

# Lineare Algebra — WS 2025/26

Sebastian Goette



# Inhaltsverzeichnis

Einleitung	1
Kapitel 1. Zahlen	3
1.1. Mengen und Abbildungen	3
1.2. Natürliche Zahlen	10
1.3. Ganze und Rationale Zahlen	14
1.4. Etwas Euklidische Geometrie	21
1.5. Komplexe Zahlen und die Geometrie der Ebene	24
1.6. Geometrie des Raumes und Quaternionen	29
1.7. Zusammenfassung	36
Kapitel 2. Vektorräume und Moduln	37
2.1. Gruppen, Ringe, Körper	37
2.2. Moduln, Vektorräume und lineare Abbildungen	47
2.3. Unterräume und Quotienten	54
2.4. Linearkombinationen, Basen und Koordinaten	62
2.5. Matrizen	69
2.6. Unendliche Indexmengen	75
2.7. Zusammenfassung	78
Kapitel 3. Vektorräume über Körpern und Schiefkörpern	79
3.1. Basen	79
3.2. Dimension und Rang	82
3.3. Lineare Gleichungssysteme	88
3.4. Die Methode der kleinsten Quadrate	96
3.5. Zusammenfassung	100
Kapitel 4. Determinanten	101
4.1. Volumina und Determinantenfunktionen	101
4.2. Die Determinante	108
4.3. Orientierung reeller Vektorräume	117
4.4. Zusammenfassung	119
Kapitel 5. Eigenwerte und Normalformen	121
5.1. Eigenvektoren	121
5.2. Das charakteristische Polynom	127
5.3. Der Satz von Cayley-Hamilton	132
5.4. Das Minimalpolynom	136
5.5. Der Satz von der eindeutigen Primfaktorzerlegung	140
5.6. Die Hauptraumzerlegung	144

5.7.	Eine allgemeine Normalform für Endomorphismen	147
5.8.	Die Jordansche Normalform	152
5.9.	Anwendungen der Jordan-Normalform	158
5.10.	Zusammenfassung	162
Kapitel 6.	Vektorräume mit Skalarprodukt	163
6.1.	Skalarprodukte	163
6.2.	Skalarprodukte als Matrizen	170
6.3.	Dualräume und adjungierte Abbildungen	179
6.4.	Normale Endomorphismen	188
6.5.	Affine Räume	198
6.6.	Bilinearformen und quadratische Funktionen	204
Notation		215
Stichwortverzeichnis		217

# Einleitung

Die Lineare Algebra ist die Lehre von Vektorräumen und linearen Abbildungen. In der Schule haben Sie bereits Vektorrechnung in der Ebene und im Raum kennengelernt; dieser Stoff wird hier ausgebaut. In vielen weiterführenden Vorlesungen werden Ihnen immer wieder Vektorräume und lineare Abbildungen begegnen, so dass es sicher sinnvoll ist, sie bereits am Anfang des Studiums kennenzulernen.

Wir beginnen im ersten Kapitel mit einer allgemeinen Einführung, bei der wir Grundlagen und erste Beispiele kennenlernen. Dazu wiederholen wir die Zahlbereiche  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  und  $\mathbb{R}$ , die Sie aus der Schule kennen. Dann führen wir die komplexen Zahlen  $\mathbb{C}$  und die Quaternionen  $\mathbb{H}$  ein. Als ersten Vorgeschmack auf den Inhalt der Vorlesung beschreiben wir die Euklidische Geometrie der Ebene  $\mathbb{R}^2$  und des Raumes  $\mathbb{R}^3$  mit Hilfe der komplexen Zahlen beziehungsweise der Quaternionen.

Im zweiten Kapitel führen wir systematisch die Grundbegriffe ein. An die Stelle konkreter Zahlbereiche treten Ringe (wie  $\mathbb{Z}$ ), Körper (wie  $\mathbb{Q}$ ,  $\mathbb{R}$  oder  $\mathbb{C}$ ) und Schiefkörper (wie  $\mathbb{H}$ ). Die Ebene  $\mathbb{R}^2$  und der Raum  $\mathbb{R}^3$  sind die einfachsten Beispiele von Vektorräumen. Abbildungen, die mit der Vektorraum-Struktur verträglich sind, heißen linear. Allgemeiner lernen wir, wie man Elemente in freien Moduln über einem gegebenen Ring durch Koordinaten und lineare Abbildungen zwischen solchen Moduln durch Matrizen beschreibt.

In jedem Kapitel ab dem zweiten werden wir nur die Grundannahmen machen, die wir für den jeweiligen Stoff benötigen. Beispielsweise müssen wir in Kapitel 2 nicht dividieren und auch die Faktoren in Produkten nicht vertauschen, so dass wir statt über Körpern auch über nichtkommutativen Ringen arbeiten können. Wir werden aber immer nur dann allgemeinere Objekte als Vektorräume über Körpern betrachten, wenn das ohne zusätzlichen technischen Aufwand möglich ist. Aus diesem Grund ist das vorliegende Skript auch nicht schwerer zu verstehen als andere Skripten zur linearen Algebra.

Im dritten Kapitel konzentrieren wir uns auf Vektorräume über Körpern und Schiefkörpern. Wir zeigen, dass jeder Vektorraum eine Basis besitzt, und dass die Dimension eine Invariante des Vektorraums ist, die ihn bis auf Isomorphie bestimmt. Außerdem betrachten wir die Struktur einer allgemeinen linearen Abbildung und lernen ein universelles Verfahren zum Lösen linearer Gleichungssysteme.

Im vierten Kapitel beschäftigen wir uns mit Endomorphismen freier Moduln über kommutativen Ringen und lernen die Determinante als wichtige Invariante

kennen. Anschließend betrachten wir Eigenwerte und das charakteristische Polynom, und lernen erste Strukturaussagen über lineare Abbildungen von einem festen Vektorraum in sich selbst kennen.

## KAPITEL 1

# Zahlen

In diesem ersten Kapitel legen wir dazu die Grundlagen. Zuerst führen wir Sprechweisen für Mengen, Abbildungen und natürliche Zahlen ein. Danach konstruieren wir ganze und rationale Zahlen, wohingegen wir die reellen Zahlen als gegeben annehmen werden — ihre Konstruktion fällt in den Bereich der Analysis. Aus den reellen Zahlen konstruieren wir die komplexen Zahlen und die Quaternionen. Zum einen sind beides wichtige Beispiele für Körper beziehungsweise Schiefkörper. Auf der anderen Seite besteht ein enger Zusammenhang zur Euklidischen Geometrie in den Dimensionen 2 und 3, und euklidische Geometrie ist sicher einer der wichtigsten Vorläufer für den Vektorraum-Kalkül, um den es in dieser Vorlesung schwerpunktmäßig gehen wird.

### 1.1. Mengen und Abbildungen

Wenn man möchte, kann man fast die gesamte Mathematik auf das Studium von Mengen und ihren Elementen zurückführen. Das ist aber leider recht mühsam, und man muss sehr sorgfältig sein, um nicht in Widersprüche zu geraten. Wenn Sie wissen möchten, wie das geht, sollten Sie später im Verlauf Ihres Studiums eine Vorlesung über Mengenlehre besuchen. Wir wollen die Mengenlehre als eine Sprache benutzen, in der man sehr elegant über mathematische Sachverhalte sprechen kann. Dazu lernen wir jetzt die ersten Vokabeln und grammatikalischen Regeln.

Georg Cantor hat den Mengenbegriff als erster eingeführt.

„Eine Menge ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unseres Denkens oder unserer Anschauung zu einem Ganzen.“

1.1. BEISPIEL. Zahlen sind Objekte unserer Anschauung, also ist  $\{1, 2, 3\}$  eine Menge. Die Menge  $\mathbb{N} = \{0, 1, 2, \dots\}$  der natürlichen Zahlen lernen wir im Abschnitt 1.2 kennen.

Die „Objekte“ in einer Menge heißen *Elemente*. Wenn ein Objekt  $a$  in einer Menge  $M$  enthalten ist, schreiben wir

$$a \in M,$$

ansonsten  $a \notin M$ .

1.2. DEFINITION. Zwei Mengen heißen gleich, wenn sie die gleichen Elemente enthalten.

1.3. BEMERKUNG. Wenn man Mengen als Aufzählung  $M = \{a_1, \dots, a_n\}$  angibt, kann es passieren, dass  $a_i = a_j$  für zwei Indizes  $i$  und  $j$ . Trotzdem ist  $a_i$  dadurch nicht „zweimal“ in  $M$  enthalten. Also zum Beispiel

$$\{1, 1, 2\} = \{2, 1\} = \{1, 2\},$$

denn alle drei Mengen enthalten die gleichen Elemente, nämlich 1 und 2. Aber natürlich gilt

$$\{1, 2\} \neq \{1, 2, 3\}.$$

1.4. BEISPIEL. Besonders wichtig ist die *leere Menge*, die gar kein Element enthält. Wir schreiben

$$\emptyset = \{ \}.$$

Inzwischen sind auch Mengen „Objekte unseres Denkens oder unserer Anschauung“ geworden. Also kann man auch Mengen betrachten, deren Elemente selbst wieder Mengen sind. In der Tat kann man ausgehend von der leeren Menge bereits sehr viele andere Mengen konstruieren, etwa

$$\emptyset = \{ \}, \quad \{ \emptyset \}, \quad \{ \{ \emptyset \}, \emptyset \} \quad \text{usw.} \dots,$$

genug, um alle Objekte dieser Vorlesung zu beschreiben.

Wir stoßen jetzt auf das erste Problem mit Cantors Mengenbegriff.

1.5. SATZ (Russellsche Antinomie). *Es gibt keine Menge  $M$ , deren Elemente genau diejenigen Mengen sind, die sich nicht selbst enthalten.*

Wir formulieren die Russellsche Antinomie hier wie selbstverständlich als einen *Satz*, also als eine bewiesene mathematische Aussage. Zu ihrer Zeit war die Russellsche Antinomie ein Widerspruch im mathematischen Denkgebäude — so etwas darf es nicht geben, denn aus einem Widerspruch lässt sich alles folgern, man könnte als Mathematiker nicht mehr zwischen „richtig“ und „falsch“ unterscheiden, und dadurch würde Mathematik als Ganzes bedeutungslos. Man hat einige Zeit gebraucht, um eine handhabbare Version der Mengenlehre zu formulieren, in der aus dem fatalen Widerspruch ein harmloser Satz wird.

BEWEIS. Würde es eine solche Menge  $M$  geben, dann müsste entweder  $M \in M$  oder  $M \notin M$  gelten. Aber nach Definition von  $M$  gilt  $M \in M$  genau dann, wenn  $M \notin M$ , und das ist ein Widerspruch. Also gibt es keine Menge  $M$ .  $\square$

1.6. BEMERKUNG. Wir haben gerade unseren ersten *indirekten Beweis* kennengelernt. Bei einem indirekten Beweis nimmt man an, dass die Aussage, die man beweisen möchte, falsch ist, und leitet daraus einen Widerspruch her. Manchmal ist das die einfachste Weise, einen Satz zu beweisen. Der Nachteil ist aber, dass man — wie im obigen Beweis — nicht auf Anhieb versteht, warum der Satz gilt. Wenn möglich, wollen wir daher indirekte Beweise vermeiden.

Zurück zu Cantors Mengenbegriff und zur Russellschen Antinomie. Wir sehen, dass nicht jede „Zusammenfassung von Objekten unseres Denkens und unserer Anschauung“ eine Menge sein kann. Wir werden daher die Existenz einiger nützlicher Mengen annehmen, und wir werden einige Konstruktionen

angeben, die neue Mengen aus alten erzeugen. Die gesamte Mathematik basiert auf der Annahme, dass man das ohne Widersprüche machen kann — aber aus prinzipiellen Gründen lässt sich die Widerspruchsfreiheit der Axiome der Mengenlehre nicht beweisen.

1.7. DEFINITION. Seien  $M$  und  $N$  Mengen, dann heißt  $M$  eine *Teilmenge* von  $N$ , wenn alle Elemente  $a$  von  $M$  auch in  $N$  enthalten sind. Dafür schreiben wir

$$M \subset N .$$

- 1.8. BEMERKUNG. (1) Die leere Menge ist Teilmenge jeder Menge  $M$ .  
 (2) Es gilt  $\{x\} \subset M$  genau dann, wenn  $x \in M$ .  
 (3) Es gilt immer  $M \subset M$ .  
 (4) Wenn  $M \subset N$  und  $M \neq N$  gilt, heißt  $M$  auch *echte Teilmenge* von  $N$ .

1.9. BEMERKUNG. Angenommen, wir wollen zeigen, dass zwei Mengen  $M$  und  $N$  gemäß Definition 1.2 gleich sind. Dazu werden wir oft erst  $M \subset N$  und dann  $N \subset M$  beweisen; aus beiden Aussagen zusammen folgt  $M = N$ .

Es sei  $\emptyset$  die leere Menge aus Beispiel 1.4. Gilt dann  $\emptyset = \{\emptyset\}$ ?

- Da  $\emptyset$  nach Konstruktion keine Elemente enthält, liegt jedes Element von  $\emptyset$  in  $\{\emptyset\}$ . Also gilt  $\emptyset \subset \{\emptyset\}$ .
- Die Menge  $\{\emptyset\}$  enthält das Element  $\emptyset$ , aber  $\emptyset \notin \emptyset$ , denn  $\emptyset$  enthält ja gar keine Elemente. Also gilt  $\{\emptyset\} \not\subset \emptyset$ .

Und somit gilt  $\emptyset \neq \{\emptyset\}$ .

In den meisten Mathebüchern wird das Symbol „ $\subset$ “ so verwendet wie hier. Es gibt zwar eine internationale Norm, nach der nur echte Teilmengen mit „ $\subset$ “ bezeichnet werden sollen, aber in der Mathematik benötigt man das Symbol für beliebige Teilmengen weitaus häufiger, und schreibt daher „ $\subset$ “. Für echte Teilmengen verwenden wir das Symbol „ $\subsetneq$ “. Falls Sie ein Mathebuch zur Hand nehmen, in dem das Symbol „ $\subset$ “ vorkommt, sollten Sie zur Sicherheit trotzdem herausfinden, ob der Autor damit beliebige oder nur echte Teilmengen bezeichnet. Genauso vorsichtig sollten Sie eigentlich mit allen Definitionen und Bezeichnungen verfahren.

Kommen wir jetzt zur Konstruktion neuer Mengen aus alten.

1.10. DEFINITION. Seien  $M$  und  $N$  Mengen.

- (1) Der *Durchschnitt*  $M \cap N$  enthält genau die Elemente, die sowohl in  $M$  als auch in  $N$  enthalten sind.
- (2) Die *Vereinigung*  $M \cup N$  enthält genau die Elemente, die in  $M$  oder in  $N$  enthalten sind.
- (3) Wenn  $M \cap N = \emptyset$  gilt, heißen  $M$  und  $N$  *disjunkt*, und  $M \cup N$  ist eine *disjunkte Vereinigung*. Um zu zeigen, dass eine Vereinigung disjunkt ist, schreiben wir  $M \dot{\cup} N$ .

- (4) Die (*Mengen-*) *Differenz*  $N \setminus M$  enthält genau die Elemente, die in  $N$ , aber nicht in  $M$  enthalten sind. Ist  $M$  Teilmenge von  $N$ , so nennt man  $N \setminus M$  auch das *Komplement* von  $M$  in  $N$ .
- (5) Das *kartesische Produkt*  $M \times N$  besteht aus allen Paaren  $(x, y)$  von Elementen  $x \in M$  und  $y \in N$ .

Insbesondere sind  $M \cap N$ ,  $M \cup N$ ,  $N \setminus M$  und  $M \times N$  auch wieder Mengen. Für den Anfang reichen uns diese Konstruktionen. Später werden wir Vereinigungen und Durchschnitte beliebig vieler Mengen benötigen.

1.11. BEMERKUNG. Die Notation  $(x, y)$  bezeichnet ein (geordnetes) *Paar*, allgemeiner bezeichnet  $(x_1, \dots, x_n)$  ein  *$n$ -Tupel*. Hierbei kommt es auf die Reihenfolge der Einträge (nicht „Elemente“!) an, und ein und derselbe Eintrag kann mehrfach auftreten. Zum Beispiel:

$$(1, 1) \in \{1, 2\} \times \{1, 2, 3\}$$

und

$$(1, 2) \neq (2, 1) \neq (2, 1, 1).$$

1.12. DEFINITION. Die Menge aller Teilmengen von  $M$  heißt *Potenzmenge*  $\mathcal{P}(M)$ . Auch die Potenzmenge einer Menge ist wieder eine Menge.

1.13. BEISPIEL. Sei  $M = \{1, 2\}$ , dann gilt

$$\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$

Es sei  $M$  eine Menge. Man betrachtet oft die Teilmenge aller Elemente  $z$  von  $M$ , die eine bestimmte Eigenschaft  $E$  haben, und schreibt dafür

$$\{z \in M \mid z \text{ hat die Eigenschaft } E\}.$$

Wenn  $E$  eine mathematisch wohldefinierte Eigenschaft ist, dann erhalten wir wieder eine Menge.

1.14. FOLGERUNG (aus der Russellschen Antinomie 1.5). *Die Gesamtheit aller Mengen ist keine Menge.*

BEWEIS. Noch ein indirekter Beweis: Wäre die Gesamtheit aller Mengen selbst eine Menge  $N$ , dann wäre auch

$$M = \{X \in N \mid X \notin X\}$$

wieder eine Menge, was nach Satz 1.5 aber nicht sein kann.  $\square$

1.15. DEFINITION. Es seien  $M$  und  $N$  Mengen. Eine *Abbildung*  $F: M \rightarrow N$  (lies „ $F$  von  $M$  nach  $N$ “) ordnet jedem Element  $x \in M$  ein Element  $F(x) \in N$  zu.

Formal betrachten wir Teilmengen  $X \subset M \times N$ . Wir fordern, dass zu jedem  $x \in M$  genau ein  $y \in N$  mit  $(x, y) \in X$  existiert, und setzen  $F(x) = y$ . Also ist  $X$  der *Graph*  $\Gamma(F) = \{(x, F(x)) \mid x \in M\}$  von  $F$ .

1.16. DEFINITION. Es sei  $F: M \rightarrow N$  eine Abbildung. Dann heißt  $M$  der *Definitionsbereich* von  $M$  und  $N$  der *Wertebereich*. Die Menge aller Abbildungen von  $M$  nach  $N$  wird mit  $\text{Abb}(M, N)$  bezeichnet .

Zwei Abbildungen sind *gleich*, wenn sie den gleichen Definitions- und den gleichen Wertebereich haben, und jedem Element des Definitionsbereichs jeweils dasselbe Element des Bildbereichs zuordnen.

1.17. DEFINITION. Es sei  $F: M \rightarrow N$  eine Abbildung. Dann heißt die Teilmenge

$$\text{im } F = \{y \in N \mid \text{Es gibt } x \in M \text{ mit } F(x) = y\} = \{F(x) \mid x \in M\}$$

das *Bild* von  $F$ .

Sei  $V \subset N$  eine Teilmenge, dann heißt

$$F^{-1}(V) = \{x \in M \mid F(x) \in V\}$$

das *Urbild* von  $V$  unter  $F$ .

Für das Urbild der einelementigen Menge  $\{y\}$  schreibt man manchmal kurz  $F^{-1}(y)$  statt  $F^{-1}(\{y\})$ . Da das zu Missverständnissen führen kann, bleiben wir erst einmal bei  $F^{-1}(\{y\})$ .

1.18. DEFINITION. Eine Abbildung  $F: M \rightarrow N$  heißt

- (1) *injektiv*, wenn für alle  $x_1, x_2 \in M$  aus  $F(x_1) = F(x_2)$  schon  $x_1 = x_2$  folgt,
- (2) *surjektiv*, wenn für alle  $y \in N$  ein  $x \in M$  existiert mit  $F(x) = y$ , und
- (3) *bijektiv*, wenn sie injektiv und surjektiv ist.

1.19. BEISPIEL. (1) Für alle Mengen  $M$  ist die Abbildung  $\text{id}_M: M \rightarrow M$  mit  $\text{id}_M(x) = x$  definiert. Sie heißt die *Identität* und ist stets bijektiv.

- (2) Die Abbildung  $F: \mathbb{R} \rightarrow \mathbb{R}$  mit  $F(x) = x^2$  ist weder injektiv noch surjektiv, denn

$$F(-2) = F(2) = 4 \quad \text{und} \quad -1 \notin \text{im}(F).$$

- (3) Die Abbildung  $F: \mathbb{N} \rightarrow \mathbb{N}$  mit  $F(x) = x^2$  ist injektiv. Die Abbildung  $G: \mathbb{N} \rightarrow \{x^2 \mid x \in \mathbb{N}\}$  mit  $G(x) = x^2$  ist bijektiv. Diese Abbildungen sind verschieden, da sie andere Wertebereiche haben.

Trotzdem werden wir später manchmal beide Abbildungen mit dem gleichen Symbol bezeichnen.

1.20. DEFINITION. Seien  $L, M, N$  Mengen und  $F: M \rightarrow N$ ,  $G: L \rightarrow M$  Abbildungen. Die *Verkettung*  $F \circ G: L \rightarrow N$  (lies „ $F$  nach  $G$ “) ist die durch

$$(F \circ G)(x) = F(G(x))$$

definierte Abbildung.

1.21. BEMERKUNG. Die Buchstaben in „ $F \circ G$ “ scheinen „falsch herum“ zu stehen, denn die Abbildungen verlaufen von links nach rechts geschrieben so:

$$\begin{array}{ccccc} L & \xrightarrow{G} & M & \xrightarrow{F} & N \\ x & \mapsto & G(x) & \longrightarrow & F(G(x)) . \end{array}$$

Aber in „ $(F \circ G)(x) = F(G(x))$ “ stimmt die Reihenfolge wieder. Beispielsweise seien  $F, G: \mathbb{R} \rightarrow \mathbb{R}$  definiert durch

$$F(x) = x^2 \quad \text{und} \quad G(x) = x + 1 ,$$

dann ist

$$(F \circ G)(x) = (x + 1)^2 \quad \text{und} \quad (G \circ F)(x) = x^2 + 1 .$$

Insbesondere gilt  $G \circ F \neq F \circ G$ .

1.22. BEMERKUNG. Sei  $F: M \rightarrow N$  eine Abbildung, und sei  $U \subset M$  eine Teilmenge. Die Abbildung  $G: U \rightarrow M$  mit  $G(x) = x$  für alle  $x \in U$  heißt *Inklusion*. Sie ist stets injektiv. Die Verkettung

$$F|_U = F \circ G: U \rightarrow N$$

(lies „ $F$  eingeschränkt auf  $U$ “) heißt *Einschränkung* von  $F$  auf  $U$ .

1.23. SATZ. Seien  $L, M, N$  Mengen und  $F, F': M \rightarrow N$ ,  $G, G': L \rightarrow M$  Abbildungen. Dann gilt

- (1) Sind  $F, G$  injektiv, so ist auch  $F \circ G$  injektiv.
- (2) Sind  $F, G$  surjektiv, so ist auch  $F \circ G$  surjektiv.
- (3) Sind  $F, G$  bijektiv, so ist auch  $F \circ G$  bijektiv.
- (4) Ist  $F \circ G$  injektiv, so auch  $G$ .
- (5) Ist  $F \circ G$  surjektiv, so auch  $F$ .
- (6) Ist  $F$  injektiv, so folgt aus  $F \circ G = F \circ G'$  bereits  $G = G'$ .
- (7) Ist  $G$  surjektiv, so folgt aus  $F \circ G = F' \circ G$  bereits  $F = F'$ .

Hierbei bezeichnen  $F'$  und  $G'$  beliebige Abbildungen und nicht die „Ableitungen“ von  $F$  und  $G$ .

BEWEIS. Zu (1) seien  $x, y \in L$ . Aus  $(F \circ G)(x) = (F \circ G)(y)$  folgt  $F(G(x)) = F(G(y))$ , also  $G(x) = G(y)$  wegen Injektivität von  $F$ , also  $x = y$  wegen Injektivität von  $G$ , also ist  $F \circ G$  ebenfalls injektiv. Der Beweis von (2) verläuft ähnlich wie (1), und (3) folgt sofort aus (1) und (2).

Die Punkte (4), (5) sind Übungsaufgaben zur Vorlesung „Analysis I“ und werden hier daher nicht bewiesen.

Aussage (6) folgt ähnlich wie (7). Zu (7) sei  $y \in M$ . Wegen Surjektivität von  $G$  existiert  $x \in L$  mit  $G(x) = y$ . Aus  $F \circ G = F' \circ G$  folgt

$$F(y) = (F \circ G)(x) = (F' \circ G)(x) = F'(y).$$

Da das für alle  $y \in M$  gilt, folgt  $F = F'$ . □

1.24. SATZ. Sei  $F: M \rightarrow N$  bijektiv. Dann existiert genau eine Abbildung  $G: N \rightarrow M$  mit  $G \circ F = \text{id}_M$  und  $F \circ G = \text{id}_N$ .

1.25. DEFINITION. Die Abbildung  $G$  aus Satz 1.24 heißt die *Umkehrabbildung* von  $F$ .

Die Umkehrabbildung von  $F$  wird manchmal mit  $F^{-1}$  bezeichnet. Auch das kann zu Missverständnissen führen, so dass wir auf diese Bezeichnung verzichten wollen.

BEWEIS VON SATZ 1.24. Wir müssen zeigen, dass  $G$  *existiert*, und dass  $G$  *eindeutig* ist.

Zur Eindeutigkeit nehmen wir an, dass  $G: N \rightarrow M$  eine Umkehrfunktion ist. Dann sei  $y \in N$  beliebig, und sei  $x \in M$  das eindeutige Element mit  $F(x) = y$ . Aus  $G \circ F = \text{id}_M$  folgt

$$G(y) = G(F(x)) = x .$$

Wenn eine Umkehrfunktion existiert, sind ihre Werte durch diese Gleichung eindeutig bestimmt. Also ist die Umkehrfunktion eindeutig.

Zur Existenz sei  $\Gamma(F)$  der Graph von  $F$ . Gemäß der obigen Überlegung betrachten wir

$$X = \{ (y, x) \in N \times M \mid (x, y) \in \Gamma(F) \} ,$$

das ist eine Menge, da  $M, N$  Mengen sind und  $(x, y) \in \Gamma(F)$  eine wohldefinierte Eigenschaft ist. Zu jedem  $y \in N$  existiert genau ein  $x \in M$  mit  $F(x) = y$ , also mit  $(x, y) \in \Gamma(F)$ , also auch mit  $(y, x) \in X$ . Also ist  $X$  nach Definition 1.15 der Graph einer Funktion  $G: N \rightarrow M$ .

Für alle  $x \in M$  ist  $(F(x), x) \in X = \Gamma(G)$ , also  $G(F(x)) = x$ , und somit  $G \circ F = \text{id}_M$ . Umgekehrt sei  $y \in N$ , und sei  $x \in M$  das eindeutige Element mit  $F(x) = y$ , also  $G(y) = x$  und  $F(G(y)) = F(x) = y$ . Somit gilt auch  $F \circ G = \text{id}_N$ . Also existiert eine Umkehrfunktion, nämlich  $G$ .  $\square$

1.26. DEFINITION. Zwei Mengen  $M$  und  $N$  heißen *gleichmächtig*, wenn es eine bijektive Abbildung  $F: M \rightarrow N$  gibt.

1.27. BEMERKUNG. Gleichmächtige Mengen haben „gleich viele“ Elemente. Für alle Mengen  $L, M$  und  $N$  gilt (Übung):

- (1)  $M$  ist gleichmächtig zu  $M$ ;
- (2)  $N$  ist genau dann gleichmächtig zu  $M$ , wenn  $M$  zu  $N$  gleichmächtig ist;
- (3) sind  $L$  zu  $M$  und  $M$  zu  $N$  gleichmächtig, so ist auch  $L$  zu  $N$  gleichmächtig.

Das heißt, Gleichmächtigkeit verhält sich wie eine Äquivalenzrelation, siehe Definition 1.42. Allerdings sollte eine Relation immer auf einer Menge definiert sein, und die Menge aller Mengen gibt es nach Folgerung 1.14 nicht.

1.28. BEISPIEL. (1) Die Mengen  $M = \{1, 2, 3\}$  und  $N = \{4, 7, 15\}$  sind gleichmächtig. Definiere z.B.  $F: M \rightarrow N$  durch

$$F(1) = 7, \quad F(2) = 4, \quad F(3) = 15.$$

- (2) Sei  $M = \{n^2 \mid n \in \mathbb{N}\} \subset \mathbb{N}$  die Menge der Quadratzahlen. Da  $F: \mathbb{N} \rightarrow M$  mit  $F(n) = n^2$  bijektiv ist, sind  $M$  und  $\mathbb{N}$  gleichmächtig, obwohl  $M$  eine echte Teilmenge von  $\mathbb{N}$  ist.

## 1.2. Natürliche Zahlen

Die natürlichen Zahlen sind uns bereits seit unserer Kindheit vertraut — wir benutzen sie zum Zählen. Für den Fall, dass es nichts zu zählen gibt, haben wir die Zahl 0. Es ist erstaunlich, dass die Zahl 0 selbst erst spät als eigenständige Zahl eingeführt wurde. Wenn wir schon ein Stück weit gezählt haben, etwa bis zu einer Zahl  $n$ , und weiterzählen wollen, brauchen wir die nächste Zahl. Wir nennen Sie den Nachfolger von  $n$  und schreiben  $\mathcal{N}f(n) = n+1$ . Schließlich wollen wir, dass die natürlichen Zahlen eine Menge bilden, die sonst keine weiteren Objekte enthält.

1.29. ANNAHME (Peano-Axiome). *Wir nehmen an, dass es eine Menge  $\mathbb{N}$  mit einem ausgezeichneten Element  $0 \in \mathbb{N}$  und einer Abbildung  $\mathcal{N}f: \mathbb{N} \rightarrow \mathbb{N}$  gibt, die die folgenden Peano-Axiome erfüllt:*

- (P1) *Die Nachfolger-Abbildung ist bijektiv als Abbildung  $\mathcal{N}f: \mathbb{N} \rightarrow \mathbb{N} \setminus \{0\}$ .*  
 (P2) *Prinzip der vollständigen Induktion. Sei  $M \subset \mathbb{N}$  mit  $0 \in M$ , so dass für alle  $m \in M$  auch  $\mathcal{N}f(m) \in M$  gilt, dann ist bereits  $M = \mathbb{N}$ .*

Axiom (P1) besagt, dass jede Zahl genau einen Nachfolger hat, und jede Zahl außer 0 selbst Nachfolger genau einer anderen Zahl ist. Axiom (P2) besagt, dass die Menge  $\mathbb{N}$  die “kleinste” Menge ist, die (P1) erfüllt. Trotzdem bestimmen die Peano-Axiome die natürlichen Zahlen nicht eindeutig — warum das so ist, lernen Sie aber erst in einer Vorlesung über Logik. Wir wollen immerhin annehmen, dass  $\mathbb{N}$  nur die Zahlen  $0, 1, 2, \dots$  enthält, aber keine weiteren Elemente. Übrigens gibt es Autoren, für die 0 nicht zu  $\mathbb{N}$  gehört. Zur Sicherheit können Sie beide Versionen mit  $\mathbb{N}_0$  und  $\mathbb{N}_>$  bezeichnen.

1.30. BEMERKUNG. Wir können natürliche Zahlen als Mengen  $\underline{0}, \underline{1}, \underline{2}, \dots$  konstruieren. Dazu setzen wir  $\underline{0} = \emptyset$  und konstruieren Nachfolger als

$$\mathcal{N}f(\underline{n}) = \underline{n+1} = \{\underline{0}, \dots, \underline{n}\} = \underline{n} \cup \{\underline{n}\}.$$

Diese Definition ist *rekursiv*, das heißt, man muss alle Zahlen bis  $\underline{n}$  kennen, um den Nachfolger  $\underline{n+1}$  zu konstruieren. Wir schreiben  $\underline{\mathbb{N}} = \{\underline{0}, \underline{1}, \underline{2}, \dots\}$ .

Die ersten „Zahlen“ sehen so aus:

$$\begin{aligned} \underline{0} &= \emptyset \\ \underline{1} &= \{\underline{0}\} = \{\emptyset\} \\ \underline{2} &= \{\underline{0}, \underline{1}\} = \{\emptyset, \{\emptyset\}\} \\ \underline{3} &= \{\underline{0}, \underline{1}, \underline{2}\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \end{aligned}$$

Auf diese Weise erhalten wir alle Zahlen mit elementaren Konstruktionen aus der leeren Menge. Da das recht mühselig ist, werden wir natürliche Zahlen meistens als Zahlen und nicht als Mengen betrachten. Nur in diesem Abschnitt werden wir die obigen Mengen manchmal benutzen.

1.31. DEFINITION. Eine Menge  $M$  heißt *endlich*, wenn sie zu einer Menge  $\underline{n}$  gleichmächtig ist. In diesem Fall heißt die Zahl  $n$  die *Mächtigkeit* von  $M$ , geschrieben  $n = \#M$ . Ansonsten heißt  $M$  *unendlich*.

1.32. BEMERKUNG. (1) Man kann sich überlegen, dass zwei Mengen  $\underline{n}$  und  $\underline{m}$  genau dann gleichmächtig sind, wenn  $\underline{n} = \underline{m}$ . Wegen Bemerkung 1.27 kann jede Menge  $M$  zu höchstens einer Menge  $\underline{n}$  gleichmächtig sein. Die Schreibweise  $\#M$  für endliche Mengen ist also sinnvoll.

(2) Endliche Mengen kann man immer als Aufzählung angeben. Sei etwa  $F: \underline{n} \rightarrow M$  bijektiv, dann schreibe

$$M = \{F(0), \dots, F(n-1)\}$$

Ist  $M$  umgekehrt als  $\{x_1, \dots, x_n\}$  gegeben, dann hat  $M$  höchstens  $n$  Elemente, ist also endlich.

(3) Für unendliche Mengen  $M$  führen wir die Schreibweise „ $\#M = \infty$ “ nicht ein, da nicht alle unendlichen Mengen gleichmächtig sind. Wir schreiben aber „ $\#M < \infty$ “, wenn  $M$  endlich ist.

1.33. DEFINITION. Es seien  $m, n \in \mathbb{N}$ , dann gilt  $m \leq n$  genau dann, wenn  $\underline{m} \subset \underline{n}$ . Es ist  $m$  *kleiner* als  $n$ , kurz  $m < n$ , wenn  $m \leq n$  und  $m \neq n$  gilt.

1.34. BEMERKUNG. Aus Bemerkung 1.30 folgt auch, dass  $m < n$  genau dann gilt, wenn  $\underline{m} \in \underline{n}$ . Man beachte den Unterschied in der Notation. Bei „ $\subset$ “ ist Gleichheit erlaubt, bei „ $<$ “ jedoch ausgeschlossen.

Der Vergleich von Zahlen führt uns auf den Begriff der Ordnung. Eine Ordnung einer Menge  $M$  ist eine *Relation*, das heißt, eine Teilmenge  $R \subset M \times M$ , die einige zusätzliche Eigenschaften besitzt. Wir sagen „es gilt  $xRy$ “ für  $x, y \in M$ , wenn  $(x, y) \in R$ .

1.35. DEFINITION. Eine Relation  $R$  auf eine Menge  $M$  heißt *Halbordnung*, wenn für alle  $x, y, z \in M$  gilt:

- (O1)  $xRx$  (Reflexivität),  
(O2)  $xRy$  und  $yRx \implies x = y$  (Antisymmetrie),  
(O3)  $xRy$  und  $yRz \implies xRz$  (Transitivität).

Eine Halbordnung heißt *Ordnung*, wenn ausserdem für alle  $x, y \in M$  gilt:

- (O4)  $xRy$  oder  $yRx$  (Totalität).

Die Eigenschaften (O1)–(O4) heißen auch *Ordnungsaxiome*.

1.36. BEISPIEL. (1) Sei  $M$  eine Menge, dann definiert „ $\subset$ “ eine Halbordnung auf der Potenzmenge  $\mathcal{P}(M)$ , denn für alle  $A, B, C \subset M$  gilt

$$\begin{aligned} A &\subset A, \\ A \subset B \text{ und } B \subset A &\implies A = B, \\ A \subset B \text{ und } B \subset C &\implies A \subset C. \end{aligned}$$

- (2) Die Relation „ $\in$ “ ist nicht transitiv und daher keine Halbordnung, denn es gilt zum Beispiel  $a \in \{a, b\}$  und  $\{a, b\} \in \{\{a\}, \{a, b\}\}$ , aber nicht  $a \in \{\{a\}, \{a, b\}\}$ .
- (3) Die Relation „ $\leq$ “ auf  $\mathbb{N}$  ist eine Ordnung. Nach Bemerkung 1.30 gilt  $\underline{\mathbb{N}} \subset \mathcal{P}(\underline{\mathbb{N}})$ , und nach Definition 1.33 entspricht „ $\leq$ “ der Einschränkung von „ $\subset$ “ auf  $\underline{\mathbb{N}}$ . Wegen (1) ist „ $\subset$ “ eine Halbordnung auf  $\underline{\mathbb{N}}$ , also ist „ $\leq$ “ eine Halbordnung auf  $\mathbb{N}$ . Zu zeigen wäre noch, dass für alle  $m, n \in \mathbb{N}$  gilt

$$n \leq m \text{ und } m \leq n \implies m = n .$$

- (4) Sei  $M$  eine Menge. Die Relation „hat höchstens so viele Elemente wie“ ist keine Ordnung auf der Potenzmenge  $\mathcal{P}(M)$ , denn sei  $M = \{1, 2, 3\}$ , dann hat  $\{1, 2\}$  höchstens so viele Elemente wie  $\{2, 3\}$  und umgekehrt, aber beide Mengen sind nicht gleich. Also ist die Antisymmetrie verletzt.

Das zweite Peano-Axiom 1.29 (P2) führt uns zur Beweismethode durch vollständige Induktion. Wir benötigen sie bald zum Rechnen.

1.37. SATZ (Vollständige Induktion). *Für jedes  $n \in \mathbb{N}$  sei  $A(n)$  eine Aussage. Wenn gilt*

- (1)  $A(0)$  ist wahr, und
- (2) aus  $A(n)$  folgt  $A(n+1)$  für alle  $n \in \mathbb{N}$ ,

dann ist  $A(n)$  für alle  $n \in \mathbb{N}$  wahr.

BEWEIS. Betrachte

$$M = \{n \in \mathbb{N} \mid \text{die Aussage } A(n) \text{ ist wahr}\}.$$

Nach unseren Annahmen in Abschnitt 1.1 ist das wieder eine Menge, also  $M \subset \mathbb{N}$ . Aus den Voraussetzungen folgt

- (1)  $0 \in M$ , und
- (2) für alle  $n \in M$  gilt  $n+1 \in M$ .

Aus dem Axiom (P2) folgt dann  $M = \mathbb{N}$ . Nach Definition von  $M$  gilt  $A(n)$  also für alle  $n \in \mathbb{N}$ .  $\square$

Eine andere Art der vollständigen Induktion funktioniert so: Wenn gilt

- (1)  $A(0)$  ist wahr, und
- (2) aus  $A(0) \wedge \dots \wedge A(n)$  folgt  $A(n+1)$  für alle  $n \in \mathbb{N}$ ,

dann gilt  $A(n)$  für alle  $n \in \mathbb{N}$ . Das zeigt man, indem man die Aussage

$$B(n) = A(0) \wedge \dots \wedge A(n)$$

induktiv mit Satz 1.37 beweist.

Wir haben in Bemerkung 1.30 Zahlen als Mengen rekursiv eingeführt. *Rekursive Definitionen* funktionieren ähnlich wie vollständige Induktion: um eine

Abbildung  $F$  von  $\mathbb{N}$  in eine Menge  $M$  anzugeben, reicht es  $F(0) \in M$  festzulegen und eine Vorschrift anzugeben, die  $F(n+1)$  aus  $F(0), \dots, F(n)$  bestimmt.

Wir führen jetzt die Grundrechenarten auf  $\mathbb{N}$  rekursiv ein. Hierbei handelt es sich um *Verknüpfungen*, das heißt, um Abbildungen  $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ , etwa

$$+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} \quad \text{mit} \quad +(m, n) = m + n.$$

1.38. DEFINITION. Die *Addition*, *Multiplikation* und *Potenzierung* sind für  $m, n \in \mathbb{N}$  definiert durch

- (1)  $m + 0 = m$                       und                       $m + \mathcal{N}f(n) = \mathcal{N}f(m + n)$ ,
- (2)  $m \cdot 0 = 0$                       und                       $m \cdot \mathcal{N}f(n) = m \cdot n + m$ ,
- (3)  $m^0 = 1$                             und                       $m^{\mathcal{N}f(n)} = m^n \cdot m$ .

BEISPIEL. Zwei einfache Rechnungen:

$$3 + 2 = 3 + \mathcal{N}f(1) = \mathcal{N}f(3 + 1) = \mathcal{N}f(3 + \mathcal{N}f(0)) = \mathcal{N}f(\mathcal{N}f(3)) = \mathcal{N}f(4) = 5,$$

$$3 \cdot 2 = 3 \cdot \mathcal{N}f(1) = 3 \cdot 1 + 3 = 3 \cdot \mathcal{N}f(0) + 3 = 3 \cdot 0 + 3 + 3 = 0 + 3 + 3 = 6.$$

1.39. PROPOSITION. *Seien  $M, N$  endliche Mengen.*

- (1) Falls  $M \cap N = \emptyset$  ist, gilt  $\#(M \dot{\cup} N) = \#M + \#N$ .
- (2) Es gilt  $\#(M \times N) = \#M \cdot \#N$ .
- (3) Es gilt  $\#\text{Abb}(N, M) = \#M^{\#N}$ .

BEWEIS. Wir beweisen (1) zur Illustration durch vollständige Induktion über die Mächtigkeit  $n = \#N$ . Es sei  $m = \#M$ .

*Induktionsanfang:* Es sei  $n = 0$ . Nach den Definitionen 1.26 und 1.31 existiert eine bijektive Abbildung von  $\emptyset = \underline{0}$  nach  $N$ , also gilt  $N = \emptyset$ . Somit

$$\#(M \dot{\cup} N) = \#(M \dot{\cup} \emptyset) = \#M = m = m + 0 = \#M + \#N.$$

*Induktionsschritt:* Es sei  $\#N = n + 1$ . Dann existiert eine bijektive Abbildung  $F: \underline{n+1} = \underline{n} \dot{\cup} \{\underline{n}\} \rightarrow N$ . Setze

$$N' = \text{im}(F|_{\underline{n}}) = \{F(\underline{0}), \dots, F(\underline{n-1})\} \quad \text{und} \quad x = F(\underline{n}),$$

so dass  $\#N' = n$ . Nach Induktionsvoraussetzung gilt  $\#(M \dot{\cup} N') = m + n$ , also existiert eine bijektive Abbildung  $G': \underline{m+n} \rightarrow M \dot{\cup} N'$ . Wir definieren  $G: \underline{(m+n)+1} \rightarrow M \dot{\cup} N$  durch

$$G(\underline{k}) = \begin{cases} G'(\underline{k}) & \text{falls } \underline{k} \in \underline{m+n}, \text{ also } k < m+n, \text{ und} \\ x & \text{falls } \underline{k} = \underline{m+n}, \text{ also } k = m+n. \end{cases}$$

Man überzeugt sich leicht, dass  $G$  bijektiv ist. Mit Definition 1.38 (1) folgt

$$\#(M \dot{\cup} N) = (m+n) + 1 = m + (n+1) = \#M + \#N. \quad \square$$

1.40. BEMERKUNG. Die Grundrechenarten hätten wir auch über die Eigenschaften (1)–(3) definieren können. Außerdem folgt aus (1), dass  $m \leq \ell$  genau dann gilt, wenn ein  $n \in \mathbb{N}$  mit  $m + n = \ell$  existiert.

Bevor wir das Assoziativgesetz kennenlernen, überlegen wir uns, was „Klammern“ eigentlich bewirken. Fassen wir  $+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  als Abbildung auf, dann bedeutet  $(\ell + m) + n$  gerade  $+(+(\ell, m), n)$ ,  $\ell + (m + n)$  bedeutet  $+(\ell, +(m, n))$ .

1.41. SATZ. Für  $\ell, m, n \in \mathbb{N}$  gelten die Rechenregeln

(1) Assoziativgesetze

$$(\ell + m) + n = \ell + (m + n)$$

$$(\ell \cdot m) \cdot n = \ell \cdot (m \cdot n)$$

(2) Neutrale Elemente

$$n + 0 = n$$

$$n \cdot 1 = n$$

(3) Kommutativgesetze

$$n + m = m + n$$

$$n \cdot m = m \cdot n$$

(4) Distributivgesetz

$$\ell \cdot (m + n) = \ell \cdot m + \ell \cdot n$$

(5) Kürzungsregeln

$$\ell + n = m + n \quad \implies \quad \ell = m$$

$$\ell \cdot n = m \cdot n \quad \implies \quad \ell = m \text{ oder } n = 0.$$

BEWEIS. Die Aussagen (2) folgen leicht aus Definition 1.38. Alle anderen lassen sich durch vollständige Induktion beweisen. Der Beweis von (5) ist Übung.  $\square$

### 1.3. Ganze und Rationale Zahlen

In diesem Abschnitt „lösen“ wir zwei Probleme: man kann in  $\mathbb{N}$  nicht subtrahieren, und man kann in  $\mathbb{N}$  auch nicht durch Zahlen  $n \neq 0$  dividieren. Um diese „Grundrechenarten“ einführen zu können, werden wir  $\mathbb{N}$  erst zu den ganzen Zahlen  $\mathbb{Z}$ , und dann zu den rationalen Zahlen  $\mathbb{Q}$  erweitern. Dazu ist zunächst etwas Vorarbeit nötig.

1.42. DEFINITION. Eine Relation  $R$  auf einer Menge  $M$  heißt *Äquivalenzrelation*, wenn für alle  $x, y, z$  gilt:

$$(\text{Ä1}) \quad xRx \quad (\text{Reflexivität}),$$

$$(\text{Ä2}) \quad xRy \implies yRx \quad (\text{Symmetrie}),$$

$$(\text{Ä3}) \quad xRy \text{ und } yRz \implies xRz \quad (\text{Transitivität}).$$

Im Unterschied zu Halbordnungen (Definition 1.35) sind Äquivalenzrelationen symmetrisch und nicht antisymmetrisch. Das erlaubt uns, Äquivalenzklassen und Quotientenmengen zu definieren. Wir erinnern uns an die Potenzmenge  $\mathcal{P}(M)$  von  $M$  aus Definition 1.12.

1.43. DEFINITION. Es sei  $R$  eine Äquivalenzrelation auf  $M$ . Für alle  $x \in M$  definieren wir die ( $R$ -) Äquivalenzklasse  $[x]$  von  $x$  als

$$[x] = \{ y \in M \mid xRy \} .$$

Die Gesamtheit aller Äquivalenzklassen bildet die *Quotientenmenge* (kurz: den *Quotienten*)  $M/R$ , also

$$M/R = \{ [x] \mid x \in M \} \subset \mathcal{P}(M) ,$$

und alle Elemente  $y \in [x]$  heißen *Repräsentanten* von  $[x] \in M/R$ . Die Abbildung  $p: M \rightarrow M/R$  mit  $p(x) = [x]$  heißt *Quotientenabbildung*.

Das einfachste Beispiel für eine Äquivalenzrelation ist die Gleichheit „ $=$ “ auf einer beliebigen Menge  $M$ . Die Axiome (Ä1)–(Ä3) gelten offensichtlich. In diesem Fall ist die Äquivalenzklasse von  $x \in M$  gerade  $[x] = \{x\}$ , und die Quotientenabbildung  $p: M \rightarrow M/=$  ist bijektiv mit  $x \mapsto \{x\}$ . Allerdings gilt strenggenommen nicht  $M = M/=$ , zum Beispiel ist

$$\{1, 2, 3\}/= = \{\{1\}, \{2\}, \{3\}\} .$$

Sei  $M$  eine beliebige Menge. Nach Bemerkung 1.27 definiert Gleichmächtigkeit eine Äquivalenzrelation  $R$  auf der Potenzmenge  $\mathcal{P}(M)$ .

1.44. PROPOSITION. *Es sei  $R$  eine Äquivalenzrelation auf  $M$ .*

- (1) *Für alle  $x \in M$  und alle  $y \in [x]$  gilt  $[x] = [y]$ , insbesondere liegt jedes  $x \in M$  in genau einer Äquivalenzklasse von  $R$ .*
- (2) *Die Abbildung  $p: M \rightarrow M/R$  ist surjektiv, und es gilt  $p(x) = p(y)$  genau dann, wenn  $xRy$  gilt.*
- (3) *Es sei  $F: M \rightarrow N$  eine Abbildung. Dann existiert genau dann eine Abbildung  $\bar{F}: M/R \rightarrow N$  mit  $F = \bar{F} \circ p$ , wenn für alle  $x, y \in M$  aus  $xRy$  folgt, dass  $F(x) = F(y)$ . In diesem Fall ist  $\bar{F}$  eindeutig.*

Die Aussage (3) heißt auch die *universelle Eigenschaft des Quotienten*. Wir nennen  $\bar{F}$  die *von  $F$  induzierte Abbildung*. Wir stellen (3) als Diagramm dar:

$$\begin{array}{ccc} M & \xrightarrow{F} & N \\ p \downarrow & \nearrow \bar{F} & \\ M/R & & \end{array}$$

BEWEIS. Zu (1) seien  $y \in [x]$  und  $z \in [y]$  beliebig, dann gilt  $xRy$  und  $yRz$ . Aus Transitivität folgt  $xRz$ , also gilt  $z \in [x]$  für alle  $z \in [y]$ , es folgt  $[y] \subset [x]$ .

Aus  $xRy$  folgt  $yRx$  wegen der Symmetrie von  $R$ , also folgt  $x \in [y]$  aus  $y \in [x]$ . Nach obigem Argument gilt also auch  $[x] \subset [y]$ , und somit  $[x] = [y]$ .

Die Surjektivität von  $p$  ist klar nach Definition von  $M/R$ , und aus (1) folgt, dass  $p(x) = [x] = [y] = p(y)$  genau dann, wenn  $xRy$  gilt. Also stimmt (2).

In (3) beginnen wir mit „ $\implies$ “. Sei also  $\bar{F}: M/R \rightarrow N$  gegeben mit  $F = \bar{F} \circ p$ , und seien  $x, y \in M$  gegeben mit  $xRy$ . Aus (2) folgt  $p(x) = p(y)$ , also erst recht

$$F(x) = \bar{F}(p(x)) = \bar{F}(p(y)) = F(y) .$$

Zu „ $\impliedby$ “ gelte  $F(x) = F(y)$  für alle  $x, y \in M$  mit  $xRy$ , also für alle  $x \in M$  und alle  $y \in [x]$ . Seien also  $[x] \in M/R$  und  $y \in [x]$  beliebig, dann dürfen wir  $\bar{F}([x]) = F(y)$  setzen. Diese Konstruktion hängt nach Voraussetzung nicht von der Wahl von  $y \in [x]$  ab. Dazu sagen wir,  $\bar{F}$  ist *wohldefiniert*.

Die Eindeutigkeit von  $\bar{F}$  folgt mit Satz 1.23 (7) aus der Surjektivität von  $p$ .  $\square$

In der Schule definiert man  $\mathbb{Z}$ , indem man zu  $\mathbb{N}$  noch negative Zahlen hinzunimmt:

$$\mathbb{Z} = \mathbb{N} \cup \{ -n \mid n \in \mathbb{N} \setminus \{0\} \} .$$

Anschließend definiert man Addition, Subtraktion und Multiplikation. Dabei muss man immer einige Fälle unterscheiden. Wir beschreiben ganze Zahlen stattdessen als Differenzen natürlicher Zahlen, also als  $m - n$  für  $m, n \in \mathbb{N}$ .

1.45. BEMERKUNG. Um die folgenden Konstruktionen zu verstehen, hier ein paar Vorüberlegungen. Für alle  $m, n, p, q \in \mathbb{N}$  gilt in  $\mathbb{Z}$ :

- (1)  $(m - n) = (p - q) \in \mathbb{Z} \iff m + q = n + p \in \mathbb{N} ,$
- (2)  $(m - n) + (p - q) = (m + p) - (n + q) ,$
- (3)  $-(m - n) = n - m ,$
- (4)  $(m - n) \cdot (p - q) = (m \cdot p + n \cdot q) - (m \cdot q + n \cdot p) ,$
- (5)  $(m - n) \leq (p - q) \iff m + q \leq n + p .$

Für eine Menge  $M$  und  $n \in \mathbb{N}$  bezeichne  $M^n$  das  $n$ -fache kartesische Produkt von  $M$  mit sich selbst, etwa  $M^2 = M \times M$ . Anstelle von  $m - n \in \mathbb{Z}$  betrachten wir das Paar  $(m, n) \in \mathbb{N}^2$ . Gemäß Bemerkung 1.45 (1) definieren wir eine Relation  $\sim$  auf der Menge  $\mathbb{N}^2$  durch

$$(m, n) \sim (p, q) \iff m + q = n + p \in \mathbb{N} .$$

Außerdem definieren wir Addition, Negatives, Multiplikation und eine Relation  $\leq$  gemäß Bemerkung 1.45 (2)–(5) durch

$$\begin{aligned} (m, n) + (p, q) &= (m + p, n + q), \\ -(m, n) &= (n, m), \\ (m, n) \cdot (p, q) &= (m \cdot p + n \cdot q, m \cdot q + n \cdot p), \\ (m, n) \leq (p, q) &\iff m + q \leq n + p. \end{aligned}$$

1.46. PROPOSITION. *Es seien  $m, n, p, q, r, s, t, u \in \mathbb{N}$ . Dann gilt*

- (1) „ $\sim$ “ *ist eine Äquivalenzrelation.*

(2) Aus  $(m, n) \sim (p, q)$  und  $(r, s) \sim (t, u)$  folgt

$$(m, n) + (r, s) \sim (p, q) + (t, u),$$

$$(m, n) \cdot (r, s) \sim (p, q) \cdot (t, u)$$

$$\text{und } -(m, n) \sim -(p, q).$$

(3) Aus  $(m, n) \sim (p, q)$  und  $(r, s) \sim (t, u)$  folgt

$$(m, n) \leq (r, s) \implies (p, q) \leq (t, u).$$

BEWEIS. Zu (1): „ $\sim$ “ ist reflexiv und symmetrisch nach Konstruktion und dem Kommutativgesetz 1.41 (3). Zur Transitivität benutzen wir zusätzlich die Kürzungsregel 1.41 (5):

$$\begin{aligned} & (m, n) \sim (p, q) \text{ und } (p, q) \sim (r, s) \\ \implies & m + q = n + p \text{ und } p + s = q + r \\ \implies & m + q + p + s = n + p + q + r \\ \implies & m + s = n + r \\ \implies & (m, n) \sim (r, s). \end{aligned}$$

Zu (2): Seien  $(m, n) \sim (p, q)$  und  $(r, s) \sim (t, u)$ , also  $m + q = n + p$  und  $r + u = s + t$ . Wegen  $m + q + r + u = n + p + s + t$  folgt

$$(m, n) + (r, s) = (m + r, n + s) \sim (p + t, q + u) = (p, q) + (t, u).$$

Wir zeigen als nächstes  $(m, n) \cdot (r, s) \sim (p, q) \cdot (r, s)$  mit der Rechnung

$$\begin{aligned} mr + ns + ps + qr &= (m + q) \cdot r + (n + p) \cdot s \\ &= (n + p) \cdot r + (m + q) \cdot s = pr + qs + ms + nr, \end{aligned}$$

also

$$(m, n)(r, s) = (mr + ns, ms + nr) \sim (pr + qs, ps + qr) = (p, q)(r, s).$$

Genauso zeigt man  $(p, q)(r, s) \sim (p, q)(t, u)$ , und wegen Transitivität gilt  $(m, n)(r, s) \sim (p, q)(t, u)$ . Die Behauptung  $-(m, n) = (n, m) \sim (q, p) = -(p, q)$  ist leicht einzusehen.

Zu (3): Mit  $(m, n) \sim (p, q)$  und  $(r, s) \sim (t, u)$  wie oben: Aus  $(m, n) \leq (r, s)$  folgt  $m + s \leq n + r$ , also existiert nach Bemerkung 1.40 ein  $k \in \mathbb{N}$  mit

$$\begin{aligned} m + s + k &= n + r \\ \implies m + p + s + u + k &= n + p + r + u = m + q + s + t \\ \implies p + u + k &= q + t \implies p + u \leq q + t \\ \implies (p, q) &\leq (t, u). \quad \square \end{aligned}$$

Wir definieren also  $\mathbb{Z}$  als Quotienten

$$\mathbb{Z} = \mathbb{N}^2 / \sim = \{ [(m, n)] \mid (m, n) \in \mathbb{N}^2 \}.$$

Proposition 1.46 garantiert wegen der universellen Eigenschaft aus 1.44 (3), dass wir mit Äquivalenzklassen rechnen dürfen:

$$\begin{aligned} [(m, n)] + [(p, q)] &= [(m + p, n + q)], \\ [(m, n)] \cdot [(p, q)] &= [(mp + nq, mq + np)], \\ -[(m, n)] &= [(n, m)], \end{aligned}$$

unabhängig von den Repräsentanten  $(m, n) \in [(m, n)]$ ,  $(p, q) \in [(p, q)]$ . Auch  $[(m, n)] \leq [(p, q)]$  ist wohldefiniert.

Konkreter sei  $p: \mathbb{N}^2 \rightarrow \mathbb{Z}$  die Quotientenabbildung. Wir halten zunächst das Paar  $(r, s) \in \mathbb{N}^2$  fest und betrachten die Abbildung  $F = p \circ (\cdot + (r, s))$  wie im folgenden Diagramm:

$$\begin{array}{ccc} \mathbb{N}^2 & \xrightarrow{\cdot + (r, s)} & \mathbb{N}^2 \\ p \downarrow & \searrow F & \downarrow p \\ \mathbb{Z} & \xrightarrow{\bar{F}} & \mathbb{Z} \end{array}$$

Also können wir zu einer ganzen Zahl ein festes Paar  $(r, s)$  addieren. Jetzt halten wir die ganze Zahl  $[(m, n)]$  fest und betrachten die Abbildung  $G = [(m, n)] + \cdot: \mathbb{N}^2 \rightarrow \mathbb{Z}$  wie im folgenden Diagramm:

$$\begin{array}{ccc} \mathbb{N}^2 & & \\ p \downarrow & \searrow [(m, n)] + \cdot & \\ \mathbb{Z} & \xrightarrow{\quad} & \mathbb{Z} \end{array}$$

Wir setzen das zu einem größeren Diagramm zusammen

$$\begin{array}{ccc} \mathbb{N}^2 \times \mathbb{N}^2 & \xrightarrow{+} & \mathbb{N}^2 \\ p \times \text{id}_{\mathbb{N}^2} \downarrow & & \downarrow p \\ \mathbb{Z} \times \mathbb{N}^2 & \xrightarrow{+} & \mathbb{Z} \\ \text{id}_{\mathbb{Z}} \times p \downarrow & & \downarrow \\ \mathbb{Z} \times \mathbb{Z} & \xrightarrow{+} & \mathbb{Z} \end{array}$$

dabei bleibt im oberen Parallelogramm das zweite Argument in  $\mathbb{N}^2$  unverändert, und im unteren das erste Argument, jetzt aber in  $\mathbb{Z}$ . Also dürfen wir zwei ganze Zahlen addieren. Mit den analogen Diagrammen erhalten wir auch die Multiplikation.

1.47. DEFINITION. Die Menge  $\mathbb{Z} = \mathbb{N}^2 / \sim$  heißt Menge der *ganzen Zahlen*.

Wir identifizieren  $n \in \mathbb{N}$  mit  $[(n, 0)] \in \mathbb{Z}$  und schreiben  $-n$  für  $[(0, n)] \in \mathbb{Z}$ . Insbesondere schreiben wir  $0 = [(0, 0)]$  und  $1 = [(1, 0)]$ .

1.48. SATZ. In  $\mathbb{Z}$  gelten Assoziativ- und Kommutativgesetz sowohl für die Addition als auch für die Multiplikation. Neutrale Elemente sind 0 für die Addition und 1 für die Multiplikation. Es gilt das Distributivgesetz. Jedes Element

$[(m, n)]$  besitzt ein additives Inverses  $-[(m, n)] = [(n, m)]$ , das heißt, es gilt

$$[(m, n)] + (-[(m, n)]) = [(m, n)] + [(n, m)] = [(0, 0)].$$

Es gilt die Kürzungsregel für die Multiplikation.

Die Relation „ $\leq$ “ auf  $\mathbb{Z}$  ist eine Ordnung, und für alle  $a, b, c \in \mathbb{Z}$  gilt:

$$\begin{aligned} a \leq b &\implies a + c \leq b + c, \\ 0 \leq a \text{ und } 0 \leq b &\implies 0 \leq ab. \end{aligned}$$

BEWEIS. Das meiste folgt direkt aus Satz 1.41 und den obigen Definitionen. Die neue Gleichung

$$[(m, n)] + (-[(m, n)]) = [(m, n)] + [(n, m)] = [(0, 0)]$$

ergibt sich aus

$$(m, n) + (n, m) = (m + n, n + m) \sim (0, 0).$$

Ähnlich zeigt man die Eigenschaften von „ $\leq$ “. □

Wir haben die natürlichen Zahlen  $\mathbb{N}$  zu den ganzen Zahlen  $\mathbb{Z}$  erweitert, um additive Inverse zu finden, also Zahlen  $-n$  mit  $n + (-n) = 0$ . Dazu haben wir natürliche Zahlen durch Paare  $(m, n) \in \mathbb{N} \times \mathbb{N}$  ersetzt, die für die Zahl  $m - n \in \mathbb{Z}$  stehen. Die Zahlen  $-n = [(0, n)]$  sind gerade die negativen Zahlen aus der Schule. Der Einfachheit halber schreiben wir ab sofort  $a, b, c, \dots \in \mathbb{Z}$ , nicht mehr  $[(m, n)]$ .

Um nun auch multiplikative Inverse  $\frac{1}{n}$  mit  $n \cdot \frac{1}{n} = 1$  für alle  $n \in \mathbb{Z} \setminus \{0\}$  zu erhalten, ersetzen wir ganze Zahlen durch Paare  $(p, q) \in \mathbb{Z} \times (\mathbb{N} \setminus \{0\})$ , die für Brüche  $\frac{p}{q}$  stehen. Das ist die Bruchrechnung, wie wir sie aus der Schule kennen.

Dazu definieren wir für  $(p, q), (r, s) \in \mathbb{Z} \times (\mathbb{N} \setminus \{0\})$ :

$$\begin{aligned} (p, q) \approx (r, s) &\iff p \cdot s = q \cdot r && \left( \iff \frac{p}{q} = \frac{r}{s} \right), \\ (p, q) + (r, s) &= (ps + qr, qs) && \left( \text{da } \frac{p}{q} + \frac{r}{s} = \frac{ps + qr}{qs} \right), \\ (p, q) \cdot (r, s) &= (pr, qs) && \left( \text{da } \frac{p}{q} \cdot \frac{r}{s} = \frac{pr}{qs} \right), \\ -(p, q) &= (-p, q) && \left( \text{da } -\frac{p}{q} = \frac{-p}{q} \right), \\ (p, q) \leq (r, s) &\iff p \cdot s \leq q \cdot r && \left( \iff \frac{p}{q} \leq \frac{r}{s}, \text{ da } q, s > 0 \right). \end{aligned}$$

Beachte, dass  $qs \in \mathbb{N} \setminus \{0\}$ , denn aus  $qs = 0 = 0 \cdot s$  würde mit der Kürzungsregel entweder  $q = 0$  oder  $s = 0$  folgen. Für  $p \neq 0$  definieren wir:

$$(p, q)^{-1} = \begin{cases} (q, p) & \text{falls } p > 0, \\ (-q, -p) & \text{falls } p < 0. \end{cases}$$

Beachte: die rechte Seite  $(\pm q, \pm p)$  liegt immer in  $\mathbb{Z} \times (\mathbb{N} \setminus \{0\})$ .

1.49. PROPOSITION. (1) Die Relation „ $\approx$ “ ist eine Äquivalenzrelation.

(2) Es seien  $(m, n), (p, q), (r, s), (t, u) \in \mathbb{Z} \times (\mathbb{N} \setminus \{0\})$  mit  $(m, n) \approx (p, q)$  und  $(r, s) \approx (t, u)$  gegeben, dann gilt

$$(m, n) + (r, s) \approx (p, q) + (t, u),$$

$$(m, n) \cdot (r, s) \approx (p, q) \cdot (t, u),$$

und es gilt  $m \neq 0 \Rightarrow p \neq 0$ , und in diesem Fall

$$(m, n)^{-1} \approx (p, q)^{-1}.$$

(3) Unter den gleichen Voraussetzungen wie in (2) gilt

$$(m, n) \leq (r, s) \Rightarrow (p, q) \leq (t, u).$$

BEWEIS. Die Beweismethode ist die gleiche wie bei Proposition 1.46, wir lassen den Beweis daher aus, ein Teil ist Übung.  $\square$

1.50. DEFINITION. Der Quotient  $\mathbb{Z} \times (\mathbb{N} \setminus \{0\}) / \approx$  heißt Menge der *rationalen Zahlen* und wird mit  $\mathbb{Q}$  bezeichnet. Für die Äquivalenzklasse  $[(p, q)]$  schreiben wir  $\frac{p}{q}$ .

Wie zuvor schließen wir aus Proposition 1.49 (2), dass wir mit Brüchen so rechnen dürfen, wie wir es aus der Schule kennen. Proposition 1.49 (3) besagt, dass wir zwei Brüche vergleichen können.

Wir identifizieren eine ganze Zahl  $n \in \mathbb{Z}$  mit dem Bruch  $\frac{n}{1} \in \mathbb{Q}$  und fassen  $\mathbb{Z}$  als Teilmenge von  $\mathbb{Q}$  auf. Insbesondere liegen  $0 = \frac{0}{1}$  und  $1 = \frac{1}{1}$  in  $\mathbb{Q}$ .

1.51. SATZ. In  $\mathbb{Q}$  gelten die folgenden Rechenregeln:

(1) Assoziativgesetz für Addition und Multiplikation

(2) neutrale Elemente:  $\frac{p}{q} + 0 = \frac{p}{q}$ ,  $\frac{p}{q} \cdot 1 = \frac{p}{q}$  für alle  $\frac{p}{q} \in \mathbb{Q}$ ;

(3) inverse Elemente:  $\frac{p}{q} + \frac{-p}{q} = 0$  für alle  $\frac{p}{q} \in \mathbb{Q}$ ,  $\frac{p}{q} \cdot \left(\frac{p}{q}\right)^{-1} = 1$  für alle  $\frac{p}{q} \in \mathbb{Q} \setminus \{0\}$ ;

(4) Kommutativgesetz für Addition und Multiplikation;

(5) Distributivgesetz;

(6) Die Relation „ $\leq$ “ ist eine Ordnung;

(7) Aus  $\frac{p}{q} \leq \frac{r}{s}$  folgt  $\frac{p}{q} + \frac{t}{u} \leq \frac{r}{s} + \frac{t}{u}$ ;

(8) Aus  $0 \leq \frac{p}{q}$  und  $0 \leq \frac{r}{s}$  folgt  $0 \leq \frac{p}{q} \cdot \frac{r}{s}$ .

BEWEIS. Diese Aussagen folgen aus den Sätzen 1.41 und 1.48, und aus der Konstruktion von  $\mathbb{Q}$ . Seien etwa  $p, r, t \in \mathbb{Z}$ ,  $q, s, u \in \mathbb{N} \setminus \{0\}$ , dann ergibt sich das Assoziativgesetz für die Addition aus

$$\begin{aligned} \left(\frac{p}{q} + \frac{r}{s}\right) + \frac{t}{u} &= \frac{ps + qr}{qs} + \frac{t}{u} = \frac{(ps + qr) \cdot u + qst}{qsu} = \frac{psu + qru + qst}{qsu} \\ &= \frac{psu + q(ru + st)}{qsu} = \frac{p}{q} + \frac{ru + st}{su} = \frac{p}{q} + \left(\frac{r}{s} + \frac{t}{u}\right). \end{aligned}$$

Betrachten wir das *multiplikative Inverse*  $(\frac{p}{q})^{-1}$  von  $\frac{p}{q} \in \mathbb{Q} \setminus \{0\}$ . Wir unterscheiden zwei Fälle:

Falls  $0 < p$ , gilt  $(\frac{p}{q})^{-1} = \frac{q}{p}$  und  $\frac{p}{q} \cdot (\frac{p}{q})^{-1} = \frac{pq}{qp} = 1$ .

Falls  $p < 0$ , gilt  $(\frac{p}{q})^{-1} = \frac{-q}{-p}$  und  $\frac{p}{q} \cdot (\frac{p}{q})^{-1} = \frac{p(-q)}{q(-p)} = \frac{-pq}{-qp} = 1$ .

Alle anderen Aussagen lassen sich ähnlich beweisen.  $\square$

### 1.4. Etwas Euklidische Geometrie

Der nächste Schritt wäre jetzt die Einführung der reellen Zahlen  $\mathbb{R}$ . In der Schule definiert man reelle Zahlen als Dezimalbrüche. Diese Konstruktion hat einige Probleme, eines davon ist  $0,99\dots = 1$ . Andere Konstruktionen benutzen (Äquivalenzklassen von) Cauchy-Folgen oder Dedekindsche Schnitte, jeweils in  $\mathbb{Q}$ . Die reellen Zahlen haben folgende Eigenschaften.

- (1) Die reellen Zahlen bilden einen *angeordneten Körper*, das heißt, es gelten alle Rechenregeln aus Satz 1.51.
- (2) Die reellen Zahlen sind *archimedisch angeordnet*, das heißt, die natürlichen Zahlen  $\mathbb{N}$  sind in  $\mathbb{R}$  enthalten, und zu jeder reellen Zahl  $r \in \mathbb{R}$  gibt es eine natürliche Zahl  $n \in \mathbb{N}$  mit  $r \leq n$ .
- (3) Die reellen Zahlen sind *vollständig*, das heißt, es ist der größte Körper, für den (1) und (2) gelten. Genauer: wenn es einen anderen Körper  $\mathbb{k}$  gibt, der (1) und (2) erfüllt, dann ist  $\mathbb{k}$  zu einem Teilkörper von  $\mathbb{R}$  isomorph. Noch genauer: es existiert eine eindeutige Abbildung  $f: \mathbb{k} \rightarrow \mathbb{R}$ , so dass für alle  $x, y$  gilt, dass  $f(x + y) = f(x) + f(y)$ ,  $f(xy) = f(x) \cdot f(y)$ ,  $f(1) = 1$  und  $f(x) \leq f(y)$  genau dann, wenn  $x \leq y$ , und diese Abbildung ist injektiv.
- (4) Die rationalen Zahlen  $\mathbb{Q}$  liegen *dicht* in  $\mathbb{R}$ , das heißt, zu  $r, s \in \mathbb{R}$  mit  $r < s$  existiert  $\frac{p}{q} \in \mathbb{Q}$  mit  $r \leq \frac{p}{q} \leq s$ .
- (5) Addition, Subtraktion, Multiplikation und Division sind *stetig*.

Die Eigenschaften (1)–(3) definieren  $\mathbb{R}$  eindeutig (modulo der Probleme, die wir mit der Eindeutigkeit von  $\mathbb{N}$  hatten). Es ist nicht offensichtlich, dass Eigenschaft (3) zu der Definition von Vollständigkeit aus der Analysis äquivalent ist. Aber es ist eine Möglichkeit, Vollständigkeit zu definieren, ohne analytische Begriffe zu verwenden.

In der Schule haben Sie Vektorrechnung möglicherweise wie folgt kennengelernt (meist mit  $n = 2$  oder  $n = 3$ ). Es sei

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_{n \text{ Faktoren}} = \{ x = (x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{R} \},$$

dann definiert man eine Vektoraddition  $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ , eine skalare Multiplikation  $\mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  für  $x = (x_1, \dots, x_n)$ ,  $y = (y_1, \dots, y_n) \in \mathbb{R}^n$  und  $a \in \mathbb{R}$  und

einen Nullvektor  $0$  durch

$$\begin{aligned}x + y &= (x_1 + y_1, \dots, x_n + y_n), \\ax &= (ax_1, \dots, ax_n), \\0 &= (0, \dots, 0).\end{aligned}$$

Übrigens werden wir in dieser Vorlesung nicht zwischen Orts- und Richtungsvektoren unterscheiden — ein  $n$ -Tupel kann immer beides bedeuten.

Um Euklidische Geometrie zu betreiben, definiert man ein Skalarprodukt. Daraus kann man Längen von Vektoren und Winkel zwischen Vektoren ableiten. Für die folgende Definition erinnern wir uns daran, dass die Cosinus-Funktion invertierbar ist als Funktion  $\cos: [0, \pi] \rightarrow [-1, 1]$  mit Umkehrfunktion  $\arccos: [-1, 1] \rightarrow [0, \pi]$ . Hierbei messen wir Winkel grundsätzlich in Bogenmaß. Insbesondere gilt

$$1^\circ = \frac{\pi}{180}.$$

1.52. DEFINITION. Wir definieren das *Standard-Skalarprodukt* auf  $\mathbb{R}^n$  als Abbildung  $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  für Vektoren  $x$  und  $y \in \mathbb{R}^n$  durch

$$(1) \quad \langle x, y \rangle = x_1 y_1 + \dots + x_n y_n.$$

Die *Euklidische Norm*  $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$  auf dem  $\mathbb{R}^n$  ist definiert durch

$$(2) \quad \|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \dots + x_n^2}.$$

Für zwei Vektoren  $x, y \in \mathbb{R}^n \setminus \{0\}$  definieren wir den *Winkel* durch

$$(3) \quad \angle(x, y) = \arccos \frac{\langle x, y \rangle}{\|x\| \|y\|} \in [0, \pi].$$

Wir sammeln einige wichtige Eigenschaften und Rechenregeln.

1.53. BEMERKUNG. Seien  $x, y, z \in \mathbb{R}^n$  sowie  $a, b \in \mathbb{R}$ , dann gilt

$$(1) \quad \langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle;$$

$$(2) \quad \langle x, y \rangle = \langle y, x \rangle;$$

$$(3) \quad \langle x, x \rangle \geq 0 \quad \text{und} \quad \langle x, x \rangle = 0 \iff x = 0.$$

All das rechnet man leicht nach; für (3) nutzen wir aus, dass  $x_1^2, \dots, x_n^2 \geq 0$ . Man sagt, das Skalarprodukt ist *linear* in der ersten Variablen (1), *symmetrisch* (2) und *positiv definit* (3). Aus (1) und (2) folgt, dass das Skalarprodukt auch in der zweiten Variable linear ist, denn

$$(1') \quad \langle x, ay + bz \rangle = \langle ay + bz, x \rangle = a \langle y, x \rangle + b \langle z, x \rangle = a \langle x, y \rangle + b \langle x, z \rangle.$$

Für den folgenden Satz benötigen wir den reellen *Absolutbetrag*  $|\cdot|: \mathbb{R} \rightarrow \mathbb{R}$ , definiert durch

$$|r| = \begin{cases} r & \text{falls } r \geq 0, \text{ und} \\ -r & \text{falls } r < 0. \end{cases}$$

Insbesondere gilt immer  $|r| \geq 0$ , und  $|r| = \sqrt{r^2}$ .

1.54. SATZ (Cauchy-Schwarz-Ungleichung). Für alle Vektoren  $x, y \in \mathbb{R}^n$  gilt

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| .$$

Gleichheit gilt genau dann, wenn Zahlen  $a, b \in \mathbb{R}$  existieren, die nicht beide Null sind, so dass

$$ax + by = 0 .$$

BEWEIS. Wir machen eine Fallunterscheidung.

Fall 1: Es sei  $x = 0$ . Dann gilt  $\|x\| = 0$  und

$$\langle x, y \rangle = 0 = 0 \cdot \|y\| = \|x\| \cdot \|y\| .$$

Also gilt sogar Gleichheit, und mit  $a = 1$  und  $b = 0$  gilt ebenfalls

$$ax + by = 1 \cdot 0 + 0 \cdot y = 0 .$$

Fall 2: Es sei  $x \neq 0$ , dann ist auch  $\|x\|^2 \neq 0$ , und wir berechnen

$$\begin{aligned} 0 \leq \left\| y - \frac{\langle x, y \rangle}{\|x\|^2} x \right\|^2 &= \left\langle y - \frac{\langle x, y \rangle}{\|x\|^2} x, y - \frac{\langle x, y \rangle}{\|x\|^2} x \right\rangle \\ &= \|y\|^2 - 2 \frac{\langle x, y \rangle}{\|x\|^2} \langle x, y \rangle + \frac{\langle x, y \rangle^2}{\|x\|^4} \|x\|^2 = \|y\|^2 - \frac{\langle x, y \rangle^2}{\|x\|^2} . \end{aligned}$$

Da  $\|x\|^2 > 0$ , folgt mit elementaren Umformungen

$$\langle x, y \rangle^2 \leq \|x\|^2 \cdot \|y\|^2 .$$

Wurzelziehen liefert die Behauptung.

Wegen  $x \neq 0$  ist Gleichheit in der Cauchy-Schwarz-Ungleichung äquivalent zu

$$\left\| y - \frac{\langle x, y \rangle}{\|x\|^2} x \right\| = 0 ,$$

aufgrund von Bemerkung 1.53 (3) also auch zu

$$y - \frac{\langle x, y \rangle}{\|x\|^2} x = 0 .$$

Daraus folgt  $ax + by = 0$  mit  $b = \|x\|^2 \neq 0$  und  $a = -\langle x, y \rangle$ .

Umgekehrt sei  $ax + by = 0$ . Wäre  $b = 0$ , so würde aus  $ax = 0$  und  $x \neq 0$  bereits  $a = 0$  folgen, aber  $a$  und  $b$  dürfen nicht beide verschwinden. Also folgt  $b \neq 0$  und

$$y = -\frac{a}{b} x = -\frac{\langle x, \frac{a}{b} x \rangle}{\|x\|^2} x = \frac{\langle x, y \rangle}{\|x\|^2} x ,$$

und es gilt Gleichheit in der Cauchy-Schwarz-Ungleichung.  $\square$

Der Vektor  $y - \frac{\langle x, y \rangle}{\|x\|^2} x$  im obigen Beweis entspricht dem Lot vom Punkt  $y$  auf die Gerade durch  $0$  mit Richtung  $x$ . Insbesondere gilt Gleichheit, wenn der Punkt  $y$  auf dieser Geraden liegt.

1.55. BEMERKUNG. Aus der Cauchy-Schwarz-Ungleichung 1.54 folgt

$$\frac{\langle x, y \rangle}{\|x\| \|y\|} \in [-1, 1] \subset \mathbb{R},$$

also ist der Arcuscosinus in Definition 1.52 (3) erklärt und der Winkel wohldefiniert. Umgekehrt gilt also

$$(1) \quad \langle x, y \rangle = \|x\| \|y\| \cos \angle(x, y).$$

Zur geometrischen Interpretation betrachten wir das Dreieck mit den Endpunkten  $0$ ,  $x$  und  $y$ . Die dritte Seite ist  $x - y$ , und wir erhalten den Cosinussatz der Euklidischen Geometrie:

$$(2) \quad \|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\| \|y\| \cos \angle(x, y).$$

### 1.5. Komplexe Zahlen und die Geometrie der Ebene

In den reellen Zahlen können wir Wurzeln aus positiven Zahlen ziehen, beispielsweise aus 2, was in  $\mathbb{Q}$  nicht möglich ist. Man kann aber keine Wurzeln aus negativen Zahlen ziehen. Diesen Missstand wollen wir jetzt beheben, indem wir die reellen Zahlen zu den komplexen Zahlen erweitern.

Die Idee ist, eine neue Zahl  $i$  einzuführen, deren Quadrat  $-1$  ist. Wir möchten mit Zahlen  $a + bi$  mit  $a, b \in \mathbb{R}$  rechnen, und alle von  $\mathbb{R}$  vertrauten Rechenregeln sollen gelten. Zum Beispiel sollten die folgenden Rechnungen richtig sein:

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

$$\text{und} \quad (a + bi) \cdot (c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

Um das rigoros zu machen, betrachten wir eine komplexe Zahl als Paar aus zwei reellen Zahlen, und definieren Addition und Multiplikation wie oben.

1.56. DEFINITION. Die *komplexen Zahlen* sind definiert als  $\mathbb{C} = \mathbb{R}^2$ , mit

$$(a, b) + (c, d) = (a + c, b + d)$$

$$\text{und} \quad (a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

für alle  $a, b, c, d \in \mathbb{R}$ .

1.57. SATZ. In  $\mathbb{C}$  gelten Assoziativ- und Kommutativgesetz sowohl für die Addition als auch für die Multiplikation. Neutrale Elemente sind  $0_{\mathbb{C}} = (0, 0)$  für die Addition und  $1_{\mathbb{C}} = (1, 0)$  für die Multiplikation. Es gilt das Distributivgesetz. Jedes Element  $(a, b)$  besitzt ein additives Inverses

$$-(a, b) = (-a, -b)$$

und, falls  $(a, b) \neq 0_{\mathbb{C}}$ , ein multiplikatives Inverses

$$(a, b)^{-1} = \left( \frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right).$$

BEWEIS. Alle Behauptungen lassen sich direkt mit den Formeln aus Definition 1.56 nachrechnen. Beispielsweise gilt

$$(a, b) \cdot \left( \frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2} \right) = \left( \frac{a^2 + b^2}{a^2 + b^2}, \frac{-ab + ba}{a^2 + b^2} \right) = (1, 0) = 1_{\mathbb{C}}. \quad \square$$

Wir sehen, dass die Abbildung  $\mathbb{R} \rightarrow \mathbb{C}$  mit  $a \mapsto (a, 0)$  verträglich mit  $+$  und  $\cdot$  ist, und  $0$  und  $1 \in \mathbb{R}$  auf  $0_{\mathbb{C}}$  und  $1_{\mathbb{C}}$  abbildet. Wir dürfen also  $\mathbb{R}$  mit den komplexen Zahlen der Form  $(\cdot, 0)$  identifizieren. Wenn wir außerdem noch  $i = (0, 1)$  definieren, können wir uns überzeugen, dass

$$(a, b) = (a, 0) + b \cdot (0, 1) = a + bi$$

für alle  $a, b \in \mathbb{R}$  gilt. Damit haben wir unsere Idee vom Anfang des Abschnitts verwirklicht. Außerdem dürfen wir jetzt auch  $0$  und  $1$  für  $0_{\mathbb{C}}$  und  $1_{\mathbb{C}}$  schreiben.

1.58. BEMERKUNG. Auf  $\mathbb{C}$  gibt es keine Ordnung „ $\leq$ “, die zu Satz 1.51 (7) und (8) analoge Eigenschaften hat. Denn gäbe es solch eine Ordnung, dann gälte entweder  $0 < x$  oder  $0 > x$  für alle  $x \neq 0$  wegen Totalität, aber wegen (7) gälte  $0 > x$  genau dann, wenn  $-x > 0$ . Also gälte  $x^2 = (-x)^2 > 0$  für alle  $x \neq 0$  wegen (8), aber dann erhielten wir wegen (7) und Transitivität einen Widerspruch:

$$0 = 1^2 + i^2 \geq 1^2 > 0.$$

1.59. DEFINITION. Sei  $z = a + bi \in \mathbb{C}$  mit  $a, b \in \mathbb{R}$ , dann heißt  $a$  der *Realteil*  $\operatorname{Re}(z)$  von  $z$  und  $b$  der *Imaginärteil*  $\operatorname{Im}(z)$  von  $z$ .

Der Imaginärteil ist also immer eine reelle Zahl, und es gilt

$$z = \operatorname{Re}(z) + \operatorname{Im}(z) \cdot i.$$

1.60. DEFINITION. Die Abbildung  $\mathbb{C} \rightarrow \mathbb{C}$  mit  $z \mapsto \bar{z} = \operatorname{Re}(z) - \operatorname{Im}(z) \cdot i$  heißt *komplexe Konjugation*,  $\bar{z}$  heißt das (*komplex*) *Konjugierte* von  $z$ .

1.61. BEMERKUNG. Die komplexe Konjugation ist verträglich mit allen Rechenoperationen, das heißt, es gilt

$$\begin{aligned} \bar{z} + \bar{w} &= \overline{z + w}, & \bar{z} \cdot \bar{w} &= \overline{z \cdot w}, \\ \overline{-z} &= -\bar{z}, & \overline{z^{-1}} &= \bar{z}^{-1}, \\ \overline{0} &= 0, & \overline{1} &= 1, \end{aligned}$$

auch das rechnet man leicht nach.

Es gilt  $\bar{\bar{z}} = z$  für alle  $z$ , also ist die komplexe Konjugation ihre eigene Umkehrabbildung. Für eine komplexe Zahl  $z$  gilt  $z = \bar{z}$  genau dann, wenn  $z \in \mathbb{R} \subset \mathbb{C}$ .

Man kann die komplexen Zahlen dadurch charakterisieren, dass sie die kleinste Erweiterung der reellen Zahlen  $\mathbb{R}$  ist, so dass alle Rechenregeln aus Satz 1.57 gelten und eine Zahl  $i$  mit  $i^2 = -1$  existiert. Aber  $i$  ist dadurch nicht eindeutig bestimmt, denn offensichtlich sind  $i$  und  $\bar{i} = -i$  gleichberechtigt.

Die Zahl  $z = i$  löst die Gleichung  $z^2 + 1 = 0$ . In den Übungen werden Sie sehen, dass man  $z^2 = w$  für alle komplexen Zahlen  $w$  lösen kann. All das sind Spezialfälle des folgenden Satzes.

1.62. SATZ (Fundamentalsatz der Algebra). *Es seien  $n \geq 1$  und  $a_1, \dots, a_n \in \mathbb{C}$ , dann existiert  $z \in \mathbb{C}$ , so dass*

$$z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n = 0 .$$

Mit rein algebraischen Methoden lässt sich dieser Satz nicht beweisen. Das liegt daran, dass die reellen Zahlen, die den komplexen ja zugrundeliegen, mit analytischen Mitteln konstruiert wurden. Einen Beweis für diesen Satz lernen Sie daher erst später, zum Beispiel in einer Vorlesung über Funktionentheorie oder Topologie.

Für  $z = a + bi$  mit  $a, b \in \mathbb{R}$  ist

$$z \cdot \bar{z} = (a + bi)(a - bi) = a^2 - (bi)^2 = a^2 + b^2 \geq 0$$

reell. Das ermöglicht folgende Definition.

1.63. DEFINITION. Wir definieren den *Absolutbetrag* (die *Norm* oder die *Länge*) einer komplexen Zahl  $z \in \mathbb{C}$  als die reelle Zahl

$$|z| = \sqrt{z \cdot \bar{z}} \geq 0 .$$

1.64. BEMERKUNG. Wir sammeln ein paar Eigenschaften des Absolutbetrages.

- (1) Da  $|a + bi|^2 = a^2 + b^2$ , entspricht  $|z| = \|z\|$  der euklidischen Norm auf  $\mathbb{C} = \mathbb{R}^2$  aus Definition 1.52 (1).
- (2) Unsere Konstruktion von  $z^{-1} = \frac{\bar{z}}{|z|^2}$  wird jetzt etwas klarer, denn

$$z \cdot \frac{\bar{z}}{|z|^2} = \frac{|z|^2}{|z|^2} = 1 .$$

- (3) Der Absolutbetrag ist *multiplikativ*, das heißt, für alle  $z$  und  $w$  gilt

$$|zw| = \sqrt{zw \overline{zw}} = \sqrt{(z\bar{z})(w\bar{w})} = \sqrt{z\bar{z}} \cdot \sqrt{w\bar{w}} = |z| |w| .$$

- (4) Der Absolutbetrag ist *subadditiv* wegen (1) und der Cauchy-Schwarz-Ungleichung 1.54, das heißt, für alle  $z, w \in \mathbb{C}$  gilt

$$|z + w| \leq |z| + |w| ,$$

denn

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

- (5) Komplexe Konjugation ist mit dem Absolutbetrag verträglich, denn

$$|\bar{z}| = \sqrt{\bar{z}z} = \sqrt{z\bar{z}} = |z| .$$

Wir wollen uns Addition und Multiplikation in  $\mathbb{C}$  jetzt mit Hilfe der zweidimensionalen Euklidischen Geometrie veranschaulichen. Dazu machen wir einige Anleihen aus der Schulmathematik und identifizieren  $\mathbb{C}$  mit dem Vektorraum  $\mathbb{R}^2$ .

Die Addition in  $\mathbb{C}$  entspricht der Vektoraddition in  $\mathbb{R}^2$ . Die komplexe Konjugation ist eine Spiegelung an der reellen Achse (also an der  $x$ -Achse).

Wir schreiben einen Vektor  $z \in \mathbb{C} \setminus \{0\}$  als

$$z = |z| \cdot \frac{z}{|z|}.$$

Dann misst  $|z| = \|z\|$  die Länge von  $z$ . Multiplikation mit  $|z| \in \mathbb{R} \subset \mathbb{C}$  entspricht offenbar der Streckung im  $\mathbb{R}^2$  mit dem Faktor  $|z|$ , denn

$$|z| \cdot (a + bi) = (|z| + 0i)(a + bi) = |z|a + |z|bi.$$

Der Vektor  $\frac{z}{|z|}$  hat Länge 1, der zugehörige Punkt in  $\mathbb{C} = \mathbb{R}^2$  liegt also auf dem Einheitskreis

$$S^1 = \{ w \in \mathbb{C} \mid |w| = 1 \}.$$

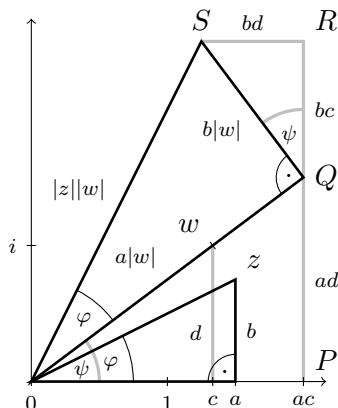
und beschreibt die Richtung von  $z$ . Es sei  $\varphi \in [0, 2\pi)$  der Winkel zwischen der reellen Achse ( $x$ -Achse) und  $\frac{z}{|z|}$  entgegen dem Uhrzeigersinn, dann folgt

$$\frac{z}{|z|} = \cos \varphi + i \sin \varphi \quad \text{und} \quad z = |z| (\cos \varphi + i \sin \varphi).$$

Später lernen Sie, dass  $\cos \varphi + i \sin \varphi = e^{i\varphi}$ . Man nennt  $z = r e^{i\varphi}$  mit  $r = |z|$  auch die *Polardarstellung* der Zahl  $z \neq 0$ . Der Winkel  $\varphi$  heißt auch das *Argument* von  $z$ , geschrieben  $\varphi = \arg(z)$ . Das Argument von 0 ist nicht definiert.

**1.65. LEMMA** (Geometrische Interpretation der komplexen Multiplikation). *Es sei  $z \in \mathbb{C} \setminus \{0\}$ , dann entspricht die Multiplikation mit  $z$  einer Streckung um den Faktor  $|z|$  gefolgt von einer Drehung um den Winkel  $\varphi = \arg(z)$  gegen den Uhrzeigersinn.*

**BEWEIS.** Wir beweisen die Aussage durch Ähnlichkeitsüberlegungen in der Ebene. Dazu sei  $z = a + bi \neq 0$  gegeben und  $w = c + di \neq 0$  ein beliebiger Punkt. Wir betrachten das folgende Bild.



Wir strecken das Dreieck  $\Delta 0cw$  um den Faktor  $a = \operatorname{Re}(z)$  und erhalten das Dreieck  $\Delta 0PQ$ . Anschließend strecken wir  $\Delta 0cw$  um den Faktor  $b = \operatorname{Im}(z)$ , drehen um einen rechten Winkel gegen den Uhrzeigersinn, und verschieben, so dass wir das Dreieck  $\Delta QRS$  erhalten. Dann hat der Punkt  $S$  die Koordinaten  $ac - bd$  und  $ad + bc$ , folglich ist  $S$  der gesuchte Punkt  $zw$ .

Nach Konstruktion liegen die Punkte  $P$ ,  $Q$  und  $R$  auf einer Geraden. Da sich die drei Winkel bei  $Q$  zu  $\pi$  ergänzen, hat das Dreieck  $\Delta 0QS$  bei  $Q$  einen rechten Winkel. Die beiden Katheten haben die Längen  $a|w|$  beziehungsweise  $b|w|$ , folglich ist  $\Delta 0QS$  ähnlich zum Dreieck  $\Delta 0az$  mit Streckfaktor  $|w|$ . Insbesondere hat es bei  $0$  den Winkel  $\varphi = \arg(z)$ .

Wir sehen also, dass der Punkt  $S = zw$  einen um den Faktor  $|z|$  größeren Absolutbetrag (d.h., Abstand zum Nullpunkt) hat als  $w$ , und ein um  $\varphi = \arg(z)$  größeres Argument als  $w$ . Der Punkt  $w = 0$  hingegen bleibt unter Multiplikation mit  $z$  unverändert.  $\square$

Wir können also komplexe Zahlen in Polardarstellung multiplizieren durch

$$r e^{i\varphi} \cdot s e^{i\psi} = rs e^{i(\varphi+\psi)} .$$

Wir können auch Wurzeln ziehen (wobei wir uns auf das Vorzeichen einigen müssen):

$$\sqrt{r e^{i\varphi}} = \pm \sqrt{r} e^{i\frac{\varphi}{2}} .$$

Andererseits lassen sich komplexe Zahlen in Polardarstellung nicht so leicht addieren.

Es fällt auf, dass der oben benutzte Winkelbegriff nicht ganz mit dem aus dem letzten Abschnitt übereinstimmt. Hier betrachten wir Drehungen gegen den Uhrzeigersinn um beliebige Winkel, wobei der Winkel  $\varphi$  und der Winkel  $\varphi + 2\pi n$  für alle  $n \in \mathbb{Z}$  die gleiche Drehung beschreiben. Alle Winkel im Intervall

$$(-\pi, \pi] = \{ x \in \mathbb{R} \mid -\pi < x \leq \pi \}$$

stehen für verschiedene Drehungen, insbesondere entsprechen Winkel  $\varphi \in (-\pi, 0)$  Drehungen im Uhrzeigersinn um  $|\varphi|$ .

In Definition 1.52 (3) hingegen haben wir nur „ungerichtete“ Winkel im Intervall  $[0, \pi]$  betrachtet. Besser ging es nicht, da die Winkel  $\varphi$  und  $-\varphi$  den gleichen Cosinus haben, und der Arcus Cosinus sich nach unserer Definition für Winkel in  $[0, \pi]$  entscheidet.

1.66. BEMERKUNG. Unter einer *Isometrie* verstehen wir eine abstandserhaltende Abbildung  $f$ . Eine Isometrie  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  der Ebene muss für alle  $x, y \in \mathbb{R}^2$  also

$$\|f(x) - f(y)\| = \|x - y\|$$

erfüllen. Aufgrund des Cosinussatzes 1.55 (2) erhält  $f$  auch (unorientierte) Winkel, für alle  $x, y, z \in \mathbb{R}^2$  gilt also

$$\angle f(x)f(y)f(z) = \angle (f(x) - f(y), f(z) - f(y)) = \angle (x - y, z - y) = \angle xyz .$$

Die Isometrien der Ebene werden erzeugt von

- (1) Verschiebungen  $w \mapsto a + w$  mit  $a \in \mathbb{C}$ ,
- (2) Drehungen um den Ursprung,  $w \mapsto zw$ , wobei  $z \in \mathbb{C}$  mit  $|z| = 1$ , und
- (3) der Spiegelung an der  $x$ -Achse,  $w \mapsto \bar{w}$ .

Insgesamt können wir also jede Isometrie  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit Hilfe komplexer Zahlen schreiben als

$$F(w) = a + zw \quad \text{oder} \quad F(w) = a + z\bar{w} ,$$

wobei  $a$  und  $z \in \mathbb{C}$  mit  $|z| = 1$  durch  $F$  eindeutig festgelegt sind.

## 1.6. Geometrie des Raumes und Quaternionen

Wir geben einen kurzen Abriss der Euklidischen Geometrie des Raumes, insbesondere führen wir das Kreuzprodukt ein. In Analogie zu den komplexen Zahlen definieren wir die Quaternionen, bei denen sowohl Kreuz- als auch Skalarprodukt auf dem  $\mathbb{R}^3$  eine wichtige Rolle spielen. Die wichtigsten Eigenschaften der Quaternionen lernen wir später kennen.

1.67. DEFINITION. Das *Kreuzprodukt* (*Vektorprodukt*) auf dem  $\mathbb{R}^3$  ist eine Abbildung  $\times: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$(u_1, u_2, u_3) \times (v_1, v_2, v_3) = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1) .$$

Beachten Sie, dass das Symbol “ $\times$ ” sowohl das kartesische Produkt von Mengen ( $\mathbb{R}^3 \times \mathbb{R}^3$ ) aus Definition 1.10 (5) als auch das Kreuzprodukt von Vektoren bezeichnet. Missverständnisse wird es deswegen voraussichtlich nicht geben.

1.68. BEMERKUNG. Für alle  $u, v, w \in \mathbb{R}^3$  und alle  $a, b \in \mathbb{R}$  gilt

- (1)  $(au + bv) \times w = a(u \times w) + b(v \times w) ,$
- (2)  $u \times v = -v \times u .$

All dies folgt unmittelbar aus Definition 1.67. Man sagt, das Kreuzprodukt ist linear im ersten Argument (1) und *antisymmetrisch* (2).

Wegen (1) und (2) ist das Kreuzprodukt auch im zweiten Argument linear, denn

$$(1') \quad u \times (av + bw) = -(av + bw) \times u \\ = -a(v \times u) - b(w \times u) = a(u \times v) + b(u \times w) .$$

1.69. SATZ. Für alle  $u, v, w, t \in \mathbb{R}^3$  gilt

- (1)  $\langle u \times v, w \rangle = \langle v \times w, u \rangle = \langle w \times u, v \rangle ,$
- (2)  $(u \times v) \times w = \langle u, w \rangle \cdot v - \langle v, w \rangle \cdot u = w \times (v \times u) ,$
- (3)  $0 = (u \times v) \times w + (v \times w) \times u + (w \times u) \times v ,$
- (4)  $\langle u \times v, w \times t \rangle = \langle u, w \rangle \langle v, t \rangle - \langle u, t \rangle \langle v, w \rangle$

Die Gleichung (2) heißt auch *Graßmann-Identität*, und (3) heißt *Jacobi-Identität*. Den Ausdruck  $\langle u \times v, w \rangle$  in (1) nennt man auch das *Spatprodukt* der Vektoren  $u, v, w$ .

