19. Mai 2000

2 Ähnlichkeit von Matrizen, Eigenwerte und Eigenvektoren

Motivation. Es seien:

- V ein K-Vektorraum mit $\dim V = n < \infty$ und $F \in \operatorname{End} V$,
- Φ, $\Psi : K^n \to V$ Koordinatensysteme für V,
- $K^{n,n} \ni A = \Phi^{-1} \circ F \circ \Phi$ die darstellende Matrix von F bezüglich Φ,
- $K^{n,n} \ni B = \Psi^{-1} \circ F \circ \Psi$ die darstellende Matrix von F bezüglich Ψ.

Gesucht ist der Zusammenhang zwischen A und B!

Es gilt $F = \Phi \circ A \circ \Phi^{-1}$, also

$$B = \Psi^{-1} \circ \Phi \circ A \circ \Phi^{-1} \circ \Psi = S \cdot A \cdot S^{-1}$$

mit $S = \Psi^{-1} \circ \Phi \in GL(n, K)$.

Definition. Die Matrizen A, $B \in K^{n,n}$ heißen zueinander *ähnlich,* wenn es eine Matrix $S \in GL(n, K)$ mit $B = S \cdot A \cdot S^{-1}$ gibt.

Bemerkungen. 1. Die Ähnlichkeit von Matrizen ist eine Äquivalenzrelation.
Reflexivität: Wegen $A = E_n \cdot A \cdot E_n = E_n \cdot A \cdot E_n^{-1}$ ist jede Matrix A zu sich selbst ähnlich.
Symmetrie: $B = S \cdot A \cdot S^{-1} \Rightarrow A = S^{-1} \cdot B \cdot S = T \cdot B \cdot T^{-1}$ mit $T = S^{-1} \in GL(n, K)$.
Transitivität: $B = S \cdot A \cdot S^{-1} \wedge C = T \cdot B \cdot T^{-1} \Rightarrow C = T \cdot S \cdot A \cdot S^{-1} \cdot T^{-1} = (T \cdot S) \cdot A \cdot (T \cdot S)^{-1}$ mit $T \cdot S \in GL(n, K)$.
2. Ähnliche Matrizen haben gleiche Determinante:

$$\begin{aligned} \det B &= \det(S \cdot A \cdot S^{-1}) = \det S \cdot \det A \cdot \det(S^{-1}) = \\ &= \det S \cdot (\det S)^{-1} \cdot \det A = \det A \,. \end{aligned}$$

3. Sind A und B ähnliche Matrizen, so gibt es einen Endomorphismus F von K^n und Koordinatensysteme Φ, Ψ für K^n, derart dass A und B die darstellenden Matrizen von F bezüglich Φ beziehungsweise Ψ sind:
Ist $B = S \cdot A \cdot S^{-1}$, so setzen wir $F = B$, $\Phi = S$ und $\Psi = E_n$. Dann gilt tatsächlich

$$\begin{aligned} \Phi^{-1} \circ F \circ \Phi &= S^{-1} \cdot B \cdot S = A \,, \\ \Psi^{-1} \circ F \circ \Psi &= E_n^{-1} \cdot B \cdot E_n = B \,. \end{aligned}$$

Problem. *Normalformen* für die Ähnlichkeitsklassen von Matrizen.

Definitionen. Es seien V ein Vektorraum und $F \in \operatorname{End} V$.

- Ein Skalar $\lambda \in K$ ist ein *Eigenwert* von F, wenn es einen Vektor $v \in V \setminus \{0\}$ gibt, derart dass gilt: $F(v) = \lambda \cdot v$.
- Ein Vektor $v \in V \setminus \{0\}$ ist ein *Eigenvektor* von F, wenn es einen Skalar $\lambda \in K$ gibt, derart dass gilt: $F(v) = \lambda \cdot v$; man spricht dann auch von einem *Eigenvektor zum Eigenwert* λ.

Bemerkungen. 1. Die Vektoren in $\operatorname{Kern} F \setminus \{0\}$ sind Eigenvektoren, und zwar genau die Eigenvektoren zum Eigenwert 0. Eigenwerte können Null sein, Eigenvektoren nie!

2. Unter den Eigenwerten und Eigenvektoren einer quadratischen Matrix versteht man die Eigenwerte und Eigenvektoren der zugehörigen linearen Abbildung. In diesem Sinn gilt: Ist A darstellende Matrix eines Endomorphismus F, so stimmen die Eigenwerte von A und F überein. Um das einzusehen, bezeichne Φ ein Koordinatensystem, bezüglich dessen A den Endomorphismus F darstellt; dann gilt:

$$\begin{aligned} A &= \Phi^{-1} \circ F \circ \Phi\,, \\ \Phi \circ A &= F \circ \Phi\,, \\ A \circ \Phi^{-1} &= \Phi^{-1} \circ F\,. \end{aligned}$$

Ist nun λ ein Eigenwert von A und w ein zugehöriger Eigenvektor, so ist $\Phi(w) \neq 0$ und es gilt

$$F(\Phi(w)) = \Phi(A(w)) = \Phi(\lambda w) = \lambda \cdot \Phi(w)\,;$$

das heißt, λ ist ein Eigenwert von F mit zugehörigem Eigenvektor $\Phi(w)$.

Ist umgekehrt λ ein Eigenwert von F und v ein zugehöriger Eigenvektor, so ist $\Phi^{-1}(v) \neq 0$ und es gilt

$$A(\Phi^{-1}(v)) = \Phi^{-1}(F(v)) = \Phi^{-1}(\lambda v) = \lambda \cdot \Phi^{-1}(v)\,;$$

das heißt, λ ist ein Eigenwert von A mit zugehörigem Eigenvektor $\Phi^{-1}(v)$.

Lemma. *Es seien V ein Vektorraum mit* $\dim V = n < \infty$ *und* $F \in \operatorname{End} V$.
Der Endomorphismus F ist genau dann durch eine Diagonalmatrix darstellbar, wenn V eine Basis aus Eigenvektoren von F besitzt.

