Lineare Algebra II

Ein Kurzskript

A. Martin-Pizarro

Albert-Ludwigs-Universität Freiburg
Akademisches Jahr 20/21

pizarro@math.uni-freiburg.de

15. Juli 2021

Anmerkungen.

Dieses Kurzskript ist während der im Akademischen Jahr an der Albert-Ludwigs-Universität in Freiburg gehaltenen Vorlesungen „Lineare Algebra I & II“ entstanden und stark geprägt von den Skripten meiner Kollegen Martin Ziegler und Tobias Kaiser.

Zu meinem eigenen Beitrag gehören sicherlich die zahlreichen Fehler, welche es im Skript definitiv geben wird. Ich bin sehr dankbar über die Mitteilung solcher Fehler und Ungenauigkeiten.

Insbesondere bedanke ich mich bei Herrn Michael Lösch für das aufmerksame Korrekturlesen aber vor allem für die Geduld.

Inhaltsverzeichnis

Kapitel 1

Die jordansche Normalform

Im Folgenden sei $\mathbb{K}$ ein Körper.

1.1 Eigenvektoren und -werte

Definition 1.1. Zwei Unterräume U und W eines $\mathbb{K}$-Vektorraumes V liegen *transversal* zueinander, falls $U \cap W = \{0_V\}$. Wenn U und W transversal liegen, schreiben wir $U \oplus W$ für den Unterraum $U + W$ und bezeichnen es als die *direkte Summe* von U und W. Allgemein ist eine Familie $(U_i)_{i\in I}$ von Unterräumen von V *transversal*, falls für jedes i aus I die Unterräume U_i und $\sum_{j\in I\setminus\{i\}} U_j$ transversal zueinander liegen. Wir schreiben dann die direkte Summe $\bigoplus_{i\in I} U_i$ für den Unterraum $\sum_{i\in I} U_i$.

Bemerkung 1.2. Die Familie $(U_i)_{i\in I}$ von Unterräumen des $\mathbb{K}$-Vektorraumes V ist genau dann transversal, wenn für alle natürlichen Zahlen $m \geq 1$, Indizes $i_1, \ldots, i_m$ aus I sowie Vektoren $u_1, \ldots, u_m$ aus V mit u_j aus U_{i_j} für $1 \leq j \leq m$, gilt

$$\sum_{j=1}^{m} u_j = 0_V \quad \Longrightarrow \quad u_1 = \ldots = u_m = 0_V$$

Beweis. Wir nehmen zuerst an, dass die Familie $(U_i)_{i\in I}$ transversal ist. Gegeben Vektoren $u_1, \ldots, u_m$ wie oben mit $0_V = u_1 + \ldots + u_m$, beweisen wir induktiv über m, dass alle Vektoren $u_1, \ldots, u_m$ Null sind. Für $m = 1$ ist die Behauptung trivial. Angenommen nun, dass $m > 1$, beachte, dass

$$u_1 + \ldots + u_{m-1} = -u_m \text{ in } U_{i_m} \cap \sum_{\substack{i\in I\\ i\neq i_m}} U_i = \{0_V\} \text{ liegt}$$

und somit $u_m = 0_V = u_1 + \ldots + u_{m-1}$. Induktiv folgt $u_1 = \ldots = u_{m-1} = 0_V$, wie gewünscht.

Für die Rückrichtung sei v ein Vektor aus $U_j \cap \sum_{i\in I\setminus\{j\}} U_i$. Wegen der [**Skript LAI**, Bemerkung 1.47] lässt sich v für verschiedene Indizes $i_1, \ldots, i_N$ aus $I \setminus \{j\}$ als $v = u_1 + \cdots + u_N$ schreiben, mit u_k in U_{i_k}. Also

$$u_1 + \cdots + u_N + (-v) = 0_V.$$

Aus unserer Annahme (mit $m = N + 1$) folgt $v = 0_V$, wie gewünscht. $\square$

Lemma 1.3. *Gegeben eine transversale Familie $(U_i)_{i\in I}$ von Unterräumen des $\mathbb{K}$-Vektorraumes V und Basen B_i von U_i für jedes i aus I, ist $\bigcup_{i\in I} B_i$ eine Basis der direkten Summe $\bigoplus_{i\in I} U_i$.*

Beweis. Beachte, dass $\bigcup_{i\in I} B_i$ immer ein Erzeugendensystem des Unterraumes $\sum_{i\in I} U_i$ ist (auch wenn die Familie nicht unbedingt transversal liegt). Wir müssen also nur zeigen, dass die Elemente aus $\bigcup_{i\in I} B_i$ (bezüglich einer festen Aufzählung) linear unabhängig bleiben. Wir nehmen also Indizes $i_1, \ldots, i_m$ aus I, sowie für $1 \leq j \leq m$ Vektoren $v_{i_j 1}, \ldots, v_{i_j r_j}$ aus B_{i_j} und Skalare $\lambda_{i_j 1}, \ldots, \lambda_{i_j r_j}$ derart, dass

$$\sum_{k=1}^{r_1} \lambda_{i_1 k} v_{i_1 k} + \cdots + \sum_{k=1}^{r_m} \lambda_{i_m k} v_{i_m k} = 0_V.$$

Weil das Element $u_j = \sum_{k=1}^{r_j} \lambda_{i_j k} v_{i_j k}$ in U_{i_j} liegt, folgt $u_1 = \ldots = u_m = 0_V$ aus der Bemerkung 1.2. Da jede Teilmenge von B_{i_j} linear unabhängig ist, schließen wir daraus, dass die Skalare $\lambda_{i_j 1} = \ldots = \lambda_{i_j r_j} = 0_{\mathbb{K}}$ für jedes $1 \leq j \leq m$, wie gewünscht. □

Definition 1.4. Gegeben einen Endomorphismus $F : V \to V$ des $\mathbb{K}$-Vektorraumes V, ist der Skalar λ aus $\mathbb{K}$ ein *Eigenwert* von F, falls es einen Vektor $v \neq 0_V$ mit $F(v) = \lambda v$ gibt. Dieser Vektor v ist ein *Eigenvektor* zum Eigenwert λ.

Beachte, dass ein nicht-trivialer Vektor nicht Eigenvektor zu verschiedenen Eigenwerten sein kann.

Definition 1.5. Gegeben einen Endomorphismus $F : V \to V$ des $\mathbb{K}$-Vektorraumes V und einen Skalar λ aus $\mathbb{K}$, ist der *Eigenraum* zu λ die Menge

$$V_\lambda = \{v \in V \mid F(v) = \lambda v\}.$$

Der Eigenraum zu λ ist klarerweise ein Unterraum des Vektorraumes V. Beachte, dass λ genau dann ein Eigenwert von F ist, wenn $\dim_{\mathbb{K}} V_\lambda \neq 0$.

Bemerkung 1.6. Für einen Endomorphismus $F : V \to V$ des $\mathbb{K}$-Vektorraumes V ist die Familie der Eigenräume $(V_\lambda)_{\lambda\in\mathbb{K}}$ transversal.

Beweis. Wir benutzen die Charakterisierung aus der Bemerkung 1.6. Seien also $\lambda_1, \ldots, \lambda_m$ verschiedene Elemente aus $\mathbb{K}$ sowie u_i Vektoren aus V_{λ_i} derart, dass $\sum_{i=0}^m u_i = 0_V$. Wir beweisen induktiv über $m \geq 1$ aus $\mathbb{N}$, dass jedes u_i der Nullvektor ist. Für $m = 1$ ist es trivial, also sei ohne Beschränkung der Allgemeinheit $m > 1$. Wir wenden F an und erhalten die Gleichung $\sum_{i=0}^m \lambda_i u_i = F(0_V) = 0_V$. Insbesondere ist

$$\sum_{i=0}^m \lambda_i u_i - \lambda_m \sum_{i=0}^m u_i = 0_V = \sum_{i=0}^{m-1} (\lambda_i - \lambda_m) u_i.$$

Weil jeder Eigenraum ein Unterraum ist, liegt der Vektor $(\lambda_i - \lambda_m)u_i$ in V_{λ_i} und wir folgern induktiv, dass $(\lambda_i - \lambda_m)u_i = 0_V$. Aus $\lambda_i \neq \lambda_m$ folgt nun $u_i = 0_V$ für jedes $1 \leq i \leq m - 1$. Insbesondere ist u_n auch der Nullvektor, wie gewünscht. □

Bemerkung 1.7. Der Skalar λ ist genau dann ein Eigenwert vom Endomorphismus $F : V \to V$, wenn der Endomorphismus $\lambda \cdot \mathrm{Id}_V - F$ nicht injektiv ist. Für endlichdimensionale Vektorräume V ist dies äquivalent dazu, dass $\lambda \cdot \mathrm{Id}_V - F$ nicht surjektiv ist, wegen [**Skript LAI**, Rangssatz 2.34].

Wenn A die Darstellungsmatrix von F bezüglich einer Basis B des n-dimensionalen Vektorraumes V ist, dann hat $\lambda \cdot \mathrm{Id}_V - F$ die Darstellungsmatrix $\lambda \cdot \mathbf{Id}_n - A$ bezüglich der Basis B. Insbesondere ist λ genau dann ein Eigenwert von F, wenn die Determinante $\det(\lambda \cdot \mathbf{Id}_n - A) = 0_{\mathbb{K}}$.

Definition 1.8. Für einen Endomorphismus $F : V \to V$ des n-dimensionalen $\mathbb{K}$-Vektorraumes V mit Darstellungsmatrix A bezüglich einer Basis B von V ist

$$\chi_F(T) = \det(T \cdot \mathbf{Id}_n - A)$$

das *charakteristische Polynom* von F.

Auf Blatt 11 der Vorlesung Lineare Algebra I haben wir mit Hilfe der Produktformel [**Skript LAI**, Proposition 3.12] gesehen, dass das charakteristische Polynom unabhängig von der Wahl der Basis B (und somit unabhängig von der Darstellungsmatrix) ist. Aus der Leibniz Formel [**Skript LAI**, Aufgabe S. 56] folgerten wir in jenem Blatt, dass das charakteristische Polynom von F normiert vom Grad n ist:

$$\chi_F(T) = T^n + b_{n-1}T^{n-1} + \cdots + b_0.$$

Der konstante Koeffizient b_0 des charakteristischen Polynoms gleicht $(-1_{\mathbb{K}})^n \det(A)$, und $-b_{n-1}$ entspricht der Spur der Matrix $A = (a_{ij})$, wodurch

$$\mathrm{Spur}(A) = \sum_{i=0}^{n} a_{ii}$$

insbesondere nicht von der Wahl der Basis abhängt: Ähnliche quadratische Matrizen haben dieselbe Spur.

Korollar 1.9. *Der Skalar λ ist genau dann ein Eigenwert vom Endomorphismus F eines nichttrivialen endlichdimensionalen Vektorraumes V, wenn λ eine Nullstelle des charakteristischen Polynomes ist.*

Insbesondere besitzt der Endomorphismus nur endlich viele Eigenwerte, wegen des Korollars A.7 im Appendix A.

Im Folgenden ist V ein n-dimensionaler $\mathbb{K}$-Vektorraum.

Proposition 1.10. *Für einen Eigenwert λ des Endomorphismus $F : V \to V$ kann die Dimension $\dim_{\mathbb{K}} V_\lambda$ höchstens die Vielfachheit $\mathrm{ord}_\lambda(\chi_F)$ (siehe Korollar A.6 im Appendix A) von λ sein.*

Den Wert $\mathrm{ord}_\lambda(\chi_F)$ nennen wir die *algebraische Vielfachheit*, wobei der Wert $\dim_{\mathbb{K}} V_\lambda$ die *geometrische Vielfachheit* des Eigenwertes λ ist. Insbesondere ist die algebraische Vielfachheit eine obere Schranke der geometrischen Vielfachheit.

Beweis. Sei $k = \dim_{\mathbb{K}} V_\lambda$. Wir müssen nur zeigen, dass $(T-\lambda)^k$ das charakteristische Polynom $\chi_F(T)$ teilt. Wähle eine Basis $\{v_1, \ldots, v_k\}$ des Eigenraumes V_λ und ergänze sie zu einer Basis $B = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ von V. Sei A die Darstellungsmatrix von F bezüglich B. Das charakteristiche Polynom von F ist dann gegeben durch

$$\chi_F(T) = \det(T \cdot \mathbf{Id}_n - A).$$

Weil die Vektoren v_i im Eigenraum zu λ liegen, d.h. $F(v_i) = \lambda v_i$ für $1 \leq i \leq k$, hat die Matrix A die Form:

$$A = \begin{pmatrix} \lambda \cdot \mathbf{Id}_k & * \\ \mathbf{0}_{n-k} & B \end{pmatrix},$$

wobei B sowie die Nullmatrix $\mathbf{0}_{n-k}$ quadratische Matrizen der Größe $n-k$ sind. Insbesondere ist

$$\chi_F(T) = \det(T \cdot \mathbf{Id}_n - A) = \det \begin{pmatrix} (T-\lambda) \cdot \mathbf{Id}_k & * \\ \mathbf{0}_{n-k} & T \cdot \mathbf{Id}_{n-k} - B \end{pmatrix} = (T-\lambda)^k \det(T \cdot \mathbf{Id}_{n-k} - B),$$

wegen der Kästchenregel [**Skript LAI**, Korollar 3.20]. Insbesondere teilt $(T-\lambda)^k$ das charakteristische Polynom $\chi_F(T)$, wie gewünscht. Beachte, dass wir nicht folgern können, dass k genau dem Wert $\operatorname{ord}_\lambda \chi_F$ entspricht, weil λ eine Nullstelle des Polynoms $\det(T \cdot \mathbf{Id}_{n-k} - B)$ sein könnte. $\square$

1.2 Diagonalisierbarkeit

Im Folgenden seien $\mathbb{K}$ ein Körper und $F : V \to V$ ein Endomorphismus des n-dimensionalen $\mathbb{K}$-Vektorraumes V.

Definition 1.11. Der Endomorphismus $F : V \to V$ ist *diagonalisierbar*, falls $V = \bigoplus_{\lambda \in \mathbb{K}} V_\lambda$ oder äquivalent dazu, wenn V eine Basis aus Eigenvektoren besitzt (wegen des Lemmas 1.3).

Weil V endlichdimensional ist, ist es auch äquivalent dazu, dass es eine Basis B von V derart gibt, dass die Darstellungsmatrix von F bezüglich B (im Definition- und im Bildbereich) eine Diagonalmatrix ist.

Eine quadratische $n \times n$-Matrix A ist *diagonalisierbar*, wenn der induzierte Endomorphismus $F_A : \mathbb{K}^n \to \mathbb{K}^n$ diagonalisierbar ist, oder äquivalent dazu, wenn A ähnlich zu einer $n \times n$-Diagonalmatrix ist (siehe [**Skript LAI**, Korollar 2.64]).

Aufgabe. Ist jeder Endomorphismus $F : V \to V$ eines 1-dimensionalen $\mathbb{K}$-Vektorraumes V diagonalisierbar?

Lemma 1.12. *Der Endomorphismus $F : V \to V$ ist genau dann diagonalisierbar, wenn*

$$n = \sum_{\lambda \text{ Eigenwert von } F} \dim_{\mathbb{K}} V_\lambda$$

Beachte, dass die obige Summe wegen des Korollars 1.9 nur aus endlich vielen Summanden besteht und somit wohldefiniert ist.

Beweis. Wir müssen lediglich feststellen, dass die direkte Summe $\bigoplus_{\lambda \text{ Eigenwert}} V_\lambda$ Dimension

$$\sum_{\lambda \text{ Eigenwert von } F} \dim_{\mathbb{K}} V_\lambda$$

hat, weil die Dimensionsformel [**Skript LAI**, Satz 2.16] gilt. $\square$

Aufgabe. Ist die Matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ diagonalisierbar?

Korollar 1.13. *Der Endomorphismus $F : V \to V$ ist genau dann diagonalisierbar, wenn das charakteristische Polynom $\chi_F(T)$ in (möglicherweise nicht paarweise verschiedene) Linearfaktoren*

$$\chi_F(T) = (T - \lambda_1) \cdots (T - \lambda_n)$$

zerfällt und für jedes $1 \leq i \leq n$ die algebraische Vielfachheit von λ_i gleich der geometrischen Vielfachheit ist, d.h.

$$\mathrm{ord}_{\lambda_i}(\chi_F) = \dim_{\mathbb{K}} V_{\lambda_i}.$$

Beweis. Wenn F diagonalisierbar ist, besitzt V eine Basis B von Eigenvektoren von F. Seien $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von F und $d_j = \dim_{\mathbb{K}} V_{\lambda_j}$. Für alle $0 \leq j \leq k-1$ setze $d'_j = \sum_{r \leq j} d_r$, wobei die leere Summe Null ist, und ordne die Basis B so an, dass die Vektoren $v_{d'_j+1}, \ldots, v_{d'_{j+1}}$ eine Basis des Eigenraumes $V_{\lambda_{j+1}}$ bilden. Die Darstellungsmatrix von F bezüglich der Basis B ist in Diagonalform

$$\begin{pmatrix} \underbrace{\lambda_1 \ \ddots \ \lambda_1}_{d_1 - \text{mal}} & & & \\ & \underbrace{\lambda_2 \ \ddots \ \lambda_2}_{d_2 - \text{mal}} & & \\ & & \ddots & \\ & & & \underbrace{\lambda_k \ \ddots \ \lambda_k}_{d_k - \text{mal}} \end{pmatrix}$$

und es folgt, dass das charakteristische Polyonom in Linearfaktoren zerfällt, in der Tat ist $\chi_F(T) = (T - \lambda_1)^{d_1} \cdots (T - \lambda_k)^{d_k}$. Weil die Eigenwerte paarweise verschieden sind, gilt insbesondere $\mathrm{ord}_{\lambda_i}(\chi_F) = d_i = \dim_{\mathbb{K}} V_{\lambda_i}$, wie gewünscht.

Die andere Richtung folgt sofort aus der Proposition 1.10 und dem Lemma 1.3, weil der Grad n des charakteristischen Polynom genau der Dimension des $\mathbb{K}$-Vektorraumes V entspricht. $\square$

Etliche Resultate in diesem Abschnittes lassen sich induktiv beweisen. Hierfür brauchen wir die folgende Definition sowie das Hilfslemma, welches uns den Induktionsschritt ermöglichen wird.

Definition 1.14. Ein Unterraum U des Vektorraumes V ist *F-invariant*, falls $F(U) \subset U$, oder äquivalent dazu, wenn die Einschränkungsabbildung $F_{\restriction U}$ einen Endomorphismus von U definiert.

Aufgabe. Sei λ ein beliebiges Element des Körpers $\mathbb{K}$ und U ein Unterraum von V. Zeige, dass U genau dann F-invariant ist, wenn U invariant unter dem Endomorphismus $F - \lambda \cdot \mathrm{Id}_V$ ist.

Lemma 1.15. *Ist U ein F-invarianter Unterraum von V, dann induziert F einen wohldefinierten Endomorphismus $\widetilde{F} : V/U \to V/U$ auf dem Quotientenraum V/U (siehe [**Skript LAI**, Proposition 2.41]) derart, dass*

$$\chi_F(T) = \chi_{F\restriction_U}(T) \cdot \chi_{\widetilde{F}}(T).$$

Beweis. Die Abbildung

$$\begin{array}{rccc} \widetilde{F}: & V/U & \to & V/U \\ & v+U & \mapsto & F(v)+U \end{array}$$

ist wohldefiniert, weil der Unterraum U invariant unter F ist. Sie ist klarerweise linear, wegen der Linearität von F sowie der Vektorraumstruktur des Quotientenraumes. Weil das charakteristische Polynom unabhängig von der Wahl der Basis ist, wähle eine Basis $B = \{v_1, \ldots, v_n\}$ von V derart, dass $\{v_1, \ldots, v_k\}$ eine Basis von U ist. Der Beweis des [**Skript LAI**, Korollars 2.43], welches besagt, dass jeder Komplementärraum von U zum Quotientenraum V/U isomorph ist, liefert, dass die Kollektion der Nebenklassen $\{v_{k+1} + U, \ldots, v_n + U\}$ eine Basis von V/U bilden.

Beachte nun, dass die Darstellungsmatrix des Endomorphismus F bezüglich der Basis B von der Form

$$\begin{pmatrix} A & * \\ \mathbf{0}_{n-k} & C \end{pmatrix}$$

ist, wobei die $k \times k$-Matrix A die Darstellungsmatrix der Einschränkung $F\restriction_U$ bezüglich der Basis $\{v_1, \ldots, v_k\}$ von U ist, da $F(v_j)$ in U liegt für alle $1 \leq j \leq k$.

Weil wegen der Kästchenregel

$$\chi_F(T) = \det\left(T \cdot \mathbf{Id}_n - \begin{pmatrix} A & * \\ \mathbf{0}_{n-k} & C \end{pmatrix}\right) = \det\begin{pmatrix} T \cdot \mathbf{Id}_k - A & * \\ \mathbf{0}_{n-k} & T \cdot \mathbf{Id}_{n-k} - C \end{pmatrix} =$$

$$= \det(T \cdot \mathbf{Id}_k - A) \cdot \det(T \cdot \mathbf{Id}_{n-k} - C) = \chi_{F\restriction_U}(T) \cdot \det(T \cdot \mathbf{Id}_{n-k} - C),$$

müssen wir nur noch zeigen, dass C die Darstellungsmatrix der Abbildung $\widetilde{F}$ bezüglich der Basis $\{v_{k+1} + U, \ldots, v_n + U\}$ des Quotientenraumes V/U ist: Sei j in $\{k+1, \ldots, n\}$. Wenn

$$F(v_j) = \sum_{1 \leq i \leq n} \lambda_{ij} v_i,$$

so ist

$$\widetilde{F}(v_j + U) = F(v_j) + U = \sum_{k+1 \leq i \leq n} \lambda_{ij}(v_i + U),$$

wie gewünscht. $\square$

Wir werden nun sehen, dass sich der Endomorphismus gut darstellen lässt, wenn das charakteristische Polyonom in Linearfaktoren zerfällt, auch wenn die algebraischen und geometrischen Vielfachheiten nicht übereinstimmen.

Definition 1.16. Der Endomorphismus $F : V \to V$ ist *trigonalisierbar*, falls es eine Basis B von V derart gibt, dass die Darstellungsmatrix von F bezüglich B (sowohl im Definitions- als auch im Bildbereich) in oberer Dreiecksform

$$\begin{pmatrix} \lambda_1 & * & \cdots & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}$$

ist.

Satz 1.17. *Der Endomorphismus $F : V \to V$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom χ_F in Linearfaktoren zerfällt.*

Insbesondere lässt sich jeder Endomorphismus eines endlichdimensionalen $\mathbb{C}$-Vektorraumes trigonalisieren.

Beweis. Wenn der Endomorphimus F eine Darstellungsmatrix (bezüglich einer Basis B) in oberer Dreiecksform besitzt, zerfällt das charakteristische Polyonom wegen der Kästchenregel klarerweise in Linearfaktoren, also müssen wir nur die andere Richtung induktiv über die Dimension $n = \dim_{\mathbb{K}} V \geq 1$ beweisen. Für $n = 1$ ist die Aussage trivial, weil jede 1×1-Matrix bereits in oberer Dreiecksform ist. Beachte, dass hier dann auch jeder nicht-triviale Vektor ein Eigenvektor ist.

Wir nehmen also an, dass $n = \dim_{\mathbb{K}} V \geq 2$. Das charakteristische Polynom $\chi_F(T)$ zerfällt in Linearfaktoren, d.h.

$$\chi_F(T) = (T - \lambda_1) \cdots (T - \lambda_n).$$

Insbesondere ist λ_1 ein Eigenwert von F nach dem Korollar 1.9. Wähle nun einen Eigenvektor $v_1 \neq 0_V$ in V zu λ_1 und setze $U = \mathrm{Lin}(\{v_1\})$. Der Unterraum U hat Dimension 1 und ist F-invariant, weil $F(\mu v_1) = (\mu \cdot \lambda_1)v_1 = \lambda_1(\mu v_1)$. Da die Einschränkung $F_{\restriction U}$ gerade der Skalarmultiplikation mit λ_1 entspricht, ist $\chi_{F_{\restriction U}}(T) = T - \lambda_1$.

Aus dem Hilfslemma 1.15 folgt $\chi_F(T) = \chi_{F_{\restriction U}}(T) \cdot \chi_{\widetilde{F}}(T) = (T - \lambda_1) \cdot \chi_{\widetilde{F}}(T)$, wobei

$$\begin{array}{rccc} \widetilde{F}: & V/U & \to & V/U \\ & v + U & \mapsto & F(v) + U \end{array}$$

ein Endomorphimus des Quotientenraumes V/U ist. Da Polynomringe Integritätsbereiche sind, erhalten wir

$$\chi_{\widetilde{F}}(T) = (T - \lambda_2) \cdots (T - \lambda_n),$$

was in Linearfaktoren zerfällt. Beachte, dass $\dim_{\mathbb{K}} V/U = n - 1 < n$, nach dem noetherschen Isomorphiesatz [**Skript LAI**, Korollar 2.46]. Wähle induktiv eine Basis $\{v_2 + U, \ldots, v_n + U\}$ vom Quotientenraum V/U derart, dass sich $\widetilde{F}$ bezüglich der Basis $\{v_2 + U, \ldots, v_n + U\}$ in oberer Dreiecksform

$$\begin{pmatrix} \lambda_2 & * & \cdots & \cdots & * \\ 0 & \lambda_3 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}$$

darstellen lässt. Wir müssen nur noch zeigen, dass die Familie $\{v_1, \ldots, v_n\}$ eine Basis von V bildet, da die dazugehörige Darstellungsmatrix von F klarerweise in oberer Dreiecksform ist. Weil die Dimension von V genau n ist, genügt es zu zeigen, dass die Vektoren linear unabhängig sind. Seien die Skalare $\mu_1, \ldots, \mu_n$ derart, dass $\sum_{i=1}^{n} \mu_i v_i = 0_V$. Dann

$$\mu_2 v_2 + \cdots + \mu_n v_n = -\mu_1 v_1 \in U,$$

also gilt

$$\mu_2(v_2 + U) + \cdots + \mu_n(v_n + U) = 0_V + U.$$

Es folgt $\mu_2 = \ldots = \mu_n = 0_K$ und somit auch $\mu_1 = 0_K$, wie gewünscht. □

1.3 Der Satz von Cayley-Hamilton

Im Folgenden seien $\mathbb{K}$ ein Körper und $F : V \to V$ ein Endomorphismus des n-dimensionalen $\mathbb{K}$-Vektorraumes V.

Notation. Definiere rekursiv über m aus $\mathbb{N}$ die Abbildungen

$$F^0 = \mathrm{Id}_V \text{ und } F^m = \underbrace{F \circ \cdots \circ F}_{m\text{-mal}} \text{ für } m > 0.$$

Beachte, dass jede Potenz F^m ein Endomorphismus von V ist.

Bemerkung 1.18. Für jeden nicht-trivialen Vektor $v \neq 0_V$ aus V muss die Familie $\{F^m(v)\}_{m\in\mathbb{N}}$ linear abhängig sein (weil V endlichdimensional ist). Wähle also k aus $\mathbb{N}$ minimal derart, dass

$$F^k(v) = \sum_{j=0}^{k-1} a_j F^j(v)$$

für $a_0, \ldots, a_{k-1}$ aus $\mathbb{K}$. Beachte, dass $k \geq 1$.

Es lässt sich leicht zeigen, dass die Menge $\{v, F(v), \ldots, F^{k-1}(v)\}$ eine Basis des F-invarianten Unterraumes $U = \mathrm{Lin}(\{F^m(v)\}_{m\in\mathbb{N}})$ bildet. Insbesondere ist das charakteristische Polynom der Einschrankung $F_{\restriction U}$ gleich

$$\chi_{F_{\restriction U}}(T) = T^k - a_{k-1}T^{k-1} - \ldots - a_0.$$

Ein solcher Unterraum ist ein *zyklischer Unterraum (bezüglich F) der Ordnung k.*

Definition 1.19. Gegeben ein Polynom $P(T) = \sum_{j=0}^{D} a_j T^j$ mit Koeffizienten aus $\mathbb{K}$ definiert

$$\begin{aligned} P(F): \ V &\to V \\ v &\mapsto \sum_{j=0}^{D} a_j F^j(v) \end{aligned}$$

den Endomorphismus $P(F) = \sum_{j=0}^{D} a_j F^j$. Wir ersetzen also jedes Monom T^j durch den Endomorphimus F^j. Weil lineare Abbildungen unter Summen und Skalarmultiplikationen abgeschlossen sind, ist $P(F)$ wiederum ein Endomorphismus von V.

Aufgabe. Zeige induktiv über den Grad des Polynomes P, dass $Q(F) \circ P(F) = (Q \cdot P)(F)$ (als Abbildung) für jedes Polynom $Q(T)$ mit Koeffizienten aus $\mathbb{K}$. Insbesondere kommutieren je zwei solche Endomorphismen, obwohl beliebige Endomorphismen von V allgemein nicht kommutieren müssen.

