Lineare Algebra und analytische Geometrie II (MIIB)

Helmut Schwichtenberg

Mathematisches Institut der Ludwig-Maximilians-Universität München
Sommersemester 1998

Inhaltsverzeichnis

Vorwort

Dieses Skriptum gibt den Inhalt der Vorlesung Lineare Algebra und analytische Geometrie II (MIIB) wieder, die ich im Sommersemester 1998 am Mathematischen Institut der Universität München gehalten habe. Quellen waren Vorlesungen meines Kollegen Hans-Jürgen Schneider sowie das Buch von Gerd Fischer [1] über lineare Algebra.

Bedanken möchte ich mich bei Ulrich Berger und Klaus Weich für ihre Unterstützung bei der Vorbereitung der Vorlesung, und bei Martin Kübler und Anke Wanger für ihre Hilfe bei der Herstellung des Manuskripts. Für die gute Zusammenarbeit und viele hilfreiche Anmerkungen bedanke ich mich bei den Übungsgruppenleitern und Korrektoren Mathias Flurl, Stefanie Kuntze, Wolfgang Sagmeister und Christian Wittmann.

München, im Juli 1998
Helmut Schwichtenberg

Kapitel 7

Eigenwerte und Eigenvektoren, Diagonalisierung

In Abschnitt 5.4 hatten wir im Kontext der euklidischen Geometrie zu einer gegebenen Matrix $A \in \mathbb{R}^{n\times n}$ nach einer *orthogonalen* Matrix $C \in \mathrm{O}(n)$ gesucht, so daß C^tAC Diagonalgestalt hat. Das Ergebnis war, daß zu einer symmetrischen Matrix A ein solches C stets gefunden werden kann. Die Hauptachsentransformation nach JACOBI lieferte ein Verfahren, ein solches $C \in \mathrm{O}(n)$ zu konstruieren. (Vorher hatten wir in 4.2.2 mit einem einfacheren Argument ein im Fall $K = \mathbb{R}$ schwächeres Resultat bewiesen, nämlich daß man zu jeder symmetrischen Matrix $A \in K^{n\times n}$ ein $C \in \mathrm{GL}(n, K)$ konstruieren kann, so daß C^tAC Diagonalgestalt hat). Offenbar kann ein solches orthogonales C auch nur für symmetrisches A existieren.

Wir kommen jetzt wieder auf die affine Geometrie zurück und wollen uns mit der Frage befassen, ob und gegebenenfalls wie man eine nicht notwendig symmetrische Matrix $A \in K^{n\times n}$ mittels einer *invertierbaren* Matrix $C \in \mathrm{GL}(n, K)$ durch Bilden von CAC^{-1} in Diagonalgestalt bringen kann. Es wird sich zeigen, daß es dafür wichtig ist, die 'Richtungen' zu identifizieren, in denen der Endomorphismus f_A als Streckung wirkt (sie werden durch die sogenannten Eigenvektoren angegeben), und ferner die Streckungsfaktoren (sogenannte Eigenwerte) zu bestimmen. Die Eigenwerte kann man als Nullstellen eines gewissen, für den Endomorphismus f_A charakteristischen Polynoms ausrechnen. Es ist deshalb nützlich, mit einigen elementaren Betrachtungen über den Polynombegriff zu beginnen.

7.1 Polynome

In 2.1.2 hatten wir $\mathbb{R}[t]$, also die Menge aller Polynome in einer Variablen t mit Koeffizienten aus $\mathbb{R}$, als Beispiel eines Vektorraums eingeführt. Damals hatten wir Polynome $\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0$ als formale Ausdrücke aufgefaßt, und Addition und Skalarmultiplikation definiert durch

$$(\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0) + (\beta_n t^n + \cdots + \beta_1 t + \beta_0) = (\alpha_n{+}\beta_n)t^n + \cdots + (\alpha_1{+}\beta_1)t + (\alpha_0{+}\beta_0),$$
$$\alpha(\alpha_n t^n + \cdots + \alpha_1 t + \alpha_0) = (\alpha\alpha_n t^n + \cdots + \alpha\alpha_1 t + \alpha\alpha_0).$$

Wir wollen jetzt den Polynombegriff etwas genauer fassen, und gleichzeitig auch Polynome (in einer Veränderlichen) über einem beliebigen Ring einführen. Es wird sich zeigen, daß man wieder einen Ring erhält, und zwar den sogenannten Polynomring $K[t]$.

Jedem Polynom $f = \sum_{i=0}^n \alpha_i t^i \in K[t]$ kann man eine Abbildung von K nach K zuordnen durch $\gamma \mapsto \sum_{i=0}^n \alpha_i \gamma^i \in K$. Für allgemeine Körper bestimmt eine solche Abbildung jedoch *nicht* die Koeffizienten. Zum Beispiel für $\mathbb{F}_2$ und $f = t^2 + t$ ist $f(0) = f(1) = 0$, also die zugeordnete Abbildung die Nullabbildung, obwohl die Koeffizienten nicht Null sind. Dies ist ein Grund, Polynome als formale Ausdrücke einzuführen.

7.1.1. *Der Polynomring $K[t]$.* Unser Ansatz zu einer genauen Fassung des Polynombegriffs ist es, daß ein formaler Ausdruck $\alpha_n t^n + \dots + \alpha_1 t + \alpha_0$ eindeutig festgelegt ist durch die Folge $(\alpha_0, \alpha_1, \dots, \alpha_n, 0, 0, \dots)$ der Koeffizienten.

Definition. Sei K ein Ring und

$K^{(\mathbb{N})} := \{ (\alpha_0, \alpha_1, \alpha_2, \dots) \in K^{\mathbb{N}} \mid \alpha_i \in K, \text{ fast alle } \alpha_i \text{ (d.h. alle bis auf endlich viele) sind } 0 \}.$

Die Elemente $f \in K^{(\mathbb{N})}$ nennt man *Polynome* über K. Man schreibt auch

$$K[t] := \{ f \mid f \text{ Polynom über } K \} = K^{(\mathbb{N})}.$$

Für $f, g \in K[t]$, etwa $f = (\alpha_0, \alpha_1, \dots)$ und $g = (\beta_0, \beta_1, \dots)$ erklären wir Addition $f + g$ und Multiplikation $f \cdot g$ durch

$$\begin{aligned} f + g &= (\alpha_0 + \beta_0, \alpha_1 + \beta_1, \dots), \\ f \cdot g &= (\gamma_0, \gamma_1, \dots) \qquad\qquad \text{mit } \gamma_i = \textstyle\sum_{k+l=i} \alpha_k \cdot \beta_l. \end{aligned}$$

Offenbar sind $f + g, f \cdot g \in K[t]$.

Satz. *Sei K ein (kommutativer) Ring.*

1. *$K[t]$ wird mit obiger Addition und Multiplikation zu einem (kommutativen) Ring mit Null $(0, 0, 0, \dots)$ und Eins $(1, 0, 0, \dots)$. Man nennt $K[t]$ den* Polynomring *über K in einer Veränderlichen.*
2. *Definiert man $t := (0, 1, 0, 0, \dots) \in K[t]$ und identifiziert man $\alpha \in K$ mit $(\alpha, 0, 0, \dots) \in K[t]$, so läßt sich jedes $f \in K[t]$ schreiben als*
$$f = \sum_{i=0}^{n} \alpha_i t^i \quad \text{mit } \alpha_i \in K.$$
(Addition und Multiplikation sind hier in $K[t]$ zu verstehen).
3. *Diese Darstellung ist im folgenden Sinn eindeutig. Gilt $\sum_{i=0}^{n} \alpha_i t^i = \sum_{j=0}^{m} \beta_j t^j$ mit (etwa) $n \le m$, so folgt $\alpha_k = \beta_k$ für $0 \le k \le n$ und $\beta_k = 0$ für $n < k \le m$.*

Beweis. 1. a. $(K[t], +)$ ist abelsche Gruppe, da die Addition komponentenweise definiert ist.

b. Ist K kommutativ, so ist die Multiplikation in K kommutativ, denn für $f = (\alpha_i)_{i \in \mathbb{N}} \in K^{(\mathbb{N})}$ und $g = (\beta_i)_{i \in \mathbb{N}} \in K^{(\mathbb{N})}$ gilt

$$\begin{aligned} f \cdot g &=: (\gamma_i) \quad \text{mit } \gamma_i = \textstyle\sum_{k+l=i} \alpha_k \beta_l \\ g \cdot f &=: (\gamma_i') \quad \text{mit } \gamma_i' = \textstyle\sum_{k+l=i} \beta_k \alpha_l = \sum_{k+l=i} \alpha_l \beta_k = \sum_{l+k=i} \alpha_k \beta_l = \gamma_i. \end{aligned}$$

c. Die Multiplikation ist assoziativ, denn ist noch $h = (\gamma_i)_{i \in \mathbb{N}} \in K^{(\mathbb{N})}$, so gilt ($[c]_i$ bezeichne die i-te Komponente von $c \in K^{(\mathbb{N})}$)

$$\begin{aligned}
[(f \cdot g) \cdot h]_i &= \sum_{k+l=i} [f \cdot g]_k \cdot \gamma_l \\
&= \sum_{k+l=i} \sum_{r+s=k} \alpha_r \beta_s \gamma_l \\
&= \sum_{r+s+l=i} \alpha_r \beta_s \gamma_l \\
&= \sum_{r+q=i} \alpha_r [g \cdot h]_q \\
&= [f \cdot (g \cdot h)]_i.
\end{aligned}$$

d. Das Distributivgesetz ergibt sich aus

$$\begin{aligned}
[f \cdot (g+h)]_i &= \sum_{k+l=i} \alpha_k(\beta_l + \gamma_l) \\
&= \sum_{k+l=i} \alpha_k \beta_l + \sum_{k+l=i} \alpha_k \gamma_l \\
&= [f \cdot g]_i + [f \cdot h]_i \\
&= [f \cdot g + f \cdot h]_i.
\end{aligned}$$

Im Fall eines nicht kommutativen Rings K beweist man $(f+g)h = fh + gh$ genauso.

e. $(1, 0, 0, \dots)$ ist Einselement, denn für alle i gilt

$$[f \cdot (1, 0, 0, \dots)]_i = \alpha_i \cdot 1 + \alpha_{i-1} \cdot 0 + \dots + \alpha_0 \cdot 0 = \alpha_i.$$

Also ist $f \cdot (1, 0, 0, \dots) = f$.

2. Setze $t^0 := (1, 0, 0, \dots)$. Durch Induktion über i sieht man dann leicht, daß

$$t^i = (\underbrace{0, \dots, 0}_{i}, 1, 0, 0, \dots).$$

Es folgt

$$\alpha \cdot t^i = (\alpha, 0, 0, \dots) \cdot t^i = (\underbrace{0, \dots, 0}_{i}, \alpha, 0, 0, \dots) = t^i \cdot \alpha.$$

Sei jetzt $f = (\alpha_i)_{i \in \mathbb{N}}$ und $\alpha_i = 0$ für $i > n$. Dann gilt

$$(\alpha_0, \alpha_1, \dots) = (\alpha_0, 0, 0, \dots) + (0, \alpha_1, 0, 0, \dots) + (0, \dots, 0, \alpha_n, 0, 0, \dots) = \sum_{i=0}^{n} \alpha_i t^i.$$

Damit ist die Existenz der im Satz angegebenen Darstellung von $f \in K[t]$ gezeigt.

3. Die Eindeutigkeit dieser Darstellung ergibt sich mit demselben Argument. QED

7.1.2. *Der Grad eines Polynoms.* Sei $f \in K[t]$, etwa $f = \sum_{i=0}^{n} \alpha_i t^i$ mit $\alpha_n \neq 0$. Dann nennt man die (nach dem eben bewiesenen Satz eindeutig bestimmte) Zahl n den *Grad* des Polynoms f und bezeichnet ihn mit $\deg(f)$. Für das Nullpolynom $0 \in K[t]$ setzt man $\deg(0) := -1$.

LEMMA. *Sei K ein Körper.*

1. *Für $f, g \in K[t]$ mit $f, g \neq 0$ gilt* $\deg(fg) = \deg(f) + \deg(g)$.

2. *Für $f, g, h \in K[t]$ mit $f \neq 0$ folgt aus $fg = fh$ stets $g = h$ (*Kürzungsregel*).*

BEWEIS. 1. Sei $f = \sum_{i=0}^{n} \alpha_i t^i$ mit $\alpha_n \neq 0$ und $g = \sum_{j=0}^{m} \beta_j t^j$ mit $\beta_m \neq 0$. Dann ist $fg = \sum_{k=0}^{n+m} \gamma_k t^k$ mit $\gamma_k = \sum_{i+j=k} \alpha_i \beta_j$, also $\gamma_{n+m} = \alpha_n \beta_m \neq 0$ (hier wird verwendet, daß wir uns in einem Körper befinden).

2. Seien $f, g, h \in K[t]$ mit $f \neq 0$ und nehmen wir an, daß $fg = fh$. Dann ist $f(g - h) = 0$. Wegen $f \neq 0$ folgt $g - h = 0$. QED

7.1.3. *Nullstellen von Polynomen.* Sei K ein Ring, $f = \sum_{i=0}^{n} \alpha_i t^i \in K[t]$ und $\gamma \in K$. Man definiert dann

$$f(\gamma) := \sum_{i=0}^{n} \alpha_i \gamma^i \in K$$

und nennt dieses Ringelement das Ergebnis der *Einsetzung* von γ in f. (Hierbei ist wie üblich $\gamma^0 := 1$ gesetzt für alle $\gamma \in K$). γ heißt *Nullstelle* von f, wenn $f(\gamma) = 0$. Schließlich nennt man das Polynom f *konstant*, wenn $f = 0$ oder $\deg(f) = 0$ ist.

BEMERKUNGEN. 1. Für $f, g \in K[t]$ und $\gamma \in K$ gilt offenbar

$$\begin{aligned}(f + g)(\gamma) &= f(\gamma) + g(\gamma), \\ (f \cdot g)(\gamma) &= f(\gamma) \cdot g(\gamma).\end{aligned}$$

2. Ist $f \in K[t]$ konstant, so hat f genau dann eine Nullstelle, wenn $f = 0$ ist. Ferner hat zum Beispiel $t^2 + 1 \in \mathbb{R}[t]$ keine Nullstelle.

LEMMA. *Sei K ein Ring, $f \in K[t]$ und $\gamma \in K$. Dann ist $f(\gamma) = 0$ genau dann, wenn es ein Polynom $g \in K[t]$ gibt mit $f = (t - \gamma)g$.*

BEWEIS. Gibt es ein Polynom $g \in K[t]$ mit $f = (t - \gamma)g$, so ist $f(\gamma) = (\gamma - \gamma) \cdot g(\gamma) = 0$. Seien nun umgekehrt $f = \sum_{i=0}^{n} \alpha_i t^i \in K[t]$ und $\gamma \in K$ mit $f(\gamma) = \sum_{i=0}^{n} \alpha_i \gamma^i = 0$ gegeben. Dann gilt

$$\begin{aligned} f &= f - \sum_{i=0}^{n} \alpha_i \gamma^i \\ &= \sum_{i=0}^{n} \alpha_i (t^i - \gamma^i) \\ &= \sum_{i=1}^{n} \alpha_i (t - \gamma)(t^{i-1} + \gamma t^{i-2} + \cdots + \gamma^{i-1}) + \alpha_0 (t^0 - \gamma^0) \\ &= (t - \gamma) g + \alpha_0 (1 - 1) \\ &= (t - \gamma) g \end{aligned}$$

für $g = \sum_{i=1}^{n} \alpha_i (t^{i-1} + \gamma t^{i-2} + \cdots + \gamma^{i-1}) \in K[t]$. QED

Folgerung. *Sei K ein Körper, $f \in K[t]$ und $\deg(f) = n \geq 0$. Dann hat f höchstens n Nullstellen.*

Beweis. Induktion nach n. *Basis* $n = 0$. Dann ist f ein konstantes, von 0 verschiedenes Polynom und hat also keine Nullstelle. *Schritt* $n - 1 \Longrightarrow n$ mit $n \geq 1$. Nehmen wir an, daß f verschiedene Nullstellen $\alpha_1, \alpha_2, \dots, \alpha_{n+1}$ besitzt. Nach dem Lemma ist $f = (t - \alpha_1)g$ für ein $g \in K[t]$ mit $\deg(g) = n - 1$. Wegen $\alpha_i - \alpha_1 \neq 0$ für $i = 2, \dots, n + 1$ sind $\alpha_2, \dots, \alpha_{n+1}$ verschiedene Nullstellen von g. Dies ist aber nach Induktionsvoraussetzung nicht möglich. Qed

Definition. Sei K ein Körper, $\lambda \in K$, $f \in K[t]$ und e eine natürliche Zahl ≥ 1. λ heißt *e-fache Nullstelle* von f, wenn sich f schreiben läßt in der Form $f = (t - \lambda)^e g$ mit $g \in K[t]$, $g(\lambda) \neq 0$.

Bemerkung. e ist eindeutig bestimmt.

Beweis. Nehmen wir an $f = (t - \lambda)^e g = (t - \lambda)^{e'} g'$ mit $g, g' \in K[t]$ und $g(\lambda) \neq 0 \neq g'(\lambda)$. Wäre etwa $e < e'$, so hätte man nach der Kürzungsregel $g = (t - \lambda)^{e'-e} g'$, was wegen $g(\lambda) \neq 0$ nicht sein kann. Qed

7.2 Eigenwerte und Eigenvektoren

Gegeben sei ein endlichdimensionaler K-Vektorraum V und eine lineare Abbildung $f \colon V \to V$. Gesucht ist eine Basis $(x_1, \dots, x_n)$ von V so daß f bezüglich $x_1, \dots, x_n$ durch eine Diagonalmatrix dargestellt wird, also

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

7.2.1. Definition. Sei V ein K-Vektorraum und $f \colon V \to V$ linear.

1. Ein $x \in V$ mit $x \neq 0$ heißt *Eigenvektor* von f, wenn es ein $\lambda \in K$ gibt mit
$$f(x) = \lambda x.$$
2. Ein $\lambda \in K$ heißt *Eigenwert* von f, wenn es ein $x \in V$ mit $x \neq 0$ gibt so daß $f(x) = \lambda x$. Man nennt dann x einen Eigenvektor zum Eigenwert λ.
3. Sei λ Eigenwert von f.
$$\operatorname{Eig}(f, \lambda) := \{\, x \in V \mid f(x) = \lambda x \,\}$$
heißt *Eigenraum* von f zum Eigenwert λ.

Bemerkungen. Sei V ein K-Vektorraum und $f \colon V \to V$ linear.

1. Man beachte, daß $\operatorname{Eig}(f, \lambda) = \operatorname{Kern}(f - \lambda \mathrm{id})$ Unterraum von V ist, und nach Definition von Eigenwerten auch $\neq 0$ ist.

2. Hat V eine Basis aus Eigenvektoren $x_1, \dots, x_n$ von f, so wird f bezüglich der Basis $x_1, \dots, x_n$ durch eine Diagonalmatrix dargestellt, und zwar

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

falls x_i Eigenvektor zum Eigenwert λ_i ist. Es gilt nämlich $f(x_i) = \lambda_i x_i$. Man nennt einen solchen Endomorphismus f *diagonalisierbar* (siehe auch 7.4).

3. Im allgemeinen ist eine solche Darstellung eines beliebigen linearen $f\colon V \to V$ nicht möglich; man kommt nur 'beinahe' auf eine Diagonalmatrix. Wir werden dies im nächsten Abschnitt 7.3 genauer klären, wo wir die JORDANsche Normalform behandeln werden.

7.2.2. *Summen von Eigenräumen.* Wir wollen uns mit dem Problem befassen, wie man die Eigenwerte und die zugehörigen Eigenräume eines gegebenen Endomorphismus f berechnen kann. Zunächst können wir feststellen, daß eine Summe von Eigenräumen immer eine direkte Summe ist.

SATZ. *Sei V ein K-Vektorraum und $f\colon V \to V$ linear. Sind $\lambda_1, \dots, \lambda_n$ verschiedene Eigenwerte von f und ist $V_i := \mathrm{Eig}(f, \lambda_i)$ der Eigenraum zu λ_i (also $V_i \neq 0$), so gilt*

$$V_1 + \dots + V_n = V_1 \oplus \dots \oplus V_n \subseteq V,$$

insbesondere also $n \leq \dim(V)$.

BEWEIS. Zu zeigen ist, daß für beliebige $x_1 \in V_1, \dots, x_n \in V_n$ aus $\sum_{i=1}^n x_i = 0$ stets folgt $x_1 = \dots = x_n = 0$. Wir zeigen dies durch Induktion über n. *Basis* $n = 1$. Hier ist nichts zu zeigen. *Schritt* $n-1 \Longrightarrow n$. Gelte $\sum_{i=1}^n x_i = 0$ mit $x_1 \in V_1, \dots, x_n \in V_n$. Dann ist

$$0 = f(\sum_{i=1}^n x_i) = \sum_{i=1}^n f(x_i) = \sum_{i=1}^n \lambda_i x_i.$$

Andererseits hat man

$$0 = \lambda_n \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n \lambda_n x_i.$$

Subtraktion ergibt

$$\begin{aligned} 0 &= \sum_{i=1}^{n-1} (\lambda_n - \lambda_i) x_i \\ 0 &= (\lambda_n - \lambda_i) x_i && \text{für } 1 \leq i \leq n-1 \text{ (Induktionsvoraussetzung)} \\ 0 &= x_i && \text{für } 1 \leq i \leq n-1, \text{ da } \lambda_n \neq \lambda_i \\ 0 &= x_n. && \text{QED} \end{aligned}$$

Insbesondere sind also Eigenvektoren $x_1, \dots, x_n$ zu verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$ immer linear unabhängig. Im Fall eines endlichdimensionalen Vektorraums erhalten wir damit ein einfaches hinreichendes Kriterium für die Diagonalisierbarkeit eines Endomorphismus.

FOLGERUNG. *Sei V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Hat f n verschiedene Eigenwerte, so ist f diagonalisierbar.* QED

Um dieses Kriterium anwenden zu können, braucht man eine Methode, die Eigenwerte eines Endomorphismus f zu berechnen. Wir werden zeigen, daß die Eigenwerte von f gerade die Nullstellen eines gewissen, durch f bestimmten Polynoms sind.

7.2.3. *Das charakteristische Polynom.* Sei V ein n-dimensionaler K-Vektorraum und f ein Endomorphismus, also $f\colon V \to V$ linear. Sei ferner $A \in K^{n\times n}$ eine darstellende Matrix von f. Wir gehen jetzt zum Polynomring $K[t]$ über und betrachten $A - tE$ als Element von $K[t]^{n\times n}$. Man nennt

$$p_A := \det(A - tE) \in K[t]$$

das *charakteristische Polynom* von A. Ist noch $B \in K^{n\times n}$ eine weitere darstellende Matrix von f, so gilt $A \sim B$ nach dem Satz in 3.4.3, d.h. $B = CAC^{-1}$ für ein $C \in \mathrm{GL}(n, K)$. Also ergibt sich für das charakteristische Polynom p_B von B (im Ring $K[t]^{n\times n}$)

$$\begin{aligned}\det(B - tE) &= \det(CAC^{-1} - tCEC^{-1})\\ &= \det(C(A - tE)C^{-1})\\ &= \det(C)\det(A - tE)\det(C)^{-1}\\ &= \det(A - tE),\end{aligned}$$

d.h. $p_A = p_B$. Wir können also das charakteristische Polynom p_f des Endomorphismus f definieren als p_A, wobei A eine beliebige darstellende Matrix von f ist. Im Grenzfall eines Endomorphismus f des Nullraums $V = 0$ setzen wir $p_f := 1$.

SATZ. *Sei V ein endlichdimensionaler K-Vektorraum, $f\colon V \to V$ linear und $\lambda \in K$. Folgende Aussagen sind äquivalent.*

1. *λ ist Eigenwert von f.*
2. *$g := f - \lambda\mathrm{id}$ ist kein Monomorphismus.*
3. *$g := f - \lambda\mathrm{id}$ ist kein Epimorphismus.*
4. *λ ist Nullstelle des charakteristischen Polynoms von f.*

BEWEIS. Äquivalent sind die Aussagen

λ ist Eigenwert von f	
$f(x) = \lambda x$ für ein $x \in V$, $x \neq 0$	
$(f - \lambda\mathrm{id})(x) = 0$ für ein $x \in V$, $x \neq 0$	
$\mathrm{Kern}(f - \lambda\mathrm{id}) \neq 0$	
$f - \lambda\mathrm{id}$ ist kein Monomorphismus	
$f - \lambda\mathrm{id}$ ist kein Epimorphismus	nach 3.2.3
$f - \lambda\mathrm{id}$ ist kein Isomorphismus	nach 3.2.3
$\det(f - \lambda\mathrm{id}) = 0$	siehe 6.3
λ ist Nullstelle des charakteristischen Polynoms von f.	QED

BEMERKUNGEN. Sei V ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear.

1. Die Eigenwerte von f lassen sich also als Nullstellen des charakteristischen Polynoms von f ausrechnen.

2. Sei λ Eigenwert von f. Den zugehörigen Eigenraum kann man wie folgt berechnen. Man fixiere eine Basis $x_1, \ldots, x_n$ von V. Sei $A = (\alpha_{ij})$ die darstellende Matrix von f bezüglich $x_1, \ldots, x_n$, also $f(x_j) = \sum_{i=1}^{n} \alpha_{ij} x_i$ für $1 \le j \le n$, oder in Matrizenschreibweise

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Dann sind folgende Aussagen äquivalent.

$$\begin{aligned}
x := \sum_{i=1}^{n} \xi_i x_i &\in \operatorname{Eig}(f, \lambda) \\
f(x) &= \lambda x \\
f(\begin{pmatrix} \xi_1 & \ldots & \xi_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) &= \lambda \begin{pmatrix} \xi_1 & \ldots & \xi_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
\begin{pmatrix} \xi_1 & \ldots & \xi_n \end{pmatrix} A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &= \lambda \begin{pmatrix} \xi_1 & \ldots & \xi_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
\begin{pmatrix} \xi_1 & \ldots & \xi_n \end{pmatrix} A^t &= \lambda \begin{pmatrix} \xi_1 & \ldots & \xi_n \end{pmatrix} \\
A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} &= \lambda E \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \\
(A - \lambda E) \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} &= 0.
\end{aligned}$$

Die Berechnung des zugehörigen Eigenraums läuft also auf die Lösung eines linearen Gleichungssystems hinaus.

3. Wir zeigen noch eine Eigenschaft des charakteristischen Polynoms, die uns später bei der Herstellung der JORDANschen Normalform nützlich sein wird. Seien $U, W \subseteq V$ Unterräume mit $V = U \oplus W$. Wir nehmen ferner an, daß U und W f-invariant sind, d.h. daß $f(U) \subseteq U$ und $f(W) \subseteq W$. (Man sagt in diesem Fall, daß der Endomorphismus f durch das Paar (U, W) *reduziert* ist; siehe [2, p.72]). Dann läßt sich das charakteristische Polynom von f zerlegen in das Produkt der charakteristischen Polynome der Einschränkungen von f auf die beiden Unterräume, also

$$p_f = p_{f|U} \cdot p_{f|W}.$$

BEWEIS. Man wähle Basen $x_1, \ldots, x_n$ von U und $y_1, \ldots, y_m$ von W. Dann gilt für geeignete

Matrizen $A \in K^{n\times n}$ und $B \in K^{m\times m}$

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Also

$$\begin{aligned} p_f &= \begin{vmatrix} A - tE & 0 \\ 0 & B - tE \end{vmatrix} \\ &= \det(A - tE)\det(B - tE) \quad \text{nach 6.2.3} \\ &= p_{f|U} \cdot p_{f|W}. \end{aligned}$$

Beispiel. Wir berechnen das charakteristische Polynom $\det(A - tE)$ für

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

Es ist

$$A - tE = \begin{pmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{pmatrix},$$

also durch Entwicklung nach der ersten Zeile

$$\begin{aligned} \begin{vmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{vmatrix} &= -t[(-2-t)(3-t)+6] - (-1)[(-3)(3-t)+6] + [6 - 2(2+t)] \\ &= -t[-6 - t + t^2 + 6] - (-1)[-9 + 3t + 6] + [2 - 2t] \\ &= -t^3 + t^2 + t - 1 \end{aligned}$$

Man sieht sofort, daß 1 eine Nullstelle ist. Es gibt also ein g mit $-t^3 + t^2 + t - 1 = (t-1)g$; durch Division findet man $g = -t^2 + 1$, d.h. $-t^3 + t^2 + t - 1 = -(t-1)^2(t+1)$. Es gibt also zwei Eigenwerte $+1$ und -1, und der Eigenwert $+1$ tritt doppelt auf.

7.2.4. *Algebraische und geometrische Vielfachheit.* Sei V ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. Jedem Eigenwert λ von f sind in natürlicher Weise zwei Zahlen zugeordnet.

1. Die Vielfachheit $\mu(p_f, \lambda)$ von λ als Nullstelle des charakteristischen Polynoms p_f von f; diese Zahl nennt man die *algebraische Vielfachheit* des Eigenwerts λ.

2. Die Dimension des zum Eigenwert λ gehörigen Eigenraums $\mathrm{Eig}(f, \lambda)$; diese Zahl nennt man die *geometrische Vielfachheit* des Eigenwerts λ.

Wir wollen uns überlegen, daß die geometrische Vielfachheit stets kleiner oder gleich der algebraischen Vielfachheit ist.

LEMMA. *Sei V ein endlichdimensionaler K-Vektorraum, $f\colon V \to V$ linear und λ Eigenwert von f. Dann gilt* $\dim(\mathrm{Eig}(f,\lambda)) \le \mu(p_f,\lambda)$.

BEWEIS. Sei $x_1,\dots,x_r$ eine beliebige Basis von $\mathrm{Eig}(f,\lambda)$. Man ergänze sie zu einer Basis $x_1,\dots,x_r,x_{r+1},\dots,x_n$ von V. Dann gilt $f(x_i)=\lambda x_i$ für $i=1,\dots,r$, also

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda & & & & & \\ & \ddots & & & & \\ & & \lambda & & & \\ * & \dots & * & * & \dots & * \\ \vdots & & & & & \\ * & \dots & * & * & \dots & * \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix}.$$

Also ist $\dim(\mathrm{Eig}(f,\lambda)) \le \mu(p_f,\lambda)$. QED

BEMERKUNG. In dem Fall, daß das charakteristische Polynom von f in Linearfaktoren zerfällt (und das ist etwa über dem Körper $\mathbb{C}$ der komplexen Zahlen stets so), werden wir in 7.4 zeigen können, daß f genau dann diagonalisierbar ist, wenn für jeden Eigenwert λ von f die geometrische Vielfachheit gleich der algebraischen Vielfachheit ist, also wenn gilt $\dim(\mathrm{Eig}(f,\lambda)) = \mu(p_f,\lambda)$. Wir wollen uns jetzt mit dem allgemeinen Fall befassen, daß nämlich die Dimension eines Eigenraums $\mathrm{Eig}(f,\lambda)$ eventuell zu klein ist. Es ist $\mathrm{Eig}(f,\lambda) = \mathrm{Kern}(g)$ mit $g := f - \lambda\mathrm{id}$. Die Idee hier ist, die eventuell zu kleine Dimension zu vergrößern und statt des Eigenraums $\mathrm{Kern}(g)$ den sogenannten Hauptraum $U_{f,\lambda} := \bigcup_{m=0}^{\infty} \mathrm{Kern}(g^m)$ zu betrachten. Dies wird uns auf die JORDANsche Normalform führen.

7.3 Jordansche Normalform

In diesem Abschnitt sei K stets ein Körper. Für jedes $\lambda \in K$ und $s \ge 1$ sei die JORDAN-*Matrix* $J_{s,\lambda}$ definiert durch

$$J_{s,\lambda} := \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix} \in K^{s\times s}$$

und insbesondere $J_{1,\lambda} := (\lambda)$.

7.3.1. SATZ. *(*JORDAN*sche Normalform). Gegeben sei ein n-dimensionaler K-Vektorraum V und $f\colon V \to V$ linear. Ferner sei p_f Produkt linearer Faktoren. Dann existiert eine Basis $x_1,\dots,x_n$ von V mit*

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

wobei jedes A_i eine JORDAN-*Matrix $J_{s,\lambda}$ ist zu einem Eigenwert λ von f.*

BEMERKUNGEN. 1. Im Satz können eventuell mehrere JORDAN-Matrizen A_i zu einem Eigenwert λ gehören.

2. Die A_i sind bis auf die Reihenfolge eindeutig bestimmt, und zwar durch f, also unabhängig von der Basis; dies wird hier nicht bewiesen.

BEWEIS. Sei etwa $p_f = (-1)^n \prod_{i=1}^m (t-\lambda_i)^{e_i}$ mit verschiedenen λ_i und $e_i \geq 1$. Wir führen den Beweis durch Induktion über die Anzahl m der verschiedenen Eigenwerte von f.

Schritt 1: Zerlegung von V. Betrachte $\lambda := \lambda_1$. Setze $g := f - \lambda\mathrm{id}$. Dann ist $\mathrm{Kern}(g)$ der Eigenraum von f bezüglich λ. Wir untersuchen den sogenannten *Hauptraum* von f bezüglich λ, definiert als

$$U := U_{f,\lambda} := \bigcup_{m=0}^{\infty} \mathrm{Kern}(g^m);$$

hierbei ist wie üblich $g^m := g \circ g \circ \cdots \circ g$ mit m Vorkommen von g.

Wir zeigen zunächst

$$U \text{ ist ein } f\text{-invarianter Unterraum, d.h. } U \text{ ist Unterraum mit } f(U) \subseteq U. \tag{7.1}$$

Zum Beweis von (7.1) stellen wir zunächst fest, daß U ein Unterraum ist. Es genügt zu zeigen, daß aus $x, y \in U$ stets folgt $x + y \in U$. Ist nämlich $x \in \mathrm{Kern}(g^m)$ und $y \in \mathrm{Kern}(g^k)$ mit etwa $m \leq k$, so folgt $x \in \mathrm{Kern}(g^k)$ und damit $x + y \in \mathrm{Kern}(g^k) \subseteq U$. $f(U) \subseteq U$ ergibt sich wie folgt. Offenbar genügt $g(U) \subseteq U$, da $f = g + \lambda\mathrm{id}$. Sei also $x \in U$, etwa $g^m(x) = 0$. Es genügt zu zeigen $g(x) \in U$. Es ist $0 = g(g^m(x)) = g^m(g(x))$, also $g(x) \in \mathrm{Kern}(g^m) \subseteq U$.

Wir setzen jetzt

$$W := \bigcap_{m=0}^{\infty} \mathrm{Bild}(g^m)$$

und zeigen

$$W \text{ ist ein } f\text{-invarianter Unterraum, d.h. } W \text{ ist Unterraum mit } f(W) \subseteq W. \tag{7.2}$$

Zum Beweis von (7.2) beachte man, daß W als Durchschnitt von Unterräumen wieder ein Unterraum ist. $f(W) \subseteq W$ ergibt sich wie folgt. Offenbar genügt wieder $g(W) \subseteq W$, da $f = g + \lambda\mathrm{id}$. Sei also $y \in W$, d.h. $y = g^m(x_m)$ für alle m, mit geeigneten $x_m \in V$. Es genügt zu zeigen $g(y) \in W$. Man erhält $g(y) = g^{m+1}(x_m) \in \mathrm{Bild}(g^{m+1}) \subseteq \mathrm{Bild}(g^m)$ für alle m.

Unser nächstes Ziel ist $V = U \oplus W$. Dazu werden wir von der Voraussetzung $\dim(V) < \infty$ Gebrauch machen. Wir zeigen zunächst, daß es ein $k \leq \dim(V)$ gibt mit $1 \leq k$ und

$$\begin{aligned} &V \supsetneq \mathrm{Bild}(g) \supsetneq \mathrm{Bild}(g^2) \supsetneq \cdots \supsetneq \mathrm{Bild}(g^k) = \mathrm{Bild}(g^{k+1}) = \ldots, \text{ also } \mathrm{Bild}(g^k) = W. \\ &0 \subsetneq \mathrm{Kern}(g) \subsetneq \mathrm{Kern}(g^2) \subsetneq \cdots \subsetneq \mathrm{Kern}(g^k) = \mathrm{Kern}(g^{k+1}) = \ldots, \text{ also } \mathrm{Kern}(g^k) = U. \end{aligned} \tag{7.3}$$

Zum Beweis von (7.3) beachte man zunächst, daß es aus Dimensionsgründen stets ein k mit $\mathrm{Bild}(g^k) = \mathrm{Bild}(g^{k+1})$ geben muß. Man wähle k minimal mit $\mathrm{Bild}(g^k) = \mathrm{Bild}(g^{k+1})$. Dann gilt

$$\begin{aligned} g^k(V) &= g^{k+1}(V) \\ g(g^k(V)) &= g(g^{k+1}(V)) \\ g^{k+1}(V) &= g^{k+2}(V) \qquad \text{usw.} \end{aligned}$$

Die zweite Aussage in (7.3) ergibt sich daraus wie folgt. Offenbar ist $0 \subseteq \mathrm{Kern}(g) \subseteq \mathrm{Kern}(g^2) \subseteq \ldots$. Wegen $\dim(\mathrm{Kern}(g^m)) + \dim(\mathrm{Bild}(g^m)) = n$ (nach 3.2.3) folgt die Behauptung aus der ersten Aussage. $1 \leq k$ ergibt sich aus $\mathrm{Kern}(g) \neq 0$ (denn $\mathrm{Kern}(g)$ ist Eigenraum zu λ).

Jetzt erhält man $V = U \oplus W$ wie folgt. $\mathrm{Kern}(g^k) \cap \mathrm{Bild}(g^k) = 0$: Sei $x \in \mathrm{Kern}(g^k) \cap \mathrm{Bild}(g^k)$. Es ist $g^k\colon \mathrm{Bild}(g^k) \to \mathrm{Bild}(g^{2k})$ wohldefiniert und surjektiv, also ein Isomorphismus. Wegen $x \in \mathrm{Bild}(g^k)$ und $g^k(x) = 0$ folgt $x = 0$. $\mathrm{Kern}(g^k) + \mathrm{Bild}(g^k) = V$ folgt aus

$$\begin{aligned}\dim(\mathrm{Kern}(g^k) + \mathrm{Bild}(g^k)) &= \dim(\mathrm{Kern}(g^k)) + \dim(\mathrm{Bild}(g^k)) \quad \text{dritte Folgerung in 3.2.3}\\ &= n\end{aligned}$$

Schritt 2: Konstruktion einer Basis von U, bezüglich derer sich g (und damit auch $f = g + \lambda\mathrm{id}$) übersichtlich verhält. Dazu ein

LEMMA. *Sei*

$$\mathrm{Kern}(g^i) = \mathrm{Kern}(g^{i-1}) \oplus \langle y_1, \ldots, y_m \rangle$$

mit $y_1, \ldots, y_m$ *linear unabhängig. Dann existieren* $z_1, \ldots, z_q \in \mathrm{Kern}(g^{i-1})$ *(mit* $q \geq 0$*) so daß*

$$\mathrm{Kern}(g^{i-1}) = \mathrm{Kern}(g^{i-2}) \oplus \langle g(y_1), \ldots, g(y_m), z_1, \ldots, z_q \rangle.$$

BEWEIS. a. Es ist $g(y_j) \in \mathrm{Kern}(g^{i-1})$, da $g^{i-1}(g(y_j)) = g^i(y_j) = 0$ wegen $y_j \in \mathrm{Kern}(g^i)$.

b. Wir zeigen gleichzeitig

$$\begin{aligned}&\mathrm{Kern}(g^{i-2}) \cap \langle g(y_1), \ldots, g(y_m) \rangle = 0\\ &g(y_1), \ldots, g(y_m) \text{ sind linear unabhängig.}\end{aligned}$$

Zum Beweis genügt es zu zeigen, daß aus $\sum_{j=1}^m \alpha_j g(y_j) \in \mathrm{Kern}(g^{i-2})$ folgt $\alpha_1 = \cdots = \alpha_m = 0$. Sei also $\sum_{j=1}^m \alpha_j g(y_j) \in \mathrm{Kern}(g^{i-2})$. Man erhält

$$\begin{aligned}&0 = g^{i-2}\Big(\sum_{j=1}^m \alpha_j g(y_j)\Big) = g^{i-1}\Big(\sum_{j=1}^m \alpha_j y_j\Big)\\ &\sum_{j=1}^m \alpha_j y_j \in \mathrm{Kern}(g^{i-1})\\ &\sum_{j=1}^m \alpha_j y_j = 0 && \text{da } \mathrm{Kern}(g^i) = \mathrm{Kern}(g^{i-1}) \oplus \langle y_1, \ldots, y_m \rangle\\ &\alpha_1 = \cdots = \alpha_m = 0 && \text{da } y_1, \ldots, y_m \text{ linear unabhängig sind.}\end{aligned}$$

c. Jetzt läßt sich das Lemma leicht beweisen. Man wähle eine Basis $u_1, \ldots, u_p$ von $\mathrm{Kern}(g^{i-2})$, und ergänze $u_1, \ldots, u_p, g(y_1), \ldots, g(y_m)$ zu einer Basis von ganz $\mathrm{Kern}(g^{i-1})$, etwa

$$u_1, \ldots, u_p, g(y_1), \ldots, g(y_m), z_1, \ldots, z_q.$$

Damit ist das Lemma bewiesen. – Man erhält aus dem Lemma

$$\begin{aligned}\mathrm{Kern}(g^k) &= \mathrm{Kern}(g^{k-1}) \oplus \langle x_1, \ldots, x_{m_1} \rangle\\ \mathrm{Kern}(g^{k-1}) &= \mathrm{Kern}(g^{k-2}) \oplus \langle g(x_1), \ldots, g(x_{m_1}), x_{m_1+1}, \ldots, x_{m_2} \rangle\\ \mathrm{Kern}(g^{k-2}) &= \mathrm{Kern}(g^{k-3}) \oplus \langle g^2(x_1), \ldots, g^2(x_{m_1}), g(x_{m_1+1}), \ldots, g(x_{m_2}), x_{m_2+1}, \ldots, x_{m_3} \rangle\end{aligned}$$

und schließlich

$$\mathrm{Kern}(g) = 0 \oplus \langle g^{k-1}(x_1), \ldots, g^{k-1}(x_{m_1}), g^{k-2}(x_{m_1+1}), \ldots, g^{k-2}(x_{m_2}), \ldots, x_{m_{k-1}+1}, \ldots, x_{m_k} \rangle$$

mit $1 \le m_1 \le m_2 \le \cdots \le m_k$. Einsetzen ergibt folgende Basis von $\mathrm{Kern}(g^k) = U$.

$$\begin{array}{ll} x_{i_1}, g(x_{i_1}), \ldots, g^{k-1}(x_{i_1}) & \text{für } 1 \le i_1 \le m_1 \\ x_{i_2}, g(x_{i_2}), \ldots, g^{k-2}(x_{i_2}) & \text{für } m_1 + 1 \le i_2 \le m_2 \\ \vdots & \\ x_{i_k} & \text{für } m_{k-1} + 1 \le i_k \le m_k. \end{array} \tag{7.4}$$

Außerdem ist $x_{i_1} \in \mathrm{Kern}(g^k)$, $x_{i_2} \in \mathrm{Kern}(g^{k-1})$ und schließlich $x_{i_k} \in \mathrm{Kern}(g)$, d.h.

$$\begin{aligned} g^k(x_{i_1}) &= 0 \\ g^{k-1}(x_{i_2}) &= 0 \\ &\vdots \\ g(x_{i_k}) &= 0 \end{aligned}$$

Jede Zeile von (7.4) hat also die Form

$$y, g(y), \ldots, g^{p-1}(y) \quad \text{mit } g^p(y) = 0 \text{ und } p \le k.$$

Aus $g^p(y) = 0$ folgt

$$g \begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix} = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 0 & & & & 0 \end{pmatrix} \begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix}$$

und wegen $f = g + \lambda \mathrm{id}$

$$f \begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix} = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix} \begin{pmatrix} y \\ g(y) \\ \vdots \\ g^{p-1}(y) \end{pmatrix}.$$

Insgesamt ergibt sich für die in (7.4) gefundene Basis $z_1, \ldots, z_d$ von U

$$f \begin{pmatrix} z_1 \\ \vdots \\ z_d \end{pmatrix} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_d \end{pmatrix},$$

wobei jedes A_i eine JORDAN-Matrix $J_{p,\lambda}$ ist.