Beweis. „$\Rightarrow$“: Es sei Φ ein Koordinatensystem für V, derart dass die darstellende Matrix A von F bezüglich Φ eine Diagonalmatrix ist, das heißt, die Form

$$A = (\lambda_1 e^1, \ldots, \lambda_n e^n)$$

mit Skalaren $\lambda_1, \ldots, \lambda_n \in K$ besitzt. Nach Definition der darstellenden Matrix haben wir

$$A = \Phi^{-1} \circ F \circ \Phi\,,$$

das heißt,

$$\Phi \circ A = F \circ \Phi\,.$$

Wir setzen $v_j = \Phi(e^j)$ für alle $j \in \{1, \ldots, n\}$; da Φ ein Isomorphismus ist, ist $\{v_1, \ldots, v_n\}$ eine Basis von V und wir berechnen

$$\begin{aligned} F(v_j) &= F \circ \Phi(e^j) = \Phi \circ A(e^j) = \Phi(\lambda_j e^j) = \\ &= \lambda_j \Phi(e^j) = \lambda_j v_j\,. \end{aligned}$$

Da eine Basis den Nullvektor nicht enthält, ergibt sich, dass jeder Basisvektor v_j ein Eigenvektor zum Eigenwert λ_j ist.

„$\Leftarrow$“: Es sei $\mathfrak{B} = \{v_1, \ldots, v_n\}$ eine Basis von V, die aus Eigenvektoren von F besteht, und $\lambda_1, \ldots, \lambda_n$ seien die zugehörigen Eigenwerte. Bezeichnet A die darstellende Matrix von F bezüglich $\mathfrak{B}$, so gilt – in den Spalten stehen die Bilder der Einheitsvektoren – für alle $j \in \{1, \ldots, n\}$:

$$\begin{aligned} A(e^j) &= \Phi_{\mathfrak{B}}^{-1} \circ F \circ \Phi_{\mathfrak{B}}(e^j) = \Phi_{\mathfrak{B}}^{-1} \circ F(v_j) = \\ &= \Phi_{\mathfrak{B}}^{-1}(\lambda_j v_j) = \lambda_j \Phi_{\mathfrak{B}}^{-1}(v_j) = \lambda_j e^j\,. \end{aligned}$$

Damit ergibt sich $A = (\lambda_1 e^1, \ldots, \lambda_n e^n)$, das heißt, A ist eine Diagonalmatrix. $\square$

Definitionen. Ein Endomorphismus F eines endlich-dimensionalen Vektorraumes V heißt *diagonalisierbar,* wenn V eine Basis aus Eigenvektoren von F besitzt.
– Eine quadratische Matrix heißt *diagonalisierbar,* wenn sie als lineare Abbildung diagonalisierbar, also ähnlich zu einer Diagonalmatrix, ist.

23. Mai 2000

Lemma. *Es seien V ein Vektorraum und $F \in \operatorname{End} V$.*
Eigenvektoren zu verschiedenen Eigenwerten von F sind linear unabhängig.
Beweis. Es seien $v_1, \ldots, v_k$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$. Wir führen den Beweis durch Induktion nach k. Da ein Eigenvektor immer vom Nullvektor verschieden ist, ist der Induktionsanfang klar.
„$k \Rightarrow k+1$“: Es sei

$$v_{k+1} = \sum_{j=1}^{k} \mu_j v_j \,;$$

dann gilt:

$$\lambda_{k+1} v_{k+1} = \sum_{j=1}^{k} \lambda_{k+1} \mu_j v_j$$

und

$$\begin{aligned} \lambda_{k+1} v_{k+1} &= F(v_{k+1}) = F(\sum_{j=1}^{k} \mu_j v_j) = \\ &= \sum_{j=1}^{k} \mu_j F(v_j) = \sum_{j=1}^{k} \mu_j \lambda_j v_j \end{aligned}$$

Aus der Induktionsannahme – das ist die lineare Unabhängigkeit der Vektoren $v_1, \ldots, v_k$ – folgt nun

$$\lambda_{k+1} \mu_j = \mu_j \lambda_j \text{ für alle } j \in \{1, \ldots, k\}$$

Wegen $v_{k+1} \neq 0$ finden wir einen Index $j_0 \in \{1, \ldots, k\}$ mit $\mu_{j_0} \neq 0$. Für diesen Index gilt dann

$$\lambda_{k+1} = \lambda_{j_0}$$

im Widerspruch zur vorausgesetzten paareweisen Verschiedenheit der auftretenden Eigenwerte. □

Folgerung. *Ein Endomorphismus eines n-dimensionalen Vektorraumes mit n paarweise verschiedenen Eigenwerten ist diagonalisierbar.* □

Bezeichnung und Definition. Es seien V ein Vektorraum und $F \in \operatorname{End} V$. Die Teilmenge

$$\operatorname{Eig}(F; \lambda) = \{v \in V \,:\, F(v) = \lambda \cdot v\}$$

heißt *Eigenraum von F bezüglich λ.*

Lemma. *Es seien V ein Vektorraum und $F \in \operatorname{End} V$.*

a) Eigenräume von F sind Untervektorräume von V.

Für alle Skalare $\lambda \in K$ gilt:

b) $\operatorname{Eig}(F; \lambda) = \operatorname{Kern}(F - \lambda \cdot \operatorname{Id}_V)$*;*

c) λ Eigenwert von $F \Leftrightarrow \mathrm{Eig}(F;\lambda) \neq 0 \Leftrightarrow \det(A - \lambda E_n) = 0$;

d) Für alle Skalare λ_1, $\lambda_2 \in K$ gilt:

$$\lambda_1 \neq \lambda_2 \Rightarrow \mathrm{Eig}(F;\lambda_1) \cap \mathrm{Eig}(F;\lambda_2) = 0\,.$$

Beweis. a) Wegen $F(0) = 0 = \lambda \cdot 0$ ist $\mathrm{Eig}(F;\lambda) \neq \emptyset$.
Für v, $w \in \mathrm{Eig}(F;\lambda)$ gilt:

$$F(v + w) = F(v) + F(w) = \lambda v + \lambda w = \lambda(v + w)\,,$$

also $v + w \in \mathrm{Eig}(F;\lambda)$.
Für $v \in \mathrm{Eig}(F;\lambda)$ und beliebiges $\mu \in K$ gilt:

$$F(\mu v) = \mu F(v) = \mu\lambda v = \lambda(\mu v)\,,$$

also $\mu v \in \mathrm{Eig}(F;\lambda)$.

