

# Lineare Algebra

**Tammo tom Dieck**

Mathematisches Institut  
Universität Göttingen

Version vom 7. Oktober 2014

# Inhaltsverzeichnis

<b>1</b>	<b>Vektorräume und lineare Abbildungen</b>	<b>4</b>
1.1	Vorbereitungen . . . . .	4
1.2	Vektorräume . . . . .	10
1.3	Basis. Dimension . . . . .	13
1.4	Lineare Abbildungen . . . . .	19
1.5	Summen und Produkte . . . . .	24
1.6	Quotientenräume . . . . .	28
<b>2</b>	<b>Matrizen</b>	<b>32</b>
2.1	Matrizen . . . . .	32
2.2	Spezielle Matrizen . . . . .	38
2.3	Basiswechsel. Koordinatentransformation . . . . .	41
2.4	Lineare Gleichungen . . . . .	47
<b>3</b>	<b>Determinanten</b>	<b>54</b>
3.1	Determinantenformen . . . . .	54
3.2	Determinante einer Matrix . . . . .	57
3.3	Volumen. Orientierung . . . . .	61
3.4	Das charakteristische Polynom . . . . .	63
<b>4</b>	<b>Bilineare Abbildungen</b>	<b>66</b>
4.1	Hilbert-Räume . . . . .	66
4.2	Isometrien . . . . .	73
4.3	Selbstadjungierte Endomorphismen . . . . .	79
4.4	Bilineare Abbildungen . . . . .	83
4.5	Adjungierte Endomorphismen . . . . .	86
4.6	Geometrische Bilinearformen . . . . .	87
4.7	Bilinearformen auf reellen Vektorräumen . . . . .	89
4.8	Die Lorentz-Gruppe . . . . .	92
<b>5</b>	<b>Affine und projektive Räume</b>	<b>101</b>
5.1	Affine Räume . . . . .	101
5.2	Projektive Räume . . . . .	106
5.3	Projektivitäten . . . . .	111

<b>6 Normalformen</b>	116
6.1 Zyklische Endomorphismen . . . . .	116
6.2 Nilpotente Endomorphismen . . . . .	117
6.3 Jordansche Normalform . . . . .	119
6.4 Normalformen für Zwei-mal-zwei Matrizen . . . . .	120
<b>7 Beispiele und Anwendungen</b>	124
7.1 Vektorprodukt . . . . .	124
7.2 Quaternionen . . . . .	128
7.3 Quaternionen und orthogonale Gruppen . . . . .	131
7.4 Quadriken . . . . .	135
7.5 Quadratische Formen . . . . .	140
7.6 Reelle und komplexe Strukturen . . . . .	142
<b>8 Anhang</b>	148
8.1 Die Mengensprache . . . . .	148
<b>Index</b>	152

# Kapitel 1

## Vektorräume und lineare Abbildungen

### 1.1 Vorbereitungen

Ein Grundbegriff der linearen Algebra ist der Begriff eines Vektorraums über einem Körper. Um ihn zu formulieren, verwenden wir den Begriff einer Gruppe und den eines Körpers. Obgleich es sich dabei um wichtige Begriffe der Mathematik handelt, mit eigenen umfangreichen Theorien, werden in diesem einleitenden Abschnitt lediglich die Vokabeln bereitgestellt.

**(1.1.1) Zahlen.** Die Algebra ist aus dem Zahlenrechnen und seinen Gesetzen entstanden. Eine elementare Kenntnis der Zahlen wird unterstellt.

Einige Mengen von Zahlen werden im vorliegenden Text mit Standardsymbolen notiert:

$\mathbb{N}$	=	$\{1, 2, 3, \dots\}$	Menge der natürlichen Zahlen
$\mathbb{N}_0$	=	$\{0, 1, 2, \dots\}$	Menge der natürlichen Zahlen einschließlich der Null
$\mathbb{Z}$	=	$\{0, \pm 1, \pm 2, \dots\}$	Menge der ganzen Zahlen
$\mathbb{Q}$	=		Menge der rationalen Zahlen (Brüche)
$\mathbb{R}$	=		Menge der reellen Zahlen
$\mathbb{C}$	=		Menge der komplexen Zahlen.

Die historische Entwicklung des Zahlbegriffs und seine axiomatische Grundlegung sind es jedoch wert, genauer angeschaut zu werden. Dazu sei ein Blick in das reichhaltige und anregende Buch Zahlen<sup>1</sup> empfohlen.  $\diamond$

---

<sup>1</sup>Ebbinghaus et. al: Zahlen (Grundwissen Mathematik 1). Springer-Verlag, Berlin Heidelberg 1983.

**(1.1.2) Komplexe Zahlen.** Wir notieren noch einige Einzelheiten über die komplexen Zahlen. Für die gesamte Mathematik ist bekanntlich die Erweiterung des Bereichs der reellen Zahlen zu den komplexen Zahlen von grundlegender Bedeutung.

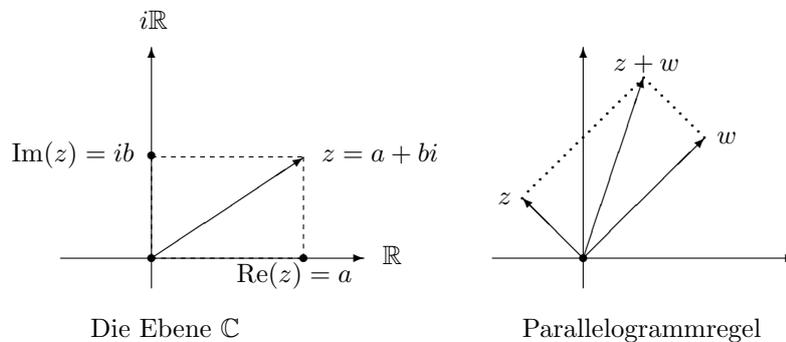
Keine reelle Zahl  $x$  genügt der Gleichung  $x^2 + 1 = 0$ . Man erfindet eine neue Zahl  $i$ , für die  $i^2 = -1$  gilt, bildet mit reellen Zahlen  $a$  und  $b$  Ausdrücke  $a + bi$ , addiert zwei solche gemäß der Regel

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

und multipliziert sie gemäß der Regel

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Man rechnet also einfach wie gewohnt, wobei jedoch immer wenn  $i^2$  vorkommt, dafür  $-1$  gesetzt wird. Man nennt  $z = a + bi$  eine *komplexe Zahl*,  $a = \operatorname{Re}(z)$  ihren *Realteil*,  $b = \operatorname{Im}(z)$  ihren *Imaginärteil* und  $i$  die *imaginäre Einheit*. Die Gesamtheit der komplexen Zahlen wird mit  $\mathbb{C}$  bezeichnet. Geometrisch wird  $z = a + bi$  als Punkt  $(a, b) \in \mathbb{R}^2$  der Zahlenebene veranschaulicht. Geometrisch wird die Addition durch die berühmte Parallelogrammregel dargestellt. Die Zahl  $\bar{z} = a - bi$  heißt die zu  $z = a + bi$  *konjugierte Zahl* (Spiegelung an der reellen Achse). Es gelten die Regeln  $\overline{u + v} = \bar{u} + \bar{v}$  und  $\overline{u \cdot v} = \bar{u} \cdot \bar{v}$ . Wegen  $z \cdot \bar{z} = (a + bi)(a - bi) = a^2 + b^2$  setzen wir  $z \cdot \bar{z} = |z|^2$  und nennen  $|z| \geq 0$  die *Norm* oder den *Betrag* von  $z$ . Diese Zahl ist der elementar-geometrische Abstand von  $z$  vom Nullpunkt. Es gelten die *Dreiecksungleichung*  $|u + v| \leq |u| + |v|$  sowie  $|uv| = |u||v|$ .  $\diamond$

Die Ebene  $\mathbb{C}$ 

Parallelogrammregel

Für die vorstehend genannten Zahlenmengen verfügt man jeweils über die Addition und die Multiplikation. Die zugehörigen Rechenregeln formulieren wir nun ganz allgemein.

**(1.1.3) Verknüpfungen.** Eine *Verknüpfung* auf einer Menge  $G$  ist eine Abbildung  $m: G \times G \rightarrow G$ . Eine Verknüpfung  $m$  auf  $G$  heißt *assoziativ*, wenn für

je drei Elemente  $a, b, c$  aus  $G$  die Gleichheit  $m(a, m(b, c)) = m(m(a, b), c)$  besteht. Formeln dieser Art lassen sich schlecht lesen. Wenn wir eine Verknüpfung abgekürzt als Multiplikation  $m(a, b) = a \cdot b = ab$  notieren, so lautet also das Assoziativgesetz  $a(bc) = (ab)c$ . Statt „verknüpfen“ sagen wir bei dieser Bezeichnungsweise auch „multiplizieren“. Eine Verknüpfung  $m$  heißt *kommutativ*, wenn für je zwei Elemente  $a, b$  aus  $G$  die Gleichheit  $m(a, b) = ab = ba = m(b, a)$  besteht. Ein Element  $e \in G$  heißt *neutrales Element* der Verknüpfung  $m$ , wenn für alle  $x \in G$  die Gleichungen  $m(e, x) = ex = x = xe = m(x, e)$  gelten. Sind  $e$  und  $f$  neutral, so folgt  $e = ef = f$ . Also hat eine Verknüpfung höchstens ein neutrales Element.  $\diamond$

**(1.1.4) Beispiel.** Sei  $G$  eine der sechs in (1.1.1) genannten Mengen. Dann ist jeweils die Addition und die Multiplikation eine Verknüpfung auf dieser Menge, und diese Verknüpfungen sind assoziativ und kommutativ.

Auf der Menge  $\mathbb{N} = \{1, 2, 3, 4, \dots\}$  hat die Multiplikation ein neutrales Element, die Eins; die Addition hat jedoch kein neutrales Element. Wie verhalten sich diesbezüglich Addition und Multiplikation auf der Menge  $\mathbb{N}_0$ ?  $\diamond$

**(1.1.5) Gruppen.** Eine *Gruppe* besteht aus einer Menge  $G$  und einer Verknüpfung  $m$  auf  $G$ , so daß die folgenden Axiome gelten:

- (1) Die Verknüpfung ist assoziativ.
- (2) Die Verknüpfung besitzt ein neutrales Element  $e \in G$ .
- (3) Zu jedem  $a \in G$  gibt es ein  $b \in G$ , so daß  $m(b, a) = e = m(a, b)$  ist.

Ist  $(G, m)$  eine Gruppe, so heißt  $m$  eine *Gruppenstruktur* auf der Menge  $G$ . Wenn die Verknüpfung als Multiplikation geschrieben wird, so nennen wir  $m$  auch *Gruppenmultiplikation*. Eine Gruppe mit einer kommutativen Multiplikation wird *kommutative Gruppe* oder *abelsche<sup>2</sup> Gruppe* genannt. In einer kommutativen Gruppe wird die Verknüpfung oft als Addition geschrieben,  $m(a, b) = a + b$ ; in diesem Fall sprechen wir natürlich von der *Gruppenaddition*. Wir unterscheiden diese beiden Notationen auch dadurch, daß wir von additiven und multiplikativen Gruppen reden. Es ist üblich, eine Gruppe  $(G, m)$  nur durch die zugrundeliegende Menge  $G$  zu bezeichnen, insbesondere dann, wenn man nicht mehrere Verknüpfungen unterscheiden muß.  $\diamond$

In einer multiplikativen Gruppe schreiben wir das neutrale Element oft als 1 und nennen es das *Einselement* der Gruppe. Das bei gegebenem  $a \in G$  eindeutig bestimmte Element  $b$  mit der Eigenschaft  $ab = 1 = ba$  nennen wir das *Inverse* von  $a$  und bezeichnen es mit  $a^{-1}$ . In einer additiv geschriebenen Gruppe ist das Inverse  $b$  von  $a$  durch die Gleichung  $a + b = 0$  bestimmt; wir schreiben in diesem Fall  $b = -a$ . Mit diesen Bezeichnungen gelten dann die Rechenregeln  $1 \cdot a = a = a \cdot 1$  und  $a^{-1} \cdot a = a \cdot a^{-1} = 1$ . Ferner ist  $(a^{-1})^{-1} = a$ , und es gilt  $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$ . Die Linksmultiplikation mit

<sup>2</sup>Niels Henrik Abel 1802 – 1829

$a \in G$  ist die Abbildung  $l(a): G \rightarrow G, x \mapsto ax$ . Die Gruppenaxiome besagen, daß  $l(a)l(b) = l(ab)$  und  $l(e) = \text{id}$  (die identische Abbildung) ist. Es folgt, daß  $l(a^{-1})$  die inverse Abbildung zu  $l(a)$  ist. Insbesondere ist  $l(a)$  immer bijektiv. Da  $x \mapsto ax$  bijektiv ist, so hat die Gleichung  $ax = b$  immer genau eine Lösung, nämlich  $x = a^{-1}b$ ; und aus  $ac = bc$  folgt  $a = b$ , das heißt man kann „kürzen“.

In einer additiv geschriebenen abelschen Gruppe bezeichnen wir das neutrale Element mit  $0$  und nennen es das *Nullelement* der Gruppe. Man verwendet die abkürzende Bezeichnung  $a + (-b) = a - b$ . Es gelten dann  $-(-a) = a$ ,  $-b + a = a - b$  und  $-(a - b) = -a + b$ .

Eine nichtleere Teilmenge  $U$  einer Gruppe  $G$  heißt *Untergruppe* von  $G$ , wenn aus  $x, y \in U$  immer  $xy \in U$  und  $x^{-1} \in U$  folgt. Dann ist nämlich offenbar die Einschränkung  $U \times U \rightarrow U, (x, y) \mapsto xy$  definiert und eine Gruppenstruktur auf  $U$ .

Ein *Homomorphismus* der Gruppe  $G$  in die Gruppe  $H$  ist eine Abbildung  $f: G \rightarrow H$ , die  $f(ab) = f(a)f(b)$  erfüllt; es folgt dann  $f(1) = f(1 \cdot 1) = f(1) \cdot f(1)$  also  $f(1) = 1$  und  $f(1) = f(a \cdot a^{-1}) = f(a) \cdot f(a^{-1})$  also  $f(a^{-1}) = (f(a))^{-1}$ . (Man sagt dann auch salopp, ein Homomorphismus „respektiert“ Einselemente und Inverse.) Ein bijektiver Homomorphismus heißt *Isomorphismus*. Die Gruppen  $G$  und  $H$  heißen *isomorph*, wenn es einen Isomorphismus  $f: G \rightarrow H$  gibt.

Wir verwenden die bekannte Schreibweise mit Potenzen in beliebigen multiplikativen Gruppen. Wir setzen also  $a^0 = 1, a^1 = a, a^2 = a \cdot a, a^{n+1} = a \cdot a^n, a^{-1} = (a^{-1})^n$  für  $n \in \mathbb{N}_0$ . Es gilt dann die Regel  $a^m \cdot a^n = a^{m+n}$  für beliebige  $m, n \in \mathbb{Z}$ . Das kann auch so formuliert werden: Die Zuordnung  $\mathbb{Z} \rightarrow G, n \mapsto a^n$  ist ein Homomorphismus der additiven Gruppe  $\mathbb{Z}$  in die multiplikative Gruppe  $G$ .

Der trockenen Definition (1.1.3) ist natürlich nicht anzusehen, daß der Gruppenbegriff einer der wichtigsten in der Mathematik ist. Wir werden im Verlauf dieses Textes etliche Gruppen kennenlernen. (Eine Aufgabe der Gruppen ist es übrigens, Symmetriestrukturen von mathematischen Objekten zu beschreiben, was wir aber an dieser Stelle nicht erläutern.)

### (1.1.6) Beispiele.

(1) Die ganzen Zahlen  $\mathbb{Z}$  sind zusammen mit der Addition eine kommutative Gruppe. Die Vielfachen  $\{cz \mid z \in \mathbb{Z}\}$  einer ganzen Zahl  $c$  sind eine Untergruppe darin.

(2) Die komplexen Zahlen vom Betrag 1 sind zusammen mit der Multiplikation eine kommutative Gruppe. In der geometrischen Interpretation der komplexen Zahlen als Punkte der Ebene besteht diese Gruppe aus den Punkten des Kreises vom Radius 1 um den Nullpunkt. In dieser Gruppe gibt es die Untergruppe  $C_n = \{\zeta \mid \zeta^n = 1\}$  der Zahlen, deren  $n$ -te Potenz gleich 1 ist und die deshalb  $n$ -te *Einheitswurzeln* heißen ( $n \in \mathbb{N}$ ). Mit Hilfe der komplexen Exponentialfunktion lassen sich diese Zahlen als  $\exp(2\pi ik/n), 0 \leq k < n$

darstellen. Zum Beispiel besteht  $C_4$  aus den Zahlen  $\pm 1, \pm i$ .

(3) Die Abbildung  $\mathbb{Z} \rightarrow C_n, k \mapsto \exp(2\pi ik/n)$  ist ein Homomorphismus der additiven Gruppe  $\mathbb{Z}$  auf die multiplikative Gruppe  $C_n$ .

(4) Die wesentliche Eigenschaft der berühmten Logarithmus-Funktion ist es, ein Isomorphismus zu sein, und zwar von der multiplikativen Gruppe der positiven reellen Zahlen auf die additive Gruppe aller reellen Zahlen („Verwandlung der Multiplikation in die Addition“).

In vielen mathematischen Kontexten gibt es den Begriff Isomorphismus. Isomorphe Objekte sind dann im wesentlichen dieselben Dinge („Dasselbe in Grün“). Unser Beispiel zeigt aber, daß inhaltlich doch ein beträchtlicher Unterschied bestehen kann und man nicht blindlings ein Objekt durch ein isomorphes ersetzen kann.

(5) Nun noch eine „abstrakte“ Gruppe. Sei  $X$  eine Menge. Die Menge  $S(X)$  der bijektiven Abbildungen  $X \rightarrow X$  ist mit der Verkettung von Abbildungen eine Gruppe, nicht kommutativ, wenn  $X$  mehr als 2 Elemente enthält.  $\diamond$

**(1.1.7) Körper.** Ein *Körper* ist eine Menge  $K$  zusammen mit zwei Verknüpfungen

$$\begin{aligned} A: K \times K &\rightarrow K, & (x, y) &\mapsto x + y \\ M: K \times K &\rightarrow K, & (x, y) &\mapsto xy = x \cdot y \end{aligned}$$

mit den Eigenschaften (Körperaxiome)

- (1)  $(K, A)$  ist eine abelsche Gruppe mit dem neutralen Element 0.
- (2)  $(K \setminus \{0\}, M)$  ist eine abelsche Gruppe mit dem neutralen Element 1.
- (3) Es gilt das *Distributivgesetz*  $a(b + c) = ab + ac$ .

(Wie üblich geht „Punktrechnung“ (auch ohne Punkt) vor „Strichrechnung“, zur Vermeidung von Klammern.) Die Verknüpfung  $A$  ist die *Addition* und die Verknüpfung  $M$  die *Multiplikation* des Körpers. Das Axiom (2) fordert unter anderem, daß das Produkt zweier von Null verschiedener Elemente wieder ungleich Null ist!

Einige einfache Folgerungen aus den Axiomen. Wegen der Kommutativität der Multiplikation gilt auch das Distributivgesetz  $(b + c)a = ba + ca$ . Aus den Gleichungen  $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$  folgt  $0 = 0 \cdot a$  und  $0 = a \cdot 0$ . Aus  $0 = a \cdot 0 = a \cdot (b + (-b)) = a \cdot b + a \cdot (-b)$  folgt  $a \cdot (-b) = -(a \cdot b)$ . Damit ergibt sich  $(-a) \cdot (-b) = -(a \cdot (-b)) = -(-ab) = ab$ . Wir haben damit aus den Körperaxiomen die Regel „Minus mal Minus gleich Plus“ hergeleitet.

Eine nichtleere Teilmenge  $L$  von  $K$  heißt *Unterkörper* von  $K$ , wenn sie mit  $a, b$  auch  $a + b$  und  $ab$  enthält und die dadurch bestimmten Verknüpfungen  $L \times L \rightarrow L, (a, b) \mapsto a + b, L \times L \rightarrow L, (a, b) \mapsto ab$  die Körperaxiome erfüllen. Aus  $a \in L$  folgt dann  $-a \in L$  und aus  $0 \neq b \in L$  folgt  $b^{-1} \in L$ . Das Einselement von  $K$  ist auch das Einselement von  $L$ .  $\diamond$

**(1.1.8) Beispiele.** Die rationalen Zahlen  $\mathbb{Q}$ , die reellen Zahlen  $\mathbb{R}$  und die komplexen Zahlen  $\mathbb{C}$ , jeweils mit der üblichen Addition und Multiplikation versehen, sind Körper. Für die ganzen Zahlen erfüllen die Addition und Multiplikation fast alle Körperaxiome, aber  $\pm 1$  sind die einzigen ganzen Zahlen, die ein Inverses bezüglich der Multiplikation haben.  $\diamond$

**(1.1.9) Beispiel.** Wir wollen zumindest einen etwas exotischen Körper beschreiben, weil er in der linearen Algebra (und auch sonst) oft aus der Rolle fällt. Ein Körper enthält die Elemente 0 und 1, und zwei so bezeichnete Elemente bilden mit den Rechenregeln, die sich aus den Axiomen ergeben, schon einen Körper. Natürlich ist dann  $1 + 1$  ein Körperelement; aus  $1 + 1 = 1$  würde  $1 = 0$  folgen, was in einem Körper nicht sein kann. Also ist  $1 + 1 = 0$ . Wollte man  $1 + 1$  mit 2 bezeichnen, so wäre  $2 = 0$ . Oh weh!, aber warum nicht?

Diesen Körper kann man sich auch aus dem Zahlenrechnen erklären. Bekanntlich zerfallen die ganzen Zahlen in gerade und ungerade Zahlen. Mit den Worten „gerade“ =  $g$  und „ungerade“ =  $u$  kann man rechnen, etwa  $g + u = u$  oder  $u \cdot u = u$ . Es entspricht dann  $g$  der Null und  $u$  der Eins.  $\diamond$

## Ergänzungen und Aufgaben

1. Gegeben sei eine Menge von Wandfarben. Eine Verknüpfung zweier Farben sei die Mischung von gleichen Mengen dieser beiden Farben. Diese Verknüpfung ist kommutativ aber nicht assoziativ. Man könnte eine Verknüpfung mit drei Variablen definieren, indem man gleiche Teile dieser drei Farben mischt. Und so weiter.

2. In einem vollständig formalen Aufbau der Mathematik würde man über das Assoziativgesetz beweisen müssen, daß jede sinnvolle Klammerung in einem Produkt von vielen Faktoren dasselbe Ergebnis liefert. (So gibt es für vier Faktoren 5 Klammerungen und für fünf Faktoren 14 Klammerungen. Ausprobieren!)

Ebenso müßte man für eine assoziative und kommutative Verknüpfung beweisen, daß das Produkt von vielen Faktoren von der Reihenfolge der Faktoren nicht abhängt. Wir wollen darüber an dieser Stelle nicht tiefsinnig werden.

3. Kann man von einer sinnvollen Klammerung von  $k$  Faktoren zu jeder anderen dadurch gelangen, daß man das Assoziativgesetz mehrmals anwendet?

4. Die formalen Rechenregeln, um die wir uns meist nicht kümmern, haben durchaus ihre Geheimnisse. Auf der Kugeloberfläche wähle man 14 Punkte geeignet und gebe ihnen als Namen die sinnvollen Klammerungen von 5 Faktoren. Man verbinde zwei Punkte durch eine sich selbst nicht kreuzende Linie, wenn sie auseinander durch das Assoziativgesetz hervorgehen. Man kann dann die Gesamtheit dieser Linien so zeichnen, daß sich je zwei nicht kreuzen. Zeichnung dazu?

5. Eine Untergruppe  $U$  der additiven Gruppe  $\mathbb{Z}$  hat die Form  $\{nc \mid n \in \mathbb{Z}\}$  für ein eindeutig bestimmtes  $c \in \mathbb{N}_0$ . Ist nämlich  $U \neq \{0\}$ , so enthält  $U$  positive Zahlen; sei  $c$  die kleinste davon. Ist  $k \in U$ , so dividieren wir  $k$  mit Rest  $0 \leq r < c$ , also  $k = qc + r$ . Dann liegt  $r = k - qc$  in  $U$  und  $r > 0$  würde der Minimalität von  $c$  wider sprechen.

6. Ist  $G$  eine Gruppe und  $a \in G$ , so bilden die  $n \in \mathbb{Z}$  mit  $a^n = 1$  eine Untergruppe  $U$  von  $\mathbb{Z}$ . Ist  $U \neq \{0\}$  und  $c \in U$  die kleinste Positive Zahl in  $U$ , so heißt  $c$  die *Ordnung*

des Elementes  $a$ . Es gilt  $a^m = a^n$  genau dann, wenn  $m - n$  durch die Ordnung  $c$  teilbar ist.

7. Die reellen Zahlen der Form  $a + b\sqrt{2}$  mit rationalen Zahlen  $a, b$  bilden einen Körper. Einzig fraglich ist die Existenz von multiplikativen Inversen;

$$\frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}$$

ist inverse zu  $a + b\sqrt{2} \neq 0$ . Der Nenner ist ungleich Null, weil 2 kein Quadrat in  $\mathbb{Q}$  ist.

8. Die Grundrechenarten in einem Körper sind Addition und Multiplikation. Die Subtraktion ist eine Konsequenz der Axiome. Durch ein von Null verschiedenes Körperelement kann man dividieren: Man könnte  $ab^{-1} = b^{-1}a$  mit  $\frac{a}{b}$  bezeichnen. Für die üblichen Zahlen sind wir daran gewöhnt, aber auch allgemein gelten die Regeln des Bruchrechnens

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \quad \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}.$$

Wegen  $a \cdot 0 = 0$  kann (darf!) man durch Null nicht dividieren.

## 1.2 Vektorräume

(1.2.1) **Vektorraum.** Ein *Vektorraum über einem Körper  $K$*  besteht aus einer Menge  $V$  und zwei Abbildungen

$$\begin{aligned} a: V \times V &\rightarrow V, & (v, w) &\mapsto a(v, w) = v + w \\ m: K \times V &\rightarrow V, & (\lambda, v) &\mapsto \lambda \cdot v = \lambda v \end{aligned}$$

mit den Eigenschaften (Vektorraumaxiome):

- (1)  $(V, a)$  ist eine additive Gruppe.
- (2) Für  $\lambda_1, \lambda_2 \in K$  und  $v \in V$  gilt  $\lambda_1(\lambda_2 v) = (\lambda_1 \lambda_2)v$ .
- (3) Für  $\lambda \in K$  und  $v_1, v_2 \in V$  gilt  $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$ .
- (4) Für  $\lambda_1, \lambda_2 \in K$  und  $v \in V$  gilt  $(\lambda_1 + \lambda_2)v = \lambda_1 v + \lambda_2 v$ .
- (5) Für  $1 \in K$  und  $v \in V$  gilt  $1 \cdot v = v$ .

Die Elemente aus  $V$  heißen *Vektoren*, die Elemente aus  $K$  (im Kontext der Vektorräume) *Skalare*. Die Abbildung  $a$  heißt *Addition von Vektoren*. Die Abbildung  $m$  heißt *Multiplikation mit Skalaren*. Die dritte und vierte Regel sind ein *Distributivgesetz*. Das Paar  $(a, m)$  von Abbildungen wird *Vektorraumstruktur* auf der Menge  $V$  genannt. Ein Vektorraum  $(V, a, m)$  wird meist nur durch die zugrundeliegende Menge  $V$  bezeichnet. Wir sprechen auch von einem  $K$ -Vektorraum, wenn ein Vektorraum über  $K$  gemeint ist. Der Hinweis auf den Körper wird weggelassen, wenn er aus dem Kontext erkenntlich ist oder wenn auf ihn nicht weiter hingewiesen werden muß. Ein Vektorraum über dem Körper

der reellen (komplexen) Zahlen wird auch *reeller (komplexer)* Vektorraum genannt. Das Paar  $(V, a)$  ist die *additive Gruppe* des Vektorraumes  $(V, a, m)$ .  $\diamond$

Der Körper  $K$  und die abelsche Gruppe  $(V, a)$  haben beide ihr eigenes Nullelement; geschrieben wird dafür aber beidemal  $0$ ; das ist ungenau, aber meistens bequem. Es gelten nämlich die Rechenregeln  $0 \cdot v = 0$  und  $\lambda \cdot 0 = 0$ . Die erste folgt so: Mit einem Distributivgesetz folgt  $0 \cdot v = (0+0) \cdot v = 0 \cdot v + 0 \cdot v$  und dann durch Kürzung in der Gruppe  $V$  die Behauptung. Ähnlich wird auf die zweite Regel geschlossen. Wir schreiben  $x - y$  für  $x + (-y)$ . Es gilt  $(-\lambda)x = \lambda(-x) = -(\lambda x)$  und speziell  $(-1)x = -x$ . Zum Beweis:  $\lambda x + (-\lambda x) = (\lambda + (-\lambda))x = 0 \cdot x = 0$ , und das bedeutet  $-(\lambda x) = (-\lambda)x$ . Analog folgt die Gleichung  $\lambda(-x) = -(\lambda x)$  mit dem anderen Distributivgesetz.

Wir erläutern die Definition eines Vektorraumes mit dem wichtigsten Beispiel:

**(1.2.2) Der Standardvektorraum.** Sei  $K$  ein Körper. Auf der Menge  $K^n$  der  $n$ -Tupel  $(x_1, \dots, x_n)$  von Elementen  $x_i \in K$  haben wir die Struktur einer abelschen Gruppe durch komponentenweise Addition

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

und eine Multiplikation mit Skalaren

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n).$$

Aus den Körperaxiomen werden die Vektorraumaxiome unmittelbar nachgewiesen. Dieser Vektorraum wird kurz und bündig „Der  $K^n$ “ gerufen. Unter  $K^0$  wollen wir den Nullraum verstehen.

Nach demselben Verfahren kann man noch größere Vektorräume herstellen. Sei  $K^\infty$  der Vektorraum, der aus allen Folgen  $(x_1, x_2, x_3, \dots)$ ,  $x_i \in K$ , besteht, die nur endlich viele von Null verschiedene Einträge haben, und sei  $K^\omega$  der Raum der aus allen Folgen dieser Art besteht (Addition und Skalarmultiplikation weiterhin komponentenweise).

Siehe zu diesen Definitionen auch (1.2.5).  $\diamond$

**(1.2.3) Unterraum.** Sei  $(V, a, m)$  ein Vektorraum über dem Körper  $K$ . Sei  $U \subset V$  eine nichtleere Teilmenge mit den Eigenschaften  $a(U \times U) \subset U$  und  $m(K \times U) \subset U$ .

$$\begin{array}{lll} U \times U \rightarrow U, & (x, y) \mapsto a(x, y) = x + y & \text{Addition} \\ K \times U \rightarrow U, & (\lambda, x) \mapsto m(\lambda, x) = \lambda x & \text{Skalarmultiplikation} \end{array}$$

auf  $U$  eine Vektorraumstruktur definiert, wie eine kurze Überlegung zeigt. Eine Teilmenge zusammen mit dieser Struktur heißt *Untervektorraum* oder kurz *Unterraum* von  $V$ . Wir sagen in diesem Fall, die Vektorraumstruktur auf  $U$

entstehe durch Einschränkung der auf  $V$  gegebenen Struktur, das heißt man verwendet immer noch die gleiche Addition und Skalarmultiplikation, aber nur für die Vektoren in  $U$ .  $\diamond$

#### (1.2.4) Beispiele.

(1) Jeder Vektorraum  $V$  hat den nur aus dem *Nullvektor* bestehenden Unterraum, den sogenannten *Nullraum* oder *trivialen Unterraum*  $\{0\}$ , der auch einfach mit  $0$  bezeichnet wird.

(2) Jeder Vektorraum ist ein Unterraum von sich selbst.

(3) Im Standardraum  $\mathbb{R}^3$  sind die Geraden und die Ebenen durch den Nullpunkt Unterräume, und es gibt keine weiteren außer  $0$  und  $\mathbb{R}^3$ , wie wir alsbald sehen werden.

(4) Im Vektorraum  $\mathbb{R}^\omega$  aller reellen Folgen bilden die konvergenten Folgen einen Unterraum. Auch  $\mathbb{R}^\infty$  ist ein Unterraum.  $\diamond$

#### (1.2.5) Beispiele (Funktionsräume).

(1) Sei  $A$  eine beliebige Menge und  $\text{Abb}(A, K)$  die Menge aller Abbildungen von  $A$  in den Körper  $K$ . Auf dieser Menge wird die Struktur eines  $K$ -Vektorraumes durch die Addition  $(f + g)(x) = f(x) + g(x)$  und die Skalarmultiplikation  $(\lambda f)(x) = \lambda(f(x))$  definiert.

Indem man Abbildungen mit geeigneten weiteren Eigenschaften betrachtet, erhält man zahlreiche Untervektorräume.

(2) Der Unterraum  $\text{Abb}_0(A, K)$  besteht aus allen Funktionen  $f: A \rightarrow K$ , die nur endlich viele von Null verschiedene Werte annehmen.

(3) Ein  $n$ -Tupel  $(x_1, \dots, x_n)$  von Elementen  $x_i \in K$  ist eigentlich eine Funktion  $x: \{1, \dots, n\} \rightarrow K$ , und zwar ist  $x_i$  der Wert an der Stelle  $i$ . Also ist  $K^n$  in diesem Sinne ein Funktionsraum.

(4) Ebenso können wir  $K^\infty$  als  $\text{Abb}_0(\mathbb{N}, K)$  und  $K^\omega$  als  $\text{Abb}(\mathbb{N}, K)$  auffassen.

(5) Im Fall  $A = \mathbb{R}$  und  $K = \mathbb{R}$  haben wir den Vektorraum  $C^k(\mathbb{R}, \mathbb{R})$  der  $k$ -mal stetig differenzierbaren Funktionen  $\mathbb{R} \rightarrow \mathbb{R}$  oder den Raum  $C([0, 1], \mathbb{R})$  der stetigen Funktionen  $[0, 1] \rightarrow \mathbb{R}$ . Vektorräume dieser Art spielen in der Analysis eine große Rolle.

(6) Die Gesamtheit  $P_n(\mathbb{R})$  aller Polynomfunktionen

$$x \mapsto a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

mit Koeffizienten  $a_i \in \mathbb{R}$  ist ein Unterraum von  $C^k(\mathbb{R}, \mathbb{R})$ . Er wird Vektorraum aller Polynome vom Grad kleinergleich  $n$  genannt.

(7) Noch etwas allgemeiner kann man auf der Menge  $\text{Abb}(A, V)$  der Abbildungen von  $A$  in einen Vektorraum  $V$  eine Vektorraumstruktur wie unter (1) definieren.  $\diamond$

## Ergänzungen und Aufgaben

1. Für den Standardraum etwa über den reellen Zahlen braucht man natürlich keine Gelehrsamkeit. Ein Vektor ist dann einfach ein Datensystem von Zahlen und man rechnet mit diesen Systemen in bekannter Weise. Betrachten wir Kochrezepte. Ein Rezeptvektor ordnet jeder Zutat ihre Mengenangabe zu. Will man für zwei Rezepte einkaufen, so muß man die Zutatenvektoren addieren. Hat man ein Rezept für zwei Personen, so muß man für sechs Personen den Zutatenvektor mit dem Skalar 3 multiplizieren.
2. Polynomfunktionen  $x \mapsto a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$  können wir über jedem Körper  $K$  betrachten, die Koeffizienten  $a_i$  sind dabei Elemente aus  $K$ . Aber ach!, für den Körper mit zwei Elementen ist  $x \mapsto x + x^2$  die Nullfunktion. Das bedeutet: Eigentlich (!) ist ein Polynom keine Funktion sondern ein formaler Ausdruck („Buchstabenrechnen“), bestimmt durch die Koeffizienten  $a_i$ , und die Potenzen von  $x$  sind zwar nützlich, aber nur Platzhalter und gewissermaßen überflüssig.
3. **Schnittträume.** Seien  $U_1$  und  $U_2$  Unterräume eines Vektorraums  $V$ . Dann ist  $U_1 \cap U_2$  ein Unterraum von  $V$ . Allgemein ist der Schnitt  $\bigcap_{j \in J} U_j$  einer beliebigen Familie  $(U_j \mid j \in J)$  von Unterräumen wieder ein Unterraum.
4. Sei  $K$  ein Unterkörper des Körpers  $L$ . Dann ist  $L$  mit der Körperaddition und der Skalarmultiplikation  $K \times L \rightarrow L, (\lambda, \mu) \mapsto \lambda\mu$  ein  $K$ -Vektorraum. Zum Beispiel ist  $\mathbb{C}$  ein  $\mathbb{R}$ -Vektorraum und  $\mathbb{R}$  ein  $\mathbb{Q}$ -Vektorraum. Jeder  $L$ -Vektorraum wird in derselben Weise durch Einschränkung der Skalarmultiplikation auf  $K$  ein  $K$ -Vektorraum. Der Fall  $K = L$  ist erlaubt. Dann haben wir aber den Standardvektorraum  $K^1$  vorliegen.
5. Sei  $V$  ein  $\mathbb{C}$ -Vektorraum. Wir erhalten auf der Menge  $V$  eine neue Vektorraumstruktur, indem wir zwar dieselbe Addition verwenden aber die neue Skalarmultiplikation  $\mathbb{C} \times V \rightarrow V, (z, v) \mapsto \bar{z}v$  (Multiplikation mit der konjugiert-komplexen Zahl). Wir nennen diesen Vektorraum den zu  $V$  *konjugierten* Vektorraum und bezeichnen ihn mit  $\bar{V}$ .

## 1.3 Basis. Dimension

Aus dem Vektorraumbegriff werden einige Begriffe abgeleitet, die für das weitere Arbeiten von grundsätzlicher Bedeutung sind. Die wichtigsten sind die in der Überschrift genannten. In diesem Abschnitt sei  $V$  ein Vektorraum über dem Körper  $K$ .

**(1.3.1) Lineare Hülle.** Sei  $S \subset V$  eine beliebige Teilmenge. Der Schnitt aller Unterräume von  $V$ , die  $S$  enthalten, heie *lineare Hülle* von  $S$  und werde mit  $L(S)$  bezeichnet. Das ist ein Unterraum. Gilt  $U = L(S)$ , so sagen wir:  $S$  *erzeugt*  $U$ ,  $S$  *spannt*  $U$  *auf*,  $S$  ist ein *Erzeugendensystem* von  $U$ . Der Vektorraum  $L(S)$  ist der (bezüglich Mengeneinklusion) kleinste Unterraum, der  $S$  enthält. Wir

beschreiben diesen Unterraum sogleich in (1.3.3) in weniger formaler Weise.  $\diamond$

Ist  $S = \emptyset$ , so ist  $L(S) = \{0\}$  der Nullraum. Genau dann ist  $S$  selbst ein Unterraum, wenn  $L(S) = S$  ist.

**(1.3.2) Linearkombination.** Ist  $(v_j \mid j \in J)$  eine endliche Familie von Vektoren und  $(\lambda_j \mid j \in J)$  eine Familie von Skalaren, so heißt die Summe

$$v = \sum_{j \in J} \lambda_j v_j$$

eine *Linearkombination* der Familie der  $v_j$ . Wir sagen von der Summe auch,  $v$  sei durch die Familie der  $v_j$  *linear dargestellt* und nennen die Summe eine *lineare Darstellung* von  $v$  durch die  $v_j$ . Ist die Indexmenge  $J$  leer, so verstehen wir unter der Summe den Nullvektor. Es wird hier nicht vorausgesetzt, daß die Vektoren  $v_j$  paarweise verschieden sind. Gehören die Vektoren  $v_j$  alle einem Unterraum  $U$  an, so liegt auch jede ihrer Linearkombinationen in diesem Unterraum.

Das Bilden von Linearkombinationen ist der fundamentale Rechenprozeß der linearen Algebra.  $\diamond$

**(1.3.3) Notiz.** Sei  $S \subset V$  eine beliebige Teilmenge. Die Menge  $L(S)$  ist gleich der Menge aller Linearkombinationen, die sich mit Vektoren aus  $S$  formen lassen, oder, wie wir auch sagen wollen, die sich durch  $S$  linear darstellen lassen.

*Beweis.* Sei  $L'(S)$  die Menge aller genannten Linearkombinationen. Dann ist  $L'(S)$  ein Unterraum, da die Summe zweier Linearkombinationen aus Elementen von  $S$  wieder eine solche Linearkombination ist; und ebenso ein skalares Vielfaches. Nach Konstruktion ist  $S \subset L'(S)$  und folglich gilt, nach Definition von  $L(S)$ , die Inklusion  $L(S) \subset L'(S)$ . Andererseits ist  $S \subset L(S)$ , und aus der Definition eines Unterraumes folgt, daß dann auch jede Linearkombination aus Elementen von  $S$  in diesem Unterraum liegt. Also gilt auch  $L'(S) \subset L(S)$ .  $\square$

Wir bemerken zur letzten Notiz: Die Menge  $S$  darf unendlich sein, jedoch ist jede Linearkombination eine endliche Summe.

**(1.3.4) Linear abhängig. Linear unabhängig.** Sei  $S \subset V$  eine beliebige Teilmenge. Ist  $x \in L(S)$ , so nennen wir  $x$  *linear abhängig von  $S$* . Ist  $x \notin L(S)$ , so heißt  $x$  *linear unabhängig von  $S$* .

Wir nennen  $S$  *linear unabhängig*, wenn jedes  $x \in S$  von  $S \setminus \{x\}$  linear unabhängig ist. Verneinung:  $S$  ist *linear abhängig*. Eine Menge ist also linear abhängig genau dann, wenn mindestens eines ihrer Elemente durch die restlichen linear dargestellt werden kann.  $\diamond$

**(1.3.5) Basis.** Eine Teilmenge  $B$  eines Vektorraumes  $V$  heißt *Basis* von  $V$ , wenn  $B$  linear unabhängig ist und  $V$  erzeugt.  $\diamond$

**(1.3.6) Notiz.** Eine Menge  $\{x_1, \dots, x_n\} \subset V$  ist genau dann linear unabhängig, wenn aus einer linearen Relation  $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$  folgt, daß alle  $\lambda_j = 0$  sind.

*Beweis.* Sei die Menge linear unabhängig. Angenommen, es gebe eine Relation, in der (etwa)  $\lambda_1 \neq 0$  ist. Dann ist

$$x_1 = -\lambda_1^{-1}(\lambda_2 x_2 + \dots + \lambda_n x_n) \in L(S \setminus \{x_1\}).$$

Widerspruch zur linearen Unabhängigkeit.

Sei die Menge linear abhängig. Dann wird eines ihrer Elemente, etwa  $x_1$ , durch die restlichen linear dargestellt  $x_1 = \sum_{j=2}^n \mu_j x_j$ . Das widerspricht offenbar der vorausgesetzten Bedingung.  $\square$

**(1.3.7) Standardbasis.** Der Vektor  $e_i \in K^n$ , der an der  $i$ -ten Stelle eine 1 hat und sonst überall Nullen, heißt der  $i$ -te *Standardvektor* des Vektorraumes  $K^n$ . (In anderen Zusammenhängen nennen wir diese Vektoren auch Einheitsvektoren.) Die Standardvektoren sind linear unabhängig. Es gilt nämlich  $\sum_{i=1}^n \lambda_i e_i = (\lambda_1, \dots, \lambda_n)$ ; wenn diese Summe also der Nullvektor  $(0, \dots, 0) \in K^n$  ist, so sind alle  $\lambda_i = 0$ . Dieselbe Summe zeigt, daß die Menge der Standardvektoren den  $K^n$  erzeugt. Diese Basis nennen wir die *Standardbasis* des  $K^n$ .  $\diamond$

Hier noch eine Bemerkung zur Sprechweise: Wir werden später oft sagen, seien die Vektoren  $x_1, \dots, x_n$  linear unabhängig oder etwas Ähnliches. Das klingt so, als käme diese Eigenschaft jedem einzelnen Vektor zu. Gemeint ist jedoch eine Eigenschaft der Menge  $\{x_1, \dots, x_n\}$ . Allgemein nennen wir das Bestehen von  $\lambda_1 x_1 + \dots + \lambda_n x_n = 0$  eine *lineare Relation* zwischen den Vektoren  $(x_1, \dots, x_n)$ , und eine *echte*, wenn nicht alle Koeffizienten  $\lambda_j$  gleich Null sind.

Ein einzelner Vektor  $x$  ist genau dann linear unabhängig, wenn er nicht der Nullvektor ist. Beweis: Ist  $x = 0$ , so ist  $1 \cdot x = 0$  eine nichttriviale lineare Relation. Ist  $\lambda x = 0$  und  $\lambda \neq 0$ , so folgt durch Multiplikation mit dem Inversen von  $\lambda$ , daß  $x = 0$  ist.

Die leere Menge ist eine Basis des Nullraumes.

**(1.3.8) Notiz.** Seien  $A$  und  $S$  Teilmengen von  $V$ . Ist jedes Element von  $A$  linear abhängig von  $S$ , so gilt  $L(S) = L(S \cup A)$ .

*Beweis.* Die Inklusion  $L(S) \subset L(S \cup A)$  ist klar. Sei  $x \in L(S \cup A)$ . Dann stellen wir  $x$  zunächst durch  $S \cup A$  linear dar. Alle Elemente von  $A$ , die in dieser linearen Darstellung vorkommen, werden sodann durch eine lineare Darstellung mittels  $S$  ersetzt. Das Resultat ist eine lineare Darstellung durch  $S$ . Also gilt  $L(S \cup A) \subset L(S)$ .  $\square$

**(1.3.9) Satz.** Sei  $M \subset S \subset V$ , sei  $M$  linear unabhängig und  $S$  ein Erzeugendensystem von  $V$ . Dann gibt es eine Basis  $B$  von  $V$  mit der Eigenschaft  $M \subset B \subset S$ . Insbesondere hat jeder Vektorraum eine Basis.

*Beweis.* Wir beweisen den Satz nur für den Fall, daß  $S$  endlich ist. Das reicht für die meisten Anwendungen und erspart uns einen genaueren Blick in die Mengenlehre.

Unter allen linear unabhängigen Teilmengen  $X$  von  $V$  mit  $M \subset X \subset S$  wählen wir eine maximale aus, etwa  $B = \{b_1, \dots, b_n\}$ . Wir zeigen, daß  $B$  den Raum  $V$  erzeugt. Falls  $B = S$  ist, bleibt nichts mehr zu zeigen. Andernfalls gibt es ein  $x \in S \setminus B$ . Wegen der Maximalität von  $B$  ist  $B \cup \{x\}$  linear abhängig. Es gibt also eine echte lineare Relation

$$\lambda x + \lambda_1 x_1 + \dots + \lambda_n x_n = 0.$$

Wäre  $\lambda = 0$ , so würde aus der linearen Unabhängigkeit von  $B$  folgen, daß alle  $\lambda_i = 0$  sind. Also ist  $\lambda \neq 0$  und damit  $x$  in dem von  $B$  erzeugten Unterraum enthalten. Somit ist  $S \subset L(B)$  und mit  $L(S) = V$  folgt  $V = L(S) \subset L(B) \subset V$  also  $L(B) = V$ .  $\square$

Die Endlichkeit der Menge  $S$  wurde im vorstehenden Beweis nur dazu benutzt, die Existenz einer maximalen Menge  $B$  sicherzustellen. Ein Axiom der Mengenlehre, etwa das sogenannte Zornsche Lemma, liefert, daß es die maximalen Mengen immer gibt. In Wahrheit wird man hier zu subtilen Problemen der Logik und Mengenlehre geführt.

**(1.3.10) Satz.** *Sei  $M \subset V$  linear unabhängig und  $T$  ein (endliches) Erzeugendensystem von  $V$ . Dann kann  $M$  durch Elemente aus  $T$  zu einer Basis ergänzt werden.*

*Beweis.* Wir wenden den vorstehenden Satz auf  $S = M \cup T$  an.  $\square$

Es gibt viele Basen. Wesentlich für die gesamte (lineare) Algebra ist der folgende erste Hauptsatz der Algebra:

**(1.3.11) Satz (Basissatz).** *In einem von einer endlichen Menge erzeugten Vektorraum haben je zwei Basen dieselbe Mächtigkeit.*  $\square$

Der Basissatz folgt aus dem mehr technisch formulierten nächsten. Er verschärft (1.3.10) durch eine Aussage über die Anzahl der zur Ergänzung nötigen Elemente. Er läuft unter dem Namen *Austauschsatz von Steinitz*<sup>3</sup>. Für einen weiteren Beweis des Basissatzes siehe Aufgabe 3.

**(1.3.12) Satz.** *Sei  $S$  eine endliche linear unabhängige Menge in  $V$  und  $T$  ein endliches Erzeugendensystem von  $V$ . Dann gilt  $|S| \leq |T|$ , und  $S$  läßt sich durch  $|T| - |S|$  Elemente aus  $T$  zu einem Erzeugendensystem ergänzen.*

*Beweis.* Wir beweisen den Satz durch vollständige Induktion nach der Mächtigkeit  $|S|$  von  $S$ . Ist  $S$  leer, also  $|S| = 0$ , so folgt die Behauptung aus (1.3.10). Sei

---

<sup>3</sup>Ernst Steinitz 1871 – 1928

$|S| > 0$  und  $S = P \cup \{b\}$ . Dann ist  $P$  als Teil von  $S$  auch linear unabhängig. Nach Induktionsvoraussetzung können wir  $P$  durch eine Teilmenge  $Q$  von  $T$  mit  $|T| - |S| + 1$  Elementen zu einem Erzeugendensystem  $P \cup Q = T'$  ergänzen. Der Vektor  $b$  ist also in  $L(T')$  enthalten aber nicht in  $L(P)$ , da  $S$  linear unabhängig ist. Eine lineare Darstellung von  $b$  durch  $T'$  muß deshalb Elemente aus  $Q$  wirklich „benutzen“, das heißt es gibt ein  $c \in Q$ , so daß  $b$  die Form

$$b = v + \lambda c, \quad \lambda \neq 0, \quad v \in L(P \cup (Q \setminus \{c\}))$$

hat. Also liegt  $c = \lambda^{-1}(b - v)$  in  $L(S \cup (Q \setminus \{c\}))$ . Da  $P \cup Q$  und also auch  $S \cup Q$  ein Erzeugendensystem ist, so nach (1.3.8) auch  $S \cup (Q \setminus \{c\})$ .  $\square$

Es ist klar, wie (1.3.11) aus (1.3.12) folgt: Sind  $A$  und  $B$  Basen, so besagt (1.3.12) sowohl  $|A| \leq |B|$  als auch  $|B| \leq |A|$ .

**(1.3.13) Dimension.** Die Mächtigkeit der Basis eines endlich erzeugten Vektorraumes  $V$  heißt die *Dimension* von  $V$ . Wir schreiben  $\dim V$  für diese Zahl (oder auch  $\dim_K V$ , um die Abhängigkeit von  $K$  zu betonen). Ist  $V$  nicht endlich erzeugt, so schreiben wir  $\dim V = \infty$ . Ist  $U$  ein Unterraum von  $V$ , so wird  $\dim V - \dim U$  die *Kodimension* von  $U$  in  $V$  genannt.  $\diamond$

Wir haben in (1.3.7) gesehen, daß  $K^n$  die Dimension  $n$  hat. Der Nullraum hat die Dimension Null.

**(1.3.14) Notiz.** Sei  $V$  ein  $n$ -dimensionaler Vektorraum. Die Menge  $S \subset V$  habe  $n$  Elemente ( $n \geq 1$ ). Dann gilt:

- (1) Ist  $S$  linear unabhängig, so ist  $S$  eine Basis.
- (2) Ist  $S$  ein Erzeugendensystem, so ist  $S$  eine Basis.

*Beweis.* (1) Wäre  $S$  keine Basis, so könnte  $S$  zu einer Basis ergänzt werden (1.3.10), im Widerspruch zu (1.3.12).

(2) Eine geeignete Teilmenge  $T$  von  $S$  ist eine Basis (1.3.10). Wegen (1.3.12) hat  $T$   $n$  Elemente, ist also gleich  $S$ .  $\square$

**(1.3.15) Notiz.** Ist  $U$  Unterraum von  $V$  und gilt  $\dim U = \dim V$ , so ist  $U = V$ .  $\square$

Sei  $b_1, \dots, b_n$  eine Basis von  $V$ . Eine Basis ist nach (1.3.5) eine Teilmenge von  $V$ . Wenn wir die Basis in der Form  $b_1, \dots, b_n$  schreiben, so haben wir die Basisvektoren durch die angehängten Nummern (= Indices) in eine Reihenfolge gebracht. Wir sprechen in diesem Fall von einer *geordneten* Basis.

**(1.3.16) Koordinaten.** Sei  $b_1, \dots, b_n$  eine geordnete Basis. Jeder Vektor  $v \in V$  läßt sich in der Form  $v = \lambda_1 b_1 + \dots + \lambda_n b_n$  schreiben. Durch  $v$  ist  $(\lambda_1, \dots, \lambda_n)$  eindeutig bestimmt, denn wäre auch  $v = \mu_1 b_1 + \dots + \mu_n b_n$ , so würde durch Subtraktion

$$0 = (\mu_1 - \lambda_1)b_1 + \dots + (\mu_n - \lambda_n)b_n$$

folgen und dann nach (1.3.6) aus der linearen Unabhängigkeit  $\mu_i - \lambda_i = 0$ . Wir nennen  $\lambda_i$  die *i-te Koordinate von  $v$  bezüglich der geordneten Basis  $b_1, \dots, b_n$* . Allgemein haben wir für jede Basis  $B$  und jedes  $b \in B$  die zugehörige *b-Koordinate*: Das ist der skalare Koeffizient von  $b$  in einer Linearkombination von Elementen aus  $B$ . Ist  $a = (a_1, \dots, a_n) \in K^n$ , so ist  $a_j$  die *j-te Koordinate von  $a$  bezüglich der Standardbasis*.  $\diamond$

## Ergänzungen und Aufgaben

1. Eine Menge ist genau dann linear unabhängig, wenn jede endliche Teilmenge linear unabhängig ist.
2. Sei  $K[S]$  der Vektorraum der Funktionen  $S \rightarrow K$ , die nur an endlich vielen Stellen ungleich Null sind. Sei  $e_s$  die Funktion, die an der Stelle  $s$  den Wert 1 hat und sonst überall den Wert Null. Die Menge der  $e_s$ ,  $s \in S$  ist eine Basis von  $K[S]$ . Diese Basis verallgemeinert die Standardbasis des  $K^n$ .
3. Der Basissatz (1.3.11) folgt offenbar, wenn man zeigt: Ist  $v_1, \dots, v_n$  eine Basis, so sind Systeme  $y_1, \dots, y_{n+1}$  immer linear abhängig. Es ist nützlich, einmal aufzuschreiben, was diese Aussage bedeutet. Zunächst gibt es nach Voraussetzung lineare Darstellungen  $y_i = \sum_{j=1}^n \lambda_{ij} v_j$  mit gewissen Skalaren  $\lambda_{ij}$ . Wir wollen einsehen: Es gibt Skalare  $x_1, \dots, x_{n+1}$ , die nicht sämtlich Null sind, so daß  $\sum_{i=1}^{n+1} x_i y_i = 0$  ist. Wir setzen die Darstellungen der  $y_i$  in diese hypothetische Relation ein und erhalten  $\sum_{i=1}^{n+1} \sum_{j=1}^n x_i \lambda_{ij} v_j = 0$ . Da die  $v_j$  eine Basis bilden, bedeutet das nach (1.3.6): Die  $x_i$  müssen alle Gleichungen  $\sum_{i=1}^{n+1} \lambda_{ij} x_i = 0$  erfüllen. Das ist, wie man sagt, ein System von  $n$  linearen Gleichungen für die zu findenden  $x_i$ . Wir führen es an dieser Stelle nicht aus, werden aber bald im Abschnitt über lineare Gleichung einsehen, daß ein solches System immer einer Lösung hat, die nicht nur aus Nullen besteht. Damit erhalten wir einen weiteren Beweis für den Basissatz.
4. Die Polynome  $1, x, x^2, \dots, x^n$  bilden eine Basis von  $P_n(\mathbb{R})$ . Ebenso die Polynome  $1, x - a, \dots, (x - a)^n$  für  $a \in \mathbb{R}$ .
5. Sei  $A$  eine Menge reeller Zahlen. Wir betrachten für  $a \in A$  die Funktionen  $f_a(x) = \exp(ax)$  aus  $\text{Abb}(\mathbb{R}, \mathbb{R})$ . Diese Menge ist linear unabhängig. Der Vektorraum  $\text{Abb}(\mathbb{R}, \mathbb{R})$  hat also überabzählbare linear unabhängige Teilmengen.
6. Die komplexen Zahlen  $\mathbb{C}$  sind ein  $\mathbb{R}$ -Vektorraum. Eine Basis ist zum Beispiel  $\{1, i\}$ .
7. Sei  $\{b_1, \dots, b_n\}$  eine Basis eines  $\mathbb{C}$ -Vektorraumes  $V$ . Durch Einschränkung der Skalarmultiplikation wird daraus ein  $\mathbb{R}$ -Vektorraum  $V_{\mathbb{R}}$ . Er hat die Basis  $b_1, ib_1, \dots, b_n, ib_n$ . Insbesondere gilt also  $\dim_{\mathbb{R}} V_{\mathbb{R}} = 2 \dim_{\mathbb{C}} V$ .
8. Prüfungsfrage: Was ist ein Vektor? Antwort: Ein Element eines Vektorraumes. So weit, so gut.

Aber doch irgendwie unbefriedigend. Man denke etwa an den Gebrauch des Wortes „Vektor“ in der Physik, etwa Kraftvektor. Da wird ja eigentlich nicht erst ein Vektorraum definiert. Man könnte an einen Standardraum denken, aber in der Natur gibt es keine Koordinaten! In Lehrbüchern der Physik findet man darüber hinaus verschiedene „Sorten“ von Vektoren. Es gibt mathematische Auswege, doch lasse ich es hier mal zum Grübeln so stehen.

## 1.4 Lineare Abbildungen

Wir kommen nun zum zweiten Grundbegriff der linearen Algebra.

**(1.4.1) Lineare Abbildung.** Seien  $V$  und  $W$  Vektorräume über dem Körper  $K$ . Eine *lineare Abbildung*  $f: V \rightarrow W$  von  $V$  nach  $W$  ist eine Abbildung der Menge  $V$  in die Menge  $W$ , so daß für  $x, y \in V$  und  $\lambda \in K$  gilt:

$$f(x + y) = f(x) + f(y) \qquad f(\lambda x) = \lambda f(x).$$

Die identische Abbildung ist linear; die Verkettung linearer Abbildungen ist linear.

Aus den Axiomen für eine lineare Abbildung  $f$  folgt durch Induktion nach  $n$  die Verträglichkeit mit Linearkombinationen:

$$\sum_{i=1}^n \lambda_i f(x_i) = f\left(\sum_{i=1}^n \lambda_i x_i\right), \quad \lambda_i \in K, x_i \in V.$$

Kennt man also die Werte von  $f$  an den Stellen  $x_i$ , so kennt man  $f$  für alle Vektoren der Form  $\sum_{i=1}^n \lambda_i x_i$ .  $\diamond$

Lineare Abbildungen sind hier also zunächst formal algebraisch definiert. Viele Begriffe der linearen Algebra haben aber einen geometrischen Hintergrund, der natürlich für ein intuitives Arbeiten oft nützlich ist. So sollte man ein Verständnis darüber gewinnen, wie lineare Abbildungen der Standardebene  $\mathbb{R}^2$  in sich „aussehen“. Das werden wir später erläutern, nachdem wir weitere Hilfsmittel wie etwa die Matrizenrechnung entwickelt haben. Hier nur einige Vokabeln: Drehungen um den Nullpunkt, Spiegelungen an einer Geraden, Dehnungen und Stauchungen, Scherungen, Parallelprojektionen auf eine Gerade.

Grundlegend unter anderem für die rechnerische Analyse von linearen Abbildungen ist der nächste Satz.

**(1.4.2) Satz.** *Sei  $B$  eine Basis des  $K$ -Vektorraumes  $V$  und  $W$  ein weiterer  $K$ -Vektorraum. Zu jeder Familie  $(w_b \mid b \in B)$  gibt es genau eine lineare Abbildung  $f: V \rightarrow W$ , die  $b$  auf  $w_b$  abbildet.*

*Beweis.* Die Eindeutigkeit folgt aus der Formel  $f\left(\sum_{b \in B} \lambda_b b\right) = \sum_{b \in B} \lambda_b f(b)$ .

Zu jedem  $v \in V$  gibt es genau eine lineare Darstellung  $v = \sum_{b \in B} \lambda_b b$ . Wir definieren eine Abbildung  $f$  durch die Vorschrift  $f(v) = \sum_{b \in B} \lambda_b w_b$ . Dann gilt jedenfalls  $f(b) = w_b$ . Die definierenden Eigenschaften einer linearen Abbildung werden nachgerechnet: Ist  $u = \sum_{b \in B} \mu_b b$ , so folgt  $u + v = \sum_{b \in B} (\mu_b + \lambda_b) b$ . Nach Definition von  $f$  ist dann  $f(u + v) = \sum_{b \in B} (\mu_b + \lambda_b) w_b$ , und letzteres ist offenbar gleich  $f(u) + f(v)$ . Ähnlich verifiziert man  $f(\lambda v) = \lambda f(v)$ .

Wir haben im vorstehenden Beweis so getan, als sei  $B$  endlich. Im allgemeinen Fall ist  $v = \sum_{b \in J} \lambda_b b$  für eine endliche Teilmenge  $J \subset B$ , und es ist dann zweckmäßig, dafür  $\sum_{b \in B} \lambda_b b$  zu schreiben (obgleich es ja eigentlich unendliche Summe nicht gibt), wobei die  $\lambda_b$  für  $b \notin J$  als Null anzusehen sind. Zwar ist

durch  $v$  die endliche Teilmenge  $J$  nicht eindeutig bestimmt, doch möge man sich davon überzeugen, daß diese Vereinbarung nicht zu Problemen führt.  $\square$

**(1.4.3) Beispiele.**

(1) Die Multiplikation mit einem Skalar  $\lambda$  liefert eine lineare Abbildung  $l_\lambda: V \rightarrow V, v \mapsto \lambda v$ . Es gilt dann  $l_{\lambda+\mu} = l_\lambda + l_\mu$  und  $l_{\lambda\mu} = l_\lambda l_\mu$ . Diese Aussagen sind ein Teil der Vektorraumaxiome.

(2) Integration von stetigen Funktionen liefert eine lineare Abbildung  $C([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$ .

(3) Differentiation ist eine lineare Abbildung  $C^k(\mathbb{R}, \mathbb{R}) \rightarrow C^{k-1}(\mathbb{R}, \mathbb{R})$ .

(4) Indem man einer konvergenten Folge ihre Summe zuordnet, erhält man eine lineare Abbildung.  $\diamond$

Für lineare Abbildungen  $f: V \rightarrow W$  mit speziellen Eigenschaften verwenden wir die folgende Terminologie:

$f$ Epimorphismus	$\Leftrightarrow$	$f$ surjektiv
$f$ Monomorphismus	$\Leftrightarrow$	$f$ injektiv
$f$ Isomorphismus	$\Leftrightarrow$	$f$ bijektiv
$f$ Endomorphismus	$\Leftrightarrow$	$V = W$
$f$ Automorphismus	$\Leftrightarrow$	$V = W, f$ bijektiv.

Vektorräume heißen *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt. Die folgende Notiz ergibt sich unmittelbar aus den beteiligten Definitionen.

**(1.4.4) Notiz.** Sei  $f: V \rightarrow W$  eine lineare Abbildung. Dann gilt:

- (1) Ist  $V_1$  Unterraum von  $V$ , so ist  $f(V_1)$  Unterraum von  $W$ .
- (2) Ist  $W_1$  Unterraum von  $W$ , so ist  $f^{-1}(W_1)$  Unterraum von  $V$ .  $\square$

Der Unterraum  $f^{-1}(0)$  heißt *Kern* von  $f$ , der Unterraum  $f(V)$  *Bild* von  $f$ .

**(1.4.5) Notiz.** Eine lineare Abbildung  $f$  ist genau dann injektiv, wenn ihr Kern der Nullraum ist.

*Beweis.* Ist  $0 \neq v \in \text{Kern}(f)$ , so ist  $f(v) = 0 = f(0)$  und folglich  $f$  nicht injektiv. Ist  $f(u) = f(v)$  für  $u \neq v$ , so liegt wegen  $f(u - v) = f(u) - f(v) = 0$  das Element  $u - v \neq 0$  im Kern von  $f$ .  $\square$

**(1.4.6) Satz.** Sei  $f: V \rightarrow W$  eine lineare Abbildung.

- (1) Ist  $f$  injektiv und  $S$  linear unabhängig, so ist auch  $f(S)$  linear unabhängig.
- (2) Ist  $f$  surjektiv und  $S$  ein Erzeugendensystem, so ist auch  $f(S)$  ein Erzeugendensystem.
- (3) Ist  $f$  ein bijektiv und  $S$  eine Basis, so ist auch  $f(S)$  eine Basis.

