Inhaltsverzeichnis

Kapitel 1

Vektorräume

Der wichtigste Begriff der linearen Algebra, und einer der fundamentalsten in der Mathematik überhaupt, ist der des *Vektorraumes.* Vektorräume sind immer *über Körpern definiert*, und daher muß zuerst die algebraische Struktur „Körper“ erklärt werden. Jeder Körper ist auch ein *Ring*, und jeder Ring ist eine *Gruppe.* Wir beginnen daher mit der Definition einer Gruppe, einer algebraischen Struktur, an die gewisse Minimalforderungen gestellt sind, um sinnvoll rechnen zu können.

1.1 Gruppen

Definition einer Gruppe

Wir benutzen im folgenden einige elementare Begriffe der Mengenlehre, wie z.B. das *direkte Produkt* $M_1 \times M_1$ zweier Mengen M_1, M_2. Der Anhang A.1 stellt diese Dinge zusammen.

Definition 1.1.1 *Eine* **Gruppe** *ist eine Menge G, zusammen mit einer Abbildung*

$$\cdot : \; G \times G \longrightarrow G,$$

so daß folgende Axiome erfüllt sind:

i) Assoziativgesetz: *Für alle $a, b, c \in G$ gilt $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.*

ii) Neutrales Element: *Es gibt ein $e \in G$, genannt das* **neutrale Element***, so daß $e \cdot a = a \cdot e = a$ für alle $a \in G$.*

iii) Inverse Elemente: *Zu jedem $a \in G$ gibt es ein $a' \in G$ mit $a \cdot a' = e$.*

Die Gruppe G heißt **abelsch**[1] *oder* **kommutativ***, wenn zusätzlich das* Kommutativgesetz *gilt:*

$$a \cdot b = b \cdot a \qquad \text{für alle } a, b \in G.$$

[1] Niels Henrik Abel, 1802–1829

Die Abbildung „$\cdot$“ heißt auch **Verknüpfung** auf der Gruppe G. Soll betont werden, welche Verknüpfung auf G gemeint ist, so schreibt man genauer $(G, \cdot)$ für die Gruppe. Bei manchen Gruppen läßt man das Zeichen für die Verknüpfung auch weg und schreibt schlicht ab für $a \cdot b$.

Der Leser kann sofort nachprüfen, daß folgende Mengen mit ihren Verknüpfungen Gruppen definieren:

1. $(\mathbb{R}, +)$, d.h. die Menge der reellen Zahlen mit der Addition als Verknüpfung. Die Null ist das neutrale Element, und das Inverse zu a ist $-a$.

2. $(\mathbb{R}^*, \cdot)$, d.h. die Menge der reellen Zahlen außer Null mit der Multiplikation als Verknüpfung. Das neutrale Element ist die Eins, und das Inverse zu a ist a^{-1}.

3. Analog zu Beispiel 1 hat man die additiven Gruppen $(\mathbb{Q}, +)$ und $(\mathbb{C}, +)$, und analog zu Beispiel 2 die multiplikativen Gruppen $(\mathbb{Q}^*, \cdot)$ und $(\mathbb{C}^*, \cdot)$.

4. $(\mathbb{Z}, +)$.

5. Die zweielementige Gruppe $\{1, -1\}$, mit Multiplikation als Verknüpfung.

6. Die **triviale Gruppe**, die nur aus einem einzigen Element (dem neutralen Element) besteht.

Dies sind alles Beispiele *abelscher* Gruppen. Beispiele nicht-abelscher Gruppen werden wir etwa mit den *symmetrischen Gruppen* kennenlernen, siehe Abschnitt 3.1.

Feststellung 1.1.2 *Sei G eine Gruppe.*

i) Es gibt nur ein neutrales Element e in G.

ii) Zu jedem $a \in G$ gibt es genau ein Inverses, bezeichnet mit a^{-1}. Es gilt $a \cdot a^{-1} = a^{-1} \cdot a = e$.

Beweis: i) Wenn e und e' neutrale Elemente sind, so gilt $e = e \cdot e' = e'$.
ii) Sei a' ein Inverses zu a, d.h. $a \cdot a' = 1$. Sei a'' ein Inverses zu a', d.h. $a' \cdot a'' = 1$. Unter Verwendung der Gruppenaxiome rechnet man

$$a' \cdot a = a' \cdot a \cdot e = a' \cdot a \cdot a' \cdot a'' = a' \cdot e \cdot a'' = a' \cdot a'' = e,$$

d.h., das *Rechts*inverse a' zu a ist auch ein *Links*inverses. Wenn $\tilde{a}$ ein weiteres Inverses zu a ist, so gilt

$$a' = a' \cdot e = a' \cdot a \cdot \tilde{a} = e \cdot \tilde{a} = \tilde{a}.$$

Also ist das Inverse eindeutig bestimmt. ∎

Verwendet man die additive Schreibweise „+“ für die Verknüpfung in einer Gruppe G, so wird das Inverse zu $a \in G$ nicht mit a^{-1}, sondern mit $-a$ bezeichnet.

Gruppenhomomorphismen

Definition 1.1.3 *Seien $(G, \cdot)$ und $(H, \circ)$ Gruppen. Eine Abbildung $f : G \to H$ heißt* **Gruppenhomomorphismus** *(oder kurz Homomorphismus[2]), wenn*

$$f(a \cdot b) = f(a) \circ f(b) \qquad \text{für alle } a, b \in G.$$

[2] gr. *homos*: gleichartig; *morphe*: Gestalt, Form

Homomorphismen sind also diejenigen Abbildungen zwischen Gruppen, die mit den Verknüpfungen „verträglich“ sind.

Beispiele: 1. Die Funktionalgleichung der e-Funktion $e^{a+b} = e^a e^b$ besagt gerade, daß

$$\exp : \mathbb{C} \longrightarrow \mathbb{C}^*$$

ein Homomorphismus der (additiven) Gruppe $\mathbb{C}$ auf die (multiplikative) Gruppe $\mathbb{C}^*$ ist.

2. Die Abbildung, die jeder von Null verschiedenen reellen Zahl ihr Vorzeichen zuordnet, ist ein Homomorphismus $\mathbb{R}^* \to \{1, -1\}$ von $\mathbb{R}^*$ auf die endliche Gruppe $\{1, -1\}$.

3. Für jede Menge M hat man die **Identität** $\mathrm{id}_M : M \to M$, die dadurch definiert ist, daß sie jedes Element auf sich selbst abbildet. Ist $M = G$ eine Gruppe, so ist id_G ein Gruppenhomomorphismus.

Feststellung 1.1.4 *Seien G und H Gruppen mit neutralem Element e_G bzw. e_H. Für jeden Homomorphismus $f : G \to H$ gilt:*

i) $f(e_G) = e_H$.

ii) $f(a)^{-1} = f(a^{-1})$ *für alle* $a \in G$.

Beweis: i) Offenbar gilt $f(e_G) = f(e_G e_G) = f(e_G)f(e_G)$. Nun multipliziere man beide Seiten mit $f(e_G)^{-1}$.
ii) $f(a)f(a^{-1}) = f(aa^{-1}) = f(e_G) = e_H$ nach i). Dies besagt gerade, daß $f(a^{-1})$ das Inverse zu $f(a)$ ist, d.h. $f(a^{-1}) = f(a)^{-1}$. ■

Bemerkung 1.1.5 Die Hintereinanderschaltung zweier Gruppenhomomorphismen ist wieder ein solcher. Genauer: Sind

$$G_1 \xrightarrow{f} G_2 \xrightarrow{g} G_3$$

Gruppenhomomorphismen, so ist die Komposition $g \circ f : G_1 \to G_3$, definiert durch $(g \circ f)(a) = g(f(a))$, $a \in G_1$, wieder ein Gruppenhomomorphismus. Denn $(g \circ f)(ab) = g(f(ab)) = g(f(a)f(b)) = g(f(a))g(f(b)) = (g \circ f)(a)(g \circ f)(b)$.

Wenn G und H Gruppen sind mit $G \subset H$, und die Verknüpfung auf G stimmt mit der Einschränkung der Verknüpfung auf H überein, so heißt G eine **Untergruppe** von H. Man schreibt dann auch $G < H$. So ist etwa $\mathbb{Q} < \mathbb{R} < \mathbb{C}$ (mit $+$ als Verknüpfung). Jede Gruppe besitzt die triviale Untergruppe $\{e\}$, die nur aus dem neutralen Element besteht. Außerdem ist jede Gruppe Untergruppe von sich selbst.

Ist $f : G \to H$ ein Gruppenhomomorphismus, so definieren wir

$$\begin{aligned} \ker(f) &:= \{a \in G : \ f(a) = e_H\}, \\ \mathrm{im}(f) &:= \{b \in H : \text{ es gibt ein } a \in G \text{ mit } f(a) = b\}. \end{aligned} \tag{1.1}$$

$\ker(f)$ heißt der **Kern** von f, und $\mathrm{im}(f)$ das **Bild** von f. Statt $\mathrm{im}(f)$ schreibt man auch $f(G)$.

Feststellung 1.1.6 *Wenn $f : G \to H$ ein Gruppenhomomorphismus ist, so gilt:*

i) $\ker(f)$ *ist eine Untergruppe von* G.

ii) $\mathrm{im}(f)$ *ist eine Untergruppe von* H.

Beweis: i) Zunächst ist $\ker(f)$ eine Teilmenge von G. Wenn $a, b \in \ker(f)$, dann gilt $f(ab) = f(a)f(b) = e_H e_H = e_H$, und somit $ab \in \ker(f)$. Die Verknüpfung auf G induziert also eine Verknüpfung auf $\ker(f)$, und wir müssen zeigen, daß $\ker(f)$ mit dieser Verknüpfung die Gruppenaxiome erfüllt. Assoziativität ist klar, denn diese ist für alle $a, b, c \in G$ erfüllt, erst recht also für $a, b, c \in \ker(f)$. Das neutrale Element e_G von G liegt in $\ker(f)$ nach Feststellung 1.1.4 i), und ist offenbar auch ein neutrales Element in $\ker(f)$. Es bleibt somit nur zu zeigen, daß mit a auch a^{-1} in $\ker(f)$ liegt. Doch dies ist richtig nach Feststellung 1.1.4 ii): $f(a^{-1}) = f(a)^{-1} = e_H^{-1} = e_H$.

ii) Seien $b_1, b_2 \in \mathrm{im}(f)$, und $a_1, a_2 \in G$ mit $f(a_i) = b_i$. Dann ist $f(a_1 a_2) = f(a_1)f(a_2) = b_1 b_2$, also ist $b_1 b_2$ im Bild. Die Verknüpfung auf H induziert somit eine Verknüpfung auf $\mathrm{im}(f)$, und es ist zu zeigen, daß $\mathrm{im}(f)$ mit dieser Verknüpfung die Gruppenaxiome erfüllt. Assoziativität ist wieder klar, da diese sogar in der größeren Menge H gilt. Das neutrale Element e_H liegt im Bild nach Feststellung 1.1.4 i). Und wenn $b = f(a)$ im Bild liegt, so auch $b^{-1} = f(a)^{-1} = f(a^{-1})$ nach Feststellung 1.1.4 ii).■

Bemerkung 1.1.7 Der eben geführte Beweis verwendete zweimal die folgende Beobachtung: Sei S eine Teilmenge der Gruppe G. S ist bereits dann eine Untergruppe von G, wenn folgendes erfüllt ist:

$$a \in S \text{ und } b \in S \implies ab \in S,$$
$$e \in S \text{ (das neutrale Element von } G),$$
$$a \in S \implies a^{-1} \in S.$$

Denn die Gruppenaxiome für S folgen dann aus den Gruppenaxiomen für G.

Der Leser sei an die Begriffe *injektiv*, *surjektiv* und *bijektiv* erinnert, die sich in Anhang A.1 finden.

Feststellung 1.1.8 *Sei* $f : G \to H$ *ein Gruppenhomomorphismus.*

i) f *ist injektiv genau dann, wenn* $\ker(f) = \{e_G\}$ *(die triviale Untergruppe).*

ii) f *ist surjektiv genau dann, wenn* $\mathrm{im}(f) = H$.

Beweis: i) Sei zunächst f injektiv. Da $f(e_G) = e_H$, kann kein weiteres Element von G auf e_H abgebildet werden (Injektivität). Deshalb besteht der Kern von f nur aus e_G. — Sei nun $\ker(f) = \{e_G\}$ vorausgesetzt. Seien $a, b \in G$ mit $f(a) = f(b)$. Wir müssen zeigen, daß $a = b$. Es gilt $f(ab^{-1}) = f(a)f(b^{-1}) = f(a)f(b)^{-1} = f(a)f(a)^{-1} = e_H$, also $ab^{-1} \in \ker(f)$. Da der Kern nur aus e_G besteht, folgt $ab^{-1} = e_G$. Dies ist equivalent mit $a = b$.

ii) $\mathrm{im}(f) = H$ ist gerade die Definition der Surjektivität. ■

Definition 1.1.9 *Wenn* $f : G \to H$ *ein bijektiver Gruppenhomomorphismus ist, so heißt* f *ein* **Isomorphismus** [3]. *Die Gruppen* G *und* H *heißen dann* **isomorph**.

[3]gr. *isos*: gleich

Isomorphe Gruppen sind gewissermaßen „gleich“: Ihre Elemente stehen in einer ein-eindeutigen Beziehung, die mit den Verknüpfungen verträglich ist. Allein durch ihre Eigenschaften als Gruppe lassen sich G und H nicht unterscheiden.

Beispiel: Der Logarithmus $\log : \mathbb{R}_{>0} \to \mathbb{R}$ ist ein Isomorphismus der Gruppe der positiven reellen Zahlen (mit Multiplikation als Verknüpfung) auf die additive Gruppe aller reellen Zahlen.

Bemerkung 1.1.10 In diesem Abschnitt wurden gewisse *Objekte* (nämlich Gruppen) behandelt, zusammen mit gewissen *Morphismen* (den Gruppenhomomorphismen) zwischen diesen Objekten. In ähnlicher Weise werden wir in Kürze *Ringe* und *Ringhomomorphismen* betrachten, oder *Vektorräume* und *Vektorraumhomomorphismen*. Diese Situation, daß eine gewisse Klasse von Objekten vorliegt, und eine dazu „passende“ Klasse von Morphismen, tritt so häufig in der Mathematik auf, daß man ihr durch den Begriff der *Kategorie* einen abstrakten Rahmen verleiht. Z.B. spricht man von der Kategorie **Gr** der Gruppen und Gruppenhomomorphismen. Eine genaue Definition findet sich in Anhang B.1.

1.2 Ringe und Körper

Ringe und Ringhomomorphismen

Definition 1.2.1 *Sei $(R, +)$ eine abelsche Gruppe mit neutralem Element 0. R heißt ein* **Ring***, wenn eine weitere Verknüpfung*

$$\cdot : R \times R \longrightarrow R,$$

die „Multiplikation“, definiert ist, so daß gilt:

i) Assoziativgesetz der Multiplikation: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ *für alle* $a, b, c \in R$.

ii) Distributivgesetze: *Für alle $a, b, c \in R$ gilt*

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c),$$
$$(b + c) \cdot a = (b \cdot a) + (c \cdot a).$$

Der Ring R heißt **kommutativ***, wenn zusätzlich*

$$a \cdot b = b \cdot a \qquad \text{für alle } a, b \in R$$

gilt. Gibt es in R ein **Einselement** *1 mit $1 \cdot a = a \cdot 1 = a$ für alle $a \neq 0$ (d.h., 1 ist ein neutrales Element der Multiplikation), so spricht man von einem* **Ring mit Eins***.*

Die beiden Verknüpfungen $+$ und $\cdot$ auf R heißen „Addition“ bzw. „Multiplikation“, obwohl sie nicht notwendig etwas mit der gewöhnlichen Addition und Multiplikation zu tun haben müssen. Manchmal schreibt man kurz $a \cdot b = ab$.

Die Distributivgesetze drücken die Verträglichkeit von Addition und Multiplikation aus. Wenn der Ring kommutativ ist, dann folgt das linksseitige Distributivgesetz aus dem rechtsseitigen, und umgekehrt. Um Klammern zu sparen, einigt man sich auf die „Punkt–vor–Strich“–Regel; die Klammern auf der rechten Seite der Distributivgesetze könnten also weggelassen werden.

Beispiele: 1. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ mit der gewöhnlichen Addition und Multiplikation sind kommutative Ringe mit Eins.

2. Sei $n \geq 2$ eine natürliche Zahl. Die Menge $n\mathbb{Z} = \{nk : k \in \mathbb{Z}\}$ aller Vielfachen von n ist ein kommutativer Ring ohne Eins.

3. Für jede natürliche Zahl $n > 1$ konstruieren wir einen kommutativen Ring mit n Elementen. Die Menge der ganzen Zahlen $\mathbb{Z}$ zerlegt sich in die n disjunkten **Restklassen**

$$i + n\mathbb{Z} = \{i + nk : k \in \mathbb{Z}\}, \qquad i = 0, \ldots, n-1.$$

Die Menge, deren Elemente diese n Restklassen sind, wird mit $\mathbb{Z}/n\mathbb{Z}$ bezeichnet (sprich: „Z modulo n Z“). Für $a \in \mathbb{Z}$ sei $\bar{a}$ diejenige Restklasse, in der a liegt. Dies definiert offenbar eine surjektive Abbildung, die *Projektion*,

$$\begin{aligned} \mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z}, \\ a &\longmapsto \bar{a}. \end{aligned} \tag{1.2}$$

Es ist leicht einzusehen, daß $\overline{a+b}$ und $\overline{a \cdot b}$ nur von den Restklassen abhängen, in denen a und b liegen (d.h., ist $\bar{a'} = \bar{a}$ und $\bar{b'} = \bar{b}$, so ist $\overline{a' + b'} = \overline{a+b}$ und $\overline{a' \cdot b'} = \overline{a \cdot b}$). Daher sind auf $\mathbb{Z}/n\mathbb{Z}$ eine Addition und eine Multiplikation durch

$$\bar{a} + \bar{b} := \overline{a+b}, \qquad \bar{a} \cdot \bar{b} := \overline{a \cdot b}.$$

wohldefiniert. Der Leser prüft nun leicht nach, daß $\mathbb{Z}/n\mathbb{Z}$ mit diesen Verknüpfungen ein kommutativer Ring mit 1 wird (mit den neutralen Elementen $\bar{0}$ und $\bar{1}$). Er heißt der **Restklassenring mod** n.

Definition 1.2.2 *Seien R und S Ringe. Eine Abbildung $f : R \to S$ heißt* **Ringhomomorphismus**, *wenn gilt:*

i) $f(a+b) = f(a) + f(b)$ *für alle* $a, b \in R$.

ii) $f(ab) = f(a)f(b)$ *für alle* $a, b \in R$.

Ein bijektiver Ringhomomorphismus heißt ein **Isomorphismus**.

Eigenschaft i) besagt gerade, daß f ein Homomorphismus der unterliegenden additiven Gruppen ist. Eigenschaft ii) bedeutet, daß f mit den Multiplikationen verträglich ist.

Beispiele: 1. Die komplexe Konjugation $\mathbb{C} \to \mathbb{C}$ ist ein Ringhomomorphismus.

2. Die Projektion (1.2) ist ein Ringhomomorphismus.

Bemerkung 1.2.3 Genau wie im Falle von Gruppen und Gruppenhomomorphismen haben wir es auch hier wieder mit *Objekten* und *Morphismen* zu tun, nämlich Ringen und Ringhomomorphismen (vgl. Bemerkung 1.1.10). Diese bilden zusammen die *Kategorie* **Ring** der Ringe. Diese Begriffe werden in Anhang B.1 genauer definiert.

Ein Ringhomomorphismus $f : R \to S$ ist insbesondere ein Homomorphismus abelscher Gruppen, und deshalb ist der **Kern** $\ker(f)$ und das **Bild** $\mathrm{im}(f)$ definiert:

$$\ker(f) = \{a \in R : \; f(a) = 0\},$$
$$\mathrm{im}(f) = f(R) = \{b \in S : \text{ es gibt ein } a \in R \text{ mit } f(a) = b\}$$

(siehe (1.1)). Nach Feststellung 1.1.6 handelt es sich um additive Untergruppen von R bzw. S. Der Leser kann leicht nachprüfen, daß es sich sogar um Unter*ringe* handelt.

Beispiel: Der Kern der Projektion (1.2) sind die ganzzahligen Vielfachen $n\mathbb{Z} = \{nk : \; k \in \mathbb{Z}\}$ von n.

Definition 1.2.4 *Ein Ring R heißt* **nullteilerfrei**, *wenn für alle $a, b \in R$ gilt:*

$$a \cdot b = 0 \implies a = 0 \textit{ oder } b = 0.$$

Ein kommutativer, nullteilerfreier Ring mit Eins heißt ein **Integritätsring**.

Beispiele: 1. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sind Integritätsringe.

2. Sei $n = rs$ mit natürlichen Zahlen $r, s \geq 2$ (d.h., n ist keine Primzahl). Dann ist $\mathbb{Z}/n\mathbb{Z}$ nicht nullteilerfrei, denn $\bar{r} \neq \bar{0}$ und $\bar{s} \neq \bar{0}$, aber $\bar{a}\bar{b} = \overline{ab} = \bar{n} = \bar{0}$.

Für jedes Element a des Ringes R gilt $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$. Addiert man auf beiden Seiten das (additive) Inverse von $0 \cdot a$, so folgt $0 \cdot a = 0$. Ebenso zeigt man $a \cdot 0 = 0$. Die 0 besitzt also kein multiplikatives inverses Element. Folglich ist ein Ring niemals eine Gruppe bezüglich der Multiplikation. Man erhält aber eine Gruppe, wenn man alle multiplikativ invertierbaren Elemente sammelt:

Definition und Feststellung 1.2.5 *Sei R ein Ring mit Eins. Ein Element $a \in R$ heißt* **invertierbar** *oder eine* **Einheit**, *wenn ein $a' \in R$ existiert mit $a \cdot a' = a' \cdot a = 1$. Die Menge R^* aller invertierbaren Elemente ist eine Gruppe bezüglich der Multiplikation, die* **Einheitengruppe** *des Ringes R.*

Beweis: Mit $a, b \in R^*$ ist auch $ab \in R^*$, denn $(ab)(b'a') = 1$. Also induziert die Multiplikation eine Verknüpfung auf R^*. Wegen $1 \cdot 1 = 1$ ist $1 \in R^*$; also enthält R^* ein neutrales Element. Per definitionem enthält R^* auch inverse Elemente. Schließlich erbt R^* die Assoziativität (der Multiplikation) von R. Damit sind alle Gruppenaxiome erfüllt. ∎

Beispiele: 1. Die Einheitengruppe von $\mathbb{Z}$ ist $\{1, -1\}$.

2. Die Einheitengruppe von $\mathbb{Q}$ (bzw. $\mathbb{R}$ bzw. $\mathbb{C}$) besteht aus allen von Null verschiedenen Elementen:

$$\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$$

(das Zeichen „\“ bedeutet „ohne“, siehe A.1). Die Notation $\mathbb{Q}^*$ für die Einheiten ist also konsistent mit der üblichen Bedeutung dieses Symbols.

3. Wir betrachten wieder die Projektion (1.2). Ein Element $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ ist genau dann eine Einheit, wenn a zu n teilerfremd ist. Denn sei letzteres der Fall. Nach einem Satz der elementaren Zahlentheorie gibt es dann $x, y \in \mathbb{Z}$ mit

$$xa + yn = 1.$$

Anwenden der Projektion ergibt $\bar{x}\bar{a} = \bar{1}$, was zeigt, daß $\bar{a}$ eine Einheit ist. Ist umgekehrt $\bar{a}\bar{x} = \bar{1}$ für ein $x \in \mathbb{Z}$, so ist $xa - 1$ durch n teilbar, also a zu n teilerfremd.

Körper

Definition 1.2.6 *Ein kommutativer Ring R mit Eins heißt ein* **Körper**, *falls $R^* = R \setminus \{0\}$.*

Körper sind also diejenigen kommutativen Ringe, bei denen jedes von Null verschiedene Element invertierbar ist.

Nicht-kommutative Ringe mit $R^* = R \setminus \{0\}$ heißen **Schiefkörper**. Sie spielen in diesem Buch keine Rolle.

Beispiele: 1. $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ sind Körper. $\mathbb{Z}$ ist kein Körper.

2. Der Restklassenring $\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Körper, wenn $n = p$ eine Primzahl ist. Dies folgt aus obigem Beispiel 3.
Insbesondere ist $\mathbb{Z}/2\mathbb{Z}$ ein Körper. Er besteht nur aus den beiden Elementen 0 und 1, und ist damit der kleinste Körper überhaupt.
Der endliche Körper $\mathbb{Z}/p\mathbb{Z}$ wird auch mit $\mathbb{F}_p$ bezeichnet. Ohne Beweis sei mitgeteilt, daß es zu jeder Primzahlpotenz $q = p^r$ einen und (bis auf Isomorphie) nur einen Körper $\mathbb{F}_q$ mit genau q Elementen gibt, und daß damit alle endlichen Körper erfaßt sind (siehe Standardtexte zur Algebra, etwa [La2]).

Feststellung 1.2.7 *Körper sind nullteilerfrei.*

Beweis: Seien a, b Elemente des Körpers K mit $a \cdot b = 0$. Angenommen, $b \neq 0$. Da K ein Körper ist, gibt es ein multiplikatives Inverses b^{-1} zu b. Es folgt $a = a \cdot 1 = a \cdot b \cdot b^{-1} = 0 \cdot b^{-1} = 0$. Aus $a \cdot b = 0$ folgt also $a = 0$ oder $b = 0$. ∎

Definition 1.2.8 *Sei K ein Körper. Falls es eine natürliche Zahl p gibt mit*

$$\underbrace{1 + \ldots + 1}_{p\text{-mal}} = 0,$$

so heißt das kleinste p mit dieser Eigenschaft die **Charakteristik** *von K. Falls kein p mit dieser Eigenschaft existiert, so bekommt K die Charakteristik 0.*

Beispiele: $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ haben die Charakteristik 0. Der endliche Körper $\mathbb{F}_p$ hat die Charakteristik p.

Feststellung 1.2.9 *Sei K ein Körper mit Charakteristik $p > 0$. Dann ist p eine Primzahl.*

Beweis: Sei $p = mn$ mit natürlichen Zahlen m, n. Wir haben zu zeigen, daß $m = p$ oder $n = p$. Offenbar gilt

$$0 = (\underbrace{1 + \ldots + 1}_{p\text{-mal}}) = (\underbrace{1 + \ldots + 1}_{m\text{-mal}})(\underbrace{1 + \ldots + 1}_{n\text{-mal}}).$$

Aus der Nullteilerfreiheit von K folgt

$$(\underbrace{1 + \ldots + 1}_{m\text{-mal}}) = 0 \qquad \text{oder} \qquad (\underbrace{1 + \ldots + 1}_{n\text{-mal}}) = 0.$$

Da p die kleinste Zahl ist, so daß eine p-fache Summe von Einsen verschwindet, folgt $m = p$ oder $n = p$. ∎

1.3 Vektorräume

Definition und Beispiele

Wir kommen jetzt zum zentralen Begriff der linearen Algebra.

Definition 1.3.1 *Ein Vektorraum ist eine abelsche Gruppe, auf der ein Körper linear operiert. Genauer: Sei K ein Körper und $(V, +)$ eine abelsche Gruppe. Dann heißt V ein* **Vektorraum über** K *(oder ein K-Vektorraum), falls eine Abbildung*

$$\begin{aligned} K \times V &\longrightarrow V, \\ (a, v) &\longmapsto a.v \end{aligned}$$

die **Skalarmultiplikation**, *definiert ist, so daß folgendes gilt:*

i) $(a + b).v = a.v + b.v$ *für alle* $a, b \in K$, $v \in V$.

ii) $a.(v + w) = a.v + a.w$ *für alle* $a \in K$, $v, w \in V$.

iii) $(a \cdot b).v = a.(b.v)$ *für alle* $a, b \in K$, $v \in V$.

iv) $1.v = v$ *für alle* $v \in V$ *(1 ist das neutrale Element der Multiplikation in K).*

Die Elemente eines Vektorraumes V heißen **Vektoren**.

Die rechten Seiten von i) und ii) sind wieder gemäß der „Punkt–vor–Strich"–Regel zu lesen. Man beachte, daß wir es hier mit vier verschiedenen Verknüpfungen zu tun haben: Der Addition $+$ in K, der Multiplikation $\cdot$ in K, der Addition $+$ in V, und der skalaren Multiplikation. Daß die Additionen in K und V mit dem gleichen Symbol bezeichnet werden, sollte für keine Verwirrung sorgen; z.B. ist klar, daß auf der linken Seite von i) die Addition in K gemeint ist, während auf der rechten Seite nur die Addition in V Sinn macht.

Oft werden die Symbole für die Multiplikationen weggelassen; dann liest sich iii) als $(ab)v = a(bv)$. Es gibt keine Möglichkeit, zwei Elemente von V miteinander zu multiplizieren[4]. Die Elemente von K heißen auch **Skalare**[5]; dies erklärt den Namen Skalarmultiplikation.

---

[4]Außer, eine solche Multiplikation ist zusätzlich definiert; der Vektorraum heißt dann eine *Algebra*, siehe 1.3.14.

[5]lat. *scalaris*: zur Leiter, Treppe gehörend. Die „Leiter" kann z.B. der geordnete Körper $K = \mathbb{R}$ sein.

Beispiel 1.3.2 Das folgende Beispiel für einen Vektorraum ist fundamental. Sei n eine feste natürliche Zahl, und K^n die Menge der **Spaltenvektoren**

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \qquad a_1, \dots, a_n \in K.$$

Wir definieren eine Addition auf K^n durch Addition der Komponenten:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

($a_i + b_i$ bedeutet natürlich die Addition in K). Der Leser prüft leicht nach, daß K^n mit dieser Verknüpfung eine abelsche Gruppe ist; das neutrale Element ist der **Nullvektor**

$$\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Wir wollen K^n zum K-Vektorraum machen. Dazu definieren wir die skalare Multiplikation durch komponentenweise Multiplikation:

$$a.\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} a \cdot a_1 \\ \vdots \\ a \cdot a_n \end{pmatrix},$$

wobei $a \cdot a_i$ die Multiplikation im Körper K bedeutet. Um nachzuprüfen, daß K^n damit in der Tat zum K-Vektorraum wird, sind die Eigenschaften i) bis iv) aus Definition 1.3.1 zu verifizieren; dies sei dem Leser überlassen.

Statt mit Spaltenvektoren hätten wir auch mit **Zeilenvektoren** $(a_1, \dots, a_n)$ arbeiten können. Dies liefert „denselben“ (einen *isomorphen*) Vektorraum, da die Vektoren lediglich anders notiert werden. In der Regel fassen wir K^n als Raum von Spaltenvektoren auf, werden jedoch manchmal aus notationstechnischen Gründen auch Zeilenvektoren verwenden.

Offenbar ist $K^1 = K$. Der Körper K (genauer: seine additive Gruppe) ist also ein Vektorraum über sich selbst.

Sei speziell $K = \mathbb{R}$. Bekanntlich können die reellen Zahlen durch die Punkte auf einer Geraden veranschaulicht werden. Interpretiert man die Komponenten eines Vektors in $\mathbb{R}^2$ als x- und y-Koordinate in einem Koordinatensystem der Ebene, so erhält man eine Visualisierung des $\mathbb{R}^2$ als Punkte in einer Ebene. Analog können die Elemente des $\mathbb{R}^3$ als Punkte des dreidimensionalen Raumes veranschaulicht werden. Der Leser überlege sich, was Addition und skalare Multiplikation in diesen geometrischen Interpretationen der Räume $\mathbb{R}^2$ und $\mathbb{R}^3$ bedeuten („Vektoraddition“, „Streckung“, „Punktspiegelung“). Diese geometrischen Bilder sind eine gewisse Quelle der Intuition für die Lineare Algebra.

Beispiel 1.3.3 Weitere typische Beispiele von Vektorräumen entstehen als Räume von Funktionen: Sei $I \subset \mathbb{R}$ ein Intervall, und $\mathcal{F}$ die Menge aller Funktionen $f : I \to \mathbb{R}$. Für zwei Funktionen $f, g \in \mathcal{F}$ definieren wir eine neue Funktion $f + g$ durch

$$(f+g)(x) := f(x) + g(x).$$

Man sieht sofort, daß $\mathcal{F}$ mit dieser Verknüpfung eine abelsche Gruppe wird (das neutrale Element ist die Nullfunktion ($f(x) = 0$ für alle $x \in I$), und das Inverse zu f ist $-f : x \mapsto -f(x)$). Für eine reelle Zahl a und $f \in \mathcal{F}$ sei die Funktion $af \in \mathcal{F}$ definiert durch

$$(af)(x) := a \cdot f(x).$$

Damit ist eine skalare Multiplikation

$$\begin{aligned} \mathbb{R} \times \mathcal{F} &\longrightarrow \mathcal{F}, \\ (a, f) &\longmapsto af, \end{aligned}$$

erklärt, durch die $\mathcal{F}$ zum $\mathbb{R}$-Vektorraum wird. Die Analysis kennt zahllose Beispiele dieser Art, wo reelle und komplexe Vektorräume (d.h., Vektorräume über $\mathbb{R}$ oder $\mathbb{C}$) als Funktionenräume entstehen.

Ein weiteres Beispiel für einen K-Vektorraum ist der **Nullvektorraum** $\{0\}$, der nur aus einem einzigen Element 0 besteht. Es handelt sich also um die triviale Gruppe, zusammen mit der dementsprechend trivialen Skalarmultiplikation $a.0 = 0$ für alle $a \in K$. Wir vereinbaren, unter K^0 immer diesen Nullvektorraum zu verstehen.

Feststellung 1.3.4 *Sei V ein K-Vektorraum, und* $\mathbf{0}$ *das neutrale Element der abelschen Gruppe V. Dann gilt:*

i) $0.v = \mathbf{0}$ *für alle* $v \in V$.

ii) $a.\mathbf{0} = \mathbf{0}$ *für alle* $a \in K$.

iii) $(-1).v = -v$ *für alle* $v \in V$.

Beweis: i) $0.v = (0+0).v = 0.v + 0.v$. Nun addiere man auf beiden Seiten das Inverse von $0.v \in V$.
ii) $a.\mathbf{0} = a.(\mathbf{0} + \mathbf{0}) = a.\mathbf{0} + a.\mathbf{0}$.
iii) $v + (-1).v = 1.v + (-1).v = (1 + (-1)).v = 0.v = \mathbf{0}$ nach i). Also ist $(-1).v$ das Inverse zu v. ∎

Definition und Feststellung 1.3.5 *Sei V ein K-Vektorraum und $W \subset V$ eine nicht-leere Teilmenge. W heißt ein* **Untervektorraum** *von V, falls gilt:*

i) $w_1, w_2 \in W \implies w_1 + w_2 \in W$.

ii) $a \in K$ *und* $w \in W \implies aw \in W$.

W ist dann selbst ein K-Vektorraum.

Beweis: i) besagt, daß die Addition auf V eine Verknüpfung auf W induziert. Wir müssen zunächst zeigen, daß W damit eine abelsche Gruppe wird. Sei $w \in W$ beliebig (W ist nicht leer). Wegen ii) ist dann auch $-w = (-1).w \in W$. Es folgt $\mathbf{0} = w + (-w) \in W$ nach i). Also enthält W ein neutrales Element. Wir haben eben gesehen, daß mit w auch $-w$ in W liegt; daher enthält W inverse Elemente. Schließlich erbt W die Assoziativität und Kommutativität von V, ist also eine abelsche Gruppe.

ii) besagt nun, daß die Skalarmultiplikation auf V eine Skalarmultiplikation auf W induziert. Damit wird W zum K-Vektorraum, denn die benötigten Eigenschaften gelten ja sogar in der größeren Menge V. ∎

Beispiele: 1. Seien $m < n$ natürliche Zahlen. Wir betrachten die Menge W derjenigen Elemente

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K^n,$$

für die $a_{m+1} = \ldots = a_n = 0$. Dieses W ist offenbar abgeschlossen unter Addition und skalarer Multiplikation, d.h., es erfüllt die Eigenschaften i) und ii) aus 1.3.5. Damit ist W ein Untervektorraum des K^n. Durch Weglassen der $n - m$ unteren Nullkomponenten lassen sich die Elemente von W mit denen des K^m identifizieren. Auf diese Weise kann man K^m als Untervektorraum des K^n auffassen (es gibt viele andere Möglichkeiten, dies zu tun; man wähle z.B. $n - m$ andere Komponenten, und setze diese 0).

2. Sei $\mathcal{F}$ der $\mathbb{R}$-Vektorraum aus Beispiel 1.3.3. Die Teilmenge $C^0(I)$ der *stetigen* Funktionen $I \to \mathbb{R}$ ist ein Untervektorraum von $\mathcal{F}$. Sei das Intervall I als offen vorausgesetzt, so daß wir problemlos differenzierbare Funktionen betrachten können. Für $r \in \mathbb{N}$ ist die Menge

$$C^r(I) := \{f \in \mathcal{F} : \ f \text{ ist } r\text{-mal stetig differenzierbar}\}$$

ein Untervektorraum von $C^0(I)$ (und auch von $\mathcal{F}$). In der Tat hat man eine ganze Kette von Untervektorräumen

$$\ldots \subset C^2(I) \subset C^1(I) \subset C^0(I) \subset \mathcal{F}. \tag{1.3}$$

Feststellung 1.3.6 *Der Schnitt beliebig vieler Untervektorräume ist wieder ein Untervektorraum. Genauer: Sei V ein K-Vektorraum, und I eine beliebige (endliche oder unendliche) Indexmenge. Für jedes $i \in I$ sei ein Untervektorraum V_i von V gegeben. Dann ist*

$$W := \bigcap_{i \in I} V_i$$

wieder ein Untervektorraum von V.

Beweis: Seien $w_1, w_2 \in W$. Dann liegen w_1 und w_2 auch in jedem V_i, also liegt $w_1 + w_2$ in jedem V_i. Mit anderen Worten, $w_1 + w_2 \in W$. Also ist Bedingung i) in Definition 1.3.5 erfüllt. Bedingung ii) sieht man ähnlich: Sei $a \in K$ und $w \in W$. Für jedes i liegt w in V_i, also auch aw. Das heißt gerade, daß $aw \in W$. ∎

Beispiel: Wir betrachten die Untervektorräume $C^j(I)$, $j \geq 0$, von $\mathcal{F}$ wie in (1.3). Der Schnitt $\bigcap_{j \geq 0} C^j(I)$ ist der Vektorraum $C^\infty(I)$ der glatten (beliebig oft differenzierbaren) Funktionen auf dem Intervall I.

Bemerkung 1.3.7 Wenn man in Definition 1.3.1 den Körper K durch einen Ring R ersetzt, so spricht man nicht von einem Vektorraum, sondern von einem **Modul** über R. Wir werden in Kapitel 3 Anlaß haben, über Moduln zu reden, siehe Definition 3.2.1.

Lineare Abbildungen

Definition 1.3.8 *Seien V und W Vektorräume über dem Körper K. Eine Abbildung $f : V \to W$ heißt* **Vektorraumhomomorphismus** *oder K-***linear** *oder einfach* **linear***, falls gilt:*

i) $f(v_1 + v_2) = f(v_1) + f(v_2)$ *für alle $v_1, v_2 \in V$.*

ii) $f(av) = af(v)$ *für alle $a \in K$, $v \in V$.*

Ist $V = W$, so spricht man von einem **Endomorphismus** [6] *von V. Eine bijektive lineare Abbildung heißt* **Isomorphismus***.*

Eigenschaft i) besagt, daß f ein Gruppenhomomorphismus zwischen den abelschen Gruppen V und W ist. Insbesondere gilt $f(\mathbf{0}_V) = \mathbf{0}_W$. Eigenschaft ii) drückt die Verträglichkeit mit den Skalarmultiplikationen auf V und W aus.

Beispiele: 1. Sei V ein K-Vektorraum, und $a \in K$ fest. Dann ist die Abbildung

$$\begin{aligned} h_a : V &\longrightarrow V, \\ v &\longmapsto av, \end{aligned}$$

linear, also ein Endomorphismus von V. Dieser wird manchmal die **Homothetie** mit a genannt. Man überlege sich, daß h_a für $a \neq 0$ sogar ein Isomorphismus ist. Offenbar ist $h_1 = \mathrm{id}_V$ die **Identität** auf V (die dadurch definiert ist, daß sie jedes Element auf sich selbst abbildet).

2. Sei $r \in \mathbb{N}$. Wie oben sei $C^r(I)$ der $\mathbb{R}$-Vektorraum der r-mal stetig differenzierbaren Funktionen auf dem offenen Intervall I. Wir bezeichnen mit f' die Ableitung von $f \in \mathcal{C}^r(I)$. Dann ist

$$\begin{aligned} C^r(I) &\longrightarrow C^{r-1}(I), \\ f &\longmapsto f', \end{aligned} \tag{1.4}$$

eine lineare Abbildung, denn bekanntlich gilt $(f + g)' = f' + g'$ und $(af)' = af'$, $a \in \mathbb{R}$. Dieser Vektorraumhomomorphismus ist nicht injektiv, denn alle konstanten Funktionen haben das gleiche Bild (die Nullfunktion). Schränkt man (1.4) auf den Untervektorraum $C^\infty(I)$ ein, so erhält man einen Endomorphismus von $C^\infty(I)$.

3. Sei $C^0([a,b])$ der $\mathbb{R}$-Vektorraum der stetigen Funktionen auf dem Intervall $[a,b] \subset \mathbb{R}$. Bekannte Eigenschaften des Riemann-Integrals besagen, daß die Abbildung

$$\begin{aligned} C^0([a,b]) &\longrightarrow \mathbb{R}, \\ f &\longmapsto \int_a^b f(x)\,dx, \end{aligned}$$

[6] gr. *endos*: innen

linear ist. Natürlich fassen wir hier $\mathbb{R} = \mathbb{R}^1$ als Vektorraum über sich selbst auf. Eine lineare Abbildung von einem K-Vektorraum V in den K-Vektorraum K wird auch eine **Linearform** auf V genannt. Also: *Das Riemann-Integral ist eine Linearform auf den stetigen Funktionen.*

Definition und Feststellung 1.3.9 *Seien V und W zwei K-Vektorräume. Die Menge aller linearen Abbildungen $V \to W$ wird mit*

$$\mathrm{Hom}_K(V, W)$$

bezeichnet, und ist in natürlicher Weise selbst wieder ein K-Vektorraum. Insbesondere gilt dies für die Menge

$$\mathrm{End}_K(V)$$

aller Endomorphismen von V.

Beweis: Für $f, g \in \mathrm{Hom}_K(V, W)$ definieren wir die Summe $f + g$ als die Abbildung

$$(f + g)(v) = f(v) + g(v), \qquad v \in V$$

(die Addition rechts ist die in W). Man prüft sofort nach, daß $f + g$ linear ist, also ein Element von $\mathrm{Hom}_K(V, W)$ ist. Für $a \in K$ und $f \in \mathrm{Hom}_K(V, W)$ definieren wir die Skalarmultiplikation durch

$$(a.f)(v) = a.f(v), \qquad (v \in V)$$

(die Skalarmultiplikation rechts ist die in W). Es ist wieder sehr leicht einzusehen, daß $a.f$ in der Tat ein Element von $\mathrm{Hom}_K(V, W)$ ist. Wir überlassen dem Leser den Nachweis, daß $\mathrm{Hom}_K(V, W)$ mit diesen Operationen tatsächlich zu einem K-Vektorraum wird. ∎

Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Ähnlich wie im Falle von Gruppen definieren wir den **Kern** und das **Bild** von f:

$$\ker(f) := \{v \in V : f(v) = 0\},$$
$$\mathrm{im}(f) := \{w \in W : \text{Es gibt ein } v \in V \text{ mit } f(v) = w\}.$$

Statt $\mathrm{im}(f)$ schreibt man auch $f(V)$. In Analogie zu Feststellung 1.1.6 gilt:

Feststellung 1.3.10 *Sei $f : V \to W$ ein K-Vektorraumhomomorphismus.*

i) $\ker(f)$ *ist ein Untervektorraum von V.*

ii) $\mathrm{im}(f)$ *ist ein Untervektorraum von W.*

Beweis: i) Seien $v_1, v_2 \in \ker(f)$. Aus $f(v_1) = \mathbf{0}$ und $f(v_2) = \mathbf{0}$ folgt $f(v_1 + v_2) = f(v_1) + f(v_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$, also $v_1 + v_2 \in \ker(f)$. Aus $a \in K$ und $v \in \ker(f)$ folgt $f(av) = af(v) = a\mathbf{0} = \mathbf{0}$, also $av \in \ker(f)$.
ii) Seien $w_1, w_2 \in \mathrm{im}(f)$, und $v_1, v_2 \in V$ mit $f(v_i) = w_i$. Es folgt $f(v_1 + v_2) = f(v_1) + f(v_2) = w_1 + w_2$,

also $w_1 + w_2 \in \operatorname{im}(f)$. Sei $a \in K$ und $w \in \operatorname{im}(f)$. Für $v \in V$ mit $f(v) = w$ gilt $f(av) = af(v) = aw$, also ist $aw \in \operatorname{im}(f)$. ∎

Beispiele: 1. Der Kern der Linearform

$$\begin{aligned} K^n &\longrightarrow K, \\ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} &\longmapsto a_1 + \ldots + a_n, \end{aligned}$$

besteht aus allen Spaltenvektoren, für die die Summe der Komponenten null ist. Das Bild ist ganz K.

2. Der Kern der Abbildung (1.4) ist der aus den konstanten Funktionen bestehende Untervektorraum von $C^r(I)$.

Ein Vektorraumhomomorphismus ist insbesondere ein Homomorphismus abelscher Gruppen. Die folgende Feststellung folgt deshalb aus Feststellung 1.1.8 (oder kann vom Leser leicht direkt bewiesen werden).

Feststellung 1.3.11 *Sei $f : V \to W$ ein Homomorphismus von K-Vektorräumen.*

i) f ist injektiv genau dann, wenn $\ker(f) = \{0\}$.

ii) f ist surjektiv genau dann, wenn $\operatorname{im}(f) = W$.

Die Aussage i) wird sehr häufig als Test für Injektivität benutzt.

Komposition linearer Abbildungen

Feststellung 1.3.12 *Die Komposition (Hintereinanderschaltung) linearer Abbildungen ist wieder linear: Ist $g \in \operatorname{Hom}_K(U, V)$ und $f \in \operatorname{Hom}_K(V, W)$,*

$$U \xrightarrow{g} V \xrightarrow{f} W,$$

so ist $f \circ g \in \operatorname{Hom}_K(U, W)$. Insbesondere haben wir eine Abbildung

$$\begin{aligned} \operatorname{End}_K(V) \times \operatorname{End}_K(V) &\longrightarrow \operatorname{End}_K(V), \\ (f, g) &\longmapsto f \circ g. \end{aligned} \tag{1.5}$$

Beweis: Dies folgt einfach aus

$$\begin{aligned} (f \circ g)(u_1 + u_2) &= f(g(u_1 + u_2)) = f(g(u_1) + g(u_2)) = f(g(u_1)) + f(g(u_2)) \\ &= (f \circ g)(u_1) + (f \circ g)(u_2) \end{aligned}$$

und

$$(f \circ g)(\lambda u) = f(g(\lambda u)) = f(\lambda g(u)) = \lambda f(g(u)) = \lambda((f \circ g)(u)).$$

∎

Bemerkung 1.3.13 Oft schreibt man einfach fg anstatt $f \circ g$ für die Komposition zweier linearer Abbildungen f und g.

Die Abbildung (1.5) definiert (neben Addition und skalarer Multiplikation) eine weitere Verknüpfung auf $\mathrm{End}_K(V)$, eine „Multiplikation“, durch die dieser Vektorraum eine reichhaltigere algebraische Struktur bekommt. Dazu folgende Definition.

Definition 1.3.14 *Sei V ein K-Vektorraum, auf dem zusätzlich eine Multiplikation*

$$\cdot : \; V \times V \longrightarrow V$$

*definiert ist, durch die V zu einem Ring wird. Dann heißt V eine K-**Algebra**, wenn die Multiplikation mit der Skalarmultiplikation in folgendem Sinne verträglich ist:*

$$\lambda.(v \cdot w) = (\lambda.v) \cdot w = v \cdot (\lambda.w) \qquad \text{für alle } v, w \in V \text{ und } \lambda \in K. \tag{1.6}$$

*Ein **Einselement** der K-Algebra V ist ein Einselement des* Ringes *V.*

Feststellung 1.3.15 *Der Vektorraum $\mathrm{End}_K(V)$ aller Endomorphismen des K-Vektorraumes V wird mit der Verknüpfung (1.5) zu einer K-Algebra. Diese besitzt ein Einselement, nämlich die Identität id_V.*

Beweis: Zunächst ist nachzuprüfen, daß $\mathrm{End}_K(V)$ mit „$+$“ und „$\circ$“ zu einem Ring mit Eins wird. Wir überlassen die einfachen Details dem Leser, und verifizieren stattdessen die Verträglichkeitseigenschaft (1.6): Seien $\lambda \in K$ und $f, g \in \mathrm{End}_K(V)$ gegeben. Für alle $v \in V$ gilt dann

$$(\lambda(f \circ g))(v) = \lambda((f \circ g)(v)) = \lambda(f(g(v))) = (\lambda f)(g(v)) = ((\lambda f) \circ g)(v),$$

also $\lambda(f \circ g) = (\lambda f) \circ g$. Ebenso

$$(\lambda(f \circ g))(v) = \lambda(f(g(v))) = f(\lambda g(v)) = f((\lambda g)(v)) = (f \circ (\lambda g))(v),$$

also $\lambda(f \circ g) = f \circ (\lambda g)$. ∎

Definition 1.3.16 *Sei V ein K-Vektorraum. Die Einheitengruppe (siehe 1.2.5) des Ringes $\mathrm{End}_K(V)$ wird mit $\mathrm{GL}(V)$ (**G**eneral **L**inear group) oder $\mathrm{Aut}(V)$ bezeichnet. Die Elemente von $\mathrm{GL}(V)$ heißen auch **Automorphismen** von V.*

Man überlegt sich leicht, daß $\mathrm{GL}(V)$ die Gruppe aller *bijektiven* linearen Abbildungen $V \to V$ ist, mit der Komposition als Verknüpfung (siehe Feststellung B.1.3 in Anhang B.1).

Exakte Sequenzen

Definition 1.3.17 *Seien A, B, C Gruppen (oder Ringe, oder K-Vektorräume), und seien*

$$A \xrightarrow{f} B \xrightarrow{g} C \tag{1.7}$$

Homomorphismen. Die Sequenz (1.7) heißt **exakt**, *falls* $\mathrm{im}(f) = \ker(g)$. *Allgemeiner heißt eine beliebige (endliche oder unendliche) Sequenz von Homomorphismen*

$$\ldots \longrightarrow A_{i-1} \xrightarrow{f_i} A_i \xrightarrow{f_{i+1}} A_{i+1} \longrightarrow \ldots$$

exakt, falls $\mathrm{im}(f_i) = \ker(f_{i+1})$ *für alle* i *(außer am linken und rechten Rand der Sequenz).*

Beispiel: Sei $n \in \mathbb{N}$, und sei $f : \mathbb{Z} \to \mathbb{Z}$ der Homomorphismus abelscher Gruppen $f(k) = nk$. Sei g die Projektion (1.2). Dann ist

$$\mathbb{Z} \xrightarrow{f} \mathbb{Z} \xrightarrow{g} \mathbb{Z}/n\mathbb{Z} \tag{1.8}$$

eine exakte Sequenz.

Feststellung 1.3.18 *Es sei* 0 *entweder die triviale Gruppe, der triviale Ring, oder der Nullvektorraum. Sei* $f : A \to B$ *ein Homomorphismus.*

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

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

exakt ist.

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

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

exakt ist.

iii) f *ist genau dann ein Isomorphismus, wenn die Sequenz*

$$0 \longrightarrow A \xrightarrow{f} B \longrightarrow 0.$$

exakt ist.

Beweis: i) Die Exaktheit der Sequenz bedeutet, daß ker(f) gleich dem Bild der Abbildung $0 \to A$ ist, welches nur aus dem neutralen Element in A besteht. Wir haben früher gesehen, daß dies gleichbedeutend mit der Injektivität von f ist (siehe Feststellung 1.1.8 und Feststellung 1.3.11).

ii) Der Kern von $B \to 0$ ist B. Die Exaktheit der Sequenz bedeutet also $\mathrm{im}(f) = B$. Dies heißt gerade, daß f surjektiv ist.

iii) folgt aus i) und ii). ∎

Beispiel: Die exakte Sequenz (1.8) läßt sich zu der exakten Sequenz

$$0 \longrightarrow \mathbb{Z} \xrightarrow{f} \mathbb{Z} \xrightarrow{g} \mathbb{Z}/n\mathbb{Z} \longrightarrow 0. \tag{1.9}$$

verlängern, denn f ist injektiv und g ist surjektiv.

Exakte Sequenzen der Form

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0 \tag{1.10}$$

treten sehr häufig auf; man nennt sie **kurze exakte Sequenzen**. (1.10) ist genau dann exakt, wenn f injektiv, g surjektiv, und $\mathrm{im}(f) = \ker(g)$ ist.

Beispiel: Der Kern des surjektiven Gruppenhomomorphismus $\exp : \mathbb{C} \to \mathbb{C}^*$ ist $2\pi i\mathbb{Z}$. Dies läßt sich durch die kurze exakte Sequenz

$$0 \longrightarrow 2\pi i\mathbb{Z} \longrightarrow \mathbb{C} \xrightarrow{\exp} \mathbb{C}^* \longrightarrow 1$$

abelscher Gruppen ausdrücken.

Die formalen Überlegungen dieses Unterabschnitts sind nicht auf Gruppen, Ringe und Vektorräume beschränkt, sondern lassen sich in allgemeineren *Kategorien* durchführen. Dem Leser sei an dieser Stelle die Lektüre des Anhangs B.1 nahegelegt.

1.4 Basis und Dimension

Wir nehmen jetzt das systematische Studium von Vektorräumen und linearen Abbildungen auf. Die Definitionen und Aussagen dieses Abschnitts sind grundlegend für alles Folgende.

Um im folgenden präzise sein zu können, benötigen wir den Begriff der *Familie*. Man führt diesen technischen Begriff zu dem einzigen Zweck ein, um von „Mengen mit Wiederholungen von Elementen“ reden zu können.

Sei M eine beliebige Menge. Eine **Familie** in M ist formal eine Abbildung $I \to M$, wobei I eine weitere Menge, die *Indexmenge*, ist. In der Regel schreiben wir $(m_i)_{i\in I}$ für eine Familie, wobei m_i das Bild von $i \in I$ ist.

Jeder Familie in M ist eine Teilmenge von M zugeordnet, nämlich ihr Bild. Man führt Familien ein, um Wiederholungen von Elementen zuzulassen, was in einer Menge nicht möglich ist. Wenn z.B. m, m' unterschiedliche Elemente von M sind, so ist

$$(m, m', m)$$

eine Familie in M (die Indexmenge ist $I = \{1, 2, 3\}$), und die ihr zugeordnete Menge ist $\{m, m'\}$.

Jede Teilmenge N von M läßt sich als Familie auffassen, nämlich als die Inklusion $N \hookrightarrow M$. In diesem Sinne ist „Familie in M“ eine Verallgemeinerung des Begriffs „Teilmenge von M“.

Lineare Hüllen

Es sei K ein beliebiger Körper.

Definition 1.4.1 *Sei V ein K-Vektorraum, und sei $S = (v_i)_{i\in I}$ eine Familie in V. Der Schnitt aller Untervektorräume von V, die jedes v_i enthalten, ist selbst ein K-Vektorraum nach Feststellung 1.3.6.*

Er heißt der von S (oder der von den v_i) **aufgespannte Raum**, *oder die* **lineare Hülle** *der v_i, und wird mit* $\mathrm{span}_K(S)$ *bezeichnet. Falls klar ist, daß man vom Körper K spricht, schreibt man einfach* $\mathrm{span}(S)$. *Es wird auch das Symbol*

$$\langle v_i : i \in I\rangle = \mathrm{span}(S)$$

verwendet. Gilt $\mathrm{span}(S) = V$, *so heißt S ein* **Erzeugendensystem** *von V.*

Der von einer Familie aufgespannte Raum läßt sich explizit beschreiben:

Feststellung 1.4.2 *Für eine Familie $S = (v_i)_{i\in I}$ in dem K-Vektorraum V gilt*

$$\mathrm{span}(S) = \{a_1 v_{i_1} + \ldots + a_n v_{i_n} : a_j \in K,\ i_j \in I,\ n \in \mathbb{N}\}.$$

Beweis: Sei W die Menge auf der rechten Seite. „$\subset$“: Man sieht sofort, daß W ein Untervektorraum von V ist. Überdies gilt $v_i \in W$ für alle i. Da $\mathrm{span}(S)$ der Schnitt aller Untervektorräume ist, die jedes v_i enthalten, folgt $\mathrm{span}(S) \subset W$.

„$\supset$“: Als Untervektorraum von V ist $\mathrm{span}(S)$ abgeschlossen unter Bildung von Summen und skalaren Vielfachen. Aus $v_{i_1}, \ldots, v_{i_n} \in \mathrm{span}(S)$ folgt daher $a_1 v_{i_1} + \ldots a_n v_{i_n} \in \mathrm{span}(S)$. Also ist $W \subset \mathrm{span}(S)$.■

Falls $S = (v_1, \ldots, v_n)$ endlich ist, so schreibt man aufgrund dieser Feststellung auch

$$\mathrm{span}(S) = Kv_1 + \ldots + Kv_n.$$

Definition 1.4.3 *Sei V ein K-Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren in V. Ein Ausdruck der Form*

$$a_1 v_{i_1} + \ldots + a_n v_{i_n}, \qquad a_j \in K,\ i_j \in I,$$

heißt eine **Linearkombination** *der v_i.*

Feststellung 1.4.2 besagt gerade, daß $\langle v_i : i \in I\rangle$ aus allen Linearkombinationen der v_i besteht.

Beispiel 1.4.4 Man betrachte die (Familie der) Vektoren

$$v_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \qquad v_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \qquad v_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}, \qquad v_4 = \begin{pmatrix}1\\1\\0\end{pmatrix}$$

im K^3.

- v_1, v_2, v_3 spannen den ganzen K^3 auf (sind also ein Erzeugendensystem), denn ein beliebiger Vektor läßt sich als Linearkombination schreiben:

$$\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix} = a_1 v_1 + a_2 v_2 + a_3 v_3.$$

- v_1, v_3, v_4 spannen ebenfalls den ganzen K^3 auf, denn

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = (a_1 - a_2)v_1 + a_3v_3 + a_2v_4.$$

  Man überlege sich, daß auch v_2, v_3, v_4 ein Erzeugendensystem des K^3 ist.

- v_1, v_2, v_4 spannen nicht ganz K^3 auf, denn jede Linearkombination dieser drei Vektoren hat eine Null in der dritten Komponente.

- Je zwei der Vektoren $v_1, \ldots, v_4$ spannen einen Raum auf, der isomorph zu K^2 ist, aber niemals den ganzen K^3.

