwei mit algebraischer und somit auch geometrischer Vielfachheit eins und den Eigenwert drei mit algebraischer Vielfachheit zwei. In der Matrix
$$A - 2E = \begin{pmatrix} 0 & -1 & -1 \\ 1 & 3 & 2 \\ -1 & -2 & -1 \end{pmatrix}$$
ist die mittlere Spalte gleich der Summe der beiden äußeren,
$$\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$
erzeugt also den Eigenraum. In
$$A - 3E = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 2 & 2 \\ -1 & -2 & -2 \end{pmatrix}$$
stimmen die zweite und die dritte Spalte miteinander überein,
$$\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$
ist also ein Eigenvektor und erzeugt auch den Eigenraum, denn da die erste Spalte kein Vielfaches der zweiten ist, hat die Matrix den Rang zwei. Die geometrische Vielfachheit des Eigenwerts drei ist also nur eins: Um zu einer Dreiecksmatrix zu kommen, müssen wir einen Hauptvektor zweiter Stufe berechnen. Aus
$$(A - 3E)^2 = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$
sieht man, daß sich
$$\vec{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

als von $\vec{v}_2$ linear unabhängiger Kandidat anbietet. Da

$$A\vec{v}_3 = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix} = 3\vec{v}_3 - \vec{v}_2$$

ist, hat A bezüglich der Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ die Form

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

der erste „Kasten" ist also einfach eine 1×1-Matrix und der zweite ist

$$\begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}.$$

Da das Quadrat des zweiten Summanden verschwindet, ist

$$e^{\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} t} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} t = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}$$

und

$$e^{\begin{pmatrix} 3 & -1 \\ 0 & 3 \end{pmatrix} t} = e^{3t} \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{3t} & -te^{3t} \\ 0 & e^{3t} \end{pmatrix}.$$

Damit ist

$$e^{Mt} = \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{3t} & -te^{3t} \\ 0 & 0 & e^{3t} \end{pmatrix}.$$

Um daraus e^{At} zu berechnen, müssen wir die Standardbasis des $\mathbb{R}^3$ durch die Hauptvektoren ausdrücken; man überzeugt sich leicht, daß

$$\vec{e}_1 = \vec{v}_1 + \vec{v}_2, \quad \vec{e}_2 = \vec{v}_1 + \vec{v}_2 - \vec{v}_3 \quad \text{und} \quad \vec{e}_3 = \vec{v}_1 - \vec{v}_3$$

ist. Bezüglich der Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ ist also

$$e^{At}\vec{e}_1 = \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^{3t} & -te^{3t} \\ 0 & 0 & e^{3t} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{2t} \\ e^{3t} \\ 0 \end{pmatrix},$$

was bezüglich der Standardbasis der Vektor

$$e^{2t}\vec{v}_1 + e^{3t}\vec{v}_2 = \begin{pmatrix} e^{2t} \\ -e^{2t} \\ e^{2t} \end{pmatrix} + \begin{pmatrix} 0 \\ e^{3t} \\ -e^{3t} \end{pmatrix} = \begin{pmatrix} e^{2t} \\ e^{3t} - e^{2t} \\ e^{2t} - e^{3t} \end{pmatrix}.$$

Also ist, bezüglich der Standardbasis ausgedrückt,

$$e^{At} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{2t} \\ e^{3t} - e^{2t} \\ e^{2t} - e^{3t} \end{pmatrix}.$$

Genauso überlegt man sich, daß

$$e^{At} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = e^{2t}\vec{v}_1 + \left(e^{3t} + te^{3t}\right)\vec{v}_2 - e^{3t}\vec{v}_3 = \begin{pmatrix} e^{2t} - e^{3t} \\ 2e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - e^{3t} - te^{3t} \end{pmatrix}$$

und

$$e^{At} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{2t} - e^{3t} \\ e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - te^{3t} \end{pmatrix}$$

ist. Da in den Spalten einer Matrix die Bilder der Basisvektoren stehen, ist somit

$$e^{At} = \begin{pmatrix} e^{2t} & e^{2t} - e^{3t} & e^{2t} - e^{3t} \\ e^{3t} - e^{2t} & 2e^{3t} - e^{2t} + te^{3t} & e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - e^{3t} & e^{2t} - e^{3t} - te^{3t} & e^{2t} - te^{3t} \end{pmatrix}.$$

Die Lösung, die den Anfangsbedingungen

$$x(0) = x_0, \quad y(0) = y_0 \quad \text{und} \quad z(0) = z_0$$

genügt, ist also

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} e^{2t} & e^{2t} - e^{3t} & e^{2t} - e^{3t} \\ e^{3t} - e^{2t} & 2e^{3t} - e^{2t} + te^{3t} & e^{3t} - e^{2t} + te^{3t} \\ e^{2t} - e^{3t} & e^{2t} - e^{3t} - te^{3t} & e^{2t} - te^{3t} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

$$= \begin{pmatrix} (x_0 + y_0 + z_0)e^{2t} - (y_0 + z_0)e^{3t} \\ \left(x_0 + (t+2)y_0 + (t+1)z_0\right)e^{3t} - (x_0 + y_0 + z_0)e^{2t} \\ (x_0 + y_0 + z_0)e^{2t} - \left(x_0 + (t+1)y_0 + tz_0\right)e^{3t} \end{pmatrix}.$$

h) Ergänzung: Die Jordan-Normalform

Die im vorigen Abschnitt konstruierte Normalform für Abbildungsmatrizen wird für alle Zwecke dieser Vorlesung ausreichen. Trotzdem ist sie nicht ganz befriedigend, da die Dreiecksmatrizen immer noch sehr willkürlich und damit komplizierter als notwendig sind. Für Interessenten sei in diesem Abschnitt gezeigt, wie sich die bislang erreichte

Dreiecksgestalt noch weiter vereinfachen läßt. Für das folgende werden wir die Ergebnisse dieses Abschnitts nicht benötigen; er kann also auch gefahrlos überlesen werden.

Die Potenzen einer oberen Dreiecksmatrix mit Nullen in der Hauptdiagonalen verschwinden, wie wir gesehen haben, ab einem meist überschaubar kleinen Exponenten, aber die Potenzen bis dahin muß man doch mühsam von Hand ausrechnen. Eine Ausnahme, bei der alles klar ist, bilden Matrizen der Form

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

bei denen direkt oberhalb der Hauptdiagonale lauter Einsen stehen, während alle anderen Einträge verschwinden, d.h.

$$N = (n_{ij}) \quad \text{mit} \quad n_{ij} = \begin{cases} 1 & \text{falls } j - i = 1 \\ 0 & \text{sonst} \end{cases}.$$

Bei der sukzessiven Potenzierung von N verschiebt sich einfach die Reihe von Einsen jeweils um eins weiter nach außen, d.h.

$$N^\ell = (n_{ij}^{(\ell)}) \quad \text{mit} \quad n_{ij}^{(\ell)} = \begin{cases} 1 & \text{falls } j - i = \ell \\ 0 & \text{sonst} \end{cases},$$

denn die zu N gehörige lineare Abbildung ψ bildet einfach den i-ten Basisvektor auf den $(i-1)$-ten ab oder auf den Nullvektor, falls es keinen $(i-1)$-ten Basisvektor mehr gibt, und entsprechend ist $\psi^{(\ell)}(\vec{b}_i) = \vec{b}_{i-\ell}$ beziehungsweise $\vec{0}$.