Satz 1.20. (*Der Satz von Cayley-Hamilton*) *Der Endomorphismus $\chi_F(F)$ ist die Nullabbildung auf V.*

Beweis. Wir müssen zeigen, dass $\chi_F(F)(v) = 0_V$ für alle v aus V. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $v \neq 0_V$. Weil V endlichdimensional ist, gibt es k aus $\mathbb{N}$ minimal derart, dass

$$F^k(v) = \sum_{j=0}^{k-1} a_j F^j(v)$$

für $a_0, \ldots, a_{k-1}$ aus $\mathbb{K}$, wie in der Bemerkung 1.18. Weil der Unterraum $U = \mathrm{Lin}(\{F^m(v)\}_{m\in\mathbb{N}})$ invariant unter F ist, folgt aus dem Hilfslemma 1.15, dass

$$\chi_F(T) = \chi_{F_{\restriction U}}(T) \cdot \chi_{\widetilde{F}}(T).$$

Beachte, dass $\chi_{F_{\restriction U}}(F)(v) = 0_V$ nach Wahl der Koeffizienten $a_0, \ldots, a_{k-1}$. Insbesondere ist

$$\chi_F(F)(v) = \Big(\chi_{F_{\restriction U}}(F) \circ \chi_{\widetilde{F}}(F)\Big)(v) = \Big(\chi_{\widetilde{F}}(F) \circ \chi_{F_{\restriction U}}(F)\Big)(v) = \chi_{\widetilde{F}}(F)(0_V) = 0_V,$$

wie gewünscht. □

Korollar 1.21. *Für eine $n \times n$-Matrix A mit Koeffizienten aus $\mathbb{K}$ und charakteristischem Polynom $T^n + b_{n-1}T^{n-1} + \cdots + b_0$ gilt*

$$A^n + b_{n-1}A^{n-1} + \cdots + b_0 \cdot \mathbf{Id}_n = \mathbf{0}_n,$$

wobei $\mathbf{0}_n$ die Nullmatrix der Größe $n \times n$ ist.

Wir können nun ein normiertes Polynom kleinsten Grades derart wählen, dass F eine *Nullstelle* (im Sinne der Definition 1.19) ist. Dieses Polynom, genannt das *Minimalpolynom* von F, enthält bereits alle Informationen, um zu überprüfen, ob der Endomorphismus F diagonalisierbar ist.

Satz 1.22. *Es existiert genau ein normiertes Polynom $m_F(T)$ mit Koeffizienten aus $\mathbb{K}$ derart, dass $m_F(F)$ der triviale Endomorphismus von V ist und für alle Polynome $P(T)$ mit Koeffizienten aus $\mathbb{K}$*

$$P(F) = 0_{\mathrm{End}(V)} \iff m_F(T) \text{ teilt } P(T) \text{ (als Polynome).}$$

Das Polynom $m_F(T)$ heißt das Minimalpolynom *von F.*

Beweis. Wir zeigen zuerst die Existenz des Minimalpolynoms: Nach dem Satz von Cayley-Hamilton 1.20 ist die Menge

$$\mathcal{P} = \{P(T) \in \mathbb{K}[T] \mid P(T) \text{ normiert und } P(F) = 0_{\mathrm{End}(V)}\}$$

nicht leer. Wähle also ein Polynom $m_F(T)$ in $\mathcal{P}$ kleinsten Grades. Dann ist $m_F(F)$ der triviale Endomorphismus von V und wir müssen nur die letzte Äquivalenz zeigen: Eine Richtung ist klar, denn wenn $P(T)$ ein Vielfaches von m_F ist, schreibe $P(T) = m_F(T) \cdot Q(T)$ für ein Polynom $Q(T)$ mit Koeffizienten aus $\mathbb{K}$. Nach der vorigen Aufgabe ist

$$P(F) = m_F(F) \circ Q(F) = Q(F) \circ m_F(F) = Q(F) \circ 0_{\mathrm{End}(V)} = 0_{\mathrm{End}(V)},$$

wie gewünscht. Wir nehmen nun an, dass das Polynom $P(T)$ den Endomorphismus F als Nullstelle (im Sinne der Definition 1.19) besitzt. Wenn $P(T)$ das triviale Polynom ist, dann ist $P(T)$ ein Vielfaches von $m_F(T)$, also können wir ohne Beschränkung der Allgemeinheit annehmen, dass $P(T)$ nicht-trivialen Führungskoeffizient a_D besitzt, wobei $D = \deg(P)$. Beachte, dass $P(F)$ genau dann trivial ist, wenn $(a_D^{-1} \cdot P)(F)$ trivial ist. Wir können also annehmen, dass $P(T)$ normiert ist. Insbesondere liegt $P(T)$ in $\mathcal{P}$, also $\deg(P) \geq \deg(m_F)$, wegen der Minimalität des Grades des Minimalpolynoms. Mit Hilfe der Division mit Rest A.4 im Appendix A schreibe

$$P(T) = m_F(T) \cdot Q(T) + R(T),$$

wobei $Q(T)$ und $R(T)$ Polynome in $\mathbb{K}[T]$ mit $\deg(R) < \deg(m_F)$ sind. Wir müssen nur noch folgern, dass $R(T)$ das Nullpolynom ist. Hierfür genügt es zu zeigen, dass $R(F)$ der triviale Endomorphismus ist, denn wäre R nicht trivial, könnten wir R normieren, was der Minimalität des Grades von $m_F(T)$ als Element in $\mathcal{P}$ widerspricht. Nun ist

$$0_{\mathrm{End}(V)} = P(F) = Q(F) \circ m_F(F) + R(F) = 0_{\mathrm{End}(V)} + R(F) = R(F),$$

wie gewünscht.

Für die Eindeutigkeit nehmen wir an, dass das Polynom $m'_F(T)$ die obigen Eigenschaften besitzt. Insbesondere ist $m'_F(F) = 0_{\mathrm{End}(V)}$, also $m'_F(T) = m_F(T) \cdot Q(T)$ für ein Polynom $Q(T)$, welches normiert sein muss (weil $m_F(T)$ und $m'_F(T)$ beide normiert sind). Analog ist $m_F(T) = m'_F(T) \cdot H(T)$ für ein normiertes Polynom $H(T)$. Insbesondere ist

$$m_F(T) = m'_F(T) \cdot H(T) = m_F(T) \cdot Q(T) \cdot H(T).$$

Da der Polynomring ein Integritätsbereich ist, folgt, dass $Q(T) \cdot H(T)$ das konstanten Polynom $1_{\mathbb{K}} = T^0$ sein muss. Insbesondere sind sowohl $Q(T)$ als auch $H(T)$ konstant, und somit gleich 1 (weil sie normiert sind). Es folgt $m_F(T) = m'_F(T)$, wie gewünscht. □

Korollar 1.23. *Jede $n \times n$-Matrix A über $\mathbb{K}$ besitzt ein* Minimalpolynom $m_A(T)$, *das heißt, ein normiertes Polynom $m_A(T) = T^k + b_{k-1}T^{k-1} + \cdots + b_0$ derart, dass*

- *die Matrix $A^k + b_{k-1}A^{k-1} + \cdots + b_0 \cdot \mathbf{Id}_n = \mathbf{0}_n$ und*
- *das Polynom m_A genau dann ein Polynom $P(T) = \sum_{i=0}^{D} c_i T^D$ teilt, wenn*

$$c_D A^D + \cdots + c_0 \mathbf{Id}_n = \mathbf{0}_n.$$

Lemma 1.24. *Das charakteristische und das Minimalpolynom besitzen dieselben Nullstellen. Insbesondere ist λ genau dann eine Nullstelle von $m_F(T)$, wenn λ ein Eigenwert von F ist.*

Beweis. Weil das Minimalpolynom $m_F(T)$ das charakteristische Polynom χ_F teilt, ist jede Nullstelle von $m_F(T)$ eine Nullstelle des charakteristischen Polynoms. Sei nun λ eine Nullstelle des charakteristischen Polynoms. Insbesondere ist der Skalar λ ein Eigenwert bezüglich F und es gibt einen Eigenvektor $v \neq 0_V$ zu λ.

Wir wollen zeigen, dass $m_F(\lambda) = 0_{\mathbb{K}}$. Mit Hilfe des Korollars A.6 schreibe

$$m_F(T) = (T - \lambda)Q(T) + m_F(\lambda),$$

für ein Polynom $Q(T)$ aus $\mathbb{K}[T]$ (es wird hier nicht verlangt, dass $Q(\lambda) \neq 0_{\mathbb{K}}$). Aus dem Satz 1.22 folgt

$$0_V = m_F(F)(v) = Q(F)(F(v) - \lambda \cdot v) + m_F(\lambda) \cdot v = Q(F)(0_V) + m_F(\lambda) \cdot v = m_F(\lambda) \cdot v,$$

also $m_F(\lambda) = 0_K$, weil v verschieden vom Nullvektor ist. □

Mit Hilfe des Minimalpolynoms können wir ein Kriterium für die Diagonalisierbarkeit von Endomorphismen (bzw. von Matrizen) liefern. Dafür benötigen wir zuerst folgenden Hilfssatz.

Lemma 1.25. *Wenn das Minimalpolynom in lauter verschiedene Linearfaktoren zerfällt, ist*

$$V = V_\lambda \oplus \operatorname{Im}(F - \lambda \cdot \operatorname{Id}_V)$$

für jeden Eigenwert λ.

Beachte, dass die obige Behauptung auch gilt, wenn λ kein Eigenwert ist, denn in diesem Fall ist $\operatorname{Ker}(F - \lambda \cdot \operatorname{Id}_V)$ der triviale Unterraum und somit $F - \lambda \cdot \operatorname{Id}_V$ ein Automorphismus (hierfür haben wir nicht gebraucht, dass $m_F(T)$ in lauter verschiedene Linearfaktoren zerfällt).

Beweis. Nach dem [**Skript LAI**, Rangssatz 2.34] müssen wir nur zeigen, dass die Unterräume V_λ und $\operatorname{Im}(F - \lambda \cdot \operatorname{Id}_V)$ von V transversal liegen. Wir nehmen $m_F(T) = (T - \lambda_1) \cdots (T - \lambda_k)$ an, wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von F sind (nach dem Lemma 1.24). Wenn $k = 1$, folgt aus $m_F(F) = 0_{\operatorname{End}(V)}$, dass $F(v) = \lambda_1 \cdot v$ für alle v aus V, also $\operatorname{Im}(F - \lambda_1 \cdot \operatorname{Id}_V) = \{0_V\}$. Die Behauptung folgt in diesem Fall trivialerweise.

Ohne Beschränkung der Allgemeinheit ist $k \geq 2$. Setze $R(T) = (T - \lambda_2) \cdots (T - \lambda_k)$, dann $m_F(T) = R(T) \cdot (T - \lambda_1)$. Sei nun

$$v \text{ in } V_{\lambda_1} \cap \operatorname{Im}(F - \lambda_1 \cdot \operatorname{Id}_V),$$

also $F(v) = \lambda_1 \cdot v$ und es gibt einen Vektor w aus V mit $v = (F - \lambda_1 \cdot \operatorname{Id}_V)(w) = F(w) - \lambda_1 \cdot w$. Wir müssen zeigen, dass $v = 0_V$. Mit Hilfe der Division mit Rest A.6 schreibe

$$R(T) = Q(T) \cdot (T - \lambda_1) + R(\lambda_1),$$

für ein Polynom $Q(T)$ mit Koeffizienten aus $\mathbb{K}$. Beachte, dass der Skalar $R(\lambda_1) \neq 0_{\mathbb{K}}$, weil die Eigenwerte alle verschieden sind. Nun ist

$$R(F)(v) = Q(F) \circ (F - \lambda_1 \operatorname{Id}_V)(v) + R(\lambda_1) \cdot v = Q(F)(0_V) + R(\lambda_1) \cdot v = R(\lambda_1) \cdot v.$$

Andererseits ist

$$R(F)(v) = R(F)\left((F - \lambda_1 \cdot \operatorname{Id}_V)(w)\right) = R(F) \circ (F - \lambda_1 \cdot \operatorname{Id}_V)(w) = m_F(F)(w) = 0_{\operatorname{End}(V)}(w) = 0_V,$$

und somit folgt $v = 0_V$, wie gewünscht. □

Satz 1.26. *Der Endomorphismus $F : V \to V$ ist genau dann diagonalisierbar, wenn das Minimalpolynom m_F in lauter verschiedene Linearfaktoren zerfällt.*

Beweis. $\Longrightarrow$: Wenn F diagonalisierbar ist, schreibe

$$\chi_F(T) = (T-\lambda_1)^{d_1} \cdots (T-\lambda_k)^{d_k},$$

wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von F sind und $d_i = \dim_{\mathbb{K}} V_{\lambda_i}$ für $1 \leq i \leq k$ nach dem Korollar 1.13. Um zu zeigen, dass das Minimalpolynom $m_F(T)$ in verschiedene Linearfaktoren zerfällt, genügt es zu zeigen, dass die Einsetzung von F in das Polynom

$$P(T) = (T-\lambda_1) \cdots (T-\lambda_k)$$

den Nullendomorphismus auf V liefert: In diesem Fall teilt $m_F(T)$ das Polynom $P(T)$, aber sie sind beide normiert und haben denselben Grad (weil sie dieselben Nullstellen haben), also $P(T) = m_F(T)$, wie gewünscht.

Da V die direkte Summe der Eigenräume ist, genügt es zu zeigen, dass die Einschränkung von $P(F)$ auf den Eigenraum V_{λ_i} trivial ist. Sei also v ein Vektor aus dem Eigenraum V_{λ_i}. Beachte, dass

$$\begin{aligned} P(F)(v) &= (F-\lambda_1\cdot \mathrm{Id}_V)\circ\cdots\circ(F-\lambda_k\cdot\mathrm{Id}_V)(v) = \\ &= (F-\lambda_1\cdot \mathrm{Id}_V)\circ\cdots\circ(F-\lambda_{i-1}\cdot\mathrm{Id}_V)\circ(F-\lambda_{i+1}\cdot\mathrm{Id}_V)\circ\cdots\circ(F-\lambda_k\cdot\mathrm{Id}_V)\circ(F-\lambda_i\cdot\mathrm{Id}_V)(v) = \\ &(F-\lambda_1\cdot \mathrm{Id}_V)\circ\cdots\circ(F-\lambda_{i-1}\cdot\mathrm{Id}_V)\circ(F-\lambda_{i+1}\cdot\mathrm{Id}_V)\circ\cdots\circ(F-\lambda_k\cdot\mathrm{Id}_V)(0_V) = 0_V, \end{aligned}$$

und somit gilt die Behauptung.

$\Longleftarrow$: Wir nehmen nun an, dass $m_F(T) = (T-\lambda_1)\cdots(T-\lambda_k)$, wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von F sind (nach dem Lemma 1.24). Wir zeigen induktiv über $n = \dim_{\mathbb{K}} V$, dass V die (direkte) Summe der Eigenräume $V_{\lambda_1}, \ldots, V_{\lambda_k}$ ist, siehe Bemerkung 1.6. Der Fall $n = 1$ folgt sofort (siehe die Aufgabe vor Lemma 1.12), also sei ohne Beschränkung der Allgemeinheit $n \geq 2$. Falls $k = 1$, folgt aus $m_F(F) = 0_{\mathrm{End}(V)}$, dass $V = V_{\lambda_1}$ und wir sind fertig.

Wir nehmen also an, dass $k \geq 2$. Nach dem Lemma 1.25 ist

$$V = V_{\lambda_1} \oplus \mathrm{Im}(F - \lambda_1 \cdot \mathrm{Id}_V).$$

Weil F einen Eigenvektor zum Eigenwert λ_1 besitzt, ist V_{λ_1} nicht trivial und somit hat der Unterraum $W = \mathrm{Im}(F - \lambda_1 \cdot \mathrm{Id}_V)$ Dimension höchstens $n-1$. Wir werden zeigen, dass die Einschränkung von F auf W einen Endomorphismus von W derart definiert, dass das entsprechende Minimalpolynom das Produkt

$$(T-\lambda_2)\cdots(T-\lambda_k)$$

ist. Nach der Induktionsannahme ist W die direkte Summe der entsprechenden Eigenräume und somit folgt die Behauptung auch für V.

Um zu zeigen, dass W invariant unter F ist, sei w ein Vektor aus W. Dann ist $w = F(v) - \lambda_1 \cdot v$ für ein v aus V und

$$F(w) = F(F(v) - \lambda_1\cdot v) = F^2(v) - \lambda_1 \cdot F(v) = (F - \lambda_1\cdot\mathrm{Id}_V)(F(v))$$

liegt klarerweise in W, wie gewünscht.

Wir müssen nur noch das Minimalpolynom von $F' = F_{\restriction W} : W \to W$ bestimmen. Hierfür setzen wir $R(T) = (T-\lambda_2)\cdots(T-\lambda_k)$, also $m_F(T) = R(T)\cdot(T-\lambda_1)$, und betrachten den

Endomorphismus $R(F')$ von W. Jeder Vektor w aus W lässt sich als $w = (F - \lambda_1 \cdot \mathrm{Id}_V)(v)$ für einen Vektor v aus V schreiben. Wegen

$$R(F')(w) = R(F)\big((F - \lambda_1 \cdot \mathrm{Id}_V)(v)\big) = (R(T) \cdot (T - \lambda_1))\,(F)(v) = m_F(F)(v) = 0_V = 0_W,$$

teilt $m_{F'}(T)$ das Polynom $R(T)$. Schreibe also

$$R(T) = m_{F'} \cdot H$$

für ein normiertes Polynom H mit Koeffizienten aus $\mathbb{K}$. Wir zeigen nun, dass $m_F(T)$ das Polynom $m_{F'}(T)\cdot(T-\lambda_1)$ teilen muss. Es genügt zu zeigen, dass der von $m_{F'}(T)\cdot(T-\lambda_1)$ induzierte Endomorphismus trivial auf V ist: Für v aus V ist $w = (T - \lambda_1)(F)(v) = (F - \lambda_1 \cdot \mathrm{Id}_V)(v)$ ein Element aus W, also

$$m_{F'} \cdot (T - \lambda_1)(F)(v) = m_{F'}(F)(w) = m_{F'}(F')(w) = 0_W = 0_V,$$

wie gewünscht.

Es folgt, dass $m_F(T)$ das normierte Polynom $m_{F'}(T) \cdot (T - \lambda_1)$ teilt und somit

$$m_{F'}(T) \cdot (T - \lambda_1) = Q(T) \cdot m_F(T),$$

für ein Polynom $Q(T)$ mit Koeffizienten aus $\mathbb{K}$. Beachte, dass $Q(T)$ auch normiert sein muss. Nun ist

$$\begin{aligned} m_{F'}(T) \cdot (T - \lambda_1) &= Q(T) \cdot (T - \lambda_1) \cdots (T - \lambda_k) = \\ &= Q(T) \cdot R(T) \cdot (T - \lambda_1) = Q(T) \cdot H(T) \cdot m_{F'}(T) \cdot (T - \lambda_1). \end{aligned}$$

Da Polynomringe Integritätsbereiche sind, folgt, dass das normierte Polynom $Q(T) \cdot H(T)$ das konstante Polynom $1_{\mathbb{K}} = T^0$ sein muss und somit $Q = H = 1_{\mathbb{K}}$. Insbesondere zerfällt das Minimalpolynom

$$m_{F'} = (T - \lambda_2) \cdots (T - \lambda_k)$$

von $F' = F_{\restriction W}$ in Linearfaktoren. Nach der Induktionsannahme ist $F_{\restriction W}$ diagonalisierbar, also

$$W = \bigoplus_{i=2}^{k} W_{\lambda_i}.$$

Da W_{λ_i} für $2 \leq i \leq k$ ein Unterraum von V_{λ_i} ist, folgt

$$V = V_{\lambda_1} \oplus W = V_{\lambda_1} \oplus \bigoplus_{i=2}^{k} W_{\lambda_i} \subset V_{\lambda_1} \oplus \bigoplus_{i=2}^{k} V_{\lambda_i} = \bigoplus_{i=1}^{k} V_{\lambda_1} \subset V$$

und wir schließen, dass F diagonalisierbar ist, weil V die direkte Summe der Eigenräume ist.

□

1.4 Nilpotente Endomorphismen

Im Folgenden seien $\mathbb{K}$ ein Körper und $F : V \to V$ ein Endomorphismus des n-dimensionalen $\mathbb{K}$-Vektorraumes V.

Bemerkung 1.27. Weil V endlichdimensional ist, muss es für die steigende Kette

$$\mathrm{Ker}(F) \subset \mathrm{Ker}(F^2) \subset \mathrm{Ker}(F^3) \subset \ldots \subset \mathrm{Ker}(F^r) \subset \ldots$$

eine natürliche Zahl m derart geben, dass $\mathrm{Ker}(F^m) = \mathrm{Ker}(F^{m+1})$. Beachte, dass der kleinstmögliche solche Wert höchstens n ist.

Es lässt sich leicht induktiv über k aus $\mathbb{N}$ zeigen, dass $\mathrm{Ker}(F^m) = \mathrm{Ker}(F^{m+k})$. Ferner ist

$$V = \mathrm{Ker}(F^m) \oplus \mathrm{Im}(F^m)$$

für den kleinstmöglichen solchen Wert m: Analog zum Beweis von 1.25 müssen wir nur zeigen, dass $\mathrm{Ker}(F^m)$ und $\mathrm{Im}(F^m)$ transversal zueinander liegen. Nun, wenn der Vektor v im Durchschnitt $\mathrm{Ker}(F^m) \cap \mathrm{Im}(F^m)$ liegt, schreibe $v = F^m(w)$ für einen Vektor w aus V. Da

$$0_V = F^m(v) = F^{2m}(w),$$

liegt der Vektor w im Unterraum $\mathrm{Ker}(F^{2m})$. Nach unserer Annahme ist $\mathrm{Ker}(F^{2m}) = \mathrm{Ker}(F^m)$ und somit $v = F^m(w) = 0_V$, wie gewünscht.

Beachte, dass in der obigen Situation die Einschränkung von F^m auf den Unterraum $\mathrm{Ker}(F^m)$ (welcher offensichtlich F-invariant ist) der triviale Endomorphismus ist. Diese Beobachtung motiviert den folgenden Begriff.

Definition 1.28. Der Endomorphimus $F : V \to V$ ist *nilpotent*, falls es eine natürliche Zahl m so gibt, dass F^m der triviale Endomorphismus auf V ist.

Die kleinstmögliche solche Zahl $m \geq 1$ ist der *Nilpotenzindex oder -grad* des nilpotenten Endomorphismus F.

Eine $n \times n$-Matrix A ist *nilpotent*, falls die von A induzierte Abbildung $F_A : \mathbb{K}^n \to \mathbb{K}^n$ nilpotent ist, oder äquivalent dazu, wenn es eine natürliche Zahl m so gibt, dass

$$A^m = \underbrace{A \cdots A}_{m\text{-mal}}$$

die Nullmatrix ist.

Nilpotenz lässt sich mit Hilfe des charakteristischen Polynoms untersuchen.

Proposition 1.29. *Folgende Aussagen sind für den Endomorphismus $F : V \to V$ äquivalent:*

(a) Der Endomorphismus F ist nilpotent (vom Nilpotenzindex m).

(b) Für jeden Vektor v aus V gibt es eine natürliche Zahl m_v derart, dass $F^{m_v}(v) = 0_V$.

(c) *Es gibt eine Basis B von V derart, dass die Darstellungsmatrix von F bezüglich B in strikter oberer Dreiecksform*

$$\begin{pmatrix} 0 & * & \cdots & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

ist.

(d) *Das charakteristische Polynom von F ist $\chi_F(T) = T^n$.*

(e) *Das Minimalpolynom von F ist $m_F(T) = T^m$ für eine natürliche Zahl m, welche dann der Nilpotenzindex ist.*

Beweis. Die Implikation $(a) \Longrightarrow (b)$ ist trivial (setze $m_v = m$).

Wir beweisen nun die Implikation $(b) \Longrightarrow (c)$ induktiv über $n = \dim_{\mathbb{K}} V$. Falls $n = 1$, ist $V = \mathrm{Lin}(\{v\})$ und F^{m_v} ist klarerweise trivial auf dem erzeugten Unterraum $\mathrm{Lin}(\{v\})$.

Wir nehmen also $n \geq 2$ an. Sei $v \neq 0_V$ in V und wähle m_v kleinstmöglich mit $F^{m_v}(v) = 0_V$, also $v_1 = F^{m_v - 1}(v) \neq 0_V$.

Insbesondere ist $U = \mathrm{Lin}(\{v_1\})$ klarerweise F-invariant (weil F trivial auf U ist). Nach dem Hilfslemma 1.15 induziert F einen Endomorphismus $\widetilde{F} : V/U \to V/U$, welcher auch die Eigenschaft (b) erfüllt. Aus dem noetherschen Isomorphiesatz [**Skript LAI**, Korollar 2.46] folgt, dass V/U Dimension $n-1$ besitzt und wir finden induktiv eine Basis $\{v_2 + U, \ldots, v_n + U\}$ vom Quotientenraum V/U derart, dass die Darstellungsmatrix von $\widetilde{F}$ bezüglich dieser Basis in strikter oberer Dreiecksform ist. Wie im Beweis von Satz 1.17 schließen wir, dass die Darstellungsmatrix von F bezüglich der Basis $\{v_1, v_2, \ldots, v_n\}$ in strikter oberer Dreiecksform ist.

Die Implikation $(c) \Longrightarrow (d)$ folgt klarerweise aus der Kästchenregel. Die Implikation $(d) \Longrightarrow (e)$ ist trivial, weil alle Teiler von T^n von der Form $a \cdot T^r$ sind, mit a aus $\mathbb{K}$ und r eine natürliche Zahl. Aus den Eigenschaften des Minimalpolynoms $m_F(T) = T^m$ folgt, dass F^m der triviale Endomorphismus auf V und somit F nilpotent (vom Index m) ist, was die Richtung $(d) \Longrightarrow (a)$ liefert. □

Lemma 1.30. *Sei $v \neq 0_V$ ein Vektor derart, dass $F^k(v) = 0_V$ für eine natürliche Zahl k, welche kleinstmöglich mit dieser Eigenschaft bezüglich v ist. Dann ist $U = \mathrm{Lin}(\{v, F(v), \ldots, F^{k-1}(v)\})$ ein zyklischer F-invarianter Unterraum der Dimension k mit Basis $\{F^{k-1}(v), \ldots, F(v), v\}$. Die Darstellungsmatrix von $F_{\restriction U}$ bezüglich dieser Basis ist*

$$N_k = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & & 0 \end{pmatrix}$$

Beweis. Wir müssen nur zeigen, dass die Familie $\{v, F(v), \ldots, F^{k-1}(v)\}$ linear unabhängig ist. Gegeben Skalare $\lambda_0, \ldots, \lambda_{k-1}$ aus $\mathbb{K}$ mit

$$\sum_{j=0}^{k} \lambda_j F^j(v) = 0_V,$$

schließen wir nach sukzessivem Anwenden von F (von F^{k-1} bis F), dass $\lambda_0 = \ldots = \lambda_{k-1} = 0_{\mathbb{K}}$, weil der Vektor $F^{k-1}(v)$ nicht trivial ist. $\square$

Definition 1.31. Eine Basis $\{v_1, \ldots, v_n\}$ des Vektorraumes V ist *adaptiert* bezüglich des Endomorphismus $F : V \to V$, falls $F(v_1) = 0_V$ und

$$F(v_j) = \begin{cases} v_{j-1} \text{ oder} \\ 0_V \end{cases}$$

für $2 \leq j \leq n$.

Beachte, dass der Endomorphimus F nilpotent sein muss, wenn V eine adaptierte Basis bezüglich F besitzt, weil F^n der triviale Endomorphismus auf V ist. In der Tat ist für den Vektor $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ aus V das Bild

$$F^n\Big(\sum_{j=1}^{n} \lambda_j v_j\Big) = \sum_{j=1}^{n} \lambda_j F^n(v_j) = 0_V,$$

weil die Basis adaptiert ist.

Wir können nun die Rückrichtung des obigen zeigen: Ein nilpotenter Endomorphismus besitzt immer adaptierte Basen.

Satz 1.32. *Wenn V nicht trivial ist und der nilpotente Endomorphismus $F : V \to V$ Nilpotenzindex $m \geq 1$ besitzt, kann jede Basis des Unterraumes U von V, welcher transversal zu* $\mathrm{Ker}(F^{m-1})$ *liegt, (bis auf Permutation) zu einer adaptierten Basis von V ergänzt werden.*

Insbesondere besitzt V adaptierte Basen (bezüglich F).

Beweis. Die letzte Behauptung folgt aus der ersten mit U ein Komplementärraum zum Unterraum $\mathrm{Ker}(F^{m-1})$. Beachte, dass $\mathrm{Ker}(F^{m-1}) \neq V$, weil m der Nilpotenzindex von F ist. Somit ist U nicht trivial und besitzt eine Basis.