BEWEIS. Zu (1) berechnen wir

$$\langle u \times v, w \rangle = u_2 v_3 w_1 - u_3 v_2 w_1 + u_3 v_1 w_2 - u_1 v_3 w_2 + u_1 v_2 w_3 - u_2 v_1 w_3,$$

und dieser Ausdruck ist invariant unter zyklischer Vertauschung von  $u$ ,  $v$  und  $w$ .

Die Graßmann-Identität (2) überprüfen wir nur in der ersten Komponente der ersten Gleichung:

$$\begin{aligned} ((u \times v) \times w)_1 &= (u \times v)_2 \cdot w_3 - (u \times v)_3 \cdot w_2 \\ &= u_3 \cdot v_1 \cdot w_3 - u_1 \cdot v_3 \cdot w_3 - u_1 \cdot v_2 \cdot w_2 + u_2 \cdot v_1 \cdot w_2 \\ &= (u_1 \cdot w_1 + u_2 \cdot w_2 + u_3 \cdot w_3) \cdot v_1 \\ &\quad - (v_1 \cdot w_1 + v_2 \cdot w_2 + v_3 \cdot w_3) \cdot u_1 \\ &= \langle u, w \rangle \cdot v_1 - \langle v, w \rangle \cdot u_1; \end{aligned}$$

die zweite und dritte Komponente ergeben sich, indem man oben die Indizes 1, 2 und 3 zyklisch vertauscht. Die zweite Gleichung folgt aus der ersten mit Antisymmetrie.

Die Jacobi-Identität (3) folgt, indem man  $u$ ,  $v$  und  $w$  in (2) zyklisch permutiert und dann alle drei Gleichungen addiert.

Behauptung (4) folgt aus (1) und (2) durch folgende Rechnung:

$$\begin{aligned} \langle u \times v, w \times t \rangle &= \langle (w \times t) \times u, v \rangle \\ &= \langle \langle w, u \rangle \cdot t - \langle t, u \rangle \cdot w, v \rangle = \langle u, w \rangle \langle v, t \rangle - \langle u, t \rangle \langle v, w \rangle. \quad \square \end{aligned}$$

1.70. BEMERKUNG. Wir geben eine geometrische Interpretation.

(1) Satz 1.69 (4) und Bemerkung 1.55 (1) implizieren, dass

$$\begin{aligned} \|u \times v\| &= \sqrt{\|u\|^2 \|v\|^2 - \langle u, v \rangle^2} \\ &= \sqrt{\|u\|^2 \|v\|^2 (1 - \cos^2 \angle(u, v))} = \|u\| \|v\| \sin \angle(u, v), \end{aligned}$$

da  $\sin^2 + \cos^2 = 1$  und  $\sin \varphi \geq 0$  für alle  $\varphi \in [0, \pi]$ . Also ist  $\|u \times v\|$  gerade der Flächeninhalt des von  $u$  und  $v$  aufgespannten Parallelogramms. Aus Bemerkung 1.68 (2) und Satz 1.69 (1) folgt

$$\langle u \times v, u \rangle = \langle u \times u, v \rangle = 0 \quad \text{und} \quad \langle u \times v, v \rangle = \langle v \times v, u \rangle = 0.$$

Also steht  $u \times v$  senkrecht auf der Fläche dieses Parallelogramms. Damit haben wir eine geometrische Beschreibung des Kreuzproduktes *bis auf das Vorzeichen*. Das Vorzeichen ergibt sich durch die Wahl einer Orientierung, wie wir später in Beispiel 4.28 lernen werden.

(2) Das Spatprodukt können wir nun als Volumen des Parallelotops mit den Kanten  $u$ ,  $v$  und  $w$  interpretieren. Da  $u \times v$  senkrecht auf der Grundfläche steht, wird die Höhe dieses Parallelotops gerade gegeben durch

$$\|w\| |\cos \angle(u \times v, w)| = \|w\| \frac{|\langle u \times v, w \rangle|}{\|u \times v\| \|w\|} = \frac{|\langle u \times v, w \rangle|}{\|u \times v\|}.$$

Als Produkt aus Grundfläche  $\|u \times v\|$  und Höhe erhalten wir das Volumen also als Absolutbetrag  $|\langle u \times v, w \rangle|$  des Spatproduktes. Das Vorzeichen des Spatproduktes ist wiederum eine Frage der Orientierung.

Wir erinnern uns an unsere Definition 1.56 der komplexen Zahlen. Dort wurde eine Multiplikation auf  $\mathbb{R} \times \mathbb{R}$  erklärt durch

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Wir führen jetzt die etwas kompliziertere Quaternionen-Multiplikation ein. Die Quaternionen wurden von Hamilton entdeckt, daher der Buchstabe  $\mathbb{H}$ .

1.71. DEFINITION. Die *Quaternionen* sind definiert als  $\mathbb{H} = \mathbb{R} \times \mathbb{R}^3$ , mit

$$\begin{aligned} (a, u) + (b, v) &= (a + b, u + v), \\ (a, u) \cdot (b, v) &= (a \cdot b - \langle u, v \rangle, a \cdot v + b \cdot u + u \times v) \\ \text{und} \quad \overline{(a, u)} &= (a, -u) \end{aligned}$$

für alle  $a, b \in \mathbb{R}$  und alle  $u, v \in \mathbb{R}^3$ . Wir identifizieren  $a \in \mathbb{R}$  mit  $(a, 0) \in \mathbb{H}$  und  $u \in \mathbb{R}^3$  mit  $(0, u) \in \mathbb{H}$ , und definieren Real- und Imaginärteil von  $(a, u)$  durch

$$\begin{aligned} \operatorname{Re}(a, u) &= \frac{1}{2} ((a, u) + \overline{(a, u)}) = a \in \mathbb{R} \\ \text{und} \quad \operatorname{Im}(a, u) &= \frac{1}{2} ((a, u) - \overline{(a, u)}) = u \in \mathbb{R}^3. \end{aligned}$$

1.72. SATZ. In  $\mathbb{H}$  gelten Assoziativ- und Kommutativgesetz für die Addition. Die Multiplikation ist assoziativ aber nicht kommutativ. Es gilt das Distributivgesetz

$$(1) \quad p \cdot (q + r) = p \cdot q + p \cdot r$$

für alle  $p, q, r \in \mathbb{H}$ . Neutrale Elemente sind  $0_{\mathbb{H}} = (0, 0)$  für die Addition und  $1_{\mathbb{H}} = (1, 0)$  für die Multiplikation. Jedes Element  $(a, u)$  besitzt ein additives Inverses

$$(2) \quad -(a, u) = (-a, -u)$$

und, falls  $(a, u) \neq 0_{\mathbb{H}}$ , ein multiplikatives Inverses

$$(3) \quad (a, u)^{-1} = \left( \frac{a}{a^2 + \|u\|^2}, -\frac{u}{a^2 + \|u\|^2} \right).$$

Für ein Quaternion  $p = (a, u)$  gilt

$$(4) \quad p \cdot q = q \cdot p$$

für alle  $q \in \mathbb{H}$  genau dann, wenn  $p \in \mathbb{R}$ , das heißt, wenn  $u = 0$ .

Für die Quaternionen-Konjugation gilt

$$(5) \quad \overline{p + q} = \bar{p} + \bar{q}, \quad \overline{-p} = -\bar{p},$$

$$(6) \quad \overline{p \cdot q} = \bar{q} \cdot \bar{p}, \quad \overline{p^{-1}} = \bar{p}^{-1}$$

für alle  $p, q \in \mathbb{H}$ , und für  $p = (a, u) \in \mathbb{H}$  gilt

$$(7) \quad \bar{p} \cdot p = a^2 + \|u\|^2 = p \cdot \bar{p}.$$

Die Quaternionen-Konjugation ist ein *Anti-Automorphismus*, das heißt, sie respektiert alle Verknüpfungen bis auf die Multiplikation, bei der sie die Reihenfolge der Faktoren vertauscht. Daher können wir aus (1) und (6) auch Distributivität im ersten Faktor folgern.

Anstelle von  $p/q$  schreiben wir sicherheitshalber  $pq^{-1}$ , was ja nicht das gleiche wie  $q^{-1}p$  sein muss. Wir erlauben Brüche von Quaternionen nur, wenn der Nenner reell ist.

BEWEIS. Die Rechenregeln für die Addition sind leicht zu überprüfen. Das Distributivgesetz (1) folgt aus den Bemerkungen 1.53 (1) und 1.68 (1):

$$\begin{aligned} (a, u) \cdot ((b, v) + (c, w)) &= (a, u) \cdot (b + c, v + w) \\ &= (a(b + c) - \langle u, v + w \rangle, a(v + w) + (b + c)u + u \times (v + w)) \\ &= (ab - \langle u, v \rangle, av + bu + u \times v) + (ac - \langle u, w \rangle, aw + cu + u \times w) \\ &= (a, u) \cdot (b, v) + (a, u) \cdot (c, w). \end{aligned}$$

Das Assoziativgesetz für die Multiplikation folgt aus Satz 1.69 (1) und (2). Außerdem überprüft man leicht, dass

$$(a, u) + (0, 0) = (a, u) = (a, u) \cdot (1, 0) = (1, 0) \cdot (a, u).$$

Auch die Formel (2) für das additive Inverse ist klar.

Es gelte (4) für ein Quaternion  $(a, u)$ . Aus der Symmetrie des Skalarproduktes und der Antisymmetrie des Kreuzproduktes folgt

$$\begin{aligned} 0 &= (a, u) \cdot (b, v) - (b, v) \cdot (a, u) \\ &= (0, u \times v - v \times u) = (0, 2u \times v). \end{aligned}$$

Wir setzen für  $v$  die drei Einheitsvektoren  $e_1, e_2, e_3$  ein und erhalten  $u_1 = u_2 = u_3 = 0$  aus Definition 1.71. Also gilt  $u = 0$ , das heißt  $(a, u) \in \mathbb{R}$ .

Es gilt

$$\begin{aligned} \overline{(a, u)} \cdot (a, u) &= (a, -u) \cdot (a, u) \\ &= (a^2 + \langle u, u \rangle, au - au - u \times u) = a^2 + \|u\|^2 \in \mathbb{R}, \end{aligned}$$

und es folgt die erste Gleichung in (7). Die zweite erhalten wir, indem wir  $u$  durch  $-u$  ersetzen. Aus (7) folgt (3), denn

$$\left( \frac{a}{a^2 + \|u\|^2}, -\frac{u}{a^2 + \|u\|^2} \right) \cdot (a, u) = \frac{1}{\overline{(a, u)} \cdot (a, u)} \overline{(a, u)} \cdot (a, u) = 1.$$

Gleichung (5) ist wiederum klar, und (6) folgt aus der Antisymmetrie des Kreuzproduktes, denn

$$\begin{aligned} \overline{(a, u)} \cdot \overline{(b, v)} &= (ab - \langle u, v \rangle, -av - bu - u \times v) \\ &= (ab - \langle -v, -u \rangle, b(-u) + a(-v) + (-v) \times (-u)) \\ &= \overline{(b, v)} \cdot \overline{(a, u)}. \end{aligned} \quad \square$$

Im Beweis sieht man, dass das Kreuzprodukt wichtig ist für das Assoziativgesetz der Multiplikation.

1.73. DEFINITION. Wir definieren den *Absolutbetrag* eines Quaternions  $q \in \mathbb{H}$  als die reelle Zahl

$$|q| = \sqrt{\bar{q}q}.$$

Wegen Satz 1.72 (7) ist das möglich, und für  $q = (a, u_1, u_2, u_3) \in \mathbb{H}$  gilt

$$|q|^2 = a^2 + u_1^2 + u_2^2 + u_3^2,$$

also stimmt  $|q|$  wiederum mit der Euklidischen Norm  $\|q\|$  auf  $\mathbb{R}^4$  überein.

1.74. BEMERKUNG. So, wie wir den komplexen Zahlen  $(1, 0)$  und  $(0, 1)$  die Namen 1 und  $i$  gegeben haben, wollen wir hier die folgenden Bezeichnungen einführen:

$$1 = (1, 0), \quad i = (0, e_1), \quad j = (0, e_2) \quad \text{und} \quad k = (0, e_3).$$

Wir erhalten die Multiplikationstabelle

$\cdot$	$i$	$j$	$k$
$i$	$-1$	$k$	$-j$
$j$	$-k$	$-1$	$i$
$k$	$j$	$-i$	$-1$

Zusammen mit den Distributivgesetzen und  $1_{\mathbb{H}} = (1, 0)$  können wir jetzt alle Quaternionen miteinander multiplizieren. Wir sehen, dass alle Einträge außerhalb der Diagonalen vom Kreuzprodukt in der Definition der Multiplikation in 1.71 herrühren.

So wie die komplexen Zahlen die Geometrie der Ebene beschreiben, beschreiben die imaginären Quaternionen die Geometrie des dreidimensionalen Raumes. Wir sehen, dass sowohl das Standard-Skalarprodukt als auch das Kreuzprodukt in der Definition auftauchen, und in der Tat erhalten wir diese zurück als

$$\langle u, v \rangle = \operatorname{Re}(\overline{(0, u)} \cdot (0, v)) \quad \text{und} \quad u \times v = \operatorname{Im}(\overline{(0, u)} \cdot (0, v)).$$

Jetzt wollen wir Isometrien des  $\mathbb{R}^3$  mit Hilfe von Quaternionen beschreiben.

1.75. SATZ. *Es sei  $q = (\cos \varphi, v \sin \varphi) \in \mathbb{H}$ , wobei  $v \in \mathbb{R}^3$  mit  $\|v\| = 1$  und  $\varphi \in \mathbb{R}$ . Für ein imaginäres  $w \in \mathbb{R}^3 \subset \mathbb{H}$  ist  $qw\bar{q}$  wieder imaginär. Die Abbildung  $F_q: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit  $w \mapsto qw\bar{q}$  beschreibt eine Drehung um die Achse durch 0 in Richtung  $v$  um den Winkel  $2\varphi$ .*

BEWEIS. Ein Quaternion  $w$  ist imaginär genau dann, wenn  $\bar{w} = -w$  gilt. Wenn  $w$  imaginär ist, ist auch  $qw\bar{q}$  imaginär, denn

$$\overline{qw\bar{q}} = \bar{\bar{q}}\bar{w}\bar{q} = -qw\bar{q} .$$

Die Abbildung  $F_q$  ist  $\mathbb{R}$ -linear wegen Satz 1.72 (1) und (4), das heißt, sie bildet Summen auf Summen ab und ist mit Streckungen verträglich. Das gleiche gilt für die Drehung  $R_{v,2\varphi}$  um die Achse durch 0 in Richtung  $v$  um den Winkel  $2\varphi$ . Wir zerlegen  $w \in \mathbb{R}^3$  wie im Beweis der Cauchy-Schwarz-Ungleichung 1.54 als

$$w = \langle v, w \rangle v + (w - \langle v, w \rangle v) ,$$

so dass der zweite Vektor wegen  $\|v\| = 1$  senkrecht auf  $v$  steht. Wegen Linearität reicht es,  $F_q v = R_{v,2\varphi} v$  und  $F_q w = R_{v,2\varphi} w$  für alle Vektoren  $w$  mit  $|w| = 1$  und  $\langle v, w \rangle = 0$  zu zeigen.

Betrachte zunächst  $v$ . Wegen  $\langle v, v \rangle = 1$  und  $v \times v = 0$  gilt in diesem Fall

$$\begin{aligned} qw\bar{q} &= (\cos \varphi, v \sin \varphi) \cdot (0, v) \cdot (\cos \varphi, -v \sin \varphi) \\ &= (-\sin \varphi, v \cos \varphi) \cdot (\cos \varphi, -v \sin \varphi) \\ &= (-\cos \varphi \sin \varphi + \cos \varphi \sin \varphi, v \sin^2 \varphi + v \cos^2 \varphi) = (0, v) , \end{aligned}$$

da  $\cos^2 \varphi + \sin^2 \varphi = 1$ . Auch die Drehung  $R_{v,2\varphi}$  hält  $v$  fest, es gilt also  $F_q v = v = R_{v,2\varphi} v$ .

Es gelte jetzt  $\langle v, w \rangle = 0$  und  $\|w\| = 1$ . Wegen  $\langle v \times w, v \rangle = 0$  und der Graßmann-Identität gilt in diesem Fall

$$\begin{aligned} qw\bar{q} &= (\cos \varphi, v \sin \varphi) \cdot (0, w) \cdot (\cos \varphi, -v \sin \varphi) \\ &= (0, w \cos \varphi + v \times w \sin \varphi) \cdot (\cos \varphi, -v \sin \varphi) \\ &= (0, w \cos^2 \varphi + v \times w \cos \varphi \sin \varphi - w \times v \cos \varphi \sin \varphi - (v \times w) \times v \sin^2 \varphi) \\ &= (0, w(\cos^2 \varphi - \sin^2 \varphi) + v \times w \cdot 2 \cos \varphi \sin \varphi) . \end{aligned}$$

Cosinus und Sinus des doppelten Winkels berechnen sich als

$$\cos(2\varphi) = \cos^2 \varphi - \sin^2 \varphi \quad \text{und} \quad \sin(2\varphi) = 2 \cos \varphi \sin \varphi .$$

Wenn wir  $\|w\| = 1$  annehmen, dann folgt aus Bemerkung 1.70, dass die Vektoren  $v$ ,  $w$  und  $v \times w$  aufeinander senkrecht stehen, und dass auch

$$\|v \times w\| = \|v\| \cdot \|w\| \cdot \sin \angle(v, w) = 1 .$$

Insbesondere bilden  $w$  und  $v \times w$  eine Orthonormalbasis der zu  $v$  senkrechten Ebene. Die Drehung  $R_{v,2\varphi}$  bildet den Vektor  $w$  also ab auf

$$R_{v,2\varphi} w = \cos(2\varphi) w + \sin(2\varphi) v \times w = F_q w .$$

Wenn wir in der obigen Rechnung  $w$  durch  $v \times w$  ersetzen, wird  $v \times w$  wegen der Graßmann-Identität zu  $v \times (v \times w) = -w$ . Wir sehen jetzt, dass auch  $v \times w$  in der zu  $v$  senkrechten Ebene um den Winkel  $2\varphi$  gedreht wird. Wegen Linearität gilt das also für den gesamten  $\mathbb{R}^3$ .  $\square$

Man beachte, dass  $\varphi$  und  $\varphi + \pi$  die gleiche Drehung beschreiben, da  $2\pi$  ja einer vollen Umdrehung entspricht. Zu einer Drehung gehören also genau zwei Quaternionen  $q$  und  $-q$ ; dieses Phänomen nennt man „Spin“. Es hat sowohl in der Mathematik als auch in der Physik eine Bedeutung.

Die Drehrichtung ergibt sich aus einer „Rechte-Faust-Regel“. Sei  $0 < \varphi < \pi$ , so dass wir um  $2\varphi \in (0, 2\pi)$  drehen. Zeigt der Daumen der rechten Hand in die Richtung von  $\operatorname{Im} q = v \sin \varphi$ , dann erfolgt die Drehung in Richtung der gekrümmten Finger. Ist  $q$  rein imaginär, also beispielsweise  $\varphi = \frac{\pi}{2}$ , dann wird um  $\pi = 180^\circ$  gedreht, so dass es auf die Drehrichtung nicht mehr ankommt.

Wir haben gesehen, dass Quaternionenmultiplikation nicht kommutativ ist. Im Allgemeinen erschwert das den Umgang mit  $\mathbb{H}$ . Aber Satz 1.75 funktioniert gerade, weil  $\mathbb{H}$  nicht kommutativ ist. Wäre  $\mathbb{H}$  kommutativ, dann wäre auch  $qw\bar{q} = q\bar{q}w = |q|^2 w = w$  wegen  $|q| = 1$ , und  $F_q$  wäre einfach die Identität.

1.76. BEMERKUNG. Die Isometrien des Raumes werden erzeugt von

- (1) Verschiebungen  $w \mapsto u + w$  mit  $u \in \mathbb{R}^3$ ,
- (2) Drehungen um die Achse durch den Ursprung in Richtung  $v$  mit Winkel  $\varphi$ , also  $w \mapsto F_q w$ , wobei jetzt

$$q = \cos \frac{\varphi}{2} + v \sin \frac{\varphi}{2},$$

- (3) Die Punktspiegelung  $w \mapsto -w$ .

In Analogie zu Bemerkung 1.66 können wir also jede Isometrie schreiben als

$$F(w) = u + qw\bar{q} \quad \text{oder} \quad F(w) = u + q\bar{w}\bar{q}.$$

Dabei sind  $u \in \operatorname{Im} \mathbb{H}$  und  $q \in \mathbb{H}$  mit  $|q| = 1$  durch  $F$  fast eindeutig festgelegt — man kann nach wie vor  $q$  durch  $-q$  ersetzen.

Die obige Darstellung hat zwei interessante Eigenschaften.

- Sei  $G(w) = v + rw\bar{r}$  eine weitere Isometrie, dann hat auch die Verkettung  $F \circ G$  die gleiche Form:

$$(F \circ G)(w) = u + q(v + rw\bar{r})\bar{q} = (u + qv\bar{q}) + qr w \bar{q}\bar{r}.$$

- Anhand der obigen Formel kann man  $u$  und  $q$  leicht bestimmen, wenn man Drehachse und -winkel kennt. Umgekehrt kann man Drehachse und -winkel ablesen, wenn  $u$  und  $q$  bekannt sind.

Aufgrunddessen lassen sich Quaternionen in der Praxis einsetzen, zum Beispiel in der Robotersteuerung und in der dreidimensionalen Bildverarbeitung.

1.77. BEMERKUNG. Analog zu den Bemerkungen 1.66 und 1.76 können wir auch alle Isometrien des  $\mathbb{R}^4$  beschreiben durch

$$F(w) = v + pw\bar{q} \quad \text{oder} \quad F(w) = v + p\bar{w}\bar{q}.$$

Hierbei ist  $w \in \mathbb{R}^4 = \mathbb{H}$ , und die Quaternionen  $v, p, q \in \mathbb{H}$  mit  $|p| = |q| = 1$  sind durch  $F$  fast eindeutig festgelegt — man darf nur das Tripel  $(v, p, q)$

durch das Tripel  $(v, -p, -q)$  ersetzen. Es gibt also auch hier einen „Spin“. Der Zusammenhang zwischen dem Paar  $(p, q)$  und der Gestalt der Isometrie ist nicht so einfach zu erklären wie in Satz 1.75 und Bemerkung 1.76.

Für  $\mathbb{R}^n$  mit  $n \geq 5$  gibt es leider keine so schönen Beschreibungen der Isometrien mehr. Wir werden später sehen, wie man Isometrien generell durch Matrizen darstellen kann.

### 1.7. Zusammenfassung

In diesem Kapitel haben wir noch einmal den Aufbau des Zahlensystems Revue passieren lassen — von der Mengenlehre als Grundlage der natürlichen Zahlen bis hin zu komplexen Zahlen und Quaternionen. Gleichzeitig haben wir gesehen, dass Euklidische Geometrie — in ihrer analytischen Ausprägung als Vektorgeometrie — zumindest in kleineren Dimensionen — eng mit Erweiterungen der reellen Zahlen verknüpft ist.

In der folgenden Tabelle gehen wir die Zahlbereiche noch einmal durch. Ganz links steht jeweils der Zahlbereich. Direkt daneben geben wir ein „Modell“ an, das heißt, eine Konstruktion des jeweiligen Zahlbereichs aus „bekanntem“ Objekten. Es folgen in Spalte 3 die entscheidenden Neuerungen, und in der letzten Spalte Stichworte zu wichtigen Konzepten.

$\mathbb{N}$	$\underline{\mathbb{N}}$		Rekursive Definitionen, vollständige Induktion
$\mathbb{Z}$	$\mathbb{N} \times \mathbb{N} / \sim$	–	„Differenzrechnung“
$\mathbb{Q}$	$\mathbb{Z} \times (\mathbb{N} \setminus \{0\}) / \approx$	/	Bruchrechnung
$\mathbb{R}$	???		Vollständigkeit
$\mathbb{C}$	$\mathbb{R}^2$	$i$	Fundamentalsatz der Algebra, ebene Geometrie
$\mathbb{H}$	$\mathbb{R} \times \mathbb{R}^3$	$j, k$	Nichtkommutativ, Geometrie des Raumes

## KAPITEL 2

# Vektorräume und Moduln

In diesem Kapitel lernen wir mit Vektoren zu rechnen, indem wir Koordinaten angeben und lineare Abbildungen als Matrizen schreiben. Einem Vektor in Koordinaten entspricht ein Element in einem freien Modul, und einer Matrix entspricht eine lineare Abbildung zwischen freien Moduln. Anschließend überlegen wir uns, warum und wie Matrixrechnung funktioniert.

Für das Rechnen mit Matrizen reicht uns zunächst einmal ein Ring, obwohl wir später meistens einen Körper, zum Beispiel  $\mathbb{R}$ , zugrunde legen werden. Die etwas größere Allgemeinheit verursacht keinen zusätzlichen Aufwand; außerdem müssen wir später gelegentlich mit Matrizen über Ringen arbeiten. Die zahlreichen Vorteile, die die Arbeit über Körpern (auch Schiefkörpern) mit sich bringt, lernen wir dann im nächsten Kapitel kennen.

Als erstes führen wir ein paar algebraische Grundbegriffe ein: Vektoren sind Elemente von Vektorräumen über Körpern oder Schiefkörpern. Etwas allgemeiner ist der Begriff eines Moduls über einem Ring. Und sowohl Ringen als auch Moduln liegen abelsche Gruppen zugrunde, mit denen wir daher beginnen werden. Nachdem wir Moduln eingeführt haben, betrachten wir spezielle „strukturerhaltende“ Abbildungen.

### 2.1. Gruppen, Ringe, Körper

Wir definieren eine Reihe wichtiger algebraischer Strukturen. Unser Hauptziel sind Körper. Aber auch Gruppen und Ringe werden uns noch häufiger begegnen.

2.1. DEFINITION. Eine *Gruppe*  $(G, *)$  ist eine Menge  $G$  mit einer Verknüpfung  $*$ :  $G \times G \rightarrow G$ , für die ein neutrales Element  $e \in G$  und für alle  $g \in G$  ein inverses Element  $g^{-1} \in G$  existiert, so dass für alle  $g, h$  und  $k$  die folgenden Gruppenaxiome gelten:

- (G1)  $g * (h * k) = (g * h) * k$  (*Assoziativgesetz*),
- (G2)  $e * g = g$  (*linksneutrales Element*),
- (G3)  $g^{-1} * g = e$  (*linksinverse Elemente*).

Eine Gruppe heißt *kommutativ* oder *abelsch*, wenn außerdem für alle  $g, h \in G$  gilt

- (G4)  $g * h = h * g$  (*Kommutativgesetz*).

2.2. BEISPIEL. Wir kennen schon Beispiele von abelschen Gruppen. Dazu ersetzen wir „ $*$ “ durch eine bekannte Operation, hier „ $+$ “.

- (1) Die ganzen Zahlen  $\mathbb{Z}$  bilden eine abelsche Gruppe  $(\mathbb{Z}, +)$ , genannt die *unendliche zyklische Gruppe*, siehe auch Satz 1.48.
- (2) Sei  $\mathbb{k} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ . Dann ist  $(\mathbb{k}, +)$  eine abelsche Gruppe, die sogenannte *additive Gruppe* von  $\mathbb{k}$ , siehe dazu die Sätze 1.51, 1.57 und 1.72, sowie Punkt (1) am Anfang von Abschnitt 1.4.
- (3) Die natürlichen Zahlen  $\mathbb{N}$  bilden keine Gruppe, denn es fehlen die inversen Elemente.

Die Gruppenaxiome sind bewusst sparsam formuliert. Dadurch hat man relativ wenig zu tun, um nachzuweisen, dass eine bestimmte Verknüpfung auf einer Menge eine Gruppe definiert. Beim Rechnen in Gruppen hilft die folgende Proposition.

2.3. PROPOSITION. *Sei  $(G, *)$  eine Gruppe, dann sind das neutrale Element  $e$  und das Inverse  $g^{-1}$  zu jedem  $g \in G$  eindeutig bestimmt. Außerdem gilt für alle  $g \in G$ , dass*

$$(G2') \quad g * e = g ,$$

$$(G3') \quad g * g^{-1} = e .$$

Insbesondere muss man das neutrale Element und die Abbildung, die einem Gruppenelement sein Inverses zuordnet, in der Notation „ $(G, *)$ “ nicht mit angeben, da beide eindeutig festgelegt sind. Das spart etwas Schreibarbeit. Und wir dürfen tatsächlich von neutralen und inversen Elementen reden, nicht von linksneutralen und linksinversen Elementen.

BEWEIS. Wir leiten aus den Gruppenaxiomen der Reihe nach einige interessante Rechenregeln ab. Für alle  $g, h, k \in G$  gilt

- (1) Linkskürzungsregel: aus  $g * h = g * k$  folgt  $h = k$ , denn

$$\begin{aligned} h &= e * h = (g^{-1} * g) * h = g^{-1} * (g * h) \\ &= g^{-1} * (g * k) = (g^{-1} * g) * k = e * k = k . \end{aligned}$$

- (2) Die Aussage (G2') folgt aus der Linkskürzungsregel (1) und

$$g^{-1} * (g * e) = (g^{-1} * g) * e = e * e = e = g^{-1} * g .$$

- (3) Eindeutigkeit des neutralen Elements: Es gelte  $f * g = g$  für alle  $g \in G$ , dann folgt aus (G2') insbesondere

$$f = f * e = e .$$

Umgekehrt gelte  $g * f = g$  für alle  $g \in G$ , dann folgt aus (G2) ebenfalls

$$f = e * f = e .$$

- (4) Aussage (G3') folgt aus der Linkskürzungsregel (1) und

$$g^{-1} * (g * g^{-1}) = (g^{-1} * g) * g^{-1} = e * g^{-1} = g^{-1} = g^{-1} * e .$$

- (5) Rechtskürzungsregel: aus
- $h * g = k * g$
- folgt
- $h = k$
- , denn

$$\begin{aligned} h &= h * e = h * (g * g^{-1}) = (h * g) * g^{-1} \\ &= (k * g) * g^{-1} = k * (g * g^{-1}) = k * e = k . \end{aligned}$$

- (6) Eindeutigkeit des Inversen: aus
- $g * h = e$
- folgt
- $h = g^{-1}$
- wegen der Linkskürzungsregel (1) und

$$g * h = e = g * g^{-1} ,$$

umgekehrt folgt  $k = g^{-1}$  aus  $k * g = e$  wegen der Rechtskürzungsregel (2) und

$$k * g = e = g^{-1} * g . \quad \square$$

2.4. BEMERKUNG. Wir erinnern uns an die Verkettung „ $\circ$ “ von Abbildungen aus Definition 1.20, an die Identität  $\text{id}_M$  aus Beispiel 1.19 (1) und an die Umkehrabbildungen aus Satz 1.24.

- (1) Es seien
- $K, L, M, N$
- Mengen und
- $F: M \rightarrow N, G: L \rightarrow M$
- und
- $H: K \rightarrow L$
- Abbildungen,

$$K \xrightarrow{H} L \xrightarrow{G} M \xrightarrow{F} N .$$

Dann gilt  $F \circ (G \circ H) = (F \circ G) \circ H$ , denn für alle  $k \in K$  ist

$$\begin{aligned} (F \circ (G \circ H))(k) &= F((G \circ H)(k)) = F(G(H(k))) \\ &= (F \circ G)(H(k)) = ((F \circ G) \circ H)(k) . \end{aligned}$$

- (2) Für
- $F: M \rightarrow N$
- gilt
- $\text{id}_N \circ F = F = F \circ \text{id}_M$
- , denn für alle
- $m \in M$
- gilt
- $(\text{id}_N \circ F)(m) = \text{id}_N(F(m)) = F(m) = F(\text{id}_M(m)) = (F \circ \text{id}_M)(m)$
- .

- (3) Es sei
- $F$
- bijektiv. Dann existiert eine Umkehrabbildung
- $S$
- nach Satz 1.24, und es gilt

$$S \circ F = \text{id}_M \quad \text{und} \quad F \circ S = \text{id}_N .$$

Diese Beziehungen sehen fast so aus wie die Gruppenaxiome (G1)–(G3). Man sollte aber beachten, dass die Abbildungen  $F, G, H, \text{id}_M, \text{id}_N$  und  $S$  im Allgemeinen von verschiedenen Typen sind. Das heißt, wenn die Mengen  $K, L, M, N$  paarweise verschieden sind, gehören keine zwei dieser Abbildungen zur gleichen Grundmenge, etwa  $F \in \text{Abb}(M, N), \text{id}_M \in \text{Abb}(M, M)$ , und so weiter.

2.5. BEISPIEL. Es sei  $M$  eine Menge. Wir definieren die Menge der *Automorphismen* von  $M$  als

$$\text{Aut}(M) = \{ F: M \rightarrow M \mid F \text{ ist bijektiv} \} .$$

Dann bildet  $(\text{Aut}(M), \circ)$  eine Gruppe. Dazu überlegen wir uns

- (1) Seien  $F$  und  $G$  bijektiv, dann ist  $F \circ G$  bijektiv nach Satz 1.23 (3). Also ist die Verknüpfung „ $\circ$ “ auf  $\text{Aut}(M)$  wohldefiniert.
- (2) Es gilt das Assoziativgesetz (G1) nach Bemerkung 2.4 (1).
- (3) Die Identität  $\text{id}_M$  aus Beispiel 1.19 (1) ist bijektiv. Nach Bemerkung 2.4 (2) ist  $\text{id}_M$  das neutrale Element in  $(\text{Aut}(M), \circ)$ .

- (4) Das Inverse zu  $F \in \text{Aut}(M)$  ist die Umkehrabbildung  $G$  aus Satz 1.24. Aus Satz 1.23 (4) und (5) folgt, dass  $G$  wieder bijektiv ist, und das Axiom (G3) folgt aus Bemerkung 2.4 (3).

Später werden uns häufiger Gruppen begegnen, die aus speziellen bijektiven Abbildungen  $F$  einer Menge  $M$  in sich bestehen.

2.6. DEFINITION. Ein *Ring*  $(R, +, \cdot)$  besteht aus einer Menge  $R$  mit einer *Addition*  $+: R \times R \rightarrow R$  und einer *Multiplikation*  $\cdot: R \times R \rightarrow R$ , so dass  $(R, +)$  eine abelsche Gruppe bildet, und so dass für alle  $r, s, t \in R$  die folgenden Ringaxiome gelten:

$$\begin{aligned} \text{(R1)} \quad & (r \cdot s) \cdot t = r \cdot (s \cdot t) && \text{(Assoziativgesetz),} \\ \text{(R2)} \quad & \begin{cases} r \cdot (s + t) = r \cdot s + r \cdot t \\ (r + s) \cdot t = r \cdot t + s \cdot t \end{cases} && \text{(Distributivgesetze).} \end{aligned}$$

Ein Ring heißt *unitär* oder *Ring mit Eins*, wenn es ein neutrales Element oder *Einelement*  $1_R$  gibt, so dass für alle  $r \in R$  gilt:

$$\text{(R3)} \quad 1_R \cdot r = r \cdot 1_R = r \quad \text{(Einsselement).}$$

Ein Ring heißt *kommutativ*, wenn für alle  $r, s \in R$  gilt:

$$\text{(R4)} \quad r \cdot s = s \cdot r \quad \text{(Kommutativgesetz).}$$

Man beachte, dass die Axiome (R3) und (R4) unabhängig voneinander erfüllt sein können. Wir werden in dieser Vorlesung fast nur Ringe mit Eins betrachten.

In allgemeinen Ringen haben wir kein Kommutativgesetz, daher brauchen wir beide Gleichungen in (R2) und (R3). Wir haben auch keine Links- oder Rechtskürzungsregeln für die Multiplikation, da uns die multiplikativen Inversen fehlen.

Die Gruppe  $(R, +)$  heißt die additive Gruppe des Rings  $(R, +, \cdot)$ . Ihr neutrales Element heißt *Nullelement* (*Null*) und wird mit  $0$  oder  $0_R$  bezeichnet, und das additive Inverse von  $r \in R$  wird  $-r$  geschrieben. Die Bezeichnung  $r^{-1}$  ist für multiplikative Inverse reserviert (wenn sie existieren). Das Symbol für die Multiplikation wird häufig weggelassen, somit steht  $rs$  kurz für  $r \cdot s$ .

2.7. BEISPIEL. Wir kennen bereits einige Ringe.

- (1) Die ganzen Zahlen  $(\mathbb{Z}, +, \cdot)$  bilden einen kommutativen Ring mit Eins, siehe Satz 1.48.
- (2) Sei  $\mathbb{k} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ . Dann ist  $(\mathbb{k}, +, \cdot)$  ein Ring mit Eins; siehe dazu die Sätze 1.51, 1.57 und 1.72, sowie Punkt (1) am Anfang von Abschnitt 1.4. Bis auf  $\mathbb{H}$  sind diese Ringe auch kommutativ.
- (3) Auf den natürlichen Zahlen  $\mathbb{N}$  sind zwar Addition und Multiplikation erklärt, und (R1)–(R4) gelten. Aber da  $(\mathbb{N}, +)$  keine Gruppe ist, ist  $(\mathbb{N}, +, \cdot)$  kein Ring, siehe Beispiel 2.2 (3).

Auch aus den Ringaxiomen lassen sich Folgerungen ziehen.

2.8. PROPOSITION. *Es sei  $(R, +, \cdot)$  ein Ring. Dann gilt für alle  $r, s \in R$ , dass*

$$(1) \quad 0_R \cdot r = r \cdot 0_R = 0_R,$$

$$(2) \quad r \cdot (-s) = (-r) \cdot s = -r \cdot s.$$

*In einem Ring mit Eins ist die Eins eindeutig, und es gilt entweder  $0_R \neq 1_R$ , oder aber  $R = \{0_R\}$ .*

Aufgrund der letzten Aussage wird bei einem Ring mit Eins manchmal zusätzlich  $0_R \neq 1_R$  gefordert.

BEWEIS. Aus dem Distributivgesetz (R2) folgt

$$0_R \cdot r = (0_R + 0_R) \cdot r = 0_R \cdot r + 0_R \cdot r,$$

also  $0_R = 0_R \cdot r$  nach Kürzungsregel für die Addition. Genauso folgt  $r \cdot 0_R = 0_R$ .

Aussage (2) folgt aus

$$0_R = r \cdot 0_R = r \cdot (s + (-s)) = r \cdot s + r \cdot (-s),$$

genauso erhält man die zweite Gleichung.

Die Eindeutigkeit der Eins folgt wie in Proposition 2.3.

Wenn in einem Ring mit Eins  $0_R = 1_R$  gilt, folgt aus (R3) und (1) für alle  $r \in R$ , dass

$$r = 1_R \cdot r = 0_R \cdot r = 0_R. \quad \square$$

Der Ring  $R = \{0\}$  heißt auch *Nullring* oder „trivialer Ring“.

2.9. BEISPIEL. Sei  $n \in \mathbb{N}$ ,  $n \geq 1$ . Wir definieren eine Relation „ $\equiv \text{ mod } n$ “ auf  $\mathbb{Z}$  durch

$$a \equiv b \pmod{n} \iff \text{es gibt } k \in \mathbb{Z} \text{ mit } a - b = kn,$$

lies: „ $a$  ist kongruent zu  $b$  modulo  $n$ “.

Wir wollen zeigen, dass es sich um eine Äquivalenzrelation handelt. Die Relation ist reflexiv (Ä1), denn  $a - a = 0 \cdot n$  für alle  $a \in \mathbb{Z}$ . Für  $a, b \in \mathbb{Z}$  gelte  $a - b = kn$  mit  $k \in \mathbb{Z}$ , dann folgt  $b - a = (-k) \cdot n$ , also ist die Relation symmetrisch (Ä2). Schließlich ist sie auch transitiv (Ä3), denn gelte  $a - b = kn$  und  $b - c = \ell n$  für  $a, b, c, k, \ell \in \mathbb{Z}$ , dann folgt  $a - c = (\ell + k) \cdot n$ .

Die Äquivalenzklasse von  $a \in \mathbb{Z}$  heißt *Restklasse von  $a$*  und hat die Form

$$[a] = \{a + k \cdot n \mid k \in \mathbb{Z}\} = \{\dots, a - n, a, a + n, \dots\}.$$

Der Quotient heißt *Menge der Restklassen modulo  $n$*  und wird mit  $\mathbb{Z}/n$  oder  $\mathbb{Z}/n\mathbb{Z}$  bezeichnet. Indem wir  $a \in \mathbb{Z}$  mit Rest durch  $n$  dividieren, erhalten wir  $b, k \in \mathbb{Z}$  mit  $0 \leq b < n$ , so dass  $a = kn + b$ . Es folgt

$$\mathbb{Z}/n = \{[0], \dots, [n-1]\},$$

insbesondere hat  $\mathbb{Z}/n$  die Mächtigkeit  $n$ .

Analog zu Abschnitt 1.3 wollen wir zeigen, dass Addition und Multiplikation in  $\mathbb{Z}$  auf dem Quotienten  $\mathbb{Z}/n\mathbb{Z}$  wohldefinierte Rechenoperationen definieren. Es sei etwa  $a - b = kn$  und  $c - d = \ell n$ , dann folgt

$$\begin{aligned}(a + c) - (b + d) &= (k + \ell) \cdot n, \\ (a \cdot c) - (b \cdot d) &= (a - b) \cdot c + b \cdot (c - d) = (kc + b\ell) \cdot n \\ \text{und} \quad (-a) - (-b) &= (-k) \cdot n.\end{aligned}$$

Somit erhalten wir Verknüpfungen  $+$ ,  $\cdot$ :  $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) \rightarrow \mathbb{Z}/n\mathbb{Z}$  sowie  $- \cdot$ :  $\mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$  mit

$$[a] + [c] = [a + c], \quad [a] \cdot [c] = [a \cdot c] \quad \text{und} \quad -[a] = [-a].$$

Schließlich wollen wir die Axiome (G1)–(G4) und (R1)–(R4) überprüfen, um zu zeigen, dass  $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$  ein kommutativer Ring mit Eins ist. Dazu setzen wir  $0_{\mathbb{Z}/n\mathbb{Z}} = [0]$  und  $1_{\mathbb{Z}/n\mathbb{Z}} = [1]$ . Jetzt folgt jedes einzelne der obigen Axiome aus der entsprechenden Rechenregel für  $(\mathbb{Z}, +, \cdot)$ , zum Beispiel

$$\begin{aligned}([a] + [b]) + [c] &= [a + b] + [c] = [(a + b) + c] \\ &= [a + (b + c)] = [a] + [b + c] = [a] + ([b] + [c]), \\ [a] \cdot ([b] + [c]) &= [a] \cdot [b + c] = [a \cdot (b + c)] \\ &= [ab + ac] = [ab] + [ac] = [a] \cdot [b] + [a] \cdot [c] \\ \text{und} \quad [1] \cdot [a] &= [1 \cdot a] = [a] = [a \cdot 1] = [a] \cdot [1].\end{aligned}$$

Somit ist  $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$  ein kommutativer Ring mit Eins. Seine additive Gruppe  $(\mathbb{Z}/n\mathbb{Z}, +)$  heißt auch die *zyklische Gruppe der Ordnung  $n$* .

2.10. DEFINITION. Ein *Schiefkörper*  $(K, +, \cdot)$  ist ein Ring mit Eins  $1_K$  und Null  $0_K$ , in dem für alle  $k \in K \setminus \{0_K\}$  ein  $k^{-1} \in K$  existiert, so dass für alle  $k \in K \setminus \{0_K\}$  die folgenden Körperaxiome gelten:

$$\begin{aligned}(\text{K1}) \quad k^{-1} \cdot k &= 1_K && (\text{multiplikatives linksinverses Element}), \\ (\text{K2}) \quad 1_K &\neq 0_K && (\text{Nichttrivialität}).\end{aligned}$$

Ein Schiefkörper heißt *Körper*, wenn die Multiplikation kommutativ ist.

2.11. BEISPIEL. Wir kennen bereits einige Körper und Schiefkörper.

- (1) Es sei  $\mathbb{k} = \mathbb{Q}, \mathbb{R}$  oder  $\mathbb{C}$ , dann ist  $(\mathbb{k}, +, \cdot)$  ein Körper, siehe dazu die Sätze 1.51, 1.57 sowie Punkt (1) am Anfang von Abschnitt 1.4. Insbesondere sind  $\mathbb{Q}, \mathbb{R}$  und  $\mathbb{C}$  auch Schiefkörper.
- (2) Die Quaternionen bilden einen “echten”, also nichtkommutativen Schiefkörper, siehe Satz 1.72.
- (3) Die natürlichen Zahlen  $(\mathbb{N}, +, \cdot)$  sind kein (Schief-) Körper, da sie noch nicht einmal einen Ring bilden. Die ganzen Zahlen  $(\mathbb{Z}, +, \cdot)$  sind zwar ein kommutativer Ring mit Eins, aber kein (Schief-) Körper, da multiplikative Inverse fehlen.

Die Körperaxiome werden in der Literatur oft unterschiedlich formuliert. Manchmal fasst man (G1)–(G4), (R1)–(R4), (K1) und (K2) (oder kleine Variationen davon) zu Axiomen (K1)–(K10) zusammen. Es folgt eine weitere Möglichkeit.

2.12. PROPOSITION. *Eine Menge  $K$  mit Verknüpfungen  $+, \cdot : K \times K \rightarrow K$  und Elementen  $0_K, 1_K \in K$  bildet genau dann einen Schiefkörper  $(K, +, \cdot)$  mit Nullelement  $0_K$  und Einselement  $1_K$ , wenn*

- (1)  $(K, +)$  eine Gruppe mit neutralem Element  $0_K$  bildet,
- (2)  $(K \setminus \{0_K\}, \cdot)$  eine Gruppe mit neutralem Element  $1_K$  bildet, und
- (3) die Distributivgesetze (R2) gelten.

Falls die Gruppe  $(K \setminus \{0_K\}, \cdot)$  abelsch ist, ist  $(K, +, \cdot)$  ein Körper.

BEWEIS.  $\implies$ : Sei  $(K, +, \cdot)$  ein Schiefkörper, dann ist  $(K, +)$  nach den Definitionen 2.6 und 2.10 eine abelsche Gruppe. Auch die Distributivgesetze (R2) haben wir vorausgesetzt, somit gelten (1) und (3).

Zu (2) betrachte  $a, b \neq 0_K$ . Es gilt  $a^{-1} \neq 0$ , denn ansonsten wäre

$$1_K = a^{-1} \cdot a = 0_K,$$

im Widerspruch zu (K2). Es gilt auch  $a \cdot b \neq 0_K$ , denn sonst wäre

$$b = (a^{-1} \cdot a) \cdot b = a^{-1} \cdot (a \cdot b) = 0_K.$$

Somit definiert die Multiplikation eine Verknüpfung auf der Menge  $K \setminus \{0_K\}$ , und auch  $1_K$  und die Inversen  $a^{-1}$  liegen in  $K \setminus \{0_K\}$ . Die Gruppenaxiome für  $(K \setminus \{0_K\}, \cdot)$  folgen jetzt aus (R1), (R3) und (K1).

$\impliedby$ : Wenn (1)–(3) erfüllt sind, gelten zunächst einmal (G1)–(G3) und (R2) wegen (1) und (3). Außerdem folgt  $0_K \cdot k = 0_K = k \cdot 0_K$  für alle  $k \in K$  mit dem gleichen Beweis wie für Proposition 2.8 (1). Es gilt  $1_K \cdot a = a = a \cdot 1_K$  für alle  $a \in K$  nach der obigen Überlegung, falls  $a = 0$ , und nach (G2) und (G2') aus Proposition 2.3 für  $(K \setminus \{0_K\}, \cdot)$ , falls  $a \neq 0_K$ . Also gilt (R3).

Aus (R3) und dem Axiom (G2) folgt für  $(K \setminus \{0_K\}, \cdot)$  mit den Distributivgesetzen, dass

$$\begin{aligned} a + a + b + b &= (1_K + 1_K) \cdot a + (1_K + 1_K) \cdot b = (1_K + 1_K) \cdot (a + b) \\ &= 1_K \cdot (a + b) + 1_K \cdot (a + b) = a + b + a + b. \end{aligned}$$

Die Kürzungsregeln in  $(K, +)$  aus dem Beweis von Proposition 2.3 liefern (G4).

Das Assoziativgesetz (R1) folgt aus (G1) für die Gruppe  $(K \setminus \{0_K\}, \cdot)$ , falls  $r, s, t \in K \setminus \{0_K\}$ . Falls mindestens eines der drei Elemente  $0_K$  ist, sind rechte und linke Seite von (R1) auch  $0_K$ , siehe oben.

Das Axiom (K1) ist gerade (G3) für  $(K \setminus \{0_K\}, \cdot)$ , und (K2) folgt, da  $1_K \in K \setminus \{0_K\}$ . Also ist  $(K, +, \cdot)$  ein Schiefkörper.  $\square$

Wir schreiben  $K^\times = K \setminus \{0_K\}$  und nennen  $(K^\times, \cdot)$  die *multiplikative Gruppe* von  $K$ . Manche Autoren schreiben auch  $K^*$ ; wir wollen uns das Sternchen aber für andere Zwecke aufsparen. Aus (G3') folgt für  $k \in K^\times$  auch, dass  $k \cdot k^{-1} = 1_K$ .

2.13. BEMERKUNG. In jedem Körper oder Schiefkörper  $(K, +, \cdot)$  gilt Proposition 2.3 für die additive Gruppe  $(K, +)$  sowie für die multiplikative Gruppe  $(K^\times, \cdot)$ . Im Fall  $(K^\times, \cdot)$  gelten manche der Aussagen in Proposition 2.3 und ihrem Beweis immer noch, wenn einzelne Elemente  $0_K$  sind. Zur Begründung benutzen wir wieder Proposition 2.8 (1).

- (1) *Kürzungsregeln*: Aus  $a \cdot b = a \cdot c$  oder  $b \cdot a = c \cdot a$  folgt  $b = c$  oder  $a = 0_K$ , genau wie in Satz 1.41 (5).
- (2) *Nullteilerfreiheit*: Aus  $a \cdot b = 0_K$  folgt  $a = 0_K$  oder  $b = 0_K$ . Das ist äquivalent zu (1).
- (3) *neutrales Element*: Es gilt  $1_K \cdot a = a \cdot 1_K = a$  für alle  $a \in K$ ;
- (4) *Eindeutigkeit der Eins*: aus  $a \cdot b = a$  oder  $b \cdot a = a$  für ein  $a \in K^\times$  und ein  $b \in K$  folgt  $b = 1_K$ ;
- (5) *Eindeutigkeit des Inversen*: aus  $a \cdot b = 1_K$  oder  $b \cdot a = 1_K$  für  $a, b \in K$  folgen  $a, b \in K^\times$  und  $b = a^{-1}$ .

Unter *Nullteilern* in einem Ring  $(R, +, \cdot)$  versteht man Elemente  $r, s \in R \setminus \{0\}$  mit  $r \cdot s = 0$ . Körper sind also *nullteilerfrei* nach (2). In Ringen kann es Nullteiler geben, zum Beispiel gilt

$$[2] \cdot [3] = [6] = [0] \quad \in \mathbb{Z}/6\mathbb{Z}.$$

2.14. DEFINITION. Sei  $R$  ein Ring mit Eins. Falls es eine Zahl  $n \in \mathbb{N} \setminus \{0\}$  gibt mit

$$(*) \quad \underbrace{1_R + \cdots + 1_R}_{n \text{ Summanden}} = 0_R,$$

dann heißt die kleinste solche Zahl die *Charakteristik*  $\chi(R)$  von  $R$ . Andernfalls ist  $\chi(R) = 0$ .

Man beachte, dass aus  $\chi(R) = n$  bereits für alle  $r \in R$  folgt:

$$\underbrace{r + \cdots + r}_{n \text{ Summanden}} = \underbrace{(1_R + \cdots + 1_R)}_{n \text{ Summanden}} \cdot r = 0.$$

2.15. BEISPIEL. Für einige Ringe kennen wir die Charakteristik.

- (1) Aus dem ersten Peano-Axiom 1.29 (P1) folgt für alle  $n \in \mathbb{N} \setminus \{0\}$ , dass

$$\underbrace{1 + \cdots + 1}_{n \text{ Summanden}} = n \neq 0.$$

Da  $\mathbb{N}$  eine Teilmenge von  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ ,  $\mathbb{C}$  und  $\mathbb{H}$  ist, folgt

$$\chi(\mathbb{Z}) = \chi(\mathbb{Q}) = \chi(\mathbb{R}) = \chi(\mathbb{C}) = \chi(\mathbb{H}) = 0.$$

- (2) Der Ring  $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$  aus Beispiel 2.9 hat Charakteristik  $n$ .

Aus der Schule kenne wir den Begriff der *Primzahl*. Es sei  $1 \leq n \in \mathbb{N}$ . Wir nennen  $a \in \mathbb{N}$  einen *Teiler* von  $n$ , kurz  $a \mid n$ , wenn es  $b \in \mathbb{N}$  mit  $ab = n$  gibt. Wir nennen eine Zahl  $p \in \mathbb{Z}$  mit  $p > 1$  eine Primzahl, wenn für alle  $a, b \in \mathbb{N}$  aus  $p \mid ab$  folgt, dass  $p \mid a$  oder  $p \mid b$ . Hieraus folgt, dass  $p$  keine Teiler außer 1 und sich selbst hat. Die Zahl 1 selbst ist keine Primzahl.

2.16. PROPOSITION. *Die Charakteristik eines Körpers ist entweder 0 oder eine Primzahl.*

BEWEIS. Wir wollen annehmen, dass  $\chi(K) \neq 0$ . Aus (K2) folgt  $1_K \neq 0_K$ , also ist  $\chi(K) \neq 1$ . Falls jetzt  $\chi(K) = a \cdot b$  mit  $a, b > 1$  gilt, betrachte die Gleichung

$$0_K = \underbrace{1_K + \cdots + 1_K}_{a \cdot b \text{ Summanden}} = \underbrace{(1_K + \cdots + 1_K)}_a \text{ Summanden} \cdot \underbrace{(1_K + \cdots + 1_K)}_b \text{ Summanden}.$$

Da  $K$  als Körper nullteilerfrei ist, muss bereits einer der beiden Faktoren oben  $0_K$  sein. Ohne Einschränkung dürfen wir annehmen, dass es sich um den ersten handelt (ansonsten vertausche  $a$  und  $b$ ). Nun ist aber  $a < a \cdot b$  da  $1 < b$ , und gleichzeitig ist  $a \cdot b$  nach Definition 2.14 die kleinste Zahl mit der Eigenschaft (\*). Aufgrund dieses Widerspruchs kann  $\chi(K)$  kein echtes Produkt sein.  $\square$

2.17. BEISPIEL. Der Ring  $\mathbb{Z}/n\mathbb{Z}$  aus Beispiel 2.9 kann also nur ein Körper sein, wenn  $n$  eine Primzahl ist.

Sei also  $p$  eine Primzahl und  $K = \mathbb{Z}/p\mathbb{Z}$ . Wir wissen schon, dass  $\mathbb{Z}/p\mathbb{Z}$  ein kommutativer Ring mit Eins  $[1] \neq [0]$  ist. Wir wollen noch die Existenz multiplikativer Inverser beweisen (K1). Jedes Element  $[a] \in \mathbb{Z}/p\mathbb{Z} \setminus \{[0]\}$  hat genau  $p$  verschiedene Vielfache in  $\mathbb{Z}/p\mathbb{Z}$ , denn sonst gäbe es  $[b], [c] \in \mathbb{Z}/p\mathbb{Z}$  mit  $[b] \neq [c]$  aber  $[a] \cdot [b] = [a] \cdot [c]$ , also  $a \cdot (b - c) = k \cdot p$  für ein  $k \in \mathbb{Z}$ , aber weder  $a$  noch  $b - c$  enthalten den Primteiler  $p$ , Widerspruch. Also ist die Abbildung  $F: \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}$  mit  $F([b]) = [a][b]$  injektiv, und daher auch surjektiv (Übung), somit existiert  $[b] \in \mathbb{Z}/p\mathbb{Z}$  mit  $[a][b] = [1]$ , das heißt,  $[a]$  hat ein multiplikatives Inverses. Es gibt also *endliche Körper*, das heißt, Körper mit endlich vielen Elementen.

Man kann (K1) auch expliziter beweisen, indem man ein Inverses angibt. Sei dazu  $1 \leq a < p$ , dann gibt es keine Zahl  $c > 1$ , die  $a$  und  $p$  teilt. Nach Satz 2.18 (2) unten für  $a_0 = p > a_1 = a$  existieren Zahlen  $d_0$  und  $d_1 \in \mathbb{Z}$  mit

$$1 = d_1 a_0 + d_0 a_1 = d_1 p + d_0 a.$$

Dann ist  $d_0 a \equiv 1$  modulo  $p$ , also ist  $[d_0] = [a]^{-1} \in \mathbb{Z}/p\mathbb{Z}$  das multiplikative Inverse von  $[a]$ .

Für den folgenden Satz brauchen wir *Division mit Rest*: Zu je zwei Zahlen  $m, n \in \mathbb{N}$  mit  $n \neq 0$  gibt es eindeutige Zahlen  $q, r \in \mathbb{N}$  mit  $0 \leq r < n$ , so dass

$$m = qn + r.$$

2.18. SATZ (Erweiterter Euklidischer Algorithmus). *Es seien  $a_0, a_1 \in \mathbb{N} \setminus \{0\}$  mit  $a_1 \leq a_0$ , Dann existieren eindeutige Zahlen  $i_0 \in \mathbb{N}$ ,  $a_1 > a_2 > \dots > a_{i_0} > a_{i_0+1} = 0$  und  $b_2, \dots, b_{i_0+1} \in \mathbb{N}$ , so dass*

$$(1) \quad a_{i-1} = b_{i+1}a_i + a_{i+1} \quad \text{für alle } 1 \leq i \leq i_0 .$$

Die Zahl  $a_{i_0}$  ist die größte Zahl in  $\mathbb{N}$ , die  $a_0$  und  $a_1$  teilt.

Setze  $d_{i_0} = 0$ ,  $d_{i_0-1} = 1$  und bestimme  $d_{i_0-2}, \dots, d_1, d_0 \in \mathbb{Z}$  so, dass

$$(2) \quad d_{i-1} = d_{i+1} - d_i b_{i+1} \quad \text{für } i_0 > i \geq 1 .$$

Dann gilt eine sogenannte Bézout-Identität  $a_{i_0} = d_1 a_0 + d_0 a_1$ .

Die Zahl  $a_{i_0}$  heißt der *größte gemeinsame Teiler* von  $a_0$  und  $a_1$ , kurz  $a_{i_0} = \text{ggT}(a_0, a_1)$ . Beachte, dass es auch andere  $d_0, d_1 \in \mathbb{Z}$  geben kann, die (2) erfüllen.

BEWEIS. Nach Definition der Division mit Rest existieren die Zahlen  $a_i$  und  $b_i$ , sind eindeutig bestimmt durch (1) und werden immer kleiner. Also erreichen wir  $a_{i_0+1} = 0$  nach  $i_0 \leq a_1$  vielen Schritten.

Es sei  $0 < c \in \mathbb{N}$  eine Zahl, die  $a_0$  und  $a_1$  teilt, dann teilt  $c$  auch alle Zahlen  $a_2, \dots, a_{i_0}$  wegen (1). Also kann es keine Zahl größer als  $a_{i_0}$  geben, die  $a_0$  und  $a_1$  teilt. Aus (1) für  $i_0$  folgt, dass  $a_{i_0}$  auch  $a_{i_0-1}$  teilt. Indem wir (1) für immer kleinere  $i$  benutzen, folgt, dass  $a_{i_0}$  auch  $a_{i_0-2}, \dots, a_1$  und  $a_0$  teilt. Also ist  $a_{i_0} = \text{ggT}(a_0, a_1)$ .

Seien jetzt  $d_i$  wie in (2) gegeben. Betrachte die Gleichung

$$(3) \quad a_{i_0} = d_{i+1}a_i + d_i a_{i+1} .$$

Wegen  $a_{i_0+1} = 0$  und  $d_{i_0+1} = 1$  gilt (3) für  $i = i_0$ . Aus den Gleichungen (1)–(3) für  $i$  erhalten wir

$$\begin{aligned} a_{i_0} &= d_{i+1}a_i + d_i(a_{i-1} - b_{i+1}a_i) \\ &= d_i a_{i-1} + (d_{i+1} - d_i b_{i+1})a_i = d_i a_{i-1} + d_{i-1} a_i . \end{aligned}$$

Also gilt (3) auch für  $i - 1$ . Für  $i = 0$  erhalten wir die Behauptung.  $\square$

2.19. BEMERKUNG. Es gibt einen Körper mit  $n$  Elementen genau dann, wenn sich  $n = p^a$  schreiben lässt, wobei  $p$  eine Primzahl ist und  $a \geq 1$ . Dieser Körper wird  $F_{p^a}$  genannt und hat die Charakteristik  $p$ . Sie lernen ihn in der Algebra-Vorlesung kennen. Es gibt auch Körper der Charakteristik  $p$  mit unendlich vielen Elementen.

Wir sollten in der linearen Algebra immer vor Augen haben, dass es diese endlichen Körper gibt; insbesondere Körper der Charakteristik 2 erfordern ein wenig zusätzliche Aufmerksamkeit.

## 2.2. Moduln, Vektorräume und lineare Abbildungen

Gruppen, Ringe und Körper begegnen uns oft dadurch, dass sie auf anderen Strukturen “wirken”. Uns interessiert hier zunächst der Fall von Ring- und Körperwirkungen; Gruppenwirkungen lernen wir später auch noch kennen.

2.20. DEFINITION. Sei  $(R, +, \cdot)$  ein Ring. Ein (*Rechts-*) *R-Modul*  $(M, +, \cdot)$  besteht aus einer abelschen Gruppe  $(M, +)$  und einer *skalaren Multiplikation*  $\cdot : M \times R \rightarrow M$ , so dass für alle  $m, n \in M$  und alle  $r, s \in R$  die folgenden Modulaxiome gelten

- (M1)  $m \cdot (r \cdot s) = (m \cdot r) \cdot s$  (*Verträglichkeit der Multiplikation*),  
 (M2)  $m \cdot (r + s) = m \cdot r + m \cdot s$  (*Erstes Distributivgesetz*),  
 (M3)  $(m + n) \cdot r = m \cdot r + n \cdot r$  (*Zweites Distributivgesetz*).

Sei  $(R, +, \cdot)$  ein Ring mit Eins 1. Ein *unitärer (Rechts-) R-Modul*  $(M, +, \cdot)$  ist ein Rechtsmodul  $(M, +, \cdot)$ , so dass zusätzlich gilt:

- (M4)  $m \cdot 1 = m$  (*Wirkung der Eins*).

Ist der Ring  $R = K$  ein Schiefkörper oder Körper, so heißen unitäre Rechts-*K-Moduln* auch (*Rechts-*) *K-Vektorräume* oder (*Rechts-*) *Vektorräume über K*.

Man beachte, dass das Symbol „+“ in (M2) zwei verschiedene Bedeutungen hat. Die Punkte für die Multiplikation kann man oft weglassen. Wir sprechen von Rechts-*R-Moduln*, weil *R* durch skalare Multiplikation „von rechts“ auf *M* wirkt. Analog definiert man Links-*R-Moduln* mit einer skalaren Multiplikation  $\cdot : R \times M \rightarrow M$ . In diesem Fall dreht sich in (M1)–(M4) jeweils die Reihenfolge der Faktoren um, beispielsweise würde (M1) zu

$$(r \cdot s) \cdot m = r \cdot (s \cdot m).$$

Auf der anderen Seite folgt Kommutativität der Addition bei einem unitären *R-Modul* aus den restlichen Axiomen wie in Proposition 2.12.

2.21. BEISPIEL. Wir können einige Moduln und Vektorräume angeben.

- (1)  $(M, +, \cdot) = (R, +, \cdot)$  ist ein Rechts-*R-Modul*, wobei “+” und “.” die gleichen Verknüpfungen sind wie in *R*, aufgefasst als  $+: M \times M \rightarrow M$  und  $\cdot : M \times R \rightarrow M$ . Nach Definition 2.6 ist nämlich  $(R, +)$  eine abelsche Gruppe, (R1) liefert (M1), und die Distributivgesetze (R2) liefern (M2) und (M3). Falls *R* eine Eins 1 besitzt, ist *M* auch unitär, denn (M4) folgt dann aus (R3). Völlig analog kann man *R* zu einem Linksmodul machen.
- (2) Der „kleinste“ Rechts-*R-Modul* ist  $(\{0\}, +, \cdot)$  mit  $0 \cdot r = 0$  für alle  $r \in R$ . Er heißt der *Nullmodul*.
- (3) Jede abelsche Gruppe *A* wird zu einem Rechts *R-Modul* mit  $a \cdot r = 0_A$  für alle  $a \in A$  und alle  $r \in R$ . Damit reduzieren sich (M1)–(M3) zur trivialen Aussage  $0_A = 0_A$ . Dieser Modul ist allerdings nicht unitär, es sei denn, er wäre bereits der Nullmodul aus (2).

- (4) Die Vektorräume  $\mathbb{R}^n$ , speziell  $\mathbb{R}^2$  und  $\mathbb{R}^3$  aus den Abschnitten 1.4–1.6 sind Vektorräume über  $\mathbb{R}$ .
- (5) In der Analysis lernen Sie viele  $\mathbb{R}$ -Vektorräume kennen. So sind die Räume der Folgen und der Nullfolgen mit Werten in  $\mathbb{R}$  Vektorräume über  $\mathbb{R}$ . Auch die Räume der stetigen oder der differenzierbaren Funktionen auf einem Intervall  $I \subset \mathbb{R}$  sind  $\mathbb{R}$ -Vektorräume.

2.22. PROPOSITION. *Es sei  $(M, +, \cdot)$  ein  $(R, +, \cdot)$ -Rechtsmodul. Dann gilt für alle  $m \in M$  und  $r \in R$ , dass*

- (1)  $0_M \cdot r = m \cdot 0_R = 0_M$ ,
- (2)  $m \cdot (-s) = (-m) \cdot s = -m \cdot s$ .

Analoge Aussagen gelten für Linksmoduln.

BEWEIS. All das folgt aus den Distributivgesetzen (M2), (M3) wie im Beweis von Proposition 2.8.  $\square$

Wir nennen  $0_M$  das *Nullelement* oder die *Null*, bei Vektorräumen auch den *Nullvektor* von  $M$ .

2.23. BEMERKUNG. Sei  $(R, +, \cdot)$  ein kommutativer Ring, zum Beispiel ein Körper. Dann kann man aus jedem Rechts- $R$ -Modul  $(M, +, \cdot)$  einen Links- $R$ -Modul  $(M, +, \cdot)$  machen und umgekehrt, indem man  $r \cdot m = m \cdot r$  für alle  $r \in R$  und  $m \in M$  setzt. Das einzige fragliche Axiom ist (M1), und wir rechnen nach, dass

$$s \cdot (r \cdot m) = (m \cdot r) \cdot s = m \cdot (r \cdot s) = (r \cdot s) \cdot m = (s \cdot r) \cdot m$$

für alle  $r, s \in R$  und  $m \in M$ . Wir dürfen in diesem Fall also einfach von *Moduln* reden.

Da wir im letzten Schritt das Kommutativgesetz (R4) benutzt haben, zeigt diese Rechnung aber auch, dass wir bei einem nicht kommutativen Ring genau zwischen Links- und Rechtsmoduln unterscheiden müssen.

Abbildung 1 gibt einen Überblick über die bis jetzt definierten algebraischen Strukturen. Der Übersicht halber haben wir nicht-unitäre Moduln von (Schief-) Körpern und Ringen mit Eins weggelassen.

In den Abschnitten 1.4–1.6 haben wir uns viel mit linearen Abbildungen eines Vektorraums in sich beschäftigt. Dieser Begriff ist bereits sinnvoll für Abbildungen zwischen Moduln.

2.24. DEFINITION. Sei  $(R, +, \cdot)$  ein Ring und seien  $(M, +, \cdot)$  und  $(N, +, \cdot)$  Rechts- $R$ -Moduln, dann heißt eine Abbildung  $F: M \rightarrow N$  ein (*Rechts- $R$ -*) *Modulhomomorphismus* oder (*rechts-*)  *$R$ -linear* (kurz: *linear*), falls für alle  $\ell, m \in M$  und alle  $r \in R$  gilt

- (L1)  $F(\ell + m) = F(\ell) + F(m)$  (*Additivität*),
- (L2)  $F(m \cdot r) = F(m) \cdot r$  (*Homogenität*).

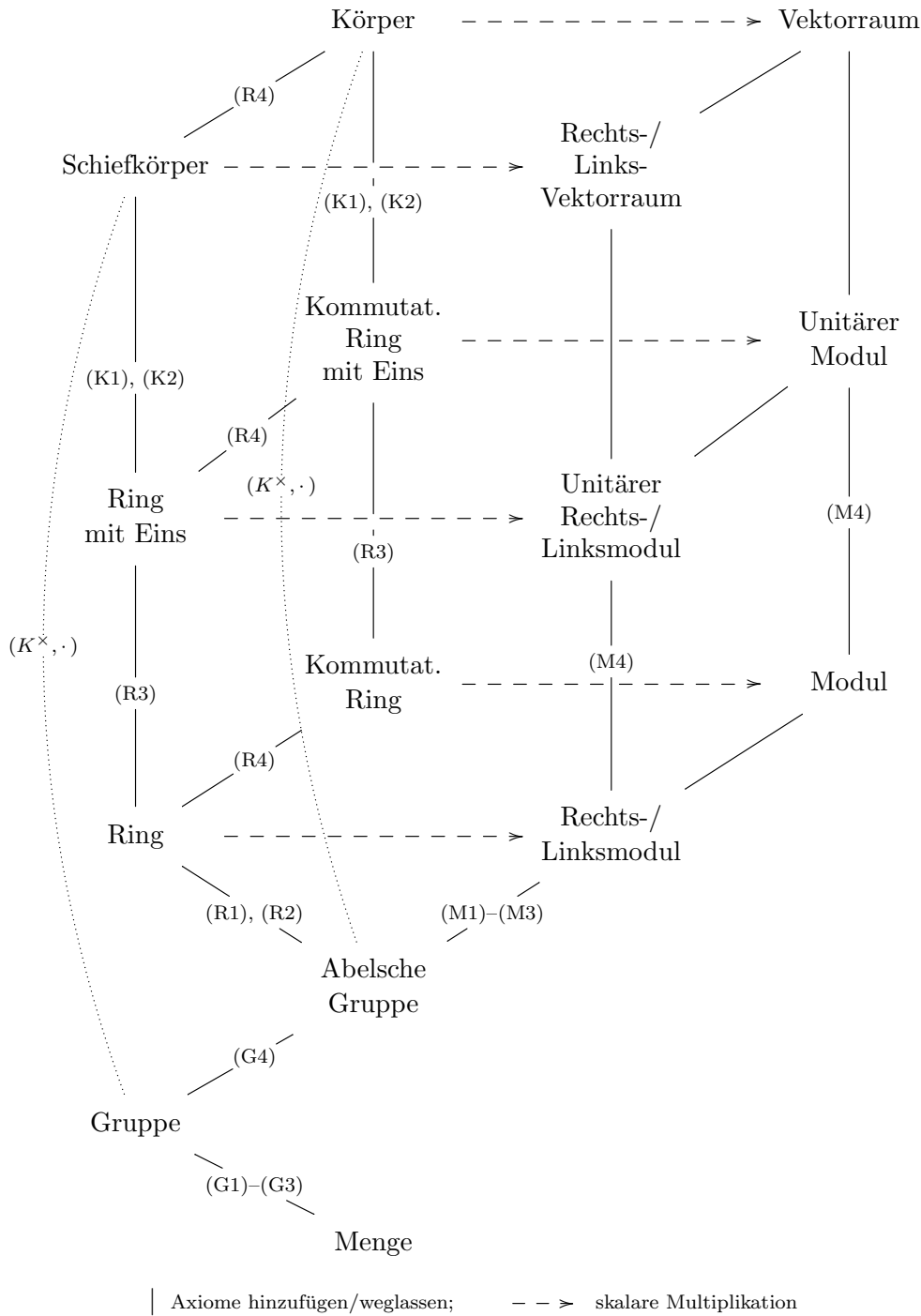


ABBILDUNG 1. Strukturen aus den Abschnitten 2.1 und 2.2

Falls  $R$  ein (Schief-) Körper ist, nennt man lineare Abbildungen zwischen (Rechts- $R$ -) Vektorräumen auch *Vektorraumhomomorphismen*. Die Menge aller (rechts-)  $R$ -linearer Abbildungen von  $M$  nach  $N$  wird mit  $\text{Hom}_R(M, N)$  bezeichnet. Analog definieren wir *Links- $R$ -Modulhomomorphismen*. Die Menge aller Links- $R$ -Modulhomomorphismen von  $A$  nach  $B$  wird mit  ${}_R\text{Hom}(M, N)$  bezeichnet.

Wir bemerken, dass die Addition in (L1) einmal in  $M$  und einmal in  $N$  stattfindet. Genauso wird in (L2) einmal in  $M$  und einmal in  $N$  skalar multipliziert. Aus diesem Grund ist es wichtig, dass beide Moduln über demselben Ring  $R$  definiert sind. Wenn  $R$  kommutativ ist, gibt es nach Bemerkung 2.23 keinen Unterschied zwischen Links- und Rechts- $R$ -Moduln. Wir sprechen dann nur noch von Modulhomomorphismen, und schreiben  $\text{Hom}(M, N)$  oder  $\text{Hom}_R(M, N)$  für die Menge aller linearer Abbildungen.

2.25. BEMERKUNG. Für lineare Abbildungen gilt wegen Proposition 2.22 (1) insbesondere immer

$$(1) \quad F(0_M) = F(0_M \cdot 0_R) = F(0_M) \cdot 0_R = 0_N .$$

Außerdem sind lineare Abbildungen verträglich mit Linearkombinationen: Seien  $m, n \in R$  und  $r, s \in R$ , dann gilt

$$(2) \quad F(m \cdot r + n \cdot s) = F(m \cdot r) + F(n \cdot s) = F(m) \cdot r + F(n) \cdot s .$$

2.26. BEISPIEL. Wir kennen bereits Beispiele linearer Abbildungen.

- (1) Wir haben bereits in den Abschnitten 1.5 und 1.6 benutzt (aber noch nicht bewiesen), dass Isometrien des  $\mathbb{R}^2$  und des  $\mathbb{R}^3$ , die den Nullpunkt festhalten,  $\mathbb{R}$ -linear sind. Dazu gehören Drehungen um den Nullpunkt und Spiegelungen an Achsen durch den Nullpunkt im  $\mathbb{R}^2$ , siehe Bemerkung 1.66, sowie Drehungen um Achsen durch den Nullpunkt, die Punktspiegelung am Ursprung, sowie Spiegelungen an Ebenen durch den Nullpunkte im  $\mathbb{R}^3$ , siehe Bemerkung 1.76.
- (2) Wir betrachten  $M = N = \mathbb{C}$  zunächst als Modul über  $\mathbb{C}$ . Die komplexe Konjugation entspricht der Spiegelung an der reellen Achse. Wir überprüfen die Axiome (L1), (L2). Nach Bemerkung 1.61 gilt

$$\overline{z + w} = \bar{z} + \bar{w} \quad \text{und} \quad \overline{z \cdot w} = \bar{z} \cdot \bar{w} .$$

Also ist komplexe Konjugation additiv, aber nicht homogen, da im allgemeinen  $w \neq \bar{w}$ . Wenn wir aber  $\mathbb{C}$  als  $\mathbb{R}$ -Modul auffassen, dann gilt auch (L2), da  $w = \bar{w}$  genau dann, wenn  $w \in \mathbb{R}$ . Also kommt es bei Linearität auf den zugrundeliegenden Ring oder (Schief-) Körper an.

- (3) Sei  $M = R$  ein unitärer Rechts- $R$ -Modul wie in Beispiel 2.21 (1), so dass  $m \cdot r = mr$  für  $m, r \in R$ . Es sei  $f: M \rightarrow M$  rechts  $R$ -linear und  $p = f(1)$ . dann folgt

$$f(m) = f(1 \cdot m) = f(1) \cdot m = pm ,$$

also wird  $f$  durch Linksmultiplikation mit  $p = f(1)$  gegeben. Umgekehrt ist Linksmultiplikation mit einem beliebigen  $r \in R$  eine rechts- $R$ -lineare Abbildung, denn für alle  $m, n, s \in R$  gilt

$$r \cdot (m + n) = r \cdot m + r \cdot n \quad \text{und} \quad r \cdot (m \cdot s) = (r \cdot m) \cdot s .$$

2.27. BEMERKUNG. Auch in der Analysis spielen lineare Abbildungen eine wichtige Rolle. Beispielsweise dient die Ableitung einer Funktion  $f: I \rightarrow \mathbb{R}$  auf einem offenen Intervall  $I \subset \mathbb{R}$  dazu, die Funktion an einer Stelle  $x_0 \in I$  zu beschreiben als

$$(1) \quad f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + o(x - x_0) ,$$

dabei ist der zweite Term linear in  $x - x_0$ , und der Rest  $o(x - x_0)$  geht für  $x \rightarrow x_0$  schneller gegen 0 als jede lineare Funktion in  $x - x_0$  außer der konstanten Funktion 0. Viele wichtige Eigenschaften von  $f$  lassen sich bereits von der „Linearisierung“  $f(x_0) + f'(x_0) \cdot (x - x_0)$  (die in unserem Sinne im Allgemeinen nicht linear, sondern Summe einer konstanten und einer linearen Abbildung ist) ablesen: wenn  $f'(x_0) \neq 0$  ist, ist  $x_0$  keine lokale Extremstelle von  $f$ , und  $f$  besitzt nahe  $x_0$  sogar eine differenzierbare Umkehrfunktion.

Außerdem rechnet man in der Analysis nach, dass Ableiten selbst eine lineare Abbildung ist, beispielsweise vom Vektorraum der differenzierbaren Funktionen auf einem Intervall  $I$  in den Raum aller Funktionen auf  $I$ , denn für differenzierbare Funktionen  $f$  und  $g$  und beliebige reelle Zahlen  $r \in \mathbb{R}$  gilt

$$(f + g)' = f' + g' \quad \text{und} \quad (rf)' = r f' .$$

Auch aufgrund dieser späteren Anwendungen lohnt es sich, lineare Abbildungen und ihre Eigenschaften genauer zu studieren.

2.28. BEISPIEL. Es sei  $M, N$  Rechts- $R$ -Moduln. Dann sind die folgenden Abbildungen immer  $R$ -linear:

(1) Die Identität aus Beispiel 1.19 (1), denn

$$\begin{aligned} \text{id}_M(\ell + m) &= \ell + m = \text{id}_M(\ell) + \text{id}_M(m) , \\ \text{und} \quad \text{id}_M(m \cdot r) &= m \cdot r = \text{id}_M(m) \cdot r . \end{aligned}$$

(2) Die Nullabbildung  $0: M \rightarrow N$  mit  $0(m) = 0_N$  für alle  $m \in M$ , denn

$$\begin{aligned} 0(\ell + m) &= 0_N = 0_N + 0_N = 0(\ell) + 0(m) , \\ \text{und} \quad 0(m \cdot r) &= 0_N = 0_N \cdot r = 0(m) \cdot r . \end{aligned}$$

2.29. PROPOSITION. *Die Hintereinanderausführung von linearen Abbildungen ist linear. Die Umkehrabbildung einer bijektiven linearen Abbildung ist linear. Die Summe linearer Abbildungen ist linear. Das Vielfache einer linearen Abbildung ist linear, wenn  $R$  kommutativ ist.*

BEWEIS. Seien  $L, M$  und  $N$  Rechts- $R$ -Moduln, und seien  $F: M \rightarrow N$  und  $G: L \rightarrow M$   $R$ -linear. Dann folgt aus der Linearität von  $F$  und  $G$  für

alle  $\ell, m \in L$  und alle  $r \in R$ , dass

$$\begin{aligned}(F \circ G)(\ell + m) &= F(G(\ell + m)) = F(G(\ell) + G(m)) \\ &= F(G(\ell)) + F(G(m)) = (F \circ G)(\ell) + (F \circ G)(m), \\ \text{und } (F \circ G)(\ell \cdot r) &= F(G(\ell \cdot r)) = F(G(\ell) \cdot r) \\ &= F(G(\ell)) \cdot r = (F \circ G)(\ell) \cdot r.\end{aligned}$$

Also ist auch  $F \circ G$  linear.

Sei jetzt  $F: M \rightarrow N$  eine bijektive lineare Abbildung und  $G: N \rightarrow M$  ihre Umkehrabbildung, siehe Satz 1.24. Es seien  $p, q \in N$  beliebig und  $\ell = G(p)$ ,  $m = G(q) \in M$ , so dass  $F(\ell) = p$  und  $F(m) = q$ . Außerdem sei  $r \in R$ . Aus der Linearität von  $F$  folgt

$$\begin{aligned}G(p + q) &= G(F(\ell) + F(m)) = G(F(\ell + m)) = \ell + m = G(p) + G(q), \\ \text{und } G(q \cdot r) &= G(F(m) \cdot r) = G(F(m \cdot r)) = m \cdot r = G(q) \cdot r.\end{aligned}$$

Also ist die Umkehrabbildung  $G$  linear.

Seien jetzt  $F, G: M \rightarrow N$  linear, dann ist auch  $F + G$  linear, denn für alle  $\ell, m \in M$  und alle  $r \in R$  gilt

$$\begin{aligned}(F + G)(\ell + m) &= F(\ell + m) + G(\ell + m) = F(\ell) + F(m) + G(\ell) + G(m) \\ &= (F + G)(\ell) + (F + G)(m), \\ (F + G)(m \cdot r) &= F(m \cdot r) + G(m \cdot r) = F(m) \cdot r + G(m) \cdot r \\ &= (F + G)(m) \cdot r.\end{aligned}$$

Sei schließlich  $F: M \rightarrow N$  linear,  $m \in M$  und  $r, s \in R$ . Dann gilt

$$\begin{aligned}(F \cdot r)(m \cdot s) &= F(m \cdot s) \cdot r = F(m) \cdot s \cdot r = F(m) \cdot (sr), \\ (F \cdot r)(m) \cdot s &= F(m) \cdot r \cdot s = F(m) \cdot (rs).\end{aligned}$$

□

Achtung: wenn  $R$  nicht kommutativ ist, zeigt die obige Rechnung, dass  $F \cdot r$  nicht automatisch linear sein muss.

2.30. DEFINITION. Es seien  $M, N$  Rechts- $R$ -Moduln. Bijektive lineare Abbildungen  $F: M \rightarrow N$  heißen (*Rechts- $R$ -*) *Modulisomorphismen*. Lineare Abbildungen  $F: M \rightarrow M$  heißen (*Rechts- $R$ -*) *Modulendomorphismen*, und wenn sie bijektiv sind, (*Rechts- $R$ -*) *Modulautomorphismen*. Falls  $R$  ein Körper ist, sprechen wir von *Vektorraumiso-, -endo- und -automorphismen*. Die Menge aller Modul- oder Vektorraumisomorphismen von  $M$  nach  $N$  wird mit  $\text{Iso}_R(M, N) \subset \text{Hom}_R(M, N)$  bezeichnet, die Menge aller Modul- oder Vektorraumendo- oder -automorphismen von  $M$  mit  $\text{End}_R(M)$  beziehungsweise  $\text{Aut}_R M \subset \text{End}_R M$ . Analoge Bezeichnungen  ${}_R \text{Iso}(M, N)$ ,  ${}_R \text{End } M$  und  ${}_R \text{Aut } M$  führen wir für Links- $R$ -Moduln oder -Vektorräume ein.

Bei  $\text{Aut}_R$  und  $\text{End}_R$  lässt man gelegentlich die Klammern weg, es ist also  $\text{End}_R M = \text{End}_R(M)$ . Analoge Bezeichnungen (Hom, End, Iso und Aut) werden in der Mathematik häufig für Abbildungen benutzt, die eine bestimmte „Struktur“ (hier die eines Moduls beziehungsweise Vektorraums) erhalten.

2.31. FOLGERUNG (aus Proposition 2.29). *Es sei  $R$  ein Ring, und  $M$  und  $N$  seien Rechts- $R$ -Moduln.*

- (1) *Die Automorphismen von  $M$  bilden eine Gruppe  $(\text{Aut}_R M, \circ)$ , die Automorphismengruppe von  $M$ .*
- (2) *Die Endomorphismen von  $M$  bilden einen Ring  $(\text{End}_R M, +, \circ)$  mit Eins  $\text{id}_M$ , den Endomorphismenring von  $M$ .*
- (3) *Der Modul  $M$  ist ein Links- $\text{End}_R M$ -Modul, die skalare Multiplikation wirkt für alle  $F \in \text{End}_R M$  und alle  $m \in M$  durch  $F \cdot m = F(m) \in M$ .*
- (4) *Die Homomorphismen  $\text{Hom}_R(M, N)$  bilden einen unitären Rechts- $\text{End}_R M$ -Modul, und einen unitären Links- $\text{End}_R N$ -Modul.*

Analoge Aussagen gelten, wenn  $M$  und  $N$  Links- $R$ -Moduln sind.

BEWEIS. Der Beweis von (1) orientiert sich am Beispiel 2.5 der Automorphismengruppe einer Menge. Zunächst einmal ist die Verknüpfung zweier Automorphismen ein Automorphismus nach Proposition 2.29, genauso wie die Umkehrabbildung eines Automorphismus. Nach Beispiel 2.28 (1) ist auch die Identität ein Automorphismus. Die Gruppenaxiome ergeben sich wieder aus Bemerkung 2.4 (1)–(3).

Die Addition auf  $\text{End}_R(M)$  in (2) ist die gleiche wie in Proposition 2.29, insbesondere ist die Summe zweier Endomorphismen wieder ein Endomorphismus, und das neutrale Element ist die Nullabbildung aus Beispiel 2.28 (2). Man überprüft leicht die Axiome (G1)–(G4). Aus Bemerkung 2.4 (1) und (2) folgen (R1), (R3). Als nächstes seien  $F, G, H \in \text{End}_R(M)$ , dann gilt

$$(*) \quad (F + G) \circ H = F \circ H + G \circ H \quad \text{und} \quad F \circ (H + K) = F \circ H + F \circ K,$$

wie man durch Einsetzen von  $m \in M$  leicht überprüft. Es folgt (R2) in (2). Also bildet  $(\text{End}_R M, +, \circ)$  einen Ring mit Eins  $1_{\text{End}_R M} = \text{id}_M$ .

Wir lassen den Beweis von (3) und (4) als Übung. □

Es sei  $R$  ein Ring mit Eins. Wir betrachten  $\text{Hom}_R(M, R)$  als Spezialfall von (4), mit  $N = R$  als unitärem Rechts- $R$ -Modul. In Beispiel 2.26 (3) haben wir gesehen, dass  $\text{End}_R R = R$ , wobei  $r \in R = \text{End}_R R$  durch Multiplikation von links wirkt. Also ist  $\text{Hom}_R(M, R)$  ein Links- $R$ -Modul mit  $(r \cdot f)(m) = r \cdot f(m) \in R$ . Wir überprüfen (M1): für  $f \in \text{Hom}_R(M, R)$ ,  $m$  in  $M$  und  $r, s \in R$  gilt

$$((r \cdot s) \cdot f)(m) = (r \cdot s) \cdot f(m) = r \cdot (s \cdot f(m)) = r \cdot (s \cdot f)(m) = (r \cdot (s \cdot f))(m).$$

Umgekehrt ist  ${}_R \text{Hom}(M, R)$  ein Rechts- $R$ -Modul mit  $(f \cdot r)(m) = f(m) \cdot r \in R$ .

2.32. DEFINITION. Sei  $M$  ein Rechts- $R$ -Modul, dann ist  $M^* = \text{Hom}_R(M, R)$  der zu  $M$  duale Links- $R$ -Modul, beziehungsweise der zu  $M$  duale Links- $R$ -Vektorraum, falls  $R$  ein (Schief-) Körper ist. Analog definieren wir den dualen Rechts  $R$ -Modul  ${}^*N$  zu einem Links- $R$ -Modul  $N$ .

### 2.3. Unterräume und Quotienten

In diesem Abschnitt lernen wir, wie man aus gegebenen Moduln neue konstruieren kann. Der Einfachheit halber werden wir ab jetzt meistens nur noch über Ringe mit Eins und unitäre Moduln sprechen.

2.33. DEFINITION. Es sei  $R$  ein Ring mit Eins,  $M$  ein unitärer Rechts- $R$ -Modul und  $U \subset M$  eine Teilmenge. Dann heißt  $U$  ein (*Rechts- $R$ -*) *Unterm modul*, falls für alle  $u, v \in U$  und alle  $r \in R$  die folgenden Untermodulaxiome gelten:

- (U1)  $0_M \in U$  (*Neutrales Element*),
- (U2)  $u + v \in U$ , (*abgeschlossen unter Addition*),
- (U3)  $u \cdot r \in U$  (*abgeschlossen unter skalarer Multiplikation*).

Analog definieren wir Links- $R$ -Unterm oduln von Links- $R$ -Moduln. Falls  $R$  ein (Schief-) Körper ist, sprechen wir stattdessen von (*Rechts-/Links-*) *Untervektorräumen*, kurz *Unterräumen*.

Anstelle von (U1) hätte es gereicht zu fordern, dass  $U \neq \emptyset$ . Denn sei  $u \in U$ , dann folgt  $0_M = u \cdot 0_R \in U$  aus (U3) und Proposition 2.22. Außerdem ist mit  $u \in U$  stets auch

$$-u = u \cdot (-1) \in U.$$

Falls  $R$  keine Eins besitzt oder  $M$  nicht unitär ist, muss man in (U2) zusätzlich  $-u \in U$  fordern.

2.34. BEISPIEL. Wir kennen bereits Beispiele von Untervektorräumen.

- (1) Wir fassen die Quaternionen  $\mathbb{H}$  als  $\mathbb{R}$ -Vektorraum auf. In Abschnitt 1.6 haben wir die Unterräume  $\mathbb{R} \subset \mathbb{H}$  der reellen und  $\mathbb{R}^3 \subset \mathbb{H}$  der imaginären Quaternionen betrachtet.
- (2) In der Analysis trifft man häufig auf Untervektorräume. Beispielsweise bilden die Nullfolgen einen Unterraum des Vektorraums aller Folgen. Für ein offenes Intervall  $I$  bilden die stetigen Funktionen auf  $I$  einen Unterraum des Raumes aller Funktionen auf  $I$ , und die differenzierbaren Funktionen einen Unterraum des Raumes der stetigen Funktionen auf  $I$ .

2.35. BEMERKUNG. Jeder Untermodul  $U$  eines unitären Rechts- $R$ -Moduls  $(M, +, \cdot)$  ist selbst ein unitärer Rechts- $R$ -Modul. Zunächst einmal existiert ein Nullelement  $0_M$  und die Verknüpfungen  $+: U \times U \rightarrow U$ ,  $-: U \rightarrow U$  und  $\cdot: U \times R \rightarrow U$  sind wohldefiniert dank (U1)–(U3). Da die Axiome (G1)–(G4) und (M1)–(M4) gelten, wenn man für die Variablen Elemente aus  $M$  einsetzt, gelten sie erst recht, wenn man nur Elemente aus  $U$  zulässt. Beispielsweise gilt  $0_M + u = u$  in  $M$  für alle  $u \in U$ , also auch in  $U$ .

Die Inklusion  $U \rightarrow M$  aus Bemerkung 1.22 ist linear, da (L1) und (L2) offensichtlich gelten.

Auf völlig analoge Weise kann man *Untergruppen* und *Unterringe* definieren. Entscheidend ist, dass eine Teilmenge der ursprünglichen Struktur mit der

Einschränkung der vorgegebenen Verknüpfungen wieder alle Gruppen- beziehungsweise Ringaxiome erfüllt. Beispielsweise sollte ein Unterring  $U \subset R$  das Element  $0_R$  enthalten, und die Summe, das Produkt und das additive Inverse von Elementen von  $U$  sollten wieder in  $U$  liegen. Bei Körpern bevorzugt man aus sprachlichen Gründen den Begriff *Teilkörper*.

2.36. DEFINITION. Es seien  $M$  und  $N$  Rechts- $R$ -Moduln, und es sei  $F: M \rightarrow N$  rechts- $R$ -linear. Dann definieren wir den *Kern*  $\ker F$  durch

$$\ker F = F^{-1}(\{0_N\}) = \{ m \in M \mid F(m) = 0 \} .$$

Wir erinnern uns auch an das Bild im  $F$ , siehe Definition 1.16.

2.37. PROPOSITION. *Es seien  $M$  und  $N$  Rechts- $R$ -Moduln, und  $F: M \rightarrow N$  sei rechts- $R$ -linear.*

- (1) *Der Kern  $\ker F$  ist ein Untermodul von  $M$ , und  $F$  ist genau dann injektiv, wenn  $\ker F = \{0_M\}$ .*
- (2) *Das Bild  $\text{im } F$  ist ein Untermodul von  $N$ , und  $F$  ist genau dann surjektiv, wenn  $\text{im } F = N$ .*

Die letzte Aussage in (2) ist klar nach Definition 1.18.

BEWEIS. Die Untermodulaxiome für der Kern folgen aus der Linearität von  $F$ , denn für alle  $m, n \in M$  und alle  $r \in R$  gilt

$$\begin{aligned} F(0_M) &= 0_N , \\ F(m) = F(n) = 0_N &\implies F(m+n) = F(m) + F(n) = 0 , \\ F(m) = 0_N &\implies F(m \cdot r) = F(m) \cdot r = 0 \end{aligned}$$

Wenn  $F$  injektiv ist, hat insbesondere  $\ker F = F^{-1}(\{0\})$  höchstens ein Element. Aus  $F(0_M) = 0_N$  folgt dann  $\ker F = \{0_M\}$ .

Sei umgekehrt  $\ker F = \{0_M\}$  und  $F(m) = F(n) \in N$ , dann folgt

$$F(m - n) = F(m) - F(n) = 0_N$$

aus der Additivität (L1) von  $F$ , somit ist  $m - n \in \ker F$ , also nach Voraussetzung  $m - n = 0$ , das heißt  $m = n$ . Also ist  $F$  injektiv, und (1) ist gezeigt.

Die Untermodulaxiome für  $\text{im } F \subset N$  folgen wieder aus der Linearität von  $F$ : für alle  $m, n \in N$ ,  $p, q \in N$  und  $r \in R$  gilt

$$\begin{aligned} 0_N &= F(0_M) , \\ p = F(m) , \quad q = F(n) &\implies p + q = F(m + n) , \\ p = F(m) &\implies p \cdot r = F(p \cdot r) . \quad \square \end{aligned}$$

Wir wollen nun Quotientenmoduln in Analogie zu Beispiel 2.9 konstruieren. Dazu sei  $(M, +, \cdot)$  ein Rechts- $R$ -Modul und  $U \subset M$  ein Untermodul. Dann definieren wir eine Relation „ $\sim$ “ auf  $M$  für alle  $m, n \in M$  durch

$$m \sim n \iff n - m \in U .$$

Das ist eine Äquivalenzrelation, denn (Ä1)–(Ä3) folgen für  $\ell, m, n \in M$  aus

$$\begin{aligned} m - m = 0 \in U, \quad n - m \in U &\implies m - n = -(n - m) \in U, \\ \text{sowie } m - \ell \in U \text{ und } n - m \in U &\implies n - \ell = (n - m) + (m - \ell) \in U. \end{aligned}$$

2.38. DEFINITION. Der Quotient  $M/U = M/\sim$  heißt der *Quotientenmodul* von  $M$  nach  $U$  (lies „ $M$  modulo  $U$ “). Falls  $R$  ein Körper ist heißt  $M/U$  der *Quotientenvektorraum*, kurz *Quotientenraum*.

Man beachte hier, dass wir zur Definition der Äquivalenzrelation „ $\sim$ “ und der Menge  $M/U$  nur die additive Struktur des Moduls  $M$  benutzt haben. Es sei  $p: M \rightarrow M/\sim$  die Quotientenabbildung, siehe Definition 1.43.

2.39. PROPOSITION. *Es sei  $(M, +, \cdot)$  ein unitärer Rechts- $R$ -Modul und  $U \subset M$  ein Untermodul. Dann induzieren „+“ und „ $\cdot$ “ Verknüpfungen*

$$+ : M/U \times M/U \rightarrow M/U \quad \text{und} \quad \cdot : M/U \times R \rightarrow M/U,$$

und  $(M/U, +, \cdot)$  ist ein unitärer Rechts- $R$ -Modul. Außerdem ist die Quotientenabbildung  $p: M \rightarrow M/U$  rechts- $R$ -linear.

BEWEIS. Wir gehen vor wie in Beispiel 2.9. Seien  $m, n, p, q \in M$  mit  $[m] = [n]$  und  $[p] = [q] \in M/U$ , also  $n - m \in U$  und  $q - p \in U$ , und  $r \in R$ , dann folgt

$$\begin{aligned} (n + q) - (m + p) &= (n - m) + (q - p) && \in U, \\ (n \cdot r) - (m \cdot r) &= (n - m) \cdot r && \in U \\ \text{und } (-n) - (-m) &= -(n - m) && \in U, \end{aligned}$$

also sind Addition und skalare Multiplikation auf  $M/U$  wohldefiniert durch

$$[m] + [p] = [m + p], \quad -[m] = [-m] \quad \text{und} \quad [m] \cdot r = [m \cdot r].$$

Wir setzen  $0_{M/U} = [0_M]$ . Jetzt können wir (G1)–(G4), (M1)–(M4) auf die entsprechenden Axiome in  $M$  zurückführen. Beispielsweise gilt (M1), denn

$$([m] \cdot r) \cdot s = [m \cdot r] \cdot s = [(m \cdot r) \cdot s] = [m \cdot (r \cdot s)] = [m] \cdot (r \cdot s).$$

Schließlich zur Linearität der Quotientenabbildung: für alle  $m, n \in M$  und  $r, s \in R$  gilt

$$p(m \cdot r + n \cdot s) = [m \cdot r + n \cdot s] = [m] \cdot r + [n] \cdot s = p(m) \cdot r + p(n) \cdot s. \quad \square$$

2.40. BEISPIEL. Wir betrachten  $M = \mathbb{Z}$  als  $\mathbb{Z}$ -Modul und

$$U = n\mathbb{Z} = \{an \mid a \in \mathbb{Z}\} = \{\dots, -n, 0, n, \dots\}.$$

Dann ist  $U$  ein Untermodul, und der Quotient  $M/U = \mathbb{Z}/n\mathbb{Z}$  entspricht als abelsche Gruppe dem Ring aus Beispiel 2.9. Wir dürfen  $\mathbb{Z}/n\mathbb{Z}$  also auch als  $\mathbb{Z}$ -Modul auffassen.