In diesem speziellen Fall sind die Potenzen von N also ohne jeden Aufwand zu berechnen, und tatsächlich genügen solche Matrizen N schon vollständig für eine Normalform der Abbildungsmatrix, die sogenannte JORDAN-Normalform:

Satz: Falls das charakteristische Polynom von $\varphi\colon V \to V$ als Produkt von Linearfaktoren geschrieben werden kann, gibt es eine Basis von V,

bezüglich derer die Abbildungsmatrix von φ die Form

$$A = \begin{pmatrix} \boxed{J_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boxed{J_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \boxed{J_s} \end{pmatrix}$$

hat mit oberen Dreiecksmatrizen

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

zu den Eigenwerten von φ. Die Anzahl der Kästchen J_i zu einem festen Eigenwert ist die geometrische Vielfachheit dieses Eigenwerts, die Summe ihrer Zeilenzahlen die algebraische.

Beweis: Wir gehen aus von der Zerlegung von V in die Haupträume zu den Eigenwerten von φ und betrachten einen festen Hauptraum H_λ. Die Einschränkung von φ auf diesen Untervektorraum läßt sich zerlegen in eine Summe

$$\varphi_{|H_\lambda} = \lambda\,\mathrm{id} + \psi\,;$$

dabei ist die Abbildungsmatrix von $\lambda\,\mathrm{id}$ bezüglich jeder beliebigen Basis gleich dem λ-fachen der Einheitsmatrix, und zumindest bezüglich der im vorigen Abschnitt konstruierten Basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ ist die Abbildungsmatrix von ψ eine obere Dreiecksmatrix N mit Nullen in der Hauptdiagonalen.

ψ bildet den Basisvektor $\vec{b}_i$ daher ab in das Erzeugnis der Basisvektoren $\vec{b}_1$ bis $\vec{b}_{i-1}$; insbesondere geht $\vec{b}_1$ auf den Nullvektor. Wiederholte Anwendung von ψ zeigt, daß für jeden Basisvektor $\vec{b}_i$ gilt: $\psi^{(i-1)}(\vec{b}_i) = \vec{0}$, wobei der Exponent von ψ für die wiederholte Anwendung der Abbildung stehen soll. Insbesondere ist also $\psi^{(r)}(\vec{v}) = \vec{0}$ für alle $\vec{v} \in H_\lambda$.

Es könnte sein, daß es schon eine kleinere Zahl s gibt, so daß $\psi^{(s)}$ die Nullabbildung ist; die kleinste solche Zahl bezeichnen wir als den *Nilpotenzgrad* von ψ.

Hat der Nilpotenzgrad seinen größtmöglichen Wert r, so sind die Vektoren

$$\vec{b}_r,\ \psi(\vec{b}_r),\ \dots,\ \psi^{(s-1)}(\vec{b}_r)$$

allesamt ungleich dem Nullvektor. Sie sind auch linear unabhängig, denn ist

$$\alpha_0\vec{b}_r + \alpha_1\psi(\vec{b}_r) + \cdots + \alpha_{r-1}\psi^{(r-1)}(\vec{b}_r) = \vec{0}\,,$$

so ist auch für jedes j

$$\begin{aligned}&\psi^{(j)}\big(\alpha_0\vec{b}_s + \alpha_1\psi(\vec{b}_s) + \cdots + \alpha_{r-1}\psi^{(r-1)}(\vec{b}_r)\big)\\ &= \alpha_0\psi^{(j)}(\vec{b}_r) + \alpha_1\psi^{(j+1)}(\vec{b}_r) + \cdots + \alpha_{r-1}\psi^{(j+r-1)}(\vec{b}_r)\big) = \vec{0}\,.\end{aligned}$$

Da $\psi^{(s)}$ für $s \geq r$ die Nullabbildung ist, treten hier nur die Summanden $\alpha_i\psi^{(i+j)}(\vec{b}_s)$ mit $i < r-j$ wirklich auf, für $j = r-1$ also nur der Summand $\alpha_0\psi^{(r-1)}(\vec{b}_r)$. Da $\psi^{(r-1)}(\vec{b}_r)$ ungleich dem Nullvektor ist, muß also $\alpha_0 = 0$ sein. Anwendung von $\psi^{(r-2)}$ zeigt als nächstes, daß $\alpha_1 = 0$ ist, und genauso zeigt man sukzessive das Verschwinden aller α_i. Also können wir

$$\vec{c}_1 = \psi^{(r-1)}(\vec{b}_r), \quad \vec{c}_2 = \psi^{(r-2)}(\vec{b}_r), \quad \dots, \quad \vec{c}_r = \vec{b}_r$$

als Basisvektoren von H_λ wählen, und bezüglich dieser Basis ist

$$\psi(\vec{c}_i) = \begin{cases} c_{i-1} & \text{für } i > 1\\ \vec{0} & \text{für } i = 1\end{cases} \quad\text{und}\quad \varphi(\vec{c}_i) = \begin{cases} \lambda\vec{c}_i + \vec{c}_{i-1} & \text{für } i > 1\\ \lambda\vec{c}_1 & \text{für } i = 1\end{cases}\,.$$

Die Abbildungsmatrix von φ hat somit die einfache Gestalt

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0\\ 0 & \lambda & 1 & \dots & 0 & 0\\ 0 & 0 & \lambda & \dots & 0 & 0\\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots\\ 0 & 0 & 0 & \dots & \lambda & 1\\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix},$$

und das ist gerade eines der JORDAN-Kästchen aus der Formulierung des Satzes.

Falls der Nilpotenzgrad s von ψ kleiner als r ist, können wir nicht so argumentieren. Wir können aber immerhin einen Vektor $\vec{v} \in H_\lambda$ finden, so daß $\psi^{(s-1)}(\vec{v}) \neq \vec{0}$ ist, denn erst $\psi^{(s)}$ ist die Nullabbildung. Genau wie oben folgt, daß

$$\vec{c}_1 = \psi^{(s-1)}(\vec{v}), \quad \vec{c}_2 = \psi^{(s-2)}(\vec{v}), \quad \dots, \quad \vec{c}_s = \vec{v}$$

linear unabhängig sind, allerdings spannen sie nur einen s-dimensionalen Teilraum U von H_λ auf. Dieser Teilraum ist ψ-invariant und damit auch φ-invariant, denn ψ bildet einfach die Basisvektoren aufeinander beziehungsweise auf den Nullvektor ab, und die Abbildungsmatrizen bezüglich dieser Basis sehen genauso aus wie oben; auch zu U gehört also ein JORDAN-Kästchen.

Um weitere Kästchen zu bekommen, brauchen wir ein invariantes Komplement von U in H_λ. Dazu wählen wir irgendeine lineare Abbildung $\omega\colon V \to k$, für die $\omega(\vec{v}) \neq 0$ ist und setzen

$$W = \left\{\vec{w} \in H_\lambda \;\middle|\; \omega(\vec{w}) = \omega\big(\psi(\vec{w})\big) = \cdots = \omega\big(\psi^{(s-1)}(\vec{w})\big) = 0i\right\}.$$

Der Durchschnitt $U \cap W$ besteht nur aus dem Nullvektor, denn jeder Vektor aus U läßt sich als

$$\vec{w} = \alpha_1\vec{c}_1 + \ldots + \alpha_s\vec{c}_s$$

schreiben, und wenn $\vec{w}$ auch in W liegt, ist

$$\omega\big(\psi^{(j)}(\vec{w})\big) = \alpha_1\psi^{(j+s-1)}(\vec{v}) + \ldots + \alpha_{s-1}\psi^{j+1}(\vec{v}) + \alpha_s\psi^{(j)}(\vec{v}) = 0$$

für $j = 0, \dots, s-1$. Da $\psi^{(\ell)}(\vec{v})$ für $\ell \geq s$ gleich dem Nullvektor ist, folgt für $j = s-1$, daß $\alpha_{s-1} = 0$ ist, und erniedrigt man j immer weiter, folgt nacheinander das Verschwinden aller Koeffizienten α_i. Somit ist $U \cap W$ in der Tat der Nullraum.

