LINEARE ALGEBRA II

OLIVER C. SCHNÜRER

Zusammenfassung. Bei diesem Manuskript handelt es sich um Notizen zu einer Vorlesung Lineare Algebra II (B2) an der Universität Konstanz im Sommersemester 2011.

Vielen Dank an M. Kummer für Korrekturen.

Inhaltsverzeichnis

Wir benutzen [1–3, 6, 9].

6. Vektorräume mit Skalarprodukt – Fortsetzung

6.10. Normale Matrizen.

Definition 6.10.1. Sei A eine quadratische Matrix über $\mathbb{R}$ oder $\mathbb{C}$. Dann heißt A normal, falls

$$A^*A = AA^*$$

gilt.

Ein Endomorphismus f eines Skalarproduktraumes heißt normal, wenn $f \circ f^* = f^* \circ f$ gilt.

Beispiele 6.10.2.

(i) Unitäre ($A^{-1} = A^*$), hermitesche ($A = A^*$) und Matrizen mit $A = -A^*$ sind normal.

(ii) Orthogonale ($A^{-1} = A^T$), symmetrische ($A = A^T$) und schiefsymmetrische Matrizen ($A = -A^T$) sind normal.

Resultate für solche Matrizen werden wir einheitlicher für normale Matrizen behandeln.

(iii) Es gibt weitere normale Matrizen, z. B.

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

Date: 17. Dezember 2012.

denn es gilt

$$\begin{aligned} AA^T &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = A^T A. \end{aligned}$$

Theorem 6.10.3. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei f ein normaler Endomorphismus. Dann gelten*

$$\ker f = \ker f^* \quad \textit{und} \quad \operatorname{im} f = \operatorname{im} f^*.$$

Insbesondere folgt $V = \ker f \oplus \operatorname{im} f$, wobei die Summanden senkrecht zueinander sind, $\ker f \perp \operatorname{im} f$.

Beweis. Aus $v \in \ker f$ erhalten wir

$$0 = \langle f(v), f(v) \rangle = \langle v, f^*(f(v)) \rangle = \langle v, f(f^*(v)) \rangle = \langle f^*(v), f^*(v) \rangle.$$

Somit gilt auch $v \in \ker f^*$.

Aus [7] wissen wir, dass $\ker f^* = (\operatorname{im} f)^\perp$ gilt. Somit folgt

$$(\operatorname{im} f)^\perp = \ker f^* = \ker f = (\operatorname{im} f^*)^\perp.$$

Da V endlichdimensional ist, folgt auch $\operatorname{im} f = \operatorname{im} f^*$. □

Korollar 6.10.4. *Sei V ein endlichdimensionaler Skalarproduktraum (über $\mathbb{R}$ oder $\mathbb{C}$). Sei $f\colon V \to V$ normal. Sei $\lambda \in \mathbb{C}$ bzw. $\in \mathbb{R}$. Dann gilt für die zugehörigen Eigenräume*

$$E_\lambda(f) = E_{\bar\lambda}(f^*).$$

Beweis. Setze $g := f - \lambda \operatorname{id}$. Dann ist $g^* = f^* - \bar\lambda \operatorname{id}$. g ist ebenfalls normal, denn es gilt

$$g \circ g^* = \underbrace{f \circ f^*}_{=f^* \circ f} - \lambda f^* - \bar\lambda f + \lambda\bar\lambda \operatorname{id} = g^* \circ g.$$

Somit erhalten wir

$$E_\lambda(f) = \ker g = \ker g^* = E_{\bar\lambda}(f^*).$$ □

Wir können normale Endomorphismen nicht nur diagonalisieren; dies ist auch eine notwendige Bedingung für die Diagonalisierbarkeit. Der folgende Satz funktioniert aber nur für komplexe Matrizen. Drehmatrizen sind Beispiele für reelle orthogonale und daher auch normale aber nicht reell diagonalisierbare Matrizen.

Theorem 6.10.5. *Sei V ein endlichdimensionaler unitärer Skalarproduktraum mit* $\dim V = n$. *Sei $f\colon V \to V$. Dann sind die folgenden Bedingungen äquivalent:*

(i) Es gibt eine Orthonormalbasis aus Eigenvektoren von f.
(ii) f ist normal.

Beweis.
„$\Longrightarrow$“: Sei $a_1, \ldots, a_n$ eine Orthonormalbasis aus Eigenvektoren von f mit $f(a_i) = \lambda_i a_i$. Wir erhalten

$$\langle a_j, f^*(a_i) \rangle = \langle f(a_j), a_i \rangle = \lambda_j \langle a_j, a_i \rangle = \left\langle a_j, \overline{\lambda_i} a_i \right\rangle.$$

Somit ist $f^*(a_i) = \bar\lambda_i a_i$. Für diese Basisvektoren gilt

$$f(f^*(a_i)) = \bar\lambda_i f(a_i) = \bar\lambda_i \lambda_i a_i = \lambda_i f^*(a_i) = f^*(f(a_i)).$$

Da es genügt, Normalität für eine Basis nachzurechnen, ist f normal.

„$\Longleftarrow$“: Sei umgekehrt f normal. Über $\mathbb{C}$ zerfällt das charakteristische Polynom in Linearfaktoren. Wähle einen normierten Eigenvektor a_1 zum Eigenvektor λ_1 und definiere $E_{\lambda_1} := \langle a_1 \rangle$. Setze $W := E_{\lambda_1}^{\perp}$. Es gilt $\dim W = n-1$. Können wir zeigen, dass W ein f-invarianter Unterraum von V ist und dass $f|_W$ normal ist, so folgt die Behauptung per Induktion nach der Dimension des Vektorraumes. Es gilt für $w \in W$

$$\langle f(w), a_1 \rangle = \langle w, f^*(a_1) \rangle = \langle w, \bar{\lambda}_1 a_1 \rangle = \lambda_1 \langle w, a_1 \rangle = 0.$$

Somit ist W ein f-invarianter Unterraum. Da E_{λ_1} auch in einem Eigenraum der normalen Abbildung f^* enthalten ist, gilt auch $f^*(W) \subset W$. Da W für f und für f^* ein invarianter Unterraum ist, ist $f|_W$ ebenfalls normal und die Behauptung folgt. □

Für Matrizen umformuliert erhalten wir

Korollar 6.10.6. *Sei $A \in \mathbb{C}^{n\times n}$. Dann ist A genau dann normal, wenn es eine unitäre Matrix $S \in \mathbb{C}^{n\times n}$ gibt, so dass SAS^{-1} diagonal ist.*

Theorem 6.10.7. *Seien $A, B \in \mathbb{C}^{n\times n}$ normal und gelte $[A,B] = AB - BA = 0$. Dann gibt es eine unitäre Matrix U, so dass UAU^* und UBU^* diagonal sind.*

Beweis. Sei λ_1 ein Eigenwert von A und sei V_1 der zugehörige Eigenraum. Sei $v_1 \in V_1$. Aus $AB = BA$ folgt

$$A(Bv_1) = BAv_1 = \lambda_1(Bv_1).$$

Somit gilt $Bv_1 \in V_1$. V_1 ist also ein invarianter Unterraum für A und B.

Definiere $W_1 := V_1^{\perp}$. Wir behaupten, dass W_1 auch ein invarianter Unterraum für A und B ist. Wir argumentieren wie in Theorem 6.10.5. Nach Korollar 6.10.4 gilt $A^*v = \bar{\lambda}_1 v$ für alle $v \in V_1$. Seien $v \in V_1$ und $w \in W_1$ beliebig. Dann gilt

$$\langle Aw, v \rangle = \langle w, A^*v \rangle = \langle w, \bar{\lambda}_1 v \rangle = \lambda_1 \langle w, v \rangle = 0.$$

Wir erhalten $AW_1 \subset V_1^{\perp} = W_1$.

Zu $BW_1 \subset W_1$: Nach Definition von V_1 ist $V_1 = \ker(A - \lambda_1 \mathbb{1})$, wobei wir Matrizen und zugehörige Abbildungen identifizieren. Wir haben gezeigt, dass $AW_1 \subset W_1$ gilt. Somit folgt auch $(A - \lambda_1\mathbb{1})W_1 \subset W_1$. Da $V_1 \oplus W_1 = V$ gilt und W_1 endlichdimensional ist, ist $(A-\lambda_1\mathbb{1})|_{W_1} \colon W_1 \to W_1$ bijektiv. Also gelten $(A - \lambda_1\mathbb{1})V = W_1$ und $W_1 = (A - \lambda_1\mathbb{1})W_1$. Wir erhalten

$$BW_1 = B(A - \lambda_1\mathbb{1})W_1 = (A - \lambda_1\mathbb{1})BW_1 \subset (A - \lambda_1\mathbb{1})V = W_1.$$

Die Zerlegung $V = V_1 \oplus W_1$ in orthogonale unter A und B invariante Unterräume impliziert, dass A und B (nach einer unitären Transformation) in Blockgestalt sind. Der zu V_1 gehörige Block ist für A ein Vielfaches der Einheitsmatrix. Die Normalität überträgt sich aufgrund der Blockgestalt auf jeden der einzelnen Blöcke, ebenso $[A,B] = 0$. Mit Theorem 6.10.5 können wir den zu V_1 gehörigen Block von B diagonalisieren: Die zugehörige unitäre Transformationsmatrix läßt sich zu einer unitären Matrix von V erweitern. Der zu W_1 gehörige Block ist kleiner als die Matrizen A und B. Somit folgt die Aussage per Induktion. □

Korollar 6.10.8. *Seien $A, B \in \mathbb{C}^{n\times n}$ normale Matrizen, die kommutieren, d. h. gelte $AB = BA$. (Oder der Kommutator $[A,B] := AB - BA$ verschwindet: $[A,B] = 0$) Dann sind AB und $A+B$ ebenfalls normal.*

Beweis. Die Eigenschaften, normal zu sein und zu kommutieren gelten genau dann für A und B, wenn sie für UAU^{-1} und UBU^{-1} gelten. Verwende nun Theorem 6.10.7. Für Diagonalmatrizen ist die Aussage klar. □

Theorem 6.10.9. *Sei $A \in \mathbb{C}^{n\times n}$. Dann sind die folgenden Aussagen sind äquivalent:*

(i) *A ist normal.*
(ii) *A lässt sich durch eine unitäre Matrix diagonalisieren.*
(iii) *Es gibt eine Orthonormalbasis aus Eigenvektoren von A.*
(iv) *Es gilt $\|Ax\| = \|A^*x\|$ für alle $x \in \mathbb{C}^n$.*
(v) *Der „hermitesche Anteil“ $\frac{1}{2}(A+A^*)$ und der “anti-hermitesche Anteil“ $\frac{1}{2}(A-A^*)$ kommutieren.*

Beweis. Die Äquivalenz von (i), (ii) und (iii) haben wir bereits gezeigt.

Klar ist (i) $\Longrightarrow$ (iv). Gelte umgekehrt (iv). Wir erhalten nacheinander

$$\begin{aligned} \langle Ax, Ax\rangle &= \langle A^*x, A^*x\rangle, \\ \langle x, A^*Ax\rangle &= \langle x, AA^*x\rangle, \\ 0 &= \langle x, [A, A^*]x\rangle, \end{aligned}$$

jeweils für alle $x \in \mathbb{C}^n$. Da $[A, A^*]^* = [A, A^*]$ gilt, ist $[A, A^*]$ normal und daher diagonalisierbar. Somit folgt aus der letzten Zeile $[A, A^*] = 0$, also (i).

(i) $\Longleftrightarrow$ (v) folgt direkt aus

$$[(A + A^*), (A - A^*)] = 2[A^*, A].$$

Von dieser Gleichheit überzeugt man sich durch eine einfache direkte Rechnung. □

Definition 6.10.10. Eine Matrix $A = \left(a^i_j\right)_{1\le i\le j} \in \mathbb{C}^{n\times n}$ heißt positiv semidefinit, falls sie hermitesch ist und

$$\langle x, Ax\rangle \ge 0 \quad \text{für alle } x \in \mathbb{C}^n$$

gilt. Schreibweise: $A \ge 0$, $\left(a^i_j\right) \ge 0$ oder $A \succcurlyeq 0$, $\left(a^i_j\right) \succcurlyeq 0$, auch $a^i_j \ge 0$ oder $a^i_j \succcurlyeq 0$, falls dies nicht mit der Positivität der Einträge verwechselt werden kann.

Gilt $\langle x, Ax\rangle > 0$ für alle $x \in \mathbb{C}^n \setminus \{0\}$, so heißt A positiv definit. Wir schreiben $A > 0$, $\left(a^i_j\right) \ge 0$ oder $A \succ 0$, $\left(a^i_j\right) \succ 0$ bzw. $a^i_j > 0$ oder $a^i_j \succ 0$.

Lemma 6.10.11. *Sei $A \in \mathbb{C}^{n\times n}$ positiv semidefinit. Dann gibt es eine eindeutig bestimmte positiv semidefinite Quadratwurzel, d. h. eine positiv semidefinite Matrix $\sqrt{A} \in \mathbb{C}^{n\times n}$, so dass $\sqrt{A}\sqrt{A} = A$ gilt.*

Beweis. Da A normal ist, gibt es eine unitäre Matrix U, so dass U^*AU diagonal ist. Die Diagonaleinträge von U^*AU sind nicht negativ. Somit gibt es eine (reelle) Diagonalmatrix B mit $BB = U^*AU$. Die gesuchte Matrix ist $\sqrt{A} = UBU^*$. Da B eine Diagonalmatrix mit nichtnegativen Einträgen ist, sieht man direkt, dass $\sqrt{A}$ positiv semidefinit ist. $\sqrt{A}\sqrt{A} = A$ ist einfach nachzurechnen.

Sei C eine weitere Quadratwurzel von A. Da C hermitesch ist, gibt es eine Orthonormalbasis $\{a_i\}$ aus Eigenvektoren von C zu Eigenwerten λ_i. Wir erhalten

$$Aa_i = CCa_i = \lambda_i^2 a_i.$$

Somit sind die Werte λ_i^2 als Eigenwerte von A eindeutig bestimmt. Aufgrund der positiven Semidefinitheit sind auch die Werte $\lambda_i \ge 0$ eindeutig bestimmt. Die obigen Gleichungen liefern, dass ein Eigenraum von A zum Eigenwert λ_i^2 ein Eigenraum von C zum Eigenwert λ_i ist. Somit ist C eindeutig bestimmt. □

Das folgende Theorem verallgemeinert die Darstellung einer komplexen Zahl in der Form $re^{i\varphi}$ auf Matrizen.

Theorem 6.10.12 (Polarzerlegung)**.** *Sei $A \in \mathbb{C}^{n\times n}$. Dann gibt es eine unitäre Matrix U und eine positiv semidefinite Matrix S, so dass $A = SU$ gilt.*

S ist eindeutig bestimmt. Ist A invertierbar, so ist die Zerlegung eindeutig.

Beweis. Die Matrix AA^* ist hermitesch und positiv semidefinit. Setze $S := \sqrt{AA^*}$. Gibt es solch eine Zerlegung, so ist S eindeutig bestimmt, denn es gilt

$$AA^* = SUU^*S^* = SS^* = S^2$$

und die Wurzel ist eindeutig bestimmt.

Wähle eine Orthogonalbasis $\{a_i\}_i$ von Eigenvektoren, so dass

$$A^*Aa_i = \lambda_i a_i$$

gilt. Es folgt

$$(AA^*)Aa_i = \lambda_i Aa_i.$$

Somit ist Aa_i (falls $\neq 0$) ein Eigenvektor zu AA^* und es gilt

$$SAa_i = \sqrt{\lambda_i}Aa_i.$$

Wegen

$$\langle Aa_i, Aa_j\rangle = \langle a_i, A^*Aa_j\rangle = \overline{\lambda}_j\delta_{ij} = \lambda_j\delta_{ij}$$

sind die Vektoren $\{Aa_i\}_i$ paarweise orthogonal zueinander. Durch Normieren und ergänzen dieser Vektoren erhalten wir eine Orthonormalbasis $\{w_i\}_i$. Ist $Aa_i \neq 0$, so gilt aufgrund der obigen Rechnung

$$w_i = \frac{Aa_i}{\|Aa_i\|} = \frac{Aa_i}{\sqrt{\langle Aa_i, Aa_i\rangle}} = \frac{Aa_i}{\sqrt{\lambda_i}}.$$

Definiere U als die unitäre Matrix, die die Orthonormalbasis $\{a_i\}_i$ auf die Orthonormalbasis $\{w_i\}_i$ abbildet. Es ist nun noch nachzuweisen, dass $A = SU$ gilt. Wir zeigen dazu, dass $Aa_i = SUa_i$ für alle a_i gilt.

(i) Sei zunächst $Aa_i \neq 0$. Dann gilt

$$SUa_i = Sw_i = S\frac{Aa_i}{\|Aa_i\|} = \sqrt{\lambda_i}\frac{Aa_i}{\sqrt{\lambda_i}} = Aa_i.$$

(ii) Sei nun $Aa_i = 0$. Wir wollen zeigen, dass dann auch $Sw_i = SUa_i = 0$ gilt. Dies ist äquivalent zu $AA^*w_i = 0$. Da $AA^*w_i \in \operatorname{im} A$ ist, genügt es, $\langle Aa_j, AA^*w_i\rangle = 0$ für alle j zu zeigen. Es gilt

$$\langle Aa_j, AA^*w_i\rangle = \langle AA^*Aa_j, w_i\rangle = \lambda_j\langle Aa_j, w_i\rangle = 0.$$

Falls $Aa_j \neq 0$ ist, folgt $Aa_j = \sqrt{\lambda_j}w_j$. Nun bilden die Vektoren w_i eine Orthogonalbasis, aus $Aa_i = 0$ folgt somit insbesondere $i \neq j$. Andernfalls ist $Aa_j = 0$ und die Behauptung ist trivial. □

Die folgende Zerlegung wird häufig ebenfalls als Polarzerlegung bezeichnet.

Korollar 6.10.13. *Sei $A \in \mathbb{C}^{n\times n}$. Dann gibt es eine unitäre Matrix U und eine positiv semidefinite Matrix S, so dass $A = US$ gilt. S ist eindeutig bestimmt. Ist A invertierbar, so ist die Zerlegung eindeutig.*

Beweis. Die Polarzerlegung $A^* = \tilde{S}\tilde{U}$ liefert eine Zerlegung $A = \tilde{U}^*\tilde{S}^*$, also von der angegebenen Form. Die Eindeutigkeit von S folgt aus $A^*A = SU^*US = S^2$. □

Lemma 6.10.14. *Sei $A \in \mathbb{C}^{n\times n}$. Dann sind die folgenden Aussagen äquivalent:*

(i) A ist normal.
(ii) In der Polarzerlegung $A = SU$ kommutieren S und U.
(iii) In der Zerlegung $A = US$ kommutieren S und U.
(iv) Es gibt eine unitäre Matrix W mit $A^* = AW$.

Beweis.
„(i)$\Longrightarrow$ (ii)“: Sei W unitär, so dass WAW^* diagonal ist,

$$WAW^* = \operatorname{diag}\left\{r_1e^{i\varphi_1}, \dots, r_1e^{i\varphi_1}, r_2e^{i\varphi_2}, \dots, r_2e^{i\varphi_2}, \dots, 0, \dots, 0\right\}$$

mit $r_i > 0$, $0 \le \varphi_i < 2\pi$ und paarweise verschiedenen Tupeln (r_i, φ_i). Die zugehörige Polarzerlegung hat dann die Gestalt $WAW^* = \tilde{S}\tilde{U}$, wobei

$$\tilde{S} = \operatorname{diag}\{r_1, \dots, r_1, r_2, \dots, r_2, \dots, 0, \dots 0\}$$

und

$$\tilde{U} = \operatorname{diag}\left\{e^{i\varphi_1}, \dots, e^{i\varphi_1}, e^{i\varphi_2}, \dots, e^{i\varphi_2}, \dots, U_0\right\}$$

sind und U_0 eine unitäre Matrix bezeichnet, der letzte Eintrag also als Block zu verstehen ist. Dann gilt $A = \big(W^*\tilde{S}W\big)\big(W^*\tilde{U}W\big)$ und die beiden Ausdrücke in Klammern entsprechen gerade der Polarzerlegung. Man sieht direkt, dass diese Matrizen vertauschen.

„(ii)$\Longrightarrow$ (i)“: Sei $A = SU = US$. Dann folgen $A^* = S^*U^* = SU^*$, $A^* = U^*S^* = U^*S$ und

$$A^*A - AA^* = SU^*US - SUU^*S = S^2 - S^2 = 0.$$

„(ii)$\Longleftrightarrow$ (iii)“: Ist offensichtlich.

„(iii)$\Longrightarrow$ (iv)“: Schreibe $A = SU$ mit $[U, S] = 0$. Wir erhalten $A = US$ und daraus $A^* = S^*U^* = SU^* = SUU^*U^* = AU^*U^*$. U^* ist unitär, da U unitär ist: $U^* = U^{-1}$ liefert $U^{**} = (U^*)^{-1}$.

„(iv)$\Longrightarrow$ (i)“: Aus $A^* = AU$ folgt einerseits

$$A = AUU^* = A^*U^*$$

und andererseits

$$A = (A^*)^* = (AU)^* = U^*A^*,$$

somit ergeben sich also

$$A^*U^* = U^*A^*$$

und

$$UA = AU.$$

Zusammengenommen erhalten wir

$$A^*A - AA^* = (AU)A - A(AU) = AAU - AAU = 0.$$

Somit ist A normal. □

7. Minimalpolynome und Jordansche Normalform

7.1. **Minimalpolynome.** Sei $f\colon V \to V$ ein Endomorphismus, durch A dargestellt. Wir setzen $A^n := \underbrace{A \cdot A \cdot \ldots \cdot A}_{n-\text{mal}}$ und $A^0 := \mathbb{1}$. Damit können wir eine Matrix auch in ein Polynom

$$p(X) = \sum_{i=0}^{n} a_iX^i, \quad p \in F[X],$$

einsetzen,

$$p(A) := \sum_{i=0}^{n} a_iA^i.$$

Die Abbildung

$$\Phi_A\colon F[X] \to \mathrm{End}(V),$$
$$p(X) \mapsto p(A)$$

heißt Einsetzungshomomorphismus. Sie ist ein Homomorphismus von Ringen (es gilt $\Phi_A(p+q) = \Phi_A(p) + \Phi_A(q)$ und $\Phi_A(p \cdot q) = \Phi_A(p) \cdot \Phi_A(q)$) und von F-Vektorräumen. (Beachte, dass dabei nicht nur X durch A, sondern auch $a \in F$ durch $a \cdot \mathbb{1} \equiv a \cdot A^0$ ersetzt wird. Bei Polynomringen in mehreren (kommutierenden) Variablen wird das Einsetzen von Matrizen für die Variablen problematisch wenn die Matrizen nicht vertauschen.)

Das Bild

$$F[A] \equiv \{p(A)\colon p(X) \in F[X]\} \subset \mathrm{End}(V)$$

ist ein kommutativer Unterring des i. a. nicht kommutativen Ringes $\mathrm{End}(V)$.

Der Kern

$$I_A := \{p(X) \in F[X]\colon p(A) = 0\} \subset F[X]$$

heißt Ideal von A.

Theorem 7.1.1 (Cayley-Hamilton)**.** *Sei $f\colon V \to V$ ein Endomorphismus,* $\dim V < \infty$*. Sei χ_f das charakteristische Polynom zu f, A die f bezüglich einer beliebigen Basis darstellende Matrix und χ_A das charakteristische Polynom von A. Dann gelten*

$$\chi_f(f) = 0 \quad \textit{und} \quad \chi_A(A) = 0.$$

Achtung, bei einer Rechnung

$$\chi_A(A) \neq \det(A - A \cdot \mathbb{1}) = \det(A - A) = \det 0 = 0$$

hat man an der Stelle „$\neq$“ die Polynomvariable X in $X \cdot \mathbb{1}$ durch A ersetzt; richtig wäre aber gewesen, det als Polynom aufzufassen. Wie man dies in der Determinante aufschreibt, ergibt sich aus dem folgenden Beweis.

Beweis. Es genügt, den Satz für Matrizen zu beweisen. Definiere $B(X) := (A - X\mathbb{1})^T$. Es gilt $\det B(X) = \chi_A(X)$. Sei $B^{\mathrm{adj}}(X)$ die adjunkte Matrix zu $B(X)$ wie in [7, Kap. Determinanten]. Dann folgt

$$B^{\mathrm{adj}}(X) \cdot B(X) = \det B(X)\mathbb{1} = \chi_A(X) \cdot \mathbb{1}. \tag{7.1}$$

Setzen wir nun statt X die Matrix A ein, wenden also den Einsetzungshomomorphismus auf jede Komponente an, so erhalten wir eine Matrix mit Matrizen als Einträgen

$$B(A) = \begin{pmatrix} a_1^1\mathbb{1} - A & a_1^2\mathbb{1} & \dots & a_1^n\mathbb{1} \\ a_2^1\mathbb{1} & a_2^2\mathbb{1} - A & \dots & a_2^n\mathbb{1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n^1\mathbb{1} & a_n^2\mathbb{1} & \dots & a_n^n\mathbb{1} - A \end{pmatrix}.$$

Wir rechnen nun weiter mit Matrizen, deren Einträge in $F[A]$ liegen. Wir können solch eine Matrix mit einer Matrix aus Spaltenvektoren multiplizieren

$$B(A)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} a_1^1 e_1 - Ae_1 + a_1^2 e_2 + \ldots + a_1^n e_n \\ a_2^1 e_1 + a_2^2 e_2 - Ae_2 + \ldots + a_2^n e_n \\ \vdots \qquad\qquad \vdots \\ a_n^1 e_1 + a_n^2 e_2 + \ldots + a_n^n e_n - Ae_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Setzen wir in die Gleichung (7.1) polynomialer Ausdrücke A statt $\mathbb{1}$ (und $\lambda \cdot \mathbb{1}$ statt eines Körperelementes λ) ein, so folgt

$$B^{\text{adj}}(A) \cdot B(A) = \begin{pmatrix} \chi_A(A) & 0 & \dots & 0 \\ 0 & \chi_A(A) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \chi_A(A) \end{pmatrix}.$$

Wir multiplizieren wie oben von rechts mit einem Vektor aus Spaltenvektoren und erhalten

$$\begin{aligned} \begin{pmatrix} \chi_A(A) & 0 & \dots & 0 \\ 0 & \chi_A(A) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \chi_A(A) \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} &= B^{\text{adj}}(A) \cdot B(A) \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \\ &= B^{\text{adj}}(A) \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \end{aligned}$$

Komponentenweise gelesen folgt $\chi_A(A)e_i = 0$ für $1 \le i \le n$. Somit ist $\chi_A(A) = 0$. □

Bemerkung 7.1.2. Mit etwas Algebra, kann man den Satz von Cayley-Hamilton wie folgt einfacher beweisen: Wir betrachten das Problem in einem Körper $K \supset F$, in dem $\chi_A(X)$ in Linearfaktoren zerfällt. Solch einen Körper gibt es für jedes Polynom (Algebravorlesung). Es genügt dann, die Aussage in diesem Körper zu zeigen. Wir nehmen also ohne Einschränkung an, dass $\chi_A(X)$ in F bereits in Linearfaktoren zerfällt. Dann können wir ([7]) eine Basis finden, so dass A trigonal wird. Nun ist der Satz eine einfache Übungsaufgabe, insbesondere, wenn wir bereits Theorem 7.2.2 benutzen, das wir noch unabhängig hiervon zeigen werden.