Bemerkung 1.4.5 Wenn $K = \mathbb{R}$ und $V = \mathbb{R}^2$ oder $V = \mathbb{R}^3$ ist, so läßt sich der von einer Familie von Vektoren aufgespannte Untervektorraum leicht geometrisch veranschaulichen.

i) Der von einem einzigen Vektor $v \neq 0$ aufgespannte Unterraum $\langle v \rangle$ ist die Gerade durch den Ursprung und den Punkt v.

ii) Bei zwei Vektoren v_1, v_2, die nicht ausgerechnet beide Null sein sollen, können zwei Fälle auftreten: Entweder liegen v_1, v_2 und der Ursprung auf einer Geraden; dann ist $\langle v_1, v_2 \rangle$ gleich dieser Geraden. Oder dies ist nicht der Fall; dann ist $\langle v_1, v_2 \rangle$ die eindeutig bestimmte Ebene, die v_1, v_2 und den Ursprung enthält.

iii) Falls drei Vektoren $v_1, v_2, v_3 \in \mathbb{R}^3$ auf einer gemeinsamen Ursprungsgeraden liegen (und nicht alle drei gleich Null sind), spannen sie diese Gerade auf. Falls dies nicht der Fall ist, könnte es eine Ebene geben, die alle drei Vektoren und den Ursprung enthält; diese wird dann von v_1, v_2, v_3 aufgespannt. Ist dies auch nicht der Fall, so sind v_1, v_2, v_3 ein Erzeugendensystem des $\mathbb{R}^3$.

Es ist wichtig, diese geometrischen Interpretationen zu beherrschen. Der Leser setze in Beispiel 1.4.4 $K = \mathbb{R}$, und mache sich die Verhältnisse anhand von Bildern klar.

Lineare Unabhängigkeit und Basen

Definition 1.4.6 *Sei $(v_i)_{i \in I}$ eine Familie in einem K-Vektorraum V. Die Familie heißt* **linear unabhängig** *(über K), wenn sich die Null nur als triviale Linearkombination der v_i darstellen läßt, wenn also für alle $a_1, \ldots, a_n \in K$ und alle $i_1, \ldots, i_n \in I$ gilt: Aus*

$$a_1 v_{i_1} + \ldots + a_n v_{i_n} = 0$$

folgt $a_1 = \ldots = a_n = 0$. Eine Familie, die nicht linear unabhängig ist, heißt **linear abhängig**.

Beispiel: Je zwei der Vektoren $v_1, \ldots, v_4$ aus Beispiel 1.4.4 sind linear unabhängig. Mit der einen Ausnahme (v_1, v_2, v_4) sind je drei der Vektoren linear unabhängig. Die Familie (v_1, v_2, v_3, v_4) ist linear abhängig.

Eine Familie $(v_i)_{i \in I}$ ist offenbar immer dann linear abhängig, wenn sie den Nullvektor enthält (d.h. $v_i = 0$ für ein $i \in I$), oder zwei gleiche Vektoren (d.h., falls es $i \neq j$ in I gibt mit $v_i = v_j$).

Definition und Feststellung 1.4.7 *Sei V ein K-Vektorraum. Für eine Familie $S = (v_i)_{i\in I}$ in V sind die folgenden Bedingungen äquivalent:*

i) S ist ein linear unabhängiges Erzeugendensystem von V.

ii) S ist eine maximal linear unabhängige Familie (maximal heißt dabei: Ist $I \subsetneq J$, und $S' : J \to V$ eine Familie, deren Einschränkung auf I gleich S ist, so ist S' nicht mehr linear unabhängig).

iii) S ist ein minimales Erzeugendensystem von V (minimal heißt dabei: Ist $J \subsetneq I$, und $S' : J \to V$ die Einschränkung von S auf J, so ist S' kein Erzeugendensystem von V mehr).

iv) Jedes Element von V läßt sich eindeutig als Linearkombination der Familie $(v_i)_{i\in I}$ schreiben.

Besitzt S diese Eigenschaften, so heißt S eine **Basis** *von V.*

Beweis: i) $\Rightarrow$ iii): Sei S ein linear unabhängiges Erzeugendensystem von V. Sei S' die Familie, die durch weglassen eines Vektors v_i, $i \in I$, entsteht. Wir müssen zeigen, daß S' kein Erzeugendensystem mehr ist. Angenommen, dies wäre doch der Fall. Dann ließe sich v_i als Linearkombination von S' schreiben:

$$v_i = a_1 v_{i_1} + \ldots + a_n v_{i_n}, \qquad i_j \in I \setminus \{i\}.$$

Dies liefert aber eine nicht-triviale Linearkombination in S, die den Nullvektor darstellt, im Widerspruch zur linearen Unabhängigkeit von S.

iii) $\Rightarrow$ iv): Da S ein Erzeugendensystem ist, läßt sich jedes $v \in V$ als Linearkombination in S schreiben. Angenommen, dies wäre nicht eindeutig, wir hätten also für ein $v \in V$ zwei Möglichkeiten:

$$a_1 v_{i_1} + \ldots + a_n v_{i_n} = v = b_1 v_{i_1} + \ldots + b_n v_{i_n}$$

(wir können die gleichen v_{i_j} verwenden, indem wir gegebenenfalls mit Koeffizienten ergänzen, die Null sind). Die beiden Linearkombinationen sollen nicht gleich sein; sei etwa $a_1 \neq b_1$. Dann folgt

$$v_{i_1} = -(a_1 - b_1)^{-1}((a_2 - b_2)v_{i_2} + \ldots + (a_n - b_n)v_{i_n}).$$

(Hier benutzen wir die Tatsache, daß K ein Körper ist, und nicht nur ein Ring). v_{i_1} ist also eine Linearkombination der Familie S', die aus S durch Weglassen des Elementes v_{i_1} entsteht. Dann ist aber offenbar S' bereits ein Erzeugendensystem von V, im Widerspruch zur Minimalität von S.

iv) $\Rightarrow$ ii): Annahme: S ist linear abhängig. Dann gibt es eine nicht-triviale Linearkombination, die Null darstellt: $a_1 v_{i_1} + \ldots + a_n v_{i_n} = 0$. Da auch die triviale Linearkombination Null darstellt, ist dies ein Widerspruch zur Eindeutigkeitsaussage in iv). Also ist S doch linear unabhängig. – Angenommen, es gäbe eine größere linear unabhängige Familie $S' : J \to V$, $J \supsetneq I$. Ein v_j mit $j \in J \setminus I$ läßt sich aber nach iv) als Linearkombination in S darstellen. Dies steht im Widerspruch zur linearen Unabhängigkeit von S'.

ii) $\Rightarrow$ i): Wenn S kein Erzeugendensystem von V wäre, so gäbe es einen Vektor $v \in V$, der sich nicht als Linearkombination von S schreiben läßt. Dann ließe sich aber durch Hinzunahme von v zu S eine größere linear unabhängige Familie konstruieren, im Widerspruch zur Maximalitätsaussage in ii). ∎

Beispiel 1.4.8 In K^n betrachten wir die Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Dann ist $(e_1, \ldots, e_n)$ eine Basis des K^n, denn jeder Vektor in K^n läßt sich eindeutig als Linearkombination der e_i darstellen:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1 e_1 + \ldots + a_n v_n.$$

Diese Basis heißt die **kanonische Basis** des K^n.

In einer Basis $(v_i)_{i \in I}$ dürfen offenbar keine Wiederholungen von Vektoren auftreten. Die *Familie* $(v_i)_{i \in I}$ kann also mit der *Menge* $\{v_i\}_{i \in I}$ identifiziert werden.

Sehr häufig benutzt wird folgende Feststellung, die ein Licht auf die Bedeutung von Basen wirft.

Feststellung 1.4.9 *Seien V und W zwei K-Vektorräume, und $(v_i)_{i \in I}$ eine Basis von V. Für jedes $i \in I$ sei ein $w_i \in W$ gegeben. Dann gibt es genau eine lineare Abbildung $f : V \to W$ mit der Eigenschaft*

$$f(v_i) = w_i \qquad \text{für alle } i \in I.$$

Kurz: Man darf die Bilder einer Basis beliebig vorgeben.

Beweis: Wir zeigen zuerst die Eindeutigkeit von f. Sei f' eine weitere lineare Abbildung $V \to W$ mit der geforderten Eigenschaft. Ein beliebiger Vektor $v \in V$ läßt sich als Linearkombination der Basis schreiben: $v = a_1 v_1 + \ldots + a_n v_n$. Dann folgt

$$f(v) = a_1 f(v_1) + \ldots + a_n f(v_n) = a_1 f'(v_1) + \ldots + a_n f'(v_n) = f'(v).$$

Also ist $f = f'$. Dies zeigt die Eindeutigkeit.

Zur Existenz: Es bleibt keine andere Wahl, als f folgendermaßen zu definieren:

$$f(v) := a_1 w_1 + \ldots + a_n w_n, \qquad \text{falls } v = a_1 v_1 + \ldots + a_n v_n.$$

Da sich jedes $v \in V$ eindeutig als Linearkombination der Basis schreiben läßt, ist f eine wohldefinierte Abbildung. Der Leser kann leicht bestätigen, daß f linear ist. Schließlich gilt $f(v_i) = w_i$ nach Konstruktion. ∎

Der Beweis des folgenden wichtigen Satzes benutzt das Zornsche Lemma aus der Mengenlehre; siehe Anhang A.2. Der Leser nehme die Aussage und insbesondere die anschließenden Folgerungen zur Kenntnis, und mag später auf den Beweis zurückkommen.

Satz 1.4.10 *Sei S ein Erzeugendensystem des K-Vektorraumes V, und T eine linear unabhängige Teilmenge von S. Dann gibt es eine Basis von V, die T enthält und in S enthalten ist.*

Beweis: Sei $\mathcal{M}$ die Menge aller linear unabhängigen Teilfamilien (notwendig Teilmengen) von S, die T enthalten. Dann ist $\mathcal{M}$ nicht leer, denn $T \in \mathcal{M}$. Falls $S', S'' \in \mathcal{M}$, und S' eine Teilfamilie von S'' ist, so schreiben wir $S' \leq S''$. Dadurch wird $\mathcal{M}$ zu einer teilweise geordneten Menge.

Sei $M \subset \mathcal{M}$ eine linear geordnete Teilmenge. Wir konstruieren eine obere Schranke von M in $\mathcal{M}$. Sei dazu $I_0 \subset I$ die Vereinigung aller Indexmengen I', wobei $S' : I' \to V$ die Elemente von M durchläuft. Wir definieren eine Familie $S_0 : I_0 \to V$. Jedes $i \in I_0$ liegt in einem I', wobei $S' = (v_i)_{i \in I'}$ ein Element von M ist. Dann sei $S_0(i) := v_i$. Man sieht leicht, daß $S_0(i)$ von der Wahl von S' unabhängig ist (der Grund ist die *lineare* Ordnung von M).

Wir haben eine Familie S_0 konstruiert, und zeigen, daß sie linear unabhängig ist. Sei

$$a_1 v_{i_1} + \ldots + a_n v_{i_n} = 0 \tag{1.11}$$

eine Linearkombination in S_0. Jedes i_j liegt in einer der Indexmengen I', und da M linear geordnet ist, finden wir ein I', das alle i_j enthält. Dieses I' ist die Indexmenge einer Familie $S' \in M$, und (1.11) ist eine Linearkombination von S'. Da S' linear unabhängig ist, folgt $a_i = 0$ für alle i. Dies zeigt die lineare Unabhängigkeit von S_0.

Da überdies $T \subset S_0$, ist S_0 ein Element von $\mathcal{M}$. Nach Konstruktion ist S_0 eine obere Schranke von M. Wir haben insgesamt gezeigt, daß jede linear geordnete Teilmenge von $\mathcal{M}$ eine obere Schranke besitzt. Das Zornsche Lemma A.2.1 besagt dann, daß es in $\mathcal{M}$ ein maximales Element B gibt. Wir zeigen, daß B eine Basis von V ist, womit der Beweis beendet ist.

Nach Konstruktion ist B linear unabhängig. Es bleibt zu zeigen, daß B ein Erzeugendensystem von V ist. Dazu genügt zu zeigen, daß jedes Element von S eine Linearkombination von B ist (denn jedes Element von V ist Linearkombination von S). Angenommen aber, ein $v \in S$ wäre keine Linearkombination von B. Dann ist $B \cup \{v\}$ eine größere linear unabhängige Familie in S als B. Dies steht im Widerspruch zur Maximalität von B. ■

Folgerung 1.4.11 (Basisauswahlsatz) *Jedes Erzeugendensystem eines K-Vektorraumes V enthält eine Basis von V.*

Folgerung 1.4.12 (Basisergänzungssatz) *Jede linear unabhängige Teilmenge eines K-Vektorraumes V läßt sich zu einer Basis von V ergänzen.*

Folgerung 1.4.13 (Basisexistenzsatz) *Jeder Vektorraum besitzt eine Basis.*

Dimension

Satz 1.4.15 weiter unten besagt, daß je zwei Basen eines Vektorraumes gleich viele Elemente besitzen, wobei „gleich viel" für unendliche Mengen bedeutet, daß sie die gleiche *Kardinalität* besitzen; siehe Anhang A.2. Wir benötigen die Aussage nur im endlichen Fall, und der Leser mag den Teil b) des Beweises von 1.4.15 auf einen späteren Zeitpunkt verschieben. Für Teil a) des Beweises schicken wir folgendes Lemma voraus.

Lemma 1.4.14 *Sei V ein K-Vektorraum, der eine endliche Basis $(v_1, \dots, v_n)$ besitzt. Sei $1 \leq m < n$, und $w \in V$ ein Vektor, der von $(v_1, \dots, v_m)$ linear unabhängig ist (d.h., $(v_1, \dots, v_m, w)$ ist linear unabhängig). Dann gibt es ein $j \in \{m+1, \dots, n\}$, so daß*

$$(v_1, \dots, v_m, \dots v_{j-1}, w, v_{j+1}, \dots, v_n)$$

eine Basis von V ist.

Beweis: Da $(v_1, \dots, v_n)$ eine Basis ist, gibt es eine eindeutig bestimmte Linearkombination

$$w = a_1 v_1 + \dots + a_n v_n. \tag{1.12}$$

Es kann nicht sein, daß $a_{m+1} = \dots = a_n = 0$, denn dann wäre w von $(v_1, \dots, v_m)$ linear abhängig. Nach Umnumerierung können wir annehmen, daß $a_n \neq 0$. Wir behaupten, daß

$$S = (v_1, \dots, v_{n-1}, w)$$

eine Basis von V ist. Es ist $v_n = a_n^{-1} w - a_n^{-1} a_1 v_1 - \dots - a_n^{-1} a_{n-1} v_{n-1}$ eine Linearkombination von S. Da $(v_1, \dots, v_n)$ ein Erzeugendensystem von V ist, ist daher auch S ein solches.

Es bleibt zu zeigen, daß S linear unabhängig ist. Sei

$$b_1 v_1 + \dots + b_{n-1} v_{n-1} + bw = 0$$

eine Linearkombination von S. Hier muß $b = 0$ sein, denn ansonsten ließe sich w als Linearkombination von $(v_1, \dots, v_{n-1})$ schreiben, im Widerspruch zu $a_n \neq 0$ in (1.12). Wenn $b = 0$, so folgt aber $b_1 = \dots = b_{n-1} = 0$ wegen der linearen Unabhängigkeit der v_i. Wir haben gezeigt, daß nur die triviale Linearkombination von S die Null darstellt. ∎

Satz 1.4.15 *Je zwei Basen eines K-Vektorraumes haben gleiche Kardinalität.*

Beweis: a) Wir nehmen zunächst an, daß V eine endliche Basis $S = (v_1, \dots, v_n)$ besitzt. Es genügt dann zu zeigen, daß jede andere Basis B nicht mehr als n Elemente besitzt. Nehmen wir im Widerspruch dazu an, es gäbe $n+1$ linear unabhängige Vektoren $w_1, \dots, w_{n+1}$. Nach Lemma 1.4.14 kann man aus S eine neue Basis S_1 gewinnen, indem man eines der v_i durch w_1 austauscht. Aus S_1 gewinnt man dann eine weitere Basis S_2, indem man ein weiteres v_i durch w_2 austauscht. So fortfahrend erhält man eine Basis S_n aus n der w_j's. Dann kann aber $w_1, \dots, w_{n+1}$ nicht linear unabhängig sein, Widerspruch.

b) Wir nehmen jetzt an, daß jede Basis von V unendlich viele Elemente besitzt. Seien S und T zwei Basen von V. Jedes $v \in S$ kann eindeutig als Linearkombination von T dargestellt werden; sei T_v die endliche Menge der daran beteiligten Elemente von T. Dann ist

$$T = \bigcup_{v \in S} T_v, \tag{1.13}$$

denn angenommen, dies wäre nicht der Fall. Es gibt dann ein $w \in T$, das in keinem T_v enthalten ist. Da S eine Basis ist, gibt es eine Linearkombination

$$w = a_1 v_1 + \dots a_n v_n, \qquad a_i \in K,\ v_i \in S.$$

Jedes v_i ist wiederum Linearkombination von T_{v_i}. Insgesamt ist w also eine Linearkombination von $T_{v_1} \cup \ldots \cup T_{v_n}$. Dies ist ein Widerspruch zur linearen Unabhängigkeit von T. Damit ist (1.13) gezeigt.

Satz A.2.2 aus Anhang A liefert nun $\mathrm{card}(S) \geq \mathrm{card}(T)$. Wir können die Rollen von S und T vertauschen, und erhalten die Gleichheit der Kardinalitäten. ∎

Aufgrund der Sätze 1.4.13 und 1.4.15 kann man jedem Vektorraum eine Invariante zuordnen, nämlich die **Länge** einer Basis:

Definition 1.4.16 *Sei V ein Vektorraum über dem Körper K. Die* **Dimension** *(genauer: die K-Dimension) von V ist folgendermaßen definiert.*

i) Der Nullvektorraum $V = 0$ bekommt die Dimension Null. In Zeichen: $\dim(V) = 0$.

ii) Besitzt V eine Basis aus n Elementen, $n \in \mathbb{N}$, so bekommt V die Dimension n. In Zeichen: $\dim(V) = n$.

iii) Hat eine (dann jede) Basis von V unendlich viele Elemente, so bekommt V die Dimension „unendlich". In Zeichen: $\dim(V) = \infty$.

Manchmal schreiben wir genauer $\dim_K(V)$ *anstatt* $\dim(V)$.

Beispiele: 1. $\dim(K^n) = n$, siehe Beispiel 1.4.8.

2. Jede komplexe Zahl läßt sich eindeutig als $x + iy$ mit $x, y \in \mathbb{R}$ schreiben. Dies zeigt, daß $\mathbb{C}$ ein zwei-dimensionaler $\mathbb{R}$-Vektorraum ist, mit Basis $(1, i)$.

3. Funktionenräume wie $C^k(I)$, $I \subset \mathbb{R}$ ein offenes Intervall, sind in der Regel unendlich-dimensional. Der Leser kann leicht unendlich viele linear unabhängige Elemente finden.

1.5 Summen und Quotienten

Endliche direkte Summen

Seien V und W zwei K-Vektorräume. Die Menge $V \times W$ aller Paare (v, w) mit $v \in V$, $w \in W$ ist in natürlicher Weise wieder ein K-Vektorraum: Zunächst ist $V \times W$ eine abelsche Gruppe durch komponentenweise Addition,

$$(v, w) + (v', w') := (v + v', w + w') \qquad (v, v' \in V,\ w, w' \in W),$$

und wird durch komponentenweise Skalarmultiplikation

$$a(v, w) := (av, aw) \qquad (a \in K,\ v \in V,\ w \in W)$$

zum K-Vektorraum.

Definition 1.5.1 *Die Menge $V \times W$ mit obiger K-Vektorraumstruktur wird mit $V \oplus W$ bezeichnet, und heißt die* **direkte Summe** *von V und W.*

Beispiele: 1. Seien $m, n \in \mathbb{N}$. Dann gibt es einen natürlichen Isomorphismus

$$\begin{aligned} K^m \oplus K^n &\xrightarrow{\sim} K^{m+n}, \\ ((a_1, \ldots, a_m), (b_1, \ldots, b_n)) &\longmapsto (a_1, \ldots, a_m, b_1, \ldots, b_n). \end{aligned}$$

2. Die direkte Summe mit dem Nullvektorraum liefert nichts neues:

$$V \oplus \{0\} \simeq V \simeq \{0\} \oplus V.$$

Bemerkung 1.5.2 Man kann V als Untervektorraum von $V \oplus W$ auffassen, indem man $v \in V$ mit $(v, 0) \in V \oplus W$ identifiziert. Ebenso ist W ein Untervektorraum von $V \oplus W$, via der Identifikation $w = (0, w)$. Wir werden dies häufig verwenden.

Die linearen Abbildungen

$$\begin{aligned} V \oplus W &\longrightarrow V \\ (v, w) &\longmapsto v \end{aligned} \qquad\qquad \begin{aligned} V \oplus W &\longrightarrow W \\ (v, w) &\longmapsto w \end{aligned}$$

heißen die **Projektionen** auf V bzw. W. Zusammen mit den Inklusionen von V bzw. W haben wir exakte Sequenzen

$$0 \longrightarrow V \longrightarrow V \oplus W \longrightarrow W \longrightarrow 0, \tag{1.14}$$

$$0 \longrightarrow W \longrightarrow V \oplus W \longrightarrow V \longrightarrow 0. \tag{1.15}$$

Etwas allgemeiner kann die direkte Summe von mehr als zwei Vektorräumen definiert werden. Wenn $V_1, \ldots, V_n$ Vektorräume über dem Körper K sind, so ist die direkte Summe

$$V_1 \oplus \ldots \oplus V_n$$

der K-Vektorraum, dessen Elemente n-Tupel $(v_1, \ldots, v_n)$ mit $v_i \in V_i$ sind, die komponentenweise addiert und mit Skalaren multipliziert werden. Es gibt dann eine Projektion auf jedes V_i, und n exakte Sequenzen

$$0 \longrightarrow V_1 \oplus \ldots \oplus V_{i-1} \oplus V_{i+1} \oplus \ldots \oplus V_n \longrightarrow V_1 \oplus \ldots \oplus V_n \longrightarrow V_i \longrightarrow 0.$$

Für die n-fache direkte Summe eines Vektorraumes V mit sich selbst schreiben wir auch V^n:

$$V^n := \underbrace{V \oplus \ldots \oplus V}_{n\text{-mal}}.$$

Im Falle $V = K$ liegt dabei kein Bezeichnungskonflikt mit der üblichen Bedeutung des Symbols K^n vor, denn offenbar gilt

$$\underbrace{K \oplus \ldots \oplus K}_{n\text{-mal}} = K^n.$$

Bemerkung 1.5.3 Sind $V_1, \dots, V_n$ Vektorräume über K, und ist $(v_i^j)_{i \in I_j}$ eine Basis von V_j, so ist die Menge der Vektoren

$$(0, \dots, 0, v_i^j, 0, \dots, 0) \in V_1 \oplus \dots \oplus V_n, \qquad i \in I_j,\ j = 1, \dots, n,$$

eine Basis von $V_1 \oplus \dots \oplus V_n$. Wählt man im ein-dimensionalen K-Vektorraum K die Basis 1, so erhält man die kanonische Basis des K^n, siehe 1.4.8.

Seien V_1 und V_2 Untervektorräume des K-Vektorraumes W. Die Menge

$$V_1 + V_2 := \{v_1 + v_2 :\ v_1 \in V_1,\ v_2 \in V_2\} \subset W$$

ist dann offenbar ein Untervektorraum von W, der V_1 und V_2 enthält (und er ist der kleinste mit dieser Eigenschaft). Er heißt die **Summe** von V_1 und V_2. Wir untersuchen den Zusammenhang mit der *direkten* Summe $V_1 \oplus V_2$. Der Kern der linearen Abbildung

$$\begin{aligned} s :\ V_1 \oplus V_2 &\longrightarrow V_1 + V_2, \\ (v_1, v_2) &\longmapsto v_1 + v_2, \end{aligned}$$

besteht aus allen Elementen $(v, -v)$ mit $v \in V_1 \cap V_2$. Mit anderen Worten, wenn $f :\ V_1 \cap V_2 \to V_1 \oplus V_2$ durch $f(v) = (v, -v)$ definiert ist, so haben wir eine exakte Sequenz

$$0 \longrightarrow V_1 \cap V_2 \xrightarrow{f} V_1 \oplus V_2 \xrightarrow{s} V_1 + V_2 \longrightarrow 0. \tag{1.16}$$

Die Abbildung s ist also genau dann ein Isomorphismus, wenn $V_1 \cap V_2 = 0$. In diesem Falle identifizieren wir $V_1 + V_2$ mit $V_1 \oplus V_2$ via s, und sagen, die Summe $V_1 + V_2$ sei *direkt*:

$$V_1 + V_2 = V_1 \oplus V_2, \qquad \text{falls } V_1 \cap V_2 = 0. \tag{1.17}$$

Manchmal wird dann auch von der *inneren* direkten Summe von V_1 und V_2 gesprochen.

Wir verallgemeinern diese Überlegungen auf mehr als zwei Untervektorräume. Seien also $V_1, \dots, V_n$ Untervektorräume von W. Die **Summe** von $V_1, \dots, V_n$ ist der Untervektorraum

$$V_1 + \dots + V_n = \{v_1 + \dots + v_n :\ v_i \in V_i \text{ für } i = 1, \dots, n\}.$$

Wir sagen, die Summe sei *direkt*, wenn die Summationsabbildung

$$\begin{aligned} V_1 \oplus \dots \oplus V_n &\longrightarrow V_1 + \dots + V_n, \\ (v_1, \dots, v_n) &\longmapsto v_1 + \dots + v_n, \end{aligned}$$

ein Isomorphismus ist. Dies ist genau dann der Fall, wenn sich jedes Element von $V_1 + \dots + V_n$ *eindeutig* in der Form $v_1 + \dots + v_n$ mit $v_i \in V_i$ schreiben läßt. Im allgemeinen genügt dafür *nicht*, daß alle Schnitte $V_i \cap V_j$ Null sind. In einer speziellen Situation ist dies aber doch richtig [7]:

[7] Feststellung 1.5.4 wird nur einmal in diesem Buch Verwendung finden, nämlich im Beweis von Satz 4.6.2.

Feststellung 1.5.4 *Seien $f_1, \dots, f_n$ Endomorphismen des K-Vektorraumes W, die paarweise miteinander kommutieren:*

$$f_i f_j = f_j f_i \qquad \text{für alle } i, j = 1, \dots, n.$$

Sei $V_i := \ker(f_i)$. Die Summe $V_1 + \dots + V_n$ ist genau dann direkt, wenn

$$V_i \cap V_j = 0 \qquad \text{für alle } i, j = 1, \dots, n, \ i \neq j.$$

Beweis: Die Bedingung $V_i \cap V_j = 0$ ist offenbar notwendig für die Direktheit der Summe. Wir zeigen, daß sie auch hinreichend ist. Angenommen also, sie sei erfüllt, aber die Summe wäre nicht direkt. Dann gibt es eine nicht-triviale Linearkombination

$$v_1 + \dots + v_n = 0, \qquad v_i \in V_i.$$

Sei die Linearkombination so, daß $\sharp\{i : \ v_i \neq 0\}$ minimal ist. Offenbar muß es mindestens zwei Indizes i, j geben, so daß $v_i \neq 0$ und $v_j \neq 0$. Wir wenden dann f_i auf die Linearkombination an, und erhalten

$$f_i v_1 + \dots + f_i v_n = 0.$$

Hier gilt $f_i v_k \in V_k = \ker(f_k)$ wegen der Voraussetzung $f_i f_k = f_k f_i$. Ausserdem gilt $f_i v_i = 0$ und $f_i v_j \neq 0$, letzteres wegen $V_i \cap V_j = 0$. Wir erhalten also wieder eine nicht-triviale Linearkombination, in der nun aber mehr Nullvektoren vorkommen als in der ursprünglichen. Dies steht im Widerspruch zur ursprünglichen Minimalitätsforderung. ∎

Bemerkung 1.5.5 Ohne genaue Definitionen zu geben merken wir an, daß die direkte Summe $V \oplus W$ die Eigenschaften eines *Produktes* in der *Kategorie* $\mathbf{Vec}_K$ erfüllt (vgl. Anhang B.1). Man kann sich in jeder Kategorie fragen, ob zu je zwei Objekten A und B ein Produkt $A \times B$ existiert. Dies ist z.B. in der Kategorie **Set** der Mengen der Fall: Das Produkt zweier Mengen A und B ist das *kartesische* Produkt. Auch in den Kategorien **Gr**, **Ab** und **Ring** existieren Produkte, die in natürlicher Weise gebildet werden.

Unendliche direkte Summen und Produkte

Sei I irgendeine Indexmenge, und V_i ein K-Vektorraum für jedes $i \in I$. Das **direkte Produkt**

$$\prod_{i \in I} V_i$$

ist der K-Vektorraum, dessen Elemente die Familien

$$(v_i)_{i \in I} \qquad \text{mit } v_i \in V_i$$

sind, die komponentenweise addiert und mit Skalaren multipliziert werden. (Formal ist ein solches Element eine Abbildung $\varphi : I \to \coprod V_i$ mit $\varphi(i) \in V_i$; das Symbol $\coprod$ bedeutet „disjunkte Vereinigung".) Wenn alle V_i gleich einem festen Vektorraum V sind so schreibt man auch

$$V^I := \prod_{i \in I} V.$$

Die **direkte Summe**

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

ist der Untervektorraum von $\prod_{i\in I} V_i$ aller Elemente

$$(v_i)_{i\in I} \qquad \text{mit } v_i = 0 \text{ für fast alle } i \in I.$$

Falls $V_i = V$ für alle i, so schreibt man auch

$$V^{(I)} = \bigoplus_{i\in I} V.$$

Falls I endlich ist, so gibt es keinen Unterschied zwischen der direkten Summe und dem direkten Produkt.

Beispiel: $\mathbb{R}^{\mathbb{N}}$ ist der $\mathbb{R}$-Vektorraum aller Folgen $(a_1, a_2, \dots)$ reeller Zahlen. $\mathbb{R}^{(\mathbb{N})}$ ist der Untervektorraum aller Folgen, in denen nur endlich viele Folgenglieder ungleich Null sind.

Wenn für jedes V_i eine Basis gegeben ist, so kann der Leser daraus leicht eine Basis für $\bigoplus V_i$ gewinnen, siehe 1.5.3. Für $\prod V_i$ ist dies nicht so leicht möglich (wir wissen aber nach Satz 1.4.13 zumindest, daß eine Basis existiert).

Quotienten

Sei V ein K-Vektorraum und $W \subset V$ ein Untervektorraum. Für jedes $v \in V$ betrachten wir die **Nebenklasse**

$$\bar{v} := v + W = \{v + w : w \in W\} \subset V.$$

Sei V/W die Menge aller dieser Nebenklassen. Wie man leicht sieht gilt

$$\bar{v}_1 = \bar{v}_2 \iff v_1 - v_2 \in W.$$

Die Nebenklasse $\overline{v_1 + v_2} = v_1 + v_2 + W$ hängt nur von $\bar{v}_1$ und $\bar{v}_2$ ab, so daß folgende Definition einer Addition auf V/W Sinn macht:

$$\bar{v}_1 + \bar{v}_2 := \overline{v_1 + v_2}. \tag{1.18}$$

Man prüft ohne Schwierigkeiten nach, daß V/W dadurch zu einer abelschen Gruppe wird (das neutrale Element ist $\bar{0}$, und $\overline{-v}$ ist das Inverse von $\bar{v}$).

Wir wollen V/W zu einem K-Vektorraum machen, und definieren dazu die Skalarmultiplikation durch

$$a.\bar{v} := \overline{a.v}, \qquad v \in V,\ a \in K. \tag{1.19}$$

Dies macht Sinn, da $\overline{a.v} = a.v + W$ nur von $\bar{v}$ abhängt. Es ist leicht nachzuprüfen, daß die abelsche Gruppe V/W mit dieser Skalarmultiplikation in der Tat zu einem K-Vektorraum wird.

Definition 1.5.6 *Die Menge V/W aller Nebenklassen $\bar{v} = v + W$ mit der durch (1.18) und (1.19) definierten K-Vektorraumstruktur heißt der* **Quotientenvektorraum** *(oder kurz* **Quotient***) von V modulo W.*

Die Verknüpfungen auf V/W sind so erklärt, daß die surjektive Abbildung

$$\begin{aligned} p: V &\longrightarrow V/W, \\ v &\longmapsto \bar{v}, \end{aligned}$$

linear ist. Der Kern dieser Abbildung, die auch **Projektion** genannt wird, ist gerade W. Mit anderen Worten, wir haben eine kurze exakte Sequenz

$$0 \longrightarrow W \longrightarrow V \xrightarrow{p} V/W \longrightarrow 0. \tag{1.20}$$

Feststellung 1.5.7 (Homomorphiesatz) *Sei $f: V \to V'$ eine lineare Abbildung von K-Vektorräumen. Zu jedem Untervektorraum W von $\ker(f)$ gibt es genau eine lineare Abbildung $\bar{f}: V/W \to V'$, so daß das Diagramm*

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "p"'] & & V' \\ & V/W \arrow[ur, "\bar{f}"'] & \end{tikzcd}$$

kommutativ ist. Falls $W = \ker(\mathrm{f})$, so ist $\bar{f}$ injektiv. Also gilt

$$V/\ker(f) \simeq \mathrm{im}(f).$$

Beispiel: Die Kommutativität des Diagramms bedeutet

$$\bar{f}(\bar{v}) = f(v) \qquad \text{für alle } v \in V. \tag{1.21}$$

Falls $\bar{f}$ überhaupt existiert, ist es also durch f eindeutig festgelegt. Umgekehrt bleibt keine andere Wahl, als $\bar{f}$ durch (1.21) zu *definieren.* Wir müssen die Wohldefiniertheit nachprüfen:

$$\bar{v}_1 = \bar{v}_2 \;\Rightarrow\; v_1 - v_2 \in W \subset \ker(f) \;\Rightarrow\; f(v_1 - v_2) = 0 \;\Rightarrow\; f(v_1) = f(v_2).$$

Also ist (1.21) eine sinnvolle Definition. Man prüft sofort nach, daß $\bar{f}$ dann linear ist. Dies beweist die Existenz.

Sei $W = \ker(f)$. Dann gilt

$$\bar{f}(\bar{v}) = 0 \;\Rightarrow\; f(v) = 0 \;\Rightarrow\; v \in \ker(f) \;\Rightarrow\; \bar{v} = \bar{0}.$$

Der Kern von $\bar{f}$ ist also Null, was äquivalent zur Injektivität von $\bar{f}$ ist. Die letzte Behauptung ist klar, da eine injektive Abbildung immer einen Isomorphismus auf ihr Bild vermittelt. ∎

Beispiele: 1. Projektion auf die ersten n Koordinaten liefert eine Surjektion $K^{m+n} \to K^n$, deren Kern isomorph zu K^m ist. Also gilt $K^{m+n}/K^m \simeq K^n$. Der Leser schreibe diesen Isomorphismus explizit in Koordinaten auf.

2. Sei $(a,b) \subset \mathbb{R}$ ein offenes Intervall, und $x \in (a,b)$ ein Punkt. Sei $C^k(a,b)$ der $\mathbb{R}$-Vektorraum der k-mal stetig differenzierbaren Funktionen. Wir betrachten die Auswertungsabbildung

$$\begin{aligned} f:\ C^k(a,b) &\longrightarrow \mathbb{R}, \\ \varphi &\longmapsto \varphi(x), \end{aligned}$$

die offenbar linear ist. Der Kern W von f besteht aus allen Funktionen $\varphi \in C^k(a,b)$, die bei x den Wert Null annehmen. Also folgt

$$C^k(a,b)/W \simeq \mathbb{R}.$$

Dies ist ein Beispiel für zwei unendlich-dimensionale Vektorräume, deren Quotient endlich-dimensional ist.

Feststellung 1.5.8 *Für K-Vektorräume V_1, V_2 und W sind die folgenden Aussagen äquivalent.*

i) $W \simeq V_1 \oplus V_2$.

ii) Es gibt eine exakte Sequenz

$$0 \longrightarrow V_1 \longrightarrow W \longrightarrow V_2 \longrightarrow 0.$$

iii) Es gibt eine exakte Sequenz

$$0 \longrightarrow V_2 \longrightarrow W \longrightarrow V_1 \longrightarrow 0.$$

iv) V_1 *ist isomorph zu einem (ebenfalls* V_1 *genannten) Untervektorraum von* W*, und* $W/V_1 \simeq V_2$.

v) V_2 *ist isomorph zu einem (ebenfalls* V_2 *genannten) Untervektorraum von* W*, und* $W/V_2 \simeq V_1$.

Beweis: i) $\Rightarrow$ ii) und i) $\Rightarrow$ iii) haben wir oben schon gesehen, siehe (1.14) und (1.15). Falls die exakte Sequenz in ii) gegeben ist, so haben wir eine surjektive Abbildung $W \to V_2$ mit einem Kern, der isomorph zu V_1 ist. Aus dem Homomorphiesatz folgt dann $W/V_1 \simeq V_2$. Ist umgekehrt dieses erfüllt, so haben wir eine exakte Sequenz wie in ii) wegen (1.20). Dies zeigt ii) $\Leftrightarrow$ iv), und iii) $\Leftrightarrow$ v) ist die gleiche Aussage mit vertauschten Bezeichnungen.

ii) $\Rightarrow$ i) und iii) $\Rightarrow$ i) sind symmetrisch; wir zeigen ersteres. Sei also eine exakte Sequenz

$$0 \longrightarrow V_1 \xrightarrow{\iota} W \xrightarrow{f} V_2 \longrightarrow 0$$

gegeben. Sei $(v_i)_{i\in I}$ eine Basis von V_2. Wegen der Surjektivität von f können wir zu jedem $i \in I$ ein $\varphi(v_i) \in W$ finden mit

$$f(\varphi(v_i)) = v_i.$$

Nach Feststellung 1.4.9 läßt sich φ zu einer linearen Abbildung $\varphi : V_2 \to W$ fortsetzen. Es gilt dann sogar

$$f(\varphi(v)) = v \qquad \text{für alle } v \in V_2.$$

Insbesondere ist φ injektiv, vermittelt also einen Isomorphismus $V_2 \xrightarrow{\sim} \varphi(V_2)$. Es genügt daher zu zeigen, daß $W = V_1 \oplus \varphi(V_2)$ (wobei V_1 mittels ι als Untervektorraum von W aufgefaßt wird). Ein beliebiges $w \in W$ läßt sich schreiben als

$$w = w - \varphi(f(w)) + \varphi(f(w)).$$

Natürlich ist $\varphi(f(w)) \in \varphi(V_2)$. Es gilt aber $w - \varphi(f(w)) \in \ker(f) = V_1$, denn $f(w - \varphi(f(w))) = f(w) - f(\varphi(f(w))) = f(w) - f(w) = 0$. Dies zeigt $W = V_1 + \varphi(V_2)$. Gemäß (1.17) müssen wir für die Direktheit der Summe noch $V_1 \cap \varphi(V_2) = 0$ zeigen. Da $V_1 = \ker(f)$ ist, gilt für $w = \varphi(v) \in V_1 \cap \varphi(V_2)$

$$0 = f(w) = f(\varphi(v)) = v,$$

also $w = f(0) = 0$. ∎

Bemerkung 1.5.9 Wenn in einer geeigneten Kategorie $\mathbf{C}$ (siehe Anhang B.1) eine kurze exakte Sequenz

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

gegeben ist, so sagt man, diese Sequenz **spaltet** oder **zerfällt**, wenn B isomorph ist zum *Produkt* von A und C (vgl. Bemerkung 1.5.5). Die Hauptaussage von Feststellung 1.5.8 ist, daß *jede* kurze exakte Sequenz in der Kategorie $\mathbf{Vec}_K$ spaltet. Dies ist falsch in der Kategorie der Moduln über einem kommutativen Ring, der kein Körper ist. (In obigem Beweis haben wir Feststellung 1.4.9 benutzt, deren Aussage für allgemeine Moduln nicht gilt.)

Kapitel 2

Endlich-dimensionale Vektorräume

Die Gesamtheit aller *endlich-dimensionalen* K-Vektorräume ist sehr übersichtlich: Nach dem sogenannten „Hauptsatz“ ist jeder n-dimensionale Raum V isomorph zum Standardraum K^n. Allerdings ist dieser Isomorphismus nicht *kanonisch*, sondern hängt von der Wahl einer Basis in V ab. Der Übergang von einer Basis zur anderen macht das eigentlich interessante in der Theorie der endlich-dimensionalen Vektorräume aus. Mit anderen Worten: Anstelle der Struktur „endlich-dimensionaler K-Vektorraum“ betrachtet man besser „endlich-dimensionaler K-Vektorraum plus Basis“.

Da es nur „wenige“ Vektorräume endlicher Dimension gibt, lassen sich auch die linearen Abbildungen zwischen solchen in übersichtlicher Weise durch *Matrizen* beschreiben. Doch auch hier gilt, daß diese Beschreibung von der Wahl von Basen abhängt, und der Übergang von einer Basis zur anderen macht einen wesentlichen Aspekt der Theorie aus.

Matrizen treten auch in natürlicher Weise bei *linearen Gleichungssystemen* auf, die anschließend behandelt werden. Schließlich führen wir den *Dualraum* eines endlich-dimensionalen Vektorraumes V ein. Dieser besitzt die gleiche Dimension wie V, ist nach dem „Hauptsatz“ also isomorph zu V – aber eben nicht kanonisch.

2.1 Dimensionsformeln und der Hauptsatz

In diesem Kapitel ist K ein beliebiger Körper, und $V, W, \ldots$ sind K-Vektorräume, in aller Regel endlich-dimensional. Nach Definition 1.4.16 bedeutet $\dim(V) = n$, $n \in \mathbb{N}$, daß V eine Basis $(v_1, \ldots, v_n)$ aus n Elementen besitzt; jeder Vektor $v \in V$ läßt sich dann eindeutig als Linearkombination der v_i darstellen:

$$v = a_1 v_1 + \ldots + a_n v_n \qquad \text{mit eindeutig bestimmten } a_i \in K.$$

Im allgemeinen besitzt ein Vektorraum sehr viele verschiedene Basen. Wir wollen die Begriffe Basis und Dimension auch für den Nullvektorraum $\{0\}$ zur Verfügung haben: Er bekommt die Dimension 0 und als (einzige) Basis die leere Menge zugeordnet.

Feststellung 2.1.1 *Seien V, V' und W Vektorräume über K.*

i) Sei $0 \to V \to W \to V' \to 0$ *eine exakte Sequenz von* K*-Vektorräumen. Falls zwei der Vektorräume endlich-dimensional sind, so auch der dritte, und es gilt*

$$\dim(W) = \dim(V) + \dim(V').$$

ii) Wenn V *und* V' *endlich-dimensional sind, so auch* $V \oplus V'$*, und es gilt*

$$\dim(V \oplus V') = \dim(V) + \dim(V').$$

iii) Sind V *und* V' *Untervektorräume des endlich-dimensionalen Vektorraumes* W*, so gilt*

$$\dim(V + V') = \dim(V) + \dim(V') - \dim(V \cap V').$$

iv) Ist V *endlich-dimensional und* $W \subset V$ *ein Untervektorraum, so gilt*

$$\dim(V/W) = \dim(V) - \dim(W).$$

v) Ist $f : V \to W$ *eine lineare Abbildung und* V *endlich-dimensional, so gilt*

$$\dim(V) = \dim(\ker(f)) + \dim(\operatorname{im}(f)).$$

Beweis: i) Nach Feststellung 1.5.8 gilt $W \simeq V \oplus V'$. Ist $(v_i)_{i\in I}$ eine Basis von V und $(v'_j)_{j\in J}$ eine Basis von V', so ist die *Vereinigung* dieser beiden Mengen eine Basis von $V \oplus V'$; vgl. Bemerkung 1.5.3 (wir haben V und V' wie üblich als Untervektorräume von $V \oplus V'$ aufgefaßt). Daraus folgt die Behauptung.

ii) folgt aus $0 \to V \to V \oplus V' \to V' \to 0$ und i).

iii) folgt aus (1.16) und i) und ii).

iv) folgt aus $0 \to W \to V \to V/W \to 0$ und i).

v) folgt aus $0 \to \ker(f) \to V \xrightarrow{f} \operatorname{im}(f) \to 0$ und i). ∎

Folgerung 2.1.2 *Sei* V *ein endlich-dimensionaler* K*-Vektorraum, und* $W \subset V$ *ein Untervektorraum. Dann sind die folgenden Aussagen äquivalent:*

i) $W = V$.

ii) $\dim(W) = \dim(V)$.

iii) $V/W = 0$.

Beweis: Aus i) folgt natürlich ii). Aus ii) folgt iii) nach 2.1.1 iv). Aus iii) folgt i) direkt nach Definition von V/W: Dieser Quotient besteht aus den Nebenklassen $v + W$, und wenn er nur das einzige Element $0 + W$ besitzt, so bedeutet dies $v + W = 0 + W$ für alle $v \in V$, also $v \in W$ für alle $v \in V$. ∎

Feststellung 2.1.1 i) läßt sich zu der folgenden Aussage verallgemeinern.

Feststellung 2.1.3 *Sei*

$$0 \longrightarrow V_1 \longrightarrow V_2 \longrightarrow \ldots \longrightarrow V_n \longrightarrow 0 \tag{2.1}$$

eine exakte Sequenz von endlich-dimensionalen K-Vektorräumen. Dann verschwindet die alternierende Summe der Dimensionen:

$$\sum_{i=1}^{n} (-1)^i \dim(V_i) = 0.$$

Beweis: Sei die exakte Sequenz

$$0 \longrightarrow V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} \ldots \xrightarrow{f_{n-1}} V_n \xrightarrow{f_n} 0 \tag{2.2}$$

gegeben. Nach Feststellung 2.1.1 v) gilt

$$\dim(V_i) = \dim(\mathrm{im}(f_i)) + \dim(\ker(f_i)) \qquad i = 1, \ldots, n,$$

und unter Ausnutzung der Exaktheit

$$\dim(V_i) = \dim(\ker(f_{i+1})) + \dim(\ker(f_i)) \qquad i = 1, \ldots, n-1,$$

Nun multipliziere man diese Gleichung mit $(-1)^i$ und summiere über $i = 1, \ldots, n-1$. Nach kurzer Rechnung ergibt sich die Behauptung. ∎

Feststellung 2.1.4 *Es seien V und W zwei K-Vektorräume der gleichen endlichen Dimension. Für eine lineare Abbildung $f : V \to W$ sind die folgenden Aussagen äquivalent:*

i) f ist injektiv.

ii) f ist surjektiv.

iii) f ist bijektiv.

Beweis: Es genügt zu zeigen, daß i) und ii) äquivalent sind. Wir betrachten die exakte Sequenz

$$0 \longrightarrow \ker(f) \longrightarrow V \xrightarrow{f} W \longrightarrow W/\mathrm{im}(f) \longrightarrow 0.$$

Feststellung 2.1.3 liefert die Dimensionsformel

$$\dim(\ker(f)) + \dim(W) = \dim(V) + \dim(W/\mathrm{im}(f)),$$

nach Voraussetzung also

$$\dim(\ker(f)) = \dim(W/\mathrm{im}(f)).$$

Deshalb gilt

$$\begin{aligned} f \text{ injektiv } &\iff \ker(f) = 0 \iff \dim(\ker(f)) = 0 \\ &\iff \dim(W/\mathrm{im}(f)) = 0 \iff \mathrm{im}(f) = W \iff f \text{ surjektiv.} \end{aligned}$$

∎

Feststellung 2.1.5 („Hauptsatz für endlich-dimensionale Vektorräume") *Sei V ein K-Vektorraum der Dimension $n \in \mathbb{N}$. Dann ist V isomorph zum K^n:*

$$V \simeq K^n.$$

Beweis: Sei $(v_1, \ldots, v_n)$ eine Basis von V. Wir betrachten die lineare Abbildung

$$\begin{aligned} K^n &\longrightarrow V, \\ (a_1, \ldots, a_n) &\longmapsto a_1 v_1 + \ldots + a_n v_n. \end{aligned}$$

Sie ist surjektiv, da jedes Element von V eine Linearkombination der Basis ist. Also ist sie ein Isomorphismus nach Feststellung 2.1.4. ∎

Bemerkung 2.1.6 Diese einfache Aussage nennt sich *Hauptsatz*, weil sie einen vollständigen Überblick über alle endlich-dimensionalen Vektorräume gibt: Bis auf Isomorphie gibt es nur einen einzigen n-dimensionalen K-Vektorraum, den K^n. Wie aus dem Beweis hervorgeht, ist der Isomorphismus $V \simeq K^n$ allerdings nicht *kanonisch*, sondern hängt von der Wahl einer Basis von V ab. Im folgenden werden wir u.a. untersuchen, was beim Übergang zu einer anderen Basis geschieht („Basiswechsel").

Man könnte kurz so sagen: Anstatt der Struktur „endlich-dimensionaler K-Vektorraum", über die es nach 2.1.5 nicht viel zu sagen gibt, beschäftigen wir uns mit der Struktur „endlich-dimensionaler K-Vektorraum plus gegebene Basis".

2.2 Matrizen

Grundlagen

Mit den *Matrizen* führen wir nun weitere fundamentale Objekte der linearen Algebra ein. Wie üblich sei K ein beliebiger Körper. Seien $m, n \in \mathbb{N}$. Eine $m \times n$-**Matrix** A mit Koeffizienten aus K ist ein rechteckiges Schema

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}, \qquad a_{ij} \in K,\ i = 1, \ldots, m,\ j = 1, \ldots, n,$$

von Elementen a_{ij} aus K, die man die **Koeffizienten** der Matrix nennt. Wenn das **Format** $m \times n$ klar ist, so schreiben wir oft abkürzend

$$A = (a_{ij})_{i,j} \qquad \text{oder} \qquad A = (a_{ij}).$$

Eine $1 \times n$–Matrix ist ein Zeilenvektor, und eine $m \times 1$–Matrix ist ein Spaltenvektor. Für die Menge aller $m \times n$-Matrizen mit Koeffizienten aus K verwenden wir die Symbole

$$M(m \times n, K) \qquad \text{oder} \qquad K^{m,n}.$$

Im Falle $m = n$, d.h. für **quadratische** Matrizen, schreiben wir kurz

$$M(n, K) = M(n \times n, K).$$

In natürlicher Weise, nämlich durch komponentenweise Addition und Skalarmultiplikation, ist $M(m \times n, K)$ ein K-Vektorraum:

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

Das neutrale Element in $M(m \times n, K)$ ist die **Nullmatrix**, in der jeder Koeffizient Null ist. Offenbar gilt

$$K^{1,n} \simeq K^n \simeq K^{n,1}.$$

Die Matrizen E_{ij}, deren (i,j)–Komponente 1 ist, während alle anderen Komponenten 0 sind, bilden eine Basis von $M(m \times n, K)$, die **kanonische** Basis:

$$E_{ij} = \begin{array}{cl} \left(\begin{array}{cccc} 0 & \dots & \dots & 0 \\ \vdots & & 1 & \vdots \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{array}\right) & \begin{array}{l} \\ \leftarrow i \\ \\ \\ \end{array} \\ \begin{array}{cccc} & & \uparrow & \\ & & j & \end{array} & \end{array} \qquad 1 \le i \le m,\ 1 \le j \le n.$$

Daraus folgt sofort:

$$\dim_K(M(m \times n, K)) = mn. \tag{2.3}$$

Nach dem „Hauptsatz“ 2.1.5 ist also $M(m \times n, K) \simeq K^{mn}$; einen Isomorphismus erhält man z.B. dadurch, daß man alle Zeilen einer Matrix nebeneinander hinschreibt.

Die **Hauptdiagonale** einer quadratischen $n \times n$–Matrix A besteht aus den Koeffizienten an den Positionen $(1,1), (2,2), \dots, (n,n)$. Sind alle anderen Koeffizienten Null, so heißt A eine **Diagonalmatrix**. Eine besondere Rolle spielt die Diagonalmatrix, deren Diagonalkoeffizienten alle 1 sind; sie heißt die $n \times n$–**Einheitsmatrix** und bekommt das Symbol $\mathbf{1}_n$ (oder einfach $\mathbf{1}$, oder E_n, oder E):

$$\mathbf{1}_n = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}.$$

Matrixmultiplikation

Als bedeutsam erweist sich die Tatsache, daß man Matrizen geeigneten Formats miteinander multiplizieren kann. Wir geben hier einfach die Definition der Matrixmultiplikation; in Kürze wird ihr tieferer Sinn klar werden.

Seien also $A = (a_{ij}) \in M(l \times m, K)$ und $B = (b_{ij}) \in M(m \times n, K)$; die Zeilenlänge von A stimmt also mit der Spaltenlänge von B überein. Deshalb können wir die Summen

$$c_{ij} := \sum_{k=1}^{m} a_{ik} b_{kj}, \qquad 1 \le i \le l,\ 1 \le j \le n \tag{2.4}$$

bilden. Alle diese bilden zusammen eine $l \times n$-Matrix $C = (c_{ij})$. Wir definieren das **Produkt** von A und B als diese Matrix C:

$$A \cdot B = C \qquad \text{mit } C = (c_{ij}) \text{ und } c_{ij} \text{ wie in (2.4).} \tag{2.5}$$

(Manchmal schreiben wir AB anstelle von $A \cdot B$). Dadurch ist eine Abbildung

$$M(l \times m, K) \times M(m \times n, K) \longrightarrow M(l \times n, K) \tag{2.6}$$

definiert, deren grundlegende Eigenschaften wir in Feststellung 2.2.1 weiter unten geben. Zuvor einige Beispiele.

Beispiele: 1. Für $l = m = n = 1$ ist (2.6) einfach die gewöhnliche Multiplikation $K \times K \to K$ im Körper K.

2. Für 2×2-Matrizen ist die Multiplikation explizit wie folgt gegeben:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}.$$

3. „Matrix mal Spaltenvektor gibt Spaltenvektor“:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + \ldots + a_{1n}b_n \\ \vdots \\ a_{m1}b_1 + \ldots + a_{mn}b_n \end{pmatrix}.$$

4. „Zeilenvektor mal Spaltenvektor gibt Skalar“:

$$(a_1, \ldots, a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \ldots + a_n b_n.$$

5. „Spaltenvektor mal Zeilenvektor gibt Matrix“:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} (b_1, \ldots, b_n) = \begin{pmatrix} a_1 b_1 & \dots & a_1 b_n \\ \vdots & & \vdots \\ a_n b_1 & \dots & a_n b_n \end{pmatrix}.$$

Feststellung 2.2.1 *Für Matrizen geeigneten Formats gelten folgende Rechenregeln:*

i) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

ii) $A \cdot (B + C) = A \cdot B + A \cdot C$.

iii) $(A + B) \cdot C = A \cdot C + B \cdot C$.

iv) $A \cdot \mathbf{1}_n = A = \mathbf{1}_m \cdot A$ *für* $A \in M(m \times n, K)$.

Diese Gleichungen sind bezüglich der „Punkt–vor–Strich“–Regel zu lesen.

Beweis: Dies zeigt man durch direkte Rechnung. Exemplarisch beweisen wir i). Sei

$$A = (a_{ij}), \qquad B = (b_{jk}), \qquad C = (c_{kl}).$$

Der (i,k)-Koeffizient von $A \cdot B$ ist $\sum_j a_{ij} b_{jk}$ (wir unterdrücken die Summationsgrenzen, d.h. die Formate der Matrizen). Der (i,l)-Koeffizient von $(A \cdot B) \cdot C$ ist daher

$$\sum_k \Big(\sum_j a_{ij} b_{jk} \Big) c_{kl}.$$

Da endliche Summen beliebig umgeordnet und vertauscht werden dürfen, ist dieser Ausdruck gleich

$$\sum_k \sum_j a_{ij} b_{jk} c_{kl} = \sum_j \sum_k a_{ij} b_{jk} c_{kl} = \sum_j a_{ij} \Big(\sum_k b_{jk} c_{kl} \Big).$$

Der Ausdruck in Klammern ist der (j,l)-Koeffizient von $B \cdot C$. Der ganze Ausdruck ist deshalb der (i,l)-Koeffizient von $A \cdot (B \cdot C)$. Wir haben gezeigt, daß die (i,l)-Koeffizienten von $(A \cdot B) \cdot C$ und $A \cdot (B \cdot C)$ gleich sind; deshalb sind diese beiden Matrizen gleich. ∎

Wir erinnern an die Definition 1.3.14 einer K-Algebra. Man vergleiche die folgende Feststellung mit 1.3.15; der Zusammenhang wird im nächsten Abschnitt genauer untersucht.

Feststellung 2.2.2 *Der Vektorraum $M(n,K)$ aller $n \times n$-Matrizen mit Koeffizienten in K wird mit der Matrixmultiplikation zu einer K-Algebra. Diese besitzt ein Einselement, nämlich die Einheitsmatrix $\mathbf{1}_n$.*

Beweis: Zunächst definiert Matrixmultiplikation tatsächlich eine Abbildung

$$M(n,K) \times M(n,K) \longrightarrow M(n,K),$$

d.h., Multiplikation von $n \times n$-Matrizen ist uneingeschränkt möglich. Feststellung 2.2.1 besagt gerade, daß $M(n,K)$ dadurch zu einem Ring mit Eins wird. Die Verträglichkeitseigenschaft (1.6) ergibt sich durch eine leichte Rechnung. ∎

Bemerkungen: 1. Die Algebra $M(1,K)$ ist der Körper K.

2. Außer im Extremfall $n = 1$ ist $M(n,K)$ ein nicht-kommutativer Ring.

Invertierbare Matrizen

Definition 2.2.3 *Die* Einheitengruppe *(siehe 1.2.5) von $M(n,K)$ heißt die* **allgemeine lineare Gruppe** *und wird mit*

$$\mathrm{GL}(n,K)$$

*bezeichnet (**G**eneral **L**inear group). Die Elemente von* $\mathrm{GL}(n,K)$ *heißen* **invertierbare** *Matrizen. Ein* $A \in M(n,K)$ *ist also genau dann invertierbar, wenn es ein* $A^{-1} \in M(n,K)$ *gibt mit*

$$A \cdot A^{-1} = A^{-1} \cdot A = \mathbf{1}_n.$$

A^{-1} *heißt dann die zu* A **inverse** *Matrix. Nicht-invertierbare Matrizen heißen auch* **singulär***.*

Bemerkungen: 1. Die inverse Matrix A^{-1} ist, falls existent, eindeutig bestimmt; siehe 1.1.2 ii).

2. Wir werden im nächsten Abschnitt sehen, daß $A \in M(n,K)$ schon dann invertierbar ist, wenn es ein $B \in M(n,K)$ gibt mit $AB = \mathbf{1}$ (oder $BA = \mathbf{1}$); in diesem Fall ist $B = A^{-1}$ (siehe 2.3.5).