2.41. BEMERKUNG. In Bemerkung 2.35 haben wir gesehen, dass geeignete Teilmengen von Gruppen, Ringen oder (Schief-) Körpern selbst wieder Gruppen, Ringe beziehungsweise Körper sind. Die Quotientenkonstruktion ist leider nicht so allgemein: Der Quotient einer Gruppe nach einer Untergruppe  $U$  beziehungsweise eines Ringes nach einem Unterring ist nur dann wieder eine Gruppe

beziehungsweise ein Ring, wenn  $U$  gewisse zusätzliche Bedingungen erfüllt (siehe Übungen). Körper und Schiefkörper haben keine Quotienten.

Der folgende Satz entspricht Proposition 1.44 (3).

2.42. PROPOSITION (Universelle Eigenschaft des Quotienten). *Es seien  $M$  und  $N$  Rechts- $R$ -Moduln, es sei  $U \subset M$  ein Untermodul mit Quotientenabbildung  $p: M \rightarrow M/U$ , und es sei  $F: M \rightarrow N$  eine rechts- $R$ -lineare Abbildung. Dann existiert genau dann eine Abbildung  $\bar{F}: M/U \rightarrow N$  mit  $F = \bar{F} \circ p$ , wenn  $U \subset \ker F$ . In diesem Fall ist  $\bar{F}$  eindeutig bestimmt und rechts- $R$ -linear. Es gilt*

$$\operatorname{im} \bar{F} = \operatorname{im} F \quad \text{und} \quad \ker \bar{F} = \ker F/U .$$

Es gilt  $U \subset \ker F$  genau dann, wenn  $F|_U = 0$ . In diesem Fall erhalten wir folgendes Diagramm:

$$\begin{array}{ccccc} U & \xrightarrow{\iota} & M & \xrightarrow{p} & M/U \\ & \searrow & \downarrow F & \exists! & \nearrow \bar{F} \\ & 0 & N & & \end{array}$$

BEWEIS. Zu „ $\implies$ “ nehmen wir an, dass  $\bar{F}$  existiert. Für alle  $u \in U$  gilt  $[u] = 0_{M/U}$ , somit

$$F(u) = \bar{F}([u]) = \bar{F}(0_{M/U}) = F(0_M) = 0_N ,$$

es folgt  $U \subset \ker F$ .

Zu „ $\impliedby$ “ nehmen wir an, dass  $U \subset \ker F$ . Seien  $m, n \in M$  mit  $[m] = [n] \in M/U$ , dann folgt

$$m - n \in U \subset \ker F \implies F(m) - F(n) = F(m - n) = 0_N ,$$

also gilt  $F(m) = F(n)$ , und  $\bar{F}([m]) = F(m)$  ist wohldefiniert.

Die Eindeutigkeit von  $\bar{F}$  folgt aus Proposition 1.44 (3). Außerdem ist  $\bar{F}$  linear, denn

$$\begin{aligned} \bar{F}([m] + [n]) &= F(m + n) = F(m) + F(n) = \bar{F}([m]) + \bar{F}([n]) , \\ \bar{F}([m] \cdot r) &= F(m \cdot r) = F(m) \cdot r = \bar{F}([m]) \cdot r \end{aligned}$$

für alle  $m, n \in M$  und alle  $r \in R$ .

Wir sehen leicht, dass  $\operatorname{im} \bar{F} = \operatorname{im} F$ . Es gilt  $[m] \in \ker \bar{F} \subset M/U$  genau dann, wenn  $m \in \ker F$ , somit folgt

$$\ker \bar{F} = \ker F/U . \quad \square$$

2.43. FOLGERUNG (Homomorphiesatz). *Es seien  $M$  und  $N$  Rechts- $R$ -Moduln und  $F: M \rightarrow N$  linear. Dann induziert  $F$  einen Isomorphismus*

$$\bar{F}: M/\ker F \rightarrow \operatorname{im} F .$$

BEWEIS. Wir wenden Proposition 2.42 an mit  $U = \ker F$ . Da  $\text{im } \bar{F} = \text{im } F$  gilt, dürfen wir  $\bar{F}$  als Abbildung mit Bildbereich  $\text{im } F$  auffassen. Dann ist  $\bar{F}$  linear. Da  $\ker \bar{F} = \ker F / \ker F = \{[0_M]\}$ , ist  $\bar{F}$  injektiv nach Proposition 2.37 (1). Außerdem ist  $\bar{F}$  surjektiv, da  $\text{im } \bar{F} = \text{im } F$ . Also ist  $\bar{F}$  ein Isomorphismus.  $\square$

Wir können also jede lineare Abbildung  $F: M \rightarrow N$  wie folgt zerlegen:

$$\begin{array}{ccc} M & \xrightarrow{F} & N \\ & \searrow p & \nearrow \iota \\ & M/\ker F & \xrightarrow[\cong]{\bar{F}} \text{im } F \end{array}$$

Dabei ist  $p$  die Quotientenabbildung und  $\iota$  die Inklusion. Die Abbildung  $\bar{F}$  ist eindeutig dadurch bestimmt, dass das Diagramm kommutiert. Um  $F$  zu verstehen, bieten sich die folgenden Schritte an.

- (1) Bestimme  $\ker F$  als Untermodul von  $M$ .
- (2) Bestimme  $\text{im } F$  als Untermodul von  $N$ .
- (3) Bestimme den Isomorphismus  $\bar{F}: M/\ker F \rightarrow \text{im } F$ .

2.44. BEISPIEL. Wir betrachten eine Ebene  $V \subset \mathbb{R}^3$  und eine Gerade  $U \subset \mathbb{R}^3$ , so dass sich  $U$  und  $V$  nur in einem Punkt schneiden. Wir wollen annehmen, dass das der Nullpunkt ist; dann sind  $U$  und  $V$  Unterräume. Unsere Anschauung sagt uns, dass es durch jeden Punkt  $x \in \mathbb{R}^3$  genau eine zu  $V$  parallele Gerade gibt, und dass diese Gerade die Ebene  $U$  genau in einem Punkt schneidet. Wir definieren  $F: \mathbb{R}^3 \rightarrow U$  so, dass  $F(x)$  gerade dieser Schnittpunkt ist.

Mit anderen Worten schreiben wir  $x = u + v$  mit  $u \in U$  und  $v \in V$ , und definieren  $F(x) = u$ . Folglich existiert  $F$ , wenn sich jeder Vektor im  $\mathbb{R}^3$  als Summe von Elementen aus  $U$  und  $V$  schreiben lässt. Und  $F$  ist eindeutig bestimmt, denn sei  $x = u + v = w + z$  mit  $u, w \in U$  und  $v, z \in V$ , dann folgt

$$U \ni u - w = z - v \in V.$$

Da wir aber  $U \cap V = \{0\}$  angenommen haben, gilt  $u - w = 0$ , also  $u = w$ .

Man überprüft jetzt leicht, dass die Abbildung  $F$  auch linear ist. Nach Konstruktion werden genau die Punkte auf der Geraden  $V$  auf den Schnittpunkt  $0$  von  $U$  und  $V$  abgebildet, also ist  $\ker F = V$ . Jeder Punkt in der Ebene  $U$  wird auf sich abgebildet, also ist  $F$  insbesondere surjektiv. Nach dem Homomorphiesatz 2.43 induziert  $F$  einen Isomorphismus

$$\mathbb{R}^3/V = \mathbb{R}^3/\ker F \cong \text{im } F = U.$$

Das Besondere hier ist, dass  $U$  selbst ein Unterraum von  $\mathbb{R}^3$  ist mit  $F|_U = \text{id}_U$ .

Sei jetzt wieder  $x \in \mathbb{R}^3$  beliebig. Nach Konstruktion ist  $x - F(x) \in V$ , da eine zu  $V$  parallele Gerade durch  $x$  und  $F(x)$  geht. Es folgt

$$x = u + v \quad \text{mit} \quad u = F(x) \in U \quad \text{und} \quad v = x - F(x) \in V,$$

und wir haben oben gesehen, dass diese Darstellung eindeutig ist. Somit liefern die Unterräume  $U$  und  $V$  ein Beispiel für die folgende Definition.

2.45. DEFINITION. Es sei  $M$  ein Rechts- $R$ -Modul und  $U, V \subset M$  Untermoduln. Die *Summe* von  $U$  und  $V$  ist gegeben durch

$$U + V = \{ u + v \mid u \in U, v \in V \} \subset M .$$

Falls  $U \cap V = \{0\}$ , heißt die Summe *direkt*, und wir schreiben statt  $U + V$  auch  $U \oplus V$ . Falls  $M$  die direkte Summe  $U \oplus V$  ist, sagen wir, dass  $V$  ein *Komplement* von  $U$  in  $M$  ist (und umgekehrt), oder, dass  $U$  und  $V$  *komplementäre Untermoduln* sind. Wenn  $R$  ein (Schief-) Körper ist, sprechen wir analog von *komplementären Unterräumen*.

Man beachte, dass wegen (U1) stets  $0_M \in U \cap V$  gilt. Einen kleineren Durchschnitt als  $\{0_M\}$  können zwei Untermoduln also nicht haben.

2.46. BEISPIEL. Wir geben Beispiele von direkten Summen und komplementären Untermoduln an.

- (1) In den Übungen zeigen Sie, dass  $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/6\mathbb{Z}$ . Also sind die Untermoduln

$$U = 2\mathbb{Z}/6\mathbb{Z} = \{[0], [2], [4]\} \cong \mathbb{Z}/3\mathbb{Z}$$

und  $V = 3\mathbb{Z}/6\mathbb{Z} = \{[0], [3]\} \cong \mathbb{Z}/2\mathbb{Z}$

von  $M = \mathbb{Z}/6\mathbb{Z}$  zueinander komplementär.

- (2) Ähnlich wie in (1) betrachte

$$V = 2\mathbb{Z}/4\mathbb{Z} = \{[0], [2]\} \subset M = \mathbb{Z}/4\mathbb{Z} .$$

Dann ist  $V \cong \mathbb{Z}/2\mathbb{Z}$ . Es gibt keinen komplementären Untermodul  $U$ , denn dieser müsste mindestens ein Element aus  $M \setminus V$  enthalten, also entweder  $[1]$  oder  $[3]$ . In beiden Fällen wäre  $[2] \in U$ , denn  $[2] = [1] + [1] = [3] + [3]$ , und somit  $U \cap V \neq \{[0]\}$ . Also existiert nicht immer ein komplementärer Untermodul.

Es sei  $V \subset M$  ein Untermodul. Wir erinnern uns an die Quotientenabbildung  $p: M \rightarrow M/V$  aus Proposition 2.39.

2.47. PROPOSITION. *Es seien  $U, V$  Untermoduln eines Rechts- $R$ -Moduls  $M$ .*

- (1) *Die Summe  $U + V \subset M$  ist ein Untermodul.*  
 (2) *Wenn die Summe direkt ist, existiert eine bijektive Abbildung*

$$U \times V \rightarrow U \oplus V \quad \text{mit} \quad (u, v) \mapsto u + v .$$

- (3) *Es sei  $p: M \rightarrow M/V$  die Quotientenabbildung. Wenn  $U$  und  $V$  komplementäre Untermoduln sind, dann ist  $p|_U: U \rightarrow M/V$  ein Modulisomorphismus.*

BEWEIS. Die Unterraumaxiome für  $U + V$  gelten, da

$$\begin{aligned} 0_M &= 0_M + 0_M && \in U + V , \\ (t + v) + (u + w) &= (t + u) + (v + w) && \in U + V \\ \text{und} \quad (u + v) \cdot r &= u \cdot r + v \cdot r && \in U + V \end{aligned}$$

für alle  $t, u \in U, v, w \in V$  und  $r \in R$ .

Die Abbildung in (2) ist immer surjektiv nach Definition der Summe. Wenn die Summe direkt ist, ist für jedes Element  $s \in U \oplus V$  die Zerlegung  $s = u + v$  mit  $u \in U$  und  $v \in V$  eindeutig, denn aus  $s = u' + v'$  mit  $u' \in U, v' \in V$  folgt

$$u' - u = v - v' \in U \cap V \implies u' - u = v - v' = 0_M .$$

Also ist die Abbildung in (2) auch injektiv.

Die Quotientenabbildung  $p: M \rightarrow M/V$  ist linear nach Proposition 2.39. Die Inklusion  $\iota: U \rightarrow M$  ist linear nach Bemerkung 2.35. Also ist auch die Abbildung  $p|_U = p \circ \iota$  in (3) linear nach Proposition 2.29.

Sei  $[m] \in M/V$  mit  $m \in M$ , dann existieren  $u \in U, v \in V$  mit  $m = u + v$ , da  $M = U \oplus V$ . Da  $p(u) = [u] = [m]$ , ist  $p|_U$  immer surjektiv.

Aus  $p(u) = p(u') \in M/V$  folgt, dass ein  $v \in V$  existiert mit  $u' = u + v$ . Wie in (2) folgt aus  $v = u - u' \in U \cap V$ , dass  $u = u'$ , wenn die Summe direkt ist. Also ist  $p|_U$  injektiv.  $\square$

2.48. BEMERKUNG. Wir können also den Quotientenmodul  $M/V$  mit Hilfe von  $p|_U$  mit einem komplementären Untermodul  $U$  identifizieren, falls ein solcher existiert. Wenn  $U$  ein zu  $V$  komplementärer Untermodul ist, gibt es meistens noch andere komplementäre Untermoduln, siehe etwa Beispiel 2.44, wo man in Richtung von  $V$  auf verschiedene Ebenen in  $\mathbb{R}^3$  projizieren kann. Das bedeutet, dass diese Beschreibung von  $M/V$  als Untermodul von  $M$  von der Wahl des Komplements  $U$  abhängt. Obwohl man oft leichter mit einem komplementären Untermodul  $U$  als mit dem Quotienten  $M/V$  arbeiten kann, ist es daher manchmal sinnvoll, den Quotienten  $M/V$  zu betrachten.

Wenn  $U$  und  $V$  gegeben sind, können wir  $U \times V$  wie in Proposition 2.47 (2) als Modul betrachten und  $U$  und  $V$  mit den komplementären Untermoduln  $U \times \{0\}, \{0\} \times V \subset M$  identifizieren. Wir nennen  $U \times V$  die *direkte Summe*  $U \oplus V$  von  $U$  und  $V$ .

Die direkte Summe erfüllt gleich zwei „universelle Eigenschaften“. Sei dazu  $M = U \oplus V$ , dann betrachten wir die Inklusionsabbildungen  $\iota_U: U \rightarrow M$  und  $\iota_V: V \rightarrow M$ . Wenn wir wie oben  $M/U \cong V$  und  $M/V \cong U$  identifizieren, erhalten wir auch Projektionen  $p_U: M \rightarrow U$  und  $p_V: M \rightarrow V$ , so dass insbesondere  $m = p_U(m) + p_V(m)$  für alle  $m \in M$ .

2.49. PROPOSITION (Universelle Eigenschaften der direkten Summe). *Es sei  $M$  ein Rechts- $R$ -Moduln und  $U, V \subset M$  Untermoduln, so dass  $M = U \oplus V$ .*

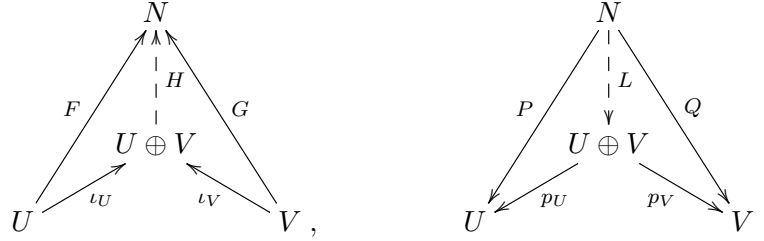
- (1) *Die Inklusions- und Projektionsabbildungen erfüllen*

$$\begin{aligned} p_U \circ \iota_U &= \text{id}_U , & p_U \circ \iota_V &= 0: V \rightarrow U , \\ p_V \circ \iota_U &= 0: U \rightarrow V & \text{und} & & p_V \circ \iota_V &= \text{id}_V . \end{aligned}$$

- (2) *Universelle Eigenschaft des Koproduktes: Sei  $N$  ein weiterer Rechts- $R$ -Modul und seien  $F: U \rightarrow N$  und  $G: V \rightarrow N$  linear, dann existiert genau eine lineare Abbildung  $H: U \oplus V \rightarrow N$ , so dass  $F = H \circ \iota_U$  und  $G = H \circ \iota_V$ .*

- (3) Universelle Eigenschaft des Produktes: Sei  $N$  ein weiterer Rechts- $R$ -Modul und seien  $P: N \rightarrow U$  und  $Q: N \rightarrow V$  linear, dann existiert genau eine lineare Abbildung  $L: N \rightarrow U \oplus V$ , so dass  $P = p_U \circ L$  und  $Q = p_V \circ L$ .

Wie eng diese beiden Eigenschaften miteinander verwandt sind, zeigen die folgenden Diagramme, die sich nur in der Richtung der Pfeile unterscheiden. Man sagt auch, die Diagramme sind zueinander *dual*.



BEWEIS. Zu (1) sei  $u \in U$ , dann folgt

$$\begin{aligned} (p_U \circ \iota_U)(u) &= p_U(u + 0_M) = u = \text{id}_U(u), \\ (p_V \circ \iota_U)(u) &= p_V(u + 0_M) = 0 = 0(u) \in V. \end{aligned}$$

Also gilt  $p_U \circ \iota_U = \text{id}_U$  und  $p_V \circ \iota_U = 0$ . Die beiden anderen Gleichungen folgen genauso.

Zu (2) zeigen wir zunächst die Eindeutigkeit. Sei also eine lineare Abbildung  $H$  gegeben mit  $H \circ \iota_U = F$  und  $H \circ \iota_V = G$ . Für  $m = u + v$  folgt

$$\begin{aligned} H(m) &= H(u) + H(v) = H(\iota_U(u)) + H(\iota_V(v)) \\ &= F(u) + G(v) = F(p_U(m)) + G(p_V(m)), \end{aligned}$$

also ist  $H$  eindeutig bestimmt.

Auf der anderen Seite ist die Abbildung  $F \circ p_U + G \circ p_V$  linear nach Proposition 2.29. Sie leistet das Gewünschte, denn wegen (1) gilt

$$\begin{aligned} (F \circ p_U + G \circ p_V) \circ \iota_U &= F \circ \underbrace{p_U \circ \iota_U}_{=\text{id}_U} + G \circ \underbrace{p_V \circ \iota_U}_{=0} = F, \\ (F \circ p_U + G \circ p_V) \circ \iota_V &= F \circ p_U \circ \iota_V + G \circ p_V \circ \iota_V = G. \end{aligned}$$

Der Beweis zu (3) verläuft analog. Es sei  $n \in N$  und  $m = L(n) = u + v$  mit  $u \in U$  und  $v \in V$ , dann folgt  $u = p_U(L(n)) = P(n)$  und  $v = p_V(L(n)) = Q(n)$ , also ist  $L$  eindeutig bestimmt.

Umgekehrt ist die Abbildung

$$\iota_U \circ P + \iota_V \circ Q: N \rightarrow M = U \oplus V$$

linear nach Proposition 2.29. Mithilfe von (1) überprüft man wieder, dass

$$p_U \circ (\iota_U \circ P + \iota_V \circ Q) = P \quad \text{und} \quad p_V \circ (\iota_U \circ P + \iota_V \circ Q) = Q. \quad \square$$

## 2.4. Linearkombinationen, Basen und Koordinaten

Bisher haben wir Moduln sehr abstrakt betrachtet. Auf der anderen Seite haben wir uns in den Abschnitten 1.4–1.6 die konkreten Vektorräume  $\mathbb{R}^n$  angeschaut, speziell für  $n = 2$  und  $n = 3$ . Im Rest dieses Kapitels wollen wir eine Brücke zwischen beiden Welten schlagen. Damit das alles reibungslos funktioniert, betrachten wir nur noch Ringe mit Eins und unitäre Moduln.

2.50. BEMERKUNG. Es sei  $I$  eine Menge und  $M$  ein Rechts- $R$ -Modul. Dann bilden die Abbildungen  $\text{Abb}(I, M)$  wieder einen Rechts- $R$ -Modul  $M^I$ . Für alle  $f, g: I \rightarrow M$  und alle  $r \in R$  definieren wir  $f + g, f \cdot r \in M^I$  durch

$$(f + g)(i) = f(i) + g(i) \quad \text{und} \quad (f \cdot r)(i) = f(i) \cdot r \in M.$$

Die Gruppenaxiome (G1)–(G4) und die Modulaxiome (M1)–(M4) lassen sich für jedes Element  $i \in I$  einzeln überprüfen. Beispielsweise folgt (M2) für  $M^I$  aus (M2) für  $M$ , da

$$a(i) \cdot (r + s) = a(i) \cdot r + a(i) \cdot s.$$

Der folgende Spezialfall ist zentral für die Vorlesung.

2.51. BEISPIEL. Es sei  $R$  ein Ring mit Eins und  $n \in \mathbb{N}$ . Wir fassen  $R$  als Rechts- $R$ -Modul auf wie in Beispiel 2.21 (1) und bezeichnen das  $n$ -fache kartesische Produkt von  $R$  mit  $R^n = \underbrace{R \times \cdots \times R}_{n \text{ Faktoren}}$ . Traditionell schreibt man Elemente von  $R^n$  als *Spaltenvektoren*

$$(1) \quad R^n = \left\{ \left( \begin{array}{c} a_1 \\ \vdots \\ a_n \end{array} \right) \mid a_1, \dots, a_n \in R \right\}.$$

Wir identifizieren  $a \in R^n$  mit einer Abbildung  $a: \{1, \dots, n\} \rightarrow M$ , indem wir jedem  $i \in \{1, \dots, n\}$  das Element  $a_i \in M$  zuordnen. Wie oben erhalten wir einen Rechts- $R$ -Modul mit den Rechenoperationen

$$(2) \quad \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot r = \begin{pmatrix} a_1 \cdot r \\ \vdots \\ a_n \cdot r \end{pmatrix}.$$

Für alle  $1 \leq j \leq n$  sei

$$(3) \quad e_j = \begin{pmatrix} \delta_{1j} \\ \vdots \\ \delta_{nj} \end{pmatrix} = j \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in R^n$$

der  $i$ -te *Standardbasisvektor*, hierbei heißt

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j, \text{ und} \\ 0 & \text{falls } i \neq j \end{cases}$$

das *Kroneckersymbol*. Der Vektor  $e_j$  hat also als  $j$ -ten Eintrag die 1, und sonst überall 0. Wir nennen  $(e_1, \dots, e_n)$  die *Standardbasis* des  $R^n$ . Wir können jedes

Element  $a \in R^n$  als Linearkombination der Standardbasis schreiben:

$$(4) \quad a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdot a_1 + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \cdot a_n = \sum_{i=1}^n e_i \cdot a_i .$$

Man überzeugt sich leicht, dass diese Darstellung auch eindeutig ist, das heißt, aus  $a = \sum_{i=1}^n e_i \cdot r_i$  mit  $r_1, \dots, r_n \in R$  folgt  $r_i = a_i$ .

2.52. BEISPIEL. Wir können auf der Menge  $R^n$  analog die Addition wie oben und eine Linksmultiplikation  $r \cdot (a_1, \dots, a_n) = (ra_1, \dots, ra_n)$  definieren, wobei wir die Elemente jetzt als *Zeilenvektoren* schreiben. Dann erhalten wir einen Links- $R$ -Modul  ${}^nR$ . Die Gruppen- und Modulaxiome werden wie oben bewiesen. Als abelsche Gruppen sind  $R^n$  und  ${}^nR$  isomorph, als Moduln jedoch nur dann, wenn  $R$  kommutativ ist, siehe Bemerkung 2.23. Als Standardbasisvektoren wählen wir entsprechend

$$\varepsilon_1 = (1, 0, \dots, 0), \quad \dots, \quad \varepsilon_n = (0, \dots, 0, 1) .$$

Auch in diesem Fall lässt sich jedes Element  $a \in {}^nR$  eindeutig darstellen als

$$(a_1, \dots, a_n) = \sum_{i=1}^n a_i \cdot \varepsilon_i .$$

Es sei  $(G, +)$  eine abelsche Gruppe,  $n \in \mathbb{N}$ , und  $a_1, \dots, a_n \in G$ . Wir setzen  $s_0 = 0 \in G$  und definieren rekursiv

$$s_i = s_{i-1} + a_i \in G \quad \text{für } i = 1, \dots, n .$$

Dann ist die *Summe der  $a_n$  für  $i$  von 1 bis  $n$*  definiert als

$$(2.1) \quad \sum_{i=1}^n a_i = s_n = a_1 + \cdots + a_n \in G .$$

Diese Konstruktion ist die Grundlage für die folgende Definition.

2.53. DEFINITION. Sei  $M$  ein Rechts- $R$ -Modul,  $n \in \mathbb{N}$ , und seien  $m_1, \dots, m_n \in M$ . Für alle  $r_1, \dots, r_n \in R$  heißt

$$(1) \quad \sum_{i=1}^n m_i \cdot r_i \in M .$$

eine *Linearkombination* der Elemente  $m_1, \dots, m_n$ .

2.54. BEMERKUNG. Linearkombinationen werden uns regelmäßig begegnen. Zum Beispiel haben wir im Beweis der Cauchy-Schwarz-Ungleichung 1.54 eine Linearkombination  $y - (\langle x, y \rangle / \|x\|^2) x$  der Vektoren  $x$  und  $y$  betrachtet, die senkrecht auf  $x$  steht. Im Beweis von Satz 1.75 haben wir einen Vektor  $w \in \mathbb{R}^3$  mit  $\langle v, w \rangle = 0$  um die Achse durch  $v$  gedreht und das Ergebnis als Linearkombination der Vektoren  $w$  und  $v \times w$  geschrieben.

Die folgende Überlegung ist die Grundlage für das Rechnen mit Matrizen.

2.55. SATZ (Universelle Eigenschaft des freien Moduls). *Es sei  $R$  ein Ring mit Eins,  $n \in \mathbb{N}$ , und  $M$  sei ein unitärer Rechts- $R$ -Modul. Dann existiert eine bijektive Abbildung*

$$(1) \quad \Phi: \text{Hom}_R(R^n, M) \xrightarrow{\cong} M^n$$

mit  $f \mapsto (f(e_1), \dots, f(e_n))$ .

Die Umkehrabbildung  $\Psi$  ordnet einem Tupel  $A = (a_1, \dots, a_n) \in M^n$  eine lineare Abbildung  $\Psi_A: R^n \rightarrow M$  zu mit

$$(2) \quad \Psi_A \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right) = \sum_{i=1}^n a_i \cdot r_i.$$

Mit anderen Worten ist eine lineare Abbildung  $R^n \rightarrow M$  genau durch die Bilder der Standardbasisvektoren  $e_1, \dots, e_n$  bestimmt, die wir frei vorgeben dürfen. Das ist die *universelle Eigenschaft des freien Moduls  $R^n$* . Wir schreiben das als Diagramm

$$\begin{array}{ccc} \{e_1, \dots, e_n\} & \hookrightarrow & R^n \\ & \searrow & \downarrow \\ & & M \end{array} \quad \begin{array}{c} \exists! \Psi_A \\ \downarrow \end{array}$$

Das Symbol „ $\exists!$ “ bedeutet „es gibt genau ein“. Wir hatten  $\text{Hom}_R(R^n, M)$  in Folgerung 2.31 (4) als Links- $\text{End}_R(M)$ - und Rechts- $\text{End}_R(R^n)$ -Modul aufgefasst. Auch  $M^n$  trägt solche eine Struktur, uns interessiert  $M^n$  aber im Moment nur als Menge.

BEWEIS. Um die Bijektivität in (1) zu prüfen, definieren wir  $\Psi$  wie in (2). Für ein beliebiges Tupel  $A = (a_1, \dots, a_n) \in M^n$ ,  $r, s \in R^n$  und  $t \in R$  gilt nach Beispiel 2.51 (2), dass

$$\begin{aligned} \Psi_A \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} + \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \right) &= \sum_{i=1}^n a_i \cdot (r_i + s_i) = \sum_{i=1}^n a_i \cdot r_i + \sum_{i=1}^n a_i \cdot s_i \\ &= \Psi_A \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right) + \Psi_A \left( \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \right) \\ \text{und} \quad \Psi_A \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \cdot t \right) &= \sum_{i=1}^n a_i \cdot (r_i \cdot t) = \sum_{i=1}^n (a_i \cdot r_i) \cdot t \\ &= \Psi_A \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right) \cdot t. \end{aligned}$$

Das mittlere Gleichheitszeichen folgt jeweils durch vollständige Induktion. Also ist  $\Psi_A: R^n \rightarrow M$  linear, und wir haben  $\Psi: M^n \rightarrow \text{Hom}_R(R^n, M)$  konstruiert.

Wir starten mit  $A \in M^n$ . Aus der Definition von  $e_i$  in Beispiel 2.51 (3) und der Konstruktion von  $\Psi_A: R^n \rightarrow M$  folgt, dass

$$\Psi_A(e_j) = \sum_{i=1}^n a_i \cdot \delta_{ij} = a_j,$$

also erhalten wir das ursprüngliche Tupel zurück.

Sei jetzt  $f: R^n \rightarrow M$  rechts- $R$ -linear und  $A = \Phi(f) = (f(e_1), \dots, f(e_n))$ . Um  $\Psi_A = f$  zu zeigen, benutzen wir Beispiel 2.51 (4). Durch Induktion über  $n$  folgt aus Linearität von  $\Psi_A$  und  $f$ , dass

$$\begin{aligned} \Psi_A \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right) &= \Psi_A \left( \sum_{i=1}^n e_i \cdot r_i \right) = \sum_{i=1}^n \Psi_A(e_i) \cdot r_i = \sum_{i=1}^n f(e_i) \cdot r_i \\ &= f \left( \sum_{i=1}^n e_i \cdot r_i \right) = f \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right), \end{aligned}$$

wir erhalten also  $f$  zurück. Somit sind  $\Phi$  und  $\Psi$  zueinander invers.  $\square$

**2.56. FOLGERUNG.** *Es sei  $f: M \rightarrow N$  eine lineare Abbildung zwischen unitären Rechts- $R$ -Moduln, und es sei  $A = (a_1, \dots, a_n)$  ein  $n$ -Tupel in  $M$ . Für alle  $r_1, \dots, r_n \in R$  gilt dann*

$$f \left( \sum_{i=1}^n a_i \cdot r_i \right) = \sum_{i=1}^n f(a_i) \cdot r_i.$$

Wir sagen auch „lineare Abbildungen sind mit Linearkombinationen verträglich“, siehe Bemerkung 2.25 (2).

**BEWEIS.** Wir bezeichnen das  $n$ -Tupel  $(f(a_1), \dots, f(a_n))$  in  $N$  mit  $B$ . Nach Satz 2.55 erhalten wir eine lineare Abbildungen  $\Psi_B: R^n \rightarrow N$ . Nach Proposition 2.29 ist  $f \circ \Psi_A: R^n \rightarrow N$  auch linear. Da beide Abbildungen die Standardbasisvektoren auf dieselben Elemente  $f(a_i) \in N$  abbilden, folgt  $\Psi_B = f \circ \Psi_A: R^n \rightarrow N$ . Für ein konkretes Element von  $R^n$  bedeutet das, dass

$$f \left( \sum_{i=1}^n a_i \cdot r_i \right) = (f \circ \Psi_A) \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right) = \Psi_B \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right) = \sum_{i=1}^n f(a_i) \cdot r_i. \quad \square$$

Mit Hilfe des Begriffs „Linearkombination“ können wir jetzt ganz klassisch sagen, wann ein Tupel  $A$  von Elementen ein Modul  $M$  erzeugt, linear unabhängig ist, oder eine Basis bildet. Im Anschluss überlegen wir uns, was das mit der Abbildung  $\Psi_A$  zu tun hat.

**2.57. DEFINITION.** Es sei  $M$  ein unitärer Rechts- $R$ -Modul,  $n \in \mathbb{N}$ , und  $A = (a_1, \dots, a_n)$  ein  $n$ -Tupel in  $M$ . Ein Element  $m \in M$  heißt *als Linearkombination*

der  $a_1, \dots, a_n$  darstellbar, wenn es  $r_1, \dots, r_n \in R$  gibt, so dass  $\sum_{i=1}^n a_i \cdot r_i = m$ . Das Erzeugnis von  $A$  (über  $R$ ) in  $M$  ist die Menge

$$\langle A \rangle = \langle A \rangle_R = \left\{ \sum_{i=1}^n a_i \cdot r_i \mid r_1, \dots, r_n \in R \right\} \subset M.$$

Falls  $M = \langle A \rangle$ , heißt  $A$  ein Erzeugendensystem von  $M$ , und  $A$  erzeugt  $M$  (über  $R$ ). Wenn es ein  $A$  wie oben gibt, das  $M$  erzeugt, heißt  $M$  endlich erzeugt (über  $R$ ).

Das Erzeugnis einer Menge  $E$  wird manchmal auch mit  $\text{span}(E)$  bezeichnet.

2.58. BEISPIEL. Als Vektorraum über  $\mathbb{C}$  wird  $\mathbb{C}$  selbst erzeugt von der Menge  $\{1\}$ . Wir können  $\mathbb{C} \cong \mathbb{R}^2$  aber auch als Vektorraum über  $\mathbb{R}$  auffassen. Über  $R$  erzeugt  $\{1\}$  nur die Teilmenge  $\mathbb{R} \subset \mathbb{C}$ , während  $\{1, i\}$  eine Erzeugermenge über  $\mathbb{R}$  ist. Aus diesem Grund ist es manchmal wichtig, den zugrundeliegenden Ring oder Körper mit anzugeben.

Noch schlimmer wird es, wenn wir  $\mathbb{C}$  als Vektorraum über  $\mathbb{Q}$  auffassen. Da  $\mathbb{Q}$  abzählbar und  $\mathbb{C}$  überabzählbar ist, ist  $\mathbb{C}$  über  $\mathbb{Q}$  nicht endlich erzeugt.

2.59. DEFINITION. Es sei  $M$  ein unitärer Rechts- $R$ -Modul,  $n \in \mathbb{N}$  und  $A = (a_1, \dots, a_n)$  ein  $n$ -Tupel in  $M$ . Falls für alle  $r_1, \dots, r_n \in R$  gilt, dass

$$0 = \sum_{i=1}^n a_i \cdot r_i \quad \implies \quad r_1 = \dots = r_n = 0,$$

dann heißt  $A$  linear unabhängig. Andernfalls heißt  $A$  linear abhängig.

Sei  $M$  ein Rechts- $R$ -Modul. Eine (endliche) Basis von  $M$  ist ein linear unabhängiges Erzeugendensystem  $A$  von  $M$ . Ein endlich erzeugter unitärer Rechts- $R$ -Modul  $M$  heißt frei (über  $R$ ), wenn er eine Basis besitzt.

Beachte, dass wir Basen immer als Tupel, nicht als Teilmengen von  $M$  auffassen. Manche Autoren sprechen daher von „angeordneten Basen“. Später benutzen wir für Basen den Buchstaben  $B$  anstelle von  $A$ .

2.60. BEISPIEL. Es sei  $n \geq 1$  und  $M = \mathbb{Z}/n\mathbb{Z}$  der  $\mathbb{Z}$ -Modul aus Beispiel 2.40. Für alle  $[a] \in \mathbb{Z}/n\mathbb{Z}$  gilt  $[a] \cdot n = [an] = [0]$ , also ist jede nichtleere Teilmenge  $E \subset \mathbb{Z}/n\mathbb{Z}$  linear abhängig. Genauer: sei  $f = [a] \in E$ , dann wähle  $(r_e)_{e \in E} = (\delta_{ef} \cdot n)_{e \in E}$ ; es folgt

$$\sum_{e \in E} e \cdot (\delta_{ef} \cdot n) = [a] \cdot n = [0],$$

da der Faktor  $\delta_{ef}$  in einer Summe über  $e$  nach Definition des Kronecker-Symbols nur den Summanden mit  $e = f$  übriglässt.

Auf der anderen Seite erzeugt die leere Menge den Modul  $\mathbb{Z}/n\mathbb{Z}$  nur dann, wenn  $n = 1$ . Somit hat  $\mathbb{Z}/n\mathbb{Z}$  keine Basis über  $\mathbb{Z}$ , wenn  $n > 1$ .

Allerdings ist  $M = \mathbb{Z}/n\mathbb{Z}$  ein freier Modul über dem Ring  $R = \mathbb{Z}/n\mathbb{Z}$  mit Basis  $E = \{[1]\}$ , denn  $E$  erzeugt  $M$ . Aus  $[0] = [1] \cdot r$  folgt  $r = [0]$ , da  $[1]$  gleichzeitig das Einselement von  $R$  ist. Also ist  $E$  auch linear unabhängig über  $R$ . Es ist also auch bei linearer Abhängigkeit wichtig, den Grundring mit anzugeben.

2.61. BEISPIEL. Als Vektorraum über  $\mathbb{C}$  hat  $\mathbb{C}$  zum Beispiel die Basis (1). Als Vektorraum über  $\mathbb{R}$  bildet  $(1, i)$  eine Basis.

2.62. FOLGERUNG (aus Satz 2.55). *Es sei  $M$  ein unitärer Rechts- $R$ -Modul,  $n \in \mathbb{N}$ , es sei  $A = (a_1, \dots, a_n)$  ein  $n$ -Tupel in  $M$  und  $\Psi_A: R^n \rightarrow M$  die zugehörige Abbildung aus Satz 2.55. Dann gilt*

- (1)  $\Psi_A: R^n \rightarrow M$  ist surjektiv  $\iff A \in M^n$  erzeugt  $M$ ,
- (2)  $\Psi_A: R^n \rightarrow M$  ist injektiv  $\iff A \in M^n$  ist linear unabhängig,
- (3)  $\Psi_A: R^n \rightarrow M$  ist bijektiv  $\iff A \in M^n$  ist eine Basis von  $M$ .

BEWEIS. Aussage (1) folgt unmittelbar aus Definition 2.57. Zu (2) benutzen wir, dass  $\Psi_A$  nach Proposition 2.37 (1) genau dann injektiv ist, wenn  $\ker \Psi_A = \{0\} \subset R^n$  gilt, was nach Definition 2.59 äquivalent zur linearen Unabhängigkeit von  $A$  ist. Schließlich folgt (3) aus (1) und (2).  $\square$

2.63. DEFINITION. Es sei  $M$  ein freier Rechts- $R$ -Modul mit Basis  $B = (b_1, \dots, b_n)$ , und es sei  $m \in M$ . Dann heißt die Abbildung  $\Psi_B$  aus Satz 2.55 die *Basisabbildung* von  $M$  zur Basis  $B$ , und ihre Umkehrabbildung die *Koordinatenabbildung* zur Basis  $B$ . Für  $m \in M$  heißen  $r_1, \dots, r_n \in R$  mit

$$\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} = \Psi_B^{-1}(m)$$

die *Koordinaten* von  $m$  bezüglich der Basis  $B$ .

Äquivalent zu  $r = \Psi_B^{-1}(m) \in R^n$  ist somit die Koordinatendarstellung

$$m = \sum_{i=1}^n b_i \cdot r_i.$$

2.64. BEMERKUNG. Wir kommen noch einmal auf Satz 2.55 zurück. Um eine lineare Abbildung  $F$  von einem freien Modul  $M$  mit Basis  $B = (b_1, \dots, b_n)$  in einen beliebigen Modul  $N$  anzugeben, können wir die Bilder  $F(b_i) = a_i \in N$  der Basiselemente *frei* vorgeben. Sei nämlich  $A = (a_1, \dots, a_n)$  das zugehörige  $n$ -Tupel und  $\Psi_A: R^n \rightarrow N$  die entsprechende lineare Abbildung aus Satz 2.55. Dann ist  $f = \Psi_A \circ \Psi_B^{-1}$  linear und bildet jeweils  $b_i$  auf  $a_i$  ab. Die *universelle Eigenschaft* eines freien Moduls gilt in diesem Sinne also für jeden Modul, der eine endliche Basis besitzt.

2.65. BEMERKUNG. Es sei  $R$  Ring mit Eins und  $M$  ein Rechts- $R$ -Modul. In Definition 2.32 haben wir den dualen Links- $R$ -Modul  $M^* = \text{Hom}_R(M; R)$  eingeführt. Im Spezialfall  $M = R^m$  folgt aus Satz 2.55, dass

$$(R^m)^* = \text{Hom}_R(R^m, R) = {}^m R$$

- (1) mit  $(a_1, \dots, a_m) \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} = \sum_{i=1}^m a_i \cdot r_i \in R.$

Für  $a \in {}^mR$  und  $s \in R$  ist die Linkswirkung definiert durch  $(s \cdot a)(r) = s \cdot (a(r))$  für alle  $r \in R^m$ , genau wie in Beispiel 2.52. Somit ist der Links- $R$ -Modul der  $m$ -elementigen Zeilen dual zum Rechts- $R$ -Modul der  $m$ -elementigen Spalten.

Es sei wieder  $(\varepsilon_i)_i$  die Standardbasis des Raumes  ${}^mR$  aus Beispiel 2.52. Zwischen den Basen  $(e_j)_j$  von  $R^m$  und  $(\varepsilon_i)_i$  von  ${}^mR$  besteht die folgenden Beziehung:

$$(2) \quad \varepsilon_i(e_j) = \sum_{k=1}^m \delta_{ik} \delta_{kj} = \delta_{ij};$$

wir sagen dazu, dass die Basis  $(\varepsilon_i)_i$  *dual* zur Basis  $(e_j)_j$  ist.

Für Links-Moduln  $N$  haben wir analog den Dualraum  ${}^*N = {}_R \text{Hom}(N, R)$  definiert. Analog zu oben folgt  ${}^*({}^mR) = R^m$ , und wiederum ist die Basis  $(e_j)_j$  zur Basis  $(\varepsilon_i)_i$  dual.

**2.66. PROPOSITION.** *Es sei  $R$  ein Ring mit Eins und  $M$  ein freier Modul mit Basis  $B = (b_1, \dots, b_m)$ . Dann bilden die einzelnen Komponentenfunktionen  $\beta_i = \varepsilon_i \circ \beta: M \rightarrow R$  der Koordinatenabbildung  $\beta = \Psi_B^{-1}: M \rightarrow R^m$  zu  $B$  eine Basis  $\beta = (\beta_i)_i$  des dualen Moduls  $B^*$ . Sie ist dual zur Basis  $B$ , das heißt, für alle  $i, j$  gilt*

$$(1) \quad \beta_i(b_j) = \delta_{ij}.$$

Wir dürfen die Koordinatenabbildung  $\beta = \Psi_B^{-1}$  zu einer Basis  $B$  also als Basisabbildung des dualen Moduls auffassen.

**BEWEIS.** Die Abbildungen  $\beta_i$  sind als Komponenten von  $\beta$  wieder linear und somit Elemente des dualen Moduls  $M^*$ . Aus  $b_j = \Psi_B(e_j)$  und  $\beta = \Psi_B^{-1}$  folgt (1), denn

$$\beta_i(b_j) = (\varepsilon_i \circ \beta)(\Psi_B(e_j)) = \varepsilon_i(e_j) = \delta_{ij}$$

nach Bemerkung 2.65 (2).

Für ein beliebiges  $\alpha \in M^*$  und  $j \in \{1, \dots, m\}$  folgt

$$\alpha(b_j) = \sum_{i=1}^m \alpha(b_i) \cdot \delta_{ij} = \sum_{i=1}^m \alpha(b_i) \cdot \beta_i(b_j) = \left( \sum_{i=1}^m \alpha(b_i) \cdot \beta_i \right) (b_j).$$

Nach Bemerkung 2.64 wird  $\alpha$  also dargestellt als Linearkombination

$$\alpha = \sum_{i=1}^m \alpha(b_i) \cdot \beta_i,$$

und die  $i$ -te Koordinate von  $\alpha$  ist gerade  $\alpha(b_i) \in R$ . Das Tupel  $(\beta_1, \dots, \beta_m)$  erzeugt also  $M^*$ . Es ist auch linear unabhängig, denn aus  $0 = r_1 \cdot \beta_1 + \dots + r_m \cdot \beta_m$  folgt für jedes  $j$ , dass

$$r_j = \sum_{i=1}^m r_i \cdot \delta_{ij} = \left( \sum_{i=1}^m r_i \cdot \beta_i \right) (b_j) = 0. \quad \square$$

## 2.5. Matrizen

Wir wollen jetzt lineare Abbildungen durch Matrizen beschreiben. Das ist zum Beispiel dann wichtig, wenn man numerische Berechnungen durchführen will (also Berechnungen mit „echten“ Zahlen, nicht abstrakte Überlegungen). Es gibt Bücher, die Matrizen als den Hauptgegenstand der linearen Algebra darstellen. Wir wollen Matrizen eher als nützliche Rechenschemata verstehen. Im Vordergrund des Interesses werden weiterhin lineare Abbildungen stehen.

2.67. DEFINITION. Es sei  $R$  ein Ring und  $m, n \in \mathbb{N}$ . Eine  $m \times n$ -Matrix über  $R$  ist eine Familie  $F = (f_{ij})_{i=1\dots m, j=1\dots n}$  in  $R$ , geschrieben

$$(1) \quad F = \begin{pmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & & \vdots \\ f_{m1} & \cdots & f_{mn} \end{pmatrix}.$$

Die Menge aller  $m \times n$ -Matrizen über  $R$  wird mit  $M_{m,n}(R)$  bezeichnet.

Wir definieren die *Matrixaddition*  $+: M_{m,n}(R) \times M_{m,n}(R) \rightarrow M_{m,n}(R)$  durch

$$(2) \quad F + G = (f_{ij} + g_{ij})_{i=1\dots m, j=1\dots n} \in M_{m,n}(R)$$

für alle  $G = (g_{ij})_{i=1\dots m, j=1\dots n} \in M_{m,n}(R)$ , und die *Matrizenmultiplikation*  $\cdot: M_{\ell,m}(R) \times M_{m,n}(R) \rightarrow M_{\ell,n}(R)$  mit  $\ell \in \mathbb{N}$  durch

$$(3) \quad H \cdot F = \left( \sum_{j=1}^m h_{ij} \cdot f_{jk} \right)_{i=1\dots \ell, k=1\dots n} \in M_{\ell,n}(R)$$

für alle  $H = (h_{ij})_{i=1\dots \ell, j=1\dots m} \in M_{\ell,m}(R)$ .

Wenn die Größe einer Matrix bekannt ist, schreiben wir  $(f_{ij})_{i,j} \in M_{m,n}(R)$  — daraus ergibt sich, dass  $1 \leq i \leq m$  und  $1 \leq j \leq n$ .

Die Matrixaddition erfolgt komponentenweise, genau wie in Beispiel 2.51. Zwei Matrizen kann man nur addieren, wenn sie die gleiche Anzahl von Zeilen und die gleiche Anzahl von Spalten haben.

Zwei Matrizen lassen sich multiplizieren, wenn die erste soviele Spalten hat wie die zweite Zeilen. Die Matrixmultiplikation lässt sich am besten am folgenden Schema verdeutlichen:

$$\begin{pmatrix} \cdot & \cdots & \cdot \\ \vdots & & \vdots \\ f_{i1} & \cdots & f_{in} \\ \vdots & & \vdots \\ \cdot & \cdots & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdots & g_{1k} & \cdots & \cdot \\ \vdots & & \vdots & & \vdots \\ \cdot & \cdots & g_{nk} & \cdots & \cdot \\ \vdots & & \vdots & & \vdots \\ \cdot & \cdots & \cdot & & \cdot \end{pmatrix} = \begin{pmatrix} \cdot & \cdots & \cdot \\ \vdots & & \vdots \\ \cdot & \cdots & f_{i1}g_{1k} + \cdots + f_{in}g_{nk} & \cdots & \cdot \\ \vdots & & \vdots & & \vdots \\ \cdot & \cdots & \cdot & & \cdot \end{pmatrix}$$

Hierbei steht  $F$  links,  $G$  oben, und das Produkt  $F \cdot G$  unten rechts.

2.68. BEMERKUNG. Wir betrachten die folgenden Spezialfälle.

- (1) Wenn  $m = 0$  oder  $n = 0$  ist, enthält  $M_{m,n}(R)$  nur ein Element, die leere Matrix  $(\ )$ .
- (2) Für  $m = 1 = n$  identifizieren wir  $M_{1,1}(R)$  mit  $R$ . Addition und Multiplikation von  $1 \times 1$ -Matrizen entsprechen genau der Addition und Multiplikation in  $R$ :

$$(r) + (s) = (r + s) \quad \text{und} \quad (r) \cdot (s) = (r \cdot s) .$$

- (3) Es sei  $n = 1$ , dann ist  $M_{m,1}(R) = R^m$  der „Raum der Spalten“ der Länge  $m$ , und Addition funktioniert genau wie in Beispiel 2.51. Wir können von rechts mit einer  $1 \times 1$ -Matrix aus (2) multiplizieren und erhalten die skalare Multiplikation aus Beispiel 2.51 (2), denn

$$\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \cdot (s) = \begin{pmatrix} r_1 \cdot s \\ \vdots \\ r_m \cdot s \end{pmatrix} = \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \cdot s .$$

Aus diesem Grund ist es sinnvoll, Spalten von rechts mit Skalaren zu multiplizieren.

- (4) Für  $m = 1$  ist  $M_{1,n}(R) = {}^nR$  der „Raum der Zeilen“ der Länge  $n$  aus Beispiel 2.52. Multiplikation mit einer  $1 \times 1$ -Matrix von links entspricht der Multiplikation mit einem Skalar.

Wir kommen jetzt zu allgemeinen Matrizen.

2.69. FOLGERUNG (aus Satz 2.55). *Es sei  $R$  ein Ring mit Eins und  $m, n \in \mathbb{N}$ . Dann existiert eine natürliche Bijektion*

$$(1) \quad \Phi: \text{Hom}_R(R^n, R^m) \rightarrow M_{m,n}(R) .$$

*Dabei steht das Bild des Standardbasisvektors  $e_j$  von  $R^n$  unter  $f: R^n \rightarrow R^m$  in der  $j$ -ten Spalte der Matrix  $(f_{ij})_{i,j} = \Phi(f)$ . Matrixmultiplikation  $\cdot: M_{m,n}(R) \times R^n \rightarrow R^m$  entspricht dem Anwenden einer linearen Abbildung, genauer*

$$(2) \quad f \left( \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \right) = \Phi(f) \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \in R^m .$$

*Die Matrixaddition entspricht der Addition linearer Abbildungen, für  $f, g \in \text{Hom}_R(R^n, R^m)$  gilt also*

$$(3) \quad \Phi(f + g) = \Phi(f) + \Phi(g) .$$

*Für  $\ell, m, n \in \mathbb{N}$  seien  $f: R^m \rightarrow R^\ell$  und  $g: R^n \rightarrow R^m$  rechts- $R$ -linear. Dann gilt*

$$(4) \quad \Phi(f \circ g) = \Phi(f) \cdot \Phi(g) \in M_{\ell,n}(R) ,$$

*die Matrixmultiplikation entspricht also der Verkettung linearer Abbildungen.*

BEWEIS. Der Modul  $R^n$  ist frei mit der Standardbasis  $\{e_1, \dots, e_n\}$ , siehe Beispiel 2.51. Nach Satz 2.55 (1) existiert eine bijektive Abbildung

$$\Phi: \text{Hom}_R(R^n, R^m) \longrightarrow (R^m)^n .$$

Wir identifizieren  $(R^m)^n$  mit  $M_{m,n}(R)$ , indem wir das  $j$ -te Element des Tupels in die  $j$ -te Spalte der Matrix  $(f_{ij})_{i,j}$  eintragen, und erhalten (1).

Sei umgekehrt  $F = (f_{ij})_{i,j} = \Phi(f) \in M_{m,n}(R)$  gegeben, das heißt, die  $j$ -te Spalte von  $F$  ist genau  $f(e_j)$ . Mit Folgerung 2.56 erhalten wir

$$f\left(\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}\right) = f\left(\sum_{j=1}^n e_j \cdot r_j\right) = \left(\sum_{j=1}^n f_{ij} \cdot r_j\right)_{i=1,\dots,m} = (f_{ij})_{i,j} \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}.$$

Die letzte Gleichung ist gerade die Definition 2.67 (3) der Matrixmultiplikation in dem Fall, dass der zweite Faktor  $(r_j)_{j=1,\dots,n}$  eine Spalte ist.

Zu (3) seien  $F = (f_{ij})_{i,j} = \Phi(f)$ ,  $G = (g_{ij})_{i,j} = \Phi(g) \in M_{m,n}(R)$ . Wir bestimmen  $\Phi(f+g)$  wie in (1), indem wir die Bilder der Vektoren  $e_k \in R^n$  berechnen. Nach Definition von  $f+g$  in Proposition 2.29 gilt

$$(f+g)(e_k) = f(e_k) + g(e_k) = \begin{pmatrix} f_{1k} \\ \vdots \\ f_{mk} \end{pmatrix} + \begin{pmatrix} g_{1k} \\ \vdots \\ g_{mk} \end{pmatrix} = \begin{pmatrix} f_{1k} + g_{1k} \\ \vdots \\ f_{mk} + g_{mk} \end{pmatrix} \in R^m,$$

und das ist genau die  $k$ -te Spalte der Matrix  $F+G = \Phi(f) + \Phi(g)$ .

Zu (4) sei  $F = (f_{ij})_{i,j} = \Phi(f) \in M_{\ell,m}(R)$  und  $G = (g_{jk})_{j,k} = \Phi(g) \in M_{m,n}(R)$ . Um die Matrix  $\Phi(f \circ g)$  zu erhalten, müssen wir wegen (1) die Bilder der Vektoren  $e_k \in R^n$  bestimmen. Nach (1) ist  $g(e_k)$  die  $k$ -te Spalte von  $G$ . Nach (2) gilt

$$f(g(e_k)) = \left(\sum_{j=1}^m f_{ij} \cdot g_{jk}\right)_{i=1,\dots,\ell} = \begin{pmatrix} f_{11} \cdot g_{1k} + \dots + f_{1m} \cdot g_{mk} \\ \vdots \\ f_{\ell 1} \cdot g_{1k} + \dots + f_{\ell m} \cdot g_{mk} \end{pmatrix} \in R^\ell,$$

Also hat  $\Phi(f \circ g)$  die gleiche  $k$ -te Spalte wie das Matrixprodukt  $\Phi(f) \cdot \Phi(g)$ . Daraus folgt unsere Behauptung.  $\square$

2.70. BEMERKUNG. Nach Folgerung 2.69 bietet es sich an, Matrizen  $F = (f_{ij})_{i,j} \in M_{m,n}(R)$  mit Hilfe von  $\Phi$  mit den zugehörigen Abbildungen  $f: R^n \rightarrow R^m$  zu identifizieren. Somit sei ab sofort

$$\text{Hom}_R(R^n, R^m) = M_{m,n}(R).$$

Man beachte: auf die Rechts- $R$ -Moduln  $R^n$  wirken Matrizen von links. Auf diese Weise kommt die skalare Multiplikation der Matrix nicht „in die Quere“. Homogenität (L2) folgt jetzt aus dem Assoziativgesetz (R1) der Multiplikation, siehe auch Beispiel 2.26 (3) und Folgerung 2.31 (3).

Bei Zeilen ist es genau spiegelbildlich: hier wirken Skalare von links wegen Bemerkung 2.68 (4) und Matrizen von rechts, es folgt also

$${}_R \text{Hom}({}^m R, {}^n R) = M_{m,n}(R).$$

Wenn man die Verkettung linearer Abbildungen als Matrixprodukt schreibt, dreht sich die Reihenfolge der Faktoren um. Aus diesem Grund ist es einfacher, mit Rechts- $R$ -Moduln zu arbeiten.

Tatsächlich kann man einige Fehler vermeiden, wenn man Skalare konsequent von rechts wirken lässt — selbst dann, wenn man über einem kommutativen Ring oder Körper arbeitet, bei dem es nach Bemerkung 2.23 eigentlich keinen Unterschied zwischen Rechts- und Linksmoduln gibt.

2.71. FOLGERUNG. *Die Matrixmultiplikation ist assoziativ.*

BEWEIS. Diese Behauptung könnte man beispielsweise mit Hilfe der Definition 2.67 (3) der Matrixmultiplikation mit etwas Aufwand nachrechnen.

Einfacher ist es, Matrizen  $F \in M_{\ell,m}(R) = \text{Hom}_R(R^m, R^\ell)$ ,  $G \in M_{m,n}(R) = \text{Hom}_R(R^n, R^m)$  und  $H \in M_{n,p}(R) = \text{Hom}_R(R^p, R^n)$  mit den entsprechenden linearen Abbildungen zu identifizieren. Da die Verkettung von Abbildungen assoziativ ist nach Bemerkung 2.4 (1), folgt aus Folgerung 2.69 (4), dass

$$F \cdot (G \cdot H) = F \circ (G \circ H) = (F \circ G) \circ H = (F \cdot G) \cdot H. \quad \square$$

2.72. DEFINITION. Es sei  $R$  ein Ring mit Eins. Eine  $m \times n$ -Matrix über  $R$  heißt *quadratisch*, wenn  $m = n$ . Der Raum der quadratischen  $n \times n$ -Matrizen über  $R$  wird mit  $M_n(R)$  bezeichnet. Die quadratische Matrix  $E_n = (\delta_{ij})_{i,j} \in M_n(R)$  heißt *Einheitsmatrix*. Eine quadratische Matrix  $F \in M_n(R)$  heißt *invertierbar*, wenn es eine Matrix  $G \in M_n(R)$  mit  $G \cdot F = E_n = F \cdot G$  gibt. In diesem Fall heißt  $G$  die zu  $F$  *inverse Matrix*; sie wird auch mit  $F^{-1}$  bezeichnet. Die Menge aller invertierbaren  $n \times n$ -Matrizen heißt *allgemeine lineare Gruppe* und wird mit  $GL(n, R)$  bezeichnet.

Wir übersetzen jetzt Folgerung 2.31 in die Matrizensprache.

2.73. FOLGERUNG (aus Folgerungen 2.31 und 2.69). *Es sei  $R$  ein Ring mit Eins und  $m, n \geq 1$ .*

- (1) *Die allgemeine lineare Gruppe  $(GL(n, R), \cdot)$  ist eine Gruppe, und es gilt  $GL(n, R) \cong \text{Aut}_R R^n$ .*
- (2) *Die quadratischen  $n \times n$ -Matrizen bilden einen Ring  $(M_n(R), +, \cdot)$  mit Eins  $E_n$ , den Matrixring, und es gilt  $M_n(R) \cong \text{End}_R R^n$ .*
- (3) *Der Raum der Spalten  $R^n$  wird durch Matrixmultiplikation zu einem unitären  $M_n(R)$ -Linksmodul.*
- (4) *Der Raum  $M_{m,n}(R)$  wird durch Matrixmultiplikation zu einem unitären Rechts- $M_n(R)$ -Modul und zu einem unitären Links- $M_m(R)$ -Modul.*

BEWEIS. Nach Folgerung 2.31 (1) und (2) bilden die Endomorphismen von  $R^n$  einen Ring  $(\text{End}_R(R^n), +, \circ)$  und die Automorphismen eine Gruppe  $(\text{Aut}_R(R^n), \circ)$ . Folgerung 2.69 liefert einen Ringisomorphismus

$$\Phi: (\text{End}_R(R^n), +, \circ) \xrightarrow{\cong} (M_n(R), +, \cdot).$$

Die Einheitsmatrix entspricht  $\text{id}_{R^n}$ , denn für alle  $m = (r_j)_j \in R^n$  gilt

$$E_n \cdot m = \left( \sum_{j=1}^n \delta_{ij} r_j \right)_i = (r_i)_i = m.$$

Sie ist die Eins in  $M_n(R)$  und das neutrale Element in  $GL(n, R)$ , es folgt (2).

Wegen Folgerung 2.69 (4) ist  $F$  genau dann als lineare Abbildung umkehrbar, also ein Automorphismus, wenn  $F$  als Matrix invertierbar ist. In diesem Fall wird die Umkehrabbildung von  $F$  genau durch die inverse Matrix  $F^{-1}$  beschrieben. Einschränken von  $\Phi$  liefert zu (2) den Gruppenisomorphismus

$$\Phi: (\text{Aut}_R(R^n), \circ) \xrightarrow{\cong} (GL(n, R), \cdot).$$

Die Punkte (3) und (4) folgen aus den entsprechenden Punkten in Folgerung 2.31 und Folgerung 2.69 (2) und (4).  $\square$

Wir merken uns:

- (1) Die Einheitsmatrix  $E_n$  entspricht der Identität des  $R^n$ .
- (2) Inverse Matrizen entsprechen Umkehrabbildungen. Aus Proposition 2.3 folgt, dass die inverse Matrix eindeutig bestimmt ist.

Wir haben in Definition 2.72 zur Invertierbarkeit von  $F$  verlangt, dass sowohl  $F \cdot G = E_n$  als auch  $G \cdot F = E_n$  gilt. In einer Gruppe reicht es, wenn  $F$  ein Linksinverses besitzt, also  $G \cdot F = E_n$  für ein  $G$  gilt. Aber  $G$  muss selbst invertierbar sein, um zu  $GL(n, R)$  zu gehören, also kommen wir um die Forderung  $F \cdot G = E_n$  nicht herum. Erst, wenn wir später mit quadratischen Matrizen über (Schief-) Körpern arbeiten, wird es reichen, nur  $G \cdot F = E_n$  zu verlangen.

Zum Schluss dieses Abschnitts wollen wir auch in freien Moduln mit festen Basen mit Koordinaten und Matrizen rechnen.

2.74. BEMERKUNG. Es sei  $M$  ein freier Rechts- $R$ -Modul mit Basis  $B = (b_1, \dots, b_m)$ . Für die Basisabbildung  $\Psi_B: R^m \rightarrow M$  aus Definition 2.63 schreiben wir kurz

$$\Psi_B \left( \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \right) = B \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} = \sum_{i=1}^m b_i \cdot r_i \in M.$$

Wir benutzen hier den gleichen Buchstaben für die Basisabbildung wie für die Basis, und tatsächlich verhält sich die Basisabbildung oben formal wie die Matrixmultiplikation der „Zeile“  $B$  aus Modulelementen mit der Spalte  $(r_i)_i \in R^m$ .

Nach Folgerung 2.62 ist die Basisabbildung invertierbar. Ihre Umkehrabbildung nennen wir die Koordinatenabbildung  $\beta = \Psi_B^{-1}$  und schreiben der Einfachheit halber

$$\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} = \beta(v) = {}_B v \in R^m.$$

Beide Abbildungen sind linear nach Proposition 2.29. Das bedeutet, dass wir mit den Koordinaten genauso rechnen dürfen wie mit den Modulelementen selbst. Es ist also egal, ob wir erst Vektoren addieren und mit Skalaren multiplizieren und dann Koordinaten bilden, oder erst Koordinaten der einzelnen Modulelemente nehmen und dann mit ihnen weiterrechnen.

2.75. FOLGERUNG. Es sei  $R$  ein Ring mit Eins, es sei  $M$  ein freier Rechts- $R$ -Modul mit Basis  $B = (b_1, \dots, b_m)$  und  $N$  ein freier Rechts- $R$ -Modul mit Basis  $C = (c_1, \dots, c_n)$ . Dann entspricht jeder linearen Abbildung  $f: N \rightarrow M$  genau eine Matrix  $A = {}_B f_C \in M_{m,n}(R)$ , die Abbildungsmatrix oder darstellende Matrix von  $f$  bezüglich  $B$  und  $C$ , so dass das folgende Diagramm kommutiert.

$$(1) \quad \begin{array}{ccc} N & \xrightarrow{f} & M \\ C \uparrow & & \uparrow B \\ R^n & \xrightarrow{A = {}_B f_C} & R^m \end{array}$$

Dabei stehen in der  $j$ -ten Spalte von  $A$  die  $B$ -Koordinaten  ${}_B(f(c_j))$  des Bildes des  $j$ -ten Basisvektors  $c_j$ . Für jedes Element  $v = C((r_i)_i) \in N$  hat das Bild  $f(v)$  also die Koordinaten  $A \cdot (r_i)_i$ , somit

$$(2) \quad {}_B(f(v)) = {}_B f_C \cdot C v .$$

Wir können uns das anhand des kommutativen Diagramms (1) oder der Formel (2) merken. Sei jetzt  $P$  ein weiterer  $R$ -Modul mit Basis  $D$  und  $G: P \rightarrow N$  linear, dann gilt völlig analog

$${}_B(f \circ g)_D = {}_B f_C \cdot C g_D .$$

Wichtig ist hier wie in (2), dass wir auf beiden Seiten der Matrixmultiplikation die gleiche Basis von  $N$  verwenden.

BEWEIS. Wir bezeichnen die Umkehrabbildung der Basisabbildung  $B$  mit  $\beta$  und setzen

$$A = \beta \circ f \circ C ,$$

dann kommutiert das Diagramm offensichtlich. Die restlichen Aussagen ergeben sich aus Folgerung 2.69.  $\square$

2.76. BEMERKUNG. Der Spezialfall  $M = N$  und  $f = \text{id}_M$  ist interessant. In diesem Fall erhalten wir das kommutative Diagramm

$$\begin{array}{ccc} & M & \\ C \nearrow & & \nwarrow B \\ R^n & \xrightarrow{A = {}_B \text{id}_C} & R^n . \end{array}$$

Multiplikation mit der Matrix  $A$  macht aus  $C$ -Koordinaten  $B$ -Koordinaten. Also besteht die  $j$ -te Spalte von  $A = (f_{ij})_{i,j}$  aus den  $B$ -Koordinaten des Vektor  $c_j$ , das heißt

$$c_j = \sum_{i=1}^n b_i \cdot f_{ij} .$$

Anders formuliert erhalten wir die Vektoren der Basis  $C$ , indem wir die „Zeile“  $B$  mit den Spalten von  $A$  multiplizieren. Aus diesem Grund nennt man die Matrix  $A$  auch *Basiswechselformel*. Die obigen Sachverhalte sind zwei Lesarten der „Gleichung“  $C = B \cdot A$ . Man beachte, dass die „Richtung“ des Basiswechsels für die Koordinaten („von  $C$  nach  $B$ “) und für die Basisvektoren („von  $B$

nach  $C$ ) genau umgekehrt ist. Um Fehler zu vermeiden, sollte man daher immer das obige kommutative Diagramm vor Augen haben.

2.77. PROPOSITION. Sei  $R$  Ring mit Eins, sei  $M$  ein  $R$ -Modul und  $m \in \mathbb{N}$ .

- (1) Es besteht eine Bijektion zwischen der Menge der angeordneten Basen  $(b_1, \dots, b_m)$  von  $M$  und den  $R$ -Modulisomorphismen  $R^m \rightarrow M$ .
- (2) Sei  $M$  frei mit Basis  $B = (b_1, \dots, b_m)$ . Dann besteht eine Bijektion zwischen der Menge der  $m$ -elementigen Basen von  $M$  und der allgemeinen linearen Gruppe  $GL(m, R)$ , die jeder Basis  $C = B \cdot A$  die Basiswechselmatrix  $A$  zuordnet.

Wenn  $M$  keine Basis der Länge  $m$  besitzt, sind insbesondere beide Mengen in (1) leer.

BEWEIS. Nach Satz 2.55 gibt es eine Bijektion  $\Psi: M^m \cong \text{Hom}_R(R^m, M)$ , und nach Folgerung 2.62 (3) ist  $B = (b_1, \dots, b_m)$  genau dann eine Basis, wenn  $\Psi_B: R^m \rightarrow M$  ein Isomorphismus ist. Es folgt (1).

Zu (2) sei  $\beta: M \rightarrow R^m$  die Koordinatenabbildung zu  $B$ , und  $C = (c_j)_j$  sei eine weitere Basis von  $M$ . Dann erhalten wir eine Basiswechselmatrix  $A = \beta \circ C$  wie in Bemerkung 2.76. Da  $\beta$  und  $C$  Isomorphismen sind, ist  $A$  invertierbar nach Folgerung 2.73 (1).

Sei umgekehrt  $A$  eine invertierbare Matrix, dann ist  $C = B \circ A: R^m \rightarrow M$  ein Isomorphismus, und  $(c_1, \dots, c_m) = \Phi(C)$  eine Basis nach (1). Die zugehörige Basiswechselmatrix ist  $\beta \circ (B \circ A) = A$ .  $\square$

2.78. BEMERKUNG. Wir können jetzt auch überlegen, wie sich die Abbildungsmatrix aus Folgerung 2.75 verhält, wenn wir eine der beiden Basen durch eine andere ersetzen. Wir betrachten dazu die kommutativen Diagramme

$$\begin{array}{ccc}
 & N & \xrightarrow{f} & M \\
 C \nearrow & & & \nwarrow D \\
 R^n & \xrightarrow{A=BfC} & R^m & \xrightarrow{D \text{ id}_B} & R^m
 \end{array}
 \quad \text{und} \quad
 \begin{array}{ccc}
 & N & \xrightarrow{f} & M \\
 E \nearrow & & & \nwarrow B \\
 R^n & \xrightarrow{C \text{ id}_E} & R^n & \xrightarrow{A=BfC} & R^m
 \end{array}$$

Hier ist  $D$  eine neue Basis von  $M$  und  $D \text{ id}_B \in GL(m, R)$  die zugehörige Basiswechselmatrix, und  $E$  ist eine Basis von  $N$  und  $C \text{ id}_E \in GL(n, R)$  die zugehörige Basiswechselmatrix. Es folgt

$$DfC = D \text{ id}_B \cdot BfC \quad \text{und} \quad BfE = BfC \cdot C \text{ id}_E .$$

Auch hier ist wieder wichtig, dass links und rechts vom Matrixmultiplikationszeichen „ $\cdot$ “ die gleiche Basis benutzt wird.

## 2.6. Unendliche Indexmengen

In den Abschnitten 2.4, 2.5 haben wir uns nur mit endlich erzeugten freien Moduln beschäftigt. Einige, aber nicht alle Resultate gelten analog für freie Moduln zu unendlichen Mengen  $I$ . Wichtig ist dabei aber, dass wir immer nur

endliche Linearkombinationen bilden, denn mehr gibt die Konstruktion in (2.1) nicht her. Für unendliche Summen bräuchten wir Methoden aus der Analysis.

2.79. DEFINITION. Es sei  $R$  ein Ring mit Eins und  $I$  eine beliebige Menge. Der *freie Modul zur Indexmenge  $I$*  ist definiert als

$$R^{(I)} = \{ (r_i)_{i \in I} \mid \text{die Menge } \{ i \in I \mid r_i \neq 0 \} \text{ ist endlich} \} \subset R^I .$$

Man überzeugt sich leicht, dass  $R^{(I)}$  ein Untermodul des Moduls  $R^I$  aus Bemerkung 2.50, also insbesondere selbst ein Rechts- $R$ -Modul ist. Elemente von  $R^I$  heißen auch Familien in  $R$  mit Indexmenge  $I$ . Elemente von  $R^{(I)}$  nennt man entsprechend *endliche Familien*. Die Notation  $R^{(I)}$  ist nicht Standard.

Es sei  $j \in I$ , dann sei  $e_j = (\delta_{ij})_{i \in I} \in R^{(I)}$  die Familie, deren einziger nichtverschwindender Eintrag eine Eins an der Stelle  $j$  ist. Wir nennen  $(e_i)_{i \in I}$  die *Standardbasis* von  $R^{(I)}$ . Eine lineare Abbildung  $f: R^{(I)} \rightarrow M$  bildet eine endliche Familie  $(r_i)_{i \in I}$  auf die endliche Linearkombination

$$f((r_i)_{i \in I}) = \sum_{i \in I} f(e_i) \cdot r_i$$

ab. Das sind alle Linearkombinationen der  $f(e_i) \in M$ , die wir mit unseren Mitteln bilden können. Zur Berechnung der rechten Seite wählen wir zunächst eine endliche Teilmenge  $J \subset I$ , so dass  $r_i = 0$  falls  $i \notin J$ . Dann suchen wir gemäß der Definitionen 1.26 und 1.31 eine Bijektion  $\alpha: \underline{n} \rightarrow J$  für  $n = \#J$  und setzen

$$\sum_{i \in I} a_i \cdot r_i = \sum_{k=0}^{n-1} a_{\alpha(k)} \cdot r_{\alpha(k)} ,$$

das ist jetzt eine endliche Summe wie in (2.1). Wenn wir  $J$  vergrößern, kommen nur Summanden der Form  $a_j \cdot 0 = 0$  hinzu. Und wenn wir die Abbildung  $\alpha$  abändern, vertauschen wir die Reihenfolge der Summation, was wegen der Kommutativität der Addition in  $M$  aber keine Rolle spielt. Also ist der Wert der Summe unabhängig von den Wahlen von  $J$  und  $\alpha$ .

2.80. BEISPIEL. Es sei  $R$  ein kommutativer Ring mit Eins. Ein *Polynom* über  $R$  in der Variablen  $X$  ist ein Ausdruck der Form

$$P(X) = \sum_{i=0}^n a_i X^i \quad \text{für ein } n \in \mathbb{N} \text{ und } a_0, \dots, a_n \in R .$$

Der Raum aller Polynome wird mit  $R[X]$  bezeichnet. Als  $R$ -Modul gilt  $R[X] \cong R^{(\mathbb{N})}$ . Wir dürfen  $r \in R$  in  $P$  einsetzen und erhalten eine lineare Abbildung

$$R[X] \longrightarrow R \quad \text{mit} \quad P(X) \longmapsto P(r) = \sum_{i=0}^n a_i \cdot r^i .$$

In Analogie dazu entspricht  $R^{\mathbb{N}}$  dem Raum  $R[[X]]$  der *formalen Potenzreihen* in  $X$  über dem Ring  $R$ . Wir dürfen mit formalen Potenzreihen zwar wie in jedem Modul rechnen, aber wir dürfen keine Elemente von  $R$  mehr einsetzen, da wir keine unendlichen Summen ausrechnen können. In der Analysis lernen

Sie den Raum der „konvergenten Potenzreihen“ über  $\mathbb{R}$  oder  $\mathbb{C}$  kennen, in die man zumindest hinreichend kleine Zahlen einsetzen darf.

In Analogie zu Satz 2.55 gilt folgendes Resultat.

2.81. SATZ (Universelle Eigenschaft des freien Moduls). *Es sei  $R$  ein Ring mit Eins,  $I$  eine Menge, und  $M$  ein Rechts- $R$ -Modul. Dann gibt es eine Bijektion*

$$\Phi: \text{Hom}_R(R^{(I)}, M) \xrightarrow{\cong} M^I \quad \text{mit} \quad f \longmapsto (f(e_i))_{i \in I}. \quad \square$$

Als Diagramm:

$$\begin{array}{ccc} I & \xrightarrow{e \cdot} & R^{(I)} \\ & \searrow A & \downarrow \exists! \Psi_A \\ & & M. \end{array}$$

BEWEIS. Für  $A = (a_i)_{i \in I} \in M^I$  und  $(r_i)_{i \in I} \in R^{(I)}$  setzen wir

$$\Psi_A((r_i)_{i \in I}) = \sum_{i \in I} a_i \cdot r_i \in M$$

wie in Satz 2.55 (2). Dann ist  $\Psi_A$  wieder eine lineare Abbildung, und es folgt

$$\Phi(\Psi_A) = (\Psi_A(e_i))_{i \in I} = \left( \sum_{j \in I} a_j \cdot \delta_{ji} \right)_{i \in I} = (a_i)_{i \in I} = A.$$

Umgekehrt sei  $f: R^{(I)} \rightarrow M$  rechts- $R$ -linear und  $A = \Phi(f) = (f(e_i))_{i \in I}$ , dann folgt

$$\Psi_A((r_i)_{i \in I}) = \sum_{i \in I} f(e_i) \cdot r_i = f\left(\sum_{i \in I} e_i \cdot r_i\right) = f((r_i)_{i \in I}). \quad \square$$

In Analogie zu den Definitionen 2.57 und 2.59 definieren wir das *Erzeugnis*  $\langle A \rangle \subset M$  einer Familie  $A$  in  $M$  als  $\text{im}(\Psi_A) \subset M$ . Wir nennen  $A$  *linear unabhängig*, wenn  $\ker(\Psi_A) = \{0\}$ . Eine Basis ist wieder ein linear unabhängiges Erzeugendensystem von  $M$ . Dann gelten Folgerung 2.62 und Bemerkung 2.64 analog. Insbesondere dürfen wir einen Modul mit Basis  $A \in M^I$  nach wie vor *frei* nennen.

2.82. BEMERKUNG. Anstelle von Bemerkung 2.65 gilt  $(R^{(I)})^* \cong {}^I R$ , dabei steht  ${}^I R$  für den  $R$ -Linksmodul auf der Menge  $\text{Abb}(I, R)$ . Sei  $(a_i)_{i \in I} \in {}^I R$  eine beliebige und  $(r_i)_{i \in I} \in R^{(I)}$  eine endliche Familie, dann erhalten wir eine endliche Summe

$$\Psi_{(a_i)_{i \in I}}((r_i)_{i \in I}) = \sum_{i \in I} a_i \cdot r_i \in R,$$

und aus Satz 2.81 folgt, dass jedes Element von  $(R^{(I)})^*$  von dieser Form ist.

Proposition 2.66 gilt daher für unendliche Indexmengen im Allgemeinen nicht mehr. Sei etwa  $B \in M^I$  eine Basis und  $\beta: M \rightarrow R^{(I)}$  die zugehörige

Koordinatenabbildung. Dann erzeugen die Komponentenfunktionen  $(\beta_i)_{i \in I} \in M^*$  den dualen Modul im Allgemeinen nicht. Dazu betrachte

$$\alpha = \sum_{j \in I} \beta_j \quad \text{mit} \quad \alpha \left( \sum_{i \in I} b_i \cdot r_i \right) = \sum_{i, j \in I} \underbrace{\beta_j(b_i)}_{=\delta_{ij}} \cdot r_i = \sum_{j \in I} r_j \in R.$$

Jede endliche Linearkombination der  $\beta_j$  würde auf einem  $b_i$  verschwinden, im Gegensatz zu  $\alpha$ , somit  $\alpha \notin \langle (\beta_j)_{j \in I} \rangle$ .

Wir können später beispielsweise zeigen, dass  $\mathbb{R}[X] \cong \mathbb{R}^{(\mathbb{N})}$  eine abzählbare Basis besitzt,  $R[[X]] \cong \mathbb{R}^{\mathbb{N}}$  jedoch nicht.

## 2.7. Zusammenfassung

Es folgt eine kurze Zwischenbilanz zum Ende des Abschnitts: In den Abschnitten 2.1 und 2.2 haben wir die Grundbegriffe kennengelernt: Gruppen, Ringe, (Schief-) Körper, Moduln beziehungsweise Vektorräume und lineare Abbildungen. Hier ging es vor allem um den Umgang mit Axiomen und Folgerungen daraus. Zunächst sind wir dabei in der Reihenfolge der Abschnitte 1.3 und 1.4 vorgegangen, später (über Ringen mit Eins) dann umgekehrt:

	abstrakt	konkret
Modul	$M$	$R^n$
Ring mit Eins	$\text{End}_R(M)$	$M_n(R)$
Gruppe	$\text{Aut}_R(M)$	$GL(n, R)$

Thema von Abschnitt 2.3 waren Unterräume, Quotienten und (direkte) Summen. Diese Konstruktionen schauen wir uns näher an, sobald wir mehr über Basen von Vektorräumen wissen. Wichtig sind hier ein Kriterium für Injektivität, die universelle Eigenschaft des Quotienten, und als Konsequenz der Homomorphiesatz 2.43.

In Abschnitt 2.4 haben wir Linearkombinationen, Erzeugendensysteme, lineare Unabhängigkeit und Basen betrachtet. Die universelle Eigenschaft 2.55 freier Moduln hat uns dann im Abschnitt 2.5 erlaubt, lineare Abbildungen zwischen freien Moduln als Matrizen zu schreiben. Matrizen ermöglichen zum einen konkrete Rechnungen mit linearen Abbildungen. Zum anderen verstehen wir mit ihrer Hilfe den Raum aller linearen Abbildungen besser. Die Existenz von Basen ist hierfür essentiell — im nächsten Kapitel lernen wir als erstes, dass endlich erzeugt Vektorräume stets Basen besitzen.

Abschnitt 2.6 gibt einen Ausblick auf unendlich erzeugte Moduln und Vektorräume. Er wird im Folgenden keine große Rolle mehr spielen.

## KAPITEL 3

# Vektorräume über Körpern und Schiefkörpern

In diesem Kapitel lernen wir typische Eigenschaften von Vektorräumen über (Schief-) Körpern kennen. Insbesondere hat jeder Vektorraum eine Basis, ist also als Modul frei. Außerdem lernen wir das Gauß-Verfahren zum Lösen linearer Gleichungssysteme kennen. Solche linearen Gleichungssysteme treten sowohl in der Praxis als auch in der Theorie häufig auf. Wir können das Gauß-Verfahren auch benutzen, um festzustellen, ob eine Matrix invertierbar ist, und gegebenenfalls die inverse Matrix zu bestimmen.

Alles, was in diesem Abschnitt passiert, beruht darauf, dass wir in einem Schiefkörper dividieren können. Auf der anderen Seite benötigen wir das Kommutativgesetz in diesem Abschnitt (noch) nicht. Für den Rest dieses Kapitels sei  $\mathbb{k}$  ein Schiefkörper, also zum Beispiel  $\mathbb{Q}$ ,  $\mathbb{R}$ ,  $\mathbb{C}$ ,  $\mathbb{H}$  oder  $\mathbb{Z}/p\mathbb{Z}$  für  $p$  prim. Wenn nichts anderes angegeben wird, seien alle  $\mathbb{k}$ -Vektorräume nach wie vor Rechts-Vektorräume, und alle Basen seien Tupel von Vektoren wie in Definition 2.59.

### 3.1. Basen

Wir haben spätestens im Abschnitt 2.5 gesehen, dass wir in freien Moduln weitaus leichter rechnen können als in beliebigen. Und wir haben auch gesehen, dass wir dadurch die Struktur dieser Moduln und der linearen Abbildungen gut beschreiben und verstehen können. Das soll diesen Abschnitt motivieren, in dem wir uns Gedanken über die Existenz von Basen machen wollen. Die beiden Sätze von Steinitz gehören zu den wichtigsten Ergebnissen dieser Vorlesung.

Wir erinnern uns an das Erzeugnis  $\langle \dots \rangle$  eines Tupels von Vektoren aus Definition 2.57 und an lineare Unabhängigkeit aus Definition 2.59. Wir arbeiten mit Rechts- $\mathbb{k}$ -Vektorräumen, aber analoge Aussagen gelten auch für Links- $\mathbb{k}$ -Vektorräume.

**3.1. LEMMA.** *Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum,  $(v_1, \dots, v_n)$  ein linear unabhängiges Tupel von Vektoren aus  $V$  und  $w \in V$ . Es gilt  $w \in \langle (v_1, \dots, v_n) \rangle$  genau dann, wenn das  $(n+1)$ -Tupel  $(v_1, \dots, v_n, w)$  linear abhängig ist.*

**BEWEIS.** Zu „ $\implies$ “ gelte  $w \in \langle (v_1, \dots, v_n) \rangle$ . Nach Definition 2.57 existieren  $a_1, \dots, a_n \in \mathbb{k}$ , so dass

$$w = \sum_{i=1}^n v_i \cdot a_i .$$

Nach Definition 2.59 ist  $(v_1, \dots, v_n, w)$  linear abhängig, da

$$0 = \sum_{i=1}^n v_i \cdot a_i + w \cdot \underbrace{(-1)}_{\neq 0}.$$

Zu „ $\Leftarrow$ “ sei  $(v_1, \dots, v_n, w)$  linear abhängig. Nach Definition 2.59 existieren  $a_1, \dots, a_n, b \in \mathbb{k}$ , die nicht alle 0 sind, so dass

$$0 = \sum_{i=1}^n v_i \cdot a_i + w \cdot b.$$

Es folgt  $b \neq 0$ . Denn andernfalls wäre ein  $a_i \neq 0$ , und wir hätten

$$0 = \sum_{i=1}^n v_i \cdot a_i$$

im Widerspruch zur linearen Unabhängigkeit von  $(v_1, \dots, v_n)$ . Es folgt

$$w = \sum_{i=1}^n v_i \cdot (-a_i b^{-1}) \in \langle (v_1, \dots, v_n) \rangle. \quad \square$$

**3.2. BEMERKUNG.** Wir haben im Beweis dividiert und können daher nicht erwarten, dass das Lemma für Moduln über beliebigen Ringen gilt. Als Gegenbeispiel betrachte  $\mathbb{Z}$  als  $\mathbb{Z}$ -Modul. Das Tupel  $(2)$  ist linear unabhängig, aber  $(2, 3)$  ist linear abhängig, da  $2 \cdot 3 - 3 \cdot 2 = 0$ . Dennoch ist 3 kein ganzzahliges Vielfaches von 2, also  $2 \notin \langle (3) \rangle$ . Die Voraussetzung, dass  $\mathbb{k}$  ein (Schief-) Körper ist, ist also notwendig. Für die meisten Aussagen in diesem und im nächsten Abschnitt finden wir Gegenbeispiele in Moduln über beliebigen Ringen.

**3.3. SATZ (Basisergänzungssatz von Steinitz).** *Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum. Es sei  $(v_1, \dots, v_r)$  ein Tupel linear unabhängiger Vektoren, und  $\{w_1, \dots, w_s\} \subset V$  sei ein endliches Erzeugendensystem. Dann gibt es  $n \geq r$  und Zahlen  $i(r+1), \dots, i(n) \in \{1, \dots, s\}$ , so dass das Tupel*

$$(1) \quad (v_1, \dots, v_r, w_{i(r+1)}, \dots, w_{i(n)})$$

eine Basis von  $V$  bildet.

**BEWEIS.** Wir beginnen mit  $n = r$  und dem  $n$ -Tupel  $B_0 = (v_1, \dots, v_r)$ . Wir setzen der Reihe nach  $j = 1, \dots, s$ , und nehmen an, dass  $w_i \in \langle B_{j-1} \rangle$  für alle  $i \leq j - 1$  gilt. Im Fall  $j = 1$  ist das trivialerweise wahr.

Wenn  $w_i \in \langle (v_1, \dots, v_r) \rangle$  gilt, setzen wir  $B_j = B_{j-1}$ . Es gilt  $w_i \in \langle B_j \rangle$  für alle  $i \leq j$ , und wir machen mit dem nächsten  $j$  weiter.

Andernfalls hängen wir  $w_j$  an das Tupel  $B_{j-1}$  an, erhöhen  $n$  um 1, und nennen dann das neue  $n$ -Tupel  $B_j$ . Nach Lemma 3.1 ist  $B_j$  linear unabhängig. Außerdem gilt wieder  $w_i \in \langle B_j \rangle$  für alle  $i \leq j$ . Wir merken uns, dass wir  $w_j$  angehängt haben, indem wir  $i(n) = j$  setzen.

Nach  $s$  Schritten haben wir ein linear unabhängiges  $n$ -Tupel  $B_s$  der Form (1) konstruiert. Da  $\langle B_s \rangle$  jeden der Erzeuger  $w_j$  enthält, folgt  $\langle B_s \rangle = V$ , also ist  $B_s$  eine Basis.  $\square$

Endlich erzeugte Vektorräume  $V$  haben also immer Basen. Als nächstes wollen wir zeigen, dass alle Basen von  $V$  gleich viele Elemente haben.

3.4. SATZ (Basisaustauschsatz von Steinitz). *Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum, es sei  $(v_1, \dots, v_r)$  ein linear unabhängiges Tupel, und  $(w_1, \dots, w_s)$  sei ein Erzeugendensystem. Dann gilt  $r \leq s$ .*

Der Name des Satzes ergibt sich aus dem „Austauschschritt“ im folgenden Beweis.

BEWEIS. Wir werden das Tupel  $(v_1, \dots, v_r)$  in mehreren Schritten zu einer Basis aus Einträgen des Tupels  $(w_1, \dots, w_s)$  umbauen. Dabei wird unser Tupel in keinem Schritt kürzer. Da das resultierende Tupel insbesondere linear unabhängig ist, kann es keinen Eintrag doppelt enthalten, also folgt die Behauptung.

Im ersten Schritt ergänzen wir  $(v_1, \dots, v_r)$  gemäß des Basisergänzungssatzes 3.3 mit Einträgen des Tupels  $(w_1, \dots, w_s)$  zu einer Basis  $B_0$ . Dann hat  $B_0$  mindestens  $r$  Einträge.

Wir setzen der Reihe nach  $i = 1, \dots, r$  und starten mit einer Basis  $B_{i-1}$ , die mit  $v_i, \dots, v_r$  anfängt. Wir entfernen  $v_i$  und erhalten ein linear unabhängiges Tupel  $B'_i$ . Da  $B_{i-1}$  linear unabhängig ist, folgt  $v_i \notin \langle B'_i \rangle$  aus Lemma 3.1, insbesondere ist  $B'_i$  kein Erzeugendensystem mehr. Also müssen wir mindestens einen Eintrag aus dem Tupel  $(w_1, \dots, w_s)$  hinzufügen, um mit Satz 3.3 eine neue Basis  $B_i$  zu bekommen. Sie ist somit nicht kürzer als  $B_{i-1}$ .

Am Ende besteht  $B_r$  aus mindestens  $r$  verschiedenen Einträgen des Tupels  $(w_1, \dots, w_s)$ . Es folgt  $r \leq s$ .  $\square$

3.5. FOLGERUNG. *Es sei  $V$  ein endlich erzeugter  $\mathbb{k}$ -Vektorraum.*

- (1) *Dann existiert  $n \in \mathbb{N}$  und eine Basis  $B = (v_1, \dots, v_n)$  von  $V$ .*
- (2) *Jede andere Basis von  $V$  hat ebenfalls  $n$  Elemente.*
- (3) *Jedes  $n$ -elementige Tupel linear unabhängiger Vektoren in  $V$  ist eine Basis von  $V$ .*
- (4) *Jedes  $n$ -elementige Erzeugendensystem von  $V$  ist eine Basis von  $V$ .*

BEWEIS. Es sei  $(w_1, \dots, w_s)$  ein Erzeugendensystem von  $V$ . Der Basisergänzungssatz 3.3 liefert uns, ausgehend vom leeren Tupel  $()$  mit  $r = 0$ , eine Basis  $B = (v_1, \dots, v_n)$  von  $V$ , deren Länge  $n \leq s$  endlich ist, also gilt (1).

Zu (2) sei  $C$  eine beliebige Basis von  $V$ , möglicherweise sogar mit unendlicher Indexmenge  $I$ . Wir können jeden Vektor  $w_i$  als Linearkombination von Vektoren aus  $C$  darstellen, dazu benötigen wir aber nur endlich viele. Da  $(w_1, \dots, w_s)$  Erzeugendensystem ist, ist jeder Vektor  $v \in V$  als Linearkombination der  $w_i$  darstellbar. In diese Linearkombination setzen wir die obigen Darstellungen der  $w_i$  ein. Insgesamt erhalten wir  $v$  als Linearkombination der Vektoren aus  $C$ , wobei wir nur endlich viele Vektoren der Familie  $C$  benötigen, nämlich nur die, die in einer der Darstellungen der  $w_i$  mit Koeffizient  $\neq 0$

vorkommen. Es sei  $I_0$  die entsprechende endliche Teilmenge der Indexmenge  $I$  von  $C$  und  $C_0: I_0 \rightarrow V$  die zugehörige Teilfamilie. Alle anderen Vektoren  $c_i$  mit  $i \in I \setminus I_0$  lassen sich als Linearkombination der  $w_i$  darstellen, also auch als Linearkombination der Vektoren aus  $C_0$ . Wäre  $C \neq C_0$ , so wäre  $C$  insbesondere linear abhängig. Wir schließen also, dass die Basis  $C$  endlich ist.

Jetzt können wir  $C$  anordnen zu  $(u_1, \dots, u_s)$ . Indem wir  $B$  als linear unabhängiges Tupel und  $C$  als Erzeugendensystem auffassen, erhalten wir  $n \leq s$  aus dem Basisaustauschsatz 3.4. Wir können die Rolle der beiden Basen auch vertauschen, und erhalten  $s \leq n$ . Also haben  $B$  und  $C$  gleich viele Elemente.

Zu (3) sei  $(w_1, \dots, w_n)$  linear unabhängig, dann können wir mit Satz 3.3 zu einer Basis von  $V$  ergänzen, die nach (2) wieder Länge  $n$  hätte. Also ist  $(w_1, \dots, w_n)$  bereits eine Basis.

Zu (4) sei analog  $(w_1, \dots, w_n)$  ein Erzeugendensystem. Eine Teilmenge davon bildet nach Satz 3.3 eine Basis, die aber wieder Länge  $n$  hätte. Also ist  $(w_1, \dots, w_n)$  bereits eine Basis.  $\square$

3.6. BEMERKUNG. Man kann analoge Sätze auch für beliebige, nicht notwendig endlich erzeugte  $\mathbb{k}$ -Vektorräume beweisen. Dazu braucht man allerdings ein weiteres Axiom für die zugrundeliegende Mengenlehre, das *Auswahlaxiom*.

### 3.2. Dimension und Rang

Wir benutzen die Basissätze, um ein paar interessante Aussagen über Vektorräume und ihre Unterräume, Quotienten und über lineare Abbildungen zu beweisen.

Aufgrund von Folgerung 3.5 ist die folgende Definition sinnvoll.

3.7. DEFINITION. Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum. Wenn  $V$  endlich erzeugt ist, ist die *Dimension*  $\dim V$  von  $V$  die Länge  $n$  einer Basis  $(v_1, \dots, v_n)$  von  $V$ , und wir nennen  $V$  *endlichdimensional*. Wenn  $V$  keine Basis endlicher Länge besitzt, heißt  $V$  *unendlichdimensional*.

Die Begriffe „endlichdimensional“ und „endlich erzeugt“ für Vektorräume sind nach Folgerung 3.5 äquivalent, und wir schreiben dafür auch „ $\dim V < \infty$ “. Wie in Bemerkung 1.32 (3) führen wir die Schreibweise „ $\dim V = \infty$ “ nicht ein, da nicht alle unendlichen Basen die gleiche Mächtigkeit haben.

3.8. FOLGERUNG. *Zwei endlichdimensionale  $\mathbb{k}$ -Vektorräume  $V$  und  $W$  sind genau dann isomorph, wenn  $\dim V = \dim W$ .*

Man sagt auch, endlichdimensionale  $\mathbb{k}$ -Vektorräume werden durch ihre Dimension *klassifiziert*. Analog werden endliche Mengen durch ihre Mächtigkeit klassifiziert, siehe Definitionen 1.26 und 1.31.

BEWEIS. Zu „ $\implies$ “ sei  $F: V \rightarrow W$  ein Isomorphismus. Wir wählen eine Basis  $C = (c_1, \dots, c_n)$  von  $V$ , wobei  $n = \dim V$ , und identifizieren wieder  $C$  mit

der zugehörigen Basisabbildung. Dann ist die Abbildung  $B = F \circ C: \mathbb{k}^n \rightarrow W$  ein Isomorphismus, und das Diagramm

$$(3.1) \quad \begin{array}{ccc} V & \xrightarrow{F} & W \\ C \uparrow & & \uparrow B \\ \mathbb{k}^n & \xrightarrow[\text{E}_n]{\text{id}_{\mathbb{k}^n}} & \mathbb{k}^n \end{array}$$

kommutiert. Wegen Proposition 2.77 (1) bilden  $b_1 = B(e_1), \dots, b_n = B(e_n)$  eine Basis von  $W$ , so dass insbesondere  $\dim W = n = \dim V$ .

Zu „ $\Leftarrow$ “ sei  $n = \dim V = \dim W$ . Wir wählen Basen von  $V$  und  $W$  mit Basisabbildungen  $B: \mathbb{k}^n \rightarrow W$  und  $C: \mathbb{k}^n \rightarrow V$ . Nach Folgerung 2.62 (3) sind Basisabbildungen Isomorphismen. Wir erhalten also einen Isomorphismus  $F = B \circ C^{-1}: V \rightarrow W$ , so dass das obige Diagramm wieder kommutiert.  $\square$

Wir erinnern uns an die Begriffe „direkte Summe“ und „komplementärer Unterraum“ aus Definition 2.45.

**3.9. PROPOSITION.** *Sei  $V$  ein endlichdimensionaler  $\mathbb{k}$ -Vektorraum und  $U \subset V$  ein Unterraum. Dann besitzt  $U$  ein Komplement  $W \subset V$ , und es gilt die Dimensionsformel*

$$\dim V = \dim U + \dim W .$$

**BEWEIS.** Jedes linear unabhängige Tupel von Vektoren in  $U$  ist in  $V$  ebenfalls linear unabhängig. Nach dem Basisaustauschsatz 3.4 kann solch ein Tupel also höchstens  $\dim V$  viele Elemente haben, insbesondere ist es endlich. Aus den Übungen wissen wir auch, dass ein maximal linear unabhängiges Tupel in  $U$  eine Basis von  $U$  bildet. Also finden wir eine Basis  $B = (v_1, \dots, v_r)$  der Länge  $r = \dim U \leq n = \dim V$  von  $U$ . Wir ergänzen  $B$  zu einer Basis  $(v_1, \dots, v_n)$  von  $V$  mit dem Basisergänzungssatz 3.3.