Nicht nur I_A ist ein Ideal. Allgemein definiert man

Definition 7.1.3 (Ideal)**.** Sei I eine Teilmenge eines kommutativen Ringes R mit Eins. Dann heißt I Ideal von R, wenn

(I1) $a + b \in I$ für alle $a, b \in I$,
(I2) $a \cdot r \in I$ für alle $a \in I$, $r \in R$.
(I3) $\mathfrak{p} \subset R$ heißt Primideal, falls $\mathfrak{p} \neq R$ ist und falls aus $a \cdot b \in \mathfrak{p}$ folgt, dass $a \in \mathfrak{p}$ oder $b \in \mathfrak{p}$ gilt.
(I4) $\mathfrak{m} \subset R$ heißt maximales Ideal, wenn es keine größeren nichttrivialen Ideale gibt, d. h. wenn jedes Ideal $\mathfrak{a}$ mit $\mathfrak{m} \subsetneq \mathfrak{a} \subset R$ bereits $\mathfrak{a} = R$ erfüllt.

Beispiele 7.1.4.
(i) $I_A \subset F[X]$, die Polynome f mit $f(A) = 0$.
(ii) Sei $m \in \mathbb{Z}$. Dann bilden die durch m teilbaren Zahlen ein Ideal, (m).
(iii) Sei $p \in \mathbb{N} \setminus \{0\}$ prim. Dann ist (p) ein Primideal in $\mathbb{Z}$.

Theorem 7.1.5. *Sei R ein kommutativer Ring mit Eins. Dann ist*
(i) das Ideal $\mathfrak{p} \subset R$ genau dann prim, wenn $R/\mathfrak{p}$ nullteilerfrei ist, d. h. falls aus $a \cdot b = 0$ folgt, dass $a = 0$ oder $b = 0$ gilt.
(ii) das Ideal $\mathfrak{m} \subset R$ genau dann maximal, wenn $R/\mathfrak{m}$ ein Körper ist.
(iii) jedes maximale Ideal prim.

Beweis. Übung. □

Definition 7.1.6 (Minimalpolynom)**.** Sei $I \subset F[X]$ ein Ideal. Dann heißt μ Minimalpolynom von I, falls

(i) μ normiert ist, d. h. von der Form $\mu(X) = \sum\limits_{i=0}^{n} a_i X^i$ mit $a_n = 1$ und
(ii) es für jedes $q \in I$ ein $p \in F[X]$ mit $q = p \cdot \mu$ gibt.

Das Minimalpolynom μ_A ist als das Minimalpolynom des Ideals I_A definiert.

Theorem 7.1.7. *Sei $I \subset F[X]$ ein Ideal mit $I \neq \{0\}$. Dann gibt es ein Minimalpolynom von I.*

Beweis. Sei $0 \neq \mu \in I$ ein Polynom minimalen Grades in I, ohne Einschränkung normiert. Sei $p \in I$ beliebig. Mit Polynomdivision erhalten wir

$$p(X) = q(X) \cdot \mu(X) + r(X)$$

mit $\deg r < \deg \mu$. Aus dieser Gleichung folgt aber $r(X) \in I$. Da wir $\mu(X)$ mit minimalem Grad gewählt haben, muss $r(X) = 0$ gelten. Die Behauptung folgt. $\square$

Beispiel 7.1.8. Sei

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

oder allgemeiner eine $n \times n$-Matrix mit d Einsen wie angedeutet auf der Nebendiagonalen. Dann gilt $\chi_A(X) = (-1)^n X^n$ für das charakteristische Polynom und $\mu_A(X) = X^{d+1}$ für das Minimalpolynom.

Theorem 7.1.9. *Sei $f \colon V \to V$ ein Endomorphismus mit $\dim V = n$. Sei μ_f das Minimalpolynom von f. Dann gilt*

(i) μ_f teilt χ_f, d. h. es gibt ein $r \in F[X]$ mit $\mu_f \cdot r = \chi_f$.
(ii) χ_f teilt μ_f^n.

Beweis.

(i) Dies folgt aus der Definition des Minimalpolynoms, Theorem 7.1.7 und dem Satz von Cayley-Hamilton: $\chi_A \in I_A$, wobei A eine f darstellende Matrix ist.
(ii) Wir wollen das Resultat der Algebra benutzen, dass es einen Körper gibt, in dem χ_f und μ_f in Linearfaktoren zerfallen und ohne Einschränkung annehmen, dass dies über F bereits der Fall ist. (Im Falle $F = \mathbb{C}$ wissen wir bereits, dass beide Polynome in Linearfaktoren zerfallen.)

Sei $(X - \lambda)$ ein Linearfaktor von χ_f und ξ ein zugehöriger Eigenvektor. Ist $\mu_A = (X - \mu_1) \cdots (X - \mu_k)$ mit $\mu_i \neq \lambda$ für $1 \leq i \leq k$, so gilt $(A - \mu_k \mathbb{1})\xi = (\lambda - \mu_k)\xi$ und somit per Induktion

$$\mu_A(A)\xi = (\lambda - \mu_1) \cdots (\lambda - \mu_k)\xi \neq 0.$$

Widerspruch zur Definition des Minimalpolynoms. Somit kommt mindestens ein Faktor $(X-\lambda)$ in $\mu_A(X)$ vor, also teilt $(X-\lambda)^n$ das Polynom $(\mu_A(X))^n$. Da $\chi_A(X)$ aber höchstens Grad n hat, enthält es auch den Linearfaktor $(X - \lambda)$ höchstens n-fach. Somit teilt $\chi_A(X)$ das Polynom $(\mu_A(X))^n$. Die Behauptung folgt. $\square$

Definition 7.1.10. Sei $f \colon V \to V$ ein Endomorphismus. Dann heißt f nilpotent, falls es ein $k \geq 1$ mit $f^k = 0$ gibt. (Entsprechend definieren wir nilpotente $(n \times n)$-Matrizen.)

Theorem 7.1.11. *Sei $f \colon V \to V$ ein Endomorphismus, $\dim V = n$. Dann sind die folgenden Aussagen äquivalent*

(i) *f ist nilpotent.*
(ii) *$f^d = 0$ für ein d mit $1 \le d \le n$.*
(iii) *$\chi_f(X) = \pm X^n$.*
(iv) *Es gibt eine Basis, so dass f durch eine obere Dreiecksmatrix mit Nullen auf der Diagonale dargestellt wird.*

Beweis.

(a) Sei f nilpotent. Dann gilt $f^k = 0$ für ein $k \in \mathbb{N}$. Also ist X^k im Ideal I_f der Polynome enthalten, die verschwinden, wenn man f einsetzt. Da das Minimalpolynom $\mu_f(X)$ das Polynom X^k teilt, folgt $\mu_f(X) = X^l$ für ein $l \in \mathbb{N}$. Der Grad des Minimalpolynomes ist $\le n$. Also ist $l \le n$, $\mu_f(f) = 0$ und daher $f^l = 0$ für ein $l \le n$.
(b) Gelte $f^d = 0$ für ein $1 \le d \le n$. Sei d minimal mit dieser Eigenschaft. Dann ist X^d das Minimalpolynom. Da das charakteristische Polynom das Polynom $\mu_f^n = X^{dn}$ teilt und den Grad n besitzt, folgt $\chi_f(X) = \pm X^n$.
(c) Sei $\chi_f(X) = \pm X^n$. Dann zerfällt das charakteristische Polynom in Linearfaktoren, f lässt sich also mit Hilfe eine trigonalisierbaren Matrix darstellen. Nur wenn diese ausschließlich die Diagonaleinträge 0 besitzt, hat $\chi_f(X)$ die Gestalt $\pm X^n$.
(d) Sei die zu f gehörige Matrix wie angegeben. Wir wollen per Induktion zeigen, dass f^k in dieser Basis durch eine obere Dreiecksmatrix dargestellt wird, bei der $(k-1)$ Nebendiagonalen ebenfalls nur Nullen enthalten: Sei f^k durch $\left(b^i_j\right)$ dargestellt. Dann gilt $b^i_j = 0$ für alle (i,j) mit $j \le i + (k-1)$ und $a^i_j = 0$ für alle (i,j) mit $j \le i$. Wir erhalten für die Matrix $\left(c^i_j\right)$, die f^{k+1} darstellt

$$c^i_j = \sum_{l=1}^{n} b^i_l a^l_j.$$

Aufgrund der Bedingung an a^l_j genügt es, über $l < j$ oder $l + 1 \le j$ zu summieren. Ebenso liefert die Bedingung an b^i_l, dass es genügt, über $l > i + (k-1)$ oder $l \ge i + k$ zu summieren. $c^i_j \ne 0$ ist also nur möglich, wenn es ein l mit $j \ge l+1 \ge i+k+1$ gibt. Dies ist äquivalent zu $j > i+k$. Also ist wie behauptet $c^i_j = 0$ für $j \le i + k$. □

7.2. **Jordansche Normalform.** Für Abbildungen oder Matrizen, deren charakteristisches Polynom zwar in Linearfaktoren zerfällt, deren geometrische und algebraische Vielfachheiten aber nicht übereinstimmen, die also nicht diagonalisierbar sind, gibt es eine Standarddarstellung, die Jordansche Normalform.

Sei hier stets $f\colon V \to V$ ein Endomorphismus, $\dim V = n < \infty$, $\chi_f(X) = \pm(X-\lambda_1)^{r_1} \cdots (X-\lambda_k)^{r_k}$ mit $\lambda_i \ne \lambda_j$ für $i \ne j$ und $\sum\limits_{l=1}^{k} r_l = n$. Sei f in einer festen Basis durch die Matrix A dargestellt. Da das charakteristische Polynom in Linearfaktoren zerfällt, lässt sich f trigonalisieren, siehe [7].

Definition 7.2.1. Sei $f\colon V \to V$ ein Endomorphismus, $\dim V < \infty$. Sei λ eine Nullstelle von $\chi_f(X)$ mit algebraischer Vielfachheit r. Dann ist der Hauptraum von f zum Eigenwert λ als

$$\ker(f - \lambda \mathbb{1})^r$$

definiert.

Theorem 7.2.2. *Nach Basiswechsel hat A die Gestalt*

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix},$$

wobei A_i Blöcke der Größe $r_k \times r_k$ sind. Weiterhin gilt

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

jedes A_i ist also eine obere Dreiecksmatrix mit λ_i's auf der Diagonalen.

Beweis. Da A nach Voraussetzung trigonalisierbar ist, dürfen wir ohne Einschränkung annehmen, dass A bereits eine obere Dreiecksmatrix mit Diagonaleinträgen wie angegeben ist.

Die Haupträume zu λ_i haben die Dimension r_i. Wir schreiben A in der Form

$$A = \begin{pmatrix} B & * & * \\ 0 & A_l & * \\ 0 & 0 & D \end{pmatrix},$$

wobei auf der Diagonalen quadratische Blöcke stehen. Wir erhalten

$$A^2 = \begin{pmatrix} B^2 & * & * \\ 0 & A_l^2 & * \\ 0 & 0 & D^2 \end{pmatrix} \quad \text{und per Induktion} \quad A^m = \begin{pmatrix} B^m & * & * \\ 0 & A_l^m & * \\ 0 & 0 & D^m \end{pmatrix}.$$

Wegen $(A_l - \lambda_l \mathbb{1})^{r_l} = 0$, vergleiche Theorem 7.1.11, liefert eine solche Rechnung für $A - \lambda_l \mathbb{1}$, dass $(A - \lambda_l \mathbb{1})^{r_l}$ die Gestalt

$$(A - \lambda_l \mathbb{1})^{r_l} = \begin{pmatrix} (B - \lambda_l \mathbb{1})^{r_l} & * & * \\ 0 & 0 & * \\ 0 & 0 & (D - \lambda_l \mathbb{1})^{r_l} \end{pmatrix}.$$

hat. Man betrachtet insbesondere die vier Blöcke links oben und sieht, dass die Dimension von $\ker(A - \lambda_l \mathbb{1})^{r_l} = r_l$ ist. Definiere V_i als den Hauptraum von A zum Eigenwert λ_i. (Beachte auch, dass dieser Hauptraum in $\langle e_1, \dots, e_{r_1 + \dots + r_l} \rangle$ enthalten ist.) Somit gilt

$$\dim V_1 + \ldots + \dim V_k = r_1 + \ldots + r_k = n.$$

Direkte Summenzerlegung. Wir behaupten, dass sogar

$$V_1 \oplus \ldots \oplus V_k = V$$

gilt. Wir beweisen per Induktion nach l:

$$V_1 \oplus \ldots \oplus V_l = \langle e_1, \dots, e_{r_1 + \dots + r_l} \rangle,$$

der zu den ersten l Kästchen gehörige Vektorraum, wird also bereits von den ersten l Haupträumen aufgespannt. Für $l = 1$ ist dies klar. Sei $0 \neq v \in V_{l+1}$ beliebig. Es genügt zu zeigen, dass $v \notin V_1 \oplus \ldots \oplus V_l$ oder $(V_1 \oplus \ldots \oplus V_l) \cap V_{l+1} = \{0\}$ gilt. Dann ist die Summe direkt und nach der Dimensionsformel für Unterräume gilt $\dim(V_1 + \ldots + V_{l+1}) = r_1 + \ldots + r_{l+1}$. Falls die Behauptung nicht richtig wäre hätten wir in Blockgestalt

$$0 = (A - \lambda_{l+1})^{r_{l+1}} v \equiv \begin{pmatrix} (B - \lambda_{l+1} \mathbb{1})^{r_{l+1}} & * & * \\ 0 & 0 & * \\ 0 & 0 & (D - \lambda_{l+1} \mathbb{1})^{r_{l+1}} \end{pmatrix} \begin{pmatrix} v_1 \\ 0 \\ 0 \end{pmatrix}.$$

Wie man an der untersten nicht verschwindenden Komponente von v_1 direkt abliest, kann dies nicht sein, da $(B-\lambda_{l+1}\mathbb{1})^{r_{l+1}}$ eine Trigonalmatrix mit nichtverschwindenden Diagonalelementen ist. Somit gilt

$$V = V_1 \oplus \ldots \oplus V_k.$$

Invarianz der Haupträume. Sei $g\colon V \to V$ ein beliebiger Endomorphismus. Dann folgt

$$\ker g \subset \ker g^2 \subset \ker g^3 \subset \ldots \subset V$$

und die Inklusion ist genau dann strikt, wenn sich die Dimensionen der Räume unterscheiden. Für die Dimensionen gilt

$$0 \le \dim\ker g \le \dim\ker g^2 \le \dim\ker g^3 \le \ldots \le \dim V = n.$$

Da es nur endlich viele mögliche Dimensionen gibt, muss in beiden Ketten ab einem gewissen $\tilde{l}$ Gleichheit gelten. Man kann aber so etwas auch direkt an der Blockdarstellung ablesen:

Betrachte $g = f - \lambda_l\mathbb{1}$. An der Blockdarstellung von $(A-\lambda_l\mathbb{1})^{r_l}$ liest man ab, dass

$$\ker(A-\lambda_l\mathbb{1})^{r_l} = \ker(A-\lambda_l\mathbb{1})^{r_l+1} = \ldots$$

gilt.

Sei $v \in V_i$. Wir behaupten, dass auch $Av \in V_i$ gilt, dass V_i also unter f invariant ist. Nach Voraussetzung ist $0 = (A-\lambda_i\mathbb{1})^{r_i}v$. Daher ist auch

$$0 = (A-\lambda_i\mathbb{1})^{r_i+1}v = (A-\lambda_i\mathbb{1})^{r_i}\cdot(A-\lambda_i\mathbb{1})v.$$

Somit ist auch $(A-\lambda_i\mathbb{1})v = Av - \lambda_i v \in V_i$. Wegen $v \in V_i$ und da V_i ein Vektorraum ist, erhalten wir $Av \in V_i$. Wählen wir eine Basis, die aus Basen der Haupträume besteht, so dass die Einschränkung auf jeden Hauptraum durch eine obere Dreiecksmatrix dargestellt ist, hat A also die Blockgestalt wie behauptet. □

Theorem 7.2.3. *Sei $A \in F^{n\times n}$ von der Gestalt*

$$\begin{pmatrix} \lambda & * & \ldots & * \\ 0 & \lambda & & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda \end{pmatrix}.$$

Dann gibt es einen Basiswechsel, so dass A die Gestalt

$$\begin{pmatrix} \lambda & \delta_1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & \delta_2 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & \delta_3 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & \ldots & 0 & \lambda & \delta_{n-2} & 0 \\ 0 & 0 & \ldots & 0 & 0 & \lambda & \delta_{n-1} \\ 0 & 0 & \ldots & 0 & 0 & 0 & \lambda \end{pmatrix},$$

hat, wobei $\delta_i \in \{0,1\}$, $1 \le i \le n-1$, gilt.

Ist $\delta_i = 1$ für alle i, so heißt A Jordankästchen der Größe n, $J_n = J_n(\lambda)$.

Beweis. Nach Subtraktion von $\lambda\cdot\mathbb{1}$ dürfen wir ohne Einschränkung $\lambda = 0$ annehmen. Dann ist A nilpotent. Sei $k \in \mathbb{N}$ maximal mit $A^k \neq 0$. Wähle ein Basis $\{r_l\}_{1\le l\le \dim\ker A^k}$ von $\ker A^k$ und ergänze sie mit a_i, $1 \le i \le i_0(k)$, zu einer Basis von V.

Wir behaupten, dass die Vektoren $\left\{a_i, Aa_i, \ldots, A^k a_i\right\}_{1\le i\le i_0(k)}$ linear unabhängig sind: Zunächst einmal sind die $\{A^l a_i\}_i$ für festes $0 \le l \le k$ linear unabhängig, denn aus $\sum\limits_i \lambda^i A^l a_i = 0$ folgt $\sum\limits_i \lambda^i A^k a_i = 0$. Die a_i's waren aber gerade so gewählt, dass sie eine Basis von $\ker A^k$ zu einer Basis von V ergänzen. Aus $\sum\limits_i \lambda^i a_i \in \ker A^k$ folgt also $\sum\limits_i \lambda^i a_i = 0$ und daher $\lambda_i = 0$ für alle i, da die a_i die Basis eines Unterraumes bilden.

Setze $W_k := \langle a_i \rangle_i$. Sei $b_l \in A^l W_k$ und gelte $\sum\limits_{l=0}^{k} b_l = 0$. Wähle $\alpha_l \in W_k$ mit $b_l = A^l \alpha_l$. Wir erhalten $\sum\limits_{l=0}^{k} A^l \alpha_l = 0$. Auf diese Gleichung wenden wir nacheinander $A^k, A^{k-1}, \ldots, A$ an und erhalten daraus nacheinander $A^k \alpha_0 = 0$, woraus $\alpha_0 = 0$ folgt, $A^k \alpha_1 = 0$, woraus $\alpha_1 = 0$ folgt, ... Somit gilt $b_l = 0$ für $0 \le l \le k$ und die gesamte Familie ist wie behauptet linear unabhängig.

Wir wollen nun A bezüglich der Basis

$$A^k a_1, A^{k-1} a_1, \ldots, A a_1, a_1, A^k a_2, \ldots, A a_2, a_2, \ldots, a_{i_0(k)}, r_1, \ldots, r_{n-(k+1)i_0(k)}$$

darstellen (es ist zwar richtig, dass man die obigen Vektoren r_i zur Basisergänzung verwenden kann, jedoch geht es auch mit einer anderen Ergänzung) und erhalten in Blockgestalt

$$\begin{pmatrix} J_{k+1}(0) & 0 & \ldots & 0 & * \\ 0 & J_{k+1}(0) & & 0 & * \\ \vdots & & \ddots & \vdots & * \\ 0 & 0 & \ldots & J_{k+1}(0) & * \\ 0 & 0 & \ldots & 0 & R \end{pmatrix}.$$

Dabei sind die Anzahl der Kästchen $J_k(0)$ und k eindeutig durch die Wahl des maximalen k und durch die Dimension des Raumes $\ker A^k$ bestimmt.

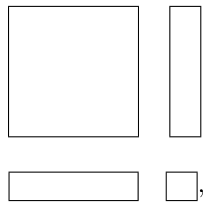
Wir wollen nun die Basis so abändern, dass die Sterne $*$ zu Nullen werden. Wir illustrieren das Vorgehen nur an einem Beispiel, im allgemeinen Fall muss man das Vorgehen hier lediglich auf jeden Block und jede Spalte über R anwenden. Im Beispiel betrachten wir $k = 3$ und die Situation, dass R nur ein 1×1-Block ist. Sei also ohne Einschränkung

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & p_1 \\ 0 & 0 & 1 & 0 & p_2 \\ 0 & 0 & 0 & 1 & p_3 \\ 0 & 0 & 0 & 0 & p_4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

bezüglich der Basis e_1, e_2, e_3, e_4, e_5. Ersetze nun e_5 durch $v := e_5 - p_1 e_2 - p_2 e_3 - p_3 e_4$. Es gilt $Av = \left(\begin{smallmatrix} -p_1+p_1 \\ -p_2+p_2 \\ -p_3+p_3 \\ p_4 \\ 0 \end{smallmatrix}\right) = p_4 e_4$. In dieser neuen Basis hat daher die Matrix die Gestalt

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & p_4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es gilt $B^4 e_5 = p_4 B^3 e_4 = p_4 B^2 e_3 = p_4 B e_2 = p_4 e_1$. Da aber das größte Jordankästchen in der obigen Blockgestalt nach Voraussetzung die Größe 4×4 hat, ist $B^4 = 0$. Somit ist $p_4 = 0$ und B hat die gewünschte Gestalt. Der allgemeine Fall funktioniert analog mit Induktion und der Betrachtung von symmetrisch angeordneten Untermatrizen der Form



wobei das Kästchen unten rechts die Größe 1×1 hat. $\square$

Wir kombinieren diese beiden Resultate und erhalten

Theorem 7.2.4. *Sei $f\colon V \to V$ ein Endomorphismus zwischen endlichdimensionalen Vektorräumen für den $\chi_f(X)$ in Linearfaktoren zerfällt. Dann gibt es eine Basis, so dass f in dieser Basis durch eine Blockmatrix dargestellt wird, die auf der Diagonalen Jordankästchen und sonst Nullen hat. Eine Matrix in dieser Form heißt in Jordanscher Normalform.*

Bis auf Permutation der Jordankästchen ist diese Matrix eindeutig bestimmt.

Beweis. Die Existenzaussage folgt aus den Theoremen 7.2.2 und 7.2.3.

Mögliche Werte der Diagonaleinträge und deren Häufigkeit stimmen mit den Nullstellen des charakteristischen Polynoms und deren Vielfachheit überein. Sei $j_k(\lambda)$ die Anzahl der Jordanblöcke $J_k(\lambda)$ in der Matrix A in Jordanscher Normalform. Es gilt

$$\dim\ker(J_k(\lambda) - \lambda\mathbb{1})^l = \begin{cases} l, & 1 \le l \le k, \\ k, & l \ge k. \end{cases}$$

Sei $n = \dim V$. Es folgt für $l \ge 1$

$$\dim\ker(A-\lambda\mathbb{1})^l = 1\cdot j_1(\lambda)+2\cdot j_2(\lambda)+\ldots+l\cdot j_l(\lambda)+l\cdot j_{l+1}(\lambda)+l\cdot j_{l+2}(\lambda)+\ldots+l\cdot j_n(\lambda).$$

Da

$$\begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & 2 & 2 & \ldots & 2 \\ 1 & 2 & 3 & \ldots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \ldots & n \end{pmatrix}$$

regulär ist (subtrahiere Zeile $(n-1)$ von Zeile n, dann Zeile $(n-2)$ von Zeile $(n-1)$, ... und erhalte eine obere Dreiecksmatrix mit Einsen auf der Diagonalen), sind die Zahlen $j_i(\lambda)$ durch $\dim\ker(A - \lambda\mathbb{1})^l$, $1 \le l \le n$, eindeutig bestimmt. Hieraus folgt die Eindeutigkeit bis auf Permutationen der Jordanblöcke. $\square$

8. Bilinearformen

8.1. **Bilinearformen, Quadratische Formen.**

Definition 8.1.1 (Bilinearform)**.**

(i) Seien $V_1, \ldots, V_n, W$ Vektorräume über F. Eine Abbildung

$$f\colon V_1 \times \ldots \times V_n \to W$$

heißt multilinear, falls für jedes $i \in \{1, \ldots, n\}$ und festes

$$(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \in V_1 \times \ldots \times V_{i-1} \times V_{i+1} \times \ldots \times V_n$$

die Abbildung

$$V_i \ni v_i \mapsto f(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)$$

linear ist.

(ii) Sei V ein F-Vektorraum. Eine multilineare Abbildung $\varphi\colon V \times V \to F$ heißt Bilinearform auf V. φ heißt symmetrisch, falls $\varphi(a,b) = \varphi(b,a)$ für alle $a, b \in V$ gilt.

Beispiele 8.1.2.

(i) Reelle Skalarprodukte sind symmetrische Bilinearformen.
(ii) $\varphi\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit $\varphi(\xi,\eta) := \xi^1\eta^1 - \xi^2\eta^2$ ist eine symmetrische Bilinearform auf $\mathbb{R}^2$, aber kein Skalarprodukt.
(iii) Sei V ein euklidischer Vektorraum und $f \in \operatorname{End}(V)$. Dann definiert $\varphi(\xi,\eta) := \langle \xi, f(\eta)\rangle$ für $\xi, \eta \in V$ eine Bilinearform.
(iv) Sei $A \in F^{n\times n}$. Dann ist $F^n \times F^n \ni (x,y) \mapsto x^T A y$ eine Bilinearform.
(v) Sei $V = C^0([a,b])$ und $k \in C^0\left([a,b]^2\right)$. Dann ist $\varphi\colon V \times V \to \mathbb{R}$, definiert durch
$$\varphi(f,g) := \int\limits_a^b \int\limits_a^b k(s,t)\cdot f(s)\cdot g(t)\; ds\, dt$$
eine Bilinearform auf V. φ ist genau dann symmetrisch, wenn $k(s,t) = k(t,s)$ für alle $a \le s, t \le b$ gilt.

Bemerkung 8.1.3.

