Westfälische Wilhelms-Universität Münster
Institut für Mathematische Logik und Grundlagenforschung

Lineare Algebra

Vorlesung von

Wolfram Pohlers

Wintersemester 1997/98
Sommersemester 1998

Typeset by Wolfram Pohlers

Vorwort

Das vorliegende Skript ist eine Mitschrift der Vorlesung, die ich im Wintersemester 1997/98 und Sommersemester 1998 an der Westfälischen Wilhelms-Universität gehalten habe. Ein Teil der Mitschrift – etwa bis Kapitel 4 – beruht auf Aufzeichnungen der von mir vorgetragenen Definitionen und Sätze, die mir Herr stud. math. PETER HORN bereits in geTEXter Form überlassen hat. Allerdings habe ich diesen Text so weit verändert und ergänzt, daß alle eventuellen Ungenauigkeiten zu meinen Lasten gehen. Die Niederschrift des übrigen Textes erfolgte vorlesungsbegleitend.
Mein ganz besonderer Dank gebührt den Studierenden Frau MAIKE WALTHER, Herrn HENRIK BUCHHOLZ, Herrn KEVIN BUCHIN und Herrn HENDRIK KLÄVER, die den vorliegenden Text Korrektur gelesen und einen Großteil der Korrekturen auch eigenhändig eingearbeitet haben.
Ich hoffe, daß dieses Skript sich als hilfreich für die Prüfungsvorbereitungen erweisen wird.

Im Januar 1999

Wolfram Pohlers

Inhaltsverzeichnis

0. Vorbemerkungen, Notationen, Konventionen

0.1 Etwas Aussagenlogik

Beispiele von Aussagen

- *Die Sonne scheint.*
- *Es regnet.*
- *Ich spanne den Schirm auf.*

Eine Aussage besteht also aus einem (oder mehreren) Objekten, *die Sonne*, *Es*, *ich*, *der Schirm* und Prädikaten, die auf die Objekte angewendet werden. Die *Sonne* scheint; *es* regnet, *ich* spanne den *Schirm* auf.
Mehrere Aussagen lassen sich zu einer neuen Aussage verknüpfen. Z.B.:

- *Die Sonne scheint* und *es ist warm.*
- *Die Sonne scheint* oder *es regnet.*
- Wenn *es regnet,* dann *spanne ich den Schirm auf.*

Durch Negation lassen sich Aussagen in ihr Gegenteil umformen. Dies ist bei *wenn... dann* Aussagen nicht immer ganz einfach. So erhalten wir beispielsweise als Negation der Aussage

- Wenn *es regnet,* dann *spanne ich den Schirm auf*

die Aussage

- *Es regnet* und *ich spanne den Schirm* nicht *auf.*

Um den umgangssprachlichen Begriffen wie *und*, *oder*, *wenn ... dann* und *nicht* eine präzise mathematische Bedeutung zu geben, führen wir aussagenlogische Junktoren ein. Aussagen sind dadurch gekennzeichnet, daß sie wahr oder falsch sein können. Die Verknüpfung von Aussagen ist also dadurch festgelegt, daß wir vereinbaren, welcher Wahrheitswert der verknüpften Aussage in Abhängigkeit der Wahrheitswerte der zu verknüpfenden Aussagen zukommen soll. Eine n-stellige aussagenlogische Operation ist also eine Funktion, die n Wahrheitswerten $\{w, f\}$ einen Wahrheitswert zuordnet. Wir kommen mit den folgenden 1- und 2–stelligen aussagenlogischen Operationen aus. (Mit Mitteln der Mathematischen Logik läßt sich zeigen, daß jede aussagenlogische Operation schon aus den folgenden aufgebaut werden kann.)

0.1.1 Definition Die *aussagenlogischen Operationen* (Junktoren)

- und ($\wedge$)
- oder ($\vee$)
- nicht ($\neg$)
- wenn-dann ($\Rightarrow$) *und*
- genau dann wenn ($\Leftrightarrow$)

werden durch die folgenden Wahrheitstafeln definiert:

$\neg$	
w	f
f	w

$\wedge$	w	f
w	w	f
f	f	f

$\vee$	w	f
w	w	w
f	w	f

$\Rightarrow$	w	f
w	w	f
f	w	w

$\Leftrightarrow$	w	f
w	w	f
f	f	w

Dabei sind die Wahrheitstafeln so zu lesen, daß in der Spalte unter dem Junktor dessen erstes Argument und in der Zeile neben dem Junktor dessen zweites Argument steht. Der Wahrheitswert der verknüpften Aussage steht dann im Schnittpunkt von Zeile und Spalte.

Legen wir die Wahrheitstafeln zugrunde, so erhalten wir durch einfaches Nachrechnen die folgenden

0.1.2 Rechenregeln für Aussagenlogische Operationen

(i) $(A \wedge B) = (B \wedge A), (A \vee B) = (B \vee A)$

(ii) $(A \wedge B) \wedge C = A \wedge (B \wedge C)$

(iii) $(A \vee B) \vee C = A \vee (B \vee C)$

(iv) $(A \wedge B) \vee C = (A \vee C) \wedge (B \vee C)$

(v) $(A \vee B) \wedge C = (A \wedge C) \vee (B \wedge C)$

(vi) $\neg\neg A = A$

(vii) $\neg(A \wedge B) = \neg A \vee \neg B$

(viii) $\neg(A \vee B) = \neg A \wedge \neg B$

(xi) $(A \Rightarrow B) = \neg A \vee B$

(x) $\neg(A \Rightarrow B) = A \wedge \neg B$

(xi) $(A \Rightarrow f) = \neg A$

(xii) $(\neg A \Rightarrow f) = A$

(xiii) $(A \Leftrightarrow B) = (A \Rightarrow B) \wedge (B \Rightarrow A)$

(xiv) $(A \Rightarrow B) = (\neg B \Rightarrow \neg A)$

0.2 Etwas Quantorenlogik

Neben reinen Aussagen wie "die Sonne scheint" spielen auch quantifizierte Aussagen eine Rolle, wie z.B. "*alle* Studenten sind fleißig" oder "*es gibt* Studenten, die fleißig sind", die etwas über alle Studenten aussagen oder die Existenz (von fleißigen Studenten) behaupten. Solche Aussagen nennt man *quantifizierte Aussagen*. Die Theorie, die solche Aussagen untersucht, ist die Quantorenlogik (oder Prädikatenlogik, wie sie häufiger genannt wird). Man kommt für die üblichen mathematischen Anwendungen mit den beiden folgenden Quantoren aus.

- $\forall$ (für Alle)
- $\exists$ (Es gibt)

Als Beispiel wollen wir eine Formel angeben, die die Stetigkeit einer reellen Funktion in einem Punkte x_0 ausdrückt.

$$(\forall\epsilon)[\epsilon > 0 \;\Rightarrow\; (\exists\delta)(\delta > 0 \wedge (\forall x)(|x - x_0| < \delta \;\Rightarrow\; |f(x) - f(x_0)| < \epsilon))]$$

Oft steht man vor dem Problem, quantifizierte Aussagen negieren zu müssen. Schon das obige Beispiel der Stetigkeit zeigt, daß dies im allgemeinen nicht einfach ist. Folgt man der inhaltlichen Bedeutung der Aussage " Es gilt nicht, daß für alle x die Aussage $F(x)$ zutrifft", so kommt man zu dem Schluß, daß es dann zumindest ein x geben muß, für das $F(x)$ nicht zutrifft. Ebenso wird die Aussage "für alle x gilt $F(x)$" falsch, wenn es ein x mit $\neg F(x)$ gibt. Damit erhält man die folgenden "Rechenregeln" für die Negation quantifizierter Aussagen (die sich im Rahmen der Mathematischen Logik genauer begründen lassen, ohne daß wir dies hier tun wollen).

- $\neg(\forall x F(x)) = \exists x \neg F(x)$
- $\neg(\exists x F(x)) = \forall x \neg F(x)$

Wenden wir dies auf das Beispiel der Stetigkeit an, so erhalten wir, daß eine Funktion in einem Punkt x_0 nicht stetig ist, wenn gilt:

$$(\exists \epsilon)[\epsilon > 0 \wedge (\forall \delta)(\delta > 0 \quad \Rightarrow \quad (\exists x)(|x - x_0| < \delta \wedge \epsilon \leq |f(x) - f(x_0)|))].$$

Um $\forall x F(x)$ zu zeigen, müssen wir nachweisen, daß $F(n)$ für ein beliebiges Element n gilt.
Um $\exists x F(x)$ zu zeigen, genügt es, ein Objekt mit $F(x)$ zu finden. Man kann $\exists x F(x)$ jedoch auch zeigen, indem wir $(\forall x)\neg F(x)$ annehmen und daraus etwas Falsches schließen. Damit haben wir

$$(\forall x)\neg F(x) \quad \Rightarrow \quad f$$

und mit 0.1.2 (xi) folgt $\neg(\forall x)\neg F(x)$, was $(\exists x)\neg\neg F(x)$ und damit

$$(\exists x)F(x)$$

bedeutet.

0.3 Etwas Mengenlehre

Alle in der Mathematik betrachteten Objekte lassen sich als Mengen auffassen. Dabei verstehen wir eine Menge naiv als eine Zusammenfassung von Objekten zu einem Oberbegriff. Diese Objekte bezeichnen wir dann als die *Elemente* der Menge, die durch den Oberbegriff repräsentiert wird. Es mag einem zunächst merkwürdig vorkommen, daß sich sogar natürliche Zahlen als Mengen darstellen lassen, obwohl wir uns gar nicht so recht vorstellen können, was wohl die Elemente einer natürlichen Zahl sein sollen. Wir wollen uns daher kurz klarmachen, daß dies tatsächlich geht. Es sei $\emptyset$ das Symbol für die leere Menge, d.h. die Menge, die keine weiteren Elemente hat. Dann setzen wir

- $0 = \emptyset$
- $1 = \{\emptyset\}$
- $2 = \{\emptyset, \{\emptyset\}\} = \{0, 1\}$
- $3 = \{0, 1, 2\}$

- *usw. . . .*

und erhalten so alle natürlichen Zahlen als Mengen. Es ist ein Ergebnis der *Mengenlehre*, daß sich tatsächlich alle in der Mathematik betrachteten Objekte als Mengen darstellen lassen. Daher hat sich die Sprache der Mengenlehre als eine in der Mathematik üblicherweise gebrauchte Sprache eingebürgert. Dies zwingt uns, uns hier kurz mit einigen Grundbegriffen der Mengenlehre auseinanderzusetzen. Dabei wollen wir nicht definieren, was eine Menge ist (das wäre ein viel zu schwieriges Unterfangen, bis heute können wir nicht genau sagen, was eigentlich Mengen sind). Wir wollen uns daher nur mit den Beziehungen auseinandersetzen, in denen Mengen untereinander stehen können.
Die wesentlichste Beziehung der Mengenlehre ist

$$a \in b,$$

die ausdrücken soll, daß die Menge a ein Element von b ist. Die Klasse aller Mengen (die selbst keine Menge sein kann) bezeichnen wir als das mathematische Universum $\mathcal{U}$, in dem alle mathematischen Objekte zu finden sind.
Die Notation $M \in \mathcal{U}$ bedeutet daher "M ist eine Menge" und damit ein mathematisch vernünftiges Objekt. Zusammenfassungen von Mengen, von denen wir noch nicht so genau wissen, ob sie mathematisch vernünftige Objekte, also Mengen, sind, bezeichnen wir vorsichtshalber zunächst als Klassen. Insbesondere sind alle Mengen auch Klassen. Obwohl wir Klassen nicht als die eigentlichen Objekte unseres Interesses betrachten, wollen wir einige ihrer grundlegenden Eigenschaften charakterisieren. So wollen wir zwei Klassen M und N als gleich betrachten, wenn sie die gleiche Extension haben, d.h. die gleichen Elemente besitzen. In Formelschreibweise bedeutet dies:

(Extensionalität) $\quad M = N :\Leftrightarrow \forall x(x \in M \Leftrightarrow x \in N).$

Beachte, daß alles, was für Klassen definiert wird, auch für Mengen gilt, die ja selbst alle auch Klassen sind. So sind zwei Mengen gleich, wenn sie als Klassen gleich sind, d.h. die gleichen Elemente besitzen.
Wir sagen, daß M eine Teilklasse von N oder N eine Oberklasse von M ist, wenn jedes Element von M auch ein Element von N ist. Wissen wir, daß M und N Mengen sind, so sprechen wir auch von Teil– und Obermengen. In Formelschreibweise

$$M \subseteq N \quad :\Leftrightarrow \quad (\forall x)(x \in M \quad \Rightarrow \quad x \in N).$$

Wir definieren

$$M \subsetneq N \quad :\Leftrightarrow \quad M \subset N \wedge M \neq N$$

und nennen M eine *echte* Teilklasse (Teilmenge) von N.

0.3.1 Lemma $M = N \Leftrightarrow M \subseteq N \wedge N \subseteq M.$

Beweis: Nach 0.1.2 gilt $x \in M \Leftrightarrow x \in N$ genau dann, wenn $(x \in M \Rightarrow x \in N)$ $\wedge$ $(x \in N \Rightarrow x \in M)$ gilt. Daraus erhalten wir

$$\begin{aligned} M = N &\iff (\forall x)[x \in M \;\;\Leftrightarrow\;\; x \in N] \\ &\iff (\forall x)[x \in M \Rightarrow x \in N] \wedge (\forall x)[x \in N \Rightarrow x \in M] \\ &\iff M \subseteq N \wedge N \subseteq M. \end{aligned}$$

□

Um Klassen bilden zu können, müssen wir ihre Objekte beschreiben können. Dazu benötigen wir eine Sprache. Wir wollen diese Sprache hier nicht exakt festlegen (das wäre für diese Vorlesung übertrieben pingelig), sondern unsere übliche Umgangssprache, angereichert um die bereits bekannten mathematischen Symbole benutzen. Wir drücken durch die Schreibweise $F(a)$ aus, daß dem Objekt a die Eigenschaft $F(a)$ (die in unserer Sprache beschrieben ist) zukommt. Haben wir so eine Eigenschaft, so bilden wir den Ausdruck $\{x|\; F(x)\}$ und sprechen von der Klasse der x mit der Eigenschaft $F(x)$. Eine Menge a ist genau dann ein Element dieser Klasse, wenn es die Eigenschaft $F(a)$ hat. In Formeln ausgedrückt schreiben wir dies als

$$a \in \{x|\; F(x)\} \quad :\Leftrightarrow \quad a \in \mathcal{U} \wedge F(a).$$

Insbesondere definieren wir die leere Klasse durch

$$\emptyset := \{x|\; \neg(\exists y)(y \in x)\}.$$

Wir wollen nun einige Prinzipien angeben, wie wir Mengen (also Elemente des Universums $\mathcal{U}$) bilden können. Wir beginnen mit

(Leere Menge) $\qquad \emptyset \in \mathcal{U}.$

Haben wir bereits eine Menge M, so wollen wir aus dieser die Elemente $x \in M$ aussondern können, denen die Eigenschaft $F(x)$ zukommt. Wir fordern daher

(Aussonderungsaxiom) $\qquad M \in \mathcal{U} \;\;\Rightarrow\;\; \{x \in M|\; F(x)\} \in \mathcal{U}.$

Wir werden es in dieser Vorlesung nie mit echten Klassen zu tun haben, d.h. mit Klassen, die keine Mengen sind. Da es darüber hinaus oft klar oder unerheblich ist, aus welcher Menge M ausgesondert wird, schreiben wir einfach salopp $\{x|\; F(x)\}$ anstatt $\{x \in M|\; F(x)\}$, falls M festgelegt ist.
Als Beispiel wollen wir endlich viele Objekte $a_1, \ldots, a_n$ zu einer Klasse M zusammenfassen. Dies läßt sich beschreiben durch:

$$M = \{x|\; x = a_1 \vee x = a_2 \vee \;\ldots\; \vee x = a_n\}$$

oder kürzer durch

$$M = \{a_1, \dots, a_n\}.$$

Dann gilt

$$x \in M \quad \Leftrightarrow \quad x \in \mathcal{U} \wedge x = a_i \text{ für ein } i \text{ mit } 1 \le i \le n.$$

Natürlich wollen wir M wieder als Menge haben, wenn $a_1, \dots, a_n$ Mengen sind. Das erreichen wir durch die folgende Vereinbarung

(Paarmengenaxiom) $$a \in \mathcal{U} \wedge b \in \mathcal{U} \quad \Rightarrow \quad \{a, b\} \in \mathcal{U}.$$

Um weitere Mengenbildungen zu beschreiben, führen wir zunächst die folgenden Operationen auf Klassen ein.

Klassenoperationen

Für zwei Klassen M und N sei:

(Vereinigung zweier Klassen) $$M \cup N := \{x |\ x \in M \vee x \in N\}$$

(Durchschnitt zweier Klassen) $$M \cap N := \{x |\ x \in M \wedge x \in N\}$$

(Klassendifferenz) $$M \backslash N := \{x |\ x \in M \wedge x \notin N\}$$

Allgemeiner definieren wir für eine Klasse $\{M_\iota |\ \iota \in I\}$ – eine solche Klasse nennt man auch eine *Mengenfamilie* mit *Indexmenge* I – die Klassen

(Vereinigung) $$\bigcup_{\iota \in I} M_\iota := \{x |\ (\exists \iota \in I)[x \in M_\iota]\}$$

und

(Durchschnitt) $$\bigcap_{\iota \in I} M_\iota := \{x |\ (\forall \iota \in I)[x \in M_\iota]\}.$$

Wir fordern nun, daß die Vereinigung über eine Mengenfamilie, deren Indexmenge eine Menge ist, ebenfalls wieder eine Menge ist. Dies wird formuliert im

(Vereinigungsmengenaxiom) $$I \in \mathcal{U} \wedge (\forall \iota \in I)[M_\iota \in \mathcal{U}] \quad \Rightarrow \quad \bigcup_{\iota \in I} M_\iota \in \mathcal{U}.$$

Man beachte, daß man jede Menge M als eine Mengenfamilie $\{x |\ x \in M\}$ auffassen kann. Daher schreibt man oft kurz $\bigcup M$ und $\bigcap M$ anstatt $\bigcup\limits_{x \in M} x$ bzw. $\bigcap\limits_{x \in M} x$.

Das Vereinigungsmengenaxiom vereinfacht sich dann zu

$$M \in \mathcal{U} \;\Rightarrow\; \bigcup M \in \mathcal{U}.$$

Eine weiteres Axiom zur Mengenbildung ist das

(Potenzmengenaxiom) $\quad M \in \mathcal{U} \;\Rightarrow\; \mathsf{Pow}(M) := \{x \mid\ x \subseteq M\} \in \mathcal{U}.$

Dabei nennt man $\mathsf{Pow}(M)$ die Potenzmenge (*pow*erset) von M. Das sind die wesentlichen Mengenbildungsoperationen, mit denen wir in dieser Vorlesung auskommen werden. Zu erwähnen ist noch, daß wir die Existenz einer unendlichen Menge aus den bisher eingeführten Axiomen nicht folgern können. Daher fordert man im Unendlichkeitsaxiom die Existenz einer unendlichen Menge. Bezeichnen wir mit $\mathbb{N}$ die Menge der natürlichen Zahlen, so genügt es zu fordern:

(Unendlichkeitsaxiom) $\quad \mathbb{N} \in \mathcal{U}.$

Wir geben das folgende Lemma ohne weiteren Beweis an. Die Beweise werden in der Vorlesung über Mengenlehre geführt. Hier können wir die Aussagen des Lemmas einfach als Grundtatsachen über Mengen voraussetzen.

0.3.2 Lemma *Es gelten*

(i) $\quad M, N \in \mathcal{U} \;\Rightarrow\; M \cup N \in \mathcal{U}$ *und* $M \cap N \in \mathcal{U}$

(ii) $\quad M \in \mathcal{U} \wedge M \neq \emptyset \;\Rightarrow\; \bigcap M \in \mathcal{U}$

Eine häufige Anfängerunsicherheit resultiert aus der extensionalen Gleichheit von Mengen. So sollte man beachten, daß

$$\{a, a\} = \{a\} \tag{1}$$

und

$$\{a, b\} = \{b, a\} = \{a, b, a\} = \{a, a, b\} = \dots \tag{2}$$

gilt.

0.4 Relationen

Will man die Reihenfolge der Elemente in einer Menge respektiert wissen, so muß man geordnetet Paare einführen. In der Sprache der Mengenlehre läßt sich das geordnete Paar definieren.

0.4.1 Definition Wir definieren

$$(a, b) := \{\{a\}, \{a, b\}\}$$

und nennen (a, b) das *geordnete Paar* von a und b.

0.4.2 Lemma $\quad (a_1, b_1) = (a_2, b_2) \Leftrightarrow a_1 = a_2 \wedge b_1 = b_2.$

Beweis: Wir zeigen zunächst die einfache Richtung "$\Leftarrow$:" Aus $a_1 = a_2$ und $b_1 = b_2$ erhalten wir zunächst $\{a_1\} = \{a_2\}$ und $\{a_1, b_1\} = \{a_2, b_2\}$ und damit auch $\{\{a_1\}, \{a_1, b_1\}\} = \{\{a_2\}, \{a_2, b_2\}\}$.
Die Gegenrichtung "$\Rightarrow$" ist deutlich aufwendiger zu zeigen. Hier beginnen wir mit

$$\begin{aligned} u = a_i \quad &\Leftrightarrow \quad u \in \{a_i\} \\ &\Leftrightarrow \quad u \in \{a_i\} \wedge u \in \{a_i, b_i\} \end{aligned} \tag{i}$$

für $i = 1, 2$. Wegen

$$y \in (a_1, b_1) \quad \Leftrightarrow \quad y = \{a_1\} \vee y = \{a_1, b_1\} \tag{ii}$$

folgt aus (i) und (ii)

$$u = a_i \quad \Leftrightarrow \quad (\forall y)[y \in (a_i, b_i) \Rightarrow u \in y]. \tag{iii}$$

Aus der Voraussetzung $(a_1, b_1) = (a_2, b_2)$ erhalten wir aus (iii)

$$\begin{aligned} u = a_1 \quad &\Leftrightarrow \quad (\forall y)[y \in (a_1, b_1) \Rightarrow u \in y] \\ &\Leftrightarrow \quad (\forall y)[y \in (a_2, b_2) \Rightarrow u \in y] \\ &\Leftrightarrow \quad u = a_2. \end{aligned} \tag{iv}$$

Damit folgt $a_1 = a_2$ und wir haben $(a_1, b_1) = (a_1, b_2)$, was

$$\{a_1, b_1\} \in \{\{a_1\}, \{a_1, b_2\}\} \tag{v}$$

und

$$\{a_1, b_2\} \in \{\{a_1\}, \{a_1, b_1\}\} \tag{vi}$$

nach sich zieht. Nun haben wir zwei Fälle zu unterscheiden:
1. Fall: $a_1 = b_1$. Dann ist nach (vi) $\{a_1, b_2\} \subseteq \{a_1, b_1\} = \{a_1\}$ und damit $b_2 = a_1 = b_1$.
2.Fall: $a_1 \neq b_1$. Dann ist nach (v) $\{a_1, b_2\} = \{a_1, b_1\}$ und mit Extensionalität folgt $b_2 = b_1$. $\square$

0.4.3 Definition Seien M und N Mengen, dann definieren wir

$$M \times N = \{(a, b) \mid a \in M \wedge b \in N\}$$

und nennen $M \times N$ das *kartesische Produkt* von M und N.

Wegen $(a,b) \in \mathsf{Pow}(\mathsf{Pow}(M \cup N))$ erhalten wir $M \times N \subseteq \mathsf{Pow}(\mathsf{Pow}(M \cup N))$. Mit Lemma 0.3.2 erhalten wir daher

$$M \in \mathcal{U} \wedge N \in \mathcal{U} \quad \Rightarrow \quad M \times N \in \mathcal{U}. \tag{3}$$

0.4.4 Definition Eine Menge $R \subseteq M \times N$ heißt eine *Relation* zwischen M und N. Wir sagen, daß $m \in M$ und $n \in N$ in Relation R stehen, wenn $(m,n) \in R$ ist. Im allgemeinen notieren wir dies in der “Infixnotation”, d.h. wir schreiben $m\ R\ n$ statt $(m,n) \in R$. Ist $M = N$, also $R \subseteq M \times M$, so sprechen wir von einer Relation auf M.

0.4.5 Definition Sei $R \subseteq M \times M$ eine Relation auf M. Wir nennen R

symmetrisch :⇔ $(\forall x \in M)(\forall y \in M)[x\ R\ y \ \Rightarrow\ y\ R\ x]$;

antisymmetrisch :⇔ $(\forall x \in M)(\forall y \in M)[x\ R\ y \wedge y\ R\ x \ \Rightarrow\ y = x]$;

reflexiv :⇔ $(\forall x \in M)[x\ R\ x]$;

antireflexiv :⇔ $(\forall x \in M)[\neg(x\ R\ x)]$;

transitiv :⇔ $(\forall x \in M)(\forall y \in M)(\forall z \in M)[x\ R\ y \wedge y\ R\ z \ \Rightarrow\ x\ R\ z]$;

linear oder konnex :⇔ $(\forall x \in M)(\forall y \in M)[x\ R\ y \vee y\ R\ x \vee x = y]$
⇔ $(\forall x \in M)(\forall y \in M)[x \neq y \ \Rightarrow\ x\ R\ y \vee y\ R\ x]$.

Die Relation R heißt eine

Halbordnung, wenn R reflexiv, antisymmetrisch und transitiv ist;

strikte Halbordnung, wenn R antireflexiv und transitiv ist;

Ordnung, wenn R eine lineare Halbordnung ist;

strikte Ordnung, wenn R eine lineare strikte Halbordnung ist;

Äquivalenzrelation, wenn R reflexiv, symmetrisch und transitiv ist.

0.4.6 Beispiele Für folgende, von der Schule her bekannten Relationen $<$, $\leq$, $=$ auf der Menge $\mathbb{N}$ der natürlichen Zahlen gilt:

$<$ ist lineare strikte Ordnung,

$\leq$ ist lineare Ordnung,

$=$ ist Äquivalenzrelation.

Als Beispiel einer vielleicht weniger bekannten Äquivalenzrelation führen wir die Kongruenzrelation “modulo k” auf der Menge $\mathbb{Z}$ der ganzen Zahlen wie folgt ein:

$$a \equiv b \mod k \quad :\Leftrightarrow \quad k|(a-b),$$

wobei $k|(a-b)$ für die Aussage steht, daß die ganze Zahl k die Differenz $a-b$ teilt. Man rechnet leicht nach, daß dies eine Äquivalenzrelation definiert.

Die Relation $n\ R\ m \quad :\Leftrightarrow \quad n|m$ definiert dann eine Halbordnung auf der Menge $\mathbb{Z}$ der ganzen Zahlen.

Ein weiteres Beipiel für eine Halbordnung ist die Teilmengenbeziehung auf der Potenzmenge $\mathsf{Pow}(M)$ einer Menge M.

0.4.7 Definition Ist $R \subseteq M \times N$ eine Relation zwischen M und N, so sei $\breve{R} := \{(b,a) |\ (a,b) \in R\}$ (gesprochen als “R *konvers*”) eine Relation zwischen N und M, d.h. $\breve{R} \subseteq N \times M$. Wir nennen $\breve{R}$ die zu R konverse Relation.

0.4.8 Definition Wir führen die folgenden Bezeichnungen und Redeweisen ein. Es sei $R \subseteq M \times N$ eine Relation. Dann definieren wir

$$\mathit{dom}\,(R) := \{x \in M |\ (\exists y \in N)[(x,y) \in R]\}.$$

Wir nennen *dom* (R) den *Definitionsbereich*, den *Vorbereich* oder den *Domain* von R;

$$\mathit{rng}\,(R) := \{y \in N |\ (\exists x \in M)[(x,y) \in R]\}.$$

Wir nennen *rng* (R) den *Nachbereich* (englisch *range*) von R.

$$\mathrm{feld}\,(R) := \mathit{dom}\,(R) \cup \mathit{rng}\,(R)$$

heißt das *Feld* der Relation R.

0.4.9 Definition Sind $R_1 \subseteq M \times N$ und $R_2 \subseteq N \times L$ Relationen, so sei

$$R_2 \circ R_1 := \{(a,b) \in M \times L |\ (\exists y \in N)[(a,y) \in R_1 \land (y,b) \in R_2]\}.$$

Wir nennen $R_2 \circ R_1$ die *Komposition* der Relationen R_1 und R_2.

0.5 Funktionen

0.5.1 Definition Eine Relation $R \subseteq Q \times Z$ heißt *rechtseindeutig*, wenn gilt:

$$(\forall x \in Q)(\forall y \in Z)(\forall z \in Z)[(x,y) \in R \land (x,z) \in R \quad \Rightarrow \quad y = z].$$

Eine rechtseindeutige Relation $f \subseteq Q \times Z$ heißt eine *partielle Funktion* mit Quelle Q und Ziel Z.

Wir notieren durch $f\colon Q \rightarrow_p Z$, daß F eine partielle Funktion mit Quelle Q und Ziel Z ist.
Ist $f\colon Q \rightarrow_p Z$ und ist $x \in$ *dom* (f), so gibt es genau ein Element y mit $(x, y) \in f$. Dieses y bezeichnen wir mit $f(x)$, also $(x, f(x)) \in f$. Die Menge *dom* (f) heißt der *Definitionsbereich* von f. Es ist

$$\mathit{dom}\,(f) = \{x \in Q \mid (\exists y)[(x, y) \in f]\}.$$

Üblicherweise betrachten wir Funktionen $f\colon Q \rightarrow_p Z$ mit *dom* $(f) = Q$. Solche Funktionen nennt man auch *total* und schreibt sie als $f\colon Q \longrightarrow Z$. Anstelle von totalen Funktionen sprechen wir auch von *Abbildungen*.
Die Menge $\mathrm{Im}\,(f) :=$ *rng* (f) heißt das *Bild* von der Funktion f. Es gilt

$$\mathrm{Im}\,(f) = \{f(x) \mid x \in \mathit{dom}\,(f)\}.$$

Oft schreibt man Funktionen in der Form

$$\begin{aligned} f\colon Q &\to Z \\ x &\mapsto f(x). \end{aligned}$$

0.5.2 Definition Eine (totale) Funktion $f : D \to Z$ heißt

injektiv $\quad :\Leftrightarrow \quad (\forall x)(\forall y)[x, y \in \mathit{dom}\,(f) \land f(x) = f(y) \Rightarrow x = y]$

surjektiv $\quad :\Leftrightarrow \quad (\forall y \in Z)(\exists x \in D)[y = f(x)]$;

bijektiv $\quad :\Leftrightarrow \quad$ *f ist injektiv und surjektiv.*

Durch Kontraposition der Definition der Injektivität erhalten wir für $f\colon D \longrightarrow Z$

$$f \text{ injektiv} \quad \Leftrightarrow \quad (\forall x \in D)(\forall y \in D)[x \neq y \Rightarrow f(x) \neq f(y)]. \tag{4}$$

Schreibweisen

Für $f\colon D \longrightarrow Z$ und $M \subseteq D$ sei

$$f[M] := \{f(x) \mid x \in M\}.$$

Wir nennen $f[M]$ *das Bild der Menge* M unter der Abbildung f. Mit dieser Schreibweise ist $f\colon D \longrightarrow Z$ surjektiv, wenn $f[D] = Z$ ist.
Für $N \subseteq Z$ sei

$$f^{-1}[N] := \{x \in D \mid f(x) \in N\}.$$

Die Menge $f^{-1}[N]$ heißt die Urbildmenge von N.

0.5.3 Lemma *Ist $f\colon M \longrightarrow N$ injektiv, so ist $\breve{f}\colon N \to_p M$ eine partielle Funktion mit dom$(\breve{f}) = \mathrm{Im}(f)$. Ist $f\colon M \longrightarrow N$ bijektiv, so ist $\breve{f} : N \to M$ eine (totale) bijektive Funktion, die üblicherweise mit f^{-1} bezeichnet wird und die zu f* inverse *Funktion genannt wird. Diese Bezeichnung ist von der Schreibweise $f^{-1}[M]$ wohl zu unterscheiden.*

Beweis: Wir haben für die erste Behauptung zu zeigen, daß $\breve{f}$ auf $\mathrm{Im}(f)$ definiert und rechtseindeutig ist. Aus der Injektivität von f folgt sofort die Rechtseindeutigkeit von $\breve{f}$. Zu $y \in \mathrm{Im}(f)$ gibt es ein $x \in M$ mit $y = f(x)$. Damit ist $(x, y) \in f$ und somit $(y, x) \in \breve{f}$, d.h. $y \in$ *dom*$(\breve{f})$. Damit ist $\mathrm{Im}(f) \subseteq$ *dom*$(\breve{f})$. Ist $y \notin \mathrm{Im}(f)$, so folgt $(\forall x \in M)[\neg(x, y) \in f]$ und damit auch $(\forall x \in M)[\neg(y, x) \in \breve{f}]$. Damit ist aber $y \notin$ *dom*$(\breve{f})$, und wir haben auch *dom*$(\breve{f}) \subseteq \mathrm{Im}(f)$ gezeigt. Ist f darüber hinaus auch noch surjektiv, so ist *dom*$(\breve{f}) = \mathrm{Im}(f) = N$ und $\breve{f}$ somit total. $\square$

0.5.4 Lemma *Sind $f\colon M \longrightarrow N$ und $g\colon N \longrightarrow L$ Funktionen, so ist auch $g \circ f\colon M \longrightarrow L$ eine Funktion.*

Beweis: Nach Definition 0.4.9 ist $g \circ f \subseteq M \times L$. Nachdem f und g Funktionen sind, ist $f(x)$ durch x und $g(y)$ durch y eindeutig bestimmt. Damit ist aber $g(f(x))$ durch x eindeutig bestimmt und es folgt $(g \circ f)(x) = g(f(x))$. Somit ist $g \circ f$ rechtseindeutig und damit eine Funktion. $\square$

0.5.5 Lemma *Ist $f\colon M \longrightarrow N$ bijektiv und bezeichne $id_M = \{(x, x) | x \in M\}$ (Identität auf M), so gilt $f^{-1} \circ f = id_M$ und $f \circ f^{-1} = id_N$.*

Beweis: Nach Lemma 0.5.3 und Lemma 0.5.4 sind sowohl $f \circ f^{-1}$ als auch $f^{-1} \circ f$ Funktionen. Für $x \in M$ erhalten wir $(x, f(x)) \in f$ und $(f(x), x) \in f^{-1}$. Damit ist $(x, x) \in f^{-1} \circ f$ und damit ist $f^{-1} \circ f = id_M$. Ganz analog zeigt man $f \circ f^{-1} = id_N$. $\square$

0.5.6 Lemma *Ist $f\colon M \longrightarrow N$ bijektiv, und ist $g : N \to M$ derart, daß $f \circ g = id_N$, so ist $g = f^{-1}$.*

Beweis: Wir haben zu zeigen, daß $g(y) = f^{-1}(y)$ für alle $y \in N$ gilt. Dies tun wir indirekt und nehmen dazu

$$(\exists y \in N)[g(y) \neq f^{-1}(y)] \qquad \text{(i)}$$

an. Wegen der Injektivität von f folgt daraus $y = f(g(y)) \neq f(f^{-1}(y)) = y$, was absurd ist. $\square$

0.5.7 Lemma *Sind* $f\colon M \longrightarrow N$ *und* $g : N \to L$ *bijektive Funktionen, so ist auch* $(g \circ f)$ *bijektiv, und es gilt* $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Beweis: Wir zeigen zunächst, daß $g \circ f$ injektiv ist. Seien dazu $x, y \in M$ mit $x \neq y$. Dann folgt wegen der Injektivität von f zunächst $f(x) \neq f(y)$ und wegen der Injektivität von g dann auch $(g \circ f)(y) = g(f(y)) \neq g(f(x)) = (g \circ f)(x)$.
Um die Surjektivität von $g \circ f$ zu erhalten, beginnen wir mit einem $y \in L$ und finden wegen der Surjektivität von g dazu ein $u \in N$ mit $y = g(u)$. Wegen der Surjektivität von f erhalten wir ein $x \in M$ mit $u = f(x)$. Setzen wir dies zusammen, so folgt $y = g(f(x)) = (g \circ f)(x)$, und wir haben ein Urbild von y unter $g \circ f$ gefunden.
Zum Schluß berechnen wir $(f^{-1} \circ g^{-1})(g \circ f)(x) = f^{-1}(g^{-1}(g(f(x)))) = f(f^{-1}(x)) = x$ und erhalten mit Lemma 0.5.6 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. □

0.6 Natürliche Zahlen

Eine wesentliche Rolle in mathematischen Betrachtungen kommt der Menge $\mathbb{N}$ der natürlichen Zahlen zu. Wir wollen daher einige Grundtatsachen über diese Menge zusammenstellen. Alle hier aufgelisteten Tatsachen lassen sich im Rahmen einer Mengenlehre mit Unendlichkeitsaxiom folgern. Wir wollen darauf hier aber nicht näher eingehen, sondern diese Tatsachen als Grundeigenschaften natürlicher Zahlen ansehen.
Natürliche Zahlen werden über den Prozeß des Zählens eingeführt. Eine Menge zu zählen, heißt ihre Elemente zu ordnen. Eine Element a kommt vor dem Element b wenn a vor b gezählt wurde. Umgekehrt können die Elemente einer geordneten Menge ihrer Ordnung nach gezählt werden. Allerdings eignet sich nicht jede Ordnung zum Zählen. So muß, damit wir weiterzählen können, die Ordung gewährleisten, daß jede verbleibende Restmenge von noch nicht gezählten Elementen ein kleinstes Element besitzt. Mit diesem Element können wir dann weiterzählen. Ordnungen mit dieser Eigenschaft heißen *Wohlordnungen*, die wir im folgenden formal korrekt einführen wollen.

0.6.1 Definition Eine strikte lineare Ordnung $\prec$ auf einer Menge M heißt eine *Wohlordnung* von M, wenn gilt:

- *feld* $(\prec) = M$
- $(\forall X)[X \subseteq M \wedge X \neq \emptyset \;\Rightarrow\; (\exists x \in X)(\forall y \in M)(y \prec x \;\Rightarrow\; y \notin X)]$,

d.h. wenn jede nicht leere Teilmenge des Feldes von $\prec$ ein bezüglich $\prec$ kleinstes Element besitzt.

Wir wollen hier die $<$–Relation auf den natürlichen Zahlen als bekannt voraussetzen. Es ist die Ordnung, die sich durch den Zählprozess ergibt. (Im Rahmen einer Mengenlehre würde sich die Ordnung durch die $\in$-Beziehung ergeben.) Wir wollen die folgende Vereinbarung treffen.

0.6.2 Vereinbarung Die Menge der natürlichen Zahlen $\mathbb{N} := \{0, 1, 2, \ldots\}$ wird durch ihre kanonische Ordnung $<$ wohlgeordnet.

0.6.3 Eigenschaften natürlicher Zahlen Ohne weitere Begründung halten wir die folgenden Eigenschaften natürlicher Zahlen fest:

($\mathbb{N}$0) $0 \in \mathbb{N}$;

($\mathbb{N}$1) *Jedes* $n \in \mathbb{N}$ *besitzt einen Nachfolger* $n+1 \in \mathbb{N}$ *mit* $n < n+1$ *und* $(\forall k < n+1)(k \leq n)$;

($\mathbb{N}$2) *Ist* $n+1 = m+1$, *so ist* $n = m$;

($\mathbb{N}$3) *Ist* $n \in \mathbb{N}$, *so ist* $n = 0$ *oder der Nachfolger einer natürlichen Zahl, d.h.* $(\forall n \in \mathbb{N})[n = 0 \vee (\exists m \in \mathbb{N})(n = m+1)]$.