Es sei  $W = \langle v_{r+1}, \dots, v_n \rangle$ , dann ist das Tupel  $(v_{r+1}, \dots, v_n)$  eine Basis von  $W$ , denn es erzeugt  $W$  und ist als Teil einer Basis von  $V$  auch linear unabhängig. Insbesondere gilt

$$\dim V = \dim U + \dim W .$$

Außerdem gilt

$$U + W = \langle v_1, \dots, v_n \rangle = V .$$

Sei nun  $v \in U \cap W$ . Dann existieren  $k_1, \dots, k_r \in \mathbb{k}$  und  $\ell_{r+1}, \dots, \ell_n \in \mathbb{k}$  mit

$$\sum_{i=1}^r v_i k_i = v = \sum_{j=r+1}^n v_j \ell_j .$$

Beides sind Darstellung als Linearkombination der Basis  $(v_1, \dots, v_n)$ . Nach Folgerung 2.62 (2) sind die Koordinaten von  $v$  eindeutig, also gilt  $k_1 = \dots = k_r = 0 = \ell_{r+1} = \dots = \ell_n$ . Insbesondere folgt  $U \cap W = \{0\}$ , also  $V = U \oplus W$ .  $\square$

3.10. FOLGERUNG. *Es seien  $U$  und  $W$  zwei Unterräume eines endlichdimensionalen  $\mathbb{k}$ -Vektorraums  $V$ . Dann sind äquivalent*

- (1)  $V = U \oplus W$ ,
- (2)  $V = U + W$  und  $\dim U + \dim W \leq \dim V$ ,
- (3)  $U \cap W = \{0\}$  und  $\dim U + \dim W \geq \dim V$ .

BEWEIS. Die Richtungen „(1)  $\implies$  (2)“ und „(1)  $\implies$  (3)“ folgen sofort aus der Definition 2.45 der direkten Summe und Proposition 3.9.

In den Übungen beweisen Sie die Dimensionsformel für Summen

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W).$$

Aus (2) schließen wir, dass

$$0 \leq \dim(U \cap W) = \dim U + \dim W - \dim V \leq 0,$$

aber also wird  $U \cap W$  von einer Basis der Länge 0 erzeugt, das heißt  $U \cap W = \{0\}$ , und es folgt (1).

Aus (3) schließen wir, dass

$$\dim V \geq \dim(U + W) = \dim U + \dim W - \dim\{0\} \geq \dim V,$$

also hat  $U + W$  eine Basis der Länge  $\dim V$ . Wäre  $U + W$  eine echte Teilmenge von  $V$ , so könnten wir zu einer Basis von  $V$  der Länge  $\geq \dim V + 1$  ergänzen, im Widerspruch zu Folgerung 3.5. Also gilt  $U + W = V$ , und wieder folgt (1).  $\square$

3.11. FOLGERUNG. *Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum von endlicher Dimension und  $U \subset V$  ein Unterraum. Dann gilt*

$$\dim(V/U) = \dim V - \dim U.$$

BEWEIS. Wir wählen einen zu  $U$  komplementären Unterraum  $W \subset V$ . Aus den Propositionen 2.47 und 3.9 folgt

$$\dim(V/U) = \dim W = \dim V - \dim U. \quad \square$$

3.12. BEMERKUNG. Wenn  $V$  unendlichdimensional ist, können wir mit dem allgemeineren Basisergänzungssatz aus Bemerkung 3.6 immer noch zu jedem Unterraum einen komplementären Unterraum konstruieren. Da man aber unendliche Dimensionen nicht subtrahieren kann, ist die Dimensionsformel in Proposition 3.9 nicht geeignet, um die Dimension des Komplements zu bestimmen. Als Beispiel betrachten wir den Raum  $V = \mathbb{R}^{(\mathbb{N})}$  der endlichen reellwertigen Folgen mit der Basis  $(e_j)_{j \in \mathbb{N}}$ , wobei wieder  $e_j = (\delta_{ij})_{i \in \mathbb{N}}$ . Wir betrachten zwei unendlichdimensionale Unterräume

$$U = \langle e_r, e_{r+1}, e_{r+2}, \dots \rangle \quad \text{und} \quad W = \langle e_0, e_2, e_4, \dots \rangle.$$

Beide sind als Vektorräume isomorph, denn wir können einen Isomorphismus  $F: U \rightarrow W$  angeben mit  $F(e_{r+j}) = e_{2j}$  für alle  $j \in \mathbb{N}$ . Aber  $U$  besitzt ein endlichdimensionales Komplement  $\langle e_0, \dots, e_{r-1} \rangle$ , während  $W$  ein unendlichdimensionales Komplement  $\langle e_1, e_3, e_5, \dots \rangle$  hat. Und da nach Proposition 2.47 alle Komplemente von  $U$  zu  $V/U$  isomorph sind, und alle Komplemente von  $W$

zu  $V/W$ , können wir die Dimension des Komplementes nun nicht mehr aus der Dimension der Räume selbst ablesen.

Übrigens hat auch  $\mathbb{R}^{(\mathbb{N})}$  selbst im Raum  $\mathbb{R}^{\mathbb{N}}$  aller reellwertigen Folgen ein Komplement. Da wir das aber wieder mit Hilfe des Zornschen Lemma beweisen müssen, können wir das Komplement nicht explizit angeben.

Mit den gleichen Methoden wie oben können wir auch lineare Abbildungen studieren. Unter einer *Blockmatrix* verstehen wir eine Matrix, die durch das Neben- und Untereinanderschreiben von Matrizen passender Größe gebildet wird. Seien etwa  $A \in M_{p,r}(\mathbb{k})$ ,  $B \in M_{p,s}(\mathbb{k})$ ,  $C \in M_{q,r}(\mathbb{k})$  und  $D \in M_{q,s}(\mathbb{k})$ , dann ist

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1r} & b_{11} & \dots & b_{1s} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p1} & \dots & a_{pr} & b_{p1} & \dots & b_{ps} \\ c_{11} & \dots & c_{1r} & d_{11} & \dots & d_{1s} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{q1} & \dots & c_{qr} & d_{q1} & \dots & d_{qs} \end{pmatrix} \in M_{p+q,r+s}(\mathbb{k}).$$

**3.13. SATZ (Rangsatz).** *Es seien  $V$  und  $W$  endlich-dimensionale  $\mathbb{k}$ -Vektorräume, und es sei  $F: V \rightarrow W$  linear. Dann existieren Basen  $B$  von  $W$  und  $C$  von  $V$ , so dass die Abbildungsmatrix  $A$  von  $F$  bezüglich dieser Basen die Normalform*

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

als Blockmatrix hat, wobei  $r = \dim \operatorname{im} F$ . Insbesondere gilt die Dimensionsformel

$$\dim \ker F + \dim \operatorname{im} F = \dim V.$$

**BEWEIS.** Es sei  $n = \dim V$  und  $r = n - \dim \ker F$ . Wie im Beweis von Proposition 3.9 wählen wir zunächst eine Basis  $(c_{r+1}, \dots, c_n)$  von  $\ker F$  und ergänzen dann zu einer Basis  $(c_1, \dots, c_n)$  von  $V$ . Dann ist  $U = \langle c_1, \dots, c_r \rangle$  ein Komplement von  $\ker F$  in  $V$ . Nach Proposition 2.47 (3) und dem Homomorphiesatz 2.43 erhalten wir einen Isomorphismus

$$\begin{array}{ccc} U & \xrightarrow{\cong} & V/\ker F & \xrightarrow{\cong} & \operatorname{im} F \\ & & \searrow & \nearrow & \\ & & & & \operatorname{im} F \end{array}$$

$F|_U$

Somit induziert die Basis  $(c_1, \dots, c_r)$  von  $U$  eine Basis  $(b_1, \dots, b_r)$  von  $\operatorname{im} F$  mit  $b_i = F(c_i)$  für alle  $1 \leq i \leq r$ . Schließlich ergänzen wir zu einer Basis  $(b_1, \dots, b_m)$  von  $W$ . Für die Abbildung  $F$  gilt also

$$F(c_j) = \begin{cases} b_j & \text{falls } j \leq r, \text{ und} \\ 0 & \text{falls } j > r. \end{cases}$$

Daraus ergibt sich die angegebene Form der Abbildungsmatrix. Außerdem folgt

$$\dim V = \dim \ker F + \dim U = \dim \ker F + \dim \operatorname{im} F. \quad \square$$

3.14. DEFINITION. Es sei  $A = (a_{ij})_{i,j} \in M_{m,n}(R)$  eine Matrix, dann definieren wir die zu  $A$  *transponierte Matrix*  $A^t \in M_{n,m}(R)$  durch  $A^t = (a_{ij})_{ji}$ .

Transponieren macht zum Beispiel aus Zeilen Spalten und umgekehrt. In Büchern wird häufig  $(r_1, \dots, r_n)^t$  für die Spalte  $\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$  geschrieben.

3.15. DEFINITION. Es sei  $F: V \rightarrow W$  linear, dann definieren wir den *Rang* von  $F$  durch  $\text{rg } F = \dim \text{im } F$ , falls  $\text{im } F$  endlichdimensional ist, ansonsten nennen wir  $F$  *von unendlichem Rang*.

Es sei  $A \in M_{m,n}(\mathbb{k})$  eine Matrix mit den Spalten  $a_1, \dots, a_n \in \mathbb{k}^m$  und den Zeilen  $\alpha_1, \dots, \alpha_m \in {}^m\mathbb{k}$ , dann definieren wir den *Spaltenrang* von  $A$  durch  $\text{rg}_S A = \dim \langle a_1, \dots, a_n \rangle$ . Analog definieren wir den *Zeilenrang* von  $A$  durch  $\text{rg}_Z A = \dim \langle \alpha_1, \dots, \alpha_m \rangle$ .

In der Definition des Zeilenrangs fassen wir das Erzeugnis der Zeilen als Links- $\mathbb{k}$ -Vektorraum auf.

3.16. BEISPIEL. Wir schauen uns Spalten- und Zeilenrang an Beispielen an.

- (1) Die „Telefonmatrix“  $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in M_3(\mathbb{Q})$  hat Rang 2, denn je zwei Zeilen oder Spalten sind linear unabhängig, aber

$$\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} 2 - \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} - \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} = 0 \quad \text{und} \quad 2(4, 5, 6) - (1, 2, 3) - (7, 8, 9) = 0.$$

- (2) Die Matrix  $\begin{pmatrix} 1 & j \\ i & k \end{pmatrix} \in M_2(\mathbb{H})$  hat Rang 1, denn

$$\begin{pmatrix} 1 \\ i \end{pmatrix} \cdot j = \begin{pmatrix} j \\ k \end{pmatrix} \quad \text{und} \quad i \cdot (1, j) = (i, k).$$

- (3) Die Transponierte der obigen Matrix hat Rang 2, denn sie ist invertierbar mit inverser Matrix  $\begin{pmatrix} 1 & i \\ j & k \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & -j/2 \\ -i/2 & -k/2 \end{pmatrix}$ .

Manchmal heißt auch die folgende Proposition „Rangatz“.

3.17. PROPOSITION. *Es sei  $A \in M_{m,n}(\mathbb{k})$ .*

- (1) *Spalten- und Zeilenrang von  $A$  ändert sich nicht, wenn man von links oder rechts mit einer invertierbaren Matrix multipliziert.*
- (2) *Es gilt  $\text{rg}_S A = \text{rg } A = \text{rg}_Z A$ .*
- (3) *Es sei  $A = {}_B F_C$  Matrixdarstellung einer linearen Abbildung  $F: V \rightarrow W$  bezüglich Basen  $B$  von  $W$  und  $C$  von  $V$ , dann gilt  $\text{rg } F = \text{rg } A$ .*

BEWEIS. Sei  $A$  eine Matrix. Wie in Bemerkung 2.70 bezeichne  $L_A: \mathbb{k}^n \rightarrow \mathbb{k}^m$  die (rechts- $\mathbb{k}$ -lineare) Multiplikation mit  $A$  von Links, und  $R_A: {}^m\mathbb{k} \rightarrow {}^m\mathbb{k}$  die (links- $\mathbb{k}$ -lineare) Multiplikation mit  $A$  von Rechts, also

$$L_A(v) = A \cdot v \quad \text{und} \quad R_A(\omega) = \omega \cdot A.$$

Es seien wieder  $a_1, \dots, a_n \in \mathbb{k}^m$  die Spalten von  $A$ . Dann gilt

$$\langle a_1, \dots, a_n \rangle = \langle L_A(e_1), \dots, L_A(e_n) \rangle = \text{im } L_A,$$

also gilt  $\operatorname{rg}_S A = \operatorname{rg} L_A$ . Für  $\operatorname{rg}_Z A$  betrachten wir analog die Basis  $\varepsilon_1, \dots, \varepsilon_m$  von  ${}^m\mathbb{k}$  aus Beispiel 2.52 und erhalten

$$\langle \alpha_1, \dots, \alpha_m \rangle = \langle \varepsilon_1 \cdot A, \dots, \varepsilon_m \cdot A \rangle = \langle R_A(\varepsilon_1), \dots, R_A(\varepsilon_m) \rangle = \operatorname{im} R_A,$$

also gilt  $\operatorname{rg}_Z A = \operatorname{rg} R_A$ .

Es sei zunächst  $B \in GL(m, \mathbb{k})$  eine invertierbare Matrix. Die zugehörige lineare Abbildung  $L_B: \mathbb{k}^m \rightarrow \mathbb{k}^m$  ist ein Automorphismus, insbesondere bijektiv, siehe Folgerung 2.73 (1). Es gilt

$$\operatorname{im}(L_{BA}) = \{ B \cdot A \cdot v \mid v \in \mathbb{k}^n \} = \operatorname{im}(B|_{\operatorname{im} L_A}).$$

Die Abbildung  $L_B|_{\operatorname{im} L_A}: \operatorname{im} L_A \rightarrow \operatorname{im} L_{BA}$  ist sicherlich immer noch injektiv und linear. Sie ist auch surjektiv, da wir das Bild entsprechend eingeschränkt haben. Somit sind  $\operatorname{im} L_A$  und  $\operatorname{im} L_{BA}$  isomorph, und es folgt

$$\operatorname{rg}_S(BA) = \dim \operatorname{im} L_{BA} = \dim \operatorname{im} L_A = \operatorname{rg}_S A.$$

Andererseits bilden  $\varepsilon_1 \cdot B, \dots, \varepsilon_m \cdot B$  nach wie vor eine Basis von  ${}^m\mathbb{k}$ , da auch  $R_B$  ein Isomorphismus ist. Die Abbildung  $R_{BA}: {}^m\mathbb{k} \rightarrow {}^m\mathbb{k}$  mit  $\omega \mapsto \omega \cdot BA$  hat also das gleiche Bild wie  $R_A$ , und es gilt

$$\operatorname{rg}_Z(BA) = \dim \operatorname{im}(R_{BA}) = \dim \operatorname{im}(R_A) = \operatorname{rg}_Z A.$$

Also sind sowohl Spalten- als auch Zeilenrang unter Multiplikation mit einer invertierbaren Matrix von links invariant.

Sei jetzt  $C \in GL(n, \mathbb{k})$  invertierbar. Ähnlich wie oben gilt jetzt

$$\operatorname{im} L_{AC} = \operatorname{im} L_A \quad \text{und} \quad \operatorname{im} R_A \cong \operatorname{im}(R_C|_{\operatorname{im} R_A}) = \operatorname{im} R_{AC},$$

und es folgt  $\operatorname{rg}_S(AC) = \operatorname{rg}_S A$  und  $\operatorname{rg}_Z(AC) = \operatorname{rg}_Z A$ , also gilt (1).

Wir wählen jetzt Basen  $B$  von  $\mathbb{k}^m$  und  $C$  von  $\mathbb{k}^n$  wie in Satz 3.13. Es folgt (2), da

$$\begin{aligned} \operatorname{rg}_S(A) &= \operatorname{rg}_S(B^{-1} \cdot A \cdot C) = \operatorname{rg}_S \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r \\ &= \operatorname{rg}_Z \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = \operatorname{rg}_Z(B^{-1} \cdot A \cdot C) = \operatorname{rg}_Z A. \end{aligned}$$

Matrixdarstellungen in (3) zu verschiedenen Basen unterscheiden sich um Multiplikation mit einer Basiswechselmatrix von links oder von rechts. Nach (1) ist es also egal, wie wir  $B$  und  $C$  wählen, daher nehmen wir Basen wie im Rangsatz (3.13). Es folgt  $\operatorname{im} F = \langle b_1, \dots, b_r \rangle$ , also  $\operatorname{rg} F = r = \operatorname{rg} A$ .  $\square$

**3.18. FOLGERUNG.** *Es sei  $f: V \rightarrow W$  eine lineare Abbildung zwischen endlich-dimensionalen  $\mathbb{k}$ -Vektorräumen.*

- (1) *Es gilt  $\operatorname{rg} f = \dim W$  genau dann, wenn  $f$  surjektiv ist.*
- (2) *Es gilt  $\operatorname{rg} f = \dim V$  genau dann, wenn  $f$  injektiv ist,*
- (3) *Es gilt  $\operatorname{rg} f = \dim V = \dim W$  genau dann, wenn  $f$  bijektiv ist.*

*Analoge Aussagen gelten auch für Matrizen.*

BEWEIS. Es gelte  $\operatorname{rg} f = \dim W$ , dann ist  $\operatorname{im} f \subset W$  ein Unterraum voller Dimension. Nach Folgerung 3.5 (3) ist jede Basis von  $\operatorname{im} f$  bereits eine Basis von  $W$ , somit gilt  $W = \operatorname{im} f$ , und  $f$  ist surjektiv. Die Gegenrichtung ist klar.

Sei jetzt  $\operatorname{rg} f = \dim V$ , dann folgt aus der Dimensionsformel im Rangsatz 3.13, dass  $\dim \ker f = \dim V - \operatorname{rg} f = 0$ . Es folgt  $\ker f = 0$ , also ist  $f$  injektiv nach Proposition 2.37 (1). Umgekehrt folgt  $\operatorname{rg} f = \dim V$  aus  $\ker f = \{0\}$ .

Aussage (3) folgt aus (1) und (2). □

### 3.3. Lineare Gleichungssysteme

3.19. DEFINITION. Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum. Eine Teilmenge  $A \subset V$  heißt *affiner Unterraum* von  $V$ , wenn es einen Untervektorraum  $U \subset V$  und ein Element  $a_0 \in A$  gibt, so dass

$$A = a_0 + U = \{ a_0 + u \mid u \in U \} .$$

Ein affiner Unterraum  $A = a + U$  heißt *endlichdimensional* mit  $\dim A = \dim U$ , wenn  $U$  endlichdimensional ist, sonst *unendlichdimensional*. Seien  $U, W \subset V$  Untervektorräume, dann heißen zwei affine Unterräume  $a+U$  und  $b+W$  *parallel*, wenn  $U = W$ .

Man beachte, dass in manchen Büchern auch die leere Menge  $\emptyset$  als affiner Unterraum der Dimension  $\dim \emptyset = -\infty$  betrachtet wird. Wir wollen die leere Menge hier separat betrachten. Außerdem ist bei uns ein affiner Unterraum auch zu sich selbst parallel, dadurch wird Parallelität eine Äquivalenzrelation.

3.20. BEMERKUNG. Ein affiner Unterraum ist also das Bild eines Untervektorraums unter der Verschiebung um  $a_0$ .

- (1) In der Definition kommt es nicht darauf an, welches  $a_0 \in A$  wir wählen. Denn sei  $a_1 = a_0 + u_1 \in A$ , dann gilt nach dem Unterraumaxiom (U2), dass

$$a_1 + U = a_0 + (u_1 + U) = a_0 + U .$$

- (2) Ein affiner Unterraum ist genau dann ein Untervektorraum, wenn  $0 \in A$ . Die Richtung „ $\implies$ “ folgt aus (U1), und „ $\impliedby$ “ folgt aus (1), denn aus  $0 \in A$  folgt  $A = 0 + U = U$  für einen Untervektorraum  $U \subset V$ . Insbesondere ist jeder Untervektorraum auch ein affiner Unterraum.
- (3) Es sei  $U \subset V$  ein Untervektorraum. Die Menge aller zu  $U$  parallelen affinen Unterräume von  $V$  ist gerade der Quotientenraum  $V/U$  aus Definition 2.38.
- (4) In der Euklidischen Geometrie betrachtet man affine Unterräume des  $\mathbb{R}^3$  der Dimensionen 0 (Punkte), 1 (Geraden) und 2 (Ebenen).

Wir kommen zu *linearen Gleichungssystemen*. Gegeben eine Matrix  $A \in M_{m,n}(\mathbb{k})$ , die sogenannte *linke Seite* und ein Vektor  $b \in \mathbb{k}^m$ , die *rechte Seite*, suchen wir alle Vektoren  $x \in \mathbb{k}^n$ , so dass  $A \cdot x = b$ . Das heißt, wir suchen die *Lösungsmenge*

$$L = \{ x \in \mathbb{k}^n \mid A \cdot x = b \} .$$

Wenn wir die Gleichung  $A \cdot x = b$  ausschreiben, erhalten wir tatsächlich ein System linearer Gleichungen, nämlich

$$(*) \quad \begin{array}{ccccccc} a_{11} \cdot x_1 & + & \dots & + & a_{1n} \cdot x_n & = & b_1, \\ \vdots & & & & \vdots & & \vdots \\ a_{m1} \cdot x_1 & + & \dots & + & a_{mn} \cdot x_n & = & b_m. \end{array}$$

Wir nennen das Gleichungssystem  $(*)$  *homogen*, wenn  $b = 0$ , und *inhomogen*, wenn  $b \neq 0$ . Das zu  $A \cdot x = b$  gehörige homogene Gleichungssystem ist  $A \cdot x = 0$ .

Etwas allgemeiner können wir eine lineare Abbildung  $F: V \rightarrow W$  und eine „rechte Seite“  $w \in W$  betrachten, und nach der „Lösungsmenge“

$$L = \{ v \in V \mid F(v) = w \} = F^{-1}(\{w\}),$$

also dem Urbild von  $w$  unter  $F$ , fragen. Wenn  $V$  und  $W$  endlichdimensional sind, können wir Basen wählen und  $F$  als Matrix schreiben, und erhalten ein lineares Gleichungssystem vom obigen Typ.

**3.21. BEMERKUNG.** Lineare Gleichungssysteme treten zum Beispiel beim Lösen der folgenden Probleme auf.

- (1) Betrachte  $A \in M_{m,n}(\mathbb{k})$ , dann ist der Kern  $\ker A$  von  $A$  gerade die Lösungsmenge des homogenen Gleichungssystems  $A \cdot x = 0$ .
- (2) Sei  $A$  wie oben, dann liegt  $b \in \mathbb{k}^m$  genau dann im Bild im  $A$  von  $A$ , wenn das Gleichungssystem  $A \cdot x = b$  (mindestens) eine Lösung hat.
- (3) Es sei  $B \in M_n \mathbb{k}$  eine Basis des  $\mathbb{k}^n$ . Um die Koordinaten  $x$  eines Vektors  $v \in \mathbb{k}^n$  bezüglich  $B$  zu bestimmen, müssen wir nach Bemerkung 2.74 das lineare Gleichungssystem  $B \cdot x = v$  lösen. Für Orthonormalbasen geht es einfacher, siehe Proposition 3.34.
- (4) Eine quadratische Matrix  $A \in M_n(\mathbb{k})$  ist genau dann invertierbar, wenn eine Matrix  $B \in M_n(\mathbb{k})$  mit  $A \cdot B = E_n$  existiert (Übung). Um die Spalten  $b_1, \dots, b_n$  von  $B$  zu bestimmen, müssen wir die  $n$  Gleichungssysteme  $A \cdot b_i = e_i$  lösen.
- (5) Das Bestimmen von Schnittpunkten von Geraden und Ebenen im Euklidischen Raum führt oft auf lineare Gleichungssysteme. Seien etwa eine Gerade  $G$  und eine Ebene  $E \subset \mathbb{R}^3$  gegeben durch

$$E = \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \cdot r + \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \cdot s \mid r, s \in \mathbb{R} \right\}$$

und

$$G = \left\{ \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \cdot t \mid t \in \mathbb{R} \right\},$$

dann bestimmen wir  $G \cap E$  durch Lösen des Gleichungssystems

$$\begin{array}{rcl} 2 + r + s = 3 + 2t & & r + s - 2t = 1, \\ -r & = & 2 + t & \iff & -r & - & t = 2, \\ -s = 1 + t & & & & -s & - & t = 1. \end{array}$$

Die einzige Lösung dieses Systems ist  $r = -1$ ,  $s = 0$ ,  $t = -t$ ; sie führt auf den einzigen Schnittpunkt

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

- (6) In der Numerik werden viele Probleme zunächst durch lineare Probleme approximiert. Anschließend sind dann lineare Gleichungssysteme zu lösen. Oftmals sind die auftretenden Matrizen „dünn besetzt“, das heißt, in jeder Zeile und/oder jeder Spalte stehen nur sehr wenige von 0 verschiedene Zahlen. Für solche Gleichungssysteme gibt es spezielle, effiziente, teils approximative Lösungsverfahren. Sie lernen sie in den entsprechenden Vorlesungen kennen.

Es folgen einfache, grundsätzliche Überlegungen zum Lösungsverhalten linearer Gleichungssysteme.

3.22. PROPOSITION. *Es sei  $A \in M_{m,n}(\mathbb{k})$  und  $b \in \mathbb{k}^m$ .*

- (1) *Die Lösungsmenge des homogenen Gleichungssystems  $A \cdot x = 0$  ist gerade  $\ker A$ .*
- (2) *Das inhomogene Gleichungssystem  $A \cdot x = b$  hat genau dann Lösungen, wenn  $b \in \operatorname{im} A$ .*
- (3) *Es sei  $A \cdot x_0 = b$ , dann ist die Lösungsmenge des inhomogenen Gleichungssystems  $A \cdot x = b$  der affine Unterraum*

$$\{ x \in \mathbb{k}^n \mid A \cdot x = b \} = x_0 + \ker A .$$

BEWEIS. Die Aussagen (1) und (2) sind gerade die Punkte (1) und (2) aus Bemerkung 3.21. Zu (3) beachten wir, dass aus  $A \cdot x_0 = b$  folgt, dass

$$A \cdot x = b \iff A \cdot (x - x_0) = b - b = 0 \iff x - x_0 \in \ker A . \quad \square$$

Punkt (3) wird gern so umformuliert: Die *allgemeine Lösung*  $x$  des inhomogenen Gleichungssystems  $A \cdot x + b$  ist die Summe aus einer *speziellen Lösung*  $x_0$  des inhomogenen Gleichungssystems und der allgemeinen Lösung  $v = x - x_0$  des zugehörigen homogenen Gleichungssystems  $A \cdot v = 0$ .

3.23. PROPOSITION. *Es seien  $A \in M_{m,n}(\mathbb{k})$  und  $b \in \mathbb{k}^m$ . Die Lösungsmenge des linearen Gleichungssystems  $A \cdot x = b$  verändert sich nicht, wenn man  $A$  und  $b$  von links mit der gleichen invertierbaren Matrix  $B \in GL(m, \mathbb{k})$  multipliziert.*

BEWEIS. Es sei  $x \in \mathbb{k}^n$  mit  $A \cdot x = b$ , dann folgt

$$(B \cdot A) \cdot x = B \cdot (A \cdot x) = B \cdot b .$$

Gelte umgekehrt  $(B \cdot A) \cdot x = B \cdot b$ , und sei  $B^{-1}$  die Inverse von  $B$ , dann folgt

$$A \cdot x = B^{-1} \cdot (B \cdot A) \cdot x = B^{-1} \cdot B \cdot b = b .$$

Also haben das alte und das neue Gleichungssystem die gleichen Lösungen.  $\square$



Ein Gleichungssystem  $A \cdot x = b$  heißt *in (strenger) Zeilenstufenform*, wenn die Matrix  $A$  in (strenger) Zeilenstufenform ist.

Eine Matrix  $A = (a_{ij})_{i,j}$  in Zeilenstufenform hat also folgende Gestalt:

$$r \begin{pmatrix} 0 & \dots & 0 & 1 & a_{1,j_1+1} & \dots & a_{1,j_2-1} & * & a_{1,j_2+1} & \dots & a_{1,j_r-1} & * & a_{1,j_r+1} & \dots & a_{1,n} \\ 0 & & & & \dots & & 0 & 1 & a_{2,j_2+1} & \dots & a_{2,j_r-1} & * & a_{2,j_r+1} & \dots & a_{2,n} \\ \vdots & & & & & & & & & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & & & \dots & & & & & a_{r-1,j_r-1} & * & a_{r-1,j_r+1} & \dots & a_{r-1,n} \\ 0 & & & \dots & & & & & 0 & 1 & a_{r,j_r+1} & \dots & a_{r,n} \\ 0 & & & & & & \dots & & & & & & 0 \\ \vdots & & & & & & & & & & & & \vdots \\ 0 & & & & & & \dots & & & & & & 0 \end{pmatrix}.$$

Die „\*“ sind beliebig, verschwinden aber, wenn  $A$  in *strenger* Zeilenstufenform ist. Die Zahlen  $r$  und  $j_1, \dots, j_r$  sind durch  $A$  eindeutig bestimmt. Wir sehen in Proposition 3.27 unten, dass man bei einem Gleichungssystem in Zeilenstufenform die Lösungsmenge leicht ablesen kann.

3.26. SATZ (Gauß-Verfahren). *Jedes lineare Gleichungssystem lässt sich mit Hilfe elementarer Zeilenumformungen in (strenge) Zeilenstufenform bringen.*

Andere Namen sind *Gauß-Algorithmus*, *Gauß-Elimination*, sowie *Gauß-Jordan-Verfahren* für die strenge Zeilenstufenform.

BEWEIS. Das Gauß-Verfahren ist ein induktiver Algorithmus, bei dem man eine Reihe elementarer Zeilenumformungen auf die Matrix  $A$  und die rechte Seite  $b$  anwendet und so die Matrix  $A$  Spalte für Spalte in strenge Zeilenstufenform bringt.

*Induktionsannahme.* Es seien  $r \geq 0$  und  $1 \leq j_1 < \dots < j_r \leq n$  sowie  $q$  mit  $j_r \leq q \leq n$  (beziehungsweise  $q \geq 0$ , falls  $r = 0$ ) gegeben, so dass die Bedingungen (1) und (2) (beziehungsweise (1)–(3) für strenge Zeilenstufenform) in Definition 3.25 für alle  $i \leq n$  und für alle  $j \leq q$  gelten. Das heißt, die Matrix  $A$  ist bis einschließlich Spalte  $q$  bereits in strenger Zeilenstufenform.

*Induktionsanfang.* Wir beginnen mit  $q = r = 0$ . Dann sind die obigen Annahmen trivialerweise erfüllt.

*Induktionsschritt.* Falls  $r = m$  oder  $q = n$  gilt, sind wir fertig. Ansonsten setzen wir  $j = q + 1 \leq n$  und unterscheiden zwei Fälle.

1. *Fall:* Falls es kein  $i$  mit  $r < i \leq m$  und  $a_{ij} \neq 0$  gibt, ist die Matrix bereits bis zur  $j$ -ten Spalte in strenger Zeilenstufenform. In diesem Fall erhöhen wir  $q$  um 1, so dass  $q = j$ , und führen den nächsten Induktionsschritt durch.

2. *Fall:* Ansonsten gibt es ein kleinstes  $i > r$  mit  $a_{ij} \neq 0$ .

*Schritt 1 („Tauschen“):* Falls  $i \neq r + 1$ , vertauschen wir die  $i$ -te und die  $(r + 1)$ -te Zeile mit einer elementaren Zeilenumformung vom Typ (1). Anschließend erhöhen wir  $r$  um 1, so dass jetzt also  $a_{rj} \neq 0$ .

*Schritt 2 („Normieren“):* Falls  $a_{rj} \neq 1$ , multiplizieren wir die  $r$ -te Zeile mit  $a_{rj}^{-1}$ , so dass anschließend  $a_{rj} = 1$ , das ist eine elementare Zeilenumformung vom Typ (2). Jetzt setzen wir  $j_r = j$ , so dass jetzt  $a_{rj_r} = 1$ , das heißt, Punkt (2) in Definition 3.25 ist für  $i = r$  erfüllt.

*Schritt 3 („Ausräumen“):* Schließlich subtrahieren wir von der  $i$ -ten Zeile das  $a_{ij_r}$ -fache der  $r$ -ten Zeile für alle  $i > r$  (beziehungsweise für alle  $i \neq r$  für die strenge Zeilenstufenform), das ist eine elementare Zeilenumformung vom Typ (3), so dass hinterher  $a_{ij_r} = 0$  für alle  $i > r$  (beziehungsweise für alle  $i \neq r$ ).

Danach erhöhen wir  $q$  um 1, so dass jetzt  $q = j$ , und haben nun auch Punkt (1) (und gegebenenfalls auch (3)) in Definition 3.25 für alle  $j \leq q$  erfüllt. Anschließend wiederholen wir den Induktionsschritt.

Am Ende erhalten wir eine Matrix in Zeilenstufenform, beziehungsweise in strenger Zeilenstufenform, je nachdem, ob wir in Schritt 3 die gesamte Spalte oder nur unterhalb vom jeweiligen  $r$  ausgeräumt haben.  $\square$

Man beachte, dass wir in einem Schritt eine ganze Zeile durch  $a_{rj_r}$  dividieren mussten, um  $a_{rj_r} = 1$  zu erreichen. Aus diesem Grund lässt sich das Gauß-Verfahren nicht auf Matrizen über Ringen anwenden, in denen nicht alle Elemente außer 0 invertierbar sind, die also keine (Schief-) Körper sind.

**3.27. PROPOSITION.** *Sei  $A \in M_{m,n}(\mathbb{k})$  eine Matrix in Zeilenstufenform, und sei  $b \in \mathbb{k}^m$ .*

- (1) *Eine Basis des Bildes  $\text{im } A = \mathbb{k}^r \times \{0\} \subset \mathbb{k}^m$  von  $A$  besteht aus den Spalten  $a_{j_i} = A(e_{j_i})$  für  $i = 1, \dots, r$ , insbesondere ist  $\text{rg } A = r$ .*
- (2) *Das Gleichungssystem (\*) ist genau dann lösbar, wenn  $b_{r+1} = \dots = b_m = 0$ ; in diesem Fall hat die Lösungsmenge die Gestalt*

$$\begin{aligned} & \{ x \in \mathbb{k}^n \mid A \cdot x = b \} \\ & = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{k}^n \mid x_{j_i} = b_i - \sum_{j=j_i+1}^n a_{ij} x_j \text{ für alle } i = 1, \dots, r \right\}, \end{aligned}$$

*jede Lösung ist also eindeutig bestimmt durch die Angabe der Koordinaten  $x_j$  für alle  $j \in \{1, \dots, n\} \setminus \{j_1, \dots, j_r\}$ .*

- (3) *Es sei  $A$  in strenger Zeilenstufenform, und es sei  $\{k_{r+1}, \dots, k_n\} = \{1, \dots, n\} \setminus \{j_1, \dots, j_r\}$  eine Aufzählung der restlichen Spaltenindizes, dann erhalten wir eine Basis  $(c_{r+1}, \dots, c_n)$  von  $\ker A$  aus Vektoren der Form*

$$c_\ell = e_{k_\ell} - \sum_{i=1}^r e_{j_i} \cdot a_{ik_\ell} \in \ker A \subset \mathbb{k}^n \quad \text{für } \ell = r+1, \dots, n.$$

Für die Basis von  $\ker A$  in (2) benutzen wir die gleichen Buchstaben wie im Beweis des Rangsatzes 3.13.

BEWEIS. Zu Aussage (1) überlegen wir uns zunächst, dass im  $A \subset \mathbb{k}^r \times \{0\} \subset \mathbb{k}^n$ , da alle Spalten von  $A$  in diesem Unterraum liegen.

Sei umgekehrt  $b \in \mathbb{k}^r \times \{0\}$ , dann hat die Lösungsmenge die in (2) angegebene Gestalt. Wenn wir  $x_j$  für  $j \in \{1, \dots, n\} \setminus \{j_1, \dots, j_r\}$  beliebig vorgeben, bestimmen die Zeilen  $i = r, \dots, 1$  in umgekehrter Reihenfolge die fehlenden Koordinaten  $x_{j_r}, \dots, x_{j_1}$  eindeutig. Daraus folgt (2) sowie im  $A = \mathbb{k}^r \times \{0\}$  und insbesondere  $r = \operatorname{rg} A$ , also gilt auch (1).

Zu (3) wählen wir  $b = 0$  und bestimmen Elemente  $c_\ell$  der Lösungsmenge  $\ker A$ , indem wir für  $p = r + 1, \dots, n$  die Koordinaten  $x_{k_p} = \delta_{\ell p}$  vorgeben. Die restlichen Koordinaten sind gerade die  $x_{j_i}$ ,  $i = 1, \dots, r$ . Wenn  $A$  in strenger Zeilenstufenform ist, bestimmen wir  $x_{j_i}$  durch die  $i$ -te Gleichung und erhalten

$$x_{j_i} = - \sum_{p=r+1}^n a_{ik_p} \cdot x_{k_p} = -a_{ik_\ell}.$$

Man überprüft anhand der Koordinaten  $x_{k_{r+1}}, \dots, x_{k_n}$ , dass diese Vektoren linear unabhängig sind. Sie erzeugen den Kern, bilden also eine Basis, da

$$\dim \ker A = n - \dim \operatorname{im} A = n - r. \quad \square$$

Wenn man das Gauß-Verfahren konkret anwendet, schreibt man gern die jeweilige linke Seite als Matrix ohne runde Klammern, macht rechts daneben einen senkrechten Strich, und schreibt die rechte Seite rechts neben diesen Strich. Dann führt man den obigen Algorithmus durch, wobei man sich nur an der linken Seite orientiert, aber alle Zeilenumformungen immer auf die linke und die rechte Seite simultan anwendet. Dabei reicht es, für jeden Induktionsschritt ein neues System aufzuschreiben. Unter Umständen kann es sinnvoll sein, auf der rechten Seite mehr als nur einen Vektor stehen zu haben, zum Beispiel, wenn man ein Gleichungssystem simultan für mehrere rechte Seiten zu lösen hat.

3.28. BEMERKUNG. Das Gauß-Verfahren kann benutzt werden, um viele verschiedene Probleme zu lösen. Einige davon haben wir in Bemerkung 3.21 bereits angeführt.

- (1) Um das Gleichungssystem (\*), also  $A \cdot x = b$  zu lösen, bringen wir es zunächst mit dem Gauß-Verfahren in Zeilenstufenform. Nach Bemerkung 3.24 entsprechen elementare Zeilenumformungen gerade der Multiplikation mit invertierbaren Matrizen von links. Da wir alle Zeilenumformungen sowohl auf die linke als auch auf die rechte Seite des Gleichungssystems angewandt haben, ist das neue Gleichungssystem nach Proposition 3.23 zum alten äquivalent, und wir können die Lösungsmenge nach Proposition 3.27 (2) ablesen.

Zur Sicherheit sei daran erinnert, dass man ein Gleichungssystem löst, indem man die *gesamte* Lösungsmenge angibt (eventuell, indem man feststellt, dass diese leer ist), und nicht nur ein einzelnes Element der Lösungsmenge. Wenn die Lösungsmenge nicht leer ist, reicht es allerdings nach Proposition 3.22 (3), eine spezielle Lösung  $x_0$  und

den Unterraum  $\ker A$  zu bestimmen, da die Lösungsmenge dann gerade  $x_0 + \ker A$  ist.

- (2) Es sei  $A \in M_{m,n}(\mathbb{k})$ , dann können wir Basen von  $\ker A \subset \mathbb{k}^n$  und im  $A \subset \mathbb{k}^m$  bestimmen. Wir bringen dazu  $A$  mit dem Gauß-Verfahren in strenge Zeilenstufenform  $BA$ , mit  $B \in GL_m(\mathbb{k})$ . Proposition 3.27 (3) liefert eine Basis von  $\ker A = \ker(BA)$ .

Nach Proposition 3.27 (1) bilden die Spalten  $j_1, \dots, j_r$  von  $BA$  eine Basis von  $\text{im}(BA) = B \cdot \text{im} A$ . Das sind aber genau die Bilder der Spalten  $j_1, \dots, j_r$  von  $A$  unter Multiplikation mit  $B$ . Da letztere ein linearer Isomorphismus ist, bilden die Spalten  $j_1, \dots, j_r$  von  $A$  eine Basis von  $\text{im} A$ .

Übrigens können wir die Basis  $(c_{r+1}, \dots, c_n)$  von  $\ker A$  wie im Beweis des Rangsatzes 3.13 zu einer Basis  $(c_1, \dots, c_n)$  von  $\mathbb{k}^n$  mit  $c_i = e_{j_i}$  für  $i = 1, \dots, r$  ergänzen. Wenn wir wie dort fortfahren, erhalten wir ebenfalls die obige Basis  $(A(e_{j_1}), \dots, A(e_{j_r}))$  von  $\text{im} A$ .

- (3) Es seien Vektoren  $v_1, \dots, v_n \in \mathbb{k}^m$  gegeben. Wir möchten wissen, ob diese Vektoren linear unabhängig sind, und ob sie  $\mathbb{k}^m$  erzeugen. Dazu schreiben wir die Vektoren als Spalten in eine Matrix  $A$  und bringen  $A$  in Zeilenstufenform. Dann bilden  $(v_1, \dots, v_n)$  genau dann ein Erzeugendensystem, wenn  $r = \text{rg} A = m$  gilt.

Und sie sind linear unabhängig, wenn  $A \cdot x = 0$  nur eine Lösung besitzt. Nach Proposition 3.27 (2) ist das genau dann der Fall, wenn  $\{j_1, \dots, j_r\} = \{1, \dots, n\}$ , das heißt, wenn  $r = \text{rg} A = n$  gilt.

- (4) Um eine Matrix  $A \in M_n(\mathbb{k})$  zu invertieren, wenden wir das Gauß-Verfahren diesmal mit der rechten Seite  $E_n$  an, das heißt, wir lösen  $n$  lineare Gleichungssysteme mit der gleichen linken Seite simultan. Wenn wir während des Verfahrens nie eine Spalte überspringen (Fall 1 im Beweis tritt nicht ein) und  $A$  in strenge Zeilenstufenform bringen, dann gilt  $j_i = i$  für alle  $i = 1, \dots, n$ . Also bleibt auf der linken Seite die Einheitsmatrix  $E_n$  stehen.

Rechts steht das Produkt  $B$  aller Elementarmatrizen, die wir im Laufe des Verfahrens angewendet haben, also

$$A \mid E_n \quad \rightsquigarrow \quad E_n \mid B.$$

Es gilt also  $B \cdot A = E_n$ . Da beide Matrizen quadratisch waren, ist  $A$  invertierbar, und  $B$  ist die inverse Matrix; dazu interpretiere  $A$  und  $B$  als lineare Abbildungen und wende eine Übungsaufgabe an.

Falls wir im Laufe des Gauß-Verfahrens eine Spalte überspringen, so dass  $j_{i_0+1} > j_{i_0} + 1$  für ein  $i_0$  (oder  $j_1 > 1$  für  $i_0 = 0$ ), folgt  $i < j_i$  für alle  $i > i_0$ , insbesondere  $r < j_r \leq n$ , so dass  $\text{rg} A < n$  gilt und  $A$  daher nicht invertierbar sein kann. Das bedeutet, dass wir das Verfahren abbrechen können, sobald Fall 1 eintritt, und feststellen können, dass  $A$  nicht invertierbar ist. Aus diesem Grund ist es geschickter, zunächst nur auf Zeilenstufenform hinzuarbeiten, und erst dann, wenn man weiß, dass die Matrix invertierbar ist, auch oberhalb der Diagonalen auszuräumen.

Für sehr große Matrizen ist das Gauß-Verfahren zu rechenaufwendig. Auch wenn die ursprüngliche Matrix nur wenige von 0 verschiedene Einträge pro Zeile und/oder Spalte enthielt, kann sich das bereits nach einigen Zwischenschritten ändern. Es gibt aber noch ein anderes Problem, sobald man nicht mit exakten Zahlen rechnet, sondern in jedem Zwischenschritt nach einer bestimmten Anzahl von Dual- oder Dezimalstellen rundet oder abschneidet: Sobald man zwei annähernd gleich große Zahlen mit kleinen prozentualen Fehlern voneinander abzieht, erhält man einen wesentlich größeren prozentualen Fehler im Ergebnis. Daher benutzt man für große Matrizen andere, effizientere Verfahren.

### 3.4. Die Methode der kleinsten Quadrate

Eigentlich besprechen wir Skalarprodukte erst in Kapitel 6. Wir führen das Standardskalarprodukt auf  $\mathbb{k}^n$  für  $\mathbb{k} = \mathbb{C}$  oder  $\mathbb{H}$  bereits hier ein, da es uns einige nützliche Rechentechniken an die Hand gibt. Außerdem können wir damit die von Gauß und Legendre gefundene Methode der kleinsten Quadrate herleiten. Tatsächlich war diese Methode für Gauß möglicherweise der Hauptgrund, sich überhaupt mit linearen Gleichungssystemen zu beschäftigen.

Für die nächste Definition brauchen wir den Begriff der adjungierten Matrix. Wir erinnern uns an die komplexe und die quaternionische Konjugation aus den Definitionen 1.60 und 1.71, und die Rechenregeln aus Bemerkung 1.61 und Satz 1.72 (5), (6). Für  $a \in \mathbb{R}$  sei wieder  $\bar{a} = a$ .

3.29. DEFINITION. Falls  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ , definieren wir die zu  $A$  *adjungierte Matrix*  $A^* \in M_{n,m}(\mathbb{k})$  durch  $A^* = (\bar{a}_{ij})_{j,i}$ .

Für  $\mathbb{k} = \mathbb{R}$  ist Adjungieren das gleiche wie Transponieren, siehe Definition 3.14.

3.30. DEFINITION. Es sei  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ . Wir definieren das *Standardskalarprodukt*  $\langle \cdot, \cdot \rangle: \mathbb{k}^n \times \mathbb{k}^n \rightarrow \mathbb{k}$  durch

$$\langle v, w \rangle = \sum_{i=1}^n \bar{v}_i \cdot w_i = v^* \cdot w .$$

3.31. BEMERKUNG. Im Fall  $\mathbb{k} = \mathbb{R}$  entspricht das genau dem Standardskalarprodukt aus Definition 1.52. Für  $\mathbb{k} = \mathbb{C}$  und  $\mathbb{H}$  hat das Standardskalarprodukt für alle  $v, w \in \mathbb{k}^n$  und alle  $r \in \mathbb{k}$  die folgenden Eigenschaften.

- (1) Es ist *sesquilinear*. Das heißt, es ist linear im zweiten Argument und additiv im ersten Argument, aber anstelle von Homogenität gilt

$$\langle v \cdot r, w \rangle = \bar{r} \cdot \langle w, v \rangle .$$

- (2) Es ist *Hermiteisch*, das heißt, es gilt

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

Hieraus folgt insbesondere, dass  $2 \operatorname{Re}\langle v, w \rangle = \langle v, w \rangle + \langle w, v \rangle$ .

(3) Es ist *positiv definit*, das heißt, es gilt

$$\mathbb{R} \ni \langle v, v \rangle \geq 0 \quad \text{und} \quad \langle v, v \rangle = 0 \iff v = 0 .$$

Im Falle  $\mathbb{k} = \mathbb{R}$  wird aus (1) Bilinearität, und aus (2) Symmetrie.

3.32. PROPOSITION. *Es sei  $R$  ein kommutativer Ring und  $\ell, m, n \in \mathbb{N}$ . Für alle  $A \in M_{m,n}(R)$  und alle  $B \in M_{\ell,m}(R)$  gilt*

$$(1) \quad (A^t)^t = A \quad \text{und} \quad A^t \cdot B^t = (B \cdot A)^t .$$

*Es sei  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$  und  $\ell, m, n \in \mathbb{N}$ . Für alle  $A \in M_{m,n}(\mathbb{k})$ , alle  $B \in M_{\ell,m}(\mathbb{k})$  und alle  $v \in \mathbb{k}^n, w \in \mathbb{k}^m$  gilt*

$$(2) \quad (A^*)^* = A, \quad \langle w, Av \rangle = \langle A^*w, v \rangle \quad \text{und} \quad A^* \cdot B^* = (B \cdot A)^* .$$

Wir lassen den Beweis als Übung.

3.33. DEFINITION. Eine *Orthonormalbasis* ( $\mathbb{k} = \mathbb{R}$ ) beziehungsweise eine (*quaternionisch*) *unitäre Basis* ( $\mathbb{k} = \mathbb{C}, \mathbb{H}$ ) des  $\mathbb{k}^n$  ist ein Tupel  $B = (b_1, \dots, b_n)$  von Vektoren im  $\mathbb{k}^n$ , so dass für alle  $i, j$  gilt

$$\langle b_i, b_j \rangle = \delta_{ij} .$$

3.34. PROPOSITION. *Ein Tupel  $B = (b_1, \dots, b_n)$  ist genau dann eine Orthonormal- beziehungsweise unitäre Basis des  $\mathbb{k}^n$ , wenn die Matrix  $B$  mit den Spalten  $b_1, \dots, b_n$  invertierbar ist mit  $B^{-1} = B^*$ . Wenn das so ist, ist  $B$  eine Basis, und für jeden Vektor  $v \in \mathbb{R}^n$  gilt*

$$v = \sum_{i=1}^n b_i \cdot \langle b_i, v \rangle .$$

Man beachte, dass die Matrix  $B$  jetzt tatsächlich die Matrix der Basisabbildung  $B$  ist, so dass die Merkregel aus Bemerkung 2.74 hier wirklich richtig ist. Im Spezialfall einer Orthonormalbasis wird die Koordinatenabbildung also gegeben durch  $B^{-1} = B^*$ . Im Allgemeinen lässt sich das Inverse einer Matrix nicht so leicht bestimmen, und allgemeine Verfahren zum Invertieren von Matrizen lernen wir später in diesem und im nächsten Kapitel kennen.

BEWEIS. In den Übungen haben wir bereits gezeigt, dass orthonormale Tupel von Vektoren linear unabhängig sind, und die Darstellung der Koordinatenabbildung überprüft. Aus Folgerung 3.5 (3) folgt, dass  $B$  eine Basis ist. Wir berechnen noch

$$B^* \cdot B = \left( \sum_{k=1}^n \bar{b}_{ki} b_{kj} \right)_{i,j} = (\langle b_i, b_j \rangle)_{i,j}$$

und schließen daraus, dass  $B^{-1} = B^*$  genau dann, wenn  $B$  eine Orthonormalbasis ist.  $\square$

Wenn man etwas ausmessen muss, dann macht man oft viele Messungen, in der Hoffnung, dass sich die Fehler der einzelnen Messungen dabei in etwa ausgleichen. Danach muss man den tatsächlichen Wert „schätzen“.

3.35. BEISPIEL. Wir wollen den Energieverbrauch eines Autos bei konstanter Geschwindigkeit pro zurückgelegter Strecke messen (physikalisch gesehen ist das eine Kraft). Dass wir überhaupt Energie verbrauchen, selbst wenn wir nicht beschleunigen, liegt an verschiedenen Typen von Reibung. Beispielsweise gibt es die Rollreibung der Räder auf der Straße und der Achsen in ihren Kugellagern, sowie den Strömungswiderstand der Karosserie an der umgebenden Luft. Wir modellieren das durch ein Polynom zweiten Grades für die Reibungskraft

$$F(v) = a_1 + a_2v + a_3v^2,$$

dabei sei  $v$  die Geschwindigkeit. Jetzt geht es darum,  $a_1$ ,  $a_2$  und  $a_3$  zu bestimmen ( $a_3$  ist übrigens der  $c_w$ -Wert).

Wenn wir die Reibungskraft  $F_1$ ,  $F_2$ ,  $F_3$  bei drei Geschwindigkeiten  $v_1$ ,  $v_2$ ,  $v_3$  messen, erhalten wir ein Gleichungssystem in den drei Unbekannten  $a_1$ ,  $a_2$ ,  $a_3$  mit drei Gleichungen

$$\begin{aligned} 1 \cdot a_1 + v_1 \cdot a_2 + v_1^2 \cdot a_3 &= F_1 \\ 1 \cdot a_1 + v_2 \cdot a_2 + v_2^2 \cdot a_3 &= F_2 \\ 1 \cdot a_1 + v_3 \cdot a_2 + v_3^2 \cdot a_3 &= F_3 \end{aligned}$$

Das können wir mit dem Gaußverfahren eindeutig lösen, wenn keine zwei Geschwindigkeiten gleich sind.

Wenn wir stattdessen hundertmal messen, bekommen wir hundert Gleichungen in nach wie vor nur drei Unbekannten. In Matrixschreibweise  $V \cdot a = F$  mit

$$V = \begin{pmatrix} 1 & v_1 & v_1^2 \\ \vdots & \vdots & \vdots \\ 1 & v_{100} & v_{100}^2 \end{pmatrix} \in \mathbb{R}^{100 \times 3} \quad \text{und} \quad F = \begin{pmatrix} F_1 \\ \vdots \\ F_{100} \end{pmatrix} \in \mathbb{R}^{100}.$$

Dieses Gleichungssystem ist in der Regel unlösbar.

Es sei jetzt allgemeiner  $Ax = b$  ein Gleichungssystem mit  $A \in \mathbb{k}^{m \times n}$  und  $b \in \mathbb{k}^m$  für  $\mathbb{k} = \mathbb{R}$ ,  $\mathbb{C}$  oder  $\mathbb{H}$ . Wir suchen nicht mehr nach einer exakten Lösung, sondern versuchen, die Norm von  $Ax - b$  zu minimieren. Geometrisch gesehen suchen wir denjenigen Punkt  $Ax$  im Unterraum  $\text{im}(A) \subset \mathbb{k}^m$ , der am nächsten am Punkt  $b \in \mathbb{k}^m$  liegt bezüglich des Euklidischen Abstands, siehe Abschnitt 1.4 im Fall  $\mathbb{k} = \mathbb{R}$ .

3.36. SATZ (Methode der kleinsten Quadrate). *Es  $A \in M_{m,n}(\mathbb{k})$  und  $b \in \mathbb{k}^m$  für  $\mathbb{k} = \mathbb{R}$ ,  $\mathbb{C}$  oder  $\mathbb{H}$ . Dann ist das Gleichungssystem  $A^*Ax = A^*b$  immer lösbar, je zwei Lösungen unterscheiden sich um ein Element von  $\ker A$ , und  $\|Ax - b\|$  nimmt genau auf der Lösungsmenge das Minimum an.*

Im obigen Beispiel müssten wir also  $V^*V \cdot a = V^* \cdot F$  lösen. Das ist jetzt nur noch ein Gleichungssystem mit drei Gleichungen in drei Unbekannten. Das Minimum von  $\|V \cdot a - F\|$  — also der Wert für ein  $a$  aus der Lösungsmenge — ist eine untere Schranke für die Messgenauigkeit.

BEWEIS. Zunächst nehmen wir an, es gäbe eine Lösung  $x_0 \in \mathbb{k}^n$  des Gleichungssystems

$$(*) \quad A^*Ax = A^*b,$$

wobei wir die Multiplikationszeichen der Kürze halber weglassen. Sei  $x \in \mathbb{k}^n$  beliebig, dann rechnen wir mit Hilfe von Bemerkung 3.31 (2), (3) und Proposition 3.32 (2) nach, dass

$$\begin{aligned} \|Ax - b\|^2 &= (A(x - x_0) + Ax_0 - b)^*(A(x - x_0) + Ax_0 - b) \\ &= \|Ax_0 - b\|^2 + 2\operatorname{Re}\langle A(x - x_0), Ax_0 - b \rangle + \|A(x - x_0)\|^2 \\ &= \|Ax_0 - b\|^2 + 2\operatorname{Re}\langle x - x_0, \underbrace{A^*Ax_0 - A^*b}_{=0} \rangle + \|A(x - x_0)\|^2 \\ &\geq \|Ax_0 - b\|^2. \end{aligned}$$

Gleichheit gilt offensichtlich genau dann, wenn  $x - x_0 \in \ker A$ .

Zur Lösbarkeit sei  $B$  ein Produkt von Elementarmatrizen, so dass  $BA^*$  in Zeilenstufenform ist. Da  $B$  invertierbar ist, ist auch  $B^*$  invertierbar mit inverser Matrix  $(B^*)^{-1} = (B^{-1})^*$ . Also ist  $A^*Ax = A^*b$  genau dann lösbar, wenn  $(BA^*)(BA^*)^*y = (BA^*)b$  lösbar ist, denn  $y$  ist eine Lösung des neuen Systems genau dann, wenn  $x = B^*y$  eine Lösung von (\*) ist.

Ohne Einschränkung sei also  $A^*$  bereits in Zeilenstufenform vom Rang  $r$ . Dann ist  $A$  in „Spaltenstufenform“, als Blockmatrix geschrieben also  $A = (C \ 0)$  mit  $C \in M_{m,r}(\mathbb{k})$  und  $\operatorname{rg} C = r$ . Wir erhalten

$$\begin{aligned} A^*A &= \begin{pmatrix} C^* \\ 0 \end{pmatrix} \cdot (C \ 0) = \begin{pmatrix} C^*C & 0 \\ 0 & 0 \end{pmatrix} \\ \text{und} \quad A^*b &= \begin{pmatrix} C^* \\ 0 \end{pmatrix} \cdot b = \begin{pmatrix} C^*b \\ 0 \end{pmatrix}. \end{aligned}$$

Wenn wir  $x = \begin{pmatrix} y \\ z \end{pmatrix}$  mit  $y \in \mathbb{k}^r$  und  $z \in \mathbb{k}^{n-r}$  schreiben, ist (\*) jetzt äquivalent zum Gleichungssystem

$$(\dagger) \quad C^*C \cdot y = C^*b$$

— an  $z$  wird also keine Bedingung gestellt.

Da  $C \in M_{m,r}(\mathbb{k})$  Rang  $r$  hat, gilt  $\dim \ker C = \dim \mathbb{k}^r - \operatorname{rg} C = 0$  nach Satz 3.13. Also ist  $C$  injektiv. Dann ist  $C^*C$  invertierbar, denn für alle  $v \in \mathbb{k}^r$  gilt wegen Proposition 3.32 und Bemerkung 3.31 (3), dass

$$\langle v, C^*Cv \rangle = \langle Cv, Cv \rangle = \|Cv\|^2 \geq 0,$$

mit Gleichheit genau dann, wenn  $Cv = 0$ . Da  $C$  injektiv ist, ist das zu  $v = 0$  äquivalent. Nun kann  $\langle v, C^*Cv \rangle > 0$  nur gelten, wenn  $C^*Cv \neq 0$ , somit ist auch  $C^*C$  injektiv. Da  $C^*C$  eine quadratische Matrix ist, ist  $C^*C$  nach Satz 3.13 auch surjektiv, also ist  $(\dagger)$  lösbar. Aber dann ist auch das System (\*) lösbar.  $\square$

3.37. BEMERKUNG. Wenn das ursprüngliche Gleichungssystem  $Ax = b$  lösbar ist, findet die Methode der kleinsten Quadrate genau die Lösungsmenge, denn die Lösungen von  $Ax = b$  minimieren dann den Ausdruck  $\|Ax - b\|$ .

Wenn das ursprüngliche Gleichungssystem  $Ax = b$  nicht lösbar ist, können wir das Minimum von  $\|Ax - b\|$  als ein Maß für die Unlösbarkeit auffassen. Man beachte, dass sich zwei Lösungen des neuen Systems  $A^*Ax = A^*b$  wieder genau um ein Element aus  $\ker A$  unterscheiden, siehe Proposition 3.22 (3).

Wir haben hier die einfachste Variante der Methode der kleinsten Quadrate vorgestellt. Wenn die Zielfunktion  $F$  nicht mehr linear von den Parametern  $a_i$  abhängt, kann man das Problem nicht mehr mit Methoden der linearen Algebra lösen.

### 3.5. Zusammenfassung

Wir ziehen wieder Bilanz. Im ersten Abschnitt haben wir die Basissätze von Steinitz kennengelernt. Sie sind sehr mächtige Hilfsmittel, um abstrakte Probleme der linearen Algebra zu lösen, etwa Existenz von Basen, Existenz komplementärer Unterräume, und so weiter. Im zweiten Abschnitt haben wir diese Methoden benutzt, um Dimensionsformeln zu zeigen. Gleichzeitig liefern die Beweise der Steinitz-Sätze auch Algorithmen zur Konstruktion von Basen, die man für explizite Rechnungen nutzen kann.

Im zweiten Abschnitt ging es um die Dimension — die entscheidende Invariante für endlich erzeugte Vektorräume — und den Rang — die entscheidende Invariante für Abbildungen zwischen ihnen. Dimensionsformeln beschreiben das Verhalten diverser Konstruktionen in der linearen Algebra, etwa Summen von Unterräumen, Quotienten, oder Kern und Bild linearer Abbildungen.

Im dritten Abschnitt haben wir vordergründig lineare Gleichungssysteme kennengelernt und mit dem Gauß-Verfahren gelöst. Das Gauß-Verfahren kann aber noch mehr: es ist unser „Schweizer Messer“ für viele kleine bis mittelgroße Probleme der linearen Algebra. Wir kennen bereits weitere „Klingen“ zur Bestimmung von Kern und Bild linearer Abbildungen und zum Invertieren von Matrizen.

Der letzte Abschnitt behandelt die Methode der kleinsten Quadrate in einfachen Fällen. Sie ermöglicht es uns beispielsweise, Parameter anhand von Messungen zu schätzen. Weitere Schätz- und Approximationsverfahren lernen Sie in Stochastik und Numerik kennen. Historisch gesehen war die Methode der kleinsten Quadrate vermutlich der Grund für Gauß, sich überhaupt mit linearen Gleichungssystemen zu beschäftigen — das sogenannte Gauß-Verfahren selbst war übrigens in China schon wenige Jahrhunderte nach Christi Geburt bekannt, und in Europa ebenfalls schon vor seiner Beschreibung durch Gauß.

## KAPITEL 4

# Determinanten

In den nächsten Kapiteln wollen wir Endomorphismen von Vektorräumen beziehungsweise freien  $R$ -Moduln  $V$  verstehen, also lineare Abbildungen  $F: V \rightarrow V$ . Endomorphismen endlich erzeugter freier Moduln werden durch quadratische Matrizen  $A \in M_n(R)$  dargestellt. In diesem Kapitel lernen wir eine erste, wichtige Invariante quadratischer Matrizen kennen, die Determinante. Über den reellen Zahlen hat die Determinante etwas mit Volumina von Parallelotopen zu tun, und etwas mit Orientierung. Wir benötigen in den folgenden Kapiteln aber auch Determinanten von Matrizen über dem Polynomring eines Körpers.

Wir beginnen in Abschnitt 4.1 mit der Beschreibung von Volumina, benutzen die dort gewonnenen Erkenntnisse in Abschnitt 4.2 zur Definition der Determinante, und führen in Abschnitt 4.3 den Begriff der Orientierung ein. Im ganzen Kapitel benötigen wir das Kommutativgesetz für die Multiplikation. Insbesondere wird  $R$  in diesem Kapitel immer einen kommutativen Ring mit Eins und  $\mathbb{k}$  immer einen Körper bezeichnen.

### 4.1. Volumina und Determinantenfunktionen

In Bemerkung 1.70 (2) haben wir die Volumina von Parallelotopen im  $\mathbb{R}^3$  ausgerechnet. Im  $\mathbb{R}^n$  wollen wir entsprechend das  $n$ -dimensionale Volumen

$$\text{vol}(v_1, \dots, v_n)$$

eins von Vektoren  $v_1, \dots, v_n \in \mathbb{R}^n$  aufgespannten Parallelotops bestimmen. Wir möchten, dass dieser Volumenbegriff zwei Eigenschaften hat, nämlich *positive Homogenität* und *Scherungsinvarianz*: Für alle  $n$ -Tupel  $(v_1, \dots, v_n)$ , alle  $i, j \in \{1, \dots, n\}$  mit  $i \neq j$  und alle  $\ell \in \mathbb{R}$  soll gelten

$$(1) \quad \text{vol}(v_1, \dots, v_{i-1}, v_i \cdot \ell, v_{i+1}, \dots, v_n) = \text{vol}(v_1, \dots, v_n) \cdot |\ell| ;$$

$$(2) \quad \text{vol}(v_1, \dots, v_{i-1}, v_i + v_j \cdot \ell, v_{i+1}, \dots, v_n) = \text{vol}(v_1, \dots, v_n) .$$

Bedingung (2) lässt sich mit dem Cavalierischen Prinzip begründen: die Querschnitte von beiden Parallelotopen mit affinen Unterräumen parallel zu  $\langle v_1, \dots, \widehat{v}_i, \dots, v_n \rangle$  haben jeweils dasselbe Volumen, wenn man  $v_i$  um ein Vielfaches von  $v_j$  abändert.

Da für allgemeine Körper kein Absolutbetrag definiert ist, ist Bedingung (1) im allgemeinen nicht sinnvoll. Wir ersetzen sie daher durch Homogenität in

jedem einzelnen Argument und fordern die Eigenschaften

$$(1') \quad \omega(v_1, \dots, v_{i-1}, v_i \cdot \ell, v_{i+1}, \dots, v_n) = \omega(v_1, \dots, v_n) \cdot \ell,$$

$$(2) \quad \omega(v_1, \dots, v_{i-1}, v_i + v_j \cdot \ell, v_{i+1}, \dots, v_n) = \omega(v_1, \dots, v_n)$$

für ein „Volumen  $\omega(v_1, \dots, v_n)$  mit Vorzeichen“, wobei  $v_1, \dots, v_n \in \mathbb{k}^n$ . Falls  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  und  $\omega$  die Bedingungen (1') und (2) erfüllt, erfüllt  $\text{vol} = |\omega|$  die Bedingungen (1) und (2) und liefert daher einen Volumenbegriff.

Außerdem lassen sich aus den Bedingungen (1') und (2) zwei weitere Bedingungen ableiten, mit denen wir später besser arbeiten können. Zum einen gilt  $\omega(v_1, \dots, v_n) = 0$  falls  $v_i = v_j$  für zwei Indizes  $i \neq j$ . Dazu „scheren“ wir  $v_i$  um  $-v_j$  und erhalten 0 als  $i$ -tes Argument, was wegen Homogenität (2) impliziert, dass das Ergebnis 0 wird.

Zum anderen ist  $\omega$  in jedem Argument additiv (und daher wegen (2) in jedem Argument linear). Dazu betrachten wir  $\omega(v_1 + w, v_2, \dots, v_n)$  für  $v_1, \dots, v_n$  und  $w \in \mathbb{k}^n$  und unterscheiden zwei Fälle: wenn  $(v_1, \dots, v_n)$  linear abhängig sind, sei einer der Vektoren eine Linearkombination der restlichen, etwa  $v_1$ , dann folgt aus Scherungsinvarianz und Homogenität, dass

$$\omega(v_1, \dots, v_n) = \omega\left(\sum_{j=2}^n v_j \cdot \ell_j, v_2, \dots, v_n\right) = \omega(0, v_2, \dots, v_n) = 0.$$

Wenn  $(w, v_1, \dots, v_2)$  linear unabhängig sind, tauschen wir  $v_1$  und  $w$  und betrachten den zweiten Fall. Andernfalls überprüft man, dass auch  $(v_1 + w, v_2, \dots, v_n)$  linear abhängig sind, und erhält

$$\omega(v_1 + w, v_2, \dots, v_n) = 0 = \omega(v_1, \dots, v_n) + \omega(w, v_2, \dots, v_n).$$

Wir dürfen also annehmen, dass  $(v_1, \dots, v_n)$  eine Basis bilden, und schreiben

$$w = \sum_j v_j \cdot \ell_j.$$

Aus Scherungsinvarianz (2) und Homogenität (1') folgt

$$\begin{aligned} \omega(v_1 + w, v_2, \dots, v_n) &= \omega\left(v_1 + \sum_j v_j \cdot \ell_j, v_2, \dots, v_n\right) \\ &= \omega(v_1 \cdot (1 + \ell_1), v_2, \dots, v_n) \\ &= \omega(v_1, \dots, v_n) + \omega(v_1, \dots, v_n) \cdot \ell_1 \\ &= \omega(v_1, \dots, v_n) + \omega\left(\sum_j v_j \cdot \ell_j, v_2, \dots, v_n\right) \\ &= \omega(v_1, \dots, v_n) + \omega(w, v_2, \dots, v_n). \end{aligned}$$

Die Linearität in den anderen Argumenten ergibt sich genauso.

Soviel zur Motivation. Ab jetzt sei  $R$  ein kommutativer Ring mit Eins,  $M$  ein  $R$ -Modul und  $k \in \mathbb{N}$ .

4.1. DEFINITION. Es sei  $M$  ein  $R$ -Modul,  $k \in \mathbb{N}$ , und  $\alpha: M^k \rightarrow R$  eine Abbildung. Dann heißt  $\alpha$  *multilinear*, wenn für alle  $i \in \{1, \dots, k\}$  und alle  $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k \in M$  die Abbildung

$$(1) \quad M \rightarrow R \quad \text{mit} \quad w \mapsto \alpha(v_1, \dots, v_{i-1}, w, v_{i+1}, \dots, v_k)$$

linear ist. Sie heißt *alternierend* oder auch *alternierende Form*, wenn für alle  $i = 1, \dots, k-1$  gilt, dass

$$(2) \quad \alpha(v_1, \dots, v_k) = 0 \quad \text{falls } v_{i+1} = v_i.$$

Die Menge aller alternierenden multilinearen Abbildungen  $\alpha: M^k \rightarrow R$  wird mit  $\Lambda^k M^*$  bezeichnet. Falls  $V$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum ist, heißt eine alternierende multilineare Abbildung  $\omega: V^n \rightarrow \mathbb{k}$  eine *Determinantenfunktion*.

Man beachte, dass wir für  $k = 1$  gerade den dualen Modul  $\Lambda^1 M^* = M^*$  erhalten. Für  $k = 0$  setzt man sinnvollerweise  $\Lambda^0 M^* = R$ .

4.2. BEISPIEL. Wir betrachten das Spatprodukt  $\mathbb{R}^3 \rightarrow \mathbb{R}$  mit  $(x, y, z) \mapsto \langle x \times y, z \rangle$  aus Satz 1.69. Wegen Bemerkungen 1.53 (1) und 1.68 (1), (1') ist das Spatprodukt multilinear, und wegen Bemerkung 1.68 (2) und Satz 1.69 (1) ist es alternierend. Also ist das Spatprodukt eine Determinantenfunktion.

4.3. PROPOSITION. *Es sei  $M$  ein  $R$ -Modul und  $\alpha: M^k \rightarrow R$  multilinear. Dann sind die folgenden Aussagen äquivalent.*

- (1) *Die Abbildung  $\alpha$  ist alternierend,*
- (2) *Aus  $v_i = v_j$  für zwei Indizes  $i \neq j$  folgt  $\alpha(v_1, \dots, v_k) = 0$ .*

*Die Aussagen (1) und (2) implizieren die Eigenschaft*

- (3) *Die Abbildung  $\alpha$  ist antisymmetrisch, das heißt, für alle  $(v_1, \dots, v_k) \in M^k$  und alle  $i, j \in \{1, \dots, k\}$  mit  $i < j$  gilt*

$$\alpha(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_k) = -\alpha(v_1, \dots, v_k).$$

Wenn  $v_i = v_j$  für  $i \neq j$  gilt, folgt aus Behauptung (3) durch „Vertauschen“ der beiden gleichen Einträge zunächst nur  $\alpha(v_1, \dots, v_n) = -\alpha(v_1, \dots, v_n)$ , also  $2\alpha(v_1, \dots, v_n) = 0$ . Nur wenn wir in  $R$  durch 2 teilen dürfen, zum Beispiel in einem Körper der Charakteristik  $\chi(\mathbb{k}) \neq 2$ , ist (3) zu (1) und (2) äquivalent.

BEWEIS. Da wir Aussage (3) gleich brauchen, beginnen wir mit „(1)  $\implies$  (3)“ und betrachten zunächst den Fall  $j = i + 1$ . Dann folgt

$$\begin{aligned} \alpha(v_1, \dots, v_k) &= \alpha(v_1, \dots, v_k) + \underbrace{\alpha(v_1, \dots, v_i, v_i, v_{i+2}, \dots, v_k)}_{=0} \\ &= \alpha(v_1, \dots, v_i, v_i + v_{i+1}, v_{i+2}, \dots, v_k) \\ &\quad - \underbrace{\alpha(v_1, \dots, v_{i-1}, v_i + v_{i+1}, v_i + v_{i+1}, v_{i+2}, \dots, v_k)}_{=0} \end{aligned}$$

$$\begin{aligned}
&= \alpha(v_1, \dots, v_{i-1}, -v_{i+1}, v_i + v_{i+1}, v_{i+2}, \dots, v_k) \\
&\quad + \underbrace{\alpha(v_1, \dots, v_{i-1}, -v_{i+1}, -v_{i+1}, v_{i+2}, \dots, v_k)}_{=0} \\
&= -\alpha(v_1, \dots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \dots, v_k) .
\end{aligned}$$

Also ändert sich das Vorzeichen, wenn man zwei benachbarte Vektoren vertauscht. Der allgemeine Fall folgt durch Induktion über  $p = j - i$ , denn

$$\begin{aligned}
\alpha(v_1, \dots, v_k) &= -\alpha(v_1, \dots, v_{i-1}, v_{j-1}, v_{i+1}, \dots, v_{j-2}, v_i, v_j, v_{j+1}, \dots, v_k) \\
&= \alpha(v_1, \dots, v_{i-1}, v_{j-1}, v_{i+1}, \dots, v_{j-2}, v_j, v_i, v_{j+1}, \dots, v_k) \\
&= -\alpha(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-2}, v_{j-1}, v_i, v_{j+1}, \dots, v_k) .
\end{aligned}$$

Dabei haben wir nur Argumente im Abstand von weniger als  $p$  vertauscht.

Zu „(1)  $\implies$  (2)“ benutzen wir (3). Es gelte  $v_i = v_j$  für  $i + 1 < j$ , so dass die Behauptung nicht unmittelbar aus (1) folgt. Wegen (3) gilt dann

$$\begin{aligned}
\alpha(v_1, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) \\
= -\alpha(v_1, \dots, v_{i-1}, v_{j-1}, v_{i+1}, \dots, v_{j-2}, v_i, v_i, v_{j+1}, \dots, v_n) = 0 .
\end{aligned}$$

Die Richtung „(2)  $\implies$  (1)“ ist klar.  $\square$

4.4. BEMERKUNG. Wir haben in der Motivation von einem „Volumen mit Vorzeichen“ Homogenität (1') und Scherungsinvarianz gefordert. Daraus haben wir die Bedingungen an eine Determinantenfunktion in Definition 4.1 hergeleitet. Umgekehrt sind Determinantenfunktionen per definitionem homogen. Sie sind wegen Proposition 4.3 (2) auch scherungsinvariant, denn

$$\begin{aligned}
\alpha(v_1, \dots, v_{i-1}, v_i + v_j \cdot \ell, v_{i+1}, \dots, v_n) \\
= \alpha(v_1, \dots, v_n) + \alpha(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_n) \cdot \ell = \alpha(v_1, \dots, v_n) .
\end{aligned}$$

Also sind Determinantenfunktionen gerade „Volumina mit Vorzeichen“.

4.5. BEMERKUNG. Wir überlegen uns leicht, dass die Summe zweier Determinantenfunktionen und auch ein skalares Vielfaches einer Determinantenfunktion wieder eine solche ist. Also ist  $\Lambda^n M^*$  ein  $R$ -Modul. An dieser Stelle braucht man Kommutativität von  $R$ , siehe dazu die Bemerkung vor Definition 2.30. Aber man braucht Kommutativität von  $R$  bereits, um überhaupt multilineare Abbildungen mit zwei oder mehr Argumenten zu bekommen, wie die folgende Rechnung zeigt:

$$\begin{aligned}
\alpha(v_1, \dots, v_k) \cdot r \cdot s &= \alpha(v_1 \cdot r, v_2, \dots, v_k) \cdot s = \alpha(v_1 \cdot r, v_2 \cdot s, v_3, \dots, v_k) \\
&= \alpha(v_1, v_2 \cdot s, v_3, \dots, v_k) \cdot r = \alpha(v_1, \dots, v_k) \cdot s \cdot r .
\end{aligned}$$

Dass es überhaupt verschiedene Determinantenfunktionen auf demselben Modul oder Vektorraum gibt, sollte uns nicht erstaunen; schließlich kann man auch das Volumen im „uns umgebenden  $\mathbb{R}^3$ “ mit verschiedenen Volumenbegriffen messen — etwa in Litern, Kubikmetern, flüssigen Unzen, Fässern, etc.

Wir wollen jetzt für alle kommutativen Ringe  $R$  mit Eins ein spezielles Element  $\omega_n \in \Lambda^n(R^n)^*$ , die *Standard-Determinantenfunktion*, durch Induktion über  $n \in \mathbb{N}$  konstruieren. Für  $n = 1$  setzen wir  $\omega_0(r) = r \in R$  und sind fertig.

Sei  $\omega_{n-1}$  bereits konstruiert. Wir fassen Vektoren  $x \in R^n$  durch Weglassen der letzten Koordinate als Vektoren  $x' \in R^{n-1}$  auf, und nennen die letzte Koordinate  $\varepsilon_n(x)$ , siehe Bemerkung 2.65. Wir definieren  $\omega_n$  rekursiv durch

$$(*) \quad \omega_n(x_1, \dots, x_n) = \sum_{i=1}^n (-1)^{i+n} \varepsilon_n(x_i) \omega_{n-1}(x'_1, \dots, \widehat{x'_i}, \dots, x'_n) \in R,$$

wobei ein Dach über einem Eintrag gerade „Weglassen“ bedeutet. Diese Konstruktion liefert zugleich ein erstes Verfahren zur Berechnung von  $\omega_n$ , die Laplace-Entwicklung nach der letzten Zeile, siehe Satz 4.16 unten.

4.6. PROPOSITION. *Es sei  $R$  ein kommutativer Ring mit Eins, dann ist die oben konstruierte Abbildung  $\omega_n: R^n \rightarrow R$  alternierend, multilinear, und erfüllt*

$$\omega_n(e_1, \dots, e_n) = 1.$$

BEWEIS. Wir beweisen die Aussage wieder durch Induktion über  $n$ . Für  $n = 1$  ist die Behauptung klar.

Sei die Proposition für  $\omega_{n-1}$  bereits bewiesen. Linearität von  $\omega_n$  an der  $i$ -ten Stelle folgt für den  $i$ -ten Summand in  $(*)$  aus der Linearität von  $\varepsilon_n$ , für die restlichen Summanden aus der Multilinearität von  $\omega_{n-1}$ .

Sei jetzt  $x_{i+1} = x_i$  für ein  $i \in \{1, \dots, n-1\}$ . Dann sind der  $i$ -te und der  $(i+1)$ -te Summand in  $(*)$  bis auf das Vorzeichen gleich und heben sich weg, bei allen anderen Summanden werden zwei gleiche Vektoren nebeneinander in  $\omega_{n-1}$  eingesetzt, was nach Induktionsvoraussetzung 0 ergibt.

Außerdem ist  $\varepsilon_n(e_i) = 0$  für  $i < n$ , und die Vektoren  $e'_i$  für  $i < n$  sind gerade die Standardbasisvektoren des  $R^{n-1}$ . Also gilt

$$\omega_n(e_1, \dots, e_n) = (-1)^{n+n} \varepsilon_n(e_n) \omega_{n-1}(e'_1, \dots, e'_{n-1}) = 1. \quad \square$$

Als nächstes überlegen wir uns, dass der Raum  $\Lambda^n V^*$  genau eindimensional ist. Zunächst erinnern wir uns an die Automorphismengruppe  $\text{Aut}(M)$  einer Menge aus Beispiel 2.5. Es sei  $\omega_n$  die obige Determinantenfunktion.

4.7. DEFINITION. Es sei  $n \in \mathbb{N}$ . Die *symmetrische Gruppe  $S_n$  in  $n$  Elementen* ist definiert als  $S_n = \text{Aut}(\{1, \dots, n\})$ , ihre Elemente  $S_n \ni \sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  heißen *Permutationen*.

Es sei  $R$  ein kommutativer Ring mit Eins. Wir definieren das *Vorzeichen* oder auch *Signum* einer Permutation  $\sigma \in S_n$  durch

$$\text{sign}(\sigma) = \omega_n(e_{\sigma(1)}, \dots, e_{\sigma(n)}).$$

Unter einer *Transposition* verstehen wir eine Permutation  $\tau = \tau_{ij} \in S_n$ , die nur zwei Elemente  $i$  und  $j$  mit  $1 \leq i < j \leq n$  vertauscht, also

$$\tau_{ij}(k) = \begin{cases} j & \text{falls } k = i, \\ i & \text{falls } k = j, \text{ und} \\ k & \text{sonst.} \end{cases}$$

4.8. PROPOSITION. Jede Permutation  $\sigma \in S_n$  kann als Produkt  $\sigma = \tau_1 \circ \dots \circ \tau_k$  von Transpositionen  $\tau_1, \dots, \tau_k \in S_n$  geschrieben werden, und es gilt

$$(1) \quad \text{sign}(\sigma) = (-1)^k .$$

Für  $\rho, \sigma \in S_n$  gilt dann

$$(2) \quad \text{sign}(\rho \circ \sigma) = \text{sign}(\rho) \cdot \text{sign}(\sigma) \quad \text{und} \quad \text{sign}(\sigma^{-1}) = \text{sign}(\sigma) .$$

Dabei fassen wir die Identität als „leeres Produkt“ mit  $k = 0$  auf. Im allgemeinen sind weder  $k$  noch  $\tau_1, \dots, \tau_k$  durch  $\sigma$  eindeutig bestimmt, allein das Vorzeichen ist eine Invariante. Nach (1) gilt stets  $\text{sign}(\sigma) \in \{1, -1\}$ , unabhängig vom Ring  $R$ . Allerdings könnte  $1 = -1$  in  $R$  gelten (Beispiel:  $R = \mathbb{Z}/2\mathbb{Z}$ ); in diesem Fall verliert das Vorzeichen seine Information.

BEWEIS. Wir beweisen die erste Aussage durch Induktion über  $n$ . Für  $n = 1$  gibt es nur eine Permutation, die Identität, mit

$$\text{sign}(\text{id}_{\{1\}}) = \omega_1(e_1) = 1 .$$

Sei die Aussage für alle  $\sigma' \in S_{n-1}$  bewiesen, und sei  $\sigma \in S_n$ . Falls  $\sigma(n) = n$ , sei  $\sigma' = \sigma|_{\{1, \dots, n-1\}} \in S_{n-1}$ . Da  $\sigma'$  ein Produkt von Transpositionen aus  $S_{n-1}$  ist, ist  $\sigma$  das Produkt von Transpositionen aus  $S_n$ , die jeweils die gleichen Elemente vertauschen. Falls  $\sigma(n) \neq n$ , sei  $\tau$  die Transposition, die  $\sigma(n)$  und  $n$  vertauscht, so dass

$$(\tau \circ \sigma)(n) = n .$$

Nach dem obigen Argument ist  $\tau \circ \sigma$  ein Produkt von Transpositionen  $\tau_1 \circ \dots \circ \tau_k$ . Da  $\tau = \tau^{-1}$ , folgt

$$\sigma = \tau \circ \tau_1 \circ \dots \circ \tau_k .$$

Wir beweisen (1) für  $\sigma = \tau_1 \circ \dots \circ \tau_k = \sigma' \circ \tau_k$  durch Induktion über  $k$ . Der Induktionsanfang für  $k = 0$  ist klar, denn  $\text{sign}(\text{id}) = \omega_n(e_1, \dots, e_n) = 1$  nach Propositionen 4.6. Angenommen,  $\sigma = \sigma' \circ \tau_k$ , und  $\tau_k$  vertausche  $i$  und  $j$ . Aus Proposition 4.3 (3) folgt dann

$$\begin{aligned} \text{sign}(\sigma) &= \omega_n(e_{\sigma'(1)}, \dots, \underbrace{e_{\sigma'(j)}}_i, \dots, \underbrace{e_{\sigma'(i)}}_j, \dots, e_{\sigma'(n)}) \\ &= -\omega_n(e_{\sigma'(1)}, \dots, e_{\sigma'(n)}) = -(-1)^{k-1} = (-1)^k . \end{aligned}$$

Zu (2) stellen wir  $\rho$  und  $\sigma$  als Produkte von  $j$  und  $k$  Transpositionen dar, dann erhalten wir eine Darstellung von  $\rho \circ \sigma$  als Produkt von  $j + k$  Transpositionen, und die erste Behauptung folgt. Die letzte ergibt sich dann aus

$$\text{sign}(\sigma) \cdot \text{sign}(\sigma^{-1}) = \text{sign}(\sigma \cdot \sigma^{-1}) = \text{sign}(\text{id}) = 1 . \quad \square$$

In der folgenden Proposition zeigen wir die Eindeutigkeit einer bestimmten Determinantenfunktion. Der Beweis liefert uns eine zweite Berechnungsmethode der Standard-Determinantenfunktion  $\omega_n$ , die sogenannte Leibniz-Formel, siehe Satz 4.13 unten.

4.9. PROPOSITION. *Es sei  $R$  ein kommutativer Ring mit Eins,  $r \in R$  und  $M$  ein freier  $R$ -Modul mit Basis  $B = (b_1, \dots, b_n)$ . Dann existiert genau eine Determinantenfunktion  $\omega \in \Lambda^n M^*$  mit*

$$\omega(b_1, \dots, b_n) = r .$$

Sei  $\omega_B$  die obige Determinantenfunktion zu  $r = 1$ , dann ist  $\Lambda^n M^*$  ein freier  $R$ -Modul mit Basis  $(\omega_B)$ .

BEWEIS. Zur Eindeutigkeit bestimmen wir den Wert von  $\omega(v_1, \dots, v_n)$  für Modulelemente

$$v_j = \sum_{i=1}^n b_i \cdot a_{ij} \in R^n \quad \text{für } j = 1, \dots, n .$$

Als erstes schließen wir aus Multilinearität, dass

$$\begin{aligned} \omega(v_1, \dots, v_n) &= \omega\left(\sum_{i=1}^n b_i \cdot a_{i1}, \dots, \sum_{i=1}^n b_i \cdot a_{in}\right) \\ &= \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \omega(b_{i_1}, \dots, b_{i_n}) \cdot a_{i_1 1} \cdots a_{i_n n} . \end{aligned}$$

Als nächstes dürfen wir wegen 4.3 (2) wir alle Summanden weglassen, bei denen  $i_j = i_k$  für  $j \neq k$ . Somit ist die Abbildung von der Menge  $\{1, \dots, n\}$  in sich mit  $j \mapsto i_j$  injektiv, und da die Menge endlich ist, auch surjektiv. Wir können die Indizes  $i_1, \dots, i_n$  also durch eine Permutation beschreiben und erhalten

$$\omega(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \omega(b_{\sigma(1)}, \dots, b_{\sigma(n)}) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n} .$$

Nach Proposition 4.8 können wir  $\sigma$  als Produkt von  $k$  Transpositionen schreiben. Wie im Beweis von Proposition 4.8 (1) geht  $\omega(b_{\sigma(1)}, \dots, b_{\sigma(n)})$  aus  $\omega(b_1, \dots, b_n)$  hervor, indem man  $k$ -fach je zwei Argumente vertauscht. Wegen Proposition 4.3 (3) ist  $\omega$  eindeutig bestimmt durch

$$\begin{aligned} \omega(v_1, \dots, v_n) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \omega(b_1, \dots, b_n) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\ (*) \quad &= \sum_{\sigma \in S_n} \text{sign}(\sigma) r \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n} . \end{aligned}$$

Es sei  $B: R^n \rightarrow M$  die Basisabbildung, siehe Bemerkung 2.74, und es sei wieder  $v_j = B(a_j)$  mit  $a_j = (a_{ij})_i \in R^n$ . Wir definieren zunächst eine alternierende multilineare Abbildung  $\omega_B$  mit

$$\omega_B(v_1, \dots, v_n) = \omega_n(a_1, \dots, a_n) , \quad \text{so dass} \quad \omega_B(b_1, \dots, b_n) = 1 .$$

Nach Proposition 4.6 ist  $\omega_n$  multilinear und alternierend, also auch  $\omega_B$ . Wir erhalten die gesuchte Form  $\omega$  als

$$\omega = \omega_B \cdot r = \omega_B \cdot \omega(b_1, \dots, b_n).$$

Aufgrund der obigen Eindeutigkeitsaussage sind alle  $\omega \in \Lambda^n M^*$  von dieser Gestalt, also bildet  $(\omega_B)$  eine Basis von  $\Lambda^n M^*$ .  $\square$

## 4.2. Die Determinante

Ausgehend von den Überlegungen im letzten Kapitel führen wir jetzt Determinanten von Endomorphismen und quadratischen Matrizen ein. Während Determinantenfunktionen dazu dienen, Volumina von Parallelotopen in einem Vektorraum zu beschreiben, misst die Determinante, um welchen Faktor ein Endomorphismus das Volumen einzelner Parallelotope vergrößert oder verkleinert.

4.10. BEMERKUNG. Es seien  $M, N$  Moduln über  $R$  und  $F: M \rightarrow N$  linear, dann definieren wir für alle  $k$  eine Abbildung  $F^*: \Lambda^k N^* \rightarrow \Lambda^k M^*$  durch

$$(1) \quad (F^*(\alpha))(v_1, \dots, v_k) = \alpha(F(v_1), \dots, F(v_k))$$

für alle  $\alpha \in \Lambda^k N^*$  und alle  $v_1, \dots, v_k$ . Die rechte Seite ist sinnvoll, da wir  $\alpha$  auf  $k$  Elemente  $F(v_1), \dots, F(v_k)$  von  $N$  anwenden, und entsprechend erhalten wir eine Abbildung  $F^*(\alpha): M^k \rightarrow R$ . Man nennt  $F^*(\alpha)$  auch die mit  $F$  zurückgeholte Form.

Wir zeigen, dass  $F^*(\alpha)$  im ersten Argument linear ist; für die anderen Argumente zeigt man Linearität genauso. Es seien  $x, y$  und  $v_2, \dots, v_k \in M$  und  $r, s \in R$ , dann folgt

$$\begin{aligned} (F^*(\alpha))(x \cdot r + y \cdot s, v_2, \dots, v_k) &= \alpha(F(x \cdot r + y \cdot s), F(v_2), \dots, F(v_k)) \\ &= \alpha(F(x) \cdot r + F(y) \cdot s, F(v_2), \dots, F(v_k)) \\ &= \alpha(F(x), F(v_2), \dots, F(v_k)) \cdot r + \alpha(F(y), F(v_2), \dots, F(v_k)) \cdot s \\ &= (F^*(\alpha))(x, v_2, \dots, v_k) \cdot r + (F^*(\alpha))(y, v_2, \dots, v_k) \cdot s. \end{aligned}$$

Also ist  $F^*(\alpha)$  multilinear.

Und  $F^*(\alpha)$  ist auch alternierend, denn

$$(F^*(\alpha))(v_1, \dots, v_i, v_i, \dots, v_k) = \alpha(F(v_1), \dots, F(v_i), F(v_i), \dots, F(v_k)) = 0.$$

Es folgt  $F^*(\alpha) \in \Lambda^k M^*$  wie behauptet.

In Bemerkung 4.5 haben wir uns überlegt, dass  $\Lambda^k M^*$  und  $\Lambda^k N^*$  Moduln über  $R$  sind. Die Abbildung  $F^*: \Lambda^k N^* \rightarrow \Lambda^k M^*$  ist linear, denn für alle  $\alpha,$

$\beta \in \Lambda^k N^*$ , alle  $r, s \in R$  und alle  $v_1, \dots, v_k \in M$  gilt

$$\begin{aligned}
 (2) \quad & (F^*(\alpha \cdot r + \beta \cdot s))(v_1, \dots, v_k) \\
 &= (\alpha \cdot r + \beta \cdot s)(F(v_1), \dots, F(v_k)) \\
 &= \alpha(F(v_1), \dots, F(v_k)) \cdot r + \beta(F(v_1), \dots, F(v_k)) \cdot s \\
 &= (F^*(\alpha) \cdot r + F^*(\beta) \cdot s)(v_1, \dots, v_k) .
 \end{aligned}$$

Schließlich seien  $F: M \rightarrow N$  und  $G: L \rightarrow M$  lineare Abbildungen, dann gilt  $(F \circ G)^* = G^* \circ F^*: \Lambda^k N^* \rightarrow \Lambda^k L^*$ , denn

$$\begin{aligned}
 (3) \quad & ((F \circ G)^*(\alpha))(\ell_1, \dots, \ell_k) = \alpha(F(G(\ell_1)), \dots, F(G(\ell_k))) \\
 &= (F^*(\alpha))(G(\ell_1), \dots, G(\ell_k)) = (G^*(F^*(\alpha)))(\ell_1, \dots, \ell_k) .
 \end{aligned}$$

Man beachte, dass Zurückholen die Reihenfolge der beteiligten Abbildungen vertauscht.

Es sei  $M$  ein freier  $R$ -Modul mit Basis  $B = (b_1, \dots, b_n)$ . In Proposition 4.9 haben wir gesehen, dass  $\Lambda^n M^* \cong R$  ein freier Modul mit einelementiger Basis ( $\omega_B$ ) ist. Dabei ist  $\omega_B \in \Lambda^n M^*$  das Element mit  $\omega_B(b_1, \dots, b_n) = 1$ . Sei jetzt  $F \in \text{End}_R(M)$ , dann ist  $F^* \in \text{End}_R(\Lambda^n M^*)$  nach der obigen Bemerkung, aber  $\text{End}_R(\Lambda^n M^*) \cong R$ , da  $\Lambda^n V^* \cong R$ . Also existiert zu jedem  $F \in \text{End } M$  ein Skalar  $a = \det F \in R$ , so dass

$$F^* \omega = \omega \cdot a \quad \text{für alle } \omega \in \Lambda^n M^* .$$

Um  $a$  zu bestimmen, wählen wir eine Basis  $(b_1, \dots, b_n)$ , definieren  $\omega_B$  wie in Proposition 4.9, und überlegen uns, dass

$$\begin{aligned}
 \omega_B(F(b_1), \dots, F(b_n)) &= (F^*(\omega_B))(b_1, \dots, b_n) \\
 &= (\omega_B \cdot a)(b_1, \dots, b_n) = (\omega_B)(b_1, \dots, b_n) \cdot a = a .
 \end{aligned}$$

Im Spezialfall  $M = R^n$  mit der Standardbasis sind die Vektoren  $F(e_1), \dots, F(e_n)$  nach Folgerung 2.75 genau die Spalten der Abbildungsmatrix  $A \in M_n(R)$  von  $F$ , und  $\omega_B$  ist gerade die Standarddeterminantenfunktion  $\omega_n$  aus Proposition 4.6. Das motiviert die folgende Definition.

**4.11. DEFINITION.** Es sei  $R$  ein kommutativer Ring mit Eins,  $M$  ein freier  $R$ -Modul mit einer  $n$ -elementigen Basis, wobei  $n \geq 1$ , und  $F \in \text{End}_R(M)$  ein Endomorphismus. Dann ist die *Determinante* von  $F$  der eindeutige Skalar  $\det F \in R$ , so dass

$$(1) \quad F^* \omega = \omega \cdot \det F \quad \text{für alle } \omega \in \Lambda^n M^* .$$

Wir definieren die *Determinante* einer Matrix  $A \in M_n(R)$  mit den Spalten  $a_1, \dots, a_n \in R^n$  durch

$$(2) \quad \det A = \omega_n(a_1, \dots, a_n) .$$

Im Falle  $n = 0$  folgt  $\det() = 1$ , da  $\omega_0() = 1$ . In Gleichung (1) haben wir für jeden Endomorphismus  $F \in \text{End}_R M$  die Determinante definiert, ohne eine Basis fixiert und  $F$  als Matrix geschrieben zu haben; diese Definition ist also *basisunabhängig*. Wenn wir eine Basis  $B$  wählen und  $A$  die Abbildungsmatrix von  $F$  bezüglich der Basis  $B$  (sowohl vom Definitions- als auch vom Wertebereich) darstellen, ist die Determinante von  $A$  durch (2) definiert. Unsere obige Vorüberlegung besagt, dass

$$\det A = \det F .$$

Auf diese Weise hängen (1) und (2) zusammen. Wichtig ist dabei immer, dass wir nur Determinanten von Endomorphismen definieren können. Abbildungen zwischen verschiedenen Vektorräumen haben keine wohldefinierte Determinante, es sei denn, wir geben auf beiden Räumen eine Basis vor — nur in diesem Fall können wir überhaupt Volumina vergleichen.

Unsere Definition der Determinante auf dem Umweg über das Zurückziehen von Determinantenfunktionen hat Vorteile: sie ist basisunabhängig und erlaubt es uns, relativ einfach die Multiplikativität der Determinante zu verstehen.

4.12. SATZ. *Sei  $V$  ein freier  $R$ -Modul mit einer  $n$ -elementigen Basis, und es seien  $F, G \in \text{End } V$ , dann gilt*

$$(1) \quad \det(F \circ G) = \det F \cdot \det G ,$$

*d.h., die Determinante ist multiplikativ. Für Matrizen  $A, B \in M_n(R)$  gilt entsprechend*

$$(2) \quad \det(A \cdot B) = \det A \cdot \det B .$$

BEWEIS. Die Multiplikativität von  $\det$  über einem Körper  $\mathbb{k}$  folgt direkt aus der Kompositionsregel in Bemerkung 4.10 (3), denn für alle  $\omega \in \Lambda^n V^*$  gilt

$$\omega \cdot \det(F \circ G) = (F \circ G)^* \omega = G^* \circ F^* \omega = F^* \omega \cdot \det G = \omega \cdot \det G \cdot \det F .$$

Indem wir  $\omega = \omega_n \neq 0$  wählen, folgt (1). Sei  $F \in \text{Aut } V$ , dann ist  $F$  invertierbar nach Definition 2.30, also existiert eine Umkehrabbildung  $F^{-1}$  mit

$$\det F \cdot \det F^{-1} = \det(F \circ F^{-1}) = \det(\text{id}_V) = 1 .$$

Insbesondere folgt  $\det F \in \mathbb{k}^\times = \mathbb{k} \setminus \{0\}$  mit  $(\det F)^{-1} = \det(F^{-1})$ , und außerdem ist  $\det: \text{Aut } V \rightarrow \mathbb{k}^\times$  ein Gruppenhomomorphismus.