(i) Seien $\varphi, \psi\colon V_1 \times \ldots \times V_n \to W$ zwei Multilinearformen. Der Raum aller Multilinearformen ist ein Vektorraum, $\operatorname{Hom}(V_1 \times \ldots \times V_n, W)$, wobei wir
$$(\varphi+\psi)(v_1,\ldots,v_n) := \varphi(v_1,\ldots,v_n) + \psi(v_1,\ldots,v_n)$$
und
$$(\alpha\cdot\varphi)(v_1,\ldots,v_n) := \alpha\cdot\varphi(v_1,\ldots,v_n)$$
setzen.
(ii) Sei $(a_1,\ldots,a_n)$ eine Basis von V. Sei φ eine Bilinearform auf V. Dann folgt für
$$\xi = \sum_{i=1}^n \xi^i a_i \quad \text{und} \quad \eta = \sum_{j=1}^n \eta^j a_j$$
wie bei linearen Abbildungen
$$\varphi(\xi,\eta) = \sum_{i,j=1}^n \xi^i \eta^j \varphi(a_i,a_j).$$
Daher ist φ durch $(c_{ij})_{1\le i,j\le n}$ mit $c_{ij} := \varphi(a_i,a_j)$ eindeutig bestimmt.
(iii) Umgekehrt definiert eine Matrix (c_{ij}) unter Verwendung der obigen Formel mit c_{ij} statt $\varphi(a_i,a_j)$ eine eindeutig bestimmte Bilinearform.
(iv) Wie bei linearen Abbildungen ist die Zuordnung $\varphi \mapsto (c_{ij})$ bei fixierter Basis $(a_1,\ldots,a_n)$ ein Vektorraumisomorphismus zwischen dem Vektorraum der Bilinearformen auf V und den $(n\times n)$-Matrizen über F. Insbesondere hat der Vektorraum der Bilinearformen auf einem n-dimensionalen Vektorraum V die Dimension n^2.
(v) Eine Bilinearform φ ist genau dann symmetrisch, wenn die zugehörige Matrix symmetrisch ist.

Bemerkung 8.1.4 (Transformationsverhalten bilinearer Abbildungen)**.**
Seien $(a_1,\ldots,a_n)$ und $(b_1,\ldots,b_n)$ Basen eines Vektorraumes V. Sei φ eine Bilinearform auf V. Sei $A = (\alpha_{ij})$ die zu φ bezüglich der Basis aus den a_i's gehörige Matrix und $B = (\beta_{ij})$ die zu den b_i's. Gelte
$$b_k = \sum_{i=1}^n d_k^i a_i$$

für $1 \le k \le n$. Setze $D := (d^i_j)$. Dann folgt

$$\beta_{kl} = \varphi(b_k, b_l) = \varphi\left(\sum_{i=1}^n d^i_k a_i, \sum_{j=1}^n d^j_l a_j\right) = \sum_{i,j=1}^n d^i_k d^j_l \varphi(a_i, a_j) = \sum_{i,j=1}^n d^i_k d^j_l \alpha_{ij}.$$

In Matrixform gilt also $B = D^T AD$. Dies ist dasselbe Transformationsverhalten wie beim reellen Skalarprodukt.

Definition 8.1.5 (Quadratische Form)**.** Sei $\varphi\colon V \times V \to F$ eine Bilinearform. Dann heißt $Q\colon V \to F$, definiert durch

$$Q(v) := \varphi(v, v) \quad \text{für } v \in V,$$

die zu φ zugehörige quadratische Form.

Beispiele 8.1.6.
(i) Sei φ die zur Matrix $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ gehörige Bilinearform. Dann gilt für die zugehörige quadratische Form Q die Identität $Q(v) = 0$ für beliebige $v \in V$.
(ii) Sei $\varphi = \langle \cdot, \cdot \rangle$. Dann gilt $Q = \|\cdot\|^2$ für die zugehörige quadratische Form Q.

Theorem 8.1.7. *Sei F ein Körper mit Charakteristik* $\operatorname{char} F \neq 2$. *Sei $Q\colon V \to F$ die zur Bilinearform $\varphi\colon V \times V \to F$ gehörige quadratische Form. Definiere $\psi\colon V \times V \to F$ durch*

$$\psi(\xi, \eta) := \tfrac{1}{2}(Q(\xi + \eta) - Q(\xi) - Q(\eta)) \quad \textit{für alle } \xi, \eta \in V.$$

(i) Es gilt

$$\psi(\xi, \eta) = \tfrac{1}{2}(\varphi(\xi, \eta) + \varphi(\eta, \xi)).$$

(ii) ψ ist eine weitere Bilinearform, zu der die quadratische Form Q gehört.
(iii) Ist φ symmetrisch, so gilt $\varphi = \psi$.
(iv) Ist φ bezüglich einer Basis die Matrix (a_{ij}) zugeordnet, so ist ψ bezüglich dieser Basis die Matrix $D = (d_{ij}) = \left(\frac{1}{2}(a_{ij} + a_{ji})\right)_{1 \le i,j \le n}$ zugeordnet.

Beweis. Wir zeigen nur die ersten beiden Teile.
(i) Es gilt

$$\begin{aligned} Q(\xi + \eta) &= \varphi(\xi + \eta, \xi + \eta) = \varphi(\xi, \xi) + \varphi(\xi, \eta) + \varphi(\eta, \xi) + \varphi(\eta, \eta) \\ &= Q(\xi) + \varphi(\xi, \eta) + \varphi(\eta, \xi) + Q(\eta). \end{aligned}$$

Wir erhalten daraus

$$\psi(\xi, \eta) = \tfrac{1}{2}(\varphi(\xi, \eta) + \varphi(\eta, \xi)).$$

(ii) Dies folgt aus

$$\psi(\xi, \xi) = \tfrac{1}{2}(\varphi(\xi, \xi) + \varphi(\xi, \xi)) = \varphi(\xi, \xi) = Q(\xi). \qquad \square$$

Korollar 8.1.8. *In einem reellen Skalarproduktraum V gilt die Polarisationsformel*

$$\langle x, y \rangle = \tfrac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right) \quad \textit{für alle } x, y \in V.$$

8.2. **Bilinearformen in euklidischen Vektorräumen.** Beliebige Bilinearformen lassen sich mit Hilfe des Skalarproduktes und mit linearen Abbildungen wie folgt darstellen:

Theorem 8.2.1. *Sei φ eine Bilinearform auf einem endlichdimensionalen reellen Vektorraum V. Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V, so gibt es eindeutig bestimmte lineare Abbildungen $f, g\colon V \to V$ mit*

$$\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \langle g(\xi), \eta \rangle$$

für alle $\xi, \eta \in V$. Die Abbildung f (und damit auch g) ist genau dann selbstadjungiert, wenn φ symmetrisch ist.

Beweis. Sei $\{a_1, \ldots, a_n\}$ eine Orthonormalbasis von V. Definiere

$$c_{ik} := \varphi(a_i, a_k) \quad \text{für } 1 \le i, k \le n.$$

Definiere die linearen Abbildungen $f, g\colon V \to V$ durch

$$f(a_k) := \sum_{j=1}^{n} c_{jk} a_j = \sum_{j,l=1}^{n} c_{lk} \delta^{lj} a_j, \quad 1 \le k \le n,$$

$$g(a_i) := \sum_{j=1}^{n} c_{ij} a_j = \sum_{j,l=1}^{n} c_{il} \delta^{lj} a_j, \quad 1 \le i \le n.$$

Dann folgt, da die Vektoren a_i eine Orthonormalbasis bilden,

$$\langle a_i, f(a_k) \rangle = \left\langle a_i, \sum_{j=1}^{n} c_{jk} a_j \right\rangle = \sum_{j=1}^{n} c_{jk} \delta_{ij} = c_{ik} = \varphi(a_i, a_k),$$

$$\langle g(a_i), a_k \rangle = \left\langle \sum_{j=1}^{n} c_{ij} a_j, a_k \right\rangle = \sum_{j=1}^{n} c_{ij} \delta_{jk} = c_{ik} = \varphi(a_i, a_k).$$

Da beide Seiten in beiden Gleichungen bilinear sind, erhalten wir

$$\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \langle g(\xi), \eta \rangle$$

für alle $\xi, \eta \in V$.

Die Abbildung f ist eindeutig bestimmt; sei nämlich $\tilde{f}$ eine weitere solche Abbildung, so erhalten wir $\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \left\langle \xi, \tilde{f}(\eta) \right\rangle$ für alle $\xi, \eta \in V$. Weiterhin folgt $0 = \left\langle \xi, f(\eta) - \tilde{f}(\eta) \right\rangle$ für alle $\xi, \eta \in V$ und damit $f(\eta) - \tilde{f}(\eta) = 0$ für alle $\eta \in V$. Also gilt $f = \tilde{f}$. Ebenso folgt die Eindeutigkeit für g.

Sei φ symmetrisch. Dann erhalten wir für $\xi, \eta \in V$

$$\langle \xi, f(\eta) \rangle = \varphi(\xi, \eta) = \varphi(\eta, \xi) = \langle \eta, f(\xi) \rangle = \langle f(\xi), \eta \rangle.$$

Somit ist f selbstadjungiert. Sei umgekehrt f selbstadjungiert. Es folgt

$$\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \langle f(\xi), \eta \rangle = \langle \eta, f(\xi) \rangle = \varphi(\eta, \xi).$$

Somit ist φ symmetrisch. $\square$

Bemerkung 8.2.2.

(i) Den obigen Satz kann man umformulieren, indem man statt eines Skalarproduktes eine Orthogonalbasis vorgibt.

(ii) f und g hängen vom Skalarprodukt ab. Die Eigenschaft, selbstadjungiert zu sein oder nicht, ist jedoch davon unabhängig, da sie äquivalent zur Symmetrie von φ ist und um die Symmetrie von φ zu überprüfen, benötigt man kein Skalarprodukt.

(iii) Die f und die das Skalarprodukt darstellenden Matrizen stimmen überein. Zu g gehört die dazu transponierte Matrix.

Definition 8.2.3. Sei $\varphi\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann definieren wir den Nullraum von φ durch

$$N(\varphi) := \{\xi \in V : \varphi(\xi, \eta) = 0 \quad \text{für alle } \eta \in V\}.$$

Eine symmetrische Bilinearform heißt ausgeartet, falls $N(\varphi) \ne \{0\}$ gilt.

Bemerkung 8.2.4.

(i) $N(\varphi)$ ist ein Unterraum von V.

(ii) Da φ symmetrisch ist, gilt auch

$$N(\varphi) = \{\xi \in V : \varphi(\eta, \xi) = 0 \quad \text{für alle } \eta \in V\}.$$

(iii) Eine symmetrische Bilinearform φ ist genau dann nicht ausgeartet, wenn es zu jedem $v \in V$ mit $v \neq 0$ ein $w \in V$ mit $\varphi(v, w) \neq 0$ gibt. (Kleine Übung)

Beispiele 8.2.5.
(i) Fassen wir ein Skalarprodukt als Bilinearform auf, so gilt $N = \{0\}$.
(ii) Sei die bilineare Abbildung $\varphi\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ bezüglich der Standardbasis durch die Matrix
$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$
gegeben. Dann gilt $N(\varphi) = \langle \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \rangle$.

Theorem 8.2.6. *Sei $\varphi\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf einem euklidischen Vektorraum V. Gilt $\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle$ für alle $\xi, \eta \in V$, so folgt auch $N = N(\varphi) = \ker f$.*

Beweis. Sei $0 = \varphi(\xi, \eta)$ für alle $\eta \in V$. Dann folgt $0 = \varphi(\xi, \eta) = \varphi(\eta, \xi) = \langle \eta, f(\xi) \rangle$ auch $f(\xi) = 0$. Die umgekehrte Inklusion funktioniert analog. □

Der Kern einer einer symmetrischen Bilinearform zugeordneten linearen Abbildung f ist daher im Gegensatz zur Abbildung f selbst vom verwendeten Skalarprodukt unabhängig. Daher ist die folgende Definition gerechtfertigt

Definition 8.2.7. Sei φ eine symmetrische Bilinearform. Dann definieren wir den Rang von φ, $\operatorname{rang} \varphi$ als den Rang einer zugehörigen linearen Abbildung f (bezüglich eines beliebigen Skalarproduktes). Wir nennen φ regulär (bzw. singulär), wenn f regulär (bzw. singulär) ist.

Bemerkung 8.2.8.
(i) Es gilt $\operatorname{rang} \varphi = \dim V - \dim N(\varphi)$.
(ii) Ist φ durch C dargestellt, so gilt $\operatorname{rang} \varphi = \operatorname{rang} C$.

Beispiele 8.2.9.
(i) Die zu einem Skalarprodukt gehörige Bilinearform ist regulär.
(ii) Bezüglich der Standardbasis des $\mathbb{R}^3$ definiert die Matrix
$$C = \begin{pmatrix} 10 & 4 & 8 \\ 4 & -8 & -4 \\ 8 & -4 & 1 \end{pmatrix}$$
eine Bilinearform φ. Es gilt $\operatorname{rang} C = 2 = \operatorname{rang} \varphi$ sowie
$$N(\varphi) = \left\langle \left(\begin{smallmatrix} 2 \\ 3 \\ -4 \end{smallmatrix}\right) \right\rangle.$$

8.3. **Hauptachsentransformation.**

Sei V ein euklidischer Raum. Das Gram-Schmidtsche Orthogonalisierungsverfahren liefert eine Orthogonalbasis von V. Bezüglich dieser Basis ist das Skalarprodukt durch die Einheitsmatrix dargestellt.

Ziel dieses Abschnittes ist es zu zeigen, dass es zu jeder symmetrischen Bilinearform auf V eine Basis von V gibt, so dass die Bilinearform durch eine Diagonalmatrix dargestellt wird. Wir können sogar eine Orthogonalbasis wählen.

Theorem 8.3.1. *Sei $\varphi\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf einem endlichdimensionalen euklidischen Vektorraum V. Sei $\dim V = n$. Dann existiert eine Orthonormalbasis $\{b_1, \ldots, b_n\}$ von V, bezüglich der φ durch eine Diagonalmatrix dargestellt wird. Die Diagonalelemente sind gerade die Eigenwerte λ_i der Matrix von φ (als Endomorphismus betrachtet) bezüglich einer beliebigen Orthonormalbasis von V.*

Beweis. Wir benutzen die lineare Abbildung $f\colon V \to V$, so dass $\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle$ für alle $\xi, \eta \in V$ gilt. Da φ symmetrisch ist, ist f selbstadjungiert. Somit gibt es eine Orthonormalbasis $\{b_1, \ldots, b_n\}$ von V aus Eigenvektoren von f zu Eigenwerten $(\lambda_i)_i$. Es gilt

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

Somit ist φ bezüglich dieser Basis durch eine Diagonalmatrix C dargestellt und die Diagonaleinträge sind die Eigenwerte von f.

Bezüglich einer weiteren Orthogonalbasis ist φ durch $C' = E^T C E$ dargestellt, wobei E eine Basiswechselmatrix ist. Da beide Basen Orthonormalbasen sind, ist E orthogonal und es gilt $E^T = E^{-1}$. Wegen $C' = E^{-1} C E$ sind C und C' ähnlich und besitzen daher die gleichen Eigenwerte. □

Beispiel 8.3.2. Betrachte die symmetrische Bilinearform φ auf $\mathbb{R}^3$, die bezüglich der Standardbasis durch die Matrix

$$C = \begin{pmatrix} 4 & -5 & -2 \\ -5 & 4 & -2 \\ -2 & -2 & -8 \end{pmatrix}$$

gegeben ist. Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ die durch C bezüglich der Standardbasis dargestellte lineare Abbildung. Die Eigenwerte von f sind die Nullstellen des zugehörigen charakteristischen Polynoms (etwas Rechenarbeit)

$$\chi_f(X) = \chi_C(X) = -(X+9)(X-9)X.$$

Bezüglich einer Orthonormalbasis $\{b_1, b_2, b_3\}$ aus Eigenvektoren ist f also durch

$$C' = \begin{pmatrix} 9 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

dargestellt. Orthonormale Eigenvektoren sind

$$b_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad b_2 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} \quad \text{und} \quad b_3 = \frac{1}{3} \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}.$$

In diesem Fall hat die Gleichung $C' = E^T C E$ die Gestalt

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 4 & -5 & -2 \\ -5 & 4 & -2 \\ -2 & -2 & -8 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{pmatrix}.$$

Lemma 8.3.3. *Sei $\varphi\colon V \times V$ eine symmetrische Bilinearform auf einem endlichdimensionalen euklidischen Vektorraum. Dann gibt es eine Basis, so dass die darstellende Matrix die folgende Diagonalgestalt hat:*

$$\begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& 0 && \\ &&&&&&& \ddots & \\ &&&&&&&& 0 \end{pmatrix}$$

Beweis. Benutze zunächst eine Basis $\{b_1, \ldots, b_n\}$ wie in Theorem 8.3.1. Nummeriere die Eigenwerte ohne Einschränkung so, dass $\lambda_1, \ldots, \lambda_s$ positiv, $\lambda_{s+1}, \ldots, \lambda_t$ negativ und $\lambda_{t+1}, \ldots, \lambda_n$ Null sind. Definiere die Vektoren

$$\begin{aligned} a_i &:= \frac{1}{\sqrt{\lambda_i}} b_i, && 1 \le i \le s, \\ a_i &:= \frac{1}{\sqrt{-\lambda_i}} b_i, && s+1 \le i \le t, \\ a_i &:= b_i, && t+1 \le i \le n. \end{aligned}$$

Die Menge $\{a_1, \ldots, a_n\}$ besteht aus paarweise zueinander orthogonalen Vektoren, von denen keiner der Nullvektor ist. Somit handelt es sich um eine Basis von V. Bezüglich dieser Basis hat φ aufgrund des Transformationsverhaltens von Bilinearformen darstellenden Matrizen die gewünschte Darstellung. □

Die Zahlen s und t hängen nicht von der speziellen Basiswahl ab.

Theorem 8.3.4 (Sylvesterscher Trägheitssatz)**.** *Sei Q eine quadratische Form in einem reellen endlich dimensionalen Vektorraum V. Dann gibt es eine Basis $\{a_1, \ldots, a_n\}$ von V, so dass Q in dieser Basis durch*

$$Q(\xi) = \sum_{i=1}^{s} \left(\xi^i\right)^2 - \sum_{i=s+1}^{t} \left(\xi^i\right)^2$$

mit $\xi = \sum\limits_{i=1}^{n} \xi^i a_i$ gegeben ist. s und t hängen dabei nur von Q ab.

Beweis. Die Existenz einer solchen Darstellung haben wir gerade in Skalarprodukträumen gezeigt. Wir können hier ein beliebiges Skalarprodukt auf V benutzen um das letzte Resultat anzuwenden. Es bleibt also, die letzte Aussage zu zeigen. Der Nullraum einer Bilinearform ist von der Basis unabhängig. Somit hängt $t = \dim V - \dim N$ nur von Q und nicht von der Basis ab.

Sei $\alpha_1, \ldots, \alpha_n$ eine weitere Basis von V zu der es σ mit

$$Q(\eta) = \sum_{i=1}^{\sigma} \left(\eta^i\right)^2 - \sum_{i=\sigma+1}^{t} \left(\eta^i\right)^2$$

für $\eta = \sum\limits_{i=1}^{n} \eta^i \alpha_i$ gibt. Wir behaupten, dass $s = \sigma$ gilt. Setze

$$\begin{aligned} U &:= \langle a_{s+1}, \ldots, a_t \rangle, \\ W &:= \langle \alpha_1, \ldots, \alpha_\sigma, \alpha_{t+1}, \ldots, \alpha_n \rangle. \end{aligned}$$

Es gilt

$$\begin{aligned} Q(\xi) &< 0 && \text{für } \xi \in U \setminus \{0\}, \\ Q(\eta) &\ge 0 && \text{für } \eta \in W. \end{aligned}$$

Somit ist $U \cap W = \{0\}$. Wir erhalten

$$n \ge \dim(U + W) = \dim U + \dim W = (t - s) + n - (t - \sigma) = n - s + \sigma.$$

Es folgt $s \ge \sigma$. Analog folgt $s \le \sigma$. Daraus erhält man die Behauptung. □

Definition 8.3.5. In der Notation des Sylvesterschen Trägheitssatzes nennen wir eine quadratische Form Q bzw. die (nach Theorem 8.1.7) zugehörige symmetrische Bilinearform φ mit $Q(\xi) = \varphi(\xi, \xi)$

(i) positiv definit, wenn $s = t = n$ gilt. Alternativ definiert man Q als positiv definit, wenn $Q(\xi) > 0$ für alle $\xi \in V$ mit $\xi \neq 0$ gilt.

(ii) negativ definit, wenn $s = 0$ und $t = n$ gelten. Alternativ: Wenn $-Q$ positiv definit ist.
(iii) positiv semidefinit, wenn $s = t$ gilt. Alternativ: Wenn $Q(\xi) \geq 0$ für alle $\xi \in V$ gilt.
(iv) negativ semidefinit, wenn $s = 0$ gilt. Alternativ: Wenn $-Q$ positiv semidefinit ist.
(v) indefinit, wenn $0 < s < t$ gilt. Alternativ, wenn es $\xi, \eta \in V$ mit $Q(\xi) < 0 < Q(\eta)$ gibt.

Beispiele 8.3.6.
(i) Eine symmetrische Bilinearform φ ist genau dann ein Skalarprodukt, wenn φ positiv definit ist.
(ii) Betrachte die symmetrische Bilinearform auf $\mathbb{R}^4$ mit zugehöriger quadratischer Form

$$Q(\xi) = 2\xi^1\xi^2 + 2\xi^1\xi^3 + 2\xi^1\xi^4 + 2\xi^2\xi^3 + 2\xi^2\xi^4 + 2\xi^3\xi^4.$$

Die zugehörige Matrix ist

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Das zugehörige charakteristische Polynom lautet (eine Seite Rechnung)

$$\chi(X) = X^4 - 6X^2 - 8X - 3 = (X+1)^3(X-3).$$

Somit ist die Matrix der Bilinearform bezüglich einer wie im Sylvesterschen Trägheitssatz geeignet normierten Basis durch

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

gegeben. Q ist also indefinit.

Theorem 8.3.7. *Seien Q_1 und Q_2 zwei quadratische Formen auf einem reellen Vektorraum V. Sei Q_1 positiv definit. Dann gibt es eine Basis $a_1, \ldots, a_n$ von V und $c_1, \ldots, c_n \in \mathbb{R}$, so dass*

$$Q_1(\xi) = \sum_{i=1}^n \left(\xi^i\right)^2, \quad \textit{und} \quad Q_2(\xi) = \sum_{i=1}^n c_i \left(\xi^i\right)^2$$

für $\xi = \sum\limits_{i=1}^n \xi^i a_i$ gilt.

Beweis. Die zu Q_1 gehörige symmetrische Bilinearform ist positiv definit und daher ein Skalarprodukt auf V. Nach Theorem 8.3.1 gibt es daher eine bezüglich dieses Skalarproduktes orthonormale Basis $\{a_1, \ldots, a_n\}$, in der die symmetrische Bilinearform zu Q_2 durch eine Diagonalmatrix dargestellt wird. □

8.4. **Extremaleigenschaften der Eigenwerte.** Vermöge

$$\varphi(\xi, \xi) = \langle \xi, f(\xi) \rangle$$

übertragen sich die folgenden Ergebnisse auch auf symmetrische Bilinearformen.

Wir setzen $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\| = 1\}$ bezüglich der vom Standardskalarprodukt auf $\mathbb{R}^n$ induzierten Norm.

Für das folgende Theorem benötigen wir einen kleinen Hilfssatz:

Lemma 8.4.1. *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch mit Eigenwerten $\lambda_1 \le \ldots \le \lambda_n \in \mathbb{R}$. Sei U ein k-dimensionaler Unterraum von $\mathbb{R}^n$. Wir verwenden das Standardskalarprodukt auf $\mathbb{R}^n$. Dann gibt es ein $x \in U$ mit $\langle x, x\rangle = \|x\|^2 = 1$ und $\langle x, Ax\rangle \ge \lambda_k$.*

Beweis. Wähle eine Orthonormalbasis $\{a_i\}_{1\le i\le n}$ aus Eigenvektoren von A mit $Aa_i = \lambda_i a_i$. Setze $W := \langle a_k, \ldots, a_n\rangle$. Aus Dimensionsgründen ist $U \cap W \neq \{0\}$. Sei $x \in U \cap W$ mit $\|x\| = 1$. Aus $x \in W$ folgt $\langle Ax, x\rangle \ge \lambda_k$. Somit folgt die Behauptung. □

Theorem 8.4.2. *Sei f ein selbstadjungierter Endomorphismus auf $\mathbb{R}^n$. Sei $\mathcal{U}_k$ der Raum aller k-dimensionalen Unterräume von $\mathbb{R}^n$. Dann gilt für die nach Größe geordneten und entsprechend ihrer Vielfachheit aufgeführten Eigenwerte $\lambda_1 \le \ldots \le \lambda_n$*

$$\lambda_k = \inf_{U\in\mathcal{U}_k} \sup_{x\in\mathbb{S}^{n-1}\cap U} \langle x, f(x)\rangle.$$

Beweis. Das Supremum wird stets angenommen, da $\mathbb{S}^{n-1} \cap U$ kompakt (= beschränkt und abgeschlossen) ist und da $x \mapsto \langle x, f(x)\rangle$ stetig ist.

Sei $\{a_i\}_{1\le i\le n}$ eine Orthogonalbasis von $\mathbb{R}^n$ mit $f(a_i) = \lambda_i a_i$. Setze $U_1 := \langle a_1, \ldots, a_k\rangle$. Dann gilt für $x = \sum\limits_{i=1}^{k} x^i a_i \in U_1$ mit $\|x\| = 1$, also $x \in \mathbb{S}^{n-1}$

$$\begin{aligned}\langle x, f(x)\rangle &= \sum_{i,j=1}^{k} x^i x^j \langle a_i, f(a_j)\rangle = \sum_{i,j=1}^{k} x^i x^j \lambda_j \underbrace{\langle a_i, a_j\rangle}_{=\delta_{ij}} = \sum_{i=1}^{k} \left(x^i\right)^2 \lambda_i \\ &\le \lambda_k \sum_{i=1}^{k} \left(x^i\right)^2 = \lambda_k \|x\|^2 = \lambda_k.\end{aligned}$$

Insbesondere folgt also

$$\langle a_k, f(a_k)\rangle = \lambda_k = \sup_{x\in\mathbb{S}^{n-1}\cap U_1} \langle x, f(x)\rangle.$$

Lemma 8.4.1 zeigt, dass

$$\inf_{U\in\mathcal{U}_k} \sup_{x\in\mathbb{S}^{n-1}\cap U} \langle x, f(x)\rangle \ge \lambda_k = \sup_{x\in\mathbb{S}^{n-1}\cap U_1} \langle x, f(x)\rangle$$

gilt. Somit folgt Gleichheit in der obigen Ungleichung und damit die Behauptung. □

Bemerkung 8.4.3. Sei U ein Unterraum von $\mathbb{R}^n$. Sei f ein symmetrischer Endomorphismus von $\mathbb{R}^n$ mit Standardskalarprodukt. Dann gilt

$$\sup_{x\in\mathbb{S}^{n-1}\cap U} \langle x, f(x)\rangle = \sup_{\substack{x\in U\\ x\neq 0}} \frac{\langle x, f(x)\rangle}{\|x\|^2}.$$

Dieser letzte Quotient heißt Rayleigh-Quotient. Wir hätten das obige Theorem auch mit Hilfe des Rayleigh-Quotienten formulieren können.