Wir beweisen den Satz mit Hilfe mehrerer Zwischenbehauptungen induktiv über den Nilpotenzindex $m \geq 1$. Falls $m = 1$ ist F der triviale Endomorphismus und somit jede Basis von V adaptiert. Wenn $m > 1$, beachte, dass die steigende Kette der Länge m

$$\{0_V\} \subsetneq \mathrm{Ker}(F) \subsetneq \mathrm{Ker}(F^2) \subsetneq \cdots \subsetneq \mathrm{Ker}(F^{m-1}) \subsetneq V$$

echter Unterräume von V bezeugt, dass $n \geq m > 1$, nach der Bemerkung 1.27).

Sei $\{u_1, \ldots, u_s\}$ eine Basis von U und $\{v_1, \ldots, v_k\}$ eine Basis von $V' = \mathrm{Ker}(F^{m-1}) \subsetneq V = \mathrm{Ker}(F^m)$. Da U und V' transversal liegen, ist die Vereinigung

$$\{u_1, \ldots, u_s\} \cup \{v_1, \ldots, v_k\}$$

eine Basis der direkten Summe $U \oplus V'$. Erweitere diese Basis zu einer Basis von V mit der Aufzählung

$$\{u_1, \ldots, u_s, u_{s+1}, \ldots, u_r, v_1, \ldots, v_k\}$$

und betrachte den Unterraum $W = \mathrm{Lin}(\{u_1, \ldots, u_s, u_{s+1}, \ldots, u_r\})$, welcher U enthält. Nach Konstruktion ist W auch transversal zu V'.

Behauptung 1. *Die Menge $\{F(u_1), \ldots, F(u_r)\}$ ist eine Basis des Unterraumes $F(W)$, welcher ein Unterraum von V' ist.*

Beweis der Behauptung 1. Für den ersten Teil müssen wir nur zeigen, dass die Einschränkungsabbildung $F_{\restriction W}$ injektiv ist: Wenn w aus W im Kern von F ist, dann liegt w auch in V', also $w = 0_V$, weil W und V' transversal zueinander sind.

Es ist leicht zu zeigen, dass $F(W)$ ein Unterraum von $V' = \mathrm{Ker}(F^{m-1})$ ist, weil der Nilpotenzindex von F genau m ist. $\square_{\text{Beh 1}}$

Der Unterraum V' ist klarerweise F-invariant. Insbesondere definiert die Einschränkungsabbildung $F_{\restriction V'}$ einen Endomorphismus $F_{\restriction V'} : V' \to V'$, welcher nilpotent vom Index $m-1$ ist. Um den Induktionsschritt anwenden zu können, müssen wir also zeigen, dass $F(W)$ transversal zu $\mathrm{Ker}({F_{\restriction V'}}^{m-2})$ liegt: Sei hierfür v in $F(W) \cap \mathrm{Ker}({F_{\restriction V'}}^{m-2})$, also $v = F(w)$ für ein Element w aus W. Aus $0_V = F^{m-2}(v) = F^{m-1}(w)$ folgt, dass w in $\mathrm{Ker}(F^{m-1}) = V'$ liegt und somit muss $w = 0_V$ und dann auch $v = 0_V$.

Wir können also induktiv die Basis $\{F(u_1), \ldots, F(u_r)\}$ bis auf Permutation zu einer adaptierten Basis

$$\{v'_1, \ldots, v'_k\}$$

von V' bezüglich $F_{\restriction V'}$ ergänzen. Die Elemente aus $\{u_1, \ldots, u_r\}$ können wir so permutieren, dass $F(u_j) = v'_{i_j}$ mit $1 \leq i_1 < \ldots < i_r \leq k$. Um eine Basis von V zu gewinnen, brauchen wir noch r viele neue passende Vektoren, da $r + k = n$. Sei nun $\{v_1, \ldots, v_n\}$ die Folge, welche wir aus der Basis $\{v'_1, \ldots, v'_k\}$ gewinnen, indem wir für $1 \leq j \leq r$ nach jedem Vektor $v'_{i_j} = F(u_j)$ den Vektor u_j auflisten.

Behauptung 2. *Die Familie $\{v'_1, \ldots, v'_k\}$ bildet eine adaptierte Basis von V bezüglich F und ergänzt (bis auf Permutation) die ursprüngliche Basis $\{u_1, \ldots, u_s\}$ von U.*

Beweis der Behauptung 2. Es ist offensichtlich, dass die Familie die Basis $\{u_1, \ldots, u_s\}$ ergänzt. Da $V = W \oplus V'$ die direkte Summe von V' und W ist, ist die Familie

$$\{v_1, \ldots, v_n\} = \{u_1, \ldots, u_r\} \cup \{v'_1, \ldots, v'_k\}$$

eine Basis von V. Wir müssen nur noch zeigen, dass die Basis mit der angegebenen Aufzählung F-adaptiert ist: Der erste Vektor v_1 ist gleich v'_1 und somit ist $F(v_1) = F(v'_1) = 0_V$. Aus der Konstruktion ist die Basis adaptiert an allen Stellen außer möglicherweise an der Stelle direkt nach dem Vektor u_j. An dieser Stelle kommt der alte Vektor v'_{i_j+1} vor. Wenn $F(v'_{i_j+1}) = 0_V$ ist, sind wir fertig. Ansonsten ist $F(v'_{i_j+1}) = v'_{i_j}$, weil die Basis $\{v'_1, \ldots, v'_k\}$ adaptiert bezüglich $F_{\restriction V'}$ ist. Nun ist

$$F(v'_{i_j+1}) = v'_{i_j} = F(u_j),$$

also liegt $u_j - v'_{i_j+1}$ in $\mathrm{Ker}(F) \subset V'$. Daraus folgt, dass der Vektor u_j aus W in V' liegt, was widerspricht, dass W und V' transversal zueinander sind. $\square_{\text{Beh 2}}$

$\square$

Aufgabe. Sei N_k die $k \times k$-Matrix aus dem Lemma 1.30. Zeige, dass die von N_k induzierte lineare Abbildung nilpotent ist. Was ist ihr Nilpotenzindex?

Satz 1.33. *Wenn der nilpotente Endomorphismus $F : V \to V$ eine adaptierte Basis B besitzt, ist die Matrixdarstellung von F bezüglich B von der Form*

$$\begin{pmatrix} N_{k_1} & 0 & \cdots & 0 \\ 0 & N_{k_2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & N_{k_r} \end{pmatrix}$$

für Teilmatrizen (oder Blöcke) $N_{k_1}, \ldots, N_{k_r}$ wie im Lemma 1.30 derart, dass $n = \sum_{j=1}^{r} k_j$. Ferner ist der Nilpotenzindex von F das Maximum der k_j's.

Insbesondere lässt sich jeder nilpotente Endomorphismus bezüglich einer geeigneten Basis als Blockmatrix

$$\begin{pmatrix} \boxed{\begin{smallmatrix} 0 & 1 & \cdots & 0 \\ & 0 & \cdots & 0 \\ & & \ddots & 1 \\ & & & 0 \end{smallmatrix}} & & & \\ & \ddots & & \\ & & \boxed{\begin{smallmatrix} 0 & 1 & \cdots & 0 \\ & 0 & \cdots & 0 \\ & & \ddots & 1 \\ & & & 0 \end{smallmatrix}} & \\ & & & \ddots \\ & & & \ddots \\ & & & & \boxed{\begin{smallmatrix} 0 & 1 & \cdots & 0 \\ & 0 & \cdots & 0 \\ & & \ddots & 1 \\ & & & 0 \end{smallmatrix}} \end{pmatrix}$$

darstellen.

Beweis. Für die F-adaptierte Basis $B = \{v_1, \ldots, v_n\}$ sei $1 \leq i_1 < i_2 < \ldots < i_r \leq n$ eine aufsteigende Aufzählung der Menge

$$\{j \in \{1, \ldots, n-1\} \mid F(v_{j+1}) = 0_V\}.$$

Definiere nun

$$k_j = \begin{cases} i_1, \text{ falls } j = 1 \\ i_j - i_{j-1}, \text{ falls } 1 < j < r \\ n - i_r, \text{ falls } j = r \end{cases}$$

Beachte, dass die Menge $\{v_1, \ldots, v_{i_1}\}$ einen Block der Länge k_1 definiert. Dementsprechend ist $\{v_{i_1+1}, \ldots, v_{i_2}\}$ ein Block der Länge k_2 und so weiter. $\square$

Bemerkung 1.34. Es folgt aus dem Beweis des Satzes 1.32, dass die Anzahl der Blöcke der Größe s genau der Differenz

$$\big(\dim_{\mathbb{K}} \operatorname{Ker}(F^s) - \dim_{\mathbb{K}} \operatorname{Ker}(F^{s-1})\big) - \big(\dim_{\mathbb{K}} \operatorname{Ker}(F^{s+1}) - \dim_{\mathbb{K}} \operatorname{Ker}(F^s)\big)$$

entspricht, weil die Elemente aus $\operatorname{Ker}(F^j) \setminus \operatorname{Ker}(F^{j-1})$ Blöcke der Größe mindestens j bestimmen, nach dem Lemma 1.30.

1.5 Die Jordan'sche Normalform

Wir werden in diesem Abschnitt sehen wie wir die Beschreibung nilpotenter Endomorphismen anwenden können, um eine vollständige Charakterisierung von Endomorphismen $F : V \to V$ eines n-dimensionalen $\mathbb{K}$-Vektorraumes V, deren charakteristisches Polynom in Linearfaktoren (möglicherweise nicht paarweise verschieden) zerfällt, anzugeben. Solche Endomorphismen lassen sich eindeutig als Summe eines diagonalisierbaren Teils und eines nilpotenten Teils schreiben.

Definition 1.35. Für einen Skalar λ aus $\mathbb{K}$ ist der *Hauptraum* zum Wert λ die Menge

$$V(\lambda) = \bigcup_{k \geq 0} \operatorname{Ker}\big((F - \lambda \cdot \mathrm{Id}_V)^k\big),$$

wobei $(F - \lambda \cdot \mathrm{Id}_V)^k = \underbrace{(F - \lambda \cdot \mathrm{Id}_V) \circ \cdots \circ (F - \lambda \cdot \mathrm{Id}_V)}_{k\text{-mal}}$ mit $(F - \lambda \cdot \mathrm{Id}_V)^0 = \mathrm{Id}_V$.

Beachte, dass jeder Hauptraum ein Unterraum ist, da die Kette von Unterräumen

$$\{0_V\} = \operatorname{Ker}\big((F - \lambda \cdot \mathrm{Id}_V)^0\big) \subset \operatorname{Ker}\big((F - \lambda \cdot \mathrm{Id}_V)^1\big) \subset \cdots \subset \operatorname{Ker}\big((F - \lambda \cdot \mathrm{Id}_V)^k\big) \subset \cdots$$

aufsteigend ist. Ferner ist $V(\lambda)$ invariant unter $F - \lambda \cdot \mathrm{Id}_V$ und somit auch $(F - \mu \cdot \mathrm{Id}_V)$-invariant (siehe die Aufgabe nach der Definition 1.14) für alle μ aus $\mathbb{K}$.

Bemerkung 1.36. Analog zu der Bemerkung 1.27 folgt, dass $V(\lambda) = \operatorname{Ker}\big((F - \lambda \cdot \mathrm{Id}_V)^m\big)$ für m kleinstmöglich mit

$$\operatorname{Ker}\big((F - \lambda \cdot \mathrm{Id}_V)^m\big) = \operatorname{Ker}\big((F - \lambda \cdot \mathrm{Id}_V)^{m+1}\big).$$

Ein solches m muss existieren, weil V endlichdimensional ist.

Insbesondere ist der Hauptraum zu λ genau dann trivial, wenn λ kein Eigenwert von F ist. Somit ist der Raum $\sum_{\lambda \in \mathbb{K}} V(\lambda)$ nur eine endliche Summe nicht-trivialer Haupträume.

Lemma 1.37. *Für Skalare $\lambda \neq \mu$ aus $\mathbb{K}$ definiert (die Einschränkung von) $F - \lambda \cdot \mathrm{Id}_V$ einen Automorphismus auf $V(\mu)$.*

Insbesondere ist die Familie der Haupträume $\{V(\lambda)\}_{\lambda \in \mathbb{K}}$ transversal und somit

$$\sum_{\lambda \in \mathbb{K}} V(\lambda) = \bigoplus_{\lambda \in \mathbb{K}} V(\lambda).$$

Beweis. Wir zeigen zuerst, dass $(F - \lambda \cdot \mathrm{Id}_V)_{\restriction V(\mu)}$ einen Automorphismus von $V(\mu)$ definiert. Beachte hierfür zunächst, dass die Einschränkung $(F - \lambda \cdot \mathrm{Id}_V)_{\restriction V(\mu)}$ einen Endomorphismus des Hauptraumes $V(\mu)$ definiert. Nach dem [**Skript LAI**, Rangssatz 2.34] müssen wir nur zeigen, dass die Abbildung $(F - \lambda \cdot \mathrm{Id}_V)_{\restriction V(\mu)}$ injektiv ist. Sei also $v \neq 0_V$ aus $V(\mu)$ und wähle $k \geq 1$ aus $\mathbb{N}$ kleinstmöglich derart, dass v in $\operatorname{Ker}\big((F - \mu \cdot \mathrm{Id}_V)^k\big)$ liegt. Dann ist

$$\begin{aligned}
0_V \neq (F - \mu \cdot \mathrm{Id}_V)^{k-1}(v) &= (\lambda - \mu)^{-1}(F - \lambda \cdot \mathrm{Id}_V)^{k-1}(\lambda v - \mu v) = \\
&= (\lambda - \mu)^{-1}(F - \mu \cdot \mathrm{Id}_V)^{k-1}(\lambda v - F(v) + F(v) - \mu v \\
&= \\
(\lambda - \mu)^{-1}(F - \mu \cdot \mathrm{Id}_V)^{k-1} &\circ (\lambda \cdot \mathrm{Id}_V - F)(v) + (\lambda - \mu)^{-1}(F - \mu \cdot \mathrm{Id}_V)^{k}(v) = \\
= (\lambda - \mu)^{-1}(F - \mu \cdot \mathrm{Id}_V)^{k-1} \circ (\lambda \cdot \mathrm{Id}_V - F)(v) + 0_V &= (\lambda - \mu)^{-1}(F - \mu \cdot \mathrm{Id}_V)^{k-1} \circ (\lambda \cdot \mathrm{Id}_V - F)(v)
\end{aligned}$$

und somit $(F - \lambda \cdot \mathrm{Id}_V)(v) = -(\lambda \cdot \mathrm{Id}_V - F)(v) \neq 0_V$, wie gewünscht. Weil die Komposition von Automorphismen wiederum ein Automorphismus ist, ist die Einschränkung jeder Potenz $(F - \lambda \cdot \mathrm{Id}_V)^k$ auch ein Automorphismus von $V(\mu)$, falls $\mu \neq \lambda$.

Wir zeigen als letztes, dass die Familie $\{V(\lambda)\}_{\lambda \in \mathbb{K}}$ transversal ist und benutzen hierfür die Bemerkung 1.2. Betrachte also paarweise verschiedene Skalare $\lambda_1, \ldots, \lambda_m$ und Vektoren $u_1, \ldots, u_r$ mit u_i aus $V(\lambda_i)$ für $1 \leq i \leq m$ derart, dass

$$u_1 + \ldots + u_m = 0_V.$$

Wir müssen zeigen, dass $u_1 = \cdots = u_m = 0_V$. Induktiv können wir annehmen, dass wir die Behauptung für alle Werte $m' < m$ bereits gezeigt haben. Schreibe nun $u_1 = \sum_{2\le i\le m} -u_i$. Da $-u_j$ im Hauptraum $V(\lambda_j)$ liegt, finde eine Potenz $(F - \lambda_j \cdot \mathrm{Id}_V)^{k_j}$ mit

$$(F - \lambda_j \cdot \mathrm{Id}_V)^{k_j}(-u_j) = 0_V.$$

Weil die Potenzen miteinander kommutieren (siehe die Aufgabe nach der Definition 1.19) folgt, dass

$$(F-\lambda_2\cdot\mathrm{Id}_V)^{k_2}\circ\cdots\circ(F-\lambda_m\cdot\mathrm{Id}_V)^{k_m}(u_1) = (F-\lambda_2\cdot\mathrm{Id}_V)^{k_2}\circ\cdots\circ(F-\lambda_m\cdot\mathrm{Id}_V)^{k_m}\left(\sum_{2\le i\le m} -u_i\right) =$$

$$= (F-\lambda_2\cdot\mathrm{Id}_V)^{k_2}\circ\cdots\circ(F-\lambda_m\cdot\mathrm{Id}_V)^{k_m}(-u_2)+\cdots+(F-\lambda_2\cdot\mathrm{Id}_V)^{k_2}\circ\cdots\circ(F-\lambda_m\cdot\mathrm{Id}_V)^{k_m}(-u_m) =$$

$$= \underbrace{0_V + \cdots + 0_V}_{m-1} = 0_V.$$

Da u_1 aus $V(\lambda_1)$ kommt und $\lambda_1 \neq \lambda_j$ für $2 \le j \le m$, schließen wir $u_1 = 0_V$ und dann

$$u_2 + \cdots + u_m = 0_V.$$

Aus der Induktionsannahme über m folgt jetzt $u_2 = \cdots = u_m = 0_V$. Somit ist die Familie $\{V(\lambda)\}_{\lambda\in\mathbb{K}}$ der Haupträume transversal. □

Proposition 1.38. *Für jeden Skalar λ aus $\mathbb{K}$ entspricht die Dimension des Hauptraumes $V(\lambda)$ der algebraischen Vielfachheit $\mathrm{ord}_\lambda(\chi_F)$ von λ bezüglich F.*

Beweis. Wir müssen die Behauptung nur zeigen falls λ ein Eigenwert von χ_F ist. Schreibe also

$$\chi_F(T) = (T - \lambda)^k \cdot Q(T)$$

mit $k = \mathrm{ord}_\lambda(\chi_F) \ge 1$ und $Q(\lambda) \neq 0_\mathbb{K}$. Der Beweis folgt sofort aus der folgenden Behauptung.

Behauptung. *Es gibt einen F-invarianten Unterraum U von V der Dimension k und eine Basis $B_U = \{u_1, \ldots, u_k\}$ von U derart, dass die Einschränkungsabbildung $F_{\restriction U}$ eine Darstellungsmatrix in oberer Dreiecksform*

$$\begin{pmatrix} \lambda & \cdots & \cdots & \cdots \\ 0 & \lambda & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda \end{pmatrix}$$

bezüglich B_U besitzt.

Wir zeigen zuerst, warum die obige Behauptung den Beweis liefert. Mit der obigen Darstellungsmatrix für $F_{\restriction U}$ folgt, dass $\chi_{F_{\restriction U}}(T) = (T - \lambda)^k$, also $(F - \lambda \cdot \mathrm{Id}_V)^k(u) = 0_V$ für alle u aus U, nach dem Satz von Cayley-Hamilton 1.20. Insbesondere ist U eine Teilmenge des Hauptraumes $V(\lambda)$ und somit $\dim_\mathbb{K} V(\lambda) \ge k$. Wir müssen zeigen, dass $\dim_\mathbb{K} V(\lambda) = k$, oder äquivalent dazu, dass $U = V(\lambda)$. Nach dem Hilfslemma 1.15 induziert F einen Endomorphismus $\widetilde{F} : V/U \to V/U$ mit

$$\chi_{F_{\restriction U}}(T) \cdot \chi_{\widetilde{F}}(T) = \chi_F(T) = (T - \lambda)^k \cdot Q(T).$$

Da Polynomringe Integritätsbereiche sind, folgt $\chi_{\widetilde{F}}(T) = Q(T)$ und somit ist λ kein Eigenwert des Endomorphismus

$$\begin{array}{rccc} \widetilde{F}: & V/U & \to & V/U \\ & v + U & \mapsto & F(v) + U \end{array},$$

so der Hauptraum zu λ von $\widetilde{F}$ ist trivial.
Gegeben nun v in $V(\lambda)$, gibt es eine natürliche Zahl m mit $(F - \lambda \cdot \mathrm{Id}_V)^m(v) = 0_V$. Aus

$$\underbrace{(\widetilde{F} - \lambda \cdot \mathrm{Id}_{V/U}) \circ \cdots \circ (\widetilde{F} - \lambda \cdot \mathrm{Id}_{V/U})}_{m\text{-mal}}(v + U) = (F - \lambda \cdot \mathrm{Id}_V)(v) + U = 0_V + U$$

folgt $v + U = 0_V + U$ und somit ist v in U, wie gewünscht.

Beweis der Behauptung. Sei $u_1 \neq 0_V$ ein Eigenvektor des Eigenwerts λ von F. Falls $k = 1$ sind wir fertig, weil der eindimensionale Unterraum $\mathrm{Lin}(\{u_1\})$ klarerweise F-invariant ist. Wir nehmen also $k > 1$ an. Setze $U_1 = \mathrm{Lin}(\{u_1\})$ und schreibe

$$\chi_F(T) = \chi_{F_{\restriction U_1}}(T) \cdot \chi_{\widetilde{F}}(T),$$

mit

$$\begin{array}{rccc} \widetilde{F}: & V/U_1 & \to & V/U_1 \\ & v + U_1 & \mapsto & F(v) + U_1 \end{array}$$

Insbesondere ist $\chi_{\widetilde{F}}(T) = (T - \lambda)^{k-1} \cdot Q(T)$ und somit ist die geometrische Vielfachheit $\mathrm{ord}_\lambda(\chi_{\widetilde{F}}) = k - 1$. Wir finden also induktiv über k einen $\widetilde{F}$-invarianten Unterraum $\widetilde{U}$ von V/U_1 der Dimension $k-1$ zusammen mit einer Basis $B_{\widetilde{U}} = \{u_2 + U_1, \ldots, u_k + U_1\}$ derart, dass die Darstellungsmatrix (a_{ij}) von $\widetilde{F}_{\restriction \widetilde{U}}$ die gesuchte Form hat. Für $2 \leq j \leq k$ ist also

$$F(v_j) + U_1 = \widetilde{F}(u_j) = \lambda(u_j + U_1) + \sum_{2 \leq i \leq j} a_{ij}(u_i + U_1).$$

Dies bedeutet, dass es einen Skalar a_{1j} in $\mathbb{K}$ gibt mit

$$F(v_j) = \lambda u_j + \sum_{1 \leq i \leq j} a_{ij} u_i.$$

Analog zum Ende des Beweis von Satz 1.17 lässt sich zeigen, dass $\{u_1, u_2, \ldots, u_k\}$ linear unabhängig ist und somit der klarerweise F-invariante Unterraum $U = \mathrm{Lin}(\{u_1, \ldots, u_k\})$ die Dimension k hat. Die Darstellungsmatrix von $F_{\restriction U}$ ist in der gewünschten Form, nach der Wahl der a_{1j}'s. $\square_{\text{Beh.}}$

$\square$

Satz 1.39. (*Jordansche Normalform & Zerlegung von Jordan-Chevalley*)

Falls das charakteristische Polynom des Endomorphismus $F : V \to V$ in Linearfaktoren (möglicherweise mit Wiederholungen) zerfällt, ist V die direkte Summe $\bigoplus_{i=1}^k V(\lambda_i)$ der nicht-trivialen Haupträume der verschiedenen Eigenwerte $\lambda_1, \ldots, \lambda_k$ von F.

Insbesondere lässt sich F durch eine geeignete Basis als Blockmatrix

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & A_k \end{pmatrix}$$

darstellen, wobei jede Teilmatrix A_i von folgender Form ist

$$A_i = \begin{pmatrix} \boxed{\begin{matrix} \lambda_i & 1 & \cdots & 0 \\ & \lambda_i & \cdots & 0 \\ & & \ddots & 1 \\ & & & \lambda_i \end{matrix}} & & & \\ & \ddots & & \\ & & \boxed{\begin{matrix} \lambda_i & 1 & \cdots & 0 \\ & \lambda_i & \cdots & 0 \\ & & \ddots & 1 \\ & & & \lambda_i \end{matrix}} & \\ & & & \ddots \\ & & & & \ddots \\ & & & & & \boxed{\begin{matrix} \lambda_i & 1 & \cdots & 0 \\ & \lambda_i & \cdots & 0 \\ & & \ddots & 1 \\ & & & \lambda_i \end{matrix}} \end{pmatrix}$$

Ferner lässt sich F eindeutig als $F = G + H$ schreiben für einen diagonalisierbaren Endomorphismus G (mit Eigenwerten $\lambda_1, \ldots, \lambda_k$) und einen nilpotenten Endomorphismus H derart, dass $G \circ H = H \circ G$.

Der größte λ_i-Block in der obigen Darstellungsmatrix von F hat die Größe $\mathrm{ord}_{\lambda_i}(m_F)$, wobei $m_F(T)$ das Minimalpolynom von F ist. Es gibt genau

$$\begin{aligned} &\big(\dim_{\mathbb{K}} \mathrm{Ker}((F - \lambda_i \cdot \mathrm{Id}_V)^s) - \dim_{\mathbb{K}} \mathrm{Ker}((F - \lambda_i \cdot \mathrm{Id}_V)^{s-1})\big) - \\ &\qquad\qquad \big(\dim_{\mathbb{K}} \mathrm{Ker}((F - \lambda_i \cdot \mathrm{Id}_V)^{s+1}) - \dim_{\mathbb{K}} \mathrm{Ker}((F - \lambda_i \cdot \mathrm{Id}_V)^s)\big) \end{aligned}$$

viele λ_i-Blöcke der Größe s.

Beweis. Falls das charakteristische Polynom χ_F in Linearfaktoren zerfällt, schreibe

$$\chi_F(T) = (T - \lambda_1)^{r_1} \cdots (T - \lambda_k)^{r_k},$$

wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von F sind. Insbesondere ist $n = \deg \chi_F = r_1 + \cdots + r_k$. Nach der Proposition 1.38 hat für $1 \leq i \leq k$ der Hauptraum $V(\lambda_i)$ genau Dimension r_i. Weil die Familie $\{V(\lambda_i)\}_{1 \leq i \leq k}$ transversal ist (nach dem Lemma 1.37), hat der Unterraum $\bigoplus_{i=1}^k V(\lambda_i)$ die Dimension $r_1 + \cdots + r_k = n$, also ist $V = \bigoplus_{i=1}^k V(\lambda_i)$.

Jeder Endomorphismus wird also durch die Einschränkungen auf seine Haupträume $V(\lambda_i)$ eindeutig bestimmt. Es gibt genau einen Endomorphismus $G : V \to V$ derart, dass $G_{\restriction V(\lambda_i)} = \lambda_i \cdot \mathrm{Id}_{V(\lambda_i)}$. Klarerweise ist G diagonalisierbar und $H = F - G$ ist nilpotent, weil $H_{\restriction V(\lambda_i)}$ nilpotent für jedes $1 \leq i \leq k$ ist. Mit Hilfe des Satzes 1.32 finden wir für jeden $V(\lambda_i)$ eine adaptierte Basis bezüglich H. Die Vereinigung dieser Basen liefert eine bezüglich H adaptierte Basis von V. Beachte, dass die Elemente dieser Basis Eigenwerte von G sind und wir somit die gewünschte Darstellungsmatrix haben. Weil die Einschränkung $G_{\restriction V(\lambda_i)}$ Skalarmultiplikation mit λ_i ist, kommutiert sie mit $H_{\restriction V(\lambda_i)}$. Wegen der Linearität kommutieren somit auch G und H miteinander.

Wir müssen also nur noch zeigen, dass die Zerlegung $F = G + H$ eindeutig ist. Hierfür nehmen wir an, dass $F = G' + H'$ für zwei miteinander kommutierende Endomorphismen G' und H', wobei G' diagonalisierbar und H' nilpotent ist. Es genügt zu zeigen, dass $G = G'$ eindeutig bestimmt ist.

Nun lässt sich V als direkte Summe der Eigenräume $V_1, \ldots, V_s$ zu den Eigenwerten $\mu_1, \ldots, \mu_s$ von G' schreiben. Beachte, dass jeder Unterraum V_i invariant unter F ist: Gegeben v aus V_i, ist $G'(v) = \mu_i v$, also

$$G'(H'(v)) = G' \circ H'(v) = H' \circ G'(v) = \mu_i H'(v).$$

Das bedeutet, dass $H'(v)$ in V_i liegt und somit V_i invariant unter $F = G' + H'$ ist.

Sei $V(\mu_j)$ der zum Wert μ_j gehörigen Hauptraum von F. Da $G'_{\restriction V_j}$ Skalarmultiplikation mit μ_j ist, wirken die Einschränkungen von $F - \mu_j \cdot \mathrm{Id}_V$ und $F - G'$ identisch auf V_j (welcher invariant unter F ist!), also $V_j \subset V(\mu_j)$. Insbesondere muss jeder Eigenwert μ_j einer der Eigenwerte λ_i sein. Bis auf Permutation folgt, dass $V_i \subset V(\lambda_i)$ für $1 \leq i \leq s$. Aus

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_k) = V_1 \oplus \cdots \oplus V_s,$$

folgt $s = k$ und G' ist einfach Skalarmultiplikation mit λ_i auf $V(\lambda_i)$, genauso wie unsere Abbildung G und somit ist $G = G'$, wie gewünscht. $\square$

Kapitel 2

Dualität und Paarungen

Im Folgenden sei $\mathbb{K}$ ein Körper.