Eine Matrix $A \in M(n,K)$ heißt **Diagonalmatrix**, falls alle Koeffizienten außerhalb der Hauptdiagonale verschwinden (gleich Null sind). Oft wird das Symbol

$$A = \mathrm{diag}(a_1, \dots, a_n) = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}, \qquad a_1, \dots, a_n \in K,$$

verwendet. Das Produkt zweier Diagonalmatrizen ist wieder diagonal:

$$\mathrm{diag}(a_1, \dots, a_n)\mathrm{diag}(b_1, \dots, b_n) = \mathrm{diag}(a_1 b_1, \dots, a_n b_n). \tag{2.7}$$

Feststellung 2.2.4 *Die invertierbaren Diagonalmatrizen bilden eine Untergruppe von* $\mathrm{GL}(n,K)$*, die isomorph zu* $K^* \times \ldots \times K^*$ *(n mal) ist. Genauer: Die Abbildung*

$$\begin{aligned} K^* \times \ldots \times K^* &\longrightarrow \mathrm{GL}(n,K), \\ (a_1, \dots, a_n) &\longmapsto \mathrm{diag}(a_1, \dots, a_n) \end{aligned} \tag{2.8}$$

ist ein injektiver Gruppenhomomorphismus.

Beweis: $K^* \times \ldots \times K^*$ ist eine Gruppe bezüglich komponentenweiser Multiplikation:

$$(a_1, \dots, a_n)(b_1, \dots, b_n) = (a_1 b_1, \dots, a_n b_n).$$

Gleichung (2.7) besagt gerade, daß die Abbildung (2.8) ein Gruppenhomomorphismus ist. Die Injektivität ist trivial. Schließlich sieht man leicht, daß eine Diagonalmatrix $\mathrm{diag}(a_1, \dots, a_n)$ genau dann invertierbar ist, wenn alle a_i ungleich Null sind. ∎

Definition und Feststellung 2.2.5 *Sei* G *eine Gruppe. Das* **Zentrum** $Z(G)$ *von* G *ist die Menge derjenigen Elemente von* G*, die mit jedem Element kommutieren:*

$$Z(G) = \{g \in G : \; gh = hg \text{ für alle } h \in G\}.$$

$Z(G)$ *ist eine kommutative Untergruppe von* G*.*

Beweis: Aus $g \in Z(G)$ und $g' \in Z(G)$ folgt $gg' \in Z(G)$, denn

$$(gg')h = g(g'h) = g(hg') = (gh)g' = (hg)g' = h(gg').$$

Aus $g \in Z(g)$ folgt $g^{-1} \in Z(G)$: Man multipliziere $gh = hg$ von beiden Seiten mit g^{-1}. Also ist $Z(G)$ eine Untergruppe von G. Da jedes Element von $Z(G)$ mit jedem Element vertauscht, ist $Z(G)$ kommutativ. ∎

Zu jedem Skalar $a \in K^*$ betrachten wir die **skalare Matrix** $\operatorname{diag}(a, \ldots, a) \in \mathrm{GL}(n, K)$. Die skalaren Matrizen bilden offenbar eine Untergruppe von $\mathrm{GL}(n, K)$, die isomorph zu K^* ist.

Feststellung 2.2.6 *Das Zentrum von* $\mathrm{GL}(n, K)$ *besteht genau aus den skalaren Matrizen, ist also isomorph zu* K^*.

Beweis: Multiplikation einer Matrix $B \in \mathrm{GL}(n, K)$ mit der skalaren Matrix $\operatorname{diag}(a, \ldots, a)$ von links *oder* von rechts bewirkt, daß jeder Koeffizient von B mit a multipliziert wird. Insbesondere liegt $\operatorname{diag}(a, \ldots, a)$ im Zentrum von $\mathrm{GL}(n, K)$.

Sei umgekehrt (a_{ij}) im Zentrum gegeben; wir müssen zeigen, daß (a_{ij}) skalar ist. Sei B_{12} die invertierbare Matrix mit Einsen auf der Hauptdiagonalen, einer Eins an der Position $(1, 2)$, und allen anderen Koeffizienten Null. Da (a_{ij}) im Zentrum ist, gilt

$$(a_{ij})B_{12} = B_{12}(a_{ij}). \tag{2.9}$$

Der $(1, 1)$-Koeffizient links ist a_{11}, und der rechts ist $a_{11} + a_{21}$. Es folgt $a_{21} = 0$. Indem man anstatt B_{12} andere Matrizen hernimmt, die jeweils genau eine 1 abseits der Hauptdiagonalen haben, zeigt man schließlich

$$a_{ij} = 0 \qquad \text{für alle } i \neq j.$$

Mit anderen Worten: (a_{ij}) ist eine Diagonalmatrix. Es bleibt zu zeigen, daß alle Diagonalkoeffizienten gleich sind. Dazu betrachte man wieder (2.9): Der $(1, 2)$-Koeffizient links ist a_{11}, der rechts ist a_{22}. Also sind diese beiden Koeffizienten gleich, und so ähnlich fährt man fort. ∎

2.3 Matrizen und lineare Abbildungen

In diesem Abschnitt untersuchen wir den Vektorraum $\mathrm{Hom}_K(V, W)$ (siehe Definition 1.3.9) für zwei *endlich-dimensionale* K-Vektorräume V und W. Wir beginnen mit dem Fall $V = K^n$ und $W = K^m$, wobei die Elemente dieser Räume hier immer als *Spaltenvektoren* aufzufassen sind.

Lineare Abbildungen zwischen den Standardräumen

Feststellung 2.3.1 *Der Vektorraum* $\mathrm{Hom}_K(K^n, K^m)$ *identifiziert sich in natürlicher Weise mit* $M(m \times n, K)$. *Genauer:*

i) Jede Matrix $A \in M(m \times n, K)$ *definiert eine lineare Abbildung*

$$K^n \longrightarrow K^m, \tag{2.10}$$

$$v \longmapsto Av$$

(v ist eine $n \times 1$-Matrix, und Av ist das Matrixprodukt; siehe Beispiel 3 auf Seite 40).

ii) Zu jedem $f \in \mathrm{Hom}_K(K^n, K^m)$ gibt es genau ein $A \in M(m \times n, K)$, so daß f mit der Abbildung (2.10) übereinstimmt.

iii) Die Abbildung

$$M(m \times n, K) \longrightarrow \mathrm{Hom}_K(K^n, K^m), \tag{2.11}$$

die jeder Matrix die lineare Abbildung (2.10) zuordnet, ist ein Isomorphismus von K-Vektorräumen.

iv) $\dim_K(\mathrm{Hom}_K(K^n, K^m)) = mn$.

Beweis: i) Man rechnet sofort nach, daß

$$A(v_1 + v_2) = Av_1 + Av_2, \qquad A(\lambda.v) = \lambda.(Av) \qquad (\lambda \in K).$$

Also ist (2.10) eine lineare Abbildung.

ii) Sei $f : K^n \to K^m$ linear gegeben. Seien $w_1, \dots, w_n$ die Bilder der kanonischen Basisvektoren des K^n:

$$f(e_i) = w_i, \qquad i = 1, \dots, n.$$

Sei A die Matrix mit den Spalten $w_1, \dots, w_n$. Dann ist $A \in M(m \times n, K)$, und wir behaupten, daß die lineare Abbildung (2.10) mit f übereinstimmt. Da eine lineare Abbildung durch ihre Werte auf einer Basis festgelegt ist (siehe Feststellung 1.4.9), müssen wir dazu nur

$$f(e_i) = Ae_i \qquad \text{für } i = 1, \dots, n$$

zeigen. Nach Definition des Matrixproduktes ist Ae_i die i-te Spalte von A, also gleich $w_i = f(e_i)$, und die Behauptung folgt.

Stellt auch A' die Abbildung f dar, so ist $Ae_i = f(e_i) = A'e_i$, also die i-te Spalte von A gleich der i-ten Spalte von A', also $A = A'$.

iii) Sei φ die Abbildung (2.11). Wir zeigen zunächst

$$\varphi(A + B) = \varphi(A) + \varphi(B) \qquad \text{für } A, B \in M(m \times n, K).$$

Dies ist eine Identität von linearen Abbildungen, die genau dann richtig ist, wenn sie auf einer Basis gilt:

$$\varphi(A + B)(e_i) = (\varphi(A) + \varphi(B))(e_i) \qquad \text{für } i = 1, \dots, n.$$

Die linke Seite ist gleich $(A + B)e_i$, was die i-te Spalte von $A + B$ ist. Die rechte Seite ist gleich $\varphi(A)(e_i) + \varphi(B)(e_i) = Ae_i + Be_i$, also gleich der i-ten Spalte von A plus der i-ten Spalte von B. Somit sind beide Seiten gleich.

Ähnlich zeigt man $\varphi(\lambda A) = \lambda\varphi(A)$, und also ist φ linear. Es ist sogar ein Isomorphismus wegen den unterstrichenen Wörtern in ii).

iv) folgt aus iii). ∎

Bemerkung 2.3.2 Der Isomorphismus (2.11) ist so natürlich, daß wir in Zukunft nicht mehr zwischen einer $m \times n$-Matrix und der durch sie definierten linearen Abbildung $K^n \to K^m$ unterscheiden:

$$M(m \times n, K) = \mathrm{Hom}_K(K^n, K^m).$$

Für $m = n$ erhalten wir insbesondere die Identifizierung

$$M(n, K) = \mathrm{End}(K^n).$$

Jetzt wird die Bedeutung der Matrixmultiplikation (2.6) klar werden:

Feststellung 2.3.3 *Sei $A \in M(l \times m, K)$ und $B \in M(m \times n, K)$. Diese Matrizen definieren lineare Abbildungen*

$$K^n \xrightarrow{B} K^m \xrightarrow{A} K^l.$$

Das Kompositum dieser beiden linearen Abbildungen ist dann durch die Produktmatrix $AB \in M(l \times n, K)$ gegeben:

$$A \circ B = AB.$$

Beweis: Wir werten die lineare Abbildung $A \circ B$ auf dem Einheitsvektor $e_i \in K^n$ aus:

$$(A \circ B)(e_i) = A(Be_i) = A \begin{pmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1j} b_{ji} \\ \vdots \\ \sum_{j=1}^m a_{lj} b_{ji} \end{pmatrix}.$$

Dies ist gerade die i-te Spalte der Produktmatrix AB, also gleich $(AB)e_i$. Die linearen Abbildungen $A \circ B$ und AB stimmen daher auf der kanonischen Basis überein, sind folglich gleich. ■

Folgerung 2.3.4 *Der Isomorphismus*

$$M(n, K) = \mathrm{End}_K(K^n)$$

von K-Vektorräumen ist sogar ein Isomorphismus von K-Algebren (d.h., ein Ringhomomorphismus). Überdies gilt

$$\mathrm{GL}(n, K) = \mathrm{Aut}(K^n)$$

(vgl. Definition 1.3.16), d.h., eine Matrix $A \in M(n, K)$ ist genau dann invertierbar, wenn der durch sie definierte Endomorphismus des K^n ein Isomorphismus ist.

Beweis: Wir müssen zeigen, daß ein Produkt von Matrizen sich mit dem Produkt (d.h., dem Kompositum) der entsprechenden Endomorphismen identifiziert. Das ist aber gerade die Aussage von Feststellung 2.3.3.

Die zweite Aussage folgt aus der ersten: Sind zwei Ringe isomorph, so auch ihre Einheitengruppen. ■

Insbesondere gilt also: Die Umkehrabbildung der bijektiven linearen Abbildung $A \in \mathrm{GL}(n, K)$ ist durch die inverse Matrix $A^{-1} \in \mathrm{GL}(n, K)$ gegeben.

Folgerung 2.3.5 *Eine Matrix $A \in M(n, K)$ ist schon dann invertierbar, wenn es ein $B \in M(n, K)$ gibt mit $AB = \mathbf{1}_n$ (oder $BA = \mathbf{1}_n$). In diesem Fall ist $B = A^{-1}$.*

Beweis: Aus $AB = \mathbf{1}$ (bzw. $BA = \mathbf{1}$) folgt, daß die lineare Abbildung A surjektiv (bzw. injektiv) ist. Nach Feststellung 2.1.4 ist A dann schon bijektiv, also ein Element von $\mathrm{GL}(n, K)$ (siehe 2.3.4). Die Matrix B stellt die inverse Abbildung zu A dar, und somit ist $B = A^{-1}$. ∎

Lineare Abbildungen zwischen allgemeinen Vektorräumen

Wir untersuchen jetzt $\mathrm{Hom}_K(V, W)$ für zwei beliebige endlich-dimensionale Vektorräume V und W. Sei etwa

$$n = \dim(V), \qquad m = \dim(W).$$

Wir wählen irgendeine Basis

$$\mathcal{B} = (v_1, \ldots, v_n)$$

von V, und

$$\mathcal{C} = (w_1, \ldots, w_m)$$

eine Basis von W. Diese Basen definieren Isomorphismen, manchmal **Koordinatenisomorphismen** genannt,

$$\begin{aligned} &\varphi_{\mathcal{B}} : K^n \xrightarrow{\sim} V, && \varphi_{\mathcal{C}} : K^m \xrightarrow{\sim} W, \\ &\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \longmapsto a_1 v_1 + \ldots + a_n v_n, && \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \longmapsto b_1 w_1 + \ldots + b_m w_m \end{aligned} \tag{2.12}$$

(vergleiche den „Hauptsatz" 2.1.5). Wir sagen auch, $(a_1, \ldots, a_n)$ sind die *Koordinaten* des Vektors $a_1 v_1 + \ldots + a_n v_n \in V$ bezüglich der Basis $\mathcal{B}$, und ähnlich für W.

Definition und Feststellung 2.3.6 *Es gelten obige Bezeichnungen. Sei $f : V \to W$ eine lineare Abbildung. Die* **Matrix von f bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$** *ist die Matrix $A = (a_{ij}) \in M(m \times n, K)$, so daß*

$$f(v_i) = a_{1i} w_1 + \ldots + a_{mi} w_m. \tag{2.13}$$

A ist die eindeutig bestimmte Matrix, so daß das Diagramm

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ K^n \arrow[u, "\varphi_{\mathcal{B}}"] \arrow[r, "A"] & K^m \arrow[u, "\varphi_{\mathcal{C}}"'] \end{tikzcd}$$

kommutativ ist. (Im Falle $W = V$ und $\mathcal{C} = \mathcal{B}$ heißt A die Matrix von f bezüglich $\mathcal{B}$.)

Beweis: Das Diagramm ist genau dann kommutativ, wenn

$$f(\varphi_{\mathcal{B}}(e_i)) = \varphi_{\mathcal{C}}(Ae_i) \qquad \text{für } i = 1, \dots, n, \tag{2.14}$$

wobei e_i die kanonischen Basisvektoren des K^n sind. Die linke Seite ist gleich $f(v_i)$. Die rechte Seite ist gleich

$$\varphi_{\mathcal{C}} \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = a_{1i}w_1 + \ldots + a_{mi}w_m.$$

Also sind beide Seiten gleich wegen (2.13). Außerdem ist A durch die Bedingung (2.14) festgelegt, denn $\varphi_{\mathcal{C}}$ ist ein Isomorphismus. ■

Bemerkung 2.3.7 i) Das Rezept für die Bestimmung der Matrix A einer linearen Abbildung f bezüglich Basen $(v_1, \dots, v_n)$ und $(w_1, \dots, w_m)$ ist also wie folgt: Man nehme den ersten Basisvektor v_1 her, betrachte sein Bild $f(v_1)$, und schreibe dieses als Linearkombination der Basis $w_1, \dots, w_m$:

$$f(v_1) = a_{11}w_1 + \ldots + a_{m1}w_m.$$

Die dabei auftretenden Koeffizienten $a_{11}, \dots, a_{m1}$ bilden dann die erste Spalte von A. Genauso verfährt man mit dem zweiten Basisvektor v_2, und erhält die zweite Spalte von A, und so fort.

ii) Faßt man $A \in M(m \times n, K)$ wie in 2.3.1 als lineare Abbildung $K^n \to K^m$ auf, und wählt in K^n und in K^m die kanonische Basis, so ist die Matrix von A bezüglich dieser Basen A selbst.

Folgerung 2.3.8 *Ordnet man jeder linearen Abbildung $f \in \mathrm{Hom}_K(V, W)$ ihre Matrix (bezüglich fest gewählter Basen) zu, so erhält man einen Isomorphismus*

$$\mathrm{Hom}_K(V, W) \xrightarrow{\sim} M(m \times n, K) \tag{2.15}$$

(der von der Wahl der Basen abhängt). Insbesondere gilt

$$\dim(\mathrm{Hom}_K(V, W)) = \dim(V)\dim(W).$$

Beweis: Ist A_i die Matrix von f_i, $i = 1, 2$, so haben wir zwei kommutative Diagramme wie in 2.3.6. Daraus folgt leicht, daß das entsprechende Diagramm mit $f_1 + f_2$ und $A_1 + A_2$ kommutativ ist. Mit anderen Worten, $A_1 + A_2$ ist die Matrix von $f_1 + f_2$. Ähnlich sieht man, daß λA_1 die Matrix von λf_1 ist, wobei $\lambda \in K$. Also ist die Abbildung (2.15) linear.

Sie ist überdies injektiv und surjektiv, was beides leicht aus dem kommutativen Diagramm in 2.3.6 folgt. ■

Feststellung 2.3.9 *Wir betrachten das kommutative Diagramm*

$$\begin{tikzcd} U \arrow[r, "g"] & V \arrow[r, "f"] & W \\ K^l \arrow[u, "\varphi_{\mathcal{A}}"] \arrow[r, "B"] & K^n \arrow[u, "\varphi_{\mathcal{B}}"] \arrow[r, "A"] & K^m. \arrow[u, "\varphi_{\mathcal{C}}"'] \end{tikzcd}$$

Dabei sind U, V, W Vektorräume über K der Dimension l bzw. n bzw. m, und mit Basen $\mathcal{A}$ bzw. $\mathcal{B}$ bzw. $\mathcal{C}$. Letztere definieren die Isomorphismen $\varphi_\mathcal{A}, \varphi_\mathcal{B}, \varphi_\mathcal{C}$. Schließlich ist A die Matrix von f (bzgl. $\mathcal{B}$ und $\mathcal{C}$) und B die Matrix von g (bzgl. $\mathcal{A}$ und $\mathcal{B}$). In dieser Situation ist die Matrix von $f \circ g$ (bzgl. $\mathcal{A}$ und $\mathcal{C}$) durch die Produktmatrix AB gegeben.

Beweis: Zunächst ist das Diagramm tatsächlich kommutativ, denn das linke und das rechte Quadrat sind kommutativ nach 2.3.6. Wir schreiben das Diagramm noch einmal ab unter Weglassung des mittleren Teils:

$$\begin{array}{ccc} U & \xrightarrow{f\circ g} & W \\ {\scriptstyle \varphi_\mathcal{A}}\uparrow & & \uparrow{\scriptstyle \varphi_\mathcal{C}} \\ K^l & \xrightarrow{A\circ B} & K^m. \end{array}$$

Nach 2.3.3 können wir die lineare Abbildung $A \circ B$ durch die Matrix AB ersetzen. Unsere Behauptung folgt dann aus der Eindeutigkeitsaussage in 2.3.6. ∎

Folgerung 2.3.10 *Sei V ein K-Vektorraum der Dimension n mit fest gewählter Basis. Ordnet man jedem $f \in \mathrm{End}_K(V)$ seine Matrix bzgl. dieser Basis zu, so erhält man einen Isomorphismus*

$$\mathrm{End}_K(V) \xrightarrow{\sim} M(n, K) \tag{2.16}$$

von K-Algebren. Einschränkung auf die Einheitengruppen induziert einen Gruppenisomorphismus

$$\mathrm{GL}(V) \xrightarrow{\sim} \mathrm{GL}(n, K).$$

Beweis: Das meiste folgt aus 2.3.8; die einzige zusätzliche Information ist, daß der Vektorraumisomorphismus (2.16) Produkte auf Produkte abbildet (links und rechts steht ein Ring). Dies folgt aber aus 2.3.9. ∎

Übersicht über die Hauptergebnisse dieses Unterabschnitts: Für zwei Vektorräume V und W der Dimension n bzw. m gilt nach Wahl von Basen:

	ist isomorph zu	als
$\mathrm{Hom}_K(V, W)$	$M(m \times n, K)$	K-Vektorräume.
$\mathrm{End}_K(V)$	$M(n, K)$	K-Algebren.
$\mathrm{GL}(V)$	$\mathrm{GL}(n, K)$	Gruppen.

Bemerkung 2.3.11 Dieser Unterabschnitt ist kaum mehr als eine Wiederholung des vorherigen. Der Hauptunterschied ist wie folgt: In einem beliebigen Vektorraum V gibt es, im Gegensatz zum K^n, keine kanonische Basis. Isomorphismen wie in der Tabelle hängen daher von der *Wahl* von Basen in V und W ab. Wir untersuchen jetzt, was beim Übergang zu anderen Basen geschieht.

Basiswechsel

Sei V ein K-Vektorraum der Dimension n, und seien $\mathcal{B} = (b_1, \ldots, b_n)$ und $\mathcal{B}' = (b'_1, \ldots, b'_n)$ zwei Basen von V. Zu diesen Basen haben wir die Koordinatenisomorphismen

$$\varphi_\mathcal{B} : K^n \longrightarrow V, \qquad \varphi_{\mathcal{B}'} : K^n \longrightarrow V,$$

charakterisiert durch $\varphi_{\mathcal{B}}(e_i) = b_i$ und $\varphi_{\mathcal{B}'}(e_i) = b'_i$. Es gibt dann eine eindeutig bestimmte Matrix $S \in \mathrm{GL}(n, K)$, so daß das Diagramm

$$\begin{tikzcd} & V & \\ K^n \arrow[ur, "\varphi_{\mathcal{B}}"] \arrow[rr, "S"'] & & K^n \arrow[ul, "\varphi_{\mathcal{B}'}"'] \end{tikzcd} \tag{2.17}$$

kommutativ ist. Wir nennen S die **Übergangsmatrix** von $\mathcal{B}$ zu $\mathcal{B}'$. Natürlich ist dann S^{-1} die Übergangsmatrix von $\mathcal{B}'$ zu $\mathcal{B}$. Explizit gilt, falls $S = (s_{ij})$,

$$s_{1i}b'_1 + \ldots + s_{ni}b'_n = b_i \qquad \text{für } i = 1, \ldots, n.$$

Feststellung 2.3.12 *Seien V und W Vektorräume über K der Dimensionen n bzw. m. Seien $\mathcal{B}, \mathcal{B}'$ Basen von V und $\mathcal{C}, \mathcal{C}'$ Basen von W. Sei S die Übergangsmatrix von $\mathcal{B}$ zu $\mathcal{B}'$ und T die Übergangsmatrix von $\mathcal{C}$ zu $\mathcal{C}'$. Dann gilt: Ist $A \in M(m \times n, K)$ die Matrix von $f \in \mathrm{Hom}_K(V, W)$ bzgl. $\mathcal{B}$ und $\mathcal{C}$, so ist*

$$TAS^{-1} \text{ die Matrix von } f \text{ bzgl. } \mathcal{B}' \text{ und } \mathcal{C}'.$$

Beweis: Wir haben das kommutative Diagramm

$$\begin{tikzcd} & V \arrow[r, "f"] & W & \\ K^n \arrow[ur, "\varphi_{\mathcal{B}'}"] \arrow[r, "S^{-1}"'] & K^n \arrow[u, "\varphi_{\mathcal{B}}"'] \arrow[r, "A"'] & K^m \arrow[u, "\varphi_{\mathcal{C}}"] \arrow[r, "T"'] & K^m \arrow[ul, "\varphi_{\mathcal{C}'}"'] \end{tikzcd}$$

(das mittlere Quadrat ist kommutativ nach Definition von A, die beiden äußeren Dreiecke sind kommutativ nach Definition von S und T). Die Komposition der unteren Abbildungen ist TAS^{-1}. Also folgt die Behauptung aus der Eindeutigkeitsaussage in 2.3.6. ∎

Folgerung 2.3.13 *Sei $\dim_K(V) = n$, und seien $\mathcal{B}, \mathcal{B}'$ zwei Basen von V. Sei $S \in \mathrm{GL}(n, K)$ die Übergangsmatrix von $\mathcal{B}$ zu $\mathcal{B}'$. Ist $A \in M(n, K)$ die Matrix von $f \in \mathrm{End}_K(V)$ bzgl. $\mathcal{B}$, so ist*

$$SAS^{-1} \text{ die Matrix von } f \text{ bzgl. } \mathcal{B}'.$$

Definition 2.3.14 *Zwei Matrizen $A, B \in M(n, K)$ heißen* **konjugiert** *(oder auch* **ähnlich***), wenn es ein $S \in \mathrm{GL}(n, K)$ gibt mit*

$$SAS^{-1} = B.$$

Dies ist eine Äquivalenzrelation, und die Äquivalenzklassen heißen **Konjugationsklassen** *quadratischer $n \times n$-Matrizen.*

Aus 2.3.13 folgt also: Die Matrizen eines Endomorphismus f bezüglich verschiedener Basen sind konjugiert. Insbesondere kann man jedem Endomorphismus eindeutig eine Konjugationsklasse zuordnen (die unabhängig von der Wahl einer Basis ist).

Bemerkung 2.3.15 Über gewissen Körpern K (den *algebraisch abgeschlossenen* Körpern, z.B. $K = \mathbb{C}$) wird es uns später möglich sein, die Konjugationsklassen quadratischer Matrizen genau zu bestimmen. Es stellt sich nämlich heraus, daß jede Konjugationsklasse eine Matrix einer besonders einfachen Form enthält, eine sogenannte *Jordanmatrix*; siehe Abschnitt 4.6.

2.4 Lineare Gleichungssysteme

Ein **lineares Gleichungssystem** mit Koeffizienten im Körper K ist eine Gleichung der Gestalt

$$Ax = b, \qquad A \in M(m \times n, K),\ x \in K^n,\ b \in K^m.$$

Hierbei sind die **Koeffizientenmatrix** A und der Spaltenvektor b gegeben, und wir suchen alle Lösungen dieser Gleichung, d.h., alle Spaltenvektoren $x \in K^n$, die diese Gleichung erfüllen. Im Falle $b = 0$ spricht man von einem **homogenen** Gleichungssystem, ansonsten von einem **inhomogenen**. Die Lösungsmenge des homogenen Gleichungssystems

$$Ax = 0$$

ist offenbar genau der *Kern* der Matrix A (aufgefaßt als lineare Abbildung $K^n \to K^m$), ist also ein Untervektorraum von K^n. Sei x_0 eine feste Lösung des inhomogenen Systems $Ax = b$, also

$$Ax_0 = b,$$

und sei $W \subset K^n$ der Lösungsraum des homogenen Systems $Ax = 0$. Dann ist der „verschobene Vektorraum"

$$x_0 + W = \{x_0 + x :\ x \in W\}$$

genau die Lösungsmenge des inhomogenen Systems $Ax = b$, wie man leicht sieht. Die Lösungen des Systems $Ax = b$ kann man also in zwei Schritten finden:

- Man bestimme irgendeine Lösung x_0.
- Man bestimme den Lösungsraum W des homogenen Systems $Ax = 0$.

Die Lösungen von $Ax = b$ sind dann genau die Vektoren $x_0 + W$. Diese Menge ist kein Vektorraum mehr (außer für $x_0 \in W$, also $b = 0$).

Homogene Systeme

Wir kümmern uns zunächst um die praktische Bestimmung des Lösungsraums des homogenen Systems $Ax = 0$. Ausgeschrieben lautet dieses System

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + \ldots + & a_{1n}x_n & = & 0, \\ a_{21}x_1 & + & a_{22}x_2 & + \ldots + & a_{2n}x_n & = & 0, \\ \vdots & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + \ldots + & a_{mn}x_n & = & 0. \end{array} \tag{2.18}$$

Es liegen also m Gleichungen in den n „Unbekannten“ $x_1, \dots, x_n$ vor. Wir beschreiben die drei sogenannten **elementaren Zeilenoperationen**, die auf das System (2.18) angewendet werden können, ohne den Lösungsraum zu ändern:

Typ I:	Vertauschung zweier Zeilen.
Typ II:	Addition eines Vielfachen einer Zeile zu einer anderen Zeile.
Typ III:	Multiplikation einer Zeile mit einem Skalar ungleich Null.

Jede dieser drei Operationen läßt sich rückgängig machen, ändert den Lösungsraum also in der Tat nicht. Die Idee ist nun, durch fortgesetzte Anwendung von elementaren Zeilenoperationen das System (2.18) auf eine einfache Form zu bringen, an der man den Lösungsraum direkt ablesen kann. Der zu beschreibende Algorithmus heißt das **Eliminationsverfahren von Gauß**[1].

Wir nehmen zunächst an, in der ersten Spalte von (2.18) ist einer der Koeffizienten $a_{11}, \dots, a_{m1}$ nicht Null, etwa a_{i1}. Dann multiplizieren wir die i-te Zeile mit a_{i1}^{-1} (Operation vom Typ III), und erhalten ein neues System mit $a_{i1} = 1$. Sodann wenden wir Operationen vom Typ II an, um alle Koeffizienten a_{j1} mit $j \neq i$ zu eliminieren: Dazu braucht man nur das $(-a_{j1})$-fache der i-ten Zeile zur j-ten Zeile zu addieren. Das Ergebnis ist ein neues System (2.18), dessen erste Spalte nur einen einzigen nichtverschwindenden Koeffizienten hat, nämlich $a_{i1} = 1$. Schließlich vertauschen wir noch die erste und i-te Zeile, so daß $a_{11} = 1$ und $a_{i1} = 0$ für $i > 1$.

Waren alle Koeffizienten in der ersten Spalte von vornherein Null, so hätten wir das gleiche Verfahren mit der zweiten Spalte durchgeführt, oder eben, falls diese auch Null ist, mit der ersten nichtverschwindenden Spalte (von links). Sei dies die k-te, so daß $a_{1k} = 1$.

Nun vergessen wir die erste Zeile und die ersten k Spalten, und wenden das eben beschriebene Verfahren auf das kleinere System

$$\begin{array}{ccccccc} a_{2,k+1}x_1 & + & a_{2,k+2}x_2 & + \dots + & a_{2n}x_n & = & 0, \\ a_{3,k+1}x_1 & + & a_{3,k+2}x_2 & + \dots + & a_{3n}x_n & = & 0, \\ \vdots & & & & & & \\ a_{m,k+1}x_1 & + & a_{m,k+2}x_2 & + \dots + & a_{mn}x_n & = & 0. \end{array} \tag{2.19}$$

an: In der ersten nicht völlig verschwindenden Spalte von links nehme man einen Koeffizienten ungleich Null, normiere ihn auf 1 (Typ III), eliminiere alle anderen Koeffizienten in dieser Spalte (Typ II), und bringe den einzigen verbleibenden Koeffizienten nach oben (Typ I).

Man sieht, daß es unnötig ist, die Unbekannten x_i jedesmal mit abzuschreiben, ebenso wie die Nullen auf der rechten Seite. Praktisch schreibt man nur die Matrix der Koeffizienten (a_{ij}) hin, und wendet die elementaren Zeilenoperationen auf diese an.

[1] Carl Friedrich Gauss, 1777-1855

Nach endlich vielen der eben beschriebenen Eliminationsschritte hat man schließlich die ursprüngliche $m \times n$-Koeffizientenmatrix auf **Zeilenstufenform** gebracht:

$$\begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & & & & & & & * \\ 0 & \dots & & & & 0 & 1 & * & \dots & & & & * \\ \vdots & & & & & & & & & & & & \\ 0 & \dots & & & & & & & 0 & 1 & * & \dots & * \\ 0 & \dots & & & & & & & & & & & 0 \\ \vdots & & & & & & & & & & & & \\ 0 & \dots & & & & & & & & & & & 0 \end{pmatrix}. \tag{2.20}$$

Die $*$ stehen für Koeffizienten, über die wir nichts aussagen können.

Wir vereinfachen die Zeilenstufenform weiter, indem wir *Spalten* vertauschen. Diese Operation führt das Gleichungssystem (2.18) zwar nicht in ein äquivalentes Gleichungssytem über (wie die elementaren Zeilenoperationen), entspricht aber lediglich einer *Umbennenung* der Unbekannten. Wenn wir z.B. die erste und zweite Spalte der Koeffizientenmatrix vertauschen, müssen wir x_1 in x_2 umbennen und x_2 in x_1. Führt man nun geeignete Spaltenvertauschungen und die entsprechenden Umbennenungen durch, so erreicht man eine Koeffizientenmatrix der speziellen Zeilenstufenform

$$\begin{pmatrix} 1 & * & \dots & & & & * \\ 0 & 1 & * & \dots & & & * \\ \vdots & & & & & & \\ 0 & \dots & 0 & 1 & * & \dots & * \\ 0 & \dots & & & & & 0 \\ \vdots & & & & & & \\ 0 & \dots & & & & & 0 \end{pmatrix}, \tag{2.21}$$

also mit Einsen an den Positionen $(1,1), \dots, (r,r)$, und in den restlichen $m-r$ Zeilen lauter Nullen. (Spaltenvertauschungen sind nicht wirklich notwendig; sie erleichtern aber die Notation im folgenden erheblich.)

Über die Lösungen des homogenen Systems mit Koeffizientenmatrix (2.21) kann man nun aber sofort einiges aussagen. Offenbar sind nämlich die Unbekannten $x_{r+1}, \dots, x_n$ *freie Variablen*, d.h., sie können beliebig vorgegeben werden. Dadurch ist dann x_r bestimmt (man betrachte die r-te Gleichung). Mit $x_r, \dots, x_n$ ist dann auch x_{r-1} bestimmt (durch die $(r-1)$-te Gleichung), und so weiter. Also: Die letzten $n-r$ Variablen sind beliebig, der Rest ist dann bestimmt. Insbesondere ist die Dimension des Lösungsraums (also des Kerns der Matrix) gleich $n-r$.

Die Lösungen von (2.21) sind aber nach Konstruktion genau die Lösungen des ursprünglichen Systems $Ax = 0$, wenn man noch die eventuell vorgenommenen Umbennenungen der Unbekannten berücksichtigt. Es folgt, daß die Zahl r in der Zeilenstufenform (2.21) oder (2.20) eindeutig durch A bestimmt ist:

$$\dim(\ker(A)) = n - r.$$

Aus der Dimensionsformel 2.1.1 v) folgt

$$r = \dim(\mathrm{im}(A)).$$

Definition 2.4.1 *Für* $A \in M(m \times n, K)$ *heißt* $r = \dim(\mathrm{im}(A))$ *der* **Rang** *von* A*, und wird mit* $\mathrm{rang}(A)$ *(oder* $\mathrm{rg}(A)$ *oder* $\mathrm{rank}(A)$ *oder* $\mathrm{rk}(A)$*) bezeichnet.*

Wir fassen die obigen Überlegungen zusammen:

Feststellung 2.4.2 *Gegeben sei das homogene lineare Gleichungssystem* $Ax = 0$ *mit der Koeffizientenmatrix* $A \in M(m \times n, K)$*. Durch fortgesetzte elementare Zeilenoperationen läßt sich* A *auf die Zeilenstufenform (2.20) oder sogar, nach eventueller Umbennenung der Unbekannten, auf die Form (2.21) bringen. Die Lösungen des entstehenden Systems, die sich leicht ablesen lassen, sind auch die Lösungen des ursprünglichen Systems* $Ax = 0$*. Die in der Zeilenstufenform auftretende Zahl* r *ist eindeutig bestimmt, nämlich* $r = \mathrm{rang}(A)$*.*

Zeilenrang und Spaltenrang

Wenn $A \in M(m \times n, K)$, so heißt die Dimension des von den *Spalten* von A aufgespannten Untervektorraums von K^m der **Spaltenrang** von A. Die Dimension des von den *Zeilen* von A aufgespannten Untervektorraums von K^n heißt der **Zeilenrang** von A. Unser Ziel ist zu zeigen, daß

$$\text{Zeilenrang} = \text{Spaltenrang} = \text{Rang}.$$

Die zweite Gleichung ist dabei klar: $\mathrm{rang}(A)$ ist die Dimension des Bildes der linearen Abbildung $A : K^n \to K^m$, also gleich der Dimension des Raumes

$$\langle Ae_1, \ldots, Ae_n \rangle, \qquad (e_1, \ldots, e_n) \text{ die kanonische Basis von } K^n,$$

und Ae_i ist die i-te Spalte von A.

Definition 2.4.3 *Sei* $A = (a_{ij}) \in M(m \times n, K)$*. Die zu* A **transponierte** *Matrix ist die Matrix*

$$^tA = (a_{ji}) = \begin{pmatrix} a_{11} & \ldots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \ldots & a_{mn} \end{pmatrix} \in M(n \times m, K).$$

Feststellung 2.4.4 *i)* $^t(A + B) = {}^tA + {}^tB$ *für* $A, B \in M(m \times n, K)$*.*

ii) $^t(\lambda A) = \lambda\, {}^tA$ *für* $A \in M(m \times n, K)$ *und* $\lambda \in K$*.*

iii) $^t({}^tA) = A$ *für alle* $A \in M(m \times n, K)$*.*

iv) $^t(AB) = {}^tB\, {}^tA$*, wann immer das Produkt* AB *definiert ist.*

v) Falls $A \in M(n, K)$ *invertierbar ist, so auch* ^tA*, und es gilt* $({}^tA)^{-1} = {}^t(A^{-1})$*.*

Beweis: i), ii) und iii) sind unmittelbar klar. iv) folgt nach einer leichten Rechnung. v) ergibt sich aus iv) (man beachte dabei Folgerung 2.3.5). ■

Bemerkung 2.4.5 Aus den Eigenschaften i) – iii) folgt insbesondere, daß die Transponierung einen Vektorraum–Isomorphismus $M(m \times n, K) \xrightarrow{\sim} M(n \times m, K)$ vermittelt.

Wir schreiben kurz $\mathrm{ZR}(A)$ (bzw. $\mathrm{SR}(A)$) für den Zeilenrang (bzw. Spaltenrang) der Matrix A. Da die Zeilen von A gerade die Spalten von ${}^tA$ sind und umgekehrt, gilt

$$\mathrm{ZR}(A) = \mathrm{SR}({}^tA) \qquad \text{und} \qquad \mathrm{SR}(A) = \mathrm{ZR}({}^tA).$$

Feststellung 2.4.6 *Der Spaltenrang ist invariant unter Multiplikation mit invertierbaren Matrizen. Genauer: Ist $A \in M(m \times n, K)$, $S \in \mathrm{GL}(m, K)$ und $T \in \mathrm{GL}(n, K)$, so gilt*

$$\mathrm{SR}(SAT) = \mathrm{SR}(A).$$

Das gleiche gilt für den Zeilenrang.

Beweis: Wir betrachten das kommutative Diagramm

$$\begin{array}{ccc}
K^n & \xrightarrow{A} & K^m \\
{\scriptstyle T}\uparrow{\scriptstyle \sim} & & {\scriptstyle \sim}\downarrow{\scriptstyle S} \\
K^n & \xrightarrow{SAT} & K^m.
\end{array}$$

Da S und T Isomorphismen sind, haben die Abbildungen A und SAT isomorphe Bilder. Insbesondere ist der Rang der beiden Matrizen gleich, und wie wir oben schon festgestellt haben, ist der Rang gleich dem Spaltenrang.

Für die Behauptung über den Zeilenrang rechnen wir

$$\mathrm{ZR}(SAT) = \mathrm{SR}({}^t(SAT)) = \mathrm{SR}({}^tT\,{}^tA\,{}^tS) = \mathrm{SR}({}^tA) = \mathrm{ZR}(A)$$

(hier wurde 2.4.4 iv) und v) benutzt). ■

Wie auf Seite 51 betrachten wir nun wieder *elementare Zeilenoperationen*, angewandt auf $m \times n$-Matrizen. Ganz analog erklären wir *elementare Spaltenoperationen* auf Matrizen. Diese Operationen lassen sich durch Multiplikation mit geeigneten invertierbaren Matrizen bewirken. Dazu definieren wir die **Elementarmatrizen**

$$S_{ij} = \left(\begin{array}{ccccccccccc}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 && \cdots && 1 &&& \\
&&&& 1 &&&&&& \\
&&& \vdots && \ddots && \vdots &&& \\
&&&&&& 1 &&&& \\
&&& 1 && \cdots && 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{array}\right)
\begin{array}{l}
\\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\
\end{array}$$

$$Q_{ij}(\lambda) = \begin{array}{c}
\left(\begin{array}{ccccccc}
1 &&&&&& \\
& \ddots &&&&& \\
&& 1 && \lambda && \\
&&& \ddots &&& \\
&&&& 1 && \\
&&&&& \ddots & \\
&&&&&& 1
\end{array}\right)
\begin{array}{l}
\\ \\ \leftarrow i \\ \\ \\ \\ \\
\end{array} \\
\uparrow \\
j
\end{array} \qquad\qquad (\lambda \in K)$$

$$D_i(\lambda) = \mathrm{diag}(1, \ldots, 1, \underset{\substack{\uparrow \\ i}}{\lambda}, 1, \ldots, 1) \qquad\qquad (\lambda \neq 0),$$

Diese drei Typen von quadratischen Matrizen sind alle invertierbar, denn

$$S_{ij}S_{ij} = \mathbf{1}, \qquad\qquad Q_{ij}(\lambda)Q_{ij}(-\lambda) = \mathbf{1}, \qquad\qquad D_i(\lambda)D_i(\lambda^{-1}) = \mathbf{1}.$$

Feststellung 2.4.7 *Sei $A \in M(m \times n, K)$. Elementare Zeilenumformungen (bzw. Spaltenumformungen) werden durch Multiplikation mit Elementarmatrizen bewirkt:*

- *Multiplikation mit S_{ij} von links (rechts) bewirkt die Vertauschung der i-ten und j-ten Zeile (Spalte) von A.*
- *Multiplikation mit $Q_{ij}(\lambda)$ von links (rechts) bewirkt die Addition des λ-fachen der j-ten Zeile (i-ten Spalte) zur i-ten (j-ten).*
- *Multiplikation mit $D_i(\lambda)$ von links (rechts) bewirkt die Multiplikation der i-ten Zeile (Spalte) mit λ.*

(Dabei sind die Elementarmatrizen vom Format $m \times m$ (bzw. $n \times n$)).

Beweis: Man rechnet die Produkte der Matrizen einfach aus. ∎

Folgerung 2.4.8 *Die Matrix $A \in M(m \times n, K)$ habe den Rang r. Dann gibt es $S \in \mathrm{GL}(m, K)$ und $T \in \mathrm{GL}(n, K)$, so daß*

$$SAT = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \tag{2.22}$$

(Nullen außerhalb der Hauptdiagonalen).

Beweis: Durch elementare Zeilentransformationen läßt sich A zunächst auf Zeilenstufenform bringen. Diese Umformungen werden nach 2.4.7 durch Multiplikation mit invertierbaren Matrizen von links bewirkt. Also gibt es ein $S \in \mathrm{GL}(m, K)$, so daß

$$SA$$

Zeilenstufenform hat. Eine Zeilenstufenform läßt sich aber durch elementare Spaltenumformungen auf die Form (2.22) bringen, wie man leicht sieht. Diese Spaltenumformungen werden gemäß 2.4.7 durch Multiplikation mit invertierbaren Matrizen von rechts bewirkt. Also gibt es ein $T \in \mathrm{GL}(n, K)$, so daß SAT die gewünschte Form hat.

Die dabei auftretende Zahl r ist gleich dem Spaltenrang von SAT, also gleich dem Spaltenrang von A (Feststellung 2.4.6). Wir haben schon festgestellt, daß der Rang gleich dem Spaltenrang ist. ∎

Folgerung 2.4.9 *Für jede Matrix $A \in M(m \times n, K)$ ist*

$$\text{Zeilenrang}(A) = \text{Spaltenrang}(A) = \text{rang}(A).$$

Beweis: Wir wählen S und T wie in Folgerung 2.4.8. Der Zeilenrang (bzw. Spaltenrang) von A ist gleich demjenigen von SAT nach 2.4.6. Der Zeilenrang sowie der Spaltenrang von SAT ist aber offenbar gleich r, dem Rang von A. ∎

Bemerkungen: 1. Aufgrund der Aussage von 2.4.9 sprechen wir in Zukunft kaum mehr von Zeilenrang und Spaltenrang, sondern schlicht vom Rang einer Matrix.

2. Der maximale Rang, den eine Matrix $A \in M(m \times n, K)$ haben kann, ist nach der letzten Folgerung gleich $\min(m, n)$. In diesem Fall sagt man, A habe **vollen Rang**.

3. Eine quadratische Matrix hat genau dann vollen Rang, wenn sie invertierbar ist.

4. Ein wesentlich abstrakterer Beweis der Aussage „Zeilenrang = Spaltenrang“ wird sich im nächsten Abschnitt ergeben, siehe Seite 62.

Inhomogene Systeme

Wir betrachten nun das *inhomogene* lineare Gleichungssystem

$$Ax = b, \qquad A \in M(m \times n, K),\ b \in K^m.$$

Um zu entscheiden, ob dieses System überhaupt eine Lösung $x \in K^m$ besitzt, dient wieder das Gaußsche Eliminationsverfahren, mit dessen Hilfe der Lösungsraum des zugehörigen homogenen Systems $Ax = 0$ bestimmt werden konnte. Im positiven Falle liefert das Verfahren auch gleich eine Lösung.

Wir bilden die **erweiterte Koeffizientenmatrix** $\tilde{A}$, indem wir an A noch die Spalte b rechts anhängen:

$$\tilde{A} = (A, b) \in M(m \times (n+1), K).$$

Man mache sich nun folgendes klar: Wendet man auf $\tilde{A}$ eine elementare Zeilentransformation an, und interpretiert die entstehende Matrix wieder als inhomogenes lineares Gleichungssytem, so ist dieses neue System zum ursprünglichen System $Ax = b$ äquivalent (d.h., $x \in K^n$ ist genau dann eine Lösung des neuen Systems, wenn es eine Lösung des alten Systems ist). Dies überlegt man sich leicht, indem man die Gleichungen ausschreibt.

Genau wie im Falle des homogenen Systems bringt man nun die ersten n Spalten von $\tilde{A}$ auf Zeilenstufenform. Man erhält also eine Matrix der Gestalt

$$\begin{pmatrix}
0 & \dots & 0 & 1 & * & \dots & & & & & & * & b'_1 \\
0 & \dots & & & & 0 & 1 & * & \dots & & & * & \vdots \\
\vdots & & & & & & & & & & & & \\
0 & \dots & & & & & & & 0 & 1 & * \ \dots & * & b'_r \\
0 & \dots & & & & & & & & & & 0 & \vdots \\
\vdots & & & & & & & & & & & & \\
0 & \dots & & & & & & & & & & 0 & b'_m
\end{pmatrix} \tag{2.23}$$

mit gewissen $b'_1, \dots, b'_m$, und r als dem Rang der Matrix A. Der Einfachheit halber benennen wir die Unbekannten x_i jetzt noch so um, daß diese Matrix die Gestalt

$$\begin{pmatrix}
1 & * & \dots & & & & & * & b'_1 \\
0 & 1 & * & \dots & & & & * & \vdots \\
\vdots & & & & & & & & \\
0 & \dots & 0 & 1 & * & \dots & & * & b'_r \\
0 & \dots & & & & & & 0 & \vdots \\
\vdots & & & & & & & & \\
0 & \dots & & & & & & 0 & b'_m
\end{pmatrix} \tag{2.24}$$

annimmt. Schreibt man sich das entstehende lineare Gleichungssystem aus, so kann man folgende Aussagen über die Lösbarkeit daran ablesen:

- Ist einer der Koeffizienten $b'_{r+1}, \ldots, b'_m$ ungleich Null, so hat das Gleichungssytem keine Lösung.
- Angenommen, es gilt $b'_{r+1} = \ldots = b'_m = 0$. Dann besitzt das Gleichungssytem eine Lösung. In der Tat: Man kann die Unbekannten $x_{r+1}, \ldots, x_n$ beliebig wählen, und dann nach $x_1, \ldots, x_r$ auflösen.
- Insbesondere gibt es genau dann eine *eindeutige* Lösung, wenn $b'_{r+1} = \ldots = b'_m = 0$ und $r = n$.

2.5 Der Dualraum

In Abschnitt 2.3 haben wir den Vektorraum $\mathrm{Hom}_K(V, W)$ für zwei endlich-dimensionale K-Vektorräume V und W untersucht. Jetzt betrachten wir speziell den Fall $W = K$.

Dualraum und duale Basis

Definition 2.5.1 *Für einen K-Vektorraum V heißt*

$$V^* := \mathrm{Hom}_K(V, K)$$

der **Dualraum** *von V. Seine Elemente, also die linearen Abbildungen $V \to K$, heißen* **Linearformen** *auf V.*

In der nächsten Feststellung findet das **Kroneckersymbol** [2] δ_{ij} Verwendung, definiert durch

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} \qquad (i, j \in \mathbb{Z}).$$

Feststellung und Definition 2.5.2 *Sei V ein n-dimensionaler K-Vektorraum.*

i) Der Dualraum V^* hat ebenfalls die Dimension n.

ii) Sei $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basis von V. Dann gibt es genau eine Basis $\mathcal{B}^* = (b_1^*, \ldots, b_n^*)$ von V^* mit der Eigenschaft

$$b_j^*(b_i) = \delta_{ij} \qquad \text{für alle } i, j \in \{1, \ldots, n\}. \tag{2.25}$$

$\mathcal{B}^*$ heißt die zu $\mathcal{B}$ **duale Basis***.*

Beweis: i) ergibt sich aus Folgerung 2.3.8.

ii) Jedes b_j^* ist durch die Werte auf der Basis $\mathcal{B}$ festgelegt; also kann es höchstens eine Basis mit der Eigenschaft (2.25) geben. Umgekehrt kann b_j^* durch (2.25) definiert werden, siehe 1.4.9. Man hat dann nur noch zu zeigen, daß $\mathcal{B}^*$ tatsächlich eine Basis von V^* ist. Da $\dim_K(V^*) = n$ schon bekannt ist, genügt es, die lineare Unabhängigkeit der b_j^* zu zeigen. Sei also

$$a_1 b_1^* + \ldots + a_n b_n^* = 0$$

[2] Leopold Kronecker, 1823-1891

eine Linearkombination, die Null darstellt (dies ist eine Gleichheit linearer Abbildungen; rechts steht die Nullabbildung). Auswerten am Element b_i liefert wegen (2.25) sofort $a_i = 0$. Also ist die Linearkombination trivial. ∎

Beispiel: Der Standardraum K^n sei als Raum von Spaltenvektoren notiert. Gemäß Feststellung 2.3.1 identifiziert sich der Dualraum $(K^n)^*$ mit $M(1 \times n, K)$, also mit n-komponentigen *Zeilenvektoren.* Der Zeilenvektor $(x_1, \ldots, x_n)$ identifiziert sich dabei mit der Linearform

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \longmapsto (x_1, \ldots, x_n) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1 x_1 + \ldots + a_n x_n$$

(Matrixmultiplikation des Zeilenvektors mit dem Spaltenvektor). Die zur kanonischen Basis des K^n duale Basis ist die kanonische Basis von $M(1 \times n, K)$.

Die duale Abbildung

Definition 2.5.3 *Sei $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Die zu f* **duale Abbildung** *ist die Abbildung*

$$\begin{aligned} f^* : W^* &\longrightarrow V^*, \\ \varphi &\longmapsto \varphi \circ f. \end{aligned}$$

Jede Linearform φ auf W wird also zu einer Linearform $f^*(\varphi) = \varphi \circ f$ auf V „zurückgeholt“:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "f^*(\varphi)"'] & W \arrow[d, "\varphi"] \\ & K \end{tikzcd}$$

Feststellung 2.5.4 *Die Zuordnung $f \mapsto f^*$ hat die folgenden Eigenschaften:*

i) f^* *ist eine* lineare *Abbildung* $W^* \to V^*$.

ii) $(\mathrm{id}_V)^* = \mathrm{id}_{V^*}$.

iii) Falls $U \xrightarrow{g} V \xrightarrow{f} W$ lineare Abbildungen sind, und $U^* \xleftarrow{g^*} V^* \xleftarrow{f^*} W^*$ die zugehörigen dualen Abbildungen, so gilt

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

Beweis: i) Für Linearformen $\varphi_1, \varphi_2 \in W^*$ gilt

$$f^*(\varphi_1 + \varphi_2) = (\varphi_1 + \varphi_2) \circ f = \varphi_1 \circ f + \varphi_2 \circ f = f^*(\varphi_1) + f^*(\varphi_2),$$

wobei man sich das mittlere Gleichheitszeichen völlig klarmache, indem man beide Seiten auf einem $v \in V$ auswertet. Ganz ähnlich gilt für $\varphi \in W^*$ und $a \in K$

$$f^*(a\varphi) = (a\varphi) \circ f = a(\varphi \circ f) = a(f^*\varphi).$$

ii) ist klar.

iii) Für $\varphi \in W^*$ gilt

$$(f \circ g)^*(\varphi) = \varphi \circ (f \circ g) = (\varphi \circ f) \circ g = (f^*(\varphi)) \circ g = g^*(f^*(\varphi)) = (g^* \circ f^*)(\varphi).$$

∎

Bemerkung 2.5.5 Die Bildung des Dualraums und der dualen Abbildung ist ein Beispiel für einen *Funktor*, einer Art Abbildung zwischen *Kategorien*. Dies wird in Anhang B.2 erläutert.

Feststellung 2.5.6 *Sei* $\dim_K(V) = n$, $\dim_K(W) = m$, *und* $f : V \to W$ *eine lineare Abbildung. Sei* $\mathcal{B}$ *eine Basis von* V, $\mathcal{C}$ *eine Basis von* W, *und* $A \in M(m \times n, K)$ *die Matrix von* f *bzgl. dieser Basen. Dann ist die transponierte Matrix* ${}^tA$ *die Matrix von* $f^* : W^* \to V^*$ *bezüglich der dualen Basen* $\mathcal{C}^*$ *und* $\mathcal{B}^*$*:*

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ K^n \arrow[u, "\varphi_{\mathcal{B}}", "\sim"'] \arrow[r, "A"] & K^m \arrow[u, "\sim", "\varphi_{\mathcal{C}}"'] \end{tikzcd} \quad \Longrightarrow \quad \begin{tikzcd} V^* & W^* \arrow[l, "f^*"'] \\ K^n \arrow[u, "\varphi_{\mathcal{B}^*}", "\sim"'] & K^m \arrow[l, "{}^tA"'] \arrow[u, "\sim", "\varphi_{\mathcal{C}^*}"'] \end{tikzcd}$$

Beweis: Sei

$$\mathcal{B} = (v_1, \dots, v_n), \qquad \mathcal{C} = (w_1, \dots, w_m), \qquad A = (a_{ij}).$$

Dann gilt

$$f(v_k) = \sum_{l=1}^{m} a_{lk} w_l \qquad \text{für } k = 1, \dots, n.$$

nach Definition von A. Sei $B = (b_{ij}) \in M(n \times m, K)$ die Matrix von f^* bezüglich der dualen Basen, so daß

$$f^*(w_j^*) = \sum_{i=1}^{n} b_{ij} v_i^* \qquad \text{für } j = 1, \dots, m. \tag{2.26}$$

Wir haben zu zeigen, daß $b_{ij} = a_{ji}$. Dazu werten wir (2.26) an der Stelle v_k aus. Die rechte Seite ergibt

$$\sum_i b_{ij} v_i^*(v_k) = \sum_i b_{ij} \delta_{ik} = b_{kj}.$$

Die linke Seite ergibt

$$(f^*(w_j^*))(v_k) = w_j^*(f(v_k)) = w_j^*\Big(\sum_l a_{lk} w_l\Big) = \sum_l a_{lk} \delta_{jl} = a_{jk}.$$

Also ist tatsächlich $b_{kj} = a_{jk}$. ∎

Exaktheit

Eine bemerkenswerte Eigenschaft des Dualisierens ist die sogenannte *Exaktheit*:

Feststellung 2.5.7 *Sei* $U \xrightarrow{g} V \xrightarrow{f} W$ *eine exakte Sequenz endlich-dimensionaler* K*-Vektorräume. Dann ist die duale Sequenz*

$$U^* \xleftarrow{g^*} V^* \xleftarrow{f^*} W^*$$

ebenfalls exakt.

Beweis: Wir haben zu zeigen, daß

$$\mathrm{im}(f^*) = \ker(g^*).$$

„$\subset$" ist einfach, denn

$$g^*(f^*(\varphi)) = \varphi \circ (f \circ g) = \varphi \circ 0 = 0 \qquad \text{für } \varphi \in W^*.$$

Wir zeigen „$\supset$". Sei also $\psi \in V^*$ im Kern von g^*, d.h. $\psi \circ g = 0$:

$$\begin{tikzcd}
U \arrow[r, "g"] \arrow[dr, "0"'] & V \arrow[r, "f"] \arrow[d, "\psi"] & W \\
& K &
\end{tikzcd}$$

Die Bedingung $\psi \circ g = 0$ ist äquivalent mit $\mathrm{im}(g) \subset \ker(\psi)$. Wegen der Exaktheit der Ausgangssequenz haben wir also

$$\ker(f) \subset \ker(\psi).$$

Nach dem Homomorphiesatz 1.5.7 induziert ψ eine lineare Abbildung $\bar{\psi} : V/\ker(f) \to K$, und f eine *injektive* lineare Abbildung $\bar{f} : V/\ker(f) \to W$. Sei $\varphi : \mathrm{im}(f) \to K$ die eindeutig bestimmte lineare Abbildung, die das Diagramm

$$\begin{tikzcd}
V/\ker(f) \arrow[r, "\bar{f}", "\sim"'] \arrow[d, "\bar{\psi}"'] & \mathrm{im}(f) \arrow[dl, "\varphi"] \\
K &
\end{tikzcd}$$

kommutativ macht. φ läßt sich (auf verschiedene Weisen) zu einer Linearform auf ganz W fortsetzen (wie etwa aus dem Basisergänzungssatz und 1.4.9 folgt). Wir erhalten dann ein kommutatives Diagramm

$$\begin{tikzcd} U \arrow[r, "g"] \arrow[dr, "0"'] & V \arrow[r, "f"] \arrow[d, "\psi"] & W \arrow[dl, "\varphi"] \\ & K & \end{tikzcd}$$

Es folgt $f^*(\varphi) = \psi$, also $\varphi \in \mathrm{im}(f^*)$, was zu beweisen war. ∎

Folgerung 2.5.8 *Es gilt* $\dim_K(\mathrm{im}(f)) = \dim_K(\mathrm{im}(f^*))$.