($\mathbb{N}$4) $(\forall n \in \mathbb{N})[0 \neq n+1]$

0.6.4 Vollständige Induktion *Gilt*

$$(\forall n \in \mathbb{N})[(\forall m)(m < n \Rightarrow F(m)) \quad \Rightarrow \quad F(n)],$$

so folgt

$$(\forall n \in \mathbb{N})[F(n)].$$

Das Prinzip der vollständigen Induktion besagt also, daß wir, um $(\forall n \in \mathbb{N})F(n)$ zu zeigen, aus der *Induktionsvoraussetzung* $(\forall k < n)F(k)$ die *Induktionsbehauptung* $F(n)$ folgern müssen.
Beweis: Wir zeigen die Behauptung indirekt und nehmen

$$(\exists x \in \mathbb{N})[\neg F(x)] \tag{i}$$

an. Dann ist

$$M := \{x \in \mathbb{N} \mid \neg F(x)\} \neq \emptyset \tag{ii}$$

und hat, da $<$ eine Wohlordnung ist, ein kleinstes Element n_0. Damit gilt aber $(\forall k < n_0)[k \notin M]$, d.h. $(\forall k < n_0)[F(k)]$. Nach Voraussetzung folgt daraus aber $F(n_0)$ im Widerspruch zu $n_0 \in M$. $\square$

Äquivalent und vielleicht auch vertrauter ist eine Umformulierung der vollständigen Induktion, wie wir sie im folgenden angeben:

0.6.5 Vollständige Induktion *Gilt der* Induktionsanfang $F(k)$ *und der* Induktionsschritt $(\forall n \in \mathbb{N})[k \leq n \wedge F(n) \Rightarrow F(n+1)]$, *so erhalten wir bereits* $(\forall n \in \mathbb{N})[k \leq n \Rightarrow F(n)]$.

Beweis: Wir gehen wieder indirekt vor und nehmen $(\exists x \in \mathbb{N})[k \leq x \wedge \neg F(x)]$ an. Dann gibt es wieder ein kleinstes n_0 mit $k \leq n_0 \wedge \neg F(n_0)$. Wegen der Voraussetzung $F(k)$ ist $n_0 \neq k$ und damit ist $n_0 = m + 1$ mit $k \leq m < n_0$ nach $(\mathbb{N}3)$ und $(\mathbb{N}1)$. Wegen $m < n_0$ haben wir aber $F(m)$ und schließen nach Voraussetzung daraus auf $F(m+1)$, d.h. $F(n_0)$, was absurd ist. □

Man wird sich als Anfänger Gedanken darüber machen, wo denn in der ersten Fassung (0.6.4) der Induktionsanfang geblieben ist. Sehen Sie genauer hin, so werden Sie bemerken, daß die Formel $(\forall m)[m < 0 \Rightarrow F(m)]$ aus logischen Gründen wahr wird. Um die Implikation $(\forall m)[m < n \Rightarrow F(m)] \Rightarrow F(n)$ also auch für $n = 0$ zur Verfügung zu haben, muß man $F(0)$ haben. Dies ist der Induktionsbeginn.

0.6.6 Definition durch Rekursion *Sind die Abbildungen* $g\colon M \longrightarrow N$ *und* $h\colon \mathbb{N} \times N \times M \longrightarrow N$ *gegeben, so ist die durch die Rekursionsvorschrift*

$$\begin{aligned} &f(k,x) = g(x) \\ &(\forall n \in \mathbb{N})[k \leq n \Rightarrow f(n+1,x) = h(n, f(n,x), x)] \end{aligned} \tag{5}$$

definierte Funktion $f\colon \mathbb{N} \times M \longrightarrow N$ *eindeutig bestimmt.*

Beweis: Vorbemerkung: Die durch die obige Rekursionsvorschrift definierte Funktion ist wieder eine Menge, d.h. ein mathematisch sinnvolles Objekt. Im Rahmen einer Mengenlehre läßt sich die Existenz von f auch beweisen. Wir wollen hier auf einen solchen Beweis verzichten, denn wir "sehen die Funktion" ja als sinnvolles Objekt. Allerdings werden wir nachweisen, daß es nur eine Funktion geben kann, die den obigen Rekursionsgleichungen gehorcht.
Sie also $\tilde{f}$ eine weitere Funktion, die die Rekursionsgleichungen (5) erfüllt. Wir zeigen

$$(\forall n \in \mathbb{N})[\tilde{f}(n,x) = f(n,x)] \tag{i}$$

durch vollständige Induktion. Für den Induktionsbeginn hat man zunächst

$$\tilde{f}(k,x) = f(k,x) \tag{ii}$$

und für den Induktionsschritt haben wir die Induktionsvoraussetzung

$$\tilde{f}(n,x) = f(n,x) \tag{iii}$$

woraus sich, da $\tilde{f}$ und f beide die Rekursionsgleichungen (5) erfüllen,

$$\tilde{f}(n+1,x) = h(n,\tilde{f}(n,x),x) = h(n,f(n,x),x) = f(n+1,x) \tag{iv}$$

ergibt. Mit vollständiger Induktion erhalten wir daraus aber (i). □

Als Beispiel für eine Anwendung der Definition durch Rekursion wollen wir n-tupel für alle natürlichen Zahlen n definieren.

0.6.7 Definition Wir beginnen die Rekursion bei 1, d.h. wir setzen $k := 1$ in 0.6.6 und definieren

$$(a_1) := a_1$$

und

$$(a_1,\ldots,a_{n+1}) := ((a_1,\ldots,a_n),a_{n+1}),$$

wobei wir von der in Definition 0.4.1 definierten Funktion ausgehen, die zwei Mengen a, b das geordnete Paar (a,b) zuordnet.

0.6.8 Lemma $\quad (a_1,\ldots,a_n) = (b_1,\ldots,b_n) \;\Leftrightarrow\; a_1 = b_1 \wedge \ldots \wedge a_n = b_n.$

Beweis: Wir führen Induktion nach n. Für $n = 1$ ist die Behauptung klar. Gehen wir von der Induktionsvoraussetzung

$$(a_1,\ldots,a_n) = (b_1,\ldots,b_n) \;\Leftrightarrow\; a_1 = b_1 \wedge \ldots \wedge a_n = b_n \tag{i}$$

aus so erhalten wir mit 0.4.2

$$\begin{aligned}
(a_1,\ldots,a_{n+1}) = (b_1,\ldots,b_{n+1}) &\Leftrightarrow ((a_1,\ldots,a_n),a_{n+1}) = ((b_1,\ldots,b_n),b_{n+1}) \\
&\Leftrightarrow (a_1,\ldots,a_n) = (b_1,\ldots,b_n) \wedge a_{n+1} = b_{n+1} \\
&\Leftrightarrow a_1 = b_1 \wedge \ldots \wedge a_n = b_n \wedge a_{n+1} = b_{n+1},
\end{aligned}$$

wobei wir bei der letzten Äquivalenz die Induktionsvoraussetzung eingesetzt haben. □

Als Anwendung der n-tupel wollen wir n-fache kartesische Produkte einführen.

0.6.9 Kartesisches Produkt *Sind $A_1,\ldots,A_n$ Mengen, so definieren wir*

$$A_1 \times \ldots \times A_n := \{(a_1,\ldots,a_n) \mid a_1 \in A_1 \wedge \ldots \wedge a_n \in A_n\}.$$

Ebenso definieren wir

$$\begin{aligned}
A^1 &:= A \\
A^{n+1} &:= A^n \times A
\end{aligned}$$

und nennen A^n das n–fache kartesische Produkt von A. Es ist dann

$$A^n := \underbrace{A \times \ldots \times A}_{n-\textit{fach}}.$$

1. Vektorräume

1.1 Gruppen, Ringe und Körper

1.1.1 Definition Sei $H \neq \emptyset$ und $\cdot \colon H \times H \longrightarrow H$ eine Abbildung. Wir schreiben $a \cdot b$ anstatt $\cdot(a, b)$ und nennen $\cdot$ eine *Verknüpfung* auf H.
Das Paar $\mathcal{H} = (H, \cdot)$ heißt eine *Halbgruppe*, wenn die Verknüpfung $\cdot$ die folgende Identität erfüllt.

$$(\forall a \in H)(\forall b \in H)(\forall c \in H)[a \cdot (b \cdot c) = (a \cdot b) \cdot c].$$

Wir sprechen dann von einer *assoziativen* Verknüpfung.

1.1.2 Bemerkung Ist $\mathcal{H} = (H, \cdot)$ eine Halbgruppe und sind $a_1, \ldots, a_n \in H$, so hängt wegen der Assoziativität von $\cdot$ der Wert des Produkts $a_1 \cdots a_n$ nicht von der Beklammerung ab. Wir definieren daher

$$a^1 = a \text{ und } a^{n+1} = a \cdot a^n$$

und erhalten

$$a^{n+m} = a^n \cdot a^m \text{ und } (a^n)^m = a^{n \cdot m}. \tag{1.1}$$

1.1.3 Definition Sei $\mathcal{H} = (H, \cdot)$ eine Halbgruppe. Ein $e_L \in H$ heißt *linksneutral* in $\mathcal{H}$, wenn

$$(\forall a \in H)[e_L \cdot a = a]$$

gilt. Analog heißt $e_R \in H$ *rechtsneutral* in $\mathcal{H}$, wenn

$$(\forall a \in H)[a \cdot e_R = a]$$

gilt. Wir nennen $e \in H$ *neutral*, wenn e sowohl links- als auch rechtsneutral ist.

1.1.4 Satz *Eine Halbgruppe besitzt höchstens ein neutrales Element.*

Beweis: Sind $e_1, e_2 \in H$ neutral, so folgt sofort $e_1 = e_1 \cdot e_2 = e_2$. $\square$

1.1.5 Definition Sei $\mathcal{H} = (H, \cdot)$ eine Halbgruppe und $e_L \in H$ ein linksneutrales Element in $\mathcal{H}$. Gilt $a \cdot b = e_L$ für $a, b \in H$, so heißt a ein *Linksinverses* von b. Analog heißt $a \in H$ ein *Rechtsinverses* für $b \in H$, wenn $b \cdot a = e_R$ für ein rechtsneutrales $e_R \in H$ gilt.

1.1.6 Definition Eine *Gruppe* ist ein Paar $\mathcal{G} = (G, \cdot)$, wobei $\cdot$ eine Verknüpfung auf G ist, die den folgenden Bedingungen genügt:

(G1) $\mathcal{G} = (G, \cdot)$ *ist eine Halbgruppe.*

(G2) *Es gibt ein linksneutrales Element* $1_L \in G$.

(G3) *Zu jedem* $a \in G$ *existiert ein Linksinverses bzgl.* 1_L.

In Formelschreibweise lauten *(G2)* und *(G3)*

$$(\exists 1_L \in G)[(\forall a \in G)(1_L \cdot a = a) \wedge (\forall a \in G)(\exists b \in G)(b \cdot a = 1_L)].$$

1.1.7 Satz *Ist* $\mathcal{G} = (G, \cdot)$ *eine Gruppe mit linksneutralem Element* 1_L*, so gelten:*

(i) *Jedes linksinverse Element bezüglich* 1_L *ist auch rechtsinvers bzgl.* 1_L.

(ii) 1_L *ist das eindeutig bestimmte neutrale Element in* $\mathcal{G}$.

(iii) *Das linksinverse Element jedes* $a \in G$ *ist eindeutig bestimmt.*

Beweis: *(i)* Seien $a, b \in G$ mit $b \cdot a = 1_L$. Wir wählen $c \in G$ so, daß $c \cdot b = 1_L$ gilt. Dann ist $a \cdot b = 1_L \cdot a \cdot b = c \cdot b \cdot a \cdot b = c \cdot 1_L \cdot b = c \cdot b = 1_L$.
(ii) Sei $a \in G$. Nach *(i)* gibt es ein $b \in G$ mit $a \cdot b = b \cdot a = 1_L$. Dann gilt $a \cdot 1_L = a \cdot (a \cdot b) = (a \cdot b) \cdot a = 1_L \cdot a = a$. Nach 1.1.4 ist dann 1_L eindeutig bestimmt.
(iii) Sei $a \in G$ und $b \cdot a = a \cdot c = 1_L$. Dann folgt $b = b \cdot 1_L = b \cdot (a \cdot c) = (b \cdot a) \cdot c = 1_L \cdot c = c$. □

Mit 1_G bezeichnen wir das nach 1.1.7 eindeutig bestimmte neutrale Element einer Gruppe $\mathcal{G} = (G, \cdot)$, zu $a \in G$ sei a^{-1} das eindeutig bestimmte Inverse. Ist aus dem Zusammenhang klar, um welche Gruppe es sich handelt, so schreiben wir im allgemeinen 1 anstatt 1_G. Anstelle von $a \cdot b$ schreiben wir in Zukunft oft kurz ab.

1.1.8 Satz *Ein Halbgruppe* $\mathcal{G} = (G, \cdot)$ *ist genau dann eine Gruppe, wenn die Gleichungen* $a \cdot y = b$ *und* $y \cdot a = b$ *für alle* $a, b \in G$ *eine Lösung in* G *haben.*

Beweis: Ist $\mathcal{G}$ eine Gruppe, so ist $x := a^{-1}b$ eine Lösung von $ax = b$ und $y := ba^{-1}$ eine Lösung von $ya = b$.

Für die Gegenrichtung gehen wir davon aus, daß $\mathcal{G}$ eine Halbgruppe ist und die o.a. Gleichungen in G lösbar sind. Dann ist $G \neq \emptyset$ und wir wählen daher ein $b \in G$ und suchen eine Lösung der Gleichung $xb = b$. Sei diese e. Zu jedem $a \in G$ gibt es dann eine Lösung von $by = a$, d.h. ein $c \in G$ mit $bc = a$. Damit folgt $ea = ebc = bc = a$. Also ist e linksneutral. Da wir die Gleichung $ya = e$ für alle $a \in G$ lösen können, hat jedes $a \in G$ bezüglich e ein linksinverses Element. Somit ist $\mathcal{G}$ eine Gruppe. $\square$

1.1.9 Definition Sei $\mathcal{G} = (G, \cdot)$ eine Gruppe und $U \subseteq G$. Wir nennen U eine *Untergruppe von* $\mathcal{G}$, wenn $(U, \cdot)$ eine Gruppe ist, d.h. wenn

$$(\forall x \in U)(\forall y \in U)[xy \in U]$$

gilt und alle Gruppenaxiome erfüllt sind.

1.1.10 Satz *Sei $\mathcal{G} = (G, \cdot)$ eine Gruppe. Eine nicht-leere Menge $U \subseteq G$ ist genau dann eine eine Untergruppe von $\mathcal{G}$, wenn $(\forall a \in U)(\forall b \in U)[a \cdot b^{-1} \in U]$ gilt.*

Beweis: Ist $(U, \cdot)$ eine Untergruppe und sind $a, b \in U$, so ist auch $a, b^{-1} \in U$ und damit auch $ab^{-1} \in U$. Gelte nun umgekehrt $(\forall a \in U)(\forall b \in U)[ab^{-1} \in U]$. Dann folgt insbesondere $1 = aa^{-1} \in U$ und für $b \in U$ damit auch $1 \cdot b^{-1} = b^{-1} \in U$. Wegen $(b^{-1})^{-1} \cdot b^{-1} = 1 = b \cdot b^{-1}$ folgt $(b^{-1})^{-1} = b$ und für $a, b \in U$ erhalten wir demnach $a, b^{-1} \in U$ und schließlich $a \cdot (b^{-1})^{-1} = a \cdot b \in U$. Damit ist U abgeschlossen unter der Verknüpfung $\cdot$, enthält das neutrale Element 1 und mit jedem Element b auch sein inverses. Also ist $(U, \cdot)$ eine Gruppe und somit U eine Untergruppe von $\mathcal{G}$. $\square$

1.1.11 Definition Ist $\mathcal{G} = (G, \cdot)$ eine Gruppe und gilt

$$(\forall a \in G)(\forall b \in G)[a \cdot b = b \cdot a],$$

so heißt $\mathcal{G}$ *kommutativ* oder *abelsch*. Oft werden abelsche Gruppen additiv geschrieben, d.h. wir schreiben $a + b$ statt $a \cdot b$, 0 statt 1 und $-a$ statt a^{-1}. Anstelle von $a + (-b)$ schreibt man $a - b$.

1.1.12 Beispiele *(i)* Wir sehen leicht, daß die Menge $\mathbb{N}$ der natürlichen Zahlen zusammen mit der Addition ein Halbgruppe bilden. Wegen $0 + n = n$ für alle $n \in \mathbb{N}$ ist $(\mathbb{N}, +)$ eine Halbgruppe mit neutralen Element 0. Jedoch ist $(\mathbb{N}, +)$ keine Gruppe, da wir zu einem $n \in \mathbb{N}$ kein $m \in \mathbb{N}$ mit $n + m = 0$ finden können.
(ii) Wollen wir $(\mathbb{N}, +)$ zu einer Gruppe machen, so müssen wir zu jedem $n \in \mathbb{N}$ sein Inverses $(-n)$ hinzunehmen. Dies ergibt die Menge $\mathbb{Z} := \mathbb{N} \cup \{-n |\ n \in \mathbb{N}\}$

der *ganzen Zahlen*, die mit der Addition eine abelsche Gruppe bildet. Die Motivation zur Einführung ganzer Zahlen liegt natürlich darin, Gleichungen der Form $n + x = m$ lösen zu können.
(iii) Die Menge $\mathbb{Z}$ der ganzen Zahlen zusammen mit der Multiplikation bildet wieder eine Halbgruppe mit neutralem Element 1. Diese läßt sich nicht so leicht zu einer Gruppe machen, da sich für 0 keine inverses Element definieren läßt. Dennoch läßt sich $\mathbb{Z}$ durch Hinzunahme der Lösungen der Gleichungen $a \cdot x = b$ für $a \neq 0$ zu der Menge $\mathbb{Q}$ der *rationalen Zahlen* erweitern, deren von 0 verschiedenen Elemente zusammen mit der Multiplikation eine Gruppe bilden.
(iv) Ein Ihnen vielleicht weniger vertrautes Beispiel einer Gruppe erhalten wir, wenn wir *Permutationen* betrachten. Dazu bezeichne $\mathbb{N}_n$ die n-elementige Menge $\{1, \ldots, n\}$, die wir als paradigmatisch für eine beliebige n-elementige Menge betrachten wollen. Sei nun

$$S_n := \{\pi \,|\; \pi\colon \mathbb{N}_n \longrightarrow \mathbb{N}_n \textit{ bijektiv}\}. \tag{1.2}$$

Ein $\pi \in S_n$ heißt eine *Permutation* der Zahlen $1, \ldots, n$. Wir beobachten, daß $(S_n, \circ)$ eine Gruppe bildet, wobei $\circ$ wie im Abschnitt 0.5 die Komposition von Abbildungen bedeutet. Zunächst folgt aus Lemma 0.5.4, daß mit $\pi, \sigma \in S_n$ auch $\pi \circ \sigma \in S_n$ gilt. Aus Lemma 0.5.5 folgt, daß id_{S_n} ein neutrales Element für die Komposition ist und zu jedem $\pi \in S_n$ ein inverses Element π^{-1} existiert. Man nennt S_n die *symmetrische Gruppe*. Überzeugen Sie sich, daß diese Gruppe **nicht** kommutativ ist.

1.1.13 Definition Seien $R \neq \emptyset$, $+\colon R \times R \longrightarrow R$ und $\cdot\colon R \times R \longrightarrow R$ Verknüpfungen auf R. Das Tripel $\mathcal{R} = (R, +, \cdot)$ heißt ein *Ring*, wenn die folgenden Bedingungen erfüllt sind:

(R1) $(R, +)$ *ist eine additive abelsche Gruppe,*

(R2) $(R, \cdot)$ *ist eine Halbgruppe,*

(R3) *es gelten die* Distributivgesetze

$$(\forall a \in R)(\forall b \in R)(\forall c \in R)[a \cdot (b + c) = a \cdot b + a \cdot c \wedge (a + b) \cdot c = a \cdot c + b \cdot c]$$

Hat $(R, \cdot)$ ein neutrales Element, so heißt $\mathcal{R}$ ein *Ring mit* 1. Nach Satz 1.1.4 hat ein Ring höchstens eine 1.
Gilt $(\forall a \in R)(\forall b \in R)[a \cdot b = b \cdot a]$, so heißt $\mathcal{R}$ *kommutativ*.

1.1.14 Einige Rechenregeln für Ringe

(o) $-(-a) = a$

(i) $0 \cdot a = a \cdot 0 = 0$

(ii) $(-a) \cdot b = -(a \cdot b) = a \cdot (-b)$

(iii) $(-a) \cdot (-b) = a \cdot b$

Beweis: (o) Wegen $-(-a) + (-a) = 0$ und $a + (-a) = 0$ ist $-(-a) = a$ nach Satz 1.1.7 (iii).
(i) $0 \cdot a + b \cdot a = (0 + b) \cdot a = b \cdot a$. Also ist $0 \cdot a = 0$. Analog erhält man $a \cdot 0 + a \cdot b = a \cdot (0 + b) = a \cdot b$ und damit $a \cdot 0 = 0$.
(ii) $ab + (-a)b = (a - a)b = 0 \cdot b = 0$. Also ist $-ab = (-a)b$. Analog gilt $a(-b) + ab = a(b - b) = a \cdot 0 = 0$ und damit $-ab = a(-b)$.
(iii) $(-a)(-b) = -((-a)b) = (-(-(ab)) = ab$. □

1.1.15 Definition Sei $\mathcal{R} = (R, +, \cdot)$ ein Ring. Ein Element $a \in R$ heißt *linker Nullteiler* in $\mathcal{R}$, wenn $a \neq 0$ ist und es ein $b \in R$ gibt mit $b \neq 0$ und $a \cdot b = 0$.
Analog heißt $a \in R$ *rechter Nullteiler* in $\mathcal{R}$, wenn $a \neq 0$ ist und es ein $b \in R$ gibt mit $b \neq 0$ und $b \cdot a = 0$.
Ein Ring ohne Nullteiler heißt *nullteilerfrei*.
Ein nullteilerfreier Ring mit 1 heißt *Integritätsring*.

1.1.16 Lemma *Ist $\mathcal{R}$ nullteilerfrei, so gelten die Kürzungsregeln*

$$b \cdot a = b \cdot c \wedge b \neq 0 \;\Rightarrow\; a = c$$

und

$$a \cdot b = c \cdot b \wedge b \neq 0 \;\Rightarrow\; a = c.$$

Beweis: Aus $ba = bc$ und $b \neq 0$ folgt $b(a - c) = 0$. Da $b \neq 0$ und $\mathcal{R}$ keine Nullteiler hat, muß $a - c = 0$ sein. Damit folgt aber $a = c$. Die zweite Kürzungsregel folgt analog. □

1.1.17 Definition Sei $\mathcal{R} = (R, +, \cdot)$ ein Ring. Ist $U \subseteq R$, so heißt $(U, +, \cdot)$ ein *Unterring* oder *Teilring* von $\mathcal{R}$, wenn $(U, +, \cdot)$ ein Ring ist, d.h. wenn gilt

- $(U, +)$ *ist eine Untergruppe von* $(R, +)$

und

- $(\forall a, b \in U)[a \cdot b \in U]$.

1.1.18 Beispiele *(i)* Die Menge $\mathbb{Z}$ ist mit den Verknüpfungen $+$ und $\cdot$ ein Ring. Der Ring $\mathbb{Z}$ ist nullteilerfrei und besitzt eine 1, damit ist $\mathbb{Z}$ ein Integritätsring.

(ii) Eine weiteres Beispiel ist die endliche Menge $\mathbb{Z}/4 \cdot \mathbb{Z} := \{\underline{0}, \underline{1}, \underline{2}, \underline{3}\}$, dessen Verknüpfungen durch die Verknüpfungstafeln

$+$	$\underline{0}$	$\underline{1}$	$\underline{2}$	$\underline{3}$
$\underline{0}$	$\underline{0}$	$\underline{1}$	$\underline{2}$	$\underline{3}$
$\underline{1}$	$\underline{1}$	$\underline{2}$	$\underline{3}$	$\underline{0}$
$\underline{2}$	$\underline{2}$	$\underline{3}$	$\underline{0}$	$\underline{1}$
$\underline{3}$	$\underline{3}$	$\underline{0}$	$\underline{1}$	$\underline{2}$

$\cdot$	$\underline{1}$	$\underline{2}$	$\underline{3}$
$\underline{1}$	$\underline{1}$	$\underline{2}$	$\underline{3}$
$\underline{2}$	$\underline{2}$	$\underline{0}$	$\underline{2}$
$\underline{3}$	$\underline{3}$	$\underline{2}$	$\underline{1}$

gegeben sind. Man rechnet leicht nach, daß es sich mit den angebenen Verknüpfungen um einen Ring handelt. Man erhält $\mathbb{Z}/4 \cdot \mathbb{Z}$, indem man in den ganzen Zahlen "modulo 4" rechnet. Das heißt, man addiert und multipliziert die Elemente in $\{\underline{0}, \underline{1}, \underline{2}, \underline{3}\}$ als ob sie ganze Zahlen wären, teilt dann durch 4 und behält nur den Rest. Dies ergibt dann immer eine Zahl zwischen 0 und 3.
Beachte, daß $\mathbb{Z}/4 \cdot \mathbb{Z}$ wegen $\underline{2} \neq \underline{0}$ aber $\underline{2} \cdot \underline{2} = \underline{0}$ Nullteiler besitzt.
(iii) Das obige Beispiel läßt sich verallgemeinern zu $\mathbb{Z}/k \cdot \mathbb{Z} = \{\underline{0}, \underline{1}, \dots, \underline{k-1}\}$ indem man "modulo k" rechnet, d.h. die Elemente von $\{\underline{0}, \underline{1}, \dots, \underline{k-1}\}$ wie ganze Zahlen addiert und multipliziert, durch k teilt und nur den Rest behält. Versuchen Sie einmal, sich davon zu überzeugen, daß dieser Ring genau dann ein Integritätsring wird, wenn k eine Primzahl oder $k = 1$ ist.

1.1.19 Definition Ein Ring $\mathcal{K} = (K, +, \cdot)$ heißt ein *(Schief-) Körper*, wenn gilt

(K1) *$\mathcal{K}$ ist ein Ring,*

(K2) *$K^* = \{k \in K \mid \ k \neq 0\}$ bildet zusammen mit $\cdot$ eine Gruppe.*

1.1.20 Folgerungen *(i)* *Jeder Körper hat mindestens zwei Elemente.*

(ii) *Jede Gleichung der Form $a + x = b$ ist in einem Körper eindeutig lösbar.*

(iii) *Jede Gleichung der Form $a \cdot x = b$ und $y \cdot a = b$ ist für $a \neq 0$ in einem Körper eindeutig lösbar.*

(iv) *Jeder Körper ist ein Integritätsring.*

Einen Schiefkörper mit kommutativer multiplikativer Gruppe heißt ein *kommutativer Körper*. Oft spricht man nur von Körpern und meint damit kommutative Körper im Gegensatz zu Schiefkörpern.
Im folgenden betrachten wir nur noch Körper, d.h. Schiefkörper mit kommutativer multiplikativer Gruppe.

1.2 Vektorräume

Sei im Folgenden K stets ein kommutativer Körper. Die Elemente von K bezeichnen wir durch kleine griechische Buchstaben α, β, γ, $\ldots$.

1.2.1 Definition Eine Menge V heißt *Vektorraum* über dem Körper K – oder K-Vektorraum –, wenn gilt:

(V1) *Es gibt eine Verknüpfung* $+\colon V \times V \longrightarrow V$ *so, daß* $(V, +)$ *eine additive abelsche Gruppe ist.*

(V2) *Es gibt eine Abbildung* $\cdot\colon K \times V \longrightarrow V$ *(skalare Multiplikation) mit folgenden Eigenschaften:*

(V21) $(\alpha + \beta) \cdot a = \alpha \cdot a + \beta \cdot a$.

(V22) $\alpha \cdot (a + b) = \alpha \cdot a + \alpha \cdot b$.

(V23) $\alpha \cdot (\beta \cdot a) = (\alpha \cdot \beta) \cdot a$.

(V24) $1 \cdot a = a$.

Die Elemente von V heißen *Vektoren*. Anstatt $\alpha \cdot a$ schreiben wir oft kurz αa.

1.2.2 Folgerung *Sei* V *ein* K*-Vektorraum. Es gilt* $\alpha a = 0_V$ *genau dann, wenn* $a = 0_V$ *oder* $\alpha = 0_K$ *ist, wobei* 0_V *das neutrale Element (Nullelement) der additiven abelschen Gruppe* $(V, +)$ *und* 0_K *das Nullelement des Körpers* K *bedeuten.*

Beweis: Wir zeigen zuerst die Richtung $\Leftarrow$: Wir erhalten $\alpha a = (\alpha + 0_K)a = \alpha a + 0_K a$. Damit ist $0_K a = 0_V$. Analog ist $\alpha a = \alpha(a + 0_V) = \alpha a + \alpha 0_V$ und somit $\alpha 0_V = 0_V$.
Zum Beweis der Gegenrichtung $\Rightarrow$ sei $\alpha a = 0_V$ für $\alpha \neq 0$. Dann folgt $1 \cdot a = \alpha^{-1}\alpha a = \alpha^{-1} 0_V = 0_V$. Also ist $a = 0_V$ nach *(V24)*. □

1.2.3 Einige Rechenregeln

(i) $(-\alpha)a = -(\alpha a)$

(ii) $\alpha(-a) = -(\alpha a)$

Ist $f\colon \mathbb{N} \longrightarrow V$ *eine Abbildung, so definieren wir für* $n \geq m$ *rekursiv*

$$\begin{aligned}
&\sum_{i=m}^{m} f(i) := f(m) \\
&\sum_{i=m}^{n+1} f(i) := (\sum_{i=m}^{n} f(i)) + f(n+1).
\end{aligned} \tag{1.3}$$

Dann gelten:

(iii) $\quad \alpha(\sum_{i=m}^{n} a_i) = \sum_{i=m}^{n} (\alpha a_i)$

(iv) $\quad (\sum_{i=m}^{n} \alpha_i)a = \sum_{i=m}^{n} (\alpha_i a)$

(v) $\quad \sum_{i=m}^{n} (\alpha_i a_i) + \sum_{i=m}^{n} (\beta_i a_i) = \sum_{i=m}^{n} ((\alpha_i + \beta_i)a_i).$

Beweis: (i) $\alpha a + (-\alpha)a = (\alpha - \alpha)a = 0 \cdot a = 0$. Damit ist $-(\alpha a) = (-\alpha)a$.
(ii) $\alpha(-a) + \alpha a = \alpha(-a + a) = \alpha \cdot 0 = 0$. Also ist $-(\alpha a) = \alpha(-a)$.
(iii) Wir führen Induktion nach n. Den Induktionsbeginn liefert $n = m$. Dann erhalten wir $\alpha \sum_{i=m}^{m} a_i = \alpha a_m = \sum_{i=m}^{m} \alpha a_i$. Für den Induktionsschritt haben wir die Induktionsvoraussetzung $\alpha \sum_{i=m}^{n} a_i = \sum_{i=m}^{n} \alpha a_i$. Damit erhalten wir $\alpha \sum_{i=m}^{n+1} a_i = \alpha(\sum_{i=m}^{n} a_i + a_{n+1}) = \alpha \sum_{i=m}^{n} a_i + \alpha a_{n+1} = \sum_{i=m}^{n} \alpha a_i + \alpha a_{n+1} = \sum_{i=m}^{n+1} \alpha a_i$.
Analog zeigt man *(iv)*.
(v) Auch hier führen wir Induktion nach n. Für $n = m$ erhalten wir $\sum_{i=m}^{m} \alpha_i a_i + \sum_{i=m}^{m} \beta_i a_i = \alpha_m a_m + \beta_m a_m = (\alpha_m + \beta_m)a_m = \sum_{i=m}^{m} (\alpha_i + \beta_i)a_i$. Der Induktionsschritt ergibt sich aus $\sum_{i=m}^{n+1} \alpha_i a_i + \sum_{i=m}^{n+1} \beta_i a_i = \sum_{i=m}^{n} \alpha_i a_i + \alpha_{n+1}a_{n+1} + \sum_{i=m}^{n} \beta_i a_i + \beta_{n+1}a_{n+1} = \sum_{i=m}^{n} (\alpha_i + \beta_i)a_i + \alpha_{n+1}a_{n+1} + \beta_{n+1}a_{n+1} = \sum_{i=m}^{n} (\alpha_i + \beta_i)a_i + (\alpha_{n+1} + \beta_{n+1})\alpha_{n+1} = \sum_{i=m}^{n+1} (\alpha_i + \beta_i)a_i$. $\square$

1.2.4 Definition Sei V ein K-Vektorraum. Eine Teilmenge $U \subseteq V$ mit $U \neq \emptyset$ heißt ein *Teilraum* (oder *Untervektorraum*) von V, wenn

(U1) $\quad (\forall u \in U)(\forall v \in U)[u + v \in U]$

(U2) $\quad (\forall \alpha \in K)(\forall u \in U)[\alpha u \in U]$

gilt.

1.2.5 Satz *Sei V ein K-Vektorraum. Dann ist $U \subseteq V$, $U \neq \emptyset$ genau dann ein Teilraum von V, wenn U mit den Operationen $+, \cdot$ von V ein Vektorraum ist.*

Beweis: $\Rightarrow$: Sei U ein Teilraum von V. Dann ist $U \neq \emptyset$ und für $a, b \in U$ folgt $(-1)b = -b \in U$. Dann ist auch $a - b \in U$ und somit nach Satz 1.1.10 U eine Untergruppe der abelschen Gruppe. Da mit $\alpha \in K$ und $a \in U$ auch $\alpha a \in U$ ist, übertragen sich die übrigen Vektorraum-Axiome von V, wo sie ja nach Voraussetzung gelten, auf U.
$\Leftarrow$: Ist U ein Vektorraum, so ist $U \neq \emptyset$, da zumindest $0 \in U$ gelten muß. Darüber hinaus muß U als Vektorraum gegenüber Addition und skalarer Multiplikation abgeschlossen sein. □

1.2.6 Satz *Sei V ein K-Vektorraum und $\{U_i \mid i \in I\}$ eine Familie von Teilräumen von V. Dann ist $\bigcap \{U_i \mid i \in I\}$ wieder ein Teilraum von V.*

Beweis: Sei $U := \bigcap \{U_i \mid i \in I\}$. Da jedes der U_i zumindest die 0 enthalten muß, folgt $0 \in U$. Damit ist $U \neq \emptyset$. Gilt $a, b \in U$ und $\alpha \in K$ so erhalten wir $(\forall i \in I)[a, b \in U_i]$ und, da alle U_i Teilräume sind, $(\forall i \in I)[a + b \in U_i]$ sowie $(\forall i \in I)[\alpha a \in U_i]$. Damit ist $a + b \in U$ als auch $\alpha a \in U$ und U ist ein Teilraum. □

1.2.7 Definition Sei V ein K-Vektorraum und $M \subseteq V$. Dann definieren wir

$$\langle M \rangle := \bigcap \{U \subseteq V \mid M \subseteq U \wedge U \text{ ist Teilraum von } V\}.$$

Wir nennen $\langle M \rangle$ das *Erzeugnis* oder den *Span* von M oder auch den von M *aufgespannten Raum*. Die Menge M heißt ein *Erzeugendensystem* für $\langle M \rangle$.

Nach Satz 1.2.6 ist $\langle M \rangle$ ein Teilraum von V. Als Beispiele erhalten wir den Nullvektorraum, dessen einziger Vektor der Nullvektor ist als

$$\langle \emptyset \rangle = \mathbf{0} \quad \text{(Null-Vektorraum)}$$

und für $v \in V$

$$\langle \{v\} \rangle = \{\alpha v \mid \alpha \in K\}$$

den von einem einzigen Vektor v aufgespannten Raum. Anstelle von $\langle \{v_1, \ldots, v_n\} \rangle$ schreiben wir in Zukunft einfach $\langle v_1, \ldots, v_n \rangle$. Wir wollen dieses Beispiel im folgenden Satz noch vertiefen.

1.2.8 Satz *Sei V ein K-Vektorraum und $M \subseteq V$, dann gilt*

$$\langle M \rangle = \{\sum_{i=1}^{n} \alpha_i a_i \mid n \in \mathbb{N} \wedge \alpha_i \in K \wedge a_i \in M\}.$$

Man nennt einen Ausdruck der Form $\sum_{i=1}^n \alpha_i a_i$ eine *Linearkombination* der Vektoren $a_1, \ldots, a_n$. Der Span von M ist damit die Menge aller Linearkombinationen von Vektoren in M.
Beweis: Sei

$$U := \{\sum_{i=1}^n \alpha_i a_i |\ n \in \mathbb{N} \wedge \alpha_i \in K \wedge a_i \in M\},$$

d.h. U ist die Menge aller Linearkombinationen von Vektoren aus M. Zunächst ist klar, daß die Summe zweier Linearkombinationen von Vektoren aus M wieder eine Linearkombinationen von Vektoren aus M ist. Damit ist U gegenüber der Addition abgeschlossen. Offensichtlich ist für $a = \sum_{i=0}^n \alpha_i a_i \in U$ auch $\beta a = \sum_{i=0}^n (\beta\alpha_i) a_i \in U$. Damit ist U ein Teilraum von V. Wegen $a = 1 \cdot a$ gilt natürlich auch $a \in U$ für alle $a \in M$. Damit ist $M \subseteq U$. Um auch die umgekehrte Inklusion zu erhalten, müssen wir zeigen, daß jeder Teilraum von V, der M umfaßt, auch U umfaßt. Sei also $M \subseteq W$ und W ein Teilraum von V. Wählen wir $a \in U$, so gilt $a = \sum_{i=0}^n \alpha_i a_i$ mit $\alpha_i \in K$ und $a_i \in M$. Wegen $M \subseteq W$ erhalten wir dann aber sofort $a = \sum_{i=0}^n \alpha_i a_i \in W$, da W als Teilraum gegenüber skalarer Multiplikation und Addition abgeschlossen ist. □

1.2.9 Definition Ein K-Vektorraum heißt *endlich erzeugt*, wenn es eine endliche Teilmenge $M \subseteq V$ mit $\langle M \rangle = V$ gibt.

1.2.10 Satz *Ist V endlich erzeugt und gilt $V = \langle M \rangle$ für $M \subseteq V$, so gibt es eine endliche Teilmenge $M_0 \subseteq M$ mit $V = \langle M_0 \rangle$.*

Beweis: Sei $V = \langle v_1, \ldots, v_n \rangle$. Dann gibt es zu jedem v_i für $1 \leq i \leq n$ ein $n_i \in \mathbb{N}$ und Vektoren $a_{i1}, \ldots, a_{in_i} \in M$ mit $v_i = \sum_{k=1}^{n_i} \alpha_{ik} a_{ik}$. Dann erhalten wir $M_0 := \{a_{ik} |\ i = 1, \ldots, n; k = 1, \ldots, n_i\}$ als eine endliche Teilmenge von M und für jedes $v \in V$ folgt nach Satz 1.2.8

$$v = \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \sum_{k=1}^{n_i} \alpha_i \alpha_{ik} a_{ik} \in \langle M_0 \rangle.$$

Damit ist $V = \langle M_0 \rangle$. □

1.2.11 Definition Sei V ein K-Vektorraum. Eine Teilmenge $M \subseteq V$ heißt *linear abhängig* genau dann, wenn es eine endliche Teilmenge $\{a_1, \ldots, a_n\} \subseteq M$ und Elemente $\alpha_1, \ldots, \alpha_n \in K$ gibt mit

$$(\alpha_1 \neq 0 \vee \ \ldots \ \vee \alpha_n \neq 0) \wedge \sum_{i=1}^n \alpha_i a_i = 0.$$

Anderenfalls heißt M *linear unabhängig*. D.h. M ist genau dann linear unabhängig, wenn für jede endliche Teilmenge $\{a_1, \ldots, a_n\} \subseteq M$ und für alle Tupel $\alpha_1, \ldots, \alpha_n$ von Elementen aus K gilt:

$$\sum_{i=1}^{n} \alpha_i a_i = 0 \quad \Rightarrow \quad \alpha_1 = \ldots = \alpha_n = 0.$$

Der Begriff der linearen Unabhängigkeit ist **der** zentrale Begriff der linearen Algebra. Ist nun $\{v_1, \ldots, v_n\}$ eine endliche Menge linear abhängiger Vektoren eines Vektorraumes V, so bedeutet das, daß sich 0 aus $v_1, \ldots, v_n$ als nicht triviale Linearkombination darstellen läßt, d.h. daß es $\alpha_1, \ldots, \alpha_n \in K$ gibt, die nicht alle verschwinden und doch $\sum_{i=1}^{n} \alpha_i v_i = 0$ gilt. Sind $v_1, \ldots, v_n$ jedoch linear unabhängig, so läßt sich die 0 nur trivial darstellen, d.h. aus $\sum_{i=1}^{n} \alpha_i v_i = 0$ folgt $\alpha_1 = \ldots = \alpha_n = 0$.

