hat es den Nullvektor als Lösung, der allerdings nach Definition *nie* ein Eigenvektor ist. Weitere Lösungen gibt es genau dann, wenn die Matrix $A - \lambda E$ des Gleichungssystems singulär ist, wenn also $\det(A - \lambda E)$ verschwindet.

Somit ist $\lambda \in k$ genau dann ein Eigenwert, wenn $\det(A - \lambda E) = 0$ ist; die zugehörigen Eigenvektoren sind die nichttrivialen Lösungen des homogenen linearen Gleichungssystems $(A - \lambda E)\vec{v} = \vec{0}$.

Damit ist klar, wie man Eigenwerte und Eigenvektoren berechnen kann: Man löse die Gleichung $\det(A - \lambda E) = 0$ und dann für jede Wurzel λ_i dieser Gleichung das lineare Gleichungssystem $(A - \lambda E)\vec{v} = \vec{0}$. Dieses lineare Gleichungssystem hat *nie* maximalen Rang, da es nach Definition eines Eigenwerts nichttriviale Lösungen geben muß; kommt man also auf ein eindeutig lösbares Gleichungssystem, ist das immer ein Zeichen für einen Rechenfehler.

Dies zeigt auch, daß die numerische Bestimmung von Eigenvektoren ihre Probleme haben kann: Falls man die Wurzeln der charakterischen Gleichung nur näherungsweise kennt, kann die Matrix des linearen Gleichungssystems eine leicht von null verschiedene Determinante haben, so daß es nur die triviale Lösung hat. In der Numerik werden Eigenwerte und Eigenvektoren daher simultan berechnet – sofern man beides braucht. Es gibt auch numerische Algorithmen, die nur die (angenäherten) Eigenwerte bestimmen ohne $\det(A - \lambda E)$ zu berechnen; dies ist vor allem nützlich für große n, wo die Berechnung einer Determinanten sehr aufwendig wäre. Für Genaueres sei auf die Numerikvorlesung verwiesen.

b) Ein erstes Beispiel

Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{pmatrix} .$$

Hier ist

$$A - \lambda E = \begin{pmatrix} 1-\lambda & 2 & 3 & 4 \\ 5 & 6-\lambda & 7 & 8 \\ 8 & 7 & 6-\lambda & 5 \\ 4 & 3 & 2 & 1-\lambda \end{pmatrix} ,$$

und nach dem Entwicklungssatz ist (bei Entwicklung nach der ersten Zeile)

$$\begin{aligned} \det(A - \lambda E) = (1-\lambda) & \begin{vmatrix} 6-\lambda & 7 & 8 \\ 7 & 6-\lambda & 5 \\ 3 & 2 & 1-\lambda \end{vmatrix} \\ & - 2 \begin{vmatrix} 5 & 7 & 8 \\ 8 & 6-\lambda & 5 \\ 4 & 2 & 1-\lambda \end{vmatrix} \\ & + 3 \begin{vmatrix} 5 & 6-\lambda & 8 \\ 8 & 7 & 5 \\ 4 & 3 & 1-\lambda \end{vmatrix} - 4 \begin{vmatrix} 5 & 6-\lambda & 7 \\ 8 & 7 & 6-\lambda \\ 4 & 3 & 2 \end{vmatrix} \\ & = (1-\lambda)(-\lambda^3 + 13\lambda^2 + 35\lambda) - 2(5\lambda^2 + 53\lambda) \\ & \quad + 3(-8\lambda^2 + \lambda) - 4(4\lambda^2 - 17\lambda) \\ & = \lambda^4 - 14\lambda^3 - 72\lambda^2 = \lambda^2(\lambda^2 - 14\lambda - 72) . \end{aligned}$$

Dieser Ausdruck verschwindet genau dann, wenn entweder λ verschwindet oder der Klammerausdruck ganz hinten. Letzterer ist eine quadratische Gleichung für λ, die man entweder nach der üblichen Formel lösen kann, oder aber man beachtet den Wurzelsatz von VIÈTE, wonach die Summe der beiden Lösungen 14 und ihr Produkt -72 sein muß; da $72 = 4 \times 18$ ist, können die beiden Lösungen daher nur -4 und 18 sein. Die Nullstellen des Polynoms sind also $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = -4$ und $\lambda_4 = 18$.

Die Eigenwerte von A sind damit 0, -4 und 18; für jeden dieser Werte müssen wir über ein lineares Gleichungssystem die Eigenvektoren bestimmen.

Beginnen wir mit $\lambda_1 = \lambda_2 = 0$. Hier ist das Gleichungssystem einfach

$A\vec{x} = \vec{0}$, d.h.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ 5x_1 + 6x_2 + 7x_3 + 8x_4 &= 0 \\ 8x_1 + 7x_2 + 6x_3 + 5x_4 &= 0 \\ 4x_1 + 3x_2 + 2x_3 + x_4 &= 0\,. \end{aligned}$$

Aufgrund der speziellen Struktur der Matrix A ergänzen sich die erste und die vierte Gleichung zu

$$5(x_1 + x_2 + x_3 + x_4) = 0\,;$$

entsprechend addieren sich die beiden mittleren Gleichungen zu

$$13(x_1 + x_2 + x_3 + x_4) = 0\,.$$

Da außerdem noch die Differenz zwischen den ersten beiden Gleichungen auf

$$4(x_1 + x_2 + x_3 + x_4) = 0$$

führt, genügt es, außer der ersten Gleichung noch als einzige weitere Gleichung

$$x_1 + x_2 + x_3 + x_4 = 0 \quad \text{oder} \quad x_1 = -(x_2 + x_3 + x_4)$$

zu betrachten. Subtrahiert man diese Gleichung von der ersten oder, was dasselbe ist, setzt man sie ein, so bleibt

$$x_2 + 2x_3 + 3x_4 = 0$$

als einzige Relation zwischen x_2, x_3 und x_4. Das gegebene Gleichungssystem ist also äquivalent zu

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ x_2 + 2x_3 + 3x_4 &= 0\,; \end{aligned}$$

insbesondere hat es also den Rang zwei und damit einen Lösungsraum der Dimension $4 - 2 = 2$.