Schritt 3. Abschluß. Wir wissen bereits $V = U \oplus W$, $f(U) \subseteq U$ und $f(W) \subseteq W$. Aus der eben gefundenen darstellenden Matrix für $f{\restriction}U$ ergibt sich

$$p_{f{\restriction}U} = \pm(t - \lambda_1)^d$$

mit $d = \dim(U)$. Nach Bemerkung 3 in 7.2.3 gilt

$$p_f = p_{f\restriction U} \cdot p_{f\restriction W}.$$

Ist $W = 0$, so sind wir fertig (dann ist $p_{f\restriction W} = 1$). Sei also W nicht der Nullraum. Dann ist $p_{f\restriction W}(\lambda_1) \neq 0$ (denn sonst gäbe es ein $x \in W$, $x \neq 0$ mit $f(x) = \lambda_1 x$, also $x \in \mathrm{Kern}(g)$ im Widerspruch zu $V = U \oplus W$ und $\mathrm{Kern}(g) \subseteq U$). Daraus ergibt sich, daß in

$$p_f = \pm(t-\lambda_1)^{e_1}\prod_{i=2}^{m}(t-\lambda_i)^{e_i} = \pm(t-\lambda_1)^d p_{f\restriction W}$$

$e_1 = d$ sein muß (denn kürzt man gemeinsame Faktoren $t - \lambda_1$ heraus, so kann rechts und links nichts übrig bleiben). Also ist

$$p_{f\restriction W} = \pm\prod_{i=2}^{m}(t-\lambda_i)^{e_i}.$$

$p_{f\restriction W}$ hat einen Eigenwert weniger als p_f. Nach Induktionsvoraussetzung gibt es also für $f\restriction W$ eine Basis mit der gewünschten Eigenschaft. Für $f\restriction U$ hatten wir dies in Schritt 2 bewiesen. Damit ergibt sich der Satz. QED

BEMERKUNGEN. 1. Die Bedingung 'p_f ist Produkt linearer Faktoren' im Satz ist notwendig.

BEWEIS. Sei

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix}$$

mit JORDAN-Matrizen A_i darstellende Matrix von f bezüglich einer Basis. Dann gilt

$$p_f = p_{A_1}p_{A_2}\cdots p_{A_r},$$
$$p_{A_i} = \pm(t-\lambda_i)^{n_i}.$$

2. In der Algebra zeigt man: Zu jedem $f \in K[t]$ existiert ein Erweiterungskörper $L \supseteq K$, in dem f in Linearfaktoren zerfällt.

3. In 'algebraisch abgeschlossenen Körpern' zerfällt jedes Polynom in Linearfaktoren. Z.B. ist $\mathbb{C}$ algebraisch abgeschlossen (Fundamentalsatz der Algebra; Beweis in der Funktionentheorie).

7.3.2. *Umformulierung für Matrizen.* Aus den Überlegungen des vorangehenden Abschnitts ergibt sich eine nützliche Fassung der JORDANschen Normalform für Matrizen.

SATZ. *(*JORDAN*sche Normalform für Matrizen). Sei K ein Körper und $A \in K^{n\times n}$. Ferner sei p_A Produkt linearer Faktoren. Dann findet man ein $S \in \mathrm{GL}(n,K)$ so daß*

$$SAS^{-1} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix},$$

wobei jedes A_i eine JORDAN*-Matrix $J_{s,\lambda}$ ist zu einem Eigenwert λ von f_A.*

BEWEIS. Wir betrachten noch einmal den Beweis zum Satz über die JORDANsche Normalform, und zwar für $V = K^n$ und $f = f_B$, $B := A^t$. In (7.4) hatten wir für $g = f - \lambda\mathrm{id}$ die folgende

Basis von $\mathrm{Kern}(g^k) = U$ erhalten.

$$\begin{array}{ll} x_{i_1}, g(x_{i_1}), \dots, g^{k-1}(x_{i_1}) & \text{für } 1 \le i_1 \le m_1 \\ x_{i_2}, g(x_{i_2}), \dots, g^{k-2}(x_{i_2}) & \text{für } m_1 + 1 \le i_2 \le m_2 \\ \vdots & \\ x_{i_k} & \text{für } m_{k-1} + 1 \le i_k \le m_k. \end{array}$$

Zur Vereinfachung nehmen wir $1 = m_1 = m_2 = \cdots = m_k$ an und schreiben x für x_{i_1}. Sei T die Matrix mit den Spalten $x, g(x), \dots, g^{k-1}(x)$. Dann gilt für $1 \le i \le k$

$$\begin{aligned} Bg^{i-1}(x) &= f_B(g^{i-1}(x)) = (\lambda \mathrm{id} + g)g^{i-1}(x) = \lambda g^{i-1}(x) + g^i(x) \\ &= i\text{-te Spalte von } T \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix} \end{aligned}$$

$$BT = T \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix}$$

$$T^{-1}BT = \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix},$$

also

$$T^t A (T^t)^{-1} = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}. \quad \text{QED}$$

BEISPIEL. Sei

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}, \text{ also } B = A^t = \begin{pmatrix} 3 & -1 & 1 \\ 4 & 0 & 2 \\ 3 & -1 & 3 \end{pmatrix}.$$

Dann ist

$$\begin{aligned} p_B = \begin{vmatrix} 3-t & -1 & 1 \\ 4 & -t & 2 \\ 3 & -1 & 3-t \end{vmatrix} &= (3-t)[-3t + t^2 + 2] + \underbrace{[12 - 4t - 6] + [-4 + 3t]}_{2-t} \\ &= (3-t)(t-1)(t-2) - (t-2) \\ &= (t-2)[-t^2 + 4t - 4] \\ &= -(t-2)^3. \end{aligned}$$

Sei $g = f_B - 2\mathrm{id}$. Dann ist

$$0 \subsetneq \mathrm{Kern}(g) \subsetneq \mathrm{Kern}(g^2) \subsetneq \mathrm{Kern}(g^3) = \mathbb{R}^3.$$

Wir folgen jetzt dem Beweis des Satzes und konstruieren eine Basis von $\mathbb{R}^3$, bezüglich derer sich g übersichtlich verhält. Die Zahl k im Satz ist hier 3. Wir müssen Basen der Räume $\mathrm{Kern}(g)$ und $\mathrm{Kern}(g^2)$ berechnen. Es ist $g = f_C$ mit

$$C = \begin{pmatrix} 1 & -1 & 1 \\ 4 & -2 & 2 \\ 3 & -1 & 1 \end{pmatrix}, \quad \text{also} \quad C^2 = \begin{pmatrix} 0 & 0 & 0 \\ 2 & -2 & 2 \\ 2 & -2 & 2 \end{pmatrix}.$$

$\mathrm{Kern}(g)$ ergibt sich als Lösungsraum des linearen Gleichungssystems $Cx = 0$:

$$\begin{cases} \xi_1 - \xi_2 + \xi_3 = 0 \\ 4\xi_1 - 2\xi_2 + 2\xi_3 = 0 \end{cases}$$

$$\begin{cases} \xi_1 - \xi_2 + \xi_3 = 0 \\ 2\xi_2 - 2\xi_3 = 0 \end{cases}$$

$$\begin{cases} \xi_1 - \xi_2 + \xi_3 = 0 \\ \xi_2 - \xi_3 = 0 \end{cases}$$

Also ist

$$\mathrm{Kern}(g) = \langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \rangle.$$

Ferner ist

$$\mathrm{Kern}(g^2) = \{ \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mid \xi_1 - \xi_2 + \xi_3 = 0 \} = \langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \rangle.$$

Man erhält

$$\mathbb{R}^3 = \mathrm{Kern}(g^3) = \underbrace{\mathrm{Kern}(g^2)}_{\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \rangle} \oplus \langle \underbrace{x_1}_{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} \rangle, \quad \text{also } m_1 = 1$$

$$\mathrm{Kern}(g^2) = \underbrace{\mathrm{Kern}(g)}_{\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \rangle} \oplus \langle \underbrace{g(x_1)}_{\begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}} \rangle, \quad \text{also } m_2 = m_1$$

$$\mathrm{Kern}(g) = 0 \oplus \langle \underbrace{g^2(x_1)}_{\begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}} \rangle, \quad \text{also } m_3 = m_2.$$

Einsetzung ergibt die folgende Basis von $\mathrm{Kern}(g^3) = \mathbb{R}^3$:

$$x_1, g(x_1), g^2(x_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

Sei T die Matrix mit den Spalten $x_1, g(x_1), g^2(x_1)$, also

$$T = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 4 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Dann ist, wie wir aus der Theorie wissen,

$$T^t A (T^t)^{-1} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Man kann dies (etwa mit MAPLE) leicht kontrollieren.

BEISPIEL. Sei

$$A = \begin{pmatrix} 2 & & & & \\ -1 & 2 & & & \\ -1 & -1 & 2 & & \\ -1 & -1 & -1 & 2 & \\ -1 & -1 & -1 & -1 & 2 \end{pmatrix}, \text{ also } B = \begin{pmatrix} 2 & -1 & -1 & -1 & -1 \\ & 2 & -1 & -1 & -1 \\ & & 2 & -1 & -1 \\ & & & 2 & -1 \\ & & & & 2 \end{pmatrix}$$

und damit $p_B = (2-t)^5$. Sei wieder $g = f_B - 2\mathrm{id}$. Dann ist

$$0 \subsetneq \mathrm{Kern}(g) \subsetneq \mathrm{Kern}(g^2) \subsetneq \mathrm{Kern}(g^3) \subsetneq \mathrm{Kern}(g^4) \subsetneq \mathrm{Kern}(g^5) = \mathbb{R}^5.$$

Es ist $g = f_C$ mit

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

also

$$C^2 = \begin{pmatrix} 0 & 0 & 1 & 2 & 3 \\ & 0 & 0 & 1 & 2 \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix}, C^3 = \begin{pmatrix} 0 & 0 & 0 & -1 & -3 \\ & 0 & 0 & 0 & -1 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix} \text{ und } C^4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 \\ & & & & 0 \end{pmatrix}.$$

Man erhält

$$\mathbb{R}^5 = \mathrm{Kern}(g^5) = \underbrace{\mathrm{Kern}(g^4)}_{\langle e_1, e_2, e_3, e_4 \rangle} \oplus \langle \underbrace{x_1}_{e_5} \rangle, \qquad \text{also } m_1 = 1$$

$$\mathrm{Kern}(g^4) = \underbrace{\mathrm{Kern}(g^3)}_{\langle e_1, e_2, e_3 \rangle} \oplus \langle \underbrace{g(x_1)}_{\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ 0 \end{pmatrix}} \rangle, \qquad \text{also } m_2 = m_1$$

$$\operatorname{Kern}(g^3) = \underbrace{\operatorname{Kern}(g^2)}_{\langle e_1, e_2\rangle} \oplus \langle \underbrace{g^2(x_1)}_{\begin{pmatrix}3\\2\\1\\0\\0\end{pmatrix}} \rangle, \quad \text{also } m_3 = m_2$$

$$\operatorname{Kern}(g^2) = \underbrace{\operatorname{Kern}(g)}_{\langle e_1\rangle} \oplus \langle \underbrace{g^3(x_1)}_{\begin{pmatrix}-3\\-1\\0\\0\\0\end{pmatrix}} \rangle, \quad \text{also } m_4 = m_3$$

$$\operatorname{Kern}(g) = 0 \oplus \langle \underbrace{g^4(x_1)}_{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}} \rangle, \quad \text{also } m_5 = m_4.$$

Einsetzung ergibt die folgende Basis von $\operatorname{Kern}(g^5) = \mathbb{R}^5$:

$$x_1, g(x_1), g^2(x_1), g^3(x_1), g^4(x_1) = \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}-1\\-1\\-1\\-1\\0\end{pmatrix}, \begin{pmatrix}3\\2\\1\\0\\0\end{pmatrix}, \begin{pmatrix}-3\\-1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}.$$

Sei T die Matrix mit diesen Spalten, also

$$T^t = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \\ 3 & 2 & 1 & 0 & 0 \\ -3 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Dann ist wieder, wie wir aus der Theorie wissen, $T^t A (T^t)^{-1}$ die JORDAN-Matrix $J_{5,2}$. Man kann dies mit MAPLE leicht kontrollieren. Die Eingaben lauten:

```
with(linalg):
T:=matrix([[0,-1,3,-3,1],[0,-1,2,-1,0],[0,-1,1,0,0],[0,-1,0,0,0],[1,0,0,0,0]]):
A:=matrix([[2,0,0,0,0],[-1,2,0,0,0],[-1,-1,2,0,0],[-1,-1,-1,2,0],[-1,-1,-1,-1,2]]):
evalm(transpose(T)*A*inverse(transpose(T)));
```

und die Ausgabe ist

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

7.4 Diagonalisierung

Sei V ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. f hatten wir *diagonalisierbar* genannt, wenn es eine Basis $x_1, \ldots, x_n$ von V und $\lambda_1, \ldots, \lambda_n \in K$ gibt so daß

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Wir beweisen ein einfaches Kriterium für die Diagonalisierbarkeit eines Endomorphismus.

7.4.1. SATZ. *Sei V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Ferner sei das charakteristische Polynom p_f Produkt linearer Faktoren, etwa $p_f = (-1)^n \prod_{i=1}^m (t - \lambda_i)^{e_i}$ mit verschiedenen λ_i und $e_i \geq 1$. f ist genau dann diagonalisierbar, wenn für alle Eigenwerte λ_i die geometrische Vielfachheit $\dim(\mathrm{Eig}(f, \lambda_i))$ gleich der algebraischen Vielfachheit e_i von λ_i ist.*

BEWEIS. 1. Sei f diagonalisierbar, also $V = V_1 \oplus \cdots \oplus V_m$ mit $V_i = \mathrm{Eig}(f, \lambda_i)$. Sei $(x_{ij})_{1 \leq j \leq e_i}$ Basis von V_i. Dann gilt

$$f\begin{pmatrix} x_{11} \\ \vdots \\ x_{1e_1} \\ \vdots \\ x_{n1} \\ \vdots \\ x_{ne_n} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_n & & \\ & & & & & \ddots & \\ & & & & & & \lambda_n \end{pmatrix}\begin{pmatrix} x_{11} \\ \vdots \\ x_{1e_1} \\ \vdots \\ x_{n1} \\ \vdots \\ x_{ne_n} \end{pmatrix} \tag{7.5}$$

mit e_i Vorkommen von λ_i. Man erhält für das charakteristische Polynom

$$p_f = \det(A - tE) = \prod_{i=1}^{n} (\lambda_i - t)^{e_i}$$

Also ist e_i sowohl die Vielfachheit von λ_i als auch die Dimension von V_i.

2. Für jeden Eigenwert λ_i von f gelte $\dim(V_i) = e_i$ mit $V_i = \mathrm{Eig}(f, \lambda_i)$. Sei $(x_{ij})_{1 \leq j \leq e_i}$ Basis von V_i. Wegen $\sum_{i=1}^n e_i = n$ ist dann $(x_{ij})_{1 \leq i \leq m, 1 \leq j \leq e_i}$ eine Basis von V, und wir haben wieder (7.5). QED

BEMERKUNG. Man kann auch wie folgt aus dem obigen Beweis von Satz 7.3.1 der JORDANschen Normalform entnehmen, daß f diagonalisierbar ist, falls für jeden Eigenwert λ_i von f gilt $\dim(\mathrm{Eig}(f, \lambda_i)) = e_i$. Setze $g_i := f - \lambda_i \mathrm{id}$. Führt man die im Beweis beschriebene Konstruktion für λ_i durch, so ergibt sich für den Hauptraum $U = U_{f,\lambda_i} = \mathrm{Kern}(g_i^k)$, daß $\dim(U) = e_i$ ist. Da auch $\dim(\mathrm{Eig}(f, \lambda)) = e_i$ ist für den Eigenraum $\mathrm{Eig}(f, \lambda_i) = \mathrm{Kern}(g_i)$, haben wir $U = \mathrm{Eig}(f, \lambda_i)$ und $k = 1$. Alle JORDAN-Matrizen in der durchgeführten Konstruktion für den Eigenwert λ_i haben also die Form (λ_i). Da dies für alle λ_i gilt, ist f diagonalisierbar.

7.4.2. Wir wollen wieder eine Fassung der gerade durchgeführten Überlegungen für Matrizen angeben.

SATZ. *(Diagonalisierung von Matrizen). Sei K ein Körper und $A \in K^{n\times n}$. Ferner sei das charakteristische Polynom p_A Produkt linearer Faktoren, und für jeden Eigenwert sei die geometrische Vielfachheit gleich der algebraischen Vielfachheit. Dann findet man ein $S \in \mathrm{GL}(n,K)$ so daß*

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

BEWEIS. Sei $B := A^t$. Aufgrund der Voraussetzung gibt es eine Basis $x_1,\ldots,x_n$ des K^n aus Eigenvektoren von f_B. Sei x_i Eigenvektor zum Eigenwert λ_i, und T die Matrix mit den Spalten $x_1,\ldots,x_n$. Dann gilt

$$\begin{aligned} Bx_i &= f_B(x_i) = \lambda_i x_i = i\text{-te Spalte von } T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \\ BT &= T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \\ T^{-1}BT &= \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \\ T^tA(T^t)^{-1} &= \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \end{aligned}$$

FOLGERUNG. *(Praktisches Verfahren zur Diagonalisierung). Sei $A \in K^{n\times n}$ und $B := A^t$.*

1. *Man bestimme p_B und versuche, p_B in Linearfaktoren zu zerlegen. Wenn dies unmöglich ist, so ist f_B nicht diagonalisierbar. Andernfalls:*
2. *Für jeden Eigenwert λ bestimme man $\dim(\mathrm{Eig}(f_B,\lambda))$. Wenn für einen Eigenwert λ gilt $\dim(\mathrm{Eig}(f_B,\lambda)) < \mu(p_B,\lambda)$, so ist f_B nicht diagonalisierbar. Andernfalls:*
3. *Für jeden Eigenwert λ bestimme man eine Basis von $\mathrm{Eig}(f_B,\lambda)$ (durch Lösen des linearen Gleichungssystems $(B-\lambda E)x = 0$). Dann ist A diagonalisierbar mit Transformationsmatrix gemäß dem Beweis des Satzes.*

BEISPIEL. Sei

$$A = \begin{pmatrix} 0 & -3 & -2 \\ -1 & -2 & -2 \\ 1 & 3 & 3 \end{pmatrix}, \text{ also } B = A^t = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

In Abschnitt 7.2 hatten wir das charakteristische Polynom $p_A = p_B = -(t-1)^2(t+1)$ bestimmt. Zu den Eigenwerten $+1$ und -1 berechnen wir jetzt die Eigenvektoren.

Für den Eigenwert -1 ergeben sich die Eigenvektoren als Lösungen des linearen Gleichungssystems

$$\begin{pmatrix} 0+1 & -1 & 1 \\ -3 & -2+1 & 3 \\ -2 & -2 & 3+1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = 0.$$

Wir bringen diese Matrix durch elementare Zeilenumformungen in die Form einer Zeilenstufenmatrix:

$$\begin{pmatrix}1 & -1 & 1\\ -3 & -1 & 3\\ -2 & -2 & 4\end{pmatrix} \mapsto_{\substack{\text{zweite Zeile } +3\cdot \text{ erste Zeile}\\ \text{dritte Zeile } +2\cdot \text{ erste Zeile}}} \begin{pmatrix}1 & -1 & 1\\ 0 & -4 & 6\\ 0 & -4 & 6\end{pmatrix}$$
$$\mapsto \begin{pmatrix}1 & -1 & 1\\ 0 & -2 & 3\\ 0 & 0 & 0\end{pmatrix}.$$

Zu lösen ist also das Gleichungssystem

$$\begin{aligned}\xi_1 - \xi_2 + \xi_3 &= 0\\ -2\xi_2 + 3\xi_3 &= 0.\end{aligned}$$

Nach dem Verfahren aus Abschnitt 1.4 erhält man als Lösungsraum alle skalaren Vielfachen von $\begin{pmatrix}\frac{1}{2}\\ \frac{3}{2}\\ 1\end{pmatrix}$. Insbesondere ist also $\begin{pmatrix}1\\3\\2\end{pmatrix}$ Eigenvektor zum Eigenwert -1.

Für den Eigenwert 1 ergeben sich die Eigenvektoren als Lösungen des linearen Gleichungssystems

$$\begin{pmatrix}0-1 & -1 & 1\\ -3 & -2-1 & 3\\ -2 & -2 & 3-1\end{pmatrix}\begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix} = 0.$$

Wir bringen diese Matrix durch elementare Zeilenumformungen in die Form einer Zeilenstufenmatrix:

$$\begin{pmatrix}-1 & -1 & 1\\ -3 & -3 & 3\\ -2 & -2 & 2\end{pmatrix} \mapsto \begin{pmatrix}-1 & -1 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Das zugehörige Gleichungssystem besteht nur aus der Gleichung $-\xi_1 - \xi_2 + \xi_3 = 0$. Nach dem Verfahren aus Abschnitt 1.4 hat der Lösungsraum (also $\mathrm{Eig}(f_B, 1)$) die Basis $\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}$.

Es ist $\dim(\mathrm{Eig}(f_B, -1)) = 1 = \mu(p_B, -1)$ und $\dim(\mathrm{Eig}(f_B, 1)) = 3-1 = 2 = \mu(p_B, 1)$. Daher ist f_B diagonalisierbar.

$\begin{pmatrix}1\\3\\2\end{pmatrix}, \begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}$ ist eine Basis des $\mathbb{R}^3$ aus Eigenvektoren von f_B. Wir setzen gemäß dem Beweis des Satzes

$$T := \begin{pmatrix}1 & -1 & 1\\ 3 & 1 & 0\\ 2 & 0 & 1\end{pmatrix}.$$

Dann gilt, wie man leicht nachrechnet,

$$T^t A (T^t)^{-1} = \begin{pmatrix}-1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Kapitel 8

Komplexifizierung und unitäre Vektorräume

Im Satz über die JORDAN-Normalform mußten wir voraussetzen, daß das charakteristische Polynom p_f des gegebenen Endomorphismus $f\colon V \to V$ in Linearfaktoren zerfällt. Dies ist stets für den Körper $\mathbb{C}$ der komplexen Zahlen der Fall, aber nicht für allgemeine Körper und insbesondere nicht für den Körper $\mathbb{R}$ der reellen Zahlen. Wir zeigen jetzt, wie man diesen 'Defekt' des Körpers der reellen Zahlen dadurch umgehen kann, daß man einen gegebenen $\mathbb{R}$-Vektorraum in kanonischer Weise zu einem $\mathbb{C}$-Vektorraum macht. Als Anwendung erhalten wir eine Variante des Satzes über die JORDAN-Normalform für $\mathbb{R}$-Vektorräume. Ferner führen wir eine ebenfalls kanonische Erweiterung der Theorie der euklidischen Vektorräume (die ja die reellen Zahlen als Grundkörper voraussetzten) auf den wichtigen Fall der komplexen Zahlen als Grundkörper durch. Man spricht dann von unitären Vektorräumen.

8.1 Komplexifizierung

8.1.1. *Komplexe Zahlen.* In 1.2.5 hatten wir den Körper $\mathbb{C}$ der komplexen Zahlen als Spezialfall einer Körpererweiterung $K(\sqrt{\xi})$ für $\xi = -1$ eingeführt. Mit einer geeigneten Identifikation hatten wir $\mathbb{R} \subseteq \mathbb{C}$. Offenbar kann man $\mathbb{C}$ als $\mathbb{R}$-Vektorraum mit Basis $1, i$ auffassen. Ferner gilt $(\alpha + i\beta)(\gamma + i\delta) = \alpha\gamma - \beta\delta + i(\alpha\delta + \beta\gamma)$.

SCHREIBWEISE. Sei $\xi = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$.

$$\begin{aligned} \Re(\xi) &:= \alpha && \textit{Realteil} \\ \Im(\xi) &:= \beta && \textit{Imaginärteil} \\ \overline{\xi} &:= \alpha - i\beta && \textit{konjugiert komplexe } \text{Zahl} \\ |\xi| &:= \sqrt{\alpha^2 + \beta^2} && \textit{Betrag.} \end{aligned}$$

Es gilt:

$$\begin{aligned} \Re(\xi) &= \frac{1}{2}(\xi + \overline{\xi}) \\ \Im(\xi) &= \frac{1}{2i}(\xi - \overline{\xi}) \\ |\xi|^2 &= \xi\overline{\xi} \\ \frac{1}{\xi} &= \frac{\overline{\xi}}{\xi\overline{\xi}} = \frac{1}{|\xi|^2}\overline{\xi} \end{aligned}$$

Geometrische Deutung (GAUSSsche Zahlenebene): $\xi = \alpha + i\beta$, $r = |\xi|$, $\cos\varphi = \frac{\alpha}{r}$, $\sin\varphi = \frac{\beta}{r}$. Also: $\xi = \alpha + i\beta = r(\cos\varphi + i\sin\varphi)$. Addition in $\mathbb{C}$: Vektoraddition in $\mathbb{R}^2$. Multiplikation in $\mathbb{C}$:

Beträge multiplizieren, Winkel addieren, denn

$$\begin{aligned} &r(\cos\varphi + i\sin\varphi)\cdot r'(\cos\varphi' + i\sin\varphi') \\ &= rr'(\underbrace{\cos\varphi\cos\varphi' - \sin\varphi\sin\varphi'}_{\cos(\varphi+\varphi')} + i(\underbrace{\cos\varphi\sin\varphi' + \sin\varphi\cos\varphi'}_{\sin(\varphi+\varphi')})). \end{aligned}$$

Formale Definition (Erinnerung): Es war $\mathbb{C} := \mathbb{R}\times\mathbb{R}$ als additive Gruppe definiert. Die Multiplikation hatten wir definiert durch

$$(\alpha,\beta)\cdot(\gamma,\delta) = (\alpha\gamma - \beta\delta, \alpha\delta + \beta\gamma)$$

für $\alpha,\beta,\gamma,\delta \in \mathbb{R}$.

BEMERKUNGEN. 1. Für $\xi,\xi' \in \mathbb{C}$ gilt:

$$\begin{aligned} &|\xi\xi'| = |\xi||\xi| \\ &|\xi + \xi'| \le |\xi| + |\xi'| \\ &|\xi| = 0 \quad \text{genau dann, wenn} \quad \xi = 0 \\ &\overline{\xi\xi'} = \overline{\xi}\,\overline{\xi'} \\ &\overline{\xi+\xi'} = \overline{\xi} + \overline{\xi'} \\ &\overline{\overline{\xi}} = \xi \\ &\overline{\xi} = \xi \quad \text{genau dann, wenn} \quad \xi \in \mathbb{R}. \end{aligned}$$

BEWEIS. Durch Ausrechnen. QED

2. Sei $\sigma\colon \mathbb{C}\to\mathbb{C}$ ein $\mathbb{R}$-*Algebra-Automorphismus*, d.h.

$$\begin{aligned} &\sigma(\alpha) = \alpha && \text{für alle } \alpha\in\mathbb{R} \\ &\sigma(\xi+\xi') = \sigma(\xi) + \sigma(\xi') && \text{für alle } \xi,\xi'\in\mathbb{C} \\ &\sigma(\xi\cdot\xi') = \sigma(\xi)\cdot\sigma(\xi') && \text{für alle } \xi,\xi'\in\mathbb{C} \\ &\sigma \text{ bijektiv.} \end{aligned}$$

(Die letzten drei Eigenschaften besagen, daß σ ein *Ringisomorphismus* ist). Dann ist σ die Identität, d.h. $\sigma(\xi) = \xi$ für alle $\xi\in\mathbb{C}$, oder σ ist die Konjugation, d.h. $\sigma(\xi) = \overline{\xi}$ für alle $\xi\in\mathbb{C}$.

BEWEIS. Sei $\xi = \alpha + i\beta$ mit $\alpha,\beta\in\mathbb{R}$. Dann ist

$$\sigma(\xi) = \underbrace{\sigma(\alpha)}_{\alpha} + \sigma(i)\underbrace{\sigma(\beta)}_{\beta} = \alpha + \sigma(i)\beta$$

Es genügt zu zeigen $\sigma(i) = i$ (dann ist σ die Identität) oder $\sigma(i) = -i$ (dann ist σ die Konjugation). Dies folgt aus:

$$\begin{aligned} &i^2 = -1 \\ &-1 = \sigma(-1) = \sigma(i^2) = \sigma(i)\cdot\sigma(i) \\ &\sigma(i)^2 + 1 = 0 \\ &(\sigma(i)+i)(\sigma(i)-i) = 0 \\ &\sigma(i) = i \quad \text{oder} \quad \sigma(i) = -i. \quad \text{QED} \end{aligned}$$

3. Fortsetzung der Konjugation auf den Polynomring $\mathbb{C}[t]$. Für

$$p = \sum_{j=0}^{n} \alpha_j t^j \in \mathbb{C}[t] \text{ mit } \alpha_j \in \mathbb{C} \text{ sei}$$

$$\overline{p} := \sum_{j=0}^{n} \overline{\alpha_j} t^j .$$

Wieder gilt für $p, q \in \mathbb{C}[t]$

$$\begin{aligned} &\overline{p \cdot q} = \overline{p} \cdot \overline{q}, \\ &\overline{p + q} = \overline{p} + \overline{q}, \\ &\overline{\overline{p}} = p, \\ &\overline{p} = p \quad \text{genau dann, wenn} \quad p \in \mathbb{R}[t]. \end{aligned}$$

Beweis. Durch Ausrechnen. Zum Beispiel für $p = \sum_{j=0}^{n} \alpha_j t^j$ und $q = \sum_{l=0}^{m} \beta_l t^l$ mit $\alpha_j, \beta_l \in \mathbb{C}$ ergibt sich

$$\begin{aligned} p \cdot q &= \sum_{k=0}^{n+m} \Big(\sum_{l+j=k} \alpha_j \beta_l \Big) t^k, \\ \overline{p \cdot q} &= \sum_{k=0}^{n+m} \underbrace{\overline{\Big(\sum_{l+j=k} \alpha_j \beta_l \Big)}}_{=\sum_{l+j=k} \overline{\alpha_j}\,\overline{\beta_l}} t^k = \overline{p} \cdot \overline{q}. \quad \text{QED} \end{aligned}$$

8.1.2. *Fundamentalsatz der Algebra.* Eine wesentliche Eigenschaft des Körpers $\mathbb{C}$ der komplexen Zahlen ist seine *algebraische Abgeschlossenheit*, d.h. daß jedes nichtkonstante Polynom in $\mathbb{C}[t]$ Produkt linearer Faktoren ist. Man nennt diese Aussage auch den Fundamentalsatz der Algebra. Einen Beweis werden wir im Rahmen dieser Vorlesung nicht geben können; er wird in Vorlesungen über Funktionentheorie geführt (eventuell auch in einem Proseminar im Anschluß an diese Vorlesung). Jedes Polynom $p \in \mathbb{C}[t]$ hat also eine Darstellung

$$p = \alpha \prod_{j=1}^{n} (t - \alpha_j) \quad \text{mit } \alpha, \alpha_j \in \mathbb{C}.$$

$\alpha_1, \ldots, \alpha_n$ sind die Nullstellen in $\mathbb{C}$ von p.

Folgerung. *Jedes nicht konstante Polynom $p \in \mathbb{R}[t]$ ist Produkt von linearen und quadratischen Faktoren über $\mathbb{R}$. Genauer gilt*

$$p = \alpha \prod_{j=1}^{r} (t - \alpha_j)^{e_j} \prod_{j=r+1}^{n} \Big(\underbrace{(t - \alpha_j)(t - \overline{\alpha_j})}_{\in \mathbb{R}[t] \text{ vom Grad } 2} \Big)^{e_j},$$

wobei $e_j \geq 1$, $\alpha \in \mathbb{R}$, $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$, $\alpha_{r+1}, \ldots, \alpha_n \in \mathbb{C} \setminus \mathbb{R}$, $\alpha_j \neq \alpha_l$ für $j \neq l$. ($\alpha_1, \ldots, \alpha_r$ sind die reellen, $\alpha_{r+1}, \overline{\alpha_{r+1}}, \ldots, \alpha_n, \overline{\alpha_n}$ die komplexen Nullstellen von p).

Beweis. Nach dem Fundamentalsatz der Algebra gibt es ein $\lambda \in \mathbb{C}$ mit

$$p = (t - \lambda)^e q \quad \text{mit } q \in \mathbb{C}[t],\ e \geq 1 \text{ und } q(\lambda) \neq 0.$$

Nach Voraussetzung ist

$$p = \overline{p} = (t - \overline{\lambda})^e \overline{q} \quad \text{mit } \overline{q}(\overline{\lambda}) = \overline{q(\lambda)} \neq 0, \text{ da } q(\lambda) \neq 0.$$

Damit ist gezeigt: ist λ Nullstelle von p mit Vielfachheit e, so ist $\overline{\lambda}$ Nullstelle von p ebenfalls mit Vielfachheit e. Außerdem gilt

$$(t - \alpha_j)(t - \overline{\alpha_j}) = t^2 - t\underbrace{(\alpha_j + \overline{\alpha_j})}_{2\cdot\Re(\alpha_j)} + \underbrace{\alpha_j\overline{\alpha_j}}_{|\alpha_j|^2} \in \mathbb{R}[t].$$

Fortsetzung dieses Verfahrens ergibt die Behauptung. QED

FOLGERUNG. *Sei $p \in \mathbb{R}[t]$, p nicht konstant und* $\deg(p)$ *ungerade. Dann hat p eine reelle Nullstelle.*

BEWEIS. Nach Voraussetzung ist $\deg(p) = \sum_{j=1}^{r} e_j + 2\sum_{j=r+1}^{n} e_j$ ungerade. Es folgt $r \geq 1$ und damit die Behauptung. QED

8.1.3. *Komplexifizierung.* Gegeben sei ein $\mathbb{R}$-Vektorraum V. Wir wollen aus V einen $\mathbb{C}$-Vektorraum $\tilde{V}$ konstruieren, seine sogenannte *Komplexifizierung*. Dazu setzen wir $\tilde{V} := V \times V$ als additive Gruppe. Die Skalarmultiplikation $\mathbb{C} \times \tilde{V} \to \tilde{V}$ sei definiert durch

$$(\alpha + i\beta)(v_1, v_2) := (\alpha v_1 - \beta v_2, \alpha v_2 + \beta v_1).$$

BEMERKUNGEN. 1. $\tilde{V}$ mit der angegebenen Struktur ist ein $\mathbb{C}$-Vektorraum.

BEWEIS. Durch Nachrechnen, wie für $V = \mathbb{R}$. QED

2. Die Abbildung

$$\begin{aligned} V &\to \tilde{V} \\ v &\mapsto (v, 0) \end{aligned}$$

ist ein $\mathbb{R}$-*Monomorphismus*, d.h. es gilt

$$\begin{aligned} v + v' &\mapsto (v + v', 0) = (v, 0) + (v', 0), \\ \alpha \cdot v &\mapsto (\alpha v, 0) = \alpha \cdot (v, 0) \quad \text{für } \alpha \in \mathbb{R}. \end{aligned}$$

Weiter gilt

$$i \cdot (v, 0) = (0, v),$$

also

$$(v_1, v_2) = (v_1, 0) + i(v_2, 0) \quad \text{in } \tilde{V}.$$

Wir identifizieren jetzt v mit $(v, 0)$; in diesem Sinn haben wir dann $V \subseteq \tilde{V}$. Jedes $x \in \tilde{V}$ läßt sich also eindeutig schreiben in der Form

$$x = v_1 + iv_2 \quad \text{mit } v_1, v_2 \in V,$$

und es gilt

$$(\alpha + i\beta)(v_1 + iv_2) = \alpha v_1 - \beta v_2 + i(\alpha v_2 + \beta v_1).$$

8.1.4. *Komplexifizierung eines Homomorphismus.* Seien V, W $\mathbb{R}$-Vektorräume und $f\colon V \to W$ linear, also $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$. Wir definieren die Komplexifizierung $\tilde{f}$ von f durch

$$\begin{aligned}\tilde{f}\colon \tilde{V} &\to \tilde{W}\\ \tilde{f}(v_1 + iv_2) &= f(v_1) + if(v_2) \quad \text{für } v_1, v_2 \in V.\end{aligned}$$

Bemerkungen. 1. $\tilde{f}$ ist ein $\mathbb{C}$-Homomorphismus.

Beweis. Durch Nachrechnen, z.B.

$$\begin{aligned}\tilde{f}((\alpha + i\beta)(v_1 + iv_2)) &= f(\alpha v_1 - \beta v_2) + if(\alpha v_2 + \beta v_1) && \text{nach Definition}\\ &= \alpha f(v_1) - \beta f(v_2) + i(\alpha f(v_2) + \beta f(v_1))\\ &= (\alpha + i\beta)\tilde{f}(v_1 + iv_2). && \text{QED}\end{aligned}$$

2. $\tilde{f}$ ist eindeutig bestimmte Fortsetzung von f auf $\tilde{V}$. ('Fortsetzung' meint $\tilde{f}\restriction V = f$).

Beweis. Sei $f'\colon \tilde{V} \to \tilde{W}$ ein $\mathbb{C}$-Homomorphismus, $f'\restriction V = f$. Dann gilt

$$\begin{aligned}f'(v_1 + iv_2) &= f'(v_1) + if'(v_2)\\ &= f(v_1) + if(v_2)\\ &= \tilde{f}(v_1 + iv_2). \quad \text{QED}\end{aligned}$$

3. Seien V, W, U $\mathbb{R}$-Vektorräume und $V \xrightarrow{f} W \xrightarrow{g} U$ $\mathbb{R}$-Homomorphismen. Dann gilt offenbar $\tilde{V} \xrightarrow{\tilde{f}} \tilde{W} \xrightarrow{\tilde{g}} \tilde{U}$. Wir zeigen

a. $\widetilde{g \circ f} = \tilde{g} \circ \tilde{f}$.

b. $\widetilde{\mathrm{id}} = \mathrm{id}$.

Man sagt in diesem Fall, daß ˜ ein '*Funktor*' ist.

Beweis. a.

$$\begin{aligned}\tilde{g}(\tilde{f}(v_1 + iv_2)) &= \tilde{g}(f(v_1) + if(v_2))\\ &= g(f(v_1)) + ig(f(v_2))\\ &= (g \circ f)(v_1) + i(g \circ f)(v_2)\\ &= (\widetilde{g \circ f})(v_1 + iv_2)\end{aligned}$$

b. $\widetilde{\mathrm{id}}(v_1 + iv_2) = v_1 + iv_2$. QED

8.1.5. Lemma. *Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum und $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$.*

1. *Ist $x_1, \ldots, x_n$ eine Basis von V über $\mathbb{R}$, so ist $x_1, \ldots, x_n$ Basis von $\tilde{V}$ über $\mathbb{C}$.*
2. *Es gilt $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(\tilde{V})$ und $\dim_{\mathbb{R}}(\tilde{V}) = 2\dim_{\mathbb{R}}(V)$.*
3. *Das charakteristische Polynom p_f von f ist gleich dem charakteristischen Polynom $p_{\tilde{f}}$ von $\tilde{f}$.*

Beweis. 1. Offenbar ist $V \subseteq \tilde{V}$. $x_1, \dots, x_n$ sind linear unabhängig über $\mathbb{C}$: Seien $\alpha_j, \beta_j \in \mathbb{R}$ und gelte $\sum_{j=1}^n (\alpha_j + i\beta_j)x_j = 0$. Dann folgt

$$\sum_{j=1}^n \alpha_j x_j + i \sum_{j=1}^n \beta_j x_j = 0$$

$$\sum_{j=1}^n \alpha_j x_j = 0 \quad \text{und} \quad \sum_{j=1}^n \beta_j x_j = 0$$

$$\alpha_1 = \dots = \alpha_n = 0 \quad \text{und} \quad \beta_1 = \dots = \beta_n = 0, \quad \text{da } x_1, \dots, x_n \ \mathbb{R}\text{-linear unabhängig sind.}$$

$x_1, \dots, x_n$ bilden ein Erzeugendensystem über $\mathbb{C}$ von $\tilde{V}$: Sei $v + iw \in \tilde{V}$ mit $u, w \in V$. Dann gilt $v = \sum_{j=1}^n \alpha_j x_j$ und $w = \sum_{j=1}^n \beta_j x_j$ mit $\alpha_j, \beta_j \in \mathbb{R}$, also $\sum_{j=1}^n (\alpha_j + i\beta_j)x_j = v + iw$.

2. $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(\tilde{V})$ gilt nach Teil 1. Es ist $\dim_{\mathbb{R}}(\tilde{V}) = 2n$, da $x_1, \dots, x_n, ix_1, \dots, ix_n$ eine $\mathbb{R}$-Basis von $\tilde{V}$ bildet.

3. Sei $x_1, \dots, x_n$ Basis von V über $\mathbb{R}$ und A sei darstellende Matrix von f bezüglich $x_1, \dots, x_n$.

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Also gilt $p_f = \det(A - tE)$. Nach Teil (1) ist wegen

$$\tilde{f} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^t \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

A auch darstellende Matrix von $\tilde{f}$. Also hat man $p_{\tilde{f}} = \det(A - tE) = p_f$. Qed

Folgerung. *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $f \in \operatorname{Hom}_{\mathbb{R}}(V, V)$. Dann gilt:*

1. *Ist $\dim(V)$ ungerade, so besitzt f mindestens einen Eigenvektor, d.h. es gibt einen Unterraum $U \subseteq V$ der Dimension 1, welcher f-invariant ist (also $f(U) \subseteq U$).*
2. *Ist $\dim(V)$ beliebig (aber ≥ 1), so gibt es einen f-invarianten Unterraum $U \subseteq V$ der Dimension 1 oder 2.*

Beweis. 2. Betrachte $\tilde{f} \in \operatorname{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{V})$. Wegen $\dim(V) \geq 1$ hat $p_{\tilde{f}}$ eine Nullstelle λ über $\mathbb{C}$. Sei etwa $\lambda =: \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$. λ ist dann Eigenwert von $\tilde{f}$. Sei $x = v + iw \in \tilde{V}$ Eigenvektor von $\tilde{f}$ zum Eigenwert λ (also $x \neq 0$ und $v, w \in V$). Man erhält

$$\tilde{f}(x) = \lambda x$$

$$\tilde{f}(v + iw) = (\alpha + i\beta)(v + iw)$$

$$f(v) + if(w) = \alpha v - \beta w + i(\alpha w + \beta v)$$

$$\begin{cases} f(v) = \alpha v - \beta w \\ f(w) = \alpha w + \beta v \end{cases}$$

$$f \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}$$

Der Unterraum $U := \mathbb{R}v + \mathbb{R}w$ leistet also das Verlangte.

1. p_f hat eine reelle Nullstelle $\alpha \in \mathbb{R}$ (nach einer Folgerung in 8.1.2). Also ist α ein Eigenwert von f. Sei $v \neq 0$, $v \in V$ Eigenvektor von f zum Eigenwert α. Dann ist $f(v) = \alpha v$. Der Unterraum $U := \mathbb{R}v$ leistet also das Verlangte. Qed

8.1.6. *Konjugation.* Sei V ein $\mathbb{R}$-Vektorraum. Die Abbildung

$$\begin{aligned} &\tilde{V} \to \tilde{V} \\ &v + iw \mapsto v - iw \quad \text{für } v, w \in V \end{aligned}$$

heißt *Konjugation* auf $\tilde{V}$.

Bemerkungen. 1. Für alle $x, y \in \tilde{V}$ und $\alpha \in \mathbb{C}$ gilt

$$\begin{aligned} &\overline{x + y} = \overline{x} + \overline{y} \\ &\overline{\alpha x} = \overline{\alpha}\,\overline{x} \\ &\overline{\overline{x}} = x \\ &\overline{x} = x \quad \text{genau dann, wenn} \quad x \in V. \end{aligned}$$

Beweis. Wie für $\mathbb{C} = \tilde{\mathbb{R}}$. Zum Beispiel zeigt man die letzte Äquivalenz wie folgt. Sei $x = v + iw$ mit $v, w \in V$, also $\overline{x} = v - iw$. Dann sind folgende Aussagen äquivalent.

$$\begin{aligned} &x = \overline{x} \\ &v + iw = v - iw \\ &2iw = 0 \\ &w = 0 \\ &x \in V. \quad \text{QED} \end{aligned}$$

2. Sei wieder $x = v + iw$ mit $v, w \in V$. Dann gilt

$$v = \frac{1}{2}(x + \overline{x}), \quad w = \frac{1}{2i}(x - \overline{x}).$$

8.1.7. *Abstieg von $\mathbb{C}$ nach $\mathbb{R}$.* Wir zeigen, wann man aus dem Vorhandensein von Homomorphismen oder Unterräumen in der Komplexifizierung $\tilde{V}$ schließen kann auf das Vorhandensein entsprechender Objekte in V.

