

# Vorlesung Lineare Algebra und Analytische Geometrie II

Andreas Knauf\*

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>2</b>
<b>2</b>	<b>Der Quotientenraum nach einem Unterraum</b>	<b>4</b>
<b>3</b>	<b>Der Dualraum eines Vektorraums</b>	<b>8</b>
<b>4</b>	<b>Direkte Summen von Vektorräumen</b>	<b>17</b>
<b>5</b>	<b>Die Jordansche Normalform</b>	<b>22</b>
5.1	Äquivalenz und Ähnlichkeit . . . . .	22
5.2	Nilpotente Endomorphismen . . . . .	26
5.3	Verallgemeinerte Eigenräume . . . . .	32
5.4	Jordanzerlegung und Jordansche Normalform . . . . .	37
5.5	Ein Beispiel . . . . .	41
<b>6</b>	<b>Reellifizierung und Komplexifizierung</b>	<b>43</b>
6.1	Skalarwechsel bei Vektorräumen . . . . .	43
6.2	Die reelle Jordansche Normalform . . . . .	47
<b>7</b>	<b>Lineare Differentialgleichungen</b>	<b>50</b>
7.1	Berechnung der Lösung . . . . .	51
7.2	Ein Beispiel . . . . .	53

---

\*Department Mathematik, Universität Erlangen-Nürnberg, Bismarckstr. 1 $\frac{1}{2}$ , D-91054 Erlangen, Germany. [knauf@mi.uni-erlangen.de](mailto:knauf@mi.uni-erlangen.de), web: [www.mathematik.uni-erlangen.de/~knauf](http://www.mathematik.uni-erlangen.de/~knauf)

<b>8 Euklidische und unitäre Vektorräume: Fortsetzung</b>	<b>56</b>
8.1 Orthonormalisierung . . . . .	58
8.2 Die adjungierte Abbildung . . . . .	60
8.3 Normale Operatoren . . . . .	62
8.4 Orthogonale und unitäre Operatoren . . . . .	65
8.5 (Anti-) Selbstadjungierte Endomorphismen . . . . .	71
8.6 Anwendung: Quantenmechanik . . . . .	74
<b>9 Quadriken</b>	<b>75</b>
<b>10 Projektive Räume</b>	<b>82</b>
<b>Literatur</b>	<b>89</b>
<b>Index</b>	<b>89</b>

**Vorbemerkung:** Dieses Skript kann kein Lehrbuch ersetzen. Einige Lehrbücher zur Linearen Algebra sind im Literaturverzeichnis erwähnt.

**Danksagung:** Ich danke Frau I. Moch für die sorgfältige Erstellung dieses Manuskriptes und Frau Dr. T. Dierkes und Herrn Dr. St. Weis für ihre Korrekturvorschläge.

# 1 Einleitung

In diesem zweiten Teil der Vorlesung werden wir zunächst die Lineare Algebra in verschiedenen Richtungen vertiefen:

1. Die Objekte der Linearen Algebra sind die Vektorräume, und die betrachteten Abbildungen zwischen ihnen sind linear.

Wir werden Methoden kennen lernen, durch geeignete Kombinationen von Vektorräumen (bzw. linearen Abbildungen) *neue* Vektorräume (bzw. lineare Abbildungen) zu konstruieren. In grober Analogie können wir diese Konstruktionen mit den Grundrechenarten Addition, Subtraktion und Multiplikation, mit deren Hilfe man ja Zahlen zu neuen Zahlen kombiniert, vergleichen. Konkret sollen direkte Summe, Quotient und Tensorprodukt von Vektorräumen sowie der Dualraum behandelt werden.

2. Wollen wir den geometrischen Gehalt eines Endomorphismus verstehen, so empfiehlt es sich, diesen durch Wahl einer geeigneten Basis in eine möglichst einfache Form zu überführen. Beispielsweise werden wir bei einer Drehung im dreidimensionalen Anschauungsraum einen Basisvektor

in Achsrichtung und die beiden anderen senkrecht dazu und zueinander wählen.

Eine besonders einfache Form besitzt eine Matrix, wenn sie diagonal ist. Schon das Beispiel der Drehungen zeigt zwar, dass man nicht jeden Endomorphismus diagonalisieren kann. Allerdings nimmt in der so genannten *Jordanschen Normalform* der Nichtdiagonalanteil eine einfache Form an.

3. Der Normalformalgorithmus ist besonders nützlich bei der Berechnung von nichtlinearen Funktionen  $f(\varphi)$  von Endomorphismen  $\varphi$ . Wir nehmen dabei an, dass die komplexe Funktion  $f$  eine konvergente Potenzreihendarstellung  $f(x) = \sum_{i=0}^{\infty} a_i x^i$  besitzt, z.B.  $f(x) = \exp(x) = \sum_{i=0}^{\infty} \frac{x^i}{i!}$ , und definieren  $f(\varphi)$  als  $\sum_{i=0}^{\infty} a_i \varphi^i$ . Dieser Ausdruck lässt sich dann leicht berechnen, wenn wir die Jordansche Normalform von  $\varphi$  einsetzen. Eine Anwendung von enormer praktischer Bedeutung stellen dabei die so genannten *Linearen Differentialgleichungen* der Form

$$\frac{d}{dt}x(t) = Ax(t) \quad (t \in \mathbb{R})$$

mit Zeitparameter  $t$ ,  $x(t) \in \mathbb{R}^n$  und  $A \in \mathcal{M}(n, \mathbb{R})$  dar. Diese besitzen die Lösungen

$$x(t) = \exp(At)x(0) \quad (t \in \mathbb{R}),$$

denn  $\frac{d}{dt} \exp(At) = A \exp(At)$ .

4. Sesquilinearformen und speziell das *Skalarprodukt* haben Sie schon im letzten Semester kennen gelernt. Sie bilden eine zusätzliche Struktur auf dem Vektorraum, die uns im Fall des Skalarproduktes ermöglicht, Längen und Winkel zu messen.

Wir werden die Frage untersuchen, welche Endomorphismen diese Struktur *invariant* lassen. Im Fall des Skalarproduktes auf dem  $\mathbb{R}^3$  sind dies gerade die Drehspiegelungen. Diese bilden eine Gruppe (im Beispielfall die Gruppe  $O(3)$ ), deren Struktur wir besser verstehen wollen.

5. Die Oberfläche einer Kugel um den Nullpunkt des  $\mathbb{R}^3$  wird durch alle Drehungen invariant gelassen, und beim Übergang zu anderen Bilinearformen gilt Entsprechendes für Kegel, Paraboloid und Hyperboloid. Wir werden als *Nullstellen quadratischer Gleichungen* gegebene Teilmengen des  $\mathbb{R}^n$  betrachten, speziell im  $\mathbb{R}^2$  die Kegelschnitte und im  $\mathbb{R}^3$  die genannten Flächen.
6. Zwei verschiedene Geraden in der Ebene besitzen genau einen Schnittpunkt, außer, wenn sie parallel sind. Dann "schneiden sie sich im Unendlichen", was normalerweise eine vornehme Umschreibung dafür ist, dass sie

sich nie schneiden. Was aber, wenn wir diesen Ausdruck wörtlich verstehen? Wieviele Punkte im unendlich Fernen müssen wir zum  $\mathbb{R}^2$  hinzufügen, und welches Gebilde entsteht dann? Diese Frage führt uns zur *projektiven Geometrie*, die unserer (an die euklidische Geometrie des Anschauungsraumes angepassten) Vorstellung zunächst fremd, aber ebenso schön wie nützlich ist.

7. Die Ergebnisse der Linearen Algebra werden universell in Mathematik, Naturwissenschaften und Technik genutzt. Daher werden in dieser Vorlesung exemplarisch *Anwendungen* besprochen, von der die Quantenmechanik bis zur Codierungstheorie.

## 2 Der Quotientenraum nach einem Unterraum

Im Zusammenhang der Lösungsmenge eines linearen Gleichungssystems haben wir den Begriff des affinen Unterraumes  $L \subset V$  eines Vektorraumes  $V$  kennen gelernt. War  $L \neq 0$ , dann existierte ein (eindeutiger) Unterraum  $U \subset V$ , sodass  $L = l + U$  für  $l \in L$ .

Gewissermaßen ist also  $L$  der um  $l$  verschobene Unterraum  $U$ .

Wir wählen nun einen Unterraum  $U \subset V$ . Dieser ermöglicht es uns, auf  $V$  die folgende Relation zu definieren:

$$R_U := \{(v_1, v_2) \in V \times V \mid v_1 - v_2 \in U\}.$$

Es handelt sich um eine Äquivalenzrelation, denn  $U$  ist ja insbesondere eine additive Gruppe. Die Äquivalenzklassen dieser Relation sind die affinen Unterräume der Form  $l + U$  für  $l \in V$ .

Die Menge der Äquivalenzklassen schreiben wir wie gewohnt in der Form  $V/R_U$ , oder kurz  $V/U$ .

Die Abbildung

$$\pi : V \rightarrow V/U, \quad v \mapsto v + U$$

ordnet jedem Vektor  $v \in V$  den durch ihn gehenden affinen Unterraum  $v + U$  zu.

**2.1 Beispiel**  $V := \mathbb{R}^3$ ,  $U := \text{span}(b)$  für einen Vektor  $b \in \mathbb{R}^3 \setminus \{0\}$ . Wir können den Vektor  $b_1 := b$  zu einer Basis  $\{b_1, b_2, b_3\}$  des  $\mathbb{R}^3$  ergänzen. Dann schneiden sich  $U$  und der zweidimensionale Unterraum  $W := \text{span}(b_2, b_3)$  nur in der Null. Ist nun  $L = v + U$  affiner Unterraum, also die zu  $U$  parallele durch  $v \in \mathbb{R}^3$  gehende Gerade, dann schneidet  $L$  den Unterraum  $W$  in genau einem Punkt, nämlich  $v' := \lambda_2 b_2 + \lambda_3 b_3$ , falls  $v = \lambda_1 b_1 + \lambda_2 b_2 + \lambda_3 b_3$ , siehe Abb. 2.1.

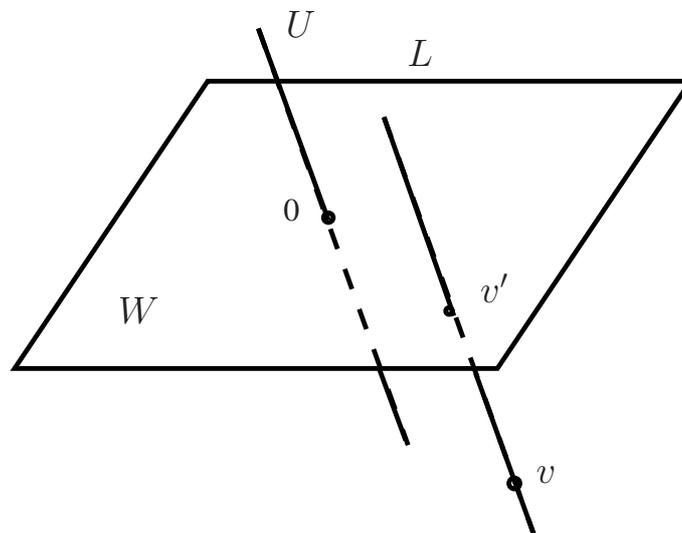


Abbildung 2.1:

Damit ist in  $W$  genau ein Repräsentant jeder Äquivalenzklasse von  $V/U$  enthalten, und wir haben eine Isomorphie  $V/U \rightarrow W, v + U \mapsto v'$  konstruiert.

Nun ist  $W \subset \mathbb{R}^3$  ein Unterraum, und wir können damit auch  $U/V$  zu einem Vektorraum machen.

Auf  $V/U$  definieren wir nun allgemein, wie bei Faktorgruppen üblich, eine Addition, und zusätzlich eine Skalarmultiplikation mit Elementen aus  $K$ :

$$\begin{aligned} [v] + [w] &:= [v + w] & (v, w \in V), \\ k[v] &:= [k \cdot v] & (v \in V, k \in K). \end{aligned}$$

**2.2 Satz** Mit diesen Verknüpfungen wird  $V/U$  zu einem  $K$ -Vektorraum, dem so genannten **Quotientenraum** oder **Faktorraum** von  $V$  nach  $U$ .

Die Abbildung  $\pi : V \rightarrow V/U$  ist linear,

$$U = \text{Kern}(\pi) \quad \text{und} \quad \text{Im}(\pi) \cong \pi(V) = V/U.$$

**Bew.:**

- Da  $V$  eine abelsche Gruppe ist, ist der Unterraum  $U$  automatisch ein Normalteiler, sodass nach Satz 3.27 der Linearen Algebra I  $V/U$  ebenfalls eine abelsche Gruppe ist. Da alle affinen Unterräume  $v + U$  nicht leer sind, gilt  $\text{Im}(\pi) = V/U$ .

- Wegen  $k \cdot [v_1] = k \cdot v_1 + U = k \cdot v_2 + U = k \cdot [v_2]$ , falls  $[v_1] = [v_2]$ , also  $v_1 - v_2 \in U$ , ist auch die Multiplikation der Äquivalenzklassen mit Zahlen repräsentantenunabhängig definiert.
- Die übrigen Vektorraumaxiome übertragen sich direkt von  $V$  auf den Faktorraum  $V/U$ .
- Ist  $u \in U$ , dann ist  $u + U = 0 + U$ , also  $[u] \in V/U$  der Nullvektor des Quotientenraumes, sodass  $U \subset \text{Kern}(\pi)$ . Ist andererseits  $u \in \text{Kern}(\pi)$ , dann ist  $[u] = U$ , also  $u \in U$ .  $\square$

**2.3 Korollar** Ist  $\text{Dim}(V) < \infty$  und  $U \subset V$  ein Unterraum, dann ist

$$\text{Dim}(V/U) = \text{Dim}(V) - \text{Dim}(U).$$

**Bew.:** Nach dem letzten Satz ist für die Quotientenabbildung  $\pi : V \rightarrow V/U$

$$\text{Rang}(\pi) \equiv \text{Dim}(\text{Im}(\pi)) = \text{Dim}(V/U)$$

und der Defekt

$$\text{Def}(\pi) = \text{Dim}(\text{Kern}(\pi)) = \text{Dim}(U).$$

Nach Satz 8.13 der Linearen Algebra I ist aber  $\text{Rang}(\pi) = \text{Dim}(V) - \text{Def}(\pi)$ .  $\square$

**2.4 Bemerkung** Das Beispiel 2.1 könnte dazu verleiten, den Faktorraum  $V/U$  für einen Unterraum  $W \subset V$  zu halten, denn tatsächlich war dort  $V/U$  zu  $W = \text{span}(b_2, b_3)$  isomorph. Diese Isomorphie hängt aber von der Wahl der Basisvektoren  $b_2$  und  $b_3$  ab, anders gesagt gibt es (unendlich) viele zweidimensionale Unterräume  $W \subset \mathbb{R}^3$  zu denen  $V/U$  in der beschriebenen Weise isomorph ist.

**Anwendung: Codes.** In Definition 5.14 der Linearen Algebra I haben wir  $(n, k)$ -Codes, d.h.  $k$ -dimensionale Unterräume  $C \subset B^n$  des arithmetischen Vektorraumes über dem zweielementigen Körper  $B$  (dem Bit) kennen gelernt. Der Quotientenraum  $B^n/C$  ist damit  $(n - k)$ -dimensional.

**2.5 Definition** Ist  $c_1, \dots, c_k \in C$  eine Basis, dann heißt die Matrix  $G := \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} \in \mathcal{M}(k \times n, B)$  eine Generatormatrix von  $C$ .

Mithilfe von  $G$  können wir jedes Codewort  $c \in C$  eindeutig als Linearkombination

$$(l_1, \dots, l_k) \cdot G = \sum_{i=1}^k l_i c_i$$

der Basisvektoren  $c_i$  darstellen. Dabei ist  $(l_1, \dots, l_k) \in B^k$ . In diesem Sinn *generiert* also  $G$  den Code  $C$ , und der Codierung der Nachricht  $(l_1, \dots, l_k)$  entspricht die Multiplikation mit  $G$ . Für die lineare Abbildung

$$\hat{G} : B^k \rightarrow B^n \quad , \quad \hat{G}(l) := l \cdot G$$

gilt dann:

$$C = \text{Im}(\hat{G}). \tag{2.1}$$

Andererseits können wir nach dem Basisergänzungssatz die Basis  $c_1, \dots, c_k$  von  $C$  zu einer Basis  $c_1, \dots, c_n$  von  $B^n$  ergänzen. Es sei

$$W := \text{span}(c_{k+1}, \dots, c_n).$$

Dieser  $(n - k)$ -dimensionale Unterraum von  $B^n$  ist durch

$$I : B^n/C \rightarrow W \quad , \quad I \left( \sum_{i=1}^n l_i c_i + C \right) := \sum_{i=k+1}^n l_i c_i$$

zu  $W$  isomorph.

Wir betrachten die lineare Abbildung

$$\hat{\Delta} : B^n \rightarrow B^{n-k} \quad , \quad \sum_{i=1}^n l_i c_i \rightarrow (l_{k+1}, \dots, l_n).$$

Komplementär zu (2.1) ist diesmal

$$C = \text{Kern}(\hat{\Delta}).$$

**2.6 Definition** Es sei  $\Delta \in \mathcal{M}((n - k) \times n, B)$  die Matrix, für die  $\hat{\Delta}(b) = b \cdot \Delta^t$ . Dann heißt  $\Delta$  eine Kontrollmatrix des Codes  $C$ .

**2.7 Beispiel** Der  $(n, n - 1)$ -Paritätscode  $C = \{(b_1, \dots, b_n) \in B^n \mid \sum_{i=1}^n b_i = 0\}$  besitzt die Generatormatrix

$$G = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 1 \\ 0 & 1 & 0 & \dots & 0 & 1 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 1 \end{pmatrix} \in \mathcal{M}((n - 1) \times n, B)$$

und die Kontrollmatrix  $\Delta = (1 \dots 1) \in \mathcal{M}(1 \times n, B)$ . Wird das Wort  $b = (b_1 \dots b_n) \in B^n$  empfangen, dann erlaubt die Berechnung von

$$b \cdot \Delta^t = (b_1 \dots b_n) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^n b_i \in B$$

zu erkennen, ob ein (einzelner) Fehler bei der Übertragung aufgetreten ist, denn dann ist  $b \cdot \Delta^t = 1$ .

### 3 Der Dualraum eines Vektorraums

**3.1 Definition** Der Dualraum eines  $K$ -Vektorraums  $V$  ist der  $K$ -Vektorraum

$$V^* := L(V, K).$$

Die Elemente  $\varphi \in V^*$  heißen Linearformen auf  $V$ .

Eine Linearform (auch *Lineares Funktional* genannt)  $\varphi \in V^*$  ist also eine lineare Abbildung

$$\varphi : V \rightarrow K$$

in den Skalkörper  $K$ .

**3.2 Beispiel** 1.  $V := K^n$ ,  $w = (w_1, \dots, w_n) \in K^n$ . Dann ist die Abbildung

$$V \rightarrow K \quad , \quad v = (v_1, \dots, v_n) \mapsto \sum_{i=1}^n v_i w_i$$

linear. Es gibt (mindestens) so viele Linearformen auf  $V$ , wie es Vektoren  $w \in V$  gibt, denn falls  $w' \neq w$ , ist das  $w'$  zugeordnete Funktional verschieden von dem von  $w$  erzeugten.

2.  $V := C([0, 1], \mathbb{R})$ , der Vektorraum der stetigen reellen Funktionen auf dem Intervall  $[0, 1]$ .

Ist  $w \in V$ , dann ist die Abbildung

$$V \rightarrow K \quad , \quad v \mapsto I_w(v) := \int_0^1 v(x)w(x)dx$$

linear, also ein Element von  $V^*$ . Wieder ist die Abbildung

$$V \rightarrow V^* \quad , \quad w \mapsto I_w$$

injektiv.

3. Eine andere Klasse von Funktionalen auf  $V = C([0, 1], \mathbb{R})$  ist durch die Auswertung an einer Stelle  $a \in [0, 1]$  gegeben:

$$\delta_a \in V^* \quad \delta_a(v) := v(a).$$

Wir wollen nun der Frage nachgehen, "wie groß" der Dualraum  $V^*$  im Vergleich zu  $V$  ist.

**3.3 Definition**  $B := (b_1, \dots, b_n)$  sei eine Basis eines  $K$ -Vektorraums  $V$ . Dann heißt  $B^* := (b_1^*, \dots, b_n^*)$  mit  $b_i \in V^*$ ,

$$b_i^*(b_j) := \delta_{ij} \quad (i, j = 1, \dots, n)$$

die zu  $B$  duale Basis von  $V^*$ .

**3.4 Satz** Die duale Basis  $B^*$  ist tatsächlich eine Basis von  $V^*$ , und somit ist  $\text{Dim}(V^*) = \text{Dim}(V)$ .

**Bew.:**

- **Lineare Unabhängigkeit:**  
Für die Koeffizienten  $\lambda_1, \dots, \lambda_n \in K$  ist die Linearform  $v^* := \sum_{i=1}^n \lambda_i b_i^* \in V^*$ , ausgewertet auf dem  $j$ -ten Basiselement von  $V$ , gleich  $v^*(b_j) = \sum_{i=1}^n \lambda_i b_i^*(b_j) = \lambda_j$ .  
Daher ist  $v^*$  nur dann der Nullvektor des Dualraums  $V^*$ , wenn  $\lambda_1 = \dots = \lambda_n = 0$ .
- $\text{span}(b_1^*, \dots, b_n^*) = V^*$ :  
Es sei  $v^* \in V^*$  eine beliebige Linearform auf  $V$  und  $\lambda_j := v^*(b_j)$ . Dann ist  $v^* = \sum_{i=1}^n \lambda_i b_i^*$ .  $\square$

Mithilfe der Basis  $B$  von  $V$  und der dualen Basis  $B^*$  von  $V$  erhalten wir also den Isomorphismus

$$V \rightarrow V^* \quad , \quad \sum_{i=1}^n \lambda_i b_i \mapsto \sum_{i=1}^n \lambda_i b_i^* .$$

Allerdings ist dieser nicht *kanonisch*, d.h. ohne Basiswahl definierbar, denn für eine andere Basis von  $V$  erhalten wir einen anderen Isomorphismus.

Außerdem ist es wichtig zu beachten, dass die Linearform  $b_i^*$  nicht nur vom Basisvektor  $b_i$ , sondern von der Wahl *aller* Basisvektoren  $b_1, \dots, b_n$  abhängt.

**3.5 Bemerkung** Ist  $V := K^n$  aber ein arithmetischer Vektorraum, dann besitzt dieser schon die kanonische Basis  $(e_1, \dots, e_n)$ . Wenn wir in diesem Fall das Element  $e_i^*$  der dualen Basis mit dem Element  $e_i$  identifizieren, erhalten wir durch lineare Fortsetzung einen 'kanonischen' Isomorphismus zwischen  $K^n$  und  $(K^n)^*$ .

Zwar können wir einen Vektorraum  $V$  nicht ohne Zuhilfenahme einer Basis in  $V^*$  abbilden, wohl aber in den Bidualraum  $V^{**}$ , den Dualraum des Dualraums:

**3.6 Satz** Die lineare Abbildung, genannt **kanonischer Homomorphismus**

$$\iota : V \rightarrow V^{**} \quad \iota(v)(\varphi) := \varphi(v) \quad (v \in V, \varphi \in V^*)$$

ist injektiv, und für  $\text{Dim}(V) < \infty$  bijektiv.

**Bew.:**

- Injektivität:

Es sei  $v \in V$ ,  $v \neq 0$ . Dann existiert nach dem Basisergänzungssatz eine Basis  $(b_i)_{i \in I}$  von  $V$  mit  $v = b_j$  für ein Basiselement. Ist nun  $\varphi \in V^*$  die Linearform mit

$$\varphi(b_i) = \delta_{ij},$$

dann ist  $\iota(v)(\varphi) = 1$ , also  $\iota(v) \in V^{**} \setminus \{0\}$ .

- Surjektivität für  $\text{Dim}(V) < \infty$ :

Es ist nach Satz 3.4

$$\text{Dim}(V^{**}) = \text{Dim}(V^*) = \text{Dim}(V) < \infty$$

und  $\iota : V \rightarrow V^{**}$  ist injektiv. Aus Satz 10.2 der LA I folgern wir, dass  $\iota$  bijektiv ist.  $\square$

Ist  $V$  aber unendlich-dimensional, dann ist im Allgemeinen  $V^{**} \supsetneq V$ . Dass  $V$  und  $V^*$  nicht isomorph sind, zeigt das folgende Beispiel.

### 3.7 Beispiel $V := K[x]$ , der Vektorraum der Polynome über $K$ .

Ist  $\lambda = (\lambda_i)_{i \in \mathbb{N}_0}$  eine Folge mit Koeffizienten  $\lambda_i \in K$ , dann ist

$$\varphi_\lambda : V \rightarrow K \quad , \quad \sum_{i=0}^n a_i x^i \mapsto \sum_{i=0}^n a_i \lambda_i$$

ein lineares Funktional, also  $\varphi_\lambda \in V^*$ .

Für  $\lambda' \neq \lambda$  ist auch  $\varphi_{\lambda'} \neq \varphi_\lambda$ , denn es gibt ein  $j \in \mathbb{N}_0$  mit  $\lambda'_j \neq \lambda_j$ , also  $\varphi_\lambda(x^j) \neq \varphi_{\lambda'}(x^j)$ .

Andererseits können wir ein beliebiges Funktional  $\varphi \in V^*$  in der Form  $\varphi = \varphi_\lambda$  mit  $\lambda_i := \varphi(x^i)$  schreiben.

Damit ist  $V^*$  isomorph zum  $K$ -Vektorraum  $K^{\mathbb{N}_0}$  der Folgen  $\lambda$ , während der Polynomraum selbst isomorph zum Unterraum

$$\{\lambda \in K^{\mathbb{N}_0} \mid \exists n_0 \in \mathbb{N}_0 : \lambda_n = 0 \quad \text{für} \quad n > n_0\}$$

der abbrechenden Folgen ist.

Dieser Unterraum ist viel kleiner als  $K^{\mathbb{N}_0}$ .

Sei z.B.  $K := (\{0, 1\}, +)$  der zweielementige Körper. Dann gibt es genau  $2^n$  Polynome  $n \in \mathbb{N}_0$ -ten Grades. Zusammen mit dem Nullpolynom ergibt das immer noch abzählbar viele Polynome. Andererseits ist  $K^{\mathbb{N}_0}$  nicht abzählbar, denn die Abbildung  $K^{\mathbb{N}_0} \rightarrow [0, 2]$ ,  $\lambda \mapsto \sum_{n=0}^{\infty} \lambda_n 2^{-n}$  ist surjektiv<sup>1</sup>,  $[0, 2] \subset \mathbb{R}$  aber nicht abzählbar.

<sup>1</sup>Hier wurde in schlampiger Schreibweise die Null bzw. Eins aus  $K$  mit der aus  $\mathbb{R}$  identifiziert.

Funktionale sind so wichtig, dass eine mathematische Disziplin, die Funktionalanalysis, nach ihnen benannt ist. Der Ansatz ist dabei der, einen Vektor  $v \in V$  möglichst genau dadurch zu bestimmen, indem man Funktionale  $\varphi_i \in V^*$  auf  $v$  auswertet, d.h.  $\varphi_i(v) \in K$  berechnet. Dabei ist für  $k_i \in K$

$$\varphi_i^{-1}(k_i) = \{v \in V \mid \varphi(v) = k\}$$

ein affiner Unterraum, und Schnitte affiner Unterräume sind affine Unterräume im Allgemeinen kleinerer Dimension.

**3.8 Beispiel** Beispielsweise lässt sich das von einem Mikrophon in einem Zeitintervall  $[0, T]$  aufgenommene Signal als Element  $v$  des Vektorraums  $V = C([0, T], \mathbb{R})$  auffassen. Soll nun ein Telefongespräch digital übertragen werden, dann wird im Abstand  $\Delta t$  von ca.  $1/10\,000$  sec. das Signal gemessen, d.h. es werden die Funktionale  $\varphi_k \in V^*$ ,  $\varphi_k(v) := v(k \cdot \Delta t) \in \mathbb{R}$  mit  $k = 1, \dots, [T/\Delta t]$  ausgewertet, und die Folge dieser (gerundeten) reellen Zahlen wird übertragen.

Natürlich lässt sich  $v$  aus diesen Zahlenwerten nicht mehr exakt, aber doch näherungsweise rekonstruieren, wenn man annimmt, dass der Anteil hoher Frequenzen am Tonsignal gering ist. Nach dem so genannten Abtast-Theorem von Shannon kann man auf diese Weise Töne mit einer Frequenz bis zu  $1/(2\Delta t) \approx 5000\text{Hz}$  übertragen.

Die lineare Algebra behandelt neben den Vektorräumen die linearen Abbildungen zwischen ihnen. Wie verhalten sich diese beim Übergang zu den Dualräumen?

**3.9 Definition** Es sei  $\varphi : V \rightarrow W$  eine lineare Abbildung.

Dann heißt die lineare Abbildung

$$\varphi^t : W^* \rightarrow V^* \quad \varphi^t(w^*) := w^* \circ \varphi \quad (w^* \in W^*)$$

die zu  $\varphi$  transponierte Abbildung.

**3.10 Satz** Sind  $V$  und  $W$  endlich-dimensionale  $K$ -Vektorräume und  $\varphi \in L(V, W)$ , dann gilt für die darstellende Matrix  $A := \mathcal{M}_C^B(\varphi)$  von  $\varphi$  bezüglich der Basen  $B$  von  $V$  und  $C$  von  $W$ :

$$A^t = \mathcal{M}_{B^*}^{C^*}(\varphi^t),$$

d.h.  $A^t$  ist die darstellende Matrix der transponierten Abbildung bezüglich der dualen Basen.

**Bew.:** Ist  $B = (b_1, \dots, b_m)$  und  $C = (c_1, \dots, c_n)$ , dann ist

$$\varphi = \sum_{i=1}^n \sum_{k=1}^m (A)_{ik} c_i b_k^*$$

oder auch  $(A)_{ik} = c_i^*(\varphi(b_k))$ .

Andererseits gilt für  $D := \mathcal{M}_{B^*}^{C^*}(\varphi^t)$

$$\varphi^t = \sum_{k=1}^m \sum_{i=1}^n (D)_{ki} b_k^* c_i$$

oder  $(D)_{ki} = b_k(\varphi^t(c_i^*)) = c_i^*(\varphi(b_k))$ . □

Offensichtlich gilt allgemein für  $\varphi, \psi \in L(V, W)$  und  $\lambda \in K$

$$(\varphi + \psi)^t = \varphi^t + \psi^t \quad \text{und} \quad (\lambda\varphi)^t = \lambda\varphi^t.$$

Aber da bei der Transposition die Richtung umgekehrt wird, gilt:

**3.11 Satz** • Für  $\varphi \in L(U, V)$  und  $\psi \in L(V, W)$  ist  $(\psi \circ \varphi)^t = \varphi^t \circ \psi^t$

•  $\text{Id}_W^t = \text{Id}_{W^*}$

**Bew.:** Für alle  $w^* \in W^*$  ist

•  $(\psi \circ \varphi)^t(w^*) = w^* \circ (\psi \circ \varphi) = (w^* \circ \psi) \circ \varphi = (\psi^t(w^*)) \circ \varphi = \varphi^t(\psi^t(w^*)) = (\varphi^t \circ \psi^t)(w^*)$ .

•  $\text{Id}_W^t(w^*) = w^*(\text{Id}_W) = w^* = \text{Id}_{W^*}(w^*)$ . □

Ist  $v \in V$ , dann können wir diesen Vektor als lineares Funktional  $v : V^* \rightarrow K$  auffassen;  $\text{Kern}(v) \subset V^*$  ist dann ein Unterraum, den wir als orthogonal zum Unterraum  $\text{span}(v) \subset V$  bezeichnen können. Allgemein definieren wir:

**3.12 Definition** Ist  $W \subset V$  ein Unterraum, so heißt

$$W^\perp := \{u \in V^* \mid u(w) = 0 \quad \text{für alle} \quad w \in W\}$$

das zu  $W$  orthogonale Komplement.

Ist  $\text{Dim}(V) < \infty$ , dann identifizieren wir nach Satz 3.6 den Bidualraum  $V^{**} = {}_i(V)$  mit  $V$ :

**3.13 Satz**  $W^\perp$  ist ein Unterraum, und für  $\text{Dim}(V) < \infty$  ist  $W^{\perp\perp} = W$ .

**Bew.:**

• Es ist  $0 \in W^\perp$ , und mit  $u_1, u_2 \in W^\perp$ ,  $\lambda \in K$  sind auch  $u_1 + u_2 \in W^\perp$  und  $\lambda u_1 \in W^\perp$ .