(Es ist klar, daß der Rang des Gleichungssystems kleiner als vier sein muß, da es nichttriviale Lösungen gibt. Bei der Berechnung von Eigenvektoren über lineare Gleichungssysteme gibt es also immer Abhängigkeiten zwischen den einzelnen Gleichungen.)

Im noch verbliebenen Gleichungssystem kann beispielsweise x_4 beliebig gewählt werden. Ist $x_4 = 0$, so bleibt noch die Relation $x_2 + 2x_3 = 0$ übrig, in der etwa x_3 beliebig gewählt werden kann; danach sind $x_2 = -2x_3$ und $x_1 = x_3$ eindeutig festgelegt; die entsprechenden Lösungen sind also die Vielfachen der Lösung $(1, -2, 1, 0)$.

Genausogut könnte etwa $x_3 = 0$ gesetzt werden; dann wäre $x_2 = -3x_4$ und $x_1 = 2x_4$; dies führt auf die Vielfachen der Lösung $(2, -3, 0, 1)$.

Damit haben wir zwei linear unabhängige Lösungen gefunden; da der Lösungsraum zweidimensional ist, sind die sämtlichen Lösungen des Gleichungssystems die Linearkombinationen

$$x_1 = -\lambda - 2\mu, \quad x_2 = 2\lambda - 3\mu, \quad x_3 = \lambda \quad \text{und} \quad x_4 = \mu$$

dieser beiden Lösungen; die beiden Vektoren

$$\begin{pmatrix} -1 \\ 2 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

bilden also eine Basis des Eigenraums von A zum Eigenwert Null.

Genauso lassen sich auch die Eigenwerte -4 und 18 behandeln: -4 führt auf das lineare Gleichungssystem

$$\begin{aligned} 5x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ 5x_1 + 10x_2 + 7x_3 + 8x_4 &= 0 \\ 8x_1 + 7x_2 + 10x_3 + 5x_4 &= 0 \\ 4x_1 + 3x_2 + 2x_3 + 5x_1 &= 0\,, \end{aligned}$$

dessen sämtliche Lösungen die Vielfachen von $(-1, -1, 1, 1)$ sind, und 18 führt auf

$$\begin{aligned} -17x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \\ 5x_1 - 12x_2 + 7x_3 + 8x_4 &= 0 \\ 8x_1 + 7x_2 - 12x_3 + 5x_4 &= 0 \\ 4x_1 + 3x_2 + 2x_3 - 17x_4 &= 0 \end{aligned}$$

mit den Vielfachen von $(5, 13, 13, 5)$ als Lösungsraum.

Was haben wir nun erreicht? Betrachten wir die lineare Abbildung

$$\varphi\colon \mathbb{R}^4 \to \mathbb{R}^4; \quad \vec{v} \mapsto A \cdot \vec{v}\,.$$

Für die Vektoren

$$\vec{b}_1 = \lambda \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix},\ \vec{b}_2 = \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix},\ \vec{b}_3 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \text{ und } \vec{b}_4 = \begin{pmatrix} 5 \\ 13 \\ 13 \\ 5 \end{pmatrix}$$

ist

$$\varphi(\vec{b}_1) = 0\,\vec{b}_1, \quad \varphi(\vec{b}_2) = 0\,\vec{b}_2, \quad \varphi(\vec{b}_3) = -4\,\vec{b}_3 \quad \text{und} \quad \varphi(\vec{b}_4) = 18\,\vec{b}_4\,;$$

bezüglich der neuen Basis $\{\vec{b}_1, \vec{b}_2, \vec{b}_3, \vec{b}_4\}$ hat φ daher die Abbildungsmatrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 18 \end{pmatrix},$$

bei der die Eigenwerte von A in der Hauptdiagonalen stehen und alle sonstigen Einträge verschwinden. Mit dieser Matrix läßt sich sehr viel angenehmer rechnen als mit der ursprünglichen Matrix A; insbesondere kann ihre Exponentialfunktion

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-4} & 0 \\ 0 & 0 & 0 & e^{18} \end{pmatrix}$$

sehr einfach berechnet werden.