Wir wollen noch eine weitere Möglichkeit kennen lernen, Eigenwerte als Extrema wiederzufinden. Das folgende Theorem kann auch aus Theorem 8.4.2 gefolgert werden. Wir geben trotzdem einen unabhängigen Beweis.

Theorem 8.4.4. *Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ bezüglich des Standardskalarproduktes selbstadjungiert. Seien $\lambda_1 \le \ldots \le \lambda_n$ die Eigenwerte von f. Dann gilt*

$$\lambda_n = \sup_{x\in\mathbb{S}^{n-1}} \langle x, f(x)\rangle = \sup_{\substack{x\in\mathbb{R}^n\\ x\neq 0}} \frac{\langle x, f(x)\rangle}{\|x\|^2}.$$

Beweis. Wir zeigen nur die erste Gleichheit.

Aufgrund der Kompaktheit von $\mathbb{S}^n$ wird das Supremum in einem Punkt $x_0 \in \mathbb{S}^{n-1}$ angenommen. Sei $\varepsilon > 0$ und $\gamma\colon (-\varepsilon, \varepsilon) \to \mathbb{S}^{n-1}$ eine C^1-Kurve mit $\gamma(0) = x_0$. Dann nimmt die Funktion $t \mapsto \langle \gamma(t), f(\gamma(t))\rangle$ in $t = 0$ ein Maximum an. Somit gilt $0 = \langle \gamma'(0), f(x_0)\rangle + \langle x_0, f(\gamma'(0))\rangle$. Sei $x_1 \in \mathbb{S}^{n-1}$ mit $\langle x_0, x_1\rangle = 0$. Dann ist $\gamma(t) := \cos t \cdot x_0 + \sin t \cdot x_1$ eine Kurve in $\mathbb{S}^{n-1}$ wie oben betrachtet mit $\gamma'(0) = x_1$. Da f selbstadjungiert ist erhalten wir

$$0 = \langle x_1, f(x_0)\rangle + \langle x_0, f(x_1)\rangle = 2\langle x_1, f(x_0)\rangle.$$

Somit ist $f(x_0)$ ein Vielfaches von x_0 und daher ist x_0 ein Eigenvektor. Da für jeden Eigenvektor $y \in \mathbb{S}^{n-1}$ von f zum Eigenwert λ

$$\lambda = \langle y, f(y)\rangle$$

gilt, liefert die obige Formel auch den größten Eigenwert und das Maximum wird an einem zugehörigen Eigenvektor angenommen. □

Auch diese Methode läßt sich auf die anderen Eigenwerte verallgemeinern.

Theorem 8.4.5. *Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ bezüglich des Standardskalarproduktes selbstadjungiert. Seien $\lambda_1 \le \ldots \le \lambda_n$ die Eigenwerte von f. Seien $a_{k+1}, \ldots, a_n$ Eigenvektoren zu den Eigenwerten $\lambda_{k+1}, \ldots, \lambda_n$. Setze $U := \langle a_{k+1}, \ldots, a_n\rangle$. Dann gilt*

$$\lambda_k = \sup_{x \in \mathbb{S}^{n-1} \cap U^\perp} \langle x, f(x)\rangle = \sup_{\substack{x \in U^\perp \\ x \neq 0}} \frac{\langle x, f(x)\rangle}{\|x\|^2}.$$

Beweis. Werde das Supremum in $a_k \in U^\perp$ angenommen. Sei $b \in U^\perp$ mit $\langle b, a_k\rangle = 0$ und $\|b\| = 1$. Betrachte eine Kurve wie oben, hier $\mathbb{R} \ni t \mapsto \cos t \cdot a_k + \sin t \cdot b$. Wiederum folgt $0 = \langle b, f(a_k)\rangle$ für alle Vektoren b wie angegeben. Sei $v \in U$ ein Eigenvektor zum Eigenwert λ. Dann gilt $\langle v, f(a_k)\rangle = \langle f(v), a_k\rangle = \lambda\langle v, a_k\rangle = 0$. Zusammengenommen folgt $\langle w, f(a_k)\rangle = 0$ für alle $w \in \mathbb{S}^{n-1} \cap \langle a_k\rangle^\perp$. Daher ist a_k ein Eigenvektor von f. Nach Konstruktion kann a_k nicht im Erzeugnis der Eigenvektoren $a_{k+1}, \ldots, a_n$ liegen. Für jeden Eigenvektor $x \in \mathbb{S}^{n-1}$ liefert $\langle x, f(x)\rangle$ gerade den zugehörigen Eigenwert. Somit folgt die obige Formel. □

Das folgende Theorem gilt auch für hermitesche Matrizen, siehe [5, Theorem 7.2.5]. Für positiv semidefinite Matrizen gilt es nicht mit „≥ 0“.

Theorem 8.4.6. *Sei $A = \left(a^i_j\right)_{1 \le i,j \le n} \in \mathbb{R}^{n \times n}$ symmetrisch. Dann sind die folgenden beiden Aussagen äquivalent:*

(i) A ist positiv definit,
(ii) alle Unterdeterminanten $\det\left(a^i_j\right)_{1 \le i,j \le k} \equiv \det A_k$, $k = 1, \ldots, n$, erfüllen $\det A_k > 0$ (und sind im hermiteschen Fall reell).

Beweis.
„(i) $\Longrightarrow$ (ii)“: Sei A positiv definit. Betrachte $A_k = \left(a^i_k\right)_{1 \le i,j \le k}$ für ein k mit $1 \le k \le n$. Dann ist auch A_k positiv definit. A_k ist diagonalisierbar und hat positive Eigenwerte. Somit ist auch das Produkt dieser Eigenwerte, $\det A_k$, positiv.

„(ii) $\Longrightarrow$ (i)“: Per Induktion. Sei die Aussage bereits für $(n-1) \times (n-1)$-Matrizen gezeigt. Nach Theorem 8.4.2, mit Infimum und Supremum vertauscht um die größeren Eigenwerte zuerst zu bekommen, besitzt A mindestens $(n-1)$ positive Eigenwerte. Das Produkt der Eigenwerte ist aber ebenfalls positiv. Somit sind alle Eigenwerte positiv. □

8.5. **Flächen im $\mathbb{R}^3$.** Wir geben nur die wichtigste Definition an.

Definition 8.5.1. Sei $\varphi\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ nicht ausgeartet. Seien $\lambda_1 \le \lambda_2 \le \lambda_3$ die Eigenwerte der zugehörigen selbstadjungierten linearen Abbildung. Dann heißt die Menge

$$\left\{x \in \mathbb{R}^3 \colon \varphi(x,x) = 1\right\}$$

(i) Ellipsoid, falls $\lambda_i > 0$ für $i = 1, 2, 3$ gilt. Die Werte $\frac{1}{\sqrt{\lambda_i}}$ heißen Halbachsen. Ein Ellipsoid heißt genau dann Rotationsellipsoid, wenn $\lambda_1 = \lambda_2$ oder $\lambda_2 = \lambda_3$ gilt. Ein Ellipsoid heißt Sphäre, falls $\lambda_1 = \lambda_2 = \lambda_3$ gilt.
(ii) einschaliges Hyperboloid, falls $\lambda_1 < 0 < \lambda_2 \le \lambda_3$ gilt. Es ist genau dann um den Eigenraum zu λ_1 rotationssymmetrisch, falls $\lambda_2 = \lambda_3$ gilt.
(iii) zweischaliges Hyperboloid, falls $\lambda_1 \le \lambda_2 < 0 < \lambda_3$ gilt. Es ist genau dann um den Eigenraum zu λ_3 rotationssymmetrisch, falls $\lambda_1 = \lambda_2$ gilt.

9. Teilbarkeit

9.1. **Teilbarkeit.** In diesem Kapitel werden wir insbesondere alle bereits in den Übungen benutzten Eigenschaften von Primzahlen zeigen. Wir benutzen [2].

Definition 9.1.1.
(i) Ein Integritätsring ist ein kommutativer Ring R mit Eins in dem aus $a \cdot b = 0$, $a, b \in R$, stets $a = 0$ oder $b = 0$ folgt (nullteilerfrei).
(ii) Ein Element $a \in R$ heißt Einheit, falls es $b \in R$ mit $ab = e$ gibt. Wir schreiben R^* für die Menge der Einheiten in R.

Beispiele 9.1.2.
(i) Sei F ein Körper. Dann ist $F[X]$ ein Integritätsring.
(ii) $\mathbb{Z}$ ist ein Integritätsring.

Definition 9.1.3 (Euklidischer Ring)**.** Ein Integritätsring heißt euklidischer Ring, wenn es eine Gradfunktion $\delta\colon R \setminus \{0\} \to \mathbb{N}$ gibt, so dass es zu $a, b \in R$ mit $b \neq 0$ (nicht notwendigerweise eindeutig bestimmte) Elemente $q, r \in R$ mit $a = qb + r$ und ($r = 0$ oder $\delta(r) < \delta(b)$) gibt.

Beispiele 9.1.4.
(1) Sei F ein Körper. Dann ist $F[X]$ ein euklidischer Ring mit der Gradfunktion $\delta = \deg$. Dies ist eine Folgerung aus der Polynomdivision.
(2) $\mathbb{Z}$ mit Absolutbetrag ist ein euklidischer Ring. Übung.

Definition 9.1.5 (Hauptidealring)**.**
(i) Sei R ein kommutativer Ring. Sei $a \in R$. Dann ist

$$(a) = Ra := \{ra \colon r \in R\}$$

ein Ideal (Übung), das von a erzeugte Hauptideal. Ein Ideal $I \subset R$ mit $(a) = I$ für ein $a \in R$ heißt (das von a erzeugte) Hauptideal.
(ii) Ein Integritätsring R heißt Hauptidealring, wenn jedes Ideal $I \subset R$ ein Hauptideal ist.

Theorem 9.1.6. *Sei R ein euklidischer Ring. Dann ist R ein Hauptidealring.*

Beweis. Sei $\mathfrak{a} \subset R$ ein Ideal. Wegen $(0) = \{0\}$ dürfen wir ohne Einschränkung annehmen, dass $\mathfrak{a}$ nicht das Nullideal ist. Sei $a \in \mathfrak{a} \setminus \{0\}$ ein beliebiges Element mit minimalem Grad. Wir behaupten, dass dann $\mathfrak{a} = (a)$ gilt. Die Inklusion $(a) \subset \mathfrak{a}$ ist klar. Sei also $b \in \mathfrak{a}$ beliebig. Da R ein euklidischer Ring ist, gibt es $p, r \in R$ mit $b = pa + r$ und ($r = 0$ oder $\delta(r) < \delta(a)$). Der zweite Fall ist ausgeschlossen, da $r = b - pa$ der Minimalität von a widerspricht. Also gilt $b = pa$. Die Behauptung folgt. $\square$

Korollar 9.1.7.

(i) *$\mathbb{Z}$ ist ein Hauptidealring.*
(ii) *Sei F ein Körper. Dann ist $F[X]$ ein Hauptidealring.*

Definition 9.1.8. Sei R ein Integritätsring. Seien $a, b \in R$.

(i) a teilt b oder a ist ein Teiler von b, wenn es ein $c \in R$ mit $ac = b$ gibt. Wir schreiben $a|b$. Ist a kein Teiler von b so schreiben wir $a \nmid b$.
(ii) a und b heißen assoziiert, wenn es eine Einheit $e \in R^*$ mit $ae = b$ gibt.

Lemma 9.1.9. *Sei R ein Integritätsring. Seien $a, b \in R$. Dann sind die folgenden Aussagen äquivalent:*

(i) $a|b$,
(ii) $b \in (a)$,
(iii) $(b) \subset (a)$.

Beweis. Übung. □

Lemma 9.1.10. *Sei R ein Integritätsring. Seien $a, b \in R$. Dann sind die folgenden Aussagen äquivalent:*

(i) $a|b$ *und* $b|a$,
(ii) $(a) = (b)$,
(iii) a *und* b *sind assoziiert.*

Beweis. Wir zeigen nur „(ii)$\Longrightarrow$(iii)“ und lassen den Rest als Übungsaufgabe.

Gelte $(a) = (b)$. Dann gibt es $c, d \in R$ mit $ac = b$ und $a = bd$. Ist $a = 0$, so folgt auch $b = 0$. Nehme daher ab jetzt $a \neq 0$ an. Es folgt $a = bd = acd$ oder $a(1 - cd) = 0$. Wegen $a \neq 0$ folgt $cd = 1$, da R ein Integritätsring ist. Also sind c und d Einheiten, $c, d \in R^*$. Somit sind a und b assoziiert. □

Definition 9.1.11. Sei R ein Integritätsring und $p \in R \setminus (R^* \cup \{0\})$.

(i) p heißt irreduzibel, wenn aus $p = ab$ mit $a, b \in R$ stets folgt, dass a oder b eine Einheit in R ist. Sonst heißt p reduzibel.
(ii) p heißt prim oder Primelement, wenn aus $p|ab$ stets $p|a$ oder $p|b$ folgt. Dies ist äquivalent dazu, dass für $\mathfrak{p} := (p)$ aus $ab \in \mathfrak{p}$ stets $a \in \mathfrak{p}$ oder $b \in \mathfrak{p}$ folgt. (Somit ist p genau dann ein Primelement, wenn das Ideal (p) ein Primideal ist.)

Bemerkung 9.1.12.

(i) $p \in R \setminus (R^* \cup \{0\})$ ist genau dann irreduzibel, wenn aus $p \in (a)$ für $a \in R$ stets $(a) = (p)$ oder $(a) = R$ folgt.
(ii) Sei $p \in R$ prim. Dann teilt p ein Produkt $a_1 \cdot \ldots \cdot a_k$ von Elementen $a_i \in R$ genau dann, wenn es einen der Faktoren teilt.

Beweis.

(i) „$\Longrightarrow$“: Sei p irreduzibel und gelte $p \in (a)$. Dann gibt es $r \in R$ mit $p = ar$. Somit ist a oder r eine Einheit. Ist a eine Einheit, so folgt $(a) = R$. Ist r eine Einheit, so folgt $(a) = (p)$.
   „$\Longleftarrow$“: Sei $p \in R \setminus (R^* \cup \{0\})$ beliebig. Gelte $p = ab$. Es folgt $p \in (a)$. Somit gilt $(a) = (p)$ oder $(a) = R$. Im ersten Fall erhalten wir $p = ar$ mit $r \in R^*$. Wegen $a \neq 0$ (sonst wäre auch $p = 0$) erhalten wir $b = r \in R^*$. Im zweiten Fall gilt $(a) = R$. Also ist a eine Einheit.
(ii) Induktion nach der Anzahl der Faktoren. □

Lemma 9.1.13. *Sei R ein Integritätsring. Dann ist jedes Primelement auch irreduzibel.*

Beweis. Sei $p \in R$ prim und seien $a, b \in R$ mit $p = ab$. Aus $p|ab$ folgt $p|a$ oder $p|b$, ohne Einschränkung $p|a$. Daher gibt es $c \in R$ mit $a = pc$. Es folgt $p = ab = pcb$ oder $p(1 - cb) = 0$. Da R ein Integritätsring ist, folgt $1 = cb$. Somit ist b eine Einheit. Daher ist p irreduzibel. □

Üblicherweise ist die Definition einer Primzahl die Definition einer irreduziblen Zahl. Dies ist aber folgenlos, da in Hauptidealringen Primelemente gerade irreduzible Elemente sind.

Theorem 9.1.14. *Sei R ein Hauptidealring. Sei $p \in R \setminus (R^* \cup \{0\})$. Dann sind die folgenden Aussagen äquivalent:*

(i) p ist irreduzibel,
(ii) p ist prim.

Beweis. Nach Lemma 9.1.13 brauchen wir nur „(i)$\Longrightarrow$(ii)“ zu zeigen.

Sei $p \in R$ irreduzibel. Seien $a, b \in R$. Gelte $p|ab$ und $p \nmid a$. Wir wollen $p|b$ zeigen. Betrachte das Ideal

$$Ra + Rp := \{ra + sp \colon r, s \in R\}.$$

Dies ist ein Ideal: Übung. Da R ein Hauptidealring ist, gilt $Ra + Rp = Rd = (d)$ für ein $d \in R$. Es folgt $a, p \in Rd$ und daher $d|a$ sowie $d|p$. Wegen $d|p$ gibt es $c \in R$ mit $p = cd$. p ist irreduzibel. Somit ist c oder d eine Einheit.

Ist c eine Einheit, so folgt $d = c^{-1}p$. Wegen $d|a$ erhalten wir auch $p|a$. Widerspruch.

Also ist d eine Einheit. Es folgt $Ra + Rp = Rd = R$. Daher gibt es $r, s \in R$ mit $ra + sp = 1$. Weiterhin gilt $rab + spb = b$. Wegen $p|ab$ und $p|spb$ folgt $p|b$. □

Korollar 9.1.15. *Sei R ein Hauptidealring. Sei $a \in R \setminus (R^* \cup \{0\})$. Dann ist a ein endliches Produkt von Primelementen.*

Beweis. Nach Theorem 9.1.14 genügt es, eine Faktorisierung von a in irreduzible Elemente herzuleiten. Nehme an, die Behauptung wäre falsch, d. h. wir könnten a nicht als endliches Produkt von irreduziblen Elementen schreiben. Dann ist a selber insbesondere nicht irreduzibel. Somit gibt es Elemente $a_1, a_1' \in R \setminus R^*$ mit $a = a_1 a_1'$. Wir haben angenommen, dass sich a nicht als endliches Produkt von irreduziblen Elementen darstellen lässt. Somit muss dies mindestens für einen der Faktoren a_1 oder a_1' ebenfalls gelten, ohne Einschränkung für a_1. Wir können nun a_1 als Produkt von zwei nicht-Einheiten $a_2, a_2' \in R \setminus R^*$ schreiben. Dies kann man fortsetzen und erhält eine Folge

$$a = a_0, a_1, a_2, \ldots \in R.$$

Dabei gilt jeweils $a_{i+1}|a_i$, aber a_{i+1} und a_i sind nicht assoziiert. Wir erhalten also eine aufsteigende Folge von Idealen

$$(a) = (a_0) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \ldots,$$

wobei die strikte Inklusion nach Lemma 9.1.10 gilt, da a_{i+1} und a_i nicht assoziiert sind. Setze nun $\mathfrak{b} := \bigcup\limits_{i=1}^{\infty} (a_i)$. Da eine beliebige Vereinigung von aufsteigenden Idealen wieder ein Ideal ist und da R ein Hauptidealring ist, folgt $\mathfrak{b} = (b)$ für ein $b \in R$. Nach Definition von $\mathfrak{b}$ gilt $b \in (a_{i_0})$ für ein $i_0 \in \mathbb{N}$. Wir erhalten

$$\mathfrak{b} = (b) \subset (a_{i_0}) \subsetneq (a_{i_0+1}) \subset \mathfrak{b}.$$

Widerspruch. Die Behauptung folgt. □

Für $R = \mathbb{Z}$ kann man auch einen direkten Beweis führen.

Lemma 9.1.16. *Sei R ein Integritätsring. Seien $p_1, \dots, p_r \in R \setminus (R^* \cup \{0\})$ prim und $q_1, \dots, q_s \in R \setminus (R^* \cup \{0\})$ irreduzibel. Gilt*

$$p_1 \cdot \ldots \cdot p_r = q_1 \cdot \ldots \cdot q_s,$$

so folgt $r = s$ und nach Umnummerierung der Elemente $q_1, \dots, q_r$ gibt es Einheiten $\varepsilon_1, \dots, \varepsilon_r \in R^*$, so dass $q_i = \varepsilon_i p_i$ für alle $1 \le i \le r$ gilt.

Beweis. Aus $p_1 \cdot \ldots \cdot p_r = q_1 \cdot \ldots \cdot q_s$ folgt insbesondere $p_1 | q_1 \cdot \ldots \cdot q_s$. Da p_1 prim ist, gibt es ein i, $1 \le i \le s$, mit $p_1 | q_i$, ohne Einschränkung nach Umnummerierung $i = 1$: $p_1 | q_1$. Es gilt also $q_1 = \varepsilon_1 p_1$ für ein $\varepsilon_1 \in R$. Da q_1 irreduzibel ist, ist ε_1 eine Einheit. Da R ein Integritätsring ist, folgt

$$p_2 \cdot \ldots p_r = \varepsilon_1 \cdot q_2 \cdot \ldots \cdot q_s.$$

Fahren wir wie oben fort, so erhalten wir

$$1 = \varepsilon_1 \cdot \ldots \cdot \varepsilon_r \cdot q_{r+1} \cdot \ldots \cdot q_s.$$

Alle Faktoren $q_{r+1}, \dots, q_s$ müssen somit Einheiten sein. Somit gilt $r = s$ und die Behauptung folgt. $\square$

Definition 9.1.17. Ein Integritätsring R heißt faktoriell, wenn sich jedes Element von $R \setminus (R^* \cup \{0\})$ als Produkt von Primelementen in R schreiben lässt.

Bemerkung 9.1.18. Die Produktdarstellung eines Elementes in $R \setminus (R^* \cup \{0\})$ ist nach Lemma 9.1.16 bis auf Einheiten eindeutig, da Primelemente irreduzibel sind.

Lemma 9.1.19. *Sei R ein faktorieller Ring. Sei $d \in R$ irreduzibel. Dann ist d prim.*

Beweis. Gelte $d | ab$. Schreibe $d = d_1 \cdots d_r$, $a = a_1 \cdots a_s$ und $b = b_1 \cdots b_t$ jeweils als Produkt von Primelementen. Dies ist nach Definition eines faktoriellen Ringes möglich. Da d irreduzibel ist, tritt in der Produktdarstellung von d mit Primelementen, also auch irreduziblen Elementen (Lemma 9.1.13), nur ein Faktor auf, der keine Einheit ist. Sei also ohne Einschränkung $r = 1$. Lemma 9.1.13 liefert, dass alle d_i, a_i und b_i's auch irreduzibel sind, da sie prim sind. Nach Lemma 9.1.16 ist d_1 zu einem a_i oder b_i assoziert. Daher folgt entweder $d | a$ oder $d | b$. Somit ist d prim. $\square$

Korollar 9.1.15 liefert

Theorem 9.1.20. *Jeder Hauptidealring ist faktoriell, insbesondere also*
(i) $\mathbb{Z}$,
(ii) $F[X]$ *für einen Körper* F.

Bemerkung 9.1.21. Sei R ein faktorieller Ring. Zerlege die Menge der Primelemente in $R \setminus (R^* \cup \{0\})$ in Äquivalenzklassen assoziierter Elemente und wähle jeweils einen Vertreter. Sei P die Menge dieser Vertreter. Dann gibt es zu jedem $a \in R \setminus \{0\}$ eine Darstellung der Form

$$a = \varepsilon \prod_{p \in P} p^{n(p)}$$

mit $\varepsilon \in R^*$ und $n(p) \in \mathbb{N}$, wobei das Produkt endlich ist.

In $\mathbb{Z}$ wählt man für P die Menge der Primelemente größer als Eins, die Primzahlen, und erhält die übliche Primfaktorzerlegung. ± 1 sind die einzigen Einheiten.

Die Einheiten in $F[X]$ sind $F \setminus \{0\} \subset F[X]$. Wir wählen daher P als die Menge der normierten Primpolynome, d. h. die Menge der Primpolynome in $F[X] \setminus ((F[X])^* \cup \{0\})$, deren höchster Koeffizient 1 ist.

Nach Wahl einer Vertretermenge P können wir zu

$$0 \neq a = \varepsilon \prod_{p \in P} p^{\mu_p} \quad \text{und} \quad 0 \neq b = \delta \prod_{p \in P} p^{\nu_p}$$

den größten gemeinsamen Teiler

$$\operatorname{ggT}(a,b) = \prod_{p\in P} p^{\min\{\mu_p,\nu_p\}}$$

und das kleinste gemeinsame Vielfache

$$\operatorname{kgV}(a,b) = \prod_{p\in P} p^{\max\{\mu_p,\nu_p\}}$$

definieren. Da die Elemente von P nur bis auf Einheiten eindeutig sind, sind ggT und kgV ebenfalls nur bis auf Einheiten wohldefiniert.

Den größten gemeinsamen Teiler kann man auch mit Hilfe von Idealen charakterisieren. Setze

$$Ra + Rb := \{ra + sb \colon r, s \in R\}.$$

Man überprüft direkt, dass dies ein Ideal ist. Der Bezug zum ggT folgt aus dem folgenden Theorem.

Theorem 9.1.22. *Sei R ein Hauptidealring. Seien $a, b \in R \setminus \{0\}$. Dann gilt*

$$Ra + Rb = Rd \quad \textit{mit} \quad d = \operatorname{ggT}(a,b).$$

Insbesondere gibt es Elemente $r, s \in R$ mit $ra + sb = d$. Es gilt $\operatorname{ggT}(r,s) = 1$.

Beweis. Das Ideal $Ra + Rb \subset R$ ist ein Hauptideal. Gelte ohne Einschränkung $Ra + Rb = Rd'$. Aus $a, b \in Rd'$ folgen $d'|a$ und $d'|b$. Somit gilt auch $d'|d$. Wegen $Ra + Rb = Rd'$ gibt es andererseits $r, s \in R$ mit $ra + sb = d'$. Somit teilt jeder gemeinsame Teiler von a und b auch d', es gilt also insbesondere $d|d'$. Aus $d'|d$ und $d|d'$ folgt, dass d und d' assoziiert sind. Also gilt die behauptete Gleichung $Ra + Rb = Rd' = Rd$. Somit gibt es $r, s \in R$ mit $ra + sb = d$. Sei t ein Teiler von a und b. Dann teilt $t \cdot \operatorname{ggT}(r,s)$ auch d. Setzt man insbesondere $t = d$, so erhält man, dass $\operatorname{ggT}(r,s)$ eine Einheit ist. Also folgt $\operatorname{ggT}(r,s) = 1$. □

Korollar 9.1.23. *Sei R ein Hauptidealring. Sei $p \in R \setminus (R^* \cup \{0\})$. Dann sind die folgenden Aussagen äquivalent:*

(i) p ist prim.
(ii) Der Restklassenring $R/(p)$ ist ein Körper.