- Mit  $w \in W$  ist auch  $w \in W^{\perp\perp}$ , denn wegen Satz 3.6 ist  $V^{**} \cong V$ , sodass

$$\begin{aligned} W^{\perp\perp} &\equiv \{v^{**} \in V^{**} \mid v^{**}(u^*) = 0 \text{ für alle } u^* \in W^\perp\} \\ &= \{\iota(v) \mid v \in V, \iota(v)(u^*) \equiv u^*(v) = 0 \text{ für alle } u^* \in W^\perp\} \\ &\cong \{v \in V \mid u^*(v) = 0 \text{ für alle } u^* \in W^\perp\}. \end{aligned}$$

- Gäbe es ein  $w \in W^{\perp\perp} \setminus W$ , dann könnten wir nach dem Basisergänzungssatz diesen von einer Basis  $(b_1, \dots, b_k)$  von  $W$  linear unabhängigen Vektor zu dieser hinzufügen und zu einer Basis  $(b_1, \dots, b_n)$  von  $V$  ergänzen, sodass  $b_{k+1} = w$ . Wir definieren nun  $u^* \in V^*$  durch

$$u^* \left( \sum_i \lambda_i b_i \right) := \lambda_{k+1}.$$

Damit ist  $u^* \in W^\perp$ , denn auf der Basis  $(b_1, \dots, b_k)$  von  $W$  verschwindet  $u^*$ . Andererseits ist  $\iota(b_{k+1})(u^*) = u^*(b_{k+1}) = 1$  und, also  $b_{k+1} \cong \iota(b_{k+1}) \notin W^{\perp\perp}$ . Widerspruch!  $\square$

**3.14 Satz** Ist  $W \subset V$  ein Unterraum und  $\text{Dim}(V) < \infty$ , dann gilt

$$\text{Dim}(W^\perp) = \text{Dim}(V) - \text{Dim}(W).$$

**Bew.:**

- Es sei  $(b_1, \dots, b_k)$  eine Basis von  $W$ , die wir zu einer Basis  $B := (b_1, \dots, b_n)$  von  $V$  ergänzen. Dann ist mit der dualen Basis  $(b_1^*, \dots, b_n^*)$  von  $B$  zu zeigen:  $W^\perp = \text{span}(b_{k+1}^*, \dots, b_n^*)$ .
- Wegen  $b_l^*(b_m) = 0$  für  $m \in \{1, \dots, k\}$  und  $l \in \{k+1, \dots, n\}$  gilt jedenfalls  $W^\perp \supset \text{span}(b_{k+1}^*, \dots, b_n^*)$ .
- Für die umgekehrte Inklusion sei  $w^* = \sum_{l=1}^n \lambda_l b_l^* \in W^\perp$ , also  $b_m(w^*) = \lambda_m = 0$ ,  $m \in \{1, \dots, k\}$ . Damit ist  $w^* \in \text{span}(b_{k+1}^*, \dots, b_n^*)$ .  $\square$ .

**3.15 Satz** Für  $\varphi \in L(V, W)$  gilt

$$\text{Im}(\varphi^t) = (\text{Kern } \varphi)^\perp \quad \text{und} \quad \text{Kern}(\varphi^t) = (\text{Im } \varphi)^\perp.$$

**Bew.:**

- Nach Definition von  $\varphi^t$  ist

$$\text{Im}(\varphi^t) = \{v^* \in V^* \mid \exists w^* \in W^* \text{ mit } v^* = w^* \circ \varphi\}.$$

Andererseits ist

$$(\text{Kern } \varphi)^\perp = \{v^* \in V^* \mid \forall v \text{ mit } \varphi(v) = 0 \text{ ist } v^*(v) = 0\}.$$

Für  $v^* = w^* \circ \varphi \in \text{Im}(\varphi^t)$  gilt daher

$$v^*(v) = w^* \circ \varphi(v) = 0 \quad (v \in \text{Kern } \varphi),$$

sodass  $\text{Im}(\varphi^t) \subset (\text{Kern } \varphi)^\perp$  ist.

- Für die umgekehrte Inklusion ist zu zeigen:

Für alle  $v^* \in (\text{Kern } \varphi)^\perp$  existiert ein  $w^* \in W^*$  mit  $v^* = w^* \circ \varphi$ .

Sei nun  $(b_j)_{j \in J}$  eine Basis von  $\text{Im}(\varphi) \subset W$  und (für eine Indexmenge  $I \supset J$ )  $(b_i)_{i \in I}$  eine Basis von  $W$ . Wir definieren die Linearform  $w^* : W \rightarrow K$  durch ihre Werte auf dieser Basis:

$$w^*(b_i) := \begin{cases} v^*(\varphi^{-1}(b_i)) & , \quad i \in J \\ 0 & , \quad i \notin J. \end{cases}$$

Obwohl  $\varphi^{-1}(b_i)$  ein affiner Unterraum von  $V$  ist, ist die Linearform  $v^*$  auf ihm konstant, denn für  $v_1, v_2 \in \varphi^{-1}(b_i)$  ist  $v_1 - v_2 \in \text{Kern } \varphi$ .

Nach dieser Definition hat  $w^*$  die gewünschte Eigenschaft  $v^* = w^* \circ \varphi$ .

- Nach Definition von  $\varphi^t$  ist

$$\begin{aligned} \text{Kern}(\varphi^t) &= \{w^* \in W^* \mid w^* \circ \varphi = 0\} \\ &= \{w^* \in W^* \mid \forall v \in V : w^*(\varphi(v)) = 0\} = (\text{Im } \varphi)^\perp. \end{aligned}$$

□

Da wir jetzt eine geometrische Interpretation der von der transponierten Matrix induzierten linearen Abbildung besitzen, können wir ein offen gebliebenes Problem des letzten Semesters zu den Akten legen:

**3.16 Satz** Sei  $A \in \mathcal{M}(m \times n, K)$ . Dann sind Zeilenrang und Spaltenrang von  $A$  gleich.

**Bew.:** Es sei  $\varphi : K^n \rightarrow K^m$  die lineare Abbildung  $\varphi(x) := Ax$ , und  $\varphi^t : K^m \rightarrow K^n$  die transponierte Abbildung, also  $\varphi^t(y) = A^t y$ . Dann ist der Spaltenrang von  $A$  gleich  $\text{Rang}(\varphi) \equiv \text{Dim}(\text{Im } \varphi)$ , und der Zeilenrang von  $A^t$  gleich  $\text{Rang}(\varphi^t)$ . Nach Satz 8.13 LA 1 ist

$$\text{Rang}(\varphi) = \text{Dim}(V) - \text{Def}(\varphi).$$

Nach Satz 3.14 ist

$$\text{Def}(\varphi) \equiv \text{Dim}(\text{Kern } \varphi) = \text{Dim}(V) - \text{Dim}((\text{Kern } \varphi)^\perp),$$

und nach Satz 3.15 ist

$$\text{Dim}((\text{Kern } \varphi)^\perp) = \text{Dim}(\text{Im}(\varphi^t)) \equiv \text{Rang}(\varphi^t).$$

Zusammen ergibt dies  $\text{Rang}(\varphi^t) = \text{Rang}(\varphi)$ . □

**Anwendung: Codes.** Wir bezeichnen wieder den zweielementigen Körper mit  $B$ .

**3.17 Definition** Es sei  $C \subset B^n$  ein linearer Code.

- Dann heißt  $C^\perp \subset (B^n)^* \cong B^n$  der zu  $C$  duale Code.
- Gilt  $C^\perp = C$ , dann heißt  $C$  selbstdual.

Nach Satz 3.13 ist  $C^\perp$  wieder ein Unterraum und damit ein linearer Code. Außerdem ist immer  $C^{\perp\perp} = C$ .

**3.18 Satz** Eine Matrix  $A \in \mathcal{M}(k \times n, B)$  ist genau dann Generatormatrix des linearen Codes  $C$ , wenn sie Kontrollmatrix von  $C^\perp$  ist.

**Bew.:** Es sei  $\hat{A} : B^k \rightarrow B^n$  die lineare Abbildung  $\hat{A}(l) := l \cdot A$ . Falls  $A$  Generatormatrix von  $C$  ist, dann gilt

$$C = \text{Im}(\hat{A}) = \text{Kern}(\hat{A}^t)^\perp,$$

also ist der zu  $C$  duale Code  $C^\perp$  von der Form

$$C^\perp = \text{Kern}(\hat{A}^t).$$

Damit ist  $A$  Kontrollmatrix von  $C^\perp$ . Die umgekehrte Implikation folgt analog. □

Um nun effektiv zu decodieren, zerlegen wir den Vektorraum  $B^n$  in Nebenklassen  $w_1^{(i)} + C$  nach dem Unterraum  $C$ . Dabei wählen wir als Repräsentanten  $w^{(i)} \in B^n$  der Nebenklasse einen Vektor minimalen Gewichts, den sog. *Nebenklassenführer*. Nun besitzen alle Elemente einer Nebenklasse unter der Surjektion  $\hat{\Delta}$  das gleiche Bild. Empfangen wir also am Ende eines Kanals das Wort  $w \in B^n$ , dann können wir seine Nebenklasse  $w + C$  bestimmen, indem wir sein sog. *Syndrom*

$$w\Delta^t \in B^{n-k}$$

berechnen. Wir vergleichen dieses Syndrom mit einer Liste der Syndrome  $w^{(i)}\Delta^t$  der Repräsentanten  $w^{(i)}$  und decodieren, indem wir  $w$  den Code

$$c := w - w^{(j)} \quad \text{für} \quad w\Delta^t = w^{(j)}\Delta^t$$

zuordnen. Da  $w^{(j)}$  in seiner Nebenklasse minimales Gewicht besitzt, hat  $c$  minimalen Abstand von  $w$ . Es kann aber noch weitere Codewörter  $c' \in C$  mit gleichem Hammingabstand  $d(w, c') = d(w, c)$  geben.

### 3.19 Beispiel Simplex- und Hammingcodes

Zur Konstruktion des binären Simplex-Codes  $Sim(k) \subset B^n$  betrachten wir die  $(k \times n)$ -Matrix  $G$  mit  $n := 2^k - 1$ , deren  $n$  Spalten die (beliebig angeordneten) von Null verschiedenen Vektoren aus  $B^k$  sind. Für  $k = 3$  ergibt sich (in lexikalischer Ordnung)

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

Jede Zeile von  $G$  besitzt  $2^{k-1}$  Einsen und  $2^{k-1} - 1$  Nullen, denn es fehlt ja nur der Null(spalten)vektor aus  $B^k$ .

Das gleiche gilt für jede von Null verschiedene Linearkombination dieser Zeilen.

$G$  ist nun Generatormatrix von  $Sim(k)$ ; wir erkennen, dass die Codewörter äquidistant sind mit Abstand  $2^{k-1}$ .

Fassen wir nun  $\Delta := G$  als Kontrollmatrix auf, dann nennen wir den entsprechenden zum Simplexcode dualen Code  $Sim(k)^\perp$  den Hamming-Code  $Ham(k) \subset B^n$ . Der Minimalabstand zweier Wörter in diesem Code ist

$$Dist(Ham(k)) = 3.$$

Dies sehen wir durch Angabe eines Wortes  $c \in Ham(k) \setminus \{0\}$  minimalen Gewichts. Die den Einsen in diesem Wort entsprechenden Spalten von  $G$  sind linear abhängig, denn  $Gc^t = 0$  und ihre Zahl ist die Gesamtzahl solcher abhängiger Spalten, also gleich 3.

Der Hamming-Code gestaltet damit die Korrektur eines Fehlers, entsprechend den Hammingkugeln vom Radius 1 um die Codewörter. Diese Kugeln umfassen jeweils neben dem Codewort noch  $n = 2^k - 1$  weitere Wörter, also insgesamt  $2^k$  Wörter. Es gibt insgesamt  $|Ham(k)| = 2^{n-k}$  Codewörter. Damit ist der Hammingcode *perfekt*, d.h. die Hammingkugeln füllen  $B^n$  aus.

Um für den Hammingcode die Decodierung durchzuführen, erstellen wir zunächst die Liste der Nebenklassenführer und ihrer Syndrome. Wegen der Perfektheit von  $Ham(k)$  können und müssen wir die Kugel vom Radius 1 um die Null als Menge der Nebenklassenführer wählen, also

$$w^{(0)} := (0, \dots, 0) \in B^n \quad ,$$

$$w^{(i)} := (0, \dots, 0, \quad \underset{i\text{-te Stelle}}{1}, \quad 0, \dots, 0) \quad , \quad (i = 1, \dots, n).$$

Damit ist das Syndrom  $w^{(0)}\Delta^t = (0, \dots, 0) \in B^k$  und  $w^{(i)}\Delta^t$  ist für  $i = 1, \dots, n$  Transponierte des  $i$ -ten Spaltenvektors von  $G$ .

Ist z.B.  $k = 3$  und empfangen wir das Wort  $w = 0010101$ , so ist sein Syndrom  $w\Delta^t = (1, 0, 0)$ , entsprechend dem des Nebenklassenführers  $w^{(1)} = 1000000$ . Wir korrigieren also, indem wir  $w$  als das Codewort

$$c = w - w^{(1)} = 1010101$$

interpretieren.

Da wir die Spalten der Kontrollmatrix von  $Ham(k)$  lexikalisch geordnet haben, können wir sogar aus  $w\Delta^t$  direkt ablesen, an welcher Stelle von  $w$  wir gegebenenfalls ein Bit korrigieren müssen, hier also wegen  $w\Delta^t = (1, 0, 0)$  an der Stelle

$$1 \cdot 2^0 + 0 \cdot 2^1 + 0 \cdot 2^2 = 1.$$

## 4 Direkte Summen von Vektorräumen

In diesem Kapitel geht es darum, mehrere Vektorräume zu einem neuen Vektorraum zusammenzufügen. Dies scheint eine einfache Operation zu sein; allerdings gilt es, vier verwandte Begriffe auseinander zu halten:

### 1. Die *Summe*

$$\sum_{i \in I} W_i := \{v \in V \mid v = \sum_{i \in I} w_i \text{ mit } w_i \in W_i\}$$

von Unterräumen  $W_i \subset V$ . In der Darstellung von  $v$  sind nur endlich viele  $w_i \neq 0$ . Falls nur endlich viele Unterräume summiert werden, schreibt man kurz

$$W_1 + \dots + W_n.$$

### 2. Ist die Darstellung in 1. für alle $v \in \sum_{i \in I} W_i$ eindeutig, sprechen wir von einer *inneren direkten Summe* und schreiben

$$\bigoplus_{i \in I} W_i \text{ statt } \sum_{i \in I} W_i.$$

### 3. Für ein System $(V_i)_{i \in I}$ von $K$ -Vektorräumen können wir auf dem kartesischen Produkt

$$\prod_{i \in I} V_i := \{(v_i)_{i \in I} \mid v_i \in V_i\}$$

durch

$$(v_i)_{i \in I} + (w_i)_{i \in I} := (v_i + w_i)_{i \in I}$$

und

$$k(v_i)_{i \in I} := (kv_i)_{i \in I}$$

für

$$(v_i)_{i \in I}, (w_i)_{i \in I} \in \prod_{i \in I} V_i \text{ und } k \in K$$

eine  $K$ -Vektorraumstruktur definieren, die das *direkte Produkt* der  $V_i$  heißt.

4. Während hier unendlich viele Faktoren  $v_i \in V_i$  ungleich Null sein können, ist dies bei der *äußeren direkten Summe* der  $V_i$  anders. Diese ist der Untervektorraum

$$\bigoplus_{i \in I} V_i := \{(v_i)_{i \in I} \in \prod_{i \in I} V_i \mid |\{i \in I \mid v_i \neq 0\}| < \infty\}.$$

Ist die Indexmenge  $I$  endlich, dann stimmen (äußere) direkte Summe und direktes Produkt überein,<sup>2</sup> im Allgemeinen aber nicht.

Im endlichen Fall schreiben wir oft

$$V_1 \oplus \dots \oplus V_n \quad \text{statt} \quad \bigoplus_{i \in I} V_i.$$

Der Unterschied zwischen 1. und 2. ist uns schon bekannt. Insbesondere ist  $W_1 + W_2$  keine direkte Summe, wenn  $\text{Dim}(W_1 \cap W_2) \neq 0$ .

**4.1 Beispiel** Der Unterschied zwischen 3. und 4. wird am Beispiel des Polynomraums  $K[x]$  deutlich. Dieser ist isomorph zur äußeren direkten Summe  $\bigoplus_{i \in \mathbb{N}_0} K$  eindimensionaler Vektorräume, wobei dem Polynom  $\sum_{i=0}^n a_i x^i$  die Folge

$$(a_0, a_1, \dots, a_n, 0, 0, \dots)$$

zugeordnet wird.

Andererseits haben wir schon gesehen, dass der Dualraum  $K[x]^*$  isomorph zum direkten Produkt  $\prod_{i \in \mathbb{N}_0} K$  ist.

Ein Unterschied zwischen 2. und 4., also der inneren und der äußeren direkten Summe, ist zunächst einmal die Tatsache, dass bei 2. die Vektorräume  $W_i$  schon Unterräume eines ihnen gemeinsamen Vektorraums  $V$  sind. Allerdings gilt dies mit den *Injektionen*

$$I_j : V_j \rightarrow \bigoplus_{i \in I} V_i \quad , \quad (I_j(v))_i := \begin{cases} v & , \quad j = i \\ 0 & , \quad \text{sonst} \end{cases} \quad (4.1)$$

( $j \in I$ ) auch für die *äußere* direkte Summe. Dann ist nämlich der Unterraum  $I_j(V_j) \subset \bigoplus_{i \in I} V_i$  zu  $V_j$  isomorph, und die äußere direkte Summe  $\bigoplus_{i \in I} V_i$  ist gleich der inneren direkten Summe

$$\bigoplus_{i \in I} I_i(V_i).$$

---

<sup>2</sup>Dass dasselbe Objekt einmal Summe und einmal Produkt genannt wird, liegt an der unterschiedlichen Perspektive: Das *Produkt* zweier (endlicher) Mengen  $M_1$  und  $M_2$  hat die Kardinalität  $|M_1 \times M_2| = |M_1| \cdot |M_2|$ , daher der Name. Sind diese Mengen gleichzeitig Vektorräume, dann haben wir die Möglichkeit, Vektoren  $v_i \in M_i$  als Vektoren  $(v_1, 0)$  bzw.  $(0, v_2)$  aus  $M_1 \times M_2$  aufzufassen und zu addieren.

Daraus schließt man für  $\text{Dim}(V_i) < \infty$  insbesondere, dass die Dimension der äußeren direkten Summe

$$\text{Dim} \left( \bigoplus_{i=1}^n V_i \right) = \sum_{i=1}^n \text{Dim}(V_i) \quad (4.2)$$

ist. Insbesondere gilt für die äußere direkte Summe eines Vektorraums mit sich selbst  $\text{Dim}(V \oplus V) = 2 \text{Dim}(V)$  (während für die *Summe von Unterräumen*  $\text{Dim}(V + V) = \text{Dim}(V)$  gilt, denn  $V + V = V!$ ).

Direktes Produkt und direkte Summe linearer Abbildungen lassen sich in der folgenden Form einführen:

**4.2 Definition** *Es seien die linearen Abbildungen*

$$\varphi_i : V_i \rightarrow W_i \quad (i \in I)$$

*gegeben. Dann heißt die Abbildung*

$$\prod_{i \in I} \varphi_i : \prod_{i \in I} V_i \rightarrow \prod_{i \in I} W_i \quad , \quad \left( \prod_{i \in I} \varphi_i \right) \left( \prod_{i \in I} V_i \right) := (\varphi_i(V_i))_{i \in I}$$

*das direkte Produkt der  $\varphi_i$ . Die Einschränkung*

$$\bigoplus_{i \in I} \varphi_i : \bigoplus_{i \in I} V_i \rightarrow \bigoplus_{i \in I} W_i$$

*des direkten Produktes auf  $\bigoplus_{i \in I} V_i$  heißt die direkte Summe der  $\varphi_i$ .*

Eine Basis für die äußere direkte Summe von Vektorräumen erhält man durch *disjunkte* Vereinigung der Basen der Summanden unter Benutzung der Injektionen (4.1).

Bei einer endlichen äußeren direkten Summe  $V_1 \oplus \dots \oplus V_n$  endlich-dimensionaler Vektorräume  $V_i$  mit Basen  $B^{(i)} = (b_1^{(i)}, \dots, b_{m_i}^{(i)})$  ist dann

$$B := \left( I_1(b_1^{(1)}), \dots, I_1(b_{m_1}^{(1)}), I_2(b_1^{(2)}), \dots, I_2(b_{m_2}^{(2)}), \dots, I_n(b_1^{(n)}), \dots, I_n(b_{m_n}^{(n)}) \right)$$

eine geordnete Basis von  $V_1 \oplus \dots \oplus V_n$ .

Ist  $M_i := \mathcal{M}_{B_i}(\varphi_i)$  die darstellende Matrix des Endomorphismus  $\varphi_i \in L(V_i)$  bezüglich der Basis  $B^{(i)}$ , dann besitzt die darstellende Matrix der Summenabbildung bezüglich  $B$  die *Blockdiagonalform*

$$\mathcal{M}_B(\varphi_1 \oplus \dots \oplus \varphi_n) = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & M_n \end{pmatrix}.$$

Eng verbunden mit dem Begriff der direkten Summe ist der der Projektion.

**4.3 Definition** Ein Endomorphismus  $P \in L(V)$  heißt Projektion oder Projektor, wenn  $P \circ P = P$  gilt.

**4.4 Satz** Es sei  $\text{Dim}(V) < \infty$ . Dann ist mit  $P \in L(V)$  auch  $Q := \text{Id} - P$  eine Projektion, und

$$V = \text{Kern}(P) \oplus \text{Im}(P) \quad (4.3)$$

mit  $\text{Im}(P) = \text{Kern}(Q)$  und  $\text{Spek}(P) \subset \{0, 1\}$ .

**Bew.:**

- $Q^2 = (\text{Id} - P)^2 = \text{Id} - 2P + P^2 = \text{Id} - P = Q$ .
- Ist  $v \in \text{Im}(P)$ , dann existiert ein  $u \in V$  mit  $v = P(u)$ , also  $P(v) = P^2(u) = P(u) = v$ , also  $v = 0$  oder  $v \notin \text{Kern}(P)$ . Also ist die Summe

$$\text{Kern}(P) + \text{Im}(P)$$

direkt. Andererseits ist nach dem Dimensionssatz  $\text{Def}(P) + \text{Rang}(P) = \text{Dim}(V)$ , sodass wegen (4.2)  $\text{Dim}(\text{Kern}(P) \oplus \text{Im}(P)) = \text{Dim}(V)$ .

Wegen  $\text{Dim}(V) < \infty$  muss dann auch schon (4.3) gelten.

- Wegen  $QP = (\text{Id} - P)P = P - P^2 = 0$  ist  $\text{Kern}(Q) \supset \text{Im}(P)$ . Andererseits ist  $v \in \text{Kern}(Q)$  wegen  $Pv = v$  auch im Bild von  $P$ .
- Ist  $v \in \text{Kern}(P)$ , dann ist  $Pv = 0$ ; ist  $v \in \text{Im}(P)$ , dann ist  $v \in \text{Kern}(Q)$ , also  $Pv = v$ . Aus (4.3) ergibt sich daher  $\text{Spek}(P) \subset \{0, 1\}$ .  $\square$

**4.5 Beispiel** 1. Die Identische und die Nullabbildung sind Projektionen.

2. Ist  $V$  ein  $\mathbb{R}$ -Vektorraum mit dem Skalarprodukt  $(\cdot, \cdot)$  und  $b \in V$  ein Vektor der Länge  $\|b\| \equiv \sqrt{(b, b)} = 1$ , dann ist  $P : V \rightarrow V$ ,  $P(x) := (b, x)b$  Projektor auf den eindimensionalen Unterraum  $\text{span}(b) \subset V$ , denn

$$P \circ P(x) = (b, x)(b, b)b = (b, x)b = P(x)$$

und  $\text{Im}(P) = \text{span}(b)$ . Hier ist  $\text{Kern}(P) = \{v \in V \mid (b, v) = 0\}$  der zu  $b$  senkrechte Unterraum.

Etwa für  $b := \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2}\sqrt{3} \end{pmatrix} \in \mathbb{R}^2$  ergibt sich  $P(x) = \frac{1}{4} \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix} x$ .

Ist

$$V = \bigoplus_{i \in I} V_i$$

direkte Summe der Unterräume  $V_i$ , dann besitzt nach Definition jeder Vektor  $v \in V$  die eindeutige Zerlegung

$$v = \{(v_i)_{i \in I}\},$$

wobei nur endlich viele  $v_i \neq 0$  sind. Für jedes  $j \in I$  ist die Abbildung  $v \mapsto v_j$  linear.

Definieren wir daher  $P_j \in \text{End}(V)$  durch

$$P_j(v) := \{(w_i)_{i \in I}\} \quad \text{mit} \quad w_i := \begin{cases} v_j & , \quad i = j \\ 0 & , \quad i \neq j, \end{cases}$$

dann ist  $\text{Im}(P_j) = V_j$ ,  $P_j^2 = P_j$  und  $P_i P_j = 0$  für  $i \neq j$  oder kurz

$$P_i P_j = \delta_{ij} P_i.$$

Für endliche Summen ( $|I| < \infty$ ) gilt

$$\sum_{i \in I} P_i = \text{Id}$$

In diesem Sinn ist die direkte Summe mit einer Familie von Projektoren verknüpft.

Ein Endomorphismus  $f \in L(V)$  heißt *diagonalisierbar*, wenn eine Basis von Eigenvektoren von  $f$  existiert.

**4.6 Satz** Ein diagonalisierbarer Endomorphismus  $f \in L(V)$  eines endlichdimensionalen Vektorraumes besitzt die Form

$$f = \sum_{\lambda \in \text{Spek}(f)} \lambda P_\lambda, \quad (4.4)$$

wobei

$$P_\lambda := \prod_{\mu \in \text{Spek}(f) \setminus \{\lambda\}} \frac{f - \mu \cdot \text{Id}}{\lambda - \mu}$$

Projektor auf den Eigenraum  $E(\lambda)$  zum Eigenwert  $\lambda \in \text{Spek}(f)$  ist. Es gilt

$$P_\lambda P_{\lambda'} = \delta_{\lambda\lambda'} P_\lambda \quad , \quad \sum_{\lambda \in \text{Spek}(f)} P_\lambda = \text{Id} \quad \text{und} \quad V = \bigoplus_{\lambda \in \text{Spek}(f)} E(\lambda). \quad (4.5)$$

**Bew.:** • Da  $E(\lambda)$  von den Basisvektoren aufgespannt wird, die Eigenvektoren zum Eigenwert  $\lambda$  sind, ergibt sich die Darstellung von  $V$  als direkte Summe.

• Es ist für einen Eigenvektor  $v_\nu \in E(\nu)$  zum Eigenwert  $\nu \in \text{Spek}(f)$

$$\begin{aligned} P_\lambda P_{\lambda'} v_\nu &= \prod_{\mu \in \text{Spek}(f) \setminus \{\lambda\}} \prod_{\mu' \in \text{Spek}(f) \setminus \{\lambda'\}} \frac{(f - \mu \cdot \text{Id})(f - \mu' \cdot \text{Id})}{(\lambda - \mu)(\lambda' - \mu')} v_\nu \\ &= \prod_{\mu \in \text{Spek}(f) \setminus \{\lambda\}} \prod_{\mu' \in \text{Spek}(f) \setminus \{\lambda'\}} \frac{(\nu - \mu)(\nu - \mu')}{(\lambda - \mu)(\lambda' - \mu')} v_\nu. \end{aligned}$$

Dieses Produkt verschwindet für  $\nu \neq \lambda$  (denn dann tritt  $\mu = \nu$  auf) und analog für  $\nu \neq \lambda'$ . Im Fall  $\nu = \lambda = \lambda'$  sind alle Faktoren des Produktes Eins.

• Die mittlere Identität in (4.5) folgt aus der ersten, wenn wir die Summe der Projektoren auf die (eindeutige) Summendarstellung  $v = \sum_{\nu \in \text{Spek}(f)} v_\nu$  mit  $v_\nu \in E(\nu)$  eines Vektors  $v \in V$  anwenden. Analog folgt (4.4).  $\square$

## 5 Die Jordansche Normalform

### 5.1 Äquivalenz und Ähnlichkeit

Wie wir wissen, besteht zwischen Matrizen aus  $\mathcal{M}(n \times m, K)$  und linearen Abbildungen  $\varphi : V \rightarrow W$   $m$ - bzw.  $n$ -dimensionaler  $K$ -Vektorräume  $V$  bzw.  $W$  ein enger Zusammenhang.

Ist  $B = (b_1, \dots, b_m)$  eine Basis von  $V$  und  $C = (c_1, \dots, c_n)$  eine Basis von  $W$ , dann ist die *darstellende Matrix*  $\mathcal{M}_C^B(\varphi) \in \mathcal{M}(n \times m, K)$  diejenige Matrix, bezüglich derer der Vektor  $\sum_{i=1}^m v_i b_i \in V$  unter  $\varphi$  auf den Vektor  $\sum_{j=1}^n w_j c_j$  mit

$$w_j = \sum_{i=1}^m (\mathcal{M}_C^B(\varphi))_{ji} v_i$$

abgebildet wird.

Wir stellen uns nun die Frage, wie wir durch geeignete Wahl der Basen eine besonders einfache Gestalt der darstellenden Matrix erreichen können. Basiswechsel entspricht der Multiplikation mit einer regulären Matrix. Wir definieren daher:

**5.1 Definition** Zwei Matrizen  $A, B \in \mathcal{M}(n \times m, K)$  heißen äquivalent, wenn es Matrizen  $S \in GL(n, K)$  und  $T \in GL(m, K)$  mit

$$B = SAT$$

gibt.

Offensichtlich ist diese Relation eine Äquivalenzrelation auf  $\mathcal{M}(n \times m, K)$ , denn aus der Gruppeneigenschaft von  $GL(r, K)$  folgen die Eigenschaften von Reflexivität, Symmetrie und Transitivität.

Unsere Aufgabe besteht nun darin, aus jeder Äquivalenzklasse einen eindeutigen Repräsentanten auszuwählen. Dieser *Normalform* soll man die Charakteristika der Äquivalenzklasse möglichst sofort ansehen.

Wir kennen schon ein solches Charakteristikum, den Rang, also die Dimension des Bildes der linearen Abbildung  $\varphi : K^m \rightarrow K^n, \varphi(x) := Ax$ . Denn wegen der

Invertierbarkeit von  $S$  und  $T$  gilt

$$\text{Rang}(SAT) = \text{Rang}(A).$$

Da Zeilenrang gleich Spaltenrang ist, gilt  $0 \leq \text{Rang}(A) \leq \min(m, n)$ . Besonders einfache Matrizen vom Rang  $r \in \{0, \dots, \min(m, n)\}$  sind sicher die

$$D_r \in \mathcal{M}(n \times m, K) \quad \text{mit} \quad (D_r)_{ij} = \delta_{ij} \cdot \begin{cases} 1 & , \quad i \leq r \\ 0 & , \quad i > r. \end{cases}$$

Beispielsweise ist  $D_r \in \mathcal{M}(4 \times 3, K)$  gleich

$$D_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad D_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad D_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

**5.2 Satz** Ist  $r := \text{Rang}(A)$ , dann ist  $A \in \mathcal{M}(n \times m, K)$  zu  $D_r$  äquivalent, aber nicht zu  $D_s$ , falls  $s \neq r$ .

**Bew.:**

- Der Defekt der linearen Abbildung  $\varphi(x) = Ax$  ist  $\text{Def}(\varphi) = m - r$ . Wir ergänzen eine Basis  $b_{r+1}, \dots, b_m$  von  $\text{Kern}(\varphi) \subset K^m$  zu einer Basis  $B = (b_1, \dots, b_m)$  und setzen  $c_i := \varphi(b_i)$ ,  $i = 1, \dots, r$ . Die  $c_i \in K^n$  sind linear unabhängig, denn andernfalls wäre  $\text{Def}(\varphi) > m - r$ . Wir ergänzen sie zu einer Basis  $C = (c_1, \dots, c_n)$  des  $K^n$ .

Es gilt dann  $\mathcal{M}_C^B(\varphi) = D_r$ .

- Da  $\text{Rang}(D_s) = s \neq r = \text{Rang}(A)$ , ist  $A$  nicht zu  $D_s$  äquivalent.  $\square$

Solange wir also keine Zusatzstrukturen wie z.B. Skalarprodukte auf den endlich-dimensionalen Vektorräumen  $V$  und  $W$  berücksichtigen, hat bezüglich geeigneter Basen die lineare Abbildung  $\varphi: V \rightarrow W$  die Form  $D_r$ .  $\text{Rang}(\varphi)$  ist mit anderen Worten die einzige Invariante von  $\varphi$  unter Basistransformationen.

**5.3 Korollar** Sind  $V$  und  $W$  endlich-dimensionale  $K$ -Vektorräume und  $\varphi, \psi \in L(V, W)$  mit  $\text{Rang}(\varphi) = \text{Rang}(\psi)$ , dann existieren  $\alpha \in GL(V)$  und  $\beta \in GL(W)$  mit

$$\psi = \beta \circ \varphi \circ \alpha.$$

Ist nun  $\varphi \in L(V)$ , also ein Endomorphismus, dann wird man sich im Allgemeinen nicht den Luxus zweier Basen auf  $V$  leisten. Die Frage, wie durch geschickte Wahl nur *einer* Basis  $B$  von  $V$  die darstellende Matrix  $\mathcal{M}_B(\varphi) \equiv \mathcal{M}_B^B(\varphi)$  auf eine möglichst einfache Normalform (die so genannte Jordansche Normalform) gebracht werden kann, ist nun schwieriger als die eben untersuchte Problematik.

Insbesondere können wir nicht mehr erwarten, dass diese Normalform immer eine Diagonalmatrix aus  $\mathcal{M}(n, K)$  ist, und zwar aus zwei, in den der Definition folgenden Beispielen illustrierten, Gründen.

**5.4 Definition** • Sind  $\varphi, \psi \in L(V)$  und existiert  $\beta \in GL(V)$  mit  $\psi = \beta \circ \varphi \circ \beta^{-1}$ , dann heißen  $\varphi$  und  $\psi$  ähnlich oder konjugiert.

- Sind  $A, B \in \mathcal{M}(n, K)$  und existiert  $S \in GL(n, K)$  mit  $B = SAS^{-1}$ , dann heißen  $A$  und  $B$  ähnlich oder konjugiert.

Ähnlichkeit ist wieder eine Äquivalenzrelation, und der Rang verändert sich unter Konjugationen nicht. Allerdings ist diese Zerlegung in Äquivalenzklassen feiner als die vorher untersuchte, denn das Spektrum ist unter Konjugationen invariant.