Damit läßt sich auch e^A berechnen: Sind nämlich

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

die vier Einheitsvektoren des $\mathbb{R}^4$, so ist

$$\vec{b}_i = B\vec{e}_i \quad \text{mit} \quad B = \begin{pmatrix} 1 & 2 & -1 & 5 \\ -2 & -3 & -1 & 13 \\ 1 & 0 & 1 & 13 \\ 0 & 1 & 1 & 5 \end{pmatrix},$$

und die Gleichung

$$A\vec{b}_i = \lambda_i \vec{b}_i$$

mit $\lambda_{1/2} = 0, \lambda_3 = -4$ und $\lambda_4 = 18$ wird zu

$$AB\vec{e}_i = \lambda_i B\vec{b}_i \quad \text{oder} \quad B^{-1}AB\,\vec{e}_i = \lambda_i \vec{e}_i\,.$$

Also ist

$$B^{-1}AB = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 18 \end{pmatrix}$$

oder

$$A = B \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 18 \end{pmatrix} B^{-1}\,.$$

Für eine beliebige Matrix M ist

$$e^{BMB^{-1}} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(BMB^{-1}\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!} BM^n B^{-1} = Be^M B^{-1}\,,$$

also ist

$$e^A = B \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-4} & 0 \\ 0 & 0 & 0 & e^{18} \end{pmatrix} B^{-1}\,,$$

was sich zumindest im Prinzip berechnen läßt – auch wenn das Ergebnis

$$\begin{pmatrix} 27e^{-4}+35+10e^{18} & 10e^{18}-19+9e^{-4} & 10e^{18}-1-9e^{-4} & 10e^{18}+17-27e^{-4} \\ -53+27e^{-4}+26e^{18} & 9e^{-4}+26e^{18}+37 & 26e^{18}-17-9e^{-4} & 26e^{18}+1-27e^{-4} \\ 26e^{18}+1-27e^{-4} & 26e^{18}-17-9e^{-4} & 9e^{-4}+26e^{18}+37 & -53+27e^{-4}+26e^{18} \\ 10e^{18}+17-27e^{-4} & 10e^{18}-1-9e^{-4} & 10e^{18}-19+9e^{-4} & 27e^{-4}+35+10e^{18} \end{pmatrix}$$

alles andere als angenehm ist. Dies zeigt wieder einmal, wieviel man sich ersparen kann, wenn man *vor* Beginn einer Rechnung eine gute Basis *bzw.* ein gutes Koordinatensystem wählt.

c) Das charakteristische Polynom einer Matrix

Es ist kein Zufall, daß im obigen Beispiel die Gleichung $\det(A-\lambda E)=0$ auf ein Polynom vierten Grades führte: Ist $A=(a_{ij})$ eine $n\times n$-Matrix, so hat die Matrix $A-\lambda E$ in der Diagonalen die Einträge $a_{ii}-\lambda$, ansonsten stimmen alle Einträge mit denen von A überein. Berechnet man daher $\det(A-\lambda E)$ gemäß der definierenden Formel, so gibt es genau ein Produkt, in dem n mit λ behaftete Faktoren vorkommen, nämlich das Produkt

$$(a_{11}-\lambda)\cdots(a_{nn}-\lambda)=(-1)^n\lambda^n+\textit{Terme niedrigerer Ordnung}$$

der Diagonaleinträge. Die restlichen Produkte, die zur Determinante aufsummiert werden, enthalten zwischen null und $n-1$ mit λ behaftete Faktoren, die Summe ist also ein Polynom vom Grad n mit höchstem Term $(-1)^n\lambda^n$.

Definition: Das Polynom $\det(A-\lambda E)$ heißt *charakteristisches Polynom* der Matrix A.

Demgemäß sind also die Eigenwerte von A gleich den Nullstellen des charakteristischen Polynoms von A.

Ist x eine Nullstelle eines Polynoms $f(X)$, so kann $f(X)$ bekanntlich durch $(X-x)$ geteilt werden, und x heißt r-*fache Nullstelle* von $f(X)$, wenn $f(X)$ durch $(X-x)^r$ teilbar ist, nicht aber durch $(X-x)^{r+1}$.

Definition: Wir sagen, der Eigenwert λ von φ *bzw.* A habe die *algebraische Vielfachheit* r, wenn λ eine r-fache Nullstelle des charakteristischen Polynoms ist.

Im obigen Beispiel hatte also der Eigenwert Null die algebraische Vielfachheit zwei, die anderen beiden hatten algebraische Vielfachheit eins. Die Dimension des jeweiligen Eigenraums, die geometrische Vielfachheit also, war genauso groß, jedoch muß dies im allgemeinen nicht der Fall sein: Für die Matrix

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

etwa hat das charakteristische Polynom

$$\det(A-\lambda E)=\begin{vmatrix}1-\lambda&1\\0&1-\lambda\end{vmatrix}=(1-\lambda)^2$$

die doppelte Nullstelle eins, $\lambda=1$ ist also ein Eigenwert mit algebraischer Vielfachheit zwei. Der zugehörige Eigenraum ist die Lösungsmenge des linearen Gleichungssystems

$$0x_1+1x_2=0$$
$$0x_1+0x_2=0\,,$$

also gerade die Menge aller Vektoren der Form $\binom{x}{0}$ und somit eindimensional. Die geometrische Vielfachheit des Eigenwerts eins ist daher nur eins.

Das Beispiel der Abbildung

$$\varphi\colon\mathbb{R}^2\to\mathbb{R}^2;\qquad\begin{pmatrix}x\\y\end{pmatrix}\mapsto\begin{pmatrix}x\cos\vartheta-y\sin\vartheta\\y\cos\vartheta+x\sin\vartheta\end{pmatrix}$$

mit Abbildungsmatrix

$$A=\begin{pmatrix}\cos\vartheta&-\sin\vartheta\\\sin\vartheta&\cos\vartheta\end{pmatrix}\in\mathbb{R}^{2\times2}$$

zeigt, daß es überhaupt keine Eigenwerte geben muß, denn hier ist das charakteristische Polynom gleich

$$\begin{vmatrix}\cos\vartheta-\lambda&-\sin\vartheta\\\sin\vartheta&\cos\vartheta-\lambda\end{vmatrix}=(\cos\vartheta-\lambda)^2+\sin^2\vartheta\,,$$