*Beweis.* (1) Aus  $\sum_{b \in J} \lambda_b b = 0$  folgt  $f(\sum_b \lambda_b b) = 0$ , also  $\sum_b \lambda_b f(b) = 0$ , da  $f$  injektiv ist, und dann folgt nach (1.3.6)  $\lambda_b = 0$ .

(2) Sei  $w \in W$ , also  $w = f(v)$  für ein  $v \in V$ . Aus  $v = \sum_b \lambda_b b$  folgt dann  $w = f(v) = \sum_b \lambda_b f(b)$ .

(3) folgt aus (1) und (2).  $\square$

**(1.4.7) Satz.** *Sei  $f: V \rightarrow W$  eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Bilden die Vektoren  $f(v_1), \dots, f(v_n)$  eine Basis des Bildes von  $f$  und  $a_1, \dots, a_m$  eine Basis des Kernes von  $f$ , so ist  $a_1, \dots, a_m, v_1, \dots, v_n$  eine Basis von  $V$ .*

In der Formulierung des Satzes haben wir stillschweigend  $n \geq 1$  und  $m \geq 1$  angenommen. Die Fälle  $n = 0$ ,  $m = 0$  sind klar.

*Beweis.* Sei  $x \in V$ . Dann gibt es eine lineare Relation  $f(x) = \sum_i \lambda_i f(v_i)$  mit gewissen  $\lambda_i \in K$ , da  $f(v_1), \dots, f(v_n)$  eine Basis von  $W$  ist. Da  $f$  linear ist, liegt  $x - \sum_i \lambda_i v_i$  im Kern von  $f$  und folglich gibt es eine lineare Relation der Form  $x - \sum_i \lambda_i v_i = \sum_j \mu_j a_j$  für geeignete  $\mu_j \in K$ . Somit bilden die genannten Elemente ein Erzeugendensystem.

Umgekehrt: Aus einer linearen Relation  $0 = \sum_j \mu_j a_j + \sum_i \lambda_i v_i$  folgt durch Anwendung von  $f$

$$0 = f(0) = \sum_j \mu_j f(a_j) + \sum_i \lambda_i f(v_i) = \sum_i \lambda_i f(v_i).$$

Weil die  $f(v_i)$  linear unabhängig sind, folgt  $\lambda_1 = \dots = \lambda_n = 0$ , und weil die  $a_j$  linear unabhängig sind, folgt dann  $\mu_1 = \dots = \mu_m = 0$ .  $\square$

Ist  $f: V \rightarrow W$  eine lineare Abbildung und  $V$  endlichdimensional, so ist auch  $f(V)$  endlichdimensional. Die Dimension des Bildes  $f(V)$  heißt *Rang* von  $f$ .

**(1.4.8) Dimensionsformel.** Für jede lineare Abbildung  $f: V \rightarrow W$  eines endlichdimensionalen Vektorraumes  $V$  gilt

$$\dim \text{Kern } f + \dim \text{Bild } f = \dim V, \quad \dim \text{Kern } f + \text{Rang } f = \dim V.$$

*Beweis.* Das ist eine unmittelbare Folgerung aus (1.4.8).  $\square$

**(1.4.9) Satz.** *Sei  $f: V \rightarrow W$  eine lineare Abbildung zwischen Vektorräumen gleicher endlicher Dimension. Dann sind folgende Aussagen äquivalent:*

- (1)  $\text{Kern } f = 0$ .
- (2)  $f$  ist injektiv.
- (3)  $f$  ist surjektiv.
- (4)  $f$  ist bijektiv (ein Isomorphismus).

*Beweis.* (1)  $\Leftrightarrow$  (2) haben wir schon in (1.2.3) bemerkt.

(2)  $\Rightarrow$  (3). Ist  $S$  eine Basis von  $V$ , so ist  $f(S)$  linear unabhängig (1.4.6) und wegen (1.3.15) eine Basis von  $W$ .

(3)  $\Rightarrow$  (4). Wegen (1.4.8) ist  $\dim \text{Kern } f = 0$ , also  $f$  auch injektiv, also bijektiv.

(4)  $\Rightarrow$  (1), (2), (3) ist klar.  $\square$

Für ein Verständnis von linearen Selbstabbildungen (= Endomorphismen) eines Vektorraumes  $V$  möchte man wissen, welche Wirkung sie auf Unterräumen haben. Insbesondere fragt man, ob es Unterräume gibt, die in sich selbst abgebildet werden.

**(1.4.10) Eigenwerte, Eigenvektoren, Eigenräume.** Sei  $f: V \rightarrow V$  eine lineare Abbildung. Wir sagen  $0 \neq v \in V$  ist ein *Eigenvektor* von  $f$  zum *Eigenwert*  $\lambda$ , wenn  $f(v) = \lambda v$  gilt. Zu gegebenem Skalar  $\lambda$  heißt  $\{x \in V \mid f(x) = \lambda x\}$  der *Eigenraum* von  $f$  zum Eigenwert  $\lambda$ , wenn dieser Raum nicht der Nullraum ist. Ein Unterraum  $W \subset V$  heiße *f-stabil*, wenn  $f(W) \subset W$  gilt. Die Eigenräume von  $f$  sind *f-stabil*.  $\diamond$

**(1.4.11) Räume aus linearen Abbildungen.** Seien  $V$  und  $W$  Vektorräume über  $K$ . Wir bezeichnen mit

$$\text{Hom}_K(V, W) \quad \text{oder} \quad \text{Hom}(V, W)$$

die Menge aller  $K$ -linearen Abbildungen  $f: V \rightarrow W$ . Auf dieser Menge definieren wir wie in (1.1.9) eine Vektorraumstruktur durch  $(f+g)(v) = f(v) + g(v)$  und  $(\lambda f)(v) = \lambda(f(v))$ , wobei  $f, g \in \text{Hom}(V, W)$ ,  $v \in V$  und  $\lambda \in K$ . (Die Notation rührt daher, daß eine lineare Abbildung in einem allgemeinen algebraischen Kontext auch als Homomorphismus bezeichnet wird. Wir könnten hier auch  $L(V, W)$  für diesen Raum schreiben.)

Mit linearen Abbildungen kann man rechnen. Der Vektorraum  $\text{Hom}(V, V)$  hat zunächst einmal eine Addition und eine Skalarmultiplikation. Er trägt aber auch ein Produkt, nämlich die Verkettung von Abbildungen, und wir schreiben deshalb auch  $f \circ g = f \cdot g = fg$  in diesem Kontext. Sind  $a, b, c, d$  Skalare und  $A, B, C, D$  Elemente aus  $\text{Hom}(V, V)$ , so gilt

$$(aA + bB)(cC + dD) = acAC + adAD + bcBC + bdBD,$$

wie man es vom Rechnen gewohnt ist. Das wird später noch deutlicher, wenn wir lineare Abbildungen durch die sogenannten Matrizen beschreiben.  $\diamond$

**(1.4.12) Dualraum.** Ist  $W = K$ , so nennen wir  $\text{Hom}(V, K) = V^*$  den *Dualraum* von  $V$ . Ein Element  $f \in V^*$ , also eine lineare Abbildung  $V \rightarrow K$ , wird *Linearform* auf  $V$  genannt.

Ist  $f: V \rightarrow W$  eine lineare Abbildung, so gehört dazu die *duale Abbildung* (auch *transponierte Abbildung* genannt)  $f^*$  in umgekehrter Richtung

$$f^*: W^* \rightarrow V^*, \quad \alpha \mapsto \alpha \circ f.$$

Sind  $f: U \rightarrow V$  und  $g: V \rightarrow W$  lineare Abbildungen, so gelten die Regeln  $(gf)^* = f^*g^*$  und  $(\text{id}(V))^* = \text{id}(V^*)$ .

Sei  $b_1, \dots, b_n$  Basis von  $V$ . Die  $k$ -te Koordinatenabbildung ist die lineare Abbildung  $b^k$ , die  $b_k$  auf 1 und  $b_l$  für  $k \neq l$  auf Null abbildet. Die  $b^k$  sind Elemente des Dualraums  $V^*$ . Sie bilden eine Basis von  $V^*$ , genannt die *Dualbasis* zur gegebenen Basis. Ist nämlich  $f \in V^*$ , so gilt  $f = \sum_k f(k)b^k$ , und diese Darstellung ist eindeutig.  $\diamond$

## Ergänzungen und Aufgaben

1. Zu einer geordneten Basis  $b_1, \dots, b_n$  von  $V$  gehört ein eindeutig bestimmter Isomorphismus  $K^n \rightarrow V$ , der den Standardvektor  $e_j$  auf  $b_j$  abbildet.

2. Sei  $0 < \rho = \sqrt[5]{2} \in \mathbb{R}$  die fünfte Wurzel aus 2. Die reellen Zahlen der Form  $\sum_{i=0}^4 a_i \rho^i$  mit rationalen  $a_i$  bilden einen Unterkörper  $\mathbb{Q}(\rho)$  von  $\mathbb{R}$ . Das einzige Axiom, das nicht unmittelbar aus den Rechenregeln folgt, ist die Existenz eines multiplikativen Inversen für  $0 \neq x \in \mathbb{Q}(\rho)$ . Man könnte versuchen, ein Inverses explizit auszurechnen; das ist aber nicht so leicht. Nach einer Formel ist aber auch nicht gefragt, sondern nur nach der Existenz. Es ist  $\mathbb{Q}(\rho)$  ein  $\mathbb{Q}$ -Vektorraum, und die Multiplikation  $y \mapsto xy$  ist  $\mathbb{Q}$ -linear und für  $x \neq 0$  injektiv und deshalb sogar bijektiv. Also liegt die 1 im Bild.

Man wird vermuten, daß  $\mathbb{Q}(\rho)$  als  $\mathbb{Q}$ -Vektorraum die Dimension 5 hat; das ist aber nicht ganz selbstverständlich. Solche Probleme werden in der Körpertheorie systematisch behandelt.

3. Für den Vektorraum  $V = K^\infty$  gibt es injektive lineare Abbildungen  $V \rightarrow V$ , die nicht bijektiv sind.

4. Zwei endlichdimensionale Vektorräume  $V$  und  $W$  sind genau dann isomorph, wenn sie gleiche Dimension haben.

5. Gilt für lineare Abbildungen  $f: V \rightarrow W$  und  $g: W \rightarrow V$  die Gleichung  $gf = \text{id}$ , so heißt  $g$  *linksinvers* zu  $f$  und  $f$  *rechtsinvers* zu  $g$ .

Eine lineare Abbildung  $f: V \rightarrow W$  ist genau dann injektiv (surjektiv), wenn sie ein Linksinverses (Rechtsinverses) hat. Sie ist genau dann ein Isomorphismus, wenn sie Links- und Rechtsinverse hat, die dann notwendig beide gleich der Umkehrabbildung sind.

6. Die Begriffsbildungen der linearen Algebra sind in vielen Bereichen nützlich, zum Beispiel bei linearen Differentialgleichungen. Die zweifach Ableitung  $D^2$  ist eine lineare Abbildung auf dem Raum der 2-mal stetig differenzierbaren Funktionen  $C^2(\mathbb{R}, \mathbb{R})$ . Der Kern der Abbildung  $D^2 + \text{id}$ , also die Menge der Funktionen, die der Differentialgleichung  $f'' = -f$  genügen, ist ein Unterraum. Er ist zweidimensional und hat die Funktionen  $\cos$  und  $\sin$  als Basis. Das kann mit elementaren Hilfsmitteln der Diffe-

rentialrechnung bewiesen werden<sup>4</sup>.

7. Eine Sequenz  $U \xrightarrow{u} V \xrightarrow{v} W$  von linearen Abbildungen heißt *exakt* an der Stelle  $V$ , wenn  $\text{Bild } u = \text{Kern } v$  ist. Eine Sequenz

$$U_0 \xrightarrow{u_0} U_1 \dots \xrightarrow{u_1} \dots \xrightarrow{u_n} U_{n+1}$$

heißt *exakt*, wenn sie an allen Stellen  $U_j$ ,  $1 \leq j \leq n$  exakt ist. Eine exakte Sequenz der Form

$$0 \rightarrow U \rightarrow V \rightarrow W \rightarrow 0$$

heißt auch *kurze exakte Sequenz*. In einer solchen exakten Sequenz ist  $U \rightarrow V$  injektiv und  $V \rightarrow W$  surjektiv und  $U$  isomorph zum Kern von  $V \rightarrow W$ .

## 1.5 Summen und Produkte

Aus gegebenen Vektorräumen lassen sich in mancherlei Weise neue herstellen. Wir behandeln in diesem Abschnitt Summen und Produkte. Der Terminus „Summe“ wird hier allerdings mit verschiedener Bedeutung benutzt.

**(1.5.1) Summenräume.** Die Vereinigung  $U_1 \cup U_2$  von Unterräumen  $U_j \subset V$  eines Vektorraumes  $V$  ist im allgemeinen kein Unterraum. Sei  $U_1 + U_2$  die Menge aller Vektoren der Form  $u_1 + u_2$ , worin  $u_j \in U_j$  ist. Diese Menge wird leicht als Unterraum nachgewiesen; er heißt die *Summe* der Unterräume  $U_1$  und  $U_2$ . Ebenso wird die Summe  $\sum_{j=1}^r U_j = U_1 + \dots + U_r$  von Unterräumen  $U_1, \dots, U_r$  als der Unterraum aller Vektoren der Form  $u_1 + \dots + u_r$ ,  $u_j \in U_j$  definiert.  $\diamond$

**(1.5.2) Dimensionsformel.** Seien  $U_1$  und  $U_2$  endlichdimensionale Unterräume eines Vektorraums. Dann gilt die Dimensionsformel

$$\dim U_1 + \dim U_2 = \dim(U_1 + U_2) + \dim(U_1 \cap U_2).$$

*Beweis.* Wir wählen eine Basis  $B_0$  von  $U_0 = U_1 \cap U_2$  und ergänzen sie zu einer Basis  $B_j$  von  $U_j$ . Wir behaupten:  $B_0 \cup B_1 \cup B_2 = B$  ist eine Basis von  $U = U_1 + U_2$ . Es ist nämlich  $B$  offenbar ein Erzeugendensystem. Ist  $W_j$  der von  $B_j$  erzeugte Unterraum, so ist  $U = W_0 \oplus W_1 \oplus W_2$  und  $U_j = W_0 \oplus W_j$ , woraus mit der Dimensionsformel für direkte Summen die Behauptung folgt.  $\square$

**(1.5.3) Satz.** Seien  $V_1, \dots, V_r$  Unterräume von  $V$ . Folgende Aussagen sind äquivalent:

<sup>4</sup>Siehe etwa Th. Bröcker: Analysis I. (2. Auflage). Spektrum Akademischer Verlag, Heidelberg 1995

- (1) Jeder Vektor  $v \in V$  besitzt eine eindeutig bestimmte Darstellung der Form  $v = v_1 + \dots + v_r$  mit  $v_i \in V_i$ .  
 (2)  $V = \sum_{j=1}^r V_j$ , und für jedes  $i$  ist  $V_i \cap (\sum_{j \neq i} V_j) = \{0\}$ .

*Beweis.* (1)  $\Rightarrow$  (2). Da jeder Vektor  $v$  eine Darstellung  $v = v_1 + \dots + v_r$  hat, ist  $\sum_{j=1}^r V_j = V$ . Sei  $v \in V_i \cap (\sum_{j \neq i} V_j)$ . Dann gibt es Darstellungen  $v = v_i$  und  $v = \sum_{j \neq i} v_j$  mit  $v_k \in V_k$ . Aus  $0 = -v_i + \sum_{j \neq i} v_j$  folgt wegen der vorausgesetzten Eindeutigkeit einer Darstellung (hier: des Nullvektors), daß  $v_k = 0$  ist für alle  $k$ ; also ist  $v = 0$ .

(2)  $\Rightarrow$  (1). Wegen  $V = \sum_{j=1}^r V_j$  gibt es mindestens eine Darstellung der genannten Art. Seien  $v = v_1 + \dots + v_r$  und  $v = v'_1 + \dots + v'_r$  zwei solche Darstellungen. Dann ist  $v_i - v'_i \in V_i \cap \sum_{j \neq i} V_j = \{0\}$ , also gilt  $v_i = v'_i$ .  $\square$

**(1.5.4) Interne direkte Summe.** Gilt eine der Aussagen (1), (2) in (1.5.3), so sagen wir:  $V$  sei in die Unterräume  $V_1, \dots, V_r$  *direkt zerlegt* und schreiben

$$V = V_1 \oplus \dots \oplus V_r = \bigoplus_{j=1}^r V_j$$

für das Bestehen dieses Sachverhalts, den wir auch eine *direkte Zerlegung* von  $V$  nennen. Statt direkte Zerlegung ist auch der Terminus *direkte Summe* in Gebrauch.  $\diamond$

**(1.5.5) Komplement.** Sei  $U$  ein Unterraum von  $W$ . Dann heißt ein Unterraum  $V$  von  $W$  *Komplement* von  $U$  in  $W$ , wenn  $W$  in  $U$  und  $V$  direkt zerlegt ist.

Im allgemeinen gibt es verschiedene Komplemente: So ist etwa jede Gerade durch den Nullpunkt, die nicht in  $U = \mathbb{R}^2 \times \{0\}$  liegt, ist ein Komplement von  $U$  in  $\mathbb{R}^3$ .  $\diamond$

**(1.5.6) Satz.** Jeder Unterraum eines  $U$  (endlichdimensionalen) Vektorraumes  $W$  besitzt ein Komplement.

*Beweis.* Eine Basis  $(b_1, \dots, b_m)$  eines Unterraumes  $U$  werde durch  $(c_1, \dots, c_n)$  zu einer Basis von  $W$  ergänzt. Sei  $V$  der von  $(c_1, \dots, c_n)$  erzeugte Unterraum. Dann ist nach dem soeben Bemerkten  $V$  ein Komplement von  $U$ .  $\square$

**(1.5.7) Projektion.** Ein Endomorphismus  $p: V \rightarrow V$  mit der Eigenschaft  $p^2 = p \circ p = p$  heißt *Projektion* oder *Projektionsoperator* von  $V$ .  $\diamond$

**(1.5.8) Satz.** Sei  $p: V \rightarrow V$  eine Projektion. Dann ist auch  $q = \iota - p$  eine Projektion ( $\iota$  die identische Abbildung), und es gelten die Relationen  $\iota = p + q$  und  $qp = 0 = pq$ .

Seien  $p, q: V \rightarrow V$  lineare Abbildungen mit den Eigenschaften  $\iota = p + q$  und  $pq = 0$ . Dann sind  $p$  and  $q$  Projektionen, und es gelten

$$V = \text{Bild } p \oplus \text{Bild } q = \text{Kern } p \oplus \text{Kern } q$$

sowie  $\text{Bild } p = \text{Kern } q$ ,  $\text{Bild } q = \text{Kern } p$ .

*Beweis.* Es ist  $qp = (\iota - p)p = p - p^2 = 0$  und  $q^2 = (1 - p)(1 - p) = 1 - p - p + p^2 = 1 - p = q$ . Diese Rechnungen zeigen die erste Behauptung.

Wegen  $q = \iota - p$  und  $pq = 0$  ist  $0 = pq = p(\iota - p) = p - p^2$ , also  $p$  eine Projektion. Wegen  $\iota = p + q$  wird  $V$  von den Bildern von  $p$  und  $q$  erzeugt. Sei  $x$  im Schnitt der beiden Bilder, also  $p(y) = x = (1 - p)(z)$  für gewisse  $y, z$ . Es folgt  $x = p(y) = p^2(y) = p(1 - p)(z) = 0$ . Also ist die Summe direkt (1.5.3).

Wegen  $pq = 0$  ist  $\text{Bild } q \subset \text{Kern } p$ . Ist  $p(x) = 0$ , so folgt  $q(x) = (\iota - p)x = x$ , also gilt  $\text{Kern } p \subset \text{Bild } q$ .  $\square$

**(1.5.9) Beispiel.** Das Rechnen mit linearen Abbildungen demonstrieren wir an einem instruktiven Beispiel. Sei  $A: V \rightarrow V$  eine lineare Abbildung eines reellen Vektorraumes, deren Quadrat  $A^2 = A \circ A$  die identische Abbildung  $\iota$  ist. Wir verwenden die Abbildungen

$$A_+ = \frac{1}{2}(A + \iota), \quad A_- = \frac{1}{2}(A - \iota).$$

Dann rechnet man leicht die Relationen

$$\iota = A_+ + A_-, \quad A_+^2 = A_+, \quad A_-^2 = A_-, \quad A_+A_- = 0$$

nach. Also sind  $A_\pm$  beides Projektionen. Sei  $V_\pm$  das Bild von  $A_\pm$ . Dann gilt nach (1.5.9)  $V = V_+ \oplus V_-$ . Es gilt ferner  $AA_\pm = A_\pm$ , so daß die Vektoren in  $V_\pm$  Eigenvektoren zum Eigenwert  $\pm 1$  sind. Das Bild von  $A_\pm$  ist der Kern von  $A_\mp$ .

Diese Rechnungen benutzen nicht wirklich die reellen Zahlen. Man kann jeden Körper verwenden, in dem  $1 + 1 \neq 0$  ist. Es ist dann  $\frac{1}{2}$  als das Inverse von  $1 + 1$  zu interpretieren.  $\diamond$

**(1.5.10) Summen und Produkte.** Sei  $(V_j \mid j \in J)$  eine Familie von  $K$ -Vektorräumen. Das mengentheoretische Produkt  $\prod_{j \in J} V_j$  machen wir zu einem Vektorraum durch komponentenweise Addition und Skalarmultiplikation. Darin haben wir den Unterraum  $\bigoplus_{j \in J} V_j$  der Familien  $(v_j \in V_j \mid j \in J)$ , die nur an endlich vielen Stellen ungleich Null sind. Ist die Indexmenge endlich, so unterscheiden sich die beiden Vektorräume nicht. Es gibt aber, wie wir gleich sehen, formale Gründen, sie trotzdem getrennt zu behandeln.

Wir haben die Projektion auf den  $k$ -ten Faktor  $p_k: \prod_{j \in J} V_j \rightarrow V_k, (v_j \mid j \in J) \mapsto v_k$ . Die Injektion des  $k$ -ten Faktor  $i_k: V_k \rightarrow \prod_{j \in J} V_j$  ordnet  $v_k$  die Familie zu, die den Wert  $v_k$  an der Stellen  $k$  hat und sonst überall Null ist. Dasselbe gilt für  $\bigoplus_{j \in J} V_j$ .

Wir nennen  $\bigoplus_{j \in J} V_j$  die *Summe* der gegebenen Familie und  $\prod_{j \in J} V_j$  das *Produkt* der gegebenen Familie.

Die Zuordnungen

$$\text{Hom}(\bigoplus_{j \in J} V_j, W) \rightarrow \prod_{j \in J} \text{Hom}(V_j, W), \quad f \mapsto (f \circ i_j)$$

$$\text{Hom}(U, \prod_{j \in J} V_j) \rightarrow \prod_{j \in J} \text{Hom}(U, V_j), \quad f \mapsto (p_j \circ f)$$

sind für jeden Vektorraum  $U$  und  $W$  bijektiv.  $\diamond$

Wir haben das Summensymbol  $\bigoplus$  auch schon für die interne direkte Summe benutzt. Ist das schädlich? Nein. Ist nämlich  $V_1, \dots, V_n$  eine Familie von Unterräumen von  $V$ , so gibt es genau eine lineare Abbildung  $f: \bigoplus_{j=1}^n V_j \rightarrow V$ , so daß  $f \circ i_j$  die Inklusion  $V_j \subset V$  ist. Genau dann ist  $f$  ein Isomorphismus, wenn  $V$  die interne direkte Summe der  $V_j$  ist, wie aus (1.5.3) folgt. Es ist dann übrigens auch  $\bigoplus_{j=1}^n V_j$  die interne direkte Summe der Bilder  $i_j(V_j)$ . Insofern unterscheiden sich also die beiden Aspekte kaum.

Die Summe von  $n$  Exemplaren  $K$  ist der Standardraum  $K^n$ .

**(1.5.11) Satz.** Sei  $V = \bigoplus_{j=1}^r V_j$  eine interne direkte Summe und  $B_j$  eine Basis von  $V_j$ . Dann ist  $B = \bigcup_j B_j$  eine Basis von  $V$  und  $B_i \cap B_j = \emptyset$  für  $i \neq j$ . Ist eine Basis  $B$  von  $V$  disjunkte Vereinigung von  $B_1, \dots, B_r$  und  $V_j$  der von  $B_j$  erzeugte Unterraum, so ist  $\bigoplus_{j=1}^r V_j = V$ . Es gilt deshalb die Dimensionsformel für direkte Summen

$$\dim \bigoplus_{j=1}^r V_j = \sum_{j=1}^r \dim V_j.$$

*Beweis.* Da  $V$  von den  $V_j$  erzeugt wird und  $V_j$  von  $B_j$ , so wird  $V$  von  $B$  erzeugt. Gibt es eine echte lineare Relation zwischen den Elementen von  $B$ , so erhält man daraus durch Sortieren für ein  $j$  ein von Null verschiedenes Element in  $V_j \cap (\sum_{i \neq j} V_i)$  im Widerspruch zu (1.5.3). Durch Umkehrung der Argumentation erhält man die zweite Aussage des Satzes.  $\square$

**(1.5.12) Direkte Zerlegungen von Abbildungen.** Direkte Zerlegungen gibt es auch für lineare Abbildungen. Seien  $\bigoplus_{j=1}^r V_j = V$  und  $\bigoplus_{j=1}^r W_j = W$  direkte Zerlegungen von  $K$ -Vektorräumen. Seien  $f_j: V_j \rightarrow W_j$  lineare Abbildungen. Dann gibt es genau eine lineare Abbildung  $f: V \rightarrow W$ , mit der Eigenschaft

$$f(v_1 + \dots + v_r) = f_1(v_1) + \dots + f_r(v_r),$$

für alle  $v_j \in V_j$ , denn durch diese Formel wird eine wohldefinierte Abbildung beschrieben, und man rechnet sie leicht als linear nach. Wir schreiben diese Abbildung als  $f = f_1 \oplus \dots \oplus f_r$  und nennen sie die *direkte Summe* der  $f_j$ . Auch sagen wir,  $f$  sei in die  $f_j$  direkt zerlegt.  $\diamond$

Direkte Zerlegungen von Endomorphismen  $f: V \rightarrow V$  sind von grundsätzlicher Bedeutung. Sei  $V$  wie oben direkt zerlegt und bestehe für alle  $j$  die Inklusion  $f_j(V_j) \subset V_j$ . Dann wird durch Einschränkung eine lineare Abbildung  $f_j: V_j \rightarrow V_j$ ,  $v \mapsto f(v)$  induziert. Es gilt  $f = f_1 \oplus \dots \oplus f_r$ .

Der  $\lambda$ -Eigenraum  $V(\lambda)$  von  $f: V \rightarrow V$  ist  $f$ -stabil. Die Summe  $f$ -stabiler Unterräume ist  $f$ -stabil. Ist  $V$  die Summe von Eigenräumen von  $f$ , so wird  $f$  in die zugehörigen Endomorphismen der Eigenräume direkt zerlegt. Genauer gilt für Endomorphismen endlichdimensionaler Vektorräume  $V$ :

**(1.5.13) Satz.** *Seien  $\lambda_1, \dots, \lambda_r$  paarweise verschiedene Eigenwerte des Endomorphismus  $f: V \rightarrow V$ , sei  $V(\lambda_j)$  der Eigenraum von  $\lambda_j$  und gelte  $V_j \subset V(\lambda_j)$ . Sei ferner  $V$  die Summe der  $V_j$ . Dann ist  $V_j = V(\lambda_j)$ , und es gibt keine weiteren Eigenwerte.*

*Beweis.* Sei  $\lambda(b)$  der Eigenwert von  $b \in B$ . Sei  $v = \sum_{b \in B} \mu_b b \neq 0$  Eigenvektor zum Eigenwert  $\lambda$ . Dann gilt

$$\sum_b \lambda \mu_b b = f(v) = \sum_b \mu_b \lambda(b) b.$$

Ist  $\mu_b \neq 0$ , so folgt  $\lambda = \lambda(b)$  und  $\mu_c = 0$  für  $\lambda(c) \neq \lambda(b)$ , daß heißt  $v \in V_j$ , falls  $\lambda_j = \lambda(b)$ . Also ist  $V_j = V(\lambda_j)$ , und es gibt keine weitere Eigenwerte.  $\square$

Der Beweis des letzten Satzes zeigt übrigens:

**(1.5.14) Satz.** *Sind  $v_1, \dots, v_r$  Eigenvektoren zu paarweise verschiedenen Eigenwerten einer linearen Abbildung, so sind sie linear unabhängig.*  $\square$

## Ergänzungen und Aufgaben

1. Einen mathematischen Begriff soll man nicht durch eine „Konstruktion“ sondern durch eine „Eigenschaft“ definieren. Obgleich es für diesen Text noch nicht so wichtig ist, sollte man zu den Vokabeln der Mengensprache auch die Vokabeln der Kategorientheorie zur Kenntnis nehmen<sup>5</sup>. In der Kategorientheorie werden eine Summe und ein Produkt durch sogenannte universelle Eigenschaften definiert (und durch diese Eigenschaft bis auf eindeutige Isomorphie bestimmt). Für die Summe lautet diese Eigenschaft in unserem Kontext:

Sei  $(\iota_j: V_j \rightarrow V \mid j \in J)$  eine Familie von linearen Abbildungen, so daß

$$\text{Hom}(V, W) \rightarrow \prod_{j \in J} \text{Hom}(V_j, W), \quad f \mapsto (f \circ \iota_j)$$

für jeden Vektorraum  $W$  bijektiv ist. Dann gibt es genau einen Isomorphismus  $\varphi: V \rightarrow \bigoplus_{j \in J} V_j$ , so daß  $\varphi \circ \iota_j = i_j$  ist (vergleiche (1.5.10)). Im Fall des Produktes arbeitet man analog mit den Projektionen  $p_k$ .

## 1.6 Quotienträume

**(1.6.1) Satz.** *Sei  $f: U \rightarrow V$  eine surjektive und  $\gamma: U \rightarrow W$  eine weitere lineare Abbildung. Genau dann gibt es eine lineare Abbildung  $\Gamma: V \rightarrow W$  mit der Eigenschaft  $\Gamma f = \gamma$ , wenn  $\text{Kern } f \subset \text{Kern } \gamma$  ist.*

<sup>5</sup>Siehe etwa: S. MacLane: Kategorien. Springer-Verlag, Berlin Heidelberg 1972

*Beweis.* Wir deuten die Situation durch ein Diagramm an.

$$\begin{array}{ccc} U & \xrightarrow{\gamma} & W \\ f \downarrow & \nearrow \Gamma & \\ V & & \end{array}$$

Gilt  $\Gamma f = \gamma$ , so folgt die genannte Inklusion der Kerne unmittelbar.

Umgekehrt gibt es  $\Gamma$  genau dann als Mengenabbildung, wenn aus  $f(x) = f(y)$  immer  $\gamma(x) = \gamma(y)$  folgt. Nun ist aber  $f(x) = f(y)$  gleichwertig zu  $x - y \in \text{Kern } f$  und ebenso für  $\gamma$ . Die somit eindeutig bestimmte Abbildung  $\Gamma$  mit der Eigenschaft  $\Gamma f = \gamma$  ist leicht als linear nachzuweisen: Sind  $x, y \in V$  gegeben, so wählen wir  $x = f(a)$ ,  $y = f(b)$  und erhalten  $\Gamma(x) + \Gamma(y) = \gamma(a) + \gamma(b) = \gamma(a + b) = \Gamma f(a + b) = \Gamma(x + y)$ . Analog für die Skalarmultiplikation.  $\square$

Eine surjektive Abbildung  $f: U \rightarrow V$  mit dem Kern  $L$  nennen wir eine *Quotientabbildung* zum Unterraum  $L$  von  $V$ . Eine surjektive Abbildung  $f: U \rightarrow V$  liefert die Äquivalenzrelation:  $x \sim y \Leftrightarrow f(x) = f(y)$ . Ist  $f$  linear mit dem Kern  $L$ , so gilt also  $x \sim y \Leftrightarrow x - y \in L$ . Die Äquivalenzklassen sind die Urbilder von Punkten und haben die Form  $x + L = \{z \mid z = x + y, y \in L\}$ . Eine Teilmenge dieser Form heißt *affiner Unterraum* von  $U$  mit der *Richtung*  $L$ . Ist  $\dim L = 1$  so sprechen wir von einer affinen Geraden, ist die Kodimension von  $L$  in  $V$  gleich 1 von einer *affinen Hyperebene* (*Ebene* im Fall  $\dim V = 3$ ).

Wir drehen den Spieß nun um! Zu jedem Unterraum  $L$  von  $U$  haben wir die Äquivalenzrelation  $x \sim y \Leftrightarrow x - y \in L$  mit den Klassen der Form  $x + L$ . Die Menge der Klassen bezeichnen wir mit  $U/L$  (gelesen  $U$  modulo  $L$ ). Sei  $p: U \rightarrow U/L$  die *Quotientabbildung*, die jedem  $x$  seine Klasse  $x + L$  zuordnet. Dann gilt:

**(1.6.2) Notiz.** *Es gibt genau eine Vektorraumstruktur auf  $U/L$ , so daß  $p: U \rightarrow U/L$  eine lineare Abbildung wird.*

*Beweis.* Da  $p$  surjektiv ist, ist die Struktur eindeutig bestimmt; damit  $p$  linear wird, müssen wir die Addition nämlich durch

$$U/L \times U/L \rightarrow U/L, \quad (x + L, y + L) \mapsto (x + y) + L$$

definieren. Zu zeigen ist, daß damit eine wohldefinierte Abbildung gegeben ist. Das bedeutet: Aus  $x + L = x' + L$ ,  $y + L = y' + L$  folgt  $x + y + L = x' + y' + L$  (Unabhängigkeit von der Auswahl der Repräsentanten). Nun ist aber  $x + L = x' + L$  gleichwertig mit  $x - x' = a \in L$ , und damit folgt  $x + y + L = x' + a + y + L = x' + y + L$ , da  $a \in L$  ist. Ebenso muß die Skalarmultiplikation durch  $\lambda(x + L) = \lambda x + L$  definiert werden, und eine Rechnung wie eben  $\lambda x + L = \lambda(x' + a) + L = \lambda x' + \lambda a + L = \lambda x' + L$  zeigt die Unabhängigkeit vom Repräsentanten. Damit

haben wir die Verknüpfungen, also die Daten für eine Vektorraumstruktur, auf  $U/L$  definiert. Die Vektorraumaxiome folgen unmittelbar aus denen für  $U$ .  $\square$

Der Raum  $U/L$  heißt *Quotientenraum*<sup>6</sup> von  $U$  nach  $L$ . Die vorstehende Überlegung zeigt insbesondere, daß jeder Unterraum Kern einer linearen Abbildung ist.

Ist  $f: U \rightarrow V$  eine lineare Abbildung, die einen Unterraum  $U_0 \subset U$  in einen Unterraum  $V_0 \subset V$  abbildet, so gibt es genau eine lineare Abbildung  $F: U/U_0 \rightarrow V/V_0$ , genannt die durch  $f$  *induzierte*, die das Diagramm

$$\begin{array}{ccccc} U_0 & \xrightarrow{\subset} & U & \xrightarrow{p_U} & U/U_0 \\ \downarrow f_0 & & \downarrow f & & \downarrow F \\ V_0 & \xrightarrow{\subset} & V & \xrightarrow{p_V} & V/V_0 \end{array}$$

kommutativ macht ( $p_U, p_V$  Quotientabbildungen,  $f_0$  Einschränkung von  $f$ ). Die Voraussetzung besagt nämlich, daß  $U_0$  im Kern von  $p_V \circ f$  liegt, und deshalb läßt sich (1.2.5) anwenden.

**(1.6.3) Notiz.** Seien  $f_0$  und  $F$  im letzten Diagramm Isomorphismen. Dann ist auch  $f$  ein Isomorphismus.

*Beweis.* Sei  $f(x) = 0$ . Dann ist  $0 = p_V f(x) = F p_U(x)$ . Weil  $F$  injektiv ist, folgt  $p_U(x) = 0$ . Mithin liegt  $x$  in  $U_0$ . Dann ist aber  $f_0(x) = f(x) = 0$ , und da  $f_0$  injektiv ist, gilt  $x = 0$ . Also ist  $f$  injektiv.

Sei  $y \in V$  gegeben. Da  $F p_U$  surjektiv ist, gibt es  $x \in U$  mit  $F p_U(x) = p_V(y)$ . Also liegt  $y - f(x)$  im Kern von  $p_V$ , also in  $V_0$ . Da  $f_0$  surjektiv ist, gibt es  $z \in U_0$  mit  $f_0(z) = y - f(x)$ . Also ist  $y = f(x + z)$  und somit  $f$  auch surjektiv.  $\square$

Aus der letzten Notiz folgt durch Induktion nach  $r$ :

**(1.6.4) Satz.** Sei  $f: V \rightarrow W$  eine lineare Abbildung, seien  $0 = V_0 \subset V_1 \subset \dots \subset V_r = V$  und  $0 = W_0 \subset W_1 \subset \dots \subset W_r = W$  Unterräume und gelte  $f(V_j) \subset W_j$ . Dann induziert  $f$  Abbildungen  $f_j: V_j/V_{j-1} \rightarrow W_j/W_{j-1}$ ,  $1 \leq j \leq r$ . Sind alle diese Abbildungen Isomorphismen, so ist  $f$  ein Isomorphismus.  $\square$

Wir benutzten schließlich die Quotientenräume dazu, einen weiteren Beweis für die Existenz der Dimension zu führen. Dazu genügt es zu zeigen:

**(1.6.5) Satz.** Der Vektorraum  $V$  habe eine Basis aus  $n$  Elementen. Dann ist jede Menge mit  $n + 1$  Elementen linear abhängig.

<sup>6</sup>Manche sagen auch Quotientenraum; das wäre aber ein Raum, der aus Quotienten besteht.

*Beweis.* Induktion nach  $n$ . Ist  $n = 1$ ,  $b$  ein Basisvektor, und sind  $x \neq 0$ ,  $y \neq 0$  aus  $V$ , so gilt  $x = \lambda b$ ,  $y = \mu b$  mit  $\lambda \neq 0 \neq \mu$ , und  $\mu x - \lambda y = 0$  ist eine lineare Relation.

Wir nehmen nun an, daß der Satz für jeden Raum mit höchstens  $n-1$  Basis-elementen gilt. Sei  $(b_1, \dots, b_n)$  eine Basis von  $V$ , und sei  $S = \{x_1, \dots, x_{n+1}\} \subset V$  gegeben. Ist  $x_{n+1} = 0$ , so ist  $S$  linear abhängig. Sei also  $x_{n+1} \neq 0$  und  $L$  der von  $x_{n+1}$  aufgespannte Unterraum. Wir benutzen die Quotientabbildung  $p: V \rightarrow V/L$ . Da die  $p(b_1), \dots, p(b_n)$  den Raum  $V/L$  erzeugen, enthält diese Menge eine Basis (1.3.10). Es gibt eine Darstellung  $x_{n+1} = \sum_{j=1}^n \lambda_j b_j$ . Wegen  $0 = p(x_{n+1}) = \sum_j \lambda_j p(b_j)$  sind die  $p(b_j)$  linear abhängig, und deshalb hat  $V/L$  nach (1.3.10) eine Basis mit weniger als  $n$  Elementen. Nach Induktion gibt es daher eine echte Relation  $\sum_k \mu_k p(x_k) = 0$ . Das bedeutet:  $\sum_{k=1}^n \mu_k x_k$  liegt im Kern von  $p$  und hat somit die Form  $\mu_{n+1} x_{n+1}$ .  $\square$

## Ergänzungen und Aufgaben

**1.** Aus der Analysis ist der Begriff einer Cauchy-Folge bekannt. Sei  $C(\mathbb{Q})$  der Vektorraum aller Cauchy-Folgen  $(x_i \mid i \in \mathbb{N})$  aus rationalen Zahlen  $x_i$ . In der Menge der reellen Zahlen hat diese Folge einen Grenzwert. Indem man jeder Cauchy-Folge ihren Grenzwert zuordnet, erhält man eine surjektive lineare Abbildung  $C(\mathbb{Q}) \rightarrow \mathbb{R}$ , deren Kern  $N(\mathbb{Q})$  aus allen Nullfolgen besteht. Auf diese Weise wird der Quotientraum  $C(\mathbb{Q})/N(\mathbb{Q})$  isomorph zu  $\mathbb{R}$ . Man kann also, wenn man die reellen Zahlen noch nicht kennt, diese algebraisch als Quotientraum aus den rationalen Zahlen konstruieren. (Weitere Strukturen, wie etwa die Multiplikation, lassen sich aus dieser Konstruktion ebenfalls unschwer herleiten.)

**2.** Die Definition des Quotientraumes kann für manche anderen algebraischen Objekte imitiert werden. Sei  $G$  eine abelsche Gruppe und  $U$  eine Untergruppe, dann ist  $x \sim y \Leftrightarrow x - y \in U$  eine Äquivalenzrelation. Auf der Menge  $G/U$  der Klassen gibt es genau eine Struktur einer abelschen Gruppe, so daß die Quotientabbildung  $G \rightarrow G/U$  ein Homomorphismus ist. (Im Falle nicht-abelscher Gruppen muß man von  $U$  eine weitere Eigenschaft verlangen: Normalteiler.)

# Kapitel 2

## Matrizen

Eine Hauptaufgabe der linearen Algebra ist die geometrische und algebraische Untersuchung der linearen Abbildungen. Dieser Abschnitt ist der Beschreibung linearer Abbildungen durch Matrizen gewidmet. Eine Matrix ist eine Tabelle. Tabellen werden als rechteckiges Schema notiert. Sie bestehen aus Zeilen und Spalten. Eine Tabelle bietet auf kleinstem Raum eine Fülle an Information übersichtlich dar. In der Mathematik ist eine Matrix ein rechteckiges Zahlenschema. Die Matrizenrechnung ist ein Algorithmus mit außerordentlichen praktischen und theoretischen Möglichkeiten. Ein wichtiges methodisches Prinzip der linearen Algebra ist die Übersetzung geometrischer und begrifflicher Aussagen über lineare Abbildungen in algebraische, algorithmische und rechnerische Aussagen über Matrizen.

Das Rechnen mit Matrizen ist aber auch ein Kalkül, der in weiten Teilen unabhängig von Vektorräumen und linearen Abbildungen entwickelt werden kann. Doch bleibt dann Einiges ohne die begriffliche Deutung unklar oder unmotiviert.

### 2.1 Matrizen

Wir beginnen mit ein wenig Terminologie über die Datensysteme der Matrizen. Die Nützlichkeit dieser Begriffsbildungen wird sich alsbald erweisen.

**(2.1.1) Matrizen.** Eine  $(n, m)$ -Matrix mit *Einträgen* oder *Elementen*  $a_{kl}$  aus

dem Körper  $K$  wird notiert als rechteckiges Schema

$$(1.1) \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{pmatrix} = A.$$

Der erste Index ist der *Zeilenindex*, der zweite Index der *Spaltenindex*. Demgemäß nennen wir  $z_i = (a_{i1}, \dots, a_{im})$  die  $i$ -te *Zeile* der Matrix  $A$  und

$$s_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \dots \\ a_{nk} \end{pmatrix}$$

die  $k$ -te *Spalte* der Matrix  $A$ . Wir sprechen in diesem Kontext auch von *Zeilenvektoren*  $z_i$  und *Spaltenvektoren*  $s_k$ .

Formal mathematisch ist eine  $(n, m)$ -Matrix eine Abbildung

$$a: \{1, \dots, n\} \times \{1, \dots, m\} \rightarrow K$$

mit dem Wert  $a_{kl}$  an der Stelle  $(k, l)$ . Deshalb sprechen wir auch von einer  $n \times m$ -Matrix.

Natürlich kann man diese Definition mengentheoretisch ganz allgemein fassen: Eine Abbildung  $a: S \times T \rightarrow M$  ist eine  $S \times T$ -Matrix mit Werten in der Menge  $M$ . Ist speziell  $S \subset \{1, \dots, n\}$  und  $T \subset \{1, \dots, m\}$ , so heißt die eingeschränkte Abbildung  $a: S \times T \rightarrow K$  eine *Untermatrix* der  $n \times m$ -Matrix  $a: \{1, \dots, n\} \times \{1, \dots, m\} \rightarrow K$ . Überhaupt werden in der Praxis die Zeilen und Spalten nicht immer beginnend mit 1 durchnummeriert. Eine Untermatrix entsteht also durch Streichung von einigen Zeilen und Spalten.  $\diamond$

Wo treten nun Matrizen auf und warum sind sie nützlich? Wir nennen in Schlagworten einige Anwendungsbereiche.

- (1) Lineare Abbildungen.
- (2) Lineare Gleichungen.
- (3) Basiswechsel.
- (4) Koordinatenwechsel.
- (5) Bilineare Abbildungen.
- (6) Gruppentheorie.
- (7) Matrizenkalkül.

**(2.1.2) Matrix einer linearen Abbildung.** Sei  $f: V \rightarrow W$  eine lineare Abbildung zwischen Vektorräumen über dem Körper  $K$ . Wir wählen geordnete Basen  $B = \{v_1, \dots, v_m\}$  von  $V$  und  $C = \{w_1, \dots, w_n\}$  von  $W$ . Wie wir wissen,

ist eine lineare Abbildung  $f$  durch die Bildvektoren  $f(v_1), \dots, f(v_m)$  bestimmt. Wir schreiben diese als Linearkombinationen der Basisvektoren von  $w$

$$(1.2) \quad f(v_i) = \sum_{j=1}^n x_{ji} w_j, \quad x_{ji} \in K.$$

Man achte in dieser Formel auf die Reihenfolge der Indices.

Die Matrix

$$(1.3) \quad \begin{pmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1m} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & x_{n3} & \cdots & x_{nm} \end{pmatrix} = X = m_B^C(f).$$

heißt die *Matrix von  $f$  bezüglich  $B, C$* . Ist  $V = W$ , so arbeiten wir im allgemeinen mit gleichen Basen  $B = C$  und sprechen von der Matrix von  $f$  bezüglich  $B$ . Matrizen werden wir meist durch große Buchstaben bezeichnen.

Die Abbildung  $f$  ist durch die Matrix  $m_B^C(f)$  eindeutig bestimmt, wenn die Basen  $B$  und  $C$  einmal fixiert sind. Die  $k$ -te Spalte ist der Koordinatenvektor von  $f(v_k)$  bezüglich der Basis  $C$ . Hierin steckt, wie schon gesagt, eine Vereinbarung: Es wäre auch möglich, den Koordinatenvektor als  $k$ -te Zeile zu notieren.  $\diamond$

**(2.1.3) Räume von Abbildungen und Matrizen.** Auf der Menge  $M(n, m)$  der  $(n, m)$ -Matrizen über  $K$  definieren wir eine Addition und Skalarmultiplikation komponentenweise:

$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij}), \quad \lambda(a_{ij}) = (\lambda a_{ij}).$$

Damit wird  $M(n, m)$  zu einem Vektorraum der Dimension  $nm$ , denn  $M(n, m)$  ist offenbar im wesentlichen ein Standardvektorraum, nur in etwas anderer Schreibweise.

In diesem Kontext hat übrigens der Standardraum  $K^n$  zwei Erscheinungsformen, nämlich  $M(1, n)$  als „Zeilenvektoren“ und  $M(n, 1)$  als „Spaltenvektoren“. Wir werden künftig beide Versionen verwenden, darauf möchte man achten.

Wir hatten früher schon den Vektorraum  $\text{Hom}(V, W)$  der  $K$ -linearen Abbildungen erklärt. Beide Räume passen nun zusammen.

Die Abbildung

$$m_B^C: \text{Hom}(V, W) \rightarrow M(n, m)$$

ist ein Isomorphismus von Vektorräumen. Daraus folgt

$$\dim \text{Hom}(V, W) = \dim V \cdot \dim W.$$

Die Bijektivität ist nichts anderes als die Aussage, daß eine lineare Abbildung durch die Bilder der Basisvektoren bestimmt ist und diese Bilder beliebig

vorgeschrieben werden können. Die Verträglichkeit mit Addition und Skalarmultiplikation folgt unmittelbar durch Einsetzen der Definitionen.  $\diamond$

**(2.1.4) Matrizenprodukt.** Eine wichtige Konsequenz aus der Zuordnung einer Matrix zu einer linearen Abbildung ist eine Formel für die Matrix einer Verkettung von linearen Abbildung. Das Resultat nennen wir das Produkt der Matrizen.

Seien  $U, V, W$  Vektorräume mit Basen  $A = (u_1, \dots, u_m)$ ,  $B = (v_1, \dots, v_n)$ ,  $C = (w_1, \dots, w_p)$ . Sei  $g \in \text{Hom}(V, W)$  und  $f \in \text{Hom}(U, V)$ . Wir setzen

$$g(v_j) = \sum_k x_{kj} w_k, \quad f(u_i) = \sum_j y_{ji} v_j.$$

Diese Daten liefern

$$\begin{aligned} (g \circ f)(u_i) &= g\left(\sum_j y_{ji} v_j\right) = \sum_j y_{ji} g(v_j) = \sum_j y_{ji} \left(\sum_k x_{kj} w_k\right) \\ &= \sum_k \left(\sum_j x_{kj} y_{ji}\right) w_k. \end{aligned}$$

Deshalb ist  $m_A^C(g \circ f) = Z = (z_{ki})$  mit  $z_{ki} = \sum_j x_{kj} y_{ji}$ . Wir nennen diese Matrix  $Z$  das Produkt  $XY$  von  $(X, Y)$ .

Wir sammeln das Resultat in einem Diagramm.

$$\begin{array}{ccccc} (g, f) \in & \text{Hom}(V, W) \times \text{Hom}(U, V) & \longrightarrow & \text{Hom}(U, W) & \ni g \circ f \\ & \downarrow m_B^C \times m_A^B & & \downarrow m_A^C & \\ (X, Y) \in & M(p, n) \times M(n, m) & \longrightarrow & M(p, m) & \ni XY = Z. \end{array}$$

Das Produkt  $XY = (x_{kj})(y_{ji})$  der Matrizen  $X = (x_{kj}) \in M(p, n)$  und  $Y = (y_{ji}) \in M(n, m)$  ist die Matrix  $Z = (z_{ki}) \in M(p, m)$  mit den Einträgen

$$z_{ki} = \sum_{j=1}^n x_{kj} y_{ji}.$$

In Worten: An der Stelle  $(k, i)$  der Produktmatrix steht eine Summe, die sich aus der  $k$ -ten Zeile der ersten (= linken) und der  $i$ -ten Spalte der zweiten (= rechten) Matrix wie angegeben berechnet. Wir notieren ein Produkt von Matrizen auch als Produkt in der üblichen Weise

$$M(p, n) \times M(n, m) \rightarrow M(p, m), \quad (X, Y) \mapsto XY = X \cdot Y$$

und benutzen die Standardsymbole wie etwa  $A \cdot A = A^2$ .

$$\begin{array}{ccc}
 \begin{array}{|c|c|} \hline x_{i1} & x_{in} \\ \hline \end{array} & \times & \begin{array}{|c|} \hline y_{1j} \\ \hline y_{nj} \\ \hline \end{array} \\
 X & \times & Y \\
 \hline
 & = & \begin{array}{|c|} \hline z_{ij} \\ \hline \end{array} \\
 & = & Z
 \end{array}$$

Es sei darauf hingewiesen, daß das Produkt von Matrizen nur dann definiert ist, wenn sie der Größe nach zusammenpassen. Immer wenn wir ein Matrizenprodukt hinschreiben, so sei dieses Passen unterstellt.  $\diamond$

Hier sind einige einfache Beispiele für Matrizenprodukte.

$$(a \ b) \begin{pmatrix} c \\ d \end{pmatrix} = (ac + bd), \quad \begin{pmatrix} c \\ d \end{pmatrix} (a \ b) = \begin{pmatrix} ca & cb \\ da & db \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}$$

**(2.1.5) Notiz.** Die oben durchgeführte Rechnung besagt, daß Verkettung von linearen Abbildungen in das Produkt von Matrizen überführt wird, in Formeln:

$$m_C^D(g)m_B^C(f) = m_B^D(gf).$$

Verkettung von Abbildungen ist assoziativ, also ist das Matrizenprodukt assoziativ  $X(YZ) = (XY)Z$ .  $\square$

**(2.1.6) Bemerkung.** Es ist nützlich, sich folgendes klarzumachen: Die Spalten einer Produktmatrix  $AB$  bestehen aus Linearkombinationen der Spaltenvektoren von  $A$ . In der  $i$ -ten Spalte wird mit den Einträgen der  $i$ -ten Spalte von  $B$  linear kombiniert. Analog sind die Zeilen der Produktmatrix Linearkombinationen der Zeilen von  $B$ . Wir schreiben diese Aussage in Formeln hin. Sei  $Z_i$  die  $i$ -te Zeile und  $S_i$  die  $i$ -Spalte einer Matrix. Dann sind

$$\begin{aligned}
 Z_i(AB) &= \sum_k a_{ik} Z_k(B) \\
 S_i(AB) &= \sum_k b_{ki} S_k(A)
 \end{aligned}$$

die genannten Relationen.  $\diamond$

## Ergänzungen und Aufgaben

1. Die Multiplikation  $\mathbb{C} \rightarrow \mathbb{C}, w \mapsto zw$  hat als  $\mathbb{R}$ -lineare Abbildung bezüglich der Basis  $1, i$  die Matrix  $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ , wenn  $z = a+bi$  die Zerlegung in Real- und Imaginärteil ist.

2. Sei  $V$  ein komplexer Vektorraum und  $f: V \rightarrow V$  eine  $\mathbb{C}$ -lineare Abbildung. Sei  $L = (\lambda_{k\ell})$  die Matrix von  $f$  bezüglich einer  $\mathbb{C}$ -Basis  $v_1, \dots, v_n$  von  $V$ . Wir setzen  $\lambda_{k\ell} = \alpha_{k\ell} + i\beta_{k\ell}$  mit reellen  $\alpha_{k\ell}$  und  $\beta_{k\ell}$  an. Aus

$$f(v_\ell) = \sum_k \lambda_{k\ell} v_k = \sum_k \alpha_{k\ell} v_k + \sum_k \beta_{k\ell} (iv_k)$$

$$f(iv_\ell) = -\sum_k \beta_{k\ell} v_k + \sum_k \alpha_{k\ell} (iv_k)$$

erkennt man, daß  $f$ , aufgefaßt als  $\mathbb{R}$ -lineare Abbildung  $f_{\mathbb{R}}: V_{\mathbb{R}} \rightarrow V_{\mathbb{R}}$ , bezüglich der  $\mathbb{R}$ -Basis  $v_1, \dots, v_n, iv_1, \dots, iv_n$  die Matrix  $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$  mit  $A = (\alpha_{k\ell}), B = (\beta_{k\ell})$  hat.

3. Nachweis der Assoziativität des Matrizenprodukts ohne die Interpretation durch lineare Abbildungen direkt aus der Formel.

4. Ist das Produkt zweier  $(n, n)$ -Matrizen invertierbar, so auch jeder Faktor.

5. Aus der Formel

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}$$

folgere man, daß die erste Matrix genau dann invertierbar ist, wenn  $ad - bc \neq 0$  ist.

6. Hat eine lineare Abbildung die Matrix  $A$ , so hat die duale Abbildung bezüglich der Dualbasen die Matrix  $A^t$ .

7. Die Verkettung

$$\text{Hom}(V, W) \times \text{Hom}(U, V) \rightarrow \text{Hom}(U, W), \quad (g, f) \mapsto g \circ f = gf$$

ist bilinear: Es gelten die Verträglichkeiten mit der Vektorraumstruktur  $g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2$ ,  $(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f$ ,  $(\lambda g) \circ f = \lambda(g \circ f)$ ,  $g \circ (\lambda f) = \lambda(g \circ f)$ .

8. Das Matrizenprodukt

$$M(p, n) \times M(n, m) \rightarrow M(p, m), \quad (X, Y) \mapsto XY$$

ist bilinear: Es gelten die Rechenregeln  $X(Y_1 + Y_2) = XY_1 + XY_2$ ,  $(X_1 + X_2)Y = X_1Y + X_2Y$ ,  $(\lambda X)Y = X(\lambda Y) = \lambda(XY)$ .

9. Besonders in der physikalischen Literatur schreibt man die Formel (1.1) gerne in der Form  $f(v_i) = \sum_{j=1}^n x_i^j w_j$ . Der Grund dafür ist, daß man in komplizierteren Formeln dieser Art („Tensorrechnung“) die sogenannte *Summenkonvention* anwendet; sie besagt, daß über einen gleichen oberen und unteren Index summiert wird — und dadurch spart man sich das Summenzeichen.

Diese Schreibweise hat auch andere Vorteile. Sind nämlich  $v^i$  und  $w^j$  die Dualbasen von  $V^*$  und  $W^*$ , so hat die duale Abbildung die Form  $f^*(w^j) = \sum_i x_i^j v^i$ , und auch hier kann man die Summenkonvention anwenden.

Mit der Summenkonvention geschrieben ist zum Beispiel das Produkt der Matrizen  $(x_i^k)$  und  $(y_i^l)$  die Matrix mit dem Eintrag  $x_i^k y_i^l$  an der Stelle  $(k, l)$ .

## 2.2 Spezielle Matrizen

**(2.2.1) Invertierbare Matrizen.** Die identische Abbildung  $\text{id}: V \rightarrow V$  hat bezüglich jeder Basis  $B$  von  $V$  die *Einheitsmatrix*  $m_B^B(\text{id}) = I_n = (\delta_{ij})$ , deren Einträge  $\delta_{ii} = 1$  und  $\delta_{ij} = 0$  für  $i \neq j$  sind. Man nennt eine Familie  $\delta_{ij}$  mit diesen Eigenschaften manchmal *Kronecker-Symbol*<sup>1</sup>.

Eine  $(n, n)$ -Matrix (also eine *quadratische Matrix*)  $A$  heißt *invertierbar* oder *regulär*, wenn sie bezüglich der Matrizenmultiplikation eine inverse Matrix hat, wenn es also eine  $(n, n)$ -Matrix  $B$  so gibt, daß  $AB = BA$  die Einheitsmatrix  $I_n$  ist.

Ist  $f: V \rightarrow W$  ein Isomorphismus, gibt es also eine inverse Abbildung  $g: W \rightarrow V$ , so ist  $gf = \text{id}(V)$  und  $fg = \text{id}(W)$ . Bezüglich beliebiger Basen  $B$  und  $C$  ist deshalb die Matrix  $m_B^C(f)$  invertierbar.

Ist, umgekehrt  $m_B^C(f)$  invertierbar, so ist  $f$  ein Isomorphismus.  $\diamond$

**(2.2.2) Matrizengruppen.** Die regulären  $(n, n)$ -Matrizen über eine Körper  $K$  bilden bezüglich der Multiplikation eine Gruppe, die *allgemeine lineare Gruppe* genannt und  $GL(n, K)$  bezeichnet wird<sup>2</sup>. Das neutrale Element ist  $I_n$ . Die linearen Gruppen und einige ihrer Untergruppen gehören zu den wichtigsten Gegenständen der Mathematik. Sie sind von einer außerordentlichen Komplexität, und ihre inneren Geheimnisse sind noch lange nicht enträtselt. Der weitere Verlauf der linearen Algebra dient zum Teil dem Studium dieser Gruppen.  $\diamond$

**(2.2.3) Diagonalmatrizen.** Ist  $A = (a_{ij})$  eine  $(n, n)$ -Matrix, so heißen die Elemente  $a_{11}, a_{22}, \dots, a_{nn}$  die *Diagonalelemente* und ihre Gesamtheit bildet die *Diagonale* oder auch *Hauptdiagonale* der Matrix  $A$ . Es gibt noch eine zweite Diagonale, die weniger wichtig ist, deshalb *Nebendiagonale* heißt und aus den Elementen  $a_{ij}$  mit  $i + j = n + 1$  besteht. Sind alle Elemente einer Matrix außerhalb der Hauptdiagonale gleich Null, so heißt sie *Diagonalmatrix*. Wir schreiben  $\text{Dia}(d_1, \dots, d_n)$  für eine Diagonalmatrix  $(a_{ij})$  mit  $a_{ii} = d_i$ . Es gilt

$$\text{Dia}(a_1, \dots, a_n) \cdot \text{Dia}(b_1, \dots, b_n) = \text{Dia}(a_1 b_1, \dots, a_n b_n).$$

Die Nullabbildung, die jeden Vektor auf den Nullvektor wirft, besitzt die *Nullmatrix*, deren sämtliche Einträge Null sind.

Später werden wir uns mit dem Problem beschäftigen, zu einer linearen Abbildung durch geschickte Basiswahl eine möglichst einfache Matrix zu finden. Diagonalmatrizen werden unter diesem Gesichtspunkt als die einfachsten angesehen.

<sup>1</sup>Leopold Kronecker 1823 – 1891

<sup>2</sup>Die Bezeichnung kommt von General Linear

Eine Diagonalmatrix  $\text{Dia}(a_1, \dots, a_n)$  ist offenbar genau dann invertierbar, wenn alle  $a_i \neq 0$  sind, wie die Produktformel für Diagonalmatrizen lehrt. Die invertierbaren Diagonalmatrizen bilden eine Untergruppe  $D(n, K)$  von  $GL(n, K)$ .

Aus der Definition einer Matrix zu einer linearen Abbildung ist unmittelbar klar:

**(2.2.4) Notiz.** Genau dann hat  $f: V \rightarrow V$  bezüglich der Basis  $B$  eine Diagonalmatrix, wenn  $B$  aus Eigenvektoren von  $f$  besteht.  $\square$

**(2.2.5) Dreiecksmatrizen.** Eine  $(n, n)$ -Matrix  $B = (b_{ij})$  heißt *obere Dreiecksmatrix*, wenn die Elemente unterhalb der Diagonale (das heißt die Elemente  $b_{ij}$  mit  $i > j$ ) gleich Null sind. Analog wird eine *untere Dreiecksmatrix* erklärt. Das Produkt oberer (unterer) Dreiecksmatrizen ist wieder eine Matrix dieser Form.

Sei  $d$  die Abbildung, die jeder oberen Dreiecksmatrix die Diagonalmatrix mit derselben Diagonale zuordnet. Dann gilt  $d(AB) = d(A)d(B)$ ; eine invertierbare Matrix wird also auf eine invertierbare abgebildet. Folglich hat eine invertierbare obere Dreiecksmatrix auf der Diagonale keine Nullen.

Sei  $V$  ein  $n$ -dimensionaler Vektorraum. Eine *Fahne* in  $V$  ist ein System

$$0 = V_0 \subset V_1 \subset V_2 \subset \dots \subset V_{n-1} \subset V_n = V$$

von Unterräumen der Dimension  $\dim V_j = j$ . Zu einer solchen Fahne gibt es eine Basis  $\{b_1, \dots, b_n\}$ , so daß  $\{b_1, \dots, b_j\}$  eine Basis von  $V_j$  ist. Eine lineare Abbildung  $f: V \rightarrow V$  mit der Eigenschaft  $f(V_j) \subset V_j$  hat bezüglich dieser Basis eine obere Dreiecksmatrix. Der von  $f$  auf dem Quotientenraum  $V_j/V_{j-1}$  induzierte Endomorphismus ist die Multiplikation mit dem  $j$ -ten Diagonalelement. Sind alle diese Elemente von Null verschieden, so ist  $f$  nach (1.6.4) ein Isomorphismus, also die Dreiecksmatrix invertierbar. Das Inverse  $g$  hat wieder die Eigenschaft  $g(V_j) \subset V_j$  und wird deshalb ebenfalls durch eine obere Dreiecksmatrix beschrieben. Damit haben wir gezeigt: Die oberen Dreiecksmatrizen ohne Nullen auf der Diagonale bilden eine Gruppe.