**5.5 Beispiel** 1.  $n = 1$ . Zwei Körperelemente  $A, B \in \mathcal{M}(1, K) = K$  sind nur dann ähnlich, wenn sie gleich sind.

2.  $n = 2, A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Zwar ist  $\text{Rang}(A) = \text{Rang}(B) = 1$ ,  $A$  und  $B$  sind also äquivalent, aber nicht ähnlich, denn  $\text{Spek}(A) = \{0, 1\} \neq \text{Spek}(B) = \{0\}$ .

$B$  ist zu überhaupt keiner Diagonalmatrix konjugiert, denn diese hätte das gleiche Spektrum, müsste also die Nullmatrix sein, was wegen  $\text{Rang}(B) = 1$  nicht sein kann.

3.  $A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \in \mathcal{M}(2, \mathbb{R})$  ist für  $\sin \varphi \neq 0$  zu keiner (reellen) Diagonalmatrix konjugiert, denn das charakteristische Polynom  $\chi_A(x) = x^2 - 2x \cos(\varphi) + 1$  besitzt dann keine reellen Nullstellen.

Fasst man allerdings  $A$  als Element von  $\mathcal{M}(2, \mathbb{C})$  auf, dann ist  $A$  zur Diagonalmatrix  $\begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}$  konjugiert, wobei  $\lambda = \cos \varphi + i \sin \varphi$  und die konjugierende Matrix z.B.  $S = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$ , also  $S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$ .

Das letzte Beispiel zeigt einmal wieder, dass die komplexen Zahlen sich algebraisch einfacher verhalten als die reellen.

**5.6 Definition** 1. Zerfällt jedes (normierte) Polynom  $p \in K[x]$   $n$ -ten Grades ( $n \geq 1$ ) in Linearfaktoren, d.h.

$$p(x) = (x - \lambda_1) \cdot \dots \cdot (x - \lambda_n),$$

dann heißt der Körper  $K$  algebraisch abgeschlossen.

2. Ein Körper  $\overline{K} \supset K$  heißt algebraischer Abschluss des Unterkörpers  $K$ , wenn  $\overline{K}$  algebraisch abgeschlossen ist, es aber keinen algebraisch abgeschlossenen Körper  $K' \supset K$  mit  $K' \subsetneq \overline{K}$  gibt.

**5.7 Satz** Jeder Körper  $K$  besitzt einen algebraischen Abschluss  $\overline{K}$ , und dieser ist (bis auf Isomorphie) eindeutig.

**Bew.:** Siehe z.B. [6]. □

**5.8 Beispiel** Der algebraische Abschluss  $\overline{K}$  des Körpers  $K := \mathbb{R}$  der reellen Zahlen muss  $\mathbb{C}$  enthalten, denn in  $\mathbb{C}$  zerfällt das Polynom  $x^2 + 1 \in \mathbb{R}[x]$  in die Linearfaktoren  $(x + i)(x - i)$ .

Im Prinzip könnte es nun komplexe Polynome  $p \in \mathbb{C}[x]$  geben, die in  $\mathbb{C}$  nicht in Linearfaktoren zerfallen. Der (in der Analysis bewiesene!) *Fundamentalsatz der Algebra* besagt aber, dass  $\mathbb{C}$  algebraisch abgeschlossen ist. Also ist  $\overline{\mathbb{R}} = \mathbb{C}$ .

Wir wollen nun zunächst das Normalformproblem für Endomorphismen bzw. quadratische Matrizen über abgeschlossenen Körpern behandeln und danach auf allgemeine Körper, insbesondere  $\mathbb{R}$ , zurückkommen:

**5.9 Definition** Ein Endomorphismus  $\varphi \in L(V)$  eines  $K$ -Vektorraumes  $V$  heißt

- nilpotent, wenn es ein  $n \in \mathbb{N}$  mit  $\varphi^n = 0$  gibt, und das kleinste solche  $n$  heißt dann Nilpotenzindex,
- diagonalisierbar, wenn  $V$  eine Basis besitzt, die aus Eigenvektoren von  $\varphi$  besteht.

Eine Matrix  $A \in \mathcal{M}(n, K)$  heißt

- nilpotent, wenn es ein  $n \in \mathbb{N}$  mit  $A^n = 0$  gibt, und das kleinste solche  $n$  heißt dann Nilpotenzindex,
- diagonalisierbar, wenn sie konjugiert zu einer Diagonalmatrix ist.

**5.10 Beispiel** 1. Für  $V := \{p \in \mathbb{R}[x] \mid \deg(p) \leq n\}$  ist der Endomorphismus  $\varphi(p) := p'$ , der einem Polynom  $p$  seine Ableitung  $p'$  zuordnet, nilpotent mit Nilpotenzindex  $n + 1$ , denn  $\varphi^{n+1} = 0$ , aber  $\varphi^n(p) = n!$  für  $p(x) = x^n$ .  $\varphi$  ist aber (außer für  $n = 0$ ) nicht diagonalisierbar, denn nur die konstanten Polynome  $p(x) = a \in \mathbb{R} \setminus \{0\}$  sind Eigenvektoren von  $\varphi$ .

2. Über dem gleichen Vektorraum ist der Endomorphismus  $\varphi(p)(x) := x^n p(1/x)$  nicht nilpotent, aber diagonalisierbar, denn  $\varphi$  besitzt die Basis von Eigenvektoren der Form  $x^l \pm x^{n-l}$  mit den Eigenwerten  $\pm 1$  (dabei ist  $l \in \{0, \dots, [n/2]\}$ ; für gerade  $n$  lässt man den Nullvektor  $x^{n/2} - x^{n/2}$  weg).

3. Über dem gleichen Vektorraum ist der Endomorphismus  $\varphi(p)(x) := p(x + 1)$  weder nilpotent (denn er erhält den Grad von  $p$ ) noch für  $n > 0$  diagonalisierbar (denn die konstanten Polynome sind die einzigen Eigenvektoren von  $\varphi$ ).

Die nilpotenten und die diagonalisierbaren Endomorphismen bzw. Matrizen werden die Bausteine sein, aus denen wir die Jordan-Normalform konstruieren.

Als Normalform einer diagonalisierbaren Matrix  $A$  wählen wir eine Diagonalmatrix, zu der  $A$  konjugiert ist. Diese besitzt die Eigenwerte von  $A$  als Diagonalelemente und ist somit bis auf Permutation der Diagonaleinträge eindeutig durch  $A$  bestimmt.

## 5.2 Nilpotente Endomorphismen

Um Normalformen für nilpotente Matrizen zu konstruieren, definieren wir für  $r \in \mathbb{N}$  die nilpotente Matrix

$$N_r \in \mathcal{M}(r, K) \quad , \quad (N_r)_{ik} := \delta_{i, k-1}.$$

Diese besitzt also in der oberen Nebendiagonale lauter Einsen, und alle anderen Einträge sind Null.

Der Nilpotenzindex von  $N_r$  ist  $r$  denn  $(N_r)^l$  besitzt außer Einsen in der  $l$ -ten oberen Nebendiagonale nur Nullen, z.B.

$$N_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad (N_3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad (N_3)^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Durch direkte Summenbildung können wir nilpotente Matrizen der Form

$$\bigoplus_{i=1}^j N_{r_i} \in \mathcal{M}(r, K) \tag{5.1}$$

mit  $r_1, \dots, r_j \in \mathbb{N}$  bilden, wobei  $r := \sum_{i=1}^j r_i$  ist.

**5.11 Beispiel**  $N_3 \oplus N_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad , \quad (N_3 \oplus N_2)^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$

Deren Nilpotenzindex ist aber nicht etwa  $r$ , sondern nur  $\max(r_1, \dots, r_j)$ , denn

$$\left( \bigoplus_{i=1}^j A_i \right)^k = \bigoplus_{i=1}^j (A_i^k).$$

Wir wollen nun zeigen, dass jede nilpotente Matrix zu einer Matrix der Form (5.1) konjugiert ist, sodass wir letztere als Normalformen wählen können. Um die Geometrie des Problems herauszuarbeiten, wechseln wir zu nilpotenten Endomorphismen über.

Schreiben wir

$$\tilde{N}_r \in L(K^r) \quad , \quad \tilde{N}_r(x) := N_r x$$

für die der Matrix  $N_r$  zugeordnete lineare Abbildung, dann fällt auf, dass wegen  $\tilde{N}_r(e_1) = 0$  und  $\tilde{N}_r(e_s) = e_{s-1}$  ( $s = 2, \dots, r$ ) für die kanonische Basis  $(e_1, \dots, e_r)$  des  $K^r$  dieser Vektorraum die Basis

$$(e_r, \tilde{N}_r(e_r), \tilde{N}_r^2(e_1), \dots, \tilde{N}_r^{r-1}(e_r))$$

besitzt, wir also eine Basis des Vektorraumes durch wiederholte Anwendung des nilpotenten Endomorphismus auf einen einzigen Vektor  $w \in K^r$  gewinnen können. Dabei hätten wir statt  $w = e_r$  einen beliebigen Vektor  $w \in K^r \setminus \text{Im}(\tilde{N}_r)$  mit nicht verschwindender  $r$ -ter Komponente zum Ausgangspunkt wählen können.

**5.12 Definition** Ein  $r$ -dimensionaler Unterraum  $W \subset V$  heißt zyklisch bezüglich des Endomorphismus  $\varphi \in L(V)$ , wenn es einen Vektor  $w \in W$  gibt, sodass

- $(w, \varphi(w), \varphi^2(w), \dots, \varphi^{r-1}(w))$  eine Basis von  $W$  ist und
- $\varphi^r(w) = 0$  ist.

$w$  heißt dann erzeugender Vektor.

**5.13 Beispiel** 1. Bezüglich des Ableitungsoperators  $\varphi(p) := p'$  auf dem Polynomraum  $V := \mathbb{R}[x]$  ist jeder Unterraum  $W := \{p \in V \mid \deg(p) \leq n\}$  zyklisch, wobei jeder Vektor  $w$  mit  $\deg(w) = n$  erzeugend ist.

Wählen wir als erzeugenden Vektor  $v$  das Monom  $v(x) = x^n$ , dann ergibt sich als Basis

$$(v^{(n)}, v^{(n-1)}, \dots, v^{(1)}, v) = \left( n!, \frac{n!}{1!}x, \dots, \frac{n!}{(n-1)!}x^{n-1}, x^n \right).$$

In dieser Basis besitzt  $\varphi|_W$  die Form  $\tilde{N}_n$ , also die Normalform.

2. Bezüglich des Operators  $\bigoplus_{i=1}^j \tilde{N}_{r_i}$  auf dem Summenraum  $V := \bigoplus_{i=1}^j W_i$  mit  $W_i := K^{r_i}$  ist der Vektor  $0 \oplus \dots \oplus 0 \oplus e_{r_i} \oplus 0 \oplus \dots \oplus 0$  (mit  $e_{r_i}$  an der  $i$ -ten Stelle) erzeugender Vektor von  $W_i$ .

Nehmen wir einmal an, wir wüssten schon, dass jede nilpotente Matrix  $A \in \mathcal{M}(r, K)$  konjugiert zu einer Normalform (5.1) ist, dann müssten wir nur  $j$  zyklische Vektoren finden, um die konjugierende Matrix  $S$  zu berechnen. Wir gehen wieder zum geometrischen Problem nilpotenter Endomorphismen  $\varphi \in L(V)$  über. Was wir zunächst berechnen können, sind die Kerne der iterierten Abbildungen  $\varphi^k$ , und tatsächlich reicht das aus.

So ist jeder Vektor  $w \in W \setminus \text{Kern}(\varphi^{n-1})$  in Beispiel 5.10.1. erzeugender Vektor.

Im allgemeinen Fall ist zunächst zu beweisen, dass  $V$  direkte Summe von Unterräumen  $V = \bigoplus_{i=1}^j W_i$  ist, sodass die  $W_i$  bezüglich  $\varphi$  zyklisch sind.

Die vorhandenen Daten sind zunächst die Dimensionen

$$d_p := \text{Def}(\varphi^p) \equiv \text{Dim}(V_p) \quad (p = 0, \dots, r+1)$$

der Kerne  $V_p := \text{Kern}(\varphi^p)$ . Aus diesen lassen sich die Dimensionen der zyklischen Unterräume  $W_i$  berechnen.

Ist  $q$  der Nilpotenzindex von  $\varphi$ , dann gilt

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_q = V,$$

und  $d_{p+1} > d_p$  für  $p < q$ , denn wären für ein  $p < q$  die Dimensionen von  $V_p$  und von  $V_{p+1}$  gleich, dann wäre auch  $V_p = V_{p+1}$ , also auch  $V_p = \dots = V_q$ .

**5.14 Lemma** • *Es ist  $V_{p+1} = \varphi^{-1}(V_p)$  ( $p = 0, \dots, q-1$ ).*

•  *$\varphi$  erzeugt injektive Abbildungen*

$$\bar{\varphi}_p : \bar{V}_{p+1} \rightarrow \bar{V}_p \quad v + V_p \mapsto \varphi(v) + V_{p-1}$$

der Quotientenräume  $\bar{V}_p := V_p/V_{p-1}$ .

Mit anderen Worten: Sind  $v, v' \in V_{p+1}$  und  $v - v' \notin V_p$ , dann ist auch  $\varphi(v) - \varphi(v') \notin V_{p-1}$ .

**Bew.:**

- $V_{p+1} \equiv \text{Kern}(\varphi^{p+1}) = \text{Kern}(\varphi \circ \varphi^p) = \{v \in V \mid \varphi(v) \in V_p\} = \varphi^{-1}(V_p)$ .
- Es sei  $w \in V_{p+1}$ . Wir müssen zeigen: Ist  $\varphi(w) \in V_{p-1}$ , dann ist  $w \in V_p$ .

Der erste Teil des Lemmas besagt aber gerade  $V_p = \varphi^{-1}(V_{p-1})$ . □

Aus  $V_{p+1} = \varphi^{-1}(V_p)$  folgt nur  $\varphi(V_{p+1}) \subset V_p$  und nicht etwa  $\varphi(V_{p+1}) = V_p$ . Die Abbildungen  $\bar{\varphi}_p$  sind also im Allgemeinen nicht surjektiv, bilden also  $\bar{V}_{p+1}$  nur injektiv in einen Unterraum von  $\bar{V}_p$  ab.

Setzen wir nun

$$\bar{d}_p := d_p - d_{p-1} = \text{Dim}(V_p) - \text{Dim}(V_{p-1}) = \text{Dim}(\bar{V}_p), \quad (5.2)$$

dann ist

$$k_p := \bar{d}_p - \bar{d}_{p+1} = \text{Dim}(\bar{V}_p) - \text{Dim}(\bar{V}_{p+1}) \quad (5.3)$$

die Kodimension dieses Bildes  $\bar{\varphi}_p(\bar{V}_{p+1})$  in  $\bar{V}_p$ .

Diese Dimensionszählerei nützt uns jetzt bei der Konstruktion einer Basis von  $V$ , bezüglich derer  $\varphi$  die nilpotente Normalform (5.1) besitzt.

Wir wählen nämlich für  $p = 1, \dots, q$  Vektoren

$$e_{p,1}, \dots, e_{p,k_p} \in V_p$$

so, dass sie Repräsentanten einer Basis von  $\overline{V}_p/\overline{\varphi}_p(\overline{V}_{p+1})$  sind. Dies ist möglich, denn ihre Anzahl  $k_p$  ist gleich der Dimension von  $\overline{V}_p/\overline{\varphi}_p(\overline{V}_{p+1})$ .

Diese Vektoren  $e_{p,j}$  nennen wir *primitiv*, und leiten aus ihnen die Vektoren

$$\varphi^r(e_{p,j}) \quad (r = 0, \dots, p-1; j = 1, \dots, k_p; p = 1, \dots, q)$$

ab. Deren Zahl ist gleich

$$\sum_{p=1}^q p \cdot k_p = \sum_{p=1}^q p \cdot (\overline{d}_p - \overline{d}_{p+1}) = \sum_{p=1}^q \overline{d}_p = \sum_{p=1}^q (d_p - d_{p-1}) = d_q = \text{Dim}(V),$$

und tatsächlich bilden sie eine Basis von  $V$ .

**5.15 Satz** 1. Für  $p = 1, \dots, q$  repräsentieren die Vektoren

$$\varphi^{p'-p}(e_{p',j}) \quad (p' = p, \dots, q; j = 1, \dots, k_{p'})$$

eine Basis von  $\overline{V}_p \equiv V_p/V_{p-1}$ .

2. Die Vektoren

$$\varphi^r(e_{p,j}) \quad (r = 0, \dots, p-1; j = 1, \dots, k_p; p = 1, \dots, q)$$

bilden eine Basis von  $V$ .

3. Die Unterräume

$$W_{p,j} := \text{span}(e_{p,j}, \dots, \varphi^{p-1}(e_{p,j})) \quad (j = 1, \dots, k_p; p = 1, \dots, q)$$

sind zyklisch bezüglich  $\varphi$  mit erzeugendem Vektor  $e_{p,j}$  und besitzen Dimension  $p$ .

4.  $V$  ist die innere direkte Summe dieser Unterräume

$$V = \bigoplus_{\substack{p=1, \dots, q \\ j=1, \dots, k_p}} W_{p,j}.$$

**Bew.:** (Siehe auch Brieskorn, p. 15 [1]):

- Für  $p = q$  ist  $\overline{V}_{p+1} = V/V$  nulldimensional; also repräsentieren  $e_{p,1}, \dots, e_{p,k_p}$  eine Basis von  $\overline{V}_p/\overline{\varphi}_p(\overline{V}_{p+1}) = \overline{V}_p = V_p/V_{p-1}$ .

- Behauptung 1. sei schon für  $p + 1 \leq q$  bewiesen, und wir wollen sie für  $p$  beweisen. Wir wissen also, dass die Vektoren

$$\varphi^{p'-(p+1)}(e_{p',j}) \quad (p' = p + 1, \dots, q; j = 1, \dots, k_{p'})$$

eine Basis von  $\overline{V}_{p+1}$  repräsentieren. Wegen der im letzten Lemma bewiesenen Injektivität von  $\overline{\varphi}_p$  repräsentieren ihre Bilder unter  $\varphi$  eine Basis von  $\overline{\varphi}_p(\overline{V}_{p+1})$ . Die primitiven Vektoren  $e_{p,1}, \dots, e_{p,k_p} \in V_p$  sind nun gerade so gewählt, dass sie diese zu einer Basis von  $\overline{V}_p$  ergänzen.

2. Wären diese Vektoren linear abhängig, gäbe es also eine Linearkombination der  $\varphi^r(e_{p,j})$ , die eine nichttriviale Darstellung des Nullvektors wäre, dann würden wir einen Koeffizienten  $\lambda_{p,j}^{(r)} \neq 0$  mit maximalem  $p-r =: p'$  finden. Wir könnten also diese Darstellung in der Form

$$\sum_{p-r=p'} \lambda_{p,j}^{(r)} \varphi^r(e_{p,j}) = - \sum_{p-r < p'} \lambda_{p,j}^{(r)} \varphi^r(e_{p,j})$$

schreiben. Die linke Seite repräsentiert einen Vektor aus  $\overline{V}_{p'}$ , die rechte Seite ist ein Vektor aus  $V_{p'-1}$ . Da die Vektoren  $\varphi^r(e_{p,j})$  der linken Seite eine Basis von  $\overline{V}_{p'} = V_{p'}/V_{p'-1}$  repräsentieren, müssen ihre Koeffizienten alle verschwinden. Widerspruch!

3. Nach 2. sind die Vektoren  $e_{p,j}, \dots, \varphi^{p-1}(e_{p,j})$  linear unabhängig und bilden daher eine Basis von  $W_{p,j}$ .
4. folgt, da für  $(p, j) \neq (p', j')$  die Unterräume  $W_{p,j}$  und  $W_{p',j'}$  von verschiedenen Basisvektoren aufgespannt werden.  $\square$

Da die auf die Unterräume  $W_{p,j}$  restringierte Abbildung  $\varphi|_{W_{p,j}}$  in der geordneten Basis

$$(\varphi^{p-1}(e_{p,j}), \varphi^{p-2}(e_{p,j}), \dots, \varphi(e_{p,j}), e_{p,j})$$

die Matrixdarstellung  $N_p$  besitzt, haben wir nebenbei gezeigt, dass  $\varphi$  zur nilpotenten Jordannormalform

$$\bigoplus_{\substack{p=1, \dots, q \\ j=1, \dots, k_p}} \tilde{N}_p$$

konjugiert ist.

**5.16 Beispiel** Auf dem Vektorraum  $\mathbb{R}[x]$  der Polynome betrachten wir den Operator

$$\hat{\varphi}(p)(x) := p'(x) - xp''(x) + p^{(3)}(x),$$

wobei  $p^{(n)}$  die  $n$ -te Ableitung des Polynoms  $p$  bezeichnet. Ist eine Lösung  $p$  der Gleichung  $\hat{\varphi}(p) = 0$  gesucht (wobei  $p$  kein Polynom mehr sein muss), dann spricht man von einer linearen Differentialgleichung. Lineare Differentialgleichungen, allerdings mit konstanten Koeffizienten, werden uns noch in dieser Vorlesung begegnen.  $\hat{\varphi}$  selbst ist nicht nilpotent, aber da

$$\hat{\varphi}^2(p)(x) = (x^2 - 1)p^{(4)}(x) - 2xp^{(5)}(x) + p^{(6)}(x)$$

ist, verschwindet schon die dritte Iterierte  $\varphi^3$  von  $\varphi := \hat{\varphi} | V$  auf  $V := \{p \in \mathbb{R}[x] \mid \deg(p) \leq 4\}$ . Mit den Basisvektoren  $e_i(x) := x^i$  von  $V$  ergibt sich

$$V_0 = \{0\}, V_1 = \text{span}(e_0, e_1), V_2 = \text{span}(e_0, e_1, e_2, e_3), V_3 = V.$$

Daher repräsentiert  $e_4$  eine Basis von  $\bar{V}_3 = V_3/V_2$ ,  $e_1, e_3$  eine Basis von  $\bar{V}_2 = V_2/V_1$  und  $e_0, e_2$  eine Basis von  $\bar{V}_1 = V_1$ . Wir wählen also als primitive Vektoren  $e_{3,1} := e_4, e_{2,1} := e_1$ , sodass  $\varphi(e_{3,1}) = 24e_1 - 8e_3, \varphi^2(e_{3,1}) = 24(e_2 - e_0)$  und  $\varphi(e_{2,1}) = e_0$  ist. Somit besitzt  $\varphi$  die Normalform  $\tilde{N}_3 \oplus \tilde{N}_2$ .

Wir können durch Umnummerieren der Basisvektoren immer erreichen, dass in der nilpotenten Normalform

$$\tilde{N}_{r_1} \oplus \tilde{N}_{r_2} \oplus \dots \oplus \tilde{N}_{r_j} \quad r_1 \geq r_2 \geq \dots \geq r_j$$

gilt. Die Summe  $r := r_1 + r_2 + \dots + r_j$  dieser Nilpotenzindizes ist gleich der Dimension des Vektorraumes  $V$ . Eine solche Zerlegung einer natürlichen Zahl  $r \in \mathbb{N}$  in absteigend geordnete Summanden  $r_i \in \mathbb{N}$  heißt eine *Partition* von  $r$ , und  $p(r)$  bezeichnet die Anzahl voneinander verschiedener Partitionen von  $r$ .

**5.17 Beispiel** •  $p(1) = 1$ .

- Da  $2 = 1 + 1$ , ist  $p(2) = 2$ .
- Da  $3 = 2 + 1 = 1 + 1 + 1$ , ist  $p(3) = 3$ .
- Da  $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$ , ist  $p(4) = 5$ .

**5.18 Satz** Auf einem  $r$ -dimensionalen Vektorraum  $V$  ist die Zahl der Äquivalenzklassen konjugierter nilpotenter Endomorphismen gleich  $p(r)$ , und diese besitzen die Normalformen

$$\tilde{N}_{r_1} \oplus \tilde{N}_{r_2} \oplus \dots \oplus \tilde{N}_{r_j} \tag{5.4}$$

mit  $r_1 \geq r_2 \geq \dots \geq r_j \geq 1$  und  $\sum_{i=1}^j r_i = r$ .

**Bew.:**

- Dass jeder nilpotente Endomorphismus  $\varphi \in L(V)$  zu einer dieser Normalformen konjugiert ist, haben wir schon gezeigt.
- Es ist also nur noch zu zeigen, dass zwei verschiedene solche Normalformen (5.4) nicht konjugiert sind. Nun läßt sich aber die Zahl  $k_p$  der Jordanblöcke  $\tilde{N}_{r_i}$  mit  $r_i = p$  nach (5.2) und (5.3) aus den Dimensionen  $d_p$  der Kerne  $V_p$  von (5.4) berechnen. Umgekehrt ist

$$d_p = \sum_{l=1}^p \bar{d}_l \quad \text{und} \quad \bar{d}_l = \sum_{s=l}^q k_s,$$

wir können die Defekte  $d_p$  also aus den Multiplizitäten  $k_s$  der Jordanblöcke der Größe  $s$  berechnen. Unter Konjugation bleiben die  $d_p$ , also auch die  $k_s$  invariant. Die absteigende Anordnung der Blockgrößen  $r_i$  impliziert daher die Eindeutigkeit.  $\square$

### 5.3 Verallgemeinerte Eigenräume

Bis jetzt haben wir das Normalformproblem für zwei Klassen von Endomorphismen  $\varphi \in L(V)$  eines  $r$ -dimensionalen  $K$ -Vektorraumes  $V$  gelöst:

- für diagonalisierbare  $\varphi$ ,
- für nilpotente  $\varphi$ .

Nun wollen wir diese Informationen nutzen, um einen beliebigen Endomorphismus  $\varphi$  (eindeutig) als Summe

$$\varphi = \varphi_D + \varphi_N \tag{5.5}$$

eines diagonalisierbaren und eines nilpotenten Endomorphismus darzustellen, die miteinander kommutieren ( $\varphi_D \circ \varphi_N = \varphi_N \circ \varphi_D$ ).

Da nilpotente  $\varphi \in L(V)$  die Normalform (5.1) besitzen, ist ihr charakteristisches Polynom gleich

$$\chi_\varphi(x) = x^r.$$

Dagegen besitzen diagonalisierbare  $\varphi \in L(V)$  das charakteristische Polynom

$$\chi_\varphi(x) = \prod_{\lambda \in \text{Spek}(\varphi)} (x - \lambda)^{m_\lambda}, \tag{5.6}$$

wobei  $m_\lambda$  die Multiplizität des Eigenwertes  $\lambda$  ist, d.h.  $m_\lambda := \text{Def}(\varphi - \lambda \cdot \text{Id})$ .

Wir benutzen daher das charakteristische Polynom dazu, die Zerlegung (5.5) zu finden.

**5.19 Definition** Es sei  $\varphi \in L(V)$  Endomorphismus des  $K$ -Vektorraumes  $V$  und  $\lambda \in K$  Eigenwert von  $\varphi$ .

- Dann heißt  $v \in V \setminus \{0\}$  verallgemeinerter Eigenvektor von  $\varphi$  zum Eigenwert  $\lambda$ , wenn ein  $q \in \mathbb{N}$  existiert sodass

$$(\varphi - \lambda \cdot \text{Id})^q v = 0.$$

- $\hat{E}(\lambda) := \bigcup_{q \in \mathbb{N}} \text{Kern}((\varphi - \lambda \cdot \text{Id})^q)$   
heißt verallgemeinerter Eigenraum von  $\varphi$  zum Eigenwert  $\lambda$ .

Für diagonalisierbare  $\varphi$  stimmen die Eigenräume  $E(\lambda) = \text{Kern}(\varphi - \lambda \text{Id})$  mit den verallgemeinerten Eigenräumen  $\hat{E}(\lambda)$  überein, für nilpotente  $\varphi$  dagegen ist ganz  $V$  verallgemeinerter Eigenraum (zum Eigenwert 0).

**5.20 Definition** Es sei  $\varphi \in L(V)$ . Dann heißt

- ein Unterraum  $U \subset V$   $\varphi$ -invariant, wenn  $\varphi(U) \subset U$ ,
- eine direkte Zerlegung  $V = \bigoplus_{i \in I} U_i$   $\varphi$ -invariant, wenn alle Unterräume  $U_i$   $\varphi$ -invariant sind.

**5.21 Satz** Es sei  $\psi \in L(V)$ . Gilt für ein  $q \in \mathbb{N}_0$

$$\text{Kern}(\psi^q) = \text{Kern}(\psi^{q+1}) \quad (5.7)$$

(mit  $\psi^0 = \text{Id}$ ), dann gilt für alle  $p \geq q$

$$\text{Kern}(\psi^p) = \text{Kern}(\psi^{p+1}), \quad (5.8)$$

und falls  $\text{Dim}(V) < \infty$

$$V = \text{Kern}(\psi^q) \oplus \text{Im}(\psi^q).$$

Diese Zerlegung ist dann  $\psi$ -invariant.

**Bew.:**

- Wir zeigen, dass aus (5.8) auch

$$\text{Kern}(\psi^{p+1}) = \text{Kern}(\psi^{p+2})$$

folgt. Ist  $v \in \text{Kern}(\psi^{p+2})$ , dann ist  $w := \psi(v) \in \text{Kern}(\psi^{p+1}) = \text{Kern}(\psi^p)$ , also  $\psi^p(w) = \psi^{p+1}(v) = 0$ .

- Es sei  $D := \text{Kern}(\psi^q) \cap \text{Im}(\psi^q)$ . Dann ist  $\psi^q(D) = \{0\}$ , und  $D = \psi^q(D')$  für  $D' := (\psi^q)^{-1}(D)$ . Damit ist  $D' \subset \text{Kern}(\psi^{2q}) = \text{Kern}(\psi^q)$ , also  $D = \{0\}$ .

- Da die Summe von Kern und Bild also direkt ist, gilt nach dem Dimensionssatz

$$\begin{aligned} \text{Dim}(\text{Kern}(\psi^q) \oplus \text{Im}(\psi^q)) &= \text{Dim}(\text{Kern}(\psi^q)) + \text{Dim}(\text{Im}(\psi^q)) \\ &= \text{Dim}(V) < \infty, \end{aligned}$$

sodass die direkte Summe schon gleich dem gesamten Vektorraum  $V$  ist.

- Für beliebige  $i \in \mathbb{N}$  gilt

$$\begin{aligned} \psi(\text{Kern}(\psi^i)) &= \{\psi(v) \mid \psi^{i-1}(\psi(v)) = 0\} \subset \{u \in V \mid \psi^{i-1}(u) = 0\} \\ &= \text{Kern}(\psi^{i-1}) \subset \text{Kern}(\psi^i) \end{aligned}$$

und

$$\psi(\text{Im}(\psi^i)) = \psi(\psi^i(V)) = \psi^i(\psi(V)) \subset \psi^i(V) = \text{Im}(\psi^i).$$

□

## 5.22 Bemerkungen

1. Ist  $\text{Dim}(V) = \infty$ , dann kann  $\text{Kern}(\psi^q) \oplus \text{Im}(\psi^q)$  ein echter Unterraum von  $V$  sein. Z.B. ist dies für  $V := \mathbb{R}[x]$  und  $\varphi \in L(V)$ ,  $\varphi(p)(x) := xp(x)$  der Fall, denn  $\varphi$  ist injektiv, also ist (5.7) für  $q = 0$  und damit auch für  $q = 1$  erfüllt, aber  $\text{Im}(\varphi)$  ist ein echter Unterraum von  $V$ , da er die Monome vom Grad 0 nicht enthält.
2. Den kleinsten Index  $q \in \mathbb{N}_0$  (soweit vorhanden), für den (5.7) erfüllt ist, nennt man den *Fittingindex* von  $\varphi$ .

**5.23 Korollar** Ist  $r := \text{Dim}(V) < \infty$ , dann ist der Fittingindex  $q \leq r$ .

**Bew.:**  $0 = \text{Def}(\psi^0) < \text{Def}(\psi^1) < \dots < \text{Def}(\psi^q) \leq r$ . □

Wir wollen nun zunächst zeigen, dass (in Verallgemeinerung von (4.5)) endlich-dimensionale Vektorräume  $V$  sich als direkte Summe

$$V = \bigoplus_{\lambda \in \text{Spek}(\varphi)} \hat{E}(\lambda)$$

der verallgemeinerten Eigenräume zerlegen lassen.

Dies ist nur möglich, wenn das charakteristische Polynom in Linearfaktoren zerfällt, also die Form (5.6) besitzt. Wir wissen schon, dass dies immer der Fall ist, wenn der Körper  $K$  algebraisch abgeschlossen ist, also z.B. für  $K = \mathbb{C}$ .

**5.24 Satz** Wenn das charakteristische Polynom  $\chi_\varphi \in K[x]$  des Endomorphismus  $\varphi \in L(V)$  auf dem endlich-dimensionalen  $K$ -Vektorraum  $V$  in Linearfaktoren zerfällt, dann ist

$$V = \bigoplus_{\lambda \in \text{Spek}(\varphi)} \hat{E}(\lambda), \quad (5.9)$$

und die Zerlegung (5.9) ist  $\varphi$ -invariant (es gilt also  $\varphi(\hat{E}(\lambda)) \subset \hat{E}(\lambda)$ ).

**Bew.:** • Für  $r := \text{Dim}(V)$  und  $v \in \hat{E}(\lambda)$  gilt

$$(\varphi - \lambda \text{Id})^r \circ \varphi(v) = \varphi \circ (\varphi - \lambda \text{Id})^r(v) = \varphi(0) = 0,$$

also  $\varphi(\hat{E}(\lambda)) \subset \hat{E}(\lambda)$ .