Die Dimension von W läßt sich zumindest nach unten leicht abschätzen: Bezüglich einer Basis von H_λ wird jede Gleichung $\omega\big(\psi^{(j)}(\vec{w})\big) = 0$ zu einer linearen Gleichung in den Koeffizienten von $\vec{w}$, der Untervektorraum W ist also die Lösungsmenge eines homogenen linearen Gleichungssystems aus s Gleichungen in $\dim H_\lambda$ Variablen. Daher ist

$$\dim W \geq \dim H_\lambda - s \quad\text{und}\quad \dim U \oplus W = \dim U + \dim W \geq \dim H_\lambda\,.$$

Da $U \oplus W$ Untervektorraum von H_λ ist, geht das nur, wenn das Gleichheitszeichen gilt, d.h. $H_\lambda = U \oplus W$.

Wir müssen uns noch überlegen, daß W unter ψ invariant ist. Dazu müssen wir zeigen, daß für alle $\vec{w} \in W$ gilt

$$\omega\big(\psi(\vec{w})\big) = \omega\Big(\psi\big(\big(\psi(\vec{w})\big)\Big) = \cdots = \omega\Big(\psi^{(s-1)}\big((\vec{w})\big)\Big) = 0\,,$$

d.h.

$$\omega\big(\psi(\vec{w})\big) = \omega\big(\psi^{(2)}(\vec{w})\big) = \cdots = \omega\big(\psi^{(s)}(\vec{w})\big) = 0$$

falls

$$\omega(\vec{w}) = \omega\big(\psi(\vec{w})\big) = \cdots = \omega\big(\psi^{(s-1)}(\vec{w})\big) = 0$$

ist. Die einzige neue Bedingung ist $\omega\big(\psi^{(s)}(\vec{w})\big) = 0$, und die ist trivialerweise erfüllt, da $\psi^{(s)}$ die Nullabbildung ist. Also ist W invariant unter ψ und somit ein invariantes Komplement von U.

Auch $\psi_{|W}$ ist eine nilpotente Abbildung von einem Nilpotenzgrad $s' \le s$; wenn wir also einen Vektor $\vec{w} \in W$ hernehmen, für den $\psi^{(s')}(\vec{w}) \ne \vec{0}$ ist, können wir die gleiche Kontruktion wie oben mit $\vec{v}$ noch einmal durchführen und erhalten einen neuen invarianten Unterraum $U' \le W$ mit einer Basis, bezüglich derer φ ein JORDAN-Kästchen als Abbildungsmatrix hat.

Falls $U' = W$ ist, sind wir damit fertig; anderfalls können wir wieder wie oben ein invariantes Komplement W' von U' in W finden und einen weiteren Teilraum abspalten, *usw.* Jeder solche Teilraum führt auf ein JORDAN-Kästchen, und das Verfahren bricht schließlich ab, da wir in einem endlichdimensionalen Vektorraum arbeiten.

In jedem der konstruierten Teilräume liegt genau ein eindimensionaler Teilraum aus Eigenvektoren (und dem Nullvektor), nämlich der vom ersten Basisvektor aufgespannte. Der Eigenraum zu λ wird also von diesen ersten Basisvektoren aufgespannt und seine Dimension, die geometrische Vielfachheit von λ, ist damit gleich der Anzahl der JORDAN-Kästchen zu λ. Die algebraische Vielfachheit ist wegen der speziellen Gestalt der Abbildungsmatrix natürlich die Anzahl der λ in der Hauptdiagonalen, d.h. gleich der Summe der Zeilenzahlen der JORDAN-Kästchen zu λ. ■

Um wenigstens ein ganz einfaches Beispiel zu sehen, betrachten wir die Matrix

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

vom Ende des vorigen Abschnitts. Links oben steht schon ein JORDAN-Kästchen zum Eigenwert zwei, rechts unten müssen wir noch etwas arbeiten.

Die Basis des $\mathbb{R}^5$ sei $\{\vec{e}_1, \dots, \vec{e}_5\}$; davon können wir $\vec{b}_1 = \vec{e}_1$ und $\vec{b}_2 = \vec{e}_2$ als Basis von H_2 gleich übernehmen. H_1 wird von $\vec{e}_3, \vec{e}_4$ und $\vec{e}_5$ aufgespannt; ein Vektor aus diesem dreidimensionalen Raum, der unter

$$N = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

den maximalen Nilpotenzgrad hat, ist etwa $\vec{e}_5$, denn $\vec{e}_5$ wird abgebildet auf $3\vec{e}_3 + 4\vec{e}_4$; da e_3 auf den Nullvektor geht und $\vec{e}_4$ auf $2\vec{e}_3$, wird dieser Vektor weiter abgebildet auf $8\vec{e}_3$, was schließlich auf den Nullvektor abgebildet wird. Mit

$$\vec{b}_3 = 2\vec{e}_3, \quad \vec{b}_4 = 3\vec{e}_3 + 4\vec{e}_4 \quad \text{und} \quad \vec{b}_5 = \vec{e}_5$$

geht also $\vec{b}_5$ unter N auf $\vec{b}_4$ und weiter auf $\vec{b}_3$; daher ist

$$A\vec{b}_3 = \vec{b}_3, \quad A\vec{b}_4 = \vec{b}_4 + \vec{b}_3 \quad \text{und} \quad A\vec{b}_5 = \vec{b}_5 + \vec{b}_4\,,$$

und die Matrix A hat bezüglich der Basis $\{\vec{b}_1, \dots, \vec{b}_5\}$ die Gestalt

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

mit den beiden JORDAN-Kästchen

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

§3: Anwendung auf lineare Differentialgleichungen

a) Systeme homogener linearer Differentialgleichungen mit konstanten Koeffizienten

Beliebige Systeme homogener linearer Differentialgleichungen mit konstanten Koeffizienten können genauso behandelt werden wie das Beispiel aus §2*g)*, und eigentlich wissen wir bereits alles, was wir brauchen. Trotzdem seien die wesentlichen Punkte hier noch einmal zusammenfassend dargestellt:

Wir betrachten ein Differentialgleichungssystem

$$\dot{\vec{y}}(t) = A\vec{y}(t)\,,$$

wobei A eine reelle oder komplexe $n \times n$-Matrix ist. Zumindest über den komplexen Zahlen zerfällt ihr charakteristisches Polynom in Linearfaktoren, es gibt also eine Basis $\mathcal{B}$ von $\mathbb{C}^n$, bezüglich derer A als obere Dreiecksmatrix geschrieben werden kann.

Ist B die komplexe $n \times n$-Matrix, deren Spalten die Vektoren aus $\mathcal{B}$ sind, ist dann also

$$A = BCB^{-1}$$

mit einer Dreiecksmatrix $C \in \mathbb{C}^{n\times n}$. Dann ist auch

$$e^{At} = Be^{Ct}B^{-1}\,,$$

und e^{Ct} kann berechnet werden, da $C = D + N$ Summe einer Diagonalmatrix und einer damit kommutierenden oberen Dreiecksmatrix mit Nullen in der Hauptdiagonalen ist. Insgesamt ist also

$$e^{At} = Be^{Dt}e^{Nt}B^{-1}\,,$$

und jeder der vier Faktoren rechts ist ein endlich vielen Schritten berechenbar.