Wir erhalten (2) als Spezialfall für  $M = R^n$ , da  $\text{End}_R M = M_n(R)$ . □

4.13. SATZ (Leibniz-Formel). *Für jede Matrix  $A \in M_n(R)$  mit  $n \geq 1$  gilt*

$$\det A = \sum_{\sigma \in S(n)} \text{sign}(\sigma) \cdot \prod_{j=1}^n a_{\sigma(j), j} = \sum_{\rho \in S(n)} \text{sign}(\rho) \cdot \prod_{i=1}^n a_{i, \rho(i)} .$$

BEWEIS. Die erste Formel ist (\*) aus dem Beweis von Proposition 4.9. Sei  $\rho$  die Umkehrabbildung von  $\sigma$ , dann erhalten wir die zweite Formel, indem wir  $j$  durch  $\rho(i)$  ersetzen. □

4.14. BEMERKUNG. Für  $n = 2$  gibt es genau zwei Permutationen, die Identität und die Transposition  $\tau_{12}$ . Wir erhalten die einfache Formel

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} a_{22} - a_{12} a_{21} .$$

Für  $n = 3$  gibt es schon sechs Permutationen. Mit Hilfe des Schemas

$$\begin{array}{cccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} & \\ & \diagdown & \diagup & \diagdown & \diagup & \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} & \\ & \diagup & \diagdown & \diagup & \diagdown & \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} & \end{array}$$

ergibt sich aus der Leibniz-Formel die *Sarrussche Regel* für  $3 \times 3$ -Matrizen:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} \\ - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} - a_{13} a_{22} a_{31} .$$

Hierbei werden die Elemente entlang der drei durchgezogenen Linien jeweils aufmultipliziert und zusammenaddiert, und die Elemente entlang der unterbrochenen Linien werden ebenfalls aufmultipliziert und danach subtrahiert.

Für  $n = 4$  gibt es bereits  $4! = 24$  Permutationen; zuviele, um die Leibniz-Formel durch ein einprägsames Rechenschema darzustellen. Zur Berechnung größerer Determinanten ist die Leibniz-Formel nicht zu empfehlen (Übung). Sie erlaubt aber einige interessante Schlussfolgerungen.

4.15. FOLGERUNG. *Es sei  $R$  ein kommutativer Ring mit Eins und  $A \in M_n(R)$ .*

- (1) *Es gilt  $\det(A^t) = \det A$ .*
- (2) *Die Determinante  $\det A$  ist multilinear und alternierend in den Zeilen der Matrix  $A$ .*
- (3) *Sei  $R = \mathbb{k}$  ein Körper, dann verschwindet die Determinante  $\det A$ , wenn die Zeilen von  $A$  linear abhängig sind.*
- (4) *Die Determinante  $\det A$  ändert sich nicht, wenn man ein Vielfaches einer Zeile zu einer anderen dazuaddiert.*
- (5) *Die Determinante  $\det A$  wechselt das Vorzeichen, wenn man zwei Zeilen vertauscht.*

BEWEIS. Aussage (1) ergibt sich durch Vergleich der Leibniz-Formeln aus Satz 4.13 für  $A$  und  $A^t$ , denn

$$\det(A^t) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \prod_{i=1}^n a_{\sigma(i),i} = \det(A) .$$

Jetzt folgen (2)–(5) für  $A$  jeweils aus Definition 4.1 und Proposition 4.3, angewandt auf die Matrix  $A^t$ .  $\square$

Wir wollen jetzt ein etwas effizienteres Verfahren zum Berechnen von Determinanten kennenlernen, genauer gesagt, eine Verallgemeinerung der Formel (\*) für  $\omega_n$ . Sei  $A = (a_{ij})_{i,j} \in M_n(R)$  eine Matrix, dann bezeichnen wir die Matrix  $A$  ohne die  $i$ -te Zeile und die  $j$ -te Spalte mit

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix} \in M_{n-1}(R).$$

4.16. SATZ (Laplace-Entwicklung). *Es sei  $A \in M_n(R)$  mit  $n \geq 1$ . Entwicklung nach der  $i$ -ten Zeile. Für alle  $i \in \{1, \dots, n\}$  gilt*

$$(1) \quad \det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \cdot \det(A_{ij}).$$

Entwicklung nach der  $j$ -ten Spalte. Für alle  $j \in \{1, \dots, n\}$  gilt

$$(2) \quad \det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \cdot \det(A_{ij}).$$

BEWEIS. Wir betrachten die durch die rechte Seite von Formel (1) induktiv definierte Abbildung  $\omega: M_n(R) \rightarrow R$  und zeigen wie im Beweis von Proposition 4.6, dass sie multilinear und alternierend in den Spalten von  $A$  ist. Aufgrund der Eindeutigkeitsaussage in Proposition 4.9 und Definition 4.11 reicht es zu zeigen, dass die rechte Seite für die Einheitsmatrix den Wert 1 annimmt, um die Behauptung (1) zu beweisen.

Es sei  $n \geq 1$  und  $i \in \{1, \dots, n\}$ , und  $\omega(A)$  bezeichne die rechte Seite von (1). Linearität von  $\omega$  an der  $k$ -ten Stelle folgt für den  $k$ -ten Summand, da  $a_{ik}$  linear von  $a_k \in R^n$  abhängt. Für die restlichen Summanden folgt sie, da  $\det A_{ij}$  multilinear in den Spalten von  $A_{ij}$  ist.

Sei jetzt  $a_{k+1} = a_k$  für ein  $k \in \{1, \dots, n-1\}$ . Dann sind der  $k$ -te und der  $(k+1)$ -te Summand in (1) bis auf das Vorzeichen gleich und heben sich weg; bei allen anderen Summanden stimmen zwei benachbarte Spalten von  $A_{ij}$  überein, so dass  $\det(A_{ij}) = 0$ . Also ist  $\omega$  multilinear und alternierend.

Für die Einheitsmatrix erhalten wir

$$\omega(E_n) = \sum_{j=1}^n (-1)^{i+j} \delta_{ij} \cdot \det((E_n)_{ij}) = \det(E_{n-1}) = 1,$$

da nur der Summand mit  $i = j$  beiträgt, und da nach Streichen der  $i$ -ten Spalte und Zeile aus der Einheitsmatrix  $E_n$  die Einheitsmatrix  $E_{n-1}$  wird. Damit ist (1) bewiesen.

Wir beweisen (2), indem wir die Transponierte  $A^t$  in (1) einsetzen und Folgerung 4.15 benutzen.  $\square$

4.17. DEFINITION. Eine Matrix  $A = (a_{ij})_{i,j} \in M_n(\mathbb{k})$  heißt *in oberer (unterer) Dreiecksgestalt*, oder kurz *obere (untere) Dreiecksmatrix*, wenn  $a_{ij} = 0$  für alle  $i, j \in \{1, \dots, n\}$  mit  $i > j$  ( $i < j$ ). Eine Matrix heißt *in strikter oberer/unterer Dreiecksgestalt*, wenn zusätzlich  $a_{ii} = 0$  für alle  $i \in \{1, \dots, n\}$ .

Somit ist die linke Matrix unten eine obere Dreiecksmatrix, und die rechte sogar in strikter Dreiecksgestalt:

$$\begin{pmatrix} a_{11} & & \dots & a_{1n} \\ 0 & a_{22} & & \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ & & & a_{n-1,n} \\ 0 & \dots & & 0 \end{pmatrix}.$$

Außerdem erinnern wir uns an Blockmatrizen, siehe Satz 3.13.

4.18. FOLGERUNG. *Es sei  $R$  ein kommutativer Ring mit Eins.*

- (1) *Seien  $A \in M_k(R)$ ,  $B \in M_{k,\ell}(R)$ ,  $C \in M_{\ell,k}(R)$  und  $D \in M_{\ell,\ell}(R)$ . Dann gilt*

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

$$\text{und} \quad \det \begin{pmatrix} B & A \\ D & 0 \end{pmatrix} = (-1)^{k\ell} \det(A) \cdot \det(D) = \det \begin{pmatrix} 0 & A \\ D & C \end{pmatrix}.$$

- (2) *Es sei  $A$  eine obere oder untere Dreiecksmatrix, dann gilt*

$$\det(A) = \prod_{i=1}^n a_{ii}.$$

BEWEIS. Es reicht, eine der vier Gleichungen in (1) zu beweisen. Die anderen lassen sich daraus ableiten, indem man die ersten  $k$  Zeilen mit den letzten  $\ell$  vertauscht, und/oder die ersten  $k$  Spalten mit den letzten  $\ell$ . Die zugehörigen Permutationen haben jeweils das Vorzeichen  $(-1)^{k\ell}$ , was auch die Vorzeichen in der zweiten Zeile erklärt.

Es sei also  $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ . Wir zeigen durch Induktion über  $k$ , dass  $\det(M) = \det(A) \cdot \det(D)$ . Als Induktionsanfang wählen wir  $k = 0$ , so dass  $M = A$ , und vereinbaren, dass  $\det() = 1$ .

Für den Induktionsschritt entwickeln wir nach der ersten Spalte und erhalten

$$\det M = \sum_{i=1}^{k+\ell} (-1)^{i+1} m_{i1} \det(M_{i1}) = \sum_{i=1}^k (-1)^{i+1} a_{i1} \det(M_{i1}),$$

da  $m_{i1} = 0$  für  $i > k$ . Für  $1 \leq i \leq k$  erhalten wir eine Matrix der Form

$$M_{i1} = \begin{pmatrix} A_{i1} & * \\ 0 & D \end{pmatrix}.$$

Nach Induktionsvoraussetzung gilt  $\det(M_{i1}) = \det(A_{i1}) \cdot \det(D)$ . Das heißt, der obere rechte Block in  $M_{i1}$  trägt nicht zur Determinanten bei. Also erhalten wir

$$\det M = \left( \sum_{i=1}^k (-1)^{i+1} a_{i1} \det(A_{i1}) \right) \cdot \det(D) = \det(A) \cdot \det(D),$$

wobei wir im letzten Schritt die Laplace-Entwicklung der Matrix  $A$  benutzt haben.

Auch Behauptung (2) beweisen wir nur für obere Dreiecksmatrizen, und zwar induktiv über  $n$  mit Hilfe von (1) für  $k = 1$ . Diesmal trägt nur  $i = 1$  bei, und die Matrix  $A_{11}$  ist wieder eine obere Dreiecksmatrix. Induktiv folgt

$$\det A = a_{11} \cdot \det(A_{11}) = \prod_{i=1}^n a_{ii}. \quad \square$$

4.19. BEMERKUNG. In Folgerung 4.15 haben wir gesehen, wie sich die Determinante unter Zeilenumformungen verhält, also können wir Determinanten jetzt auch mit dem Gauß-Verfahren aus Satz 3.26 berechnen. Wegen Folgerung 4.18 müssen wir unsere Matrix nicht auf strenge Zeilenstufenform bringen; es reicht obere Dreiecksgestalt.

Wir modifizieren das im Beweis von Satz 3.26 beschriebene Verfahren, angewandt auf eine Matrix  $A$ , wie folgt. Wir beginnen mit einem Vorfaktor  $a_0 = 1$  und erhalten nach dem  $r$ -ten Schritt

$$\begin{aligned} \det A &= \dots = a_r \cdot \det \begin{pmatrix} 1 & a_{12} & & \dots & & a_{1,n} \\ 0 & \ddots & \ddots & & & \vdots \\ & \ddots & 1 & a_{r,r+1} & \dots & a_{r,n} \\ \vdots & & 0 & a_{r+1,r+1} & \dots & a_{r+1,n} \\ & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{n,r+1} & \dots & a_{n,n} \end{pmatrix} \\ &= a_r \cdot \det \begin{pmatrix} a_{r+1,r+1} & \dots & a_{r+1,n} \\ \vdots & & \vdots \\ a_{n,r+1} & \dots & a_{n,n} \end{pmatrix} \end{aligned}$$

Hierbei haben wir Folgerung 4.18 (1) und (2) ausgenutzt und geschlossen, dass nur der untere rechte Block einen Beitrag leistet. Sie müssen also bei einer größeren Matrix gegen Ende des Verfahrens nicht mehr die ganze Matrix mit-schleppen.

Falls wir im Laufe des Verfahrens eine Spalte überspringen („1. Fall“ im Beweis von Satz 3.26) ist am Ende des Verfahrens die letzte Zeile 0, folgt  $\det A$  aus Folgerung 4.15 (3), und wir können das Gauß-Verfahren an dieser Stelle abbrechen. Genauso sind wir beim Invertieren in Bemerkung 3.28 (4) verfahren.

Ansonsten ändern wir dann, wenn wir im ersten Schritt tauschen müssen, das Vorzeichen der Vorfaktors wegen Folgerung 4.15 (5). Vor dem Normieren multiplizieren wir den Vorfaktor mit  $a_{rr}$  und erhalten  $a_r = a_{rr} a_{r-1}$  wegen

Folgerung 4.15 (2). Anschließend räumen wir unterhalb der aktuellen Zeile aus, wobei sich der Vorfaktor wegen Folgerung 4.15 (4) nicht ändert.

Am Schluss des Verfahrens erhalten wir einen Vorfaktor  $a_n$ , multipliziert mit der Determinante einer oberen Dreiecksmatrix mit Einsen auf der Diagonalen (also  $a_{11} = \cdots = a_{nn} = 1$ ). Nach Folgerung 4.18 ist diese Determinante 1, also ist  $a_n$  die Determinante der ursprünglichen Matrix.

Wir betrachten den Rechenaufwand. Dabei zählen wir nur Multiplikationen und Divisionen, da die anderen Rechenschritte weniger aufwändig sind.

- (1) In der Leibniz-Formel summieren wir über  $n!$  Permutationen, und bei jeder müssen wir  $(n-1)$ -mal multiplizieren. Der Aufwand ist insgesamt mit  $n! \cdot (n-1)$  Multiplikationen sehr hoch.
- (2) Bei der Laplace-Entwicklung benötigt man eigentlich mehr als  $n!$  Multiplikationen. Wenn wir alle Determinanten von Untermatrizen zwischenspeichern, reichen  $n(2^{n-1} - 1)$  Multiplikationen, der Speicheraufwand ist aber ähnlich hoch. Dafür wird das Verfahren deutlich effizienter, wenn die Matrix viele Nullen enthält.
- (3) Das Gaußverfahren kommt mit  $\frac{n^3-3}{3} + n - 1$  Multiplikationen und Divisionen aus. Man kann es etwas abkürzen, indem man die Determinante der letzten  $2 \times 2$ - oder  $3 \times 3$ -Matrix mit Laplace ausrechnet — das spart genau einen Rechenschritt.

Für kleine  $n$  ergeben sich folgende Zahlen:

$n$	2	3	4	5	6	...	10
Leibniz	2	12	72	480	3 600	...	32 659 200
Laplace	2	9	28	75	186	...	5 110
Gauß	2	9	22	43	74	...	338

4.20. DEFINITION. Es sei  $A \in M_n(R)$ . Die *Adjunkte* von  $A$  ist definiert als

$$\text{adj } A = \left( (-1)^{i+j} \det(A_{ji}) \right)_{i,j} \in M_n(R).$$

Trotz der ähnlichen Namen hat die Adjunkte nichts mit der adjungierten Matrix aus Definition 3.29 zu tun. Die nächste Folgerung ergibt sich aus dem Laplaceschen Entwicklungssatz.

4.21. FOLGERUNG (Cramersche Regeln). *Es sei  $R$  ein kommutativer Ring mit Eins. Eine Matrix  $A \in M_n(R)$  ist genau dann invertierbar, wenn*

$$\det A \in R^\times = \{ r \in R \mid \text{es gibt ein } s \in R \text{ mit } rs = 1 \},$$

und in diesem Fall gilt

$$(1) \quad A^{-1} = (\det A)^{-1} \text{adj } A.$$

Wenn  $\det A \in R^\times$ , ist das Gleichungssystem  $A \cdot x = b$  für alle  $b \in R^n$  eindeutig lösbar mit

$$(2) \quad x_i = \frac{\det A_i}{\det A}, \quad \text{wobei } A_i = (a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n) \in M_n(R).$$

Wir nennen  $R^\times$  auch die *Einheitengruppe* oder *multiplikative Gruppe* von  $R$ , und ihre Elemente *Einheiten*. Wegen Satz 4.12 liefert die Determinante einen Gruppenhomomorphismus  $\det: \text{Aut}(V) \rightarrow R^\times$ .

BEWEIS. Sei  $A'_{ij} \in M_n(\mathbb{k})$  diejenige Matrix, die wir erhalten, indem wir in  $A$  die  $i$ -te Spalte durch eine Kopie der  $j$ -ten ersetzen. Außerdem sei  $\text{adj } A = (c_{ij})_{i,j}$ . Wir berechnen

$$\begin{aligned} \text{adj } A \cdot A &= \left( \sum_{k=1}^n c_{ik} a_{kj} \right)_{i,j} = \left( \sum_{k=1}^n (-1)^{i+k} \det(A_{ki}) \cdot a_{kj} \right)_{i,j} \\ &= (\det A'_{ij})_{i,j} = \det A \cdot E_n. \end{aligned}$$

Im letzten Schritt haben wir zum einen ausgenutzt, dass  $A'_{ij}$  zwei gleiche Spalten hat und daher  $\det A'_{ij} = 0$ , falls  $i \neq j$ . Zum anderen ist  $A'_{ii} = A$  für alle  $i$ , und die obige Formel folgt aus der Laplace-Entwicklung nach Satz 4.16 (2).

Wenn  $A \in M_n(R)$  in  $R$  invertierbar ist, folgt  $\det A \cdot \det A^{-1} = 1$  aus Satz 4.12 (2), also ist  $\det A$  in  $R$  invertierbar. Umgekehrt, wenn  $\det A$  in  $R$  invertierbar ist, existiert nach obiger Rechnung eine Inverse  $A^{-1}$  wie in (1).

Zu (2) multiplizieren wir  $b$  mit der Inversen  $A^{-1}$  aus (1) und erhalten mit der Laplace-Entwicklung nach der  $i$ -ten Spalte insbesondere

$$x_i = \det A^{-1} (\text{adj } A \cdot b)_i = \frac{1}{\det A} \sum_{j=1}^n (-1)^{i+j} \det(A_{ji}) \cdot b_j = \frac{\det A_j}{\det A}. \quad \square$$

4.22. BEISPIEL. Das Inverse einer  $2 \times 2$ -Matrix ist nach der 1. Cramerschen Regel gerade

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Bereits für  $n \geq 3$  empfiehlt es sich jedoch nicht mehr, Matrizen über Körpern mit der Cramerschen Regel zu invertieren. Das Gauß-Verfahren aus Bemerkung 3.28 (4) ist schneller. Nur über Ringen funktioniert das Gauß-Verfahren in der Regel nicht.

Anhand der obigen Formel sieht man ein Problem mit Determinanten über Schiefkörpern: die quaternionische Matrix  $A = \begin{pmatrix} 1+i & 1+j \\ 1-j & 1-i \end{pmatrix}$  ist invertierbar (Übung), aber es gilt

$$ad - bc = (1+i)(1-i) - (1+j)(1-j) = 2 - 2 = 0.$$

Als letztes wollen wir die Ableitung der Determinante berechnen.

4.23. DEFINITION. Es sei  $A \in M_n(R)$ , dann definieren wir die *Spur* von  $A$  durch

$$\text{tr } A = \sum_{i=1}^n a_{ii} \in R.$$

4.24. FOLGERUNG. *Es sei  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  und  $A: (a, b) \rightarrow M_n(\mathbb{k})$  eine differenzierbare Abbildung, dann gilt*

$$(\det A)' = \operatorname{tr}(A' \cdot \operatorname{adj} A) = \det A \cdot \operatorname{tr}(A' \cdot A^{-1}).$$

Der letzte Ausdruck ist wegen der Cramerschen Regel 4.21 (1) offensichtlich nur sinnvoll, wenn  $\det A \neq 0$ .

BEWEIS. Es sei  $t \in (a, b)$ . Wir setzen  $A = A(t)$  und  $B = A'(t)$ . Mit Hilfe der Leibniz-Formel aus Satz 4.13 sehen wir, dass die Determinante eine Summe von Produkten ist, die aus jeder Spalte genau einen Matrixeintrag enthalten. Nach der Produktregel müssen wir in jedem Produkt jeden einzelnen Faktor einmal ableiten und mit den anderen Faktoren zusammenmultiplizieren. Sei wieder  $\operatorname{adj} A = (c_{ij})_{i,j}$ , dann gilt

$$\begin{aligned} (\det A)'(t) &= \sum_{j=1}^n \det((a_1, \dots, a_{j-1}, b_j, a_{j+1}, \dots, a_n)) \\ &= \sum_{i,j=1}^n (-1)^{i+j} b_{ij} \cdot \det(A_{ij}) = \sum_{i,j=1}^n b_{ij} c_{ji} = \operatorname{tr}(B \cdot \operatorname{adj} A). \end{aligned}$$

Dabei haben in der zweiten Zeile nach der  $j$ -ten Spalte entwickelt und dann die Definition der Adjunkten ausgenutzt. Mit der Cramerschen Regel 4.21 (1) folgt auch die zweite Behauptung.  $\square$

### 4.3. Orientierung reeller Vektorräume

Eine einfache Folgerung aus Satz 4.12 ist die Möglichkeit, endlich erzeugte Vektorräume zu „orientieren“. Wir lassen nur Körper  $\mathbb{k} \subset \mathbb{R}$  zu, damit wir vom „Vorzeichen“ eines Elements von  $\mathbb{k}$  sprechen können.

4.25. DEFINITION. Sei  $\mathbb{k} \subset \mathbb{R}$  ein Körper, und sei  $V$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum. Seien  $(x_1, \dots, x_n)$  und  $(y_1, \dots, y_n)$  zwei Basen von  $V$  mit  $y_j = \sum_{i=1}^n x_i a_{ij}$ . Dann heißen die Basen *gleich orientiert*, wenn die Basiswechselmatrix  $A = (a_{ij})_{i,j} \in \operatorname{End}(\mathbb{k}^n)$  positive Determinante hat.

4.26. FOLGERUNG. *Sei  $V$  ein  $\mathbb{k}$ -Vektorraum der Dimension  $n \geq 1$ . Der Begriff „gleich orientiert“ definiert eine Äquivalenzrelation mit zwei Äquivalenzklassen auf der Menge aller Basen von  $V$ .*

*Sei  $0 \neq \omega \in \Lambda^n V^*$  eine Determinantenfunktion, dann bestehen diese Äquivalenzklassen aus allen Basen  $(b_1, \dots, b_n)$  für die  $\omega(b_1, \dots, b_n) > 0$  beziehungsweise  $\omega(b_1, \dots, b_n) < 0$  gilt.*

BEWEIS. Es seien  $B, C$  Basen von  $V$ . Wir betrachten den Basiswechsel

$$\begin{array}{ccc} & V & \\ C \nearrow & & \nwarrow B \\ \mathbb{k}^n & \xrightarrow{A} & \mathbb{k}^n \end{array}.$$

Da Basiswechsel nach Proposition 2.77 invertierbar sind, hat  $A$  ein Inverses  $A^{-1} \in \text{End}(\mathbb{k}^n)$ . Also folgt  $\det A \neq 0$ .

Wenn wir  $\omega \in \Lambda^n V^* \neq 0$ , dann folgt

$$\omega(c_1, \dots, c_n) = (A^* \omega)(b_1, \dots, b_n) = \det A \cdot \omega(b_1, \dots, b_n).$$

Also sind die Basen  $B$  und  $C$  genau dann gleich orientiert, wenn  $\omega(b_1, \dots, b_n)$  und  $\omega(c_1, \dots, c_n)$  das gleiche Vorzeichen haben. Da „hat das gleiche Vorzeichen wie“ eine Äquivalenzrelation auf  $\mathbb{k}^\times = \mathbb{k} \setminus \{0\} \subset \mathbb{R}^\times$  definiert, erhalten wir die gesuchte Äquivalenzrelation auf der Menge aller Basen.

Da es nur zwei mögliche Vorzeichen gibt, finden wir höchstens zwei Äquivalenzklassen. Dass es zwei gibt, sieht man daran, dass  $(-b_1, b_2, \dots, b_n)$  und  $(b_1, \dots, b_n)$  verschieden orientiert sind.  $\square$

4.27. DEFINITION. Sei  $\mathbb{k} \subset \mathbb{R}$  ein Körper. Eine *Orientierung* eines endlich erzeugten  $\mathbb{k}$ -Vektorraums  $V$  ist eine Äquivalenzklasse gleich orientierter Basen. Sei  $\omega \neq 0$  eine Determinantenfunktion, die genau auf dieser Äquivalenzklasse positiv ist, dann heißt  $\omega$  *positiv* bezüglich der gegebenen Orientierung, und umgekehrt heißt obige Orientierung *durch  $\omega$  induziert*.

Ein Automorphismus  $F \in \text{Aut } V$  heißt *orientierungserhaltend* (*orientierungsumkehrend*), wenn  $\det F > 0$  ( $\det F < 0$ ).

Aus dem obigen Beweis folgt, dass die Begriffe „orientierungserhaltend“ und „orientierungsumkehrend“ nicht von der Wahl einer Orientierung auf  $V$  abhängen.

4.28. BEISPIEL. Auf dem Vektorraum  $\mathbb{R}^n$  definieren wir die *Standard-Orientierung* so, dass die Standard-Basis  $e_1, \dots, e_n$  positiv orientiert ist. Für die Standard-Determinantenfunktion gilt

$$\omega_n(e_1, \dots, e_n) = 1 > 0,$$

also ist sie positiv bezüglich der Standard-Orientierung.

In Bemerkung 1.70 haben wir eine geometrische Interpretation des Kreuz- und des Spatproduktes gegeben. Nur das Vorzeichen hatten wir nicht klären können. Mit Hilfe der Sarrusschen Regel können wir nachrechnen, dass

$$\omega_3(u, v, w) = \langle u \times v, w \rangle$$

gilt. Also ist das Spatprodukt nach 1.70 (2) die (eindeutige) positive Determinantenfunktion, deren Absolutbetrag das Volumen von Parallelotopen angibt. Da

$$\omega_3(u, v, u \times v) = \|u \times v\|^2 \geq 0$$

gilt, ist das Kreuzprodukt  $u \times v$  nach 1.70 (1) der (eindeutige) Vektor im  $\mathbb{R}^3$ , der senkrecht auf  $u$  und  $v$  steht, dessen Länge den Flächeninhalt des von  $u$  und  $v$  aufgespannten Parallelogramms angibt, und der (falls  $u$  und  $v$  nicht linear abhängig sind) mit  $u$  und  $v$  eine positiv orientierte Basis des  $\mathbb{R}^3$  bildet.

4.29. BEMERKUNG. In den Übungen haben Sie die *orthogonale Gruppe*  $O(n)$  der linearen Isometrien des  $\mathbb{R}^n$  kennengelernt, das heißt, der linearen Abbildungen, die das Standardskalarprodukt  $\langle \cdot, \cdot \rangle$  aus Definition 1.52 erhalten. Für alle  $A \in O(n)$  gilt  $\det A \in \{\pm 1\}$ , da

$$O(n) = \{ A \in M_n(\mathbb{R}) \mid A^t \cdot A = E_n \},$$

siehe auch Proposition 3.34. Außerdem haben wir die *spezielle orthogonale Gruppe*  $SO(n)$  der Elemente  $A \in O(n)$  mit  $\det A = 1$  definiert, das ist also die Untergruppe der orientierungserhaltenden Isometrien.

Genauso haben wir die Untergruppe  $SL(n, \mathbb{R}) \subset GL(n, \mathbb{R})$  der Elemente mit Determinante 1 kennengelernt. Wir betrachten zunächst die Gruppe

$$GL(n, \mathbb{R})^+ = \{ A \in GL(n, \mathbb{R}) \mid \det A > 0 \} \subset GL(n, \mathbb{R})$$

der orientierungserhaltenden Automorphismen. Als nächstes gibt es auch eine Untergruppe

$$\{ A \in GL(n, \mathbb{R}) \mid |\det A| = 1 \} \subset GL(n, \mathbb{R})$$

der *volumenerhaltenden Automorphismen*. Dabei erinnern wir uns daran, dass das Volumen durch den Absolutbetrag einer Determinantenfunktion gemessen wird, siehe dazu den Beginn von Abschnitt 4.1. Der Durchschnitt der beiden obigen Untergruppen ist genau  $SL(n, \mathbb{R})$ , somit ist  $SL(n, \mathbb{R})$  die Gruppe der orientierungs- und volumenerhaltenden Automorphismen des  $\mathbb{R}^n$ . Wie bereits am Anfang von Abschnitt 4.1 gesagt, ist über anderen Körpern wie  $\mathbb{C}$  oder  $\mathbb{Z}/p\mathbb{Z}$  nicht möglich, Volumina „ohne Vorzeichen“ zu erklären. Aus dem gleichen Grund ist von den obigen Untergruppen der  $GL(n, \mathbb{k})$  nur  $SL(n, \mathbb{k})$  für alle  $\mathbb{k}$  sinnvoll definiert.

Schließlich können wir mit Hilfe der Determinante (und der ersten Cramerschen Regel) die allgemeine lineare Gruppe aus Definition 2.72 auch über beliebigen kommutativen Ringen mit Eins (oder Körpern) leicht charakterisieren, und die spezielle lineare Gruppe wie folgt definieren:

$$\begin{aligned} GL(n, R) &= \{ A \in M_n(R) \mid \det A \in R^\times \}, \\ SL(n, R) &= \{ A \in M_n(R) \mid \det A = 1 \}. \end{aligned}$$

#### 4.4. Zusammenfassung

Determinanten sind wichtige Invarianten linearer Endomorphismen. Im nächsten Kapitel werden wir sie benutzen, um Eigenwerte von Endomorphismen zu bestimmen. Im Zusammenhang mit der mehrdimensionalen Integral-Transformationsformel wird sie auch wieder benötigt.

Um eine Anschauung für die Determinante zu bekommen und einen Zusammenhang zum Spatprodukt im  $\mathbb{R}^3$  herzustellen, haben wir im Abschnitt 4.1 zunächst Determinantenfunktionen als „orientierte Volumina“ eingeführt (wobei wir aber erst im letzten Abschnitt gesehen haben, dass das Vorzeichen eines Volumens etwas mit seiner Orientierung zu tun hat). Die Axiome für Determinanten gehen auf Weierstraß zurück. Den Räumen  $\Lambda^k V^*$  mit  $k < \dim V$

können Sie später im Zusammenhang mit dem Satz von Stokes oder der de Rham-Kohomologie wieder begegnen.

Im zweiten Abschnitt haben wir Determinanten von Endomorphismen als „Skalierungsfaktoren“ für Volumina eingeführt. Viele Lehrbücher führen Determinanten stattdessen zunächst für Matrizen explizit mit Hilfe der Leibniz-Formel oder des Laplaceschen Entwicklungssatzes ein, müssen dann aber den Produktsatz 4.12 etwas umständlicher beweisen. Bei uns wirkt die Definition zwar etwas komplizierter, hilft uns aber dafür eher, die Bedeutung der Determinanten zu verstehen.

Wir haben verschiedene Rechenverfahren für Determinanten kennengelernt, dabei ist das Gauß-Verfahren für mittelgroße Matrizen über einem Körper das schnellste, falls viele Einträge nicht 0 sind. Wenn wir (wie im nächsten Abschnitt) über Ringen arbeiten müssen, oder wenn eine Matrix nur wenige von 0 verschiedene Einträge enthält, bevorzugen wir die Laplace-Entwicklung.

Die Cramerschen Regeln zum Invertieren von Matrizen und zum Lösen von Gleichungssystemen sind aufgrund ihres hohen Rechenaufwandes eher von theoretischem Interesse, wie wir bei der Ableitung der Determinanten gesehen haben. Das gleiche gilt für die Leibniz-Formel. Immerhin motiviert die Cramersche Regel den Namen „Determinante“ als Invariante, die die eindeutige Lösbarkeit eines quadratischen linearen Gleichungssystems bestimmt („determiniert“).

Im letzten Abschnitt haben wir gesehen, wie das Vorzeichen der Determinanten über  $\mathbb{R}$  uns die Möglichkeit gibt, von Orientierungen zu sprechen. Auch dieser Begriff wird Ihnen noch häufiger begegnen. Außerdem haben wir einige typische Matrixgruppen kennengelernt.

## KAPITEL 5

# Eigenwerte und Normalformen

Wir versuchen, Endomorphismen durch möglichst einfache Matrizen darzustellen, im Idealfall durch Diagonalmatrizen. Dazu studieren wir Eigenwerte und Eigenvektoren. Wir lernen feinere Invarianten von Endomorphismen endlich-dimensionaler Vektorräume kennen, das charakteristische und das Minimalpolynom. Beide helfen, Eigenwerte zu finden und die Struktur von Endomorphismen besser zu verstehen. Am Schluss des Kapitels beweisen wir einige Struktursätze und betrachten die Jordan-Normalform.

Wir benötigen nach wie vor das Kommutativgesetz für die Multiplikation, und arbeiten daher meist über Körpern oder über kommutativen Ringen mit Eins. Da wir viel mit Polynomen arbeiten müssen, beinhaltet dieses Kapitel auch einiges an Ringtheorie. Unter anderem beweisen wir auch den Satz über die eindeutige Primfaktorzerlegung und den chinesischen Restsatz.

### 5.1. Eigenvektoren

Ab sofort sei  $\mathbb{k}$  stets ein Körper, insbesondere kommutativ.

5.1. DEFINITION. Es sei  $F \in \text{End}_{\mathbb{k}} V$  ein Endomorphismus eines  $\mathbb{k}$ -Vektorraums  $V$  und  $\lambda \in \mathbb{k}$ . Der  $\lambda$ -Eigenraum von  $F$  ist die Menge

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

Wenn  $V_{\lambda} \neq \{0\}$ , ist  $\lambda$  ein *Eigenwert* von  $F$ , und die Elemente von  $V_{\lambda} \setminus \{0\}$  heißen *Eigenvektoren* von  $F$  zum Eigenwert  $\lambda$ .

Genauso definieren wir Eigenvektoren, Eigenräume und Eigenwerte quadratischer Matrizen über  $\mathbb{k}$ .

Der Eigenraum  $V_{\lambda}$  besteht somit aus dem Nullvektor und allen Eigenvektoren zu  $\lambda$ . Man beachte, dass der Nullvektor als Element des Eigenraumes zugelassen ist, aber selbst nicht als Eigenvektor betrachtet wird. Das liegt daran, dass die Gleichung  $F(0) = 0 \cdot \lambda$  für alle  $\lambda$  und alle  $F$  erfüllt ist und somit keine Information über  $F$  und  $\lambda$  enthält.

5.2. BEISPIEL. Eigenvektoren und Eigenwerte spielen in vielen Situationen eine Rolle.

- (1) Der Eigenraum zum Eigenwert 0 ist gerade der Kern, siehe Definition 2.36.

- (2) Eigenvektoren zum Eigenwert 1 sind gerade Fixpunkte des Endomorphismus. In den Übungen schauen wir uns dazu bestimmte Isometrien der Ebene  $\mathbb{R}^2$  und des Raumes  $\mathbb{R}^3$  an.
- (3) Es sei

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \in M_2(\mathbb{Q}), \quad v = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{Q}^2 \quad \text{und} \quad w = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in \mathbb{Q}^2,$$

dann ist  $v$  ein Eigenvektor von  $A$  zum Eigenwert 3 und  $w$  ein Eigenvektor zum Eigenwert 1, wie man leicht nachrechnet. Anschaulich bedeutet das, dass  $F$  in Richtung von  $v$  um den Faktor 3 streckt, und die Gerade durch 0 in Richtung  $w$  punktwise festgehalten wird.

- (4) Auch Endomorphismen unendlich-dimensionaler Vektorräume können Eigenvektoren haben. Sei etwa  $C^\infty(\mathbb{R})$  der Raum der unendlich oft differenzierbaren reellwertigen Funktionen auf  $\mathbb{R}$ , dann ist die Ableitung  $\frac{d}{dx}$  ein Endomorphismus von  $C^\infty(\mathbb{R})$ , und jedes  $\lambda \in \mathbb{R}$  ist Eigenwert; die zugehörigen Eigenvektoren („Eigenfunktionen“) haben die Form

$$x \mapsto c e^{\lambda x} \quad \text{mit } c \neq 0.$$

5.3. BEMERKUNG. Da  $\mathbb{k}$  ein Körper ist, gilt  $v \cdot \lambda = (\lambda \cdot \text{id}_V)(v)$ , so dass wir den Eigenraum  $V_\lambda$  auch schreiben können als

$$V_\lambda = \{ v \in V \mid (F - \lambda \cdot \text{id}_V)(v) = 0 \} = \ker(F - \lambda \cdot \text{id}_V).$$

Nach Proposition 2.29 ist  $F - \lambda \cdot \text{id}_V$  wieder eine lineare Abbildung; dazu muss die Multiplikation in  $\mathbb{k}$  kommutativ sein. Insbesondere ist  $V_\lambda \subset V$  ein Unterraum nach Proposition 2.37 (1).

Wir nehmen jetzt an, dass  $V$  endlich-dimensional ist. Um den Eigenraum  $V_\lambda$  zu einem vorgegebenen  $\lambda \in \mathbb{k}$  zu berechnen, müssen wir also nur eine Basis  $B$  von  $V$  wählen, die Abbildungsmatrix  $A = {}_B F_B$  von  $F$  bezüglich der Basis  $B$  (sowohl für den Definitions- als auch für den Wertebereich  $V$ ) bestimmen, und dann das Gleichungssystem

$$(5.1) \quad (A - \lambda E_n)(x) = 0.$$

lösen. Die Lösungsmenge liefert genau die  $B$ -Koordinaten der Elemente von  $V_\lambda$ .

Wenn  $A \in M_n(\mathbb{k})$  gegeben ist, berechnen wir mit (5.1) die Eigenräume von  $A$  als Unterräume von  $\mathbb{k}^n$ .

Wenn ein Endomorphismus  $F \in \text{End}_{\mathbb{k}} V$  genug Eigenvektoren hat, können wir unter Umständen eine Basis von  $V$  finden, bezüglich der  $F$  eine besonders einfache Abbildungsmatrix hat. Wir erinnern uns an den Begriff der Dreiecksmatrix aus Definition 4.17.

5.4. DEFINITION. Eine *Diagonalmatrix* ist eine quadratische Matrix  $A = (a_{ij})_{i,j} \in M_n(R)$ , so dass  $a_{ij} = 0$  für alle  $i, j \in \{1, \dots, n\}$  mit  $i \neq j$ .

Sei  $M$  ein endlich erzeugter freier  $R$ -Modul. Ein Endomorphismus  $F \in \text{End}_R M$  heißt *diagonalisierbar*, wenn es eine Basis  $B$  von  $M$  gibt, bezüglich

der die Abbildungsmatrix von  $F$  eine Diagonalmatrix ist. Wir nennen  $F$  *trigonalisierbar*, wenn es eine Basis gibt, bezüglich der die Abbildungsmatrix von  $F$  eine Dreiecksmatrix ist.

Eine Matrix  $A \in M_n(R)$  heißt trigonalisierbar (diagonalisierbar), wenn es eine invertierbare Matrix  $G \in GL(n, R)$  gibt, so dass  $G^{-1} \cdot A \cdot G$  eine Dreiecks- (Diagonal-) matrix ist.

Wir betrachten dazu die Diagramme

$$(5.2) \quad \begin{array}{ccc} M & \xrightarrow{F} & M \\ B \uparrow \cong & & \cong \uparrow B \\ R^n & \xrightarrow[C_{BF_B}]{} & R^n \end{array}, \quad \begin{array}{ccc} R^n & \xrightarrow{A} & R^n \\ G \uparrow \cong & & \cong \uparrow G \\ R^n & \xrightarrow[C_{G^{-1}AG}]{} & R^n \end{array}.$$

Gesucht ist eine Basis  $B$  beziehungsweise eine invertierbare Matrix  $G$ , so dass  $C$  eine Diagonal- beziehungsweise Dreiecksmatrix wird. Im Vergleich zu Diagramm (3.1) fällt auf, dass links und rechts dieselbe Basis  $B$  beziehungsweise dieselbe invertierbare Matrix  $G$  steht, siehe Bemerkung 5.10 unten.

Eine typische Diagonalmatrix hat die Gestalt

$$(5.3) \quad \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} \in M_n(R).$$

Da jede Diagonalmatrix insbesondere eine Dreiecksmatrix ist, folgt trigonalisierbar aus diagonalisierbar. Übrigens spielt es keine Rolle, ob wir Trigonalisierbarkeit mit oberen oder mit unteren Dreiecksmatrizen definieren: sei die Abbildungsmatrix  $A = (a_{ij})_{i,j}$  von  $F$  bezüglich  $B = (b_1, \dots, b_n)$  eine obere Dreiecksmatrix, dann ist die Abbildungsmatrix  $(a_{n+1-i, n+1-j})_{i,j}$  von  $F$  bezüglich der Basis  $(b_n, \dots, b_1)$  eine untere Dreiecksmatrix, und umgekehrt.

Mit Diagonalmatrizen kann man besonders einfach rechnen. Daher wäre es schön, wenn jeder Endomorphismus diagonalisierbar wäre. Das ist aber leider nicht der Fall — am Ende dieses Kapitels werden wir einige Kriterien kennen dafür lernen, dass eine Matrix oder Abbildung diagonalisierbar beziehungsweise trigonalisierbar ist. Außerdem werden wir in Gestalt der Jordan-Normalform eine besonders einfache Form von Dreiecksmatrizen kennenlernen, in die wir jede Dreiecksmatrix umformen können.

**5.5. PROPOSITION.** *Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum mit Basis  $B = (b_1, \dots, b_n)$ , und sei  $F \in \text{End}_{\mathbb{k}} V$  ein Endomorphismus mit Abbildungsmatrix  $A$  bezüglich  $B$ . Dann ist  $b_j$  genau dann ein Eigenvektor von  $F$  zum Eigenwert  $\lambda_j$ , wenn die  $j$ -te Spalte von  $A$  gerade  $e_j \cdot \lambda$  ist. Insbesondere ist  $F$  genau dann diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt.*

**BEWEIS.** Die  $j$ -te Spalte von  $A$  enthält die  $B$ -Koordinaten von  $F(b_j)$  nach Folgerung 2.75. Also ist  $b_j$  genau dann ein Eigenvektor zum Eigenwert  $\lambda_j$ ,

wenn  $F(b_j) = b_j \cdot \lambda_j$ , und somit die  $j$ -te Spalte  $a_j$  von  $A$  die Zahl  $\lambda_j$  an der  $j$ -ten Stelle und sonst nur Nullen enthält, das heißt, wenn  $a_j = e_j \cdot \lambda_j$ . Es folgt die erste Behauptung. Die zweite ist jetzt offensichtlich.  $\square$

Um eine Basis aus Eigenvektoren zu bekommen, brauchen wir also eine Familie von Eigenvektoren, die linear unabhängig sind und  $V$  erzeugen. Lineare Unabhängigkeit garantiert uns in Spezialfällen die folgende Überlegung.

5.6. PROPOSITION. *Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum. Dann sind Eigenvektoren eines Endomorphismus von  $V$  zu verschiedenen Eigenwerten linear unabhängig.*

BEWEIS. Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}} V$  ein Endomorphismus. Es sei  $(v_i)_{i \in I}$  eine Familie in  $V \setminus \{0\}$  und  $(\lambda_i)_{i \in I}$  eine Familie in  $\mathbb{k}$  mit  $\lambda_i \neq \lambda_j$  für alle  $i, j \in I$  mit  $i \neq j$ , so dass  $v_i$  jeweils Eigenvektor zum Eigenwert  $\lambda_i$  ist. Zu zeigen ist, dass die Familie  $(v_i)_{i \in I}$  linear unabhängig ist.

Wir beginnen mit dem Fall  $I = \{1, \dots, k\}$ , das heißt, wir betrachten nur endlich viele Eigenvektoren. In diesem Fall verläuft der Beweis durch vollständige Induktion über  $k$ . Im Fall  $k = 0$  ist nichts zu zeigen.

Sei  $k \geq 1$  und seien  $a_i \in \mathbb{k}$  gegeben, so dass

$$0 = \sum_{i=1}^k v_i \cdot a_i .$$

Dann dürfen wir den Endomorphismus  $F - \lambda_k \text{id}_V$  anwenden und erhalten

$$\begin{aligned} 0 &= (F - \lambda_k \text{id}_V) \left( \sum_{i=1}^k v_i \cdot a_i \right) = \sum_{i=1}^k v_i \cdot (\lambda_i - \lambda_k) \cdot a_i \\ &= \sum_{i=1}^{k-1} v_i \cdot \underbrace{((\lambda_i - \lambda_k) \cdot a_i)}_{\neq 0} . \end{aligned}$$

Nach Induktionsvoraussetzung verschwinden die Zahlen  $(\lambda_i - \lambda_k) \cdot a_i \in \mathbb{k}$ , es folgt  $a_i = 0$  für  $i = 1, \dots, k-1$ . Aus der ursprünglichen Gleichung wird also

$$0 = v_k \cdot a_k ,$$

und da  $v_k \neq 0$  folgt  $a_k = 0$ . Also sind  $(v_1, \dots, v_k)$  linear unabhängig.

Es bleibt der Fall einer unendlichen Indexmenge. Wir hatten lineare Unabhängigkeit in Abschnitt 2.6 kurz eingeführt. Zu zeigen ist, dass die Abbildung  $\Psi_E$  aus Satz 2.81 zur Familie  $E = (v_i)_{i \in I}$  injektiv ist. Nach Definition 2.79 ist das äquivalent zu folgender Aussage. Wenn  $a_i = 0$  für alle bis auf endlich viele  $i \in I$  gilt, folgt aus

$$0 = \sum_{i \in I} v_i \cdot a_i$$

bereits  $a_i = 0$  für alle  $i \in I$ . Aber diese Aussage haben wir gerade bewiesen.  $\square$

Wir erinnern uns an die Definition 2.45 einer direkten Summe zweier Unterräume. Wir wollen die direkte Summe vieler Unterräume analog dazu definieren.

5.7. DEFINITION. Es sei  $(U_i)_{i \in I}$  eine Familie von Untermoduln eines  $R$ -Moduls  $M$ . Dann definieren wir ihre *Summe* als

$$\sum_{i \in I} U_i = \left\{ \sum_{i \in I_0} u_i \mid I_0 \subset I \text{ endlich und } u_i \in U_i \text{ für alle } i \in I_0 \right\}.$$

Wir nennen die Summe *direkt*, wenn für alle endlichen Teilmengen  $I_0 \subset I$  und alle Familien  $(u_i)_{i \in I_0}$  mit  $u_i \in U_i$  für alle  $i \in I_0$  gilt

$$\sum_{i \in I_0} u_i = 0 \quad \implies \quad u_i = 0 \text{ für alle } i \in I_0.$$

Man überprüft leicht, dass die Summe von Unterräumen wieder ein Unterraum ist, und zwar der kleinste Unterraum von  $V$ , der alle  $U_i$  umfasst. Wie in Proposition 2.47 (2) ist die Summe genau dann direkt, wenn die Darstellung  $\sum_{i \in I_0} u_i$  eindeutig ist bis auf das Hinzufügen oder Weglassen von Summanden  $u_j = 0$ .

5.8. FOLGERUNG. *Es sei  $V$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}} V$  ein Endomorphismus.*

- (1) *Dann hat  $F$  höchstens  $n$  verschiedene Eigenwerte.*
- (2) *Seien  $V_{\lambda_1}, \dots, V_{\lambda_k}$  Eigenräume von  $F$  zu verschiedenen Eigenwerten von  $F$ , dann ist die Summe  $V_{\lambda_1} + \dots + V_{\lambda_k} \subset V$  direkt.*
- (3) *Der Endomorphismus  $F$  ist genau dann diagonalisierbar, wenn  $V$  eine direkte Summe aus Eigenräumen von  $F$  ist.*

Teil (2) gilt analog auch für unendlich viele Eigenräume zu paarweise verschiedenen Eigenwerten in einem unendlich-dimensionalen Unterraum.

BEWEIS. Nach dem Basisaustauschsatz 3.4 kann es kein Tupel aus mehr als  $n$  linear unabhängigen Vektoren geben, und (1) folgt aus Proposition 5.6.

Seien jetzt  $V_{\lambda_1}, \dots, V_{\lambda_k}$  Eigenräume zu verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_k$  von  $F$ , und seien  $v_j \in V_{\lambda_j}$  für  $j = 1, \dots, k$ . Angenommen es gilt

$$\sum_{j=1}^k v_j = 0,$$

dann verschwände eine nichttriviale Linearkombination von Eigenvektoren im Widerspruch zu Proposition 5.6, es sei denn,  $v_j = 0$  für alle  $j = 1, \dots, k$ . Also ist die Summe direkt, und es folgt (2).

Als nächstes beweisen wir (3), „ $\implies$ “. Sei also  $B = (b_1, \dots, b_n)$  eine Basis von  $V$ , so dass  $F$  bezüglich  $B$  durch eine Diagonalmatrix  $A$  dargestellt wird. Nach Proposition 5.6 ist jeder Basisvektor ein Eigenvektor. Seien  $\lambda_1, \dots, \lambda_k$  die Eigenwerte von  $F$ . Zu jedem Eigenwert  $\lambda_j$  von  $F$  sei  $U_{\lambda_j} \subset V_{\lambda_j}$  der Unterraum, der von den Basisvektoren  $b_i$  mit  $F(b_i) = b_i \cdot \lambda_j$  aufgespannt wird. Dann

folgt aus (2) und der Tatsache, dass  $B$  eine Basis ist und jeder Basisvektor in einem  $U_{\lambda_j}$  vorkommt, dass

$$V = \sum_{j=1}^k U_{\lambda_j} \subset \bigoplus_{j=1}^k V_{\lambda_j} \subset V .$$

Daher gilt „ $=$ “ anstelle von „ $\subset$ “ in der obigen Ungleichung, insbesondere ist  $V$  die direkte Summe der Eigenräume von  $F$ .

Sei zu (3), „ $\Leftarrow$ “, schließlich  $V = \sum_{i=1}^n V_{\lambda_i}$ , dann ist die Summe direkt nach (2). Wir wählen Basen  $(b_1^i, \dots, b_{r_i}^i)$  von  $V_{\lambda_i}$ . Nach Voraussetzung bildet das Tupel  $(b_1^1, b_2^1, \dots, b_{r_1}^1, b_1^2, \dots, b_{r_k}^k)$  ein Erzeugendensystem von  $V$ . Da die Summe direkt ist, bildet es sogar eine Basis. Denn seien jetzt  $a_j^i \in \mathbb{k}$ , so dass

$$0 = \sum_{i=1}^k \sum_{j=1}^{r_i} b_j^i \cdot a_j^i .$$

Da die Summe direkt ist, folgt für jedes  $i \in \{1, \dots, k\}$ , dass

$$\sum_{j=1}^{r_i} b_j^i \cdot a_j^i = 0 ,$$

also auch  $a_1^i = \dots = a_{r_i}^i = 0$ , da  $(b_1^i, \dots, b_{r_i}^i)$  als Basis von  $V_{\lambda_i}$  linear unabhängig ist. Insgesamt ist das Tupel  $(b_1^1, b_2^1, \dots, b_{r_1}^1, b_1^2, \dots, b_{r_k}^k)$  eine Basis aus Eigenvektoren, also ist  $F$  diagonalisierbar nach Proposition 5.5.  $\square$

**5.9. FOLGERUNG.** *Es sei  $V$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}}(V)$ . Wenn es  $n$  verschiedene Eigenwerte gibt, besitzt  $V$  eine Basis aus Eigenvektoren; somit ist  $F$  dann diagonalisierbar.*

*Es sei  $W$  ebenfalls ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum und  $G \in \text{End}_{\mathbb{k}} W$ . Dann existiert genau dann ein Isomorphismus  $P: V \rightarrow W$  mit  $P \circ F = G \circ P$ , wenn  $G$  die gleichen Eigenwerte wie  $F$  hat.*

Die zweite Aussage entspricht dem Diagramm

$$\begin{array}{ccc} V & \xrightarrow{F} & V \\ P \downarrow \cong & & \cong \downarrow P \\ W & \xrightarrow{G} & W . \end{array}$$

BEWEIS. Übung.  $\square$

**5.10. BEMERKUNG.** Als Fazit dieses Abschnitts halten wir fest, dass es sich lohnt, Eigenwerte und Eigenvektoren von Endomorphismen zu bestimmen, um eine Abbildungsmatrix in möglichst einfacher Form zu finden. Für Abbildungen zwischen verschiedenen endlich-dimensionalen Vektorräumen, haben wir das im Rangsatz 3.13 bereits getan. Dort haben wir gezeigt, dass es reicht, den Rang der Abbildung zu kennen, um eine besonders einfache darstellende Matrix anzugeben.

Hier ist die Situation komplizierter, da wir in Definition 5.4 dieselbe Basis für Definitions- und Wertebereich der Abbildung wählen müssen, vergleiche dazu Diagramm (5.2) mit Diagramm (3.1). In der Situation von Folgerung 5.9 ist die Diagonalmatrix (5.3) bis auf die Reihenfolge der Einträge eindeutig festgelegt. Es gibt also bereits hier weitaus mehr mögliche darstellende Matrizen als beim Rangsatz.

Unser Projekt für den zweiten Teil dieses Kapitels wird darin bestehen, für jeden Endomorphismus eines endlich-dimensionalen  $\mathbb{k}$ -Vektorraums eine darstellende Matrix in einer gewissen Blockgestalt zu finden, die bis auf die Reihenfolge der Blöcke eindeutig ist.

## 5.2. Das charakteristische Polynom

Wir betrachten Bemerkung 5.3, speziell Gleichung (5.1), um Eigenwerte zu beschreiben. In Proposition 5.14 unten sehen wir, dass  $\lambda \in \mathbb{k}$  genau dann ein Eigenwert von  $F \in \text{End}_{\mathbb{k}}(V)$  ist, wenn  $\det(\lambda \text{id}_V - F) = 0$ . Wir können also die Funktion

$$\chi_F: \mathbb{k} \rightarrow \mathbb{k} \quad \text{mit} \quad \chi_F(\lambda) = \det(\lambda \text{id}_V - F)$$

betrachten und ihre Nullstellen suchen, um Eigenwerte von  $F$  zu finden. Ohne etwas über die Struktur dieser Funktion zu wissen, kann es allerdings schwierig werden, Nullstellen zu finden. Wir wollen die Funktion  $\chi_F$  daher etwas algebraischer betrachten, was uns in vieler Hinsicht mehr Informationen liefert.

Es sei  $R$  ein kommutativer Ring mit Eins, und es sei  $R^{(\mathbb{N})}$  der Raum der endlichen  $R$ -wertigen Folgen, siehe Definition 2.79. Wir erinnern uns an die Menge

$$R[X] = \left\{ \sum_{i=0}^{\infty} p_i X^i \mid (p_i)_{i \in \mathbb{N}} \in R^{(\mathbb{N})} \right\}$$

der Polynome über  $R$  aus Beispiel 2.80. Sei  $P = P(X) = \sum_{i=0}^{\infty} p_i X^i$  ein Polynom, dann können wir Elemente  $r \in R$  einsetzen via

$$P(r) = \sum_{i=0}^{\infty} p_i r^i \in R.$$

Man beachte, dass wir bei einem Polynom die Variable  $X$  in der Notation  $P = P(X)$  mitschreiben oder weglassen dürfen. In der Regel lassen wir die Variable  $X$  nur dann weg, wenn klar ist, dass  $P$  ein Polynom in der Variablen  $X$  ist. Als nächstes machen wir aus  $R[X]$  einen Ring.

5.11. DEFINITION. Es seien

$$P = P(X) = \sum_{i=0}^{\infty} p_i X^i \quad \text{und} \quad Q = Q(X) = \sum_{j=0}^{\infty} q_j X^j \in R[X]$$

zwei Polynome, dann definieren wir ihre *Summe* und ihr *Produkt* durch

$$(1) \quad (P + Q)(X) = \sum_{i=0}^{\infty} (p_i + q_i) X^i \quad \in R[X],$$

$$(2) \quad \text{und} \quad (P \cdot Q)(X) = \sum_{i=0}^{\infty} \sum_{j=0}^i (p_{i-j} \cdot q_j) X^i \quad \in R[X].$$

Außerdem betrachten wir  $R$  als Teilmenge von  $R[X]$ , dabei wird aus  $r \in R$  das *konstante Polynom*

$$r = r \cdot X^0 = \sum_{i=0}^{\infty} r \delta_{i0} X^i.$$

5.12. PROPOSITION. *Es sei  $R$  ein kommutativer Ring mit Eins  $1 \in R$ .*

- (1) *Dann ist  $(R[X], +, \cdot)$  ein kommutativer Ring mit Eins  $1 \in R \subset R[X]$ .*  
 (2) *Für alle  $r \in R$  gilt*

$$(P + Q)(r) = P(r) + Q(r) \quad \text{und} \quad (P \cdot Q)(r) = P(r) \cdot Q(r) \in R.$$

Für Rechnungen mit Polynomen gelten nach (1) also die gleichen Regeln wie im Ring  $R$ , nämlich Assoziativ-, Kommutativ- und Distributivgesetze. Wegen (2) bleiben alle Rechnungen mit Polynomen gültig, wenn wir für  $X$  Elemente aus  $R$  in  $P$  einsetzen.

BEWEIS. Zu (1) müssen wir zunächst zeigen, dass  $(R[X], +)$  eine abelsche Gruppe bildet. Da die Addition von Polynomen der Addition im  $R$ -Modul  $R^{(\mathbb{N})}$  entspricht, folgt das wie in Abschnitt 2.6 nach Definition 2.79. Die Ringaxiome (R1)–(R4) folgen durch Nachrechnen aus den entsprechenden Axiomen für  $R$ . Wir machen das hier für das erste Distributivgesetz vor. Seien

$$P(X) = \sum_{i=1}^k p_i X^i, \quad Q(X) = \sum_{i=1}^{\ell} q_i X^i \quad \text{und} \quad R(X) = \sum_{i=1}^m r_i X^i$$

Polynome über  $R$ , dann gilt

$$\begin{aligned} (P \cdot (Q + R))(X) &= \sum_{i=0}^{k+\max(\ell,m)} \sum_{j=0}^i (p_{i-j} (q_j + r_j)) X^i \\ &= \sum_{i=0}^{\max(k+\ell, k+m)} \left( \sum_{j=0}^i p_{i-j} q_j + \sum_{j=0}^i p_{i-j} r_j \right) X^i \\ &= (P \cdot Q + P \cdot R)(X). \end{aligned}$$

Die anderen Axiome folgen durch ähnliche Rechnungen.

Zu (2) zeigen wir

$$(P + Q)(r) = \sum_{i=0}^{\max(k,\ell)} (p_i + q_i) r^i = \sum_{i=0}^k p_i r^i + \sum_{i=0}^{\ell} q_i r^i = P(r) + Q(r),$$

$$\begin{aligned}
(P \cdot Q)(r) &= \sum_{i=0}^{k+\ell} \sum_{j=0}^i p_{i-j} q_j r^i = \sum_{i=0}^{k+\ell} \sum_{j=0}^i p_{i-j} r^{i-j} \cdot q_j r^j \\
&= \left( \sum_{i=0}^k p_i r^i \right) \cdot \left( \sum_{j=0}^{\ell} q_j r^j \right) = P(r) \cdot Q(r). \quad \square
\end{aligned}$$

Im Beweis von (2) haben wir unter anderem benutzt, dass  $R$  kommutativ ist. Über nichtkommutativen Ringen gäbe es keine schöne Auswertungsabbildung.

Es sei  $\mathbb{k}$  ein Körper,  $V$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}} V$  ein Endomorphismus. Dann können wir eine Basis  $B$  von  $V$  wählen und erhalten die Abbildungsmatrix  $A \in M_n(\mathbb{k})$  von  $F$  bezüglich  $B$ , siehe Folgerung 2.75. Wir hatten in Kapitel 4 die Determinante quadratischer Matrizen über einem Ring definiert, siehe Definition 4.11. Also können wir die Determinante von  $X E_n - A \in M_n(\mathbb{k}[X])$  bilden und erhalten

$$\chi_A(X) = \det(X \cdot E_n - A) \in \mathbb{k}[X].$$

Wir können zeigen, dass  $\chi_A(X)$  nicht von der Wahl der Basis  $B$  abhängt. Sei nämlich  $C$  eine weitere Basis und  $G \in GL(n, \mathbb{k})$  die Basiswechselmatrix mit  $C = B \cdot G$ , siehe Bemerkung 2.76. Dann hat  $F$  bezüglich der Basis  $C$  die Abbildungsmatrix  $G^{-1} \cdot A \cdot G$ , vergleiche das mit Diagramm (5.2). Da  $\mathbb{k} \subset \mathbb{k}[X]$ , dürfen wir  $G$  als Matrix in  $M_n(\mathbb{k}[X])$  auffassen. Da  $G$  mit der Einheitsmatrix  $E_n$  kommutiert, folgt aus Satz 4.12, dass

$$\begin{aligned}
\det(X \cdot E_n - G^{-1} \cdot A \cdot G) &= \det(G^{-1} \cdot (X \cdot E_n - A) \cdot G) \\
&= \det G^{-1} \cdot \det(X \cdot E_n - A) \cdot \det G = \det(X \cdot E_n - A).
\end{aligned}$$

Somit hängt  $\chi_A(X)$  nicht von der Wahl der Basis  $B$  ab, wir dürfen es also als Invariante des Endomorphismus  $F$  betrachten.

**5.13. DEFINITION.** Es sei  $R$  ein kommutativer Ring und  $n \in \mathbb{N}$ . Sei  $A \in M_n(R)$ , dann heißt

$$\chi_A(X) = \det(X \cdot E_n - A) \in R[X]$$

das *charakteristische Polynom* der Matrix  $A$ .

Sei  $V$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum mit Basis  $B$  und  $F \in \text{End}_{\mathbb{k}} V$  ein Endomorphismus mit Abbildungsmatrix  $A$  bezüglich  $B$ . Dann heißt  $\chi_F(X) = \chi_A(X)$  das charakteristische Polynom von  $F$ .

**5.14. PROPOSITION.** *Die Eigenwerte eines Endomorphismus eines endlich-dimensionalen  $\mathbb{k}$ -Vektorraums sind genau die Nullstellen des charakteristischen Polynoms.*

**BEWEIS.** Es sei  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}}(V)$ . Eine Zahl  $\lambda \in \mathbb{k}$  ist nach Bemerkung 5.3 genau dann ein Eigenwert von  $F$ , wenn  $\ker(F - \lambda \text{id}_V) \neq \{0\}$ , das heißt, wenn  $F - \lambda \text{id}_V$  nicht injektiv ist. Aus Folgerung 3.18 ergibt sich, dass  $F - \lambda \text{id}_V$  genau dann injektiv ist, wenn  $F - \lambda \text{id}_V$  surjektiv ist. Also ist  $\lambda$  genau dann ein Eigenwert von  $F$ ,

wenn  $F - \lambda \operatorname{id}_V$  nicht invertierbar ist. Nach Folgerung 4.21 (1) ist das genau dann der Fall, wenn  $\chi_F(\lambda) = \det(\lambda \operatorname{id}_V - F) = 0$ . Dabei haben wir benutzt, dass Einsetzen von  $\lambda$  mit allen Rechenoperationen in der Leibniz-Formel für die Determinante verträglich ist.  $\square$

Nach dem Fundamentalsatz 1.62 der Algebra hat jedes komplexe Polynom eine Nullstelle, also hat jeder Endomorphismus eines endlich-dimensionalen komplexen Vektorraums einen Eigenwert. Später werden wir sehen, dass noch einiges mehr gilt.

Zur Berechnung des charakteristischen Polynoms empfiehlt sich zum Beispiel die Laplace-Entwicklung 4.16. Der Gauß-Algorithmus funktioniert nicht, da  $\mathbb{k}[X]$  kein Körper ist.

5.15. BEMERKUNG. Es lohnt sich, das charakteristische Polynom etwas detaillierter zu betrachten.

- (1) Die Einträge der Matrix  $X \cdot E_n - A$  sind Polynome vom Grad  $\leq 1$ , das heißt, die Variable  $X$  kommt höchstens einmal vor. Aufgrund der Leibniz-Formel 4.13 lässt sich  $\chi_A(X)$  als Summe von Produkten aus je  $n$  Matrixeinträgen schreiben. Also kann  $X$  in  $\chi_A(X)$  höchstens in der  $n$ -ten Potenz vorkommen.

Wir multiplizieren alle Klammern  $(X - a_{ii})$  in der Leibniz-Formel in Gedanken aus. Zum Koeffizienten von  $X^n$  tragen dann nur diejenigen Produkte bei, bei denen jeder Faktor  $X$  ist. Da alle  $X$  auf der Diagonalen stehen, trägt also nur das Produkt zur Permutation  $\operatorname{id} \in S(n)$  bei. Das gleiche gilt immer noch, wenn nur ein Faktor nicht gerade  $X$  ist. Es folgt

$$\begin{aligned} \chi_A(X) &= (X - a_{11}) \cdots (X - a_{nn}) + \dots \\ (*) \quad &= X^n - (a_{11} + \dots + a_{nn}) X^{n-1} + \dots, \end{aligned}$$

- (2) Da  $\chi_F(X)$  nur von  $F$  abhängt, sind die Koeffizienten von  $\chi_F(X)$  Invarianten des Endomorphismus  $F$ . Schreibe also

$$\chi_F(X) = X^n - \sigma_1(F) X^{n-1} \pm \dots + (-1)^n \sigma_n(F) X^0,$$

dann nennt man  $\sigma_i(F)$  die *elementarsymmetrischen Funktionen* von  $F$ . Analog definieren wir  $\sigma_i(A)$ . Aus der Leibniz-Formel folgt, dass  $\sigma_i(A)$  eine Summe von Produkten von je  $i$  Matrixeinträgen von  $A$  und einem Vorfaktor ist, denn jeder Summand in der Leibniz-Formel 4.13 ist ein Produkt aus  $n$  Faktoren. Nach Ausmultiplizieren ist jeder der Faktoren entweder  $X$  oder ein Matrixeintrag.

- (3) Als erstes wollen wir den konstanten Term  $(-1)^n \sigma_n(F)$  des charakteristischen Polynoms bestimmen. Sei also  $A$  wieder die Abbildungsmatrix von  $F$  bezüglich einer Basis  $B$ . Dann setzen wir  $X = 0$  und erhalten sofort

$$\chi_A(0) = \det(-A) = (-1)^n \det A,$$

und analog für Endomorphismen  $F$ . Somit gilt  $\sigma_n(F) = \det F$ .

- (4) Wir schauen uns noch  $\sigma_1(A)$  an. Für jede Permutation  $\rho \in S(n) \setminus \{\text{id}\}$  gibt es wenigstens zwei verschiedene Indizes  $i, j \in \{1, \dots, n\}$  mit  $\rho(i) \neq i$  und  $\rho(j) \neq j$ . Also liefert  $\rho$  einen Summanden, in dem nur  $X^k$  mit  $k \leq n-2$  vorkommt. Der einzige Beitrag zu  $\sigma_1(A)$  in der Leibniz-Formel 4.13 kommt also wieder von  $\rho = \text{id} \in S(n)$ . Aus (\*) folgt

$$\sigma_1(A) = a_{11} + \dots + a_{nn} = \text{tr}(A),$$

siehe Definition 4.23.

5.16. BEISPIEL. Als Anwendungsbeispiel betrachten wir ein ungedämpftes Federpendel, beschrieben durch die Differentialgleichung

$$(*) \quad \ddot{u}(t) = -c u(t).$$

Hierbei ist  $u(t)$  die Auslenkung des Pendels aus der Ruhelage in Abhängigkeit von der Zeit  $t$ , und  $c > 0$  hängt von der Federkonstante und der Masse des Pendels ab. Der doppelte Punkt bezeichnet die zweite Ableitung nach der Zeit. Solche Differentialgleichungen gehören eigentlich in den Bereich der Analysis. Wenn sie linear sind, kann lineare Algebra aber bei der Lösung helfen.

Der erste Trick bei der Lösung besteht darin, die Geschwindigkeit  $v$  als zusätzliche Variable einzuführen. Jetzt kann man die obige Differentialgleichung zweiter Ordnung umschreiben als System erster Ordnung, nämlich

$$(\dagger) \quad \begin{aligned} \dot{u}(t) &= v(t), \\ \dot{v}(t) &= -c u(t), \end{aligned} \quad \text{beziehungsweise} \quad \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -c & 0 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix}.$$

Da die Koeffizientenmatrix  $C$  nicht von der Zeit abhängt, können wir versuchen, sie mit Hilfe eines konstanten Basiswechsels zu diagonalisieren. Ihr charakteristisches Polynom

$$\chi_C(X) = \det \begin{pmatrix} X & -1 \\ c & X \end{pmatrix} = X^2 + c$$

hat die Nullstellen  $\pm i\sqrt{c} \in \mathbb{C}$ . An dieser Stelle gehen wir also zu komplexen Koeffizienten über. Am Ende des Verfahrens müssen wir sicherstellen, dass unsere physikalisch relevanten Lösungen wieder rein reell sind. Als zugehörige Eigenvektoren bestimmen wir  $\begin{pmatrix} \mp i \\ \sqrt{c} \end{pmatrix} \in \mathbb{C}^2$ . Das liefert uns eine Basiswechsellmatrix und die Darstellung

$$\begin{pmatrix} -i & i \\ \sqrt{c} & \sqrt{c} \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 & 1 \\ -c & 0 \end{pmatrix} \cdot \begin{pmatrix} -i & i \\ \sqrt{c} & \sqrt{c} \end{pmatrix} = \begin{pmatrix} i\sqrt{c} & 0 \\ 0 & -i\sqrt{c} \end{pmatrix}.$$

Wir lösen das neue Gleichungssystem

$$(\ddagger) \quad \begin{pmatrix} \dot{f} \\ \dot{g} \end{pmatrix} = \begin{pmatrix} i\sqrt{c} & 0 \\ 0 & -i\sqrt{c} \end{pmatrix} \cdot \begin{pmatrix} f \\ g \end{pmatrix} \quad \text{und erhalten} \quad \begin{aligned} f(t) &= z e^{it\sqrt{c}}, \\ g(t) &= w e^{-it\sqrt{c}}, \end{aligned}$$

für beliebige Konstanten  $z, w \in \mathbb{C}$ . Mit der obigen Basiswechsellmatrix liefert das komplexe Lösungen von  $(\dagger)$  der Form

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} -i & i \\ \sqrt{c} & \sqrt{c} \end{pmatrix} \cdot \begin{pmatrix} z e^{it\sqrt{c}} \\ w e^{-it\sqrt{c}} \end{pmatrix} = \begin{pmatrix} -iz e^{it\sqrt{c}} + iw e^{-it\sqrt{c}} \\ z\sqrt{c} e^{it\sqrt{c}} + w\sqrt{c} e^{-it\sqrt{c}} \end{pmatrix}.$$

Wie erwartet gilt  $\dot{u} = v$  und  $\ddot{u} = -cu$ . Wir erhalten reelle Lösungen, falls  $w = \bar{z}$ . Wir setzen  $z = \frac{ri}{2} e^{i\varphi}$  mit  $r, \varphi \in \mathbb{R}$  und  $w = \bar{z} = -\frac{ri}{2} e^{-i\varphi}$  und schreiben

$$u(t) = r \frac{e^{i(\varphi+t\sqrt{c})} + e^{-i(\varphi+t\sqrt{c})}}{2} = r \cos(\varphi + t\sqrt{c})$$

und

$$v(t) = \dot{u}(t) = r\sqrt{c} \frac{-e^{i(\varphi+t\sqrt{c})} + e^{-i(\varphi+t\sqrt{c})}}{2i} = -r\sqrt{c} \sin(\varphi + t\sqrt{c}).$$

Die Lösungen sind also Schwingungen, wobei die Amplitude  $r$  und die Phase  $\varphi$  als freie Parameter auftreten. Wenn Ort und Geschwindigkeit zur Zeit  $t_0$  gegeben sind, kann man  $z$  und  $w$  leicht mit Hilfe eines linearen Gleichungssystems bestimmen und erhält daraus  $r$  und  $\varphi$ .

Es sind also mehrere Schritte nötig, um die Differentialgleichung (\*) zu lösen. Als erstes wandeln wir sie in ein System von Differentialgleichungen erster Ordnung um. Als nächstes „diagonalisieren“ wir dieses System. Das geht so einfach nur über  $\mathbb{C}$  und bei konstanter Koeffizientenmatrix — sonst müssten wir nämlich auch die Ableitung der Basiswechselmatrix mit berücksichtigen. Jetzt haben wir entkoppelte Differentialgleichungen an zwei Funktionen, hier  $f$  und  $g$ , die wir einzeln lösen können. Zum Schluss müssen wir aus allen möglichen Lösungen für das ursprüngliche System die physikalisch sinnvollen herausuchen — hier die reellwertigen.

### 5.3. Der Satz von Cayley-Hamilton

Wir müssen später immer wieder Matrizen in Polynome einsetzen. Wir überlegen uns zunächst, was das bedeutet, und wie man damit rechnet. Das führt auf den Begriff der Algebra. Danach zeigen wir, dass  $\chi_A(A) = 0 = \chi_F(F)$ . Das führt uns im nächsten Abschnitt auf den Begriff des Minimalpolynoms.

5.17. DEFINITION. Es sei  $R$  ein kommutativer Ring. Eine *Algebra*  $(A, \cdot)$  über  $R$  besteht aus einem  $R$ -Modul  $A$  und einem *bilinearen Produkt*  $\cdot : A \times A \rightarrow A$ , das heißt, für alle  $w \in M$  sind die beiden Abbildungen  $A \rightarrow A$  mit

$$v \longmapsto v \cdot w \quad \text{und} \quad v \longmapsto w \cdot v$$

linear. Eine Algebra  $(A, \cdot)$  heißt *assoziativ (kommutativ)*, wenn ihr Produkt assoziativ (kommutativ) ist. Wenn es ein neutrales Element für das Produkt gibt, nennen wir  $(A, \cdot)$  eine *Algebra mit Eins*.

In der Notation lassen wir das Produkt oft weg, wenn klar ist, welches Produkt wir meinen.

5.18. BEMERKUNG. Wir werden im Folgenden fast immer Algebren mit Eins  $1_A \in A$  betrachten. In diesem Fall existiert eine Abbildung

$$(1) \quad \varphi: R \longrightarrow A \quad \text{mit} \quad \varphi(r) = 1_A \cdot r.$$

Für alle  $r, s \in R$  rechnet man nach, dass

$$(2) \quad \varphi(r + s) = 1_A \cdot (r + s) = 1_A \cdot r + 1_A \cdot s = \varphi(r) + \varphi(s),$$

$$(3) \quad \varphi(rs) = (1_A \cdot r) \cdot s = (1_A \cdot r) \cdot (1_A \cdot s) = \varphi(r) \cdot \varphi(s)$$

$$(4) \quad \text{und} \quad \varphi(1) = 1_A \cdot 1 = 1_A.$$

Die Bildelemente  $\varphi(r)$  kommutieren immer mit allen Elementen aus  $A$ , denn wegen Bilinearität des Produkts gilt

$$(5) \quad \varphi(r) \cdot a = (1_A \cdot r) \cdot a = (1_A \cdot a) \cdot r = a \cdot r = (a \cdot 1_A) \cdot r = a \cdot \varphi(r),$$

vergleiche Bemerkung 4.5.

Wenn  $A$  eine Algebra mit Eins über einem Körper  $\mathbb{k}$  ist, ist die Abbildung  $\varphi$  injektiv, es sei denn, es wäre  $A = \{0\}$ . Denn aus  $\varphi(y) = \varphi(z)$  folgt  $\varphi(y - z) = 0$ . Und aus  $\varphi(x) = 0$  für  $x \neq 0$  folgt

$$1_A = \varphi(1) = \varphi\left(\frac{1}{x} x\right) = \varphi\left(\frac{1}{x}\right) \cdot \varphi(x) = \varphi\left(\frac{1}{x}\right) \cdot 0 = 0,$$

und für beliebige  $a \in A$  somit

$$a = 1_A \cdot a = 0 \cdot a = 0.$$

5.19. BEISPIEL. Wir kennen schon einige Algebren.

- (1) Der Euklidische Raum  $\mathbb{R}^3$  mit dem Kreuzprodukt ist eine Algebra, allerdings ohne Eins. Sie ist weder assoziativ noch kommutativ.
- (2) Wir können sowohl  $\mathbb{C}$  als auch  $\mathbb{H}$  als assoziative Algebren mit Eins über  $\mathbb{R}$  auffassen. Dann ist  $\mathbb{C}$  kommutativ,  $\mathbb{H}$  aber nicht.
- (3) Der Matrixring  $(M_n(R), \cdot)$  über einem kommutativen Ring  $R$  mit Eins aus Folgerung 2.73 (3) ist eine assoziative Algebra mit Eins  $E_n$ .
- (4) Sei  $V$  ein  $\mathbb{k}$ -Vektorraum, dann ist  $(\text{End}_{\mathbb{k}}(V), \circ)$  eine assoziative Algebra mit Eins  $\text{id}_V$  über  $\mathbb{k}$ , siehe Folgerung 2.31 (2). Falls  $n = \dim V < \infty$  gilt, ist sie zur Algebra  $M_n(\mathbb{k})$  aus (3) isomorph.
- (5) Allgemeiner sei  $R$  ein kommutativer Ring mit Eins und  $S$  ein beliebiger Ring mit Eins. Wenn eine Abbildung  $\varphi: R \rightarrow S$  die Bedingungen (2)–(5) aus Bemerkung 5.18 erfüllt, dann ist  $(S, \cdot)$  eine assoziative Algebra mit Eins über  $R$ , wobei  $s \cdot r = s \cdot \varphi(r)$ . (Übung).
- (6) Insbesondere ist der Polynomring  $R[X]$  eine Algebra mit Eins  $1 = 1 X^0$  über  $R$ .

5.20. DEFINITION. Es seien  $A$  und  $B$  Algebren über einem kommutativen Ring  $R$ . Ein *Algebrenhomomorphismus* ist eine  $R$ -lineare Abbildung  $f: A \rightarrow B$ , die *multiplikativ* ist, das heißt, für alle  $a, b \in A$  gilt die Axiome

$$(1) \quad f(a \cdot b) = f(a) \cdot f(b).$$

Seien  $A$  und  $B$  Algebren mit Eins, dann heißt  $f$  *unitär*, wenn zusätzlich

$$(2) \quad f(1_A) = 1_B.$$

Ganz analog würde man auch Ringhomomorphismen  $f: R \rightarrow S$  definieren, allerdings würden wir statt  $R$ -linear dann *additiv* fordern, das heißt, für alle  $r, s \in R$  muss  $f(r + s) = f(r) + f(s)$  gelten, vergleiche (L1).

5.21. PROPOSITION (Universelle Eigenschaft des Polynomrings). *Es sei  $R$  ein kommutativer Ring mit Eins, und  $A$  eine Algebra über  $R$  mit Eins. Dann existiert zu jedem  $a \in A$  ein eindeutiger Algebren-Homomorphismus  $\text{ev}_a: R[X] \rightarrow A$  mit  $\text{ev}_a|_R = \varphi$  und  $\text{ev}_a(X) = a$ .*

$$\begin{array}{ccc} R[X] & \longleftarrow & \{X\} \\ \uparrow & \text{ev}_a \swarrow & \downarrow a \\ R & \xrightarrow{\varphi} & A \end{array}$$

Später schreiben wir für  $\text{ev}_a(P)$  einfach  $P(a)$ . Der Fall  $R = A$  ist ein Spezialfall nach Beispiel 5.19 (5) mit  $\varphi = \text{id}_R: R \rightarrow R$  und beschreibt das Einsetzen von  $r \in R$  aus Beispiel 2.80.

BEWEIS. Wir beweisen wie üblich zunächst die Eindeutigkeit. Dazu sei  $P = \sum_{i=0}^n p_i X^i$  ein Polynom und  $a \in A$ . Wir benutzen die Abbildung  $\varphi$  und die Rechenregeln aus Bemerkung 5.18. Wenn  $\text{ev}_a: R[X] \rightarrow A$  existiert, dann folgt

$$(*) \quad \text{ev}_a(P) = \sum_{i=0}^n \text{ev}_a(p_i X^i) = \sum_{i=0}^n \text{ev}_a(1_A \cdot p_i) \cdot \text{ev}_a(X)^i = \sum_{i=0}^n \varphi(p_i) \cdot a^i.$$

Zur Existenz definieren  $\text{ev}_a(P)$  durch (\*). Dann müssen wir zeigen, dass  $\text{ev}_a$  linear ist und (1) und (2) aus Definition 5.20 erfüllt. Es sei  $P$  wie oben und  $Q(X) = \sum_{i=0}^m q_j X^j$  ein weiteres Polynom. Wir setzen  $p_i = q_j = 0$  für  $i > n$  und  $j > m$ . Dann gilt

$$\begin{aligned} \text{ev}_a(P + Q) &= \sum_{i=0}^{\infty} \varphi(p_i + q_i) a^i = \sum_{i=0}^n \varphi(p_i) a^i + \sum_{i=0}^m \varphi(q_i) a^i \\ &= \text{ev}_a(P) + \text{ev}_a(Q), \\ \text{ev}_a(P \cdot Q) &= \sum_{k=0}^{\infty} \varphi\left(\sum_{i=0}^k p_i q_{k-i}\right) a^k = \sum_{i=0}^n \varphi(p_i) a^i \cdot \sum_{j=0}^m \varphi(q_j) a^j \\ &= \text{ev}_a(P) \cdot \text{ev}_a(Q), \\ \text{ev}_a(1_{R[X]}) &= \sum_{i=0}^{\infty} \varphi(\delta_{i0}) a^i = a^0 = 1. \end{aligned}$$

Die Homogenität (L2) von  $\text{ev}_a$  zeigen wir, indem wir das Produkt mit  $Q = r \in R$  betrachten. Somit ist  $\text{ev}_a$  tatsächlich ein Algebrenhomomorphismus.  $\square$

5.22. SATZ (Cayley-Hamilton). *Es sei  $R$  ein kommutativer Ring mit Eins und  $A \in M_n(R)$ . Dann gilt  $\chi_A(A) = 0$ .*

*Es sei  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}}(V)$ . Dann gilt  $\chi_F(F) = 0$ .*

An manchen Stellen findet sich zu diesem Satz die folgende Heuristik: „Einsetzen von  $A$  in  $\chi_A$  liefert  $\det(A \cdot E_n - A) = 0$ .“ So einfach ist es leider

nicht, denn die beiden  $A$ s in der obigen Formel leben in verschiedenen Ringen. Um das zu verdeutlichen, betrachten wir stattdessen das einfachere Polynom  $P_A(X) = \text{tr}(X \cdot E_n - A) = nX - \text{tr} A$ . Es gilt  $0 = P_A(A) = nA - \text{tr} A \cdot E_n$  genau dann, wenn  $A$  ein Vielfaches der Einheitsmatrix ist, im allgemeinen also nicht. Die obige Heuristik würde aber immer  $P_A(A) = 0$  liefern.

BEWEIS. Wir beweisen die erste Formulierung. Die zweite ergibt sich daraus, indem wir  $F$  bezüglich einer Basis  $B$  durch eine Matrix  $A = {}_B F_B$  darstellen, denn dann gilt  $\chi_A = \chi_F$  und  $\chi_A(A) = {}_B \chi_F(F)_B$ . Zu letzterem betrachte (\*) im Beweis von Proposition 5.21 und benutze, dass  ${}_B (F^k)_B = A^k$ .

Es sei zunächst  $B \in M_n(R[X])$ . Wir erinnern uns an die Adjunkte  $\text{adj } B \in M_n(R[X])$  aus Definition 4.20. Im Beweis der Cramerschen Regel 4.21 (1) haben wir gezeigt, dass

$$\text{adj } B \cdot B = E_n \cdot \det B \in M_n(R[X]).$$

Wir betrachten die spezielle Matrix  $B = X \cdot E_n - A \in M_n(R[X])$  und erhalten

$$\text{adj}(X E_n - A) \cdot (X E_n - A) = E_n \cdot \det(X E_n - A) = E_n \cdot \chi_A(X).$$

Nach Definition 4.20 sind die Einträge der Adjunkten Determinanten von  $(n-1)$ -reihigen Untermatrizen, in diesem Fall also Polynome in  $X$ , in denen  $X^0$  bis  $X^{n-1}$  auftreten können. Wir fassen die Koeffizienten von  $X^i$  jeweils zu einer Matrix  $B_i \in M_n(R)$  zusammen und erhalten

$$\text{adj}(X E_n - A) = \sum_{i=0}^{\infty} B_i \cdot X^i,$$

wobei  $B_i = 0$  für  $i \geq n$ . Außerdem seien  $c_0, \dots, c_n$  die Koeffizienten von  $\chi_A(X)$ , und  $c_i = 0$  für alle  $i > n$ . Dann gilt also

$$\sum_{i=0}^{\infty} B_i \cdot (X E_n - A) \cdot X^i = \sum_{i=0}^n E_n \cdot c_i X^i.$$

Da  $(X^i)_{i \in \mathbb{N}}$  eine Basis von  $R[X]$  über  $R$  bildet, können wir die Koeffizienten von  $X^i$  vergleichen und erhalten in  $M_n(R)$  die Identitäten

$$B_{i-1} - B_i \cdot A = E_n \cdot c_i \quad \text{für alle } i \geq 0,$$

wobei  $B_{-1} = 0$ .

Wir definieren  $\chi_A(A)$  wie in Proposition 5.21 mit  $A = M_n(R)$  und erhalten

$$\begin{aligned} \chi_A(A) &= \sum_{i=0}^{\infty} \varphi(c_i) A^i = \sum_{i=0}^{\infty} (E_n \cdot c_i) \cdot A^i \\ &= \sum_{i=0}^{\infty} (B_{i-1} - B_i \cdot A) A^i = \sum_{i=1}^{\infty} B_{i-1} A^i - \sum_{i=0}^{\infty} B_i A^{i+1} = 0. \quad \square \end{aligned}$$

### 5.4. Das Minimalpolynom

In diesem Abschnitt lernen wir, im Polynomring mit Rest zu dividieren, analog zur Division mit Rest in  $\mathbb{Z}$ . Mit dieser Methode können wir zu jedem Endomorphismus  $F$  eines endlich-dimensionalen  $\mathbb{k}$ -Vektorraums  $V$  ein Minimalpolynom  $\mu_F$  finden, das alle Polynome  $P$  teilt, für die  $P(F) = 0 \in \text{End}_{\mathbb{k}}(V)$ .

5.23. DEFINITION. Es sei  $R$  ein kommutativer Ring mit Eins und

$$P = P(X) = \sum_{i=0}^{\infty} p_i X^i$$

ein Polynom über  $R$  mit  $(p_i)_i \in R^{\mathbb{N}}$ . Der größte Index  $i \in \mathbb{N}$  mit  $p_i \neq 0$  heißt der *Grad*  $\deg P$  von  $P$ ; falls  $p_i = 0$  für alle  $i \in \mathbb{N}$ , setzen wir  $\deg P = -\infty$ . Polynome  $P \in R[X]$  vom Grad  $\deg P \leq 1$  heißen *linear*.

Es sei  $P(X) = \sum_{i=0}^k p_i X^i \in R[X] \setminus \{0\}$  mit  $p_k \neq 0$  ein Polynom vom Grad  $\deg P = k$ , dann heißt  $p_k$  der *Leitkoeffizient* von  $P$ . Ein *normiertes Polynom* ist ein Polynom  $P \neq 0$  mit Leitkoeffizient  $p_k = 1$ .

In Definition 5.11 haben wir  $R$  bereits mit der Teilmenge der konstanten Polynome in  $R[X]$  identifiziert. Für das konstante Polynom  $r \in R$  gilt also  $\deg r = 0$ , falls  $r \neq 0$ , sonst  $\deg r = -\infty$ . Beachte aber: ein lineares Polynom  $P = aX + b$  beschreibt nicht notwendigerweise eine lineare Abbildung  $P: R \rightarrow R$  — das gilt nur, wenn  $b = 0$ .

5.24. BEMERKUNG. Da nur endlich viele  $p_i \neq 0$  sind, ist der Grad  $p_i \in \mathbb{N} \cup \{-\infty\}$  wohldefiniert, und das einzige Polynom vom Grad  $-\infty$  ist das *Nullpolynom*  $0 \in R \subset R[X]$ . Es seien jetzt  $P, Q \in R[X]$  wie oben, mit  $k = \deg P$  und  $\ell = \deg Q$ .

- (1) Das *Maximum*  $\max(k, \ell)$  zweier natürlicher Zahlen ist die größere von beiden. Außerdem setzen wir  $\max(n, -\infty) = \max(-\infty, n) = n$  für alle  $n \in \mathbb{N} \cup \{-\infty\}$ . Dann gilt

$$\deg(P + Q) \leq \max(\deg P, \deg Q) ,$$

denn für alle  $m > \max(k, \ell)$  erhalten wir  $p_m + q_m = 0$  als Koeffizienten von  $X^m$  in  $P + Q$  nach Definition 5.11 (1).

- (2) Wenn

$$\deg(P + Q) < \max(\deg P, \deg Q) ,$$

dann haben beide Polynome den gleichen Grad und die Summe der Leitkoeffizienten von  $P$  und  $Q$  verschwindet. Als Beispiel betrachte  $P = X - 3$  und  $Q = -X + 5$ , dann ist

$$P + Q = (X - 3) + (-X + 5) = 2 .$$

- (3) Wir setzen  $(-\infty) + n = n + (-\infty) = -\infty$  für alle  $n \in \mathbb{N} \cup \{-\infty\}$ . Für das Produkt  $P \cdot Q$  gilt dann

$$\deg(P \cdot Q) \leq \deg P + \deg Q ,$$

denn nach Definition 5.11 (2) ist der Koeffizient von  $X^{k+\ell}$  in  $P \cdot Q$  gerade  $p_k \cdot q_\ell$ , und höhere Potenzen von  $X$  kommen nicht vor.

(4) Der Fall

$$\deg(P \cdot Q) < \deg P + \deg Q$$

kann nur eintreten, wenn das Produkt der Leitkoeffizienten  $p_k \cdot q_\ell$  verschwindet. Nach Voraussetzung ist  $p_k \neq 0 \neq q_\ell$ , also sind  $p_k$  und  $q_\ell$  Nullteiler, siehe Bemerkung 2.13 (2).

(5) Wenn  $R$  nullteilerfrei ist, also nach Bemerkung 2.13 insbesondere, wenn  $R = \mathbb{k}$  ein Körper ist, ist  $R[X]$  wieder nullteilerfrei.

5.25. SATZ (Polynomdivision mit Rest). *Es sei  $R$  ein kommutativer Ring mit Eins, es seien  $P, Q \in R[X]$  Polynome über  $R$ , und  $Q$  sei normiert. Dann existieren eindeutige Polynome  $S, T \in R[X]$ , so dass*

$$(1) \quad P = S \cdot Q + T$$

$$(2) \quad \text{und} \quad \deg T < \deg Q .$$

Dieser Satz ist völlig analog zur Division mit Rest in  $\mathbb{N}$  oder  $\mathbb{Z}$ , siehe Abschnitt 2.1 vor Satz 2.18. Division mit Rest hat nur für wenige Ringe gute Eigenschaften. Diese Ringe heißen „Euklidisch“, da in ihnen der Euklidische Algorithmus funktioniert — siehe kommender Abschnitt.

BEWEIS. Die Existenz von  $S$  und  $T$  beweisen wir durch Induktion über  $k = \deg P$ . Im Falle  $\deg P < \deg Q$  setzen wir  $T = P$  und  $S = 0$  und sind fertig.

Wir nehmen jetzt an, dass  $k = \deg P \geq \deg Q$ , und dass wir alle Polynome vom Grad  $< k$  mit Rest durch  $Q$  dividieren können. Es sei  $\ell = \deg Q$ . Da  $Q$  normiert ist, schreiben wir  $Q$  als

$$Q(X) = X^\ell + \sum_{j=0}^{\ell-1} q_j X^j .$$

Es sei  $p_k \neq 0$  der Leitkoeffizient von  $P$ , dann betrachten wir das Polynom

$$P'(X) = P(X) - p_k X^{k-\ell} \cdot Q(X) \in R[X] .$$

Dann gilt

$$p_k X^{k-\ell} \cdot Q(X) = \sum_{i=0}^{\ell} p_k q_i X^{k-\ell+i} = p_k X^k + \sum_{j=k-\ell}^{k-1} p_k q_{j+\ell-k} X^j ,$$

also verschwindet der Koeffizient vom Grad  $k$  in  $P'$ , so dass  $\deg P' < k$ . Nach Induktionsvoraussetzung existieren  $S', T \in R[X]$  mit  $\deg T < \ell = \deg Q$ , so dass

$$P' = S' \cdot Q + T .$$

Wir setzen  $S(X) = S'(X) + p_k X^{\ell-k}$  und erhalten

$$\begin{aligned} P(X) &= P'(X) + p_k X^{k-\ell} \cdot Q(X) \\ &= (S'(X) + p_k X^{k-\ell}) \cdot Q(X) + T(X) = S(X) \cdot Q(X) + T(X) , \end{aligned}$$

womit die Existenz von  $S$  und  $T$  gezeigt ist.

Um die Eindeutigkeit zu beweisen, nehmen wir an, dass

$$P = S \cdot Q + T = S' \cdot Q + T' \in R[X] \quad \text{mit} \quad \deg T, \quad \deg T' < \deg Q .$$

Dann gilt

$$(*) \quad \deg((S - S') \cdot Q) = \deg(T' - T) < \deg Q ,$$

denn aus Bemerkung 5.24 (1) folgt  $\deg(T - T') \leq \max(\deg T, \deg T') < \deg Q$ . Wir nehmen an, dass  $S - S' \neq 0$ , dann sei  $n = \deg(S - S')$ , und  $s_n$  sei der Leitkoeffizient. Aus Bemerkung 5.24 (3) folgt

$$\deg((S - S') \cdot Q) \leq \deg(S - S') + \deg Q = n + \ell ,$$

und der Koeffizient vor  $X^{\ell+n}$  wird gegeben durch  $s_n \cdot q_\ell = s_n \neq 0$ , so dass

$$\deg((S - S') \cdot Q) = \deg(S - S') + \deg Q \geq \deg Q$$

im Widerspruch zu (\*). Also gilt  $S = S'$ , und Eindeutigkeit folgt, da

$$T' - T = (S - S') \cdot Q = 0 \in R[X] . \quad \square$$

Wenn wir über einem Körper arbeiten, können wir durch beliebige Polynome  $Q \neq 0$  dividieren. Dazu dividieren wir erst alle Koeffizienten von  $Q$  durch den Leitkoeffizienten und erhalten ein normiertes Polynom. Durch dieses Polynom dividieren wir mit Rest wie oben. Am Ende müssen wir dann das Ergebnis  $S$  noch durch den Leitkoeffizienten von  $Q$  teilen. Als Beispiel betrachte  $P = X^2 - 1$  und  $Q = 2X + 1$ . Wir dividieren zunächst durch  $\frac{1}{2}Q$  und erhalten

$$X^2 - 1 = \left(X - \frac{1}{2}\right) \cdot \left(X + \frac{1}{2}\right) - \frac{3}{4} ,$$

also gilt

$$X^2 - 1 = \left(\frac{X}{2} - \frac{1}{4}\right) \cdot (2X + 1) - \frac{3}{4} .$$

Im Folgenden werden wir allerdings meistens durch normierte Polynome dividieren.

5.26. BEISPIEL. Polynomdivision lässt sich schriftlich durchführen, ganz analog zur Division im Dezimalsystem. Wir dividieren  $X^5 - X^4 + 2X^2 - 2X + 1$  durch  $X^4 + X^3 - X^2 + 1$ . Die Rechnung

$$\begin{array}{r} X^5 - X^4 + 2X^2 - 2X + 1 \\ X^5 + X^4 - X^3 + X \\ \hline - 2X^4 + X^3 + 2X^2 - 3X + 1 \\ - 2X^4 - 2X^3 + 2X^2 - 2 \\ \hline 3X^3 - 3X + 3 \end{array}$$

liefert einen Rest von kleinerem Grad als der Divisor und zeigt, dass

$$X^5 - X^4 + 2X^2 - 2X + 1 = (X^4 + X^3 - X^2 + 1) \cdot (X - 2) + (3X^3 - 3X + 3) .$$

Wir schreiben  $Q \mid P$  und sagen „ $Q$  teilt  $P$ “ oder „ $P$  ist Vielfaches von  $Q$ “, wenn die obige Division ohne Rest möglich ist, das heißt, wenn  $S \in R[X]$  existiert, so dass  $P = S \cdot Q$ . Andernfalls schreiben wir  $Q \nmid P$ .

5.27. SATZ. *Es sei  $F \in \text{End}_{\mathbb{k}}(V)$  Endomorphismus eines  $\mathbb{k}$ -Vektorraums  $V$ . Dann gibt es eindeutiges Polynom  $\mu_F \in \mathbb{k}[X]$ , das entweder normiert oder 0 ist, so dass*

$$\{ P \in \mathbb{k}[X] \mid P(F) = 0 \in \text{End}_{\mathbb{k}}(V) \} = \mathbb{k}[X] \cdot \mu_F \subset \mathbb{k}[X] .$$

BEWEIS. Es sei

$$M = \{ P \in \mathbb{k}[X] \mid P(F) = 0 \in \text{End}_{\mathbb{k}}(V) \} \subset \mathbb{k}[X] .$$

Falls  $M = \{0\}$  gilt, setzen wir  $\mu_F = 0$ .

Andernfalls gibt es ein Polynom  $Q \in M \setminus \{0\}$  von kleinstem Grad  $\deg Q \in \mathbb{N}$ . Wir dürfen annehmen, dass  $Q$  normiert ist. Denn sei  $r \in \mathbb{k}^\times = \mathbb{k} \setminus \{0\}$  (siehe Bemerkung 2.13) der Leitkoeffizient von  $Q$ , dann gilt

$$0 = r^{-1} \cdot Q(F) = (r^{-1}Q)(F) \in \text{End}_{\mathbb{k}}(V) ,$$

da Einsetzen nach Proposition 5.21 eine  $\mathbb{k}$ -lineare Abbildung ist. Also ist  $r^{-1}Q$  ein normiertes Polynom vom kleinsten Grad in  $M \setminus \{0\}$ .

Sei also  $Q$  wie oben und normiert. Für ein weiteres Polynom  $P \in M$  liefert Division mit Rest

$$P = S \cdot Q + T \in \mathbb{k}[X]$$

mit  $\deg T < \deg Q$ . Da Einsetzen von  $F$  ein Algebrenhomomorphismus  $\mathbb{k}[X] \rightarrow \text{End}_{\mathbb{k}}(V)$  ist, folgt

$$T(F) = P(F) - S(F) \circ Q(F) = 0 - S(F) \cdot 0 = 0 \in \text{End}_{\mathbb{k}}(V) ,$$

also  $T \in M$ . Wäre  $T \neq 0$ , dann hätte  $Q$  nicht den kleinstmöglichen Grad in  $M \setminus \{0\}$ , im Widerspruch zur Wahl von  $Q$ . Es folgt  $Q \mid P$  für alle  $P \in M$ , also  $M \subset \mathbb{k}[X] \cdot Q$ .