• Für  $\lambda, \mu \in \text{Spek}(\varphi)$  mit  $\lambda \neq \mu$  gilt

$$\hat{E}(\lambda) \cap \hat{E}(\mu) = \{0\}.$$

Denn der Unterraum  $U := \hat{E}(\lambda) \cap \hat{E}(\mu)$  ist, wie wir gerade gesehen haben, invariant unter  $\varphi$ .

- Nun besitzt für  $\text{Dim}(U) > 0$  der restringierte Endomorphismus  $\hat{\varphi} := \varphi|_U$  den Eigenwert  $\lambda$ , denn  $(\hat{\varphi} - \lambda \text{Id})^r U = \{0\}$ .
  - Andererseits besitzt  $\hat{\varphi}$  keinen weiteren Eigenwert  $\sigma \neq \lambda$ , denn mit  $(\hat{\varphi} - \sigma \text{Id})v = 0$  für  $v \in U \setminus \{0\}$  würde  $(\hat{\varphi} - \lambda \text{Id})^r v = (\hat{\varphi} - \lambda \text{Id})^{r-1} \cdot (\sigma - \lambda)v = \dots = (\sigma - \lambda)^r v \neq 0$  folgen.
  - Aus Symmetriegründen müsste aber auch  $\mu$  der einzige Eigenwert von  $\hat{\varphi}$  sein, was die Annahme  $\text{Dim}(U) > 0$  widerlegt.
- Es seien nun  $\lambda_1, \dots, \lambda_k \in \text{Spek}(\varphi)$  voneinander verschiedene Eigenwerte. Wir wollen zeigen, dass die Summe

$$\hat{E}(\lambda_1) + \dots + \hat{E}(\lambda_k)$$

der verallgemeinerten Eigenräume direkt ist, d.h. aus  $v_{\lambda_1} + \dots + v_{\lambda_k} = 0$  für  $v_{\lambda_i} \in \hat{E}(\lambda_i)$  folgt, dass  $v_{\lambda_i} = 0$ . Dies wurde gerade für  $k = 2$  bewiesen.

Der Induktionsschluss von  $k - 1$  auf  $k$  geht wie folgt: Es ist

$$(\varphi - \lambda_k \text{Id})^r v_{\lambda_k} = 0,$$

also auch

$$\sum_{i=1}^{k-1} (\varphi - \lambda_k \text{Id})^r v_{\lambda_i} = 0. \quad (5.10)$$

Wegen  $\varphi(\hat{E}(\lambda_i)) \subset \hat{E}(\lambda_i)$  ist auch  $(\varphi - \lambda_k \text{Id})^r \hat{E}(\lambda_i) \subset \hat{E}(\lambda_i)$ . Damit müssen für  $i = 1, \dots, k-1$  diese Vektoren in (5.10) die Nullvektoren sein, denn die Summe  $\hat{E}(\lambda_1) + \dots + \hat{E}(\lambda_{k-1})$  ist nach Induktionsvoraussetzung direkt. Es ist also

$$v_{\lambda_i} \in \hat{E}(\lambda_i) \cap \hat{E}(\lambda_k) = \{0\} \quad (i = 1, \dots, k-1)$$

und damit auch  $v_k = -(v_1 + \dots + v_{s-1}) = 0$ .

• Wir müssen noch zeigen, dass die direkte Summe

$$\bigoplus_{\lambda \in \text{Spek}(\varphi)} \hat{E}(\lambda) \subset V$$

gleich  $V$  ist. Dies ist für Dimension  $r = 1$  der Fall und sei allgemein für Dimensionen  $\leq r-1$  gezeigt.

Nach Satz 5.21 ist für  $\lambda \in \text{Spek}(\varphi)$  die Zerlegung

$$V = \hat{E}(\lambda) \oplus \text{Im}((\varphi - \lambda \text{Id})^r) \quad (5.11)$$

$\varphi$ -invariant. Daher können wir (falls nicht schon  $\hat{E}(\lambda) = V$  gilt)  $\hat{\varphi} := \varphi|_I$  auf dem Bildraum  $I := \text{Im}((\varphi - \lambda \cdot \text{Id})^r)$  betrachten. Da  $\text{Dim}(I) < \text{Dim}(V)$ , besitzt  $I$  nach Induktionsvoraussetzung eine  $\hat{\varphi}$ -invariante Zerlegung

$$I = \bigoplus_{\mu \in \text{Spek}(\hat{\varphi})} I_\mu \quad (5.12)$$

in verallgemeinerte Eigenräume  $I_\mu$  von  $\hat{\varphi}$ .

Es ist nach dem folgenden Lemma  $\text{Spek}(\hat{\varphi}) = \text{Spek}(\varphi) \setminus \{\lambda\}$ .

Daraus folgt aber

$$I_\mu \subset \hat{E}(\mu) \quad (\mu \in \text{Spek}(\varphi)), \quad (5.13)$$

denn wegen der  $\hat{\varphi}$ -Invarianz von  $I$  ist

$$\{0\} = (\hat{\varphi} - \mu \cdot \text{Id})^r(I_\mu) = (\varphi - \mu \cdot \text{Id})^r(I_\mu).$$

Insgesamt folgt aus (5.11), (5.12) und (5.13)

$$\begin{aligned} V &= \hat{E}(\lambda) \oplus I = \hat{E}(\lambda) \oplus \bigoplus_{\mu \in \text{Spek}(\hat{\varphi})} I_\mu = \\ &= \hat{E}(\lambda) \oplus \bigoplus_{\mu \in \text{Spek}(\varphi) \setminus \{\lambda\}} \hat{E}(\mu) = \bigoplus_{\mu \in \text{Spek}(\varphi)} \hat{E}(\mu). \end{aligned}$$

□

**5.25 Lemma** Es sei  $\varphi \in L(V)$ ,  $\dim(V) < \infty$  und  $U \subset V$  ein  $\varphi$ -invarianter Unterraum.

Dann ist für  $\hat{\varphi} := \varphi|_U$  das charakteristische Polynom  $\chi_{\hat{\varphi}}$  ein Teiler von  $\chi_{\varphi}$ . Zerfällt also  $\chi_{\varphi}$ , dann auch  $\chi_{\hat{\varphi}}$ .

**Bew.:** Ist  $(b_1, \dots, b_k)$  eine Basis von  $U$  und  $B = (b_1, \dots, b_k, \dots, b_r)$  eine Basis von  $V$ , dann ist die darstellende Matrix von  $\varphi$  von der Blockform

$$\mathcal{M}_B(\varphi) = M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad \text{mit} \quad M_{21} = 0$$

und  $M_{11} \in \mathcal{M}(k, K)$ ,  $M_{22} \in \mathcal{M}(r - k, K)$ .

Die gleiche Blockstruktur besitzt  $M - x \cdot \mathbb{1}_r$  für  $x \in K$ . Daher gilt (Übungsaufgabe)

$$\chi_M(x) = \chi_{M_{11}}(x) \cdot \chi_{M_{22}}(x) \quad (x \in K).$$

Wegen  $\chi_M = \chi_{\varphi}$  und  $\chi_{M_{11}} = \chi_{\hat{\varphi}}$  folgt daraus die Behauptung.  $\square$

## 5.4 Jordanzerlegung und Jordansche Normalform

Satz 5.24 ermöglicht es uns,  $\varphi \in L(V)$  in der Form (5.5) als Summe  $\varphi = \varphi_D + \varphi_N$  eines diagonalisierbaren und eines nilpotenten Endomorphismus zu schreiben.

Es sei dazu  $f_{\lambda}$  der Fittingindex von  $\varphi - \lambda \text{Id}$  und

$$R_{\lambda} \in K[x] \quad \text{das Polynom} \quad R_{\lambda}(x) := \prod_{\mu \in \text{Spek}(\varphi) \setminus \{\lambda\}} (x - \mu)^{f_{\mu}}. \quad (5.14)$$

Setzen wir  $\varphi$  in diese Polynome ein, dann ist

$$R_{\lambda}(\varphi)(V) \subset \hat{E}(\lambda),$$

denn  $V$  besitzt die  $\varphi$ -invariante Zerlegung (5.12), und  $(\varphi - \mu \text{Id})^{f_{\mu}}(\hat{E}(\mu)) = \{0\}$ . Andererseits sind die Polynome  $R_{\lambda} \in K[x]$  teilerfremd, d.h. es existiert kein Polynom  $R \in K[x]$  vom Grad  $> 0$ , das alle  $R_{\lambda}$  ohne Rest teilt.

Mithilfe des Euklidischen Algorithmus beweist man (Übungsaufgabe) die Existenz von Polynomen  $Q_{\lambda} \in K[x]$  mit

$$\sum_{\lambda \in \text{Spek}(\varphi)} Q_{\lambda} \cdot R_{\lambda} = 1. \quad (5.15)$$

**5.26 Satz (Jordanzerlegung)** Das charakteristische Polynom von  $\varphi \in L(V)$  zerfällt in Linearfaktoren. Dann bilden die Endomorphismen

$$P_{\lambda} := Q_{\lambda}(\varphi) \cdot R_{\lambda}(\varphi)$$

(mit den Polynomen  $R_\lambda$  aus (5.14) und  $Q_\lambda$  aus (5.15)) eine Familie von Projektoren auf die verallgemeinerten Eigenräume  $\hat{E}(\lambda)$  mit

$$P_\lambda P_\mu = \delta_{\lambda\mu} P_\lambda \quad \text{und} \quad \sum_{\lambda \in \text{Spek}(\varphi)} P_\lambda = \text{Id}_V. \quad (5.16)$$

Weiter gilt:

1. Der Endomorphismus

$$\varphi_D := \sum_{\lambda \in \text{Spek}(\varphi)} \lambda P_\lambda$$

ist diagonalisierbar, und seine Eigenräume zu den Eigenwerten  $\lambda$  stimmen mit den verallgemeinerten Eigenräumen  $\hat{E}(\lambda)$  von  $\varphi$  zu den gleichen Eigenwerten überein.

2. Der Endomorphismus  $\varphi_N := \varphi - \varphi_D$  lässt die Zerlegung (5.9) von  $V$  in die verallgemeinerten Eigenräume  $\hat{E}(\lambda)$  invariant, und  $\varphi_N \upharpoonright_{\hat{E}(\lambda)}$  besitzt Nilpotenzindex  $f_\lambda$ .

3. Es gilt  $\varphi_D \circ \varphi_N = \varphi_N \circ \varphi_D$ , d.h.  $\varphi_D$  und  $\varphi_N$  kommutieren.

4.  $\varphi = \varphi_D + \varphi_N$  ist die einzige Zerlegung von  $\varphi$  in einen diagonalisierbaren und einen nilpotenten Endomorphismus, die miteinander kommutieren.

**Bew.:** • Die zweite Identität aus (5.16) folgt aus (5.15).

• Dass  $P_\lambda(V) \subset \hat{E}(\lambda)$  ist, folgt wegen  $R_\lambda(\varphi)(V) \subset \hat{E}(\lambda)$  und der  $\varphi$ -Invarianz der verallgemeinerten Eigenräume. Da andererseits

$$\sum_{\lambda \in \text{Spek}(\varphi)} P_\lambda = \text{Id}_V$$

ist, ist nur  $P_\lambda(V) = \hat{E}(\lambda)$  mit der direkten Zerlegung (5.9) verträglich.

• Wegen  $P_\mu \circ P_\lambda = P_\lambda \circ P_\mu$  ist

$$P_\mu P_\lambda(V) = \hat{E}(\lambda) \cap \hat{E}(\mu) = \begin{cases} \hat{E}(\lambda) & , \quad \mu = \lambda \\ \{0\} & , \quad \mu \neq \lambda. \end{cases}$$

Daher ist  $P_\lambda = P_\lambda \text{Id}_V = P_\lambda \left( \sum_{\mu} P_\mu \right) = P_\lambda P_\lambda$ , sodass auch die erste Identität aus (5.16) bewiesen ist.

1. Wegen

$$\text{Im}(P_\lambda) = \hat{E}(\lambda) \quad \text{und} \quad P_\lambda P_\lambda = P_\lambda,$$

ist nach Satz (4.4)  $\hat{E}(\lambda)$  Eigenraum von  $P_\lambda$  zum Eigenwert Eins. Also ist  $\hat{E}(\lambda)$  Eigenraum von  $\varphi_D$  zum Eigenwert  $\lambda$ .

2.  $\varphi_N$  lässt die Zerlegung (5.13) invariant, weil dies auch  $\varphi$  und  $\varphi_D$  tun. Wegen

$$\varphi_N \upharpoonright_{\hat{E}(\lambda)} = (\varphi - \lambda \text{Id}) \upharpoonright_{\hat{E}(\lambda)}$$

ist  $\varphi_N^{f_\lambda} \upharpoonright_{\hat{E}(\lambda)} = 0$ , und der Fittingindex ist die kleinste solche Potenz.

3. folgt, da  $\varphi_D$  und  $\varphi_N$  Polynome in  $\varphi$  sind.
4. Sei  $\varphi = \psi_D + \psi_N$  mit  $\psi_D, \psi_N \in L(V)$  diagonalisierbar, bzw. nilpotent und  $\psi_D \circ \psi_N = \psi_N \circ \psi_D$ . Dann kommutiert  $\psi_D$  auch mit  $\varphi$ , also auch mit dem Polynom  $\varphi_D$  von  $\varphi$ .

Auch  $\varphi_N$  und  $\psi_N$  kommutieren, denn

$$\begin{aligned} \varphi_N \circ \psi_N &= (\varphi - \varphi_D) \circ (\varphi - \psi_D) = \varphi \circ \varphi - \varphi_D \circ \varphi - \varphi \circ \psi_D + \varphi_D \circ \psi_D \\ &= \varphi \circ \varphi - \varphi \circ \varphi_D - \psi_D \circ \varphi + \psi_D \circ \varphi_D = (\varphi - \psi_D) \circ (\varphi - \psi_D) \\ &= \psi_N \circ \varphi_N. \end{aligned}$$

Nach dem folgenden Satz ist

$$\varphi_D - \psi_D = \psi_N - \varphi_N$$

damit diagonalisierbar *und* nilpotent, und damit die Nullabbildung.  $\square$

**5.27 Satz** Für endlich viele Endomorphismen eines endlich-dimensionalen Vektorraumes gilt:

1. Die Summe kommutierender nilpotenter Endomorphismen ist nilpotent.
2. Kommutierende diagonalisierbare Endomorphismen besitzen eine *gemeinsame* Basis aus Eigenvektoren.

**Bew.:**

1. Es seien  $\varphi, \psi \in L(V)$  nilpotent und  $\varphi^n = \psi^n = 0$ . Gilt außerdem  $\varphi \circ \psi = \psi \circ \varphi$ , dann ist

$$(\varphi + \psi)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} \varphi^k \psi^{2n-k} = 0,$$

denn falls  $k \leq n$ , dann ist  $\psi^{2n-k} = 0$ , andernfalls  $\varphi^k = 0$ . Iteration dieses Arguments zeigt die Aussage für endlich viele kommutierende nilpotente Endomorphismen.

2. Es seien  $\varphi, \psi \in L(V)$  diagonalisierbar, dann können wir sie mithilfe der Projektoren  $P_\lambda$  bzw.  $Q_\mu$  auf die Eigenräume von  $\varphi$  bzw.  $\psi$  in der Form

$$\varphi = \sum_{\lambda \in \text{Spek}(\varphi)} \lambda P_\lambda \quad , \quad \psi = \sum_{\mu \in \text{Spek}(\psi)} \mu Q_\mu \quad (5.17)$$

schreiben. Ist außerdem  $[\varphi, \psi] = 0$ , dann ist, da die Projektoren  $P_\lambda$  bzw.  $Q_\mu$  nach Satz 4.6 Funktionen von  $\varphi$  bzw.  $\psi$  sind, auch

$$[P_\lambda, Q_\mu] = 0 \quad (\lambda \in \text{Spek}(\varphi), \mu \in \text{Spek}(\psi)).$$

Daher erfüllen die Operatoren

$$R(\lambda, \mu) := P_\lambda \circ Q_\mu$$

nicht nur die Relation

$$\sum_{\lambda, \mu} R(\lambda, \mu) = \text{Id}, \quad (5.18)$$

sondern sind ebenfalls Projektoren mit

$$R(\lambda, \mu) R(\lambda', \mu') = \delta_{\lambda\lambda'} \delta_{\mu\mu'} R(\lambda, \mu). \quad (5.19)$$

Setzen wir  $E(\lambda, \mu) := R(\lambda, \mu)(V)$ , dann wegen (5.18) und (5.19)

$$V = \bigoplus_{\lambda, \mu} E(\lambda, \mu),$$

und wegen (5.17) sind die Restriktionen von  $\varphi$  bzw.  $\psi$  auf den Unterraum  $E(\lambda, \mu)$  das  $\lambda$ -fache bzw.  $\mu$ -fache der Identität.

Iteration dieses Arguments zeigt die Aussage für endlich viele kommutierende diagonalisierbare Endomorphismen.  $\square$

Wir benutzen jetzt die Abkürzung

$$J_r(\lambda) := \lambda \mathbb{1}_r + N_r.$$

Beispielsweise ist  $J_4(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$ , und der nilpotente Anteil  $J_r(0) = N_r$  von  $J_r(\lambda)$  kommutiert mit dem Diagonalanteil  $\lambda \mathbb{1}_r$ .

**5.28 Definition** Eine Matrix aus  $\mathcal{M}(r, K)$  heißt Jordanmatrix (oder ist in Jordannormalform), wenn sie die Gestalt

$$\bigoplus_{i=1}^j J_{r_i}(\lambda_i)$$

(mit  $\sum_{i=1}^j r_i = r$  und  $\lambda_i \in K$ ) besitzt. Die Matrizen  $J_{r_i}(\lambda_i)$  heißen dann die Jordanblöcke der Matrix.

**5.29 Satz** Das charakteristische Polynom von  $\varphi \in L(V)$  zerfällt in Linearfaktoren. Dann existiert eine Basis  $B$  von  $V$  (genannt **Jordanbasis**), sodass die darstellende Matrix  $\mathcal{M}_B(\varphi)$  eine Jordanmatrix ist.

**Bew.:** Wir gehen von der  $\varphi$ -invarianten Zerlegung (5.9) von  $V$  in die verallgemeinerten Eigenräume  $\hat{E}(\lambda)$  aus, und betrachten die nilpotenten Endomorphismen

$$(\varphi - \lambda \cdot \text{Id})|_{\hat{E}(\lambda)} \quad (\lambda \in \text{Spek}(\varphi))$$

dieser Unterräume. Diese besitzen nach Theorem 5.15 Jordanbasen  $B^{(\lambda)}$ . schreiben wir diese hintereinander, erhalten wir eine Jordanbasis  $B$  von  $\varphi$ .  $\square$

**5.30 Bemerkung** Im allgemeinen gibt es keine kanonische Wahl der Anordnung der Jordanbasen  $B^{(\lambda)}$  (eine Ausnahme ist der Fall  $K = \mathbb{R}$ , wo wir die Eigenwerte  $\lambda \in \text{Spek}(\varphi)$  und entsprechend die  $B^{(\lambda)}$  nach aufsteigender Größe anordnen können).

Wir sprechen daher von *einer*, nicht *der* Jordannormalform von  $\varphi$ .

## 5.5 Ein Beispiel

Wir wollen eine Jordan-Normalform der reellen Matrix

$$M := \begin{pmatrix} 1/4 & 1 & -3/4 & 0 & -3/4 & 1 \\ 1/2 & 4 & 1/2 & 0 & 7/2 & -3 \\ 9/4 & 3 & 13/4 & 0 & 25/4 & -6 \\ 3/4 & 0 & 3/4 & 2 & 7/4 & -1 \\ -1/2 & 0 & -1/2 & 0 & 1/2 & 1 \\ 0 & 2 & 0 & 0 & 2 & 0 \end{pmatrix}$$

bzw. des durch  $\varphi(x) := Mx$  gegebenen Endomorphismus von  $V = \mathbb{R}^6$  bestimmen.

Als ersten Schritt berechnen wir das charakteristische Polynom

$$\chi_M(x) = x^6 - 10x^5 + 41x^4 - 88x^3 + 104x^2 - 64x + 16.$$

Dieses besitzt die Nullstelle 2, was man daraus ablesen kann, dass die vierte Spalte von  $M$  außer der 2 an der Diagonalstelle nur aus Nullen besteht. Fortgesetzte Division durch  $x - 2$  zeigt, dass diese Nullstelle die Ordnung vier besitzt, und dass

$$\chi_M(x) = (x - 2)^4(x - 1)^2$$

ist. Als nächstes bestimmen wir die verallgemeinerten Eigenräume  $\hat{E}(\lambda)$  dieser Nullstellen.

Der vierdimensionale Kern von  $(\varphi - 2\text{Id})^3$  mit darstellender Matrix

$$(M - 2 \cdot \mathbb{1})^3 = \begin{pmatrix} -11/4 & 0 & -7/4 & 0 & -19/4 & 4 \\ 1/2 & 0 & 1/2 & 0 & 3/2 & -1 \\ 13/4 & 0 & 9/4 & 0 & 25/4 & -5 \\ 1/2 & 0 & 1/2 & 0 & 3/2 & -1 \\ -1/2 & 0 & -1/2 & 0 & -3/2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

enthält den Vektor  $e_{3,1}^{(2)} := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$  (wobei in der Notation  $e_{p,j}^{(\lambda)}$  der obere Index  $\lambda$  den Eigenwert bezeichnet, und  $e_{p,j}^{(\lambda)}$  primitiver Vektor von  $\varphi - \lambda\text{Id}$  ist, vergleiche Kapitel 5.2). Der verallgemeinerte Eigenraum  $\hat{E}(2)$  wird daher von den Vektoren

$$b_3 := e_{3,1}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad b_2 := (\varphi - 2 \cdot \text{Id}) \left( e_{3,1}^{(2)} \right) = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix},$$

$$b_1 := (\varphi - 2 \cdot \text{Id})^2 \left( e_{3,1}^{(2)} \right) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

und dem zweiten Eigenvektor  $b_4 := e_{1,1}^{(2)} := \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \\ 1 \end{pmatrix}$  von  $\varphi$  zum Eigenwert 2 aufgespannt.

Analog behandeln wir den zweidimensionalen verallgemeinerten Eigenraum  $\hat{E}(1)$ . Der Kern von  $(\varphi - \text{Id})^2$  mit darstellender Matrix

$$(M - \mathbb{1})^2 = \begin{pmatrix} -1/4 & 2 & -1/4 & 0 & 7/4 & -1 \\ 1/2 & 5 & 1/2 & 0 & 11/2 & -5 \\ 7/4 & 6 & 7/4 & 0 & 31/4 & -8 \\ 1 & 1 & 1 & 1 & 3 & -2 \\ -1/2 & 0 & -1/2 & 0 & -1/2 & -1 \\ 0 & 4 & 0 & 0 & 4 & -3 \end{pmatrix}$$

enthält den Vektor  $b_6 := e_{2,1}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}$ , und  $b_5 := (\varphi - \text{Id}) \left( e_{2,1}^{(1)} \right) = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ .

Schreiben wir diese sechs Vektoren  $b_1, \dots, b_6$  nebeneinander, so erhalten wir die Matrix

$$S^{-1} := \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 & 0 & 1 \\ 0 & 3 & 0 & -2 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 2 & 0 & 1 & 0 & 0 \end{pmatrix}$$

mit Inverser

$$S = \begin{pmatrix} 1/2 & 0 & 1/2 & 1 & 3/2 & -1 \\ 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ -1/2 & 0 & -1/2 & 0 & -1/2 & 1 \\ 7/4 & 0 & 3/2 & 0 & 7/4 & -2 \\ -1/2 & 0 & -1/2 & 0 & -3/2 & 1 \end{pmatrix},$$

sodass die Jordansche Normalform

$$SMS^{-1} = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = J_3(2) \oplus J_1(2) \oplus J_2(1)$$

ist.

## 6 Reellifizierung und Komplexifizierung

### 6.1 Skalarwechsel bei Vektorräumen

Bisher haben wir nur Vektorräume über dem gleichen Skalarkörper zueinander in Beziehung gesetzt. Nun ist aber Voraussetzung für die Bildung der Jordanschen Normalform eines Endomorphismus das Zerfallen seines charakteristischen Polynoms in Linearfaktoren. Dies können wir allgemein nur erwarten, wenn der Körper algebraisch abgeschlossen ist. Andererseits besitzt jeder Körper eine algebraisch abgeschlossene Körpererweiterung, zum Beispiel  $\mathbb{C}$  im Fall von  $\mathbb{R}$ .

Wir werden daher jetzt die Beziehungen von Vektorräumen über einem Körper  $\tilde{K}$  und denen über einem Unterkörper  $K \subset \tilde{K}$  untersuchen. Wir werden uns hauptsächlich das für die Anwendungen wichtigste Beispiel  $\mathbb{R} \subset \mathbb{C}$  anschauen.

Zunächst können wir einen  $\tilde{K}$ -Vektorraum  $V$  in einen  $K$ -Vektorraum  $V_K$  umwandeln, indem wir nur noch mit Elementen des Unterkörpers  $K$  multiplizieren, d.h. die Skalarmultiplikation  $\tilde{K} \times V \rightarrow V$  auf  $K \times V$  restringieren, aber weder den Vektorraum als Menge noch die Addition  $V \times V \rightarrow V$  von Vektoren verändern.

**6.1 Definition** Ist  $V$  ein  $\mathbb{C}$ -Vektorraum, dann heißt der  $\mathbb{R}$ -Vektorraum  $V_{\mathbb{R}}$  die Reellifizierung von  $V$ .

**6.2 Satz** Hat der  $\mathbb{C}$ -Vektorraum  $V$  endliche Dimension, dann gilt

$$\dim_{\mathbb{R}}(V_{\mathbb{R}}) = 2 \dim_{\mathbb{C}}(V).$$

**Bew.:** Ist  $B = (b_1, \dots, b_n)$  eine  $\mathbb{C}$ -Basis von  $V$ , dann ist

$$B_{\mathbb{R}} := (b_1, \dots, b_n, ib_1, \dots, ib_n)$$

eine  $\mathbb{R}$ -Basis von  $V_{\mathbb{R}}$ . Denn

•

$$\begin{aligned} \text{span}_{\mathbb{R}}(B_{\mathbb{R}}) &:= \left\{ \sum_{k=1}^n \lambda_k b_k + \mu_k (ib_k) \mid \lambda_k, \mu_k \in \mathbb{R} \right\} \\ &= \left\{ \sum_{k=1}^n \nu_k b_k \mid \nu_k \in \mathbb{C} \right\} := \text{span}_{\mathbb{C}}(B) = V, \end{aligned}$$

da wir nur  $\nu_k := \lambda_k + i\mu_k$  wählen müssen.

• Ist  $0 = \sum_{k=1}^n \lambda_k b_k + \mu_k (ib_k)$  für  $\lambda_k, \mu_k \in \mathbb{R}$ , dann gilt dies auch für  $\sum_{k=1}^n \nu_k b_k$  mit  $\nu_k = \lambda_k + i\mu_k$ . Da  $B$  eine  $\mathbb{C}$ -Basis von  $V$  ist, muss  $\nu_1 = \dots = \nu_n = 0$  sein, also auch  $\lambda_k = \mu_k = 0$  für alle  $k$ .  $\square$

**6.3 Satz** Ist  $\varphi : V \rightarrow W$  eine  $\mathbb{C}$ -lineare Abbildung der  $\mathbb{C}$ -Vektorräume  $V$  und  $W$ , dann ist  $\varphi : V_{\mathbb{R}} \rightarrow W_{\mathbb{R}}$  eine  $\mathbb{R}$ -lineare Abbildung.

**Bew.:** Dies folgt, da für alle  $k, l \in \mathbb{R}$ ,  $v, w \in V = V_{\mathbb{R}}$  (Gleichheit als Menge!) gilt:  $\varphi(kv + lw) = k\varphi(v) + l\varphi(w)$ .  $\square$

Lassen wir nur noch Elemente von  $\mathbb{R}$  als Skalare zu, dann bleibt die bisher auf  $V$  definierte Multiplikation mit nicht reellen Zahlen auf  $V_{\mathbb{R}}$  immerhin noch ein  $\mathbb{R}$ -linearer Endomorphismus.

Speziell gilt:

**6.4 Satz** 1. Ist  $V_{\mathbb{R}}$  die Reellifizierung des  $\mathbb{C}$ -Vektorraumes  $V$ , dann ist die Abbildung

$$J : V_{\mathbb{R}} \rightarrow V_{\mathbb{R}} \quad , \quad J(v) := iv$$

ein Element von  $L(V_{\mathbb{R}})$  und  $J^2 = -\text{Id}$ .

2. Ist  $W$  ein  $\mathbb{R}$ -Vektorraum, und ist  $J \in L(W)$  eine **komplexe Struktur**, d.h. gilt  $J^2 = -\text{Id}$ , dann ist  $W$  mit gleicher Vektoraddition, aber neuer Skalarmultiplikation

$$\mathbb{C} \times W \rightarrow W \quad , \quad k \cdot w := \text{Re}(k) \cdot w + \text{Im}(k) \cdot J(w)$$

ein  $\mathbb{C}$ -Vektorraum.

**Bew.:**

1. Es gilt  $J(kv + lw) = kJ(v) + lJ(w)$  für  $k, l \in \mathbb{R}$  und  $v, w \in V$ . Außerdem gilt  $J^2(v) = -v$ .

2. Für  $k, l \in \mathbb{C}$  und  $w \in W$  gilt

$$\begin{aligned} k \cdot (l \cdot w) &= k \cdot (\text{Re}(l)w + \text{Im}(l)J(w)) \\ &= (\text{Re}(k)\text{Re}(l) - \text{Im}(k)\text{Im}(l))w + \\ &\quad (\text{Re}(k)\text{Im}(l) + \text{Im}(k)\text{Re}(l))J(w) \\ &= \text{Re}(k \cdot l)w + \text{Im}(k \cdot l)J(w) = (k \cdot l)w. \end{aligned}$$

Die anderen  $\mathbb{C}$ -Vektorraumaxiome überprüft man analog.  $\square$

**6.5 Beispiel** Im einfachsten Fall ist  $V := \mathbb{C}$ , also  $V_{\mathbb{R}} = \mathbb{R}^2$ , und  $J : V_{\mathbb{R}} \rightarrow V_{\mathbb{R}}$  ist eine Drehung um  $\pi/2$ :

$$J(v) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot v.$$

Wir benutzen den zweiten Teil des Satzes 6.4, um aus reellen Vektorräumen komplexe zu konstruieren:

**6.6 Definition** Ist  $V$  ein  $\mathbb{R}$ -Vektorraum, dann heißt der  $\mathbb{C}$ -Vektorraum

$$V_{\mathbb{C}} := V \oplus V$$

mit der komplexen Struktur

$$J(v_1, v_2) := (-v_2, v_1) \quad ((v_1, v_2) \in V \oplus V)$$

die Komplexifizierung von  $V$ .

Wir schreiben kurz  $v_1 + iv_2$  für  $v = (v_1, v_2) \in V_{\mathbb{C}}$ , und nennen  $\operatorname{Re}(v) := v_1$  den Realteil,  $\operatorname{Im}(v) := v_2$  den Imaginärteil von  $v$ .

Da wir aus einer  $\mathbb{R}$ -Basis  $(b_1, \dots, b_n)$  eines endlich-dimensionalen  $\mathbb{R}$ -Vektorraums  $V$  die  $\mathbb{C}$ -Basis  $((b_1, 0), \dots, (b_n, 0))$  von  $V_{\mathbb{C}}$  konstruieren können, gilt

$$\operatorname{Dim}_{\mathbb{C}}(V_{\mathbb{C}}) = \operatorname{Dim}_{\mathbb{R}}(V). \quad (6.1)$$

## 6.7 Bemerkungen

1. Offensichtlich verallgemeinert die Komplexifizierung  $V_{\mathbb{C}}$  von  $V$  die Einführung der komplexen Zahlen  $\mathbb{C} := \mathbb{R} \times \mathbb{R}$  mit Addition  $(v_1, v_2) + (w_1, w_2) := (v_1 + w_1, v_2 + w_2)$  und Multiplikation

$$(k_1, k_2)(v_1, v_2) := (k_1v_1 - k_2v_2, k_1v_2 + k_2v_1).$$

2. Die Reellifizierung ist *nicht* die Umkehrung der Komplexifizierung. Insbesondere gilt nach Satz 6.2 und (6.1):  $\operatorname{Dim}((V_{\mathbb{C}})_{\mathbb{R}}) = 2\operatorname{Dim}(V)$  und  $\operatorname{Dim}((W_{\mathbb{R}})_{\mathbb{C}}) = 2\operatorname{Dim}(W)$ .
3. Ist  $V$  ein  $K$ -Vektorraum und  $\tilde{K} \supset K$  ein Erweiterungskörper, dann können wir analog die Skalarerweiterung  $V_{\tilde{K}}$  von  $V_K$  einführen, indem wir zunächst den Körper  $\tilde{K}$  als  $K$ -Vektorraum auffassen, in diesem eine Basis  $(b_j)_{j \in J}$  von  $\tilde{K}$  wählen, und

$$V_{\tilde{K}} := \bigoplus_{j \in J} \{b_j\} \times V$$

setzen. Die Skalarmultiplikation erklärt man dann analog zum Fall  $K = \mathbb{R}$  und  $\tilde{K} = \mathbb{C}$ , also  $V_{\mathbb{C}} = (\{1\} \times V) \oplus (\{i\} \times V)$ .