Wie wir weiterhin wissen, hat jede Lösung der Differentialgleichung $\dot{\vec{y}}(t) = A\vec{y}(t)$ die Form

$$\vec{y}(t) = e^{At}\vec{y}_0 \quad \text{mit} \quad \boldsymbol{y}_0 = \vec{y}(0) \in \mathbb{C}^n\ ;$$

wir kennen also alle Lösungen des Systems. Falls Anfangsbedingungen vorgegeben sind, die sich auf einen beliebigen Zeitpunkt $t = t_0$ beziehen, können wir die Lösung entsprechend schreiben als

$$\vec{y}(t) = e^{A(t-t_0)}\vec{y}(t_0)\,.$$

Solange wir in der Lage sind, die Nullstellen des charakteristischen Polynoms von A zu berechnen, können wir also jedes Anfangswertproblem in Gestalt eines Systems linearer Differentialgleichungen mit konstanten Koeffizienten lösen.

b) Langzeitverhalten der Lösung

Nicht immer ist es notwendig oder auch nur möglich, ein Differentialgleichungssystem zur exakten numerischen Vorhersage der weiteren Entwicklung eines Systems zu verwenden; gelegentlich reicht auch ein qualitativer Überblick. Dabei geht es vor allem um das Langzeitverhalten des Systems: Nähert es sich einem Gleichgewicht, „explodiert" es, oder wird es auf lange Sicht periodisch, wie wir es etwa vom Fall der erzwungenen Schwingung her kennen.

Für solche Aussagen reicht es im Falle eines linearen homogenen Differentialgleichungssystems $\dot{\vec{y}}(t) = A\vec{y}(t)$, die Eigenwerte der Matrix A zu kennen: Bezüglich einer Basis aus Hauptvektoren läßt sich A in der Form $D + N$ schreiben mit einer Diagonalmatrix D, für die e^{Dt} Diagonalmatrix ist mit den Funktionen $e^{\lambda t}$ als Einträgen, wobei λ die Eigenwerte von A durchläuft. Die Matrix e^{Nt} hat Polynome in t als Einträge; das Produkt $e^{Dt}e^{Nt}$ hat also wegen der speziellen Formen von D und N Produkte von Polynomen in t mit e^{λ^t} als Einträge, wobei der Grad des Polynoms höchstens die um eins verminderte größte Stufe eines Hauptvektors zu λ ist. Bei der Rücktransformation auf die Ausgangsbasis entstehen Linearkombinationen solcher Funktionen, die bei der Multiplikation mit dem Vektor der Anfangswerte selbst wieder linear kombiniert werden. Insgesamt ist also jede Lösungsfunktion eine Linearkombination von Termen der Form $t^j e^{\lambda t}$.

Falls nur ein Eigenwert λ von A positiven Realteil hat, muß das System fast unweigerlich explodieren, da der Betrag von $e^{\lambda t}$ für hinreichend großes t jede vorgegebene Grenze überschreitet. Zwar sind eventuell Anfangsbedingungen möglich, bei denen $e^{\lambda t}$ nicht in der Lösung des Anfangswertproblems auftritt, aber da die Anfangswerte in der Praxis nie durch Naturgesetze gegeben sind, sondern durch fehlerbehaftete Meßwerte, kann schon eine kleine Störung den Term $e^{\lambda t}$ wieder ins Spiel bringen, und auch für sehr kleines c dominiert $ce^{\lambda t}$ langfristig jede beschränkte Funktion.

Haben dagegen *alle* Eigenwerte von A negativen Realteil, geht jeder Term $t^j e^{\lambda t}$ gegen Null für $t \to \infty$ und damit auch jede Lösungsfunktion; der Lösungsvektor nähert sich also immer mehr der Gleichgewichtslösung $\vec{y}(t) \equiv 0$.

Bleibt noch der Fall von Eigenwerten mit Realteil null. Hier werden die bei mehrfachen Eigenwerten möglichen Polynome wichtig, denn während der Betrag von $e^{\lambda t}$ konstant ist, geht der Betrag der mit einem nichtkonstanten Polynom multiplizierten Funktion gegen $\pm\infty$.

Schließlich sollten wir auch die Imaginärteile der Eigenwerte nicht ganz vergessen: Falls A, wie in den meisten Anwendungen, eine reelle Matrix ist, ist mit jedem nichtreellen Eigenwert λ auch die konjugiert komplexe Zahl $\overline{\lambda}$ ein Eigenwert derselben Vielfachheit wie λ. Mit $\lambda = a + ib$ ist

$$e^{\lambda t} = e^{at}(\cos bt + i \sin bt) \quad \text{und} \quad e^{\overline{\lambda} t} = e^{at}(\cos bt - i \sin bt)\,,$$

jede Linearkombination von $e^{\lambda t}$ und $e^{\overline{\lambda} t}$ läßt sich also auch als Linearkombination der beiden reellen Funktionen

$$e^{at} \cos bt \quad \text{und} \quad e^{at} \sin bt$$

schreiben. Die Imaginärteile bringen also Schwingungsanteile in die Lösungsfunktionen.

Um einen ersten Eindruck vom Lösungsverhalten linearer homogener Differentialgleichungssysteme mit reellen Koeffizienten zu bekommen, wollen wir uns überlegen, was in niedrigen Dimensionen möglich ist.

Im Eindimensionalen besteht das „System“ einfach aus der Differentialgleichung $\dot{y}(t) = ay(t)$ mit Lösung $y(t) = y(0)e^{at}$; für $a > 0$ geht dies je

nach Vorzeichen von $y(0)$ gegen plus oder minus unendlich für $t \to \infty$, für $a < 0$ gegen null, und für $a = 0$ ist $y(t) = y(0)$ konstant.

Etwas interessanter ist es im Zweidimensionalen: Hier hat die Matrix A einen oder zwei Eigenwerte.

Falls es nur einen gibt und dieser die geometrische Vielfachheit zwei hat, ändert sich nichts wesentliches gegenüber der eindimensionalen Situation: Die Lösungsfunktion ist $\binom{x(0)}{y(0)} \cdot e^{\lambda t}$.

Falls der Eigenwert nur die algebraische Vielfachheit eins hat, sind die Lösungsfunktionen Linearkombinationen von $e^{\lambda t}$ und $te^{\lambda t}$, also Produkte lineare Funktionen mit $e^{\lambda t}$. Jetzt können die Lösungen auch im Fall $\lambda = 0$ unbeschränkt werden.

Wenn es zwei verschiedene Eigenwerte gibt, können diese entweder beide reell oder aber konjugiert komplex sein.

Im reellen Fall können beide positiv sein; dann geht jede nichttriviale Lösung ins Unendliche, oder aber beide können negativ sein; dann nähert sich jede Lösung immer mehr dem Nullpunkt. Alternativ könnte einer positiv und einer negativ sein; in diesem Fall nähern sich alle Lösungskurven einer Geraden, gehen aber mit dieser ins Unendliche. Schließlich könnte noch ein Eigenwert null sein; dann liegen alle Lösungskurven auf Geraden und gehen dort je nach Vorzeichen des anderen Eigenwerts gegen einen festen Punkt oder aber ins Unendliche.

Bleibt noch der Fall zweier konjugiert komplexer Eigenwerte $a \pm ib$; in diesem Fall sind die Lösungsfunktionen Linearkombinationen von $e^{at} \cos t$ und $e^{at} \sin t$.

Für $a = 0$ haben wir einfach reine Schwingungen; falls die beiden Eigenvektoren senkrecht aufeinander stehen, bekommen wir als Lösungskurven Kreislinien, ansonsten affine Verzerrungen davon, also Ellipsen.