b) Für alle $v \in V$ gilt:

$$\begin{aligned} v \in \mathrm{Eig}(F;\lambda) &\Leftrightarrow F(v) = \lambda v = \lambda \mathrm{Id}(v) \Leftrightarrow \\ &\Leftrightarrow F(v) - (\lambda \mathrm{Id})(v) = 0 \Leftrightarrow \\ &\Leftrightarrow (F - \lambda \mathrm{Id})(v) = 0 \Leftrightarrow \\ &\Leftrightarrow v \in \mathrm{Kern}(F - \lambda \mathrm{Id})\,. \end{aligned}$$

c) Die erste Äquivalenz ist nach der Definition der Eigenwerte trivial. Ferner gilt: λ Eigenwert von $A \in K^{n,n} \Leftrightarrow \mathrm{Eig}(A;\lambda) \neq 0 \overset{b)}{\Leftrightarrow} \mathrm{Kern}(A - \lambda E_n) \neq 0 \Leftrightarrow \det(A - \lambda E_n) = 0$.

d) Für $v \in \mathrm{Eig}(F;\lambda_1) \cap \mathrm{Eig}(F;\lambda_2)$ gilt:

$$\begin{aligned} v &= (\lambda_1 - \lambda_2)^{-1}(\lambda_1 - \lambda_2)v = (\lambda_1 - \lambda_2)^{-1}(\lambda_1 v - \lambda_2 v) = \\ &= (\lambda_1 - \lambda_2)^{-1}(F(v) - F(v)) = 0\,. \end{aligned}$$

□

Bezeichnung und Definition. Es sei $A \in K^{n,n}$. Das Polynom

$$P_A = \det(A - t \cdot E_n) \in K[t]$$

heißt *charakteristisches Polynom* von A.

Bemerkung. Die Eigenwerte einer Matrix sind gerade die *Nullstellen* oder *Wurzeln* des charakteristischen Polynoms (= Nullstellen der zugehörigen Polynomfunktion).

Beispiele. 1. Drehungen in der Ebene (Zentrum im Ursprung, Drehwinkel α):

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

charakteristisches Polynom: $P_A = (\cos\alpha - t)^2 + \sin^2\alpha$
Nullstellen existieren nur bei $\alpha = 0, \pi$. Es gibt jeweils nur einen Eigenwert, entweder 1 oder -1. Im ersten Fall ist A als Abbildung die Identität, im zweiten die Punktspiegelung am Ursprung.

2. Achsenspiegelung an einer Achse mit der Steigung $\tan\alpha$:

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

charakteristisches Polynom: $P_A = t^2 - 1$
Eigenwerte $1, -1$;
Eigenvektor zum Eigenwert 1: $(\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2})$;
Eigenvektor zum Eigenwert -1: $(-\sin\frac{\alpha}{2}, \cos\frac{\alpha}{2})$.

Bemerkung. Ähnliche Matrizen haben das gleiche charakteristische Polynom: Dazu muß man nur beachten, dass das Matrizenprodukt homogen in jeder Variablen ist. Wir zeigen: Sind A und B zueinander ähnliche Matrizen in $K^{n,n}$, so sind auch die Matrizen $A - tE_n$ und $B - tE_n$ in $K(t)^{n,n}$ zueinander ähnlich. Zunächst eine Nebenrechnung: Für alle $S \in GL(n, K)$ gilt:

$$tE_n = t(S \cdot E_n \cdot S^{-1}) = S \cdot t(E_n \cdot S^{-1}) = S \cdot tE_n \cdot S^{-1}\,.$$

Ist nun $B = S \cdot A \cdot S^{-1}$, so berechnen wir weiter

$$\begin{aligned} B - tE_n &= S \cdot A \cdot S^{-1} - S \cdot tE_n \cdot S^{-1} = \\ &= S \cdot (A - tE_n) \cdot S^{-1}\,. \end{aligned}$$

Definition. Es seien V ein Vektorraum mit $\dim V = n < \infty$ und $F \in \operatorname{End} V$. Das *charakteristische Polynom* P_F von F wird definiert als das charakterische Polynom einer darstellenden Matrix von F; da alle darstellenden Matrizen von F zueinander ähnlich sind, ist diese Festsetzung unabhängig von der Auswahl der darstellenden Matrix.

26. Mai 2000

Lemma und Definition. *Es sei $A \in K^{n,n}$. Ist $P_A = \det(A - t \cdot E_n) = \sum_{k=0}^{n} \alpha_k t^k$, so gilt:*

$$\begin{aligned} \alpha_n &= (-1)^n\,, \\ \alpha_{n-1} &= (-1)^{n-1} \sum_{i=1}^{n} \alpha_{ii} \text{ und} \\ \alpha_0 &= \det A\,. \end{aligned}$$

Der Skalar $\sum_{i=1}^{n} \alpha_{ii} \in K$ heißt Spur von A.

Beweis. Wir setzen $B = (\beta_{ij}) = A - tE_n$. Die Komponenten β_{ij} von B sind Polynome in t vom Grad kleiner-gleich 1; der Grad 1 liegt genau dann vor, wenn $i = j$ ist. Die zur Berechnung von $\det B$ benötigten Produkte $\prod_{i=1}^{n} \beta_{i\pi(i)}$ sind dann Polynome vom Grad kleiner-gleich n (für alle $\pi \in \mathfrak{S}_n$). Ist $\pi \neq \mathrm{id}$, so haben mindestens zwei Faktoren $\beta_{i\pi(i)}$ einen Grad kleiner-gleich 0, also hat das Produkt $\prod_{i=1}^{n} \beta_{i\pi(i)}$ einen Grad kleiner-gleich $n-2$. Damit ergeben sich die beiden höchsten Koeffizienten von P_A allein aus dem Produkt $\prod_{i=1}^{n} \beta_{ii} = \prod_{i=1}^{n} (\alpha_{ii} - t)$, was unmittelbar auf $\alpha_n = (-1)^n$ und $\alpha_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} \alpha_{ii}$ führt. Den niedrigsten Koeffizienten α_0 erhält man, in dem man $t = 0$ setzt; das liefert $\alpha_0 = \det A$. $\square$