Umgekehrt sei  $P = S \cdot Q$ , dann folgt  $P \in M$ , da

$$P(F) = S(F) \circ Q(F) = S(F) \circ 0 = 0 \in \text{End}_{\mathbb{k}}(V) ,$$

also gilt auch  $\mathbb{k}[X] \cdot Q \subset M$ . □

Wir haben ausgenutzt, dass die Menge  $M \subset \mathbb{k}[X]$  ein Untermodul des eindimensionalen  $\mathbb{k}[X]$ -Moduls  $\mathbb{k}[X]$  ist. Solche Teilmengen eines Ringes nennt man Ideale. In einer Vorlesung über Algebra lernen Sie, dass Euklidische Ringe (wie  $\mathbb{Z}$  oder  $\mathbb{k}[X]$ ) Hauptidealringe sind, das heißt, jedes Ideal wird von einem einzelnen Element erzeugt. Das Argument benutzt die Grad-Funktion und Division mit Rest genau wie im obigen Beweis.

5.28. DEFINITION. Das Polynom  $\mu_F \in \mathbb{k}[X]$  heißt das *Minimalpolynom* von  $F$ .

5.29. FOLGERUNG (aus dem Satz 5.22 von Cayley-Hamilton). *Sei  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}}(V)$ , dann gilt  $\mu_F \mid \chi_F$ , insbesondere ist  $\mu_F \neq 0$  und  $\deg \mu_F \leq \dim V$ .*

BEWEIS. Nach dem Satz von Cayley-Hamilton ist  $\chi_F$  in der Menge  $M$  aus obigem Beweis enthalten. Also gilt  $\mu_F \mid \chi_F$ . Daraus folgt  $\mu_F \neq 0$  und  $\deg \mu_F \leq \deg \chi_F = \dim V$ . □

### 5.5. Der Satz von der eindeutigen Primfaktorzerlegung

Wir wiederholen den Euklidischen Algorithmus 2.18 und benutzen ihn, um Polynome auf eindeutige Weise in Primfaktoren zu zerlegen. All das werden wir im folgenden Abschnitt benutzen, um einen Vektorraum  $V$  mit Endomorphismus  $F$  in eine direkte Summe invarianter Unterräume zu zerlegen, auf denen wir  $F$  besser beschreiben können. Wir erinnern uns an die Teilmenge  $R^\times \subset R$  der Einheiten, das heißt, der multiplikativ invertierbaren Elemente, siehe Folgerung 4.21. In einem Körper gilt  $\mathbb{k}^\times = \mathbb{k} \setminus \{0\}$ . In  $\mathbb{Z}$  analog  $\mathbb{Z}^\times = \{\pm 1\}$ .

5.30. SATZ (Erweiterter Euklidischer Algorithmus für Polynome). *Es seien  $P_0, P_1 \in \mathbb{k}[X]$  normierte Polynome mit  $\deg P_1 \leq \deg P_0$ , Dann existieren eine eindeutige Zahl  $i_0 \in \mathbb{N}$ , normierte Polynome  $P_2, \dots, P_{i_0}$  und  $S_2, \dots, S_{i_0+1} \in \mathbb{k}[X]$  mit  $\deg P_1 > \dots > \deg P_{i_0}$  und Zahlen  $r_2, \dots, r_{i_0} \in \mathbb{k}^\times$ , so dass*

$$(1) \quad P_{i-1} = S_{i+1}P_i + r_{i+1}P_{i+1} \quad \text{für } 1 \leq i < i_0 \quad \text{und} \quad P_{i_0-1} = S_{i_0+1}P_{i_0}.$$

*Dann ist  $P_{i_0} = \text{ggT}(P_0, P_1)$  das eindeutige normierte Polynom vom höchsten Grad, das sowohl  $P_0$  als auch  $P_1$  teilt.*

*Setze  $Q_{i_0} = 0$ ,  $Q_{i_0-1} = r_{i_0}^{-1}$  und  $r_1 = 1$ , und bestimme der Reihe nach  $Q_{i_0-2}, \dots, Q_0 \in \mathbb{k}[X]$  so, dass*

$$(2) \quad Q_{i-1} = \frac{1}{r_i} (Q_{i+1} - Q_i S_{i+1}) \quad \text{für } i_0 > i \geq 1.$$

*Dann gilt eine Bézout-Identität  $P_{i_0} = Q_1 P_0 + Q_0 P_1$ .*

BEWEIS. Wir passen den Beweis von Satz 2.18 an (Übung). Die induktive Behauptung zu Teil (2) lautet  $P_{i_0} = Q_i P_{i-1} + r_i Q_{i-1} P_i$  für  $i_0 \geq i \geq 1$ .  $\square$

Dass die Grad-Funktion Werte in  $\mathbb{N}$  hat, ist im Beweis wichtig, denn jede strikt fallende Folge natürlicher Zahlen muss nach endlich vielen Schritten aufhören. Für  $\mathbb{Z}$  haben wir in Satz 2.18 als Grad-Funktion den Absolutbetrag  $|\cdot|: \mathbb{Z} \rightarrow \mathbb{N}$  verwendet.

5.31. BEISPIEL. Wir führen Beispiel 5.26 weiter. Dabei sehen wir auch, dass wir die Zahlen  $r_i$  benötigen, wenn wir mit normierten Polynomen arbeiten wollen. Es seien

$$P_0 = X^5 - X^4 + 2X^2 - 2X + 1 \quad \text{und} \quad P_1 = X^4 + X^3 - X^2 + 1.$$

Dann rechnen wir

$$\begin{aligned} X^5 - X^4 + 2X^2 - 2X + 1 &= (X^4 + X^3 - X^2 + 1)(X - 2) + 3(X^3 - X + 1), \\ X^4 + X^3 - X^2 + 1 &= (X^3 - X + 1)(X + 1) + 0, \end{aligned}$$

also erhalten wir  $i_0 = 2$ ,  $r_2 = 3$  und

$$P_2 = X^3 - X + 1, \quad P_3 = 0, \quad S_2 = X - 2, \quad S_3 = X + 1.$$

Mit  $Q_2 = 0$ ,  $Q_1 = \frac{1}{r_2} = \frac{1}{3}$  und  $r_1 = 1$  berechnen wir

$$Q_0 = \frac{1}{r_1} (Q_2 - Q_1 S_2) = -\frac{1}{3} (X - 2) = -\frac{1}{3} X + \frac{2}{3}.$$

In der Tat gilt

$$\begin{aligned} Q_1 P_0 + Q_0 P_1 &= \frac{1}{3} (X^5 - X^4 + 2X^2 - 2X + 1) \\ &\quad + \left( -\frac{1}{3} X + \frac{2}{3} \right) (X^4 + X^3 - X^2 + 1) \\ &= X^3 - X + 1 = P_2 . \end{aligned}$$

Wir hatten im Zusammenhang mit Proposition 2.16 eine etwas andere Definition einer Primzahl kennengelernt als in der Schule. Wir wollen diese zwei Begriffe einander gegenüberstellen.

5.32. DEFINITION. Es sei  $R$  ein nullteilerfreier Ring. Ein Element  $0 \neq r \in R \setminus R^\times$  heißt

- (1) *prim* oder *Primelement*, wenn für alle  $s, t \in R$  aus  $r \mid st$  folgt, dass  $r \mid s$  oder  $r \mid t$ , und
- (2) *irreduzibel*, wenn für alle  $s, t \in R$  aus  $r = st$  folgt, dass  $s \in R^\times$  oder  $t \in R^\times$ .

In Körpern gilt  $\mathbb{k}^\times = \mathbb{k} \setminus \{0\}$ , also gibt es weder prime noch irreduzible Elemente. Es gilt  $\mathbb{Z}^\times = \{\pm 1\}$ . Sei  $n \in \mathbb{Z}$  irreduzibel, dann gilt zunächst  $n \notin \{0, \pm 1\}$ . Falls  $n = st$  und  $s \in \mathbb{Z}^\times$ , folgt  $t = \pm n$ . Das ist genau der Begriff einer Primzahl, wie wir ihn aus der Schule kennen. Euklids Lemma zeigt, dass die obigen Begriffe in Euklidischen Ringen wie  $\mathbb{Z}$  und  $\mathbb{k}[X]$  äquivalent sind.

5.33. LEMMA (Euklid). *Es sei  $R = \mathbb{Z}$  oder  $R = \mathbb{k}[X]$ , dann ist  $r \in R$  genau dann prim, wenn es irreduzibel ist.*

BEWEIS. Zu „ $\implies$ “ gelte  $r = st$ . Da  $r \mid st$  und  $r$  prim ist, teilt  $r$  einen der Faktoren, ohne Einschränkung  $r \mid s$ . Also existiert  $u \in R$  mit  $s = ru$ , somit  $r = st = rut$ . Da  $R$  nullteilerfrei ist, gilt nach Bemerkung 2.13 die Kürzungsregel. Es folgt  $1 = ut$ , das heißt,  $u, t \in R^\times$ , was zu zeigen war.

Zu „ $\impliedby$ “ nehmen wir an, dass  $r$  irreduzibel ist und  $st$  teilt, aber nicht  $s$ . Es sei  $q$  der größte gemeinsame Teiler von  $r$  und  $s$ . Schreibe  $r = pq$  mit  $p \in R$ . Da  $r$  irreduzibel ist, gibt es zwei Möglichkeiten.

Entweder gilt  $p \in R^\times$ . Dann folgt  $r \mid q = p^{-1}r$ .

Oder es gilt  $q \in R^\times$ . Falls  $q = 1$ , sei  $1 = ar + bs$  eine Bézout-Identität nach dem erweiterten Euklidischen Algorithmus 2.18 beziehungsweise 5.30. Falls  $q \neq 1$ , multiplizieren wir die entsprechende Darstellung von  $q$  mit  $q^{-1}$  und erhalten ebenfalls  $1 = ar + bs$ . Aus  $r \mid st$  folgt offensichtlich

$$r \mid art + bst = t . \quad \square$$

Im Folgenden definieren wir das leere Produkt als 1.

5.34. SATZ (Eindeutige Primfaktorzerlegung). *Jede ganze Zahl  $n \neq 0$  lässt sich schreiben als*

$$(1) \quad n = \pm p_1^{k_1} \cdots p_\ell^{k_\ell},$$

dabei sind  $p_1, \dots, p_\ell \in \mathbb{N}$  Primzahlen und  $k_1, \dots, k_\ell \in \mathbb{N} \setminus \{0\}$ .

Jedes Polynom  $Q \in \mathbb{k}[X] \setminus \{0\}$  lässt sich schreiben als

$$(2) \quad Q = r \cdot P_1^{k_1} \cdots P_\ell^{k_\ell},$$

dabei ist  $r \in \mathbb{k}^\times$ ,  $P_1, \dots, P_\ell$  sind irreduzible normierte Polynome und  $k_1, \dots, k_\ell \in \mathbb{N} \setminus \{0\}$ .

In beiden Fällen ist die Darstellung bis auf die Reihenfolge der Faktoren eindeutig. Wir schreiben  $\text{ord}_{p_i}(n) = k_i$  ( $\text{ord}_{P_i}(Q) = k_i$ ), und  $\text{ord}_p(n) = 0$  ( $\text{ord}_P(Q) = 0$ ), falls der Primfaktor  $p$  ( $P$ ) in der Zerlegung nicht vorkommt.

Der folgende Beweis ist nicht konstruktiv. Tatsächlich gibt es kein Verfahren, dass die Zerlegung (2) immer findet, das lernen Sie später in der Vorlesung Algebra und Zahlentheorie. Die Zerlegung (1) findet man durch systematisches Probieren. Auch hier ist kein effizienter Algorithmus bekannt — darauf beruht das RSA-Verfahren in der Kryptographie.

BEWEIS. Beide Beweise sind formal identisch, wir führen den Beweis daher nur für Polynome über  $\mathbb{k}$  vor. In (2) ist  $r$  offensichtlich der Leitkoeffizient von  $Q$ . Wir nehmen ab jetzt an, dass  $Q$  normiert ist, so dass wir  $r = 1$  weglassen dürfen.

Zur Existenz der Zerlegung schreiben wir  $Q$  als Produkt von Polynomen vom Grad  $\geq 1$ , falls möglich. Wir dürfen annehmen, dass die Faktoren wieder normiert sind. Wegen Bemerkung 5.24 (3), (4) haben beide Faktoren kleineren Grad als  $Q$ . Wir können also jeden Faktor weiter zerlegen, falls möglich. Da wir insgesamt höchstens  $\deg Q$  viele Faktoren bekommen können, erhalten wir nach endlich vielen Schritten eine Zerlegung (2) in irreduzible normierte Polynome.

Zur Eindeutigkeit sei etwa

$$Q = S_1^{n_1} \cdots S_m^{n_m}$$

eine weitere Zerlegung in irreduzible normierte Polynome. Nach Euklids Lemma sind die  $S_i$  auch prim. Schreibe  $Q = P_1^{k_1} \cdot R$ , dann folgt  $S_i \mid P_1^{k_1}$  oder  $S_i \mid R$  aus  $S_i \mid Q$ . Aus  $S_i \mid P_1^{k_1}$  folgt  $S_i \mid P_1$ , da  $S_i$  prim ist. Da  $P_1$  auch irreduzibel ist, unterscheiden sich die beiden Polynome höchstens um eine Einheit  $s \in \mathbb{k}^\times$ . Da beide normiert sind, folgt  $s = 1$ , also  $S_i = P_1$ . Wenn  $S_i \mid R$  gilt, machen wir mit  $R = P_2^{k_2} \cdots P_\ell^{k_\ell}$  an Stelle von  $Q$  weiter. Aus  $S_i \mid Q$  folgt also  $S_i = P_j$  für ein  $j$ .

Nun ist der Polynomring nullteilerfrei, das heißt, nach Bemerkung 2.13 gilt die Kürzungsregel. Wir erhalten also zwei Zerlegungen des normierten Polynoms  $Q'$  mit  $Q = Q'S_i = Q'P_j$ . Es hat kleineren Grad als  $Q$ . Wir fahren solange fort, bis nur noch ein normiertes Polynom vom Grad 0 übrig ist, also 1. Das zeigt, dass beide Zerlegungen die gleichen Faktoren mit den gleichen Exponenten enthalten, also bis auf die Reihenfolge der Faktoren gleich sind.  $\square$

5.35. BEMERKUNG. Normierte Polynome vom Grad 1 nennt man *Linearfaktoren*. Sie haben die Gestalt  $X - x$  für ein  $x \in \mathbb{k}[X]$  und verschwinden genau an der Stelle  $x$ . Für  $\text{ord}_{X-\lambda}(Q)$  schreiben wir auch kurz  $\text{ord}_\lambda(Q)$ .

- (1) Linearfaktoren sind irreduzibel. Denn sei  $X - x = RS$  für  $R, S \in \mathbb{k}[X]$ , dann hat wegen Bemerkung 5.24 (3), (4) einer der Faktoren Grad 0, liegt also in  $\mathbb{k}^\times$  und ist somit auch in  $\mathbb{k}[X]$  invertierbar. Also ist  $X - x$  irreduzibel und nach Euklids Lemma 5.33 auch prim.
- (2) Es sei  $P$  ein beliebiges Polynom. Wir dividieren mit Rest durch  $X - x$  und erhalten

$$P = S \cdot (X - x) + T \quad \text{mit} \quad \deg T \leq 0,$$

insbesondere gilt  $T = c \in \mathbb{k}$ . Einsetzen von  $x$  liefert

$$P(x) = S(x) \cdot 0 + T = c,$$

somit gilt  $X - x \mid P$  genau dann, wenn  $P(x) = 0$ . Tatsächlich gibt uns Polynomdivision ein effizientes Verfahren, um  $P(x)$  auszurechnen — die Berechnung geht schneller als bei naivem Einsetzen, siehe Übungen.

- (3) Wegen des Satzes 5.41 von der eindeutigen Primfaktorzerlegung kann ein Polynom  $P$  vom Grad  $n \geq 0$  durch höchstens  $n$  Linearfaktoren teilbar sein. Also kann  $P$  höchstens  $n$  paarweise verschiedene Nullstellen haben. Auch das haben wir in den Übungen schon auf anderem Wege bewiesen.

5.36. BEISPIEL. Wir betrachten speziell  $\mathbb{k}[X]$  für  $\mathbb{k} = \mathbb{R}$  und  $\mathbb{C}$ .

- (1) Nach dem Fundamentalsatz 1.62 der Algebra hat jedes Polynom  $P \in \mathbb{C}[X]$  vom Grad  $\deg P \geq 1$  mindestens eine Nullstelle  $z_0$ . Nach obiger Bemerkung 5.35 (2) gilt  $P = (X - z_0)S$  für ein  $S \in \mathbb{C}[X]$ . Induktiv folgt, dass sich  $P$  als Produkt von Linearfaktoren (und gegebenenfalls einem Leitkoeffizienten aus  $\mathbb{C}^\times$ ) schreiben lässt. Insbesondere sind die normierten irreduziblen Polynome über  $\mathbb{C}$  genau die Linearfaktoren.
- (2) Wenn ein reelles Polynom  $P$  Nullstellen in  $\mathbb{R}$  hat, können wir wie oben Linearfaktoren abspalten. Es sei also  $P \in \mathbb{R}[X]$  vom Grad  $\deg P \geq 1$  ein Polynom ohne reelle Nullstellen.

Wir fassen  $P$  als komplexes Polynom mit reellen Koeffizienten auf, dann hat  $P$  mindestens eine Nullstelle  $z_0 \in \mathbb{C} \setminus \mathbb{R}$ . Da  $P$  reelle Koeffizienten hat, folgt aber auch

$$P(\bar{z}_0) = \overline{P(z_0)} = 0,$$

also existiert  $Q \in \mathbb{C}[X]$ , so dass  $P = (X - z_0)(X - \bar{z}_0)Q$ . Sei  $z_0 = x + iy$ , dann existieren  $a, b \in \mathbb{R}$ , so dass

$$\begin{aligned} (X - z_0)(X - \bar{z}_0) &= (X - x_0 - iy_0)(X - x_0 + iy_0) = (X - x_0)^2 + y_0^2 \\ &= X^2 - 2x_0X + (x_0^2 + y_0^2) = X^2 + aX + b, \end{aligned}$$

wobei  $b > a^2/4$ . Division mit Rest durch obiges Polynom zeigt, dass  $Q \in \mathbb{R}[X]$ .

Somit gibt es über  $\mathbb{R}$  zwei Sorten irreduzibler Polynome:

- (a) Linearfaktoren, und
- (b) Quadratische Polynome  $X^2 + aX + b$  mit  $b > a^2/4$ .

Letzteres sind genau die quadratischen Polynome, die nach der „Mitternachtsformel“ keine reelle Nullstellen haben.

5.37. BEMERKUNG. Wir haben in Beispiel 2.17 endliche Körper  $\mathbb{Z}/p\mathbb{Z}$  konstruiert, indem wir  $Z$  durch die Äquivalenzrelation

$$m \sim n \iff n - m \in p\mathbb{Z}$$

geteilt haben, wobei wir  $p \in \mathbb{Z}$  prim gewählt haben. Völlig analog sei  $P \in \mathbb{k}[X]$  prim, dann ist  $\mathbb{k}[X]/P \cdot \mathbb{k}[X]$  ebenfalls ein Körper (Übung). In einer Algebra-Vorlesung lernen Sie mehr über solche Körper.

- (1) Wenn  $P = X - x$  ein Linearfaktor ist, ist  $\mathbb{k}[X]/P \cdot \mathbb{k}[X] \cong \mathbb{k}$ . Insbesondere können wir aus  $\mathbb{C}$  auf diese Weise keine neuen Körper konstruieren.
- (2) Sei  $\mathbb{k} = \mathbb{R}$  und  $P$  irreduzibel und quadratisch wie in Beispiel 5.36 (2b). Dann gilt  $\mathbb{R}[X]/P \cdot \mathbb{R}[X] \cong \mathbb{C}$ .
- (3) Man erhält alle endlichen Körper der Charakteristik  $p$  auf diese Weise aus  $F_p = \mathbb{Z}/p\mathbb{Z}$ , siehe Algebra-Vorlesung.

## 5.6. Die Hauptraumzerlegung

Wir betrachten wieder einen Endomorphismus  $F$  eines endlich-dimensionalen Vektorraums  $V$ . Ausgehend von der Primfaktorzerlegung des Minimalpolynoms  $\mu_F$  zerlegen wir  $V$  in eine direkte Summe invarianter Unterräume, der sogenannten Haupträume. Im nächsten Abschnitt stellen wir den Endomorphismus dann auf jedem Hauptraum durch eine standardisierte Matrix dar.

5.38. BEMERKUNG. In Beispiel 2.21 (3) haben wir gesehen, dass ein  $\mathbb{Z}$ -Modul genau das gleiche ist wie eine abelsche Gruppe. Analog dazu ist ein  $\mathbb{k}[X]$ -Modul genau das gleiche wie ein  $\mathbb{k}$ -Vektorraum  $V$  mit einem Endomorphismus  $F \in \text{End}_{\mathbb{k}}(V)$ .

Zum einen ist  $V$  ein Links- $\text{End}_{\mathbb{k}}(V)$ -Modul, und die Abbildung  $\text{ev}_F: \mathbb{k}[X] \rightarrow \text{End}_{\mathbb{k}}(V)$  aus Proposition 5.21 macht aus  $V$  einen  $\mathbb{k}[X]$ -Modul via

$$P \cdot v = \text{ev}_F(P)(v) = P(F)(v) \quad \text{für alle } P \in \mathbb{k}[X] \text{ und alle } v \in V.$$

Zum anderen sei  $M$  ein  $\mathbb{k}[X]$ -Modul. Indem wir die Multiplikation mit Skalaren auf  $\mathbb{k} \subset \mathbb{k}[X]$  einschränken, wird daraus ein  $\mathbb{k}$ -Modul, also ein  $\mathbb{k}$ -Vektorraum. Und wir definieren  $F(v) = X \cdot v$  für alle  $v \in V$ . Aus der Eindeutigkeitsaussage in Proposition 5.21 folgt, dass diese beiden Konstruktionen zueinander invers sind, und wir schreiben fortan  $M = (V, F)$ .

Seien  $M = (V, F)$  und  $N = (W, G)$  zwei  $\mathbb{k}[X]$ -Moduln. Dann ist eine  $\mathbb{k}[X]$ -lineare Abbildung von  $M$  nach  $N$  gerade eine  $\mathbb{k}$ -lineare Abbildung  $L: M \rightarrow N$ ,

so dass zusätzlich  $L \circ F = L(X \cdot) = X \cdot L(\cdot) = G \circ L$  gilt, mit anderen Worten kommutiert das Diagramm

$$\begin{array}{ccc} V & \xrightarrow{L} & W \\ F \downarrow & & \downarrow G \\ V & \xrightarrow{L} & W. \end{array}$$

Ein Untervektorraum  $U \subset V$  ist genau dann ein  $\mathbb{k}[X]$ -Untermodul von  $M = (V, F)$ , wenn  $U$  außerdem abgeschlossen ist unter Multiplikation mit Elementen aus  $\mathbb{k}[X]$ . Das gilt bereits, wenn  $Xu \in U$ , das heißt  $F(u) \in U$ , für alle  $u \in U$ . Solche Unterräume nennt man *F-invariant*.

Die Multiplikation  $m_P = P(F): M \rightarrow M$  mit  $P \in \mathbb{k}[X]$  ist eine  $\mathbb{k}[X]$ -lineare Abbildung, da  $\mathbb{k}[X]$  kommutativ ist. Insbesondere ist  $m_P$  homogen (L2), denn  $m_P(Qv) = PQ \cdot v = QP \cdot v = Q \cdot m_P(v)$ . Also sind  $\ker(m_P) = \ker(P(F))$  und  $\operatorname{im}(m_P) = \operatorname{im}(P(F))$  Untermoduln von  $M$ , siehe Proposition 2.37, das heißt, *F*-invariante Untervektorräume von  $V$ .

Zu guter Letzt sei  $U \subset V$  ein invarianter Untervektorraum, also ein  $\mathbb{k}[X]$ -Untermodul. Wir betrachten den Quotienten  $V/U$ , dazu brauchen wir nur die zugrundeliegenden abelsche Gruppe  $(V, +)$ . Insbesondere hängt der Quotient als abelsche Gruppe nicht davon ab, ob wir mit  $\mathbb{k}$ -Vektorräumen oder mit  $\mathbb{k}[X]$ -Moduln arbeiten, siehe Definition 2.38. Nach Proposition 2.39 trägt der Quotient  $V/U$  ebenfalls eine  $\mathbb{k}[X]$ -Modulstruktur. Mit anderen Worten induziert  $F \in \operatorname{End}_{\mathbb{k}}(V)$  einen Endomorphismus  $\bar{F} \in \operatorname{End}_{\mathbb{k}}(V/U)$ .

**5.39. PROPOSITION.** *Es sei  $M = (V, F)$  ein  $\mathbb{k}[X]$ -Modul und  $P = QR$  ein Produkt teilerfremder Polynome  $Q$  und  $R$ , so dass  $Pv = 0$  für alle  $v \in V$ . Dann gilt*

- (1)  $\ker(m_Q) = \operatorname{im}(m_R)$ ,  $\operatorname{im}(m_Q) = \ker(m_R)$ ,
- (2) und  $U = \ker(m_Q) \oplus \ker(m_R)$ .

**BEWEIS.** Sei  $Rw \in \operatorname{im}(m_R)$  für ein  $w \in V$ . Dann gilt  $Q(Rw) = Pw = 0$ , somit  $\operatorname{im}(m_R) \subset \ker(m_Q)$ . Mit dem erweiterten Euklidischen Algorithmus finden wir eine Bézout-Identität  $1 = SQ + TR \in \mathbb{k}[X]$  mit  $S, T \in \mathbb{k}[X]$ . Für  $v \in \ker(m_Q)$  folgt

$$v = (SQ + RT)v = S(Qv) + R(Tv) = R(Tv) \in \operatorname{im}(m_R).$$

Somit gilt auch  $\ker(m_Q) \subset \operatorname{im}(m_R)$ . Die zweite Aussage in (1) folgt analog.

Sei jetzt  $v \in V$  beliebig. Mit der obigen Bézout-Identität erhalten wir eine Zerlegung

$$v = (QS + RT)v = Q(Sv) + R(Tv) \in \operatorname{im}(m_Q) + \operatorname{im}(m_R),$$

somit ist  $V$  die Summe der genannten Unterräume. Für  $v \in \operatorname{im}(m_Q) \cap \operatorname{im}(m_R) = \ker(m_R) \cap \ker(m_Q)$  folgt hingegen

$$v = (SQ + TR)v = S(Qv) + T(Rv) = 0,$$

also ist die Summe in (2) auch direkt. □

Wir übertragen dieses Resultat auf  $\mathbb{Z}$ -Moduln, also abelsche Gruppen. Es sei also  $A$  eine abelsche Gruppe und  $n = qr \in \mathbb{N}$  ein Produkt teilerfremder natürlicher Zahlen, so dass die Multiplikation  $m_n$  mit  $n$  auf  $A$  als Nullabbildung wirkt. Dann liefert das obige Argument eine völlig analoge Zerlegung

$$(*) \quad A = \ker(m_q) \oplus \ker(m_r) .$$

Wir betrachten jetzt den Spezialfall  $A = \mathbb{Z}/n\mathbb{Z}$ , siehe Beispiel 2.40, und leiten aus (\*) den chinesischen Restsatz her. Er geht zurück auf ein chinesisches Manuscript, vermutlich aus dem dritten Jahrhundert.

5.40. SATZ (Chinesischer Restsatz). *Es seien  $n_1, \dots, n_\ell$  paarweise teilerfremde natürliche Zahlen, dann existiert für jedes Tupel ganzer Zahlen  $a_1, \dots, a_\ell$  eine ganze Zahl  $a$ , die die folgende simultane Kongruenz erfüllt:*

$$a \equiv a_i \pmod{n_i} \quad \text{für } i = 1, \dots, \ell .$$

*Alle Lösungen dieser Kongruenz sind kongruent modulo  $n_1 \cdots n_\ell$ .*

BEWEIS. Es sei  $n = n_1 \cdots n_\ell$ . Wir erinnern uns an die direkte Summe mehrerer Untermoduln aus Definition 5.7 und zeigen induktiv, dass

$$(\dagger) \quad \mathbb{Z}/n\mathbb{Z} \cong \ker(m_{n_1}) \oplus \cdots \oplus \ker(m_{n_\ell}) .$$

Falls  $\ell = 0$  gilt  $n = 1$  und somit  $\mathbb{Z}/n\mathbb{Z} = \{0\}$ . Also ist nichts zu zeigen.

Für  $\ell \geq 1$  nehmen wir an, dass die Aussage für  $\ell - 1$  bereits gezeigt ist. Dann schreiben wir  $n = n_1 \cdot r \in \mathbb{k}[X]$  mit  $r = n_2 \cdots n_\ell$ . Nach Voraussetzung sind  $n_1$  und  $r$  teilerfremd, und (\*) liefert eine Zerlegung  $\mathbb{Z}/n\mathbb{Z} = A_1 \oplus B$  mit  $A_1 = \ker(m_{n_1})$ . Nach Induktionsvoraussetzung erhalten wir eine analoge Zerlegung  $B = A_2 \oplus \cdots \oplus A_\ell$ .

Dann ist auch die Summe  $\mathbb{Z}/n\mathbb{Z} = A_1 \oplus \cdots \oplus A_\ell$  direkt. Denn sei  $0 = [b_1] + \cdots + [b_\ell] \in \mathbb{Z}/n\mathbb{Z}$  mit  $[b_i] \in A_i$  für alle  $i$ , dann folgt  $[b_1] = 0 \in \mathbb{Z}/n_1\mathbb{Z}$  und  $[b_2] + \cdots + [b_\ell] \in \mathbb{Z}/r\mathbb{Z}$  aus (\*) und  $[b_i] = 0 \in \mathbb{Z}/n_i\mathbb{Z}$  für  $i \geq 2$  nach Induktionsvoraussetzung.

Sei jetzt  $[b] \in A_1 = \ker(m_{n_1})$ , das heißt  $bn_1 = cn$  für ein  $c \in \mathbb{Z}$ . Mit der Kürzungsregel folgt  $b = cn_2 \cdots n_\ell$ , somit  $A_1 = (n_2 \cdots n_\ell)\mathbb{Z}/n\mathbb{Z}$ . Da  $n_1$  zu  $n_2 \cdots n_\ell$  teilerfremd ist, erhalten wir eine Verkettung von Isomorphismen

$$\mathbb{Z}/n_1\mathbb{Z} \xrightarrow{\cdot n_2 \cdots n_\ell} A_1 \xrightarrow{\text{mod } n_1} \mathbb{Z}/n_1\mathbb{Z} .$$

Somit existiert zu jedem  $[a_1] \in \mathbb{Z}/n_1\mathbb{Z}$  genau ein Element  $[b_1] \in A_1 \subset \mathbb{Z}/n\mathbb{Z}$ , so dass  $b_1 \equiv a_1 \pmod{n_1}$ . Außerdem gilt  $b_1 \equiv 0 \pmod{n_i}$  für  $i \neq 1$ .

Entsprechend bestimmen wir  $[b_2] \in A_2, \dots, [b_\ell] \in A_\ell$ . Da (\dagger) eine direkte Summe ist, folgen Existenz und Eindeutigkeit modulo  $n$  des gesuchten  $a \in \mathbb{Z}$  mit  $[a] = [b_1] + \cdots + [b_\ell] \in \mathbb{Z}/n\mathbb{Z}$ .  $\square$

Im nächsten Satz können wir  $P = \mu_F$  wählen.

5.41. SATZ (verallgemeinerte Hauptraumzerlegung). *Es sei  $M = (V, F)$  ein  $\mathbb{k}[X]$ -Modul, es sei  $P \in \mathbb{k}[X]$  ein normiertes Polynom, so dass  $Pv = 0$  für alle  $v \in V$ , und es sei*

$$P = P_1^{k_1} \cdots P_\ell^{k_\ell}$$

*seine Primfaktorzerlegung. Dann existiert eine Zerlegung*

$$V = \ker(m_{P_1}^{k_1}) \oplus \cdots \oplus \ker(m_{P_\ell}^{k_\ell}).$$

Der Unterraum  $\ker(m_{P_i}^{k_i})$  heißt der verallgemeinerte *Hauptraum* von  $F$  zum irreduziblen Polynom  $P_i$ . Wenn  $P_i = X - x_i$  ein Linearfaktor ist, sprechen wir auch vom Hauptraum zum Eigenwert  $x_i$ .

BEWEIS. Dieser Satz folgt durch Induktion über  $\ell$ . Falls  $\ell = 0$  gilt  $P = 1$  und somit  $V = \ker(P) = \{0\}$ . Also ist nichts zu zeigen.

Für  $\ell \geq 1$  nehmen wir an, dass die Aussage für  $\ell - 1$  bereits gezeigt ist. Dann schreiben wir  $P = P_1^{k_1} R \in \mathbb{k}[X]$  mit  $R = P_2^{k_2} \cdots P_\ell^{k_\ell}$ . Nach Voraussetzung sind  $P_1^{k_1}$  und  $R$  teilerfremd, und Proposition 5.39 liefert eine Zerlegung  $V = U_1 \oplus W$  mit  $U_1 = \ker(m_{P_1}^{k_1})$ . Nach Induktionsvoraussetzung erhalten wir eine analoge Zerlegung  $W = U_2 \oplus \cdots \oplus U_\ell$ . Wie oben ist dann auch die Summe  $V = U_1 \oplus \cdots \oplus U_\ell$  direkt.  $\square$

In den Übungen geben wir noch zwei Anwendungen des chinesischen Restsatzes. Zum einen zerlegen wir die abelsche Gruppe  $\mathbb{Q}/\mathbb{Z}$  in eine direkte Summe. Zum anderen betrachten wir die sogenannte Partialbruchzerlegung rationaler Funktionen.

### 5.7. Eine allgemeine Normalform für Endomorphismen

In diesem Abschnitt stellen wir einen Endomorphismus auf jedem Hauptraum einzeln durch eine spezielle, angepasste Matrix dar. Damit lösen wir das Problem aus Bemerkung 5.10: zwei Endomorphismen  $F \in \text{End}_{\mathbb{k}}(V)$  und  $G \in \text{End}_{\mathbb{k}}(W)$  endlich-dimensionaler Vektorräume sind genau dann isomorph, wenn sie durch die gleiche Matrix in Normalform dargestellt werden. Wir werden sehen, dass die Matrix in Normalform durch deutlich weniger Parameter beschrieben werden kann als die darstellenden Matrizen von  $F$  und  $G$  bezüglich beliebiger Basen.

Die Konstruktion einer darstellenden Matrix hängt vor allem an der Wahl einer passenden Basis. In diesem Kontext ist das in erster Linie eine Aufgabe in Buchführung. Um später nicht völlig die Übersicht zu verlieren, beginnen wir einigen Vorüberlegungen.

5.42. BEMERKUNG. Wir hatten in Bemerkung 5.38 gesehen, dass ein  $\mathbb{k}[X]$ -Modul  $M$  das gleiche ist wie ein  $\mathbb{k}$ -Vektorraum  $V$  mit einem Endomorphismus  $F \in \text{End}_{\mathbb{k}}(V)$ , und  $M = (V, F)$  geschrieben. Wenn es ein Polynom  $P \in \mathbb{k}[X]$  gibt, so dass  $Pv = P(F)(v) = 0$  für alle  $v \in V$ , dann wirkt offensichtlich auch der Ring  $\mathbb{k}[X]/P \cdot \mathbb{k}[X]$  auf  $V$ .

In Bemerkung 5.37 haben wir den Fall betrachtet, dass  $P$  irreduzibel ist, und festgestellt, dass  $\mathbb{K} = \mathbb{k}[X]/P \cdot \mathbb{k}[X]$  dann ein Körper ist. Insbesondere ist  $M = (V, F)$  mit  $V = \ker(m_P)$  dann ein  $\mathbb{K}$ -Vektorraum, und wir haben die Werkzeuge aus Kapitel 3 zur Verfügung.

Sei umgekehrt  $M = (V, F)$  ein  $\mathbb{K}$ -Vektorraum, dann erhalten wir einen  $\mathbb{k}$ -Vektorraum  $V$ , indem wir die Multiplikation auf  $\mathbb{k} \subset \mathbb{K}$  einschränken.

5.43. PROPOSITION. *Es sei  $P \in \mathbb{k}[X]$  ein irreduzibles Polynom vom Grad  $n = \deg P$  und  $V$  ein Vektorraum über dem Körper  $\mathbb{K} = \mathbb{k}[X]/P \cdot \mathbb{k}[X]$ . Es sei  $B = (v_1, \dots, v_m)$  eine  $\mathbb{K}$ -Basis von  $V$ , dann ist  $C = (X^j v_i)_{1 \leq i \leq m, 0 \leq j < n}$  eine  $\mathbb{k}$ -Basis von  $V$ . Insbesondere folgt  $\dim_{\mathbb{k}} V = n \dim_{\mathbb{K}} V$ .*

BEWEIS. Repräsentiere ein beliebiges Element von  $\mathbb{K}$  durch  $Q \in \mathbb{k}[X]$  und schreibe  $Q = S \cdot P + T$  mit  $\deg T < n$ . Da  $[P] = 0 \in \mathbb{K}$ , gilt  $[Q] = [T] \in \mathbb{K}$ . Also können wir jedes Element von  $\mathbb{K}$  durch ein eindeutiges Polynom  $Q$  vom Grad  $\deg Q < n$  darstellen.

Wir stellen einen beliebigen Vektor  $v \in V$  in der gegebenen  $\mathbb{K}$ -Basis dar als  $\mathbb{K}$ -Linearkombination mit Koeffizienten  $[Q_i]$ , wobei  $\deg Q_i < n$ , etwa  $Q_i = q_{i,0}X^0 + \dots + q_{i,n-1}X^{n-1}$  und erhalten

$$(*) \quad v = \sum_{i=1}^m Q_i v_i = \sum_{i=1}^m \sum_{j=0}^{n-1} q_{i,j} X^j v_i .$$

Insbesondere erhalten wir eine  $\mathbb{k}$ -Linearkombination der Elemente  $X^j v_i$  von  $C$ . Somit bildet  $C$  ein  $\mathbb{k}$ -Erzeugendensystem.

Sei jetzt  $v = 0$  in (\*) für Koeffizienten  $q_{i,j} \in \mathbb{k}$ . Dann bilden wir Polynome  $Q_i = \sum_{j=0}^{n-1} q_{i,j} X^j$  und erhalten eine  $\mathbb{K}$ -Darstellung der 0 bezüglich  $B$ . Da  $B$  linear unabhängig über  $\mathbb{K}$  ist, folgt  $[Q_i] = 0 \in \mathbb{K}$  für alle  $i$ . Da  $Q_i$  der eindeutige Repräsentant von  $[Q_i]$  vom Grad  $\deg Q_i < n$  ist, gilt  $q_{i,j} = 0$  für alle  $i, j$ . Aber dann ist  $C$  auch linear unabhängig über  $\mathbb{k}$ , also eine  $\mathbb{k}$ -Basis von  $V$ .  $\square$

Wir betrachten jetzt einen einzelnen Hauptraum  $\ker(M_P^k) \subset V$  für ein irreduzibles Polynom  $P \in \mathbb{k}[X]$ . Dann gibt es eine Kette von Inklusionen

$$\{0\} = \ker(m_P^0) \subset \ker(m_P^1) \subset \dots \subset \ker(m_P^{k-1}) \subset \ker(m_P^k) = V .$$

5.44. PROPOSITION. *Es sei  $V = \ker(m_P^k)$  für ein irreduzibles Polynom  $P \in \mathbb{k}[X]$ , und es sei  $\mathbb{K} = \mathbb{k}[X]/P \cdot \mathbb{k}[X]$ . Für alle  $1 \leq j \leq k$  ist*

$$(1) \quad W_j = \ker(m_P^j) / \ker(m_P^{j-1})$$

*ein  $\mathbb{K}$ -Vektorraum, und  $m_P \in \text{End}_{\mathbb{k}}(V)$  induziert eine  $\mathbb{K}$ -lineare, injektive Abbildung*

$$(2) \quad \bar{m}_P: W_{j+1} \longrightarrow W_j .$$

BEWEIS. Nach Bemerkung 5.38 sind  $\ker(m_P^j)$  und  $\ker(m_P^{j-1})$   $F$ -invariante Unterräume, also  $\mathbb{k}[X]$ -Untermoduln von  $M = (V, F)$ . Somit ist der Quotient  $W_j = \ker(m_P^j) / \ker(m_P^{j-1})$  nach Proposition 2.39 ebenfalls ein  $\mathbb{k}[X]$ -Modul.

Es sei  $j \geq 1$  beliebig und  $v \in \ker(m_P^j)$ , dann gilt  $m_P^{j-1}(Pv) = m_P^j(v) = 0$ , also folgt  $Pv \in \ker(m_P^{j-1})$ , das heißt, wir erhalten eine  $\mathbb{k}[X]$ -lineare Abbildung

$$(*) \quad m_P: \ker(m_P^j) \longrightarrow \ker(m_P^{j-1}).$$

Insbesondere wirkt  $P$  wie 0 auf  $\ker(m_P^j)/\ker(m_P^{j-1})$ , somit ist  $W_j$  in (1) ein  $\mathbb{K}$ -Vektorraum nach Bemerkung 5.42.

Mit (\*) für  $j+1$  konstruieren wir  $m_P: \ker(m_P^{j+1}) \rightarrow \ker(m_P^j) \rightarrow W_j$ . Wir betrachten das Diagramm

$$\begin{array}{ccccc} \ker(m_P^j) & \hookrightarrow & \ker(m_P^{j+1}) & \twoheadrightarrow & W_{j+1} = \ker(m_P^{j+1})/\ker(m_P^j) \\ & \searrow 0 & \downarrow m_P & \swarrow \bar{m}_P & \\ & & W_j = \ker(m_P^j)/\ker(m_P^{j-1}) & & \end{array}$$

von  $\mathbb{k}[X]$ -Moduln. Wegen (\*) gilt  $0 = m_P: \ker(m_P^j) \rightarrow W_j$ . Nach der universellen Eigenschaft 2.42 des Quotienten induziert  $m_P$  eine eindeutige  $\mathbb{k}[X]$ -lineare Abbildung  $\bar{m}_P: W_{j+1} \rightarrow W_j$ . Da beides  $\mathbb{K}$ -Vektorräume sind, ist  $\bar{m}_P$  sogar  $\mathbb{K}$ -linear.

Sei jetzt  $v \in \ker(m_P^{j+1})$  ein Element, so dass  $w = [v] \in \ker(\bar{m}_P) \subset W_{j+1}$ . Dann folgt  $Pv = m_P(v) \in \ker(m_P^{j-1})$ , somit  $m_P^j(v) = 0$  und  $v \in \ker(m_P^j)$ . Es folgt  $w = 0 \in W_{j+1}$ , das heißt, die Abbildung  $\bar{m}_P$  in (2) ist injektiv.  $\square$

5.45. PROPOSITION. *Es gelte  $\dim_{\mathbb{k}} V < \infty$ . Unter den Annahmen von Proposition 5.44 existieren Vektoren  $e_i^{(j)} \in \ker(m_P^j) \setminus \ker(m_P^{j-1})$  für  $1 \leq j \leq k$  und  $d_{j+1} < i \leq d_j$ , wobei  $d_j = \dim_{\mathbb{K}} W_j$ , so dass  $V$  eine Basis besitzt der Form*

$$B = \left( (F^p \circ P(F)^q) (e_i^{(j)}) \right)_{1 \leq j \leq k, d_{j+1} < i \leq d_j, 0 \leq p < n, 0 \leq q < j}.$$

BEWEIS. Aus  $\dim_{\mathbb{k}} V < \infty$  folgt  $\dim_{\mathbb{k}} W_j < \infty$  für alle  $j$ , und wegen Proposition 5.43 gilt das auch über  $\mathbb{K}$ . Wir bestimmen als erstes  $\mathbb{K}$ -Basen von  $W_k, W_{k-1}, \dots, W_1$  wie folgt. Wir beginnen mit einer  $\mathbb{K}$ -Basis  $(w_1^{(k)}, \dots, w_{d_k}^{(k)})$  von  $W_k$ . In jedem weiteren Schritt betrachten wir das Bild der bereits konstruierten  $d_{j+1}$ -elementigen  $\mathbb{K}$ -Basis von  $W_{j+1}$  unter der injektiven Abbildung  $\bar{m}_P$ . Nach dem Basisergänzungssatz 3.3 können wir Elemente  $(w_{d_{j+1}+1}^{(j)}, \dots, w_{d_j}^{(j)})$  ergänzen und erhalten eine  $\mathbb{K}$ -Basis von  $W_j$ :

$$(1) \quad \left( (P(F)^{\ell-j}) (w_i^{(\ell)}) \right)_{j \leq \ell \leq k, d_{\ell+1} < i \leq d_{\ell}}.$$

Mit Hilfe von Proposition 5.43 gehen wir über zu einer  $\mathbb{k}$ -Basis von  $W_j$  der Form

$$(2) \quad \left( (F^p \circ P(F)^{\ell-j}) (w_i^{(\ell)}) \right)_{j \leq \ell \leq k, d_{\ell+1} < i \leq d_{\ell}, 0 \leq p < n}.$$

Anschließend wählen wir für  $1 \leq j \leq k$  und  $d_{j+1} < i \leq d_j$  Repräsentanten  $e_i^{(j)} \in \ker(m_P^j)$  von  $w_i^{(j)} \in W_j$ . Wir behaupten, dass dann das Tupel  $B$  aus der Proposition eine  $\mathbb{k}$ -Basis von  $V$  ist.

Um einen Vektor  $v \in V$  als Linearkombination darzustellen, stellen wir zunächst seine Äquivalenzklasse in  $W_k = V/\ker(m_P^{k-1})$  dar als

$$[v] = \sum_{i=1}^{d_k} \sum_{p=0}^{n-1} a_{i,p}^{(k)} F^p(w_i^{(k)}) \in W_k .$$

Dann ersetzen wir die  $w_i^{(k)}$  durch  $e_i^{(k)}$  und betrachten jetzt

$$v_{k-1} = v - \sum_{i=1}^{d_k} \sum_{p=0}^{n-1} a_{i,p}^{(k)} F^p(e_i^{(k)}) \in \ker(m_P^{k-1}) \in \ker(m_P^{k-1}) .$$

Wir stellen  $[v_{k-1}] \in W_{k-1}$  bezüglich der Basis (2) für  $j = k - 1$  dar, ersetzen wieder  $w_i^{(k-1)}$  durch  $v_i^{(k-1)}$ , und ziehen die entsprechende Linearkombination von  $v_{k-1}$  ab. Auf diese Weise arbeiten wir uns bis  $\ker(m_P)$  vor und haben schließlich eine Linearkombination für ganz  $v$  gefunden.

Um lineare Unabhängigkeit zu beweisen, stellen wir 0 dar und betrachten der Reihe nach das Bild der jeweiligen Linearkombination in  $W_k, W_{k-1}, \dots$ , um zu zeigen, dass die entsprechenden Koeffizienten alle verschwinden.  $\square$

Zum Polynom  $P = X^n + \sum_{q=0}^{n-1} p_{n-q} X^q$ . betrachten wir die sogenannte *Begleitmatrix*  $M_P$  zum Polynom  $P$  und die Matrix  $N_n$ , gegeben durch

$$M_P = \begin{pmatrix} 0 & & -p_n \\ 1 & \ddots & \vdots \\ & \ddots & 0 \\ & & 1 & -p_1 \end{pmatrix} \quad \text{und} \quad N_n = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ & & & 0 \\ & & & \vdots \\ & & & 0 \end{pmatrix} \in M_n(\mathbb{k}) .$$

5.46. PROPOSITION. *Unter den Annahmen von Proposition 5.45 spannt das Tupel*

$$B_i^{(j)} = \left( e_i^{(j)}, \dots, F^{n-1}(e_i^{(j)}), P(F)(e_i^{(j)}), \dots, (F^{n-1} P(F)^{j-1})(e_i^{(j)}) \right)$$

aus Elementen von  $B$  für alle  $1 \leq j \leq k$  und alle  $d_{j+1} < i \leq d_j$  einen  $F$ -invarianten Unterraum  $V_i^{(j)} \subset V$  auf, und bezüglich  $B_i^{(j)}$  wird  $F$  dargestellt durch die Blockmatrix

$$(*) \quad B_i^{(j)}(F|_{V_i^{(j)}})_{B_i^{(j)}} = \begin{pmatrix} M_P & & & \\ N_n & M_P & & \\ & \ddots & \ddots & \\ & & N_n & M_P \end{pmatrix} \in M_{jn}(\mathbb{k}) ,$$

Alle nicht bezeichneten Blöcke sind 0. Beachte, dass die untere Nebendiagonale durchgehend aus 1 besteht. Die Matrix  ${}_B F_B$  ist dann eine Blockdiagonalmatrix aus Blöcken der Form (\*).

BEWEIS. Falls  $q < n - 1$ , gilt

$$F((F^q P(F)^r)(e_i^{(j)})) = (F^{q+1} P(F)^r)(e_i^{(j)}) .$$

Falls  $q = n - 1$  und  $r < j - 1$ , erhalten wir

$$F((F^{n-1}P(F)^r)(e_i^{(j)})) = P(F)^{r+1}(e_i^{(j)}) - \sum_{q=0}^{n-1} p_{n-q}(F^q P(F)^r)(e_i^{(j)}).$$

Falls  $q = n - 1$  und  $r = j - 1$ , verschwindet  $P(F)^j(e_i^{(j)})$ , und es bleibt

$$F((F^{n-1}P(F)^r)(e_i^{(j)})) = - \sum_{q=0}^{n-1} p_{n-q}(F^q P(F)^r)(e_i^{(j)}). \quad \square$$

5.47. SATZ (Weierstraß-Normalform für Endomorphismen). *Es sei  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}}(V)$ . Dann existieren irreduzible Polynome  $P_1, \dots, P_\ell$  und  $k_1, \dots, k_\ell \in \mathbb{N} \setminus \{0\}$ , so dass  $V$  als direkte Summe  $F$ -invarianter Unterräume  $V_i$  geschrieben werden kann, auf denen  $F$  jeweils durch eine Matrix der Form (\*) zu  $P_i$  und  $k_i$  dargestellt wird. Die Matrixdarstellung ist bis auf die Reihenfolge der Blöcke eindeutig.*

*Sei  $W$  ein weiterer endlich-dimensionaler  $\mathbb{k}$ -Vektorraum und  $G \in \text{End}_{\mathbb{k}}(W)$ . Dann existiert genau dann Isomorphismus  $\Phi: V \rightarrow W$  mit  $\Phi \circ F = G \circ \Phi$ , wenn  $F$  und  $G$  durch die gleiche Matrix in Normalform dargestellt werden.*

Die Polynom  $P_i$  müssen nicht paarweise verschieden sein, und die zugrundeliegende Basis  $B$  von  $V$  ist nicht eindeutig bestimmt.

BEWEIS. Wir benutzen zunächst die verallgemeinerte Hauptraumzerlegung aus Satz 5.41, um  $V$  in eine direkte Summe zu verschiedenen irreduziblen Polynomen  $P_i$  zu zerlegen. Anschließend konstruieren wir auf jedem Hauptraum eine Basis wie in Proposition 5.45. Proposition 5.46 liefert uns die einzelnen Blöcke der Form (\*).

Zur Eindeutigkeit überlegen wir uns zunächst, dass für jedes irreduzible Polynom  $F$  der Hauptraum  $\ker(m_P^k)$  nur von  $F$  und  $P$  abhängt, wenn  $k$  hinreichend groß gewählt wurde. Die Anzahl der Blöcke (\*) der Größe  $j n \times j n$  ergibt sich als

$$d_j - d_{j+1} = \frac{1}{n} \left( 2 \dim \ker(m_P^j) - \dim \ker(m_P^{j+1}) - \dim \ker(m_P^{j-1}) \right)$$

und hängt ebenfalls nur von  $F$  und  $P$  ab. Umgekehrt kann man die obigen Zahlen direkt aus der Matrix  ${}_B F_B$  in Normalform ablesen. Hieraus ergibt sich die Eindeutigkeit.

Zu guter Letzt folgt aus der Existenz eines Isomorphismus  $\Phi$  mit  $\Phi \circ F = G \circ \Phi$ , dass die darstellenden Matrizen übereinstimmen. Wenn andererseits die darstellenden Matrizen übereinstimmen, dann gibt es Basen wie in Proposition 5.45 für jeden Hauptraum. Indem man entsprechende Basisvektoren aufeinander abbildet, erhält man den gesuchten Isomorphismus  $\Phi$ .  $\square$

5.48. FOLGERUNG (vergleiche Satz 5.22 von Cayley-Hamilton). *Es sei  $F$  Endomorphismus eines endlichen  $\mathbb{k}$ -Vektorraums, und es sei*

$$\chi_F = P_1^{m_1} \dots P_\ell^{m_\ell}$$

die Zerlegung des charakteristischen Polynom in irreduzible Faktoren. Dann gilt

$$\mu_F = P_1^{k_1} \cdots P_\ell^{k_\ell},$$

wobei  $1 \leq k_i \leq m_i$  für alle  $1 \leq i \leq \ell$ . Insbesondere ist  $\lambda \in \mathbb{k}$  genau dann ein Eigenwert von  $F$ , wenn  $\mu_F(\lambda) = 0$ .

BEWEIS. Nach Konstruktion hat der Block (\*) in Proposition 5.46 das Minimalpolynom  $P^j$ . In den Übungen haben wir  $\chi_{M_P} = P$  nachgerechnet. Mit Folgerung 4.18 über die Determinante von Block-Dreiecksmatrizen sehen wir leicht, dass das charakteristische Polynom von (\*) ebenfalls  $P^j$  ist. Für  $F$  auf dem gesamten Hauptraum  $\ker(P(F)^k)$  erhalten wir

$$\chi_{F|_{\ker(P(F)^k)}} = \prod_{j=1}^k \prod_{i=m_{j+1}+1}^{m_j} P^j = P^{\frac{\dim V}{n}} \quad \text{und} \quad \mu_{F|_{\ker(P(F)^k)}} = P^k.$$

Ein irreduzibles Polynom kann in  $\chi_F$  nur vorkommen, wenn es auch in  $\mu_F$  vorkommt, allerdings hat es in  $\chi_F$  möglicherweise einen größeren Exponenten. Der Spezialfall  $P = X - \lambda$  liefert die letzte Aussage.  $\square$

Tatsächlich können wir den Satz 5.22 von Cayley-Hamilton sogar auf diesem Wege beweisen. Wir haben ihn bis jetzt nur bei der Definition des Minimalpolynoms in Satz 5.27 benutzt, um zu zeigen, dass die dortige Menge  $M$  nicht nur 0 enthält. Aber das können wir auch einsehen, indem wir uns überlegen, dass  $\dim_{\mathbb{k}} \text{End}_{\mathbb{k}}(V) = n^2$ , wenn  $n = \dim_{\mathbb{k}} V$ . Eine nichttriviale Darstellung der 0 als Linearkombination der  $(n^2 + 1)$  Endomorphismen  $F^0, \dots, F^{n^2}$  liefert ein Element von  $M \setminus \{0\}$ , und es folgt  $\mu_F \neq 0$ .

## 5.8. Die Jordansche Normalform

Wir betrachten jetzt den Spezialfall, dass das Minimalpolynom in Linearfaktoren zerfällt. Damit können wir endlich die Begriffe „diagonalisierbar“ und „trigonalisierbar“ aus Definition 5.4 charakterisieren.

Für  $P = X - \lambda$  vereinfachen sich die Überlegungen aus dem letzten Abschnitt. Zum Beispiel gilt  $\mathbb{K} = \mathbb{k}$  nach Bemerkung 5.37 (1). Als Begleitmatrix erhalten wir  $M_{X-\lambda} = (\lambda) \in M_1(\mathbb{k})$ . Wenn wir die Basis  $B_i^{(j)}$  aus Proposition 5.46 in umgekehrter Reihenfolge aufschreiben, also

$$\left( (F - \lambda)^{j-1}(e_i^{(j)}), \dots, (F - \lambda)(e_i^{(j)}), e_i^{(j)} \right),$$

dreht sich in der Matrix (\*) aus Proposition 5.46 die Reihenfolge der Blöcke gerade um, und wir erhalten einen sogenannten *Jordan-Block* der Form

$$J_j(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in M_j(\mathbb{k}).$$

5.49. DEFINITION. Es sei  $R$  ein Ring. Ein Element  $r \in R$  *nilpotent*, wenn es ein  $k \in \mathbb{N}$  gibt, so dass  $r^k = 0 \in R$ . Das minimale solche  $k$  heißt auch die (Nilpotenz-) *Ordnung* von  $F$ .

5.50. BEISPIEL. Wir erinnern uns an die Definition 4.17 von Dreiecksmatrizen. Jede strikte obere Dreiecksmatrix  $A \in M_n(\mathbb{k})$  ist ein nilpotentes Element von  $M_n(\mathbb{k})$  und beschreibt einen nilpotenten Endomorphismus von  $\mathbb{k}^n$ . Zur Begründung schreibe  $A^k = (a_{ij}^{(k)})_{i,j} \in M_n(\mathbb{k})$ , dann hat  $A^k$  die Eigenschaft, dass  $a_{ij}^{(k)} = 0$  für alle  $i, j$  mit  $j < i + k$ . In anderen Worten gilt

$$A^k = \begin{pmatrix} 0 & \dots & 0 & * & \dots & * \\ & & & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & * \\ & & & & & 0 \\ & & & & & \vdots \\ 0 & & \dots & & & 0 \end{pmatrix}.$$

Wenn das stimmt, ist spätestens  $A^n = 0$ , also ist  $A$  nilpotent.

Da  $A$  strikte obere Dreiecksmatrix ist, gilt obige Behauptung für  $k = 1$ . Wenn wir sie für  $k$  bereits überprüft haben, berechnen wir  $A^{k+1} = A \cdot A^k$ , also

$$a_{ij}^{(k+1)} = \sum_{l=1}^n a_{il} a_{lj}^{(k)}.$$

Sei  $j < i + k + 1$ . Die Summanden für  $l < i + 1$  verschwinden, da dann  $a_{il} = 0$ . Die Summanden für  $l \geq i + 1$  verschwinden nach Induktionsvoraussetzung, da  $j < i + k + 1 \leq l + k$ . Also gilt  $a_{ij}^{(k+1)} = 0$  wie gefordert. Damit folgt die Behauptung per Induktion.

5.51. SATZ (Trigonalisierbarkeit, Jordan-Normalform). *Es sei  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \text{End}_{\mathbb{k}} V$ . Dann sind die folgenden Aussagen äquivalent.*

- (1) *Der Endomorphismus  $F$  ist trigonalisierbar.*
- (2) *Das charakteristische Polynom  $\chi_F$  zerfällt vollständig in Linearfaktoren.*
- (3) *Das Minimalpolynom  $\mu_F$  zerfällt vollständig in Linearfaktoren.*
- (4) *Der Vektorraum  $V$  zerfällt in eine direkte Summe von Haupträumen zu Eigenwerten von  $F$ .*
- (5) *Jordan-Normalform. Der Endomorphismus  $F$  lässt sich als Jordan-Matrix schreiben. Das heißt, es existieren  $\lambda_1, \dots, \lambda_\ell \in \mathbb{k}$  und  $k_1, \dots, k_\ell \in \mathbb{N} \setminus \{0\}$ , so dass  $F$  bezüglich einer geeigneten Basis  $B$  dargestellt wird durch*

$${}_B F_B = \begin{pmatrix} J_{k_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{k_\ell}(\lambda_\ell) \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \lambda_1 & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_1 \\ & & & & \ddots & & & \\ & & & & & \lambda_\ell & 1 & & 0 \\ & & & & & & \lambda_\ell & \ddots & \\ & & & & & & & \ddots & 1 \\ & & & & & 0 & & & \lambda_\ell \end{pmatrix}.$$

(6) Jordan-Chevalley-Zerlegung. *Es existiert eine Darstellung  $F = D + N$ , bei der  $D$  diagonalisierbar und  $N$  nilpotent ist, mit  $DN = ND$ .*

*Dabei ist die Jordan-Matrix in (5) bis auf die Reihenfolge der einzelnen Jordan-Blöcke eindeutig. In (6) sind die Endomorphismen  $D$  und  $N$  eindeutig bestimmt.*

*Analoge Aussagen gelten für quadratische Matrizen über  $\mathbb{k}$ .*

Insbesondere ist jede obere Dreiecksmatrix zu einer Matrix in Jordan-Normalform konjugiert. Beide Matrizen haben die gleichen Diagonaleinträge. In den Übungen sehen wir, dass man mit Jordan-Matrizen leichter rechnen kann als mit allgemeinen Dreiecksmatrizen.

Die Zahlen  $\lambda_1, \dots, \lambda_\ell$  müssen nicht paarweise verschieden sein. Die Basis  $B$  in (5) ist nicht eindeutig. Das ergibt sich aus der Konstruktion in Proposition 5.45, die gewisse Wahlmöglichkeiten zulässt.

Die Forderung  $DN = ND$  in (6) ist wesentlich, andernfalls könnte es andere Zerlegungen mit anderen Eigenwerten geben. Es gibt für alle Endomorphismen eine Jordan-Chevalley-Zerlegung, wenn man „diagonalisierbar“ durch den allgemeineren Begriff „halbeinfach“ ersetzt. Das wollen wir hier aber nicht weiter vertiefen.

**BEWEIS.** Wir zeigen hier nur die Aussagen über Endomorphismen.

Zu „(1)  $\implies$  (2)“ sei  $F$  dargestellt durch eine Dreiecksmatrix  $A \in M_n(\mathbb{k})$ . Wir wenden Folgerung 4.18 (2) auf die Dreiecksmatrix  $X \cdot E_n - A$  an und erhalten die Zerlegung

$$\chi_F(X) = \det(X \cdot E_n - A) = \prod_{i=1}^n (X - a_{ii}).$$

Zu „(2)  $\implies$  (3)“ benutzen wir Folgerung 5.48 zum Satz von Cayley-Hamilton, wonach  $\mu_F \mid \chi_F$ , und Satz 5.34 über die eindeutige Primfaktorzerlegung, woraus folgt, dass mit  $\chi_F$  auch  $\mu_F$  in Linearfaktoren zerfällt.

Die Richtung „(3)  $\implies$  (4)“ folgt dann unmittelbar aus Satz 5.41.

Der Schritt „(4)  $\implies$  (5)“ folgt — wie oben geschildert — aus Satz 5.47 zur Weierstraß-Normalform. Für jeden Hauptraum wählen wir eine Basis wie in Proposition 5.45 (da  $n = \deg(X - \lambda_i) = 1$ , müssen wir auf die Vektoren  $(F - \lambda)^q(e_i^{(j)})$  nicht noch zusätzlich  $F$  anwenden). Wie oben angemerkt schreiben wir diese Basen in umgekehrter Reihenfolge, um Jordanblöcke  $J_j(\lambda_i)$  zu erhalten.

Klar ist außerdem „(5)  $\implies$  (1)“, da die Jordan-Matrix eine obere Dreiecksmatrix ist.

Zu „(5)  $\implies$  (6)“ sei  $D$  der Endomorphismus, der durch den Diagonalanteil der Jordan-Matrix  ${}_B F_B$  in (5) bezüglich  $B$  dargestellt wird, und  $N = F - D$ . Als strikte obere Dreiecksmatrix ist  $N$  nilpotent, und man sieht leicht, dass  $N$  mit  $D$  vertauscht.

Zu „(6)  $\implies$  (4)“ bemerken wir zunächst, dass  $V$  nach Folgerung 5.8 (3) in die Eigenräume  $V_\lambda$  von  $D$  zerfällt. Da  $D$  mit  $N$  kommutiert, sind diese Eigenräume  $N$ -invariant (Übung). Es folgt, dass  $V_\lambda$  auch  $F$ -invariant ist, und dass  $F|_{V_\lambda} - \lambda \operatorname{id}_{V_\lambda} = N|_{V_\lambda}$  nilpotent ist. Mithin ist  $\lambda$  ein Eigenwert von  $F$ , und  $V_\lambda$  ist ein Unterraum des  $\lambda$ -Haupttraumes von  $F$ . Aber da die Summe aller Haupträume von  $F$  direkt ist, gilt sogar Gleichheit.

Die Eindeutigkeit der Jordan-Chevalley-Zerlegung folgt jetzt aus der Eindeutigkeit der Hauptraumzerlegung von  $F$  in Satz 5.41. Die Eindeutigkeit der Jordan-Normalform folgt aus der entsprechenden Aussage in Satz 5.47.  $\square$

**5.52. FOLGERUNG.** *Es sei  $\mathbb{k}$  ein algebraisch abgeschlossener Körper,  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum und  $F \in \operatorname{End}_{\mathbb{k}} V$ . Dann lässt sich  $F$  durch eine Matrix in Jordan-Normalform darstellen, und  $F$  besitzt auch eine Jordan-Chevalley-Zerlegung.*

**BEWEIS.** Das folgt aus Satz 5.51 (2) oder auch (3), da  $\chi_F$  und  $\mu_F$  nach Beispiel 5.36 (1) in Linearfaktoren zerfallen.  $\square$

**5.53. BEISPIEL.** Wir betrachten jetzt ein gedämpftes Federpendel, beschrieben durch die lineare Differentialgleichung

$$(*) \quad \ddot{u}(t) = -b\dot{u}(t) - c u(t);$$

im Vergleich zu Beispiel 5.16 kommt also ein Reibungsterm  $-b\dot{u}(t)$  hinzu. Wir führen wieder  $v = \dot{u}$  als zusätzliche Variable ein und erhalten das System

$$(\dagger) \quad \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -c & -b \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix}.$$

Das charakteristische Polynom der Koeffizientenmatrix  $C$  hat die Form

$$\chi_C(X) = \det \begin{pmatrix} X & -1 \\ c & X + b \end{pmatrix} = X^2 + bX + c.$$

Wir unterscheiden drei Fälle.

- (1) Falls  $b^2 < 4c$ , hat  $\chi_C$  zwei zueinander konjugierte komplexe Nullstellen  $-\frac{b}{2} \pm i\frac{\sqrt{4c-b^2}}{2}$ . Wenn wir wie in Beispiel 5.16 weiterrechnen, erhalten wir reelle Lösungen der Form

$$u(t) = r e^{-\frac{bt}{2}} \cos\left(\varphi + t \frac{\sqrt{4c-b^2}}{2}\right).$$

Eine Winkelfunktion mit exponentiell abfallender Amplitude nennt man auch *gedämpfte Schwingung*.

- (2) Falls  $b^2 > 4c$ , hat  $\chi_C$  die reellen Nullstellen  $-\frac{b \pm \sqrt{b^2-4c}}{2} < 0$ . Wir erhalten die allgemeine Lösung

$$u(t) = r e^{-\frac{b+\sqrt{b^2-4c}}{2}t} + s e^{-\frac{b-\sqrt{b^2-4c}}{2}t}.$$

- (3) Im Grenzfall  $b^2 = 4c$  gilt  $\chi_F = (X + \frac{b}{2})^2$ . Da  $C$  keine Diagonalmatrix ist, muss sie sich als Jordanblock  $J_2(-\frac{b}{2})$  schreiben lassen. Wir beginnen dazu mit dem Vektor  $v_2 = e_2 \notin \ker(C + \frac{b}{2})$  und fügen als ersten Basisvektor  $v_1(C + \frac{b}{2})(e_2) = (\frac{1}{-\frac{b}{2}})$  hinzu. Bezüglich dieser Basis erhalten wir den Jordanblock

$$\begin{pmatrix} 1 & 0 \\ -\frac{b}{2} & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 & 1 \\ -\frac{b^2}{4} & -b \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -\frac{b}{2} & 1 \end{pmatrix} = \begin{pmatrix} -\frac{b}{2} & 1 \\ 0 & -\frac{b}{2} \end{pmatrix} = J_2\left(-\frac{b}{2}\right).$$

Wir lösen das neue Gleichungssystem

$$\begin{pmatrix} \dot{f} \\ \dot{g} \end{pmatrix} = \begin{pmatrix} -\frac{b}{2} & 1 \\ 0 & -\frac{b}{2} \end{pmatrix} \cdot \begin{pmatrix} f \\ g \end{pmatrix}, \quad \text{also} \quad \begin{aligned} \dot{f} &= -\frac{b}{2}f + g, \\ \dot{g} &= -\frac{b}{2}g. \end{aligned}$$

Wir erhalten sofort eine Lösung  $g = se^{-\frac{bt}{2}}$  mit  $s \in \mathbb{R}$ . Für  $f$  raten wir die Lösung  $f = (r + st)e^{-\frac{bt}{2}}$  mit  $r \in \mathbb{R}$ , mehr dazu später. Rücktransformation liefert

$$\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{b}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} (r + st)e^{-\frac{bt}{2}} \\ se^{-\frac{bt}{2}} \end{pmatrix} = \begin{pmatrix} (r + st)e^{-\frac{bt}{2}} \\ -\frac{b}{2}(r + st)e^{-\frac{bt}{2}} + se^{-\frac{bt}{2}} \end{pmatrix},$$

und wieder gilt  $v = \dot{u}$ .

Wir stellen uns jetzt vor, dass die Gleichung (\*) die Federung eines Kraftwagens beschreibt, und dass wir den Parameter  $b > 0$  einstellen können. Im Fall  $b < 2\sqrt{c}$  würde das Auto nach einem Stoß zu schwingen beginnen und dann allmählich in Ruhelage zurückkehren. Dieses Verhalten ist sicher nicht erwünscht. Im Fall  $b > 2\sqrt{c}$  kehrt das Auto nach einem Stoß ohne zu schwingen in Ruhelage zurück, das ist schon besser. Man kann sich überzeugen, dass  $b - \sqrt{b^2 - 4c}$  in Abhängigkeit von  $b$  monoton fällt. Das heißt, je größer  $b$  ist, desto langsamer wird das Auto in Ruhelage zurückkehren, da für große  $t$  die Funktion  $e^{-\frac{b-\sqrt{b^2-4c}}{2}t}$  dominiert. Somit wird man versuchen,  $b$  so klein wie möglich einzustellen, ohne dass es zu Schwingungen kommt, das heißt, man versucht genau dem Parameter  $b = 2\sqrt{c}$  zu treffen, bei dem die Koeffizientenmatrix zu einem Jordanblock ähnlich ist. Der zusätzliche Faktor  $t$  in der Lösung wird von der fallenden Exponentialfunktion dominiert und stört daher nicht.

Wir erinnern uns an die Ordnung  $\text{ord}_\lambda(Q) = \text{ord}_{X-\lambda}(Q)$  aus Satz 5.34, siehe auch Bemerkung 5.35.

5.54. DEFINITION. Es sei  $\mathbb{k}$  ein Körper,  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum,  $F \in \text{End}_{\mathbb{k}}(V)$  und  $\lambda \in \mathbb{k}$ . Sei  $V_\lambda$  der  $\lambda$ -Eigenraum von  $F$ , dann nennen wir  $\dim_{\mathbb{k}} V_\lambda$  die *geometrische Vielfachheit* und  $\text{ord}_\lambda \chi_F = \text{ord}_{X-\lambda} \chi_F$  die *algebraische Vielfachheit* von  $\lambda$  als Eigenwert von  $F$ .

5.55. BEMERKUNG. Jeder Jordanblock  $J_k(\lambda)$  hat einen eindimensionalen Eigenraum, auf gespannt vom ersten Basisvektor. Somit ist die geometrische Vielfachheit genau die Anzahl der Jordanblöcke zum Eigenwert  $\lambda$ .

Die algebraische Vielfachheit ist die Dimension des  $\lambda$ -Haupttraums, also die Summe der Größen der Jordanblöcke zum Eigenwert  $\lambda$ . Sie ist mindestens so groß wie die geometrische Vielfachheit.

Sei andererseits  $j$  die Potenz, mit der  $X - \lambda$  im Minimalpolynom  $\mu_F$  vorkommt, dann ist  $j$  die maximale Größe eines Jordanblocks zum Eigenwert  $\lambda$ . Wir schließen daraus, dass die algebraische Vielfachheit um mindestens  $j - 1$  größer ist als die geometrische.

Wir erinnern uns jetzt an Diagonalisierbarkeit, siehe Definition 5.4. Eine hinreichende, aber nicht notwendige Bedingung für Diagonalisierbarkeit hatten wir in Folgerung 5.9 bereits kennengelernt.

5.56. SATZ (Diagonalisierbarkeit). *Es sei  $\mathbb{k}$  ein Körper und  $V$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum. Dann sind für einen Endomorphismus  $F \in \text{End}_{\mathbb{k}} V$  die folgenden Aussagen äquivalent.*

- (1)  $F$  ist diagonalisierbar;
- (2)  $V$  besitzt eine Basis aus Eigenvektoren von  $F$ ;
- (3)  $V$  zerfällt in eine direkte Summe von Eigenräumen von  $F$ ;
- (4) Das charakteristische Polynom zerfällt in Linearfaktoren, und für jeden Eigenwert  $\lambda \in \mathbb{k}$  stimmen algebraische und geometrische Vielfachheit überein, das heißt, es gilt

$$\text{ord}_\lambda \chi_A = \dim \ker(\lambda \text{id}_V - F) ;$$

- (5) Das Minimalpolynom zerfällt in paarweise verschiedene Linearfaktoren.

*Die Diagonaleinträge sind gerade die Eigenwerte von  $F$ , und kommen entsprechend ihrer Vielfachheit in  $\chi_F$  oft vor. Insbesondere ist die Diagonalmatrix bis auf Reihenfolge der Diagonaleinträge eindeutig durch  $F$  bestimmt.*

*Analoge Aussagen gelten für quadratische Matrizen über  $\mathbb{k}$ .*

BEWEIS. Die Äquivalenz von (1)–(3) haben wir in Proposition 5.5 und Folgerung 5.8 (3) gezeigt, und daraus ergibt sich auch die Eindeutigkeit der Diagonalmatrix bis auf Reihenfolge der Einträge.

Zu „(3)  $\implies$  (4)“ sei  $V_\lambda \subset V$  der  $\lambda$ -Eigenraum von  $F$ . Nach Folgerung 4.18 (1) ist  $\chi_F$  das Produkt der charakteristischen Polynome von  $F|_{V_\lambda}$

über alle  $\lambda$ , und diese sind gerade  $(X - \lambda)^{\dim V_\lambda}$ . Also zerfällt  $\chi_F$  in Linearfaktoren, und algebraische und geometrische Vielfachheit stimmen für alle  $\lambda$  überein.

Zu „(4)  $\implies$  (5)“. Nach dem Satz 5.22 von Cayley-Hamilton gilt  $\mu_F \mid \chi_F$ , also zerfällt  $\mu_F$  ebenfalls in Linearfaktoren. Wegen Bemerkung 5.55 kann kein Linearfaktor mehr als einmal in  $\mu_F$  vorkommen.

Zu „(5)  $\implies$  (3)“. Wenn  $\mu_F$  in Linearfaktoren zerfällt, zerfällt  $V$  in eine direkte Summe von Haupträumen zu Eigenwerten  $\lambda \in \mathbb{k}$  nach Satz 5.41. Wenn  $X - \lambda$  nur einmal in  $\mu_F$  vorkommt, ist der Eigenraum  $\ker(X - \lambda)$  bereits der ganze Hauptraum, also zerfällt  $V$  in eine direkte Summe von Eigenräumen.  $\square$

### 5.9. Anwendungen der Jordan-Normalform

Zum Schluss des Kapitels skizzieren wir zwei Anwendungen der Jordan-Normalform im Zusammenhang mit Analysis.

Es sei  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$ , und es sei  $(p_n)_{n \in \mathbb{N}}$  eine Folge in  $\mathbb{k}$ . Eine *Potenzreihe* in einer Variablen  $X$  ist ein Ausdruck der Form

$$P(X) = \sum_{n=0}^{\infty} p_n X^n,$$

vergleiche Beispiel 2.80. Im Unterschied zu einem Polynom ist es erlaubt, dass beliebig viele  $p_i$  von 0 verschieden sind.

In der Analysis definiert man den Begriff der *Konvergenz* einer Potenzreihe an einer Stelle  $x \in \mathbb{k}$ ; das Gegenteil davon ist *Divergenz*. Man zeigt dann, dass es einen Konvergenzradius  $\rho \in [0, \infty]$  in Abhängigkeit von den Koeffizienten  $(p_n)_n$  gibt, so dass

$$\begin{aligned} |x| < \rho &\implies P(x) \text{ konvergiert, und} \\ |x| > \rho &\implies P(x) \text{ divergiert.} \end{aligned}$$

Im Fall  $|x| = \rho$  ist sowohl Konvergenz als auch Divergenz möglich.

Ähnlich wie in Proposition 5.21 ist es möglich, Matrizen  $A \in M_n(\mathbb{k})$  in Potenzreihen einzusetzen. Es sei  $C = B^{-1} \cdot A \cdot B \in M_n(\mathbb{C})$  die Jordan-Normalform über  $\mathbb{C}$  von  $A$ . Die Rechnung

$$\begin{aligned} P(A) &= \sum_{i=0}^{\infty} p_i (B \cdot C \cdot B^{-1})^i = \sum_{i=0}^{\infty} p_i B \cdot C^i \cdot B^{-1} \\ &= B \cdot \left( \sum_{i=0}^{\infty} p_i C^i \right) \cdot B^{-1} = B \cdot P(C) \cdot B^{-1} \end{aligned}$$

zeigt, dass  $P(A)$  genau dann konvergiert, wenn  $P(C)$  konvergiert.

Als nächstes überlegt man sich, dass man jeden Jordanblock einzeln behandeln kann, da für jedes Polynom  $Q$  gerade

$$Q \begin{pmatrix} J(\lambda_1, \ell_1) & & 0 \\ & \ddots & \\ 0 & & J(\lambda_k, \ell_k) \end{pmatrix} = \begin{pmatrix} Q(J(\lambda_1, \ell_1)) & & 0 \\ & \ddots & \\ 0 & & Q(J(\lambda_k, \ell_k)) \end{pmatrix}.$$

In den Übungen sehen Sie, dass

$$Q \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} = \begin{pmatrix} \frac{1}{0!} Q(\lambda) & \frac{1}{1!} Q'(\lambda) & \cdots & \frac{1}{(\ell-1)!} Q^{(\ell-1)}(\lambda) \\ & \frac{1}{0!} Q(\lambda) & \ddots & \vdots \\ & & \ddots & \frac{1}{1!} Q'(\lambda) \\ 0 & & & \frac{1}{0!} Q(\lambda) \end{pmatrix}.$$

Die höheren Ableitungen einer Potenzreihe sind wieder Potenzreihen mit dem gleichen Konvergenzradius. Falls  $|\lambda| < \rho$  konvergiert  $P(J(\lambda, \ell))$ , und es gilt

$$P(J(\lambda, \ell)) = \begin{pmatrix} \frac{1}{0!} P(\lambda) & & & 0 \\ \frac{1}{1!} P'(\lambda) & \frac{1}{0!} P(\lambda) & & \\ \vdots & \ddots & \ddots & \\ \frac{1}{(\ell-1)!} P^{(\ell-1)}(\lambda) & \cdots & \frac{1}{1!} P'(\lambda) & \frac{1}{0!} P(\lambda) \end{pmatrix}.$$

Damit das alles auch im Reellen funktioniert, betrachten wir  $A \in M_n(\mathbb{R})$  als komplexe Matrix  $A \in M_n(\mathbb{C})$ , bevor wir die Eigenwerte bestimmen; diese heißen dann die *komplexen Eigenwerte* von  $A$ .

**5.57. PROPOSITION.** *Es sei  $P$  eine Potenzreihe über  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  mit Konvergenzradius  $\rho > 0$  und  $A \in M_n(\mathbb{k})$ . Wenn alle komplexen Eigenwerte von  $A$  vom Betrag kleiner als  $\rho$  sind, dann konvergiert die Reihe  $P(A)$ . Hat ein komplexer Eigenwert größeren Betrag als  $\rho$ , dann divergiert sie.  $\square$*

Es ist also nicht einmal nötig, die Eigenwerte exakt zu bestimmen. Es reicht, den Betrag der Nullstellen des charakteristischen Polynoms  $\chi_A$  gegen  $\rho$  abzuschätzen. Das kann in Spezialfällen deutlich leichter sein. Auch die Jordan-Normalform selbst taucht in der Formulierung der Proposition nicht auf.

**5.58. BEISPIEL.** Der *Arcustangens* ist die Umkehrfunktion der Funktion

$$\tan = \frac{\sin}{\cos} : \mathbb{C} \setminus \left\{ (2n+1) \frac{\pi}{2} \mid n \in \mathbb{Z} \right\} \longrightarrow \mathbb{C}$$

und wird dargestellt durch die Reihe

$$\arctan(X) = X - \frac{1}{3} X^3 + \frac{1}{5} X^5 - \frac{1}{7} X^7 + \dots$$

mit Konvergenzradius 1. Das heißt, für alle  $z \in \mathbb{C}$  mit  $|z| < 1$  konvergiert

$$z - \frac{1}{3} z^3 + \frac{1}{5} z^5 - \frac{1}{7} z^7 + \dots$$

gegen den Wert  $\arctan z$ .

Wir betrachten speziell die Matrix

$$A = \begin{pmatrix} \frac{37}{\sqrt{3}} & -16 \\ 27 & -\frac{35}{\sqrt{3}} \end{pmatrix}.$$

Ihre Einträge sind so groß, dass man erst einmal nicht glaubt, dass die Reihe  $\arctan(A)$  konvergiert. Wir bestimmen das charakteristische Polynom von  $A$  und erhalten

$$\chi_F(X) = \det \begin{pmatrix} X - \frac{37}{\sqrt{3}} & 16 \\ -27 & X + \frac{35}{\sqrt{3}} \end{pmatrix} = X^2 - \frac{2}{\sqrt{3}}X + \frac{1}{3} = \left(X - \frac{1}{\sqrt{3}}\right)^2.$$

Da der einzige Eigenwert von  $A$  gleich  $\frac{1}{\sqrt{3}} < 1$  ist, konvergiert die Reihe  $\arctan(A)$ . In der Tat ist

$$A = \begin{pmatrix} \sqrt{3} & 4 \\ 2 & 3\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 1 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \begin{pmatrix} 3\sqrt{3} & -4 \\ -2 & \sqrt{3} \end{pmatrix}$$

$$\begin{aligned} \text{und} \quad \arctan(A) &= \begin{pmatrix} \sqrt{3} & 4 \\ 2 & 3\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} \arctan \frac{1}{\sqrt{3}} & 0 \\ \arctan' \frac{1}{\sqrt{3}} & \arctan \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \begin{pmatrix} 3\sqrt{3} & -4 \\ -2 & \sqrt{3} \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{3} & 4 \\ 2 & 3\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} \frac{\pi}{6} & 0 \\ \frac{3}{4} & \frac{\pi}{6} \end{pmatrix} \cdot \begin{pmatrix} 3\sqrt{3} & -4 \\ -2 & \sqrt{3} \end{pmatrix} \\ &= \begin{pmatrix} 9\sqrt{3} - \frac{3\pi}{2} & \frac{\sqrt{3}\pi}{2} - 12 \\ \frac{81}{4} & \frac{\pi}{6} - 9\sqrt{3} \end{pmatrix}. \end{aligned}$$

Dabei haben wir benutzt, dass

$$\tan \frac{\pi}{6} = \sin \frac{\pi}{6} / \cos \frac{\pi}{6} = \frac{1}{2} / \frac{\sqrt{3}}{2} = \frac{1}{\sqrt{3}},$$

somit  $\arctan \frac{1}{\sqrt{3}} = \frac{\pi}{6}$ , und dass  $\arctan'(z) = \frac{1}{1+z^2}$ , somit  $\arctan' \frac{1}{\sqrt{3}} = \frac{3}{4}$ .

Wir kommen zu einer zweiten Anwendung der Jordan-Normalform. Diesmal geht es um die Lösung von *gewöhnlichen linearen Differentialgleichungssystemen* mit konstanten Koeffizienten. Dabei sei eine Matrix  $A \in M_n(\mathbb{R})$  gegeben. Gesucht sind Funktionen  $f_1, \dots, f_n: \mathbb{R} \rightarrow \mathbb{R}$ , so dass gilt

$$(*) \quad \begin{pmatrix} f_1' \\ \vdots \\ f_n' \end{pmatrix} = A \cdot \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}.$$

Nach dem Satz von Picard-Lindelöf gibt es zu jedem Anfangsvektor  $v \in \mathbb{R}^n$  und zu jeder Startzeit  $t_0 \in \mathbb{R}$  eine eindeutige Lösung

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}: \mathbb{R} \longrightarrow \mathbb{R}^n$$

von (\*) mit  $f(t_0) = v$ . Diese Lösung ist auf ganz  $\mathbb{R}$  definiert und beliebig oft differenzierbar. Außerdem ist das Differentialgleichungssystem zeitunabhängig, das heißt, für alle  $s \in \mathbb{R}$  erhalten wir weitere Lösungen

$$f(\cdot + s) \in L \quad \text{mit} \quad t \longmapsto f(t + s) \in \mathbb{R}^n.$$

Unter einer *Fundamentallösung* versteht man eine Abbildung  $F: t \rightarrow M_n(\mathbb{R})$  mit  $F(0) = E_n$ , so dass für alle  $v \in \mathbb{R}^n$  die Abbildung

$$t \mapsto F(t) \cdot v \in \mathbb{R}^n$$

eine Lösung von (\*) mit Anfangswert  $v$  bei  $t = 0$  ist. Die Fundamentallösung erfüllt die Gleichung

$$F'(t) = A \cdot F(t)$$

mit  $F(0) = E_n$ . Wegen des Satzes von Picard-Lindelöf existiert sie und ist eindeutig bestimmt. Dann ist für alle  $t_0$  die Abbildung

$$t \mapsto F(t - t_0) \cdot v \in \mathbb{R}^n$$

die eindeutig bestimmte Lösung von (\*), die zur Zeit  $t_0$  den Wert  $v$  annimmt.

Als Ansatz wählen wir  $F(t) = \exp(tA)$ . Die Exponentialreihe hat Konvergenzradius  $\rho = \infty$ , wegen Proposition 5.57 konvergiert  $\exp(tA)$  also für alle  $t \in \mathbb{R}$ . Für  $t = 0$  gilt

$$\exp(0A) = A^0 = E_n .$$

Außerdem erwarten wir, dass

$$\frac{d}{dt} \exp(tA) = A \cdot \exp(tA) .$$

Im folgenden gehen wir zu Funktionen  $f: \mathbb{R} \rightarrow \mathbb{C}$  über, damit wir mit komplexen Matrizen und ihren Jordan-Normalformen rechnen können. Es sei  $B \in GL(n, \mathbb{C})$  invertierbar, und es sei  $f \in L$  eine Lösung von (\*), dann ist  $g = B \cdot f: \mathbb{R} \rightarrow \mathbb{C}^n$  eine Lösung des Differentialgleichungssystems

$$g' = B \cdot f' = B \cdot A \cdot f = (B \cdot A \cdot B^{-1}) \cdot g .$$

Sei  $F$  eine Fundamentallösung von (\*), dann ist entsprechend  $B \cdot F(t) \cdot B^{-1}$  eine Fundamentallösung des obigen Systems. Wir können also die Jordan-Normalform von  $A$  einsetzen, um die Fundamentallösung für (\*) zu bestimmen.

Für festes  $t$  gilt

$$\frac{d}{dx} e^{tx} = t e^{tx} .$$

Für einen Jordanblock  $J(\lambda, \ell)$  erhalten wir also

$$\exp(tJ(\lambda, \ell)) = \begin{pmatrix} \frac{1}{0!} e^{t\lambda} & \frac{t}{1!} e^{t\lambda} & \dots & \frac{t^{\ell-1}}{(\ell-1)!} e^{t\lambda} \\ & \frac{1}{0!} e^{t\lambda} & \ddots & \vdots \\ & & \ddots & \frac{t}{1!} e^{t\lambda} \\ 0 & & & \frac{1}{0!} e^{t\lambda} \end{pmatrix} .$$

Man überprüft jetzt leicht, dass dann tatsächlich

$$\frac{d}{dt} \exp(tJ(\lambda, \ell)) = J(\lambda, \ell) \cdot \exp(tJ(\lambda, \ell)) .$$

Da wir jede reelle Matrix über  $\mathbb{C}$  in Jordan-Normalform bringen können, erhalten wir das folgende Ergebnis.

5.59. PROPOSITION. *Es sei  $A \in M_n(\mathbb{R})$ . Dann ist*

$$t \longmapsto \exp(tA)$$

*die Fundamentallösung für das Differentialgleichungssystem (\*).*

### 5.10. Zusammenfassung

Wir fassen noch einmal die wichtigsten Aspekte dieses Kapitels zusammen. Wir haben uns zum einen mit Endomorphismen von Vektorräumen, Eigenwerten und Normalformen beschäftigt. Der Begriff des Eigenvektors ist in vielen Bereichen der Mathematik und auch in der Physik sehr wichtig, daher sollten wir möglichst viele verschiedene Charakterisierungen kennen. Außerdem haben wir gesehen, dass etwas allgemeinere Konzepte wie das eines invarianten Unterraums oder eines Hauptraumes fast ebenso wichtig sind.

Diagonalisierbare Endomorphismen sind, was Eigenwerte und Fragen des Rechenaufwandes angeht, der Optimalfall, trigonalisierbare Endomorphismen sind für manche praktischen Zwecke fast genauso gut. Daher sollten wir möglichst viele Kriterien für Diagonalisierbarkeit und Trigonalisierbarkeit kennen. In jedem Fall ist es wichtig, sowohl über Endomorphismen als auch über ihre darstellenden Matrizen bezüglich geschickter gewählter Basen reden zu können.

Auf der anderen Seite haben wir auch einige Fakten über Ringe kennengelernt. Division mit Rest in Euklidischen Ringen, die eindeutige Primfaktorzerlegung in  $\mathbb{Z}$  und  $\mathbb{k}[X]$  und der chinesische Restsatz gehören zur mathematischen Allgemeinbildung. In einer Algebra-Vorlesung wird dieses Gebiet noch vertieft.

In gewisser Weise ist dieses Kapitel aber auch prototypisch für (reine) Mathematik überhaupt. Wir haben uns am Ende von Abschnitt 5.1 zwei Fragen gestellt.

- Wann lassen sich Endomorphismen oder Matrizen diagonalisieren?
- Wie können wir einem Endomorphismus  $F \in \text{End}_{\mathbb{k}}(V)$  eine Matrix zuordnen, die ihn besonders gut beschreibt?

Und dann haben wir uns auf die Suche nach Antworten gemacht, und auf dem Weg viele interessante neue Begriffe und Methoden kennengelernt, die mit dem ursprünglichen Problem scheinbar nichts zu tun haben. Zum Beispiel Polynome:

- Die Polynome  $\chi_F$  und  $\mu_F \in \mathbb{k}[X]$  sind wichtige Invarianten von  $F$ .
- Das Paar  $(V, F)$  beschreibt selbst einen Modul über  $\mathbb{k}[X]$ .

Um mit Polynomen arbeiten zu können, haben wir (zum Teil bekannte) Techniken aus einem völlig anderen Kontext ( $\mathbb{Z}$ ) auf unser Problem übertragen. Und zu guter Letzt ist alles recht abstrakt geworden. Aber vielleicht reicht es, sich die Primfaktorzerlegung für ganze Zahlen vorstellen zu können, um mit der von Polynomen zu arbeiten. Oder es reicht, sich Vektorräume über  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  vorzustellen, um später auch mit Vektorräumen über anderen Körpern wie  $\mathbb{K}$  zu arbeiten.

## KAPITEL 6

# Vektorräume mit Skalarprodukt

In diesem Kapitel betrachten wir Vektorräume über  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$  mit Skalarprodukt. Wir haben bereits in Abschnitt 1.4 über Euklidische Geometrie gesehen, dass man mit Hilfe des Standard-Skalarproduktes Längen und Winkel bestimmen kann. Auch in der Physik spielen Skalarprodukte eine große Rolle.

In Abschnitt 3.4 haben wir gesehen, dass man mit Orthonormalbasen besonders gut rechnen kann, insbesondere braucht man die Inverse der Basisabbildung nicht umständlich auszurechnen. In diesem Kapitel konstruieren wir systematisch Orthogonalbasen und entsprechende Basen für Vektorräume mit Skalarprodukt über  $\mathbb{C}$  oder  $\mathbb{H}$ , mit denen man entsprechend einfach arbeiten kann.

Am Ende von Abschnitt 2.2 haben wir den Dualraum  $V^*$  eines Vektorraums  $V$  kennengelernt. Wir führen hier auch noch den sogenannten „Antidualraum“ ein und zeigen, wie beide über ein Skalarprodukt mit  $V$  identifiziert werden können, wenn  $V$  endlich-dimensional ist.

Lineare Abbildungen, die ein Skalarprodukt invariant lassen, heißen „lineare Isometrien“. Lineare Isometrien sind ein Spezialfall sogenannter „normaler Abbildungen“. Wir zeigen, dass normale Abbildungen bezüglich geeigneter Orthogonalbasen durch spezielle Matrizen dargestellt werden können. Dadurch erhalten wir einen einfacheren Zugang zur Klassifikation von Isometrien Euklidischer Vektorräume, vergleiche dazu die Überlegungen am Ende der Abschnitte 1.5 und 1.6.

### 6.1. Skalarprodukte

Im folgenden sei stets  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ , wenn nicht anders angegeben. In Definition 1.52 (1) haben wir das Standard-Skalarprodukt auf dem Vektorraum  $\mathbb{R}^n$  eingeführt als

$$\langle x, y \rangle = \sum_{a=1}^n x_a y_a \quad \text{für alle } x, y \in \mathbb{R}^n .$$

Wir haben in 1.52 (2) und (3) gesehen, wie man mit dem Skalarprodukt Längen von Vektoren und Winkel zwischen Vektoren definieren kann. Insbesondere ist

$$\|x\|^2 = \langle x, x \rangle = \sum_{a=1}^n x_a^2 \geq 0 ,$$

da die rechte Seite eine Summe von Quadraten ist.

Wenn wir diese Definition unverändert auf  $\mathbb{C}^n$  oder  $\mathbb{H}^n$  übertragen, haben wir ein Problem, denn für  $z \in \mathbb{C}$  gilt  $z^2 \in \mathbb{R}$  mit  $z^2 \geq 0$  nur dann, wenn bereits  $z$  eine reelle Zahl ist, also eine Zahl mit  $\operatorname{Im} z = 0$ . Wir erinnern uns daher an die Überlegung, die zur Definition 1.63 des komplexen Absolutbetrags geführt hat, siehe auch Bemerkung 1.64 (1). Es gilt

$$\bar{z} \cdot z = (x - yi)(x + yi) = x^2 + y^2 = \|z\|^2 \geq 0 \quad \text{für alle } z \in \mathbb{C}.$$

Völlig analog gilt nach Satz 1.72 (7), dass

$$\bar{q} \cdot q = \|q\|^2 \geq 0 \quad \text{für alle } q \in \mathbb{H},$$

dabei bezeichnet  $\|\cdot\|$  in beiden Gleichungen die Euklidische Norm auf  $\mathbb{C} \cong \mathbb{R}^2$  beziehungsweise  $\mathbb{H} \cong \mathbb{R}^4$ .

6.1. BEMERKUNG. Wir führen auch auf  $\mathbb{R}$  eine „Konjugation“ ein durch

$$\bar{t} = t \quad \text{für alle } t \in \mathbb{R}.$$

Wegen Bemerkung 1.61 und Satz 1.72 gilt für  $r, s \in \mathbb{k} = \mathbb{R}, \mathbb{C}$  und  $\mathbb{H}$  gleichermaßen

- (1)  $\overline{r + s} = \bar{r} + \bar{s},$
- (2)  $\overline{r \cdot s} = \bar{s} \cdot \bar{r},$
- (3)  $\bar{\bar{r}} = r,$
- (4)  $\bar{r} = r \iff r \in \mathbb{R} \subset \mathbb{k},$
- (5)  $\bar{r} \cdot r \geq 0 \quad \text{und} \quad \bar{r} \cdot r = 0 \iff r = 0,$

Wegen (1) und (2) nennen wir die Konjugation einen *Antiautomorphismus* von  $\mathbb{k}$ , da sie die Reihenfolge der Faktoren in einem Produkt umdreht. Wegen (5) können wir den Absolutbetrag

$$|r| = \sqrt{\bar{r} \cdot r} \in \mathbb{R}$$

definieren, siehe auch Definitionen 1.63 und 1.73 für  $\mathbb{k} = \mathbb{C}$  und  $\mathbb{H}$ . Eigenschaften des komplexen Absolutbetrages haben wir in Bemerkung 1.64 zusammengestellt. Die entsprechenden Aussagen über den quaternionischen Absolutbetrag lassen sich analog mit Hilfe von Satz 1.72 zeigen.

6.2. DEFINITION. Es sei  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ . Es sei  $V$  ein Rechts- $\mathbb{k}$ -Vektorraum und  $W$  ein Links- $\mathbb{k}$ -Vektorraum. Eine Abbildung  $\varphi: V \rightarrow W$  heißt ( $\mathbb{k}$ -) *semilinear* oder *antilinear*, wenn für alle  $u, v \in V$  und alle  $r \in \mathbb{k}$  gilt

- (L1)  $\varphi(u + v) = \varphi(u) + \varphi(v) \quad (\text{Additivität}),$
- ( $\bar{\text{L}}2$ )  $\varphi(v \cdot r) = \bar{r} \cdot \varphi(v) \quad (\text{Antihomogenität}).$

Analog definieren wir anti- oder semilineare Abbildungen von einem Links- in einen Rechtsvektorraum.

Wir benutzen die obigen Begriffe, um Axiome für Skalarprodukte anzugeben.