und abgesehen vom Fall $\sin\vartheta=0$, wenn A gleich der positiven oder negativen Einheitsmatrix ist, hat dieses Polynom keine reelle Nullstelle, da es nur positive Werte annimmt. Es hat aber natürlich die beiden komplexen Nullstellen

$$\lambda_{1/2}=\cos\vartheta\pm i\sin\vartheta=e^{\pm i\vartheta}\;;$$

fassen wir φ als Abbildung von $\mathbb{C}^2$ nach $\mathbb{C}^2$ auf, gibt es also zwei Eigenwerte. Beide haben die algebraische und geometrische Vielfachheit eins; zugehörige Eigenvektoren sind etwa $\binom{1}{i}$ und $\binom{1}{-i}$. Wählen wir diese beiden Vektoren als Basis, so wird die Abbildungsmatrix von φ bezüglich dieser neuen Basis zur Diagonalmatrix

$$\begin{pmatrix} e^{i\vartheta} & 0 \\ 0 & e^{-i\vartheta} \end{pmatrix} .$$

Allgemein gilt für die algebraischen und geometrischen Vielfachheiten von Eigenvektoren

Satz: *a)* Die geometrische Vielfachheit eines Eigenwerts ist stets kleiner oder gleich der algebraischen Vielfachheit.
b) Die Summe der algebraischen Vielfachheiten der verschiedenen Eigenwerte einer linearen Abbildung ist kleiner oder gleich der Dimension des Vektorraums.

Beweis: a) Der Eigenwert λ der $n \times n$-Matrix A habe die geometrische Vielfachheit r, d.h. der zugehörige Eigenraum habe die Dimension r. Wir wählen eine Basis $\vec{b}_1, \ldots, \vec{b}_r$ dieses Eigenraums und ergänzen sie zu einer Basis des gesamten Vektorraums; bezüglich dieser Basis sei C die Abbildungsmatrix der linearen Abbildung

$$\varphi : \begin{cases} k^n \to k^n \\ \vec{v} \mapsto A\vec{v} \end{cases} .$$

Da $\vec{b}_1, \ldots, \vec{b}_r$ Eigenvektoren zum Eigenwert λ sind, ist $\varphi(\vec{b}_i) = \lambda \vec{b}_i$. In den ersten r Spalten von C steht also jeweils in der Diagonalen das Element λ und ansonsten überall die Null. A hat somit die Form

$$A = \begin{pmatrix} \lambda & 0 & \ldots & 0 & * & \ldots & * \\ 0 & \lambda & \ldots & 0 & * & \ldots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda & * & \ldots & * \\ 0 & 0 & \ldots & 0 & & & \\ \vdots & \vdots & \ddots & \vdots & & \mathsf{M} & \\ 0 & 0 & \ldots & 0 & & & \end{pmatrix} ,$$

wobei uns weder die mit $*$ bezeichneten Körperelemente noch die $(n-r) \times (n-r)$-Matrix M weiter zu interessieren brauchen.

Für $C - xE$ gilt dasselbe, nur daß jetzt $\lambda - x$ in der Diagonalen steht, d.h. diese Matrix hat die Form

$$\begin{pmatrix} \lambda - x & 0 & \ldots & 0 & * & \ldots & * \\ 0 & \lambda - x & \ldots & 0 & * & \ldots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda - x & * & \ldots & * \\ 0 & 0 & \ldots & 0 & & & \\ \vdots & \vdots & \ddots & \vdots & & \mathsf{M} - \mathsf{x}\mathsf{E}_{n-r} & \\ 0 & 0 & \ldots & 0 & & & \end{pmatrix} ,$$

wobei E_{n-r} die $(n-r) \times (n-r)$-Einheitsmatrix bezeichnet.

Zur Berechnung ihrer Determinanten verwenden wir den LAPLACEschen Entwicklungssatz: Da in der ersten Zeile (oder Spalte) nur an der ersten Stelle ein von Null verschiedener Eintrag steht, ist diese Determinante gleich $(\lambda - x)$ mal der Determinante jener Matrix, die durch Streichen der ersten Zeile und Spalte entsteht. Falls $r > 1$ ist, hat diese neue Matrix dieselbe Form, wir können den LAPLACEschen Entwicklungssatz also noch einmal anwenden *usw.;* wir erhalten schließlich

$$\det(C - xE) = (\lambda - x)^r \det(M - xE_{n-r}) .$$

Somit ist $\det(C - xE)$ durch $(x - \lambda)^r$ teilbar.

Was uns wirklich interessiert, ist aber nicht $\det(C - xE)$, sondern $\det(A - xE)$. Ist B die Matrix des Basiswechsels von der Standardbasis des k^n auf die Basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$, jene Matrix also, deren Spaltenvektoren die $\vec{b}_i$ sind, so ist $C = B^{-1}AB$ und

$$\begin{aligned} \det(C - xE) &= \det(B^{-1}AB - xE) = \det(B^{-1}AB - xB^{-1}EB) \\ &= \det\bigl(B(A - xE)B^{-1}\bigr) = \det B \det(A - xE)(\det B)^{-1} \\ &= \det(A - xE) . \end{aligned}$$

A und C haben also dasselbe charakteristische Polynom, und somit ist auch das charakteristische Polynom von A durch $(x - \lambda)^r$ teilbar. Die algebraische Vielfachheit von λ ist daher mindestens r.