Folgerungen. 1. *Das charakteristische Polynom einer $n \times n$-Matrix hat den Grad n.*
2. *Ähnliche Matrizen haben gleiche Spur.*

Erinnerung an den Grad eines Polynoms: Es sei $P = \sum_{k=0}^{\infty} \alpha_k t^k \in K[t]$. Man setzt

$$\operatorname{Grad} P = \begin{cases} \max\{i \mid i \in \mathbb{N}_0 \wedge \alpha_i \neq 0\}, & P \neq 0, \\ -\infty, & P = 0. \end{cases}$$

Dann gilt für alle P, $Q \in K[t]$:

$$\begin{aligned} \operatorname{Grad}(P \cdot Q) &= \operatorname{Grad} P + \operatorname{Grad} Q\,, \\ \operatorname{Grad}(P \pm Q) &\leq \max\{\operatorname{Grad} P, \operatorname{Grad} Q\}\,, \end{aligned}$$

wobei im zweiten Fall Gleichheit sicher dann auftritt, wenn $\operatorname{Grad} P \neq \operatorname{Grad} Q$ ist. Mit dieser Konvention ist $P \neq 0 \Leftrightarrow \operatorname{Grad} P \geq 0$ für alle $P \in K[t]$. Auch ergibt sich, dass Polynome vom Grad 0 keine Nullstellen haben; die zugehörigen Polynomfunktionen sind konstante Funktionen mit einem Wert ungleich Null.

Polynomdivision. Satz von der Division mit Rest: *Zu P_1, $P_2 \in K[t]$ mit $P_2 \neq 0$ gibt es eindeutige bestimmte Polynome Q, $R \in K[t]$ mit*

$$P_1 = P_2 \cdot Q + R \text{ und } \operatorname{Grad} R < \operatorname{Grad} P_2\,.$$

Beweis. *Eindeutigkeit*: Sei $P_2 \cdot Q + R = P_2 \cdot \tilde{Q} + \tilde{R}$ mit $\operatorname{Grad} R$, $\operatorname{Grad} \tilde{R} < \operatorname{Grad} P_2$. Umrechnung ergibt

$$P_2(Q - \tilde{Q}) = \tilde{R} - R\,.$$

Wäre nun $Q \neq \tilde{Q}$, also $Q - \tilde{Q} \neq 0$, so hätten wir einerseits

$$\operatorname{Grad}(P_2(Q - \tilde{Q})) = \operatorname{Grad} P_2 + \operatorname{Grad}(Q - \tilde{Q}) \geq \operatorname{Grad} P_2$$

und andererseits

$$\operatorname{Grad}(\tilde{R} - R) \leq \max\{\operatorname{Grad} \tilde{R}, \operatorname{Grad} R\} < \operatorname{Grad} P_2\,.$$

was aber nicht sein kann. Damit haben wir $Q = \tilde{Q}$ und im Hinblick auf die eben angegebene Gleichung auch $R = \tilde{R}$.
Existenz: Ist $\operatorname{Grad} P_1 < \operatorname{Grad} P_2$, so erfüllen $Q = 0$ und $R = P_1$ die gewünschten Bedingungen. Sei nun $\operatorname{Grad} P_2 = m$ und $\operatorname{Grad} P_1 = m + n$ mit $n \in \mathbb{N}_0$, also $P_1 = \sum_{i=0}^{m+n} \alpha_i t^i$, $P_2 = \sum_{i=0}^{m} \beta_i t^i$ mit $\alpha_{m+n} \neq 0$, $\beta_m \neq 0$. Der weitere Beweis wird durch Induktion nach n geführt.
Induktionsanfang bei $n = 0$. Wir setzen $Q = \alpha_m / \beta_m$ und $R = P_1 - P_2 Q$. Dann gilt zunächst

$$\operatorname{Grad} R \leq \max\{\operatorname{Grad} P_1, \operatorname{Grad} P_2 + \operatorname{Grad} Q\} = m$$

und wir berechnen für den Koeffizienten von t^m in R:

$$\alpha_m - \beta_m \cdot \frac{\alpha_m}{\beta_m} = 0\,;$$

also ist sicher $\operatorname{Grad} R < m = \operatorname{Grad} P_2$.
Induktionsschluß für $n > 0$ unter der Annahme, dass die Behauptung für $\operatorname{Grad} P_1 - \operatorname{Grad} P_2 < n$ bereits als richtig erkannt ist. Wir setzen zunächst

$$\begin{aligned} Q_1 &= \frac{\alpha_{m+n}}{\beta_m} \cdot t^n \\ P_3 &= P_1 - P_2 Q_1\,. \end{aligned}$$

Dann gilt zunächst

$$\operatorname{Grad} P_3 \leq \max\{\operatorname{Grad} P_1, \operatorname{Grad} P_2 + \operatorname{Grad} Q_1\} = m + n$$

und wir berechnen für den Koeffizienten von t^{m+n} in P_3:

$$\alpha_{m+n} - \beta_m \cdot \frac{\alpha_{m+n}}{\beta_m} = 0\,;$$

also ist $\operatorname{Grad} P_3 < m + n$. Ist sogar $\operatorname{Grad} P_3 < m$, so setzen wir $Q = Q_1$, $R = P_3$ und sind fertig. Andernfalls finden wir nach Induktionsvoraussetzung Polynome Q_2 und R mit $P_3 = P_2Q_2 + R$ und $\operatorname{Grad} R < \operatorname{Grad} P_2$. Setzen wir $Q = Q_1 + Q_2$, so ergibt sich auch

$$P_1 = P_2Q_1 + P_3 = P_2Q_1 + P_2Q_2 + R = P_2Q + R\,. \;\square$$

Bemerkungen. $P \in K[t] \setminus \{0\}$, $\lambda \in K$, $P(\lambda) = 0 \Rightarrow$ Es gibt genau ein $Q \in K[t]$ mit $P = (t - \lambda)Q$ (mit $\operatorname{Grad} Q = \operatorname{Grad} P - 1$).
Beweis. *Existenz*: Durch Polynomdivision erhalten wir Q, $R \in K[t]$ mit $P = (t-\lambda)Q+R$ und $\operatorname{Grad} R < \operatorname{Grad}(t - \lambda) = 1$. Setzen wir λ ein, so ergibt sich

$$0 = P(\lambda) = (\lambda - \lambda)Q(\lambda) + R(\lambda) = R(\lambda)\,.$$

Damit kann $\operatorname{Grad} R$ nicht 0 sein, also bleibt nur $\operatorname{Grad} R = -\infty$, das heißt, $R = 0$ übrig.
Eindeutigkeit: $(t - \lambda)(Q - \tilde{Q}) = 0$ ist nur für $Q = \tilde{Q}$ möglich (Nullteilerfreiheit des Polynomringes). $\square$