Eine obere  $(n, n)$ -Dreiecksmatrix  $A$  habe auf der Diagonale nur Nullen. Dann ist  $A^n$  die Nullmatrix. Eine Matrix, für die eine Potenz die Nullmatrix ist, heißt *nilpotent*.  $\diamond$

**(2.2.6) Blockmatrizen.** Eine Matrix wird oft in kleinere Matrizen aufgeteilt, indem man sie durch senkrechte und waagerechte Linien in Rechtecke zerlegt. Die Teilmatrizen heißen dann die *Blöcke* der großen Matrix. Es sieht dann das Ergebnis so aus, als sei die große Matrix eine Matrix, deren Einträge selbst wieder Matrizen sind, nämlich die Blöcke. Betrachten wir eine direkte Summe

$f \oplus g: U \oplus U' \rightarrow V \oplus V'$  mit Basen  $B, B', C, C'$  in  $U, U', V, V'$ , so ergibt sich eine Blockzerlegung

$$m_{B \cup B'}^{C \cup C'}(f \oplus g) = \left( \begin{array}{c|c} m_B^C(f) & 0 \\ \hline 0 & m_{B'}^{C'}(g) \end{array} \right).$$

Ist  $U \subset V$  ein  $f$ -stabiler Unterraum und ergänzen wir eine Basis von  $U$  zu einer von  $V$ , so hat die Matrix von  $f$  eine Blockform

$$\begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}.$$

Sind allgemeiner  $U_1 \subset \dots \subset U_t$   $f$ -stabile Unterräume und ergänzen wir nacheinander Basen von  $U_1, U_2, \dots$ , so wird die Matrix von  $f$  eine obere Blockdreiecksmatrix. Aus (1.6.4) entnehmen wir, daß eine solche Matrix invertierbar ist, wenn die Diagonalblöcke invertierbar sind. Ein direkter Beweis durch Induktion nach der Anzahl der Blöcke ist eine instruktive Aufgabe. Auch mit Hilfe der Determinantentheorie kann man den Beweis führen.  $\diamond$

**(2.2.7) Transponierte Matrizen.** Werden in einer Matrix  $A$  Zeilen und Spalten vertauscht, so entsteht die *transponierte Matrix*  $A^t$  von  $A$ . Genauer: Ist  $A = (a_{ij})$  eine  $(m, n)$ -Matrix, so ist  $A^t$  eine  $(n, m)$ -Matrix, die an der Stelle  $(i, j)$  das Element  $a_{ji}$  trägt. Das Transponieren ist für viele Rechenprozesse nützlich. Es gelten die folgenden Rechenregeln:  $(A + B)^t = A^t + B^t$  und  $(AB)^t = B^t A^t$ . Da die Matrizenmultiplikation im allgemeinen nicht kommutativ ist, muß die letzte Regel sorgsam beachtet werden.  $\diamond$

## Ergänzungen und Aufgaben

1. Sei  $N$  eine nilpotente quadratische Matrix mit  $N^n = 0$ . Dann ist

$$(I - N)(I + N + N^2 + \dots + N^{n-1}) = I - N^n = I.$$

Also ist  $I - N$  invertierbar. Eine obere Dreiecksmatrix mit lauter Einsen auf der Diagonale ist also invertierbar. Ist  $B$  eine obere Dreiecksmatrix ohne Nullen auf der Diagonale, so gibt es eine Diagonalmatrix  $D$ , so daß  $DB$  nur Einsen auf der Diagonale hat. Also ist  $B$  invertierbar.

**2. Verallgemeinerte Blockmatrizen.** Sei  $U_1, \dots, U_m$  eine Familie von Vektorräumen mit der direkten Summe  $U = \bigoplus_s U_s$ . Dazu gibt es die kanonischen Injektionen  $i_k: U_k \rightarrow U$  und die kanonischen Projektionen  $p_l: U \rightarrow U_l$ . Ist  $V_1, \dots, V_n$  eine weitere Familie mit der direkten Summe  $V = \bigoplus_t V_t$ , so benutzen wir  $U$  und  $V$  als oberen Index, um die zu  $U$  und  $V$  gehörenden Injektionen und Projektionen zu unterscheiden. Die lineare Abbildung  $\varphi: U \rightarrow V$  besitzt die Teilabbildungen

$$\varphi_{lk} = p_l^V \circ \varphi \circ i_k^U: U_k \rightarrow V_l.$$

Aus diesen gewinnt man  $\varphi$  zurück

$$\varphi(u_1, \dots, u_m) = (\sum_j \varphi_{1j} u_j, \dots, \sum_j \varphi_{nj} u_j).$$

Die Wirkung von  $\varphi$  als Resultat der Teilabbildungen läßt sich nun ebenfalls als Matrizenprodukt schreiben

$$\begin{pmatrix} \varphi_{11} & \cdots & \varphi_{1m} \\ \vdots & \cdots & \vdots \\ \varphi_{n1} & \cdots & \varphi_{nm} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix},$$

worin aber die Einträge der Matrix jetzt nicht mehr Elemente des Körpers sind sondern lineare Abbildungen. Mit dieser Notation erhält auch die Verkettung von linearen Abbildungen dasselbe Aussehen wie ein Matrizenprodukt, es ist dann nämlich  $(\psi \circ \varphi)_{kl} = \sum_j \psi_{kj} \circ \varphi_{jl}$ .

Wenn wir speziell  $K^m$  als direkte Summe  $\bigoplus_{s=1}^m K$  ansehen, so wird  $\varphi_{lk}: K \rightarrow K$  formal gesehen eine lineare Abbildung, aber eine lineare Abbildung  $K \rightarrow K$  ist die Multiplikation mit einem Körperelement, und wir können dann ohne Schaden die Matrix  $\varphi_{lk}$  als eine Matrix mit Einträgen aus  $K$  ansehen.

## 2.3 Basiswechsel. Koordinatentransformation

Wir haben lineare Abbildungen durch Matrizen beschrieben. Dazu mußten wir in den beteiligten Vektorräumen Basen auswählen. Wir überlegen nun, wie sich die Matrixdarstellung einer linearen Abbildung ändert, wenn wir die Basen ändern. Für den gesamten Übersetzungsprozeß ist diese Abhängigkeit immer zu berücksichtigen. Die Basiswahl wird als etwas Zufälliges und Unwichtiges angesehen. Jedoch erlaubt die Freiheit der Basiswahl auf der rechnerischen Seite, durch geschickte Auswahl ein vorgegebenes Problem zu vereinfachen oder übersichtlicher zu gestalten.

**(2.3.1) Basiswechsel.** Ein Vektorraum mit einer geordneten Basis  $B$  sei vorgegeben. Wählen wir eine zweite Basis  $B'$ , so wollen wir vorübergehend  $B$  die alte und  $B'$  die neue Basis nennen. Wir schreiben auch  $(V, B)$  und  $(V, B')$  für diese Daten. Die Vektoren der alten Basis sind Linearkombinationen der neuen Basis, und wir notieren diesen Sachverhalt gemäß unserer Vereinbarungen

$$b_j = \sum_k w_{kj} b'_k.$$

Die dadurch bestimmte Matrix gehört zur identitischen Abbildung

$$m_B^{B'}(\text{id}) = (w_{kl}) = S.$$

Wir nennen sie die *Basiswechselmatrix* „Alt durch Neu“. Die Matrix ist invertierbar, und das Inverse beschreibt „Neu durch Alt“. Aus der Verkettungsformel (2.1.5) folgt

$$m_{B'}^{B''}(\text{id})m_B^{B'}(\text{id}) = m_B^{B''}(\text{id}).$$

Die zu einer linearen Abbildung  $f: (V, B) \rightarrow (W, C)$  gehörende Matrix ist  $m_B^C(f) = X$ . Wählen wir neue Basen  $B'$  und  $C'$ , so ändert sich die Matrix  $m_{B'}^{C'}(f) = X'$  gemäß

$$m_{B'}^{C'}(f) = m_{B'}^{C'}(\text{id}_W \circ f \circ \text{id}_V) = m_C^{C'}(\text{id}_W)m_B^C(f)m_{B'}^B(\text{id}_V).$$

Mit den Abkürzungen

$$T = m_C^{C'}(\text{id}_W) \quad \text{und} \quad S = m_{B'}^B(\text{id}_V)$$

erhalten wir also die Formel  $TXS^{-1} = X'$ .

Ist  $V = W$ , so benutzt man zweckmäßig in  $V$  und  $W$  dieselbe Basis. In diesem Fall hat man natürlich nur einen Basiswechsel zur Verfügung. Wir notieren gesondert diesen Spezialfall:

**(2.3.2) Satz (Basiswechsel).** *Mit der oben vereinbarten Sprechweise gilt: Ist  $f: V \rightarrow V$  ein Endomorphismus,  $X$  seine Matrix in der alten und  $Y$  seine Matrix in der neuen Basis, und ist  $T$  die Matrix, die die alten Basisvektoren durch die neuen ausdrückt, so gilt  $T^{-1}XT = Y$ .  $\square$*

Im folgenden werden wir bei der Untersuchung von Endomorphismen immer nur mit einer Basis arbeiten, wenn nicht ausdrücklich etwas anderes gesagt wird.  $\diamond$

**(2.3.3) Äquivalente Matrizen. Konjugierte (ähnliche) Matrizen.** Die vorstehenden Formeln kann man als Rechenprozesse mit Matrizen ansehen, ohne dabei an Basiswechsel zu denken. Darauf bezieht sich die folgende Terminologie.

Wir nennen  $(n, m)$ -Matrizen  $X$  und  $Y$  *äquivalent* oder, zur Verdeutlichung, *grob-äquivalent*, wenn es eine invertierbare  $(n, n)$ -Matrix  $A$  und eine invertierbare  $(m, m)$ -Matrix  $B$  gibt, so daß die Gleichung  $A^{-1}XB = Y$  besteht.

Zwei  $(n, n)$ -Matrizen  $X$  und  $Y$  heißen *ähnlich* oder *konjugiert*, wenn es eine invertierbare  $(n, n)$ -Matrix  $T$  gibt, so daß die Gleichung  $T^{-1}XT = Y$  besteht.

Äquivalenz ist eine Äquivalenzrelation auf  $(n, m)$ -Matrizen, Ähnlichkeit ist eine Äquivalenzrelation auf  $(n, n)$ -Matrizen.  $\diamond$

Zu einem Endomorphismus gehört also eine Konjugationsklasse von Matrizen. Die Mehrdeutigkeit, die durch die Basiswahl entsteht, haben wir jedenfalls algebraisch etwas in den Griff bekommen. Eine naheliegende Frage ist nun, ob man einer Matrix gewisse „Invarianten“ zuordnen kann, die nur von der

Konjugationsklasse abhängen. Wir präsentieren eine einfach zu berechnende Invariante, die sogenannte Spur. Ihre Bedeutung wird allerdings an dieser Stelle nicht erläutert.

**(2.3.4) Die Spur.** Die Summe der Diagonalelemente einer  $(n, n)$ -Matrix  $A = (a_{ij})$  heißt *Spur* der Matrix, in Zeichen:

$$\text{Sp}(A) = \sum_{j=1}^n a_{jj}.$$

Ist  $M_n(K)$  der Vektorraum der  $(n, n)$ -Matrizen über  $K$ , so ist die Spur eine lineare Abbildung

$$\text{Sp}: M_n(K) \rightarrow K.$$

Eine simple Rechnung zeigt für zwei  $(n, n)$ -Matrizen  $A$  und  $B$  die Relation

$$\text{Sp}(AB) = \text{Sp}(BA).$$

Es folgt: Konjugierte Matrizen haben dieselbe Spur. Damit ist die Spur auch für einen Endomorphismus definiert, nämlich als die Spur irgendeiner repräsentierenden Matrix (siehe (2.3.3)).  $\diamond$

**(2.3.5) Satz.** Sei  $T: M_n(K) \rightarrow K$  eine lineare Abbildung mit der Eigenschaft  $T(AB) = T(BA)$  für alle  $A, B \in M_n(K)$ . Dann ist  $T$  ein skalares Vielfaches von  $\text{Sp}$ .

*Beweis.* Sei  $E_{ij} \in M_n(K)$  die Matrix, die an der Stelle  $(i, j)$  eine 1 hat und sonst überall Nullen. Die  $E_{ij}$  bilden eine Basis von  $M_n(K)$ . Es gilt  $E_{ij}E_{kl} = \delta_{jk}E_{il}$ . Für  $i \neq l$  ist deshalb

$$E_{ij}E_{jl} - E_{jl}E_{ij} = E_{il}$$

und folglich  $T(E_{il}) = 0$ . Wegen

$$E_{ij}E_{ji} - E_{ji}E_{ij} = E_{ii} - E_{jj}$$

gilt außerdem  $T(E_{ii}) = T(E_{jj})$ . Damit ist der Wert von  $T$  auf  $(a_{ij})$  gleich  $T(a_{11}) \text{Sp}(a_{ij})$ .  $\square$

**(2.3.6) Koordinaten.** Hat man in einem Vektorraum  $V$  eine Basis  $B = \{b_1, \dots, b_m\}$  gewählt, so kann man die Elemente von  $V$  durch Koordinaten beschreiben; die Koordinaten von  $v \in V$  sind die Koeffizienten in einer Linearkombination  $v = \sum_{j=1}^m \lambda_j b_j$  durch die Basisvektoren. Wir wollen die Koordinaten für die weiteren Zwecke als Spaltenvektor auffassen. Der dadurch bestimmte lineare Isomorphismus

$$\kappa = \kappa_B: V \rightarrow K^m, \quad v \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$$

heißt *Koordinatenabbildung* von  $V$  bezüglich  $B$ . Wahl einer Basis entspricht so der Wahl eines Isomorphismus  $V \rightarrow K^m$ . Auch hier fragt man natürlich, wie sich die Koordinaten ändern, wenn eine andere Basis gewählt wird. Eine zweite Frage ist, wie eine lineare Abbildung durch die Koordinaten beschrieben wird. Zum letzteren erklären wir, wie lineare Abbildungen zwischen Standardvektorräumen mittels Matrizen hergestellt werden.  $\diamond$

**(2.3.7) Standardabbildung zu einer Matrix.** Der Vektorraum  $K^n$  hat die Standardbasis  $\{e_1, \dots, e_n\}$ . Deshalb haben wir auch eine Standardmethode, lineare Abbildungen zwischen diesen Räumen durch Matrizen zu beschreiben. Sei  $A = (a_{ij})$  eine  $(n, m)$ -Matrix mit Eintragungen aus  $K$ . Wir fassen die Elemente aus  $K^m$  als  $(m, 1)$ -Matrizen auf, das heißt als „Spaltenvektoren“. Für  $x \in K^m$  ist dann das Matrizenprodukt  $Ax$  definiert und ein „Spaltenvektor“ aus  $K^n$ . Die Abbildung

$$L: K^m \rightarrow K^n, \quad x \mapsto Ax$$

ist eine lineare Abbildung, und die Matrix von  $L$  bezüglich der Standardbasen ist  $A$  (oh Wunder!). Wir bezeichnen diese Abbildung deshalb besser mit  $A: K^m \rightarrow K^n$ , denn glücklicherweise haben das Abbildungssymbol  $A(x)$  und das Matrizenprodukt dasselbe Aussehen, und die Linearität der Abbildung führt äußerlich zu denselben Formeln wie das Matrizenrechnen, wie wir im letzten Abschnitt gesehen haben. Also noch einmal ausgebreitet hingeschrieben:  $A$  ist die lineare Abbildung

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j}x_j \\ \sum_{j=1}^m a_{2j}x_j \\ \dots \\ \sum_{j=1}^m a_{nj}x_j \end{pmatrix}.$$

Man merke sich die Regel: In den Spalten der Matrix stehen die Bilder der Standardbasisvektoren.  $\diamond$

**(2.3.8) Lineare Abbildungen durch Koordinaten.** Seien  $V$  und  $W$  Vektorräume mit Basen  $B$  und  $C$  wie vordem, und sei  $f: V \rightarrow W$  eine lineare Abbildung mit der zugehörigen Matrix  $m_B^C(f)$ . Wie eben gesehen, gehört zu dieser Matrix auch eine Abbildung der Standardvektorräume, und dieses ist nun genau die Beschreibung der linearen Abbildung durch die Koordinaten. Formal könne wir diese Aussage so formulieren: Das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow \kappa_B & & \downarrow \kappa_C \\ K^m & \xrightarrow{m_B^C(f)} & K^n \end{array}$$

ist kommutativ, das heißt es gilt  $m_B^C(f) \circ \kappa_B = \kappa_C \circ f$ .  $\diamond$

**(2.3.9) Koordinatentransformation.** Zusätzlich zum Vorhergehenden wählen wir Basen  $B'$  von  $V$  und  $C'$  von  $W$ . Das folgende Diagramm versammelt die uns zur Verfügung stehenden Daten.

$$\begin{array}{ccc}
 K^m & \xrightarrow{A'} & K^n \\
 \uparrow \kappa_{B'} & & \uparrow \kappa_{C'} \\
 V & \xrightarrow{f} & W \\
 \downarrow \kappa_B & & \downarrow \kappa_C \\
 K^m & \xrightarrow{A} & K^n
 \end{array}
 \qquad
 \begin{array}{l}
 A' = m_{B'}^{C'}(f) \\
 \\ \\ \\
 A = m_B^C(f)
 \end{array}$$

Wir setzen

$$\kappa_{B'} \circ \kappa_B^{-1} = S, \qquad \kappa_{C'} \circ \kappa_C^{-1} = T.$$

Das sind invertierbare lineare Abbildungen (Matrizen)  $S$  und  $T$ . Die Kommutativität des Diagramms besagt

$$A'S = TA \quad \text{und} \quad A' = TAS^{-1}.$$

Die Matrizen und linearen Abbildungen  $S$  und  $T$  heißen *Koordinatentransformationen*. Die Matrix  $S$  beschreibt nämlich, wie aus den Koordinaten bezüglich  $B$  diejenigen bezüglich  $B'$  entstehen (neu aus alt).

Oder, wenn wir explizit mit Elementen arbeiten: Ist  $b_i = \sum_k w_{ki} b'_k$  der Basiswechsel alt-aus-neu, so ist  $\lambda'_k = \sum_j w_{kj} \lambda_j$  die Koordinatentransformation neu-aus-alt. Die lineare Abbildung  $S$  wird also in Matrizenform durch die Matrix  $(w_{kj})$  beschrieben.  $\diamond$

**(2.3.10) Bemerkung.** Man muß zwischen einem Vektor aus  $K^n$  — ein  $n$ -Tupel — und seinen Komponenten bezüglich einer Basis — auch ein  $n$ -Tupel — begrifflich unterscheiden. Bezüglich der Standardbasis sieht man jedoch diesen Unterschied nicht! Ebenso unterscheiden muß man eine Matrix, die einen Basiswechsel beschreibt, von einer Matrix, die eine lineare Abbildung beschreibt.  $\diamond$

**(2.3.11) Rang einer Matrix.** Die Dimension des von den Zeilenvektoren einer Matrix  $A$  aufgespannten Unterraums nennen wir den *Zeilenrang* der Matrix. Analog wird der *Spaltenrang* definiert. Der Zeilenrang einer Matrix ist gleich dem Spaltenrang der transponierten Matrix.  $\diamond$

**(2.3.12) Satz.** *Beschreibt die Matrix  $A$  die lineare Abbildung  $f$ , so gilt*

$$\dim(\text{Bild } f) = \text{Spaltenrang } A.$$

*Beweis.* Wir betrachten das Diagramm aus (2.2.3). Die Bilder von  $f$  und von  $A$  haben dieselbe Dimension, weil  $\kappa_C$  einen Isomorphismus zwischen diesen Bildräumen vermittelt. Das Bild von  $A$  wird von den Spaltenvektoren erzeugt.  $\square$

**(2.3.13) Satz.** *Sei  $f: V \rightarrow W$  eine lineare Abbildung vom Spaltenrang  $r$ . Die Dimensionen von  $V, W$ , Kern  $f$  seien  $m, n, s$ . Dann gibt es Basen  $\{b_1, \dots, b_m\}$  von  $V$  und  $\{c_1, \dots, c_n\}$  von  $W$  mit der Eigenschaft:  $f(b_i) = c_i$  für  $1 \leq i \leq r$  und  $g(b_j) = 0$  für  $j > r$ . Bezüglich dieser Basen hat  $f$  die Matrix*

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

*Jede Matrix ist also zu einer dieser Form grob-äquivalent. Zwei  $(n, m)$ -Matrizen sind genau dann grob-äquivalent, wenn sie denselben Spaltenrang haben.*

*Beweis.* Eine Basis  $\{c_1, \dots, c_r\}$  des Bildes von  $f$  werde zu einer Basis von  $W$  ergänzt. Wir wählen  $\{b_1, \dots, b_r\}$ , so daß für  $1 \leq j \leq r$  die Relation  $f(b_j) = c_j$  besteht. Dann sind diese Vektoren linear unabhängig, und wir ergänzen sie durch eine Basis des Kerns von  $f$  zu einer Basis von  $V$ . Bezüglich dieser Basen hat dann die Matrix von  $f$  die genannte Form. Zur letzten Aussage bemerken wir, daß die grobe Äquivalenz den Spaltenrang nicht ändert, da ein Isomorphismus einen Unterraum auf einen gleichdimensionalen Unterraum abbildet.  $\square$

**(2.3.14) Satz.** *Für jede Matrix stimmen Zeilen- und Spaltenrang überein.*

*Beweis.* Ist  $A$  eine Matrix, so wählen wir invertierbare Matrizen  $P$  und  $Q$  so, daß  $PAQ$  die Normalform des letzten Satzes hat. In dieser Normalform sind offenbar Zeilen- und Spaltenrang gleich, da in Zeilen und Spalten nur Vektoren einer Standardbasis und Nullvektoren stehen. Die Matrizen  $PAQ$  und  $(PAQ)^t$  haben also denselben Spaltenrang. Es ist  $(PAQ)^t = Q^t A^t P^t$ , und  $Q^t$  und  $P^t$  sind wieder invertierbar, Multiplikation mit ihnen ändert also nicht den Spaltenrang.  $\square$

Wegen des letzten Satzes sprechen wir vom *Rang einer Matrix* schlechthin. Der Satz ist insofern überraschend, als die Zeilen- und Spaltenvektoren äußerlich wenig miteinander zu tun haben. Wir heben noch hervor:

**(2.3.15) Notiz.** *Eine  $(n, n)$ -Matrix ist genau dann regulär, wenn sie den Rang  $n$  hat.*  $\square$

## Ergänzungen und Aufgaben

1. Sei  $f: V \rightarrow W$  eine lineare Abbildung und  $f^*: W^* \rightarrow V^*$  die duale Abbildung. Die Aussage „Zeilenrang = Spaltenrang“ ist äquivalent zu „Rang  $f = \text{Rang } f^*$ “.
2.  $f: V \rightarrow W$  ist genau dann injektive (surjektiv), wenn die duale Abbildung  $f^*: W^* \rightarrow V^*$  surjektiv (injektiv) ist. Daraus folgert man, daß das Bild von  $f$  isomorph zum Bild von  $f^*$  ist.
3. Sei  $U \rightarrow V \rightarrow W$  exakt. Dann ist auch die duale Sequenz  $W^* \rightarrow V^* \rightarrow U^*$  exakt.

## 2.4 Lineare Gleichungen

Das „Lösen von Gleichungen“ ist ein Grundproblem der Mathematik. Innerhalb der linearen Algebra wird die Aufgabe gestellt,  $k$  lineare Gleichungen in  $n$  Unbekannten zu lösen. Was heißt das? Nachfolgend sind solche Gleichungen in einer systematischen Weise aufgeschrieben:

$$\begin{aligned} G_1: & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ G_2: & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ & \cdots \\ G_k: & a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k. \end{aligned}$$

Darin sind  $a_{11}, a_{12}, \dots, a_{kn}$  und  $b_1, \dots, b_k$  gegebene Elemente eines Körpers  $K$ . Eine „Lösung“ des Systems von Gleichungen ist ein Systemen  $(x_1, \dots, x_n)$  von  $x_j \in K$ , die die  $k$  Gleichungen  $G_1, \dots, G_k$  gleichzeitig „erfüllen“, für die also die aufgeschriebenen Daten eine Gleichheit von Elementen des Körpers sind. Der Index  $i$  in  $a_{ij}$  gibt an, in welcher Gleichung wir uns befinden, der Index  $j$  gibt an, vor welcher Unbekannten  $x_j$  das Körperelement steht. Es dürfen natürlich einige der Koeffizienten  $a_{ij}$  gleich Null sein; in der Praxis schreiben wir dann diesen Summanden gar nicht erst hin. Drei Punkte bedeuten „und so weiter“.

Sei  $L(G_1, \dots, G_k)$  die Lösungsmenge des Gleichungssystems. Wir beschreiben jetzt ein Verfahren nach Art der Schulmathematik zur Bestimmung der Lösungsmenge. Dazu ist keinerlei Theorie oder gelehrte Vorbildung nötig, lediglich etwas systematische „Buchführung“. Und natürlich gibt es das Verfahren schon lange vor unserer heutigen linearen Algebra.

Danach schauen wir uns das Verfahren genauer an und benutzen dazu die Sprache der Vektorräume und Matrizen.

**(2.4.1) Der Gaußsche Algorithmus.** Wir beginnen damit, aus den gegebenen Gleichungen neue so zu bilden, daß sich dabei die Lösungsmenge nicht ändert, symbolisch

$$L(G_1, \dots, G_k) = L(G'_1, \dots, G'_k).$$

Wir werden dann diese Umformungen geeignet solange wiederholen, bis wir zu einem Gleichungssystem gelangt sind, dem wir die Lösungsmenge unmittelbar ansehen. Hier nun *elementare Umformungen*, die die Lösungsmenge nicht ändern:

- (1) Vertauschen der Reihenfolge von Gleichungen.
- (2) Multiplikation einer Gleichung mit einer von Null verschiedenen Zahl  $\lambda$ ; symbolisch:

$$L(G_1, \dots, G_k) = L(\lambda G_1, \dots, G_k)$$

bei Multiplikation von  $G_1$  mit  $\lambda \neq 0$ .

- (3) Addition des  $\lambda$ -fachen einer Gleichung zu einer anderen; symbolisch:

$$L(G_1, G_2, \dots) = L(G_1, \lambda G_1 + G_2, \dots)$$

bei Addition von  $\lambda G_1$  zu  $G_2$ .

Da wir alle diese Schritte zurückrechnen können, ändert sich bei ihnen die Lösungsmenge nicht.

Wir formen jetzt das System schrittweise um, indem wir nacheinander auf die Koeffizienten von  $x_1, x_2, \dots$  achten. Sind alle Koeffizienten  $a_{j1}$  von  $x_1$  gleich Null, so betrachten wir  $x_1$  als erledigt. Andernfalls sei etwa  $a_{j1} \neq 0$ . Wir setzen dann die Gleichung  $G_j$  an die erste Stelle, sodann machen wir nach (2) dort den Koeffizienten von  $x_1$  zu 1 und subtrahieren nach (3) geeignete Vielfache der nunmehr ersten Gleichung von den folgenden, um dortselbst den Koeffizienten von  $x_1$  zu Null zu machen. Jetzt lassen wir die erste Gleichung in Ruhe und behandeln die weiteren Gleichungen nach demselben Verfahren, aber beginnend mit  $x_2$ , da ja  $x_1$  nicht mehr „vorhanden“ ist. Schließlich erreichen wir ein System, bei dem die letzte Gleichung die folgende Form hat: Entweder

$$x_j + b_{k,j+1}x_{j+1} + \dots + b_{k,n}x_n = c_k,$$

für ein  $j \in \{1, \dots, n\}$  oder eventuell  $0 = c_k$ .

Im letzteren Fall ist für  $c_k \neq 0$  ein Widerspruch erreicht, mit anderen Worten: Es gibt keine Lösung. Beispielsweise widersprechen sich offenbar die Gleichungen  $G_1 : x_1 + x_2 = 1$  und  $G_2 : x_1 + x_2 = 2$  und nach dem Verfahren würden wir  $0 \cdot x_1 + 0 \cdot x_2 = 1$  erhalten.

Im ersten Fall können wir  $x_{j+1}, \dots, x_n$  beliebig vorgeben und dann  $x_j$  durch die Gleichung ausrechnen. Nun steigen wir die Gleichungen von unten nach oben auf, setzen die schon festgelegten Werte der  $x_i$  ein und rechnen das jeweils am Anfang stehende  $x_i$  durch die restlichen aus, wobei gegebenenfalls vorher noch nicht festgelegte  $x_r$  willkürlich gewählt werden dürfen.

Wir haben ein Rechenverfahren zur Lösung linearer Gleichungssysteme beschrieben. Schematische Rechenverfahren bezeichnet man als *Algorithmus*. Unser Verfahren wird manchmal *Eliminationsverfahren* genannt, weil nacheinander die Unbekannten  $x_j$  beseitigt (= eliminiert) werden. Das Verfahren wird auch *Gaußscher<sup>3</sup> Algorithmus* genannt.  $\diamond$

<sup>3</sup>Carl Friedrich Gauß 1777 – 1855

**(2.4.2) Beispiel.** Ein Beispiel mag die Methode verdeutlichen (etwa  $K = \mathbb{R}$ ). Angenommen, das Endergebnis der elementaren Umformungen habe die Form:

$$\begin{aligned}x_1 + x_2 + 3x_3 + x_4 + x_6 &= 0 \\x_2 - x_3 - x_5 - 2x_6 &= 1 \\x_5 &= 2.\end{aligned}$$

Damit ist zunächst  $x_5$  festgelegt. Wir steigen auf, setzen  $x_5 = 2$  ein, wählen  $x_3, x_4, x_6$  willkürlich und rechnen  $x_2$  aus. Wir steigen zur ersten Gleichung auf, haben  $x_2$  bis  $x_6$  schon festgelegt und rechnen  $x_1$  aus. Die Lösungsmenge ist also

$$(-(3 + 4x_3 + x_4 + 3x_6), 3 + x_3 + 2x_6, x_3, x_4, 2, x_6)$$

mit beliebigen  $x_3, x_4, x_6 \in \mathbb{R}$ . (Ein dreidimensionaler affiner Unterraum des  $\mathbb{R}^6$ .)  $\diamond$

**(2.4.3) Beispiel.** Wir erinnern an die Elementargeometrie der Ebene  $\mathbb{R}^2$ . Die Punkte  $(x, y) \in \mathbb{R}^2$ , die einer Gleichung  $ax + by = c$  genügen, füllen eine Gerade aus (wenn  $a$  und  $b$  nicht beide Null sind). Hat man eine zweite Gleichung dieser Art, so gibt es eine gemeinsame Lösung genau dann, wenn sich die Geraden schneiden, der Schnittpunkt ist die eindeutig bestimmte Lösung. Allerdings können die Geraden auch parallel verlaufen, dann gibt es keine Lösung. Formelmäßig tritt dieser Fall auf, wenn sich die Gleichungen widersprechen. Wenn man schließlich drei Gleichungen hat, so werden die geometrischen Möglichkeiten schon unübersichtlich.

Analoge Diskussion der geometrischen Lage von drei Ebenen im Raum  $\mathbb{R}^3$  als Aufgabe.  $\diamond$

Die vorstehende Überlegung hat eine wichtige Konsequenz qualitativer Art. Ein Gleichungssystem wie anfangs vorgestellt, bei dem alle  $b_j$  gleich Null sind, heißt *homogen*. Elementare Umformungen machen aus einem homogenen System wieder ein solches. Ein homogenes System führt niemals zu Widersprüchen, denn alle  $x_j = 0$  gesetzt ist eine Lösung. Angenommen nun, es gibt in einem homogenen System weniger Gleichungen als Unbekannte (also  $k < n$ ). Dann können wir in der Lösungsmenge ein geeignetes  $x_j$  beliebig vorgeben; es gibt also eine Lösung, die nicht nur aus Nullen besteht. Wir haben eingesehen:

**(2.4.4) Satz.** *Ein homogenes lineares Gleichungssystem aus  $k$  Gleichungen zwischen  $n$  Unbekannten hat im Fall  $k < n$  eine Lösung, die nicht nur aus Nullen besteht.*  $\square$

Ich habe schon früher erläutert, wie man aus diesem Resultat den Rangsatz herleiten kann. Und das ist auch der Beweis, der sich ohne Idee von selbst ergibt und den man sich deshalb merken soll.

Wir werden nun den Algorithmus formaler beschreiben, genauer analysieren und auf weitere Probleme der linearen Algebra anwenden.

**(2.4.5) Gleichungssysteme matrizentheoretisch.** Lineare Gleichungssysteme lassen sich mit Hilfe der Matrizenmultiplikation in Kurzform schreiben. Sei  $A = (a_{ij})$  eine  $(m, n)$ -Matrix über  $K$  und  $b \in K^m$  ein Spaltenvektor. Das System linearer Gleichungen

$$\sum_{j=1}^n a_{ij}x_j = b_i \quad \text{für } 1 \leq i \leq m$$

schreiben wir dann  $Ax = b$  und nennen  $A$  die *Koeffizientenmatrix* des Systems. Das System heißt *homogen*, wenn  $b = 0$  ist, andernfalls *inhomogen*. Wir fassen  $A$  als lineare Abbildung auf

$$L_A: K^n \rightarrow K^m, \quad x \mapsto Ax.$$

Die Gleichung zu lösen, heißt dann, das Urbild  $L_A^{-1}(b)$  zu bestimmen. Danach wird die Theorie der Gleichungen ein Spezialfall der Theorie der linearen Abbildungen und Matrizen.

Sei zunächst  $b = 0$ , liege also ein homogenes System vor. Dann ist die Lösungsmenge  $L_A^{-1}(0) = U \subset K^n$  ein Unterraum, der Kern von  $L_A$ . Im inhomogenen Fall ist das System genau dann lösbar, wenn  $b$  im Bild von  $L_A$  liegt. Ist  $y$  irgendeine Lösung des Systems, so ist

$$\{y + u \mid u \in \text{Kern } L_A\} = y + U$$

die Lösungsgesamtheit. Wir haben eine Menge dieser Form einen *affinen Unterraum* von  $K^n$  mit der *Richtung*  $U$  genannt. Genau dann liegt  $b$  im Bild von  $L_A$ , wenn  $b$  eine Linearkombination der Spaltenvektoren von  $A$  ist; und letzteres ist genau dann der Fall, wenn die Matrix  $A$  und die sogenannte *erweiterte Matrix*

$$(A, b) = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

denselben Rang haben. Interessant ist der Fall  $m = n$ : Ist  $A$  regulär, so hat  $Ax = b$  genau eine Lösung  $x = A^{-1}b$ .  $\diamond$

**(2.4.6) Eliminationsverfahren und Elementarmatrizen.** Das Eliminationsverfahren zur Lösung linearer Gleichungen hat eine Interpretation durch Matrizen. Sei  $E_{ij}$  die Matrix, die an der Stelle  $ij$  eine 1 und sonst überall Nullen hat. Eine  $(m, m)$ -Matrix der Form

$$I_m + \lambda E_{ij}, \quad i \neq j, \lambda \in K$$

heißt *Elementarmatrix*. Sei  $A$  eine  $(m, n)$ -Matrix mit der  $k$ -ten Zeile  $a_k$ . Das Produkt  $(I_m + \lambda E_{ij})A$  ist eine Matrix mit den folgenden Zeilen: Die  $i$ -te Zeile ist  $a_i + \lambda a_j$  und die  $k$ -te Zeile für  $k \neq i$  ist  $a_k$ . Die Elementarmatrix ist invertierbar

und hat das Inverse  $I_m - \lambda E_{ij}$ . Eine invertierbare Diagonalmatrix wollen wir ebenfalls *Elementarmatrix* nennen. Das Produkt  $\text{Dia}(\lambda_1, \dots, \lambda_m)A$  hat  $\lambda_i a_i$  als  $i$ -te Zeile. Durch Multiplikation von links mit geeigneten Elementarmatrizen können wir also die folgenden *Zeilenumformungen* realisieren:

- (1) Multiplikation einer Zeile mit einem Skalar  $\lambda \neq 0$ .
- (2) Addition einer Zeile zu einer anderen.

Die folgende Sequenz von Umformungen dieser Art zeigt, daß dann auch Zeilenvertauschungen bewirkt werden können:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \begin{pmatrix} a_1 + a_2 \\ a_2 \end{pmatrix}, \quad \begin{pmatrix} a_1 + a_2 \\ -a_1 \end{pmatrix}, \quad \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}, \quad \begin{pmatrix} a_2 \\ a_1 \end{pmatrix}.$$

Multiplikation von links mit  $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  vertauscht bei einer  $(2, n)$ -Matrix die Zeilen. Die Sequenz wird nach dem Genannten durch Multiplikation von links mit

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

bewirkt.

Da Zeilenumformungen durch Produktbildung mit invertierbaren Matrizen bewirkt werden, ändert sich dabei nicht der Rang einer Matrix.  $\diamond$

**(2.4.7) Zeilenstufenform einer Matrix.** Das Endergebnis des Eliminationsverfahrens ist eine Art Normalform für Matrizen. Eine Matrix kann durch Zeilenumformungen in *Zeilenstufenform* (siehe unten) überführen kann. Die Spalten mit den Nummern  $(j_1, \dots, j_k)$  heißen die *Stufen*. Unter der „Treppe“ stehen Nullen. Die Fragezeichen sind nicht näher bezeichnete Elemente. Wird die Matrix  $A$  des Systems auf Zeilenstufenform  $S$  gebracht, so haben  $Ax = 0$  und  $Sx = 0$  denselben Lösungsraum. Es gibt ein Produkt  $U$  von Elementarmatrizen, so daß  $UA = S$  ist. Aus  $Ax = b$  wird also  $Sx = UAx = Ub$ , das heißt man muß den Spaltenvektor  $b$  denselben Zeilenumformungen unterwerfen, um das neue (zum alten äquivalente) inhomogene System zu erhalten.  $\diamond$

$$\begin{array}{c}
 \left( \begin{array}{cccc|cccc|cc}
 0 & \dots & 0 & 1 & ??? & 0 & ??? & 0 & ??? & 0 & ??? \\
 & & & 0 & & 1 & ??? & 0 & ??? & 0 & ??? \\
 & & & & & & & 1 & ??? & 0 & ??? \\
 & & & & & & & & & \cdot & \dots \\
 & & & & & & & & & \cdot & \dots \\
 & & & & & & & & & 0 & ??? \\
 & & & & & & & & & 1 & ??? \\
 0 & \dots & 0 & & & & & & & 0 & 0
 \end{array} \right) \\
 \begin{array}{cccc}
 j_1 & & j_2 & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 j_k & & & 
 \end{array}
 \end{array}$$

Zeilenstufenform

**(2.4.8) Notiz.** Die ausgeschriebene Zeilenstufenmatrix hat den Rang  $k$ . Die Spalten mit den Nummern  $(j_1, \dots, j_k)$  bilden eine Basis des von den Spaltenvektoren erzeugten Raumes.  $\square$

**(2.4.9) Satz.** Eine Matrix ist genau dann invertierbar, wenn sie sich durch Zeilenumformungen in die Einheitsmatrix überführen läßt, denn die Einheitsmatrix ist die einzige invertierbare Zeilenstufenmatrix. Eine invertierbare Matrix ist Produkt von Elementarmatrizen.  $\square$

**(2.4.10) Korollar.** Eine obere (untere) Dreiecksmatrix ist genau dann invertierbar, wenn alle Diagonalelemente von Null verschieden sind.  $\square$

Der letzte Satz liefert:

**(2.4.11) Verfahren zur Berechnung der inversen Matrix.** Man schreibe nebeneinander  $A \mid I_m$ , bringe  $A$  durch Zeilenumformungen auf Zeilenstufenform und wende dieselben Zeilenumformungen auf  $I_m$  an; das letztere liefert  $A^{-1}$ .  $\diamond$

**(2.4.12) Satz.** Eine  $(m, n)$ -Matrix  $A = (a_{ij})$  hat genau dann den Rang größer-gleich  $k$ , wenn es

$$1 \leq i_1 < i_2 < \dots < i_k \leq m, \quad 1 \leq j_1 < j_2 < \dots < j_k \leq n$$

gibt, so daß die  $(k, k)$ -Matrix

$$\tilde{A} = (a_{ij} \mid i \in \{i_1, \dots, i_k\}, j \in \{j_1, \dots, j_k\})$$

regulär ist. Der Rang von  $A$  ist also das maximale  $k$ , so daß aus  $A$  durch Streichen von Zeilen und Spalten eine reguläre  $(k, k)$ -Untermatrix entsteht.

*Beweis.* Hat  $\tilde{A}$  den Rang  $k$ , also  $k$  linear unabhängige Spalten, so sind erst recht die Spalten zum Index  $(j_1, \dots, j_k)$  von  $A$  linear unabhängig, und deshalb hat  $A$  mindestens den Rang  $k$ . Ist umgekehrt der Rang mindestens  $k$ , so wählen wir  $(j_1, \dots, j_k)$  so aus, daß die zugehörigen Spalten linear unabhängig sind, streichen die übrigen Spalten und wählen dann  $(i_1, \dots, i_k)$  so, daß bei der verbliebenen Matrix die Zeilen zu den Indices  $i_1, \dots, i_k$  linear unabhängig sind.  $\square$

**(2.4.13) Auswahl einer Basis.** Seien  $(v_1, \dots, v_n)$  Vektoren des  $K^m$ . Sie erzeugen einen Unterraum  $V$ . Wie berechnet man eine Basis von  $V$ ? Wir schreiben die Vektoren  $(v_1, \dots, v_n)$  als *Spalten* einer Matrix  $A$  auf und bringen diese durch *Zeilenumformungen* auf Zeilenstufenform  $S$ , wie oben ausgebreitet. Seien  $(j_1, \dots, j_k)$  die Stufen. Dann ist  $\{v_j \mid j \in \{j_1, \dots, j_k\}\}$  eine Basis von  $V$ . Der Beweis ergibt sich so: Es ist  $A = US$  mit einem Produkt  $U$  von Elementarmatrizen. Ist  $s_k$  die  $k$ -te Spalte von  $S$ , so gilt  $v_k = Us_k$ . Jetzt wenden wir (2.4.8) an: Eine Basis des Spaltenraums von  $S$  wird durch  $U$  in eine Basis des Spaltenraums von  $A$  transportiert.  $\diamond$

**(2.4.14) Berechnung einer Basis des Lösungsraumes.** Sei  $A$  gegeben und  $S$  Zeilenstufenform von  $A$  mit den Stufen  $(j_1, \dots, j_k)$ . Dann ist  $Ax = 0$  äquivalent zu  $Sx = 0$ . Sei  $B$  die Matrix, die aus  $S$  durch Weglassen der Spalten zum Index  $(j_1, \dots, j_k)$  und der letzten  $n - k$  Zeilen entsteht. Die verbleibenden Spalten mögen die Nummern  $r_1, \dots, r_{n-k}$  tragen, das heißt es ist  $\{j_1, \dots, j_k\} \cup \{r_1, \dots, r_{n-k}\} = \{1, \dots, n\}$ . Dann ist  $Sx = 0$  äquivalent zu

$$\begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_k} \end{pmatrix} = -B \begin{pmatrix} x_{r_1} \\ \vdots \\ x_{r_{n-k}} \end{pmatrix}.$$

Setzt man rechts die Standardbasis des  $K^{n-k}$  ein, so erhält man eine Basis des Lösungsraums.  $\diamond$

Alles bisher über Zeilen Gesagte gilt mutatis mutandis für Spalten.

## Ergänzungen und Aufgaben

1. Die Zeilenstufenmatrix zu einer Matrix  $A$  ist durch den Zeilenraum von  $A$  eindeutig bestimmt.

*Beweis.* Sei  $U$  dieser Unterraum, und sei  $p_j: U \rightarrow K^j$ ,  $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_j)$  die Projektion auf die ersten  $j$  Koordinaten. Dann sind die Stufen an denjenigen Stellen  $j$ , wo  $\dim p_j(U) > \dim p_{j-1}(U)$  ist. Also ist die Sequenz  $(j_1, \dots, j_k)$  durch  $U$  bestimmt. Die ersten  $k$  Zeilen der großen Matrix  $S$  oben bilden eine Basis von  $U$ . Sind  $(a_1, \dots, a_k)$  bzw.  $(b_1, \dots, b_k)$  Basen von  $U$  in Zeilenstufenform und ist  $a_r = \sum_i \lambda_i b_i$ , so ist  $\lambda_i$  der Koeffizient von  $a_r$  an der Stelle  $j_i$ . Das ist aber  $\delta_{ri}$ . Also ist  $a_r = b_r$ .

# Kapitel 3

## Determinanten

### 3.1 Determinantenformen

**(3.1.1) Determinantenform.** Sei  $V$  ein  $n$ -dimensionaler Vektorraum über dem Körper  $K$  und  $V \times V \times \dots \times V = V^n$  das  $n$ -fache cartesische Produkt von  $V$  mit sich, also die Menge aller  $n$ -Tupel  $(v_1, \dots, v_n)$ ,  $v_i \in V$ . Eine Abbildung  $D: V^n \rightarrow K$  heißt  $n$ -linear oder  $n$ -Linearform, wenn sie in jeder Variablen linear ist, wenn also immer gilt:

$$\begin{aligned} D(v_1, \dots, v_i + v'_i, \dots) &= D(v_1, \dots, v_i, \dots) + D(v_1, \dots, v'_i, \dots) \\ D(v_1, \dots, \lambda v_i, \dots) &= \lambda D(v_1, \dots, v_i, \dots). \end{aligned}$$

Eine  $n$ -Linearform  $D$  heißt *alternierend*, wenn aus  $v_i = v_j$  für  $i \neq j$  immer  $D(v_1, \dots, v_n) = 0$  folgt. Sei  $\text{Alt}^n(V)$  die Menge der alternierenden  $n$ -Linearformen auf dem  $n$ -dimensionalen Vektorraum  $V$ . Sie trägt eine Vektorraumstruktur durch Addition und skalare Multiplikation der Funktionswerte

$$(\lambda_1 D_1 + \lambda_2 D_2)(v_1, \dots, v_n) = \lambda_1 D(v_1, \dots, v_n) + \lambda_2 D_2(v_1, \dots, v_n).$$

Ein von Null verschiedenes Element aus  $\text{Alt}^n(V)$  heißt *Determinantenform* auf dem Vektorraum  $V$ .  $\diamond$

**(3.1.2) Satz.** Sei  $D$  eine Determinantenform auf  $V$ . Dann gilt:

- (1) Ist  $i \neq k$  und  $\lambda \in K$ , so gilt

$$D(\dots, v_i, \dots, v_k \dots) = D(\dots, v_i + \lambda v_k, \dots, v_k, \dots).$$

(An den punktierten Stellen stehen jeweils dieselben Vektoren.)

- (2) Werden zwei Vektoren vertauscht, so ändert sich das Vorzeichen.  
(3) Sind  $v_1, \dots, v_n$  linear abhängig, so ist  $D(v_1, \dots, v_n) = 0$ .

*Beweis.* (1) Wir wenden auf die rechte Seite die Linearität in der  $i$ -ten Variablen an und benutzen dann, daß  $D$  alternierend ist.

(2) Wir schreiben der Einfachheit halber nur zwei Variable hin und rechnen wie folgt:  $0 = D(v_1+v_2, v_1+v_2) = D(v_1, v_1) + D(v_1, v_2) + D(v_2, v_1) + D(v_2, v_2) = D(v_1, v_2) + D(v_2, v_1)$ .

(3) Sei etwa  $v_1 = \sum_{j=2}^n \lambda_j v_j$ . Linearität in der ersten Variablen liefert

$$D(v_1, \dots, v_n) = \sum_{j=2}^n \lambda_j D(v_j, v_2, \dots, v_n).$$

Da  $D$  alternierend ist, so ist jeder Summand rechts gleich Null.  $\square$

Für den ersten Existenzbeweis für Determinantenformen verwenden wir das Vorzeichen oder Signum einer Permutation.

**(3.1.3) Permutationen.** Eine bijektive Abbildung einer Menge  $M$  heißt *Permutation* von  $M$ . Wir bezeichnen mit  $S_n$  die Menge der Permutationen von  $\{1, \dots, n\}$ . Bezüglich Verkettung ist  $S_n$  eine Gruppe, genannt *symmetrische Gruppe* oder *Permutationsgruppe*.

Sei  $P$  die Menge der Paare  $(i, j)$ ,  $i \neq j$ ,  $1 \leq i, j \leq n$ . Ein *Halbsystem*  $H$  in  $P$  sei eine Teilmenge, die von den Paaren  $(i, j)$  und  $(j, i)$  jeweils genau eines enthält. Für  $\sigma \in S_n$  ist

$$\varepsilon(\sigma) = \prod_{(i,j) \in H} \frac{\sigma(i) - \sigma(j)}{i - j} \in \{\pm 1\}$$

unabhängig vom Halbsystem, denn wird  $(i, j)$  durch  $(j, i)$  ersetzt, so ändert sich im Zähler und Nenner das Vorzeichen. Ist  $H$  ein Halbsystem, so ist auch  $\tau(H) = \{(\tau(i), \tau(j)) \mid (i, j) \in H\}$  ein Halbsystem. Es gilt

$$\begin{aligned} \varepsilon(\sigma)\varepsilon(\tau) &= \prod_{(i,j) \in H} \frac{\sigma(i) - \sigma(j)}{i - j} \cdot \prod_{(i,j) \in H} \frac{\tau(i) - \tau(j)}{i - j} \\ &= \prod_{(i,j) \in H} \frac{\sigma\tau(i) - \sigma\tau(j)}{\tau(i) - \tau(j)} \cdot \prod_{(i,j) \in H} \frac{\tau(i) - \tau(j)}{i - j} \\ &= \varepsilon(\sigma\tau). \end{aligned}$$

Diese Formel besagt, daß  $\varepsilon: S_n \rightarrow \{\pm 1\}$  ein Homomorphismus in die multiplikative Gruppe  $\{\pm 1\}$  ist.

Wie nennen  $\varepsilon(\sigma)$  das *Vorzeichen* oder das *Signum* der Permutation  $\sigma$ . Eine Permutation heißt *gerade*, wenn ihr Vorzeichen gleich 1 ist, andernfalls *ungerade*. Die geraden Permutationen bilden eine Untergruppe von  $S_n$ , die *alternierende Gruppe*  $A_n$ . Eine Permutation, die genau zwei Elemente vertauscht und die anderen festläßt, eine sogenannte *Transposition*, hat das Vorzeichen  $-1$ . Da sich jede Permutation als Verkettung von Transpositionen schreiben läßt, so ist durch die letzte Eigenschaft das Vorzeichen eindeutig bestimmt.  $\diamond$

**(3.1.4) Satz.** *Der Vektorraum  $\text{Alt}^n(V)$  ist eindimensional, das heißt es gibt, bis auf skalare Vielfache, genau eine Determinantenform.*

*Beweis.* Sei eine Determinantenform  $D$  gegeben. Wir wählen eine Basis  $(b_1, \dots, b_n)$  von  $V$  und schreiben  $v_i = \sum_j \lambda_{ij} b_j$ . Dann rechnen wir mittels der Linearität in jeder Variablen ein Summenmonster aus

$$D(v_1, \dots, v_n) = \sum \lambda_{1j_1} \cdots \lambda_{nj_n} D(b_{j_1}, \dots, b_{j_n}).$$

Die Summe läuft dabei zunächst über alle Systeme  $(j_1, \dots, j_n)$  mit  $1 \leq j_k \leq n$ . Da  $D$  alternierend ist, ergibt sich aber höchstens dann ein von Null verschiedener Summand, wenn  $(j_1, \dots, j_n)$  eine Permutation von  $\{1, \dots, n\}$  ist. Wegen (3.2.1) gilt  $D(b_{1,\sigma(1)}, \dots, b_{n,\sigma(n)}) = \varepsilon(\sigma)D(b_1, \dots, b_n)$ . Insgesamt erhalten wir die *Leibnizsche<sup>1</sup> Formel*

$$D(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \varepsilon(\sigma) \lambda_{1,\sigma(1)} \cdots \lambda_{n,\sigma(n)} D(b_1, \dots, b_n),$$

die zeigt, daß  $D$  durch den Wert auf einer Basis bestimmt ist. Also ist  $\text{Alt}^n(V)$  höchstens eindimensional.

Es bleibt zu zeigen, daß es von Null verschiedene Determinantenformen überhaupt gibt. Zu diesem Zweck definieren wir eine Abbildung  $D$  durch die eben hergeleitete Formel, indem wir den Wert von  $D(b_1, \dots, b_n)$  irgendwie, etwa als 1, fixieren. Es bleibt dann zu zeigen, daß die so gewonnene Abbildung  $n$ -linear und alternierend ist. Die Linearität ist unmittelbar klar. Das Alternierende folgt leicht mit der oben genannten Eigenschaft von  $\varepsilon$ . Wir lassen die genaue Ausführung als Aufgabe, weil wir gleich noch einen zweiten Existenzbeweis für die Determinante führen.  $\square$

Für den zweiten Existenzbeweis verwenden wir Determinanten für Matrizen und leiten gleichzeitig eine induktive Berechnungsmethode her.

## Ergänzungen und Aufgaben

1. Sei  $\alpha: V^n \rightarrow K$  eine Abbildung mit den Eigenschaften

- (1)  $\alpha(v_1, \dots, \lambda v_j, \dots, v_n) = \lambda \alpha(v_1, \dots, v_j, \dots, v_n)$  (Homogenität),
- (2)  $\alpha(v_1, \dots, v_j, \dots) = \alpha(v_1, \dots, v_i + v_j, \dots)$  für  $i \neq j$  (Scherungsinvarianz).

Dann ist  $\alpha$  eine alternierende  $n$ -Linearform auf  $V$ . Im Fall  $n = 2$  lassen sich diese Axiome als elementargeometrische Eigenschaften des Flächeninhalts von Parallelogrammen veranschaulichen (für  $\lambda \geq 0$ ).

2. Die Gruppe  $S_n$  hat die Ordnung  $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$ , gelesen *n-Fakultät*.

<sup>1</sup>Gottfried Wilhelm Leibniz 1646 – 1716

## 3.2 Determinante einer Matrix

Sei  $V = K^n$ . Wir fassen Vektoren  $(a_1, \dots, a_n)$  aus  $K^n$  als Zeilen einer  $(n, n)$ -Matrix  $A = (a_{ij})$  auf, indem wir  $a_i = (a_{i1}, \dots, a_{in})$  schreiben. Auf diese Weise wird  $(K^n)^n$  mit  $M(n, n; K)$  identifiziert.

**(3.2.1) Determinante.** Eine *Determinantenfunktion* oder kurz eine *Determinante* für  $(n, n)$ -Matrizen über  $K$  ist eine Abbildung

$$\det: M(n, n; K) \rightarrow K,$$

die bezüglich jeder Zeile linear ist, die bezüglich der Zeilen alternierend ist und die durch die Bedingung  $\det(I_n) = 1$  normiert ist.  $\diamond$

Die Definitionen (3.1.1) und (3.2.1) sind im wesentlichen gleichwertig.

Ist  $(e_j)$  wie üblich die Standardbasis des  $K^n$ , so ist  $a_i = \sum_j a_{ij} e_j$  die  $i$ -te Zeile der Matrix  $A = (a_{ij})$ . Für eine Determinantenform  $D$  auf  $K^n$  mit  $D(e_1, \dots, e_n) = 1$  wird dann durch die Festsetzung  $\det(A) = D(a_1, \dots, a_n)$  eine Determinantenfunktion gegeben. Es gilt also die Leibniz-Formel

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1, \sigma(1)} \dots a_{n, \sigma(n)}.$$

Es sollte vielleicht einmal betont werden, daß jeder Summand ein Produkt ist, worin aus jeder Zeile und Spalte der Matrix genau ein Element benutzt wird.

Sei umgekehrt  $\det$  eine Determinantenfunktion auf  $M(n, n; K)$  und  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum. Wir wählen eine Basis  $(b_1, \dots, b_n)$  von  $V$ , schreiben  $v_i = \sum_j a_{ij} b_j$  und setzen  $D(v_1, \dots, v_n) = \det(a_{ij})$ . Dann ist  $D$  eine Determinantenform auf  $V$ . Für sie gilt  $D(b_1, \dots, b_n) = 1$ .

Der Eindeigkeitsatz (3.1.4) ist beweistechnisch sehr nützlich, wie wir beim Beweis der Produktformel sehen werden.

Ist  $A$  eine  $(n, n)$ -Matrix, so bezeichne  $A_{ik}$  die daraus durch Streichen der  $i$ -ten Zeile und  $k$ -ten Spalte entstehende  $(n-1, n-1)$ -Matrix.

**(3.2.2) Entwicklung nach einer Spalte.** *Es gibt genau eine Determinantenfunktion auf  $M(n, n; K)$ . Sie läßt sich durch die folgende Formel, genannt Entwicklung nach der  $k$ -ten Spalte induktiv berechnen:*

$$\det(A) = \sum_{i=1}^n (-1)^{i+k} a_{ik} \det(A_{ik}).$$

Die Vorzeichen  $(-1)^{i+k}$  bilden ein Schachbrettmuster

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

*Beweis.* Eindeutigkeit. Sei  $D$  eine Determinantenfunktion. Sind die Zeilen  $(v_1, \dots, v_n)$  linear abhängig, so ist  $D(v_1, \dots, v_n) = 0$ . Andernfalls können wir die Matrix mit den Zeilen  $(v_1, \dots, v_n)$  durch elementare Zeilenumformungen in die Einheitsmatrix überführen (2.4.3). Wir wissen, wie sich  $D$  bei elementaren Zeilenumformungen verhält: Zum Beispiel gelten  $D(\lambda v_1, \dots) = \lambda_1 D(v_1, \dots)$  sowie  $D(v_1 + v_2, v_2, \dots, v_n) = D(v_1, \dots, v_n)$ . Also läßt sich  $D(v_1, \dots, v_n)$  durch Zeilenumformungen berechnen. Das ist übrigens auch ein praktisches Berechnungsverfahren.

Existenz. Wir beweisen die Existenz durch Induktion nach  $n$ . Im Fall  $n = 1$  setzen wir  $\det(a) = a$ . Ist  $n > 1$ , so wählen wir ein  $k \in \{1, \dots, n\}$  und definieren  $\det(A)$  durch die im Satz genannte Formel. Wir zeigen, daß die damit gegebene Abbildung  $M(n, n; K) \rightarrow K$   $n$ -linear, alternierend und normiert ist.

Zunächst ist  $\det(I) = \sum_i (-1)^{i+k} \delta_{i,k} \det(I_{i,k}) = \delta_{kk} \det(I_{k,k}) = 1$  wie gewünscht (hier ist  $I = I_n$ ).

Weiterhin ist  $\det$  linear in jeder Zeile, weil dieses offenbar für jeden Summanden in der Entwicklungformel gilt.

Es bleibt das Alternierende zu zeigen. Sei  $v_r = v_s$  für  $r < s$ . Falls  $i \notin \{r, s\}$  so ist  $\det(A_{ik}) = 0$ , weil  $A_{ik}$  zwei gleiche Zeilen hat. Also bleiben nur die Summanden

$$\begin{aligned} & (-1)^{r+k} a_{rk} \det(A_{rk}) + (-1)^{s+k} a_{sk} \det(A_{sk}) \\ = & (-1)^k a_{rk} ((-1)^r \det(A_{rk}) + (-1)^s \det(A_{sk})). \end{aligned}$$

Für  $s = r + 1$  ist  $A_{rk} = A_{sk}$ , und deshalb annullieren sich die beiden Summanden in der Klammer. Für  $s > r + 1$  hat  $A_{rk}$  Zeilen der Form

$$v'_1, \dots, v'_{r-1}, v'_{r+1}, \dots, v'_{s-1}, v'_r, v'_{s+1}, \dots$$

und  $A_{sk}$  Zeilen der Form

$$v'_1, \dots, v'_{r-1}, v'_r, v'_{r+1}, \dots, v'_{s-1}, v'_{s+1}, \dots,$$

wobei  $v'_j$  aus  $v_j$  durch Streichen der  $k$ -ten Komponente entsteht. Also wird beim Übergang von  $A_{rk}$  nach  $A_{sk}$  die Zeile  $v'_r$  mit den  $s - r - 1$  Zeilen  $v'_{r+1}, \dots, v'_{s-1}$  vertauscht. Bei einer solchen Vertauschung ändert sich die Determinante nach (3.1.2) um das Vorzeichen  $(-1)^{s-r-1}$ , so daß sich auch in diesem Fall die beiden Summanden annullieren.  $\square$

Ist  $D \in \text{Alt}^n(V)$  und  $f: V \rightarrow V$  ein Endomorphismus, so ist  $(v_1, \dots, v_n) \mapsto D(fv_1, \dots, fv_n)$  wieder aus  $\text{Alt}^n(V)$ . Ist  $D \neq 0$ , so gibt es nach (3.1.4) genau ein  $d(f) \in K$ , so daß die Gleichung

$$D(fv_1, \dots, fv_n) = d(f)D(v_1, \dots, v_n)$$

für alle  $(v_1, \dots, v_n)$  gilt, und außerdem ist der Faktor  $d(f)$  von dem gewählten  $D \neq 0$  unabhängig. Wir setzen  $\det(f) = d(f)$  und nennen dieses Element die

*Determinante* der linearen Abbildung  $f$ . Ist  $V = K^n$  und wird  $f$  bezüglich der Standardbasis  $(e_1, \dots, e_n)$  durch die Matrix  $A$  gegeben, so gilt

$$\det(f) = \det(f) \det(e_1, \dots, e_n) = \det(fe_1, \dots, fe_n) = \det(A^t).$$

Wir betonen, daß  $\det(f)$  basisfrei definiert worden ist, das heißt ohne vorherige Auswahl einer Basis.

**(3.2.3) Produktsatz.** Für Endomorphismen  $f$  und  $g$  eines Vektorraumes  $V$  gilt  $\det(f \circ g) = \det(f) \det(g)$ , und für  $(n, n)$ -Matrizen  $A$  und  $B$  gilt ebenso  $\det(AB) = \det(A) \det(B)$ .

*Beweis.* Wir wenden dreimal die Definition an

$$\begin{aligned} d(fg)D(v_1, \dots, v_n) &= D(fg(v_1), \dots, fg(v_n)) \\ &= d(f)D(g(v_1), \dots, g(v_n)) \\ &= d(f)d(g)D(v_1, \dots, v_n). \end{aligned}$$

Analog durch Übersetzung in die Matrizen. □

Der kurze Beweis des Produktsatzes ist ein gutes Beispiel für die Kraft der begrifflichen Argumentation. Man schreibe einmal die Formel für  $(2, 2)$ -Matrizen explizit auf. Und für  $(3, 3)$ -Matrizen besser nicht?

**(3.2.4) Satz.** Die  $(n, n)$ -Matrix  $A$  ist genau dann regulär, wenn  $\det(A) \neq 0$  ist. Der Endomorphismus  $f: V \rightarrow V$  ist genau dann ein Isomorphismus, wenn  $\det(f) \neq 0$  ist.

*Beweis.* Gibt es eine zu  $A$  inverse Matrix  $B$ , so folgt aus dem Produktsatz  $1 = \det(I_n) = \det(AB) = \det(A) \det(B)$ , und das ist nur möglich, wenn  $\det(A) \neq 0$  ist. Ist umgekehrt  $\det(A) \neq 0$ , so sind die Zeilen von  $A$  nach (2.4.1) linear unabhängig, und deshalb ist  $A$  regulär. Für Endomorphismen wird der Beweis genauso geführt. □

Eine weitere Konsequenz des Produktsatzes ist:

$$SL(n, K) = \{A \in GL(n, K) \mid \det A = 1\}$$

ist eine Untergruppe von  $GL(n, K)$ , die sogenannte *spezielle lineare Gruppe*.

**(3.2.5) Satz.** Eine Matrix und ihre Transponierte haben dieselbe Determinante.

*Beweis.* Ist  $A$  eine  $(n, n)$ -Matrix vom Rang kleiner als  $n$ , so hat nach auch  $A^t$  einen Rang kleiner als  $n$ , und die Determinanten von  $A$  und  $A^t$  sind nach (2.4.2) beide gleich Null.

Sei also  $A$  invertierbar. Dann ist  $A$  nach (2.4.3) Produkt von Elementarmatrizen, etwa  $A = S_1 S_2 \cdots S_r$ . Folglich ist  $A^t = S_r^t \cdots S_1^t$ , und nach dem

Produktsatz genügt es deshalb, die Gleichung  $\det(S) = \det(S^t)$  für Elementarmatrizen  $S$  zu zeigen. Ist  $S$  eine Diagonalmatrix, so gilt  $S = S^t$ . Für die anderen Elementarmatrizen ist nach (2.4.1)  $\det(S) = 1$ .  $\square$

Durch Übergang zur transponierten Matrix erhalten wir jetzt aus dem schon Bewiesenen:

**(3.2.6) Entwicklung nach einer Zeile.** Die Zuordnung  $A \mapsto \det(A)$  ist  $n$ -linear und alternierend in den Spaltenvektoren. Wir haben die zu (3.2.2) analoge Entwicklung

$$\det(A) = \sum_{i=1}^n (-1)^{i+k} a_{ki} \det A_{ki}.$$

nach der  $k$ -ten Zeile.  $\square$

Wir setzen  $a_{ji}^\# = (-1)^{i+j} \det(A_{ij})$  und nennen die Matrix  $A^\# = (a_{ij}^\#)$  die Adjunkte von  $A$ .

**(3.2.7) Satz.** Für eine  $(n, n)$ -Matrix  $A$  gilt  $A^\# A = A A^\# = \det(A) \cdot I_n$ .

*Beweis.* Der Eintrag von  $A A^\#$  an der Stelle  $(i, i)$  ist

$$\sum_j a_{ij} a_{ji}^\# = \sum_i a_{ij} \cdot (-1)^{i+j} \det(A_{ij}) = \det(A).$$

Der Eintrag an der Stelle  $(i, j)$  für  $i \neq j$  ist  $\sum_k (-1)^{j+k} a_{ik} \det(A_{jk})$ , also gleich der Entwicklung nach der  $j$ -ten Zeile einer Matrix, die aus  $A$  entsteht, indem die  $j$ -te Zeile durch die  $i$ -te Zeile ersetzt wird. Also ist dieser Eintrag als Determinante einer Matrix mit zwei gleichen Zeilen gleich Null. Die andere Behauptung folgt ebenso durch Entwicklung nach Spalten.  $\square$

**(3.2.8) Inverse Matrix.** Ist  $A$  regulär, so liefert der letzte Satz die Formel

$$A^{-1} = \det(A)^{-1} A^\#$$

für das Inverse von  $A$ .  $\square$

Nach dieser Formel wird man das Inverse fast niemals praktisch berechnen; aber es ist bemerkenswert, daß es überhaupt eine Formel (rationale Funktion in den Einträgen) gibt. (Zum Beispiel besagt diese Formel, daß das Inverse einer reellen  $(n, n)$ -Matrix  $A$  stetig von  $A$  abhängt.) Wir erhalten daraus auch eine Formel für die Lösung eines linearen Gleichungssystems.