Die Komplexkonjugation  $z \equiv \operatorname{Re}(z) + i \operatorname{Im}(z) \mapsto \bar{z} = \operatorname{Re}(z) - i \operatorname{Im}(z)$  ist ein Körperautomorphismus von  $\mathbb{C}$ . Diese Abbildung verallgemeinert sich folgendermaßen:

**6.8 Definition** Ist  $V_{\mathbb{C}}$  die Komplexifizierung des  $\mathbb{R}$ -Vektorraumes  $V$ , dann heißt  $\bar{v} := (v_1, -v_2) \equiv v_1 - iv_2 \in V_{\mathbb{C}}$  der zu  $v = (v_1, v_2) \equiv v_1 + iv_2 \in V_{\mathbb{C}}$  komplex konjugierte Vektor und die Abbildung  $v \mapsto \bar{v}$  Konjugation.

Die Rechenregeln der Konjugation auf  $\mathbb{C}$  übertragen sich auf  $V_{\mathbb{C}}$ :

$$\bar{\bar{v}} = v \quad , \quad \overline{v+w} = \bar{v} + \bar{w} \quad , \quad \overline{kv} = \bar{k}\bar{v} \quad (v, w \in V_{\mathbb{C}}, k \in \mathbb{C}).$$

Die reellen Zahlen zeichnen sich gegenüber den anderen komplexen Zahlen dadurch aus, dass sie sich bei der Konjugation nicht ändern. Es liegt daher nahe, den reellen Vektorraum  $V$  analog in  $V_{\mathbb{C}}$  einzubetten, indem man

$$I : V \rightarrow V_{\mathbb{C}} \quad , \quad v \mapsto (v, 0) = v$$

setzt. Es gilt dann  $I(V) = \{v \in V_{\mathbb{C}} \mid \bar{v} = v\}$ . Wir schreiben daher einfach  $V$  statt  $I(V)$ .

Auf eine Teilmenge  $M \subset V_{\mathbb{C}}$  wenden wir die Konjugation an, indem wir  $\bar{M} := \{\bar{v} \mid v \in M\}$  setzen.

**6.9 Lemma** 1. Ist  $U \subset V$  ein Unterraum des  $\mathbb{R}$ -Vektorraumes  $V$ , dann ist  $U_{\mathbb{C}} \subset V_{\mathbb{C}}$  ein komplexer Unterraum von  $V_{\mathbb{C}}$  mit

$$\overline{U_{\mathbb{C}}} = U_{\mathbb{C}}.$$

2. Ist  $W \subset V_{\mathbb{C}}$  ein komplexer Unterraum von  $V_{\mathbb{C}}$  mit

$$\overline{W} = W,$$

dann existiert ein eindeutiger reeller Unterraum  $U$  von  $V$  mit

$$W = U_{\mathbb{C}} \quad , \quad \text{und zwar} \quad U := V \cap W.$$

**Bew.:**

1.  $U_{\mathbb{C}} = \{u_1 + iu_2 \mid u_1, u_2 \in U\} = \{u_1 - iu_2 \mid u_1, u_2 \in U\} = \overline{U_{\mathbb{C}}}$ . Ebenso ist  $U_{\mathbb{C}}$  abgeschlossen bezüglich der Multiplikation mit komplexen Zahlen.
2. Für beliebige  $\mathbb{C}$ -Unterräume  $W \subset V_{\mathbb{C}}$  ist  $U := V \cap W$  ein  $\mathbb{R}$ -Unterraum von  $V$ , aber wegen Teil 1. des Lemmas kann höchstens dann  $U_{\mathbb{C}} = W$  gelten, wenn  $W = \overline{W}$ .

Ist aber  $W = \overline{W}$ , dann sind für jeden Vektor  $w \in W$  auch *Realteil*  $\operatorname{Re}(w) := \frac{1}{2}(w + \overline{w})$  und sein *Imaginärteil*  $\operatorname{Im}(w) := -\frac{i}{2}(w - \overline{w})$  Elemente von  $W$ .

Andererseits sind Real- und Imaginärteil beliebiger Vektoren  $w \in V_{\mathbb{C}}$  reell (d.h.  $\overline{\operatorname{Re}(w)} = \operatorname{Re}(w) \in V$  und  $\overline{\operatorname{Im}(w)} = \operatorname{Im}(w) \in V$ ), sodass

$$w = \operatorname{Re}(w) + i \operatorname{Im}(w) \quad \text{mit} \quad \operatorname{Re}(w), \operatorname{Im}(w) \in U.$$

□

Die meisten komplexen Unterräume  $W \subset V_{\mathbb{C}}$  sind aber nicht Komplexifizierungen von Unterräumen  $U \subset V$  (wenn man von den Trivialfällen  $\operatorname{Dim}(V) \leq 1$  absieht).

**6.10 Beispiel**  $V := \mathbb{R}^2$ , also  $V_{\mathbb{C}} = \mathbb{C}^2$ . Der Unterraum  $W := \{(\lambda, i\lambda) \mid \lambda \in \mathbb{C}\}$  von  $V_{\mathbb{C}}$  hat die Dimension  $\operatorname{Dim}_{\mathbb{C}}(W) = 1$ , denn  $W = \operatorname{span}_{\mathbb{C}}((1, i))$ .

Andererseits ist  $W \cap V = \{0\}$ , also ist  $W$  nicht Komplexifizierung eines Unterraumes von  $V$ .

Die Komplexifizierung eines reellen Vektorraumes ist nützlich, weil  $\mathbb{C}$  der algebraische Abschluss des Körpers  $\mathbb{R}$  der reellen Zahlen ist, also insbesondere jedes komplexe Polynom in Linearfaktoren zerfällt.

Wir wenden dies auf das charakteristische Polynom eines reellen Endomorphismus an, um dessen Jordansche Normalform zu verstehen.

## 6.2 Die reelle Jordansche Normalform

Etwas allgemeiner definieren wir

**6.11 Definition** Die Komplexifizierung einer  $\mathbb{R}$ -linearen Abbildung  $\varphi : V \rightarrow W$  ist die  $\mathbb{C}$ -lineare Abbildung

$$\varphi_{\mathbb{C}} : V_{\mathbb{C}} \rightarrow W_{\mathbb{C}} \quad , \quad \varphi_{\mathbb{C}}(v_1, v_2) := (\varphi(v_1), \varphi(v_2)) \quad ((v_1, v_2) \in V_{\mathbb{C}}).$$

In unserer Schreibweise  $v_1 + iv_2$  für  $(v_1, v_2)$  ist  $\varphi_{\mathbb{C}}$  gerade durch

$$\varphi_{\mathbb{C}}(v_1 + iv_2) := \varphi(v_1) + i\varphi(v_2)$$

definiert, also durch die Eigenschaft, dass die imaginäre Einheit  $i$  (und damit eine beliebige komplexe Zahl) nach Belieben vor oder nach Ausführung der Abbildung heranzumultiplizieren kann.

Ist daher  $M := \mathcal{M}_{\mathbb{C}}^B(\varphi) \in \mathcal{M}(n \times m, \mathbb{R})$  darstellende Matrix der reell-linearen Abbildung  $\varphi$ , dann ist dieselbe Matrix  $M$  auch darstellende Matrix von

$\varphi_{\mathbb{C}}$ , nur dass sie diesmal als Element von  $\mathcal{M}(n \times m, \mathbb{C}) \supset \mathcal{M}(n \times m, \mathbb{R})$  aufgefasst wird.

Um nun die Jordansche Normalform eines reellen Endomorphismus  $\varphi$  zu finden, verallgemeinern wir das Beispiel 5.5.3.

Dort war die Matrix  $A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \in \mathcal{M}(2, \mathbb{R})$  (die für  $\varphi \notin \pi\mathbb{Z}$  zu keiner reellen Diagonalmatrix konjugiert ist) als Element von  $\mathcal{M}(2, \mathbb{C})$  aufgefasst worden, und es galt

$$S^{-1}AS = \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \quad \text{mit} \quad \lambda = \cos \varphi + i \sin \varphi \quad \text{und} \quad S = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \in GL(2, \mathbb{C}), \quad (6.2)$$

$A$  war also komplex diagonalisierbar.

**6.12 Definition** Sind  $V$  und  $W$   $\mathbb{R}$ -Vektorräume und  $\varphi : V_{\mathbb{C}} \rightarrow W_{\mathbb{C}}$  eine (nicht notwendig lineare) Abbildung, dann heißt

$$\overline{\varphi} : V_{\mathbb{C}} \rightarrow W_{\mathbb{C}} \quad , \quad \overline{\varphi}(v) := \overline{\varphi(\overline{v})}$$

die zu  $\varphi$  konjugiert komplexe Abbildung.

Für beliebige Abbildungen  $\varphi, \varphi_1, \varphi_2 : V_{\mathbb{C}} \rightarrow W_{\mathbb{C}}$  und  $\psi : U_{\mathbb{C}} \rightarrow V_{\mathbb{C}}$  gilt

$$\overline{\overline{\varphi}} = \varphi \quad \text{und} \quad \overline{\varphi \circ \psi} = \overline{\varphi} \circ \overline{\psi},$$

$$\overline{\varphi_1 + \varphi_2} = \overline{\varphi_1} + \overline{\varphi_2} \quad , \quad \overline{k\varphi} = \overline{k}\overline{\varphi} \quad (k \in \mathbb{C}).$$

während für  $\mathbb{C}$ -lineare Abbildungen die Konjugation wieder eine lineare Abbildung liefert:  $\overline{\varphi}(kv + lw) = k\overline{\varphi}(v) + l\overline{\varphi}(w)$ .

Auch die durch  $\varphi \in L(V_{\mathbb{C}}, W_{\mathbb{C}})$  definierten Unterräume Kern und Bild verhalten sich vernünftig unter Konjugation:

$$\overline{\text{Kern}(\varphi)} = \text{Kern}(\overline{\varphi}) \quad , \quad \overline{\text{Im}(\varphi)} = \text{Im}(\overline{\varphi}).$$

All diese Identitäten ergeben sich unmittelbar aus der Definition der Konjugation.

**6.13 Lemma**  $\varphi \in L(V_{\mathbb{C}}, W_{\mathbb{C}})$  ist genau dann gleich der Komplexifizierung einer reellen linearen Abbildung, wenn  $\varphi = \overline{\varphi}$  gilt. In diesem Fall ist  $\varphi = \psi_{\mathbb{C}}$  für  $\psi := \varphi|_V$ .

**Bew.:**

- Ist  $\varphi = \psi_{\mathbb{C}}$  mit  $\psi \in L(V, W)$ , dann ist für  $v_1, v_2 \in V$

$$\overline{\varphi}(v_1 + iv_2) = \overline{\varphi(v_1 - iv_2)} = \overline{\psi(v_1) - i\psi(v_2)} = \psi(v_1) + i\psi(v_2) = \varphi(v_1 + iv_2).$$

- Ist  $\varphi = \overline{\varphi} \in L(V_{\mathbb{C}}, W_{\mathbb{C}})$ , dann ist für  $v \in V$ , also  $\overline{v} = v$ ,

$$\varphi(v) = \varphi(\overline{v}) = \overline{\overline{\varphi}(v)} = \overline{\varphi(v)},$$

sodass  $\psi \in L(V, W)$  für  $\psi := \varphi|_V$  gilt.  $\square$

Es sei nun  $V$  ein  $\mathbb{R}$ -Vektorraum der Dimension  $n < \infty$ . Dann ist das charakteristische Polynom von  $\varphi \in L(V)$  ein normiertes reelles Polynom  $n$ -ten Grades.

**6.14 Satz** Ist  $p \in \mathbb{R}[x]$  ein von Null verschiedenes Polynom, dann existieren  $k, l \in \mathbb{N}_0$  mit  $k + 2l = \deg(p)$  und  $a_1, \dots, a_k, b_1, \dots, b_l, c_1, \dots, c_l, d \in \mathbb{R}$  mit  $4c_j > b_j^2$ , sodass

$$p(x) = d \prod_{j=1}^k (x - a_j) \cdot \prod_{j=1}^l (x^2 + b_j x + c_j).$$

**Bew.:** Aufgefasst als komplexes Polynom, zerfällt  $p$  in ein Produkt von Linearfaktoren:  $p(x) = d \cdot \prod_{i=1}^n (x - \lambda_i)$  mit  $\lambda_i \in \mathbb{C}$ . Wegen  $\overline{p} = p$  ist  $d$  reell und mit  $\lambda_i$  auch  $\overline{\lambda_i}$  Nullstelle.

Die reellen Nullstellen  $\lambda_i = \overline{\lambda_i}$  bezeichnen wir mit  $a_j$ .

Ist  $\text{Im}(\lambda_i) > 0$ , dann setzen wir  $b_j := -2\text{Re}(\lambda_i)$  und  $c_j := |\lambda_i|^2$ , sodass  $(x - \lambda_i)(x - \overline{\lambda_i}) = x^2 + b_j x + c_j^2$  gilt.  $\square$

Das charakteristische Polynom von  $\varphi_{\mathbb{C}}$  ist gleich dem von  $\varphi$ , denn eine Basis  $B$  von  $V$  ist gleichzeitig eine Basis von  $V_{\mathbb{C}}$ , und die darstellenden Matrizen stimmen überein:

$$\mathcal{M}_B(\varphi_{\mathbb{C}}) = \mathcal{M}_B(\varphi)$$

Wir benutzen nun eine komplexe Jordanbasis von  $\varphi_{\mathbb{C}} \in L(V_{\mathbb{C}})$ , um möglichst viel über den Endomorphismus  $\varphi \in L(V)$  eines  $r$ -dimensionalen  $\mathbb{R}$ -Vektorraums zu erfahren.

Wir unterscheiden zwischen reellen und nicht reellen Eigenwerten  $\lambda \in \text{spek}(\varphi_{\mathbb{C}})$ .

- Ist  $\lambda \in \mathbb{R}$ , dann ist der verallgemeinerte Eigenraum  $\hat{E}_{\varphi_{\mathbb{C}}}(\lambda) = \text{Kern}((\varphi_{\mathbb{C}} - \lambda \text{Id})^r)$  invariant unter Komplex-Konjugation, denn  $\overline{\varphi_{\mathbb{C}} - \lambda \text{Id}} = \varphi_{\mathbb{C}} - \lambda \text{Id}$ . Nach Lemma 6.9 ist daher  $\hat{E}_{\varphi_{\mathbb{C}}}(\lambda)$  Komplexifizierung des reellen Unterraums  $\hat{E}_{\varphi_{\mathbb{C}}}(\lambda) \cap V$  von  $V$ , und auf diesem Unterraum finden wir eine reelle Jordanbasis für die Einschränkung von  $\varphi$ .
- Ist dagegen  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ , dann ist  $\overline{\lambda} \neq \lambda$  und auch  $\overline{\lambda} \in \text{spek}(\varphi_{\mathbb{C}})$ . In einer Jordannormalform von  $\varphi$  gibt es genauso viele Jordanblöcke  $J_{r_i}(\lambda)$  wie  $J_{r_i}(\overline{\lambda})$ , denn wegen deren Eindeutigkeit bis auf Permutation der Blöcke und der aus Lemma 6.13 folgenden Invarianz  $\overline{\varphi_{\mathbb{C}}} = \varphi_{\mathbb{C}}$  unter Komplexkonjugation treten diese paarweise auf.

Besitzt nun (z.B. für  $\text{Im}(\lambda) > 0$ ) ein Jordanblock  $J_k(\lambda)$  in  $\hat{E}_{\varphi_{\mathbb{C}}}(\lambda)$  die Jordanbasis  $(c_1, \dots, c_k)$ , dann benutzen wir die konjugierte Jordanbasis  $(\bar{c}_1, \dots, \bar{c}_k)$  für den Jordanblock  $J_k(\bar{\lambda})$  in  $\hat{E}_{\varphi_{\mathbb{C}}}(\bar{\lambda}) = \hat{E}_{\varphi_{\mathbb{C}}}(\lambda)$ .

Von diesen ( $\mathbb{C}$ -linear unabhängigen!) Basen ausgehend, konstruieren wir die  $\mathbb{R}$ -Basis

$$(a_1, b_1, a_2, b_2, \dots, a_k, b_k) \quad \text{mit} \quad a_j := \text{Re}(c_j) \quad , \quad b_j := \text{Im}(c_j). \quad (6.3)$$

Die Jordan-Beziehung

$$\varphi_{\mathbb{C}}(c_1) = \lambda c_1 \quad , \quad \varphi_{\mathbb{C}}(c_j) = \lambda c_j + c_{j-1} \quad (j = 2, \dots, k)$$

übersetzt sich dabei in

$$\varphi(a_1) = \text{Re}(\lambda)a_1 + \text{Im}(\lambda)b_1 \quad , \quad \varphi(a_j) = \text{Re}(\lambda)a_j + \text{Im}(\lambda)b_j + a_{j-1} \quad (j = 2, \dots, k)$$

und

$$\varphi(b_1) = -\text{Im}(\lambda)b_1 + \text{Re}(\lambda)a_1 \quad , \quad \varphi(b_j) = -\text{Im}(\lambda)b_j + \text{Re}(\lambda)a_j + b_{j-1} \quad (j = 2, \dots, k).$$

Im einfachsten Fall  $k = 1$  ist die darstellende Matrix der Restriktion von  $\varphi$  bezüglich der Basis (6.3) gleich  $\begin{pmatrix} \text{Re}(\lambda) & \text{Im}(\lambda) \\ -\text{Im}(\lambda) & \text{Re}(\lambda) \end{pmatrix}$ , den wir von den Drehungen im  $\mathbb{R}^2$  kennen (dort ist  $|\lambda| = 1$ ).

Für  $k = 2$  ist die darstellende Matrix  $\begin{pmatrix} \text{Re}(\lambda) & \text{Im}(\lambda) & 1 & 0 \\ -\text{Im}(\lambda) & \text{Re}(\lambda) & 0 & 1 \\ 0 & 0 & \text{Re}(\lambda) & \text{Im}(\lambda) \\ 0 & 0 & -\text{Im}(\lambda) & \text{Re}(\lambda) \end{pmatrix}$ .

## 7 Lineare Differentialgleichungen

Differentialgleichungen (DGLn) modellieren Naturvorgänge. Ein Standardbeispiel ist die Newtonsche Kraftgleichung

$$F(q, \dot{q}) = m\ddot{q}, \quad (7.1)$$

wobei die Kraft  $F$  von Ort  $q \in \mathbb{R}^d$  und Geschwindigkeit  $\dot{q} \in \mathbb{R}^d$  abhängt, und gleich Masse  $m > 0$  mal Beschleunigung  $\ddot{q}$  ist. Gesucht sind Lösungen, d.h. im idealen Fall glatte Funktionen  $q: \mathbb{R} \rightarrow \mathbb{R}^d$ , die diese Gleichung erfüllen.

**7.1 Beispiel**  $F(q, \dot{q}) = 0$ ;  $q^{(0)} \in \mathbb{R}^d$ ,  $v^{(0)} \in \mathbb{R}^d$  beliebig.  $q(t) := q^{(0)} + v^{(0)}t$  erfüllt die Newtonsche Kraftgleichung (freie Bewegung mit Startpunkt  $q^{(0)}$  und Geschwindigkeit  $v^{(0)}$  zum Zeitpunkt  $t = 0$ ).

In der Linearen Algebra interessieren wir uns für *lineare* (homogene, zeitunabhängige) Differentialgleichungen, die in der Form

$$\dot{x} = Ax \quad \text{mit} \quad A \in \mathcal{M}(n, \mathbb{R})$$

geschrieben werden können. Gesucht ist die *Lösung des Anfangswertproblems* für  $x^{(0)} \in \mathbb{R}^d$ , d.h. eine differenzierbare Funktion

$$x : \mathbb{R} \rightarrow \mathbb{R}^n \quad \text{mit} \quad \frac{d}{dt}x(t) = Ax(t) \quad (t \in \mathbb{R}) \quad \text{und} \quad x(0) = x^{(0)}.$$

Die *Ordnung* einer DGL ist die höchste vorkommende Ableitungsstufe. Während die Newtonsche DGL (7.1) von 2. Ordnung ist, betrachten wir also hier DGLn 1. Ordnung.

**7.2 Beispiel** Im Fall der freien Bewegung setzen wir  $n := 2d$  und schreiben  $x = \begin{pmatrix} q \\ v \end{pmatrix}$  mit  $q, v \in \mathbb{R}^d$  (Ort bzw. Geschwindigkeit). Dann gilt für  $A := \begin{pmatrix} 0 & \mathbb{1} \\ 0 & 0 \end{pmatrix} \in \mathcal{M}(n, \mathbb{R})$

$$\dot{x} = Ax,$$

und die (eindeutige) Lösung des Anfangswertproblems mit  $x^{(0)} = \begin{pmatrix} q^{(0)} \\ v^{(0)} \end{pmatrix}$  ist

$$x(t) = \begin{pmatrix} q^{(0)} + v^{(0)}t \\ v^{(0)} \end{pmatrix} \quad (t \in \mathbb{R}).$$

## 7.1 Berechnung der Lösung

Im Gegensatz zu allgemeinen Differentialgleichungen können wir die linearen DGLn mit algebraischen Mitteln lösen.

Wir definieren dazu die Anwendung der Exponentialfunktion auf eine Matrix  $A \in \mathcal{M}(n, \mathbb{C})$  durch

$$e^A := \exp(A) := \lim_{k \rightarrow \infty} \sum_{l=0}^k \frac{A^l}{l!},$$

also durch die Einsetzung von  $A$  in die Potenzreihe der Exponentialfunktion.

Dabei ist der Limes-Begriff zu klären. Wir benutzen auf dem  $\mathbb{R}^n$  die Euklidische Norm  $\|v\| := \sqrt{|v_1|^2 + \dots + |v_n|^2}$  und auf  $\mathcal{M}(n, \mathbb{C})$  die *Operatornorm*

$$\|A\| := \max_{v \in \mathbb{R}^n \setminus \{0\}} \frac{\|Av\|}{\|v\|} \quad (A \in \mathcal{M}(n, \mathbb{C})).$$

$\|A\|$  ist also der maximale Streckungsfaktor der linearen Abbildung  $v \mapsto Av$ .<sup>3</sup>

Es gilt nun (ohne Beweis):

---

<sup>3</sup>Da für  $A \in \mathcal{M}(n, \mathbb{R})$   $\|A\| = \max_{v \in \mathbb{R}^n, \|v\|=1} \|Av\|$  ist, und die Sphäre  $S^{n-1} := \{v \in \mathbb{R}^n \mid \|v\| = 1\}$  kompakt ist, wird das Supremum angenommen. Eine analoge Aussage gilt für  $A \in \mathcal{M}(n, \mathbb{C})$ .

**7.3 Satz** 1. Die Reihe  $e^A$  ist für alle  $a > 0$  und  $A$  aus

$$\{A \in \mathcal{M}(n, \mathbb{C}) \mid \|A\| \leq a\}$$

gleichmäßig konvergent, und es gilt

$$\exp(t_1 A) \exp(t_2 A) = \exp((t_1 + t_2)A) \quad (t_1, t_2 \in \mathbb{R}).$$

2.  $t \mapsto x(t) := \exp(At)x_0$  ist die eindeutige Lösung des Anfangswertproblems  $\dot{x} = Ax$ ,  $x(0) = x^{(0)}$ .

Wir können insbesondere die Reihe  $\exp(At) = \sum_{l=0}^{\infty} \frac{(At)^l}{l!}$  gliedweise differenzieren und erhalten

$$\frac{d}{dt} \exp(At) = A \exp(At) \quad (t \in \mathbb{R}).$$

**7.4 Beispiel** 1.  $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Dann ist  $A^l = \text{diag}(\lambda_1^l, \dots, \lambda_n^l)$ , also

$$\exp(At) = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \quad \text{und} \quad x(t) = \begin{pmatrix} e^{\lambda_1 t} x_1^{(0)} \\ \vdots \\ e^{\lambda_n t} x_n^{(0)} \end{pmatrix}.$$

2.  $A = N_n$  (nilpotente Blockmatrix). Dann ist  $\exp(At) = \sum_{l=0}^n \frac{(At)^l}{l!}$ , also z.B. für  $n = 4$

$$\exp(N_4 t) = \begin{pmatrix} 1 & t & t^2/2 & t^3/6 \\ 0 & 1 & t & t^2/2 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

3.  $A = A_1 \oplus A_2$ . Dann ist  $A^l = A_1^l \oplus A_2^l$ , also  $\exp(At) = \exp(A_1 t) \oplus \exp(A_2 t)$ .

Diese Beispiele ermöglichen zusammengenommen die Exponentialfunktion einer Jordanmatrix

$$J_{r_1}(\lambda_1) \oplus \dots \oplus J_{r_k}(\lambda_k) \in \mathcal{M}(r, \mathbb{C})$$

zu bilden, denn  $J_r(\lambda) = \lambda \mathbb{1}_r + N_r$ , also

$$\begin{aligned} \exp(J_r(\lambda)t) &= \exp(\lambda t \mathbb{1} + N_r t) \\ &= \exp(\lambda t \mathbb{1}) \cdot \exp(N_r t) = e^{\lambda t} \exp(N_r t). \end{aligned}$$

Die Exponentialfunktion einer beliebigen reellen Matrix  $A \in \mathcal{M}(n, \mathbb{R})$  bilden wir, indem wir sie mittels  $S \in GL(n, \mathbb{C})$  zu einer komplexen Jordanmatrix

$$J = S^{-1}AS$$

konjugieren. Dann ist

$$\exp(At) = \exp(SJS^{-1}t) = S \exp(Jt)S^{-1},$$

denn bei der Bildung der Potenzen

$$(SJS^{-1})^l = SJS^{-1} \cdot \dots \cdot SJS^{-1}$$

fallen die inneren Paare  $S^{-1}S = \mathbb{1}_n$  heraus.

## 7.2 Ein Beispiel

Es soll die lineare Differentialgleichung

$$\dot{q}_1 = v_1 \quad , \quad \dot{q}_2 = v_2 \quad , \quad \dot{v}_1 = q_2 - q_1 \quad , \quad \dot{v}_2 = q_1 - q_2$$

gelöst werden. Diese beschreibt die Bewegung zweier Massenpunkte auf der Geraden mit Orten  $q_i$  und Geschwindigkeiten  $v_i$ , die sich gegenseitig mit einer zu ihrem Abstand proportionalen Kraft anziehen.

Schreiben wir  $(x_1, x_2, x_3, x_4) = (v_1, v_2, q_1, q_2)$ , dann entspricht diese DGL 2. Ordnung  $\dot{x} = Mx$  mit

$$M := \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Wir stellen die Matrix  $m$  des Dgl-Systems auf.

Die Koordinaten sind die Geschwindigkeiten und Orte  $(v_1, v_2, q_1, q_2)$  der beiden Massenpunkte.

$m = \{\{0, 0, -1, 1\}, \{0, 0, 1, -1\}, \{1, 0, 0, 0\}, \{0, 1, 0, 0\}\};$  **MatrixForm**[ $m$ ]

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

Wir berechnen die Eigenwerte und Eigenvektoren von  $m$ .

Die erste Klammer von **Eigensystem**[ $m$ ] enthält die Nullstellen des charakteristischen Polynoms von  $m$  mit ihren Multiplizitäten.

Die zweite Klammer enthält die dazugehörigen Eigenvektoren.

**Eigensystem**[ $m$ ]

$\{\{0, 0, -i\sqrt{2}, i\sqrt{2}\},$   
 $\{0, 0, 1, 1\}, \{0, 0, 0, 0\}, \{i\sqrt{2}, -i\sqrt{2}, -1, 1\}, \{-i\sqrt{2}, i\sqrt{2}, -1, 1\}\}$

Der erste Eigenvektor  $\{0,0,1,1\}$  zum Eigenwert 0 entspricht einer Verschiebung der beiden Orte der Massenpunkte um den gleichen Wert.

Da  $m$  auf dem verallgemeinerten Eigenraum  $\hat{E}(0)$  nicht diagonalisierbar ist, wird kein zweiter Eigenvektor zum Eigenwert 0 gefunden.

Daher berechnen wir die (Eigenwerte und) Eigenvektoren von  $m^2$ :

**Eigensystem**[ $m.m$ ]

$\{\{-2, -2, 0, 0\}, \{0, 0, -1, 1\}, \{-1, 1, 0, 0\}, \{0, 0, 1, 1\}, \{1, 1, 0, 0\}\}$

Dabei ändert sich die Reihenfolge der (quadrierten) Eigenwerte.

Der zweite Eigenvektor  $\{1,1,0,0\}$  zum Eigenwert 0 entspricht einer Bewegung beider Massen mit gleicher Anfangsgeschwindigkeit vom Nullpunkt weg.

Die Linearkombination  $\{0,0,-1,1\}+\{-1,1,0,0\}=\{-1,1,-1,1\}$  der beiden gefundenen Eigenvektoren zum Eigenwert  $-2$  entspricht Anfangsbedingungen,

bei denen sich die beiden Massenpunkte von ihrem Schwerpunkt bei 0 entfernen,

während sie sich bei der Linearkombination  $\{0,0,-1,1\}-\{-1,1,0,0\}=\{1,-1,-1,1\}$  diesem annähern.

Wir wenden die Exponentialfunktion auf die Matrix  $m$  an:

**em = FullSimplify**[**MatrixExp**[ $m t$ ]]; **MatrixForm**[**em**]

$$\begin{pmatrix} \cos\left[\frac{t}{\sqrt{2}}\right]^2 & \sin\left[\frac{t}{\sqrt{2}}\right]^2 & -\frac{\sin[\sqrt{2} t]}{\sqrt{2}} & \frac{\sin[\sqrt{2} t]}{\sqrt{2}} \\ \sin\left[\frac{t}{\sqrt{2}}\right]^2 & \cos\left[\frac{t}{\sqrt{2}}\right]^2 & \frac{\sin[\sqrt{2} t]}{\sqrt{2}} & -\frac{\sin[\sqrt{2} t]}{\sqrt{2}} \\ \frac{1}{4} (2t + \sqrt{2} \sin[\sqrt{2} t]) & \frac{1}{4} (2t - \sqrt{2} \sin[\sqrt{2} t]) & \cos\left[\frac{t}{\sqrt{2}}\right]^2 & \sin\left[\frac{t}{\sqrt{2}}\right]^2 \\ \frac{1}{4} (2t - \sqrt{2} \sin[\sqrt{2} t]) & \frac{1}{4} (2t + \sqrt{2} \sin[\sqrt{2} t]) & \sin\left[\frac{t}{\sqrt{2}}\right]^2 & \cos\left[\frac{t}{\sqrt{2}}\right]^2 \end{pmatrix}$$

Nun wollen wir die allgemeine Lösung des Differentialgleichungssystems bestimmen. Diese hängt von den Konstanten  $C[1], C[2], C[3]$  und  $C[4]$  ab.

$$\text{FullSimplify[DSolve[{q1''[t] == q2[t] - q1[t], q2''[t] == q1[t] - q2[t]}, {q1[t], q2[t]}, t]]$$

$$\left\{ \left\{ q1[t] \rightarrow \frac{1}{4} \left( 2 (t (C[1] + C[2]) + C[3] + C[4]) + 2 (C[3] - C[4]) \cos[\sqrt{2} t] + \sqrt{2} (C[1] - C[2]) \sin[\sqrt{2} t] \right), \right. \right.$$

$$\left. \left. q2[t] \rightarrow \frac{1}{4} \left( 2 (t (C[1] + C[2]) + C[3] + C[4]) + 2 (-C[3] + C[4]) \cos[\sqrt{2} t] + \sqrt{2} (-C[1] + C[2]) \sin[\sqrt{2} t] \right) \right\} \right\}$$

Jetzt geben wir Anfangsorte und Anfangsgeschwindigkeiten an und lassen das Dgl-System lösen:

$$\text{DSolve[{q1''[t] == q2[t] - q1[t], q2''[t] == q1[t] - q2[t],$$

$$q1[0] == 0, q2[0] == 1, q1'[0] == 1/10, q2'[0] == 0}, {q1[t], q2[t]}, t]$$

$$\left\{ \left\{ q1[t] \rightarrow \frac{1}{80} e^{-i\sqrt{2} t} (-20 + i\sqrt{2} + 40 e^{i\sqrt{2} t} - 20 e^{2i\sqrt{2} t} - i\sqrt{2} e^{2i\sqrt{2} t} + 4 e^{i\sqrt{2} t} t), \right. \right.$$

$$\left. \left. q2[t] \rightarrow \frac{1}{80} e^{-i\sqrt{2} t} (20 - i\sqrt{2} + 40 e^{i\sqrt{2} t} + 20 e^{2i\sqrt{2} t} + i\sqrt{2} e^{2i\sqrt{2} t} + 4 e^{i\sqrt{2} t} t) \right\} \right\}$$

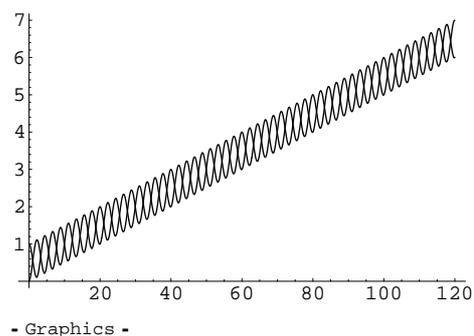
Wir haben vergessen, die Trigonometrischen Additionstheoreme anzuwenden, um die Ausdrücke zu vereinfachen:

$$\mathbf{x} = \text{FullSimplify[Evaluate[{q1[t], q2[t]} /. \%]]$$

$$\left\{ \left\{ \frac{1}{40} (20 + 2 t - 20 \cos[\sqrt{2} t] + \sqrt{2} \sin[\sqrt{2} t]), \frac{1}{40} (20 + 2 t + 20 \cos[\sqrt{2} t] - \sqrt{2} \sin[\sqrt{2} t]) \right\} \right\}$$

$q_1$  und  $q_2$  werden bis zur Zeit  $t_{\max}$  gezeichnet.