30. Mai 2000

Dies Ergebnis läßt sich auch so ausdrücken: Ist der Skalar $\lambda \in K$ Wurzel des Polynoms $P \in K[t]$, so ist die rationale Funktion (=Element des Körpers $K(t)$) $P/(t - \lambda)$ ein Polynom, dessen Grad um genau 1 kleiner ist als der Grad von P. Daraus folgern wir, dass ein Polynom vom Grad n höchstens n Nullstellen hat.
Beweis. Es seien $\lambda_1, \ldots, \lambda_k$ paarweise verschiedene Nullstellen von $P \in K[t]$ mit $\operatorname{Grad} P = n$. Wir setzen $Q = P/(t - \lambda_1)$ und finden für $i \in \{2, \ldots, k\}$

$$0 = P(\lambda_i) = (\lambda_i - \lambda_1)Q(\lambda_i)\,;$$

also wegen $\lambda_i - \lambda_1 \neq= 0$ auch $Q(\lambda_i) = 0$. Wegen $\operatorname{Grad} Q = n - 1$ können wir nun die Induktionsvoraussetzung anwenden und erhalten $k - 1 \leq n - 1$, das heißt, $k \leq n$, wie behauptet. $\square$
Da also ein Polynom n-ten Grades höchstens n Nullstellen hat, hat eine $n \times n$-Matrix auch aus diesem Grund höchstens n Eigenwerte. Bei der Existenz von n (paarweise verschiedenen) Eigenwerten ist – wie schon festgestellt – die Matrix diagonalisierbar.

Lemma, Bezeichnung und Definition. *Zu jedem Paar* (P, λ), *bestehend aus einem Polynom* $P \in K[t] \setminus \{0\}$ *und einem Skalar* $\lambda \in K$, *gibt es ein eindeutig bestimmtes Paar* (m, Q), *bestehend aus einer natürlichen Zahl* $m \in \mathbb{N}_0$ *und einem Polynom* $Q \in K[t]$, *derart dass gilt*:

$$\begin{aligned} P &= (t - \lambda)^m \cdot Q \qquad \text{und} \\ Q(\lambda) &\neq 0\,. \end{aligned}$$

Die Zahl m wird im folgenden durch $m(P; \lambda)$ bezeichnet; sie heißt *Vielfachheit* (oder *Multiplizität*) von λ als Nullstelle von P.
Beweis. *Existenz*: Wir setzen

$$m = \max\{i \mid i \in \mathbb{N}_0 \wedge \frac{P}{(t - \lambda)^i} \text{ ist Polynom}\}$$

und $Q = P/(t-\lambda)^m$. Wäre $Q(\lambda) = 0$, so gäbe es $Q_1 \in K[t]$ mit $Q = (t-\lambda)Q_1$ und damit wäre $P/(t-\lambda)^{m+1} = Q/(t-\lambda) = Q_1$ auch ein Polynom, im Widerspruch zur Maximalität von m.
Eindeutigkeit: $(t-\lambda)^m Q = (t-\lambda)^{\tilde{m}}\tilde{Q}$ mit impliziert $m < \tilde{m}$ impliziert $Q = (t-\lambda)^{\tilde{m}-m}\tilde{Q}$ und damit

$$Q(\lambda) = (\lambda-\lambda)^{\tilde{m}-m}\tilde{Q} = 0\,;$$

also hat Q nicht die gewünschte Eigenschaft. Im Fall $m > \tilde{m}$ ergibt sich analog, dass $\tilde{Q}$ nicht die gewünschte Eigenschaft hat. □

Lemma. *Es seien V ein Vektorraum mit* $\dim V = n < \infty$ *und* $F \in \operatorname{End} V$. *Dann gilt für alle* $\lambda \in K$ *und* $F \in L(V,V)$:

$$m(P_F;\lambda) \geq \dim \operatorname{Eig}(F;\lambda)\,.$$

(In anderen Worten: die *algebraische* Vielfachheit eines Skalars in Bezug auf das charakteristische Polynom eines Endomorphismus eines endlich-dimensionalen Vektorraumes ist größer-gleich der *geometrischen* Vielfachheit des Skalars, worunter die Dimension des zugehörigen Eigenraumes verstanden wird.)

Beweis. Falls λ kein Eigenwert ist, sind beide Seiten der Ungleichung 0. Ist λ ein Eigenwert, so ist jedenfalls $k = \dim \operatorname{Eig}(F;\lambda) \geq 1$ und wir finden eine Basis $\{v_1,\ldots,v_k\}$ von $\operatorname{Eig}(F;\lambda)$. Diese ergänzen durch Vektoren v_{k+1}, …, v_n zu einer Basis $\mathfrak{B}$ von V und setzen zur Abkürzung $\Phi = \Phi_{\mathfrak{B}}$. Dann gilt für die ersten k Spalten der darstellenden Matrix A:

$$A(e^j) = \Phi^{-1} \circ f \circ \Phi(e^j) = \Phi^{-1}(\lambda v_j) = \lambda e^j\,.$$

Damit gilt:

$$A = \begin{pmatrix} \lambda E_k & * \\ 0 & B \end{pmatrix};$$

also erhalten wir

$$A - tE_n = \begin{pmatrix} (\lambda - t)E_k & * \\ 0 & B - tE_{n-k} \end{pmatrix}.$$

