Lineare Algebra II und Geometrie

LVA 405.080

C. Fuchs

Rumpfskriptum zur Vorlesung **26.01.2018**

Inhaltsübersicht

Ziel des zweiten Teils der Linearen Algebra ist es, Vektorräume zu untersuchen, welche als zusätzliche Struktur ein Skalarprodukt besitzen. Aus der Schule ist bereits bekannt, dass sich damit sehr viele Aufgaben sehr schön lösen lassen (man denke z.B. an den Normalvektor einer Ebene oder die Hesse Normalform und die Extremwertaufgaben, welche sich dadurch lösen lassen). Wir werden allgemeiner versuchen alle möglichen Skalarprodukte zu klassifizieren. In der Vorlesung gibt es zwei große Highlights: Zum einen werden wir die Klassifikation von Endomorphismen mit Hilfe der Jordan-Normalform kennenlernen; zum anderen werden wir den Spektralsatz für normale Abbildungen formulieren und beweisen. Weitere Themen umfassen:

§1. Eigenwerte und Eigenvektoren

§2. Die Jordansche Normalform

§3. Euklidische und Unitäre Vektorräume

§4. Adjungierte Abbildungen

§5. Symmetrische Bilinearformen

Bei Fragen oder Bemerkungen (speziell Hinweise auf Fehler aller Art sind willkommen) schicken Sie ein Email an clemens.fuchs@math.ethz.ch.

§1. Eigenwerte und Eigenvektoren

1.1 Eigenvektoren

1.1.1 Es sei nun V ein endlich-dimensionaler Vektorraum über K, $\varphi : V \to V$ ein Endomorphismus. Ein $\lambda \in K$ heisst *Eigenwert* von φ, falls es mindestens einen Vektor $v \in V\backslash\{0\}$ gibt mit der Eigenschaft

$$\varphi(v) = \lambda \cdot v.$$

v heisst ein zum Eigenwert λ gehöriger *Eigenvektor*. Umgekehrt heisst $v \neq 0$ Eigenvektor, falls es einen Eigenwert λ von φ gibt mit v ein zugehöriger Eigenvektor. Da $\varphi(v) = \lambda v$ äquivalent zu $(\varphi - \lambda\mathrm{id})v = 0$ ist, ist die Menge $\mathrm{Ker}(\varphi - \lambda\mathrm{id})$ genau die Menge der zu λ gehörigen Eigenvektoren von φ zusammen mit dem Nullvektor. Wir halten das fest:

1.1.2 **Satz** Es ist $\lambda \in K$ genau dann Eigenwert von $\varphi \in \mathrm{End}(V)$, falls $\mathrm{Ker}(\varphi - \lambda\mathrm{id}) \neq \{0\}$. Offenbar ist φ genau dann injektiv, wenn $0 \in K$ kein Eigenwert von φ ist.

Beweis. Folgt sofort aus den Vorbemerkungen. □

1.1.3 Wir setzen $V_\lambda := \mathrm{Ker}(\varphi - \lambda \mathrm{id})$ und nennen diesen Vektorraum den *Eigenraum* von φ zum Eigenwert λ. Die Menge der Eigenwerte von φ nennen wir das *Spektrum* von φ, in Zeichen $\mathrm{Spec}(\varphi)$.

1.1.4 Betrachte $V = K^n, \varphi : K^n \to K^n$ von der Gestalt $\varphi = \varphi_A$ für $A \in M_n(K), n = \dim V$. Dann heissen die Eigenwerte, Eigenvektoren und Eigenräume von φ_A auch die *Eigenwerte, Eigenvektoren, Eigenräume* der Matrix A.

Beispiele a) Für $\dim V = 1$ hat jede Abbildung φ genau einen Eigenwert λ und $v \mapsto \lambda v$ für alle $v \in V$.
b) Mit Hilfe des Fortsetzungssatzes lassen sich für $2 \leq \dim V$ sofort lineare Abbildungen angeben, die zwei verschiedene Eigenwerte haben.
c) Sei $C^\infty(\mathbb{R})$ der Vektorraum aller beliebig oft stetig differenzierbaren Funktionen $\mathbb{R} \to \mathbb{R}$. Wir untersuchen die lineare Abbildung d/dx und die verallgemeinerte Exponentialfunktion $x \mapsto e^{\lambda x}$ mit $\lambda \in \mathbb{R}$ beliebig. Wegen $de^{\lambda x}/dx = \lambda e^{\lambda x}$ ist jede reelle Zahl Eigenwert von d/dx.
c) Die Abbildung $\varphi \in \mathrm{GL}_2(\mathbb{R})$ mit $(x, y) \mapsto (-y, x)$ hat keine Eigenwerte, da für alle $(x, y) \in \mathbb{R}^2 \backslash \{(0, 0)\}$ gilt

$$\begin{vmatrix} x & -y \\ y & x \end{vmatrix} \neq 0.$$

d) Bei $\dim V = 0$ gibt es für ein lineares φ keine Eigenwerte.
e) Sei $V = \mathbb{R}^2$ und φ eine Drehung um den Winkel α; somit gilt

$$M_{\eta\eta}(\varphi) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Anschaulich ist klar, dass es mit Ausnahme der Fälle $\alpha = 0, \pi$ keinen Eigenvektor geben kann. Ist φ jedoch eine Drehspiegelung um den Winkel α (in dem wir die Richung von $\varphi(e_2)$ umkehren), so gibt es die beiden Eigenvektoren ${}^t(\cos\alpha/2, \sin\alpha/2)$ zum Eigenwert 1 und ${}^t(\cos(\alpha+\pi)/2, \sin(\alpha+\pi)/2)$ zum Eigenwert -1. Führen wir diese beiden Vektoren als Basis β ein, so erhalten wir

$$M_{\beta\beta}(\varphi) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

als Koordinatenmatrix.

1.2 Das charakteristische Polynom

Sei $A \in M_n(K)$. Wir müssen zur Bestimmung der Eigenwerte jene $\lambda \in K$ bestimmen für die $\mathrm{Ker}(\varphi_A - \lambda \mathrm{id}) \neq \{0\}$ ist. Daraus folgt aber $\det(A - \lambda E_n) = 0$. Dies wollen wir nun genauer untersuchen. Somit definieren wir:

1.2.1 **Definition** Das Polynom $\det(A - XE_n) =: P_A(X) \in K[X]$ heisst das *charakteristische Polynom* von A.

1.2.2 **Satz** Es ist $\lambda \in K$ genau dann Eigenwert von A, wenn λ Nullstelle des charakteristischen Polynoms $P_A(X)$ ist.

Beweis. Es ist λ Eigenwert von A genau für $\mathrm{rg}(A - \lambda E_n) < n$, was gleichbedeutend zu $\det(A - \lambda E_n) = P_A(\lambda) = 0$ ist. $\square$

Beispiele a) Die Matrix
$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$
hat keine Eigenwerte. b) Die Matrix aus a) aufgefasst als Element von $M_2(\mathbb{C})$ hat die Eigenwerte $\pm i$. c) Jede Matrix $A \in M_n(\mathbb{R})$ mit n ungerade hat mindestens einen Eigenwert.

Der letzte Satz gibt uns grundsätzlich die Möglichkeiten die Eigenwerte von A zu bestimmen, ohne den Umweg über die zugehörigen Eigenvektoren gehen zu müssen.

1.2.3 Durch Entwicklung der Determinante erhalten wir
$$P_A(X) = (-1)^n X^n + (-1)^{n-1}\mathrm{tr}(A)X^{n-1} + \ldots + \det A,$$
wobei die Zahl $\mathrm{tr}(A) = \alpha_{11} + \ldots + \alpha_{nn} \in K$ die *Spur* der Matrix $A = (\alpha_{ij})$ genannt wird.

1.2.4 **Methode zur Berechnung der Eigenwerte und Eigenvektoren** Gegeben sei $A \in K^{n\times n}$.

1. Ermittle das charakteristische Polynom $P_A(X)$.
2. Bestimme die Nullstellen von $P_A(X)$; dies sind die Eigenwerte von A.
3. Zu jeder Nullstelle $\lambda \in K$ von $P_A(X)$ gehört der Eigenraum $\mathrm{Ker}(A - \lambda E_n)$. Eine Basis des Eigenraums ergibt sich dann mittels Gauß-Jordan aus der allgemeinen Lösung des homogenen $n \times n$-LGS
$$(A - \lambda E_n) \cdot x = 0.$$

1.2.5 Nun wenden wir uns der Berechnung der Eigenwerte eines Endomorphismus φ zu. Sei β eine Basis von V, $\Phi_\beta : K^n \to V$ die zugehörige Koordinatisierung und $A = M_{\beta\beta}(\varphi)$ eine Koordinatenmatrix von φ. Es ist $v \in V$ genau dann ein Eigenvektor von φ, falls $\Phi_\beta^{-1}(v)$ Eigenvektor der Matrix A ist. Die betreffenden Eigenwerte sind gleich! Wir bestimmen also die Eigenwerte der Matrix A. Welchen Einfluss hat ein Basiswechsel? Sei γ eine weitere Basis, dann folgt $M_{\gamma\gamma}(\varphi) = BM_{\beta\beta}(\varphi)B^{-1}$, wobei $B = M_{\gamma\beta}(\mathrm{id}_V) \in \mathrm{GL}_n$ ist; solche Matrizen haben wir ähnlich genannt.

1.2.6 **Satz** Ähnliche Matrizen haben dasselbe charakteristische Polynom.

Beweis. Sei $B = PAP^{-1}$ mit $P \in \mathrm{GL}_n(K)$. Es ist $P(A - XE_n)P^{-1} = PAP^{-1} - P(XE_n)P^{-1} = B - XE_n$ und daher gilt $P_B(X) = \det(B - XE_n) = \det P \det(A - XE_n) \det P^{-1} = \det(A - XE_n) = P_A(X)$. □

1.2.7 **Definition** Ist $\varphi \in \mathrm{End}(V), 1 \leq \dim V < \infty$ und β eine beliebige Basis von V, so ist das *charakteristische Polynom* $P_\varphi(X)$ erklärt als das charakteristische Polynom der Koordinatenmatrix $M_{\beta\beta}(\varphi)$. Aufgrund von Satz 1.2.6 ist diese Definition sinnvoll.

1.2.8 **Methode zur Berechnung der Eigenwerte und Eigenvektoren** Gegeben sei $\varphi \in \mathrm{End}(V)$.

1. Wähle eine geordnete Basis $\beta = (v_1, \ldots, v_n)$ von V.
2. Bestimme $A = M_{\beta\beta}(\varphi)$.
3. Berechne die Eigenwerte und Eigenvektoren von A mit 1.2.4.
4. Die Eigenwerte von A sind die Eigenwerte von φ.
5. Die Eigenvektoren von φ ergeben sich durch Anwenden der Koordinatisierungsabbildung Φ_β auf die Eigenvektoren von A.

1.2.9 Sei $\lambda \in K$ ein Eigenwert von φ. Wir nennen $\dim V_\lambda$ die *geometrische Vielfachheit* von λ. Die *algebraische Vielfachheit* von λ ist die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms (siehe 1.3.4).

1.3 Polynome

Sei K ein Körper. Die Menge der Folgen über K mit endlichem Träger (d.h. mit nur endlich vielen von 0 verschiedenen Folgengliedern), bilden einen Untervektorraum der Menge aller Folgen $K^{\mathbb{N}}$. Durch die Identifikation $X := (0, 1, 0, 0, \ldots)$ erhalten wir damit den K-Vektorraum der Polynome in der "Unbestimmten" X, den wir mit $K[X]$ bezeichnen. Jedes $f \in K[X]$ besitzt somit eine eindeutige Darstellung der Form $f = f_0 + f_1 X + \cdots + f_n X^n$ mit $f_n \neq 0$; $n = \deg(f)$ heißt der Grad von f, f_n der Führungskoeffizient und f_0 das konstante Glied. Polynome können zudem multipliziert werden. Mit dieser zusätzlichen Operation wird $K[X]$ zu einer assoziativen und kommutativen K-Algebra. Wir benötigen noch einige (sicher schon aus der Schule vertraute) Fakten über Polynome; diese Fakten waren Teil der LV "Zahlentheorie" aus dem SS13, die wir somit für alle LV-Teilnehmer nun als bekannt voraussetzen wollen.

1.3.1 **Satz** (Divisionsalgorithmus) Seien $f, g \in K[X]$ mit $\deg f = n, \deg g = m$. Dann gibt es eindeutig bestimmte Polynome s, r, sodass $f = sg + r$ mit $r = 0$ oder $\deg r < m$.

Beweis. Sei $f = a_0 X^n + \cdots + a_n$ und $g = b_0 X^m + \cdots + b_m$. Es sind zwei Fälle zu unterscheiden: $n < m$: Setze $s := 0, r := f$, so folgt die Aussage. $n \geq m$: Ansatz $s := c_0 X^{n-m} + \cdots + c_{n-m}, r := d_0 X^{m-1} + \cdots + d_{m-1}$. Koeffizientenvergleich liefert dann das folgende lineare Gleichungssystem in den Unbekannten $c_0, \ldots, c_{n-m}, d_0, \ldots, d_{m-1}$: $b_0 c_0 = a_0, b_1 c_0 + b_0 c_1 = a_1, \ldots, \cdots + b_1 c_{n-m-1} + b_0 c_{n-m} = a_{n-m}, \cdots + b_2 c_{n-m-1} + b_1 c_{n-m} + d_0 = a_{n-m+1}, \ldots, b_m c_{n-m} + d_{m-1} = a_n$. Die Koeffizientenmatrix ist eine untere Dreiecksmatrix und wegen $b_0 \neq 0$ besitzt diese Matrix den Rang $n+1$. Somit gibt es genau eine Lösung. $\square$

Der Divisionsalgorithmus entspricht dem aus der Schule geläufigen Division mit Rest in der Menge $\mathbb{Z}$. Wir übernehmen von dort auch die folgende geläufige Sprechweise:

1.3.2 **Definition** Es seien $f, g \in K[X]$ gegeben. Falls es ein $h \in K[X]$ mit der Eigenschaft $f = gh$ gibt, so heisst f ein *Vielfaches* von g und g ein *Teiler* von f. Ein Polynom $f \in K[X]$ mit $\deg f \geq 1$ heisst *irreduzibel*, falls aus $f = gh$ mit $g, h \in K[X]$ stets $g \in K$ oder $h \in K$ folgt.

1.3.3 **Satz** Sei $f \in K[X]$. Es ist $\lambda \in K$ genau dann Nullstelle von f, falls $X - \lambda$ ein Teiler von f ist.

Beweis. Sei λ eine Nullstelle von f. Falls $f = 0$ so ist $X - \lambda$ ein Teiler von f. Angenommen $f \neq 0$; dann gibt es wegen 1.3.1 Polynome s, r mit $f(X) = s(X)(X - \lambda) + r(X)$. Dabei gilt $r = d_0 \in K$. Indem wir X durch λ ersetzen folgt $0 = f(\lambda) = s(\lambda)(\lambda - \lambda) + d_0$ und somit $d_0 = r = 0$. Gilt umgekehrt $f(X) = s(X)(X - \lambda)$, so ist λ Nullstelle von f. $\square$

1.3.4 **Definition** Sei $\lambda \in K$ eine Nullstelle von $f \in K[X]$ mit $\deg f \geq 1$. Unter der *Vielfachheit* der Nullstelle λ verstehen wir die grösste natürliche Zahl m mit der Eigenschaft, dass $(X - \lambda)^m$ ein Teiler von f ist.