Beweis.
(a) Sei p prim. Sei $a \notin (p)$, gelte also $p \nmid a$. Direkt nach Definition folgt dann $\operatorname{ggT}(p,a) = 1$. Nach Theorem 9.1.22 folgt also, dass $(p) \subset R$ maximal ist. Somit ist $R/(p)$ ein Körper.
(b) Sei p nicht prim. Dann gibt es $a, b \in R$ mit $p|ab$ aber $p \nmid a$ und $p \nmid b$. Daher folgt für die Restklassen in $R/(p)$ sowohl $\bar{a}\bar{b} = 0$ als auch $\bar{a} \neq 0$ und $\bar{b} \neq 0$. In einem Körper gibt es aber keine Nullteiler. Somit ist $R/(p)$ kein Körper. □

Bemerkung 9.1.24. Sei F ein Körper. Sei $\varphi \colon \mathbb{Z} \to F$ der Ringhomomorphismus mit $\varphi(1) = 1$. Definiere die Charakteristik von F als die kleinste Zahl c größer als 0 mit $\varphi(c) = 0$ und als 0, falls es keine solche Zahl gibt.

Analog zum obigen Beweis sieht man, das die Charakteristik eines Körpers 0 oder eine Primzahl ist.

Die folgenden Beispiele werden in der Algebra noch vertieft behandelt.

Beispiele 9.1.25.
(i) Sei $p \in \mathbb{Z}$. Dann ist der Restklassenring $\mathbb{Z}/p\mathbb{Z}$ genau dann ein Körper, wenn p prim ist.

(ii) Sei F ein Körper. Sei $f \in F[X]$. Dann ist $F[X]/(f)$ genau dann ein Körper, wenn f prim ist.

$$F[X]/(X-1) \cong F,$$
$$\mathbb{R}[X]/\left(X^2+1\right) \cong \mathbb{C}.$$

Es gibt einen zu $\mathbb{Q}[X]/(X^2-3)$ isomorphen Körper K (isomorph heißt hier: als Ring mit Eins isomorph) mit

$$\mathbb{Q} \subsetneq K \subsetneq \mathbb{R}.$$

Beweis: Übung.

10. Das Tensorprodukt

Tensorprodukte definiert man nicht nur für Vektorräume, sondern z. B. auch für Moduln über Ringen. Sei F ein fixierter Körper und alle Vektorräume seien F-Vektorräume.

10.1. **Definition und Existenz.**

Definition 10.1.1. Seien $E_1, \ldots, E_n, V$ Vektorräume. Sei $L^n(E_1, \ldots, E_n; V)$ der Vektorraum der multilinearen Abbildungen

$$\varphi \colon E_1 \times \ldots \times E_n \to V.$$

Wir schreiben $L^1(E_1; V) \equiv L(E_1, V) \equiv \operatorname{Hom}(E_1, V)$. Ein Vektorraum W mit einer multilinearen Abbildung $\tau \colon E_1 \times \ldots \times E_n \to W$ heißt Tensorprodukt von $E_1, \ldots, E_n$,

$$W = E_1 \otimes \ldots \otimes E_n,$$

falls die folgende Eigenschaft erfüllt ist: Sei V ein beliebiger Vektorraum. Zu jeder multilinearen Abbildung $\varphi \in L^n(E_1, \ldots, E_n; V)$, $\varphi \colon E_1 \times \ldots \times E_n \to V$, gibt es eine eindeutig bestimmte lineare Abbildung $\Phi \colon E_1 \otimes \ldots \otimes E_n \to V$, so dass das Diagramm

$$\begin{tikzcd} E_1 \times \ldots \times E_n \arrow[rr, "\tau"] \arrow[dr, "\varphi"'] & & E_1 \otimes \ldots \otimes E_n \arrow[dl, "\Phi"] \\ & V & \end{tikzcd}$$

kommutiert.

Theorem 10.1.2. *Das Tensorprodukt ist, falls es existiert, bis auf einen Isomorphismus eindeutig bestimmt.*

Beweis. Wäre $\langle \tau(E_1 \times \ldots \times E_n) \rangle \subsetneq E_1 \otimes \ldots \otimes E_n$, so könnte Φ nicht eindeutig bestimmt sein.

Seien $\tau \colon E_1 \times \ldots \times E_n \to E_1 \otimes \ldots \otimes E_n$ und $\vartheta \colon E_1 \times \ldots \times E_n \to W$ zwei Tensorprodukte. Wenden wir die definierende Eigenschaft des Tensorproduktes an, so erhalten wir lineare Abbildungen $\Theta \colon E_1 \otimes \ldots \otimes E_n \to W$ mit $\vartheta = \Theta \circ \tau$ und $T \colon W \to E_1 \otimes \ldots \otimes E_n$ mit $\tau = T \circ \vartheta$. Also erhalten wir $\vartheta = \Theta \circ T \circ \vartheta$ und $\tau = T \circ \Theta \circ \tau$. Ein beliebiges Element in $E_1 \otimes \ldots \otimes E_n$ können wir wegen $\langle \tau(E_1 \times \ldots \times E_n) \rangle = E_1 \otimes \ldots \otimes E_n$ in der Form $\sum\limits_{i=1}^{m} \lambda^i \tau(v_i)$ mit $\lambda_i \in F$, $v_i \in E_1 \times \ldots \times E_n$ für ein $m \in \mathbb{N}$ darstellen. Da T und Θ linear sind, erhalten wir

$$\sum_{i=1}^{m} \lambda^i \tau(v_i) = \sum_{i=1}^{m} \lambda^i T \circ \Theta \circ \tau(v_i) = T \circ \Theta \left(\sum_{i=1}^{m} \lambda^i \tau(v_i) \right).$$

Da $\sum \lambda^i \tau(v_i)$ ein beliebiges Element in $E_1 \otimes \ldots \otimes E_n$ ist, erhalten wir $T \circ \Theta = \mathrm{id}$. Analog erhalten wir $\Theta \circ T = \mathrm{id}$. T und Θ sind damit die gesuchten Isomorphismen. $\square$

Somit werden wir nachfolgend stets von dem Tensorprodukt sprechen und das nachfolgend Konstruierte meinen.

Theorem 10.1.3. *Seien $E_1, \ldots, E_n$ Vektorräume. Dann existiert das Tensorprodukt*

$$\tau\colon E_1 \times \ldots \times E_n \to W.$$

Beweis. Seien $e_{v_1,\ldots,v_n}$ für beliebige $v_i \in E_i$, $1 \le i \le n$, die Basis eines (hochdimensionalen) Vektorraumes $\mathcal{W}$. Auf $\mathcal{W}$ definieren wir eine Äquivalenzrelation R, die durch

$$e_{v_1,\ldots,v_i+w_i,\ldots,v_n} \sim e_{v_1,\ldots,v_i,\ldots,v_n} + e_{v_1,\ldots,w_i,\ldots,v_n}, \quad 1 \le i \le n,$$

und

$$\lambda e_{v_1,\ldots,v_i,\ldots,v_n} \sim e_{v_1,\ldots,\lambda v_i,\ldots,v_n}, \quad 1 \le i \le n$$

erzeugt ist. Dann ist $W := \mathcal{W}/R$ ein Vektorraum (Übung). Wir schreiben $v_1 \otimes \ldots \otimes v_n$ für $[e_{v_1,\ldots,v_n}]$. Setze $\tau(v_1, \ldots, v_n) := v_1 \otimes \ldots \otimes v_n$. Es gilt (exemplarisch nur für den Fall $n = 2$ aufgeschrieben)

$$\begin{aligned}(v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w,\\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2,\\ (\lambda v) \otimes w &= \lambda(v \otimes w) = v \otimes (\lambda w).\end{aligned}$$

Somit bilden die Elemente

$$a_{1,k_1} \otimes \ldots \otimes a_{n,k_n}$$

eine Basis von $E_1 \otimes \ldots \otimes E_n$, falls die Vektoren $(a_{i,k})_k$ eine Basis von E_i, $1 \le i \le n$, bilden. Dies folgt im Fall von zwei Vektorräumen mit Hilfe der von den multilinearen Abbildungen

$$\begin{aligned}\varphi^{i_0,j_0}\colon E_1 \times E_2 &\to W\\ (a_i, a_j) &\mapsto \delta_i^{i_0} \delta_j^{j_0}\end{aligned}$$

nach Definition des Tensorproduktes induzierten linearen Abbildung

$$\Phi^{i_0,j_0}\colon E_1 \otimes E_2 \to W.$$

Beachte dabei, dass die Basisaussage für den Rest des Existenzbeweises nicht gebraucht wird.

Achtung: Allgemeine Elemente in $E_1 \otimes \ldots \otimes E_n$ sind endliche Summen von Elementen der Form

$$v_1 \otimes \ldots \otimes v_n,$$

$v_i \in E_i$, $1 \le i \le n$, sogenannten reinen oder einfachen Tensoren, und nicht nur solche Elemente.

Sei V ein beliebiger Vektorraum. Sei $\varphi\colon E_1 \times \ldots \times E_n \to V$ eine beliebige multilineare Abbildung. Definiere $\Phi\colon E_1 \otimes \ldots \otimes E_n \to V$ durch

$$\Phi(v_1 \otimes \ldots \otimes v_n) := \varphi(v_1, \ldots, v_n)$$

und lineare Fortsetzung. Da die Äquivalenzrelation R entsprechend der Multilinearität definiert ist, ist Φ eine wohldefinierte lineare Abbildung (Übung). $\varphi = \Phi \circ \tau$ gilt nach Konstruktion. Da $E_1 \otimes \ldots \otimes E_n$ von $\operatorname{im} \tau$ erzeugt ist, ist Φ eindeutig definiert. $\square$

Theorem 10.1.4. *Seien U, V, W Vektorräume. Dann gibt es die folgenden Isomorphismen*

$$L(U, L(V, W)) \cong L^2(U, V; W) \cong L(U \otimes V, W).$$

Beweis.

(i) Sei $f \in L^2(U, V; W)$. Wir wollen f zunächst auf ein Element in $L(U, L(V, W))$ abbilden. Sei $x \in U$. Dann setzen wir $(g(x))(y) := f(x, y)$ für $y \in V$. Die Abbildung $y \mapsto (g(x))(y)$ ist linear. Weiterhin ist $U \ni x \mapsto g(x) \in L(V, W)$, linear.

(ii) Sei nun $\varphi \in L(U, L(V, W))$. Dann ordnen wir der Abbildung φ eine Abbildung $f \in L^2(U, V; W)$ vermöge

$$f(x, y) := (\varphi(x))(y)$$

zu. f ist bilinear. Die Abbildung $\varphi \mapsto f$ ist linear. Außerdem ist sie zu der obigen Abbildung $f \mapsto g$ invers.

(iii) Definiere $L^2(U, V; W) \to L(U \otimes V; W)$ durch $\varphi \mapsto \Phi$ mit den Bezeichnungen aus Definition 10.1.1. $\varphi \mapsto \Phi$ ist eine lineare Abbildung. Die Injektivität folgt aus der Definition des Tensorproduktes: Φ kann in dem kommutativen Diagramm nur die Nullabbildung sein, wenn $\varphi \equiv 0$ gilt. Sei $\Phi \in L(U \otimes V; W)$ beliebig. Dann ist $\Phi \circ \tau \colon U \times V \to W$ eine multilineare Abbildung. Ihr wird unter der obigen Abbildung gerade Φ zugeordnet. Somit ist die angegebene Zuordnung auch surjektiv. □

10.2. **Basisdarstellung.** Um die Darstellung einfacher zu halten, beschränken wir uns hier auf den Fall mit wenigen Vektorräumen, E und E^*. Dies ist für Anwendungen häufig völlig ausreichend. Der allgemeine Fall ist nur von der Notation her komplizierter. Wir wollen hier stets annehmen, dass E endlichdimensional ist.

Definition 10.2.1. Sei E ein Vektorraum. Sei $(r, s) \in \mathbb{N} \times \mathbb{N} \setminus \{(0, 0\}$. Ein Tensor T der Stufe (r, s) ist ein Element

$$T \in \underbrace{E \otimes \ldots \otimes E}_{r \text{ Faktoren}} \otimes \underbrace{E^* \otimes \ldots \otimes E^*}_{s \text{ Faktoren}}.$$

Wir schreiben auch $T \in E^{(r,s)}$.

Bemerkung 10.2.2. Ein Element $(v_1, \ldots, v_r, \omega_1, \ldots, \omega_s) \in E^r \times (E^*)^s$ induziert eine Abbildung

$$f \in L^{r+s}(\underbrace{E^*, \ldots, E^*}_{r \text{ Stück}}, \underbrace{E, \ldots, E}_{s \text{ Stück}}; F)$$

vermöge

$$f(\eta_1, \ldots, \eta_r, u_1, \ldots, u_s) := \eta_1(v_1) \cdot \ldots \cdot \eta_r(v_r) \cdot \omega_1(u_1) \cdot \ldots \cdot \omega_s(u_s).$$

Mit derselben Vorschrift können wir einem reinen Tensor $v_1 \otimes \ldots \otimes v_r \otimes \omega_1 \otimes \ldots \otimes \omega_s$ ein $f \in L^{r+s}(E^*, \ldots, E^*, E, \ldots, E; F)$ in wohldefinierter Weise zuordnen. Durch Addition von reinen Tensoren und linearen Abbildungen erhalten wir also eine lineare Abbildung Φ, die jedem Tensor der Stufe (r, s) eine lineare Abbildung $f \in L^{r+s}(E^*, \ldots, E^*, E, \ldots, E; F)$ zuordnet: $\Phi(v_1 \otimes \ldots \otimes v_r \otimes \omega_1 \otimes \ldots \otimes \omega_s) := f$.

Sei umgekehrt $f \in L^{r+s}(E^*, \ldots, E^*, E \ldots, E; F)$. Sei $(a_i)_{1 \le i \le n}$ eine Basis von E und sei $\left((a^*)^i\right)_{1 \le i \le n}$ die zugehörige duale Basis. Für $\omega^k = \sum\limits_{i_k=1}^{n} \omega^k_{i_k} (a^*)^{i_k} \in E^*$, $1 \le k \le r$, und $\xi_l = \sum\limits_{j_l=1}^{n} \xi_l^{j_l} a_{j_l}$, $1 \le l \le s$, erhalten wir aufgrund der Multilinearität von f

$$\begin{aligned}
&f\left(\omega^1,\dots,\omega^r,\xi_1,\dots,\xi_s\right)\\
&=\sum_{i_1=1}^n\cdots\sum_{i_r=1}^n\sum_{j_1=1}^n\cdots\sum_{j_s=1}^n \omega^1_{i_1}\cdots\omega^r_{i_r}\cdot\xi^{j_1}_1\cdots\xi^{j_s}_s\cdot\\
&\qquad\cdot f\left((a^*)^{i_1},\dots,(a^*)^{i_r},a_{j_1},\dots,a_{j_s}\right)\\
&=\sum_{i_1=1}^n\cdots\sum_{i_r=1}^n\sum_{j_1=1}^n\cdots\sum_{j_s=1}^n f\left((a^*)^{i_1},\dots,(a^*)^{i_r},a_{j_1},\dots,a_{j_s}\right)\cdot\\
&\qquad\cdot\Phi\left(a_{i_1}\otimes\ldots\otimes a_{i_r}\otimes(a^*)^{j_1}\otimes\ldots\otimes(a^*)^{j_s}\right)\left(\omega^1,\dots,\omega^r,\xi_1,\dots,\xi_s\right)
\end{aligned}$$

mit Φ wie oben definiert. Indem wir $f(\ldots)$ (Argument wie oben) noch mit dem Tensor $a_{i_1}\otimes\ldots\otimes(a^*)^{j_s}$ multiplizieren und aufsummieren, können wir umgekehrt jeder multilinearen Abbildung einen Tensor zuordnen. Diese beiden Zuordnungen sind invers zueinander. Daher werden wir nun Tensoren und multilineare Abbildungen identifizieren.

Definition 10.2.3. Sei E ein F-Vektorraum. Sei $(a_i)_{1\le i\le n}$ eine Basis von E. Sei $\left((a^*)^i\right)_{1\le i\le n}$ die zugehörige duale Basis. Sei T ein (r,s)-Tensor. Wir definieren $\left(T^{i_1\ldots i_r}_{j_1\ldots j_s}\right)_{1\le j_1,\dots,j_s,i_1,\dots,i_r\le n}$ durch

$$T^{i_1\ldots i_r}_{j_1\ldots j_s}:=T\left((a^*)^{i_1},\dots,(a^*)^{i_r},a_{j_1},\dots,a_{j_s}\right).$$

Bemerkung 10.2.4. Sei E ein Vektorraum mit Basis $(a_i)_{1\le i\le n}$, Dualraum E^* und zugehöriger dualer Basis $\left((a^*)^j\right)_{1\le j\le n}$. Sei T ein (r,s)-Tensor. Für $\omega^k=\sum\limits_{i_k=1}^n\omega^k_{i_k}(a^*)^{i_k}\in E^*$, $1\le k\le r$, und $\xi_l=\sum\limits_{j_l=1}^n\xi^{j_l}_l a_{j_l}$, $1\le l\le s$, erhalten wir aufgrund der Multilinearität von T

$$\begin{aligned}
&T\left(\omega^1,\dots,\omega^r,\xi_1,\dots,\xi_s\right)\\
&=\sum_{i_1=1}^n\cdots\sum_{i_r=1}^n\sum_{j_1=1}^n\cdots\sum_{j_s=1}^n \omega^1_{i_1}\cdots\omega^r_{i_r}\xi^{j_1}_1\cdots\xi^{j_s}_s T\left((a^*)^{i_1},\dots,(a^*)^{i_r},a_{j_1},\dots,a_{j_s}\right)\\
&=\sum_{i_1=1}^n\cdots\sum_{i_r=1}^n\sum_{j_1=1}^n\cdots\sum_{j_s=1}^n \omega^1_{i_1}\cdots\omega^r_{i_r}\xi^{j_1}_1\cdots\xi^{j_s}_s T^{i_1\ldots i_r}_{j_1\ldots j_s}.
\end{aligned}$$

Wir sagen, dass T bezüglich der Basis $(a_i)_{1\le i\le n}$ durch $\left(T^{i_1\ldots i_r}_{j_1\ldots j_s}\right)$ dargestellt sei.

Wir wollen nun die Darstellung $\left(\tilde{T}^{i_1\ldots i_r}_{j_1\ldots j_s}\right)$ von T bezüglich einer weiteren Basis $(b_i)_{1\le i\le n}$ bestimmen. Sei dazu $(d^j_i)_{1\le i,j\le n}$ die invertierbare Matrix mit

$$b_i=\sum_{j=1}^n d^j_i a_j.$$

Sei $(D^j_i)_{1\le i,j\le n}$ die Inverse zu $(d^j_i)_{1\le i,j\le n}$. Dann gilt für die zu $(b_i)_{1\le i\le n}$ duale Basis

$$(b^*)^i=\sum_{j=1}^n D^i_j\,(a^*)^j,$$

denn es gilt

$$\delta^j_i\overset{!}{=}(b^*)^j\,(b_i)=\sum_{k,l=1}^n D^j_k\,(a^*)^k\,d^l_i a_l=\sum_{k,l=1}^n D^j_k\delta^k_l d^l_i=\sum_{k=1}^n D^j_k d^k_i.$$

Wir erhalten

$$\begin{aligned}\tilde{T}^{i_1\ldots i_r}_{j_1\ldots j_s} &= T\left((b^*)^{i_1},\ldots (b^*)^{i_r}, b_{j_1},\ldots,b_{j_s}\right)\\ &= \sum_{k_1=1}^n \cdots \sum_{k_r=1}^n \sum_{l_1=1}^n \cdots \sum_{l_s=1}^n D^{i_1}_{k_1}\cdots D^{i_r}_{k_r}\cdot d^{l_1}_{j_1}\cdots d^{l_s}_{j_s}\cdot\\ &\qquad\cdot T\left((a^*)^{k_1},\ldots (a^*)^{k_r}, a_{l_1},\ldots,a_{l_s}\right)\\ &= \sum_{k_1=1}^n \cdots \sum_{k_r=1}^n \sum_{l_1=1}^n \cdots \sum_{l_s=1}^n D^{i_1}_{k_1}\cdots D^{i_r}_{k_r}\cdot d^{l_1}_{j_1}\cdots d^{l_s}_{j_s}\cdot T^{k_1\ldots k_r}_{l_1\ldots l_s}.\end{aligned}$$

Im Falle eines $(1,1)$-Tensors entspricht dies gerade dem Transformationsverhalten einer Matrix.

Physiker definieren Tensoren über ihr Transformationsverhalten.

Bemerkung 10.2.5 (Heben und Senken von Indices)**.** Wir illustrieren dies beispielhaft nur für $(1,1)$-Tensoren. Die Übertragung auf allgemeine Tensoren ist einfach. Sei E ein endlichdimensionaler Vektorraum, $\dim E = n$. Sei $T \in E^{(1,1)}$. Seien $X, Y \in E$ und $\omega \in E^*$. Die Abbildung $(\omega, X) \to T(\omega, X)$ ist multilinear. Dann ist $T(\cdot, X)$ eine Linearform auf E^*, also vermöge $E^{**} \cong E$ ein Element von E. Somit ist die Abbildung

$$Y \mapsto \langle T(\cdot, X), Y\rangle$$

linear. Wegen $X, Y \in E$ fassen wir somit T mit Hilfe des Skalarproduktes als Element in $E^{(0,2)}$ auf.

In Koordinaten entspricht das dem Senken eines Indexes. Sei das Skalarprodukt $\langle\cdot,\cdot\rangle$ durch die Matrix $(g_{ij})_{1\le i,j\le n}$ dargestellt, eine in der Differentialgeometrie übliche Notation. Sei $(g^{ij})_{1\le i,j\le n}$ die zu $(g_{ij})_{1\le i,j\le n}$ in dem Sinn inverse Matrix, dass eine der beiden äquivalenten Bedingungen,

$$\sum_{i=1}^n g^{ki} g_{il} = \delta^k_l \quad \text{oder} \quad \sum_{i=1}^n g_{ki} g^{il} = \delta^l_k,$$

gilt. Sei T bezüglich einer festen Basis durch die Matrix $(T^i{}_j)_{1\le i,j\le n}$ dargestellt. Wollen wir T wie oben mit Hilfe des Skalarproduktes (g_{ij}) als Element in $\tilde{T} \in E^{(0,2)}$ auffassen, so lautet die entsprechende Koordinatendarstellung $\tilde{T}_{ij} = \sum\limits_{k=1}^n T^k{}_j g_{ki}$. Dies wird als Senken eines Indexes bezeichnet. Mit der Anordnung der Indices kann man verdeutlichen, welcher Index gesenkt wurde.

Als (später in der Differentialgeometrie relevantes) Beispiel betrachten wir den Tensor $R = \left(R_{ij}{}^k{}_l\right)_{1\le i,j,k,l\le n} \in E^{(1,3)}$, der durch Senken des dritten Indexes zu einem $(0,4)$-Tensor wird

$$\tilde{R}_{ijkl} = \sum_{r=1}^n R_{ij}{}^r{}_l g_{rk}.$$

Die dazu inverse Operation ist das Heben eines Indexes. Aus einem Tensor $\tilde{R} = \left(\tilde{R}_{ijkl}\right)_{1\le i,j,k,l\le n} \in E^{(0,4)}$ können wir so wieder einen $(1,3)$-Tensor machen

$$R_{ij}{}^k{}_l = \sum_{r=1}^n \tilde{R}_{ijrl} g^{rk}.$$

10.3. **Eigenschaften.**

Lemma 10.3.1. *Seien U, V, W und V_i, $i \in I$, F-Vektorräume. Dann induzieren die auf Erzeugendensystemen angegebenen Abbildungen Isomorphismen:*

$$\begin{aligned} F \otimes V &\xrightarrow{\sim} V, & \lambda \otimes v &\mapsto \lambda v,\\ V \otimes W &\xrightarrow{\sim} W \otimes V, & v \otimes w &\mapsto w \otimes v,\\ (U \otimes V) \otimes W &\xrightarrow{\sim} U \otimes (V \otimes W), & (u \otimes v) \otimes w &\mapsto u \otimes (v \otimes w),\\ \left(\bigoplus_{i\in I} V_i\right) \otimes W &\xrightarrow{\sim} \bigoplus_{i\in I} (V_i \otimes W), & (v_i)_{i\in I} \otimes w &\mapsto (v_i \otimes w)_{i\in I}. \end{aligned}$$

Beweis. Wir weisen nur nach, dass die erste Abbildung einen Isomorphismus liefert und lassen die übrigen Aussagen als Übungsaufgabe.

Die Abbildung $\varphi\colon F \times V \ni (\lambda, v) \mapsto \lambda v \in V$ ist bilinear und induziert daher eine wohldefinierte lineare Abbildung $\Phi\colon F \otimes V \ni \lambda \otimes v \mapsto \lambda v \in V$. Definiere eine lineare Abbildung durch $\psi\colon V \ni v \mapsto 1 \otimes v \in F \otimes V$. Es folgt $\Phi \circ \psi(v) = \Phi(1 \otimes v) = v$ für alle $v \in V$. Weiterhin gilt $\psi \circ \Phi(\lambda \otimes v) = \psi(\lambda v) = 1 \otimes (\lambda v) = \lambda \otimes v$. Somit erhalten wir $\varphi \circ \psi = \mathrm{id}$ und $\psi \circ \varphi = \mathrm{id}$. φ ist ein Isomorphismus. □

Definition 10.3.2 (Tensorprodukt von Abbildungen)**.** Seien $\varphi\colon M \to M'$ und $\psi\colon N \to N'$ lineare Abbildungen. Dann definieren wir die Abbildung

$$\varphi \otimes \psi\colon M \otimes N \to M' \otimes N'$$

durch

$$x \otimes y \mapsto \varphi(x) \otimes \psi(y)$$

und lineare Fortsetzung. Dies ist aufgrund der Bilinearität der Abbildung

$$M \times N \ni (x, y) \mapsto \varphi(x) \otimes \psi(y)$$

und der definierenden Eigenschaft des Tensorproduktes möglich.

Sei $\varphi\colon M \to M'$ linear. Dann heißt

$$\varphi \otimes \mathrm{id}\colon M \otimes N \to M' \otimes N$$

die mit N tensorierte Abbildung.

Wir bemerken, dass bei dieser Konstruktion das Diagramm

$$\begin{tikzcd} (m,n) \arrow[d] & M \times N \arrow[r] \arrow[d, "\varphi\times\psi"'] & M \otimes N \arrow[d, "\varphi\otimes\psi"] & m \otimes n \arrow[d] \\ (\varphi(m), \psi(n)) & M' \times N' \arrow[r] & M' \otimes N' & \varphi(m) \otimes \varphi(n) \end{tikzcd}$$

kommutativ wird.

Diese Konstruktion funktioniert auch mit mehr als zwei Abbildungen.

Definition 10.3.3 (Exakte Sequenzen)**.** Sei

$$M_n \xrightarrow{\varphi_n} M_{n-1} \xrightarrow{\varphi_{n-1}} \cdots\cdots \xrightarrow{\varphi_3} M_2 \xrightarrow{\varphi_2} M_1$$

eine Folge linearer Abbildungen (möglicherweise auch in eine oder beide Richtungen unendlich fortgesetzt). Eine solche Folge heißt Sequenz.