Damit berechnen wir

$$P_A = \det(\lambda - t)E_k \cdot P_B = (t-\lambda)^k \cdot (-1)^k P_B\,,$$

woraus $k \leq m(P_F;\lambda)$ folgt. □

Definitionen. Es seien V ein Vektorraum und $\mathcal{W} = (W_j)_{j\in J}$ eine Familie von Untervektorräumen von V.
a) V ist *Summe* der Familie $\mathcal{W}$ - symbolisch

$$V = \sum_{j\in J} W_j \quad -$$

wenn sich jeder Vektor $v \in V$ als Summe

$$v = \sum_{j\in J} w_j$$

mit $w_j \in W_j$ für alle $j \in J$ und $w_j = 0$ für fast alle $j \in J$ darstellen läßt, das heißt, wenn gilt:

$$V = \operatorname{Span}\Big(\bigcup_{j\in J} W_j\Big)\,.$$

b) V ist *direkte Summe* der Familie $\mathcal{W}$ - symbolisch

$$V = \bigoplus_{j \in J} W_j \quad -$$

wenn sich jeder Vektor $v \in V$ **eindeutig** als Summe

$$v = \sum_{j \in J} w_j$$

mit $w_j \in W_j$ für alle $j \in J$ und $w_j = 0$ für fast alle $j \in J$ darstellen läßt.
Im Falle $J = \{1, \ldots, k\}$ $(k \in \mathbb{N})$ schreibt man auch

$$V = \bigoplus_{j=1}^{k} W_j = W_1 \oplus W_2 \oplus \ldots \oplus W_k \,.$$

Sprechweise. Ein Polynom *zerfällt in Linearfaktoren,* wenn es ein Produkt von *linearen* Polynomen, das heißt, Polynomen mit dem Grad 1 ist.

Satz. *Es seien V ein Vektorraum mit $\dim V = n < \infty$ und $F \in \operatorname{End} V$. Dann sind die folgenden Aussagen äquivalent:*
a) der Endomorphismus F ist diagonlisierbar;
b) das charakteristische Polynom P_F von F zerfällt in Linearfaktoren und es gilt für alle $\lambda \in K$:

$$m(P_F; \lambda) = \dim \operatorname{Eig}(F; \lambda)\,;$$

c) der Vektorraum V ist direkte Summe der nichttrivialen Eigenräume von F, das heißt, sind $\lambda_1, \ldots, \lambda_k$ die Eigenwerte von F (paarweise verschieden), so ist

$$V = \operatorname{Eig}(F; \lambda_1) \oplus \ldots \oplus \operatorname{Eig}(F; \lambda_k)\,.$$

Beweis. a) $\Rightarrow$ b): Es seien $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Eigenwerte von F. Wir wählen eine Basis $\mathfrak{B} = \{v_1, \ldots, v_n\}$ aus Eigenvektoren von F. Für jedes $j \in \{1, \ldots, k\}$ bezeichne m_j die Anzahl der Eigenvektoren zum Eigenwert λ_j in $\mathfrak{B}$ und $\bar{m}_j = \sum_{\bar{j}=1}^{j} m_{\bar{j}}$. Dann gilt $\dim \operatorname{Eig}(F; \lambda_j) \geq m_j \geq 1$ für alle $j \in \{1, \ldots, n\}$, sowie $\bar{m}_1 = m_1$ und $\bar{m}_k = n$. Durch Umnumerieren der Basisvektoren können wir erreichen, dass die Vektoren $v_1, \ldots, v_{m_1}$ Eigenvektoren zum Eigenwert λ_1 und für $j \in \{2, \ldots, k\}$ die Vektoren $v_{\bar{m}_{j-1}+1}, \ldots, v_{\bar{m}_j}$ Eigenvektoren zum Eigenwert λ_j sind. Die darstellende Matrix A von F bezüglich $\mathfrak{B}$ hat dann die Form

$$A = \begin{pmatrix} \lambda_1 E_{m_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 E_{m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k E_{m_k} \end{pmatrix},$$

woraus sich

$$P_F = (t - \lambda_1)^{m_1} \cdot (t - \lambda_2)^{m_2} \cdot \ldots \cdot (t - \lambda_k)^{m_k} \cdot (-1)^n$$

ergibt. Aus dieser Darstellung von P_F lassen sich sofort die Vielfachheiten der Eigenwerte von F ablesen: für alle $j \in \{1, \ldots, k\}$ gilt:

$$m(P_F; \lambda_j) = m_j\,.$$

Mit Hilfe des vorigen Lemmas und der eben angegebenen Abschätzung erhalten wir nun

$$\dim \operatorname{Eig}(F; \lambda_j) \leq m(P_F; \lambda_j) = m_j \leq \dim \operatorname{Eig}(F; \lambda_j)\,,$$

woraus sich die behauptete Gleichung für die Eigenwerte von F ergibt; für die übrigen $\lambda \in K$ ist sie trivial.
b) $\Rightarrow$ c): Wir setzen zur Abkürzung

$$m_j = m(P_F; \lambda_j) = \dim \operatorname{Eig}(F; \lambda_j)$$

für alle $j \in \{1, \ldots, k\}$ und auch wieder $\bar{m}_j = \sum_{\bar{j}=1}^{j} m_{\bar{j}}$. Weil P_F in Linearfaktoren zerfällt, gilt $\bar{m}_k = \sum_{j=1}^{k} m_j = n$.
Wir wählen nun eine Basis $\{v_1, \ldots, v_{m_1}\}$ von $\operatorname{Eig}(F; \lambda_1)$ und für $j \in \{2, \ldots, k\}$ Basen $\{v_{\bar{m}_{j-1}+1}, \ldots, v_{\bar{m}_j}\}$ von $\operatorname{Eig}(F; \lambda_j)$. Dann behaupten wir, dass die n Vektoren $v_1, \ldots, v_n$ linear unabhängig sind, also eine Basis von V bilden. Sei dazu $\sum_{i=1}^{n} \mu_i v_i = 0$ angenommen. Wir setzen $w_1 = \sum_{i=1}^{m_1} \mu_i v_i$ und $w_j = \sum_{i=\bar{m}_{j-1}+1}^{\bar{m}_j} \mu_i v_i$ für $j \in \{2, \ldots, k\}$, dann gilt $\sum_{j=1}^{k} w_j = 0$. Wären nun gewisse dieser w_j von Null verschieden, so handelte es sich um Eigenvektoren zu paarweise verschiedenen Eigenwerten; also wären sie linear unabhängig und damit könnte ihre Summe nicht Null sein. Also gilt $w_j = 0$ für alle $j \in \{1, \ldots, k\}$. Die Vektoren, die zur Bildung eines w_j herangezogen werden, sind Basis eines Eigenraums, also linear unabhängig; damit kann durch diese Vektoren der Nullvektor nur trivial dargestellt werden, das heißt, es gilt $\mu_i = 0$ für alle $i \in \{1, \ldots, n\}$. Damit ist die behauptete lineare Unabhängigkeit bewiesen und $\{v_1, \ldots, v_n\}$ ist eine Basis von V.