Unabhängig von diesem Ergebnis wollen wir noch festhalten, daß nach der gerade durchgeführten Rechnung für eine beliebige Matrix A und eine invertierbare Matrix B die beiden Matrizen A und BAB^{-1} dasselbe charakteristische Polynom haben; insbesondere haben also die Abbildungsmatrizen einer linearen Abbildung zu verschiedenen Basen dasselbe charakteristische Polynom.

b) Sind $\lambda_1, \ldots, \lambda_\ell$ die verschiedenen Eigenwerte von φ und sind $r_1, \ldots, r_\ell$ ihre algebraischen Vielfachheiten, so ist das charakteristische Polynom $\det(A - xE)$ teilbar durch

$$(x-\lambda_1)^{r_1} \cdots (x-\lambda_\ell)^{r_\ell} .$$

Dies ist ein Polynom vom Grad $r_1 + \cdots + r_\ell$, wohingegen das charakteristische Polynom Grad n hat; daher ist

$$r_1 + \cdots + r_\ell \le n ,$$

denn der Grad eines Teilers kann nicht größer sein als der des Polynoms selbst. ■

Zum Abschluß dieses Abschnitts sei noch ein Kriterium angegeben, wann es für eine lineare Abbildung φ eine Basis aus Eigenwerten gibt, wann also die Abbildungsmatrix bezüglich einer geeigneten Basis Diagonalgestalt hat:

Satz: Zur linearen Abbildung $\varphi: V \to V$ eines n-dimensionalen Vektorraums gibt es genau dann eine Basis aus Eigenvektoren von φ, wenn

1.) das charakteristische Polynom von φ als Produkt von Linearfaktoren geschrieben werden kann

2.) die geometrische Vielfachheit eines jeden Eigenwerts gleich der algebraischen ist.

Beweis: Zunächst sei $\varphi: V \to V$ eine lineare Abbildung derart, daß V eine Basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$ aus Eigenvektoren von φ habe. Wir müssen zeigen, daß *1.)* und *2.)* erfüllt sind.

Da die Basisvektoren $\vec{b}_i$ Eigenvektoren sind, gibt es zu jedem $\vec{b}_i$ ein Körperelement λ_i, so daß $\varphi(\vec{b}_i) = \lambda\vec{b}_i$ ist; bezüglich dieser Basis hat die

Abbildungsmatrix A von φ daher Diagonalgestalt, und das charakteristische Polynom

$$\det(A - \lambda E) = \begin{vmatrix} \lambda_1 - \lambda & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n - \lambda \end{vmatrix} = \prod_{i=1}^{n} (\lambda_i - \lambda)$$

zerfällt in der Tat in Linearfaktoren. Die algebraische Vielfachheit des Eigenwerts λ_i ist gleich der Anzahl jeder Indizes $j \in \{1, \ldots, n\}$, für die $\lambda_j = \lambda_i$ ist; dies ist auch die geometrische Vielfachheit, denn der Eigenraum wird aufgespannt von den Vektoren $\vec{b}_j$ zu diesen j. Also sind *1)* und *2)* erfüllt.

Umgekehrt erfülle die Abbildung φ die Bedingungen *1)* und *2);* wir müssen zeigen, daß es eine Basis aus Eigenvektoren von φ gibt.

Wegen *1)* läßt sich das charakteristische Polynom in der Form

$$(\lambda_1 - \lambda)^{r_1} \cdot \cdots \cdot (\lambda_s - \lambda)^{r_s}$$

schreiben, wobei wir annehmen können, daß die λ_i paarweise verschieden sind. Dann ist r_i die algebraische Vielfachheit von λ_i. Da das charakteristische Polynom den Grad n hat, folgt, daß

$$r_1 + \cdots + r_s = n$$

ist. Außerdem gibt es wegen *2.)* zu jedem λ_i einen r_i-dimensionalen Eigenraum, also r_i linear unabhängige Eigenvektoren. Da Eigenvektoren zu verschiedenen Eigenwerten nach dem Lemma vom Anfang dieses Abschnitts stets linear unabhängig sind, ist auch das System all dieser Eigenvektoren linear unabhängig und somit eine Basis, denn es besteht aus $n = \dim V$ Vektoren. Damit ist eine Basis aus Eigenvektoren von φ gefunden. ■

d) Eigenwerte symmetrischer und Hermitescher Matrizen

Wie wir im letzten Paragraphen gesehen haben, kann die geometrische Vielfachheit eines Eigenwerts kleiner sein als die algebraische, und im Falle einer reellen Matrix müssen nicht auch die Eigenwerte reell sein. In diesem Abschnitt wollen wir sehen, daß solche Dinge bei symmetrischen (und auch den noch zu definierenden HERMITEschen) Matrizen nicht möglich sind.

Symmetrische und HERMITEsche Matrizen hängen eng mit (HERMITEschen) Skalarprodukten zusammen: Ist V ein endlichdimensionaler EUKLIDischer Vektorraum und $\{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis von V, so gilt für zwei Vektoren

$$\vec{v} = \sum_{i=1}^{n} v_i \vec{b}_i \quad \text{und} \quad \vec{w} = \sum_{i=1}^{n} w_i \vec{b}_i$$

wegen der Linearität des Skalarprodukts in beiden Argumenten

$$\vec{v} \cdot \vec{w} = \left(\sum_{i=1}^{n} v_i \vec{b}_i \right) \cdot \left(\sum_{j=1}^{n} w_j \vec{b}_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i w_j \, \vec{b}_i \cdot \vec{b}_j \, .$$

Setzen wir

$$c_{ij} \underset{\text{def}}{=} \vec{b}_i \cdot \vec{b}_j \, ,$$

so ist wegen der Symmetrie des Skalarprodukts $c_{ij} = c_{ji}$, wir haben also eine symmetrische $n \times n$-Matrix C.