Eine Sequenz heißt exakte Sequenz, falls $\operatorname{im} \varphi_k = \ker \varphi_{k-1}$ für alle $n \geq k \geq 3$ (allgemeiner: an jeder Stelle) gilt.

Theorem 10.3.4. *Sei*

$$M' \xrightarrow{\varphi} M \xrightarrow{\psi} M'' \longrightarrow 0$$

eine exakte Sequenz von Vektorräumen. Sei N ein weiterer Vektorraum. Dann ist auch die Sequenz

$$M' \otimes N \xrightarrow{\varphi \otimes \mathrm{id}} M \otimes N \xrightarrow{\psi \otimes \mathrm{id}} M'' \otimes N \longrightarrow 0$$

exakt.

Beweis. Aus der Exaktheit folgt $\psi \circ \varphi = 0$. Hieraus erhalten wir $(\psi \otimes \mathrm{id}) \circ (\varphi \otimes \mathrm{id}) = (\psi \circ \varphi) \otimes \mathrm{id} = 0$, da $0 \otimes n$, $n \in N$, der Nullvektor in $M'' \otimes N$ ist. Somit ist $\mathrm{im}(\varphi \otimes \mathrm{id}) \subset \ker(\psi \otimes \mathrm{id})$. Daher induziert $\psi \otimes \mathrm{id}$ eine lineare Abbildung

$$\bar{\psi} \colon (M \otimes N)/\operatorname{im}(\varphi \otimes \mathrm{id}) \to M'' \otimes N.$$

Wäre $\mathrm{im}(\varphi \otimes \mathrm{id}) = \ker(\psi \otimes \mathrm{id})$, so wäre die Abbildung $\bar{\psi}$ injektiv. Daher genügt es, nachzuweisen, dass $\bar{\psi}$ ein Isomorphismus ist.

ψ ist nach Voraussetzung surjektiv. Wähle also zu jedem $x'' \in M''$ ein $\imath(x'') \in M$ mit $\psi(\imath(x'')) = x''$. Es ist nicht klar, ob $\imath$ linear ist. Definiere

$$\begin{aligned} \sigma \colon M'' \times N &\to (M \otimes N)/\operatorname{im}(\varphi \otimes \mathrm{id}), \\ (x'', y) &\mapsto \overline{\imath(x'') \otimes y}. \end{aligned}$$

$\overline{\imath(x'') \otimes y}$ steht hier für die Restklasse von $\imath(x'') \otimes y$. Wir behaupten, dass σ eine bilineare Abbildung ist. Die Linearität im zweiten Argument ist klar. Wir zeigen zunächst, dass das Element $\overline{\imath(x'') \otimes y}$ von der speziellen Wahl von $\imath(x'') \in M$ zu $x'' \in M''$ unabhängig ist. Seien $x_1, x_2 \in M$ zwei solche Wahlen. Dann folgt aus $\psi(x_1 - x_2) = 0$ und der vorausgesetzten Exaktheit $x_1 - x_2 \in \operatorname{im} \varphi$, also $x_1 - x_2 = \varphi(x')$ für ein $x' \in M'$. Wir erhalten

$$\overline{x_1 \otimes y} - \overline{x_2 \otimes y} = \overline{\varphi(x') \otimes y} = \overline{(\varphi \otimes \mathrm{id})(x' \otimes y)} = 0$$

nach Definition des Quotientenraumes. Somit ist σ wohldefiniert.

Die Abbildung ψ induziert eine lineare Abbildung $M/\operatorname{im} \varphi = M/\ker \psi \to M''$. Zu dieser gibt es eine lineare Inverse. Bis auf die Auswahl eines Vertreters in $M/\operatorname{im} \varphi$ ist i gerade diese Inverse. Da wir die Wohldefiniertheit überprüft haben, ist σ somit bilinear.

Nach Definition des Tensorraumes induziert σ eine Abbildung

$$M'' \otimes N \to (M \otimes N)/\operatorname{im}(\varphi \otimes \mathrm{id}).$$

Nach Definition von σ ist klar, dass diese induzierte Abbildung auf Erzeugern der Form $x'' \otimes y$ zu $\overline{\psi}$ invers ist.

Die Surjektivität der Abbildung $\psi \otimes \mathrm{id}$ folgt direkt aus der Surjektivität von ψ.

Die Behauptung folgt. □

Beispiel 10.3.5. Sei $U \subset V$ ein Unterraum. Dann ist die Sequenz

$$U \xrightarrow{i} V \xrightarrow{\pi} V/U \longrightarrow 0$$

mit der Inklusionsabbildung i und der Projektionsabbildung π exakt. Sei W ein weiterer Vektorraum. Dann ist

$$U \otimes W \xrightarrow{i \otimes \mathrm{id}} V \otimes W \xrightarrow{\pi \otimes \mathrm{id}} (V/U) \otimes W \longrightarrow 0$$

ebenfalls exakt. Somit liefert der Homomorphiesatz für $\pi \otimes \mathrm{id}$ zusammen mit der Exaktheit den Isomorphismus

$$\begin{aligned}(V/U) \otimes W &\xrightarrow{\sim} (V \otimes W)/\operatorname{im}(i \otimes \mathrm{id}),\\ \bar{x} \otimes y &\mapsto \overline{x \otimes y}.\end{aligned}$$

Im Gegensatz zu den bisherigen Resultaten zu Tensorprodukten lässt sich Theorem 10.3.7 i. a. nicht auf allgemeinere Situationen als Vektorräume und lineare Abbildungen übertragen, denn das folgende Theorem ist spezifisch für Vektorräume.

Theorem 10.3.6. *Sei*

$$0 \longrightarrow A \xrightarrow{\varphi} C \xrightarrow{\psi} B \longrightarrow 0$$

eine exakte Sequenz von Vektorräumen und linearen Abbildungen. Dann ist C isomorph zu $A \oplus B$.

Beweis. φ ist injektiv und somit ein Isomorphismus auf $\operatorname{im}\varphi$. Der Unterraum $\varphi(A) \subset C$ besitzt ein Komplement D mit $\varphi(A) \oplus D = C$. Da $\psi \circ \varphi = 0$ gilt und ψ surjektiv ist, ist auch $\psi|_D \colon D \to B$ surjektiv. $\psi|_D$ ist aber auch injektiv, da $D \cap \ker\psi = D \cap \operatorname{im}\varphi = \{0\}$ gilt. Somit ist $\psi|_D$ bijektiv. □

Wir wenden dies nun auf die Exaktheit auf der linken Seite an.

Theorem 10.3.7. *Sei*

$$0 \longrightarrow A \xrightarrow{\varphi} A \oplus B \xrightarrow{\psi} B \longrightarrow 0$$

eine exakte Sequenz von Vektorräumen und linearen Abbildungen mit $\varphi(a) = (a, 0)$ sowie $\psi(a, b) = b$. Sei V ein weiterer Vektorraum. Dann ist

$$0 \longrightarrow A \otimes V \xrightarrow{\varphi\otimes\mathrm{id}} (A \oplus B) \otimes V \xrightarrow{\psi\otimes\mathrm{id}} B \otimes V \longrightarrow 0$$

ebenfalls exakt.

Beweis. Die Exaktheit von

$$A \otimes V \to (A \oplus B) \otimes V \to B \otimes V \to 0$$

haben wir bereits gezeigt. Benutze den Isomorphismus

$$\begin{aligned}(A \oplus B) \otimes V &\to (A \otimes V) \oplus (B \otimes V),\\ (a, b) \otimes v &\mapsto (a \otimes v, b \otimes v).\end{aligned}$$

Nach Verkettung mit diesem Isomorphismus bekommt $\varphi \otimes \mathrm{id}$ die Gestalt $a \otimes v \mapsto (a \otimes v, 0)$, ist also offensichtlicherweise injektiv. Somit erhalten wir die behauptete Exaktheit. □

11. Kettenkomplexe und exakte Sequenzen

Wir benutzen in diesem Kapitel [4].

11.1. **Kettenkomplexe.** Wir wiederholen zunächst eine Definition zu Gruppenhomomorphismen und beweisen ein kleines Lemma dazu.

Lemma 11.1.1. *Seien G, H Gruppen. Dann heißt $\varphi\colon G \to H$ Gruppenhomomorphismus, falls $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ für alle $a, b \in G$ gilt. Es gelten*

(i) $\varphi(1) = 1$,
(ii) *φ ist genau dann injektiv, wenn $\varphi^{-1}(\{1\}) = \{1\}$ gilt.*

Beweis.

(i) Dies folgt aus $\varphi(a) = \varphi(1 \cdot a) = \varphi(1) \cdot \varphi(a)$.

(ii) Sei φ nicht injektiv und gelte $\varphi(a) = \varphi(b)$ für $a \neq b$. Dann folgt $\varphi(ab^{-1}) = 1$ mit $ab^{-1} \neq 1$.

Die andere Richtung ist trivial. $\square$

Definition 11.1.2 (Kettenkomplexe und exakte Sequenzen)**.** Sei

$$C_n \xrightarrow{d_n} C_{n-1} \xrightarrow{d_{n-1}} \cdots\cdots \xrightarrow{d_3} C_2 \xrightarrow{d_2} C_1$$

eine Folge von Gruppen und Gruppenhomomorphismen (möglicherweise auch in eine oder beide Richtungen unendlich fortgesetzt). Eine solche Folge heißt Sequenz.

(i) Eine Sequenz heißt Kettenkomplex, falls $\operatorname{im}\varphi_k \subset \ker\varphi_{k-1}$ für alle $n \geq k \geq 3$ (allgemeiner: an jeder Stelle) gilt.
(ii) Eine Sequenz heißt exakte Sequenz, falls $\operatorname{im}\varphi_k = \ker\varphi_{k-1}$ für alle $n \geq k \geq 3$ (allgemeiner: an jeder Stelle) gilt.
(iii) Gilt $\operatorname{im} d_{k+1} = \ker d_k$, so heißt die Sequenz an der Stelle C_k exakt.
(iv) Eine exakte Sequenz der Form

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

heißt kurze exakte Sequenz. (Hierbei steht 0 für die triviale Gruppe, die nur aus einem Element besteht. Die nicht bezeichneten Pfeile stehen für die trivialen Homomorphismen. Abelsche Gruppen schreiben wir meist additiv mit dem neutralen Element 0.)
(v) Eine exakte Sequenz, die in beiden Richtungen unendlich lang ist, heißt lange exakte Sequenz.

Lemma 11.1.3. *Sei*

$$A \xrightarrow{f} B \xrightarrow{g} C$$

eine Sequenz von Gruppen und Gruppenhomomorphismen. Dann sind die folgenden beiden Aussagen äquivalent:

(i) $\operatorname{im} f \subset \ker g$.
(ii) $g \circ f = 0$.

Beweis. Übung. $\square$

Lemma 11.1.4. *Für Gruppen G, H und einen Gruppenhomomorphismen $f\colon G \to H$ gelten:*

(i) f ist genau dann injektiv, wenn die Sequenz

$$0 \longrightarrow G \xrightarrow{f} H$$

(an der Stelle G) exakt ist.

(ii) f ist genau dann surjektiv, wenn die Sequenz

$$G \xrightarrow{f} H \longrightarrow 0$$

(an der Stelle H) exakt ist.

(iii) f ist genau dann bijektiv, wenn die Sequenz

$$0 \longrightarrow G \xrightarrow{f} H \longrightarrow 0$$

exakt ist.

Beweis.

(i) Die Exaktheit ist äquivalent zu $\operatorname{im} 0 = \{0\} = \ker f$.
(ii) Die Exaktheit ist äquivalent zu $\operatorname{im} f = \ker 0 = H$.
(iii) Folgt aus (i) und (ii). $\square$

Definition 11.1.5 (Direkte Summe)**.** Seien G, H Gruppen. Dann ist $G \times H$ mit der Verknüpfung

$$(g_1, h_1) \cdot (g_2, h_2) := (g_1 \cdot g_2, h_1 \cdot h_2),$$

dem neutralen Element $e = (e_G, e_H)$ und der Inversen $(g, h)^{-1} := (g^{-1}, h^{-1})$ eine Gruppe, die direkte Summe der Gruppen G und H: $G \oplus H$.

Im folgenden Theorem ist die Kommutativität der beteiligten Gruppen wichtig.

Theorem 11.1.6. *Sei*

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{j} C \longrightarrow 0$$

eine exakte Sequenz abelscher Gruppen. Dann sind die folgenden Bedingungen äquivalent:

(i) Es gibt einen (Gruppen-)Homomorphismus $s\colon B \to A$ mit $s \circ i = \mathrm{id}_A$.
(ii) Es gibt einen (Gruppen-)Homomorphismus $t\colon C \to B$ mit $j \circ t = \mathrm{id}_C$.
(iii) Es gibt einen Isomorphismus $B \cong A \oplus C$, so dass das folgende Diagramm kommutativ ist:

$$\begin{tikzcd} 0 \arrow[r] & A \arrow[r, "i"] \arrow[dr] & B \arrow[r, "j"] \arrow[d, "\cong"] & C \arrow[r] \arrow[dl] & 0, \\ & & A \oplus C & & \end{tikzcd}$$

wobei die fehlenden Abbildungen durch $i_A\colon A \to A \oplus C$ mit $a \mapsto (a, 0)$ und $i_C\colon C \to A \oplus C$ mit $c \mapsto (0, c)$ gegeben sind. Definiere weiterhin die Projektionen auf den ersten Faktor $\pi_A\colon A \oplus C \to A$ bzw. auf den zweiten Faktor $\pi_C\colon A \oplus C \to C$. Es gilt $\pi_A i_A = \mathrm{id}_A$ und $\pi_C i_C = \mathrm{id}_C$.

Ist eine dieser Bedingungen erfüllt, so sagen wir, dass die exakte Sequenz spaltet.

Beweis. „(iii)$\Longrightarrow$(i)“: Sei $\Phi\colon B \to A \oplus C$ der angegebene Isomorphismus. Definiere $s\colon B \to A$ durch $s(b) := \pi_A(\Phi(b))$. Es folgt $si(a) = \pi_A(\Phi(i(a))) = \pi_A(a, 0) = a$, wobei die mittlere Gleichheit aus der Kommutativität des angegebenen Diagrammes folgt.

„(iii)$\Longrightarrow$(ii)“: Sei $\Phi\colon B \to A \oplus C$ der angenommene Isomorphismus. Setze $t(c) := \Phi^{-1}(i_C(c))$. Aus der Kommutativität des angegebenen Diagrammes folgt $i_C j = \Phi$ und daraus $\pi_C \Phi = j$ oder $\pi_C = j\Phi^{-1}$. Wir erhalten $jt = j\Phi^{-1} i_C = \pi_C i_C = \mathrm{id}$.

„(i)$\Longrightarrow$(iii)“: Die gesuchte Abbildung Φ ist $B \ni b \mapsto (s(b), j(b)) \in A \oplus C$. Dies ist ein Gruppenhomomorphismus.

Wir zeigen zunächst, dass Φ surjektiv ist: Sei $(a, c) \in A \oplus C$. Da j surjektiv ist, gibt es ein $b_1 \in B$ mit $jb_1 = c$ (b_1 ist bis auf ein Element in $\operatorname{im} i = \ker j$ wohldefiniert. Wir werden b_1 nun noch um ein solches Element abändern um die Surjektivität von Φ zu zeigen.). Setze $b := i(a - sb_1) + b_1$. Es gilt $jb = jb_1 = c$. Weiterhin gilt nach Voraussetzung $sb = sia - (si)sb_1 + sb_1 = a - sb_1 + sb_1 = a$. Somit ist $\Phi(b) = (a, c)$ und Φ ist surjektiv.

Zur Injektivität von b: Gelte $(s(b), j(b)) = (0, 0)$. Aus $j(b) = 0$ folgt zunächst $b \in \ker j$ und dann aufgrund der Exaktheit der kurzen exakten Sequenz $b \in \operatorname{im} i$. Wähle daher $a \in A$ mit $ia = b$. Benutze nun $s(b) = 0$. Wir erhalten $0 = s(b) = sia = a$, also $a = 0$. Somit gilt $b = ia = i0 = 0$ und Φ ist injektiv.

„(ii)$\Longrightarrow$(iii)“: Betrachte $A \oplus C \to B$ mit $(a, c) \mapsto ia + tc$.

Zur Surjektivität: Sei $b \in B$. Setze $c := jb$. Es gilt $j(b - tc) = jb - jtc = c - c = 0$. Also gilt $b - tc \in \ker j$. Aufgrund der Exaktheit gibt es somit ein $a \in A$ mit $ia = b - tc$. Nun wird (a, c) auf $ia + tc = b - tc + tc = b$ abgebildet. Dies zeigt die Surjektivität.

Zur Injektivität: Gelte $ia + tc = 0$. Es folgt $0 = j0 = jia + jtc = jtc = c$. Somit erhalten wir $c = 0$ und $ia = 0$. Da i injektiv ist, gilt auch $a = 0$. $\square$

Solch ein Beweis wird auch als Diagrammjagd bezeichnet.

Beispiele 11.1.7.

(i)

$$0 \to A \to A \oplus B \to B \to 0$$

mit den Abbildungen $A \ni a \mapsto (a, 0) \in A \oplus B$ und $A \oplus B \ni (a, b) \mapsto b \in B$ spaltet.

(ii) Nicht jede exakte Sequenz spaltet:

$$0 \to \mathbb{Z}/2\mathbb{Z} \xrightarrow{\cdot 2} \mathbb{Z}/4\mathbb{Z} \xrightarrow{\text{mod } 2} \mathbb{Z}/2\mathbb{Z} \to 0$$

spaltet nicht, da $\mathbb{Z}/4\mathbb{Z}$ und $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ nicht isomorph sind (betrachte $a + a$).

Theorem 11.1.8. *Sei*

$$\cdots \longrightarrow C_{n+2} \xrightarrow{d_{n+2}} C_{n+1} \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \xrightarrow{d_{n-1}} C_{n-2} \longrightarrow \cdots$$

eine lange exakte Sequenz. Dann erhalten wir für jedes n eine kurze exakte Sequenz

$$0 \longrightarrow \operatorname{im} d_{n+1} \xrightarrow{i_n} C_n \xrightarrow{d_n} \operatorname{im} d_n \longrightarrow 0,$$

wobei i_n die Inklusionsabbildung bezeichnet.

Beweis. Die Injektivität von i_n und die Surjektivität von d_n sind klar. Es gilt $\operatorname{im} i_n = \operatorname{im} d_{n+1} = \ker d_n$ aufgrund der Exaktheit der langen exakten Sequenz. Somit ist die kurze Sequenz exakt. □

11.2. **Homologie.**

Definition 11.2.1 (Homologiegruppen)**.** Sei

$$\cdots \xrightarrow{d_{n+2}} C_{n+1} \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} C_{n-1} \xrightarrow{d_{n-1}} \cdots$$

ein Kettenkomplex abelscher Gruppen. Wir wollen ab jetzt annehmen, dass jeder Kettenkomplex ein Kettenkomplex abelscher Gruppen ist. Es gilt $\operatorname{im} d_{n+1} \subset \ker d_n$. $\operatorname{im} d_{n+1}$ und $\ker d_n$ sind Untergruppen von C_n. Da C_n abelsch ist, ist $\operatorname{im} d_{n+1}$ automatisch auch ein Normalteiler von $\ker d_n$ und C_n. Definiere die n-te Homologiegruppe des Kettenkomplexes als die Quotientengruppe $H_n := \ker d_n / \operatorname{im} d_{n+1}$. Wir verwenden folgende Bezeichnungen:

(i) Elemente in $\ker d_n$ heißen Zykel.
(ii) Elemente in $\operatorname{im} d_{n+1}$ heißen Ränder.
(iii) Die Nebenklassen von $\operatorname{im} d_{n+1}$, $a + \operatorname{im} d_{n+1}$, in H_n heißen Homologieklassen.
(iv) Zwei Zykel, die dieselbe Homologieklasse repräsentieren, heißen homolog. (Daher unterscheiden sich zwei homologe Zykel um einen Rand.)

Definition 11.2.2. Ein **kurze exakte Sequenz von Kettenkomplexen**

$$0 \to A \to B \to C \to 0$$

ist ein kommutatives Diagramm der Form

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
\cdots \arrow[r, "d"] & A_{n+1} \arrow[r, "d"] \arrow[d, "i"] & A_n \arrow[r, "d"] \arrow[d, "i"] & A_{n-1} \arrow[r, "d"] \arrow[d, "i"] & \cdots \\
\cdots \arrow[r, "d"] & B_{n+1} \arrow[r, "d"] \arrow[d, "j"] & B_n \arrow[r, "d"] \arrow[d, "j"] & B_{n-1} \arrow[r, "d"] \arrow[d, "j"] & \cdots \\
\cdots \arrow[r, "d"] & C_{n+1} \arrow[r, "d"] \arrow[d] & C_n \arrow[r, "d"] \arrow[d] & C_{n-1} \arrow[r, "d"] \arrow[d] & \cdots \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

in dem die Zeilen Kettenkomplexe und die Spalten kurze exakte Sequenzen sind.

Lemma 11.2.3. *Sei $(i_n)_{n\in\mathbb{N}}$ eine Folge von Abbildungen zwischen den Gruppen der Kettenkomplexe $\left(A, d^A\right)$ und $\left(B, d^B\right)$, so dass das Diagramm*

$$\begin{tikzcd}
\cdots \arrow[r] & A_{n+1} \arrow[r, "d^A_{n+1}"] \arrow[d, "i_{n+1}"] & A_n \arrow[r] \arrow[d, "i_n"] & \cdots \\
\cdots \arrow[r] & B_{n+1} \arrow[r, "d^B_{n+1}"] & B_n \arrow[r] & \cdots
\end{tikzcd}$$

kommutiert. Dann ist die vertreterweise definierte Abbildung

$$\begin{aligned}(i_n)_*\colon H_n(A) &\to H_n(B),\\ [a] &\mapsto [i_n a]\end{aligned}$$

wohldefiniert. Solch eine Abbildung $i = (i_n)_n$ nennen wir eine Kettenabbildung. i_* heißt die von i induzierte Abbildung.

Ist $j\colon \left(B, d^B\right) \to \left(C, d^C\right)$ eine weitere Kettenabbildung, so gilt $(j \circ i)_* = j_* \circ i_*$. Schließlich gilt $\mathrm{id}_* = \mathrm{id}$.

Beweis. Wir zeigen nur den nichttrivialen Teil, die Wohldefiniertheit von $(i_n)_*$. Seien $a, a' \in A_n$ mit $a - a' \in \operatorname{im} d^A_{n+1}$. Gelte also $a' = a + d^A_{n+1}\alpha$ für ein $\alpha \in A_{n+1}$. Dann folgt aufgrund der Kommutativität des Diagrammes

$$\begin{aligned}(i_n)_*(a) &= [i_n(a)],\\ (i_n)_*\left(a + d^A_{n+1}\alpha\right) &= \left[i_n a + i_n d^A_{n+1}\alpha\right] = \left[i_n a + d^B_{n+1}(i_{n+1}\alpha)\right] = [i_n a].\end{aligned} \qquad \square$$

Theorem 11.2.4. *Sei $0 \to A \to B \to C \to 0$ eine kurze exakte Sequenz von Kettenkomplexen mit Bezeichnungen wie in Definition 11.2.2. Seien*

$$H_n(\cdot) := \ker d / \operatorname{im} d$$

für $\cdot = A, B, C$ die zugehörigen Homologiegruppen. Dann gibt es eine Abbildung $\partial : H_n(C) \to H_{n-1}(A)$ für alle n, so dass

$$\begin{tikzcd}
\ldots \arrow[r] & H_n(A) \arrow[r, "i_*"] & H_n(B) \arrow[r, "j_*"] & H_n(C) \arrow[r, "\partial"] & {} \\
{} \arrow[r, "\partial"] & H_{n-1}(A) \arrow[r, "i_*"] & H_{n-1}(B) \arrow[r, "j_*"] & H_{n-1}(C) \arrow[r] & \ldots
\end{tikzcd}$$

eine lange exakte Sequenz von Homologiegruppen ist.

Beweis. Zunächst einmal konstruieren wir die Abbildung ∂. (Dann werden wir die Exaktheit der Sequenz beweisen.)

Die Randabbildung ∂ soll eine Abbildung $\partial : H_n(C) \to H_{n-1}(A)$ sein. Sei also $c \in C_n$ ein Zykel. Da j surjektiv ist, gibt es ein $b \in B_n$ mit $c = j(b)$. (Beachte, dass es hier zunächst einmal nur darum geht, dem Element $c \in C_n$ ein Element $\partial c \in A_{n-1}$ zuzuordnen. Erst später wollen wir uns um die Wohldefiniertheit kümmern.) Aufgrund der Kommutativität des Diagramms in der kurzen exakten Sequenz von Kettenkomplexen gilt $jdb = djb = dc = 0$, wobei wir ausgenutzt haben, dass $c = jb$ ist und dass c ein Zykel ist, also $dc = 0$ erfüllt. Also gilt $B_{n-1} \ni db \in \ker j$.

Da die Sequenz exakt ist, gilt $\operatorname{im} i = \ker j$ und somit finden wir ein $a \in A_{n-1}$, so dass $ia = db$ ist. Das Element a ist ein Zykel, denn es gilt $ida = dia = ddb = 0$ aufgrund der Kommutativität des Diagrammes, wegen $ia = db$ und da $d^2 = 0$ ist. $da = 0$ folgt nun, denn i ist injektiv.

Wir definieren nun $d : H_n(C) \to H_{n-1}(A)$ durch $d[c] := [a]$, wobei $[\cdot]$ die jeweiligen Homologieklassen bezeichnet.

Nachweis der Wohldefiniertheit:

- a ist eindeutig durch db bestimmt, denn i ist injektiv.
- Wählt man statt b ein anderes Element $b' \in B_n$, so dass $j(b') = j(b)$ gilt, so ist $b' - b \in \ker j = \operatorname{im} i$. Also gibt es ein $a' \in A_n$ mit $b' - b = ia'$ oder $b' = b + ia'$. Dann erhält man statt $a \in A_{n-1}$ das Element $a + da'$, denn es gilt $i(a + da') = ia + ida' = db + dia' = d(b + ia')$ und somit erfüllt $a + da'$ die oben angegebene Vorschrift, um aus $b \in B_n$ ein Element $a \in A_{n-1}$ zu gewinnen. Bei Betrachtung der Homologiegruppen ist aber $[a] = [a + da']$, da sich die Repräsentanten nur um ein Element in $\operatorname{im} d$ unterscheiden.
- Sei $c + dc'$ ein anderer Vertreter für die Homologieklasse $[c]$ mit $c' \in C_{n+1}$. Da j surjektiv ist, finden wir $b' \in B_{n+1}$ mit $jb' = c'$. Dann gilt $c + dc' = c + djb' = c + jdb' = j(b + db')$ aufgrund der Kommutativität des Diagramms. Also wird b in der obigen Konstruktion durch $b + db'$ ersetzt, aber $db = d(b + db')$ wegen $d^2 = 0$ und somit sind wir in B_{n-1} wieder bei db, was auch wieder dasselbe $a \in A_{n-1}$ liefert.