1.3.5 **Satz** Sei $f \in K[X]$ ein Polynom mit $\deg f = n \geq 1$ und r verschiedenen Nullstellen $\lambda_1, \ldots, \lambda_r \in K$; es bezeichne m_i die Vielfachheit von λ_i für $i \in \{1, \ldots, r\}$. Dann ist das Polynom f durch

$$(X - \lambda_1)^{m_1} \cdot \ldots \cdot (X - \lambda_r)^{m_r}$$

teilbar. Es folgt daraus $m_1 + \ldots + m_r \leq n$. Insbesondere hat f höchstens n verschiedene Nullstellen.

Beweis. Mehrfache Anwendung von 1.3.1 liefert $f(X) = (X-\lambda_1)^{n_1} \cdots (X-\lambda_r)^{n_r} s(X)$ mit $0 \leq n_i \leq m_i$ und einem Polynom s ohne Nullstellen bei $\lambda_1, \dots, \lambda_r$. Wir dividieren beide Seiten durch $(X - \lambda_1)^{n_1}$. Das ergibt $s_1(X) = (X - \lambda_2)^{n_2} \cdots (X - \lambda_r)^{n_r} s(X)$. Wegen $s_1(\lambda_1) \neq 0$ ist λ_1 keine Nullstelle von s_1 also $n_1 = m_1$. Diese Überlegung ist für $\lambda_2, \dots, \lambda_r$ zu wiederholen. Der Rest ist klar. □

1.3.6 Lässt sich ein Polynom $f \in K[X], \deg f = n \geq 1$ in der Form

$$f = a_0(X - \lambda_1)^{m_1} \cdot \ldots \cdot (X - \lambda_r)^{m_r}$$

mit $\lambda_1, \dots, \lambda_r \in K$ schreiben, so sagen wir, dass f über K in Linearfaktoren zerfällt. OBdA sind $\lambda_1, \dots, \lambda_r$ verschieden, dann sind $m_1, \dots, m_r$ die Vielfachheiten und es gilt $m_1 + \ldots + m_r = n$. Die Darstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

1.3.7 Wir nennen einen Körper K *algebraisch abgeschlossen*, falls jedes nicht-konstante Polynom über K in Linearfaktoren zerfällt. Wir verwenden hier ohne Beweis (siehe Algebra-Vorlesung): Zu jedem Körper K existiert ein Körper $\overline{K} \supseteq K$, der algebraisch abgeschlossen ist (genannt der algebraische Abschluss von K). Der Körper der komplexen Zahlen $\mathbb{C}$ ist algebraisch abgeschlossen (Fundamentalsatz der Algebra). Der Körper $\mathbb{R}$ ist nicht algebraisch abgeschlossen, z.B. $X^2 + 1$ zerfällt nicht in Linearfaktoren; jedes Polynom $f \in \mathbb{R}[X]$ zerfällt aber in Linearfaktoren oder quadratische Faktoren, d.h.

$$f = a_0 \prod_{i=1}^{l} (X - \lambda_i) \cdot \prod_{i=l+1}^{n} q_i$$

mit $q_i = X^2 + a_i X + b_i, a_i, b_i \in \mathbb{R}$ irreduzibel.

1.4 Algebraische und geometrische Vielfachheit

1.4.1 Die Dimension des Lösungsraumes dieses LGS wird auch als *geometrische Vielfachheit* von λ, die Vielfachheit von λ als Nullstelle im charakteristischen Polynom von φ als *algebraische Vielfachheit* von λ bezeichnet.

1.4.2 **Satz** Sei $\lambda \in K$ Eigenwert von $\varphi \in \mathrm{End}(V)$. Dann ist die geometrische Vielfachheit stets kleiner oder gleich der algebraischen Vielfachheit von λ.

Beweis. Sei $v_1, \dots, v_r$ Basis von V_λ. Ergänze diese durch $v_{r+1}, \dots, v_n$ zu einer Basis β von V. Dann gilt $\varphi(v_i) = \lambda v_i$ für $i = 1 \dots, r$. Indem wir die daraus resultierende Darstellungsmatrix $M_{\beta\beta}(\varphi)$ betrachten, folgt $(X-\lambda)^r | P_\varphi$ und daher ist $\dim V_\lambda = r$ kleiner oder gleich der algebraischen Vielfachheit von λ. □

1.4.3 Falls V endlich-dimensional ist, so besitzt φ also höchstens endlich viele Eigenwerte, $\mathrm{Spec}(\varphi)$ ist also endlich, und es gilt

$$\dim \sum_{\lambda \in \mathrm{Spec}(\varphi)} V_\lambda \leq \sum_{\lambda \in \mathrm{Spec}(\varphi)} \dim V_\lambda \leq \dim V.$$

Es ist sogar die Anzahl der Eigenwerte von φ gezählt mit Vielfachheiten beschränkt durch $\dim V$.

1.4.4 **Satz** Sei $\varphi \in \mathrm{End}(V)$ und $M \subset V$ eine linear abhängige Menge von Eigenvektoren von φ. Dann gehören mindestens zwei verschiedene Vektoren von M zum selben Eigenwert. Somit gilt: Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Beweis. Wähle aus M eine Basis $m_1, \ldots, m_s$ von $\langle M \rangle$ und bezeichne die zugehörigen Eigenwerte mit $\lambda_1, \ldots, \lambda_s$. Da M linear abhängig ist, folgt $s \geq 1$ und es existiert ein $m_{s+1} \in M \backslash \{m_1, \ldots, m_s\}$ zum Eigenwerte λ_{s+1}. Es folgt $m_{s+1} = \sum_{i=1}^s \mu_i m_i$ und daher $\sum_{i=1}^s \mu_i \lambda_i m_i = \varphi(m_{s+1}) = \lambda_{s+1} m_{s+1} = \sum_{i=1}^s \mu_i \lambda_{s+1} m_i$. Da wenigstens für ein $i \in \{1, \ldots, s\}$ gilt $\mu_i \neq 0$, folgt $\lambda_{s+1} = \lambda_i$. $\square$

1.4.5 Insbesondere folgt

$$\sum_{\lambda \in \operatorname{Spec}(\varphi)} V_\lambda = \bigoplus_{\lambda \in \operatorname{Spec}(\varphi)} V_\lambda.$$

Es folgt ausserdem: Falls $|\operatorname{Spec}(\varphi)| = \dim V$, so gilt wegen $\dim V \geq \dim \bigoplus_\lambda V_\lambda = \sum_\lambda \dim V_\lambda \geq |\operatorname{Spec}(\varphi)| = n$ die Identität

$$V = \bigoplus V_\lambda, \quad V_\lambda = K v_\lambda, \quad \lambda \in \operatorname{Spec}(\varphi).$$

Sind die Nullstellen von P_φ alle paarweise verschieden und liegen sie alle in K, so besitzt V eine Basis aus Eigenvektoren.

Beispiele Wir berechnen die Eigenvektoren und Eigenräume von

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{C}), \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(K), \quad \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \in M_2(\mathbb{R})$$

Für Details siehe Vorlesung.

§2. Die Jordansche Normalform

Sei K ein Körper und V ein n-dimensionaler Vektorraum über K mit $1 \leq n < \infty$. Bisher: $\varphi \in \operatorname{Hom}(V, W)$ mit $M_{\gamma\beta}(\varphi) = A \in M_{m,n}(K)$. Gesucht waren Basen β, γ, sodass A eine möglichst einfache Form hat! Es gilt: Es gibt Basen β', γ' (welche durch elementare Zeilen- und Spaltenumformungen gefunden werden können), sodass $M_{\gamma'\beta'}(\varphi) = I$ mit

$$I = \left(\begin{array}{ccc|c} & & & 0 \\ & E_r & & \vdots \\ & & & 0 \\ \hline 0 & \cdots & 0 & 0 \end{array}\right)$$

mit $r = \operatorname{rg}(\varphi) = \operatorname{rg} A$. Somit gibt es Basen $v'_1, \ldots, v'_n, w'_1, \ldots, w'_m$ von V bzw. W mit

$$\varphi(v'_i) = \begin{cases} w'_i, & 1 \leq i \leq \operatorname{rg}(\varphi), \\ 0, & \text{sonst.} \end{cases}$$

Jetzt $\varphi \in \operatorname{End}(V)$ mit $M_{\beta\beta}(\varphi) \in M_n(K)$. Gesucht: Eine Basis γ von V mit $M_{\gamma\gamma}(\varphi) = J$ möglichst einfach! Fragen: Wie sieht J aus? Wie lautet γ und wie berechnet man γ und J?

2.1 Diagonalisierbarkeit

Wir sehen uns nun die Situation vom Ende des letzten Kapitels etwas genauer an. Wir diskutieren also Endomorphismen φ von V für die es eine Koordinatenmatrix $M_{\beta\beta}(\varphi)$ gibt, welche Diagonalgestalt mit Diagonaleinträgen $\lambda_1, \ldots, \lambda_n$ hat. Wir nennen eine Matrix $D = (\lambda_i \delta_{ij}) \in M_n(K)$ eine *Diagonalmatrix*; wir schreiben $D =: \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Es sind dann die λ_i die Eigenwerte von φ und die zur Basis β gehörenden Vektoren b_i die zugehörigen Eigenvektoren.

2.1.1 **Definition** $\varphi \in \mathrm{End}(V)$ heisst *diagonalisierbar*, wenn V eine Basis bestehend aus Eigenvektoren von φ besitzt. Eine Matrix $A \in M_n(K)$ heisst *diagonalisierbar*, falls sie ähnlich zu einer Diagonalmatrix ist.

2.1.2 **Satz** Eine Abbildung $\varphi \in \mathrm{End}(V)$ ist genau dann diagonalisierbar, wenn eine Koordinatenmatrix von φ diagonalisierbar ist.

Beweis. Ist φ diagonalisierbar, so existiert eine Basis aus Eigenvektoren β von V und die Koordinatenmatrix $M_{\beta\beta}(\varphi)$ ist dann eine Diagonalmatrix. Sei umgekehrt $M_{\gamma\gamma}(\varphi)$ diagonalisierbar für eine Basis γ von V. Dann gibt es eine reguläre Matrix P mit $PM_{\gamma\gamma}(\varphi)P^{-1} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Somit existiert eine Basis β von V mit Matrix des Basiswechsels P. Die Abbildung φ ist daher diagonalisierbar. $\square$

Beispiel Die Matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(K)$$

zeigt, dass nicht jede Matrix diagonalisierbar ist.

Notwendig für die Diagonalisierbarkeit ist offenbar, dass das charakteristische Polynom in Linearfaktoren zerfällt; die Umkehrung ist jedoch i.A. nicht richtig. Eine hinreichende Bedingung für Diagonalisierbarkeit liefert der folgende:

2.1.3 **Satz** Sei $\varphi \in \mathrm{End}(V)$ und zerfalle das charakteristische Polynom $P_\varphi(X)$ in Linearfaktoren mit n verschiedenen Nullstellen, dann ist φ diagonalisierbar. Sei $A \in M_n(K)$ und zerfalle das charakteristische Polynom $P_A(X)$ in Linearfaktoren mit n verschiedenen Nullstellen, dann ist A diagonalisierbar.

Beweis. Sei $P_\varphi(X) = (-1)^n(X-\lambda_1)\cdots(X-\lambda_n)$ und $v_1, \ldots, v_n \neq 0$ zugehörige Eigenvektoren. Nach 1.4.4 ist $\{v_1, \ldots, v_n\}$ linear unabhängig und somit eine Basis bestehend aus Eigenvektoren. Der zweite Teil ist klar. $\square$

Die Umkehrung des letzten Satzes gilt ebenfalls nicht, wie z.B. $\varphi = \mathrm{id}_V$ zeigt. Wir zeigen nun ein Kriterium für Diagonalisierbarkeit:

2.1.4 **Satz** Sei $\varphi \in \mathrm{End}(V)$. Dann ist φ diagonalisierbar $\Leftrightarrow$ $P_\varphi(X) = (-1)^n(X-\lambda_1)^{n_1}\cdots(X-\lambda_r)^{n_r}$ mit verschiedenen $\lambda_1, \ldots, \lambda_r \in K$ und $n_i = \dim V_{\lambda_i}$ für $i = 1, \ldots, r$ (also stets algebraische gleich geometrische Vielfachheit gilt). Dies ist also genau dann der Fall, wenn $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$ gilt.

Beweis. Ist φ diagonalisierbar, so gibt es eine Basis aus Eigenvektoren; seien $\lambda_1, \ldots, \lambda_r$ die verschiedenen Eigenvektoren von φ und $v_1^{(i)}, \ldots, v_{s_i}^{(i)}$ eine Basis des zu λ_i zugehörigen Eigenraumes V_{λ_i} für $i = 1, \ldots, r$. Es gilt also $s_1 + \cdots + s_r = n = \dim V$. Andererseits gilt für die algebraischen Vielfachheiten $n_1, \ldots, n_r$ von $\lambda_1, \ldots, \lambda_r$, dass $n_1 + \cdots + n_r = n$. Nach 1.4.2 ist $s_i \leq n_i$ und daher muss $s_i = n_i$ für $i = 1, \ldots, r$ gelten. Jetzt zur Umkehrung: Da $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r} \subseteq V$ und nach Voraussetzung dieser Unterraum die Dimension $n_1 + \cdots + n_r = n$ hat, gibt es eine Basis von V bestehend aus Eigenvektoren, womit φ diagonalisierbar ist. $\square$

2.1.5 Die analoge Aussage von 2.1.4 für Matrizen lässt sich leicht entsprechend formulieren: eine Matrix $A \in M_n(K)$ ist diagonalisierbar genau dann, wenn das charakteristische Polynom von A über

K in Linearfaktoren zerfällt und für alle Eigenwerte die algebraische und geometrische Vielfachheit gleich ist.

2.2 Der Satz von Cayley-Hamilton

2.2.1 $\mathrm{End}(V)$ ein n^2-dimensionaler Vektorraum. Wir können auf $\mathrm{End}(V)$ ausserdem eine Multiplikation definieren, indem wir $(\varphi\psi)(v) := (\varphi \circ \psi)(v) = \varphi(\psi(v))$ setzen (also die Komposition verwenden). Es zeigt sich, dass $\mathrm{End}(V)$ mit der Vektoraddition und dieser Multiplikation einen Ring mit Einselement (nämlich id_V) bildet, wobei die Multiplikation gemäss $\lambda(\varphi\psi) = (\lambda\varphi)\psi = \varphi(\lambda\psi)$ mit der skalaren Multiplikation verträglich ist. Man sagt, dass $\mathrm{End}(V)$ eine assozitative K-Algebra mit Einselement ist. Sei $\varphi \in \mathrm{End}(V)$. Es gibt dann für die Folge $\varphi^0 = \mathrm{id}_V, \varphi, \varphi^2, \ldots$ einen kleinsten Exponenten s mit $1 \leq s \leq n^2$ sodass φ^s durch $\varphi^0, \varphi, \varphi^2, \ldots, \varphi^{s-1}$ linearkombiniert werden kann. Es ist dann $\varphi^0, \varphi, \varphi^2, \ldots, \varphi^{s-1}$ linear unabhängig und daher existieren eindeutig bestimmte Skalare mit $\varphi^s + \eta_1\varphi^{s-1} + \ldots + \eta_s\varphi^0 = 0 = 0_V$. Bilden wir das Polynom

$$M_\varphi(X) := X^s + \eta_1 X^{s-1} + \ldots + \eta_s,$$

so ist $M_\varphi(\varphi) = 0$.