2.1 Dualräume

Definition 2.1. Der *Dualraum* eines $\mathbb{K}$-Vektorraumes V ist der $\mathbb{K}$-Vektorraum $V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ aller linearen Abbildungen $F : V \to \mathbb{K}$ mit der Summe

$$\begin{array}{ccccc} V^* \times V^* & \to & V^* & & \\ (F,G) & \mapsto & F+G: & V & \to & \mathbb{K} \\ & & & v & \mapsto & F(v)+G(v) \end{array}$$

und der Skalarmultiplikation

$$\begin{array}{cccccc} \mathbb{K} \times V^* & \to & V^* & & & \\ (\lambda, F) & \mapsto & \lambda F: & V & \to & \mathbb{K} \\ & & & v & \mapsto & \lambda \cdot F(v) \end{array}$$

Elemente aus V^* heißen auch *Funktionale* oder *Linearformen.*

Wenn V endlichdimensional ist, sind V und V^* isomorph, aber der Isomorphismus ist nicht kanonisch, er hängt von der Wahl einer Basis von V ab.

Definition 2.2. Sei $B = (b_i)_{i \in I}$ eine Basis des $\mathbb{K}$-Vektorraumes V. Die *Dualbasis* $B^* = (b_i^*)_{i \in I}$ zu B ist die Kollektion der linearen Abbildungen $b_i^* : V \to \mathbb{K}$ mit

$$b_i^*(b_j) = \delta_{ij} = \begin{cases} 1_{\mathbb{K}}, & \text{falls } i = j \\ 0_{\mathbb{K}}, & \text{sonst.} \end{cases}$$

Bemerkung 2.3. Das Symbol δ_{ij} heißt Kronecker-Delta.

Beachte, dass die lineare Abbildung b_i^* durch die Wirkung auf den Basisvektoren eindeutig bestimmt ist: Für v aus V beliebig ist $b_i^*(v)$ die Koordinate von v bezüglich des Basisvektors b_i.

Lemma 2.4. *Die Dualbasis $B^* = (b_i^*)_{i \in I}$ zu der Basis $B = (b_i)_{i \in I}$ bildet eine linear unabhängige Familie des Dualraums V^*.*

Wenn V endlichdimensional ist, so ist V^* endlichdimensional und B^* ist eine Basis von V^*. Insbesondere sind V und V^* isomorph.

Beweis. Wir zeigen zuerst, dass die Familie $B^* = (b_i^*)_{i\in I}$ linear unabhängig ist. Seien also $i_1, \dots, i_n$ aus I und $\lambda_1, \dots, \lambda_n$ Skalare derart, dass das Funktional $\sum_{1\le k\le n} \lambda_k b_{i_k}^* : V \to \mathbb{K}$ die Nullabbildung ist.

Für $1 \le j \le n$ ist $0_{\mathbb{K}} = (\sum_{1\le k\le n} \lambda_k b_{i_k}^*)(b_{i_j}) = \lambda_j$ und somit B^* linear unabhängig, wie gewünscht.

Wir nehmen nun an, dass V endlichdimensional ist. Sei $\{b_1, \dots, b_n\}$ eine Basis von V und $\{b_1^*, \dots, b_n^*\}$ die entsprechende Dualbasis auf V^*. Wir müssen nur noch zeigen, dass sich jedes Funktional $F : V \to \mathbb{K}$ als Linearkombination der b_i^* schreiben lässt. Setze $\lambda_i = F(b_i)$ für $1 \le i \le n$. Das Funktional

$$F - \sum_{1\le i\le n} \lambda_i b_i^*$$

ist klarerweise Null auf jedem Basisvektor und somit die Nullabbildung. Es folgt, dass V^* dieselbe Dimension wie V hat und somit sind diese beiden Vektorräume (nicht unbedingt kanonisch) isomorph. □

Lemma 2.5. *Seien B und B' zwei Basen des endlichdimensionalen $\mathbb{K}$-Vektorraumes V und $A = (a_{ij})$ die Übergangsmatrix von B' nach B. Für die entsprechenden Dualbasen B^* und $(B')^*$ des Dualraumes V^* ist $({}^tA)^{-1}$ die Übergangsmatrix von $(B')^*$ nach B^*.*

Beweis. Schreibe $B = \{b_1, \dots, b_n\}$ sowie $B' = \{b'_1, \dots, b'_n\}$. Wir wollen die Koordinaten des Funktionales $(b'_j)^*$ bezüglich $\{b_1^*, \dots, b_n^*\}$ bestimmen:

$$(b'_j)^* = c_{1j} b_1^* + \cdots + c_{nj} b_n^*,$$

oder in symbolischer Notation

$$\begin{pmatrix} (b'_1)^* \\ \vdots \\ (b'_n)^* \end{pmatrix} = C \cdot \begin{pmatrix} b_1^* \\ \vdots \\ b_n^* \end{pmatrix}$$

mit $C = (c_{ij})$. Nun ist $b'_k = a_{1k} b_i + \cdots + a_{nk} b_n$ für $1 \le k \le n$, also

$$\begin{aligned}(b'_j)^*(b'_k) = \delta_{jk} = \left(\sum_{1\le i\le n} c_{ij} b_i^*\right)\left(\sum_{1\le t\le n} a_{tk} b_t\right) &= \sum_{1\le i\le n}\sum_{1\le t\le n} c_{ij} a_{tk} b_i^*(b_t) = \\ &= \sum_{1\le i\le n}\sum_{1\le t\le n} c_{ij} a_{tk} \delta_{it} = \sum_{1\le i\le n} c_{ij} a_{ik} = \sum_{1\le i\le n} d_{ki} c_{ij},\end{aligned}$$

wobei ${}^tA = (d_{r\ell})$ mit $d_{r\ell} = a_{\ell r}$. Wenn wir alle Werte für j und k einsetzen bedeutet es, dass

$$\mathbf{Id}_n = D \cdot C$$

und somit $C = D^{-1} = ({}^tA)^{-1}$, wie gewünscht. □

Trotz der fehlenden Kanonizität des Isomorphismus zwischen V und seinem Dualraum lässt sich V kanonisch als Unterraum in den zu V^* dualen Raum V^{**} einbetten. Wir werden im nächsten Abschnitt sehen, dass dies allgemein aus der Existenz einer *dualen Paarung* folgt.

Lemma 2.6. *Für jeden $\mathbb{K}$-Vektorraum V existiert ein kanonischer Monomorphismus*

$$\begin{aligned} \varphi : \ V &\hookrightarrow V^{**} \\ v &\mapsto \varphi_v : \ V^* \to \mathbb{K} \\ &\qquad\qquad F \mapsto F(v) \end{aligned}$$

Beweis. Beachte, dass φ wohldefiniert ist, weil φ_v für jeden Vektor v aus V linear ist:

$$\varphi_v(\lambda F + \mu G) = \lambda F(v) + \mu G(v) = \lambda \varphi_v(F) + \mu \varphi_v(G).$$

Wir zeigen zuerst, dass φ linear ist: Seien also λ und μ aus $\mathbb{K}$ sowie v und w aus V. Für F aus V^* ist

$$\varphi(\lambda v + \mu w)(F) = F(\lambda v + \mu w) = \lambda F(v) + \mu F(w) = (\lambda \varphi_v + \mu \varphi_w)(F) = (\lambda \varphi(v) + \mu \varphi(w))\,(F).$$

Für die Injektivität von φ müssen wir also nur zeigen, dass $\varphi_v : V^* \to \mathbb{K}$ nicht trivial ist, wenn $v \neq 0_V$. Der Unterraum $U = \mathrm{Lin}(\{v\})$ besitzt einen Komplementärraum W in V, also $V = U \oplus W$ mit $W \neq V$. Also gibt es eine lineare Abbildung $F : V \to \mathbb{K}$ derart, dass $F_{\restriction W}$ die Nullabbildung ist und $F(\lambda v) = \lambda$ für λv aus U. Wegen $\varphi_v(F) = 1_{\mathbb{K}} \neq 0_{\mathbb{K}}$ ist φ_v ist nicht trivial, wie gewünscht. $\square$

Korollar 2.7. *Jeder endlichdimensionale $\mathbb{K}$-Vektorraum V ist isomorph zum Bidualraum V^{**}.*

Aufgabe. Falls der $\mathbb{K}$-Vektorraum V die direkte Summe der Unterräume U und W ist, so ist V^* kanonisch isomorph zu $U^* \oplus W^*$. Wie wird der Isomorphismus definiert?

Definition 2.8. Für eine lineare Abbildung $F : V \to W$ ist

$$\begin{array}{rccccl} F^*: & W^* & \to & V^* & & \\ & \psi & \mapsto & \psi \circ F: & V & \to \mathbb{K} \\ & & & & v & \mapsto \psi(F(v)) \end{array}$$

die *duale* Abbildung.

Bemerkung 2.9. Es lässt sich leicht zeigen, dass F^* wiederum linear ist. Ferner ist

$$(F+G)^* = F^* + G^*,$$

sowie

$$(\lambda F)^* = \lambda F^*.$$

Die zur Identitätsabbildung Id_V duale Abbildung ist die Identitätsabbildung Id_{V^*} auf V^*. Ferner ist

$$(F \circ G)^* = G^* \circ F^*$$

für kompatible lineare Abbildungen

$$U \xrightarrow{G} V \xrightarrow{F} W.$$

In der Sprache der Kategorien bedeutet dies, dass die Korrespondenz

$$\begin{array}{ccc} \{\mathbb{K} - \text{Vektorräume}\} & \to & \{\mathbb{K} - \text{Vektorräume}\} \\ V & \to & V^* \end{array}$$

einen *kontravarianten Funktor* definiert.

Aufgabe. Zeige mit Hilfe des Lemmas 2.6, dass F trivial sein muss, falls die duale Abbildung F^* trivial ist.

Schließe daraus, dass gegeben endlichdimensionale $\mathbb{K}$-Vektorräume V und W sowie eine lineare Abbildung $G : W^* \to V^*$ eine lineare Abbildung $F : V \to W$ mit $G = F^*$ existiert.
Hinweis: Warum ist die Dimension des $\mathbb{K}$-Vektorraumes $\mathrm{Hom}_{\mathbb{K}}(V, W)$ genau $\dim_{\mathbb{K}} V \dim_{\mathbb{K}} W$?

Lemma 2.10. *Für lineare Abbildungen $F : V \to W$ zwischen $\mathbb{K}$-Vektorräumen gelten die folgenden Äquivalenzen:*

(a) Die Abbildung F ist genau dann injektiv, wenn die duale Abbilldung $F^* : W^* \to V^*$ surjektiv ist.

(b) Die Abbildung F ist genau dann surjektiv, wenn die duale Abbilldung $F^* : W^* \to V^*$ injektiv ist.

Insbesondere ist der Homomorphismus F genau dann ein Isomorphismus, wenn F^* einen Isomorphismus von W^* nach V^* definiert.

Beweis. Für (a) nehmen wir zuerst an, dass die lineare Abbildung F injektiv ist. Um zu zeigen, dass $F^* : W^* \to V^*$ ein Epimorphismus ist, sei ψ ein Funktional aus V^*. Wir suchen ein Funktional θ aus W^* derart, dass $F^*(\theta) = \psi$ (als Abbildungen), oder äquivalent dazu,

$$\theta\left(F(v)\right) = \psi(v)$$

für alle v aus V. Das heißt, dass die gesuchte Abbildung θ auf $\operatorname{Im}(F)$ eindeutig bestimmt ist. Um θ auf dem gesamten Vektorraum W zu definieren, wähle einen zu $\operatorname{Im}(F)$ komplementären Raum Z in W, also $W = \operatorname{Im}(F) \oplus Z$.

Da F injektiv ist, lässt sich jedes Element w aus $\operatorname{Im}(F)$ für einen eindeutig bestimmten Vektor v aus V als $F(v)$ schreiben. Wir definieren also

$$\begin{array}{rcll} \theta : & W & \to & \mathbb{K} \\ & w & \mapsto & \psi(v), \text{wobei } v \text{ der einzige Vektor aus } V \text{ so ist, dass} \\ & & & \qquad\qquad F(v) - w \text{ in } Z \text{ liegt.} \end{array}$$

Aus der obigen Diskussion folgt, dass die Abbildung θ wohldefiniert ist, weil $W = \operatorname{Im}(F) \oplus Z$ und es somit für jeden Vektor w aus W ein eindeutig bestimmtes Element w' aus $\operatorname{Im}(F)$ gibt mit $w' - w$ in Z.

Wir müssen nur zeigen, dass die obige Abbildung θ linear ist: Gegeben Skalare λ_1 und λ_2 aus $\mathbb{K}$ sowie Vektoren w_1 und w_2 aus W, finde eindeutig bestimmte Vektoren v_1 und v_2 aus V derart, dass $w_i - F(v_i)$ in Z liegt. Dann ist

$$\lambda_1 w_1 + \lambda_2 w_2 - F(\lambda_1 v_1 + \lambda_2 v_2) \text{ in } Z,$$

wegen der Linearität von F und somit

$$\theta(\lambda_1 w_1 + \lambda_2 w_2) = \psi(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \psi(v_1) + \lambda_2 \psi(v_2) = \lambda_1 \theta(w_1) + \lambda_2 \theta(w_2),$$

wie gewünscht.

Für die andere Richtung der Äquivalenz, nehmen wir nun an, dass $F^* : W^* \to V^*$ surjektiv ist. Um zu zeigen, dass F injektiv ist, sei v ein Element von $\operatorname{Ker}(F)$. Wenn v nicht der Nullvektor wäre, fänden wir leicht ein Funktional $\psi : V \to \mathbb{K}$ mit $\psi(v) = 1_{\mathbb{K}}$. Wegen der Surjektivität von F^* gäbe es ein Funktional θ aus W^* mit

$$\theta\left(F(v)\right) = \psi(v) \neq 0_V,$$

was nicht möglich ist, weil $F(v) = 0_W$.

Für (*b*) nehmen wir zuerst an, dass F surjektiv ist. Sei θ ein Funktional in $\operatorname{Ker}(F^*)$, d.h. $\theta \circ F : V \to \mathbb{K}$ ist die Nullabbildung. Wir wollen zeigen, dass $\theta(w) = 0_{\mathbb{K}}$ für alle w aus W. Jeder Vektor w aus W lässt sich schreiben als $w = F(v)$ für einen Vektor v aus V (es wird nicht behauptet, dass v eindeutig bestimmt ist). Nun gilt

$$\theta(w) = \theta\left(F(v)\right) = 0_{\mathbb{K}},$$

also ist θ die Nullabbildung, wie gewünscht.

Für die andere Richtung nehmen wir als letztes an, dass F^* injektiv ist. Falls $\operatorname{Im}(F) \neq W$, wähle einen Komplementärraum Z zu $\operatorname{Im}(F)$ in W. Da Z nicht trivial ist, besitzt Z eine Basis C. Wir können nun ein Funktional $G : W \to \mathbb{K}$ so definieren, dass die Einschränkung von G auf $\operatorname{Im}(F)$ trivial ist und $G(w) = 1_{\mathbb{K}}$ für jeden Basisvektor w aus C gilt. Beachte, dass $F^*(G) = G \circ F : V \to \mathbb{K}$ die Nullabbildung ist, nach unserer Konstruktion von G, und somit G trivial sein muss, weil die Abbildung F^* injektiv ist. Es folgt, dass C die leere Menge ist, also $Z = \{0_W\}$, wie gewünscht. □

Lemma 2.11. *Sei $F : V \to W$ eine lineare Abbildung zwischen den endlichdimensionalen $\mathbb{K}$-Vektorräumen V und W mit Darstellungsmatrix A bezüglich der Basen $\{v_1, \ldots, v_n\}$ von V und $\{w_1, \ldots, w_m\}$ von W. Die duale Abbildung $F^* : W^* \to V^*$ hat die Darstellungsmatrix ${}^tA$ bezüglich der dualen Basen $\{w_1^*, \ldots, w_m^*\}$ von W^* und $\{v_1^*, \ldots, v_n^*\}$ von V^*.*

Insbesondere ist der Rang von F^* gleich $\operatorname{Rg}(F)$ *und die Determinante* $\det(F^*) = \det(F)$.

Beweis. Für $A = (a_{ij})$ ist

$$F(v_j) = \sum_{1 \le i \le m} a_{ij} w_i \text{ für } 1 \le j \le n.$$

Die Darstellungsmatrix $X = (\lambda_{ij})$ von F^* bezüglich der Basen $\{w_1^*, \ldots, w_m^*\}$ von W^* und $\{v_1^*, \ldots, v_n^*\}$ von V^* wird durch die Identitäten

$$F^*(w_i^*) = \sum_{1 \le k \le n} \lambda_{ki} v_k^*, \text{ für } 1 \le i \le m,$$

bestimmt. Nun ist für $1 \le j \le n$

$$F^*(w_i^*)(v_j) = w_i^* \circ F(v_j) = w_i^* \left(\sum_{1 \le r \le m} a_{rj} w_r \right) = \sum_{1 \le r \le m} a_{rj} w_i^*(w_r) = a_{ij}.$$

Andererseits gilt

$$a_{ij} = F^*(w_i^*)(v_j) = \left(\sum_{1 \le k \le n} \lambda_{ki} v_k^* \right)(v_j) = \sum_{1 \le k \le n} \lambda_{ki} v_k^*(v_j) = \lambda_{ji}$$

und somit ist die Matrix X die transponierte ${}^tA$ von A, wie gewünscht.

Die letzten zwei Behauptungen folgen sofort aus [**Skript LAI**, Satz 2.52 & Korollar 3.15]. □

2.2 Duale Paarungen

Im Folgenden seien V und W Vektorräume über dem Körper $\mathbb{K}$.

Definition 2.12. Eine Abbildung $\varphi : V \times W \to \mathbb{K}$ ist *bilinear*, falls sie linear in jeder Koordinate ist, das heißt,

$$\varphi(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 \varphi(v_1, w) + \lambda_2 \varphi(v_2, w)$$

und

$$\varphi(v, \lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 \varphi(v, w_1) + \lambda_2 \varphi(v, w_2)$$

für alle Skalare λ_1 und λ_2 aus $\mathbb{K}$, Vektoren v, v_1 und v_2 aus V, sowie w, w_1 und w_2 aus W.

Beispiel 2.13. Die Abbildung $\begin{array}{rccc} \varphi : & V \times V^* & \to & \mathbb{K} \\ & (v, F) & \mapsto & F(v) \end{array}$ ist klarerweise bilinear.

Bemerkung 2.14. Gegeben eine bilineare Abbildung $\varphi : V \times W \to \mathbb{K}$ und einen Vektor v aus V, ist die Abbildung

$$\begin{array}{rccc} \varphi(v, -) : & W & \to & \mathbb{K} \\ & w & \mapsto & \varphi(v, w) \end{array}$$

ein Funktional aus W^*, weil φ linear in der zweiten Koordinate ist.

Proposition 2.15. *Die Abbildung*

$$\begin{array}{rccc} \Phi : & \{\varphi : V \times W \to \mathbb{K} \mid \varphi \text{ bilinear}\} & \to & \{F : V \to W^* \text{ linear}\} \\ & \varphi & \mapsto & \begin{array}{rccc} F_\varphi : & V & \to & W^* \\ & v & \mapsto & \varphi(v, -) \end{array} \end{array}$$

ist eine Bijektion mit inverser Abbildung

$$\begin{array}{rccc} \Phi^{-1} : & \{F : V \to W^* \text{ linear}\} & \to & \{\varphi : V \times W \to \mathbb{K} \mid \varphi \text{ bilinear}\} \\ & F & \mapsto & \begin{array}{rccc} \varphi_F : & V \times W & \to & \mathbb{K} \\ & (v, w) & \mapsto & F(v)(w) \end{array} \end{array}$$

Beweis. Beachte zuerst, dass ϕ wohldefiniert ist, da $F_\varphi : V \to W^*$ linear ist, weil φ in der ersten Koordinate linear ist. Dementsprechend ist Φ^{-1} auch wohldefiniert, weil φ_F bilinear ist:

$$\begin{aligned} \varphi_F(\lambda_1 v_1 + \lambda_2 v_2, w) = F(\lambda_1 v_1 + \lambda_2 v_2)(w) = (\lambda_1 F(v_1) + \lambda_2 F(v_2))\,(w) = \\ = \lambda_1 F(v_1)(w) + \lambda_2 F(v_2)(w) = \lambda_1 \varphi_F(v_1, w) + \lambda_2 \varphi_F(v_2, w) \end{aligned}$$

und

$$\begin{aligned} \varphi_F(v, \lambda_1 w_1 + \lambda_2 w_2) = F(v)(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 F(v)(w_1) + \lambda_2 F(v)(w_2) = \\ = \lambda_1 \varphi_F(v, w_1) + \lambda_2 \varphi_F(v, w_2). \end{aligned}$$

Es lässt sich sofort zeigen, dass Φ und Φ^{-1} inverse zueinander sind, denn $F_\varphi(v)(w) = \varphi(v, w)$ und $\varphi_F(v, -) = F$. □

Definition 2.16. Ein *duales Paar* besteht aus zwei endlichdimensionalen Vektorräumen V und W zusammen mit einer bilinearen Abbildung $\varphi : V \times W \to \mathbb{K}$ derart, dass die entsprechende Abbildung $F_\varphi = \Phi(\varphi) : V \to W^*$ ein Isomorphismus ist.

Insbesondere ist $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W^* = \dim_{\mathbb{K}} W$, nach dem Lemma 2.4.

Aus dem Korollar 2.7 folgt, dass die Vektorräume V und V^* zusammen mit der bilinearen Abbildung φ aus dem Beispiel 2.13 ein duales Paar bilden, wenn V endlichdimensional ist.

Wir werden nun eine Charakterisierung dualer Paare angeben, welche direkt aus der bilinearen Abbildung gelesen werden kann. Hierfür müssen wir den *Rang* einer bilinearen Abbildung einführen. Im Folgenden sind die Vektorräume V und W endlichdimensional.

Lemma 2.17. *Gegeben Basen $B = \{v_1, \ldots, v_n\}$ von V und $C = \{w_1, \ldots, w_m\}$ von W, sei $a_{ij} = \varphi(v_i, w_j)$ für $1 \leq i \leq n$ und $1 \leq j \leq m$.*

Die Abbildung φ wird durch die Matrix $A = (a_{ij})$ eindeutig bestimmt. Falls B' eine andere Basis von V und C' eine andere Basis von W ist, dann gilt für die entsprechende Matrix A', dass

$$A' = {}^t T \cdot A \cdot S,$$

wobei T die Übergangsmatrix von B' nach B und S die Übergangsmatrix von C' nach C ist.

Insbesondere hängt der Rang der Matrix A nicht von der Wahl der Basen B und C ab, denn ${}^t(U^{-1}) = ({}^tU)^{-1}$ gilt für jede reguläre Matrix U. Wir bezeichnen diesen Wert als der *der Rang* $\mathrm{Rg}(\varphi)$ *der Bilinearform* φ. Die Matrix A ist die *Darstellungsmatrix* der Bilinearform φ bezüglich der Basen B von V und C von W.

Beweis. Gegeben v aus V und w aus W, schreibe $v = \sum_{1\leq i\leq n} \lambda_i v_i$ sowie $w = \sum_{1\leq j\leq m} \mu_j w_j$. Aus der Bilinearität folgt

$$\varphi(v, w) = \varphi\left(\sum_{1\leq i\leq n} \lambda_i v_i, \sum_{1\leq j\leq m} \mu_j w_j\right) = \sum_{1\leq i\leq n}\sum_{1\leq j\leq m} \lambda_i \mu_j \varphi(v_i, w_j) = (\lambda_1, \ldots, \lambda_n) \cdot A \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix}.$$

Insbesondere wird φ durch A eindeutig bestimmt.

Wir nehmen nun an, dass wir neue Basen $B' = \{v'_1, \ldots, v'_n\}$ von V und $C' = \{w'_1, \ldots, w'_m\}$ von W bekommen haben. Der Vektor v hat die Koordinaten $(\lambda'_1, \ldots, \lambda'_n)$ bezüglich B', wobei

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = T \cdot \begin{pmatrix} \lambda'_1 \\ \vdots \\ \lambda'_n \end{pmatrix}.$$

Dementsprechend sind $(\mu'_1, \ldots, \mu'_m)$ die Koordinaten von w bezüglich C' mit

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = S \cdot \begin{pmatrix} \mu'_1 \\ \vdots \\ \mu'_m \end{pmatrix}.$$

Nun ist

$$\varphi(v, w) = (\lambda_1, \ldots, \lambda_n) \cdot A \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = (\lambda'_1, \ldots, \lambda'_n) \cdot {}^tT \cdot A \cdot S \cdot \begin{pmatrix} \mu'_1 \\ \vdots \\ \mu'_m \end{pmatrix}$$

und somit ist ${}^tT \cdot A \cdot S$ die Matrix von φ bezüglich der Basen B' und C'. □

Korollar 2.18. *Gegeben eine Bilinearform $\varphi : V \times W \to \mathbb{K}$ mit Darstellungsmatrix A bezüglich der Basen B von V und C von W, sei C^* die zu C duale Basis aus W^*. Die Darstellungsmatrix der Abbildung F_φ (wie in der Proposition 2.15) bezüglich der Basen B von V und C^* von W^* ist ${}^tA$.*

Insbesondere ist (V, W, φ) genau dann ein duales Paar, wenn $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W = \mathrm{Rg}(\varphi)$.

Beweis. Um die Darstellungsmatrix von F_φ zu bestimmen, müssen wir die Koordinaten der Bilder $F_\varphi(v_j)$ bezüglich der zu C dualen Basis $C^* = \{w_1^*, \ldots, w_m^*\}$ bestimmen, also

$$F_\varphi(v_j) = \lambda_{1j} w_1^* + \cdots + \lambda_{mj} w_m^*,$$

wobei $F_\varphi(v_j) = \varphi(v_j, -) : W \to \mathbb{K}$. Nun ist

$$a_{ji} = \varphi(v_j, w_i) = F_\varphi(v_j)(w_i) = (\lambda_{1j} w_1^* + \cdots + \lambda_{mj} w_m^*)\,(w_i) = \lambda_{ij}$$

und somit ist ${}^tA$ die Darstellungsmatrix von F_φ bezüglich B und C^*, wie gewünscht.

Beachte, dass F_φ genau dann ein Isomorphismus ist, wenn $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W^* = \mathrm{Rg}(F_\varphi) = \mathrm{Rg}({}^tA)$. Da $\dim_{\mathbb{K}} W = \dim_{\mathbb{K}} W^*$ und $\mathrm{Rg}({}^tA) = \mathrm{Rg}(A)$, folgt die äquivalente Charakterisierung dualer Paare. $\square$

Aufgabe. Es folgt aus dem obigen Korollar, dass (V, W, φ) genau dann ein duales Paar ist, wenn (W, V, φ^{op}) es ist, wobei $\varphi^{op}(w, v) = \varphi(v, w)$.

Was ist der Bezug zwischen der Darstellungsmatrix von φ^{op} bezüglich der Basen C von W und B von V und der entsprechenden Matrix für φ?

Lemma 2.19. *Seien V und W zwei endlichdimensionale Vektorräume zusammen mit einer bilinearen Abbildung $\varphi : V \times W \to \mathbb{K}$. Das Tripel (V, W, φ) ist genau dann ein duales Paar, wenn φ* nicht-ausgeartet *ist, das heißt:*

- *Gegeben einen Vektor v aus V derart, dass $\varphi(v, w) = 0_{\mathbb{K}}$ für alle w aus W, ist $v = 0_V$.*
- *Gegeben einen Vektor w aus W derart, dass $\varphi(v, w) = 0_{\mathbb{K}}$ für alle v aus V, ist $w = 0_W$.*

Beweis. Wenn (V, W, φ) ein duales Paar bildet, so ist (W, V, φ^{op}) ein duales Paar. Insbesondere müssen wir nur den ersten Punkt zeigen, weil der andere für die Bilinearform φ^{op} folgt. Sei also v in V mit $\varphi(v, w) = 0_{\mathbb{K}}$ für alle w aus W, d.h. $F_\varphi(v)$ ist die Nullabbildung. Da F_φ ein Isomorphismus ist, muss somit v der Nullvektor aus V sein, wie gewünscht.

Wir nehmen nun an, dass φ nicht ausgeartet ist und müssen zeigen, dass F_φ ein Isomorphismus ist. Nach dem ersten Punkt ist die Abbildung F_φ injektiv, also $\dim_{\mathbb{K}} V \leq \dim_{\mathbb{K}} W^* = \dim_{\mathbb{K}} W$. Nach dem zweiten Punkt ist die Abbildung $F_{\varphi^{op}}$ auch injektiv, also $\dim_{\mathbb{K}} W \leq \dim_{\mathbb{K}} V^* = \dim_{\mathbb{K}} V$, und daher gilt Gleichheit. Insbesondere ist der Monomorphismus F_φ ein Isomorphismus wegen dem Rangsatz. $\square$

Aufgabe. Gegeben ein duales Paar (V, W, φ) sind die Basen $B = \{v_1, \ldots, v_n\}$ von V und $\{w_1, \ldots, w_n\}$ von W *dual zueinander*, falls $\varphi(v_i, w_j) = \delta_{ij}$, wobei δ_{ij} das Kronecker-Delta ist.