Beweis: Wir betrachten die exakte Sequenz

$$0 \longrightarrow \ker(f) \xrightarrow{\iota} V \xrightarrow{f} W.$$

Nach 2.5.7 ist die duale Sequenz

$$0 \longleftarrow \ker(f)^* \xleftarrow{\iota^*} V^* \xleftarrow{f^*} W^*$$

ebenfalls exakt. Deshalb gilt

$$\begin{aligned} \dim(\mathrm{im}(f^*)) &= \dim(\ker(\iota^*)) \\ &= \dim(V^*) - \dim(\mathrm{im}(\iota^*)) \\ &= \dim(V^*) - \dim(\ker(f)^*) \\ &= \dim(V) - \dim(\ker(f)) = \dim(\mathrm{im}(f)), \end{aligned}$$

wobei wir zweimal 2.1.1 v) benutzt haben. ∎

Als Anwendung der Begriffe dieses Abschnitts erhalten wir einen neuen Beweis der Aussage „Zeilenrang = Spaltenrang“, siehe Folgerung 2.4.9. Denn sei $A \in M(m \times n, K)$. Nach 2.5.6 und 2.5.8 gilt

$$\dim(\mathrm{im}({}^tA)) = \dim(\mathrm{im}(A^*)) = \dim(\mathrm{im}(A)),$$

und folglich

$$\text{Zeilenrang}(A) = \text{Spaltenrang}({}^tA) = \dim(\mathrm{im}({}^tA)) = \dim(\mathrm{im}(A)) = \text{Spaltenrang}(A).$$

Der Bidualraum

Sei V ein n-dimensionaler K-Vektorraum. Wir untersuchen den Raum

$$V^{**} = (V^*)^*,$$

also den Dualraum des Dualraumes von V. Dieser heißt der **Bidualraum** von V. Nach 2.5.2 gilt $\dim_K(V^{**}) = n$.

Jedem Vektor $v \in V$ ordnen wir die lineare Abbildung

$$\varphi_v : V^* \longrightarrow K,$$

$$f \longmapsto f(v),$$

zu (man prüfe nach, daß tatsächlich $\varphi_v(f_1 + f_2) = \varphi_v(f_1) + \varphi_v(f_2)$ und $\varphi_v(af) = a\varphi_v(f)$ gilt). Dies ist die „Auswertungsabbildung an der Stelle v". Offenbar ist φ_v eine Linearform auf V^*, also ein Element des Bidualraums V^{**}.

Feststellung 2.5.9 *Es gelten obige Notationen.*

i) Die Räume V und V^{**} sind <u>kanonisch</u> isomorph via

$$\begin{aligned} V &\xrightarrow{\sim} V^{**}, \\ v &\longmapsto \varphi_v. \end{aligned} \tag{2.27}$$

ii) Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V, und $\mathcal{B}^* = (v_1^*, \ldots, v_n^*)$ die dazu duale Basis (von V^*). Dann ist $\mathcal{B}^{**} = (\varphi_{v_1}, \ldots, \varphi_{v_n})$ die zu $\mathcal{B}^*$ duale Basis (von V^{**}).

Beweis: i) Zunächst ist die Abbildung (2.27) linear, wie man leicht nachrechnet. Wegen 2.1.4 müssen wir nur noch zeigen, daß sie injektiv ist. Sei also $v \in V$ im Kern, d.h. $\varphi_v = 0$. Dies bedeutet

$$f(v) = \varphi_v(f) = 0 \qquad \text{für alle } f \in V^*.$$

Dann muß aber $v = 0$ sein, denn sonst könnte man v zu einer Basis von V ergänzen, und (mittels 1.4.9) ein $f \in V^*$ konstruieren, das auf v nicht den Wert Null annimmt.

ii) Nach i) ist $\mathcal{B}^{**}$ eine Basis von V^{**}, und offenbar gilt $\varphi_{v_i}(v_j^*) = v_j^*(v_i) = \delta_{ij}$. ∎

Aufgrund des *kanonischen* Isomorphismus 2.27 kann man V mit V^{**} identifizieren; ein $v \in V$ entspricht dabei der *Auswertungsabbildung* in v. Wird diese Identifizierung vorgenommen, so läßt sich ii) auch so ausdrücken: Die duale Basis der dualen Basis zu $\mathcal{B}$ ist $\mathcal{B}$ selbst.

Bemerkung 2.5.10 Die Vektorräume V, V^* und V^{**} haben alle die gleiche Dimension n, sind nach dem „Hauptsatz" 2.1.5 also alle paarweise isomorph. Während es jedoch, wie eben gesehen, einen kanonischen Isomorphismus $V \xrightarrow{\sim} V^{**}$ gibt, gibt es *keine* kanonische Identifizierung von V mit V^*.

Kapitel 3

Determinanten

Jeder quadratischen Matrix $A \in M(n, K)$ ist ein Skalar $\det(A) \in K$ zugeordnet, ihre *Determinante*. Die Abbildung $A \mapsto \det(A)$ hat zahlreiche bemerkenswerte Eigenschaften, wie etwa $\det(AB) = \det(A)\det(B)$, was die Determinante zu einem wichtigen Hilfsmittel macht. Für die Theorie der Determinanten wird jedoch nur von der *Ring*struktur des Körpers K Gebrauch gemacht, und wir arbeiten daher gleich allgemeiner über einem kommutativen Ring R.

Vorbereitend bringen wir einiges Material über symmetrische Gruppen, die auch in anderen Zusammenhängen wichtig sind. Abschnitt 3.2 formuliert die wichtigsten Eigenschaften der Determinante, und die beiden folgenden Abschnitte dienen deren Beweis. Schließlich definieren wir die Determinante (und Spur) eines Endomorphismus eines endlich-dimensionalen Vektorraumes, was sich basisunabhängig tun läßt.

3.1 Die symmetrische Gruppe

Sei M eine beliebige Menge. Eine **Permutation** von M ist eine bijektive Selbstabbildung

$$\sigma : M \longrightarrow M.$$

Die Menge aller Permutationen von M bildet eine Gruppe mit der Komposition (Hintereinanderschaltung) als Verknüpfung, wie leicht einzusehen ist.

Definition 3.1.1 *Die* **symmetrische Gruppe** $\mathcal{S}_n$ *vom Grade* n *ist die Gruppe aller Permutationen der Menge* $\{1, \ldots, n\}$.

Beispiele: $\mathcal{S}_1$ besteht aus nur einem Element und $\mathcal{S}_2$ aus zwei Elementen. Diese beiden Gruppen sind natürlich kommutativ. Für $n \geq 3$ dagegen ist $\mathcal{S}_n$ nicht mehr kommutativ.

Feststellung 3.1.2 *Die symmetrische Gruppe* $\mathcal{S}_n$ *hat* $n! = 1 \cdot 2 \cdot \ldots \cdot n$ *Elemente.*

Beweis: Es gibt n Möglichkeiten, wohin das Element 1 abgebildet wird, dann noch $n-1$ Möglichkeiten, wohin 2 abgebildet wird, usw. ∎

Spezielle Elemente von $\mathcal{S}_n$ sind die sogenannten *Zyklen*. Seien $a_1, \ldots, a_m \in \{1, \ldots, n\}$, und sei σ die Permutation mit

$$\begin{aligned} &\sigma(a_i) = a_{i+1} \quad \text{für } i = 1, \ldots, m-1, \qquad \sigma(a_m) = a_1, \\ &\text{und } \sigma(b) = b \quad \text{für } b \in \{1, \ldots, n\} \setminus \{a_1, \ldots, a_m\}. \end{aligned}$$

Dann heißt σ aus naheliegenden Gründen ein **Zyklus** der Länge m. Diesen Zyklus beschreibt man auch durch das Symbol

$$\sigma = (a_1 \ \ldots \ a_m)$$

(oder ebensogut durch $\sigma = (a_2 \ \ldots \ a_m \ a_1)$, usw.). Zyklen der Länge 2 heißen **Transpositionen**; dies sind also die Permutationen

$$\tau = (ij) \qquad \text{mit } i, j \in \{1, \ldots, n\},\ i \neq j.$$

Offenbar gilt $\tau^2 = 1 (= \mathrm{id})$ für jede Transposition τ.

Lemma 3.1.3 *Alle Transpositionen sind zueinander konjugiert, d.h., sind $\tau, \tau' \in \mathcal{S}_n$ Transpositionen, so gibt es ein $\sigma \in \mathcal{S}_n$ mit*

$$\tau' = \sigma \tau \sigma^{-1}.$$

Beweis: Wir dürfen annehmen, daß $\tau' = (12)$. Sei $\tau = (ij)$. Für $\{i, j\} = \{1, 2\}$ nehme man $\sigma = \mathrm{id}$. Ansonsten können zwei Fälle auftreten:

<u>1. Fall:</u> $\{1, 2\} \cap \{i, j\} = \emptyset$. Dann gilt

$$(12) = (2j)(1i)(ij)(1i)(2j),$$

wie man leicht nachrechnet. Man kann also $\sigma = (2j)(1i)$ nehmen.

<u>2. Fall:</u> $\{1, 2\} \cap \{i, j\} \neq \emptyset$, etwa $i = 1$. Es gilt

$$(12) = (2j)(1j)(2j),$$

also kann man $\sigma = (2j)$ nehmen. ∎

Gerade und ungerade Permutationen

Die **Parität**[1] einer ganzen Zahl $r \in \mathbb{Z}$ ist *gerade*, falls $r \in 2\mathbb{Z}$, und *ungerade*, falls $r \in 2\mathbb{Z} + 1$.

Feststellung und Definition 3.1.4 *Sei n eine natürliche Zahl.*

i) Es gibt einen eindeutig bestimmten Homomorphismus

$$\mathrm{sgn} : \ \mathcal{S}_n \longrightarrow \{\pm 1\},$$

so daß $\mathrm{sgn}(\tau) = -1$ *für jede Transposition τ. Dieser heißt der* **Vorzeichenhomomorphismus** *(signum).*

---

[1] lat. *paritas*: Gleichheit, Gleichstellung

ii) Jedes Element $\sigma \in \mathcal{S}_n$ läßt sich als Produkt von Transpositionen darstellen:

$$\sigma = \tau_1 \ldots \tau_r, \qquad \tau_i \text{ Transpositionen.}$$

Diese Darstellung ist nicht eindeutig, aber die Parität der Zahl r ist eindeutig bestimmt. Es gilt $\operatorname{sgn}(\sigma) = (-1)^r$.

Der Kern von sgn *heißt die* **alternierende Gruppe** *vom Grade n. Symbol:*

$$A_n = \{\sigma \in \mathcal{S}_n : \operatorname{sgn}(\sigma) = 1\}.$$

Beweis: Wir definieren sgn durch

$$\operatorname{sgn}(\sigma) = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Da σ eine Bijektion ist, treten im Zähler die gleichen Differenzen auf wie im Nenner, aber eventuell mit unterschiedlichen Vorzeichen. Deshalb ist tatsächlich $\operatorname{sgn}(\sigma) \in \{\pm 1\}$. Wir zeigen, daß sgn ein Homomorphismus ist. In der Tat, für $\sigma, \tau \in \mathcal{S}_n$ gilt

$$\begin{aligned}\operatorname{sgn}(\sigma\tau) &= \prod_{i<j} \left(\frac{(\sigma\tau)(j) - (\sigma\tau)(i)}{\tau(j) - \tau(i)} \cdot \frac{\tau(j) - \tau(i)}{j - i} \right) \\ &= \left(\prod_{i<j} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \right) \left(\prod_{i<j} \frac{\tau(j) - \tau(i)}{j - i} \right) \\ &= \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau).\end{aligned}$$

(Zum linken Produkt in der zweiten Zeile: Mit (i, j) durchläuft auch $(\tau(i), \tau(j))$ alle Paare mit nichtgleichen Komponenten. Für die Paare mit $\tau(i) > \tau(j)$ multipliziere man Zähler und Nenner mit -1.)

Man sieht sofort, daß $\operatorname{sgn}((12)) = -1$, und dann, mit Hilfe von Lemma 3.1.3, daß $\operatorname{sgn}(\tau) = -1$ für jede Transposition τ. Damit ist die Existenzaussage in i) bewiesen.

Wir zeigen die erste Aussage in ii) durch Induktion nach der Anzahl der Elemente von $\{1, \ldots, n\}$, die von σ *nicht* festgelassen werden. Falls diese Anzahl Null ist, ist nichts zu beweisen. Es gebe also ein $i \in \{1, \ldots, n\}$, daß von σ nicht festgelassen wird. Sei $j = \sigma(i)$ und $\tau = (ij)$. Die Permutation

$$\tilde{\sigma} := \tau\sigma$$

läßt dann mehr Elemente fest als σ. Nach Induktionsvoraussetzung läßt sich $\tilde{\sigma}$ als Produkt von Transpositionen schreiben; damit aber auch $\sigma = \tau\tilde{\sigma}$.

Anwenden von sgn auf $\sigma = \tau_1 \ldots \tau_r$ ergibt $\operatorname{sgn}(\sigma) = (-1)^r$. Also ist die Parität von r durch σ bestimmt, und hängt nicht von der konkreten Darstellung ab. Außerdem ist jetzt klar, daß es nur diese eine Funktion sgn mit den gewünschten Eigenschaften geben kann. ∎

Es gibt also zwei Typen von Elementen von $\mathcal{S}_n$: Die *geraden* Permutationen, die sich als Produkt einer geraden Anzahl von Transpositionen schreiben lassen, und die *ungeraden* Permutationen, die ein Produkt einer ungeraden Anzahl von Transpositionen sind. Die geraden Permutationen bilden die Untergruppe A_n.

Permutationsmatrizen

Definition 3.1.5 *Eine* **Permutationsmatrix** *ist eine Matrix $A \in M(n, K)$, die in jeder Zeile und jeder Spalte genau eine* 1 *besitzt, und sonst nur Nullen.*

Feststellung 3.1.6 *Sei $n \in \mathbb{N}$.*

i) Jede Permutationsmatrix $A \in M(n, K)$ ist invertierbar mit $A^{-1} = {}^tA$.

ii) Es gibt genau $n!$ Permutationsmatrizen in $\mathrm{GL}(n, K)$.

iii) Die Permutationsmatrizen bilden eine Untergruppe von $\mathrm{GL}(n, K)$.

Beweis: All dies ist leicht einzusehen. ∎

Wir ordnen jeder Permutation $\sigma \in \mathcal{S}_n$ eine Permutationsmatrix A_σ zu, definiert durch

$$A_\sigma e_i = e_{\sigma(i)}$$

(e_i die kanonischen Basisvektoren des K^n). Explizit: Falls $A_\sigma = (a_{ij})$, so ist

$$a_{ij} = \begin{cases} 1, & \text{falls } \sigma(j) = i, \\ 0, & \text{falls } \sigma(j) \neq i. \end{cases}$$

Feststellung 3.1.7 *Die Abbildung $\sigma \mapsto A_\sigma$ ist ein Gruppenisomorphismus von $\mathcal{S}_n$ auf die Gruppe der Permutationsmatrizen in $\mathrm{GL}(n, K)$.*

Beweis: Die Abbildung ist ein Gruppenhomomorphismus, denn

$$A_\sigma A_\tau e_i = A_\sigma e_{\tau(i)} = e_{\sigma(\tau(i))} = A_{\sigma\tau} e_i,$$

und also $A_\sigma A_\tau = A_{\sigma\tau}$. Falls $A_\sigma = \mathbf{1}$, so muß $\sigma = \mathrm{id}$ sein. Also ist der Homomorphismus injektiv. Aber beide Gruppen haben die gleiche Ordnung $n!$, und folglich handelt es sich um einen Isomorphismus. ∎

3.2 Die Determinante

Die Determinante ist eine Abbildung, die jeder Matrix $A \in M(n, K)$ ein Element $\det(A) \in K$ zuordnet. Wir legen jedoch anstelle des Körpers K allgemeiner einen kommutativen Ring mit 1 zugrunde. Obwohl wir in diesem Buch fast ausschließlich Determinanten über Körpern verwenden, rechtfertigt sich diese größere Allgemeinheit aus folgenden Gründen:

- Der Determinantenbegriff über kommutativen Ringen ist in anderen Zusammenhängen ein wichtiges Hilfsmittel.
- Die Sätze und Beweise lassen sich praktisch ohne Mehraufwand für Ringe formulieren.
- Auch wir werden den allgemeineren Determinantenbegriff an einer (allerdings unproblematischen) Stelle verwenden: Bei der Definition des charakteristischen Polynoms bildet man die Determinante einer Matrix mit Koeffizienten in einem Polynomring (siehe 4.2.1).

- Die Determinante ist ihrem Wesen nach „ringartig“. Wir wollten nicht durch die Beschränkung auf den Körperfall den Eindruck hervorrufen, daß die Körpereigenschaften wesentlich benutzt werden.

Sei also hier und im folgenden

R ein kommutativer Ring mit 1.

Ein *R-Modul* ist das Gegenstück zu einem K-Vektorraum:

Definition 3.2.1 *Ein R-**Modul** oder **Modul über** R ist eine abelsche Gruppe $(V,+)$, auf der R linear operiert. Genauer: Es ist eine Abbildung*

$$\begin{aligned} R \times V &\longrightarrow V, \\ (a,v) &\longmapsto a.v \end{aligned}$$

definiert, so daß folgendes gilt:

i) $(a+b).v = a.v + b.v$ *für alle* $a,b \in R$, $v \in V$.

ii) $a.(v+w) = a.v + a.w$ *für alle* $a \in R$, $v,w \in V$.

iii) $(a \cdot b).v = a.(b.v)$ *für alle* $a,b \in R$, $v \in V$.

iv) $1.v = v$ *für alle* $v \in V$.

Die für uns wichtigsten R-Moduln sind R^n, die Menge aller Spaltenvektoren (oder Zeilenvektoren) mit den natürlichen Verknüpfungen, und $M(n,R)$, die Menge aller $n \times n$-Matrizen mit Koeffizienten aus R. Letzteres ist nicht nur ein R-Modul, sondern auch ein Ring, indem die Multiplikation durch die übliche Matrixmultiplikation definiert wird. Dies ist alles analog zum Körperfall.

Bemerkung 3.2.2 Der Hauptunterschied zwischen Vektorräumen und Moduln ist, daß letztere im allgemeinen keine Basis besitzen. Deshalb gibt es auch keinen Dimensionsbegriff für Moduln. Die im vorherigen Kapitel entwickelte Theorie beruhte wesentlich auf der Existenz von Basen, und nur wenig davon besitzt ein Analogon für Moduln über einem Ring.

In diesem Abschnitt wird der Existenzsatz und die wichtigsten Eigenschaften der Determinante formuliert. Die Beweise dieser Aussagen werden in den beiden folgenden Abschnitten nachgeliefert.

Wir sagen, eine Abbildung

$$\mu: \ M(m \times n, R) \longrightarrow R$$

ist *linear in der ersten Spalte*, falls

$$\mu(\underbrace{v_1 + v_1', v_2, \dots, v_n}_{m \times n\text{-Matrix}}) = \mu(\underbrace{v_1, v_2, \dots, v_n}_{m \times n\text{-Matrix}}) + \mu(\underbrace{v_1', v_2, \dots, v_n}_{m \times n\text{-Matrix}})$$

für alle Spaltenvektoren $v_1', v_1, \ldots, v_n \in R^m$, und falls

$$\mu(av_1, v_2, \ldots, v_n) = a\mu(v_1, v_2, \ldots, v_n)$$

für alle $a \in R$ und $v_1, \ldots, v_n \in R^m$. Analog ist „linear in der zweiten Spalte“ usw. definiert.

Satz und Definition 3.2.3 *Es gibt genau eine Abbildung* $\det : M(n, R) \to R$ *mit folgenden Eigenschaften:*

i) $\det$ *ist linear in jeder Spalte.*

ii) Bei Vertauschung zweier Spalten ändert $\det$ *das Vorzeichen.*

iii) $\det(\mathbf{1}_n) = 1$ *(*$\mathbf{1}_n$ *ist die* $n \times n$*-Einheitsmatrix).*

$\det$ *heißt die* **Determinantenfunktion** *vom Grade* n *oder kurz die* **Determinante***.*

Der **Beweis** erfolgt im nächsten Abschnitt; siehe Feststellung 3.3.4 auf Seite 74. Ein weiterer Beweis wird sich in Abschnitt 6.3 ergeben, vgl. Bemerkung 6.3.11.

Satz 3.2.4 (Leibnizsche Determinantenformel) *Für* $A = (a_{ij}) \in M(n, R)$ *gilt*

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)}$$

(summiert wird über alle Elemente der symmetrischen Gruppe, siehe Abschnitt 3.1).

Dies ist mit in Feststellung 3.3.4 enthalten.

Beispiel: Für 2×2-Matrizen gilt: $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$. Man prüfe die Eigenschaften i), ii), iii) aus Definition 3.2.3 explizit nach.

Bemerkung 3.2.5 Die Bedeutung der Leibnizschen Determinantenformel liegt nicht in der praktischen Berechnung von Determinanten, wofür es wesentlich effektivere Methoden gibt (wir werden gleich eine kennenlernen). Vielmehr hat sie theoretische Bedeutung. Ist z.B. $R = \mathbb{R}$ oder $R = \mathbb{C}$, so hat $M(n, R)$ eine natürliche *Topologie*, und an der Leibnizformel liest man ab, daß $\det : M(n, R) \to R$ eine *stetige* Abbildung ist. Eine weitere Anwendung ist die folgende wichtige Eigenschaft der Determinante:

Folgerung 3.2.6 *Für jede Matrix* $A \in M(n, R)$ *gilt*

$$\det(A) = \det({}^t\!A).$$

Beweis: Mit $A = (a_{ij})$ gilt

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)}$$

gemäß Satz 3.2.4. Nun gilt offenbar

$$a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} = a_{\sigma^{-1}(1),1} \cdot \ldots \cdot a_{\sigma^{-1}(n),n},$$

und damit

$$\begin{aligned} \det(A) &= \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1),1} \cdot \ldots \cdot a_{\sigma^{-1}(n),n} \\ &= \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n} = \det({}^tA), \end{aligned}$$

denn ${}^tA = (a_{ji})_{i,j}$. ∎

Folgerung 3.2.7 *Die Determinante ist nicht nur linear in jeder Spalte, sondern auch in jeder Zeile. Ebenso wie für die Spalten gilt: Bei Vertauschung zweier Zeilen ändert sich das Vorzeichen.*

Satz 3.2.8 *Für $A, B \in M(n,R)$ gilt* $\det(AB) = \det(A)\det(B)$.

Der **Beweis** findet sich auf Seite 75.

Genau wie im Körperfall bezeichne $\mathrm{GL}(n,R)$ die Einheitengruppe des Ringes $M(n,R)$. Wie aus 3.2.8 und 3.2.12 weiter unten hervorgeht, induziert det einen Gruppenhomomorphismus

$$\det: \ \mathrm{GL}(n,R) \longrightarrow \mathrm{GL}(1,R) = R^*. \tag{3.1}$$

Der Kern dieses Homomorphismus bekommt einen Namen:

Definition 3.2.9 *Der Kern des Homomorphismus (3.1) heißt die* **spezielle lineare Gruppe** $\mathrm{SL}(n,R)$. *Also:*

$$\mathrm{SL}(n,R) = \{A \in \mathrm{GL}(n,R) : \ \det(A) = 1\}.$$

Es ist leicht zu sehen, daß der Homomorphismus (3.1) surjektiv ist. Deshalb haben wir eine kurze exakte Sequenz von Gruppen

$$1 \longrightarrow \mathrm{SL}(n,R) \longrightarrow \mathrm{GL}(n,R) \longrightarrow \mathrm{GL}(1,R) \longrightarrow 1. \tag{3.2}$$

Satz 3.2.10 (Kästchenformel) *Seien $n_1, \ldots, n_k$ natürliche Zahlen und $n = n_1 + \ldots + n_k$. Seien $A_i \in M(n_i, R)$, und A eine $n \times n$-Matrix der Form*

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

wobei oberhalb der Diagonalkästchen irgendwelche Koeffizienten stehen, und unterhalb Null. Dann gilt

$$\det(A) = \det(A_1) \cdot \ldots \cdot \det(A_k).$$

Insbesondere gilt

$$\det \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} = a_{11} \cdot \ldots \cdot a_{nn}$$

für obere Dreiecksmatrizen.

Dies wird auf Seite 75 bewiesen.

Satz 3.2.11 (Laplacescher Entwicklungssatz) *Sei $A = (a_{ij}) \in M(n, R)$, und $A_{ij} \in M(n-1, R)$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.*

i) (Entwicklung nach einer Zeile) Für jedes $i \in \{1, \ldots, n\}$ gilt

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

ii) (Entwicklung nach einer Spalte) Für jedes $j \in \{1, \ldots, n\}$ gilt

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

Zum **Beweis** siehe Seite 76.

Feststellung 3.2.12 *Für eine Matrix $A \in M(n, R)$ gilt:*

$$A \text{ ist eine Einheit im Ring } M(n, R) \quad \Longleftrightarrow \quad \det(A) \text{ ist eine Einheit in } R.$$

Insbesondere gilt im Falle eines Körpers $R = K$:

$$A \in M(n, K) \text{ ist invertierbar} \quad \Longleftrightarrow \quad \det(A) \neq 0.$$

Wir geben den **Beweis** im übernächsten Abschnitt, siehe Seite 77.

Bemerkung 3.2.13 Sei $R = K$ ein Körper. Zur praktischen Berechnung der Determinante einer vorgelegten Matrix A kann man dann systematisch so vorgehen: Zunächst wissen wir, wie sich die Determinante unter elementaren Zeilen- und Spaltentransformationen (siehe Seite 51) verhält. Bei Typ I nämlich bekommt man ein Minuszeichen, Typ II ändert nichts, bei Typ III bekommt man einen skalaren Faktor (dies ist alles leicht zu sehen). Durch forgesetzte elementare Transformationen bringt man die Matrix nun auf eine Form, an der man die Determinante leicht ablesen kann, z.B. obere Dreiecksform. Hat man über alle auftretenden Vorzeichen und skalaren Faktoren Buch geführt, erhält man so die Determinante der ursprünglichen Matrix.

3.3 Alternierende Multilinearformen

Definition 3.3.1 *Sei V ein R-Modul, und $k \in \mathbb{N}$. Eine* **Multilinearform** *vom Grade $k \in \mathbb{N}$, oder kurz eine k-***Form***, auf V ist eine Abbildung*

$$\mu : \underbrace{V \times \ldots \times V}_{k\text{-}mal} \longrightarrow R,$$

die in jeder Komponente linear ist, d.h.,

$$\mu(v_1, \ldots, v_i + v_i', \ldots, v_k) = \mu(v_1, \ldots, v_i, \ldots, v_k) + \mu(v_1, \ldots, v_i', \ldots, v_k)$$

für alle $v_1, \ldots, v_k, v_i' \in V$, und

$$\mu(v_1, \ldots, av_i, \ldots, v_k) = a\mu(v_1, \ldots, v_i, \ldots, v_k) \qquad \text{für alle } v_1, \ldots, v_k \in V,\ a \in R.$$

Bemerkungen: 1. Üblicherweise wird zusätzlich definiert: Eine 0-Form auf V ist einfach ein Element von R.

2. Eine 1-Form ist eine R-lineare Abbildung $V \to R$, im Körperfall also eine Linearform auf V, siehe Definition 2.5.1.

3. Eine 2-Form auf V heißt auch eine **Bilinearform**. Im Körperfall widmen wir diesen später ein eigenes Kapitel.

Feststellung und Definition 3.3.2 *Für eine k-Form $\mu : V^k \to R$ sind die folgenden Bedingungen äquivalent:*

i) $\mu(v_1, \ldots, v_k) = 0$, *falls einer der Vektoren eine Linearkombination der anderen ist.*

ii) $\mu(v_1, \ldots, v_k) = 0$, *falls es $i \neq j$ gibt mit $v_i = v_j$.*

In diesem Falle heißt μ **alternierend***, und es gilt zusätzlich:*

iii) $\mu(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\mu(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)$, *d.h., bei Vertauschung zweier Vektoren kehrt sich das Vorzeichen um.*

iv) $\mu(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \operatorname{sgn}(\sigma)\mu(v_1, \ldots, v_k)$ *für jede Permutation $\sigma \in \mathcal{S}_k$.*

Hat R die Eigenschaft, daß $a = -a$ nur für $a = 0$ gilt, so folgen i) und ii) auch aus iii) oder iv), d.h., alle vier Bedingungen sind äquivalent. Die Menge aller alternierenden k-Formen auf V ist in natürlicher Weise ein R-Modul, und wird mit

$$\mathrm{Alt}^k V$$

bezeichnet.

Beweis: i) $\Rightarrow$ ii) ist trivial. ii) $\Rightarrow$ i) folgt sofort mit Hilfe der Multilinearität.

ii) $\Rightarrow$ iii): Nach Multilinearität und ii) ist

$$\mu(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_k)$$

$$= \mu(v_1, \dots, v_i, \dots, v_j, \dots, v_k) + \mu(v_1, \dots, v_j, \dots, v_i, \dots, v_k).$$

Wiederum wegen ii) ist aber die linke Seite gleich Null.

iv) ist offenbar allgemeiner als iii), und iii) $\Rightarrow$ iv) folgt leicht aus 3.1.4 ii).

Falls in R gilt: Aus $a = -a$ folgt $a = 0$, so ergibt sich ii) sofort aus iii), und alle vier Bedingungen sind äquivalent.

Es ist klar, daß sich alternierende k-Formen in natürlicher Weise addieren und mit Elementen aus R multiplizieren lassen, und die Moduleigenschaften sind leicht nachzuprüfen. ■

Bemerkung 3.3.3 Im Falle $n = 1$ ist das Alternieren keine Bedingung. Für einen Körper $R = K$ ist $\mathrm{Alt}^1 V = V^*$ der Dualraum von V.

Wir betrachten nun n-Formen auf dem R^n. Indem man die Spalten einer Matrix $A \in M(n, R)$ als Elemente des R^n auffaßt, kann man eine solche n-Form als Funktion auf $M(n, R)$ interpretieren, die linear in jeder Spalte ist.

Obwohl R-Moduln, wie weiter oben erwähnt, im allgemeinen keine R-Basis besitzen, ist dies doch für die Standardmoduln R^n der Fall. Genau wie im Körperfall haben wir hier die *kanonischen Basisvektoren* $e_1, \dots, e_n$, und jedes Element $v \in R^n$ läßt sich in eindeutiger Weise als Linearkombination

$$v = \sum_{i=1}^{n} a_i e_i \qquad \text{mit } a_i \in R$$

schreiben.

Feststellung 3.3.4 *Es gibt genau eine alternierende n-Form, genannt* det, *auf dem R^n, so daß*

$$\det(e_1, \dots, e_n) = 1.$$

Explizit gilt

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma_n} \qquad \text{für } A = (a_{ij}) \in M(n, R). \tag{3.3}$$

Für eine weitere alternierende n-Form μ auf dem R^n gilt

$$\mu(v_1, \dots, v_n) = \det(A)\mu(e_1, \dots, e_n),$$

wobei $v_i = \sum_{i=1}^n a_{ij} e_i$ und $A = (a_{ij}) \in M(n, R)$.

Beweis: Sei μ eine alternierende n-Form. Aufgrund der Multilinearität gilt mit $v_j = \sum_{i=1}^n a_{ij} e_i$

$$\begin{aligned}
\mu(v_1, \dots, v_n) &= \mu\Big(\sum_{i_1=1}^{n} a_{i_1 1} e_{i_1}, \dots, \sum_{i_n=1}^{n} a_{i_n n} e_{i_n} \Big) \\
&= \sum_{i_1, \dots, i_n = 1}^{n} a_{i_1 1} \cdot \ldots \cdot a_{i_n n}\, \mu(e_{i_1}, \dots, e_{i_n}).
\end{aligned}$$

Befinden sich unter den $i_1, \ldots, i_n$ zwei gleiche Indizes, so ist $\mu(e_{i_1}, \ldots, e_{i_n}) = 0$ aufgrund des Alternierens. In der Summe kann man sich deshalb auf n-Tupel $i_1, \ldots, i_n$ beschränken, in denen die i_k paarweise verschieden sind. Dies läßt sich auch so schreiben:

$$\begin{aligned}\mu(v_1, \ldots, v_n) &= \sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}\, \mu(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}\, \mu(e_1, \ldots, e_n),\end{aligned}$$

wobei wir 3.3.2 iv) benutzt haben. Statt σ kann man auch über alle $\sigma^{-1} \in \mathcal{S}_n$ summieren, und beachtet man noch $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$ und $a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} = a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$, so folgt

$$\mu(v_1, \ldots, v_n) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}\, \mu(e_1, \ldots, e_n).$$

Fordert man noch $\mu(e_1, \ldots, e_n) = 1$, so ergibt sich notwendig die Formel (3.3) für det. Dies zeigt die Eindeutigkeit von det. Zur Existenz: Wir *definieren* die Funktion det durch (3.3), und müssen nur noch zeigen, daß es sich um eine alternierende n-Form handelt. Die Multilinearität ist sehr einfach zu sehen; wir zeigen nur das Alternieren. Sei $v_i = v_j$. Sei $\tau = (ij) \in \mathcal{S}_n$ die Permutation, die i und j vertauscht. In der Summe über $\sigma \in \mathcal{S}_n$ fassen wir nun jeweils die beiden Permutationen σ und $\tau\sigma$ zusammen. Wegen $v_i = v_j$ ist $a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} = a_{1,\tau\sigma(1)} \cdot \ldots \cdot a_{1,\tau\sigma(n)}$. Aber $\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\tau\sigma)$, und somit heben sich diese beiden Terme weg. ∎

Bemerkung 3.3.5 Sei $R = K$ ein Körper. Aus der letzten Feststellung folgt $\dim_K(\mathrm{Alt}^n(K^n)) = 1$. Allgemeiner läßt sich zeigen (und wir werden dies mit Satz 6.3.10 tun): Ist V ein n-dimensionaler K-Vektorraum, so gilt $\dim_K(\mathrm{Alt}^k(V)) = \binom{n}{k}$, mit dem **Binomialkoeffizienten**

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k}.$$

(Dieser ist immer eine ganze Zahl, und hat die folgende kombinatorische Bedeutung: $\binom{n}{k}$ ist die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge.) Insbesondere ist $\dim_K(\mathrm{Alt}^k(V)) = 0$ für $k > n$, was klar ist, da mehr als n Vektoren von V immer linear abhängig sind.

Beweis von Satz 3.2.8: Bei festem A sei $\mu : M(n, R) \to R$ die Abbildung $\mu(B) := \det(AB)$. Es ist leicht zu sehen, daß μ eine alternierende n-Form ist. Aus 3.3.4 folgt deshalb

$$\det(AB) = \mu(B) = \det(B)\mu(e_1, \ldots, e_n).$$

Aber für $B = \mathbf{1}$ folgt $\mu(e_1, \ldots, e_n) = \det(A\mathbf{1}) = \det(A)$. ∎

Beweis von Satz 3.2.10: Mit vollständiger Induktion reduziert man das Problem leicht auf den Fall von *zwei* Kästchen. Sei also $n = n_1 + n_2$, $A_1 \in M(n_1, R)$, $A_2 \in M(n_2, R)$, und

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

Das Produkt $a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)}$ in der Leibnizformel ist immer dann Null, wenn es ein $i > n_1$ gibt mit $\sigma(i) \le n_1$. Die Summation über $\sigma \in \mathcal{S}_n$ kann man also auf die Teilmenge

$$\mathcal{T} = \{\sigma \in \mathcal{S}_n : \sigma(\{1,\ldots,n_1\}) = \{1,\ldots,n_1\},\ \sigma(\{n_1+1,\ldots,n\}) = \{n_1+1,\ldots,n\}\}$$

beschränken. $\mathcal{T}$ besteht also aus den $\sigma \in \mathcal{S}_n$, die eine Permutation der ersten n_1 und der letzten n_2 Zahlen bewirken. In der Tat ist $\mathcal{T}$ eine Untergruppe von $\mathcal{S}_n$, und dem Leser wird es leichtfallen, einen Isomorphismus von $\mathcal{T}$ mit $\mathcal{S}_{n_1} \times \mathcal{S}_{n_2}$ hinzuschreiben. Deshalb läßt sich die Leibnizformel schreiben als

$$\begin{aligned}
\det(A) &= \sum_{\sigma_1 \in \mathcal{S}_{n_1}} \sum_{\sigma_2 \in \mathcal{S}_{n_2}} \operatorname{sgn}(\sigma_1)\operatorname{sgn}(\sigma_2) a_{1,\sigma_1(1)} \cdot \ldots \cdot a_{n_1,\sigma_1(n_1)} \cdot a_{n_1+1,\sigma_2(1)+n_1} \cdot \ldots \cdot a_{n,\sigma_2(n_2)+n_1} \\
&= \Bigg(\sum_{\sigma \in \mathcal{S}_{n_1}} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} \cdot \ldots \cdot a_{n_1,\sigma(n_1)} \Bigg) \Bigg(\sum_{\sigma \in \mathcal{S}_{n_2}} \operatorname{sgn}(\sigma) a_{n_1+1,\sigma(1)+n_1} \cdot \ldots \cdot a_{n,\sigma(n_2)+n_1} \Bigg) \\
&= \det(A_1)\det(A_2).
\end{aligned}$$

■

3.4 Laplaceentwicklung und Cramersche Regel

Wir geben zunächst den **Beweis** des Laplaceschen Entwicklungssatzes 3.2.11. Wegen Folgerung 3.2.6 genügt es, die Aussage i) zu beweisen. Man überlegt sich leicht, daß der Ausdruck

$$(-1)^{i+j} a_{ij} \det(A_{ij})$$

als Funktion von $A = (a_{ij})$ die drei Eigenschaften aus Satz 3.2.3 erfüllt. Aufgrund der Eindeutigkeitsaussage in 3.2.3 muß daher

$$\det(A) = (-1)^{i+j} a_{ij} \det(A_{ij}) \qquad \text{für alle } A \in M(n,K)$$

gelten. ■

Lemma 3.4.1 *Sei $A \in M(n,R)$, und (wie im Laplaceschen Entwicklungssatz) A_{ij} die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht. Dann gilt*

$$\sum_{k=1}^{n} a_{ik}(-1)^{j+k} \det(A_{jk}) = \det(A) \cdot \delta_{ij} \qquad \textit{für alle } i,j \in \{1,\ldots,n\}.$$

Beweis: Für i=j ist dies gerade die Entwicklung von $\det(A)$ nach der i-ten Zeile. Sei $j \ne i$, und A' die Matrix, die entsteht, wenn man die j-te Zeile von A durch die i-te Zeile ersetzt. Dann besitzt A' zwei gleiche Zeilen, also ist $\det(A') = 0$. Laplace-Entwicklung von $\det(A')$ nach der j-ten Zeile liefert gerade die gewünschte Summe. ■

Zu $A \in M(n,R)$ definiert man die **komplementäre Matrix** $\tilde{A} = (\tilde{a}_{kj}) \in M(n,R)$ durch

$$\tilde{a}_{kj} = (-1)^{j+k} \det(A_{jk}).$$

Das soeben formulierte Lemma läßt sich dann so interpretieren:

Feststellung 3.4.2 (Cramersche Regel über die komplementäre Matrix) [2] *Ist $\tilde{A}$ die zu $A \in$*

[2] Gabriel Cramer, 1704-1752

$M(n,R)$ komplementäre Matrix, so gilt

$$A\tilde{A} = \tilde{A}A = \det(A)\mathbf{1}_n. \tag{3.4}$$

Beweis: Bedenkt man, wie Matrizen miteinander multipliziert werden, so ist $A\tilde{A} = \det(A)\mathbf{1}$ genau die Aussage von Lemma 3.4.1. Die andere Aussage bekommt man analog, wenn man mit Spalten anstatt Zeilen arbeitet. ∎

Beweis von Feststellung 3.2.12: Sei $A \in M(n,R)$ eine Einheit. Dann gibt es ein $B \in M(n,R)$ mit $AB = \mathbf{1}_n$. Anwenden der Determinante ergibt $\det(A)\det(B) = \det(AB) = \det(\mathbf{1}) = 1$, und folglich ist $\det(A)$ eine Einheit in R.

Ist umgekehrt $\det(A)$ als Einheit in R vorausgesetzt, so ergibt sich aus (3.4), daß $\det(A)^{-1}\tilde{A} \in M(n,R)$ ein Inverses zu A in $M(n,R)$ ist. ∎

Die in diesem Beweis gemachte Beobachtung halten wir gesondert fest:

Satz 3.4.3 (Cramersche Regel für die inverse Matrix) *Ist $A \in M(n,R)$ invertierbar, und $A^{-1} = (a'_{ij})$ die inverse Matrix, so gilt*

$$a'_{ij} = \det(A)^{-1}(-1)^{i+j}\det(A_{ji}).$$

Bemerkung 3.4.4 Ist $K = \mathbb{R}$ oder $K = \mathbb{C}$, so haben wir schon in Bemerkung 3.2.5 festgestellt, daß die Determinante *stetig* von den Koeffizienten abhängt. Aus der Cramerschen Regel folgt dann, daß die Koeffizienten von A^{-1} stetig von den Koeffizienten von A abhängen. Mit anderen Worten: Die Matrixinversion

$$\begin{aligned} \mathrm{GL}(n,K) &\longrightarrow \mathrm{GL}(n,K), \\ A &\longmapsto A^{-1}, \end{aligned}$$

ist eine stetige Abbildung.

3.5 Determinante und Spur von Endomorphismen

Ab jetzt sei wieder K ein Körper.

Determinante eines Endomorphismus

Wir erinnern an Definition 2.3.14: Zwei Matrizen $A, A' \in M(n,K)$ heißen *konjugiert*, falls es ein $P \in \mathrm{GL}(n,K)$ gibt mit $A' = PAP^{-1}$. Aus dem Multiplikationssatz 3.2.8 folgt sofort:

> Konjugierte Matrizen haben die gleiche Determinante.

Sei jetzt $f : V \to V$ ein Endomorphismus eines endlich-dimensionalen K-Vektorraumes V. Sei A die Matrix von f bezüglich einer beliebigen Basis von V. Aufgrund von Folgerung 2.3.13 und der eben gemachten Bemerkung ist $\det(A)$ unabhängig von der gewählten Basis. Deshalb macht es Sinn, die **Determinante von f** als die Determinante von A zu definieren. In Zeichen:

$$\det(f) := \det(A).$$

Feststellung 3.5.1 *Sei* $\dim_K(V) = n$ *und* $f, g \in \operatorname{End}_K(V)$. *Es gilt:*

i) $\det(fg) = \det(gf) = \det(f)\det(g)$.

ii) $\det(\mathrm{id}_V) = 1$.

iii) $\det(f) \neq 0 \iff f$ *ist ein Isomorphismus.*

iv) Für jede alternierende n*-Form* $\mu : V^n \to K$ *gilt*

$$\mu(fv_1, \dots, fv_n) = \det(f)\mu(v_1, \dots, v_n) \qquad \text{für alle } v_1, \dots, v_n \in V.$$

Beweis: Ist A die Matrix von f und B die Matrix von g, so ist AB die Matrix von fg. Daraus ergibt sich i). Die Matrix von id_V ist die Einheitsmatrix, woraus ii) folgt. Zu iii) beachte man: f ist ein Isomorphismus genau dann, wenn A ein Isomorphismus ist, genau dann, wenn $\det(A) \neq 0$ ist.

Wir zeigen iv). Wir dürfen annehmen, daß $\mu \neq 0$ ist. Es ist leicht zu sehen, daß die Abbildung

$$\begin{aligned} &V \times \ldots \times V \longrightarrow K, \\ &(v_1, \dots, v_n) \longmapsto \mu(fv_1, \dots, fv_n), \end{aligned}$$

eine alternierende n-Form auf V ist, ebenso wie μ. Der Raum $\operatorname{Alt}^n(V)$ ist aber eindimensional nach Feststellung 3.3.4. Somit muß es eine Konstante $c \in K$ geben, mit der gilt

$$\mu(fv_1, \dots, fv_n) = c\,\mu(v_1, \dots, v_n) \qquad \text{für alle } v_1, \dots, v_n \in V,$$

und die Behauptung ist, daß diese Konstante gerade $\det(f)$ ist. Zur Berechnung von c genügt es, irgendeine Basis $w_1, \dots, w_n$ einzusetzen, denn dann ist $\mu(w_1, \dots, w_n) \neq 0$ (vgl. den Beweis von 3.3.4). Wenn A die Matrix von f bezüglich dieser Basis ist, so gilt

$$f(w_j) = \sum_{i=1}^{n} a_{ij} w_i$$

nach Definition von $A = (a_{ij})$. Somit gilt

$$\begin{aligned} \mu(fw_1, \dots, fw_n) &= \mu\Big(\sum_{i_1=1}^{n} a_{i_1 1} w_{i_1}, \dots, \sum_{i_n=1}^{n} a_{i_n n} w_{i_n}\Big) \\ &= \sum_{i_1, \dots, i_n = 1}^{n} a_{i_1 1} \cdot \ldots \cdot a_{i_n n}\, \mu(w_{i_1}, \dots, w_{i_n}) \\ &= \sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}\, \mu(w_{\sigma(1)}, \dots, w_{\sigma(n)}) \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}\, \mu(w_1, \dots, w_n) \\ &= \det(A)\mu(w_1, \dots, w_n) = \det(f)\mu(w_1, \dots, w_n), \end{aligned}$$

wobei wir die Leibnizregel verwendet haben. Nun folgt $c = \det(f)$ wegen $\mu(w_1, \dots, w_n) \neq 0$. ∎

Bemerkung 3.5.2 Eigenschaft iv) kann auch zur *Definition* von $\det(f)$ verwendet werden. Diese Definition ist *intrinsisch*, d.h., sie nimmt keinen Rückgriff auf einen Koordinatenisomorphismus (nach Wahl einer Basis) und die Determinante von Matrizen.

Spur von Matrizen und Endomorphismen

Die **Spur**[3] einer quadratischen Matrix $A = (a_{ij}) \in M(n, K)$ ist die Summe ihrer Diagonalkoeffizienten:

$$\operatorname{tr}(A) := \sum_{i=1}^{n} a_{ii}.$$

Feststellung 3.5.3 *Die Spur hat die folgenden grundlegenden Eigenschaften:*

i) Die Spur ist eine lineare Abbildung $M(n, K) \to K$, d.h., tr *ist eine Linearform auf $M(n, K)$.*

ii) Für $A, B \in M(n, K)$ gilt $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

iii) Konjugierte Matrizen haben die gleiche Spur.

Beweis: i) ist unmittelbar klar, ii) ergibt sich nach leichter Rechnung, und iii) folgt aus ii). ∎

Sei f ein Endomorphismus eines endlich-dimensionalen K-Vektorraumes V. Sei A die Matrix von f bezüglich irgendeiner Basis von V. Wir definieren die **Spur von** f als die Spur von A:

$$\operatorname{tr}(f) := \operatorname{tr}(A).$$

Aufgrund von iii) in 3.5.3 ist diese Definition unabhängig von der gewählten Basis.

Feststellung 3.5.4 *Sei V ein endlich-dimensionaler K-Vektorraum.*

i) $\operatorname{tr} : \operatorname{End}_K(V) \to K$ *ist eine lineare Abbildung.*

ii) Für $f, g \in \operatorname{End}_K(V)$ gilt $\operatorname{tr}(fg) = \operatorname{tr}(gf)$.

Beweis: Dies folgt sofort aus den entsprechenden Eigenschaften für die Spur von Matrizen. ∎

[3] engl. *trace*

Kapitel 4

Endomorphismen und Polynome

Dieses Kapitel setzt das Studium der Endomorphismen eines endlich-dimensionalen Vektorraumes fort. Es stellt sich heraus, daß es fundamentale Zusammenhänge mit *Polynomen* über dem Grundkörper gibt. Z.B. ist jedem Endomorphismus ein *charakteristisches Polynom* zugeordnet, dessen Nullstellen die *Eigenwerte* des Endomorphismus sind – ein grundlegender Begriff. Gegen Ende des Kapitels werden wir in der Lage sein, die Endomorphismen eines vorgelegten endlich-dimensionalen Vektorraumes weitgehend zu *klassifizieren*. Dabei stützen wir uns auf einige Eigenschaften von Polynomringen, die daher zunächst studiert werden sollen. Ein Aufwand, der sich lohnt: Polynome spielen in vielen Bereichen der Mathematik (Algebra, Analysis, Geometrie, Zahlentheorie, ...) unverzichtbare Rollen.

4.1 Polynome

Grundbegriffe

In diesem Abschnitt sei immer

R ein kommutativer Ring mit Eins;

insbesondere kann $R = K$ ein Körper sein, was später für uns hauptsächlich der Fall sein wird. Wir verzichten auf die formale Definition einer **R-Algebra**, da diese völlig analog zu der einer K-Algebra ist (siehe Definition 1.3.14): Es handelt sich um einen R-Modul (Definition 3.2.1), in dem man zusätzlich multiplizieren kann. Beispiel für eine R-Algebra ist $M(n, R)$ (mit Matrixmultiplikation).

Ein **Polynom**[1] mit Koeffizienten in R ist ein Ausdruck der Form

$$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n, \qquad a_i \in R,$$

wobei X eine „Unbekannte“ ist. Polynome kann man addieren und multiplizieren (letzteres ganz natürlich unter Benutzung von $X^iX^j = X^{i+j}$), sie bilden daher einen Ring, den **Polynomring** über R in der Unbestimmten X. Wir wollen etwas formaler sein:

[1] gr. *polys*: viel; lat. *nomen*: Name

Definition 4.1.1 *Sei $\mathcal{B}$ eine R-Algebra, in der es ein Element X gibt, so daß sich jedes Element von $\mathcal{B}$ eindeutig in der Form*

$$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n, \qquad \textit{mit } a_i \in R \tag{4.1}$$

schreiben läßt. Dann heißt $\mathcal{B}$ der **Polynomring**[2] *in der Unbestimmten X über R, und wird mit dem Symbol*

$$\mathcal{B} = R[X]$$

bezeichnet. Die Elemente von $R[X]$ heißen **Polynome***, und die Elemente $a_i \in R$ wie oben die* **Koeffizienten** *des Polynoms.*

Ist $R = K$ ein Körper, so ist $K[X]$ insbesondere ein K-Vektorraum; als solcher besitzt er definitionsgemäß die Basis

$$1,\ X,\ X^2,\ X^3, \ldots ,$$

ist also unendlich-dimensional. Polynomringe sind immer *kommutativ*, was einfach an $X^iX^j = X^{i+j} = X^jX^i$ liegt. Ist $P \in R[X]$ in der Form (4.1) dargestellt mit $c_n \neq 0$, so heißt n der **Grad**[3] von P, und wir schreiben

$$n = \deg(P).$$

Das Nullpolynom $P = 0$ bekommt den Grad $-\infty$. Ist $n = \deg(P)$ und $c_n = 1$, so heißt das Polynom P **normiert**. Falls R *nullteilerfrei* ist, so gilt für $P, Q \in R[X]$ offenbar

$$\deg(PQ) = \deg(P)\deg(Q)$$

(im allgemeinen gilt nur „$\leq$“). Daraus folgt leicht:

$R[X]$ ist nullteilerfrei, falls R es ist.

Polynome vom Grad 0 sind einfach Elemente von R; insbesondere ist R ein *Unterring* von $R[X]$ (weshalb wir vom Polynomring *über* R sprechen). Man überlegt sich im nullteilerfreien Fall:

R^* ist die Einheitengruppe von $R[X]$.

Bemerkung 4.1.2 Es ist leicht zu sehen, daß der Polynomring $R[X]$ in der Tat *existiert*. Im wesentlichen *definiert* man ihn als die Menge aller formalen Ausdrücke der Form (4.1), mit der gewünschten Multiplikation. Ferner ist leicht zu sehen, daß zwei Polynomringe $R[X]$ und $R[Y]$ immer *isomorph* sind. Es gibt nämlich einen (eindeutig bestimmten) R-Algebren-Isomorphismus $R[X] \to R[Y]$, der X auf Y abbildet (siehe die gleich folgende Feststellung 4.1.3). Kurz gesagt: Polynomringe existieren und sind eindeutig. Wir sprechen daher von *dem* Polynomring über R in einer Unbestimmten.

[2] weniger gebräuchlich, obwohl genauer, ist *Polynomalgebra*
[3] engl. *degree*

Prominente Beispiele für Polynomringe sind $\mathbb{Z}[X]$, $\mathbb{Q}[X]$, $\mathbb{R}[X]$, $\mathbb{C}[X]$. Man beachte, daß die Elemente dieser Ringe für uns lediglich formale Ausdrücke der Form (4.1) sind, und nicht etwa Funktionen auf $\mathbb{C}$ (indem man für X komplexe Zahlen einsetzt). Man könnte jedoch sagen, daß jedes solche Polynom P eine Funktion f_P auf $\mathbb{C}$ *definiert*. In den obigen Fällen ist die Abbildung $P \mapsto f_P$ *injektiv*; dies ist der Grund, warum Polynome häufig mit den durch sie definierten Funktionen identifiziert werden.

Man kann auch den Polynomring über einem Polynomring bilden; es ist dann $R[X][Y] = R[X, Y]$ der Polynomring in *zwei* Unbekannten über R, und seine Elemente sind eben alle polynomialen Ausdrücke in den Unbekannten X und Y mit Koeffizienten in R. Noch allgemeiner hat man den Polynomring $R[X_1, \dots, X_n]$ in n Unbekannten über R.

Universelle Eigenschaft

Feststellung 4.1.3 *Der Polynomring $R[X]$ besitzt folgende „universelle Eigenschaft“: Ist $\mathcal{B}$ eine weitere R-Algebra, und $B \in \mathcal{B}$ ein beliebiges Element, so gibt es genau einen R-Algebren-Homomorphismus $R[X] \to \mathcal{B}$, der X auf B abbildet.*

Beweis: Sei $\varphi : R[X] \to \mathcal{B}$ ein Homomorphismus von R-Algebren mit $\varphi(X) = B$. Für ein beliebiges Polynom $P = \sum_{i=1}^n a_i X^i$ gilt dann

$$\varphi(P) = \sum_{i=1}^n \varphi(a_i X^i) = \sum_{i=1}^n a_i \varphi(X^i) = \sum_{i=1}^n a_i \varphi(X)^i = \sum_{i=1}^n a_i B^i.$$

Also kann es höchstens *ein* φ mit den gewünschten Eigenschaften geben. Umgekehrt können wir eine Abbildung $\varphi : R[X] \to \mathcal{B}$ durch

$$\varphi\Big(\sum_{i=1}^n a_i X^i\Big) := \sum_{i=1}^n a_i B^i$$

definieren. Da die Darstellung $\sum a_i X^i$ eindeutig ist, ist φ jedenfalls eine wohldefinierte Abbildung. Dem Leser sei der Nachweis überlassen, daß φ tatsächlich ein R-Algebren-Homomorphismus ist. ∎

Wie im Beweis dieser Feststellung gesehen, ist der eindeutig bestimmte R-Algebren-Homomorphismus $R[X] \to \mathcal{B}$ mit $X \mapsto B$ dadurch gegeben, daß in einem Polynom $\sum a_i X^i$ das X durch B ersetzt wird. Er heißt deshalb auch der zu $B \in \mathcal{B}$ gehörige **Einsetzungshomomorphismus**.

Das Polynom $P = \sum a_i X^i \in R[X]$ wird oft auch mit dem Symbol $P(X)$ bezeichnet, und das Bild von P unter dem obigen Einsetzungshomomorphismus dann mit $P(B)$. Dies ist besonders üblich im Fall $\mathcal{B} = R$.

Beispiel 4.1.4 Sei V ein endlich-dimensionaler K-Vektorraum, und $f \in \mathrm{End}_K(V)$. Wir betrachten den durch f definierten Einsetzungshomomorphismus

$$\begin{aligned} K[X] &\longrightarrow \mathrm{End}_K(V), \\ X &\longmapsto f. \end{aligned}$$

Dieser ist eine lineare Abbildung eines unendlich-dimensionalen K-Vektorraumes in einen endlich-dimensionalen, muß also einen nicht-trivialen Kern besitzen. Mit anderen Worten, es gibt ein Polynom $P = \sum a_i X^i$, das nicht Null ist, so daß $\sum a_i f^i$ Null in $\mathrm{End}_K(V)$ ist. Sei μ_f das normierte Polynom kleinsten Grades mit dieser Eigenschaft. Dann heißt μ_f das **Minimalpolynom** von f. Wir werden es später genauer untersuchen.

Ideale

Feststellung 4.1.5 (Division mit Rest) *Seien $f, g \in R[X]$, und der höchste Koeffizient von g sei eine Einheit in R. Dann gibt es eindeutig bestimmte $q, r \in R[X]$ mit*

$$f = qg + r \qquad \text{und} \qquad \deg(r) < \deg(g). \tag{4.2}$$

Beweis: Eine solche Darstellung ist eindeutig: Ist $f = qg + r = q'g + r'$ mit $\deg(r) < \deg(g)$ und $\deg(r') < \deg(g)$, so folgt $(q - q')g = r' - r$. Hier muß $q = q'$ sein, denn sonst hätte $(q - q')g$ einen größeren Grad als $r - r'$. Aus $q = q'$ folgt dann auch $r = r'$.

Zur Existenz kann man Induktion nach dem Grad von f führen. Für $\deg(f) < \deg(g)$ setze man $q = 0$ und $r = f$. Für $\deg(f) \geq \deg(g)$ betrachte man $h = f - agX^k$ mit $a \in R^*$ und $k \in \mathbb{N}_0$. Für geeignete Wahl von a und k gilt $\deg(h) < \deg(f)$. Nach Induktionsvoraussetzung gibt es deswegen eine Darstellung der Gestalt (4.2) für h. Daraus gewinnt man dann eine Darstellung für f. ∎

Definition 4.1.6 *Seien $f, g \in R[X]$. Gibt es ein $h \in R[X]$ mit $f = gh$, so heißt g ein* **Teiler** *von f, und f ein* **Vielfaches** *von g; man schreibt dann $g|f$.*

Genau dann ist also $g|f$, wenn in der Darstellung (4.2) $r = 0$ ist.

Definition 4.1.7 *Sei $\mathcal{B}$ ein kommutativer Ring. Eine Teilmenge $I \subset \mathcal{B}$ heißt ein* **Ideal***, falls gilt:*

i) $a, b \in I \implies a + b \in I$,

ii) $a \in R,\ b \in I \implies ab \in I$

(kurz gesagt, es muß $I + I \subset I$ und $RI \subset I$ gelten).

Man sieht leicht, daß ein Ideal $I \subset \mathcal{B}$ insbesondere eine Untergruppe der additiven Gruppe ist. Genau dann ist eine solche Untergruppe ein Ideal, wenn Multiplikation mit Ringelementen nicht aus der Untergruppe herausführt.

Jeder Kern eines Ringhomomorphismus von $\mathcal{B}$ in einen anderen Ring ist ein Ideal von $\mathcal{B}$. Umgekehrt kann man zu jedem Ideal I einen Ringhomomorphismus konstruieren, dessen Kern gerade I ist (nämlich die Projektion auf den *Quotientenring* B/I, dessen Definition der Leser vielleicht selbst versuchen mag). Also: *Ideale sind genau die Kerne von Ringhomomorphismen.*

Der ganze Ring $\mathcal{B}$ ist ein Ideal, ebenso wie das **Nullideal** $\{0\}$. Für jedes Element $b \in B$ ist

$$(b) := \{ab : a \in R\},$$

also die Menge aller Vielfachen von b, ein Ideal von $\mathcal{B}$. Es heißt das von b erzeugte **Hauptideal**. Genau dann ist $(b) = \mathcal{B}$, wenn b eine Einheit in $\mathcal{B}$ ist. Wenn $b_1, \ldots, b_n \in \mathcal{B}$ irgendwelche Elemente sind, so ist

$$(b_1, \ldots, b_n) := \{a_1 b_1 + \ldots + a_n b_n : \ a_1, \ldots, a_n \in \mathcal{B}\}$$

das von den b_i **erzeugte** Ideal. Hauptideale sind also diejenigen Ideale, die von einem einzigen Element erzeugt werden.