Für $a \neq 0$ ergeben sich entsprechend Spiralen, die je nach Vorzeichen von a entweder ins Unendliche gehen oder aber sich auf den Nullpunkt zusammenziehen.

Betrachten wir als Beispiel das System

$$\dot{x}(t) = -10{,}2x(t) - 25y(t) \quad \text{und} \quad \dot{y}(t) = 5x(t) + 9{,}8y(t)\,.$$

Die Matrix

$$A = \begin{pmatrix} -10\frac{1}{5} & -25 \\ 5 & 9\frac{4}{5} \end{pmatrix}$$

hat das charakteristische Polynom

$$\lambda^2 + \frac{2}{5}\lambda + 25\frac{1}{5}$$

mit Nullstellen $-\frac{1}{5} \pm 5i$, die Lösungskurven sind also sich nach innen zusammenziehende Spiralen. Mit den Anfangsbedingungen $x(0) = 0$ und $y(0) = 1$ erhalten wir die Lösung

$$x(t) = -5e^{-t/5} \sin 5t \quad \text{und} \quad y(t) = e^{-t/5}(\cos 5t + 2\sin 5t)\,,$$

die in Abbildung 29 zu sehen ist.

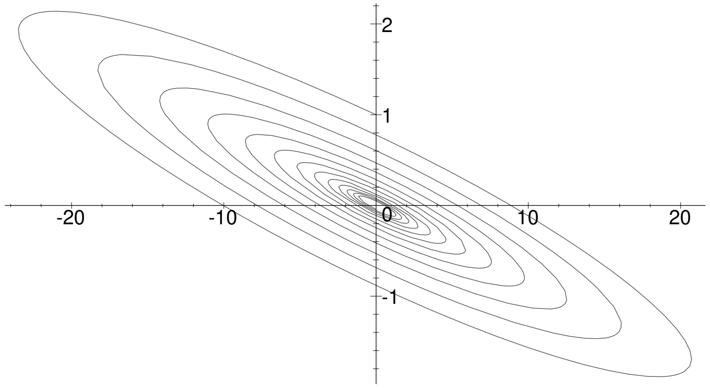

Abb. 29: Spiralförmige Annäherung an den Nullpunkt

Im Dreidimensionalen gibt es entsprechend mehr Möglichkeiten; ich möchte auf die weniger interessanten Fälle mit lauter reellen Eigenwerten verzichten und nur den betrachten, daß zwei konjugiert komplexe auftreten, etwa $a \pm ib$. Der dritte Eigenwert λ muß dann natürlich reell sein.

Falls er null ist, sind wir im wesentlichen in der zweidimensionalen Situation: Da dann in Richtung des dritten Eigenvektors alles konstant

ist, spielt sich alles in einer Ebenen ab, die parallel zu der von den ersten beiden Eigenvektoren aufgespannten liegt.

Falls er negativ ist, nähert sich in Richtung des dritten Eigenvektors alles der Ebenen, in der dessen Koordinate null ist, die also von den beiden anderen Eigenvektoren aufgespannt wird, und dort ist die Dynamik je nach Vorzeichen von a spiralförmig nach innen oder außen oder einfach ellipsenförmig. Abbildung 30 zeigt den Fall $a = -1/10, b = 2$ und $\lambda = -1/6$; die Eigenvektoren zeigen in Richtung der Koordinatenachsen.

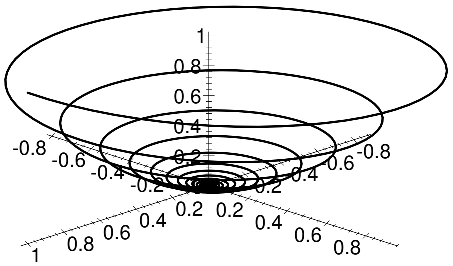

Abb. 30: Spiralförmige Annäherung an den Nullpunkt im Dreidimensionalen

In Abbildung 31 ist $a = +1/10$ und ansonsten alles unverändert; wie man sieht, nähert sich nun die Lösungskurve zwar der (x, y)-Ebenen, geht in dieser aber spiralförmig ins Unendliche.

Falls der dritte Eigenwert positiv ist, geht in Richtung seines Eigenvektors alles ins Unendliche; je nach Vorzeichen von a entfernen sich die Lösungskurven dabei spiralförmig von der durch diesen Eigenvektor aufgespannten Geraden oder aber gehen auf sie zu. Abbildung 32 zeigt den letzteren Fall mit $\lambda = +1/6$ und a, b wie in Abbildung 29.

Man beachte, daß die Abbildungen 31 und 32 zwar auf den ersten Blick sehr ähnlich aussehen, daß aber die Dynamik in Abbildung 31 in z-Richtung nach unten geht, in Abbildung 32 aber nach oben. Der Rotationssinn der Spiralen ist in beiden Fällen der Gegenuhrzeigersinn.

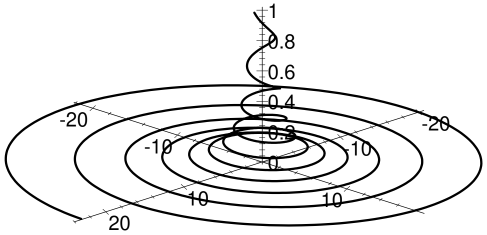

Abb. 31: Spirale, die sich der (x, y)-Ebenen nähert

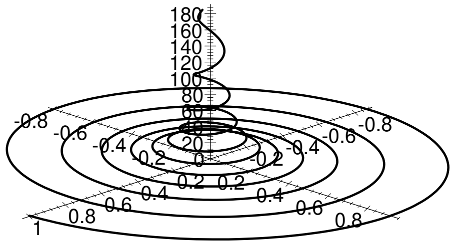

Abb. 32: Spirale, die sich der z-Achse nähert

c) Lineare homogene Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten

Eine der wichtigsten Anwendungen der Theorie aus den letzten Abschnitten in der Elektrotechnik sind lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten. Dieser Abschnitt soll zeigen, daß man in diesem Spezialfall weder Determinanten noch Eigen- und Hauptvektoren berechnen muß, um die Lösungsfunktionen zu finden.

Um zu sehen, was der allgemeine Ansatz in diesem Spezialfall ergibt, schreiben wir die Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) = 0 \quad \text{mit} \quad a_i \in \mathbb{C}\,.$$

als ein System

$$\begin{aligned}
\dot{y}_0(t) &= y_1(t)\\
\dot{y}_1(t) &= y_2(t)\\
&\vdots\\
\dot{y}_{n-2}(t) &= y_{n-1}(t)\\
\dot{y}_{n-1}(t) &= -a_{n-1}y_{n-1}(t) - \cdots - a_1\dot{y}_1(t) - a_0 y_0(t)
\end{aligned}$$

oder kurz $\dot{\vec{y}}(t) = A\vec{y}(t)$ mit

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0\\ 0 & 0 & 1 & \dots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \dots & 1 & 0\\ 0 & 0 & 0 & \dots & 0 & 1\\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

und

$$\vec{y}(t) = \begin{pmatrix} y_0(t)\\ \vdots\\ y_{n-1}(t) \end{pmatrix}.$$

Die Lösungen $\vec{y}(t)$ dieses Systems definieren Lösungen $y(t) = y_0(t)$ der obigen Gleichung, und umgekehrt ist auch für jede solche Lösung $y(t)$ der Vektor

$$\vec{y}(t) = \begin{pmatrix} y(t)\\ \dot{y}(t)\\ \ddot{y}(t)\\ \vdots\\ y^{(n-2)}(t)\\ y^{(n-1)}(t) \end{pmatrix}$$

eine Lösung des Systems $\dot{\vec{y}}(t) = A\vec{y}(t)$.