Die Matrix C legt das Skalarprodukt eindeutig fest, denn für zwei beliebige Vektoren $\vec{v}, \vec{w}$ wie oben ist

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i w_j \cdot c_{ij} \, .$$

Diese Formel definiert umgekehrt auch für jede symmetrische Matrix $C \in \mathbb{R}^{n \times n}$ eine bilineare Abbildung $V \times V \to \mathbb{R}$, allerdings muß diese nicht positiv definit und damit kein Skalarprodukt sein.

Ist V ein HERMITEscher Vektorraum, wieder mit Basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$, so ist jetzt für zwei Vektoren

$$\vec{v} = \sum_{i=1}^{n} v_i \vec{b}_i \quad \text{und} \quad \vec{w} = \sum_{i=1}^{n} w_i \vec{b}_i \quad \text{mit} \quad v_i, w_i \in \mathbb{C}$$

$$\vec{v} \cdot \vec{w} = \left(\sum_{i=1}^{n} v_i \vec{b}_i \right) \cdot \left(\sum_{j=1}^{n} w_j \vec{b}_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} v_i \overline{w}_j \, \vec{b}_i \cdot \vec{b}_j \, .$$

Setzen wir auch hier wieder

$$c_{ij} \underset{\text{def}}{=} \vec{b}_i \cdot \vec{b}_j \, ,$$

so ist nun $c_{ij} = \overline{c_{ji}}$. Matrizen mit dieser Eigenschaft wollen wir als HERMITEsch bezeichnen.

Um dies etwas kompakter ausdrücken zu können, definieren wir

Definition: *a)* Für eine Matrix $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ bezeichnen wir die Matrix $\overline{A} = \left(\overline{a_{ij}}\right)$ als die zu A *konjugiert komplexe Matrix.*
b) $A \in \mathbb{C}^{n \times n}$ heißt HERMITEsch, falls ${}^tA = \overline{A}$ ist.
c) Zu einem Vektor $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ heißt $\overline{\vec{v}} = \begin{pmatrix} \overline{v_1} \\ \vdots \\ \overline{v_n} \end{pmatrix}$ der *konjugiert komplexe* Vektor.

(Letztere Schreibweise sieht zwar grausam aus, vor allem wann man sie mit der sehr viel schöneren Alternative $\overline{\mathfrak{v}}$ vergleicht; sie läßt sich aber nicht vermeiden, wenn man Vektoren mit Pfeilen kennzeichnen will.)

Schließlich wollen wir Vektoren hier mit $1 \times n$-Matrizen identifizieren; insbesondere rechnen wir mit dem „transponierten Vektor"

$${}^t\vec{v} = (v_1, \ldots, v_n) \, .$$

Mit dieser Bezeichnung kann das Standardskalarprodukt zweier Vektoren $\vec{v}, \vec{w} \in \mathbb{R}^n$ als Matrixprodukt ${}^t\vec{v}\vec{w}$ geschrieben werden; das Standard-HERMITEsche Produkt in $\mathbb{C}^n$ ist entsprechend ${}^t\vec{v}\,\overline{\vec{w}}$.

Da die komplexe Konjugation auf $\mathbb{R}$ keine Wirkung hat, ist eine HERMITEsche Matrix mit reellen Einträgen einfach eine symmetrische Matrix; wir können uns im folgenden bei den Beweisen daher auf HERMITEsche Matrizen beschränken und erhalten trotzdem Ergebnisse, die auch für reelle symmetrische Matrizen gelten.

Das Hauptziel dieses Abschnitts ist

Satz: A sei eine symmetrische reelle oder HERMITEsche (komplexe) Matrix.
a) Dann sind alle Eigenwerte von A reell.

b) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal bezüglich des Standard- *bzw.* HERMITEschen Skalarprodukts.
c) Für jeden Eigenwert von A ist die geometrische Vielfachheit gleich der algebraischen Vielfachheit.
d) $\mathbb{R}^n$ *bzw.* $\mathbb{C}^n$ hat eine Orthonormalbasis aus Eigenvektoren von A.

Beweis: a) Ist $\lambda \in \mathbb{C}$ ein Eigenwert von A, so gibt es nach Definition einen Vektor $\vec{v} \neq \vec{0}$, so daß $A\vec{v} = \lambda\vec{v}$ ist. Da die komplexe Konjugation mit sämtlichen Grundrechenarten vertauschbar ist, folgt, daß

$$\overline{A}\,\overline{\vec{v}} = \overline{\lambda}\,\overline{\vec{v}}\,,$$

d.h.

$${}^t\vec{v}\overline{A}\,\overline{\vec{v}} = \overline{\lambda}\,{}^t\vec{v}\overline{\vec{v}}\,.$$

Andererseits ist ${}^t(A\vec{v}) = {}^t\vec{v}\;{}^tA$, und da A als HERMITEsch vorausgesetzt wird, ist ${}^tA = \overline{A}$. Also ist ${}^t\vec{v}\,\overline{A} = \lambda\,{}^t\vec{v}$ und

$${}^t\vec{v}\,\overline{A}\,\overline{\vec{v}} = \lambda\,{}^t\vec{v}\overline{\vec{v}}\,.$$