Definition 4.1.8 *Ein Ring heißt* **Hauptidealring**, *wenn jedes seiner Ideale ein Hauptideal ist.*

Feststellung 4.1.9 *Der Polynomring $K[X]$ über einem Körper K ist ein Hauptidealring.*

Beweis: Sei $I \subset K[X]$ ein Ideal. Unter allen Polynomen in I, die ungleich Null sind, sei g eines von minimalem Grad. Wir behaupten, daß $I = (g)$. Klar ist $I \supset (g)$ wegen Eigenschaft ii) in der Definition eines Ideals. Umgekehrt sei $f \in I$. Nach Feststellung 4.1.5 gibt es $q, r \in K[X]$ mit $\deg(r) < deg(f)$ und

$$f = qg + r.$$

Nun ist $r = f - qg \in I$; wegen $\deg(r) < \deg(g)$ und der Wahl von g muß daher $r = 0$ sein. Also ist $f = qg \in (g)$. ∎

Folgerung 4.1.10 *Zu jedem Ideal $I \neq (0)$ in $K[X]$ gibt es genau ein normiertes Polynom $f \in K[X]$ mit $I = (f)$.*

Beweis: Nach der vorherigen Feststellung gibt es zumindest ein f mit $I = (f)$. Falls auch $I = (g)$, so ist f ein Vielfaches von g, und g ein Vielfaches von f, und somit haben f und g den gleichen Grad, und unterscheiden sich nur um eine Einheit (ein Element von K^*). Also ist der Erzeuger f eindeutig bis auf Multiplikation mit Einheiten, und für eine geeignete Einheit erhält man ein normiertes f. ∎

Eindeutige Primfaktorzerlegung

Definition 4.1.11 *Sei R nullteilerfrei.*

i) Ein Polynom $f \in R[X]$ heißt **irreduzibel**, *falls f keine Einheit ist, und falls für alle $g, h \in R[X]$ gilt:*

$$f = gh \implies g \in R^* \text{ oder } h \in R^*.$$

ii) Ein Polynom $f \in R[X]$ heißt ein **Primelement**, *falls f keine Einheit ist, und falls für alle $g, h \in R[X]$ gilt:*

$$f | gh \implies f | g \text{ oder } f | h.$$

Feststellung 4.1.12 *Sei R nullteilerfrei.*

i) Jedes Primelement in $R[X]$ ist irreduzibel.

ii) Falls $R = K$ ein Körper ist, so ist jedes irreduzible Polynom auch ein Primelement. In diesem Fall gibt es also keine Unterschiede zwischen irreduziblen Elementen und Primelementen.

Beweis: i) Sei f ein Primelement und $f = gh$. Dann ist insbesondere $f|gh$, also $f|g$ oder $f|h$. Sei etwa $f|g$, d.h., $g = fq$ mit einem $q \in R[X]$. Es folgt $f = gh = fqh$, und, da $R[X]$ nullteilerfrei ist, $qh = 1$. Also ist h eine Einheit.

ii) Sei f irreduzibel und $f|gh$. Wenn I das von f und g erzeugte Ideal ist, so gibt es nach Feststellung 4.1.9 ein $p \in K[X]$ mit

$$I = (f, g) = (p).$$

Insbesondere ist f ein Vielfaches von p. Da f irreduzibel ist, können nur zwei Fälle auftreten:

<u>1. Fall</u>: $p = af$ mit einer Einheit $a \in K^*$. Dann ist g ein Vielfaches von f, also $f|g$, und wir sind fertig.

<u>2. Fall</u>: p ist eine Einheit. Dann ist $(p) = K[X]$, also gibt es $q_1, q_2 \in K[X]$ mit

$$q_1 f + q_2 g = 1.$$

Multiplikation mit h liefert $h = q_1 fh + q_2 gh$. Aber gh ist ein Vielfaches von f nach Voraussetzung, und damit auch h. Mit anderen Worten, $f|h$. ■

Satz 4.1.13 *Sei K ein Körper. Der Polynomring $K[X]$ ist faktoriell, d.h.: Jedes $f \in K[X]$, welches keine Einheit und nicht Null ist, besitzt eine Darstellung*

$$f = p_1 \cdot \ldots \cdot p_n \qquad \text{mit irreduziblen } p_i \in K[X],$$

wobei die Polynome p_i bis auf die Reihenfolge und bis auf Multiplikation mit Einheiten eindeutig bestimmt sind.

Beweis: Die Existenz einer solchen Darstellung folgt mit Induktion nach dem Grad von f: Ist f nicht schon selbst irreduzibel, so kann man $f = gh$ schreiben, wobei g und h kleineren Grad als f haben, und nach Induktionsvoraussetzung damit Produkte von irreduziblen Elementen sind.

Zur Eindeutigkeit: Seien

$$f = p_1 \cdot \ldots \cdot p_n = q_1 \cdot \ldots \cdot q_m$$

zwei Darstellungen mit irreduziblen p_i und q_j. Dann gilt $p_1 | q_1 \cdot \ldots \cdot q_m$, und eine wiederholte Anwendung von 4.1.12 ii) zeigt $p_1 | q_j$ für ein j. Da p_1 und q_j beides irreduzible Elemente sind, können sie sich nur um eine Einheit unterscheiden. Man kann also p_1 auf beiden Seiten kürzen, wobei die eventuell übrigbleibende Einheit zu einem der q_k geschlagen werden kann. Man hat dann zwei Darstellungen kleinerer Länge, und die Behauptung folgt etwa durch Induktion nach $\max(m, n)$. ■

Nullstellen

Definition 4.1.14 *Sei $f \in R[X]$. Ist $a \in R$ und $X-a$ ein Teiler von f, so heißt a eine* **Nullstelle** *von f, und $X-a$ ein* **Linearfaktor** *von f. In diesem Fall heißt das Maximum der Zahlen $k \in \mathbb{N}$ mit $(X-a)^k | f$ die* **Vielfachheit** *oder* **Ordnung** *der Nullstelle a in f, und man schreibt*

$$k =: \operatorname{ord}_a(f).$$

Ist a keine Nullstelle von f, so setzen wir $\operatorname{ord}_a(f) = 0$.

Feststellung 4.1.15 *i) $a \in R$ ist genau dann eine Nullstelle von $f = f(X)$, wenn $f(a) = 0$.*

ii) Sei R nullteilerfrei. Sind $a_1, \dots, a_m$ paarweise verschiedene Nullstellen von $f \in R[X]$, und $k_i = \operatorname{ord}_{a_i}(f)$, so gilt

$$f = (X-a_1)^{k_1} \cdot \ldots \cdot (X-a_m)^{k_m} g,$$

wobei $g \in R[X]$ ein Polynom ist mit $g(a_i) \neq 0$ für $i = 1, \dots, m$.

iii) Falls R nullteilerfrei ist, so besitzt ein Polynom vom Grade n in $R[X]$ höchstens n Nullstellen.

Beweis: i) Man wende 4.1.5 an mit $g = X - a$.

ii) Wir führen Induktion nach dem Grad von f. Für $\deg(f) = 1$, also $f = bX + c$, und $f(a) = 0$, kann man $f = (X-a)b$ schreiben. Sei $\deg(f) > 1$. Nach Definition der Ordnung gilt $f = (X-a_1)^{k_1} h$ mit einem Polynom h, das sich nicht mehr durch $X - a_1$ teilen läßt. Für $i = 2, \dots, m$ gilt nun: a_i ist eine Nullstelle von h, und $k_i = \operatorname{ord}_{a_i}(h)$ (dies ist leicht zu sehen: gemeinsame Faktoren $(X - a_i)$ von f und h lassen sich wegen der vorausgesetzten Nullteilerfreiheit kürzen). Man benutze nun die Induktionsvoraussetzung für h (es ist $\deg(h) < \deg(f)$ wegen Nullteilerfreiheit).

iii) folgt sofort aus ii). ∎

Sei $R = K$ ein Körper. Wir sagen, ein Polynom $f \in K[X]$ vom Grade n **zerfällt über K vollständig in Linearfaktoren**, falls es $a_1, \dots, a_n \in K$ gibt, nicht notwendig paarweise verschieden, und ein $a \in K^*$, mit

$$f = a(X-a_1) \cdot \ldots \cdot (X-a_n).$$

In diesem Fall ist jedes a_i eine Nullstelle von f, und es gibt keine weiteren. Durch Zusammenfassen der Linearfaktoren $X - a_i$ mit gleichem a_i erhält man eine Darstellung wie in 4.1.15 ii) mit $g = a$ vom Grade 0.

Feststellung und Definition 4.1.16 *Für einen Körper K sind die folgenden Aussagen äquivalent:*

i) Jedes Polynom $f \in K[X]$ besitzt eine Nullstelle in K.

ii) Jedes Polynom $f \in K[X]$ zerfällt über K vollständig in Linearfaktoren.

Körper mit dieser Eigenschaft heißen **algebraisch abgeschlossen**.

Beweis: Aus ii) folgt natürlich i). Die umgekehrte Richtung beweist man leicht durch Induktion nach dem Grad von f (durch fortgesetztes „Abspalten von Linearfaktoren“). ∎

Beispiele: 1. $\mathbb{Q}$ ist nicht algebraisch abgeschlossen, da das Polynom $X^2 - 2$ keine Nullstelle in $\mathbb{Q}$ besitzt.

2. $\mathbb{R}$ ist nicht algebraisch abgeschlossen, da das Polynom $X^2 + 1$ keine reelle Nullstelle besitzt.

3. Der sogenannte *Fundamentalsatz der Algebra* besagt, daß $\mathbb{C}$ algebraisch abgeschlossen ist.

4.2 Charakteristisches Polynom und Minimalpolynom

In diesem Abschnitt sei V stets ein n-dimensionaler K-Vektorraum. Wir ordnen jedem Endomorphismus $f \in \mathrm{End}_K(V)$ zwei Polynome in $K[X]$ zu, das *charakteristische Polynom* χ_f, und das *Minimalpolynom* μ_f. Die hauptsächliche Bedeutung dieser Polynome wird erst in den nächsten Abschnitten klarer werden.

Definitionen

Feststellung und Definition 4.2.1 *Sei f ein Endomorphismus eines n-dimensionalen K-Vektorraumes V. Sei A die Matrix von f bezüglich irgendeiner Basis von V. Man betrachte die Matrix $X\mathbf{1} - A \in M(n, K[X])$ (hierbei ist $X\mathbf{1}$ die skalare Matrix $\mathrm{diag}(X, \ldots, X)$).*

i) Die Determinante

$$\chi_f := \det(X\mathbf{1} - A) \in K[X]$$

ist unabhängig von der gewählten Basis, und heißt das **charakteristische Polynom** *von f.*

ii) χ_f ist ein normiertes Polynom vom Grad n.

iii) Ist $\chi_f = X^n + s_{n-1}X^{n-1} + \ldots + s_1 X + s_0$, so gilt

$$s_{n-1} = -\mathrm{tr}(f) = -\mathrm{tr}(A), \qquad s_0 = (-1)^n \det(f) = (-1)^n \det(A).$$

Beweis: i) Bezüglich einer anderen Basis hat f die Matrix PAP^{-1} mit $P \in \mathrm{GL}(n, K)$ (Folgerung 2.3.13). Es gilt dann

$$\begin{aligned}\det(X\mathbf{1} - PAP^{-1}) &= \det(P(X\mathbf{1})P^{-1} - PAP^{-1}) = \det(P(X\mathbf{1} - A)P^{-1}) \\ &= \det(P)\det(X\mathbf{1} - A)\det(P^{-1}) = \det(X\mathbf{1} - A)\end{aligned}$$

(für die erste Gleichheit beachte man, daß P mit $X\mathbf{1}$ vertauscht; vgl. Feststellung 2.2.6).

ii) und iii) In der Leibnizformel für $\det(X\mathbf{1} - A)$ gehört zur Permutation $\sigma = \mathrm{id}$ der Term

$$(X - a_{11}) \cdot \ldots \cdot (X - a_{nn}) = X^n - \mathrm{tr}(A)X^{n-1} + \text{ niedrigere Potenzen von } X.$$

Alle anderen Terme in der Leibnizformel liefern ebenfalls nur niedrigere Potenzen von X (höchstens vom Grad $n-2$). Es folgt, daß χ_f ein normiertes Polynom n-ten Grades ist, sowie die Behauptung über s_{n-1}.

In der Gleichung $\det(X\mathbf{1} - A) = X^n + s_{n-1}X^{n-1} + \ldots + s_1X + s_0$ setze man $X = 0$ (d.h., man wende den Einsetzungshomomorphismus $K[X] \to K$ an, der X auf 0 abbildet). Es folgt die Behauptung über s_0. ■

Bemerkungen: 1. Insbesondere ist jeder quadratischen Matrix $A \in M(n,K) = \mathrm{End}_K(K^n)$ ein charakteristisches Polynom χ_A zugeordnet; es ist $\chi_A = \det(X\mathbf{1}_n - A)$.

2. Manchmal wird das charakteristische Polynom als $\det(A - X\mathbf{1})$ definiert. Der Zusammenhang mit unserer Definition ist $\det(A - X\mathbf{1}) = (-1)^n\chi_f$.

3. Es ist diese Definition des charakteristischen Polynoms, bei der wir den Begriff der Determinante über kommutativen Ringen benötigen; ansonsten treten in diesem Buch nur Determinanten über Körpern auf.

Die Bedeutung des charakteristischen Polynoms χ_f liegt hauptsächlich darin, daß seine Nullstellen in K die im nächsten Abschnitt zu definierenden *Eigenwerte* von f sind.

Beispiel 4.2.2 Wir betrachten den Endomorphismus $r(\vartheta) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$ des $\mathbb{C}^2$, wobei $\theta \in \mathbb{R}$. Sein charakteristisches Polynom ist

$$\begin{aligned}\chi_{r(\vartheta)} &= \det\begin{pmatrix} X - \cos(\vartheta) & -\sin(\vartheta) \\ \sin(\vartheta) & X - \cos(\vartheta) \end{pmatrix} = (X - \cos(\vartheta))^2 + \sin(\vartheta)^2 \\ &= X^2 - 2X\cos(\vartheta) + 1 = (X - e^{i\vartheta})(X - e^{-i\vartheta}).\end{aligned}$$

Wie gesagt sind die Nullstellen $e^{\pm i\vartheta}$ gerade die *Eigenwerte* von $r(\vartheta)$; siehe dazu den nächsten Abschnitt.

Feststellung und Definition 4.2.3 *Sei $f \in \mathrm{End}_K(V)$. Es gibt ein eindeutig bestimmtes normiertes Polynom μ_f mit folgenden Eigenschaften:*

i) $\mu_f(f) = 0$.

ii) Ist $P \in K[X]$ mit $P(f) = 0$, so gilt $\mu_f | P$.

μ_f heißt das **Minimalpolynom** *von f.*

Beweis: Man betrachte den zu f gehörigen Einsetzungshomomorphismus

$$\begin{aligned} K[X] &\longrightarrow \mathrm{End}_K(V), \\ X &\longmapsto f. \end{aligned}$$

Sein Kern ist ein Ideal $I \subset K[X]$, das definitionsgemäß aus allen Polynomen P mit $P(f) = 0$ besteht. Wir definieren μ_f als den nach Folgerung 4.1.10 eindeutig bestimmten normierten Erzeuger von I. ■

Satz von Cayley–Hamilton

Nach Definition des Minimalpolynoms gilt $\mu_f(f) = 0$. Es gilt aber auch $\chi_f(f) = 0$, was der Inhalt des folgenden Satzes ist.

Satz 4.2.4 (Satz von Cayley–Hamilton) *Sei f ein Endomorphismus des endlich-dimensionalen K-Vektorraumes V, mit charakteristischem Polynom χ_f und Minimalpolynom μ_f. Dann gilt $\mu_f | \chi_f$, oder äquivalent, $\chi_f(f) = 0$.*

Beweis: Sei $A = (a_{ij})$ die Matrix von f bezüglich einer Basis $v_1, \ldots, v_n$. Wir arbeiten über dem Ring $R = \mathrm{End}_K(V)$, und betrachten folgende $n \times n$-Matrix über R:

$$B = \begin{pmatrix} f - a_{11}\mathrm{id} & -a_{12}\mathrm{id} & \ldots & -a_{1n}\mathrm{id} \\ -a_{21}\mathrm{id} & f - a_{22}\mathrm{id} & \ldots & -a_{2n}\mathrm{id} \\ \vdots & & \ddots & \vdots \\ -a_{n1}\mathrm{id} & -a_{n2}\mathrm{id} & \ldots & f - a_{nn}\mathrm{id} \end{pmatrix}.$$

Offenbar ist $\det(B) = \chi_f(f) \in R$. Wenn $\tilde{B}$ die zu B komplementäre Matrix ist (siehe Feststellung 3.4.2), so gilt also

$$B\tilde{B} = \chi_f(f)\mathbf{1}_n,$$

oder äquivalent, mit $B = (B_{ij})$ und $\tilde{B} = (\tilde{B}_{ij})$,

$$\sum_{k=1}^{n} B_{ik}\tilde{B}_{kj} = \chi_f(f)\delta_{ij}.$$

Dies ist eine Gleichung in $R = \mathrm{End}_K(V)$. Wir wenden beide Seiten auf den Basisvektor v_i an und summieren über i; das ergibt

$$\sum_{k=1}^{n} \tilde{B}_{kj} \sum_{i=1}^{n} B_{ik} v_i = \chi_f(f) v_j.$$

Die linke Seite ist aber Null, denn

$$\sum_{i=1}^{n} B_{ik} v_i = 0$$

wegen $f v_k = \sum_{i=1}^{n} a_{ik} v_i$. Also ist $\chi_f(f) v_j = 0$ für alle j, und damit $\chi_f(f) = 0$. ∎

Folgerung 4.2.5 *Das Minimalpolynom μ_f hat höchstens den Grad $n = \dim_K(V)$ (mit $\deg(\mu_f) = n$ genau dann, wenn $\mu_f = \chi_f$).*

Nullstellen von χ_f und μ_f

Feststellung 4.2.6 *Sei $\dim_K(V) = n$ und $f \in \mathrm{End}_K(V)$. Dann gilt*

$$\mu_f | \chi_f \qquad \text{und} \qquad \chi_f | \mu_f^n.$$

Insbesondere haben χ_f und μ_f die gleichen Nullstellen in K (wobei aber natürlich $\mathrm{ord}_\lambda(\mu_f) < \mathrm{ord}_\lambda(\chi_f)$ sein kann).

Beweis: $\mu_f | \chi_f$ ist gerade die Aussage von Cayley-Hamilton. Für die zweite Aussage sei A die Matrix von f bezüglich irgendeiner Basis. Sei $R = K[A]$ das Bild das durch A definierten Einsetzungshomomorphismus $K[X] \to M(n,K)$ (welches einfach aus allen Polynomausdrücken in A besteht). Wenn $\mu_f = X^m + a_{m-1}X^{m-1} + \ldots + a_0$, so betrachten wir das Polynom

$$P = \mathbf{1}\,X^m + a_{m-1}\mathbf{1}\,X^{m-1} + \ldots + a_1\mathbf{1}\,X + a_0\mathbf{1} \in R[X].$$

Es besitzt in $R[X]$ die Nullstelle A (wegen Cayley-Hamilton), und also gilt

$$P = (X - A)Q \qquad \text{für ein } Q \in R[X]$$

nach Feststellung 4.1.15 i). Auf diese Gleichung wenden wir nun den Einsetzungshomomorphismus

$$\begin{aligned} R[X] &\longrightarrow M(n, K[X]), \\ X &\longmapsto \begin{pmatrix} X & & \\ & \ddots & \\ & & X \end{pmatrix}, \end{aligned}$$

an; das Ergebnis ist

$$\begin{pmatrix} \mu_f & & \\ & \ddots & \\ & & \mu_f \end{pmatrix} = (X\mathbf{1} - A)q$$

für irgendeine Matrix $q \in M(n, K[X])$. Wir bilden die Determinante auf beiden Seiten:

$$\mu_f^n = \chi_f \ \det(q).$$

Dies ist nun eine Gleichung in $K[X]$, und es folgt $\chi_f | \mu_f^n$. ∎

Beispiel: Wir betrachten den Endomorphismus $r(\vartheta)$ des $\mathbb{C}^2$ aus Beispiel 4.2.2. Ist $\vartheta \notin \pi\mathbb{Z}$, so gilt $e^{i\vartheta} \neq e^{-i\vartheta}$, und aus 4.2.6 folgt $\mu_{r(\vartheta)} = \chi_{r(\vartheta)} = (X - e^{i\vartheta})(X - e^{-i\vartheta})$. Ist dagegen $\vartheta \in \pi\mathbb{Z}$, so gilt $r(\vartheta) = \pm\mathbf{1}$, und deshalb $\mu_{r(\vartheta)} = X \mp 1$; in diesem Fall ist $\mu_{r(\vartheta)}$ also ein *echter* Teiler von $\chi_{r(\vartheta)}$.

4.3 Eigenwerte

Eigenwerte und Eigenvektoren

Definition 4.3.1 *Sei f ein Endomorphismus eines (nicht notwendig endlich-dimensionalen) K-Vektorraumes V. Ein Skalar $\lambda \in K$ heißt* **Eigenwert** *von f, falls es ein $v \in V$ gibt mit $v \neq 0$ und*

$$fv = \lambda v. \tag{4.3}$$

Jeder Vektor $v \in V$ mit (4.3) heißt **Eigenvektor** *von f zum Eigenwert λ. Alle Eigenvektoren von f zum gleichen Eigenwert λ bilden offenbar einen Untervektorraum von V, den* **Eigenraum** *von f zum Eigenwert λ.*

Beispiele: 1. Der Eigenraum zum Eigenwert 0 ist der Kern von f.

2. id_V hat den einzigen Eigenwert 1, und der zugehörige Eigenraum ist ganz V.

3. Man betrachte den Endomorphismus $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ des K^2. Dieser besitzt die Eigenwerte 1 und 2. Ein Eigenvektor zum Eigenwert 1 ist $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, ein Eigenvektor zum Eigenwert 2 ist $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

4. Für $\vartheta \in \mathbb{R}$ betrachte man den Endomorphismus $r(\vartheta) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$ des $\mathbb{R}^2$. Wenn ϑ nicht gerade ein Vielfaches von π ist, so besitzt $r(\vartheta)$ keine Eigenwerte in $\mathbb{R}$, denn es handelt sich um eine Drehung um den Winkel ϑ. Faßt man $r(\vartheta)$ dagegen als Endomorphismus des $\mathbb{C}$-Vektorraumes $\mathbb{C}^2$ auf, so besitzt $r(\vartheta)$ die Eigenwerte $e^{\pm i\vartheta}$ mit den Eigenvektoren $\begin{pmatrix} 1 \\ \pm i \end{pmatrix}$; vgl. mit Beispiel 4.2.2.

5. Sei $V = C^\infty(a,b)$ der $\mathbb{R}$-Vektorraum der beliebig häufig differenzierbaren Funktionen auf dem Intervall (a,b). Sei $f = \frac{d}{dx} : V \to V$ der Differentiationsoperator. Dann ist jede reelle Zahl Eigenwert von f, und alle Eigenräume sind eindimensional. Denn die Eigenwertgleichung

$$\frac{d}{dx}\varphi = \lambda\varphi$$

ist eine Differentialgleichung für φ, deren Lösung genau die Vielfachen der Funktion $\varphi(x) = e^{\lambda x}$ sind.

Eigenwerte und Polynome

Es ist ein fundamentales Problem, die Eigenwerte – und eventuell zu jedem Eigenwert den zugehörigen Eigenraum – eines vorgelegten Endomorphismus zu bestimmen. Für endlich-dimensionale Vektorräume sind dabei die im letzten Abschnitt definierten Polynome χ_f und μ_f von Bedeutung:

Satz 4.3.2 *Sei V ein endlich-dimensionaler K-Vektorraum, und $f \in \mathrm{End}_K(V)$. Für ein Skalar $\lambda \in K$ sind die folgenden Aussagen äquivalent:*

i) λ ist ein Eigenwert von f

ii) λ ist Nullstelle des charakteristischen Polynoms χ_f.

iii) λ ist Nullstelle des Minimalpolynoms μ_f.

Beweis: Sei A die Matrix von f bezüglich irgendeiner Basis. Wir haben

$$\begin{aligned}
\lambda \text{ Eigenwert von } f &\iff \text{Es gibt } v \neq 0 \text{ mit } fv = \lambda v && \text{(Definition von „Eigenwert“)} \\
&\iff \text{Es gibt } v \neq 0 \text{ mit } (f - \lambda\mathrm{id})v = 0 \\
&\iff f - \lambda\,\mathrm{id} \text{ ist kein Isomorphismus} \\
&\iff A - \lambda\,\mathbf{1} \text{ ist nicht invertierbar} && \text{(Folgerung 2.3.10)} \\
&\iff \det(A - \lambda\mathbf{1}) = 0 && \text{(Feststellung 3.2.12)} \\
&\iff \chi_f(\lambda) = 0 && \text{(Definition von } \chi_f\text{).}
\end{aligned}$$

Also sind i) und ii) äquivalent. Die Äquivalenz von ii) und iii) haben wir schon in 4.2.6 gesehen. ∎

Um die Eigenwerte eines vorgelegten Endomorphismus f (eines endlich-dimensionalen K-Vektorraumes) zu finden, kann man also die Nullstellen des charakteristischen Polynoms χ_f bestimmen (dieses ist als $\det(X\mathbf{1} - A)$ mechanisch ausrechenbar und in aller Regel leichter zu finden als μ_f). Zumindest bei kleinen Dimensionen führt dies oft zum Erfolg, während sich die Nullstellen von Polynomen höheren Grades im allgemeinen nicht bestimmen lassen.

Illustriert wird Satz 4.3.2 durch obiges Beipiel 4 und Beispiel 4.2.2.

Folgerung 4.3.3 *Falls $\dim_K(V) = n$, so hat $f \in \mathrm{End}_K(V)$ höchstens n verschiedene Eigenwerte in K.*

Beweis: Dies folgt aus Feststellung 4.1.15. ∎

Beispiel 4.3.4 Wir betrachten den $\mathbb{R}$-Vektorraum $\mathbb{R}^n$. Für gerades n muß ein Endomorphismus f des $\mathbb{R}^n$ keine (reellen) Eigenwerte besitzen, wie Beispiel 4 auf Seite 92 zeigt. Sei dagegen n ungerade. Das charakteristische Polynom χ_f ist normiert vom Grade n, geht als Funktion von $x \in \mathbb{R}$ also gegen $\pm\infty$, wenn x gegen $\pm\infty$ geht. Aus Stetigkeitsgründen besitzt χ_f daher mindestens eine reelle Nullstelle. Es folgt: Ein Endomorphismus des $\mathbb{R}^n$ besitzt bei ungeradem n mindestens einen (reellen) Eigenwert.

Algebraische und geometrische Vielfachheit

Definition 4.3.5 *Sei V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$, und $\lambda \in K$ ein Eigenwert von f.*

i) Sei $V_\lambda = \{v \in V : fv = \lambda v\}$ der Eigenraum zu λ. Die Dimension $\dim_K(V_\lambda)$ heißt die **geometrische Vielfachheit** *des Eigenwertes λ.*

ii) Sei $\chi_f \in K[X]$ das charakteristische Polynom von f. Die Nullstellenordnung $\mathrm{ord}_\lambda(\chi_f)$ heißt die **algebraische Vielfachheit** *des Eigenwertes λ.*

Beispiel: Der Eigenwert λ des Endomorphismus $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ des K^2 hat die algebraische Vielfachheit 2, aber die geometrische Vielfachheit 1.

Feststellung 4.3.6 *Die geometrische Vielfachheit ist höchstens so groß wie die algebraische. Genauer: Sei f ein Endomorphismus des endlich-dimensionalen Vektorraumes V, und λ ein Eigenwert von f mit Eigenraum V_λ. Dann gilt:*

$$\dim_K(V_\lambda) \le \mathrm{ord}_\lambda(\chi_f).$$

Beweis: Sei $v_1, \ldots, v_r$ eine Basis von V_λ. Wir ergänzen sie zu einer Basis $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$ von V. Bezüglich dieser Basis hat die Matrix von f die Gestalt

$$A = \begin{pmatrix} \lambda\mathbf{1}_r & * \\ \mathbf{0} & B \end{pmatrix},$$

wobei $\mathbf{0}$ die $(n-r)\times r$-Nullmatrix ist, und B vom Format $(n-r)\times(n-r)$. Nach der Kästchenformel gilt

$$\chi_f = \det(X\mathbf{1}_r - \lambda\mathbf{1}_r)\det(X\mathbf{1}_{n-r} - B) = (X-\lambda)^r \det(X\mathbf{1}_{n-r} - B).$$

Es folgt $\mathrm{ord}_\lambda(\chi_f) \geq r = \dim(V_\lambda)$. ∎

Beispiel: Das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} 2 & 6 & 0 & -2 \\ 0 & 4 & 0 & -1 \\ 0 & 3 & 2 & -1 \\ 0 & 6 & 0 & -1 \end{pmatrix} \in M(4,\mathbb{Q})$$

(d.h., das charakteristische Polynom des durch A definierten Endomorphismus des $\mathbb{Q}^4$) berechnet sich zu $\chi_A = (X-1)(X-2)^3$. Also besitzt A die Eigenwerte 1 und 2. Aus der letzten Feststellung folgt, daß der Eigenraum zum Eigenwert 1 eindimensional ist, während man über die Dimension des anderen Eigenraums zunächst nur Aussagen kann, daß sie kleiner oder gleich 3 ist. Der Ansatz $Av = 2v$ führt auf ein lineares Gleichungssystem, dessen Lösungsraum sich als zweidimensional erweist. Dies ist die Dimension des Eigenraums zum Eigenwert 2.

4.4 Diagonalisierbare Endomorphismen

Diagonalisierbarkeit

Feststellung 4.4.1 *Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. Genauer: Sei V ein beliebiger K-Vektorraum, $\lambda_1,\dots,\lambda_r$ paarweise verschiedene Eigenwerte eines Endomorphismus f von V, und v_i ein Eigenvektor zu λ_i $(i = 1,\dots,r)$. Dann sind $v_1,\dots,v_r$ linear unabhängig.*

Beweis: Wir führen Induktion nach r, wobei zum Induktionsanfang $r = 1$ nichts zu sagen ist. Sei $r > 1$. Auf eine Linearkombination $\sum_{i=1}^r a_i v_i = 0$ wende man den Endomorphismus $f - \lambda_r \mathrm{id}$ an; das Ergebnis ist

$$\sum_{i=1}^{r-1} a_i(\lambda_i - \lambda_r)v_i = 0.$$

Die Induktionsvoraussetzung liefert $a_i(\lambda_i - \lambda_r) = 0$ für $i = 1,\dots,r-1$. Da die λ_i paarweise verschieden sind, folgt $a_i = 0$ für $i = 1,\dots,r-1$, und dann auch $a_r = 0$. ∎

Satz und Definition 4.4.2 *Sei f ein Endomorphismus eines endlich-dimensionalen K-Vektorraumes V mit den Eigenwerten $\lambda_1,\dots,\lambda_r$. Die folgenden Aussagen sind äquivalent:*

i) V besitzt eine Basis aus Eigenvektoren von f.

ii) V ist die direkte Summe aller Eigenräume: $V = \bigoplus_{i=1}^r V_{\lambda_i}$.

iii) Die Matrix von f bezüglich einer geeigneten Basis von V ist eine Diagonalmatrix.

iv) Das charakteristische Polynom χ_f zerfällt über K vollständig in Linearfaktoren, und die geometrischen Vielfachheiten sind gleich den algebraischen:

$$\dim_K(V_{\lambda_i}) = \mathrm{ord}_{\lambda_i}(\chi_f) \qquad \text{für } i = 1, \dots, r.$$

v) Das Minimalpolynom von f zerfällt über K vollständig in Linearfaktoren und besitzt nur einfache Nullstellen.

Sind diese Bedingungen erfüllt, so heißt f **diagonalisierbar**.

Beweis: i) $\Rightarrow$ iii) ist klar.

iii) $\Rightarrow$ iv) Sei $A = \mathrm{diag}(a_1, \dots, a_n)$ die Matrix von f bezüglich einer geeigneten Basis. Die Koeffizienten a_i sind offenbar Eigenwerte von f, und wenn der Eigenwert λ_j auf der Diagonalen m_j-mal vorkommt, so ist die geometrische Vielfachheit von λ_j mindestens m_j. Andererseits kann man χ_f direkt an A ablesen. Erstens erkennt man, daß χ_f über K in Linearfaktoren zerfällt, und zweitens, daß die algebraische Vielfachheit von λ_j genau m_j ist. Zusammen mit 4.3.6 folgt $\dim(V_{\lambda_j}) = m_j$.

iv) $\Rightarrow$ ii) Sei $\chi_f = \prod_{i=1}^r (X - \lambda_i)^{m_i}$. Innerhalb von V haben wir den Untervektorraum $V_{\lambda_1} \oplus \dots \oplus V_{\lambda_r}$ (die Summe ist direkt nach Feststellung 4.4.1). Dessen Dimension ist nach Voraussetzung gleich $\sum_{i=1}^r m_i = n = \dim(V)$, und somit ist er schon ganz V.

ii) $\Rightarrow$ i) ist klar.

ii) $\Rightarrow$ v) Jeder Vektor $v \in V$ ist nach Voraussetzung Summe von Eigenvektoren, und wird daher vom Endomorphismus

$$(f - \lambda_1 \mathrm{id}) \cdot \ldots \cdot (f - \lambda_r \mathrm{id})$$

annuliert. Dieser Endomorphismus ist also 0 in $\mathrm{End}_K(V)$, oder mit anderen Worten, das Polynom

$$p = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_r) \in K[X]$$

liegt im Kern des durch f definierten Einsetzungshomomorphismus $K[X] \to \mathrm{End}_K(V)$. Nach Definition des Minimalpolynoms μ_f folgt $\mu_f | p$. Da p über K in Linearfaktoren zerfällt und nur einfache Nullstellen besitzt, gilt das gleiche dann für μ_f.

v) $\Rightarrow$ ii) wird durch Induktion nach $\dim(V)$ bewiesen, wobei der Fall $\dim(V) = 1$ trivial ist. Sei $\dim(V) > 1$. Nach Voraussetzung gilt

$$\mu_f = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_r) \tag{4.4}$$

mit gewissen, paarweise verschiedenen $\lambda_i \in K$. Division mit Rest (Feststellung 4.1.5) liefert eine Darstellung

$$(X - \lambda_2) \cdot \ldots \cdot (X - \lambda_r) = q(X)(X - \lambda_1) + c \qquad \text{mit } q \in K[X] \text{ und } c \in K^*.$$

Einsetzen von f und Anwenden auf einen Vektor $v \in V$ ergibt

$$cv = (f - \lambda_2 \mathrm{id}) \cdot \ldots \cdot (f - \lambda_r \mathrm{id})v - (f - \lambda_1 \mathrm{id})q(f)v.$$

Der erste Summand $(f-\lambda_2\mathrm{id})\cdot\ldots\cdot(f-\lambda_r\mathrm{id})v$ liegt wegen (4.4) im Kern von $f-\lambda_1\mathrm{id}$, der zweite Summand liegt im Bild von $f-\lambda_1\mathrm{id}$. Wegen $c\neq 0$ schließen wir auf $V=\ker(f-\lambda_1\mathrm{id})+\mathrm{im}(f-\lambda_1\mathrm{id})$, und aus Dimensionsgründen (Feststellung 2.1.1 v)) dann auf

$$V=\ker(f-\lambda_1\mathrm{id})\oplus\mathrm{im}(f-\lambda_1\mathrm{id}). \tag{4.5}$$

Man überlege sich nun, daß $(X-\lambda_2)\cdot\ldots\cdot(X-\lambda_r)$ das Minimalpolynom des Endomorphismus $f|_W$ des Vektorraumes $W:=\mathrm{im}(f-\lambda_1\mathrm{id})$ ist. Man darf deshalb die Induktionsvoraussetzung auf W und $f|_W$ anwenden, und erhält eine direkte Summe für W der gewünschten Gestalt. Zusammen mit (4.5) erhält man die Behauptung wegen $\ker(f-\lambda_1\mathrm{id})=V_{\lambda_1}$. ∎

Beispiele: 1. Eine *Matrix* $A\in M(n,K)$ (wie üblich aufgefaßt als Endomorphismus des K^n) ist genau dann diagonalisierbar, wenn es ein $S\in\mathrm{GL}(n,K)$ gibt, so daß SAS^{-1} eine Diagonalmatrix ist; siehe 2.3.13.

2. Sei $\dim_K(V)=n$. Falls das charakteristische Polynom von $f\in\mathrm{End}_K(V)$ über K in n *verschiedene* Linearfaktoren zerfällt,

$$\chi_f=(X-\lambda_1)\cdot\ldots\cdot(X-\lambda_n)\qquad\text{mit paarweise verschiedenen }\lambda_i,$$

so ist f diagonalisierbar, denn Bedingung iv) ist erfüllt. Ist v_i ein Eigenvektor zu λ_i, so ist $v_1,\ldots,v_n$ eine Basis von V, und bezüglich dieser besitzt f die Matrix $\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$.

Simultane Diagonalisierbarkeit

Seien gleich mehrere Endomorphismen $f_1,\ldots,f_r$ eines n-dimensionalen K-Vektorraumes V gegeben, die alle diagonalisierbar sind. Wir stellen die Frage, ob sie *simultan* diagonalisierbar sind, d.h., ob es eine Basis $v_1,\ldots v_n$ von V gibt, so daß jedes v_i Eigenvektor für *jeden* der Endomorphismen f_j ist. Im allgemeinen ist dies nicht der Fall, wie einfache Gegenbeispiele zeigen. Es erweist sich jedoch als richtig, wenn man die f_j als *paarweise kommutierend* voraussetzt, d.h., $f_if_j=f_jf_i$ für alle $i,j\in\{1,\ldots,r\}$.

Sei $\mathcal{A}$ eine kommutative K-Unteralgebra der K-Algebra $\mathrm{End}_K(V)$. Unter einem **Charakter** von $\mathcal{A}$ verstehen wir einen K-Algebren-Homomorphismus $\mathcal{A}\to K$, d.h., eine Linearform λ auf $\mathcal{A}$, die zusätzlich $\lambda(AB)=\lambda(A)\lambda(B)$ für alle $A,B\in\mathcal{A}$ erfüllt. Für einen solchen Charakter λ betrachten wir den Unterraum

$$V_\lambda=\{v\in V:\ Av=\lambda(A)v\quad\text{für alle }A\in\mathcal{A}\}.$$

Ist $V_\lambda\neq 0$, so nennen wir λ eine **Wurzel** von $\mathcal{A}$, und V_λ den zugehörigen **Wurzelraum**. Man beachte, daß dies eine Verallgemeinerung der vorherigen Begriffe „Eigenwert“ und „Eigenraum“ ist: Ist $\mathcal{A}=K[f]$ von einem einzigen Endomorphismus f erzeugt, und λ eine Wurzel, so ist $\lambda(f)$ ein Eigenwert von f, und der zugehörige Eigenraum $V_{\lambda(f)}$ stimmt mit dem Wurzelraum V_λ überein.

Die Elemente von V_λ, die nicht Null sind, nennen wir auch **Eigenvektoren** von $\mathcal{A}$ bezüglich λ. Wir verallgemeinern zunächst Feststellung 4.4.1:

Lemma 4.4.3 *Seien $\lambda_1,\ldots,\lambda_r$ paarweise verschiedene Wurzeln von $\mathcal{A}$, und v_i ein Eigenvektor von $\mathcal{A}$ bezüglich λ_i, $i=1,\ldots,r$. Dann sind $v_1,\ldots,v_r$ linear unabhängig.*

Beweis: Wir führen Induktion nach r, wobei für $r = 1$ nichts zu beweisen ist. Sei $\sum_{i=1}^{r} a_i v_i = 0$ eine Linearkombination. Ein $i \in \{1, \dots, r-1\}$ sei festgehalten. Nach Voraussetzung gibt es ein $f \in \mathcal{A}$ mit $\lambda_i(f) \neq \lambda_r(f)$. Wir wenden den Endomorphismus $f - \lambda_r(f)\mathrm{id}$ auf die Linearkombination an, und erhalten

$$\sum_{j=1}^{r-1} a_j(\lambda_j(f) - \lambda_r(f))v_j = 0.$$

Nach Induktionsvoraussetzung folgt $a_i(\lambda_i(f) - \lambda_r(f)) = 0$, und dann $a_i = 0$. Also ist $a_1 = \dots = a_{r-1} = 0$, und dann auch $a_r = 0$. ∎

Lemma 4.4.4 *Seien $f_1, \dots, f_m$ paarweise kommutierende Endomorphismen von V, und $\mathcal{A}$ die (kommutative) K-Unteralgebra von $\mathrm{End}_K(V)$, die von den f_i erzeugt wird. Dann sind äquivalent:*

i) Es gibt eine gemeinsame Eigenbasis $v_1, \dots, v_n$ für die f_j, d.h., jedes v_i ist Eigenvektor für jedes f_j.

ii) Sind $\lambda_1, \dots, \lambda_r$ die Wurzeln von $\mathcal{A}$, so ist V die direkte Summe aller Wurzelräume: $V = \bigoplus_{i=1}^{r} V_{\lambda_i}$.

Beweis: i) ⇒ ii) Wir betrachten ein festes $i \in \{1, \dots, n\}$. Nach Voraussetzung gibt es $\lambda_1, \dots, \lambda_m \in K$ mit

$$f_j v_i = \lambda_j v_i \qquad \text{für alle } j \in \{1, \dots, m\}. \tag{4.6}$$

Die Algebra $\mathcal{A}$ besteht aus allen Endomorphismen der Form $P(f_1, \dots, f_m)$, wobei $P(X_1, \dots, X_m)$ ein Polynom in den m Unbekannten $X_1, \dots, X_m$ ist. Aus (4.6) folgt

$$P(f_1, \dots, f_m)v_i = P(\lambda_1, \dots, \lambda_m)v_i.$$

Also ist v_i Eigenvektor für *jedes* Element von $\mathcal{A}$, d.h., es gibt eine Abbildung $\lambda : \mathcal{A} \to K$ mit $fv = \lambda(f)v$ für alle $f \in \mathcal{A}$. Es ist leicht zu sehen, daß λ ein Charakter von $\mathcal{A}$ ist. Also ist λ eine Wurzel, und v_i liegt im Wurzelraum V_λ. Wir haben gezeigt, daß V die Summe aller Wurzelräume ist. Die Summe ist sogar direkt nach Lemma 4.4.3.

ii) ⇒ i) ist klar. ∎

Satz 4.4.5 *Sei V eine n-dimensionaler K-Vektorraum. Seien $f_1, \dots, f_m$ paarweise kommutierende Endomorphismen von V, und jedes f_i sei diagonalisierbar. Dann sind die f_i simultan diagonalisierbar, d.h., es gibt eine Basis $v_1, \dots, v_n$ von V, so daß jedes v_i Eigenvektor zu jedem f_j ist.*

Beweis: Wir führen Induktion nach m, und nehmen gleich $m > 1$ an. Sei $\mathcal{A}$ die von $f_1, \dots, f_{m-1}$ erzeugte Unteralgebra von $\mathrm{End}_K(V)$. Nach Induktionsvoraussetzung und Lemma 4.4.4 gilt

$$V = \bigoplus_{i=1}^{r} V_{\lambda_i}, \tag{4.7}$$

wobei $\lambda_1, \ldots, \lambda_r$ die Wurzeln von $\mathcal{A}$ sind, und V_{λ_i} die zugehörigen Wurzelräume. Aufgrund der Voraussetzung des Kommutierens gilt

$$f_r(V_{\lambda_i}) \subset V_{\lambda_i} \qquad \text{für alle } i = 1, \ldots, r.$$

Seien $\mu_1, \ldots, \mu_s$ die Eigenwerte von f_r auf V, und $W_1, \ldots, W_s$ die zugehörigen Eigenräume, so daß $V = W_1 \oplus \ldots \oplus W_s$. Wir behaupten, daß

$$V_{\lambda_i} = (V_{\lambda_i} \cap W_1) \oplus \ldots \oplus (V_{\lambda_i} \cap W_s) \tag{4.8}$$

für jedes i. Dabei ist „$\supset$“ trivial. Für „$\subset$“ sei $v \in V_{\lambda_i}$. Auf jeden Fall gibt es eine (eindeutige) Zerlegung $v = v_1 + \ldots + v_s$ mit $v_j \in W_j$. Für $f \in \mathcal{A}$ gilt

$$\lambda_i(f)v_1 + \ldots \lambda_i(f)v_s = \lambda_i(f)v = fv = fv_1 + \ldots + fv_s.$$

Wiederum aufgrund des Kommutierens gilt $fv_j \in W_j$, und wir schließen auf $\lambda_i(f)v_j = fv_j$. Also ist $v_j \in V_{\lambda_i}$, und (4.8) ist gezeigt.

Jeder von Null verschiedene Vektor in $V_{\lambda_i} \cap W_j$ ist offenbar Eigenvektor für jedes f_k, $k = 1, \ldots, r$. Aus (4.7) und (4.8) folgt daher die Behauptung. ∎

Wir geben drei äquivalente Formulierungen von Satz 4.4.5:

1. Sei $\mathcal{A}$ eine kommutative Unteralgebra von $\mathrm{End}_K(V)$, die von diagonalisierbaren Elementen erzeugt wird. Dann ist V die direkte Summe aller Wurzelräume von $\mathcal{A}$.

2. Seien $A_1, \ldots, A_m$ paarweise kommutierende Matrizen in $M(n, K)$. Dann gibt es ein $S \in \mathrm{GL}(n, K)$, so daß SA_iS^{-1} eine Diagonalmatrix ist für jedes $i = 1, \ldots, m$.

3. Sei $\mathcal{A}$ eine kommutative Unteralgebra von $M(n, K)$, die von diagonalisierbaren Matrizen erzeugt wird. Dann gibt es ein $S \in \mathrm{GL}(n, K)$, so daß die Algebra $S\mathcal{A}S^{-1}$ nur aus Diagonalmatrizen besteht.

4.5 Nilpotente Endomorphismen

Trigonalisierung

Feststellung und Definition 4.5.1 *Für einen Endomorphismus f eines endlich-dimensionalen K-Vektorraumes V sind äquivalent:*

i) Es gibt eine Basis von V, bezüglich derer die Matrix von f eine obere Dreiecksmatrix ist.

ii) Das charakteristische Polynom χ_f zerfällt über K vollständig in Linearfaktoren.

Ist dies erfüllt, so heißt f **trigonalisierbar**.

Beweis: i) $\Rightarrow$ ii) ist sofort einzusehen. Für die andere Richtung führen wir Induktion nach $n = \dim_K(V)$, und setzen gleich $n > 1$ voraus. Sei $\chi_f = (X - \lambda_1) \cdot \ldots \cdot (X - \lambda_n)$. Aufgrund von Satz 4.3.2 ist λ_1 ein Eigenwert von f. Wir ergänzen einen zugehörigen Eigenvektor v_1 zu einer Basis $v_1, v_2, \ldots, v_n$ von V. Bezüglich dieser besitzt f die Matrix

$$A = \begin{pmatrix} \lambda_1 & * \\ 0 & B \end{pmatrix} \qquad \text{mit } B \in M(n-1, K).$$

Sei $W = \langle v_2, \dots, v_n \rangle$. Für jedes $w \in W$ gibt es eine eindeutige Zerlegung

$$fw = a(w)v_1 + g(w) \qquad \text{mit } a(w) \in K \text{ und } g(w) \in W.$$

Man sieht leicht, daß die so definierte Abbildung $g : W \to W$ ein Endomorphismus von W ist, der bezüglich der Basis $v_2, \dots, v_n$ die Matrix B besitzt. Aufgrund der Kästchenformel gilt daher

$$\chi_f = (X - \lambda_1)\det(X\mathbf{1}_{n-1} - B) = (X - \lambda_1)\chi_g,$$

und folglich $\chi_g = (X - \lambda_2) \cdot \ldots \cdot (X - \lambda_n)$. Auf χ_g können wir somit die Induktionsvoraussetzung anwenden, und erhalten eine Basis $v_2', \dots, v_n'$ von W, bezüglich derer g obere Dreiecksform hat. Mit $v_1, v_2', \dots, v_n'$ haben wir dann eine Basis von f wie gewünscht. ∎

Für eine Matrix $A \in M(n, K)$ gilt also: Zerfällt χ_A über K in Linearfaktoren, so gibt es ein $S \in \mathrm{GL}(n, K)$, so daß SAS^{-1} eine obere Dreiecksmatrix ist.

Bemerkung 4.5.2 Ist die Matrix von f eine obere Dreiecksmatrix, so müssen auf der Hauptdiagonalen ersichtlich die Eigenwerte von f stehen, denn dies sind die Nullstellen des charakteristischen Polynoms.

Folgerung 4.5.3 *Ist K algebraisch abgeschlossen, so läßt sich jedes $f \in \mathrm{End}_K(V)$ auf obere Dreiecksgestalt bringen. Dies ist insbesondere für $K = \mathbb{C}$ der Fall.*

Nilpotente Endomorphismen

Satz und Definition 4.5.4 *Sei f ein Endomorphismus des n-dimensionalen K-Vektorraumes V. Die folgenden Bedingungen sind äquivalent:*

i) $f^k = 0$ *für ein* $k \in \mathbb{N}$.

ii) $f^n = 0$.

iii) $\chi_f = X^n$.

iv) Das Minimalpolynom μ_f ist eine Potenz von X.

v) Bezüglich einer geeigneten Basis ist die Matrix von f eine obere Dreiecksmatrix mit Nullen auf der Hauptdiagonalen.

vi) χ_f zerfällt über K in Linearfaktoren, und f besitzt den einzigen Eigenwert Null.

Sind diese Bedingungen erfüllt, so heißt f **nilpotent**.

Beweis: iii) $\Rightarrow$ v) folgt aus Feststellung 4.5.1 und der anschließenden Bemerkung 4.5.2.

v) $\Rightarrow$ ii) Sei $A \in M(n, K)$ eine obere Dreiecksmatrix mit Nullen auf der Hauptdiagonalen. Wir müssen nur zeigen, daß $A^n = 0$. Dies sieht man leicht durch direkte Rechnung: A^2 hat Nullen auf der ersten Nebendiagonalen, A^3 auf der zweiten, usw.

ii) $\Rightarrow$ iv) Nach Voraussetzung und Definition des Minimalpolynoms gilt $\mu_f | X^n$. Dann muß μ_f aber selber eine Potenz von X sein.

iv) $\Rightarrow$ i) ist klar.

i) $\Rightarrow$ vi) Das Minimalpolynom μ_f muß eine Potenz von X sein, und wegen Feststellung 4.2.6 gilt das gleiche für χ_f. Insbesondere zerfällt χ_f über K in Linearfaktoren, und besitzt nur die Nullstelle 0.

vi) $\Rightarrow$ iii) Klar, da die Nullstellen von χ_f die Eigenwerte von f sind. ∎

Klassifikation

Das Paradebeispiel eines nilpotenten Endomorphismus ist die n-te **Jordanmatrix**, definiert durch

$$J_n := \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & \dots & & 0 \end{pmatrix} \in M(n,K) \tag{4.9}$$

(Einsen auf der ersten Nebendiagonalen und sonst Nullen); es ist $J_1 = (0)$. Man mache sich klar, wie die Potenzen J_n^k aussehen: Mit wachsendem k rückt die Diagonale mit den Einsen höher. Insbesondere ist $J_n^n = 0$, und $J_n^k \neq 0$ für $k < n$. Der Hauptsatz über nilpotente Endomorphismen (Satz 4.5.6) besagt, daß die Matrix eines solchen, bei geeigneter Basiswahl, aus Jordanblöcken zusammengesetzt ist.

Lemma 4.5.5 *Sei* $\dim_K(V) = n$ *und* $f \in \operatorname{End}_K(V)$ *nilpotent. Für ein* $v \in V$ *sind die folgenden Aussagen äquivalent:*

i) Es gilt $f^n v = 0$*, und die Vektoren* $v, fv, \dots, f^{n-1}v$ *bilden eine Basis von* V*.*

ii) Es gilt $f^n v = 0$*, aber* $f^{n-1}v \neq 0$*.*

Ist dies erfüllt, so ist J_n *die Matrix von* f *bezüglich der Basis* $f^{n-1}v, f^{n-2}v, \dots, fv, v$*, und wir nennen* v *einen* **f-zyklischen Vektor**.

Beweis: Aus i) folgt natürlich ii). Sei umgekehrt ii) erfüllt. Wir müssen zeigen, daß $v, fv, \dots, f^{n-1}v$ linear unabhängig sind, daß es also kein Polynom $P \in K[X]$ mit

$$P(f)v = 0 \qquad \text{und} \qquad \deg(P) < n$$

geben kann außer dem Nullpolynom. Sei $W \subset V$ der von $v, fv, f^2v, \dots$ aufgespannte Unterraum. Es gilt $f(W) \subset W$, und $f^n = 0$ auf W wegen $f^n v = 0$. Das Minimalpolynom von $f|_W$ teilt daher X^n, ist also selbst eine X-Potenz. Andererseits annuliert nach Voraussetzung keine kleinere Potenz als f^n den Vektor v. Somit ist das Minimalpolynom gleich X^n, und es folgt $P(f) \neq 0$ auf W für jedes Polynom $P \neq 0$ von einem Grad kleiner als n. Dann muß auch $P(f)v \neq 0$ sein.

Die letzte Aussage ist klar. ∎

Satz 4.5.6 *Sei* $\dim_K(V) = n$ *und* $f \in \operatorname{End}_K(V)$ *ein nilpotenter Endomorphismus. Sei* $d \in \mathbb{N}$ *die kleinste Zahl mit* $f^d = 0$*. Bezüglich einer geeigneten Basis (einer sogenannten* **Jordanbasis***) besitzt* f *dann eine Matrix der Form*

$$
\begin{pmatrix}
J_d & & & & & & & & \\
& \ddots & & & & & & & \\
& & J_d & & & & & & \\
& & & J_{d-1} & & & & & \\
& & & & \ddots & & & & \\
& & & & & J_{d-1} & & & \\
& & & & & & \ddots & & \\
& & & & & & & J_1 & \\
& & & & & & & & \ddots \\
& & & & & & & & & J_1
\end{pmatrix}
\quad
\begin{array}{l}
\left.\begin{array}{l} \\ \\ \\ \end{array}\right\} s_d \text{ Blöcke} \\
\left.\begin{array}{l} \\ \\ \\ \end{array}\right\} s_{d-1} \text{ Blöcke} \\
\quad\vdots \\
\left.\begin{array}{l} \\ \\ \\ \end{array}\right\} s_1 \text{ Blöcke}
\end{array}
$$

Die Zahlen $s_1, \ldots, s_d \in \mathbb{N}_0$ *sind eindeutig bestimmt, und es gilt* $s_d \neq 0$.

Beweis: Wir beweisen die Existenz einer Basis mit der behaupteten Eigenschaft durch Induktion nach $n = \dim(V)$. Der Unterraum $f(V) \subset V$ ist echt kleiner als V, da f^{d-1} auf diesem Unterraum Null ist (nicht aber auf V). Nach Induktionsvoraussetzung gilt

$$f(V) = W_1 \oplus \ldots \oplus W_r$$

mit Unterräumen W_i, die einen zyklischen Vektor w_i besitzen. Sei $m_i = \dim(W_i)$, so daß $f^{m_i-1}w_i \neq 0$, aber $f^{m_i}w_i = 0$. Wir wählen irgendein $v_i \in V$ mit $fv_i = w_i$. Dieses v_i ist dann ein zyklischer Vektor für den Raum $V_i := W_i + \langle v_i \rangle$, der somit die Dimension $m_i + 1$ besitzt (Lemma 4.5.5). Man beachte, daß $m_i \leq d - 1$ nach Induktionsvoraussetzung, also $\dim(V_i) \leq d$. Wir behaupten:

$$V' := V_1 + \ldots + V_r \qquad \text{ist eine direkte Summe,} \tag{4.10}$$

und

$$V = V' + \ker(f). \tag{4.11}$$

Daraus folgt dann die Existenzaussage, denn V' besitzt eine Jordanbasis, die man durch Elemente von $\ker(f)$ zu einer Jordanbasis von V ergänzen kann.

Beweis von (4.10): Jedes Element von V_i ist von der Form $P(f)v_i$ mit einem Polynom $P \in K[X]$. Angenommen, wir haben eine Linearkombination

$$P_1(f)v_1 + \ldots + P_r(f)v_r = 0. \tag{4.12}$$

Anwenden von f ergibt $P_1(f)w_1 + \ldots + P_r(f)w_r = 0$, und dann

$$P_i(f)w_i = 0 \qquad \text{für alle } i = 1, \ldots, r,$$

da die Summe $W_1 + \ldots + W_r$ direkt ist. Auf W_i besitzt $f|_{W_i}$ das Minimalpolynom X^{m_i}, also folgt $X^{m_i}|P_i$, und insbesondere $X|P_i$. Es ist daher $P_i = XQ_i$ für ein Polynom Q_i, und aus (4.12) ergibt sich

$$Q_1(f)w_1 + \ldots + Q_r(f)w_r = 0.$$

Wie eben folgt $Q_i(f)w_i = 0$ für alle i, dann $X^{m_i}|Q_i$, und daher $X^{m_i+1}|P_i$. Also ist $P_i(f)v_i = 0$, was zu zeigen war.

Beweis von (4.11): Jedes Element von $f(V)$ ist von der Form $P_1(f)w_1 + \ldots + P_r(f)w_r = f(P_1(f)v_1 + \ldots + P_r(f)v_r) \in f(V')$, also $f(V) = f(V')$. Zu gegebenem $v \in V$ gibt es daher ein $v' \in V'$ mit $fv = fv'$, und es folgt

$$v = v' + (v - v') \in V' + \ker(f).$$

Dies beschließt den Beweis von (4.11), und damit der Existenzaussage.

Zur Eindeutigkeit der Zahlen $s_1, \ldots, s_d$: Hält man sich vor Augen, wie die Potenzen der Jordankästchen aussehen, so erkennt man

$$\begin{aligned}
&\dim(\mathrm{im}(f^{d-1})) = s_d,\\
&\dim(\mathrm{im}(f^{d-2})) = 2s_d + s_{d-1},\\
&\quad\vdots\\
&\dim(\mathrm{im}(f)) = (d-1)s_d + (d-2)s_{d-1} + \ldots + 2s_3 + s_2.
\end{aligned}$$

Also sind $s_1, \ldots, s_d$ durch f bestimmt, und überdies muß $s_d \neq 0$ gelten, da sonst schon $f^{d-1} = 0$ wäre. ∎

4.6 Allgemeine Endomorphismen

In den beiden vorangegangenen Abschnitten haben wir *diagonalisierbare* bzw. *nilpotente* Endomorphismen untersucht und klassifiziert. Die *Jordan-Chevalley-Zerlegung* besagt, daß sich ein beliebiger Endomorphismus, dessen charakteristisches Polynom in Linearfaktoren zerfällt, als Summe eines diagonalisierbaren und eines nilpotenten Endomorphismus schreiben läßt (sogar eindeutig, wenn man noch eine Zusatzbedingung fordert). Bringt man den nilpotenten Anteil noch auf die in Satz 4.5.6 angegebene Form, so erhält man die *Jordansche Normalform* eines Endomorphismus, was das Hauptergebnis dieses Kapitels darstellt.