2.2.2 **Definition** Sei $f \in K[X]$. Wir nennen f ein *Annullatorpolynom* von φ, falls $f(\varphi) = 0$. Das *Minimalpolynom* $M_\varphi(X)$ von φ ist jenes normierte Annullatorpolynom von φ, welches den kleinsten Grad besitzt.

2.2.3 Die Begriffe Annullator- und Minimalpolynom lassen sich sofort auf $M_n(K)$ übertragen; ähnliche Matrizen haben diesselben Annullatorpolynome und dasselbe Minimalpolynom.

2.2.4 **Satz** Sei $M_\varphi(X)$ das Minimalpolynom von $\varphi \in \mathrm{End}(V)$. Dann sind die Annullatorpolynome von φ genau die Vielfachen von $M_\varphi(X)$.

Beweis. Sei f ein Annulatorpolynom. Der Fall $f = 0$ ist trivial, sodass wir $f \neq 0$ annehmen können. Dann ist $\deg f \geq \deg M_\varphi$ und es gibt Polynome s, r mit $p = sM_\varphi + r, \deg r < \deg M_\varphi$ oder $r = 0$. Aus $f(\varphi) = M_\varphi(\varphi) = 0$ folgt $r(\varphi) = 0$, was nach Definition von M_φ nur mit $r = 0$ möglich ist. Die Umkehrung ist trivial. □

Beispiele a) $\varphi = \mathrm{id}_V \Rightarrow M_\varphi = X - 1$, b) $\varphi = 0 \Rightarrow M_\varphi = X$, c) $\varphi \in \mathrm{End}(V)$ mit $\varphi \circ \varphi = \varphi, \varphi \neq 0, \mathrm{id}_V \Rightarrow M_\varphi = X(X-1)$.

2.2.5 **Satz** (Satz von Cayley-Hamilton) Für $\varphi \in \mathrm{End}(V)$ ist das charakteristische Polynom $P_\varphi(X)$ ein Annullatorpolynom von φ. Das Minimalpolynom $M_\varphi(X)$ ist ein Teiler des charakteristischen Polynoms von φ.

Beweis. Sei β eine Basis von V und $A = (\alpha_{ij}) = M_{\beta\beta}(\varphi)$. In den Spalten der Matrix A sind die Koordinaten der φ-Bilder der Basisvektoren $v_i = \beta^{-1}(e_i)$ gemäss $\varphi(v_j) = \sum_{i=1}^n \alpha_{ij}v_i$ abzulesen. Somit gilt (wobei wir $\psi \cdot v := \psi(v)$ setzen)

$$\begin{pmatrix} \varphi & 0 & \cdots & 0 \\ 0 & \varphi & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & & \cdots & \varphi \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \alpha_{11}\mathrm{id}_V & \cdots & \alpha_{n1}\mathrm{id}_V \\ \alpha_{12}\mathrm{id}_V & & \alpha_{n2}\mathrm{id}_V \\ \vdots & & \vdots \\ \alpha_{1n}\mathrm{id}_V & \cdots & \alpha_{nn}\mathrm{id}_V \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Beachte, dass die Koeffizienten der transponierten Matrix ${}^tA$ auftreten. Umformung ergibt

$$\begin{pmatrix} \alpha_{11}\mathrm{id}_V - \varphi & \cdots & \alpha_{n1}\mathrm{id}_V \\ \alpha_{12}\mathrm{id}_V & & \alpha_{n2}\mathrm{id}_V \\ \vdots & & \vdots \\ \alpha_{1n}\mathrm{id}_V & \cdots & \alpha_{nn}\mathrm{id}_V - \varphi \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = 0.$$

Betrachte die Matrix $B := {}^tA - \varphi E_n = (\beta_{ij}) \in M_n(K[\varphi])$. Ihre Determinante ist das charakteristische Polynom von φ, da transponierte Matrizen dieselbe Determinante haben. Multiplizieren wir beide Seiten der obigen Gleichung mit der Kofaktormatrix $\hat{B}$ von links, so folgt $P_\varphi(\varphi)(v_1) = \cdots = P_\varphi(\varphi)(v_n) = 0$, woraus $\mathrm{Ker}(P_\varphi(\varphi)) = V$ und somit P_φ Annulatorpolynom von φ folgt. Der Rest ist klar. □

Somit folgt zumindest $\deg M_\varphi \leq n$. Wie man das Minimalpolynom berechnet werden wir im nächsten Kapitel sehen.

Beispiele Die Matrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M_2(K)$$

haben dasselbe charakteristische Polynom $X^2 - 2X + 1$ aber die Minimalpolynome $X - 1$ bzw. $X^2 - 2X + 1$.

2.3 Die Normalform von Jordan

Wie am Anfang des Kapitels beschrieben ist das Ziel, alle Endomorphismen zu klassifizeren. Wir beginnen mit einer Vorbemerkung über gewissen Unterräume von V:

2.3.1 **Definition** Ein Unterraum $U \subseteq V$ heisst *φ-invariant*, falls $\varphi(U) \subseteq U$.

2.3.2 Summe und Durchschnitt φ-invarianter Unterräume sind klarerweise wieder φ-invariant. Insbesondere sind V und $\{0\}$ φ-invariant. Lässt sich V als direkte Summe φ-invarianter Unterräume $U_1 \oplus \cdots \oplus U_r$ mit $U_j \neq \{0\}$ für $j \in \{1, \ldots, r\}$ schreiben, so besitzt φ eine Koordinatenmatrix B in der Form einer *erweiterten Diagonalmatrix*, nämlich eine Matrix, die aus Diagonalblöcken $A_1, \ldots, A_r$ aufgebaut ist, wobei die $A_j \in M_{n_j}(K)$ mit $n_j = \dim U_j$ sind; man muss ja nur Basen der Unterräume U_i in der richtigen Reihenfolge zu einer Basis von V zusammensetzen (und die Spalten der Koordinatenmatrix wie üblich deuten). Ausserdem gilt $P_B(X) = P_{A_1}(X) \cdots P_{A_r}(X)$.

2.3.3 Wir betrachten nun für $\varphi \in \mathrm{End}(V)$ mit Eigenwert $\lambda \in K$ die Vektorräume

$$V_m = \mathrm{Ker}(\varphi - \lambda\mathrm{id})^m.$$

Es gilt $\{0\} = V_0 \subseteq V_1 \subseteq \ldots \subseteq V_m \subseteq V_{m+1} \subseteq \ldots \subseteq V$.

2.3.4 **Definition** Ein Vektor $w \in V$ heisst ein *Hauptvektor* von φ zum Eigenwert λ, falls mindestens ein $m \in \mathbb{N}$ mit $w \in V_m$ existiert. Das kleinste m mit dieser Eigenschaft heisst die *Stufe* von w. Der zum Eigenwert λ gehörige *Hauptraum* $V_\varphi(\lambda)$, ist die Menge aller Hauptvektoren von φ zum Eigenwert λ, also

$$V_\varphi(\lambda) = \bigcup_{m \geq 0} V_m \subseteq V.$$

Beispiele Die Hauptvektoren der Stufe 1 sind genau die Eigenvektoren von φ zum Eigenwert λ; der einzige Hauptvektor der Stufe 0 ist der Nullvektor.

2.3.5 **Satz** Der Hauptraum $V_\varphi(\lambda)$ ist ein φ-invarianter Unterraum von V.

Beweis. Es gilt $0 \in V_\varphi(\lambda) \neq \emptyset$. Seien $w_1, w_2 \in V_\varphi(\lambda)$ Hauptvektoren der Stufen m_1, m_2 und $\mu \in K$. Beachte die Inklusionskette $\{0\} = \operatorname{Ker}(\varphi - \lambda \mathrm{id}_V)^0 \subset \operatorname{Ker}(\varphi - \lambda \mathrm{id}_V)^1 \subset \cdots \subset \operatorname{Ker}(\varphi - \lambda \mathrm{id}_V)^m \subset \cdots \subset V_\varphi(\lambda)$. Für hinreichend grosses m ist daher $w_1, w_2 \in \operatorname{Ker}(\varphi - \lambda \mathrm{id}_V)^m$ und somit $w_1 + w_2 \in V_\varphi(\lambda)$. Es ist klarerweise $\mu w \in V_\varphi(\lambda)$. Also ist $V_\varphi(\lambda)$ ein Unterraum von V. Wähle $w \in V_\varphi(\lambda)$; sei m die Stufe von w. Dann zeigt $(\varphi - \lambda \mathrm{id}_V)^m(\varphi(w)) = (\varphi \circ (\varphi - \lambda \mathrm{id}_V)^m)(w) = \varphi(0) = 0$ die Gültigkeit von $\varphi(w) \in V_\varphi(\lambda)$, womit es sich um eine φ-invarianten Unterraum handelt. □

2.3.6 **Satz** Sei w ein Hauptvektor von φ zum Eigenwert $\lambda \in K$ und $m \geq 1$ die Stufe von w. Dann ist $(\varphi - \lambda \mathrm{id})^{m-1}(w)$ ein Eigenvektor von φ zum Eigenwert λ.

Beweis. Wegen $((\varphi - \lambda \mathrm{id}_V) \circ (\varphi - \lambda \mathrm{id}_V)^{m-1})(w) = 0$ ist $(\varphi - \lambda \mathrm{id}_V)^{m-1}(w) \neq 0$ Eigenvektor von φ zum Eigenwert λ. □

2.3.7 **Satz** Seien k verschiedene Hauptvektoren $w_1, \ldots, w_k$ von φ zum Eigenwert λ mit Stufen $m_1 \geq m_2 \geq \ldots \geq m_k \geq 1$ derart gegeben, dass die Vektoren $(\varphi - \lambda \mathrm{id})^{m_j - 1}(w_j), j \in \{1, \ldots, k\}$ genau k linear unabhängige Vektoren sind. Dann sind die Vektoren

$$v_{ji} := (\varphi - \lambda \mathrm{id})^{m_j - i}(w_j) \text{ mit } j \in \{1, \ldots, k\}, i \in \{1, \ldots, m_j\}$$

verschieden und linear unabhängig.

Beweis. Es ist jeder Vektor v_{ji} ein Hauptvektor von φ zum Eigenwert λ mit Stufe i. Wir erhalten also folgendes Schema; dabei bedeutet $\mapsto$ die Anwendung der Abbildung $\varphi - \lambda\mathrm{id}$ (und wir haben zur Demonstration $m_1 = m_2 > m_3 \geq \cdots > m_k \geq 1$ gesetzt):

$$\begin{array}{ccccccccccccccccc}
w_1 & = & v_{1m_1} & \mapsto & v_{1m_1-1} & \mapsto & \cdots & \mapsto & v_{1m_3} & \mapsto & \cdots & \mapsto & v_{1m_k} & \mapsto & \cdots & \mapsto & v_{11} & \mapsto & 0 \\
w_2 & = & v_{2m_2} & \mapsto & v_{1m_2-1} & \mapsto & \cdots & \mapsto & v_{2m_3} & \mapsto & \cdots & \mapsto & v_{2m_k} & \mapsto & \cdots & \mapsto & v_{21} & \mapsto & 0 \\
w_3 & = & & & & & & & v_{3m_3} & \mapsto & \cdots & \mapsto & v_{3m_k} & \mapsto & \cdots & \mapsto & v_{31} & \mapsto & 0 \\
 & \vdots & & & & & & & & & & & & & & & & & \\
w_k & = & & & & & & & & & & & v_{km_k} & \mapsto & \cdots & \mapsto & v_{k1} & \mapsto & 0
\end{array}$$

Die Vektoren v_{ji} sind paarweise verschieden: Für Vektoren verschiedener Stufe ist das trivial; Vektoren derselben Stufe, etwa v_{1i} und v_{2i}, können auch nicht gleich sein, da sie unter $(\varphi - \lambda \mathrm{id}_V)^{i-1}$ in die verschiedenen Eigenvektoren v_{11} und v_{21} übergehen. Betrachte nun eine beliebige Linearkombination

$$\sum_{j=1}^{k} \sum_{i=1}^{m_j} \mu_{ji} v_{ji} = 0$$

mit $\mu_{ji} \in K$. Wende auf diese Linearkombination $(\varphi - \lambda \mathrm{id}_V)^{m_1 - 1}$ and. Alle v_{ji} mit $i < m_1$ gehen in den Nullvektor über. Das zeigt $\sum_j \mu_{jm_1} v_{j1} = 0$, wobei j jene Teilmenge von $\{1, \ldots, k\}$ durchläuft für die w_j Hauptvektor der Stufe m_1 ist. Nach Voraussetzung sind daher alle μ_{jm_1} gleich 0. Nun wird $(\varphi - \lambda \mathrm{id}_V)^{m_1 - 2}$ angewendet. Das zeigt $\mu_{jm_1 - 1} = 0$, usw. Damit liegt insgesamt eine triviale Linearkombination vor. □

2.3.8 Wir übertragen dies nun auf Matrizen. Ein obere Dreiecksmatrix der Gestalt

$$J_m(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \in M_m(K)$$

mit $m \geq 1, \lambda \in K$ heisst *Jordan-Matrix*. Die Bedeutung der Jordan-Matrix sieht man so: Sei w ein Hauptvektor der Stufe m zum Eigenwert λ. Setzen wir

$$v_i := (\varphi - \lambda \mathrm{id})^{m-i}(w) \text{ für } i \in \{1, \dots, m\},$$

dann ist $v_1, \dots, v_m$ eine Basis β eines φ-invarianten Unterraumes $T \subseteq V$, denn es ist

$$\begin{aligned} \varphi(v_1) &= \lambda \mathrm{id}(v_1) + (\varphi - \lambda \mathrm{id})(v_1) = 0 + \lambda v_1 = \lambda v_1 \\ \varphi(v_2) &= \lambda \mathrm{id}(v_2) + (\varphi - \lambda \mathrm{id})(v_2) = v_1 + \lambda v_2 \\ &\vdots \\ \varphi(v_3) &= \lambda \mathrm{id}(v_3) + (\varphi - \lambda \mathrm{id})(v_3) = v_2 + \lambda v_3 \\ &\vdots \\ \varphi(v_m) &= \lambda \mathrm{id}(v_m) + (\varphi - \lambda \mathrm{id})(v_m) = v_{m-1} + \lambda v_m \end{aligned}$$

Die Einschränkung von φ auf T wird also bzgl. β durch $J_m(\lambda)$ beschrieben. Umgekehrt bestimmt $J_m(\lambda)$ eine lineare Abbildung bei der die kanonische Basis die Rolle der v_i von oben übernehmen. Geometrisch bedeutet $J_2(\lambda)$ im $\mathbb{R}^2$ eine sogenannte Scherung.