Zeige, dass V und W Basen besitzen, welche dual zueinander sind.

Proposition 2.20. *Sei $\varphi : V \times W \to \mathbb{K}$ ein duales Paar. Für jeden Endomorphismus $G : W \to W$ gibt es einen Endomorphismus $G^t : V \to V$, genannt* der zu G adjungierte Endomorphismus, *welcher durch das kommutative Diagramm*

$$\begin{tikzcd} V \arrow[r, "F_\varphi"] \arrow[d, "G^t"] \arrow[dr, phantom, "\square"] & W^* \arrow[d, "G^*"] \\ V \arrow[r, "F_\varphi"] & W^* \end{tikzcd}$$

gegeben wird, wobei $G^* : W^* \to W^*$ die zu G duale Abbildung ist. Der zu G adjungierte Endomophismus G^t ist durch die Identität

$$\varphi(v, G(w)) = \varphi(G^t(v), w) \ \forall v \in V \ \forall w \in W$$

eindeutig bestimmt.

Beweis. Wir definieren $G^t : V \to V$ als $G^t = F_\varphi^{-1} \circ G^* \circ F_\varphi$, weil die Abbildung $F_\varphi : V \to W^*$ ein Isomorphismus und somit invertierbar ist.

Wir zeigen zuerst, dass G^t die obige Identität erfüllt. Seien also v aus V und w aus W. Nun ist $F_\varphi(v)$ ein Element aus W^*, also liegt $G^* \circ F_\varphi(v)$ in W^*. Weil F_φ ein Isomorphismus ist, gibt es einen (einzigen) Vektor v_1 aus V derart, dass $F_\varphi(v_1)$ gleich dem Funktional $G^* \circ F_\varphi(v)$ ist. Aus der Definition des adjungierten Endomorphismus G^t folgt nun $v_1 = G^t(v)$.

Somit ergibt sich

$$\varphi(v, G(w)) = F_\varphi(v)(G(w)) = G^* \circ F_\varphi(v)(w) = F_\varphi(v_1)(w) = \varphi(v_1, w) = \varphi(G^t(v), w),$$

wie gewünscht.

Wir zeigen noch, dass G durch die obige Identität eindeutig bestimmt wird. Sei hierfür $G_1 : V \to V$ ein Endomorphismus derart, dass für alle v aus V und w aus W

$$\varphi(v, G(w)) = \varphi(G_1(v), w).$$

Aus der Bilinearität (und weil G^t die obige Identität auch erfüllt) folgt, dass

$$\varphi(G_1(v) - G^t(v), w) = 0_{\mathbb{K}} \text{ für alle } w \text{ aus } W.$$

Da φ nicht-ausgeartet ist, muss $G_1(v) - G^t(v) = 0_V$ für alle v aus V und somit $G_1 = G^t$, wie gewünscht. $\square$

Aufgabe. Seien im dualen Paar (V, W, φ) zwei zueinander duale Basen B von V und C von W gegeben. Wenn A die Darstellungsmatrix des Endomorphismus $G : W \to W$ bezüglich der Basis C ist, was ist die Darstellungsmatrix von G^t bezüglich der Basis B?

2.3 Orthogonalität in dualen Paaren

Bemerkung 2.21. Betrachte das duale Paar (V, W) bezüglich der bilinearen Abbildung (oder Form) $\varphi : V \times W \to \mathbb{K}$. Gegeben einen Unterraum U von V, definiere das *orthogonale Komplement* $U^\perp$ als die Menge

$$U^\perp = \{w \in W \mid \varphi(u, w) = 0_\mathbb{K} \forall u \in U\}.$$

Es lässt sich leicht zeigen, dass $U^\perp$ ein Unterraum von W ist. Ferner gilt $V^\perp = \{0_W\}$, weil φ nicht-ausgeartet ist, und $\{0_V\}^\perp = W$ wegen Linearität.

Wenn wir orthogonale Komplementation zweimal anwenden, bekommen wir die Identität: $(U^\perp)^\perp = U$. Es ist klar, dass $U \subset (U^\perp)^\perp$. Falls es einen Vektor v in $(U^\perp)^\perp \setminus U$ gäbe, könnten wir ein Funktional H in V^* so finden, dass $H_{\restriction U}$ trivial ist, aber $H(v) = 1_\mathbb{K}$. Weil (W, V, φ^{op}) auch ein duales Paar ist, gäbe es einen Vektor w mit $F_{\varphi^{op}}(w) = H$.

Für jedes u aus U wäre $H(u) = 0_\mathbb{K} = \varphi^{op}(w, u) = \varphi(u, w)$ und somit w in $U^\perp$. Insbesondere hätten wir $\varphi(v, w) = 0_\mathbb{K}$, aber $H(v) = 1_\mathbb{K} \neq 0_\mathbb{K}$, was den gewünschten Widerspruch liefert.

Wir werden nun zeigen, dass das orthogonale Komplement tatsächlich eine Art *Komplementärraum* ist (obwohl er nicht im selben Vektorraum liegt).

Lemma 2.22. *Gegeben ein duales Paar (V, W) bezüglich der Bilinearform $\varphi : V \times W \to \mathbb{K}$, bilden für jeden Unterraum U von V die Vektorräume V/U und $U^\perp$ ein duales Paar bezüglich der bilinearen Abbildung*

$$\begin{array}{rccc} \widetilde{\varphi}: & V/U \times U^\perp & \to & \mathbb{K} \\ & (v + U, w) & \mapsto & \varphi(v, w) \end{array}$$

Insbesondere ist

$$\dim_\mathbb{K} V = \dim_\mathbb{K} U + \dim_\mathbb{K} U^\perp.$$

Beweis. Beachte, dass die Abbildung $\widetilde{\varphi}$ wohldefiniert ist: Falls die Nebenklassen $v + U$ und $v_1 + U$ gleich sind, liegt die Differenz $v - v_1$ in U, also $\varphi(v - v_1, w) = 0_\mathbb{K}$, weil w aus dem Unterraum $U^\perp$ kommt. Wegen der Linearität gilt

$$\widetilde{\varphi}(v_1 + u, w) = \varphi(v_1, w) = \varphi(v_1, w) + 0_K = \varphi(v_1, w) + \varphi(v - v_1, w) = \varphi(v, w) = \widetilde{\varphi}(v + U, w).$$

Die Abbildung $\widetilde{\varphi}$ ist klarerweise bilinear, weil φ bilinear ist und die Vektorraumstruktur auf dem Quotientenraum V/U kompatibel mit der Summe der Repräsentanten ist.

Nach dem Lemma 2.19 müssen wir nur zeigen, dass $\widetilde{\varphi}$ nicht-ausgeartet ist: Eine Bedingung ist klar, weil $\widetilde{\varphi}(v + U, w) = 0_\mathbb{K}$ für alle Nebenklassen $v + U$ bedeutet, dass $\varphi(v, w) = 0_\mathbb{K}$ für alle Vektoren v aus V und somit w der Nullvektor sein muss.

Für die andere Bedingung sei nun $v + U$ eine Nebenklasse derart, dass $\widetilde{\varphi}(v + u, w) = 0_\mathbb{K}$ für alle w aus $U^\perp$. Dies bedeutet, dass v in $(U^\perp)^\perp = U$ liegt und somit $v + U$ der Nullvektor im Quotientenraum V/U ist.

Die letzte Behauptung folgt sofort, da $\dim_\mathbb{K} V/U = \dim_\mathbb{K} V - \dim_\mathbb{K} U$. □

Definition 2.23. Eine *symmetrische Bilinearform* auf dem $\mathbb{K}$-Vektorraum V ist eine Bilinearform $\varphi : V \times V \to \mathbb{K}$ derart, dass $\varphi(v, u) = \varphi(u, v)$ für alle v und u aus V.

Aufgabe. Zeige, dass eine Bilinearform genau dann symmetrisch ist, wenn ihre Darstellungsmatrix A bezüglich der Basis B *symmetrisch* ist, das heißt, $A = {}^tA$.

Für symmetrische Bilinearformen ist der Begriff der Orthogonalität zwischen Vektoren eine symmetrische Relation: Der Vektor u ist orthogonal zum Vektor v, notiert mit $u \perp v$, falls $\varphi(v, u) = 0_{\mathbb{K}}$, oder äquivalent dazu, falls u in $\operatorname{Lin}(\{v\})^{\perp}$ liegt.

Im nächsten Abschnitt werden wir eine wichtige Klasse symmetrischer Bilinearformen untersuchen: *Skalarprodukte* in euklidischen Räumen.

Kapitel 3

Skalarprodukte

Im Folgenden sei $\mathbb{K}$ ein Körper.

3.1 Euklidische Räume

Definition 3.1. Für eine symmetrische Bilinearform auf einem $\mathbb{K}$-Vektorraum V der Dimension n ist ihre assoziierte *quadratische Form* die Abbildung

$$\begin{array}{rccc} Q: & V & \to & \mathbb{K} \\ & v & \mapsto & \varphi(v,v) \end{array}$$

Sei $A = (a_{ij})$ die Darstellungsmatrix von φ bezüglich der Basis $\{v_1, \ldots, v_n\}$ von V. Wir können jeden Vektor v aus V durch seine Koordinaten $(\lambda_1, \ldots, \lambda_n)$ repräsentieren.

Mit dieser Identifikation ist

$$Q(v) = \varphi(v,v) = (\lambda_1, \ldots, \lambda_n) \cdot A \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \sum_{1 \le i,j \le n} a_{ij} \lambda_i \lambda_j.$$

Der obige Ausdruck ist ein multivariables Polynom (in den Variablen $\lambda_1, \ldots, \lambda_n$), in dem nur Monome vom Grad 2 vorkommen. Dies erklärt, warum Q quadratische Form heißt.

Für den euklidischen n-dimensionalen Raum $\mathbb{R}^n$ mit dem *Standardskalarprodukt*

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{1 \le i \le n} x_i y_i$$

ist die quadratische Form gegeben durch

$$\sum_{1 \le i \le n} x_i^2.$$

Bemerkung 3.2. (Polarisationssatz) Jede symmetrische Bilinearform über einem Körper mit Charakteristik verschieden von 2 wird durch ihre quadratische Form eindeutig bestimmt, denn

$$Q(u+v) = \varphi(u,u) + \varphi(v,v) + 2\varphi(u,v)$$

und

$$Q(u-v) = \varphi(u,u) + \varphi(v,v) - 2\varphi(u,v)$$

also $\varphi(u,v) = {}^{Q(u+v)-Q(u-v)}/_{4}$.

Für die folgende Definition benötigen wir, dass der Grundkörper angeordnet ist. Daher werden alle Vektorräume in diesem Abschnitt Vektorräume über dem Körper $\mathbb{R}$ der reellen Zahlen sein.

Definition 3.3. Eine symmetrische Bilinearform $\varphi : V \times V \to \mathbb{R}$ auf dem $\mathbb{R}$-Vektorraum V ist:

- *positiv semidefinit*, falls $\varphi(v, v) \geq 0$ für alle v aus V.
- *negativ semidefinit*, falls $\varphi(v, v) \leq 0$ für alle v aus V.
- *positiv definit*, falls $\varphi(v, v) > 0$ für alle $v \neq 0_V$ aus V.
- *negativ definit*, falls $\varphi(v, v) < 0$ für alle $v \neq 0_V$ aus V.
- *indefinit*, falls keiner der obigen Fälle eintritt.

Beachte, dass φ genau dann positiv (semi)definit ist, wenn die Bilinearform $-\varphi$ negativ (semi)definit ist.

Beispiel 3.4. Das Standardskalarprodukt auf dem $\mathbb{R}$-Vektorraum $\mathbb{R}^n$

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ ((x_1, \ldots, x_n), (y_1, \ldots, y_n)) & \mapsto & \sum\limits_{1 \leq i \leq n} x_i y_i \end{array}$$

ist positiv definit, weil die Summe von nicht-trivialen Quadraten echt positiv ist.

Die Bilinearform

$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ ((x_1, \ldots, x_n), (y_1, \ldots, y_n)) & \mapsto & x_1 y_1 \end{array}$$

ist positiv semidefinit, aber nicht positiv definit, falls $n \geq 2$.

Die Bilinearform

$$\begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ ((x_1, x_2), (y_1, y_2)) & \mapsto & x_1 y_1 - x_2 y_2 \end{array}$$

ist indefinit.

Definition 3.5. Ein *euklidischer Raum* ist ein $\mathbb{R}$-Vektorraum V zusammen mit einer positiv definiten symmetrischen Bilinearform, gennant *Skalarprodukt*, welche wir mit der Notation $\langle -, - \rangle$ bezeichnen werden.

Lemma 3.6. (*Cauchy-Schwarz'sche Ungleichung*) *Gegeben einen euklidischen Raum* $(V, \langle -, - \rangle)$, *setze*

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Die Ungleichung

$$|\langle v, w \rangle| \leq \|v\| \, \|w\|$$

gilt für alle v *und* w *aus* V.

Beachte, dass der Wert $\|v\|$ eine reelle Zahl ist, weil das Skalarprodukt positiv semidefinit ist.

Beweis. Für beliebiges λ aus $\mathbb{R}$ gilt

$$0 \le \langle v - \lambda w, v - \lambda w\rangle = \langle v, v\rangle + \lambda^2\langle w, w\rangle - 2\lambda\langle v, w\rangle = \|v\|^2 + \lambda^2 \|w\|^2 - 2\lambda\langle v, w\rangle.$$

Falls $w = 0_V$, gilt die Behauptung des Lemmas trivialerweise wegen der Linearität der Bilinearform. Sei also ohne Beschränkung der Allgemeinheit $w \neq 0_W$. Weil das Skalarprodukt positiv definit ist, gilt $\|w\|^2 = \langle w, w\rangle > 0$. Setze nun $\lambda = \|w\|^{-2}\langle v, w\rangle$ in die obige Ungleichung ein, also

$$0 \le \|v\|^2 + \langle v, w\rangle^2 \|w\|^{-2} - 2\langle v, w\rangle^2 \|w\|^{-2} = \|v\|^2 - \langle v, w\rangle^2 \|w\|^{-2},$$

oder äquivalent dazu

$$\langle v, w\rangle^2 \le \|v\|^2 \|w\|^2.$$

Nun ziehen wir die Quadratwurzeln und folgern, dass

$$|\langle v, w\rangle| \le \|v\| \|w\|,$$

wie gewünscht. □

Definition 3.7. Ein $\mathbb{R}$-Vektorraum V ist *normiert*, falls er eine Abbildung $\|-\| : V \to \mathbb{R}$, gennant *Norm*, mit folgenden Eigenschaften besitzt:

- $\|v\| \ge 0$ für alle v aus V. Ferner ist
$$\|v\| = 0 \iff v = 0_V.$$
- $\|\lambda v\| = |\lambda| \|v\|$ für alle v aus V und λ aus $\mathbb{R}$.
- *Dreiecksungleichung:* $\|v + w\| \le \|v\| + \|w\|$ für alle v und w aus V.

Beispiel 3.8. Die Abbildungen

$$\begin{array}{cccc} \|\|_{\infty_1}: & \mathbb{R}^n & \to & \mathbb{R} \\ & (x_1, \dots, x_n) & \mapsto & \sum\limits_{1\le i\le n} |x_i| \end{array} \quad \text{und} \quad \begin{array}{cccc} \|\|_{\infty_2}: & \mathbb{R}^n & \to & \mathbb{R} \\ & (x_1, \dots, x_n) & \mapsto & \sup\limits_{1\le i\le n} |x_i| \end{array}$$

sind Normen auf dem $\mathbb{R}$-Vektorraum $\mathbb{R}^n$. Des Weiteren definiert die Abbildung

$$f \mapsto \int_0^1 |f(t)|dt$$

eine Norm auf dem $\mathbb{R}$-Vektorraum aller stetigen Funktionen $f : [0, 1] \to \mathbb{R}$.

Proposition 3.9. *Jeder euklidischer Raum* $(V, \langle -, -\rangle)$ *ist ein normierter Raum bezüglich der* euklidischen *Norm*

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

Beweis. Es ist offensichtlich, dass die obige Norm keine negativen Werte annimmt. Ferner ist die Norm eines nicht-trivialen Vektors nicht Null, weil das Skalarprodukt positiv definit ist. Die Kompatibilität mit dem Skalarprodukt ist auch klar.

Für die Dreiecksungleichung folgt aus

$$\|v + w\|^2 = \langle v + w, v + w\rangle = \|v\|^2 + \|w\|^2 + 2\langle v, w\rangle \overset{3.6}{\le} \|v\|^2 + \|w\|^2 + 2 \|v\| \|v\| = (\|v\| + \|w\|)^2,$$

dass

$$\|v + w\| \le \|v\| + \|w\|,$$

wie gewünscht. □

3.2 Orthogonalität und Orthonormalität

Wie im Abschnitt 2.3 bereits eingeführt, können wir in jedem euklidischen Raum mit Hilfe des Skalarproduktes über Orthogonalität sprechen. Im folgenden sei $(V, \langle -, - \rangle)$ ein euklidischer Raum.

Definition 3.10. Zwei Vektoren v und w aus V sind *orthogonal zueinander* (bezeichnet mit $v \perp w$), falls $\langle v, w \rangle = 0$.

Lemma 3.11. (*Der Satz von Pythagoras*) *Die Vektoren v und w sind genau dann orthogonal zueinander, falls*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 .$$

Beweis. Beachte, dass

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \|v\|^2 + \|w\|^2 + 2\langle v, w \rangle.$$

Insbesondere gilt

$$v \perp w \iff \|v + w\|^2 = \|v\|^2 + \|w\|^2 .$$

□

Aufgabe. Aus der Cauchy-Schwarz'schen Ungleichung folgt

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \, \|w\|} \leq 1$$

für alle nicht-triviale Vektoren v und w aus dem euklidischen Raum $(V, \langle -, - \rangle)$. Daher gibt es genau einen eindeutigen Wert θ im Intervall $[0, \pi]$ derart, dass

$$\cos \theta = \frac{\langle v, w \rangle}{\|v\| \, \|w\|}.$$

Der Wert θ heißt der *Winkel* zwischen v und w.

Zeige, dass v und w genau dann orthogonal zueinander sind, wenn der Winkel zwischen ihnen genau $\pi/2$ beträgt.

Definition 3.12. Ein *orthogonales System* ist eine Familie $(\{v_i\}_{i \in I})$ nicht-trivialer paarweise orthogonaler Vektoren aus V, d.h.

$$v_i \perp v_j \quad \text{für} \quad i \neq j.$$

Ein *orthonormales System* ist orthogonales System $(\{v_i\}_{i \in I})$ derart, dass jeder Vektor v_i Norm $\|v_i\| = 1$ hat.

Eine *Orthonormalbasis* (ONB) ist eine Basis von V, welche ein orthonormales System ist.

Es lässt sich leicht zeigen, dass jedes orthogonale System eine linear unabhängige Familie von Vektoren bildet.

Bemerkung 3.13. Gegeben eine Orthonormalbasis $B = (b_i)_{i\in I}$ des euklidischen Raumes $(V, \langle -,-\rangle)$, können wir die Koordinaten von Vektoren sowie die Darstellungsmatrizen von Endomorphismen leicht berechnen: Gegeben einen Vektor v aus V, schreibe

$$v = \sum_{i\in I} \lambda_i b_i,$$

für Skalare λ_i aus $\mathbb{R}$ mit $\lambda_i = 0$ für fast alle i aus I (so ist die obige Summe endlich und somit wohldefiniert). Dann ist das Produkt $\langle v, b_j\rangle$ die j-te Koordinate λ_j von v bezüglich B.

Insbesondere ist die Darstellungsmatrix $A = (a_{ij})$ des Endomorphismus $F : V \to V$ bezüglich B leicht zu gewinnen, falls V endlichdimensional ist. Denn der Eintrag a_{ij} ist die i-te Koordinate des Vektors $F(b_j)$, also $a_{ij} = \langle F(b_j), b_i\rangle$.

Des Weiteren ist das Skalarprodukt zweier Vektoren v und w aus V genau die Summe

$$\langle v, w\rangle = \sum_{i\in I} \langle v, b_i\rangle\langle w, b_i\rangle,$$

wegen der Bilinearität des Skalarprodukt.

Satz 3.14. (*Gram-Schmidt'sches Orthogonalisierungsverfahren*) *Seien für eine natürliche Zahl $n \geq 1$ linear unabhängige Vektoren $v_1, \dots, v_n$ aus dem euklidischen Raum $(V, \langle -,-\rangle)$ gegeben. Es gibt eine Orthonormalbasis $\{b_1, \dots, b_n\}$ des Unterraumes* $\mathrm{Lin}(\{v_1, \dots, v_n\})$. *Falls die Vektoren $v_1, \dots, v_k$ bereits ein orthonormales System für $k \leq n$ bilden, finden wir eine Orthonormalbasis $\{b_1, \dots, b_n\}$ mit $b_i = v_i$ für $1 \leq i \leq k$.*

Insbesondere besitzt jeder nicht-triviale endlichdimensionale euklidische Raum eine Orthonormalbasis.

Beweis. Die letzte Behauptung folgt sofort aus dem Orthogonalisierungsverfahren, welches wir nun induktiv über n beweisen. Für $n = 1$ ist $v_1 \neq 0$ und

$$b_1 = \frac{v_1}{\|v_1\|}$$

hat klarerweise Norm 1 und spannt denselben Unterraum wie v_1 auf.

Wir nehmen nun an, dass das Verfahren für die natürliche Zahl n bereits bewiesen wurde und zeigen es für $n+1$. Gegeben eine linear unabhängige Familie $\{v_1, \dots, v_{n+1}\}$ von Vektoren, sei $\{b_1, \dots, b_n\}$ eine Orthonormalbasis des n-dimensionalen Unterraumes $\mathrm{Lin}(\{v_1, \dots, v_n\})$. Wir definieren zuerst den Hilfsvektor

$$b'_{n+1} = v_{n+1} - \sum_{1\leq i\leq n} \langle v_{n+1}, b_i\rangle b_i.$$

Beachte, dass b'_{n+1} nicht der Nullvektor sein kann, weil sonst v_{n+1} im Unterraum $\mathrm{Lin}(\{v_1, \dots, v_n\})$ wäre, was der linearen Unabhängigkeit der Familie widerspricht. Also $b'_{n+1} \neq 0_V$ und wir können ihn normieren durch

$$b_{n+1} = \frac{b'_{n+1}}{\|b'_{n+1}\|}.$$

Aus der Konstruktion folgt, dass

$$\mathrm{Lin}(\{b_1, \dots, b_{n+1}\}) \subset \mathrm{Lin}(\{v_1, \dots, v_{n+1}\}).$$

Wenn wir zeigen, dass die Vektoren $b_1, \ldots, b_{n+1}$ paarweise orthogonal zueinander sind, muss oben aus Dimensionsgründen Gleichheit gelten. Wir müssen also nur noch zeigen, dass $b_{n+1} \perp b_i$ für $1 \leq 1 \leq n$, oder äquivalent dazu, dass b'_{n+1} orthogonal zu jedem b_j ist. In der Tat gilt

$$\langle b'_{n+1}, b_j \rangle = \langle v_{n+1}, b_j \rangle - \sum_{1 \leq i \leq n} \langle v_{n+1}, b_i \rangle \langle b_i, b_j \rangle = \langle v_{n+1}, b_j \rangle - \langle v_{n+1}, b_j \rangle \, \|b_j\|^2 = 0,$$

wie gewünscht.

Wir nehmen nun an, dass die Vektoren $v_1, \ldots, v_k$ orthonormal sind und zeigen induktiv über $1 \leq i \leq k$, dass $v_i = b_i$. Für $i = 1$ folgt es sofort aus dem Induktionsanfang für $n = 1$. Angenommen nun, dass die Behauptung für alle $i' \leq i$ gilt, müssen wir zeigen, dass $v_{i+1} = b_{i+1}$. Zuerst betrachte den konstruierten Hilfsvektor

$$b'_{i+1} = v_{i+1} - \sum_{1 \leq j \leq i} \langle v_{i+1}, b_j \rangle b_j.$$

Nach unserer Induktionsannahme ist $b_j = v_j$ für $1 \leq j \leq i$, also $\langle v_{i+1}, b_j \rangle = \langle v_{i+1}, v_j \rangle = 0$ und somit

$$b'_{i+1} = v_{i+1} - \sum_{1 \leq j \leq i} \langle v_{i+1}, v_j \rangle v_j = v_{i+1}.$$

Da v_{i+1} bereits normiert ist, gilt $b_{i+1} = v_{i+1}$, wie gewünscht. □

Definition 3.15. Zwei Teilmengen A und B des euklidischen Raumes $(V, \langle -, - \rangle)$ sind *orthogonal* (bezeichnet mit $A \perp B$), falls $v \perp w$ für alle v aus A und w aus B.

Das *orthogonale Komplement* der Teilmenge A von V ist die Menge

$$A^\perp = \{v \in V \mid \{v\} \perp A\}.$$

Die leere Menge sowie der triviale Unterraum $\{0_V\}$ sind orthogonal zu jeder anderen Teilmenge von V. Aus der Linearität des Skalarproduktes folgt sofort die folgende Behauptung.

Lemma 3.16. $A \perp B \iff \mathrm{Lin}(A) \perp \mathrm{Lin}(B)$.

Aufgabe. Zeige, dass das orthogonale Komplement $A^\perp$ der Teilmenge A der größte Unterraum von V ist, welcher zu A orthogonal ist.

Bemerkung 3.17. Die orthogonalen Komplemente von A und von $\mathrm{Lin}(A)$ stimmen überein: Es ist klar, dass $\mathrm{Lin}(A)^\perp$ in $A^\perp$ enthalten ist. Aus $A \perp A^\perp$ folgt $\mathrm{Lin}(A) \perp \mathrm{Lin}(A^\perp)$ und somit $\mathrm{Lin}(A) \perp A^\perp$, was die andere Inklusion liefert.

Proposition 3.18. *Gegeben einen endlichdimensionalen Unterraum U des euklidischen Raumes $(V, \langle -, - \rangle)$, ist V die direkte Summe von U und seinem orthogonalen Komplement $U^\perp$.*

Insbesondere ist $U = (U^\perp)^\perp$.

Beachte, dass das orthogonale Komplement $U^\perp$ von U nicht unbedingt endlichdimensional sein muss (gerade, wenn V unendlichdimensional ist).

Beweis. Es lässt sich leicht zeigen, dass U und $U^\perp$ transversal zueinander liegen (auch wenn U unendlichdimensional wäre). Wir müssen also nur zeigen, dass $V = U + U^\perp$. Wenn U trivial ist, gilt $U^\perp = V$ und die Behauptung folgt.

Ohne Beschränkung der Allgemeinheit können wir also annehmen, dass U nicht trivial ist und somit eine (endliche) Basis $\{b_1, \ldots, b_n\}$ besitzt, welche mit Hilfe des Gram-Schmidt'schen Orthogonalisierungsverfahren 3.14 orthonormal gewählt werden kann.

Gegeben einen Vektor v aus V, liegt die Linearkombination

$$u = \sum_{1 \le i \le n} \langle v, b_i \rangle b_i$$

in U. Wir müssen also nur noch zeigen, dass $v - u$ in $U^\perp$ liegt, oder äquivalent dazu, dass $\langle v - u, b_j \rangle = 0$ für alle $1 \le j \le n$, nach dem Lemma 3.16. Aus der Konstruktion folgt

$$\langle v - u, b_j \rangle = \langle v, b_j \rangle - \langle u, b_j \rangle = \langle v, b_j \rangle - \sum_{1 \le i \le n} \langle v, b_i \rangle \langle b_i, b_j \rangle = \langle v, b_j \rangle - \langle v, b_j \rangle \left\| b_j \right\|^2 = 0,$$

wie gewünscht.

Für die Gleichheit $U = (U^\perp)^\perp$ beachte, dass U in $(U^\perp)^\perp$ liegt, weil U orthogonal zum Komplement $U^\perp$ ist. Gegeben nun einen Vektor v aus $(U^\perp)^\perp \subset V = U \oplus U^\perp$, schreibe v eindeutig als $v = u + w$ mit u aus U und w aus $U^\perp$. Der Vektor w ist genau dann Null, wenn $v = u$ in U liegt. Es ist

$$w = v - u \text{ in } (U^\perp)^\perp \cap U^\perp = \{0_V\}$$

und somit $(U^\perp)^\perp \subset U$, wie gewünscht. $\square$

Mit Hilfe der vorigen Proposition können wir nun die orthogonale Projektion auf einen (endlichdimensionalen) Unterraum definieren.

Bemerkung 3.19. Gegeben einen Unterraum U des euklidischen Raumes $(V, \langle -, - \rangle)$ und einen Vektor v aus V, welcher in $U \oplus U^\perp$ liegt (was für endlichdimensionale U immer der Fall ist, nach der Proposition 3.18), sei u aus U derart, dass $v - u$ orthogonal zu allen Vektoren aus U ist, oder äquivalenz dazu, dass $v - u$ in $U^\perp$ liegt.

Der Vektor u wird von v eindeutig bestimmt (weil U und $U^\perp$ transversal sind) und heißt die *orthogonale Projektion* $\mathrm{Pr}_U^\perp(v)$ von v auf V.