1.2.12 Bemerkung *Ist $M \subseteq V$ mit $0 \in M$, so ist M linear abhängig wegen*

$$a_1, \ldots, a_n \in M \wedge a_1 = 0 \Rightarrow \sum_{i=1}^{n} \alpha_i a_i = 0 \textit{ für } \alpha_1 = 1, \alpha_2 = \ldots \alpha_n = 0.$$

1.2.13 Satz *Ist V ein K-Vektorraum und $a_1, \ldots, a_n \in V$ mit $n > 1$, so sind äquivalent:*

(i) $a_1, \ldots, a_n$ *sind linear abhängig,*

(ii) *eines der $a_i, i \in \{1, \ldots, n\}$ ist eine Linearkombination der übrigen, d.h. der Vektoren in $\{a_j \mid j \in \{1, \ldots, n\} \wedge i \neq j\}$,*

(iii) $\langle a_1, \ldots, a_n \rangle$ *wird bereits durch $n-1$ viele Elemente $a_{i_1}, \ldots, a_{i_{n-1}}$ erzeugt.*

Beweis: (i) ⇒ (ii): Sind $a_1, \ldots, a_n$ linear abhängig, so gibt es ein $j \in \{1, \ldots, n\}$ und ein $\alpha_j \in K$ mit $\alpha_j \neq 0$ und $\sum_{i=1}^{n} \alpha_i a_i = 0$. Also $a_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} {\alpha_j}^{-1}(-\alpha_i) a_i$.
(ii) ⇒ (iii): Sei $a_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} \alpha_i a_i$. Für $b \in \langle a_1, \ldots, a_n \rangle$ gilt dann $b = \sum_{k=1}^{n} \beta_k a_k = \sum_{\substack{k=1 \\ k \neq j}}^{n} \beta_k a_k + \beta_j \sum_{\substack{i=1 \\ i \neq j}}^{n} \alpha_i a_i = \sum_{\substack{k=1 \\ k \neq j}}^{n} (\beta_k + \beta_j \alpha_k) a_k \in \langle a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n \rangle$.
(iii) ⇒ (i): Ohne die Allgemeinheit einzuschränken, dürfen wir $\langle a_1, \ldots, a_n \rangle = \langle a_2, \ldots, a_n \rangle$ annehmen. Dann ist $a_1 = \sum_{i=2}^{n} \alpha_i a_i$ und damit $1 \cdot a_1 + \sum_{i=2}^{n} -\alpha_i a_i = 0$. Da $1 \neq 0$ ist, folgt daraus, daß $a_1, \ldots, a_n$ linear abhängig sind. □

1.2.14 Korollar *Jeder endlich erzeugte K-Vektorraum V besitzt ein linear unabhängiges Erzeugendensystem.*

Beweis: Wir führen den Beweis durch Induktion nach n. Sei $\{a_1, \ldots, a_n\}$ Erzeugendensystem von V. Für $n = 0$ ist $V = \mathbf{0}$, und wir sind fertig, da die leere Menge linear unabhängig ist. Ist $n = 1$, so folgt die Behauptung aus der Tatsache, daß ein einzelner Vektor $v \neq 0$ immer linear unabhängig ist. Sei also $n > 1$. Sind $a_1, \ldots, a_n$ linear unabhängig, so sind wir fertig. Anderenfalls können wir nach Satz 1.2.13 (iii) einen der Vektoren im Erzeugendensystems weglassen und erhalten so $\langle a_1, \ldots, a_n\rangle = \langle a_{i_1}, \ldots, a_{i_{n-1}}\rangle$. Nach Induktionsvoraussetzung enthält $\langle a_{i_1}, \ldots, a_{i_{n-1}}\rangle$ ein linear unabhängiges Erzeugendensystem und wir sind fertig. □

1.2.15 Definition Ein linear unabhängiges Erzeugendensystem B für einen K-Vektorraum V heißt eine *Basis* von V.

1.2.16 Satz *Ist V ein K–Vektorraum und B eine Basis von V, so läßt sich jedes von 0 verschiedene $v \in V$ eindeutig (bis auf Umnummerierung) als Linearkombination von Elementen aus B darstellen.*

Beweis: Da B ein Erzeugendensystem von V ist, finden wir zu $v \in V$ mit $v \neq 0$ Vektoren $b_1, \ldots, b_n \in B$ und Körperelemente $\alpha_1, \ldots, \alpha_n$ mit $v = \sum_{i=1}^n \alpha_i b_i$. Um die Eindeutigkeit der Darstellung zu zeigen, nehmen wir an, daß auch $v = \sum_{j=1}^m \beta_i b'_i$ mit $b'_i \in B$ und $\beta_i \in K$ gelte. Sei nun $\{c_1, \ldots, c_k\} := \{b_1, \ldots, b_n\} \cap \{b'{}_1, \ldots, b'{}_m\}$. Nach geeigneter Umnummerierung erhalten wir dann

$$v = \sum_{i=1}^{k} \alpha_i c_i + \sum_{i=k+1}^{n} \alpha_i b_i \quad \text{und} \quad v = \sum_{i=1}^{k} \beta_i c_i + \sum_{i=k+1}^{m} \beta_i b'_i.$$

Damit folgt $0 = \sum_{i=1}^k (\alpha_i - \beta_i) c_i + \sum_{i=k+1}^n \alpha_i b_i + \sum_{i=k+1}^m -\beta_i b'_i$. Da B eine linear unabhängige Menge ist, folgern wir daraus $\alpha_i = \beta_i = 0$ für $k < i$ und damit $m = n = k$ und $\alpha_i = \beta_i$ für $1 \leq i \leq k$. □

Wir wollen eine Menge $B \subseteq V$ von Vektoren *maximal linear unabhängig* nennen, wenn B linear unabhängig ist, aber die Menge $B \cup \{v\}$ für jedes $v \in V \setminus B$ linear abhängig ist.
Analog nennen wir eine Menge $E \subseteq V$ ein *minimales Erzeugendensystem*, wenn $V = \langle E\rangle$ ist, aber $\langle E \setminus \{v\}\rangle \neq V$ für jedes $v \in E$ gilt.

1.2.17 Satz *Für einen K–Vektorraum V und $B \subseteq V$ sind äquivalent:*

(i) B *ist eine Basis von* V.

(ii) B *ist eine maximale linear unabhängige Menge.*

(iii) B *ist ein minimales Erzeugendensystem.*

Beweis: (i) $\Rightarrow$ (ii): Ist $v \in V$ und $v \notin B$, so folgt $v = \sum_{i=1}^{n} \alpha_i a_i$ und damit $\sum_{i=1}^{n} \alpha_i a_i - v = 0$. Also ist $B \cup \{v\}$ linear abhängig. Damit ist B eine maximal linear unabhängige Menge.
(ii) $\Rightarrow$ (iii): Sei B maximal linear unabhängig. Dann gibt es zu $v \in V \setminus B$ Körperelemente $\alpha, \alpha_1, \dots, \alpha_n \in K$ sowie $b_1, \dots, b_n \in B$ mit $\alpha \neq 0$ und $\alpha v + \sum_{i=1}^{n} \alpha_i b_i = 0$. Damit ist $v = \sum_{i=1}^{n}(-\alpha_i \alpha^{-1}) b_i$. Damit ist jedes $v \in V$ eine Linearkombination von Elementen in B und B ist ein Erzeugendensystem. Wäre B nicht minimal, so fänden wir ein $b \in B$ so, daß $B \setminus \{b\}$ noch immer ein Erzeugensystem wäre. Insbesondere wäre dann $b = \sum_{i=1}^{n} \alpha_i b_i$ mit $b_i \in B \setminus \{b\}$. Das steht aber im Widerspruch zur linearen Unabhängigkeit von B.
(iii) $\Rightarrow$ (i): Nehmen wir an, daß B noch keine Basis ist, so muß B linear abhängig sein. Dann gibt es $b_1, \dots, b_n \in B$ und $\alpha_1, \dots, \alpha_n \in K$, die nicht alle verschwinden mit $0 = \sum_{i=1}^{n} \alpha_i b_i$. Ist $\alpha_j \neq 0$, so folgt $b_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} (-\alpha_i {\alpha_j}^{-1}) b_j$ und damit $V = \langle B \setminus \{b_j\} \rangle$. Dies steht aber im Widerspruch zur Minimalität von B als Erzeugendensystem. □

Unser Ziel wird es nun sein, den folgenden Satz zu beweisen.

1.2.18 Satz *Jeder K-Vektorraum V mit $V \neq \mathbf{0}$ besitzt eine Basis.*

Die Schwierigkeit im Beweis dieses Satzes liegt darin, daß es nicht endlich erzeugte Vektorräume gibt. Für endlich erzeugte Vektorräume ist Satz 1.2.18 eine sofortige Folgerung aus Korollar 1.2.14. Man kann ihn im Vereine mit Satz 1.2.10 sogar in der folgenden Art verschärfen.

1.2.19 Satz *Jeder endlich erzeugte Vektorraum, der nicht der Nullraum ist, besitzt eine endliche Basis.*

Im Falle von unendlich erzeugten Vektorräumen wird sich der Beweis schwieriger gestalten, da wir etwas mehr über die zugrunde liegende Mengenlehre wissen müssen. Wem dies für den Anfang zu schwierig erscheint, sollte den folgenden Abschnitt überschlagen und im Beweis von Satz 1.2.20 den “Limesfall” einfach auslassen.

Ein kleiner Exkurs ins Unendliche

Wir haben die natürlichen Zahlen als eine durch die kanonische Ordnung wohlgeordnete Menge eingeführt. Man abstrahiert nun von der zugrunde liegenden Menge und betrachtet nur den Typus der Ordnung (Mathematisch exakt ist der Ordnungstypus die Äquivalenzklasse von Ordnungen, wobei zwei Ordnungen äquivalent heißen, wenn sie sich ordnungstreu und bijektiv aufeinander abbilden lassen.). Die Ordnungstypen von Wohlordnungen bezeichnet man als *Ordinalzahlen*. Endliche Mengen lassen sich bis auf Äquivalenz nur auf eine Art ordnen. Den Ordnungstypus einer n-elementigen Menge identifiziert man mit der natürlichen Zahl n. Natürliche Zahlen lassen sich also als endliche Ordinalzahlen auffassen. Den Ordnungstypus der Ordnung aller natürlichen Zahlen in ihrer kanonischen Ordnung

$$\underbrace{0, 1, \dots}_{\omega}$$

bezeichnet man als ω. Dies ist der einfachste unendliche Ordnungstypus. Für unendliche Mengen gilt nun aber nicht mehr, daß sie sich bis auf Äquivalenz nur auf eine Art ordnen lassen. So erhalten wir eine neue Ordnung, wenn wir $m \prec_1 n \quad :\Leftrightarrow \quad (m \neq 0 \wedge n \neq 0 \wedge m < n) \vee (m \neq 0 \wedge n = 0)$ setzen. Dann haben wir eine Ordnung

$$\underbrace{1, 2, \dots}_{\omega}, 0, \tag{1.4}$$

die sich offensichtlich nicht mehr bijektiv ordnungstreu auf eine Menge vom Ordnungstypus ω abbilden läßt, da sie ein größtes Element, nämlich 0, besitzt, was bei ω nicht der Fall ist. Andererseits sind aber der Ordnungstypus von $0, 1, \dots$ und $1, 2, \dots$ offensichtlich gleich. Damit ist der Ordnungstypus von $1, 2, \dots, 0$ offenbar größer als ω, da er einen Ordnungstypus ω als ein echtes Anfangsstück enthält. Anschaulich kann man sich das so vorstellen, daß man eine Menge vom Ordnungstypus ω fertiggezählt hat und dann um eines weiter zählt. Es handelt sich also um eine Fortsetzung des Zählprozesses in das Unendliche. Die in (1.4) gezeigte Ordnung ist offensichtlich wieder eine Wohlordnung. Da sie um genau ein Element über ω hinauszählt, sagen wir, daß sie den Ordnungstypus $\omega + 1$ hat. Wir können den Prozeß fortsetzen, indem wir eine Wohlordnung $2, 3, \dots, 0, 1$ vom Ordnungstypus $\omega + 2$, eine Wohlordnung $\underbrace{k, k+1, \dots}_{\omega}, \underbrace{0, 1, \dots, k-1}_{k}$ vom Ordnungstypus $\omega + k$, eine Wohlordnung $\underbrace{0, 2, \dots}_{\omega}, \underbrace{1, 3 \dots}_{\omega}$ vom Ordnungstypus $\omega + \omega$ angeben, in der alle geraden Zahlen vor den ungeraden kommen und so fort.

Was man aus diesen Beispielen lernen sollte ist die Tatsache, daß drei Arten unendlicher Ordinalzahlen auftreten können:

- *Die Ordinalzahl* 0 *als der Ordnungstypus der leeren Menge, die der natürlichen Zahl* 0 *entspricht;*

- *Ordinalzahlen die einen Ordnungstypus repräsentieren, der ein größtes Element besitzt, wie z.B. jede endliche Ordinalzahl, die Ordinalzahlen* $\omega + 1$*,* $\omega + k$*,* $\dots$*; diese Ordinalzahlen heißen* Nachfolgerzahlen*, da zwischen ihrem Vorgänger, den man als Ordnungstypus der Ordnung erhält, in der man das größte Element einfach fortläßt, und der Ordinalzahl selbst keine weiteren Ordinalzahlen liegen.*

- *Ordinalzahlen, die einen Ordnungstypus repräsentieren, der kein größtes Element besitzt, wie z.b. ω, $\omega + \omega$, etc. Diese Ordinalzahlen nennt man* Limeszahlen. *Hat man nämlich eine Limeszahl λ und eine Ordinalzahl σ kleiner als λ, d.h. der Ordnungstypus σ ist äquivalent zu einem echten Anfangsstück einer durch λ repräsentierten Ordnung, so ist der Nachfolger $\sigma + 1$ von σ noch immer äquivalent zu einem echten Anfangsstück der durch λ repräsentierten Ordnung und damit kleiner als λ.*

Die Klasse On der Ordinalzahlen erweist sich selbst als wohlgeordnet. Damit folgt das Prinzip der Induktion über Ordinalzahlen genauso, wie wir es für natürliche Zahlen gefolgert haben. Wir erhalten also in Verallgemeinerung von 0.6.4:

$$\textit{Gilt } (\forall \xi \in \mathrm{On})[((\forall \zeta)(\zeta < \xi \Rightarrow F(\zeta))) \Rightarrow F(\xi)], \textit{ so folgt } (\forall \xi \in \mathrm{On})F(\xi). \tag{1.5}$$

Wollen wir 0.6.5 verallgemeinern, so habe wir die drei Typen von Ordinalzahlen zu beachten.

$$\begin{aligned}
&\textit{Gilt der Induktionsanfang } F(0),\\
&\textit{der Induktionsschritt } (\forall \xi \in \mathrm{On})[F(\xi) \Rightarrow F(\xi + 1)]\\
&\textit{und der Limesschritt } (\forall \xi < \lambda)F(\xi) \Rightarrow F(\lambda) \textit{ für Limeszahlen } \lambda,\\
&\textit{so folgt } (\forall \xi \in \mathrm{On})F(\xi).
\end{aligned} \tag{1.6}$$

Ebenso lassen sich Funktionen rekursiv über Ordinalzahlen definieren. Wir erhalten die folgende Verallgemeinerung von 0.6.6

Definition durch transfinite Rekursion *Sind die Abbildungen $g\colon M \longrightarrow N$ und $h\colon On \times N \times M \longrightarrow N$ gegeben, so ist die durch die Rekursionsvorschrift*

$$\begin{aligned}
&f(0, x) = g(x)\\
&(\forall \xi \in \mathit{On})[f(\xi + 1, x) = h(\xi, f(\xi, x), x)]\\
&f(\lambda, x) = \textstyle\bigcup_{\xi < \lambda} f(\xi, x) \textit{ für Limeszahlen } \lambda
\end{aligned}$$

definierte Funktion $f\colon On \times M \longrightarrow N$ eindeutig bestimmt.

Natürlich muß hierbei sicher gestellt sein, daß $\bigcup_{\xi<\lambda} f(\xi, x)$ immer wieder ein Element von N ist. Man kann die Rekursion wie in 0.6.6 statt bei 0 auch bei jeder beliebigen Ordinalzahl beginnen lassen.
Die Annahme, die wir über unsere Mengenwelt nun machen, ist der

Wohlordnungssatz *Jede Menge $M \in \mathcal{U}$ läßt sich wohlordnen.*

Der Ordnungstyp der kürzesten Wohlordnung, die sich auf einer Menge M definieren läßt, heißt die *Kardinalzahl* der Menge M und wird im allgemeinen mit $|M|$ bezeichnet. Man kann daher die Forderung des Wohlordnungssatzes auf die Form *"Jede Menge besitzt eine Kardinalzahl"* bringen.
Der Wohlordnungssatz mag einem zunächst merkwürdig anmuten. Eine Wohlordnung der Menge der reellen Zahlen kann man sich beispielsweise kaum vorstellen. Vor dem Hintergrund der übrigen Annahmen über unser Mengenuniversum ist der Wohlordnungssatz jedoch äquivalent zu einem anderen Prinzip, dem Auswahlaxiom, welches besagt:

Auswahlaxiom *Zu jeder Indexmenge I und Familie $\{M_\iota \mid \iota \in I\}$ von nicht leeren Mengen M_ι gibt es eine Auswahlfunktion $f\colon I \longrightarrow \bigcup_{\iota \in I} M_\iota$, so daß*

$$(\forall \iota \in I)(f(\iota) \in M_\iota)$$

gilt.

Das Auswahlaxiom besagt also, daß sich aus jedem Mitglied einer Familie nicht leerer Mengen in uniformer Weise ein Element auswählen läßt. Das klingt schon viel plausibler als der Wohlordnungssatz und ist auch ein in der Mathematik allgemein akzeptiertes Prinzip. Auf der anderen Seite leuchtet es ein, daß sich das Auswahlaxiom aus dem Wohlordnungssatz folgern läßt. Wir ordnen die Vereinigung der Mengen M_ι und wählen dann aus jedem Familienmitglied das kleinste Element aus. Auf der anderen Seite ist es auch möglich, mit Hilfe des Auswahlaxioms jede nicht leere Menge M zu zählen. Wir fangen einfach mit einem beliebigen Element an, zählen und entfernen dies aus M und setzten diese Prozedur mit Hilfe einer Auswahlfunktion auf der Potenzmenge von M so lange fort, bis wir zur leeren Menge gelangen.
Algebraikern ist der Begriffsapparat, der mit der Einführung von Ordinalzahlen verbunden ist, in der Regel viel zu aufwendig. (Obwohl dieses Gebiet für sich gesehen äußerst faszinierend ist. David Hilbert hat die Einführung der transfiniten Zahlen als eine der größten Leistungen der Mathematik bezeichnet.) Daher machen sie sich oft eine weitere Tatsache, das Lemma von Zorn, zu Nutze, die ihrerseits zum Auswahlaxiom und damit auch zum Wohlordnungssatz äquivalent ist. Um das Lemma zu formulieren, erinnern wir uns an den Begriff der Halbordnung. Das ist eine reflexive, antisymmetrische und transitive Relation, die nicht unbedingt linear sein muß, d.h. zwei Elemente des Feldes der Relation müssen nicht unbedingt vergleichbar sein. Haben wir eine Teilmenge $K \subseteq \text{feld}(R)$ in der je zwei Elemente vergleichbar sind, d.h. gilt $(\forall x \in K)(\forall y \in K)[x\ R\ y \vee x = y \vee y\ R\ x]$, so nennen wir K eine *Kette* in R. Ein Element $r \in \text{feld}(R)$ ist eine obere Schranke für eine Kette K, wenn $(\forall k \in K)[k\ R\ r]$ gilt. Wir nennen ein Element $m \in \text{feld}(R)$ maximal, wenn $m\ R\ r$ immer schon $m = r$ nach sich zieht.

Lemma von Zorn *Ist R eine Halbordnung, in der jede Kette eine obere Schranke besitzt, so hat R maximale Elemente.*

Die Äquivalenz von Zornschen Lemma und Auswahlaxiom plausibel zu machen, ist nicht ganz so einfach, und deshalb will ich hier auch nicht weiter darauf eingehen. Dies ist die Angelegenheit einer Vorlesung über Mengenlehre. Wir wollen daher hier unseren Exkurs ins Unendliche beenden.

1.2.20 Satz *Sei V ein K-Vektorraum und A ein Erzeugendensystem für V. Zu jeder linear unabhängigen Menge $B \subseteq V$ gibt es eine Teilmenge $C \subseteq A$ mit $B \cap C = \emptyset$ und $B \cup C$ ist eine Basis von V.*

Beweis: Sei $\kappa := |A|$. Dann können wir A schreiben als $A = \{a_\xi \mid \xi < \kappa\}$. (Wer den Exkurs in das Unendliche nicht mitgemacht hat, soll hier A als endlich annehmen. Dann ist κ eine natürliche Zahl und er kann den Limesfall übergehen, da unterhalb natürlicher Zahlen keine Limeszahlen liegen. Der Beweis bleibt aber unter der Voraussetzung, daß A ein endliches Erzeugendensystem ist, völlig korrekt.) Durch Rekursion nach $\xi < \kappa$ definieren wir uns Mengen M_ξ.

$$M_0 := B$$
$$M_{\xi+1} := \begin{cases} M_\xi \cup \{a_\xi\} & \text{falls diese Menge linear unabhängig ist} \\ M_\xi & \text{sonst} \end{cases}$$
$$M_\lambda := \bigcup_{\xi<\lambda} M_\xi \text{ für Limeszahlen } \lambda$$

und setzen

$$M := \bigcup_{\xi<\kappa} M_\xi.$$

Dann zeigen wir durch Induktion nach ξ, daß alle Mengen M_ξ linear unabhängig sind. Für $M_0 = B$ gilt dies nach Voraussetzung. Für $M_{\xi+1}$ gilt dies nach Konstruktion, und haben wir schließlich $a_1, \ldots, a_n \in M_\xi$ für eine Limeszahl ξ und nehmen wir $\sum_{i=1}^n \alpha_i a_i = 0$ an, so gibt es ein $\eta < \xi$ mit $a_1, \ldots, a_n \in M_\eta$. Nach Induktionsvoraussetzung ist aber M_η linear unabhängig und damit erhalten wir $\alpha_1 = \ldots = \alpha_n = 0$. Wir behaupten nun weiter, daß die Menge M maximal linear unabhängig ist. Da $M = \bigcup_{\xi<\kappa} M_\xi$ und alle M_ξ linear unabhängig sind, erhalten wir wie im vorherigen Limesfall, daß auch M linear unabhängig ist. Sei nun $v \in A$ und $v \notin M$. Dann gibt es ein $\xi < \kappa$ mit $v = a_\xi$. Damit ist $a_\xi \notin M_{\xi+1} \subseteq M$. Nach Konstruktion von $M_{\xi+1}$ muß daher $M_\xi \cup \{a_\xi\}$ linear abhängig sein. Dann gibt es aber Vektoren $a_1, \ldots, a_n \in M_\xi \subseteq M$ und Körperelemente $\alpha_1, \ldots, \alpha_n$ und β so, daß $\sum_{i=1}^n \alpha_i a_i + \beta a_\xi = 0$ ist, aber nicht alle α_i und β gleichzeitig 0 sind. Damit ist v eine Linearkombination von Elementen in $M_\xi \subseteq M$. Für $v \in M \cap A$ gilt dies trivialerweise. Für ein beliebiges $v \in V$ finden wir, da A ein Erzeugendensystem ist, Vektoren $a_{\xi_1}, \ldots, a_{\xi_n} \in A$ und Körperelemente $\alpha_1, \ldots, \alpha_n$ mit $v = \sum_{i=1}^n \alpha_i a_{\xi_i}$. Wir haben aber gerade gesehen, daß jedes der $a_{\xi_i} \in A$ eine Linearkombination von Elementen in M ist. Dann ist aber auch v eine Linearkombination von Elementen in M, und M ist auch ein Erzeugendensystem. Damit ist M eine Basis. Setzen wir $C := M \setminus B$, so ist $C \subseteq A$, $B \cap C = \emptyset$ und $M = B \cup C$, und der Satz ist bewiesen. □

Wir wollen, einfach weil man es in der Literatur oft so findet, den Satz nochmals mit Hilfe des Lemma von Zorn beweisen. Dazu betrachten wir die Menge $\mathfrak{P} := \{B \cup D \mid D \subseteq A \wedge B \cup D \text{ ist linear unabhängig}\}$. Diese Menge wird durch die Mengeninklusion $\subseteq$ halbgeordnet. Sei nun $\{E_\iota \mid \iota \in I\}$ eine Kette in dieser Halbordnung. Jedes E_ι hat die Gestalt $B \cup D_\iota$ und damit hat $E := \bigcup_{\iota \in I} E_\iota$ die Gestalt $B \cup \bigcup_{\iota \in I} D_\iota$. Wie im vorherigen Beweis folgt, daß E als Vereinigung linear unabhängiger Mengen selbst linear unabhängig ist. Damit hat jede Kette in $(\mathfrak{P}, \subseteq)$

eine obere Schranke und besitzt nach dem Lemma von Zorn maximale Elemente. Sei M ein solches maximales Element. Wegen $M \in \mathfrak{P}$ ist dann M linear unabhängig. Ist $v \in A$, so folgt aus der Maximalität von M, daß $M \cup \{v\} = M$ oder $M \cup \{v\}$ linear abhängig ist. In beiden Fällen läßt sich v als Linearkombination von Elementen in M darstellen. Da A ein Erzeugendensystem war, gilt $v = \sum_{i=1}^{n} \alpha_i a_i$ mit $a_1, \ldots, a_n \in A$ für jedes $v \in V$. Da sich jedes $a_i \in A$ als Linearkombination von Elementen in M darstellen läßt, erhalten wir auch v als Linearkombination von Elementen aus M. Damit ist M auch ein Erzeugendensystem und somit eine Basis, und wir setzen wie im vorherigen Beweis $C := M \setminus A$. □

Beide Beweise unterscheiden sich nur in der Konstruktion der Menge M. Die sich daran anschließende Argumentation ist im wesentlichen identisch. Die Konstruktion der Menge M ist im ersten Beweis durchsichtiger (beim Zornschen Lemma fällt sie vom Himmel), und der Vorteil des ersten Beweises durch transfinite Rekursion liegt auch darin, daß sich der endlich erzeugte Fall einfach durch Weglassen des Limesfalles ergibt. Der Beweis über das Zornsche Lemma hat seinerseits den Vorteil, daß er neben dem Zornschen Lemma keinen zusätzlichen Begriffsapparat benötigt.
Wie bereits gesagt werden in dieser Vorlesung endlich erzeugte Vektorräume die Hauptrolle spielen. Für diese Vektorräume erhalten wir Satz 1.2.20 ohne den Exkurs ins Unendliche.

1.2.21 Korollar *Jeder Vektorraum besitzt eine Basis.*

Beweis: Jeder Vektorraum V besitzt ein Erzeugendensystem, z.B. V selbst. Da die leere Menge linear unabhängig ist, können wir in Satz 1.2.20 $B = \emptyset$ setzen und erhalten so die Behauptung. □

1.2.22 Satz *Ist V ein K-Vektorraum und sind B_1 und B_2 Basen von V, so gilt einer der folgenden Fälle:*

(i) $V = \mathbf{0}$ *und* $B_1 = B_2 = \emptyset$,

(ii) B_1 *und* B_2 *sind beide unendlich,*

(iii) V *ist endlich erzeugt und* $|B_1| = |B_2|$.

Beweis: Sei $V \neq \mathbf{0}$. Ist V nicht endlich erzeugt, so kann V keine endliche Basis besitzen. Sei daher V endlich erzeugt und $B_1 = \{a_1, \ldots, a_n\}$ und $B_2 = \{b_1, \ldots, b_m\}$ Basen. Dann ist die Menge $A_1 := \{a_2, \ldots, a_n\}$ linear unabhängig und nach Satz 1.2.17 keine Basis. Nach Satz 1.2.20 finden wir eine Menge $C_1 \subseteq B_2$ mit $A_1 \cap C_1 = \emptyset$ so, daß $A_1 \cup C_1$ eine Basis ist. Damit ist insbesondere $C_1 \neq \emptyset$.

Nun definieren wir $A_2 := \{a_3, \dots, a_n\} \cup C_1$. Diese Menge ist wieder linear unabhängig und kein Erzeugendensystem. Nach Satz 1.2.20 finden wir eine Menge $C_2 \subseteq B_2$ so, daß $A_2 \cap C_2 = \emptyset$ und $A_2 \cup C_2$ eine Basis ist. Wir iterieren nun dieses Verfahren und erhalten nach n–vielen Schritten eine Basis $C_1 \cup \ldots \cup C_n \subseteq B_2$, so daß alle C_i paarweise verschieden sind. Also ist $|B_1| = n \leq |B_2|$. Vertauschen wir die Rollen von B_1 und B_2, so erhalten wir mit dem gleichen Verfahren $|B_2| \leq |B_1|$. □

Diese Stelle erscheint mir geeignet zu sein, einmal zu präzisieren, wie das obige Argument "*Wir iterieren nun dies Verfahren...* " zu verstehen ist. Dahinter steckt eine Definition durch Rekursion und ein Beweis durch Induktion. Zunächst definieren wir durch Rekursion Mengen $A_i = \{a_{i+1}, \dots, a_n\}$ und $C_i \subseteq B_2$ derart, daß $A_i \cup \bigcup_{j \leq i} C_j$ eine Basis bilden und $(A_i \cup \bigcup_{j<i} C_j) \cap C_i = \emptyset$ gilt. Wir setzen $C_0 := \emptyset$. Als Induktionvoraussetzung gehen wir davon aus, daß $A_i = \{a_{i+1}, \dots, a_n\}$ und C_i definiert sind und die geforderten Eigenschaften besitzen. Dann ist die Menge $A_{i+1} \cup \bigcup_{j \leq i} C_j$ linear unabhängig und, da sie eine echte Teilmenge der Basis $A_i \cup \bigcup_{j \leq i} C_j$ ist, nach Satz 1.2.17 keine Basis. Da B_2 ein Erzeugendensystem ist, gibt es nach Satz 1.2.20 eine Menge $C_{i+1} \subseteq B_2$, so daß $(A_{i+1} \cup \bigcup_{j \leq i} C_j) \cup C_{i+1}$ eine Basis und $(A_{i+1} \cup \bigcup_{j \leq i} C_j) \cap C_{i+1} = \emptyset$ ist. Damit sind A_i und C_i für $i = 0, \dots, n$ definiert, es ist $A_n = \emptyset$, $C_1 \cup \ldots \cup C_n$ eine Basis und alle C_i sind paarweise verschieden.
Wann immer Konstruktionen durch "*Iteration*" fortgesetzt werden, sind diese analog wie im obigen Beispiel zu präzisieren.

1.2.23 Definition (**Dimension eines Vektorraumes**) Sei V ein K-Vektorraum. Wir definieren dessen Dimension durch

$$\dim_K V := \begin{cases} 0, & \text{falls } V = \mathbf{0} \text{ ist;} \\ \infty, & \text{falls } V \text{ nicht endlich erzeugt ist;} \\ |B|, & \text{falls } B \text{ eine Basis von } V \text{ ist.} \end{cases}$$

Ist es klar oder unerheblich, um welchen Grundkörper K es sich handelt, so schreiben wir kurz $\dim V$ statt $\dim_K V$.

1.2.24 Satz *Ist V ein K-Vektorraum und sind $a_1, \dots, a_n \in V$ linear unabhängige Vektoren, so sind äquivalent:*

(i) $\dim V = n$.

(ii) $a_1, \dots, a_n$ *ist eine Basis.*

Beweis: (i) ⇒ (ii): Ist B eine Basis von V, so gibt es nach Satz 1.2.20 eine Menge $C \subseteq B$, so daß $\{a_1, \dots, a_n\} \cup C$ eine Basis von V und $\{a_1, \dots, a_n\} \cap C = \emptyset$

ist. Nach Satz 1.2.22 müssen dann aber B und $\{a_1, \dots, a_n\} \cup C$ genau n Elemente besitzen, woraus sich $C = \emptyset$ ergibt, und daher ist $\{a_1, \dots, a_n\}$ schon eine Basis.
(ii) $\Rightarrow$ (i): Ist $\{a_1, \dots, a_n\}$ eine Basis, so ist $\dim V = |\{a_1, \dots, a_n\}| = n$. $\square$

1.2.25 Satz *Sei V ein endlich erzeugter Vektorraum. Dann gelten folgende Aussagen:*

(i) *Sind $a_1, \dots, a_n \in V$ linear unabhängig, so ist $n \leq \dim V$.*

(ii) *Wird V von $b_1, \dots, b_m \in V$ erzeugt, so ist $\dim V \leq m$.*

Beweis: (i): Sind $a_1, \dots, a_n$ linear unabhängig, so lassen sie sich nach Satz 1.2.20 zu einer Basis $\{a_1, \dots, a_n\} \cup C$ mit $C \subseteq V$ ergänzen. Also ist $n \leq \dim V$.
(ii): Ist $V = \langle b_1, \dots, b_m \rangle$, so enthält $\{b_1, \dots, b_m\}$ nach Satz 1.2.10 ein minimales Erzeugendensystem. Also ist $\dim V \leq m$. $\square$

1.2.26 Bemerkung *Man beachte, daß die Kontraposition der Aussage (i) in Satz 1.2.25 besagt, daß für $m > \dim V$ je m viele Vektoren stets linear abhängig sind.*

1.2.27 Satz *Ist V ein endlich dimensionaler Vektorraum und U ein Teilraum von V, so gelten:*

(i) $\dim U \leq \dim V$.

(ii) $\dim U = \dim V \quad \Leftrightarrow \quad U = V$.

Beweis: (i): Ist B eine Basis von U, so ist B eine in V linear unabhängige Menge. Damit ist $\dim U = |B| \leq \dim V$.
(ii): $\Rightarrow$: Ist B eine Basis von U, so ist B linear unabhängig in V. Aber B ist auch maximal linear unabhängig in V, da jede Erweiterung von B mehr als $\dim U = \dim V$ viele Elemente enthalten müßte, was nach Bemerkung 1.2.26 unmöglich ist. Damit ist B eine Basis von V und wir erhalten $U = \langle B \rangle = V$.
Die Gegenrichtung $\Leftarrow$ gilt offensichtlich.

1.3 Einige Beispiele

1.3.1 Der Vektorraum K^n

Sei K ein Körper. Wir definieren

$$K^n := \{(\alpha_1, \dots, \alpha_n) \bigm| \alpha_i \in K \text{ für } i = 1, \dots, n\} \tag{1.7}$$

und nennen K^n die Menge der n-tupel über K. Wir wollen Verknüpfungen auf dem K^n so definieren, daß K^n mit diesen Verknüpfungen zu einem Vektorraum wird. Wir beginnen mit der Addition und setzen

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n). \tag{1.8}$$

Man nennt diese Art der Definition, die sich in jeder Komponente des n-tupels auf Addition im Grundkörper bezieht, auch eine *komponentenweise* oder *punktweise* Definition. Die Addition im K^n wird also komponentenweise festgesetzt. Ebenso definieren wir die skalare Multiplikation punkt- (oder komponenten)weise durch

$$\alpha \cdot (\alpha_1, \ldots, \alpha_n) := (\alpha \cdot \alpha_1, \ldots, \alpha \cdot \alpha_n). \tag{1.9}$$

Die Tatsache, daß $(K, +)$ eine Gruppe ist, überträgt sich dank der komponentenweise Definition unmittelbar auf $(K^n, +)$. Man rechnet dies leicht nach. Das neutrale Element ist das n-tupel, das aus lauter Nullen besteht. Das inverse Element zu $(\alpha_1, \ldots, \alpha_n)$ ergibt sich sofort als $(-\alpha_1, \ldots, -\alpha_n)$. Die Axiome *(V21)–(V24)* rechnet man ebenfalls sofort nach. Damit ist K^n ein Vektorraum. Dies ist natürlich auch für $n = 1$ richtig. Jeder Körper kann also als Vektorraum über sich selbst aufgefaßt werden.
Setzen wir $K = \mathbb{R}$, so erhalten wir die bekannten Vektorräume $\mathbb{R}^1$, die Gerade, $\mathbb{R}^2$, die euklidische Ebene, und $\mathbb{R}^3$, den euklidischen Raum. Diese mag man sich als anschauliche Beispiele für Vektorräume vor Augen halten. Dabei soll man sich aber stets bewußt sein, daß diese Vektorräume noch viel zusätzliche Struktur tragen (z.B. rechtwinkelige Koordinatensysteme), die wir im bisher eingeführten allgemeinen Fall noch nicht zur Verfügung haben.
Kehren wir daher wieder zum allgemeinen Fall eines beliebigen Körpers K zurück. Eine nützliche Notationshilfe ist das von Kronecker eingeführte Symbol

$$\delta_{ij} := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases} \tag{1.10}$$

Mit Hilfe dieses Kroneckersymbols definieren wir die Vektoren

$$\mathbf{e}_i^n := (\delta_{i1}, \ldots, \delta_{in}). \tag{1.11}$$

Die Komponenten des Vektors $\mathbf{e}_i^n$ sind daher alle 0 bis auf die i-te Stelle, an der eine 1 steht. Das Superskript n lassen wir im allgemeinen weg, da es sich im Regelfall aus dem Zusammenhang ergibt. Gilt nun $\sum_{i=1}^n \alpha_i \mathbf{e}_i = 0$, so erhalten wir $\alpha_i \cdot 1 = 0$ für $i = 1, \ldots n$. Da jeder Körper nullteilerfrei ist, folgt daraus $\alpha_1 = \ldots = \alpha_n = 0$. Damit sind $\mathbf{e}_1, \ldots, \mathbf{e}_n$ linear unabhängig. Anderenfalls gilt für $(\alpha_1, \ldots, \alpha_n) \in K^n$ stets $(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^n \alpha_i \mathbf{e}_i$. Damit ist $\mathbf{e}_1, \ldots, \mathbf{e}_n$ auch ein Erzeugendensystem und somit eine Basis. Also haben wir

1.3.1 Satz *Die Menge K^n der n-tupel über einen Körper K bilden mit der komponentenweisen Addition und skalaren Multiplikation einen Vektorraum der Dimension n. Die Menge $\mathbb{E}^n := \{\mathbf{e}^n{}_1, \ldots, \mathbf{e}^n{}_n\}$ ist eine Basis des K^n. Sie heißt dessen* kanonische *Basis.*

1.3.2 Der Vektorraum der Abbildungen von einer Menge in einen Körper

Ein n-tupel $(\alpha_1, \ldots, \alpha_n) \in K^n$ läßt sich auch als eine Abbildung der Menge $f\colon \{1, \ldots, n\} \longrightarrow K$ auffassen, die gegeben ist durch $f(i) := \alpha_i$. Die naheliegende Verallgemeinerung ist, nun die Menge aller Abbildungen einer Menge X in einen Körper zu betrachten. Sei also

$$\text{Abb}(X, K) := K^X := \{f \mid f\colon X \longrightarrow K\}. \tag{1.12}$$

Wir definieren auf $\text{Abb}(X, K)$ eine Addition und skalare Multiplikation durch

$$\begin{aligned}(f+g)(x) &:= f(x) + g(x)\\ (\alpha \cdot f)(x) &:= \alpha \cdot f(x).\end{aligned} \tag{1.13}$$

Die Definition (1.13) ist so zu verstehen, daß zwei Abbildungen f und g zu einer neuen Abbildung $f + g$ verknüpft werden, die wie in (1.13) definiert ist. Da $f(x)$ und $g(x)$ Elemente von K sind, ist $f(x)+g(x)$ durch die Addition im Körper wohldefiniert. Ebenso werden $\alpha \in K$ und $f \in K^X$ die in (1.13) definierte Abbildung $\alpha \cdot f$ zugeordnet. Man rechnet wieder sofort nach, daß $(K^X, +)$ eine additive abelsche Gruppe bildet, deren neutrales Element durch die Nullabbildung $\mathbf{0}(x) := 0$ für alle $x \in X$ gegeben ist. Das inverse Element einer Abbildung f erhalten wir als $-f = -1 \cdot f$.
Wir erhalten einen Unterraum von K^X durch die Festsetzung

$$\text{Abb}[X, K] := \{f \in K^X \mid f(x) = 0 \text{ für fast alle } x \in X\}. \tag{1.14}$$

Dabei bedeutet *“für fast alle $x \in X$”*, daß dies für alle $x \in X$ bis auf endlich viele Ausnahmen gilt. Wir wollen uns als erstes davon überzeugen, daß $\text{Abb}[X, K]$ auch tatsächlich ein Unterraum ist. Sind f und g in $\text{Abb}[X, K]$, so gibt es endliche Mengen $I_1, I_2 \subseteq X$ mit $(\forall x \in I_1)[f(x) \neq 0]$ und $(\forall x \in I_2)[g(x) \neq 0]$. Dann gilt aber $(f+g)(x) \neq 0 \;\Rightarrow\; x \in I_1 \cup I_2$, d.h. $(f+g)(x) = 0$ für fast alle $x \in X$, denn mit I_1 und I_2 ist auch $I_1 \cup I_2$ endlich. Ebenso erhalten wir, daß mit $f \in \text{Abb}[X, K]$ und $\alpha \in K$ auch $\alpha f \in \text{Abb}[X, K]$ ist.
Ist X bereits eine endliche Menge, so sind $\text{Abb}(X, K)$ und $\text{Abb}[X, K]$ offenbar identisch. Ist $|X| = n$, so ist, wie bereits eingangs erwähnt, $\text{Abb}(X, K) = \text{Abb}[X, K] = K^n$, denn für $X = \{x_1, \ldots, x_n\}$ kann $f \in \text{Abb}(X, K)$ mit dem n-tupel $(f(x_1), \ldots, f(x_n)) \in K^n$ identifiziert werden.