Die Abbildung $\partial : H_n(C) \to H_{n-1}(A)$ ist ein Homomorphismus, denn wenn $\partial[c_1] = [a_1]$ und $\partial[c_2] = [a_2]$ vermöge b_1 bzw. b_2 gelten, dann ist $j(b_1 + b_2) = jb_1 + jb_2 = c_1 + c_2$ und $i(a_1 + a_2) = ia_1 + ia_2 = db_1 + db_2 = d(b_1 + b_2)$. Somit gilt auch $\partial([c_1] + [c_2]) = [a_1] + [a_2]$. Hier haben wir (öfters) benutzt, dass wir schon wissen, dass alle Abbildungen, außer der neu konstruierten Randabbildung ∂, Homomorphismen sind.

Nun wollen wir nachrechnen, dass die angegebene Sequenz (an jeder Stelle) exakt ist.

$$H_n(A) \xrightarrow{i_*} H_n(B) \xrightarrow{j_*} H_n(C) \xrightarrow{\partial} H_{n-1}(A) \xrightarrow{i_*} H_{n-1}(B)$$

Es sind die folgenden sechs Inklusionen nachzurechnen:

- $\operatorname{im} i_* \subset \ker j_*$: Es gilt $ji = 0$, da das Diagramm exakt ist. Somit gilt auch $j_* i_* = 0$ und die Behauptung folgt.
- $\operatorname{im} j_* \subset \ker \partial$: Sei b ein Zykel. Dann ist $j_*[b]$ ein beliebiges Element in $\operatorname{im} j_*$ und es gilt $d^B b = 0$. Somit ist nach Definition $a = 0$ in der Konstruktion von ∂ und die Behauptung folgt.
- $\operatorname{im} \partial \subset \ker i_*$: Mit den Bezeichnungen aus der Konstruktion von ∂ gilt $i_* \partial[c] = [db] = 0$, da in $H.(B)$ Äquivalenzklassen modulo $\operatorname{im} d$ betrachtet werden.
- $\ker j_* \subset \operatorname{im} i_*$: Nach Definition besitzt eine Homologieklasse in $\ker j_*$ als Vertreter einen Zykel $b \in B_n$, so dass jb ein Rand ist (und daher $j_*[b] = 0$ gilt). Also gibt es $c' \in C_{n+1}$ mit $dc' = jb$. Da j surjektiv ist, finden wir

$b' \in B_{n+1}$, so dass $c' = j(b')$ gilt. Es gilt $djb' = dc' = jb$ und daher folgt $j(b-db') = jb-jdb' = jb-djb' = 0$. Also ist $b-db' \in \ker j = \operatorname{im} i$ aufgrund der Exaktheit der kurzen exakten Sequenzen. Daher gibt es $a \in A_n$ mit $ia = b-db'$. Es gilt $ida = dia = d(b-db') = db = 0$, da b ein Zykel ist. Nun ist i injektiv und daher ist auch a ein Zykel. Somit folgt $i_*[a] = [b-db'] = [b]$. Somit haben wir zu einer Homologieklasse $[b] \in \ker j_*$ ein a gefunden, so dass $i_*[a] = [b]$ gilt, also folgt $\ker j_* \subset \operatorname{im} i_*$ wie behauptet.

- $\ker\partial \subset \operatorname{im} j_*$: Wir verwenden wieder die Bezeichnungen aus der Konstruktion von ∂. Sei c ein Vertreter einer Homologieklasse in $\ker\partial$. Dann gilt $a = da'$ für ein $a' \in A_n$. Dann ist $b - ia'$ ein Zykel, denn es gilt $d(b-ia') = db-dia' = db-ida' = db-ia = 0$ nach Wahl von a. Weiterhin gilt $j(b-ia') = jb-jia' = jb = c$. Beachte, dass j_* auf $[b-ia']$ definiert ist, da $b-ia'$ ein Zykel ist. Es gilt $j_*[b-ia'] = [c]$ und wir schließen, dass $\ker\partial \subset \operatorname{im} j_*$ gilt.
- $\ker i_* \subset \operatorname{im}\partial$: Sei $a \in A_{n-1}$ ein Zykel, so dass $[a] \in \ker i_*$ gilt. Dann gibt es $b \in B_n$ mit $ia = db$. Es gilt $djb = jdb = jia = 0$ und somit ist jb ein Zykel. Nach Definition von ∂ gilt schließlich noch $\partial[jb] = [a]$ und die Behauptung folgt. $\square$

Lemma 11.2.5 (Fünferlemma)**.** *Sei*

$$\begin{array}{ccccccccc}
A & \xrightarrow{i} & B & \xrightarrow{j} & C & \xrightarrow{k} & D & \xrightarrow{l} & E \\
\downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} & & \downarrow{\scriptstyle\delta} & & \downarrow{\scriptstyle\varepsilon} \\
A' & \xrightarrow{i'} & B' & \xrightarrow{j'} & C' & \xrightarrow{k'} & D' & \xrightarrow{l'} & E'
\end{array}$$

ein kommutatives Diagramm abelscher Gruppen mit exakten Zeilen. Sind die Homomorphismen α, β, δ und ε Isomorphismen, so ist auch γ ein Isomorphismus.

Beweis. Wir werden zeigen, dass

(i) γ surjektiv ist, wenn β und δ surjektiv und ε injektiv sind,
(ii) γ injektiv ist, wenn β und δ injektiv sind und α surjektiv ist.

Damit können wir die Voraussetzungen an α und ε abschwächen.

(i) Sei $c' \in C'$. Da δ surjektiv ist, gibt es ein $d \in D$, so dass $\delta d = k'c'$ gilt. Es gilt $\varepsilon ld = l'\delta d = l'k'c' = 0$, da insbesondere die letzte Zeile exakt ist. Aus $\varepsilon ld = 0$ und der Injektivität von ε folgt $ld = 0$. Da die erste Zeile exakt ist, finden wir ein $c \in C$, so dass $d = kc$ gilt. Es gilt $k'(c' - \gamma c) = k'c' - k'\gamma c = k'c' - \delta kc = k'c' - \delta d = 0$ aufgrund von schon hergeleiteten Identitäten. Aufgrund der Exaktheit gibt es ein $b' \in B'$, so dass $c' - \gamma c = j'b'$ gilt. Weil β surjektiv ist, existiert $b \in B$ mit $b' = \beta b$. Es gilt nun $\gamma(c + jb) = \gamma c + \gamma jb = \gamma c + j'\beta b = \gamma c + j'b' = c'$ und daher ist γ surjektiv.

(ii) Sei nun $c \in C$ mit $\gamma c = 0$. Es gilt $\delta kc = k'\gamma c = 0$. Da δ injektiv ist, folgt auch $kc = 0$. Daher gibt es ein $b \in B$ mit $c = jb$. Es gilt $j'\beta b = \gamma jb = \gamma c = 0$. Aufgrund der Exaktheit finden wir also $a' \in A'$, so dass $\beta b = i'a'$ gilt. Nun ist α surjektiv, es gibt also ein $a \in A$, so dass $a' = \alpha a$ ist. Es gilt $\beta(ia - b) = \beta ia - \beta b = i'\alpha a - \beta b = i'a' - \beta b = 0$. Da β injektiv ist, folgt auch $ia - b = 0$ oder, äquivalent dazu, $b = ia$. Weiterhin gilt $c = jb = jia = 0$ aufgrund der Exaktheit ($ji = 0$). Also ist $\ker\gamma$ trivial und γ injektiv. $\square$

12. Darstellungstheorie

Wir folgen [8].

Sei V hier stets ein endlichdimensionaler $\mathbb{C}$-Vektorraum und $GL(V)$ die Gruppe der Isomorphismen $V \to V$. Isomorphismen $V \to V$ heißen auch Automorphismen,

$\mathrm{Aut}(V)$. Wir nehmen der Einfachheit halber stets an, dass alle Gruppen G, die dargestellt werden, endliche Gruppen sind.

12.1. **Darstellungen endlicher Gruppen.**

Definition 12.1.1 (Darstellung)**.**

(i) Sei G eine Gruppe und V ein Vektorraum. Dann heißt ein Gruppenhomomorphismus $\rho\colon G \to GL(V)$ eine (lineare) Darstellung, d. h. es gilt

$$\rho(st) = \rho(s)\rho(t) \quad \text{für alle } s, t \in G.$$

Wir schreiben auch $\rho_s \equiv \rho(s)$. (Es folgen $\rho(1) = 1$ sowie $\rho(s^{-1}) = (\rho(s))^{-1}$.) Wir sagen auch, dass V eine Darstellung sei.

(ii) $\dim V =: n$ heißt der Rang der Darstellung.

(iii) In Matrixform ist eine Darstellung durch $G \ni s \mapsto R_s \in \mathbb{C}^{n\times n}$ mit $\det R_s \neq 0$ und $R_{st} = R_s R_t$ gegeben. R_s ist dabei die ρ_s bezüglich einer fixierten Basis darstellende Matrix.

(iv) Seien $\rho\colon G \to V$ und $\rho'\colon G \to V'$ zwei Darstellungen. Dann heißen ρ und ρ' ähnlich (oder isomorph), falls sie durch einen linearen Isomorphismus $\tau\colon V \to V'$ vermöge

$$\tau\rho(s) = \rho'(s)\tau \quad \text{für alle } s \in G$$

verknüpft sind. In Matrixform bedeutet dies, dass es eine invertierbare Matrix $T \in \mathbb{C}^{n\times n}$ mit $TR_s = R'_s T$ oder $R'_s = TR_sT^{-1}$ für alle $s \in G$ gilt. (Dabei ist R'_s die $\rho'(s)$ bezüglich einer festen Basis von V' zugeordnete Matrix.)

Beispiele 12.1.2.

(i) Eine Darstellung vom Grad 1 ist ein Gruppenhomomorphismus $\rho\colon G \to \mathbb{C}^* \equiv \mathbb{C}\setminus\{0\}$. Da in einer endlichen Gruppe G jedes Element $s \in G$ endliche Ordnung hat (d. h. es gibt ein $k \in \mathbb{N}$ mit $s^k = 1$), ist $\rho(s)$ eine Einheitswurzel, also von der Form $e^{\frac{2\pi i l}{k}}$ für ein $l \in \mathbb{N}$. Insbesondere folgt $|\rho(s)| = 1$ für alle $s \in G$.

(ii) Die Darstellung mit $\rho(s) = 1$ für alle $s \in G$ heißt trivial.

(iii) Sei $g = |G|$ die Ordnung (= die Anzahl der Elemente) von G. Sei V ein Vektorraum der Dimension g mit Basiselementen $(e_t)_{t\in G}$, indiziert durch die Gruppenelemente. Sei $\rho(s)\colon V \to V$ die lineare Abbildung mit $e_t \mapsto e_{st}$ (Beachte die Reihenfolge der Indices.). Dies ist eine Darstellung, die reguläre Darstellung. Der Grad dieser Darstellung ist g. Wegen $e_s = \rho_s(e_1)$ bilden die Bilder von e_1 eine Basis von V.

Sei umgekehrt $\tau\colon G \to W$ eine Darstellung von G mit $w \in W$, so dass die Vektoren $\{\tau_s(w)\}_{s\in G}$ eine Basis von W bilden. Dann ist τ vermöge

$$\Phi\colon V \to W, \quad \Phi(e_s) = \tau_s(w),$$

zur regulären Darstellung isomorph, denn es gilt für alle $s, t \in G$

$$\Phi\rho_s\Phi^{-1}(\tau_t(w)) = \tau_s(\tau_t(w)).$$

(iv) Allgemeiner ist folgendes Beispiel: Operiere die Gruppe G auf einer endlichen Menge X durch Permutation, d. h. $x \mapsto sx$ ist für alle s eine Permutation mit den Eigenschaften

$$1x = x \quad \text{und} \quad s(tx) = (st)x$$

für alle $x \in X$, $s, t \in G$.

Sei V ein Vektorraum mit einer Basis $\{e_x\}_{x\in X}$, indiziert durch die Elemente von X. Für $s \in X$ definieren wir eine lineare Abbildung $\rho_s \in \mathrm{Aut}(V)$ durch $e_x \mapsto e_{sx}$. Dies ist eine Darstellung. Sie heißt Permutationsdarstellung.

12.2. **Teildarstellungen.**

Definition 12.2.1. Sei $\rho\colon G \to GL(V)$ eine Darstellung. Sei $W \subset V$ ein Unterraum, der für alle ρ_s invariant (= stabil) ist. Dann heißt W unter G invariant oder stabil.

Somit ist $\rho_s^W := \rho_s|_W \colon W \to W$ in $\mathrm{Aut}(W)$ und es gilt $\rho_{st}^W = \rho_s^W \rho_t^W$. Daher ist $\rho^W \colon G \to GL(W)$ eine lineare Darstellung von G in W. W heißt Teildarstellung von V.

Beispiel 12.2.2. Sei V die reguläre Darstellung von G. Setze

$$x := \sum_{s\in G} e_s, \quad W := \langle x\rangle.$$

Dann gilt $\rho_s x = x$ für alle $s \in G$. Somit ist W eine Teildarstellung von V. Diese Darstellung ist isomorph zur trivialen Darstellung.

Bemerkung 12.2.3. Seien W, W' Unterräume von V mit $V = W \oplus W'$. Dann lässt sich jedes $v \in V$ in eindeutiger Weise als $w + w'$ mit $w \in W$ und $w' \in W'$ schreiben. Wir erhalten also eine Projektion $p\colon V \to V$ (= eine Abbildung mit $p|_{\mathrm{im}\, p} = \mathrm{id}$) von V nach $W \subset V$ mit $v \mapsto w$.

Umgekehrt liefert jede Projektion $p\colon V \to V$ eine Zerlegung von V als direkte Summe: $V = \mathrm{im}\, p \oplus \ker p$.

Wir erhalten eine bijektive Beziehung zwischen den Projektionen p von V nach W und den Komplementen von W in V.

Theorem 12.2.4. *Sei $\rho\colon G \to GL(V)$ eine lineare Darstellung von G in V und sei $W \subset V$ ein unter G invarianter Unterraum. Dann besitzt W in V ein unter G invariantes Komplement.*

Beweis. Wähle ein beliebiges Komplement von W in V: W'. Sei $p\colon V \to V$ die zu $V = W \oplus W'$ gehörige Projektion von V nach W. Sei $g = |G|$. Definiere

$$p^0 := \tfrac{1}{g} \sum_{t\in G} \rho_t p \rho_t^{-1}.$$

Behauptung: Die Abbildung p ist auf W die Identität und bildet V nach W ab: Sei $x \in W$. Dann ist $\rho_t^{-1}x \in W$ nach Voraussetzung. Dort ist p aber die Identität, also $p\rho_t^{-1}x = \rho_t^{-1}x$, $\rho_t p \rho_t^{-1} x = x$ und es folgt, dass p^0 auf W die Identität ist. $\mathrm{im}\, p^0 \subset W$ folgt, da p auf W projiziert und ρ_t den Unterraum W invariant lässt.

Behauptung: Für alle $s \in G$ gilt $\rho_s p^0 = p^0 \rho_s$, die beiden Abbildungen kommutieren also, denn es gilt

$$\rho_s p^0 \rho_s^{-1} = \frac{1}{g}\sum_{t\in G} \rho_s\rho_t p \rho_t^{-1}\rho_s^{-1} = \frac{1}{g}\sum_{t\in G} \rho_{st} p \rho_{st}^{-1} = p^0.$$

Setze $W^0 := \ker p^0$. Wir behaupten, dass W^0 unter G invariant ist: Seien $x \in W^0$ und $s \in G$. Dann folgt $p^0 x = 0$, also $p^0\rho_s x = \rho_s p^0 x = 0$. Somit ist $\rho_s x \in W^0$. W^0 ist also der gesuchte Unterraum. □

Bemerkung 12.2.5.

(i) Ist V ein Skalarproduktraum, so können wir Theorem 12.2.4 auch wie folgt beweisen: Erfüllt das Skalarprodukt nicht $\langle \rho_s x, \rho_s y\rangle = \langle x, y\rangle$ für alle $s \in G$ und $x \in V$, so ersetzen wir es durch $\langle x, y\rangle_\rho := \frac{1}{g}\sum\limits_{t\in G} \langle \rho_t x, \rho_t y\rangle$. Dies erfüllt dann $\langle \rho_s x, \rho_s y\rangle_\rho = \langle x, y\rangle_\rho$. Wir sagen, das Skalarprodukt $\langle\cdot,\cdot\rangle_\rho$ sei unter ρ invariant. (Wir bemerken, dass die Matrixdarstellungen R_t einer Darstellung mit einem invarianten Skalarprodukt bezüglich einer Orthonormalbasis unitär sind.) Dann ist das orthogonale Komplement W^0 von W bezüglich $\langle\cdot,\cdot\rangle_\rho$: Sei

$x \in W^0$ und $y \in W$. Sei $t \in G$ beliebig. Setze $z := \rho_t^{-1} y$. Dann ist $z \in W$ und es folgt
$$0 \stackrel{!}{=} \langle \rho_t x, y \rangle = \langle \rho_t x, \rho_t z \rangle = \langle x, z \rangle = 0.$$

(ii) Ist $V = W \oplus W^0$ eine Zerlegung in unter G invariante Unterräume, so wird ρ_t in einer Basis von V, die aus den Elementen einer Basis von W und von W^0 (in geeigneter Reihenfolge) zusammengesetzt ist, durch eine Matrix der Form
$$\begin{pmatrix} R_t & 0 \\ 0 & R_t^0 \end{pmatrix}$$
dargestellt. R_t und R_t^0 gehören zu Darstellungen W bzw. W^0 von G. Wir sagen, dass V eine direkte Summe der Darstellungen W und W^0 (bzw. ρ^W und ρ^{W^0}) sei. Die direkte Summe von mehr als zwei Darstellungen ist analog definiert.

Definition 12.2.6. Eine Darstellung heißt irreduzibel, wenn sie nicht die direkte Summe von zwei Darstellungen (außer mit dem Vektorraum $\{0\}$) ist.

Per Induktion folgt aus Theorem 12.2.4 direkt

Korollar 12.2.7. *Jede Darstellung ist die direkte Summe von irreduziblen Darstellungen.*

Im Wesentlichen ist diese Zerlegung eindeutig.

12.3. **Tensorprodukt von Darstellungen.**

Definition 12.3.1. Seien $\rho^1 \colon G \to GL(V_1)$ und $\rho^2 \colon G \to GL(V_2)$ zwei lineare Darstellungen. Dann definieren wir das Tensorprodukt $\rho := \rho^1 \otimes \rho^2 \colon G \to GL(V_1 \otimes V_2)$ durch
$$\rho_s(x \otimes y) := \rho_s^1(x) \otimes \rho_s^2(y)$$
für $s \in G$ und $(x, y) \in V_1 \times V_2$. $V_1 \otimes V_2$ heißt das Tensorprodukt der beiden Darstellungen.

Bemerkung 12.3.2. Sei $\{a_i^1\}$ eine Basis von V_1 und $\{a_j^2\}$ eine Basis von V_2. Sei ρ_s^k durch die Matrix $\left({}^k_s m^i_j\right)$ dargestellt, gelte also
$$\rho_s^k\left(a_j^k\right) = \sum_i {}^k_s m^i_j a_i^k.$$
Dann folgt aufgrund der Multilinearität des Tensorproduktes
$$\rho_s\left(a_i^1 \otimes a_j^2\right) = \sum_{k,l} {}^1_s m^k_i \cdot {}^2_s m^l_j \cdot a_k^1 \otimes a_l^2.$$
Nummeriert man die Basiselemente $\left\{a_k^1 \otimes a_k^2\right\}_{k,l}$ mit einem Index, so sieht man, dass der Faktor ${}^1_s m^k_i \cdot {}^2_s m^l_j$ wieder als Eintrag einer größeren Matrix aufgefasst werden kann. Diese bezeichnet man als Tensorprodukt der beiden Matrizen $\left({}^1_s m^k_i\right)_{k,i}$ und $\left({}^2_s m^l_j\right)_{l,j}$.

Bemerkung 12.3.3.
Das Tensorprodukt von zwei irreduziblen Darstellungen braucht selbst nicht wieder irreduzibel zu sein.

Bemerkung 12.3.4 (Symmetrisches Quadrat und alternierendes Quadrat)**.** Sei V ein Vektorraum. Definiere den Automorphismus $\vartheta \colon V \to V$ durch
$$\vartheta(x \otimes y) = y \otimes x.$$
Es gilt $\vartheta^2 = 1$. Den Eigenraum $E_1(\vartheta)$ bezeichnen wir mit $\mathrm{Sym}^2(V)$, den Eigenraum $E_{-1}(\vartheta)$ mit $\mathrm{Alt}^2(V)$. Sei $\{a_i\}_{1 \le i \le n}$ eine Basis von V. Elemente der Form $\{a_i \otimes a_j +$

$a_j \otimes a_i\}_{i \leq j}$ bilden eine Basis von $\mathrm{Sym}^2(V)$ und Elemente der Form $\{a_i \otimes a_j - a_j \otimes a_i\}_{i<j}$ bilden eine Basis von $\mathrm{Alt}^2(V)$. Es gilt

$$\dim \mathrm{Sym}^2(V) = \frac{n(n+1)}{2} \quad \text{und} \quad \dim \mathrm{Alt}^2(V) = \frac{n(n-1)}{2}.$$

Sei $\rho\colon G \to GL(V)$ eine Darstellung. Unter der Darstellung $\rho \otimes \rho\colon G \to GL(V \otimes V)$ sind die Unterräume $\mathrm{Sym}^2(V)$ und $\mathrm{Alt}^2(V)$ invariant. Die zugehörigen Einschränkungen heißen das symmetrische Quadrat bzw. das alternierende Quadrat der Darstellung ρ.

12.4. **Charaktere.**

Definition 12.4.1 (Charakter)**.** Sei $\rho\colon G \to GL(V)$ eine Darstellung. Dann heißt $\chi_\rho\colon G \to \mathbb{C}$ mit

$$\chi_\rho(s) := \operatorname{tr} R_s = \operatorname{tr} \rho_s$$

der Charakter der Darstellung ρ.

Proposition 12.4.2. *Sei χ der Charakter einer Darstellung ρ vom Grad n. Dann gilt:*

(i) $\chi(1) = n$.
(ii) $\chi\left(s^{-1}\right) = \overline{\chi(s)}$ *für* $s \in G$.
(iii) $\chi\left(tst^{-1}\right) = \chi(s)$ *für* $s, t \in G$.

Beweis.

(i) Folgt aus $\rho(1) = \mathbb{1}$.
(ii) Wir haben gesehen, dass R_s bei einem geeigneten Skalarprodukt eine unitäre Matrix ist. Somit besitzt es n komplexe Eigenwerte $\lambda_1, \ldots, \lambda_n$ (mit Vielfachheiten gezählt) vom Betrag 1. Es folgt

$$\overline{\chi(s)} = \overline{\operatorname{tr} \rho_s} = \sum_{i=1}^n \overline{\lambda_i} = \sum_{i=1}^n \lambda_i^{-1} = \operatorname{tr}\left(\rho_s^{-1}\right) = \operatorname{tr} \rho_{s^{-1}} = \chi\left(s^{-1}\right).$$

(iii) Wir setzen $u = ts$ und $v = t^{-1}$. Dann ist die Behauptung äquivalent zu $\chi(uv) = \chi(vu)$ und folgt aus $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. □

Proposition 12.4.3. *Seien $\rho^1\colon G \to GL(V_1)$ und $\rho^2\colon G \to GL(V_2)$ zwei lineare Darstellungen von G und χ_1 und χ_2 ihre Charaktere. Dann gilt:*

(i) Der Charakter χ der direkten Summendarstellung $V_1 \oplus V_2$ erfüllt $\chi = \chi_1 + \chi_2$.
(ii) Der Charakter ψ der Tensorproduktdarstellung $V_1 \otimes V_2$ erfüllt $\psi = \chi_1 \cdot \chi_2$.

Beweis.

(i) In Matrixform ist die Darstellung durch

$$R_s = \begin{pmatrix} R_s^1 & 0 \\ 0 & R_s^2 \end{pmatrix}$$

gegeben. Somit folgt $\operatorname{tr} R_s = \operatorname{tr} R_s^1 + \operatorname{tr} R_s^2$ und somit $\chi = \chi_1 + \chi_2$.
(ii) Wir benutzen die Matrixdarstellung aus Bemerkung 12.3.2 und erhalten

$$\psi(s) = \sum_{i,j} {}^1_s m_i^i \cdot {}^2_s m_j^j = \chi_1(s) \cdot \chi_2(s). \qquad \square$$

Proposition 12.4.4. *Sei $\rho\colon G \to GL(V)$ eine lineare Darstellung mit Charakter χ. Sei χ_σ^2 der Charakter der Darstellung $\mathrm{Sym}^2(V)$ und χ_α^2 der Charakter der Darstellung $\mathrm{Alt}^2(V)$. Dann gilt für alle $s \in G$*

$$\chi_\sigma^2(s) = \tfrac{1}{2}\left(\chi(s)^2 + \chi\left(s^2\right)\right),$$
$$\chi_\alpha^2(s) = \tfrac{1}{2}\left(\chi(s)^2 - \chi\left(s^2\right)\right)$$

und

$$\chi^2_\sigma(s) + \chi^2_\alpha(s) = \chi(s)^2.$$

Beweis. Sei $s \in G$. Da sich ρ_s als unitäre Matrix darstellen lässt, können wir eine Basis $\{a_i\}_i$ aus Eigenvektoren von ρ_s zu Eigenwerten $\{\lambda_i\}_i$ wählen.

Es folgt

$$\chi(s) = \sum \lambda_i \quad \text{und} \quad \chi\left(s^2\right) = \sum \lambda_i^2.$$

Weiterhin erhalten wir aus

$$(\rho_s \otimes \rho_s)(a_i \otimes a_j + a_j \otimes a_i) = \lambda_i\lambda_j(a_i \otimes a_j + a_j \otimes a_i)$$

und

$$(\rho_s \otimes \rho_s)(a_i \otimes a_j - a_j \otimes a_i) = \lambda_i\lambda_j(a_i \otimes a_j - a_j \otimes a_i)$$

die Beziehungen

$$\chi^2_\sigma(s) = \sum_{i\le j} \lambda_i\lambda_j = \sum_i \lambda_i^2 + \sum_{i<j} \lambda_i\lambda_j = \tfrac{1}{2}\left(\sum_i \lambda_i\right)^2 + \tfrac{1}{2}\sum_i \lambda_i^2$$

und

$$\chi^2_\alpha(s) = \sum_{i<j} \lambda_i\lambda_j = \tfrac{1}{2}\left(\sum_i \lambda_i\right)^2 - \tfrac{1}{2}\sum_i \lambda_i^2.$$

Die Behauptung folgt. (Beachte, dass

$$V \otimes V = \operatorname{Sym}^2(V) \oplus \operatorname{Alt}^2(V)$$

mit Proposition 12.4.3 einen weiteren Beweis von $\chi^2 = \chi^2_\sigma + \chi^2_\alpha$ liefert.) □

12.5. **Lemma von Schur.**

Proposition 12.5.1 (Lemma von Schur)**.** *Seien $\rho^1\colon G \to GL(V_1)$ und $\rho^2\colon G \to GL(V_2)$ zwei irreduzible Darstellungen von G. Sei $f\colon V_1 \to V_2$ linear mit $\rho^2_s f = f\rho^1_s$ für alle $s \in G$. Dann gilt:*

(i) Sind ρ^1 und ρ^2 nicht isomorph, so gilt $f = 0$.
(ii) Ist $V_1 = V_2$ und $\rho^1 = \rho^2$, so ist f ein skalares Vielfaches der Identiät (= eine Homothetie).