Somit ist

$$\lambda\,{}^t\vec{v}\overline{\vec{v}} = \overline{\lambda}\,{}^t\vec{v}\overline{\vec{v}}\,,$$

und da ${}^t\vec{v}\overline{\vec{v}}$ für einen Vektor $\vec{v} \neq \vec{0}$ nicht verschwindet, folgt $\lambda = \overline{\lambda}$, d.h. λ ist reell.

b) $\vec{v}$ sei Eigenvektor zum Eigenwert λ, und $\vec{w}$ sei Eigenvektor zum davon verschiedenen Eigenwert μ, d.h.

$$A\vec{v} = \lambda\vec{v} \quad \text{und} \quad A\vec{w} = \mu\vec{w}\,.$$

Dann ist

$$\lambda\,{}^t\vec{v}\overline{\vec{w}} = {}^t(\lambda\vec{v})\overline{\vec{w}} = {}^t(A\vec{v})\overline{\vec{w}} = {}^t\vec{v}\;{}^tA\overline{\vec{w}} = {}^t\vec{v}\overline{A}\,\overline{\vec{w}} = {}^t\vec{v}\overline{A\vec{w}} = {}^t\vec{v}\overline{\mu\vec{w}} = \overline{\mu}\,{}^t\vec{v}\overline{\vec{w}}\,.$$

Da wir schon wissen, daß alle Eigenwerte reell sind, ist $\overline{\mu} = \mu \neq \lambda$. Die obige Gleichungskette kann daher nur richtig sein, wenn ${}^t\vec{v}\overline{\vec{w}}$ verschwindet, d.h. wenn $\vec{v}$ und $\vec{w}$ orthogonal sind.

Beim Beweis von *c)* gehen wir im wesentlichen genauso vor wie im vorigen Abschnitt, als wir zeigten, daß die geometrische Vielfachheit eines Eigenwerts stets kleiner oder gleich der algebraischen ist; die zusätzliche Annahme über die Matrix A wird zeigen, daß hier die beiden Vielfachheiten sogar gleich sind.

λ sei also ein Eigenwert von A mit geometrischer Vielfachheit r, d.h. der zugehörige Eigenraum habe die Dimension r. Wir wählen eine Basis $\{\vec{b}_1, \dots, \vec{b}_r\}$ und ergänzen diese zu einer Basis $\mathcal{B} = \{\vec{b}_1, \dots, \vec{b}_n\}$ des gesamten Vektorraums $V = \mathbb{R}^n$ oder $\mathbb{C}^n$. Da wir es hier nur mit diesen beiden speziellen Vektorräumen zu tun haben, können wir, indem wir nötigenfalls das GRAM-SCHMIDTsche Orthogonalisierungsverfahren anwenden und anschließend die Längen aller Vektoren auf eins normieren, annehmen, daß es sich dabei um eine Orthonormalbasis handelt.

Nun betrachten wir die lineare Abbildung

$$\varphi\colon V \to V; \quad \vec{v} \mapsto A\vec{v}\,.$$

Bezüglich der Standardbasis hat sie natürlich A als Abbildungsmatrix; für uns interessanter ist aber die Abbildungsmatrix C bezüglich der neuen Basis $\mathcal{B}$. Dazu sei B die Matrix, deren Spalten die Vektoren $\vec{b}_i$ sind; da der Eintrag an der Stelle (i,j) eines Matrixprodukts das (Standard-)Skalarprodukt des i-ten Zeilenvektors des ersten Faktors mit dem j-ten Spaltenvektor des zweiten Faktors ist, steht an der Stelle (i,j) der Matrix ${}^tB\overline{B}$ das (Standard) HERMITEsche Produkt der Vektoren $\vec{b}_i$ und $\vec{b}_j$. Da $\mathcal{B}$ als Orthonormalbasis gewählt wurde, ist daher

$${}^tB\overline{B} = E \quad \text{und damit} \quad {}^tB = \overline{B}^{\,-1} = \overline{B^{-1}}\,.$$

Aus dieser Formel folgt, daß mit A auch C eine HERMITEsche Matrix ist, denn

$$C = {}^t(B^{-1}AB) = {}^tB\;{}^tA\;{}^tB^{-1} = \overline{B^{-1}}\,\overline{A}\,\overline{B} = \overline{B^{-1}AB} = \overline{C}\,.$$

Die ersten r Basisvektoren $\vec{b}_i$ sind Eigenvektoren von A zum Eigenwert λ; für $i \leq r$ ist daher $\varphi(\vec{b}_i) = \lambda\vec{b}_i$, d.h. in der i-ten Spalte von C steht an der i-ten Stelle die reelle Zahl λ und ansonsten überall die Null, genau wie auch im vorigen Abschnitt. Im Gegensatz zu dort haben wir nun aber eine HERMITEsche Matrix; da in der i-ten Spalte abgesehen von λ auf der Hauptdiagonalen nur Nullen stehen, muß daher dasselbe auch für die i-te Zeile gelten; die Matrix C hat also die Form

$$C = \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & & & \\ \vdots & \vdots & \ddots & \vdots & & M & \\ 0 & 0 & \dots & 0 & & & \end{pmatrix},$$

wobei M eine $(n-r)\times(n-r)$-Matrix ist, die uns nicht weiter zu interessieren braucht. Damit hat $C - xE$ die Form

$$\begin{pmatrix} \lambda - x & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \lambda - x & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda - x & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & & & \\ \vdots & \vdots & \ddots & \vdots & & M - xE_{n-r} & \\ 0 & 0 & \dots & 0 & & & \end{pmatrix},$$

wobei E_{n-r} die $(n-r)\times(n-r)$-Einheitsmatrix bezeichnet.