**(3.2.9) Cramersche Regel.** Sei  $A$  eine reguläre Matrix. Die Lösung des Gleichungssystems  $Ax = b$  ist  $x = A^{-1}b = |A|^{-1} A^\# b$ , also

$$x_i = |A|^{-1} \sum_j (-1)^{i+j} b_j |A_{ji}|.$$

Die letzte Summe ist aber nach dem Entwicklungssatz die Determinante einer Matrix, die aus  $A$  entsteht, wenn die  $i$ -te Spalte durch  $b$  ersetzt wird. Die damit gewonnene Formel für die Lösung wird *Cramersche<sup>2</sup> Regel* genannt.  $\diamond$

<sup>2</sup>Gabriel Cramer 1704 – 1752

**(3.2.10) Determinante von Blockmatrizen.** Haben wir eine Blockmatrix

$$U = \begin{pmatrix} I_k & 0 \\ B & C \end{pmatrix}$$

mit quadratischer Matrix  $C$ , so liefert Entwicklung nach der ersten Zeile und Induktion nach  $k$  die Gleichung  $\det(U) = \det(C)$ . Ähnlich wird eine Matrix behandelt, in der  $C$  und  $I_k$  vertauscht vorkommen. Aus dem Produktsatz erhalten wir damit

$$\det \begin{pmatrix} A & 0 \\ B & C \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ B & I \end{pmatrix} \cdot \det \begin{pmatrix} I & 0 \\ 0 & C \end{pmatrix} = \det(A) \det(C),$$

wenn  $A$  und  $C$  quadratisch sind. Durch Transposition erhalten wir das analoge Ergebnis für eine Blockmatrix mit der Null links unten. Daraus entnehmen wir durch Induktion nach  $n$  für  $(n, n)$ -Dreiecksmatrizen  $D$  mit Diagonale  $(d_1, \dots, d_n)$  die Gleichung  $\det(D) = d_1 \cdot d_2 \cdot \dots \cdot d_n$ . Wir erkennen noch einmal, daß eine solche Matrix genau dann regulär ist, wenn alle  $d_k$  von Null verschieden sind.  $\diamond$

### Ergänzungen und Aufgaben

1. Man leite den Entwicklungssatz (3.2.2) aus der Leibniz-Formel her. (Das ist nur ein Problem der Organisation und erfordert keine neue Idee.)
2. Das Resultat von (3.2.8) haben wir für  $(2, 2)$ -Matrizen schon früher als Rechenaufgabe gestellt, nämlich wo?
3. Sei  $f: V \rightarrow V$  eine  $\mathbb{C}$ -lineare Abbildung und  $f_{\mathbb{R}}: V_{\mathbb{R}} \rightarrow V_{\mathbb{R}}$  dieselbe Abbildung der zugrundeliegenden  $\mathbb{R}$ -Vektorräume. Dann gilt  $\det(f_{\mathbb{R}}) = |\det(f)|^2$ .

## 3.3 Volumen. Orientierung

Wir benutzen die Determinante, um zwei geometrische Begriffe für reelle Vektorräume zu definieren.

**(3.3.1) Volumen.** Seien  $(b_1, \dots, b_n)$  Vektoren des  $\mathbb{R}^n$ . Die Teilmenge

$$P(b_1, \dots, b_n) = \left\{ \sum_{i=1}^n \lambda_i b_i \mid 0 \leq \lambda_i \leq 1 \right\}$$

wird als der von  $(b_1, \dots, b_n)$  aufgespannte  $n$ -dimensionale Quader bezeichnet. Im Fall  $n = 2$  handelt es sich um ein Parallelogramm. Der Absolutbetrag

$$|\det(b_1, \dots, b_n)| = \text{Vol } P(b_1, \dots, b_n)$$

wird als *Volumen* von  $P(b_1, \dots, b_n)$  definiert. Im Fall  $n = 2$  bedeuten die Grundeigenschaften einer Determinante elementargeometrische Eigenschaften des Flächeninhaltes von Parallelogrammen.  $\diamond$

**(3.3.2) Orientierung.** Was bedeutet das Vorzeichen der Determinante in diesem Zusammenhang? Es dient zur Definition einer subtilen geometrischen Eigenschaft: der Orientierung. Sei  $B$  die Menge aller Basen eines  $n$ -dimensionalen Vektorraumes  $V$  über  $\mathbb{R}$ . Wir sagen, ein Paar  $B_1, B_2$  von Basen sei *gleich orientiert*, in Zeichen  $B_1 \simeq B_2$ , wenn die Basiswechselmatrix von  $B_1$  nach  $B_2$  positive Determinante hat. Aus dem Produktsatz für Determinanten folgt sofort, daß  $\simeq$  eine Äquivalenzrelation auf  $B$  ist. Eine Äquivalenzklasse von Basen heißt *Orientierung* von  $V$ . Ein *orientierter* Vektorraum ist ein Paar  $(V, o)$  aus einem Vektorraum  $V$  und einer Orientierung  $o$  von  $V$ . Ist eine Orientierung  $o$  von  $V$  festgelegt, so heißt eine Basis in der Klasse  $o$  *positiv* bezüglich dieser Orientierung; und natürlich *negativ*, wenn sie zur anderen Klasse gehört (Okzidentierung).  $\diamond$

Im Eindimensionalen ist eine Orientierung anschaulich eine Durchlaufrichtung der Geraden. Im Zweidimensionalen ist eine Orientierung anschaulich ein Drehsinn (Uhrzeigersinn). Üblicherweise wird die durch  $e_1, e_2$  festgelegte Orientierung des  $\mathbb{R}^2$  als positiv angesehen, der Uhrzeigersinn gehört dann zur negativen Orientierung. Im Dreidimensionalen ist eine Orientierung anschaulich ein Schraubensinn.

Ferner dient die Orientierung dazu, die Begriffe Rechts und Links festzulegen.

Eine Orientierung ist eine zusätzliche Struktur, eine „Vereinbarung“. Vereinbarungen über Rechts und Links kann man nicht durch das Telefon mitteilen sondern nur durch Vorzeigen (oder durch Vorurteile)<sup>3</sup>.

Sei  $V$  ein komplexer Vektorraum mit Basis  $B = \{b_1, \dots, b_n\}$ . Dann ist die Orientierung des unterliegenden reellen Raumes  $V_{\mathbb{R}}$ , die durch  $b_1, ib_1, \dots, b_n, ib_n$  gegeben wird, unabhängig von der Wahl von  $B$  und ihrer Anordnung. Sie heißt die durch die komplexe Struktur induzierte *kanonische Orientierung*.

Wie erinnerlich wird das Integral einer Funktion  $f: [a, b] \rightarrow \mathbb{R}$  oft mit dem Flächeninhalt motiviert. Aber auch dabei kommt es auf die Orientierung an, es gibt negative Integrale (Integration ist eine lineare Abbildung).

---

<sup>3</sup>Bekannt ist die Scherzfrage: Warum vertauscht ein Spiegel rechts und links aber nicht oben und unten? Nun, irgendwas vertauscht ja ein Spiegel, aber was?

### 3.4 Das charakteristische Polynom

Wir unterstellen eine gewisse Bekanntschaft mit dem Rechnen mit Polynomen. Dabei wird jetzt definitiv ein Polynom nicht als eine Funktion angesehen sondern als ein formales Objekt. Polynome werden also addiert und multipliziert nach den Regeln des „Buchstabenrechnens“. Ein Polynom  $p = p(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$  ist durch seine Koeffizienten bestimmt. Natürlich kann man auch weiterhin zu einem formalen Polynom die zugehörige Funktion verwenden. Die „Unbestimmte“ wird durch ein Körperelement  $\lambda \in K$  ersetzt, und man erhält  $p(\lambda) = a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_n \in K$ . Man sagt auch,  $\lambda$  wird in  $p$  eingesetzt. Ist  $p(\lambda) = 0$ , so heißt  $\lambda$  eine *Nullstelle* von  $p$ .

Analog zu den natürlichen und ganzen Zahlen kann man auch für Polynome eine Teilbarkeitslehre entwickeln, was aber hier zunächst nicht im einzelnen erläutert werden soll.

Ist  $\lambda$  eine Nullstelle des Polynoms  $p$ , so ist  $p$  durch das lineare Polynom  $x - \lambda$  teilbar, es gilt also  $p = (x - \lambda) \cdot q$ , worin das Polynom  $q$  einen um 1 kleineren Grad als  $p$  hat.

Interessant ist der Fall, daß  $p$  vollständig in Linearfaktoren zerfällt,  $p = \prod_{j=1}^n (x - \lambda_j)$ . Bekanntlich gilt das für Polynome über den komplexen Zahlen. Kommt der Faktor  $(x - \lambda_j)$  im Produkt  $k$ -mal vor, so heißt  $k$  die *Vielfachheit* der Nullstelle  $\lambda_j$ . Bis auf die Reihenfolge sind die Linearfaktoren durch  $p$  bestimmt. Ist  $q$  ein Teiler von  $p$ , gilt also eine Relation  $p = q \cdot r$ , so ist  $q$  ebenfalls ein Produkt von Linearfaktoren.

**(3.4.1) Charakteristisches Polynom.** Sei  $A = (a_{ij})$  eine  $(n, n)$ -Matrix mit Einträgen aus dem Körper  $K$ . Wir betrachten die mit einer weiteren Veränderlichen  $x$  gebildete Matrix  $xI_n - A$ . Auf der Diagonale stehen also die linearen Polynome  $x - a_{ii}$ . Rechnen wir die Determinante  $\det(xI_n - A)$  nach der Leibnizschen Formel aus, so sehen wir: Jeder Summand ist ein Polynom höchstens vom Grad  $n$ . Der einzige Summand vom Grad  $n$  erscheint in

$$(x - a_{11})(x - a_{22}) \cdot \dots \cdot (x - a_{nn}).$$

Deshalb ist

$$c_A(x) = \det(xI_n - A) = x^n + c_1x^{n-1} + \dots + c_n$$

ein Polynom vom Grad  $n$ , genannt das *charakteristische Polynom*  $c_A$  von  $A$ . Setzen wir  $x = 0$  ein, so erhalten wir  $\det(-A) = (-1)^n \det(A)$  als konstantes Glied des charakteristischen Polynoms. Der Koeffizient von  $x^{n-1}$  ist  $-\text{Sp}(A)$  (siehe Aufgabe 1).  $\diamond$

**(3.4.2) Satz.** *Genau dann ist  $\lambda \in K$  ein Eigenwert von  $A$ , wenn  $\lambda$  Nullstelle des charakteristischen Polynoms  $\det(xI - A)$  ist.*

*Beweis.* Sei  $\lambda$  Eigenwert von  $A$ . Dann gilt für ein  $x \in K^n \setminus \{0\}$  die Gleichung  $Ax = \lambda x$ . Wegen  $(\lambda I - A)x = 0$  ist  $\lambda I - A$  nicht invertierbar und deshalb ist  $\det(\lambda I - A) = 0$ , also  $\lambda$  Nullstelle. Ist umgekehrt  $\lambda$  Nullstelle, also  $\lambda I - A$  nicht invertierbar, so hat  $x \mapsto (\lambda I - A)x$  einen von 0 verschiedenen Kern, also  $A$  einen Eigenvektor zum Eigenwert  $\lambda$ .  $\square$

Es gilt  $\det(xI - SAS^{-1}) = \det(xSIS^{-1} - SAS^{-1}) = \det(S(xI - A)S^{-1}) = \det(S) \det(xI - A) \det(S^{-1}) = \det(xI - A)$ . Deshalb haben  $A$  und  $SAS^{-1}$  dasselbe charakteristische Polynom. Ist  $f: V \rightarrow V$  ein Endomorphismus eines  $n$ -dimensionalen  $K$ -Vektorraumes und  $A$  eine Matrix von  $f$  bezüglich einer Basis von  $V$ , so ist  $\det(xI - A)$  unabhängig von der gewählten Basis. Wir nennen deshalb  $\det(xI - A)$  auch das *charakteristische Polynom*  $c_f$  des Endomorphismus  $f$ . Ist  $A$  diagonalisierbar, das heißt gibt es  $S \in GL(n, K)$  mit  $SAS^{-1} = D = \text{Dia}(\lambda_1, \dots, \lambda_n)$ , so ist offenbar

$$\det(xI - A) = \det(xI - D) = (x - \lambda_1) \dots (x - \lambda_n);$$

das charakteristische Polynom zerfällt also in Linearfaktoren. Die Vielfachheit einer Nullstelle, eines Eigenwertes  $\lambda_j$ , ist in diesem Fall gleich der Dimension des zugehörigen Eigenraumes.

**(3.4.3) Satz.** Sei  $A \in M(n, n; K)$ . Folgende Aussagen sind äquivalent:

- (1)  $A$  ist diagonalisierbar.
- (2) Das charakteristische Polynom von  $A$  zerfällt in Linearfaktoren. Für jeden Eigenwert  $\lambda$  der Vielfachheit  $k$  hat der Eigenraum von  $\lambda$  die Dimension  $k$ .

*Beweis.* (1)  $\Rightarrow$  (2) haben wir schon bemerkt.

(2)  $\Rightarrow$  (1). Sei  $\prod_{j=1}^r (x - \lambda_j)^{n(j)}$  das charakteristische Polynom von  $A$ . Es ist dann  $n = \sum_{j=1}^r n(j)$ . Der Eigenraum  $V(\lambda_j)$  hat nach Voraussetzung die Dimension  $n(j)$ . Ferner gilt

$$V(\lambda_i) \cap (\sum_{j \neq i} V(\lambda_j)) = \{0\}$$

wie immer für Eigenräume. Also hat  $\sum_{j=1}^r V(\lambda_j) = V$  als direkte Summe der  $V(\lambda_j)$  die Dimension  $n = \sum n(j)$ . Deshalb ist  $V = K^n$ , und  $V$  hat eine Basis aus Eigenvektoren.  $\square$

**(3.4.4) Notiz.** Für eine Blockmatrix

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$$

mit quadratischen  $B$  und  $C$  gilt  $c_A = c_B c_C$ . Ist  $f: V \rightarrow V$  gegeben, ist  $U$  ein  $f$ -stabiler Unterraum und sind  $g: U \rightarrow U$  und  $h: V/U \rightarrow V/U$  die durch  $f$  induzierten Endomorphismen, so gilt  $c_f = c_g c_h$ .  $\square$

**(3.4.5) Satz.** *Ein Endomorphismus besitzt genau dann eine obere Dreiecksmatrix bezüglich einer geeigneten Basis, wenn sein charakteristisches Polynom in Linearfaktoren zerfällt.*

*Beweis.* Die eine Richtung haben wir eben schon bemerkt. Zerfalle also das charakteristische Polynom von  $f: V \rightarrow V$  in Linearfaktoren. Dann gibt es jedenfalls einen Eigenvektor  $v$ . Er erzeuge den Unterraum  $U$ . Dieser ist  $f$ -stabil. Das charakteristische Polynom der auf  $V/U$  induzierten Abbildung zerfällt als Faktor von  $c_f$  ebenfalls in Linearfaktoren. Nun wende man Induktion nach der Dimension an.  $\square$

### Ergänzungen und Aufgaben

1. Der Koeffizient von  $x^{n-1}$  des charakteristischen Polynoms einer  $(n, n)$ -Matrix kann mit der Leibniz-Formel bestimmt werden. Dazu schaue man auf die Summanden der Formel. Damit  $x^{n-1}$  vorkommt, müssen mindestens  $n-1$  Faktoren des Summanden die Form  $x - a_{jj}$  haben und dann auch notwendig der noch fehlende Faktor. Der Koeffizient von  $x^{n-1}$  in  $\prod_{j=1}^n (x - a_{jj})$  ist aber das Negative der Spur.

# Kapitel 4

## Bilineare Abbildungen

### 4.1 Hilbert-Räume

Die Strecke von  $(0, 0)$  nach  $(x, y)$  in der Zahlenebene  $\mathbb{R}^2$  hat in der elementaren euklidischen Geometrie nach dem Satz des Pythagoras eine Länge mit dem Quadrat  $x^2 + y^2$ . Ebenso ergibt sich das Längenquadrat der Strecke von  $(0, 0, 0)$  nach  $(x, y, z)$  im dreidimensionalen euklidischen Raum zu  $x^2 + y^2 + z^2$  durch zweimalige Anwendung des Satzes von Pythagoras.

Aus diesem Grund liegt es nahe, für eine allgemeine Längenmessung

$$q: \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto \sum_{j=1}^n x_j^2$$

zu betrachten. Als quadratische Abbildung ist  $q$  nicht linear. Die *Abweichung von der Linearität* ist

$$b(x, y) = q(x + y) - q(x) - q(y),$$

denn bei einer linearen Abbildung  $q$  wäre dieser Wert gleich Null. In Koordinaten errechnen wir  $b((x_j), (y_j)) = 2 \sum_{j=1}^n x_j y_j$ . Die Abbildung  $b$  hat den Vorteil, in jeder Variablen linear zu sein. Der Faktor 2 ist unwesentlich. Diese Vorbetrachtungen führen uns zur nächsten Definition.

**(4.1.1) Standardskalarprodukt.** Wir verwenden Bezeichnungen der Art  $x = (x_j) = (x_1, \dots, x_n)$ . Die Abbildung

$$\langle -, - \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad ((x_j), (y_j)) \mapsto \sum_{j=1}^n x_j y_j = \langle x, y \rangle$$

nennen wir das *Standardskalarprodukt* auf dem  $\mathbb{R}^n$  (das ist ein „Produkt“ mit einem Skalar als Resultat). Für komplexe Vektorräume gibt es eine ähnliche Abbildung

$$\langle -, - \rangle: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad ((z_j), (w_j)) \mapsto \sum_{j=1}^n \bar{z}_j w_j = \langle z, w \rangle,$$

die ebenfalls *Standardskalarprodukt* genannt werden soll. Der Grund für das Konjugiert-Komplexe in der letzten Definition ist natürlich die Relation  $\bar{z}z = a^2 + b^2$  für  $z = a + bi$  mit  $a, b \in \mathbb{R}$ ; also wieder das Längenquadrat.  $\diamond$

Wie oft in der Algebra ist nicht die explizite Formel entscheidend, wichtig sind vielmehr ihre formalen Eigenschaften. Das führt zur nächsten Definition. Wir behandeln reelle und komplexe Vektorräume gleichzeitig und vereinbaren dazu:  $\mathbb{F}$  sei einer der Körper  $\mathbb{R}$  oder  $\mathbb{C}$ . Mit  $\bar{z}$  bezeichnen wir die zu  $z \in \mathbb{F}$  konjugiert-komplexe Zahl; ist  $z \in \mathbb{R}$ , so ist also  $\bar{z} = z$ .

**(4.1.2) Skalarprodukt.** Die Standardabbildungen (4.1.1) haben die folgenden Eigenschaften: Für  $v_1, v_2, v, w_1, w_2, w$  aus  $\mathbb{F}^n$  und  $\lambda$  aus  $\mathbb{F}$  gilt

- (1)  $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$
- (2)  $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
- (3)  $\langle \lambda v, w \rangle = \bar{\lambda} \langle v, w \rangle$
- (4)  $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$
- (5)  $\langle v, w \rangle = \overline{\langle w, v \rangle}$
- (6)  $\langle v, v \rangle > 0$ , falls  $v \neq 0$ .

Sei  $V$  ein  $\mathbb{F}$ -Vektorraum. Eine Abbildung  $\langle -, - \rangle: V \times V \rightarrow \mathbb{F}$  mit den Eigenschaften (1) – (6) heißt *Skalarprodukt* auf  $V$ . Ein Paar  $(V, \langle -, - \rangle)$  heißt im Fall  $\mathbb{F} = \mathbb{R}$  ein *euklidischer Vektorraum* und im Fall  $\mathbb{F} = \mathbb{C}$  ein *unitärer Vektorraum*. Statt Skalarprodukt ist auch der Terminus *inneres Produkt* gebräuchlich. Wollen wir  $\mathbb{R}$  und  $\mathbb{C}$  gemeinsam behandeln, so nennen wir  $(V, \langle -, - \rangle)$  einen *Hilbert-Raum*<sup>1</sup>.  $\diamond$

Der Begriff des Hilbert-Raumes ist wichtig in der Analysis. Es wird dann allerdings noch zusätzlich verlangt, daß Cauchy-Folgen konvergieren (Vollständigkeit von  $V$ ). Das spielt für unsere algebraischen Überlegungen keine Rolle. In der Analysis gebraucht man ohne die Bedingung der Vollständigkeit das Wortungetüm Prä-Hilbert-Raum.)

Ein reelles Skalarprodukt ist in beiden Variablen linear. Ein komplexes Skalarprodukt wird in beiden Variablen linear, wenn wir es als auf  $\bar{V} \times V$  definiert auffassen. Abbildungen von diesem Typ nennen wir *bilinear*. Sie werden später allgemein untersucht. Werden nur die Eigenschaften (1) – (5) gefordert, so sprechen wir im Fall  $\mathbb{F} = \mathbb{R}$  von einer *symmetrischen Bilinearform* und im Fall  $\mathbb{F} = \mathbb{C}$  von einer *hermiteschen Form*<sup>2</sup> auf  $V$ . Gilt dann für diese Formen zusätzlich (6), so heißen sie *positiv definit*.

Allgemein folgt aus (1) – (4)

$$\langle \sum_j \lambda_j v_j, \sum_k \mu_k w_k \rangle = \sum_{j,k} \bar{\lambda}_j \mu_k \langle v_j, w_k \rangle,$$

genannt „bilineares Ausrechnen“. Insbesondere ist die Abbildung  $\langle -, - \rangle$  bekannt, wenn man die Matrix  $(\langle v_j, v_k \rangle)$  für eine Basis  $\{v_1, \dots, v_n\}$  kennt. Ist

<sup>1</sup>David Hilbert 1862 – 1943

<sup>2</sup>Charles Hermite 1822 – 1901

$A = (a_{jk})$  irgendeine  $(n, n)$ -Matrix über  $\mathbb{F}$  und  $B = \{v_1, \dots, v_n\}$  eine Basis von  $V$ , so wird durch

$$\langle \sum_j \lambda_j v_j, \sum_k \mu_k v_k \rangle = \sum_{j,k} \bar{\lambda}_j a_{jk} \mu_k = \bar{\lambda}^t A \mu$$

mit den Koordinatenvektoren  $\lambda^t = (\lambda_1, \dots, \lambda_n)$  und  $\mu^t = (\mu_1, \dots, \mu_n)$  und den entsprechenden Spaltenvektoren  $\lambda$  und  $\mu$  eine Abbildung  $V \times V \rightarrow \mathbb{F}$  mit den Eigenschaften (1) – (4) definiert. Genau dann gilt (5), wenn  $\bar{A}^t = A$  ist; eine Matrix mit diesen Eigenschaften heißt im Fall  $\mathbb{F} = \mathbb{R}$  *symmetrisch* und im Fall  $\mathbb{F} = \mathbb{C}$  *hermitesch*. Einer Matrix ist jedoch nicht so leicht anzusehen, wann (6) gilt; ist das der Fall, so nennen wir sie *positiv definit*.

**(4.1.3) Beispiel.** Eine symmetrische reelle  $(2, 2)$ -Matrix hat die Gestalt

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Ist sie positiv definit, so muß

$$(x, y) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2 > 0$$

sein, sofern  $(x, y) \neq (0, 0)$  ist. Demnach ist  $0 < a$  und  $0 < c$  notwendig für positive Definitheit ( $x$  oder  $y$  gleich Null setzen). Aus der Formel

$$ax^2 + 2bxy + cy^2 = (\sqrt{a}x + \sqrt{c}y)^2 + 2(b - \sqrt{ac})xy$$

erkennen wir, daß  $b < \sqrt{ac}$  oder äquivalent  $\det A > 0$  ebenfalls notwendig ist ( $x = \sqrt{c}$  und  $y = -\sqrt{a}$  setzen). Es ist eine kleine Aufgabe, zu zeigen, daß unter den eben hergeleiteten Bedingungen die Form wirklich positiv definit ist. Später werden wir diese Tatsache ohne Rechnung einsehen lernen (4.7.4).  $\diamond$

Sei im folgenden  $(V, \langle -, - \rangle)$  ein Hilbert-Raum. Wie üblich wird er meist nur durch die zugrundeliegende Menge  $V$  bezeichnet. Skalarprodukte dienen unter anderem zur Längendefinition. Wir nennen  $|v| = \sqrt{\langle v, v \rangle}$  die *Norm* oder *Länge* von  $v$  und  $|v - w|$  den *Abstand von  $v, w$*  (bezüglich  $\langle -, - \rangle$ ). Vektoren der Länge 1 heißen *Einheitsvektoren*. Ist  $x \neq 0$ , so ist  $|x|^{-1}x$  ein Einheitsvektor. Gilt  $\langle u, v \rangle = 0$ , so heißen  $u, v$  *orthogonal*. Wegen Axiom (5) ist  $\langle u, v \rangle = 0$  zu  $\langle v, u \rangle = 0$  äquivalent.

Für orthogonale  $u, v$  gilt  $|u|^2 + |v|^2 = |u + v|^2$  (eine ab jetzt auch *Satz des Pythagoras* genannte Relation).

Unterräume  $U_1$  und  $U_2$  von  $V$  heißen *orthogonal*, wenn für  $u_j \in U_j$  immer  $\langle u_1, u_2 \rangle = 0$  ist. Ist  $V$  die direkte Summe paarweise orthogonaler Unterräume  $(U_j \mid 1 \leq j \leq t)$ , so sagen wir,  $V$  sei die *orthogonale Summe* der  $U_j$  und notieren diesen Sachverhalt durch

$$V = U_1 \perp U_2 \perp \dots \perp U_t.$$

**(4.1.4) Satz.** Für  $x, y \in V$  gilt:

- (1)  $|\langle x, y \rangle| \leq |x||y|$  (Cauchy-Schwarzsche Ungleichung<sup>34</sup>).  
 (2)  $|x + y| \leq |x| + |y|$  (Dreiecksungleichung).

*Beweis.* (1) Seien  $x, y \in V$  gegeben, und sei  $x \neq 0$ . Wir suchen zunächst eine orthogonale Zerlegung der Form

$$y = \lambda x + z, \quad \lambda \in \mathbb{F}, \quad \langle x, z \rangle = 0.$$

Eine solche Zerlegung ist auf genau eine Weise möglich; ferner ist notwendig  $\lambda = |x|^{-2} \langle x, y \rangle$ . Aus  $z = y - \lambda x$  und  $\langle x, z \rangle = 0$  folgt nämlich

$$0 = \langle x, z \rangle = \langle x, y - \lambda x \rangle = \langle x, y \rangle - \lambda \langle x, x \rangle.$$

Deshalb hat  $\lambda$  die behauptete Gestalt. Umgekehrt rechnet man leicht nach, daß mit diesem  $\lambda$  der Vektor  $y - \lambda x$  orthogonal zu  $x$  ist.

Für  $x = 0$  gilt die behauptete Ungleichung. Andernfalls folgt mit dem soeben Überlegten und dem Satz des Pythagoras

$$|y|^2 = |\lambda x|^2 + |z|^2 \geq |\lambda|^2 |x|^2 = \frac{\langle x, y \rangle^2}{|x|^4} \cdot |x|^2,$$

und das ist zu (1) gleichwertig.

(2) wird durch die folgende Rechnung belegt:

$$\begin{aligned} |x + y|^2 &= |x|^2 + |y|^2 + \langle x, y \rangle + \langle y, x \rangle \\ &\leq |x|^2 + |y|^2 + 2|\langle x, y \rangle| \leq |x|^2 + |y|^2 + 2|x||y| \\ &= (|x| + |y|)^2. \end{aligned}$$

Wir haben darin (1) verwendet. □

Im vorstehenden Satz gilt in (1) genau dann die Gleichheit, wenn  $x$  und  $y$  linear abhängig sind. Das sieht man der letzten Formel im Beweis von (1) an, weil die Gleichheit genau im Fall  $z = 0$  besteht.

In einem Dreieck ist die Summe der Längen zweier Seiten mindestens gleich der Länge der dritten Seite; das ist der Grund für die Bezeichnung „Dreiecksungleichung“.

Wir benutzen nun die Funktion Cosinus und setzen im Falle  $x \neq 0 \neq y$  und  $\mathbb{F} = \mathbb{R}$

$$\cos w(x, y) = \frac{\langle x, y \rangle}{|x||y|}.$$

Das ist möglich, weil nach dem letzten Satz die rechte Seite im Intervall  $[-1, 1]$  liegt. Wir normieren  $0 \leq w(x, y) \leq \pi$  und nennen  $w(x, y)$  den von  $x$  und  $y$  eingeschlossenen *Winkel*.

<sup>3</sup>Augustin Cauchy 1789 – 1857

<sup>4</sup>Hermann Amandus Schwarz 1843 – 1921

Sei  $M \subset V$  eine Menge von Null verschiedener Vektoren. Wir definieren:

- $M$  Orthogonalsystem  $\Leftrightarrow$  Je zwei verschiedene Vektoren aus  $M$  sind orthogonal  
 $M$  Orthonormalsystem  $\Leftrightarrow$   $M$  Orthogonalsystem, und jeder Vektor aus  $M$  ist Einheitsvektor  
 $M$  Orthonormalbasis  $\Leftrightarrow$   $M$  Basis und Orthonormalsystem

**(4.1.5) Notiz.** Sei  $M$  Orthogonalsystem. Dann ist  $M$  linear unabhängig.

*Beweis.* Seien  $a_1, \dots, a_n$  paarweise verschiedene Vektoren aus  $M$ . Sei  $\sum \lambda_i a_i = 0$ . Dann ist  $0 = \langle a_j, \sum_i \lambda_i a_i \rangle = \sum_i \lambda_i \langle a_j, a_i \rangle = \lambda_j \langle a_j, a_j \rangle$ ; also ist  $\lambda_j = 0$ .  $\square$

**(4.1.6) Satz.** Sei  $M = \{x_1, \dots, x_n\} \subset V$  linear unabhängig. Es gibt ein Orthonormalsystem  $\{e_1, \dots, e_n\}$  mit der Eigenschaft: Für jedes  $j \in \{1, \dots, n\}$  spannen  $\{x_1, \dots, x_j\}$  und  $\{e_1, \dots, e_j\}$  denselben Unterraum auf.

*Beweis.* Die Vektoren  $e_j$  werden induktiv definiert. Es ist  $x_1 \neq 0$ , also  $|x_1| \neq 0$  und somit  $e_1 = |x_1|^{-1} x_1$  ein Einheitsvektor. Sei  $\{e_1, \dots, e_k\}$  ein Orthonormalsystem, so daß für alle  $j \in \{1, \dots, k\}$  die im Satz genannte Bedingung erfüllt ist. Der Vektor

$$b_{k+1} = x_{k+1} - \sum_{j=1}^k \langle e_j, x_{k+1} \rangle e_j$$

ist zu  $e_1, \dots, e_k$  orthogonal. Er ist auch von Null verschieden, da  $x_{k+1}, e_1, \dots, e_k$  linear unabhängig sind. Wir setzen  $e_{k+1} = |b_{k+1}|^{-1} b_{k+1}$ .  $\square$

Das im Beweis von (4.1.6) angegebene Verfahren heißt *Orthonormalisierungsverfahren* von Gram<sup>5</sup> und Schmidt<sup>6</sup>. Es läßt sich auch für abzählbar-unendlich viele Vektoren  $x_1, x_2, x_3, \dots$  durchführen. Als unmittelbare Folgerung erhalten wir:

**(4.1.7) Satz.** Sei  $V$  endlichdimensional. Jedes Orthonormalsystem kann zu einer Orthonormalbasis ergänzt werden. Insbesondere gibt es Orthonormalbasen.  $\square$

**(4.1.8) Notiz.** Sei  $e_1, \dots, e_n$  Orthonormalbasis. Dann ist  $\langle e_j, v \rangle$  die  $j$ -te Komponente von  $v$  bezüglich dieser Basis.

*Beweis.* Aus  $v = \sum_j \lambda_j e_j$  folgt  $\langle e_k, v \rangle = \lambda_k$ .  $\square$

**(4.1.9) Satz.** Sei  $B = \{v_1, \dots, v_n\}$  ein Orthonormalsystem. Sei  $v \in V$ ,  $c_j = \langle v_j, v \rangle$ . Seien  $a_1, \dots, a_n \in \mathbb{F}$ . Dann gelten:

- (1)  $|v - \sum_{k=1}^n c_k v_k| \leq |v - \sum_{k=1}^n a_k v_k|$  (Besselsche Ungleichung<sup>7</sup>).  
 (2)  $\sum_{k=1}^n |c_k|^2 \leq |v|^2$  (Parsevalsche Ungleichung<sup>8</sup>).

<sup>5</sup>Jørgen Pedersen Gram 1850 – 1916

<sup>6</sup>Erhard Schmidt 1876 – 1959

<sup>7</sup>Friedrich Wilhelm Bessel 1784 – 1846

<sup>8</sup>Marc-Antoine Parseval 1755 – 1836

(3) Ist  $B$  eine Orthonormalbasis, so gilt  $\sum_{k=1}^n |c_k|^2 = |v|^2$ . Ist  $B$  Orthonormalsystem und gilt diese Gleichung immer, so ist  $B$  eine Basis.

*Beweis.* (1) Es ist  $v_j$  orthogonal zu  $v - \sum c_k v_k$ ; also ist  $v - \sum c_k v_k$  orthogonal zu jeder Linearkombination der  $v_j$  (bilineares Ausrechnen). Nach dem Satz des Pythagoras gilt somit

$$\begin{aligned} |v - \sum a_k v_k|^2 &= |v - \sum c_k v_k + \sum (c_k - a_k) v_k|^2 \\ &= |v - \sum c_k v_k|^2 + |\sum (c_k - a_k) v_k|^2 \geq |v - \sum c_k v_k|^2. \end{aligned}$$

(2) folgt aus  $0 \leq \langle v - \sum c_k v_k, v - \sum c_k v_k \rangle$  durch bilineares Ausrechnen des Skalarproduktes. Ebenso folgt die Gleichheit in (3), weil dann  $v = \sum c_k v_k$  ist.  $\square$

Sei  $U$  ein endlichdimensionaler Unterraum von  $V$ . Der Raum

$$U^\perp = \{v \in V \mid \text{für alle } u \in U \text{ ist } \langle u, v \rangle = 0\}$$

heißt *orthogonales Komplement* von  $U$  in  $V$ , denn es gilt:

**(4.1.10) Satz.** Für jeden endlichdimensionalen Unterraum  $U$  von  $V$  besteht die orthogonale Zerlegung  $V = U \oplus U^\perp$ .

*Beweis.* Sei  $v \in U \cap U^\perp$ . Dann ist  $\langle v, v \rangle = 0$ , also  $v = 0$ . Sei  $(e_1, \dots, e_n)$  eine Orthonormalbasis von  $U$ . Dann liegt für jedes  $v$  der Vektor

$$p(v) = v - \sum_{k=1}^n \langle e_k, v \rangle e_k$$

in  $U^\perp$ . Folglich gilt  $V = U + U^\perp$ .  $\square$

Sei  $p: V \rightarrow V$  eine Projektion, das heißt es gelte  $p \circ p = p$ . Wir nennen  $p$  *orthogonale Projektion* auf  $U = \text{Bild } p$ , wenn  $\text{Kern } p = U^\perp$  ist; es ist dann also

$$U = \text{Bild } p \perp \text{Kern } p.$$

**(4.1.11) Satz.** Sei  $U$  ein endlichdimensionaler Unterraum von  $V$ . Dann gilt:

- (1) Es gibt genau eine orthogonale Projektion auf  $U$ .
- (2) Ist  $p$  die orthogonale Projektion auf  $U$ , so gilt für  $u \in U$ ,  $v \in V$  immer  $|u - v| \geq |pv - v|$ . Aus  $|u - v| = |pv - v|$  folgt  $u = pv$ .

*Beweis.* (1) Es ist  $V = U \oplus U^\perp$ . Jeder Vektor  $v \in V$  hat also eine eindeutige Darstellung der Form  $v = u + u'$ ,  $u \in U$ ,  $u' \in U^\perp$ . Es ist  $v \mapsto u$  eine orthogonale Projektion auf  $U$ .

Ist umgekehrt  $p$  eine orthogonale Projektion auf  $U = \text{Bild } p$ , so gilt  $p(v) = p(u) = u$ , da  $u' \in U^\perp = \text{Kern } p$  ist.

(2) Wir betrachten die folgende Kette von Gleichungen und Ungleichungen

$$|u - v|^2 = |pu - v|^2 = |pu - pv + pv - v|^2 \stackrel{(*)}{=} |pu - pv|^2 + |pv - v|^2 \geq |pv - v|^2.$$

Gilt Gleichheit, so folgt  $|pu - pv|^2 = 0$ , also  $pu - pv = u - pv = 0$ . Die Gleichheit (\*) gilt nach dem Satz des Pythagoras, weil  $pv - v \in \text{Kern } p$  orthogonal zu  $pu - pv \in \text{Bild } p$  ist.  $\square$

Die Aussage (2) des letzten Satzes läßt sich so lesen: Die orthogonale Projektion  $pv$  gibt denjenigen Vektor in  $U$  an, der von  $v$  den kleinsten Abstand hat (das *Lot* von  $v$  auf  $U$ ).

## Ergänzungen und Aufgaben

**1.** Eine leichte Rechnung zeigt  $\langle a + b, a + b \rangle - \langle a - b, a - b \rangle = 4\text{Re} \langle a, b \rangle$ . In der ebenen Geometrie sind  $a + b$  und  $a - b$  die beiden Diagonalen des von  $a$  und  $b$  aufgespannten Parallelogramms. Ein Parallelogramm ist also genau dann ein Rechteck (das heißt  $a, b$  sind orthogonal), wenn seine Diagonalen gleich lang sind.

**2. Normierte Vektorräume.** Wir erwähnen der Vollständigkeit halber einen weiteren, insbesondere in der Analysis wichtigen, Begriff. Eine *Norm* auf einem  $\mathbb{F}$ -Vektorraum  $V$  ist eine Abbildung  $N: V \rightarrow \mathbb{R}$  mit den Eigenschaften:

- (1)  $N(\lambda x) = |\lambda|N(x)$
- (2)  $N(x + y) \leq N(x) + N(y)$
- (3)  $N(x) > 0 \Leftrightarrow x \neq 0$ .

Ein Paar  $(V, N)$  aus einem Vektorraum  $V$  und einer Norm  $N$  auf  $V$  heißt *normierter Vektorraum*. Ein Skalarprodukt führt also in natürlicher Weise zu einem normierten Vektorraum.

**3. Hessesche Normalform**<sup>9</sup>. Sei  $V$  ein endlichdimensionaler euklidischer Raum. Jede Hyperebene  $L$  von  $V$  läßt sich in der Form

$$L = \{x \mid \langle e, x \rangle = c\}$$

mit einem Einheitsvektor  $e$  und einer reellen Zahl  $c \geq 0$  schreiben. Die Daten  $e$  und  $c$  sind unter diesen Bedingungen durch  $L$  eindeutig bestimmt. Die Hyperebenengleichung  $\langle e, x \rangle = c$  heißt *Hessesche Normalform*.

*Beweis.* Eine Hyperebene  $L$  ist eine Menge der Form  $L = v + U$  mit einem Unterraum  $U$  der Kodimension 1 in  $V$ . Sei  $e$  ein Einheitsvektor im orthogonalen Komplement von  $U$ . Dann hat  $v$  die Form  $v = ce + u$  mit einem  $u \in U$ , und folglich ist  $L = ce + U$ . Indem wir  $e$  eventuell durch  $-e$  ersetzen, können wir  $c \geq 0$  annehmen. Dann ist aber  $L = \{x \mid \langle e, x \rangle = c\}$ . Ist umgekehrt  $L$  in dieser Form gegeben, so ist  $e$  orthogonal zur Richtung von  $L$ .  $\square$

Sei  $\langle e, x \rangle = c$  eine Hessesche Normalform. Für  $v = ce + u \in L$ ,  $u \in (\mathbb{R}e)^\perp$  ist  $|v|^2 = |u|^2 + c^2$ . Deshalb ist  $c$  der minimale Abstand eines Punktes von  $L$  vom Nullpunkt. Wir nennen  $c$  den *Abstand* von  $L$  zum Nullpunkt.

Ist  $\langle e, y \rangle = d$ , so heißt  $d - c$  der *orientierte Abstand* von  $y$  und  $L$  und  $|d - c|$  der *Abstand* von  $y$  und  $L$ .

„Orientiert“ soll besagen, daß man durch das Vorzeichen von  $\langle e, y \rangle - c$  die beiden Halbräume unterscheiden kann, in die  $V$  durch  $L$  zerlegt wird; es gilt  $V \setminus L = H^+ \cup$

<sup>9</sup>Ludwig Otto Hesse 1811 – 1874

$H^-$ , positives Vorzeichen für  $y \in H^+$ , negatives für  $y \in H^-$ . Setzt man  $y = de + w$ ,  $w \in (\mathbb{R}e)^\perp$ , so ist

$$|y - (ce + v)|^2 = |de + w - ce - v|^2 = |d - c|^2 + |w - v|^2;$$

dieser Wert wird für  $w = v$  minimal. Also ist  $|d - c|$  der minimale Abstand von  $y$  zu einem Punkt von  $L$ .

## 4.2 Isometrien

Sei  $(V, \langle -, - \rangle)$  ein  $n$ -dimensionaler Hilbert-Raum. Ein Endomorphismus  $f$  von  $V$  heißt *Isometrie* von  $V$ , wenn immer  $\langle u, v \rangle = \langle fu, fv \rangle$  gilt. Eine Isometrie bewahrt Längen und Abstände und Winkel. Im Fall  $\mathbb{F} = \mathbb{R}$  folgt aus

$$|u + v|^2 - |u|^2 - |v|^2 = 2\langle u, v \rangle,$$

daß eine längenerhaltende lineare Abbildung eine Isometrie ist. Eine Isometrie ist ein Isomorphismus, denn aus  $fu = 0$  folgt  $0 = \langle fu, fu \rangle = \langle u, u \rangle$  und damit  $u = 0$ ; also ist  $f$  injektiv und damit bijektiv. Die Isometrien bilden bezüglich Verkettung eine Gruppe, die im Fall  $\mathbb{F} = \mathbb{R}$  die *orthogonale Gruppe*  $O(V)$  und im Fall  $\mathbb{F} = \mathbb{C}$  die *unitäre Gruppe*  $U(V)$  von  $V$  genannt wird. Die Abbildungen in  $O(V)$  (und in  $U(V)$ ) heißen *orthogonal* (und *unitär*).

**(4.2.1)** Sei  $A = (a_{ij})$  die Matrix eines Endomorphismus  $f$  bezüglich einer Orthonormalbasis  $(e_1, \dots, e_n)$ . Dann ist  $f$  genau dann eine Isometrie, wenn  $\bar{A}^t A = I_n$  ist.

*Beweis.* Genau dann ist  $f$  eine Isometrie, wenn immer  $\langle e_k, e_l \rangle = \langle fe_k, fe_l \rangle$  gilt (Linearität von  $f$ , bilineares Ausrechnen). Wir setzen  $fe_k = \sum_\mu a_{\mu k} e_\mu$  ein und sehen, daß diese Gleichung zu  $\bar{A}^t A = I_n$  äquivalent ist.  $\square$

Eine reelle (komplexe)  $(n, n)$ -Matrix  $A$  heißt *orthogonal* (*unitär*), wenn  $\bar{A}^t A = I_n$  gilt. Sei  $O(n)$  bzw.  $U(n)$  die Gruppe der orthogonalen bzw. unitären  $(n, n)$ -Matrizen. Eine orthogonale Matrix ist auch unitär, wir haben also eine Untergruppe  $O(n) \subset U(n)$ . Aus  $\bar{A}^t A = I_n$  folgt  $1 = \det(I_n) = \det(\bar{A}^t) \det(A) = |\det(A)|^2$ . Deshalb hat eine orthogonale Matrix eine Determinante  $\pm 1$  und eine unitäre Matrix eine Determinante vom Betrag 1. Die Untergruppen

$$SO(n) = \{A \in O(n) \mid \det A = 1\}, \quad SU(n) = \{A \in U(n) \mid \det A = 1\}$$

heißen *spezielle orthogonale Gruppen* orthogonale (unitäre) Gruppen.

Eine Rechnung wie im Beweis von (4.2.1) zeigt:

**(4.2.2) Notiz.** Sei  $(e_1, \dots, e_n)$  eine Orthonormalbasis und  $(b_1, \dots, b_n)$  eine andere Basis. Die Basiswechselmatrix ist genau dann orthogonal (unitär), wenn  $(b_1, \dots, b_n)$  eine Orthonormalbasis ist.  $\square$

Allgemein ist das Element von  $\bar{A}^t A$  an der Stelle  $(k, l)$  gleich  $\langle a_k, a_l \rangle$ , wenn  $a_l$  die  $l$ -te Spalte von  $A$  ist (hier ist jetzt  $\langle -, - \rangle$  das Standardskalarprodukt auf  $\mathbb{F}^n$ ). Also bedeutet  $\bar{A}^t A = I_n$ , daß die Spalten eine Orthonormalbasis bezüglich des Standardskalarproduktes bilden. Aus der Gleichung  $\bar{A}^t A = I_n$  folgt  $\bar{A} A^t = I_n$ , indem man sie von links (rechts) mit dem Inversen von  $\bar{A}^t$  (von  $A$ ) multipliziert und das Ergebnis invertiert. Also bilden auch die Zeilen eine Orthonormalbasis.

**(4.2.3) Die Gruppe  $O(2)$ .** Ein zu  $(a, b) \neq (0, 0)$  orthogonaler Vektor hat die Form  $(-cb, ca)$ . Die Matrix

$$\begin{pmatrix} a & -cb \\ b & ca \end{pmatrix}$$

hat die Determinante  $c(a^2 + b^2)$ . Ist sie aus  $O(2)$ , so ist  $a^2 + b^2 = 1$  und folglich  $c = \pm 1$ . Demnach besteht  $SO(2)$  aus den Matrizen

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a^2 + b^2 = 1.$$

Sie lassen sich in der Form

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} = D(\varphi)$$

schreiben. Solche Matrizen heißen *Drehmatrizen*. (Drehung um den Winkel  $\varphi$  gegen den Uhrzeigersinn in der üblichen Veranschaulichung der ebenen Geometrie.) Die anderen Matrizen aus  $O(2)$  sind

$$S = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \quad a^2 + b^2 = 1.$$

Mit dieser Matrix  $S$  gilt

$$S \begin{pmatrix} a+1 \\ b \end{pmatrix} = \begin{pmatrix} a+1 \\ b \end{pmatrix}, \quad S \begin{pmatrix} -b \\ a+1 \end{pmatrix} = - \begin{pmatrix} -b \\ a+1 \end{pmatrix}.$$

Also ist  $S$  die Spiegelung an dem durch  $\begin{pmatrix} a+1 \\ b \end{pmatrix}$  erzeugten Unterraum. Insbesondere gilt  $S^2 = I_2$ . Wir schreiben Spiegelmatrizen in der Form

$$\begin{pmatrix} \cos 2\varphi & \sin 2\varphi \\ \sin 2\varphi & -\cos 2\varphi \end{pmatrix} = S(\varphi).$$

Der Grund ist aus den Relationen

$$S(\varphi) \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad S(\varphi) \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} = - \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$$

ersichtlich;  $S(\varphi)$  ist die Spiegelung an der von  $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$  aufgespannten Geraden. Es gelten die Formeln:

$$\begin{aligned} D(\varphi)D(\psi) &= D(\varphi + \psi) \\ S(0)D(\varphi) &= D(-\varphi)S(0) \\ D(\varphi)S(0)D(-\varphi) &= S(\varphi) \\ S(\varphi)S(\psi) &= D(2\varphi - 2\psi) \\ S(\varphi) &= D(2\varphi)S(0) \\ S(\varphi)D(\psi) &= D(2\varphi - \psi)S(0). \end{aligned}$$

Sie geben an, wie in der Gruppe  $O(2)$  gerechnet wird. Insbesondere sehen wir: Jede Drehung ist Produkt zweier geeigneter Spiegelungen, und das Produkt zweier Spiegelungen ist immer eine Drehung.

Ein unitäre  $(1, 1)$ -Matrix „ist“ eine komplexe Zahl vom Betrag eins. Multiplikation mit  $z = a + bi$  ist eine  $\mathbb{R}$ -lineare Abbildung  $\mathbb{C} \rightarrow \mathbb{C}$ , die bezüglich der  $\mathbb{R}$ -Basis  $1, i$  von  $\mathbb{C}$  die Matrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

hat. Hat  $z$  den Betrag 1, so handelt es sich um eine Drehmatrix. Damit erkennen wir die geometrische Bedeutung der Multiplikation komplexer Zahlen.  $\diamond$

**(4.2.4) Notiz.** Ist  $f$  eine Isometrie von  $V$  und  $U$  ein  $f$ -stabiler Unterraum, so gilt  $f(U^\perp) \subset U^\perp$ .

*Beweis.* Es gilt für  $u \in U, v \in U^\perp$  und den zu  $f$  inversen Endomorphismus  $f^{-1}$

$$\langle f v, u \rangle = \langle v, f^{-1}(u) \rangle = 0,$$

weil wegen  $f(U) = U$  auch  $f^{-1}(U) = U$  ist. Also ist  $f v \in U^\perp$ , und deshalb gilt  $f(U^\perp) = U^\perp$ .  $\square$

Ist  $\lambda$  ein Eigenwert der Isometrie  $f$  mit dem Eigenvektor  $v$ , so gilt

$$\langle v, v \rangle = \langle f v, f v \rangle = \langle \lambda v, \lambda v \rangle = |\lambda|^2 \langle v, v \rangle,$$

also  $|\lambda| = 1, \lambda \in S^1$ . Ist  $V(\lambda)$  Eigenraum zu  $\lambda$ , so gilt  $f(V(\lambda)) \subset V(\lambda)$  und, da  $f$  injektiv ist, auch  $f(V(\lambda)) = V(\lambda)$ . Also läßt sich die vorstehende Notiz auf  $U = V(\lambda)$  anwenden.

**(4.2.5) Satz.** *Sei  $f$  ein unitärer Endomorphismus des unitären Raumes  $V$ . Dann ist  $V$  direkte Summe der Eigenräume von  $f$ . Eigenräume zu verschiedenen Eigenwerten sind orthogonal. Es besitzt  $V$  eine Orthonormalbasis aus Eigenvektoren.*

*Beweis.* Da  $f$  jedenfalls einen Eigenwert hat, weil ein charakteristisches Polynom über den komplexen Zahlen Nullstellen hat, folgt mittels (4.2.4) durch Induktion nach  $\dim V$  die erste Aussage des Satzes. Sei  $\lambda \neq \mu$ ,  $x \in V(\lambda)$ ,  $y \in V(\mu)$ . Es folgt  $\lambda\langle x, y \rangle = \langle x, \lambda y \rangle = \langle x, fy \rangle = \langle f^{-1}x, y \rangle = \langle \mu^{-1}x, y \rangle = \mu\langle x, y \rangle$ , denn wegen  $fy = \mu y$  ist  $\mu^{-1}y = f^{-1}y$  und wegen  $\mu \in S^1$  ist  $\mu^{-1} = \bar{\mu}$ . Die dritte Aussage folgt, indem man eine Orthonormalbasis in jedem  $V(\lambda)$  wählt.  $\square$

In die Matrixsprache übersetzt besagt (4.2.5) zusammen mit (4.2.2) und der allgemeinen Basiswechselformel:

**(4.2.6) Satz.** *Ist  $A \in U(n)$ , so gibt es  $S \in U(n)$ , so daß  $SAS^{-1} = \text{Dia}(\lambda_1, \dots, \lambda_n)$  mit  $\lambda_j \in S^1$ . Es sind die  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A$ . Deshalb ist die Diagonalmatrix bis auf Permutation der Diagonalelemente durch  $A$  eindeutig bestimmt.*  $\square$

**Zusatz.** Indem man  $S$  mit einer geeigneten Diagonalmatrix  $\text{Dia}(\lambda, \lambda, \dots, \lambda)$  multipliziert, kann man erreichen, daß  $S$  in (4.2.6) aus  $SU(n)$  ist.  $\square$

**(4.2.7) Satz.** *Sei  $V$  ein unitärer Raum und  $M$  eine Menge vertauschbarer unitärer Endomorphismen von  $V$  (das heißt  $f, g \in M \Rightarrow fg = gf$ ). Dann gibt es eine orthogonale Zerlegung  $V = V_1 \perp \dots \perp V_n$  in eindimensionale Unterräume  $V_j$ , die  $M$ -stabil sind (das heißt für  $f \in M$  gilt  $f(V_j) \subset V_j$ ).*

*Beweis.* Das ist klar, falls  $M$  nur aus der Identität besteht. Sei also  $f \in M$  nicht die Identität und  $V(\lambda)$  ein Eigenraum von  $f$  zu einem Eigenwert  $\lambda \neq 1$ . Falls  $V(\lambda) = V$  ist, so ist jede orthogonale Zerlegung von  $V$   $f$ -stabil, und wir können  $M \setminus \{f\}$  betrachten. Sei also  $V(\lambda) \neq V$ . Für  $g \in M$  und  $x \in V(\lambda)$  gilt dann wegen  $f(g)x = gf(x) = g(\lambda x) = \lambda g(x)$  auch  $g(x) \in V(\lambda)$ . Also sind  $V(\lambda)$  und  $V(\lambda)^\perp$   $M$ -stabil, und man kann Induktion nach  $\dim V$  anwenden.  $\square$

Matrizentheoretisch gewendet besagt (4.2.7), daß für eine Menge  $M \subset U(n)$  paarweise vertauschbarer Matrizen ein  $S \in U(n)$  existiert, so daß alle  $SAS^{-1}$ ,  $A \in M$ , Diagonalmatrizen sind (*simultane Diagonalisierung*).

Wir leiten nun analoge Sätze für orthogonale Endomorphismen her. Sei  $V$  ein  $n$ -dimensionaler euklidischer Raum und  $\rho: V \rightarrow V$  ein orthogonaler Endomorphismus, also Isomorphismus. Wissen wir schon: Ein Eigenwert  $\lambda$  von  $\rho$  erfüllt  $\lambda^2 = 1$ , und die Eigenräume  $V(1)$  und  $V(-1)$  zu den Eigenwerten 1 und  $-1$  sind zueinander orthogonal. Der Unterraum  $V_0 = (V(1) \oplus V(-1))^\perp$  ist  $\rho$ -stabil.

**(4.2.8) Lemma.** *Die Dimension von  $V_0$  ist gerade.*

*Beweis.* Andernfalls hätte das charakteristische Polynom von  $\psi = \rho | V_0: V_0 \rightarrow V_0$  ungeraden Grad und nach dem Zwischenwertsatz eine reelle Nullstelle. Dann hätte aber  $\psi$  einen Eigenraum zum Eigenwert 1 oder  $-1$  im Widerspruch zur Definition von  $V_0$ .  $\square$

**(4.2.9) Satz.** *Der Unterraum  $V_0$  ist die orthogonale Summe zweidimensionaler  $\rho$ -stabiler Unterräume.*

*Beweis.* Die Abbildung  $\rho + \rho^{-1}$  ist selbstadjungiert:

$$\langle (\rho + \rho^{-1})x, y \rangle = \langle \rho x, y \rangle + \langle \rho^{-1}x, y \rangle = \langle x, \rho^{-1}y \rangle + \langle x, \rho y \rangle = \langle x, (\rho + \rho^{-1})y \rangle.$$

Sie hat deshalb einen reellen Eigenwert, etwa  $\lambda$ , zum Eigenvektor  $e$ . Aus  $\rho e + \rho^{-1}e = \lambda e$  folgt  $\rho^2 e = \lambda \rho e - e$ . Die Vektoren  $\rho e$  und  $e$  sind linear unabhängig, da  $\rho$  in  $V_0$  keine Eigenvektoren hat. Also erzeugen  $e, \rho e$  einen zweidimensionalen Unterraum  $W$ , der wegen  $\rho^2 e = \lambda \rho e - e$  auch  $\rho$ -stabil ist. Wir betrachten nun das orthogonale Komplement von  $W$  in  $V_0$  und wenden Induktion nach der Dimension an.  $\square$

Sei  $W$  ein zweidimensionaler  $\rho$ -stabiler Unterraum in  $V_0$ , und sei  $A$  die Matrix von  $\rho: W \rightarrow W$  bezüglich einer Orthonormalbasis. Dann ist  $A$  orthogonal. Wäre  $\det A = -1$ , so hätte  $A$  die Eigenwerte 1 und  $-1$  und das widerspräche der Definition von  $V_0$ . Also ist  $A$  eine Drehmatrix.

Übersetzen wir die voranstehenden Ergebnisse in die Matrizensprache (Basiswechsel), so erhalten wir:

**(4.2.10) Satz.** *Jede Matrix  $A \in O(n)$  ist in  $O(n)$  konjugiert zu einer Matrix der folgenden Form (Blockdiagonalmatrix):*

$$\begin{pmatrix} D & & & & \\ & A_1 & & & \\ & & A_2 & & \\ & & & \ddots & \\ & & & & A_k \end{pmatrix}.$$

*Darin ist  $D$  eine quadratische Diagonalmatrix mit Elementen  $\pm 1$  auf der Diagonale und jedes  $A_j$  ist eine  $(2, 2)$ -Drehmatrix.*  $\square$

Allgemein wollen wir eine Matrix  $A$  aus  $SO(n)$  eine *Drehmatrix* und die dadurch vermittelte lineare Abbildung eine *Drehung* des  $\mathbb{R}^n$  nennen.

Wir betrachten jetzt speziell den Fall einer Drehung  $\rho$  des  $\mathbb{R}^3$ . Bezüglich einer geeigneten Orthonormalbasis wird die zugehörige lineare Abbildung  $\rho$  durch eine Matrix der Form

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}$$

beschrieben. Ist  $\rho \neq \text{id}$ , so ist der Eigenraum zum Eigenwert 1 eindimensional. Er wird *Drehachse* von  $\rho$  genannt. Jede Drehung des  $\mathbb{R}^3$  ist eine Drehung um eine Achse im üblichen anschaulich-geometrischen Sinn. In der Matrix von  $B$  ist  $\alpha$  der Drehwinkel und es gilt

$$\cos \alpha = \frac{1}{2}(\text{Sp}(B) - 1).$$

Da sich die Spur einer Matrix bei Basiswechsel nicht ändert, haben wir damit eine Formel gewonnen, mit der  $\cos \alpha$  ohne Bestimmung einer Normalform bestimmt werden kann.

Eine orthogonale Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  heie *Spiegelung* an der Hyperebene  $H$ , wenn der Eigenraum  $V_+$  zum Eigenwert 1 gleich  $H$  ist. Das orthogonale Komplement von  $V_+$  ist dann ein eindimensionaler Raum  $V_-$  zum Eigenwert  $-1$ . Ist  $a \in V$  von Null verschieden, so ist durch die Formel

$$s_a(v) = v - 2 \frac{\langle a, v \rangle}{\langle a, a \rangle} a$$

die Spiegelung  $s_a$  an der zu  $a$  orthogonalen Hyperebene gegeben. Wir hatten in (4.2.3) gesehen, da jedes Element von  $SO(2)$  Produkt zweier Spiegelungen ist. Daraus folgt leicht mittels (4.2.10):

**(4.2.11) Satz.** *Jedes Element von  $SO(2k)$  oder  $SO(2k + 1)$  ist Produkt von hchstens  $2k$  Spiegelungen. Jedes Element von  $O(n)$  ist Produkt von hchstens  $n$  Spiegelungen.*  $\square$

Sei  $A \in O(3) \setminus SO(3)$ . Dann hat  $A$  den Eigenwert  $-1$ . Es gibt also eine  $A$ -invariante Gerade. Im dazu orthogonalen Komplement wird gedreht.

## Erganzungen und Aufgaben

1. Wir interpretieren das Verfahren (4.1.6) durch Matrizen, indem wir es auf den  $\mathbb{R}^n$  mit dem Standardskalarprodukt anwenden. Sei also  $(x_1, \dots, x_n)$  eine Basis und  $(f_1, \dots, f_n)$  das Resultat der Orthonormalisierung. Setzen wir  $f_i = \sum_{k=1}^i b_{ki} x_k$ , so ist  $B = (b_{ki})$  eine obere Dreiecksmatrix, deren Diagonalelemente positiv sind. Sei  $X$  die Matrix mit den Spalten  $S_k(X) = x_k$  und  $F$  die Matrix mit den Spalten  $S_k(F) = f_k$ . Dann ist  $F$  eine orthogonale Matrix. Da auch  $B^{-1}$  eine obere Dreiecksmatrix mit positiver Diagonale ist, so sehen wir, da

$$Q(n) \times O(n) \rightarrow GL(n, \mathbb{R}), \quad (C, X) \rightarrow CX$$

surjektiv ist, wenn  $Q(n)$  die Gruppe der genannten Dreiecksmatrizen ist. Da  $Q(n) \cap O(n)$  nur die Einheitsmatrix enthalt, so ist diese Abbildung auch injektiv.

### 4.3 Selbstadjungierte Endomorphismen

Kreis, Ellipse und Hyperbel werden durch Gleichungen der Form  $ax^2 + cy^2 = r$  beschrieben. Setzen wir eine Koordinatentransformation  $x = \alpha u + \beta v$ ,  $y = \gamma u + \delta v$  in diese Gleichung ein, so erhält sie die Form  $Au^2 + 2Buv + Cv^2 = r$ . Analoge höherdimensionale Gebilde werden durch quadratische Gleichungen der Art  $\sum_{i,j} a_{ij}x_i x_j = r$  beschrieben, zum Beispiel liefert  $x^2 + y^2 + z^2 = 1$  die Oberfläche der Kugel vom Radius 1. Wenn wir  $a_{ij}$  durch  $\frac{1}{2}(a_{ij} + a_{ji})$  ersetzen, ändert sich die Gleichung nicht. Wir können also ohne Einschränkung mit einer symmetrischen Matrix  $A = (a_{ij})$  arbeiten. Dann läßt sich die fragliche Gleichung in der uns schon geläufigen Form

$$r = \sum_{i,j} a_{ij}x_i x_j = x^t A x = \langle Ax, x \rangle = \langle x, Ax \rangle$$

schreiben. Wir untersuchen zunächst derartige Formeln vom Standpunkt der linearen Algebra und kehren danach zu den geometrischen Figuren zurück.

Sei  $V$  ein endlichdimensionaler Hilbert-Raum mit Skalarprodukt  $\langle -, - \rangle$ . Ein Endomorphismus  $f$  von  $V$  heißt *selbstadjungiert*, wenn immer  $\langle f(u), v \rangle = \langle u, f(v) \rangle$  gilt. Eine reelle (komplexe)  $(n, n)$ -Matrix  $A$  heißt, wie erinnerlich, *symmetrisch (hermitesch)*, wenn  $\bar{A} = A^t$  gilt.

**(4.3.1) Notiz.** *Ein Endomorphismus ist genau dann selbstadjungiert, wenn seine Matrix bezüglich einer Orthonormalbasis symmetrisch (hermitesch) ist.*

*Beweis.* Sei  $A$  die Matrix von  $f$  bezüglich einer Orthonormalbasis. Daß  $f$  selbstadjungiert ist, bedeutet für beliebige Koordinatenvektoren  $x, y$  die Gleichheit  $x^t \bar{A}^t y = x^t A y$ . Allgemein folgt aus einer Gleichheit  $x^t C y = x^t D y$  für alle  $x, y$  die Gleichheit der Matrizen  $C, D$ : man setze für  $x, y$  die Standardbasis ein.  $\square$

**(4.3.2) Notiz.** *Sei  $f$  selbstadjungiert. Dann gilt:*

- (1) *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*
- (2) *Ist  $U \subset V$  Unterraum und  $f(U) \subset U$ , so ist auch  $f(U^\perp) \subset U^\perp$ .*
- (3) *Ein Eigenwert von  $f$  ist reell.*

*Beweis.* (1) Sei  $f(v) = \lambda v$ ,  $f(w) = \mu w$ ,  $\lambda \neq \mu$ . Es folgt

$$\lambda \langle v, w \rangle = \langle f(v), w \rangle = \langle v, f(w) \rangle = \mu \langle v, w \rangle$$

und daraus  $\langle v, w \rangle = 0$ .

(2) Sei  $u \in U$ ,  $v \in U^\perp$ . Dann ist  $\langle f(v), u \rangle = \langle v, f(u) \rangle = 0$ .