Satz. *Seien V, W $\mathbb{R}$-Vektorräume.*

1. *Für alle $x \in \tilde{V}$ gilt $x \in V$ genau dann, wenn $x = \overline{x}$.*

2. *Ein beliebiges $g \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{W})$ hat die Form $g = \tilde{f}$ für ein $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$ genau dann, wenn $\overline{g(x)} = g(\overline{x})$ für alle $x \in \tilde{V}$. In diesem Fall ist $f = g|V$.*

3. *Für einen beliebigen $\mathbb{C}$-Unterraum U' von $\tilde{V}$ gilt: Es gibt einen $\mathbb{R}$-Unterraum U von V mit $U' = \tilde{U}$ genau dann, wenn für alle $x \in U'$ gilt $\overline{x} \in U'$, d.h. wenn U' invariant ist unter Konjugation. In diesem Fall ist $U = \{\, x \in U' \mid x = \overline{x} \,\}$.*

Beweis. 1. Dies war oben als Bemerkung 1 bewiesen.

2. Sei $g \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{W})$. Nehmen wir zunächst an, daß $g = \tilde{f}$ für ein $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$. Dann gilt für ein beliebiges $x = v_1 + iv_2$ mit $v_1, v_2 \in V$

$$\begin{aligned} \overline{\tilde{f}(x)} &= \overline{f(v_1) + if(v_2)} \\ &= f(v_1) - if(v_2) \\ &= f(v_1) + if(-v_2) \\ &= \tilde{f}(v_1 - iv_2) \\ &= \tilde{f}(\overline{x}). \end{aligned}$$

Sei nun umgekehrt $\overline{g(x)} = g(\overline{x})$ für alle $x \in \tilde{V}$. Für alle $x \in V$ gilt dann $\overline{g(x)} = g(\overline{x}) = g(x)$ (da $x = \overline{x}$), also $g(x) \in W$ nach Teil 1. Daher ist $g{\restriction}V =: f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$. Weiter gilt $\tilde{f} = g$, da die Fortsetzung von f zu einem $\tilde{f} \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{W})$ eindeutig bestimmt ist.

3. Sei $U' \subseteq \tilde{V}$ ein $\mathbb{C}$-Unterraum. Nehmen wir zunächst an, daß $U' = \tilde{U}$ für einen $\mathbb{R}$-Unterraum U von V. Dann läßt sich jedes $x \in \tilde{U}$ schreiben in der Form $x = u_1 + iu_2$ mit $u_1, u_2 \in U$. Also ist auch $\overline{x} = u_1 - iu_2 \in \tilde{U}$. Sei nun umgekehrt U' invariant unter Konjugation. Gesucht ist ein $\mathbb{R}$-Unterraum U von V mit $U' = \tilde{U}$. Wir setzen $U := \{\, x \in U' \mid x = \overline{x} \,\}$.

Zu $U' \subseteq \tilde{U}$. Sei $x = v_1 + iv_2 \in U'$ mit $v_1, v_2 \in V$. Dann ist $v_1 = \frac{1}{2}(x + \overline{x})$ und $v_2 = \frac{1}{2i}(x - \overline{x})$. Beide Vektoren sind in U', da mit x auch $\overline{x} \in U'$ ist. Wegen $v_1 = \overline{v_1}$ und $v_2 = \overline{v_2}$ haben wir $v_1, v_2 \in U$ und damit $x \in \tilde{U}$.

Zu $\tilde{U} \subseteq U'$. Sei $x \in \tilde{U}$, also $x = u_1 + iu_2$ mit $u_1, u_2 \in U$. Dann sind $u_1, u_2 \in U'$, also auch $x \in U'$.

Wir müssen noch zeigen, daß U eindeutig bestimmt ist. Sei also $U_0 \subseteq V$ ein weiterer $\mathbb{R}$-Unterraum mit $U' = \tilde{U}_0$. Gezeigt wird $U_0 = U := \{\, x \in U' \mid x = \overline{x} \,\}$. Sei zunächst $x \in U_0$. Dann ist $x = \overline{x}$ und $x \in \tilde{U}_0 = U'$, also $x \in U$. Sei nun umgekehrt $x \in U$. Dann ist $x \in U'$ mit $x = \overline{x}$, und ferner $x = v_1 + iv_2$ mit $v_1, v_2 \in U_0$ (da $x \in U' = \tilde{U}_0$). Wegen $x = \overline{x}$ folgt $v_2 = 0$, also $x = v_1 \in U_0$. QED

BEMERKUNGEN. 1. Wir haben die Komplexifizierung bezüglich $\mathbb{R} \subseteq \mathbb{C}$ durchgeführt. Allgemeiner kann man eine entsprechende Konstruktion für eine beliebige Körpererweiterung $K \subseteq L$ und einen K-Vektorraum V durchführen. Man kann dann $L \otimes_K V$ als L-Vektorraum auffassen (siehe den Abschnitt 10.4 über Tensorprodukte).

2. Ist eine Körpererweiterung $K \subseteq L$ eine sogenannte *Galoiserweiterung* (siehe Algebra-Vorlesung), so läßt sich der Abstieg von L nach K wie eben der von $\mathbb{C}$ nach $\mathbb{R}$ durchführen.

8.2 Anwendung auf die reelle Jordansche Normalform

Zunächst sei noch einmal an den Beweis des Satzes über die JORDANsche Normalform erinnert (wir schreiben jetzt V^λ anstelle von $U_{f,\lambda}$). Für jeden Eigenwert λ von f hatten wir $g := f - \lambda\mathrm{id}$ gesetzt und den Hauptraum V^λ zum Eigenwert λ definiert durch

$$V^\lambda := \bigcup_{m=0}^{\infty} \mathrm{Kern}(g^m).$$

Es war

$$0 \subsetneq \mathrm{Kern}(g) \subsetneq \cdots \subsetneq \underbrace{\mathrm{Kern}(g^k)}_{V^\lambda} = \mathrm{Kern}(g^{k+1}) = \ldots .$$

In V^λ wurde dann eine geeignete Basis gewählt. Die Dimension von V^λ war die Vielfachheit von λ als Nullstelle von p_f. Ferner hatten wir eine Zerlegung

$$V = \underbrace{\mathrm{Kern}(g^k)}_{V^\lambda} \oplus \mathrm{Bild}(g^k)$$

in f-invariante Unterräume erreicht. Das volle Resultat ergab sich durch Iteration. Aus dem damaligen Überlegungen erhält man insbesondere:

8.2.1. SATZ. *(Hauptraumzerlegung). Sei V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Ferner sei p_f Produkt linearer Faktoren, etwa $p_f = (-1)^n \prod_{j=1}^{s}(t-\lambda_j)^{e_j}$ mit verschiedenen λ_j und $e_j \geq 1$. Für jedes $\lambda \in K$ setzen wir $V^\lambda := \bigcup_{m=0}^{\infty} \mathrm{Kern}(g^m)$. Dann gilt*

$$V = \bigoplus_{\lambda \in K} V^\lambda.$$

BEWEIS. Offenbar ist $V^\lambda \neq 0$ genau dann, wenn λ Eigenwert von f ist. Es seien $\lambda_1, \dots, \lambda_s$ die verschiedenen Eigenwerte von f. Wir betrachten die Abbildung

$$\begin{aligned} h\colon V^{\lambda_1} \times \cdots \times V^{\lambda_s} &\to V \\ (x_1, \dots, x_s) &\mapsto x_1 + \cdots + x_s, \end{aligned}$$

Man sieht sofort, daß h linear ist. h ist surjektiv aufgrund der Konstruktion der Basis im Beweis des Satzes über die JORDANsche Normalform (Basiselemente, die zum Eigenwert λ gehören, liegen in V^λ). Ferner gilt nach dem Dimensionssatz

$$\begin{aligned} \dim(\mathrm{Kern}(h)) + \dim(V) &= \dim(V^{\lambda_1} \times \cdots \times V^{\lambda_s}) \\ &= \dim(V^{\lambda_1}) + \cdots + \dim(V^{\lambda_s}) \\ &= e_1 + \cdots + e_s \\ &= \dim(V). \end{aligned}$$

Demnach ist $\mathrm{Kern}(h) = 0$, also h ein Isomorphismus und damit $V = \bigoplus_{j=1}^{s} V^{\lambda_j}$. QED

8.2.2. SATZ. *(Reelle* JORDAN*sche Normalform). Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum und $f\colon V \to V$ linear. Dann gibt es eine Basis $v_1, \dots, v_n$ von V mit*

$$f\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

wobei jedes $A_j \in \mathbb{R}^{n_j \times n_j}$ die Form

$$\begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix} \quad \text{λ reeller Eigenwert von f}$$

oder die Form

$$\begin{pmatrix} \boxed{\begin{matrix} \alpha & -\beta \\ \beta & \alpha \end{matrix}} & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} \alpha & -\beta \\ \beta & \alpha \end{matrix}} & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & & \\ & & \ddots & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \\ 0 & & & \boxed{\begin{matrix} \alpha & -\beta \\ \beta & \alpha \end{matrix}} & \end{pmatrix} \quad \text{$\lambda := \alpha + i\beta$ ($\beta \neq 0$) komplexer Eigenwert von $\tilde{f}$}$$

(mit $\alpha, \beta \in \mathbb{R}$) hat.

BEWEIS. Wir betrachten die eindeutige Fortsetzung $\tilde{f}\colon \tilde{V} \to \tilde{V}$ von f auf $\tilde{V}$. Nach Satz 8.2.1 gilt $\tilde{V} = \bigoplus_{\lambda \in \mathbb{C}} \tilde{V}^\lambda$. Für $\tilde{V}$ können wir jetzt den Satz über die JORDANsche Normalform anwenden, da $\mathbb{C}$ algebraisch abgeschlossen ist. Er liefert eine Basis, so daß die darstellende Matrix von f JORDAN-Normalform hat. Unser Ansatz ist die Konstruktion einer solchen Basis, und anschließend ein Abstieg von $\mathbb{C}$ nach $\mathbb{R}$.

1. Sei $\lambda \in \mathbb{C}$, λ Eigenwert von $\tilde{f}$. Wir setzen $g := \tilde{f} - \lambda\mathrm{id}$ und betrachten $\overline{g} := \tilde{f} - \overline{\lambda}\mathrm{id} \in \mathrm{Hom}_{\mathbb{C}}(\tilde{V}, \tilde{V})$. Dann gilt

$$\overline{g^l(x)} = \overline{g}^l(\overline{x}) \quad \text{für alle } x \in \tilde{V}, l \geq 0.$$

BEWEIS. Es ist $\overline{g(x)} = \overline{\tilde{f}(x) - \lambda x} = \tilde{f}(\overline{x}) - \overline{\lambda}\overline{x} = \overline{g}(\overline{x})$. Wir führen den Beweis durch Induktion über l. *Basis* $l = 0$. Wegen $g^0 = \overline{g}^0 = \mathrm{id}$ ist hier nichts zu zeigen. *Schritt* $l \mapsto l + 1$.

$$\begin{aligned}
\overline{g^{l+1}(x)} &= \overline{g(g^l(x))} \\
&= \overline{g}(\overline{g^l(x)}) && \text{nach der Vorbemerkung} \\
&= \overline{g}(\overline{g}^l(\overline{x})) && \text{nach Induktionsvoraussetzung} \\
&= \overline{g}^{l+1}(\overline{x}).
\end{aligned}$$

2. Sei λ komplexer Eigenwert von $\tilde{f}$. Dann ist

$$\begin{aligned}
\tilde{V}^\lambda &\to \tilde{V}^{\overline{\lambda}} \\
x &\mapsto \overline{x}
\end{aligned}$$

bijektiv.

BEWEIS. Sei $x \in \tilde{V}^\lambda$, also $g^l(x) = 0$ für ein l. Nach (1) ist $\overline{g^l(x)} = \overline{g}^l(\overline{x}) = 0$, also $\overline{x} \in \tilde{V}^{\overline{\lambda}}$. Genauso zeigt man, daß aus $y \in \tilde{V}^{\overline{\lambda}}$ folgt $\overline{y} \in \tilde{V}^\lambda$.

3. Sei λ reeller Eigenwert von $\tilde{f}$ (also auch von f). Nach (2) ist $\tilde{V}^\lambda$ invariant bei Konjugation. Nach Teil 3 des Satzes in 8.1.7 ist dann $\tilde{V}^\lambda = \tilde{U}$ für einen $\mathbb{R}$-Unterraum $U \subseteq V$, und es gilt

$$U = \{\, x \in \tilde{V}^\lambda \mid x = \overline{x} \,\} = \{\, x \in V \mid \exists l \underbrace{(\tilde{f} - \lambda\mathrm{id})^l(x)}_{=(f-\lambda\mathrm{id})^l(x)} = 0 \,\} = V^\lambda.$$

Also ist $\widetilde{V^\lambda} = \tilde{V}^\lambda$. Man wähle man nun wie im Beweis der JORDAN-Normalform eine 'gute' Basis von V^λ bezüglich f. Sie ist dann auch Basis von $\tilde{V}^\lambda$ bezüglich $\tilde{f}$.

4. Sei λ komplexer Eigenwert von $\tilde{f}$ mit $\lambda \neq \overline{\lambda}$. Sei $v_1 + iw, \ldots, v_m + iw_m$ eine 'gute' Basis von $\tilde{V}^\lambda$ mit $v_j, w_j \in V$, so daß $\tilde{f}$ bezüglich dieser Basis durch eine Matrix in JORDANscher Normalform dargestellt wird. Da $\tilde{V}^\lambda \to \tilde{V}^{\overline{\lambda}}$, $x \mapsto \overline{x}$ ein Isomorphismus ist, muß $v_1 - iw_1, \ldots, v_m - iw_m$ Basis von $\tilde{V}^{\overline{\lambda}}$ sein. Also ist $v_1, w_1, \ldots, v_m, w_m$ eine $\mathbb{C}$-Basis von $\tilde{V}^\lambda \oplus \tilde{V}^{\overline{\lambda}}$ (da es sich um ein Erzeugendensystem aus $\dim(\tilde{V}^\lambda \oplus \tilde{V}^{\overline{\lambda}}) = 2m$ Elementen handelt). OBdA sei

$$\tilde{f}\begin{pmatrix} v_1 + iw_1 \\ \vdots \\ v_m + iw_m \end{pmatrix} = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix} \begin{pmatrix} v_1 + iw_1 \\ \vdots \\ v_m + iw_m \end{pmatrix}$$

(man betrachte jede JORDAN-Matrix einzeln). Also gilt für $j < m$ mit $\lambda =: \alpha + i\beta$

$$\underbrace{\tilde{f}(v_j + iw_j)}_{f(v_j)+if(w_j)} = \underbrace{(\alpha + i\beta)(v_j + iw_j)}_{\alpha v_j - \beta w_j + i(\beta v_j + \alpha w_j)} + v_{j+1} + iw_{j+1}$$

$$\begin{cases} f(v_j) = \alpha v_j - \beta w_j + v_{j+1} \\ f(w_j) = \beta v_j + \alpha w_j + w_{j+1} \end{cases}$$

$$f\begin{pmatrix} v_j \\ w_j \end{pmatrix} = \begin{pmatrix} \alpha & -\beta & 1 & 0 \\ \beta & \alpha & 0 & 1 \end{pmatrix} \begin{pmatrix} v_j \\ w_j \\ v_{j+1} \\ w_{j+1} \end{pmatrix}.$$

Ebenso erhält man für $j = m$

$$f\begin{pmatrix} v_m \\ w_m \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} v_m \\ w_m \end{pmatrix}.$$

5. Konstruiert wurde eine $\mathbb{C}$-Basis von $\tilde{V} = \bigoplus_{\lambda \in \mathbb{C}} \tilde{V}^\lambda$, bestehend aus Vektoren aus V. Da $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(\tilde{V})$, und die Familie der $\mathbb{C}$-Basis auch $\mathbb{R}$-linear unabhängig ist, hat man damit eine $\mathbb{R}$-Basis von V konstruiert. f hat bezüglich dieser Basis die gewünschte Form. QED

8.3 Unitäre Vektorräume

Die komplexen Zahlen kann man als den euklidischen Vektorraum $\mathbb{R}^2$ auffassen, indem man $\xi = \alpha + i\beta$ mit dem Paar (α, β) identifiziert (in 1.2.5 waren die komplexen Zahlen so definiert worden). Für den Betrag gilt dann

$$\|(\alpha, \beta)\| = \sqrt{\alpha^2 + \beta^2} = |\xi| = \sqrt{\xi\overline{\xi}}.$$

Wir betrachten jetzt die folgende Abbildung

$$\begin{aligned} \sigma\colon \mathbb{C} \times \mathbb{C} &\to \mathbb{C} \\ (\xi, \xi') &\mapsto \xi\overline{\xi'}. \end{aligned}$$

Dann gilt offenbar

$$\begin{aligned} &\sigma(\xi, \xi) \geq 0, \\ &+\sqrt{\sigma(\xi, \xi)} = |\xi|, \\ &\sigma(\xi, \xi') = \xi\overline{\xi'} = \overline{\overline{\xi}\,\xi'} = \overline{\sigma(\xi', \xi)}. \end{aligned}$$

Entsprechend erweitern wir jetzt den Begriff eines euklidischen Vektorraums (über $\mathbb{R}$) zu dem eines unitären Vektorraums (über $\mathbb{C}$).

8.3.1. DEFINITION. Sei V ein $\mathbb{C}$-Vektorraum und $\sigma\colon V \times V \to \mathbb{C}$ eine Abbildung. σ heißt *hermitesches Skalarprodukt*, wenn σ linear im ersten Argument ist, d.h. wenn für alle $x, y, z \in V$ und alle $\alpha, \beta \in \mathbb{C}$ gilt

$$\sigma(\alpha x + \beta y, z) = \alpha\sigma(x, z) + \beta\sigma(y, z),$$

ferner σ eine *hermitesche Abbildung* ist, d.h. für alle $x, y \in V$ gilt

$$\sigma(x, y) = \overline{\sigma(y, x)},$$

und σ positiv definit ist, d.h. für alle $x \in V$ gilt

$$\sigma(x, x) > 0 \quad \text{für } x \neq 0.$$

(V, σ) heißt *unitärer Vektorraum*, wenn σ ein hermitesches Skalarprodukt auf V ist.

BEMERKUNGEN. Sei (V, σ) ein unitärer Vektorraum.

1. Es ist $\sigma(x, x) \in \mathbb{R}$ für alle $x \in V$ (wegen $\sigma(x, x) = \overline{\sigma(x, x)}$). Deshalb ist die Forderung $\sigma(x, x) > 0$ in der Definition sinnvoll.

2. Offenbar gilt $\sigma(x, x) = 0$ genau dann, wenn $x = 0$. Beweis: Aus $\sigma(x, x) = 0$ folgt $x = 0$ wegen der positiven Definitheit. Umgekehrt ist $\sigma(0, y) = 0$ für alle y, da $\sigma(x, y)$ linear im ersten Argument ist.

3. Es gilt $\sigma(x, \alpha y + \beta z) = \overline{\alpha}\sigma(x, y) + \overline{\beta}\sigma(x, z)$, denn

$$\begin{aligned} \sigma(x, \alpha y + \beta z) &= \overline{\sigma(\alpha y + \beta z, x)} \\ &= \overline{\alpha}\overline{\sigma(y, x)} + \overline{\beta}\,\overline{\sigma(z, x)} \\ &= \overline{\alpha}\sigma(x, y) + \overline{\beta}\sigma(x, z). \end{aligned}$$

Man sagt deshalb, daß σ *semilinear* im zweiten Argument ist.

4. Man nennt Abbildungen $\sigma\colon V \times V \to \mathbb{C}$, die linear im ersten und semilinear im zweiten Argument sind, auch *Sesquilinearformen* (lat. "anderthalb").

BEISPIELE. 1. Das Standardbeispiel ist $V = \mathbb{C}^n$, wobei für $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ und $y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}$ gesetzt wird

$$\sigma(x, y) := x^t\overline{y} = \xi_1\overline{\eta_1} + \cdots + \xi_n\overline{\eta_n}.$$

Man zeigt leicht, daß damit ein hermitesches Skalarprodukt definiert ist: Offenbar ist σ linear im ersten Argument. σ ist hermitesche Abbildung, denn

$$\overline{\sigma(x, y)} = \overline{x^t\overline{y}} = \overline{x}^t y = y^t\overline{x} = \sigma(y, x).$$

σ ist positiv definit, denn

$$\sigma(x, x) = x^t\overline{x} = \xi_1\overline{\xi_1} + \cdots + \xi_n\overline{\xi_n} = |\xi_1|^2 + \cdots + |\xi_n|^2 > 0,$$

falls ein $|\xi_j| \neq 0$. Man nennt σ das *kanonische Skalarprodukt* im $\mathbb{C}^n$.

2. Sei (V, σ) ein unitärer Vektorraum, W ein $\mathbb{C}$-Vektorraum und $f\colon W \to V$ ein $\mathbb{C}$-Monomorphismus (z.B. $W \subseteq V$ ein Unterraum und f die Einbettung). Setze

$$\sigma_f(x, y) := \sigma(f(x), f(y))$$

für alle $x, y \in W$. Dann ist (W, σ_f) ein unitärer Vektorraum.

BEWEIS. Offenbar ist σ_f linear im ersten Argument. σ_f ist hermitesche Abbildung, denn

$$\overline{\sigma_f(x, y)} = \overline{\sigma(f(x), f(y))} = \sigma(f(y), f(x)) = \sigma_f(y, x).$$

σ_f ist positiv definit, denn für $x \in W$ mit $x \neq 0$ ist $f(x) \neq 0$ (da f ein Monomorphismus ist), also $0 < \sigma(f(x), f(x)) = \sigma_f(x, x)$. QED

8.3.2. *Übertragung der Theorie euklidischer Vektorräume.* Viele Definitionen und Sätze mit ihren Beweisen lassen sich leicht von euklidischen auf unitäre Vektorräume übertragen. Wir wollen dies im folgenden durchführen.

Sei (V, σ) ein unitärer Vektorraum und $x, y \in V$. Dann nennt man $\|x\| := \sqrt{\sigma(x,x)} \geq 0$ die *Länge* von x. Ferner nennt man x *orthogonal* zu y und schreibt $x \perp y$, wenn gilt $\sigma(x,y) = 0$. Ist noch $y \neq 0$, so nennt man ein $p \in V$ *orthogonale Projektion* von x auf y, wenn gilt

1. $p \in \mathbb{C}y$ und
2. $x - p \perp y$.

$h := x - p$ heißt *Lot* von x auf y.

SATZ. *Sei (V, σ) ein unitärer Vektorraum, $x, y \in V$ und $y \neq 0$. Es gibt genau eine orthogonale Projektion von x auf y, und zwar*

$$p = \frac{\sigma(x,y)}{\sigma(y,y)} y.$$

BEWEIS. Wie in Abschnitt 5.1.

SATZ. *Sei (V, σ) ein unitärer Vektorraum und $x, y \in V$. Dann gilt*

1. *$\|x\| = 0$ genau dann, wenn $x = 0$.*
2. *$\|\alpha x\| = |\alpha| \cdot \|x\|$ für alle $\alpha \in \mathbb{C}$.*
3. *(SCHWARZsche Ungleichung). $|\sigma(x,y)| \leq \|x\| \cdot \|y\|$. Die Gleichheit gilt genau dann, wenn x und y linear abhängig sind.*
4. *(Dreiecksungleichung). $\|x + y\| \leq \|x\| + \|y\|$.*

BEWEIS. Wie in Abschnitt 5.1. Nur für die SCHWARZsche Ungleichung und die Dreiecksungleichung muß die frühere Argumentation leicht verändert werden.

3. OBdA sei $y \neq 0$. Sei $p = \frac{\sigma(x,y)}{\sigma(y,y)} y$ die orthogonale Projektion von x auf y und $h := x - p$ das Lot von x auf y. Dann gilt

$$\begin{aligned}
0 &\leq \sigma(h,h) \\
&= \sigma(x-p, x-p) \\
&= \sigma(x,x) - \sigma(p,x) - \sigma(x,p) + \sigma(p,p) \\
&= \sigma(x,x) - \frac{\sigma(x,y)}{\sigma(y,y)}\sigma(y,x) - \frac{\overline{\sigma(x,y)}}{\overline{\sigma(y,y)}}\sigma(x,y) + \frac{\sigma(x,y)\overline{\sigma(x,y)}}{\sigma(y,y)\overline{\sigma(y,y)}}\sigma(y,y) \\
&= \frac{\sigma(x,x)\sigma(y,y) - \sigma(x,y)\overline{\sigma(x,y)}}{\sigma(y,y)},
\end{aligned}$$

also $|\sigma(x,y)|^2 \leq \sigma(x,x)\sigma(y,y)$. Für den Zusatz (Gleichheit gilt genau dann, wenn x und y linear abhängig sind) kann man den Beweis in 5.1 unverändert übernehmen.

4. Man erhält aufgrund der SCHWARZschen Ungleichung

$$\begin{aligned}\sigma(x+y,x+y) &= \sigma(x,x)+\sigma(x,y)+\sigma(y,x)++\sigma(y,y)\\ &= \sigma(x,x)+\sigma(x,y)+\overline{\sigma(x,y)}+\sigma(y,y)\\ &= \sigma(x,x)+2\Re(\sigma(x,y))+\sigma(y,y)\\ &\le \sigma(x,x)+2|\sigma(x,y)|+\sigma(y,y)\\ &\le \|x\|^2+2\|x\|\cdot\|y\|+\|y\|^2\\ &= (\|x\|+\|y\|)^2. \quad \text{QED}\end{aligned}$$

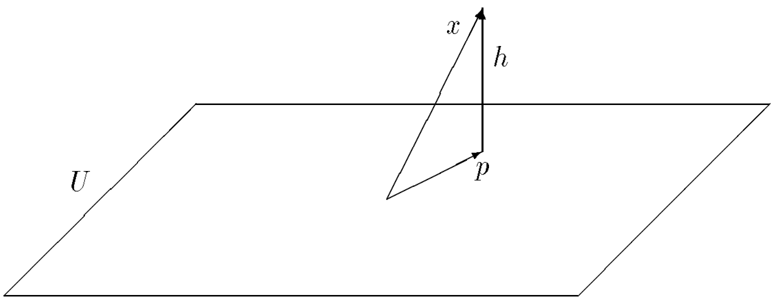
Eine Familie $(e_i)_{i\in I}$ von Vektoren aus einem unitären Vektorraum heißt *Orthonormalsystem* (ONS), wenn gilt

$$\begin{aligned}\sigma(e_i,e_j) &= 0 \quad \text{für } i,j\in I,\ i\neq j,\\ \sigma(e_i,e_i) &= 1 \quad \text{für } i\in I.\end{aligned}$$

$(e_i)_{i\in I}$ heißt *Orthonormalbasis* (ONB), wenn zusätzlich $(e_i)_{i\in I}$ Basis ist. Im Standardbeispiel $(\mathbb{C}^n,\sigma)$ mit $\sigma(x,y) = x^t\overline{y}$ ist $(e_1,\dots,e_n)$ Orthonormalbasis. Ferner gilt wieder: Ist $(e_i)_{i\in I}$ Orthonormalbasis, so ist $(e_i)_{i\in I}$ linear unabhängig. Ist $x\in V$ und $U\subseteq V$ Unterraum, so heißt $p\in V$ *orthogonale Projektion* von x auf U, wenn gilt

1. $p\in U$ und
2. $x-p\perp U$ (d.h. $x-p\perp u$ für alle $u\in U$).

$h := x-p$ heißt *Lot* von x auf U. (Für $U=\mathbb{C}y$ ist das die alte Definition).



SATZ. *Sei (V,σ) ein unitärer Vektorraum, $x\in V$, ferner $U\subseteq V$ ein Unterraum und $(e_1,\dots,e_n)$ eine Orthonormalbasis von U. Dann gibt es genau eine orthogonale Projektion von x auf U, und zwar*

$$p=\sum_{i=1}^{n}\sigma(x,e_i)e_i.$$

BEWEIS. Wie in Abschnitt 5.1. QED

SATZ. *(SCHMIDTsches Orthonormalisierungsverfahren). Sei (V,σ) ein unitärer Vektorraum, $x_1,x_2,x_3,\dots$ linear unabhängig (endlich oder abzählbar unendlich viele Vektoren). Dann kann man ein Orthonormalsystem $e_1,e_2,e_3,\dots$ konstruieren (soviele wie die x_n's) so daß für jedes n gilt: $(e_1,\dots,e_n)$ ist Orthonormalbasis von $U_n := \langle x_1,\dots,x_n\rangle$.*

BEWEIS. Wie in Abschnitt 5.1. QED

Jeder endlichdimensionale unitäre Vektorraum besitzt also eine Orthonormalbasis.

Satz. *Sei V ein unitärer Vektorraum und $U \subseteq V$ ein endlichdimensionaler Unterraum. Dann gilt*

1. $V = U \oplus U^{\perp}$.
2. $U = U^{\perp\perp}$.

Beweis. 1. $V = U + U^{\perp}$. Sei $x \in V$ und p die orthogonale Projektion von x auf U. Dann ist

$$x = \underbrace{p}_{\in U} + \underbrace{(x-p)}_{\in U^{\perp}} \in U + U^{\perp}.$$

$U \cap U^{\perp} = 0$. Sei $x \in U \cap U^{\perp}$. Dann ist $x \perp x$, also $\sigma(x,x) = 0$ und damit $x = 0$.

2. $U \subseteq U^{\perp\perp}$ gilt stets (auch für eine Menge M statt U). $U^{\perp\perp} \subseteq U$: Sei $x \in U^{\perp\perp}$. Dann gilt

$$\begin{aligned}
&x = \underbrace{p}_{\in U} + \underbrace{(x-p)}_{\in U^{\perp}} && \text{nach Teil (1)}\\
&\begin{cases} x - p \perp x & \text{da } x - p \in U^{\perp} \text{ und } x \in U^{\perp\perp}\\ x - p \perp p & \text{da } x - p \in U^{\perp} \text{ und } p \in U \end{cases}\\
&\sigma(x-p, x-p) = 0\\
&x - p = 0\\
&x = p \in U. && \text{Qed}
\end{aligned}$$

8.3.3. *Komplexifizierung.* Einen gegebenen euklidischen Vektorraum kann man stets in eindeutiger Weise zu einem unitären Vektorraum machen.

Satz. *Sei (V,σ) euklidischer Vektorraum und $\tilde{V}$ die Komplexifizierung von V. Dann gibt es genau ein hermitesches Skalarprodukt $\tilde{\sigma}$ auf $\tilde{V}$ mit $\tilde{\sigma}{\restriction}V = \sigma$.*

Beweis. Es war $V \subseteq \tilde{V} = \{\, x + iy \mid x, y \in V \,\}$.

Eindeutigkeit von $\tilde{\sigma}$. Sei $\tilde{\sigma}$ hermitesches Skalarprodukt mit $\tilde{\sigma}{\restriction}V = \sigma$. Dann gilt für alle $x, y, u, v \in V$

$$\begin{aligned}
\tilde{\sigma}(x+iy, u+iv) &= \tilde{\sigma}(x,u) + i\tilde{\sigma}(y,u) - i\tilde{\sigma}(x,v) + \tilde{\sigma}(y,v)\\
&= \sigma(x,u) + \sigma(y,v) + i(\sigma(y,u) - \sigma(x,v)).
\end{aligned}$$

Also ist $\tilde{\sigma}$ nur von σ abhängig.

Existenz von $\tilde{\sigma}$. Für $x, y, u, v \in V$ setze

$$\tilde{\sigma}(x+iy, u+iv) := \sigma(x,u) + \sigma(y,v) + i(\sigma(y,u) - \sigma(x,v)).$$

Durch Nachrechnen zeigt man, daß dann $\tilde{\sigma}$ ein hermitesches Skalarprodukt ist:

1. $\tilde{\sigma}$ ist linear im ersten Argument: Die Additivität im ersten Argument ist klar. Ferner gilt

$$\begin{aligned}
&\tilde{\sigma}((\alpha + i\beta)(x+iy), u+iv)\\
&= \tilde{\sigma}(\alpha x - \beta y + i(\alpha y + \beta x), u + iv)\\
&= \sigma(\alpha x - \beta y, u) + \sigma(\alpha y + \beta x, v) + i(\sigma(\alpha y + \beta x, u) - \sigma(\alpha x - \beta y, v)).
\end{aligned}$$

Andererseits ist

$$
\begin{aligned}
&(\alpha+i\beta)\tilde\sigma(x+iy,u+iv)\\
&=\alpha\sigma(x,u)+\alpha\sigma(y,v)-\beta\sigma(y,u)+\beta\sigma(x,v)+i(\beta\sigma(x,u)+\beta\sigma(y,v)+\alpha\sigma(y,u)-\alpha\sigma(x,v)).
\end{aligned}
$$

2. $\tilde\sigma$ ist hermitesch: Es gilt

$$
\begin{aligned}
\tilde\sigma(u+iv,x+iy) &= \sigma(u,x)+\sigma(v,y)+i(\sigma(v,x)-\sigma(u,y))\\
&= \sigma(x,u)+\sigma(y,v)+i(\sigma(x,v)-\sigma(y,u))\\
&= \overline{\tilde\sigma(x+iy,u+iv)}.
\end{aligned}
$$

3. $\tilde\sigma$ ist positiv definit: Es gilt

$$\tilde\sigma(x+iy,x+iy)=\sigma(x,x)+\sigma(y,y)+i(\underbrace{\sigma(y,x)-\sigma(x,y)}_{=0})>0,\quad \text{falls } x\neq 0 \text{ oder } y\neq 0. \quad \text{QED}$$

Bemerkung. Unter den Voraussetzungen des Satzes gilt: Ist $(e_1,\dots,e_n)$ eine Orthonormalbasis von V, so ist $(e_1,\dots,e_n)$ auch eine Orthonormalbasis von $\tilde V$. Beweis. Nach Lemma 8.1.5 ist $(e_1,\dots,e_n)$ ist eine $\mathbb{C}$-Basis von $\tilde V$, und außerdem gilt $\tilde\sigma(e_r,e_s)=\sigma(e_r,e_s)=\delta_{rs}$.

Beispiel. Auf $V=\mathbb{R}^n$ sei $\sigma(x,y)=x^ty$ für $x,y\in\mathbb{R}^n$. Auf $\tilde V=\mathbb{C}^n$ gilt dann $\tilde\sigma(a,b)=a^t\overline{b}$ für $a,b\in\mathbb{C}^n$. Dies kann man wie früher durch Verifikation beweisen, indem man zeigt, daß $\tilde\sigma$ ein hermitesches Skalarprodukt ist. Eine andere Möglichkeit ist die folgende. Nach dem Satz existiert $\tilde\sigma$ so daß $\tilde\sigma|V=\sigma$. Für $a=\sum_{j=1}^n\alpha_je_j$ und $b=\sum_\ell\beta_\ell e_\ell$ ist dann

$$\tilde\sigma(a,b)=\sum_{j,\ell}\alpha_j\overline{\beta_\ell}\sigma(e_j,e_\ell)=a^t\overline{b}.$$

8.3.4. Satz. *(Darstellungssatz). Sei (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $g\in V^*=\mathrm{Hom}_{\mathbb{C}}(V,\mathbb{C})$ (bzw. $=\mathrm{Hom}_{\mathbb{R}}(V,\mathbb{R})$). Dann gibt es genau ein $y\in V$ so daß für alle $x\in V$ gilt*

$$g(x)=\sigma(x,y).$$

Ist $(e_1,\dots,e_n)$ eine Orthonormalbasis von V, so gilt $y=\sum_{j=1}^n\overline{g(e_j)}e_j$.

Beweis. Sei (V,σ) ein unitärer Vektorraum (ebenso geht es im Fall eines euklidischen Vektorraums).

Existenz von y. Sei $(e_1,\dots,e_n)$ eine Orthonormalbasis von V und $x\in V$. Dann gilt

$$
\begin{aligned}
x &= \sum_{j=1}^n\sigma(x,e_j)e_j\\
g(x) &= \sum_{j=1}^n\sigma(x,e_j)g(e_j)=\sigma(x,\underbrace{\sum_{j=1}^n\overline{g(e_j)}e_j}_{=:y}).
\end{aligned}
$$

Eindeutigkeit von y. Seien y,y' mit den im Satz formulierten Eigenschaften gegeben. Dann gilt

$$
\begin{aligned}
&\sigma(x,y)=\sigma(x,y') && \text{für alle } x\in V\\
&\sigma(x,y-y')=0 && \text{für alle } x\in V\\
&\sigma(y-y',y-y')=0 &&\\
&y-y'=0. && \text{QED}
\end{aligned}
$$

8.3.5. *Adjungierter Endomorphismus.* Für jede Matrix $A = (\alpha_{ij}) \in \mathbb{C}^{n\times n}$ definieren wir

$$\overline{A} := (\overline{\alpha_{ij}}) \quad \text{und} \quad A^* := \overline{A}^t = (\overline{\alpha_{ji}})_{ij}.$$

$\overline{A}$ heißt die *konjugierte* Matrix und A^* die *adjungierte* Matrix zu A.

SATZ. *Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$).*

1. *Es gibt genau ein $f^* \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$) so daß für alle $x, y \in V$ gilt*

$$\sigma(f(x), y) = \sigma(x, f^*(y)).$$

   Dieses f^* heißt der zu f adjungierte *Endomorphismus. Ist $(e_1, \ldots, e_n)$ eine Orthonormalbasis von V, so gilt*

$$f^*(y) = \sum_{j=1}^{n} \sigma(y, f(e_j))e_j.$$

2. *Ist $A \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) die darstellende Matrix von f bzgl. einer Orthonormalbasis $\mathcal{A} = (e_1, \ldots, e_n)$ von V, so ist A^* die darstellende Matrix von f^* bzgl. $(e_1, \ldots, e_n)$, also*

$$\mathcal{M}_{\mathcal{A}}(f^*) = \mathcal{M}_{\mathcal{A}}(f)^*.$$

BEWEIS. 1. Existenz von f^*. Sei $y \in V$ fest. Betrachte

$$\begin{aligned} g\colon V &\to \mathbb{C} \quad (\text{bzw. } \mathbb{R}) \\ x &\mapsto \sigma(f(x), y). \end{aligned}$$

Es ist $g \in V^*$. Sei $(e_1, \ldots, e_n)$ eine Orthonormalbasis von V. Nach dem Darstellungssatz gilt

$$\sigma(f(x), y) = g(x) = \sigma(x, \sum_{j=1}^{n} \overline{g(e_j)}e_j) = \sigma(x, \sum_{j=1}^{n} \overline{\sigma(f(e_j), y)}e_j).$$

Setze

$$f^*(y) := \sum_{j=1}^{n} \overline{\sigma(f(e_j), y)}e_j = \sum_{j=1}^{n} \sigma(y, f(e_j))e_j.$$

Dann ist offenbar $f^* \in \mathrm{Hom}_{\mathbb{C}}(V, V)$, und für alle $x, y \in V$ gilt $\sigma(f(x), y) = \sigma(x, f^*(y))$.

Eindeutigkeit von f^*. $f^*(y)$ ist eindeutig bestimmt nach dem Darstellungssatz.

2. Sei $A = (\alpha_{ij}) \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) darstellende Matrix von f bzgl. der Orthonormalbasis $(e_1, \ldots, e_n)$, also

$$f(e_j) = \sum_{i=1}^{n} \alpha_{ij} e_i.$$

Dann gilt

$$f^*(e_j) = \sum_{s} \sigma(e_j, \underbrace{f(e_s)}_{\sum_i \alpha_{is} e_i})e_s = \sum_{s,i} \overline{\alpha_{is}}\,\underbrace{\sigma(e_j, e_i)}_{\delta_{ji}}\,e_s = \sum_{s} \overline{\alpha_{js}}e_s. \quad \text{QED}$$

Bemerkung. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum, $f, g \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$) und $\alpha \in \mathbb{C}$ (bzw. $\in \mathbb{R}$). Dann gilt

1. $(f+g)^* = f^* + g^*$.
2. $(f \circ g)^* = g^* \circ f^*$.
3. $(\alpha f)^* = \overline{\alpha} f^*$.
4. $f^{**} = f$.
5. Ist f ein Isomorphismus, so gilt $(f^{-1})^* = (f^*)^{-1}$.

Die Abbildung $f \mapsto f^*$ für $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$) hat also ähnliche Eigenschaften wie die Konjugation $\alpha \mapsto \overline{\alpha}$ für $\alpha \in \mathbb{C}$.

Beweis. 1. Für alle $x, y \in V$ gilt

$$\begin{aligned}\sigma((f+g)(x), y) &= \sigma(f(x) + g(x), y) \\ &= \sigma(f(x), y) + \sigma(g(x), y) \\ &= \sigma(x, f^*(y)) + \sigma(x, g^*(y)) \\ &= \sigma(x, f^*(y) + g^*(y)) \\ &= \sigma(x, (f^* + g^*)(y)).\end{aligned}$$

Also ist $(f+g)^* = f^* + g^*$.

2. Für alle $x, y \in V$ gilt

$$\sigma((f \circ g)(x), y) = \sigma(g(x), f^*(y)) = \sigma(x, (g^* \circ f^*)(y)).$$

Also ist $(f \circ g)^* = g^* \circ f^*$.

3. Für alle $x, y \in V$ gilt

$$\sigma(\alpha f(x), y) = \alpha\sigma(f(x), y) = \alpha\sigma(x, f^*(y)) = \sigma(x, \overline{\alpha} f^*(y)).$$

Also ist $(\alpha f)^* = \overline{\alpha} f^*$.

4. Zu zeigen ist

$$\sigma(f^*(x), y) = \sigma(x, f(y))$$

für alle $x, y \in V$, denn nach der Definition adjungierter Abbildungen folgt daraus die Behauptung. Es gilt aber

$$\sigma(f^*(x), y) = \overline{\sigma(y, f^*(x))} = \overline{\sigma(f(y), x)} = \sigma(x, f(y)).$$

5. Sei f ein Isomorphismus. Dann ist $f \circ f^{-1} = \mathrm{id}$, also wegen der Eindeutigkeit des adjungierter Endomorphismus auch $(f \circ f^{-1})^* = \mathrm{id}^* = \mathrm{id}$. Nach (2) ist aber $(f \circ f^{-1})^* = (f^{-1})^* \circ f^*$. Daraus folgt die Behauptung $(f^{-1})^* = (f^*)^{-1}$. Qed

Bemerkung. Teil (4) kann man auch leicht über darstellende Matrizen beweisen: Ist $A \in \mathbb{C}^{n \times n}$ darstellende Matrix von f bzgl. einer Orthonormalbasis $(e_1, \ldots, e_n)$, so ist nach Teil (2) des Satzes 8.3.5 über adjungierte Endomorphismen A^* die darstellende Matrix von f^* bzgl. derselben Orthonormalbasis $(e_1, \ldots, e_n)$, und auch $A^{**} = A$ die darstellende Matrix von f^{**} bzgl. $(e_1, \ldots, e_n)$. Also muß $f^{**} = f$ sein.