Aus dem vorigen Abschnitt wissen wir, daß die Lösungen dieses Systems einen n-dimensionalen Vektorraum V von vektorwertigen Funktionen $t \mapsto \vec{y}(t)$ bilden. Jede einzelne Komponente jeder Lösung aus diesem Vektorraum ist darstellbar als Linearkombination von Funktionen $t^j e^{\lambda_i t}$, wobei $\lambda_1, \dots, \lambda_r$ die verschiedenen Eigenwerte von A sind und die nichtnegative ganze Zahl j kleiner ist als die größte Stufe eines Hauptvektors zu λ_i, erst recht also kleiner als die algebraische Vielfachheit von λ_i. Insbesondere stehen daher bei einfachen Eigenwerten oder allgemeiner bei Eigenwerten, deren geometrische Vielfachheit gleich der algebraischen ist, überhaupt keine echten t-Potenzen vor den Exponentialfunktionen.

Außerdem wissen wir, daß jedes Anfangswertproblem eindeutig lösbar ist: Bei Anfangsbedingungen

$$y_0(t_0) = c_0, \quad \dots, \quad y_{n-1}(t_0) = c_{n-1}$$

ist

$$\vec{y}(t) = e^{A(t-t_0)}$$

die Lösung. Vornehm ausgedrückt können wir auch sagen, daß für jedes $t_0 \in \mathbb{R}$ die Abbildung

$$V \to \mathbb{C}^n; \quad \vec{y}(t) \mapsto \vec{y}(t_0)$$

ein Isomorphismus ist.

Für jeden Lösungsvektor $\vec{y}(t)$ des Differentialgleichungssystems ist die nullte Komponenten $y_0(t)$ Lösung der Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) = 0\,,$$

und diese Komponente bestimmt auch die anderen Komponenten $y_i(t)$, die ja gerade die Ableitungen von $y_0(t)$ sind. Daher bilden auch die Lösungsfunktionen dieser Gleichung einen n-dimensionalen Vektorraum, d.h. auch die Abbildung $\vec{y}(t) \mapsto y_0(t)$ ist ein Isomorphismus, und wenn wir Anfangsbedingungen mit ins Spiel bringen folgt auch, daß es für jede Vorgabe von Werten $y(t_0), \dot{y}(t_0), \ddot{y}(t_0), \dots, y^{(n-1)}(t_0)$ genau eine Lösung gibt.

Wie wir uns gerade überlegt haben, ist $y_0(t)$ Linearkombination von Funktionen der Art $t^j e^{\lambda_i t}$, wobei λ_i die Eigenwerte von A durchläuft und j kleiner als die algebraische Vielfachheit des Eigenwerts λ_i ist.

Die Summe der algebraischen Vielfachheiten der (komplexen) Eigenwerte einer komplexen $n \times n$-Matrix ist gleich dem Grad des charakteristischen Polynoms, also gleich n; damit gibt es genau n solche Funktionen. Andererseits wissen wir, daß der Lösungsraum die Dimension n hat; somit werden alle diese Funktionen wirklich gebraucht und sie bilden eine Basis des Lösungsraums.

Für eine wirklich befriedigende Kenntniss des Lösungsraums fehlt uns nun nur noch ein Verfahren, wie wir die Eigenwerte λ_i und deren algebraische Vielfachheiten α_i direkt aus den Koeffizienten der Gleichung berechnen können.

Zumindest die Eigenwerte λ_i lassen sich leicht aus der Gleichung ablesen: Für die Funktion $y(t) = e^{\lambda t}$ ist $y^{(i)}(t) = \lambda^i e^{\lambda t}$, sie ist also genau dann eine Lösung der Differentialgleichung, wenn

$$\begin{aligned} &\lambda^n e^{\lambda t} + a_{n-1}\lambda^{n-1} e^{\lambda t} + \cdots + a_1\lambda e^{\lambda t} + a_0 e^{\lambda t} \\ &= (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)e^{\lambda t} \end{aligned}$$

verschwindet; da $e^{\lambda t}$ nirgends verschwindet, ist dies genau dann der Fall, wenn

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0 \qquad (*)$$

ist.

Definition: Die Gleichung $(*)$ heißt *charakteristische Gleichung* der Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) = 0\,.$$

Die Eigenwerte λ_i von A sind also genau die Nullstellen der charakteristischen Gleichung der Differentialgleichung. Da diese denselben Grad hat wie das charakteristische Polynom von A liegt die Vermutung nahe, daß sie (eventuell bis auf eine Konstante) damit übereinstimmt, und das ist auch tatsächlich der Fall:

Lemma: Die charakteristische Gleichung der Differentialgleichung ist das mit $(-1)^n$ multiplizierte charakteristische Polynom von A.

Beweis: Für $n = 1$ ist die Differentialgleichung $\dot{y}(t) = ay(t)$ identisch mit dem zugehörigen System; die charakteristische Gleichung ist $\lambda - a$, und die „Matrix“ $A = (a)$ hat das charakteristische Polynom $a - \lambda$. Für $n = 1$ stimmt die Behauptung also.

Für $n > 1$ entwickeln wir das charakteristische Polynom

$$\det(A - \lambda E) = \begin{vmatrix} -\lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & -\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\lambda & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} - \lambda \end{vmatrix}$$

nach der ersten Spalte und erhalten

$$\det(A - \lambda E) = -\lambda \begin{vmatrix} -\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & -\lambda & 1 \\ -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} - \lambda \end{vmatrix}$$

$$+ (-1)^{n-1}(-a_0) \begin{vmatrix} 1 & 0 & \dots & 0 & 0 \\ -\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & -\lambda & 1 \end{vmatrix} .$$

Die erste Determinate auf der rechten Seite ist von derselben Form wie das betrachtete charakteristische Polynom, hat aber nur die Größe $(n-1) \times (n-1)$. Wir können daher induktiv schließen, daß sie gleich

$$(-1)^{n-1}\left(a_n\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \cdots + a_2\lambda + a_1\right)$$

ist. Die zweite Determinante läßt sich direkt ausrechnen: Wie man sich leicht überlegt (oder in **[HM1]**, Kapitel I, §4*f)* nachliest), ist die Determinante einer Dreiecksmatrix gleich dem Produkt ihrer Diagonalelemente;

die zweite Determinante ist also gleich eins. Damit folgt

$$\begin{aligned} \det(A - \lambda E) &= (-\lambda)(-1)^{n-1}\left(a_n\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \cdots + a_2\lambda + a_1\right) \\ &\quad + (-1)^{n-1}(-a_0) \\ &= (-1)^n\left(a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_2\lambda^2 + a_1\lambda + a_0\right) , \end{aligned}$$

wie behauptet. ■

Insbesondere haben also die charakteristische Gleichung einer Differentialgleichung und das charakteristische Polynom der zugehörigen Matrix nicht nur dieselben Nullstellen, sondern auch die Vielfachheiten dieser Nullstellen sind gleich; die algebraischen Vielfachheiten der Eigenwerte lassen sich direkt aus der charakteristischen Gleichung ablesen.