(3) Aus  $f v = \lambda v$ ,  $v \neq 0$  folgt nämlich  $\lambda \langle v, v \rangle = \langle v, \lambda v \rangle = \langle v, f v \rangle = \langle f v, v \rangle = \langle \lambda v, v \rangle = \bar{\lambda} \langle v, v \rangle$ , also  $\lambda = \bar{\lambda} \in \mathbb{R}$ .  $\square$

**(4.3.3) Satz.** *Ein selbstadjungierter Endomorphismus besitzt eine Orthonormalbasis aus Eigenvektoren.*

*Beweis.* Wir betrachten zunächst den Fall  $\mathbb{F} = \mathbb{C}$ . Dann gibt es jedenfalls einen Eigenvektor  $v$ , da das charakteristische Polynom über den komplexen Zahlen Nullstellen hat. Es erzeuge  $v$  den Unterraum  $U$ . Wegen (4.3.2) können wir den Satz für  $U^\perp$  durch Induktion nach der Dimension annehmen. Im Fall  $\mathbb{F} = \mathbb{R}$  stellen wir uns den Endomorphismus als durch eine symmetrische Matrix beschrieben vor. Eine solche ist aber auch hermitesch und hat deshalb einen reellen Eigenwert. Nun argumentieren wir wie im komplexen Fall.  $\square$

Wir notieren einige Konsequenzen des letzten Satzes. Ist  $V = \mathbb{F}^n$  und  $\langle -, - \rangle$  das Standardskalarprodukt, so liefert eine symmetrische (hermitesche)  $(n, n)$ -Matrix  $G = (g_{ij})$  einen selbstadjungierten Endomorphismus  $x \mapsto Gx$ . Ist  $T$  die Matrix, die die Standardbasis in die Orthonormalbasis aus Eigenvektoren überführt, so ist  $T$  eine orthogonale (unitäre) Matrix. Demnach impliziert (4.3.3) den nächsten Satz.

**(4.3.4) Satz.** *Ist  $G$  eine symmetrische (hermitesche)  $(n, n)$ -Matrix, so existiert  $T \in O(n)$  ( $T \in U(n)$ ), so daß  $T^{-1}GT = \bar{T}^tGT$  eine Diagonalmatrix ist.*  $\square$

Man kann übrigens in (4.3.4) offenbar  $T$  aus  $SO(n)$  bzw. aus  $SU(n)$  wählen. Eine Matrix aus  $SO(n)$  heißt *Drehmatrix* und die zugehörige lineare Abbildung des  $\mathbb{R}^n$  eine *Drehung*.

Eine symmetrische (hermitesche) Matrix  $P$  heißt, wie erinnerlich, *positiv definit*, wenn  $(x, y) \mapsto x^tPy$  ein Skalarprodukt ist.

**(4.3.5) Simultane Diagonalisierung.** Sei  $P$  eine positiv definite und  $G$  eine symmetrische reelle  $(n, n)$ -Matrix. Dann gibt es  $T \in GL(n, \mathbb{R})$ , so daß gilt:

- (1)  $T^tPT = I_n$ .
- (2)  $T^tGT$  ist Diagonalmatrix.

*Beweis.* Sei  $T_1$  eine Matrix, die die Standardbasis des  $\mathbb{R}^n$  in eine Orthonormalbasis für das Skalarprodukt  $(x, y) \mapsto x^tPy$  transformiert. Dann ist  $(x, y) \mapsto x^tT_1^tPT_1y$  ein Skalarprodukt, mit  $e_1, \dots, e_n$  als Orthonormalbasis. Also gilt  $T_1^tPT_1 = I_n$ . Es ist  $T_1^tGT_1$  symmetrisch. Man wähle  $T_2 \in O(n)$ , so daß  $T_2^tT_1^tGT_1T_2$  eine Diagonalmatrix ist und setze  $T = T_1T_2$ .  $\square$

Einen analogen Satz gibt es natürlich für komplexe, positiv definite  $P$  und hermitesche  $G$ .

Wir besprechen nun eine weitere Methode, reelle Eigenwerte zu finden. Die *Norm* eines Endomorphismus  $f$  wird durch

$$\|f\| = \text{Maximum } \{|f(v)| \mid 1 = |v|\}$$

definiert. (Eine Anleihe aus der Analysis besagt: Dieses Maximum existiert. Zunächst überlegt man sich, daß die Menge  $\{|f(v)| \mid 1 = |v|\}$  beschränkt ist. Sie hat also ein Supremum  $\|f\|$ . Man wählt eine Folge  $(w_k)$  mit  $\lim_{k \rightarrow \infty} |f(w_k)| = \|f\|$ . Man schreibt  $w_k$  als Linearkombination einer Orthonormalbasis; die Komponenten sind beschränkt und konvergieren nach Übergang zu einer Teilfolge.

Also können wir annehmen, daß  $(w_k)$  gegen  $w$  konvergiert. Aus Gründen der Stetigkeit folgt  $|w| = 1$  und  $|f(w)| = \|f\|$ . Ebenso schließt man, daß die Menge  $\{|\langle f(v), v \rangle| \mid |v| = 1\}$  ein Maximum enthält.)

Die Norm  $\|f\|$  läßt sich auch als das kleinste reelle  $C$  charakterisieren, so daß für alle  $v \in V$  die Ungleichung

$$|f(v)| \leq C|v|$$

gilt. Aus dieser Ungleichung folgt nämlich  $|f(v)| \leq C$  für  $|v| = 1$ , also  $\|f\| \leq C$ . Ist  $v \neq 0$  gegeben, so ist ein  $v|v|^{-1}$  Einheitsvektor, also gilt  $|f(v|v|^{-1})| \leq \|f\|$  und deshalb auch  $|f(v)| \leq \|f\||v|$ .

**(4.3.6) Notiz.** Sei  $f$  selbstadjungiert. Dann gilt

$$\|f\| = \text{Max}\{|\langle f v, v \rangle| \mid |v| = 1\} = M.$$

*Beweis.* Die Schwarzsche Ungleichung liefert für  $|v| = 1$

$$|\langle f v, v \rangle| \leq |f v| |v| \leq \|f\|.$$

Also gilt  $M \leq \|f\|$ . Wir wollen nun  $\|f\| \leq M$ , also  $|f(v)| \leq M$  für  $|v| = 1$  zeigen. Wir dürfen  $f(v) \neq 0$  annehmen und setzen  $w = |f(v)|^{-1} f(v)$ . Dann ist

$$\langle f v, w \rangle = |f(v)| = \langle v, f w \rangle$$

also

$$\langle f(v+w), v+w \rangle = \langle f v, v \rangle + 2|f v| + \langle f w, w \rangle$$

$$\langle f(v-w), v-w \rangle = \langle f v, v \rangle - 2|f v| + \langle f w, w \rangle$$

durch Subtraktion dieser Gleichungen ergibt sich

$$4|f v| = \langle f(v+w), v+w \rangle - \langle f(v-w), v-w \rangle \leq M|v+w|^2 + M|v-w|^2 = 4M,$$

wie gewünscht (die Ungleichung folgt mit der Definition von  $M$  und die letzte Gleichung durch bilineares Ausrechnen).  $\square$

**(4.3.7) Satz.** Ist  $f$  ein selbstadjungierter Endomorphismus, so ist  $\|f\|$  oder  $-\|f\|$  ein Eigenwert von  $f$ .

*Beweis.* Wir wählen einen Einheitsvektor  $v \in V$ , so daß  $\langle v, f v \rangle = \alpha$  und  $|\alpha| = \|f\|$  ist, das heißt das Maximum von  $\langle f x, x \rangle$  werde (nach (4.3.6)) in  $x = v$  angenommen. Dann folgt

$$\begin{aligned} 0 &\leq |f(v) - \alpha v|^2 = |f(v)|^2 - \alpha \langle f v, v \rangle - \bar{\alpha} \langle v, f v \rangle + |\alpha|^2 |v|^2 \\ &= |f(v)|^2 - 2\bar{\alpha} \alpha + |\alpha|^2 |v|^2 \leq 0, \end{aligned}$$

mit der letzten Ungleichung nach (4.3.6). Also gilt  $f(v) = \alpha v$ .  $\square$

Wir geben schließlich (für den reellen Fall) noch einen weiteren analytischen Beweis dafür, daß ein Einheitsvektor, der  $\langle v, fv \rangle$  maximiert, ein Eigenvektor ist. Sei zu diesem Zweck  $w$  irgendein Einheitsvektor, der zu  $v$  orthogonal ist. Wir wählen eine stetig differenzierbare Kurve  $a: [-1, 1] \rightarrow S^{n-1}$  mit den Eigenschaften:  $a(0) = v$ ,  $a'(0) = w$ , zum Beispiel  $a(t) = \cos t \cdot v + \sin t \cdot w$ . Dann hat die Funktion  $t \mapsto \langle a(t), fa(t) \rangle$  an der Stelle  $t = 0$  ein Maximum und folglich die Ableitung Null. Diese Ableitung ist  $\langle fv, w \rangle$ . Demnach ist  $fv$  zu allen  $w$  orthogonal und daher ein Vielfaches von  $v$ . Ebenso kann man mit dem Minimum argumentieren.

Es gibt eine gemeinsame Verallgemeinerung der Diagonalisierbarkeit unitärer und hermitescher Matrizen. Eine Matrix  $A$  heie *normal*, wenn gilt

$$\bar{A}^t \cdot A = A \cdot \bar{A}^t.$$

**(4.3.8) Satz.** Eine Matrix  $A \in M(n, n; \mathbb{C})$  ist genau dann normal, wenn es  $S \in U(n)$  so gibt, da  $SAS^{-1}$  eine Diagonalmatrix ist.

*Beweis.* Ist  $A$  normal und  $S \in U(n)$ , so ist  $SAS^{-1}$  normal. Eine Diagonalmatrix ist normal. Sei also  $A$  normal. Fr einen Eigenraum  $V(\lambda)$  gilt dann  $\bar{A}^t V(\lambda) \subset V(\lambda)$ , denn fr  $v \in V(\lambda)$  ist  $A(\bar{A}^t v) = \bar{A}^t AV = \bar{A}^t \lambda v = \lambda(\bar{A}^t v)$ , also  $\bar{A}^t v \in V(\lambda)$ . Man betrachte die Einschrnkung  $\bar{A}^t|_{V(\lambda)}: V(\lambda) \rightarrow V(\lambda)$ , die selbst einen Eigenwert  $\mu$  und einen Eigenraum  $V(\lambda, \mu)$  hat. Es gilt  $AV(\lambda, \mu) \subset V(\lambda, \mu)$ , denn fr  $v \in V(\lambda, \mu)$  ist  $\bar{A}^t(Av) = A\bar{A}^t v = A\mu v = \mu(Av)$ . Es gilt ebenfalls  $AV(\lambda, \mu)^\perp \subset V(\lambda, \mu)^\perp$ , denn aus  $v \in V(\lambda, \mu)$  und  $\langle v, u \rangle = 0$  folgt  $\langle v, Au \rangle = \langle \bar{A}^t v, u \rangle = \langle \mu v, u \rangle = \mu \langle v, u \rangle = 0$ . Ebenso gilt  $\bar{A}^t V(\lambda, \mu)^\perp \subset V(\lambda, \mu)^\perp$ . Also ist die Einschrnkung von  $A$  auf  $V(\lambda, \mu)^\perp$  ein normaler Endomorphismus, und man kommt durch Induktion nach  $\dim V$  ans Ziel.  $\square$

Wir kommen nun zu den eingangs erwhnten Anwendungen.

**(4.3.9) Hauptachsen von Quadriken.** Zunchst einige allgemeine Vorbemerkungen. Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  eine Abbildung und

$$N(f, c) = \{x \in \mathbb{R}^n \mid f(x) = c\} = f^{-1}(c)$$

die  $c$ -Stellenmenge von  $f$ . Sei  $A \in GL(n, \mathbb{R})$  und  $L_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x \mapsto Ax$  der zugehrige lineare Automorphismus. Die Wahl eines neuen Koordinatensystems bedeutet, da wir eine Abbildung der Form  $f \circ L_A = g$  betrachten. In den neuen Koordinaten wird  $N(f, c)$  also  $N(g, c)$ . Wir knnen aber auerdem das Bild der Menge  $N(f, c)$  bei  $L_A$  betrachten, und dann gilt

$$L_A N(f, c) = N(f \circ L_A^{-1}, c).$$

Sei  $C = (c_{kl})$  eine symmetrische reelle Matrix. Damit betrachten wir die Funktion

$$f: \mathbb{R}^n \rightarrow \mathbb{R}, \quad x = (x_1, \dots, x_n) \mapsto \sum_{k,l}^n c_{kl} x_k x_l = x^t C x.$$

Eine Menge der Form  $f^{-1}(c)$  nennen wir eine *Quadrik*. In neuen Koordinaten haben wir die Funktion  $f \circ L_A: x \mapsto x^t A^t C A x$  zu betrachten, also wiederum eine Quadrik. Wir wissen nach (4.3.4), daß wir eine Drehung  $A$  so finden können, daß  $A^t C A = \text{Dia}(\lambda_1, \dots, \lambda_n)$  ist. In den neuen Koordinaten hat die Gleichung der Quadrik die Form  $\sum_j \lambda_j x_j^2 = r$ , hat sich also beträchtlich vereinfacht.

Ebenso sehen wir, daß das Bild einer Quadrik bei einem linearen Automorphismus wieder eine Quadrik ist. Wir können aus demselben Grund durch eine Drehung eine Quadrik in eine Lage bringen, in der sie als Punktmenge durch eine Gleichung der Form  $\sum_j \lambda_j x_j^2 = r$  beschrieben wird. Durch Zurückdrehen der Koordinatenachsen erhalten wir die sogenannten *Hauptachsen* der Quadrik, und die Drehung in die genannte Normalform heißt die *Hauptachsentransformation* der Quadrik.  $\diamond$

## 4.4 Bilineare Abbildungen

Wir stellen einige allgemeine Aussagen über bilineare Abbildungen zusammen. Spezialfälle sind uns schon begegnet. Seien  $U, V$  und  $W$  Vektorräume über dem Körper  $K$ . Eine Abbildung  $s: U \times V \rightarrow W$  heißt *bilinear*, wenn für jedes  $v \in V$  die Abbildung  $U \rightarrow W, u \mapsto s(u, v)$  und für jedes  $u \in U$  die Abbildung  $V \rightarrow W, v \mapsto s(u, v)$  linear ist. Ist  $W = K$ , so sprechen wir auch von einer *Bilinearform*.

Eine bilineare Abbildung  $s: U \times V \rightarrow W$  hat zwei *adjungierte lineare Abbildungen*

$$\begin{array}{lll} s_L: U \rightarrow \text{Hom}(V, W), & u \mapsto s(u, ?) & \text{links} \\ s_R: V \rightarrow \text{Hom}(U, W), & v \mapsto s(? , v) & \text{rechts.} \end{array}$$

Dabei ist  $s(u, ?)$  das Symbol für  $v \mapsto s(u, v)$ .

Wir bezeichnen mit  $\text{Bil}(U \times V, W)$  die Menge der bilinearen Abbildungen  $U \times V \rightarrow W$ . Im Falle der Bilinearformen lassen wir das Symbol  $W$  weg. Durch Addition und skalare Multiplikation von Funktionswerten trägt diese Menge wie üblich eine Vektorraumstruktur. Aus den Definitionen verifiziert man sofort:

**(4.4.1) Notiz.** Die Zuordnung  $s \mapsto s_L$  liefert einen Isomorphismus

$$\Lambda: \text{Bil}(U \times V, W) \rightarrow \text{Hom}(U, \text{Hom}(V, W))$$

von Vektorräumen.  $\square$

Im weiteren Verlauf dieses Abschnitts betrachten wir nur noch Bilinearformen. Eine Bilinearform  $U \times V \rightarrow K$  nennen wir auch eine *Paarung* von  $U$  mit  $V$ . Zunächst geben wir eine Beschreibung durch Matrizen an. Sei  $B = \{b_1, \dots, b_m\}$

eine Basis von  $U$  und  $C = \{c_1, \dots, c_n\}$  eine Basis von  $V$ . Dann gilt für eine Bilinearform  $s: U \times V \rightarrow K$

$$s(\sum_i \lambda_i b_i, \sum_j \mu_j c_j) = \sum_{i,j} \lambda_i \mu_j s(b_i, c_j),$$

und daraus entnehmen wir, daß die zu  $s$  gehörende Matrix  $S = (s(b_i, c_j))$  die Form  $s$  bestimmt und jede  $(m, n)$ -Matrix auf diese Weise aus einer Bilinearform entsteht. Wir haben somit einen Isomorphismus

$$m_{B,C}: \text{Bil}(U \times V) \rightarrow M(m, n), \quad s \mapsto (s(b_i, c_j))$$

von Vektorräumen. Fassen wir  $K^m$  und  $K^n$  als Räume von Spaltenvektoren auf, so liefert eine  $(m, n)$ -Matrix  $S$  die Bilinearform

$$\sigma: K^m \times K^n \rightarrow K, \quad (x, y) \mapsto x^t S y,$$

und bezüglich der Standardbasen  $B$  und  $C$  ist  $m_{B,C}(\sigma) = S$ . Sind  $\Phi = (\varphi_{kl})$  und  $\Psi = (\psi_{rs})$  Basiswechselmatrizen in  $U$  und  $V$  (neue Basis durch die alte!), so rechnen wir

$$s(b'_i, c'_j) = s(\sum_k \varphi_{ki} b_k, \sum_l \psi_{lj} c_l) = \sum_{k,l} \varphi_{ki} s(b_k, c_l) \psi_{lj}$$

und erkennen, daß die neue Matrix  $S'$  sich aus der alten  $S$  durch

$$S' = \Phi^t S \Psi$$

ausrechnen läßt.

Die adjungierte Abbildung  $s_L$  zu einer Bilinearform  $s: U \times V \rightarrow K$  ist eine lineare Abbildung  $U \rightarrow V^*$  in den Dualraum und ebenso  $s_R: V \rightarrow U^*$ . Wir wählen Dualbasen  $b^1, \dots, b^m$  zu  $(b_1, \dots, b_m)$  und  $c^1, \dots, c^n$  zu  $(c_1, \dots, c_n)$ . Dann gilt

$$s_L(b_i) = \sum_j s(b_i, c_j) c^j, \quad s_R(c_i) = \sum_j s(b_j, c_i) b^j.$$

Zum Beweis (der ersten Gleichung) zeigen wir, daß beide Linearformen auf  $c_k$  denselben Wert annehmen:

$$(s_L(b_i))(c_k) = s(b_i, c_k)$$

$$(\sum_j s(b_i, c_j) c^j)(c_k) = \sum_j s(b_i, c_j) c^j(c_k) = \sum_j s(b_i, c_j) \delta_k^j = s(b_i, c_k).$$

Damit haben wir gezeigt:

**(4.4.2) Notiz.** Die Matrix von  $s_L$  bezüglich der Basen  $(b_i)$  und  $(c^j)$  ist die transponierte Matrix  $S^t$ . Die Matrix von  $s_R$  bezüglich der Basen  $(c_i)$  und  $(b^j)$  ist  $S$ .  $\square$

Eine Bilinearform  $s: U \times V \rightarrow K$  heißt *regulär*, wenn die beiden adjungierten Abbildungen  $s_L: U \rightarrow V^*$  und  $s_R: V \rightarrow U^*$  injektiv sind. Eine reguläre Form heißt auch *Dualität* zwischen  $U$  und  $V$ . Im Falle endlichdimensionaler  $U$  und  $V$  gilt für reguläres  $s$  dann

$$\dim U \leq \dim V^* = \dim V \leq \dim U^* = \dim U;$$

also besteht überall Gleichheit, und die adjungierten Abbildungen sind Isomorphismen. Aus den Matrizenrechnungen folgt:

**(4.4.3) Notiz.** *Eine Bilinearform ist genau dann regulär, wenn die zugehörige Matrix regulär ist.*  $\square$

**(4.4.4) Beispiele.**

(1) Das Standardskalarprodukt auf dem  $\mathbb{R}^n$ .

(2) Das Einsetzen in Linearformen  $V^* \times V \rightarrow K$ ,  $(\alpha, x) \mapsto \alpha(x)$ . Durch Bilinearformen dieser Art werden Funktionssymbol und Variablensymbol gleichwertig behandelt. Man kann Vektoren in ein Funktional einsetzen; man kann Funktionale an einem Vektor auswerten.

(3) Sei  $C([a, b], \mathbb{R}) = V$  der Vektorraum der stetigen Funktionen  $[a, b] \rightarrow \mathbb{R}$  und  $p: [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion. Dann ist  $(f, g) \mapsto \int_a^b f(x)g(x)p(x)dx$  eine Bilinearform auf  $V$ . Ist  $p \geq 0$  nicht die Nullfunktion, so ist die Form ein Skalarprodukt.

(4) Verkettung von linearen Abbildungen und Matrizenprodukt.  $\diamond$

Für Bilinearformen  $s: V \times V \rightarrow K$  verwenden wir natürlich in beiden Variablen dieselbe Basis. Eine solche Form heißt *symmetrisch*, wenn immer  $s(u, v) = s(v, u)$  gilt; *alternierend*, wenn immer  $s(x, x) = 0$  gilt; *schiefsymmetrisch*, wenn immer  $s(x, y) = -s(y, x)$  gilt. Eine alternierende Form ist schief-symmetrisch, wie durch bilineares Ausrechnen von  $s(x + y, x + y)$  folgt. Sei  $(b_1, \dots, b_n)$  eine Basis von  $V$  und  $S = (s(b_k, b_l))$  die Matrix von  $s$ . Dann ist  $s$  genau dann symmetrisch, wenn  $S^t = S$  gilt; genau dann schief-symmetrisch, wenn  $S^t = -S$  gilt; genau dann alternierend, wenn  $S^t = -S$  gilt und die Diagonale von  $S$  nur Nullen enthält. Wir nennen Matrizen mit diesen Eigenschaften entsprechend symmetrisch, schief-symmetrisch oder alternierend.

Ein Isomorphismus  $f: V \rightarrow V$  heißt *Isometrie* von  $s$ , wenn immer  $s(x, y) = s(fx, fy)$  gilt. Die Isometrien von  $s$  bilden bezüglich Verkettung eine Gruppe. Sei  $S$  die Matrix von  $s$  bezüglich  $\{b_1, \dots, b_n\}$  und  $A$  die Matrix von  $f$  bezüglich ebendieser Basis. Dann ist  $f$  genau dann eine  $s$ -Isometrie, wenn die Gleichung

$$S = A^t S A$$

gilt. Ist  $S$  regulär, so folgt aus dieser Gleichung, daß  $A$  regulär ist und die Determinante  $\pm 1$  hat. Die Gruppe der  $s$ -Isometrien ist also isomorph zur Gruppe der invertierbaren Matrizen  $A$ , die die Gleichung  $S = A^t S A$  erfüllen. Eine solche Situation hatten wir schon bei den orthogonalen Matrizen angetroffen. Die

unitären Matrizen fügen sich diesem Schema nicht ganz, da das Konjugieren der komplexen Zahlen gebraucht wird. In einem formalen Kontext kann man stattdessen einen Körperautomorphismus der Ordnung zwei verwenden (Sesquilinearformen).

## 4.5 Adjungierte Endomorphismen

Seien  $U$  und  $V$  endlichdimensionale Vektorräume über dem Körper  $K$ , und sei  $s: U \times V \rightarrow K$  eine reguläre Bilinearform.

Für jeden Endomorphismus  $f: U \rightarrow U$  ist

$$\rho(f, s): U \times V \rightarrow K, \quad (u, v) \mapsto s(f(u), v)$$

wieder eine Bilinearform. Ist  $s_L: U \rightarrow V^*$  die Linksadjungierte zu  $s$ , so ist die Zusammensetzung  $s_L \circ f: U \rightarrow V^*$  die Linksadjungierte zu der neuen Form  $\rho(f, s)$ . Da  $s$  regulär ist und folglich  $s_L$  ein Isomorphismus, hat jede lineare Abbildung  $g: U \rightarrow V^*$  die Form  $s_L \circ f$  für genau einen Endomorphismus  $f$  von  $U$ . Da  $s \mapsto s_L$  ein Isomorphismus  $\text{Bil}(U \times V) \rightarrow \text{Hom}(U, V^*)$  ist, so ist auch

$$\text{Hom}(U, U) \rightarrow \text{Bil}(U \times V), \quad f \mapsto \rho(f, s)$$

ein Isomorphismus.

Analog hat man für einen Endomorphismus  $g$  von  $V$  die Bilinearform

$$\lambda(s, g): U \times V \rightarrow K, \quad (u, v) \mapsto s(u, g(v))$$

und die Zuordnung

$$\text{Hom}(V, V) \rightarrow \text{Bil}(U \times V), \quad g \mapsto \lambda(s, g)$$

ist ein Isomorphismus.

Zu jedem Endomorphismus  $f$  von  $U$  gibt es deshalb genau einen Endomorphismus von  $V$ , so daß für alle  $u \in U$  und  $v \in V$  die Gleichung

$$s(f(u), v) = s(u, g(v))$$

gilt. Wir setzen  $g = f^\#$  und nennen  $f^\#$  die bezüglich  $s$  zu  $f$  *adjungierte Abbildung*. Analog setzen wir  $f = g_\#$ . Es gelten dann Relationen  $(f^\#)_\# = f$  und  $(g_\#)^\# = g$ .

**(4.5.1) Satz.** *Das Diagramm*

$$\begin{array}{ccc} V & \xrightarrow{s_R} & U^* \\ \uparrow f^\# & & \uparrow f^* \\ V & \xrightarrow{s_R} & U^* \end{array}$$

*ist kommutativ.*

*Beweis.* Die Verifikation aus den Definitionen ist leicht und durch das folgende Schicksal von  $v \in V$  angedeutet:

$$\begin{array}{ccc} f^\#v & \longrightarrow & s(?, f^\#v) = s(f(?), v) \\ \uparrow & & \uparrow \\ v & \longrightarrow & s(?, v) \end{array}$$

Das Diagramm klärt den Zusammenhang zwischen der dualen Abbildung und der adjungierten Abbildung.  $\square$

Es gibt ein analoges Diagramm für  $s_L$  und  $g_\#$ .

Wir wählen nun eine Basis  $B = \{b_1, \dots, b_n\}$  von  $U$  und  $C = \{c_1, \dots, c_n\}$  von  $V$ . In  $U^*$  verwenden wir die duale Basis  $B^* = \{b_1^*, \dots, b_n^*\}$ . Die Matrix von  $s_R$  bezüglich dieser Basen ist  $S = (s(b_i, c_j))$ , wie wir in (4.4.2) gesehen haben.

Hat  $f$  die Matrix  $A$  bezüglich  $B$ , so hat nach (4.5.1)  $f^\#$  die Matrix von  $s_R^{-1} \circ f^* \circ s_R$ , also die Matrix  $S^{-1}A^tS$  bezüglich  $C$ . Aus (4.5.1) entnehmen wir  $(f_2f_1)^\# = f_1^\#f_2^\#$  und  $(\text{id})^\# = \text{id}$ . Analog für die Kreuze unten. Bedeutsam ist der Fall  $U = V$ ,  $B = C$ . Wir nennen in diesem Fall  $f$  *selbstadjungiert* bezüglich  $s$ , wenn  $f = f^\#$  ist, wenn also immer  $s(f(u), v) = s(u, f(v))$  gilt. Mit den eben verwendeten Matrizen ist das genau dann der Fall, wenn  $SA = A^tS$  gilt. Wir notieren:

**(4.5.2) Notiz.** *Ist  $s$  symmetrisch, so ist die Bilinearform  $(u, v) \mapsto s(f(u), v)$  genau dann symmetrisch, wenn  $f$  selbstadjungiert ist.*

*Beweis.* Ist die Form symmetrisch, so hat man  $s(f(u), v) = s(f(v), u)$ . Ersteres ist gleich  $s(u, f^\#(v))$  und letzteres wegen der Symmetrie von  $s$  gleich  $s(u, f(v))$ . Es folgt  $f = f^\#$ . Analog folgt aus  $f = f^\#$  die Symmetrie der Form  $\rho(f, s)$ .  $\square$

Ist  $B$  eine Orthonormalbasis, also  $S$  die Einheitsmatrix, so ist  $f$  genau dann selbstadjungiert, wenn die Matrix  $A$  bezüglich dieser Basis symmetrisch ist, das heißt  $A = A^t$  gilt. Allgemein ist bei Verwendung von Orthonormalbasen  $A^t$  die Matrix von  $f^\#$ , wenn  $A$  die Matrix von  $f$  ist.

## 4.6 Geometrische Bilinearformen

Sei  $V$  ein endlichdimensionaler Vektorraum über  $K$  und  $s: V \times V \rightarrow K$  eine Bilinearform. Ein geordnetes Paar  $(u, v)$  von Vektoren aus  $V$  heißt *s-orthogonal* oder *orthogonal* bezüglich  $s$ , wenn  $s(u, v) = 0$  ist. Die Form  $s$  nennen wir *geometrisch*, wenn aus  $s(u, v) = 0$  immer  $s(v, u) = 0$  folgt, wenn also Orthogonalität eine symmetrische Relation ist. Wir betrachten im weiteren nur geometrische Formen. Symmetrische und schiefsymmetrische Formen sind geometrisch.

Ist  $U \subset V$  ein Unterraum, so setzen wir

$$U^\perp = \{v \in V \mid \text{für alle } u \in U \text{ ist } s(u, v) = 0\}.$$

Im Gegensatz zum Fall der Skalarprodukte gibt es jetzt im allgemeinen keine direkte Zerlegung  $V = U \oplus U^\perp$ .

Ein Vektor  $v$  heißt *isotrop* bezüglich  $s$ , wenn  $s(v, v) = 0$  ist. Ein Unterraum heißt *isotrop*, wenn er nur isotrope Vektoren enthält. Gegenteil: *anisotrop*. Der Unterraum  $V^\perp$  wird *Radikal* der Form  $s$  auf  $V$  genannt. Eine Form ist genau dann regulär, wenn ihr Radikal gleich Null ist.

**(4.6.1) Satz.** *Sei  $U$  ein Unterraum, der keinen von Null verschiedenen isotropen Vektor enthält. Dann besteht eine direkte Zerlegung  $V = U \oplus U^\perp$ .*

*Beweis.* Die Voraussetzung besagt  $U \cap U^\perp = 0$ . Der Raum  $U^\perp$  ist der Kern der Abbildung

$$V \xrightarrow{s_R} V^* \xrightarrow{j^*} U^*,$$

worin  $j: U \rightarrow V$  die Inklusion ist. Nach den Dimensionsformeln gilt

$$\dim V \geq \dim(U + U^\perp) = \dim U^\perp + \dim U^* \geq \dim V.$$

Es folgt  $V = U + U^\perp$ . □

Wie im ersten Abschnitt definieren wir: Unterräume  $U_1$  und  $U_2$  heißen *orthogonal*, wenn für  $u_i \in U_i$  immer  $s(u_1, u_2) = 0$  gilt. Ist  $V$  die direkte Summe paarweise orthogonaler Unterräume ( $U_j \mid 1 \leq j \leq t$ ), so sagen wir,  $V$  sei die *orthogonale Summe* der  $U_j$ . Wir notieren diesen Sachverhalt durch

$$V = U_1 \perp \dots \perp U_t.$$

**(4.6.2) Satz.** *Sei  $s$  regulär. Seien  $U, U_0, U_1$  Unterräume von  $V$ . Dann gilt:*

- (1)  $U_0 \subset U_1 \Rightarrow U_0^\perp \supset U_1^\perp$ .
- (2)  $\dim U + \dim U^\perp = \dim V$ .
- (3)  $(U_0 \cap U_1)^\perp = U_0^\perp + U_1^\perp$ .
- (4)  $(U_0 + U_1)^\perp = U_0^\perp \cap U_1^\perp$ .
- (5)  $(U^\perp)^\perp = U$ .

*Beweis.* (1) ist klar.

(2) Wir verwenden wieder die Abbildung

$$V \xrightarrow{s_R} V^* \xrightarrow{j^*} U^*.$$

Da  $j$  injektiv ist, ist  $j^*$  surjektiv. Da  $s_R$  bijektiv ist, folgt die Behauptung aus  $\dim \text{Kern } f + \dim \text{Bild } f = \dim V$  für  $f = j^* \circ s_R$ .

(5) Aus den Definitionen entnimmt man  $U \subset (U^\perp)^\perp$ . Aus (2) entnimmt man  $\dim U = \dim(U^\perp)^\perp$ .

(3), (4) Aus  $U_0 \cap U_1 \subset U_i$  folgt  $(U_0 \cap U_1)^\perp \supset U_i^\perp$  und deshalb  $(U_0 \cap U_1)^\perp \supset U_0^\perp + U_1^\perp$ . Aus  $U_i \subset U_0 + U_1$  folgt  $U_i^\perp \supset (U_0 + U_1)^\perp$  und deshalb  $U_0^\perp \cap U_1^\perp \supset (U_0 + U_1)^\perp$  sowie  $(U_0^\perp \cap U_1^\perp)^\perp \subset U_0 + U_1$ . Wenden wir die letzte Inklusion auf das Paar  $U_0^\perp, U_1^\perp$  statt auf  $U_0, U_1$  an, so folgt  $(U_0 \cap U_1)^\perp \subset U_0^\perp + U_1^\perp$ . Die umgekehrte Inklusion hatten wir aber schon oben gezeigt. Damit haben wir (3) gezeigt. Analog folgt (4).  $\square$

Eine Basis  $b_1, \dots, b_n$  heißt *Orthogonalbasis*, falls für  $i \neq j$  immer  $s(b_i, b_j) = 0$  ist. Die Matrix  $S$  von  $s$  bezüglich einer Orthogonalbasis ist eine Diagonalmatrix (und umgekehrt).

Wir sagen, ein Körper habe die *Charakteristik zwei*, wenn  $1 + 1 = 0$  ist.

**(4.6.3) Satz.** *Der Körper  $K$  habe nicht die Charakteristik 2. Dann gibt es zu jeder symmetrischen Bilinearform  $s$  eine Orthogonalbasis. (Es muß  $s$  hierzu nicht regulär sein.)*

*Beweis.* Induktion nach  $\dim V$ . Für  $\dim V = 1$  können wir irgendeine Basis nehmen. Falls es ein  $b \in V$  mit  $s(b, b) \neq 0$  gibt, so sei  $U$  der von  $b$  erzeugte Unterraum. Ist  $b_2, \dots, b_m$  eine Orthogonalbasis von  $U^\perp$  (bezüglich der durch  $s$  auf  $U^\perp$  gegebenen Form), so ist  $b, b_2, \dots, b_m$  eine Orthogonalbasis. Es bleibt der Fall, daß für alle  $b \in V$  die Gleichheit  $s(b, b) = 0$  gilt. Dann ist

$$s(b + c, b + c) = s(b, b) + s(b, c) + s(c, b) + s(c, c) = 2s(b, c),$$

also  $s(b, c) = 0$  für irgend zwei Vektoren  $b$  und  $c$ , da  $K$  nicht die Charakteristik zwei hat. Jede Basis ist dann eine Orthogonalbasis.  $\square$

Interpretieren wir (4.6.3) durch Matrizen und Basiswechsel, so erhalten wir:

**(4.6.4) Satz.** *Hat  $K$  nicht die Charakteristik 2, und ist  $S$  eine symmetrische Matrix, so gibt es eine invertierbare Matrix  $T$  derart, daß  $T^t S T$  eine Diagonalmatrix ist.*  $\square$

## 4.7 Bilinearformen auf reellen Vektorräumen

Wir betrachten symmetrische Bilinearformen  $s: V \times V \rightarrow \mathbb{R}$  auf Vektorräumen  $V$  über  $\mathbb{R}$ . Eine solche Form heißt *positiv definit* (*negativ definit*), wenn für alle  $v \in V \setminus \{0\}$  stets  $s(v, v) > 0$  (stets  $s(v, v) < 0$ ) gilt. Eine positiv definite Form  $s$  ist ein Skalarprodukt auf  $V$ . Eine Orthogonalbasis  $b_1, \dots, b_n$  heißt *orthonormal*, wenn immer  $s(b_i, b_i) \in \{\pm 1\}$  ist. Ist  $\{b_1, \dots, b_n\}$  Orthogonalbasis, so gibt

es  $\lambda_j > 0$ , so daß  $\{\lambda_1 b_1, \dots, \lambda_n b_n\}$  orthonormal ist. Nach (4.6.4) hat jede reguläre symmetrische Form auf einem endlichdimensionalen  $\mathbb{R}$ -Vektorraum eine Orthonormalbasis.

Sei im folgenden  $V$  endlichdimensional und  $s$  regulär, soweit nichts anderes gesagt wird. Ein Unterraum  $U$  von  $V$  heie *positiv* (bezüglich  $s$ ), wenn für alle  $x \in U \setminus \{0\}$  die Ungleichung  $s(x, x) > 0$  gilt. Analog: *negativ*. Ein positiver Unterraum  $U$  heie *maximal*, wenn jeder Unterraum, der  $U$  echt umfat, nicht mehr positiv ist.

**(4.7.1) Satz.** Sei  $\{b_1, \dots, b_r, c_1, \dots, c_s\}$  eine Orthonormalbasis von  $V$ . Es gelte  $s(b_i, b_i) = 1$  und  $s(c_j, c_j) = -1$ . Sei  $V(+)$  bzw.  $V(-)$  der von den  $b_1, \dots, b_r$  bzw.  $c_1, \dots, c_s$  erzeugte Unterraum. Dann gilt:  $V(+)$  ist ein maximaler positiver und  $V(-)$  ist ein maximaler negativer Unterraum von  $V$ . Ferner ist  $V$  die orthogonale Summe von  $V(+)$  und  $V(-)$ .

*Beweis.* Es ist  $s(\sum \lambda_i b_i, \lambda_i b_i) = \sum \lambda_i^2$ . Deshalb ist  $V(+)$  positiv. Analog folgt, da  $V(-)$  negativ ist. Ist  $u \in V(+)$  und  $v \in V(-)$ , so folgt  $s(u, v) = 0$ , weil die Basisvektoren  $b_i$  und  $c_j$  paarweise zueinander orthogonal sind. Da offenbar  $V = V(+) \oplus V(-)$  gilt, folgt  $V = V(+) \perp V(-)$ .

Sei nun  $U \supset V(+)$ ,  $U \neq V(+)$  ein Unterraum. Ein Vektor  $u \in U \setminus V(+)$  hat eine Darstellung  $u = u_+ + u_-$ ,  $u_+ \in V(+)$ ,  $u_- \in V(-)$  und  $u_- \neq 0$ . Es ist dann  $u_- = u - u_+ \in U$  und folglich auch jeder Vektor  $u_+ + \lambda u_-$ ,  $\lambda \in \mathbb{R}$ , in  $U$  enthalten. Aus  $s(u, u) = s(u_+, u_+) + \lambda^2 s(u_-, u_-)$  und  $s(u_-, u_-) < 0$  folgt  $s(u, u) < 0$  für genügend groe  $\lambda$ . Also ist  $U$  nicht positiv und damit  $V(+)$  maximal.  $\square$

**(4.7.2) Satz.** Sei  $V^+$  ein maximaler positiver Raum, und sei  $V^- = (V^+)^\perp$ . Dann ist  $V^-$  ein negativer Raum und  $V$  die orthogonale Summe von  $V^+$  und  $V^-$ .

*Beweis.* Wir haben  $V = V^+ + V^-$  und  $V^+ \cap V^- = 0$  zu zeigen.

Sei  $v \in V^+ \cap V^-$ . Dann ist nach Definition von  $V^-$  als  $(V^+)^\perp$  auch  $s(v, v) = 0$ . Da  $s$  auf  $V^+$  positiv definit ist, folgt  $v = 0$ . Wegen  $\dim V^+ + \dim (V^+)^\perp = \dim V$  folgt  $V = V^+ + V^-$ .

Sei  $z \neq 0$  aus  $V^-$ . Wäre  $s(z, z) > 0$ , so wäre für  $\lambda \in \mathbb{R}$  und  $v \in V^+$  auch

$$s(v + \lambda z, v + \lambda z) = s(v, v) + \lambda^2 s(z, z) > 0$$

und deshalb  $s$  auf dem von  $V^+$  und  $z$  erzeugten Raum positiv definit, im Widerspruch zur Wahl von  $V^+$ . Also gilt  $s(z, z) \leq 0$  für alle  $z \in V^-$ . Ist  $s(z, z) = 0$  und  $z, w \in V^-$ , so folgt, weil  $w + \lambda z \in V^-$  ist,

$$0 \geq s(w + \lambda z, w + \lambda z) = s(w, w) + 2\lambda s(w, z).$$

Das kann für alle  $\lambda \in \mathbb{R}$  nur gelten, wenn  $s(w, z) = 0$  ist. Da für alle  $u \in V^+$  ebenfalls  $s(u, z) = 0$  ist, folgt aus der Regularität von  $s$ , da  $z = 0$  ist. Mithin ist  $s$  auf  $V^-$  negativ definit.  $\square$

Wir wollen nun Eindeutigkeitsaussagen über die Zerlegung von  $V$  in positive und negative Unterräume herleiten.

**(4.7.3) Satz.** *Sei  $s$  eine reguläre Form. Dann gilt:*

- (1) *Je zwei maximale positive Unterräume  $V_1$  und  $V_2$  von  $V$  haben dieselbe Dimension. (Analog für maximale negative.)*
- (2) *Sei  $V = V_+ \oplus V_-$ , und sei  $V_+$  bzw.  $V_-$  positiv bzw. negativ. Dann sind  $V_+$  und  $V_-$  maximal.*
- (3) *Sei  $V_+$  bzw.  $V_-$  ein maximaler positiver bzw. negativer Unterraum. Dann gilt  $V = V_+ \oplus V_-$ .*

*Beweis.* (1) Ist  $U$  positiv und  $W$  negativ, so gilt  $U \cap V = \{0\}$ , denn  $x \in U \cap V$  impliziert  $0 \geq s(x, x) \geq 0$ , und wegen der positiven Definitheit von  $s$  auf  $U$  ist deshalb  $x = 0$ .

Sei  $\dim V_1 \leq \dim V_2$ . Es ist  $V_1^\perp$  nach (4.7.2) negativ. Also gilt  $V_2 \cap V_1^\perp = \{0\}$  und deshalb  $\dim V = \dim V_1 + \dim V_1^\perp \leq \dim V_2 + \dim V_1^\perp = \dim(V_2 + V_1^\perp) + \dim(V_2 \cap V_1^\perp) = \dim(V_2 + V_1^\perp) \leq \dim V$ . Somit steht überall die Gleichheit und wir erhalten  $\dim V_1 = \dim V_2$ . (In der Un-Gleichungskette wurde (4.7.2) und die Dimensionsformel benutzt.)

(2) Sei  $U \supset V_+$ ,  $U \neq V_+$  ein Unterraum. Sei  $u \in U \setminus V_+$ ,  $u = u_+ + u_-$ ,  $u_\pm \in V_\pm$ . Wie im Beweis von (4.7.1) sieht man, daß  $U$  Vektoren der Form  $u_+ + \lambda u_-$ ,  $s(u_-, u_-) < 0$ ,  $\lambda \in \mathbb{R}$  beliebig, enthält. Es kann aber

$$s(u_+ + \lambda u_-, u_+ + \lambda u_-) = s(u_+, u_+) + 2\lambda s(u_+, u_-) + \lambda^2 s(u_-, u_-)$$

nicht für alle  $\lambda$  positiv sein, also ist  $U$  nicht positiv und demnach  $V_+$  maximal.

(3) Wegen (4.7.2) gilt  $V = V_+ \oplus V_+^\perp$ . Wegen (1) und (2) gilt  $\dim V_+^\perp = \dim V_-$ . Wegen  $V_+ \cap V_- = \{0\}$  ist also  $\dim(V_+ + V_-) = \dim V_+ + \dim V_- = \dim V_+ + \dim V_+^\perp = \dim V$  und deshalb  $V = V_+ \oplus V_-$ .  $\square$

Wegen (4.7.1) und (4.7.3) ist in einer Orthonormalbasis  $B$  die Anzahl der Basisvektoren  $b \in B$  mit  $s(b, b) = 1$  unabhängig von der Wahl der Orthonormalbasis (Trägheitssatz von Sylvester<sup>10</sup>).

Ist  $s: V \times V \rightarrow \mathbb{R}$  symmetrisch, aber vielleicht nicht regulär, so betrachten wir das Radikal  $V_0$  von  $s$ . Sei  $V = V_0 \oplus W$  mit geeignetem  $W$ . Auf  $W$  ist  $s$  regulär: Sei nämlich  $w \in W$  und  $s(w, v) = 0$  für alle  $v \in W$ . Dann gilt auch  $s(w, v) = 0$  für alle  $v \in V$ , also  $w \in V_0$ . Wegen  $V_0 \cap W = \{0\}$  folgt  $w = 0$ . Aus (4.7.2) erhält man eine Zerlegung  $V = V_0 \oplus V^+ \oplus V^-$ . Wir nennen das Tripel  $(\dim V^+, \dim V^-, \dim V_0)$  den *Index* von  $s$  und  $\dim V^+ - \dim V^-$  die *Signatur* von  $s$ . (Die Terminologie in diesem Kontext ist nicht einheitlich.)

Eine reguläre Form wird bezüglich einer geeigneten Orthonormalbasis durch eine Matrix der Gestalt

$$I_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \quad n = p + q$$

<sup>10</sup> James Joseph Sylvester 1814 – 1897

beschrieben. Es ist  $(p, q)$  der Index und  $p - q$  die Signatur. Die Gruppe

$$O(p, q) = \{A \in GL(n, \mathbb{R}) \mid A^t I_{p,q} A = I_{p,q}\}$$

heißt *orthogonale Gruppe* dieser Form.

Für  $q = 0$  ist  $O(n, 0) = O(n)$ . Ist  $p = 3$ ,  $q = 1$ , so heißt  $(V, s)$  *Minkowski-Raum*<sup>11</sup> und  $O(3, 1)$  die *Lorentz-Gruppe*<sup>12</sup>. Sie regiert die spezielle Relativitätstheorie von Einstein<sup>13</sup>.

Eine symmetrische reelle  $(n, n)$ -Matrix  $P$  heißt, wie erinnerlich *positiv definit*, wenn  $(x, y) \mapsto x^t P y$  ein Skalarprodukt auf dem  $\mathbb{R}^n$  ist. Jedes Skalarprodukt auf dem  $\mathbb{R}^n$  ist von dieser Gestalt,  $P$  ist nämlich die Matrix von  $s$  bezüglich der Standardbasis.

Wir werden nun ein Verfahren angeben, mit dem man rechnerisch entscheiden kann, ob eine Matrix positiv definit ist. Sei  $A = (a_{ij})$  eine reelle symmetrische  $(n, n)$ -Matrix, und sei  $A(k)$  die Untermatrix  $(a_{ij} \mid 1 \leq i, j \leq k)$ . Dann gilt:

**(4.7.4) Satz.** *Eine reelle symmetrische  $(n, n)$ -Matrix  $A$  ist genau dann positiv definit, wenn alle Determinanten  $\det(A(k))$  positiv sind.*

*Beweis.* Ist  $A$  positiv definit, so gibt es  $T$  mit  $T^t A T = I_n$ , woraus wir sehen, daß  $\det A$  positiv ist. Die Matrix  $A(k)$  beschreibt ein auf  $\mathbb{R}^k$  eingeschränktes Skalarprodukt und hat aus demselben Grund eine positive Determinante.

Seien umgekehrt die Determinantenbedingungen erfüllt. Wir betrachten die durch  $A$  auf  $\mathbb{R}^n$  gegebene Bilinearform. Durch Induktion nach der Dimension erkennen wir, daß auf  $\mathbb{R}^{n-1} \times \{0\}$  eine positive Form induziert wird. Also ist der negative Raum null- oder eindimensional. Wäre er eindimensional, so wäre  $\det A < 0$ .  $\square$

## 4.8 Die Lorentz-Gruppe

Wir betrachten ausführlich die symmetrische Bilinearform

$$s: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, \quad (x, y) \mapsto x^t H y,$$

worin

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

<sup>11</sup>Hermann Minkowski 1864 – 1909

<sup>12</sup>Hendrik Antoon Lorentz 1853 – 1928

<sup>13</sup>Albert Einstein 1879 – 1955

ist. Wir benutzen die theoretischen Einsichten der früheren Abschnitte. Die Form heißt *hyperbolische Form*, weil die Menge

$$H(d) = \{x \in \mathbb{R}^2 \mid s(x, x) = d\} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 - x_2^2 = d\}$$

für  $d \neq 0$  eine Hyperbel ist. Die Menge  $H(0)$  besteht aus den beiden Winkelhalbierenden der Koordinatenachsen und ist die Menge der isotropen Vektoren. Das Komplement  $\mathbb{R}^2 \setminus H(0)$  zerfällt in vier Stücke (Kegel), von denen das rechte und linke die positiven und das obere und untere die negativen Vektoren enthält. Die Kegel werden formelmäßig durch

$$P(+) = \{(x, y) \mid x > |y| > 0\}, \quad P(-) = \{(x, y) \mid -x > |y| > 0\}$$

$$N(+) = \{(x, y) \mid y > |x| > 0\}, \quad N(-) = \{(x, y) \mid -y > |x| > 0\}$$

beschrieben ( $P$  für positiv,  $N$  für negativ).

Wählen wir als  $V^+$  den von  $(1, t)$  für  $-1 < t < 1$  aufgespannten Raum, so ist  $V^-$  der von  $(t, 1)$  aufgespannte. Also liegen  $V^+$  und  $V^-$  spiegelbildlich zu den isotropen Geraden. Wir erkennen, wie sich bei Annäherung von  $V^+$  und  $V^-$  notwendig isotrope Vektoren ergeben.

Die Isometrien von  $s$  werden durch die reellen  $(2, 2)$ -Matrizen  $B$  gegeben, die der Gleichung  $B^t H B = H$  genügen. Sie bilden die Gruppe  $O(1, 1)$ , die wir die *kleine Lorentz-Gruppe* nennen wollen. Setzen wir

$$B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix},$$

so bedeutet  $B^t H B = H$  das Bestehen der Gleichungen

$$b_1^2 - b_3^2 = 1, \quad b_2^2 - b_4^2 = -1, \quad b_1 b_2 - b_3 b_4 = 0.$$

Quadrieren der letzten Gleichung und Einsetzen der beiden ersten führt zu  $b_1^2 = b_4^2$ . Außerdem sehen wir, daß diese Quadrate größergleich 1 sind. Falls  $b_1 = b_4 = b$  ist, erhalten wir die Matrizen

$$\begin{pmatrix} b & \pm\sqrt{b^2 - 1} \\ \pm\sqrt{b^2 - 1} & b \end{pmatrix},$$

der Determinante 1 und, falls  $b_1 = -b_4 = b$  ist, die Matrizen

$$\begin{pmatrix} b & \pm\sqrt{b^2 - 1} \\ \mp\sqrt{b^2 - 1} & -b \end{pmatrix}.$$

der Determinante  $-1$ . Wir betrachten den ersten Fall. Es muß  $b^2 \geq 1$  sein, damit eine reelle Quadratwurzel existiert. Im Fall  $b \geq 1$  setzen wir

$$b = \frac{1}{\sqrt{1 - \beta^2}}, \quad -1 < \beta < 1$$

und erhalten die Matrix

$$B(\beta) = \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{\beta}{\sqrt{1-\beta^2}} \\ \frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{pmatrix}.$$

Im physikalischen Kontext betrachtet man die Form  $x_1^2 - c^2 x_2^2$  ( $c$  die Lichtgeschwindigkeit). Die Matrizen  $B(\beta)$  sind dann durch

$$A(\beta) = \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & \frac{\beta c}{\sqrt{1-\beta^2}} \\ \frac{\beta/c}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} \end{pmatrix}$$

zu ersetzen. Verwenden wir die Matrix  $A(\beta)$  als Koordinatentransformation von  $(z, t)$ -Koordinaten, so ergeben sich die neuen  $(z', t')$ -Koordinaten als

$$z' = \frac{z + \beta ct}{\sqrt{1-\beta^2}}, \quad t' = \frac{t + (\beta/c)z}{\sqrt{1-\beta^2}}.$$

Das ist die berühmte *Lorentz-Transformation*. Die erste Gleichung wird physikalisch so interpretiert: Der Ursprung des  $(z', t')$ -Systems hat gegenüber dem Ursprung des  $(z, t)$ -Systems die Geschwindigkeit  $v = -\beta c$  (was  $z' = 0$  entspricht). Also ist  $\beta = -v/c$ . Damit erhalten wir typische Formeln der speziellen Relativitätstheorie.

Zurück zu den Matrizen  $B(\beta)$ . Die Formel

$$B(\beta)B(\gamma) = B\left(\frac{\beta + \gamma}{1 + \beta\gamma}\right)$$

wird durch Nachrechnen verifiziert.

**(4.8.1) Satz.** Die Matrizen  $B(\beta)$  mit  $-1 < \beta < 1$  bilden eine Untergruppe  $SO(1, 1)_+$  von  $O(1, 1)$ . Sie ist isomorph zur additiven Gruppe  $(\mathbb{R}, +)$ .

*Beweis.* Wir verwenden die Bijektion

$$\rho: \mathbb{R} \rightarrow ]-1, 1[, \quad x \mapsto \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

Sie genügt dem Additionstheorem

$$\rho(\alpha + \beta) = \frac{\rho(\alpha) + \rho(\beta)}{1 + \rho(\alpha)\rho(\beta)}.$$

Deshalb ist  $\alpha \mapsto B(\rho(\alpha))$  der gewünschte Isomorphismus.  $\square$

Die vorstehend benutzte Funktion  $\rho$  ist der *tangens hyperbolicus*. Er ist Quotient des *sinus hyperbolicus*  $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$  und des *cosinus hyperbolicus*  $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$ . Für diese Funktionen gelten die *Additionstheoreme*

$$\begin{aligned}\cosh(x+y) &= \cosh(x)\cosh(y) + \sinh(x)\sinh(y), \\ \sinh(x+y) &= \sinh(x)\cosh(y) + \cosh(x)\sinh(y)\end{aligned}$$

sowie die Identität  $\cosh^2(x) - \sinh^2(x) = 1$ . Die Matrizen

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

bilden eine kommutative Untergruppe  $G$  von  $O(1,1)$  der Ordnung 4; jedes Element darin hat höchstens die Ordnung zwei (*Kleinsche Vierergruppe*<sup>14</sup>).

**(4.8.2) Notiz.** *Es gibt einen surjektiven Homomorphismus  $k: O(1,1) \rightarrow G$  mit dem Kern  $SO(1,1)_+$ . Er bildet ein Element  $xB(\beta)$ ,  $x \in G$ , auf  $x$  ab. Eine andere Version dieses Homomorphismus ist*

$$O(1,1) \rightarrow \mathbb{Z}^* \times \mathbb{Z}^*, \quad \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \mapsto \left( \frac{b_1}{|b_1|}, \frac{b_4}{|b_4|} \right).$$

□

Zum letzteren bemerke man die Äquivalenzen:

$b_1 > 0 \Leftrightarrow$  die positiven Kegel werden nicht vertauscht

$b_4 > 0 \Leftrightarrow$  die negativen Kegel werden nicht vertauscht.

Ferner ist  $b_1 b_4 > 0$  äquivalent zu  $\det = 1$ .

Die Gruppe  $O(1,1)$  besteht aus den vier Stücken (Zusammenhangskomponenten im Sinne der Topologie)  $xSO(1,1)_+$ ,  $x \in G$ . Welche Bedeutung haben diese Stücke? Zunächst kann man nach der Determinante unterscheiden; die Untergruppe  $SO(1,1)$  der Elemente mit Determinante 1 besteht aus zwei Stücken. Ferner besteht der Kegel der positiven Vektoren selbst aus zwei Stücken: dem Teil mit  $x_1 > 0$  und dem mit  $x_1 < 0$ . Der positive Kegel wird durch Elemente von  $SO(1,1)$  in sich abgebildet. Die Elemente von  $SO(1,1)_+$  bilden alle vier Stücke von  $\mathbb{R}^2 \setminus H(0)$  in sich ab. Man kann also die vier Stücke der Gruppe durch ihre Vertauschungseigenschaften bezüglich der vier Stücke von  $\mathbb{R}^2 \setminus H(0)$  charakterisieren.

Im weiteren vergleichen wir die Gruppen  $O(2)$  und  $O(1,1)$ . Die Matrizen aus  $SO(1,1)_+$  lassen sich auch in der Form

$$\begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix} = L(t)$$

---

<sup>14</sup>Felix Klein 1849 – 1925

schreiben. Die Additionstheoreme von  $\sinh$  und  $\cosh$  zeigen die Gleichung  $L(t)L(u) = L(t+u)$ . Folglich ist

$$\mathbb{R} \rightarrow SO(1,1)_+, \quad t \mapsto L(t)$$

ein Isomorphismus, der überdies die Analogie zu den Drehmatrizen zeigt.

Die Funktionen  $\sin, \cos, \sinh, \cosh$  sind auch für komplexe Argumente erklärt, und es gelten die Eulerschen Formeln

$$\cosh(iz) = \cos(z), \quad \sinh(iz) = i \sin(z)$$

für alle  $z \in \mathbb{C}$ .

Die Gruppen

$$\begin{aligned} O(2; \mathbb{C}) &= \{A \in GL(2, \mathbb{C}) \mid A^t A = I\} \\ O(1, 1; \mathbb{C}) &= \{A \in GL(2, \mathbb{C}) \mid A^t H A = H\} \end{aligned}$$

sind isomorph, weil die zugehörigen Bilinearformen über den komplexen Zahlen isomorph sind. Explizit wird ein Isomorphismus folgendermaßen gegeben:

$$O(2; \mathbb{C}) \rightarrow O(1, 1; \mathbb{C}), \quad A \mapsto \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} A \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix}.$$

Bei diesem Isomorphismus wird die Drehmatrix  $D(iz)$  auf  $L(-z)$  abgebildet ( $z \in \mathbb{C}$  beliebig). Wir erkennen, daß die Lorentz-Matrizen  $L$  komplexe Formen der Drehmatrizen  $D$  sind: "Drehungen" um einen rein imaginären Winkel.

Wir beschreiben nun genauer die Struktur der Gruppe  $O(2; \mathbb{C})$ . Wir haben die Determinantenabbildung nach  $\mathbb{Z}^*$ . Der Kern ist  $SO(2; \mathbb{C})$ . Wir setzen für  $z \in \mathbb{C}^*$

$$M(u) = \begin{pmatrix} \frac{1}{2}(u + u^{-1}) & \frac{1}{2i}(u - u^{-1}) \\ -\frac{1}{2i}(u - u^{-1}) & \frac{1}{2}(u + u^{-1}) \end{pmatrix}.$$

Damit gilt:

**(4.8.3) Satz.**  $\mathbb{C}^* \rightarrow SO(2; \mathbb{C}), \quad u \mapsto M(u)$  ist ein Isomorphismus.

*Beweis.* Durch Nachrechnen bestätigt man, daß ein injektiver Homomorphismus vorliegt. (Übrigens, setzt man  $u = \exp(z)$ , so ist  $M(u) = D(z)$ , und man ist wieder bei den Additionstheoremen von  $\cos$  und  $\sin$  angelangt). Für die Surjektivität sieht man zunächst, daß man alle Matrizen der Form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad a^2 + b^2 = 1$$

bestimmen muß. Sodann schreibt man

$$1 = a^2 + b^2 = (a + bi)(a - bi) = uv, \quad u = a + bi, \quad v = a - bi$$

und löst nach  $a$  und  $b$  auf.  $\square$

Wir betrachten schließlich die Gruppen  $SO(2)$  und  $SO(1,1)_+$  vom dynamischen Standpunkt, indem wir die Bahnen bei der natürlichen Operation auf  $\mathbb{R}^2$  beschreiben (die Operation durch Matrizenmultiplikation). Die Bahnen von  $SO(2)$  sind leicht beschrieben: Die konzentrischen Kreise um den Nullpunkt; der Nullpunkt selbst ist auch eine Bahn. Nun zur anderen Gruppe. Die Gleichung

$$\begin{pmatrix} \cosh(t) & \sinh(t) \\ \sinh(t) & \cosh(t) \end{pmatrix} \begin{pmatrix} a \\ \pm a \end{pmatrix} = e^{\pm t} \begin{pmatrix} a \\ \pm a \end{pmatrix}$$

zeigt, daß die Menge der isotropen von Null verschiedenen Vektoren in vier Bahnen zerfällt: Die vier Halbgeraden der Menge  $H(0) \setminus 0$ . Die Bahnen durch nicht-isotrope Vektoren sind Hyperbeläste in den entsprechenden Kegeln, und zwar besteht jede Hyperbel  $H(d)$  aus vier Bahnen (= Ästen). Instruktiv ist es, die Bewegung eines Bahnpunktes bei der Wirkung von  $L(t)$  zu verfolgen, wenn  $t$  von  $-\infty$  nach  $\infty$  läuft: Die Punkte kommen aus dem Unendlichen in der Nähe der Asymptote  $y = -x$  angerast und verschwinden im Unendlichen in der Nähe der Asymptote  $y = x$ .

## Die allgemeine Lorentz-Gruppe

Wir betrachten jetzt die allgemeine Lorentz-Gruppe. Sei  $V$  ein  $n$ -dimensionaler reeller Vektorraum und  $s: V \times V \rightarrow \mathbb{R}$  eine reguläre symmetrische Bilinearform vom Index  $(n-1, 1)$ . Sei  $q: V \rightarrow \mathbb{R}$ ,  $x \mapsto s(x, x)$  die zu  $s$  gehörende quadratische Form. Das Paar  $(V, S)$  heißt  $n$ -dimensionaler *Minkowski-Raum*. Für  $n = 4$  erhalten wir den Raum der speziellen Relativitätstheorie. Für einen Vektor  $x \in V$  führen wir die folgenden Bezeichnungen ein:

$$\begin{aligned} x \text{ raumartig} &\Leftrightarrow q(x) > 0 \\ x \text{ lichtartig} &\Leftrightarrow q(x) = 0 \\ x \text{ zeitartig} &\Leftrightarrow q(x) < 0. \end{aligned}$$

Diese Terminologie ist natürlich der Physik entlehnt. Wir haben früher diese Fälle positiv, isotrop und negativ genannt. Ist  $x$  zeitartig, so ist das orthogonale Komplement von  $x$  ein  $(n-1)$ -dimensionaler raumartiger Unterraum.

**(4.8.4) Notiz.** Die Relation  $x \sim y \Leftrightarrow s(x, y) < 0$  ist eine Äquivalenzrelation auf der Menge der zeitartigen Vektoren. Es gibt zwei Äquivalenzklassen.

*Beweis.* Reflexivität und Symmetrie sind klar. Sei also  $s(x, y) < 0$ ,  $s(y, z) < 0$ . Wir haben  $s(x, z) < 0$  zu zeigen. Ohne wesentliche Einschränkung sei  $s(y, y) = -1$ . Sei  $N$  das orthogonale Komplement von  $y$ . Dann gibt es Darstellungen der Form

$$x = \alpha y + u, \quad y = \beta y + v$$

mit  $u, v \in N$  und

$$\alpha = -s(x, y), \quad \beta = -s(z, y).$$

Wir rechnen

$$\begin{aligned} s(x, x) &= -\alpha^2 + s(u, u) < 0 \\ s(z, z) &= -\beta^2 + s(v, v) < 0 \\ s(x, z) &= -\alpha\beta + s(u, v). \end{aligned}$$

Auf  $N$  ist  $s$  positiv definit, und wir haben deshalb die Cauchy-Schwarzsche Ungleichung

$$s(u, v)^2 \leq s(u, u)s(v, v) < \alpha^2\beta^2.$$

Nach Voraussetzung sind  $\alpha$  und  $\beta$  positiv. Also folgt  $s(u, v) < \alpha\beta$  und damit  $s(x, z) < 0$ .

Ist  $T(x)$  die Klasse von  $x$ , so gilt für jeden zeitartigen Vektor  $y$  entweder  $s(x, y) < 0$  oder  $s(-x, y) < 0$ , denn  $s(x, y) = 0$  würde  $y \in N$  bedeuten. Also gibt es nur die Klassen  $T(x)$  und  $T(-x)$ .  $\square$

**(4.8.5) Notiz.** Die Äquivalenzklassen der zeitartigen Vektoren sind konvex.

*Beweis.* Seien  $x, y, z$  in derselben Klasse, und sei  $t \in [0, 1]$ . Dann gilt

$$s(x, ty + (1-t)z) = ts(x, y) + (1-t)s(x, z) < 0.$$

Also liegen alle  $ty + (1-t)z$  in der Klasse von  $x$ .  $\square$

Wir bezeichnen die Klassen mit  $T_+$  und  $T_-$  und nennen  $T_+$  ( $T_-$ ) den *Zukunftskegel* (*Vergangenheitskegel*).