2. Juni 2000

Nun kann jeder beliebige Vektor $v \in V$ als Linearkombination $v = \sum_{i=1}^{n} \mu_i v_i$ dargestellt werden. Wir setzen wieder $w_1 = \sum_{i=1}^{m_1} \mu_i v_i$ und $w_j = \sum_{i=\bar{m}_{j-1}+1}^{\bar{m}_j} \mu_i v_i$ für $j \in \{2, \ldots, k\}$, dann gilt $v = \sum_{j=1}^{k} w_j$ mit $w_j \in \operatorname{Eig}(F; \lambda_j)$. Also ist V Summe der nichttrivialen Eigenräume von F. Es bleibt zu zeigen, dass die Summe direkt, das heißt, die angegebene Zerlegung des Vektors v eindeutig ist. Dazu sei $v = \sum_{j=1}^{k} \tilde{w}_j$ eine weitere Zerlegung von v mit $\tilde{w}_j \in \operatorname{Eig}(F; \lambda_j)$. Dann haben wir Skalare $\tilde{\mu}_1, \ldots, \tilde{\mu}_n$ derart, dass gilt: $\tilde{w}_1 = \sum_{i=1}^{m_1} \tilde{\mu}_i v_i$ und $\tilde{w}_j = \sum_{i=\bar{m}_{j-1}+1}^{\bar{m}_j} \tilde{\mu}_i v_i$ für $j \in \{2, \ldots, k\}$, also $v = \sum_{i=1}^{n} \tilde{\mu}_i v_i$. Da aber v nur eine Darstellung als Linearkombination der v_i besitzt, folgt $\tilde{\mu}_i = \mu_i$ für alle $i \in \{1, \ldots, n\}$, und damit $\tilde{w}_j = w_j$ für alle $j \in \{1, \ldots, k\}$.
c) $\Rightarrow$ a): Wir wählen für jedes $j \in \{1, \ldots, k\}$ eine Basis $\mathfrak{B}_j$ von $\operatorname{Eig}(F; \lambda_j)$. Da sich jedes $v \in V$ in der Form $v = \sum_{j=1}^{k} w_j$ mit $w_j \in \operatorname{Eig}(F; \lambda_j)$ zerlegen läßt und jeder Vektor w_j eine Linearkombination aus $\mathfrak{B}_j$ ist, ist $\bigcup_{j=1}^{k} \mathfrak{B}_j$ ein Erzeugendensystem von V. Nach Konstruktion besteht dieses Erzeugendensystem nur aus Eigenvektoren von F. Wenn es nicht linear unabhängig ist, können wir es zu einer Basis verkleinern. Damit haben wir auf jeden Fall eine Basis aus Eigenvektoren von F. Den möglichen Nachweis, dass die Menge $\bigcup_{j=1}^{k} \mathfrak{B}_j$ bereits selbst linear unabhängig ist, können wir uns deshalb sparen. $\square$

Definitionen. Es sei V ein Vektorraum mit $\dim V = n < \infty$:

1. Ein $(n+1)$-tupel $(V_0, \ldots, V_n)$ von Untervektorräumen eines n-dimensionalen Vektorraumes V heißt *Fahne* in V, wenn gilt: $V_{i-1} \underset{\neq}{\subset} V_i$ für $i = 1, \ldots, n$ ($\Rightarrow \dim V_i = i$ für $i = 0, \ldots, n$, $V_0 = 0$, $V_n = V$).
2. Ist ein Endomorphismus $F \in \operatorname{End}(V)$ gegeben, so heißt eine Fahne $(V_0, \ldots, V_n)$ in V *F-invariant,* wenn für alle $i \in \{0, \ldots, n\}$ gilt: $F(V_i) \subset V_i$.
3. Ein Endomorphismus $F \in \operatorname{End}(V)$ heißt *trigonalisierbar,* wenn es eine F-invariante Fahne gibt.
4. Eine quadratische Matrix heißt *trigonalisierbar,* wenn der zugehörige Endomorphismus trigonalisierbar ist.

Bemerkung. Ein Endomorphismus ist genau dann trigonalisierbar, wenn er durch eine Dreiecksmatrix dargestellt werden kann. Damit ist eine quadratische Matrix genau dann trigonalisierbar, wenn sie zu einer Dreiecksmatrix ähnlich ist.

Beweis. „$\Rightarrow$“: Es seien $F \in \operatorname{End} V$ und $(V_0, \ldots, V_n)$ in V eine F-invariante Fahne. Wir wählen der Reihe nach Vektoren v_1, v_2, $\ldots$, $v_n \in V$ derart, dass für jedes $j \in \{1, \ldots, n\}$ die Familie $(v_1, \ldots, v_j)$ eine Basis von V_j ist. Aus der F-Invarianz folgt dann $F(v_j) \in \operatorname{Span}(v_1, \ldots, v_j)$, das heißt,

$$F(v_j) = \sum_{i=1}^{j} \alpha_{ij} v_i$$

mit eindeutig bestimmten $\alpha_{ij} \in K$. Setzen wir noch $\alpha_{ij} = 0$ für $1 \leq j < i \leq n$, so ist $A = (\alpha_{ij})$ die darstellende Matrix von F bezüglich der Basis $(v_1, \ldots, v_n)$ von V und dabei handelt es sich um eine obere Dreiecksmatrix.