`Plot[Evaluate[x], {t, 0, tmax}, PlotPoints -> 300]`



Warum bemühen wir eigentlich die Lineare Algebra, statt die Differentialgleichungen gleich numerisch zu lösen? Weil wir dann auf einen Schlag die Lösung für alle Zeiten finden. Numerik aber verschlechtert sich mit der Zeit. Hier explodiert aus mir unbekanntem Gründen die numerische Lösung zur Zeit  $t=109$ :

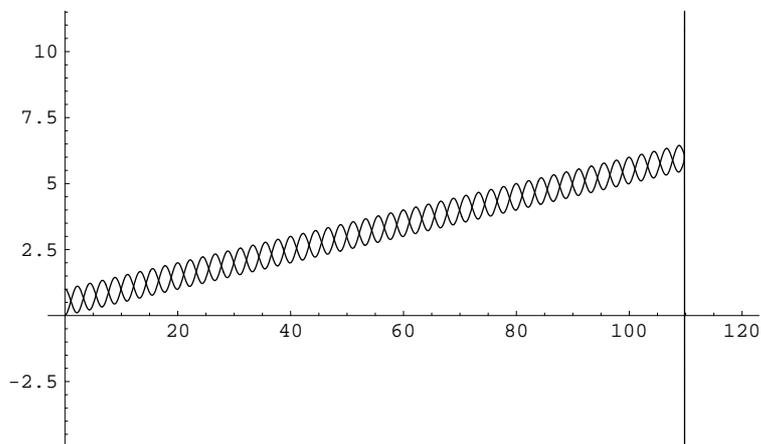
```
tmax = 120;
NDSolve[{Q1''[t] == Q2[t] - Q1[t], Q2''[t] == Q1[t] - Q2[t],
  Q1[0] == 0, Q2[0] == 1, Q1'[0] == 0.1, Q2'[0] == 0}, {Q1, Q2}, {t, 0, tmax}]

NDSolve::mxst : Maximum number of 1000 steps reached at the point t == 109.68734614586384`.

{{Q1 -> InterpolatingFunction[{{0., 109.687}}, <>],
  Q2 -> InterpolatingFunction[{{0., 109.687}}, <>]}}

Plot[Evaluate[{Q1[t], Q2[t]} /. %], {t, 0, tmax}, PlotPoints -> 300]

InterpolatingFunction::dmval : Input value {109.754} lies
  outside the range of data in the interpolating function. Extrapolation will be used.
InterpolatingFunction::dmval : Input value {110.169} lies
  outside the range of data in the interpolating function. Extrapolation will be used.
InterpolatingFunction::dmval : Input value {109.955} lies
  outside the range of data in the interpolating function. Extrapolation will be used.
General::stop : Further output of InterpolatingFunction::dmval will be suppressed during this calculation.
```



- Graphics -

## 8 Euklidische und unitäre Vektorräume: Fortsetzung

Im Folgenden greifen wir wieder das Thema auf, mit dem der erste Teil der Vorlesung endete: Skalarprodukte über  $K$ -Vektorräumen  $V$  mit  $K = \mathbb{R}$  oder  $\mathbb{C}$ , d.h. reelle bzw. unitäre Vektorräume.

Wir führen also zusätzlich zur Vektorraumstruktur ein zweites Strukturelement ein. Skalarprodukte sind *positiv definite* Sesquilinearformen, d.h. insbesondere gilt  $\langle v, v \rangle > 0$  für  $v \in V \setminus \{0\}$ . Wir nutzen also die Ordnung der reellen Zahlen.

Die Beschränkung auf reelle oder komplexe Vektorräume eröffnet uns eine

neue Welt, denn sie ermöglicht die Verschmelzung von Linearer Algebra und Analysis, d.h. z.B. die Betrachtung konvergenter Reihen  $\sum_{n=1}^{\infty} v_n$  von Vektoren  $v_n \in V$ . Die Konkretisierung dieser Perspektive bleibt aber der *Funktionalanalysis* vorbehalten, also einer weiterführenden Vorlesung.

Die Cauchy–Schwarz–Ungleichung

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\| \quad (v, w \in V)$$

ermöglicht die Definition des Winkels  $\varphi \in [0, \pi/2]$  zwischen den von  $v \neq 0$  und  $w \neq 0$  aufgespannten eindimensionalen Unterräumen:

$$\cos(\varphi) = \frac{|\langle v, w \rangle|}{\|v\| \cdot \|w\|}$$

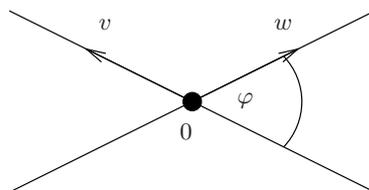


Abbildung 8.1: Winkel zwischen eindimensionalen Unterräumen

Besonders wichtig ist der Fall orthogonaler Vektoren mit  $\varphi = \pi/2$  bzw.  $\langle v, w \rangle = 0$ .

Für ein Skalarprodukt auf einem  $\mathbb{R}$ -Vektorraum  $V$  ist

$$\langle v, w \rangle = \frac{1}{4} (\langle v + w, v + w \rangle - \langle v - w, v - w \rangle) \quad (v, w \in V),$$

wir können also das Skalarprodukt aus der Norm rekonstruieren. Analoges gilt wegen der *Polarisationsidentität*

$$\langle v, w \rangle = \frac{1}{4} \left( \langle v + w, v + w \rangle - \langle v - w, v - w \rangle - i \langle v + iw, v + iw \rangle + i \langle v - iw, v - iw \rangle \right) \quad (v, w \in V)$$

auf  $\mathbb{C}$ -Vektorräumen  $V$ .

Insbesondere können wir die Winkelmessung durch die Längenmessung ersetzen.

## 8.1 Orthonormalisierung

Es sei  $V$  ein  $K$ -Vektorraum der Dimension<sup>4</sup>  $1 \leq \dim(V) < \infty$  mit einem Skalarprodukt.

**8.1 Definition** • Eine nicht leere Teilmenge  $M \subset V$  heißt Orthogonalsystem, wenn  $0 \notin M$  und  $\langle v, w \rangle = 0$  ( $v \neq w \in M$ ).

- Sie heißt Orthonormalsystem, wenn zusätzlich  $\|v\| = 1$  ( $v \in M$ ) gilt.
- Eine Basis  $B$  von  $V$  heißt Orthogonalbasis (Orthonormalbasis), wenn sie gleichzeitig ein Orthogonalsystem (Orthonormalsystem) ist.

Insbesondere mit Orthonormalbasen  $B = (b_1, \dots, b_n)$  lässt sich einfach rechnen, denn es gilt die Basisdarstellung

$$v := \sum_{i=1}^n \lambda_i b_i \quad \text{mit} \quad \lambda_i := \langle b_i, v \rangle \quad (v \in V).$$

Dies lässt sich folgendermaßen geometrisch deuten:

**8.2 Lemma** Ist  $B = (b_1, \dots, b_n)$  eine Orthonormalbasis, dann ist

$$P_i \in L(V) \quad , \quad P_i(v) := \langle b_i, v \rangle b_i \quad (i = 1, \dots, n)$$

eine Familie von Projektoren, d.h.

$$P_i P_j = \delta_{ij} P_j \quad \text{und} \quad \sum_{i=1}^n P_i = \text{Id}_V.$$

**Bew.:**

- $P_i P_j v = \langle b_i, \langle b_j, v \rangle b_j \rangle b_i = \langle b_j, v \rangle \langle b_i, b_j \rangle b_i = \langle b_j, v \rangle \delta_{ij} b_i$  ( $v \in V$ ).
- Weiter sind für  $k = 1, \dots, n$  die  $Q_k := \sum_{i=1}^k P_i$  Projektoren auf  $V_k := \text{span}(b_1, \dots, b_k)$ , denn  $Q_k^2 = \sum_{i,j=1}^k P_i P_j = \sum_{i,j=1}^k \delta_{ij} P_j = Q_k$ , und  $Q_k(\sum_{i=1}^n \lambda_i b_i) = \sum_{j=1}^k \sum_{i=1}^n \lambda_i P_j(b_i) = \sum_{i=1}^k \lambda_i b_i$ . Für  $k = n$  ist damit  $Q_k = \text{Id}_V$ .  $\square$

**8.3 Definition** Eine Projektion  $P = P^2 \in L(V)$  heißt Orthogonalprojektion, wenn  $\text{Kern}(P) = \text{Im}(P)^\perp$ .

---

<sup>4</sup>Ist  $\dim(V) = \infty$ , dann wird der Basisbegriff oft verändert, und man verlangt nicht mehr, dass jeder Vektor *endliche* Linearkombination von Basisvektoren sein muss. Siehe z.B. Reed und Simon, Functional Analysis, Kapitel II.3.

Wir wissen schon (Satz 4.4), dass für eine Projektion  $V = \text{Kern}(P) \oplus \text{Im}(P)$  gilt, wir können also jeden Vektor  $v \in V$  eindeutig in  $v = v_K + v_I$  mit  $v_K := v - P(v) \in \text{Kern}(P)$  und  $v_I := P(v) \in \text{Im}(P)$  zerlegen.

Umgekehrt ist  $P = P^2$  genau dann eine Orthogonalprojektion, wenn gilt:

$$\langle v - P(v), P(v) \rangle = 0 \quad (v \in V)$$

Weiter ist eine Projektion genau dann Orthogonalprojektion, wenn

$$\langle v, P(w) \rangle = \langle P(v), w \rangle \quad (v, w \in V),$$

denn

$$\begin{aligned} \langle v, P(w) \rangle &= \langle P(v), P(w) \rangle + \langle v - P(v), P(w) \rangle \\ &= \langle P(v), w \rangle + \langle v - P(v), P(w) \rangle + \langle P(v), P(w) - w \rangle. \end{aligned}$$

**8.4 Beispiel** 1. Die Projektionen  $P_b(v) := \langle b, v \rangle b$  für  $b \in V$ ,  $\|b\| = 1$  sind Orthogonalprojektionen mit  $\text{Im}(P_b) = \text{span}(b)$  und  $\text{Kern}(P_b) = \text{span}(b)^\perp$ .

2. Die Projektion  $P(v) := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} v$  auf dem Vektorraum  $\mathbb{R}^2$  mit Skalarprodukt  $\langle v, w \rangle = v_1 w_1 + v_2 w_2$  ist keine Orthogonalprojektion, denn  $\text{Kern}(P) = \text{span}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$  steht nicht senkrecht auf  $\text{Im}(P) = \text{span}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$ .

Wie können wir nun aus einer beliebigen Basis  $B = (b_1, \dots, b_n)$  von  $V$  eine Orthonormalbasis (ONB) konstruieren? Dies wird im Beweis des folgenden Satzes vorgeführt. Das Verfahren heißt *Gram-Schmidt-Orthonormalisierung*.

**8.5 Satz (E. Schmidt)** Ist  $B = (b_1, \dots, b_n)$  eine Basis von  $V$ , dann existiert eine ONB  $C = (c_1, \dots, c_n)$  von  $V$  mit

$$\text{span}(c_1, \dots, c_k) = \text{span}(b_1, \dots, b_k) \quad (k = 1, \dots, n).$$

**Bew.:**

- Wir setzen  $c_1 := b_1 / \|b_1\|$ , sodass  $\|c_1\| = 1$ .
- Haben wir für  $k < n$  schon eine Orthonormalbasis  $(c_1, \dots, c_k)$  von  $V_k := \text{span}(b_1, \dots, b_k)$  gefunden, dann ist der Projektor  $Q_k = \sum_{i=1}^k P_i$  wegen  $\langle Q_k(v), w \rangle = \sum_{i=1}^k \langle P_i(v), w \rangle = \sum_{i=1}^k \langle v, P_i(w) \rangle = \langle v, Q_k(w) \rangle$  Orthogonalprojektor auf  $V_k$ . Wir setzen

$$c_{k+1} := \frac{(\text{Id} - Q_k)b_{k+1}}{\|(\text{Id} - Q_k)b_{k+1}\|},$$

projizieren also  $b_{k+1}$  auf den zu  $V_k$  orthogonalen Unterraum, und normieren auf Länge 1. Wegen

$$\langle c_i, (\text{Id} - Q_k)b_{k+1} \rangle = \langle (\text{Id} - Q_k)c_i, b_{k+1} \rangle = \langle 0, b_{k+1} \rangle \quad (i = 1, \dots, k)$$

und  $\|c_{k+1}\| = 1$  bildet  $(c_1, \dots, c_{k+1})$  eine ONB von  $V_{k+1}$ .  $\square$

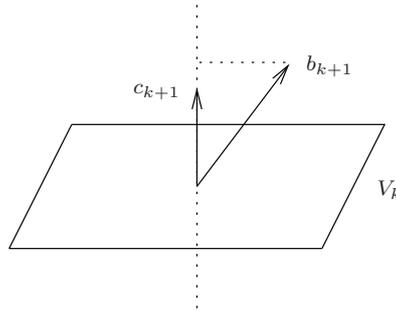


Abbildung 8.2: Gram-Schmidt-Orthonormalisierung

## 8.2 Die adjungierte Abbildung

Es seien  $V$  und  $W$  euklidische bzw. unitäre Vektorräume.

**8.6 Definition** Eine lineare Abbildung  $\varphi^* \in L(W, V)$  heißt zu  $\varphi \in L(V, W)$  adjungiert, wenn

$$\langle \varphi^*(w), v \rangle = \langle w, \varphi(v) \rangle \quad (v \in V, w \in W).$$

Gibt es eine solche adjungierte Abbildung, dann ist sie eindeutig.

**8.7 Satz** Ist  $\text{Dim}(V) < \infty$ , dann existiert  $\varphi^*$  und besitzt bezüglich einer ONB  $(b_1, \dots, b_n)$  die Form

$$\varphi^*(w) = \sum_{i=1}^n \langle \varphi(b_i), w \rangle b_i \quad (w \in W).$$

**Bew.:** Der Fall  $\text{Dim}(V) = 0$  ist trivial, und aus der Identität  $v = \sum_{j=1}^n \langle b_j, v \rangle b_j$  ergibt sich

$$\begin{aligned} \left\langle \sum_{i=1}^n \langle \varphi(b_i), w \rangle b_i, v \right\rangle &= \sum_{i,j=1}^n \langle w, \varphi(b_i) \rangle \langle b_i, \langle b_j, v \rangle b_j \rangle \\ &= \sum_{i,j=1}^n \langle w, \varphi(b_i) \rangle \langle b_j, v \rangle \delta_{ij} = \sum_{j=1}^n \langle w, \varphi(\langle b_j, v \rangle b_j) \rangle \\ &= \langle w, \varphi(v) \rangle. \end{aligned}$$

Wir erinnern uns an die Diskussion des Dualraumes  $V^*$  eines Vektorraumes (Kapitel 3).

Dessen Elemente sind die linearen Abbildungen  $V \rightarrow K$ . Ist nun  $w \in V$  und besitzt  $V$  ein Skalarprodukt, dann ist die Abbildung

$$I(w) \in V^* \quad , \quad I(w)(v) := \langle w, v \rangle$$

ein lineares Funktional auf  $V$ , und für  $w \neq 0$  ist auch  $I(w) \neq 0$ , denn  $I(w)(w) = \langle w, w \rangle > 0$ . Wir erhalten also eine Injektion

$$I : V \rightarrow V^* .$$

Diese ist *konjugiert-linear*, es gilt also

$$I(w_1 + w_2) = I(w_1) + I(w_2) \quad , \quad I(k \cdot w) = \bar{k} \cdot I(w) .$$

Damit ist für  $K = \mathbb{C}$  die Injektion *nicht* linear, für  $K = \mathbb{R}$  wegen  $\bar{k} = k$  aber schon.

Ist nun  $\text{Dim}(V) < \infty$  und  $B = (b_1, \dots, b_n)$  eine Orthonormalbasis, dann ist  $(I(b_1), \dots, I(b_n))$  gleich der dualen Basis, also der Basis  $(b_1^*, \dots, b_n^*)$  von  $V^*$  mit  $b_i^*(b_k) = \delta_{ik}$ .

Damit ist  $I_V : V \rightarrow V^*$  surjektiv und auch  $I_V^{-1} : V^* \rightarrow V$  konjugiert-linear.

**8.8 Satz** Sind  $V$  und  $W$  endlich-dimensionale euklidische oder unitäre Vektorräume, und ist  $\varphi \in L(V, W)$ , dann gilt für die adjungierte Abbildung

$$\varphi^* = I_V^{-1} \circ \varphi^t \circ I_W . \tag{8.1}$$

Ist  $B$  eine ONB von  $V$  und  $C$  eine ONB von  $W$ , dann gilt für die darstellende Matrizen  $A := \mathcal{M}_C^B(\varphi)$  und  $A^* := \mathcal{M}_B^C(\varphi)$

$$A^* = \overline{A^t} \quad , \quad \text{d.h.} \quad (A^*)_{i,k} = \overline{(A)_{k,i}} \quad \begin{array}{l} (i = 1, \dots, \text{Dim}(V), \\ j = 1, \dots, \text{Dim}(W)). \end{array}$$

**Bew.:** Nach Definition der transponierten Abbildung  $\varphi^t : W^* \rightarrow V^*$  gilt

$$\varphi^t(w^*)(v) = w^*(\varphi(v)) \quad (v \in V, w^* \in W^*) .$$

Daher ist für alle  $v \in V$  und  $w \in W$

$$\begin{aligned} \varphi^t(I_W(w))(v) &= I_W(w)(\varphi(v)) = \langle w, \varphi(v) \rangle \\ &= \langle \varphi^*(w), v \rangle = I_V(\varphi^*(w))(v) \end{aligned}$$

oder kurz

$$\varphi^t \circ I_W = I_V \circ \varphi^* .$$

Wegen der Invertierbarkeit von  $I_V$  impliziert dies Gleichung (8.1).

Nach Satz 3.10 ist

$$A^t = \mathcal{M}_{B^*}^{C^*}(\varphi^t)$$

die darstellende Matrix der transponierten Abbildung bezüglich der dualen Basen. Da  $I_V^{-1}$  und  $I_W$  konjugiert-linear sind und  $I_V^{-1}(B^*) = B$ ,  $I_W(C) = C^*$  gilt, ergibt sich daraus  $A^* = \overline{A^t}$ .  $\square$

Die adjungierte Abbildung hat also den gleichen Rang wie  $\varphi$ .

**8.9 Korollar** Sind  $V$  und  $W$  endlich-dimensionale euklidische oder unitäre Vektorräume, und ist  $\varphi \in L(V, W)$ , dann gilt

$$\text{Im}(\varphi^*) = \text{Kern}(\varphi)^\perp \subset V$$

und

$$\text{Kern}(\varphi^*) = (\text{Im}\varphi)^\perp \subset W.$$

**Bew.:** Dies ergibt sich aus dem entsprechenden Satz für die transponierte Abbildung, denn für einen Unterraum  $U \subset V$  ist der zu  $U$  in  $V^*$  orthogonale Unterraum

$$\{v^* \in V^*, v^*(u) = 0 \quad \forall u \in U\}$$

gleich dem Bild von

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \quad \forall u \in U\}$$

unter der Abbildung  $I_V : V \rightarrow V^*$ , und analog für  $W$ .  $\square$

Damit ist  $\varphi^*$  genau dann surjektiv, wenn  $\varphi$  injektiv ist, und umgekehrt.

**8.10 Beispiel** Die adjungierte Abbildung zur Projektion  $\varphi(v) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} v$  im  $\mathbb{R}^2$  ist  $\varphi^*(v) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} v$ .

$\varphi^*$  projiziert auf  $\text{Kern}(\varphi)^\perp = \text{span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$  entlang der Richtung  $\text{Im}(\varphi)^\perp = \text{span}\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$ .

Wir haben gesehen, dass eine Projektion  $P \in L(V)$  genau dann eine Orthogonalprojektion ist, wenn  $P^* = P$  ist.

### 8.3 Normale Operatoren

Diese so genannte Selbstadjungiertheit eines Operators ist Spezialfall der folgenden Eigenschaft.

**8.11 Definition** Existiert die adjungierte Abbildung  $\varphi^*$  zu  $\varphi \in L(V)$  und gilt  $\varphi^* \circ \varphi = \varphi \circ \varphi^*$ , dann heißt  $\varphi$  normal.

**8.12 Beispiel**  $\varphi(v) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} v$  ist nicht normal, denn  $\varphi^*(v) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} v$  und

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

**8.13 Satz** Für normale Endomorphismen  $\varphi \in L(V)$  gilt

$$\text{Kern}(\varphi) = \text{Kern}(\varphi^*) \quad \text{und} \quad \text{Im}(\varphi) = \text{Im}(\varphi^*).$$

**Bew.:**

- Allgemein gilt für einen zweiten Endomorphismus  $\psi$ :  $\text{Kern}(\psi \circ \varphi) \supset \text{Kern}(\varphi)$ . Für  $\psi = \varphi^*$  ist aber nach Corollar 8.9  $\text{Kern}(\varphi^*) = \text{Im}(\varphi)^\perp$ , sodass  $\text{Kern}(\varphi^* \circ \varphi) = \text{Kern}(\varphi)$ . Analog gilt  $\text{Kern}(\varphi \circ \varphi^*) = \text{Kern}(\varphi^*)$ , woraus wegen der Normalität von  $\varphi$  die Beziehung  $\text{Kern}(\varphi) = \text{Kern}(\varphi^*)$  folgt.
- Allgemein gilt  $\text{Im}(\varphi \circ \psi) \subset \text{Im}(\varphi)$ , aber für  $\psi = \varphi^*$  wegen  $\text{Kern}(\varphi) = \text{Im}(\varphi^*)^\perp$  sogar  $\text{Im}(\varphi \circ \varphi^*) = \text{Im}(\varphi)$ . Analog folgt  $\text{Im}(\varphi^* \circ \varphi) = \text{Im}(\varphi^*)$  und wegen der Normalität von  $\varphi$  die zweite Beziehung.  $\square$

**8.14 Satz** Ein Endomorphismus  $\varphi \in L(V)$  eines endlich-dimensionalen unitären Vektorraums  $V$  ist genau dann normal, wenn  $V$  eine ONB aus Eigenvektoren von  $\varphi$  besitzt.

**Bew.:**

- Besitzt  $V$  eine solche ONB, dann ist  $\varphi$  von der Form  $\varphi = \sum_{\lambda \in \text{Spek}(\varphi)} \lambda P_\lambda$  mit der Familie von Projektoren  $P_\lambda$  auf die Eigenräume zu den Eigenwerten  $\lambda$ . Da diese orthogonal sind, gilt  $P_\lambda^* = P_\lambda$ , also

$$\varphi^* \circ \varphi = \varphi \circ \varphi^* = \sum_{\lambda \in \text{Spek}(\varphi)} |\lambda|^2 P_\lambda.$$

- Ist  $\varphi$  normal, also  $[\varphi, \varphi^*] = 0$ , dann gilt auch  $[\varphi_D, \varphi_D^*] = [\varphi_N, \varphi_N^*] = 0$  für den diagonalisierbaren bzw. nilpotenten Anteil von  $\varphi$ , denn diese sind Polynome in  $\varphi$ . Daraus folgt zunächst  $\varphi_N = 0$ . Denn wegen Satz 8.13 gilt  $\text{Kern}(\varphi_N) = \text{Kern}(\varphi_N^*)$ . Damit ist  $\text{Kern}(\varphi_N) = \text{Im}(\varphi_N)^\perp$ , sodass  $\text{Kern}(\varphi_N^2) = \text{Kern}(\varphi_N)$ , also  $\text{Kern}(\varphi_N) = V$  ist.
- Gemäß Satz 5.26 können wir also  $\varphi = \varphi_D$  in der Form

$$\varphi = \sum_{\lambda \in \text{Spek}(\varphi)} \lambda P_\lambda$$

schreiben, wobei die Projektoren  $P_\lambda$  als Polynome von  $\varphi$  ebenfalls normal sind. Normale Projektoren sind aber Orthogonalprojektoren, denn nach Satz 8.13 projiziert  $P^*$  ebenso wie  $P$  entlang  $\text{Kern}(P) = \text{Kern}(P^*)$  auf  $\text{Im}(P) = \text{Im}(P^*)$ . Da andererseits nach Corollar 8.9  $\text{Kern}(P_\lambda) = \text{Im}(P_\lambda)^\perp$ , und  $\text{Im}(P_\mu) \subset \text{Kern}(P_\lambda)$  für  $\mu \in \text{Spek}(\varphi) \setminus \{\lambda\}$ , stehen je zwei Eigenräume  $P_\lambda$  und  $P_\mu$  aufeinander senkrecht.  $\square$

Insbesondere sind also normale Endomorphismen unitärer Vektorräume diagonalisierbar.

Auch wenn das charakteristische Polynom eines normalen Endomorphismus eines *euklidischen* Vektorraumes nicht in Linearfaktoren zerfällt, besitzt dieser eine einfache Normalform:

**8.15 Lemma** *Ist  $V$  ein euklidischer Vektorraum und  $\varphi \in L(V)$  normal, dann ist auch seine Komplexifizierung  $\varphi_{\mathbb{C}} \in L(V_{\mathbb{C}})$  normal bezüglich des komplexifizierten Skalarproduktes*

$$\langle v_1 + iv_2, w_1 + iw_2 \rangle := \langle v_1, w_1 \rangle + \langle v_2, w_2 \rangle + i(\langle v_1, w_2 \rangle - \langle v_2, w_1 \rangle).$$

**Bew.:**

$$\begin{aligned} & \langle v_1 + iv_2, \varphi_{\mathbb{C}}(w_1 + iw_2) \rangle \\ &= \langle v_1, \varphi(w_1) \rangle + \langle v_2, \varphi(w_2) \rangle + i(\langle v_1, \varphi(w_2) \rangle - \langle v_2, \varphi(w_1) \rangle) \\ &= \langle \varphi^*(v_1), w_1 \rangle + \langle \varphi^*(v_2), w_2 \rangle + i(\langle \varphi^*(v_1), w_2 \rangle - \langle \varphi^*(v_2), w_1 \rangle) \\ &= \langle (\varphi^*)_{\mathbb{C}}(v_1 + iv_2), w_1 + iw_2 \rangle \end{aligned}$$

für alle  $v_1 + iv_2, w_1 + iw_2 \in V_{\mathbb{C}}$  impliziert

$$(\varphi_{\mathbb{C}})^* = (\varphi^*)_{\mathbb{C}},$$

also

$$[\varphi_{\mathbb{C}}, (\varphi_{\mathbb{C}})^*] = [\varphi_{\mathbb{C}}, (\varphi^*)_{\mathbb{C}}] = [\varphi, \varphi^*]_{\mathbb{C}} = 0_{\mathbb{C}} = 0.$$

$\square$

Wir können also für  $\text{Dim}(V) < \infty$  eine reelle Normalform von  $\varphi$  dadurch berechnen, dass wir die komplexe Normalform von  $\varphi_{\mathbb{C}}$  benutzen, die nach Satz 8.14 bezüglich einer ONB eine Diagonalmatrix ist.

**8.16 Satz** *Ist  $V$  ein endlich-dimensionaler euklidischer Vektorraum und  $\varphi \in L(V)$  normal, dann besitzt  $\varphi$  bezüglich einer geeigneten ONB von  $V$  die Normal-*



Während bijektive lineare Abbildungen  $\varphi : V \rightarrow W$  die Vektorraumstrukturen ineinander überführen, werden durch orthogonale bzw. unitäre Transformationen auch die Skalarprodukte ineinander übergeführt.

Dies ist für  $K$ -Vektorräume  $V, W$  immer möglich, wenn  $V$  und  $W$  die gleiche endliche Dimension  $n$  besitzen, denn dann können wir ja ONBs  $(b_1, \dots, b_n)$  von  $V$  und  $(c_1, \dots, c_n)$  von  $W$  konstruieren und aufeinander abbilden:

$$\varphi(b_i) := c_i \quad (i = 1, \dots, n).$$

Lineare Fortsetzung ergibt eine orthogonale bzw. unitäre Transformation  $\varphi \in L(V, W)$ .

Wir können also, wenn wir diese Abbildungen untersuchen wollen, ohne Verlust an Allgemeinheit orthogonale Endomorphismen des  $\mathbb{R}^n$  bzw. unitäre Endomorphismen des  $\mathbb{C}^n$  studieren, wobei wir jeweils das Standard-Skalarprodukt verwenden.

**8.18 Definition** •  $O(n) := \{O \in \mathcal{M}(n, \mathbb{R}) \mid O \text{ ist orthogonal}\}$  heißt die orthogonale Gruppe des  $\mathbb{R}^n$ .

- $SO(n) := \{O \in O(n) \mid \text{Det}(O) = 1\}$  heißt die speziell orthogonale Gruppe des  $\mathbb{R}^n$ .
- $U(n) := \{U \in \mathcal{M}(n, \mathbb{C}) \mid U \text{ ist unitär}\}$  heißt die unitäre Gruppe des  $\mathbb{C}^n$ .
- $SU(n) := \{U \in U(n) \mid \text{Det}(U) = 1\}$  heißt die speziell unitäre Gruppe des  $\mathbb{C}^n$ .

Tatsächlich ist  $O(n)$  eine Untergruppe von  $GL(n, \mathbb{R})$ , denn  $O \in \mathcal{M}(n, \mathbb{R})$  ist genau dann orthogonal, wenn  $O^*O = \mathbb{1}$ , was wiederum gleichbedeutend mit  $O \in GL(n, \mathbb{R})$  und  $O^* = O^{-1}$  ist.  $O(n)$  ist also bezüglich Inversenbildung abgeschlossen, und wegen  $(O_1O_2)^* = O_2^*O_1^* = O_2^{-1}O_1^{-1} = (O_1O_2)^{-1}$  auch bezüglich Multiplikation. Zudem gilt  $\mathbb{1} \in O(n)$ .

Wegen des Determinantenproduktsatzes ist  $SO(n) \subset O(n)$  ebenfalls eine Untergruppe, und analog argumentiert man im unitären Fall.

Da das Skalarprodukt unter orthogonalen Abbildungen invariant gelassen wird, sind die Spaltenvektoren von  $O \in O(n)$ , also die Bilder der Standard-ONB  $(e_1, \dots, e_n)$  des euklidischen  $\mathbb{R}^n$ , eine ONB. Gleiches gilt für  $U \in U(n)$ .

Nach Satz 8.14 besitzt  $U$  bezüglich einer ONB des  $\mathbb{C}^n$  die Normalform  $\text{diag}(\lambda_1, \dots, \lambda_n)$ , wobei die  $\lambda_j$  die Eigenwerte von  $U$  sind, also wegen  $U^*U = \mathbb{1}$   $|\lambda_j| = 1$  gilt. Für  $U \in SU(n)$  ist zusätzlich  $\prod_{j=1}^n \lambda_j = 1$ .



Dagegen ist  $O(2)$  nicht abelsch, denn beispielsweise ist

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

**8.20 Beispiel** Die orthogonale Gruppe  $O(3)$  des  $\mathbb{R}^3$ .

• Ist  $O \in SO(3)$ , dann ist die reelle Normalform von  $O$  von der Gestalt

$$S^{-1}OS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix},$$

wobei die Transformationsmatrix  $S$  aus  $SO(3)$  gewählt werden kann. Wir können daher den Betrag des *Drehwinkels*  $\varphi$  aus

$$\operatorname{tr}(O) = \operatorname{tr}(S^{-1}OS) = 1 + 2 \cos(\varphi)$$

ablesen.

• Die *Drehachse* von  $O \neq \mathbb{1}$  ist der Eigenraum zum Eigenwert 1 und damit eindimensional. Die zu diesem eindimensionalen Unterraum orthogonale Ebene bleibt unter  $O$  als Menge invariant, die Vektoren dieser Ebene werden aber um den Winkel  $\varphi$  gedreht.

•  $O \in O(3) \setminus SO(3)$  besitzt die reelle Normalform

$$S^{-1}OS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}.$$

Dies entspricht einer Drehung um den Winkel  $\varphi$ , gefolgt von einer Spiegelung an der Drehebene. Letztere ist für  $\varphi \neq \pi$  eindeutig definiert.

• Wir können folgendermaßen eine stetige Abbildung

$$\phi : B \rightarrow SO(3)$$

von der Einheitskugel  $B := \{x \in \mathbb{R}^3 \mid \|x\| \leq 1\}$  des  $\mathbb{R}^3$  auf die Drehgruppe definieren:  $\phi(0) := \mathbb{1}$ , und für  $x \in B \setminus \{0\}$  sei  $\phi(x)$  die Drehung mit Achse  $\operatorname{span}(x) \subset \mathbb{R}^3$  um den Winkel  $\pi \cdot |x|$  in positiver Richtung bezüglich  $x$ . Es gilt  $\phi(x) = \phi(y)$  für  $x \neq y$  genau, wenn

$$y = -x \in \partial B := \{x \in \mathbb{R}^3 \mid \|x\| = 1\}.$$

Denn Drehungen um die gleiche Achse mit Winkel  $\pi$  bzw.  $-\pi$  sind einander gleich.  $\partial B$  ist die Oberfläche der Vollkugel  $B$ , und die Punkte  $x$  und  $-x \in \partial B$  sind Antipoden auf dieser zweidimensionalen Sphäre.

Die Vollkugel  $B$  mit identifizierten Antipoden auf  $\partial B$  ist damit ein topologisches Modell von  $SO(3)$  (siehe auch Knörrer [3], Kapitel 6).

- Schon  $SO(3)$  ist keine abelsche Gruppe mehr, denn Drehungen um verschiedene Achsen kommutieren nicht. Beispiel: Drehung mit Winkel  $\pi$  um  $x$ -Achse und  $z$ -Achse

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Die Isometrien des unitären Vektorraumes  $\mathbb{C}^n$  bilden die Gruppe  $U(n)$ . Die Determinante  $\text{Det}(U)$  einer unitären Matrix  $U$  ist eine komplexe Zahl, die wegen  $\overline{\text{Det}(U)} = \text{Det}(U^*) = \text{Det}(U^{-1}) = \text{Det}(U)^{-1}$  vom Betrag 1 ist. Die Isometrien  $U$  von  $\mathbb{C}$  beispielsweise bestehen aus der Multiplikation mit der komplexen Zahl  $U \in U(1) = \{c \in \mathbb{C} \mid |c| = 1\}$ . Wegen des Determinantenproduktsatzes ist

$$\text{Det} : U(n) \rightarrow U(1)$$

ein Gruppenhomomorphismus, und damit  $SU(n) = \text{Det}^{-1}(1)$  eine normale Untergruppe von  $U(n)$ .