Die *Lorentz-Gruppe*  $O(V, s)$  von  $(V, s)$  ist die Gruppe der Isometrien von  $s$ , das heißt sie besteht aus den Endomorphismen  $f$  von  $V$ , genannt *Lorentz-Transformationen*, mit der Eigenschaft:  $s(x, y) = s(fx, fy)$  für alle  $x, y \in V$ . Die Abbildungen  $f$  aus  $O(V, s)$  bilden also raumartige (lichtartige, zeitartige) Vektoren wieder in solche ab. Sind ferner  $x, y$  äquivalente zeitartige Vektoren, so auch  $fx, fy$ . Die  $f \in O(V, s)$ , die  $T_+$  in sich abbilden, formen eine Untergruppe  $O(V, s)_+$  vom Index 2 in  $O(V, s)$ , also einen Normalteiler; seine Elemente heißen *orthochrone Lorentz-Transformationen*.

Es gibt eine Orthonormalbasis (Minkowski-Basis)  $v_1, \dots, v_n \in V$  bezüglich  $s$ , das heißt eine Basis mit der Eigenschaft  $s(v_j, v_j) = 1$  für  $j \leq n-1$  und  $s(v_n, v_n) = -1$ . Genau dann ist  $A = (a_{ij})$  die Matrix einer Lorentz-Transformation bezüglich dieser Basis, wenn  $AI_{n-1,1}A^t = I_{n-1,1}$  ist, wobei  $I_{n-1,1} = \text{Dia}(1, \dots, 1, -1)$  ist. Eine Lorentz-Transformation hat die Determinante  $\pm 1$ . Wir setzen:

$$\begin{aligned} O(n-1, 1) &= \{A \in GL(n, \mathbb{R}) \mid AI_{n-1,1}A^t = I_{n-1,1}\} \\ SO(n-1, 1) &= \{A \in O(n-1, 1) \mid \det A = 1\} \\ O(n-1, 1)_+ &= \{A \in O(n-1, 1) \mid A \text{ orthochron}\} \\ SO(n-1, 1)_+ &= SO(n-1, 1) \cap O(n-1, 1)_+. \end{aligned}$$

Wir nennen  $O(n-1, 1)$  *Lorentz-Gruppe*. Als Standardmodell des Minkowski-Raumes verwenden wir den  $\mathbb{R}^n$  mit der quadratischen Form  $q(x_1, \dots, x_n) = x_1^2 + \dots + x_{n-1}^2 - x_n^2$ . Die Standardbasis ist dann eine Minkowski-Basis.

**(4.8.6) Notiz.** Eine Matrix  $(a_{ij}) \in O(n-1, 1)$  ist genau dann orthochron, wenn  $a_{n,n} < 0$  ist. Die Zuordnung  $(a_{ij}) \mapsto a_{n,n}/|a_{n,n}|$  ist ein Homomorphismus  $\varepsilon: O(n-1, 1) \rightarrow \mathbb{Z}^*$  mit dem Kern  $O(n-1, 1)_+$ .

*Beweis.* Der Vektor  $(0, \dots, 0, 1) = e_n$  ist zeitartig. Genau dann ist  $(a_{ij})$  orthochron, wenn  $Ae_n$  zu  $e_n$  äquivalent ist. Es ist  $s(e_n, Ae_n) = -a_{n,n}$ . Da jedenfalls  $Ae_n$  wieder zeitartig ist, so ist  $a_{n,n} \neq 0$ .  $\square$

Im Standardmodell setzen wir  $T_\pm = \{x \in \mathbb{R}^n \mid q(x) < 0, \pm x_n > 0\}$ . Die Menge  $Q = \{x \in \mathbb{R}^n \mid q(x) = -1\}$  besteht aus den beiden Stücken  $Q_\pm = \{x \in Q \mid \pm x_n \geq 1\}$ . Die Lorentz-Transformationen bilden  $Q$  in sich ab.

**(4.8.7) Satz.**  $SO(n-1, 1)_+$  operiert transitiv auf  $Q_+$ . Die Standgruppe von  $e_n$  besteht aus den Matrizen  $\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$ ,  $A \in SO(n-1)$ .

*Beweis.* Seien  $y_1, y_2$  Vektoren aus  $Q_+$  und  $N_1, N_2$  ihre orthogonalen Komplemente. Wir haben eine orthogonale Zerlegung bezüglich  $s$

$$\mathbb{R}y_1 \perp N_1 = \mathbb{R}^n = \mathbb{R}y_2 \perp N_2.$$

Auf  $N_1$  und  $N_2$  ist  $s$  positiv definit. Es gilt deshalb eine Isometrie  $\varphi: (N_1, s) \rightarrow (N_2, s)$ . Diese Isometrie ergänzen wir durch  $y_1 \mapsto y_2$  zu einer Isometrie  $f$  des Minkowski-Raumes. Da  $y_1$  und  $y_2$  in  $Q_+ \subset T_+$  liegen, ist  $f$  orthochron. Hat  $f$  nicht Determinante 1, so ändern wir  $\varphi$  um eine Isometrie von  $(N_2, s)$  mit der Determinante  $-1$  ab.  $\square$

Der homogene Raum  $SO(n-1, 1)_+/SO(n-1)$  ist also zu  $Q_+$  homöomorph und  $Q_+$  ist vermöge der Projektion  $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_{n-1})$  zu  $\mathbb{R}^{n-1}$  homöomorph.

Wir wollen jetzt noch genauere Information über die klassischen Lorentz-Gruppen  $O(3, 1)$  und  $O(2, 1)$  gewinnen. Dazu benutzen wir ein anderes Modell des Minkowski-Raumes.

Sei  $H(2, \mathbb{C})$  die Menge der komplexen, hermiteschen  $(2, 2)$ -Matrizen. Eine Matrix aus  $H(2, \mathbb{C})$  hat die Form

$$A = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix}, \quad a, c \in \mathbb{R}, \quad b \in \mathbb{C}.$$

Die Menge  $H(2, \mathbb{C})$  ist ein vierdimensionaler reeller Vektorraum. Eine Basis ist

$$\tau_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Die Matrizen  $\tau_1, \tau_2, \tau_3$  heißen *Pauli-Matrizen*. Die Abbildung

$$q = -\det: H(2, \mathbb{C}) \rightarrow \mathbb{R}, \quad \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix} = |b|^2 - ac$$

ist eine quadratische Form auf  $H(2, \mathbb{C})$ . Die zugehörige symmetrische Bilinearform ist

$$\langle \sigma, \tau \rangle = \frac{1}{2}(Sp(\sigma\tau) - Sp(\sigma)Sp(\tau)),$$

weil für jede  $(2,2)$ -Matrix  $\sigma$  die Relation

$$Sp(\sigma^2) - (Sp\sigma)^2 = -2 \det \sigma$$

besteht. Die Form  $q$  nimmt auf den Basisvektoren  $\tau_1, \dots, \tau_4$  die Werte  $1, 1, 1, -1$  an. Also ist  $(H(2, \mathbb{C}), q)$  ein vierdimensionaler Minkowski-Raum.

Sei  $A \in GL(2, \mathbb{C})$  und  $X \in H(2, \mathbb{C})$ . Dann zeigt eine leichte Rechnung, daß auch  $AX\bar{A}^t \in H(2, \mathbb{C})$  ist. Wir erhalten somit eine lineare Abbildung

$$\Lambda_A: H(2, \mathbb{C}) \rightarrow H(2, \mathbb{C}), \quad X \mapsto AX\bar{A}^t$$

und es gilt

$$\Lambda_A \Lambda_B = \Lambda_{AB}.$$

Ist  $A \in SL(2, \mathbb{C})$ , so zeigt die Produktformel für Determinanten die Relation  $q(\Lambda_A X) = q(X)$ . Folglich ist dann  $\Lambda_A$  eine Lorentz-Transformation unseres Modells.

**(4.8.8) Satz.** Die Zuordnung  $A \mapsto \Lambda_A$  ist ein surjektiver Homomorphismus  $\Lambda: SL(2, \mathbb{C}) \rightarrow SO(3, 1)_+$  mit dem Kern  $\{\pm I\}$ .

*Beweis.* Wir bestimmen den Kern. Gilt  $AX\bar{A}^t = X$  für alle  $X \in H(2, \mathbb{C})$ , so folgt zunächst für  $X = I$  die Relation  $A\bar{A}^t = I$ . Also gilt  $AX = XA$  für alle  $X \in H(2, \mathbb{C})$  und dann auch  $A(iX) = (iX)A$ . Ist  $X$  hermitesch, so ist  $iX$  schieferhermitesch und umgekehrt. Jede Matrix ist Summe einer hermiteschen und einer schieferhermiteschen. Also ist  $A$  mit allen  $(2, 2)$ -Matrizen vertauschbar und deshalb ein Vielfaches der Einheitsmatrix mit der Determinante 1.

Analoge Überlegungen lassen für den 3-dimensionalen Raum reellen symmetrischen  $(2, 2)$ -Matrizen und die Gruppe  $SL(2, \mathbb{R})$  durchführen. Dieselbe Form wie oben macht daraus einen dreidimensionalen Minkowski-Raum. Jedes  $A \in SL(2, \mathbb{R})$  liefert durch  $X \mapsto AXA^t$  eine Isometrie dieses Raumes. Wir erhalten einen surjektiven Homomorphismus  $\lambda: SL(2, \mathbb{R}) \rightarrow SO(2, 1)_+$  mit dem Kern  $\pm I$ .  $\square$

# Kapitel 5

## Affine und projektive Räume

### 5.1 Affine Räume

Wir beginnen damit, zwei verschiedene Definitionen des Begriffs „affiner Raum“ zu geben.

**(5.1.1) Affiner Raum.** Ein *affiner Raum*  $A = (TX, \rho, X)$  über dem Körper  $K$  besteht aus einer Menge  $X$ , einem  $K$ -Vektorraum  $TX$  und einer Operation

$$\rho: TX \times X \rightarrow X, \quad (v, x) \mapsto \rho(v, x) = v + x$$

der additiven Gruppe von  $TX$  auf der Menge  $X$ , so daß folgendes gilt:  $0 + x = x$ ,  $(v_1 + v_2) + x = v_1 + (v_2 + x)$ . Die Operation ist *einfach transitiv*; das soll heißen: Zu jedem geordneten Paar  $x_1, x_2$  von Elementen aus  $X$  gibt es genau ein  $v \in TX$ , so daß  $\rho(v, x_1) = x_2$  ist. Die Elemente von  $X$  heißen die *Punkte* des affinen Raumes. Die Dimension von  $TX$  wird als die Dimension des affinen Raumes bezeichnet.  $\diamond$

**(5.1.2) Der Standardraum.** Sowohl  $TX$  als auch  $X$  sind ein Vektorraum  $V$ , und  $\rho$  ist die Vektorraumaddition. Das Beispiel ist insofern verwirrend, als die Elemente der Menge  $V$  hier zweierlei Funktion haben: Zum einen „Vektoren“ und zum anderen „Punkte“ zu sein.  $\diamond$

Die Abbildung  $t_v: X \rightarrow X$ ,  $x \mapsto v + x$  heißt *Translation* mit  $v$ . Die Abbildung  $a_x: TX \rightarrow X$ ,  $v \mapsto v + x$  heißt *Abtragen der Vektoren* von  $x$ .

Die Abbildungen  $a_x$  sind bijektiv, weil die Operation einfach transitiv ist. Sind  $x_1, x_2 \in X$  gegeben, so heißt der durch  $v + x_1 = x_2$  eindeutig bestimmte Vektor  $v$  der *Verbindungsvektor*  $v = \overrightarrow{x_1 x_2}$  von  $x_1$  nach  $x_2$ . Wir bezeichnen ihn

auch mit  $v = x_2 - x_1$ ; das Minuszeichen ist hierbei nur symbolisch zu verstehen, denn auf  $X$  ist keine Verknüpfung erklärt.

In einem affinen Raum über  $\mathbb{R}$  heißt die Menge  $\{tv + x_1 \mid t \in [0, 1]\}$ ,  $v = \overrightarrow{x_1 x_2}$  die (Verbindungs-) *Strecke* von  $x_1$  nach  $x_2$ . Eine Teilmenge  $K \subset X$  heißt *konvex*, wenn sie mit je zwei Punkten auch deren Verbindungsstrecke enthält. Der Durchschnitt konvexer Mengen ist konvex. Jede Teilmenge  $S \subset X$  liegt in einer kleinsten konvexen Teilmenge, der *konvexen Hülle*  $C(S)$  von  $S$ ; das ist der Durchschnitt aller  $S$  enthaltenden konvexen Teilmengen von  $X$ ; solche gibt es, weil  $X$  konvex ist.

Wir üblich sagen wir auch:  $X$  ist ein affiner Raum; der Vektorraum  $TX$  und die Operation  $\rho$  werden dann als gegeben unterstellt. Die Bezeichnung  $TX$  soll an „Tangentialraum von  $X$ “ oder an „Translationsraum von  $X$ “ erinnern.

Wir können die Eigenschaften des Verbindungsvektors zu einer zweiten Definition eines affinen Raumes benutzen.

**(5.1.3) Affiner Raum.** Ein *affiner Raum* ist ein Tripel  $(X, a, TX)$ , das aus einer Menge  $X$ , einem Vektorraum  $TX$  und einer Abbildung  $a: X \times X \rightarrow TX$  besteht, so daß gilt:

- (1) Für je drei Elemente  $x_1, x_2, x_3 \in X$  ist  $a(x_1, x_2) + a(x_2, x_3) = a(x_1, x_3)$ .
- (2) Für jedes  $x \in X$  ist die Abbildung  $X \mapsto TX, y \mapsto a(x, y)$  bijektiv.

Aus (1) folgt, daß  $a(x, x)$  für jedes  $x$  der Nullvektor ist. Ferner gilt  $a(x, y) = -a(y, x)$ . Wir setzen  $a(x, y) = \overrightarrow{xy}$ .  $\diamond$

Die Definitionen (5.1.1) und (5.1.3) hängen folgendermaßen zusammen. Ist  $(TX, \rho, X)$  wie in (5.1.1) gegeben, so ist  $a(x, y)$  durch  $\overrightarrow{xy}$  definiert. Ist  $(X, a, TX)$  wie in (5.1.3) gegeben, so wird  $\rho: TX \times X \rightarrow X$  durch  $\rho(v, x) = y$  definiert, wobei  $y$  durch  $a(x, y) = v$  eindeutig bestimmt ist.

**(5.1.4) Beispiel.** Sei  $U$  Unterraum des Vektorraums  $V$ , und sei  $b \in V$ . Die Nebenklasse  $U + b = b + U = \{b + u \mid u \in U\}$  liefert den affinen Raum  $(U, \rho, b + U)$  mit der Operation  $\rho(u_1, b + u_2) = u_1 + b + u_2$ .  $\diamond$

**(5.1.5) Beispiel.** Sei  $f: V \rightarrow W$  eine lineare Abbildung zwischen Vektorräumen  $V$  und  $W$ , und sei  $w \in f(V) \subset W$ . Dann ist  $(\text{Kern } f, \rho, f^{-1}(w))$  ein affiner Raum, wobei  $\rho$  wieder durch die Vektorraumaddition in  $V$  erklärt wird. Dieses Beispiel ist ein Spezialfall des vorigen.  $\diamond$

**(5.1.6) Satz.** Sei  $E$  ein endlich-dimensionaler Vektorraum und  $H \subset E$  ein Unterraum. Sei  $E_H = \{W \mid W \text{ Komplement von } H \text{ in } E\}$ . Die Menge  $E_H$  hat eine kanonische Struktur eines affinen Raumes zum Vektorraum  $\text{Hom}(E/H, H)$ .

*Beweis.* Sei  $\text{Pr}(E, H) = \{f: E \rightarrow H \mid f \text{ linear, } f|_H = \text{id}\}$  die Menge der Projektionen auf  $H$ . Wir erhalten eine Abbildung

$$\text{Pr}(E, H) \rightarrow E_H, \quad f \mapsto \text{Kern } f.$$

Diese Abbildung ist bijektiv, denn jedes  $W \in E_H$  liefert wegen  $E = W \oplus H$  eine Projektion  $E \rightarrow H$  mit Kern  $W$ . Wir erklären deshalb die Struktur des affinen Raumes auf  $Pr(E, H)$  statt auf  $E_H$ . Sind  $p_1, p_2 \in Pr(E, H)$ , so liegt  $H$  im Kern von  $p_1 - p_2$  und liefert deshalb eine lineare Abbildung  $a(p_1, p_2): E/H \rightarrow H \in \text{Hom}(E/H, H)$ . Es gilt offenbar (5.1.3) (1).

Ist  $\varphi \in \text{Hom}(E/H, H)$  und  $\tilde{\varphi}: E \rightarrow H$  die Zusammensetzung mit der kanonischen Abbildung  $E \rightarrow E/H$ , so ist für  $p \in Pr(E, H)$  auch  $p + \tilde{\varphi} \in Pr(E, H)$  und  $a(p + \tilde{\varphi}, p) = \varphi$ . Man erkennt (5.1.3) (2).  $\square$

**(5.1.7) Affine Abbildungen.** Seien  $(TX, \rho, X) = A$  und  $(TY, \sigma, Y) = B$  affine Räume zu Vektorräumen über demselben Körper. Eine *affine Abbildung*  $A \rightarrow B$  besteht aus einer Mengenabbildung  $f: X \rightarrow Y$  und einer linearen Abbildung  $T(f): TX \rightarrow TY$ , so daß folgendes Diagramm kommutativ ist:

$$\begin{array}{ccc} TX \times X & \xrightarrow{\rho} & X \\ \downarrow Tf \times f & & \downarrow f \\ TY \times Y & \xrightarrow{\sigma} & Y. \end{array}$$

Es ist  $Tf$  durch  $f$  bestimmt. Ist nämlich  $v + x = y$  für  $v \in TX$  und  $x, y \in X$ , so folgt

$$Tf(v) + f(x) = f(v + x) = f(y),$$

also ist  $Tf(v)$  der Verbindungsvektor von  $f(y)$  nach  $f(x)$ . Es ist also das Diagramm

$$\begin{array}{ccc} X \times X & \xrightarrow{a} & TX \\ \downarrow f \times f & & \downarrow Tf \\ Y \times Y & \xrightarrow{a} & TY \end{array}$$

(mit Daten im Sinne von (5.1.3)) kommutativ. Deshalb nennen wir in diesem Zusammenhang  $f$  *affine Abbildung*, wenn es eine lineare Abbildung  $Tf$  so gibt, daß das letzte Diagramm kommutativ wird.  $\diamond$

**(5.1.8) Satz.** *Seien  $X$  und  $Y$  affine Räume, und sei  $x_0 \in X$  gegeben. Eine affine Abbildung  $f: X \rightarrow Y$  ist durch den Punkt  $f(x_0) = y_0$  und die lineare Abbildung  $Tf: TX \rightarrow TY$  vermöge  $f(v + x_0) = Tf(v) + y_0$  bestimmt, und zu gegebenem  $y_0 \in Y$  und linearer Abbildung  $\alpha: TX \rightarrow TY$  wird durch  $f(v + x_0) = \alpha(v) + y_0$  eine affine Abbildung definiert.*

*Beweis.* Da  $TX \rightarrow X, v \mapsto v + x_0$  bijektiv ist, so liefert die Formel  $f(v + x_0) = \alpha(v) + y_0$  jedenfalls eine Abbildung  $f: X \rightarrow Y$ . Sei  $x_1 = v + x_0, u \in TX$ . Dann gilt

$$f(u + x_1) = f(u + v + x_0) = \alpha(u + v) + y_0$$

$$Tf(u) + f(x_1) = \alpha(u) + f(v + x_0) = \alpha(u) + \alpha(v) + y_0$$

und damit die gewünschte Kommutativität.  $\square$

Ein *affines Koordinatensystem*  $(x_0; v_1, \dots, v_n)$  von  $X$  besteht aus einem Punkt  $x_0 \in X$  (dem Ursprung) und einer Basis  $(v_1, \dots, v_n)$  von  $TX$ .

In affinen Räumen  $X$  kann man in geeigneter Weise mit Punkten rechnen. Das führt zu den *baryzentrischen Koordinaten* (Schwerpunkt-Koordinaten). Sei  $(TX, \rho, X)$  ein affiner Raum mit einem  $K$ -Vektorraum  $TX$  der Dimension  $n$ . Seien  $p_0, \dots, p_k$  aus  $X$  und  $\lambda_0, \dots, \lambda_k$  aus  $K$  mit der Summe  $\lambda_0 + \dots + \lambda_k = 1$  gegeben. Mit

$$(1.1) \quad \lambda_0 p_0 + \lambda_1 p_1 + \dots + \lambda_k p_k = p$$

bezeichnen wir den durch

$$(1.2) \quad \lambda_1(p_1 - p_0) + \dots + \lambda_k(p_k - p_0) + p_0$$

definierten Punkt. Hier sind wiederum  $p_j - p_0$  Verbindungsvektoren. Formal ergibt sich (1.2) aus (1.1), indem  $\lambda_0$  durch  $1 - \sum_{j=1}^k \lambda_j$  ersetzt wird. Wenn statt  $p_0$  ein anderer der Punkte  $p_j$  ausgezeichnet wird, ergibt sich nach der Vorschrift (1.2) dasselbe Element. Diese und ähnliche Aussagen beweist man am besten, indem man den affinen Raum in einen Vektorraum einbettet, wie es im folgenden erläutert wird.

Sei  $a_x: TX \rightarrow X$ ,  $v \mapsto v + x$ . Wir betrachten die Einbettung

$$\alpha_x: X \rightarrow TX \oplus K, \quad z \mapsto (a_x^{-1}(z), 1) = (z - x, 1).$$

Sei  $L(y, x): TX \oplus K \rightarrow TX \oplus K$  der lineare Automorphismus, der auf  $TX$  die Identität ist und  $(0, 1) \in TX \times K$  auf  $(x - y, 1)$  abbildet. Dann gilt  $L(y, x)d_x = d_y$ .

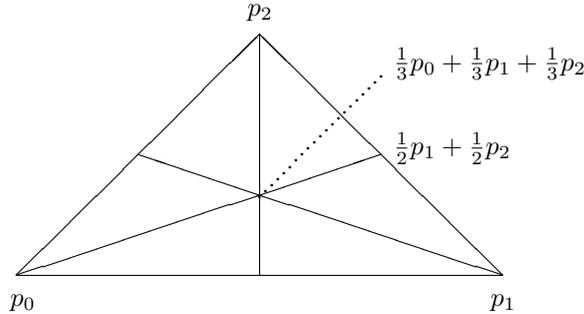
Durch  $\alpha_x$  wird  $X$  auf die Hyperbene  $TX \times \{1\}$  von  $TX \oplus K$  abgebildet. Ist  $\sum \lambda_j = 1$  und sind  $p_0, \dots, p_k \in X$ , so liegt  $\sum \lambda_j \alpha_x(p_j)$  im Bild von  $\alpha_x$ , hat also die Form  $\alpha_x(p)$  für ein  $p \in X$ , das mit  $\sum \lambda_j p_j$  bezeichnet werde. Für jedes  $y \in X$  gilt dann auch  $\sum \lambda_j \alpha_j(p_j) = \alpha_y(p)$ , so daß die Definition von  $p$  unabhängig von der Wahl von  $x$  ist. Mit anderen Worten: Der Punkt

$$\sum_{j=0}^k \lambda_j (p_j - x) + x, \quad \text{für } \sum_{j=0}^k \lambda_j = 1$$

ist unabhängig von der Wahl von  $x$ . Wir nennen  $\sum \lambda_j p_j$  den *Schwerpunkt* der  $p_0, \dots, p_k$  bezüglich der *Gewichte*  $\lambda_0, \dots, \lambda_k$ . Sind  $p_0, p_1, p_2$  drei Punkte der reellen affinen Ebene, versehen mit der Masse 1, so ist

$$\frac{1}{3}p_0 + \frac{1}{3}p_1 + \frac{1}{3}p_2 = \frac{1}{3}(p_0 + p_1 + p_2)$$

der gewöhnliche Schwerpunkt. Ferner ist  $\frac{1}{2}p_0 + \frac{1}{2}p_1$  der Mittelpunkt der Strecke von  $p_0$  nach  $p_1$ .



Für die Schwerpunktrechnung gilt ein Assoziativgesetz.

**(5.1.9) Satz.** Sei  $I = I_1 \cup \dots \cup I_s$  eine disjunkte Vereinigung einer endlichen Menge  $I$ . Seien  $(p_i \mid i \in I)$  Punkte aus dem affinen Raum  $X$ . Seien  $(\mu_j(k) \mid j \in I_k)$  Skalare mit  $\sum_{j \in I_k} \mu_j(k) = 1$ . Seien  $\lambda_1, \dots, \lambda_s$  Skalare mit  $\sum_{i=1}^s \lambda_i = 1$ . Dann gilt

$$\sum_{i=1}^s \lambda_i \left( \sum_{j \in I_i} \mu_j(i) p_j \right) = \sum_{i,j} \lambda_i \mu_j(i) p_j.$$

*Beweis.* Anwendung einer Abbildung  $\alpha_x$ . □

Die Schwerpunkte können zur Koordinatenbeschreibung eines affinen Raumes  $X$  dienen. Sei  $X$   $n$ -dimensional. Wir nennen  $k + 1$  Punkte  $p_0, \dots, p_k$  aus  $X$  *unabhängig*, wenn die Vektoren  $p_1 - p_0, \dots, p_k - p_0$  in  $TX$  linear unabhängig sind.

**(5.1.10) Satz.** Die Punkte  $p_0, \dots, p_k$  sind genau dann unabhängig, wenn die Abbildung

$$\sigma: \left\{ (\lambda_0, \dots, \lambda_k) \in K^{k+1} \mid \sum_{i=0}^k \lambda_i = 1 \right\} \rightarrow X, \quad (\lambda_0, \dots, \lambda_k) \mapsto \sum_{i=0}^k \lambda_i p_i$$

injektiv ist.

*Beweis.* Seien die Punkte unabhängig. Es ist

$$\sum_{i=0}^k \lambda_i p_i = \sum_{i=1}^k \lambda_i (p_i - p_0) + p_0.$$

Die Zusammensetzung von  $\sigma$  mit dem Inversen von  $a_q, q = p_0$ , ist

$$(\lambda_0, \dots, \lambda_k) \mapsto \sum_{i=1}^k \lambda_i (p_i - p_0).$$

Diese Abbildung ist wegen der linearen Unabhängigkeit der  $p_j - p_0$  injektiv, denn die  $\lambda_1, \dots, \lambda_k$  sind beliebig und  $\lambda_0$  ist durch  $\lambda_0 = 1 - \sum_{i=1}^k \lambda_i$  bestimmt. Da  $a_q^{-1}\sigma$  injektiv ist, so auch  $\sigma$ . Analog folgt aus der Injektivität von  $\sigma$  die Unabhängigkeit der Punkte.  $\square$

Wir nennen  $(p_0, \dots, p_n)$  ein *affines Koordinatensystem* des  $n$ -dimensionalen affinen Raumes  $X$ , wenn diese Punkte unabhängig sind. Jeder Punkt  $p \in X$  besitzt dann genau eine Darstellung  $p = \lambda_0 p_0 + \dots + \lambda_n p_n$  mit  $\lambda_0 + \dots + \lambda_n = 1$ . In einer solchen Darstellung heißen  $(\lambda_0, \dots, \lambda_n)$  die *baryzentrischen Koordinaten* von  $p$  bezüglich  $(p_0, \dots, p_n)$ .

**(5.1.11) Satz.** *Sei  $X$  ein affiner Raum über den reellen Zahlen. Die konvexe Hülle einer Menge  $\{p_0, \dots, p_k\} \subset X$  besteht aus allen Punkten der Form  $\sum_{i=0}^k \lambda_i p_i$  mit  $\lambda_i \geq 0$  und  $\sum_{i=0}^k \lambda_i = 1$ .*

*Beweis.* Die genannte Menge enthält alle  $p_i$  und ist nach (5.1.9) konvex. Durch Induktion nach  $\ell$  zeigt man, daß jede  $\{p_0, \dots, p_\ell\}$  enthaltende konvexe Menge auch alle Punkte  $\sum_{i=0}^{\ell} \lambda_i p_i$  mit  $\lambda_i \geq 0$ ,  $\sum_{i=0}^{\ell} \lambda_i = 1$  enthält. Für  $k = 1$  ist das die Definition der Konvexität. Der Induktionsschritt beruht für  $\lambda_{\ell+1} \neq 1$  auf der Gleichheit

$$\lambda \left( \frac{\lambda_0}{\lambda} + \dots + \frac{\lambda_\ell}{\lambda} p_\ell \right) + (1 - \lambda) p_{\ell+1}$$

mit  $\lambda = \lambda_0 + \dots + \lambda_\ell$ .  $\square$

Sind  $p_0, \dots, p_k$  unabhängig, so heißt ihre konvexe Hülle das von  $\{p_0, \dots, p_k\}$  aufgespannte *k-Simplex* ( $k = 2$  Dreieck,  $k = 3$  Tetraeder).

Wir erwähnen noch: Seien  $X$  und  $Y$  affine Räume zu Vektorräumen über demselben Körper  $K$ . Eine Abbildung  $f: X \rightarrow Y$  ist genau dann affin, wenn für  $p_0, \dots, p_k \in X$  und  $\lambda_0, \dots, \lambda_k \in K$  immer  $f(\sum \lambda_j p_j) = \sum \lambda_j f(p_j)$  gilt.

## 5.2 Projektive Räume

Der *projektive Raum*  $P(V)$  des Vektorraums  $V$  über dem Körper  $K$  ist die Menge aller eindimensionalen Unterräume von  $V$ . Ist  $V$   $n$ -dimensional, so ordnen wir  $P(V)$  die *Dimension*  $n - 1$  zu. Ist  $V$  nulldimensional, so gibt es keine eindimensionalen Unterräume, deshalb ist  $P(V)$  die leere Menge; sie bekommt die Dimension  $-1$ . Ist  $V$  eindimensional, so ist  $P(V)$  ein Punkt. Die Elemente von  $P(V)$  heißen *Punkte*. Ist  $U \subset V$  ein Unterraum, so ist  $P(U) \subset P(V)$  ein *projektiver Unterraum*. Für  $U_1, U_2$  in  $V$  gilt  $P(U_1) = P(U_2)$  genau dann, wenn  $U_1 = U_2$  ist.

Wir geben eine zweite Beschreibung von  $P(V)$  an. Sei  $p \in P(V)$  und  $v \in p$  ein von Null verschiedener Vektor der Geraden  $p \subset V$ . Dann besteht  $p$  aus

den skalaren Vielfachen von  $v$ , es ist  $p$  der von  $v$  erzeugte Unterraum,  $p = [v]$  geschrieben. Für Vektoren  $v, w \in V \setminus 0$  ist genau dann  $[v] = [w]$ , wenn für ein  $\lambda \in K^* = K \setminus 0$  die Relation  $v = \lambda w$  besteht. Deshalb betrachten wir auf  $V \setminus 0$  die Äquivalenzrelation:

$$v \sim w \Leftrightarrow \text{es ist } v = \lambda w \text{ für ein } \lambda \in K^*.$$

Die Menge der Äquivalenzklassen kann mit  $P(V)$  identifiziert werden, indem der Klasse von  $v$  das Element  $[v] \in P(V)$  zugeordnet wird. Wir nennen

$$V \setminus \{0\} \rightarrow P(V), \quad v \mapsto [v]$$

die kanonische Abbildung.

Ist  $V = K^{n+1}$  und  $0 \neq x = (x_0, \dots, x_n)$ , so schreiben wir

$$[x] = [x_0, \dots, x_n] = [x_0 : x_1 : \dots : x_n] \in P(K^{n+1})$$

und bezeichnen die  $x_j$  als die *homogenen Koordinaten* von  $x$ . Sie sind definitionsgemäß nur bis auf einen gemeinsamen Faktor bestimmt.

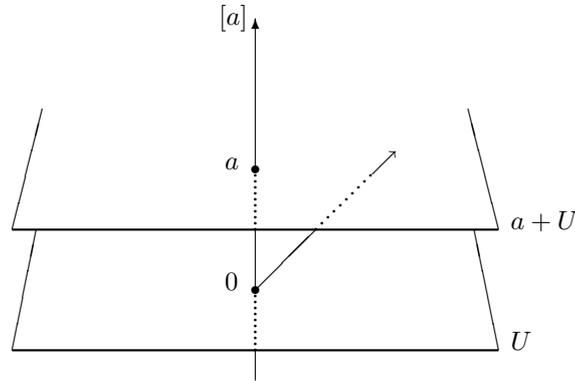
Wir wollen nun erläutern, in welcher Weise der  $n$ -dimensionale projektive Raum als Vervollständigung eines  $n$ -dimensionalen affinen Raumes verstanden werden kann.

Sei  $U \subset V$  eine Hyperebene und  $H = P(U) \subset P(V)$  die zugehörige *projektive Hyperebene*. Sei  $A = P(V) \setminus H$ . Wir wählen  $a \in V \setminus U$  aus. Dann gilt  $V = [a] \oplus U$ . Mit dieser Zerlegung lassen sich die Punkte aus  $A$  als die Punkte  $[\lambda a, u]$ ,  $\lambda \neq 0$ ,  $u \in U$  beschreiben. Da aber  $[\lambda a, u] = [a, \lambda^{-1}u]$  ist, können wir die erste Komponente zu  $a$  normieren. Wir erhalten deshalb eine Bijektion

$$\rho_{U,a}: U \rightarrow a + U \rightarrow A, \quad u \mapsto a + u \mapsto [a + u].$$

Durch Wahl von  $a$  wird also  $A$  mit einem affinen Unterraum zum Vektorraum  $U$  identifiziert. Wir nennen  $\rho_{U,a}$  die *Kartenabbildung* von  $P(V)$  zu  $(U, a)$ , ferner  $A$  den *affinen Raum zur Hyperebene  $H$*  und  $H$  die *unendlich ferne Hyperebene* von  $A$ .

Eine invariantere Beschreibung der Karte zur Hyperebene  $U$  läßt sich folgendermaßen geben. Sei wie in (5.1.6)  $V_U$  die Menge der zu  $U$  komplementären eindimensionalen Unterräumen von  $V$ . Ein solcher Unterraum ist genau dann komplementär, wenn er nicht in  $U$  enthalten ist. Also ist  $V_U = P(V) \setminus P(U)$ . Wir haben in (5.1.6) gesehen, daß  $V_U$  kanonisch ein affiner Raum zum Vektorraum  $\text{Hom}(V/U, U)$  ist.



Für den projektiven Raum  $P(K^{n+1})$  haben wir  $n + 1$  affine Hyperebenen, die durch Nullsetzen jeweils einer Koordinate erhalten werden. Sei

$$U_i = \{[x_0, \dots, x_n] \mid x_i \neq 0\} \subset P(K^{n+1}).$$

Die Bedingung  $x_i \neq 0$  ist unabhängig von der Wahl des Repräsentanten in der Äquivalenzklasse  $[x_0, \dots, x_n]$ . Wir erhalten zueinander inverse Bijektionen

$$U_i \rightarrow K^n, \quad [x_0, \dots, x_n] \mapsto \left( \frac{x_0}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \dots \right)$$

$$K^n \rightarrow U_i, \quad (a_1, \dots, a_n) \mapsto [a_1, \dots, a_{i-1}, 1, a_{i+1}, \dots].$$

Die erste Abbildung ist wohldefiniert, weil sich der frei verfügbare gemeinsame Faktor herauskürzt. Die  $U_i$  überdecken  $P(K^{n+1})$ , das heißt  $P(K^{n+1})$  ist die Vereinigung der Mengen  $U_0, \dots, U_n$ .

Gehen wir umgekehrt von einem Vektorraum  $U$  aus und möchten diesen als affinen Raum eines projektiven Raumes realisieren, so bilden wir  $V = K \oplus U$  und wählen  $H = P(U)$  als unendliche ferne Hyperebene und  $a = (1, 0)$ . Dann haben wir wie eben Bijektionen

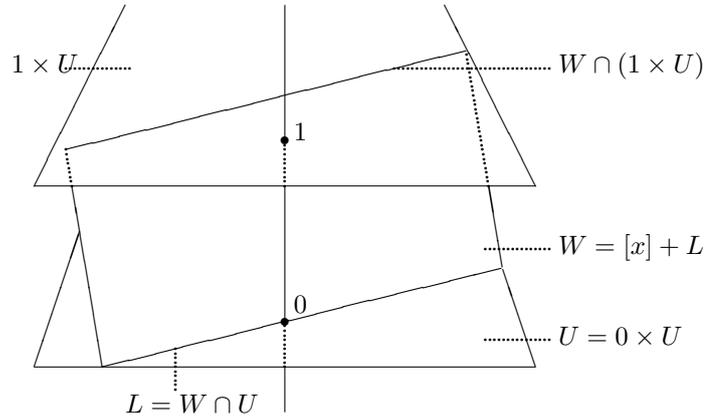
$$U \rightarrow \{1\} \times U \rightarrow P(V) \setminus P(U), \quad u \mapsto [1, u].$$

Wir wollen jetzt die Geometrie von  $U$  und  $P(V)$  vergleichen. Es gilt:

- (5.2.1) Satz.** *Es gibt eine kanonische Bijektion zwischen den Mengen:*
- (1)  $k$ -dimensionale affine Unterräume von  $1 \times U$ .

(2)  $k$ -dimensionale projektive Unterräume von  $P(K \oplus U)$ , die nicht in  $P(U)$  enthalten sind.

Die Bijektion wird durch  $r: P(W) \mapsto W \cap (1 \times U)$  gegeben.



*Beweis.* Sei  $W \not\subset U = \{0\} \times U$ . Dann gibt es in  $W$  ein Element mit erster Komponente 1, etwa  $x$ , und damit ist  $\{1\} \times U = x + U$ . Es folgt  $W \cap (1 \times U) = W \cap (x + U) = x + W \cap U$ . Nach der Dimensionsformel ist

$$\dim(W \cap U) = \dim W + \dim U - \dim(W \oplus U) = \dim W - 1,$$

weil  $W \oplus U = V$  die Dimension  $n + 1$  hat.

Sei umgekehrt  $x + L$  ein  $k$ -dimensionaler affiner Unterraum in  $1 \times U$  mit zugehöriger Richtung  $L \subset U$ . Dann ist  $W = [x] + L$  ein  $(k + 1)$ -dimensionaler Unterraum von  $V$ , denn  $L \subset U$  und  $x \notin U$ . Es ist für  $x - y \in L$  auch  $[x] + L = [y] + L$ , so daß  $W$  nicht von der Auswahl von  $x$  zur Beschreibung von  $x + L$  abhängt. Wir können deshalb eine Abbildung in der umgekehrten Richtung durch  $s: x + L \mapsto P([x] + L)$  definieren. Es handelt sich um die Umkehrabbildung zu  $r: P(W) \mapsto W \cap (1 \times U)$ . Es ist nämlich  $([x] + L) \cap (1 \times U) = x + L$  nach Konstruktion, also  $rs = \text{id}$ .

Umgekehrt:  $P(W) \mapsto W \cap (1 \times U) = x + W \cap U \mapsto P([x] + W \cap U)$ ; es ist  $W \cap U \not\subset [x] + W \cap U \subset W$ , also wegen  $\dim(W \cap U) = \dim W - 1$ , gilt  $\dim([x] + W \cap U) = \dim W$ , also  $[x] + W \cap U = W$ .  $\square$

Einen eindimensionalen projektiven Unterraum  $P(U)$  von  $P(V)$  nennen wir *projektive Gerade*, einen zweidimensionalen *projektive Ebene*. Wir schreiben auch

$$KP^n = P^n(K) = P(K^{n+1}),$$

um die Dimension zu kennzeichnen.

Wir können jetzt einsehen, daß in einer projektiven Ebene  $P^2$  je zwei verschiedene Geraden genau einen Schnittpunkt haben. Sei

$$P^2 = P(V) \supset P(U_1), P(U_2), P(U_1) \neq P(U_2)$$

und gelte  $\dim U_1 = \dim U_2 = 2$ . Wegen  $U_1 \neq U_2$  ist  $U_1 + U_2 = V$  und deshalb nach der Dimensionsformel  $\dim(U_1 \cap U_2) = 1$ . Folglich ist  $P(U_1) \cap P(U_2) = P(U_1 \cap U_2)$  ein Punkt.

Ebenfalls lassen sich je zwei verschiedene Punkte  $p_1, p_2 \in P^2$  durch genau eine Gerade verbinden. Ist  $p_i = P(U_i)$ ,  $\dim U_i = 1$ , so ist  $U_1 \oplus U_2$  zweidimensional und  $L = P(U_1 \oplus U_2)$  die gewünschte Gerade.

In diesem Zusammenhang wird der Begriff „parallel“ ein relativer Begriff: Nach Auszeichnung einer Geraden  $H$  in  $KP^2$  als „unendlich fern“ sind zwei affine Geraden in der affinen Ebene  $A = KP^2 \setminus H$  genau dann parallel, wenn sie sich in  $H$  schneiden. Die Wahl der Geraden  $H$  markiert einen Standpunkt. Die projektive Ebene umfaßt viele affine Ebenen. Es stellt sich heraus, daß man Aussagen über das Schneiden und Verbinden affiner Unterräume besser im vollständigen projektiven Raum untersucht.

Die analytische Geometrie beschreibt geometrische Figuren als Teilmengen von Vektorräumen.

Eine Teilmenge  $M$  des  $K^n$  läßt sich oft als Nullstellenmenge einer Gleichung gewinnen:

$$M = \{(x_1, \dots, x_n) \mid f(x_1, \dots, x_n) = 0\}$$

mit einer geeigneten Funktion  $f: K^n \rightarrow K$ . Für den projektiven Raum können wir, mit einer Vorsichtsmaßnahme, ebenso vorgehen: Die  $(n+1)$ -Tupel  $(x_0, \dots, x_n)$  und  $(\lambda x_0, \dots, \lambda x_n)$  beschreiben für  $\lambda \in K^*$  denselben Punkt in  $KP^n$ . Aus diesem Grunde betrachten wir Funktionen  $f: K^{n+1} \rightarrow K$ , bei denen  $f(x_0, \dots, x_n) = 0$  und  $f(\lambda x_0, \dots, \lambda x_n) = 0$  gleichwertige Relationen sind. Diese Forderung wird durch die sogenannten homogenen Funktionen erfüllt:  $f$  heißt *homogen vom Grad*  $k \in \mathbb{N}$ , wenn für  $\lambda \in K^*$

$$f(\lambda x_0, \dots, \lambda x_n) = \lambda^k f(x_0, \dots, x_n)$$

gilt. Für eine homogene Funktion  $f: K^{n+1} \rightarrow K$  ist also

$$\{(x_0, \dots, x_n) \in P(K^{n+1}) \mid f(x_0, \dots, x_n) = 0\}$$

eine wohldefinierte Teilmenge des Raumes  $P(K^{n+1})$ . Typische Beispiele für homogene Funktionen vom Grad  $k$  sind die *homogenen Polynome* vom Grad  $k$

$$p(x_0, \dots, x_n) = \sum_{i_0 + \dots + i_n = k} a(i_0, \dots, i_n) x_0^{i_0} \dots x_n^{i_n}$$

mit Koeffizienten  $a(i_0, \dots, i_n) \in K$ , zum Beispiel die homogenen quadratischen Polynome

$$q(x_0, \dots, x_n) = \sum_{i,j=0}^n a_{ij} x_i x_j.$$

Betrachten wir den affinen Teilraum  $U_0 \subset P(K^{n+1})$  und die Bijektion

$$\rho_0: K^n \rightarrow U_0 \subset P(K^{n+1}), (x_1, \dots, x_n) \mapsto [1, x_1, \dots, x_n].$$

Die Teilmenge  $U_0 \cap \{[x_0, \dots, x_n] \mid q(x_0, \dots, x_n) = 0\}$  wird vermöge  $\rho_0$  zu einer Teilmenge von  $K^n$ . Sie erscheint als Nullstellenmenge der Gleichung  $f(x_1, \dots, x_n) = 0$ , wobei

$$f(x_1, \dots, x_n) = q(1, x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij}x_i x_j + \sum_{i=1}^n (a_{0i} + a_{i0})x_i + a_{00}.$$

Wir sehen deshalb, daß die sogenannte *projektive Quadrik*

$$Q_q = \{[x_0, \dots, x_n] \mid q(x_0, \dots, x_n) = 0\}$$

als Vervollständigung der affinen Quadrik

$$Q_f = \{(x_1, \dots, x_n) \mid f(x_1, \dots, x_n) = 0\}$$

erscheint; gewisse unendlich ferne Punkte werden hinzugehügt.

Betrachten wir als Beispiel Ellipse, Parabel und Hyperbel. Der Grundkörper sei der Körper der reellen Zahlen. Wir schreiben die affinen und die zugehörigen homogenen Gleichungen nebeneinander:

$$\begin{array}{lll} \text{I:} & x_1^2 + x_2^2 - 1 = 0 & x_1^2 + x_2^2 - x_0^2 = 0 \\ \text{II:} & x_1^2 - x_2^2 - 1 = 0 & x_1^2 - x_2^2 - x_0^2 = 0 \\ \text{III:} & x_1^2 - 2x_2 = 0 & x_1^2 - 2x_0x_2 = 0. \end{array}$$

Die eventuell hinzugefügten unendlich fernen Punkte  $[x_0, x_1, x_2]$  erhalten wir durch die Bedingung  $x_0 = 0$ . Im Fall I führt das zu  $x_1^2 + x_2^2 = 0$ , also zu  $x_1 = x_2 = 0$ . Das Tripel  $[0, 0, 0]$  ist aber nicht erlaubt, d. h. die unendlich ferne Gerade hat einen leeren Schnitt mit der projektiven Quadrik, der Kreis ist schon projektiv vollständig. Im Fall II erhalten wir  $x_1^2 - x_2^2 = 0$ . Dazu gehören zwei Punkte im projektiven Raum:  $[0, 1, 1]$  und  $[0, -1, 1]$ , denn alle Lösungen  $(x_1, x_2) = (a, \pm a)$ ,  $a \neq 0$ , führen zu diesen Punkten. Die Hyperbel wird durch zwei Punkte im Unendlichen vervollständigt. Die beiden Asymptoten schneiden dort die unendlich ferne Gerade. Im Fall III erhalten wir  $x_1^2 = 0$ , also einen Punkt im Unendlichen  $[0, 0, 1]$ .

Ellipse, Parabel und Hyperbel sind, projektiv betrachtet, dieselben Objekte, sie „erscheinen“ nur verschieden, je nach Lage der unendlich fernen Geraden zur projektiven Quadrik.

## 5.3 Projektivitäten

Eine bijektive lineare Abbildung  $\alpha: V \rightarrow V$  bildet einen eindimensionalen Unterraum  $v \subset V$  wieder auf einen solchen ab. Deshalb wird eine *projektive Ab-*

*bildung*, auch *Projektivität* genannt, induziert:

$$P(\alpha): P(V) \rightarrow P(V), \quad v \mapsto \alpha(v),$$

die auch  $[v] \mapsto [\alpha(v)]$  geschrieben wird. Die Gesamtheit dieser Selbstabbildungen ist eine Gruppe  $P \operatorname{Aut}(V)$ . Der Homomorphismus

$$P: \operatorname{Aut}(V) \rightarrow P \operatorname{Aut}(V), \quad \alpha \mapsto P(\alpha)$$

ist surjektiv. Falls  $P(\alpha) = \operatorname{id}$  ist, so gilt für jedes  $v \in V \setminus \{0\}$  sicherlich  $\alpha(v) = \lambda(v)v$ , mit einem  $\lambda(v) \in K$ . Wir zeigen, daß  $\lambda(v)$  von  $v$  unabhängig ist. Seien  $v, w$  linear unabhängig. Dann ist

$$\alpha(v+w) = \lambda(v+w)v + \lambda(v+w)w, \quad \alpha(v) + \alpha(w) = \lambda(v)v + \lambda(w)w,$$

und durch Koeffizientenvergleich folgt  $\lambda(v) = \lambda(w)$ . Der Kern  $P^{-1}(\operatorname{id})$  von  $P$  besteht also aus den Vielfachen der Identität. Für  $V = K^{n+1}$  schreiben wir  $PGL(n+1, K) = P \operatorname{Aut}(K^{n+1})$ .

Die Elemente der projektiven Gruppe  $PGL(2, K)$  sind von der Form

$$[z_0, z_1] \mapsto [az_0 + bz_1, cz_0 + dz_1],$$

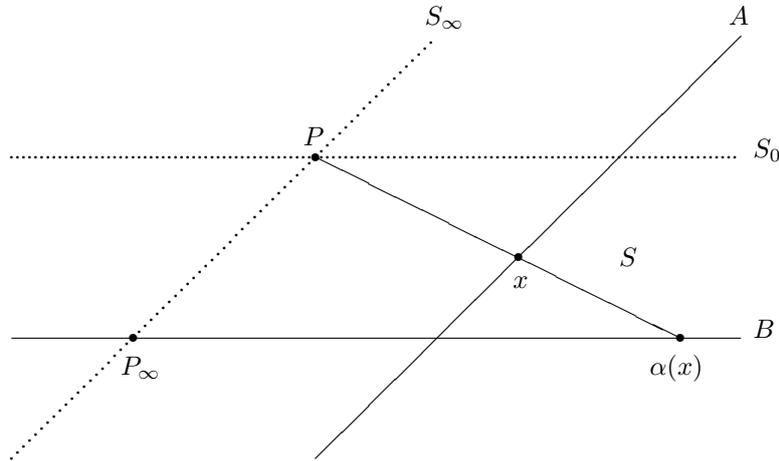
wobei die Determinante  $ad - bc \neq 0$  ist. Falls  $cz_0 + dz_1 \neq 0$  und  $z_1 \neq 0$  ist und wir  $z = z_0/z_1$  setzen, so ist der Punkt rechts gleich  $\left[\frac{az+b}{cz+d}, 1\right]$ . Wenn wir also die affine Gerade  $\{[z, 1] \mid z \in K\} \subset KP^1$  betrachten, so ist eine Projektivität eine gebrochen-lineare Abbildung  $z \mapsto \frac{az+b}{cz+d}$ .

Der Punkt  $z \in K$ , für den  $cz + d = 0$  ist (solche gibt es nur, falls  $c \neq 0$  ist), wird auf den unendlich fernen Punkt abgebildet. Die „Division durch 0“ erhält also in diesem Zusammenhang eine Sinnzuweisung. Der unendlich ferne Punkt selbst wird auf  $a/c$  abgebildet. Das entspricht für  $K = \mathbb{R}$  dem aus der Analysis Geläufigen,

$$\lim_{z \rightarrow \infty} \frac{az+b}{cz+d} = \frac{a}{c}.$$

Die projektiven Räume und Projektivitäten haben etwas mit Projektionen zu tun (daher der Name). Betrachten wir als Beispiel die *Zentralprojektion*  $\alpha$

von einem Punkt  $P \in \mathbb{R}^2$  einer Geraden  $A$  auf eine Gerade  $B$ .



Wenn sich der Projektionsstrahl  $S$  der Geraden  $S_0$  nähert, wandert  $\alpha(x)$  ins Unendliche, der Punkt  $P_0$  wird auf den unendlich fernen Punkt von  $B$  abgebildet. Umgekehrt erscheint  $P_\infty \in B$  als Bild des unendlich fernen Punktes von  $A$ . Wie sieht die Abbildung  $\alpha$  formelmäßig aus?

Sei, der Einfachheit halber,  $A$  die  $y$ - und  $B$  die  $x$ -Achse. Sei  $P = (a, b)$ . Die Gerade durch  $(a, b)$  und  $(0, x)$  ist  $\{t(a, b - x) + (0, x) \mid t \in \mathbb{R}\}$ , und sie trifft die  $x$ -Achse im Punkt  $(\frac{ax}{x-b}, 0)$ . Also ist  $\alpha$  die gebrochen-lineare Abbildung

$$\alpha(x) = \frac{ax}{x - b}$$

und somit eine Projektivität im oben definierten Sinne.

Ein  $k$ -Tupel von Punkten  $[v_1], \dots, [v_k]$  aus  $P(V)$  heißt *unabhängig* (oder, *in allgemeiner Lage*), wenn die Vektoren  $v_1, \dots, v_k$  in  $V$  linear unabhängig sind. Unabhängige  $k$ -Tupel spannen einen  $(k - 1)$ -dimensionalen projektiven Unterraum auf; er ist der kleinste, in dem sie alle liegen. Ist  $\dim P(V) = n$ , also  $\dim V = n + 1$ , so heißt ein  $(n + 2)$ -Tupel  $(p_0, p_1, \dots, p_{n+1})$  von Punkten in  $P(V)$  ein *projektives Koordinatensystem*, wenn je  $n + 1$  Punkte unabhängig sind. Diese Terminologie wird durch den folgenden Satz gerechtfertigt.

**(5.3.1) Satz.** Sind  $p_0, \dots, p_{n+1}$  und  $q_0, \dots, q_{n+1}$  projektive Koordinatensysteme von  $P(V)$ , so gibt es genau eine Projektivität  $P(\alpha)$ , die  $p_j$  auf  $q_j$  abbildet ( $j = 0, \dots, n + 1$ ).

*Beweis.* Wir wählen eine Basis  $e_0, \dots, e_n$  von  $V$  und setzen  $e_{n+1} = e_0 + \dots + e_n$ . Dann sind je  $n+1$  der Vektoren  $e_0, \dots, e_{n+1}$  linear unabhängig, es liegt also ein projektives Koordinatensystem vor.

Existenz von  $P(\alpha)$  mit  $P(\alpha)(p_j) = [e_j]$ . Da die  $p_0, \dots, p_n$  unabhängig sind, gibt es eine Basis  $v_0, \dots, v_n$  von  $V$  mit  $[v_j] = p_j$  für  $j = 0, \dots, n$ . Es gibt ferner eine lineare Relation  $v_{n+1} = \lambda_0 v_0 + \dots + \lambda_n v_n$ . Kein  $\lambda_j$  ist Null, weil je  $n+1$  der Vektoren  $v_0, \dots, v_{n+1}$  linear unabhängig ist. Wir können deshalb auch  $w_0 = \lambda_0 v_0, \dots, w_n = \lambda_n v_n$  als Basis verwenden. Dann gilt  $[w_j] = p_j$  und  $[w_0 + \dots + w_n] = [v_{n+1}] = p_{n+1}$ . Sei  $\alpha: V \rightarrow V$  der durch  $\alpha(w_j) = e_j$  für  $j \in \{0, \dots, n\}$  bestimmte lineare Automorphismus. Er bildet dann auch  $\alpha(w_{n+1})$  auf  $e_{n+1}$  ab, und folglich hat  $P(\alpha)$  die gewünschten Eigenschaften. Analog kann man die  $q_j$  auf die  $e_j$  abbilden und dann durch Zusammensetzung die  $p_j$  auf die  $q_j$ .

Eindeutigkeit. Sei  $P(\alpha)[e_j] = [e_j]$  für alle  $j$ . Es gilt dann  $\alpha(e_j) = \lambda_j e_j$  für  $j = 0, \dots, n$  und deshalb  $\alpha(e_{n+1}) = \lambda_0 e_0 + \dots + \lambda_n e_n$ . Andererseits ist  $\alpha(e_{n+1}) = \lambda e_{n+1} = \lambda e_0 + \dots + \lambda e_n$ . Folglich ist  $\lambda_0 = \dots = \lambda_n = \lambda$  und deshalb  $P(\alpha)$  die Identität.  $\square$

Ist  $p_0, \dots, p_{n+1}$  ein projektives Koordinatensystem von  $P(V)$  und  $e_0, \dots, e_n$  die Standardbasis von  $K^{n+1}$ , so gibt es genau einen projektiven Isomorphismus

$$P(\alpha): P(V) \rightarrow P(K^{n+1}), \quad p_j \mapsto [e_j]$$

für  $j = 0, \dots, n+1$ . Jedem  $[x] = [x_0, \dots, x_n] \in P(K^{n+1})$  entspricht genau ein Punkt  $P(V)$ , und in diesem Sinne ist  $p_0, \dots, p_{n+1}$  ein Koordinatensystem.

Wir betrachten speziell den Fall  $\dim P(V) = 1$ ,  $\dim V = 2$ . Der zugehörige Standardraum  $P(K^2)$  ist  $K \cup \{\infty\}$ , wobei  $K = \{[1, k] \mid k \in K\}$  und  $\infty = [0, 1]$ . Wir setzen

$$[e_0] = [1, 0] = 0, \quad [e_1] = [0, 1] = \infty, \quad [e_2] = [1, 1] = 1.$$

Sind  $p_0, p_1, p_2$  paarweise verschiedene Punkte in  $P(V)$ , so bilden sie ein projektives Koordinatensystem. Es gibt deshalb genau einen Isomorphismus

$$P(\alpha): P(V) \rightarrow P(K^2), \quad p_j \mapsto [e_j].$$

Ein vierter Punkt  $p \in P(V)$  wird dabei auf ein eindeutig bestimmtes Element  $P(\alpha)(p) \in P(K^2) = K \cup \{\infty\}$  abgebildet. Wir schreiben

$$P(\alpha)(p) = D(p_0, p_1, p_2, p) \in K \cup \{\infty\}$$

und nennen dieses Körperelement (eventuell  $\infty$ ) das *Doppelverhältnis* des Quadrupels  $(p_0, p_1, p_2, p)$ . Es läßt sich nämlich in der folgenden Weise ausrechnen (was dann den Namen erklärt).

Sei  $V = K^2$  und also  $P(V) = K \cup \{\infty\}$ . Die gebrochen-lineare Transformation

$$K \cup \{\infty\} \rightarrow K \cup \{\infty\}, \quad p \mapsto \frac{p - p_0}{p - p_1} : \frac{p_2 - p_0}{p_2 - p_1}$$

bildet  $p_0, p_1, p_2$  auf  $0, \infty, 1$  ab und ist deshalb das gewünschte  $P(\alpha)$ . Für  $p_j = \infty$  muß man die Formel (richtig) interpretieren. (Es gibt verschiedene Vereinbarungen für das Doppelverhältnis. Alle Permutationen der vier Variablen kommen vor.) Nach unserer abstrakten Definition des Doppelverhältnisses ist es auch für den Fall definiert, daß  $p$  einer der Punkte  $p_0, p_1, p_2$  ist. Das Doppelverhältnis ist invariant gegenüber Projektivitäten. Hier noch eine Aufgabe:

**(5.3.2) Satz.** *Die Projektivitäten von  $P(K \times V)$ , die die unendlich ferne Hyperebene  $H = P(V)$  in sich abbilden, bilden eine zur affinen Gruppe von  $V$  isomorphe Untergruppe.  $\square$*

# Kapitel 6

## Normalformen

Wird ein Endomorphismus  $f$  eines Vektorraumes  $V$  bezüglich zweier Basen durch Matrizen  $A$  und  $B$  beschrieben, so besagt die Basiswechselformel, daß eine invertierbare Matrix  $S$  existiert, mit der die Relation  $A = SBS^{-1}$  gilt, die Matrizen also konjugiert sind. Das Normalformenproblem besteht darin, in der Konjugationsklasse einer Matrix eine möglichst einfache zu finden und die Konjugationsklassen durch geeignete Daten, die man aus der Matrix ablesen kann, zu charakterisieren.

### 6.1 Zyklische Endomorphismen

Ist  $U$  ein  $f$ -stabiler Unterraum von  $V$  und  $v \in U$ , so liegen alle Vektoren  $v, f(v), f^2(v), \dots$  in  $U$ . Für jeden Vektor  $v \in V$  ist die lineare Hülle der Vektoren  $v, f(v), f^2(v), \dots$  ein  $f$ -stabiler Unterraum  $L(v, f)$ .

Das Paar  $(V, f)$  heiße *zyklisch* und  $f$  ein *zyklischer* Endomorphismus, wenn es einen Vektor  $v \in V$  so gibt, daß  $V = L(v, f)$  ist. Wir untersuchen zunächst zyklische Endomorphismen als Bausteine für allgemeine.

Wir setzen  $v_j = f^j(v)$ . Die Familie der  $v_j$  ist linear abhängig, da  $V$  endlichdimensional ist. Sei  $n$  minimal gewählt, so daß  $v_n$  von  $v_0, \dots, v_{n-1}$  linear abhängig ist. Dann besteht eine Relation

$$v_n + a_1 v_{n-1} + \dots + a_n v_0 = 0, \quad a_j \in K.$$

Der Vektor  $v_n$  kann darin durchaus der Nullvektor sein. Aus der Relation folgt durch Induktion nach  $k$ , daß alle  $v_{n+k}$  für  $k \geq 0$  in dem von  $v_0, \dots, v_{n-1}$  erzeugten Unterraum liegen. Die Menge dieser Vektoren ist aber auch linear unabhängig, da eine echte lineare Relation der Minimalität von  $n$  widerspräche.

Damit haben wir in  $\{v_0, \dots, v_{n-1}\}$  eine Basis von  $L(v, f)$  gewonnen. Die Matrix von  $f$  bezüglich dieser Basis ist

$$\begin{pmatrix} 0 & & & & -a_n \\ 1 & 0 & & & -a_{n-1} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & 0 & -a_2 \\ & & & 1 & -a_1 \end{pmatrix}.$$

Die letzte Spalte ist kenntlich. Unterhalb der Hauptdiagonale stehen Einsen. Alles andere sind Nullen.

**(6.1.1) Notiz.** *Das charakteristische Polynom der Matrix der vorigen Matrix ist*

$$x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n.$$

*Beweis.* Induktion nach  $n$  mittels Entwicklung nach der ersten Zeile.  $\square$

Die Relation schreiben wir in der Form (mit  $f^0 = \text{id}$ )

$$(f^n + a_1f^{n-1} + \dots + a_nf^0)(v) = 0.$$

Sie besagt:  $v$  liegt im Kern des Endomorphismus

$$f^n + a_1f^{n-1} + \dots + a_nf^0.$$

Durch wiederholte Anwendung von  $f$  sehen wir, daß alle Vektoren  $v_j$  im Kern dieses Endomorphismus liegen. Also handelt es sich um die Nullabbildung.

## 6.2 Nilpotente Endomorphismen

Der einfachste Fall des Vorhergehenden liegt vor, wenn  $v_n = 0$  ist. Dann gilt  $f^n = 0$ . Wir nennen  $f$  *nilpotent*, wenn für ein  $N \in \mathbb{N}$  die Relation  $f^N = 0$  besteht. Ist  $f$  nilpotent und  $v \in V$ , so gibt es eine kleinste natürliche Zahl  $k$  derart, daß  $f^{k-1}(v) \neq 0$  aber  $f^k(v) = 0$  ist. Wir nennen  $k$  die *Stufe* von  $v$  bezüglich  $f$ .

**(6.2.1) Satz.** *Sei  $f$  nilpotent,  $w \in V$  ein Vektor maximaler Stufe und  $U$  ein bezüglich Inklusion maximaler  $f$ -stabiler Unterraum mit  $L(w, f) \cap U = 0$ . Dann gilt  $V = L(w, f) \oplus U$ .*

*Beweis.* Zu zeigen ist, daß  $L(w, f)$  und  $U$  zusammen den Raum  $V$  erzeugen. Andernfalls gibt es ein  $v \notin L(w, f) + U$ . Es gibt ein kleinstes  $j$ , so daß  $f^j(v)$

in  $L(w, f) + U$  liegt, da ja  $f^N(v) = 0$  jedenfalls darin liegt. Indem wir  $v$  durch  $f^{j-1}(v)$  ersetzen, erhalten wir einen eventuell neuen Vektor  $v$ , der nicht in  $L(w, f) + U$  liegt, wohl aber dessen Bild bei  $f$ . Somit gilt

$$f(v) = \lambda w + f(w') + u, \quad u \in U, \quad w' \in L(w, f),$$

weil  $w, f(w), \dots, f^{k-1}(w)$  eine Basis von  $L(w, f)$  ist und deshalb jeder Vektor von  $L(w, f)$  die angegebene Form  $\lambda w + f(w')$  hat. Nach Voraussetzung über  $k$  hat  $v$  höchstens die Stufe  $k$ , also  $f(v)$  höchstens die Stufe  $k-1$ . Da aber  $L(w, f) + U$  wegen  $L(w, f) \cap U = 0$  eine direkte Zerlegung in  $f$ -stabile Unterräume ist, so ist die Stufe von  $f(v)$  mindestens gleich der von  $\lambda w + f(w')$ . Also muß  $\lambda = 0$  sein.