8.3.6. *Darstellende Matrizen von Endomorphismen.* Wir stellen einige Eigenschaften darstellender Matrizen von Endomorphismen zusammen, die wir im folgenden verwenden werden. Sei V ein n-dimensionaler K-Vektorraum und $\mathcal{A} = (x_1, \ldots, x_n)$ eine Basis von V. In 3.4.3 hatten wir gesehen, daß

$$\mathcal{M}_\mathcal{A} \colon \mathrm{Hom}_K(V,V) \to K^{n\times n}$$
$$\mathcal{M}_\mathcal{A}(f) = \text{die Matrix } A = (\alpha_{ij}) \text{ mit } f(x_j) = \sum_{i=1}^n \alpha_{ij} x_i \text{ für } 1 \le j \le n$$

ein Vektorraum-Isomorphismus ist. Nun trägt $\mathrm{Hom}_K(V,V) = \mathrm{End}_K(V)$ ebenso wie $K^{n\times n}$ sogar eine Ringstruktur: die (nicht kommutative) Multiplikation ist die Komposition $\circ$ zweier Endomorphismen, und die Identität id das Einselement. Wir wollen uns überlegen, daß der Vektorraum-Isomorphismus $\mathcal{M}_\mathcal{A}$ auch mit der Ringstruktur verträglich ist. Seien also $f, g \in \mathrm{End}_K(V)$ mit darstellenden Matrizen $A = (\alpha_{ij})$ und $B = (\beta_{js})$ bzgl. derselben Basis $\mathcal{A} = (x_1, \ldots, x_n)$ gegeben. Wir zeigen

$$\mathcal{M}_\mathcal{A}(f \circ g) = \mathcal{M}_\mathcal{A}(f) \circ \mathcal{M}_\mathcal{A}(g),$$
$$\mathcal{M}_\mathcal{A}(\mathrm{id}) = E.$$

Beweis. Die zweite Geichung gilt offensichtlich. Für die erste haben wir zunächst

$$f(x_j) = \sum_i \alpha_{ij} x_i \quad \text{und} \quad g(x_s) = \sum_j \beta_{js} x_j,$$

also

$$f(g(x_s)) = f(\sum_j \beta_{js} x_j) = \sum_j \beta_{js} f(x_j) = \sum_j \beta_{js} \sum_i \alpha_{ij} x_i = \sum_i \Big(\sum_j \alpha_{ij}\beta_{js}\Big) x_i. \quad \text{QED}$$

Ist also der Endomorphismus f oder die Matrix $\mathcal{M}_\mathcal{A}(f)$ invertierbar, so gilt

$$\mathcal{M}_\mathcal{A}(f^{-1}) = (\mathcal{M}_\mathcal{A}(f))^{-1}.$$

8.3.7. Satz. *(Unitäre Endomorphismen). Sei (V, σ) ein unitärer (bzw. euklidischer) Vektorraum, $\mathcal{A} = (e_1, \ldots, e_n)$ eine Orthonormalbasis von V und $A \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) die darstellende Matrix von f bzgl. $(e_1, \ldots, e_n)$. Dann sind folgende Aussagen äquivalent.*

1. *$f^* = f^{-1}$ (insbesondere also: f ist Isomorphismus).*
2. *$A^* = A^{-1}$ (insbesondere also: A ist invertierbar).*
3. *$\sigma(x, y) = \sigma(f(x), f(y))$ für alle $x, y \in V$.*
4. *$\|x\| = \|f(x)\|$ für alle $x \in V$.*
5. *$(f(e_1), \ldots, f(e_n))$ ist eine Orthonormalbasis von V.*

Beweis. $1 \Longrightarrow 2$. Sei f ein Isomorphismus und es gelte $f^* = f^{-1}$. Dann ist

$$\begin{aligned} \mathcal{M}_\mathcal{A}(f)^* &= \mathcal{M}_\mathcal{A}(f^*) && \text{nach Satz 8.3.5(2)} \\ &= \mathcal{M}_\mathcal{A}(f^{-1}) && \text{da } f^* = f^{-1} \\ &= \mathcal{M}_\mathcal{A}(f)^{-1} && \text{nach 8.3.6.} \end{aligned}$$

$2 \Longrightarrow 1$. Sei $A = \mathcal{M}_\mathcal{A}(f)$ invertierbar und es gelte $A^* = A^{-1}$. Dann ist

$$\begin{aligned} \mathcal{M}_\mathcal{A}(f^{-1}) &= \mathcal{M}_\mathcal{A}(f)^{-1} && \text{nach 8.3.6} \\ &= \mathcal{M}_\mathcal{A}(f)^* && \text{nach Voraussetzung} \\ &= \mathcal{M}_\mathcal{A}(f^*) && \text{nach Satz 8.3.5(2),} \end{aligned}$$

also $f^* = f^{-1}$.

$1 \equiv 3$. (3) ist äquivalent zu

$$\sigma(x,y) = \sigma(f(x), f(y)) = \sigma(x, (f^* \circ f)(y)) \quad \text{für alle } x, y \in V.$$

Nach dem Eindeutigkeitsteil des Darstellungssatzes 8.3.4 ist dies äquivalent zu $f^* \circ f = \mathrm{id}$, also zu (1).

$3 \equiv 4$. Dies folgt aus der Linearität von f und

$$\begin{aligned} \|x\| &= \sqrt{\sigma(x,x)} \\ \sigma(x,y) &= \frac{1}{4}\|x+y\|^2 - \frac{1}{4}\|x-y\|^2 + \frac{i}{4}\|x+iy\|^2 - \frac{i}{4}\|x-iy\|^2, \end{aligned}$$

was man durch Nachrechnen leicht verifiziert.

$3 \equiv 5$. Folgende Aussagen sind äquivalent.

$$\forall x, y \in V\, (\sigma(x,y) = \sigma(f(x), f(y)))$$
$$\forall s, t\, (\underbrace{\sigma(e_s, e_t)}_{\delta_{st}} = \sigma(f(e_s), f(e_t)))$$

$(f(e_1), \dots, f(e_n))$ ist Orthonormalsystem $\quad (f(e_1), \dots, f(e_n))$ ist Orthonormalbasis. QED

DEFINITION. Sei (V, σ) ein unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_\mathbb{C}(V,V)$ (bzw. $f \in \mathrm{Hom}_\mathbb{R}(V,V)$). f heißt *unitär* (bzw. *orthogonal*), wenn f eine Bedingung (und damit alle) des vorangehenden Satzes erfüllt, etwa $f^* = f^{-1}$. Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *unitär* (bzw. *orthogonal*), wenn $A^* = A^{-1}$. Folgende Bezeichnungen sind gebräuchlich.

$$\begin{aligned} \mathrm{U}(n) &:= \{ A \in \mathrm{GL}(n, \mathbb{C}) \mid \overline{A}^t A = E \} && \textit{unitäre Gruppe} \\ \mathrm{SU}(n) &:= \{ A \in \mathrm{U}(n) \mid \det(A) = +1 \} && \textit{spezielle unitäre Gruppe.} \\ \mathrm{O}(n) &:= \{ A \in \mathrm{GL}(n, \mathbb{R}) \mid A^t A = E \} && \textit{orthogonale Gruppe} \\ \mathrm{SO}(n) &:= \{ A \in \mathrm{O}(n) \mid \det(A) = +1 \} && \textit{spezielle orthogonale Gruppe} \end{aligned}$$

BEMERKUNG. Es gelten die Untergruppenbeziehungen

$$\begin{aligned} \mathrm{SO}(n) \subseteq \mathrm{O}(n) \subseteq \mathrm{GL}(n, \mathbb{R}), \\ \mathrm{SU}(n) \subseteq \mathrm{U}(n) \subseteq \mathrm{GL}(n, \mathbb{C}). \end{aligned}$$

8.3.8. *Selbstadjungierte Endomorphismen.* Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_\mathbb{C}(V,V)$ (bzw. $\in \mathrm{Hom}_\mathbb{R}(V,V)$). f heißt *selbstadjungiert*, wenn $f = f^*$ ist. Eine Matrix $A \in \mathbb{C}^{n \times n}$ (bzw. $\in \mathbb{R}^{n \times n}$) heißt *hermitesch* (bzw. *symmetrisch*), wenn $A^* = A$ ist, also $\overline{A}^t = A$ (bzw. $A^t = A$).

BEISPIEL. $\left(\begin{smallmatrix} 1 & i \\ -i & 0 \end{smallmatrix}\right)$ ist hermitesch.

SATZ. *Sei (V, σ) ein endlichdimensionaler unitärer Vektorraum, $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ und $\lambda \in \mathbb{C}$ ein Eigenwert von f. Dann gilt*

1. *Ist f selbstadjungiert, so ist $\lambda \in \mathbb{R}$.*
2. *Ist f unitär, so ist $|\lambda| = 1$.*

BEWEIS. Sei $x \in V$, $x \neq 0$ und $f(x) = \lambda x$. Dann erhält man folgendes. 1.

$$\begin{aligned}
&f = f^* \\
&\sigma(f(x), x) = \sigma(x, f(x)) \\
&\sigma(\lambda x, x) = \sigma(x, \lambda x) \\
&\lambda\sigma(x, x) = \overline{\lambda}\sigma(x, x) \\
&\lambda = \overline{\lambda}, && \text{da } x \neq 0.
\end{aligned}$$

2.

$$\begin{aligned}
&\sigma(f(x), f(x)) = \sigma(x, x) \\
&\sigma(\lambda x, \lambda x) = \sigma(x, x) \\
&\lambda\overline{\lambda}\sigma(x, x) = \sigma(x, x) \\
&\lambda\overline{\lambda} = 1, && \text{da } x \neq 0 \\
&|\lambda| = 1. && \text{QED}
\end{aligned}$$

8.3.9. *Diagonalisierung.* Im nächsten Kapitel 9 wollen wir uns mit dem Problem befassen, wann ein Endomorphismus f eines gegebenen endlichdimensionalen unitären (bzw. euklidischen) Vektorraums (V, σ) diagonalisierbar ist, d.h. eine Orthonormalbasis aus Eigenvektoren von f besitzt. Dafür zunächst einige vorbereitende Anmerkungen.

BEMERKUNGEN. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum.

1. Seien $(e_1, \ldots, e_n)$ und $(e'_1, \ldots, e'_n)$ Orthonormalbasen von V. Nach dem Satz 8.3.7 über orthogonale Endomorphismen (insbesondere der Äquivalenz von (2) und (5)) gilt

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} \quad \text{für ein } C \in \mathrm{U}(n) \text{ (bzw. } \in \mathrm{O}(n)).$$

2. Seien $A, B \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$). Dann sind folgende Aussagen äquivalent.

- A und B sind darstellende Matrizen desselben $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$) f bezüglich evtl. verschiedener Orthonormalbasen.
- $B = C^{-1}AC$ für ein $C \in \mathrm{U}(n)$ (bzw. $\in \mathrm{O}(n)$), d.h. A und B sind *unitär* (bzw. *orthogonal*) *äquivalent*.

BEWEIS. Seien A, B darstellende Matrizen desselben $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$) bezüglich der Orthonormalbasen $(e_1, \ldots, e_n)$ und $(e'_1, \ldots, e'_n)$. Dann gilt

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{und} \quad f\begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = B^t \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

Nach (1) hat man ein $C \in U(n)$ mit

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

Es folgt

$$A^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = Cf \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = CB^t \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = CB^tC^{-1} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Also ist

$$\begin{aligned} A^t &= CB^tC^{-1} \\ B^t &= C^{-1}A^tC \\ B &= C^tA(C^{-1})^t = D^{-1}AD \quad \text{mit } D = (C^t)^{-1} \in U(n) \text{ (bzw. } \in \mathrm{O}(n)). \end{aligned}$$

Sei umgekehrt$B = C^{-1}AC$ für ein $C \in \mathrm{U}(n)$ (bzw. $\in \mathrm{O}(n)$). Man wähle eine Orthonormalbasis $(e'_1, \ldots, e'_n)$ und setze

$$\begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = C^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Dann gilt für den durch

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

definierten Endomorphismus

$$f \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = C^tf \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^tA^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^tA^t(C^t)^{-1} \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = (C^{-1}AC)^t \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

Also sind A und B darstellende Matrizen von f bezüglich verschiedener Orthonormalbasen. QED

3. Sei A darstellende Matrix von f bezüglich einer Orthonormalbasis. Dann sind folgende Aussagen äquivalent.

- V hat eine Orthonormalbasis aus Eigenvektoren von f.
-

$$C^{-1}AC = \begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix} \quad \text{für ein } C \in \mathrm{U}(n) \text{ (bzw. } \in \mathrm{O}(n)), \text{ mit } \gamma_j \in \mathbb{C} \text{ (bzw. } \in \mathbb{R}).$$

Beweis. Dies folgt aus (1), da die darstellende Matrix bezüglich einer Orthonormalbasis aus Eigenvektoren eine Diagonalmatrix ist. QED

4. Nach der Hauptachsentransformation in 5.4 gilt für jeden Endomorphismus f eines endlichdimensionalen euklidischen Vektorraums V:

Ist f selbstadjungiert, so hat V eine Orthonormalbasis aus Eigenvektoren.

Für Matrizen bedeutet dies:

Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, so gibt es $C \in \mathrm{O}(n)$, $\gamma_j \in \mathbb{R}$ mit $\quad C^{-1}AC = \begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix}.$

Der damalige Beweis verwendete JACOBIsche Rotationsmatrizen und etwas Analysis. Im folgenden werden wir einen neuen Beweis dieses Satzes geben, der die Komplexifizierung und den Fundamentalsatz der Algebra benutzt. Gleichzeitig wird sich dabei eine Verallgemeinerung auf unitäre Räume ergeben.

Kapitel 9

Endomorphismen

Wir wollen in diesem Kapitel die Endomorphismen von endlichdimensionalen unitären (bzw. euklidischen) Vektorräumen genauer studieren. Als erstes erhalten wir eine Charakterisierung derjenigen Endomorphismen, die eine Orthonormalbasis aus Eigenvektoren besitzen. In Abschnitt 9.2 befassen wir uns mit der Zerlegung eines Endomorphismus in geometrisch einfach zu beschreibende Bestandteile (Drehungen oder Spiegelungen, Streckungen und orthogonale Projektionen). Es wird sich zeigen, daß eine solche Zerlegung immer eindeutig möglich ist. Im letzten Abschnitt 9.3 verschaffen wir uns eine Übersicht über die orthogonalen Endomorphismen eines endlichdimensionalen euklidischen Vektorraums.

9.1 Normale Endomorphismen

Wir geben jetzt eine Charakterisierung der diagonalisierbaren Endomorphismen f eines endlichdimensionalen unitären (bzw. euklidischen) Vektorraums (V, σ); ein Endomorphismus f heißt diagonalisierbar, wenn (V, σ) eine Orthonormalbasis aus Eigenvektoren von f besitzt. Dazu benötigen wir einen etwas technischen Begriff, nämlich den eines normalen Endomorphismus. Wir werden zeigen:

- In einem unitären Vektorraum ist ein Endomorphismus genau dann diagonalisierbar, wenn er *normal* ist.
- In einem euklidischen Vektorraum ist ein Endomorphismus genau dann diagonalisierbar, wenn er *selbstadjungiert* ist.

9.1.1. *Normale Endomorphismen.* Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$). f heißt *normal*, wenn

$$f^* \circ f = f \circ f^*.$$

Eine Matrix $A \in \mathbb{C}^{n \times n}$ (bzw. $\in \mathbb{R}^{n \times n}$) heißt *normal*, wenn $A^* A = A A^*$ ist.

Bemerkungen. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ (bzw. $\in \mathrm{Hom}_{\mathbb{R}}(V, V)$).

1. Sei A darstellende Matrix von f bzgl. einer Orthonormalbasis. Dann gilt: f ist normal (bzw. selbstadjungiert) genau dann, wenn A normal (bzw. symmetrisch) ist. Dies folgt aus dem Satz 8.3.5(2) über adjungierte Endomorphismen.
2. Ist f selbstadjungiert, so ist f auch normal, denn aus $f = f^*$ folgt offenbar $f^* \circ f = f \circ f^*$.

3. Ist f unitär (bzw. orthogonal), so ist f auch normal, denn aus $f^* = f^{-1}$ folgt wieder $f^* \circ f = f \circ f^*$.

4. Seien $U_i \subseteq V$ Unterräume mit $U_1 \perp U_2$ und $V = U_1 \oplus U_2$. Wir setzten voraus, daß $f(U_i) \subseteq U_i$ für $i = 1, 2$. Dann gilt:

f ist normal	genau dann, wenn $f\restriction U_i \colon U_i \to U_i$ für $i = 1, 2$ normal sind;
f ist selbstadjungiert	genau dann, wenn $f\restriction U_i \colon U_i \to U_i$ für $i = 1, 2$ selbstadjungiert sind.

Dies sieht man wie folgt. Sei $e_1, \ldots, e_m$ Orthonormalbasis von U_1 und $e_{m+1}, \ldots, e_n$ Orthonormalbasis von U_2. Dann ist $e_1, \ldots, e_n$ Orthonormalbasis von $V = U_1 \oplus U_2$, da $U_1 \perp U_2$. Sei A_i die transponierte darstellende Matrix von $f\restriction U_i$ bzgl. obiger Orthonormalbasis von U_i. Dann ist

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

und damit

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

die transponierte darstellende Matrix von f bzgl. $e_1, \ldots, e_n$. Weiter gilt

$$A^* = \begin{pmatrix} A_1^* & 0 \\ 0 & A_2^* \end{pmatrix}.$$

Dann sind folgende Aussagen äquivalent.

f ist normal	
A ist normal	nach Bemerkung 1
$AA^* = A^*A$	
$A_iA_i^* = A_i^*A_i$	für $i = 1, 2$
$f\restriction U_i$ sind normal	für $i = 1, 2$.

Entsprechend schließt man für selbstadjungierte Endomorphismen.

Satz. *Sei (V, σ) ein endlichdimensionaler K-Vektorraum mit $K \in \{\mathbb{R}, \mathbb{C}\}$ und $\sigma \colon V \times V \to K$ ein Skalarprodukt. Sei $f \in \mathrm{Hom}_K(V, V)$. Das charakteristische Polynom f zerfalle über K in Linearfaktoren. Dann sind folgende Aussagen äquivalent.*

1. *f ist normal.*

2. *V hat eine Orthonormalbasis aus Eigenvektoren von f.*

Beweis. $1 \Longrightarrow 2$. Für $\lambda \in K$ sei

$$V_\lambda := \{\, x \in V \mid f(x) = \lambda x \,\}$$

der Eigenraum von f zu λ.

Schritt a.

$$\text{Ist } g \in \mathrm{Hom}_K(V, V) \text{ und } U \subseteq V \text{ ein Unterraum mit } g(U) \subseteq U, \text{ so ist } g^*(U^\perp) \subseteq U^\perp; \tag{9.1}$$

$$\text{Ist } \lambda \in K \text{ und } g \in \mathrm{Hom}_K(V, V) \text{ mit } f \circ g = g \circ f, \text{ so ist } g(V_\lambda) \subseteq V_\lambda. \tag{9.2}$$

(9.1) gilt, da für alle $x \in U$ und $y \in U^\perp$

$$\sigma(\underbrace{g(x)}_{\in U}, y) = \sigma(x, g^*(y)) = 0, \quad \text{also } g^*(y) \in U^\perp \text{ für } y \in U^\perp.$$

(9.2) gilt, da man für alle $x \in V_\lambda$ hat

$$\begin{aligned}
&f(x) = \lambda x \\
&f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x) \quad \text{wegen } f \circ g = g \circ f \\
&g(x) \in V_\lambda.
\end{aligned}$$

Schritt b. Ist f normal, so gilt $f(V_\lambda^\perp) \subseteq V_\lambda^\perp$ für alle $\lambda \in K$, denn man hat

$$\begin{aligned}
f^*(V_\lambda) \subseteq V_\lambda \quad & \text{nach (9.2)} \\
f^{**}(V_\lambda^\perp) \subseteq V_\lambda^\perp \quad & \text{nach (9.1)} \\
f(V_\lambda^\perp) \subseteq V_\lambda^\perp \quad & \text{wegen } f = f^{**}.
\end{aligned}$$

Schritt c. Wir zeigen $1 \Longrightarrow 2$ durch Induktion über $\dim(V)$.

Fall $\dim(V) = 1$. Man wähle ein $e \in V$ mit $\|e\| = 1$. Dann ist (e) Basis von V, und e ist Eigenvektor von f wegen $f(e) \in Ke = V$.

Fall $\dim(V) > 1$. Da das charakteristische Polynom von f über K in Linearfaktoren zerfällt, hat f einen Eigenwert, etwa $\lambda \in K$. Es gilt $V = V_\lambda \oplus V_\lambda^\perp$, und wegen $\dim(V_\lambda) > 0$ folgt $\dim(V_\lambda^\perp) < \dim(V)$. V_λ hat trivialerweise ein Orthonormalbasis aus Eigenvektoren von f, da ja V_λ Eigenraum ist.

Weiter hat $V_\lambda^\perp$ eine Orthonormalbasis aus Eigenvektoren von f; dies sieht man wie folgt. Offenbar ist $f(V_\lambda) \subseteq V_\lambda$ und nach Schritt (b) auch $f(V_\lambda^\perp) \subseteq V_\lambda^\perp$. Nach der obigen Bemerkung (4) ist dann auch $f \restriction V_\lambda^\perp$ normal. Ferner teilt nach Bemerkung (3) in 7.2.3 das charakteristische Polynom von $f \restriction V_\lambda^\perp$ dasjenige von f. Da nach Voraussetzung p_f in Linearfaktoren zerfällt, muß dies also auch für $p_{f \restriction V_\lambda^\perp}$ gelten (dies zeigt man leicht allgemein mit Hilfe von Lemma 7.1.3, und zwar durch Induktion über den Grad). Nach Induktionsvoraussetzung hat also $V_\lambda^\perp$ eine Orthonormalbasis aus Eigenvektoren von $f \restriction V_\lambda^\perp$ (und damit von f). Wegen $V = V_\lambda \oplus V_\lambda^\perp$ hat also auch V eine Orthonormalbasis aus Eigenvektoren von f.

$2 \Longrightarrow 1$. Sei $e_1, \ldots, e_n$ eine Orthonormalbasis aus Eigenvektoren von f. Dann gilt

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

A ist normal, da

$$A^* = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix}$$

und deshalb $AA^* = A^*A$. Also ist f normal nach Bemerkung (1).

FOLGERUNG. *Sei (V, σ) ein endlichdimensionaler unitärer Vektorraum und $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$. Dann sind folgende Aussagen äquivalent.*

1. *f ist normal.*

2. *V hat eine Orthonormalbasis aus Eigenvektoren von f.*

BEWEIS. Nach dem Fundamentalsatz der Algebra zerfällt das charakteristische Polynom f über $\mathbb{C}$ in Linearfaktoren. Also ergibt sich aus dem obigen Satz die Behauptung. QED

9.1.2. *Hauptachsentransformation.* Als Folgerung aus dem eben bewiesenen Satz ergibt sich jetzt leicht ein neuer Beweis der Hauptachsentransformation für euklidische Vektorräume. Er beruht auf dem folgenden einfachen Lemma.

LEMMA. *Sei (V, σ) endlichdimensionaler euklidischer Vektorraum und $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ selbstadjungiert. Dann zerfällt das charakteristische Polynom von f über $\mathbb{R}$ in Linearfaktoren.*

BEWEIS. Sei A die darstellende Matrix von f bzgl. irgendeiner Orthonormalbasis von V. Dann ist A symmetrisch, nach Bemerkung (1). Faßt man A als $\mathbb{C}^{n\times n}$-Matrix auf, so ist A hermitesch und damit hat (nach Satz 8.3.8) das charakteristische Polynom $\det(A - tE)$ nur reelle Nullstellen über $\mathbb{C}$ (denn A ist darstellende Matrix eines selbstadjungierten Endomorphismus). Andererseits zerfällt p_f nach dem Fundamentalsatz der Algebra über $\mathbb{C}$ in Linearfaktoren. Insgesamt hat man, daß p_f über $\mathbb{R}$ in Linearfaktoren zerfällt. QED

FOLGERUNG. *(Hauptachsentransformation). Sei (V, σ) endlichdimensionaler euklidischer Vektorraum und $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$. Dann sind folgende Aussagen äquivalent.*

1. *f ist selbstadjungiert.*

2. *V hat eine Orthonormalbasis aus Eigenvektoren von f.*

BEWEIS. $1 \Longrightarrow 2$. Die Behauptung folgt sofort aus dem Satz 9.1.1 zusammen mit obigem Lemma; man braucht noch, daß jeder selbstadjungierte Endomorphismus normal ist.

$2 \Longrightarrow 1$. Sei $e_1, \ldots, e_n$ eine Orthonormalbasis aus Eigenvektoren von f. Dann gilt

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ und } \lambda_j \in \mathbb{R}.$$

Also ist $A = A^t$. QED

9.2 Zerlegung von Endomorphismen; geometrische Deutung

9.2.1. *Charakterisierung orthogonaler Projektionen.* Sei V ein K-Vektorraum und f ein Endomorphismus. f heißt *Projektion*, wenn es Unterräume $U, W \subseteq V$ gibt mit $V = U \oplus W$, so daß für alle $u \in U$ und $w \in W$ gilt: $f(u + w) = u$. Man nennt f *Parallelprojektion von U in Richtung W.*

Sei (V, σ) endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus. f heißt *orthogonale Projektion*, wenn es einen Unterraum $U \subseteq V$ gibt, so daß für alle $p \in U$ und $h \in U^\perp$ gilt $f(p + h) = p$.

LEMMA. *Sei (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus. Folgende Aussagen sind äquivalent.*

1. *f ist orthogonale Projektion.*
2. *f ist selbstadjungiert und es gilt $f^2 = f$.*

BEWEIS. $1 \Longrightarrow 2$. Sei $V = U \oplus U^\perp$ und f orthogonale Projektion auf U. Sei $e_1, \dots, e_m$ Orthonormalbasis von U und $e_{m+1}, \dots, e_n$ Orthonormalbasis von $U^\perp$. Dann ist $e_1, \dots, e_n$ Orthonormalbasis von V, und es gilt

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}.$$

Daraus folgt $f^2 = f$ sowie die Selbstadjungiertheit von f.

$2 \Longrightarrow 1$. Sei f selbstadjungiert mit $f^2 = f$. Nach der Folgerung in 9.1.1 (bzw. in 9.1.2) hat V eine Orthonormalbasis $e_1, \dots, e_n$ aus Eigenvektoren von f. Sei $x \neq 0$ Eigenvektor von f zum Eigenwert λ. Dann gilt

$$\begin{aligned} f(x) &= \lambda x \\ f(x) &= f^2(x) = \lambda f(x) = \lambda^2 x, \quad \text{wegen } f^2 = f \end{aligned}$$

also

$$\begin{aligned} \lambda &= \lambda^2 && \text{da } x \neq 0 \\ 0 &= \lambda(\lambda - 1) \\ \lambda &= 0 \text{ oder } \lambda = 1. \end{aligned}$$

Seien $e_1, \dots, e_m$ die e_j mit $\lambda_j = 1$. Dann

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}.$$

Also ist f orthogonale Projektion auf $\langle e_1, \dots, e_m \rangle$ (insbesondere $f = 0$, falls alle $\lambda_j = 0$). QED

9.2.2. *Streckungen, Drehungen und Spiegelungen.* Sei (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus von V. f heißt *Streckung* (oder *Dehnung*) genau dann, wenn f selbstadjungiert ist und alle Eigenwerte von f größer als 0 sind.

Bemerkung. Sei f eine Streckung. Dann gibt es nach der Folgerung in 9.1.1 (bzw. in 9.1.2) eine Orthonormalbasis $e_1, \ldots, e_n$ von V aus Eigenvektoren von f und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit $\lambda_i > 0$ für $1 \leq i \leq n$, so daß

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

f bewirkt also in der Richtung e_j eine Streckung, falls $\lambda_j \geq 1$, oder eine Stauchung, falls $\lambda_j \leq 1$ ist.

Definition. Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum und f ein Endomorphismus.

1. f heißt *Drehung*, falls f orthogonal und $\det(f) = 1$ ist.
2. f heißt *Spiegelung*, falls f orthogonal und $\det(f) = -1$ ist.

Bemerkungen. 1. Ist f orthogonal, so ist $\det(f) = +1$ oder $\det(f) = -1$. Dies sieht man wie folgt. Sei A die darstellende Matrix von f bzgl. der Orthonormalbasis $e_1, \ldots, e_n$. Dann gilt:

$$\begin{aligned} & A^* = A^t = A^{-1} && \text{da } f \text{ orthogonal} \\ & AA^t = E \\ & \det(AA^t) = \det(A) \cdot \underbrace{\det(A^t)}_{=\det(A)} = \det(E) = 1 \\ & \det(A) = 1 \text{ oder } \det(A) = -1. \end{aligned}$$

2. Sei (V, σ) endlichdimensionaler euklidischer Vektorraum, $U \subseteq V$ Untervektorraum und $V = U \oplus U^\perp$. Dann definiere

$$\begin{aligned} f\colon V &\to V \\ x = p + h &\mapsto p - h \quad \text{mit } p \in U \text{ und } h \in U^\perp. \end{aligned}$$

Sei etwa $e_1, \ldots, e_m$ eine Orthonormalbasis von U und $e_{m+1}, \ldots, e_n$ eine Orthonormalbasis von $U^\perp$. Dann gilt

$$f\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \\ & \underbrace{\qquad\qquad}_{m} & & & \underbrace{\qquad\qquad}_{n-m} & \end{pmatrix}}_{=:A}\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix},$$

also $\det(A) = (-1)^{n-m}$. Demnach ist f Drehung, falls $n - m$ gerade ist, und Spiegelung, falls $n - m$ ungerade ist.

Ist zum Beispiel $V = \mathbb{R}^3$ und $U = \mathbb{R}e_1$, so folgt $U^\perp = \langle e_2, e_3 \rangle$ und wir haben eine Drehung um $\mathbb{R}e_1$. Ist $U = \langle e_1, e_2 \rangle$, so folgt $U^\perp = \mathbb{R}e_3$. und wir haben eine Spiegelung an der Ebene $\langle e_1, e_2 \rangle$.

Im Spezialfall, daß U eine Hyperebene ist (d.h. $m := \dim(U) = n-1$), heißt f *Hyperebenenspiegelung*.

3. Normalformen für Drehungen und Spiegelungen folgen im nächsten Abschnitt. Beispiel: Sei V zweidimensional, e_1, e_2 eine Orthonormalbasis von V und

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}}_{A} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad \text{mit } 0 \leq \varphi < 2\pi.$$

Dann gilt $AA^t = E$, also ist f orthogonal. Es gilt $\det(A) = \cos^2\varphi + \sin^2\varphi = 1$, also ist f eine Drehung. Es handelt sich um eine Drehung gegen den Uhrzeigersinn um den Winkel φ.

9.2.3. *Zerlegung von Endomorphismen.* Wir wollen uns jetzt überlegen, daß man in einem endlichdimensionalen euklidischen Vektorraum jeden Endomorphismus darstellen kann als

- Hintereinanderausführung einer orthogonalen Projektion, einer Streckung und einer anschließenden Drehung oder Spiegelung, oder auch als
- Hintereinanderausführung einer Drehung oder Spiegelung, einer Streckung und einer anschließenden orthogonalen Projektion.

Für unitäre Vektorräume gilt eine entsprechende Aussage.

SATZ. *(Zerlegung von Endomorphismen). Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus von V. Dann gibt es Endomorphismen p, p', s, s', d, d' mit p, p' orthogonale Projektionen, s, s' selbstadjungiert mit reellen, positiven Eigenwerten und d, d' unitär (bzw. orthogonal), so daß*

$$\begin{aligned} f &= d \circ s \circ p \\ &= p' \circ s' \circ d'. \end{aligned}$$

BEWEIS. $f^* \circ f$ ist selbstadjungiert, da $(f^* \circ f)^* = f^* \circ f^{**} = f^* \circ f$. Mit der Folgerung in 9.1.1 (bzw. in 9.1.2) erhält man eine Orthonormalbasis $e_1, \ldots, e_n$ von V aus Eigenvektoren von $f^* \circ f$. Es gilt also

$$(f^* \circ f)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } \lambda_i \in \mathbb{R}.$$

OBdA seien $\lambda_1, \ldots, \lambda_m \neq 0$ und $\lambda_{m+1}, \ldots, \lambda_n = 0$ für $0 \leq m \leq n$. Für $1 \leq i, j \leq m$ gilt

$$\sigma(f(e_i), f(e_j)) = \sigma(e_i, \underbrace{(f^* \circ f)(e_j)}_{=\lambda_j e_j}) = \lambda_j \delta_{ij},$$

also

$$\begin{aligned} &f(e_i) \neq 0 && \text{für } 1 \leq i \leq m \\ &f(e_i) \perp f(e_j) && \text{für } 1 \leq i, j \leq m,\ i \neq j. \end{aligned}$$

Nun normiere man die $f(e_i)$ für $1 \leq i \leq m$; man erhält

$$e_i' := \frac{f(e_i)}{\|f(e_i)\|} \qquad \text{für } 1 \le i \le m.$$

Also ist $e_1', \dots, e_m'$ ein Orthonormalsystem. Nach dem SCHMIDTschen Orthonormalisierungsverfahren existiert eine Orthonormalbasis $e_1', \dots, e_m', e_{m+1}', \dots e_n'$ von V. Damit erhält man eine Matrix $C \in \mathbb{C}^{n\times n}$ (bzw. $\in \mathbb{R}^{n\times n}$) mit $C^* = C^{-1}$ und

$$\begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} = C \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Es gilt $f(e_i) = \gamma_i e_i'$ mit $\gamma_i := \|f(e_i)\| \in \mathbb{R}$ sogar für $1 \le i \le n$ (mit $\gamma_{m+1} = \dots = \gamma_n = 0$). Also

$$\begin{aligned}
f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} &= \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix} \\
&= \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} C \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}}_{=:A} C \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}
\end{aligned}$$

Es seien d, s, p wie folgt definiert.

$$d \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

$$s \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}$$

$$p \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ \multicolumn{3}{c}{\underbrace{\hphantom{1 \ddots 1}}_{m}} & \multicolumn{3}{c}{\underbrace{\hphantom{0 \ddots 0}}_{n-m}} \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}$$

Also ist d unitär (orthogonal), s selbstadjungiert mit positiven Eigenwerten und p die orthogonale Projektion auf $\langle e_1, \dots, e_m \rangle$. Es folgt $f = d \circ s \circ p$, denn es gilt:

$$
\begin{aligned}
(d \circ s \circ p)\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} &= (d \circ s)(\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}) \\
&= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} d(s\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}) \\
&= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} d(\begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}) \\
&= \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}\begin{pmatrix} \gamma_1 & & & & & \\ & \ddots & & & & \\ & & \gamma_m & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}\underbrace{d\begin{pmatrix} e_1 \\ \vdots \\ e_m \\ e_{m+1} \\ \vdots \\ e_n \end{pmatrix}}_{C\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}} \\
&= f\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.
\end{aligned}
$$

Die zweite Zerlegung ergibt sich durch Anwendung der ersten auf f^* anstelle von f.

$$
\begin{aligned}
f^* &= d \circ s \circ p \\
f = f^{**} &= \underbrace{p^*}_{=p} \circ \underbrace{s^*}_{=s} \circ d^*.
\end{aligned}
$$

Es genügt also zu zeigen, daß d^* unitär (bzw. orthogonal) ist; dies folgt aber aus $d^{**} = (d^{-1})^* = (d^*)^{-1}$. Qed

BEMERKUNGEN. 1. Sei (V, σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und f ein Endomorphismus von V. Dann gibt es

g	selbstadjungierter Endomorphismus mit Eigenwerten ≥ 0
d	unitärer (bzw. orthogonaler) Endomorphismus

so daß

$$f = d \circ g \quad \text{(siehe vorheriger Satz; setze } g := s \circ p\text{)}$$

Dabei ist g *eindeutig bestimmt.* (Ebenso für $f = g' \circ d'$ mit g' selbstadjungiert und alle Eigenwerte von $g' \geq 0$, sowie d' unitär bzw. orthogonal).

BEWEIS. Seien g_1, g selbstadjungiert, alle ihre Eigenwerte ≥ 0 und d_1, d unitär (bzw. orthogonal), so daß $f = d \circ g = d_1 \circ g_1$ ist. Man erhält:

$$\begin{aligned} f^* \circ f &= \underbrace{g^*}_{=g} \circ \underbrace{d^*}_{=d^{-1}} \circ d \circ g = g^2 \\ &= g_1^* \circ d_1^* \circ d_1 \circ g_1 = g_1^2. \end{aligned}$$

Da g_1 selbstadjungiert ist, existiert nach der Folgerung in 9.1.1 (bzw. in 9.1.2) eine Orthonormalbasis $e_1, \ldots, e_n$ von V aus Eigenvektoren von g_1.

$$g_1 \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } \lambda_j \geq 0 \text{ für } 1 \leq j \leq n.$$

Also gilt

$$g_1^2(e_j) = \lambda_j^2 e_j = g^2(e_j)$$

Nach (9.3) (siehe unten) folgt somit

$$\begin{aligned} &g_1(e_j) = \lambda_j e_j = g(e_j) \qquad \text{für } 1 \leq j \leq n, \\ &g = g_1 \end{aligned}$$

Zu zeigen bleibt, daß für jeden selbstadjungierten Endomorphismus g, dessen sämtliche Eigenwerte ≥ 0 sind, jedes $x \in V$ mit $x \neq 0$ und jedes $\lambda \in \mathbb{R}$ mit $\lambda \geq 0$ gilt

$$g(x) = \lambda x \quad \text{genau dann, wenn} \quad g^2(x) = \lambda^2 x. \tag{9.3}$$

Dies beweisen wir wie folgt. $\Longrightarrow$. Klar. $\Longleftarrow$. Sei $e_1, \ldots, e_n$ eine Orthonormalbasis von V aus Eigenvektoren von g, also

$$g \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \quad \text{mit } \lambda_j \geq 0 \text{ für } 1 \leq j \leq n.$$

Sei $x = \sum_{j=1}^{n} \alpha_j e_j$. Dann gilt

$$\begin{aligned}
&g^2(x) = \lambda^2 x \\
&\sum_{j=1}^{n} \alpha_j \lambda_j^2 e_j = \sum_{j=1}^{n} \alpha_j \lambda^2 e_j \\
&\alpha_j \lambda_j^2 = \alpha_j \lambda^2 && \text{für alle } 1 \le j \le n \\
&\lambda_j = \lambda && \text{für alle } 1 \le j \le n \text{ mit } \alpha_j \ne 0, \text{ da } \lambda_j, \lambda \ge 0.
\end{aligned}$$

Also ergibt sich

$$\begin{aligned}
&\sum_{j=1}^{n} \alpha_j \lambda_j e_j = \sum_{j=1}^{n} \alpha_j \lambda e_j \\
&g(x) = \lambda x. && \text{QED}
\end{aligned}$$

2. Ist in (1) speziell f ein Isomorphismus, so sind die nach (1) existierenden Endomorphismen d, g mit $f = d \circ g$ offenbar eindeutig bestimmt.

9.2.4. *Praktische Durchführung der Hauptachsentransformation und der Zerlegung von Endomorphismen.* 1. Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum, f ein selbstadjungierter Endomorphismus, $e_1, \ldots, e_n$ eine Orthonormalbasis von V und $A \in \mathbb{R}^{n \times n}$ mit

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Wir betrachten das folgende Problem.

$$\text{Gesucht ist eine Orthonormalbasis } e'_1, \ldots, e'_n \text{ von } V \text{ aus Eigenvektoren von } f. \tag{9.4}$$

Ein äquivalentes Problem ist

$$\text{Gesucht ist eine Matrix } C \in O(n) \text{ und } \gamma_1, \ldots, \gamma_n \in \mathbb{R} \text{ mit } C^t A C = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}. \tag{9.5}$$

Wir erinnern noch einmal an den Beweis der Äquivalenz: (9.4)$\Longrightarrow$(9.5). Sei $e'_1, \ldots, e'_n$ eine Orthonormalbasis von V aus Eigenvektoren von f, also

$$f \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}.$$

Es gibt ein $C \in O(n)$ mit

$$\begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = C^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Also:

$$f\begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = C^t f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t AC \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}$$

$$C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

(9.5)$\Longrightarrow$(9.4): Sei $C \in O(n)$ gegeben mit $C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$. Setze $\begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} := C^t \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$.

Dann gilt

$$f\begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = C^t f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t A \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^t AC \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix} \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}. \quad \text{QED}$$

Nun wollen wir eine Lösung für das Problem (9.4) bzw. (9.5) konstruieren:

Schritt 1. Bestimme alle Eigenwerte von A als Nullstellen von $\det(A - tE)$. (Dies ist insbesondere für $n = 2, 3$ möglich).

Schritt 2. Bestimme zu jedem Eigenwert λ von A eine Basis des Eigenraums

$$U_\lambda = \{\, x \in \mathbb{R}^n \mid Ax = \lambda x \,\}.$$

Dies entspricht der Lösung eines homogenen Gleichungssystems mit der Matrix $A - \lambda E$.

Schritt 3. Bestimme nach dem SCHMIDTschen Orthonormalisierungsverfahren eine Orthonormalbasis in jedem Eigenraum U_λ. Sei C die Matrix aus $\mathbb{R}^{n\times n}$, deren Spalten die Orthonormalbasen der U_λ sind, λ Eigenwert von A. Dann gilt:

a. $C \in O(n)$

b. $C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$ mit γ_i Eigenwert von A (wobei mehrfache Eigenwerte entsprechend mehrfach aufgeschrieben sind).

BEWEIS. a. Wir zeigen $C \in O(n)$. Nach Bemerkung (2) (s.u.) gilt $\mathbb{R}^n = \bigoplus U_\lambda$, λ Eigenwert von A, und $U_\lambda \perp U_\gamma$ für $\lambda \neq \gamma$. (Denn U_λ ist Eigenraum zum Eigenwert λ der Abbildung $\mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto Ax$). Also bilden die Spalten von C eine Orthonormalbasis von $\mathbb{R}^n$ und somit ist $C \in O(n)$.

b. Wir zeigen, daß $C^t AC$ Diagonalmatrix ist. Folgende Aussagen sind äquivalent.

$$C^t AC = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$$

$$AC = C \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}$$

$$Ac_i = c_i \gamma_i \qquad (c_i \ i\text{-te Spalte von } C) \text{ für } 1 \le i \le n.$$

Solche γ_i existieren nach Konstruktion von C. QED

2. Umformulierung der Folgerung in 9.1.1 (bzw. in 9.1.2). Seien (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum, f ein Endomorphismus von V und $\lambda_1,\dots,\lambda_m$ die verschiedenen Eigenwerte von f. (Also $\lambda_i \neq \lambda_j$ für $i \neq j$). Sei

$$V_\lambda = \{\, x \in V \mid f(x) = \lambda x \,\}$$

der Eigenraum von f zum Eigenwert λ. Dann ist f genau dann normal (bzw. selbstadjungiert), wenn $V = \bigoplus_{i=1}^m V_{\lambda_i}$ und $V_{\lambda_i} \perp V_{\lambda_j}$ für $i \neq j$ ist.

Beweis. $\Longleftarrow$. V hat dann eine Orthonormalbasis aus Eigenvektoren von f. $\Longrightarrow$. Sei $e_1,\dots,e_n$ eine Orthonormalbasis von V aus Eigenvektoren von f und seien Räume V_i definiert durch $V_i := \langle\{\, e_j \mid e_j \text{ Eigenvektor zum Eigenwert } \lambda_i \,\}\rangle$. Dann ist $V = \bigoplus_{i=1}^m V_i$ und $V_i \perp V_j$ für $i \neq j$. Außerdem gilt:

$$\begin{aligned}
&V_i \subseteq V_{\lambda_i}\\
&\dim(V_i) \leq \dim(V_{\lambda_i})\\
&\dim(V) = \sum_{i=1}^m \dim(V_i) \leq \sum_{i=1}^m \dim(V_{\lambda_i}) \leq \dim(V) \quad \text{nach dem Satz in 7.2.2}\\
&\dim(V_i) = \dim(V_{\lambda_i})\\
&V_i = V_{\lambda_i}. \qquad\qquad\qquad\qquad\qquad\qquad \text{QED}
\end{aligned}$$

3. Analog zu (1) berechnet man eine Orthonormalbasis aus Eigenvektoren eines normalen Endomorphismus f eines unitären Vektorraums (V,σ).

4. Zur Folgerung über die Zerlegung von Endomorphismen: Die Zerlegung läßt sich explizit angeben, falls eine Orthonormalbasis aus Eigenvektoren von $f^* \circ f$ (selbstadjungiert) bekannt ist.

9.2.5. *Simultanes Diagonalisieren.* Es folgt nun eine für die Anwendungen wichtige Verallgemeinerung der Folgerung in 9.1.1 (bzw. in 9.1.2).