Aufgabe. Betrachte den Unterraum $U = \mathrm{Lin}(\{e_1\})$ von $\mathbb{R}^2$ mit dem Standardskalarprodukt. Berechne für jeden Vektor (a, b) aus $\mathbb{R}^2$ die orthogonale Projektion $\mathrm{Pr}_U^\perp((a, b))$.

Lemma 3.20. *Für einen Unterraum U des euklidischen Raumes $(V, \langle -, - \rangle)$ sowie Vektoren v aus V und u aus U sind folgende Behauptungen äquivalent:*

- *Die orthogonale Projektion $\mathrm{Pr}_U^\perp(v)$ ist definiert und gleich u.*
- *Für alle u_1 aus U ist*
$$\|v - u_1\| \ge \|v - u\|,$$
wobei Gleichheit genau dann gilt, wenn $u_1 = u$.

Beweis. Falls die orthogonale Projektion $\mathrm{Pr}_U^\perp(v)$ definiert ist, schreibe $v = u + w$ mit w aus $U^\perp$. Für u_1 aus U ist

$$\|v - u_1\|^2 = \langle v - u_1, v - u_1 \rangle = \langle u - u_1 + w, u - u_1 + w \rangle \overset{3.11}{=} \|u - u_1\|^2 + \|w\|^2,$$

weil w und $u - u_1$ orthogonal zueinander sind. Insbesondere ist $\|v - u_1\|^2 \geq \|w\|^2 = \|v - u\|^2$, wie gewünscht.

Des Weiteren gilt genau dann Gleichheit, wenn $\|u - u_1\| = 0$, oder äquivalent dazu $u = u_1$.

Wir nehmen nun an, dass

$$\|v - u_1\| > \|v - u\|$$

für alle u_1 aus $U \setminus \{u\}$. Da $v = u + (v - u)$, müssen wir nur zeigen, dass der Vektor $v - u$ in $U^\perp$ liegt. Ansonsten gibt es einen Vektor u_1 aus U mit $\langle v - u, u_1 \rangle \neq 0$, also $u_1 \neq 0_V$. Ohne Beschränkung der Allgemeinheit ist u_1 normiert. Weil der Vektor $u' = u + \langle v - u, u_1 \rangle u_1$ nicht gleich u sein kann, ist

$$\begin{aligned} \|v - u\|^2 < \|v - u'\|^2 &= \|(v - u) - \langle v - u, u_1 \rangle u_1\|^2 = \\ &= \langle (v - u) - \langle v - u, u_1 \rangle u_1, (v - u) - \langle v - u, u_1 \rangle u_1 \rangle = \\ &= \|v - u\|^2 + \langle v - u, u_1 \rangle^2 \|u_1\|^2 - 2\langle v - u, u_1 \rangle \langle v - u, u_1 \rangle = \|v - u\|^2 - \langle v - u, u_1 \rangle^2 \end{aligned}$$

Weil die reelle Zahl $\langle v - u, u_1 \rangle^2$ echt positiv ist, folgt

$$\|v - u\|^2 < \|v - u\|^2 - \langle v - u, u_1 \rangle^2 < \|v - u\|^2,$$

was den gewünschten Widerspruch liefert. □

3.3 Unitäre Räume und Adjungierte Endomorphismen

Wenn wir den endlichdimensionalen Vektorraum $\mathbb{C}^n$ betrachten, induziert die Verknüpfung

$$\begin{array}{rcl} \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ ((x_1, \ldots, x_n), (y_1, \ldots, y_n)) & \mapsto & \sum\limits_{1 \leq i \leq n} x_i y_i \end{array}$$

keine Norm mit Werten im Körper $\mathbb{R}$, weil für eine komplexe Zahl z der Ausdruck z^2 nicht unbedingt reell sein muss. Jedoch ist $|z|^2 = z\bar{z}$ eine reelle Zahl, wobei $\bar{z} = a - ib$ die komplex konjugierte der komplexen Zahl $z = a + ib$ aus $\mathbb{C}$ ist, siehe [**Skript LAI**, Beispiel 1.38]. Diese Überlegung begründet die folgende Definition.

Definition 3.21. Eine *hermitesche* Bilinearform (oder eine *Sesquilinearform*) auf einem $\mathbb{C}$-Vektorraum V ist eine Abbildung $\varphi : V \times V \to \mathbb{C}$ mit folgenden Eigenschaften:

- Die Form φ ist linear in der ersten Koordinate: Für alle λ_1 und λ_2 aus $\mathbb{C}$ sowie v_1, v_2 und w aus V gilt
$$\varphi(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 \varphi(v_1, w) + \lambda_2 \varphi(v_2, w).$$
- Die Form erfüllt die *hermitesche Symmetrie*:
$$\varphi(w, v) = \overline{\varphi(v, w)} \text{ für alle } v \text{ und } w \text{ aus } V.$$

Insbesondere ist die hermitesche Bilinearform *semilinear* in der zweiten Koordinate: Für alle λ_1 und λ_2 aus $\mathbb{C}$ sowie v, w_1 und w_2 aus V gilt

$$\varphi(v, \lambda_1 w_1 + \lambda_2 w_2) = \bar{\lambda}_1 \varphi(v, w_1) + \bar{\lambda}_2 \varphi(v, w_2).$$

Die hermitesche Bilinearform ist *positiv definit*, falls $\varphi(v, v)$ für alle $v \neq 0_V$ eine positive reelle Zahl ist. Der Raum V ist *unitär*, falls er eine hermitesche positiv definite Bilinearform besitzt, welche wir auch mit $\langle -, - \rangle$ und als *Skalarprodukt* bezeichnen werden.

Beispiel 3.22. Die Abbildung

$$\begin{array}{rcl} \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ ((x_1, \dots, x_n), (y_1, \dots, y_n)) & \mapsto & \sum\limits_{1 \le i \le n} x_i \bar{y}_i \end{array}$$

ist ein Skalarprodukt auf dem $\mathbb{C}$-Vektorraum $\mathbb{C}^n$.

Bemerkung 3.23. Gegeben einen unitären Raum $(V, \langle -, - \rangle)$, ist

$$\langle v + \lambda w, v + \lambda w \rangle = \langle v, v \rangle + \lambda \bar{\lambda} \langle w, w \rangle + \big(\lambda \langle w, v \rangle + \bar{\lambda} \langle v, w \rangle \big) = \langle v, v \rangle + |\lambda|^2 \langle w, w \rangle +$$
$$+ \Big(\lambda \langle w, v \rangle + \overline{\lambda \langle w, v \rangle} \Big) = \langle v, v \rangle + |\lambda|^2 \langle w, w \rangle + 2\mathrm{Re}(\lambda \langle w, v \rangle),$$

wobei $\mathrm{Re}(z)$ der Realteil der komplexen Zahl z ist.

Insbesondere wird das Skalarprodukt von der assoziierten quadratischen Form $Q(v) = \langle v, v \rangle$ eindeutig bestimmt (mit demselben Argument wie in der Bemerkung 3.2). Des Weiteren gilt auch die Cauchy-Schwarz'sche Ungleichung und somit ist V ein normierter Vektorraum bezüglich $\|v\| = \sqrt{\langle v, v \rangle}$.

Mit Hilfe der obigen Überlegung lassen sich alle Begriffe und Ergebnisse des vorigen Abschnittes (Orthogonalität und Orthonormalität, sowie orthogonale Projektion und das Gramm-Schmidt'sches Verfahren), bis auf den Begriff des Winkels zwischen zwei Vektoren, für unitäre Räume identisch zeigen.

Wir werden ab jetzt euklidische und unitäre Räume gleichzeitig behandeln. Der Grundkörper $\mathbb{K}$ kann also $\mathbb{R}$ oder $\mathbb{C}$ sein, abhängig davon, in welcher Situation wir sind. Um die Schreibweise zu vereinheitlichen, schreiben wir für einen Skalar λ aus $\mathbb{K}$

$$\bar{\lambda} = \begin{cases} \text{das komplex Konjugierte } \bar{\lambda}, \text{ falls } \lambda \text{ aus } \mathbb{C} \text{ kommt.} \\ \lambda, \text{ falls } \lambda \text{ in } \mathbb{R} \text{ liegt.} \end{cases}$$

Bemerkung 3.24. Gegeben einen euklidischen oder unitären Raum $(V, \langle -, - \rangle)$ über dem Körper $\mathbb{K}$ (beachte, dass $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$ ist, abhängig davon ob der Vektorraum euklidisch oder unitär ist) können wir folgenderweise eine *konjugierte* Skalarmultiplikation definieren: Für λ aus $\mathbb{K}$ und v aus V setze

$$\lambda \cdot_{\text{konj}} v = \bar{\lambda} \cdot v,$$

wobei $\cdot$ die ursprüngliche Skalarmultiplikation des $\mathbb{K}$-Vektorraumes V bezeichnet. Die konjugierte Skalarmultiplikation definiert eine neue $\mathbb{K}$-Vektorraumstruktur auf V derart, dass sämtliche Begriffe wie Basen, Dimension und Endomorphismen erhalten bleiben, denn

$$v = \sum_{1 \le i \le n} \lambda_i \cdot_{\text{konj}} v_i \iff v = \sum_{1 \le i \le n} \bar{\lambda}_i \cdot v_i$$

und

$$F(\lambda \cdot_{\text{konj}} v) = F(\bar{\lambda} \cdot v) = \bar{\lambda} \cdot F(v) = \lambda \cdot_{\text{konj}} F(v).$$

Bezeichne mit V_{konj} den Vektorraum V bezüglich der konjugierten Skalarmultiplikation. Beachte, dass im Fall eines euklidischen Vektorraumes V_{konj} (als Vektorraum) genau der Vektorraum V ist. Im Fall eines unitären Raumes sind zumindest die abelschen Gruppen $(V, +)$ und $(V_{\text{konj}}, +)$ identisch.

Lemma 3.25. *Wenn der euklidische oder unitäre Raum V endlichdimensional ist, bilden V und V_{konj} ein duales Paar bezüglich des Skalarproduktes $\langle -, - \rangle : V \times V_{konj} \to \mathbb{K}$.*

Beweis. Nach dem Lemma 2.19 müssen wir nur zeigen, dass $\langle -, - \rangle$ in der ersten Koordinate nicht-ausgeartet ist, wegen der hermiteschen Symmetrie

$$\langle u, w \rangle = \overline{\langle w, u \rangle}.$$

Sei v aus V derart, dass die Abbildung $\langle v, - \rangle : V_{\text{konj}} \to \mathbb{K}$ trivial ist. Insbesondere ist $\|v\|^2 = \langle v, v \rangle = 0_{\mathbb{K}}$, also $\|v\| = 0_{\mathbb{K}}$ und somit $v = 0_V$, wie gewünscht. □

Definition 3.26. Für einen Endomorphismus $F : V \to V$ eines endlichdimensionalen euklidischen oder unitären Raumes $(V, \langle -, - \rangle)$ ist sein *adjungierter* Endomorphismus $F^t : V \to V$ der durch die Identität

$$\langle v, F(w) \rangle = \langle F^t(v), w \rangle, \text{ für alle } v \text{ und } w \text{ aus } V$$

eindeutig definierte Endomorphismus, welcher nach Proposition 2.20 und Lemma 3.25 existiert. Beachte, dass die lineare Abbildung F^t ein Endomorphismus von V_{konj} ist, aber aus der vorigen Diskussion induziert sie eine lineare Abbildung von V nach V (und wir identifizieren diese Abbildung mit F^t.)

Aufgabe. Sei U ein Unterraum des euklidischen oder unitären Raumes $(V, \langle -, - \rangle)$. Zeige, dass U genau dann invariant unter F ist, wenn $U^\perp$ invariant unter F^t ist.

Bemerkung 3.27. Die Darstellungsmatrix des zu F adjungierten Endomorphismus $F^t : V \to V$ lässt sich leicht bezüglich einer Orthonormalbasis $B = \{b_1, \ldots, b_n\}$ berechnen: Gegeben einen Vektor v aus V, schreibe

$$v = \sum_{1 \le i \le n} \langle v, b_i \rangle b_i,$$

siehe Bemerkung 3.13. Dann ist

$$F^t(v) = \sum_{1 \le i \le n} \langle v, F(b_i) \rangle b_i.$$

Wir müssen nur zeigen, dass die obige Summe die definitorische Eigenschaften des adjungierten Endomorphismus erfüllt, wegen der Eindeutigkeit der Definition des adjungierten Endomorphismus. Seien also v und w aus V gegeben. Es ist

$$\langle \sum_{1 \le i \le n} \langle v, F(b_i) \rangle b_i, w \rangle = \sum_{1 \le i \le n} \langle v, F(b_i) \rangle \langle b_i, w \rangle = \sum_{1 \le i \le n} \overline{\langle w, b_i \rangle} \langle v, F(b_i) \rangle = \sum_{1 \le i \le n} \langle v, \langle w, b_i \rangle F(b_i) \rangle =$$

$$= \langle v, \sum_{1 \le i \le n} \langle w, b_i \rangle F(b_i) \rangle = \langle v, \sum_{1 \le i \le n} F(\langle w, b_i \rangle b_i) \rangle = \langle v, F\left(\sum_{1 \le i \le n} \langle w, b_i \rangle b_i \right) \rangle = \langle v, F(w) \rangle,$$

wie gewünscht.

Definition 3.28. Sei $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$. Die *adjungierte* Matrix einer quadratischen Matrix $A = (a_{ij})$ aus $\mathcal{M}_{n \times n}(\mathbb{K})$ ist die $n \times n$-Matrix $A^* = (c_{ij})$ mit $c_{ij} = \overline{a_{ji}}$.

Bemerkung 3.29. Falls der Grundkörper $\mathbb{K} = \mathbb{R}$ ist, folgt $A^* = {}^tA$. Allgemein gilt

$$\det(A^*) = \overline{\det(A)},$$

weil die komplexe Konjugation mit der Summe und dem Produkt kompatibel ist. Insbesondere ist A genau dann regulär, wenn ihre adjungierte Matrix A^* regulär ist.

Des Weiteren gelten die folgenden Identitäten für A und B aus $\mathcal{M}_{n\times n}(\mathbb{K})$ sowie λ aus $\mathbb{K}$:

- $(A^*)^* = A$.
- $(A + B)^* = A^* + B^*$.
- $(\lambda \cdot A)^* = \bar{\lambda} \cdot A^*$.
- $(A \cdot B)^* = B^* \cdot A^*$.

Korollar 3.30. *Gegeben zwei Orthonormalbasen $B = \{b_1, \ldots, b_n\}$ und $C = \{c_1, \ldots, c_n\}$ eines endlichdimensionalen euklidischen oder unitären Vektorraum $(V, \langle -, - \rangle)$ sowie einen Endomorphismus $F : V \to V$ mit Darstellungsmatrix A bezüglich der Basen B (im Definitionsbereich) und C (im Bildbereich) von V ist die Darstellungsmatrix des adjungierten Endomorphismus $F^t : V \to V$ bezüglich der Basen C (im Definitionsbereich) und B (im Bildbereich) von V genau die adjungierte Matrix A^* von A.*

Beweis. Beachte, dass der Eintrag a_{ij} von A die i-te Koordinate des Vektors $F(b_j)$ bezüglich der Orthonormalbasis $C = \{c_1, \ldots, c_n\}$ ist, also $a_{ij} = \langle F(b_j), c_i \rangle$, nach der Bemerkung 3.13. Aus der hermiteschen Symmetrie und der definitorischen Eigenschaft des adjungierten Endomorphismus $F^t : V \to V$ folgt

$$\overline{a_{ji}} = \overline{\langle F(b_i), c_j \rangle} = \langle c_j, F(b_i) \rangle = \langle F^t(c_j), b_i \rangle,$$

was genau der i-ten Koordinate des Vektors $F^t(c_j)$ bezüglich der Basis $\{b_1, \ldots, b_n\}$ entspricht, wie gewünscht. □

3.4 Normale und Selbstadjungierte Abbildungen

In diesem Abschnitt wollen wir den Bezug zwischen einem Endomorphismus F des endlichdimensionalen euklidischen oder unitären Raumes $(V, \langle -, - \rangle)$ und seinem adjungierten Endomorphismus F^t verstehen.

Der erste Schritt dieser Analysis betrachtet den Fall, in welchem die beiden Endomorphismen bezüglich Komposition (als Elemente des Ringes $\mathrm{End}(V)$ betrachtet) miteinander kommutieren. In diesem Fall ist F ein *normaler* Endomorphismus.

Im Folgenden ist $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$ und $(V, \langle -, - \rangle)$ ein n-dimensionaler euklidischer oder unitärer Raum $(V, \langle -, - \rangle)$.

Definition 3.31. Der Endomorphismus $F : V \to V$ ist *normal*, falls $F \circ F^t = F^t \circ F$.

Lemma 3.32. *Folgende Aussagen sind für einen Endomorphismus $F : V \to V$ äquivalent:*

(a) Der Endomorphismus F ist normal.

(b) Für alle v und w aus V gilt

$$\langle F(v), F(w) \rangle = \langle F^t(v), F^t(w) \rangle.$$

Beweis. (a) $\Longrightarrow$ (b) :

$$\langle F(v), F(w)\rangle = \langle F^t(F(v)), w\rangle \overset{F \text{ normal}}{=} \langle F(F^t(v)), w\rangle = \overline{\langle w, F\circ F^t(v)\rangle} =$$
$$= \overline{\langle F^t(w), F^t(v)\rangle} = \langle F^t(v), F^t(w)\rangle.$$

(b) $\Longrightarrow$ (a) : Wir müssen nur zeigen, dass für alle v und w aus V

$$\langle (F^t\circ F - F\circ F^t)(v), w\rangle = 0_V,$$

oder äquivalent dazu

$$\langle F^t\circ F(v), w\rangle = \langle F\circ F^t(v), w\rangle,$$

weil das Skalarprodukt nicht-ausgeartet ist. Nun

$$\langle F^t\circ F(v), w\rangle = \langle F(v), F(w)\rangle \overset{(b)}{=} \langle F^t(v), F^t(w)\rangle = \overline{\langle F^t(w), F^t(v)\rangle} =$$
$$= \overline{\langle w, F(F^t(v))\rangle} = \langle F(F^t(v)), w\rangle = \langle F\circ F^t(v), w\rangle,$$

wie gewünscht. □

Definition 3.33. Eine quadratische Matrix A aus $\mathcal{M}_{n\times n}(\mathbb{K})$ ist *normal*, falls $A\cdot A^* = A^*\cdot A$.

Bemerkung 3.34. Der Endomorphismus $F : V \to V$ ist genau dann normal, wenn seine Darstellungsmatrix A bezüglich einer Orthonormalbasis B von V (sowohl im Definitions- als auch im Bildbereich) normal ist, weil $A^*\cdot A$ die Darstellungsmatrix von $F^t\circ F$ und das Produkt $A\cdot A^*$ die Darstellungsmatrix von $F\circ F^t$ bezüglich B ist.

Wir werden im Satz 3.36 sehen, dass jeder normale Endomorphismus eines endlichdimensionalen unitären Vektorraumes diagonalisierbar ist. Zuerst untersuchen wir den Bezug zwischen den Eigenwerten von F und den Eigenwerten seines adjungierten Endomorphismus F^t.

Proposition 3.35. *Der Vektor v aus V ist genau dann Eigenvektor des normalen Endomorphismus $F : V \to V$ bezüglich des Eigenwertes λ, wenn v Eigenvektor des adjungierten Endomorphismus $F^t : V \to V$ bezüglich des Eigenwertes $\bar\lambda$ ist.*

Insbesondere besitzen F und sein adjungierter Endomorphismus F^t dieselben Eigenwerte, wenn der Grundkörper $\mathbb{R}$ ist.

Beweis. Wir müssen nur die folgende Äquivalenz zeigen:

$$\|F(v) - \lambda v\| = 0_V \iff \left\|F^t(v) - \bar\lambda v\right\| = 0_V.$$

Beachte, dass

$$\langle F(v), v\rangle = \overline{\langle v, F(v)\rangle} = \overline{\langle F^t(v), v\rangle} = \langle v, F^t(v)\rangle$$

und somit

$$\|F(v) - \lambda v\|^2 = \langle F(v)-\lambda v, F(v)-\lambda v\rangle = \langle F(v), F(v)\rangle + \lambda\bar\lambda\langle v, v\rangle - \big(\lambda\langle v, F(v)\rangle + \bar\lambda\langle F(v), v\rangle\big) =$$
$$= \langle F^t(v), F^t(v)\rangle + \lambda\bar\lambda\langle v, v\rangle - \big(\lambda\langle F^t(v), v\rangle + \bar\lambda\langle v, F^t(v)\rangle\big) =$$
$$= \langle F^t(v), F^t(v)\rangle + \lambda\bar\lambda\langle v, v\rangle - \big(\bar\lambda\langle v, F^t(v)\rangle + \lambda\langle F^t(v), v\rangle\big) =$$
$$= \langle F^t(v) - \bar\lambda v, F^t(v) - \bar\lambda v\rangle = \left\|F^t(v) - \bar\lambda v\right\|^2.$$

□

Satz 3.36. *Sei $(V, \langle -, - \rangle)$ ein euklidischer oder unitärer Raum der Dimension $n \geq 1$ und $F : V \to V$ ein Endomorphismus, dessen charakteristisches Polynom $\chi_F(T)$ in Linearfaktoren über $\mathbb{K}$ zerfällt. Der Endomorphismus F ist genau dann normal, wenn V eine Orthonormalbasis B aus Eigenvektoren von F besitzt.*

Insbesondere ist jeder normale Endomorphismus eines nicht-trivialen endlichdimensionalen unitären Raumes diagonalsierbar, nach dem Korollar A.10.

Beweis. Wir nehmen zuerst an, dass F normal ist, also $F \circ F^t = F^t \circ F$ und zeigen die Existenz der gesuchten Basis induktiv über $n = \dim_{\mathbb{K}} V \geq 1$. Für $n = 1$ ist es trivial, weil Eigenvektoren nicht der Nullvektor sind und wir sie deshalb immer normieren können (beachte, dass wir hier nicht verwendet haben, dass F normal ist).

Für $n \geq 2$ sei λ aus $\mathbb{K}$ ein Eigenwert von F (ein solches λ existiert, weil χ_F über $\mathbb{K}$ in Linearfaktoren zerfällt) und wähle einen Eigenvektor v zu λ. Da $v \neq 0_V$, können wir $b_1 = \|v\|^{-1} v$ setzen. Beachte, dass b_1 normiert und auch ein Eigenvektor zu λ ist. Der Unterraum $U = \mathrm{Lin}(\{b_1\})^\perp$ hat Dimension $\dim_{\mathbb{K}} U = n - 1$ nach dem Lemma 2.22 und V ist die direkte Summe von $\mathrm{Lin}(\{b_1\})$ und U, wegen der Proposition 3.18. Da $\mathrm{Lin}(\{b_1\})$ nach der Proposition 3.35 invariant unter F^t ist, folgt, dass U invariant unter F ist (siehe die Aufgabe nach der Definition 3.26). Analog ist U auch invariant unter F^t.

Weil die Identität

$$\langle F(u_1), F(u_2) \rangle = \langle F^t(u_1), F^t(u_2) \rangle$$

auch für alle u_1 und u_2 aus U gilt, siehe Lemma 3.32, folgt, dass die Einschränkungsabbildung $F_{\restriction U}$ normal ist mit adjungiertem Endomorphismus ${F^t}_{\restriction U}$. Wegen der Induktionsannahme finden wir eine Orthonormalbasis $\{b_2, \ldots, b_n\}$ von Eigenvektoren von $F_{\restriction U}$ in U. Klarerweise ist $\{b_1, \ldots, b_n\}$ eine Orthonormalbasis von U, welche aus Eigenvektoren von F besteht.

Für die andere Richtung sei $B = \{b_1, \ldots, b_n\}$ eine Orthonormalbasis von Eigenvektoren von F bezüglich der Eigenwerte $\lambda_1, \ldots, \lambda_n$ (möglicherweise mit Wiederholungen). Es gibt genau eine lineare Abbildung $G : V \to V$ mit $G(b_i) = \overline{\lambda_i} b_i$ und klarerweise gilt $G \circ F = F \circ G$. Wir müssen also nur noch zeigen, dass $F^t = G$ und somit F normal ist.

Nun ist jeder Basisvektor b_i auch Eigenvektor zu F^t mit Eigenwert $\overline{\lambda_i}$ nach der Proposition 3.35. Somit stimmen F^t und G auf den Basisvektoren überein. Es folgt $F^t = G$, wie gewünscht. □

Wir untersuchen nun eine Sonderkollektion normaler Endomorphismen: Selbstadjungierte Endomorphismen.

Definition 3.37. Ein Endomorphismus $F : V \to V$ des euklidischen oder unitären Raumes $(V, \langle -, - \rangle)$ ist *selbstadjungiert*, wenn $F = F^t$ oder äquivalent dazu, falls

$$\langle v, F(w) \rangle = \langle F(v), w \rangle$$

für alle v und w aus V.

Eine quadratische $n \times n$-Matrix A über K ist *hermitesch*, falls $A^* = A$.

Beachte, dass eine reelle Matrix A genau dann hermitesch ist, wenn sie symmetrisch ist: $A = {}^tA$. Folgende Bemerkung lässt sich leicht mit Hilfe des Korollars 3.30 zeigen.

Bemerkung 3.38. Für endlichdimensionale Vektorräume V ist der Endomorphismus $F : V \to V$ genau dann selbstadjungiert, wenn seine Darstellungsmatrix A bezüglich einer Orthonormalbasis B (sowohl im Definitions- als auch im Bildbereich) hermitesch ist.

Klarerweise sind selbstadjungierte Endomorphismen normal.

Satz 3.39. *Für einen selbstadjungierten Endomorphismus $F : V \to V$ des endlichdimensionalen euklidischen oder unitären Raumes $(V, \langle -, - \rangle)$ gilt:*

(a) Alle Eigenwerte von F sind reell.

(b) Gegeben Eigenvektoren v und w von F zu verschiedenen Eigenwerten λ und μ, sind u und w orthogonal zueinander.

Beweis. Sei λ aus $\mathbb{K}$ ein Eigenwert von F und $v \neq 0_V$ ein Eigenvektor zum Eigenwert λ, also $F(v) = \lambda v$. Nun ist

$$\lambda \|v\|^2 = \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle F(v), v\rangle = \langle F^t(v), v\rangle = \langle v, F(v)\rangle = \langle v, \lambda v\rangle = \bar{\lambda} \|v\|^2$$

und somit liegt $\lambda = \bar{\lambda}$ in $\mathbb{R}$, wie gewünscht.

Für (b) betrachte

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle F(v), w\rangle = \langle F^t(v), w\rangle = \langle v, F(w)\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle.$$

Aus $\lambda \neq \mu$ folgt, dass $\langle v, w\rangle = 0_{\mathbb{K}}$, also sind u und v orthogonal zueinander, wie gewünscht. □

Korollar 3.40. (*Der Spektralsatz*) *Jede symmetrische reelle quadratische Matrize ist diagonalisierbar. Jeder selbstadjungierter Endomorphismus eines endlichdimensionalen euklidischen Raumes ist diagonalisierbar.*

Beweis. Sei $A = {}^tA$ aus $\mathcal{M}_{n\times n}(\mathbb{R})$ und betrachte sie als Matrix über $\mathbb{C}$ mit entsprechender linearer Abbildung $F_A : \mathbb{C}^n \to \mathbb{C}^n$, wobei wir $\mathbb{C}^n$ als unitären Raum betrachten (siehe Beispiel 3.22).

Die Matrix A ist hermitesch, also ist F_A selbstadjungiert und somit normal, nach der Bemerkung 3.38. Das charakteristische Polynom $\chi_{F_A}(T)$ zerfällt in Linearfaktoren über $\mathbb{C}$, wobei der Satz 3.39 besagt, dass alle Eigenwerte von F_A reell sind. Aus dem Satz 3.36 folgt, dass der Vektorraum eine Orthonormalbasis von Eigenvektoren der reellen linearen Abbildung

$$\begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^n \\ (x_1, \ldots, x_n) & \mapsto & A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{array}$$

besitzt, welche somit diagonalisierbar ist, wie gewünscht. □

Kapitel 4

Transformationen der Ebene und des Raumes

In diesem Abschnitt ist $\mathbb{K}$ der Körper $\mathbb{R}$ oder $\mathbb{C}$.

4.1 Hauptachsen

Satz 4.1. (*Hauptachsentransformation*) *Sei $(V, \langle -, - \rangle)$ ein euklidischer oder unitärer Raum der Dimension $n = \dim_{\mathbb{K}} V \geq 1$. Gegeben eine hermitesche Bilinearform $\varphi : V \times V \to \mathbb{K}$, gibt es eine Orthonormalbasis $B = \{b_1, \dots, b_n\}$ von V derart, dass die Darstellungsmatrix von φ bezüglich B in Diagonalform ist.*

Die Basisvektoren der Orthonormalbasis $B = \{b_1, \dots, b_n\}$ heißen *die Hauptachsen* der Bilinearform φ.