Damit haben alle Bausteile zusammen und können das Ergebnis dieses Abschnitts im folgenden Satz zusammenfassen:

Satz: Die charakteristische Gleichung

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0$$

der Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0y(t) = 0$$

habe die Nullstellen $\lambda_1, \dots, \lambda_r$; die Vielfachheit der Nullstelle λ_i sei α_i. Dann bilden die Lösungen der Differentialgleichung einen n-dimensionalen Vektorraum V mit Basis

$$\left\{e^{\lambda_1 t}, te^{\lambda_1 t}, \dots, t^{\alpha_1 - 1}e^{\lambda_1 t}, \quad \dots \quad e^{\lambda_r t}, te^{\lambda_r t}, \dots, t^{\alpha_r - 1}e^{\lambda_r t}\right\} .$$

Für jedes $t_0 \in \mathbb{R}$ ist die lineare Abbildung

$$V \to \mathbb{C}^n; \quad y \mapsto \left(y(t_0), \dot{y}(t_0), \ddot{y}(t_0), \dots, y^{(n-1)}(t_0)\right)$$

ein Isomorphismus; insbesondere ist jedes Anfangswertproblem

$$y(t_0) = c_0, \quad \dot{y}(t_0) = c_1, \quad \dots \quad y^{(n-1)}(t_0) = c_{n-1}$$

eindeutig lösbar. ■

Sofern wir uns für komplexwertige Lösungen interessieren, liefert uns dieser Satz alles was wir brauchen. Bei vielen Anwendungen hat man es aber mit Gleichungen mit reellen Koeffizienten zu tun, und von der Natur des Problems her interessieren nur reelle Lösungen. Mit unserem bisherigen Ansatz kommen wir auch bei diesen Problemen nicht um komplexe Lösungen herum; beispielsweise hat die wohlbekannte Differentialgleichung

$$\ddot{y}(t) + y(t) = 0$$

die charakteristische Gleichung

$$\lambda^2 + 1 = 0$$

mit den beiden rein imaginären Lösungen $\lambda = \pm i$, die oben angegebene Basis des Lösungsraums ist also

$$\left\{e^{it},\ e^{-it}\right\}.$$

Falls wir noch nie etwas vom obigen Satz gehört hätten, würden wir stattdessen sagen, daß die Lösungen dieser Differentialgleichung genau die Linearkombinationen von Sinus und Cosinus sind, Basis des Lösungsraums ist also

$$\left\{\sin t,\ \cos t\right\}.$$

Über $\mathbb{C}$ sind beide Aussagen äquivalent, denn auf Grund der EULERschen Formeln

$$\cos t = \frac{1}{2}\left(e^{it} + e^{-it}\right) \quad \text{und} \quad \sin t = \frac{1}{2i}\left(e^{it} - e^{-it}\right)$$

bzw.

$$e^{it} = \cos t + i \sin t \quad \text{und} \quad e^{-it} = \cos t - i \sin t$$

erzeugen beide Basen denselben $\mathbb{C}$-Vektorraum. Wenn wir an reellen Lösungen interessiert sind, ist aber die zweite Basis erheblich nützlicher, denn sie erzeugt auch den $\mathbb{R}$-Vektorraum der reellen Lösungen.

Entsprechend hätten wie auch für allgemeinere Differentialgleichungen mit reellen Koeffizienten gerne eine $\mathbb{C}$-Basis des (komplexen) Lösungsraums, die gleichzeitig $\mathbb{R}$-Basis des reellen Lösungsraums ist. Im Rest dieses Abschnitts wollen wir uns überlegen, wie wir diese Basis konstruieren können.

Wir gehen also aus von einer Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) = 0 \quad \text{mit} \quad a_i \in \mathbb{R}$$

und betrachten zunächst wieder die charakteristische Gleichung

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0\,.$$

Diese *kann* reelle Lösungen haben; diese bezeichnen wir mit λ_1 bis λ_p, und falls es keine gibt, setzen wir $p = 0$.

Falls p gleich der Gesamtzahl r der Nullstellen der charakteristischen Gleichung ist, sind wir fertig: Die Funktionen $t^j e^{\lambda_i t}$ sind allesamt reell, und die, bei denen j kleiner als die Vielfachheit der Nullstelle λ_i ist, spannen sowohl den komplexen als auch den reellen Lösungsraum auf.

Andernfalls gibt es auch noch nichtreelle Nullstellen. Für jede solche Nullstelle λ ist auch die konjugiert komplexe Zahl $\overline{\lambda}$ eine Nullstelle, denn da die Koeffizienten a_i reell sind, ist $\overline{a_i} = a_i$ und damit

$$\overline{\lambda}^n + a_{n-1}\overline{\lambda}^{n-1} + \ldots + a_1\overline{\lambda} + a_0 = \overline{\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0} = 0\,.$$

Nichtreelle Nullstellen treten also immer auf als Paare konjugiert komplexer Zahlen; außerdem haben λ und $\overline{\lambda}$ dieselbe Vielfachheit, denn eine Nullstelle ist genau dann eine r-fache Nullstelle eines Polynoms P, wenn sie Nullstelle von P und allen seinen Ableitungen bis zur $(r-1)$-ten ist. Da auch alle diese Ableitungen Polynome mit reellen Koeffizienten sind, zeigt die gleiche Rechnung wie oben, daß $P^{(j)}(\lambda)$ genau dann verschwindet, wenn auch $P^{(j)}(\overline{\lambda})$ verschwindet.

Wir fassen daher die nichtreellen Nullstellen zu Paaren $\left(\lambda_{p+j}, \overline{\lambda_{p+j}}\right)$ zusammen, wobei j von 1 bis zur Anzahl q dieser Paare läuft; die Gesamtzahl der Nullstellen ist also $r = p + 2q$.

Die Vielfachheit der Nullstelle λ_k sei weiterhin mit α_k bezeichnet; für $k > p$ ist das gleichzeitig auch die Vielfachheit der Nullstelle $\overline{\lambda}_k$.

Für $k > p$ wenden wir, genau wie im obigen Beispiel, die EULERschen Formeln an: Für

$$\lambda_k = \mu_k + i\nu_k \quad \text{mit} \quad \mu_k, \nu_k \in \mathbb{R}$$

ist

$$e^{\lambda_k t} = e^{\mu_k t}\big(\cos\nu_k t + i\sin\nu_k t\big) \quad\text{und}\quad e^{\overline{\lambda_k} t} = e^{\mu_k t}\big(\cos\nu_k t - i\sin\nu_k t\big)\,,$$

genauso ist für jedes j

$$t^j e^{\lambda_k t} = t^j e^{\mu_k t}\big(\cos\nu_k t + i\sin\nu_k t\big)$$

und

$$t^j e^{\overline{\lambda_k} t} = t^j e^{\mu_k t}\big(\cos\nu_k t - i\sin\nu_k t\big)\,.$$

Also spannt

$$\big\{t^j e^{\lambda_k t},\ t^j e^{\overline{\lambda_k} t}\big\}$$

denselben $\mathbb{C}$-Vektorraum auf wie

$$\big\{t^j e^{\mu_k t}\cos\nu_k t,\ t^j e^{\mu_k t}\sin\nu_k t\big\}\,,$$

und letztere Basis spannt gleichzeitig den $\mathbb{R}$-Vektorraum aller reeller Funktionen aus diesem $\mathbb{C}$-Vektorraum auf.