Hauptraumzerlegung

Sei f ein Endomorphismus des n-dimensionalen K-Vektorraumes V. Wir nehmen an, daß sein charakteristisches Polynom χ_f über K vollständig in Linearfaktoren zerfällt. Wenn $\lambda_1, \ldots, \lambda_r \in K$ die paarweise verschiedenen Nullstellen von χ_f sind (und damit die Eigenwerte von f), so gilt also

$$\chi_f = (X - \lambda_1)^{k_1} \cdot \ldots \cdot (X - \lambda_r)^{k_r} \tag{4.13}$$

mit $k_i = \mathrm{ord}_{\lambda_i}(\chi_f)$ als der *algebraischen Vielfachheit* des Eigenwertes λ_i.

Definition 4.6.1 *Der Unterraum*

$$\ker(f - \lambda_i\,\mathrm{id})^{k_i} \subset V$$

heißt der **Hauptraum** *von f zum Eigenwert λ_i.*

Man beachte, daß jeder Hauptraum durch f in sich selbst abgebildet wird, und daß der Hauptraum den Eigenraum enthält:

$$\ker(f - \lambda_i \,\mathrm{id}) \subset \ker(f - \lambda_i \,\mathrm{id})^{k_i}.$$

Satz 4.6.2 *Mit den obigen Bezeichnungen und Voraussetzungen gilt: V ist die direkte Summe aller Haupträume von f:*

$$V = \ker(f - \lambda_1 \,\mathrm{id})^{k_1} \oplus \ldots \oplus (f - \lambda_r \,\mathrm{id})^{k_r}.$$

Beweis: Wir betrachten die Polynome

$$p_i := \prod_{\substack{j=1 \\ j \neq i}}^{n} (X - \lambda_i)^{k_i}, \qquad i = 1, \ldots, r. \tag{4.14}$$

Diese besitzen keinen gemeinsamen Teiler; das von ihnen erzeugte Ideal muß also schon ganz $K[X]$ sein ($(p_1, \ldots, p_n)$ ist wie jedes Ideal von $K[X]$ ein Hauptideal, etwa gleich (p), und p muß dann eine Einheit sein). Es gibt also $q_1, \ldots, q_r \in K[X]$ mit

$$p_1 q_1 + \ldots + p_r q_r = 1.$$

Einsetzen von f und anwenden auf ein $v \in V$ ergibt

$$v = p_1(f) q_1(f) v + \ldots + p_r(f) q_r(f) v.$$

Nun gilt $p_i(f) q_i(f) v \in \ker(f - \lambda_i \,\mathrm{id})^{k_i}$, denn

$$(f - \lambda_i \,\mathrm{id})^{k_i} p_i(f) q_i(f) v = \chi_f(f) q_i(f) v = 0$$

nach dem Satz von Cayley–Hamilton. Damit haben wir

$$V = \ker(f - \lambda_1 \,\mathrm{id})^{k_1} + \ldots + \ker(f - \lambda_r \,\mathrm{id})^{k_r}$$

schon bewiesen, und es bleibt zu zeigen, daß diese Summe direkt ist. Nach Feststellung 1.5.4 genügt es dafür zu zeigen, daß für $i \neq j$ der Raum $W := \ker(f - \lambda_i \,\mathrm{id})^{k_i} \cap \ker(f - \lambda_j \,\mathrm{id})^{k_j}$ Null ist. Offenbar gilt $f(W) \subset W$, und das Minimalpolynom des Endomorphismus $f|_W \in \mathrm{End}_K(W)$ teilt sowohl $(X - \lambda_i)^{k_i}$ als auch $(X - \lambda_j)^{k_j}$. Dieses Minimalpolynom muß daher 1 sein, was wiederum $W = 0$ erzwingt. ∎

Jordan–Chevalley–Zerlegung

Satz 4.6.3 *Sei V ein endlich-dimensionaler K-Vektorraum, und $f \in \mathrm{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom über K in Linearfaktoren zerfällt. Es gibt dann eindeutig bestimmte Endomorphismen $f_d, f_n \in \mathrm{End}_K(V)$, die die folgenden Bedingungen erfüllen:*

$$f = f_d + f_n, \qquad f_d \text{ ist diagonalisierbar}, \qquad f_n \text{ ist nilpotent}, \qquad f_d f_n = f_n f_d.$$

Ferner gibt es Polynome $p, q \in K[X]$ mit $f_d = p(f)$ und $f_n = q(f)$.

Beweis: Es sei χ_f durch (4.13) gegeben, und die Polynome p_i seien wie in (4.14) definiert. Da $(X-\lambda_i)^{k_i}$ und p_i teilerfremd sind, gibt es Polynome $g_i, h_i \in K[X]$ mit

$$g_i(X-\lambda_i)^{k_i} + h_i p_i = 1.$$

Wir setzen

$$p := \lambda_1 h_1 p_1 + \ldots + \lambda_r h_r p_r, \qquad q := X - p,$$

sowie $f_d := p(f)$, $f_n := q(f)$. Dann gilt zumindest $f = f_d + f_n$ und $f_d f_n = f_n f_d$. Zur Abkürzung sei $V_i := \ker(f - \lambda_i \,\mathrm{id})^{k_i}$ der Hauptraum zu λ_i. Für jedes i gilt

$$p - \lambda_i = \sum_{j \neq i} \lambda_j h_j p_j + \lambda_i(h_i p_i - 1) = \sum_{j \neq i} \lambda_j h_j p_j - \lambda_i g_i (X-\lambda_i)^{k_i} \in ((X-\lambda_i)^{k_i}),$$

und es folgt, daß $p(f) - \lambda_i \,\mathrm{id}$ auf V_i der Nullendomorphismus ist. Mit anderen Worten,

$$f_d\Big|_{V_i} = \lambda_i \,\mathrm{id}_{V_i},$$

und da $V = V_1 \oplus \ldots \oplus V_r$ nach Satz 4.6.2, folgt die Diagonalisierbarkeit von f_d. Ferner ist

$$f_n\Big|_{V_i} = (f - f_d)\Big|_{V_i} = (f - \lambda_i \,\mathrm{id})\Big|_{V_i}$$

nilpotent auf V_i (die k_i-te Potenz verschwindet), und wieder wegen $V = V_1 \oplus \ldots \oplus V_r$ muß f_n auf ganz V nilpotent sein.

Es bleibt die Eindeutigkeitsaussage zu zeigen. Sei $f = f_d' + f_n'$ noch eine Zerlegung mit den entsprechenden Eigenschaften. Dann folgt zunächst

$$f_d - f_d' = f_n' - f_n. \tag{4.15}$$

Da f_d und f_n Polynome in f sind, kommutiert f_d mit f_d' und f_n mit f_n'. Deshalb ist mit f_d und f_d' auch $f_d - f_d'$ diagonalisierbar (Satz 4.4.5), und mit f_n und f_n' ist auch $f_n' - f_n$ nilpotent. In (4.15) steht deshalb ein Endomorphismus, der sowohl diagonalisierbar als auch nilpotent ist. Der einzige solche ist aber der Nullendomorphismus, und es folgt $f_d = f_d'$ und $f_n = f_n'$. ∎

Die Jordansche Normalform

In Gleichung (4.9) haben wir eine nilpotente Jordanmatrix J_n definiert. Allgemeiner nennt man jede Matrix der Form

$$J_n(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & \ldots & & \lambda \end{pmatrix} = \lambda \mathbf{1}_n + J_n \in M(n,K) \tag{4.16}$$

mit $\lambda \in K$ eine **Jordanmatrix**.

Satz 4.6.4 (Jordansche Normalform) *Sei V ein endlich-dimensionaler K-Vektorraum, und $f \in \mathrm{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom über K in Linearfaktoren zerfällt:*

$$\chi_f = (X-\lambda_1)^{k_1} \cdot \ldots \cdot (X-\lambda_r)^{k_r}$$

mit den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r$ von f. Dann gibt es eine Basis von V, bezüglich derer die Matrix A von f folgende Form hat:

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}, \qquad A_i \in M(k_i, K),$$

wobei jedes A_i wiederum eine aus Jordankästchen bestehende Blockmatrix ist:

$$A_i = \begin{pmatrix} J_{l_{i1}}(\lambda_i) & & \\ & \ddots & \\ & & J_{l_{is_i}}(\lambda_i) \end{pmatrix}, \qquad \textit{wobei } l_{i1} + \ldots + l_{is_i} = k_i.$$

Die Zahlen $l_{i1}, \ldots, l_{is_i}$ sind eindeutig bestimmt; jedes A_i ist also eindeutig bis auf die Reihenfolge der Jordankästchen. Die Matrix A ist eindeutig bis auf die Reihenfolge der A_i.

Beweis: Zur Abkürzung sei wieder $V_i := \ker(f - \lambda_i \,\mathrm{id})^{k_i}$ der Hauptraum zu λ_i.

Zunächst zur Eindeutigkeit. Sei $v_1, \ldots, v_n$ eine Basis der beschriebenen Art. Es ist dann leicht zu sehen, daß die ersten k_1 Basisvektoren in V_1 liegen, die nächsten k_2 in V_2, usw. Wegen Satz 4.6.2 bilden also die ersten k_1 Vektoren eine Basis von V_1, die nächsten k_2 Vektoren eine Basis von V_2, usw. Mit anderen Worten, A_i ist die Matrix von $f\big|_{V_i}$. Die Matrix des nilpotenten Endomorphismus $(f-\lambda_i \,\mathrm{id})\big|_{V_i}$ ist dann $A_i - \lambda_i \mathbf{1}_{k_i}$. Die Eindeutigkeit der Zahlen $l_{i1}, \ldots, l_{is_i}$ ergibt sich somit aus dem Klassifikationssatz 4.5.6 für nilpotente Endomorphismen.

Nun ist klar, wie man die Existenzfrage anzugehen hat. Wegen $V = V_1 \oplus \ldots \oplus V_r$ braucht man nur eine Basis von V_i zu finden, bezüglich derer $f\big|_{V_i}$ die Matrix A_i besitzt. Der Endomorphismus $(f-\lambda_i \,\mathrm{id})\big|_{V_i}$ von V_i ist nilpotent, und besitzt deshalb eine Matrix der in Satz 4.5.6 angegebenen Form. Dann hat aber bezüglich derselben Basis die Matrix von $f\big|_{V_i}$ die gewünschte Form. ■

Die in diesem Satz gemachte Voraussetzung über das charakteristische Polynom ist natürlich immer dann erfüllt, wenn der Körper K *algebraisch abgeschlossen* ist. Für solche Körper liefert Satz 4.6.4 also eine vollständige Klassifikation aller Endomorphismen eines vorgegebenen endlich-dimensionalen Vektorraumes. Dies ist insbesondere für $K = \mathbb{C}$ der Fall.

Bemerkung: Die Matrix A in Satz 4.6.4 ist offenbar die Summe einer Diagonalmatrix und einer nilpotenten Matrix, die miteinander vertauschbar sind. Dies beweist erneut einen Teil von Satz 4.6.3.

Kapitel 5

Bilinearformen

Im Kapitel über Determinanten hatten wir bereits *Multilinearformen* definiert. Multilinearformen vom Grade 2 sind besonders wichtig und heißen *Bilinearformen.* In Abschnitt 5.1 bringen wir die allgemeinen Grundlagen über Bilinearformen auf beliebigen Vektorräumen. Danach wird der Fall endlich-dimensionaler Vektorräume behandelt, in welchem sich Bilinearformen ähnlich wie Endomorphismen durch Matrizen beschreiben lassen.

Die tiefere Untersuchung von Bilinearformen auf einem Vektorraum hängt stark vom zugrunde liegenden Körper K ab. In Abschnitt 5.3 betrachten wir den Spezialfall $K = \mathbb{R}$, wo der *Sylvestersche Trägheitssatz* eine vollständige Klassifikation der symmetrischen Bilinearformen liefert. Im Falle $K = \mathbb{C}$ betrachtet man statt Bilinearformen besser *Hermitesche Formen.* Auch diese lassen sich vollständig klassifizieren, was in Abschnitt 5.4 erklärt wird. Ferner bringen wir die *Spektralsätze* für *euklidische* und *unitäre* Vektorräume.

5.1 Grundbegriffe

Definition und Beispiele

Definition 5.1.1 *Seien V und W beliebige K-Vektorräume. Eine Abbildung*

$$\beta :\ V \times W \longrightarrow K$$

heißt **bilinear** *oder eine* **Bilinearform**, *falls sie in jeder der beiden Variablen linear ist, falls also gilt:*

$$\beta(a_1v_1 + a_2v_2, w) = a_1\beta(v_1, w) + a_2\beta(v_2, w) \qquad \text{für alle } v_1, v_2 \in V,\ w \in W,\ a_1, a_2 \in K,$$

und

$$\beta(v, a_1w_1 + a_2w_2) = a_1\beta(v, w_1) + a_2\beta(v, w_2) \qquad \text{für alle } v \in V,\ w_1, w_2 \in W,\ a_1, a_2 \in K.$$

Die bilinearen Abbildungen $V \times W \to K$ bilden in natürlicher Weise einen K-Vektorraum, der mit

$$\mathrm{Bil}_K(V, W)$$

bezeichnet wird.

Oft wird eine Bilinearform einfach durch Klammern bezeichnet, etwa

$$(v,w) = \beta(v,w) \qquad \text{oder} \qquad \langle v,w\rangle = \beta(v,w).$$

Beispiele: 1. Die Abbildung $K \times K \to K$, $(x,y) \mapsto xy$ ist bilinear.

2. Ist allgemeiner $A = (a_{ij}) \in M(m \times n, K)$, so ist

$$\begin{aligned} \beta_A :\ K^m \times K^n &\longrightarrow K, \\ (v,w) &\longmapsto {}^t vAw \qquad \text{(Matrixmultiplikation)} \end{aligned} \tag{5.1}$$

eine Bilinearform (K^m und K^n sind hier Räume von *Spalten*vektoren). Explizit: Falls $A = (a_{ij})$, $v = (v_i)$, und $w = (w_i)$, so ist

$$\beta(v,w) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, v_i\, w_j.$$

3. Ist V^* der Dualraum von V, so ist

$$\begin{aligned} V \times V^* &\longrightarrow K, \\ (v,f) &\longmapsto f(v), \end{aligned}$$

bilinear.

Bemerkung 5.1.2 Ist $(v_i)_{i\in I}$ eine Basis von V und $(w_j)_{j\in J}$ eine Basis von W, so ist die Bilinearform β offenbar durch die Skalare

$$\beta(v_i, w_j), \qquad i \in I,\ j \in J,$$

festgelegt. Sind umgekehrt $c_{ij} \in K$ beliebig vorgegebene Elemente, so gibt es (genau) eine Bilinearform β mit $\beta(v_i, w_j) = c_{ij}$, nämlich

$$\beta\Big(\sum_i a_i v_i, \sum_j b_j w_j\Big) = \sum_{i,j} a_i b_j c_{ij} \qquad (a_i, b_j \in K,\ \text{nur endlich viele } \neq 0).$$

Dies ist das bilineare Analogon zu Feststellung 1.4.9.

Definition 5.1.3 *Sei V ein K-Vektorraum. Ist $\beta :\ V \times V \longrightarrow K$ eine bilineare Abbildung, so spricht man von einer* **Bilinearform auf V**. *Diese heißt* **symmetrisch**, *falls*

$$\beta(v_1, v_2) = \beta(v_2, v_1) \qquad \textit{für alle } v_1, v_2 \in V,$$

und **antisymmetrisch** *oder* **schiefsymmetrisch**, *falls*

$$\beta(v_1, v_2) = -\beta(v_2, v_1) \qquad \textit{für alle } v_1, v_2 \in V.$$

Sie heißt **symplektisch**, *falls*

$$\beta(v,v) = 0 \qquad \textit{für alle } v \in V.$$

Eine Bilinearform ist nichts anderes als eine 2-Form im Sinne von Definition 3.3.1. Die *symplektischen* Bilinearformen sind genau die *alternierenden* 2-Formen. Sie bilden, ebenso wie die symmetrischen bzw. antisymmetrischen Bilinearformen, einen Untervektorraum von $\mathrm{Bil}_K(V)$.

Wenn die Charakteristik von K nicht 2 ist, so gibt es keinen Unterschied zwischen schiefsymmetrischen und symplektischen Bilinearformen (vgl. Feststellung 3.3.2). In diesem Fall läßt sich jede Bilinearform wegen

$$\beta(v,w) = \frac{\beta(v,w)+\beta(w,v)}{2} + \frac{\beta(v,w)-\beta(w,v)}{2} \tag{5.2}$$

als Summe einer symmetrischen und einer antisymmetrischen Bilinearform schreiben. Es gilt sogar: $\mathrm{Bil}_K(V)$ ist die direkte Summe des Raums der symmetrischen und der antisymmetrischen Bilinearformen. Ist dagegen $\mathrm{char}(K) = 2$, also $1 = -1$, so ist eine symmetrische Bilinearform das gleiche wie eine antisymmetrische.

Beispiele: 1. Sei $A \in M(n,K)$. Die Bilinearform

$$\begin{aligned} K^n \times K^n &\longrightarrow K, \\ (v,w) &\longmapsto {}^t v A w, \end{aligned}$$

ist genau dann (anti-) symmetrisch, wenn A (anti-) symmetrisch ist (d.h., wenn ${}^tA = \pm A$), denn

$${}^t vAw = {}^t({}^t vAw) = {}^t w\, {}^t A v.$$

2. Ist im vorherigen Beispiel $K = \mathbb{R}$ und $A = \mathbf{1}$, so erhält man das übliche **Skalarprodukt** $\langle \cdot\,, \cdot \rangle$ auf dem $\mathbb{R}^n$. Dieses kann in bekannter Weise zur Winkelmessung dienen.

3. Sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall. Die Abbildung

$$(f,g) \longmapsto \int_a^b f(x)g(x)\,dx$$

definiert eine symmetrische Bilinearform auf dem Raum $C^0([a,b])$ der stetigen, reellwertigen Funktionen auf $[a,b]$.

Orthogonalität

Definition 5.1.4 *Sei $\beta : V \times W \to K$ bilinear. Zwei Vektoren $v \in V$ und $w \in W$ heißen* **orthogonal** *(bezüglich β), falls $\beta(v,w) = 0$. Für eine Teilmenge $S \subset V$ bzw. $T \subset W$ heißt*

$$S^\perp = \{w \in W : \beta(v,w) = 0 \text{ für alle } v \in S\}$$

bzw.

$${}^\perp T = \{v \in V : \beta(v,w) = 0 \text{ für alle } w \in T\}$$

das **orthogonale Komplement** *von S bzw. T.*

Feststellung 5.1.5 *Mit obigen Bezeichnungen gilt: $S^\perp$ ist ein Untervektorraum von W und ${}^\perp T$ ist ein Untervektorraum von V. Ferner gilt*

$$S \subset {}^\perp(S^\perp), \qquad T \subset ({}^\perp T)^\perp.$$

Beweis: Leichte Übung. ∎

Beispiele: 1. Ist wie oben $\langle\,\cdot\,,\,\cdot\,\rangle$ das übliche Skalarprodukt auf dem $\mathbb{R}^3$, so gilt

$$\langle v, w\rangle = 0 \qquad \Longleftrightarrow \qquad \text{Die Vektoren } v \text{ und } w \text{ stehen senkrecht aufeinander.}$$

Dies erklärt die Bezeichnung „orthogonal".

2. Den Raum $C^0([-\pi,\pi])$ versehen wir wie oben mit der Bilinearform

$$(f,g) = \int_{-\pi}^{\pi} f(x)g(x)\,dx.$$

Die Funktionen f_n, definiert durch

$$f_0(x) = \frac{1}{\sqrt{2\pi}}, \qquad f_{2n-1}(x) = \frac{\cos(nx)}{\sqrt{\pi}}, \qquad f_{2n}(x) = \frac{\sin(ns)}{\sqrt{\pi}}, \qquad n \in \mathbb{N},$$

sind paarweise orthogonal:

$$(f_n, f_m) = 0 \qquad \text{für } n \neq m.$$

Überdies gilt $(f_n, f_n) = 1$; man spricht daher von einem **Orthonormalsystem**. Dieses Orthonormalsystem ist die Grundlage der Theorie der *Fourierreihen*.

Jede bilineare Abbildung $\beta : V \times W \to K$ definiert zwei *lineare* Abbildungen, nämlich

$$\begin{aligned} V &\longrightarrow W^*, \\ v &\longmapsto (w \mapsto \beta(v,w)), \end{aligned} \tag{5.3}$$

und

$$\begin{aligned} W &\longrightarrow V^*, \\ w &\longmapsto (v \mapsto \beta(v,w)). \end{aligned} \tag{5.4}$$

Kennt man eine der Abbildungen (5.4) oder (5.3), so läßt sich daraus β rekonstruieren.

Feststellung und Definition 5.1.6 *Eine bilineare Abbildung $\beta : V \times W \to K$ heißt* **nicht ausgeartet in der ersten Variablen**, *falls die folgenden äquivalenten Bedingungen erfüllt sind:*

i) $\beta(v,w) = 0$ für alle $w \in W \quad \Longrightarrow \quad v = 0$.

ii) Die Abbildung (5.3) ist injektiv.

iii) $^{\perp}W = 0$.

β heißt **nicht ausgeartet in der zweiten Variablen**, *falls die folgenden äquivalenten Bedingungen erfüllt sind:*

i) $\beta(v, w) = 0$ *für alle* $v \in V \quad \Longrightarrow \quad w = 0$.

ii) Die Abbildung (5.4) ist injektiv.

iii) $V^{\perp} = 0$.

Beweis: Die Äquivalenz der Bedingungen ist weitgehend klar. ∎

Beispiele: 1. Die Bilinearform in Beispiel 3 auf Seite 108 ist in beiden Variablen nicht ausgeartet.

2. Die Bilinearform

$$K^n \times K^{n+1} \longrightarrow K,$$
$$((x_i),\ (y_j)) \longmapsto \sum_{i=1}^{n} x_i y_i,$$

ist nicht ausgeartet in der ersten, aber ausgeartet in der zweiten Variablen.

Im allgemeinen ist der *Kern* der Abbildung (5.3) (bzw. (5.4)) gerade $^{\perp}W$ (bzw. $V^{\perp}$). Ist β nicht ausgeartet in der ersten (bzw. zweiten) Variablen, so kann man also V (bzw. W) mit einem Unterraum von W^* (bzw. V^*) identifizieren. Man prüft leicht nach, daß β in natürlicher Weise eine *in beiden Variablen nicht ausgeartete* bilineare Abbildung

$$\bar{\beta} :\ V/{}^{\perp}W \ \times\ W/V^{\perp} \longrightarrow K$$

induziert.

Isometrien

Definition 5.1.7 *Sei β eine Bilinearform auf V. Eine* **Isometrie**[1] *oder eine* **orthogonale Abbildung** *auf V ist eine bijektive lineare Abbildung $f :\ V \to V$ mit*

$$\beta(fv, fw) = \beta(v, w) \qquad \textit{für alle } v, w \in V.$$

Die Menge aller dieser Abbildungen wird mit $\mathrm{O}(V, \beta)$ *oder einfach* $\mathrm{O}(V)$ *bezeichnet, und heißt die* **orthogonale Gruppe** *von* (V, β).

Orthogonale Abbildungen sind also diejenigen Bijektionen, die sowohl die lineare Struktur des Vektorraumes als auch die auf ihm definierte Bilinearform invariant lassen. Man mache sich klar, daß $\mathrm{O}(V)$ eine *Untergruppe* von $\mathrm{GL}(V)$ ist.

Beispiele: 1. Sei $V = K$ und $\beta(x, y) = xy$. Ein $a \in K^* = \mathrm{GL}(V)$ ist genau dann eine Isometrie, wenn

$$xy = (ax)(ay) = a^2xy \qquad \text{für alle } x, y \in K$$

[1] gr. *isometria*: gleiches Maß

ist. Dies ist gleichbedeutend mit $a^2 = 1$, und folglich ist $\mathrm{O}(V) = \{1, -1\}$ (für $\mathrm{char}(K) = 2$ ist dies die triviale Gruppe).

2. Ist $V = \mathbb{R}^n$ mit dem kanonischen Skalarprodukt, so schreibt man einfach $O(n)$ für $\mathrm{O}(V)$. Dies sind also alle bijektiven, linearen, winkeltreuen Abbildungen $\mathbb{R}^n \to \mathbb{R}^n$. Für $n = 2$ fallen hierunter z.B. die *Drehungen*

$$r(\vartheta) = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

(siehe Beispiel 4 auf Seite 92) sowie die *Spiegelung*

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

Man rechnet leicht nach, daß dies tatsächlich orthogonale Abbildungen im Sinne von Definition 5.1.7 sind.

5.2 Bilinearformen auf endlich-dimensionalen Räumen

Im endlich-dimensionalen Fall lassen sich Bilinearformen (ähnlich wie Endomorphismen) nach Wahl einer Basis durch Matrizen beschreiben. In diesem Abschnitt sei V ein n-dimensionaler K-Vektorraum, versehen mit einer Bilinearform

$$\beta : V \times V \longrightarrow K.$$

Die Matrix einer Bilinearform

Definition und Feststellung 5.2.1 *Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Die* **Strukturmatrix** *oder einfach* **Matrix** *der Bilinearform β bezüglich $\mathcal{B}$ ist*

$$B = (\beta(v_i, v_j))_{i,j} \in M(n, K).$$

B ist durch die Beziehung

$$\beta(\varphi_{\mathcal{B}}(x), \varphi_{\mathcal{B}}(y)) = {}^t xBy \qquad \text{für alle } x, y \in K^n \tag{5.5}$$

charakterisiert, also durch die Kommutativität des Diagramms

$$\begin{array}{ccc} V \times V & \xrightarrow{\beta} & K \\ {\scriptstyle \varphi_{\mathcal{B}} \times \varphi_{\mathcal{B}}} \uparrow \wr & & \uparrow {\scriptstyle \mathrm{id}} \\ K^n \times K^n & \xrightarrow{\beta_B} & K. \end{array}$$

Hierbei ist $\varphi_{\mathcal{B}}$ wie in (2.12) der durch $\mathcal{B}$ definierte Koordinatenisomorphismus, und β_B ist die Bilinearform (5.1).

Beweis: Die Gültigkeit von (5.5) folgt sofort aus den Definitionen. Ferner kann es nur *eine* Matrix B mit (5.5) geben, denn aus ${}^t xBy = {}^t xB'y$ für alle x, y folgt $B = B'$ (man setze die kanonischen Basisvektoren ein). ∎

Feststellung 5.2.2 (Matrixbeschreibung bilinearer Abbildungen) *Sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V.*

i) Ordnet man jeder Bilinearform auf V ihre Matrix bezüglich $\mathcal{B}$ zu, so erhält man einen Isomorphismus

$$\mathrm{Bil}_K(V) \xrightarrow{\sim} M(n, K).$$

ii) Insbesondere ist $\dim(\mathrm{Bil}_K(V)) = n^2$.

iii) Ist $\mathcal{B}'$ eine weitere Basis von V, ist S die Übergangsmatrix von $\mathcal{B}$ zu $\mathcal{B}'$ (siehe (2.17)), und ist B bzw. B' die Matrix von β bezüglich $\mathcal{B}$ bzw. $\mathcal{B}'$, so gilt

$$B = {}^t SB'S.$$

iv) Die Determinanten von B und B' unterscheiden sich nur um ein Quadrat aus K^*:

$$\det(B) = x^2 \det(B') \qquad \textit{mit einem } x \in K^*.$$

Der Rang von B ist gleich dem von B'.

Beweis: i) folgt sofort aus Bemerkung 5.1.2, und ii) aus i). Gemäß (5.5) ist B' durch

$$\beta(\varphi_{\mathcal{B}'}(x), \varphi_{\mathcal{B}'}(y)) = {}^t xB'y \qquad \text{für alle } x, y \in K^n$$

charakterisiert. Setzt man Sx und Sy anstatt x und y ein, so folgt

$$\beta(\varphi_{\mathcal{B}'}(Sx), \varphi_{\mathcal{B}'}(Sy)) = {}^t(Sx)B'(Sy) \qquad \text{für alle } x, y \in K^n,$$

oder also

$$\beta(\varphi_{\mathcal{B}}(x), \varphi_{\mathcal{B}}(y)) = {}^t x({}^t SB'S)y \qquad \text{für alle } x, y \in K^n.$$

Daraus folgt $B = {}^t SB'S$. Aus dieser Beziehung ergeben sich die Ausagen in iv), da S invertierbar ist und $\det(S) = \det({}^t S)$. ∎

Bemerkung 5.2.3 Wir halten ausdrücklich das unterschiedliche Transformationsverhalten für Matrizen von *Endomorphismen* (Folgerung 2.3.13) bzw. *Bilinearformen* (Feststellung 5.2.2) fest:

$$\begin{aligned} A &= S^{-1}A'S && \text{für Endomorphismen,} \\ B &= {}^t SB'S && \text{für Bilinearformen.} \end{aligned}$$

Nicht ausgeartete Bilinearformen

Feststellung 5.2.4 *Für die Bilinearform β auf dem endlich-dimensionalen Raum V sind äquivalent:*

i) β ist nicht ausgeartet in der ersten Variablen.

ii) β ist nicht ausgeartet in der zweiten Variablen.

iii) B ist invertierbar, wobei B die Matrix von β bezüglich einer beliebigen Basis ist.

Beweis: Wir zeigen die Äquivalenz von i) und iii); die Äquivalenz von ii) und iii) sieht man genauso. Wegen 5.2.1 müssen wir nur zeigen:

$$B \text{ ist invertierbar} \qquad \Longleftrightarrow \qquad \beta_B \text{ ist nicht ausgeartet in der ersten Variablen.}$$

Hier ist wie oben $\beta_B(x,y) = {}^t xBy$ für $x, y \in K^n$. Sei zunächst diese Bilinearform nicht ausgeartet in der ersten Variablen, d.h.,

$${}^t xBy = 0 \text{ für alle } y \qquad \Longrightarrow \qquad x = 0.$$

Wäre aber B singulär, so gäbe es ein $x \neq 0$ mit ${}^t xB = 0$, im Widerspruch zu dieser Aussage.

Sei umgekehrt B invertierbar und ${}^t xBy = 0$ für alle y. Dann muß ${}^t xB$ der Nullvektor sein, und aus der Invertierbarkeit von B folgt $x = 0$. ∎

Aufgrund dieser Feststellung sprechen wir bei endlich-dimensionalen Räumen einfach von **nicht ausgearteten** Bilinearformen. Im nicht ausgearteten Fall sind die Abbildungen (5.3) und (5.4), also

$$V \longrightarrow V^*, \qquad\qquad V \longrightarrow V^*, \tag{5.6}$$
$$v \longmapsto (v' \mapsto \beta(v,v')), \qquad\qquad v \longmapsto (v' \mapsto \beta(v',v)), \tag{5.7}$$

aus Dimensionsgründen Isomorphismen.

Feststellung 5.2.5 *Sei β eine nicht ausgeartete Bilinearform auf V, und $W \subset V$ ein Teilraum. Dann gelten die Dimensionsformeln*

$$\dim(W) + \dim(W^\perp) = \dim(V) = \dim(W) + \dim({}^\perp W). \tag{5.8}$$

Außerdem gilt

$${}^\perp(W^\perp) = W = ({}^\perp W)^\perp. \tag{5.9}$$

Beweis: Hinter den zweiten Isomorphismus in (5.6) schalten wir die (surjektive) Einschränkungsabbildung $V^* \to W^*$. Die entstehende lineare Abbildung hat den Kern $W^\perp$. Wir haben also eine exakte Sequenz

$$0 \longrightarrow W^\perp \longrightarrow V \longrightarrow W^* \longrightarrow 0,$$

so daß $\dim(W) + \dim(W^\perp) = \dim(V)$ aus 2.1.1 i) folgt. Ebenso beweist man die zweite Dimensionsformel. Aus diesen Formeln folgt

$$\dim({}^\perp(W^\perp)) = \dim(({}^\perp W)^\perp) = \dim(W).$$

Da offenbar $W \subset {}^\perp(W^\perp)$ und $W \subset ({}^\perp W)^\perp$, ergibt sich daraus (5.9). ∎

Ist β eine Bilinearform auf V, und $W \subset V$ ein Teilraum, so läßt sich β zu einer Bilinearform $\beta|_W$ auf W einschränken. Der Teilraum W heißt **nicht ausgeartet**, falls diese Bilinearform nicht ausgeartet ist.

Folgerung 5.2.6 *Sei β eine nicht-ausgeartete Bilinearform auf V, und $W \subset V$ ein Teilraum. Genau dann gilt*

$$V = W \oplus W^\perp,$$

wenn W nicht ausgeartet ist.

Beweis: In jedem Fall gilt $\dim(W) + \dim(W^\perp) = \dim(V)$ nach 5.2.5. Genau dann ist also $V = W \oplus W^\perp$, wenn $W \cap W^\perp = 0$. Letzteres ist aber offenbar äquivalent dazu, daß $\beta|_W$ nicht ausgeartet ist. ∎

Die Adjungierte

Feststellung und Definition 5.2.7 *Sei β eine nicht ausgeartete Bilinearform auf dem endlich-dimensionalen Raum V. Zu jedem Endomorphismus f von V gibt es genau einen Endomorphismus f^* von V, genannt die* **Rechts-Adjungierte** *von V, mit der Eigenschaft*

$$\beta(fv, w) = \beta(v, f^*w) \qquad \text{für alle } v, w \in V. \tag{5.10}$$

Die Zuordnung $\mathrm{End}_K(V) \to \mathrm{End}_K(V)$, $f \mapsto f^*$, besitzt folgende Eigenschaften:

$$(f+g)^* = f^* + g^*, \qquad (af)^* = af^*, \qquad (fg)^* = g^*f^*.$$

Ähnlich gibt es zu jedem $f \in \mathrm{End}_K(V)$ genau eine **Links-Adjungierte** *${}^*f \in \mathrm{End}_K(V)$ mit*

$$\beta(v, fw) = \beta({}^*fv, w) \qquad \text{für alle } v, w \in V, \tag{5.11}$$

und es gilt

$${}^*(f+g) = {}^*f + {}^*g, \qquad {}^*(af) = a\,{}^*f, \qquad {}^*(fg) = {}^*g\,{}^*f.$$

Es besteht der Zusammenhang

$${}^*(f^*) = f = ({}^*f)^*.$$

Ist β symmetrisch, so ist ${}^*f = f^*$, und wir sprechen einfach von der **Adjungierten**.

Beweis: Wir beweisen nur die Aussagen über die Rechts-Adjungierte. Bei festem $w \in V$ ist die Abbildung

$$v \longmapsto \beta(fv, w)$$

eine Linearform auf V. Da die erste Abbildung in (5.6) ein Isomorphismus ist, gibt es ein eindeutig bestimmtes Element f^*w von V mit (5.10). Man prüft leicht nach, daß die so definierte Abbildung

$f^* : V \to V$ linear ist. Dies zeigt bereits die Existenz und Eindeutigkeit der Rechts-Adjungierten. Der Endomorphismus $f^* + g^*$ besitzt die definierende Eigenschaft der Rechts-Adjungierten von $f + g$, muß aufgrund der Eindeutigkeit also gleich $(f+g)^*$ sein. Ebenso leicht sieht man die übrigen Eigenschaften.■

Die Abbildung $f \mapsto f^*$ ist also ein Endomorphismus von $\mathrm{End}_K(V)$, und ein *Anti-Ringhomomorphismus* wegen $(fg)^* = g^* f^*$. Sie ist also ein *Anti-K-Algebrenhomomorphismus.* Sie ist bijektiv mit der Inversen $f \mapsto {}^*f$. Im symmetrischen Fall ist sie ihre eigene Inverse.

Isometrien und klassische Gruppen

Feststellung 5.2.8 *Sei $\mathcal{B}$ eine Basis von V, und B die Matrix von β bezüglich $\mathcal{B}$. Sei $f \in \mathrm{GL}(V)$ und A die Matrix von f bezüglich der gleichen Basis $\mathcal{B}$. Dann gilt:*

$$f \in \mathrm{O}(V, \beta) \qquad \Longleftrightarrow \qquad {}^tABA = B.$$

Beweis: $f \in \mathrm{O}(V)$ bedeutet nach Definition

$$\beta(fv, fw) = \beta(v, w) \qquad \text{für alle } v, w \in V.$$

Mit dem Koordinatenisomorphismus $\varphi = \varphi_{\mathcal{B}}$ ist dies äquivalent zu

$$\beta(f(\varphi(x)), f(\varphi(y))) = \beta(\varphi(x), \varphi(y)) \qquad \text{für alle } x, y \in K^n.$$

Es ist $f(\varphi(x)) = \varphi(Ax)$ nach Definition von A. Wegen (5.5) ist die letzte Gleichung also äquivalent mit

$${}^t(Ax)B(Ay) = {}^txBy \qquad \text{für alle } x, y \in K^n.$$

Dies ist aber wiederum äquivalent mit ${}^tABA = B$. ■

Man nennt

$$\mathrm{O}(n, K) = \{A \in \mathrm{GL}(n, K) : {}^tAA = \mathbf{1}_n\}$$

die **orthogonale Gruppe** vom Grade n über K, und

$$\mathrm{SO}(n, K) = \mathrm{O}(n, K) \cap \mathrm{SL}(n, K)$$

die **spezielle orthogonale Gruppe**. Mit $J = \begin{pmatrix} 0 & \mathbf{1}_n \\ -\mathbf{1}_n & 0 \end{pmatrix}$ heißt

$$\mathrm{Sp}(2n, K) = \{A \in \mathrm{GL}(2n, K) : {}^tAJA = J\}$$

die **symplektische Gruppe** vom Grade n (oft wird $\mathrm{Sp}(n)$ statt $\mathrm{Sp}(2n)$ geschrieben). Es ist also $O(n, K)$ (bzw. $\mathrm{Sp}(2n, K)$) die Gruppe aller Isometrien des Raumes K^n, versehen mit der symmetrischen (bzw. schiefsymmetrischen) Bilinearform $\beta_{\mathbf{1}}$ (bzw. β_J). Manchmal werden andere Bilinearformen zur Definition dieser Gruppen verwendet, z.B.

$$J = \begin{pmatrix} & & & & & 1 \\ & & & & \iddots & \\ & & & 1 & & \\ & & -1 & & & \\ & \iddots & & & & \\ -1 & & & & & \end{pmatrix}$$

zur Definition von $\mathrm{Sp}(2n, K)$. Alle diese Varianten liefern jedoch isomorphe Gruppen. Die Gruppen

$$\mathrm{GL}(n, K), \qquad \mathrm{SL}(n, K), \qquad \mathrm{SO}(n, K), \qquad \mathrm{Sp}(2n, K),$$

sowie einige andere heißen auch **klassische Gruppen**.

Symplektische Bilinearformen

Wir klassifizieren jetzt alle *symplektischen* Bilinearformen auf dem n-dimensionalen Raum V, d.h., alle Bilinearformen β auf V mit

$$\beta(v, v) = 0 \qquad \text{für alle } v \in V.$$

Jede solche Bilinearform ist schiefsymmetrisch ($\beta(v, w) = -\beta(w, v)$), und ist $\mathrm{char}(K) \neq 2$, so gilt, wie nach Definition 5.1.3 erwähnt, auch die Umkehrung. Die Klassifikation der symplektischen Formen ist wesentlich einfacher als die der symmetrischen Bilinearformen, die anschließend behandelt werden.

Satz 5.2.9 (Klassifikation symplektischer Bilinearformen) *Sei β eine symplektische Bilinearform auf dem n-dimensionalen K-Vektorraum V. Dann gibt es eine Basis von V, so daß die Matrix von β durch*

$$B = \begin{pmatrix} H & & & \\ & \ddots & & \\ & & H & \\ & & & \mathbf{0} \end{pmatrix} \qquad \textit{mit } H = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{5.12}$$

gegeben ist. Die Anzahl der H-Blöcke sowie die Größe des Nullblocks $\mathbf{0}$ *sind durch β bestimmt.*

Beweis: Die letzten Aussagen ergeben sich daraus, daß der Rang von B bestimmt ist (Feststellung 5.2.2 iv)), so daß wir nur die Existenz zeigen müssen. Sei dazu $V^{\perp} \neq V$, denn sonst ist $\beta = 0$ und wir sind fertig. Zu einem beliebig gewählten Vektor $u \in V$, der nicht in $V^{\perp}$ liegt, gibt es ein $v \in V$ mit

$$\beta(u, v) = 1.$$

Dies bedeutet, daß die Matrix der Einschränkung von β auf den zweidimensionalen Raum $W := \langle u, v\rangle$ bzgl. der Basis u, v durch H gegeben ist. Nun gilt

$$V = W \oplus W^{\perp},$$

denn einerseits gilt $\dim(W) + \dim(W^{\perp}) = \dim(V)$ nach 5.2.5, und andererseits ist $W \cap W^{\perp} = 0$, wie man leicht sieht. Somit ergibt sich die Behauptung durch Induktion nach $\dim(V)$. ∎

Folgerung 5.2.10 *Ist β eine nicht-ausgeartete symplektische Bilinearform auf V, so ist die Dimension von V gerade. Die Determinante der Matrix von β bezüglich einer beliebigen Basis ist ein Quadrat in K.*

Beweis: Die zweite Behauptung ergibt sich aus 5.2.2 iv), da die Determinante der Matrix (5.12) entweder 0 oder 1 ist. ∎

Sei speziell $V = K^n$ und $\mathrm{char}(K) \neq 2$. Für jede schiefsymmetrische Matrix $A \in M(n, K)$ gilt dann: Ist n ungerade, so ist A singulär. Die Determinante $\det(A)$ ist ein Quadrat in K.

Symmetrische Bilinearformen

Wir setzen jetzt

$$\mathrm{char}(K) \neq 2 \tag{5.13}$$

voraus.

Feststellung und Definition 5.2.11 *Sei β eine symmetrische Bilinearform auf V. Dann gibt es eine Basis $v_1, \ldots, v_n$ von V mit*

$$\beta(v_i, v_j) = 0 \qquad \text{für alle } i \neq j.$$

Jede solche Basis heißt eine **Orthogonalbasis** *von V bzgl. β. Die Matrix von β bzgl. $v_1, \ldots, v_n$ ist also eine Diagonalmatrix mit den Diagonalkoeffizienten $\beta(v_i, v_i)$.*

Beweis: Wir führen Induktion nach $n = \dim(V)$, wobei der Fall $n = 1$ trivial ist. Falls $\beta(v, v) = 0$ für alle $v \in V$, so ist β sowohl symmetrisch als auch antisymmetrisch, und damit 0 (man beachte die Voraussetzung (5.13)). In diesem Fall ist nichts zu beweisen. Sei andererseits ein Vektor $v \in V$ mit $\beta(v, v) \neq 0$ vorhanden. Dann ist

$$V = \langle v \rangle \oplus \langle v \rangle^{\perp}$$

nach 5.2.6. Auf den $(n-1)$-dimensionalen Raum $\langle v \rangle^{\perp}$ kann man jetzt die Induktionsvoraussetzung anwenden. ∎

Für Matrizen formuliert besagt diese Feststellung: Zu jeder symmetrischen Matrix $A \in M(n, K)$ gibt es ein $S \in \mathrm{GL}(n, K)$, so daß

$${}^tSAS$$

eine Diagonalmatrix ist (siehe 5.2.2 iii)).

Ist eine Orthogonalbasis $v_1, \ldots, v_n$ gegeben, so ist die Nicht-Ausgeartetheit von β sehr leicht zu entscheiden: Nach Feststellung 5.2.4 ist β genau dann nicht ausgeartet, wenn die Diagonalkoeffizienten

$$\beta(v_i, v_i), \qquad i = 1, \ldots, n,$$

sämtlich ungleich Null sind.

Die Matrix einer symmetrischen Bilinearform läßt sich also auf Diagonalgestalt bringen. Die Klassifikation sämtlicher solcher Bilinearformen läuft damit auf folgende Frage hinaus: Wann gibt es für zwei gegebene Diagonalmatrizen $A = \mathrm{diag}(a_1, \ldots, a_n)$ und $B = \mathrm{diag}(b_1, \ldots, b_n)$ ein $S \in \mathrm{GL}(n, K)$, so daß

$${}^tSAS = B$$

gilt? Dies ist ein im allgemeinen schwieriges Problem, das stark vom Grundkörper K abhängt. Wir machen zumindest folgende Beobachtung: Unterscheidet sich a_i von b_i nur um ein Quadrat, gilt also

$$b_i = x_i^2 a_i \qquad \text{mit einem } x_i \in K^*$$

für alle i, so gibt es ein geeignetes S, nämlich $S = \mathrm{diag}(x_1, \ldots, x_n)$. Kurz gesagt: Die Diagonalkoeffizienten in der Matrix einer symmetrischen Bilinearform bezüglich einer Orthogonalbasis lassen sich um beliebige Quadrate in K^* abändern. Für Körper, in denen jedes Element von K^* ein Quadrat ist, lassen sich die Diagonalkoeffizienten also *beliebig* abändern. Insbesondere kann man die Einheitsmatrix herstellen. Körper, für die dies der Fall ist, sind z.B. die algebraisch abgeschlossenen Körper, denn in diesen besitzt das Polynom $X^2 - a$ für jedes $a \in K^*$ eine Nullstelle. Wir halten dies fest:

Feststellung 5.2.12 *Sei $K^{*2} = K^*$, d.h., jedes Element von K^* sei ein Quadrat. Jede nicht ausgeartete symmetrische Bilinearform besitzt dann bzgl. einer geeigneten Basis die Einheitsmatrix als Strukturmatrix. Die Bedingung an K ist für alle algebraisch abgeschlossenen Körper erfüllt, insbesondere für $K = \mathbb{C}$.*

Für $K = \mathbb{R}$ läßt sich ebenfalls eine vollständige Klassifikation erreichen; siehe Abschnitt 5.3.

Wir halten noch fest, daß sich auch die Reihenfolge der Diagonalkoeffizienten beliebig ändern läßt; ist nämlich S eine Permutationsmatrix, so bewirkt die Operation ${}^tS\,\mathrm{diag}(a_1, \ldots, a_n)S$ eine Permutation der $a_1, \ldots, a_n$ (dies entspricht einfach einer Umnumerierung der Basisvektoren).

Quadratische Formen

Definition 5.2.13 *Sei V ein K-Vektorraum. Eine Abbildung*

$$q : V \longrightarrow K$$

heißt **quadratische Form** *auf V, falls gilt:*

i) $q(av) = a^2 q(v)$ *für alle* $a \in K$ *und* $v \in V$.

ii) Die Abbildung

$$(v, w) \longmapsto q(v + w) - q(v) - q(w) \tag{5.14}$$

ist eine Bilinearform auf V.

Offenbar bilden die quadratischen Formen auf V einen K-Vektorraum.

Wir setzen im folgenden wieder

$$\mathrm{char}(K) \neq 2$$

voraus. Wenn β die Bilinearform in (5.14) bezeichnet, so gilt

$$q(v) = \frac{1}{2}\beta(v, v) \tag{5.15}$$

(man setze $w = v$ und benutze i)). Es läßt sich also q aus β wieder zurückgewinnen. Außerdem ist β offenbar symmetrisch. Ist umgekehrt eine symmetrische Bilinearform β gegeben, und wird die Funktion q durch (5.15) definiert, so ist q eine quadratische Form auf V. Unter der Voraussetzung $\mathrm{char}(K) \neq 2$

gibt es also eine *Bijektion* zwischen quadratischen Formen und symmetrischen Bilinearformen auf V (genauer: einen *Isomorphismus* zwischen den Räumen dieser Objekte).

Insbesondere lassen sich quadratische Formen nach Wahl einer Basis durch Matrizen beschreiben, nämlich die Strukturmatrizen der zugehörigen Bilinearformen. So beschreibt etwa die Einheitsmatrix auf dem K^n (mit der kanonischen Basis) die quadratische Form

$$q(x_1, \dots, x_n) = x_1^2 + \ldots + x_n^2.$$

Die Matrix $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ beschreibt die quadratische Form

$$q(x, y) = ax^2 + 2bxy + y^2$$

auf K^2.

5.3 Bilinearformen über $\mathbb{R}$

Wir betrachten jetzt speziell Bilinearformen auf endlich-dimensionalen reellen Vektorräumen. Im ganzen Abschnitt ist also $K = \mathbb{R}$ der zugrundeliegende Körper. Wegen $\mathrm{char}(\mathbb{R}) = 0$ läßt sich jede Bilinearform auf dem $\mathbb{R}$-Vektorraum V als Summe einer symmetrischen und einer antisymmetrischen Bilinearform schreiben, vgl. (5.2). Über letztere ist nach Satz 5.2.9 nicht mehr allzuviel zu sagen. Wir betrachten daher durchweg nur *symmetrische* Bilinearformen.

Definitheit

Definition 5.3.1 *Sei β eine symmetrische Bilinearform auf dem $\mathbb{R}$-Vektorraum V. Sie heißt*

- **positiv definit**, *falls $\beta(v, v) > 0$ für alle $v \in V$, $v \neq 0$.*
- **positiv semidefinit**, *falls $\beta(v, v) \geq 0$ für alle $v \in V$.*
- **negativ definit**, *falls $\beta(v, v) < 0$ für alle $v \in V$, $v \neq 0$.*
- **negativ semidefinit**, *falls $\beta(v, v) \leq 0$ für alle $v \in V$.*

Ein Unterraum W von V heißt **positiv definit** *(bzw. positiv semidefinit, usw.), falls die Einschränkung von β auf W es ist.*

Entsprechende Definitionen gelten für Matrizen, wenn man sie mit den durch sie definierten Bilinearformen auf dem $\mathbb{R}^n$ identifizert: Z.B. heißt eine symmetrische Matrix $B \in M(n, \mathbb{R})$ **positiv definit**, falls

$${}^t vBv > 0 \qquad \text{für alle } v \in \mathbb{R}^n,\ v \neq 0.$$

Offenbar ist β genau dann positiv definit (positiv semidefinit, ...), wenn die Matrix von β bezüglich irgendeiner Basis es ist.

Ist eine Orthogonalbasis für β gegeben, so läßt sich die Definitheit sofort ablesen:

Feststellung 5.3.2 *Seien $a_1, \dots, a_n$ die Diagonalkoeffizienten in der Matrix von β bezüglich einer Orthogonalbasis. Genau dann ist β*

- *positiv definit, wenn $a_i > 0$ für alle i.*
- *positiv semidefinit, wenn $a_i \geq 0$ für alle i.*
- *negativ definit, wenn $a_i < 0$ für alle i.*
- *negativ semidefinit, wenn $a_i \leq 0$ für alle i.*

Beweis: Dies ist leicht zu sehen. ∎

Folgerung 5.3.3 *Für die Matrix B einer positiv definiten Bilinearform gilt $\det(B) > 0$. Insbesondere sind positiv und negativ definite Bilinearformen nicht ausgeartet.*

Beweis: Wegen 5.3.2 ist die erste Aussage richtig, falls B eine Diagonalmatrix ist. Der allgemeine Fall folgt dann aus Feststellung 5.2.2 iv). Man beachte nun noch, daß β dann und nur dann positiv definit ist, wenn $-\beta$ negativ definit ist. ∎

Allgemeiner als diese Aussage ist das folgende **Determinantenkriterium für positive Definitheit**. Es wird folgende Notation verwendet: Ist $A \in M(n, \mathbb{R})$, und S eine Teilmenge von $\{1, \dots, n\}$, so bezeichnet A_S die Matrix, die aus A durch Streichen aller Zeilen und Spalten entsteht, deren Indizes *nicht* in S sind. Z.B. ist

$$A_{\{1,\dots,k\}} = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}.$$

Feststellung 5.3.4 *Für $A \in M(n, R)$ sind folgende Aussagen äquivalent:*

i) A ist positiv definit.

ii) Für jedes $S \subset \{1, \dots, n\}$ ist $\det(A_S) > 0$.

iii) Für jedes $k \in \{1, \dots, n\}$ ist $\det(A_{\{1,\dots,k\}}) > 0$.

Beweis: i) $\Rightarrow$ ii) A ist die Strukturmatrix der positiv definiten Bilinearform $\beta : v \mapsto {}^t vAv$ bezüglich der kanonischen Basis $e_1, \dots, e_n$ des $\mathbb{R}^n$. Die Matrix A_S ist dann die Strukturmatrix der Einschränkung $\beta|_W$ auf den Raum $W = \langle e_i : i \in S \rangle$. Diese Einschränkung bleibt natürlich positiv definit, und somit ist $\det(A_S) > 0$ nach Folgerung 5.3.3.

ii) $\Rightarrow$ iii) ist trivial.

iii) $\Rightarrow$ i) Induktion nach n. Die Behauptung ist richtig für $n = 1$, sei also $n > 1$. Sei $\beta|_W$ die Einschränkung von $\beta : v \mapsto {}^t vAv$ auf $W = \langle e_1, \dots, e_{n-1} \rangle$. Nach Induktionsvoraussetzung ist $\beta|_W$ positiv definit, und damit nicht ausgeartet. Aus 5.2.6 ergibt sich daher

$$\mathbb{R}^n = W \oplus W^\perp,$$

wobei $W^\perp$ eindimensional ist, etwa $W^\perp = \langle v \rangle$ mit $v \in \mathbb{R}^n$. Wir müssen nur noch

$$a := \beta(v, v) > 0$$

zeigen. Sei $w_1, \ldots, w_{n-1}$ eine Orthogonalbasis von $\beta|_W$. Die Matrix von β bezüglich der Basis $w_1, \ldots, w_{n-1}, v$ ist dann

$$\operatorname{diag}(a_1, \ldots, a_{n-1}, a) \qquad \text{mit positiven } a_1, \ldots, a_{n-1}.$$

Ihre Determinante $a_1 \cdot \ldots \cdot a_{n-1} a$ ist wie $\det(A)$ positiv (Feststellung 5.2.2 iv)), und folglich muß auch a positiv sein. ■

Klassifikation

Satz 5.3.5 (Trägheitssatz von Sylvester) [2] *Sei β eine nicht-ausgeartete symmetrische Bilinearform auf dem n-dimensionalen $\mathbb{R}$-Vektorraum V. Dann gibt es eine Basis von V, bezüglich derer die Matrix B von β die folgende Form hat:*

$$B = \operatorname{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{-1, \ldots, -1}_{s}) \qquad (r + s = n).$$

Die Zahlen r und s sind dabei eindeutig bestimmt.

Beweis: Auf jeden Fall gibt es eine Orthogonalbasis für β. Wir hatten schon auf Seite 119 festgestellt, daß sich die Diagonalkoeffizienten noch um Quadrate, d.h. hier um positive Zahlen, abändern lassen. Somit gibt es eine Orthogonalbasis, so daß nur die Zahlen 1 und -1 auf der Diagonale vorkommen (0 kommt nicht vor, da β nicht ausgeartet ist). Ferner kann die Reihenfolge der Diagonalkoeffizienten durch Umnumerierung der Basiselemente beliebig geändert werden. Also gibt es eine Basis $v_1, \ldots, v_n$ der beschriebenen Art, und es bleibt nur die Eindeutigkeit der Zahlen r und s zu zeigen.

Offenbar ist $V_+ = \langle v_1, \ldots, v_r \rangle$ ein positiv definiter Teilraum, und $V_- = \langle v_{r+1}, \ldots, v_n \rangle$ ein negativ definiter Teilraum. Wir behaupten, daß V_- ein negativ definiter Teilraum maximaler Dimension ist. Ist nämlich W irgendein negativ definiter Teilraum, so gilt $W \cap V_+ = 0$, also $\dim(W) + \dim(V_+) \leq n$, und somit $\dim(W) \leq n - r = s$. Die Zahl s läßt sich also charakterisieren als die maximale Dimension eines negativ definiten Teilraums, und ist somit durch β eindeutig bestimmt. ■

Traditionell nennt man die Zahl s in diesem Satz den **Trägheitsindex** von β, und die Größe $r - s$ die **Signatur** von β.

Orthogonale Gruppen

Sei

$$E_{m,n} = \operatorname{diag}(\underbrace{1, \ldots, 1}_{m}, \underbrace{-1, \ldots, -1}_{n}), \qquad m, n \in \mathbb{N}_0.$$

[2] James Sylvester, 1814-1897

Sei $\beta_{m,n}$ die durch die Matrix $E_{m,n}$ auf dem $\mathbb{R}^{m+n}$ definierte Bilinearform. Die orthogonale Gruppe von $\beta_{m,n}$ wird mit $\mathrm{O}(m,n)$ bezeichnet, also (Feststellung 5.2.8)

$$\mathrm{O}(m,n) = \{A \in \mathrm{GL}(m+n,\mathbb{R}) : {}^tAE_{m,n}A = E_{m,n}\}.$$

Insbesondere ist

$$\mathrm{O}(m,\mathbb{R}) = \mathrm{O}(m,0) = \{A \in \mathrm{GL}(m,\mathbb{R}) : {}^tAA = \mathbf{1}\},$$

siehe S. 116. Die Matrizen aus $\mathrm{O}(m,n)$ haben offenbar Determinante 1 oder -1, und man setzt

$$\mathrm{SO}(m,n) = \mathrm{O}(m,n) \cap \mathrm{SL}(m+n,\mathbb{R}).$$

Diese Untergruppe macht die „Hälfte“ von $\mathrm{O}(m,n)$ aus: Wählt man nämlich irgendein Element $A_0 \in \mathrm{O}(m,n)$, welches nicht in $\mathrm{SO}(m,n)$ enthalten ist, z.B.

$$A_0 = \mathrm{diag}(1,\dots,1,-1),$$

so gilt

$$\mathrm{O}(m,n) = \mathrm{SO}(m,n) \sqcup A_0\mathrm{SO}(m,n).$$

Der spezielle Grundkörper $\mathbb{R}$ ermöglicht es, neben den allgemeinen *algebraischen* Überlegungen auch noch *topologische* anzustellen; wir nehmen an, der Leser ist mit den Grundbegriffen der mengentheoretischen Topologie vertraut.

Der Raum $M(n,\mathbb{R})$ verfügt über eine natürliche Topologie, die von (irgend) einem Isomorphismus mit dem $\mathbb{R}^{n^2}$ herkommt. Die Determinantenabbildung

$$\det : M(n,\mathbb{R}) \longrightarrow \mathbb{R}$$

ist *stetig*, wie sofort aus der Leibnizformel 3.2.4 folgt. Als Urbild der offenen Teilmenge $\mathbb{R}^* \subset \mathbb{R}$ ist daher $\mathrm{GL}(n,\mathbb{R})$ eine *offene* Teilmenge von $M(n,\mathbb{R})$: Die invertierbaren Matrizen machen den größten Teil von $M(n,\mathbb{R})$ aus (genauer: der „Rest“ ist eine Untermannigfaltigkeit echt kleinerer Dimension).