Aus  $v - w' \in L(w, f) + U$  würde  $v \in L(w, f) + U$  folgen; also liegt  $v' = v - w'$  nicht in  $L(w, f) + U$ . Es gilt  $f(v') = f(v) - f(w') = u \in U$ . Folglich ist  $L(v', f) + U$  ein  $f$ -stabiler Unterraum, der größer als  $U$  ist. Es gilt aber auch

$$L(w, f) \cap (L(v', f) + U) = 0,$$

denn sonst gäbe es  $w_1 \in L(w, f)$  von der Form

$$w_1 = u + \lambda v', \quad \lambda \neq 0, \quad u \in U,$$

und dann wäre

$$v' = \frac{1}{\lambda}(w_1 - u) \in L(w, f) + U,$$

was wir oben schon ausgeschlossen hatten. Damit haben wir einen Widerspruch zur Maximalität von  $U$  erhalten.  $\square$

Indem wir den vorstehenden Satz induktiv anwenden, erhalten wir den folgenden *Struktursatz für nilpotente Morphismen*:

**(6.2.2) Satz.** *Sei  $f: V \rightarrow V$  nilpotent. Dann gibt es eine direkte Zerlegung von  $V$  in  $f$ -stabile zyklische Unterräume  $V = V_1 \oplus \dots \oplus V_r$ . Wählen wir die Bezeichnungen so, daß immer  $n_j = \dim V_j \geq n_{j+1} = \dim V_{j+1} > 0$  gilt, so ist  $(n_1, \dots, n_r)$  durch  $f$  eindeutig bestimmt.*

*Beweis.* Die Existenz folgt induktiv aus dem vorigen Satz. Die Anzahl  $r$  ist die Dimension des Kerns von  $f$ . Die Anzahl der Summanden mit  $n_s > j$  ist die Anzahl der zyklischen Summanden einer entsprechenden Zerlegung von  $f: f^j(V) \rightarrow f^j(V)$ .  $\square$

Wir bemerken zur Abrundung noch:

**(6.2.3) Satz.** *Ein Endomorphismus  $f: V \rightarrow V$  eines  $n$ -dimensionalen Vektorraumes ist genau dann nilpotent, wenn sein charakteristisches Polynom  $x^n$  ist.*

*Beweis.* Aus (6.1.1) und (6.2.2) entnehmen wir, daß ein nilpotenter Endomorphismus das angegebene Polynom hat. Die Umkehrung folgt mittels (3.4.5).  $\square$

## 6.3 Jordansche Normalform

Sei  $f: V \rightarrow V$  ein Endomorphismus und  $\lambda$  ein Eigenwert von  $f$ . Für die Räume  $V^t(\lambda) = \text{Kern}(f - \lambda)^t$  gilt  $V^t(\lambda) \subset V^{t+1}(\lambda)$ . (Wir schreiben  $f - \lambda$  anstatt  $f - \lambda \cdot \text{id}$ .) Es gibt deshalb ein minimales  $k \geq 1$ , so daß für alle  $l \geq k$  die Gleichheit  $V^k(\lambda) = V^l(\lambda)$  gilt. Diesen Raum bezeichnen wir mit  $V_\lambda$  und nennen ihn den *verallgemeinerten Eigenraum* von  $f$  zum Eigenwert  $\lambda$  oder den  *$\lambda$ -primären Raum* von  $f$ . Aus der Konstruktion folgt sofort:

**(6.3.1) Notiz.** Die Räume  $V^t(\lambda)$  sind  $f$ -stabil. □

Auf  $V_\lambda$  ist  $f - \lambda = g$  nilpotent. Die direkte Zerlegung von  $V_\lambda$  nach Satz (6.2.2) in  $g$ -stabile zyklische Unterräume ist dann auch eine in  $f$ -stabile Räume. Nach (6.1.1) hat  $f = g + \lambda$  in jedem  $g$ -zyklischen Summanden eine Matrix der Form

$$\begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & \ddots & \ddots & & \\ & & & 1 & \lambda \end{pmatrix},$$

auf der Diagonalen steht immer  $\lambda$ , darunter jeweils eine 1, der Rest sind Nullen. Eine Matrix dieser Form heißt *Jordan-Matrix* und eine Blockdiagonalmatrix aus Jordan-Matrizen eine *jordansche Normalform*. Wir wissen also, daß  $f$  auf jedem verallgemeinerten Eigenraum durch eine Matrix in jordanischer Normalform beschrieben werden kann.

**(6.3.2) Satz.** Das charakteristische Polynom von  $f$  zerfällt in Linearfaktoren  $\chi = \prod_{j=1}^r (x - \lambda_j)^{n(j)}$  mit paarweise verschiedenen  $\lambda_j$ . Dann ist  $V$  die direkte Summe der verallgemeinerten Eigenräume  $W_j$  zu den Eigenwerten  $\lambda_j$ . Die Dimension von  $W_j$  ist  $n(j)$ .

*Beweis.* Das charakteristische Polynom von  $f$  auf  $W_j$  hat die Form  $(x - \lambda)^{m(j)}$ , und dieses ist ein Faktor von  $\chi$ . Also ist  $m(j) \leq n(j)$ . Bestünde nicht Gleichheit, so hätte die von  $f$  auf  $V/W_j$  induzierte Abbildung einen Eigenwert  $\lambda_j$ . Sei  $U_j$  das Urbild des Eigenraumes von  $f$  auf  $V/W_j$ . Dann gilt  $(f - \lambda)(U_j) \subset W_j$ , was der Maximalität des verallgemeinerten Eigenraums widerspräche. Damit haben wir die Dimensionsaussage des Satzes gezeigt. Der Durchschnitt  $W_j \cap \sum_{k \neq j} W_k$  ist Null, denn andernfalls enthielte  $\sum_{k \neq j} U_k$  einen Eigenvektor zu  $\lambda_j$ , was nicht sein kann, da das charakteristische Polynom von  $f$  auf dieser Summe nur Nullstellen ungleich  $\lambda_j$  hat. Also ist  $\sum_k W_k$  in die  $W_k$  direkt zerlegt und aus Dimensionsgründen gleich  $V$ . □

Im Beweis des letzten Satzes haben wir benutzt, daß ein Faktor eines in Linearfaktoren zerlegten Polynoms ein Produkt eines Teils dieser Faktoren ist. Das ist für komplexe Polynome relativ klar und verlangt allgemein eine Überlegung zur Teilbarkeitslehre von Polynomen, die wir hier übergehen. Über den

komplexen Zahlen ist die Voraussetzung des letzten Satzes immer erfüllt, weil überhaupt alle Polynome in Linearfaktoren zerfallen. Jede Matrix besitzt eine jordanische Normalform und, wegen des Eindeutigkeitsatzes (6.2.2), sind zwei Matrizen genau dann konjugiert, wenn die Normalform (bis auf die Reihenfolge) dieselben Jordan-Matrizen enthält.

## 6.4 Normalformen für Zwei-mal-zwei Matrizen

Als Illustration der Normalformtheorie soll im folgenden das Normalformenproblem im einzelnen für verschiedene Situationen von  $(2,2)$ -Matrizen ohne allgemeine Theorie durchgerechnet werden.

### (6.4.1) Konjugationsklassen in $GL(2, \mathbb{C})$ .

Erster Fall:  $A \in GL(2, \mathbb{C})$  habe zwei verschiedene Eigenwerte  $\lambda_1$  und  $\lambda_2$ . Das charakteristische Polynom ist dann  $(x - \lambda_1)(x - \lambda_2)$ . Indem man Eigenvektoren als Basis nimmt, sieht man, daß  $A$  zu  $\text{Dia}(\lambda_1, \lambda_2)$  konjugiert ist. Vertauschen wir die Basisvektoren, so erhalten wir  $\text{Dia}(\lambda_2, \lambda_1)$ . Auf die Reihenfolge der Eigenwerte kommt es also nicht an. Das charakteristische Polynom und folglich die Menge der Eigenwerte sind Invarianten der Konjugationsklasse.

Das charakteristische Polynom einer  $(2,2)$ -Matrix  $A$  ist

$$\chi_A(x) = x^2 - \text{Sp}(A)x + \det(A),$$

läßt sich also durch Spur und Determinante berechnen. Genau dann hat  $\chi_A$  zwei verschiedene Nullstellen, wenn

$$\text{Sp}^2(A) - 4 \det(A) \neq 0$$

ist. Die Menge  $\{\lambda_1, \lambda_2\}$  ist durch  $\text{Sp}(A)$  und  $\det(A)$  eindeutig bestimmt.

Zweiter Fall: Die Matrix  $A \in GL(2, \mathbb{C})$  hat nur einen Eigenwert  $\lambda$ . Das charakteristische Polynom ist dann  $(x - \lambda)^2$ . Es gibt einen Eigenvektor  $w_1$  zum Eigenwert  $\lambda$ . Wir ergänzen ihn durch  $w_2$  zu einer Basis. Die lineare Abbildung  $A$  hat bezüglich dieser Basis eine Matrix der Form

$$\begin{pmatrix} \lambda & \mu \\ 0 & \lambda \end{pmatrix}.$$

Daraus sehen wir, daß  $(A - \lambda I)^2 = 0$  ist. Wir unterscheiden die Fälle  $A - \lambda I = 0$  und  $A - \lambda I \neq 0$ . Im ersten Fall gibt es eine Basis aus Eigenvektoren, und  $A$  ist zu  $\text{Dia}(\lambda, \lambda)$  konjugiert. Andernfalls gibt es einen Vektor  $v_1$ , der nicht im Kern von  $A - \lambda I$  liegt. Sei  $v_2 = (A - \lambda I)v_1$ . Dann sind  $v_1$  und  $v_2$  linear unabhängig, denn eine echte lineare Relation  $av_1 + b(A - \lambda I)v_1 = 0$  würde  $v_1$  als Eigenvektor zum Eigenwert  $\lambda - ab^{-1}$  erweisen. Da  $(A - \lambda I)^2 = 0$  ist, ist  $v_2$  ein Eigenvektor.

Bezüglich der Basis  $v_1, v_2$  hat somit die durch  $A$  vermittelte lineare Abbildung die Matrix

$$(1.1) \quad \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}.$$

Matrizen dieser Bauart werden uns alsbald bei der jordanischen Normalform begegnen.  $\diamond$

Sei  $\text{Kon}GL(2, \mathbb{C})$  die Menge der Konjugationsklassen und  $[A]$  die Klasse von  $A$ . Dann haben wir also das folgende Ergebnis erhalten:

**(6.4.2) Satz.** *Die Zuordnung  $A \mapsto (\text{Sp}(A), \det(A))$  induziert eine wohldefinierte Abbildung*

$$\kappa: \text{Kon}GL(2, \mathbb{C}) \rightarrow \mathbb{C} \times \mathbb{C}^*.$$

*Diese Abbildung ist surjektiv.*

*Genau dann hat  $A$  zwei verschiedene Eigenwerte, wenn  $\text{Sp}^2(A) \neq 4\det(A)$  ist. In diesem Fall ist  $[A]$  durch  $\kappa[A] = u$  bestimmt, und  $\kappa^{-1}(u)$  besteht aus einem Element. Die Konjugationsklasse von  $A$  enthält zwei Diagonalmatrizen; sie gehen durch Vertauschung der Diagonalelemente auseinander hervor.*

*Die Elemente  $\{(x, y) \in \mathbb{C} \times \mathbb{C}^* \mid x^2 = 4y\}$  haben bei  $\kappa$  zwei Urbilder. Diese beiden Urbilder enthalten genau eine Matrix der Form  $\text{Dia}(\lambda, \lambda)$  bzw. (1.1) und sind dadurch zu unterscheiden. Der Unterschied ist durch die Dimension des Eigenraumes gegeben (nämlich 2 bzw. 1). Der Eigenwert ist in diesem Fall  $\frac{1}{2}\text{Sp}(A)$ .  $\square$*

In den vorstehenden Überlegungen wurden die komplexen Zahlen nur dazu benutzt, um das charakteristische Polynom in Linearfaktoren zu zerlegen. Setzen wir letzteres voraus, so erhalten wir eine entsprechende Normalform für eine Matrix über einem beliebigen Körper.

**(6.4.3)  $SL(2, \mathbb{C})$ -Konjugationsklassen in  $GL(2, \mathbb{C})$ .** In diesem Fall wird also nur die Konjugation durch Matrizen der Determinante 1 erlaubt. Eine Matrix der Form  $\text{Dia}(c, c) = A(c)$  ist mit allen  $(2,2)$ -Matrizen vertauschbar. Gilt  $AXA^{-1} = Y$ , so gilt deshalb auch  $A(c)AX(A(c)A)^{-1} = Y$ . Es ist  $\det(A(c)A) = c^2 \det(A)$ . Da jede komplexe Zahl ein Quadrat ist, können wir also  $B \in SL(2, \mathbb{C})$  so finden, daß  $BXB^{-1} = Y$  ist. Mit anderen Worten: Die  $SL(2, \mathbb{C})$ - und die  $GL(2, \mathbb{C})$ -Klassen sind dieselben.  $\diamond$

**(6.4.4)  $SL(2, \mathbb{R})$ -Konjugationsklassen in  $GL(2, \mathbb{R})$ .** Erster Fall: Die Matrix  $A \in GL(2, \mathbb{R})$  habe zwei verschiedene reelle Eigenwerte  $\lambda_1, \lambda_2$ . Wir können Basisvektoren  $b_i$  in den Eigenräumen  $V(\lambda_i)$  sicherlich so wählen, daß  $\det(b_1, b_2) = 1$  ist. In der Konjugationsklasse von  $A$  liegen also genau die beiden Diagonalmatrizen  $\text{Dia}(\lambda_1, \lambda_2)$  und  $\text{Dia}(\lambda_2, \lambda_1)$ , und die Konjugationsklasse ist durch das charakteristische Polynom, also durch Spur und Determinante

bestimmt. Das charakteristische Polynom hat genau dann zwei verschiedene reelle Nullstellen, wenn

$$\operatorname{Sp}^2(A) > 4 \det(A)$$

ist.

Zweiter Fall: Die Matrix  $A$  hat einen nicht-reellen Eigenwert  $\lambda$ . Dann ist auch  $\bar{\lambda}$  ein Eigenwert, da  $\chi_A$  ein reelles Polynom ist. Dieser Fall tritt genau dann auf, wenn

$$\operatorname{Sp}^2(A) < 4 \det(A)$$

ist. Sei  $x$  ein komplexer Eigenvektor zu  $\lambda$ . Dann ist  $\bar{x}$  Eigenvektor zu  $\bar{\lambda}$ . Die Vektoren  $x$  und  $\bar{x}$  sind  $\mathbb{C}$ -linear unabhängig, da sie zu verschiedenen Eigenwerten gehören. Die Vektoren

$$c_1 = \frac{1}{\sqrt{2}}(x + \bar{x}), \quad c_2 = \frac{i}{\sqrt{2}}(x - \bar{x})$$

sind dann reelle Vektoren und ebenfalls linear unabhängig. Die Transformationsmatrix von  $(x, \bar{x})$  zu  $(c_1, c_2)$  ist

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$

Sie ist unitär. Bezüglich  $x, \bar{x}$  hat die lineare Abbildung  $A$  die Matrix  $\operatorname{Dia}(\lambda, \bar{\lambda})$ , also bezüglich  $(c_1, c_2)$  die Matrix

$$(1.2) \quad S^{-1} \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} S = \frac{1}{2} \begin{pmatrix} \lambda + \bar{\lambda} & i(\lambda - \bar{\lambda}) \\ -i(\lambda - \bar{\lambda}) & \lambda + \bar{\lambda} \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix},$$

wenn  $\lambda = \alpha + i\beta$  die Zerlegung in Real- und Imaginärteil ist. Bei komplexen Eigenwerten kann man also jedenfalls durch  $GL(2, \mathbb{R})$ -Konjugation die Normalform (1.2) erreichen. Die Eigenwerte sind  $\alpha \pm i\beta$ . Demnach sind  $\alpha$  und  $\pm\beta$  durch die Konjugationsklasse bestimmt. Wegen

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

sind die beiden Möglichkeiten konjugiert, und wir können etwa noch  $\beta > 0$  verlangen, um eine eindeutige Normalform zu bekommen. Aus einer Gleichung

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

schließt man wegen  $\beta \neq 0$  auf  $b = c$  und  $-a = d$ , also auf  $ad - bc = -a^2 - c^2 < 0$ . Die Konjugation kann also nicht mit Matrizen in  $SL(2, \mathbb{R})$  bewirkt werden.

Dritter Fall:  $A$  habe nur einen Eigenwert  $\lambda$ . Dieser ist notwendig reell, das charakteristische Polynom ist  $(x - \lambda)^2$ . Ist der Eigenraum zweidimensional, so kann man mittels  $SL(2, \mathbb{R})$ -Konjugation sicherlich die Form  $\operatorname{Dia}(\lambda, \lambda)$  erreichen. Anderfalls schließen wir wie in (6.4.1) und erhalten eine Matrix (1.1).  $\diamond$

**(6.4.5) Satz.** Die Zuordnung  $A \mapsto (\text{Sp}(A), \det A)$  induziert eine surjektive Abbildung

$$\kappa: \text{Kon } GL(2, \mathbb{R}) \rightarrow \mathbb{R} \times \mathbb{R}^*.$$

Die Menge  $\{(x, y) \mid x^2 \neq 4y\}$  hat genau ein Urbild. Für  $x^2 > 4y$  enthält eine Konjugationsklasse genau zwei Diagonalmatrizen, die sich durch Vertauschung der Diagonalelemente unterscheiden. Für  $x^2 < 4y$  enthält eine Konjugationsklasse genau zwei Matrizen der Form (1.2), die durch Transposition auseinander hervorgehen. Für  $x^2 = 4y$  gibt es zwei Urbilder bei  $\kappa$ . Diese Urbilder enthalten genau eine Matrix der Form  $\text{Dia}(\lambda, \lambda)$  bzw. (1.1).  $\square$

Sei  $\text{Kon}(SL(2, \mathbb{R}), GL(2, \mathbb{R}))$  die Menge der  $SL(2, \mathbb{R})$ -Konjugationsklassen in  $GL(2, \mathbb{R})$ . Dann haben wir wiederum eine durch  $A \mapsto (\text{Sp}(A), \det(A))$  induzierte wohldefinierte Abbildung

$$\sigma: \text{Kon}(SL(2, \mathbb{R}), GL(2, \mathbb{R})) \rightarrow \mathbb{R} \times \mathbb{R}^*.$$

Man überlegt sich:

**(6.4.6) Notiz.**  $\begin{pmatrix} \lambda & \mu \\ 0 & \lambda \end{pmatrix}$  und  $\begin{pmatrix} \lambda & \tilde{\mu} \\ 0 & \lambda \end{pmatrix}$  sind genau dann in  $SL(2, \mathbb{R})$ -konjugiert, wenn  $cd > 0$  ist, also wenn  $c$  und  $d$  dasselbe Vorzeichen haben.  $\square$

Mit diesen Vereinbarungen gilt:

**(6.4.7) Satz.** Für  $x^2 > 4y$  besteht  $\sigma^{-1}(x, y)$  aus genau einem Element, wie im letzten Satz. Für  $\text{Sp}^2 A < 4 \det A$  gibt es in der Klasse von  $A$  genau eine Matrix der Form (1.2). Die Konjugationsklassen werden durch  $(\alpha, \beta)$ ,  $\beta \neq 0$ , eindeutig parametrisiert. Die zu  $(\alpha, \beta)$  und  $(\alpha, -\beta)$  gehörenden Matrizen haben dasselbe Bild bei  $\sigma$ . Die Urbildmenge  $\sigma^{-1}(x, y)$  hat für  $x^2 < 4y$  zwei Elemente. Im Fall  $\text{Sp}^2 A = 4 \det A$  gibt es drei durch  $\sigma[A]$  bestimmte Konjugationsklassen, die durch das Enthalten der Matrizen

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & -1 \\ 0 & \lambda \end{pmatrix},$$

charakterisiert werden können. Darin ist  $\lambda$  durch  $\sigma[A]$  bestimmt.  $\square$

# Kapitel 7

## Beispiele und Anwendungen

### 7.1 Vektorprodukt

Sei  $V = \mathbb{R}^3$  mit dem Standardskalarprodukt  $\langle -, - \rangle$  versehen. Für  $x, y, z \in V$  sei  $\det(x, y, z)$  die Determinante mit den Zeilenvektoren  $x, y$  und  $z$ . Für feste  $x, y \in V$  ist  $V \rightarrow \mathbb{R}, z \mapsto \det(x, y, z)$  eine lineare Abbildung. Es gibt deshalb genau einen (von  $x$  und  $y$  abhängigen) Vektor  $w(x, y)$ , so daß diese lineare Abbildung durch  $z \mapsto \langle w(x, y), z \rangle$  gegeben wird. Wir schreiben  $w(x, y) = x \times y$  und nennen  $x \times y$  das *Vektorprodukt* des geordneten Paares  $(x, y)$  (im Gegensatz zum Skalarprodukt ist also das Resultat der Verknüpfung wiederum ein Vektor). Wir erhalten somit eine bilineare Abbildung

$$(1.1) \quad \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (x, y) \mapsto x \times y.$$

Definitionsgemäß gilt für alle  $x, y, z$  die Gleichung

$$(1.2) \quad \langle x \times y, z \rangle = \det(x, y, z).$$

Da die Determinante alternierend ist, folgt

$$(1.3) \quad x \times y = -y \times x, \quad x \times x = 0$$

für alle  $x$  und  $y$ ; die Verknüpfung ist also nicht kommutativ. Nach der Definition (1.2) folgt  $\langle x \times y, \lambda x + \mu y \rangle = 0$ . Also gilt:

**(7.1.1) Satz.** *Der Vektor  $x \times y$  ist orthogonal zu dem von  $x$  und  $y$  aufgespannten Unterraum.*  $\square$

Sei  $(x_1, x_2, x_3) = x$  und entsprechend für  $y$ . Dann gilt:

**(7.1.2) Satz.** Der Vektor  $x \times y$  hat die Komponentendarstellung

$$\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} e_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} e_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} e_3.$$

*Beweis.* Man bildet das Skalarprodukt mit  $z$  und sieht, daß die Entwicklung von

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$$

nach der dritten Zeile herauskommt.  $\square$

**(7.1.3) Korollar.**  $e_1 \times e_2 = e_3$ ,  $e_1 \times e_3 = -e_2$ ,  $e_2 \times e_3 = e_1$ .  $\square$

Jedes  $y$  liefert eine lineare Abbildung  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $z \mapsto y \times z$ . Setzen wir  $y = (a, b, c)$ , so errechnet man die Matrix dieser linearen Abbildung zu

$$(1.4) \quad \varphi(y) = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}.$$

Für diese Matrix  $A$  gilt  $A^t = -A$ . Matrizen mit dieser Eigenschaft haben wir *schiefsymmetrisch* genannt. Ersichtlich bilden die schiefsymmetrischen  $(3,3)$ -Matrizen einen dreidimensionalen Vektorraum, da sie durch die Elemente über der Hauptdiagonalen bestimmt sind. Nennen wir den Vektorraum dieser  $(3,3)$ -Matrizen  $so(3)$ , so haben wir also einen linearen Isomorphismus

$$(1.5) \quad \varphi: \mathbb{R}^3 \rightarrow so(3).$$

Wir setzen  $[A, B] = AB - BA$  für je zwei  $(n, n)$ -Matrizen  $A, B$  und nennen  $[A, B]$  die *Lie-Klammer* von  $(A, B)$ . Die Lie-Klammer liefert ebenfalls eine bilineare Verknüpfung

$$so(3) \times so(3) \rightarrow so(3), \quad (A, B) \mapsto [A, B].$$

**(7.1.4) Satz.** Die Abbildung  $\varphi: \mathbb{R}^3 \rightarrow so(3)$  erfüllt  $\varphi(x \times y) = [\varphi(x), \varphi(y)]$ .  $\square$

Das Vektorprodukt ist nicht einmal assoziativ. Stattdessen gilt die

**(7.1.5) Jacobi-Identität.**  $x \times (y \times z) + y \times (z \times x) + z \times (x \times y) = 0$  (zyklische Vertauschung von  $x, y, z$  in den Summanden).

*Beweis.* Es genügt wegen des letzten Satzes, die Identität

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$$

nachzuweisen. Es gilt  $[A[B, C]] = ABC - ACB - BCA + CBA$ . Der erste und dritte Summand sind gerade Permutationen von  $ABC$ , der zweite und vierte ungerade. Wenn wir also  $A, B$  und  $C$  zyklisch vertauschen und alles addieren, ergibt die Summe der ersten und dritten (zweiten und vierten) Summanden jeweils Null.  $\square$

**(7.1.6) Satz.** *Das Vektorprodukt hat die folgenden Eigenschaften:*

- (1)  $x \times y = 0$  genau dann, wenn  $x, y$  linear abhängig sind.
- (2) Sind  $x, y$  linear unabhängig, so ist  $x, y, x \times y$  eine positiv orientierte Basis, das heißt  $\det(x, y, x \times y) > 0$ .
- (3) Aus  $x \times y = 0$  für alle  $y$  folgt  $x = 0$ .
- (4)  $\langle x \times y, z \rangle = \langle x, y \times z \rangle$ .
- (5)  $\langle x, y \rangle^2 + |x \times y|^2 = |x|^2 |y|^2$  (Verschärfte Cauchy-Schwarzsche Ungleichung.)
- (6)  $x \times (y \times z) = \langle x, z \rangle y - \langle x, y \rangle z$  (Graßmann-Identität).
- (7)  $(a \times b) \times (c \times d) = \det(a, b, c)c - \det(a, b, d)d$ .

*Beweis.* (1) Ist  $x = \lambda y$ , so folgt  $x \times y = \lambda y \times y = 0$ . Sind  $x, y$  linear unabhängig, so gibt es eine Basis der Form  $x, y, z$ . Für diese ist  $\det(x, y, z) \neq 0$  und deshalb ist  $x \times y \neq 0$ .

(2) Ist  $x \times y \neq 0$ , also  $x, y, x \times y$  eine positiv orientierte Basis.

(3) Sei  $x \neq 0$ . Wir wählen  $x, y$  linear unabhängig. Dann ist  $x \times y \neq 0$ .

(4)  $\langle x \times y, z \rangle = \det(x, y, z) = \det(y, z, x) = \langle y \times z, x \rangle = \langle x, y \times z \rangle$ .

(6) Die Abbildungen  $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , die  $(x, y, z)$  auf  $x \times (y \times z)$  bzw. auf  $\langle x, z \rangle y - \langle x, y \rangle z$  abbilden, sind linear in jeder Variablen. Um ihre Gleichheit zu zeigen, genügt es deshalb, Vektoren der Standardbasis einzusetzen. Das ist leicht, z. B.  $e_2 \times (e_2 \times e_3) = e_2 \times e_1 = -e_3$  und  $\langle e_2, e_3 \rangle e_2 - \langle e_2, e_2 \rangle e_3 = -e_3$  etc.

(5) Nach (4) gilt  $|x \times y|^2 = \langle x \times y, x \times y \rangle = \langle x, y \times (x \times y) \rangle$ . Man setzt (6) ein und erhält das Gewünschte.

(7) Man wendet (6) an und danach (1.2). □

**(7.1.7) Satz.** *Sei  $A$  eine reelle  $(3, 3)$ -Matrix. Dann gilt*

$$A^t(Ax \times Ay) = \det(A)x \times y.$$

*Beweis.* Für alle  $z \in \mathbb{R}^3$  haben wir die Gleichungskette

$$\begin{aligned} \langle A^t(Ax \times Ay), z \rangle &= \langle Ax \times Ay, Az \rangle = \det(Ax, Ay, Az) \\ &= \det(A) \det(x, y, z) = \det(A) \langle x \times y, z \rangle. \end{aligned}$$

Es folgt die Behauptung. □

**(7.1.8) Korollar.** *Sei  $A \in SO(3)$ . Aus dem letzten Satz und  $A^t = A^{-1}$  folgt: Die Zuordnung  $x \mapsto Ax$  ist ein Isomorphismus der Verknüpfung Kreuzprodukt, das heißt es gilt  $A(x \times y) = Ax \times Ay$ . □*

Wenn man die Matrix  $A$  (1.4) quadriert, so erhält man eine Matrix, deren Diagonalelemente  $-(b^2 + c^2)$ ,  $-(a^2 + c^2)$ ,  $-(a^2 + b^2)$  sind. Folglich ist  $-\frac{1}{2}\text{Spur}(A^2) = a^2 + b^2 + c^2$ . Deshalb gilt:

**(7.1.9) Satz.**

- (1)  $(A, B) \mapsto -\frac{1}{2}\text{Sp}(AB)$  ist auf dem Vektorraum  $so(3)$  ein Skalarprodukt.
- (2) Für jedes  $U \in O(3)$  und  $A \in so(3)$  ist  $UAU^{-1} \in so(3)$ . Die lineare Abbildung  $so(3) \rightarrow so(3)$ ,  $A \mapsto UAU^{-1}$  ist orthogonal bezüglich des Skalarproduktes in (1).
- (3) Die Abbildung  $\varphi: \mathbb{R}^3 \rightarrow so(3)$  ist eine Isometrie.
- (4) Für  $U \in SO(3)$  gilt  $\varphi(Ux) = U\varphi(x)U^{-1}$ .

*Beweis.* (1), (2) und (3) ergeben sich unmittelbar aus den Definitionen und den voranstehenden Ergebnissen.

(4) Es genügt, die Gleichung für  $x = e_1$  nachzurechnen, da sie dann wegen  $\varphi(U_1U_2e_1) = U_1U_2\varphi(e_1)U_2^{-1} = U_1\varphi(U_2e_1)$  für jeden Einheitsvektor  $U_2e_1$  gilt. Wenn man für  $x = e_1$  die fraglichen Matrizen auf der rechten Seite der gewünschten Formel ausmultipliziert, muß man auch die Gleichung  $A^t = \tilde{A}$  verwenden, die für  $A \in SO(n)$  aus der Cramerschen Regel folgt.  $\square$

Sei  $B = \{b_1, b_2, b_3\}$  eine Basis des  $\mathbb{R}^3$ . Für jeden Index  $k$  wählen wir  $i, j$  so, daß  $(i, j, k)$  eine gerade Permutation ist und setzen

$$b_i \times b_j = \sum_{\ell=1}^3 \gamma_{k\ell} b_\ell.$$

Wir erhalten nach (1.2)

$$\det(b_i, b_j, b_t) = \sum_{\ell} \gamma_{k\ell} \langle b_\ell, b_t \rangle = \det(B) \delta_{kt}.$$

Sei  $g = (g_{ij})$  mit  $g_{ij} = \langle b_i, b_j \rangle$ . Dann ist also die Matrix  $\gamma = (\gamma_{k\ell})$  durch

$$\gamma = \det(B)g^{-1}$$

gegeben. Die Matrix  $\gamma$  beschreibt, wie sich das Kreuzprodukt bezüglich irgendeiner Basis ausrechnen läßt. Ist  $B$  eine positive Orthonormalbasis, so ist  $g = \gamma$  die Einheitsmatrix.

Für  $x, y \in \mathbb{R}^3$  ist  $|x \times y|$  der Flächeninhalt des von  $x$  und  $y$  aufgespannten Parallelogramms. Das läßt sich so begründen. Sind  $x$  und  $y$  linear abhängig, so sind  $|x \times y|$  und der Flächeninhalt gleich Null. Andernfalls sei  $E$  die von  $x, y$  aufgespannte Ebene und  $b_1, b_2, b_3$  eine positive Orthonormalbasis mit  $b_1, b_2 \in E$ . Sind  $x = \sum x_i b_i$ ,  $y = \sum y_i b_i$  die Komponentendarstellungen, so ist  $x_3 = y_3 = 0$  und  $x \times y = (x_1 y_2 - x_2 y_1) b_3$ . Es ist aber definitionsgemäß  $|x_1 y_2 - x_2 y_1|$  der Flächeninhalt des durch  $x$  und  $y$  aufgespannten Parallelogramms. Mittels  $\langle x, y \rangle = |x||y| \cos \alpha$  mit dem Winkel  $\alpha$  zwischen  $x$  und  $y$  folgt  $|x \times y| = |x||y| \sin \alpha$ ,  $0 \leq \alpha \leq \pi$ .

## 7.2 Quaternionen

Sei  $\mathbb{H}$  die Menge der komplexen (2,2)-Matrizen der Form

$$(2.1) \quad A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}.$$

Diese Menge ist ein vierdimensionaler reeller Vektorraum. Das Produkt zweier Matrizen der Form (2.1) hat wieder diese Form. Deshalb machen die  $\mathbb{R}$ -Vektorraumstruktur und die Multiplikation von Matrizen die Menge  $\mathbb{H}$  zu einer  $\mathbb{R}$ -Algebra, genannt die Algebra der *Quaternionen*, erfunden von Hamilton<sup>1</sup>.

Wir setzen  $A^* = \bar{A}^t$ . Die Abbildung  $\mathbb{H} \rightarrow \mathbb{H}$ ,  $A \mapsto A^*$  heißt *Konjugation* von  $\mathbb{H}$ , ist  $\mathbb{R}$ -linear und erfüllt  $(AB)^* = B^*A^*$ . Wegen der letzten Eigenschaft spricht man auch von einem *Antihomomorphismus*  $\mathbb{H} \rightarrow \mathbb{H}$ . Man hat sich daran zu erinnern, daß die Konjugation in  $\mathbb{H}$  nicht etwa der Übergang zur elementweise konjugiert-komplexen Matrix ist. Man definiert die *Norm*  $N(A)$  des Elements (2.1) als  $N(A) = |a|^2 + |b|^2 = \det(A)$ . (Hier bedeutet „Norm“ etwas anderes als die früher definierte Norm einer Matrix oder eines Vektors in einem Hilbert-Raum.) Es gilt wegen des Produktsatzes für Determinanten

$$(2.2) \quad N(A)N(B) = N(AB).$$

Ferner gilt

$$(2.3) \quad A^* \cdot A = N(A) \cdot I,$$

denn  $A^*$  ist auch die Adjunkte von  $A$ . Es ist  $N(A) = 0$  genau dann, wenn  $A$  die Nullmatrix ist. Ein von Null verschiedenes Element  $A \in \mathbb{H}$  hat deshalb das multiplikative Inverse  $A^{-1} = N(A)^{-1} \cdot A^*$ . Bis auf das Kommutativgesetz der Multiplikation erfüllt also  $\mathbb{H}$  die Körperaxiome. Deshalb nennt man  $\mathbb{H}$  einen *Schiefkörper* oder eine *Divisionsalgebra*.

Wir haben eine Einbettung

$$(2.4) \quad \iota: \mathbb{C} \rightarrow \mathbb{H}, \quad c \mapsto \begin{pmatrix} c & 0 \\ 0 & \bar{c} \end{pmatrix},$$

die  $\iota(a+b) = \iota(a) + \iota(b)$  und  $\iota(ab) = \iota(a)\iota(b)$  erfüllt. Das Bild ist also ein zu  $\mathbb{C}$  isomorpher Unterkörper. Wir fassen vermöge  $\iota$  manchmal  $\mathbb{C}$  als Teilmenge  $\mathbb{H}$  auf; man muß dabei aufpassen und die Elemente von  $\mathbb{C} \subset \mathbb{H}$  von den Koeffizienten der Matrix als komplexe Zahlen unterscheiden. Jedenfalls wird durch (2.4)  $\mathbb{H}$  zu einem zweidimensionalen  $\mathbb{C}$ -Vektorraum; die skalare Multiplikation mit  $c \in \mathbb{C}$  wird durch *Linksmultiplikation* mit der Matrix  $\iota(c)$  definiert. Da  $\mathbb{H}$  nicht

<sup>1</sup>William Rowan Hamilton 1805 – 1865

kommutativ ist, muß man auch Links- und Rechtsmultiplikation unterscheiden. Als  $\mathbb{C}$ -Vektorraum hat  $\mathbb{H}$  die *Standardbasis*

$$(2.5) \quad 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

mit den Multiplikationsregeln

$$(2.6) \quad j^2 = -1 \quad \text{und} \quad zj = j\bar{z} = \begin{pmatrix} 0 & z \\ -\bar{z} & 0 \end{pmatrix}$$

für  $z \in \mathbb{C}$ . Mit diesen Festlegungen und Identifizierungen gilt dann

$$(2.7) \quad \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} = a + bj, \quad a, b \in \mathbb{C}.$$

(Genau genommen steht also rechts in (2.7)  $\iota(a) + \iota(b)j$ .) Der komplexe Vektorraum  $\mathbb{H}$  trägt das *unitäre Skalarprodukt*

$$(2.8) \quad s(a + bj, c + dj) = a\bar{c} + b\bar{d}$$

und (2.5) ist dafür orthonormal.

Die Einbettung (2.4) liefert insbesondere die *kanonische Einbettung*  $\mathbb{R} \rightarrow \mathbb{H}$ ,  $r \mapsto r \cdot 1$ . Als reeller Vektorraum hat  $\mathbb{H}$  die *Standardbasis*

$$(2.9) \quad 1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad j = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

mit den *Multiplikationsregeln*

$$(2.10) \quad \begin{aligned} i^2 &= j^2 = k^2 = -1 \\ ij &= -ji = k \\ jk &= -kj = i \\ ki &= -ik = j. \end{aligned}$$

Es ist (2.9) eine Orthonormalbasis bezüglich des *euklidischen Skalarprodukts*  $\langle -, - \rangle$ , das durch den Realteil von (2.8) gegeben wird. Eine Quaternion kann man sich also als ein Element der Form

$$(2.11) \quad a + bi + cj + dk = q, \quad a, b, c, d \in \mathbb{R}$$

vorstellen. Quaternionen werden dann komponentenweise addiert und unter Verwendung von (2.10) multipliziert. Die Norm der Quaternion (2.11) im früher definierten Sinne ist  $a^2 + b^2 + c^2 + d^2$ . Die Elemente (2.11) mit  $a = 0$  bilden einen dreidimensionalen reellen Unterraum im  $\mathbb{H}$  von  $\mathbb{H}$ , dessen Elemente *reine Quaternionen* heißen und den wir über die Basis  $(i, j, k)$  mit  $\mathbb{R}^3$  identifizieren. Wir haben dann eine *kanonische orthogonale Zerlegung*

$$(2.12) \quad \mathbb{H} = \mathbb{R} \oplus \mathbb{R}^3, \quad \mathbb{R}^3 = \text{im } \mathbb{H}, \quad q = \text{Re}(q) + \text{im}(q) \in \mathbb{H}$$

von  $\mathbb{H}$  in den Unterkörper  $\mathbb{R}$  und den euklidischen Raum der reinen Quaternionen. Wir nennen  $\text{im}(q) = bi + cj + dk$  auch den *reinen* oder *imaginären* Teil der Quaternion (2.11).

Wir untersuchen nun Vertauschbarkeit in der Algebra  $\mathbb{H}$ . Für  $a \in \mathbb{R}$ ,  $q \in \mathbb{H}$  gilt immer  $aq = qa$ . Wir betrachten deshalb das sogenannte *Zentrum* von  $\mathbb{H}$

$$Z(\mathbb{H}) = \{z \in \mathbb{H} \mid \text{für alle } q \in \mathbb{H} \text{ gilt } zq = qz\}$$

und behaupten

**(7.2.1) Satz.**  $Z(\mathbb{H}) = \mathbb{R}$ .

Ein Beweis von (7.2.1) folgt unmittelbar aus dem nächsten Satz.

**(7.2.2) Satz.** Sei  $A \in \mathbb{H} \setminus \mathbb{R}$ ,  $X \in \mathbb{H}$  und gelte  $XA = AX$ . Dann ist  $X \in \mathbb{R} + \mathbb{R} \cdot A$ , das heißt  $X = \lambda + \mu A$  mit  $\lambda, \mu \in \mathbb{R}$ .

*Beweis.* Genau dann ist  $XA = AX$ , wenn  $X \text{im}(A) = \text{im}(A)X$  ist. Wir dürfen also  $A \in \text{im } \mathbb{H}$ ,  $A \neq 0$  voraussetzen. Mittels (2.10) errechnet man für eine reine Quaternion  $bi + cj + dk = A$

$$(2.13) \quad (bi + cj + dk)^2 = -(b^2 + c^2 + d^2) = -N(A) \in \mathbb{R}.$$

Sei  $Y = \text{im}(X)$ . Dann gilt auch  $YA = AY$ . Falls  $Y \neq 0$  ist, können wir, nach Abänderung um einen reellen Faktor, wegen (2.13) annehmen, daß  $Y^2 = A^2 = -1$  ist. Dann folgt

$$(Y - A)(Y + A) = Y^2 + YA - AY - A^2 = 0,$$

also  $Y = \pm A$ . □

Wir können auch den Teil  $\text{im } \mathbb{H}$  von  $\mathbb{H}$  durch die multiplikativen Eigenschaften von  $\mathbb{H}$  charakterisieren.

**(7.2.3) Satz.**  $A \in \text{im } \mathbb{H}$  genau dann, wenn  $A^2 \in \mathbb{R}$ ,  $A^2 \leq 0$ .

*Beweis.* Ist  $A \in \text{im } \mathbb{H}$ , so ist nach (2.13)  $A^2$  reell und nicht positiv.

Für den Beweis der Umkehrung sei  $A = r + Q$ ,  $r \in \mathbb{R}$ ,  $Q \in \text{im } \mathbb{H}$  angesetzt. Dann ist  $A^2 = r^2 + 2rQ + Q^2$  genau dann reell, wenn  $rQ = 0$  ist, also entweder wenn  $r = 0$  ist oder  $Q = 0$ . Ist  $Q = 0$ , so folgt aus  $A^2 = r^2 \leq 0$  aber auch  $r = 0$ . □

Die Anordnung in  $\mathbb{R}$  kann ebenfalls durch die Körperstruktur beschrieben werden, denn die nicht-negativen Elemente sind genau die Quadrate und  $a \leq b$  ist zu  $0 \leq b - a$  äquivalent. Folglich ist in (7.2.3) die Bedingung  $A^2 \leq 0$  durch die Ringstruktur von  $\mathbb{H}$  beschreibbar. Demnach ist auch die Zerlegung  $A = \text{Re}(A) + \text{im}(A)$  und damit die Konjugation

$$A^* = \text{Re}(A) - \text{im}(A)$$

und die euklidische Struktur ringtheoretisch (das heißt durch Aussagen über „plus“ und „mal“) bestimmt.

**(7.2.4) Satz.** *Ein Automorphismus  $\alpha$  des Körpers  $\mathbb{R}$  ist die Identität. Für einen Automorphismus  $\varphi : \mathbb{H} \rightarrow \mathbb{H}$  gelten  $\varphi(r) = r$ , falls  $r \in \mathbb{R}$ , sowie  $\varphi(*A) = \varphi(A)^*$  und  $N(\varphi(A)) = N(A)$ .*

*Beweis.* Weil  $\mathbb{R}_+ = [0, \infty[$  die Menge der Quadrate ist, gilt  $\varphi(\mathbb{R}_+) = \mathbb{R}_+$ . Ist  $a < b$ , so gilt  $b - a \in \mathbb{R}_+$  also  $\varphi(b - a) = \varphi(b) - \varphi(a) \in \mathbb{R}_+$ ; deshalb ist  $\varphi$  monoton wachsend. Aus  $\varphi(1) = 1$  folgt mittels Homomorphie  $\varphi(n) = n$  für  $n \in \mathbb{Z}$ , dann  $\varphi(r) = r$  für  $r \in \mathbb{Q}$  und dann wegen der Monotonie  $\varphi(x) = x$  für alle  $x \in \mathbb{R}$ .

Ein Automorphismus  $\varphi$  bildet das Zentrum in das Zentrum ab, also ist  $\varphi(\mathbb{R}) \subset \mathbb{R}$  und wegen der ersten Aussage  $\varphi(r) = r$  für  $r \in \mathbb{R}$ . Es folgt dann mittels (7.2.3)  $\varphi(\text{im } \mathbb{H}) = \text{im } (\mathbb{H})$ . Ist  $A = B + C$  die Zerlegung in Real- und Imaginärteil, so ist

$$\varphi(A) = \varphi(B) + \varphi(C) = B + \varphi(C)$$

also

$$\varphi(A)^* = B - \varphi(C) = \varphi(B - C) = \varphi(A^*).$$

Aus  $N(A) = A^* \cdot A \in \mathbb{R}$  folgt dann

$$\begin{aligned} N(\varphi(A)) &= \varphi(A)^* \varphi(A) = \varphi(A^*) \varphi(A) = \varphi(A^* A) \\ &= \varphi(N(A)) = N(A) \end{aligned}$$

und damit die letzte Behauptung des Satzes. □

Nach Definition der Norm ist  $N(X) = X \cdot X^* \cdot X = \langle X, X \rangle$  und folglich

$$(2.14) \quad \langle X, Y \rangle = \frac{1}{2}(X \cdot Y^* + Y \cdot X^*) = \frac{1}{2}(X^* \cdot Y + Y^* \cdot X).$$

Aus  $2\langle X^*, Y \rangle = X^* \cdot Y^* + YX$  erhält man durch Rechtsmultiplikation mit  $Y$  die

$$(7.2.5) \text{ Dreier-Identität. } YXY = 2\langle X^*, Y \rangle Y - \langle Y, Y \rangle X^*. \quad \square$$

### 7.3 Quaternionen und orthogonale Gruppen

Nach unserer Darstellung der Quaternionenalgebra  $\mathbb{H}$  durch die Matrizen der Form (2.1) ist  $SU(2) \subset \mathbb{H}$  genau die Untergruppe der Quaternionen der Norm 1. Aus der Relation  $A^t A = \det(A)I$  für  $A \in \mathbb{H}$  folgt, daß für  $\det(A) = N(A) = 1$  die Matrix  $A$  aus  $SU(2)$  ist. Ist umgekehrt  $\begin{pmatrix} a & b \\ \text{choossec} & d \end{pmatrix} \in SU(2)$ , so ist

$(c, d)$  orthogonal zu  $(a, b)$ , also ein Vielfaches vom sicherlich zu  $(a, b)$  orthogonalen Vektor  $(-\bar{b}, \bar{a})$ . Dieses Vielfache muß wegen  $ad - bc = 1$  dann  $(-\bar{b}, \bar{a})$  selbst sein. Demnach ist  $SU(2)$  als topologischer Raum die dreidimensionale Einheitssphäre. (Es ist eine bemerkenswerte Tatsache, daß nur die Sphären  $S^0, S^1, S^3$  eine stetige Gruppenmultiplikation tragen können.) Es war  $\mathbb{H}$  ein euklidischer Raum mit Skalarprodukt  $\langle -, - \rangle$ , das durch  $\langle x, x \rangle = N(x)$  für alle  $x \in \mathbb{H}$  charakterisiert war. Wir setzen wie üblich  $|x| = \langle x, x \rangle^{1/2}$  und sprechen in diesem Zusammenhang vom *Betrag* der Quaternion  $x$ .

**(7.3.1) Notiz.** Für  $A, B \in SU(2)$  sind die Abbildungen  $\mathbb{H} \rightarrow \mathbb{H}$ ,  $X \mapsto AXB$  und  $X \mapsto X^*$  orthogonal.

*Beweis.*  $|AXB| = |A||X||B| = |X|$ ,  $|X| = |X^*|$ . □

Wir werden sehen, daß auf die in (7.3.1) genannte Weise alle orthogonalen Abbildungen erhalten werden. Zur Vorbereitung dienen einige allgemeine Bemerkungen über Spiegelungen in euklidischen Räumen  $V$ ,  $\langle -, - \rangle$ . Sei  $v \in V$  und  $|v| = 1$ . Die lineare Abbildung

$$(3.1) \quad s_v: V \rightarrow V, \quad x \mapsto x - 2\langle x, v \rangle v$$

heißt *Spiegelung* zu  $v$ . Es gilt  $s_v(v) = -v$  und  $s_v(x) = x$ , falls  $\langle x, v \rangle = 0$  ist. Demnach ist  $s_v$  eine orthogonale Abbildung der Determinante  $-1$ , die als Spiegelung an der zu  $v$  orthogonalen Hyperebene angesehen werden kann.

Wir hatten die Formel (2.14)  $2\langle X, Y \rangle = X \cdot Y^* + Y \cdot X^*$ . Für  $1 \in \mathbb{H}$  liefert sie  $2\langle 1, X \rangle = X + X^*$  und deshalb

$$(3.2) \quad s_1(X) = -X^*$$

und für  $A \in SU(2)$

$$\begin{aligned} s_A s_1(X) &= -s_A(X^*) = -X^* + 2\langle A, X^* \rangle A \\ &= 2\langle A, X^* \rangle A - \langle A, A \rangle X^* = AXA = p_A(X) \end{aligned}$$

nach der Dreier-Identität (7.2.5), also

$$(3.3) \quad p_A = s_A \circ s_1.$$

Schließlich gilt

$$(3.4) \quad \begin{aligned} s_A \circ s_B &= (s_A \circ s_1) \circ (s_1 \circ s_B) \\ &= (s_A \circ s_1) \circ (s_1 \circ s_B \circ s_1^{-1}) \circ s_1 \\ &= p_A \circ p_{-B^*}. \end{aligned}$$

Für das Letztere beachte man die Gleichung  $T \circ s_v \circ T^{-1} = s_{Tv}$  für jede orthogonale Abbildung  $T: V \rightarrow V$  sowie (3.2).

Wir wollen jetzt diese Rechnungen zusammen mit der Erzeugbarkeit von orthogonalen Abbildungen durch Spiegelungen verwenden. Für  $A, B \in SU(2)$  setzen wir

$$q(A, B): \mathbb{H} \rightarrow \mathbb{H}, \quad X \mapsto A \cdot X \cdot B^*.$$

Dann gilt

$$q(A_1 B_1) q(A, B) X = A_1 (A X B^*) B_1^* = q(A_1 A, B_1 B).$$

Also ist

$$(3.5) \quad q: SU(2) \times SU(2) \rightarrow O(\mathbb{H}), \quad (A, B) \mapsto q(A, B)$$

ein Homomorphismus. Ist  $X \in \mathbb{R}$ , so auch  $A X A^* = X$ , und deshalb liefert  $q(A, A)$  eine orthogonale Abbildung

$$\pi(A): \text{im } \mathbb{H} \rightarrow \text{im } \mathbb{H}, \quad X \mapsto A X A^*,$$

und wir erhalten einen Homomorphismus

$$(3.6) \quad \pi: SU(2) \rightarrow O(\text{im } \mathbb{H}).$$

**(7.3.2) Notiz.** Die Bilder der Abbildungen (3.5) bzw. (3.6) liegen in  $SO(\mathbb{H})$  bzw.  $SO(\text{im } \mathbb{H})$ .

*Beweis.* Eine Matrix in  $SU(2)$  ist zu einer Matrix der Form  $\text{Dia}(\lambda, \lambda^{-1})$  konjugiert. Es gibt komplexe Zahlen  $\mu$ , die  $\mu^2 = \lambda$  erfüllen. Es folgt, daß jede Matrix  $A \in SU(2)$  die Form  $A = B^2$  für ein  $B \in SU(2)$  hat. Natürlich hat  $\pi(A) = \pi(B)^2$  positive Determinate. Analog wird für  $q$  geschlossen.  $\square$

**(7.3.3) Satz.** Die Homomorphismen  $q$  und  $\pi$  sind surjektiv. Es gilt  $\text{Kern } \pi = \{I, -I\}$  und  $\text{Kern } q = \{(I, I), (-I, -I)\}$ .

*Beweis.* Es gilt  $q(A, A^*) = p_A$ . Also liegen wegen (3.4) alle Elemente der Form  $s_A \circ s_B$  im Bild von  $q$ . Jedes Element von  $SO(\mathbb{H}) \cong SO(4)$  läßt sich als Produkt von vier Spiegelungen darstellen. Also ist  $q$  surjektiv.

Sei  $f \in SO(\text{im } \mathbb{H})$  gegeben. Wir setzen  $f$  zu einer Abbildung  $f \in SO(\mathbb{H})$  fort, indem wir Elemente aus  $\mathbb{R}$  identisch abbilden. Da  $q$  surjektiv ist, gibt es  $A, B \in SU(2)$  so, daß  $f = q(A, B)$ . Wegen  $f(1) = 1$  folgt aber  $AB = 1$ , also  $B = A^*$ . Daher ist  $\pi$  surjektiv.

Sei  $A \in \text{Kern } \pi$ , also  $A X A^* = X$  für alle  $X \in \text{im } \mathbb{H}$  und deshalb auch für alle  $X \in \mathbb{H}$ . Dann liegt aber wegen  $A^* = A^{-1}$  das Element  $A$  im Zentrum  $\mathbb{R}$  von  $\mathbb{H}$  und deshalb ist  $A \in SU(2) \cap \text{Zentrum} = \{I, -I\}$ .

Sei  $(A, B) \in \text{Kern } q$ , also  $A X B^* = X$  für alle  $X \in \mathbb{H}$ . Wir setzen  $X = 1$  und erkennen,  $A = B$ ,  $A \in \text{Kern } \pi$ . Demnach hat  $\text{Kern } q$  die behauptete Gestalt.  $\square$

Als leichte Folgerung erhalten wir:

**(7.3.4) Satz.** Jede Abbildung  $f \in O(\mathbb{H}) \setminus SO(\mathbb{H})$  lässt sich mit geeigneten  $A, B \in SU(2)$  in der Form  $X \mapsto A \cdot X^* \cdot B$  schreiben. Jede Abbildung  $f \in O(\text{im } \mathbb{H}) \setminus SO(\text{im } \mathbb{H})$  hat für ein geeignetes  $A \in SU(2)$  die Form  $X \mapsto -AXA^*$ .

*Beweis.* Die Abbildung  $c: \mathbb{H} \rightarrow \mathbb{H}, X \mapsto X^*$  liegt in  $O(\mathbb{H}) \setminus SO(\mathbb{H})$  und erfüllt  $c(X) = -X$  für  $X \in \text{im } \mathbb{H}$ . Man wende nun (7.3.3) auf  $f \circ c$  an.  $\square$

$\text{im}(\mathbb{H})$  ist genau der Vektorraum der *schief-hermiteschen* (2,2)-Matrizen über  $\mathbb{C}$ . Dabei heißt eine Matrix  $A$  schief-hermitesch, wenn  $A^t = -\bar{A}$  gilt. Ist  $s \in U(n)$  und  $A$  eine schief-hermitesche  $(n, n)$ -Matrix, so ist  $SAS^{-1}$  ebenfalls schief-hermitesch. Für  $S \in SU(2) \subset \mathbb{H}$  ist  $S^{-1} = S^*$ . Damit lässt sich der Homomorphismus  $\pi$  auch formulieren, ohne auf Quaternionen Bezug zu nehmen.

In der Physik werden häufig stattdessen die hermiteschen Matrizen  $A$  verwendet, die also  $A^t = \bar{A}$  erfüllen. Ist  $A$  hermitesch, so  $iA$  schief-hermitesch. So kann man die beiden Standpunkte vergleichen. Wir beschreiben die Situation für hermitesche Matrizen der Vollständigkeit und Abwechslung halber ausführlich. Der Satz (7.3.5) übersetzt ein Ergebnis aus (7.3.3).

Der weiteren Untersuchung dienen die folgenden allgemeinen Vorbemerkungen. Sei  $M_n := M(n, n; \mathbb{C})$  der Vektorraum der komplexen  $(n, n)$ -Matrizen. Durch  $(A, B) \mapsto \text{Sp}(A^t \bar{B})$  wird auf  $M_n$  eine positiv definite hermitesche Form definiert. Sei

$$H(n) = \{A \in M_n \mid A^t = \bar{A}\}$$

die Menge der hermiteschen  $(n, n)$ -Matrizen. Dann ist  $H(n)$  ein reeller Unterraum von  $M_n$ . Der Realteil  $\text{Re Sp}(A^t \bar{B})$  definiert ein Skalarprodukt auf dem reellen Vektorraum  $M_n$ . Sind  $A, B \in H(n)$ , so gilt  $\text{Sp}(A^t \bar{B}) = \text{Sp}(A^t B^t) = \text{Sp}((BA)^t) = \text{Sp}(AB)$ . Deshalb können wir auf  $H(n)$  das Skalarprodukt durch  $(A, B) \mapsto \text{Re Sp}(AB)$  erklären. Sei  $U \in U(n)$ , also  $t_U \bar{U} = I$ . Wegen

$$(UAU^{-1})^t (\overline{UBU^{-1}}) = \bar{U}^t A \bar{B} \bar{U}^{-1}$$

und der allgemeinen Relation  $\text{Sp}(XYX^{-1}) = \text{Sp}(Y)$  gilt

$$\text{Sp}((UAU^{-1})^t (\overline{UBU^{-1}})) = \text{Sp}(A^t \bar{B}).$$

Die lineare Abbildung

$$L_U: M_n \rightarrow M_n, \quad A \mapsto UAU^{-1}$$

ist deshalb unitär bezüglich der eben erklärten hermiteschen Form. Ist  $A \in H(n)$ , so gilt  $\overline{UAU^{-1}} = (UAU^{-1})^t$ , also  $UAU^{-1} \in H(n)$ . Wir haben deshalb eine induzierte lineare Abbildung  $L_U: H(n) \rightarrow H(n)$ , die bezüglich des eben definierten Skalarprodukts orthogonal ist.

Betrachten wir den Fall  $n = 2$ . Eine Matrix aus  $H(2)$  hat dann die Form

$$\begin{pmatrix} a & c \\ \bar{c} & b \end{pmatrix} \quad a, b \in \mathbb{R}, \quad c \in \mathbb{C}.$$

Also ist  $H(2)$  vierdimensional. Sei  $V \subset H(2)$  der Unterraum der Matrizen mit Spur 0, also  $-a = b$ . Dieser dreidimensionale Unterraum ist bei  $L_U$  invariant, so daß  $L_U: V \rightarrow V$  eine orthogonale Abbildung ist. Die Norm von

$$\begin{pmatrix} a & \alpha + i\beta \\ \alpha - i\beta & -a \end{pmatrix}$$

bezüglich des Skalarprodukts ist  $2(a^2 + \alpha^2 + \beta^2)$ . Es handelt sich also in dieser Koordinatendarstellung bis auf einen konstanten Faktor um das Standardskalarprodukt. Die orthogonale Gruppe von  $V$  wollen wir deshalb mit  $O(3)$  identifizieren.

**(7.3.5) Satz.** Die Zuordnung  $U \mapsto L_U$  liefert einen surjektiven Homomorphismus  $\varphi: SU(2) \rightarrow SO(3)$ . Der Kern ist  $\{I, -I\}$ .  $\square$

Als Raum (das heißt einschließlich der Stetigkeitseigenschaften) ist  $SU(2)$  die dreidimensionale Einheitssphäre  $S^3 \subset H$ , und deshalb lassen sich die Elemente von  $SU(2)$  durch Quadrupel  $(a, b, c, d) \in \mathbb{R}^4$  mit  $a^2 + b^2 + c^2 + d^2 = 1$  parametrisieren. Man kann nun ausrechnen, welche Matrix das Bild bei  $\pi$  bezüglich der Standardbasis in  $\mathbb{H}$  ist. Man erhält

$$\begin{pmatrix} a^2 + b^2 + c^2 - d^2 & -2ad + 2bc & 2ac + 2bd \\ 2ad + 2bc & a^2 - b^2 + c^2 - d^2 & -2ab + 2cd \\ -2ac + 2bd & 2ab + 2cd & a^2 - b^2 - c^2 + d^2 \end{pmatrix}.$$

Die dadurch gegebene Parametrisierung von  $SO(3)$  war schon Euler bekannt.

Für  $SO(2)$  gibt es eine analoge Parametrisierung, nämlich durch Matrizen der Form

$$\begin{pmatrix} \frac{x^2 - y^2}{x^2 + y^2} & \frac{2xy}{x^2 + y^2} \\ \frac{-2xy}{x^2 + y^2} & \frac{x^2 - y^2}{x^2 + y^2} \end{pmatrix}.$$

Diese Formel ist die Grundlage für die Transformation von Integralen über rationale Funktionen aus Sinus und Cosinus in Integrale über rationale Funktionen. Vermöge der Surjektion  $\pi: SU(2) \rightarrow SO(3)$  entsprechen die Elemente von  $SO(3)$  den Paaren  $(X, -X)$  antipodischer Vektoren  $X \in S^3$ . Diese Paare entsprechen genau den eindimensionalen Unterräumen von  $\mathbb{R}^4$ . Die Menge dieser Unterräume ist der dreidimensionale reelle projektive Raum; also kann  $SO(3)$  als dieser Raum aufgefaßt werden.

## 7.4 Quadriken

Sei  $K$  ein Körper der Charakteristik  $\neq 2$ . Wir betrachten Funktionen  $f: K^n \rightarrow K$ , die durch quadratische Polynome gegeben sind:

$$(4.1) \quad f(x) = \sum_{i,j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c$$

$x = (x_1, \dots, x_n) \in K^n$ ,  $a_{ij}, b_i$  und  $c$  aus  $K$ . Wir können  $f$  in der Form

$$f(x) = x^t Ax + Bx + c$$

schreiben, worin  $A$  die Matrix  $(a_{ij})$  ist und  $B = (b_1, \dots, b_n)$ . Die Funktion  $f$  wird nicht geändert, wenn man den Koeffizienten von  $x_i x_j$  durch  $\frac{1}{2}(a_{ij} + a_{ji})$  ersetzt. Deshalb können und wollen wir im folgenden die Matrix  $A$  als symmetrisch voraussetzen. Außerdem soll  $A$  nicht die Nullmatrix sein. Die Menge

$$M = \{x \in K^n \mid f(x) = 0\}$$

heißt die durch  $f$  bestimmte *affine Quadrik* (oder auch *Hyperfläche zweiter Ordnung, auch zweiten Grades*).

Zur geometrischen Untersuchung der Quadriken nimmt man den Standpunkt ein, daß geometrische Eigenschaften durch (geeignete) Koordinatentransformationen nicht verändert werden. Wie schon früher dargelegt, muß man auch hier unterscheiden, ob man sich durch Koordinatentransformation eine neue Funktion verschafft, deren Nullstellenquadrik man untersucht, oder ob man die Nullstellenquadrik  $M$  von  $f$  abbildet und das Bild untersucht.

Die allgemeinsten Abbildungen (bzw. Koordinatentransformationen), die wir in diesem Zusammenhang betrachten, sind die affinen:

$$\varphi: K^n \rightarrow K^n, \quad x \mapsto Sx + t,$$

$S \in GL(n, K)$ ,  $t \in K^n$ . Gelegentlich will man aber die zugelassenen Matrizen  $S$  einschränken, zum Beispiel im Fall  $K = \mathbb{R}$  nur orthogonale verwenden. Das kann wie folgt in einen allgemeinen Rahmen gestellt werden:

Sei  $G \subset GL(n, K)$  eine Untergruppe. Die Gesamtheit der affinen Abbildungen  $\varphi$  mit  $S \in G$  bildet dann bezüglich Verkettung eine Gruppe  $A(G)$ . Indem man  $\varphi$  die Matrix  $S$  zuordnet, erhält man einen surjektiven Homomorphismus  $A(G) \rightarrow G$ , dessen Kern aus der zu  $(K^n, +)$  isomorphen Untergruppe der Translationen in  $A(G)$  besteht.

Sei zunächst  $\varphi(x) = Sx$ . Die Funktion  $g(x) = f(\varphi(x)) = x^t S^t A S x + B S x + c$  gehört zur Matrix  $S^t A S$ . Wir wissen, daß  $S^t A S$  bei geeigneter Wahl von  $S$  eine Diagonalmatrix ist:  $S^t A S = \text{Dia}(\lambda_1, \lambda_2, \dots, \lambda_n)$ . Falls  $\lambda_{k+1} = \dots = \lambda_n = 0$  und  $\lambda_j \neq 0$  für  $j \leq k$  (und das können wir nach geeigneter Wahl von  $S$  annehmen), so hat  $g$  die Form

$$\begin{aligned} g(x) &= \sum_{i=1}^k \lambda_i x_i^2 + \sum_{i=1}^n d_i x_i + d \\ &= \sum_{i=1}^k \lambda_i (x_i - r_i)^2 + \sum_{i=k+1}^n d_i x_i + d'. \end{aligned}$$

(Letzteres nach geeigneter quadratischer Ergänzung). Durch Zusammensetzung mit der Translation  $(x_i) \mapsto (x_i) + (r_1, \dots, r_k, 0, \dots, 0)$  erhalten wir eine Funktion der Form

$$h(x) = \sum_{i=1}^k \lambda_i x_i^2 + \sum_{i=k+1}^n d_i x_i + d$$

(mit neuen  $d_i, d$ ). Sind  $d_{k+1}, \dots, d_n = 0$ , so hören wir auf zu denken. Andernfalls sei  $d_\ell \neq 0$  und  $\ell$  minimal mit dieser Eigenschaft. Wir wählen neue Koordinaten  $x_i = u_i$  für  $i \neq \ell$  und  $u_\ell = \sum d_i x_i$ . Die transformierte Funktion hat die Gestalt

$$q(u) = \sum_{i=1}^k \lambda_i u_i^2 + \lambda u_{k+1} + d.$$

Wir können kann noch  $\lambda_i$  um ein Quadrat abändern und  $u_{k+1}$  mit einer Konstanten  $\neq 0$  multiplizieren. Damit haben wir gewisse Normalformen von quadratischen Funktionen bei affinen Koordinatentransformationen gewonnen.