6.3. DEFINITION. Sei  $V$  ein Rechts- $\mathbb{k}$ -Vektorraum. Eine Abbildung  $S: V \times V \rightarrow \mathbb{k}$  heißt *Sesquilinearform*, wenn für alle  $u, v \in V$  die Abbildung

$$(S1) \quad \begin{array}{ll} S(u, \cdot): V \rightarrow \mathbb{k} & \text{mit } v \mapsto S(u, v) \text{ linear, und} \\ S(\cdot, v): V \rightarrow \mathbb{k} & \text{mit } u \mapsto S(u, v) \text{ antilinear ist.} \end{array}$$

Eine Sesquilinearform  $S: V \times V \rightarrow \mathbb{k}$  heißt *Hermiteische Form*, wenn für alle  $u, v \in V$  gilt

$$(S2) \quad S(v, u) = \overline{S(u, v)} \in \mathbb{k}.$$

Eine Hermiteische Form  $S: V \times V \rightarrow \mathbb{k}$  heißt *positiv semidefinit* oder kurz  $S \geq 0$ , wenn

$$S(v, v) \geq 0 \quad \text{für alle } v \in V.$$

Sie heißt *positiv definit* oder *positiv*, kurz  $S > 0$ , wenn für alle  $v \in V$  gilt

$$(S3) \quad S(v, v) \geq 0 \quad \text{und} \quad S(v, v) = 0 \iff v = 0.$$

Ein *Skalarprodukt* oder auch eine *Hermiteische Metrik* auf  $V$  ist eine positive definite Hermiteische Form  $g$  auf  $V$ . Wir nennen  $(V, g)$  einen *Euklidischen Vektorraum*, wenn  $\mathbb{k} = \mathbb{R}$ , einen *unitären Vektorraum*, wenn  $\mathbb{k} = \mathbb{C}$ , und einen *quaternionisch-unitären Vektorraum*, wenn  $\mathbb{k} = \mathbb{H}$ .

Die lateinische Vorsilbe „semi“ bedeutet „halb“. Eine semilineare Abbildung erfüllt nur die Hälfte der Axiome, daher der Name. Die Vorsilbe „sesqui“ bedeutet „anderthalb“. Eine Sesquilinearform ist in einem Argument linear, im anderen nur halb, also insgesamt nur anderthalbfach linear.

Man beachte, dass es für  $\mathbb{k} = \mathbb{R}$  keinen Unterschied zwischen semilinear und linear und zwischen sesquilinear und bilinear (also linear in beiden Argumenten) gibt, da die Konjugation auf  $\mathbb{R}$  die Identität ist. Genausowenig gibt es über  $\mathbb{R}$  einen Unterschied zwischen Hermiteisch und symmetrisch ( $S(u, v) = S(v, u)$  für alle  $u, v \in V$ ). Um eine einheitliche Notation zu haben, schreiben wir trotzdem die Konjugation auch für  $\mathbb{k} = \mathbb{R}$  immer mit.

6.4. BEMERKUNG. Wir wollen uns überlegen, dass die Definitionen 6.2 und 6.3 sinnvoll sind.

- (1) Semilineare Abbildungen sind mit den Vektorraumaxiomen verträglich. Wir prüfen insbesondere die Verträglichkeit mit dem Assoziativgesetz (M1). Für  $\varphi: V \rightarrow W$  wir oben und  $v \in V, r, s \in \mathbb{k}$  gilt

$$\begin{aligned} \varphi((v \cdot r) \cdot s) &= \bar{s} \cdot \varphi(v \cdot r) = \bar{s} \cdot \bar{r} \cdot \varphi(v) \\ \varphi(v \cdot (r \cdot s)) &= \overline{\bar{r} \cdot \bar{s}} \cdot \varphi(v) = \bar{s} \cdot \bar{r} \cdot \varphi(v). \end{aligned}$$

Dabei haben wir die Eigenschaft (2) der Konjugation aus Bemerkung 6.1 und die Antihomogenität ( $\overline{\bar{L}2}$ ) ausgenutzt. Die Verträglichkeit mit den anderen Axiomen zeigt man entsprechend.

- (2) Sei jetzt  $S$  eine Sesquilinearform auf  $V$ . Die Homogenität (L2) im zweiten Argument ist mit der Antihomogenität ( $\overline{L2}$ ) im ersten Argument verträglich, denn für  $u, v \in V$  und  $r, s \in \mathbb{k}$  gilt

$$\begin{aligned} S(u \cdot r, v \cdot s) &= \bar{r} \cdot S(u, v \cdot s) = \bar{r} \cdot S(u, v) \cdot s, \\ S(u \cdot r, v \cdot s) &= S(u \cdot r, v) \cdot s = \bar{r} \cdot S(u, v) \cdot s. \end{aligned}$$

Ohne Antihomogenität in einem der Argumente hätten wir für  $\mathbb{k} = \mathbb{H}$  Probleme bekommen, siehe Bemerkung 4.5. Aber das ist nicht der Hauptgrund dafür, Sesquilinearformen zu betrachten.

- (3) Wenn  $S$  Hermitesch und linear im zweiten Argument ist, folgt Selinearität im ersten Argument. Wir überprüfen nur Antihomogenität mit der folgenden Rechnung: Für  $u, v \in V$  und  $r \in \mathbb{k}$  gilt

$$S(u \cdot r, v) = \overline{S(v, u \cdot r)} = \overline{S(v, u) \cdot r} = \bar{r} \cdot \overline{S(v, u)} = \bar{r} \cdot S(u, v).$$

- (4) Der Hauptgrund dafür, dass wir mit Sesquilinear- statt mit Bilinearformen arbeiten, ist der folgende. Wenn wir zweimal dasselbe Argument  $v \in V$  einsetzen, gilt

$$S(v, v) = \overline{S(v, v)} \quad \implies \quad S(v, v) \in \mathbb{R}$$

nach (S2) und Bemerkung 6.1 (4). Insbesondere können wir nun verlangen, dass  $S(v, v) \geq 0$ . Hätten wir  $S(u, v) = S(v, u)$  gefordert, so erhielten wir für  $\mathbb{k} = \mathbb{C}$  ein Element in  $\mathbb{C}$ , und die Relation „ $\geq$ “ ist auf  $\mathbb{C}$  nicht definiert.

- (5) Schließlich gilt für  $v \in V$  und  $r \in \mathbb{k}$  noch, dass

$$S(v \cdot r, v \cdot r) = \bar{r} \cdot \underbrace{S(v, v)}_{\in \mathbb{R}} \cdot r = (\bar{r} \cdot r) \cdot S(v, v) = \underbrace{|r|^2}_{\geq 0} \cdot S(v, v)$$

wegen Bemerkung 6.1 (4), (5). Also verhält sich  $S(v, v)$  wie das Quadrat einer „Länge“.

In Definition 1.52 haben wir das Standard-Skalarprodukt auf  $\mathbb{R}^n$  kennengelernt. Wir wollen jetzt die Standard-Skalarprodukte auf  $\mathbb{C}^n$  und  $\mathbb{H}^n$  konstruieren. Dazu gehen wir einen kleinen Umweg über adjungierte Matrizen.

6.5. BEMERKUNG. In Definition 3.29 haben wir bereits die adjungierte Matrix  $A^* \in M_{n,m}(\mathbb{k})$  zu einer Matrix  $A \in M_{m,n}(\mathbb{k})$  definiert durch

$$A^* = (\bar{a}_{ji})_{i,j}, \quad \text{wobei} \quad A = (a_{ij})_{i,j}.$$

Wir haben sie benutzt, um die Methode der kleinsten Quadrate einzuführen. In Proposition 3.32 (2) haben wir einige Eigenschaften kennengelernt, unter anderem

$$(1) \quad (A \cdot B)^* = B^* \cdot A^* \quad \text{für alle } A \in M_{n,m}(\mathbb{k}) \text{ und alle } B \in M_{m,\ell}(\mathbb{k}),$$

außerdem gilt  $(A^*)^* = A$ . Somit hat Adjungieren ähnliche Eigenschaften wie die Konjugation, siehe Bemerkung 6.1 (1)–(3).

In Beispiel 2.51 haben wir den Rechts- $\mathbb{k}$ -Vektorraum  $\mathbb{k}^n = M_{n,1}(\mathbb{k})$  der Spalten und den Links- $\mathbb{k}$ -Vektorraum  ${}^n\mathbb{k} = M_{1,n}(\mathbb{k})$  der Zeilen definiert. Nach

Bemerkung 2.68 (3), (4) entspricht die Multiplikation mit Skalaren aus  $\mathbb{k}$  genau der Multiplikation mit  $1 \times 1$ -Matrizen. Dann ist die Abbildung

$$(2) \quad \cdot^* : \mathbb{k}^n \longrightarrow {}^n\mathbb{k} \quad \text{mit} \quad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \longmapsto v^* = (\bar{v}_1, \dots, \bar{v}_n)$$

semilinear, denn für  $1 \times 1$ -Matrizen  $r$  ist  $r^* = \bar{r}$ , und nach (1) oben gilt

$$(v \cdot r)^* = r^* \cdot v^* = \bar{r} \cdot v^* .$$

Die Umkehrabbildung  $\cdot^* : {}^n\mathbb{k} \rightarrow \mathbb{k}^n$  ist ebenfalls semilinear.

6.6. BEISPIEL. Seien  $u, v \in \mathbb{k}^n$  für  $\mathbb{k} = \mathbb{C}$  oder  $\mathbb{H}$ , dann definieren wir das komplexe und das quaternionische Standard-Skalarprodukt in Analogie zu Definition 1.52 (1) durch

$$\langle u, v \rangle = u^* \cdot v = \sum_{a=1}^n \bar{u}_a \cdot v_a \in \mathbb{k} = M_{1,1}(\mathbb{k}) .$$

Wir überprüfen die Axiome (S1)–(S3). Linearität im zweiten Argument ist leicht zu zeigen, und wegen Bemerkung 6.5 (1) ist  $\langle \cdot, \cdot \rangle$  Hermitesch. Dann folgt Sesquilinearität wie in Bemerkung 6.4 (3). Schließlich gilt

$$\langle v, v \rangle = v^* \cdot v = \sum_{a=1}^n \bar{v}_a \cdot v_a = \sum_{a=1}^n |v_a|^2 \geq 0$$

nach Bemerkung 6.1 (5), und Gleichheit ist nur möglich, wenn alle  $|v_a|^2 = 0$ , also alle  $v_a = 0$ , das heißt, wenn  $v = 0 \in \mathbb{k}^n$ . Somit ist  $\langle \cdot, \cdot \rangle$  auch positiv definit, also ein Skalarprodukt auf  $\mathbb{k}^n$ .

6.7. BEISPIEL. Wir geben ein Beispiel aus der Analysis. Dabei sei  $V = C^\infty([0, 1]; \mathbb{k})$  der Raum der unendlich oft differenzierbaren Funktionen auf dem Intervall  $[0, 1]$  mit Werten in  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ . Die folgenden Konstruktionen lassen sich auch auf anderen Intervallen an Stelle von  $[0, 1]$  durchführen.

(1) Das  $L^2$ -Skalarprodukt ist definiert durch

$$\langle f, g \rangle_{L^2} = \int_0^1 \overline{f(t)} g(t) dt \in \mathbb{k} .$$

Da  $f, g$  stetig sind, sind sie auf  $[0, 1]$  beschränkt, so dass das Riemann-Integral existiert. Das  $L^2$ -Skalarprodukt ist offensichtlich sesquilinear, Hermitesch und positiv semidefinit. Um zu sehen, dass es definit ist, sei  $f \neq 0$ . Also existiert  $t \in [0, 1]$  mit  $f(t) \neq 0$ . Wegen Stetigkeit existiert ein  $\varepsilon > 0$ , so dass  $|f(t)| \geq \varepsilon$  auf  $(t - \varepsilon, t + \varepsilon) \cap [0, 1]$ , somit

$$\langle f, f \rangle_{L^2} = \int_0^1 |f(t)|^2 dt \geq \int_{\max(0, t-\varepsilon)}^{\min(1, t+\varepsilon)} \varepsilon^2 dt \geq \varepsilon^3 > 0 .$$

(2) Wir versuchen es mit

$$\langle\langle f, g \rangle\rangle = \int_0^1 \overline{f'(t)} g'(t) dt = \langle f', g' \rangle_{L^2} \in \mathbb{k} .$$

Diese Hermitesche Sesquilinearform ist nur positiv semidefinit, denn für  $f \equiv c \in \mathbb{k}$  konstant gilt  $f'(t) \equiv 0$ , somit  $\langle\langle f, f \rangle\rangle = 0$ .

- (3) Wir addieren die beiden obigen Produkte und erhalten das (erste) Sobolev-Skalarprodukt

$$\langle f, g \rangle_{H^1} = \langle f, g \rangle_{L^2} + \langle f', g' \rangle_{L^2} \in \mathbb{k}.$$

Die Summe ist wieder positiv definit, denn für  $f \neq 0$  gilt

$$\langle f, f \rangle_{H^1} = \underbrace{\langle f, f \rangle_{L^2}}_{>0} + \underbrace{\langle f', f' \rangle_{L^2}}_{\geq 0} > 0.$$

In Analysis lernen Sie, dass zwei Skalarprodukte  $\langle \cdot, \cdot \rangle_1$  und  $\langle \cdot, \cdot \rangle_2$  auf einem endlich-dimensionalen Vektorraum  $V$  vergleichbar sind, das heißt, es gibt eine Konstante  $C > 0$ , so dass

$$\frac{1}{C} \langle v, v \rangle_1 \leq \langle v, v \rangle_2 \leq C \langle v, v \rangle_1.$$

Die obigen zwei Skalarprodukte auf  $C^\infty([0, 1]; \mathbb{k})$  sind nicht vergleichbar. Zwar gilt offensichtlich

$$\langle f, f \rangle_{L^2} \leq \langle f, f \rangle_{H^1},$$

aber für die Folge  $f_n(x) = x^n$  gilt

$$\begin{aligned} \langle f_n, f_n \rangle_{L^2} &= \int_0^1 x^{2n} dx = \left( \frac{1}{2n+1} x^{2n+1} \right) \Big|_{x=0}^1 = \frac{1}{2n+1}, \\ \langle f'_n, f'_n \rangle_{L^2} &= \int_0^1 n^2 x^{2n-2} dx = \left( \frac{n^2}{2n-1} x^{2n-1} \right) \Big|_{x=0}^1 = \frac{n^2}{2n-1}, \\ \langle f_n, f_n \rangle_{H^1} &= \langle f_n, f_n \rangle_{L^2} + \langle f'_n, f'_n \rangle_{L^2} = \frac{2n^3 + n^2 + 2n - 1}{(2n+1)(2n-1)}, \end{aligned}$$

und man sieht leicht, dass die Folge

$$\left( \frac{\langle f_n, f_n \rangle_{H^1}}{\langle f_n, f_n \rangle_{L^2}} \right)_n = \left( \frac{2n^3 + n^2 + 2n - 1}{2n - 1} \right)$$

unbeschränkt ist für  $n \rightarrow \infty$ .

Ab sofort verwenden wir für Skalarprodukte die Buchstaben  $g, h, \dots$ , da wir den Buchstaben  $B$  später wieder für Basen und Basisabbildungen benutzen wollen.

6.8. DEFINITION. Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt. Dann definieren wir die *Norm* zum Skalarprodukt  $g$  durch

$$\|v\|_g = \sqrt{g(v, v)} \in \mathbb{R}.$$

Im Falle  $\mathbb{k} = \mathbb{R}$  nennt man diese Norm auch die *Euklidische Norm* zum Skalarprodukt  $g$ , vergleiche Definition 1.52 (2).

6.9. BEMERKUNG. Es sei  $(V, g)$  ein Rechts- $\mathbb{k}$ -Vektorraum mit Skalarprodukt. Dann gelten die *Norm-Axiome*

- (N1)  $\|v\|_g \geq 0$  und  $\|v\|_g = 0 \iff v = 0$  (*Positivität*),  
 (N2)  $\|v \cdot r\|_g = |r| \cdot \|v\|_g$  (*Homogenität*),  
 (N3)  $\|v + w\|_g \leq \|v\|_g + \|w\|_g$  (*Dreiecksungleichung*),

für alle  $v, w \in V$  und  $r \in \mathbb{k}$ . Jede Abbildung  $\|\cdot\| : V \rightarrow \mathbb{R}$ , die (N1)–(N3) erfüllt heißt eine *Norm* auf  $V$ .

Da  $g$  nach (S3) positiv definit ist, folgt (N1). Aus Bemerkung 6.4 (5) ergibt sich unmittelbar (N2). Für jede Zahl  $r \in \mathbb{k}$  gilt  $r + \bar{r} \in \mathbb{R}$  wegen Bemerkung 6.1 (3) und (4). Außerdem folgt

$$|r + \bar{r}|^2 + |r - \bar{r}|^2 = (r + \bar{r})^2 - (r - \bar{r})^2 = 2r\bar{r} + 2\bar{r}r = 4|r|^2,$$

so dass insbesondere  $r + \bar{r} \leq 2|r|$  gilt. Dabei haben wir  $|\bar{r}| = |r|$  benutzt, siehe Bemerkung 1.64 (5) im Falle  $\mathbb{k} = \mathbb{C}$ . Mit der Cauchy-Schwarz-Ungleichung, siehe Satz 1.54 und Satz 6.10 unten, ergibt sich

$$\begin{aligned} \|v + w\|_g^2 &= g(v + w, v + w) = \|v\|_g^2 + g(v, w) + g(w, v) + \|w\|_g^2 \\ &\leq \|v\|_g^2 + 2|g(v, w)| + \|w\|_g^2 \\ &\leq \|v\|_g^2 + 2\|v\|_g \|w\|_g + \|w\|_g^2 = (\|v\|_g + \|w\|_g)^2. \end{aligned}$$

Wurzelziehen liefert die Dreiecksungleichung (N3).

6.10. SATZ (Cauchy-Schwarz-Ungleichung). *Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt. Dann gilt für alle Vektoren  $v, w \in V$ , dass*

$$|g(v, w)| \leq \|v\|_g \cdot \|w\|_g.$$

*Gleichheit gilt genau dann, wenn  $v$  und  $w$  linear abhängig sind.*

BEWEIS. Wir passen den Beweis von Satz 1.54 an.

Fall 1: Es sei  $v = 0$ . Dann gilt

$$g(v, w) = g(0, w) = 0 = \|0\|_g \cdot \|w\|_g = \|v\|_g \cdot \|w\|_g.$$

Also gilt Gleichheit, und  $v$  und  $w$  sind offensichtlich linear abhängig.

Fall 2: Es sei  $v \neq 0$ , dann folgt  $\|v\|_g^2 = g(v, v) > 0$ , und wir berechnen

$$\begin{aligned} 0 &\leq \left\| w - v \cdot \frac{g(v, w)}{\|v\|_g^2} \right\|_g = g\left( w - v \cdot \frac{g(v, w)}{\|v\|_g^2}, w - v \cdot \frac{g(v, w)}{\|v\|_g^2} \right) \\ &= \|w\|_g^2 - \frac{\overline{g(v, w)}}{\|v\|_g^2} g(v, w) - \underbrace{g(w, v)}_{= \overline{g(v, w)}} \frac{g(v, w)}{\|v\|_g^2} + \frac{\overline{g(v, w)}}{\|v\|_g^2} \underbrace{\|v\|_g^2}_{\in \mathbb{R}} \frac{g(v, w)}{\|v\|_g^2} \\ &= \|w\|_g^2 - \frac{|g(v, w)|^2}{\|v\|_g^2}. \end{aligned}$$

Hieraus ergibt sich die Ungleichung durch elementare Umformungen.

Wenn Gleichheit gilt, dann folgt aus (S3) (oder äquivalent aus (N1)), dass

$$0 = w - v \cdot \frac{g(v, w)}{\|v\|_g^2},$$

insbesondere sind  $v$  und  $w$  dann linear abhängig. Seien umgekehrt  $v$  und  $w$  linear abhängig, dann gilt  $w = v \cdot r$ , da  $v \neq 0$  nach Annahme. Wir erhalten also

$$|g(v, w)| = |g(v, v \cdot r)| = \left| \|v\|_g^2 r \right| = \|v\|_g^2 |r| = \|v\|_g \cdot \|v \cdot r\|. \quad \square$$

6.11. BEMERKUNG. Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt. Dann erfüllt die Norm  $\|\cdot\|_g$  für alle  $v, w \in V$  die *Parallelogramm-Identität*

$$(1) \quad \|v + w\|_g^2 + \|v - w\|_g^2 = 2\|v\|_g^2 + 2\|w\|_g^2,$$

wie man leicht nachrechnet.

Man kann das Skalarprodukt aus der Norm  $\|\cdot\|_g$  zurückgewinnen mit Hilfe der *Polarisationsformeln*

$$(2) \quad g(v, w) = \frac{1}{4} (\|v + w\|_g^2 - \|v - w\|_g^2) \quad \text{falls } \mathbb{k} = \mathbb{R},$$

$$(3) \quad g(v, w) = \frac{1}{4} (\|v + w\|_g^2 - \|v - w\|_g^2) \\ - \frac{i}{4} (\|v + w \cdot i\|_g^2 - \|v - w \cdot i\|_g^2) \quad \text{falls } \mathbb{k} = \mathbb{C},$$

$$(4) \quad g(v, w) = \frac{1}{4} (\|v + w\|_g^2 - \|v - w\|_g^2) \\ - \frac{i}{4} (\|v + w \cdot i\|_g^2 - \|v - w \cdot i\|_g^2) \\ - \frac{j}{4} (\|v + w \cdot j\|_g^2 - \|v - w \cdot j\|_g^2) \\ - \frac{k}{4} (\|v + w \cdot k\|_g^2 - \|v - w \cdot k\|_g^2) \quad \text{falls } \mathbb{k} = \mathbb{H}.$$

Darüberhinaus kann man zeigen, dass jede Norm auf einem  $\mathbb{k}$ -Vektorraum  $V$ , die die Parallelogrammidentität (1) erfüllt, von einem Skalarprodukt auf  $V$  herkommt, das man mit Hilfe der passenden Polarisationsformel berechnen kann.

## 6.2. Skalarprodukte als Matrizen

In diesem Abschnitt stellen wir Sesquilinearformen auf endlich-dimensionalen Vektorräumen bezüglich einer Basis als Matrizen dar. Wir untersuchen die Eigenschaften dieser Matrizen und geben Kriterien dafür, dass eine solche Matrix ein Hermitesches und positiv definites Skalarprodukt darstellt.

Die darstellende Matrix hat eine besonders einfache Gestalt, wenn die Basisvektoren alle Länge 1 haben und paarweise aufeinander senkrecht stehen. Solche Orthonormalbasen haben wir bereits in Abschnitt 2.5 kennengelernt. Wir lernen ein Verfahren kennen, das Orthonormalbasen mit speziellen Eigenschaften produziert. In diesem Zusammenhang beweisen wir auch ein Kriterium dafür, ob eine Matrix positiv definit ist.

6.12. DEFINITION. Es sei  $(V, g)$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum mit Skalarprodukt, und es sei  $B = (b_1, \dots, b_n)$  eine Basis von  $V$ . Dann definieren wir die *Gramsche Matrix*  $G \in M_n(\mathbb{k})$  von  $g$  durch

$$G = (g(b_i, b_j))_{i,j} \in M_n(\mathbb{k}).$$

6.13. BEISPIEL. Wir schränken das  $L^2$ -Skalarprodukt aus Beispiel 6.7 (1) auf dem Raum  $C^\infty([0, 1]; \mathbb{k})$  auf den  $(n+1)$ -dimensionalen Raum der Polynome  $P$  vom Grad  $\deg P \leq n$  ein. Als Basis wählen wir die Polynome  $f_0(x) = x^0, \dots, f_n(x) = x^n$ . Dann erhalten wir als Gramsche Matrix

$$\begin{aligned} G = (\langle f_i, f_j \rangle)_{i,j} &= \left( \int_0^1 x^i \cdot x^j dx \right)_{i,j} = \left( \frac{x^{i+j+1}}{i+j+1} \Big|_{x=0}^1 \right) = \left( \frac{1}{i+j+1} \right)_{i,j} \\ &= \begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n+1} \\ \frac{1}{2} & \frac{1}{3} & & \frac{1}{n+2} \\ \vdots & & \ddots & \vdots \\ \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n+1} \end{pmatrix}. \end{aligned}$$

6.14. DEFINITION. Eine quadratische Matrix  $A = (a_{ij})_{i,j} = M_n(R)$  heißt *symmetrisch*, wenn  $A = A^t$ , das heißt, wenn  $a_{ij} = a_{ji}$  für alle  $i, j$  gilt. Eine quadratische Matrix  $A = (a_{ij})_{i,j} = M_n(\mathbb{k})$  heißt *selbstadjungiert* oder *Hermiteisch*, wenn  $A = A^*$ , das heißt, wenn  $a_{ij} = \bar{a}_{ji}$  für alle  $i, j$  gilt.

Eine Hermiteische Matrix  $A = (a_{ij})_{i,j} = M_n(\mathbb{k})$  heißt *positiv semidefinit*, wenn

$$x^* \cdot A \cdot x \geq 0 \quad \text{für alle } x \in \mathbb{k}^n.$$

Sie heißt *positiv definit*, wenn

$$x^* \cdot A \cdot x \geq 0 \quad \text{und} \quad x^* \cdot A \cdot x = 0 \iff x = 0 \quad \text{für alle } x \in \mathbb{k}^n.$$

Die Definition von positiv (semi-) definit ist sinnvoll, da für eine Hermiteische Matrix  $A$  gilt, dass

$$x^* A x = x^* A^* (x^*)^* = (x^* A x)^* = \overline{x^* A x} \in \mathbb{R} \subset \mathbb{k}$$

nach Bemerkung 6.1 (4).

Wir beachten, dass die Begriffe „Hermiteisch“, „selbstadjungiert“ und „symmetrisch“ für  $\mathbb{k} = \mathbb{R}$  gleichbedeutend sind. Wir werden im Folgenden auch für  $\mathbb{k} = \mathbb{R}$  die Begriffe „Hermiteisch“ und „selbstadjungiert“ verwenden. Dabei benutzt man „Hermiteisch“ eher für Matrizen, die Skalarprodukte darstellen, und „selbstadjungiert“ für Matrizen, die Endomorphismen darstellen.

6.15. BEMERKUNG. Es sei  $S$  eine Sesquilinearform auf einem endlich-dimensionalen  $\mathbb{k}$ -Vektorraum, und es sei  $B = (b_1, \dots, b_n)$  eine Basis von  $V$ . Wie in Definition 6.12 betrachten wir die Matrix

$$A = (S(b_p, b_q))_{p,q} \in M_n(\mathbb{k}).$$

Wenn  $S$  ein Skalarprodukt ist, handelt es sich dabei gerade um die Gramsche Matrix.

- (1) Die Matrix  $A$  legt  $S$  eindeutig fest. Denn seien  $v, w \in V$ , dann existieren Koordinaten  $x = (x_p)_p, y = (y_q)_q \in \mathbb{k}^n$ , so dass

$$v = B(x) = \sum_{p=1}^n b_p \cdot x_p \quad \text{und} \quad w = B(y) = \sum_{q=1}^n b_q \cdot y_q,$$

siehe Definition 2.63. Da  $S$  eine Sesquilinearform ist, folgt

$$S(v, w) = S\left(\sum_{p=1}^n b_p \cdot x_p, \sum_{q=1}^n b_q \cdot y_q\right) = \sum_{p,q=1}^n \bar{x}_p S(b_p, b_q) y_q = x^* A y.$$

Umgekehrt liefert die obige Formel zu jeder Matrix  $A \in M_n(\mathbb{k})$  eine Sesquilinearform  $S$  auf  $V$ .

- (2) Die Sesquilinearform  $S$  ist genau dann Hermitesch, wenn die Matrix  $A$  Hermitesch ist. Da  $S(b_q, b_p) = \overline{S(b_p, b_q)}$ , ist die Richtung „ $\Rightarrow$ “ klar.

Zu „ $\Leftarrow$ “ seien  $v = B(x)$  und  $w = B(y) \in V$  wie oben und  $A$  sei Hermitesch. Aus (1) folgt

$$S(w, v) = y^* A x = y^* A^* (x^*)^* = (x^* A y)^* = \overline{S(v, w)}.$$

- (3) Sei  $S$  Hermitesch, dann ist  $S$  genau dann positiv (semi-) definit, wenn  $A$  positiv (semi-) definit ist. Die Basisabbildung  $B: \mathbb{k}^n \rightarrow V$  ist bijektiv, also gilt  $S(v, v) \geq 0$  wegen (1) genau dann für alle  $v \in V$ , wenn

$$x^* A x = S(B(x), B(x)) \geq 0$$

für alle  $x \in \mathbb{k}^n$  gilt. Entsprechend gilt  $S(v, v) = 0$  genau dann nur für  $v = 0$ , wenn  $x^* A x = 0$  nur für  $x = 0$  gilt. In Folgerung 6.19 (4) lernen wir noch ein etwas griffigeres Kriterium für positive Definitheit kennen, bei dem wir  $x^* A x$  nicht für alle  $x \in \mathbb{k}^n$  testen müssen.

- (4) Zum Schluss betrachten wir noch das Verhalten der darstellenden Matrix  $A$  unter Basiswechsel. Dazu seien  $B = (b_1, \dots, b_n)$  und  $C = (c_1, \dots, c_n)$  Basen von  $V$ . Dann existiert eine Matrix  $M = (m_{pq})_{p,q} \in GL(n, \mathbb{k})$ , so dass

$$c_q = \sum_{p=1}^n b_p \cdot m_{pq},$$

siehe Bemerkung 2.76. Es sei  $A$  wie oben die darstellende Matrix zur Basis  $B$ , dann erhalten wir zur Basis  $C$  die darstellende Matrix

$$\begin{aligned} (S(c_p, c_q))_{p,q} &= \left( S\left(\sum_{r=1}^n b_r \cdot m_{rp}, \sum_{s=1}^n b_s \cdot m_{sq}\right) \right)_{p,q} \\ &= \left( \sum_{r,s=1}^n \bar{m}_{rp} S(b_r, b_s) m_{sq} \right)_{p,q} = M^* A M. \end{aligned}$$

Sei  $v = C(s)$  mit  $s \in \mathbb{k}^n$ , dann folgt

$$v = \sum_{q=1}^n c_q \cdot s_q = \sum_{p,q=1}^n b_p \cdot m_{pq} \cdot s_q = \sum_{p=1}^n b_p x_p,$$

wobei  $x = M \cdot s \in \mathbb{k}^n$ . In der Tat gilt für  $v = C(s) = B(x)$  und  $w = C(t) = B(y)$  mit  $x = M s$  und  $y = M t$ , dass

$$S(v, w) = x^* A y = (M s)^* A (M t) = s^* (M^* A M) t.$$

Das Standardskalarprodukt  $\langle \cdot, \cdot \rangle$  auf  $\mathbb{k}^n$  aus Beispiel 6.6 wird bezüglich der Standardbasis durch die Einheitsmatrix dargestellt:  $\langle e_i, e_j \rangle = \delta_{ij}$ . Somit ist die Standardbasis eine Orthonormalbasis für das Standardskalarprodukt, siehe Definition 3.30. Wir wollen den Begriff der Orthonormalbasis jetzt auf beliebige Vektorräume mit Skalarprodukt ausdehnen.

6.16. DEFINITION. Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt. Ein Tupel  $(v_1, \dots, v_k)$  von Elementen von  $V$  heißt *orthogonal* oder auch (*paarweise senkrecht*), wenn

$$g(v_i, v_j) = 0 \quad \text{für alle } i, j \text{ mit } i \neq j.$$

Wenn  $(v_1, \dots, v_k)$  außerdem eine Basis bildet, nennt man diese eine *Orthonormalbasis*.

Dann heißt eine Basis  $B = (b_1, \dots, b_n)$  von  $V$  eine *Orthonormalbasis* von  $V$ , wenn

$$g(b_i, b_j) = \delta_{ij}.$$

Eine Orthonormalbasis eines  $\mathbb{k}$ -Vektorraums heißt manchmal auch *unitäre Basis* ( $\mathbb{k} = \mathbb{C}$ ), beziehungsweise *quaternionisch-unitäre Basis* ( $\mathbb{k} = \mathbb{H}$ ).

6.17. BEMERKUNG. Es sei  $(V, g)$  ein endlich-dimensionaler Vektorraum mit Skalarprodukt.

- (1) Jedes orthogonale Tupel  $(v_1, \dots, v_k)$  mit  $v_i \neq 0$  für alle  $i$  ist linear unabhängig, denn sei

$$0 = \sum_{p=1}^k v_p \cdot r_p,$$

dann folgt für alle  $q$ , dass

$$0 = g\left(v_q, \sum_{p=1}^k v_p \cdot r_p\right) = \sum_{p=1}^k g(v_p, v_q) \cdot r_p = \|v_q\|_g^2 r_q.$$

Aus  $v_q \neq 0$  folgt  $\|v_q\|_g \neq 0$ , und somit  $r_q = 0$ . Also sind  $v_1, \dots, v_k$  linear unabhängig.

- (2) Sei  $\dim V = n$ , dann bildet ein orthogonales  $n$ -Tupel von Vektoren eine Basis, wenn keiner der Vektoren verschwindet, also eine Orthogonalbasis. Das folgt aus (1) und dem Basissätzen 3.3 und 3.4 von Steinitz, siehe auch Aufgabe 2 von Blatt 11 zur Linearen Algebra I.
- (3) In Definition 2.63 hatten wir die Koordinatenabbildung als Inverse der Basisabbildung  $B: \mathbb{k}^n \rightarrow V$  eingeführt. Es sei  $B = (b_1, \dots, b_n)$  eine

Orthonormalbasis, dann wird die Koordinatenabbildung  $B^{-1}: V \rightarrow \mathbb{k}^n$  beschrieben durch die Formel

$$B^{-1}(v) = \begin{pmatrix} g(b_1, v) \\ \vdots \\ g(b_n, v) \end{pmatrix},$$

denn sei  $v = B(x)$  mit  $x \in \mathbb{k}^n$ , dann gilt

$$g(b_p, v) = g\left(b_p, \sum_{q=1}^n b_q \cdot x_q\right) = \sum_{q=1}^n g(b_p, b_q) x_q = x_p.$$

Eine ähnliche Aussage hat wir in Proposition 3.34 bereits für das Standard-Skalarprodukt bewiesen.

Im Folgenden bezeichnen wir das Erzeugnis von  $v_1, \dots, v_p$  mit  $\langle v_1, \dots, v_p \rangle$ . Für das Skalarprodukt verwenden wir wieder den Buchstaben  $g$ , um Verwechslungen zu vermeiden. Die Axiome (S1)–(S3) bleiben gültig, wenn man  $g$  auf einen Unterraum einschränkt. Insbesondere ist also  $g|_{\langle v_1, \dots, v_p \rangle \times \langle v_1, \dots, v_p \rangle}$  wieder ein Skalarprodukt, das wir der Kürze halber wieder mit  $g$  bezeichnen.

6.18. SATZ (Gram-Schmidt-Orthonormalisierungsverfahren). *Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt und  $(v_1, \dots, v_n)$  sei eine Basis von  $V$ . Dann existieren eindeutig bestimmte Vektoren  $b_1, \dots, b_n \in V$ , so dass für alle  $p = 1, \dots, n$  gilt:*

- (1)  $(b_1, \dots, b_p)$  ist eine  $g$ -Orthonormalbasis von  $\langle v_1, \dots, v_p \rangle \subset V$ , und
- (2) es gilt  $g(b_p, v_p) \in \mathbb{R}$  und  $g(v_p, b_p) > 0$ .

Dazu konstruiert man  $b_p$  induktiv durch

$$b_p = w_p \cdot \frac{1}{\|w_p\|_g}, \quad \text{wobei} \quad w_p = v_p - \sum_{q=1}^{p-1} b_q \cdot g(b_q, v_p).$$

Insbesondere erhalten wir am Ende eine Orthonormalbasis  $(b_1, \dots, b_n)$  von  $(V, g)$ . Für viele Anwendungen reicht das, aber manchmal möchten wir die volle Stärke der Eigenschaften (1) und (2) ausnutzen.

BEWEIS. Wir beweisen den Satz durch Induktion. Für  $p = 0$  ist nichts zu zeigen.

Sei also  $p \geq 1$ , und seien  $b_1, \dots, b_{p-1}$  bereits konstruiert. Wir beginnen mit der Existenzaussage und definieren  $w_p$  wie oben. Nach Voraussetzung liegen  $b_1, \dots, b_{p-1} \in \langle v_1, \dots, v_{p-1} \rangle$ , also betrachten wir

$$w_p = v_p - \sum_{q=1}^{p-1} b_q \cdot g(b_q, v_p) \in \langle v_1, \dots, v_p \rangle.$$

Da die  $v_q$  linear unabhängig sind, gilt

$$v_p \notin \langle v_1, \dots, v_{p-1} \rangle = \langle b_1, \dots, b_{p-1} \rangle,$$

also auch  $w_p \notin \langle v_1, \dots, v_{p-1} \rangle$ , insbesondere  $w_p \neq 0$ , so dass wir  $b_p$  wie oben definieren dürfen. Für  $q \leq p-1$  berechnen wir

$$g(b_q, b_p) = g\left(b_q, v_p - \sum_{r=1}^{p-1} b_r \cdot g(b_r, v_p)\right) \frac{1}{\|w_p\|_g} = \frac{g(b_q, v_p) - g(b_q, v_p)}{\|w_p\|_g} = 0.$$

Außerdem gilt  $\|b_p\| = 1$  nach Konstruktion, und die Vektoren  $b_1, \dots, b_{p-1}$  sind nach Induktionsvoraussetzung orthogonal und normiert, also ist (1) erfüllt.

Aus (1) und der Konstruktion von  $b_p$  folgern wir (2), denn es gilt

$$g(b_p, v_p) = g\left(b_p, v_p - \sum_{q=1}^{p-1} b_q \cdot g(b_q, v_p)\right) = g(b_p, w_p) = \|w_p\|_g > 0.$$

Damit ist die Existenz von  $b_p$  mit den gewünschten Eigenschaften bewiesen.

Wir kommen zur Eindeutigkeit. Da  $b_1, \dots, b_{p-1}$  durch (1) und (2) bereits eindeutig bestimmt sind, brauchen wir im Induktionsschritt nur noch die Eindeutigkeit von  $b_p$  zu beweisen. Es sei also  $v \in \langle v_1, \dots, v_p \rangle$  ein weiterer Vektor, so dass  $g(b_q, v) = 0$  für  $1 \leq q < p$ ,  $\|v\|_g = 1$  und  $g(v, v_p) > 0$ . Wir stellen  $v$  in der Orthonormalbasis  $(b_1, \dots, b_p)$  von  $(v_1, \dots, v_p)$  dar als

$$v = \sum_{q=1}^p b_q \cdot x_q.$$

Dann folgt als erstes  $x_q = g(b_q, v) = 0$  für alle  $1 \leq q < p$ , so dass  $v = b_p \cdot x_p$ . Es folgt

$$|x_p| = |x_p| \|b_p\|_g = \|v\|_g = 1.$$

Da  $g(v_p, b_p) = \overline{g(b_p, v_p)} > 0$ , gilt außerdem

$$\bar{x}_p g(b_p, v_p) = g(b_p \cdot x_p, v_p) = g(v, v_p) > 0,$$

also auch  $\bar{x}_p > 0$ , und daher  $x_p > 0$ . Aber die einzige Zahl  $x_p \in \mathbb{k}$  mit  $|x_p| = 1$ ,  $x_p \in \mathbb{R}$  und  $x_p > 0$  ist 1. Also folgt  $v = b_p$ , und die Eindeutigkeit ist ebenfalls gezeigt.  $\square$

Es sei  $A \in M_n(\mathbb{k})$  eine quadratische Matrix und  $r \leq n$ , dann schreiben wir  $A_r \in M_r(\mathbb{k}_k)$  für den oberen linken  $r \times r$ -Block

$$A_r = ((a_{p,q})_{p,q \leq r}) = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}.$$

In Folgerung 4.15 (1) haben wir gesehen, dass  $\det A^t = \det A$  für alle quadratischen Matrizen  $M_n(\mathbb{k})$  gilt, wenn  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  ein Körper ist. Da die Konjugation mit den Rechenoperationen in  $\mathbb{k}$  verträglich ist, sieht man anhand der Leibniz-Formel aus Satz 4.13, dass  $\det \bar{A} = \overline{\det A}$  gilt. Insgesamt folgt daraus

$$\det A^* = \det \bar{A}^t = \det \bar{A} = \overline{\det A}.$$

6.19. FOLGERUNG. *Es sei  $A = (a_{pq})_{p,q} \in M_n(\mathbb{k})$  eine quadratische Matrix. Dann sind die folgenden Aussagen äquivalent.*

- (1) *Die Matrix  $A$  ist Hermitesch und positiv definit.*
- (2) *Cholesky-Zerlegung. Es gibt eine obere Dreiecksmatrix  $B$  mit reellen, positiven Diagonaleinträgen, so dass  $A = B^*B$ .*
- (3) *Es gibt eine invertierbare Matrix  $B \in GL(n, \mathbb{k})$  mit  $A = B^*B$ .*

*Falls  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$ , sind die obigen Aussagen außerdem äquivalent zum*

- (4) *Sylvester- oder auch Hurwitz-Kriterium. Die Matrix  $A$  ist Hermitesch, und für alle  $r = 1, \dots, n$  gilt  $\det A_r > 0$ .*

In der Analysis wendet man (4) auf die Hesse-Matrix einer Funktion an einem kritischen Punkt an, um festzustellen, ob ein lokales Minimum vorliegt. Um ein lokales Maximum nachzuweisen, braucht man analog zu (4) ein Kriterium für negative Definitheit. Dazu betrachten wir anstelle einer Hermiteschen Matrix  $A$  die Matrix  $-A$  und sehen, dass

$$\begin{aligned} (v^*Av \leq 0 \quad \text{und} \quad v^*Av = 0 \Leftrightarrow v = 0) \\ \iff \quad (-1)^r \det A_r > 0 \quad \text{für alle } r = 1, \dots, n. \end{aligned}$$

Achtung: Das Sylvester-Kriterium funktioniert nicht für positiv semidefinite Matrizen. Beispielsweise sei

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

dann gilt  $\det A_1 = a_{11} = 0 \geq 0$  und  $\det A_2 = \det A = 0 \geq 0$ , aber die Matrix  $A$  ist nicht positiv semidefinit, denn  $e_2^*Ae_2 = -1$ .

BEWEIS. Zu „(1)  $\implies$  (2)“ fassen wir  $A$  als Gramsche Matrix eines Skalarproduktes

$$g(x, y) = x^*Ay \in \mathbb{k}$$

auf  $\mathbb{k}^n$  auf. Wir konstruieren eine Orthonormalbasis  $(v_1, \dots, v_n)$  von  $\mathbb{k}^n$  mit dem Gram-Schmidt-Verfahren, beginnend mit der Standardbasis  $(e_1, \dots, e_n)$ . Es sei  $B \in M_n(\mathbb{k})$  die Basiswechselmatrix, so dass

$$e_q = \sum_{p=1}^n v_p \cdot b_{pq}.$$

Dann ist  $B$  eine obere Dreiecksmatrix nach Satz 6.18 (1), denn aus  $e_q \in \langle e_1, \dots, e_q \rangle = \langle v_1, \dots, v_q \rangle$  folgt  $b_{pq} = 0$  für  $p > q$ . Die Diagonaleinträge sind reell und positiv nach Satz 6.18 (2), denn

$$b_{qq} = g(v_q, v_q \cdot b_{qq}) = g\left(v_q, \sum_{p=1}^n v_p \cdot b_{pq}\right) = g(v_q, e_q) > 0.$$

Schließlich gilt  $A = B^*B$  nach Bemerkung 6.15 (4).

Der Schritt „(2)  $\implies$  (3)“ folgt, da Dreiecksmatrizen mit von 0 verschiedenen Diagonaleinträgen mit dem Gauß-Verfahren 3.26 invertiert werden können.

Zu „(3)  $\implies$  (1)“ überlegen wir uns, dass  $A$  Hermitesch ist, da

$$A^* = (B^*B)^* = B^*(B^*)^* = B^*B = A.$$

Da  $B$  invertierbar ist, ist  $A$  positiv definit, denn

$$x^*Ax = x^*B^*Bx = (Bx)^*(Bx) \geq 0 \quad \text{und} \quad x^*Ax = 0 \iff Bx = 0 \iff x = 0.$$

Es sei jetzt  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  ein Körper, so dass wir Determinanten bilden können. Wir schließen „(2)  $\implies$  (4)“, denn für  $p, q \leq r$  gilt

$$a_{pq} = (Be_p)^*(Be_q) = \sum_{s=1}^p \sum_{t=1}^q \bar{b}_{sp} b_{sq} = \sum_{s,t=1}^r \bar{b}_{sp} b_{sq},$$

so dass  $A_r = B_r^*B_r$ , und da  $B_r$  wie oben invertierbar ist, folgt

$$\det A_r = \det B_r^* \det B_r = |\det B_r|^2 > 0.$$

Zu „(4)  $\implies$  (1)“ beweisen wir durch Induktion über  $r$ , dass  $A_r$  ein Skalarprodukt auf  $\mathbb{k}^r$  definiert. Für  $r = 1$  ist das klar, da  $a_{11} = \det A_1 > 0$ .

Es sei also  $r \geq 1$ , und  $A_r$  definiere ein Skalarprodukt auf  $\mathbb{k}^r$ . Wir konstruieren wie oben eine Orthonormalbasis  $(v_1, \dots, v_r)$  mit dem Gram-Schmidt-Verfahren, beginnend mit der Standardbasis. Wir definieren

$$w_{r+1} = e_{r+1} - \sum_{p=1}^r v_p \cdot (v_p^* A e_{r+1}),$$

so dass  $v_q^* A w_{r+1} = 0$  für  $q \leq r$ . Da  $A$  Hermitesch ist, gilt ebenfalls  $w_{r+1}^* A v_q = 0$  für alle  $q \leq r$ . Dann bilden  $(v_1, \dots, v_r, w_{r+1})$  eine Basis von  $\langle e_1, \dots, e_{r+1} \rangle$ . Es sei  $C_{r+1}$  die zugehörige Basiswechselmatrix, so dass

$$e_q = \sum_{p=1}^r v_p \cdot c_{pq} + w_{r+1} \cdot c_{p,r+1}.$$

Aus Bemerkung 6.15 (4) folgt

(\*)

$$A_{r+1} = C_{r+1}^* D_{r+1} C_{r+1}, \quad \text{wobei } D_{r+1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & w_{r+1}^* A w_{r+1} \end{pmatrix},$$

somit

$$0 < \det A_{r+1} = |\det C_{r+1}|^2 w_{r+1}^* A w_{r+1}.$$

Man überprüft jetzt leicht, dass  $D_{r+1}$  positiv definit ist, und damit auch  $A_{r+1}$ , denn

$$x^* A_{r+1} x = (C_{r+1} x)^* D_{r+1} (C_{r+1} x).$$

Also beschreibt  $A_{r+1}$  ebenfalls ein Skalarprodukt auf  $\mathbb{k}^{r+1}$ . Damit ist die Behauptung bewiesen.  $\square$

Wir haben oben nur gezeigt, dass die Cholesky-Zerlegung existiert. Sie ist auch eindeutig und lässt sich mit einem vergleichsweise einfachen Algorithmus berechnen, den wir hier aber nicht besprechen wollen.

6.20. BEMERKUNG. Wir geben noch eine geometrische Deutung der diversen Konstruktion in den Beweisen von Satz 6.18 und Folgerung 6.19.

- (1) Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt und  $U \subset V$  ein Unterraum. Wir wählen eine Orthonormalbasis  $(e_1, \dots, e_m)$  von  $U$  und definieren eine Abbildung

$$p: V \rightarrow U \quad \text{durch} \quad p(v) = \sum_{p=1}^m e_p \cdot g(e_p, v) \in U.$$

Für alle  $u \in U$  gilt  $p(u) = u$ , somit gilt  $p^2 = p$ . Eine Abbildung mit dieser Eigenschaft heißt *Projektion*.

Es gilt  $g(u, p(v)) = g(u, v)$  zunächst einmal für  $u = e_1, \dots, e_m$ , wie man leicht nachrechnet. Da  $(e_1, \dots, e_m)$  eine Basis von  $U$  bilden, gilt  $g(u, p(v)) = g(u, v)$  sogar für alle  $u$ , mit anderen Worten

$$g(u, v - p(v)) = 0 \quad \text{für alle } u \in U \text{ und alle } v \in V.$$

Aus diesem Grund nennt man  $p$  die *orthogonale* oder *senkrechte* Projektion von  $V$  auf den Unterraum  $U$ . Man kann zeigen, dass  $p(v)$  derjenige Punkt in  $U$  ist, für den der Abstand  $\|v - p(v)\|_g$  minimal wird (Übung).

- (2) Es sei jetzt  $\mathbb{k} = \mathbb{R}$ , und es sei  $A$  die Gramsche Matrix eines Skalarproduktes  $g$  bezüglich einer Basis  $(v_1, \dots, v_n)$  von  $V$ . Wir wollen durch Induktion über  $r$  motivieren, dass  $\det A_r$  das Quadrat des Volumens des von den Vektoren  $v_1, \dots, v_r$  aufgespannten  $r$ -dimensionalen Paralleleotops  $P_r$  ist. Dabei erinnern wir uns an den Anfang von Abschnitt 4.1, wo wir entsprechende Überlegungen für Paralleleotope maximaler Dimension  $r = n$  in  $\mathbb{R}^n$  angestellt haben.

In Dimension  $r = 1$  sollte das „Volumen“ des Vektors  $v_1$  seine Länge sein. In der Tat gilt

$$\det A_1 = a_{11} = g(v_1, v_1) = \|v_1\|_g^2.$$

Sei jetzt  $r \geq 1$ . Dann hat das von  $v_1, \dots, v_{r+1}$  aufgespannte Parallelotop  $P_{r+1}$  als „Grundfläche“ das Parallelotop  $P_r$  vom Volumen  $\text{vol}(P_r) = \sqrt{\det A_r}$  nach Induktionsvoraussetzung, und als Höhe den Vektor  $w_{r+1} = v_{r+1} - p(v_{r+1})$ , dabei ist  $p: V \rightarrow \langle v_1, \dots, v_r \rangle$  die orthogonale Projektion aus (1).

Wie im Beweis der Folgerung 6.19, Schritt (4)  $\Rightarrow$  (1), sei  $e_1, \dots, e_r$  eine Orthonormalbasis von  $\langle v_1, \dots, v_r \rangle$ , und  $C_{r+1}$  sei die dortige Basiswechsellmatrix. Dann hat  $C_{r+1}$  die Blockgestalt

$$C_{r+1} = \begin{pmatrix} B_r & * \\ 0 & 1 \end{pmatrix}.$$

Nach Induktionsvoraussetzung und (\*) hat also die Grundfläche das Volumen

$$\text{vol}(P_r) = \sqrt{\det A_r} = |\det B_r| = |\det C_{r+1}| .$$

Die Länge der Höhe ist  $\|w_{r+1}\|_g = \sqrt{w_{r+1}^* A_r w_{r+1}}$ , und somit erhalten wir mit (\*), dass

$$\text{vol}(P_{r+1}) = |\det C_{r+1}| \cdot \sqrt{w_{r+1}^* A_r w_{r+1}} = \sqrt{\det A_{r+1}} .$$

Damit ist unsere Behauptung gezeigt, allerdings unter der Annahme, dass man das Volumen mit Hilfe der Formel „Grundfläche  $\times$  Höhe“ berechnen darf.

Einen alternativen Zugang zur Behauptung  $\text{vol } P_r = \sqrt{\det A_r}$  finden Sie in den Übungen.

- (3) Im  $\mathbb{R}^n$  mit dem Standard-Skalarprodukt hat das von den Vektoren  $v_1, \dots, v_r$  aufgespannte Parallelotop also das Volumen

$$\sqrt{\det((\langle v_p, v_q \rangle)_{p,q \leq r})} = \sqrt{\det((v_p^* v_q)_{p,q \leq r})} .$$

Beispielsweise hatten wir in Bemerkung 1.70 eine geometrische Interpretation des Kreuz- und des Spatproduktes gegeben. Dazu hatten wir die Fläche des von  $u, v \in \mathbb{R}^3$  aufgespannten Parallelogramms berechnet als

$$\|u \times v\| = \sqrt{\|u\|^2 \|v\|^2 - \langle u, v \rangle^2} = \det \begin{pmatrix} \langle u, u \rangle & \langle u, v \rangle \\ \langle v, u \rangle & \langle v, v \rangle \end{pmatrix}^{\frac{1}{2}} .$$

### 6.3. Dualräume und adjungierte Abbildungen

Wir erinnern uns an die Definition 2.32 des Dualraumes  $V^* = \text{Hom}_{\mathbb{k}}(V, \mathbb{k})$  eines Rechts- $\mathbb{k}$ -Vektorraums  $V$ . Der Dualraum eines Rechtsvektorraums ist ein Linksvektorraum und umgekehrt (wobei wir anstelle von  ${}^*V$  oft einfach wieder  $V^*$  schreiben). Elemente des Dualraumes heißen auch *Linearformen*  $\alpha \in V^*$ . Linearformen auf  $V$  sind also lineare Abbildungen  $\alpha: V \rightarrow \mathbb{k}$ .

6.21. BEISPIEL. Wir kennen Beispiele von Linearformen.

- (1) Für  $p = 1, \dots, n$  ist die Abbildung  $\varepsilon_p: \mathbb{k}^n \rightarrow \mathbb{k}$ , die  $x \in \mathbb{k}^n$  die  $p$ -te Koordinate zuordnet, eine Linearform. Für die Standardbasisvektoren  $e_1, \dots, e_n$  gilt

$$\varepsilon_p(e_q) = \delta_{pq} ,$$

und man nennt  $(\varepsilon_1, \dots, \varepsilon_n)$  die zu  $(e_1, \dots, e_n)$  duale Basis von  ${}^n\mathbb{k}$ , siehe Bemerkung 2.65.

Allgemeiner sei  $V$  ein Rechts- $\mathbb{k}$ -Vektorraum und  $B = (b_1, \dots, b_n)$  eine Basis. Die dazu duale Basis  $(\beta_1, \dots, \beta_n)$  mit  $\beta_p = \varepsilon_p \circ B^{-1}: V \rightarrow \mathbb{k}$  haben wir in Proposition 2.66 aus der Koordinatenabbildung konstruiert, so dass wieder

$$\beta_p(b_q) = \delta_{pq} .$$

- (2) Wir betrachten den Raum  $C^\infty([0, 1], \mathbb{k})$  der beliebig oft differenzierbaren,  $\mathbb{k}$ -wertigen Funktionen auf dem Intervall  $[0, 1]$ . Typische Linearformen auf  $C^\infty([0, 1], \mathbb{k})$  sind zum Beispiel

$$f \mapsto f(x_0), \quad f \mapsto f'(x_0), \dots$$

für  $x_0 \in [0, 1]$ , sowie Linearkombinationen solcher Linearformen.

- (3) Auch die Abbildung

$$f \mapsto \int_0^1 f(x) dx$$

ist linear. Allgemeiner betrachten wir das  $L^2$ -Skalarprodukt aus Beispiel 6.7 (1). Die obige Linearform entspricht der Abbildung

$$\langle 1, \cdot \rangle_{L^2}: C^\infty([0, 1], \mathbb{k}) \longrightarrow \mathbb{k} \quad \text{mit} \quad f \mapsto \langle 1, f \rangle_{L^2} = \int_0^1 \bar{1} \cdot f(x) dx,$$

wobei 1 die konstante Abbildung  $x \mapsto 1$  bezeichne. Wenn wir das  $L^2$ -Skalarprodukt mit der gleichen Definition auf beschränkte und stückweise stetige Funktionen erweitern, existiert für ein beschränktes und stückweise stetiges  $g: [0, 1] \rightarrow \mathbb{k}$  die Abbildung

$$f \mapsto \langle f, g \rangle_{L^2} = \int_0^1 \overline{g(x)} \cdot f(x) dx.$$

Einschränken auf  $f \in C^\infty([0, 1], \mathbb{k})$  liefert wieder eine Linearform auf  $C^\infty([0, 1], \mathbb{k})$ .

- (4) Wenn  $g$  stetig und differenzierbar ist, dann erhalten wir entsprechend mit der Erweiterung des Sobolev-Skalarproduktes aus Beispiel 6.7 (3) auf  $C^1$ -Funktionen eine Linearform

$$g \mapsto \langle f, g \rangle_{H^1} = \langle f, g \rangle_{L^2} + \langle f', g' \rangle_{L^2}.$$

- (5) Es sei  $U \subset \mathbb{R}^n$  offen und  $f: U \rightarrow \mathbb{R}$  eine  $C^1$ -Funktion (total differenzierbar würde auch ausreichen). Dann ist die *totale Ableitung* an der Stelle  $x_0 \in U$  die lineare Abbildung  $df(x_0): \mathbb{R}^n \rightarrow \mathbb{R}$ , die jedem Vektor  $v \in \mathbb{R}^n$  die *Richtungsableitung*

$$df(x_0)(v) = \lim_{t \rightarrow 0} \frac{f(x_0 + v \cdot t) - f(x_0)}{t}$$

zuordnet. Die Koordinaten von  $df(x_0)$  bezüglich der dualen Basis  $\varepsilon_1, \dots, \varepsilon_n$  von  ${}^n\mathbb{R} = (\mathbb{R}^n)^*$  heißen auch die *partiellen Ableitungen* von  $f$ .

Wir erinnern uns auch an anti- (oder semi-) lineare Abbildungen, siehe Definition 6.2.

6.22. PROPOSITION. *Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt. Dann induziert  $g$  eine injektive antilineare Abbildung  $g: V \rightarrow V^*$  durch*

$$v \mapsto g(v) \in V^* \quad \text{mit} \quad g(v)(w) = g(v, w) \quad \text{für alle } v, w \in V.$$

Eine Linearform  $\alpha \in V^*$  heißt *darstellbar* bezüglich  $g$ , wenn es ein  $v \in V$  mit  $\alpha = g(v)$  gibt. Man sagt auch, dass  $v \in V$  die Linearform  $\alpha$  *darstellt*.

BEWEIS. Es sei  $v \in V$ , dann ist nach (S1) in Definition 6.3 die Abbildung

$$g(v) = g(v, \cdot): V \rightarrow \mathbb{k}$$

linear, also gilt  $g(v) \in V^*$ . Nach Definition 2.32 ist  $V^*$  ein Links- $\mathbb{k}$ -Vektorraum. Aus (S1) folgt, dass  $g: V \rightarrow V^*$  antilinear ist, denn für  $u, v, w \in V$  und  $r, s \in \mathbb{k}$  gilt

$$g(u \cdot r + v \cdot s)(w) = g(u \cdot r + v \cdot s, w) = \bar{r} g(u, w) + \bar{s} g(v, w) = (\bar{r} g(u) + \bar{s} g(v))(w).$$

Zur Injektivität nehmen wir an, dass  $g(u) = g(v)$ , das heißt, es gilt  $g(u)(w) = g(v)(w)$  für alle  $w \in V$ . Dann folgt

$$0 = g(u)(u - v) - g(v)(u - v) = \|u - v\|_g^2,$$

also gilt  $u = v$  wegen (S3), und  $g: V \rightarrow V^*$  ist injektiv.  $\square$

6.23. BEISPIEL. Wir wollen wissen, ob die Linearformen aus dem obigen Beispiel darstellbar sind.

- (1) Es sei  $(e_1, \dots, e_n)$  eine Orthonormalbasis, dann wird  $\varepsilon_p$  im Beispiel 6.21 (1) durch den Vektor  $e_p$  dargestellt, siehe Proposition 3.34 und Bemerkung 6.17 (3).
- (2) Im Beispiel 6.21 (3) wird die Linearform

$$f \mapsto \int_0^1 \overline{g(x)} \cdot f(x) dx$$

auf den stückweise stetigen Funktionen durch  $g$  dargestellt. Auf dem Unterraum  $C^\infty([0, 1], \mathbb{k})$  wird sie nur genau dann durch  $g$  dargestellt, wenn  $g \in C^\infty([0, 1], \mathbb{k})$ . Also gibt es viele Linearformen auf  $C^\infty([0, 1], \mathbb{k})$ , die bezüglich des  $L^2$ -Skalarproduktes nicht (das heißt, nicht durch  $C^\infty$ -Funktionen) darstellbar sind. Auch die Linearformen aus 6.21 (2) sind nicht durch  $C^\infty$ -Funktionen darstellbar.

- (3) Wie oben ist die Linearform  $\langle g, \cdot \rangle_{H^1}$  aus 6.21 (4) auf  $C^1$  bezüglich des ersten Sobolev-Skalarproduktes durch  $g$  darstellbar. Auf  $C^\infty([0, 1], \mathbb{k})$  ist sie genau dann darstellbar, wenn  $g \in C^\infty([0, 1], \mathbb{k})$ . Wenn darüberhinaus  $g'(0) = g'(1) = 0$  gilt, ist  $\langle g, \cdot \rangle_{H^1}$  sogar bezüglich des  $L^2$ -Skalarproduktes darstellbar, denn partielle Integration liefert

$$\begin{aligned} \langle g, f \rangle_{H^1} &= \int_0^1 \overline{g(x)} f(x) dx + \int_0^1 \overline{g'(x)} f'(x) dx \\ &= \int_0^1 \overline{g(x)} f(x) dx + \overline{g'(x)} f(x) \Big|_{x=0}^1 - \int_0^1 \overline{g''(x)} f(x) dx \\ &= \langle g - g'', f \rangle_{L^2}. \end{aligned}$$

- (4) Die totale Ableitung  $df(x_0) \in {}^n\mathbb{R} = (\mathbb{R}^n)^*$  aus Beispiel 6.21 (5) wird bezüglich des Standardskalarproduktes auf  $\mathbb{R}^n$  dargestellt durch den *Gradienten*

$$\text{grad } f(x_0) = (df(x_0))^* = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix}.$$

Um zu sehen, dass alle Linearformen auf einem endlich-dimensionalen Vektorraum mit Skalarprodukt darstellbar sind, führen wir als Hilfsmittel noch einen weiteren Vektorraum ein.

6.24. DEFINITION. Es sei  $V$  ein Rechts- $\mathbb{k}$ -Vektorraum, Dann ist eine *Antilinearform* eine antilineare Abbildung  $\gamma: V \rightarrow \mathbb{k}$ . Wir definieren wir den *Antidualraum* von  $V$  als

$$\overline{V}^* = \{ \gamma: V \rightarrow \mathbb{k} \mid \gamma \text{ ist } \mathbb{k}\text{-antilinear} \}.$$

Analog definieren wir den Antidualraum eines Links- $\mathbb{k}$ -Vektorraums.

6.25. BEMERKUNG. Wir sammeln einige einfache Eigenschaften.

- (1) Der Antidualraum  $\overline{V}^*$  eines Rechts- $\mathbb{k}$ -Vektorraums  $V$  ist wieder ein Rechts- $\mathbb{k}$ -Vektorraum. Sei  $\gamma: V \rightarrow \mathbb{k}$  antilinear, und seien  $v \in V$  und  $r, s \in \mathbb{k}$ , dann definieren wir

$$(\gamma \cdot r)(v) = \gamma(v) \cdot r \in \mathbb{k}.$$

mit Definition 6.2 erhalten wir

$$(\gamma \cdot r)(v \cdot s) = \gamma(v \cdot s) \cdot r = \bar{s} \cdot \gamma(v) \cdot r = \bar{s} \cdot (\gamma \cdot r)(v).$$

- (2) Wir können aus jeder Linearform  $\alpha \in V^*$  eine Antilinearform  $\gamma = \bar{\alpha} \in \overline{V}^*$  machen und umgekehrt, wobei

$$\bar{\alpha}(v) = \overline{\alpha(v)} \in \mathbb{k}.$$

Das liefert eine antilineare Abbildung  $V^* \rightarrow \overline{V}^*$  mit einer antilinearen Umkehrabbildung. Ähnlich wie in Bemerkung 6.4 geht dabei die Links- $\mathbb{k}$ -Vektorraumstruktur von  $V^*$  in die Rechts- $\mathbb{k}$ -Vektorraumstruktur von  $\overline{V}^*$  über und umgekehrt.

- (3) Wir definieren eine Abbildung  $\bar{g}: V \rightarrow \overline{V}^*$  durch

$$\bar{g}(v)(w) = \overline{g(v)}(w) = \overline{g(v, w)} = g(w, v)$$

für alle  $v, w \in V$ . Wie in Proposition 6.22 folgt, dass  $\bar{g}$  eine injektive lineare Abbildung ist. Antilinearformen im Bild von  $\bar{g}$  heißen wieder *darstellbar*.

Die Frage, welche Linearformen sich durch Elemente spezieller Funktionenräume darstellen lassen, ist ein wichtiges Thema in der Funktionalanalysis. Das folgende Lemma ist ein elementarer Spezialfall des Riesz'schen Darstellungssatzes.

6.26. LEMMA. *Es sei  $(V, g)$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum, dann sind die Abbildungen  $g: V \rightarrow V^*$  und  $\bar{g}: V \rightarrow \overline{V}^*$  bijektiv, und wir erhalten Umkehrabbildungen  $g^{-1}: V^* \rightarrow V$  und  $\bar{g}^{-1}: \overline{V}^* \rightarrow V$ .*

Insbesondere ist jede Linearform und jede Antilinearform auf einem endlich-dimensionalen Vektorraum  $V$  bezüglich  $g$  darstellbar. Dieses Lemma erklärt also insbesondere Beispiel 6.23 (1).

Der Dualraum eines unendlich-dimensionalen  $\mathbb{k}$ -Vektorraums ist (unter geeigneten mengentheoretischen Annahmen) stets mächtiger als der Vektorraum selbst, so dass  $g$  für unendlich-dimensionale Vektorräume nie invertierbar ist. Aus diesem Grund betrachtet man in der Funktionalanalysis stattdessen den Raum der stetigen Linearformen bezüglich der zum Skalarprodukt gehörigen Norm. Dadurch wird  $g$  auch für zahlreiche wichtige unendlich-dimensionale Vektorräume mit Skalarprodukt invertierbar.

BEWEIS. Es sei  $\dim V = n$ . Nach Proposition 2.66 ist  $V^*$  ein  $n$ -dimensionaler Links  $\mathbb{k}$ -Vektorraum. Nach Bemerkung 6.25 (2) ist  $\overline{V}^*$  ein  $n$ -dimensionaler Rechts- $\mathbb{k}$ -Vektorraum. Die Abbildung  $\bar{g}: V \rightarrow \overline{V}^*$  ist nach Proposition 6.22 und Bemerkung 6.25 (3) injektiv. Aus dem Rangsatz 3.13 folgt, dass  $\bar{g}$  ein Isomorphismus ist. Aber dann ist auch  $g$  bijektiv.  $\square$

6.27. BEMERKUNG. Es sei  $(e_1, \dots, e_n)$  eine Orthonormalbasis von  $V$ , siehe Satz 6.18. Wir können sie benutzen, um einen alternativen, konstruktiven Beweis von Lemma 6.26 zu geben. Nach Beispiel 6.23 (1) werden die Vektoren der dualen Basis  $(\varepsilon_1, \dots, \varepsilon_n)$  durch die Vektoren  $e_1, \dots, e_n$  dargestellt. Sei also

$$\alpha = \sum_{p=1}^n a_p \cdot \varepsilon_p \quad \text{mit} \quad a \in {}^n\mathbb{k},$$

dann wird  $\alpha$  dargestellt durch den Vektor

$$v = \sum_{p=1}^n e_p \cdot \bar{a}_p,$$

denn für alle  $w \in V$  gilt

$$\alpha(w) = \sum_{p=1}^n a_p \cdot \varepsilon_p(w) = \sum_{p=1}^n a_p \cdot g(e_p, w) = g\left(\sum_{p=1}^n e_p \cdot \bar{a}_p, w\right).$$

Also gilt  $\alpha = g(v)$  und analog  $\bar{\alpha} = \bar{g}(v)$ .

6.28. DEFINITION. Es seien  $(V, g)$  und  $(W, h)$  Vektorräume über  $\mathbb{k}$  mit Skalarprodukt. Eine lineare Abbildung  $F: V \rightarrow W$  heißt *adjungierbar*, wenn es eine Abbildung  $G: W \rightarrow V$  gibt, so dass

$$g(G(w), v) = h(w, F(v)) \quad \text{für alle } v \in V \text{ und } w \in W.$$

Dann nennen wir  $G$  die zu  $F$  *adjungierte Abbildung*, und schreiben  $G = F^*$ .

6.29. BEMERKUNG. Es seien  $(V, g)$ ,  $(W, h)$  und  $F: V \rightarrow W$  wie oben.

- (1) Falls  $F$  adjungierbar ist, ist die adjungierte Abbildung  $G$  von  $F$  eindeutig bestimmt. Denn sei  $H$  eine weitere adjungierte Abbildung von  $F$ , dann gilt für alle  $w \in W$ , dass

$$\begin{aligned} 0 &= h(w, F(G(w) - H(w))) - h(w, F(G(w) - H(w))) \\ &= g(G(w), G(w) - H(w)) - g(H(w), G(w) - H(w)) \\ &= \|G(w) - H(w)\|_g^2, \end{aligned}$$

und somit  $G(w) = H(w)$  wegen der Eigenschaft (S3) des Skalarproduktes  $g$ . Daher dürfen wir  $F^*$  für die Adjungierte Abbildung schreiben, wenn Sie existiert.

- (2) Die adjungierte Abbildung  $F^*: W \rightarrow V$  ist linear, denn für alle  $v \in V$  und alle  $u, w \in W$  und alle  $r, s \in \mathbb{k}$  gilt

$$\begin{aligned} g(F^*(u.r + w.s), v) &= h(u.r + w.s, F(v)) = \bar{r} h(u, F(v)) + \bar{s} h(w, F(v)) \\ &= \bar{r} g(F^*(u), v) + \bar{s} g(F^*(w), v) = g(F^*(u) \cdot r + F^*(w) \cdot s, v). \end{aligned}$$

Indem wir  $v = F^*(u \cdot r + w \cdot s) - F^*(u) \cdot r - F^*(w) \cdot s$  wählen, erhalten wir

$$0 = \|F^*(u \cdot r + w \cdot s) - F^*(u) \cdot r - F^*(w) \cdot s\|_g^2,$$

und wegen (S3) gilt somit  $F^*(u \cdot r + w \cdot s) = F^*(u) \cdot r + F^*(w) \cdot s$ .

- (3) Wenn  $G$  zu  $F$  adjungiert ist, ist auch  $F$  zu  $G$  adjungiert, denn wegen (S2) gilt

$$h(F(v), w) = \overline{h(w, F(v))} = \overline{g(G(w), v)} = g(v, G(w))$$

für alle  $v \in V$  und  $w \in W$ . Wegen (1) gilt also  $F = G^*$  genau dann, wenn  $G = F^*$ , insbesondere folgt  $(F^*)^* = F$ .

### 6.30. BEISPIEL. Wir geben Beispiele adjungierter Abbildungen.

- (1) Wir betrachten das Standardskalarprodukt auf den Räumen  $\mathbb{k}^m$  und  $\mathbb{k}^n$ . Sei  $F: \mathbb{k}^n \rightarrow \mathbb{k}^m$  gegeben durch eine Matrix  $C \in M_{m,n}(\mathbb{k})$ , dann wird die adjungierte Abbildung  $F^*$  gegeben durch die adjungierte Matrix, denn für alle  $x \in \mathbb{k}^m$  und alle  $y \in \mathbb{k}^n$  gilt

$$\langle F^*(x), y \rangle = \langle x, F(y) \rangle = x^* A y = x^* (A^*)^* y = (A^*)^* y = \langle A^* x, y \rangle.$$

Aus diesem Grund benutzen wir in beiden Fällen den Begriff „adjungiert“.

- (2) Etwas allgemeiner sei  $(e_1, \dots, e_n)$  eine Orthonormalbasis von  $(V, g)$  und  $(f_1, \dots, f_m)$  eine Orthonormalbasis von  $(W, h)$ . Wenn  $F: V \rightarrow W$  bezüglich dieser Basen durch eine Matrix  $A \in M_{m,n}(\mathbb{k})$  dargestellt wird, dann wird  $F^*$  durch  $A^*$  dargestellt, denn für alle  $p = 1, \dots, m$  und alle  $q = 1, \dots, n$  gilt

$$\langle e_p, F^*(f_p) \rangle = \langle F(e_p), f_q \rangle = \overline{\langle f_q, F(e_p) \rangle} = \bar{a}_{pq}.$$

- (3) Wir betrachten wieder den Raum  $V = C^\infty([0, 1]; \mathbb{k})$  mit dem  $L^2$ -Skalarprodukt. Multiplikation mit einer Funktion  $f \in V$  definiert eine lineare Abbildung Die adjungierte Abbildung ist Multiplikation mit  $\bar{f}$ , denn

$$\langle g, fh \rangle_{L^2} = \int_0^1 \overline{g(x)} f(x) h(x) dx = \int_0^1 \overline{\overline{f(x)} g(x)} h(x) dx = \langle \bar{f}g, h \rangle_{L^2}.$$

- (4) Es sei  $V$  wie oben, und es sei  $f \in V$  eine Funktion mit  $f(0) = f(1) = 0$ . Dann betrachten wir den Differentialoperator  $F \in \text{End}(V)$  mit

$$F(g) = f \cdot g'$$

und bestimmen den adjungierten Differentialoperator durch partielle Integration als

$$\begin{aligned} \langle F^*(g), h \rangle_{L^2} &= \langle g, F(h) \rangle = \int_0^1 \overline{g(x)} f(x) h'(x) dx \\ &= \left( \overline{g(x)} f(x) h(x) \right) \Big|_{x=0}^1 - \int_0^1 (\overline{g} f)'(x) h(x) dx \\ &= \langle (\overline{f} g)', h \rangle_{L^2} = \langle \overline{f} g' + \overline{f}' g, h \rangle_{L^2}. \end{aligned}$$

- (5) Wenn wir in (4) einfach nur den Ableitungsoperator  $F(g) = g'$  betrachten, zeigt eine analoge Rechnung, dass  $F$  nicht adjungierbar ist, da sich die Randterme  $(\overline{g(x)} h(x)) \Big|_{x=0}^1$  nicht durch ein Integral beschreiben lassen. In der Analysis umgeht man dieses Problem, indem man den Begriff des adjungierten Operators etwas anders definiert und dann Randbedingungen stellt wie etwa  $g(0) = g(1) = 0$ , um keine Randterme mehr zu erhalten.

Die adjungierte Abbildung ist eng verwandt mit dem folgenden Konzept.

6.31. DEFINITION. Es seien  $V$  und  $W$  Vektorräume über einem Körper  $\mathbb{k}$ , und es sei  $F: V \rightarrow W$  eine lineare Abbildung. Die zu  $F$  *duale Abbildung*  $F^*: W^* \rightarrow V^*$  ist definiert durch

$$F^* \beta = \beta \circ F \in V^* \quad \text{für alle } \beta \in W^*.$$

Man beachte, dass die duale Abbildung im Gegensatz zur adjungierten Abbildung immer existiert und nach Definition eindeutig bestimmt ist. Wir verwenden für beide die Bezeichnung  $F^*$ , man muss also aufpassen, welche der beiden Abbildungen jeweils gemeint ist:  $F^*: W^* \rightarrow V^*$  ist die duale Abbildung,  $F^*: W \rightarrow V$  die adjungierte. Aus diesem Grund verwenden manche Autoren für duale Moduln, Vektorräume und Abbildungen das Symbol  $\cdot'$  oder  $\cdot^\vee$ .

6.32. BEMERKUNG. Wir sammeln ein paar elementare Eigenschaften. Seien dazu  $U, V$  und  $W$  Vektorräume.

- (1) Es gilt stets  $\text{id}_V^* = \text{id}_{V^*}$ , denn  $\alpha \circ \text{id}_V = \alpha \in V^*$  für alle  $\alpha \in V^*$ .
- (2) Seien  $F: V \rightarrow W$  und  $G: U \rightarrow V$  linear, dann gilt  $(F \circ G)^* = G^* \circ F^*$ , denn

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

- (3) Die duale Abbildung ist linear. Das lässt sich nachrechnen, da  $\mathbb{k}$  auf  $\beta \in W^*$  durch  $(r \cdot \beta)(w) = r \cdot \beta(w)$  wirkt.
- (4) Es seien  $B = (v_1, \dots, v_n)$  und  $C = (w_1, \dots, w_m)$  Basen von  $V$  beziehungsweise  $W$ , und es seien  $B^* = (\varphi_1, \dots, \varphi_n)$  und  $C^* = (\psi_1, \dots, \psi_m)$  die dualen Basen von  $V^*$  und  $W^*$ . Es sei  $F: V \rightarrow W$  bezüglich der obigen Basen dargestellt durch die Abbildungsmatrix  $A = M_{m,n}(\mathbb{k})$ , dann gilt

$$\psi_p(F(v_q)) = \psi_p\left(\sum_{r=1}^m w_r \cdot a_{rq}\right) = \sum_{r=1}^m \psi_p(w_r) \cdot a_{rq} = a_{pq}$$

für alle  $p = 1, \dots, m$  und alle  $q = 1, \dots, n$ . Dabei geht der Vektor  $v = B(x)$  in den Vektor  $C(Ax)$  über, wobei  $x \in \mathbb{k}^n$ .

Es sei jetzt  $\eta \in {}^m\mathbb{k}$  eine Zeile. Für die duale Matrix rechnen wir

$$\begin{aligned} F^*(C^*(\eta))(v_r) &= F^*\left(\sum_{p=1}^m \eta_p \cdot \psi_p\right)(v_r) = \left(\sum_{p=1}^m \eta_p \cdot (\psi_p \circ F)\right)(v_r) \\ &= \sum_{p=1}^m \eta_p \cdot a_{pr} = \left(\sum_{p=1}^m \sum_{q=1}^n \eta_p \cdot a_{pq} \cdot \varphi_q\right)(v_r) = (B^*(\eta A))(v_r). \end{aligned}$$

Die duale Abbildung wird durch also dieselbe Matrix  $A$  dargestellt, allerdings werden jetzt Zeilen in  $\mathbb{k}^m$  von rechts mit  $A$  multipliziert.

Mit Hilfe von Lemma 6.26 können wir einen Zusammenhang zwischen der adjungierten Abbildung und der dualen Abbildung herstellen.

**6.33. PROPOSITION.** *Es seien  $(V, g)$  und  $(W, h)$  Vektorräume über  $\mathbb{k}$  mit Skalarprodukt und  $F: V \rightarrow W$  sei linear. Wenn  $F$  adjungierbar ist, kommutiert das Diagramm*

$$\begin{array}{ccc} W & \xrightarrow{F^*} & V \\ h \downarrow & & \downarrow g \\ W^* & \xrightarrow{F^*} & V^* \end{array}.$$

*Insbesondere ist  $F$  immer adjungierbar, wenn  $V$  endlich-dimensional ist.*

Man beachte, dass die beiden waagerechten Pfeile lineare Abbildungen sind, während die senkrechten Pfeile antilinear sind. Somit sind beide Wege von  $W$  nach  $V^*$  durch antilineare Abbildungen gegeben.

Die letzte Behauptung erklärt insbesondere die Beispiele 6.30 (1) und (2). Die Beispiele 6.30 (3) und (4) zeigen, dass die zusätzliche Bedingung  $\dim V < \infty$  nicht notwendig ist.

**BEWEIS.** Es seien  $v \in V$  und  $w \in W$ , dann folgt

$$g(F^*(w))(v) = g(F^*(w), v) = h(w, F(v)) = h(w)(F(v)) = F^*(h(w))(v).$$

Da das für alle  $v \in V$  gilt, folgt  $g \circ F^* = F^* \circ h$ , wobei links die adjungierte und rechts die duale Abbildung gemeint ist. Damit ist die erste Behauptung bewiesen.

Wenn  $g$  invertierbar ist, können wir die adjungierte Abbildung als  $g^{-1} \circ F^* \circ h$  schreiben. Nach Lemma 6.26 gilt das, wenn  $V$  endlich-dimensional ist.  $\square$

**6.34. DEFINITION.** Es seien  $V$  und  $W$  Vektorräume über  $\mathbb{k}$ , es sei  $F: V \rightarrow W$  linear, und  $S$  sei eine Sesquilinearform auf  $W$ . Dann definiert man die mit  $F$  zurückgeholte Sesquilinearform  $F^*S$  auf  $V$  durch

$$(F^*S)(u, v) = S(F(v), F(w)) \quad \text{für alle } u, v \in V.$$

Wir haben jetzt die Notation  $F^*$  mit einer weiteren Bedeutung versehen. Aus dem Zusammenhang muss man jeweils erkennen, wofür  $F^*$  gerade steht.

6.35. BEMERKUNG. Die ersten drei der folgenden Eigenschaften rechnet man leicht nach.

- (1) Die Form  $F^*S$  ist wieder sesquilinear (S1).
- (2) Wenn  $S$  Hermitesch ist, dann ist auch  $F^*S$  Hermitesch (S2).
- (3) Wenn  $S$  außerdem positiv semidefinit ist, dann ist auch  $F^*$  positiv semidefinit.
- (4) Wenn  $S$  positiv definit (S3) und  $F$  injektiv ist, dann ist auch  $F^*S$  ein Skalarprodukt, denn dann gilt

$$(F^*S)(v, v) = S(F(v), F(v)) = 0 \iff F(v) = 0 \iff v = 0.$$

In diesem Fall heißt  $F^*S$  auch das zurückgeholte Skalarprodukt. Auf die Injektivität von  $F$  kann man leider nicht verzichten, denn für alle  $v \in \ker F$  gilt  $(F^*S)(v) = 0$ .

- (5) Es seien  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  Basen von  $V$  und  $W$ . Sei  $A \in M_{M,n}(\mathbb{k})$  die Abbildungsmatrix von  $F$ , und sei  $S$  dargestellt durch die Gramsche Matrix  $G$ , dann wird  $F^*S$  dargestellt durch die Matrix  $A^*GA$ , denn

$$(F^*S)(v_p, v_q) = S\left(\sum_{r=1}^m w_r \cdot a_{rp}, \sum_{s=1}^m w_s \cdot a_{sq}\right) = \sum_{r,s=1}^m \bar{a}_{rp} g_{rs} a_{sq}.$$

6.36. BEMERKUNG. Genauso können wir eine Sesquilinearform  $S$  auf  $W$  mit einer antilinearen Abbildung  $F: V \rightarrow W$  zurückholen durch

$$(F^*S)(u, v) = S(F(v), F(u)) \quad \text{für alle } u, v \in V.$$

Durch das Vertauschen der Argumente stellen wir sicher, dass  $F^*S$  wieder sesquilinear ist. Die Punkte (2)–(4) aus Bemerkung 6.35 gelten analog.

Zum Beispiel sei  $(V, g)$  ein endlich-dimensionaler Vektorraum mit Skalarprodukt, dann ist die antilineare Abbildung  $g: V \rightarrow V^*$  invertierbar, und wir können das Skalarprodukt  $g$  mit der Inversen Abbildung  $g^{-1}$  auf  $V^*$  zurückholen. Dieses Skalarprodukt nennen wir das zu  $g$  *duale Skalarprodukt*  $g^*$  auf  $V^*$ . Sei dazu  $(e_1, \dots, e_n)$  eine Orthonormalbasis von  $V$ , dann ist  $(\varepsilon_1, \dots, \varepsilon_n) = (g(e_1), \dots, g(e_n))$  die duale Basis von  $V^*$  nach Beispiel 6.23 (1). Somit gilt

$$((g^{-1})^*g)(\varepsilon_p, \varepsilon_q) = g(g^{-1}(\varepsilon_q), g^{-1}(\varepsilon_p)) = g(e_q, e_p) = \delta_{pq},$$

also ist die duale Basis einer Orthonormalbasis wieder eine Orthonormalbasis. Im Allgemeinen sei  $A$  die Gramsche Matrix von  $g$ , dann kann man zeigen, dass  $g^*$  durch die Inverse Matrix  $A^{-1}$  dargestellt wird.

6.37. BEMERKUNG. Es seien  $(V, g)$  und  $(W, h)$  Vektorräume über  $\mathbb{k}$  mit Skalarprodukt. Wir nennen eine lineare Abbildung  $F: V \rightarrow W$  eine *isometrische Einbettung*, wenn  $F^*h = g$  gilt. In diesem Fall gilt also für alle  $u, v \in V$ , dass

$$g(u, v) = (F^*h)(u, v) = h(F(u), F(v)) \quad \text{und} \quad \|v\|_g = \|F(v)\|_h.$$

Insbesondere ist  $F$  immer injektiv.

Seien  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  Orthonormalbasen von  $V$  und  $W$ , und sei  $A \in M_{m,n}(\mathbb{k})$  die Abbildungsmatrix von  $F$ . Wegen Bemerkung 6.35 (5) gilt dann

$$E_n = A^* E_m A = A^* A.$$

Falls  $n = m$  ist, ist  $A$  insbesondere invertierbar, und es gilt  $A^{-1} = A^*$ . In diesem Fall nennen wir  $F$  eine *lineare Isometrie*.

#### 6.4. Normale Endomorphismen

In diesem Kapitel betrachten wir bestimmte Endomorphismen von endlich-dimensionalen  $\mathbb{k}$ -Vektorräumen mit Skalarprodukt und zeigen, dass sie sich bezüglich einer geeigneten Orthonormalbasis durch besonders einfache Matrizen darstellen lassen. Diese Resultate haben zahlreiche Anwendungen, unter anderem in Analysis, Geometrie und Physik.

6.38. DEFINITION. Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt und es sei  $F \in \text{End}_{\mathbb{k}}(V)$  adjungierbar. Dann heißt  $F$

- (1) *selbstadjungiert*, wenn  $F^* = F$ ,
- (2) *schief*, wenn  $F^* = -F$ ,
- (3) *normal*, wenn  $F^* \circ F = F \circ F^*$ , und
- (4) *lineare Isometrie*, oder *isometrischer* oder auch *unitärer Automorphismus*, wenn  $F$  invertierbar ist mit  $F^{-1} = F^*$ .

6.39. BEMERKUNG. Man sieht leicht, dass selbstadjungierte und schiefe Endomorphismen und isometrische Automorphismen allesamt normal sind. Stellt man  $F$  wie oben bezüglich einer Orthonormalbasis als Matrix  $A \in M_n(\mathbb{k})$  dar, so gilt jeweils  $A^* = A$ ,  $A^* = -A$ ,  $A^* A = A A^*$ , beziehungsweise  $A^{-1} = A^*$ .

6.40. SATZ (Hauptsatz über normale Abbildungen). *Es sei  $(V, g)$  ein endlich-dimensionaler komplexer Vektorraum mit Skalarprodukt und  $F \in \text{End}_{\mathbb{C}} V$ . Dann existiert genau dann eine unitäre Basis von  $(V, g)$  aus Eigenvektoren von  $F$ , wenn  $F$  normal ist.*

Insbesondere sind normale Endomorphismen über  $\mathbb{C}$  immer diagonalisierbar; die Aussage im Satz ist aber noch etwas stärker, da wir sogar eine unitäre (also eine Orthonormal-) Basis erhalten. Die Darstellung als Diagonalmatrix ist eindeutig bis auf die Reihenfolge der Einträge nach Satz 5.56. In der Funktionalanalysis heißt der entsprechende Satz auch der „Spektralsatz für normale Operatoren“.

BEWEIS. Zu „ $\implies$ “ sei  $(e_1, \dots, e_n)$  eine Orthonormalbasis aus Eigenvektoren von  $F$ . Nach Proposition 5.5 wird  $F$  bezüglich dieser Basis durch eine Diagonalmatrix  $A$  dargestellt. Nach Beispiel 6.30 (2) wird  $F^*$  durch  $A^*$  dargestellt, und  $A^*$  ist auch eine Diagonalmatrix. Man sieht leicht, dass  $A^* A = A A^*$  gilt, somit ist  $F$  normal.

Wir beweisen „ $\Leftarrow$ “ durch Induktion über die Dimension  $n$  von  $V$ . Im Fall  $n = 0$  ist nichts zu zeigen. Es sei also  $n \geq 1$ . Da  $\mathbb{C}$  algebraisch abgeschlossen ist, hat das charakteristische Polynom  $\chi_F$  eine Nullstelle  $\lambda$ . Es sei  $v$  ein Eigenvektor von  $F$  zum Eigenwert  $\lambda$ . Dann ist  $v$  auch ein Eigenvektor von  $F^*$  zum Eigenwert  $\bar{\lambda}$ , denn

$$\begin{aligned} \|F^*(v) - v \cdot \bar{\lambda}\|_g^2 &= g((F^* - \bar{\lambda} \operatorname{id}_V)(v), (F^* - \bar{\lambda} \operatorname{id}_V)(v)) \\ &= g((F - \lambda \operatorname{id}_V)(F^* - \bar{\lambda} \operatorname{id}_V)(v), v) \\ &= g((F^* - \bar{\lambda} \operatorname{id}_V)(F - \lambda \operatorname{id}_V)(v), v) = 0 \end{aligned}$$

Es sei

$$W = \{ w \in V \mid g(v, w) = 0 \}$$

das orthogonale Komplement vom Vektor  $v$ , dann ist  $W \subset V$  ein Untervektorraum, siehe Übungen. Der Unterraum  $W$  ist sowohl unter  $F$  als auch unter  $F^*$  invariant, denn sei  $w \in W$ , dann folgt

$$\begin{aligned} g(v, F(w)) &= g(F^*(v), w) = g(v \cdot \bar{\lambda}, w) = 0 \\ \text{und} \quad g(v, F^*(w)) &= g(F(v), w) = g(v \cdot \lambda, w) = 0, \end{aligned}$$

somit liegen mit  $w$  auch  $F(w)$  und  $F^*(w)$  wieder in  $W$ .

Insbesondere ist  $F^*|_W$  gleichzeitig die adjungierte Abbildung zu  $F|_W$  bezüglich des auf  $W$  eingeschränkten Skalarproduktes, und  $F|_W$  ist nach wie vor normal, also existiert nach Induktionsannahme eine unitäre Basis  $v_2, \dots, v_n$  von  $W$  aus Eigenvektoren von  $F|_W$ . Wir dürfen  $\|v\|_g = 1$  annehmen. Dann ist  $(v, v_2, \dots, v_n)$  eine unitäre Basis von  $V$  aus Eigenvektoren von  $F$ , und wir sind fertig.  $\square$

Über den reellen Zahlen verhalten sich normale Abbildungen etwas komplizierter. Man überprüft, dass Matrizen der Form

$$(*) \quad \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

normal sind, denn

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & 0 \\ 0 & a^2 + b^2 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \cdot \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Das charakteristische Polynom  $(X - a)^2 + b^2$  hat jedoch keine reellen Nullstellen, falls  $b \neq 0$ . Nach Lemma 5.33 hat die obige Matrix also keine Eigenvektoren.

**6.41. FOLGERUNG.** *Es sei  $(V, g)$  ein endlich-dimensionaler reeller Vektorraum mit Skalarprodukt und  $F \in \operatorname{End}_{\mathbb{R}} V$ . Dann existiert genau dann eine Orthonormalbasis von  $(V, g)$ , bezüglich der  $F$  durch eine Block-Diagonalmatrix aus  $1 \times 1$ -Blöcken und aus  $2 \times 2$ -Blöcken der Gestalt  $(*)$  mit  $b > 0$  dargestellt wird, wenn  $F$  normal ist. In diesem Fall ist die Matrix bis auf die Reihenfolge der Blöcke eindeutig.*

BEWEIS. Die Richtung „ $\implies$ “ folgt wie im Beweis von Satz 6.40 durch Nachrechnen.

Zu „ $\impliedby$ “ wählen wir zunächst eine beliebige Orthonormalbasis von  $V$  und stellen  $F$  durch eine normale Matrix  $A \in M_n(\mathbb{R})$  dar. Dann betrachten wir  $A$  als normale komplexe Matrix, also als normalen Endomorphismus von  $\mathbb{C}^n$  mit dem Standardskalarprodukt. Der Beweis geht wieder durch Induktion über  $n$ .

Wir finden einen gemeinsamen Eigenvektor  $v$  von  $A$  zum Eigenwert  $\lambda$  und von  $A^*$  zum Eigenwert  $\bar{\lambda}$ . Wenn  $\lambda$  reell ist, können wir  $v \in \mathbb{R}^n \subset \mathbb{C}^n$  wählen wegen Lemma 5.33. Danach betrachten wir das orthogonale Komplement  $W$  von  $v$  und machen weiter wie im obigen Beweis.

Wenn  $\lambda$  nicht reell ist, betrachten wir den Vektor  $\bar{v} \in \mathbb{C}^n$  und rechnen nach, dass

$$A\bar{v} = \bar{A}v = \overline{Av} = \overline{v \cdot \lambda} = \bar{v} \cdot \bar{\lambda},$$

so dass  $\bar{v}$  ein Eigenvektor von  $A$  zum Eigenwert  $\bar{\lambda} \neq \lambda$  ist. Dabei haben wir benutzt, dass  $\bar{\bar{A}} = A$ , da  $A$  eine reelle Matrix ist. Wir schreiben  $\lambda = a + bi$  mit  $a, b \in \mathbb{R}$  und  $v = w + ui$  mit  $u, w \in \mathbb{R}^n$ , und erhalten

$$\begin{aligned} Au &= \frac{i}{2} (A\bar{v} - Av) = \frac{i}{2} (\bar{v}\bar{\lambda} - v\lambda) \\ &= \frac{i}{2} ((w - ui)(a - bi) - (w + ui)(a + bi)) = ua + wb, \\ \text{und} \quad Aw &= \frac{1}{2} (A\bar{v} + Av) = \frac{1}{2} (\bar{v}\bar{\lambda} + v\lambda) \\ &= \frac{1}{2} ((w - ui)(a - bi) + (w + ui)(a + bi)) = -ub + wa. \end{aligned}$$

Wir nehmen an, dass  $b = \text{Im } \lambda > 0$ , andernfalls vertauschen wir die Rollen von  $v$  und  $\bar{v}$ . Dann hat  $A$  auf dem von  $u$  und  $w$  aufgespannten Unterraum  $U$  bezüglich der Basis  $(u, w)$  gerade die Gestalt (\*).

Als nächstes überlegen wir uns, dass  $v$  und  $\bar{v}$  aufeinander senkrecht stehen, da sie Eigenvektoren zu verschiedenen Eigenwerten sind, und somit wie im Beweis von Satz 6.40 der Vektor  $\bar{v}$  im orthogonalen Komplement von  $v$  liegt. Wir nehmen an, dass  $\|v\|_g = \|\bar{v}\|_g = 2$ , dann gilt

$$\begin{aligned} 2 &= \|v\|_g^2 = g(w + ui, w + ui) = \|w\|_g^2 + \|v\|_g^2 + (g(w, u) - g(u, w))i \\ 0 &= g(w - ui, w + ui) = \|w\|_g^2 - \|v\|_g^2 + 2g(w, u)i. \end{aligned}$$

Da  $u, w$  reell sind, sind auch alle einzelnen Skalarprodukte rechts reell. Hieraus folgt  $g(u, w) = 0$  und  $\|u\|_g^2 = \|w\|_g^2 = 1$ , so dass  $u, w$  eine Orthonormalbasis von  $U$  bilden. Wie im Beweis von Satz 6.40 ist das orthogonale Komplement

$$W = \{z \in \mathbb{C}^n \mid g(w, z) = g(u, z) = 0\} = \{z \in \mathbb{C}^n \mid g(v, z) = g(\bar{v}, z) = 0\}$$

invariant unter  $F$  und  $F^*$ , und wir können den Beweis wie oben fortsetzen.

Wir erhalten also eine Blockmatrix aus  $1 \times 1$ -Blöcken, die genau den reellen Nullstellen von  $\chi_F$  entsprechen, und aus  $2 \times 2$ -Blöcken der Gestalt (\*), so

dass  $a \pm bi$  echt komplexe Nullstellen von  $\chi_F$  sind. Somit können wir die gesuchte Matrix aus den komplexen Nullstellen des charakteristischen Polynoms ablesen, was die Eindeutigkeitsaussage beweist.  $\square$

Da die Quaternionen nicht kommutativ sind, ist der Begriff des Eigenraums nicht sinnvoll, siehe Übung 4 von Blatt 1. Dennoch erhalten können wir normale Abbildungen auch über den Quaternionen charakterisieren.

6.42. FOLGERUNG. *Es sei  $(V, g)$  ein Rechts- $\mathbb{H}$ -Vektorraum mit Skalarprodukt und  $F \in \text{End}_{\mathbb{H}} V$ . Dann existiert genau dann eine quaternionisch unitäre Basis von  $V$ , bezüglich der  $F$  durch eine Diagonalmatrix mit Einträgen der Form  $a + bi$  mit  $b \geq 0$  dargestellt wird, wenn  $F$  normal ist. Diese Matrix ist eindeutig bis auf Reihenfolge der Einträge.*

BEWEIS. Die Richtung „ $\implies$ “ überprüft man wieder durch Nachrechnen.

Wir beweisen „ $\impliedby$ “ wieder durch Induktion über  $n = \dim_{\mathbb{H}} V$ . Dazu betrachten wir  $\mathbb{C} = \mathbb{R} + i\mathbb{R} \subset \mathbb{H}$  als Teilkörper und fassen  $V$  für einen Moment als komplexen Vektorraum auf. Wir erhalten ein komplexes Skalarprodukt, indem wir die  $j$ - und  $k$ -Komponenten von  $g$  vergessen. Bezüglich dieses Skalarproduktes ist  $F$  als komplex lineare Abbildung immer noch normal mit der selben adjungierten Abbildung  $F^*$ . Also existiert wie oben ein simultaner Eigenvektor  $v$  zum Eigenwert  $\lambda = a + bi \in \mathbb{C}$  von  $F$  und zum Eigenwert  $\bar{\lambda} = a - bi$  von  $F^*$ . Wenn  $b \geq 0$  gilt, dann ist das orthogonale Komplement

$$W = \{ w \in V \mid g(v, w) = 0 \in \mathbb{H} \}$$

ein invarianter quaternionischer Unterraum der Dimension  $n-1$ , und wir fahren fort wie im Beweis von Satz 6.40.

Falls  $b < 0$ , betrachten wir den Vektor  $v \cdot j$ . Es gilt

$$F(v \cdot j) = F(v) \cdot j = v \cdot ((a + bi)j) = v \cdot (j(a - bi)) = (v \cdot j) \cdot \bar{\lambda}$$

und genauso  $F^*(v \cdot j) = (v \cdot j) \cdot \lambda$ . Anstelle von  $v$  betrachten wir also  $v \cdot j$  und machen weiter wie oben und erhalten die gesuchte quaternionisch unitäre Basis.

Zur Eindeutigkeit überlegen wir uns, dass das charakteristische Polynom  $\chi_F$  von  $F$  als Endomorphismus über dem Körper  $\mathbb{C}$  aufgrund der obigen Überlegung in Faktoren der Form

$$(X - \lambda)(X - \bar{\lambda}) = (X - a)^2 + b^2$$

zerfällt. Dadurch sind die Diagonaleinträge bis auf ihre Reihenfolge eindeutig festgelegt.  $\square$

Wir können „ $i$ “ in der Darstellung  $a + bi$  auch durch  $j$ ,  $k$  oder einen beliebigen anderen imaginären Einheitsquaternion  $q$  ersetzen. Dadurch ändern sich die Matrix und die zugehörige Basis, aber nicht die Paare  $(a, b)$  in  $a + bq$ . Im Beweis arbeiten wir dann mit einem Teilkörper  $\mathbb{R} + q\mathbb{R} \cong \mathbb{C}$ .