Satz. *Sei (V,σ) ein endlichdimensionaler unitärer (bzw. euklidischer) Vektorraum und $M \subseteq \mathrm{Hom}_{\mathbb{C}}(V,V)$ (bzw. $M \subseteq \mathrm{Hom}_{\mathbb{R}}(V,V)$) eine Teilmenge. Dann sind äquivalent:*

1. a. *Für alle $f \in M$ gilt: f ist normal (bzw. selbstadjungiert).*
   b. *Für alle $f, g \in M$ gilt: $f \circ g = g \circ f$.*
2. *Es existiert eine Orthonormalbasis $e_1,\dots,e_n$ von V, so daß für alle $f \in M$ gilt: $e_1,\dots,e_n$ sind Eigenvektoren von f. ($e_1,\dots,e_n$ ist also eine "simultane" Orthonormalbasis).*

Beweis. $2 \Longrightarrow 1$. (a) gilt nach der Folgerung 9.1.1 (bzw. in 9.1.2). b. Seien $f, g \in M$. Dann existieren γ_i, λ_i für $1 \leq i \leq n$ mit:

$$f\begin{pmatrix} e_1\\ \vdots\\ e_n\end{pmatrix} = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n\end{pmatrix}\begin{pmatrix} e_1\\ \vdots\\ e_n\end{pmatrix} \quad \text{und} \quad g\begin{pmatrix} e_1\\ \vdots\\ e_n\end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\begin{pmatrix} e_1\\ \vdots\\ e_n\end{pmatrix}.$$

Also folgt $f \circ g = g \circ f$.

$1 \Longrightarrow 2$. Sei (V,σ) unitär. (Den euklidischen Fall behandelt man analog).

Schritt 1. Sei U der von M erzeugte Unterraum von $\mathrm{Hom}_{\mathbb{C}}(V,V)$. Es ist $\dim(U) < \infty$, denn es gilt $\dim(\mathrm{Hom}_{\mathbb{C}}(V,V)) = (\dim(V))^2 < \infty$. Sei $f_1,\dots,f_m$ eine Basis von U. Es genügt, eine simultane Orthonormalbasis $e_1,\dots,e_n$ aus Eigenvektoren für $f_1,\dots,f_m$ zu finden. Ist nämlich $e_1,\dots,e_n$ eine solche Orthonormalbasis und ist $f \in M$, so ist $f = \sum_{j=1}^m \alpha_j f_j$ mit $\alpha_j \in \mathbb{C}$ und damit folgt für $1 \le i \le n$:

$$\begin{aligned} f(e_i) &= \sum_{j=1}^m \alpha_j f_j(e_i) \\ &= \sum_{j=1}^m \alpha_j \gamma_{ij} e_i \qquad \gamma_{ij} \text{ Eigenwert von } f_j \text{ zum Eigenvektor } e_i \\ &= \underbrace{\Big(\sum_{j=1}^m \alpha_j \gamma_{ij}\Big)}_{\in\mathbb{C}} e_i. \end{aligned}$$

Also ist e_i Eigenvektor von f.

Schritt 2. Nach Schritt 1 können wir also ohne Einschränkung $M = \{f_1,\dots,f_m\}$ setzen. Wir beweisen nun die Aussage des Satzes durch Induktion nach m.

$m = 1$. Siehe Folgerung 9.1.1.

$m-1 \Longrightarrow m$. Seien $\lambda_1,\dots,\lambda_t$ die verschiedenen Eigenwerte von f_1 und V_{λ_i} der Eigenraum zum Eigenwert λ_i. Nach der vorangegangenen Bemerkung (2) (Umformulierung der Folgerung in 9.1.1) gilt $V = \bigoplus_{i=1}^t V_{\lambda_i}$ und $V_{\lambda_i} \perp V_{\lambda_j}$ für $i \neq j$. Wegen $f_1 \circ f_j = f_j \circ f_1$ gilt $f_j(V_\lambda) \subseteq V_\lambda$ für jeden Eigenwert λ von f_1, denn für $x \in V_\lambda$ haben wir

$$\begin{aligned} &f_1(x) = \lambda x \\ &f_1(f_j(x)) = f_j(f_1(x)) = f_j(\lambda x) = \lambda f_j(x) \\ &f_j(x) \in V_\lambda. \end{aligned}$$

Sei nun λ ein fest gewählter Eigenwert von f_1 und $2 \le j \le t$. Setze

$$\begin{aligned} g_j := f_j{\restriction}V_\lambda \colon V_\lambda &\to V_\lambda \\ x &\mapsto f_j(x). \end{aligned}$$

Nach den Bemerkungen in 9.1.1 gilt:

- g_j ist normal, da $V = \bigoplus_{i=1}^t V_{\lambda_i}$ und $V_{\lambda_i} \perp V_{\lambda_j}$ für $i \neq j$, und
- $g_s \circ g_t = g_t \circ g_s$ für $s \neq 0$, $2 \le s,t \le n$, da $f_s \circ f_t = f_t \circ f_s$.

Nach Induktionsvoraussetzung hat V_λ eine simultane Orthonormalbasis aus Eigenvektoren von g_j, $j = 2,\dots,m$. Die Vereinigung dieser Orthonormalbasen für alle V_λ, λ Eigenwert von f_1, liefert die Behauptung. QED

BEMERKUNGEN. 1. Genauso zeigt man: Sei V ein endlichdimensionaler K-Vektorraum und $M \subseteq \mathrm{Hom}_K(V,V)$ eine Teilmenge. Dann sind äquivalent:

1. a. Für alle $f \in M$ gilt: f ist diagonalisierbar.
   b. Für alle $f, g \in M$ gilt: $f \circ g = g \circ f$.

2. Es existiert eine simultane Basis von V aus Eigenvektoren aller $f \in M$.

2. Matrizenformulierung des Satzes: Sei $M \subseteq \mathbb{C}^{n\times n}$ (bzw. $M \subseteq \mathbb{R}^{n\times n}$) eine Teilmenge. Dann sind äquivalent:

1. a. Für alle $A \in M$ gilt: A ist normal (bzw. selbstadjungiert).
   b. Für alle $A, B \in M$ gilt: $AB = BA$.
2. Es existiert ein $C \in U(n)$ (bzw. $\in O(n)$), so daß für alle $A \in M$ gilt: C^*AC ist eine Diagonalmatrix.

9.3 Orthogonale Endomorphismen

Wir wollen uns eine Übersicht über alle orthogonalen Endomorphismen eines endlichdimensionalen euklidischen Vektorraums verschaffen. Zunächst zeigen wir, daß jeder orthogonale Endomorphismus als Hintereinanderausführung von Hyperebenenspiegelungen dargestellt werden kann. Wir betrachten dann den zwei- und dreidimensionalen Fall und können dort leicht die orthogonalen Endomorphismen klassifizieren. Mit Hilfe der reellen JORDANschen Normalform ergibt sich, daß damit auch schon der allgemeine Fall einer beliebigen endlichen Dimension erledigt ist. Abschließend verschaffen wir uns noch eine Parameterdarstellung der $\mathrm{SO}(n)$, unter Verwendung der schon in Abschnitt 5.4 benutzten JACOBIschen Rotationsmatrizen.

9.3.1. *Zerlegung in Hyperebenenspiegelungen.* Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum. Wir definieren

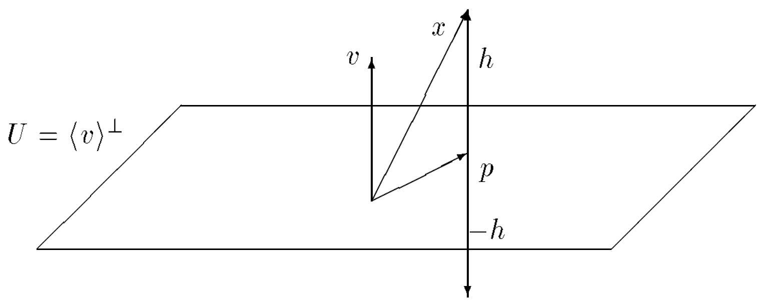
$$\begin{aligned} \mathrm{O}(V) &:= \mathrm{O}(V,\sigma) := \{\, f \in \mathrm{Hom}_{\mathbb{R}}(V,V) \mid f \text{ orthogonal} \,\}, \\ \mathrm{SO}(V) &:= \mathrm{SO}(V,\sigma) := \{\, f \in \mathrm{O}(V) \mid \det(f) = 1 \,\}. \end{aligned}$$

BEMERKUNG. Wir wollen uns zunächst überlegen, daß sich jede Hyperebenenspiegelung auf einem Unterraum von V zu einer Hyperebenenspiegelung auf ganz V fortsetzen läßt.

BEWEIS. Sei etwa $\dim(V) = n$. Unter einer Hyperebene verstehen wir einen Unterraum $U \subseteq V$ mit $\dim(U) = n - 1$. Sei $U^\perp = \langle v \rangle$, also $U = \langle v \rangle^\perp$ und $V = \langle v \rangle^\perp \oplus \langle v \rangle$. Als Hyperebenenspiegelung hatten in wir in 9.2.2 Endomorphismen der Gestalt

$$\begin{aligned} f_v \colon V &\to V \\ x = p + h &\mapsto p - h \quad \text{mit } p \in \langle v \rangle^\perp \text{ und } h \in \langle v \rangle \end{aligned}$$

bezeichnet.



Es gilt

$$f_v(x) = x - 2\frac{\sigma(x,v)}{\sigma(v,v)}v, \tag{9.6}$$

denn für $x = p + \alpha v$ mit $p \in \langle v\rangle^\perp$ und $\alpha \in \mathbb{R}$ ist

$$\begin{aligned} \sigma(x,v) &= \alpha\sigma(v,v) && \text{da } p \perp v, \\ f_v(x) &= p - \alpha v \\ &= x - 2\alpha v \\ &= x - 2\frac{\sigma(x,v)}{\sigma(v,v)}v. \end{aligned}$$

Aus (9.6) folgt offenbar unsere Behauptung, daß sich nämlich jede Hyperebenenspiegelung auf einem Unterraum von V zu einer Hyperebenenspiegelung auf ganz V fortsetzen läßt. QED

SATZ. *(Zerlegung in Hyperebenenspiegelungen). Sei (V,σ) ein euklidischer Vektorraum mit $\dim(V) = n < \infty$, sei $f \in \mathrm{O}(V)$ und $f \neq \mathrm{id}$. Dann existieren ein $r \leq n$ und Hyperebenenspiegelungen $f_1, \ldots, f_r \in \mathrm{O}(V)$ mit $f = f_1 \circ \cdots \circ f_r$.*

BEWEIS. Durch Induktion nach n.

Basis $n = 1$. Sei e eine Orthonormalbasis von V. Dann

$$\begin{aligned} f(e) &= \pm e && \text{da } f \text{ orthogonal,} \\ f(e) &= -e && \text{da } f \neq \mathrm{id}, \\ f(\alpha e) &= -\alpha e && \text{für alle } \alpha \in \mathbb{R}. \end{aligned}$$

Also ist $f = f_e$.

Schritt $n-1 \Longrightarrow n$ (für $n > 1$).

1. Annahme: Es existiert ein $v \in V$ mit $f(v) = v$ und $v \neq 0$. Es ist

$$V = \underbrace{\langle v\rangle^\perp}_{=:U} \oplus \langle v\rangle.$$

Zunächst zeigen wir: $f(\langle v\rangle^\perp) \subseteq \langle v\rangle^\perp$. Sei dazu $u \in \langle v\rangle^\perp$. Dann gilt:

$$\begin{aligned} &\sigma(f(u),v) = \sigma(f(u),f(v)) = \sigma(u,v) = 0 \quad \text{und deshalb} \\ &f(u) \in \langle v\rangle^\perp. \end{aligned}$$

Also ist $f{\restriction}U \in \mathrm{O}(U)$. Nach Induktionsvoraussetzung existieren $v_1, \ldots, v_r \in U$, $r \leq n-1$ mit

$$f{\restriction}U = (f_{v_1}{\restriction}U) \circ \cdots \circ (f_{v_r}{\restriction}U) \quad (\text{oder } f{\restriction}U = \mathrm{id},\ r = 0)$$

(siehe vorige Bemerkung). Damit folgt $f = f_{v_1} \circ \cdots \circ f_{v_r}$, denn für ein $u \in U$ gilt nach Konstruktion $f(u) = (f_{v_1} \circ \cdots \circ f_{v_r})(u)$, für v gilt $f(v) = v = f_{v_1}(\ldots f_{v_r}(v)\ldots)$, da $v \perp v_i$ für alle $1 \leq i \leq r$, und $V = U \oplus \langle v\rangle$.

2. Reduktion des allgemeinen Falls auf (1). Es ist $f \neq \mathrm{id}$, also existiert ein $v \in V$ mit $f(v) \neq v$. Setze $w := f(v) - v$. Wir zeigen $(f_w \circ f)(v) = v$.

Zum Beweis beachte man, daß $2f(v) = f(v) + v + \underbrace{f(v) - v}_{w}$. Es gilt $(f(v) + v) \perp w$, denn

$$\begin{aligned}\sigma(f(v)+v, f(v)) &= \underbrace{\sigma(f(v), f(v))}_{=\sigma(v,v)} + \sigma(v, f(v)) \\ &= \sigma(v, v + f(v)) \\ &= \sigma(v + f(v), v).\end{aligned}$$

Also folgt nach der Definition von f_w

$$\begin{aligned}f_w(2f(v)) &= f(v) + v + \underbrace{-w}_{v-f(v)} = 2v \\ f_w(f(v)) &= v.\end{aligned}$$

Nach (1) folgt nun wegen $f_w \circ f \in \mathrm{O}(V)$, daß es Hyperebenenspiegelungen $f_1, \ldots, f_r \in \mathrm{O}(V)$ gibt $(r \leq n-1)$ mit

$$\begin{aligned}f_w \circ f &= f_1 \circ \cdots \circ f_r \\ f &= f_w \circ f_1 \circ \cdots \circ f_r, \quad \text{da } f_w^2 = \mathrm{id}. \quad \text{QED}\end{aligned}$$

Definition. Sei V ein K-Vektorraum und f ein Endomorphismus. $\{ x \in V \mid f(x) = x \}$ heißt *Fixraum* von f.

Folgerung. *Sei (V, σ) ein endlichdimensionaler euklidischer Vektorraum und $f \in \mathrm{O}(V)$.*

1. *Sind $f_1, \ldots, f_r$ $(r \leq n)$ Hyperebenenspiegelungen in $\mathrm{O}(V)$ mit $f = f_1 \circ \cdots \circ f_r$, so hat der Fixraum von f eine Dimension $\geq n - r$.*

2. *Sei $\dim(V) = n$ ungerade.*

   Gilt $\det(f) = +1$, so ist $+1$ Eigenwert von f.

   Gilt $\det(f) = -1$, so ist -1 Eigenwert von f.

Beweis. 1. Sei $f_i = f_{v_i}$ mit $v_i \in V$ für $1 \leq i \leq r$. Es ist $V = \langle v_1 \ldots, v_r \rangle \oplus \langle v_1, \ldots, v_r \rangle^\perp$, also $\dim(\langle v_1, \ldots, v_r \rangle) \leq r$ und $\dim(\langle v_1, \ldots, v_r \rangle^\perp) \geq n - r$. Ferner gilt

$$\langle v_1, \ldots, v_r \rangle^\perp \subseteq \{ x \in V \mid f(x) = x \},$$

denn aus $x \perp v_i$ für alle $1 \leq i \leq r$ folgt $(f_{v_1} \circ \cdots \circ f_{v_r})(x) = x$.

2. Sei $\dim(V) = n$ ungerade. Nach dem Satz existieren Hyperebenenspiegelungen $f_1, \ldots, f_r$, $r \leq n$ mit $f = f_1 \circ \cdots \circ f_r$. Also ist $\det(f) = \det(f_1) \ldots \det(f_r) = (-1)^r$.

Fall $\det(f) = +1$. Dann ist r gerade, also $r < n$. Nach (1) hat f einen Fixpunkt $\neq 0$. Also ist 1 Eigenwert von f.

Fall $\det(f) = -1$. Dann ist $\det(-f) = (-1)^n \det(f) = (-1)(-1) = 1$. Also hat $-f$ einen Fixpunkt $x \neq 0$, d.h. $(-f)(x) = x$, d.h. $f(x) = -x$. Also ist -1 ein Eigenwert von f. QED

9.3.2. *Geometrische Veranschaulichung im Fall* $\dim(V) = 2$ *oder* $= 3$. Sei (V, σ) ein euklidischer Vektorraum mit $\dim(V) = 2$. Es seien e_1, e_2 eine Orthonormalbasis von V und $v = \alpha_1 e_1 + \alpha_2 e_2 \in V$ $(\alpha_1, \alpha_2 \in \mathbb{R})$. Es gelte $\|v\| = 1$, d. h. $\alpha_1^2 + \alpha_2^2 = 1$. Wir betrachten die Geradenspiegelung

$$\begin{aligned} &f_v \colon V \to V \\ &f_v(x) = x - 2\sigma(x, v)v \quad (\text{da } \sigma(v, v) = 1). \end{aligned}$$

Man erhält

$$\begin{aligned} f_v(e_1) &= e_1 - 2\alpha_1(\alpha_1 e_1 + \alpha_2 e_2) \\ f_v(e_2) &= e_2 - 2\alpha_2(\alpha_1 e_1 + \alpha_2 e_2) \end{aligned}$$

und damit

$$f_v \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} 1 - 2\alpha_1^2 & -2\alpha_1\alpha_2 \\ -2\alpha_1\alpha_2 & 1 - 2\alpha_2^2 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

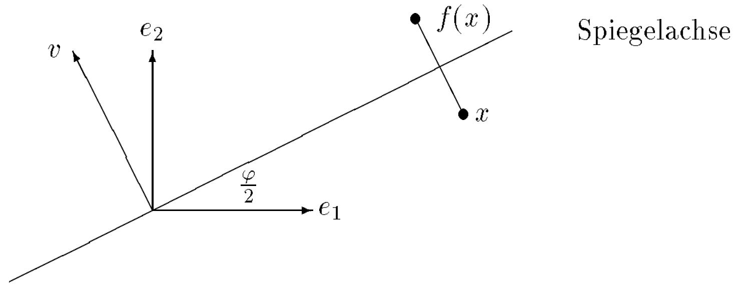
Es gilt $1 = \alpha_1^2 + \alpha_2^2$, also $1 - 2\alpha_1^2 = \alpha_2^2 - \alpha_1^2$. Wegen $\alpha_1^2 + \alpha_2^2 = 1$ kann man ein $\varphi \in \mathbb{R}$ wählen mit $\cos\frac{\varphi}{2} = \alpha_2$ und $\sin\frac{\varphi}{2} = -\alpha_1$. Nach dem Additionstheorem für sin und cos erhält man dann:

$$\begin{aligned} 1 - 2\alpha_1^2 &= \alpha_2^2 - \alpha_1^2 = \cos^2\frac{\varphi}{2} - \sin^2\frac{\varphi}{2} = \cos\varphi \\ -2\alpha_1\alpha_2 &= 2\cos\frac{\varphi}{2}\sin\frac{\varphi}{2} = \sin\varphi \\ 1 - 2\alpha_2^2 &= \alpha_1^2 - \alpha_2^2 = -\cos\varphi. \end{aligned}$$

Also ergibt sich

$$f_v \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

Geometrisch handelt es sich um eine Spiegelung an $\langle v \rangle^\perp$, $v = -\sin\frac{\varphi}{2}\, e_1 + \cos\frac{\varphi}{2}\, e_2$.



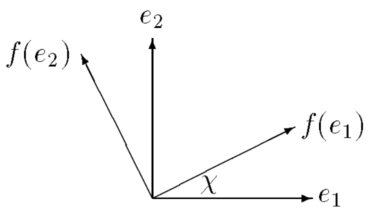
Komposition von Geradenspiegelungen: Seien $\varphi, \psi \in \mathbb{R}$.

$$\begin{aligned} \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \begin{pmatrix} \cos\psi & \sin\psi \\ \sin\psi & -\cos\psi \end{pmatrix} &= \begin{pmatrix} \cos\varphi\cos\psi + \sin\varphi\sin\psi & \cos\varphi\sin\psi - \sin\varphi\cos\psi \\ \sin\varphi\cos\psi - \cos\varphi\sin\psi & \sin\varphi\sin\psi + \cos\varphi\cos\psi \end{pmatrix} \\ &= \begin{pmatrix} \cos(\psi - \varphi) & \sin(\psi - \varphi) \\ -\sin(\psi - \varphi) & \cos(\psi - \varphi) \end{pmatrix} \end{aligned}$$

bedeutet eine Drehung um den Winkel $\psi - \varphi$ gegen den Uhrzeigersinn, denn

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\chi & \sin\chi \\ -\sin\chi & \cos\chi \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

bedeutet geometrisch



Ergebnis: Sei e_1, e_2 eine Orthonormalbasis von V und $f \in \mathrm{O}(V)$. Dann gilt:

$\det(f) = -1 \implies f$ *Geradenspiegelung*, es existiert ein $\varphi \in \mathbb{R}$ mit

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad (-\pi < \varphi \leq \pi).$$

$\det(f) = +1 \implies f$ *Drehung*, es existiert ein $\varphi \in \mathbb{R}$ mit

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad (-\pi < \varphi \leq \pi).$$

f ist Komposition zweier Geradenspiegelungen.

Bemerkung. Diese Gestalt der darstellenden Matrizen kann man auch leicht direkt einsehen.

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{O}(2) \implies A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \text{ oder } A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} \text{ für ein } \varphi \in \mathbb{R}.$$

Beweis. A ist orthogonal, also gilt

$$\begin{aligned} &\alpha^2 + \beta^2 = 1 \\ &\gamma^2 + \delta^2 = 1 \\ &\alpha\gamma + \beta\delta = 0. \end{aligned}$$

Annahme: $\alpha, \beta \neq 0$.

$$\begin{aligned} &\frac{\gamma}{\beta} = -\frac{\delta}{\alpha} =: \chi \\ &\gamma = \beta\chi \quad , \quad \delta = -\alpha\chi \\ &1 = \gamma^2 + \delta^2 = \chi^2(\beta^2 + \alpha^2) = \chi^2 \\ &\chi = \pm 1. \end{aligned}$$

Damit folgt

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} \text{ oder } A = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Wegen $\alpha^2 + \beta^2 = 1$ gibt es ein $\varphi \in \mathbb{R}$ mit $\alpha = \cos\varphi$ und $\beta = \sin\varphi$.

Falls $\alpha = 0$: $\beta = \pm 1$, $\delta = 0$, $\gamma = \pm 1$.

Falls $\beta = 0$: $\alpha = \pm 1$, $\gamma = 0$, $\delta = \pm 1$. QED

Wir behandeln jetzt den dreidimensionalen Fall. Sei also (V, σ) ein euklidischer Vektorraum, $\dim(V) = 3$ und $f \in \mathrm{O}(V)$.

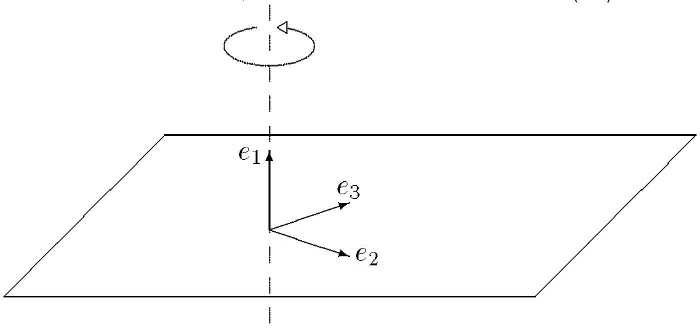
Fall $\det(f) = 1$. Nach der Folgerung in 9.3.1 ist 1 Eigenwert von f, da $\dim(V) = 3$ ungerade ist. Sei e_1 Eigenvektor zum Eigenwert 1 mit $\|e_1\| = 1$, und e_2, e_3 eine Orthonormalbasis von $\langle e_1\rangle^\perp$. Dann ist

$$f{\restriction}\langle e_2, e_3\rangle\colon \langle e_2, e_3\rangle \to \langle e_2, e_3\rangle$$

eine orthogonale Abbildung mit $\det(f{\restriction}\langle e_2, e_3\rangle) = 1$. Nach der vorangehenden Überlegung im zweidimensionalen Fall gibt es ein $\varphi \in \mathbb{R}$ mit

$$f\begin{pmatrix} e_1\\ e_2\\ e_3\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\varphi & \sin\varphi\\ 0 & -\sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix} e_1\\ e_2\\ e_3\end{pmatrix}, \quad -\pi < \varphi \le \pi.$$

f ist eine *Drehung* mit *Drehachse* $\mathbb{R}e_1$, *Drehebene* $\mathbb{R}e_2 \oplus \mathbb{R}e_3 = \langle e_1\rangle^\perp$ und *Drehwinkel* φ.



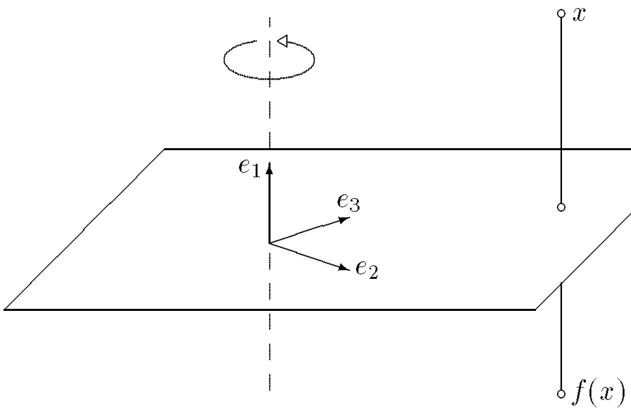
Fall $\det(f) = -1$. Nach der Folgerung in 9.3.1 ist -1 Eigenwert von f. Sei e_1 Eigenvektor von f zum Eigenwert -1 mit $\|e_1\| = 1$, und e_2, e_3 eine Orthonormalbasis von $\langle e_1\rangle^\perp$. Dann ist

$$f{\restriction}\langle e_2, e_3\rangle\colon \langle e_2, e_3\rangle \to \langle e_2, e_3\rangle$$

eine orthogonale Abbildung mit $\det(f{\restriction}\langle e_2, e_3\rangle) = 1$, da $\det(f) = -1$. Nach der vorangehenden Überlegung im zweidimensionalen Fall existiert ein $\varphi \in \mathbb{R}$ mit

$$f\begin{pmatrix} e_1\\ e_2\\ e_3\end{pmatrix} = \begin{pmatrix} -1 & 0 & 0\\ 0 & \cos\varphi & \sin\varphi\\ 0 & -\sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix} e_1\\ e_2\\ e_3\end{pmatrix}, \quad -\pi < \varphi \le \pi.$$

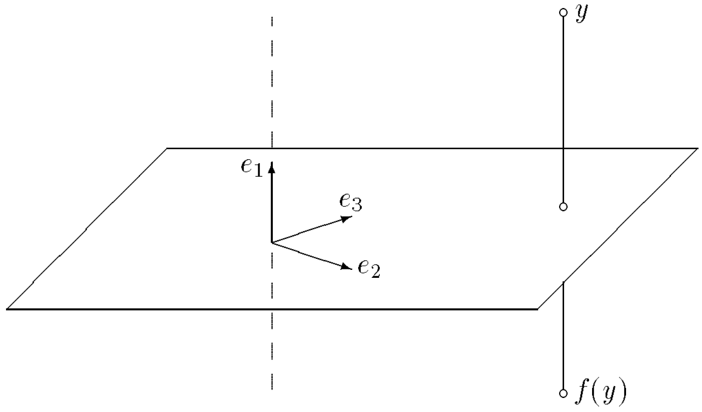
f ist eine *Drehspiegelung*, d.h. eine Drehung um die Achse $\mathbb{R}e_1$ um den Winkel φ und eine anschließende Spiegelung an der Drehebene $\mathbb{R}e_2 \oplus \mathbb{R}e_3$.



Grenzfälle: $\varphi = 0$.

$$f\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

Es handelt sich um eine Spiegelung an der Ebene $\langle e_2, e_3 \rangle$, und zwar um die Hyperebenenspiegelung f_{e_1}.



$\varphi = \pi$.

$$f\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

Hier hat man eine Spiegelung am Ursprung.

Wir zeigen jetzt, daß der Drehwinkel φ nur von der Orientierung der gewählten Orthonormalbasis abhängt.

FOLGERUNG. *Sei (V, σ) ein euklidischer, orientierter Vektorraum mit* $\dim(V) = 2$ *und* $f \in \mathrm{SO}(V)$. *Dann existiert genau ein $\varphi \in \mathbb{R}$, $-\pi < \varphi \leq \pi$, so daß für alle positiv orientierten Orthonormalbasen e_1, e_2 von V gilt*

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

φ *heißt* orientierter Drehwinkel *und* $|\varphi|$ absoluter Drehwinkel.

BEWEIS. Sei e_1, e_2 eine fest gewählte, positiv orientierte Orthonormalbasis von V. Dann existiert nach obiger Bemerkung genau ein $\varphi \in \mathbb{R}$, $-\pi < \varphi \leq \pi$, mit

$$f\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$$

(φ ist eindeutig bestimmt nach dem Verlauf von sin und cos). Sei e_1', e_2' eine beliebige, positiv orientierte Orthonormalbasis von V. Dann gibt es ein $A \in \mathrm{SO}(2)$ mit

$$\begin{pmatrix} e_1' \\ e_2' \end{pmatrix} = A \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}.$$

A vertauscht mit $\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$, da $A = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix}$ für ein $\psi \in \mathbb{R}$ ist, d.h.

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} A = \begin{pmatrix} \cos(\varphi+\psi) & \sin(\varphi+\psi) \\ -\sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} = A \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Daraus folgt

$$f\begin{pmatrix} e_1' \\ e_2' \end{pmatrix} = \underbrace{A\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}A^t}_{\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}}\begin{pmatrix} e_1' \\ e_2' \end{pmatrix}. \quad \text{QED}$$

DEFINITION. Für $A = (\alpha_{ij}) \in K^{n\times n}$ ist die *Spur* von A definiert durch $\operatorname{Spur}(A) := \sum_{i=1}^n \alpha_{ii}$.

BEMERKUNG. Sei (V, σ) ein euklidischer, orientierter Vektorraum mit $\dim(V) = 2$ und $f \in \mathrm{SO}(V)$. Sei v_1, v_2 eine beliebige Basis von V und $A \in \mathbb{R}^{2\times 2}$ mit

$$f\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = A\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Dann ist $\operatorname{Spur}(A) = 2\cos\varphi$, also ist der Drehwinkel $|\varphi|$ durch $\operatorname{Spur}(A)$ festgelegt.

BEWEIS. A ist ähnlich zu $\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$. Es genügt also zu zeigen, daß ähnliche Matrizen dieselbe Spur haben.

$$\operatorname{Spur}(CAC^{-1}) = \operatorname{Spur}(AC^{-1}C) = \operatorname{Spur}(A) \quad \text{für } A, C \in K^{n\times n} \text{ und } C \in \mathrm{GL}(n, K),$$

denn allgemein gilt für $A = (\alpha_{ij}), B = (\beta_{ij}) \in K^{n\times n}$:

$$\operatorname{Spur}(AB) = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}\beta_{ji} = \operatorname{Spur}(BA). \quad \text{QED}$$

FOLGERUNG. *Sei (V, σ) ein euklidischer Vektorraum mit $\dim(V) = 3$ und $f \in \mathrm{SO}(V)$. Sei e_1 ein Eigenvektor von f zum Eigenwert 1 mit $\|e_1\| = 1$ (e_1 existiert nach der Folgerung in 9.3.1) und sei $\langle e_1\rangle^\perp$ orientiert. Dann existiert genau ein $\varphi \in \mathbb{R}$ mit $-\pi < \varphi \leq \pi$, so daß für alle positiv orientierten Orthonormalbasen e_2, e_3 von $\mathbb{R}e_1^\perp$ gilt:*

$$f\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}.$$

BEWEIS. f induziert eine Drehung in $\langle e_1\rangle^\perp$. Also folgt die Behauptung unmittelbar aus der obigen Folgerung. QED

BEMERKUNG. Sei $A \in \mathbb{R}^{3\times 3}$ beliebige darstellende Matrix von f bzgl. einer beliebigen Basis. Dann ist $\operatorname{Spur}(A) = 1 + 2\cos\varphi$, da A ähnlich ist zu

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi \\ 0 & -\sin\varphi & \cos\varphi \end{pmatrix}$$

(Beweis wie zuvor). Also ist der absolute Drehwinkel $|\varphi|$ durch $\operatorname{Spur}(A)$ festgelegt.

9.3.3. *Geometrische Veranschaulichung im allgemeinen Fall.* Aus den bisher behandelten Spezialfällen ergibt sich jetzt mit Hilfe des Satzes 8.2.2 von der reellen JORDANschen Normalform eine entsprechende Aussage für den allgemeinen Fall.

SATZ. *Sei (V,σ) ein euklidischer Vektorraum mit $1 \leq \dim(V) < \infty$ und sei $f \in \mathrm{O}(V)$. Dann existieren eine Orthonormalbasis $e_1,\ldots,e_n$ von V, $r,s,t \in \mathbb{N}$, und $\varphi_1,\ldots,\varphi_t \in \mathbb{R}$ mit*

$$f\begin{pmatrix} e_1\\ \vdots \\ e_n\end{pmatrix} = \begin{pmatrix} \underbrace{\begin{matrix}1&&\\&\ddots&\\&&1\end{matrix}}_{r-mal} &&&& \\ & \underbrace{\begin{matrix}-1&&\\&\ddots&\\&&-1\end{matrix}}_{s-mal} &&& \\ && \boxed{\begin{matrix}\cos\varphi_1 & \sin\varphi_1\\ -\sin\varphi_1 & \cos\varphi_1\end{matrix}} && \\ &&& \ddots & \\ &&&& \boxed{\begin{matrix}\cos\varphi_t & \sin\varphi_t\\ -\sin\varphi_t & \cos\varphi_t\end{matrix}}\end{pmatrix}\begin{pmatrix} e_1\\ \vdots\\ e_n\end{pmatrix}.$$

BEWEIS. 1. Gezeigt wird: Es existieren Unterräume $U_i \subseteq V$, $1 \leq i \leq m$ mit $V = \bigoplus_{i=1}^m U_i$, $f(U_i) \subseteq U_i$, $\dim(U_i) = 1,2$ und $U_i \perp U_j$ für alle $i \neq j$. Dies beweisen wir durch Induktion nach $\dim(V)$.

$\dim(V) = 1$. Klar.

Induktionsschritt. Nach der reellen JORDANschen Normalform 8.2.2 gibt es zu jedem $f \in \mathrm{Hom}_{\mathbb{R}}(V,V)$ einen Unterraum $U \subseteq V$ mit $\dim(U) = 1,2$ und $f(U) \subseteq U$. Denn ist etwa

$$J = \begin{pmatrix} \boxed{\begin{matrix}\alpha & -\beta\\ \beta & \alpha\end{matrix}} & \boxed{\begin{matrix}1 & 0\\ 0 & 1\end{matrix}} &&&& \\ & \ddots &&&& \\ && \ddots &&& \\ &&& \boxed{\begin{matrix}\alpha & -\beta\\ \beta & \alpha\end{matrix}} & \boxed{\begin{matrix}1 & 0\\ 0 & 1\end{matrix}} \\ &&&& \boxed{\begin{matrix}\alpha & -\beta\\ \beta & \alpha\end{matrix}}\end{pmatrix},$$

eine darstellende Matrix in reeller JORDANscher Normalform und ist U der von den beiden Basisvektoren, die zu dem letzten Kästchen $\boxed{\begin{matrix}\alpha & -\beta\\ \beta & \alpha\end{matrix}}$ gehören, erzeugte Unterraum, so ist $\dim(U) = 2$ und es gilt $f(U) \subseteq U$. Es folgt $f(U^\perp) \subseteq U^\perp$, denn

$$\begin{array}{ll} f(U) = U & \text{da } f \text{ ein Isomorphismus ist. Also ist} \\ f^{-1}\restriction U\colon U \to U & \text{orthogonal.} \end{array}$$

Sei $v \in U^\perp$, $u \in U$. Dann ist

$$\begin{aligned} \sigma(f(v),u) &= \sigma(f^{-1}(f(v)), f^{-1}(u)) \\ &= \sigma(v, f^{-1}(u)) \\ &= 0, \end{aligned}$$

also $f(v) \in U^\perp$. Daraus folgt $V = U \oplus U^\perp$, $f{\restriction}U^\perp : U^\perp \to U^\perp$, $\dim(U) = 1$ oder $\dim(U) = 2$, $\dim(U^\perp) < \dim(V)$ und $U \perp U^\perp$. Auf $U^\perp$ läßt sich also die Induktionsvoraussetzung anwenden, woraus die Behauptung folgt.

2. Nach (1) genügt es wegen 9.3.2, den Satz für $\dim(V) = 1, 2$ zu beweisen. Für $\dim(V) = 1$ gilt er offensichtlich, und für $\dim(V) = 2$ beachte man, daß die Matrix $\left(\begin{smallmatrix} \cos\varphi_i & \sin\varphi_i \\ \sin\varphi_i & -\cos\varphi_i \end{smallmatrix}\right)$ nach 9.3.2 eine Spiegelung an $\langle v\rangle^\perp$, $v = -\sin\frac{\varphi}{2}\, e_1 + \cos\frac{\varphi}{2}\, e_2$ darstellt und deshalb durch Wahl einer geeigneten Orthonormalbasis in die Form $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ gebracht werden kann. QED

9.3.4. *Parameterdarstellung der* SO(n). Nach 9.3.2 ist

$$\begin{aligned} \mathbb{R} &\to \mathrm{SO}(2) \\ \varphi &\mapsto \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \end{aligned}$$

surjektiv; diese Abbildung heißt *Parameterdarstellung* von SO(2).

Gesucht ist eine ähnliche Parameterdarstellung für SO(n). Heuristisch betrachtet erkennt man: “Für SO(n) braucht man $\frac{n(n-1)}{2}$ Parameter”. Denn für $A = (\alpha_{ij}) \in \mathbb{R}^{n\times n}$ sind äquivalent

$A \in \mathrm{SO}(n)$

die Spalten von A bilden eine Orthonormalbasis von $\mathbb{R}^n$

$\sum_{i=1}^n \alpha_{ij}\alpha_{il} = \delta_{jl}$ für $j \le l$.

Das sind $\frac{n(n-1)}{2} + n$ Relationen, und es bleiben noch $\frac{n(n-1)}{2}$ “freie Parameter” übrig. Dies sieht man wie folgt. Sei x die Anzahl der Elemente von A über der Diagonalen und n die Anzahl der Elemente auf der Diagonalen. Dann gilt: $n + 2x = n^2$, also $x = \frac{n^2-n}{2} = \frac{n(n-1)}{2}$.

Sei $1 \le i < j \le n$ und $\varphi \in \mathbb{R}$. In Abschnitt 5.4 hatten wir die die JACOBImatrix $R_{ij}(\varphi)$ definiert durch

$$R_{ij}(\varphi) := \begin{array}{cc} \begin{pmatrix} 1 &&&&&&&&& \\ & \ddots &&&&&&&& \\ && 1 &&&&&&& \\ &&& \cos\varphi &&&& \sin\varphi &&& \\ &&&& 1 &&&&& \\ &&&&& \ddots &&&& \\ &&&&&& 1 &&& \\ &&& -\sin\varphi &&&& \cos\varphi && \\ &&&&&&&& 1 & \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix} & \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix} \\ \begin{matrix} \qquad\qquad\quad i & \qquad\qquad\qquad j \end{matrix} & \end{array}$$

$$= (\alpha_{st}) \in \mathbb{R}^{n\times n} \text{ mit } \alpha_{st} = \begin{cases} \delta_{st} & \text{falls } (s,t) \ne (i,i),(i,j),(j,i),(j,j) \\ \cos\varphi & \text{falls } (s,t) = (i,i),(j,j) \\ \sin\varphi & \text{falls } (s,t) = (i,j) \\ -\sin\varphi & \text{falls } (s,t) = (j,i). \end{cases}$$

Bemerkungen. 1. Jacobimatrizen wurden in Abschnitt 5.4 beim Beweis der Hauptachsentransformation nach dem Jacobischen Verfahren verwendet.

2. $R_{ij}(\varphi) \in \mathrm{SO}(n)$ beschreibt eine "ebene Drehung". Sei (V, σ) ein euklidischer Vektorraum und $e_1, \ldots, e_n$ eine Orthonormalbasis von V,

$$f \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = R_{ij}(\varphi) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}.$$

Dann ist $V = \langle e_i, e_j \rangle \oplus \langle e_i, e_j \rangle^{\perp}$, $f{\restriction}\langle e_i, e_j \rangle$ ist eine Drehung und $f{\restriction}\langle e_i, e_j \rangle^{\perp}$ ist die Identität.

3. $R_{ij}(\varphi) R_{ij}(-\varphi) = E$, d.h. $R_{ij}(\varphi)^t = R_{ij}(\varphi)^{-1} = R_{ij}(-\varphi)$.