Sei nun  $K = \mathbb{R}$  der Körper der reellen Zahlen. Dann können wir zunächst nach dem Satz über Hauptachsentransformationen die Matrix  $A$  durch eine orthogonale Matrix  $S$  diagonalisieren. Sodann können wir bei dem zuletzt durchgeführten Übergang von  $x$  nach  $u$  etwas vorsichtiger sein und ein Vielfaches von  $(d_{k+1}, \dots, d_n)$  im Raum der letzten  $(n - k)$  Koordinaten als Basisvektor einer Orthonormalbasis verwenden, um diesen Übergang als orthogonale Abbildung zu erhalten. Insgesamt erhalten wir:

**(7.4.1) Satz.** *Durch Zusammensetzen mit einer Bewegung  $\varphi(x) = Sx + t$ ,  $S \in SO(n)$  läßt sich  $f$  aus (4.1) (bis auf einen Faktor  $\neq 0$ ) in eine der Funktionen*

$$\sum_{i=1}^k \lambda_i x_i^2, \quad \sum_{i=1}^k \lambda_i x_i^2 - 1, \quad \sum_{i=1}^k \lambda_i x_i^2 - 2x_{k+1}$$

überführen ( $k \geq 1$ , weil  $(a_{ij}) \neq 0$  vorausgesetzt). □

Da ein konstanter Faktor irrelevant für die Untersuchung von  $f(x) = 0$  ist und da eine Bewegung nach allgemeiner Ansicht die geometrische Gestalt nicht verändert, ist also jede affine Quadrik durch Nullsetzen einer der Funktionen in (7.4.1) gegeben. Zur qualitativen Veranschaulichung führen wir noch Koordinatenstreckungen  $\varphi(x_i) = |\lambda_i|^{-1/2} x_i$  durch. Es bleiben Gleichungen der Gestalt

$$(4.2) \quad \begin{aligned} |x|^2 - |y|^2 &= 0 \\ |x|^2 - |y|^2 &= 1 \\ |x|^2 - |y|^2 &= 2z \end{aligned} \quad x \in \mathbb{R}^r, y \in \mathbb{R}^s, z \in \mathbb{R} .$$

Hierbei haben wir in  $x$  die Komponenten mit positivem  $\lambda$ , in  $y$  die mit negativem  $\lambda$  zusammengefaßt; es darf natürlich  $r$  oder  $s$  gleich Null sein. Die Koordinaten, die in der Gleichung nicht vorkommen, sind frei wählbar.

Nun zur Veranschaulichung einige Beispiele. Wir gehen von den Gleichungen (4.2) aus.

**(7.4.2) Quadriken im  $\mathbb{R}^2$ .** Für  $n = 2$  gibt es die Fälle:

$$\begin{array}{llll}
 (1) & -y^2 & = & 1 \\
 (2) & x^2 & = & 0 & -y^2 & = & 0 \\
 (3) & x^2 & = & 1 \\
 (4) & x^2 - 2y & = & 0 & -x^2 - 2y & = & 0 \\
 (5) & x^2 - y^2 & = & 0 \\
 (6) & x^2 - y^2 & = & 1 \\
 (7) & -x^2 + y^2 & = & 0 \\
 (8) & x^2 + y^2 & = & 1.
 \end{array}$$

Die zugehörigen Nullstellenmengen (Quadriken) sind: (1) leere Menge; (2) eine Gerade („doppelt zu zählen“); (3) ein Paar paralleler Geraden  $x = 1$ ,  $x = -1$ ; (7) ein Punkt. Alle die Fälle sind mehr oder weniger degeneriert. (5) Zwei sich schneidende Geraden. (4) Parabel. (6) Hyperbel. (8) Kreis (Ellipse). Die letzten drei Fälle sind die regulären.  $\diamond$

Für  $n = 3$  sind die Quadriken, bis auf die degenerierten Fälle, Flächen im Raum. Wir verwenden  $x, y, z$  als Koordinaten. Falls  $z$  in den Gleichungen nicht vorkommt, so hat man die  $(x, y)$ -Lösungsmenge mit der  $z$ -Achse cartesisch zu multiplizieren. So liefert  $x^2 + y^2 = 1$  einen Zylinder und  $x^2 - y^2 = 0$  zwei sich schneidende Ebenen etc.

Wir betrachten deshalb nur die Fälle, in denen alle drei Unbestimmten in den Gleichungen „vorkommen“ und deren Lösungsmengen Flächen sind.

**(7.4.3) Quadriken im  $\mathbb{R}^3$ .** Es bleiben folgende Fälle übrig:

$$\begin{array}{ll}
 (1) & x^2 + y^2 - z^2 = 0 \\
 (2) & x^2 + y^2 + z^2 = 1 \\
 (3) & x^2 + y^2 - z^2 = 1 \\
 (4) & x^2 - y^2 - z^2 = 1 \\
 (5) & x^2 + y^2 - 2z = 0 \\
 (6) & x^2 - y^2 - 2z = 0
 \end{array}$$

Falls in einer solchen Gleichung  $x^2 + y^2$  vorkommt, lassen Drehungen um die  $z$ -Achse die Figur invariant. Es handelt sich um eine *Rotationsfläche*, sie entsteht, indem man die Lösungskurve in der  $(x, y)$ -Ebene um die  $z$ -Achse rotiert. Es ergeben sich die folgenden Flächen:

- (1) Doppelkegel
- (2) Kugel (Ellipsoid)
- (3) Einseitiges Hyperboloid (Rotation einer Hyperbel)
- (4) Zweiseitiges Hyperboloid (Invariant bei Drehung um  $x$ -Achse)
- (5) Rotationsparaboloid

Im Fall (6) liegt keine Rotationssymmetrie vor. In der  $(x, y)$ -Ebene hat man eine nach oben gekrümmte Parabel, in der  $(y, z)$ -Ebene eine nach unten

gekrümmte. Die entstehende Fläche heißt *Sattelfläche*. Der Nullpunkt ist der *Sattelpunkt*. Die beiden Geraden  $x = y, z = 0$  und  $x = -y, z = 0$  liegen in der Fläche und gehen durch den Sattelpunkt. Schneidet man die Sattelfläche in der Höhe  $z$  parallel zur  $(x, y)$ -Ebene, so entstehen als *Höhenlinien* Hyperbeln (für  $z \neq 0$ ) oder die beiden genannten Geraden.

Trotz ihrer augenfälligen Krümmung enthalten einige dieser Flächen ganze Scharen von Geraden.

Die Sattelfläche  $z = x^2 - y^2 = (x + y)(x - y)$  enthält die Geradenschar I und II:

$$\begin{array}{ll} \text{I} & x + y = a, \quad a(x - y) = z \\ \text{II} & x - y = a, \quad a(x + y) = z \end{array}$$

Für festes  $a \in \mathbb{R}$  definieren die beiden Gleichungen in I eine Gerade im  $\mathbb{R}^3$ , denn die Matrix

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

hat immer den Rang 2. Und durch Multiplikation dieser Gleichungen erkennt man, daß die Geraden in der Fläche enthalten sind. Analog im Fall II. Für verschiedene  $a \in \mathbb{R}$  sind die zugehörigen Geraden disjunkt, denn zwei Gleichungen  $x + y = a, x + y = b$  können für  $a \neq b$  nicht gleichzeitig bestehen. Jeder Punkt der Fläche liegt auf genau einer der Geraden in jeder Schar, denn der Parameter  $a$  der Schar läßt sich aus den Koordinaten des Punktes berechnen. Jede Gerade der ersten Schar schneidet jede Gerade der zweiten in genau einem Punkt, denn  $x + y = a, x - y = b$  hat immer genau eine Lösung und, einmal gelöst, errechnet sich der zugehörige  $z$ -Wert. Man kann die Punkte der Fläche also durch die zugehörigen Parameter  $(a, b)$  der (ersten, zweiten) Schar beschreiben. Die Fläche läßt sich durch diese Koordinatisierung umkehrbar eindeutig (und stetig) auf  $\mathbb{R} \times \mathbb{R}$  abbilden.

In ähnlicher Weise läßt sich das einschalige Hyperboloid durch zwei Geradenscharen beschreiben. Wir beschreiben es durch die Gleichung

$$(x + z)(x - z) = x^2 - z^2 = 1 - y^2 = (1 - y)(1 + y).$$

Es enthält jedenfalls die beiden Geraden

$$\begin{array}{ll} \text{I} & x + z = 1 - y, \quad x - z = 1 + y \\ \text{II} & x + z = 1 + y, \quad x - z = 1 - y \end{array}$$

Aus jeder dieser Geraden geht durch Rotation um die  $z$ -Achse eine Gerade hervor, die auf dem Hyperboloid liegt. Rechnerisch kann man durch Einfügen von Parametern in diese Aufteilung der Faktoren ähnlich wie bei der Sattelfläche vorgehen. Betrachten wir die zu I gehörende Schar.

Für  $(\lambda, \mu) \in \mathbb{R}^2 \setminus 0$  liefern die beiden Gleichungen

$$\lambda(x + z) = \mu(1 - y), \quad \mu(x - z) = \lambda(1 + y)$$

eine Gerade, denn die Matrix

$$\begin{pmatrix} \lambda & \mu & \lambda \\ \mu & -\lambda & -\mu \end{pmatrix}$$

hat immer den Rang 2. Natürlich liefern  $(\lambda, \mu)$  und  $(\alpha\lambda, \alpha\mu)$  für  $\alpha \neq 0$  dieselbe Gerade. Für  $\lambda\mu \neq 0$  sieht man durch Multiplikation der Gleichungen sofort, daß die Geraden in der Fläche enthalten sind. Aber auch für etwa  $\mu = 0$  ist die Gerade  $x + z = 0$ ,  $1 + y = 0$  in der Fläche enthalten. Der Schnittpunkt der  $(\lambda, \mu)$ -Gerade mit der  $(x, y)$ -Ebene ist der Punkt

$$(x, y) = \left( \frac{2\lambda\mu}{\mu^2 + \lambda^2}, \frac{\mu^2 - \lambda^2}{\mu^2 + \lambda^2} \right)$$

auf dem Einheitskreis. Es liefern  $(\lambda, \mu)$  und  $(\lambda', \mu')$  genau dann denselben Punkt auf dem Einheitskreis, wenn  $\alpha(\lambda, \mu) = (\lambda', \mu')$  für ein  $\alpha \neq 0$  gilt. Demnach ist die Gerade der Schar durch diesen Punkt bestimmt. Ähnlich wie bei der Sattelfläche folgt nun: Jeder Punkt der Fläche liegt auf genau einer Geraden der Schar, denn  $(\lambda, \mu)$  kann bis auf einen konstanten Faktor durch die Koordinaten  $(x, y, z)$  aus den Schargleichungen bestimmt werden.

Zu jeder Geraden der einen Schar gibt es eine der anderen, die dazu parallel ist. Solche Ausnahmefälle kann man nur durch Hinzufügen unendlich ferner Punkte beseitigen.

## 7.5 Quadratische Formen

Sei  $V$  ein Vektorraum über dem Körper  $K$ . Eine Abbildung  $q: V \rightarrow K$  heißt *quadratische Form* auf  $V$ , wenn sie die folgenden Eigenschaften hat:

- (1) Für  $\lambda \in K$  und  $x \in V$  gilt  $q(\lambda x) = \lambda^2 q(x)$ .
- (2) Die Abbildung

$$s_q: V \times V \rightarrow K, \quad (x, y) \mapsto q(x + y) - q(x) - q(y)$$

ist bilinear.

Die quadratischen Formen auf  $V$  bilden einen Vektorraum  $QF(V)$ , als Teilraum von  $\text{Abb}(V, K)$ . Ein Paar  $(V, q)$  mit einer quadratischen Form  $q$  auf  $V$  heißt *quadratischer Raum*. Die zu  $q$  gehörende Form  $s_q$  heißt *Polarisierung* von  $q$ .

Ist  $s: V \times V \rightarrow K$  eine Bilinearform, so ist  $q_s: v \mapsto s(v, v)$  eine quadratische Form. Sei  $BF(V)$  der Vektorraum der Bilinearformen auf  $V$ . Durch  $s \mapsto q_s$  erhalten wir eine lineare Abbildung

$$\Omega: BF(V) \rightarrow QF(V).$$

Durch  $s^t(x, y) = s(y, x)$  wird die zu  $s$  *transponierte* Bilinearform erklärt.

Transposition ist eine lineare Abbildung

$$\tau: BF(V) \rightarrow BF(V), s \mapsto s^t$$

mit der Eigenschaft  $\tau^2 = \text{id}$  (Involution). Die Formen  $s$  und  $s^t$  haben bei  $\Omega$  dasselbe Bild.

Wir bezeichnen noch mit  $SF(V) \subset BF(V)$  den Unterraum der symmetrischen Bilinearformen.

**(7.5.1) Satz.** *Sei  $V$  ein endlichdimensionaler Vektorraum über  $K$ . Die Sequenz*

$$0 \longrightarrow SF(V) \xrightarrow{\subset} BF(V) \xrightarrow{1-\tau} BF(V) \xrightarrow{\Omega} QF(V) \longrightarrow 0$$

*ist exakt.*

*Beweis.* Die Exaktheit an den Stellen  $SF(V)$  und dem linken  $BF(V)$  folgt unmittelbar aus den Definitionen; ebenso die Relation  $\Omega \circ (1 - \tau) = 0$ .

Für den Beweis der Surjektivität von  $\Omega$  leiten wir eine Matrixbeschreibung einer quadratischen Form  $q$  mit zugehöriger Bilinearform  $s = s_q$  her. Seien  $v_1, \dots, v_n \in V$  und  $\lambda_1, \dots, \lambda_n \in K$ . Dann folgt aus den definierenden Eigenschaften (1), (2) einer quadratischen Form durch Induktion nach  $n$

$$q\left(\sum_{j=1}^n \lambda_j v_j\right) = \sum_{j=1}^n \lambda_j^2 q(v_j) + \sum_{i < j} \lambda_i \lambda_j s(v_i, v_j).$$

Ist  $(v_1, \dots, v_n)$  eine Basis von  $V$ , so betrachten wir zu gegebenem  $q$  die Bilinearform  $s$  mit der Matrix  $S = (a_{i,j})$

$$\begin{aligned} a_{ii} &= q(v_i) \\ a_{ij} &= s_q(v_i, v_j) \quad i < j \\ a_{ij} &= 0 \quad i > j. \end{aligned}$$

Dann ist nach der obigen Formel  $q = q_s$ . Folglich ist  $\Omega$  surjektiv. Angenommen,  $\Omega s = 0$ , d. h. für alle  $x \in V$  gilt  $s(x, x) = 0$ . Daraus folgt  $s(x, y) = -s(y, x)$  für alle  $x, y \in V$ . Definieren wir eine Bilinearform  $t$  durch die Matrix

$$\begin{aligned} t(v_i, v_i) &= 0 \\ t(v_i, v_j) &= s(v_i, v_j) \quad i < j \\ t(v_i, v_j) &= 0 \quad i > j, \end{aligned}$$

so ist  $(1 - \tau)(t) = s$ . Damit sehen wir auch die Exaktheit am rechten  $BF(V)$ .  $\square$

Das Bild von  $1 - \tau$  in  $BF(V)$  besteht aus den *alternierenden* Bilinearformen.

Oft werden quadratische Formen und symmetrische Bilinearformen miteinander verglichen. Die linearen Abbildungen

$$\Phi: QF(V) \rightarrow SF(V), q \mapsto s_q$$

$$\Psi: SF(V) \rightarrow QF(V), \quad s \mapsto q_s$$

erfüllen die Relationen

$$\Psi\Phi(q) = 2q, \quad \Phi\Psi(s) = 2s.$$

Hat  $K$  nicht die Charakteristik 2, so sind  $\Psi$  und  $\Phi$  Isomorphismen. Es besteht dann kein wesentlicher Unterschied zwischen quadratischen Formen und symmetrischen Bilinearformen. Im allgemeinen sind aber  $\Phi$  und  $\Psi$  weder injektiv noch surjektiv. Gilt  $\Phi(q) = s$ , so nennt man  $q$  eine *quadratische Verfeinerung* von  $s$ .

## 7.6 Reelle und komplexe Strukturen

Wir untersuchen in diesem Abschnitt den Zusammenhang zwischen Vektorräumen über  $\mathbb{R}$  und  $\mathbb{C}$ .

Sei  $V$  ein  $\mathbb{C}$ -Vektorraum. Indem wir dieselbe Menge und Vektoraddition verwenden und die gegebene Skalarmultiplikation nur für Zahlen  $\lambda \in \mathbb{R}$  verwenden, erhalten wir einen  $\mathbb{R}$ -Vektorraum, der mit  $V_{\mathbb{R}}$  bezeichnet werde. (Man sagt,  $V_{\mathbb{R}}$  entsteht aus  $V$  durch „Vergessen“ von Struktur.) Eine  $\mathbb{C}$ -lineare Abbildung  $f: V \rightarrow W$  induziert dann eine  $\mathbb{R}$ -lineare Abbildung  $f_{\mathbb{R}}: V_{\mathbb{R}} \rightarrow W_{\mathbb{R}}$ ; es handelt sich um dieselbe Abbildung der zugrundeliegenden Mengen.

**(7.6.1) Satz.** *Sei  $v_1, \dots, v_n$  eine  $\mathbb{C}$ -Basis von  $V$ . Dann ist  $v_1, iv_1, \dots, v_n, iv_n$  eine  $\mathbb{R}$ -Basis von  $V_{\mathbb{R}}$ .*

*Beweis.* Sei  $v \in V$ ,  $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ ,  $\lambda_j = a_j + ib_j$  mit  $a_j, b_j \in \mathbb{R}$ . Dann ist  $v = a_1 v_1 + b_1 (iv_1) + \dots + a_n v_n + b_n (iv_n)$ , also bilden die genannten Elemente ein Erzeugendensystem von  $V_{\mathbb{R}}$ . Ist  $v$  in dieser Darstellung gleich Null, so schließt man aus der linearen Unabhängigkeit der  $v_1, \dots, v_n$  auf  $a_j = b_j = 0$  für alle  $j$ . Deshalb sind die genannten Elemente auch linear unabhängig.  $\square$

Die Aussage (7.6.1) über die Dimensionen schreiben wir so:

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

Ist  $U$  ein  $\mathbb{R}$ -Vektorraum, so heißt eine lineare Abbildung  $J: U \rightarrow U$  mit  $J^2 = -\text{id}$  eine *komplexe Struktur* auf  $U$ . Wir können mittels  $J$  nämlich  $U$  zu einem  $\mathbb{C}$ -Vektorraum machen, indem wir dieselbe Vektoraddition verwenden, und die Skalarmultiplikation für  $\lambda = a + bi$  mit  $a, b \in \mathbb{R}$  und  $u \in U$  durch

$$\lambda u = au + bJu$$

definieren. Wegen  $J^2 = -\text{id}$  ist dann  $J: U \rightarrow U$  in dieser neuen Vektorraumstruktur die Multiplikation mit  $i$ . Sei  $(U, J)$  der so erhaltene (von  $J$  abhängige)  $\mathbb{C}$ -Vektorraum. Es gilt nach Konstruktion

$$(U, J)_{\mathbb{R}} = U.$$

Eine komplexe Struktur kann es auf  $U$  natürlich nur dann geben, wenn  $\dim_{\mathbb{R}} U$  gerade ist.

Ist  $V$  ein  $\mathbb{C}$ -Vektorraum, so trägt  $U = V_{\mathbb{R}}$  eine komplexe Struktur  $J$ , die durch  $Ju = iu$  gegeben wird, und es ist  $V = (U, J)$  nach Konstruktion. Sind  $(U_1, J_1)$  und  $(U_2, J_2)$   $\mathbb{R}$ -Vektorräume mit komplexer Struktur und ist  $f: U_1 \rightarrow U_2$  eine  $\mathbb{R}$ -lineare Abbildung, die  $J_2 f = f J_1$  erfüllt, so ist dieselbe Mengenabbildung eine  $\mathbb{C}$ -lineare Abbildung

$$f: (U_1, J_1) \rightarrow (U_2, J_2).$$

Ist  $V$  ein  $\mathbb{C}$ -Vektorraum, so gehört dazu der *konjugierte*  $\mathbb{C}$ -Vektorraum  $\bar{V}$ , der dieselbe Vektoraddition wie  $V$  hat, dagegen die neue Skalarmultiplikation

$$(\lambda, v) \mapsto \bar{\lambda}v.$$

Der zu  $(U, J)$  konjugierte Vektorraum ist  $(U, -J)$ .

Ist  $U$  ein  $\mathbb{R}$ -Vektorraum, so können wir auf  $V = U \oplus U$  eine komplexe Struktur durch  $J(u_1, u_2) = (-u_2, u_1)$  definieren, denn offenbar ist  $J^2(u_1, u_2) = (-u_1, -u_2)$ . Den entstehenden  $\mathbb{C}$ -Vektorraum  $(U \oplus U, J)$  bezeichnen wir mit  $U_{\mathbb{C}}$  und nennen ihn die *Komplexifizierung* von  $U$ . Ist  $f: U \rightarrow U'$  eine  $\mathbb{R}$ -lineare Abbildung, so erfüllt

$$f \oplus f: U \oplus U \rightarrow U' \oplus U'$$

$$\begin{aligned} (f \oplus f)J(u_1, u_2) &= (f \oplus f)(-u_2, u_1) = (-fu_2, fu_1) \\ &= J(fu_1, fu_2) = J(f \oplus f)(u_1, u_2) \end{aligned}$$

und ist deshalb eine  $\mathbb{C}$ -lineare Abbildung  $(U \oplus U, J) \rightarrow (U' \oplus U', J)$ , die wir mit

$$f_{\mathbb{C}}: U_{\mathbb{C}} \rightarrow U'_{\mathbb{C}}$$

bezeichnen. Es gilt  $\text{id}_{\mathbb{C}} = \text{id}$  und  $(f \circ g)_{\mathbb{C}} = f_{\mathbb{C}} \circ g_{\mathbb{C}}$ . Es gilt

$$\dim_{\mathbb{R}} U = \dim_{\mathbb{C}} U_{\mathbb{C}}.$$

Ferner gilt  $(U_{\mathbb{C}})_{\mathbb{R}} = U \oplus U$ .

Sei  $V$  ein  $\mathbb{C}$ -Vektorraum. Eine Abbildung  $f: V \rightarrow V$  heißt *konjugiert-linear*, wenn  $f_{\mathbb{R}}$  linear ist und wenn für  $\lambda \in \mathbb{C}$  immer  $f(\lambda x) = \bar{\lambda}f(x)$  gilt. Eine konjugiert-lineare Abbildung ist also dasselbe wie eine  $\mathbb{C}$ -lineare Abbildung  $V \rightarrow \bar{V}$ . Eine konjugiert lineare Abbildung  $c: V \rightarrow V$  heißt *reelle Struktur* auf

$V$ , wenn  $c^2 = \text{id}$  ist. Ist  $V = \mathbb{C}^n$ , so ist  $c(\lambda_1, \dots, \lambda_n) = (\bar{\lambda}_1, \dots, \bar{\lambda}_n)$  eine reelle Struktur. Unter einer reellen Struktur soll man sich also den Übergang zum Konjugiertkomplexen vorstellen.

Sei  $c$  eine reelle Struktur auf  $V$ . Wir betrachten die Eigenräume von  $c$  zu den Eigenwerten  $\pm 1$ :

$$V_+ = \{x \in V \mid cx = x\}, \quad V_- = \{x \in V \mid cx = -x\}.$$

Da  $c$  konjugiert-linear ist, sind  $V_+$  und  $V_-$  jedenfalls reelle Unterräume von  $V$ , das heißt Unterräume von  $V_{\mathbb{R}}$ .

**(7.6.2) Satz.** *Es gilt  $V_{\mathbb{R}} = V_+ \oplus V_-$ . Die Multiplikation mit  $i \in \mathbb{C}$  liefert Isomorphismen  $V_+ \rightarrow V_-$ ,  $V_- \rightarrow V_+$  von  $\mathbb{R}$ -Vektorräumen.*

*Beweis.* Sicherlich ist  $V_+ \cap V_- = \{0\}$ , wie immer bei verschiedenen Eigenwerten. Es ist für jedes  $v$

$$v_+ = \frac{1}{2}(v + cv) \in V_+, \quad v_- = \frac{1}{2}(v - cv) \in V_-,$$

wie man mittels  $c^2 = \text{id}$  nachrechnet. Wegen  $v = v_+ + v_-$  ist also auch  $V = V_+ + V_-$ . Damit ist die erste Behauptung gezeigt. Sei  $v \in V_+$ . Da  $c$  konjugiert-linear ist, gilt  $c(iv) = \bar{i}c(v) = -iv$ , also  $iv \in V_-$ . Analog ist für  $v \in V_-$  auch  $cv \in V_+$ . Die Abbildung  $V_+ \rightarrow V_-$ ,  $v \mapsto iv$  ist ein Isomorphismus mit Inversem  $w \mapsto -iw$ .  $\square$

Sei  $U$  ein  $\mathbb{R}$ -Vektorraum. Dann hat  $U_{\mathbb{C}}$  eine kanonische reelle Struktur, die durch  $c(u_1, u_2) = (u_1, -u_2)$  auf  $U \oplus U$  gegeben ist. Mit der oben erklärten komplexen Struktur  $J$  gilt nämlich  $cJ = -Jc$ , weshalb  $c$  als konjugiert-linear erkannt wird. Die Eigenräume  $(U_{\mathbb{C}})_+$  und  $(U_{\mathbb{C}})_-$  sind hier  $U \oplus \{0\}$  und  $\{0\} \oplus U$ . Das legt die folgende Überlegung nahe.

**(7.6.3) Satz.** *Sei  $V$  ein  $\mathbb{C}$ -Vektorraum mit reeller Struktur  $c$ . Dann gibt es einen kanonischen Isomorphismus von  $\mathbb{C}$ -Vektorräumen*

$$\alpha: (V_+)_{\mathbb{C}} \rightarrow V,$$

der die eben auf  $(V_+)_{\mathbb{C}}$  definierte reelle Struktur in die gegebene Struktur  $c$  auf  $V$  übersetzt.

*Beweis.* Wir setzen  $\alpha(v_1, v_2) = v_1 + iv_2$ , für  $v_1, v_2 \in V_+$ . Dann ist  $iv_2 \in V_-$  und  $\alpha$  offenbar ein  $\mathbb{R}$ -linearer Isomorphismus (wegen (7.6.2)). Es ist  $\alpha$  aber  $\mathbb{C}$ -linear, weil

$$\alpha J(v_1, v_2) = \alpha(-v_2, v_1) = -v_2 + iv_1 = i(v_1 + iv_2) = i\alpha(v_1, v_2)$$

ist. Es gilt  $\alpha(v_1, -v_2) = v_1 - iv_2 = cv_1 - icv_2 = cv_1 + c(iv_2) = c(v_1 + iv_2) = c\alpha(v_1, v_2)$ . Demnach ist  $\alpha$  wie behauptet mit den reellen Strukturen verträglich.  $\square$

**(7.6.4) Satz.** Für jeden komplexen Vektorraum  $V$  gibt es einen kanonischen Isomorphismus

$$(V_{\mathbb{R}})_{\mathbb{C}} \cong V \oplus \bar{V}.$$

*Beweis.* Die Abbildung

$$\beta: (V_{\mathbb{R}})_{\mathbb{C}} \rightarrow V \oplus \bar{V}, \quad (v_1, v_2) \mapsto (v_1 + iv_2, v_1 - iv_2)$$

ist  $\mathbb{R}$ -linear und offenbar injektiv, also ein  $\mathbb{R}$ -Isomorphismus. Man rechnet nach, daß sie mit den jeweiligen komplexen Strukturen verträglich ist.  $\square$

**(7.6.5) Bemerkung.** Vektorräume gleicher Dimension über einem Körper sind isomorph. Im allgemeinen kann man aber einen Isomorphismus nur durch Auswahl von Basen herstellen. Gibt es einen ausgezeichneten, basisfrei definierten Isomorphismus, so nennt man diesen *kanonisch*. Die Räume  $V$  und  $\bar{V}$  (bzw.  $V$  und  $V^*$ ) sind isomorph, aber *nicht* kanonisch.  $\diamond$

Ist  $f: V \rightarrow V'$  eine  $\mathbb{C}$ -lineare Abbildung, sind  $c, c'$  reelle Strukturen auf  $V, V'$  und gilt  $fc = c'f$ , so induziert  $f$  Abbildungen  $f_+: V_+ \rightarrow V'_+$ ,  $f_-: V_- \rightarrow V'_-$ , die  $\mathbb{R}$ -linear sind. Bezüglich der Isomorphismen in (7.6.3) wird  $(f_+)_{\mathbb{C}}$  in  $f$  überführt. Ein Unterraum  $W$  von  $V$  heißt *c-reell*, wenn  $cW \subset W$  ist. Es induziert dann  $c$  eine reelle Struktur auf  $W$ , und es gilt  $W_+ = W \cap V_+$ .

Wir geben nun noch einen zweiten Beweis für den Normalformensatz orthogonaler Matrizen, der darauf beruht, eine orthogonale Abbildung zunächst als unitäre aufzufassen, diese dann zu diagonalisieren und das Ergebnis ins Reelle zurückzurechnen. Zunächst behandeln wir den allgemeinen Fall von Eigenraumzerlegungen.

Endomorphismen reeller Vektorräume haben nicht immer Eigenwerte. In diesem Fall betrachten wir zunächst die Komplexifizierung, zerlegen sie in Eigenräume und rechnen danach auf das Reelle zurück. Wir behandeln hier nur den Fall, in dem eine Eigenraumzerlegung existiert.

**(7.6.6) Satz.** Sei  $V$  ein  $\mathbb{C}$ -Vektorraum mit reeller Struktur  $c$ . Sei  $f: V \rightarrow V$  eine  $\mathbb{C}$ -lineare Abbildung, die  $fc = cf$  erfüllt. Sei  $\lambda$  Eigenwert von  $f$  und  $V(\lambda)$  der zugehörige Eigenraum. Sei  $U = V_+$  und  $h = f_+: U \rightarrow U$ . Dann gilt:

- (1) Ist  $\lambda$  reell, so ist  $V(\lambda)$  ein  $c$ -reeller Unterraum von  $V$ ; es ist  $\lambda$  Eigenwert von  $h$  und  $U(h, \lambda) = V(\lambda)_+$ .
- (2) Ist  $\lambda$  nicht reell, so ist auch  $\bar{\lambda}$  Eigenwert. Es ist  $V(\lambda) \oplus V(\bar{\lambda})$   $c$ -reeller Unterraum von  $V$ . Es induziert  $c$  einen  $\mathbb{R}$ -linearen Isomorphismus  $V(\lambda) \rightarrow V(\bar{\lambda})$ . Der Unterraum  $(V(\lambda) \oplus V(\bar{\lambda}))_+ =: U\{\lambda, \bar{\lambda}\} \subset U$  wird durch  $h$  in sich abgebildet.
- (3) Ist  $V$  direkte Summe der  $V(\lambda_1), \dots, V(\lambda_r), V(\mu_1), V(\bar{\mu}_1), \dots, V(\mu_s), V(\bar{\mu}_s)$  mit reellen  $\lambda_j$  und nicht-reellen  $\mu_j$ , so ist  $U$  direkte Summe der

$$U(\lambda_1), \dots, U(\lambda_r), U\{\mu_1, \bar{\mu}_1\}, \dots, U\{\mu_s, \bar{\mu}_s\}.$$

*Beweis.* Sei  $x \in V(\lambda)$ . Wegen  $f(c(x)) = cf(x) = c(\lambda x) = \bar{\lambda}c(x)$  ist also  $c(x) \in V(\bar{\lambda})$ . Damit folgt (1) und (2). Ist allgemein  $V = V' \oplus V''$  eine direkte Zerlegung in  $c$ -reelle Unterräume, so wird eine direkte Zerlegung  $V_+ = V'_+ \oplus V''_+$  induziert. Damit folgt (3).  $\square$

Sei  $V$  direkte Summe von Eigenräumen. Ist  $b_1, \dots, b_t$  Basis von  $V(\lambda)$ , so ist  $c(b_1), \dots, c(b_t)$  Basis von  $V(\bar{\lambda})$ . Wir können deshalb eine Zerlegung

$$V = V(1) \oplus \dots \oplus V(k) \oplus W(1) \oplus \dots \oplus W(\ell)$$

in reelle Unterräume finden, die die folgenden Eigenschaften hat:

- (1) Die  $V(1), \dots, V(k)$  sind eindimensional. Jedes  $V(j)$  ist in einem Eigenraum zum reellen Eigenwert  $\lambda_j$  enthalten.
- (2) Die  $W(1), \dots, W(\ell)$  sind zweidimensional. Jedes  $W(j)$  hat eine Basis der Form  $b_j, cb_j$ , mit  $f(b_j) = \mu_j b_j$  und  $\mu_j \notin \mathbb{R}$ .

Der Raum  $U = V_+$  zerfällt dann dementsprechend nach (7.6.6)

$$U = V(1)_+ \oplus \dots \oplus W(\ell)_+$$

in reell ein- bzw. zweidimensionale Unterräume, die durch  $h = f_+$  in sich abgebildet werden. Es hat  $W(j)_+$  die Basis

$$\frac{1}{\sqrt{2}}(b_j + cb_j) = a_j(1), \quad \frac{1}{\sqrt{2}}(ib_j - icb_j) = a_j(2)$$

und bezüglich dieser Basis hat  $h : W(j)_+ \rightarrow W(j)_+$  die Matrix

$$\begin{pmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{pmatrix}, \quad \alpha_j = \operatorname{Re} \mu_j, \quad \beta_j = \operatorname{Im} \mu_j.$$

Der Basiswechsel von  $a(1), a(2)$  nach  $b, cb$  hat die Matrix

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = S.$$

Das ist eine unitäre Matrix.

Wir übersetzen jetzt die voranstehenden Ergebnisse in matrizentheoretische Aussagen. Sei  $A$  eine reelle  $(n, n)$ -Matrix. Wir fassen  $A$  als Endomorphismus  $h$  von  $U = \mathbb{R}^n$  und als Endomorphismus  $f$  von  $V = \mathbb{C}^n$  auf. Es trägt  $V$  die reelle Struktur  $c(z_1, \dots, z_n) = (\bar{z}_1, \dots, \bar{z}_n)$ , und es ist  $V_+ = U = \mathbb{R}^n$  der Unterraum der reellen  $n$ -Tupel. Es ist dann  $h = f_+$  im eben verwendeten Sinn. Angenommen  $V$  besitzt eine direkte Zerlegung in Eigenräume von  $f$ . Dann finden wir nach den voranstehenden Überlegungen eine Basis von  $\mathbb{R}^n$ , so daß  $h$  darin eine Matrix der Blockdiagonalform

$$\operatorname{Dia}(\lambda_1, \dots, \lambda_r, M_1, \dots, M_\ell)$$

hat, worin die  $\lambda_j$  die reellen Eigenwerte sind und

$$M_j = \begin{pmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{pmatrix}$$

ist. Mit anderen Worten: Es gibt  $S \in GL(n, \mathbb{R})$ , so daß  $SAS^{-1}$  diese Form hat.

Wir kommen nun zu den Anwendungen auf orthogonale Abbildungen. Sei  $\beta : V \times V \rightarrow \mathbb{C}$  eine hermitesche Form auf einem  $\mathbb{C}$ -Vektorraum  $V$ . Dann ist  $s_\beta = s : V_{\mathbb{R}} \times V_{\mathbb{R}} \rightarrow \mathbb{R}$ ,  $(u, v) \mapsto \operatorname{Re} \beta(u, v)$  eine symmetrische Bilinearform. Ist  $\beta$  regulär, so auch  $s_\beta$ . Ist  $\beta$  positiv definit, so auch  $s_\beta$ . Ist  $v_1, \dots, v_n$  eine Orthonormalbasis bezüglich  $\beta$ , so ist  $v_1, iv_1, \dots, v_n, iv_n$  eine Orthonormalbasis bezüglich  $s$ ; denn  $\beta(v_j, iv_j) = -i\beta(v_j, v_j)$  hat den Realteil 0. Ist  $f : V \rightarrow V$  unitär bezüglich  $\beta$ , so ist  $f_{\mathbb{R}}$  orthogonal bezüglich  $s_\beta$ . Deshalb wird ein Homomorphismus

$$r : U(n) \rightarrow SO(2n)$$

induziert. Für  $n = 1$  erhält man speziell einen Isomorphismus

$$r : U(1) \rightarrow SO(2), \quad e^{i\varphi} \mapsto \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Sei nun  $A \in O(n)$ . Wie eben fassen wir  $A$  als Endomorphismus von  $\mathbb{C}^n$  mit reeller Struktur  $c(z_1, \dots, z_n) = (\bar{z}_1, \dots, \bar{z}_n)$  auf. Wir können  $A$  auch als unitäre Matrix ansehen. Es gibt deshalb eine Orthonormalbasis des  $\mathbb{C}^n$  bezüglich der hermiteschen Standardform aus Eigenvektoren  $a_1, \dots, a_k, b_1, cb_1, \dots, b_\ell, cb_\ell$  von  $A$ . Die daraus nach dem eben beschriebenen Verfahren gebildete Basis

$$a_1, \dots, a_k, a_1(1), a_2(2), \dots, a_\ell(1), a_\ell(2)$$

ist eine Orthonormalbasis des  $\mathbb{R}^n$  bezüglich des Standardskalarprodukts. Die Kästchen sind orthogonal, also aus  $SO(2)$ . Damit haben wir erneut den Normalformensatz für orthogonale Matrizen hergeleitet.

# Kapitel 8

## Anhang

### 8.1 Die Mengensprache

Mathematik ist Sprache. Sprache wird vereinbart. Vereinbart wird das Zweckmäßige. Für die Mitteilung mathematischer Gegenstände und Gedanken hat sich weltweit die Mengensprache durchgesetzt. Die Mengensprache<sup>1</sup> ist Teil der natürlichen Umgangssprache, keiner weiteren Begründung nötig noch fähig, erlernbar durch Gebrauch und Nachahmung. Erstaunlich, daß ein Arsenal von wenigen Begriffen und wenigen Prinzipien ihrer Verarbeitung für die Mathematik ausreicht. Da die natürliche Sprache mehrdeutig ist, sammeln wir hier — ohne weitere inhaltliche Erklärung — das Vokabular und beschreiben seine mathematische Verwendung. Wie üblich lebt die mathematische Schriftform von abkürzenden Symbolen.

Eine *Menge* ist eine Zusammenfassung von *Elementen* zu einem neuen Objekt. Das Symbol  $x \in A$  besagt,  $x$  ist ein Element der Menge  $A$ . Wir sagen dafür auch,  $x$  ist in der Menge  $A$  enthalten. Wie wird eine Menge mitgeteilt? Eine Möglichkeit ist, ihre Elemente in geschweiften Klammern einfach aufzuschreiben; so ist  $\{1, 2, 3\}$  die Menge mit den Elementen 1,2,3. Bei Mengen mit unendlich vielen Elementen ist das natürlich unmöglich. Dann kann man etwa folgende Mitteilungsart wählen:  $\{x \mid x \text{ hat die Eigenschaft } E\}$  ist die Menge aller Elemente mit der Eigenschaft  $E$ . Jedes Gedankending kann Element einer Menge werden. So können Mengen selbst wieder Elemente einer Menge sein. Eine Menge ist durch die in ihr enthaltenen Elemente bestimmt.

Ist jedes Element von  $A$  auch Element von  $B$ , so sagt man,  $A$  sei in  $B$  *enthalten*,  $A$  sei *Teilmenge*, *Untermenge* von  $B$  und schreibt  $A \subset B$  oder  $B \supset A$  dafür. Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente enthalten, wenn also sowohl  $A \subset B$  als auch  $B \subset A$  gilt. Auf diesem Wege

---

<sup>1</sup>Zu unterscheiden von der Mengenlehre, einer mathematischen Disziplin.

beweist man auch die Gleichheit zweier Mengen.

Aus vorgegebenen Mengen  $A$  und  $B$  lassen sich neue Mengen bilden. Der **Durchschnitt** oder **Schnitt**  $A \cap B$  besteht aus den Elementen, die sowohl in  $A$  als auch in  $B$  enthalten sind. Die **Vereinigung**  $A \cup B$  besteht aus den Elementen, die entweder in  $A$  oder in  $B$  oder in beiden enthalten sind. Das **cartesische Produkt**  $A \times B$  hat als Elemente alle geordneten Paare  $(a, b)$ , wobei  $a \in A$  und  $b \in B$  ist. Der Terminus „geordnet“ bedeutet, daß es auf die Reihenfolge ankommt, das heißt  $(a, b)$  ist von  $(b, a)$  zu unterscheiden.

Die vorstehenden Bildungen lassen sich für ein beliebiges System  $(A_j \mid j \in J)$  von Mengen  $A_j$  durchführen. Der Index  $j$  dient hier zur Kennzeichnung und Unterscheidung der Mengen und kann zum Beispiel eine Nummer sein. Man nennt in diesem Kontext  $J$  eine **Indexmenge** und  $(A_j \mid j \in J)$  eine durch  $J$  indizierte **Familie** von Mengen.

Der **Durchschnitt**  $\bigcap_{j \in J} A_j$  besteht aus den Elementen, die in sämtlichen  $A_j$  liegen. Die **Vereinigung**  $\bigcup_{j \in J} A_j$  besteht aus den Elementen, die in mindestens einer der Mengen  $A_j$  liegen. Damit man den Durchschnitt immer bilden kann, ist es nützlich, die **leere Menge**  $\emptyset$  zuzulassen, die überhaupt kein Element hat und folglich Teilmenge jeder Menge ist. Gilt  $A \cap B = \emptyset$ , so heißen  $A$  und  $B$  **disjunkt**. Natürlich verwendet man Bezeichnungen wie  $A_1 \cap A_2 \cap A_3$  für den Durchschnitt der drei Mengen  $A_1, A_2, A_3$ . Analog für Vereinigungen und bei mehreren Mengen.

Sind  $A$  und  $B$  Mengen, so bezeichnet  $A \setminus B$  die Menge der Elemente von  $A$ , die nicht in  $B$  enthalten sind (**Differenzmenge** oder das **Komplement**). Bei dieser Bezeichnung wird nicht unterstellt, daß  $B$  Teilmenge von  $A$  ist. Besteht  $B$  nur aus dem Element  $b$ , so schreiben wir auch kurz  $A \setminus b$ .

Das **cartesische Produkt**  $\prod_{j \in J} A_j$  besteht aus allen Familien  $(a_j \mid j \in J)$ , worin  $a_j \in A_j$  ist. Auch hier gibt es Bezeichnungen wie  $A_1 \times A_2 \times A_3$ .

Seien  $A$  und  $B$  Mengen. Eine **Abbildung von  $A$  nach  $B$**  ist eine Vorschrift, die jedem  $a \in A$  ein  $b \in B$  zuordnet. Die Vorschrift wird meist durch ein Funktionssymbol  $b = f(a)$  geschrieben. Man sagt auch,  $a$  werde auf  $f(a)$  abgebildet. In Symbolen schreiben wir dann

$$f: A \rightarrow B, \quad a \mapsto f(a) = fa.$$

Manchmal schreiben wir auch  $f$  neben oder über den Pfeil. Dabei heißt  $A$  der **Definitionsbereich** oder die **Quelle** der Abbildung,  $B$  der **Bildbereich** oder das **Ziel** der Abbildung. Die „Vorschrift“ ist durch die Teilmenge

$$\{(a, b) \mid b = f(a), a \in A\} \subset A \times B$$

bestimmt und kann mengentheoretisch mit ihr gleichgesetzt werden. Zu einer Abbildung gehören also die **drei** Daten  $A$ ,  $B$  und  $f$ , wenn man präzise sein will. In der Praxis erlaubt man sich allerdings gewisse Freiheiten. Statt Abbildung sagt man auch **Funktion**, insbesondere dann, wenn die Zielmenge aus Zahlen

besteht. In einigen Fällen sind auch Worte wie **Operator** oder **Funktional** gebräuchlich. Die Abbildung  $A \times B \rightarrow A$ ,  $(a, b) \mapsto a$  heißt **Projektion auf den Faktor**  $A$ ; analog hat man für ein beliebiges cartesisches Produkt die Projektionen auf die Faktoren. Die Abbildung  $\text{id}(A) = \text{id}: A \rightarrow A$ ,  $a \mapsto a$  heißt die **identische Abbildung** von  $A$ . Zu einer Inklusion  $A \subset B$  gehört die Abbildung  $i: A \rightarrow B$ ,  $a \mapsto a$ , die wir ebenfalls **Inklusion** nennen und auch mit  $i: A \subset B$  mißbräuchlich bezeichnen. Eine gegebene Abbildung  $f: B \rightarrow Y$  besitzt die **Einschränkung**  $f \circ i: A \rightarrow Y$  auf die Teilmenge  $A \subset B$ , die auch  $f|A$  notiert wird.

Die **Verkettung** der Abbildungen  $f: A \rightarrow B$  und  $g: B \rightarrow C$  ist die Abbildung  $A \rightarrow C$ , die  $a$  auf  $g(f(a))$  abbildet; sie wird mit  $gf$  oder  $g \circ f$  bezeichnet. Oft hat man eine größere Anzahl von Abbildungen und Verkettungen zu notieren. Das geschieht übersichtlich in einem **Diagramm**, welches aus den Abbildungspfeilen, ihren Quellen und Zielen besteht. Führen alle möglichen Verkettungen zwischen festen Stellen in einem Diagramm immer zum gleichen Ergebnis, so sagen wir, das Diagramm sei **kommutativ** oder es **kommutiere**. Zum Beispiel ist ein Rechteck

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \downarrow g & & \downarrow h \\ C & \xrightarrow{i} & D \end{array}$$

kommutativ genau dann, wenn  $h \circ f = i \circ g$  gilt.

Zu  $f: A \rightarrow B$  und  $f': A' \rightarrow B'$  gehört das **cartesische Produkt**

$$f \times f': A \times A' \rightarrow B \times B', \quad (a, a') \mapsto (f(a), f'(a')).$$

Analog haben wir das cartesische Produkt

$$\prod_{j \in J} f_j: \prod_{j \in J} A_j \rightarrow \prod_{j \in J} B_j$$

einer beliebigen Familie  $(f_j: A_j \rightarrow B_j \mid j \in J)$  von Abbildungen.

Eine Abbildung  $f: A \rightarrow B$  heißt **injektiv**, wenn aus  $a_1 \neq a_2$  immer  $f(a_1) \neq f(a_2)$  folgt, und **surjektiv**, wenn zu jedem  $b \in B$  ein  $a \in A$  mit  $b = f(a)$  existiert. Sie heißt **bijektiv**, wenn sie sowohl surjektiv als auch injektiv ist. Eine Abbildung  $f: A \rightarrow B$  ist genau dann bijektiv, wenn sie eine **Umkehrabbildung**  $g: B \rightarrow A$  hat, das heißt wenn eine Abbildung  $g$  existiert mit den Eigenschaften  $gf = \text{id}(A)$  und  $fg = \text{id}(B)$ . Ist  $C \subset A$ , so heißt  $f(C) = \{b \mid \text{es gibt } c \in C \text{ mit } b = f(c)\}$  das **Bild** von  $C$  bei  $f$ . Ist  $D \subset B$ , so heißt  $f^{-1}(D) = \{a \mid f(a) \in D\}$  das **Urbild** von  $D$  bei  $f$ . Das Urbild der einelementigen Teilmenge  $\{c\}$  wird  $f^{-1}(c)$  geschrieben.

Zwei Mengen heißen **gleichmächtig**, wenn es eine Bijektion (= eine bijektive Abbildung) zwischen ihnen gibt. Die **Mächtigkeit** einer Menge  $M$  ist die „Anzahl“ ihrer Elemente (was immer das für unendliche Mengen bedeuten

mag); sie wird mit  $|M|$  bezeichnet. Eine Menge heißt **abzählbar unendlich**, wenn es eine Bijektion zur Menge der natürlichen Zahlen gibt.

Eine **Äquivalenzrelation** auf einer Menge  $A$  ist eine Teilmenge  $R \subset A \times A$  mit den folgenden Eigenschaften:

- (1) Für jedes  $a \in A$  ist  $(a, a) \in R$ .
- (2) Ist  $(a, b) \in R$ , so ist auch  $(b, a) \in R$ .
- (3) Ist  $(a, b) \in R$  und  $(b, c) \in R$ , so ist auch  $(a, c) \in R$ .

Ist  $R$  solch eine Äquivalenzrelation, so schreibt man  $a \sim b$  für  $(a, b) \in R$  und sagt in diesem Fall,  $a$  ist **äquivalent** zu  $b$  bezüglich oder modulo dieser Relation. Die Menge  $K(a) = \{b \mid a \sim b\}$  heißt die **Äquivalenzklasse** von  $a$ . Genau dann gilt  $K(a) = K(b)$ , wenn  $a \sim b$  ist. Die Menge der Äquivalenzklassen wird mit  $A/R$  bezeichnet und  $A$  modulo  $R$  gelesen. Die Abbildung  $p: A \rightarrow A/R, a \mapsto K(a)$  heißt die **Quotientabbildung** der Äquivalenzrelation. Es ist eine surjektive Abbildung. Die Menge  $A$  ist die disjunkte Vereinigung der Äquivalenzklassen. Ist irgendeine disjunkte Zerlegung von  $A$  in nichtleere Teilmengen gegeben, so gibt es genau eine Äquivalenzrelation auf  $A$ , deren Äquivalenzklassen genau die Mengen dieser Zerlegung sind. Ist  $f: A \rightarrow B$  eine surjektive Abbildung, so bilden die Urbilder der Elemente von  $B$  solch eine disjunkte Zerlegung. Ein Element einer Äquivalenzklasse wird **Repräsentant** dieser Klasse genannt. Ist  $p: A \rightarrow A/R$  die Quotientabbildung einer Äquivalenzrelation und  $f: A \rightarrow B$  eine Abbildung, so gibt es genau dann eine Abbildung  $g: A/R \rightarrow B$  mit der Eigenschaft  $g \circ p = f$ , wenn  $f$  die Äquivalenzklassen von  $R$  auf einelementige Mengen abbildet. Man definiert  $g$  durch  $g(K(a)) = f(a)$  und sagt, die Definition sei **unabhängig** von der Auswahl des Repräsentanten oder kurz und salopp:  $g$  ist **wohldefiniert**<sup>2</sup>. Es gibt also drei im wesentlichen gleichwertige Erscheinungsformen von Äquivalenzrelationen: Die Relation  $R$ , die disjunkte Zerlegung in Klassen, die surjektive Abbildung auf die Menge der Klassen.

Der Durchschnitt von Äquivalenzrelationen auf  $A$ , aufgefaßt als Teilmengen von  $A \times A$ , ist wieder eine. Zu jeder beliebigen Teilmenge  $S \subset A \times A$  gibt es deshalb eine kleinste  $S$  umfassende Äquivalenzrelation, genannt die von  $S$  **erzeugte**. Sind  $R_1$  und  $R_2$  Äquivalenzrelationen auf  $A$  und gilt  $R_1 \subset R_2$ , so heißt  $R_1$  **feiner** als  $R_2$  und  $R_2$  **gröber** als  $R_1$ . Die feinste Äquivalenzrelation ist die **Gleichheit**; die zugehörige Teilmenge von  $A \times A$  ist die **Diagonale**  $\{(a, a) \mid a \in A\}$  von  $A \times A$ .

Allgemein wird eine Teilmenge  $R \subset A \times A$  als eine (zweistellige) **Relation** auf  $A$  bezeichnet. Die Idee dabei ist, daß  $R$  diejenigen Paare  $(a, b)$  aussondert, die in einer jeweils festgelegten Beziehung (Relation) zueinander stehen. Meist haben Relationen ihre eigene Notation, zum Beispiel  $a < b$ ,  $a$  kleiner als  $b$ . In diesem Fall besteht  $R$  aus den Paaren  $(a, b)$ , für die  $a < b$  gilt.

<sup>2</sup>Damit ist „wohldefiniert“ definiert.

# Index

- $f$ -stabil, 22
- $n$ -Fakultät, 56
- Äquivalenzklasse, 151
- Äquivalenzrelation, 151
  - feiner, 151
  - gröber, 151
  - von einer Menge erzeugte, 151
- äquivalent, 151
  
- Abbildung
  - affine, 103
  - duale, 23
  - induzierte, 30
  - lineare, 19
  - orthogonal, 73
  - projektive, 112
  - selbstadjungiert, 87
  - transponierte, 23
  - unitär, 73
- Abbildung von  $A$  nach  $B$ , 149
- Abstand, 68, 72
  - orientierter, 72
- Abtragen eines Vektors, 101
- Abweichung von der Linearität, 66
- abzählbar unendlich, 151
- Additionstheoreme, 95
- adjungierte Abbildung, 86
- adjungierte lineare Abbildungen, 83
- Adjunkte, 60
- affine Abbildung, 103
- affiner Raum, 101, 102
  - Standardraum, 101
- affiner Unterraum, 50
- affines Koordinatensystem, 104, 106
- Algorithmus, 48
- allgemeine lineare
- alternierende Gruppe, 55
- anisotrop, 88
- Antihomomorphismus, 128
- Austauschsatz von Steinitz, 16
- Automorphismus, 20
  
- baryzentrische Koordinaten, 104, 106
- Basen
  - gleich orientiert, 62
- Basis, 14
  - geordnete, 17
  - negativ orientiert, 62
  - positiv orientiert, 62
- Basiswechselmatrix, 42
- Besselsche Ungleichung, 70
- Betrag
  - einer Quaternion, 132
- bijektiv, 150
- Bild, 150
- Bildbereich, 149
- bilinear, 67, 83
- Bilinearform, 83
  - alternierend, 85
  - geometrisch, 87
  - negativ definit, 89
  - positiv definit, 89
  - regulär, 85
  - schief-symmetrisch, 85
  - symmetrisch, 67, 85
- Bilinearform
  - alternierend, 141
- Blockmatrizen, 39
  
- cartesisches Produkt, 150
- cartesisches Produkt, 149
- Cauchy-Schwarzsche Ungleichung, 69
- charakteristisches
- cosinus hyperbolicus, 95
- Cramersche Regel, 60
  
- Definitionsbereich, 149
- Determinante, 57
- Determinante einer linearen Abbildung, 59
- Determinantenform, 54
- Determinantenfunktion, 57
- Diagonale, 151
- Diagonalmatrix, 38

- Diagramm, 150  
     kommutativ, 150  
 Differenzmenge, 149  
 Dimension, 17  
     eines projektiven Raumes, 106  
 Dimensionsformel, 21  
 direkte Summe, 27  
 disjunkt, 149  
 Distributivgesetz, 10  
 Divisionsalgebra, 128  
 Doppelverhältnis, 114  
 Drehachse, 78  
 Drehmatrix, 74, 77, 80  
 Drehung, 77, 80  
 Dreiecksmatrix, 39  
     obere, 39  
     untere, 39  
 Dreiecksungleichung, 5, 69  
 Dreier-Identität, 131  
 Dualbasis, 23  
 duale Abbildung, 23  
 Dualität, 85  
 Dualraum, 22  
 Durchschnitt, 149  
  
 Ebene  
     projektive, 109  
 Eigenraum, 22  
     verallgemeinerter, 119  
 einfach transitiv, 101  
 Einheitsmatrix, 38  
 Einheitsvektor, 68  
 Einheitswurzel, 7  
 Einschränkung, 150  
 Einselement, 6  
 Element, 148  
 elementare Umformungen, 48  
 Elementarmatrix, 50, 51  
 Eliminationsverfahren, 48  
 Endomorphismus, 20  
     charakteristisches Polynom, 64  
     nilpotent, 117  
     zyklischer, 116  
 Entwicklung nach einer  
 Entwicklung nach einer  
 Epimorphismus, 20  
 Erste Dimensionsformel, 24  
  
 euklidischer Vektorraum, 67  
 exakte Sequenz, 24  
     kurze, 24  
  
 Fahne, 39  
 Familie, 149  
 Funktion, 149  
 Funktional, 150  
  
 Gaußscher Algorithmus, 48  
 Gerade  
     projektive, 109  
 Gleichheit, 151  
 gleichmächtig, 150  
 Gleichung  
     homogen, 49, 50  
     inhomogen, 50  
     Koeffizientenmatrix, 50  
 Graßmann-Identität<sup>3</sup>, 126  
 Gruppe, 6  
     abelsch, 6  
     allgemeine lineare, 38  
     alternierende, 55  
     kommutativ, 6  
     orthogonale, 73, 92  
     spezielle lineare, 59  
     spezielle orthogonale, 73  
     symmetrische, 55  
     unitäre, 73  
 Gruppenaddition, 6  
 Gruppenmultiplikation, 6  
 Gruppenstruktur, 6  
  
 Höhenlinie, 139  
 Halbsystem, 55  
 Hauptachsen, 83  
 Hauptachsentransformation, 83  
 hermitesch, 79  
 hermitesche Form, 67  
 Hessesche Normalform, 72  
 Hilbert-Raum, 67  
 homogene Koordinaten, 107  
 Homomorphismus, 7  
 hyperbolische Form, 93  
 Hyperebene  
  


---

<sup>3</sup>Hermann Günther Graßmann 1809 – 1877

- affine, 29
- projektive, 107
- unendlich ferne, 107
- identische Abbildung, 150
- imaginäre Einheit, 5
- Imaginärteil, 5
- Index, 91
- Indexmenge, 149
- injektiv, 150
- Inklusion, 150
- inneres Produkt, 67
- Inverses, 6
- Isometrie, 73, 85
- isomorph, 7
- Isomorphismus, 7, 20
- isotrop, 88
- Jacobi-Identität, 125
- Jordan-Matrix, 119
- Körper, 8
  - Addition, 8
  - Charakteristik zwei, 89
  - Multiplikation, 8
- Kartenabbildung, 107
- Kleinsche Vierergruppe, 95
- Kodimension, 17
- kommutativ, 150
- Komplement, 149
- Komplement eines Unterraumes, 25
- komplexe Struktur, 142
- konjugiert-linear, 143
- konvex, 102
- konvexe Hülle, 102
- Koordinaten
  - baryzentrische, 104, 106
  - bezüglich einer geordneten Basis, 18
  - homogene, 107
- Koordinatenabbildung, 44
- Koordinatentransformationen, 45
- Kronecker-Symbol, 38
- Länge, 68
- leere Menge, 149
- Leibnizsche Formel, 56
- Lie-Klammer, 125
- linear abhängig
  - von einer Menge, 14
- linear unabhängig von einer Menge, 14
- lineare Abbildung, 19
- lineare Hülle, 13
- Linearform, 22, 54
  - alternierend, 54
- Linearkombination, 14
- linksinvers, 23
- Linksmultiplikation, 128
- Lorentz-Gruppe, 92, 98, 99
  - kleine, 93
- Lorentz-Transformation, 94, 98
  - orthochron, 98
- Lot, 72
- Mächtigkeit, 150
- Matrix
  - Adjunkte, 60
  - Block, 39
  - charakteristisches Polynom, 63
  - Diagonale, 38
  - Diagonalelemente, 38
  - diagonalisierbar, 64
  - Einträge, 32
  - Elemente, 32
  - erweiterte, 50
  - Hauptdiagonale, 38
  - hermitesche, 68
  - invertierbar, 38
  - Nebendiagonale, 38
  - nilpotent, 39
  - normal, 82
  - orthogonal, 73
  - positiv definit, 68
  - quadratisch, 38
  - Rang, 46
  - regulär, 38
  - Spalte, 33
  - Spaltenindex, 33
  - Spaltenrang, 45
  - Spaltenvektor, 33
  - Spur, 43
  - symmetrische, 68
  - transponierte, 40
  - unitär, 73
  - Zeile, 33

- Zeilenindex, 33
- Zeilenrang, 45
- Zeilenstufenform, 51
- Zeilenvektor, 33
- Matrix von  $f$  bezüglich  $B, C$ , 34
- Matrizen
  - ähnlich, 42
  - äquivalent, 42
  - grob-äquivalent, 42
  - konjugiert, 42
- Matrizenprodukt, 35
- Menge, 148
  - linear abhängig, 14
  - linear unabhängig, 14
- Minkowski-Raum, 92, 97
- Monomorphismus, 20
- Norm, 68, 72, 80
- Normalform
  - jordansche, 119
- Nullelement, 7
- Nullmatrix, 38
- Nullraum, 12
- Nullstelle, 63
- Nullvektor, 12
- Operator, 150
- Ordnung eines Gruppenelementes, 9
- Orientierung, 62
  - kanonische, 62
- orthogonal, 87, 88
- orthogonale Gruppe, 73
- orthogonale Projektion, 71
- orthogonale Summe, 88
- orthogonales Komplement, 71
- Orthogonalsystem, 70
- Orthonormalbasis, 70, 89
- Orthonormalisierungsverfahren, 70
- Orthonormalsystem, 70
- Parsevalsche Ungleichung, 70
- Permutation, 55
  - gerade, 55
  - Signum, 55
  - Transposition, 55
  - ungerade, 55
  - Vorzeichen, 55
- Permutationsgruppe, 55
- Polarsierung, 140
- positiv definit, 67, 80
- Produkt von Vektorräumen, 26
- Produktsatz für Determinanten, 59
- Projektion, 25
- Projektion auf den Faktor  $A$ , 150
- Projektionsoperator, 25
- projektive Quadrik, 111
- projektive
  - projektiver Raum, 106
- projektives Koordinatensystem, 113
- Projektivität, 112
- Punkt
  - eines affinen Raumes, 101
- Punkte
  - in allgemeiner Lage, 113
  - unabhängig, 113
- Quader, 61
- quadratische Form, 140
- quadratische Verfeinerung, 142
- quadratischer Raum, 140
- Quadrik, 83
  - projektive, 111
- Quaternionen, 128
  - $\mathbb{C}$ -Standardbasis, 129
  - $\mathbb{R}$ -Standardbasis, 129
  - euklidisches Skalarprodukts, 129
  - Konjugation, 128
  - Multiplikationsregeln, 129
  - reine, 129
  - unitäres Skalarprodukt, 129
  - Zentrum, 130
- Quelle, 149
- Quotientabbildung, 29, 151
- Quotientraum, 30
- Radikal, 88
- Rang einer linearen Abbildung, 21
- Rang einer Matrix, 46
- Raum
  - affiner, 101, 102
  - projektiver, 106
- Realteil, 5
- rechtsinvers, 23
- reelle Struktur, 143

- Relation, 151
- Repräsentant, 151
  
- Sattelfläche, 139
- Sattelpunkt, 139
- Satz des Pythagoras, 68
- schief-hermitesch, 134
- Schiefkörper, 128
- schiefsymmetrisch, 125
- Schnitt, 149
- selbstadjungiert, 79
- Signatur, 91
- Simplex, 106
- simultane Diagonalisierung, 76
- sinus hyperbolicus, 95
- Skalar, 10
- Skalarprodukt, 67
- Spaltenrang, 45
- spezielle lineare Gruppe, 59
- Spiegelung, 78, 132
- Spur, 43
- Standardbasis, 15
- Standardskalarprodukt, 66, 67
- Standardvektor, 15
- Standardvektorraum, 11
- Strecke, 102
- Struktursatz für nilpotente Morphismen, 118
- Stufe eines Vektors, 117
- Summe
  - orthogonale, 68
- Summe von Unterräumen, 24
- Summe von Vektorräume, 26
- Summenkonvention, 37
- surjektiv, 150
- symmetrisch, 79
- symmetrische Bilinearform, 67
- symmetrische Gruppe, 55
  
- tangens hyperbolicus, 95
- Teilmenge, 148
- Trägheitssatz von Sylvester, 91
- Translation, 101
- Transposition, 55
  
- Umkehrabbildung, 150
- unabhängig, 151
  
- unitäre Gruppe, 73
- unitärer Vektorraum, 67
- Untergruppe, 7
- Unterkörper, 8
- Untermatrix, 33
- Untermenge, 148
- Unterräume
  - orthogonale, 68
- Unterraum, 11
  - $c$ -reell, 145
  - affiner, 29, 50
    - Richtung, 29, 50
  - aufgespannt von einer Menge, 13
  - Erzeugendensystem, 13
  - erzeugt von einer Menge, 13
  - negativ, 90
  - positiv, 90
  - trivialer, 12
- Untervektorraum, 11
- Urbild, 150
  
- Vektor, 10
  - lineare Darstellung, 14
- Vektoren
  - orthogonale, 68
- Vektorprodukt, 124
- Vektorräume
  - isomorph, 20
- Vektorraum
  - $f$ -zyklisch, 116
  - (interne) direkte Zerlegung, 25
  - Addition, 10
  - additive Gruppe, 11
  - direkte Summe, 25
  - euklidischer, 67
  - komplexer, 11
  - konjugierter, 13
  - Multiplikation mit Skalaren, 10
  - normierter, 72
  - orientierter, 62
  - primärer Anteil zu einem reeler, 11
  - unitärer, 67
- Vektorraum über einem Körper, 10
- Vektorraumstruktur, 10
- Verbindungsstrecke, 102
- Verbindungsvektor, 101

Vereinigung, 149  
Vergangenheitskegel, 98  
Verkettung, 150  
Verknüpfung, 5  
    assoziativ, 5  
    kommutativ, 6  
    neutrales Element, 6  
Vielfachheit, 63  
Volumen, 62  
  
Winkel, 69  
wohldefiniert, 151  
  
Zahl  
    Betrag, 5  
    komplexe, 5  
        Imaginärteil, 5  
        konjugierte, 5  
        Realteil, 5  
    Norm, 5  
Zahlen, 4  
    ganze, 4  
    komplexe, 4  
    natürliche, 4  
    rationale, 4  
    reelle, 4  
Zeilenrang, 45  
Zeilenumformungen, 51  
Zentralprojektion, 112  
Zentrum, 130  
Ziel, 149  
Zukunftskegel, 98