Wie wir uns schon im vorigen Abschnitt überlegten beim Beweis, daß die geometrische Vielfachheit eines Eigenwerts immer kleiner oder gleich der algebraischen ist, haben A und A_B dasselbe charakterische Polynom; da wir die Matrix C besser kennen, rechnen wir mit ihr.

Wie in Abschnitt *c)* folgt auf Grund der obigen Form der Matrix $C - xE$ aus dem LAPLACEschen Entwicklungssatz, daß

$$\det(A - xE) = \det(C - xE) = (\lambda - x)^r \det(M - xE_{n-r})$$

ist, wobei E_{n-r} die $(n-r)\times(n-r)$-Einheitsmatrix bezeichnet. Wir müssen zeigen, daß die algebraische Vielfachheit von λ *genau* gleich r ist, daß also λ keine Nullstelle von $\det(M - xE_{n-r})$ sein kann.

Wäre λ Nullstelle von $\det(M - xE_{n-r})$, so hätte M den Eigenwert λ, es gäbe also einen $(n-r)$-dimensionalen Eigenvektor $\vec{w}$ von M. Wegen der speziellen Form der Matrix A_B ist für jeden Eigenvektor

$$\vec{w} = \begin{pmatrix} w_{r-1} \\ \vdots \\ w_n \end{pmatrix} \quad \text{von } M \text{ der Vektor} \quad \vec{v} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ w_{r-1} \\ \vdots \\ w_n \end{pmatrix}$$

ein Eigenvektor von C und damit von A – Eigenvektoren hängen schließlich nur von der linearen Abbildung ab, nicht von einer speziellen Abbildungsmatrix. Dies widerspricht aber der Voraussetzung, daß der Eigenraum zum Eigenwert λ von $\vec{b}_1, \dots, \vec{b}_r$ erzeugt wird, denn $\vec{v}$ ist linear unabhängig von diesen $\vec{b}_i$.

Also hat λ die algebraische Vielfachheit r, und *c)* ist gezeigt.

d) ist nun eine einfache Folgerung aus den übrigen Aussagen – wenn wir voraussetzen, daß jedes Polynom über den komplexen Zahlen in Linearfaktoren zerfällt; das ist der sogenannte *Fundamentalsatz der Algebra.*

Wir wissen dann, daß die Summe der algebraischen Vielfachheiten aller Eigenwerte gleich der Dimension n des Vektorraums ist und daß alle Eigenwerte reell sind; da die algebraischen gleich den geometrischen Vielfachheiten sind, gibt es also n Eigenvektoren, die eine Basis von V bilden.

Für jeden einzelnen Eigenraum können wir die Eigenvektoren nach GRAM-SCHMIDT so wählen, daß sie eine Orthonormalbasis bilden; da Eigenvektoren zu verschiedenen Eigenwerten stets orthogonal sind, ist die Vereinigungsmenge dieser Basen Orthonormalbasis von V. ∎

e) 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 cweiner 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 läßt sich – wie wir im vorigen Abschnitt anhand eines Beispiels gesehen haben – dadurch lösen, daß wir zu einem größeren Körper übergehen, also etwa von den reellen zu den komplexen Zahlen.

Ernster ist das erste Problem; dazu wollen wir uns in diesem Abschnitt überlegen, wie man, wenn schon keine Diagonalgestalt, doch wenigstens eine möglichst einfache Gestalt der Abbildungsmatrix finden kann.

Eine wesentliches Hilfsmittel dazu sind invariante Unterräume:

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

Offensichtlich sind die Eigenräume von φ invariante Unterräume im Sinne dieser Definition; wir werden uns, für den Fall, daß diese nicht ausreichen, im Laufe dieses Abschnitts noch weitere Beispiele konstruieren. Da es hier vor allem um Matrizen gehen soll, wollen wir für den Rest dieses Abschnitts annehmen, daß V endliche Dimension habe.

Ist U ein φ-invarianter Unterraum, so können wir eine Basis $\{\vec b_1, \dots, \vec b_r\}$ von U 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 $\varphi_{|U}\colon U \to U$, 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 $U \cap W = \{\vec 0\}$. (Man sagt dann, $V = U \oplus W$ sei *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

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

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 $\ell \in \mathbb{N}$ gibt, so daß $(\varphi - \lambda\,\mathrm{id})^\ell(\vec{v}) = \vec{0}$ ist. Das kleinste solche ℓ heißt *Stufe* des Hauptvektors.

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

Beispielsweise sind also die Hauptvektoren der Stufe eins abgesehen vom Nullvektor gerade die Eigenvektoren zum Eigenwert λ.

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 ansteigen, es gibt also ein maximales ℓ, das als Stufe eines Hauptvektors auftreten kann, und 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 $j-1$. Da mit $\vec{v}$ auch $\lambda\vec{v}$ in H_λ liegt, ist damit auch $\varphi(\vec{v})$ 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 $\big((\varphi - \lambda\,\mathrm{id})^\ell\big)$, 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 auch wieder die Behauptung.

Als nächstes wollen wir zeigen, daß der Durchschnitt der beiden Summanden nur aus dem Nullvektor besteht. Dazu sei $\vec{v}$ im Durchschnitt der beiden Räume. 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 aber auch $\vec{w}$ 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