Bei unendlichem X liegt die Situation jedoch völlig anders. Definieren wir ein verallgemeinertes Kroneckersymbol durch

$$\delta_a(x) := \begin{cases} 1 & \text{falls } x = a \\ 0 & \text{sonst,} \end{cases}$$

so sehen wir, daß die Menge $\{\delta_a \mid a \in X\}$ in $\text{Abb}(X, K)$ linear unabhängig ist. Aus $\sum_{i=1}^n \alpha_i \delta_{a_i} = 0$ erhalten wir nämlich $\sum_{i=1}^n \alpha_i \delta_{a_i}(a_i) = 0$ und damit $\alpha_i = 0$ für $i = 1, \ldots, n$. Andererseits ist nicht zu sehen, daß die Funktionen δ_a für $a \in X$ auch ein Erzeugendensystem bilden. Betrachten wir jedoch den Raum $\text{Abb}[X, K]$, so ist die Menge $\{\delta_a \mid a \in X\}$ auch ein Erzeugendensystem. Ist nämlich $f \in \text{Abb}[X, K]$, so finden wir eine endliche Teilmenge $\{a_1, \ldots, a_n\} \subseteq X$, so daß $f(x) = 0$ für alle $x \in X \setminus \{a_1, \ldots, a_n\}$ gilt. Dann erhalten wir offenbar $f(a_i) = f(a_i) \cdot \delta_{a_i}(a_i)$ und somit $f(x) = \sum_{i=1}^n f(a_i)\delta_{a_i}(x)$ für beliebiges $x \in X$, d.h. es ist $f = \sum_{i=1}^n f(a_i)\delta_{a_i}$. Somit erhalten wir

1.3.2 Satz *Die Menge der Funktionen* $\{\delta_a \mid a \in X\}$ *bilden eine Basis des Raumes* $\text{Abb}[X, K]$. *Damit ist* $\dim(\text{Abb}[X, K]) = |X|$. *Ist* X *eine endliche Menge mit* $|X| = n$, *so ist* $\text{Abb}(X, K) = \text{Abb}[X, K] = K^n$.

1.3.3 Polynome

Wählen wir im vorherigen Beispiel als X die Menge $\mathbb{N}$ der natürlichen Zahlen, so erhalten wir den Vektorraum der Polynome mit Koeffizienten in K. Wir definieren

$$K[x] := \text{Abb}[\mathbb{N}, K]. \tag{1.15}$$

Wie wir im Abschnitt 1.3.2 gesehen haben, ist die Menge der Kroneckersymbole $\{\delta_n \mid n \in \mathbb{N}\}$ eine Basis des $K[x]$. Zu jedem von der Nullabbildung verschiedenen $f \in K[x]$ gibt es dann ein kleinstes $n \in \mathbb{N}$ derart, daß $f(m) = 0$ für alle $m > n$ gilt. Wir nennen n den *Grad* von f und notieren ihn mit $G(f)$. Als Konvention setzen wir $G(0) := -\infty$, wobei 0 die Nullabbildung ist. Jedes Element $f \in K[x]$ ist daher durch das Tupel $(f(0), \ldots, f(n))$ mit $n := G(f)$ vollständig charakterisiert. Damit bietet sich die folgende Schreibweise an. Wir schreiben

$$f = f(0) + f(1) \cdot x + f(2) \cdot x^2 + \ldots + f(n) \cdot x^n,$$

wobei x einfach als ein Symbol – *eine Unbestimmte* – zu lesen ist. So erklärt sich auch die Notation $K[x]$. Diese Schreibweise hat den Vorteil, daß wir unsere üblichen Rechenregeln "formal" anwenden können. So ist z.B. für $f = f(0) + f(1) \cdot x + f(2) \cdot x^2 + \ldots + f(n) \cdot x^n$ und $g = g(0) + g(1) \cdot x + g(2) \cdot x^2 + \ldots + g(m) \cdot x^m$ die Summe $f + g = (f(0) + g(0)) + (f(1) + g(1)) \cdot x + \ldots + (f(k) + g(k)) \cdot x^k$ mit $k := \max\{m, n\}$ und $\alpha \cdot f = \alpha f(0) + \alpha f(1) \cdot x + \ldots + \alpha f(n) \cdot x^n$. Die eigentlichen

Vorteile dieser Schreibweise ergeben sich dann, wenn man eine Multiplikation auf $K[x]$ definieren und somit $K[x]$ zu einem Ring machen will. Hier definiert man für $f = \sum_{i=1}^{m} f(i)\delta_i$ und $g = \sum_{i=1}^{n} g(i)\delta_i$

$$f \cdot g = \sum_{k=1}^{m+n} (\sum_{i+j=k} (f(i) \cdot g(j))) \cdot \delta_k$$

und sieht sofort, daß man dieses Produkt erhält, wenn man $f(0) + f(1) \cdot x + \ldots + f(m) \cdot x^m$ und $g(0)+g(1)\cdot x+\ldots+g(n)\cdot x^n$ formal ausmultipliziert und dann nach Potenzen von x ordnet. Der Polynomring $K[x]$ spielt in der Algebra eine wichtige Rolle und wird uns auch in dieser Vorlesung noch mehrfach begegnen.

1.3.4 Produkte und Summen von Vektorräumen

Ist $\{V_\xi |\ \xi \in I\}$ eine Familie von Vektorräumen über einem Grundkörper K, so definieren wir

$$\prod_{\xi \in I} V_\xi := \{f |\ f : I \longrightarrow \bigcup_{\xi \in I} V_\xi \wedge (\forall \xi \in I)[f(\xi) \in V_\xi]\} \tag{1.16}$$

und

$$\bigoplus_{\xi \in I} V_\xi := \{f |\ f \in \prod_{\xi \in I} V_\xi \wedge f(\xi) = 0 \text{ für fast alle } \xi \in I\}. \tag{1.17}$$

Man nennt $\prod_{\xi \in I} V_\xi$ das Produkt der Vektorräume V_ξ und $\bigoplus_{\xi \in I} V_\xi$ das Koprodukt oder die direkte Summe der Vektorräume V_ξ. Warum man von einer Summe spricht, wird uns in Abschnitt 2.4 klarwerden.
Die Addition und skalare Multiplikation in beiden Räumen wird wieder punktweise definiert, d.h. durch

$$(f + g)(\xi) := f(\xi) + g(\xi)$$

und

$$(\alpha \cdot f)(\xi) := \alpha f(\xi).$$

Da $f(\xi)$ und $g(\xi)$ bei im Vektorraum V_ξ liegen, ist dies wohldefiniert.
Ist die Indexmenge endlich, z.B. $I = \{i_1, \ldots, i_n\}$, so schreiben wir $V_{i_1} \times \cdots \times V_{i_n}$ anstatt $\prod_{i \in I} V_i$ und $V_{i_1} \oplus \cdots \oplus V_{i_n}$ anstatt $\bigoplus_{i \in I} V_i$. (Hierbei ist aber zu bemerken, daß endliche Produkte und endliche Summen sich nicht unterscheiden.)
Spezialfälle erhalten wir, wenn wir $I = \{1, \ldots, n\}$ und $V_i = V$ für $i \in I$ wählen. Dann erhalten wir den Vektorraum V^n der n-tupel von Vektoren aus V. Untersu-

chungen über Basen und Dimension von $\prod_{\xi \in I} V_\xi$ und $\bigoplus_{\xi \in I} V_\xi$ wollen wir zurückstellen, bis uns mehr Hilfsmittel zur Verfügung stehen.

2. Lineare Abbildungen

2.1 Homomorphismen (am Beispiel von Gruppen)

Wir haben bereits bemerkt, daß Abbildungen zwischen Mengen eine wichtige Rolle in mathematischen Untersuchungen spielen. Tragen die Mengen darüber hinaus noch eine Struktur (handelt es sich also beispielsweise um Gruppen, Ringe, Körper, Vektorräume, ...), so sind die Abbildungen von besonderem Interesse, die diese Struktur respektieren. Solche Abbildungen nennt man *Homomorphismen*. Wir wollen den Begriff des Homomorphismus und einige seiner wichtigsten Eigenschaften am Beispiel der Struktur von Gruppen studieren.

2.1.1 Definition Es seien $(G, \circ_G)$ und $(H, \circ_H)$ Gruppen. Eine Funktion $f : G \longrightarrow H$ heißt ein *Homomorphismus* der Gruppen, wenn gilt:

$$(\forall a \in G)(\forall b \in G)[f(a \circ_G b) = f(a) \circ_H f(b)].$$

Für einen Homomorphismus definieren wir

$$\text{Kern } f = \{a \in G \mid f(a) = 1_H\},$$

wobei 1_H das neutrale Element der Gruppe $(H, \circ_H)$ bedeutet. Wir nennen Kern f den Kern des Homomorphismus f.

2.1.2 Folgerungen *Sei $f : G \longrightarrow H$ ein Homomorphismus der Gruppen. Dann gelten:*

(i) $f(1_G) = 1_H$.

(ii) $f(a^{-1}) = (f(a))^{-1}$.

(iii) Kern f *ist eine Untergruppe von* G.

(iv) Im f *ist eine Untergruppe von* H.

(v) f *ist injektiv* $\Leftrightarrow$ Kern $f = \{1_G\}$.

Beweis: (i): Es ist $f(1_G) = f(1_G \cdot 1_G) = f(1_G) \cdot f(1_G)$. Damit ist $1_H = f(1_G)^{-1} \cdot f(1_G) = f(1_G)^{-1} \cdot f(1_G) \cdot f(1_G) = f(1_G)$.
(ii): $f(a^{-1}) \cdot f(a) = f(a^{-1} \cdot a) = f(1_G) = 1_H$. Also ist $f(a)^{-1} = f(a^{-1})$.
(iii): Für $a, b \in$ Kern f gilt $f(ab^{-1}) = f(a)f(b^{-1}) = f(a)f(b)^{-1} = 1_H \cdot {1_H}^{-1} = 1_H \cdot 1_H = 1_H$. Also ist $ab^{-1} \in$ Kern f, der damit nach Satz 1.1.10 eine Untergruppe von G ist.
(iv): Sind $a, b \in$ Im f, so gibt es $x, y \in G$ mit $a = f(x)$ und $b = f(y)$. Damit erhalten wir $ab^{-1} = f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1})$. Damit ist xy^{-1} ein Urbild von ab^{-1}. Damit ist $ab^{-1} \in$ Im f, und Im f ist nach Satz 1.1.10 eine Untergruppe von H.
(v): $\Rightarrow$: Ist f injektiv und $a \in$ Kern f, so folgt wegen $f(a) = 1_H = f(1_G)$ sofort $a = 1_G$. D.h. Kern $f \subseteq \{1_G\}$, aber $\{1_G\} \subseteq$ Kern f folgt aus (iii), denn jede Untergruppe muß das neutrale Element enthalten.
$\Leftarrow$: Ist Kern $f = \{1_G\}$ und $f(a) = f(b)$, so erhalten wir $1_H = f(a)f(b)^{-1} = f(ab^{-1})$ und damit $ab^{-1} = 1_G$, woraus sofort $a = b$ durch Multiplikation mit b von rechts folgt. $\square$

2.1.3 Lemma *Sind $f\colon G_1 \longrightarrow G_2$ und $g\colon G_2 \longrightarrow G_3$ Homomorphismen, so ist auch $g \circ f\colon G_1 \longrightarrow G_3$ ein Homomorphismus.*

Beweis: Es ist $(g \circ f)(ab) = g(f(ab)) = g(f(a)f(b)) = g(f(a))g(f(b)) = (g \circ f)(a)(g \circ f)(b)$. $\square$

2.1.4 Lemma *Ist $f\colon G \longrightarrow H$ ein bijektiver Homomorphismus, so ist $f^{-1}\colon H \longrightarrow G$ ebenfalls ein bijektiver Homomorphismus.*

Beweis: Seien $a, b \in H$. Da f surjektiv ist finden wir $x, y \in G$ mit $f(x) = a$ und $f(y) = b$. Damit folgt $f^{-1}(ab) = f^{-1}(f(x)f(y)) = f^{-1}(f(xy)) = xy = f^{-1}(a)f^{-1}(b)$. $\square$

Bezeichnungen

- *Ein surjektiver Homomorphismus heißt ein* Epimorphismus.
- *Ein injektiver Homomorphismus heißt ein* Monomorphismus.
- *Ein bijektiver Homomorphismus heißt ein* Isomorphismus.
- *Ein Homomorphismus $f : G \to G$ heißt ein* Endomorphismus.
- *Ein bijektiver Endomorphismus heißt ein* Automorphismus.

2.1.5 Definition Zwei Gruppen heißen *isomorph*, wenn es einen Isomorphismus zwischen den Gruppen gibt.
Die Isomorphie der Gruppen G und H wird mit $G \cong H$ notiert.

2.1.6 Lemma *Die Isomorphie von Gruppen ist eine Äquivalenzrelation auf der Menge aller Gruppen.*

Beweis: Es gilt $G \cong G$ vermöge der Identität, die offensichtlich ein Isomorphismus der Gruppen ist. Damit ist die Relation $\cong$ reflexiv. Gilt $G \cong H$ vermöge des Isomorphismus $f\colon G \longrightarrow H$, so ist nach Lemma 2.1.4 $f^{-1}\colon H \longrightarrow G$ ebenfalls ein Isomorphismus der Gruppen, und wir haben $H \cong G$. Damit ist $\cong$ als symmetrisch nachgewiesen. Gilt schließlich $G \cong H$ und $H \cong L$ vermöge der Isomorphismen $f\colon G \longrightarrow H$ und $g\colon H \longrightarrow L$, so ist $(g \circ f)\colon G \longrightarrow L$ nach Lemma 2.1.3 ein Homomorphismus der Gruppen und nach Lemma 0.5.7 auch bijektiv. Damit ist $G \cong L$ vermöge $(g \circ f)$, und wir haben $\cong$ auch als transitiv nachgewiesen. $\square$

2.2 Kanonische Faktorisierung von Homomorphismen und der Homomorphiesatz

2.2.1 Lemma *Ist $f\colon G \longrightarrow H$ ein Homomorphismus der Gruppen $(G, \circ)$ und $(H, \circ)$, so ist die Relation $a \sim_f b \;\; :\Leftrightarrow \;\; f(a) = f(b)$ eine Äquivalenzrelation auf G.*

Beweis: Wegen $f(a) = f(a)$ ist $\sim_f$ reflexiv, wegen $f(a) = f(b) \;\Rightarrow\; f(b) = f(a)$ symmetrisch und wegen $f(a) = f(b) = f(c) \;\Rightarrow\; f(a) = f(c)$ auch transitiv. $\square$

2.2.2 Lemma *Die Äquivalenzklassen $[a]_f := \{b \in G \mid a \sim_f b\}$ bilden eine Partition von G.*

Beweis: Wir nennen eine Familie $\{M_\iota \mid \iota \in I\}$ von Teilmengen von G eine *Partition* von G, wenn $G = \bigcup_{\iota \in I} M_\iota$ gilt und die Vereinigung disjunkt ist, d.h. wenn $(\forall \iota \in I)(\forall \kappa \in I)[M_\iota \neq M_\kappa \;\Rightarrow\; M_\iota \cap M_\kappa = \emptyset]$ gilt. Wir haben demnach zu zeigen, daß $G = \bigcup_{x \in G} [x]_f$ ist und aus $[a]_f \neq [b]_f$ schon $[a]_f \cap [b]_f = \emptyset$ folgt. Nach Definition der Äquivalenzklassen ist $\bigcup_{x \in G} [x]_f \subseteq G$. Ist umgekehrt $a \in G$, so folgt $a \in [a]_f \subseteq \bigcup_{x \in G} [x]_f$. Damit haben wir auch $G \subseteq \bigcup_{x \in G} [x]_f$. Nehmen wir $[a]_f \cap [b]_f \neq \emptyset$ an, so gibt es ein $c \in [a]_f \cap [b]_f$, d.h. es gilt $f(a) = f(c) = f(b)$. Also folgt $a \sim_f b$ und damit $[a]_f = [b]_f$. Die Kontraposition dieser Aussage ist aber gerade $[a]_f \neq [b]_f \;\Rightarrow\; [a]_f \cap [b]_f = \emptyset$, und wir sind fertig. $\square$

2.2.3 Lemma *Es gilt $[a]_f = \{a \cdot x \mid x \in \text{Kern } f\} =: a \cdot \text{Kern } f$.*

Beweis: $\subseteq$: Ist $b \in [a]_f$, so gilt $f(b) = f(a)$. Die Gleichung $b = a \cdot x$ hat in G eine Lösung und wir erhalten $f(b) = f(ax) = f(a)f(x) = f(a)$. Es folgt $f(x) = 1$ und damit $x \in \text{Kern } f$. Also ist $b = ax \in a \cdot \text{Kern } f$.
$\supseteq$: Ist $b \in a \cdot \text{Kern } f$, so ist $b = ax$ für ein $x \in \text{Kern } f$, und es folgt $f(b) = f(ax) = f(a)f(x) = f(a) \cdot 1 = f(a)$. Also gilt $b \sim_f a$ und damit $b \in [a]_f$. □

2.2.4 Lemma *Die Menge* $G/\text{Kern } f := \{a \cdot \text{Kern } f \,|\, a \in G\}$ *wird mit der Verknüpfung* $(a \cdot \text{Kern } f) \circ (b \cdot \text{Kern } f) := (ab) \cdot \text{Kern } f$ *eine Gruppe. Diese Gruppe heißt die* Faktor- *oder* Quotientengruppe *von* G *modulo dem Kern von* f.

Beweis: Zunächst haben wir zu zeigen, daß die im Lemma definierte Verknüpfung tatsächlich eine Abbildung $\circ: G/\text{Kern } f \times G/\text{Kern } f \longrightarrow G/\text{Kern } f$ ist. Man sagt, daß wir die *Wohldefiniertheit* der Verknüpfung zeigen müssen. Nachzuprüfen ist nur die Rechtseindeutigkeit. Seien also $a \cdot \text{Kern } f = a' \cdot \text{Kern } f$ und $b \cdot \text{Kern } f = b' \cdot \text{Kern } f$. Wir haben $ab \cdot \text{Kern } f = a'b' \cdot \text{Kern } f$ nachzuprüfen. Mit Lemma 2.2.3 erhalten wir aber $x \in ab \cdot \text{Kern } f \Leftrightarrow f(x) = f(ab) = f(a)f(b) = f(a')f(b') = f(a'b') \Leftrightarrow x \in a'b' \cdot \text{Kern } f$. Damit ist $ab \cdot \text{Kern } f \subseteq a'b' \cdot \text{Kern } f$ und die Gegenrichtung erhalten wir aus Symmetriegründen in völlig analoger Weise. Nun überzeugen wir uns davon, daß $(1 \cdot \text{Kern } f) \circ (a \cdot \text{Kern } f) = 1a \cdot \text{Kern } f = a \cdot \text{Kern } f$ ist und haben damit ein linksneutrales Element. Wegen $(a^{-1} \cdot \text{Kern } f) \circ (a \cdot \text{Kern } f) = a^{-1}a \cdot \text{Kern } f = 1 \cdot \text{Kern } f$ besitzt jedes Element $a \cdot \text{Kern } f$ das Element $a^{-1} \cdot \text{Kern } f$ als inverses Element. Damit ist $G/\text{Kern } f$ eine Gruppe. □

2.2.5 Lemma *Die Abbildung* $\pi_f: G \longrightarrow G/\text{Kern } f$, *die definiert ist durch* $\pi_f(a) := a \cdot \text{Kern } f$, *ist ein Epimorphismus der Gruppen mit* $\text{Kern } \pi_f = \text{Kern } f$.
Man nennt π_f *den* kanonischen Epimorphismus *von* G *auf* $G/\text{Kern } f$.

Beweis: Wegen $\pi_f(ab) := ab \cdot \text{Kern } f = (a \cdot \text{Kern } f) \circ (b \cdot \text{Kern } f) = \pi_f(a) \cdot \pi_f(b)$ liegt ein Homomorphismus vor. Ist $x \in G/\text{Kern } f$, so gibt es ein $a \in G$ mit $x = a \cdot \text{Kern } f = \pi_f(a)$. Damit ist a ein Urbild von x unter π_f. Da dies für beliebige $x \in G/\text{Kern } f$ gilt, ist π_f surjektiv.
Ist $a \in \text{Kern } \pi_f$, so ist $\pi_f(a) = a \cdot \text{Kern } f = 1 \cdot \text{Kern } f$. Also ist $a \sim_f 1$, d.h. $f(a) = f(1) = 1_H$. Damit ist $\text{Kern } \pi_f \subseteq \text{Kern } f$. Ist umgekehrt $a \in \text{Kern } f$, so ist $f(a) = f(1)$ und damit $\pi_f(a) = a \cdot \text{Kern } f = 1 \cdot \text{Kern } f$, d.h. $a \in \text{Kern } \pi_f$. Somit ist $\text{Kern } f = \text{Kern } \pi_f$. □

2.2.6 Lemma *Die Abbildung* $\iota_f: G/\text{Kern } f \longrightarrow H$, *die definiert ist durch* $\iota_f(a \cdot \text{Kern } f) := f(a)$, *ist ein Monomorphismus der Gruppen.*
Man nennt ι_f *die* kanonische Einbettung *von* $G/\text{Kern } f$ *in* H.

Beweis: Zunächst wollen wir betonen, daß die Abbildung ι_f wegen $a \cdot \mathrm{Kern}\ f = a' \cdot \mathrm{Kern}\ f \;\Leftrightarrow\; f(a) = f(a') \;\Leftrightarrow\; \iota_f(a \cdot \mathrm{Kern}\ f) = \iota_f(a' \cdot \mathrm{Kern}\ f)$ wohldefiniert (d.h. rechtseindeutig) ist. Aber da wir ja überall Äquivalenzen stehen haben, ist sie ebenso linkseindeutig, d.h. aus $\iota_f(x) = \iota_f(y)$ folgt $x = y$. Damit ist ι_f injektiv. Wegen $\iota_f((a \cdot \mathrm{Kern}\ f) \circ (b \cdot \mathrm{Kern}\ f)) = \iota_f(ab \cdot \mathrm{Kern}\ f) = f(ab) = f(a)f(b) = \iota_f(a \cdot \mathrm{Kern}\ f) \cdot \iota_f(b \cdot \mathrm{Kern}\ f)$ ist ι_f auch ein Homomorphismus der Gruppen. $\square$

2.2.7 Homomorphiesatz für Gruppen *Ist $f\colon G \longrightarrow H$ ein Homomorphismus der Gruppen, so ist $G/\mathrm{Kern}\ f$ eine Gruppe, und es gibt einen surjektiven Homomorphismus $\pi_f\colon G \longrightarrow G/\mathrm{Kern}\ f$ und einen injektiven Homomorphismus $\iota_f\colon G/\mathrm{Kern}\ f \longrightarrow H$ derart, daß das folgende Diagramm kommutiert.*

$$\begin{tikzcd} G \arrow[r, "f"] \arrow[d, "\pi_f"'] & H \\ G/\mathrm{Kern}\ f \arrow[ur, "\iota_f"] & \end{tikzcd}$$

D.h. es gilt $f = \iota_f \circ \pi_f$.
Insbesondere ist $G/\mathrm{Kern}\ f \cong f[G] = \mathrm{Im}\ f$.

Beweis: Nach den Lemmata 2.2.5 und 2.2.6 erhalten wir sowohl den kanonischen Epimorphismus $\pi_f\colon G \longrightarrow G/\mathrm{Kern}\ f$ als auch die kanonische Einbettung $\iota_f\colon G/\mathrm{Kern}\ f \longrightarrow H$. Für $a \in G$ erhalten wir $(\iota_f \circ \pi_f)(a) = \iota_f(\pi_f(a)) = \iota_f(a \cdot \mathrm{Kern}\ f) = f(a)$, und wir haben die Kommutativität des Diagramms gezeigt. Ist $x \in \mathrm{Im}\ f$, so ist $x = f(a)$ für ein $a \in G$, und es folgt $x = f(a) = \iota_f(a \cdot \mathrm{Kern}\ f)$. Damit ist der Monomorphismus $\iota_f\colon G/\mathrm{Kern}\ f \longrightarrow \mathrm{Im}\ f$ surjektiv und folglich ein Isomorphismus. $\square$

Wir wollen die Situation, die zum Homomorphiesatz geführt hat, nochmals von einem anderen Standpunkt aus betrachten. Als erstes haben wir eingesehen, daß $\mathrm{Kern}\ f$ eine Untergruppe der Gruppe G ist. Dann haben wir eine Äquivalenzrelation $a \sim_f b \;:\Leftrightarrow\; f(a) = f(b)$ eingeführt. Nun gilt $f(a) = f(b) \;\Leftrightarrow\; 1_H = f(b)^{-1}f(a) = f(b^{-1}a) \;\Leftrightarrow\; b^{-1}a \in \mathrm{Kern}\ f$. Die erste Frage, die wir uns hier stellen können, ist, ob für jede Untergruppe U von G durch die Definition $a \sim b \;:\Leftrightarrow\; b^{-1}a \in U$ eine Äquivalenzrelation definiert wird. Dies wollen wir im folgenden Lemma untersuchen.

2.2.8 Lemma *Ist U eine Untergruppe einer Gruppe G, so definiert die Relation*

$$a \sim b \bmod U \quad :\Leftrightarrow \quad b^{-1}a \in U$$

eine Äquivalenzrelation auf G. Für deren Äquivalenzklassen

$$[a]_U := \{b \mid a \sim b \mod U\}$$

gilt

$$[a]_U = a \cdot U := \{a \cdot x \mid x \in U\}.$$

Man nennt $a \cdot U$ die linke Nebenklasse von a bzgl. U*. Die Nebenklassen bilden eine Partition von G, d.h. es gilt $G = \bigcup_{a \in G} a \cdot U$ und $a \cdot U \cap b \cdot U \neq \emptyset \;\Rightarrow\; a \cdot U = b \cdot U$*

Beweis: Wegen $a^{-1}a = 1 \in U$ ist $\sim \mod U$ reflexiv. Aus $a \sim b \mod U$ folgt $b^{-1}a \in U$ und daraus $(b^{-1}a)^{-1} = a^{-1}b \in U$, d.h. $b \sim a \mod U$. Damit ist $\sim \mod U$ symmetrisch. Die Transitivität erhalten wir wegen $a \sim b \mod U \wedge b \sim c \mod U \;\Rightarrow\; b^{-1}a \in U \wedge c^{-1}b \in U \;\Rightarrow\; (c^{-1}b)(b^{-1}a) = c^{-1}a \in U \;\Rightarrow\; a \sim c \mod U$.
Ist $b \in [a]_U$, so folgt $a \sim b \mod U$ und daher $b^{-1}a \in U$. Damit ist aber $b = a(a^{-1}b)$, und es ist $a^{-1}b \in U$. Damit ist $b \in a \cdot U$. Ist umgekehrt $b \in a \cdot U$, so gibt es ein $u \in U$ mit $b = au$. Dann folgt aber $b^{-1}a = u^{-1}a^{-1}a = u \in U$ und damit $b \sim a \mod U$, d.h. $b \in [a]_U$. Damit haben wir $[a]_U = a \cdot U$ gezeigt. Die Tatsache, daß die Nebenklassen eine Partition von G bilden, wollen wir etwas allgemeiner nachweisen. Wir zeigen:

Hilfssatz *Ist M eine Menge und $\sim$ eine Äquivalenzrelation auf M, so bilden die Äquivalenzklassen $[a] := \{b \in M \mid a \sim b\}$ eine Partition von M.*

Es ist klar, daß dann die letzte noch offenen Behauptung des Lemma 2.2.8 dann aus dem Hilfssatz folgt. □

Beweis des Hilfssatzes: Wir haben zu zeigen, daß

$$M = \bigcup_{a \in M} [a] \tag{i}$$

und

$$[a] \neq [b] \;\Rightarrow\; [a] \cap [b] = \emptyset \tag{ii}$$

gilt. Wegen $[a] \subseteq M$ für alle $a \in M$ ist die Inklusion $\supseteq$ in (i) klar. Umgekehrt gilt aber wegen der Reflexivität von $\sim$ immer $a \in [a]$, womit auch $M \subseteq \bigcup_{a \in M}[a]$ folgt. Damit ist (i) gezeigt. Ist nun $[a] \cap [b] \neq \emptyset$, so gibt es ein $c \in [a] \cap [b]$, und damit folgt $a \sim c \sim b$, was wegen der Transitivität von $\sim$ auch

$$a \sim b \tag{iii}$$

nach sich zieht. Wir behaupten nun

$$a \sim b \;\;\Rightarrow\;\; [a] = [b]. \tag{iv}$$

Aus (iii) und (iv) folgt dann die Behauptung. Um (iv) zu zeigen, nehmen wir $x \in [a]$, d.h. $a \sim x$ an. Dann haben wir $a \sim b \wedge a \sim x$, woraus $b \sim a \wedge a \sim x$ mit der Symmetrie und daraus $b \sim x$ mit der Transitivität von $\sim$ folgt. Also ist $[a] \subseteq [b]$. Völlig symmetrisch erhalten wir auch $[b] \subseteq [a]$, und alles ist gezeigt. □

Die Tatsache, daß die Nebenklassen bezüglich einer Untergruppe eine Partition der Gruppe bilden, gibt die Motivation zur folgenden Definition.

2.2.9 Definition Die *Ordnung* einer endlichen Gruppe ist die Anzahl ihrer Elemente.
Der *Index* einer Untergruppe U in einer endlichen Gruppe G ist die Anzahl der Nebenklassen bezüglich U.
Der Index wird notiert durch $[G : U]$. D.h. es gilt

$$[G : U] := |\{a \cdot U \mid a \in G\}|.$$

Betrachten wir die triviale Untergruppe $\{1\} \subseteq G$, die nur aus dem neutralen Element besteht, so erhalten wir die Nebenklassen $a \cdot \{1\} := \{a\}$. Damit ist $|\{a \cdot \{1\} \mid a \in G\}| = |G|$, und es folgt

$$[G : 1] = |G|. \tag{2.1}$$

Allgemein erhalten wir

2.2.10 Satz *Ist G eine endliche Gruppe und U eine Untergruppe von G, so ist $[G : 1] = [G : U] \cdot [U : 1]$. D.h. die Ordnung einer Untergruppe ist stets ein Teiler der Ordnung der Gruppe.*

Beweis: Die Abbildung $l_a : U \longrightarrow a \cdot U$ mit $l_a(u) := au$ ist offensichtlich surjektiv und wegen $au = bu \;\Rightarrow\; a = auu^{-1} = buu^{-1} = b$ auch injektiv. Damit folgt $|U| = |a \cdot U|$. Nun ist wegen Lemma 2.2.8 $[G : 1] = |G| = |\bigcup_{a \in G} a \cdot U| = [G : U]|a \cdot U| = [G : U]|U| = [G : U][U : 1]$. □

Wir haben nun gesehen, daß jede Untergruppe U einer Gruppe G eine Äquivalenzrelation $\sim$ mod U auf G und damit auch eine Partition von G in $[G : U]$ viele Nebenklassen erzeugt. Die nächste Frage, die sich daran anschließt, ist, ob sich auf der Menge der Nebenklassen wieder eine Gruppenstruktur definieren läßt. Ist $U = \text{Kern } f$ für einen Gruppenhomomorphismus f, so wissen wir bereits, daß dies möglich ist. Entscheidend für die Wohldefiniertheit der damals eingeführten Verknüpfung auf den Nebenklassen war, daß die durch den Homomorphismus f induzierte Äquivalenzrelation $\sim_f$ mit der Gruppenverknüpfung auf G verträglich war, d.h. daß aus $a \sim_f a'$ und $b \sim_f b'$ auch $ab \sim_f a'b'$ folgte. Das ist im allgemeinen für die durch eine Untergruppe U erzeugte Äquivalenzrelation $\sim \mod U$ nicht richtig. Um $ab \sim a'b' \mod U$ aus $a \sim a' \mod U$ und $b \sim b' \mod U$ zu folgern, müßten wir $(a'b')^{-1} \cdot ab \in U$, d.h. $b'^{-1}a'^{-1}ab \in U$, aus $a'^{-1}a \in U$ und $b'^{-1}b \in U$ folgern. Hinreichend dafür ist es, $b'^{-1}ub \in U$ für jedes $u \in U$ und b,b' mit $b'^{-1}b \in U$ zu haben. Nun gilt $(\forall u \in U)[b'^{-1}ub \in U] \;\Leftrightarrow\; U \cdot b = b' \cdot U$ und wegen $b \sim b' \mod U$ ist, wie (iv) im Beweis des Hilfsatzes gezeigt hat, $b' \cdot U = b \cdot U$. Hinreichend für die Verträglichkeit der Äquivalenzrelation $\sim \mod U$ mit der Gruppenverknüpfung ist daher, daß $b \cdot U = U \cdot b$ für alle $b \in G$ gilt. Dies ist der Hintergrund für die folgende Definition.

2.2.11 Definition Eine Untergruppe U einer Gruppe G heißt ein *Normalteiler* von G, falls $b \cdot U = U \cdot b$ für alle $b \in G$ gilt.

Warnung Die Aussage

$$b \cdot U = \{b \cdot x \mid x \in U\} = U \cdot b = \{x \cdot b \mid x \in U\}$$

bedeutet **nicht**, daß $bu = ub$ für $u \in U$ gilt, sondern

$$(\forall x \in U)(\exists v \in U)[b \cdot x = v \cdot b] \text{ und } (\forall y \in U)(\exists v \in U)[y \cdot b = b \cdot v].$$

Bemerkung *Ist G eine abelsche, d.h. kommutative, Gruppe, so ist jede Untergruppe von G bereits ein Normalteiler.*

2.2.12 Satz *Ist U ein Normalteiler von G, so ist $G/U := \{a \cdot U \mid a \in G\}$ mit der Verknüpfung $(a \cdot U) \circ (b \cdot U) := ab \cdot U$ eine Gruppe. Man nennt G/U die* Faktor- *oder* Quotientengruppe *von G nach dem Normalteiler U.*
Die Abbildung $\pi_U : G \longrightarrow G/U$, die definiert ist durch $\pi_U(a) := a \cdot U$, ist ein Epimorphismus der Gruppen mit Kern $(\pi) = U$. *Sie heißt der kanonische Epimorphismus von G auf G/U. Für endliche Gruppen G gilt $|G| = |G/U| \cdot |U|$.*

Beweis: Wir haben in unseren Vorbetrachtungen bereits festgestellt, daß die Eigenschaft, Normalteiler zu sein, hinreicht, um $a \sim a' \mod U \wedge b \sim b' \mod U \;\Rightarrow\; ab \sim a'b' \mod U$ zu erhalten, d.h. aus $a \cdot U = a' \cdot U$ und $b \cdot U = b' \cdot U$ folgt $ab \cdot U = a'b' \cdot U$. Das bedeutet aber, daß die Verknüpfung $(a \cdot U) \circ (b \cdot U)$ wohldefiniert ist. Offensichtlich ist $U = 1 \cdot U$ ein linksneutrales Element und $a^{-1} \cdot U$ invers zu $a \cdot U$. Damit ist G/U mit dieser Verknüpfung eine Gruppe. Wegen $\pi(ab) = (ab) \cdot U = (a \cdot U) \circ (b \cdot U) = \pi(a) \circ \pi(b)$ ist $\pi : G \longrightarrow G/U$ ein Homomorphismus. Zu $a \cdot U \in G/U$ ist $a \in G$ ein Urbild unter π. Damit ist π surjektiv und aus $\pi(a) = 1 \cdot U = U$ folgt $a = 1 \cdot a \in U$, d.h. Kern $\pi \subseteq U$. Umgekehrt ist für $u \in U$ stets $\pi(u) = u \cdot U = U$. Damit ist Kern $\pi = U$. □

2.2.13 Satz *Eine Untergruppe $U \subseteq G$ ist genau dann ein Normalteiler, wenn U der Kern eines Gruppenhomomorphismus ist.*

Beweis: Ist U ein Normalteiler, so folgt nach Satz 2.2.12, daß U Kern des kanonischen Epimorphismus ist. Ist $f\colon G \longrightarrow H$ ein Homomorphismus, so gilt $x \in a \cdot \text{Kern } f$ genau dann, wenn $x = au$ für ein $u \in \text{Kern } f$ ist. Die Gleichung $ya = x$ ist in G lösbar und es gilt $f(a) = f(x)f(u) = f(x) = f(ya) = f(y)f(a)$, woraus $f(y) = 1$ folgt. Also ist $y \in \text{Kern } f$ und damit $x \in \text{Kern } f \cdot a$. Damit haben wir $a \cdot \text{Kern } f \subseteq \text{Kern } f \cdot a$ und völlig analog ergibt sich auch $\text{Kern } f \cdot a \subseteq a \cdot \text{Kern } f$.
□

Wir wollen diesen Abschnitt mit der Bemerkung beschließen, daß die hier gemachten Untersuchungen keinesfalls auf Gruppen und deren Homomorphismen beschränkt sind, sondern für beliebige algebraische Strukturen gelten. In dieser Vorlesung werden wir jedoch vorwiegend Gruppen und Vektorräume betrachten.

2.3 Lineare Abbildungen

2.3.1 Definition Sind V und W K-Vektorräume, so nennen wir eine Abbildung $f\colon V \longrightarrow W$ *linear* oder einen *Vektorraumhomomorphismus*, wenn f ein Homomorphismus der additiven abelschen Gruppen ist und für alle $\alpha \in K$ und $a \in V$ auch $f(\alpha a) = \alpha f(a)$ gilt.

Berücksichtigen wir die Tatsache, daß wir nun additiv notierte abelsche Gruppen vorliegen haben, so erhalten wir die anschließenden Folgerungen in ähnlicher Weise wie die Folgerungen 2.1.2.

2.3.2 Folgerungen *Ist $f\colon V \longrightarrow W$ linear, so gelten:*

(i) $f(0_V) = 0_W$;

(ii) $f(-a) = -f(a)$;

(iii) $\text{Kern}(f) = \{a \in V \mid f(a) = 0_W\}$ *ist ein Unterraum von V;*

(iv) $\text{Im}(f) = \{f(a) \mid a \in V\}$ *ist ein Unterraum von W;*

(v) f *ist injektiv* $\Leftrightarrow$ $\text{Kern}(f) = \{0_V\} = \mathbf{0}$.