Satz. *Sei $n \in \mathbb{N}$ und $A \in \mathrm{SO}(n)$. Dann ist A Produkt von $\frac{n(n-1)}{2}$ Jacobimatrizen.*

Beweis. 1. Seien $A = (\alpha_{rt})$, $1 \leq i < j \leq n$, $\varphi \in \mathbb{R}$, $s := \sin \varphi$ und $c := \cos \varphi$. Betrachte

$$AR_{ij}(\varphi) = (\alpha_{rt}) \begin{array}{cc} \begin{matrix} & & & i & & & & & j & & & \end{matrix} & \\ \begin{pmatrix} 1 & & & & & & & & & & & \\ & \ddots & & & & & & & & & & \\ & & 1 & & & & & & & & & \\ & & & c & & & & & s & & & \\ & & & & 1 & & & & & & & \\ & & & & & \ddots & & & & & & \\ & & & & & & \ddots & & & & & \\ & & & & & & & 1 & & & & \\ & & & -s & & & & & c & & & \\ & & & & & & & & & 1 & & \\ & & & & & & & & & & \ddots & \\ & & & & & & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix} \end{array}$$

$$= \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1i-1} & c\alpha_{1i} - s\alpha_{1j} & \alpha_{1i+1} & \ldots & \alpha_{1j-1} & s\alpha_{1i} + c\alpha_{1j} & \alpha_{1i+1} & \ldots & \alpha_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \ldots & \alpha_{ni-1} & c\alpha_{ni} - s\alpha_{nj} & \alpha_{ni+1} & \ldots & \alpha_{nj-1} & s\alpha_{ni} + c\alpha_{nj} & \alpha_{ni+1} & \ldots & \alpha_{nn} \end{pmatrix}$$
$$= \begin{pmatrix} a_1 & \cdots & a_{i-1} & ca_i - sa_j & a_{i+1} & \cdots & a_{j-1} & sa_i + ca_j & a_{j+1} & \cdots & a_n \end{pmatrix},$$

wobei a_t die t-te Spalte von A ist. Zum Beweis des Satzes wird nun φ passend gewählt. Dazu benötigen wir folgende Aussage: Sind $\alpha, \beta \in \mathbb{R}$, so existiert ein $\varphi \in \mathbb{R}$ derart daß mit $s := \sin \varphi$ und $c := \cos \varphi$ gilt

$$\begin{aligned} c\alpha - s\beta &= 0 \\ s\alpha + c\beta &\geq 0. \end{aligned} \tag{9.7}$$

Beweis von (9.7). *Fall* $\alpha \neq 0$ oder $\beta \neq 0$. Wähle φ mit

$$\begin{aligned} \sin \varphi = s &= \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}} \\ \cos \varphi = c &= \frac{\beta}{\sqrt{\alpha^2 + \beta^2}}. \end{aligned}$$

Dies ist möglich, da $s^2 + c^2 = 1$ ist. Dann gilt

$$\begin{aligned} c\alpha - s\beta &= 0 \\ s\alpha + c\beta &= \sqrt{\alpha^2 + \beta^2} > 0. \end{aligned}$$

Fall $\alpha = \beta = 0$. Dann gilt für beliebige φ

$$\begin{aligned} c\alpha - s\beta &= 0 \\ s\alpha + c\beta &= 0. \end{aligned}$$

2. Nach (1) existiert ein $\varphi_{12} \in \mathbb{R}$, so daß mit $R_{12} := R_{12}(\varphi_{12})$ gilt:

$$AR_{12} = \begin{pmatrix} * & \dots\dots\dots\dots\dots\dots\dots\dots & & & * \\ \vdots & & & & \vdots \\ * & \dots\dots\dots\dots\dots\dots\dots\dots & & & * \\ & & & & \\ 0 & \underbrace{(\sin\varphi)\alpha_{n1} + (\cos\varphi)\alpha_{n2}}_{\geq 0} & \alpha_{n3} & \dots & \alpha_{nn} \end{pmatrix}$$

(nach (9.7) mit $\alpha = \alpha_{n1}, \beta = \alpha_{n2}$). Ebenso gibt es ein $\varphi_{23} \in \mathbb{R}$ so daß mit $R_{23} := R_{23}(\varphi_{23})$ gilt:

$$AR_{12}R_{23} = \begin{pmatrix} * & \dots\dots\dots\dots & & & & * \\ \vdots & & & & & \vdots \\ * & \dots\dots\dots\dots & & & & * \\ & & & & & \\ 0 & 0 & \underbrace{\alpha'_{n3}}_{\geq 0} & \alpha_{n4} & \dots & \alpha_{nn} \end{pmatrix}.$$

Durch Iteration erhält man $\varphi_{12}, \varphi_{23}, \varphi_{34}, \dots, \varphi_{n-1,n}$ mit

$$AR_{12}R_{23}R_{34}\dots R_{n-1n} = \begin{pmatrix} * & \dots\dots & & * \\ \vdots & & & \vdots \\ * & \dots\dots & & * \\ & & & \\ 0 & \dots & \dots\ 0 & \underbrace{\tilde{\alpha}_{nn}}_{\geq 0} \end{pmatrix} =: (\tilde{\alpha}_{rt}).$$

Wegen $A, R_{12}, \dots, R_{n-1n} \in \mathrm{SO}(n)$ folgt $(\tilde{\alpha}_{rt}) \in \mathrm{SO}(n)$. Also folgt

$$1 = \sum_{j=1}^{n} \tilde{\alpha}_{nj}^2 = \tilde{\alpha}_{nn}^2 \quad \text{und damit} \quad \tilde{\alpha}_{nn} = 1, \quad \text{da } \tilde{\alpha}_{nn} \geq 0.$$

Außerdem gilt

$$\begin{aligned} 1 = \sum_{i=1}^{n} \tilde{\alpha}_{in}^2 &= 1 + \sum_{i=1}^{n-1} \tilde{\alpha}_{in}^2 \quad && \text{da } \tilde{\alpha}_{nn} = 1, \text{ also} \\ \tilde{\alpha}_{in} &= 0 && \text{für } 1 \leq i \leq n-1. \end{aligned}$$

Als Ergebnis erhalten wir

$$AR_{12}R_{23}\dots R_{n-1n} = \begin{pmatrix} \boxed{\quad * \quad} & \begin{matrix}0\\ \vdots\\ 0\end{matrix} \\ 0\dots 0 & 1 \end{pmatrix}.$$

Wendet man nun das gleiche Verfahren auf $\boxed{*} \in \mathrm{SO}(n-1)$ an, so folgt durch Iteration

$$A\,\underbrace{R_{12}R_{23}\dots R_{n-1n}}\,\underbrace{R_{12}\dots R_{n-2n-1}}\dots\underbrace{R_{12}R_{23}}\,\underbrace{R_{12}} = \begin{pmatrix} \alpha & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix}, \alpha \in \mathbb{R},$$

und es ist $\alpha = 1$, da $A, R_{st} \in \mathrm{SO}(n)$ sind. Das abschließende Ergebnis lautet also

$$A\,\underbrace{\underbrace{R_{12}R_{23}\dots R_{n-1n}}_{n-1}\,\underbrace{R_{12}\dots R_{n-2n-1}}_{n-2}\dots\underbrace{R_{12}R_{23}}_{2}\,\underbrace{R_{12}}_{1}}_{\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}\ \text{Jacobimatrizen}} = E.$$

Damit folgt die Behauptung, da $R_{ij}(\varphi)^{-1} = R_{ij}(-\varphi)$ ist. QED

FOLGERUNG. *Ist $A \in \mathrm{SO}(3)$, so existieren $\varphi_1, \varphi_2, \varphi_3 \in \mathbb{R}$ mit*

$$A = \begin{pmatrix} \cos\varphi_1 & \sin\varphi_1 & 0 \\ -\sin\varphi_1 & \cos\varphi_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi_2 & \sin\varphi_2 \\ 0 & -\sin\varphi_2 & \cos\varphi_2 \end{pmatrix} \begin{pmatrix} \cos\varphi_3 & \sin\varphi_3 & 0 \\ -\sin\varphi_3 & \cos\varphi_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

BEWEIS. Nach dem Satz gibt es $\varphi_1, \varphi_2, \varphi_3 \in \mathbb{R}$ mit $A = R_{12}(\varphi_1)R_{23}(\varphi_2)R_{12}(\varphi_3)$. QED

Die Zahlen $\varphi_1, \varphi_2, \varphi_3$ heißen *Eulersche Winkel*. Sie werden z.B. bei der Beschreibung einer Kreiselbewegung verwendet. Geometrisch kann man dies wie folgt deuten. Seien $\varphi_1, \varphi_2, \varphi_3 \in \mathbb{R}$ und $A = R_{12}(\varphi_1)R_{23}(\varphi_2)R_{12}(\varphi_3)$. Sei (V, σ) ein euklidischer Vektorraum mit einer Orthonormalbasis e_1, e_2, e_3 und

$$f\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} = A\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}.$$

Wir können dann f schreiben als Hintereinanderausführung der Drehungen f_1, f_2, f_3, also in der Form $f = f_1 \circ f_2 \circ f_3$ mit

$$f_3\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} := R_{12}(\varphi_3)\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} =: \begin{pmatrix} e_1' \\ e_2' \\ e_3' \end{pmatrix} \qquad \text{Orthonormalbasis von } V,$$

$$f_2\begin{pmatrix} e_1' \\ e_2' \\ e_3' \end{pmatrix} := R_{23}(\varphi_2)\begin{pmatrix} e_1' \\ e_2' \\ e_3' \end{pmatrix} =: \begin{pmatrix} e_1'' \\ e_2'' \\ e_3'' \end{pmatrix} \qquad \text{Orthonormalbasis von } V,$$

$$f_1\begin{pmatrix} e_1'' \\ e_2'' \\ e_3'' \end{pmatrix} := R_{12}(\varphi_1)\begin{pmatrix} e_1'' \\ e_2'' \\ e_3'' \end{pmatrix} = \underbrace{R_{12}(\varphi_1)R_{23}(\varphi_2)R_{12}(\varphi_3)}_{A}\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

f_3 ist eine Drehung mit Drehachse $\mathbb{R}e_3$ und Drehwinkel φ_3, f_2 eine Drehung mit Drehachse $\mathbb{R}e_1'$ und Drehwinkel φ_2, und f_1 eine Drehung mit Drehachse $\mathbb{R}e_3''$ und Drehwinkel φ_1.

Bemerkungen. Es gibt hier einen Zusammenhang zur Topologie: man fasse $\mathrm{SO}(n)$, $\mathrm{O}(n)$ und $\mathrm{GL}(n,\mathbb{R})$ als Teilmengen von $\mathbb{R}^{n^2}$ mit der induzierten Topologie auf, also

$$\mathrm{SO}(n) \subseteq \mathrm{O}(n) \subseteq \mathrm{GL}(n,\mathbb{R}) \subseteq \mathbb{R}^{n\times n} \cong \mathbb{R}^{n^2}.$$

1. $\mathrm{O}(n)$ ist beschränkt und abgeschlossen in $\mathbb{R}^{n^2}$, also *kompakt*, und ebenso $\mathrm{SO}(n)$.

Beweis. Sei $A = (\alpha_{ij}) \in \mathbb{R}^{n^2}$. Dann gilt

$$A \in \mathrm{O}(n) \equiv \sum_{j=1}^{n} \alpha_{ij}\alpha_{lj} = \delta_{il} \quad \text{für alle } i,l$$
$$A \in \mathrm{SO}(n) \equiv A \in \mathrm{O}(n) \text{ und } \det(A) = 1.$$

Also sind $\mathrm{O}(n)$ (und $\mathrm{SO}(n)$) abgeschlossen als Nullstellengebilde von Gleichungen. Außerdem folgt aus $A \in \mathrm{O}(n)$, daß $\sum_{j=1}^{n} \alpha_{ij}^2 = 1$ für alle i, also $|\alpha_{ij}| \leq 1$ für alle i,j. Also ist $\mathrm{O}(n)$ beschränkt. Qed

2. Nach dem Satz ist

$$\Phi\colon \mathbb{R}^{\frac{n(n-1)}{2}} \to \mathrm{SO}(n)$$
$$(\varphi_1,\ldots,\varphi_{\frac{n(n-1)}{2}}) \mapsto R_{12}(\varphi_1)\ldots R_{n-1n}(\varphi_{n-1})R_{12}(\varphi_n)\ldots R_{n-2n-1}(\varphi_{2n-2})\ldots R_{12}(\varphi_{\frac{n(n-1)}{2}})$$

surjektiv. Φ ist stetig, da $\varphi \mapsto R_{ij}(\varphi)$ und die Matrizenmultiplikation stetig sind. Da $\mathbb{R}^{\frac{n(n-1)}{2}}$ zusammenhängend ist, ist also auch $\mathrm{SO}(n)$ zusammenhängend. Aber es gibt keine Bijektion

$$\Psi\colon \mathbb{R}^{\frac{n(n-1)}{2}} \xrightarrow{\cong} \mathrm{SO}(n),$$

so daß Ψ und Ψ^{-1} stetig sind, denn $\mathrm{SO}(n)$ ist kompakt und $\mathbb{R}^{\frac{n(n-1)}{2}}$ ist nicht kompakt.

Aber es gibt eine "lokale Isomorphie", die Cayley-Transformation (siehe Übungen).

$$\mathbb{R}^{\frac{n(n-1)}{2}} \cong \{\, A \in \mathbb{R}^{n\times n} \mid A^t = -A \,\} \xrightarrow{\cong} \{\, B \in \mathrm{SO}(n) \mid \det(B+E) \neq 0 \,\} \quad \text{Isomorphismus}$$
$$A \mapsto (E-A)(E+A)^{-1}$$
$$(E-B)(E+B)^{-1} \leftarrow B$$

d.h. $\mathrm{SO}(n)$ (und auch $\mathrm{O}(n)$, s.u.) sind "Mannigfaltigkeiten" der Dimension $\frac{n(n-1)}{2}$.

3. Sei $A \in \mathrm{O}(n)$, $\det(A) = -1$, z.B.

$$A = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Dann ist $\mathrm{O}(n)$ disjunkte Vereinigung von $\mathrm{SO}(n)$ und $A\cdot\mathrm{SO}(n)$ (also $\mathrm{O}(n) = \mathrm{SO}(n) \cup A\cdot\mathrm{SO}(n)$ und $\mathrm{SO}(n) \cap A\cdot\mathrm{SO}(n) = \emptyset$). Dies ist eine Zerlegung in *Zusammenhangskomponenten*. (Hierbei ist $A\cdot\mathrm{SO}(n) := \{\, AB \mid B \in \mathrm{SO}(n)\,\}$).

Beweis. Sei $B \in \mathrm{O}(n)$ mit $B \notin \mathrm{SO}(n)$. Dann ist $\det(B) = -1$, also $\det(A^tB) = 1$, damit $B = A\underbrace{A^tB}_{\in \mathrm{SO}(n)} \in A\cdot\mathrm{SO}(n)$. Außerdem ist $\mathrm{SO}(n) \cap A\cdot\mathrm{SO}(n) = \emptyset$ und nach (b) ist $\mathrm{SO}(n)$ zusammenhängend, also ist auch $A\cdot\mathrm{SO}(n)$ zusammenhängend. Qed

Bemerkung. Aus dem Beweis des Satzes folgt allgemein für ein $A \in \mathbb{R}^{n \times n}$ (also nicht notwendig $A \in \mathrm{SO}(n)$): Es existiert ein $B \in \mathrm{SO}(n)$, welches Produkt von $\frac{n(n-1)}{2}$ Jacobimatrizen ist, mit der Eigenschaft, daß

$$AB = C = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ & \ddots & \vdots \\ & & \gamma_{nn} \end{pmatrix}$$

eine obere Dreiecksmatrix ist. Außerdem gilt $A = CB^t$, also $A^t = BC^t$, wobei C^t eine untere Dreiecksmatrix ist. Damit erhalten wir: Es existiert ein $Q \in \mathrm{SO}(n)$ und eine untere Dreiecksmatrix R mit $A = QR$. Diese so genannte *QR-Zerlegung* spielt in der numerischen Mathematik eine wichtige Rolle.

Kapitel 10

Gruppen, Moduln, Tensorprodukt

In diesem Kapitel wollen wir die grundlegenden algebraischen Hilfsmittel ausbauen. Als Anwendung ergibt sich in Abschnitt 10.1 eine Klassifikation der endlichen Untergruppen der SO(3). In Abschnitt 10.2 definieren wir Faktorgruppen und beweisen ihre universelle Eigenschaft; als Anwendungen ergeben sich einige einfache zahlentheoretische Aussagen, z.B. $x^p = x \mod p$ für $x \in \mathbb{Z}$ und Primzahlen p. Der Begriff eines Moduls (eine Art Vektorraum über einem Ring statt eines Körpers) wird in Abschnitt 10.3 eingeführt; mit seiner Hilfe erhalten wir einen einfachen Beweis des Satzes von Cayley/Hamilton, welcher aussagt, daß das "Einsetzen" einer Matrix in ihr eigenes charakteristisches Polynom stets die Nullmatrix ergibt. In Abschnitt 10.4 konstruieren wir schließlich die in der Physik und auch in der Differentiageometrie wichtigen Tensorprodukte von Moduln; auf sie wird man in natürlicher Weise durch die Betrachtung von mehrstelligen linearen Abbildungen (sog. multilinearen Abbildungen) geführt.

10.1 Gruppen

In Abschnitt 1.2.2 hatten wir den Gruppenbegriff eingeführt und seine einfachsten Eigenschaften bewiesen. Wir wollen uns jetzt etwas mehr Kenntnisse über Gruppen und insbesondere endliche Gruppen verschaffen; Ziel ist eine Klassifikation der endlichen Untergruppen der SO(3). Im folgenden verwenden wir x, y, z (statt α, β, γ) für die Elemente einer Gruppe und schreiben xy statt $x \circ y$.

Sei G eine endliche Gruppe. Dann heißt die Anzahl der Elemente von G die *Ordnung* von G, geschrieben $\mathrm{Ord}(G)$. Für eine endliche Gruppe $G = \{g_1, \dots, g_n\}$ kann man die Multiplikation in einer Gruppentafel explizit aufschreiben:

$$\begin{array}{c|cccc} \cdot & g_1 & g_2 & \cdots & g_n \\ \hline g_1 & g_1g_1 & g_1g_2 & \cdots & g_1g_n \\ g_2 & g_2g_1 & g_2g_2 & \cdots & g_2g_n \\ \vdots & \vdots & \vdots & & \vdots \\ g_n & g_ng_1 & g_ng_2 & \cdots & g_ng_n \end{array}$$

In jeder Zeile und in jeder Spalte kommt jedes Element von G genau einmal vor.

10.1.1. *Gruppenhomomorphismen, Untergruppen, Normalteiler.* Seien G, H Gruppen und $f\colon G \to H$ eine Abbildung. f heißt *Gruppenhomomorphismus*, falls für alle $x, y \in G$ gilt

$$f(xy) = f(x)f(y).$$

f heißt *Gruppenisomorphismus*, falls f ein bijektiver Gruppenhomomorphismus ist. G ist *isomorph* zu H (geschrieben $G \cong H$), falls ein Gruppenisomorphismus $g\colon G \to H$ existiert.

Bemerkung. Seien G, H Gruppen und $f\colon G \to H$ ein Gruppenhomomorphismus. Dann gilt

$$\begin{aligned} f(e) &= e \\ f(x^{-1}) &= f(x)^{-1} \quad \text{für alle } x \in G. \end{aligned}$$

Beweis.

$$\begin{aligned} f(e) &= f(ee) = f(e)f(e), && \text{also } f(e) = e. \\ f(e) &= f(xx^{-1}) = f(x)f(x^{-1}) && \text{also } f(x)^{-1} = f(x^{-1}). \text{ QED} \end{aligned}$$

In 5.3.1 hatten wir bereits den Begriff einer Untergruppe wie folgt definiert. Eine Teilmenge U einer Gruppe G heißt *Untergruppe*, wenn die Gruppenstruktur auf G eine Gruppenstruktur auf U induziert, d.h. wenn gilt (1) $xy \in U$ für alle $x, y \in U$, (2) $e \in U$ und (3) $x^{-1} \in U$ für alle $x \in U$. Eine Untergruppe N einer Gruppe G heißt *Normalteiler*, wenn für alle $x \in G$ gilt

$$xN = Nx,$$

wobei $xN := \{\, xy \mid y \in N \,\}$ und $Nx := \{\, yx \mid y \in N \,\}$.

Bemerkung. Sei G eine Gruppe und $U \subseteq G$ eine Teilmenge. Dann ist U ist genau dann Normalteiler von G, wenn U eine Untergruppe von G ist und für alle $x \in G$ und $y \in U$ auch $xyx^{-1} \in U$ ist.

Beweis. $\Longrightarrow$. Aus $xU = Ux$ folgt $xyx^{-1} \in U$. $\Longleftarrow$. Es ist $xU \subseteq Ux$ und aus $x^{-1}U \subseteq Ux^{-1}$ folgt $Ux \subseteq xU$. Also $xU = Ux$. QED

Lemma. *Sei G eine Gruppe und $U \subseteq G$ eine Teilmenge. Dann sind äquivalent:*

1. *U ist Untergruppe von G.*
2. a. *$U \neq \emptyset$.*
   b. *Für alle $x, y \in U$ ist $xy^{-1} \in U$.*

Beweis. 1 $\Longrightarrow$ 2. klar. 2 $\Longrightarrow$ 1. Da $U \neq \emptyset$ ist, gibt es ein $x_0 \in U$. Wegen (2b) ist dann $e = x_0x_0^{-1} \in U$. Für ein beliebiges $y \in U$ ist wieder mit (2b) auch $y^{-1} = ey^{-1} \in U$. Schließlich folgt für beliebige $x, y \in U$ wegen $y^{-1} \in U$ und wegen (2b) auch $xy = x(y^{-1})^{-1} \in U$. QED

Lemma. *Seien G, H Gruppen, $f\colon G \to H$ ein Gruppenhomomorphismus, U eine Untergruppe von G und V eine Untergruppe von H. Dann gilt:*

1. *$f(U)$ ist Untergruppe von H.*
2. *$f^{-1}(V)$ ist Untergruppe von G.*
3. *Falls V Normalteiler von H ist, so ist $f^{-1}(V)$ Normalteiler von G.*

Beweis. (1) und (2) sind klar. 3. Sei $x \in G$ und $y \in f^{-1}(V)$. Dann gilt

$$f(xyx^{-1}) = f(x)f(y)f(x)^{-1} \in V,$$

da $f(y) \in V$ und V Normalteiler ist. Also ist $xyx^{-1} \in f^{-1}(V)$. QED

Bemerkung. Sei U Normalteiler in G und $f\colon G \to H$ ein Gruppenhomomorphismus. Dann muß $f(U)$ nicht Normalteiler in H sein. Dies sieht man durch folgendes Beispiel: Wähle $G \subseteq H$ Untergruppe, die kein Normalteiler von H ist und $U = G$, sowie $f\colon G \to H$ die Einbettung von G nach H.

Definition. Sei G eine Gruppe und $M \subseteq G$ eine Teilmenge.

1. $\langle M \rangle := \{\, x \in G \mid \exists n{\in}\mathbb{N}, x_1, \ldots, x_n{\in}M, \varepsilon_1, \ldots, \varepsilon_n{\in}\{1,-1\}\ x = x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n} \,\}$ heißt *von M erzeugte Untergruppe*. (Dabei sei $\langle \emptyset \rangle := \{e\}$).

2. G heißt *zyklisch*, falls es ein $x \in G$ gibt mit $G = \langle x \rangle$.

Bemerkungen. 1. $\langle M \rangle$ ist die kleinste Untergruppe von G, die M enthält. Dies beweist man wie die analoge Aussage für Vektorräume.

2. Ist $G = \langle x \rangle$ für ein $x \in G$, so ist $G = \{\, x^s \mid s \in \mathbb{Z} \,\}$.

10.1.2. Beispiele. 1. Ist V ein K-Vektorraum, so ist V mit der Addition eine Gruppe.

2. $\mathbb{Z}$ ist mit der Addition eine Gruppe. Es gilt $\mathbb{Z} = \langle 1 \rangle$. In additiven Gruppen schreibt man nx mit $n \in \mathbb{Z}$ statt x^n und $-x$ statt x^{-1}.

3. Sei R ein Ring und $\mathrm{U}(R) := \{\, r \in R \mid \exists s{\in}R : rs = sr = 1 \,\}$. Dann ist $\mathrm{U}(R)$ mit der Multiplikation von R eine Gruppe, sie heißt *Einheitengruppe* von R. Im Spezialfall eines Körpers K ist $\mathrm{U}(K) = K^* = \{\, x \in K \mid x \neq 0 \,\}$.

4. Sei M eine Menge und $S(M) := \{\, f\colon M \to M \mid f \text{ bijektive Abbildung} \,\}$. Dann ist $S(M)$ mit der Komposition von Abbildungen eine Gruppe. Sie heißt *symmetrische Gruppe* auf M.

Im Spezialfall $M = \{1, 2, \ldots, n\}$ für ein $n \in \mathbb{N}$ ist $S_n := S(\{1, \ldots, n\})$ die *Permutationsgruppe* von n Elementen, die wir bereits in 6.1.4 betrachtet hatten. Definiere den Gruppenhomomorphismus

$$\begin{aligned} \varepsilon_n\colon S_n &\to \{1, -1\} \\ \sigma &\mapsto \text{Vorzeichen der Permutation } \sigma \end{aligned}$$

und setze $A_n := \mathrm{Kern}(\varepsilon_n)$. A_n ist die Menge der geraden Permutationen und heißt *alternierende Gruppe* von n Elementen. Leicht zu erkennen ist (siehe 6.1.4)

$$\mathrm{Ord}(S_n) = n!,$$
$$\mathrm{Ord}(A_n) = \frac{n!}{2}.$$

Sei $N \subseteq M$ eine Teilmenge. Dann sind

$$\begin{aligned} G &:= \{\, f \in S(M) \mid f(N) = N \,\} \\ H &:= \{\, f \in S(M) \mid \forall x{\in}N\ f(x) = x \,\} \end{aligned}$$

Untergruppen von $S(M)$, denn etwa für G gilt: 1. $\mathrm{id}_M \in G$. 2. Sind $f, g \in G$, so ist $(fg)(N) = f(g(N)) = f(N) = N$, also ist $fg \in G$. 3. Sei $f \in G$. Dann ist $f{\restriction}N\colon N \to N$ bijektiv, also gilt für alle $x \in N$: $f^{-1}(f(x)) = x$. Damit erhält man $f^{-1}(N) = N$, also ist auch $f^{-1} \in G$. Ebenso beweist man, daß H Untergruppe von S_n ist.

Insbesondere ist $S_n \subseteq S_{n+1}$ Untergruppe bei der Identifikation

$$S_n = \{\, \sigma \in S_{n+1} \mid \sigma(n+1) = n+1 \,\}.$$

5. Sei V ein Vektorraum über einem Körper K. Dann ist $\mathrm{Aut}_K(V) := \{ f \in \mathrm{Hom}_K(V,V) \mid f \text{ Isomorphismus} \}$ mit der Komposition von Abbildungen eine Gruppe. Sie heißt *Automorphismengruppe* von V.

6. $\mathrm{GL}(n,K)$ heißt *allgemeine lineare Gruppe*. Es ist $\det\colon \mathrm{GL}(n,K) \to K^*$ ein Gruppenhomomorphismus und $\mathrm{SL}(n,K) := \mathrm{Kern}(\det)$ ist die *spezielle lineare Gruppe*. Für $K = \mathbb{R}$ ist $\mathrm{O}(n)$ die *orthogonale Gruppe* und $\mathrm{SO}(n) = \mathrm{Kern}(\det\colon \mathrm{O}(n) \to \{1,-1\})$ ist die *spezielle orthogonale Gruppe*.

7. Sei V ein endlichdimensionaler K-Vektorraum. Dann ist $\{ g\colon V \to V \mid g \text{ Affinität} \}$ die *affine Gruppe* von V. (Erinnerung: Eine Abbildung $g\colon V \to V$ hieß Affinität, wenn es ein $f \in \mathrm{Aut}_K(V)$ und ein $a \in V$ gibt, so daß für alle $x \in V$ gilt $g(x) = f(x) + a$).

8. Sei (V,σ) ein endlichdimensionaler euklidischer Vektorraum. Dann heißt $\{ g\colon V \to V \mid g \text{ Bewegung} \}$ *euklidische Gruppe* von V. (Erinnerung: Eine Abbildung $g\colon V \to V$ ist genau dann eine Bewegung (oder Kongruenz), falls es ein $f \in \mathrm{O}(V)$ und ein $a \in V$ gibt, so daß für alle $x \in V$ gilt $g(x) = f(x) + a$).

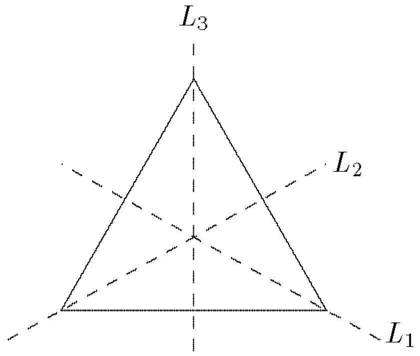
Definition. Sei (V,σ) ein endlichdimensionaler euklidischer Vektorraum und sei $F \subseteq V$ eine Teilmenge ("geometrische Figur"). Dann heißt

$$\{ g\colon V \to V \mid g \text{ Bewegung von } V,\ g(F) = F \}$$

Symmetriegruppe von F.

Bemerkung. Die Symmetriegruppe von F ist tatsächlich eine Gruppe, nämlich eine Untergruppe der euklidischen Gruppe von V. Dies zeigt man wie im obigen Beispiel bei $S(M)$.

Beispiele. (Anschaulich). 1. Sei $F \subseteq \mathbb{R}^2$ ein gleichseitiges Dreieck.



Es sei oBdA 0 im Schnittpunkt von L_1, L_1, L_3, und G sei die Symmetriegruppe von F. Leicht einzusehen ist $G \subseteq \mathrm{O}(\mathbb{R}^2)$. Sei f die Drehung um den Winkel $\frac{2\pi}{3}$ und g_i die Spiegelung an L_i für $i = 1,2,3$. Man beachte, daß $\mathrm{id}, f, f^2, g_1, g_2, g_3$ verschieden sind.

Behauptung. $G = \{\mathrm{id}, f, f^2, g_1, g_2, g_3\}$.

Beweis. Jedes $g \in G$ bewirkt eine Permutation der Eckpunkte (g ist durch die Bilder der Eckpunkte festgelegt). Also gilt

$$\{\mathrm{id}, f, f^2, g_1, g_2, g_3\} \subseteq G \hookrightarrow S_3,$$

($\hookrightarrow$ soll bedeuten: injektive Abbildung, also Einbettung) und S_3 hat die Ordnung 6. Qed

Aus dem obigen Beweis folgt sogar $G \cong S_3$. Es ist

$$G = \{\mathrm{id}, f, f^2, g, gf, gf^2\},$$

wobei g eine der Spiegelungen g_i ist (Beweis einfach). Ferner gelten folgende Relationen:

$$\begin{aligned} f^3 &= \mathrm{id} \\ g^2 &= \mathrm{id} \\ gfgf &= \mathrm{id} \qquad\qquad \text{(leicht nachzurechnen),} \end{aligned}$$

also auch

$$fg = gf^{-1} = gf^2.$$

Allgemein kann man zeigen (dies wird hier aber nicht durchgeführt): Sei G die Symmetriegruppe des regulären n-Ecks mit $n \geq 3$, f die Drehung um den Winkel $\frac{2\pi}{n}$ und g eine beliebige Spiegelung in G. Dann ist

$$G = \{\mathrm{id}, f, \ldots, f^{n-1}, g, gf, \ldots, gf^{n-1}\} =: D_{2n}.$$

$D_{2n} = \langle f, g\rangle$ heißt *Diedergruppe* der Ordnung $2n$. (Oft wird auch die Bezeichnung D_n statt D_{2n} verwendet). Es gelten folgende Relationen:

$$\begin{aligned} f^n &= \mathrm{id} \\ f^t &\neq \mathrm{id} \qquad \text{für } 1 \leq t < n \\ fg &= gf^{n-1} \end{aligned}$$

2. $\{\, f \in \mathrm{SO}(\mathbb{R}^2) \mid f$ Symmetrie des regulären n-Ecks$\,\}$ wird von der Drehung um den Winkel $\frac{2\pi}{n}$ erzeugt, ist also eine zyklische Gruppe der Ordnung n.

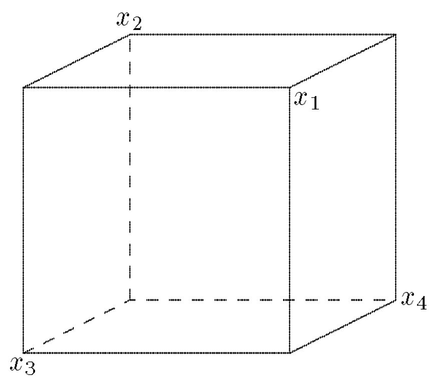
3. Für ein $F \subseteq \mathbb{R}^3$ nennt man $\{\, f \in \mathrm{SO}(\mathbb{R}^3) \mid f$ Symmetrie von $F\,\}$ die *Drehgruppe* von F.

Die Bestimmung der Drehgruppen für die so genannten *platonischen Körper* (dies sind die fünf konvexen, regulären Polyeder im $\mathbb{R}^3$) liefert folgendes Ergebnis.

Polyeder	Beschreibung	Drehgruppe	mit Ordnung
Tetraeder	4 Dreiecke	A_4	$4 \cdot 3 = 12$
Würfel	6 Quadrate	S_4	$6 \cdot 4 = 24$
Oktaeder	8 Dreiecke	S_4	$8 \cdot 3 = 24$
Dodekaeder	12 Fünfecke	A_5	$12 \cdot 5 = 60$
Ikosaeder	20 Dreiecke	A_5	$20 \cdot 3 = 60$

Wir wollen dies im folgenden begründen.

a. Sei G die Drehgruppe des Würfels



Durch jeden Eckpunkt geht genau eine Raumdiagonale. Bei Symmetrien werden Raumdiagonalen in Raumdiagonale abgebildet. Sei L_i die Raumdiagonale durch den Punkt x_i. Betrachte

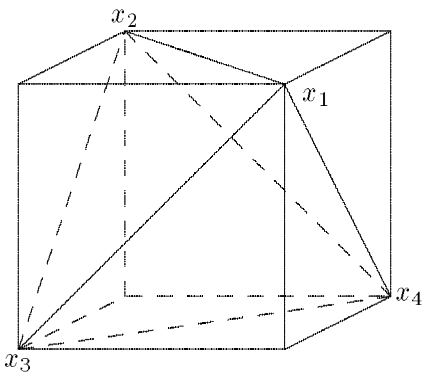
$$\begin{aligned} \varphi\colon G &\to S_4 \\ f &\mapsto (i \mapsto f'(i)) \quad \text{wobei } f' \text{ definiert ist durch } f(L_i) = L_{f'(i)}. \end{aligned}$$

φ ist ein injektiver Gruppenhomomorphismus. Zählt man die Elemente von G auf, so erhält man als Symmetrieachsen (Drehachsen von Elementen in G):

- 4 Raumdiagonalen: Je 2 Drehungen $\neq$ id.
- 3 Geraden durch die Mittelpunkte gegenüberliegender Seiten: Je 3 Drehungen $\neq$ id.
- 6 Geraden durch die Mittelpunkte gegenüberliegender Kanten: Je 1 Drehung $\neq$ id.

Also hat G mindestens $8 + 9 + 6 + 1 = 24$ Elemente. (Die letzte 1 gehört zur Identität). Da φ ein Monomorphismus ist, $\mathrm{Ord}(G) \geq 24$ und $\mathrm{Ord}(S_4) = 4! = 24$ ist, erhält man als Ergebnis $G \cong S_4$.

b. Sei H die Drehgruppe des Tetraeders



Sei f die Drehung des Würfels mit $f(x_1) = x_1$, d.h. $f(L_1) = L_1$. Jedes solche f ist eine Drehung des Dreiecks x_2, x_3, x_4. Für $f \neq$ id ergeben sich zwei Möglichkeiten, nämlich

$$\varphi(f_1) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} \quad \text{oder} \quad \varphi(f_2) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}.$$

Es sind $\varphi(f_1)$ und $\varphi(f_2)$ gerade Permutationen. Sei g_i die Drehung des Würfels, die den Tetraeder in sich selbst abbildet mit $g_i(x_1) = x_i$ für $i = 2, 3, 4$. Wieder ist $\varphi(g_i)$ eine gerade Permutation. Sei $g \in H$ beliebig. Dann gilt $g(x_1) = x_i$ für ein $1 \leq i \leq 4$.

Fall 1. $g(x_1) = x_1$. Dann ist $g = \mathrm{id}$, $g = f_1$ oder $g = f_2$.

Fall 2. $g(x_1) = x_i$ mit $i \neq 1$. Dann ergibt sich

$$\begin{aligned} &g_i^{-1} g(x_1) = x_1 \\ &g_i^{-1} g = \mathrm{id},\ g_i^{-1} g = f_1 \text{ oder } g_i^{-1} g = f_2 \\ &g = g_i,\ g = g_i f_1 \text{ oder } g = g_i f_2. \end{aligned}$$

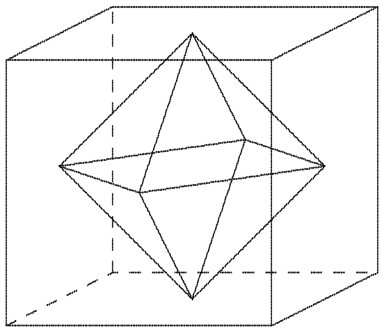
Also ist $\varphi(g)$ wieder eine gerade Permutation.

Außerdem hat H mindestens 12 Elemente, nämlich

$$\begin{aligned} &f_0 := \mathrm{id}, f_1, f_2, \\ &g_i f_0, g_i f_1, g_i f_2 \quad \text{für } i = 2, 3, 4. \end{aligned}$$

Ergebnis: $H \cong A_4$, denn $\mathrm{Ord}(H) \geq 12$ und $\mathrm{Ord}(A_4) = \frac{4!}{2} = 12$.

c. Drehgruppe des Oktaeders ist die Drehgruppe des Würfels, denn verbindet man die Mitten der Seiten des Würfels, so ergibt sich ein Oktaeder.



d. Durch ähnliche Überlegungen folgt: Der Dodekaeder und der Ikosaeder haben die gleiche Drehgruppe, und diese ist isomorph zu A_5.

10.1.3. *Operationen einer Gruppe auf einer Menge.* Unser Ziel ist eine Klassifikation der endlichen Untergruppen der SO(3). Vorbereitend dazu sollen einige nützliche Begriffe eingeführt werden.

Definition. Sei G eine Gruppe, S eine Menge und $m\colon G \times S \to S$ eine Abbildung. Wir schreiben $gx := m(g,x)$ für $g \in G$ und $x \in S$. m heißt *Operation* von G auf S, falls für alle $x \in S$ und $g, h \in G$ gilt:

$$\begin{aligned} (gh)x &= g(hx) \\ ex &= x. \end{aligned}$$

Ferner heißen

$$\begin{aligned} &Gx := \{\, gx \mid g \in G \,\} && \textit{Orbit} \text{ (oder } \textit{Bahn}\text{) von } x, \\ &G_x := \{\, g \in G \mid gx = x \,\} && \textit{Isotropiegruppe} \text{ (oder } \textit{Stabilisator}\text{) von } x. \end{aligned}$$

Bemerkung. Sei G eine Gruppe, S eine Menge und $m\colon G \times S \to S$ eine Operation von G auf S sowie $x \in S$. Dann ist G_x eine Untergruppe von G.

Beweis. 1. Wegen $ex = x$ ist $e \in G_x$. 2. Seien $g, h \in G_x$. Dann ist $(gh)x = g(hx) = gx = x$, also auch $gh \in G_x$. 3. Sei $g \in G_x$, also $gx = x$. Es ist $x = ex = (g^{-1}g)x = g^{-1}(gx) = g^{-1}x$, also auch $g^{-1} \in G_x$. QED

Beispiele. 1. Sei H eine Gruppe und $G \subseteq H$ eine Untergruppe. Dann ist

$$\begin{aligned} &G \times H \to H \\ &(g,x) \mapsto gx \qquad \text{(Produkt in } H\text{)} \end{aligned}$$

eine Operation von G auf H.

2. Sei G eine Gruppe. Dann ist

$$\begin{aligned} G \times G &\to G \\ (g, x) &\mapsto gxg^{-1} \end{aligned}$$

eine Operation von G auf G, sie wird auch *Konjugation* genannt.

BEWEIS. Seien $g, h \in G$ und $x \in G$. Es gilt $(e, x) \mapsto exe^{-1} = x$ und $((gh), x) \mapsto (gh)x(gh)^{-1} = ghxh^{-1}g^{-1} = g(hxh^{-1})g^{-1}$. QED

3. Sei V ein endlichdimensionaler euklidischer Vektorraum und G die euklidische Gruppe von V (oder eine Untergruppe davon). Dann ist

$$\begin{aligned} G \times V &\to V \\ (g, x) &\mapsto g(x) \end{aligned}$$

eine Operation von G auf V. Ist speziell $G = \mathrm{O}(V)$, so ist für ein $x \in V$ der Orbit von x

$$Gx := \{ gx \mid g \in G \} = \{ y \in V \mid \|y\| = \|x\| \}.$$

Für $\dim(V) = 2$ ist also der Orbit Gx der Kreis um den Nullpunkt mit Radius $\|x\|$, und für $\dim(V) = 3$ ist der Orbit Gx die Oberfläche der Kugel um den Nullpunkt mit Radius $\|x\|$.

SATZ. *Sei G eine Gruppe, S eine Menge und $G \times S \to S$ eine Operation. Dann gilt:*

1. *S ist disjunkte Vereinigung von Orbits.*
2. *Für endliches G, S und beliebiges $x \in S$ gilt*

$$\mathrm{Ord}(G) = \mathrm{Ord}(G_x) \cdot |Gx|.$$

Hierbei ist $|Gx|$ die Anzahl der Elemente in Gx.

BEWEIS. 1. a. Es ist $S = \bigcup_{x \in S} Gx$, denn für $x \in S$ gilt $x = ex$.

b. Es genügt, noch zu zeigen, daß aus $Gx \cap Gy \neq \emptyset$ folgt $Gx = Gy$. Seien also Gx, Gy gegeben mit $Gx \cap Gy \neq \emptyset$. Dann existieren $g, h \in G$ mit $gx = hy$, also $x = g^{-1}hy \in Gy$, woraus $Gx \subseteq Gy$ folgt. Analog zeigt man $Gy \subseteq Gx$. Also ist $Gx = Gy$.

2. Betrachte die surjektive Abbildung

$$\begin{aligned} \varphi \colon G &\to Gx \\ g &\mapsto gx. \end{aligned}$$

a. Sei $r := |Gx|$, etwa $Gx = \{f_1 x, \ldots, f_r x\}$ mit $f_i \in G$. Wir haben dann

$$\begin{aligned} G &= \bigcup_{i=1}^{r} \varphi^{-1}(f_i x), \\ \varphi^{-1}(f_i x) &\cap \varphi^{-1}(f_j x) = \emptyset \quad \text{für } i \neq j \end{aligned}$$

(dies gilt für jede Abbildung; $\varphi^{-1}(y)$ heißt *Faser* von $y \in Gx$).

b. Berechnung der Fasern. Sei $f \in G$. Dann ist

$$\varphi^{-1}(fx) = fG_x,$$

denn für $g \in G$ sind folgende Aussagen äquivalent.

$$\begin{aligned}
&g \in \varphi^{-1}(fx)\\
&gx = fx\\
&f^{-1}gx = x\\
&f^{-1}g \in G_x\\
&g \in fG_x.
\end{aligned}$$

c. Für ein $f \in G$ ist die Abbildung

$$\begin{aligned}
&G_x \to fG_x\\
&g \mapsto fg
\end{aligned}$$

bijektiv, denn sie hat die Umkehrabbildung $h \mapsto f^{-1}h$. Es folgt:

$$\begin{aligned}
\mathrm{Ord}(G) &= \sum_{i=1}^{r} \mathrm{Ord}(\varphi^{-1}(f_i x)) && \text{nach (a)}\\
&= \sum_{i=1}^{r} \mathrm{Ord}(f_i G_x) && \text{nach (b)}\\
&= r \cdot \mathrm{Ord}(G_x)\\
&= |Gx| \cdot \mathrm{Ord}(G_x) && \text{nach Definition von } r \text{ in (a).}
\end{aligned}$$
QED

SATZ. *(Satz von* LAGRANGE*). Sei H eine endliche Gruppe und $G \subseteq H$ eine Untergruppe. Es sei $[H : G]$ (der* Index *von G in H) die Anzahl der verschiedenen Teilmengen $Gx = \{\, gx \mid g \in G \,\}$, $x \in H$, von H. Dann gilt:*

$$\mathrm{Ord}(H) = \mathrm{Ord}(G) \cdot [H : G].$$

BEWEIS. Betrachte die Operation

$$\begin{aligned}
&G \times H \to H\\
&(g, x) \mapsto gx \qquad (\text{Produkt in } H).
\end{aligned}$$

Seien $Gx_1, \ldots, Gx_r$ die paarweise verschiedenen Orbits. Nach obigem Satz, Teil (1), gilt

$$H = \bigcup_{i=1}^{r} Gx_i \quad \text{und} \quad Gx_i \cap Gx_j = \emptyset \quad \text{für } i \neq j.$$

Die Abbildung $G \to Gx$, $g \mapsto gx$ ist bijektiv für alle $x \in H$, denn $y \mapsto yx^{-1}$ ist die Umkehrabbildung. Also ist $\mathrm{Ord}(H) = r \cdot \mathrm{Ord}(G)$. QED

FOLGERUNG. *Sei G eine endliche Gruppe mit $\mathrm{Ord}(G) = p$ und p prim. Dann ist G zyklisch.*

BEWEIS. Es ist $\mathrm{Ord}(G) = p >\geq 2$, also existiert ein $x \in G$ mit $x \neq e$. Also ist $\langle x \rangle$ Untergruppe von G mit $\mathrm{Ord}(\langle x \rangle) \neq 1$. Nach dem Satz 10.1.3 von LAGRANGE teilt $\mathrm{Ord}(\langle x \rangle)$ die Ordnung von G, und wegen p prim folgt $\mathrm{Ord}(\langle x \rangle) = \mathrm{Ord}(G)$. Also ist $\langle x \rangle = G$. QED

10.1.4. *Klassifikation der endlichen Untergruppen der* SO(3). Es folgt die Anwendung auf endliche Drehgruppen. Zunächst wird der Begriff der Diedergruppe eingeführt.

DEFINITION. Seien $t \in \mathbb{N}$, $t \geq 1$ und G eine Gruppe. G heißt *Diedergruppe* der Ordnung $2t$, wenn gilt:

1. $\mathrm{Ord}(G) = 2t$.
2. Es existieren $f, g \in G$ mit
$$\begin{aligned} f^t &= e \\ g^2 &= e \\ fg &= gf^{-1} \\ G &= \{e, f, \ldots, f^{t-1}, g, gf, \ldots, gf^{t-1}\}. \end{aligned}$$

(Der Nachweis der Existenz einer Diedergruppe ist Teil der Übungen.)

BEMERKUNGEN. 1. D_{2t} ist die Symmetriegruppe des regulären t-Ecks in der Ebene.

2. Sei G eine Diedergruppe der Ordnung $2t$, etwa $G = \{e, f, \ldots, f^{t-1}, g, gf, \ldots, gf^{t-1}\}$ mit g, f wie in der Definition. Dann hat jedes $h \in G$ eine *eindeutige* Darstellung
$$h = g^i f^j \quad \text{mit } 0 \leq i \leq 1,\ 0 \leq j \leq t-1.$$
Die Multiplikation in G hängt nur von den Paaren (i, j) ab, denn für $0 \leq i, k \leq 1$ und $0 \leq j, l \leq t-1$ gilt
$$g^i f^j g^k f^l = \begin{cases} g^i f^{j+l} & \text{falls } k = 0 \\ g^{i+1} f^{-j+l} & \text{falls } k = 1. \end{cases}$$
Also sind je zwei Diedergruppen der Ordnung $2t$ zueinander isomorph.