Wir kommen jetzt zu wichtigen Spezialfällen normaler Abbildungen.

6.43. FOLGERUNG (Hauptachsentransformation). *Es sei  $(V, g)$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum mit Skalarprodukt und  $F \in \text{End}_{\mathbb{k}} V$ . Dann existiert genau dann eine Orthonormalbasis von  $V$ , bezüglich der  $F$  durch eine Diagonalmatrix mit reellen Einträgen dargestellt wird, wenn  $F$  selbstadjungiert ist.*

Aus den obigen Ergebnissen folgt dann auch die Eindeutigkeit dieser Diagonalmatrix bis auf die Reihenfolge der Diagonaleinträge. In der Funktionalanalysis heißt der entsprechende Satz auch der „Spektralsatz für selbstadjungierte Operatoren“.

BEWEIS. Die Richtung „ $\implies$ “ ergibt sich wieder durch Nachrechnen.

Zu „ $\impliedby$ “ wenden wir Satz 6.40 oder eine der Folgerungen 6.41 oder 6.42 an. Da  $F$  selbstadjungiert ist, ist auch die Matrix  $A \in M_n(\mathbb{k})$ , die wir so erhalten, selbstadjungiert. Im Fall  $\mathbb{k} = \mathbb{R}$  ist ein  $2 \times 2$ -Block vom Typ \* nur dann selbstadjungiert, wenn  $b = 0$  gilt. In den Fällen  $\mathbb{k} = \mathbb{C}$  oder  $\mathbb{H}$  muss  $\lambda = \bar{\lambda}$  für jeden Diagonaleintrag  $\lambda \in \mathbb{k}$  gelten, und wegen Bemerkung 6.1 (4) folgt  $\lambda \in \mathbb{R}$ .  $\square$

6.44. BEISPIEL. Wir betrachten einen physikalischen Körper  $K$  im  $\mathbb{R}^3$ , der sich ohne Einfluss äußerer Kräfte bewegt. Dabei nehmen wir an, dass der Schwerpunkt für alle Zeiten im Nullpunkt liegt. Dann dreht sich der Körper um sich selbst.

Um diese Drehung zu beschreiben, betrachtet man zu einer festen Zeit  $t$  den Trägheitstensor  $F_t: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$F_t(v) = \int_{K_t} \rho_t(p) p \times (v \times p) d^3x,$$

wobei  $K_t \subset \mathbb{R}^3$  den Körper zur Zeit  $t$  und  $\rho_t(x)$  seine (Massen-) im Punkt  $x$  bezeichne. Dabei bezeichne die Richtung von  $v$  die Drehachse und  $\|v\|$  die Drehgeschwindigkeit, dann beschreibt  $v \times p$  die tatsächliche Geschwindigkeit im Punkt  $p$ , und  $\rho_t(p) p \times (v \times p)$  den Beitrag zum Drehimpuls. Insgesamt ist  $F_t(v_t)$  also der Drehimpuls zur Zeit  $t$ , wenn  $v_t$  wie oben die Drehung zur Zeit  $t$  beschreibt.

Die Abbildung  $F_t$  ist selbstadjungiert. Am einfachsten überlegt man sich das für den Integranden: für alle  $w \in \mathbb{R}^3$  gilt

$$\langle p \times (v \times p), w \rangle = \langle p, p \rangle \langle v, w \rangle - \langle p, v \rangle \langle p, w \rangle = \langle v, p \times (w \times p) \rangle$$

nach Satz 1.69 (2). Also existiert eine Orthonormalbasis  $(e_1(t), e_2(t), e_3(t))$  von  $\mathbb{R}^3$  aus Eigenvektoren von  $F_t$ . Wegen der Cauchy-Schwarz-Ungleichung 6.10 sind alle Eigenwerte  $\lambda_1 \leq \lambda_2 \leq \lambda_3$  positiv, falls der Körper sich in jeder Raumrichtung ausdehnt.

Die Richtungen der Eigenvektoren heißen die *Hauptachsen* des Körpers  $K$ . Bei einem achsenparallelen Quader sind das beispielsweise gerade die Koordinatenachsen. Wenn sich der Körper zu einer festen Zeit um eine der Hauptachsen dreht, dann tut er das für alle Zeit. Dreht er sich hingegen um eine andere Achse, dann verändert sich die Drehachse selbst im Laufe der Zeit; der Körper scheint

zu taumeln. Es gibt jedoch einen konstanten Drehimpulsvektor  $L = F_t(v_t) \in \mathbb{R}^3$ , so dass die Drehachse zu jedem Zeitpunkt in Richtung  $F_t^{-1}(L)$  zeigt.

6.45. FOLGERUNG. *Es sei  $(V, g)$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum, und  $S$  sei eine Sesquilinearform auf  $V$ . Dann existiert genau dann eine  $g$ -Orthonormalbasis von  $V$ , bezüglich der  $S$  durch eine Diagonalmatrix mit reellen Einträgen dargestellt wird, wenn  $S$  Hermitesch ist.*

Wie in Folgerung 6.43 sind die Diagonaleinträge bis auf ihre Reihenfolge eindeutig.

BEWEIS. Die Richtung „ $\implies$ “ ergibt sich aus Bemerkung 6.15 (2).

Zu „ $\impliedby$ “ fassen wir zunächst  $S$  als antilineare Abbildung  $S: V \rightarrow V^*$  wie in Proposition 6.22 auf. Da wir nichts über die Definitheit von  $S$  wissen, können wir allerdings nicht schließen, dass  $S$  injektiv ist. Sei  $g^{-1}: V^* \rightarrow V$  die antilineare Umkehrabbildung aus Lemma 6.26, dann setzen wir  $F = g^{-1} \circ S \in \text{End}_{\mathbb{k}} V$ , so dass

$$S(v, w) = (g \circ F)(v)(w) = g(F(v), w) \quad \text{für alle } v, w \in V.$$

Da  $S$  Hermitesch ist, ist  $F$  selbstadjungiert, und Folgerung 6.43 liefert eine Orthonormalbasis  $(e_1, \dots, e_n)$  aus Eigenwerten von  $F$ . Man überlegt sich leicht, dass  $F$  und  $S$  bezüglich  $(e_1, \dots, e_n)$  durch die selbe Matrix dargestellt werden, also durch eine Diagonalmatrix mit reellen Eigenwerten.  $\square$

6.46. BEISPIEL. Ein Brillenglas ist eine gekrümmte Fläche. Die Wirkung des Glases auf Lichtstrahlen hängt von der Krümmung ab. Wenn wir das Glas in einem Punkt  $p$  flach auf den Tisch legen, können wir eine Seite des Glases als Graph einer Funktion  $f: U \rightarrow \mathbb{R}$  mit  $U \subset \mathbb{R}^2$  darstellen. Wenn  $f$  mindestens zweimal stetig differenzierbar ist, beschreibt die zweite Ableitung bei  $p$  die Krümmung an der Stelle  $p$ . Nach dem Satz von Schwarz ist die zweite Ableitung an der Stelle  $p$  eine reelle symmetrische Bilinearform (also eine reelle Hermiteische Sesquilinearform)  $f''(p): \mathbb{R}^2 \rightarrow \mathbb{R}$ , und die Krümmung in Richtung  $v \in \mathbb{R}^2$  wird gegeben als

$$f''(p)(v, v) \quad \text{für alle } v \in \mathbb{R}^2 \text{ mit } \|v\| = 1.$$

Nach Folgerung 6.45 können wir  $f''(p)$  bezüglich einer Orthogonalbasis  $(v_1, v_2)$  des  $\mathbb{R}^2$  als Diagonalmatrix mit Einträgen  $\kappa_1, \kappa_2$  schreiben. Dann nennt man  $\kappa_1$  und  $\kappa_2$  die *Hauptkrümmungen* im Punkt  $p$ , und  $v_1, v_2$  die *Hauptkrümmungsrichtungen*. Wir dürfen  $\kappa_1 \leq \kappa_2$  annehmen. Bei einem gewöhnlichen Brillenglas sollten die Hauptkrümmungen und die Hauptkrümmungen über das ganze Glas in etwa konstant bleiben. In diesem Fall muss der Augenarzt dem Optiker die Krümmungen (als Werte in Dioptrien) und eine Hauptkrümmungsrichtung mitteilen. Die andere Hauptkrümmungsrichtung ergibt sich, da beide senkrecht aufeinander stehen.

6.47. FOLGERUNG (Singuläre Werte). *Es seien  $(V, g)$  und  $(W, h)$  endlich-dimensionale  $\mathbb{k}$ -Vektorräume mit Skalarprodukten, und es sei  $F: V \rightarrow W$  linear.*

Dann existieren Orthonormalbasen von  $V$  und von  $W$ , so dass  $F$  bezüglich dieser Basen dargestellt wird durch eine Matrix der Form

$$\begin{pmatrix} a_1 & 0 & & \cdots & & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ & & \ddots & a_{\text{rg } F} & 0 & \cdots & 0 \\ \vdots & & & 0 & 0 & \cdots & 0 \\ & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

mit eindeutig bestimmten reellen Einträgen  $a_1 \geq \cdots \geq a_{\text{rg } F} > 0$ .

Diese Folgerung ist eine Verfeinerung des Rangsatzes 3.13, in dem anstelle von Orthonormalbasen beliebige Basen erlaubt sind.

BEWEIS. Die Abbildung  $F^*F: V \rightarrow V$  ist offensichtlich selbstadjungiert und hat nicht-negative Eigenwerte, denn die zugehörige Hermitesche Bilinearform

$$S(u, v) = g(u, (F^* \circ F)(v)) = h(F(u), F(v)) = (F^*h)(u, v)$$

ist positiv semidefinit nach Bemerkung 6.35 (3). Die Hauptachsentransformation 6.43 liefert eine Orthonormalbasis  $(v_1, \dots, v_n)$  aus Eigenvektoren von  $F^* \circ F$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$ ; dabei sortieren wir die Basisvektoren so, dass  $\lambda_1 \geq \cdots \geq \lambda_\ell > 0 = \lambda_{\ell+1} = \cdots = \lambda_n$ .

Jetzt betrachten wir die Vektoren  $w_1 = F(v_1) \cdot \lambda_1^{-\frac{1}{2}}, \dots, w_\ell = F(v_\ell) \cdot \lambda_\ell^{-\frac{1}{2}}$ . Da die Faktoren  $\lambda_p^{-\frac{1}{2}}$  reell sind, folgt

$$h(w_p, w_q) = \frac{h(F(v_p), F(v_q))}{\sqrt{\lambda_p} \sqrt{\lambda_q}} = \frac{g(v_p, (F^*F)(v_q))}{\sqrt{\lambda_p \lambda_q}} = \delta_{pq} \frac{\sqrt{\lambda_q}}{\sqrt{\lambda_p}} = \delta_{pq}.$$

Nach Bemerkung 6.17 (1) sind die Vektoren  $w_1, \dots, w_\ell$  linear unabhängig. Also ergänzen wir mit dem Basisergänzungssatz 3.3 zu einer Basis von  $W$ , die wir mit dem Gram-Schmidt-Verfahren 6.18 in eine Orthonormalbasis  $(w_1, \dots, w_m)$  überführen. Bezüglich der so konstruierten Basen hat  $F$  die angegebene Abbildungsmatrix, wobei  $a_p = \sqrt{\lambda_p}$  für alle  $p = 1, \dots, \ell = \text{rg } F$ .

Die Eindeutigkeitsaussage ergibt sich, indem man aus der gegebenen Abbildungsmatrix die Abbildungsmatrix von  $F^*F$  ableitet und die Eindeutigkeitsaussage aus Folgerung 6.43 benutzt.  $\square$

6.48. BEMERKUNG. Die singulären Werte geben also an, wie stark die Abbildung  $F$  die Längen in unterschiedlichen Richtungen verzerrt. Wenn wir beispielsweise eine Gummifolie als Fläche im Raum betrachten, dann können wir das als eine Abbildung  $f: U \rightarrow \mathbb{R}^3$  mit  $U \subset \mathbb{R}^2$  betrachten. Es sei  $p \in U$ , dann gibt die Ableitung  $F = df(p): \mathbb{R}^2 \rightarrow \mathbb{R}^3$  an, wie  $f$  am Punkt  $p$  die Richtungen im  $\mathbb{R}^2$  in den  $\mathbb{R}^3$  abbildet. Die singulären Werte  $a_1 \geq a_2$  geben das Maximum und das Minimum der Längenverzerrung an. Nach Folgerung 6.47 stehen die zugehörigen Richtungen immer senkrecht aufeinander.

Die singulären Werte  $a_1, \dots, a_{\text{rg } F}$  heißen manchmal auch *verallgemeinerte Eigenwerte* von  $F$ . Diese Bezeichnung ist etwas unglücklich, da für eine Matrix  $A$  die Eigenwerte von  $A$  mit den verallgemeinerten Eigenwerten, also den Eigenwerten von  $A^*A$ , nichts zu tun haben müssen. Als Beispiel betrachte einen Jordan-Block  $M_\ell(\lambda) \in M_\ell(\mathbb{k})$  der Grösse  $\ell$  zum Eigenwert  $\lambda$ . Dann hat

$$M_\ell(\lambda)^* M_\ell(\lambda) = \begin{pmatrix} \lambda^2 + 1 & \lambda & & 0 \\ \lambda & \ddots & \ddots & \\ & \ddots & \lambda^2 + 1 & \lambda \\ 0 & & \lambda & \lambda^2 \end{pmatrix}$$

für  $\ell = 2$  die Eigenwerte  $\lambda^2 + \frac{1}{2} \pm \sqrt{\lambda^2 + \frac{1}{4}}$ , und die singulären Werte sind die positiven Wurzeln davon.

Ähnliche Aussagen wie in Folgerung 6.43 lassen sich auch für schiefe Endomorphismen  $F \in \text{End}_{\mathbb{k}} V$  eines endlich-dimensionalen  $\mathbb{k}$ -Vektorraums beweisen. Dazu muss man wieder nur untersuchen, welche der möglichen Normalformen in Satz 6.40 und den Folgerungen 6.41 und 6.42 schiefe Endomorphismen beschreiben.

6.49. FOLGERUNG. *Es sei  $(V, g)$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum mit Skalarprodukt und  $F \in \text{End}_{\mathbb{k}} V$ . Dann existiert genau dann eine Orthonormalbasis von  $V$ , bezüglich der  $F$  dargestellt wird*

- (1) *durch eine Block-Diagonalmatrix aus  $1 \times 1$ -Blöcken  $0$  und  $2 \times 2$ -Blöcken vom Typ (\*) mit  $a = 0$  falls  $\mathbb{k} = \mathbb{R}$ ,*
- (2) *durch eine Diagonalmatrix mit rein imaginären Einträgen falls  $\mathbb{k} = \mathbb{C}$ , beziehungsweise*
- (3) *durch eine Diagonalmatrix mit Einträgen der Form  $bi$  mit  $b \geq 0$  falls  $\mathbb{k} = \mathbb{H}$ ,*

wenn  $F$  schief ist.

BEWEIS. Analog zum Beweis von Folgerung 6.43. □

Besonders interessant ist der Fall, dass  $F$  eine Isometrie ist. Aus Kapitel 1 kennen wir Spiegelungen und Drehungen.

6.50. FOLGERUNG. *Es sei  $(V, g)$  ein endlich-dimensionaler  $\mathbb{k}$ -Vektorraum mit Skalarprodukt und  $F \in \text{End}_{\mathbb{k}} V$ . Dann existiert genau dann eine Orthonormalbasis von  $V$ , bezüglich der  $F$  dargestellt wird*

- (1) *durch eine Block-Diagonalmatrix aus  $1 \times 1$ -Blöcken  $\pm 1$  und  $2 \times 2$ -Blöcken vom Typ (\*) mit  $a^2 + b^2 = 1$  falls  $\mathbb{k} = \mathbb{R}$ ,*
- (2) *durch eine Diagonalmatrix mit Einträgen vom Betrag 1 falls  $\mathbb{k} = \mathbb{C}$ , beziehungsweise*
- (3) *durch eine Diagonalmatrix mit Einträgen der Form  $a+bi$  vom Betrag 1 mit  $b \geq 0$  falls  $\mathbb{k} = \mathbb{H}$ ,*

wenn  $F$  eine lineare Isometrie ist.

BEWEIS. Analog zum Beweis von Folgerung 6.43. □

6.51. BEMERKUNG. In Aufgabe 2 von Blatt 14 zur linearen Algebra I und Bemerkung 4.29 haben wir die Untergruppen

$$O(n) = \{ A \in M_n(\mathbb{R}) \mid A^t \cdot A = E_n \}$$

und  $SO(n) = \{ A \in O(n) \mid \det A = 1 \}$

der Gruppe  $GL(n, \mathbb{R})$  kennengelernt. Die Elemente von  $O(n)$  sind dadurch charakterisiert, dass sie das Standard-Skalarprodukt erhalten, somit ist die *orthogonale Gruppe*  $O(n)$  die Gruppe der linearen Isometrien des  $\mathbb{R}^n$ . Gleichzeitig ist  $O(n)$  auch die Gruppe der Basiswechsellmatrizen zwischen Orthonormalbasen, siehe Proposition 2.77 und Bemerkung 6.15 (4).

Die *spezielle orthogonale Gruppe*  $SO(n)$  ist die Gruppe der orientierungserhaltenden Isometrien. Gleichzeitig ist sie die Gruppe der Basiswechsellmatrizen zwischen gleich orientierten Orthonormalbasen.

Analog betrachten wir die *unitäre* und die *spezielle unitäre Gruppe*

$$U(n) = \{ A \in M_n(\mathbb{C}) \mid A^* \cdot A = E_n \}$$

und  $SU(n) = \{ A \in U(n) \mid \det A = 1 \}$ .

Die unitäre Gruppe  $U(n)$  ist die Gruppe der linearen Isometrien des  $\mathbb{C}^n$  mit dem Standardskalarprodukt, und gleichzeitig die Gruppe der Basiswechsellmatrizen zwischen unitären Basen. Für Elemente  $A \in U(n)$  gilt

$$1 = \det(A^* A) = |\det a|^2,$$

und das Beispiel  $(e^{it}) \in U(1)$  zeigt, dass alle komplexen Zahlen vom Betrag 1 als Determinante einer unitären Matrix auftreten können. Da wir Orientierungen für komplexe Vektorräume nicht eingeführt haben, ist  $SU(n)$  einfach nur die Untergruppe der Isometrien mit Determinante 1.

Über den Quaternionen definieren wir nur die (*kompakte*) *symplektische Gruppe*

$$Sp(n) = \{ A \in M_n(\mathbb{H}) \mid A^* \cdot A = E_n \}$$

der linearen Isometrien des  $\mathbb{H}^n$  mit Standardskalarprodukt, beziehungsweise der Basiswechsellmatrizen zwischen quaternionisch unitären Basen. Da die Quaternionen nicht kommutativ sind, gibt es keine Determinante, und wir definieren nur diese eine Gruppe.

6.52. BEMERKUNG. Wir haben wieder eine Reihe von Normalformen kennengelernt und auch ein paar Anwendungen gesehen.

- (1) Sei  $(V, g)$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum mit Skalarprodukt, dann ist für jede Orthonormalbasis  $B$  von  $V$  die Basisabbildung ein Isomorphismus  $B: \mathbb{k}^n \rightarrow V$ , so dass  $B^*g$  gerade das Standardskalarprodukt auf  $\mathbb{k}^n$  ist. Insbesondere ist die Dimension eine vollständige Invariante für endlich-dimensionale  $\mathbb{k}$ -Vektorräume mit Skalarprodukt, ähnlich wie in Folgerung 3.8 für endlich-dimensionale Vektorräume.

- (2) Es sei  $(V, g)$  ein  $\mathbb{k}$ -Vektorraum mit Skalarprodukt. Wir betrachten zwei Endomorphismen  $F, G \in \text{End}_{\mathbb{k}} V$  als *metrisch äquivalent*, wenn es eine lineare Isometrie  $U \in \text{Aut}_{\mathbb{k}} V$  gibt, so dass  $G = U^{-1}FU$ . Somit sind  $F$  und  $G$  genau dann äquivalent, wenn es Orthonormalbasen  $B$  und  $C$  gibt, so dass  $F$  bezüglich  $B$  die gleiche Darstellung hat wie  $G$  bezüglich  $C$ . Dann haben wir in Satz 6.40 und den Folgerungen 6.41–6.42 eine Normalform für normale Endomorphismen kennengelernt. Spezialfälle haben wir in den Folgerungen 6.43, 6.49 und 6.50 betrachtet. Für selbstadjungierte Matrizen beispielsweise erhalten wir als vollständige Invariante die Dimension  $\dim V$  und das Tupel der nach Größe geordneten reellen Eigenwerte.

Man beachte, dass nicht nur die Auswahl der betrachteten Endomorphismen spezieller ist als in Kapitel 5, sondern auch die Äquivalenzrelation.

- (3) Wir nennen zwei lineare Abbildungen  $F, G: V \rightarrow W$  zwischen Vektorräumen *metrisch äquivalent*, wenn es lineare Isometrien  $P \in \text{Aut}_{\mathbb{k}} V$  und  $Q \in \text{Aut}_{\mathbb{k}} W$  gibt, so dass  $Q \circ F = G \circ P$ , siehe Folgerung 3.18. In diesem Fall liefert Folgerung 6.47 eine Normalform, und  $(\dim, \dim W, \text{rg } F)$  bildet zusammen mit dem Tupel der nach Größe geordneten singulären Werte eine vollständige Invariante.

6.53. BEMERKUNG. Es sei  $F \in \text{End}_{\mathbb{R}}(V)$  eine Isometrie eines endlich-dimensionalen Euklidischen Vektorraums  $(V, g)$ .

- (1) Gemäß der Orthonormalbasis aus Folgerung 6.50 zerlegen wir  $V$  in eine direkte Summe von Unterräumen, die paarweise zueinander senkrecht stehen. Dann operiert  $F$  auf den eindimensionalen Unterräumen  $U_i$  mit Eigenwert  $\pm 1$ , also als  $\pm \text{id}_{U_i}$ , das heißt als Identität oder als Spiegelung.

Auf den zweidimensionalen Eigenräumen  $V_j$  wirkt  $V$  durch eine Matrix vom Typ (\*) mit  $a^2 + b^2 = 1$  und  $b > 0$ . Also finden wir einen Winkel  $\varphi = \arccos a \in [0, \pi]$ , so dass

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Diese Matrix beschreibt eine Drehung des Raumes  $V_j$  um den Winkel  $\varphi$  und heißt daher auch einfach *Drehmatrix*.

- (2) Gemäß Definition 4.27 heißt  $F$  genau dann orientierungserhaltend, wenn  $\det F > 0$ . Nach Folgerung 4.18 (1) ergibt sich die Determinante als das Produkt der Determinanten der einzelnen Blöcke. Ein  $1 \times 1$ -Block  $\pm 1$  hat Determinante  $\pm 1$ , während eine Drehmatrix stets Determinante  $a^2 + b^2 = \cos^2 \varphi + \sin^2 \varphi = 1$  hat. Somit ist eine Isometrie genau dann orientierungserhaltend, wenn die Anzahl der Spiegelungen in (1), also die Dimension des  $-1$ -Eigenraumes, gerade ist.
- (3) Wenn  $V$  eine feste Orientierung trägt, versuchen wir, in Folgerung 6.50 eine orientierte Basis anzugeben. Das geht immer, wenn  $\pm 1$  Eigenwert ist, da wir dann das Vorzeichen des zugehörigen Eigenvektors frei

wählen können. In einem Drehblock legt jedoch die Wahl des Winkels  $\varphi \in (0, \pi)$  eine Orientierung fest. Wenn wir also nur Drehmatrizen zu Winkeln  $\varphi_i \in (0, \pi)$  haben, müssen wir unter Umständen einen Winkel  $\varphi$  durch  $-\varphi$  ersetzen. In diesem Fall können wir sagen, dass  $F$  entgegen dem mathematischen Drehsinn wirkt. In der Ebene ist eine Drehung im mathematischen Drehsinn eine Drehung gegen den Uhrzeigersinn, und umgekehrt.

### 6.5. Affine Räume

Wenn wir bei einem Vektorraum den Nullpunkt „vergessen“, erhalten wir einen affinen Raum. In Definition 3.19 hatten wir bereits affine Unterräume von Vektorräumen kennengelernt. In diesem Abschnitt wollen wir affine Räume etwas abstrakter einführen und auch als metrische Räume betrachten.

Für den Anfang betrachten wir wieder beliebige Schiefkörper  $\mathbb{k}$ .

6.54. DEFINITION. Es sei  $V$  ein  $\mathbb{k}$ -Vektorraum. Ein *affiner Raum* über  $V$  ist eine Menge  $A$ , zusammen mit einer Abbildung  $+: A \times V \rightarrow A$ , so dass gilt:

- (1) für alle  $a \in A$  und alle  $v, w \in V$  gilt  $(a + v) + w = a + (v + w)$ ,
- (2) zu je zwei Punkten  $a, b \in A$  existiert genau ein  $v \in V$  mit  $b = a + v$ .

Die *Dimension* von  $A$  ist gerade die Dimension von  $V$ .

Es sei  $B$  ein weiterer affiner Raum über einem  $\mathbb{k}$ -Vektorraum  $W$ . Eine Abbildung  $F: A \rightarrow B$  heißt *affin*, wenn es eine lineare Abbildung  $L: V \rightarrow W$  gibt, so dass

$$F(a + v) = F(a) + L(v)$$

für alle  $a \in A$  und alle  $v \in V$ . In diesem Fall nennt man  $F$  auch lineare Abbildung *über*  $L$ .

Eine nichtleere Teilmenge  $C \subset A$  heißt *affiner Unterraum*, wenn es einen Untervektorraum  $U \subset V$  gibt, so dass für alle  $c \in C$  und alle  $v \in V$  der Punkt  $c + v$  genau dann in  $C$  liegt, wenn  $v \in U$ . Man nennt  $C$  dann auch affinen Unterraum *über*  $U$ . Zwei affine Unterräume heißen *parallel*, wenn sie über dem gleichen linearen Unterraum von  $V$  liegen.

6.55. BEISPIEL. Wir kennen schon einfache Beispiele.

- (1) Jeder Vektorraum ist ein affiner Raum über sich selbst. Die affinen Unterräume im Sinne von Definition 3.19 sind genau die affinen Unterräume im obigen Sinne. Sei  $F: V \rightarrow W$  eine affine Abbildung über der linearen Abbildung  $L: V \rightarrow W$ , dann folgt

$$A(v) = A(0 + v) = A(0) + L(v) \quad \text{für alle } v \in V .$$

Also haben affine Abbildungen zwischen Vektorräumen stets die Gestalt

$$A(v) = L(v) + w ,$$

wobei  $L$  linear ist Umgekehrt ist jede Abbildung dieser Form affin.

- (2) Es sei  $V$  ein Vektorraum und  $U \subset V$  ein Unterraum. Dann ist jeder zu  $U$  parallele affine Unterraum  $A$  selbst ein affiner Raum über dem Untervektorraum  $U$ . Beispielsweise ist die Lösungsmenge eines inhomogenen Gleichungssystems ein affiner Raum über der Lösungsmenge des zugehörigen homogenen Gleichungssystems, siehe Proposition 3.22 (3).

6.56. BEMERKUNG. Wir sammeln ein paar elementare Eigenschaften.

- (1) Für jeden Punkt  $a \in A$  ist die Zuordnung  $v \mapsto a + v$  eine Bijektion von  $V$  nach  $A$ . Die Umkehrabbildung schreiben wir als Subtraktion

$$-: A \times A \rightarrow V,$$

so dass  $b - a = v$  genau dann, wenn  $b = a + v$ . Eine andere Bezeichnung ist  $\overrightarrow{ab} = b - a$ .

- (2) Es sei  $A$  ein affiner Raum über einem  $\mathbb{k}$ -Vektorraum  $V$ . Wenn wir einen *Ursprung*  $o \in A$  wählen, können wir  $A$  und  $V$  identifizieren, indem wir  $a \in A$  mit dem Vektor  $a - o \in V$  und  $v \in V$  mit dem Punkt  $a + v \in A$  gleichsetzen.

Wir setzen wieder  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$  und erinnern uns an den Begriff einer Norm auf einem  $\mathbb{k}$ -Vektorraum, siehe Bemerkung 6.9. In Definition 6.8 haben wir speziell die Norm  $\|\cdot\|_g$  zu einem Skalarprodukt  $g$  auf  $V$  eingeführt.

6.57. DEFINITION. Es sei  $A$  ein affiner Raum über einem  $\mathbb{k}$ -Vektorraum  $V$  und  $\|\cdot\|$  eine Norm auf  $V$ . Dann definieren wir die *affine Metrik*  $d: A \times A \rightarrow \mathbb{R}$  zu  $\|\cdot\|$  auf  $A$  durch

$$d(a, b) = \|a - b\| \quad \text{für alle } a, b \in A.$$

Wenn  $\|\cdot\| = \|\cdot\|_g$  die Euklidische Norm zu einem Skalarprodukt  $g$  auf  $V$  ist, nennen wir  $d_g = d$  eine *Euklidische Metrik* auf  $V$ . Ein affiner Raum mit einer Euklidischen Metrik heißt auch *Euklidischer Raum*  $(A, d)$ .

Eine affine Abbildung  $F: A \rightarrow B$  zwischen Euklidischen Räumen  $(A, d)$  und  $(B, e)$  heißt (*affine*) *isometrische Einbettung*, wenn

$$e(F(a), F(b)) = d(a, b) \quad \text{für alle } a, b \in A,$$

und (*affine*) *Isometrie*, wenn sie darüberhinaus invertierbar ist.

Obwohl es hier nicht gefordert haben, ist es für Studium Euklidischer Räume  $(A, d)$  am sinnvollsten, anzunehmen, dass der die zugrundeliegenden Vektorräume reell sind, also über  $\mathbb{k} = \mathbb{R}$  zu arbeiten. Mehr dazu später.

6.58. BEISPIEL. In der Schule haben Sie die Geometrie der Euklidischen Ebene  $(\mathbb{R}^2, d_g)$  studiert, wobei  $d_g$  zum Standard-Skalarprodukt auf  $\mathbb{R}^2$  gehört. Analog kann man Euklidische Räume  $(\mathbb{R}^n, d_g)$  beliebiger Dimension betrachten. Wir nennen  $d_g$  später die Standardmetrik.

In der klassischen Newtonschen Mechanik geht man davon aus, dass uns ein dreidimensionaler Euklidischer Raum umgibt. In diesem Raum ist weder ein

Ursprung festgelegt (obwohl er von manchen Leuten auf der Erde, von anderen im Mittelpunkt der Sonne oder gar im Mittelpunkt der Galaxie gesehen wird), noch gibt es ausgezeichnete Richtung (wenn wir einen festen Punkt auf der Erde als Ursprung wählen, könnten wir als Richtungen zum Beispiel „Norden“, „Westen“ und „oben“ wählen, aber diese Wahl hängt dann von der Wahl unseres Ursprungs ab).

Auf der anderen Seite gibt es in der klassischen Mechanik die Vorstellung, dass es eine Euklidische Metrik  $d$  unabhängig vom Bezugspunkt gibt. Selbst, wenn sich der Ursprung entlang einer Geraden mit konstanter Geschwindigkeit bewegt, soll sich an dieser Metrik nichts ändern. Die zweite dieser Annahmen wird in Einsteins spezieller Relativitätstheorie durch die etwas komplizierteren Lorentzschen Transformationsformeln ersetzt. In der allgemeinen Relativitätstheorie schließlich wird aus dem „flachen“ Euklidischen Raum eine gekrümmte Raumzeit.

6.59. BEMERKUNG. Wir können Euklidische Räume als metrische Räume betrachten.

- (1) Eine Metrik auf einer Menge  $M$  ist eine Funktion  $d: M \times M \rightarrow \mathbb{R}$ , so dass für alle  $a, b, c \in M$  die folgenden Axiome gelten:

- (D1)  $d(a, b) \geq 0$  und  $d(a, b) = 0 \iff a = b$  (*Positivität*),  
 (D2)  $d(b, a) = d(a, b)$  (*Symmetrie*),  
 (D3)  $d(a, c) \leq d(a, b) + d(b, c)$  (*Dreiecksungleichung*).

Dann nennt man  $(M, d)$  einen *metrischen Raum*.

Für eine affine Metrik zu einer Norm  $\|\cdot\|$  auf  $V$  folgen diese Axiome jeweils aus den entsprechenden Axiomen (N1)–(N3) für  $\|\cdot\|$ .

Auf der anderen Seite kommt nicht jede Metrik auf  $A$  von einer Norm, beispielsweise gehört zu keiner Norm die „diskreten Metrik“

$$d(a, b) = \begin{cases} 0 & \text{falls } a = b, \text{ und} \\ 1 & \text{sonst.} \end{cases}$$

Also ist nicht jede Metrik auf einem affinen Raum eine affine Metrik. Das lässt sich auch dadurch erklären, dass die Homogenität (N2) zum Beweis der Symmetrie nur für die Skalare  $\pm 1$  benutzt wird.

- (2) Es sei  $d = d_g$  eine Euklidische Metrik auf  $A$ . In der Dreiecksungleichung gilt Gleichheit genau dann, wenn es reelle Zahlen  $r, s \geq 0$  gibt, die nicht beide verschwinden, so dass

$$(b - a)r = (c - b)s \in V.$$

Somit zeigen beide Vektoren „in die gleiche Richtung“. Zur Begründung schreiben wir  $v = b - a$  und  $w = c - b \in V$  und betrachten

den Beweis der Dreiecksungleichung in Bemerkung 6.9, wonach

$$\begin{aligned}\|v + w\|_g^2 &= \|v\|_g^2 + 2 \operatorname{Re} g(v, w) + \|w\|_g^2 \\ &\leq \|v\|_g^2 + 2 |g(v, w)| + \|w\|_g^2 \\ &\leq \|v\|_g^2 + 2 \|v\|_g \|w\|_g + \|w\|_g^2 = (\|v\|_g + \|w\|_g)^2.\end{aligned}$$

Wegen der Cauchy-Schwarz-Ungleichung 6.10 wird aus der zweiten Ungleichung genau dann eine Gleichung, wenn  $v$  und  $w$  linear abhängig sind. Wir wollen annehmen, dass  $r \in \mathbb{k}$  mit  $w = v \cdot r$  existiert, ansonsten vertauschen wir die Rollen von  $v$  und  $w$ . Dann gilt

$$\operatorname{Re}(g(v, v \cdot r)) = \operatorname{Re}(r) \underbrace{g(v, v)}_{\geq 0} \leq |r| g(v, v),$$

und Gleichheit gilt genau dann, wenn  $r$  eine nichtnegative reelle Zahl ist. Mit  $s = 1$  erhalten wir die obige Behauptung.

- (3) Eine *Isometrie* zwischen metrischen Räumen  $(M, d)$  und  $(N, e)$  ist eine invertierbare Abbildung  $F: M \rightarrow N$ , so dass

$$e(F(a), F(b)) = d(a, b) \quad \text{für alle } a, b \in M.$$

Es seien wieder  $(A, d)$  und  $(B, e)$  Euklidische Räume über  $\mathbb{k}$ . Wenn es eine Isometrie  $F: A \rightarrow B$  gibt, kann man daraus folgern, dass  $F$  linear über  $\mathbb{R}$  ist. Der Beweis ist nicht ganz einfach und benutzt unter anderem (2).

Die Abbildung  $F$  muss jedoch nicht  $\mathbb{k}$ -linear sein, falls  $\mathbb{k} = \mathbb{C}$  oder  $\mathbb{H}$ . Aus diesem Grund ist es vom Standpunkt der metrischen Geometrie (also der Geometrie von Mengen  $M$  mit einer Metrik  $d$  wie in (1)) nicht besonders sinnvoll, Euklidische Räume über  $\mathbb{C}$  oder  $\mathbb{H}$  zu betrachten.

6.60. PROPOSITION. *Es seien  $(A, d_g)$  und  $(B, d_h)$  endlich-dimensionale Euklidische Räume der gleichen Dimension über  $\mathbb{k}$ -Vektorräumen  $(V, g)$  und  $(W, h)$  mit Skalarprodukten. Dann gibt es eine affine Isometrie  $F: A \rightarrow B$ .*

Mit anderen Worten ist die Dimension eine vollständige Invariante für endlich-dimensionale Euklidische Räume über einem festen Körper  $\mathbb{k}$  bis auf affine Isometrie, und  $(\mathbb{k}^n, d_g)$  ist eine zugehörige Normalform, wenn  $d_g$  die Euklidische Metrik zum Standard-Skalarprodukt bezeichnet. Im Falle  $\mathbb{k} = \mathbb{R}$  ist die Dimension wegen der obigen Bemerkung 6.59 (3) sogar eine vollständige Invariante endlich-dimensionaler Euklidischer Räume bis auf Isometrie.

BEWEIS. Wir wählen jeweils einen Ursprung  $o \in A$  und  $p \in B$  und identifizieren  $A$  und  $B$  mit den zugrundeliegenden  $\mathbb{k}$ -Vektorräumen  $(V, g)$  und  $(W, h)$  mit Skalarprodukten wie in Bemerkung 6.56 (2). Wegen Bemerkung 6.52 (1) gibt es eine lineare Isometrie  $L: V \rightarrow W$ . Dann definieren wir  $F: A \rightarrow B$  durch

$$F(a) = p + L(a - o).$$

Diese Abbildung ist eine affine Isometrie, denn für alle  $a, b \in A$  gilt

$$\begin{aligned} d_h(F(a), F(b)) &= \|F(a) - F(b)\|_h = \|p + L(a + o) - p - L(b + o)\|_h \\ &= \|L(a - b)\|_h = \|a - b\|_g = d_g(a, b). \quad \square \end{aligned}$$

6.61. BEMERKUNG. Die *Euklidische Gruppe* oder auch (*Euklidische*) *Bewegungsgruppe*  $E(n, \mathbb{k})$  ist die Gruppe der affinen Isometrien von  $(\mathbb{k}^n, d)$ , wobei  $d$  die Standardmetrik sei. Für  $\mathbb{k} = \mathbb{R}$  schreiben wir kurz  $E(n) = E(n, \mathbb{R})$ .

- (1) Nach Beispiel 6.55 (1) können wir jedes Element  $F \in E(n, \mathbb{k})$  schreiben als

$$v \mapsto w + Av \quad \text{mit } A \in M_n(\mathbb{k}).$$

Da  $F$  eine Isometrie ist, muss für alle  $v \in \mathbb{k}^n$  gelten, dass

$$\|Av\| = \|F(v) - F(0)\| = d(F(v), F(0)) = d(v, 0) = \|v\|.$$

Mit Hilfe der Polarisationsformeln aus Bemerkung 6.11 (2)–(4) folgt daraus  $\langle Au, Av \rangle = \langle u, v \rangle$  für alle  $u, v \in \mathbb{k}^n$ , so dass  $A \in U(n, \mathbb{k})$  mit  $U(n, \mathbb{k}) = O(n)$ ,  $U(n)$  beziehungsweise  $Sp(n)$ , je nachdem ob  $\mathbb{k} = \mathbb{R}, \mathbb{C}$  oder  $\mathbb{H}$ . Umgekehrt sieht man leicht, dass die obige Abbildung  $F$  eine affine Isometrie, also eine *Bewegung* ist, wenn  $A \in U(n, \mathbb{k})$  gilt.

- (2) Wir schreiben  $F = (w, A)$  für die obige Abbildung  $F$ . Wenn wir zwei solche Abbildungen  $F = (w, A)$  und  $G = (x, B)$  verketten, erhalten wir

$$(F \circ G)(v) = w + A(x + Bv) = (w + Ax) + ABv,$$

also gilt  $(w, A) \circ (x, B) = (w + Ax, AB)$ . Somit werden die Matrizen in den zweiten Einträgen der Paare multipliziert, während die Vektoren im ersten Eintrag erst addiert werden, nachdem der zweite Vektor von links mit der Matrix aus dem ersten Paar multipliziert wurde.

Das heißt, als Menge gilt  $E(n, \mathbb{k}) = \mathbb{k}^n \times U(n, \mathbb{k})$ , aber für die Verknüpfung  $\circ$  wird die Wirkung von  $U(n, \mathbb{k})$  auf  $\mathbb{k}^n$  benutzt. Man nennt daher  $E(n, \mathbb{k})$  das *semidirekte Produkt* von  $\mathbb{k}^n$  und  $U(n, \mathbb{k})$  und schreibt entsprechend

$$E(n) = \mathbb{R}^n \rtimes O(n),$$

$$E(n, \mathbb{C}) = \mathbb{C}^n \rtimes U(n)$$

$$\text{und } E(n, \mathbb{H}) = \mathbb{H}^n \rtimes Sp(n).$$

- (3) Für den Fall  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  können wir auch die Untergruppen

$$SE(n) = \mathbb{R}^n \rtimes SO(n) \subset E(n)$$

$$\text{und } SE(n, \mathbb{C}) = \mathbb{C}^n \rtimes SU(n) \subset E(n, \mathbb{C})$$

betrachten. Dann ist  $SE(n)$  die Gruppe der orientierungserhaltenden Bewegungen.

Wir wollen jetzt eine möglichst geometrische Beschreibung von affinen Isometrien geben. Zusammen mit den Normalformen für Isometrien auf Folgerung 6.50 können wir hieraus leicht eine Normalform für affine Isometrien herleiten.

6.62. SATZ. *Es sei  $(A, d)$  ein affiner Raum über einem endlich-dimensionalen  $\mathbb{k}$ -Vektorraum  $(V, g)$  mit Skalarprodukt. Es sei  $F: A \rightarrow A$  eine affine Isometrie über einer linearen Isometrie  $L$ . Dann existiert ein Punkt  $o \in A$  und ein Vektor  $x$  aus dem Eigenraum  $U$  von  $L$  zum Eigenwert 1, so dass*

$$F(a) = o + x + L(a - o).$$

*Der Vektor  $x \in U$  ist eindeutig durch  $A$  bestimmt. Der Punkt  $o$  kann beliebig gewählt werden aus einem affinen,  $F$ -invarianten Unterraum  $B \subset A$  über  $U$ , auf dem  $F$  durch Addition von  $x$  wirkt.*

Wir nennen  $B$  die *Achsenmenge* von  $F$ , und alle parallel affinen Geraden der Form

$$\{a + x \cdot r \mid r \in \mathbb{k}\} \subset B$$

mit  $a \in B$  heißen *Achsen* von  $F$ .

BEWEIS. Wir wählen zunächst einen beliebigen Punkt als Ursprung, identifizieren  $A$  mit  $V$  wie in Bemerkung 6.56 (2) und schreiben  $F(v) = y + L(v)$  wie in Beispiel 6.55 (1). Es sei  $U \subset V$  der Eigenraum zum Eigenwert 1 und  $W \subset V$  das orthogonale Komplement von  $U$ . Wie im Beweis von Satz 6.40 sind  $U$  und  $W$  invariant unter  $L$ , und  $L|_U = \text{id}_U$ .

Wir schreiben  $y = (x, z) \in U \oplus W = V$ . Die Abbildung  $\text{id}_W - L|_W$  ist invertierbar, denn 1 ist kein Eigenwert mehr von  $L|_W$ . Wir bestimmen  $q \in W$  so, dass  $q - L(q) = z$ . Dann setzen wir  $o = (p, q) \in U \oplus W \cong A$  für ein beliebiges  $p \in U$ . Für alle  $v = (u, w) \in U \oplus W = V$  folgt

$$\begin{aligned} (*) \quad F(o + v) &= (x, z) + L(p + u, q + w) = (x + p + u, z + L(q) + L(w)) \\ &= (x + p + u, q + L(w)) = o + (x, 0) + L(v). \end{aligned}$$

Wir wählen also  $o$  als unseren neuen Ursprung und haben die gesuchte Darstellung von  $F$  gefunden.

Da  $F$  eine affine Abbildung über  $L$  ist, ist  $L$  durch  $F$  eindeutig bestimmt. Wir betrachten  $o' = o + (p', q')$  als neuen Ursprung und  $v = (u, w)$  mit  $p', u \in U$  sowie  $q', w \in W$ . Dann betrachten wir den Vektor

$$\begin{aligned} x' &= F(o' + v) - o' - L(v) \\ &= F(o + (p' + u, q' + w)) - o - (p', q') - L(u, w) \\ &= (p' + u + x - p' - u, L(q' + w) - q' - L(w)) = (x, L(q') - q'). \end{aligned}$$

Dann gilt

$$L(x') = (x, (L \circ L)(q') - L(q')) = (x, L(q') - q') = x'$$

genau dann, wenn

$$(\text{id}_W - L|_W) \circ (\text{id}_W - L|_W)(q') = 0.$$

Nach Konstruktion ist  $\text{id}_W - L|_W$  invertierbar, also gilt das genau dann, wenn  $q' = 0$ , das heißt, wenn  $x' = x$  ist und  $o \in B$ . Damit ist die Eindeutigkeitsaussage bewiesen.  $\square$

Wir wollen mit Hilfe dieses Satzes Normalformen von Isometrien verstehen

6.63. BEISPIEL. Es sei  $A$  ein zweidimensionaler reeller Euklidischer Raum über einem zweidimensionalen Euklidischen Vektorraum  $(V, g)$  und  $F$  eine affine Isometrie von  $A$  über eine linearen Isometrie  $L$  von  $V$ . Es sei wieder  $U$  der Eigenraum von  $L$  zum Eigenwert 1. Wir stellen  $L$  wie in Folgerung 6.50 (1) dar und unterscheiden folgende Fälle.

- (1) Es sei  $F$  orientierungserhaltend.
  - (a) Es sei  $L = \text{id}_V$ , dann ist  $U = V$ , und  $B = A$  ist die Achsenmenge. Falls  $x = 0$  ist, ist  $F = \text{id}_A$  die Identität, ansonsten ist  $F(a) = a + x$  eine *Verschiebung*.
  - (b) Ansonsten ist  $L$  eine Drehung, also ist  $U = \{0\}$  und daher  $x = 0$ . Die Achsenmenge  $B$  besteht aus einem einzigen Punkt  $o$ , und  $F$  ist eine *Drehung* um  $o$ .
- (2) Wenn  $F$  nicht orientierungserhaltend ist, sind die Eigenräume zu den Eigenwerten  $\pm 1$  nach Bemerkung 6.53 (2) jeweils eindimensional, also ist die Achsenmenge  $B$  eine Gerade. Falls  $x = 0$ , ist  $F$  die *Spiegelung* an dieser Geraden, ansonsten eine *Gleitspiegelung*.

6.64. BEISPIEL. Sei  $A$  jetzt ein dreidimensionaler reeller Euklidischer Raum und  $V, F, L$  und  $U \subset V$  wie oben.

- (1) Es sei  $F$  orientierungserhaltend.
  - (a) Es sei  $L = \text{id}_V$ , dann ist  $U = V$ , und wie oben ist  $F$  entweder die *Identität* oder eine *Verschiebung*.
  - (b) Ansonsten ist  $L$  eine Drehung, und  $U$  ist eindimensional. Also ist die Achsenmenge  $B$  eine Gerade. Falls  $x = 0$ , ist  $F$  eine *Drehung* um die Gerade  $B$ , ansonsten eine *Schraubung*.
- (2) Wenn  $F$  orientierungsumkehrend ist, ist der Eigenraum von  $L$  zum Eigenwert 1 mindestens eindimensional.
  - (a) Wenn  $L$  einen zweidimensionalen Eigenraum zum Eigenwert 1 hat, ist die Achsenmenge  $B$  eine Ebene. In diesem Fall ist  $F$  eine *Spiegelung* an  $B$ , falls  $x = 0$ , ansonsten eine *Gleitspiegelung*.
  - (b) Wenn  $L$  in der Darstellung aus Folgerung 6.50 (1) durch einen Eigenwert  $-1$  und einen Drehblock beschrieben wird, erhalten wir eine *Drehspiegelung*. Die Achsenmenge enthält nur einen Punkt  $o$ . Dabei wird zunächst an einer Ebene durch  $o$  gespiegelt, anschließend um die Gerade durch  $o$  senrecht zu dieser Ebene gedreht.
  - (c) Einen Spezialfall davon erhalten wir, wenn der Eigenwert  $-1$  Multiplizität 3 hat. In diesem Fall enthält die Achsenmenge ebenfalls nur einen Punkt  $o$ , und  $F$  ist eine *Punktspiegelung* an  $o$ .

## 6.6. Bilinearformen und quadratische Funktionen

In diesem Abschnitt betrachten wir Hermitesche Sesquilinearformen, die nicht notwendig positiv definit sind. Ein Beispiel dafür ist die Lorentz-Metrik in der speziellen Relativitätstheorie.



dargestellt wird. Die Anzahlen der Einträge 1,  $-1$  und  $0$  sind gerade  $n_+(S)$ ,  $n_-(S)$  und  $n_0(S)$ .

Man nennt das Tripel  $(n_+, n_-, n_0)$  auch die *Signatur* der Hermiteschen Sesquilinearform. Wenn  $S$  nicht ausgeartet ist, heißt das Paar  $(n_+, n_-)$  oder auch die Differenz  $n_+ - n_-$  die Signatur von  $S$ . Die Signatur  $(n_+, n_-, n_0)$  bildet eine vollständige Invariante für  $\mathbb{k}$ -Vektorräume mit Hermitescher Sesquilinearform. Sei  $A$  die obige Matrix, dann ist  $\mathbb{k}^n$  mit  $S(x, y) = x^*Ay$  die zugehörige Normalform.

BEWEIS. Die Existenz der Basis  $B$  lässt sich auf zweierlei Weisen zeigen. Zunächst, indem man ein beliebiges Skalarprodukt  $g$  auf  $V$  wählt und dann Folgerung 6.45 anwendet. Anschließend ersetzt man die Basisvektoren  $v_i \neq \ker S$  durch  $e_i = v_i \cdot \frac{1}{\sqrt{|S(v_i, v_i)|}}$  (beachte, dass  $S(v_i, v_i) \in \mathbb{R}$ ). Bezüglich dieser Basis hat die Gramsche Matrix von  $S$  die gewünschte Gestalt. Dieser Beweis ist nicht konstruktiv, da Satz 6.40 und die anschließenden Folgerungen nicht konstruktiv sind.

Alternativ wählt man zunächst eine Basis  $e_{n-n_0+1}, \dots, e_n$  von  $\ker S$  mit dem Gauß-Verfahren und fixiert ein Komplement  $W \subset V$  vom Kern. Dann wählt man eine Basis  $v_1, \dots, v_{n-n_0}$  von  $W$ . Als nächstes konstruieren wir induktiv Vektoren  $e_1, \dots, e_{n-n_0}$  mit einem modifizierten Gram-Schmidt-Verfahren.

Seien dazu  $e_1, \dots, e_{p-1}$  bereits konstruiert und  $S(e_q, v_r) = 0$  für alle  $q < p$  und alle  $p \leq r \leq n - n_0$ . Falls  $S(v_p, v_p) = 0$ , existiert  $r > p$  mit  $S(v_p, v_r) \neq 0$ , da  $v_p \notin \ker S$ . Für  $t \in \mathbb{k}$  betrachte den Vektor  $v_p + v_r t$ , dann gilt

$$S(v_p + v_r t, v_p + v_r t) = 2 \operatorname{Re}(S(v_p, v_r) t) + |t|^2 S(v_r, v_r).$$

Für hinreichend kleine  $t \neq 0$  ist  $2|S(v_p, v_r) t| > |t|^2 S(v_r, v_r)$ . Wenn wir also für  $t$  ein kleines reelles Vielfaches von  $\overline{S(v_p, v_r)}$  wählen, folgt

$$S(v_p + v_r t, v_p + v_r t) \neq 0.$$

Wir ersetzen  $v_p$  durch  $v_p + v_r t$ , dann gilt nach wie vor  $S(e_q, v_p) = 0$  für alle  $q < p$ .

Da  $S(v_p, v_p) \in \mathbb{R} \setminus \{0\}$ , können wir jetzt

$$e_p = v_p \cdot \frac{1}{\sqrt{|S(v_p, v_p)|}}$$

definieren. Anschliessend machen wir die Vektoren  $v_{p+1}, \dots, v_{n-n_0}$  orthogonal zu  $e_p$  bezüglich  $S$ , indem wir  $v_r$  für alle  $r > p$  durch

$$v_r - e_p \cdot \underbrace{S(e_p, e_p)}_{=\pm 1} S(e_p, v_r)$$

ersetzen. Falls  $p < n - n_0$ , ersetzen wir  $p$  durch  $p + 1$  und machen weiter.

Zum Schluss sortieren wir die Basisvektoren so um, dass die Diagonaleinträge in der gewünschten Reihenfolge dastehen. Wir erhalten eine Diagonalmatrix  $A$  wie in (\*). Die Eindeutigkeit von  $n_0 = n_0(S) = \dim \ker S$  ist klar.

Zur Eindeutigkeit von  $n_+$  sei  $U \subset V$  ein positiver Unterraum. Falls  $n_+ < \dim U$ , finden wir aus Dimensionsgründen einen Vektor  $v \in V_+ = \langle e_1, \dots, e_{n_+} \rangle$  mit  $S(u, v) = 0$  für alle  $u \in U$ , also  $v \in U^\perp \cap V_+$ , insbesondere  $U \oplus \langle v \rangle$  positiv und  $U$  daher nicht maximal positiv.

Sei umgekehrt  $\dim U > n_+$ , dann betrachte  $V_- \oplus V_0 = \langle e_{n_++1}, \dots, e_n \rangle$ . Aus Dimensionsgründen existiert  $u \in U \cap (V_- \oplus V_0)$ , also gilt  $S(u, u) \leq 0$ , und  $U$  ist nicht positiv. Also hat ein maximaler positiver Unterraum gerade die Dimension  $n_+ = n_+(S)$ . Analog hat ein maximaler negativer Unterraum Dimension  $n_- = n_-(S)$ .  $\square$

6.68. BEMERKUNG. Es sei  $V$  ein  $n$ -dimensionaler  $\mathbb{k}$ -Vektorraum mit einer Hermiteschen Sesquilinearform  $S$ . Nach Sylvesters Trägheitssatz 6.67 dürfen wir  $V = \mathbb{k}^n$  annehmen, wobei  $S$  durch die obige Diagonalmatrix (\*) gegeben wird. Es sei  $p = n_+(S)$  und  $q = n_-(S)$ . Wir interessieren uns für die Untergruppe der Automorphismengruppe  $GL(n, \mathbb{k})$ , die die Form  $S$  erhalten, also

$$G = \{ F \in GL(n, \mathbb{k}) \mid F^* S = S \} .$$

- (1) Falls  $p + q = n$  gilt, ist  $S$  nicht ausgeartet. In diesem Fall heißt die entsprechende Gruppe  $U(p, q; \mathbb{k})$ , beziehungsweise

$$\begin{aligned} O(p, q) &= U(p, q; \mathbb{R}) , \\ U(p, q) &= U(p, q; \mathbb{C}) \\ \text{und} \quad Sp(p, q) &= U(p, q; \mathbb{H}) . \end{aligned}$$

Im Falle  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  haben wir Determinanten zur Verfügung und definieren

$$\begin{aligned} SO(p, q) &= O(p, q) \cap SL(n, \mathbb{R}) \\ \text{und} \quad SU(p, q) &= U(p, q) \cap SL(n, \mathbb{C}) . \end{aligned}$$

Es besteht eine gewisse formale Analogie zu den Gruppen aus Bemerkung 6.51, beispielsweise gilt

$$\begin{aligned} SO(1, 1) &= \left\{ \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix} \mid t \in \mathbb{R} \right\} , \\ SO(2) &= \left\{ \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \mid t \in \mathbb{R} \right\} . \end{aligned}$$

Für diese Gruppen ist jedoch das Analogon zu Folgerung 6.50 im Allgemeinen nicht mehr richtig. Dazu betrachten wir für  $0 \neq t \in \mathbb{R}$  die Matrix

$$A = \begin{pmatrix} 1 + ti & t \\ t & 1 - ti \end{pmatrix} \in M_2(\mathbb{C}) .$$

Man rechnet nach, dass

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



Zum Beweis gehen wir analog vor wie im zweiten Beweis des Trägheitssatzes 6.67. Beim Normieren ersetzen wir allerdings einen Vektor  $v_p$  mit  $B(v_p, v_p) \neq 0$  durch

$$e_p = v_p \cdot \frac{1}{\sqrt{B(v_p, v_p)}},$$

so dass jetzt stets  $B(e_p, e_p) = 1$  gilt.

- (3) Da  $\mathbb{H}$  nicht kommutativ ist, können wir keine Bilinearformen über  $\mathbb{H}$  definieren.

6.71. DEFINITION. Es sei  $\mathbb{k}$  ein Körper und  $V$  ein  $\mathbb{k}$ -Vektorraum. Eine Abbildung  $q: V \rightarrow \mathbb{k}$  heißt *quadratische Funktion*, wenn eine symmetrische Bilinearform  $B$  auf  $V$ , eine Linearform  $\alpha \in V^*$  und eine Konstante  $c \in \mathbb{k}$  existieren, so dass

$$q(v) = B(v, v) + \alpha(v) + c.$$

Wir nennen  $q$  *nicht ausgeartet*, wenn  $B$  nicht ausgeartet ist. Die Nullstellenmenge  $Q$  einer quadratischen Funktion heißt auch *Quadrik* oder *Hyperfläche zweiten Grades*.

Eine quadratische Funktion  $q$  ist so etwas wie ein Polynom vom Grad  $\leq 2$  in einer Variablen aus dem Vektorraum  $V$ , und die Quadrik  $Q = q^{-1}(0)$  ist ihre Nullstellenmenge. In der Analysis lernt man im Fall  $\mathbb{k} = \mathbb{R}$ , dass die Nullstellenmenge eine glatte Hyperfläche ist, wenn 0 ein regulärer Wert von  $q$  ist. Man beachte, dass es entartete Fälle gibt, in denen  $Q$  nicht glatt oder noch nicht einmal eine Hyperfläche ist, siehe Beispiele 6.74, 6.75. In diesen Fällen ist 0 kein regulärer Wert von  $q$ .

Im Folgenden müssen wir durch 2 teilen können, daher erinnern wir uns an die Charakteristik  $\chi(\mathbb{k})$  eines Körpers aus Definition 2.14.

6.72. SATZ. *Es sei  $\mathbb{k}$  ein Körper der Charakteristik  $\chi(\mathbb{k}) \neq 2$ , es sei  $V$  ein  $\mathbb{k}$ -Vektorraum und  $q$  quadratische Funktion auf  $V$ . Dann existiert eine invertierbare affine Abbildung  $F: V \rightarrow V$ , eine symmetrische Bilinearform  $B$  auf  $V$ , ein Komplement  $W$  von  $U = \ker B$ , eine Linearform  $\alpha \in U^*$  und  $c \in \mathbb{k}$ , so dass*

$$(q \circ F)(u + w) = B(w, w) + \alpha(u) + c \quad \text{für alle } u \in \ker S \text{ und } w \in W.$$

BEWEIS. Sei  $q(v) = B(v, v) + \beta(v) + b$  für eine symmetrische Bilinearform  $B$ , eine Linearform  $\beta \in V^*$  und eine Konstante  $b \in \mathbb{k}$ . Wir wählen ein Komplement  $W$  von  $U = \ker B$ , so dass  $V = U \oplus W$ . Wir definieren  $\alpha \in U^*$  und  $\gamma \in W^*$  durch

$$\beta(u + w) = \alpha(u) + \gamma(w) \quad \text{für alle } u \in \ker B \text{ und } w \in W.$$

Ähnlich wie in Proposition 6.22 fassen wir  $B|_W$  als linearen Isomorphismus  $B: W \rightarrow W^*$  mit  $B(w) = B(w, \cdot)$  auf. Dann existiert  $x = B^{-1}(\gamma) \in W$

mit  $2B(x, w) = \gamma(w) = \beta(w)$  für alle  $w \in W$ . Es sei  $F: V \rightarrow V$  die Verschiebung  $F(v) = v - x$ . Für  $v = u + w$  mit  $u \in U$  und  $w \in W$  gilt dann

$$\begin{aligned}(q \circ F)(v) &= B(w - x, w - x) + \alpha(u) + \gamma(w - x) + b \\ &= B(w, w) + \alpha(u) - 2B(x, w) + \gamma(w) + b + B(x, x) - \gamma(x) \\ &= B(w, w) + \alpha(u) + c\end{aligned}$$

mit  $c = b + B(x, x) - \gamma(x)$ . □

Im Beweis haben wir als affine Abbildung also nur eine Verschiebung gewählt, um eine quadratische Ergänzung durchzuführen. Im Falle  $\mathbb{k} = \mathbb{R}$  oder  $\mathbb{C}$  und  $V = \mathbb{k}^n$  würden wir zusätzlich noch einen linearen Isomorphismus dazuschalten, so dass die Form  $B$  auf  $\mathbb{k}^n$  durch eine der speziellen Formen aus Bemerkung 6.70 (1) oder (2) dargestellt wird, und so dass entweder  $\alpha = 0$  oder  $\alpha = \varepsilon_n$  gilt.

Um die Gestalt von  $Q = q^{-1}(0) \subset V$  im Falle  $\mathbb{k} = \mathbb{R}$  darzustellen, nehmen wir an, dass  $B$  tatsächlich durch die Matrix (\*) wie im Trägheitssatz 6.67 von Sylvester dargestellt wird und die Signatur durch das Tripel  $(n_+, n_-, n_0)$  gegeben ist. Außerdem definieren wir noch folgende Mengen:

$$\begin{aligned}V^+ &= \langle e_1, \dots, e_{n_+} \rangle, \\ V^- &= \langle e_{n_++1}, \dots, e_{n_++n_-} \rangle, \\ S^+ &= \{ v \in V^+ \mid B(v, v) = 1 \}, \\ S^- &= \{ v \in V^- \mid B(v, v) = -1 \}, \\ \text{und } U' &= \langle e_{n_++n_-+1}, \dots, e_{n-1} \rangle,\end{aligned}$$

dann sind  $S^+$ ,  $S^-$  gerade die „Einheitssphären“ im positiven beziehungsweise im negativen Unterraum, und  $U' = \ker \alpha \subset U = \ker B$  falls  $\alpha \neq 0$ .

**6.73. FOLGERUNG.** *Es sei  $V$  ein endlich-dimensionaler reeller Vektorraum, es sei  $B$  eine symmetrische Bilinearform auf  $V$ , und  $W = V^+ \oplus V^- \subset V$  ein Komplement von  $U = \ker S$ . Es seien  $\alpha \in U^*$ ,  $c \in \mathbb{R}$  und*

$$q(u, w) = B(w, w) + \alpha(u) + c \quad \text{für alle } u \in \ker S \text{ und } w \in W.$$

Dann hat  $Q = q^{-1}(0)$  eine der folgenden Gestalten.

(1) Falls  $\alpha = 0$ , gilt  $Q = U \times Q'$ , und

(a) falls  $c = 0 \dots$

(i) und  $n_+ = 0$  oder  $n_- = 0$ , ist  $Q' = \{0\}$ ,

(ii) und  $n_+, n_- \geq 1$ , ist  $Q'$  ein Doppelkegel

$$Q' = \{ (v_+, v_-) \mid v_+ \in S^+, v_- \in S^-, \text{ und } 0 \leq r \in \mathbb{R} \};$$

(b) falls  $c > 0 \dots$

(i) und  $n_- = 0$ , ist  $Q' = \emptyset$ ,

(ii) und  $n_- \neq 0$ , wird  $Q'$  durch  $V^+ \times S^-$  parametrisiert, wobei

$$Q' = \{ (v_+, v_- \sqrt{c + B(v_+, v_+)}) \mid v_+ \in V^+ \text{ und } v_- \in S^- \};$$

- (c) falls  $c < 0$  hat  $Q'$  eine entsprechende Gestalt wie in (1.b), aber mit den Rollen von  $V^+$  und  $V^-$  vertauscht.
- (2) Falls  $\alpha \neq 0$ :  $Q = \ker \alpha \times \Gamma$ , dabei ist  $\Gamma$  der Graph der nicht-ausgearteten quadratischen Funktion

$$w \mapsto -(B(w, w) + c)$$

über  $W = V_+ \oplus V_-$ .

BEWEIS. Man überzeugt sich, dass die Fallunterscheidung in der Folgerung vollständig ist. Es reicht also, Fall für Fall zu betrachten. Wir betrachten auf  $V^\pm$  die Norm  $\|v_\pm\| = \sqrt{\pm B(v_\pm, v_\pm)}$ .

Im Fall (1) hängt  $q(u, w)$  nicht von  $u$  ab, also sei

$$Q' = \{ w \in W \mid q(0, w) = 0 \},$$

dann gilt  $(u, w) \in Q$  genau dann, wenn  $w \in Q'$ , also gilt  $Q = U \times Q'$ . Ab sofort betrachten wir also nur noch die nicht-ausgeartete quadratische Form

$$q'(w) = q|_W(w) = B(w, w) + c$$

auf  $W$ .

Im Fall (1.a) ist  $c = 0$ . Falls (1.a.i) mit  $n_- = 0$  vorliegt, folgt  $B(w, w) \geq 0$ , und  $q'(w) = B(w, w) = 0$  gilt genau dann, wenn  $w = 0$  ( $B|_{W \times W}$  ist also positiv definit). Analoges gilt für  $-q'$ , falls  $n_+ = 0$  gilt.

Im Fall (1.a.ii) sei  $w = (v_+ r, v_- r) \in V^+ \oplus V^- = W$  mit  $v_\pm \in S^\pm$  und  $r \geq 0$ , dann folgt

$$q'(w) = \|v_+\|^2 r^2 - \|v_-\|^2 r^2 = 0,$$

da  $\|v_+\|^2 = \|v_-\|^2 = 1$  nach Annahme, also  $(v_+ r, v_- r) \in Q'$ . Sei umgekehrt  $w = (w_+, w_-) \in Q' \subset V^+ \oplus V^-$ , dann folgt

$$0 = q'(w) = \|w_+\|^2 - \|w_-\|^2,$$

also dürfen wir  $r = \|w_+\| = \|w_-\|$  setzen. Falls  $w_+ = w_- = 0$ , dürfen wir  $v_\pm \in S^\pm$  beliebig wählen; das geht, da  $S^\pm \neq \emptyset$  falls  $n_\pm \geq 1$ . Andernfalls setzen wir  $v_\pm = w_\pm \frac{1}{n} \in S^\pm$  und erhalten  $w = (v_+ r, v_- r)$  wie oben.

Im Fall (1.b) ist  $c > 0$ . Im Fall (1.b.i) folgt  $q'(w) > 0$  für alle  $w \in W$ , also  $Q' = \emptyset$ .

Im Fall (1.b.ii) gilt entsprechend  $w_- \neq 0$  für alle  $w = (w_+, w_-) \in Q'$ , und es folgt

$$\|w_-\|^2 = \|w_+\|^2 + c,$$

also erhalten wir für jeden Vektor  $v_+ \in V^+$  und jede Richtung  $v_- \in V^-$  eine eindeutige Lösung  $(v_+, v_- r) \in Q'$  mit

$$r = \sqrt{c + \|v_+\|^2}.$$

Im Fall (1.c) ersetzen wir  $q'$  durch  $-q'$  und machen wie in (1.b) weiter. Dabei tauschen  $V^+$  und  $V^-$  ihre Rollen.

Im Fall (2) gilt  $n_0 \geq 1$ , und wir dürfen wie oben gesagt annehmen, dass  $\alpha = \varepsilon^n$ . Für einen Vektor

$$v = (v_+, v_-, u', e_n h) \in V^+ \oplus V^- \oplus U' \oplus \langle e_n \rangle$$

gilt also  $q(v) = 0$  genau dann, wenn

$$h = -B(w, w) - c = \|v_-\|^2 - \|v_+\|^2 - c.$$

Für jede Wahl von  $(v_+, v_-, u')$  gibt es also genau eine Zahl  $h \in \mathbb{R}$ , so dass  $(v_+, v_-, u', e_n h) \in Q$ , und  $h$  hängt nicht von  $u' \in U' = \ker \alpha$  ab. Also hat  $Q$  die angegebene Gestalt.  $\square$

6.74. BEISPIEL. Es sei  $Q \subset \mathbb{R}^2$  eine Quadrik. Es sei  $q: \mathbb{R}^2 \rightarrow \mathbb{R}$  eine quadratische Funktion in der obigen Normalform. Wir schreiben  $v = (x, y) \in \mathbb{R}^2$ . Wir geben im jeden einzelnen der Fälle aus Folgerung 6.73 die Gestalt von  $Q$  an.

Im Fall (1.a.i) ist  $Q' = \{0\}$  ein Punkt. Falls  $n_0 = 0$ , ist auch  $Q$  ein Punkt. Andernfalls erhalten wir eine Gerade  $Q = Q' \times \mathbb{R}$ , falls  $n_0 = 1$ , oder den gesamten  $\mathbb{R}^2 = Q' \times \mathbb{R}^2$ , falls  $n_0 = 2$ .

Im Fall (1.a.ii) folgt  $n_+ = n_- = 1$  und  $n_0 = 0$ . Die Quadrik  $Q$  besteht aus den beiden Geraden  $y = x$  und  $y = -x$ .

Im Fall (1.b.i) ist  $Q = Q' = \emptyset$ .

Im Fall (1.b.ii) gibt es drei Möglichkeiten. Falls  $n_- = 2$  und  $n_+ = n_0 = 0$ , ist

$$Q = \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = c \}$$

ein Kreis. Falls  $n_- = 1 = n_+$  und  $n_0 = 0$ , besteht

$$Q = \{ (x, y) \mid y = \pm \sqrt{c + x^2} \}$$

aus den zwei Ästen einer Hyperbel. Falls  $n_- = 1 = n_0$  und  $n_+ = 0$ , besteht  $Q$  nur aus den zwei Geraden  $y = \pm \sqrt{c}$ , da  $S^0$  nur aus den zwei Punkten  $\pm 1 \in \mathbb{R}$  besteht.

Der Fall (1.c) liefert die gleichen geometrischen Figuren wie (1.b).

Im Fall (2) sei  $\alpha(x, y) = y$ , so dass  $Q$  der Graph einer quadratischen Funktion  $q': \mathbb{R} \rightarrow \mathbb{R}$  ist. Wir unterscheiden drei Fälle. Falls  $n_0 = 2$ , ist  $q'$  konstant, und  $Q$  eine zur  $x$ -Achse parallele Gerade. Falls  $n_+ = 1 = n_0$  und  $n_- = 0$ , ist

$$Q = \{ (x, y) \mid y = -c - x^2 \}$$

eine nach unten offene Parabel. Fall  $n_- = 1 = n_0$  und  $n_+ = 0$ , ist  $Q$  entsprechend eine nach oben offene Parabel.

Man nennt alle diese Figuren auch Kegelschnitte, da sich die meisten (alle bis auf die leere Menge, den gesamten  $\mathbb{R}^2$  und die zwei parallelen Geraden) als Schnitt eines Doppelkegels im  $\mathbb{R}^3$  mit einer Ebene darstellen lassen. Man erhält umgekehrt jede Quadrik im  $\mathbb{R}^2$  aus einem der obigen Beispiele durch eine invertierbare affine Abbildung. Wenn diese Abbildung keine affine Isometrie ist, kann sich das dadurch bemerkbar machen, dass aus dem runden Kreis eine Ellipse, aus der Hyperbel mit rechtem Winkel zwischen den Asymptoten eine

Hyperbel mit einem anderen Asymptotenwinkel, aus der Einheitsparabel eine Parabel anderer Größe, und aus zwei sich rechtwinklig schneidenden Geraden zwei sich unter einem beliebigen Winkel  $\neq 0$  schneidende Geraden werden.

6.75. BEISPIEL. Wir betrachten zum Schluss Quadriken im  $\mathbb{R}^3$ . Dabei listen wir aber nur noch die verschiedenen auftretenden Formen und in Klammern die Tripel  $(n_+, n_-, n_0)$  auf.

Im Fall (1.a.i) erhalten wir einen Punkt  $((3,0,0)$  oder  $(0,3,0))$ , eine Gerade  $((2,0,1)$  oder  $(0,2,1))$ , eine Ebene  $((1,0,2)$  oder  $(0,1,2))$  oder den gesamten  $\mathbb{R}^3$   $((0,0,3))$ .

Im Fall (1.a.ii) erhalten wir einen Doppelkegel  $((2,1,0)$  oder  $(1,2,0))$  oder zwei sich schneidende Ebenen  $((1,1,1))$ .

Im Fall (1.b.i) erhalten wir die leere Menge  $((3,0,0)$ ,  $(2,0,1)$ ,  $(1,0,2)$  oder  $(0,0,3))$ .

Im Fall (1.b.ii) erhalten wir eine Kugel  $((0,3,0))$  ein einschaliges Rotationshyperboloid  $((1,2,0))$  einen Zylinder, also das Produkt aus einem Kreis und einer Geraden  $((0,2,1))$ , ein zweischaliges Rotationshyperboloid  $((2,1,0))$ , das Produkt aus einer Hyperbel und einer Geraden  $((1,1,1))$  oder zwei parallele Ebenen  $((0,1,2))$ .

Der Fall (1.c) liefert wieder die gleichen Flächen wie (1.b).

Im Fall (2) erhalten wir ein Rotationsparaboloid  $((2,0,1)$  oder  $(0,2,1))$ , ein hyperbolisches Paraboloid  $((1,1,1))$ , ein Produkt aus einer Parabel und einer Geraden  $((1,0,2)$  oder  $(0,1,2))$ , oder eine Ebene  $((0,0,3))$ .

Allgemeine Quadriken im  $\mathbb{R}^3$  entstehen aus den obigen durch invertierbare affine Abbildungen. Wenn wir nur affine Isometrien zulassen wollen, können wir die zugrundeliegende symmetrische Bilinearform  $B$  nicht auf die Normalform aus dem Sylvesterschen Trägheitssatz 6.67 bringen, aber wegen Folgerung 6.45 immerhin auf Diagonalgestalt. Hieraus folgt zum Beispiel, das ein Ellipsoid immer drei aufeinander senkrechte Hauptachsen hat, also bis auf eine affine Isometrie von der folgenden Form ist:

$$Q = \{ (x, y, z) \in \mathbb{R}^3 \mid ax^2 + by^2 + cz^2 = 1 \} \quad \text{mit } a, b, c > 0.$$

Dieser geometrische Sachverhalt ist ein weiterer Grund, das Hauptergebnis aus Abschnitt 6.4 „Hauptachsentransformation“ zu nennen.



## Notation

$\in$ , 3 $\{\dots\}$ , 4 $\emptyset$ , 4 $\subset$ , 5 $\subsetneq$ , 5 $\cap$ , 5 $\cup$ , 5 $\setminus$ , 6 $\times$ (Mengen), 6 $(\dots)$ , 6 $\mathcal{P}$ , 6 $\{\dots   \dots\}$ , 6 $F: M \rightarrow N$ , 6 Abb, 7 $\text{im}$ , 7 $F^{-1}$ , 7 $\text{id}$ , 7 $\circ$ , 7 $F _U$ , 8 $\mathbb{N}$ , 10 $\underline{n}$ , 10 $\overline{\mathbb{N}}$ , 10 $\#$ , 11 $\leq$ , 11 $\mathbb{Z}$ , 18 $\mathbb{Q}$ , 20 $\mathbb{R}$ , 21 $\mathbb{R}^n$ , 21 $\langle \cdot, \cdot \rangle$ , 22 $\ \cdot\ $ , 22 $\angle$ , 22 $i$ , 24 $\mathbb{C}$ , 24	$\text{Re}$ , 25 $\text{Im}$ , 25 $\bar{\cdot}$ , 25 $ \cdot $ , 26 $\arg$ , 27 $\times$ (Vektoren), 29 $\mathbb{H}$ , 31 $j, k$ , 33 $\text{Aut}$ , 39 $\equiv \text{ mod }$ , 41 $\mathbb{Z}/n$ , 41 $\mathbb{k}^\times$ , 44 $a   n$ , 45 $\text{ggT}$ , 46 $\text{Hom}_R, {}_R \text{Hom}$ , 50 $\text{Iso}_R, {}_R \text{Iso}$ , 52 $\text{End}_R, {}_R \text{End}$ , 52 $\text{Aut}_R, {}_R \text{Aut}$ , 52 $M^*, {}^*M$ , 53 $\ker$ , 55 $U + V$ , 59 $U \oplus V$ , 59 $M^I$ , 62 $R^n$ , 62 $\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$ , 62 $e_1, \dots, e_n$ , 62 $\delta_{ij}$ , 62 ${}^n R$ , 63 $(r_1, \dots, r_n)$ , 63 $\varepsilon_1, \dots, \varepsilon_n$ , 63 $\sum_{i=1}^n$ , 63 $\exists!$ , 64
---	--

- $\langle A \rangle$ , 66  
 $\begin{pmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & & \vdots \\ f_{m1} & \cdots & f_{mn} \end{pmatrix}$ , 69  
 $M_{m,n}(R)$ , 69  
 $M_n(R)$ , 72  
 $E_n$ , 72  
 $F^{-1}$ , 72  
 $GL(n, R)$ , 72  
 $Bv$ , 73  
 $Bf_C$ , 74  
 $B \text{id}_C$ , 74  
 $R^{(I)}$ , 76  
 $R[X]$ , 76  
 $R[[X]]$ , 76  
 ${}^I R$ , 77  
 $\dim$ , 82  
 $A^t$ , 86  
 $\text{rg}$ , 86  
 $\text{rg}_S, \text{rg}_Z$ , 86  
 $a_0 + U$ , 88  
 $P_{ij}$ , 91  
 $M_i(k)$ , 91  
 $E_{ij}(k)$ , 91  
 $A^*$ , 96  
 $\text{vol}$ , 101  
 $\Lambda^k M^*$ , 103  
 $S_n$ , 105  
 $\text{sign}$ , 105  
 $F^*$ , 108  
 $\det$ , 109  
 $R^\times$ , 115  
 $O(n), SO(n)$ , 119  
 $GL(n, \mathbb{R})^+$ , 119  
 $SL(n, \mathbb{R})$ , 119  
 $V_\lambda$ , 121  
 $\chi_A(X), \chi_F(X)$ , 129  
 $\sigma_i(A), \sigma_i(F)$ , 130  
 $\text{deg}$ , 136  
 $\max$ , 136  
 $Q | P, Q \nmid P$ , 138  
 $\mu_F, \mu_A$ , 139  
 $\text{ord}_p(n), \text{ord}_P(Q)$ , 142  
 $\text{ord}_\lambda$ , 143  
 $m.$ , 145  
 $\langle \cdot, \cdot \rangle_{L^2}$ , 167  
 $\langle \cdot, \cdot \rangle_{H^1}$ , 168  
 $\| \cdot \|_g$ , 168  
 $df$ , 180  
 $\text{grad } f$ , 181  
 $\overline{V}^*$ , 182  
 $F^*$ , 183  
 $U(n), SU(n)$ , 196  
 $Sp(n)$ , 196  
 $d, d_g$ , 199  
 $E(n), E(n, \mathbb{k})$ , 202  
 $U(n, \mathbb{k})$ , 202  
 $\rtimes$ , 202  
 $SE(n), SE(n, \mathbb{C})$ , 202  
 $U(p, q; \mathbb{k})$ , 207  
 $O(p, q), SO(p, q)$ , 207  
 $U(p, q), SU(p, q)$ , 207  
 $Sp(p, q)$ , 207

## Stichwortverzeichnis

- Abbildung, 6
  - adjungierbare, 183
  - adjungierte, 183
  - affine, 198
  - Basis-, 67, 73
  - induzierte, 15
  - Koordinaten-, 67, 73, 173, 179
  - lineare, 48
    - anti-, 164
    - semi-, 164
  - multilineare, 103
  - normale, 188, 189–191
  - Null-, 51
  - Projektion, 178
  - Quotienten-, 15, 56
  - schiefe, 188, 195
  - selbstadjungierte, 188, 193
  - Umkehr-, 9, 39
- Abbildungsmatrix, 74, 85, 123
- abgeschlossen, 54
- Ableitung, 51
  - partielle, 180
  - Richtungs-, 180
  - totale, 180
- Absolutbetrag, 22, 164
  - auf  $\mathbb{C}$ , 26
- Achse, 203
- Addition, 13, 40
  - Matrix-, 69, 70
  - Vektor-, 21, 27
- additiv, 48, 133, 164
- Adjunkte, 115, 135
- äquivalent
  - metrisch, 197
- Äquivalenzklasse, 15, 20
- Äquivalenzrelation, 9, 14, 16, 20, 41, 88
- Algebra, 132
  - assoziative, 132
  - kommutative, 132
  - mit Eins, 132
- Algorithmus
  - Euklidischer, 46
  - Polynome, 140
  - Gauß-Verfahren, 92, 94–96
  - Gram-Schmidt-
    - Orthonormalisierungsverfahren, 174
- alternierend, 103
- antisymmetrisch, 29, 103
- Arcustangens, 159
- Argument, 27
- Assoziativgesetz, 14, 18, 20, 24, 37, 40
- Automorphismus, 39
  - Anti-, 32, 164
  - isometrischer, 188, 195
  - Modul-, 52
  - orientierungserhaltender, 118
  - orientierungsumkehrender, 118
  - unitärer, 188, 195
  - Vektorraum-, 52
  - volumenerhaltender, 119
- Axiome
  - Äquivalenzrelation, 14
  - Gruppe, 37
  - Homomorphismus, 48
    - unitärer Ring-, 133
  - Körper, 42
  - Lineare Abbildung, 48
  - Mengenlehre, 5
  - Metrik, 200
  - Modul, 47
  - Norm, 169
  - Ordnung, 11
  - Peano- für  $\mathbb{N}$ , 10
  - Ring, 40
  - Skalarprodukt, 165
  - Untermodul, 54
  - Vektorraum, 47
- Basis, 66, 77, 80–82
  - angeordnete, 66
  - duale, 68, 179, 185
  - Orthogonal-, 173
  - Orthonormal-, 97, 173
  - Standard-, 62, 76, 118

- unitäre, 97, 173
  - quaternionisch, 97, 173
- basisunabhängig, 110
- Basiswechsel, 74, 117, 172
- Bewegung, 202
  - orientierungserhaltende, 202
- Beweis
  - indirekter, 4
- bijektiv, 7, 8
- Bild, 7, 55, 89, 95
- bilinear, 132, 165
- Block
  - Jordan-, 152
- Charakteristik, 44
- Chinesischer Restsatz, **146**
- Cramersche Regel, **115**, 135
- definit
  - negativ, 176
  - positiv, 22, 165, 171
  - semi-
    - positiv, 165, 171
- Definition
  - rekursive, 10, 12, 63
- Definitionsbereich, 7
- Determinante, 109, 110–119, 130
- Determinantenfunktion, 103, 117
  - Standard-, 105, 107, 118
- diagonalisierbar, 122, 126, 157
- Diagonalmatrix, 122
- Differentialgleichung
  - lineare, 131, 155
- Differentialgleichungssystem
  - gewöhnliches lineares, 160
- Differenz
  - von Mengen, 6
- Dimension, 82
- Dimensionsformel
  - Komplement, 83
  - lineare Abbildung, 85
  - Summe, 84
- disjunkt, 5
- Distributivgesetz, 14, 18, 20, 24, 40, 47
- Divergenz, 158
- Division
  - mit Rest, 41, 45, 137
  - Polynom-, 137
- Doppelkegel, 210
- Drehung, 33, 35, 204
- Dreiecksgestalt, 113
  - strikte, 113
- Dreiecksungleichung, 169, 200
- dual, 61
- Durchschnitt, 5
- Eigenraum, 121
- Eigenschaft
  - universelle
    - freier Modul, 64, 67, **77**
    - Koprodukt, 60
    - Polynomring, **134**
    - Produkt, 61
    - Quotient, 15, 57
- Eigenvektor, 121
- Eigenwert, 121, 147
  - komplexer, 159
  - verallgemeinerte, 195
- Einbettung
  - isometrische, 187
    - affine, 199
- Einheit, 116, 140
- Eins, 40
- Einschränkung, 8
- einsetzen, 127
- Element, 3
  - Eins-, 40
  - inverses, 19, 20, 37, 38
  - irreduzibles, 141, 143
  - neutrales, 14, 18, 20, 24, 37, 38, 40, 54
  - Null-, 40, 48
  - Prim-, 141, 143
- endlich, 11
- endlich erzeugt, 66
- endlichdimensional, 82, 88
- Endomorphismus
  - diagonalisierbarer, 122, 126, 157, 188
  - Modul-, 52
  - nilpotenter, 153
  - normaler, 188, 189–191
  - schiefer, 188, 195
  - selbstadjungierter, 188, 193
  - trigonalisierbarer, 123, 153
  - Vektorraum-, 52
- Erzeugendensystem, 66, 77, 80–82
- Erzeugnis, 66, 77
- Euklidischer Algorithmus, **46**
  - Polynome, 140
- Faktor
  - Linear-, 143
- Familie, 76
  - endliche, 76
- Federpendel, 131
  - gedämpftes, 155
- Form
  - alternierende, 103
  - Bilinear-, 165, 208
    - symmetrische, 208
  - Hermiteische, 165
    - ausgeartete, 205

- positive, 165
- Linear-, 179
  - darstellbare, 180
- Sesquilinear-, 165
  - zurückgeholte, 186
- zurückgeholte, 108
- Fundamentallösung, 161
- Funktion
  - elementarsymmetrische, 130
  - Grad-, 136, 140, 142
  - quadratische, 209
- Gauß-Verfahren, **92**, 94–96, 114
- Gleichheit
  - Abbildungen, 7
- gleichmächtig, 9
- Gleichungssystem
  - Differential-, 160
  - lineares, 88–96
    - homogenes, 89
    - inhomogenes, 89
- Grad
  - Polynom, 136, 140, 142
- Gradient, 181
- Gram-Schmidt-
  - Orthonormalisierungsverfahren, **174**
  - modifiziertes, 206
- Graph, 6
- Gruppe, 37
  - abelsche, 37, 40
  - additive, 38, 40
  - Automorphismen-
    - Modul, 53
  - Bewegungs-, 202
  - Einheiten-, 116, 140
  - Euklidische, 202
  - lineare
    - allgemeine, 72, 75, 119
    - spezielle, 119
  - multiplikative, 44, 116, 140
  - orthogonale, 119, 196
    - spezielle, 119, 196
  - symmetrische, 105
  - symplektische
    - kompakte, 196
  - unitäre, 196
    - spezielle, 196
  - Unter-, 55
  - zyklische
    - Ordnung  $n$ , 42
    - unendliche, 38
- Halbordnung, 11
- Hauptraum, 147
- Hermiteisch, 165, 171
- homogen, 48, 169
  - anti-, 164
  - positiv, 101
- Homomorphismus
  - Algebren-, 133
    - unitärer, 133
  - Modul-, 48
  - Ring-, 133
  - Vektorraum-, 50
- Hyperfläche
  - zweiten Grades, 209
- Ideal, 139
- Identität, 7, 39
  - Bézout-, 46, 140, 141, 145
  - Graßmann-, 29, 34
  - Jacobi-, 29
  - Parallelogramm-, 170
- Imaginärteil, 25, 31
- Index, 205
- Induktion
  - vollständige, 12
- injektiv, 7, 8
- Inklusion, 8
- Invariante
  - vollständige, 196, 206
- Inverses
  - additives, 19, 24
  - multiplikatives, 19, 21, 24
- irreduzibel, 141, 143
- Isometrie, 201
  - affine, 199
  - der Ebene, 28
  - des Raumes, 35
  - lineare, 119, 188, 195, 196
  - orientierungserhaltende, 119, 196
- isomorph, 82
- Isomorphismus
  - Modul-, 52
  - Vektorraum-, 52
- Jordan-Block, 152
- Kegelschnitt, 212
- Kern, 55, 89, 95, 205
- Klassifikation, 82
- Körper, 42
  - angeordneter, 21
    - archimedisch, 21
    - vollständig, 21
  - endlicher, 45, 144
  - Schief-, 42
  - Teil-, 55
- Kommutativgesetz, 14, 18, 20, 24, 37, 40
- Komplement, 6, 59, 83
- kongruent, 41

- Konjugation, 164
  - komplexe, 25
  - quaternionische, 31
- Konvergenz, 158
- Koordinaten, 67, 179
- Krümmung
  - Haupt-, 193
- Kürzungsregel, 14, 19, 38, 44
  
- Laplace-Entwicklung, 105, **112**, 114–116
- Leibniz-Formel, **110**, 111, 115, 117, 130
- Leitkoeffizient, 136
- Lemma
  - Euklid, **141**, 142
- linear, 22, 29, 48
  - anti-, 164, 180
  - Polynom, 136
  - semi-, 164, 180
- linear abhängig, 66, 95
- linear unabhängig, 66, 77, 79–82, 95
- Linearfaktor, 143
- Linearisierung, 51
- Linearkombination, 50, 63, 76
- Lösung
  - allgemeine, 90
  - Fundamental-, 161
  - spezielle, 90
  
- Mächtigkeit, 11
- Matrix, 69
  - Abbildungs-, 74, 85, 123, 185
  - adjungierte, 96, 184
  - Basiswechsel-, 74, 75, 117
  - Begleit-, 150
  - Block-, 85, 113
  - darstellende
    - Abbildung, 74
  - Diagonal-, 122
  - diagonalisierbare, 123, 157
  - Dreh-, 197
  - Dreiecks-, 113, 123, 153
  - Einheits-, 72
  - Elementar-, 91, 95
  - Gramsche, 171, 205
  - Hermiteische, 171
  - inverse, 72, 95, 115
  - invertierbare, 72, 95, 115
  - Jordan-, 153
    - positiv definite, 171
    - positiv semidefinite, 171
    - quadratische, 72
    - selbstadjungierte, 171
    - symmetrische, 171
    - transponierte, 86
    - trigonalisierbare, 123, 154
- Maximum, 136, 176
- Menge, 3, 4
  - endliche, 11
  - Lösungs-, 88
  - unendliche, 11
- Methode der kleinsten Quadrate, 98
- Metrik
  - affine, 199
  - diskrete, 200
  - Euklidische, 199
  - Hermiteische, 165
- Minimum, 176
- Mitternachtsformel, 144
- Modul
  - dualer, 53, 68
  - freier, 66, 67–68, 76
  - Links-, 47, 71, 144
  - Null-, 47
  - Quotienten-, 56
  - Rechts-, 47, 71
  - unitärer, 47
  - Unter-, 54, 145
- modulo, 41, 56
- multilinear, 103
- Multiplikation, 13, 40
  - komplexe, 24
    - geometrische Interpretation, 27
  - Matrix-, 69, 70–75
  - Quaternionen-, 31
  - skalare, 21, 47, 70
  - Verträglichkeit, 47
- multiplikativ, 26, 133
  
- nilpotent, 153
- Norm, 169
  - auf  $\mathbb{C}$ , 26
  - Euklidische, 22, 168
  - Skalarprodukt, 168
- Normalform, 196
  - Hermiteische Sequilinearform, 206
  - Jordan-, **153**, 158–162
  - lineare Abbildung, 85
  - normale Endomorphismen, 197
  - singuläre Werte, **193**
  - Weierstraß-, **151**
- Null, 40, 48
- Nullität, 205
- Nullmodul, 47
- Nullstelle, 143
- Nullteiler, 44, 137
  
- Ordnung, 11, 25
  - Nilpotenz-, 153
  - Nullstelle, 143
  - Primfaktor, 142

- Orientierung, 30, 118
  - Standard-, 118
- orientierungserhaltend, 118, 119
- orientierungsumkehrend, 118
- orthogonal, 173
- Paar, 6
- parallel, 88, 198
- Parallelotop, 30, 101, 178
- Peano-Axiome, 10
- Permutation, 105
- Polarardstellung, 27
- Polarisationsformel, 170
- Polynom, 76, 127
  - charakteristisches, 129, 153, 157
  - konstantes, 128
  - lineares, 136
  - Minimal-, 139, 146, 153, 157
  - normiertes, 136
  - Null-, 136
- positiv, 165, 169, 200
- Potenz, 13
- Potenzmenge, 6, 14
- Potenzreihe, 158
  - formale, 76
  - konvergente, 77
- prim, 141, 143
- Primfaktorzerlegung, 142, 147
- Primzahl, 45, 141
- Produkt
  - Algebra, 132
  - kartesisches, 6, 16
  - Kreuz-, 29, 118, 133
  - Lorentz-, 205
  - semidirektes, 202
  - Skalar-
    - Standard-, 22, 96
    - Spat-, 29, 103, 118
    - von Polynomen, 128
- Projektion, 178
  - orthogonale, 178
- Quadrate
  - Methode der kleinsten, 98
- Quadrik, 209, 210–213
- Quaternionen, 31, 32–36
- Quotientenmenge, 15, 17
- Quotientenraum, 56, 88
- Rang, 86
  - Spalten-, 86
  - Zeilen-, 86
- Raum
  - affiner, 198
  - Ausartungs-, 205
  - Eigen-, 121
  - Euklidischer, 199
  - Haupt-, 147
  - metrischer, 200
- Realteil, 25, 31
- Regel
  - Cramer, 115, 135
  - Sarrus, 111, 118
- Relation, 11
  - Äquivalenz-, 9, 14, 16, 20, 41, 88
  - antisymmetrisch, 11
  - reflexiv, 11, 14
  - symmetrisch, 14
  - transitiv, 11, 14
- Repräsentanten, 15
- Restklasse, 41
- Ring, 40
  - Endomorphismen-, 53, 133
  - Euklidischer, 137
  - Hauptideal-, 139
  - kommutativer, 40
  - Matrix-, 72, 133
  - mit Eins, 40
  - Null-, 41
  - Polynom-, 128, 133
  - unitärer, 40
  - Unter-, 55
- Russelsche Antinomie, 4
- Sarrussche Regel, 111, 118
- Satz, 4
  - Basisaustausch-, 81
  - Basisergänzungs-, 80
  - Cauchy-Schwarz-Ungleichung, 23, 169
  - Cayley-Hamilton, 134, 139, 151
  - chinesischer Rest-, 146
  - Cosinus-, 24
  - Diagonalisierbarkeit, 157
  - Euklidischer Algorithmus, 46
    - Polynome, 140
  - Fundamental- der Algebra, 26, 130, 143
  - Gauß-Verfahren, 92, 94–96, 114
  - Gram-Schmidt, 174
  - Haupt-
    - normale Abbildungen, 188
  - Hauptachsentransformation, 192, 213
  - Hauptraumzerlegung, 155
    - verallgemeinerte, 147, 151
  - Homomorphie-, 57
  - Hurwitz-Kriterium, 176
  - Laplace-Entwicklung, 105, 112,
    - 114–116
  - Leibniz-Formel, 110, 111, 115, 117, 130
  - Methode der kleinsten Quadrate, 98
  - Normalform
    - Jordan-, 153

- Weierstraß-, **151**
- Picard-Lindelöf, 160
- Primfaktorzerlegung, **142**
- Rang-, **85**, 86, 93
- Rieszscher Darstellungs-, 182
- Schwarz, 193
- singuläre Werte, **193**
- Spektral-
  - normale Operatoren, 188
  - selbstadjungierte Operatoren, 192
- Steinitz
  - Basisaustausch-, **81**
  - Basisergänzungs-, **80**
  - Sylvester-Kriterium, **176**
  - Sylvesterscher Trägheits-, **205**
  - Trigonalisierbarkeit, **153**
- scherungsinvariant, 101
- Schiefkörper, 42
- Schraubung, 204
- Schwingung, 132
  - gedampfte, 156
- Seite
  - linke, 88
  - rechte, 88
- selbstadjungiert, 171
- senkrecht, 173
- sesquilinear, 165
- Signatur, 206
- Signum, 105
- Skalarprodukt, 165
  - $L^2$ -, 167
  - Sobolev-, 168
  - Standard
    - komplexes, 167
    - quaternionisches, 167
  - Standard-, 22, 96, 163
    - komplexes, 96
    - quaternionisches, 96
  - zurückgeholtes, 187
- Spiegelung, 27, 197, 204
  - Dreh-, 204
  - Gleit-, 204
  - Punkt-, 35, 204
- Spin, 35
- Spur, 116
- subadditiv, 26
- Summe, 63
  - direkte, 84, 125, 126
  - von Polynomen, 128
  - von Untermoduln, 59, 125
    - direkte, 59, 84, 125
- surjektiv, 7, 8
- Symbol
  - Kronecker-, 62
- symmetrisch, 22, 165, 171, 200
- Teiler, 45, 138
  - größter gemeinsamer, 46, 140
- Teilmenge, 5
  - echte, 5
- total (Ordnung), 11
- Tägheitstensor, 192
- Transposition, 106
- trigonalisierbar, 123, 153
- Tupel, 6
- Umkehrabbildung, 9, 39
- unendlichdimensional, 82
- unendlichdimensional, 88
- Ungleichung
  - Cauchy-Schwarz-, 23, **169**
  - Dreiecks-, 169, 200
- Untermodul, 54
  - komplementärer, 59
- Unterraum, 54
  - affiner, 88, 198
  - invarianter, 145
  - komplementärer, 59, 83
- Urbild, 7, 89
- Ursprung, 199
- Vektor
  - Basis-
    - Standard-, 62
  - Null-, 22, 48
  - Spalten-, 62
  - Zeilen-, 63
- Vektorraum, 47
  - antidualer, 182
  - dualer, 53, 179
  - Euklidischer, 165
  - Quotienten-, 56
  - unitärer, 165
    - quaternionisch-, 165
  - Unter-, 54
    - invarianter, 145
- Vereinigung, 5
- Verkettung, 7, 39
- Verknüpfungen, 13
- Verschiebung, 35, 204
- Verträglichkeit
  - der Multiplikation, 47
- Vielfaches, 138
- Vielfachheit
  - algebraische, 157
  - geometrische, 157
- Volumen, 30, 101, 178
- Vorzeichen
  - Permutation, 105
- Wert
  - singulärer, 193–195

Wertebereich, 7  
Winkel, 22, 33  
wohldefiniert, 16  
Wurzel, 28

Zahlen

- ganze, 18
- komplexe, 24, 25–28
  - Polardarstellung, 27
- natürliche, 14
- rationale, 20
- reelle, 21

Zeilenstufenform, 91, 93–96

- strenge, 91, 95

Zeilenumformung

- elementare, 91, 92

Zerlegung

- Cholesky-, **176**
- Jordan-Chevalley, **154**