Beweis: Die Behauptungen (i), (ii) und (v) folgen nach 2.1.2 sofort aus der Tatsache, daß f ein Homomorphismus der abelschen Gruppen ist. Um (iii) zu zeigen, beachten wir zuerst, daß nach 2.1.2 Kern f eine Untergruppe der abelschen Gruppe von V ist. Ist $\alpha \in K$ und $a \in \text{Kern } f$, so folgt $f(\alpha a) = \alpha f(a) = \alpha \cdot 0_W = 0_W$, d.h. daß Kern f gegenüber skalarer Multiplikation abgeschlossen ist. Damit ist

Kern f ein Unterraum von V. Ähnlich erhalten wir (iv). Zunächst ist Im f eine Untergruppe der abelschen Gruppe von W. Für $\alpha \in K$ und $b \in$ Im f folgt aber $b = f(x)$ für ein $x \in V$, und wir erhalten $\alpha b = \alpha f(x) = f(\alpha x)$, woraus $\alpha b \in$ Im f folgt. □

2.3.3 Lemma *Sind $f: V \longrightarrow U$, $g: U \longrightarrow W$ linear, so ist $g \circ f: V \longrightarrow W$ ebenfalls linear.*

Beweis: Nach Lemma 2.1.3 ist $g \circ f$ ein Homomorphismus der Gruppen. Wegen $(g \circ f)(\alpha a) = g(f(\alpha a)) = g(\alpha f(a)) = \alpha g(f(a)) = \alpha(g \circ f)(a)$ ist $g \circ f$ dann aber auch linear. □

2.3.4 Lemma *Ist $f: V \longrightarrow W$ linear und bijektiv, so ist $f^{-1}: W \longrightarrow V$ linear.*

Beweis: Nach Lemma 2.1.4 ist $f^{-1}: W \longrightarrow V$ ein Homomorphismus der abelschen Gruppen. Durch Nachrechnen erhalten wir $f^{-1}(\alpha a) = f^{-1}(\alpha f(x)) = f^{-1}(f(\alpha x)) = \alpha x = \alpha(f^{-1}(f(x))) = \alpha f^{-1}(a)$, wobei x das Urbild von a unter f war. Damit ist f^{-1} linear. □

2.3.5 Lemma *Ist $f: V \longrightarrow W$ linear, so gelten:*

(i) $V = \langle a_1, \dots, a_n \rangle \;\Rightarrow\; f[V] = \mathrm{Im}\,(f) = \langle f(a_1), \dots, f(a_n) \rangle$;

(ii) $a_1, \dots, a_n$ *linear abhängig* $\Rightarrow$ $f(a_1), \dots, f(a_n)$ *linear abhängig;*

(iii) $f(a_1), \dots, f(a_n)$ *linear unabhängig* $\Rightarrow$ $a_1, \dots, a_n$ *linear unabhängig.*

Beweis: (i): Ist $b \in f[V]$, so gibt es $x \in V$ mit $b = f(x) = f(\sum_{i=1}^n \alpha_i a_i) = \sum_{i=1}^n \alpha_i f(a_i)$, d.h. $b \in \langle f(a_1), \dots, f(a_n) \rangle$.
Die Aussagen (ii) und (iii) sind wechselseitig Kontrapositionen. Es genügt daher, eine zu beweisen. Wir zeigen (iii). Aus $\sum_{i=1}^n \alpha_i a_i = 0$ folgt $0 = f(\sum_{i=1}^n \alpha_i a_i) = \sum_{i=1}^n \alpha_i f(a_i)$ und wegen der linearen Unabhängigkeit von $f(a_1), \dots, f(a_n)$ daraus $\alpha_1 = \ldots = \alpha_n = 0$. □

2.3.6 Satz *Ist $f: V \longrightarrow W$ linear, so ist* $\dim(\mathrm{Im}\,(f)) \leq \dim(V)$.

Beweis: Ist $\{b_1, \dots, b_n\}$ eine Basis von V, so ist $\{f(b_1), \dots, f(b_n)\}$ nach Lemma 2.3.5 (i) ein Erzeugendensystem von Im f. Damit ist $\dim(\text{Im } f) \leq n = \dim V$. □

2.3.7 Korollar *Gilt $V \cong W$, so ist* $\dim(V) = \dim(W)$.

Beweis: Ist $f\colon V \longrightarrow W$ ein Isomorphismus, so gilt nach Satz 2.3.6 $\dim V \geq \dim f[V] = \dim W$. Da nach Satz 2.3.4 auch $f^{-1}\colon W \longrightarrow V$ ein Isomorphismus ist, folgt gleichermaßen $\dim W \geq \dim V$. □

2.3.8 Lemma *Ist $f\colon V \longrightarrow W$ linear, so sind äquivalent:*

(i) f *ist injektiv.*

(ii) $a_1, \ldots, a_n$ *linear unabhängig* $\Leftrightarrow$ $f(a_1), \ldots, f(a_n)$ *linear unabhängig.*

Beweis: (i) $\Rightarrow$ (ii): Ist $f\colon V \longrightarrow W$ linear, so folgt nach Lemma 2.3.5 (iii) aus der linearen Unabhängigkeit von $f(a_1), \ldots, f(a_n)$ auch die lineare Unabhängigkeit von $a_1, \ldots, a_n$. Sind $a_1, \ldots, a_n$ linear unabhängig und gilt $\sum_{i=1}^n \alpha_i f(a_i) = 0$, so erhalten wir $f(\sum_{i=1}^n \alpha_i a_i) = 0$, d.h. $\sum_{i=1}^n \alpha_i a_i \in$ Kern f. Ist nun f injektiv, so folgt daraus $\sum_{i=1}^n \alpha_i a_i = 0$ und somit $\alpha_1 = \ldots = \alpha_n = 0$. Damit sind $f(a_1), \ldots, f(a_n)$ linear unabhängig.
(ii) $\Rightarrow$ (i): Sei $\{b_1, \ldots, b_m\}$ eine Basis des Unterraums Kern f. Für $a \in$ Kern f erhalten wir $0 = f(a) = f(\sum_{i=1}^n \alpha_i b_i) = \sum_{i=1}^n \alpha_i f(b_i)$, woraus wir $\alpha_1 = \ldots = \alpha_n = 0$ wegen der linearen Unabhängigkeit von $f(b_1), \ldots, f(b_n)$ erhalten. Damit ist $a = \sum_{i=1}^n \alpha_i b_i = 0$ und somit Kern $f = \mathbf{0}$. D.h. f ist injektiv. □

2.3.9 Satz *Sind V und W K-Vektorräume mit* $\dim(V) = \dim(W) < \infty$ *und* $f\colon V \longrightarrow W$ *eine lineare Abbildung, so sind äquivalent:*

(i) f *ist injektiv;*

(ii) f *ist surjektiv;*

(iii) f *ist bijektiv.*

Beweis: Es genügt natürlich (i) $\Leftrightarrow$ (ii) zu beweisen. Beginnen wir mit (i) $\Rightarrow$ (ii): Sei $f\colon V \longrightarrow W$ injektiv. Nach Satz 2.3.6 folgt $\dim(\text{Im } f) \leq \dim V$ und nach Lemma 2.3.8 $\dim(\text{Im } f) \geq \dim V$. Damit ist $\dim(\text{Im } f) = \dim V = \dim W$. Da Im f ein Unterraum von W ist, folgt nach Satz 1.2.27 (ii) Im $f = W$, und f ist surjektiv.
(ii) $\Rightarrow$ (i): Ist $\{a_1, \ldots, a_n\}$ eine Basis von V, so ist $\langle f(a_1), \ldots, f(a_n)\rangle =$ Im $f = W$, da f surjektiv ist. Damit sind aber $f(a_1), \ldots, f(a_n)$ linear unabhängig. Nach Lemma 2.3.8 folgt daraus die Injektivität von f. □

2.3.10 Satz *Ist $U \subseteq V$ ein Teilraum eines K-Vektorraumes V, so ist die Menge* $V/U := \{a + U \mid a \in V\}$ *mit den Verknüpfungen* $(a+U)+(b+U) := (a+b)+U$

und $\alpha \cdot (a+U) := \alpha \cdot a + U$ *ein Vektorraum. Man nennt* V/U *den* Quotientenraum *von* U *in* V.
Der kanonische Epimorphismus $\pi\colon V \longrightarrow V/U$*, der definiert ist durch* $\pi(a) := a + U$*, ist linear mit* $\operatorname{Kern}(\pi) = U$.

Beweis: Zunächst ist U als Teilraum eine Untergruppe der abelschen Gruppe von V. Damit ist U ein Normalteiler, und V/U wird nach Satz 2.2.12 mit der oben definierten Verknüpfung eine Gruppe. Es bleibt also zu zeigen, daß die skalare Multiplikation wohldefiniert ist und die Axiome in (V2) erfüllt. Wir beginnen mit der Wohldefiniertheit. Sei dazu $a + U = b + U$ und $x \in \alpha(a + U)$. Dann gibt es ein $u \in U$ mit $x = \alpha a + u$. Wegen $a + U = b + U$ ist $a = b + v$ für ein $v \in U$, und wir erhalten $x = \alpha a + u = \alpha(b + v) + u = \alpha b + (\alpha v + u) \in \alpha(b + U)$, da $\alpha v + u \in U$ ist. Damit ist $\alpha(a+U) \subseteq \alpha(b+U)$, und die umgekehrte Inklusion folgt analog. Die skalare Multiplikation ist also wohldefiniert. Nun ist $(\alpha+\beta)(a+U) = (\alpha + \beta) \cdot a + U = (\alpha \cdot a + \beta \cdot a) + U = (\alpha a + U) + (\beta a + U)$. Damit gilt (V21). Weiter erhalten wir $\alpha \cdot ((a+U) + (b+U)) = \alpha \cdot (a+b+U) = (\alpha(a+b)) + U = (\alpha a + \alpha b) + U = (\alpha a + U) + (\alpha b + U)$, womit (V22) erfüllt ist. Nun berechnen wir $\alpha(\beta(a + U)) = \alpha(\beta a + U) = \alpha(\beta a) + U = (\alpha\beta)a + U = (\alpha\beta)(a + U)$ und haben damit (V23) nachgewiesen. Schließlich erhalten wir auch (V24) wegen $1 \cdot (a + U) = (1 \cdot a) + U = a + U$.
Wir wissen bereits nach Satz 2.2.12, daß $\pi\colon V \longrightarrow V/U$ ein Epimorphismus der Gruppen mit Kern $\pi = U$ ist. Wegen $\pi(\alpha a) = (\alpha a) + U = \alpha(a + U) = \alpha\pi(a)$ ist π aber auch linear, und alles ist bewiesen. □

2.3.11 Homomorphiesatz für Vektorräume *Sind* V, W *beides* K*-Vektorräume und ist* $f\colon V \longrightarrow W$ *eine lineare Abbildung, so gibt es einen Epimorphismus* π_f *und einen Monomorphismus* ι_f *derart, daß das folgende Diagramm*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi_f"'] & W \\ V/\operatorname{Kern} f \arrow[ur, "\iota_f"] & \end{tikzcd}$$

kommutiert, d.h. daß $f = \iota_f \circ \pi_f$ *gilt. Insbesondere ist dann* $V/\operatorname{Kern} f \cong f[V]$.

Beweis: Nach dem Homomorphiesatz für Gruppen (Satz 2.2.7) wissen wir bereits, daß die Aussage des Satzes für die den Vektorräumen zugrunde liegenden Gruppen korrekt ist. Es bleibt daher nur zu zeigen, daß die Abbildungen π_f und ι_f auch linear sind. Für den kanonischen Epimorphismus haben wir das in Satz 2.3.10 bereits eingesehen. Wir müssen uns also nur noch davon überzeugen, daß sich die kanonische Einbettung $\iota_f(a + \text{Kern } f) := f(a)$ linear ist. Dazu berechnen wir $\iota_f(\alpha(a + \text{Kern } f)) = \iota_f(\alpha a + \text{Kern } f) = f(\alpha a) = \alpha f(a) = \alpha\iota_f(a + \text{Kern } f)$. Damit ist ι_f linear, und alle Behauptungen folgen aus Satz 2.2.7. □

2.4 Direkte Summen von Unterräumen

2.4.1 Definition Sei V ein K-Vektorraum und $U_1, \ldots, U_n$ Unterräume von V. Wir definieren

$$\sum_{i=1}^{n} U_i := \{\sum_{i=1}^{n} x_i |\ x_i \in U_i\}$$

und nennen $\sum_{i=1}^n U_i$ die *Summe der Unterräume U_i*.

2.4.2 Satz *Sind $U_1, \ldots, U_n$ Teilräume von V, so ist $\sum_{i=1}^n U_i$ wieder ein Teilraum von V und es gilt $\sum_{i=1}^n U_i = \langle U_1 \cup \ldots \cup U_n \rangle$.*

Beweis: Es genügt natürlich $\sum_{i=1}^n U_i = \langle U_1 \cup \ldots \cup U_n \rangle$ zu zeigen, da $\langle U_1 \cup \ldots \cup U_n \rangle$ per definitionem ein Teilraum ist. Gilt $x_i \in U_i$ für $i = 1, \ldots, n$, so ist $\sum_{i=1}^n x_i \in \langle \bigcup_{i=1}^n U_i \rangle$. Damit haben wir $\sum_{i=1}^n U_i \subseteq \langle \bigcup_{i=1}^n U_i \rangle$. Ist umgekehrt $v \in \langle \bigcup_{i=1}^n U_i \rangle$, so gibt es $y_i \in U_i$ und $\alpha_i \in K$ mit $v = \sum_{i=1}^n \alpha_i y_i$, wobei wir ggf. $\alpha_i = 0$ zulassen. Da alle U_i Unterräume sind, ist mit y_i auch $\alpha_i y_i \in U_i$. Setzen wir $x_i := \alpha_i y_i \in U_i$, so erhalten wir $v = \sum_{i=1}^n x_i \in \sum_{i=1}^n U_i$. □

Wir erinnern an die Definition von $\bigoplus_{i=1}^n U_i$ in (1.17).

2.4.3 Satz *Die Abbildung*

$$\oplus\colon \bigoplus_{i=1}^{n} U_i \longrightarrow \sum_{i=1}^{n} U_i$$
$$(x_1, \ldots, x_n) \mapsto \sum_{i=1}^n x_i$$

ist ein Epimorphismus der Vektorräume. Wir nennen die Summe der Unterräume direkt, *wenn $\oplus$ ein Isomorphismus ist. Direkte Summen von Unterräumen notieren wir als $\bigoplus_{i=1}^n U_i$ oder $U_1 \oplus \ldots \oplus U_n$.*

Beweis: Zunächst ist klar, daß die Abbildung $\oplus$ surjektiv ist. Wir haben daher nur die Linearität nachzuprüfen. Es ist $\oplus(\alpha(a_1, \ldots, a_n)) = \oplus(\alpha a_1, \ldots, \alpha a_n) = \sum_{i=1}^n \alpha a_i = \alpha \sum_{i=1}^n a_i = \alpha \cdot \oplus(a_1, \ldots, a_n)$ und $\oplus((a_1, \ldots, a_n) + (b_1, \ldots, b_n)) = \oplus(a_1 + b_1, \ldots, a_n + b_n) = \sum_{i=1}^n (a_i + b_i) = \sum_{i=1}^n a_i + \sum_{i=1}^n b_i = \oplus(a_1, \ldots, a_n) + \oplus(b_1, \ldots, b_n)$. Damit ist $\oplus$ linear. □

2.4.4 Satz *Seien $U_1, \dots, U_n$ Teilräume von V. Die Abbildung* $\oplus\colon \bigoplus_{i=1}^{n} U_i \longrightarrow \sum_{i=1}^{n} U_i$ *ist genau dann ein Isomorphismus, wenn gilt:*

$$(\forall i \in \{1, \dots, n\})[U_i \cap \sum_{\substack{j=1\\ j\neq i}}^{n} U_j = \{0\}].$$

Insbesondere gilt $U + W = U \oplus W \Leftrightarrow U \cap W = \{0\} =: \mathbf{0}$.

Beweis: Wegen seiner Surjektivität ist $\oplus$ genau dann bijektiv, wenn es injektiv ist.
$\Rightarrow$: Dazu gehen wir von $U_i \cap \sum_{\substack{j=1\\ j\neq i}}^{n} U_j \neq \{0\}$ für ein $i \in \{1, \dots, n\}$ aus und wählen $0 \neq x_i \in U_i \cap \sum_{\substack{j=1\\ j\neq i}}^{n} U_j \neq \{0\}$. Dann ist $x_i = \sum_{\substack{j=1\\ j\neq i}}^{n} y_i$ und damit $x_i + \sum_{\substack{j=1\\ j\neq i}}^{n} (-y_i) = 0$ und somit $(-y_1, \dots, -y_{i-1}, x_i, -y_{i+1}, \dots, -y_n) \in$ Kern $\oplus$. Damit ist $\oplus$ nicht injektiv.
$\Leftarrow$: Hier gehen wir von $U_i \cap \sum_{\substack{j=1\\ j\neq i}}^{n} U_j = \{0\}$ aus und betrachten $(x_1, \dots, x_n) \in$ Kern $\oplus$. Dann gilt aber $x_i = \sum_{\substack{j=1\\ j\neq i}}^{n} -x_j$ für $i = 1, \dots, n$ und damit $x_i \in U_i \cap \sum_{\substack{j=1\\ j\neq i}}^{n} U_j = \{0\}$. Also ist $x_i = 0$ für $i = 1, \dots, n$ und damit $(x_1, \dots, x_n) = \vec{0}$, d.h. Kern $\oplus = \mathbf{0}$ und $\oplus$ ist injektiv. $\square$

Beachte *Gilt* $V = U_1 \oplus \dots \oplus U_n$, *so hat jedes* $v \in V$ *eine Darstellung* $v = u_1 + \dots + u_n$. *Ist* $v = u_1' + \dots + u_n'$ *eine weitere Darstellung, so folgt wegen der Injektivität von* $\oplus$ *schon* $(c_1, \dots, c_n) = (c_1', \dots, c_n')$, *d.h.* $c_i = c_i'$ *für* $i = 1, \dots, n$. *Die Darstellung als Summe ist bei direkten Summen von Unterräumen also eindeutig.*

2.4.5 Satz *Sei V ein K-Vektorraum und U ein Teilraum von V. Dann gibt es einen Teilraum W von V mit $V = U \oplus W$. W heißt das* Komplement *von U in V.*

Beweis: Sei B eine Basis von U. Nach Satz 1.2.20 finden wir eine Menge $C \subseteq V$, so daß $B \cap C = \emptyset$ und $B \cup C$ eine Basis von V ist. Wir definieren $W := \langle C \rangle$. Ist $v \in V$, so gibt es endliche Mengen $I \subseteq B$ und $J \subseteq C$ mit $v = \sum_{b \in I} \alpha_b b + \sum_{c \in J} \beta_c c \in U + W$. Es bleibt als nur noch die Direktheit der Summe nachzuweisen. Ist $u \in U$, $w \in W$ mit $u + w = 0$, so erhalten wir wieder $0 = u + w = \sum_{b \in I} \alpha_b b + \sum_{c \in J} \beta_c c$ mit endlichen Mengen $I \subseteq B$ und $J \subseteq C$. Wegen der linearen Unabhängigkeit der Menge $B \cup C$ folgt daraus $\alpha_b = 0$ für alle $b \in I$ und $\beta_c = 0$ für alle $c \in J$. Also ist $u = \sum_{b \in I} \alpha_b b = 0$ und $w =$

$\sum_{c \in J} \beta_c c = 0$, d.h. Kern $\oplus = \{(u,w) \in U \times W \mid u + w = 0\} = \{(0,0)\}$, und $\oplus: U \oplus W \longrightarrow U + W$ ist injektiv. □

2.4.6 Satz *Ist V ein K-Vektorraum mit $\dim(V) = n$, so ist $V \cong K^n$.*

Bemerkung *Satz 2.4.6 besagt insbesondere, daß K-Vektorräume gleicher (endlicher) Dimension isomorph sind. Es gilt sogar, daß zwei endlich dimensionale K-Vektorräume genau dann isomorph sind, wenn sie die gleiche Dimension haben.*

Beweis: Wir wollen uns als erstes die Bemerkung klar machen. Sind zwei Vektorräume isomorph, so haben sie nach Korollar 2.3.7 gleiche Dimension. Haben umgekehrt zwei Vektorräume V und W die gleiche endliche Dimension n, so gilt $V \cong K^n \cong W$ und damit $V \cong W$.
Nun zum Beweis des Satzes. Wir wählen eine Basis $B := \{b_1, \ldots, b_n\}$ und definieren die *Koordinatendarstellung* $h_B(v)$ von $v = \sum_{i=1}^n \alpha_i b_i$ bezüglich der Basis B durch

$$\begin{aligned} h_B: V &\longrightarrow K^n \\ \textstyle\sum_{i=1}^n \alpha_i b_i &\mapsto (\alpha_1, \ldots, \alpha_n). \end{aligned} \tag{2.2}$$

Wir behaupten, daß h_B ein Isomorphismus der Vektorräume ist. Da B eine Basis ist, hat jedes $v \in V$ nach Satz 1.2.16 eine eindeutige Darstellung in der Form $v = \sum_{i=1}^n \alpha_i b_i$. Damit ist h_B auf ganz V wohldefiniert. Man rechnet nun leicht nach, daß h_B linear ist. Zu jedem $(\alpha_1, \ldots, \alpha_n) \in K^n$ ist $v := \sum_{i=1}^n \alpha_i b_i$ ein Urbild. Damit ist h_B surjektiv. Gilt $h_B(\sum_{i=1}^n \alpha_i b_i) = (\alpha_1, \ldots, \alpha_n) = 0$, so folgt $\alpha_1 = \ldots = \alpha_n = 0$ und damit $v = 0$. Damit ist Kern $h_B = \{0\}$ und somit h_B auch injektiv. □

2.4.7 Satz *Für endlich dimensionale Vektorräume V und W gilt* $\dim(V \times W) = \dim(V) + \dim(W)$.

Beweis: Ist $\dim V = m$ und $\dim W = n$, so folgt nach Satz 2.4.6 $V \cong K^m$ und $W \cong K^n$. Seien B und C Basen von V bzw. W und $h_B: V \longrightarrow K^m$ und $h_C: W \longrightarrow K^n$ die Koordinatendarstellungen bzgl. B und C. Für $(\alpha_1, \ldots, \alpha_m) \in K^m$ und $(\beta_1, \ldots, \beta_n) \in K^n$ definieren wir

$$(\alpha_1, \ldots, \alpha_n) \star (\beta_1, \ldots, \beta_m) := (\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n) \in K^{m+n}.$$

Die Abbildung

$$\begin{aligned} h_{B \times C}: V \times W &\longrightarrow K^{m+n} \\ (v,w) &\longmapsto h_B(v) \star h_C(w) \end{aligned}$$

definiert dann, wie man leicht nachrechnet, einen Isomorphismus der Vektorräume. Damit ist $\dim(V \times W) = \dim K^{m+n} = m + n = \dim V + \dim W$. □

2.4.8 Satz *Sind V_i für $i = 1, \ldots, n$ endlich dimensionale Vektorräume, so gilt*

$$\dim(\prod_{i=1}^{n} V_i) = \dim(\bigoplus_{i=1}^{n} V_i) = \sum_{i=1}^{n} \dim(V).$$

Beweis: Der Beweis folgt durch Induktion nach n aus Satz 2.4.7. □

2.4.9 Satz *Sind U, W endlich dimensionale Teilräume eines Vektorraumes V, so gilt:*

$$U + W = U \oplus W \Leftrightarrow \dim(U + W) = \dim(U) + \dim(W).$$

Beweis: Die Abbildung $\oplus: U \oplus W \longrightarrow U + W$ ist surjektiv. Nach Korollar 2.3.7 und Satz 2.3.9 ist $\oplus$ genau dann injektiv, wenn $\dim(U + W) = \dim(U \oplus W) = \dim U + \dim W$ ist. □

2.4.10 Satz *Sind U und W endlich dimensionale Teilräume eines Vektorraumes V, so gilt* $\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$.

Beweis: Da $U \cap W$ ein Teilraum von U und W ist, erhalten wir nach Satz 2.4.5 Teilräume U_1 und W_1 mit $U = U_1 \oplus (U \cap W)$ und $W = W_1 \oplus (U \cap W)$. Damit ist $U + W = U_1 + W_1 + (U \cap W)$. Für $x \in U_1 \cap (W_1 + (U \cap W))$ folgt wegen $U_1 \subseteq U$ schon $x \in U_1 \cap (U \cap W)$ und damit $x = 0$. Aus $x \in W_1 \cap (U_1 + (U \cap W))$ und $x \in (U \cap W) \cap (U_1 + W_1)$ folgt in analoger Weise $x = 0$. Damit ist nach Satz 2.4.4 $U + W = U_1 \oplus W_1 \oplus (U \cap W)$, und wir erhalten $\dim(U + W) = \dim U_1 + \dim W_1 + \dim(U \cap W) = \dim U - \dim(U \cap W) + \dim W - \dim(U \cap W) + \dim(U \cap W) = \dim U + \dim W - \dim(U \cap W)$. □

2.4.11 Satz *Sei V ein Vektorraum. Ist $V = U \oplus W$, so ist $W \cong V/U$.*

Beweis: Da $V = U \oplus W$ ist, gibt es zu jedem $v \in V$ eindeutig bestimmte Vektoren $u \in U$ und $w \in W$ mit $v = u + w$. Wir definieren eine Abbildung $f: V \longrightarrow W$ durch $f(u + w) := w$. Man rechnet mühelos nach, daß dann f eine lineare Abbildung mit Kern $f = U$ und Im $f = W$ ist. Nach dem Homomorphiesatz für Vektorräume folgt $V/U = V/\text{Kern } f \cong \text{Im } f = W$. □

2.4.12 Satz *Ist $f: V \longrightarrow W$ eine lineare Abbildung, so gilt* $\dim(V) = \dim(\text{Kern } f) + \dim(\text{Im}\,(f))$.

Beweis: Da Kern f ein Teilraum von V ist, erhalten wir nach Satz 2.4.5 einen Teilraum W mit $V = \text{Kern } f \oplus W$. Nach Satz 2.4.11 gilt dann $W \cong V/\text{Kern } f \cong \text{Im } f$. Damit folgt $\dim V = \dim \text{Kern } f + \dim \text{Im } f$. □

2.5 Der Vektorraum Hom(V,W)

2.5.1 Definition Es seien V und W beides K-Vektorräume. Wir definieren

$$\text{Hom}_K(V,W) := \{f \mid f\colon V \longrightarrow W \wedge f \text{ ist linear}\}.$$

Ist es klar oder unerheblich, um welchen Körper K es sich handelt, schreiben wir Hom(V,W) statt Hom$_K(V,W)$.

2.5.2 Satz *Die Menge* Hom$_K(V,W)$ *wird mit den Verknüpfungen*

- $(f+g)(v) := f(v) + g(v)$

und

- $(\alpha \cdot f)(v) := \alpha \cdot f(v)$

zu einem K-Vektorraum.

Beweis: Durch Nachrechnen erhalten wir aus der Definition von $f+g$ und αf

$$f, g \in \text{Hom}(V,W) \quad \Rightarrow \quad f+g \in \text{Hom}(V,W) \tag{i}$$

und

$$f \in \text{Hom}(V,W) \quad \Rightarrow \quad \alpha f \in \text{Hom}(V,W). \tag{ii}$$

Wegen $(f + (g+h))(v) = f(v) + (g+h)(v) = f(v) + (g(v) + h(v)) = (f(v) + g(v)) + h(v) = ((f+g)+h)(v)$ erhalten wir

$$(\text{Hom}(V,W), +) \text{ ist Halbgruppe }. \tag{iii}$$

Definieren wir $\mathbf{0}\colon V \longrightarrow W$ durch $\mathbf{0}(v) := 0_W$, so ist $\mathbf{0} \in \text{Hom}(V,W)$, und wir erhalten:

$$\mathbf{0} \text{ ist ein neutrales Element bezüglich } + . \tag{iv}$$

Ebenso folgt

$$-f := -1 \cdot f \text{ ist invers zu } f. \tag{v}$$

Wegen $(f+g)(v) = f(v) + g(v) = g(v) + f(v) = (g+f)(v)$ folgt

$$(\mathrm{Hom}(V, W), +) \text{ ist abelsch.} \tag{vi}$$

Schließlich rechnen wir die folgenden Eigenschaften mühelos nach:

$$\alpha(f+g) = \alpha f + \alpha g \text{ und } (\alpha+\beta)f = \alpha f + \beta f \tag{vii}$$

$$\alpha(\beta f) = (\alpha\beta) f \tag{viii}$$

und

$$1 \cdot f = f. \tag{ix}$$

Mit (i) – (ix) haben wir $\mathrm{Hom}(V, W)$ als Vektorraum nachgewiesen. □

2.5.3 Satz *Gilt* $f, f_1, f_2 \in \mathrm{Hom}(V, V')$ *und* $g \in \mathrm{Hom}(V', W)$*, so folgt*

(i) $(g \circ f) \in \mathrm{Hom}(V, W)$

und

(ii) $g \circ (f_1 + f_2) = (g \circ f_1) + (g \circ f_2)$.

Beweis: Die Behauptung (i) folgt aus Lemma 2.3.3. Die Behauptung (ii) rechnet man einfach nach. □

2.5.4 Satz *Der Vektorraum* $\mathrm{Hom}_K(V, V)$ *der Endomorphismen ist mit der Komposition von Abbildungen ein Ring mit Einselement.*

Beweis: Aus Satz 2.5.3 erhalten wir sofort, daß $\mathrm{Hom}(V, V)$ ein Ring ist. Die Identität auf V ist natürlich linear und liefert das Einselement des Ringes. □

2.5.5 Definition Man nennt $\mathrm{End}_K(V) := \mathrm{Hom}_K(V, V)$ den *Endomorphismenring* von V.

2.5.6 Satz *Die Menge* $\mathrm{GL}_K(V) := \{f \mid f : V \longrightarrow V \wedge f \text{ ist linear und bijektiv}\}$ *bildet mit der Komposition eine Gruppe. Sie heißt die* allgemeine lineare Gruppe von V *(*General *L*inear *Group). In der Regel lassen wir das Subskript* $_K$ *weg und schreiben kurz* $\mathrm{GL}(V)$. *Es ist* $\mathrm{GL}(V) \subseteq \mathrm{End}(V)$.

Beweis: Die Identität auf V ist linear und bijektiv. Damit besitzt $\mathrm{GL}(V)$ ein neutrales Element. Für $f \in \mathrm{GL}(V)$ ist nach Lemma 2.3.4 f^{-1} ebenfalls linear und bijektiv. Damit hat jedes Element in $\mathrm{GL}(V)$ ein Inverses. □

2.5.7 Bemerkung $\mathrm{End}_K(V)$ *ist ein Ring mit 1, der* nicht *nullteilerfrei ist.*

Beweis: Sei $V = U \oplus W$. Dann sind die Abbildungen $f_U : V \longrightarrow V$ mit $f_U(u + w) := u$ und $f_W : V \longrightarrow V$ mit $f_W(u + w) := w$ beide linear und nicht die Nullabbildung. Es gilt aber $(f_W \circ f_U)(u+w) = f_W(u) = 0$ für alle $v = u+w \in V$. Damit ist $f_W \circ f_U = \mathbf{0}$, d.h., f_U und f_W sind Nullteiler. $\square$

3. Matrizen

3.1 Basisdarstellung linearer Abbildungen

3.1.1 Beobachtung Es seien V und W endlich dimensionale K-Vektorräume und $B = \{b_1, \ldots, b_m\}$ eine Basis für V sowie $C = \{c_1, \ldots, c_n\}$ eine Basis für W. Ist $f \in \mathrm{Hom}_K(V, W)$ und $v \in V$, so gilt

$$v = \sum_{i=1}^{m} (\alpha_i b_i)$$

und somit

$$f(v) = f(\sum_{i=1}^{m} (\alpha_i b_i)) = \sum_{i=1}^{m} (\alpha_i f(b_i)).$$

Damit ist die Abbildung f durch die Werte $f(b_1), \ldots, f(b_m)$ eindeutig bestimmt. Umgekehrt definiert jede Abbildung $k\colon B \longrightarrow C$ eine lineare Abbildung $\bar{k}$ durch die Festsetzung

$$\bar{k}(\sum_{i=1}^{m} (\alpha_i b_i)) := \sum_{i=1}^{m} (\alpha_i k(b_i)) \in W.$$

Nun ist $f(b_i) \in W$ für $i = 1, \ldots, m$. Da $\{c_1 \ldots c_n\}$ Basis von W ist, gibt es $\alpha_{ij} \in K$ für $j = 1, \ldots, n$ mit

$$f(b_i) = \sum_{j=1}^{n} \alpha_{ij} c_j.$$

Damit ist f durch die Körperelemente $(\alpha_{ij})_{\substack{i=1\ldots m \\ j=1\ldots n}}$ eindeutig bestimmt.

3.1.2 Definition Eine $m \times n$ Matrix über einem Körper K ist ein Schema der Form

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} =: (\alpha_{ij})_{\substack{i=1\dots m \\ j=1\dots n}} =: (\alpha_{ij})_{\substack{m \\ n}} =: (\alpha_{ij})_{m,n}$$

mit $\alpha_{ij} \in K$. Die Menge der $m \times n$ Matrizen über K bezeichnen wir mit $K^{m\times n}$, d.h. es ist

$$K^{(m,n)} := K^{m\times n} := \{\mathfrak{A} \mid \ \mathfrak{A} \text{ ist eine } m \times n\text{-Matrix über } K\}.$$

Die n-tupel $(\alpha_{i1}, \dots, \alpha_{in})$ für $i = 1, \dots, m$ heißen die *Zeilen* der Matrix, die Vektoren

$$\begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} \quad \text{für } j = 1, \dots, n$$

heißen die *Spalten* der Matrix. Die Elemente α_{ii} für $i = 1, \dots, \min\{m, n\}$ bilden die *Hauptdiagonale*, die Elemente $\alpha_{(m-k)(k+1)}$ für $k = 0, \dots, \min\{m, n\} - 1$ die *Nebendiagonale* der Matrix.

Bemerkung *Die Schreibweise von Matrizen als rechteckige Schemata hat lediglich (sehr gute) pragmatische Gründe. Man kann eine $m \times n$-Matrix*

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$

als ein m-tupel von n-tupeln $(\alpha_{11}, \dots, \alpha_{1n}), \dots, (\alpha_{m1}, \dots, \alpha_{mn})$ also als Vektor aus $(K^n)^m$ auffassen.

Dies ist der Hintergrund des folgenden Satzes.

3.1.3 Satz *Mit den Verknüpfungen*

- $(\alpha_{ij})_{m,n} + (\beta_{ij})_{m,n} := (\alpha_{ij} + \beta_{ij})_{m,n}$

und

- $\alpha \cdot (\beta_{ij})_{m,n} := (\alpha \cdot \beta_{ij})_{m,n}$

wird der $K^{m\times n}$ zu einem Vektorraum.

Beweis: Die eben eingeführten Verknüpfungen sind genau die Verknüpfungen des $(K^n)^m = K^{n\cdot m}$. Wir wissen aber schon, daß $K^{n\cdot m}$ ein Vektorraum ist. Im wesentlichen ist die $m \times n$-Matrizen-Schreibweise nur eine bequemere Schreibweise für den Vektorraum $K^{n\cdot m}$.

Das *Nullelement* im $K^{m\times n}$ bezeichnen wir durch

$$0 := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

Das Inverse (bzgl. der Addition) zu (α_{ij}) ist $(-\alpha_{ij})$. Wir wollen in Zukunft immer vom *Negativen* eines Elementes einer additiven abelschen Gruppe sprechen, wenn wir das Inverse bezüglich der Addition meinen. Das erspart Verwirrungen, da wir auf den Matrizen auch eine multiplikative Struktur einführen werden.

3.1.4 Lemma *Ist V ein Vektorraum der Dimension n, so wird die Abbildung id$_V \in$ $\mathrm{Hom}(V,V)$ bezüglich jeder Basis B von V durch die Matrix $(\delta_{ij})_{n,n}$ dargestellt. Die Matrix*

$$(\delta_{ij})_{n,n} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

heißt Einheitsmatrix *und wird mit $\mathfrak{E}^{(n,n)}$ bezeichnet. Oft ist aus dem Zusammenhang klar, welches n gemeint ist. Dann schreiben wir kurz $\mathfrak{E}$.*

Beweis: Es ist $id_V(b_i) = b_i = \sum_{j=1}^{n} \delta_{ij} b_j$. □

Notationen

Anstatt $K^{m\times 1}$ schreiben wir K_m und nennen K_m den Vektorraum der *Spaltenvektoren* über K.

$$K_m := K^{m\times 1} := \{ \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} \bigm| \alpha_i \in K\}.$$

Statt $K^{1\times n}$ schreiben wir K^n und nennen K^n den Vektorraum der *Zeilenvektoren* über K.

$$K^n := K^{1\times n} := \{(\alpha_1,\ldots,\alpha_n) \bigm| \alpha_i \in K\}.$$

3.1.5 Satz *Seien V und W K-Vektorräume und $B = \{b_1,\ldots,b_m\}$ sowie $C = \{c_1,\ldots,c_n\}$ Basen von V bzw. W. Wir definieren eine Abbildung*

$$\mathfrak{D}_{B,C} \colon \mathrm{Hom}_K(V,W) \longrightarrow K^{m\times n}$$

durch

$$\mathfrak{D}_{B,C}(f) := (\alpha_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}$$

falls

$$f(b_i) = \sum_{j=1}^{n} \alpha_{ij} c_j$$

für alle $i = 1, \ldots, m$ *gilt.*
Die Matrix $\mathfrak{D}_{B,C}(f)$ *heißt die* Matrixdarstellung *von* f *bezüglich der Basen* B *und* C. *Die Abbildung* $\mathfrak{D}_{B,C}$ *ist ein Isomorphismus der Vektorräume.*

Beweis: Wir haben bereits in 3.1.1 beobachtet, daß die Matrix $\mathfrak{D}_{B,C}(f)$ für alle $f \in \mathrm{Hom}(V,W)$ definiert ist. Umgekehrt erzeugt jede Matrix $\mathfrak{A} = (\alpha_{ij})_{m,n}$ vermöge $k_{\mathfrak{A},B,C}(b_i) := \sum_{j=1}^{n} \alpha_{ij} c_j$ eine Abbildung $k_{\mathfrak{A},B,C}\colon B \longrightarrow C$, die sich, wie in 3.1.1 bemerkt, zu einer linearen Abbildung $\bar{k}_{\mathfrak{A},B,C}\colon V \longrightarrow W$ fortsetzen läßt. Dann gilt $\mathfrak{D}_{B,C}(\bar{k}_{\mathfrak{A},B,C}) = \mathfrak{A}$ und ist $\mathfrak{D}_{B,C}$ eine surjektive Abbildung. Wir haben ihre Linearität nachzurechnen. Dazu beobachten wir, daß mit $\mathfrak{D}_{B,C}(f) = (\alpha_{ij})$ und $\mathfrak{D}_{B,C}(g) = (\beta_{ij})$ gilt

$$(f+g)(b_i) = f(b_i) + g(b_i) = \sum_{j=1}^{n} \alpha_{ij} c_j + \sum_{j=1}^{n} \beta_{ij} c_j = \sum_{j=1}^{n} (\alpha_{ij} + \beta_{ij}) c_j$$

und damit $\mathfrak{D}_{B,C}(f+g) = \mathfrak{D}_{B,C}(f) + \mathfrak{D}_{B,C}(g)$. Analog erhalten wir $\alpha f(b_i) = \alpha \cdot \sum_{j=1}^{n} \alpha_{ij} c_j = \sum_{j=1}^{n} \alpha\alpha_{ij} c_j$ und damit $\mathfrak{D}_{B,C}(\alpha f) = \alpha \cdot \mathfrak{D}_{B,C}(f)$.
Ist $\mathfrak{D}_{B,C}(f) = 0$, so folgt $f(b_i) = 0$ für $i = 1, \ldots, m$. Damit ist f die Nullabbildung, d.h. Kern$(\mathfrak{D}_{B,C}) = \{0\}$, und $\mathfrak{D}_{B,C}$ ist injektiv. □

3.1.6 Korollar *Sind* V *und* W *endlich dimensionale Vektorräume der Dimensionen* m *und* n, *so gilt:*

$$\mathrm{Hom}(V,W) \cong K^{m \times n}$$

und

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

3.1.7 Beobachtung Seien V, V' und W Vektorräume, $f \in \mathrm{Hom}(V,V')$, $g \in \mathrm{Hom}(V',W)$, $B = \{b_1, \ldots, b_m\}$ eine Basis von V, $C = \{c_1, \ldots, c_k\}$ eine Basis von V' und $D = \{d_1, \ldots, d_n\}$ eine Basis von W.
Ist $\mathfrak{D}_{B,C}(f) = (\alpha_{ij})_{m,k}$ und $\mathfrak{D}_{C,D}(g) = (\gamma_{ij})_{k,n}$, so ist

$$\mathfrak{D}_{B,D}(g \circ f) = (\sum_{l=1}^{k} \alpha_{il}\gamma_{lj})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}}.$$

Die eben gemachte Beobachtung ist die Motivation für die folgende Definition des Produktes von Matrizen.

3.1.8 Definition (Matrixmultiplikation) Sei $\mathfrak{A} \in K^{m\times k}$, $\mathfrak{B} \in K^{k\times n}$ mit $\mathfrak{A} = (\alpha_{ij})_{\substack{i=1,\ldots,m\\ j=1,\ldots,k}}$ und $\mathfrak{B} = (\beta_{ij})_{\substack{i=1,\ldots,k\\ j=1,\ldots,n}}$. Wir definieren

$$\mathfrak{A} \cdot \mathfrak{B} := (\sum_{l=1}^{k} \alpha_{il}\beta_{lj})_{\substack{i=1,\ldots,m\\ j=1,\ldots,n}}.$$

und nennen $\mathfrak{A} \cdot \mathfrak{B}$ das *Produkt* der Matrizen $\mathfrak{A}$ und $\mathfrak{B}$.