SATZ.

1. *Sei $G \subseteq \mathrm{SO}(2)$ eine endliche Untergruppe. Dann ist G zyklisch.*
2. *Sei $G \subseteq \mathrm{SO}(3)$ eine endliche Untergruppe. Dann gilt:*

   G ist zyklisch

   oder G ist Diedergruppe, d.h. $G \cong D_{2n}$ (Symmetriegruppe des regulären n-Ecks),

   oder $\mathrm{Ord}(G) \in \{12, 24, 60\}$.

BEMERKUNG. In (2) gilt sogar: Ist G nicht zyklisch und keine Diedergruppe, also $\mathrm{Ord}(G) = 12$, 24 bzw. 60, so ist $G \cong A_4$, S_4 bzw. A_5. (Dies wird hier nicht bewiesen.)

BEWEIS. 1. Sei $f \in \mathrm{SO}(\mathbb{R}^2)$. Dann existiert genau ein φ, $0 \leq \varphi < 2\pi$ mit
$$f = f_\varphi := \text{Drehung um } \varphi \text{ gegen den Uhrzeigersinn.}$$

Sei $\mathrm{Ord}(G) = n$, etwa $G = \{f_{\varphi_1}, \ldots, f_{\varphi_n}\}$ mit $0 \leq \varphi_i < 2\pi$. Wir können oBdA $\varphi_1 = 0$ annehmen. Sei $\varphi := \min\{\, \varphi_i \mid 2 \leq i \leq n\,\}$. Dann existiert für jedes i mit $2 \leq i \leq n$ ein $m \in \mathbb{N}$ und ein ψ_i mit $0 \leq \psi_i < \varphi$, so daß gilt:

$$\begin{aligned} &\varphi_i = m\varphi + \psi_i \\ &f_{\varphi_i} = f_{m\varphi+\psi_i} = f_\varphi^m f_{\psi_i} \\ &f_{\psi_i} \in G && \text{da } G \text{ Gruppe} \\ &\psi_i = 0 && \text{da } \psi_i < \varphi. \end{aligned}$$

Also ist $G = \langle f_\varphi \rangle = \{\, f_\varphi^0, f_\varphi, \ldots, f_\varphi^{n-1}\,\}$.

2. Jedes $f \in \mathrm{SO}(\mathbb{R}^3)$, $f \neq \mathrm{id}$, bildet die Oberfläche der Einheitskugel in sich selbst ab und hat dort genau zwei Fixpunkte (nämlich wo die Drehachse die Einheitskugel durchstößt). Sei

$$S := \{\, x \in \mathbb{R}^3 \mid \|x\| = 1 \text{ und es gibt ein } f \in G,\ f \neq \mathrm{id} \text{ mit } f(x) = x\,\}$$

die Menge der *Pole* von G. Wir können oBdA $\mathrm{Ord}(G) > 1$ annehmen, also ist $S \neq \emptyset$.

a. Wir zeigen, daß

$$\begin{aligned} G \times S &\to S \\ (g, x) &\mapsto g(x) \end{aligned}$$

eine wohldefinierte Operation ist. Zu zeigen ist, daß aus $g \in G$ und $x \in S$ folgt $g(x) \in S$. Wegen $x \in S$ gibt es ein $f \in G$, $f \neq \mathrm{id}$ mit $f(x) = x$. Also ist $gfg^{-1}g(x) = g(x)$, und es gilt $gfg^{-1} \in G$, $gfg^{-1} \neq \mathrm{id}$. Daraus folgt $g(x) \in S$.

b. Behauptung: Aus $x \in S$ und $g \in G$ folgt $G_{g(x)} = gG_xg^{-1}$, insbesondere also $\mathrm{Ord}(G_{g(x)}) = \mathrm{Ord}(G_x)$.

Beweis: Folgende Aussagen sind äquivalent.

$$\begin{aligned} &f \in G_{g(x)} \\ &fg(x) = g(x) \\ &g^{-1}fg(x) = x \\ &g^{-1}fg \in G_x \\ &f \in gG_xg^{-1}. \end{aligned}$$

Weiter gilt $\mathrm{Ord}(G_{g(x)}) = \mathrm{Ord}(G_x)$, da die Abbildung

$$\begin{aligned} G &\to G \\ h &\mapsto ghg^{-1} \end{aligned}$$

bijektiv ist und gG_xg^{-1} das Bild von G_x unter dieser Abbildung ist.

c. Seien $Gx_1, \ldots, Gx_r$ die paarweise verschiedenen Orbits bzgl. dieser Operation (S ist also disjunkte Vereinigung von $Gx_1, \ldots, Gx_r$). Nach dem Satz gilt für jedes i mit $1 \leq i \leq r$

$$\underbrace{\mathrm{Ord}(G)}_{=:n} = \underbrace{|Gx_i|}_{=:m_i} \cdot \underbrace{\mathrm{Ord}(G_{x_i})}_{=:n_i}.$$

Sei etwa

$$Gx_i = \{f_1(x_i), \ldots, f_{m_i}(x_i)\},$$

Die Faser von jedem $f_j(x_i)$ ist dann eine Teilmenge von G mit $n_i = \mathrm{Ord}(G_{x_i})$-vielen Elementen.

Unser Ansatz besteht nun darin, alle Paare (f, x) mit $f \in G$, $f \neq \mathrm{id}$ und x Pol von f zu zählen.

- Einerseits gibt es $n-1$ Möglichkeiten für f mit jeweils 2 Polen.
- Andererseits: Sei i fest gewählt. Betrachte $f_j(x_i) \in Gx_i$. Wieviele $h \in G$ mit dem Pol $f_j(x_i)$ gibt es? Soviele wie
$$\operatorname{Ord}(\underbrace{\{\, h \in G \mid hf_j(x_i) = f_j(x_i)\,\}}_{G_{f_j(x_i)}} = \operatorname{Ord}(G_{x_i}) = n_i \quad \text{nach (b)}.$$
Bei festem i haben wir also m_i Möglichkeiten für x, mit jeweils $n_i - 1$ zugehörigen f.

Dies ergibt
$$\begin{aligned} 2(n-1) &= \sum_{i=1}^{r} m_i(n_i - 1) \\ 2 - \frac{2}{n} &= \sum_{i=1}^{r} \frac{m_i(n_i-1)}{n} = \sum_{i=1}^{r}(1 - \frac{1}{n_i}) \quad \text{wegen } n = m_i n_i. \end{aligned}$$
Wir diskutieren nun die Gleichung
$$\boxed{2 - \frac{2}{n} = \sum_{i=1}^{r}(1 - \frac{1}{n_i})} \quad (\text{oBdA } n_i \geq 2). \tag{10.1}$$
d. Behauptung: $r = 2$ oder $r = 3$.

Beweis: Unmöglich ist der Fall $r = 1$. Annahme: $r \geq 4$. Dann folgt aus (10.1) der Widerspruch:
$$2 - \frac{2}{n} = \sum_{i=1}^{r}(1 - \frac{1}{n_i}) \geq 4 - \frac{4}{2} = 2.$$
Fall $r = 2$. Aus (10.1) folgt
$$\begin{aligned} &2 - \frac{2}{n} = 2 - \frac{1}{n_1} - \frac{1}{n_2} \\ &\frac{2}{n} = \frac{1}{n_1} + \frac{1}{n_2} && (n_1, n_2 \leq n) \\ &n_1 = n_2 = n \quad \text{und damit} \quad m_1 = m_2 = 1. \end{aligned}$$
Also gibt es genau 2 Pole und jeder Pol kommt n-mal vor. Sei $x \in S$, also ein Pol. Man erhält
$$\begin{aligned} G &= \{\, g \in \mathrm{SO}(\mathbb{R}^3) \mid g(x) = x \,\} \\ &\cong \text{ endliche Untergruppe von } \mathrm{SO}(\mathbb{R}^2) && (\text{Ebene: } (\mathbb{R}x)^{\perp}) \\ G &\quad \text{zyklisch} && \text{nach (1)}. \end{aligned}$$
Fall $r = 3$. OBdA gelte $2 \leq n_1 \leq n_2 \leq n_3 \leq n$. Nach (10.1) haben wir
$$2 - \frac{2}{n} = 3 - \sum_{i=1}^{3} \frac{1}{n_i}.$$
Annahme $n_1 \geq 3$. Dann erhalten wir den Widerspruch
$$2 - \frac{2}{n} = 3 - \underbrace{\sum_{i=1}^{3} \frac{1}{n_i}}_{\leq 1} \geq 2.$$

Also ist $n_1 = 2$ und damit

$$2 - \frac{2}{n} = 3 - \frac{1}{2} - \frac{1}{n_2} - \frac{1}{n_3},$$
$$\frac{1}{n_2} + \frac{1}{n_3} = \frac{1}{2} + \frac{2}{n}.$$

Annahme: $n_2 \geq 4$. Dann haben wir $\frac{1}{n_2} + \frac{1}{n_3} \leq \frac{2}{4} = \frac{1}{2}$ und damit einen Widerspruch. Also ist $n_2 = 2$ oder $n_2 = 3$. Es ergeben sich die folgenden Möglichkeiten:

i. $n_1 = 2$ und $n_2 = 2$. Dann ist $n_3 = \frac{n}{2}$ und n gerade.

ii. $n_1 = 2$, $n_2 = 3$ und $n_3 = 3$. Dann gilt $\frac{2}{3} = \frac{1}{2} + \frac{2}{n}$, also $\frac{4}{6} - \frac{3}{6} = \frac{2}{n}$, also $n = 12$.

iii. $n_1 = 2$, $n_2 = 3$ und $n_3 = 4$. Dann gilt $\frac{1}{3} + \frac{1}{4} = \frac{1}{2} + \frac{2}{n}$, also $\frac{4}{12} - \frac{3}{12} = \frac{2}{n}$, also $n = 24$.

iv. $n_1 = 2$, $n_2 = 3$ und $n_3 = 5$. Dann folgt $\frac{1}{3} + \frac{1}{5} = \frac{1}{2} + \frac{2}{n}$, also $\frac{10}{30} + \frac{6}{30} - \frac{15}{30} = \frac{2}{n}$, also $n = 60$.

Andere Fälle sind nicht möglich, denn es gilt

$$\begin{aligned} &\frac{1}{3} + \frac{1}{n_3} = \frac{1}{2} + \frac{2}{n} \\ &\frac{1}{n_3} = \frac{1}{6} + \frac{2}{n} > \frac{1}{6} \\ &n_3 < 6 \qquad\qquad (\text{außerdem } n_2 = 3 \leq n_3). \end{aligned}$$

Es bleibt noch zu zeigen, daß im Fall (i) G eine Diedergruppe ist.

e. Behauptung: Aus $r = 3$, $n_1 = 2$, $n_2 = 2$, $n_3 = \frac{n}{2}$ folgt $G \cong D_{2 \cdot \frac{n}{2}}$.

Beweis: Es sei $S = Gx_1 \cup Gx_2 \cup Gx_3$ (disjunkte Vereinigung) und $m_1 = \frac{n}{2}$, $m_2 = \frac{n}{2}$ und $m_3 = \frac{n}{\frac{n}{2}} = 2$. Es ist $|Gx_3| = 2$, sei also $Gx_3 =: \{x, y\}$. Damit erhält man

$$G_x = \{\, f \in G \mid f(x) = x \,\} \cong \text{ endliche Untergruppe von } \mathrm{SO}(\mathbb{R}^2) \text{ (Ebene: } (\mathbb{R}x)^{\perp}).$$

Also ist nach (1) $G_x = \langle f \rangle$ für ein $f \in G$. Wähle ein $g \in G$ mit $g \notin G_x$ (möglich wegen $\mathrm{Ord}(G_x) = \frac{n}{2}$). Dann folgt mit $t := \frac{n}{2}$:

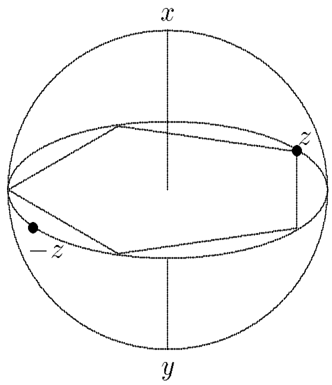
$$G = \{f^0 = \mathrm{id}, f, \ldots, f^{t-1}, g, gf, \ldots, gf^{t-1}\}, \quad G_x = \{f^0, f, \ldots, f^{t-1}\}.$$

Außerdem folgt für ein $g \in G$ mit $g \notin G_x$ (d.h. $g(x) \neq x$), daß $g^2 = \mathrm{id}$. Dies sieht man wie folgt. Zunächst gilt für jedes $g \in G$

$$g(Gx_3) = gGx_3 = Gx_3,$$

d.h. $g(\{x, y\}) = \{x, y\}$, d.h. $g(x) = x$ und $g(y) = y$, oder $g(x) = y$ und $g(y) = x$. Daraus folgt im Fall $g(x) \neq x$, daß $g(x) = y$ und damit $g^2(x) = g(y) = x$. Ebenso folgt im Fall $g(y) \neq y$, daß $g^2(y) = y$. Da g eine Drehung mit $g(x) \neq x$ und $g(y) \neq y$ ist, besitzt g einen Pol z mit $g(z) = z$ und $z \neq x, y$. Dann ist auch $g^2(z) = z$, also ist $g^2 = \mathrm{id}$, da g^2 drei Fixpunkte hat.

Ergebnis: $G = \langle f, g \rangle$, $f^t = \mathrm{id}$, $f^s \neq \mathrm{id}$ für $s < t$, $g^2 = \mathrm{id}$ und $fg = gf^{-1}$ da $(gf)^2 = gfgf = \mathrm{id}$. Also ist $G \cong D_{2t}$ mit $t \geq 1$. QED



10.2 Faktorgruppen

Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Die Faktorgruppe G/H ist dann wie der Quotientenvektorraum V/U in 2.3.5 definiert. Wir konstruieren die Faktorgruppe G/N für einen Normalteiler $N \subseteq G$, zusammen mit dem natürlichen Gruppenhomomorphismus $\mathrm{kan}\colon G \to G/N$. Es zeigt sich, daß hierfür die Eigenschaft von N, Normalteiler zu sein, notwendig ist. Als einfache Folgerung erhalten wir eine Charakterisierung von Normalteilern als Kerne von Gruppenhomomorphismen. Wir beweisen dann die universelle Eigenschaft von Faktorgruppen und als Folgerung den Homomorphiesatz (vgl. den Homomorphiesatz für Vektorräume in 3.2.2). Als einfache Anwendung erhalten wir eine Klassifikation der zyklischen Gruppen und für endliche Gruppen den Satz von Fermat, daß $x^{\mathrm{Ord}(G)} = e$ ist. Ferner ergibt sich $x^p \equiv x \mod p$ für $x \in \mathbb{Z}$ und Primzahlen p.

10.2.1. *Nebenklassen, Konstruktion der Faktorgruppe.* Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Für $x \in G$ heißt
$$xH := \{\, xh \mid h \in H \,\}$$
die von x erzeugte *Linksnebenklasse* bzgl. H. (Entsprechend heißt $Hx := \{\, hx \mid h \in H \,\}$ *Rechtsnebenklasse*). Die Menge der Linksnebenklassen von G bzgl. H, also
$$G/H := \{\, xH \mid x \in G \,\},$$
wird G *modulo* H genannt.

Bemerkungen. 1. Sei G eine Gruppe, $H \subseteq G$ eine Untergruppe und $x, y \in G$. Dann gilt
$$xH = yH \text{ genau dann, wenn } y^{-1}x \in H.$$

Beweis. $\Longrightarrow$. Gelte $xH = yH$. Dann existiert ein $h \in H$ mit $x = xe = yh$. Also $y^{-1}x = h \in H$.

$\Longleftarrow$. Sei $y^{-1}x \in H$. Dann gibt es ein $h \in H$ mit $y^{-1}x = h$, also $x = yh$. Damit gilt für alle $h' \in H$, daß $xh' = yhh' \in yH$, also $xH \subseteq yH$. Wegen $(y^{-1}x)^{-1} = x^{-1}y$ folgt genauso $yH \subseteq xH$. Qed

2. Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Für $x, y \in G$ definieren wir
$$x \sim_H y :\equiv y^{-1}x \in H.$$
Nach (1) ist $\sim_H$ eine Äquivalenzrelation auf G. Wir zeigen
$$xH = \{\, y \mid y \sim_H x \,\} =: [x] \quad \text{Äquivalenzklasse von } x.$$

Beweis. $\subseteq$. Sei $h \in H$. Zu zeigen ist $xh \sim_H x$, d.h. $x^{-1}xh \in H$. $\supseteq$. Sei $y \sim_H x$, also $y^{-1}x \in H$. Zu zeigen ist $y \in xH$. Es existiert ein $h \in H$ mit

$$\begin{aligned} y^{-1}x &= h \\ x^{-1}y &= h^{-1} \\ y &= xh^{-1} \in xH. \quad \text{Qed} \end{aligned}$$

Nach Abschnitt 1.2 (den Bemerkungen zu Äquivalenzklassen) ist also G die disjunkte Vereinigung von Linksnebenklassen.

3. Sei G eine Gruppe und $H \subseteq G$ eine Untergruppe. Wähle in jeder Linksnebenklasse ein Element (Repräsentant) x_i aus. Dann ist $G = \bigcup_{i \in I} x_i H$ (disjunkte Vereinigung) wobei $(x_i)_{i \in I}$ die Familie der Repräsentanten ist. Es ist dann auch $G = \bigcup_{i \in I} Hx_i^{-1}$ disjunkte Vereinigung von Rechtsnebenklassen, denn $\varphi\colon G \to G$, $x \mapsto x^{-1}$ ist bijektiv und es gilt $\varphi(xH) = Hx^{-1}$.

Beispiele. 1. Sei V ein Vektorraum über einen Körper K und $U \subseteq V$ ein Untervektorraum. Faßt man V als additive Gruppe und U als Untergruppe auf, so ist

$$V/U = \{\, x + U \mid x \in V \,\}.$$

2. Sei $n \in \mathbb{N}$. Faßt man $\mathbb{Z}$ auf als additive Gruppe und definiert man

$$\begin{aligned} n\mathbb{Z} &= \{\, nz \mid z \in \mathbb{Z} \,\} \\ &= \{0, \pm n, \pm 2n, \pm 3n, \dots\} \\ &= \text{die von } n \text{ erzeugte Untergruppe,} \end{aligned}$$

so gilt: $\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \dots, [n-1]\}$ hat n Elemente.

Beweis. a. Für alle $x \in \mathbb{Z}$ existieren $z, r \in \mathbb{Z}$ mit $x = zn + r$ und $0 \le r < n$ (Division mit Rest). Also ist $[x] = [r]$.

b. Noch zu zeigen: Für alle $0 \le r, s < n$ mit $r \ne s$ ist $[r] \ne [s]$. Sei dazu oBdA $r > s$. Annahme: $[r] = [s]$. Dann existiert ein $z \in \mathbb{Z}$ mit $r - s = n > 0$ und $r - s < n$. Also folgt $0 < r - s < nz$, was ein Widerspruch zu $z \in \mathbb{Z}$ ist. Qed

Schreibweise. Statt $x - y \in n\mathbb{Z}$ schreibt man $x \equiv y \mod n$ und sagt "x ist kongruent y modulo n". Rechnet man mit $\equiv$ statt mit $=$, so bedeutet dies das Rechnen in $\mathbb{Z}/n\mathbb{Z}$ statt in $\mathbb{Z}$.

10.2.2. Satz. *(Konstruktion der Faktorgruppe). Sei G eine Gruppe und $N \subseteq G$ Normalteiler. Definiere*

$$\begin{aligned} \operatorname{kan}\colon G &\to G/N \\ x &\mapsto [x]. \end{aligned}$$

Dann gibt es auf G/N genau eine Gruppenstruktur, so daß $\operatorname{kan}\colon G \to G/N$ *ein Gruppenhomomorphismus wird. In diesem Fall gilt* $\operatorname{Kern}(\operatorname{kan}) = N$.

Beweis. *Existenz.* Für alle $x, y \in G$ setzen wir

$$[x] \cdot [y] := [xy] \quad \text{Produkt in } G/N.$$

Wir zeigen zunächst, daß die Multiplikation wohldefiniert ist. Seien also $x, x', y, y' \in G$ mit $[x] = [x']$ und $[y] = [y']$ gegeben. Zu zeigen ist $[xy] = [x'y']$, d.h. die obige Definition des Produkts in G/N ist unabhängig vom Repräsentanten von $[x]$, $[y]$. Dies folgt aus

$$\begin{aligned}
(x'y')^{-1}(xy) &= y'^{-1} \underbrace{x'^{-1} x}_{\in N} y \\
&= y'^{-1} y n && \text{für ein } n \in N \text{ wegen } Ny = yN, \text{ denn } N \text{ ist Normalteiler} \\
&\in Nn && \text{wegen } [y] = [y'] \\
&= N.
\end{aligned}$$

Also ist $[x'y'] = [xy]$.

Wir zeigen jetzt, daß die Multiplikation assoziativ ist. Dies folgt aus

$$[x]([y][z]) = [x]([yz]) = [x(yz)] = [(xy)z] = [xy][z] = ([x][y])[z].$$

Ferner ist $[e]$ das neutrale Element, denn es gilt

$$[e][x] = [ex] = [x] \quad \text{für alle } x \in G.$$

Schließlich ist $[x^{-1}]$ das Inverse zu $[x]$, denn es gilt

$$[x^{-1}][x] = [x^{-1}x] = [e].$$

Demnach ist G/N mit obiger Multiplikation eine Gruppe. kan ist offenbar Gruppenhomomorphismus und es gilt $\mathrm{Kern}(\mathrm{kan}) = N$.

Eindeutigkeit. Eine Verknüpfung $\circ$ auf G/N, so daß kan Gruppenhomomorphismus ist, muß die Bedingung $xN \circ yN = xyN = xyNN = xN \cdot yN$ erfüllen. QED

BEMERKUNGEN. 1. Sei G eine Gruppe und $N \subseteq G$ Untergruppe. Sei noch

$$\begin{aligned}
\mathrm{kan}\colon G &\to G/N \\
x &\mapsto [x]
\end{aligned}$$

Gruppenhomomorphismus bzgl. irgendeiner (fest gewählten) Gruppenstruktur auf G/N. Dann ist N Normalteiler.

BEWEIS. Es ist $\mathrm{kan}(e) = [e]$ neutrales Element in G/N, da kan Gruppenhomomorphismus ist. Ferner ist

$$\mathrm{Kern}(\mathrm{kan}) = \{\, x \in G \mid [x] = [e] \,\} = N.$$

Also ist N Normalteiler, denn nach der folgenden Bemerkung gilt dies stets für Kerne von Gruppenhomomorphismen. QED

2. Sei G eine Gruppe und $N \subseteq G$ Untergruppe. Dann ist N Normalteiler von G genau dann, wenn es eine Gruppe G' und einen Gruppenhomomorphismus $f\colon G \to G'$ gibt mit $N = \mathrm{Kern}(f)$.

BEWEIS. $\Longleftarrow$. Es ist $N = \mathrm{Kern}(f) = f^{-1}(\{e\})$, und $\{e\}$ ist Normalteiler. $\Longrightarrow$ folgt aus dem eben bewiesenen Satz. QED

3. Sei $G = \mathbb{Z}$ und $N = n\mathbb{Z}$ für ein $n \in N$. Dann gilt für alle $x, y, z \in \mathbb{Z}$

$$\begin{aligned}
[x] + [y] = [z] &\quad \text{genau dann, wenn} \quad x + y = z + nu \quad \text{für ein } u \in \mathbb{Z} \\
[y] = -[x] &\quad \text{genau dann, wenn} \quad x + y \in n\mathbb{Z}.
\end{aligned}$$

Ist zum Beispiel $n = 4$, so hat man

$$\begin{aligned} &\mathbb{Z}/4\mathbb{Z} = \{[0],[1],[2],[3]\}, \\ &[1]+[2] = [3], \\ &[2]+[3] = [5] = [1], \\ &-[1] = [3]. \end{aligned}$$

10.2.3. SATZ. *(Universelle Eigenschaft der Faktorgruppe). Seien G, H Gruppen, $N \subseteq G$ Normalteiler und $f\colon G \to H$ Gruppenhomomorphismus mit $N \subseteq \mathrm{Kern}(f)$. Dann gibt es genau einen Gruppenhomomorphismus $g\colon G/N \to H$ mit $g \circ \mathrm{kan} = f$.*

$$\begin{tikzcd} G \arrow[rr, "f"] \arrow[dr, "\mathrm{kan}"'] & & H \\ & G/N \arrow[ur, "g"'] & \end{tikzcd}$$

BEWEIS. *Existenz.* Für $x \in G$ definieren wir $g([x]) := f(x)$.

a. g ist wohldefiniert, denn für $x, y \in G$ mit $[x] = [y]$ gilt

$$\begin{aligned} &y^{-1}x \in N \\ &x = yn \quad \text{für ein } n \in N \\ &f(x) = f(yn) = f(y)\underbrace{f(n)}_{=e} = f(y). \end{aligned}$$

Also ist $g([x]) := f(x)$ unabhängig von der Wahl des Repräsentanten x.

b. g ist Gruppenhomomorphismus, denn für $x, y \in G$ gilt

$$g([x][y]) = g([xy]) = f(xy) = f(x)f(y) = g([x])g([y]).$$

Eindeutigkeit. Für eine Abbildung g mit den gewünschten Eigenschaften muß gelten $g([x]) = (g \circ \mathrm{kan})(x) = f(x)$ für alle $x \in G$. QED

BEMERKUNG. Seien G, H Gruppen und $N \subseteq G$ Normalteiler. Dann ist die Abbildung

$$\begin{aligned} \{\, g\colon G/N \to H \mid g \text{ Homomorphismus} \,\} &\to \{\, f\colon G \to H \mid f \text{ Homomorphismus mit } f(N) = \{e\} \,\} \\ g &\mapsto g \circ \mathrm{kan} \end{aligned}$$

offenbar eine Bijektion.

FOLGERUNG. *(Homomorphiesatz). Seien G, H Gruppen, $f\colon G \to H$ surjektiver Gruppenhomomorphismus und $\mathrm{kan}\colon G \to G/\mathrm{Kern}(f)$ wie in Satz 10.2.2 definiert. Dann gibt es genau eine Abbildung $\overline{f}\colon G/\mathrm{Kern}(f) \to H$ mit $\overline{f} \circ \mathrm{kan} = f$. Dieses eindeutig bestimmte $\overline{f}$ ist ein Isomorphismus.*

$$\begin{tikzcd} G \arrow[rr, "f"] \arrow[dr, "\mathrm{kan}"'] & & H \\ & G/\mathrm{Kern}(f) \arrow[ur, "\overline{f}"'] & \end{tikzcd}$$

Beweis. Nach dem Satz ist $\overline{f}$ ein wohldefinierter Gruppenhomomorphismus. $\overline{f}$ ist surjektiv, denn zu jedem $y \in H$ gibt es ein $x \in G$ mit $y = f(x) = \overline{f}([x])$. $\overline{f}$ ist injektiv, denn für $x \in G$ mit $\overline{f}([x]) = e$ gilt $f(x) = \overline{f}([x]) = e$, also $x \in \mathrm{Kern}(f)$ und damit $[x] = [e]$. QED

10.2.4. *Anwendungen: zyklische Gruppen, der Satz von* Fermat. Als erste einfache Anwendung erhalten wir eine Übersicht über alle zyklischen Gruppen.

Folgerung. *Sei G eine zyklische Gruppe. Dann gilt*

$$G \cong \begin{cases} \mathbb{Z}, & \textit{falls } \mathrm{Ord}(G) = \infty \\ \mathbb{Z}/m\mathbb{Z}, & \textit{falls } \mathrm{Ord}(G) = m < \infty. \end{cases}$$

Beweis. 1. Sei G zyklisch, etwa $G = \langle x \rangle$ für ein $x \in G$. Dann ist

$$\begin{aligned} f\colon \mathbb{Z} &\to G \\ i &\mapsto x^i \end{aligned}$$

ein (offenbar surjektiver) Gruppenhomomorphismus, denn es gilt

$$f(i+j) = x^{i+j} = x^i x^j = f(i)f(j).$$

2. Noch zu zeigen ist

$$\mathrm{Kern}(f) = \begin{cases} 0 & \text{falls } \mathrm{Ord}(G) = \infty \\ n\mathbb{Z} & \text{falls } \mathrm{Ord}(G) = n < \infty. \end{cases}$$

Zum Beweis beachte man, daß $\mathrm{Kern}(f)$ Untergruppe von $\mathbb{Z}$ ist, also nach der folgenden Bemerkung von der Form $m\mathbb{Z}$ für ein $m \in \mathbb{Z}$ mit $m \geq 0$ ist. Im Fall $m \neq 0$ ist $G \cong \mathbb{Z}/\mathrm{Kern}(f) = \mathbb{Z}/m\mathbb{Z}$, und im Fall $m = 0$ ist $f\colon \mathbb{Z} \to G$ ein Isomorphismus. QED

Bemerkung. Wir haben noch nachzutragen, daß jede Untergruppe $N \subseteq \mathbb{Z}$ von der Form $N = m\mathbb{Z}$ ist für ein $m \in \mathbb{Z}$ mit $m \geq 0$. Zum Beweis können wir annehmen, daß $N \neq \{0\}$ ist. Sei m das kleinste positive Element von N. Für ein beliebiges $z \in N$ gibt es dann $r, s \in \mathbb{Z}$ mit $0 \leq r < m$ so daß $z = ms + r$, also $r = z - ms \in N$ und damit $r = 0$, denn m war minimal gewählt. Also ist $z \in m\mathbb{Z}$ und damit $N = m\mathbb{Z}$.

Als weitere Folgerung ergibt sich der Satz von Fermat.

Folgerung. *(Satz von* Fermat*). Sei G eine endliche Gruppe der Ordnung n und $x \in G$. Dann ist $x^n = e$.*

Beweis. $\langle x \rangle \subseteq G$ ist eine Untergruppe. Nach dem Satz 10.1.3 von Lagrange gilt

$$n = \mathrm{Ord}(G) = \underbrace{\mathrm{Ord}\langle x \rangle}_{=:m} \cdot \underbrace{[G : \langle x \rangle]}_{=:r}.$$

Nach der vorangehenden Folgerung ist $\langle x \rangle \cong \mathbb{Z}/m\mathbb{Z}$ für ein $m \in \mathbb{Z}$ mit $m > 0$, also $x^m = e$, denn in $\mathbb{Z}/m\mathbb{Z}$ gilt

$$m \cdot \overline{a} = \underbrace{\overline{a} + \overline{a} + \cdots + \overline{a}}_{m\text{-mal}} = \overline{m \cdot a} = 0,$$

also $x^n = x^{mr} = (x^m)^r = e$. QED

Sei G eine Gruppe und $x \in G$. Dann nennt man $\mathrm{Ord}(x) := \mathrm{Ord}(\langle x \rangle)$ die *Ordnung* von x in G.

BEMERKUNGEN. 1. Sei G eine Gruppe und $x \in G$ mit $\mathrm{Ord}(x) = n$. Dann gilt

a. $\langle x \rangle = \{x^0, x^1, \ldots, x^{n-1}\}$.

b. $n = \min\{ i \mid i > 0, x^i = e \}$.

c. Für alle $m \in \mathbb{N}$ gilt $x^m = e$ genau dann, wenn n (also die Ordnung von x) die Zahl m teilt.

BEWEIS. Nach der Folgerung ist $\langle x \rangle \cong \mathbb{Z}/n\mathbb{Z}$ vermöge $x \mapsto [1]$. Wegen

$$\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \ldots, [n-1]\}$$

gelten (a) und (b). Für (c) ist die Richtung $\Longleftarrow$ klar. $\Longrightarrow$. Sei $x^m = e$. Dann gibt es $q, z \in \mathbb{Z}$ mit $0 \le r < n$ so daß $m = qn + r$, also $e = x^m = x^{qn+r} = (x^n)^q \cdot x^r = e \cdot x^r$. Daraus folgt $r = 0$, denn n ist die kleinste positive Zahl mit $x^n = e$. QED

2. Sei p Primzahl. Dann ist (wie unten bewiesen wird) $\mathbb{Z}/p\mathbb{Z}$ ein Körper mit der folgenden Struktur.

$$[x] + [y] = [x + y],$$
$$[x] \cdot [y] = [x \cdot y].$$

Also ist $G = \{ [x] \in \mathbb{Z}/p\mathbb{Z} \mid [x] \neq 0 \}$ eine multiplikative Gruppe der Ordnung $p - 1$. Nach dem Satz von FERMAT gilt dann

$$\begin{aligned} &[x]^{p-1} = [1] && \text{für alle } x \in \mathbb{Z} \text{ mit } [x] \neq 0 \\ &[x]^p = [x] && \text{für alle } x \in \mathbb{Z} \\ &x^p \equiv x \mod p && \text{für alle } x \in \mathbb{Z}. \end{aligned}$$

Nachzutragen bleibt, daß für Primzahlen p stets $\mathbb{Z}/p\mathbb{Z}$ ein Körper ist. Offenbar genügt es zu zeigen, daß es zu jedem $x \in \mathbb{Z}$ mit $[x] \neq 0$ ein $y \in \mathbb{Z}$ mit $[y] \neq 0$ gibt so daß $[x][y] = [1]$ ist. Sei also $x \in \mathbb{Z}$ mit $[x] \neq 0$ gegeben. Dann ist der größte gemeinsame Teiler (x, p) von x und p gleich 1. Man betrachte

$$\{ ax + bp \mid a, b \in \mathbb{Z} \} =: G.$$

Offenbar ist $G \subseteq \mathbb{Z}$ eine Untergruppe, also $G = m\mathbb{Z}$ für ein $m > 0$. Es ist $m = 1$, denn aus $x = mz_1$ und $p = mz_2$ für $z_1, z_2 \in \mathbb{Z}$ folgt $m|x$ und $m|p$, also $m = 1$ wegen $(x, p) = 1$. Also ist $ax + bp = 1$ für geeignete $a, b \in \mathbb{Z}$, also $[a][x] + [0] = [1]$. Es folgt $[a] \neq 0$.

10.3 Moduln, Satz von Cayley-Hamilton

Wir definieren den Begriff eines Moduls, als eine Art Vektorraum über einem Ring statt einem Körper. Als Anwendung beweisen wir den Satz von CAYLEY/HAMILTON, welcher aussagt, daß das "Einsetzen" einer Matrix in ihr eigenes charakteristisches Polynom stets die Nullmatrix ergibt. Ferner führen wir den Begriff des Minimalpolynoms ein. Schliesslich stellen wir einige Hilfsmittel bereit (Faktormoduln, frei erzeugte Moduln, Basis und lineare Forsetzung für Moduln), die zur Einführung des Tensorprodukts im folgenden Abschnitt gebraucht werden. – In diesem Abschnitt sei R ein nicht notwendig kommutativer Ring mit Einselement.

10.3.1. DEFINITION. Sei M eine abelsche Gruppe und $m\colon R \times M \to M$ eine Abbildung; wir schreiben rx für $m(r,x)$. (M,m) oder kurz M heißt *R-Linksmodul*, wenn für alle $r,s \in R$ und alle $x,y \in M$ gilt

1. $r(sx) = (rs)x$ (Assoziativgesetz).
2. $1x = x$ (1 ist linksneutrales Element).
3. $(r+s)x = rx + sx$ und $r(x+y) = rx + ry$ (Distributivgesetze).

BEMERKUNGEN. 1. Eine analoge Definition trifft man für *Rechtsmoduln*. Im folgenden verstehen wir unter einem *Modul* stets einen Linksmodul.

2. Einen Modul kann man ansehen als eine Art Vektorraum über einem Ring (statt eines Körpers). Einfache Rechenregeln beweist man genauso wie bei Vektorräumen, etwa

$$\begin{aligned} 0x &= 0, \\ r0 &= 0, \\ (-1)x &= -x. \end{aligned}$$

DEFINITION. Seien M, M' Moduln über R und $f\colon M \to M'$ eine Abbildung. f heißt *R-Modulhomomorphismus*, wenn für alle $x,y \in M$ und $r \in R$ gilt

$$\begin{aligned} f(x+y) &= f(x) + f(y), \\ f(rx) &= rf(x). \end{aligned}$$

BEISPIELE. 1. Jeder Vektorraum V über einem Körper K ist ein K-Modul.

2. Sei R der Ring $\mathbb{Z}$ der ganzen Zahlen und M eine abelsche Gruppe. Dann kann man M als $\mathbb{Z}$-Modul auffassen, und zwar mit folgendem Skalarprodukt. Sei $n \in Z$ und $x \in M$. Im Fall $n \geq 0$ setzen wir

$$nx = \underbrace{x + x + \cdots + x}_{n\text{-mal}},$$

und im Fall $n < 0$ sei $nx := -(-n)x$. Die Eigenschaften verifiziert man durch einfaches Nachrechnen (mit Fallunterscheidungen).

3. Sei V Vektorraum über einem Körper K und R der Endomorphismenring $\operatorname{End}_K(V)$. Dann ist V ein R-(Links)modul mittels

$$f \cdot x = f(x).$$

4. Sei V ein K-Modul über einem kommutativen Ring K und $f\colon V \to V$ ein K-Modulhomomorphismus. Wir definieren V_f wie folgt als $K[t]$-Modul. $V_f := V$ als abelsche Gruppe. Ferner wird das Skalarprodukt

$$K[t] \times V \to V \quad \text{definiert durch}$$
$$\Big(\sum_{i=0}^{n} \alpha_i t^i\Big) \cdot v = \sum_{i=0}^{n} \alpha_i f^i(v).$$

Wir zeigen jetzt, daß V_f ein $K[t]$-Modul ist. Seien also $r := \sum_{i=0}^n \alpha_i t^i$ und $s := \sum_{j=0}^n \beta_j t^j$ aus $K[t]$, ferner $v, w \in V$. Dann gilt

$$\begin{aligned} r(sv) &= \Big(\sum_{i=0}^n \alpha_i t^i\Big)\Big(\sum_{j=0}^m \beta_j f^j(v)\Big) \\ &= \sum_{i=0}^n \alpha_i f^i\Big(\sum_{j=0}^m \beta_j f^j(v)\Big) \\ &= \sum_{i,j} \alpha_i\beta_j f^{i+j}(v) \\ &= \sum_l \Big(\sum_{i+j=l} \alpha_i\beta_j\Big) f^l(v) \\ &= (rs)v. \end{aligned}$$

Ferner gilt $1v = f^0(v) = v$, und $r(v+w) = \sum_i \alpha_i f^i(v+w) = \sum_i \alpha_i(f^i(v) + f^i(w)) = rv + rw$ sowie (wobei wir oBdA $n = m$ annehmen), und auch

$$(r+s)v \;= \sum_{i=0}^n (\alpha_i + \beta_i) f^i(v) = \sum_{i=0}^n \alpha_i f^i(v) + \sum_{i=0}^n \beta_i f^i(v) = rv + sv.$$

Bemerkung. Die Skalarmultiplikation im $K[t]$-Modul V_f, also

$$\Big(\sum_{i=0}^n \alpha_i t^i\Big) \cdot v := \sum_{i=0}^n \alpha_i f^i(v)$$

kann man auffassen als Abbildung

$$\begin{aligned} &\Phi_f \colon K[t] \to \mathrm{End}_K(V) \\ &\Phi_f(r)(v) = rv. \end{aligned}$$

Dieses Φ_f ist ein K-Vektorraumhomomorphismus und auch ein Ringhomomorphismus, denn es gilt

$$\Phi_f(rs)(v) = (rs)v = r(sv) = \Phi_f(r)\left(\Phi_f(s)(v)\right) = \left(\Phi_f(r) \circ \Phi_f(s)\right)(v),$$

also $\Phi_f(rs) = \Phi_f(r) \circ \Phi_f(s)$.

5. Sei K ein Ring. Dann ist $M := K^n = \{ \left(\begin{smallmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{smallmatrix}\right) \mid \alpha_1, \ldots, \alpha_n \in K \}$ ein K-Modul mit folgender Addition und Skalarmultiplikation.

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} := \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix},$$

$$\alpha \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} := \begin{pmatrix} \alpha\alpha_1 \\ \vdots \\ \alpha\alpha_n \end{pmatrix}.$$

10.3.2. *Satz von* CAYLEY/HAMILTON. Als Anwendung des eben konstruierten Moduls V_f oder des entsprechenden Ringhomomorphismus $\Phi_f\colon K[t] \to \mathrm{End}_K(V)$ wollen wir jetzt eine wichtige Eigenschaft des charakteristischen Polynoms beweisen.

Sei also K ein Körper, V ein n-dimensionaler K-Vektorraum und f ein Endomorphismus, also $f\colon V \to V$ linear. Das charakteristische Polynom p_f von f war in 7.2.3 definiert als

$$p_f := p_A := \det(A - tE) \in K[t],$$

wobei $A \in K^{n\times n}$ eine (beliebige) darstellende Matrix von f ist.

Wir wollen uns eine Übersicht über die Menge aller Polynome aus $K[t]$ verschaffen, die durch Φ_f auf $0 \in \mathrm{End}_K(V)$ abgebildet werden, also über

$$\mathcal{I}_f := \{\sum_{i=0}^{n} \alpha_i t^i \in K[t] \mid \sum_{i=0}^{n} \alpha_i f^i = 0 \in \mathrm{End}_K(V)\}.$$

Zunächst ist klar, daß $\mathcal{I}_f$ ein von 0 verschiedenes Polynom vom Grad höchstens n^2 enthält, denn $f^0, f^1, \ldots, f^{n^2}$ sind linear abhängig wegen $\dim(\mathrm{End}_K(V)) = n^2$. Der zu beweisende Satz von CAYLEY/HAMILTON sagt aus, daß sogar schon ein Polynom vom Grad n in $\mathcal{I}_f$ liegen muß, nämlich das charakteristische Polynom p_f.

SATZ. (CAYLEY/HAMILTON). *Sei K ein Körper, V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$ ein Endomorphismus und p_f das charakteristische Polynom von f. Dann ist $\Phi_f(p_f) = 0$. Das heißt: Aus $p_f = \sum_{i=0}^{n} \alpha_i t^i$ folgt $\sum_{i=0}^{n} \alpha_i f^i = 0 \in \mathrm{End}_K(V)$.*

BEWEIS. Sei $A \in K^{n\times n}$ darstellende Matrix von f bzgl. einer Basis von V, und V_f definiert wie im Beispiel oben. Nach Definition der Skalarmultiplikation in V_f gilt dann für alle $v \in V$

$$t \cdot v = f(v) = Av.$$

Wir schieben jetzt eine allgemeine Bemerkung ein, die die folgende Argumentation verständlich machen soll. Sei M ein R-Modul. Ist $A \in R^{n\times m}$ und $x \in M$, so wird Ax komponentenweise erklärt. Ist noch $B \in K^{m\times k}$, so gilt $(AB)x = A(Bx)$.

Es ist $1 \cdot v = v$, wenn man 1 als $\in K[t]$, ferner $v \in V = K^n$ und $\cdot$ als das Skalarprodukt im Modul V_f auffaßt. Also ist Ev nach der Bemerkung oben erklärt, und ebenso tEv (was klammerfrei geschrieben werden kann). Gleichfalls kann man in Av die Matrix A als $\in K[t]^{n\times n}$ und die Multiplikation als Skalarmultiplikation in V_f auffassen. Also gilt

$$tv = tEv = Av,$$
$$(A - tE)v = 0.$$

Setze $B := A - tE \in K[t]^{n\times n}$. Nach 6.2.2 gibt es zu B eine komplementäre Matrix $B^\# \in K[t]^{n\times n}$ mit

$$B^\# B = \det(B)E.$$

Aus $Bv = 0$ folgt

$$0 = B^\# Bv = \det(B)v.$$

Es war $p_f = \det(B) = \sum_{i=0}^{n} \alpha_i t^i$. Für unser beliebiges $v \in V$ ist also

$$0 = \det(B)v = \Big(\sum_{i=0}^{n} \alpha_i t^i\Big)v = \sum_{i=0}^{n} \alpha_i f^i(v) = \Big(\sum_{i=0}^{n} \alpha_i f^i\Big)(v)$$

und damit $\sum_{i=0}^{n} \alpha_i f^i = 0 \in \mathrm{End}_K(V)$. QED

BEMERKUNG. Ist $x_1, \ldots, x_n$ Basis von V und $A \in K^{n\times n}$ eine Matrix mit

$$f\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

so gilt auch

$$f^i\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A^i\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Aus $\Phi_f(p_f) = \sum_{i=0}^n \alpha_i f^i = 0$ folgt also, daß

$$\Big(\sum_{i=0}^n \alpha_i A^i\Big)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \Big(\sum_{i=0}^n \alpha_i f^i\Big)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0,$$

also $\sum_{i=0}^n \alpha_i A^i = 0$. Das "Einsetzen" der Matrix A in ihr eigenes charakteristisches Polynom p_A liefert also die Nullmatrix.