Beachte, dass im Fall $\mathbb{K} = \mathbb{R}$ die hermitesche Bilinearform $\varphi : V \times V \to \mathbb{K}$ symmetrisch ist.

Beweis. Wähle eine Orthonormalbasis $\{v_1, \dots, v_n\}$ von V und definiere den Endomorphismus

$$\begin{array}{rccc} F : & V & \to & V \\ & v & \mapsto & \sum\limits_{1 \leq i \leq n} \varphi(v, v_i) v_i \end{array}$$

Wir zeigen zuerst, dass $\langle v, F(w) \rangle = \varphi(v, w)$ für alle v und w aus V. Beachte, dass

$$\langle v_i, F(v_j) \rangle = \langle v_i, \sum_{1 \leq k \leq n} \varphi(v_j, v_k) v_k \rangle = \sum_{1 \leq k \leq n} \overline{\varphi(v_j, v_k)} \langle v_i, v_k \rangle =$$

$$= \sum_{1 \leq k \leq n} \overline{\varphi(v_j, v_k)} \delta_{ik} = \overline{\varphi(v_j, v_i)} = \varphi(v_i, v_j),$$

also folgt für

$$v = \sum_{1 \leq i \leq n} \lambda_i v_i \text{ und } w = \sum_{1 \leq j \leq n} \mu_j v_j$$

aus der Linearität, dass

$$\langle v, F(w)\rangle = \langle \sum_{1\le i\le n} \lambda_i v_i, \sum_{1\le j\le n} \mu_j F(v_j)\rangle = \sum_{1\le i,j\le n} \lambda_i\overline{\mu_j}\langle v_i, F(v_j)\rangle = \sum_{1\le i,j\le n} \lambda_i\overline{\mu_j}\varphi(v_i, v_j) =$$

$$= \sum_{1\le i\le n} \lambda_i\varphi(v_i, \sum_{1\le j\le n} \mu_j v_j) = \sum_{1\le i\le n} \lambda_i\varphi(v_i, w) = \varphi(\sum_{1\le i\le n} \lambda_i v_i, w) = \varphi(v, w).$$

Daraus folgt, dass der Endomorphismus F selbstadjungiert ist:

$$\langle v, F(w)\rangle = \varphi(v, w) = \overline{\varphi(w, v)} = \overline{\langle w, F(v)\rangle} = \langle F(v), w\rangle.$$

Aus dem Satz 3.36 oder aus dem (Beweis von) Korollar 3.40 (für $\mathbb{K} = \mathbb{R}$) schließen wir die Existenz einer Orthonormalbasis $B = \{b_1, \dots, b_n\}$ von Eigenvektoren für F bezüglich der Eigenwerte $\lambda_1, \dots, \lambda_n$ (möglicherweise mit Wiederholungen). Des Weiteren sind die Eigenwerte reell, aus dem Satz 3.39.

Die Darstellungsmatrix der hermiteschen Bilinearform φ bezüglich B ist in Diagonalform, denn

$$\varphi(b_i, b_j) = \langle b_i, F(b_j)\rangle = \langle b_i, \lambda_j b_j\rangle = \lambda_j\langle b_i, b_j\rangle = \lambda_j\delta_{ij},$$

wobei δ_{ij} das Kronecker-Delta ist. Insbesondere sind alle Einträge außerhalb der Diagonale Null. An der (i, i)-ten Stelle ist der Wert die reelle Zahl λ_i für $1 \le i \le n$. $\square$

Beachte, dass der Hauptachsentransformationssatz einen anderen Beweis des Satzes der Signatur von Sylvester B.3 liefert. Wir werden nun einen anderen Satz von Sylvester beweisen, welcher ein Kriterium liefert, wann eine symmetrische Bilinearform positiv definit ist. Hierfür betrachten wir *Hauptminoren* der Darstellungsmatrix. Gegeben eine $n \times n$-Matrix $A = (a_{ij})$ über dem Körper $\mathbb{K}$, ist ein *Hauptminor* der Matrix die Determinante

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk}\end{pmatrix}$$

für eine Zahl $1 \le k \le n$.

Satz 4.2. (*Das Kriterium von Sylvester*) *Folgende Bedingungen sind äquivalent für eine symmetrische Bilinearform $\varphi : V \times V \to \mathbb{R}$ eines endlichdimensionalen euklidischen Raumes V.*

(a) Die Bilinearform ist positiv definit.

(b) Alle Eigenwerte der Darstellungsmatrix A von φ (bezüglich einer beliebigen Basis von V) sind echt positiv.

(c) Alle Hauptminoren der Darstellungsmatrix A sind echt positiv, d.h.

$$\det\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk}\end{pmatrix} > 0 \quad \textit{für alle } 1 \le k \le n.$$

Beweis. $(a) \Longrightarrow (b)$: Mit Hilfe des Satzes 4.1 finden wir eine Orthonormalbasis $B = \{b_1, \dots, b_n\}$ derart, dass die Darstellungsmatrix von φ bezüglich der Basis B in Diagonalform ist. Die Eigenwerte der Darstellungsmatrix A sind gerade die Einträge $\lambda_i = \varphi(b_i, b_i)$. Wegen $b_i \neq 0_V$ ist $\lambda_i > 0$, weil φ positiv definit ist.

$(b) \Longrightarrow (a)$: Weil beide Bedingungen unabhängig von der Wahl der Basis sind, können wir φ bezüglich der obigen Orthonormalbasis $B = \{b_1, \dots, b_n\}$ berechnen. Die Eigenwerte $\lambda_i = \varphi(b_i, b_i)$ sind alle echt positiv. Gegeben einen beliebigen Vektor v aus V, schreibe

$$v = \sum_{1 \leq i \leq n} \mu_i b_i.$$

Es folgt, dass

$$\varphi(v, v) = \sum_{1 \leq i \leq n} \mu_i^2 \lambda_i \geq 0.$$

Ferner ist $\varphi(v, v) = 0$ genau dann, wenn alle Einträge $\mu_1 = \ldots = \mu_n = 0$ sind und somit $v = 0_V$, wie gewünscht.

$(a) \Longrightarrow (c)$: Sei $\{v_1, \dots, v_n\}$ eine Basis von V derart, dass die entsprechende Darstellungsmatrix von φ die Matrix A ist. Für $1 \leq k \leq n$, setze $U_k = \operatorname{Lin}(\{v_1, \dots, v_k\})$.

Die Einschränkung der Bilinearform auf U_k hat die Teilmatrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

als Darstellungsmatrix. Aus der Richtung $(a) \Longrightarrow (b)$ folgt, dass die Einschränkung $\varphi_{\restriction U_k \times U_k}$ auch positiv definit ist. Also sind die Eigenwerte der obigen Teilmatrix, welche diagonalisierbar ist, nach dem Spektralsatz 3.40, echt positiv. Somit ist die Determinante auch echt positiv, weil sie das Produkt der Diagonaleinträge der ähnlichen Diagonalmatrix ist.

$(c) \Longrightarrow (b)$: Wir beweisen es induktiv über $n = \dim_{\mathbb{K}} V$, wobei der Fall $n = 1$ trivial ist. Wir nehmen also $n \geq 2$ an.

Beachte, dass die Determinante $\det(A) > 0$ und somit kein Eigenwert der Matrix A, welche diagonalisierbar ist, Null sein kann. Sei $B = \{b_1, \dots, b_n\}$ die obige Orthonormalbasis mit $\lambda_i = \varphi(b_i, b_i)$. Wir nehmen nun an, dass nicht alle Eigenwerte λ_i echt positiv sind. Dann muss es also mindestens zwei Indizes $i \neq j$ derart geben, dass λ_i und λ_j beide echt negativ sind. Schreibe b_i und b_j bezüglich der Basis $\{v_1, \dots, v_n\}$, also

$$b_i = \sum_{1 \leq k \leq n} \mu_k(i) v_k \text{ und } b_j = \sum_{1 \leq k \leq n} \mu_k(j) v_k$$

und setze

$$(\alpha, \beta) = \begin{cases} (\mu_n(j), -\mu_n(i)), & \text{falls } \mu_n(i)^2 + \mu_n(j)^2 > 0 \\ (1, 1), & \text{sonst.} \end{cases}$$

Die Linearkombination $0 \neq w = \alpha b_i + \beta b_j$ liegt im Unterraum $U = \operatorname{Lin}(\{v_1, \dots, v_{n-1}\})$ und die Einschränkung der Bilinearform auf U hat als Darstellungsmatrix die Teilmatrix

$$\begin{pmatrix} a_{11} & \cdots & a_{1\,n-1} \\ \vdots & \ddots & \vdots \\ a_{n-1\,1} & \cdots & a_{n-1\,n-1} \end{pmatrix}.$$

Alle Hauptminoren dieser Teilmatrix sind echt positiv und somit ist $\varphi_{\restriction U \times U}$ positiv definit, nach unserer Induktionsannahme. Nun ist

$$0 < \varphi(w,w) = \alpha^2 \varphi(b_i, b_i) + \beta^2 \varphi(b_j, b_j) + 2\alpha\beta\varphi(b_i, b_j) = \alpha^2 \lambda_i + \beta^2 \lambda_j + 0 < 0,$$

was den gewünschten Widerspruch liefert. □

4.2 Orthogonale Abbildungen

Im Folgenden ist $\mathbb{K}$ entweder $\mathbb{R}$ oder $\mathbb{C}$ und $(V, \langle -,- \rangle)$ ein n-dimensionaler euklidischer oder unitärer Raum $(V, \langle -,- \rangle)$.

Definition 4.3. Ein Endomorphismus $F : V \to V$ ist *orthogonal*, falls

$$\langle v, w \rangle = \langle F(v), F(w) \rangle$$

für alle Vektoren v und w aus V.

In der Literatur werden orthogonale Abbildungen unitärer Räume auch *unitär* genannt.

Lemma 4.4. *Folgende Bedingungen für den Endomorphismus $F : V \to V$ sind äquivalent:*

(a) Die Abbildung F ist orthogonal.

(b) Der Endomorphimus F bildet jedes orthonormale System zu einem orthonormalen System ab.

(c) Die Norm bleibt unter F erhalten, d.h.

$$\|F(v)\| = \|v\|$$

für alle Vektoren v aus V.

Beweis. $(a) \Longrightarrow (b)$: Es ist offensichtlich, dass F Orthogonalität erhält. Ferner ist das Bild eines normierten Vektors wiederum normiert, denn

$$\|v\|^2 = \langle v, v \rangle \overset{F \text{ orth.}}{=} \langle F(v), F(v) \rangle = \|F(v)\|^2 .$$

$(b) \Longrightarrow (c)$: Sei v aus V beliebig. Fall $v = 0_V$, gilt die Gleichheit $\|F(v)\| = 0_V = \|v\|$ trivialerweise. Wenn $v \neq 0_V$, ist $\|v\| \neq 0_{\mathbb{K}}$ und somit bestimmt der Vektor $\|v\|^{-1} v$ ein orthonormales System der Größe 1. Das Bild davon muss also normiert sein, also $\left\|F(\|v\|^{-1} v)\right\| = 1_{\mathbb{K}}$ oder äquivalent dazu $\|F(v)\| = \|v\|$, wie gewünscht.

$(c) \Longrightarrow (a)$: Gegeben zwei Vektoren v und w aus V, müssen wir zeigen, dass

$$\langle v, w \rangle = \langle F(v), F(w) \rangle.$$

Wir beweisen die Behauptung mit Fallunterscheidung: Falls die Vektoren linear abhängig sind, können wir ohne Beschränkung der Allgemeinheit annehmen, dass $v \neq 0_V$ und $w = \lambda v$ für ein λ aus $\mathbb{K}$. Nun ist

$$\langle v, w \rangle = \bar{\lambda}\langle v, v \rangle = \bar{\lambda}\,\|v\|^2 = \bar{\lambda}\,\|F(v)\|^2 = \bar{\lambda}\langle F(v), F(v) \rangle = \langle F(v), \lambda F(v) \rangle = \langle F(v), F(w) \rangle,$$

wie gewünscht.

Falls v und w linear unabhängig sind, finde mit Hilfe des Gram-Schmidt'schen Orthogonalisierungsverfahren 3.14 eine Orthonormalbasis $\{b_1, b_2\}$ des Unterraumes $\mathrm{Lin}(\{v, w\})$. Dann ist

$$v = \lambda_1 b_1 + \lambda_2 b_2 \text{ und } w = \mu_1 b_1 + \mu_2 b_2$$

für Skalare λ_i, μ_i aus $\mathbb{K}$ mit $i = 1, 2$ und es gilt

$$\langle v, w\rangle = \lambda_1\mu_1 + \overline{\lambda_2\mu_2}.$$

Nun ist

$$F(v) = \lambda_1 F(b_1) + \lambda_2 F(b_2) \text{ und } F(w) = \mu_1 F(b_1) + \mu_2 F(b_2).$$

Wenn wir zeigen können, dass $F(b_1)$ und $F(b_2)$ wiederum orthogonal zueinander sind, ist F orthogonal, denn

$$\begin{aligned}\langle F(v), F(w)\rangle &= \lambda_1\mu_1 \|F(b_1)\|^2 + \overline{\lambda_2\mu_2}\, \|F(b_2)\|^2 + \big(\lambda_1\overline{\mu_2}\langle F(b_1), F(b_2)\rangle + \lambda_2\overline{\mu_1}\langle F(b_2), F(b_1)\rangle\big) = \\ &= \lambda_1\mu_1 + \overline{\lambda_2\mu_2} + \big(\lambda_1\overline{\mu_2}\langle F(b_1), F(b_2)\rangle + \lambda_2\overline{\mu_1}\langle F(b_2), F(b_1)\rangle\big) = \\ &= \langle v, w\rangle + 0_{\mathbb{K}} = \langle v, w\rangle.\end{aligned}$$

Nach dem Satz von Pythagoras 3.11 ist für jedes ξ aus $\mathbb{K}$ mit $|\xi| = 1$

$$\begin{aligned}2 &= \|b_1\|^2 + \|\xi b_2\|^2 = \|b_1 + \xi b_2\|^2 \overset{(c)}{=} \|F(b_1 + \xi b_2)\|^2 = \langle F(b_1) + \xi F(b_2), F(b_1) + \xi F(b_2)\rangle = \\ &= \|F(b_1)\|^2 + |\xi|^2 \|F(b_2)\|^2 + \big(\xi\langle F(b_2), F(b_1)\rangle + \bar{\xi}\langle F(b_1), F(b_2)\rangle\big) \overset{(c)}{=} \\ &= 2 + \big(\xi\langle F(b_2), F(b_1)\rangle + \bar{\xi}\langle F(b_1), F(b_2)\rangle\big).\end{aligned}$$

Es folgt also, dass $\xi\langle F(b_2), F(b_1)\rangle = -\overline{\xi\langle F(b_2), F(b_1)\rangle}$ für alle ξ aus $\mathbb{K}$ mit $|\xi| = 1$. Im Fall $\mathbb{K} = \mathbb{R}$ folgt (mit $\xi = 1$), dass $\langle F(b_2), F(b_1)\rangle = 0$ und somit sind $F(b_1)$ und $F(b_2)$ orthogonal zueinander, wie gewünscht. Falls $\mathbb{K} = \mathbb{C}$, setze $\xi = i$ (die komplexe Zahl i, kein Index!), also

$$i\langle F(b_2), F(b_1)\rangle = -\overline{i\langle F(b_2), F(b_1)\rangle} = i\overline{\langle F(b_2), F(b_1)\rangle}.$$

Insbesondere ist $\langle F(b_2), F(b_1)\rangle = \overline{\langle F(b_2), F(b_1)\rangle}$ und somit ist das Skalarprodukt $\langle F(b_2), F(b_1)\rangle$ eine reelle Zahl. Wir können nun mit $\xi = 1$ wie oben schließen, dass $F(b_1)$ und $F(b_2)$ orthogonal zueinander sind. □

Weil die Norm eines nicht-trivialen Vektors verschieden von $0_{\mathbb{K}}$ ist, schließen wir aus dem obigen, dass orthogonale Abbildungen $F : V \to V$ immer Isomorphismen sind.

Korollar 4.5. *Orthogonale Endomorphismen sind injektiv und somit bijektiv.*

Proposition 4.6. *Sei $F : V \to V$ ein Automorphismus des euklidischen oder unitären Raumes $(V, \langle -, -\rangle)$. Die Abbildung F ist genau dann orthogonal, wenn der zu F adjungierte Endomorphismus F^t genau die inverse Abbildung F^{-1} ist.*

Insbesondere sind orthogonale Abbildungen normal.

Beweis. Die letzte Behauptung folgt klarerweise aus der Äquivalenz in der Proposition, weil F und F^{-1} miteinander kommutieren. Angenommen, dass F orthogonal ist, müssen wir nur zeigen, dass $F^{-1} : V \to V$ die definitorische Eigenschaft des adjungierten Endomorphismus erfüllt:

$$\langle v, F(w) \rangle = \langle F(F^{-1}(v)), F(w) \rangle \overset{F \text{ orth.}}{=} \langle F^{-1}(v), w \rangle.$$

Angenommen nun, dass F^{-1} und F^t gleich sind, folgt für v und w aus V beliebig, dass

$$\langle F(v), F(w) \rangle \overset{\text{Def. von } F^t}{=} \langle F^t \circ F(v), w \rangle = \langle F^{-1} \circ F(v), w \rangle = \langle v, w \rangle.$$

Somit ist F orthogonal, wie gewünscht. □

Definition 4.7. Eine reguläre Matrix A aus $\mathrm{GL}_n(\mathbb{K})$ ist *orthogonal*, falls ihre adjungierte Matrix A^* die inverse Matrix A^{-1} ist.

Aus dem Korollar 3.30 folgt leicht die nächste Behauptung.

Korollar 4.8. *Der Automorphismus $F : V \to V$ ist genau dann* orthogonal*, wenn seine Darstellungsmatrix bezüglich einer Orthonormalbasis von V* orthogonal *ist.*

Insbesondere ist eine reguläre Matrix genau dann orthogonal, wenn ihre Inverse auch orthogonal ist.

Lemma 4.9. *Eine quadratische $n \times n$-Matrix A über $\mathbb{K}$ ist genau dann orthogonal, wenn ihre Spaltenvektoren, beziehungsweise ihre Zeilenvektoren, ein orthonormales System, und somit eine Orthonormalbasis, von $\mathbb{K}^n$ bilden.*

Beweis. Beachte, dass

$$A \cdot A^* = \mathbf{Id}_n \qquad \iff \qquad \sum_{1 \le k \le n} a_{ik} \overline{a_{jk}} = \delta_{ij},$$

wobei δ_{ij} das Kronecker-Delta ist. Insbesondere ist A regulär und $A^* = A^{-1}$ genau dann, wenn die Zeilenvektoren $a_1, \ldots, a_n$ mit $a_i = (a_{i1}, \ldots, a_{in})$ ein orthonormales System bilden.

Analog ist

$$A^* \cdot A = \mathbf{Id}_n \qquad \iff \qquad \sum_{1 \le k \le n} \overline{a_{ki}} a_{kj} = \delta_{ij}.$$

Somit ist A genau dann orthogonal, wenn die Spaltenvektoren $a^1, \ldots, a^n$, wobei $a^j = (a_{1j}, \ldots, a_{nj})$, ein orthonormales System bilden. □

Definition 4.10. Eine Matrix A aus $\mathcal{M}_{n \times n}(\mathbb{K})$ ist *orthogonal diagonalisierbar*, falls es eine Orthogonalmatrix S so gibt, dass $S^{-1}AS$ in Diagonalform ist.

Es folgt aus dem Beweis des Satzes 3.36, sowie aus der Hauptachsentransformation 4.1, dass die Darstellungsmatrizen normaler Endomorphismen und hermitescher Bilinearformen orthogonal diagonalisierbar sind, sobald das charakteristische Polynom in Linearfaktoren zerfällt.

Korollar 4.11. *Jede normale Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt, ist orthogonal diagonalisierbar.*

Proposition 4.12. *Sei $F : V \to V$ ein orthogonaler Automorphismus des euklidischen oder unitären Raumes $(V, \langle -, - \rangle)$. Alle Eigenwerte von F haben Absolutbetrag 1. Des Weiteren ist $|\det(F)| = 1$.*

Beweis. Sei λ ein Eigenwert von F und $v \neq 0_V$ ein Eigenvektor von F bezüglich λ. Aus dem Lemma 4.4 folgt
$$\|v\|^2 = \|F(v)\|^2 = \|\lambda v\|^2 = |\lambda|^2 \|v\|^2 .$$
Insbesondere ist $|\lambda|^2 = 1$ und somit muss der Absolutbetrag $|\lambda| = 1$ sein.

Für die letzte Behauptung genügt es nach dem Korollar 4.8 zu zeigen, dass orthogonale reguläre Matrizen Determinante (im Absolutbetrag) genau 1 haben. Sei $A = (a_{ij})$ eine orthogonale quadratische Matrix, also $A^* = {}^t\bar{A} = A^{-1}$. Aus der Produktformel folgt
$$1_{\mathbb{K}} = \det(\mathbf{Id}_n) = \det(A) \cdot \det(A^*) = \det(A) \cdot \det(\bar{A}) = |\det(A)|^2,$$
denn
$$\det(\bar{A}) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{1 \leq k \leq n} \overline{a_{k\sigma(k)}} = \overline{\det(A)},$$
wegen der Leibniz-Formel für Determinanten. $\square$

Wir werden nun eine partielle Rückrichtung des obigen Lemmas zeigen, unter der Annahme, dass das charakteristische Polynom des normalen Endomorphismus in Linearfaktoren zerfällt, was für Endomorphismen unitärer Räume immer der Fall ist.

Korollar 4.13. *Sei $F : V \to V$ ein normaler Endomorphimus des endlichdimensionalen euklidischen oder unitären Raumes $(V, \langle -, - \rangle)$ derart, dass das charakteristische Polynom in Linearfaktoren zerfällt. Wenn alle Eigenwerte von F Absolutbetrag* 1 *besitzen, ist F orthogonal.*

Beweis. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass die Dimension $\dim_{\mathbb{K}} V = n \geq 1$. Nach dem Satz 3.36 gibt es eine Orthonormalbasis $B = \{b_1, \ldots, b_n\}$ von V, welche aus Eigenvektoren von F bezüglich der Eigenwerte $\lambda_1, \ldots, \lambda_n$ (möglicherweise mit Wiederholungen) besteht. Die Darstellungsmatrix von F ist bezüglich der Basis $B = \{b_1, \ldots, b_n\}$ in Diagonalform
$$A = \begin{pmatrix} \lambda_1 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix}$$
Nach Annahme ist ihre Determinante nicht Null ist, also ist A invertierbar. Beachte, dass
$$A^* = \begin{pmatrix} \overline{\lambda_1} & \cdots & \cdots & 0 \\ 0 & \overline{\lambda_2} & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \overline{\lambda_n} \end{pmatrix} = \begin{pmatrix} \lambda_1^{-1} & \cdots & \cdots & 0 \\ 0 & \lambda_2^{-1} & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_n^{-1} \end{pmatrix}$$
weil $z^{-1} = |z|^{-2}\bar{z}$ für komplexe Zahlen z verschieden von Null (wenn der Grundkörper $\mathbb{K} = \mathbb{R}$ ist, gilt $z^{-1} = z^{-2}z = |z|^{-2}\bar{z}$ auch). Die letzte Matrix ist genau die Inverse A^{-1} und somit ist A orthogonal. Aus dem Korollar 4.8 folgt nun, dass F orthogonal ist, wie gewünscht. $\square$

Wir werden nun eine Sondergruppe von Automorphismen in $\mathrm{GL}(V)$ untersuchen, die Gruppe der *Drehungen*, und geben eine geometrische Interpretation der Hauptachsen dieser Transformationen für den euklidischen Raum und die euklidische Ebene. Im Folgenden ist $(V, \langle -, - \rangle)$ ein endlichdimensionaler euklidischer Raum.

Definition 4.14. Eine orthogonale Abbildung $F : V \to V$ eines euklidischen Raumes ist eine *Drehung*, falls die Determinante $\det(F) = 1$.

Aus der Produktformel folgt, dass die Komposition von Drehungen wiederum eine Drehung ist. Des Weiteren ist das Inverse einer Drehung auch eine Drehung. Somit bilden die Drehungen eine Untergruppe von $\mathrm{GL}(V)$.

Satz 4.15. *Betrachte einen zweidimensionalen euklidischen Raum $(V, \langle -, - \rangle)$ mit einer festen Orthonormalbasis $\{b_1, b_2\}$. Die Abbildung $F : V \to V$ ist genau dann eine Drehung, wenn die Darstellungsmatrix von F bezüglich der Orthonormalbasis $\{b_1, b_2\}$ von der Form*

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

ist, für einen Wert α aus $[0, 2\pi)$, genannt den Drehwinkel der Drehung *(bezüglich der Orientierung im Gegenuhrzeigersinn).*

Beweis. Aus dem Korollar 4.8 und dem Lemma 4.9 folgt, dass solche Matrizen orthogonal sind und Drehungen definieren. Wir nehmen also an, dass F eine Drehung definiert und A die Darstellungsmatrix von F bezüglich der Orthonormalbasis $\{b_1, b_2\}$ von V ist. Wegen der Bemerkung 3.13 ist $A = (a_{ij})$ mit $a_{ij} = \langle F(b_j), b_i \rangle$. Weil F orthogonal ist, muss A auch orthogonal sein. Aus dem Lemma 4.9 folgt, dass die erste Spalte normiert ist

$$\langle F(b_1), b_1 \rangle^2 + \langle F(b_1), b_2 \rangle^2 = 1$$

und es einen eindeutigen Wert α aus $[0, 2\pi)$ mit

$$\cos\alpha = \langle F(b_1), b_1 \rangle \text{ und } \sin\alpha = \langle F(b_1), b_2 \rangle$$

gibt. Weil beide Spalten orthogonal zueinander sind, folgt

$$\cos\alpha \langle F(b_2), b_1 \rangle + \sin\alpha \langle F(b_2), b_2 \rangle = 0.$$

Des Weiteren ist

$$\det(A) = 1 = \cos\alpha \langle F(b_2), b_2 \rangle - \sin\alpha \langle F(b_2), b_1 \rangle.$$

Schreiben wir beide Gleichung in Matrizenform, erhalten wir

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \langle F(b_2), b_1 \rangle \\ \langle F(b_2), b_2 \rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Insbesondere ist

$$\begin{pmatrix} \langle F(b_2), b_1 \rangle \\ \langle F(b_2), b_2 \rangle \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix},$$

wie gewünscht. □

Satz 4.16. *Betrachte den euklidischen Raum $\mathbb{R}^3$ mit dem Standardskalarprodukt (siehe Beispiel 3.4). Die Abbildung $F : \mathbb{R}^3 \to R^3$ ist genau dann eine Drehung, wenn es eine Orthonormalbasis $B = \{b_1, b_2, b_3\}$ von $\mathbb{R}^3$ so gibt, dass die Darstellungsmatrix von F bezüglich B von der Form*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

ist, für einen Wert α aus $[0, 2\pi)$. Der Unterraum $\mathrm{Lin}(\{b_1\})$ *heißt* die Drehachse *und der Wert α heißt der* Drehwinkel der Drehung.

Beweis. Analog zum vorigen Beweis müssen wir nur zeigen, dass sich jede Drehung so darstellen lässt. Sei A die Darstellungsmatrix der Drehung F bezüglich der kanonischen Basis $\{e_1, \ldots, e_3\}$ von $\mathbb{R}^3$. Schreibe

$$\chi_F = \det(T \cdot \mathbf{Id}_3 - A) = T^3 + c_2 + c_1 T + c_0,$$

mit $c_0 = (-1)^3 \det(F) = -1$.

Jedes nicht-triviale Polynom über $\mathbb{R}$ von Grad 3 besitzt mindestens eine reelle Nullstelle. Seien also λ_1, λ_2 und λ_3 die Nullstellen (im algebraischen abgeschlossenen Körper $\mathbb{C}$) von χ_F mit λ_1 in $\mathbb{R}$. Weil F orthogonal ist, müssen alle Eigenwerte Absolutbetrag 1 haben, also $\lambda_1 = \pm 1$.

Wir zeigen zuerst, dass einer der Eigenwerte genau 1 sein muss (ein normierter Eigenvektor in dieser Richtung bestimmt dann die Drehachse). Weil

$$\lambda_1 \cdot \lambda_2 \cdot \lambda_3 = \det(F) = 1 = \bar{1} = \lambda_1 \cdot \overline{\lambda_2} \cdot \overline{\lambda_3},$$

folgt, dass $\lambda_2 = \overline{\lambda_2}$ und $\lambda_3 = \overline{\lambda_3}$ beide in $\mathbb{R}$ liegen, oder $\lambda_2 = \overline{\lambda_3} = \lambda_3^{-1}$. Im ersten Fall sind alle $\lambda_j = \pm 1$ und somit muss einer der Werte 1 sein, weil das Produkt aller drei Elemente $\det(F) = 1$ ist. Im zweiten Fall ist

$$\det(F) = 1 = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 = \lambda_1$$

und somit ein Eigenwert 1, wie gewünscht.