Aus der Definition der Matrixmultiplikation und Beobachtung 3.1.7 erhalten wir sofort den folgenden Satz.

3.1.9 Satz *Es seien V, V' und W K-Vektorräume mit $\dim(V) = m, \dim(V') = k$ und $\dim(W) = n$. Sind B, C und D Basen von V, V' und W, $f \in \mathrm{Hom}(V, V')$ und $g \in \mathrm{Hom}(V', W)$, so gilt*

$$\mathfrak{D}_{B,D}(g \circ f) = \mathfrak{D}_{B,C}(f) \cdot \mathfrak{D}_{C,D}(g).$$

Warnung *Das Ergebnis des Satzes 3.1.9 wird oft als unschön empfunden, da die Abbildung $\mathfrak{D}_{B,B}\colon \mathrm{Hom}(V, V) \longrightarrow K^{n\times n}$ kein Homomorphismus der Ringe, sondern ein sogenannter* Antihomomorphismus *wird. D.h. es gilt $\mathfrak{D}_{B,B}(g \circ f) = \mathfrak{D}_{B,B}(f) \cdot \mathfrak{D}_{B,B}(g)$ anstatt $\mathfrak{D}_{B,B}(g \circ f) = \mathfrak{D}_{B,B}(g) \cdot \mathfrak{D}_{B,B}(f)$. Will man dies vermeiden, so muß man $\mathfrak{D}_{B,C}(f) := (\beta_{ij})_{m,n}$ definieren für $f(b_i) = \sum_{j=1}^{n} \alpha_{ji}c_j$ und $\beta_{ij} := \alpha_{ji}$ für $i = 1, \ldots, n$ und $j = 1, \ldots, m$. Die so erhaltene Matrix ist dann die transponierte (s.u.) der von uns eingeführten Matrix. Wir haben uns hier für diese etwas einfachere Definition entschlossen und nehmen dafür die Unschönheit in Kauf.* **Die meisten Lehrbücher gehen jedoch den anderen Weg**. *Für die Entwicklung der Theorie ist die Definition der darstellenden Matrix jedoch nicht entscheidend.*

Beachte Wenn Sie zwei Matrizen $\mathfrak{A}$ und $\mathfrak{B}$ multiplizieren wollen, so muß $\mathfrak{A} \in K^{m\times k}$ und $\mathfrak{B} \in K^{k\times n}$ sein, d.h. die Länge der Zeilen von $\mathfrak{A}$ muß mit der Länge der Spalten von $\mathfrak{B}$ übereinstimmen.
Ein Spezialfall der Matrixmultiplikation ist die Multiplikation eines Zeilenvektors $\mathfrak{a} := (\alpha_1, \ldots, \alpha_n) \in K^n = K^{1\times n}$ mit einem Spaltenvektor $\mathfrak{b} := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in K_n = K^{n\times 1}$, die das Körperelement

$$\mathfrak{a} \cdot \mathfrak{b} = \sum_{j=1}^{n} \alpha_j \beta_j$$

ergibt. Dem einen oder anderen mag dieses Produkt als “Skalarprodukt” oder “inneres Produkt” zweier Vektoren von der Schule her bekannt sein. Schreiben wir die Matrizen als Matrix $\mathfrak{A} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$ von Zeilenvektoren $\mathfrak{a}_i \in K^k$ und $\mathfrak{B} = (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$ als Matrix von Spaltenvektoren $\mathfrak{b}_j \in K_k$, so erhalten wir $\mathfrak{A} \cdot \mathfrak{B} = (\mathfrak{a}_i \mathfrak{b}_j)_{m,n}$. Die Merkregel **Zeile mal Spalte** ist hier oft sehr nützlich.
Man kann eine Matrix auch in “Kästchen” unterteilen. Schreiben wir beispielsweise

$$\mathfrak{A} = \begin{pmatrix} \mathfrak{A}_1 & \mathfrak{A}_2 \\ \mathfrak{A}_3 & \mathfrak{A}_4 \end{pmatrix} \text{ und } \mathfrak{B} = \begin{pmatrix} \mathfrak{B}_1 & \mathfrak{B}_2 \\ \mathfrak{B}_3 & \mathfrak{B}_4 \end{pmatrix}$$

mit $\mathfrak{A}_1 \in K^{p\times q}$, $\mathfrak{A}_2 \in K^{p\times(n-q)}$, $\mathfrak{A}_3 \in K^{(m-p)\times q}$, $\mathfrak{A}_4 \in K^{(m-p)\times(n-q)}$, $\mathfrak{B}_1 \in K^{q\times s}$, $\mathfrak{B}_2 \in K^{q\times(k-s)}$, $\mathfrak{B}_3 \in K^{(n-q)\times s}$, $\mathfrak{B}_4 \in K^{(n-q)\times(k-s)}$, so erhalten wir

$$\mathfrak{A} \cdot \mathfrak{B} = \begin{pmatrix} \mathfrak{A}_1 \cdot \mathfrak{B}_1 + \mathfrak{A}_2 \cdot \mathfrak{B}_3 & \mathfrak{A}_1 \cdot \mathfrak{B}_2 + \mathfrak{A}_2 \cdot \mathfrak{B}_4 \\ \mathfrak{A}_3 \cdot \mathfrak{B}_1 + \mathfrak{A}_4 \cdot \mathfrak{B}_3 & \mathfrak{A}_3 \cdot \mathfrak{B}_2 + \mathfrak{A}_4 \cdot \mathfrak{B}_4 \end{pmatrix},$$

wie man leicht nachrechnet. Auch hier gilt das Prinzip “Zeile mal Spalte”.

3.1.10 Bemerkung *Für Matrizen gelten folgende Rechenregeln:*

(i) $(\alpha \cdot \mathfrak{A}) \cdot \mathfrak{B} = \mathfrak{A} \cdot (\alpha \cdot \mathfrak{B}) = \alpha \cdot (\mathfrak{A} \cdot \mathfrak{B})$ *für* $\alpha \in K, \mathfrak{A} \in K^{m\times n}, \mathfrak{B} \in K^{n\times k}$.

(ii) $\mathfrak{A} \cdot (\mathfrak{B} + \mathfrak{C}) = \mathfrak{A} \cdot \mathfrak{B} + \mathfrak{A} \cdot \mathfrak{C}$ *für* $\mathfrak{A} \in K^{m\times n}, \mathfrak{B}, \mathfrak{C} \in K^{n\times k}$.

(iii) $(\mathfrak{A} + \mathfrak{B}) \cdot \mathfrak{C} = \mathfrak{A} \cdot \mathfrak{C} + \mathfrak{B} \cdot \mathfrak{C}$ *für* $\mathfrak{A}, \mathfrak{B} \in K^{m\times n}, \mathfrak{C} \in K^{n\times k}$.

(iv) $(\mathfrak{A} \cdot \mathfrak{B}) \cdot \mathfrak{C} = \mathfrak{A} \cdot (\mathfrak{B} \cdot \mathfrak{C})$ *für* $\mathfrak{A} \in K^{m\times n}, \mathfrak{B} \in K^{n\times k}, \mathfrak{C} \in K^{k\times l}$.

Beweis: Nach der Wahl geeigneter Basen erhalten wir $\mathfrak{A} = \mathfrak{D}_{B,C}(f)$ für ein $f \in \mathrm{Hom}(K^m, K^k)$, $\mathfrak{B} = \mathfrak{D}_{C,D}(g)$ für ein $g \in \mathrm{Hom}(K^k, K^n)$. Wegen $\alpha(g \circ f)(v) = \alpha \cdot g(f(v)) = g(\alpha f(v)) = (g \circ \alpha f)(v)$ haben wir

$$\alpha(g \circ f) = (g \circ \alpha f) \text{ und } \alpha(g \circ f) = \alpha g \circ f. \tag{i}$$

Damit erhalten wir $(\alpha\mathfrak{A})\mathfrak{B} = (\alpha\mathfrak{D}_{B,C}(f)) \cdot \mathfrak{D}_{C,D}(g) = \mathfrak{D}_{B,C}(\alpha f) \cdot \mathfrak{D}_{C,D}(g) = \mathfrak{D}_{B,D}(g \circ \alpha f) = \mathfrak{D}_{B,D}(\alpha(g \circ f)) = \alpha\mathfrak{D}_{B,D}(g \circ f) = \alpha\cdot(\mathfrak{D}_{B,C}(f)\cdot\mathfrak{D}_{C,D}(g)) = \alpha(\mathfrak{A}\mathfrak{B})$ und $(\alpha\mathfrak{A})\mathfrak{B} = \mathfrak{D}_{B,D}(g \circ \alpha f) = \mathfrak{D}_{B,D}(\alpha g \circ f) = \mathfrak{D}_{B,C}(f)\cdot\mathfrak{D}_{C,D}(\alpha g) = \mathfrak{A}(\alpha\mathfrak{B})$. Damit ist die Aussage (i) bewiesen.
Sei $\mathfrak{C} = \mathfrak{D}_{C,D}(h)$ für ein $h \in \mathrm{Hom}(K^k, K^n)$. Dann ist $\mathfrak{A}(\mathfrak{B} + \mathfrak{C}) = \mathfrak{D}_{C,D}((g + h) \circ f) = \mathfrak{D}_{C,D}(g \circ f + h \circ f) = \mathfrak{D}_{C,D}(g \circ f) + \mathfrak{D}_{C,D}(h \circ f) =$

$(\mathfrak{D}_{B,C}(f)\cdot\mathfrak{D}_{C,D}(g)) + (\mathfrak{D}_{B,C}(f)\cdot\mathfrak{D}_{C,D}(h)) = \mathfrak{A}\mathfrak{B} + \mathfrak{A}\mathfrak{C}$. Womit (ii) bewiesen ist. Die Aussage (iii) zeigt man völlig analog.
Sei nun $\mathfrak{C} = \mathfrak{D}_{D,E}(h)$ für ein $h \in \mathrm{Hom}(K^n, K^p)$. Dann folgt $(\mathfrak{A}\mathfrak{B})\mathfrak{C} = (\mathfrak{D}_{B,C}(f)\cdot\mathfrak{D}_{C,D}(g))\cdot\mathfrak{D}_{D,E}(h) = \mathfrak{D}_{B,D}(g\circ f)\cdot\mathfrak{D}_{D,E}(h) = \mathfrak{D}_{B,E}(h\circ(g\circ f)) = \mathfrak{D}_{B,E}((h\circ g)\circ f)) = \mathfrak{D}_{B,C}(f)\cdot\mathfrak{D}_{C,E}(h\circ g) = \mathfrak{A}(\mathfrak{D}_{C,D}(g)\cdot\mathfrak{D}_{D,E}(h)) = \mathfrak{A}(\mathfrak{B}\mathfrak{C})$. Damit ist dann auch (iv) bewiesen. □

3.1.11 Satz *Die quadratischen Matrizen des $K^{n\times n}$ bilden mit der Matrixaddition und der Matrixmultiplikation einen Ring mit der Einheitsmatrix als Einselement. Ist V ein Vektorraum mit $\dim V = n$ und B eine Basis von V, so ist $\mathfrak{D}_{B,B}\colon \mathrm{End}(V) \longrightarrow K^{n\times n}$ ein Antiisomorphismus der Ringe, d.h. es gilt $\mathfrak{D}_{B,B}(g\circ f) = \mathfrak{D}_{B,B}(f)\cdot\mathfrak{D}_{B,B}(g)$.*

Beweis: Die Ringaxiome wurden in 3.1.10 nachgerechnet. Da $\mathfrak{D}_{B,B}$ in Satz 3.1.5 als Vektorraumisomorphimus nachgewiesen wurde, ist diese Abbildung bijektiv. Daß ein Antihomomorphismus vorliegt, folgt aus Satz 3.1.9. Wir werden im nächsten Abschnitt sehen, daß dieser Antiisomorphismus sich mühelos zu einem Isomorphismus ergänzen läßt. □

3.1.12 Bemerkung Ein K-Vektorraum V, auf dem noch zusätzlich eine Verknüpfung $\cdot\colon V\times V \longrightarrow V$ definiert ist, heißt eine *Algebra*. Der $K^{n\times n}$ bildet mit der Matrixmultiplikation die Algebra der quadratischen Matrizen. Ebenso bilden die Endomorphismen eines Vektorraumes mit der Komposition von Abbildungen eine Algebra.

3.1.13 Definition Eine Matrix $\mathfrak{A} \in K^{n\times n}$ heißt *invertierbar,* wenn es eine Matrix $\mathfrak{B} \in K^{n\times n}$ gibt, mit $\mathfrak{A}\cdot\mathfrak{B} = \mathfrak{E}^{(n,n)}$.
Die Menge der invertierbaren Matrizen des $K^{n\times n}$ bezeichnen wir mit $\mathrm{GL}(n,K)$, d.h.

$$\mathrm{GL}(n,K) := \{\mathfrak{A}\in K^{n\times n} \mid (\exists\mathfrak{B}\in K^{n\times n})[\mathfrak{A}\cdot\mathfrak{B} = \mathfrak{E}]\}$$

3.1.14 Satz *Die Menge* $\mathrm{GL}(n,K)$ *bildet mit der Matrixmultiplikation eine Gruppe. Sie heißt die* allgemeine lineare Gruppe. *Für eine Basis B eines K-Vektorraums der Dimension n gilt* $\mathrm{GL}(n,K) = \{\mathfrak{D}_{B,B}(f) \mid f \in \mathrm{GL}(V)\}$.

Beweis: Ist $f \in \mathrm{GL}(V)$ und $\mathfrak{A} = \mathfrak{D}_{B,B}(f)$, so gibt es $f^{-1} \in Gl(V)$, und es gilt $\mathfrak{E}^{(n,n)} = \mathfrak{D}_{B,B}(f\circ f^{-1}) = \mathfrak{D}_{B,B}(f^{-1})\cdot\mathfrak{D}_{B,B}(f)$, und wir sehen, daß $\mathfrak{D}_{B,B}(f) \in \mathrm{GL}(n,K)$ ist. Ist umgekehrt $\mathfrak{A} \in \mathrm{GL}(n,K)$, so sind die Abbildungen $\bar{k}_{\mathfrak{A},B,B}$ und $\bar{k}_{\mathfrak{A}^{-1},B,B}$ (vgl. den Beweis des Satzes 3.1.5) in $\mathrm{End}(V)$ und es folgt $\mathfrak{D}_{B,B}(id_V) = \mathfrak{E} = \mathfrak{A}^{-1}\cdot\mathfrak{A} = \mathfrak{D}_{B,B}(\bar{k}_{\mathfrak{A},B,B}\circ\bar{k}_{\mathfrak{A}^{-1},B,B})$ und damit $\bar{k}_{\mathfrak{A},B,B} \in$

$\mathrm{GL}(V)$. Damit ist $\mathrm{GL}(n,K) = \{\mathfrak{D}_{B,B}(f) \mid f \in \mathrm{GL}(V)\}$ und eine Gruppe, da $\mathrm{GL}(V)$ eine Gruppe und $\mathfrak{D}_{B,B}$ ein Antiisomorphismus ist. □

Wir wollen den Abschnitt damit beschließen, eine kanonische Basis des $K^{m\times n}$ anzugeben. Dazu definieren wir die Matrizen

$$\mathfrak{E}^{m,n}_{k,l} = (\delta_{ik} \cdot \delta_{jl})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}. \tag{3.1}$$

Das bedeutet, daß $\mathfrak{E}^{m,n}_{k,l}$ die $m \times n$-Matrix ist, die genau im Schnittpunkt der k-ten Zeile und l-ten Spalte eine 1 und sonst lauter Nullen stehen hat. Wenn m und n aus dem Zusammenhang ersichtlich sind, lassen wir die oberen Indizes weg und schreiben kurz $\mathfrak{E}_{k,l}$.

3.1.15 Satz *Die Menge* $\mathbb{E}^{m,n} := \{\mathfrak{E}^{m,n}_{k,l} \mid k = 1,\dots,m; \quad l = 1,\dots,n\}$ *ist eine Basis des* $K^{m\times n}$.

Beweis: Offensichtlich ist $\mathbb{E}^{m,n}$ genau die kanonische Basis des $K^{m\cdot n}$. □

3.2 Der duale Raum und die transponierte Matrix

3.2.1 Definition Für $\mathfrak{A} = (\alpha_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in K^{m\times n}$ sei $\mathfrak{A}^t := (\beta_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} \in K^{n\times m}$ mit $\beta_{ij} = \alpha_{ji}$ für $i = 1,\dots,n; j = 1,\dots,m$.
$\mathfrak{A}^t$ heißt die zu $\mathfrak{A}$ *transponierte* Matrix. Sie entsteht durch Spiegelung an der Hauptdiagonalen.

Um die geometrische Bedeutung der transponierten Matrix genauer zu studieren, führen wir den Begriff des dualen Raumes ein.

3.2.2 Definition Sei V ein K-Vektorraum. Wir definieren $V^* := \mathrm{Hom}(V,K)$. Man nennt V^* den zu V *dualen* Raum.

In der obigen Definition wird der Körper K als ein eindimensionaler Vektorraum über sich selbst aufgefaßt. Eine lineare Abbildung von einem Vektorraum V in dessen Grundkörper bezeichnet man auch als eine *Linearform*. Der duale Raum V^* ist also die Menge der Linearformen von V in K.

3.2.3 Satz *Ist* $\dim(V) < \infty$, *so ist* $V \cong V^*$.

Beweis: Wegen $\dim V^* = \dim \mathrm{Hom}(V,K) = \dim V \cdot \dim K = \dim V$ folgt dies aus der Bemerkung zu Satz 2.4.6. □

Der Isomorphismus, der sich auf Grund von Satz 3.2.3 ergibt, ist allerdings kein sehr natürlicher. Wir erhalten ihn nach Wahl von Basen B für V und C für V^*

als Komposition ${h_C}^{-1} \circ h_B$, wobei die Basen B und C nichts miteinander zu tun haben müssen.
Ein Vektorraum und sein dualer Raum sind jedoch viel enger miteinander verknüpft. Das wollen wir im folgenden herausarbeiten.

3.2.4 Satz *Ist $B = \{b_i \mid\ i \in I\}$ eine Basis von V und setzen wir*

$$b_i^*: V \longrightarrow K \quad b_i^*(\sum_{j=1}^{n} \alpha_j b_j) := \sum_{j=1}^{n} \alpha_j \delta_{ij},$$

so ist die Menge $B^* := \{b_i^* \mid\ i \in I\} \subseteq V^*$ linear unabhängig, und es gilt $b_i^*(b_j) = \delta_{ij}$.

Beweis: Es folgt nach 3.1.1, daß durch die Festsetzung $b_i^*(\sum_{j=1}^{n} \alpha_j b_j) := \sum_{j=1}^{n} \alpha_j \delta_{ij}$ eine Linearform mit $b_i^*(b_j) = \delta_{ij}$ definiert wird. Gilt nun $\sum_{i=1}^{n} \alpha_i b_i^* = 0$, so erhalten wir $0 = (\sum_{i=1}^{n} \alpha_i b_i^*)(b_k) = \sum_{i=1}^{n} \alpha_i b_i^*(b_k) = \sum_{i=1}^{n} \alpha_i \delta_{ik} = \alpha_k$ für $k = 1, \ldots, n$. Also ist $\alpha_1 = \ldots = \alpha_n = 0$, und B^* ist linear unabhängig. □

3.2.5 Korollar *Ist $v \in V$ und $v \neq 0$, so gibt es ein $\lambda \in V^*$ mit $\lambda(v) \neq 0$.*

Beweis: Sei B eine Basis von V und $v = \sum_{i=1}^{n} \alpha_i b_i$. Dann gilt $\alpha_j \neq 0$ für mindestens ein $j \in \{1, \ldots, n\}$. Damit erhalten wir aber $b_j^*(v) = \sum_{i=1}^{n} \alpha_i b_j^*(b_i) = \alpha_j \neq 0$. □

3.2.6 Lemma *Ist $\dim(V) = \dim(W) = n$, $f: V \longrightarrow W$ eine lineare Abbildung und gibt es eine Basis B von V, so daß $f[B]$ eine Basis von W ist, so ist f ein Isomorphismus.*

Beweis: Die Abbildung f ist surjektiv, da sich jedes $w \in W$ als Linearkombination $w = \sum_{i=1}^{n} \alpha_i f(b_i) = f(\sum_{i=1}^{n} \alpha_i b_i)$ mit $b_i \in B$ darstellen läßt. Ist $0 = f(v) = f(\sum_{i=1}^{n} \alpha_i b_i) = \sum_{i=1}^{n} \alpha_i f(b_i)$, so folgt wegen der linearen Unabhängigkeit von $f(b_1), \ldots, f(b_n)$ schon $\alpha_1 = \ldots = \alpha_n = 0$ und damit $v = \sum_{i=1}^{n} \alpha_i b_i = 0$. Damit ist Kern $f = \{0\}$ und f ist injektiv. □

3.2.7 Satz *Ist V ein endlich dimensionaler Vektorraum und $B = \{b_1, \ldots, b_n\}$ eine Basis von V, so ist $B^* = \{b_1^*, \ldots, b_n^*\}$ eine Basis von V^* und die Abbildung*

$$\begin{aligned} {}^*: V &\longrightarrow V^* \\ v = \sum_{i=1}^{n} \alpha_i b_i &\longmapsto \sum_{i=1}^{n} \alpha_i b_i^* =: v^* \end{aligned}$$

ein Isomorphismus der Vektorräume. Man nennt B^* die zu B duale Basis.

Beweis: Wir wissen nach Satz 3.2.4, daß ${b^*}_1, \ldots, {b^*}_n$ linear unabhängig sind. Zu zeigen bleibt daher, daß sie auch ein Erzeugendensystem bilden. Ist $\lambda \in V^*$, so zeigen wir, daß $\lambda = \sum_{i=1}^n \lambda(b_i) b_i^*$ ist. Dies ist aber wegen $(\sum_{i=1}^n \lambda(b_i) b_i^*)(b_j) = \sum_{i=1}^n \lambda(b_i) b_i^*(b_j) = \sum_{i=1}^n \lambda(b_i)\delta_{ij} = \lambda(b_j)$ für $j = 1, \ldots, n$ sofort klar. Da die Abbildung ${}^*: V \longrightarrow V^*$ die Basis B auf die Basis B^* abbildet und V und V^* die gleiche Dimension haben, folgt nach Lemma 3.2.6, daß * ein Isomorphismus ist.

$\square$

Man sollte im vorherigen Beweis beachten, daß er nur für endlich dimensionale Vektorräume funktioniert. Ist V nämlich unendlich dimensional, so wissen wir zwar, daß für jedes $v \in V$ eine endliche Teilmenge $\{b_1, \ldots, b_n\} \subseteq B$ mit $v = \sum_{i=1}^n \alpha_i b_i$ existiert und erhalten für ein $\lambda \in V^*$ daraus $\lambda(v) = \sum_{i=1}^n \alpha_i \lambda(b_i) = \sum_{i=1}^n \alpha_i (\sum_{j=1}^n \lambda(b_j) b_j^*)(b_i)$, d.h. $\lambda(v) = (\sum_{j=1}^n \lambda(b_j) b_j^*)(v)$, aber die Menge $\{b_1, \ldots, b_n\}$ läßt sich nicht uniform angeben, sondern hängt von der Darstellung von v durch Basisvektoren ab. So ist die Abbildung * für unendlich dimensionale Vektorräume zwar injektiv, aber im allgemeinen nicht surjektiv.

3.2.8 Satz *Ist $f \in \mathrm{Hom}(V, W)$ und $\lambda \in W^*$, so ist die Abbildung $f^*: W^* \longrightarrow V^*$, die definiert ist durch $f^*(\lambda) := \lambda \circ f$ in $\mathrm{Hom}(W^*, V^*)$. Wir nennen f^* die zu f* duale Abbildung. *Sind V und W endlich dimensional und sind B und C Basen von V und W, so gilt $\mathfrak{D}_{C^*,B^*}(f^*) = \mathfrak{D}_{B,C}(f)^t$.*

Beweis: Aus der Definition von f^* folgt die Kommutativität des folgenden Diagramms.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "\lambda\circ f", "f^*(\lambda)"'] & W \arrow[d, "\lambda"] \\ & K \end{tikzcd}$$

Damit ist $f^*(\lambda) \in \mathrm{Hom}(V, K) = V^*$. Wegen $f^*(\alpha\lambda_1 + \beta\lambda_2) = (\alpha\lambda_1 + \beta\lambda_2) \circ f = \alpha(\lambda_1 \circ f) + \beta(\lambda_2 \circ f) = \alpha f^*(\lambda_1) + \beta f^*(\lambda_2)$ ist f^* linear.
Sind nun V und W endlich dimensional mit Basen $B = (b_1, \ldots, b_m)$ und $C = (c_1, \ldots, c_n)$ und $\mathfrak{D}_{B,C}(f) = (\alpha_{ij})$, so folgt $f(b_i) = \sum_{j=1}^n \alpha_{ij} c_j$ für $i = 1, \ldots, m$.
Sei $\mathfrak{D}_{C^*,B^*}(f^*) = (\beta_{ij})$. Wir behaupten, daß dann

$$\beta_{ij} = \alpha_{ji} \quad \text{für } i = 1, \ldots, n \text{ und } j = 1, \ldots, m \tag{i}$$

gilt. Zunächst erhalten wir nämlich

$$f^*(c_i^*)(b_j) = (\sum_{k=1}^m \beta_{ik} b_k^*)(b_j) = \sum_{k=1}^m \beta_{ik} b_k^*(b_j) = \beta_{ij} \tag{ii}$$

und dann auch

$$f^*(c_i^*)(b_j) = c_i^*(f(b_j)) = c_i^*(\sum_{k=1}^{n} \alpha_{jk} c_k) = \sum_{k=1}^{n} \alpha_{jk} c_i^*(c_k) = \alpha_{ji}. \qquad \text{(iii)}$$

Aus (ii) und (iii) folgt aber sofort (i). □

3.2.9 Satz *Die Abbildung* $^*\colon \mathrm{Hom}(V,W) \longrightarrow \mathrm{Hom}(W^*,V^*)$ *mit* $f \longmapsto f^*$ *ist ein Monomorphismus der Vektorräume. Sind* V *und* W *endlich dimensional, so ist* $*$ *ein Isomorphismus.*

Beweis: Sei $f \in \mathrm{Hom}(V,W)$. Nach Satz 3.2.8 gilt $f^*\colon W^* \longrightarrow V^*$. Um die Linearität nachzuweisen, berechnen wir (vgl. den Beweis von Satz 3.1.10) $(\alpha f + \beta g)^*(\lambda) = \lambda \circ (\alpha f + \beta g) = \alpha(\lambda \circ f) + \beta(\lambda \circ g) = \alpha \cdot f^*(\lambda) + \beta \cdot g^*(\lambda)$, woraus $(\alpha f + \beta g)^* = \alpha f^* + \beta g^*$ folgt. Damit ist * linear.
Sei $f \in \mathrm{Kern}\,(^*)$. Nehmen wir $f \neq 0$ an, so gibt es ein $v \in V$ mit $f(v) \neq 0$ und nach Korollar 3.2.5 damit ein $\lambda \in W^*$ mit $0 \neq \lambda(f(v)) = (\lambda \circ f)(v) = (f^*(\lambda))(v)$. Damit ist $f^*(\lambda) \neq 0$ und folglich auch $f^* \neq 0$, was unserer Annahme $f \in \mathrm{Kern}\,(^*)$ widerspricht. Damit ist * injektiv.
Ist nun $\dim V = m < \infty$ und $\dim W = n < \infty$, so gilt $\dim \mathrm{Hom}(V,W) = \dim V \cdot \dim W = \dim V^* \cdot \dim W^* = \dim \mathrm{Hom}(W^*,V^*)$ und nach Satz 2.3.9 ist dann * bereits ein Isomorphismus. □

3.2.10 Satz *Für* $f \in \mathrm{Hom}(V,V')$, $g \in \mathrm{Hom}(V',W)$ *ist* $(g \circ f)^* = f^* \circ g^*$. *Damit ist die Abbildung* $^*\colon \mathrm{End}\, V \longrightarrow \mathrm{End}\, V^*$ *ein Antimonomorphismus der Ringe. Insbesondere gilt dann* $(\mathfrak{A} \cdot \mathfrak{B})^t = \mathfrak{B}^t \cdot \mathfrak{A}^t$ *für Matrizen* $\mathfrak{A}$ *und* $\mathfrak{B}$. *Nach Wahl einer Basis* B *für einen Vektorraum* V *der Dimension* $n < \infty$ *ist die Abbildung*

$$\begin{aligned} {}^t \circ \mathfrak{D}_{B,B}\colon \mathrm{End}\, V &\longrightarrow K^{n\times n} \\ f &\longmapsto \mathfrak{D}_{B,B}(f)^t \end{aligned}$$

ein Isomorphismus der Algebren.

Beweis: Die erste Behauptung des Satzes entnehmen wir direkt dem folgenden Diagramm

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "(g\circ f)^*(\lambda)"', "f^*(g^*(\lambda))"] & V' \arrow[r, "g"] \arrow[d, "g^*(\lambda)"] & W \arrow[d, "\lambda"'] \\ K \arrow[r, equal] & K \arrow[r, equal] & K. \end{tikzcd}$$

Da * nach Satz 3.2.9 injektiv ist, folgt wegen $(gf)^* = f^* g^*$ sofort, daß ein Antimonomorphismus der Endomorphismenringe vorliegt. Ist V endlich dimensional, so ist * ein Antiisomorphismus.

Ist $\mathfrak{A} = \mathfrak{D}_{B,C}(f)$ und $\mathfrak{B} = \mathfrak{D}_{C,D}(g)$ für $f \in \mathrm{Hom}(V,V')$, $g \in \mathrm{Hom}(V',W)$ und geeigneten Basen B, C und D für V, V' und W, so erhalten wir $(\mathfrak{A}\mathfrak{B})^t = (\mathfrak{D}_{B,C}(f)\cdot\mathfrak{D}_{C,D}(g))^t = (\mathfrak{D}_{B,D}(g \circ f))^t = \mathfrak{D}_{D^*,B^*}((g \circ f)^*) = \mathfrak{D}_{D^*,B^*}(f^* \circ g^*) = \mathfrak{D}_{D^*,C^*}(g^*) \cdot \mathfrak{D}_{C^*,B^*}(f^*) = \mathfrak{D}_{C,D}(g)^t \cdot \mathfrak{D}_{B,C}(f)^t = \mathfrak{B}^t \cdot \mathfrak{A}^t$.
Wir haben das folgende kommutative Diagramm

$$\begin{tikzcd}
\mathrm{End}(V) \arrow[rr, "*"] \arrow[dr, "\mathfrak{D}_{B^*,B^*}\circ^*"'] & & \mathrm{End}(V^*) \arrow[dl, "\mathfrak{D}_{B^*,B^*}"] \\
& K^{n\times n} &
\end{tikzcd}$$

Im Zusammenhang mit Satz 3.1.9 haben wir schon bemerkt, daß die Abbildung $\mathfrak{D}_{B^*,B^*}\colon \mathrm{End}(V^*) \longrightarrow K^{n\times n}$ ein Antihomomorphismus der Ringe ist. Nach Satz 3.1.5 ist sie ein Isomorphismus der Vektorräume. Da auch * ein Isomorphismus der Vektorräume und ein Antihomomorphismus der Ringe ist und die Komposition zweier Antihomomorphismen einen Homomorphismus ergibt, erhalten wir $\mathfrak{D}_{B^*,B^*}\circ^*$ als einen Isomorphismus sowohl der Vektorräume als auch der Ringe und damit als einen Isomorphismus der Algebren. Nun gilt aber $\mathfrak{D}_{B^*,B^*}(f^*) = \mathfrak{D}_{B,B}(f)^t$, d.h. $^t \circ \mathfrak{D}_{B,B} = \mathfrak{D}_{B^*,B^*}\circ^*$. □

Bemerkung *Satz 3.2.8 ist der Grund, weswegen man üblicherweise die darstellende Matrix eines Homomorphismus $f \in \mathrm{Hom}(V,W)$ bezüglich zweier Basen B und C von V bzw. W als $(\mathfrak{D}_{B,C}(f))^t$ (in unserer Notation) definiert. Man erhält die Abbildung $f \longmapsto$* darstellende Matrix von f *dann direkt als einen Isomorphismus der Algebren.*

Sieht man sich Satz 3.2.9 nochmal genauer an, so bemerkt man, daß der Isomorphismus zwischen $\mathrm{Hom}(V,W)$ und $\mathrm{Hom}(W^*,V^*)$ einfacher anzugeben ist, als der zwischen V und V^*, zu dessen Definition wir eine Basis benötigen. Der Homomorphismus $f \longmapsto f^*$ ist dagegen unabhängig von der Wahl einer Basis. Man sagt daher, daß $\mathrm{Hom}(V,W)$ und $\mathrm{Hom}(W^*,V^*)$ *kanonisch isomorph* sind.
Zwischen V und dem dualen Raum V^* sind solche kanonische Isomorphismen nicht bekannt. Anders ist die Situation, wenn wir zum dualen des dualen Raumes, d.h. zu V^{**} übergehen. Dort erhalten wir im endlich dimensionalen Fall einen kanonischen Isomorphismus.

3.2.11 Satz *Die Abbildung $^{**}\colon V \longrightarrow V^{**}$, $V \ni v \longmapsto v^{**} \in V^{**}$ mit $v^{**}(\lambda) := \lambda(v)$ für $\lambda \in V^*$ definiert einen kanonischen, d.h. basisunabhängigen, Monomorphismus von V in V^{**}. Ist $\dim(V) < \infty$, so ist ** ein kanonischer Isomorphismus.*

Beweis: Wegen $v^{**}(\alpha\lambda + \beta\mu) = (\alpha\lambda + \beta\mu)(v) = \alpha(\lambda(v)) + \beta(\mu(v)) = \alpha \cdot v^{**}(\lambda) + \beta \cdot v^{**}(\mu)$ ist $v^{**} \in \mathrm{Hom}(V^*,K) = V^{**}$.

Weiter erhalten wir $(\alpha v + \beta w)^{**}(\lambda) = \lambda(\alpha v + \beta w) = \alpha \cdot \lambda(v) + \beta \cdot \lambda(w) = \alpha \cdot v^{**}(\lambda) + \beta \cdot w^{**}(\lambda) = (\alpha \cdot v^{**} + \beta \cdot w^{**})(\lambda)$, d.h. $(\alpha v + \beta w)^{**} = \alpha \cdot v^{**} + \beta \cdot w^{**}$, und wir haben die Abbildung ** als linear erkannt.
Ist $v \in \operatorname{Kern} {}^{**}$, so gilt $v^{**}(\lambda) = \lambda(v) = 0$ für alle $\lambda \in V^{**}$. Für $v \neq 0$ gibt es aber nach Korollar 3.2.5 ein $\lambda \in V^*$ mit $\lambda(v) \neq 0$. Damit folgt $v = 0$ und die Abbildung ** ist injektiv.
Ist V endlich dimensional, so folgt $\dim V^{**} = \dim V^* = \dim V$ und nach Satz 2.3.9 ist dann ** ein Isomorphismus. $\square$

Man nennt die Abbildung $V \ni v \mapsto v^{**} \in V^{**}$ die *kanonische Einbettung von V in V^{**}*. Es hat sich als ungemein nützlich erwiesen, die Urbilder und Bilder kanonischer Einbettung zu identifizieren, d.h. als gleiche Objekte zu betrachten. Daher werden wir im folgenden sehr oft einen Vektor $v \in V$ und die mit ihm kanonisch assoziierte Abbildung $v^{**}: V^* \longrightarrow K$ als die gleichen Objekte betrachten. Auf diese Weise läßt sich V als Teilraum von V^{**} auffassen. Ist V endlich dimensional, so betrachten wir V und V^{**} in der Regel als die gleichen Räume. Die Nützlichkeit dieser Betrachtungsweise soll im folgenden klarer werden. Eine erste Anwendung bringt bereits der folgende Satz.

3.2.12 Satz *Sind V und W endlich-dimensionale Vektorräume, so gilt $f^{**}(v^{**}) = (f(v))^{**}$ für alle $f \in \operatorname{Hom}(V, W)$. Ist B eine Basis von V und C eine Basis von W, so folgt $\mathfrak{D}_{B,C}(f) = \mathfrak{D}_{B^{**},C^{**}}(f^{**})$.*

Beweis: Für beliebiges $\lambda \in W^*$ erhalten wir $f^{**}(v^{**})(\lambda) = v^{**}(f^*(\lambda)) = (f^*(\lambda))(v) = \lambda(f(v)) = f(v)^{**}(\lambda)$, d.h. $f^{**}(v^{**}) = f(v)^{**}$.
Ist $\mathfrak{D}_{B,C}(f) = (\alpha_{ij})$, so erhalten wir $f^{**}(b_i^{**}) = f(b_i)^{**} = (\sum_{j=1}^n \alpha_{ij} c_j)^{**} = \sum_{j=1}^n \alpha_{ij} c_j^{**}$ und damit $\mathfrak{D}_{B,C}(f) = \mathfrak{D}_{B^{**},C^{**}}(f^{**})$. $\square$

Bemerkung Wir wollen einmal studieren, wie die Aussage des Satzes 3.2.12 mit der Identifikation von v und v^{**} aussieht. Die Aussage $f^{**}(v^{**}) = f(v)^{**}$ wird dann zu

$$(\forall v \in V)[f^{**}(v) = f(v)], \quad \text{d.h. } f^{**} = f. \tag{3.2}$$

Es folgt aus (3.2), daß die Abbildung *, zweimal angewandt, die Identität ergibt. Solche Abbildungen nennt man *involutorisch*. Da $B^{**} = \{b^{**} \mid b \in B\} = B$ und $C^{**} = C$ folgt auch $\mathfrak{D}_{B,C}(f^{**}) = \mathfrak{D}_{B,C}(f)$, was im Einklang zur Identifikation $f^{**} = f$ steht.
Identifizieren wir für endlich dimensionale Vektorräume V und V^{**}, so sind auch die Räume $\operatorname{End}(V)$ und $\operatorname{End}(V^{**})$ ihrer Endomorphismen gleich und die Abbildung $^*: \operatorname{End}(V) \longrightarrow \operatorname{End}(V^*)$ ist somit ein *involutorischer Antiautomorphismus der Ringe*.

In Satz 3.2.7 haben wir einen basisabhängigen Isomorphismus von V auf V^* definiert. Um zu studieren, ob auch dieser Isomorphismus unter der Identifikation $V^{**} = V$ involutorisch wird, studieren wir zunächst dessen Wirkung auf die Elemente der Basis $B = (b_1, \ldots, b_n)$. Wir erhalten

$$(b_i{}^*)^* = b_i^{**}, \tag{3.3}$$

denn es gilt $b_i^{**}(b_j{}^*) = b_j{}^*(b_i) = \delta_{ij}$, d.h. b_i^{**} erfüllt die Definitionsgleichung $(b_i{}^*)^*(b_j{}^*) = \delta_{ij}$. Aus (3.3) folgt jedoch sofort

$$(v^*)^* = ((\sum_{i=1}^{n} \alpha_i b_i)^*)^* = \sum_{i=1}^{n} \alpha_i (b_i{}^*)^* = \sum_{i=1}^{n} \alpha_i b_i = v. \tag{3.4}$$

Es folgt sofort aus der Definition des Transponierens von Matrizen, daß die Transposition eine involutorische Operation auf Matrizen ist (zweimal spiegeln an der Hauptdiagonalen). Mit Hilfe von (3.3) erhalten wir diese Tatsache auf anderem Wege. Für $\mathfrak{A} = \mathfrak{D}_{B,C}(f)$ gilt nämlich $((\mathfrak{A})^t)^t = ((\mathfrak{D}_{B,C}(f))^t)^t = (\mathfrak{D}_{C^*,B^*}(f^*))^t = \mathfrak{D}_{B^{**},C^{**}}(f^{**}) = \mathfrak{D}_{B,C}(f) = \mathfrak{A}$.

3.2.13 Definition Zwei Vektoren $v \in V$ und $w \in V^*$ heißen *orthogonal* oder *senkrecht*, wenn $w(v) = 0$ ist. Geschrieben wird dies als $w \perp v$.
Für $U \subseteq V$ definieren wir $U^\perp := \{w \in V^* \mid (\forall u \in U)[w \perp u]\}$. Wir nennen $U^\perp$ das *orthogonale Komplement* von U.