Wie eben gesehen, nimmt die Einschränkung von $\det : \mathrm{GL}(m+n,\mathbb{R}) \to \mathbb{R}^*$ auf $\mathrm{O}(m,n)$ nur die Werte ± 1 an:

$$\det : \mathrm{O}(m,n) \longrightarrow \{\pm 1\}.$$

Es folgt, daß sowohl $\mathrm{SO}(m,n)$ als auch $A_0\mathrm{SO}(m,n)$ (mit A_0 wie oben) offen in $\mathrm{O}(m,n)$ sind. Man kann zeigen, daß $\mathrm{SO}(m,n)$ überdies *zusammenhängend* ist; es folgt, daß $\mathrm{SO}(m,n)$ und $A_0\mathrm{SO}(m,n)$ die beiden Zusammenhangskomponenten von $\mathrm{O}(m,n)$ sind. Die Untergruppe $\mathrm{SO}(m,n)$ läßt sich also topologisch charakterisieren als diejenige Zusammenhangskomponente von $\mathrm{O}(m,n)$, die das neutrale Element $\mathbf{1}$ enthält (die *Einskomponente*).

Wir betrachten speziell den Fall $m+n=2$. Die Abbildungen

$$\mathbb{R}/(2\pi\mathbb{Z}) \longrightarrow \mathrm{SO}(2,\mathbb{R}), \tag{5.16}$$

$$x \longmapsto \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix},$$

und

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathrm{SO}(1,1), \\ x &\longmapsto \begin{pmatrix} \cosh(x) & \sinh(x) \\ \sinh(x) & \cosh(x) \end{pmatrix}, \end{aligned} \tag{5.17}$$

erweisen sich leicht als stetige, injektive Gruppenhomomorphismen. Man kann zeigen, daß es sich sogar um Isomorphismen handelt. Dies sind explizite *Karten*, die den Gruppen $\mathrm{SO}(2,\mathbb{R})$ und $\mathrm{SO}(1,1)$ die Struktur einer eindimensionalen differenzierbaren Mannigfaltigkeit geben.

Außerdem liest man an diesen Isomorphismen ab, daß $\mathrm{SO}(2,\mathbb{R})$ ***kompakt*** ist, und $\mathrm{SO}(1,1)$ nicht. Allgemein gilt:

$$\mathrm{SO}(m,n) \text{ ist kompakt} \qquad \Longleftrightarrow \qquad m = 0 \text{ oder } n = 0.$$

Beweis hierfür: Seien $v_1, \dots, v_m$ die Spalten einer Matrix $A \in \mathrm{SO}(m,\mathbb{R})$. Aus der Bedingung ${}^tAA = \mathbf{1}$ folgt ${}^tv_iv_i = 1$, also liegen alle Spaltenvektoren auf der kompakten Einheitskugel im $\mathbb{R}^m$. Dies zeigt die Kompaktheit von $\mathrm{SO}(m,0)$. Für die Kompaktheit von $\mathrm{SO}(0,n)$ beachte man nur

$$\mathrm{SO}(0,n) = \mathrm{SO}(n,0)$$

(allgemein gilt $\mathrm{SO}(m,n) \simeq \mathrm{SO}(n,m)$). Ist $m \neq 0$ und $n \neq 0$, so läßt sich eine Abbildung $\mathbb{R} \to \mathrm{SO}(m,n)$ der Art (5.17) angeben, was die Nicht-Kompaktheit zeigt. ∎

Wir haben den folgenden Satz bewiesen:

Satz 5.3.6 *Für eine symmetrische Bilinearform β auf dem endlich-dimensionalen $\mathbb{R}$-Vektorraum V sind äquivalent:*

i) Es gibt eine Basis von V, bezüglich derer die Matrix von β gleich $\mathbf{1}$ oder $-\mathbf{1}$ ist.

ii) β ist positiv oder negativ definit.

iii) Für alle $v \in V$ gilt: Aus $\beta(v,v) = 0$ folgt $v = 0$ (d.h., die zu β assoziierte quadratische Form stellt nicht die Null dar; siehe Seite 119).

iv) Die Gruppe $\mathrm{SO}(V,\beta)$ ist kompakt.

v) Die Gruppe $\mathrm{O}(V,\beta)$ ist kompakt.

Selbstadjungierte Operatoren auf euklidischen Räumen

Definition 5.3.7 *Ein* **euklidischer Raum** *ist ein endlich-dimensionaler $\mathbb{R}$-Vektorraum V, versehen mit einer positiv-definiten symmetrischen Bilinearform β.*

Im Falle eines euklidischen Vektorraumes wird gerne die Notation

$$\langle v, w\rangle = \beta(v, w)$$

benutzt; die Bilinearform $\langle\ ,\ \rangle$ heißt auch ein **Skalarprodukt** auf V.

Beispiel: $V = \mathbb{R}^n$ mit dem (durch die Einheitsmatrix definierten) **Standard-Skalarprodukt**

$$\langle x, y\rangle = {}^t xy = \sum_{i=1}^{n} x_i y_i \qquad (x = {}^t(x_1, \dots, x_n),\ y = {}^t(y_1, \dots, y_n)).$$

Wie auf Seite 115 ordnen wir jedem Endomorphismus f des euklidischen Vektorraumes V seine Adjungierte f^* zu; diese ist also durch

$$\langle fv, w\rangle = \langle v, f^* w\rangle \qquad \text{für alle } v, w \in V$$

charakterisiert. Da $\langle\ ,\ \rangle$ symmetrisch ist, ist $f^* = {}^*f$.

Definition 5.3.8 *Der Endomorphismus f des euklidischen Vektorraumes V heißt* **selbstadjungiert**, *falls $f = f^*$, falls also*

$$\langle fv, w\rangle = \langle v, fw\rangle \qquad \textit{für alle } v, w \in V$$

gilt.

Da $\langle\ ,\ \rangle$ positiv definit ist, gibt es eine **Orthonormalbasis** für V, d.h. eine Basis $v_1, \dots, v_n$, so daß

$$\langle v_i, v_j\rangle = \delta_{ij}.$$

Feststellung 5.3.9 *Sei $v_1, \dots, v_n$ eine Orthonormalbasis des euklidischen Vektorraumes V. Besitzt $f \in \mathrm{End}(V)$ bezüglich dieser Basis die Matrix A, so besitzt f^* die Matrix ${}^tA$. Insbesondere gilt: Die Matrix eines selbstadjungierten Endomorphismus bezüglich einer Orthonormalbasis ist symmetrisch.*

Beweis: Sei φ der durch $v_1, \dots, v_n$ definierte Isomorphismus $\mathbb{R}^n \to V$, d.h., $\varphi(e_i) = v_i$. Da es sich um eine Orthonormalbasis handelt, gilt

$$\langle \varphi(x), \varphi(y)\rangle = {}^t xy \qquad \text{für alle } x, y \in \mathbb{R}^n.$$

Nach Definition der Matrix A gilt $f(\varphi(x)) = \varphi(Ax)$ für alle $x \in \mathbb{R}^n$. Deshalb ist

$$\langle f(\varphi(x)), \varphi(y)\rangle = \langle \varphi(Ax), \varphi(y)\rangle = {}^t(Ax)y = {}^t x\, {}^t Ay. \tag{5.18}$$

Wenn andererseits B die Matrix von f^* bezeichnet, so gilt

$$\langle f(\varphi(x)), \varphi(y)\rangle = \langle \varphi(x), f^*(\varphi(y))\rangle = \langle \varphi(x), \varphi(By)\rangle = {}^t xBy. \tag{5.19}$$

Vergleich von (5.18) und (5.19) liefert nun ${}^tA = B$. ∎

Selbstadjungierte Endomorphismen entsprechen also symmetrischen Matrizen. Über letztere gilt folgender Satz.

Satz 5.3.10 *Eine reelle, symmetrische Matrix ist diagonalisierbar, und alle ihre Eigenwerte sind reell.*

Der **Beweis** könnte hier mit Methoden der reellen Analysis geführt werden. Er wird sich jedoch leicht aus den Überlegungen des nächsten Abschnitts ergeben (siehe Seite 131).

Satz 5.3.11 (Spektralsatz für euklidische Vektorräume) *Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum, und f ein selbstadjungierter Endomorphismus von V. Dann besitzt V eine Orthonormalbasis, die aus Eigenvektoren von f besteht.*

Beweis: Der Beweis erfolgt durch Induktion nach $n = \dim(V)$, wobei für $n = 1$ nichts zu zeigen ist. Aufgrund von Feststellung 5.3.9 und Satz 5.3.10 gibt es einen Eigenvektor $v \in V$ von f:

$$fv = \lambda v \qquad \text{mit } \lambda \in \mathbb{R}.$$

Wir dürfen $\langle v, v\rangle = 1$ annehmen. Sei $W = \langle v\rangle^{\perp}$ das orthogonale Komplement von v. Für jedes $w \in W$ gilt

$$\langle fw, v\rangle = \langle w, fv\rangle = \lambda\langle w, v\rangle = 0,$$

und also $fw \in W$. Dies zeigt $f(W) \subset W$. Wir können deshalb die Induktionsvoraussetzung auf den euklidischen Raum $(W, \langle\ ,\ \rangle)$ mit dem selbstadjungierten Endomorphismus $f\Big|_W$ anwenden, und erhalten eine Orthonormalbasis $w_1, \dots, w_{n-1}$ von W der gewünschten Art. Dann ist $v, w_1, \dots, w_{n-1}$ die gesuchte Orthonormalbasis von V. ∎

Folgerung 5.3.12 *Sei A eine reelle symmetrische $n \times n$-Matrix. Dann gibt es eine orthogonale Matrix $S \in O(n)$ mit*

$${}^tSAS = \operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

wobei $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ notwendig die Eigenwerte von A sind.

Beweis: Wir versehen den $\mathbb{R}^n$ mit dem Standard-Skalarprodukt $\langle x, y\rangle = \sum x_i y_i$. Die Matrix A ist dann ein selbstadjungierter Endomorphismus auf diesem euklidischen Raum, und wir wählen eine Orthonormalbasis $v_1, \dots, v_n$ wie im Spektralsatz. Sei S die Matrix mit den Spalten $v_1, \dots, v_n$. Da es sich um eine Orthonormalbasis handelt, gilt ${}^tSS = \mathbf{1}$, d.h. $S \in O(n)$. Da jedes v_i Eigenvektor für A ist, gilt

$$AS = S \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$

Nun beachte man nur noch ${}^tS = S^{-1}$. ∎

5.4 Hermitesche Formen

In diesem Abschnitt betrachten wir bilineare Abbildungen auf *komplexen* Vektorräumen. Zunächst sei an die **komplexe Konjugation**

$$\mathbb{C} \longrightarrow \mathbb{C}, \qquad z = x + iy \longmapsto x - iy =: \bar{z} \qquad (x, y \in \mathbb{R})$$

erinnert. Diese ist ein Ringhomomorphismus, d.h.

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \qquad \overline{z_1 z_2} = \bar{z}_1 \bar{z}_2.$$

Definitionen

Sei jetzt V ein n-dimensionaler $\mathbb{C}$-Vektorraum. Versieht man die abelsche Gruppe V mit der neuen Skalarmultiplikation

$$a.v := \bar{a}v, \qquad a \in \mathbb{C},\ v \in V$$

(rechts steht die ursprüngliche Skalarmultiplikation auf V), so erhält man den **konjugiert komplexen Vektorraum** $\bar{V}$; dieser ist ebenfalls ein n-dimensionaler $\mathbb{C}$-Vektorraum. Anstatt Bilinearformen auf V zu untersuchen, erweist es sich als sinnvoller, bilineare Abbildungen

$$V \times \bar{V} \longrightarrow \mathbb{C}$$

zu betrachten. Diese heißen **Sesquilinearformen**, und können in äquivalenter Weise auch wie folgt definiert werden.

Definition 5.4.1 *Sei V ein endlich-dimensionaler $\mathbb{C}$-Vektorraum. Eine* **Sesquilinearform**[3] *auf V ist eine Abbildung*

$$\beta:\ V \times V \longrightarrow \mathbb{C}$$

mit den folgenden Eigenschaften:

i) β ist linear in der ersten Variablen:

$$\beta(v_1 + v_2, w) = \beta(v_1, w) + \beta(v_2, w), \qquad \beta(av, w) = a\,\beta(v, w).$$

ii) β ist semilinear in der zweiten Variablen:

$$\beta(v, w_1 + w_2) = \beta(v, w_1) + \beta(v, w_2), \qquad \beta(v, aw) = \bar{a}\,\beta(v, w).$$

Die Sesquilinearform β heißt eine **hermitesche Form**[4]*, falls zusätzlich*

$$\beta(v, w) = \overline{\beta(w, v)} \qquad \text{für alle } v, w \in V$$

gilt.

Für eine hermitesche Form β ist $\beta(v, v)$ offenbar immer eine reelle Zahl. Deshalb macht folgende Definition Sinn.

Definition 5.4.2 *Eine hermitesche Form β auf V heißt* **positiv definit***, falls*

$$\beta(v, v) > 0 \qquad \text{für alle } v \in V,\ v \neq 0.$$

Ein **unitärer Raum** *ist ein endlich-dimensionaler $\mathbb{C}$-Vektorraum V, versehen mit einer positiv definiten hermiteschen Form β.*

[3] *sesqui-* (lat.): anderthalbfach
[4] Charles Hermite, 1822-1901

Für unitäre Räume schreibt man die hermitesche Form gerne als

$$\langle v, w\rangle = \beta(v, w),$$

und nennt sie ein **Skalarprodukt** auf V.

Beispiel: Der $\mathbb{C}^n$ wird durch das **Standard-Skalarprodukt**

$$\langle x, y\rangle = {}^t x\bar{y} = \sum_{i=1}^{n} x_i\bar{y}_i \qquad (x = {}^t(x_1, \ldots, x_n),\ y = {}^t(y_1, \ldots, y_n))$$

zu einem unitären Raum.

Klassifikation

Genau wie Bilinearformen lassen sich Sesquilinearformen durch Matrizen beschreiben: Die Strukturmatrix der Sesquilinearform β bezüglich der Basis $v_1, \ldots, v_n$ ist $B = \beta(v_i, v_j)$. Offenbar ist β genau dann hermitesch, wenn

$$B = {}^t\bar{B}.$$

Matrizen mit dieser Eigenschaft heißen **hermitesche Matrizen**. Jede hermitesche Matrix definiert eine hermitesche Form auf dem $\mathbb{C}^n$. Die Einheitsmatrix definiert das Standard-Skalarprodukt.

Wir haben gesehen (Seite 119), daß für $\operatorname{char}(K) \neq 2$ eine *symmetrische Bilinearform* β schon durch die ihr zugeordnete *quadratische Form* $q(v) = \frac{1}{2}\beta(v, v)$ bestimmt ist:

$$\beta(v, w) = \frac{1}{2}\Big(q(v + w) - q(v - w)\Big). \tag{5.20}$$

Ebenso ist eine hermitesche Form β durch die reellwertige Funktion $q(v) = \frac{1}{2}\beta(v, v)$ bestimmt:

$$\beta(v, w) = \frac{1}{2}\Big(q(v + w) - q(v - w) + iq(v + iw) - iq(v - iw)\Big). \tag{5.21}$$

Die Formeln (5.20) und (5.21) heißen auch **Polarisationsformeln**. Mit Hilfe von (5.21) beweist man nun leicht das Analogon von Feststellung 5.2.11: *Jede hermitesche Form besitzt eine Orthogonalbasis.* Genauer gilt die folgende hermitesche Version des Sylvesterschen Trägheitssatzes 5.3.5.

Satz 5.4.3 *Sei β eine nicht-ausgeartete hermitesche Form auf dem n-dimensionalen $\mathbb{C}$-Vektorraum V. Dann gibt es eine Basis von V, bezüglich derer die Matrix B von β die folgende Form hat:*

$$B = \operatorname{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{-1, \ldots, -1}_{s}) \qquad (r + s = n).$$

Die Zahlen r und s sind dabei eindeutig bestimmt.

Wir verzichten auf den **Beweis**, da er wie im reellen Fall erfolgt. Offenbar ist β genau dann positiv definit, wenn $s = 0$ ist.

Man beachte den Unterschied zur Klassifikation der *symmetrischen Bilinearformen* auf komplexen Vektorräumen: Hier gibt es stets eine *Orthonormalbasis* (Feststellung 5.2.12).

Unitäre Gruppen

Sei β eine hermitesche Form auf V. Ein Endomorphismus $f \in \mathrm{GL}(V)$ heißt **unitär**, falls

$$\beta(fv, fw) = \beta(v, w) \qquad \text{für alle } v, w \in V.$$

Die unitären Endomorphismen bilden eine Untergruppe von $\mathrm{GL}(V)$, die **unitäre Gruppe**

$$U(V, \beta) = \{f \in \mathrm{GL}(V) : \beta(fv, fw) = \beta(v, w) \text{ für alle } v, w \in V\}$$

(manchmal einfach $U(V)$).

Feststellung 5.4.4 *Sei $\mathcal{B}$ eine Basis von V, und B die Matrix von β bezüglich $\mathcal{B}$. Sei $f \in \mathrm{GL}(V)$ und A die Matrix von f bezüglich der gleichen Basis $\mathcal{B}$. Dann gilt:*

$$f \in \mathrm{U}(V, \beta) \qquad \Longleftrightarrow \qquad {}^t\!AB\bar{A} = B.$$

Beweis: Analog zu 5.2.8. ■

Durch Determinantenbildung ergibt sich aus dieser Aussage:

$$f \in \mathrm{U}(V, \beta) \quad \Longrightarrow \quad |\det(f)| = 1.$$

Wir definieren noch die **spezielle unitäre Gruppe**

$$\mathrm{SU}(V, \beta) = \mathrm{U}(V, \beta) \cap \mathrm{SL}(V).$$

Wie im Abschnitt über orthogonale Gruppen auf Seite 122 sei

$$E_{m,n} = \mathrm{diag}(\underbrace{1, \dots, 1}_{m}, \underbrace{-1, \dots, -1}_{n}), \qquad m, n \in \mathbb{N}_0.$$

Sei $\beta_{m,n}$ die durch die Matrix $E_{m,n}$ auf dem $\mathbb{C}^{m+n}$ definierte hermitesche Form. Die unitäre Gruppe von $\beta_{m,n}$ wird mit $\mathrm{U}(m, n)$ bezeichnet, also

$$\mathrm{U}(m, n) = \{A \in \mathrm{GL}(m + n, \mathbb{C}) : {}^t\!AE_{m,n}\bar{A} = E_{m,n}\},$$

und insbesondere

$$\mathrm{U}(m) := \mathrm{U}(m, 0) = \{A \in \mathrm{GL}(m, \mathbb{C}) : {}^t\!A\bar{A} = \mathbf{1}\}.$$

Diese unitären Gruppen tragen als Teilmengen von $M(n, \mathbb{C})$ eine natürliche Topologie, und es lassen sich ähnliche topologische Überlegungen wie für orthogonale Gruppen anstellen; z.B. ist leicht zu sehen, daß $U(m)$ *kompakt* ist.

Schließlich geben wir noch eine explizite Beschreibung der speziellen unitären Gruppen im Fall $m+n = 2$. Es gilt

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} : a, b \in \mathbb{C},\ a\bar{a} + b\bar{b} = 1 \right\}, \tag{5.22}$$

$$\mathrm{SU}(1,1) = \left\{ \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} : a, b \in \mathbb{C},\ a\bar{a} - b\bar{b} = 1 \right\}. \tag{5.23}$$

Dies ist leicht nachzuprüfen. Man vergleiche mit den Beschreibungen (5.16) und (5.17) der orthogonalen Gruppen.

Normale Operatoren auf unitären Räumen

Sei jetzt $(V, \langle\ ,\ \rangle)$ ein unitärer Raum. Zu einem beliebigen Endomorphismus $f \in \mathrm{End}_{\mathbb{C}}(V)$ ist die **Adjungierte** $f^* \in \mathrm{End}_{\mathbb{C}}(V)$ der durch

$$\langle fv, w\rangle = \langle v, f^*w\rangle \qquad \text{für alle } v, w \in V$$

charakterisierte Endomorphismus (die Existenz und Eindeutigkeit kann man auf 5.2.7 zurückführen, oder direkt einsehen). Da $\langle, \rangle$ konjugiert-symmetrisch ist, ist die Rechts- gleich der Links-Adjungierten, d.h. es gilt auch

$$\langle v, fw\rangle = \langle f^*v, w\rangle \qquad \text{für alle } v, w \in V.$$

Folgende Eigenschaften der Adjungierten prüft man sofort nach:

$$(f_1 + f_2)^* = f_1^* + f_2^*, \qquad (af)^* = \bar{a}f^*, \qquad (f_1f_2)^* = f_2^*f_1^*, \qquad f^{**} = f.$$

Wie in 5.3.9 sieht man: Ist $v_1, \ldots, v_n$ eine Orthonormalbasis von $\langle\ ,\ \rangle$, und ist A die Matrix von f bezüglich dieser Basis, so ist

$$A^* := {}^t\bar{A}$$

die Matrix von f^*. Man nennt A^* die zu A **adjungierte Matrix**.

Definition 5.4.5 *Der Endomorphismus f des unitären Raumes $(V, \langle\ ,\ \rangle)$ heißt*

i) **selbstadjungiert**, *falls $f = f^*$,*

ii) **normal**, *falls $ff^* = f^*f$.*

Offenbar ist jeder selbstadjungierte Endomorphismus normal. Es stellt sich heraus, daß für unitäre Räume die Normalität der wichtigere Begriff ist.

Feststellung 5.4.6 *Sei f ein normaler Endomorphismus von V, und $v \in V$ ein Eigenvektor zum Eigenwert λ: $fv = \lambda v$. Dann ist v auch Eigenvektor zu f^*, und zwar zum Eigenwert $\bar{\lambda}$:*

$$f^*v = \bar{\lambda}v.$$

Beweis: Aus der Normalität folgt sofort

$$\langle fv, fw\rangle = \langle f^*v, f^*w\rangle \qquad \text{für alle } v, w \in V.$$

Setzt man insbesondere $v = w$, so folgt wegen der Positiv-Definitheit

$$fv = 0 \qquad \Longleftrightarrow \qquad f^*v = 0,$$

also $\ker(f) = \ker(f^*)$. Angewandt auf den ebenfalls normalen Endomorphismus $f - \lambda\,\mathrm{id}$ folgt

$$\ker(f - \lambda\,\mathrm{id}) = \ker((f - \lambda\,\mathrm{id})^*) = \ker(f^* - \bar{\lambda}\,\mathrm{id}).$$

Dies ist die Behauptung. ∎

Satz 5.4.7 (Spektralsatz für unitäre Vektorräume) *Sei* $(V, \langle\ ,\ \rangle)$ *ein unitärer Raum, und* f *ein normaler Endomorphismus von* V*. Dann besitzt* V *eine Orthonormalbasis, die aus Eigenvektoren von* f *besteht.*

Beweis: Wie im Beweis von Satz 5.3.11 führen wir Induktion nach $n = \dim(V)$. Sei gleich $n \geq 2$. Da $\mathbb{C}$ algebraisch abgeschlossen ist, besitzt f einen Eigenvektor $v \in V$:

$$fv = \lambda v \qquad \text{mit } \lambda \in \mathbb{C}$$

(das charakteristische Polynom besitzt eine Nullstelle in $\mathbb{C}$). Wir dürfen $\langle v, v\rangle = 1$ annehmen. Sei $W = \langle v\rangle^{\perp}$ das orthogonale Komplement von v. Für jedes $w \in W$ gilt

$$\langle f^* w, v\rangle = \langle w, fv\rangle = \bar{\lambda}\langle w, v\rangle = 0,$$

und also $f^* w \in W$. Dies zeigt $f^*(W) \subset W$. Wegen Feststellung 5.4.6 gilt ferner für $w \in W$

$$\langle fw, v\rangle = \langle w, f^* v\rangle = \lambda\langle w, v\rangle = 0,$$

und damit $f(W) \subset W$. Sowohl f als auch f^* induzieren somit Endomorphismen von W, und man kann die Induktionsvoraussetzung auf den unitären Raum $(W, \langle\ ,\ \rangle)$ anwenden. ∎

Folgerung 5.4.8 *Sei* $A \in M(n, \mathbb{C})$ *eine normale Matrix, d.h.* $AA^* = A^*A$*. Dann gibt es eine unitäre Matrix* $S \in U(n)$ *mit*

$$S^*AS = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

wobei $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ *notwendig die Eigenwerte von* A *sind.*

Der **Beweis** verläuft ganz ähnlich wie der von Folgerung 5.3.12. ∎

Folgerung 5.4.9 *Ein selbstadjungierter Endomorphismus auf einem unitären Raum ist diagonalisierbar, und alle Eigenwerte sind reell.*

Beweis: Ein selbstadjungierter Endomorphismus f ist normal, also folgt die Diagonalisierbarkeit aus dem Spektralsatz. Ist λ ein Eigenwert und v ein zugehöriger Eigenvektor, so gilt

$$\lambda\langle v, v\rangle = \langle fv, v\rangle = \langle v, fv\rangle = \bar{\lambda}\langle v, v\rangle,$$

und damit $\lambda = \bar{\lambda}$. ∎

Angewandt auf das Standard-Skalarprodukt auf dem $\mathbb{C}^n$ besagt diese Folgerung: Jede selbstadjungierte Matrix, d.h. jedes $A \in M(n, \mathbb{C})$ mit $A = A^*$, ist diagonalisierbar, und alle Eigenwerte sind reell.

Schließlich führen wir den **Beweis** von Satz 5.3.10: Eine reelle, symmetrische Matrix besitzt jedenfalls einen komplexen Eigenwert, wenn man sie als Element von $M(n, \mathbb{C})$ auffaßt. Als solches ist sie aber *selbstadjungiert*, und der Eigenwert muß folglich reell sein. Nur diese Existenz eines reellen Eigenwertes wurde im Beweis von Satz 5.3.11 benutzt. Die Aussage von 5.3.10 ist aber in der des Spektralsatzes enthalten. ∎

Kapitel 6

Tensorprodukte

Zu zwei K-Vektorräumen V und W kann man einen weiteren K-Vektorraum $V \otimes W$ bilden, das *Tensorprodukt*, welches durch eine gewisse universelle Eigenschaft charakterisiert ist. Tensorprodukte besitzen zahlreiche Anwendungen, so z.B. in der Analysis auf Mannigfaltigkeiten, wo vielfache Tensorprodukte von Tangentialräumen und ihren Dualräumen gebildet werden.

Nach den allgemeinen Grundlagen werden in den Abschnitten 6.2 und 6.3 die Zusammenhänge zwischen Tensorprodukten und Multilinearformen erläutert. Als weitere Anwendung von Tensorprodukten betrachten wir die *Basiserweiterung*, die aus einem K-Vektorraum einen L-Vektorraum macht, wobei L ein größerer Körper als K ist. Ein Spezialfall davon ist die *Komplexifizierung* eines $\mathbb{R}$-Vektorraumes zu einem $\mathbb{C}$-Vektorraum.

6.1 Grundlagen

Definition und universelle Eigenschaft

Sei K wieder ein beliebiger Körper, und V, W seien K-Vektorräume, nicht notwendig endlich-dimensional. Sei F der K-Vektorraum aller Abbildungen $V \times W \to K$, die fast überall (d.h., mit Ausnahme endlich vieler Elemente) den Wert Null annehmen. Man darf sich F als die Menge aller formaler Ausdrücke der Form

$$\sum_{i=1}^{n} a_i(v_i, w_i), \qquad a_i \in K,\ v_i \in V,\ w_i \in W \tag{6.1}$$

vorstellen, die in natürlicher Weise addiert und mit Skalaren multipliziert werden; das Element (6.1) entspricht der Abbildung, die (v_i, w_i) auf a_i abbildet. Offenbar bilden die Elemente von $V \times W$ eine Basis von F. Man nennt F den **freien Vektorraum** in der Basis $V \times W$.

Sei U der Untervektorraum von F, der von allen Elementen der Form

$$(v_1 + v_2, w) - (v_1, w) - (v_2, w), \qquad (av, w) - a(v, w), \tag{6.2}$$

$$(v, w_1 + w_2) - (v, w_1) - (v, w_2), \qquad (v, aw) - a(v, w) \tag{6.3}$$

aufgespannt wird ($a \in K$, $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$). Wir definieren das **Tensorprodukt** von V und W als den Quotientenvektorraum

$$V \otimes W := F/U.$$

Das Bild des Elementes $(v, w) \in F$ unter der natürlichen Projektion $F \to V \otimes W$ wird mit

$$v \otimes w$$

bezeichnet.

Feststellung 6.1.1 *V und W seien beliebige K-Vektorräume.*

i) Jedes Element des Tensorprodukts $V \otimes W$ besitzt eine Darstellung der Form

$$\sum_{i=1}^{n} a_i(v_i \otimes w_i)$$

(die jedoch keineswegs eindeutig ist).

ii) In $V \otimes W$ gelten die Rechenregeln

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

Beweis: Die kanonische Projektion $F \to V \otimes W$ ist surjektiv. Wendet man sie auf das allgemeine Element (6.1) von F an, erhält man i). Aussage ii) folgt, indem man die Projektion auf die Relationen (6.2) und (6.3) anwendet, denn diese Relationen werden Null im Quotienten. ∎

In Abschnitt 5.1 haben wir Bilinearformen $V \times W \to K$ definiert. Ersetzt man K durch einen beliebigen K-Vektorraum Y, so erhält man den allgemeineren Begriff einer **bilinearen Abbildung**. Dies ist also eine Abbildung

$$\beta : \ V \times W \longrightarrow Y,$$

die in jeder der beiden Variablen linear ist. Die Menge aller dieser Abbildungen sei mit

$$\mathrm{Bil}_K(V \times W, Y)$$

bezeichnet. Diese ist selbst in natürlicher Weise ein K-Vektorraum. Das Tensorprodukt $V \otimes W$, zusammen mit der natürlichen Abbildung $V \times W \to V \otimes W$, $(v, w) \mapsto v \otimes w$, besitzt folgende „universelle Eigenschaft“:

Feststellung 6.1.2 (Universelle Eigenschaft des Tensorprodukts) *Zu jeder <u>bilinearen</u> Abbildung $\beta : \ V \times W \to Y$ gibt es genau eine <u>lineare</u> Abbildung $\varphi : \ V \otimes W \to Y$, die das folgende Diagramm kommutativ macht:*

$$\begin{tikzcd} V \times W \arrow[r, "f"] \arrow[d] & Y \\ V \otimes W \arrow[ur, "\varphi"'] & \end{tikzcd}$$

Beweis: Wenn F wieder den freien Vektorraum in der Basis $V \times W$ bezeichnet, so definieren wir zunächst eine lineare Abbildung

$$\begin{aligned} \varphi: \ F &\longrightarrow Y, \\ \sum a_i(v_i, w_i) &\longmapsto \sum a_i \beta(v_i, w_i). \end{aligned}$$

Es ist eine unmittelbare Konsequenz der Bilinearität von β, daß φ die Ausdrücke der Form (6.2) und (6.3) auf Null abbildet. Mit anderen Worten, der von diesen Elementen erzeugte Untervektorraum U liegt im Kern von φ. Aufgrund des Homomorphiesatzes 1.5.7 induziert φ daher eine (ebenfalls so genannte) lineare Abbildung

$$\varphi: \ V \otimes W = F/U \longrightarrow Y.$$

Offenbar gilt

$$\varphi(v \otimes w) = \beta(v, w) \qquad \text{für alle } v \in V,\ w \in W,$$

d.h. obiges Diagramm ist kommutativ. Ferner kann es nur *eine* Abbildung mit dieser Eigenschaft geben, denn eine lineare Abbildung auf $V \otimes W$ ist wegen Feststellung 6.1.1 i) durch ihre Werte auf den Elementen $v \otimes w$ festgelegt. ∎

Einige Isomorphismen

Feststellung 6.1.3 *Seien U, V, W Vektorräume über dem Körper K. Dann gibt es natürliche Isomorphismen*

i) $V \otimes K \simeq V$.

ii) $V \otimes W \simeq W \otimes V$.

iii) $U \otimes (V \otimes W) \simeq (U \otimes V) \otimes W$.

Beweis: Alle drei Aussagen werden mit Hilfe der universellen Eigenschaft 6.1.2 bewiesen.

i) Die Abbildung

$$\begin{aligned} V \times K &\longrightarrow V, \\ (v, a) &\longmapsto av, \end{aligned}$$

ist bilinear. Aufgrund der universellen Eigenschaft gibt es daher eine lineare Abbildung $\varphi : V \otimes K \to V$ mit

$$\varphi(v \otimes a) = av.$$

Eine Umkehrabbildung zu φ ist durch $v \mapsto v \otimes 1$ gegeben. Also ist φ ein Isomorphismus.

ii) Die Abbildung

$$\begin{aligned} V \times W &\longrightarrow W \otimes V, \\ (v, w) &\longmapsto w \otimes v, \end{aligned}$$

ist bilinear, und die universelle Eigenschaft liefert eine lineare Abbildung $\varphi : V \otimes W \to W \otimes V$ mit

$$\varphi(v \otimes w) = w \otimes v.$$

Ebenso konstruiert man eine Umkehrabbildung zu φ.

iii) Bei festem $u \in U$ betrachte man die Abbildung

$$\begin{aligned} V \times W &\longrightarrow (U \otimes V) \otimes W, \\ (v, w) &\longmapsto (u \otimes v) \otimes w. \end{aligned}$$

Aufgrund von 6.1.1 ii) ist sie bilinear. Wegen der universellen Eigenschaft gibt es daher eine lineare Abbildung $\varphi_u : V \otimes W \to (U \otimes V) \otimes W$ mit der Eigenschaft

$$\varphi_u(v \otimes w) = (u \otimes v) \otimes w.$$

Man prüft nun leicht nach, daß die Abbildung

$$\begin{aligned} U \times (V \otimes W) &\longrightarrow (U \otimes V) \otimes W, \\ (u, x) &\longmapsto \varphi_u(x), \end{aligned}$$

bilinear ist. Erneutes Ausnutzen der universellen Eigenschaft liefert deshalb eine lineare Abbildung

$$\varphi : U \otimes (V \otimes W) \longrightarrow (U \otimes V) \otimes W$$

mit der Eigenschaft

$$\varphi(u \otimes (v \otimes w)) = (u \otimes v) \otimes w.$$

Auf die gleiche Weise konstruiert man eine Umkehrabbildung zu φ. Also ist φ ein Isomorphismus. ∎

Aufgrund der Aussage iii) in dieser Feststellung dürfen die Klammern bei mehrfachen Tensorprodukten weggelassen werden. Insbesondere definieren wir

$$V^{\otimes n} := V \otimes \ldots \otimes V \qquad (n \text{ Faktoren}).$$

Bemerkung 6.1.4 Das Tensorprodukt läßt sich allgemeiner für Moduln über kommutativen Ringen definieren. In der Tat kann dies wörtlich wie oben geschehen, da wir bisher nur die Ringeigenschaften des Körpers K ausgenutzt haben. Das ändert sich im folgenden.

Tensorprodukt und Basis

Feststellung 6.1.5 *Sei $(v_i)_{i\in I}$ eine Basis von V, und $(w_j)_{j\in J}$ eine Basis von W. Dann ist*

$$(v_i \otimes w_j)_{i\in I,\, j\in J}$$

eine Basis von $V \otimes W$.

Beweis: Sei $v \otimes w \in V \otimes W$. Drückt man v in der Basis (v_i) und w in der Basis (w_j) aus, so erhält man nach „ausmultiplizieren" eine Linearkombination der Vektoren $v_i \otimes w_j$. Nach Feststellung 6.1.1 i) bilden diese Vektoren also jedenfalls ein Erzeugendensystem von $V \otimes W$. Wir zeigen jetzt die lineare Unabhängigkeit der $v_i \otimes w_j$. Sei dazu

$$\sum_{i,j} a_{ij}\, v_i \otimes w_j = 0 \tag{6.4}$$

eine Linearkombination (nur endlich viele a_{ij} sind ungleich Null). Sei $i_0 \in I$ und $j_0 \in J$ fest. Sei $\beta:\, V \times W \to K$ die bilineare Abbildung mit $\beta(v_{i_0}, w_{i_0}) = 1$ und $\beta(v_i, w_j) = 0$ für alle $(i,j) \neq (i_0, j_0)$ (Bemerkung 5.1.2). Sei φ die gemäß Feststellung 6.1.2 zu β korrespondierende lineare Abbildung $V \otimes W \to K$. Es gilt

$$\varphi(v_i \otimes w_j) = \begin{cases} 1, & \text{falls } (i,j) = (i_0, j_0), \\ 0, & \text{falls } (i,j) \neq (i_0, j_0). \end{cases}$$

Anwenden von φ auf die Gleichung (6.4) liefert daher $a_{i_0,j_0} = 0$. So zeigt man, daß alle Koeffizienten verschwinden. ∎

Folgerung 6.1.6 *$V \otimes W$ ist genau dann endlich-dimensional, wenn V und W es sind, und in diesem Fall gilt*

$$\dim(V \otimes W) = \dim(V)\, \dim(W).$$

Folgerung 6.1.7 *Sind V und W endlich-dimensionale K-Vektorräume, so gibt es einen Isomorphismus*

$$\varphi:\; V^* \otimes W \longrightarrow \mathrm{Hom}_K(V, W)$$

mit der Eigenschaft

$$\varphi(f \otimes w)(v) = f(v)w \qquad \textit{für alle } f \in V^*,\ w \in W,\ v \in V.$$

(V^* ist der Dualraum von V).

Beweis: Die Existenz von φ folgt, indem man die universelle Eigenschaft auf die bilineare Abbildung $V^* \times W \to \mathrm{Hom}(V, W)$ anwendet, die (f, w) auf den Homomorphismus

$$\begin{aligned} L_{f,w}:\; V &\longrightarrow W, \\ v &\longmapsto f(v)w \end{aligned}$$

abbildet. Sei $(v_1, \dots, v_n)$ eine Basis von V und $(w_1, \dots, w_m)$ eine Basis von W. Sei $(f_1, \dots, f_n)$ die zu $(v_1, \dots, v_n)$ duale Basis. Wir behaupten, daß die Elemente $L_{ij} := \varphi(f_i \otimes w_j)$ linear unabhängig sind. In der Tat, sei

$$\sum_{i,j} a_{ij} L_{ij} = 0$$

mit $c_{ij} \in K$. Für festes i_0 ist dann insbesondere

$$0 = \sum_{i,j} a_{ij} L_{ij}(v_{i_0}) = \sum_{i,j} a_{ij} f_i(v_{i_0}) w_j = \sum_j a_{i_0 j} w_j.$$

Da die w_j linear unabhängig sind, folgt $a_{i_0 j} = 0$ für alle j. Also sind alle Koeffizienten null, und wir haben die lineare Unabhängigkeit der L_{ij} gezeigt. Das Bild von φ ist also mindestens (mn)-dimensional. Da beide Räume $V^* \otimes W$ und $\mathrm{Hom}_K(V, W)$ die Dimension mn haben, ist φ ein Isomorphismus. ∎

Funktoreigenschaften

Feststellung 6.1.8 *Sind $f_1 : V_1 \to W_1$ und $f_2 : V_2 \to W_2$ lineare Abbildungen, so induzieren diese eine lineare Abbildung*

$$f_1 \otimes f_2 : V_1 \otimes V_2 \longrightarrow W_1 \otimes W_2$$

mit der Eigenschaft

$$(f_1 \otimes f_2)(v_1 \otimes v_2) = f_1(v_1) \otimes f_2(v_2).$$

Beweis: Dies folgt wieder aus der universellen Eigenschaft; man startet mit der bilinearen Abbildung $(v_1, v_2) \mapsto f_1(v_1) \otimes f_2(v_2)$. ∎

Im speziellen Fall $V_2 = W_2 = V$ und $f_2 = \mathrm{id}_V$ schreibt man auch $f_1 \otimes V$ anstatt $f_1 \otimes \mathrm{id}_V$.

Feststellung 6.1.9 *Sei*

$$W_1 \xrightarrow{f} W_2 \xrightarrow{g} W_3$$

eine exakte Sequenz linearer Abbildungen, und V ein weiterer K-Vektorraum. Dann ist auch die Sequenz

$$W_1 \otimes V \xrightarrow{f \otimes V} W_2 \otimes V \xrightarrow{g \otimes V} W_3 \otimes V$$

exakt.

Beweis: Es ist leicht zu sehen, daß $\mathrm{im}(f \otimes V) \subset \ker(g \otimes V)$. Sei umgekehrt $\sum w_i \otimes v_i \in \ker(g \otimes V)$, d.h.

$$\sum g(w_i) \otimes v_i = 0. \tag{6.5}$$

Wir dürfen annehmen, daß die $v_i \in V$ Elemente einer Basis von V sind (Feststellung 6.1.5). Sei (λ_i) die dazu duale Basis von V^*. Wendet man $\mathrm{id}_{W_3} \otimes \lambda_j : W_3 \otimes V \to W_3 \otimes K \simeq W_3$ auf (6.5) an, so erhält man

$$g(w_j) = 0.$$

Wegen der Exaktheit der Ausgangssequenz gibt es dann $w_i' \in W_1$ mit $f(w_i') = w_i$. Das Element $\sum w_i' \otimes v_i$ wird also unter $f \otimes V$ auf $\sum w_i \otimes v_i$ abgebildet. ∎

Bemerkung 6.1.10 Sei V ein fest gewählter K-Vektorraum. Die Operation $\_ \otimes V$ definiert dann einen *Funktor* von der Kategorie $\mathbf{Vec}_K$ nach $\mathbf{Vec}_K$, siehe Anhang B.2. Feststellung 6.1.9 drückt die sogenannte *Exaktheit* dieses Funktors aus. Für allgemeine Moduln über Ringen gilt keine solche Aussage.

6.2 Multilinearformen

Alle Vektorräume in diesem Abschnitt sind *endlich-dimensional.*

Symmetrisierung und Antisymmetrisierung

Wir betrachten das k-fache Tensorprodukt $V^{\otimes k}$ eines K-Vektorraumes V. Die symmetrische Gruppe $\mathcal{S}_k$ *operiert* auf $V^{\otimes k}$ durch Permutation der Komponenten. Genauer: Zu jedem $\sigma \in \mathcal{S}_k$ gibt es eine lineare Abbildung $V^{\otimes k} \to V^{\otimes k}$, ebenfalls mit σ bezeichnet, die durch

$$\sigma(v_1 \otimes \ldots \otimes v_k) = v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)}$$

gekennzeichnet ist (dies erkennt man mit Hilfe der universellen Eigenschaft des Tensorprodukts). Dem neutralen Element $e \in \mathcal{S}_k$ wird dabei die Identität auf V zugeordnet. Offenbar gilt

$$(\sigma\tau)(x) = \sigma(\tau(x)) \qquad \text{für alle } x \in V^{\otimes k},$$

und wir erhalten einen *Gruppenhomomorphismus*

$$\mathcal{S}_k \longrightarrow \mathrm{GL}(V^{\otimes k}).$$

Ein Vektor $x \in V^{\otimes k}$ heißt **symmetrisch** (bzw. **antisymmetrisch**), falls

$$\sigma(x) = x \qquad (\text{bzw. } \sigma(x) = \mathrm{sgn}(\sigma)\sigma(x)) \qquad \text{für alle } \sigma \in \mathcal{S}_k.$$

Für einen Vektor $x \in V^{\otimes k}$ definieren wir seine **Symmetrisierung**

$$x_{\mathrm{sym}} := \sum_{\sigma \in \mathcal{S}_k} \sigma(x),$$

sowie seine **Antisymmetrisierung**

$$x_{\mathrm{alt}} := \sum_{\sigma \in \mathcal{S}_k} \mathrm{sgn}(\sigma)\, \sigma(x).$$

Dies definiert Endomorphismen

$$\mathrm{sym}: V^{\otimes k} \longrightarrow V^{\otimes k} \qquad \text{und} \qquad \mathrm{alt}: V^{\otimes k} \longrightarrow V^{\otimes k}.$$

Das Bild von sym (bzw. alt) besteht aus symmetrischen (bzw. antisymmetrischen) Vektoren. Wir betrachten die Untervektorräume

$$I_s^k := \langle x - \sigma(x) : x \in V^{\otimes k} \rangle, \tag{6.6}$$

$$I_a^k := \langle v_1 \otimes \ldots \otimes v_k \ : \ v_i = v_j \text{ für irgendwelche } i \neq j \rangle \tag{6.7}$$

von $V^{\otimes k}$; dies sind die k-graduierten Anteile der definierenden Ideale für die *symmetrische Algebra* bzw. die *alternierende Algebra* von V, siehe (6.15) und (6.16) im nächsten Abschnitt. Wir definieren ferner

$$\mathbf{S}^k V := V^{\otimes k}/I_s^k, \qquad \bigwedge^k V := V^{\otimes k}/I_a^k;$$

dies sind die k-graduierten Anteile der symmetrischen bzw. alternierenden Algebra von V, die im nächsten Abschnitt definiert wird. Wie man sich leicht überlegt, gilt im Falle $\mathrm{char}(K) \neq 2$

$$I_a^k = \langle \mathrm{sgn}(\sigma)x - \sigma(x) : x \in V^{\otimes k} \rangle, \tag{6.8}$$

in Analogie zu (6.6); vgl. auch Feststellung 3.3.2.

Feststellung 6.2.1 *Im Falle* $\mathrm{char}(K) = 0$ *gilt*

$$\ker(\mathrm{sym}) = I_s^k, \qquad \ker(\mathrm{alt}) = I_a^k$$

sowie

$$V^{\otimes k} \ = \ (V^{\otimes k})_{\mathrm{sym}} \oplus I_s^k \ = \ (V^{\otimes k})_{\mathrm{alt}} \oplus I_a^k.$$

Folglich ist

$$\mathbf{S}^k V \simeq (V^{\otimes k})_{\mathrm{sym}} \qquad \textit{und} \qquad \bigwedge^k V \simeq (V^{\otimes k})_{\mathrm{alt}}$$

als K-Vektorräume.

Beweis: Es ist klar, daß $I_s^k \subset \ker(\mathrm{sym})$ und $I_a^k \subset \ker(\mathrm{alt})$. Ferner gilt

$$k!\,x - x_{\mathrm{sym}} = \sum_{\sigma \in \mathcal{S}_k} (x - \sigma(x)) \in I_s^k, \qquad k!\,x - x_{\mathrm{alt}} = \sum_{\sigma \in \mathcal{S}_k} (x - \mathrm{sgn}(\sigma)\sigma(x)) \in I_a^k,$$

und daher $V^{\otimes k} \ = \ (V^{\otimes k})_{\mathrm{sym}} + I_s^k \ = \ (V^{\otimes k})_{\mathrm{alt}} + I_a^k$. Daraus folgen bereits die ersten Aussagen. Die letzten sind dann klar nach Definition von $\mathbf{S}^k V$ bzw. $\bigwedge^k V$. ∎

Einige Paarungen

Unter einer **Paarung** zweier endlich-dimensionaler K-Vektorräume V, W versteht man eine in beiden Variablen nicht-ausgeartete Bilinearform

$$\langle\, , \rangle : V \times W \longrightarrow K. \tag{6.9}$$

Die Abbildung $V \to W^*$, die einem $v \in V$ die Linearform $w \mapsto \langle v, w\rangle$ zuordnet, ist wegen der Nicht-Ausgeartetheit injektiv. Ebenso bekommt man eine injektive lineare Abbildung $W \to V^*$. Also haben V und W die gleiche Dimension, und wir haben sogar *Isomorphismen*

$$V \xrightarrow{\sim} W^*, \qquad W \xrightarrow{\sim} V^*.$$

Kurz gesagt: Eine Paarung zwischen zwei endlich-dimensionalen Vektorräumen induziert immer Isomorphismen zwischen einem dieser Vektorräume und dem Dualraum des anderen.

Ist die Paarung (6.9) gegeben, und ein Untervektorraum U von V, so bezeichnen wir mit

$$U^\perp = \{w \in W : \langle u, w\rangle = 0 \text{ für alle } u \in U\}$$

das **orthogonale Komplement** von U. Die ursprüngliche Paarung induziert dann eine neue Paarung

$$U \times W/U^\perp \longrightarrow K, \tag{6.10}$$

und somit einen Isomorphismus

$$U \xrightarrow{\sim} (W/U^\perp)^*.$$

Feststellung 6.2.2 *Sei V ein endlich-dimensionaler K-Vektorraum. Für jedes $k \in \mathbb{N}$ haben wir dann eine Paarung*

$$\langle\, , \rangle : (V^*)^{\otimes k} \times V^{\otimes k} \longrightarrow K, \tag{6.11}$$

charakterisiert durch

$$\langle \lambda_1 \otimes \ldots \otimes \lambda_k, v_1 \otimes \ldots \otimes v_k\rangle = \lambda_1(v_1) \cdot \ldots \cdot \lambda_k(v_k).$$

Folglich ist $(V^*)^{\otimes k} \simeq (V^{\otimes k})^*$. Ferner gilt

$$\langle \sigma\lambda, \sigma x\rangle = \langle \lambda, x\rangle \qquad \text{für alle } \lambda \in (V^*)^{\otimes k},\ x \in V^{\otimes k},\ \sigma \in \mathcal{S}_k. \tag{6.12}$$

Beweis: Sei $v_1, \ldots, v_n$ eine Basis von V, und $\lambda_1, \ldots, \lambda_n$ die dazu duale Basis von V^*. Dann definiert die Vorschrift

$$\langle \lambda_{i_1} \otimes \ldots \otimes \lambda_{i_k}, v_{j_1} \otimes \ldots \otimes v_{j_k}\rangle = \lambda_{i_1}(v_{j_1}) \cdot \ldots \cdot \lambda_{i_k}(v_{j_k}) \tag{6.13}$$

wegen 6.1.5 und Bemerkung 5.1.2 eine Bilinearform $(V^*)^{\otimes k} \times V^{\otimes K} \to K$. Der Ausdruck auf der rechten Seite von (6.13) ist 1, falls $i_1 = j_1, \ldots, i_k = j_k$, und 0 sonst. Dieses benutzend prüft man leicht nach, daß die Bilinearform in beiden Variablen nicht ausgeartet ist.

Eigenschaft (6.12) braucht man nur auf den Basiselementen nachzuprüfen, und dort ist sie offensichtlich. ∎

Folgerung 6.2.3 *Die Paarung (6.11) induziert Paarungen*

$$(V^*)^{\otimes k}_{\mathrm{sym}} \times V^{\otimes k}/I^k_s \longrightarrow K, \qquad (V^*)^{\otimes k}_{\mathrm{alt}} \times V^{\otimes k}/I^a_s \longrightarrow K,$$

wobei wir im alternierenden Fall $\mathrm{char}(K) \neq 2$ *voraussetzen. Im Falle* $\mathrm{char}(K) = 0$ *ist folglich*

$$\mathbf{S}^k(V^*) \simeq (\mathbf{S}^k V)^*, \qquad \bigwedge^k(V^*) \simeq (\bigwedge^k V)^*.$$

Beweis: Wir beweisen nur die „symmetrischen" Aussagen; der alternierende Fall geht analog. Für die erste Behauptung ist zu zeigen, daß das orthogonale Komplement von $(V^*)^{\otimes k}_{\mathrm{sym}}$ gleich I^k_s ist, vgl. (6.10). Dies folgt aus (6.12):

$$\begin{aligned} \lambda \text{ symmetrisch} &\iff \sigma\lambda = \lambda && \text{für alle } \sigma \in \mathcal{S}_k \\ &\iff \langle \lambda, \sigma x \rangle = \langle \lambda, x \rangle && \text{für alle } \sigma \in \mathcal{S}_k,\ x \in V^{\otimes k}, \\ &\iff \langle \lambda, \sigma x - x \rangle = 0 && \text{für alle } \sigma \in \mathcal{S}_k,\ x \in V^{\otimes k}, \\ &\iff \langle \lambda, I^k_s \rangle = 0. \end{aligned}$$

Der Isomorphismus $\mathbf{S}^k(V^*) \simeq (\mathbf{S}^k V)^*$ ergibt sich nun mit Hilfe von Feststellung 6.2.1:

$$\mathbf{S}^k(V^*) \simeq (V^*)^{\otimes k}_{\mathrm{sym}} \simeq ((V^{\otimes k})/I^k_s)^* = (\mathbf{S}^k V)^*.$$

■

Universelle Eigenschaften

Eine **Multilinearform** vom Grade k auf V ist eine Abbildung

$$V^k \longrightarrow K,$$

die in jeder Variablen linear ist, siehe Definition 3.3.1. Etwas allgemeiner als Multilinearformen betrachten wir **multilineare Abbildungen**, d.h. Abbildungen

$$\mu : \ V \times \ldots \times V \longrightarrow Y,$$

die in jeder Variablen linear sind; hierbei ist Y ein beliebiger K-Vektorraum. Die multilineare Abbildung μ heißt **symmetrisch**, falls

$$\mu(v_1, \ldots, v_k) = \mu(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \qquad \text{für alle } \sigma \in \mathcal{S}_k.$$

Sie heißt **alternierend**, falls

$$\mu(v_1, \ldots, v_k) = 0 \qquad \text{wann immer } v_i = v_j \text{ für irgendwelche } i \neq j.$$

Im Falle $\mathrm{char}(K) = 0$ ist das Alternieren äquivalent mit

$$\mu(v_1, \ldots, v_k) = \mathrm{sgn}(\sigma)\, \mu(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \qquad \text{für alle } \sigma \in \mathcal{S}_k.$$

Als Symbole für den Vektorraum der (symmetrischen, alternierenden) multilinearen Abbildungen $V^k \to Y$ verwenden wir

$$\mathrm{Mult}^k(V, Y) \quad \text{bzw.} \quad \mathrm{Sym}^k(V, Y) \quad \text{bzw.} \quad \mathrm{Alt}^k(V, Y).$$

Es gelten nun die folgenden *universellen Eigenschaften.*

Feststellung 6.2.4 *Seien V und Y endlich-dimensionale K-Vektorräume.*

i) Zu jeder multilinearen Abbildung $\mu : V^k \to Y$ gibt es genau eine lineare Abbildung $\varphi : V^{\otimes k} \to Y$, die das folgende Diagramm kommutativ macht:

$$\begin{tikzcd} V^k \arrow[r, "\mu"] \arrow[d] & Y \\ V^{\otimes k} \arrow[ur, "\varphi"'] & \end{tikzcd}$$

Folglich ist

$$\mathrm{Mult}^k(V,Y) \simeq \mathrm{Hom}(V^{\otimes k}, Y) \simeq (V^{\otimes k})^* \otimes Y.$$

ii) Zu jeder symmetrischen multilinearen Abbildung $\mu : V^k \to Y$ gibt es genau eine lineare Abbildung $\varphi : \mathbf{S}^k V \to Y$, die das folgende Diagramm kommutativ macht:

$$\begin{tikzcd} V^k \arrow[r, "\mu"] \arrow[d] & Y \\ \mathbf{S}^k V \arrow[ur, "\varphi"'] & \end{tikzcd}$$

(die vertikale Abbildung ist die Komposition $V^k \to V^{\otimes k} \to \mathbf{S}^k V$). Folglich ist

$$\mathrm{Sym}^k(V,Y) \simeq \mathrm{Hom}(\mathbf{S}^k V, Y) \simeq (\mathbf{S}^k V)^* \otimes Y.$$

iii) Zu jeder alternierenden multilinearen Abbildung $\mu : V^k \to Y$ gibt es genau eine lineare Abbildung $\varphi : \bigwedge^k V \to Y$, die das folgende Diagramm kommutativ macht:

$$\begin{tikzcd} V^k \arrow[r, "\mu"] \arrow[d] & Y \\ \bigwedge^k V \arrow[ur, "\varphi"'] & \end{tikzcd}$$

(die vertikale Abbildung ist die Komposition $V^k \to V^{\otimes k} \to \bigwedge^k V$). Folglich ist

$$\mathrm{Alt}^k(V,Y) \simeq \mathrm{Hom}(\bigwedge^k V, Y) \simeq (\bigwedge^k V)^* \otimes Y.$$

Beweis: Dies sind nunmehr Standard-Argumente; im wesentlichen nutzt man wiederholt die universelle Eigenschaft 6.1.2 aus. Außerdem beachte man Feststellung 6.1.7. ∎

Unter Beachtung von 6.2.1, 6.2.2 und 6.2.3 erhält man:

Folgerung 6.2.5 *Es gilt* $\mathrm{Mult}^k(V) \simeq (V^*)^{\otimes k}$*. Für* $\mathrm{char}(K) = 0$ *gilt zusätzlich*

$$\mathrm{Sym}^k(V) \simeq (V^*)^{\otimes k}_{\mathrm{sym}} \simeq \mathbf{S}^k(V^*),$$
$$\mathrm{Alt}^k(V) \simeq (V^*)^{\otimes k}_{\mathrm{alt}} \simeq \textstyle\bigwedge^k(V^*).$$

Explizit: Das Element $\lambda_1 \otimes \ldots \otimes \lambda_k \in (V^*)^{\otimes k}$ *entspricht der Multilinearform*

$$(v_1, \ldots, v_n) \longmapsto \lambda_1(v_1) \cdot \ldots \cdot \lambda_k(v_k).$$

Symmetrische (bzw. antisymmetrische) Elemente von $(V^*)^{\otimes k}$ *entsprechen symmetrischen (bzw. antisymmetrischen) Multilinearformen.*

Das Bild des Elementes $v_1 \otimes \ldots \otimes v_k \in V^{\otimes k}$ unter der Projektion

$$V^{\otimes k} \longrightarrow V^{\otimes k}/I_a^k = \textstyle\bigwedge^k V$$

wird mit $v_1 \wedge \ldots \wedge v_k$ bezeichnet. Explizit entspricht das Element

$$\lambda_1 \wedge \ldots \wedge \lambda_k \in \textstyle\bigwedge^k(V^*)$$

dann der folgenden alternierenden Multilinearform auf V: Zunächst antisymmetrisiert man und erhält das Element

$$\sum_{\sigma \in \mathcal{S}_k} \mathrm{sgn}(\sigma)\, \lambda_{\sigma(1)} \otimes \ldots \otimes \lambda_{\sigma(k)} \in V^{\otimes k}.$$

Die entsprechende Multilinearform μ auf V ist dann gegeben durch

$$\begin{aligned}\mu(v_1, \ldots, v_k) &= \sum_{\sigma \in \mathcal{S}_k} \mathrm{sgn}(\sigma)\, \lambda_{\sigma(1)}(v_1) \cdot \ldots \cdot \lambda_{\sigma(k)}(v_k) \\ &= \sum_{\sigma \in \mathcal{S}_k} \mathrm{sgn}(\sigma)\, \lambda_1(v_{\sigma(1)}) \cdot \ldots \cdot \lambda_k(v_{\sigma(k)}).\end{aligned} \tag{6.14}$$

Z.B. entspricht $\lambda_1 \wedge \lambda_2 \in \bigwedge^2 V$ der alternierenden Multilinearform

$$\mu(v_1, v_2) = \lambda_1(v_1)\lambda_2(v_2) - \lambda_2(v_1)\lambda_1(v_2).$$

6.3 Die Tensoralgebra

Graduierte Algebren

Definition 6.3.1 *Eine* K*-Algebra* $\mathcal{A}$ *(siehe Definition 1.3.14) heißt* **graduiert**, *wenn sie sich als direkte Summe*

$$\mathcal{A} = \bigoplus_{n=0}^{\infty} \mathcal{A}_n$$

von Vektorräumen $\mathcal{A}_n$ schreiben läßt, so daß die Multiplikation auf $\mathcal{A}$ für jedes m, n eine bilineare Abbildung

$$\mathcal{A}_m \times \mathcal{A}_n \longrightarrow \mathcal{A}_{m+n}$$

induziert. Ein Element von $\mathcal{A}_n$ heißt **homogen** *vom Grad n, und $\mathcal{A}_n$ heißt auch der* **n-graduierte Anteil** *von $\mathcal{A}$.*

In einer graduierten Algebra $\mathcal{A}$ ist das Produkt eines Elementes vom Grad m mit einem Element vom Grad n also vom Grad $m+n$. Elemente, die nicht in einem $\mathcal{A}_n$ liegen, bekommen keinen Grad. Der 0-graduierte Anteil $\mathcal{A}_0$ ist offenbar selbst eine K-Algebra, während $\mathcal{A}_n$ für $n \geq 1$ kein Ring ist.