2.3.9 Um die gesuchte Normalform zu erhalten, müssen wir eine Situation wie in 2.3.5 erreichen. Dazu starten wir mit einer Basis des Eigenraumes (dies ist für jeden Eigenvektor zu machen) und suchen, durch Lösen von linearen Gleichungssystemen bzgl. der Abbildung $\varphi - \lambda \mathrm{id}$, zu jedem Eigenvektor in einer Basis des Eigenraumes eine möglichst lange Kette von (Haupt-)Vektoren wie im letzten Satz. Da die durch die Ketten gegebenen Unterräume des Hauptraumes zum Eigenwert λ φ-invariant sind, erhalten wir als Koordinatenmatrix von φ auf dem Hauptraum $V_\varphi(\lambda)$ bzgl. der so gefundenen Basis (Achtung bei der Sortierung der Vektoren; wir müssen die Vektoren in den Ketten von rechts nach links ordnen) eine erweiterte Diagonalmatrix, dessen Diagonalblöcke einzig aus Jordan-Matrizen besteht. Wir geben ein Beispiel bevor wir die Situation in einem Satz zusammenfassen:

Beispiel Sei

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

Es ist 0 der einzige Eigenwert von A mit algebraischer Vielfachheit 3. Zur Bestimmung des Eigenraumes müssen wir das LGS $Ax = 0$ lösen; der Rang von $A = A - 0E_3$ ist 2 und daher die geometrische Vielfachheit von 0 gleich 1. Somit gibt es in der Situation von 2.3.5 genau eine Kette $v_{13} \mapsto v_{12} \mapsto v_{11}$ und die Koordinatenmatrix der durch A gegebenen linearen Abbildung bzgl. der Basis v_{11}, v_{12}, v_{13} ist $J_3(0)$. Der Eigenraum wird durch $v_{11} = e_1$ aufgespannt. Wir berechnen noch die Hauptvektoren v_{12} der Stufe 2 bzw. v_{13} der Stufe 3: Dazu lösen wir das LGS $Ax = v_{11} = e_1$, was zu $v_{12} = x = e_2$

führt, bzw. das LGS $Ax = v_{12} = e_2$, was zu $x = v_{13} = {}^t(0, -3/2, 1/2)$ führt. Somit gilt mit

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

dass $P^{-1}AP = J_3(0)$.

2.3.10 **Satz von der Jordan-Normalform** Sei zunächst $P_\varphi(X) = (-1)^n(X - \lambda)^n$. Dann gilt:

a) Es gibt eine Basis β von V, sodass $M_{\beta\beta}(\varphi)$ eine erweiterte Diagonalmatrix mit Diagonalblöcken $A_1, \ldots, A_p$ ist, wobei A_q für $q \in \{1, \ldots, p\}$ selbst eine erweitere Diagonalmatrix mit Diagonalblöcken $J_{m_q}(\lambda), \ldots, J_{m_q}(\lambda)$ bestehend aus $k_q \geq 1$ gleichen Jordan-Matrizen ist. Dabei kann $m_1 > \ldots > m_p \geq 1$ vorausgesetzt werden und für die geometrische Vielfachheit von λ gilt

$$\dim \operatorname{Ker}(\varphi - \lambda \mathrm{id}) = k_1 + \cdots + k_p.$$

b) Ist A eine Koordinatenmatrix von φ in Form einer aus Jordan-Matrizen aufgebauten erweiterten Diagonalmatrix, so unterscheidet sich A von der oben angegebenen Matrix nur in der Reihenfolge der auftretenden Jordan-Matrizen.

Sei nun allgemein $P_\varphi(X) = (-1)^n(X - \lambda_1)^{n_1} \cdots (X - \lambda_r)^{n_r}$ mit $r \geq 1$ verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in K$. Dann ist $V = V_\varphi(\lambda_1) \oplus \cdots \oplus V_\varphi(\lambda_r)$ und $\dim V_\varphi(\lambda_j) = n_j$ für $j \in \{1, \ldots, r\}$, und ausserdem gilt:

c) Es gibt eine Basis von V, sodass die Koordinatenmatrix von φ bzgl. dieser Basis eine Diagonalmatrix mit Diagonalblöcken $B_1, \ldots, B_r$ ist, wobei B_j für $j \in \{1, \ldots, r\}$ eine wie im ersten Teil aufgebaute Matrix in $M_{n_j}(K)$ ist.

d) Ist B eine Koordinatenmatrix von φ in Form einer aus Jordan-Matrizen augebauten erweiterten Diagonalmatrix, so unterscheidet sich B von der oben angegebenen Matrix nur in der Reihenfolge der auftretenden Jordan-Matrizen.

Beweisskizze. Der Nachweis für die "Hauptraumzerlegung" wird mit Hilfe des Minimalpolynoms geführt und verwendet 2.2.3. Für die Einschränkung von φ auf den Hauptraum argumentiert man mit 2.3.5 indem man wieder das Minimalpolynom zu Hilfe nimmt; dieser Teil ist konstruktiv und somit ist klar, wie man die Basis und die Normalform bestimmt. Für Details siehe das Zusatzblatt. □

2.3.11 Beachte, dass das charakteristische Polynom in Linearfaktoren zerfallen muss, was sicher passiert, wenn wir in den algebraischen Abschluss von K wechseln. Der letzte Satz stellt also sicher, dass das Rechenverfahren, dass wir bereits angedeutet haben funktioniert. Ein kleineres Problem liegt noch darin, dass a priori nicht klar ist, wie die Basis des Eigenraumes gewählt werden muss; findet man nicht genügend linear unabhängig Hauptvektoren, dann muss die Basis so abgeändert werden, dass die auftretenden Gleichungssysteme lösbar werden (und im konkreten Fall sieht man, wie die Basis geändert werden muss).

2.3.12 Wir sagen, dass die Matrix aus dem letzten Satz *Jordan-Normalform* besitzt. Wir fassen noch einmal die mit φ mitbestimmten wichtigen Grössen zusammen:

- Die Zahl r der verschiedenen Nullstellen $\lambda_1, \ldots, \lambda_r$ von $P_\varphi(X)$.

- Die algebraischen Vielfachheiten $n_1, \dots, n_r$ dieser Nullstellen; sie sind die Dimensionen der zugehörigen Haupträume.

- Für jeden Hauptraum $V_\varphi(\lambda_j)$ die Zahlenpaare (k_{ji}, m_{ji}) mit $j \in \{1, \dots, r\}, i \in \{1, \dots, p_j\}$; sie geben an, dass genau k_{ji} Jordan-Matrizen $J_{m_{ji}}(\lambda_j)$ auftreten. Dabei $m_{ji} > \cdots > m_{jp_j}$. Es ist $k_{j1} + \cdots + k_{jp_j} = \dim \operatorname{Ker}(\varphi - \lambda_j \mathrm{id})$ die geometrische Vielfachheit von λ_j, $k_{j1}m_{j1} + \cdots + k_{jp_j}m_{jp_j} = n_j$ die algebraische Vielfachheit von λ_j.

2.3.13 Wir können nun ausserdem das Minimalpolynom von φ angeben. Da jeweils $m_{11}, \dots, m_{r1}$ die grössten Zeilenzahlen der zu den Eigenwerten $\lambda_1, \dots, \lambda_r$ gehörigen Jordan-Matrizen sind, ergibt sich

$$M_\varphi(X) = (X - \lambda_1)^{m_{11}} \cdots (X - \lambda_r)^{m_{r1}}.$$

Als weitere Folgerung erhalten wir: Genau für

$$m_{11} = \cdots = m_{r1} = 1$$

ist φ diagonalisierbar, denn genau in diesem Fall treten in der Jordan-Normalform nur Jordan-Matrizen J_1 auf. Dies ist äquivalent dazu, dass das Minimalpolynom über K in Linearfaktoren mit lauter verschiedenen Nullstellen zerfällt.

2.3.14 Es folgt: Jede Matrix $A \in M_n(K)$ deren charakteristisches Polynom in K in Linearfaktoren zerfällt, ist ähnlich einer Matrix in Jordan-Normalform. Zwei Matrizen, deren charakteristisches Polynom je in K in Linearfaktoren zerfallen, sind genau dann ähnlich, wenn sie ähnlich zur selben Matrix in Jordan-Normalform sind.

Beispiel Sei V ein 6-dimensionaler reeller Vektorraum und $\varphi \in \operatorname{End}(V)$ gegeben durch

$$A = M_{\beta\beta}(\varphi) = \begin{pmatrix} 1 & -1 & 0 & 0 & 2 & 1 \\ 1 & 4 & 0 & 0 & -3 & -1 \\ 1 & 1 & 1 & 1 & -2 & -1 \\ 1 & 1 & -1 & 3 & -2 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

1. *Schritt* (Bestimmung der Eigenwerte von φ mit den algebraischen Vielfachheiten): Das charakteristische Polynom lautet $P_\varphi(X) = \det(A - XE_6) = (2 - X)^6$ und somit ist $\lambda_1 = 2$ der einzige Eigenwert mit algebraischer Vielfachheit $n_1 = 6$. Insbesondere ist A nach 2.3.14 ähnlich einer Matrix in Jordan-Normalform. In diesem Beispiel gibt es nur einen Eigenwert; sonst sind die folgenden Schritte für jeden Eigenwert durchzuführen.
2. *Schritt* (Bestimmung einer Basis des Eigenraumes und somit der geometrischen Vielfachheiten): Wir müssen dazu das lineare Gleichungssystem $(A - 2E_6)x = 0$ lösen; wir tun dies durch das folgende

Schema und verwenden beim Umformen nur elementare Zeilenumformungen:

$$E_6|(A-2E_6) = \left.\begin{array}{cccccc|cccccc} 1&0&0&0&0&0&-1&-1&0&0&2&1\\ 0&1&0&0&0&0&1&2&0&0&-3&-1\\ 0&0&1&0&0&0&1&1&-1&1&-2&-1\\ 0&0&0&1&0&0&1&1&-1&1&-2&-1\\ 0&0&0&0&1&0&0&1&0&0&-1&0\\ 0&0&0&0&0&1&0&-1&0&0&1&0 \end{array}\right.$$

$$\rightsquigarrow \begin{array}{cccccc|cccccc} 0&1&0&0&0&2&1&0&0&0&-1&-1\\ 0&0&0&0&0&-1&0&1&0&0&-1&0\\ 0&1&0&-1&0&1&0&0&1&-1&0&0\\ 1&1&0&0&0&1&0&0&0&0&0&0\\ 0&0&1&-1&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&1&0&0&0&0&0&0 \end{array} = T|T^*(A-2E_6)$$

Dabei gilt $T(A-2E_6) = T^*(A-2E_6)$ ist die Matrix in Treppennormalform. Die Matrix T hilft im folgenden Rechenschritte zu sparen. Wir erhalten als Basis des Eigenraumes $\{v_{11}, v_{21}, v_{31}\}$ mit $\Phi_\beta^{-1}(v_{11}) = {}^t(0,0,1,1,0,0), \Phi_\beta^{-1}(v_{21}) = {}^t(1,1,1,1,1,0)$ (beachte, dass wir aus Demonstrationszwecken nicht die kanonische Wahl ${}^t(1,1,0,0,1,0)$ verwenden), $\Phi_\beta^{-1}(v_{31}) = {}^t(1,0,0,0,0,1)$. Die geometrische Vielfachheit ist somit 3 und daher treten in der gesuchten Jordan-Normalform von A drei Jordan-Matrizen auf deren Grösse aber noch unbekannt ist.

3. *Schritt* (Bestimmung einer geeigneten Basis des Hauptraumes, sodass die Basis die Ketten im Beweis von Satz 2.3.7 liefert): Zunächst bestimmen wir eine maximale Menge von Vektoren, welche unter $\varphi - 2\mathrm{id}_V$ auf eine l.u. Menge von Eigenvektoren abgebildet werden. Dazu setzen wir an $x \mapsto v_{11}, x \mapsto v_{21}, x \mapsto v_{31}$, was zu drei linearen Gleichungssystemen mit stets derselben Koeffizientenmatrix $A-2E_6$ führt. Wenn wir Gausselimination verwenden brauchen wir nur die Koordinaten von v_{11}, v_{21}, v_{31} mit T zu transformieren und erhalten:

x_1	x_2	x_3	x_4	x_5	x_6	v_{11}	v_{21}	v_{31}
1	0	0	0	−1	−1	0	1	2
0	1	0	0	−1	0	0	0	−1
0	0	1	−1	0	0	−1	0	1
0	0	0	0	0	0	0	2	2
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1

Wir lesen ab, dass von den drei Eigenvektoren nur v_{11} im Bildraum von $\varphi - 2\mathrm{id}_V$ liegt. Da die 3×3-Teilmatrix rechts unten den Defekt 2 hat, können wir die Eigenbasis so verändern, dass maximal zwei Gleichungssysteme lösbar sind (beachte, dass wir im Beweis ja beginnend mit den Hauptvektoren höchster Stufe argumentiert haben, welche dann die Eigenbasis induziert haben und wir jetzt gerade anders herum starten müssen; es ist nicht garantiert, dass die frei gewählte Basis des Eigenraums bereits die Gesuchte ist!). Wir vertauschen v_{21} mit $\tilde{v}_{21} := v_{21} - v_{31}$ und erhalten $\Phi_\beta^{-1}(\tilde{v}_{21}) = {}^t(0,1,1,1,1,-1)$. Das gibt

x_1	x_2	x_3	x_4	x_5	x_6	v_{11}	$\tilde{v}_{21}$	v_{31}
1	0	0	0	−1	−1	0	−1	2
0	1	0	0	−1	0	0	1	−1
0	0	1	−1	0	0	−1	−1	1
0	0	0	0	0	0	0	0	2
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1

und daher die Hauptvektoren 2. Stufe $v_{12} \mapsto v_{11}, v_{22} \mapsto \tilde{v}_{21}$ mit $\Phi_\beta^{-1}(v_{12}) = {}^t(0,0,-1,0,0,0), \Phi_\beta^{-1}(v_{22}) = {}^t(0,1,-1,0,0,1)$ (wir wählen auch hier ganz bewusst und zu Demonstrationszwecken diesen Lösungsvektor anstatt der kanonischen Wahl ${}^t(-1,1,-1,0,0,0)$). Somit ist $\{v_{11}, \tilde{v}_{21}, v_{31}, v_{12}, v_{22}\}$ eine Basis von $\operatorname{Ker}(\varphi - 2\mathrm{id}_V)^2$. Als nächstes suchen wir eine maximale Menge von Vektoren, welche unter $\varphi - 2\mathrm{id}_V$ bijektiv auf eine l.u. Menge von Hauptvektoren 2. Stufe abgebildet werden. Wir setzen an $x \mapsto v_{11}, x \mapsto \tilde{v}_{21}, x \mapsto v_{31}, x \mapsto v_{12}, x \mapsto v_{22}$. Durch $T\Phi_\beta^{-1}(v_{12}), T\Phi_\beta^{-1}(v_{22})$ erhalten wir

x_1	x_2	x_3	x_4	x_5	x_6	v_{11}	$\tilde{v}_{21}$	v_{31}	v_{12}	v_{22}
1	0	0	0	-1	-1	0	-1	2	0	3
0	1	0	0	-1	0	0	1	-1	0	-1
0	0	1	-1	0	0	-1	-1	1	-1	2
0	0	0	0	0	0	0	0	2	0	2
0	0	0	0	0	0	0	0	0	-1	-1
0	0	0	0	0	0	0	0	1	0	1

Keines der beiden letzten Gleichungssystem ist lösbar; da der Defekt der 3×5-Matrix rechts unten Defekt gleich 1 hat können wir v_{22} durch $\tilde{v}_{22} := v_{22} - v_{12} - v_{31}$ tauschen, was ebenfalls ein Hauptvektor 2. Stufe ist. Wir notieren $\Phi_\beta^{-1}(\tilde{v}_{22}) = {}^t(-1,1,0,0,0,0)$. Nun aber Achtung! Da wir v_{22} verändert haben, müssen wir prüfen, ob wir die vorherigen Basisvektoren auch abzuändern haben; dazu berechnen wir $(\varphi - 2\mathrm{id}_V)(\tilde{v}_{22}) = (\varphi - 2\mathrm{id}_V)(v_{22} - v_{12} - v_{31}) = \tilde{v}_{21} - v_{11} =: \tilde{\tilde{v}}_{21}$, welcher klarerweise ein Eigenvektor ist; es ist $\Phi_\beta^{-1}(\tilde{\tilde{v}}_{21}) = {}^t(0,1,0,0,1,-1)$. Wir erhalten

x_1	x_2	x_3	x_4	x_5	x_6	v_{11}	$\tilde{\tilde{v}}_{21}$	v_{31}	v_{12}	$\tilde{v}_{22}$
1	0	0	0	-1	-1	0	-1	2	0	1
0	1	0	0	-1	0	0	1	-1	0	0
0	0	1	-1	0	0	-1	0	1	0	1
0	0	0	0	0	0	0	0	2	0	0
0	0	0	0	0	0	0	0	0	-1	0
0	0	0	0	0	0	0	0	1	0	0

und somit den Hauptvektor 3. Stufe $v_{23} \mapsto \tilde{v}_{22}$ mit $\Phi_\beta^{-1}(v_{23}) = {}^t(1,0,1,0,0,0)$. Wegen $\dim V_\varphi(2) = 6$ liegt mit $(v_{11}, \tilde{\tilde{v}}_{21}, v_{31}, v_{12}, \tilde{v}_{22}, v_{23})$ eine (allerdings noch falsch geordnete) Basis des Hauptraumes vor. Ordnen wir die Basis γ zu $(\tilde{\tilde{v}}_{21}, \tilde{v}_{22}, v_{23}, v_{11}, v_{12}, v_{31})$, dann ist

$$M_{\beta\gamma}(\mathrm{id}_V) = \left(\begin{array}{ccc|cc|c} 0 & -1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) =: S$$

und somit

$$J = S^{-1}AS = M_{\gamma\gamma}(\varphi) = \left(\begin{array}{ccc|cc|c} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right) = \operatorname{diag}(J_3(2), J_2(2), J_1(2)).$$

Der Beweis gibt also klar die Strategie zur Berechnung der Jordan-Normalform von φ vor.