Beweis.

(i) Sind ρ^1 und ρ^2 nicht isomorph, so ist der Kern von f nicht trivial. Sei $a \in \ker f$. Die Menge $\{\rho^1_s a\}_{s\in G}$ erzeugt einen unter ρ^1 invarianten Unterraum U_1. Da ρ^1 jedoch irreduzibel ist, folgt $U_1 = V_1$. Wir erhalten $0 = \rho^2_s f a = f\rho^1_s a$. Somit ist $\{\rho^1_s a\}_{s\in G} \subset \ker f$. Da $\ker f$ ein Unterraum ist, gilt die Inklusion auch, wenn wir links zum Erzeugnis übergehen. Also ist $U_1 = V_1 \subset \ker f$ und daher $f \equiv 0$.
(ii) Sei λ ein Eigenwert von f. Da $\mathbb{C}$ der zugrunde liegende Körper ist, existiert zumindest ein Eigenwert. Setze $f' := f - \lambda$. Es gilt $\ker f' \neq \{0\}$. Nun liefert der erste Teil $f' \equiv 0$, also $f = \lambda\,\mathrm{id}$. □

Korollar 12.5.2. *Seien V_1 und V_2 zwei irreduzible Darstellungen von G mit Gruppenordnung g, d. h. $g = |G|$. Sei $h\colon V_1 \to V_2$ linear. Setze*

$$h^0 := \frac{1}{g}\sum_{t\in G} (\rho^2_t)^{-1} h \rho^1_t.$$

Dann gilt

(i) Sind ρ^1 und ρ^2 nicht isomorph, so gilt $h^0 = 0$.

(ii) Sind $V_1 = V_2$, $\rho^1 = \rho^2$ *und* $\dim V_1 = n$, *so gilt* $h^0 = \frac{1}{n} \operatorname{tr} h \cdot \mathrm{id}$.

Beweis. Die Abbildung h^0 erfüllt die Bedingungen an f aus dem Schurschen Lemma, denn es gilt

$$\left(\rho_s^2\right)^{-1} h^0 \rho_s^1 = \frac{1}{g} \sum_{t \in G} \left(\rho_s^2\right)^{-1} \left(\rho_t^2\right)^{-1} h \rho_t^1 \rho_s^1 = \frac{1}{g} \sum_{t \in G} \left(\rho_{ts}^2\right)^{-1} h \rho_{ts}^1 = h^0.$$

Für nicht isomorphe Darstellungen ρ^1 und ρ^2 liefert daher das Schursche Lemma $h^0 = 0$. Sonst liefert das Schursche Lemma $h^0 = \lambda \, \mathrm{id}$ für ein $\lambda \in F$. Es folgt

$$n\lambda = \operatorname{tr}(\lambda \, \mathrm{id}) = \operatorname{tr} h^0 = \frac{1}{g} \sum_{t \in G} \operatorname{tr} \left(\left(\rho_t^1\right)^{-1} h \rho_t^1 \right) = \operatorname{tr} h. \qquad \square$$

Mit naheliegenden Bezeichnungen erhalten wir daraus in Matrixform

$$\left(h^0\right)^i_j = \frac{1}{g} \sum_{t \in G} \sum_{k,l} \left(\rho^2_{t^{-1}}\right)^i_k h^k_l \left(\rho^1_t\right)^l_j.$$

Für nicht isomorphe irreduzible Darstellungen ρ^1 und ρ^2 gilt nach Korollar 12.5.2 $h^0 = 0$, unabhängig von h. Da die rechte Seite in der obigen Gleichung linear in h ist, ist dies nur möglich, wenn der Rest der Summe bereits verschwindet. Wir erhalten also

Korollar 12.5.3. *Seien* ρ^1, ρ^2 *zwei nicht isomorphe irreduzible Darstellungen, so gilt für beliebige* i, j, k, l

$$\frac{1}{g} \sum_{t \in G} \left(\rho^2_{t^{-1}}\right)^i_k \left(\rho^1_t\right)^l_j = 0.$$

Im Fall $V_1 = V_2$, $\rho^1 = \rho^2$, $\dim V_1 = n$ gilt $h^0 = \frac{1}{n} \operatorname{tr} h \cdot \mathrm{id}$, d. h. in Matrixform

$$\frac{1}{g} \sum_{t \in G} \sum_{k,l} \left(\rho^1_{t^{-1}}\right)^i_k h^k_l \left(\rho^1_t\right)^l_j = \frac{1}{n} \sum_k h^k_k \delta^i_j = \frac{1}{n} \sum_{k,l} \delta^i_j h^k_l \delta^l_k.$$

Da dies für jede lineare Abbildung $h \colon V_1 \to V_2$ gilt, erhalten wir

Korollar 12.5.4. *Seien* $V_1 = V_2$, $\rho^1 = \rho^2$, $\dim V_1 = n$, *so gilt für beliebige* i, j, k, l

$$\frac{1}{g} \sum_{t \in G} \left(\rho^1_{t^{-1}}\right)^i_k \left(\rho^1_t\right)^l_j = \frac{1}{n} \delta^i_j \delta^l_k.$$

Bemerkung 12.5.5.

(i) Die beiden Korollare werden wir noch als Orthogonalitätsrelationen im Falle unitärer Darstellungen (= die ρ_t darstellenden Matrizen sind unitär) interpretieren.

(ii) Für Funktionen φ, ψ auf G definieren wir

$$(\varphi, \psi) := \frac{1}{g} \sum_{t \in G} \varphi\left(t^{-1}\right) \psi(t) = \frac{1}{g} \sum_{t \in G} \varphi(t) \psi\left(t^{-1}\right).$$

Es gilt $(\varphi, \psi) = (\psi, \varphi)$ und (φ, ψ) ist in φ und ψ linear. Aus den oberen Korollaren erhalten wir mit dieser Notation

$$\left(\left(\rho^1\right)^i_k, \left(\rho^1\right)^l_j \right) = 0 \quad \text{bzw.} \quad \left(\left(\rho^1\right)^i_k, \left(\rho^1\right)^l_j \right) = \frac{1}{n} \delta^i_j \delta^l_k$$

für beliebige i, j, k, l.

12.6. **Orthogonalitätsrelationen für Charaktere.**

Definition 12.6.1. Sei G eine Gruppe der Ordnung g. Seien $\varphi, \psi\colon G \to \mathbb{C}$. Dann definieren wir

$$\langle \varphi, \psi \rangle := \frac{1}{g} \sum_{t \in G} \varphi(t) \overline{\psi(t)}.$$

Bemerkung 12.6.2.

(i) Dies ist ein Skalarprodukt, denn es ist linear in φ, sesquilinear in ψ und es gilt $\langle \varphi, \varphi \rangle > 0$ für $\varphi \not\equiv 0$.

(ii) Setze $\check{\psi}(t) := \overline{\psi\left(t^{-1}\right)}$. Dann folgt

$$\langle \varphi, \psi \rangle = \frac{1}{g} \sum_{t \in G} \varphi(t) \check{\psi}\left(t^{-1}\right) = \left(\varphi, \check{\psi}\right).$$

Nach Proposition 12.4.2 gilt für einen Charakter χ die Beziehung $\chi = \check{\chi}$. Für beliebige Funktionen φ auf G folgt daher $\langle \varphi, \chi \rangle = (\varphi, \chi)$, falls χ ein Charakter ist.

Theorem 12.6.3 (Orthogonalitätsrelationen für Charaktere)**.**

(i) Ist χ der Charakter einer irreduziblen Darstellung, so gilt $\langle \chi, \chi \rangle = 1$.

(ii) Sind χ und ψ die Charaktere von zwei nicht isomorphen irreduziblen Darstellungen, so gilt $\langle \chi, \psi \rangle = 0$.

Beweis.

(i) Sei ρ eine irreduzible Darstellung mit Charakter χ, in Matrixform $\rho_t = \left(r^i_j(t)\right)$. Aus $\chi(t) = \sum\limits_i r^i_i(t)$ folgt

$$\langle \chi, \chi \rangle = (\chi, \chi) = \sum_{i,j} \left(r^i_i, r^j_j\right).$$

Nach Korollar 12.5.4 gilt $\left(r^i_i, r^j_j\right) = \frac{1}{n} \delta^i_j \delta^j_i$, wobei n der Grad der Darstellung ist. Wir erhalten also

$$\langle \chi, \chi \rangle = \tfrac{1}{n} \sum_{i,j=1}^{n} \delta^i_j \delta^j_i = \tfrac{n}{n} = 1.$$

(ii) Seien ρ und σ zwei nicht isomorphe irreduzible Darstellungen von G mit Charakteren χ und ψ und Matrixdarstellungen $\left(r^i_j(t)\right)$ bzw. $\left(s^i_j(t)\right)$. Dann folgt

$$\langle \chi, \psi \rangle = (\chi, \psi) = \sum_{i,j} \left(r^i_i, s^j_j\right) = \sum_{i,j} \underbrace{\tfrac{1}{g} \sum_{t \in G} r\left(t^{-1}\right)^i_i s(t)^j_j}_{=0} = 0$$

nach Korollar 12.5.3. ☐

Bemerkung 12.6.4. Wir nennen den Charakter einer irreduziblen Darstellung einen irreduziblen Charakter. Theorem 12.6.3 liefert, dass irreduzible Charaktere (zu nicht isomorphen Darstellungen) ein normiertes Orthogonalsystem bilden.

Theorem 12.6.5. *Sei V eine lineare Darstellung von G mit Charakter χ. Nehme an, dass V sich als*

$$V = V_1 \oplus \ldots \oplus V_k$$

in irreduzible Darstellungen zerlegen lässt. Sei W eine irreduzible Darstellung von G mit Charakter φ. Dann ist die Anzahl der Summanden V_i, die zu W isomorph sind, durch $\langle \chi, \varphi \rangle = (\chi, \varphi)$ gegeben.

Beweis. Nach Proposition 12.4.3 erhalten wir $\chi = \chi_1 + \ldots + \chi_k$, wobei die Funktionen χ_i die Charaktere zu den Darstellungen V_i bezeichnen. Wir erhalten $\langle \chi, \varphi \rangle = \langle \chi_1, \varphi \rangle + \ldots + \langle \chi_k, \varphi \rangle$. Nach Theorem 12.6.3 sind die Summanden $\langle \chi_i, \varphi \rangle$ genau dann gleich 1, wenn V_i eine Darstellung isomorph zu W ist und sonst gleich 0. Die Behauptung folgt. □

Korollar 12.6.6. *Die Anzahl der Darstellungen V_i, die isomorph zu W sind, hängt nicht von der gewählten Zerlegung von V in irreduzible Darstellungen V_i ab.*

Wir nennen diese Zahl die Anzahl, wie oft W in V enthalten ist oder vorkommt.

In diesem Sinne ist die Zerlegung einer Darstellung in irreduzible Darstellungen eindeutig.

Korollar 12.6.7. *Zwei Darstellungen mit dem gleichen Charakter sind isomorph (und trivialerweise auch umgekehrt).*

Beweis. Da die Anzahl der zu einer gegebenen irreduziblen Darstellung isomorphen Summanden nach Korollar 12.6.6 übereinstimmt, folgt die Behauptung. □

Bemerkung 12.6.8. Aufgrund der obigen Resultate genügt es daher für die Untersuchung von Darstellungen, ihre Charaktere zu untersuchen.

Seien nämlich $\chi_1, \ldots, \chi_h$ die verschiedenen irreduziblen Charaktere von G (deren Endlichkeit wir erst im nächsten Kapitel nachweisen werden) mit zugehörigen Darstellungen $W_1, \ldots, W_h$, so ist jede Darstellung von V isomorph zu einer direkten Summe

$$V = m_1 W_1 \oplus \ldots \oplus m_h W_h,$$

wobei $m_i \in \mathbb{N}$ gilt und mW für die m-fache direkte Summe $W \oplus \ldots \oplus W$ steht. Der Charakter φ von V erfüllt $\varphi = m_1\chi_1 + \ldots + m_h\chi_h$. Es gilt $m_i = \langle \varphi, \chi_i \rangle$.

(Dies gilt insbesondere auch für das Tensorprodukt zweier irreduzibler Darstellungen $W_i \otimes W_j$ und wir erhalten $\chi_i\chi_j = \sum\limits_k m_{ij}^k \chi_k$, $m_{ij}^k \in \mathbb{N}$. Produkte von Charakteren sind also wieder Linearkombinationen von Charakteren mit $\mathbb{N}$-Koeffizienten.)

Aufgrund der Orthonormalität der irreduziblen Charaktere χ_i folgt im ersten Beispiel $\varphi = \sum\limits_{i=1}^h m_i\chi_i$ vom Anfang der Bemerkung

$$\langle \varphi, \varphi \rangle = \sum_{i=1}^h m_i^2. \tag{12.1}$$

Daraus erhalten wir direkt ein Theorem, das wir auch als Irreduzibilitätskriterium verwenden können.

Theorem 12.6.9. *Sei φ der Charakter einer Darstellung $V \neq \{0\}$. Dann ist $\langle \varphi, \varphi \rangle \in \mathbb{N} \setminus \{0\}$ und es gilt genau dann $\langle \varphi, \varphi \rangle = 1$, wenn V irreduzibel ist.*

Beweis. $\langle \varphi, \varphi \rangle \in \mathbb{N} \setminus \{0\}$ folgt direkt aus (12.1). Hieraus folgt auch, dass $\langle \varphi, \varphi \rangle$ genau dann gleich 1 ist, wenn nur genau ein m_i gleich 1 ist und die anderen verschwinden. In diesem Falle ist V aber isomorph zu einer irreduziblen Darstellung. □

12.7. **Zerlegung der regulären Darstellung.** Wir wollen ab jetzt annehmen, dass $\chi_1, \ldots, \chi_h$ die irreduziblen Charaktere von G sind. ($h < \infty$ wird bald gezeigt.) Seien $n_1, \ldots, n_h$ die Grade zugehöriger irreduzibler Darstellungen. Dann gilt nach Proposition 12.4.2 $n_i = \chi_i(1)$.

Proposition 12.7.1. *Der Charakter r_G der regulären Darstellung ist durch*

$$r_G(s) = \begin{cases} g = |G|, & s = 1, \\ 0, & s \neq 1 \end{cases}$$

gegeben.

Beweis. Wir erinnern daran, dass die reguläre Darstellung R auf dem von $\{e_t\}_{t\in G}$ erzeugten Vektorraum $\rho_s e_t = e_{st}$ erfüllt.

Ist $s \neq 1$, so gilt $st \neq t$ für alle t. Somit verschwinden alle Diagonaleinträge der ρ_s darstellenden Matrix. Also folgt $\operatorname{tr}\rho_s = 0$. Schließlich gilt im Fall $s = 1$

$$r_G(1) = \operatorname{tr}\rho_1 = \operatorname{tr}\operatorname{id} = \dim R = g. \qquad \square$$

Korollar 12.7.2. *Jede irreduzible Darstellung W_i ist in der regulären Darstellung mit Vielfachheit n_i, ihrem Grad, enthalten.*

Insbesondere gibt es also für eine gegebene endliche Gruppe G nur endlich viele irreduzible Charaktere.

Beweis. Sei χ_i der zugehörige irreduzible Charakter. Nach Theorem 12.6.5 ist die Anzahl, wie oft W_i in R vorkommt, durch $\langle r_G, \chi_i\rangle$ gegeben. Wir erhalten

$$\langle r_G, \chi_i\rangle = \tfrac{1}{g}\sum_{t\in G} r_G\left(t^{-1}\right)\chi_i(t) = \frac{1}{g}g\chi_i(1) = \chi_i(1) = n_i. \qquad \square$$

Korollar 12.7.3.

(i) Die Grade n_i der irreduziblen (paarweise nicht isomorphen) Darstellungen einer Gruppe G erfüllen $\sum\limits_{i=1}^{h} n_i^2 = g$.

(ii) Ist $s \neq 1$, so folgt $\sum\limits_{i=1}^{h} n_i\chi_i(s) = 0$.

Beweis. Nach Korollar 12.7.2 folgt $r_G(s) = \sum\limits_i n_i\chi_i(s)$ für alle $s \in G$.

(i) Wähle $s = 1$.

(ii) Wähle $s \neq 1$. $\square$

Bemerkung 12.7.4. Kennt man paarweise nichtisomorphe Darstellungen einer Gruppe mit Graden n_i, so sind dies genau dann alle solchen Darstellungen, wenn $\sum\limits_i n_i^2 = g$ gilt.

Man kann auch zeigen, dass die Grade n_i die Gruppenordnung g teilen.

12.8. **Anzahl irreduzibler Darstellungen.**

Definition 12.8.1. Eine Funktion f auf einer Gruppe G heißt Klassenfunktion, falls $f\left(tst^{-1}\right) = f(s)$ für alle $s, t \in G$ gilt.

Sei $t \in G$. Dann heißt

$$\{sts^{-1} : s \in G\}$$

die Konjugationsklasse von t in G. Die Konjugationsklassen bilden eine Partition von G. Eine Klassenfunktion ist auf jeder Konjugationsklasse von G konstant.

Proposition 12.8.2. *Sei f eine Klassenfunktion auf G. Sei $\rho\colon G \to GL(V)$ eine lineare Darstellung von G. Definiere $\rho_f\colon V \to V$ durch*

$$\rho_f := \sum_{t\in G} f(t)\rho_t.$$

Ist V eine irreduzible Darstellung vom Grad n und Charakter χ, dann gilt $\rho_f = \lambda\operatorname{id}$ mit

$$\lambda = \tfrac{1}{n}\sum_{t\in G} f(t)\chi(t) = \tfrac{g}{n}\langle f, \overline{\chi}\rangle.$$

Beweis. Es gilt

$$\rho_s^{-1}\rho_f\rho_s = \sum_{t\in G} f(t)\rho_s^{-1}\rho_t\rho_s = \sum_{t\in G} f(t)\rho_{s^{-1}ts}.$$

Mit $u = s^{-1}ts$ erhalten wir

$$\rho_s^{-1}\rho_f\rho_s = \sum_{u\in G} f\left(sus^{-1}\right)\rho_u = \sum_{u\in G} f(u)\rho_u = \rho_f.$$

Somit gilt $\rho_f\rho_s = \rho_s\rho_f$. Nach dem Schurschen Lemma ist ρ_f daher eine Homothetie, also von der Form $\rho_f = \lambda\,\mathrm{id}$. Wir erhalten daher

$$n\lambda = \mathrm{tr}(\lambda\,\mathrm{id}) = \mathrm{tr}\,\rho_f = \sum_{t\in G} f(t)\,\mathrm{tr}\,\rho_t = \sum_{t\in G} f(t)\chi(t) = g\langle f, \overline{\chi}\rangle. \qquad \square$$

Theorem 12.8.3. *Sei H der Raum der Klassenfunktionen auf G. Dann sind die irreduziblen Charaktere $\chi_1, \ldots, \chi_n \in H$ eine Orthonormalbasis von H.*

Beweis. Es ist klar, dass Charaktere Klassenfunktionen sind. Nach Theorem 12.6.3 bilden die irreduziblen Charaktere eine orthonormierte Menge. Um zu zeigen, dass sie auch H erzeugen, weisen wir nach, dass eine Klassenfunktion $\overline{f}$, die auf allen χ_i's senkrecht steht, die Nullfunktion ist. $\left\langle\overline{f}, \chi_i\right\rangle = 0$ ist äquivalent zu $\langle f, \overline{\chi_i}\rangle = 0$. f ist ebenfalls eine Klassenfunktion. Es genügt der Nachweis, dass $f = 0$ ist. Die in Proposition 12.8.2 definierte Abbildung ρ_f ist in der vorliegenden Situation Null, falls ρ irreduzibel ist. Ein beliebiges ρ ist aber die direkte Summe von irreduziblen Darstellungen. Somit ist $\rho_f = 0$ für eine beliebige Darstellung ρ. Wir wenden dies nun insbesondere auf die reguläre Darstellung R an und erhalten

$$0 = \rho_f e_1 = \sum_{t\in G} f(t)\rho_t e_1 = \sum_{t\in G} f(t)e_t.$$

Dies ist aber nur möglich, wenn $f(t) = 0$ für alle $t \in G$ gilt. Die Behauptung folgt. $\square$

Theorem 12.8.4. *Die Anzahl der irreduziblen Darstellungen von G (bis auf Isomorphismen) stimmt mit der Anzahl der Konjugationsklassen in G überein.*

Beweis. Sie $C_1, \ldots, C_k$ die Menge der Konjugationsklassen in G. Eine Funktion ist genau dann eine Klassenfunktion, wenn sie auf Konjugationsklassen konstant ist. Somit hat der Raum H der Klassenfunktionen auf G die Dimension k.

Nach Theorem 12.8.3 stimmt aber die Anzahl der irreduziblen Charaktere mit der Dimension von H überein. Somit folgt die Behauptung. $\square$

Proposition 12.8.5. *Sei $s \in G$. Sei $c(s)$ die Anzahl der Elemente in der Konjugationsklasse von s in G und h die Anzahl der Konjugationsklassen. Dann gilt*

(i) $\sum\limits_{i=1}^{h} \overline{\chi_i(s)}\chi_i(s) = \frac{g}{c(s)}$.

(ii) Ist t nicht zu s konjugiert, liegen t und s also in verschiedenen Konjugationsklassen, so gilt $\sum\limits_{i=1}^{h} \overline{\chi_i(s)}\chi_i(t) = 0$.

Beweis. Sei f_s die Klassenfunktion, die auf der Konjugationsklasse von s den Wert 1 annimmt und sonst verschwindet. Nach Theorem 12.8.3 können wir f_s in der Form

$$f_s = \sum_{i=1}^{h} \lambda_i\chi_i \quad \text{mit} \quad \lambda_i = \langle f_s, \chi_i\rangle = \frac{c(s)}{g}\overline{\chi_i(s)}$$

darstellen. Für beliebiges $t \in G$ folgt

$$f_s(t) = \frac{c(s)}{g}\sum_{i=1}^{h} \overline{\chi_i(s)}\chi_i(t).$$

Für $t = s$ ist $f_s(s) = 1$ und die erste Behauptung folgt. Ist t nicht zu s konjugiert, so erhalten wir aus $f_s(t) = 0$ die zweite Behauptung. $\square$

12.9. **Kanonische Zerlegung einer Darstellung.** Sei $\rho\colon G \to GL(V)$ eine lineare Darstellung von G. Wir wollen eine kanonische Zerlegung in direkte Summanden herleiten, die gröber als eine Zerlegung in irreduzible Teildarstellungen ist, aber dafür eindeutig bestimmt ist.

Seien $\chi_1, \ldots, \chi_h$ paarweise verschiedene irreduzible Charaktere zu irreduziblen Teildarstellungen $W_1, \ldots, W_h$ von G mit Graden $n_1, \ldots, n_h$. Sei $V = U_1 \oplus \ldots \oplus U_m$ eine Zerlegung von V als direkte Summe irreduzibler Darstellungen. Bezeichne mit V_i die direkte Summe aller U_k's, die zu W_i isomorph sind. Dann gilt

$$V = V_1 \oplus \ldots \oplus V_h.$$

(Wir haben also in der Zerlegung von V in direkte irreduzible Summanden die isomorphen zusammengefasst.) Diese Zerlegung heißt kanonische Zerlegung. Über ihre Eigenschaften beweisen wir

Theorem 12.9.1. *Unter den obigen Voraussetzungen und Bezeichnungen gilt:*

(i) Die Zerlegung $V = V_1 \oplus \ldots \oplus V_h$ hängt nicht von der ursprünglich gewählten Zerlegung in irreduzible Darstellungen ab.

(ii) Die Projektionen $p_i\colon V \to V$ mit $\operatorname{im} p_i = V_i$ zu dieser Zerlegung sind durch

$$p_i = \tfrac{n_i}{g} \sum_{t \in G} \overline{\chi_i(t)} \rho_t$$

gegeben.

Beweis. Da die Projektionen die Zerlegung definieren, zeigen wir nur den zweiten Teil. Wir schränken p_i auf eine irreduzible Darstellung W vom Grad n mit Charakter χ ein. Dann ist nach Proposition 12.8.2 $p_i|_W = \lambda \operatorname{id}$ mit

$$\lambda = \tfrac{n_i}{g} \tfrac{g}{n} \langle \overline{\chi_i}, \overline{\chi} \rangle = \tfrac{n_i}{n} \overline{\langle \chi, \chi_i \rangle}.$$

Es gilt $\lambda = 1$ für $\chi_i = \chi$ und sonst $\lambda = 0$. D. h. die Einschränkungen von p_i sind die Identität auf irreduziblen Summanden, die zu W_i isomorph sind und sonst die Nullabbildung. Schreiben wir also $x \in V$ gemäß der Zerlegung $V = V_1 \oplus \ldots \oplus V_h$ als $x = x_1 + \ldots + x_h$, so folgt $p_i(x) = p_i(x_1) + \ldots + p_i(x_h) = x_i$. Somit ist p_i die gesuchte Projektion. □

Literatur

1. Siegfried Bosch, *Algebra*, 6. Auflage, Springer-Verlag, Berlin, 2005.
2. Siegfried Bosch, *Lineare Algebra*, 4. Auflage, Springer-Verlag, Berlin, 2008.
3. Gerd Fischer, *Lineare Algebra*, fifth ed., Grundkurs Mathematik, vol. 17, Friedr. Vieweg & Sohn, Braunschweig, 1979.
4. Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002, `http://www.math.cornell.edu/~hatcher/`.
5. Roger A. Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994, Corrected reprint of the 1991 original.
6. Serge Lang, *Algebra*, third ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002.
7. Oliver Schnürer, *Lineare Algebra I*, 2010, Skript zur Vorlesung.
8. Jean-Pierre Serre, *Linear representations of finite groups*, Springer-Verlag, New York, 1977, Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42.
9. Urs Stammbach, *Lineare Algebra*, Teubner Studienskripten, vol. 82, B. G. Teubner, Stuttgart, 1980.

Oliver C. Schnürer, Fachbereich Mathematik und Statistik, Universität Konstanz, 78457 Konstanz, Germany

E-mail address: `Oliver.Schnuerer@uni-konstanz.de`