Aus Definition 3.2.13 folgt

$$U \subseteq V \quad \Rightarrow \quad U^\perp \subseteq V^* \wedge U^{\perp\perp} \subseteq V^{**} \supseteq V,$$

wobei wir in der letzten Inklusion wieder von der Identifikation $v = v^{**}$ Gebrauch gemacht haben.
Man sollte beachten, daß die Relation des “senkrecht Stehens” bislang nicht symmetrisch definiert ist. Für $w \in V^*$ und $v \in V$ ist bislang nur $w \perp v$ aber nicht $v \perp w$ definiert, d.h. $\perp \subseteq V^* \times V$. Betrachten wir aber wieder die Identifikation $v = v^{**}$, so können wir $v \perp w$ als $v^{**} \perp w$, d.h. durch $v^{**}(w) = 0$ definiert betrachten. Dann gilt $v \perp w \;\Leftrightarrow\; w \perp v$. Wir haben nun zu untersuchen, wie sich die orthogonalen Komplemente bezüglich dieser unterschiedlichen Definition verhalten. Dazu führen wir für $W \subseteq V^*$ die folgende Notation ein

$$^\perp W := \{v \in V \mid (\forall w \in W)[w \perp v]\}, \tag{3.5}$$

d.h. $^\perp W$ ist das orthogonale Komplement bezüglich der eben in Erwägung gezogenen Relation $\perp \subseteq V \times V^*$.

3.2.14 Lemma *Für $W \subseteq V^*$ ist ${}^\perp W \subseteq W^\perp$. Ist $\dim(V) < \infty$, so ist ${}^\perp W = W^\perp$.*

Beweis: Aus $v \in {}^\perp W$ folgt $w \perp v$, d.h. $w(v) = 0$ für alle $w \in W$. Wegen $0 = w(v) = v^{**}(w)$ folgt dann sofort $v \in W^\perp$, wobei wir die Identifikation $v^{**} = v$ benutzt haben. Ist V endlich dimensional, so ist ** surjektiv, und jedes Element von $W^\perp$ hat die Gestalt v^{**} für ein $v \in V$. Damit folgt $0 = v^{**}(w) = w(v)$ für alle $w \in W$ und damit auch $v = v^{**} \in {}^\perp W$. □

3.2.15 Lemma *Für $U \subseteq V$ ist $U^\perp$ ein Teilraum von V^*. Es gilt $U \subseteq U^{\perp\perp}$.*

Beweis: Für $v, w \in U^\perp$ gilt $v(u) = w(u) = 0$ für alle $v \in U$. Damit folgt $(\alpha v + \beta w)(u) = \alpha v(u) + \beta w(u) = 0$ für alle $u \in U$. Also ist $U^\perp$ ein Teilraum von V^*.
Für $u \in U$ gilt $u^{**}(v) = v(u) = 0$ für alle $v \in U^\perp$. Damit folgt $u = u^{**} \in U^{\perp\perp}$ und wir haben $U \subseteq U^{\perp\perp}$. □

3.2.16 Satz *Für $f \in \mathrm{Hom}(V, W)$ gelten*

(i) Kern $f =^\perp$ (Im f^*)

und

(ii) Kern $f^* =$ (Im $f)^\perp$.

Beweis: Wir erhalten (i) wegen

$$\begin{aligned} v \in \text{Kern } f &\Leftrightarrow f(v) = 0 \\ &\Leftrightarrow (\forall w^* \in W^*)[w^*(f(v)) = 0] \\ &\Leftrightarrow (\forall w^* \in W^*)[(f^*(w^*))(v) = 0] \\ &\Leftrightarrow (\forall w^* \in W^*)[f^*(w^*) \perp v] \\ &\Leftrightarrow (\forall u \in \text{Im } f^*)[u \perp v] \\ &\Leftrightarrow v \in^\perp \text{Im } (f^*). \end{aligned}$$

Ähnlich erhalten wir (ii) wegen

$$\begin{aligned} v \in \text{Kern } f^* &\Leftrightarrow f^*(v) = 0 \\ &\Leftrightarrow v \circ f = 0 \\ &\Leftrightarrow (\forall x \in V)[v(f(x)) = 0] \\ &\Leftrightarrow (\forall x \in V)[v \perp f(x)] \\ &\Leftrightarrow (\forall u \in \text{Im } f)[v \perp u] \\ &\Leftrightarrow v \in \text{Im } f^\perp. \end{aligned}$$

3.2.17 Satz *Ist* $\dim(V) < \infty$ *und* U *ein Teilraum von* V*, so gilt* $\dim(V) = \dim(U) + \dim(U^\perp)$ *und* $U^{\perp\perp} = U$.

Beweis: Ist U ein Teilraum von V, so erhalten wir ein Komplement W mit $V = U \oplus W$ und wählen eine Basis $B_1 \cup B_2$ mit $B_1 := (b_1, \ldots, b_m)$ und $B_2 := (b_{m+1}, \ldots, b_n)$, so daß B_1 eine Basis von U und B_2 eine Basis von W ist. Schon im Beweis von Satz 2.4.11 haben wir gesehen, daß die Abbildung $u + w \longmapsto w$ linear ist. Dann ist aber auch die Abbildung

$$\begin{aligned} f\colon V &\longrightarrow V^* \\ u + v &\longmapsto v^* \end{aligned}$$

als Komposition linearer Abbildungen linear mit Kern $f = U$. Nach Satz 2.4.12 gilt damit

$$\dim V = \dim U + \dim(\text{Im } f). \tag{i}$$

Ist nun $u \in U$ und $w^* \in \text{Im } f$, so ist $u = \sum_{i=1}^m \alpha_i b_i$ und $w^* = \sum_{i=m+1}^n \alpha_i b_i^*$. Wegen $b_i^*(b_j) = \delta_{ij}$ folgt dann $w^*(u) = 0$ und damit

$$\text{Im } f \subseteq U^\perp. \tag{ii}$$

Ist $v^* \in U^\perp$, so finden wir $u = \sum_{i=1}^m \alpha_i b_i \in U$ und $w = \sum_{i=m+1}^n \alpha_i b_i \in W$, so daß $v^* = u^* + w^*$ ist. Wegen $v^*(b_j) = 0$ für $j = 1, \ldots, m$ folgt 0$= (u^* + w^*)(b_j) = \sum_{i=1}^m \alpha_i b_i^*(b_j) + \sum_{i=m+1}^n \alpha_i b_i^*(b_j) = \alpha_j$ für $j = 1, \ldots m$. Damit ist $u = 0$ und somit $v^* = w^* \in \text{Im } f$, und wir erhalten

$$U^\perp \subseteq \text{Im } f. \tag{iii}$$

Aus (i),(ii) und (iii) folgt dann

$$\dim V = \dim U + \dim U^\perp. \tag{iv}$$

Wegen

$$\dim V = \dim U + \dim U^\perp \tag{v}$$

und

$$\dim V^* = \dim U^\perp + \dim U^{\perp\perp} \tag{vi}$$

erhalten wir

$$\dim U^{\perp\perp} = \dim V - \dim U^\perp = \dim U. \tag{vii}$$

Nach Lemma 3.2.15 haben wir $U \subseteq U^{\perp\perp}$. Dann ist U bereits ein Teilraum von $U^{\perp\perp}$, und zusammen mit (vii) folgt nach Satz 1.2.27 $U = U^{\perp\perp}$. □

3.3 Basistransformationen

3.3.1 Definition Sei $\mathfrak{A} \in K^{m\times n}$. Die Matrix $\mathfrak{A}$ kann in zweifacher Weise als eine Abbildung aufgefaßt werden

$$\begin{aligned} \overline{\mathfrak{A}}\colon K^m &\longrightarrow K^n \\ \mathfrak{x} &\longmapsto \mathfrak{x}\cdot\mathfrak{A} \end{aligned}$$

$$\begin{aligned} \widehat{\mathfrak{A}}\colon K_n &\longrightarrow K_m \\ \mathfrak{x} &\longmapsto \mathfrak{A}\cdot\mathfrak{x}. \end{aligned}$$

Wir erinnern an die in Satz 1.3.1 eingeführte kanonische Basis des Vektorraums K^n. In (1.11) definierten wir

$$\mathbf{e}_i^n = (\delta_{i1},\dots,\delta_{in}).$$

Damit ist $\mathbf{e}_i^n$ das n-tupel, das aus lauter Nullen besteht bis auf die Komponente i, an der eine 1 steht. Die kanonische Basis des K^n notierten wir als

$$\mathbb{E}^n = (\mathbf{e}_1^n,\dots,\mathbf{e}_n^n).$$

Dual dazu definieren wir

$$\mathbb{E}_n = (\mathbf{e}_1^{n\,t},\dots,\mathbf{e}_n^{n\,t}) =: \mathbb{E}^{n\,t} \tag{3.6}$$

als die kanonische Basis des K_n. Damit ist die kanonische Basis des K^n gegeben als die folgende Menge von Zeilenvektoren

$$\mathbb{E}^n = \{(1,0,\dots,0),(0,1,\dots,0),\dots,(0,\dots,0,1)\}$$

und die des K_n als die folgende Menge von Spaltenvektoren

$$\mathbb{E}_n = \left\{\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\dots,\begin{pmatrix}0\\\vdots\\1\\0\end{pmatrix},\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}\right\}.$$

Wenn aus dem Zusammenhang klar ist, zu welchem Raum die kanonischen Basisvektoren gehören, lassen wir den oberen Index weg.

3.3.2 Lemma *Für die darstellenden Matrizen der Abbildungen $\overline{\mathfrak{A}}$ und $\widehat{\mathfrak{A}}$ gilt $\mathfrak{D}_{\mathbb{E}^m,\mathbb{E}^n}(\overline{\mathfrak{A}}) = \mathfrak{A}$ und $\mathfrak{D}_{\mathbb{E}_n,\mathbb{E}_m}(\widehat{\mathfrak{A}}) = \mathfrak{A}^t$.*

Beweis: Sei $\mathfrak{A} = (\alpha_{ij})_{m,n}$. Dann folgt $\overline{\mathfrak{A}}(\mathbf{e}_i^n) = \mathbf{e}_i^n\cdot\mathfrak{A} = (\alpha_{i1},\dots,\alpha_{in}) = \sum_{j=1}^n \alpha_{ij}\mathbf{e}_j^n$, und damit folgt $\mathfrak{D}_{\mathbb{E}^m,\mathbb{E}^n}(\overline{\mathfrak{A}}) = \mathfrak{A}$.

Sei $\mathfrak{D}_{\mathbb{E}_n,\mathbb{E}_m}(\widehat{\mathfrak{A}}) = (\beta_{ij})_{n,m}$, d.h. $\widehat{\mathfrak{A}}(\mathbf{e}_i^t) = \sum_{j=1}^m \beta_{ij}\mathbf{e}_j^t$. Wegen $\widehat{\mathfrak{A}}(\mathbf{e}_i^t) = \mathfrak{A}\cdot\mathbf{e}_i^t = \begin{pmatrix}\alpha_{1i}\\ \vdots \\ \alpha_{mi}\end{pmatrix} = \sum_{j=1}^m \alpha_{ji}\mathbf{e}_j^t$ erhalten wir nach Satz 1.2.16 $\beta_{ij} = \alpha_{ji}$ für $i = 1,\dots,n$ und $j = 1,\dots,m$. □

3.3.3 Beobachtung Für $\mathfrak{x} \in K_n = K^{n\times 1}$ ist $\overline{\mathfrak{x}}\colon K^n \longrightarrow K$ und damit $\overline{\mathfrak{x}} \in \mathrm{Hom}(K^n,K) = (K^n)^*$, und für $\mathfrak{x} \in K^n = K^{1\times n}$ ist $\widehat{\mathfrak{x}} \in \mathrm{Hom}(K_n,K) = {K_n}^*$. Identifizieren wir $\mathfrak{x}$ mit der Abbildung $\overline{\mathfrak{x}}$, so sehen wir, daß der K_n der zu K^n duale Raum ist. Umgekehrt erhalten wir K^n als den dualen Raum von K_n, wenn wir $\mathfrak{x}$ und $\widehat{\mathfrak{x}}$ identifizieren. Damit sind die Räume K^n und K_n paarweise zueinander dual. Um die zu $\mathbb{E}^n$ duale Basis zu finden, berechnen wir $\overline{\mathbf{e}_i^t}(\mathbf{e}_j) = \mathbf{e}_j\cdot\mathbf{e}_i^t = \delta_{ij}$. Damit ist $\mathbb{E}_n$ die zu $\mathbb{E}^n$ duale Basis. Ebenso erhalten wir wegen $\widehat{\mathbf{e}_i}(\mathbf{e}_j^t) = \mathbf{e}_i\cdot\mathbf{e}_j^t = \delta_{ij}$, daß $\mathbb{E}^n$ die zu $\mathbb{E}_n$ duale Basis ist.
Den nach Satz 3.2.7 durch die dualen Basen $\mathbb{E}^n$ und $\mathbb{E}_n$ erzeugten Isomorphismus * zwischen den dualen Räumen erhalten wir für $\mathfrak{x} \in K^n$ als $\mathfrak{x}^* = \sum_{i=1}^n \alpha_i{\mathbf{e}_i}^* = \sum_{i=1}^n \alpha_i\mathbf{e}_i^t = \mathfrak{x}^t$ und für $\mathfrak{x} \in K_n$ als $\mathfrak{x}^* = \sum_{i=1}^n \alpha_i\mathbf{e}_i^{t*} = \sum_{i=1}^n \alpha_i\mathbf{e}_i^{tt} = \mathfrak{x}^t$.
Damit sehen wir, daß der Übergang zum transponierten Vektor ein Dualisierungsprozeß ist. Das entspricht der in Satz 3.2.8 festgestellten Tatsache, daß die duale Abbildung durch die transponierte Matrix dargestellt wird.

3.3.4 Beobachtung Ist $\mathfrak{A} \in K^{m\times n}$, so ist $\overline{\mathfrak{A}}\colon K^m \longrightarrow K^n$. Für die duale Abbildung $\overline{\mathfrak{A}}^*\colon K_n \longrightarrow K_m$ gilt dann $\mathfrak{D}_{\mathbb{E}_n,\mathbb{E}_m}(\overline{\mathfrak{A}}^*) = (\mathfrak{D}_{\mathbb{E}^m,\mathbb{E}^n}(\overline{\mathfrak{A}}))^t = \mathfrak{A}^t = \mathfrak{D}_{\mathbb{E}_n,\mathbb{E}_m}(\widehat{\mathfrak{A}})$ nach Satz 3.2.8. Damit ist $\overline{\mathfrak{A}}^* = \widehat{\mathfrak{A}}$, d.h. $\widehat{\mathfrak{A}}$ ist die zu $\overline{\mathfrak{A}}$ duale Abbildung. Ebenso erhalten wir $\mathfrak{D}_{\mathbb{E}^m,\mathbb{E}^n}(\widehat{\mathfrak{A}}^*) = (\mathfrak{D}_{\mathbb{E}_n,\mathbb{E}_m}(\widehat{\mathfrak{A}}))^t = (\mathfrak{A}^t)^t = \mathfrak{A} = \mathfrak{D}_{\mathbb{E}^m,\mathbb{E}^n}(\overline{\mathfrak{A}})$ und damit $\widehat{\mathfrak{A}}^* = \overline{\mathfrak{A}}$. Die Abbildungen $\overline{\mathfrak{A}}$ und $\widehat{\mathfrak{A}}$ sind also wechselseitig dual.
Darüber hinaus beobachten wir, daß die folgenden Diagramme kommutativ sind:

$$\begin{tikzcd}
K^m \arrow[r, "\overline{\mathfrak{A}}"] \arrow[d, "t"'] & K^n \arrow[d, "t"'] \\
K_m \arrow[r, "\widehat{\mathfrak{A}^t}"'] & K_n
\end{tikzcd}
\qquad
\begin{tikzcd}
K_n \arrow[r, "\widehat{\mathfrak{A}}"] \arrow[d, "t"'] & K_m \arrow[d, "t"'] \\
K^n \arrow[r, "\overline{\mathfrak{A}^t}"'] & K_m
\end{tikzcd}$$

Die Identifikation $\mathfrak{x} = \overline{\mathfrak{x}}$ für $\mathfrak{x} \in K_n$ und $\mathfrak{x} = \widehat{\mathfrak{x}}$ für $\mathfrak{x} \in K^n$ läßt sich noch weiter begründen. Ist $\mathfrak{x} \in K_n$, so folgt $\mathfrak{x}^{**}(\widehat{\mathfrak{y}}) = \widehat{\mathfrak{y}}(\mathfrak{x}) = \mathfrak{y}\cdot\mathfrak{x} = \overline{\mathfrak{x}}(\mathfrak{y})$, d.h. $\mathfrak{x}^{**} = \overline{\mathfrak{x}}$. Für $\mathfrak{y} \in K^m$ erhalten wir analog $\mathfrak{y}^{**}(\overline{\mathfrak{x}}) = \overline{\mathfrak{x}}(\mathfrak{y}) = \mathfrak{y}\cdot\mathfrak{x} = \widehat{\mathfrak{y}}(\mathfrak{x})$ und damit $\mathfrak{y}^{**} = \widehat{\mathfrak{y}}$. Die Identifikationen $\mathfrak{x}^{**} = \mathfrak{x}, \widehat{\mathfrak{x}} = \mathfrak{x}$ und $\overline{\mathfrak{y}} = \mathfrak{y}$ sind also die gleichen.

3.3.5 Definition Sei V ein K-Vektorraum mit $\dim(V) = n$ und $B = (b_1, \ldots, b_n)$ eine Basis von V. Wir definieren (vgl.(2.2))

$$h_B : V \longrightarrow K^n \quad \sum_{i=1}^{n} \alpha_i b_i \longmapsto (\alpha_1, \ldots, \alpha_n)$$

und

$$\hat{h}_B : V \longrightarrow K_n \quad \sum_{i=1}^{n} \alpha_i b_i \longmapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

und nennen die Abbildungen h_B und $\hat{h}_B$ die *Koordinatendarstellungen* bezüglich der Basis B. Die Vektoren $h_B(v)$ und $\hat{h}_B(v)$ nennen wir die *Koordinatenvektoren* von v bezüglich der Basis B. Offensichtlich gilt $h_B(v) = \hat{h}_B(v)^t$ und $\hat{h}_B(v) = h_B(v)^t$.

Die Koordinatendarstellungen von Vektoren entsprechen der Basisdarstellungen von Abbildungen. Wie beide zusammenhängen, klärt der folgende Satz und dessen Folgerung.

3.3.6 Satz *Die Koordinatendarstellungen h_B und $\hat{h}_B$ sind bijektiv, und das folgende Diagramm kommutiert:*

$$\begin{tikzcd}
K_m \arrow[r, "\widehat{\mathfrak{D}_{C^*,B^*}(f^*)}"] & K_n \\
V \arrow[u, "\hat{h}_B"] \arrow[r, "f"] \arrow[d, "h_B"'] & W \arrow[u, "\hat{h}_C"'] \arrow[d, "h_C"] \\
K^m \arrow[r, "\overline{\mathfrak{D}_{B,C}(f)}"] & K^n
\end{tikzcd}$$

Beweis: Gilt $h_B(v) = 0$ für $v = \sum_{i=1}^{n} \alpha_i b_i$, so folgt $\alpha_1 = \ldots = \alpha_n = 0$ und damit $v = 0$. Damit ist h_B injektiv und wegen $\dim V = \dim K^n$ damit auch bijektiv. Die analoge Aussage ergibt sich auch für $\hat{h}_B$. Zum Nachweis der Kommutativität des Diagrammes beginnen wir mit dem unteren Rechteck. Seien $B = (b_1, \ldots, b_m)$ und $C = (c_1, \ldots, c_n)$ Basen von V bzw. W. Für $v \in V$ mit $v = \sum_{i=1}^{m} \beta_i b_i$ und $\mathfrak{D}_{B,C}(f) = (\alpha_{ij})$ erhalten wir

$$\begin{aligned}
&h_C(f(v)) = h_C(\textstyle\sum_{i=1}^{m} \beta_i f(b_i)) = h_C(\sum_{i=1}^{m} \beta_i (\sum_{j=1}^{n} \alpha_{ij} c_j)) \\
&= h_C(\textstyle\sum_{j=1}^{n} (\sum_{i=1}^{m} \beta_i \alpha_{ij}) c_j) = (\sum_{i=1}^{m} \beta_i \alpha_{i1}, \ldots, \sum_{i=1}^{m} \beta_i \alpha_{in}) \\
&= (\beta_1, \ldots, \beta_n) \cdot (\alpha_{ij}) = h_B(v) \cdot \mathfrak{D}_{B,C}(f) = \overline{\mathfrak{D}_{B,C}(f)}(h_B(v)).
\end{aligned}$$

Für das obere Rechteck erhalten wir $\mathfrak{D}_{C^*,B^*}(f^*) = (\mathfrak{D}_{B,C}(f))^t = (\gamma_{ij})_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}}$ mit $\gamma_{ij} = \alpha_{ji}$ für $i = 1, \ldots, n$ und $j = 1, \ldots, m$. Also folgt

$$\hat{h}_C(f(v)) = \begin{pmatrix} \sum_{i=1}^{m} \beta_i \alpha_{i1} \\ \vdots \\ \sum_{i=1}^{m} \beta_i \alpha_{in} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m} \beta_i \gamma_{1i} \\ \vdots \\ \sum_{i=1}^{m} \beta_i \gamma_{ni} \end{pmatrix} = (\gamma_{ij}) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \widehat{\mathfrak{D}_{C^*,B^*}(f^*)}(\hat{h}_B(v)),$$

d.h.

$$\hat{h}_C \circ f = \widehat{\mathfrak{D}_{C^*,B^*}(f^*)} \circ \hat{h}_B.$$

□

3.3.7 Folgerung *Ist $f: V \longrightarrow W$ linear und $\mathfrak{A}$ die Basisdarstellung von f bezüglich zweier Basen B und C von V bzw. W, so gilt $h_C \circ f = \overline{\mathfrak{A}} \circ h_B$, d.h. $h_C(f(v)) = h_B(v) \cdot \mathfrak{A}$ und $\hat{h}_C \circ f = \widehat{\mathfrak{A}^t} \circ \hat{h}_B$, d.h. $\hat{h}_C(f(v)) = \mathfrak{A}^t \cdot \hat{h}_B(v)$.*

Der Beweis der Folgerung ergibt sich sofort aus Satz 3.3.6 und der Definition der Abbildungen $\overline{\mathfrak{D}_{B,C}(f)}$ und $\widehat{\mathfrak{D}_{C^*,B^*}(f^*)}$.

3.3.8 Lemma *Sind V und W endlich dimensionale K-Vektorräume, B und C Basen von V bzw. W und gilt $h_C \circ f = \overline{\mathfrak{A}} \circ h_B$ (bzw. $\hat{h}_C \circ f = \widehat{\mathfrak{A}^t} \circ \hat{h}_B$), so ist $\mathfrak{A} = \mathfrak{D}_{B,C}(f)$.*

Beweis: Nach Satz 3.3.6 und Voraussetzung gilt $h_C \circ f = \overline{\mathfrak{A}} \circ h_B = \overline{\mathfrak{D}_{B,C}(f)} \circ h_B$. Wegen der Bijektivität von h_B folgt daraus $\overline{\mathfrak{A}} = \overline{\mathfrak{D}_{B,C}(f)}$ und damit nach Lemma 3.3.2 $\mathfrak{A} = \mathfrak{D}_{B,C}(f)$. Analog zeigt man dann die zweite Behauptung. □

3.3.9 Definition Sei V ein K-Vektorraum mit $\dim(V) = n < \infty$, seien $B = \{b_1, \ldots, b_n\}$ und $C = \{c_1, \ldots, c_n\}$ Basen von V. Dann gilt $b_i = \sum_{j=1}^{n} \alpha_{ij} \cdot c_j$. Wir definieren $\mathfrak{T}_{B,C} := (\alpha_{ij})_{\substack{i=1,\ldots,n\\ j=1,\ldots,n}}$ und nennen $\mathfrak{T}_{B,C}$ die *Transformationsmatrix* oder *Übergangsmatrix* von B nach C.

Beachte Nach der Definition von $\mathfrak{T}_{B,C}$ haben wir $\mathfrak{T}_{B,C} = \mathfrak{D}_{B,C}(id_V)$.

3.3.10 Satz *Es gelten*

(i) $h_C = \overline{\mathfrak{T}_{B,C}} \circ h_B$

und

(ii) $\hat{h}_C = \widehat{\mathfrak{T}^t_{B,C}} \circ \hat{h}_B$.

Beweis: (i): Sei $\mathfrak{T}_{B,C} = (\alpha_{ij})$, und $v = \sum_{i=1}^n \beta_i b_i$. Dann folgt $v = \sum_{i=1}^n \beta_i \sum_{j=1}^n \alpha_{ij} c_j = \sum_{j=1}^n (\sum_{i=1}^n \beta_i \alpha_{ij}) c_j$. Damit ist $h_C(v) = (\sum_{i=1}^n \beta_i \alpha_{i1}, \ldots, \sum_{i=1}^n \beta_i \alpha_{in}) = h_B(v) \cdot \mathfrak{T}_{B,C} = (\overline{\mathfrak{T}_{B,C}} \circ h_B)(v)$.

(ii): Es ist $\hat{h}_C(v) = \begin{pmatrix} \sum_{i=1}^n \beta_i \alpha_{i1} \\ \vdots \\ \sum_{i=1}^n \beta_i \alpha_{in} \end{pmatrix} = \mathfrak{T}^t_{B,C} \cdot \hat{h}_B(v) = (\widehat{\mathfrak{T}^t_{B,C}} \circ \hat{h}_B)(v)$. □

3.3.11 Korollar $\mathfrak{T}_{B,C} \in \mathrm{GL}(n, K)$.

Beweis: Die Abbildungen h_B und h_C sind beide bijektiv. Nach Satz 3.3.10 gilt dann $\overline{\mathfrak{T}_{B,C}} = h_C \circ {h_B}^{-1}$. Nach Lemma 0.5.7 ist dann $\overline{\mathfrak{T}_{B,C}}$ bijektiv . Da $\mathfrak{T}_{B,C}$ deren darstellende Matrix bezüglich der kanonischen Basis des K^n ist, folgt nach Satz 3.1.14 $\mathfrak{T}_{B,C} \in \mathrm{GL}(n, K)$. □

3.3.12 Bemerkung *Ist* $\dim(V) = n$, $\mathfrak{A} = (\alpha_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}} \in \mathrm{GL}(n, K)$ *und* $C = (c_1, \ldots, c_n)$ *eine Basis von* V*, so definiert* $b_i := \sum_{j=1}^n \alpha_{ij} c_j$ *eine Basis* $B = (b_1, \ldots, b_n)$ *mit* $\mathfrak{T}_{B,C} = \mathfrak{A}$.

Beweis: Es genügt zu zeigen, daß die Vektoren $(b_1, \ldots, b_n)$ linear unabhängig sind. Zunächst haben wir

$$\mathfrak{A} \in \mathrm{GL}(n, K) \wedge \mathfrak{x} \cdot \mathfrak{A} = 0 \;\Rightarrow\; \mathfrak{x} = 0, \qquad \text{(i)}$$

denn es ist $0 = (\mathfrak{x}\mathfrak{A})\mathfrak{A}^{-1} = \mathfrak{x}(\mathfrak{A}\mathfrak{A}^{-1}) = \mathfrak{x}$. Sei nun $\sum_{i=1}^n \beta_i b_i = 0$. Dann folgt $0 = \sum_{i=1}^n \beta_i b_i = \sum_{i=1}^n \beta_i (\sum_{j=1}^n \alpha_{ij} c_j) = \sum_{j=1}^n (\sum_{i=1}^n \beta_i \alpha_{ij}) c_j$. Damit folgt $\sum_{i=1}^n \beta_i \alpha_{ij} = 0$ für $j = 1, \ldots, n$, d.h. $(\beta_1, \ldots, \beta_n) \cdot \mathfrak{A} = 0$. Mit (i) erhalten wir damit $(\beta_1, \ldots, \beta_n) = 0$, d.h. $\beta_1 = \ldots = \beta_n = 0$, und $\{b_1, \ldots, b_n\}$ ist linear unabhängig. □

3.3.13 Satz *(Transformationssatz) Es seien* V *und* W *endlich-dimensionale* K*-Vektorräume,* B *und* B' *Basen von* V *sowie* C *und* C' *Basen von* W *und* $f \in \mathrm{Hom}(V, W)$*. Dann kommutiert das folgende Diagramm*

$$\begin{tikzcd}
K^m \arrow[rrr, "\overline{\mathfrak{D}_{B,C}(f)}"] \arrow[dd, "\overline{\mathfrak{T}_{B,B'}}"'] & & & K^n \arrow[dd, "\overline{\mathfrak{T}_{C,C'}}"] \\
& V \arrow[ul, "h_B"'] \arrow[dl, "h_{B'}"] \arrow[r, "f"] & W \arrow[ur, "h_C"] \arrow[dr, "h_{C'}"'] & \\
K^m \arrow[rrr, "\overline{\mathfrak{D}_{B',C'}(f)}"'] & & & K^n
\end{tikzcd}$$

D.h. es gilt

$$\overline{\mathfrak{D}_{B',C'}(f)} = h_{C'} \circ f \circ h_{B'}^{-1} = \overline{\mathfrak{T}_{C,C'}} \circ \overline{\mathfrak{D}_{B,C}(f)} \circ \overline{\mathfrak{T}_{B,B'}}^{-1}$$

und damit auch

$$\mathfrak{D}_{B',C'}(f) = \mathfrak{T}_{B,B'}^{-1} \cdot \mathfrak{D}_{B,C}(f) \cdot \mathfrak{T}_{C,C'} = \mathfrak{T}_{B',B} \cdot \mathfrak{D}_{B,C}(f) \cdot \mathfrak{T}_{C,C'}.$$

Beweis: Die Kommutativität der beiden inneren Trapeze haben wir bereits in Satz 3.3.6 gezeigt. Um auch die Kommutativität des äußeren Rechtecks zu zeigen, haben wir die folgenden Gleichungen zur Verfügung

$$h_C \circ f = \overline{\mathfrak{D}_{B,C}(f)} \circ h_B \tag{i}$$

$$h_{C'} = \overline{\mathfrak{T}_{C,C'}} \circ h_C \tag{ii}$$

und

$$h_{B'} = \overline{\mathfrak{T}_{B,B'}} \circ h_B. \tag{iii}$$

Damit erhalten wir

$$\begin{aligned} h_{C'} \circ f &= \overline{\mathfrak{T}_{C,C'}} \circ h_C \circ f \\ &= \overline{\mathfrak{T}_{C,C'}} \circ \overline{\mathfrak{D}_{B,C}(f)} \circ h_B \\ &= \overline{\mathfrak{T}_{C,C'}} \circ \overline{\mathfrak{D}_{B,C}(f)} \circ \overline{\mathfrak{T}_{B,B'}}^{-1} \circ h_{B'} \\ &= \overline{\mathfrak{T}_{B,B'}^{-1} \cdot \mathfrak{D}_{B,C}(f) \cdot \mathfrak{T}_{C,C'}} \circ h_{B'}, \end{aligned} \tag{iv}$$

wobei wir uns noch davon zu überzeugen haben, daß

$$\overline{\mathfrak{A}}^{-1} = \overline{\mathfrak{A}^{-1}} \tag{v}$$

gilt. Aus der dritten Zeile von (iv) folgt die Kommutativität des äußeren Rechtecks des Diagramms. Aus der letzten Zeile von (iv) folgt $\overline{\mathfrak{D}_{B',C'}(f)} = \overline{\mathfrak{T}_{B,B'}^{-1} \cdot \mathfrak{D}_{B,C}(f) \cdot \mathfrak{T}_{C,C'}}$ und daher $\mathfrak{D}_{B',C'}(f) = \mathfrak{T}_{B,B'}^{-1} \cdot \mathfrak{D}_{B,C}(f) \cdot \mathfrak{T}_{C,C'}$. Der guten Form halber wollen wir auch (v) nachrechnen. Es ist $\overline{\mathfrak{A}^{-1}}(\overline{\mathfrak{A}}(\mathfrak{x})) = (\mathfrak{x} \cdot \mathfrak{A}) \cdot \mathfrak{A}^{-1} = \mathfrak{x} \cdot (\mathfrak{A}\mathfrak{A}^{-1}) = \mathfrak{x} \cdot \mathfrak{E} = \mathfrak{x}$. Damit ist $\overline{\mathfrak{A}^{-1}} \circ \overline{\mathfrak{A}} = id$ und folglich $\overline{\mathfrak{A}}^{-1} = \overline{\overline{\mathfrak{A}}^{-1}}$.

$\square$

3.4 Der Rang linearer Abbildungen und der Rang einer Matrix

3.4.1 Definition Sei $f \in \text{Hom}(V, W)$. Dann definieren wir $\text{Rang}\,(f) := \dim(\text{Im}\,(f))$. Für $\mathfrak{A} \in K^{m\times n}$ sei $\text{Rang}\,(\mathfrak{A}) := \text{Rang}\,(\overline{\mathfrak{A}})$. Wir nennen $\text{Rang}\,(f)$ den *Rang* der linearen Abbildung f und $\text{Rang}\,(\mathfrak{A})$ den Rang der Matrix $\mathfrak{A}$.

3.4.2 Satz *Sei* $f \in \text{Hom}(V, W)$ *mit* $\dim V, \dim W < \infty$. *Dann gilt* $\text{Rang}\,(f) = \text{Rang}\,(f^*)$. *Insbesondere ist dann* $\text{Rang}\,(\mathfrak{A}) = \text{Rang}\,(\mathfrak{A}^t)$.

Beweis: Für $f \in \text{Hom}(V, W)$ haben wir nach Satz 2.4.12 und Satz 3.2.16

$$\begin{aligned}\dim V &= \dim(\text{Kern } f) + \dim(\text{Im } f)\\ &= \dim(\text{Im } f^{*\perp}) + \dim(\text{Im } f)\end{aligned} \tag{i}$$

und nach Satz 3.2.17 auch

$$\dim V = \dim(\text{Im } f^*) + \dim(\text{Im } f^{*\perp}). \tag{ii}$$

Aus (i) und (ii) folgt dann

$$\text{Rang } f = \dim(\text{Im } f) = \dim(\text{Im } f^*) = \text{Rang } f^*. \tag{iii}$$

Für $\mathfrak{A} \in K^{m\times n}$ und $\mathfrak{x} \in K^n$ folgt

$$\overline{\mathfrak{A}^t}(\mathfrak{x}) = \mathfrak{x}\mathfrak{A}^t = (\mathfrak{A}\mathfrak{x}^t)^t = \widehat{\mathfrak{A}}(\mathfrak{x}^t)^t. \tag{iv}$$

Aus (iv) folgt also, daß das Diagramm

$$\begin{array}{ccc} K^n & \xrightarrow{\overline{\mathfrak{A}^t}} & K^m \\ \big\downarrow{\scriptstyle t} & & \big\uparrow{\scriptstyle t} \\ K_n & \xrightarrow{\widehat{\mathfrak{A}}} & K_m \end{array} \tag{3.7}$$

kommutiert. Damit folgt $\text{Rang}\,(\mathfrak{A}^t) = \text{Rang}\,(\overline{\mathfrak{A}^t}) = \text{Rang } \widehat{\mathfrak{A}} = \text{Rang } \overline{\mathfrak{A}} = \text{Rang } \mathfrak{A}$, denn $\widehat{\mathfrak{A}}$ und $\overline{\mathfrak{A}}$ sind nach Satz 3.3.4 dual. □

3.4.3 Satz *Ist* $\mathfrak{A} = \mathfrak{D}_{B,C}(f)$ *bezüglich zweier beliebiger Basen* B *und* C *des* K^m *bzw.* K^n, *so gilt* $\text{Rang}\,(\mathfrak{A}) = \text{Rang}\,(f)$.

Beweis: Nach Satz 3.3.13 gilt

$$h_{\mathbb{E}^n} \circ f \circ {h_{\mathbb{E}^m}}^{-1} = \overline{\mathfrak{T}_{C,\mathbb{E}^n}} \circ \overline{\mathfrak{D}_{B,C}(f)} \circ \overline{\mathfrak{T}^{-1}_{B,\mathbb{E}^m}}. \tag{i}$$

Nun ist $h_{\mathbb{E}^n} = id_{K^n}$ und $h_{\mathbb{E}^m} = id_{K^m}$, und damit kommutiert das folgende Diagramm

$$\begin{array}{ccc} K^m & \xrightarrow{f} & K^n \\ \scriptstyle\overline{\mathfrak{T}_{B,\mathbb{E}^m}}\uparrow & & \uparrow\scriptstyle\overline{\mathfrak{T}_{C,\mathbb{E}^n}} \\ K^m & \xrightarrow{\overline{\mathfrak{A}}} & K^n. \end{array}$$

Da die Abbildungen $\overline{\mathfrak{T}_{B,\mathbb{E}^m}}$ und $\overline{\mathfrak{T}_{C,\mathbb{E}^n}}$ beide bijektiv sind, folgt Im $f \cong$ Im $\overline{\mathfrak{A}}$ und damit auch Rang $f =$ Rang $\mathfrak{A}$. □

3.4.4 Korollar *Ist* $\mathfrak{A} \in K^{m\times n}$ *und sind* $\mathfrak{B} \in \mathrm{GL}(m, K)$ *und* $\mathfrak{C} \in \mathrm{GL}(n, K)$*, so gilt* Rang $(\mathfrak{A}) =$ Rang $(\mathfrak{B} \cdot \mathfrak{A} \cdot \mathfrak{C})$.

Beweis: Nach Bemerkung 3.3.12 können wir $\mathfrak{B}$ und $\mathfrak{C}$ als Transformationsmatrizen auffassen. Damit stellen $\mathfrak{A}$ und $\mathfrak{B} \cdot \mathfrak{A} \cdot \mathfrak{C}$ die gleiche Abbildung bezüglich verschiedener Basen dar. Die Behauptung folgt daher aus Satz 3.4.3. □

3.4.5 Satz *Sei* $f \in \mathrm{Hom}(V, W)$. *Die Abbildung* f *ist genau dann injektiv, wenn* Rang $(f) = \dim(V)$ *ist.*

Beweis: Wegen $\dim V = \dim(\text{Kern } f) + \dim(\text{Im } f)$ ist Kern f genau dann der Nullraum, wenn $\dim V = \dim(\text{Im } f) =$ Rang f ist. □

3.4.6 Korollar *Eine Matrix* $\mathfrak{A} \in K^{n\times n}$ *ist genau dann invertierbar, wenn* Rang $(\mathfrak{A}) = n$ *ist.*

Beweis: Nach Satz 3.1.14 ist $\mathfrak{A} \in \mathrm{GL}(n, K)$ genau dann, wenn $\overline{\mathfrak{A}}$ bijektiv ist. Nach Satz 2.3.9 ist aber die Abbildung $\overline{\mathfrak{A}}$ genau dann bijektiv, wenn sie injektiv ist. Nach Satz 3.4.5 ist dies genau dann der Fall, wenn Rang $\mathfrak{A} = \dim K^n = n$ ist. □

3.5 Elementare Umformungen

3.5.1 Normalformensatz *Sei* $f \in \mathrm{Hom}(V, W)$ *mit* $\dim(V) = m < \infty$ *und* $\dim(W) = n < \infty$. *Dann gibt es Basen* B *von* V *und* C *von* W*, so daß*

$$\mathfrak{D}_{B,C}(f) = \begin{pmatrix} \mathfrak{E}^{r,r} & 0 \\ 0 & 0 \end{pmatrix} \text{ mit } r = \text{Rang}\,(f)$$

ist.