§3. Euklidische und Unitäre Vektorräume

3.0 Vorbemerkungen

3.0.1 Wir setzen im folgenden stets voraus, dass $K = \mathbb{R}$ oder $K = \mathbb{C}$ gilt; V ist also stets ein reeller oder komplexer Vektorraum. Weiters bezeichnen wir mit ω immer einen Automorphismus von K, das ist eine bijektive Abbildung von K nach K mit $f(\lambda + \mu) = f(\lambda) + f(\mu), f(\lambda\mu) = f(\lambda)f(\mu)$ für alle $\lambda, \mu \in K$, mit $\omega^2 = \mathrm{id}_K$. Ein Beispiel ist die komplexe Konjugation auf $\mathbb{C}$, also $\omega(x + iy) = x - iy, x, y \in \mathbb{R}$; diese lässt sich natürlich auch auf $\mathbb{C}^n$ komponentenweise anwenden und macht auch auf $\mathbb{R}^n$ Sinn (ist dort aber nur die Identität).

3.0.2 Sei V ein n-dimensionaler Vektorraum über K und β eine geordnete Basis von V. "By abuse of language" schreiben setzen wir $\beta = \Phi_\beta^{-1}$; wir interpretieren also β als Abbildung $V \to K^n$, welche $v \in V$ die Koordinaten $\Phi_\beta^{-1}(v)$ in K^n zuordnet.

3.1 Skalarprodukte

3.1.1 Aus der Schule ist bekannt, dass man $x = {}^t(x_1, \ldots, x_n), y = {}^t(y_1, \ldots, y_n) \in \mathbb{R}^n$ eine Zahl

$$\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$$

zuordnet; diese wird das *Skalarprodukt* der Vektoren x, y genannt und hat folgende Eigenschaften: $\langle x+x', y\rangle = \langle x, y\rangle + \langle x', y\rangle, \langle x, y+y'\rangle = \langle x, y\rangle + \langle x, y'\rangle, \langle \lambda x, y\rangle = \langle x, \lambda y\rangle = \lambda\langle x, y\rangle, \langle x, y\rangle = \langle y, x\rangle$, sowie $\langle x, x\rangle \geq 0$ und $\langle x, x\rangle = 0 \Leftrightarrow x = 0$. Mit Hilfe des Skalarproduktes kann man eine Längenmessung (Norm) auf $\mathbb{R}^n$ durch $\|x\| = \sqrt{\langle x, x\rangle}$ einführen, sowie durch $\measuredangle(x, y) = \arccos(\langle x, y\rangle/(\|x\| \cdot \|y\|))$ Winkel messen. Auf $\mathbb{C}^n$ wird das Skalarprodukt durch

$$\langle x, y\rangle = x_1\overline{y_1} + \cdots + x_n\overline{y_n}$$

eingeführt; es erfüllt ähnliche Eigenschaften, nämlich $\langle x+x', y\rangle = \langle x, y\rangle + \langle x', y\rangle, \langle x, y+y'\rangle = \langle x, y\rangle + \langle x, y'\rangle, \langle \lambda x, y\rangle = \lambda\langle x, y\rangle, \langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle, \langle x, y\rangle = \overline{\langle y, x\rangle}$, sowie $\langle x, x\rangle \in \mathbb{R}_{\geq 0}$ und $\langle x, x\rangle = 0 \Leftrightarrow x = 0$, und kann ebenfalls zur Längen- und Winkelmessung eingesetzt werden.

Im folgenden wollen wir diese Begriffe verallgemeinern. Wir beginnen mit einigen Definitionen:

3.1.2 **Definition** Seien V, W Vektorräume über K und $\zeta \in \mathrm{Aut}(K)$. Eine Abbildung $\varphi : V \to W$ heisst ζ-*semilinear*, falls $\varphi(u + v) = \varphi(u) + \varphi(v), \varphi(\lambda u) = \zeta(\lambda)\varphi(u), u, v \in V, \lambda \in K$. Eine Abbildung $\sigma : V \times V \to K$ mit den Eigenschaften $\sigma(u + v, w) = \sigma(u, w) + \sigma(v, w), \sigma(u, v + w) = \sigma(u, v) + \sigma(u, w), \sigma(\lambda v, w) = \lambda\sigma(v, w), \sigma(v, \lambda w) = \zeta(\lambda)\sigma(v, w)$ für $u, v, w \in V, \lambda \in K$ heisst eine ζ-*Sesquilinearform*. σ ist also linear im ersten und ζ-semilinear im zweiten Argument.

3.1.3 Von der Theorie der linearen Abbildungen lässt sich alles auch für ζ-semilineare Abbildungen übertragen; insbesondere gilt das für die Begriffe wie Rang und Defekt. Sind etwa V, W endlichdimensional und β bzw. γ Basen von V bzw. W, dann wird ein solches φ eindeutig durch eine Matrix $A = M_{\beta\gamma}(\varphi) \in M(m, n)(K)$ beschrieben (in den Spalten stehen wieder $\gamma(\varphi(v_1)), \ldots, \gamma(\varphi(v_n))$) und es gilt $\varphi(v) = \gamma^{-1}(A\zeta(\beta(v)))$ für alle $v \in V$; der Rang von φ ist der Rang von A, etc.

3.1.4 Wir können jede ζ-Sesquilinearform mit Hilfe einer Matrix beschreiben, wenn wir eine Basis β von V vorgeben: $M_\beta(\sigma) := (\sigma(v_i, v_j))$, wobei $v_1, \ldots, v_n$ die zu β gehörende Basis von V bezeichnet. Wir nennen den Rang dieser Matrix auch den Rang von σ und sagen σ ist nicht-ausgeartet, falls $\mathrm{rg}(\sigma)$

maximal ist (dies ist gleichbedeutend damit, dass $\{v \in V; \sigma(v,V)=0\} = \{v \in V; \sigma(V,v)=0\} = \{0\}$ gilt). Wie man leicht nachrechnet erhalten wir

$$\sigma(u,v) = {}^t\beta(u) M_\beta(\sigma)\zeta(\beta(v)),$$

sowie die Transformationsformel

$$M_\gamma(\sigma) = {}^tT M_\beta(\sigma)\zeta(T),$$

wobei $T = T_{\gamma\beta} = M_{\gamma\beta}(\mathrm{id}_V) \in \mathrm{GL}_n(K)$ die Matrix des Basiswechsel bezeichnet. Wir nennen zwei Matrizen, welche so auseinander hervorgehen *ζ-kongruent*. Die Äquivalenzklassen von $M_n(K)$ bezüglich dieser Äquivalenzrelation entsprechen umgekehrt eindeutig den Klassen von ζ-Sesquilinearformen, welche durch die Äquivalenzrelation $\sigma_2 = \sigma_1 \circ (\psi,\psi)$ mit $\psi \in \mathrm{End}(V)$ bijektiv hervorgehen; in diesem Fall heissen σ_1, σ_2 *ζ-kongruent*.

3.1.5 Sei $V = K^n$. Setze

$$\sigma({}^t(x_1,\ldots,x_n), {}^t(y_1,\ldots,y_n)) := x_1\zeta(y_1) + \ldots + x_n\zeta(y_n).$$

Es handelt sich um eine ζ-Sesquilinearform, genannt die *kanonische Sesquilinearform*. Die kanonische Bilinearform (auf $\mathbb{R}$) bzw. hermitsche Sesquilinearform sind die kanonischen Skalarprodukte auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$. Beachte, dass für beliebiges $\sigma : V \times V \to K$ gilt stets $\sigma(x,y) = {}^txA\zeta(y)$ mit $x = {}^t(x_1,\ldots,x_n), y = {}^t(y_1,\ldots,y_n), A = M_\eta(\sigma) \in M_n(K)$.

3.1.6 **Definition** Wir nennen eine ω-Sesquilinearform $\sigma : V \times V \to K$ *ω-symmetrisch*, falls $\sigma(u,v) = \omega(\sigma(v,u))$ für alle $u,v \in V$. Bei $\omega = \mathrm{id}_K$ sprechen wir von einer *symmetrischen Bilinearform* und bei $K = \mathbb{C}, \omega$ die komplexe Konjugation von einer *hermiteschen Sesquilinearform*.

3.1.7 **Definition** Sei $A \in M_n(K)$ und $\omega \in \mathrm{Aut}(K)$. Die Matrix A heisst *ω-symmetrisch*, falls ${}^tA = \omega(A)$ ist. Für $\omega = \mathrm{id}_K$ auch kurz symmetrisch. Im Fall $K = \mathbb{C}, \omega$ die komplexe Konjugation heisst eine ω-symmetrische Matrix eine *hermitesche Matrix*.

3.1.8 Offenbar ist eine Bilinearform σ symmetrisch genau dann, wenn die Koordinatenmatrix $A = M_\beta(\sigma)$ symmetrisch ist. Im Fall $K = \mathbb{C}$ und ω die komplexe Konjugation ist eine ω-Sesquilinearform hermitesch genau dann, wenn die Koordinatenmatrix A hermitesch ist. Sie erfüllt also $A = {}^t\omega(A) =: {}^hA$.

3.1.9 **Definition** Eine symmetrische Bilinearform bzw. hermitesche Sesquilinearform σ heisst *positiv-definit*, falls $\sigma(v,v) > 0$ für alle $v \in V, v \neq 0$. Beachte, dass auch im komplexen Fall $\sigma(v,v) \in \mathbb{R}$. Eine symmetrische bzw. hermitesche Matrix $A \in M_n(K)$ heisst *positiv-definit*, falls ${}^tvA\omega(v) > 0$ für alle $v \in K^n, v \neq 0$. Beachte, dass aus σ positiv-definit stets σ nicht-ausgeartet folgt.

3.1.10 **Definition** Wir nennen eine positiv-definite symmetrische Bilinearform bzw. eine positiv-definite hermitesche Sesquilinearform eine *Skalarprodukt*. Wir schreiben anstatt $\sigma(v,w) = v \cdot w$, falls σ ein Skalarprodukt ist. Ein reeller Vektorraum zusammen mit einer positiv-definiten symmetrischen Bilinearform heisst ein *euklidischer Vektorraum* und ein komplexer Vektorraum mit einer positiv-definiten hermiteschen Sesquilinearform ein *unitärer Vektorraum*. Ein euklidischer oder unitärer Vektorraum wird auch *Prähilbertraum* genannt.

Beispiel Paradebeispiele sind natürlich $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ mit der kanonischen Bilinearform bzw. hermiteschen Sesquilinearform, welche klarerweise positiv-definit sind.

3.2 Normierte Vektorräume

Sei V im folgenden ein euklidischer oder unitärer Vektorraum über K mit Skalarprodukt σ (beachte, dass hier insbesondere σ positiv-definit entscheidend ist).

3.2.1 **Definition** Für alle $v \in V$ ist die *Länge* von v erklärt durch $\|v\| := \sqrt{v \cdot v} \in \mathbb{R}^+ \cup \{0\}$. Ein Vektor heisst normiert, falls $\|v\| = 1$. Die Abbildung $v \mapsto \|v\|$ nennen wir die *Längenfunktion* von V in $\mathbb{R}$. Es ist $\|v\| = 0 \Leftrightarrow v = 0$.
Es ist insbesondere $\mathbb{R}$ bzw. $\mathbb{C}$ mit dem kanonischen Skalarprodukt ein euklidischer bzw. unitärer Raum. In diesem Fall schreiben wir $|\lambda|$ anstatt $\|\lambda\|$ und sprechen vom *absoluten Betrag* von $\lambda \in K$.

3.2.2 **Satz** (Schwarzsche Ungleichung) Es gilt

$$|v \cdot w| \leq \|v\| \, \|w\|$$

für alle $v, w \in V$. Gleichheit gilt genau dann, wenn v, w linear abhängig oder $v = w$.

Beweis. Sei $x := w \cdot w = \|w\|^2 \in \mathbb{R}$ und $y := v \cdot w \in K$. Dann gilt $0 \leq \|xv - yw\|^2 = (xv - yw) \cdot (xv - yw) = x^2(v \cdot v) - x\overline{y}(v \cdot w) - yx(w \cdot v) + y\overline{y}(w \cdot w) = x(x(v \cdot v) - \overline{y}y - y\overline{y} + y\overline{y}) = \|w\|^2(\|v\|^2\|w\|^2 - |v \cdot w|^2)$. Gilt Gleichheit, so folgt aus $0 = \|xv - yw\|$ dann $xv - yw = 0$. Für $w = 0$ oder $v = w$ ist nichts zu zeigen. Andernfalls ist $\|w\|^2 = x \neq 0$ und $v = x^{-1}yw$, also $\{v, w\}$ linear abhängig. Sei umgekehrt $w = \lambda v, \lambda \in K$, dann $|v \cdot w|^2 = (v \cdot \lambda v)\overline{(v \cdot \lambda v)} = \overline{\lambda}(v \cdot v)\overline{(v \cdot \lambda v)} = (v \cdot v)\overline{(\lambda v \cdot \lambda v)} = \|v\|^2\|w\|^2$. $\square$

3.2.3 **Satz** (Dreiecksungleichung) Es gilt $\|v + w\| \leq \|v\| + \|w\|, \|\lambda v\| = |\lambda|\|v\|$ für alle $v, w \in V, \lambda \in K$.