„$\Leftarrow$“: trivial. □

Satz. *Ein Endomorphismus F eines endlich-dimensionalen Vektorraumes ist genau dann trigonalisierbar, wenn P_F in Linearfaktoren zerfällt.*

Beweis. „$\Rightarrow$“: Ist $A = (\alpha_{ij})$ eine Dreiecksmatrix, so ist auch $A - tE$ eine Dreiecksmatrix und damit ist P_A das Produkt der Diagonalelemente von $A - tE$, das heißt,

$$P_A = \prod_{j=1}^{n} (\alpha_{jj} - t) \,.$$

„$\Leftarrow$“: Wir führen den Beweis durch Induktion nach $n = \dim V$. Im Fall $n = 1$ haben wir ein Polynom vom Grad 1, also ein lineares Polynom und damit ist nichts weiter zu beweisen.

„$n \Rightarrow n+1$“: Es seien V ein Vektorraum mit $\dim V = n+1$ und $F \in \operatorname{End} V$ mit $P_F = \prod_{j=1}^{n+1} (\lambda_j - t)$ für gewisse λ_1, $\ldots$, λ_n, λ_{n+1}. Die λ_j sind die Eigenwerte von F. Wir wählen zunächst einen Eigenvektor v_1 zum Eigenwert λ_1 und dann Vektoren v_2, $\ldots$, v_{n+1} derart, dass $\mathfrak{B} = \{v_1, \ldots, v_{n+1}\}$ eine Basis von V ist. Die darstellende Matrix A von F bezüglich $\mathfrak{B}$ hat die Form

$$A = \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{A} \end{pmatrix} .$$

Es sei nun

$$\tilde{V} = \operatorname{Span}(v_2, \ldots, v_{n+1}) \,;$$

damit ist $\tilde{V}$ ein Vektorraum mit $\dim \tilde{V} = n$ und Basis $\tilde{\mathfrak{B}} = \{v_2, \ldots, v_{n+1}\}$. Weiter sei $\tilde{F}$ der Endomorphismus von $\tilde{V}$, der bezüglich $\tilde{\mathfrak{B}}$ durch die Matrix $\tilde{A}$ dargestellt wird. Aus

$$A - tE_{n+1} = \begin{pmatrix} \lambda_1 - t & * \\ 0 & \tilde{A} - tE_n \end{pmatrix}$$

erhalten wir

$$\begin{aligned} \prod_{j=1}^{n+1} (\lambda_j - t) &= P_F = \det \begin{pmatrix} \lambda_1 - t & * \\ 0 & \tilde{A} - tE_n \end{pmatrix} = \\ &= (\lambda_1 - t) \cdot P_{\tilde{F}} \,; \end{aligned}$$

Division durch $(\lambda_1 - t)$ (im Körper $K(t)$) liefert dann

$$P_{\tilde{F}} = \prod_{j=2}^{n+1} (\lambda_j - t) \,,$$

das heißt, $P_{\tilde{F}}$ zerfällt in Linearfaktoren. Damit können wir auf $\tilde{F}$ die Induktionsvoraussetzung anwenden und finden eine $\tilde{F}$-invariante Fahne $(\tilde{V}_0, \ldots, \tilde{V}_n)$ in $\tilde{V}$. Da $\tilde{V}$ ein Untervektorraum von V ist, ist auch jedes $\tilde{V}_j$ ein Untervektorraum von V.
Wir setzen nun $V_0 = \tilde{V}_0 = 0$, $V_1 = \mathrm{Span}(v_1)$ und $V_j = V_1 + \tilde{V}_{j-1}$ für alle $j \in \{2, \ldots, n+1\}$. Um einzusehen, dass $(V_0, \ldots, V_{n+1})$ eine Fahne ist, genügt es, $\dim V_j = j$ für alle $j \in \{0, \ldots, n\}$ nachzuweisen. Für $j = 0, 1$ ergibt sich das unmittelbar aus der Konstruktion. Für größere j beachten wir, dass wegen $V_1 \cap \tilde{V} = 0$ auch $V_1 \cap \tilde{V}_{j-1} = 0$ für alle $j \in \{2, \ldots, n+1\}$ gilt. Damit ergibt sich aus der Dimensionsformel für Untervektorräume wie gewünscht:

$$\dim V_j = \dim V_1 + \dim \tilde{V}_{j-1} = 1 + j - 1 = j\,.$$

Nun ist noch die F-Invarianz dieser Fahne nachzuweisen. Wir betrachten $v \in V_j \setminus \{0\}$ und zerlegen zunächst in der Form $v = \mu v_1 + \tilde{v}$ mit $\tilde{v} \in \tilde{V}_{j-1}$. Bezeichnet w den Koordinatenvektor von $\tilde{v}$ bezüglich $\tilde{\mathfrak{B}}$, so ist $\begin{pmatrix} \mu \\ w \end{pmatrix}$ der Koordinatenvektor von v bezüglich $\mathfrak{B}$. Für den Koordinatenvektor von $F(v)$ erhalten wir damit die Form

$$A \begin{pmatrix} \mu \\ w \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{A} \end{pmatrix} \begin{pmatrix} \mu \\ w \end{pmatrix} = \begin{pmatrix} \nu \\ \tilde{A}w \end{pmatrix}.$$

Nach Definition von $\tilde{F}$ ist $\tilde{A}w$ der Koordinatenvektor von $\tilde{F}(\tilde{v})$; wegen der $\tilde{F}$-Invarianz von $\tilde{V}_{j-1}$ ist $\tilde{F}(\tilde{v}) \in \tilde{V}_{j-1}$. Damit gilt

$$F(v) = \nu v_1 + \tilde{F}(\tilde{v}) \in V_1 + V_{j-1} = V_j\,. \quad \square$$

6. Juni 2000

Bemerkung. Ein Körper K heißt *algebraisch abgeschlossen,* wenn jedes Polynom mit Koeffizienten aus K in Linearfaktoren zerfällt. Der sogenannte *Fundamentalsatz der Algebra* besagt, dass der Körper $\mathbb{C}$ der komplexen Zahlen ein Beispiel für einen algebraisch abgeschlossenen Körper ist. Nach dem eben bewiesen Satz ist jede quadratische Matrix über einem algebraisch abgeschlossenen Körper, also insbesondere jede quadratische Matrix mit komplexen Komponenten, ähnlich zu einer Dreiecksmatrix. $\square$