10.3.3. *Das Minimalpolynom.* Das charakteristische Polynom p_A zu A ist i.a. keineswegs ein Polynom kleinsten Grades mit A als "Nullstelle". Ist zum Beispiel $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in K^{2\times 2}$, so ist $p_A = (1-t)^2$, aber $\mu_A := t-1$ besitzt ebenfalls A als Nullstelle, denn es gilt $\mu_A(A) = A - 1\cdot E = 0$.

Sei $V \neq 0$ ein endlichdimensionaler K-Vektorraum und $f\colon V \to V$ linear. $\mathcal{I}_f$ sei definiert wie in 10.3.2. Offenbar ist $\mathcal{I}_f$ abgeschlossen gegen Addition (d.h. aus $g, h \in \mathcal{I}_f$ folgt $g+h \in \mathcal{I}_f$) und gegen Multiplikation mit beliebigen Elementen des Rings $K[t]$ (d.h. aus $g \in K[t]$ und $h \in \mathcal{I}_f$ folgt $gh \in \mathcal{I}_f$); eine solche Teilmenge eines Ringes nennt man ein *Ideal*. Nach dem Satz von CAYLEY/HAMILTON ist $p_f \in \mathcal{I}_f$, wir kennen also bereits ein Polynom $\neq 0$ in $\mathcal{I}_f$.

DEFINITION. Sei K ein Körper. Unter den Polynomen $\neq 0$ in $\mathcal{I}_f$ gibt es eines kleinsten Grades, das wir (wegen der Abgeschlossenheit von $\mathcal{I}_f$ gegen Multiplikation mit beliebigen Körperelementen) als normiert annehmen können, d.h. der Leitkoeffizient (also der Koeffizient der höchsten t-Potenz) ist 1. Dieses Polynom ist offenbar eindeutig bestimmt; man nennt es das *Minimalpolynom* von f und bezeichnet es mit μ_f.

Wir wollen zeigen, daß das Minimalpolynom Teiler des charakteristischen Polynoms p_f ist. Dazu benötigen wir die Division mit Rest im Polynomring $K[t]$.

SATZ. (DIVISION MIT REST). *Sei K ein Körper. Zu $f, g \in K[t]$ mit $f \neq 0$ gibt es eindeutig bestimmte $q, r \in K[t]$ mit*

$$g = fq + r \quad \textit{und } \deg(r) < \deg(f).$$

BEWEIS. *Existenz.* Sei

$$\begin{aligned} f &= a_n t^n + \cdots + a_1 t + a_0 \quad \text{mit } a_n \neq 0 \\ g &= b_m t^m + \cdots + b_1 t + b_0 \quad \text{mit } b_m \neq 0. \end{aligned}$$

Den Beweis führten wir durch Induktion nach m.

Fall $m = 0$, $n > 0$. Setze $q = 0$ und $r = g$.

Fall $m \geq 0$, $n = 0$. Setze $q = \frac{1}{a_0} g$ und $r = 0$.

Fall $m > 0$, $n > 0$. OBdA sei $n \leq m$ (sonst setze $q = 0$ und $r = g$). Es gibt ein $g_1 \in K[t]$ mit

$$g = \frac{b_m}{a_n} t^{m-n} f + g_1 \quad \text{und } \deg(g_1) < m.$$

Nach Induktionsvoraussetzung gibt es $q_1, r_1 \in K[t]$ mit

$$\begin{aligned} g &= \frac{b_m}{a_n} t^{m-n} f + q_1 f + r \quad \text{und } \deg(r) < n \\ &= \underbrace{(\frac{b_m}{a_n} t^{m-n} + q_1)}_{=:q} f + r \end{aligned}$$

Eindeutigkeit. Gelte $g = q_1 f + r_1 = q_2 f + r_2$ mit $\deg(r_1), \deg(r_2) < \deg(f)$. Dann ist $f(q_1 - q_2) = r_2 - r_1$ und $\deg(r_2 - r_1) < \deg(f)$. Wäre $q_1 - q_2 \neq 0$, so würde gelten $\deg(f(q_1 - q_2)) = \deg(f) + \deg(q_1 - q_2) \geq \deg(f)$ (wegen 7.1.2, da K ein Körper ist), also ein Widerspruch (wegen $\deg(0) = -\infty$). Folglich muß $q_1 = q_2$ sein und somit auch $r_1 = r_2$. QED

FOLGERUNG. *Sei $V \neq 0$ ein endlichdimensionaler K-Vektorraum und $f \colon V \to V$ linear. Dann ist das Minimalpolynom von f ein Teiler des charakteristischen Polynoms von f, also*

$$\mu_f | p_f.$$

BEWEIS. Wegen $\mu_f \neq 0$ gibt es nach dem Satz $q, r \in K[t]$ mit

$$p_f = \mu_f q + r \quad \text{und } \deg(r) < \deg(\mu_f).$$

Wegen $r(f) = p_f(f) - \mu_f(f) q(f) = 0$ und der Minimalität des Grades des Minimalpolynoms muß $r = 0$ sein. QED

SATZ. *Sei $V \neq 0$ ein endlichdimensionaler K-Vektorraum und $f \colon V \to V$ linear. Dann haben das charakteristische Polynom p_f und das Minimalpolynom μ_f die gleichen Nullstellen (i.a. haben also die Nullstellen von μ_f kleinere Vielfachheit).*

BEWEIS. Wegen $\mu_f | p_f$ ist jede Nullstelle von μ_f auch Nullstelle von p_f. Sei umgekehrt λ Nullstelle von p_f, also Eigenwert von f, und x ein Eigenvektor zu λ, also $x \neq 0$ mit $f(x) = \lambda x$. Sei

$$\mu_f = \sum_{i=0}^{k} \alpha_i t^i$$

das Minimalpolynom von f. Dann gilt

$$0 = \Big(\sum_{i=0}^{k} \alpha_i f^i\Big)(x) = \sum_{i=0}^{k} \alpha_i f^i(x) = \sum_{i=0}^{k} \alpha_i \lambda^i x = \Big(\sum_{i=0}^{k} \alpha_i \lambda^i\Big) x.$$

Also ist $\mu_f(\lambda) = 0$. QED

Der folgende Satz liefert ein Verfahren, das Minimalpolynom aus der JORDANschen Normalform zu berechnen.

SATZ. *Sei V ein n-dimensionaler K-Vektorraum und $f\colon V \to V$ linear. Ferner sei p_f Produkt linearer Faktoren, etwa*

$$p_f = (-1)^n (t-\lambda_1)^{e_1} \cdot \dots \cdot (t-\lambda_m)^{e_m} \quad \textit{mit verschiedenen } \lambda_i \textit{ und } e_i \geq 1.$$

Sei $k_i = \min\{\, k \mid \operatorname{Kern}(g_i^k) = \operatorname{Kern}(g_i^{k+1}) \,\}$ mit $g_i := f - \lambda_i \mathrm{id}$. Dann ist

$$\mu_f = (t-\lambda_1)^{k_1} \cdot \dots \cdot (t-\lambda_m)^{k_m}$$

das Minimalpolynoms von f.

BEMERKUNG. Die Existenz von k_i wurde im Beweis des Satzes 7.3.1 von der JORDANschen Normalform bewiesen. Aus diesem Beweis geht auch hervor, daß k_i die Zeilenzahl des größten zum Eigenwert λ_i gehörigen JORDAN-Kästchens ist, und daß auch $k_i = \min\{\, k \mid (g_i{\restriction}U_i)^k = 0 \,\text{und}\, (g_i{\restriction}U_i)^{k-1} \neq 0 \,\}$ ist mit $U_i = \operatorname{Kern}(g_i^k)$. Man nennt k_i den *Index* der nilpotenten Abbildung $g_i{\restriction}U_i$. Ist insbesondere f diagonalisierbar, so folgt

$$\mu_f = (t-\lambda_1) \cdot \dots \cdot (t-\lambda_m).$$

BEWEIS. Nach der obigen Folgerung ist μ_f ein Teiler von p_f, und nach dem darauf folgenden Satz hat μ_f dieselben Nullstellen wie p_f, also

$$\mu_f = (t-\lambda_1)^{n_1} \cdot \dots \cdot (t-\lambda_m)^{n_m}$$

mit $n_i \leq e_i$ für $i \in \{1, \dots, m\}$. Zu zeigen ist $n_i = k_i$.

Nach dem Beweis von 7.3.1 ist $V = U_1 \oplus \dots \oplus U_m$ mit $U_i = \operatorname{Kern}(g_i^{k_i})$ eine Zerlegung von V in f-invariante Unterräume.

Wir zeigen $n_i \leq k_i$. Da die n_i Exponenten des Minimalpolynoms sind, genügt

$$(f-\lambda_1 \mathrm{id})^{k_1} \circ \dots \circ (f-\lambda_m \mathrm{id})^{k_m} = 0, \quad \text{also } g_1^{k_1} \circ \dots \circ g_m^{k_m} = 0.$$

Wegen $V = U_1 \oplus \dots \oplus U_m$ genügt es dafür zu zeigen, daß jedes $x_i \in U_i$ auf 0 abgebildet wird. Sei etwa $x_1 \in U_1$. Dann gilt

$$(g_1^{k_1} \circ g_2^{k_2} \circ \dots \circ g_m^{k_m})(x_1) = (g_2^{k_2} \circ \dots \circ g_m^{k_m})(g_1^{k_1}(x_1)) = 0,$$

da g_i und g_j kommutieren (denn $(f-\lambda_i \mathrm{id}) \circ (f-\lambda_j \mathrm{id}) = f^2 - (\lambda_i + \lambda_j) f + \lambda_i \lambda_j \mathrm{id} = (f-\lambda_j \mathrm{id}) \circ (f-\lambda_i \mathrm{id})$) und $x_1 \in U_1 = \operatorname{Kern}(g_1^{k_1})$.

Wir zeigen jetzt, daß die Annahme $n_i < k_i$ auf einen Widerspruch führt. Nehmen wir etwa $n_1 < k_1$ an. Dann gibt es ein $x_1 \in U_1$ mit $y := g_1^{n_1}(x_1) \neq 0$. Da U_1 f-invariant ist (und $g_j = f - \lambda_j \mathrm{id}$), ist U_1 auch g_j-invariant, also auch $g_j^{n_j}$-invariant. Für $j \neq 1$ ist

$$g_j^{n_j}{\restriction}U_1 \colon U_1 \to U_1$$

injektiv, denn aus $g_j^{n_j}(x) = 0$ für ein $x \in U_1$ folgt $x \in \operatorname{Kern}(g_j^{n_j}) \subseteq U_j$ (wegen $n_j \leq k_j$), also $x \in U_1 \cap U_j = 0$. Also ist $g_j^{n_j}{\restriction}U_1 \colon U_1 \to U_1$ ein Isomorphismus und damit auch $g_2^{n_2}{\restriction}U_1 \circ \dots \circ g_m^{n_m}{\restriction}U_1$. Wegen des schon verwendeten Kommutierens von g_i und g_j folgt

$$(g_1^{n_1} \circ g_2^{n_2} \circ \dots \circ g_m^{n_m})(x_1) = (g_2^{n_2} \circ \dots \circ g_m^{n_m})(\underbrace{g_1^{n_1}(x_1)}_{=y\neq 0}) \neq 0$$

im Widerspruch zu $\mu_f(f) = 0$. QED

10.3.4. *Faktormoduln.* Sei M ein R-Modul und $N \subseteq M$ eine Teilmenge. N heißt *R-Untermodul* von M, wenn N gegen Addition und gegen Skalarmultiplikation mit beliebigen Ringelementen abgeschlossen ist.

SATZ. *(Konstruktion des Faktormoduls). Sei M ein R-(Links)modul, $N \subseteq M$ ein Untermodul und*

$$\begin{aligned} \mathrm{kan}\colon M &\to M/N \\ x &\mapsto [x]. \end{aligned}$$

Dann gibt es genau eine R-Modulstruktur auf M/N, so daß kan *ein R-Homomorphismus wird.*

BEWEIS. M/N ist abelsche Gruppe bei $[x]+[y]=[x+y]$. Definiere

$$\begin{aligned} R \times M/N &\to M/N \\ (r,[x]) &\mapsto [rx]. \end{aligned}$$

Zu zeigen ist nur noch, daß diese Abbildung wohldefiniert ist (denn es ist klar, daß M/N damit zu einem R-Modul wird). Seien also $x, y \in M$, $r \in R$ und gelte $[x]=[y]$. Dann ist $x-y =: n \in N$, also $rx - ry = r(x-y) = rn \in N$ und damit $[rx]=[ry]$. QED

SATZ. *(Universelle Eigenschaft des Faktormoduls). Seien M, M' R-Moduln, $N \subseteq M$ Untermodul und $f\colon M \to M'$ ein R-Modulhomomorphismus mit $f(N)=\{0\}$. Dann gibt es genau einen R-Modulhomomorphismus $g\colon M/N \to M'$ mit $g \circ \mathrm{kan} = f$.*

$$\begin{tikzcd} M \arrow[rr, "f"] \arrow[dr, "\mathrm{kan}"'] & & M' \\ & M/N \arrow[ur, dashed, "g"'] & \end{tikzcd}$$

BEWEIS. Wie für Gruppen. QED

FOLGERUNG. *(Homomorphiesatz). Seien M, M' R-Moduln und $f\colon M \to M'$ ein R-Modulhomomorphismus. Dann gilt*

$$\begin{aligned} \overline{f}\colon M/\mathrm{Kern}(f) &\to \mathrm{Bild}(f) \\ [x] &\mapsto f(x) \end{aligned}$$

ist ein wohldefinierter R-Modulisomorphismus (d.h. ein bijektiver R-Modulhomomorphismus; es ist $\mathrm{Kern}(f) := f^{-1}(0)$*).*

BEWEIS. Wie für Gruppen. QED

Definition. Sei I eine Menge und R ein *Ring*.

$$\begin{aligned} R^I &:= \{\, f\colon I \to R \mid f \text{ Abbildung} \,\} \\ &= \{\, (r_i)_{i\in I} \mid r_i \in R \text{ für } i \in I \,\} \end{aligned}$$

wird zu einem R-Modul durch

$$\begin{aligned} (r_i) + (s_i) &:= (r_i + s_i) \\ r \cdot (r_i) &:= (rr_i) \end{aligned}$$

für $r_i, s_i, r \in R$. Weiter definieren wir

$$R^{(I)} := \{\, (r_i)_{i\in I} \in R^I \mid \text{für fast alle } i \in I \text{ gilt } r_i = 0 \,\}.$$

Für $j \in I$ sei

$$e_j := (r_i)_{i\in I} \quad \text{mit} \quad r_i := \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

Sei M ein R-Modul und I eine Menge. Eine Familie $(x_i)_{i\in I}$ von Elementen aus M heißt *Basis* von M wenn folgendes gilt.

- Für alle $x \in M$ gibt es eine endliche Teilmenge $J \subseteq I$ von I und $r_i \in R$ für $i \in J$ so daß $x = \sum_{i\in J} r_i x_i$.
- Für alle endlichen Teilmengen $J \subseteq I$ und $r_i, s_i \in R$ für $i \in J$

$$\text{folgt aus } \textstyle\sum_{i\in J} r_i x_i = \sum i \in J s_i x_i \text{ stets } r_i = s_i \text{ für alle } i \in J.$$

Bemerkungen. 1. Sei I eine Menge und R ein Ring. Dann ist offenbar $R^{(I)}$ ein R-Modul mit der Basis $(e_i)_{i\in I}$.

2. Man kann zeigen, daß nicht jeder Modul eine Basis besitzt (ein Beispiel ist $\mathbb{Z}/n\mathbb{Z}$). Ferner ist die Basislänge i.a. nicht eindeutig bestimmt. Diese für Vektorräume gültigen Aussagen lassen sich also nicht auf Moduln übertragen.

3. Auch für Moduln gilt der Satz von der *linearen Fortsetzung*. Seien M, M' R-Moduln, $(x_i)_{i\in I}$ eine Basis von M und $(y_i)_{i\in I}$ eine beliebige Familie von Elementen aus M'. Dann gibt es genau einen Modulhomomorphismus $f\colon M \to M'$ mit $f(x_i) = y_i$ für alle $i \in I$. Dies beweist man wie in 3.1.2 für Vektorräume.

10.4 Tensorprodukt

In diesem Abschnitt sei K ein kommutativer Ring, z.B. ein Körper.

10.4.1. Definition. Sei n eine natürliche Zahl und $M_1, \ldots, M_n, M$ R-Moduln. Eine Abbildung $f\colon M_1 \times \cdots \times M_n \to M$ heißt *n-multilinear*, wenn für jedes i und alle $x_j \in M_j$ mit $j \neq i$ die Abbildung

$$\begin{aligned} M_i &\to M \\ x_i &\mapsto f(x_1, \ldots, x_n) \end{aligned}$$

R-linear ist. Sei T ein R-Modul und $t\colon M_1 \times \cdots \times M_n \to T$ eine Abbildung. t heißt *universell n-multilinear*, wenn t n-multilinear ist und für jeden R-Modul X und jede n-multilineare Abbildung $f\colon M_1 \times \cdots \times M_n \to X$ es genau eine R-lineare Abbildung $g\colon T \to X$ gibt mit $f = g \circ t$, d.h. so daß

$$\begin{tikzcd} M_1 \times \cdots \times M_n \arrow[rr, "f"] \arrow[dr, "t"'] & & X \\ & T \arrow[ur, dashed, "g"'] & \end{tikzcd}$$

kommutativ ist.

Bemerkung. Seien

$$t\colon M_1 \times \cdots \times M_n \to T$$
$$t'\colon M_1 \times \cdots \times M_n \to T'$$

universell n-multilinear. Dann gibt es genau einen Isomorphismus $g\colon T \to T'$ mit $t' = g \circ t$.

$$\begin{tikzcd} M_1 \times \cdots \times M_n \arrow[rr, "t'"] \arrow[dr, "t"'] & & T' \\ & T \arrow[ur, dashed, "g"'] & \end{tikzcd}$$

Beweis. 1. Da t multilinear und t' universell multilinear ist, gibt es ein $h\colon T' \to T$ mit $t = h \circ t'$.

2. Da t' multilinear und t universell multilinear ist, gibt es ein $g\colon T \to T'$ mit $t' = g \circ t$.

3. Aus (1), (2) folgt $t = h \circ t' = h \circ g \circ t$, also $h \circ g = \mathrm{id}$ (da t universell ist) und ebenso $t' = g \circ t = g \circ h \circ t'$, also $g \circ h = \mathrm{id}$ (da t' universell ist). Qed.

$$\begin{tikzcd} & T \\ M_1 \times \cdots \times M_n \arrow[ur] \arrow[r] \arrow[dr] & T' \arrow[u, "h"] \\ & T \arrow[u, "g"] \arrow[uu, bend right=60, "\mathrm{id}"'] \end{tikzcd}$$

10.4.2. Satz. *(Existenz des Tensorprodukts). Seien $M_1, \ldots, M_n$ R-Moduln. Dann gibt es einen R-Modul T und eine universelle n-multilineare Abbildung $t\colon M_1 \times \cdots \times M_n \to T$.*

Beweis. Sei $F := R^{(M_1 \times \cdots \times M_n)}$ der von der Menge $M_1 \times \cdots \times M_n$ frei erzeugte R-Modul, also der R-Modul mit der aus allen $e_{(x_1,\ldots,x_n)}$ für $(x_1, \ldots, x_n) \in M_1 \times \cdots \times M_n$ bestehenden Basis. In F identifizieren wir $e_{(x_1,\ldots,x_n)}$ mit $(x_1, \ldots, x_n)$. Sei $N \subseteq F$ der Untermodul, der erzeugt wird von allen Elementen von F der Form

$$(x_1, \ldots, x_{i-1}, x_i + x_i', x_{i+1}, \ldots, x_n) - (x_1, \ldots, x_i, \ldots, x_n) - (x_1, \ldots, x_i', \ldots, x_n),$$
$$(x_1, \ldots, x_{i-1}, r x_i, x_{i+1}, \ldots, x_n) - r(x_1, \ldots, x_i, \ldots, x_n)$$

mit $x_1 \in M_1, \ldots, x_i, x_i' \in M_i, \ldots, x_n \in M_n$ und $r \in R$. Definiere

$$\begin{aligned}
&T := F/N,\\
&t\colon M_1 \times \cdots \times M_n \to T = F/N,\\
&t(x_1, \ldots, x_n) := [(x_1, \ldots, x_n)].
\end{aligned}$$

Wir zeigen zunächst, daß t multilinear ist. Zum Beispiel für $r \in R$ und $x_i \in M_i$ haben wir

$$\begin{aligned}
t(rx_1, x_2, \ldots, x_n) &= [(rx_1, x_2, \ldots, x_n)]\\
&= [r(x_1, x_2, \ldots, x_n)]\\
&= r[(x_1, x_2, \ldots, x_n)]\\
&= rt(x_1, x_2, \ldots, x_n),
\end{aligned}$$

da $(rx_1, x_2, \ldots, x_n) - r(x_1, \ldots, x_n) \in N$. Wir zeigen jetzt, daß t universell multilinear ist. Sei also $f\colon M_1 \times \cdots \times M_n \to X$ multilinear, X ein R-Modul. Wir definieren

$$\begin{aligned}
&f'\colon F \to X \quad \text{durch}\\
&f'(e_{(x_1,\ldots,x_n)}) := f(x_1, \ldots, x_n).
\end{aligned}$$

Es gibt genau einen R-Homomorphismus mit dieser Eigenschaft (lineare Fortsetzung). Es gilt $f'(N) = 0$, denn f' ist auf jedem der obigen Erzeugenden von N Null; z.B. ist

$$\begin{aligned}
f'((rx_1, x_2, \ldots, x_n) - r(x_1, \ldots, x_n)) &= f'(rx_1, x_2, \ldots, x_n) - f'(r(x_1, \ldots, x_n))\\
&= f(rx_1, x_2, \ldots, x_n) - rf(x_1, \ldots, x_n)\\
&= 0, \quad \text{da } f \text{ multilinear ist.}
\end{aligned}$$

Nach der universellen Eigenschaft des Faktormoduls gibt es einen R-Homomorphismus

$$\begin{aligned}
&g\colon F/N = T \to X\\
&\quad g([e_{(x_1,\ldots,x_n)}]) \mapsto f(x_1, \ldots, x_n).
\end{aligned}$$

Es gilt $f = g \circ t$, denn für $x_i \in M_i$ $(1 \le i \le n)$ ist

$$g(t(x_1, \ldots, x_n)) = g([e_{(x_1,\ldots,x_n)}]) = f(x_1, \ldots, x_n).$$

g ist eindeutig bestimmt durch $f = g \circ t$: Sei noch $g'\colon F/N = T \to X$ gegeben mit $f = g' \circ t$. Dann gilt

$$g(t(x_1, \ldots, x_n)) = g([e_{(x_1,\ldots,x_n)}]) = g'([e_{(x_1,\ldots,x_n)}]) = g'(t(x_1, \ldots, x_n)).$$

Die Abbildungen g und g' stimmen also auf einem Erzeugendensystem von $T = F/N$ überein (denn $(e_{(x_1,\ldots,x_n)})_{(x_1,\ldots,x_n)\in M_1\times\cdots\times M_n}$ ist Basis von F), also $g = g'$. QED

DEFINITION. Seien $M_1, \ldots, M_n$ R-Moduln.

$$M_1 \otimes_R \cdots \otimes_R M_n = M_1 \otimes \cdots \otimes M_n := T$$

mit T wie im Beweis des Satzes heißt *Tensorprodukt*. Die im Beweis ebenfalls konstruierte Abbildung

$$\mathrm{kan} := t\colon M_1 \times \cdots \times M_n \to M_1 \otimes \cdots \otimes M_n$$

heißt *kanonische Abbildung*. Für $x_i \in M_i$ $(1 \le i \le n)$ setzt man

$$x_1 \otimes \cdots \otimes x_n := t(x_1, \ldots, x_n).$$

Bemerkungen. 1. Nach dem Satz ist $t\colon M_1\times\cdots\times M_n\to M_1\otimes\cdots\otimes M_n$ universell multilinear und deshalb nach einer obigen Bemerkung bis auf Isomorphie eindeutig bestimmt.

2. Da $t\colon M_1\times\cdots\times M_n\to M_1\otimes\cdots\otimes M_n$ multilinear ist, gilt

$$\begin{aligned} x_1\otimes\cdots\otimes(x_i+x_i')\otimes\cdots\otimes x_n &= x_1\otimes\cdots\otimes x_i\otimes\cdots\otimes x_n + x_1\otimes\cdots\otimes x_i'\otimes\cdots\otimes x_n,\\ x_1\otimes\cdots\otimes rx_i\otimes\cdots\otimes x_n &= r(x_1\otimes\cdots\otimes x_i\otimes\cdots\otimes x_n)\end{aligned}$$

für $x_1\in M_1,\dots,x_i,x_i'\in M_i,\dots,x_n\in M_n$, $r\in R$ und $1\le i\le n$. Insbesondere ist

$$x_1\otimes\cdots\otimes x_n=0\quad\text{falls ein } x_i=0 \text{ ist.}$$

3. Die Elemente $x_1\otimes\cdots\otimes x_n\in M_1\otimes\cdots\otimes M_n$ bilden ein Erzeugendensystem von $M_1\otimes\cdots\otimes M_n$ als R-Modul (nach dem Beweis des Satzes) und sogar als abelsche Gruppe, denn es gilt

$$r(x_1\otimes\cdots\otimes x_n)=rx_1\otimes\cdots\otimes x_n.$$

Beispiel. $R=\mathbb{Z}$, $M_1=\mathbb{Z}/2\mathbb{Z}$, $M_2=\mathbb{Z}/3\mathbb{Z}$. Dann gilt

$$M_1\otimes M_2=\mathbb{Z}/2\mathbb{Z}\otimes\mathbb{Z}/3\mathbb{Z}=0,$$

denn für $\alpha,\beta\in\mathbb{Z}$ ist $3\cdot(\alpha+2\mathbb{Z})=3\alpha+2\mathbb{Z}=\alpha+2\mathbb{Z}$ (Fallunterscheidung α gerade/ungerade), also

$$\begin{aligned}(\alpha+2\mathbb{Z})\otimes(\beta+3\mathbb{Z}) &= 3(\alpha+2\mathbb{Z})\otimes(\beta+3\mathbb{Z})\\ &= (\alpha+2\mathbb{Z})\otimes\underbrace{3(\beta+3\mathbb{Z})}_{=3\mathbb{Z}=[0]}\\ &= 0.\end{aligned}$$

10.4.3. Satz. *Seien $M_1,\dots,M_n$ R-Moduln, $(x_{i_1})_{i_1\in I_1}$ Basis von $M_1,\dots,(x_{i_n})_{i_n\in I_n}$ Basis von M_n. Dann ist*

$$(x_{i_1}\otimes\cdots\otimes x_{i_n})_{i_1\in I_1,\dots,i_n\in I_n}$$

Basis von $M_1\otimes\dots,\otimes M_n$.

Beweis. *Erzeugendensystem.* Sei $m_j\in M_j$ für $1\le j\le n$. Dann ist m_j R-Linearkombination von Elementen aus $(x_{i_j})_{i_j\in I_j}$. Also ist $m_1\otimes\cdots\otimes m_n$ R-Linearkombination von Elementen $x_{i_1}\otimes\cdots\otimes x_{i_m}$, da kan: $M_1\times\cdots\times M_n\to M_1\otimes\cdots\otimes M_n$ multilinear war.

Lineare Unabhängigkeit. Seien $i_1\in I_1,\dots,i_n\in I_n$ fest gewählt. Betrachte

$$\begin{aligned} &f\colon M_1\times\cdots\times M_n\to R\\ &f(m_1,\dots,m_n)=r_{i_1}\cdots r_{i_n},\end{aligned}$$

wobei r_{i_j} der Koeffizient von x_{i_j} in der Basisdarstellung von m_j ist. Offensichtlich ist f multilinear. Nach der universellen Eigenschaft des Tensorprodukts gibt es einen R-Homomorphismus

$$g\colon M_1\otimes\cdots\otimes M_n\to R$$

mit $f=g\circ\text{kan}$. Also ist

$$g(x_{l_1}\otimes\cdots\otimes x_{l_n})=\begin{cases}1 & \text{falls } l_1=i_1,\dots,l_n=i_n\\ 0 & \text{sonst.}\end{cases}$$

Sei jetzt

$$\sum_{l_1,\ldots,l_n} r_{l_1\cdots l_n} x_{l_1} \otimes \cdots \otimes x_{l_n} = 0 \qquad \text{mit } r_{l_1\cdots l_n} \in R, \text{ fast alle } r_{l_1\cdots l_n} = 0.$$

Daraus folgt

$$0 = g\Big(\sum_{l_1,\ldots,l_n} r_{l_1\cdots l_n} x_{l_1} \otimes \cdots \otimes x_{l_n}\Big) = r_{i_1\cdots i_n}.$$

Dies gilt für alle $i_1, \ldots, i_n$. QED

10.4.4. DEFINITION. Sei V Vektorraum über einem Körper K und $p, q \in \mathbb{N}$.

$$V_p^q := \underbrace{V \otimes \cdots \otimes V}_{p\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{-mal}}$$

heißt *Tensorraum* der *p-fach kontravarianten* und *q-fach kovarianten Tensoren* (oder auch Tensorraum der Stufe (p, q)). Hierbei ist $V^* := \mathrm{Hom}_K(V, V)$ der zu V duale Vektorraum.

BEMERKUNG. Sei $v_1, \ldots, v_n$ Basis von V. $f_i \in V^*$ sei definiert durch $f_i(v_j) := \delta_{ij}$ für $1 \le i \le n$. Dann ist $f_1, \ldots, f_n$ Basis von V^*. (f_i) heißt die zu (v_i) *duale Basis*. Im folgenden verwenden wir die Schreibweise $f_i =: v^i \in V^*$.

BEWEIS. *Erzeugendensystem.* Sei $f \in V^*$ und $f(v_i) =: \alpha_i$. Dann gilt

$$\Big(\sum_i \alpha_i f_i\Big)(v_j) = \sum_i \alpha_i f_i(v_j) = \alpha_j,$$

also $\sum_i \alpha_i f_i = f$.

Lineare Unabhängigkeit. Gelte $\sum_i \alpha_i f_i = 0$. Also

$$0 = \Big(\sum_i \alpha_i f_i\Big)(v_j) = \sum_i \alpha_i f_i(v_j) = \alpha_j$$

für alle j mit $1 \le j \le n$. QED

FOLGERUNG. *Sei V ein K-Vektorraum, $v_1, \ldots, v_n$ Basis von V unf $p, q \in \mathbb{N}$. Dann hat*

$$V_p^q = \underbrace{V \otimes \cdots \otimes V}_{p\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{-mal}}$$

die K-Basis

$$(v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes v^{j_1} \otimes \cdots \otimes v^{j_q})_{1 \le i_1, \ldots, i_p, j_1, \ldots, j_q \le n}$$

(also $\dim(V_p^q) = n^{p+q}$). Jedes $x \in V_p^q$ hat also folgende Basisdarstellung.

$$x = \sum_{\substack{i_1,\ldots,i_p \\ j_1,\ldots,j_q}} \xi^{i_1\cdots i_p}_{j_1\cdots j_q} v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes v^{j_1} \otimes \cdots \otimes v^{j_q}$$

mit $\xi^{i_1\cdots i_p}_{j_1\cdots j_q} v_{i_1} \in K$. (Die Schreibweise der Indizes ist Konvention. Sie ist so getroffen, daß stets über oben und unten stehende Indizes summiert wird). Sei $\tilde v_1, \ldots, \tilde v_n$ eine weitere Basis von V, $(\tilde v^i)$ die duale Basis zu $(\tilde v_i)$. Seien weiter $\alpha_i^j, \beta_i^j \in K$, $1 \le i, j \le n$ mit

$$\boxed{\tilde v_i = \sum_j \alpha_i^j v_j, \qquad v_i = \sum_j \beta_i^j \tilde v_j}$$

(also $(\alpha_i^j)_{ij}(\beta_i^j)_{ij} = E$). x habe bzgl. der neuen Basis folgende Darstellung:

$$x = \sum_{\substack{s_1,\dots,s_p\\ t_1,\dots,t_q}} \tilde{\xi}^{s_1\cdots s_p}_{t_1\cdots t_q} \tilde{v}_{s_1} \otimes \cdots \otimes \tilde{v}_{s_p} \otimes \tilde{v}^{t_1} \otimes \cdots \otimes \tilde{v}^{t_q}.$$

Dann gilt die Transformationsformel:

$$\boxed{\tilde{\xi}^{s_1\cdots s_p}_{t_1\cdots t_q} = \sum_{\substack{i_1,\dots,i_p\\ j_1,\dots,j_q}} \xi^{i_1\cdots i_p}_{j_1\cdots j_q} \alpha^{j_1}_{t_1} \cdots \alpha^{j_q}_{t_q} \beta^{s_1}_{i_1} \cdots \beta^{s_p}_{i_p}.}$$

Beweis. Die erste Aussage ist klar nach dem vorherigem Satz. Für die zweite Aussage ist zu berechnen

1. (v_i) aus $(\tilde{v}_i)$.
2. (v^j) aus $(\tilde{v}^j)$.

Zu (1). Nach Voraussetzung ist

$$v_i = \sum_s \beta_i^s \tilde{v}_s.$$

Zu (2). Gelte $v^j = \sum_t \gamma_t^j \tilde{v}^t$. Man erhält

$$\begin{aligned} v^j(\tilde{v}_i) &= \Big(\sum_t \gamma_t^j \tilde{v}^t\Big)(\tilde{v}_i) \\ &= \sum_t \gamma_t^j \tilde{v}^t(\tilde{v}_i) \\ &= \gamma_i^j \end{aligned}$$

und andererseits wegen $\tilde{v}_i = \sum_s \alpha_i^s v_s$

$$\begin{aligned} v^j(\tilde{v}_i) &= v^j\Big(\sum_s \alpha_i^s v_s\Big) \\ &= \sum_s \alpha_i^s v^j(v_s) \\ &= \alpha_i^j. \end{aligned}$$

Also ist

$$v^j = \sum_t \alpha_t^j \tilde{v}^t.$$

Man erhält

$$\sum_{\substack{i_1,\dots,i_p,j_1,\dots,j_q\\ s_1,\dots,s_p,t_1,\dots,t_q}} \xi^{i_1\cdots i_p}_{j_1\cdots j_q} \beta^{s_1\cdots s_p}_{i_1\cdots i_p} \alpha^{j_1\cdots j_q}_{t_1\cdots t_q} \tilde{v}_{s_1} \otimes \cdots \otimes \tilde{v}_{s_p} \otimes \tilde{v}^{t_1} \otimes \cdots \otimes \tilde{v}^{t_q}.$$

Also

$$\tilde{\xi}^{s_1\cdots s_p}_{t_1\cdots t_q} = \sum_{\substack{i_1,\dots,i_p\\ j_1,\dots,j_q}} \xi^{i_1\cdots i_p}_{j_1\cdots j_q} \beta^{s_1\cdots s_p}_{i_1\cdots i_p} \alpha^{j_1\cdots j_q}_{t_1\cdots t_q}. \quad \text{QED}$$

Bemerkungen. 1. Die Transformationsformel für Koeffizienten von Tensoren kann man sich folgendermaßen merken.

$$\tilde{v}^i = \sum_j \alpha_i^j v_j \qquad \text{Basiswechsel,} \quad A := (\alpha_i^j)_{ij}.$$

Kovariante Indizes $j_1, \ldots, j_q$ transformieren sich mittels A, kontravariante Indizes $i_1, \ldots, i_p$ mittels $(A^{-1})^t$. Für A orthogonal ist $(A^{-1})^t = A$, d.h. kovariante und kontravariante Indizes transformieren sich gleich.

2. Häufig werden Tensoren durch das Transformationsverhalten der Koeffizienten definiert: Jeder Basis $v_1, \ldots, v_n$ von V sei eine Familie $\xi(v_1, \ldots, v_n) = (\xi^{i_1 \cdots i_p}_{j_1 \cdots j_q})_{1 \leq i_\nu, j_\mu \leq n}$ $(\in K^{n^{p+q}})$ zugeordnet:

$$(v_1, \ldots, v_n) \mapsto \xi(v_1, \ldots, v_n).$$

Bei einem Basiswechsel von (v_i) zu $(\tilde{v}_i)$ gilt zwischen

$$\begin{aligned} \xi(v_1, \ldots, v_n) &=: (\xi^{i_1 \cdots i_p}_{j_1 \cdots j_q}) \quad \text{und} \\ \xi(\tilde{v}_1, \ldots, \tilde{v}_n) &=: (\tilde{\xi}^{s_1 \cdots s_p}_{t_1 \cdots t_q}) \end{aligned}$$

der im Satz beschriebene Zusammenhang. Dann ist

$$\begin{aligned} x :&= \sum_{i_\nu, j_\mu} \xi^{i_1 \cdots i_p}_{j_1 \cdots j_q} v_{i_1} \otimes \cdots \otimes v_{i_p} \otimes v^{j_1} \otimes \cdots \otimes v^{j_q} \\ &= \sum_{s_\nu, t_\mu} \tilde{\xi}^{s_1 \cdots s_p}_{t_1 \cdots t_q} \tilde{v}_{s_1} \otimes \cdots \otimes \tilde{v}_{s_p} \otimes \tilde{v}^{t_1} \otimes \cdots \otimes \tilde{v}^{t_q} \end{aligned}$$

ein Tensor in V_p^q. Für jede Basis $v_1, \ldots, v_n$ von V ist also $\xi(v_1, \ldots, v_n)$ das System der Koeffizienten des Tensors x.

3. Für $\dim(V) < \infty$ läßt sich der Tensorraum V_p^q auch anders definieren: Für K-Vektorräume $V_1, \ldots, V_n$ sei

$$M := \operatorname{Mult}(V_1 \times \ldots V_n, K) := \{\, f \colon V_1 \times \cdots \times V_n \to K \mid f \text{ multilinear} \,\}$$

die Menge der *Multilinearformen*. M ist K-Vektorraum durch

$$\begin{aligned} (f+g)(v_1, \ldots, v_n) &:= f(v_1, \ldots, v_n) + g(v_1, \ldots, v_n) \\ (\alpha f)(v_1, \ldots, v_n) &:= \alpha f(v_1, \ldots, v_n) \end{aligned}$$

für $f, g \in M$ und $\alpha \in K$.

Außerdem gilt für $\dim(V_1) < \infty, \ldots, \dim(V_n) < \infty$:

$$\begin{aligned} t \colon V_1 \times \cdots \times V_n &\to \operatorname{Mult}(V_1^* \times \cdots \times V_n^*, K) \\ (v_1, \ldots, v_n) &\mapsto ((f_1, \ldots, f_n) \mapsto f_1(v_1) \ldots f_n(v_n)) \end{aligned}$$

ist universelle multilineare Abbildung, d.h.

$$\begin{aligned} V_1 \otimes \cdots \otimes V_n &\xrightarrow{\cong} \operatorname{Mult}(V_1^* \times \cdots \times V_n^*, K) \\ v_1 \otimes \cdots \otimes v_n &\mapsto ((f_1, \ldots, f_n) \mapsto f_1(v_1) \cdots f_n(v_n)) \end{aligned}$$

Insbesondere gilt für $\dim(V) < \infty$:

$$\begin{aligned} V_p^q &= \underbrace{V \otimes \cdots \otimes V}_{p\text{-mal}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q\text{-mal}} \\ &\cong \operatorname{Mult}(\underbrace{V^* \times \cdots \times V^*}_{p\text{-mal}} \times \underbrace{V^{**} \times \cdots \times V^{**}}_{q\text{-mal}}, K) \end{aligned}$$

Wegen

$$\begin{aligned} V &\overset{\cong}{\to} V^{**} \\ v &\mapsto (f \mapsto f(v)) \end{aligned}$$

folgt

$$V_p^q \cong \operatorname{Mult}(\underbrace{V^* \times \cdots \times V^*}_{p\text{-mal}} \times \underbrace{V \times \cdots \times V}_{q\text{-mal}}, K).$$

Beweis. t ist offensichtlich multilinear. Wir zeigenjetzt, daß t universell multilinear ist. Sei X ein K-Vektorraum und $f\colon V_1 \times \cdots \times V_n \to X$ multilinear. Sei $(v_{i_\nu})_{i_\nu \in I_\nu}$ Basis von V_ν für $1 \leq \nu \leq n$. f ist bestimmt durch die Werte auf $\{\, (v_{i_1}, \ldots, v_{i_n}) \mid i_1 \in I_1, \ldots, i_n \in I_n \,\}$. Es genügt nun zu zeigen:

$$(t(v_{i_1}, \ldots, v_{i_n}))_{i_1 \in I_1, \ldots, i_n \in I_n} \text{ ist Basis von } \operatorname{Mult}(V_1^* \times \cdots \times V_n^*, K).$$

Dies sieht man wie folgt. *Erzeugendensystem.* Sei $h \in \operatorname{Mult}(V_1^* \times \cdots \times V_n^*, K)$,

$$h(v^{i_1}, \ldots, v^{i_n}) =: \alpha_{i_1 \cdots i_n} \in K.$$

Dann gilt:

$$\begin{aligned} &\Big(\sum_{j_1, \ldots, j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n})\Big)(v^{i_1}, \ldots, v^{i_n}) \\ &= \sum_{j_1, \ldots, j_n} \alpha_{j_1 \cdots j_n} v^{i_1}(v_{j_1}) \cdots v^{i_n}(v_{j_n}) \\ &= \alpha_{i_1 \cdots i_n}, \end{aligned}$$

d.h. $h = \sum_{j_1, \ldots, j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n})$.

Lineare Unabhängigkeit. Gelte

$$\sum_{j_1, \ldots, j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n}) = 0.$$

Dann folgt

$$\begin{aligned} 0 &= \Big(\sum_{j_1, \ldots, j_n} \alpha_{j_1 \cdots j_n} t(v_{j_1}, \ldots, v_{j_n})\Big)(v^{i_1}, \ldots, v^{i_n}) \\ &= \alpha_{i_1 \cdots i_n}. \end{aligned}$$

Man erhält:

$$\begin{tikzcd} V_1 \times \cdots \times V_n \arrow[rr, "f"] \arrow[dr, "t"'] & & X \\ & \operatorname{Mult}(V_1^* \times \cdots \times V_n^*, K) \arrow[ur, dashed, "g"'] & \end{tikzcd}$$

Definiere nun g durch

$$g(t(v_{i_1}, \ldots, v_{i_n})) = f(v_{i_1}, \ldots, v_{i_n}).$$

Offensichtlich ist g wohldefiniert, K-Homomorphismus, und es gilt $f = g \circ t$. QED

BEMERKUNG. Im allgemeinen $(\dim(V) = \infty)$ ist $V^{**} \ncong V$ und $\mathrm{Mult}(V_1^* \times \cdots \times V_n^*, K) \ncong V_1 \otimes \cdots \otimes V_n$.

BEMERKUNG. In 8.1.7 haben wir die Komplexifizierung bezüglich $\mathbb{R} \subseteq \mathbb{C}$ durchgeführt. Allgemeiner kann man eine entsprechende Konstruktion für eine beliebige Körpererweiterung $K \subseteq L$ und einen K-Vektorraum V durchführen. Man bildet dann das Tensorprodukt $L \otimes_K V$ und faßt es als L-Vektorraum auf.

Index

Literaturverzeichnis

[1] Gerd Fischer. *Lineare Algebra, 11-te Auflage*. Vieweg, Braunschweig, 1997.

[2] Paul R. Halmos. *Finite-Dimensional Vector Spaces*. Springer Verlag, Berlin, Heidelberg, New York, 1974.

[3] Hannes Stoppel and Birgit Griese. *Übungsbuch zur Linearen Algebra*. Vieweg, Braunschweig, 1997.