Beweis. Sei $v \cdot w = x + iy \in K$ mit $x, y \in \mathbb{R}$. Wegen $v \cdot w + \overline{v \cdot w} = 2x \leq 2|x| = 2\sqrt{x^2} \leq 2\sqrt{x^2 + y^2} = 2|x + iy| = 2|v \cdot w|$ und 3.2.2 folgt $\|v + w\|^2 = (v + w) \cdot (v + w) = \|v\|^2 + \|w\|^2 + v \cdot w + w \cdot v \leq \|v\|^2 + \|w\|^2 + 2|v \cdot w| \leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = (\|v\| + \|w\|)^2$. Weiters ist $\|\lambda v\|^2 = (\lambda v) \cdot (\lambda v) = \lambda\overline{\lambda}(v \cdot v) = |\lambda|^2\|v\|^2$.
$\square$

3.2.4 Allgemeiner heisst eine Abbildung $V \to \mathbb{R}, v \mapsto \|v\|$ mit $\|v\| = 0 \Rightarrow v = 0, \|v + w\| \leq \|v\| + \|w\|, \|\lambda v\| = |\lambda|\|v\|$ für alle $v, w \in V, \lambda \in K$ eine *Norm* und V dann ein *normierter Raum*. Es folgt $\|v\| = 0 \Leftrightarrow v = 0$ und $\|v\| = \frac{1}{2}(\|v\| + \| - v\|) \geq \frac{1}{2}\|v - v\| = 0$. Die Norm definiert ausserdem eine *Metrik* $d(v, w) := \|w - v\|$, welche die Eigenschaften $d(u, v) \geq 0, d(u, v) = 0 \Leftrightarrow u = v, d(v, w) = d(w, v), d(v, w) \leq d(v, u) + d(u, w)$ für alle $u, v, w \in V$ erfüllt.

3.2.5 In jedem euklidischen Vektorraum gilt somit

$$-1 \leq \frac{v \cdot w}{\|v\|\|w\|} \leq 1$$

für alle $v, w \in V \backslash \{0\}$. Aus der Analysis kennen wir die Cosinusfunktion

$$\cos : \mathbb{R} \to [-1, 1], x \mapsto \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Es ist $\pi/2 \in \mathbb{R}$ die kleinste positive Nullstelle von cos und die Einschränkung von cos auf $[0, \pi]$ ist bijektiv und besitzt die inverse Funktion $\arccos : [-1, 1] \to [0, \pi]$. Wir nennen

$$\measuredangle(v, w) := \arccos\left(\frac{v \cdot w}{\|v\| \, \|w\|}\right) \in [0, \pi]$$

das *Winkelmass* der Vektoren v, w. Es gilt $\measuredangle(v,w) = \measuredangle(w,v) = \measuredangle(\lambda v, w)$ für alle $\lambda > 0$ und $\measuredangle(v,w)$ stimmt z.B. im $\mathbb{R}^2$ mit der Definition des Winkels aus der Schule überein (ist $v = (\cos\alpha, \sin\alpha), w = (\cos\beta, \sin\beta)$ so ist $\measuredangle(v,w) = \beta - \alpha$). Offenbar gilt $\|v-w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\,\|w\|\cos\alpha$ mit α das Winkelmass von v, w (das ist der *Cosinussatz*).

3.3 Orthogonalität

Wir haben gesehen, dass wir mit dem Skalarprodukt eine Winkelmessung einführen können. Insbesondere sind wir an Vektoren interessiert, die senkrecht aufeinander stehen. Wir definieren:

3.3.1 **Definition** Zwei Vektoren $u, v \in V$ heissen *orthogonal*, genau dann wenn $u \cdot v = 0 \Leftrightarrow: u \perp v$. Wir nennen zwei Unterräume $U_1, U_2 \subseteq V$ *orthogonal*, falls $u_1 \perp u_2$ für alle $u_1 \in U_1, u_2 \in U_2$; wir schreiben $U_1 \perp U_2$. Die Vektoren $v_1, \ldots, v_n \in V \backslash \{0\}$ heissen *orthogonal*, wenn $v_i \perp v_j$ für alle $i \neq j$. Sei heissen *orthonormal*, falls zusätzlich $\|v_i\| = 1, i = 1, \ldots, n$ und eine *Orthonormalbasis*, falls sie auch eine Basis von V sind.

Wir zeigen einige einfache Aussagen:

3.3.2 **Satz** Sind $v_1, \ldots, v_n$ orthogonal, so auch linear unabhängig und die Vektoren $\|v_1\|^{-1}v_1, \ldots, \|v_n\|^{-1}v_n$ sind orthonormal. Ist $v_1, \ldots, v_n$ eine Orthonormalbasis und $v \in V$, so ist $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ mit $\lambda_i = v \cdot v_i$ für $i = 1, \ldots, n$.

Beweis. Sei $\sum_{i=1}^n \lambda_i v_i = 0$, dann folgt $0 = (\sum_{i=1}^n \lambda_i v_i) \cdot v_j = \sum_{i=1}^n \lambda_i (v_i \cdot v_j) = \lambda_j \|v_j\|^2$ und somit $\lambda_j = 0$; dies gilt für alle j und daher sind $\{v_1, \ldots, v_n\}$ linear unabhängig. Die zweite Aussage folgt aus $\|\|v_i\|^{-1}v_i\| = \|v_i\|^{-1}\|v_i\| = 1$ für $i = 1, \ldots, n$. Wie oben sieht man $v \cdot v_i = \lambda_i(v_i \cdot v_i) = \lambda_i$ und somit folgt die Aussage. $\square$

Man nennt die letzte Darstellung die (formale) *Fourierreihe* von v und die λ_i die *Fourier-Koeffizienten* von v; vgl. mit den orthonormalen Funktionen $f_n(x) = e^{inx}, n \in \mathbb{Z}$ im unitären Vektorraum $C([0, 2\pi])$ mit dem Integralskalarprodukt

$$f \cdot g = \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{g(x)}dx.$$

Wir zeigen als nächstes, dass wir stets eine Orthonormalbasis von V finden können.

3.3.3 **Satz** (Gram-Schmidtsches Orthogonalisierungsverfahren) Sei V ein n-dimensionaler euklidischer bzw. unitärer Vektorraum und $v_1, \ldots, v_n$ eine bliebige Basis von V. Dann gibt es eine Orthogonalbasis $w_1, \ldots, w_n$ von V mit $\langle\{v_1, \ldots, v_k\}\rangle = \langle\{w_1, \ldots, w_k\}\rangle$ für $k = 1, \ldots, n$.

Beweis. Für $n = 0$ ist die Aussage trivial. Andernfalls verwenden wir vollständige Induktion nach k mit dem Induktionsanfang $w_1 = v_1$. Sei für $k \geq 1$ bereits eine Orthogonalbasis $\{w_1 \ldots, w_k\}$ des Unterraums $\langle\{w_1, \ldots, w_k\}\rangle$ konstruiert. Wir setzen

$$w_{k+1} = v_{k+1} - \sum_{j=1}^{k} \frac{v_{k+1} \cdot w_j}{w_j \cdot w_j} w_j.$$

Es gilt $\langle\{v_1, \ldots, v_{k+1}\}\rangle = \langle\{w_1, \ldots, w_{k+1}\}\rangle$, was $w_{k+1} \neq 0$ und $w_{k+1} \cdot w_{k+1} \neq 0$ ergibt. Weiter gilt für alle $l \in \{1, \ldots, k\}$

$$w_{k+1} \cdot w_l = \left(v_{k+1} - \sum_{j=1}^{k} \frac{v_{k+1} \cdot w_j}{w_j \cdot w_j} w_j \right) \cdot w_l = v_{k+1} \cdot w_l - \frac{v_{k+1} \cdot w_l}{w_l \cdot w_l} w_l \cdot w_l = 0,$$

sodass mit $\{w_1, \ldots, w_{k+1}\}$ eine Orthogonalsystem vorliegt. Nach 3.3.2 ist die so erhaltene Menge linear unabhängig und ein Erzeugendensystem von V nach Konstruktion. $\square$

Somit besitzt jeder endlichdimensionale euklidische oder unitäre Vektorraum eine Orthogonal- und somit auch eine Orthonormalbasis.

Beispiel Durch Anwendung von 3.3.3 erhalten wir aus $v_1 = {}^t(1,1,1), v_2 = {}^t(0,i,0), v_3 = {}^t(i,0,0)$ in $V = \mathbb{C}^3$ (ausgestattet mit dem kanonischen Skalarprodukt) die Orthogonalbasis $w_1 = {}^t(1,1,1), w_2 = {}^t(-i/3, 2i/3, -i/3), w_3 = {}^t(i/2, 0, -i/2)$. Um eine Orthonormalbasis daraus zu erhalten, müssen wir nur noch mit $\|w_1\| = \sqrt{3}, \|w_2\| = 2/\sqrt{6}, \|w_3\| = \sqrt{2}/2$ normieren.

Wir erhalten sofort noch einen weiteren wichtigen Satz, welcher einen wichtigen Einblick in Komplemente gibt:

3.3.4 **Definition** Sei $U \subseteq V$ ein Unterraum. Dann nennen wir $U^\perp := \{v \in V; v \perp u \text{ für alle } u \in U\}$ das *orthogonale Komplement* von U.

3.3.5 **Satz** (Hauptsatz über orthogonale Komplemente) Ist U ein Unterraum eines n-dimensionalen euklidischen bzw. unitären Vektorraumes V, so gilt $V = U \oplus U^\perp$ und $\dim U^\perp = n - \dim U$; beachte, dass daher $U^\perp$ eindeutig ist, da dieser Raum unabhängig von der Wahl der Basis definiert wurde.

Beweis. Wir wählen eine Orthogonalbasis $v_1, \ldots, v_r$ von U und erweitern sie mit $w_{r+1}, \ldots, w_r$ zu einer Basis von V. Mit Hilfe von 3.3.3 können wir dann aber eine Orthogonalbasis $v_1 \ldots, v_r, v_{r+1}, \ldots, v_n$ von V finden. Wir können also jedes $v \in V$ eindeutig als Linearkombination $\sum_{i=1}^n \lambda_i v_i$ schreiben; wir setzen $u = \sum_{i=1}^r \lambda_i v_i, u^\perp = \sum_{i=r+1}^n \lambda_i v_i$ und erhalten eine Zerlegung $v = u + u^\perp$ mit $u \in U$ und $u^\perp \in U^\perp$. Diese Zerlegung ist eindeutig, da die Linearkombination in der Basis eindeutig ist. Somit folgt die erste Aussage. Der Rest ist klar. $\square$

3.3.6 Mit Hilfe des letzten Satzes können wir auch sofort die sogenannte *orthogonale Projektion* π_U von V auf U (in Richtung $U^\perp$) definieren: Sei $v_1, \ldots, v_r$ Orthonormalbasis von $U^\perp$; dann definieren wir $\pi_U : V \to U, v \mapsto v - (v \cdot v_1)v_1 - \ldots - (v \cdot v_r)v_r = (v \cdot u_1)u_1 + \ldots + (v \cdot u_{n-r})u_{n-r} \in U$, wobei $u_1, \ldots, u_{n-r}, v_1, \ldots, v_r$ eine Orthonormalbasis von V ist.

Beispiel Im unendlich-dimensionalen Fall ist die Situation komplizierter. Z.B. für den reellen Vektorraum $V = C[a,b]$ (mit dem üblichen Integralskalarprodukt) ist die Menge der Polynomfunktionen ein Unterraum U; nach dem Satz von Weierstrass aus der Analysis ist aber $U^\perp = \{0\}$ denn für $f \in C([a,b])$ gibt es eine Folge von $P_n \in U$ mit $P_n \to f$ gleichmässig auf $[a,b]$ und aus $f \cdot P = 0$ für alle $P \in U$ folgt somit $0 = \lim_{n\to\infty} f \cdot P_n = f \cdot (\lim_{n\to\infty} P_n) = f \cdot f = \|f\|^2 \Rightarrow f = 0$ (wobei wir die gleichmässige Konvergenz verwendet haben um den Limes mit der Integration zu vertauschen). Es gilt aber eine schwächere Version von 9.3.5 (siehe Funktionalanalysis).

Beispiel Sei $[a,b]$ ein reelles Intervall und $w(x) \geq 0$ eine Gewichtsfunktion auf $[a,b]$. Betrachte $V = \langle\{1, x, x^2, \ldots, x^n\}\rangle$ und das Skalarprodukt

$$f \cdot g := \int_a^b f(x)g(x)w(x)dx.$$

Die Elemente einer Orthogonalbasis von V nennt man orthogonale Polynome. Spezielle (und für viele Anwendungen sehr wichtige) Orthogonalpolynome sind die Legendre-Polynome mit $a = -1, b = 1, w(x) = 1$ (z.B. $P_0 = 1, P_1 = x, P_2 = \frac{1}{2}(3x^2 - 1), P_3 = \frac{1}{2}(5x^3 - 3x), P_4 = \frac{1}{8}(35x^4 - 30x^2 + 3)$), die

Tschebychef-Polynome mit $a = -1, b = 1, w(x) = (1-x^2)^{-1/2}$ (z.B. $T_0 = 1, T_1 = x, T_2 = 2x^2-1, T_3 = 4x^3 - 3x, T_4 = 16x^5 - 20x^3 + 5x$) oder die Hermite-Polynome mit $a = -\infty, b = \infty, w(x) = e^{-x^2/2}$ (z.B. $H_0 = 1, H_1 = x, H_2 = x^2 - 1, H_3 = x^3 - 3x, H_4 = x^4 - 6x^2 + 3$).

Beispiel Im $\mathbb{C}^3$ mit dem kanonischen Skalarprodukt sei $U := \langle\{\,{}^t(1, i, 1),\ {}^t(1, 0, 0)\}\rangle$ gegeben. Man sieht

$$M_{\eta\eta}(p_U) = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{pmatrix}$$

(Details siehe Vorlesung); beachte, dass diese Matrix hermitesch ist (das ist kein Zufall - siehe später).

3.3.7 Noch einige weitere Folgerungen: Wegen der Eindeutigkeit des orthogonalen Komplementes von U folgt aus $U^\perp \oplus (U^\perp)^\perp = U^\perp \oplus U$ die Identität $U = U^{\perp\perp}$. Ausserdem gilt $U_1 \subseteq U_2 \Leftrightarrow U_1^\perp \supseteq U_2^\perp$ und somit $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp, (U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$. Da stets $U \cap U^\perp = \{0\}$ gilt, ist $\sigma|_{U\times U}$ nicht-ausgeartet (das folgt natürlich auch aus der Vorausetzung σ positiv-definit).

3.4 Isometrische Abbildungen

Wir sehen uns nun strukturverträgliche Abbildungen zwischen euklidischen bzw. unitären Vektorräumen an.

3.4.1 **Definition** Seien V, W euklidische bzw. unitäre Vektorräume. Eine lineare Abbildung $\varphi \in \mathrm{Hom}(V, W)$ heisst *isometrisch* oder eine *Isometrie*, falls $v \cdot w = \varphi(v) \cdot \varphi(w)$ für alle $v, w \in V$ ist. Die Vektorräume heissen *isometrisch-isomorph*, falls es einen isometrischen Isomorphismus gibt.

3.4.2 Jede Isometrie φ ist injektiv ($v \in \mathrm{Ker}(\varphi) \Rightarrow v \cdot w = \varphi(v) \cdot \varphi(w) = 0 \cdot \varphi(w) = 0$ für alle $w \in V$ und somit $v \in V^\perp = \{0\}$) und erhält Längen und Winkeln. Falls $\dim V = \dim W < \infty$, so ist sie auch stets surjektiv. Für eine Basis β von V gilt: β ist eine Orthonormalbasis genau dann, wenn β eine Isometrie ist. Je zwei n-dimensionale euklidische bzw. unitäre Vektorräume sind isometrisch-isomorph. Die Menge aller isometrischen Isomorphismen bilden eine Untergruppe von $\mathrm{GL}(V)$, welche die *isometrische Gruppe* genannt wird. Im euklidischen Fall spricht man auch von der *orthogonalen* Gruppe und im unitären Fall von der *unitären* Gruppe; entsprechend nennt man die isometrische Isomorphismen auch orthogonal bzw. unitär.