### 8.21 Beispiel Die unitäre Gruppe $U(2)$ des $\mathbb{C}^2$ .

- Ist  $U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \in U(2)$ , dann muss  $|u_{11}|^2 + |u_{21}|^2 = 1 = |u_{12}|^2 + |u_{22}|^2$  und  $\overline{u_{11}}u_{12} + \overline{u_{21}}u_{22} = 0$  gelten. Dies ist genau dann erfüllt, wenn der erste Spaltenvektor  $\begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix}$  die Länge 1 besitzt, und der zweite Spaltenvektor von der Form  $\begin{pmatrix} u_{12} \\ u_{22} \end{pmatrix} = \lambda \begin{pmatrix} -\overline{u_{21}} \\ \overline{u_{11}} \end{pmatrix}$  für ein  $\lambda \in U(1)$  ist. In diesem Fall ist

$$\text{Det}(U) = u_{11}u_{22} - u_{21}u_{12} = \lambda(|u_{11}|^2 + |u_{21}|^2) = \lambda.$$

- Ist also  $U \in SU(2)$ , dann ist

$$U = \begin{pmatrix} v & -\overline{w} \\ w & \overline{v} \end{pmatrix} \quad \text{mit} \quad v, w \in \mathbb{C} \quad \text{und} \quad |v|^2 + |w|^2 = 1.$$

Der erste Spaltenvektor  $\begin{pmatrix} v \\ w \end{pmatrix}$  legt also schon die Matrix  $U$  fest, und  $\begin{pmatrix} v \\ w \end{pmatrix} \in \mathbb{C}^2$  ist ein beliebiger Vektor der Länge 1.

Die Vektoren  $\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} v_1 + iv_2 \\ w_1 + iw_2 \end{pmatrix}$  der Länge  $v_1^2 + v_2^2 + w_1^2 + w_2^2 = 1$  bilden aber die Einheitskugel  $S^3 \subset \mathbb{C}^2 \cong \mathbb{R}^4$ . Die dreidimensionale Einheitskugel  $S^3$  bildet damit ein topologisches Modell der Gruppe  $SU(2)$ .

### 8.22 Beispiel (Diskrete Fouriertransformation) Es sei

$$V_N := \{f : \mathbb{Z}/N\mathbb{Z} \rightarrow \mathbb{C}\} \quad , \quad \langle f, g \rangle := \sum_{x \in \mathbb{Z}/N\mathbb{Z}} \overline{f(x)}g(x)$$

der unitäre Vektorraum der komplexen Funktionen auf der abelschen Gruppe  $\mathbb{Z}/N\mathbb{Z}$ . Diese besitzt  $N \in \mathbb{N}$  Elemente, z.B. durch  $0, 1, \dots, N-1$  repräsentiert, und in Anwendungen ist  $f \in V_N$  oft als Diskretisierung einer Funktion  $F : [0, T] \rightarrow \mathbb{C}$

definiert:  $f(x) := F(x/T) \quad (x = 0, \dots, N-1)$ . Die *diskrete Fouriertransformation*  $U_N \in L(V_N)$  ist nun durch

$$(U_N f)(k) := N^{-1/2} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} f(x) e^{-2\pi i k x / N} \quad (k \in \mathbb{Z}/N\mathbb{Z})$$

definiert, sodass die adjungierte Abbildung gleich

$$(U_N^* g)(x) = N^{-1/2} \sum_{k \in \mathbb{Z}/N\mathbb{Z}} g(k) e^{2\pi i k x / N} \quad (x \in \mathbb{Z}/N\mathbb{Z})$$

ist.<sup>5</sup> Damit ist  $U_N$  unitär, denn

$$\begin{aligned} (U_N^* U_N f)(y) &= N^{-1} \sum_{x, k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k (y-x)/N) f(x) \\ &= N^{-1} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} N \delta_{y,x} f(x) = f(y) \end{aligned}$$

folgt (mit  $z = y - x$ ) aus der Relation

$$\sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) = N \delta_{z,0} \quad (z \in \mathbb{Z}/N\mathbb{Z}). \quad (8.2)$$

Die Identität (8.2) für  $z = 0$  folgt aus  $e^0 = 1$ , während für  $z \neq 0$  die Anwendung der Bijektion  $k \mapsto k + 1$  auf  $\mathbb{Z}/N\mathbb{Z}$

$$\begin{aligned} \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) &= \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i (k+1) z / N) \\ &= e^{2\pi i z / N} \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) \end{aligned}$$

oder

$$(1 - e^{2\pi i z / N}) \sum_{k \in \mathbb{Z}/N\mathbb{Z}} \exp(2\pi i k z / N) = 0$$

ergibt. Da der Faktor  $1 - e^{2\pi i z / N} \neq 0$  ist, verschwindet die Summe.

Wir haben mit dem Nachweis der Unitarität von  $U_N$  auch gezeigt, dass  $(b_1, \dots, b_N)$  mit  $b_k(x) := \frac{\exp(2\pi i k x / N)}{\sqrt{N}}$  eine ONB von  $V_N$  ist. Anders gesagt, entspricht der Fouriertransformation von  $f \in V_N$  die "Entwicklung von  $f$  nach den Wellen  $b_k$ "

$$f = \sum_{k=0}^{N-1} \lambda_k b_k \quad \text{mit} \quad \lambda_k := (b_k, f) = (U_N f)(k).$$

---

<sup>5</sup>Tatsächlich ist  $U_N$  wohldefiniert, denn für  $k' - k \in N\mathbb{Z}$  und  $x' - x \in N\mathbb{Z}$  ist  $\exp(2\pi i k' x' / N) = \exp(2\pi i k x / N)$ .

## 8.5 (Anti-) Selbstadjungierte Endomorphismen

**8.23 Definition** Ist  $V$  ein euklidischer bzw. unitärer Vektorraum, dann heißt ein  $\varphi \in L(V)$  (dessen Adjungierte  $\varphi^* \in L(V)$  existiert)

- selbstadjungiert, wenn  $\varphi^* = \varphi$ ,
- antiselbstadjungiert, wenn  $\varphi^* = -\varphi$  gilt.

Die Begriffe (anti-) symmetrisch und (anti-) hermitesch werden ebenfalls benutzt, der erste eher bei euklidischen und der zweite eher bei unitären Vektorräumen. Analoge Definitionen gelten für Matrizen aus  $\mathcal{M}(n, K)$  über den Körpern  $K = \mathbb{R}$  oder  $\mathbb{C}$ .

(Anti-) Selbstadjungierte Endomorphismen sind normal, denn

$$\varphi^* \circ \varphi = \pm \varphi \circ \varphi = \varphi \circ \varphi^*.$$

**8.24 Satz** Ist  $V$  ein unitärer Vektorraum,  $\dim(V) < \infty$  und  $\varphi \in L(V)$ , dann ist

- $\text{spek}(\varphi) \subset \mathbb{R}$ , falls  $\varphi$  selbstadjungiert und
- $\text{spek}(\varphi) \subset i \cdot \mathbb{R}$ , falls  $\varphi$  antiselbstadjungiert ist.

**Bew.:** Wegen der Normalität von  $\varphi$  ist  $\varphi$  diagonalisierbar, und

$$\varphi = \sum_{\lambda \in \text{spek}(\varphi)} \lambda P_\lambda,$$

wobei die  $P_\lambda$  Orthogonalprojektionen sind. Damit ist  $\varphi^* = \sum_{\lambda \in \text{spek}(\varphi)} \bar{\lambda} P_\lambda$ .

Falls  $\varphi^* = \varphi$  gilt, ist damit  $\bar{\lambda} = \lambda$ , also  $\text{spek}(\varphi) \subset \mathbb{R}$ .

Dagegen ist für  $\varphi^* = -\varphi$  der Eigenwert  $\lambda$  imaginär, denn  $\bar{\lambda} = -\lambda$ .  $\square$

**8.25 Korollar** Ist  $V$  ein euklidischer Vektorraum endlicher Dimension, und  $\varphi \in L(V)$  selbstadjungiert, dann besitzt  $V$  eine ONB aus Eigenvektoren von  $\varphi$ .

**Bew.:** Auch  $\varphi_{\mathbb{C}} \in L(V_{\mathbb{C}})$  ist selbstadjungiert, das charakteristische Polynom von  $\varphi_{\mathbb{C}}$  besitzt also nur reelle Nullstellen. Daher zerfällt das charakteristische Polynom von  $\varphi$  in Linearfaktoren, sodass  $\varphi$  diagonalisierbar ist. Wegen der Normalität von  $\varphi$  sind die Eigenräume  $E(\lambda)$  zu den Eigenwerten  $\lambda \in \text{spek}(\varphi)$  orthogonal.  $\square$

Dass Eigenvektoren  $v$  und  $w$  zu den Eigenwerten  $\mu \neq \lambda$  für selbstadjungierte Endomorphismen aufeinander senkrecht stehen, sieht man auch folgendermaßen:

$$0 = (\varphi(v), w) - (v, \varphi(w)) = (\lambda - \mu)(v, w) \quad , \text{ also } (v, w) = 0.$$

Jeder Endomorphismus  $\varphi \in L(V)$ , dessen Adjungierte  $\varphi^*$  existiert, lässt sich eindeutig in die Summe

$$\varphi = \varphi_s + \varphi_a \quad (8.3)$$

eines selbstadjungierten Operators  $\varphi_s := \frac{1}{2}(\varphi + \varphi^*)$  und eines antiselbstadjungierten Operators  $\varphi_a := \frac{1}{2}(\varphi - \varphi^*)$  zerlegen. Denn ist  $\varphi = \psi_s + \psi_a$  eine zweite solche Zerlegung, dann ist der Operator  $\pi := \varphi_s - \psi_s$  selbstadjungiert, aber wegen  $\pi = \psi_a - \varphi_a$  auch antiselbstadjungiert, sodass  $\pi = 0$  folgt.

**8.26 Bemerkung** Die Zerlegung (8.3) ist dann besonders nützlich, wenn  $\varphi_s$  und  $\varphi_a$  kommutieren. Dies ist wegen

$$[\varphi, \varphi^*] = [\varphi_s + \varphi_a, \varphi_s - \varphi_a] = 2[\varphi_a, \varphi_s]$$

genau dann der Fall, wenn  $\varphi$  normal ist.

Nach Satz 5.27 können wir damit für unitäre Vektorräume  $V$  endlicher Dimension normale  $\varphi$  dadurch diagonalisieren, dass wir  $\varphi_s$  und  $\varphi_a$  diagonalisieren, und eine  $\varphi_s$  und  $\varphi_a$  gemeinsame Basis von Eigenvektoren aufsuchen.

Der Vektorraum  $L(K^n)$  der Endomorphismen von  $\mathbb{R}^n$  bzw.  $\mathbb{C}^n$  besitzt (bezüglich des Standard-Skalarproduktes auf  $K^n$ ) eine Zerlegung

$$L(K^n) = L_s(K^n) \oplus L_a(K^n) \quad (8.4)$$

in den *reellen* Unterraum  $L_s(K^n) := \{\varphi \in L(K^n) \mid \varphi^* = \varphi\}$  selbstadjungierter Endomorphismen und den *reellen* Unterraum  $L_a(K^n)$  antiselbstadjungierter Endomorphismen.

**8.27 Lemma** *Es ist*

$$\dim_{\mathbb{R}}(L_s(K^n)) = \begin{cases} \frac{1}{2}n(n+1) & , \quad K = \mathbb{R} \\ n^2 & , \quad K = \mathbb{C} \end{cases}$$

und

$$\dim_{\mathbb{R}}(L_a(K^n)) = \begin{cases} \frac{1}{2}n(n-1) & , \quad K = \mathbb{R} \\ n^2 & , \quad K = \mathbb{C} \end{cases} .$$

**Bew.:** Wir betrachten die darstellenden Matrizen bezüglich der kanonischen Basis des  $K^n$ .

- Um eine symmetrische Matrix  $A \in \mathcal{M}(n, \mathbb{R})$  zu fixieren, müssen wir nur die  $\frac{1}{2}n(n+1)$  Einträge oberhalb und einschließlich der Diagonale frei wählen.
- Für eine symmetrische Matrix  $A \in \mathcal{M}(n, \mathbb{C})$  können wir die  $\frac{1}{2}n(n-1)$  komplexen Einträge oberhalb der Diagonale frei wählen, während die Einträge auf der Diagonale reell sein müssen, aber ansonsten frei wählbar sind.

- Die Dimensionen der reellen antisymmetrischen Teilräume ergeben sich aus (8.4):

$$\dim_{\mathbb{R}}(L_a(K^n)) = \dim_{\mathbb{R}}(\mathcal{M}_n(K)) - \dim_{\mathbb{R}}(L_s(K^n)).$$

□

Beachtet werden sollte, dass  $L_s(\mathbb{C}^n)$  und  $L_a(\mathbb{C}^n)$  keine  $\mathbb{C}$ -Unterräume sind, denn Multiplikation mit  $i$  permutiert diese reellen Unterräume.

Ebenso wenig ist im Allgemeinen das Produkt zweier (anti-) selbstadjungierter wieder (anti-) selbstadjungiert.

In Kapitel 7.1 wurde die Exponentialabbildung auf quadratische Matrizen angewandt.

**8.28 Satz** Für  $A \in \mathcal{M}(n, K)$ ,  $K = \mathbb{R}$  oder  $\mathbb{C}$ , gilt

$$\text{Det}(\exp(A)) = \exp(\text{tr}(A)). \quad (8.5)$$

**Bew.:**

- Ist  $A \in \mathcal{M}(n, \mathbb{C})$  und ist  $J = SAS^{-1}$  Jordan-Normalform, dann ist  $J$  eine obere Dreiecksmatrix, sodass

$$\text{Det}(\exp(J)) = \prod_{k=1}^n \exp(J_{kk}) = \exp\left(\sum_{k=1}^n J_{kk}\right) = \exp(\text{tr}(J))$$

gilt. Andererseits ist

$$\text{Det}(\exp(A)) = \text{Det}(S^{-1} \exp(J) S) = \text{Det}(\exp(J))$$

und

$$\text{tr}(A) = \text{tr}(S^{-1} J S) = \text{tr}(J),$$

woraus (8.5) folgt.

- Ist  $B \in \mathcal{M}(n, \mathbb{R})$ , dann gilt für die komplexifizierte Matrix  $B_{\mathbb{C}}$

$$\text{Det}_{\mathbb{R}}(B) = \text{Det}_{\mathbb{C}}(B_{\mathbb{C}}) \quad \text{und} \quad \text{tr}_{\mathbb{R}}(B) = \text{tr}_{\mathbb{C}}(B_{\mathbb{C}}),$$

woraus (8.5) auch für  $A \in \mathcal{M}(n, \mathbb{R})$  folgt. □

**8.29 Satz** Ist  $A \in \mathcal{M}(n, K)$  antiselbstadjungiert, dann ist

$$\exp(A) \in \begin{cases} SO(n) & , \text{ falls } K = \mathbb{R} \\ U(n) & , \text{ falls } K \in \mathbb{C} \end{cases} .$$

**Bew.:** Für  $O := \exp(A) \in \mathcal{M}(n, K)$  gilt  $O^* = \exp(A)^* = \exp(A^*) = \exp(-A) = O^{-1}$ , denn  $\exp(A)\exp(-A) = \exp(0) = \mathbb{1}$ . Andererseits folgt für  $A = -A^t \in \mathcal{M}(n, \mathbb{R})$  wegen  $\operatorname{tr} A = -\operatorname{tr}(A^t) = -\operatorname{tr}(A)$  die Identität  $\operatorname{Det}(\exp(A)) = \exp(\operatorname{tr}(A)) = e^0 = 1$ .  $\square$

**8.30 Beispiel** Für  $K = \mathbb{R}$  und  $n = 2$  ist  $A \in \mathcal{M}_a(2, \mathbb{R})$  von der Form  $A = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$  für ein  $a \in \mathbb{R}$ , und

$$\begin{aligned} \exp(A) &= \sum_{n=0}^{\infty} \frac{A^n}{n!} = \sum_{n=0}^{\infty} \frac{a^n \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^n}{n!} \\ &= \left( \sum_{m=0}^{\infty} \frac{a^{2m} (-1)^m}{(2m)!} \right) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \left( \sum_{m=0}^{\infty} \frac{a^{2m+1} (-1)^m}{(2m+1)!} \right) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \cos(a) \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin(a) \cdot \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos(a) & -\sin(a) \\ \sin(a) & \cos(a) \end{pmatrix}. \end{aligned}$$

Es handelt sich also um eine Drehung um den Winkel  $a$ .

## 8.6 Anwendung: Quantenmechanik

Zerlegt man das Sonnenlicht mit einem Prisma, dann zeigen sich neben einem kontinuierlichen Spektrum feine Linien. Deren Existenz folgt aus der Tatsache, dass die in einem Atom oder Molekül gebundenen Elektronen nur diskrete Energiewerte annehmen können; man spricht von einer *gequantelten* Energie. Die Vermutung, dass diese reellen Energiewerte Eigenwerte eines selbstadjungierten Operators sind, ist naheliegend und zutreffend. Die Theorie dieser so genannten *Schrödingeroperatoren* ist die Quantenmechanik.

**8.31 Beispiel** Für den Fall des *Wasserstoffatoms* ist der unitäre Vektorraum des Elektrons der Raum  $L^2(\mathbb{R}^3)$  der quadratintegrierbaren Funktionen<sup>6</sup>  $f : \mathbb{R}^3 \rightarrow \mathbb{C}$  mit dem Skalarprodukt

$$\langle f, g \rangle := \int_{\mathbb{R}^3} \bar{f}(x)g(x)dx.$$

Der Schrödingeroperator  $H$  auf  $L^2(\mathbb{R}^3)$  ist von der Form

$$(Hf)(x) := -\frac{1}{2}\Delta f(x) - \frac{f(x)}{|x|},$$

wobei  $\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \frac{\partial^2 f}{\partial x_3^2}$  der so genannte Laplaceoperator ist, und die Multiplikation mit dem Potential  $-\frac{1}{|x|}$  der Kernladung die elektrostatische Anziehung

<sup>6</sup>Genauer: Äquivalenzklassen solcher Funktionen  $f_1, f_2$ , die sich nur um eine Funktion  $f_1 - f_2$  mit  $\int_{\mathbb{R}^3} |f_1 - f_2|^2 dx = 0$  unterscheiden.

des Elektrons durch den Kern beschreibt. Man kann nun das Spektrum von  $H$  berechnen:

$$\sigma(H) = [0, \infty) \cup \{-1/(2n^2) \mid n \in \mathbb{N}\}.$$

In einer geeigneten Energieeinheit entspricht diese Teilmenge von  $\mathbb{R}$  genau den Energien, die ein Elektron im Kraftfeld eines Wasserstoffatomkerns (eines Protons) annehmen kann: Für positive Energien ist das Elektron ungebunden und kann beliebige Energien annehmen, während die für das Wasserstoffatom beobachteten Spektralwerte Energiedifferenzen  $\frac{1}{2} \left( \frac{1}{m^2} - \frac{1}{n^2} \right)$  entsprechen, die dem Übergang von einem zum anderen gebundenen Zustand entsprechen.

Ein Hauptziel der Quantenmechanik ist die (meist näherungsweise) Berechnung des Spektrums und der Eigenfunktionen der Schrödingeroperatoren  $H$ . Dies ist auch deshalb wichtig, weil es der Schlüssel zur Berechnung der *Zeitentwicklungsoperatoren*

$$U(t) := \exp(-iHt) \quad (t \in \mathbb{R})$$

ist, die den quantenmechanischen Anfangszustand  $\varphi$  zur Zeit Null auf den Zustand  $U(t)\varphi$  zur Zeit  $t$  abbilden. Die Operatoren  $U(t)$  sind unitär, wenn  $H$  selbstadjungiert ist.

## 9 Quadriken

Bis jetzt wurden in dieser Vorlesung hauptsächlich solche Teilmengen  $M \subset V$  von Vektorräumen  $V$  betrachtet, die selbst Vektorräume sind, oder bei denen zumindest die Menge  $\{m_1 - m_2 \mid m_1, m_2 \in M\}$  aller Differenzvektoren einen Unterraum von  $V$  bildet (nicht leerer affiner Unterraum  $M$ ).

Jetzt werden wir verallgemeinert Quadriken untersuchen. Beispiele solcher Quadriken für  $V = \mathbb{R}^2$  sind neben den Geraden die Ellipsen, Parabeln und Hyperbeln. Um Fallunterscheidungen zu vermeiden, betrachten wir nur Körper  $K$  der Charakteristik  $\neq 2$ .

**9.1 Definition** Eine Teilmenge  $Q \subset K^n$  heißt *Quadrik*, wenn ein quadratisches Polynom  $p \in K[x_1, \dots, x_n]$  mit

$$Q = \{(x_1, \dots, x_n) \in K^n \mid p(x_1, \dots, x_n) = 0\}$$

existiert.

Wir können nun mit  $x := (x_1, \dots, x_n)^t$  dieses Polynom in der Form

$$p(x) = x^t A x + b^t x + c \quad (x \in K^n)$$

schreiben, wobei  $A \in \mathcal{M}(n, k)$ ,  $b \in K^n$  und  $c \in K$  eindeutig bestimmt sind, wenn wir  $A = A^t$  (also die Symmetrie von  $A$  fordern (Letzteres können wir wegen  $\text{char}(K) \neq 2$  immer durch Übergang von  $A$  zu  $\frac{1}{2}(A + A^t)$  erreichen).

Man sieht dem Polynom  $p$  nicht direkt die geometrische Form der Quadrik  $Q$  an. Überhaupt müssen wir uns zunächst über diese Formen ins Bild setzen. Dazu ist es naheliegend, eine Koordinatentransformation des  $K^n$  vorzunehmen, die das Polynom  $p$  vereinfacht. In Frage kommen zunächst die so genannten Affinitäten:

**9.2 Definition** Eine Abbildung  $f : V \rightarrow W$  von  $K$ -Vektorräumen heißt affin, wenn die Abbildung

$$\tilde{f} : V \rightarrow W \quad \tilde{f}(x) := f(x) - f(0)$$

linear ist. Sie heißt Affinität, wenn sie zusätzlich bijektiv ist.

Die Affinitäten  $f : K^n \rightarrow K^n$  bilden eine Gruppe, denn jedes solche  $f$  besitzt die Form  $f(x) = Ax + v$  mit  $A \in GL(n, K)$  und  $v \in K^n$ , sodass

$$f^{-1}(y) = A^{-1}y - A^{-1}v \quad \text{und} \quad f \circ g(x) = ABx + (v + Aw)$$

für  $g(x) = Bx + w$  gilt.

**9.3 Beispiel** Beispiele für Affinitäten des  $\mathbb{R}^n$  sind die Drehungen  $f(x) = Ox$  mit  $O \in SO(n)$ , die Streckungen  $f(x) = cx$  um einen Faktor  $c \in \mathbb{R} \setminus \{0\}$  oder die Translationen  $f(x) = x + v$  um den Vektor  $v \in \mathbb{R}^n$ .

Es ist nun rationell, sowohl die Affinitäten als auch die Quadriken des  $K^n$  im  $K^{n+1}$  darzustellen. Ist nämlich

$$f(x) = Ax + v \quad (x \in K^n)$$

eine Affinität, dann gilt für die erweiterte Matrix

$$E_f := \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in \mathcal{M}(n+1, K)$$

die Eigenschaft eines Gruppenhomomorphismus

$$E_g E_f = \begin{pmatrix} 1 & 0 \\ w & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ w+Bv & BA \end{pmatrix} = E_{g \circ f},$$

(wobei  $g(x) = Bx + w$  eine weitere Affinität des  $K^n$  sei).

Ordnen wir dem Vektor  $x \in K^n$  den erweiterten Vektor  $E_x := \begin{pmatrix} 1 \\ x \end{pmatrix} \in K^{n+1}$  zu, dann gilt

$$E_f E_x = \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} = \begin{pmatrix} 1 \\ Ax+v \end{pmatrix} = E_{f(x)}. \quad (9.1)$$

Ordnen wir andererseits dem quadratischen Polynom  $p \in K[x_1, \dots, x_n]$  der Form

$$p(x) = (x, Ax) + (b, x) + c$$

(mit der Bilinearform  $(b, x) = b_1x_1 + \dots + b_nx_n$ ) die Matrix

$$S_p := \begin{pmatrix} c & b^t/2 \\ b/2 & (A+A^t)/2 \end{pmatrix} \in \mathcal{M}(n+1, K)$$

zu, dann ist  $S_p$  symmetrisch ( $S_p^t = S_p$ ), und es gilt

$$p(x) = (E_x, S_p E_x). \quad (9.2)$$

Wir wollen nun durch eine Affinität  $f$  des  $K^n$  die  $p$  zugeordnete Quadrik  $p^{-1}(0) \subset K^n$  auf eine möglichst einfache Form bringen. D.h. wir substituieren  $x = f(y)$  in  $p(x)$ . Wegen (9.2) und (9.1) bedeutet dies

$$\begin{aligned} p(f(y)) &= (E_{f(y)}, S_p E_{f(y)}) = (E_f E_y, S_p E_f E_y) \\ &= (E_y, S'_p E_y) \quad \text{mit} \quad S'_p := E_f^t S_p E_f. \end{aligned} \quad (9.3)$$

Zur Lösung dieses Normalformproblem es für den Spezialfall  $K = \mathbb{R}$  schauen wir uns zunächst das verwandte Problem der Normalform unter orientierungserhaltenden Kongruenzen an:

#### 9.4 Definition Eine Affinität $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt

- Kongruenz, wenn sie Abstände erhält, d.h.

$$\|f(x) - f(y)\| = \|x - y\| \quad (x, y \in \mathbb{R}^n). \quad (9.4)$$

- Sie heißt orientierungserhaltend, wenn  $\text{Det}(A) > 0$  (mit  $f(x) = Ax + v$ ) gilt.

Eine Affinität  $f(x) = Ax + v$  ist genau dann eine Kongruenz, wenn  $A \in O(n)$  gilt, und genau dann eine orientierungserhaltende Kongruenz, wenn  $A \in SO(n)$ .

Weiter kann man zeigen, dass jede abstandserhaltende Abbildung  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  (für die also (9.4) gilt) eine Affinität ist.

**9.5 Satz** Ist  $p \in \mathbb{R}[x_1, \dots, x_n]$  ein quadratisches Polynom der Form  $p(x) = (x, Ax) + (b, x) + c$  mit  $A = A^t$  und  $m := \text{Rang}(A)$ , dann existiert eine orientierungserhaltende Kongruenz  $f$  des  $\mathbb{R}^n$  mit

$$p \circ f(y) = \begin{cases} \sum_{i=1}^m \lambda_i y_i^2 & , \text{Rang}(E_p) = m \\ \sum_{i=1}^m \lambda_i y_i^2 + c & , \text{Rang}(E_p) = m + 1 \\ \sum_{i=1}^m \lambda_i y_i^2 + 2\|b\|y_{m+1} & , \text{Rang}(E_p) = m + 2. \end{cases}$$

Dabei sind  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$  die von Null verschiedenen Eigenwerte der symmetrischen Matrix  $A$ .

**Bew.:** Gemäß (9.3) besitzt das quadratische Polynom  $p \circ f$  die Form  $p \circ f(y) = (E_y, S'_p E_y)$  mit  $S'_p = E_f^t S_p E_f$ . Wir setzen  $f$  aus vier orientierungserhaltenden Kongruenzen zusammen:  $f = f_4 \circ f_3 \circ f_2 \circ f_1$ .

1. Dabei ist  $f_1(x) := O_1 x$ , wobei wir zunächst nur  $O_1 \in O(n)$  fordern. Damit ist die Transponierte von  $E_{f_1} = \begin{pmatrix} 1 & 0 \\ 0 & O_1 \end{pmatrix}$  gleich  $E_{f_1}^t = \begin{pmatrix} 1 & 0 \\ 0 & O_1^{-1} \end{pmatrix}$  und es gibt ein  $O_1$  mit

$$E_{f_1}^t S_p E_{f_1} = \begin{pmatrix} c & \tilde{b}^t \\ \tilde{b} & D \end{pmatrix} \quad \text{und} \quad D = \text{diag}(\lambda_1, \dots, \lambda_m, 0, \dots, 0).$$

Ist  $O_1 \notin SO(n)$ , also  $f_1$  nicht orientierungserhaltend, dann können wir dies durch Übergang zur Drehmatrix  $\text{diag}(-1, 1, \dots, 1)O_1 \in SO(n)$  erzwingen. Ist  $\tilde{b} = 0$ , dann ist der Satz damit bewiesen.

2. Wir wählen  $f_2$  von der Form  $f_2(x) = x + v_2$  mit

$$(v_2)_i := \begin{cases} -\tilde{b}_i / \lambda_i & , \quad i \leq m \\ 0 & , \quad i > m. \end{cases}$$

Dann ist  $E_{f_2 \circ f_1}^t S_p E_{f_2 \circ f_1} = \begin{pmatrix} c & \hat{b}^t \\ \hat{b} & D \end{pmatrix}$ , wobei  $\hat{b}_1 = \dots = \hat{b}_m = 0$  ist.

3. Wegen der Annahme  $\tilde{b} \neq 0$  ist auch  $\hat{b} \neq 0$ . Es gibt aber eine Drehung  $O_3 \in SO(n)$  von der Form  $O_3 = \mathbb{1}_m \oplus \tilde{O}_3$  mit  $\tilde{O}_3 \in SO(n-m)$ , die  $O_3 \hat{b} = \mu \cdot e_{m+1}$  mit  $\mu := \|\hat{b}\|$  bewirkt,  $\hat{b}$  also in Richtung des  $(m+1)$ -ten Einheitsvektors dreht. Damit gilt für  $f_3(x) := O_3 x$

$$E_{f_3 \circ f_2 \circ f_1}^t S_p E_{f_3 \circ f_2 \circ f_1} = \begin{pmatrix} c & \mu \cdot e_{m+1}^t \\ \mu \cdot e_{m+1} & D \end{pmatrix},$$

die Diagonalmatrix  $D$  bleibt also unverändert.

4. Ist  $v_4 := -\frac{c}{2\mu} e_{m+1}$ , dann ist für  $f_4(x) := x + v_4$  und  $f = f_4 \circ f_3 \circ f_2 \circ f_1$

$$S'_p = E_f^t S_p E_f = \begin{pmatrix} 0 & \mu e_{m+1}^t \\ \mu e_{m+1} & D \end{pmatrix},$$

also  $p \circ f(y) = \sum_{i=1}^m \lambda_i y_i^2 + 2\|b\|y_{m+1}$ . □

**9.6 Bemerkung** Ist  $b = 0$  und  $c = 0$ , dann ist  $p$  die quadratische Form  $p(x) = (x, Ax)$ , und der Satz besagt, dass es eine Drehung  $x = f(y) = Oy$  des  $\mathbb{R}^n$  gibt, die  $p$  in Diagonalfom  $p \circ f(y) = \sum_{i=1}^m \lambda_i y_i^2$  transformiert.  $O \in O(n)$  heißt dann auch *Hauptachsentransformation*.

Wir kehren zu unserer ursprünglichen Fragestellung der Normalform von Quadriken unter Affinitäten zurück und zeigen:

**9.7 Satz** Ist  $p \in \mathbb{R}[x_1, \dots, x_n]$  ein quadratisches Polynom der Form  $p(x) = (x, Ax) + (b, x) + c$  mit  $A = A^t$  und  $m := \text{Rang}(A)$ , dann existiert eine orientierungserhaltende Affinität  $g$  des  $\mathbb{R}^n$ , für die die Quadrik  $(p \circ g)^{-1}(0) \subset \mathbb{R}^n$  durch die Gleichung

$$\begin{cases} z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_m^2 = 0 & , \text{Rang}(E_p) = m \\ z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_m^2 = 1 & , \text{Rang}(E_p) = m + 1 \\ z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_m^2 = 2z_{m+1} & , \text{Rang}(E_p) = m + 2 \end{cases}$$

beschrieben wird.

**Bew.:**

- Ist  $g_1(z) = Mz$  mit  $M := \text{diag}(|\lambda_1|^{-1/2}, \dots, |\lambda_m|^{-1/2}, 1, \dots, 1)$ , dann ist unter Verwendung der orientierungserhaltenden Kongruenz  $f$  aus Satz 9.5

$$p \circ f \circ g_1(z) = \sum_{i=1}^k z_i^2 - \sum_{i=k+1}^m z_i^2, \text{ falls } \text{Rang}(E_p) = m.$$

- Ist dagegen  $\text{Rang}(E_p) = m + 1$ , dann setzen wir  $g_1(z) = Mz$  mit  $M := \text{diag}\left(\left|\frac{c}{\lambda_1}\right|^{1/2}, \dots, \left|\frac{c}{\lambda_m}\right|^{-1/2}, 1, \dots, 1\right)$ , sodass  $p \circ f \circ g_1(z) = |c| \left(\sum_{i=1}^k z_i^2 - \sum_{i=k+1}^m z_i^2 + \text{sign}(c)\right)$  ist.

Ist  $\text{sign}(c) = +1$ , dann permutieren wir noch die  $z_1, \dots, z_k$  mit den  $z_{k+1}, \dots, z_m$ . Dies ist eine orthogonale Transformation, und wir können wieder die Positivität der Determinante durch Multiplikation einer Koordinate mit  $-1$  erzwingen.