Ohne Beschränkung der Allgemeinheit können wir also annehmen, dass der reelle Eigenwert $\lambda_1 = 1$ ist. Wähle nun einen normierten Eigenvektor b_1 von F bezüglich λ_1. Das orthogonale Komplement U von $\mathrm{Lin}(\{b_1\})$ hat Dimension 2, nach der Proposition 3.18. Beachte, dass b_1 auch ein Eigenvektor von $F^{-1} = F^t$ ist aus der Proposition 3.35, so U ist dann invariant unter dem Automorphismus F, aus der Aufgabe nach der Definition 3.26.

Wähle also mit Hilfe des Gramm-Schmidt'schen Orthonormalisierungsverfahren 3.14 eine Orthonormalbasis $\{b_2, b_3\}$ von U. die Darstellungsmatrix von F bezüglich $\{b_1, \ldots, b_3\}$ ist von der Form

$$\begin{pmatrix} 1 & 0 \quad 0 \\ 0 & \\ 0 & \raisebox{0.5em}{A_1} \end{pmatrix}$$

wobei A_1 die Darstellungsmatrix der Einschränkungsabbildung $F_{\restriction U}$ ist. Weil $F_{\restriction U}$ auch orthogonal ist, folgt aus der Kästchenregel, dass $F_{\restriction U}$ eine Drehung ist. Der zweidimensionaler Raum U mit der Orthonormalbasis $\{b_2, b_3\}$ ist klarerweise euklidisch. Nach dem vorigen Satz 4.15 hat A_1 die gewünschte Form. $\square$

Appendix

A Polynomringe

Sei $\mathbb{K}$ ein Körper.

Definition A.1. Der *Polynomring über* $\mathbb{K}$ *in der Variablen* T ist die Kollektion $\mathbb{K}[T]$ von Ausdrücken der Form

$$P = a_0 + a_1 T + \cdots + a_n T^n,$$

wobei n eine (beliebige) natürliche Zahlen ist und jeder Koeffizient a_i in $\mathbb{K}$ liegt. Das *Nullpolynom (oder das triviale Polynom)* 0 ist das Polynom, dessen Koeffizienten alle Null sind. Jedes nicht-triviale Polynom lässt sich eindeutig als

$$P = a_0 + a_1 T + \cdots + a_n T^n$$

schreiben, mit $a_n \neq 0$ für eine natürliche Zahl $n = \deg(P)$, genannt den *Grad* von P. Als Konvention setzen wir $\deg(0) = -\infty$.

Wenn P Grad n hat, ist der Koeffizient a_n in der Darstellung von P der *Führungskoeffizient* von P. Das Polynom P ist *normiert*, falls der Führungskoeffizient $1_{\mathbb{K}}$ ist.

Bemerkung A.2. Jedes Element λ von $\mathbb{K}$ lässt sich als *konstantes Polynom* vom Grad 0 auffassen.

Auf dem Polynomring können wir folgenderweise eine Summe definieren: Gegeben Polynome

$$P = \sum_{i=1}^{n} a_i T^i \text{ und } Q = \sum_{j=1}^{m} b_j T^j,$$

können wir annehmen, dass $n = m$ ist (für $d = \max(n, m)$ setze für $n < i \leq d$ und $m < j \leq d$ die neuen Koeffizienten $a_i = b_j = 0$). Dann ist

$$P + Q = \sum_{i=1}^{n} (a_i + b_i) T^i.$$

Analog definieren wir folgendermaßen eine Multiplikation auf $\mathbb{K}[T]$:

$$P \cdot Q = \sum_{k=1}^{n+m} c_k T^k \text{ mit } c_k = \sum_{i+j=k} a_i b_j.$$

Es lässt sich leicht zeigen, dass $\mathbb{K}[T]$ mit diesen beiden Verknüpfungen ein kommutativer Ring mit Eins ist und eine kompatible Struktur als $\mathbb{K}$-Vektorraum derart besitzt, dass für alle λ aus $\mathbb{K}$ sowie Polynome P und Q aus $\mathbb{K}[T]$ gilt:

$$\lambda(P \cdot Q) = (\lambda P) \cdot Q = P(\lambda Q).$$

Solche Ringe heißen *kommutative* $\mathbb{K}$*-Algebren.*

Beachte, dass der Polynomring $\mathbb{K}[T]$ ein Integritätsbereich ist: Wenn P und Q beide nicht trivial sind, dann ist $PQ \neq 0$, da

$$c_{\deg(P)+\deg(Q)} = a_{\deg(P)} b_{\deg(Q)} \neq 0_{\mathbb{K}},$$

weil der Körper $\mathbb{K}$ ein Integritätsbereich ist.

Bemerkung A.3. Der Polynomring über $\mathbb{K}$ in einer Variablen kann leicht in folgender Weise konstruiert werden. Betrachte die Menge $\mathcal{I}$ aller abzählbaren Folgen $(a_n)_{n\in\mathbb{N}}$ aus $\mathbb{K}$ derart, dass alle bis auf endlich viele a_n Null sind.

Die Folge $(a_n)_{n\in\mathbb{N}}$ aus $\mathcal{I}$ ist eindeutig in Korrespondenz mit dem Polynom

$$P(T) = a_0 + a_1 T + \cdots$$

Beachte, dass wegen der Definition von $\mathcal{I}$ der obige Ausdruck in der Tat ein Polynom ist. Die Summe von Polynomen ist in Korrespondenz mit der koordinatenweisen Summe von Folgen. Allerdings entspricht das Produkt von Polynomen nicht dem koordinatenweisen Produkt von Folgen. Die Nullfolge $(0)_{n\in\mathbb{N}}$ stellt das triviale Polynom dar.

Der Polynomring ist bis auf $\mathbb{K}$-Algebra-Isomorphismus eindeutig bestimmt. Daher reden wir vom Polynomring anstatt von einem Polynomring.

Satz A.4. (*Division mit Rest*) *Gegeben Polynome P und Q mit $\deg(Q) > 0$, existieren eindeutig bestimmte Polynome H und R mit*

$$P = HQ + R \text{ und } \deg(R) < \deg(Q).$$

Das Polynom H ist der *Quotient* und das Polynom R der *Rest* der Division mit Rest von P durch Q. Wenn $R = 0$ ist, *teilt* Q das Polynom P. Das Polynom Q ist ein *echter Teiler (oder Faktor)* von P, falls Q das Polynom P teilt und $0 < \deg(Q) < \deg(P)$.

Beweis. **Existenz:** Wenn $\deg(P)$ in der Menge $\{-\infty, 0, 1, \ldots, \deg(Q) - 2\}$ liegt, setze $H = 0$ und $R = P$. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $\deg(P) = \deg(Q) - 1 + k$ für eine natürliche Zahl k. Wir beweisen die Existenz des Quotienten und des Restes induktiv über k. Für $k = 0$ setze $H = 0$ und $R = P$. Für $k > 0$ schreibe

$$P = \sum_{i=1}^{\deg(P)} a_i T^i \text{ und } Q = \sum_{j=1}^{\deg(Q)} b_j T^j,$$

mit $a_{\deg(P)} \neq 0_{\mathbb{K}}$ und $b_{\deg(Q)} \neq 0_{\mathbb{K}}$. Setze nun

$$P' = P - a_{\deg(P)} b_{\deg(Q)}^{-1} T^{k-1} Q = c_0 + c_1 T + \cdots + c_{\deg(P)-1} T^{\deg(P)-1}$$

für gewisse Elemente c_i aus $\mathbb{K}$ (es wird nicht behauptet, dass $c_{\deg(P)-1} \neq 0_{\mathbb{K}}$). Beachte, dass $\deg(P') \leq \deg(P) - 1 = \deg(Q) - 1 + k - 1$. Wegen der Induktionsannahme gibt es H' und R' mit $\deg(R') < \deg(Q)$ und

$$P - a_{\deg(P)} b_{\deg(Q)}^{-1} T^k Q = H'Q + R',$$

also

$$P = \Big(a_{\deg(P)} b_{\deg(Q)}^{-1} T^k + H'\Big) Q + R',$$

wie gewünscht.

Eindeutigkeit: Wir nehmen an, dass $P = QH_1 + R_1 = QH_2 + R_2$ mit $\deg(R_1), \deg(R_2) < \deg(Q)$. Dann gilt

$$R_1 - R_2 = Q(H_2 - H_1).$$

Weil der Grad von $R_1 - R_2$ echt kleiner als $\deg(Q)$ ist, muss das Polynom $H_2 - H_1$ trivial sein. Das heißt, dass $H_1 = H_2$ und somit $R_1 = R_2$, wie gewünscht. $\square$

Bemerkung A.5. Jedes Polynom $P = \sum_{i=0}^{n} a_i T^i$ definiert in folgender Weise eine Abbildung auf $\mathbb{K}$:

$$\begin{array}{rccc} P: & \mathbb{K} & \to & \mathbb{K} \\ & c & \mapsto & \sum_{i=0}^{n} a_i c^i \end{array}$$

Das Element c aus $\mathbb{K}$ ist eine *Nullstelle* von P, falls $P(c) = 0$. Zum Beispiel besitzt das Polynom $T^2 - 3$ zwei Nullstellen in $\mathbb{R}$, aber das Polynom $T^2 + 1$ hat keine Nullstellen in $\mathbb{R}$.

Für das triviale Polynom ist jedes Element aus $\mathbb{K}$ eine Nullstelle, aber ein konstantes nicht-triviales Polynom besitzt keine Nullstelle.

Korollar A.6. *Gegeben ein Element c aus $\mathbb{K}$, lässt sich jedes nicht-konstante Polynom P eindeutig als*

$$P = (T-c)^k H + P(c)$$

schreiben, für ein Polynom H mit $H(c) \neq 0_{\mathbb{K}}$ und eine natürliche Zahl k. Insbesondere ist $k \geq 1$ und

$$P = (T-c)^k H,$$

wenn c eine Nullstelle von P ist.

Die Zahl $k = \operatorname{ord}_c(P)$ heißt die *Vielfachheit* der Nullstelle c.

Beweis. Wir zeigen zuerst, dass eine solche Darstellung eindeutig ist: Angenommen, dass

$$P = (T-c)^k H + P(c) = (T-c)^\ell H' + P(c),$$

mit $k \neq \ell$, können wir ohne Beschränkung der Allgemeinheit annehmen, dass $k < \ell$, also

$$(T-c)^k H = (T-c)^{k+(\ell-k)} H' = (T-c)^k (T-c)^{\ell-k} H'.$$

Weil der Polynomring ein Integritätsbereich ist, folgt aus dem Lemma **??**, dass

$$H = (T-c)^{\ell-k} H'.$$

Weil $H(c) \neq 0_{\mathbb{K}}$, jedoch $\ell - k \geq 1$, liefert dies den gewünschten Widerspruch.

Die Existenz wird induktiv über $\deg(P)$ bewiesen. Weil P nicht konstant ist, ist $\deg(P)$ eine positive natürliche Zahl. Wir wenden nun Division mit Rest A.4 für $Q = T - c$ (welches nicht-trivial ist) an. Also

$$P = (T-c)P_1 + R,$$

wobei $\deg(R) < \deg(T-c) = 1$. Dies bedeutet, dass $R = b$ ein konstantes Polynom ist (möglicherweise ist der Rest R das triviale Polynom). Wenn wir in die obige Gleichung c einsetzen, erhalten wir $P(c) = R(c) = b$.

Wir machen nun eine Fallunterscheidung: Falls $P_1(c) \neq 0_{\mathbb{K}}$, setze $H = P_1$ und $k = 1$. Wenn $P_1(c) = 0_{\mathbb{K}}$, beachte, dass P_1 nicht konstant sein kann, denn sonst wäre $P_1 = 0_{\mathbb{K}[T]}$ und somit wäre $P = P(c)$ konstant. Weil

$$\deg(P) = \deg\big((T-c)P_1 + R\big) = \max(\deg((T-c)P_1), 0) = \deg\big((T-c)P_1\big) = 1 + \deg(P_1),$$

schreibe nun induktiv P_1 als

$$P_1 = (T-c)^\ell H + P_1(c) = (T-c)^\ell H,$$

mit $H(c) \neq 0_{\mathbb{K}}$. Insbesondere ist

$$P = (T-c)P_1 + P(c) = (T-c)\big((T-c)^{\ell} H\big) + P(c) = (T-c)^{\ell+1} H + P(c),$$

also setze $k = \ell + 1$. □

Korollar A.7. *Jedes nicht-triviale Polynom P über $\mathbb{K}$ lässt sich (bis auf Permutation) eindeutig schreiben als*

$$P = (T-c_1)\cdots(T-c_k)P_0,$$

für eine natürliche Zahl $0 \leq k \leq \deg(P)$ sowie Elemente $c_1, \ldots, c_k$ aus $\mathbb{K}$ (möglicherweise mit Wiederholungen) und ein Polynom P_0, das keine Nullstelle in $\mathbb{K}$ besitzt.

Insbesondere besitzt ein nicht-triviales Polynom höchstens $\deg(P)$ viele Nullstellen im Körper $\mathbb{K}$.

Beweis. Da jede Nullstelle von P eines der Elemente c_i sein muss, folgt die zweite Behauptung sofort aus der obigen Darstellung. Wir beweisen die Existenz einer solchen Darstellung wie oben induktiv über $\deg(P)$. Wenn $\deg(P) = 0$, ist das Polynom P konstant und besitzt keine Nullstelle in $\mathbb{K}$. Setze also $k = 0$ und $P_0 = P$. Wir nehmen nun an, dass $\deg(P) \geq 1$. Wenn P keine Nullstelle in $\mathbb{K}$ besitzt, sind wir fertig: setze $k = 0$ und $P_0 = P$. Sei also c_1 eine Nullstelle von P. Mit Hilfe des Korollars A.6 schreiben wir $P = (T-c_1)^{\mathrm{ord}_c(P)} H$ für ein Polynom H mit $H(c_1) \neq 0_{\mathbb{K}}$. Weil $\mathrm{ord}_c(P) \geq 1$, ist

$$\deg\Big((T-c_1)^{\mathrm{ord}_c(P)-1} H\Big) < \deg(P)$$

und so können wir induktiv schreiben

$$(T-c_1)^{\mathrm{ord}_c(P)-1} H = (T-c_2)\cdots(T-c_k)P_0,$$

für eine natürliche Zahl k mit $k - 1 \leq \deg\Big((T-c_1)^{\mathrm{ord}_c(P)-1} H\Big)$, wobei das Polynom P_0 keine Nullstelle in $\mathbb{K}$ besitzt. Insbesondere ist $k \leq \deg(P)$ und

$$P = (T-c_1)^{\mathrm{ord}_c(P)} H = (T-c_1)\big((T-c_1)^{\mathrm{ord}_c(P)-1} H\big) = (T-c_1)\cdots(T-c_k)P_0,$$

wie gewünscht. Die Eindeutigkeit der Darstellung folgt leicht aus der Kommutativität des Polynomringes zusammen mit dem [**Skript LAI**, Lemma 1.22]. □

Aufgabe. Kann die Menge $\{P(c)\}_{c \in \mathbb{K}}$ endlich sein, wenn P ein nicht-konstantes Polynom ist?

Definition A.8. Der Körper $\mathbb{K}$ ist *algebraisch abgeschlossen*, wenn jedes nicht-konstante Polynom über $\mathbb{K}$ eine Nullstelle in $\mathbb{K}$ besitzt.

Bemerkung A.9. Weder $\mathbb{R}$ noch $\mathbb{Q}$ sind algebraisch abgeschlossen, weil das Polynom T^2+1 mit ganzzahligen Koeffizienten keine Nullstelle besitzt. Jedoch lässt sich jeder Körper als Teilkörper in einen algebraisch abgeschlossenen Körper einbetten.

Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen, wegen des Fundamentalsatzes der Algebra, dessen Beweis nicht nur algebraische Methoden verwendet (wir werden den Beweis in dieser Vorlesung nicht sehen).

Korollar A.10. *Wenn der Körper algebraisch abgeschlossen ist, zerfällt jedes nicht-konstante Polynom P in Linearfaktoren. Das heißt, dass*

$$P = \lambda(T - c_1) \cdots (T - c_{\deg(P)}),$$

für Elemente $c_1, \ldots, c_{\deg(P)}$ und $\lambda \neq 0_{\mathbb{K}}$ aus $\mathbb{K}$.

Wenn P normiert ist, ist $\lambda = 1_{\mathbb{K}}$.

Beweis. Mit Hilfe des Korollars A.7 schreibe

$$P = (T - c_1) \cdots (T - c_k)P_0,$$

für eine natürliche Zahl $k \leq \deg(P)$ und ein Polynom P_0 ohne Nullstellen in $\mathbb{K}$. Insbesondere muss P_0 konstant sein, weil $\mathbb{K}$ algebraisch abgeschlossen ist. Da P nicht-konstant ist, haben wir $P_0 = \lambda \neq 0_K$ und somit $k = \deg(P)$, wie gewünscht.

Die zweite Behauptung folgt sofort, weil der Führungskoeffizient des Produktes

$$\lambda(T - c_1) \cdots (T - c_k)$$

gerade das Element λ ist. □

B Signatur einer symmetrischen Bilinearform

Im diesem Abschnitt sei $\mathbb{K}$ ein Körper der Charakteristik verschieden von 2. Wir wollen eine kanonische Darstellung für symmetrische Bilinearformen eines endlichdimensionalen $\mathbb{K}$-Vektorraumes mit Hilfe einer geeigneten Basisauswahl. Wenn wir die Vektoren v und w bezüglich dieser Basis B als Tupel mit Koordinaten $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ repräsentieren, ist der Wert der symmetrischen Bilinearform

$$\varphi(v, w) = \sum_{1 \leq i \leq n} \mu_i x_i y_i,$$

für Skalare $\mu_1, \ldots, \mu_n$ aus $\mathbb{K}$, welche die *Gewichte* der symmetrischen Bilinearform darstellen.

Proposition B.1. (*Lagranges Diagonalisierungsmethode*) *Jede symmetrische Bilinearform $\varphi : V \times V \to \mathbb{K}$ eines n-dimensionalen $\mathbb{K}$-Vektorraumes V lässt sich nach einer Basiswechseltransformation durch eine Diagonalmatrix darstellen.*

Insbesondere ist φ genau dann ausgeartet, wenn 0 ein Eigenwert der Darstellungsmatrix ist.

Beachte, dass die Bedingung, dass 0 ein Eigenwert der Darstellungsmatrix A der symmetrischen Bilinearform φ bezüglich einer Basis B von V ist, nicht von der Wahl der Basis abhängt: Gegeben eine andere Basis B', sei S die Übergangsmatrix von B nach B'. Die Darstellungsmatrix von φ bezüglich der Basis B' ist dann ${}^tS \cdot A \cdot S$. Eine Matrix hat genau dann 0 als Eigenwert, wenn sie singulär ist, oder äquivalent dazu, wenn die Determinante Null ist. Da $\det(S) = \det({}^tS) \neq 0_{\mathbb{K}}$ ist, bleibt diese Eigenschaft bei Basiswechsel erhalten.

Beweis. Für die erste Behauptung genügt es zu zeigen, dass der Vektorraum V eine Basis $B = \{b_1, \ldots, b_n\}$ besitzt, welche aus paarweise orthogonalen Vektoren bezüglich der Bilinearform φ besteht: Es ist $\varphi(b_i, b_j) = 0$ falls $i \neq j$ und die Darstellungsmatrix von φ bezüglich B ist in Diagonalform mit den Einträgen $\varphi(b_1, b_1), \ldots, \varphi(b_n, b_n)$ auf der Diagonalen.

Wir beweisen die Existenz einer solchen Basis induktiv über $n \geq 1$. Für $n = 1$ ist die Behauptung trivial, weil jeder nicht-triviale Vektor eine orthogonale Basis bildet. Wir nehmen also an, dass die Behauptung für alle Werte $k < n$ bewiesen wurde. Sei $\varphi : V \times V \to \mathbb{K}$ eine symmetrische Bilinearform auf dem n-dimensionalen Vektorraum V mit zugehöriger quadratischen Form Q (siehe Definition 3.1). Wenn Q die Nullabbildung auf V ist, dann ist φ trivial, nach dem Polarisationssatz 3.2, und jede Basis von V besteht aus orthogonalen Vektoren. Sonst gibt es einen Vektor b_1 mit $Q(b_1) = \varphi(b_1, b_1) \neq 0_{\mathbb{K}}$. Die Abbildung

$$\begin{array}{rccc} F: & V & \to & \mathbb{K} \\ & v & \mapsto & \varphi(v, b_1) \end{array}$$

ist linear mit eindimensionalen Bild $\mathrm{Im}(F)$, weil $Q(b_1) = F(b_1)$ im Bildbereich liegt. Nach dem [**Skript LAI**, Rangssatz 2.34] ist der Kern $\mathrm{Ker}(F)$ ein Unterraum von V und hat Dimension $n - 1$. Beachte, dass alle Elemente aus $\mathrm{Ker}(F)$ orthogonal zu b_1 sind. Wir finden induktiv eine Basis $\{b_2, \ldots, b_n\}$ von $\mathrm{Ker}(F)$, welche orthogonal bezüglich der Einschränkung der Bilinearform ist. Wir müssen nur noch zeigen, dass die Menge $\{b_1, \ldots, b_n\}$ linear unabhängig ist. Sonst wäre b_1 eine Linearkombination von $\{b_2, \ldots, b_n\}$ und somit läge b_1 in $\mathrm{Ker}(F)$, was $Q(b_1) \neq 0_{\mathbb{K}}$ widerspricht.

Für die letzte Behauptung ist offensichtlich, dass φ ausgeartet ist, wenn 0 ein Eigenwert der Darstellungsmatrix ist, weil die Einträge auf der Diagonalen genau $\varphi(b_1, b_1), \ldots, \varphi(b_n, b_n)$ sind. Für die andere Richtung nehmen wir an, dass 0 kein Eigenwert der Matrix ist. Gegeben einen Vektor v aus V, schreibe

$$v = \sum_{1 \leq i \leq n} \lambda_i b_i.$$

Falls mit $\varphi(v, -) : V \to \mathbb{K}$ trivial ist, folgt aus der Linearität, dass

$$0_{\mathbb{K}} = \varphi(v, b_j) = \lambda_j \varphi(b_j, b_j),$$

also $\lambda_j = 0_{\mathbb{K}}$ für alle $1 \leq j \leq n$. Der Vektor v ist also 0_V, und somit φ nicht-ausgeartet. $\square$

Korollar B.2. *Wenn jedes Element des Körpers $\mathbb{K}$ ein Quadrat ist (z.B. wenn der Körper algebraisch abgeschlossen ist, wie $\mathbb{C}$), dann lässt sich jede symmetrische Bilinearform $\varphi : V \times V \to \mathbb{K}$ eines n-dimensionalen $\mathbb{K}$-Vektorraumes V nach einer Basiswechseltransformation durch eine Diagonalmatrix der Form*

$$\begin{pmatrix} \underbrace{1_{\mathbb{K}} \ \ddots \ 1_{\mathbb{K}}}_{d-\mathrm{mal}} & \\ & \underbrace{0_{\mathbb{K}} \ \ddots \ 0_{\mathbb{K}}}_{r-\mathrm{mal}} \end{pmatrix}$$

darstellen. Die Werte d und r hängen nur von φ ab.

Beweis. Nach der Proposition B.1 gibt es eine Basis $\{b_1, \ldots, b_n\}$ von V, welche orthogonal bezüglich φ ist. Setze

$$c_i = \begin{cases} \frac{b_i}{\sqrt{Q(b_i)}}, & \text{falls } Q(b_i) \neq 0_{\mathbb{K}} \\ b_i, & \text{sonst} \end{cases},$$

wobei $\sqrt{\lambda}$ eine beliebige Wahl der quadratischen Wurzel des Elementes λ aus $\mathbb{K}$ ist. Es ist klar, dass $\{c_1, \ldots, c_n\}$ eine orthogonale Basis von V derart bildet, dass $\varphi(c_i, c_i)$ nur $0_{\mathbb{K}}$ oder $1_{\mathbb{K}}$ sein kann. Wir können die Basis ohne Beschränkung der Allgemeinheit so umordnen, dass zuerst alle Elemente c_i mit $\varphi(c_i, c_i) = 1_{\mathbb{K}}$ vorkommen, was die obige Darstellungsmatrix liefert.

Beachte, dass die Zahl $d = \mathrm{Rg}(\varphi)$ dem Rang der Bilinearform entspricht, welcher nach dem Lemma 2.17 nicht von der Wahl der Basis abhängt. Somit ist $r = \dim_{\mathbb{K}} V - d$ auch eindeutig bestimmt. $\square$

Satz B.3. (*Der Trägheitssatz von Sylvester*) *Jede symmetrische Bilinearform $\varphi : V \times V \to \mathbb{K}$ eines n-dimensionalen $\mathbb{R}$-Vektorraumes V lässt sich nach einer Basiswechseltransformation durch eine Diagonalmatrix der Form*

$$\begin{pmatrix} \underbrace{\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}}_{p-mal} & & \\ & \underbrace{\begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix}}_{q-mal} & \\ & & \underbrace{\begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix}}_{r-mal} \end{pmatrix}$$

darstellen. Die Werte p, q und r hängen nur von φ ab.

Beweis. Wähle eine Basis $\{b_1, \ldots, b_n\}$ von V, welche orthogonal bezüglich φ ist. So eine Basis existiert nach Proposition B.1. setze nun

$$c_i = \begin{cases} \frac{b_i}{\sqrt{Q(b_i)}}, \text{ falls } Q(b_i) > 0 \\ \frac{b_i}{\sqrt{-Q(b_i)}}, \text{ falls } Q(b_i) < 0 \\ b_i, \text{ sonst.} \end{cases} .$$

Ohne Beschränkung der Allgemeinheit können wir die Basis $\{c_1, \ldots, c_n\}$ so anordnen, dass zuerst alle c_i's mit $\varphi(c_i, c_i) = 1$ vorkommen, danach die c_j's mit $\varphi(c_j, c_j) = -1$ und am Ende die mit $\varphi(c_j, c_j) = 0$. Die Darstellungsmatrix von φ bezüglich der orthogonalen Basis $\{c_1, \ldots, c_n\}$ hat die gewünschte Form.

Wir zeigen nun, dass die Werte p, q und r intrinsisch für φ sind: Da $\mathrm{Rg}(\varphi) = p + q$, hängt $r = \dim_{\mathbb{K}} V - \mathrm{Rg}(\varphi)$ nicht von der Wahl der Basis ab. Es genügt also zu zeigen, dass p nur von φ abhängt. Beachte, dass die symmetrische Bilinearform φ auf dem p-dimensionalen Unterraum $\mathrm{Lin}(\{c_1, \ldots, c_p\})$ positiv definit ist, weil

$$\varphi\left(\sum_{1 \le i \le p} \lambda_i c_i, \sum_{1 \le i \le p} \lambda_i c_i\right) = \sum_{1 \le i \le p} \lambda_i^2$$

und jede endliche Summe nicht-trivialer Quadrate in $\mathbb{R}$ echt positiv ist. Sei h das Maximum der Dimensionen aller Unterräume U von V derart, dass φ positiv definit auf U ist. Dann ist klarerweise $p \leq h$, denn $\operatorname{Lin}(\{c_1, \ldots, c_p\})$ ist ein solcher Unterraum. Wir müssen nur zeigen, dass $p = h$. Sei nun U ein Unterraum von V der Dimension h derart, dass φ positiv definit auf U ist. Es genügt zu zeigen, dass U und $\operatorname{Lin}(\{c_{p+1}, \ldots, c_n\})$ transversal liegen, denn somit gilt

$$n = p+n-p \leq \dim_{\mathbb{K}} U + \dim_{\mathbb{K}} \operatorname{Lin}(\{c_{p+1}, \ldots, c_n\}) = \dim_{\mathbb{K}} U \oplus \operatorname{Lin}(\{c_{p+1}, \ldots, c_n\}) \leq \dim_{\mathbb{K}} V = n.$$

Insbesondere gilt Gleichheit oben überall und es folgt, dass $p = h$, wie gewünscht.

Angenommen, dass der Durchschnitt

$$U \cap \operatorname{Lin}(\{c_{p+1}, \ldots, c_n\})$$

nicht trivial ist, finden wir einen Vektor u mit

$$0_V \neq u = \sum_{p+1 \leq i \leq n} \lambda_i c_i \text{ in } U,$$

ist

$$\varphi(u, u) = \sum_{p+1 \leq i \leq n} \lambda_i^2 \varphi(c_i, c_i) = \sum_{p+1 \leq i \leq p+q} -\lambda_i^2 \leq 0.$$

Der Vektor u ist nicht trivial und kommt aus U. Da φ auf U positiv definit ist, muss $\varphi(u, u) > 0$ sein, was den gewünschten Widerspruch liefert. □

Definition B.4. Die ganze Zahl $p-q$ heißt die *Signatur* der reellen symmetrischen Bilinearform $\varphi : V \times V \to \mathbb{R}$.

Literaturverzeichnis

[**Skript LAI**] A. Martin-Pizarro, *Lineare Algebra I: Ein Kurzskript*, Wintersemester 20/21, `http://home.mathematik.uni-freiburg.de/loesch/lehre/ws2021/LA.pdf`