3.4.3 **Definition** Eine reguläre Matrix $A \in M_n(\mathbb{R})$ heisst *orthogonal*, falls $A^{-1} = {}^tA$. Eine reguläre Matrix $A \in M_n(\mathbb{C})$ heisst *unitär*, falls $A^{-1} = {}^hA$.

3.4.4 Die Menge $\mathrm{O}(n)$ der orthogonale Matrizen bilden eine Untergruppe von $\mathrm{GL}_n(\mathbb{R})$, genannt die *orthogonale Gruppe*. Sie haben Determinante ± 1. Man nennt die Menge $\mathrm{SO}(n) := \{A \in \mathrm{O}(n); \det(A) = 1\}$ die *spezielle orthogonale Gruppe*. Genauso bilden die unitären Matrizen $\mathrm{U}(n)$ eine Untergruppe von $\mathrm{GL}_n(\mathbb{C})$, welche die *unitäre Gruppe* genannt wird. Für die Determinante von $A \in \mathrm{U}(n)$ gilt $|\det(A)| = 1$.

3.4.5 **Satz** Sei $A \in M_n(\mathbb{R})$ bzw. $M_n(\mathbb{C})$. Dann ist sind die folgenden Aussagen äquivalent:

a) A orthogonal bzw. unitär,

b) die Spalten von A bilden eine Orthonormalbasis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ bzgl. des kanonisch- -en Skalarproduktes,

c) die Zeilen von A bilden ein Orthonormalbasis von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ bzgl. des kanonischen

Skalarproduktes.

Beweis. Es gilt ${}^tAA = E_n$ bzw. ${}^hAA = E_n$. Mit dem kanonischen Skalarprodukt auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ ist diese Bedingung daher gleichbedeutend damit, dass die Spalten von A eine Orthonormalbasis bilden. Der letzte Teil folgt sofort durch (hermitesch) transponieren. □

Die Bedeutung von orthogonalen Matrizen zeigt der folgende Satz:

3.4.6 **Satz** Sei β eine Orthonormalbasis eines euklidischen bzw. unitären Vektorraumes V. Eine Abbildung $\varphi \in \mathrm{GL}(V)$ ist genau dann isometrisch, falls $M_{\beta\beta}(\varphi)$ eine orthogonale bzw. unitäre Matrix ist.

Beweis. Sei $v_1, \ldots, v_n$ die zu β gehörenden Basisvektoren. Sei φ isometrisch, so gilt $\beta(\varphi(v_i))\cdot\beta(\varphi(v_j)) = \varphi(v_i) \cdot \varphi(v_j) = v_i \cdot v_j = \delta_{ij}$. Somit bilden die Spalten von $M_{\beta\beta}(\varphi) = (\beta(\varphi(v_1)), \ldots, \beta(\varphi(v_n)))$ eine Orthonormalbasis und daher ist $M_{\beta\beta}(\varphi)$ nach 3.4.4 orthogonal bzw. unitär. Umgekehrt nehmen wir, dass $M_{\beta\beta}(\varphi)$ orthogonal bzw. unitär sei. Dann bilden mit 3.4.4 die Spalten $\beta(\varphi(v_1)), \ldots, \beta(\varphi(v_n))$ eine ONB und daher gilt $\varphi(v_i)\cdot\varphi(v_j) = \beta(\varphi(v_i))\cdot\beta(\varphi(v_j)) = \delta_{ij} = v_i\cdot v_j$, wobei wir verwendet haben, dass β eine ONB ist. Somit folgt die Aussage. □

3.4.7 Wir können wieder die üblichen Klassifikationsfragen stellen: Zwei Matrizen $A, B \in M_n(\mathbb{R})$ bzw. $M_n(\mathbb{C})$ heissen *orthogonal-ähnlich* bzw. *unitär-ähnlich*, falls es ein $P \in \mathrm{O}(n)$ bzw. $\mathrm{U}(n)$ mit $B = PAP^{-1}$ gibt. Offenbar handelt es sich um eine Verfeinerung der Einteilung mittels Ähnlichkeit, denn aus orthogonal- bzw. unitär-ähnlich folgt ähnlich und gleichzeitig ω-kongruent; diese beiden Begriffe fallen hier also zusammen.

§4. Adjungierte Abbildungen

Es sei im folgenden V stets ein euklidischer oder unitärer Vektorraum.

4.1 Adjungierte Abbildung

Sind V, W zwei euklidischer bzw. unitäre Vektorräume mit Skalarprodukten σ_V bzw. σ_W, welche wir beide mit $\cdot$ bezeichnen. Ausserdem sei $\varphi \in \mathrm{Hom}(V, W)$.

4.1.1 **Definition** Wir nennen $\psi \in \mathrm{Hom}(W, V)$ eine zu $\varphi \in \mathrm{Hom}(V, W)$ *adjungierte Abbildung*, falls $\varphi(v) \cdot w = v \cdot \psi(w)$ für alle $v \in V, w \in W$ gilt.

Besitzt eine Abbildung φ eine adjungierte Abbildung ψ, so ist umgekehrt auch φ eine adjungierte Abbildung zu ψ.

4.1.2 **Satz** Jede Abbildung $\varphi \in \mathrm{Hom}(V, W)$ besitzt höchstens eine adjungierte Abbildung. Sind V, W endlichdimensional, dann gibt es stets genau eine adjungierte Abbildung $\varphi^\wedge$; sind β bzw. γ Orthonormalbasen von V bzw. W so gilt $M_{\gamma\beta}(\varphi^\wedge) = {}^t\omega(M_{\beta\gamma}(\varphi))$.

Beweis. Angenommen es gibt zwei adjungierte Abbildungen ψ_1, ψ_2 zu φ. Es folgt $v\cdot\psi_1(w) = \varphi(v)\cdot w = v \cdot \psi_2(w)$ und somit $v \cdot (\psi_1 - \psi_2)(w) = 0$ für alle $v \in V, w \in W$. Daher ist $(\psi_1 - \psi_2)(w) = 0$ für alle $w \in W$ und somit $\psi_1 = \psi_2$. Sei $v_1 = \beta^{-1}(e_1), \ldots, v_n = \beta^{-1}(e_n), w_1 = \gamma^{-1}(e_1), \ldots, w_m = \gamma^{-1}(e_m)$. Dann zeigt $\gamma(\varphi(v_i))\cdot e_j = \gamma(\varphi(v_i))\cdot\gamma(w_j) = \varphi(v_i)\cdot w_j = v_i\cdot\varphi^\wedge(w_j) = \beta(v_i)\cdot\beta(\varphi^\wedge(w_j)) = e_i\cdot\beta(\varphi^\wedge(w_j))$, dass notwendigerweise $M_{\gamma\beta}(\varphi^\wedge) = {}^t\omega(M_{\beta\gamma}(\varphi))$ gelten muss. Umgekehrt folgt mit dieser Voraussetzung, dass $\varphi(v) \cdot w = v \cdot \varphi^\wedge(w)$ für alle $v \in V, w \in W$ gilt, wenn wir $\varphi^\wedge \in \mathrm{Hom}(W, V)$ durch $M_{\gamma\beta}(\varphi^\wedge) = {}^t\omega(M_{\beta\gamma}(\varphi))$ definieren. □

4.1.3 Sind $\varphi_1, \varphi_2 \in \mathrm{Hom}(V, W)$ mit adjungierten Abbildungen $\varphi_1^\wedge, \varphi_2^\wedge$, dann ist $(\varphi_1 + \varphi_2)^\wedge = \varphi_1^\wedge + \varphi_2^\wedge, (\lambda\varphi_1)^\wedge = \omega(\lambda)\varphi_1^\wedge$; für $\varphi, \psi \in \mathrm{End}(V)$ ist $(\psi \circ \varphi)^\wedge = \varphi^\wedge \circ \psi^\wedge$.

Ausserdem gilt:

4.1.4 **Satz** Seien $\varphi \in \mathrm{Hom}(V, W)$ und $\varphi^\wedge \in \mathrm{Hom}(W, V)$ adjungierte Abbildungen, dann ist $\varphi(V)^\perp = \mathrm{Ker}\varphi^\wedge$.

Beweis. Es gilt $w \in \varphi(V)^\perp \Leftrightarrow \varphi(v) \cdot w = v \cdot \varphi^\wedge(w) = 0$ für alle $v \in V \Leftrightarrow \varphi^\wedge(w) = 0 \Leftrightarrow w \in \mathrm{Ker}(\varphi^\wedge)$ für $w \in W$ und somit die Aussage. □

4.1.5 Es folgt, dass aus φ surjektiv die Injektivität von $\varphi^\wedge$ folgt; ist $\dim W < \infty$ so gilt auch die Umkehrung. Indem man die Rolle von φ und $\varphi^\wedge$ vertauscht ist ausserdem $\mathrm{Ker}\varphi = \varphi^\wedge(W)^\perp$ und aus der Surjektivität von $\varphi^\wedge$ folgt die Injektivität von φ. Es ist $W = \mathrm{Ker}(\varphi^\wedge) \oplus \varphi(V), V = \mathrm{Ker}(\varphi) \oplus \varphi^\wedge(W)$, falls V und W endliche Dimension haben (wegen 3.3.5).

4.1.6 **Satz** Eine Bijektion $\varphi \in \mathrm{Hom}(V, W)$ ist genau dann eine Isometrie, falls $\varphi^{-1} = \varphi^\wedge$ ist.

Beweis. Sei φ eine isometrische Bijektion. Dann zeigt, $\varphi(v) \cdot w = \varphi^{-1}(\varphi(v)) \cdot \varphi^{-1}(w) = v \cdot \varphi^{-1}(w)$ für alle $v \in V, w \in W$, dass φ^{-1} eine zu φ adjungierte Abbildung ist. Umgekehrt folgt aus $\varphi(v) \cdot w = v \cdot \varphi^{-1}(w)$ für alle $v \in V, w \in W$ mit $w =: \varphi(u)$ sofort $\varphi(v) \cdot \varphi(u) = v \cdot u$ für alle $v, u \in V$. □

4.2 Selbstadjungierte Abbildung

Wir wollen nun die Zerlegung einer komplexen Zahl z in Real- und Imaginärteil nachbilden; dies lässt sich auch in der Form $z = \frac{1}{2}(z + \omega(z)) + \frac{1}{2}(z - \omega(z))$ schreiben. Wir definieren:

4.2.1 **Definition** Eine Abbildung $\varphi \in \mathrm{End}(V)$ heisst *selbstadjungiert* bzw. *antiselbstadjungiert*, falls φ bzw. $-\varphi$ adjungierte Abbildung zu φ ist.

Solche Abbildungen treten z.B. in der Quantenphysik als Übertragung der klassischen Grössen auf, welche nun auf die Wellenfunktionen der Elementarteilchen angewandt werden.

4.2.2 In einem euklidischen Vektorraum nennt man eine (anti-)selbstadjungierte Abbildung φ auch *(schief-)symmetrisch* und in einem unitären Vektorraum auch *(schief-)hermitesch*. Die Erklärung dafür sieht man so (und das folgt unmittelbar aus 4.1.2): Sei β eine Orthonormalbasis von V, dann ist die Abbildung $\varphi \in \mathrm{End}(V)$ genau dann (anti-)selbstadjungiert, falls die Matrix $M_{\beta\beta}(\varphi)$ ω-(schief-)symmetrisch ist, d.h. $M_{\beta\beta}(\varphi) = \pm\, {}^t\omega(M_{\beta\beta}(\varphi))$.

Somit kann man leicht Beispiele für selbstadjungierte Abbildungen angeben. Weitere Beispiele sind die Orthogonalprojektionen, wie man durch direkte Rechnung zeigt.

4.3 Normale Abbildungen und der Spektralsatz

Wir behandeln nun selbstadjungierte, antiselbstadjungierte und bijektive isometrische Abbildungen eines euklidischen oder unitären Vektorraumes V gleichzeitig, denn sie ordnen sich der folgenden Definition unter:

4.3.1 **Definition** Eine Abbildung $\varphi \in \mathrm{End}(V)$ heisst *normal*, falls φ eine adjungierte Abbildung $\varphi^\wedge$ besitzt und mit dieser kommutiert, d.h. $\varphi \circ \varphi^\wedge = \varphi^\wedge \circ \varphi$.

4.3.2 Für eine normale Abbildung gilt: $\varphi(v) \cdot \varphi(w) = \varphi^\wedge(v) \cdot \varphi^\wedge(w)$ für alle $v, w \in V$. Insbesondere ist daher $\mathrm{Ker}(\varphi) = \mathrm{Ker}(\varphi^\wedge)$ und somit erhalten wir für endlich-dimensionales V die orthogonale Zerlegung $V = \mathrm{Ker}(\varphi) \oplus \mathrm{Im}(\varphi)$.

Beispiel Jede selbstadjungierte, antiselbstadjungierte und bijektiv-isometrische Abbildung ist normal. Jedes $\varphi \in \mathrm{End}(V)$, für welche es eine Orthonormalbasis $v_1, \ldots, v_n$ bestehend aus Eigenvektoren gibt, ist normal (denn wir erhalten $\varphi^\wedge$ durch $v_i \mapsto \omega(\lambda_i) v_i$ mit λ_i der zu v_i gehörige Eigenwert für $i = 1, \ldots, n$).

4.3.3 Wir nennen eine Matrix $A \in M_n(K)$ *ω-normal*, falls ${}^t\omega(A) \cdot A = A \cdot {}^t\omega(A)$ ist. Im euklidischen Fall sprechen wir von *normalen* und im unitären Fall von *konjugiert-normalen* Matrizen. Wie üblich gilt: Sei β eine Orthonormalbasis von V, dann ist $\varphi \in \mathrm{End}(V)$ normal genau dann, wenn $M_{\beta\beta}(\varphi)$ eine ω-normale Matrix ist.

Wir nähern uns nun dem nächsten Hauptresultat, dem sogenannten Spektralsatz für normale Abbildungen. Zunächst zeigen wir:

4.3.4 **Hilfssatz** Sei $\varphi \in \mathrm{End}(V)$ normal und U ein φ- und $\varphi^\wedge$-invarianter Unterraum von V. Dann ist $\psi : U \to U, u \mapsto \varphi(u)$ normal. Der Unterraum $U^\perp$ ist ebenfalls φ- und $\varphi^\wedge$-invariant.

Beweis. Die adjungierte Abbildung zu ψ ist durch $u \mapsto \varphi^\wedge(u)$ für alle $u \in U$ gegeben. Somit ist ψ normal. Wähle $v \in U^\perp$. Dann gilt $0 = v \cdot u = v \cdot \varphi(u) = \varphi^\wedge(v) \cdot u$ für alle $u \in U$, d.h. $\varphi^\wedge(v) \in U^\perp$. Die φ-Invarianz von $U^\perp$ folgt analog wegen $\varphi^{\wedge\wedge} = \varphi$. □

4.3.5 **Hilfssatz** Sei $\varphi \in \mathrm{End}(V)$ normal und $\lambda \in \mathrm{Spec}(\varphi)$. Dann ist $\omega(\lambda) \in \mathrm{Spec}(\varphi^\wedge)$. Der Eigenraum von φ zum Eigenwert λ stimmt mit dem Eigenraum von $\varphi^\wedge$ zum Eigenwert $\omega(\lambda)$ überein. Insbesondere gilt $\lambda = \omega(\lambda)$ falls φ selbstadjungiert, $\lambda = -\omega(\lambda)$ falls φ antiselbstadjungiert und $\lambda\omega(\lambda) = 1$ falls φ isometrisch ist.