Beispiele: 1. Die Polynomalgebra $K[X]$ ist eine graduierte Algebra. Hier ist $K[X]_n = KX^n$ eindimensional.

2. Allgemeiner ist die Polynomalgebra in n Variablen $K[X_1, \ldots, X_n]$ eine graduierte Algebra. Der k-graduierte Anteil wird von allen Monomen $X_1^{i_1} \cdot \ldots \cdot X_n^{i_n}$ mit $i_1 + \ldots + i_n = k$ aufgespannt.

Ein Homomorphismus von graduierten K-Algebren $f : \mathcal{A} \to \mathcal{B}$ heißt **graduiert**, falls $f(\mathcal{A}_n) \subset \mathcal{B}_n$, falls also homogene Elemente vom Grad n auf ebensolche abgebildet werden. Die graduierten K-Algebren, zusammen mit den graduierten Homomorphismen, bilden eine *Kategorie*, siehe Anhang B.1.

Die Tensoralgebra

Sei V ein beliebiger K-Vektorraum. Wir definieren

$$V^{\otimes 0} := K, \qquad V^{\otimes 1} := V, \qquad V^{\otimes n} = \underbrace{V \otimes \ldots \otimes V}_{n\text{-mal}},$$

und betrachten den Vektorraum

$$\mathbf{T}V := \bigoplus_{n=0}^{\infty} V^{\otimes n}.$$

Auf $\mathbf{T}V$ läßt sich wie folgt eine Multiplikation definieren. Unter Ausnutzung der universellen Eigenschaft des Tensorprodukts sieht man leicht, daß es eine lineare Abbildung

$$\varphi_{m,n} : \; V^{\otimes m} \otimes V^{\otimes n} \longrightarrow V^{\otimes (m+n)}$$

mit

$$\varphi((v_1 \otimes \ldots \otimes v_m) \otimes (v_1' \otimes \ldots \otimes v_n')) = v_1 \otimes \ldots \otimes v_m \otimes v_1' \otimes \ldots \otimes v_n'$$

gibt. $\varphi_{m,n}(x, y)$ ist bereits das Produkt der speziellen Elemente $x \in V^{\otimes m}$ und $y \in V^{\otimes n}$ in $\mathbf{T}V$. Beliebige Elemente $x, y \in \mathbf{T}V$ werden multipliziert, indem man sie in ihre homogenen Anteile zerlegt,

$$x = x_0 + x_1 + \ldots + x_n, \qquad y = y_0 + y_1 + \ldots + y_n, \qquad x_i, y_i \in V^{\otimes i},$$

und „ausmultipliziert“, also

$$x \cdot y := \sum_{i,j} \varphi_{i,j}(x_i, y_j).$$

Feststellung und Definition 6.3.2 *Der Vektorraum* $\mathbf{T}V$ *wird mit der oben definierten Multiplikation zu einer graduierten* K*-Algebra. Sie heißt die* **Tensoralgebra** *von* V*.*

Beweis: Dies ist leicht nachzuprüfen. ■

Bemerkung 6.3.3 Ist $f : V \to W$ eine lineare Abbildung, so induziert f lineare Abbildungen $V^{\otimes n} \to W^{\otimes n}$, die sich zu einer Abbildung

$$\mathbf{T}f : \mathbf{T}V \longrightarrow \mathbf{T}W$$

zusammensetzen. Man sieht leicht, daß $\mathbf{T}f$ sogar ein Homomorphismus von graduierten K-Algebren ist. Auf diese Weise erhält man einen *Funktor* von $\mathbf{Vec}_K$ in die Kategorie der graduierten K-Algebren (siehe Anhang B.2).

Ideale und Quotienten

Wir wollen gewisse Quotienten der Tensoralgebra bilden. Damit ein Quotient $\mathcal{A}/I$ einer Algebra $\mathcal{A}$ wieder eine Algebra ist, muß der Untervektorraum I ein *Ideal* sein:

Definition 6.3.4 *Sei* I *ein Untervektorraum der* K*-Algebra* $\mathcal{A}$*. Dann heißt* I *ein* **Ideal**, *falls*

$$ax \in I \quad \text{und} \quad xa \in I \qquad\qquad \text{für alle } a \in \mathcal{A},\ x \in I$$

(man vergleiche mit Definition 4.1.7 für einen kommutativen Ring).

Feststellung 6.3.5 *Ist* $\mathcal{A}$ *ein* K*-Algebra und* $I \subset \mathcal{A}$ *ein Ideal, so ist der Quotientenvektorraum* $\mathcal{A}/I$ *in natürlicher Weise wieder eine* K*-Algebra, und die Projektion* $\mathcal{A} \to \mathcal{A}/I$ *ist ein Homomorphismus von* K*-Algebren.*

Beweis: Sei $\bar{a}$ das Bild von $a \in \mathcal{A}$ unter der Projektion. Wir definieren das Produkt von $\bar{a}$ und $\bar{b}$ als

$$\bar{a} \cdot \bar{b} := \overline{ab}.$$

Unter Ausnutzung der Idealeigenschaften prüft man zunächst nach, daß dies wohldefiniert ist. Alles andere folgt dann sehr schnell. ■

Ist speziell $\mathcal{A}$ eine *graduierte* Algebra, so benötigt man eine Zusatzbedingung an I, damit die Quotientenalgebra $\mathcal{A}/I$ wieder graduiert ist:

Feststellung und Definition 6.3.6 *Sei* $\mathcal{A} = \bigoplus \mathcal{A}_n$ *eine graduierte* K*-Algebra. Ein Ideal* $I \subset \mathcal{A}$ *heißt* **graduiert**, *falls die folgenden äquivalenten Bedingungen erfüllt sind:*

i) $I = \bigoplus_{n=0}^{\infty} (I \cap \mathcal{A}_n)$.

ii) I wird von homogenen Elementen erzeugt (als Ideal).

Beweis: i) $\Rightarrow$ ii) ist klar. Für ii) $\Rightarrow$ i) beachte man zunächst, daß die Inklusion

$$I \supset \bigoplus (I \cap \mathcal{A}_n)$$

trivialerweise immer erfüllt ist. Ist $x \in I$, so läßt sich x nach Voraussetzung als

$$x = a_1 x_1 b_1 + \ldots + a_n x_n b_n \qquad \text{mit } a_i, b_i \in \mathcal{A} \text{ und homogenen } x_i \in I$$

schreiben. Jedes a_i (bzw. b_i) ist ferner eine Summe homogener Elemente a_{ij} (bzw. b_{ik}). Insgesamt ist dann $x = \sum_{i,j,k} a_{ij} x_i b_{ik}$, wobei der Summand $a_{ij} x_i b_{ik}$ sowohl homogen als auch ein Element von I ist. Dies zeigt die umgekehrte Inklusion. ∎

Ist I ein graduiertes Ideal der graduierten Algebra $\mathcal{A} = \bigoplus \mathcal{A}_n$, so ist offenbar

$$\mathcal{A}/I = \bigoplus_{n=0}^{\infty} \mathcal{A}_n/(I \cap \mathcal{A}_n).$$

Die Quotientenalgebra besitzt also wieder eine Zerlegung als direkte Summe, und es ist schnell zu sehen, daß dadurch $\mathcal{A}/I$ in der Tat selbst zu einer graduierten K-Algebra wird.

Die symmetrische Algebra

Wir betrachten wieder die Tensoralgebra $\mathbf{T}V$ des K-Vektorraumes V. Sei $I_s \subset \mathbf{T}V$ der Untervektorraum, der von den Elementen

$$v_1 \otimes \ldots \otimes v_k - v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)} \tag{6.15}$$

aufgespannt wird, wobei $\sigma \in \mathcal{S}_k$ irgendeine Permutation ist. Es ist klar, daß I_s sogar ein *Ideal* in $\mathbf{T}V$ ist, und zwar ein graduiertes Ideal, da die Elemente (6.15) homogen sind. Offenbar ist

$$I_s = \bigoplus_{k=0}^{\infty} I_s^k,$$

wobei wir I_s^k bereits in (6.6) definiert hatten.

Definition 6.3.7 *Die graduierte K-Algebra $\mathbf{S}V := \mathbf{T}V/I_s$ heißt die* **symmetrische Algebra** *des K-Vektorraumes V.*

Feststellung 6.3.8 *Die K-Algebra $\mathbf{S}V$ ist kommutativ. Ist V endlich-dimensional und $v_1, \ldots, v_n$ eine Basis von V, so ist $\mathbf{S}V$ isomorph zum Polynomring über K in n Variablen:*

$$\begin{aligned} \mathbf{S}V &\xrightarrow{\sim} K[X_1, \ldots, X_n], \\ v_i &\longleftrightarrow X_i. \end{aligned}$$

Beweis: Das Ideal I_s von $\mathbf{T}V$ war gerade so definiert, daß der Quotient $\mathbf{S}V$ kommutativ ist. Zur zweiten Behauptung: Unter Benutzung der universellen Eigenschaft 6.1.2 konstruiert man zunächst eine lineare Abbildung

$$\varphi : \mathbf{T}V \longrightarrow K[X_1, \ldots, X_n]$$

mit

$$\varphi(v_{i_1} \otimes \ldots \otimes v_{i_k}) = X_{i_1} \cdot \ldots \cdot X_{i_k}.$$

Die Elemente (6.15) liegen wegen der Kommutativität des Polynomrings im Kern von φ. Deshalb induziert φ eine (ebenfalls so genannte) lineare Abbildung

$$\varphi : \mathbf{S}V \longrightarrow K[X_1, \ldots , X_n]$$

mit $\varphi(v_i) = X_i$. Aufgrund der Definition der Multiplikation im Ring $\mathbf{S}V$ ist klar, daß φ ein Homomorphismus von K-Algebren ist. Aber man kann leicht eine Umkehrabbildung zu φ angeben. Aufgrund der universellen Eigenschaft des Polynomrings (vgl. 4.1.3) gibt es nämlich einen K-Algebren-Homomorphismus $K[X_1, \ldots , X_n] \to \mathbf{S}V$, der X_i auf v_i abbildet. ■

Die alternierende Algebra

Sei $I_a \subset \mathbf{T}V$ der Untervektorraum, der von Elementen der Form

$$v_1 \otimes \ldots \otimes v_k \qquad \text{mit } v_i = v_j \text{ für irgendwelche } i \neq j \tag{6.16}$$

aufgespannt wird. Offensichtlich ist I_a sogar ein *Ideal* in $\mathbf{T}V$, und zwar ein graduiertes Ideal. Es gilt

$$I_a = \bigoplus_{k=0}^{\infty} I_a^k$$

mit I_a^k wie in (6.7).

Definition 6.3.9 *Die graduierte K-Algebra $\bigwedge V := \mathbf{T}V/I_a$ heißt die* **alternierende Algebra** *des K-Vektorraumes V.*

Das Bild des Elementes $v_1 \otimes \ldots \otimes v_k \in \mathbf{T}V$ im Quotienten $\bigwedge V$ wird mit

$$v_1 \wedge \ldots \wedge v_k$$

bezeichnet. Wendet man auf die Gleichung $(v + w) \otimes (v + w) = v \otimes v + w \otimes w + v \otimes w + w \otimes v$ die Projektion $\mathbf{T}V \to \bigwedge V$ an, so erhält man

$$v \wedge w = -w \wedge v \qquad \text{für alle } v, w \in V. \tag{6.17}$$

Daraus gewinnt man die allgemeinere Relation

$$v_1 \wedge \ldots \wedge v_k = \operatorname{sgn}(\sigma)\, v_{\sigma(1)} \wedge \ldots \wedge v_{\sigma(k)} \tag{6.18}$$

für alle Permutationen $\sigma \in \mathcal{S}_k$; man vergleiche mit der definierenden Relation (6.15) für die symmetrische Algebra.

Satz 6.3.10 *Sei $v_1, \ldots, v_n$ eine Basis von V. Die Vektoren*

$$v_{i_1} \wedge \ldots \wedge v_{i_k} \qquad \textit{mit } 1 \leq i_1 < \ldots < i_k \leq n \tag{6.19}$$

bilden eine Basis des k-graduierten Anteils $\bigwedge^k V$ von $\bigwedge V$. Insbesondere ist $\bigwedge^k V = 0$ für $k > n$. Es gilt

$$\dim(\textstyle\bigwedge^k V) = \binom{n}{k}, \qquad \dim(\textstyle\bigwedge V) = 2^n.$$

Beweis: Der k-graduierte Anteil $V^{\otimes k}$ von $\mathbf{T}V$ besitzt nach Feststellung 6.1.5 die Basis

$$v_{i_1} \otimes \ldots \otimes v_{i_k}, \qquad i_1, \ldots, i_k \in \{1, \ldots, n\}. \tag{6.20}$$

Der k-graduierte Anteil von $\bigwedge V$ wird deshalb von den Bildern $v_{i_1} \wedge \ldots \wedge v_{i_k}$ dieser Vektoren unter der Projektion aufgespannt. Wegen (6.18) genügen dazu aber bereits die Vektoren (6.19) (ist $k > n$, so läßt es sich nicht vermeiden, daß zwei der v_{i_j} gleich sind, und folglich sind alle Vektoren $v_{i_1} \wedge \ldots \wedge v_{i_k}$ null).

Wir zeigen als nächstes, daß das Element

$$x := v_1 \otimes \ldots \otimes v_n \in V^{\otimes n} \tag{6.21}$$

nicht in dem definierenden Ideal I_a für die alternierende Algebra liegt. Angenommen, dies wäre doch der Fall. Dann ist x eine Linearkombination von Elementen der Form $v'_1 \otimes \ldots \otimes v'_n$, in denen zwei der v'_i gleich sind. Wegen (6.18) dürfen wir annehmen, daß jeweils v'_1 und v'_2 gleich sind. Somit wäre x invariant unter dem Isomorphismus von $V^{\otimes n}$, der die ersten beiden Komponenten vertauscht. Dies ist aber offenbar nicht richtig.

Aus dem bisherigen folgt, daß das Element $v_1 \wedge \ldots \wedge v_n$ nicht Null ist und den Raum $\bigwedge^n V$ aufspannt. Dies ist die Behauptung für $k = n$. Es bleibt der Fall $k < n$ zu behandeln. Sei

$$\sum_{i_1 < \ldots < i_k} a_{i_1, \ldots, i_k}\, v_{i_1} \wedge \ldots \wedge v_{i_k} = 0 \tag{6.22}$$

eine Linearkombination der Vektoren (6.19), die Null darstellt. Wir wählen $j_1 < \ldots < j_k$ fest. Sei $\{j_{k+1}, \ldots, j_n\}$ das Komplement von $\{j_1, \ldots, j_k\}$ in der Menge $\{1, \ldots, n\}$. Multipliziert man die Gleichung (6.22) (in der Algebra $\bigwedge V$) mit $v_{j_{k+1}} \wedge \ldots \wedge v_{j_n}$, so überlebt nur ein Term, nämlich

$$a_{j_1, \ldots, j_k}\, v_{j_1} \wedge \ldots \wedge v_{j_k} \wedge v_{j_{k+1}} \wedge \ldots \wedge v_{j_n} = 0.$$

Aber $v_{j_1} \wedge \ldots \wedge v_{j_n}$ ist nur eine Umordnung von $v_1 \wedge \ldots \wedge v_n$, und wir haben eben gesehen, daß dieses Element nicht null ist. Es folgt $a_{j_1, \ldots, j_k} = 0$. Dies zeigt die lineare Unabhängigkeit der Elemente (6.19).

Die Behauptung über die Dimensionen ergibt sich jetzt einfach aus der kombinatorischen Bedeutung der Binomialkoeffizienten, vgl. Bemerkung 3.3.5. ∎

Bemerkung 6.3.11 Insbesondere ist $\bigwedge^n V$ eindimensional, falls $\dim(V) = n$. Aus Feststellung 6.2.4 iii) folgt dann, daß es bis auf skalare Vielfache nur eine nicht-triviale alternierende n-Form auf V gibt. Damit haben wir erneut Satz 3.2.3 über die Existenz und Eindeutigkeit der Determinante bewiesen.

Wir setzen jetzt $\mathrm{char}(K) = 0$ voraus, und betrachten $\bigwedge(V^*)$, wobei V^* der Dualraum von V ist. Der k-graduierte Anteil

$$\bigwedge^k(V^*) = (V^*)^{\otimes k}/I_a^k$$

ist nach Feststellung 6.2.4 isomorph zum Raum aller alternierenden Multilinearformen vom Grade k auf V. Die Algebrastruktur auf $\bigwedge(V^*)$ macht es daher möglich, eine k-Form mit einer l-Form zu einer $(k+l)$-Form zu multiplizieren. Explizit: Sei

$$\lambda = \lambda_1 \wedge \ldots \wedge \lambda_k \in \bigwedge^k(V^*), \qquad \eta = \lambda_{k+1} \wedge \ldots \wedge \lambda_{k+l} \in \bigwedge^l(V^*).$$

Gleichung (6.14) beschreibt die entsprechenden Multilinearformen. Das *Produkt* $\lambda \wedge \eta$ entspricht nach derselben Gleichung der Multilinearform

$$\begin{aligned}(\lambda \wedge \eta)(v_1, \ldots v_{k+l}) &= \sum_{\sigma \in \mathcal{S}_{k+l}} \mathrm{sgn}(\sigma)\, \lambda_1(v_{\sigma(1)}) \cdot \ldots \cdot \lambda_{k+l}(v_{\sigma(k+l)}) \\ &= \frac{1}{k!\, l!} \sum_{\tau_1 \in \mathcal{S}_k} \sum_{\tau_2 \in \mathcal{S}_l} \sum_{\sigma \in \mathcal{S}_{k+l}} \mathrm{sgn}(\sigma)\, \mathrm{sgn}(\tau_1)\, \mathrm{sgn}(\tau_2) \\ &\quad \lambda_1(v_{\sigma(\tau_1(1))}) \cdot \ldots \cdot \lambda_k(v_{\sigma(\tau_1(k))}) \cdot \lambda_{k+1}(v_{\sigma(k+\tau_2(1))}) \cdot \ldots \cdot \lambda_{k+l}(v_{\sigma(k+\tau_2(l))}) \\ &= \frac{1}{k!\, l!} \sum_{\sigma \in \mathcal{S}_{k+l}} \mathrm{sgn}(\sigma)\, \lambda(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\, \eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}). \end{aligned} \tag{6.23}$$

Diese Formel findet man oft in Büchern über Analysis auf Mannigfaltigkeiten; dort ist V der Tangentialraum an einem Punkt einer Mannigfaltigkeit, und (6.23) wird zur Definition des „Dachproduktes“ zweier Differentialformen benutzt.

6.4 Basiserweiterung

Definition und grundlegende Eigenschaften

Wir nehmen an, der Grundkörper K sei in einem weiteren Körper L enthalten: $K \subset L$. In diesem Fall spricht man von einer **Körpererweiterung** und schreibt L/K (dies ist kein Quotient, sondern soll nur andeuten, daß L über K liegt). Der größere Körper L ist dann offenbar ein *Vektorraum* über K. Man nennt die Dimension des K-Vektorraumes L den **Grad** der Körpererweiterung, und schreibt

$$[L : K] = \dim_K(L).$$

Im Fall $[L : K] < \infty$ spricht man von einer **endlichen** Körpererweiterung, ansonsten von einer **unendlichen**.

Beispiel: 1. $\mathbb{C}/\mathbb{R}$ ist eine Körpererweiterung vom Grad 2; eine Basis des $\mathbb{R}$-Vektorraumes $\mathbb{C}$ ist $\{1, i\}$.

2. $\mathbb{R}/\mathbb{Q}$ ist eine Körpererweiterung unendlichen Grades.

Ist eine Körpererweiterung L/K gegeben, so ermöglicht es das Tensorprodukt, K-Vektorräume zu L-Vektorräumen zu machen:

Feststellung und Definition 6.4.1 *Sei L/K eine Körpererweiterung, und V ein beliebiger K-Vektorraum. Das Tensorprodukt*

$$V_L := V \otimes_K L$$

der K-Vektorräume V und L ist dann in natürlicher Weise ein L-Vektorraum. Die Skalarmultiplikation ist durch

$$a.(v \otimes b) = v \otimes (ab) \qquad (a, b \in L,\ v \in V)$$

charakterisiert. Man sagt, V_L entsteht aus V durch **Basiserweiterung** *mit L.*

Beweis: Bei festem $a \in L$ ist die Abbildung $V \times L \to V_L$, $(v, b) \mapsto v \otimes (ab)$ K-bilinear. Aufgrund der universellen Eigenschaft 6.1.2 gibt es eine *lineare* Abbildung $\varphi_a : V_L \to V_L$ mit $\varphi_a(v \otimes b) = v \otimes (ab)$. Man prüft dann leicht nach, daß die Abbildung

$$L \times V_L \longrightarrow V_L, \qquad (a, x) \longmapsto \varphi_a(x)$$

eine Skalarmultiplikation definiert, durch die V_L zu einem L-Vektorraum wird. ∎

Feststellung 6.4.2 *Ist $(v_i)_{i \in I}$ eine K-Basis von V, so ist $(v_i \otimes 1)_{i \in I}$ eine L-Basis von V_L. Insbesondere gilt: V ist genau dann ein endlich-dimensionaler K-Vektorraum, wenn V_L ein endlich-dimensionaler L-Vektorraum ist, und in diesem Fall ist*

$$\dim_K(V) = \dim_L(V_L).$$

Beweis: Ein beliebiges $v \in V$ läßt sich als Linearkombination $\sum a_i v_i$ mit $a_i \in K$ schreiben. Das Element $v \otimes b \in V_L$ läßt sich daher schreiben als

$$v \otimes b = (\sum a_i v_i) \otimes b = \sum((a_i v_i) \otimes b) = \sum v_i \otimes (a_i b) = \sum (a_i b).(v_i \otimes 1).$$

Die Elemente $v_i \otimes 1$ spannen daher V_L auf. Um deren lineare Unabhängigkeit zu zeigen, sei $\lambda_j \in V^*$ die Linearform mit $\lambda_j(v_i) = \delta_{ij}$. Auf eine Linearkombination

$$\sum b_i(v_i \otimes 1) = 0$$

wende man die lineare Abbildung $\lambda_j \otimes \mathrm{id}_L : V \otimes L \to K \otimes L \simeq L$ an; dies liefert $b_j = 0$. ∎

Die Abbildung $v \mapsto v \otimes 1 \in V_L$ ist injektiv und K-linear. Auf diese Weise kann man V als K-Untervektorraum von V_L auffassen (natürlich nicht als L-Untervektorraum). Oft wird dann einfach v anstatt $v \otimes 1$ geschrieben. Ist etwa (v_i) eine K-Basis von V wie in der Feststellung, so besteht V aus allen K-Linearkombinationen $\sum a_i v_i$, und V_L aus allen L-Linearkombinationen $\sum b_i v_i$.

Beispiel: Sei $M(m \times n, K)$ der Raum der $m \times n$-Matrizen mit Koeffizienten aus K. Wenn E_{ij} die Matrix bezeichnet, deren (i, j)-Koeffizient 1 und alle anderen 0 sind, so bilden die E_{ij} eine K-Basis von $M(m \times n, K)$. Nach obiger Feststellung bilden die E_{ij} dann auch eine L-Basis von $M(m \times n, K)_L$. Daher ist klar

$$M(m \times n, K) \otimes_K L \simeq M(m \times n, L) \qquad \text{als } L\text{-Vektorräume.}$$

Bemerkung 6.4.3 Ist $f : V \to W$ ein Homomorphismus von K-Vektorräumen, so haben wir dazu einen Homomorphismus von L-Vektorräumen

$$f_L : V_L \longrightarrow W_L,$$

nämlich $f_L = f \otimes \mathrm{id}_L$. Auf diese Weise wird die Basiserweiterung zu einem *Funktor* von der Kategorie $\mathbf{Vec}_K$ der K-Vektorräume zu der Kategorie $\mathbf{Vec}_L$ der L-Vektorräume; siehe Anhang B.

Basiserweiterung bei Algebren

Sei jetzt $\mathcal{A}$ eine K-Algebra, und $\mathcal{A}_L$ die Basiserweiterung zu einem L-Vektorraum. Unter Ausnutzung der universellen Eigenschaft des Tensorproduktes ist leicht zu sehen, daß es eine L-bilineare Abbildung

$$\mathcal{A}_L \times \mathcal{A}_L \longrightarrow \mathcal{A}_L$$

mit der Eigenschaft

$$(x \otimes b,\, x' \otimes b') \longmapsto xx' \otimes bb'$$

gibt (xx' ist die Multiplikation in der Algebra $\mathcal{A}$). Man prüft nach, daß diese Abbildung eine Multiplikation auf $\mathcal{A}_L$ definiert, durch die $\mathcal{A}_L$ zu einer L-Algebra wird. Kurz gesagt, die Basiserweiterung einer Algebra ist wieder eine Algebra. Via der Abbildung $x \mapsto x \otimes 1$ kann man $\mathcal{A}$ als K-Unteralgebra von $\mathcal{A}_L$ auffassen.

Beispiele: 1. Wir haben oben gesehen, daß $M(n, K) \otimes L \simeq M(n, L)$ als L-Vektorräume. Dies ist nun sogar ein Isomorphismus von L-Algebren.

2. Der Polynomring $K[X]$ besitzt die K-Basis $1, X, X^2, \ldots$, und $L[X]$ besitzt dieselben Vektoren als L-Basis. Deshalb ist klar, daß $K[X] \otimes_K L \simeq L[X]$ als L-Vektorräume. Doch mit der oben definierten L-Algebra-Struktur auf $K[X] \otimes L$ wird dies sogar zu einem Isomorphismus von L-Algebren. Allgemeiner gilt auch für Polynomringe in mehreren Variablen

$$K[X_1, \ldots, X_n] \otimes_K L \,\simeq\, L[X_1, \ldots, X_n] \qquad \text{als } L\text{-Algebren.}$$

Bemerkung 6.4.4 Ist $f : \mathcal{A} \to \mathcal{B}$ ein Homomorphismus von K-Algebren, so ist die in Bemerkung 6.4.3 erklärte Abbildung

$$f_L : \mathcal{A}_L \longrightarrow \mathcal{B}_L$$

ein Homomorphismus von L-Algebren. Man erhält so einen *Funktor* von der Kategorie $\mathbf{Alg}_K$ der K-Algebren zu der Kategorie $\mathbf{Alg}_L$ der L-Algebren.

Komplexifizierung

Als wichtigen Spezialfall von Basiserweiterungen betrachten wir die Komplexifizierung eines reellen Vektorraumes. Wenn V ein Vektorraum über $\mathbb{R}$ ist, so heißt

$$V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$$

die **Komplexifizierung** von V. Dies ist ein $\mathbb{C}$-Vektorraum mit

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

Via der $\mathbb{R}$-linearen Inklusion $v \mapsto v \otimes 1$ fassen wir V als $\mathbb{R}$-Untervektorraum von $V_{\mathbb{C}}$ auf. Jedes Element $w \in V_{\mathbb{C}}$ besitzt dann eine eindeutige Darstellung

$$w = u + iv \qquad \text{mit } u, v \in V. \tag{6.24}$$

Die komplexe Konjugation auf $\mathbb{C}$ induziert eine Konjugation auf $V_{\mathbb{C}}$. Genauer: Sei $z \mapsto \bar{z}$ die komplexe Konjugation auf $\mathbb{C}$. Dies ist eine $\mathbb{R}$-lineare Abbildung, und wir können die $\mathbb{R}$-lineare Abbildung $\tau = \mathrm{id}_V \otimes {}^-\colon V_{\mathbb{C}} \to V_{\mathbb{C}}$ betrachten. Es gelten offenbar die folgenden Eigenschaften:

$$\begin{aligned} \tau(aw) &= \bar{a}\tau(w) && \text{für alle } a \in \mathbb{C},\ w \in V_{\mathbb{C}}, \\ \tau(\tau(w)) &= w && \text{für alle } w \in V_{\mathbb{C}}. \end{aligned}$$

Explizit gilt

$$\tau(u + iv) = u - iv \qquad \text{für alle } u, v \in V.$$

Kennt man nur $V_{\mathbb{C}}$ und τ, so kann man daraus V zurückgewinnen, denn offenbar gilt für alle $w \in V_{\mathbb{C}}$

$$w \in V \qquad \Longleftrightarrow \qquad \tau(w) = w.$$

Ist umgekehrt ein $\mathbb{C}$-Vektorraum W gegeben, zusammen mit einer $\mathbb{R}$-linearen Abbildung $\tau : W \to W$, so daß

$$\tau(aw) = \bar{a}\tau(w), \qquad \tau \circ \tau = \mathrm{id}_W, \tag{6.25}$$

so betrachte man den $\mathbb{R}$-Untervektorraum

$$W_{\mathbb{R}} := \{w \in W :\ \tau(w) = w\}.$$

Jedes $w \in W$ läßt sich dann in der Form $w = u + iv$ mit $u, v \in W_{\mathbb{R}}$ schreiben, nämlich

$$w = \frac{1}{2}(w + \tau(w)) + i\frac{1}{2i}(w - \tau(w)).$$

Diese Darstellung ist eindeutig, denn $W_{\mathbb{R}} \cap iW_{\mathbb{R}} = 0$. Somit ist die natürliche Abbildung

$$W_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C} \longrightarrow W, \qquad v \otimes a \longmapsto av,$$

ein Isomorphismus von $\mathbb{C}$-Vektorräumen.

Feststellung und Definition 6.4.5 *Sei W ein $\mathbb{C}$-Vektorraum. Es gibt eine Bijektion zwischen den folgenden Objekten:*

- *$\mathbb{R}$-Untervektorräume V von W, so daß die natürliche Abbildung $V \otimes_{\mathbb{R}} \mathbb{C} \to W$ ein Isomorphismus ist.*

- *$\mathbb{R}$-lineare Abbildungen $\tau : W \to W$ mit den Eigenschaften (6.25).*

Ein solcher Untervektorraum V bzw. eine solche Abbildung τ heißt eine **reelle Struktur** *auf W.*

Beweis: Wir haben oben gesehen, wie man aus V ein τ gewinnt, und umgekehrt. Diese Konstruktionen sind zueinander invers. ∎

Beispiel: Eine reelle Struktur auf dem eindimensionalen Vektorraum $\mathbb{C}$ ist $\mathbb{R}$ (in der üblichen Weise ein $\mathbb{R}$-Unterraum von $\mathbb{C}$). Die zugehörige Abbildung τ ist die komplexe Konjugation. Eine weitere reelle Struktur auf $\mathbb{C}$ ist $\mathbb{R}(1+i)$. Das zugehörige τ ist

$$\tau(x+iy) = y + ix, \qquad x, y \in \mathbb{R}.$$

Anhang A

Mengenlehre

Ein präziser, axiomatischer Aufbau der Mengenlehre findet sich in [Bo]. Weit von einem solchen Vorgehen entfernt, werden hier lediglich in naiver Weise einige elementare Begriffe wiederholt, sowie das Zornsche Lemma vorgestellt. Außerdem zitieren wir einen Satz über Kardinalitäten, der im Beweis von Satz 1.4.15 Verwendung fand.

A.1 Elementare Begriffe

Mengen

Eine **Menge** ist einfach eine strukturlose Ansammlung von Objekten, genannt die **Elemente** der Menge. Eine Menge kann endlich oder unendlich viele Elemente besitzen. Ist a ein Element der Menge M, so wird dies durch das Symbol

$$a \in M$$

ausgedrückt. Besteht M aus den Elementen $a, b, c, \dots$, so schreibt man

$$M = \{a, b, c, \dots\};$$

z.B. ist

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

die Menge der natürlichen Zahlen. Es gibt eine Menge, die keine Elemente besitzt, die sogenannte **leere Menge** $\emptyset$. Sind M_1, M_2 zwei Mengen, und ist jedes Element von M_1 auch ein Element von M_2, gilt also

$$a \in M_1 \quad \Longrightarrow \quad a \in M_2,$$

so heißt M_1 eine **Teilmenge** von M_2, und M_2 eine **Obermenge** von M_1; man bringt dies durch die Symbole

$$M_1 \subset M_2 \qquad \text{oder} \qquad M_2 \supset M_1$$

zum Ausdruck. Zu zwei Mengen M_1, M_2 ist eine **Schnittmenge**

$$M_1 \cap M_2 = \{a : \ a \in M_1 \text{ und } a \in M_2\}$$

sowie eine **Vereinigungsmenge**

$$M_1 \cup M_2 = \{a : \ a \in M_1 \text{ oder } a \in M_2\}$$

definiert. Offenbar gilt

$$M_1 \cap M_2 \subset M_i, \qquad M_i \subset M_1 \cup M_2 \qquad (i = 1, 2).$$

M_1 und M_2 heißen **disjunkt**, falls sie kein gemeinsames Element haben, falls also

$$M_1 \cap M_2 = \emptyset.$$

Schnittmengen und Vereinigungen kann man allgemeiner für beliebig viele Mengen bilden: Sind M_i, $i \in I$ irgendwelche Mengen, indiziert durch eine weitere Menge I, so ist klar, was unter

$$\bigcap_{i \in I} M_i \qquad \text{und} \qquad \bigcup_{i \in I} M_i$$

zu verstehen ist. Ist $M_1 \subset M_2$, so ist das **Komplement** von M_1 in M_2 die Menge

$$M_2 \setminus M_1 := \{a \in M_2 : \ a \notin M_1\}$$

(sprich etwa „M_2 ohne M_1“). M_2 ist dann die disjunkte Vereinigung von M_1 und $M_2 \setminus M_1$.

Sind M_1, M_2 beliebige Mengen, so heißt die Menge aller Paare

$$M_1 \times M_2 := \{(a, b) : \ a \in M_1, \ b \in M_2\}$$

das **direkte Produkt** von M_1 und M_2. In analoger Weise definiert man das direkte Produkt endlich vieler Mengen:

$$M_1 \times \ldots \times M_n = \{(a_1, \ldots, a_n) : \ a_i \in M_i\}.$$

Abbildungen

Eine **Abbildung** f von M_1 nach M_2 ist eine Vorschrift, die jedem Element von M_1 ein Element von M_2 zuordnet; man schreibt dann

$$f : \ M_1 \longrightarrow M_2 \qquad \text{oder} \qquad M_1 \xrightarrow{f} M_2.$$

Das durch f dem Element $a \in M_1$ zugeordnete Element von M_2 wird mit $f(a)$ bezeichnet. Das **Bild** von f ist

$$\mathrm{im}(f) := f(M_1) := \{f(a) : \ a \in M_1\} \subset M_2.$$

Beispiel: Das Bild der Abbildung

$$f : \ \mathbb{N} \longrightarrow \mathbb{N},$$

$$a \longmapsto 2a,$$

ist die Menge $\{2, 4, 6, \dots\}$ der geraden natürlichen Zahlen.

Die Abbildung $f: M_1 \to M_2$ heißt **surjektiv**[1], falls $\mathrm{im}(f) = M_2$, falls also jedes Element von M_2 das Bild (mindestens) eines Elementes von M_1 ist. f heißt **injektiv**, falls jedes Element von M_2 das Bild *höchstens* eines Elementes von M_1 ist, oder anders ausgedrückt, falls für alle $a, b \in M_1$ gilt

$$f(a) = f(b) \quad \Longrightarrow \quad a = b.$$

Eine Abbildung heißt **bijektiv**, falls sie injektiv und surjektiv ist. *Jedes* Element von M_2 ist dann das Bild eines *eindeutig bestimmten* Elementes von M_1, oder kurz gesagt, M_2 wird durch f mit M_1 identifiziert. Folgende Aussage macht man sich leicht klar.

Feststellung A.1.1 *Für eine Abbildung $f: M_1 \to M_2$ zwischen endlichen Mengen M_1, M_2 sind äquivalent:*

i) f ist injektiv.

ii) f ist surjektiv.

iii) f ist bijektiv.

Ist $f: M_1 \to M_2$ eine Abbildung, und $N \subset M_2$, so heißt

$$f^{-1}(N) = \{a \in M_1 : f(a) \in N\}$$

das **Urbild** von N. Das bilden von Urbildern ist *operationstreu*, d.h.

$$f^{-1}(N_1 \cap N_2) = f^{-1}(N_1) \cap f^{-1}(N_2),$$
$$f^{-1}(N_1 \cup N_2) = f^{-1}(N_1) \cup f^{-1}(N_2),$$
$$f^{-1}(M_2 \setminus N) = M_1 \setminus f^{-1}(N).$$

Sind

$$M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3$$

zwei Abbildungen, so heißt die Abbildung

$$M_1 \longrightarrow M_3,$$
$$a \longmapsto g(f(a)),$$

die **Komposition** oder **Hintereinanderschaltung** von f und g, und wird mit dem Symbol $g \circ f$ bezeichnet.

[1] lat. *sur*: auf; *jacio*: werfen

Mengen von Zahlen

Es erfordert erheblichen Aufwand, die Zahlbereiche $\mathbb{N}, \mathbb{Z}, \ldots$ in mathematisch rigoroser Weise einzuführen, siehe etwa [Eb]. Wir setzen stattdessen eine gewisse Vertrautheit mit diesen Objekten voraus, und fassen an dieser Stelle nur noch zusammen.

Wir definieren die Menge der **natürlichen Zahlen**

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

Oft wird die Null zu den natürlichen Zahlen gezählt. Wir führen dafür jedoch das Symbol

$$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$$

ein. Die Menge der **ganzen Zahlen** ist

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

Die **rationalen Zahlen** sind Brüche ganzer Zahlen:

$$\mathbb{Q} = \left\{ \frac{a}{b} : \ a, b \in \mathbb{Z}, \ b \neq 0 \right\}.$$

Die rationalen Zahlen sind in gewisser Weise lückenhaft[2]. Man vervollständigt sie daher zu der Menge

$$\mathbb{R}$$

der **reellen Zahlen**. Die Elemente von $\mathbb{R}$ kann man sich durch die Punkte auf einer Geraden veranschaulichen. Das Bedürfnis, beliebige Polynomgleichungen $X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0 = 0$ lösen zu können, führt zur Einführung der **komplexen Zahlen**

$$\mathbb{C}.$$

Jedes Element von $\mathbb{C}$ besitzt eine eindeutige Darstellung der Gestalt $a + ib$ mit $a, b \in \mathbb{R}$, wobei $i \in \mathbb{C}$ ein spezielles Element mit der Eigenschaft $i^2 = -1$ ist. Deshalb lassen sich die Elemente von $\mathbb{C}$ mit den Punkten einer Ebene identifizieren.

Wir haben die Mengen

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

erklärt. Mit der Terminologie von Kapitel 1 ist $\mathbb{Z}$ ein *Ring*, und $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ sind *Körper*. Nur $\mathbb{C}$ ist dabei *algebraisch abgeschlossen*, siehe Definition 4.1.16.

A.2 Auswahlaxiom und Kardinalität

Das Zornsche Lemma

Eine **Ordnung** auf einer Menge M ist eine Relation $a \leq b$ für gewisse (nicht notwendig alle) $a, b \in M$, so daß gilt:

[2]Genauer: nicht jede *Cauchyfolge* besitzt einen Grenzwert in $\mathbb{Q}$.

i) $a \leq a$ für alle $a \in M$.

ii) Aus $a \leq b$ und $b \leq c$ folgt $a \leq c$.

iii) Aus $a \leq b$ und $b \leq a$ folgt $a = b$.

M heißt dann eine (teilweise) **geordnete** Menge. Falls *alle* Elemente vergleichbar sind, falls also stets $a \leq b$ oder $b \leq a$ gilt, so heißt M **total geordnet** oder **linear geordnet**. Ist N eine Teilmenge der geordneten Menge M, so ist auch N geordnet. Eine **obere Schranke** von N in M ist ein Element $b \in M$, so daß $a \leq b$ für alle $a \in N$. Ein **maximales Element** in M ist ein $b \in M$, so daß für alle $a \in M$ gilt:

$$b \leq a \qquad \Longrightarrow \qquad b = a.$$

Das folgende *Zornsche Lemma*[3] kann man entweder aus den grundlegenden Axiomen der Mengenlehre folgern (es ist dann äquivalent zum *Auswahlaxiom*), oder selbst zum Axiom erheben.

Satz A.2.1 (Lemma von Zorn) *Sei M eine teilweise geordnete nicht-leere Menge mit der Eigenschaft, daß jede total geordnete Teilmenge eine obere Schranke in M besitzt. Dann gibt es in M (mindestens) ein maximales Element.*

Kardinalität

Zwei Mengen M und N haben die gleiche **Kardinalität** oder **Mächtigkeit**, in Zeichen

$$\mathrm{card}(M) = \mathrm{card}(N),$$

falls es eine Bijektion $M \xrightarrow{\sim} N$ gibt. M hat die gleiche oder eine geringere Kardinalität als N, in Zeichen

$$\mathrm{card}(M) \leq \mathrm{card}(N),$$

falls es eine Injektion $M \hookrightarrow N$ gibt. Viel genauer sind diese Begriffe in [Bo] III.3 definiert.

Satz A.2.2 *Seien I und T unendliche Mengen. Für jedes $i \in I$ sei eine endliche Teilmenge T_i von T gegeben, und es gelte*

$$T = \bigcup_{i \in I} T_i.$$

Dann ist $\mathrm{card}(I) \geq \mathrm{card}(T)$.

Dies folgt aus Corollary 3 in [Bo] III.6.3. Wir benötigen diese Aussage im Beweis von Satz 1.4.15.

[3] Max Zorn, 1906-1993

Anhang B

Kategorien und Funktoren

Dieser Anhang bringt wenig substantielle Mathematik, stellt aber eine moderne Terminologie vor.

B.1 Kategorien

Definition von Kategorien

In den Bemerkungen 1.1.10 und 1.2.3 haben wir das Konzept der *Kategorien* bereits angedeutet. Hier ist die formale Definition.

Definition B.1.1 *Eine* **Kategorie** *besteht aus den folgenden Daten:*

- *Einer Klasse von* **Objekten**.
- *Zu je zwei Objekten A, B eine Menge von* **Morphismen** $\mathrm{Mor}(A, B)$.
- *Zu jedem Objekt A einen speziellen Morphismus* $\mathrm{id}_A \in \mathrm{Mor}(A, A)$.
- *Zu je drei Objekten A, B, C eine Abbildung*

$$\begin{aligned}\mathrm{Mor}(A,B) \times \mathrm{Mor}(B,C) &\longrightarrow \mathrm{Mor}(A,C),\\ (f,g) &\longmapsto g \circ f.\end{aligned}$$

Es müssen die folgenden Axiome gelten:

i) Ist $f \in \mathrm{Mor}(A, B)$, so gilt $f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f$.

ii) Für $f \in \mathrm{Mor}(A, B)$, $g \in \mathrm{Mor}(B, C)$ und $h \in \mathrm{Mor}(C, D)$ gilt

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Wir haben bereits mehrere Beispiele für Kategorien kennengelernt. Eines davon ist die Kategorie **Gr** der Gruppen. Die Klasse der Objekte ist dabei die Gesamtheit *aller* Gruppen, und für zwei Gruppen

A, B ist $\mathrm{Mor}(A, B)$ die Menge *aller* Gruppenhomomorphismen $A \to B$. Die Verknüpfung $\mathrm{Mor}(A, B) \circ \mathrm{Mor}(B, C) \to \mathrm{Mor}(A, C)$ ist einfach die Hintereinanderschaltung von Gruppenhomomorphismen; vgl. Bemerkung 1.1.5. Der Morphismus id_A ist die Identität auf A. Die Gültigkeit der beiden Axiome ist sofort einzusehen.

Wir geben weitere Beispiele von Kategorien, die uns bis jetzt begegnet sind. Den Nachweis, daß es sich tatsächlich um Kategorien handelt, kann der Leser ähnlich wie bei den Gruppen führen.

- Die Kategorie **Set** der Mengen (und Abbildungen).
- Die Kategorie **Gr** der Gruppen (und Gruppenhomomorphismen).
- Die Kategorie **Ab** der abelschen Gruppen (und Homomorphismen zwischen abelschen Gruppen).
- Die Kategorie **Ring** der Ringe (und Ringhomomorphismen).
- Die Kategorie $\mathbf{Vec}_K$ der Vektorräume über dem Körper K (und K-linearen Abbildungen). Für V und W zwei K-Vektorräume haben wir $\mathrm{Mor}(V, W)$ mit $\mathrm{Hom}_K(V, W)$ bezeichnet (Definition 1.3.9).
- Die Kategorie $\mathbf{Alg}_K$ der K-Algebren (und K-Algebren-Homomorphismen).
- Die Kategorie $\mathbf{Mod}_R$ der Moduln über dem Ring R.

In all diesen Kategorien besteht $\mathrm{Mor}(A, B)$ aus gewissen Abbildungen $A \to B$. Dies ist jedoch in der Definition einer Kategorie nicht gefordert; tatsächlich gibt es viele Kategorien, deren Morphismen keine Abbildungen sind. Dennoch verwendet man statt $f \in \mathrm{Mor}(A, B)$ auch die Symbole

$$f : A \longrightarrow B \qquad \text{oder} \qquad A \xrightarrow{f} B.$$

Definition B.1.2 *Ein Isomorphismus in einer Kategorie ist ein Morphismus, zu dem es ein zweiseitiges Inverses gibt. Genauer: Ist $f : A \to B$ ein Morphismus, so heißt f ein* **Isomorphismus**, *falls ein Morphismus $g : B \to A$ existiert mit $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$.*

In den Definitionen 1.1.9, 1.2.2 und 1.3.8 haben wir Isomorphismen (von Gruppen, Ringen, bzw. Vektorräumen) als bijektive Homomorphismen definiert. Die folgende Feststellung besagt, daß dies mit der kategoriellen Definition B.1.2 übereinstimmt.

Feststellung B.1.3 *Sei $f : A \to B$ ein Morphismus in einer der Kategorien* **Set**, **Gr**, **Ab**, **Ring** *oder* $\mathbf{Vec}_K$. *Genau dann ist f bijektiv, wenn ein Morphismus $g : B \to A$ existiert mit $f \circ g = \mathrm{id}_B$ und $g \circ f = \mathrm{id}_A$.*

Beweis: Wir nehmen zunächst an, daß das zweiseitige Inverse g existiert, und zeigen, daß f bijektiv ist. f injektiv: Aus $f(a_1) = f(a_2)$ folgt $g(f(a_1)) = g(f(a_2))$, also

$$a_1 = \mathrm{id}_A(a_1) = (g \circ f)(a_1) = g(f(a_1)) = g(f(a_2)) = (g \circ f)(a_2) = \mathrm{id}_A(a_2) = a_2.$$

f ist aber auch surjektiv, denn für beliebiges $b \in B$ gilt

$$f(g(b)) = (f \circ g)(b) = \mathrm{id}_B(b) = b.$$

Jedes Element von B liegt also im Bild von f.

Sei jetzt f als bijektiv vorausgesetzt. Wir müssen die Existenz eines zweiseitigen Inversen g zeigen. Für $b \in B$ definieren wir $g(b)$ als das eindeutig bestimmte Element von A, das unter f auf b abgebildet wird (ein solches Element existiert, da f surjektiv ist, und ist eindeutig, da f injektiv ist). Dies definiert eine Abbildung $g : B \to A$, für die wir die gewünschten Eigenschaften nachweisen. Für $b \in B$ gilt

$$f(g(b)) = b$$

nach Konstruktion. Also ist $f \circ g = \mathrm{id}_B$. Sei nun $a \in A$ beliebig, und $b = f(a)$. Dann besagt obige Gleichung

$$f(g(f(a))) = f(a).$$

Da f injektiv ist, folgt daraus $g(f(a)) = a$. Mit anderen Worten: $g \circ f = \mathrm{id}_A$. Damit ist gezeigt, daß g die Eigenschaften eines zweiseitigen Inversen hat.

Es bleibt zu zeigen, daß g überhaupt ein *Morphismus* in der betrachteten Kategorie ist. Dies ist nur für **Set** klar, da die Morphismen dort beliebige Abbildungen sind. Wenn wir aber z.B. in $\mathbf{Vec}_K$ sind, so muß noch bewiesen werden, daß g eine lineare Abbildung ist. Wir führen dies durch und überlassen dem Leser den analogen Beweis, daß g im Falle von **Gr** und **Ab** ein Gruppenhomomorphismus, im Falle von **Ring** ein Ringhomomorphismus ist.

Seien $b_1, b_2 \in B$. Es gilt

$$f(g(b_1 + b_2)) = b_1 + b_2 = f(g(b_1)) + f(g(b_2)) = f(g(b_1) + g(b_2)).$$

Aus der Injektivität von f folgt $g(b_1 + b_2) = g(b_1) + g(b_2)$. Sei jetzt $b \in B$ und $\lambda \in K$. Es gilt

$$f(g(\lambda b)) = \lambda b = \lambda f(g(b)) = f(\lambda g(b)).$$

Wieder aus der Injektivität von f folgt $g(\lambda(b)) = \lambda g(b)$. Wir haben gezeigt, daß g eine lineare Abbildung ist. ∎

Bemerkung B.1.4 Die Kategorie **Top** der *topologischen Räume* und *stetigen Abbildungen* ist ein prominentes Beispiel dafür, daß bijektive Morphismen nicht automatisch Isomorphismen zu sein brauchen.

Definition B.1.5 *Sei* **C** *eine Kategorie.*

i) Ein Objekt **I** *in* **C** *heißt* **initial**, *falls es zu jedem Objekt A in* **C** *genau einen Morphismus* $\mathbf{I} \to A$ *gibt.*

ii) Ein Objekt **F** *in* **C** *heißt* **final**, *falls es zu jedem Objekt A in* **C** *genau einen Morphismus* $A \to \mathbf{F}$ *gibt.*

Feststellung B.1.6 *i) Die triviale Gruppe ist gleichzeitig ein initiales und finales Objekt in* **Gr**. *Dasselbe gilt in* **Ab**.

ii) Der Nullring ist gleichzeitig ein initiales und finales Objekt in **Ring**.

iii) Der Nullvektorraum ist gleichzeitig ein initiales und finales Objekt in $\mathbf{Vec}_K$.

Beweis: Wir beweisen iii); i) und ii) gehen analog. Sei $\mathbf{0}$ der triviale K-Vektorraum, und V ein beliebiger K-Vektorraum. Die Abbildung

$$\mathbf{0} \longrightarrow V,$$

die das einzige Element von $\mathbf{0}$ auf $0 \in V$ schickt, ist offenbar linear. Sie ist die einzige lineare Abbildung $\mathbf{0} \to V$, denn jede lineare Abbildung bildet Null auf Null ab. Also ist $\mathbf{0}$ ein initiales Objekt in $\mathbf{Vec}_K$.

Die einzige Abbildung

$$V \longrightarrow \mathbf{0},$$

die überhaupt existiert, und jedes Element von V auf das einzige Element von $\mathbf{0}$ abbildet, ist linear. Dies zeigt, daß $\mathbf{0}$ ein finales Objekt in $\mathbf{Vec}_K$ ist. ∎

Die triviale Gruppe, den Nullring, und den Nullvektorraum bezeichnen wir oft einfach mit 0. Die triviale Gruppe wird auch mit 1 bezeichnet, falls die multiplikative Schreibweise verwendet wird. Wann immer Symbole wie

$$0 \longrightarrow A \qquad \text{oder} \qquad B \longrightarrow 1$$

auftauchen sind die eindeutig bestimmten Morphismen gemeint, die bei dem initialen Objekt starten bzw. in das finale Objekt gehen.

B.2 Funktoren

Definition B.2.1 *Seien* $\mathbf{C}_1$ *und* $\mathbf{C}_2$ *Kategorien. Ein* **covarianter Funktor** *von* $\mathbf{C}_1$ *nach* $\mathbf{C}_2$ *ist eine Vorschrift* $\mathcal{F}$*, die jedem Objekt* A *von* $\mathbf{C}_1$ *ein Objekt* $\mathcal{F}(A)$ *von* $\mathbf{C}_2$ *zuordnet, sowie jedem Morphismus*

$$A \xrightarrow{f} B \tag{B.1}$$

in $\mathbf{C}_1$ *einen Morphismus*

$$\mathcal{F}(A) \xrightarrow{\mathcal{F}(f)} \mathcal{F}(B)$$

in $\mathbf{C}_2$*, so daß gilt:*

i) $\mathcal{F}(\mathrm{id}_A) = \mathrm{id}_{\mathcal{F}(A)}$ *für alle Objekte* A *von* $\mathbf{C}_1$.

ii) $\mathcal{F}(g \circ f) = \mathcal{F}(g) \circ \mathcal{F}(f)$ *für zwei Morphismen* $A \xrightarrow{f} B \xrightarrow{g} C$ *in* $\mathbf{C}_1$.

Ein **kontravarianter Funktor** *von* $\mathbf{C}_1$ *nach* $\mathbf{C}_2$ *ist eine ähnliche Vorschrift, nur wird jedem Morphismus (B.1) in* $\mathbf{C}_1$ *ein Morphismus in der umgekehrten Richtung zugeordnet,*

$$\mathcal{F}(A) \xleftarrow{\mathcal{F}(f)} \mathcal{F}(B),$$

und die covariante Kettenregel in ii) ist zu ersetzen durch $\mathcal{F}(g \circ f) = \mathcal{F}(f) \circ \mathcal{F}(g)$.

Beispiel B.2.2 Sei $\mathbf{Ring}_1$ die Kategorie der Ringe mit 1 (die Morphismen sind Ringhomomorphismen, die 1 auf 1 abbilden). Jedem solchen Ring R ordnen wir seine Einheitengruppe R^* zu (siehe 1.2.5). Ist

$$R_1 \xrightarrow{f} R_2$$

ein Morphismus in $\mathbf{Ring}_1$, so bildet f, wie man leicht sieht, Einheiten auf Einheiten ab, induziert also einen Gruppenhomomorphismus

$$R_1^* \xrightarrow{f^*} R_2^*,$$

wobei f^* die Einschränkung von f auf R_1^* bezeichnet. Die Zuordnung $R \mapsto R^*$ und $f \mapsto f^*$ definiert auf diese Weise einen covarianten Funktor von $\mathbf{Ring}_1$ nach $\mathbf{Gr}$ (die Funktoreigenschaften i) und ii) sind nahezu trivial).

Beispiel B.2.3 Jeder K-Vektorraum V ist eine abelsche Gruppe bezüglich der Addition. Ordnet man V diese abelsche Gruppe zu, und einer linearen Abbildung $f: V \to W$ den Gruppenhomomorphismus f, so erhält man einen covarianten Funktor von $\mathbf{Vec}_K$ nach $\mathbf{Ab}$. Man vergißt also den Skalarkörper K und die skalare Multiplikation, und behält nur die additive Gruppe V übrig. Funktoren dieser Art heißen **Vergißfunktoren**.

Beispiel B.2.4 Der Funktor Hom. Sei W ein fest gewählter K-Vektorraum. Zu jeder linearen Abbildung

$$f: \; V_1 \longrightarrow V_2$$

bekommen wir eine lineare Abbildung

$$\begin{aligned} \operatorname{Hom}(W, f): \; \operatorname{Hom}(W, V_1) &\longrightarrow \operatorname{Hom}(W, V_2), \\ g &\longmapsto f \circ g. \end{aligned}$$

Es ist leicht zu sehen, daß dadurch ein covarianter Funktor

$$\operatorname{Hom}(W, \,\cdot\,): \; \mathbf{Vec}_K \longrightarrow \mathbf{Vec}_K$$

definiert wird. Analog bekommt man einen kontravarianten Funktor

$$\operatorname{Hom}(\,\cdot\,, W): \; \mathbf{Vec}_K \longrightarrow \mathbf{Vec}_K,$$

indem man der linearen Abbildung $f: \; V_1 \to V_2$ die lineare Abbildung

$$\begin{aligned} \operatorname{Hom}(f, W): \; \operatorname{Hom}(V_2, W) &\longrightarrow \operatorname{Hom}(V_1, W), \\ g &\longmapsto g \circ f. \end{aligned}$$

zuordnet.

Beispiel B.2.5 Setzt man im vorhergehenden Beispiel $W = K$, so erhält man den **Dualraumfunktor**, der jeder linearen Abbildung $f : V_1 \to V_2$ die duale lineare Abbildung

$$f^* : V_2^* \longrightarrow V_1^*$$

zuordnet, vgl. Definitionen 2.5.1 und 2.5.3. In Feststellung 2.5.4 hatten wir verifiziert, daß es sich tatsächlich um einen Funktor handelt.

Beispiel B.2.6 Tensorieren mit einem fest gewählten K-Vektorraum V liefert einen covarianten Funktor $\_ \otimes V : \mathbf{Vec}_K \to \mathbf{Vec}_K$. Jedem Morphismus $f : W_1 \to W_2$ wird dabei der Morphismus

$$f \otimes V := f \otimes \mathrm{id}_V : W_1 \otimes V \longrightarrow W_2 \otimes V$$

zugeordnet, vgl. Bemerkung 6.1.10. Analog bekommt man einen ebenfalls covarianten Funktor $V \otimes \_$, der wegen Feststellung 6.1.3 in einem hier nicht näher definierten Sinne zu $\_ \otimes V$ *isomorph* ist.

Beispiel B.2.7 Eine lineare Abbildung $f : V \to W$ induziert eine lineare Abbildung $f^{\otimes n} : V^{\otimes n} \to W^{\otimes n}$ auf den Tensorpotenzen, und damit auch eine lineare Abbildung

$$\mathbf{T}f : \mathbf{T}V \longrightarrow \mathbf{T}W$$

zwischen den Tensoralgebren. Diese erweist sich sofort als ein Homomorphismus von K-Algebren. Auf diese Weise erhält man einen covarianten Funktor $\mathbf{T}$ von $\mathbf{Vec}_K$ in die Kategorie $\mathbf{Alg}_K$ der K-Algebren.

Symbolverzeichnis

Mengen, Abbildungen, Zahlbereiche

Vektorräume, lineare Abbildungen

Matrizen

Ringe und Körper

Gruppen

Multilinearformen

Tensorprodukte

Sonstiges

Literaturverzeichnis

[Bo] BOURBAKI, N.: *Theory of Sets.* Hermann, 1968

[Eb] EBBINGHAUS, H.D. et al.: *Zahlen.* Springer, 1992

[F] FISCHER, G.: *Lineare Algebra.* Vieweg, 1997

[J] JÄNICH, K.: *Lineare Algebra.* Springer, 1998

[La1] LANG, S.: *Linear Algebra.* Addison-Wesley, 1971

[La2] LANG, S.: *Algebra.* Addison-Wesley, 1993

Index