Damit können wir auch eine Basis des reellen Lösungsraums angeben:

Satz: Die charakteristische Gleichung

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0$$

der Differentialgleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \cdots + a_1\dot{y}(t) + a_0 y(t) = 0$$

habe die reellen Nullstellen $\lambda_1, \ldots, \lambda_p$ sowie die Paare konjugiert komplexer Nullstellen $\big(\lambda_{p+1}, \overline{\lambda_{p+1}}\big), \ldots, \big(\lambda_{p+q}, \overline{\lambda_{p+q}}\big)$ mit $\lambda_k = \mu_k + i\nu_k$; die Vielfachheit der Nullstelle λ_i sei α_i. Dann bilden die Lösungen der Differentialgleichung einen n-dimensionalen Vektorraum V, der aufgespannt wird von den Funktionen

$$t^j e^{\lambda_i t} \quad \text{für } i = 1, \ldots, p$$

und

$$t^j e^{\mu_i t}\cos\nu_i t, \quad t^j e^{\mu_i t}\sin\nu_i t \quad \text{für } i = p+1, \ldots, p+q\,,$$

wobei j jeweils von null bis $\alpha_i - 1$ läuft. Für jedes $t_0 \in \mathbb{R}$ ist die lineare Abbildung

$$V \to \mathbb{R}^n; \quad y \mapsto \big(y(t_0), \dot{y}(t_0), \ddot{y}(t_0), \ldots, y^{(n-1)}(t_0)\big)$$

ein Isomorphismus; insbesondere ist jedes Anfangswertproblem

$$y(t_0) = c_0, \quad \dot{y}(t_0) = c_1, \quad \ldots \quad y^{(n-1)}(t_0)$$

eindeutig lösbar.

■

d) Inhomogene Differentialgleichungen

Am Beispiel der Schwingungsdifferentialgleichungen haben wir gesehen, daß zur Beschreibung interessanter Phänomene homogene Differentialgleichungen oft nicht ausreichen; sobald etwa ein elektrischer Schwingkreis eine Stromquelle enthält, haben wir eine inhomogene Differentialgleichungen.

Wie schon im Falle der Schwingungsdifferentialgleichungen genügt es zur Lösung einer allgemeinen inhomogenen linearen Differentialgleichung, den Lösungsraum der zugehörigen homogenen Differentialgleichung zu kennen und dazu noch *eine* Lösung der inhomogenen, denn wegen der Linearität der linken Seite ist die Differenz zweier Lösungen der inhomogenen Differentialgleichung Lösung der homogenen.

Wir kennen bereits eine Methode, uns eine solche spezielle Lösung zu verschaffen: Sobald wir uns Anfangswerte vorgeben, können wir die entsprechende Lösung des inhomogenen Problems mittels LAPLACE-Transformation finden – sofern wir die LAPLACE-Transformierte des inhomogenen Anteils kennen und aus der LAPLACE-Transformierten der Lösungsfunktion diese Funktion selbst bestimmen können. In vielen praktisch relevanten Fällen wird dies mit Hilfe von Tabellen möglich sein, sofern man nur die grundlegenden Regeln für den Umgang mit LAPLACE-Transformationen kennt.

Eine zweite Methode ist das *Erraten* einer Lösung. In der Praxis betrachtet man Differentialgleichungen meist, um das künftige Verhalten eines Systems vorauszusagen, und oft wird man (leider nicht immer richtige!) ungefähre Erwartungen haben, wie dieses Verhalten aussehen sollte.

Falls man diese als Ansatz mit unbestimmten Parametern in die Differentialgleichung einsetzt und die Erwartungen richtig war, kann man die Parameter bestimmen und hat eine Lösung gefunden; andernfalls muß man sich etwas neues überlegen.

Als Beispiel betrachten wir nochmals die Differentialgleichung

$$\ddot{y}(t) + \rho\dot{y}(t) + \sigma y(t) = f(t) .$$

Wir kamen auf diese Differentialgleichung bei der betrachtung eines elektrischen Schwingkreises, an den eine externe Spannung proportional $f(t)$ angelegt war, und für vernünftiges $f(t)$ sollte man erwarten, daß es Lösungen gibt, bei denen sich $y(t)$ völlig von $f(t)$ bestimmen läßt und daher von ähnlicher Gestalt ist.

Im Falle einer Gleichstromquelle $f(t) = c$ vermuten wir, daß es vielleicht eine Lösung gibt, bei der nur ein Gleichstrom fließt. Der Ansatz $y(t) = a$ führt auf

$$\sigma a = c \quad \text{oder} \quad a = \frac{c}{\sigma} ,$$

falls $\sigma \neq 0$ ist.

Für $\sigma = 0$ haben wir *de facto* eine Differentialgleichung für $\dot{y}(t)$ statt für $y(t)$; man könnte es daher mit einem Ansatz der Form $\dot{y}(t) = b$ oder $y(t) = bt + a$ versuchen; für $\rho \neq 0$ führt das zu

$$\rho b = c \quad \text{oder} \quad b = \frac{c}{\rho} ;$$

a ist hier natürlich beliebig.

Entsprechend können wir auch bei einer angelegten Wechselspannung vorgehen: Der Schwingkreis wird sicherlich die Amplitude und die Phase verändern, aber die permanent angelegte Wechselspannung sollte doch dem Schwingkreis ihre Frequenz aufzwingen, so daß es zumindest eine Lösung geben sollte, die eine reine Schwingung mit dieser Frequenz ist. Für

$$\ddot{y}(t) + \rho\dot{y}(t) + \sigma y(t) = A\cos\omega_0 t$$

versuchen wir es also mit

$$y(t) = a\cos\omega_0 t + b\sin\omega_0 t$$

und berechnen die Ableitungen:

$$\dot{y}(t) = -a\omega_0\sin\omega_0 t + b\omega_0\cos\omega_0 t$$

und

$$\ddot{y}(t) = -a\omega_0^2\cos\omega_0 t - b\omega_0^2\sin\omega_0 t .$$

Einsetzen in die linke Seite der Differentialgleichung liefert

$$\begin{aligned} &\ddot{y}(t) + \rho\dot{y}(t) + \sigma y(t) \\ &= (-a\omega_0^2 + \rho b\omega_0 + \sigma a)\cos\omega_0 t + (-b\omega_0^2 - \rho a\omega_0 + \sigma b)\sin\omega_0 t , \end{aligned}$$

und das ist gleich $A\cos\omega_0 t$, falls

$$(\sigma - \omega_0^2)a + \rho\omega_0 b = A \quad \text{und} \quad (\sigma - \omega_0^2)b - \rho\omega_0 a = 0$$

ist.

Für $\sigma = \omega_0^2$ ist dies problemlos zu lösen: Dann ist einfach

$$a = 0 \quad \text{und} \quad b = \frac{A}{\rho\omega_0} .$$

Für $\sigma \neq \omega_0^2$ können wir die zweite Gleichung durch $(\sigma - \omega_0^2)$ dividieren und erhalten

$$b = \frac{\rho\omega_0}{\sigma - \omega_0^2} a .$$

Dies können wir in die erste Gleichung einsetzen:

$$(\sigma - \omega_0^2)a + \rho\omega_0\frac{\rho\omega_0}{\sigma - \omega_0^2}a = \frac{(\sigma - \omega_0^2)^2 + \rho^2\omega_0^2}{\sigma - \omega_0^2}a = A$$

oder

$$a = \frac{A(\sigma - \omega_0^2)}{(\sigma - \omega_0^2)^2 + \rho^2\omega_0^2} \quad \text{und} \quad b = \frac{A\rho\omega_0}{(\sigma - \omega_0^2)^2 + \rho^2\omega_0^2} .$$

In beiden Fällen gibt es also in der Tat eine Lösung der postulierten Form, die wir hiermit explizit bestimmt haben.

Im Prinzip haben wir diese Lösung bereits in §1*c)* berechnet, mit dem Ansatz hier ist die Situation nun aber sehr viel übersichtlicher: Im häufigsten Falle, in dem die homogene Gleichung eine gedämpfte Schwingung beschreibt etwa wissen wir nun, daß es eine feste reine Schwingung mit