Beweis. Für alle $\mu \in K$ folgt $(\varphi - \mu \mathrm{id}_V)^\wedge = \varphi^\wedge - (\mu \mathrm{id}_V)^\wedge = \varphi^\wedge - \omega(\lambda)\mathrm{id}_V$. Mit φ ist daher auch $\varphi - \mu \mathrm{id}_V$ normal, da je zwei der Abbildungen $\varphi, \varphi^\wedge, \mu \mathrm{id}_V$ und $\omega(\mu)\mathrm{id}_V$ kommutieren. Sei nun λ Eigenwert von φ. Dann gilt $v \in \mathrm{Ker}(\varphi - \lambda \mathrm{id}_V) \Leftrightarrow (\varphi - \lambda \mathrm{id}_V)(v) = 0 \Leftrightarrow ((\varphi - \lambda \mathrm{id}_V)(v)) \cdot ((\varphi - \lambda \mathrm{id}_V)(v)) = 0 \Leftrightarrow ((\varphi^\wedge - \omega(\lambda)\mathrm{id}_V)(v)) \cdot ((\varphi^\wedge - \omega(\lambda)\mathrm{id}_V)(v)) = 0 \Leftrightarrow (\varphi^\wedge - \omega(\lambda)\mathrm{id}_V)(v) = 0 \Leftrightarrow v \in \mathrm{Ker}(\varphi^\wedge - \omega(\lambda)\mathrm{id}_V)$, was $\omega(\lambda)$ als Eigenwert von $\varphi^\wedge$ erweist und die Gleichheit der betreffenden Eigenräume liefert. Sei v ein Eigenvektor zum Eigenwert λ von φ; für $\varphi^\wedge = \pm\varphi$ folgt dann $\lambda v = \varphi(v) = \pm\varphi^\wedge(v) = \pm\omega(\lambda)v$ und da $v \neq 0$ somit $\lambda = \pm\omega(\lambda)$. Für eine isometrische Abbildung φ wähle $v \in \mathrm{Ker}(\varphi - \lambda \mathrm{id}_V) \backslash \{0\}$; wegen $v \cdot v = \varphi(v) \cdot \varphi(v) = (\lambda v) \cdot (\lambda v) = \lambda\omega(\lambda)(v \cdot v)$ und $v \cdot v \neq 0$ ist $\lambda\omega(\lambda) = 1$. □

4.3.6 **Hilfssatz** Sei $\varphi \in \mathrm{End}(V)$ normal. Dann sind je zwei zu verschiedenen Eigenwerten von φ gehörigen Eigenräume orthogonal.

Beweis. Seien λ_1 und λ_2 zwei verschiedene Eigenwerte von φ, $v_1 \in \mathrm{Ker}(\varphi - \lambda_1 \mathrm{id}_V)$ und $v_2 \in \mathrm{Ker}(\varphi - \lambda_2 \mathrm{id}_V)$. Dann gilt $\lambda_1 v_1 \cdot v_2 = \varphi(v_1) \cdot v_2 = v_1 \cdot \varphi^\wedge(v_2) = v_1 \cdot \omega(\lambda_2) v_2 = \lambda_2 v_1 \cdot v_2$. Das zeigt die Aussage. □

4.3.7 **Satz** (Spektralsatz für normale Abbildungen) Sei V endlichdimensional. Falls das charakteristische Polynom einer normalen Abbildung $\varphi \in \mathrm{End}(V)$ in Linearfaktoren zerfällt, besitzt φ eine Orthogonalbasis aus Eigenvektoren.

Beweis. Der Satz ist für $n = 1$ richtig. Wir führen den Beweis durch vollständige Induktion nach n und schliessen für $n > 1$ von $n - 1$ auf n: Sei φ normal und v_1 ein Eigenvektor von φ zu einem Eigenwert $\lambda_1 \in K$. Der Unterraum $\langle\{v_1\}\rangle$ ist φ-invariant und auch $\varphi^\wedge$-invariant. Aus 4.3.4 folgt, dass gleiches daher für $\langle\{v_1\}\rangle^\perp$ gilt und dass die Abbildung $\psi : \langle\{v_1\}\rangle^\perp \to \langle\{v_1\}\rangle^\perp$, $v \mapsto \varphi(v)$ normal ist. Ergänze eine Basis β_1 von $\langle\{v_1\}\rangle^\perp$ durch v_1 zu einer Basis β von V. Dann gilt

$$M_{\beta\beta}(\varphi) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & M_{\beta_1\beta_1}(\psi) \end{pmatrix}$$

und daher weiter $P_\varphi(X) = (\lambda_1 - X) P_\psi(X)$. Es zerfällt also auch P_ψ in Linearfaktoren. Nach Induktionsannahme besitzt daher ψ eine Orthogonalbasis von Eigenvektoren, die durch v_1 nach Orthogonalisieren des zu λ_1 zugehörigen Eigenraums mittels Gram-Schmidt und unter Beachtung von 4.3.6 zu einer Orthogonalbasis von Eigenvektoren von φ ergänzt wird. □

4.3.8 Jede normale Abbildung deren charakteristisches Polynom in Linearfaktor zerfällt ist also diagonalisierbar. Die Diagonalisierbarkeit einer ω-normalen Matrix folgt aus dem Zerfallen des charakteristischen Polynomes. Diese Basis lässt sich leicht bestimmten, indem man zunächst irgendeine Basis aus Eigenvektoren berechnet und die Eigenvektoren zu gleichen Eigenwerten mit Hilfe von 4.3.4 orthogonalisiert.

4.3.9 Sei nun V ein n-dimensionaler unitärer Vektorraum. Dann folgt unmittelbar aus dem Spektralsatz:
a) $\varphi \in \mathrm{End}(V)$ ist normal genau dann, wenn φ eine Orthonormalbasis aus Eigenvektoren besitzt. $A \in M_n(\mathbb{C})$ ist genau dann konjugiert-normal, falls sie unitär-ähnlich einer Diagonalmatrix ist.
b) $\varphi \in \mathrm{End}(V)$ ist genau dann hermitesch bzw. schiefhermitesch bzw. unitär, falls es eine Orthonormalbasis aus Eigenvektoren und nur reellen bzw. nur rein imaginären Eigenwerten bzw. nur Eigenwerten mit Absolutbetrag 1 besitzt. $A \in M_n(\mathbb{C})$ ist genau dann hermitesch bzw. schiefhermitesch bzw. unitär, falls sie unitär-ähnlich zu einer Diagonalmatrix ist, die aus reellen bzw. rein-imaginären Elementen bzw. Diagonalelementen mit Absolutbetrag 1 besteht.

4.3.10 Sei V jetzt ein n-dimensionaler euklidischer Vektorraum. Dann folgt aus dem Spektralsatz:
a) $\varphi \in \mathrm{End}(V)$ ist symmetrisch genau dann, wenn φ eine Orthonormalbasis aus Eigenvektoren besitzt. Eine Matrix $A \in M_n(\mathbb{R})$ ist genau dann symmetrisch, falls sie orthogonal-ähnlich zu einer Diagonalmatrix ist.
b) $\varphi \in \mathrm{End}(V)$ ist normal genau dann, wenn die Koordinatenmatrix von φ bzgl. einer Orthonormalbasis von V, welche nur aus reellen Diagonalelementen und Matrizen der Form

$$J_2(z, \overline{z}) = \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \in M_2(\mathbb{R})$$

für Paare konjugiert komplexer Zahlen $z = x + iy, \overline{z} = x - iy$ aufgebaut ist. Sie ist schiefsymmetrisch, falls alle Diagonalelemente dieser Blockdiagonalmatrix 0 sind und orthogonal, falls alle reellen

Diagonalelemente ± 1 und auftretenden komplexen Zahlen $|z| = 1$ haben; somit ist

$$J_2(z, \overline{z}) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \in M_2(\mathbb{R})$$

mit $0 \leq \alpha < 2\pi$. Die letzten beiden Aussagen kann man entsprechend auch auf Matrizen übertragen. Ausserdem erhalten wir auch die Klassifikation von allen Drehungen (das sind die Elemente von $\mathrm{SO}(n)$) für $n = 2$ und $n = 3$: in der Ebene haben sie die Gestalt der obigen Matrix und für $n = 3$ gibt es einen Fixvektor (das ist der Satz von Euler-d'Alembert), da 1 einer der Eigenwerte sein muss, um den eine Drehung in der Ebene erfolgt.

Beweis. Es genügt die Aussage für $A \in M_n(\mathbb{R})$ zu zeigen. Beachte, dass wir A auch als komplexe Matrix auffassen und dann 4.3.7 anwenden können. Sei $z = x + iy \in \mathbb{C}$ ein Eigenwert zum Eigenvektor $v \in \mathbb{C}^n$, so ist wegen $A\overline{v} = \overline{Av} = \overline{zv} = \overline{z}\overline{v}$ der Vektor $\overline{v}$ ein Eigenvektor von A zum Eigenwert $\overline{z} = x - iy$. Setzen wir $v_1 = \frac{1}{2}(v + \overline{v}), v_2 = \frac{1}{2i}(v - \overline{v})$, so ist $v_1, v_2 \in \mathbb{R}^n$ und $Av_1 = xv_1 + yv_2, Av_2 = xv_2 - yv_1$. Somit folgen die Aussagen aus dem Spektralsatz. □

Wir halten noch eine wichtige Folgerung fest:

4.3.11 **Satz** (Hauptachsentransformation) Sei σ eine symmetrische Bilinearform bzw. hermitesche Sesquilinearform eines n-dimensionalen euklidischen bzw. unitären Vektorraumes V. Dann existiert mindestens eine Orthonormalbasis $v_1, \ldots, v_n$ von V mit $\sigma(v_i, v_j) = 0$ für alle $i \neq j, i, j \in \{1, \ldots, n\}$.

Beweis. Sei γ bzw. $w_1, \ldots, w_n$ eine beliebige Basis von V, $A = M_\gamma(\cdot)$ eine Koordinatenmatrix des Skalarproduktes und $B = M_\gamma(\sigma) \in M_n(K)$. Nach Vorausseztung gilt ${}^t\omega(A) = A, {}^t\omega(B) = B$. Wir erhalten eine Abbildung $\varphi \in \mathrm{End}(V)$ durch $M_{\gamma\gamma}(\varphi) = \omega(A)^{-1}\omega(B)$. Es gilt $\sigma(v, w) = v \cdot \varphi(w) = \varphi(v) \cdot w$, denn dies ist gleichbedeutend zu ${}^t\gamma(v)B\omega(\gamma(w)) = {}^t\gamma(v)A\omega(M_{\gamma\gamma}(\varphi)\gamma(w)) = {}^t(M_{\gamma\gamma}(\varphi)\gamma(v))A\omega(\gamma(w))$ für alle $v, w \in V$. Somit $\varphi = \varphi^\wedge$. Als selbstadjungierte Abbildung besitzt φ nur reelle Eigenwerte und eine ONB β (bzw. $v_1, \ldots, v_n$) aus Eigenvektoren. Es ist $\varphi(v_i) = \lambda_i v_i$ für $i = 1, \ldots, n$. Wegen $\lambda_i \in \mathbb{R}$ und $\sigma(v_i, v_j) = v_i \cdot \varphi(v_j) = \varphi(v_i) \cdot v_j = \lambda_j \delta_{ij} = \lambda_i \delta_{i,j}$ für alle $i, j \in \{1, \ldots, n\}$ hat β die gewünschte Eigenschaft. □

4.3.12 Jede Orthonormalbasis von V mit den obigen Eigenschaften besteht nur aus Eigenvektoren der Abbildung $\varphi \in \mathrm{End}(V)$ mit $\sigma(v, w) = v \cdot \varphi(w)$. Es ist also leicht die gesuchte Basis zu berechnen. Die durch die v_i aufgespannten eindimensionalen Unterräume heissen die *Hauptachsen* von σ. Falls φ n verschiedene Eigenwerte besitzt, dann sind die Hauptachsen eindeutig bestimmt. Der letzte Satz lässt sich auch als simultane Transformation auf Diagonalgestalt von zwei ω-symmetrischen Sesquilinearformen auffassen, von denen mindestens eine positiv-definit ist (und somit die Rolle des Skalarproduktes übernimmt).

Beispiel Sei $V = \mathbb{R}^2$ mit dem kanonischen Skalarprodukt und σ eine symmetrische Bilineaform mit Koordinatenmatrix

$$M_\eta(\sigma) = \begin{pmatrix} 31 & -12 \\ -12 & 21 \end{pmatrix} = A \in M_2(\mathbb{R}).$$

Es gilt $\sigma(v, w) = {}^tvAw = {}^tv(Aw) = {}^t(Av)w = \varphi(v) \cdot w = v \cdot \varphi(w)$ mit der selbstadjungierten Abbildung $\varphi \in \mathrm{End}(V)$ gegeben durch $M_{\eta\eta}(\varphi) = A$. Die Eigenwerte von A lauten $39, 13$ und die zugehörigen normierten Eigenvektoren sind gegeben durch $1/\sqrt{13}\,{}^t(3, -2), 1/\sqrt{13}\,{}^t(2, 3)$. In dem wir

$$P = \frac{1}{\sqrt{13}} \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} \in \mathrm{O}(2)$$

als neue Basis einführen erhalten wir ${}^tPAP = \mathrm{diag}(39, 13)$.

Wir beantworten noch die Frage, wie man positive Definitheit für eine symmetrische Bilinearform überprüft:

4.3.13 **Satz** Sei σ eine symmetrische Bilinearform eines n-dimensionalen euklidischen Vektorraumes V. Weiters sei $A = M_\beta(\sigma)$ die Koordinatenmatrix von σ. Dann gilt: σ ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.

Beweis. Da σ symmetrisch, ist A eine symmetrische Matrix und besitzt somit eine ONB γ aus Eigenvektoren. Die Matrix $B = M_\gamma(\sigma) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ ist dann ebenfalls eine Koordinatenmatrix von σ. Die reellen Zahlen $\lambda_1, \ldots, \lambda_n$ sind die Eigenwerte der Matrix A. Offenbar gilt mit ${}^t(x_1, \ldots, x_n) = \gamma(v)$ für $v \in V$: $\sigma(v, v) > 0$ für alle $v \neq 0 \Leftrightarrow \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 > 0$ für alle ${}^t(x_1, \ldots, x_n) \in \mathbb{R}^n \setminus \{0\} \Leftrightarrow \lambda_i > 0$ für alle $i = 1, \ldots, n$. $\square$

Beachte, dass die Eigenwerte keine Invarianten von σ und daher nicht eindeutig bestimmt sind. Eine wesentlich effizientere Möglichkeit das Definitheitsverhalten zu bestimmen (und dies funktionert dann auch im hermiteschen Fall) besteht in symmetrischer Diagonalisierung. Wie man dann sieht, sind nur die Vorzeichen der Diagonaleinträge eindeutig bestimmt (das ist der Trägheitssatz von Sylvester). Auf dieses Verfahren gehen wir aus Zeitgründen nicht ein.

Literatur

1. G. Fischer, Lineare Algebra, Vieweg-Verlag, 2010
2. H. Havlicek, Lineare Algebra für Technische Mathematiker, Heldermann, 2012
3. S. Lang, Linear Algebra, Springer, 2004