Die Quadrik von  $p \circ f \circ g_1$  hängt nicht vom Vorfaktor  $|c| > 0$  ab.

- Im letzten Fall  $\text{Rang}(E_p) = m + 2$  wird durch Übergang zur Variablen  $z_{m+1} := \|b\|y_{m+1} - \frac{c}{2}$  reskaliert und die Konstante  $c$  gleichzeitig durch Null ersetzt.  $\square$

**9.8 Beispiel** 1. Für  $n = 1$  kommen die Fälle  $Q = \{0\}$ ,  $Q = \{-1, 1\}$ ,  $Q = \mathbb{R}$  und  $Q = \emptyset$  vor, entsprechend den Kardinalitäten der Lösungsmengen quadratischer Gleichungen.

2. Für  $n = 2$  können wir zunächst die Quadriken der eindimensionalen Normalformen mit  $\mathbb{R}$  multiplizieren, und erhalten die Gerade  $Q = \mathbb{R} \times \{0\}$ , die Doppelgerade  $Q = \mathbb{R} \times \{-1, 1\}$ , die Ebene  $Q = \mathbb{R}^2$  und die leere Menge.

Die Gleichung  $z_1^2 + z_2^2 = 0$  hat als Lösungsmenge den *Nullpunkt*, während für  $z_1^2 - z_2^2 = 0$  die beiden Winkelhalbierenden die Quadrik  $Q = \text{span}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cup \text{span}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$  ergeben.

Diese sind die *Asymptoten* der *Hyperbel*  $Q = \{(z_1, z_2) \in \mathbb{R}^2 \mid z_1^2 - z_2^2 = 1\}$ .

Die Quadrik zur Gleichung  $z_1^2 + z_2^2 = 1$  ist natürlich der *Kreis* mit Radius 1 um den Nullpunkt.

Zuletzt beschreibt die Gleichung  $z_1^2 = 2z_2$  eine *Parabel*.

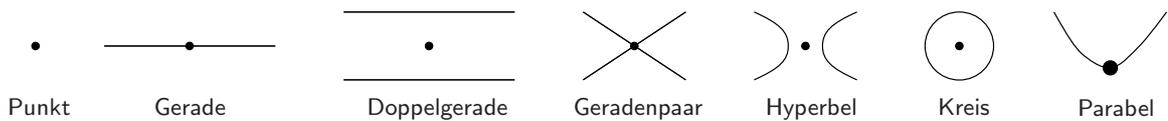
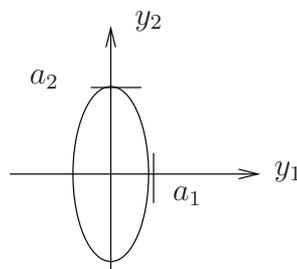


Abbildung 9.1: Affine Normalformen der Quadriken für  $n = 2$  (außer  $Q = \mathbb{R}^2$ ,  $Q = \emptyset$ ) mit eingezeichnetem Ursprung des Koordinatensystems

Die Äquivalenzklasseneinteilung bezüglich orientierungserhaltender Kongruenzen ist feiner als die bezüglich Affinitäten. Beispielsweise erhalten wir statt des Kreises die Kongruenznormalformen  $\frac{y_1^2}{a_1^2} + \frac{y_2^2}{a_2^2} = 1$  mit  $0 < a_1 \leq a_2$  der *Ellipsen*, deren Halbachsen mit den Längen  $a_i$  in Richtung der  $y_i$ -Achsen gedreht sind:



Wir sehen, dass wir sowohl im affinen als auch im Kongruenzfall Normalformen von gewissen Quadriken der Dimension  $n + 1$  erhalten, indem wir im  $\mathbb{R}^{n+1}$  (statt im  $\mathbb{R}^n$ ) die Nullstellenmenge der quadratischen Polynome von Normalformen der Dimension  $n$  betrachten. Man spricht dann von *Zylindern*:

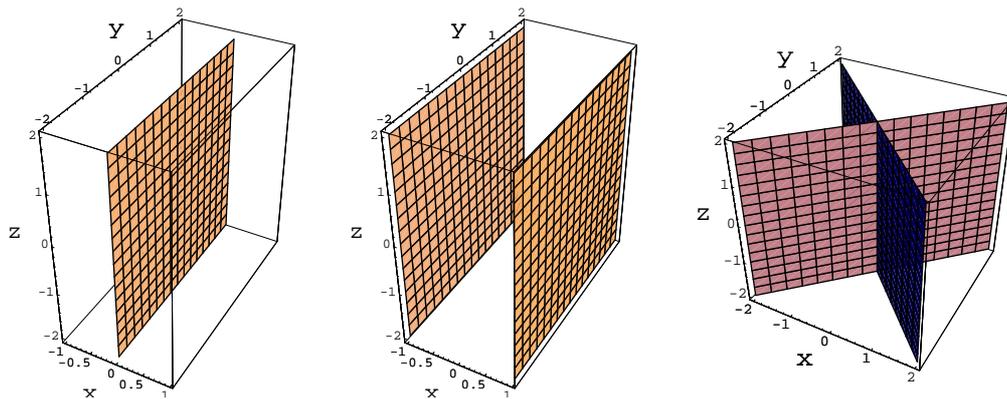


Abbildung 9.2: Affine Normalformen von Quadriken für  $n = 3$ :  $x$  (Ebene),  $x^2 - 1$  (Doppel-Ebene),  $x^2 - y^2$  (Ebenenpaar)

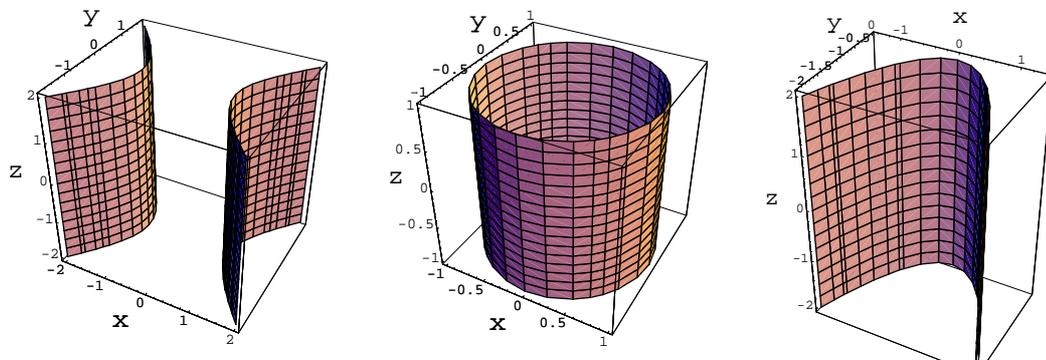


Abbildung 9.3: Affine Normalformen von Quadriken für  $n = 3$ :  $x^2 - y^2 - 1$  (Zylinder über Hyperbel),  $x^2 + y^2 - 1$  (Zylinder über Kreis),  $x^2 - 2y$  (Zylinder über Parabel)

Daneben finden wir in jeder Dimension jedoch auch Quadriken, die nicht Zylinder sind:

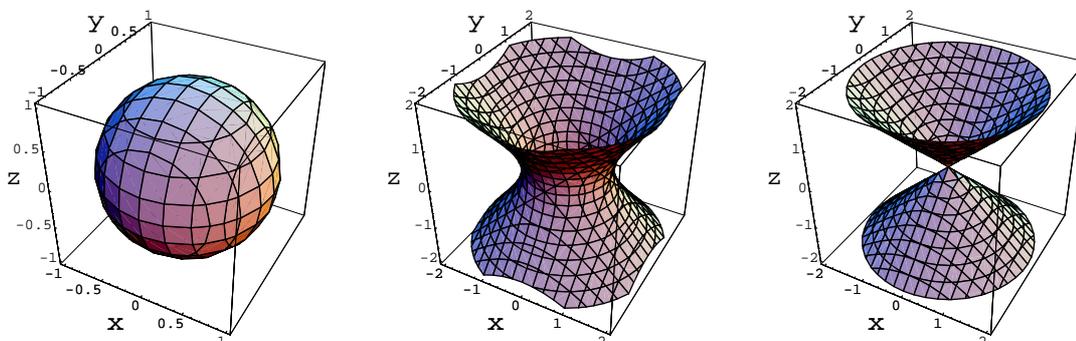


Abbildung 9.4: Affine Normalformen von Quadriken für  $n = 3$ :  $x^2 + y^2 + z^2 - 1$  (Sphäre),  $x^2 + y^2 - z^2 - 1$  (einschaliges Hyperboloid),  $x^2 + y^2 - z^2$  (Kegel)

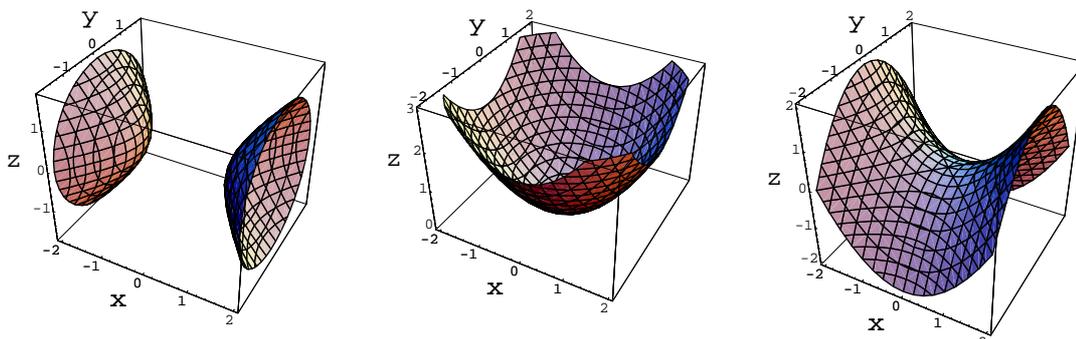


Abbildung 9.5: Affine Normalformen von Quadriken für  $n = 3$ :  $x^2 - y^2 - z^2 - 1$  (zweischaliges Hyperboloid),  $x^2 + y^2 - 2z$  (Paraboloid),  $x^2 - y^2 - 2z$  (Sattel)

## 10 Projektive Räume

**10.1 Definition** • Der projektive Raum  $\mathbb{P}(V)$  zum Vektorraum  $V$  ist die Menge der eindimensionalen Unterräume von  $V$ .

- Ist  $\text{Dim}(V) < \infty$ , dann heißt  $\dim(\mathbb{P}(V)) := \text{Dim}(V) - 1$  die Dimension des projektiven Raumes.

- Eine Teilmenge  $M \subset \mathbb{P}(V)$  der Form  $M = \mathbb{P}(U)$ ,  $U \subset V$  heißt Unterraum von  $\mathbb{P}(V)$ , und zwar projektive Gerade, falls  $\dim(M) = 1$ , projektive Ebene, falls  $\dim(M) = 2$  und projektive Hyperebene, falls  $\dim(M) = \dim(\mathbb{P}(V)) - 1$ .

Eine Sichtweise des projektiven Raumes  $\mathbb{P}_n(K) := \mathbb{P}(K^{n+1})$  zum arithmetischen Vektorraum ist die folgende. Bequemlichkeitshalber bezeichnen wir die kanonische Basis des  $K^{n+1}$  mit  $(e_0, e_1, \dots, e_n)$  statt mit  $(e_1, e_2, \dots, e_{n+1})$ , und die Koordinaten bezüglich dieser Basis mit  $(x_0, x_1, \dots, x_n)$ . Der Unterraum  $\text{span}(e_1, \dots, e_n) \subset K^{n+1}$  ist  $n$ -dimensional, und sei mit  $K^n$  bezeichnet. Dann geht durch jeden Punkt  $x \in U$  des affinen Unterraums

$$U := e_0 + K^n = \{(x_0, \dots, x_n) \in K^{n+1} \mid x_0 = 1\}$$

genau ein eindimensionaler Unterraum, nämlich  $\text{span}(x)$ . Nur die eindimensionalen Unterräume, die in  $K^n$  liegen, schneiden  $U$  nicht.

Die Abbildung

$$I : K^n \rightarrow \mathbb{P}_n(K) \quad , \quad (x_1, \dots, x_n) \mapsto \text{span}(1, x_1, \dots, x_n) \quad (10.1)$$

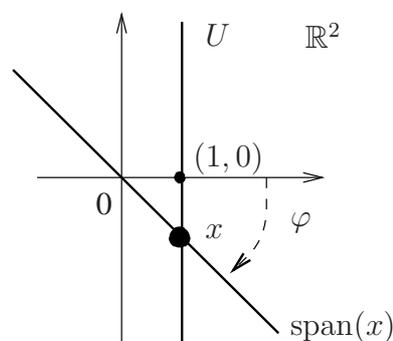
ist daher injektiv, und das Komplement besteht aus den eindimensionalen Unterräumen des  $K^n$ :

$$\mathbb{P}_n(K) \setminus I(K^n) \cong \mathbb{P}_{n-1}(K).$$

Anders gesagt, wird im projektiven Raum  $\mathbb{P}_n(K)$  der Vektorraum  $K^n$  noch um "Punkte bei Unendlich" ergänzt, und diese bilden selbst den projektiven Raum  $\mathbb{P}_{n-1}(K)$ .

**10.2 Beispiel** Am einfachsten lässt sich dies im reellen Fall  $K = \mathbb{R}$  visualisieren.

1. Im Fall des  $\mathbb{P}_1(\mathbb{R})$  ist  $U$  die affine Gerade  $(1, 0) + \text{span}(0, 1)$



Es gilt  $\mathbb{P}_1(\mathbb{R}) \cong \mathbb{R} \cup \mathbb{P}_0(\mathbb{R}) \cong \mathbb{R} \cup \{\infty\}$ , denn  $\mathbb{P}_0(\mathbb{R}) = \mathbb{P}(\mathbb{R})$  besteht aus genau einem Element, das wir hier mit  $\infty$  bezeichnen.

$x \in U$  ist von der Form  $x = (1, \tan \varphi)$  mit  $\varphi \in (-\pi/2, \pi/2)$ , und strebt  $\|x\|$  gegen unendlich, dann strebt  $|\varphi|$  gegen  $\pi/2$ , also  $e^{2i\varphi}$  gegen  $-1$ . Daher ist die Abbildung

$$\Phi : \mathbb{P}_1(\mathbb{R}) \rightarrow S^1 \quad , \quad \text{span}(x) \mapsto e^{2i\varphi} \quad , \quad \mathbb{R} \mapsto -1$$

des eindimensionalen, reell-projektiven Raumes  $\mathbb{P}_1(\mathbb{R})$  in die Kreislinie  $S^1 = \{c \in \mathbb{C} \mid |c| = 1\}$  bijektiv. Wir können uns also den  $\mathbb{P}_1(\mathbb{R})$  als Kreislinie vorstellen.

2. Dies lässt sich auf höhere Dimensionen verallgemeinern. Jeder eindimensionale Unterraum des  $\mathbb{R}^{n+1}$  schneidet die Sphäre  $S^n := \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$  in genau zwei Punkten  $x$  und  $y$ , und diese sind *Antipoden*, d.h.  $x + y = 0$ . Andererseits geht durch jedes Antipodenpaar von  $S^n$  genau ein eindimensionaler Unterraum. Bezeichnet daher  $\sim$  die Äquivalenzrelation auf  $S^n$ , die durch  $x \sim y \iff \text{span}(x) = \text{span}(y)$  definiert ist, dann ist die Abbildung

$$\psi : (S^n / \sim) \rightarrow \mathbb{P}_n(\mathbb{R}) \quad , \quad [x] \mapsto \text{span}(x)$$

bijektiv. Der  $n$ -dimensionale projektive Raum  $\mathbb{P}_n(\mathbb{R})$  kann also mit  $S^n / \sim$  identifiziert werden.

Nun ist zwar  $S^1 / \sim$  selbst wieder eine Kreislinie, aber Ähnliches gilt für höhere Dimensionen nicht.

Wenn wir uns den  $\mathbb{P}_2(\mathbb{R})$  vorstellen wollen, dann können wir auch die Nordhalbkugel  $N := \{x \in S^2 \mid x_1 \geq 0\}$  nehmen und Antipoden auf dem Äquator  $\{x \in N \mid x_1 = 0\}$  identifizieren.

Der *Schnitt* zweier projektiver Unterräume von  $\mathbb{P}(V)$  ist wieder ein projektiver Unterraum, und für die Unterräume  $U_1, U_2$  von  $V$  gilt

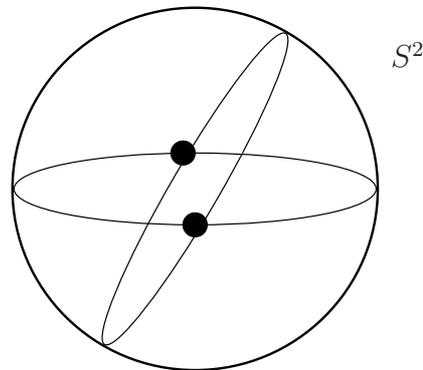
$$\mathbb{P}(U_1) \cap \mathbb{P}(U_2) = \mathbb{P}(U_1 \cap U_2).$$

Andererseits ist auch der *Verbindungsraum*

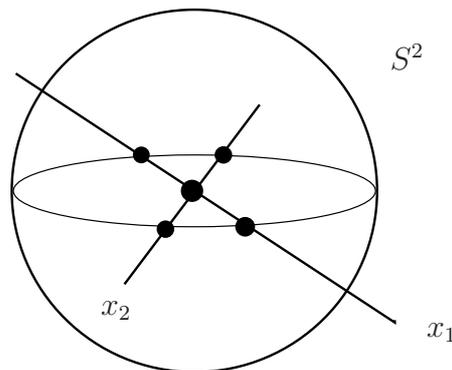
$$\mathbb{P}(U_1) \vee \mathbb{P}(U_2) := \mathbb{P}(U_1 + U_2)$$

zweier projektiver Unterräume von  $\mathbb{P}(V)$  wieder ein projektiver Unterraum, und zwar der kleinste, der sowohl  $\mathbb{P}(U_1)$  als auch  $\mathbb{P}(U_2)$  enthält.

- 10.3 Beispiel** 1. Modulo Antipodenidentifikation entsprechen die eindimensionalen projektiven Unterräume von  $\mathbb{P}_2(\mathbb{R})$  den Großkreisen auf  $S^2$ . Zwei verschiedene solche Großkreise schneiden sich genau in einem Antipodenpaar, sodass sich die projektiven Geraden des  $\mathbb{P}_2(\mathbb{R})$  in genau einem Punkt schneiden.



2. Den Punkten  $U \subset \mathbb{R}^3$  des  $\mathbb{P}_2(\mathbb{R})$  entsprechen die Antipodenpaare  $U \cap S^2$ . Der Verbindungsraum von  $U_1 \neq U_2 \in \mathbb{P}_2(\mathbb{R})$  entspricht also dem eindeutigen Großkreis auf  $S^2$ , der durch die Antipodenpaare  $U_1 \cap S^2$  und  $U_2 \cap S^2$  geht.



Während sich zwei affine Geraden in der Ebene nicht schneiden, wenn sie parallel sind, schneiden sich, wie wir gesehen haben, projektive Geraden in der projektiven Ebene  $\mathbb{P}_2(\mathbb{R})$  immer, man muss also keine Fallunterscheidung vornehmen.

Allgemein gilt:

**10.4 Satz** Sind  $U_1, U_2$  Unterräume des endlich-dimensionalen Vektorraumes  $V$ , dann gilt

$$\dim(\mathbb{P}(U_1) \vee \mathbb{P}(U_2)) + \dim(\mathbb{P}(U_1) \cap \mathbb{P}(U_2)) = \dim(\mathbb{P}(U_1)) + \dim(\mathbb{P}(U_2)).$$

**Bew.:** Es gilt  $\text{Dim}(U_1 + U_2) = \text{Dim}(U_1) + \text{Dim}(U_2) - \text{Dim}(U_1 \cap U_2)$ , woraus sich mit  $\dim(\mathbb{P}(U)) = \text{Dim}(U) - 1$  die Formel ergibt.  $\square$

Auch die leere Menge ist ein projektiver Unterraum von  $\mathbb{P}(V)$ , denn  $\phi = \mathbb{P}(\{0\})$ ; es ist  $\dim(\phi) = -1$ . Wir folgern:

**10.5 Korollar** Ist  $\dim(\mathbb{P}(U_1)) + \dim(\mathbb{P}(U_2)) \geq \dim(\mathbb{P}(V))$ , dann ist

$$\mathbb{P}(U_1) \cap \mathbb{P}(U_2) \neq \emptyset.$$

**Bew.:**  $\dim(\mathbb{P}(U_1) \cap \mathbb{P}(U_2)) = \dim(\mathbb{P}(U_1)) + \dim(\mathbb{P}(U_2)) - \dim(\mathbb{P}(U_1) \vee \mathbb{P}(U_2)) \geq 0.$  □

Nachdem wir nun verschiedene Aspekte projektiver Räume studiert haben, sollen jetzt Abbildungen zwischen ihnen eingeführt werden, die "die projektive Struktur erhalten". Es ist aber zunächst nicht ganz klar, was dies bedeutet. Denn während auf Gruppen, Ringen, Körpern oder Vektorräumen Verknüpfungen erklärt sind, können wir Punkte eines projektiven Raumes weder addieren noch mit einem Körperelement multiplizieren.

Der  $\mathbb{P}(V)$  wurde nur als *Menge* eingeführt, die Menge der eindimensionalen Unterräume von  $V$ . Als Automorphismen von  $\mathbb{P}(V)$  kommen daher zunächst alle bijektiven Abbildungen in Frage.

Tatsächlich besitzen projektive Räume  $\mathbb{P}(V)$  doch mehr Struktur, denn einerseits entstehen sie durch Äquivalenzklassenbildung auf  $V \setminus \{0\}$  aus dem Vektorraum  $V$  (mit  $x \sim y$ , falls  $\text{span}(x) = \text{span}(y)$ ), andererseits besitzen sie mit den projektiven Unterräumen ausgezeichnete Teilmengen.

Wir können also als strukturerehaltende Abbildungen

$$f : \mathbb{P}(V) \rightarrow \mathbb{P}(W)$$

diejenigen  $f$  wählen, die

1. von einer linearen Abbildung  $F : V \rightarrow W$  abstammen
2. oder die projektive Unterräume auf projektive Unterräume abbilden.

Ist  $\dim(\mathbb{P}(V)) = 1$ , dann erfüllt *jede* Bijektion  $f : \mathbb{P}(V) \rightarrow \mathbb{P}(V)$  die zweite Forderung, einfach, weil  $\mathbb{P}(V)$  außer sich nur die leere Menge als projektiven Unterraum besitzt. Man könnte also denken, dass die zweite Forderung von viel mehr Abbildungen als die erste erfüllt wird.

Dass dies für  $\dim(\mathbb{P}(V)) \geq 2$  und insbesondere  $K = \mathbb{R}$  nicht der Fall ist, ist Inhalt des so genannten *Hauptsatzes der projektiven Geometrie*.<sup>7</sup>

Wir werden jetzt die Abbildungen  $f$  untersuchen, die der (einfacheren) Bedingung 1. genügen:

---

<sup>7</sup>Allerdings ist Forderung 2. z.B. im Fall des Körpers  $K = \mathbb{C}$  auch für  $\dim(\mathbb{P}(V)) \geq 2$  schwächer als Forderung 1., denn gilt für  $f : V \rightarrow V$  die Eigenschaft

$$f(v_1 + v_2) = f(v_1) + f(v_2) \quad \text{aber} \quad f(kv) = \bar{k}f(v),$$

dann ist  $f$  zwar nicht linear, aber mit  $U \subset V$  ist auch  $f(U) \subset V$  ein Unterraum.

**10.6 Definition** Es seien  $V$  und  $W$   $K$ -Vektorräume. Dann heißt  $f : \mathbb{P}(V) \rightarrow \mathbb{P}(W)$  projektiv, wenn eine lineare Abbildung  $F : V \rightarrow W$  existiert, für die

$$f(\text{span}(v)) = F(\text{span}(v)) \quad (v \in V \setminus \{0\})$$

gilt, und man schreibt dann  $f = \mathbb{P}(F)$ .

Ist  $f$  zusätzlich bijektiv, dann heißt  $f$  Projektivität.

Um die Wirkung einer projektiven Abbildung besser zu verstehen, ist es nützlich, *Koordinaten* auf  $\mathbb{P}(V)$  einzuführen. Ist  $\dim \mathbb{P}(V) = u$ , dann erwarten wir, einen Punkt von  $\mathbb{P}(V)$  durch Angabe von  $n$  Koordinaten, d.h. Zahlen aus  $K$ , beschreiben zu können, wie wir dies von  $n$ -dimensionalen  $K$ -Vektorräumen gewohnt sind. dies ist auch möglich, allerdings nur *lokal*.

**10.7 Beispiel** Die Punkte von  $\mathbb{P}_1(\mathbb{R})$  entsprechen den Unterräumen  $\text{span}(x) \subset \mathbb{R}^2$  für  $x = (x_0, x_1) \in \mathbb{R}^2 \setminus \{0\}$ . Ist  $k \neq 0$ , dann ist  $\text{span}(y) = \text{span}(k)$  für  $y \equiv (y_0, y_1) := (kx_0, kx_1)$ .

Ist  $x_0 \neq 0$ , dann können wir  $x_1/x_0 \in \mathbb{R}$  als lokale Koordinate in  $\mathbb{P}_1(\mathbb{R})$  benutzen, denn  $y_1/y_0 = x_1/x_0$ , d.h. der Quotient hängt nicht vom Repräsentanten von  $\text{span}(x)$  ab.

Ist dagegen  $x_0 = 0$ , dann ist wegen  $x \neq 0$  der Quotient  $x_0/x_1 \in \mathbb{R}$  wohldefiniert, und wir nutzen diesen als lokale Koordinate auf  $\mathbb{P}_1(\mathbb{R}) \cong S^1$ .

Um diese Fallunterscheidungen zu vermeiden, werden zur Beschreibung des projektiven Raums  $\mathbb{P}_n(K) \equiv \mathbb{P}(K^{n+1})$  oft so genannte *homogene Koordinaten*

$$(x_0 : x_1 : \dots : x_n) \quad \text{benutzt mit} \quad (x_0, \dots, x_n) \in K^{n+1} \setminus \{0\},$$

und der Verabredung, dass  $(x_0 : x_1 : \dots : x_n) = (y_0 : y_1 : \dots : y_n)$  genau dann gilt, wenn eine Zahl  $k \in K \setminus \{0\}$  mit  $y_i = kx_i$   $i = 0, \dots, n$  existiert.

Dies sind zwar keine Koordinaten im üblichen Sinn, aber wir können sie zur Konstruktion lokaler Koordinaten verwenden: Wegen  $x \neq 0$  existiert ein  $i$  mit  $x_i \neq 0$ , und das  $n$ -Tupel  $(x_0/x_i, x_1/x_i, \dots, x_{i-1}/x_i, x_{i+1}/x_i, \dots, x_n/x_i) \in K^n$  ist repräsentantenunabhängig.

Ist nun  $F \in L(K^{n+1})$  von der Form  $F(x) = Ax$ ,  $A \in \mathcal{M}(n+1, K)$ , dann besitzt  $z := f(x)$  die homogenen Koordinaten

$$(z_0 : \dots : z_n) = (a_{00}x_0 + a_{01}x_1 + \dots + a_{0n}x_n : \dots : a_{n0}x_0 + \dots + a_{n1}x_n)$$

(wobei  $a_{kl} := (A)_{kl}$  ist).

Ist nun z.B.  $x_i \neq 0$  und  $z_k \neq 0$ , dann ergibt sich die (lokale) Koordinatentransformation

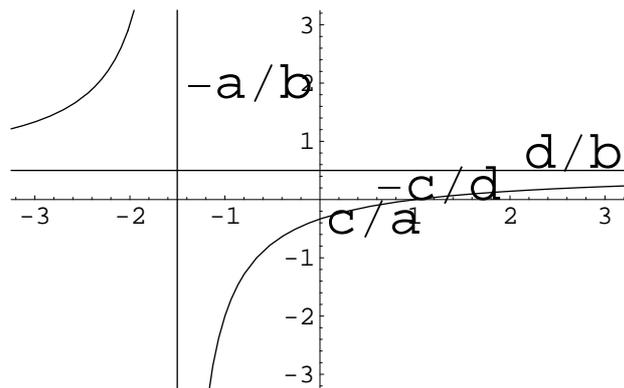
$$\frac{z_l}{z_k} = \frac{a_{li} + \sum_{j \neq i} a_{lj}(x_j/x_i)}{a_{ki} + \sum_{j \neq i} a_{kj}(x_j/x_i)} \quad (l \neq k).$$

Man spricht hier von *gebrochen linearen Transformationen*.

**10.8 Beispiel** Ist  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}(2, \mathbb{R})$ , dann ist  $f$  lokal für  $x_0 \neq 0$  und  $z_0 \neq 0$ , d.h.  $ax_0 + bx_1 \neq 0$  in den Koordinaten  $\tilde{z} := z_1/z_0, \tilde{x} := x_1/x_0$  durch

$$\tilde{z} = \frac{c + d\tilde{x}}{a + b\tilde{x}}$$

gegeben.



Dies ist genau eine affine Abbildung, wenn  $b = 0$  ist, also  $\tilde{z} = (c/a) + (d/a)\tilde{x}$ .

Allgemein bildet eine Projektivität  $f : \mathbb{P}_n(k) \rightarrow \mathbb{P}_n(k)$  den in der Form (10.1) in den projektiven Raum eingebetteten Vektorraum  $K^n \subset \mathbb{P}_n(K)$  nicht in sich ab. Das ist genau dann der Fall, wenn  $f = \mathbb{P}(F)$  mit  $F(x) = Ax$  mit  $(A)_{0k} = 0$ ,  $k = 1, \dots, n$  ist, und in diesem Fall ist  $f|_{K^n} : K^n \rightarrow K^n$  eine Affinität.

## Literatur

- [1] Brieskorn, E.: Lineare Algebra und analytische Geometrie. Vieweg 1982
- [2] Fischer, G.: Lineare Algebra. Vieweg 1995
- [3] Knörrer, H.: Geometrie. Vieweg 1996
- [4] Kostrikin, A.I.; Manin, Y.I.: Linear algebra and geometry. Gordon & Breach, 1997
- [5] Kowalsky, H.-J.: Einführung in die Lineare Algebra. de Gruyter 1971
- [6] Waerden, B.L. van der: Algebra Bd 1. Springer 1971
- [7] Walter, R.: Einführung in die lineare Algebra. Vieweg 1982
- [8] Walter, R.: Lineare Algebra und analytische Geometrie. Vieweg 1985

## Index

- Abbildung
  - adjungierte 60
  - affine 76
  - normale 62
  - transponierte 11
- Abtast-Theorem 11
- Affinität 76
- ähnlich 24
- algebraisch abgeschlossen 24
- Anfangswertproblem 51
- antiselbstadjungiert 71
- Blockdiagonalform 19
- Boysche Fläche 92
- Cauchy–Schwarz–Ungleichung 57
- Code
  - dualer 15
  - Hamming- 16
  - Paritäts- 7
  - perfekter 16
  - Simplex- 16
- darstellende Matrix 22
- diagonalisierbar 21, 25
- Differentialgleichung 31, 51
- Dimension 82
- direkte Summe 19
  - äußere 18
  - innere 17
- direktes Produkt 19
- Doppelgerade 79
- duale Basis 9
- Eigenraum 21
  - verallgemeinerter 33
- Eigenvektor 21, 63
  - verallgemeinerter 33
- Ellipse 80
- Euklidische Norm 51
- Faktorraum 5
- Fittingindex 34
- Fouriertransformation 69
- Generatormatrix 6
- Gram-Schmidt-Verfahren 59
- Gruppe
  - der Affinitäten 76
  - orthogonale 66
  - speziell orthogonale 66
  - speziell unitäre 66
  - unitäre 66
- Hauptsatz der projektiven Geometrie 86
- Hyperbel 80
- Hyperboloid
  - einschaliges 82
  - zweischaliges 82
- Imaginärteil 45
- Isometrie 65
- Jordanbasis 41
- Jordanmatrix 40
- Jordanzerlegung 37
- kanonischer Homomorphismus 9
- Kegel 82
- komplexe Struktur 44
- Komplexifizierung 45, 47
- Kongruenz 77
- Konjugation 46
- konjugiert-linear 61
- Kontrollmatrix 7
- Koordinaten, homogene 87
- lineare Differentialgleichung 31
- Linearform 8
- Matrizen
  - ähnliche 24
  - äquivalente 22
  - nilpotent 25
  - Nilpotenzindex 25
  - Normalform 22
  - Operatornorm 51
  - orthogonal 65

Orthogonalbasis 58  
orthogonales Komplement 12  
Orthonormalsystem 58  
Parabel 80  
Paraboloid 82  
Partition 31  
Polarisationsidentität 57  
Projektion 20, 21  
    Orthogonal- 58  
projektiv 87  
projektiver Raum 82  
Projektivität 87  
Quadrik 75  
Quotientenraum 5  
Rang 15  
Realteil 45  
Reellifizierung 43  
Sattel 82  
Schrödingeroperator 74  
selbstadjungiert 71  
Skalarerweiterung 45  
Summe 17  
transponierte Abbildung 11  
unitär 65  
Vektor  
    erzeugender 27  
    primitiver 29  
Verbindungsraum 84  
zyklisch 27  
Zylinder 81



**Abb.:** Boysche Fläche (Immersion des  $\mathbb{P}_2(\mathbb{R})$  in  $\mathbb{R}^3$ ), Mathematisches Forschungsinstitut Oberwolfach.<sup>8</sup>

---

<sup>8</sup>Weitere Informationen unter <http://www.mfo.de/boysurface/boysurface.html>