Beweis: Nach Satz 2.4.5 erhalten wir einen Unterraum U von V, so daß $V = \text{Kern } f \oplus U$ ist. Wegen $\dim U = \dim V - \dim(\text{Kern } f) = \dim(\text{Im } f) = r$ erhalten wir eine Basis $(b_1, \ldots, b_r, b_{r+1}, \ldots, b_n)$ von V, so daß $(b_1, \ldots, b_r)$ eine Basis von U und $(b_{r+1}, \ldots, b_n)$ eine Basis von Kern f ist. Wir behaupten, daß dann $(f(b_1), \ldots, f(b_r))$ eine Basis von Im f ist. Dazu genügt es zu zeigen, daß die Menge $\{f(b_1), \ldots, f(b_r)\}$ linear unabhängig ist. Sei also $0 = \sum_{i=1}^r \alpha_i f(b_i) = f(\sum_{i=1}^r \alpha_i b_i)$. Dann ist $\sum_{i=1}^r \alpha_i b_i \in \text{Kern } f \cap U$ und daher $\sum_{i=1}^r \alpha_i b_i = 0$. Wegen der linearen Unabhängigkeit von $\{b_1, \ldots, b_r\}$ folgt daher $\alpha_1 = \ldots = \alpha_r = 0$. Damit ist$(f(b_1), \ldots, f(b_r))$ eine Basis von Im f, die wir zu einer Basis $C = (c_1, \ldots, c_n)$ von W ergänzen. D.h. wir haben $c_i = f(b_i)$ für $i = 1, \ldots, r$. Damit erhalten wir für $\mathfrak{D}_{B,C}(f) = (\alpha_{ij})$

$$f(b_i) = \sum_{j=1}^n \alpha_{ij} c_j = \begin{cases} c_i & \text{für } i = 1, \ldots, r \\ 0 & \text{für } i = r+1, \ldots, n, \end{cases}$$

woraus $\alpha_{ij} = \begin{cases} \delta_{ij} & \text{für } i = 1, \ldots, r \\ 0 & \text{für } i = r+1, \ldots, n \end{cases}$ folgt. Damit hat $\mathfrak{D}_{B,C}(f)$ die behauptete Gestalt. □

3.5.2 Definition Die elementaren Basistransformationen sind:

(I) *Addieren eines Basisvektors zu einem anderen,*

(II) *Multiplikation eines Basisvektors mit $\alpha \in K$, $\alpha \neq 0$.*

3.5.3 Lemma *Gehen C und D aus einer Basis B durch eine elementare Basistransformation des Typs (I) bzw. (II) hervor, so sind C und D wieder Basen, und es gelten für $k, l \in \{1, \ldots, n\}$*

$$\mathfrak{T}_I := \mathfrak{T}_{B,C} = \mathfrak{E}^{n,n} - \mathfrak{E}_{k,l}^{n,n}$$

und

$$\mathfrak{T}_{II} := \mathfrak{T}_{B,D} = \mathfrak{E}^{n,n} + (\frac{1}{\alpha} - 1)\mathfrak{E}_{k,k}^{n,n}.$$

Beweis: Seien $B = (b_1, \ldots, b_n)$, $C = (c_1, \ldots, c_n)$ und $D = (d_1, \ldots, d_n)$ mit

$$c_i = \begin{cases} b_i & \text{für } i \neq k \\ b_k + b_l & \text{für } i = k \end{cases} \quad \text{und} \quad d_i = \begin{cases} b_i & \text{für } i \neq k \\ \alpha \cdot b_k & \text{für } i = k, \end{cases}$$

wobei $k \neq l$ ist. Um einzusehen, daß C und D wieder Basen sind, genügt es, die lineare Unabhängigkeit von $\{c_1, \ldots, c_n\}$ und $\{d_1, \ldots, d_n\}$ zu zeigen. Ist $0 = \sum_{i=1}^n \alpha_i c_i = \sum_{\substack{i=1\\i\neq k}}^n \alpha_i b_i + \alpha_k(b_k + b_l) = \sum_{\substack{i=1\\i\neq l}}^n \alpha_i b_i + (\alpha_l + \alpha_k) b_l$, so folgt $\alpha_i = 0$

für $i \in \{1, \dots, n\} \setminus \{l\}$ und $(\alpha_l + \alpha_k) = 0$. Wegen $\alpha_k = 0$ impliziert das letztere auch $\alpha_l = 0$. Ist $0 = \sum_{i=1}^n \alpha_i d_i = \sum_{\substack{i=1\\i\neq k}}^n \alpha_i b_i + \alpha \cdot \alpha_k b_k$, so folgt $\alpha_i = 0$ für $i \in \{1, \dots, n\} \setminus \{k\}$ und $\alpha \cdot \alpha_k = 0$. Da $\alpha \neq 0$ und jeder Körper nullteilerfrei ist, erhalten wir auch $\alpha_k = 0$. Damit sind C und D Basen.
Um die Transformationsmatrix $\mathfrak{T}_{B,C}$ zu erhalten, berechnen wir

$$b_i = \sum_{j=1}^n \alpha_{ij} c_j = \sum_{\substack{j=1\\j\neq k}}^n \alpha_{ij} b_j + \alpha_{ik}(b_k + b_l) = \sum_{\substack{j=1\\j\neq l}}^n \alpha_{ij} b_j + (\alpha_{il} + \alpha_{ik}) b_l.$$

Daraus folgt zunächst $\alpha_{ij} = \delta_{ij}$ für $i \in \{1, \dots, n\} \setminus \{k\}$ und $j \in \{1, \dots, n\} \setminus \{l\}$. Wegen $b_l = c_l$ ist $\alpha_{ll} = 1$, und für $i \neq k, l$ ist $\alpha_{il} = 0$. Wegen $b_k = \alpha_{kk} b_k + (\alpha_{kl} + \alpha_{kk}) b_l$ folgt $\alpha_{kk} = 1$ und $\alpha_{kl} = -\alpha_{kk} = -1$. Für $j \neq l$, $j \neq k$ ist schließlich $\alpha_{kj} = 0$. Damit ist $\mathfrak{T}_I = \mathfrak{T}_{B,C} = (\alpha_{ij}) = \mathfrak{E}^{n,n} - \mathfrak{E}^{n,n}_{k,l}$.
Um $\mathfrak{T}_{B,D}$ zu berechnen, setzen wir

$$b_i = \sum_{j=1}^n \alpha_{ij} d_j = \sum_{\substack{j=1\\j\neq k}}^n \alpha_{ij} b_j + \alpha\alpha_{ik} b_k.$$

Daraus folgt $\alpha_{ij} = \delta_{ij}$ für $i \neq k$ oder $j \neq k$. Wegen $b_k = \alpha\alpha_{kk} b_k$ ist $\alpha_{kk} = \frac{1}{\alpha}$. Damit ist $\mathfrak{T}_{II} = \mathfrak{T}_{B,D} = (\alpha_{ij}) = \mathfrak{E}^{n,n} - \mathfrak{E}^{n,n}_{k,k} + \frac{1}{\alpha}\mathfrak{E}^{n,n}_{k,k} = \mathfrak{E}^{n,n} + (\frac{1}{\alpha} - 1)\mathfrak{E}^{n,n}_{k,k}$. □

3.5.4 Bemerkung *Es folgt aus Korollar 3.3.11, daß die Matrizen* $\mathfrak{E}^{n,n} - \mathfrak{E}^{n,n}_{k,l}$ *und* $\mathfrak{E}^{n,n} + (\frac{1}{\alpha} - 1)\mathfrak{E}^{n,n}_{k,k}$ *beide umkehrbar sind, und man rechnet sofort nach, daß für deren Inversen* $(\mathfrak{E}^{n,n} - \mathfrak{E}^{n,n}_{k,l})^{-1} = \mathfrak{E}^{n,n} + \mathfrak{E}^{n,n}_{k,l}$ *und* $(\mathfrak{E}^{n,n} + (\frac{1}{\alpha} - 1)\mathfrak{E}^{n,n}_{k,k})^{-1} = \mathfrak{E}^{n,n} + (\alpha - 1)\mathfrak{E}^{n,n}_{k,k}$ *gilt.*
Die Matrizen der Gestalt $\mathfrak{E} + \mathfrak{E}_{k,l}$ *und der Gestalt* $\mathfrak{E} + (\alpha - 1) \cdot \mathfrak{E}_{k,k}$ *nennen wir* Elementarmatrizen.

Um die Wirkung der Multiplikation von Elementarmatrizen zu studieren, stellen wir einige Rechnungen an.

3.5.5 Lemma *Schreiben wir die Matrix* $\mathfrak{A} = (\alpha_{ij})_{\substack{i=1,\dots,m\\j=1,\dots,n}}$ *in der Form* $\mathfrak{A} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$ *mit* $\mathfrak{a}_i = (\alpha_{i1}, \dots, \alpha_{in})$*, so gilt* $\mathbf{e}_i^m \cdot \mathfrak{A} = \mathfrak{a}_i$.
Dual erhalten wir für $\mathfrak{A} = (\mathfrak{b}_1, \dots, \mathfrak{b}_n)$ *mit* $\mathfrak{b}_i = \begin{pmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{mi} \end{pmatrix}$ *auch* $\mathfrak{A} \cdot (\mathbf{e}_i^n)^t = \mathfrak{b}_i$.

Beweis: Es ist $\mathbf{e}_i^m \cdot \mathfrak{A} = (\sum_{j=1}^{m} \delta_{ij}\alpha_{j1}, \dots, \sum_{j=1}^{m} \delta_{nj}\alpha_{jn}) = (\alpha_{i1}, \dots, \alpha_{in}) = \mathfrak{a}_i$. Der Beweis der dualen Behauptung ergibt sich durch Transposition. Es ist $\mathfrak{A} \cdot (\mathbf{e}_i^n)^t = (\mathbf{e}_i^n \cdot \mathfrak{A}^t)^t = (\mathfrak{a}_i)^{tt} = \mathfrak{a}_i$.

3.5.6 Lemma *Schreiben wir* $\mathfrak{A} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$*, so folgt* $(\mathfrak{E} + \mathfrak{E}_{k,l}) \cdot \mathfrak{A} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_k + \mathfrak{a}_l \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$

und $(\mathfrak{E} + (\alpha - 1)\mathfrak{E}_{k,k}) \cdot \mathfrak{A} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \alpha \cdot \mathfrak{a}_k \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$.

Für $\mathfrak{A} = (\mathfrak{b}_1, \dots, \mathfrak{b}_n)$ *gilt dual* $\mathfrak{A} \cdot (\mathfrak{E} + \mathfrak{E}_{k,l}) = (\mathfrak{b}_1, \dots, \mathfrak{b}_k + \mathfrak{b}_l, \dots, \mathfrak{b}_n)$ *und* $\mathfrak{A} \cdot (\mathfrak{E} + (\alpha - 1)\mathfrak{E}_{k,k}) = (\mathfrak{b}_1, \dots, \alpha \cdot \mathfrak{b}_k, \dots, \mathfrak{b}_n)$.

Beweis: Schreiben wir $\mathfrak{E} + \mathfrak{E}_{k,l}$ als Matrix von Zeilenvektoren, so erhalten wir $\mathfrak{E} + \mathfrak{E}_{k,l} = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_k + \mathbf{e}_l \\ \vdots \\ \mathbf{e}_m \end{pmatrix}$ und damit $(\mathfrak{E} + \mathfrak{E}_{k,l}) \cdot \mathfrak{A} = \begin{pmatrix} \mathbf{e}_1 \cdot \mathfrak{A} \\ \vdots \\ (\mathbf{e}_k + \mathbf{e}_l) \cdot \mathfrak{A} \\ \vdots \\ \mathbf{e}_m \cdot \mathfrak{A} \end{pmatrix} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_k + \mathfrak{a}_l \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$.

Analog folgt $(\mathfrak{E} + (\alpha - 1)\mathfrak{E}_{k,k}) \cdot \mathfrak{A} = \begin{pmatrix} \mathbf{e}_1 \cdot \mathfrak{A} \\ \vdots \\ \alpha\mathbf{e}_k \cdot \mathfrak{A} \\ \vdots \\ \mathbf{e}_m \cdot \mathfrak{A} \end{pmatrix} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \alpha\mathfrak{a}_k \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$. $\square$

3.5.7 Definition Die **elementaren Zeilen- und Spaltenumformungen** einer $m \times n$–Matrix sind:

(ZI) *Addition einer Zeile zu einer anderen (d.h. die Multiplikation mit der Elementarmatrix* $(\mathfrak{E} + \mathfrak{E}_{k,l})$ *von links.)*

(SI) *Addition einer Spalte zu einer anderen (d.h. die Multiplikation mit der Elementarmatrix* $(\mathfrak{E} + \mathfrak{E}_{l,k})$ *von rechts.)*

(ZII) *Multiplikation einer Zeile mit $\alpha \in K$, $\alpha \neq 0$ (d.h. die Multiplikation mit $(\mathfrak{E} + (\alpha - 1) \cdot \mathfrak{E}_{k,k})$ von links.)*

(SII) *Multiplikation einer Spalte mit $\alpha \in K$, $\alpha \neq 0$ (d.h. die Multiplikation mit $(\mathfrak{E} + (\alpha - 1) \cdot \mathfrak{E}_{k,k})$ von rechts.)*

Durch sukzessives Ausführen elementarer Zeilen- bzw. Spaltenumformungen erhalten wir Umformungen der folgenden Typen:

(ZIII) *Addition des Vielfachen einer Zeile zu einer anderen.*

(SIII) *Addition des Vielfachen einer Spalte zu einer anderen.*

(ZIV) *Vertauschen zweier Zeilen.*

$$\begin{pmatrix} \vdots \\ \mathfrak{a} \\ \vdots \\ \mathfrak{b} \\ \vdots \end{pmatrix} \xrightarrow{\text{ZI}} \begin{pmatrix} \vdots \\ \mathfrak{a} \\ \vdots \\ \mathfrak{a} + \mathfrak{b} \\ \vdots \end{pmatrix} \xrightarrow{\text{ZIII}} \begin{pmatrix} \vdots \\ \mathfrak{a} - (\mathfrak{a} + \mathfrak{b}) \\ \vdots \\ \mathfrak{a} + \mathfrak{b} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ -\mathfrak{b} \\ \vdots \\ \mathfrak{a} + \mathfrak{b} \\ \vdots \end{pmatrix} \xrightarrow{\text{ZI}}$$

$$\begin{pmatrix} \vdots \\ -\mathfrak{b} \\ \vdots \\ \mathfrak{a} + \mathfrak{b} - \mathfrak{b} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ -\mathfrak{b} \\ \vdots \\ \mathfrak{a} \\ \vdots \end{pmatrix} \xrightarrow{\text{ZII}} \begin{pmatrix} \vdots \\ \mathfrak{b} \\ \vdots \\ \mathfrak{a} \\ \vdots \end{pmatrix}.$$

(SIV) *Vertauschen zweier Spalten.*

(ZV) *Addition einer Linearkombination von r ($1 \leq r < n$) Zeilen zu einer anderen.*

$$\begin{pmatrix} \vdots \\ \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_r \\ \vdots \\ \mathfrak{b} \\ \vdots \end{pmatrix} \overset{\text{ZII}}{\rightarrow} \dots \overset{\text{ZII}}{\rightarrow} \begin{pmatrix} \vdots \\ \alpha_1 \mathfrak{a}_1 \\ \vdots \\ \alpha_r \mathfrak{a}_r \\ \vdots \\ \mathfrak{b} \\ \vdots \end{pmatrix} \overset{\text{ZIII}}{\rightarrow} \dots \overset{\text{ZIII}}{\rightarrow} \begin{pmatrix} \vdots \\ \alpha_1 \mathfrak{a}_1 \\ \vdots \\ \alpha_r \mathfrak{a}_r \\ \vdots \\ \sum_{i=1}^{r} \alpha_i \mathfrak{a}_i + \mathfrak{b} \\ \vdots \end{pmatrix}.$$

(SV) *Addition einer Linearkombination von r ($1 \leq r < n$) Spalten zu einer anderen.*

Beachte Elementare Zeilenumformungen entsprechen der Multiplikation mit einem Produkt von Elementarmatrizen von links, elementare Spaltenumformungen der Multiplikation mit einem Produkt von Elementarmatrizen von rechts. Aus ZIII bzw. SIII folgt, daß auch die zu einer Elementarmatrix inverse Matrix ein Produkt von Elementarmatrizen ist.

3.5.8 Satz *Ist $f \in \mathrm{Hom}(V, W)$ und B' eine Basis von V, die aus B durch elementare Basistransformationen entsteht, so entsteht $\mathfrak{D}_{B',C}(f)$ aus $\mathfrak{D}_{B,C}(f)$ durch elementare Zeilenumformungen.*
Dual entsteht $\mathfrak{D}_{B,C'}(f)$ aus $\mathfrak{D}_{B,C}(f)$ durch elementare Spaltenumformungen, wenn C' aus einer Basis C von W durch elementare Basistransformationen hervorgeht.

Beweis: Nach Satz 3.3.13 ist $\mathfrak{D}_{B',C}(f) = \mathfrak{T}_{B,B'}^{-1} \cdot \mathfrak{D}_{B,C}(f)$. Die Inversen der Transformationsmatrizen elementarer Basistransformationen sind aber genau die Elementarmatrizen, und Multiplikation mit Produkten von Elementarmatrizen von links entspricht elementaren Zeilenumformungen. Analog ist $\mathfrak{D}_{B,C'}(f) = \mathfrak{D}_{B,C}(f) \cdot \mathfrak{T}_{C,C'}$, und $\mathfrak{T}_{C,C'}$ ist als Inverses einer Elementarmatrix ein Produkt von Elementarmatrizen. Multiplikation mit Produkten von Elementarmatrizen von rechts entspricht aber elementaren Spaltenumformungen. □

3.5.9 Satz *Ist $r = \mathrm{Rang}\,(\mathfrak{A})$, so ist r die maximale Anzahl linear unabhängiger Zeilenvektoren – der* Zeilenrang *– von $\mathfrak{A}$. Ebenso ist r die maximale Anzahl linear unabhängiger Spaltenvektoren – der* Spaltenrang *– von $\mathfrak{A}$.*

Beweis: Für $\mathfrak{A} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$ ist $\overline{\mathfrak{A}}(\alpha_1, \ldots, \alpha_n) = (\alpha_1, \ldots, \alpha_n) \cdot \mathfrak{A} = \sum_{i=1}^m \alpha_i \mathfrak{a}_i$.
Damit ist $\mathrm{Im}\ \overline{\mathfrak{A}} = \langle \mathfrak{a}_1, \ldots, \mathfrak{a}_m \rangle$. Wegen $\dim(\mathrm{Im}\ \overline{\mathfrak{A}}) = \mathrm{Rang}\ \mathfrak{A} = r$ sind daher höchstens r Zeilenvektoren linear unabhängig. Nach Satz 3.4.2 ist aber auch $\mathrm{Rang}\ \mathfrak{A}^t = r$, und die zweite Behauptung folgt durch Übergang zur transponierten Matrix. □

3.5.10 Bemerkung *Aus Satz 3.5.9 folgt, daß eine Matrix genausoviele linear unabhängige Zeilen- wie Spaltenvektoren enthält. Der* Zeilenrang *einer Matrix ist also stets gleich ihrem* Spaltenrang.

3.5.11 Lemma *Ist $\mathfrak{A} \in K^{m\times n}$ mit $\mathrm{Rang}\,(\mathfrak{A}) = r$, dann gibt es Produkte $\mathfrak{V}$ und $\mathfrak{W}$ von Elementarmatrizen derart, daß $\mathfrak{A} \cdot \mathfrak{W} = (\mathfrak{B}, 0)$ mit $\mathfrak{B} \in K^{m\times r}$ und $\mathfrak{V} \cdot \mathfrak{A} = \begin{pmatrix} \mathfrak{C} \\ 0 \end{pmatrix}$ mit $\mathfrak{C} \in K^{r\times n}$.*

Beweis: Die elementaren Spaltenumformungen entsprechen genau der Multiplikation mit einem Produkt von Elementarmatrizen von rechts. Durch elementare Spaltenumformungen vom Typ SIV bringen wir $\mathfrak{A}$ in die Form $(\mathfrak{a}_1, \dots, \mathfrak{a}_r, \mathfrak{a}_{r+1}, \dots, \mathfrak{a}_n)$, wobei $\mathfrak{a}_1, \dots, \mathfrak{a}_r$ linear unabhängig sind und jedes $\mathfrak{a}_i$ für $i \in \{r+1, \dots, n\}$ eine Linearkombination von $\mathfrak{a}_1, \dots, \mathfrak{a}_r$ ist. Dann bringen wir mit elementaren Spaltenumformungen vom Typ SV die $\mathfrak{a}_{r+1}, \dots, \mathfrak{a}_n$ zum Verschwinden. Durch Übergang zur transponierten Matrix folgt aus der ersten Behauptung auch sofort die zweite. □

3.5.12 Satz *Zu jeder Matrix $\mathfrak{A} \in K^{m \times n}$ mit* Rang $(\mathfrak{A}) = r$ *gibt es Produkte $\mathfrak{W}$ und $\mathfrak{V}$ von Elementarmatrizen, so daß $\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{W} = \begin{pmatrix} \mathfrak{E}^{r,r} & 0 \\ 0 & 0 \end{pmatrix}$ ist.*

Beweis: Mit Lemma 3.5.11 können wir $\mathfrak{A}$ durch elementare Spalten- und Zeilenumformungen auf die Gestalt $\mathfrak{C} := \begin{pmatrix} \mathfrak{B} & 0 \\ 0 & 0 \end{pmatrix}$ mit $\mathfrak{B} \in K^{r \times r}$ und Rang $\mathfrak{B} = r$ bringen. Sei $\mathfrak{C} = (\beta_{ij})_{m,n}$. Dann sind nicht alle $\beta_{1j} = 0$. Durch Vertauschen von Spalten (SIV) erreichen wir $\beta^1_{11} \neq 0$ und durch SII $\beta^2_{11} = 1$. Durch Zeilenumformungen ZIII erhalten wir $\beta^3_{i1} = 0$ für $i = 2, \dots, m$. Durch Spaltenumformungen SIII erhalten wir $\beta^4_{1j} = 0$ für $j = 2, \dots, n$. Damit haben wir $\mathfrak{A}$ durch elementare Zeilen- und Spaltenumformungen auf die Gestalt $\mathfrak{C}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathfrak{B}_1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ gebracht, wobei $\mathfrak{B}_1 \in K^{(r-1)\times(r-1)}$ mit Rang $(\mathfrak{B}_1) = r-1$ ist. Nun wenden wir das gleiche Verfahren auf $\mathfrak{B}_1$ an und erhalten durch Iteration die gewünschte Gestalt. □

Eine Anwendung von Satz 3.5.12 ist die explizite Berechnung von Basen des Kernes und des Bildes einer Abbildung $f \in \mathrm{Hom}(K^m, K^n)$. Wir berechnen zunächst $\mathfrak{A} := \mathfrak{D}_{\mathbb{E}^m,\mathbb{E}^n}(f)$. Ist Rang $\mathfrak{A} = r$ und $(b_{r+1}, \dots, b_m)$ eine Basis des Kernes, so daß $B = (b_1, \dots, b_r, b_{r+1}, \dots, b_m)$ eine Basis des K^m und $C = (f(b_1), \dots, f(b_r), c_{r+1}, \dots, c_n)$ eine Basis des K^n ist, so ist nach dem Normalformensatz (Satz 3.5.1) $(f(b_1), \dots, f(b_r))$ eine Basis des Bildes, und es gilt $\mathfrak{D}_{B,C}(f) = \begin{pmatrix} \mathfrak{E}^{r,r} & 0 \\ 0 & 0 \end{pmatrix}$. Nach Satz 3.5.12 gibt es umkehrbare Matrizen $\mathfrak{V}$ und $\mathfrak{W}$, so daß $\begin{pmatrix} \mathfrak{E}^{r,r} & 0 \\ 0 & 0 \end{pmatrix} = \mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{W}$ ist. Nun gilt aber nach Satz 3.3.13 auch $\mathfrak{D}_{B,C}(f) = \mathfrak{T}_{B,\mathbb{E}^m} \cdot \mathfrak{D}_{\mathbb{E}^m,\mathbb{E}^n}(f) \cdot \mathfrak{T}_{\mathbb{E}^n,C}$. (Hierbei haben wir benutzt, daß $\mathfrak{T}_{B,\mathbb{E}^m} = \mathfrak{T}^{-1}_{\mathbb{E}^m,B}$ ist.) Also können wir $\mathfrak{T}_{B,\mathbb{E}^m} = \mathfrak{V}$ setzen. Kennen wir $\mathfrak{V} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_m \end{pmatrix}$ mit $\mathfrak{a}_i = (\alpha_{i1}, \dots, \alpha_{im})$, so ist $b_i = \sum_{j=1}^m \alpha_{ij} \mathbf{e}^m_j = \mathfrak{a}_i$. Damit sind $\mathfrak{a}_{r+1}, \dots, \mathfrak{a}_m$ eine Basis des Kernes und $(f(\mathfrak{a}_1), \dots, f(\mathfrak{a}_r)) = (\overline{\mathfrak{A}}(\mathfrak{a}_1), \dots, \overline{\mathfrak{A}}(\mathfrak{a}_r)) = (\mathfrak{a}_1\mathfrak{A}, \dots, \mathfrak{a}_r\mathfrak{A})$ eine Basis des Bildes. Es geht also nur darum, $\mathfrak{V}$ zu berechnen. Wegen $\mathfrak{V} = \mathfrak{V} \cdot \mathfrak{E}$ erhalten wir

$\mathfrak{B}$ aber dadurch, daß wir die gleichen Zeilenumformungen auf die Einheitsmatrix $\mathfrak{E}^{m,m}$ anwenden, die wir anzuwenden haben, um $\mathfrak{A}$ auf Normalform zu transformieren. Diese Transformation liefert uns gleichzeitig den Rang von $\mathfrak{A}$, so daß wir alle wesentlichen Daten dann von $\mathfrak{B}$ ablesen können.

In der eben skizzierten Anwendung waren nur die Zeilenumformungen von Interesse. Wir werden sehen, daß dies auch in anderen Situationen der Fall ist. Daher wollen wir uns Zeilentransformationen genauer ansehen.

3.5.13 Definition Eine Matrix $\mathfrak{A}$ ist in *Zeilenstufenform*, wenn sie die Gestalt

$$\mathfrak{A} = \begin{pmatrix} 0 & \cdots & 0 & \boxed{\alpha_{1j_1}} & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & 0 & \boxed{\alpha_{2j_2}} & \cdots & \cdots & \cdots \\ \vdots & \cdots & \cdots & \cdots & & \ddots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & & 0 & \boxed{\alpha_{r-1j_{r-1}}} & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & & 0 & \boxed{\alpha_{rj_r}} & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

hat, wobei die eingerahmten Elemente α_{ij_i} in der *Stufenlinie* $\alpha_{1j_1}, \ldots, \alpha_{rj_r}$ ungleich 0, die Elemente unterhalb der Stufenlinie 0 und die übrigen Elemente beliebig sind. Die Elemente in den Kästchen heißen *Pivots* oder *Angelpunkte* der Matrix in Zeilenstufenform.

Beachte *Da bei der Matrix $\mathfrak{A}$ in Zeilenstufenform die ersten r Zeilen linear unabhängig sind, hat $\mathfrak{A}$ den Zeilenrang r und damit gilt* Rang $\mathfrak{A} = r$.

3.5.14 Satz *Zu jeder Matrix $\mathfrak{A} \in K^{m\times n}$ gibt es ein Produkt $\mathfrak{B}$ von Elementarmatrizen, so daß $\mathfrak{B} \cdot \mathfrak{A}$ in Zeilenstufenform ist.*

Beweis: Die Zeilenstufenform wird durch das *Gaußsche Eliminationsverfahren* gefunden, das wie folgt abläuft:

Erster Schritt: *Man bestimme j_1 als den Index der ersten Spalte von $\mathfrak{A}$, die nicht nur aus Nullen besteht. Mit* ZIV *bringt man ein von 0 verschiedenes Element dieser Spalte in die erste Zeile. Dieses ist schon α_{1j_1}.*

Zweiter Schritt: *Durch $m-1$ Umformungen vom Typ* ZIII *macht man die Elemente unterhalb von α_{1j_1} zu 0.*

Iteration: *Jetzt haben wir eine Matrix der Gestalt*

$$\begin{pmatrix} 0 & \cdots & 0 & \alpha_{1j_1} & \cdots \\ 0 \cdots\cdots\cdots & & 0 & & \mathfrak{A}' \end{pmatrix}$$

und wenden den ersten und zweiten Schritt auf die unten rechts stehende Restmatrix $\mathfrak{A}' \in K^{(m-1)\times(n-j_1)}$ an. Alles, was außerhalb von $\mathfrak{A}'$ steht, bleibt dabei unverändert. Dies wiederholen wir so lange, bis die verbleibende Restmatrix $\mathfrak{A}'$ die Nullmatrix ist.

Da wir im Gaußschen Eliminationsverfahren nur Zeilenumformungen vornehmen, erhalten wir die Zeilenstufenform als $\mathfrak{B} \cdot \mathfrak{A}$, wobei $\mathfrak{B}$ ein Produkt von Elementarmatrizen ist. □

Zur Erläuterung des Gaußschen Eliminationsverfahren wollen wir ein Beispiel rechnen.

Ausgangsmatrix

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

Erster Schritt: Nach ZIV: $Z_1 \leftrightarrow Z_2$

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

Zweiter Schritt: Nach ZIII: $Z_3 + Z_1$

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

Erster Schritt': Nach ZIV: $Z_4 \leftrightarrow Z_2$

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

Zweiter Schritt': Nach ZIII: $Z_3 - 2Z_2$ und $Z_4 - 2Z_2$

$$\begin{pmatrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{pmatrix}$$

und $Z_4 - Z_3$

$$\begin{pmatrix} 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Also ist $r = 3$, $j_1 = 2$, $j_2 = 4$, $j_3 = 5$ und Rang $(\mathfrak{A}) = 3$.

3.5.15 Korollar *Jede Matrix* $\mathfrak{A} \in \mathrm{GL}(n, K)$ *ist Produkt von Elementarmatrizen.*

Beweis: Ist $\mathfrak{A} \in \mathrm{GL}(n, K)$, so ist nach Korollar 3.4.6 Rang $\mathfrak{A} = n$. Nach Satz 3.5.12 gibt es Produkte $\mathfrak{V}$ und $\mathfrak{W}$ von Elementarmatrizen, so daß $\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{W} = \mathfrak{E}^{n,n}$ ist. Damit ist $\mathfrak{A} = \mathfrak{V}^{-1} \cdot \mathfrak{W}^{-1}$. Wie wir nach Definition 3.5.7 bemerkt haben, ist aber $\mathfrak{V}^{-1} \cdot \mathfrak{W}^{-1}$ wieder ein Produkt von Elementarmatrizen. □

3.5.16 Korollar *Ist* $\mathfrak{A} \in \mathrm{GL}(n, K)$, *so kann* $\mathfrak{A}$ *bereits allein durch elementare Zeilen-* oder *Spaltenumformungen in die Einheitsmatrix überführt werden.*

Beweis: Sei $\mathfrak{A} = \prod_{i=1}^{n} \mathfrak{V}_i$ mit Elementarmatrizen $\mathfrak{V}_i$. Dann gilt $\mathfrak{E} = \mathfrak{A}^{-1} \cdot \mathfrak{A} = (\prod_{i=1}^{n} \mathfrak{V}_i)^{-1} \cdot \mathfrak{A}$. Nun ist aber $(\prod_{i=1}^{n} \mathfrak{V}_i)^{-1} = \prod_{i=1}^{n} \mathfrak{V}_{n-i+1}^{-1}$ wieder ein Produkt von Elementarmatrizen. Die Multiplikation mit Produkten von Elementarmatrizen von links entspricht aber elementaren Zeilenumformungen. Die zweite Behauptung folgt durch Übergang zur transponierten Matrix. □

3.5.17 Satz *Sei* $\mathfrak{A} \in \mathrm{GL}(n, K)$. *Man erhält* $\mathfrak{A}^{-1}$, *indem man dieselben Zeilen- und Spaltenumformungen, die* $\mathfrak{A}$ *in* $\mathfrak{E}$ *überführen, an* $\mathfrak{E}$ *vornimmt.*

Beweis: Nach Korollar 3.5.16 gibt es Elementarmatrizen $\mathfrak{V}_i$ mit $\mathfrak{E} = \prod_{i=1}^{n} \mathfrak{V}_i \cdot \mathfrak{A}$. Dann ist $\mathfrak{A}^{-1} = \prod_{i=1}^{n} \mathfrak{V}_i \cdot \mathfrak{E}$. □

3.6 Lineare Gleichungssysteme und affine Räume

3.6.1 Definition Sei K ein Körper. Ein System von m Gleichungen

$$\begin{array}{ccccccc} \alpha_{11}x_1 & + & \cdots & + & \alpha_{1n}x_n & = & \beta_1 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \cdots & + & \alpha_{mn}x_n & = & \beta_m \end{array} \tag{3.8}$$

mit $\alpha_{ij} \in K$ und $\beta_i \in K$ heißt ein *lineares Gleichungssystem* mit n Unbekannten $x_1, \ldots, x_n$.

Setzen wir $\mathfrak{A} = (\alpha_{ij})_{\substack{i=1,\ldots,m,\\ j=1,\ldots,n}}$, $\mathfrak{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $\mathfrak{b} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$, so läßt sich (3.8) schreiben als

$$\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}. \tag{3.9}$$

Wir untersuchen die folgenden Fragen:

(A) Unter welchen Bedingungen ist (3.9) lösbar, d.h. wann gibt es ein $\mathfrak{x} \in K_n$ mit $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$?

(B) Unter welchen Bedingungen ist (3.9) eindeutig lösbar, d.h. wann gibt es genau ein $\mathfrak{x} \in K_n$ mit $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$?

(C) Unter welchen Bedingungen an $\mathfrak{A}$ ist (3.9) universell lösbar, d.h. lösbar für jedes $\mathfrak{b} \in K_m$?

(D) Wie lassen sich die Lösungen von (3.9) beschreiben?

Fassen wir (3.9) in der Form $\widehat{\mathfrak{A}}(\mathfrak{x}) = \mathfrak{b}$ als Abbildung auf, so hilft uns die bislang entwickelte Theorie linearer Abbildungen bei der Beantwortung dieser Fragestellungen.

3.6.2 Beobachtung Sei $\mathfrak{A} \in K^{m\times n}$, $\mathfrak{b} \in K_m$. Das Gleichungssystem $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ ist genau dann lösbar, wenn $\mathfrak{b} \in \mathrm{Im}\,(\widehat{\mathfrak{A}})$ ist. Damit ist $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ universell lösbar, wenn $\mathrm{Im}\,(\widehat{\mathfrak{A}}) = K_m$ ist, d.h. wenn $\mathrm{Rang}\,(\mathfrak{A}) = m$ ist.

Diese Beobachtung läßt sich in zwei Sätze fassen.

3.6.3 Satz *Das Gleichungssystem $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ für $\mathfrak{A} \in K^{m\times n}$ und $\mathfrak{b} \in K_m$ ist genau dann universell lösbar, wenn $m = \mathrm{Rang}\,(\mathfrak{A}) \leq n$ ist.*

Beweis: Wie wir bereits beobachtet haben, ist das Gleichungssystem genau dann universell lösbar, wenn $\mathrm{Rang}\ \mathfrak{A} = m$ ist. Da $\widehat{\mathfrak{A}}\colon K_n \longrightarrow K_m$ ist $m = \dim(\mathrm{Im}\ \overline{\mathfrak{A}}) \leq n$. □

3.6.4 Satz *Das Gleichungssystem* $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ *für* $\mathfrak{A} \in K^{m \times n}$ *und* $\mathfrak{b} \in K_m$ *ist genau dann lösbar, wenn* Rang $(\mathfrak{A}) =$ Rang $(\mathfrak{A}, \mathfrak{b})$ *ist, wobei* $(\mathfrak{A}, \mathfrak{b}) \in K^{m \times (n+1)}$ *durch Hinzufügen der Spalte* $\mathfrak{b}$ *aus* $\mathfrak{A}$ *entsteht. Man nennt* $(\mathfrak{A}, \mathfrak{b})$ *die* geränderte Matrix *des Gleichungssystems.*

Beweis: Wie wir bereits beobachtet haben, ist das Gleichungssystem genau dann lösbar, wenn $\mathfrak{b} \in \mathrm{Im}\,(\widehat{\mathfrak{A}})$ ist. Für $\mathfrak{A} = (\mathfrak{a}_1, \ldots, \mathfrak{a}_n)$ ist aber $\mathrm{Im}\,(\widehat{\mathfrak{A}}) = \langle \mathfrak{a}_1, \ldots, \mathfrak{a}_n \rangle$. Nun ist $\mathfrak{b} \in \langle \mathfrak{a}_1, \ldots, \mathfrak{a}_n \rangle$ genau dann, wenn $\dim\langle \mathfrak{a}_1, \ldots, \mathfrak{a}_n, \mathfrak{b} \rangle = \dim\langle \mathfrak{a}_1, \ldots, \mathfrak{a}_n \rangle$ ist. Wegen Rang $(\mathfrak{A}, \mathfrak{b}) = \dim\langle \mathfrak{a}_1, \ldots, \mathfrak{a}_n, \mathfrak{b} \rangle$ folgt daraus die Behauptung. □

3.6.5 Satz *Das Gleichungssystem* $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ *für* $\mathfrak{A} \in K^{m \times n}$*, und* $\mathfrak{b} \in K_m$ *ist genau dann eindeutig lösbar, wenn* $n =$ Rang $(\mathfrak{A}) =$ Rang $(\mathfrak{A}, \mathfrak{b})$ *ist.*

Beweis: $\Rightarrow$: Ist Rang $\mathfrak{A} <$ Rang $(\mathfrak{A}, \mathfrak{b})$, so ist das Gleichungssystem nach Satz 3.6.4 nicht lösbar. Ist $n >$ Rang $\mathfrak{A} =$ Rang $(\mathfrak{A}, \mathfrak{b})$, so ist $\dim(\text{Kern } \widehat{\mathfrak{A}}) = n -$ Rang $\mathfrak{A} \neq 0$. Nach Satz 3.6.4 gibt es ein $\mathfrak{x} \in K_n$ mit $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$. Da $\dim(\text{Kern } \widehat{\mathfrak{A}}) \neq 0$ ist, gibt es ein $\mathfrak{y} \in$ Kern $\widehat{\mathfrak{A}}$ mit $\mathfrak{y} \neq 0$. Dann folgt $\mathfrak{x} \neq \mathfrak{x} + \mathfrak{y}$ und $\mathfrak{A} \cdot (\mathfrak{x} + \mathfrak{y}) = \mathfrak{A} \cdot \mathfrak{x} + \mathfrak{A} \cdot \mathfrak{y} = \mathfrak{b} + 0 = \mathfrak{b}$. Das Gleichungssystem hat daher mindestens zwei Lösungen.
$\Leftarrow$: Ist Rang $\mathfrak{A} =$ Rang $(\mathfrak{A}, \mathfrak{b})$, so ist das Gleichungssystem nach Satz 3.6.4 lösbar. Aus $\dim(\text{Kern } \widehat{\mathfrak{A}}) = n -$ Rang $\mathfrak{A} = 0$ erhalten wir Kern $\widehat{\mathfrak{A}} = \mathbf{0}$. Damit ist $\widehat{\mathfrak{A}}$ injektiv, und $\mathfrak{b}$ besitzt genau ein Urbild. □

3.6.6 Satz *Das Gleichungssystem* $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ *für* $\mathfrak{A} \in K^{m \times n}$, $\mathfrak{b} \in K_m$ *ist genau dann eindeutig und universell lösbar, wenn* Rang $(\mathfrak{A}) = m = n$ *(und damit* $\mathfrak{A} \in \mathrm{GL}(n, K)$*) ist. Die Lösung ist dann durch* $\mathfrak{x} = \mathfrak{A}^{-1} \cdot \mathfrak{b}$ *gegeben.*

Beweis: Das Gleichungssystem $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ ist genau dann universell und eindeutig lösbar, wenn $\widehat{\mathfrak{A}}$ bijektiv ist. Nach Satz 3.1.14 ist dies genau dann der Fall, wenn $\mathfrak{A}^t$ invertierbar ist. Nach Korollar 3.4.6 ist dies äquivalent dazu, daß Rang $\mathfrak{A}^t =$ Rang $\mathfrak{A} = n$ ist. Durch Multiplikation des Gleichungssystems $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ mit $\mathfrak{A}^{-1}$ von links erhalten wir die Lösung $\mathfrak{x} = \mathfrak{A}^{-1} \cdot \mathfrak{b}$. □

Damit haben wir die Fragen (A)–(C) beantwortet und wollen nun die Struktur der Lösungsmenge studieren.

3.6.7 Satz *Ist* $\mathfrak{A} \in K^{m \times n}$, $\mathfrak{b} \in K_m$ *und* $\mathfrak{x} \in \widehat{\mathfrak{A}}^{-1}(\mathfrak{b})$*, so erhalten wir die Lösungsmenge durch* $\widehat{\mathfrak{A}}^{-1}(\mathfrak{b}) = \mathfrak{x} +$ Kern $(\widehat{\mathfrak{A}})$.

Beweis: Sei $\mathfrak{x} \in \widehat{\mathfrak{A}}^{-1}(b)$.
$\supseteq$: Wir haben bereits im Beweis zu Satz 3.6.5 gesehen, daß dann für $\mathfrak{y} \in$ Kern $(\widehat{\mathfrak{A}})$ auch $\mathfrak{x} + \mathfrak{y}$ eine Lösung des Gleichungssystems ist. Also gilt $\mathfrak{x} +$ Kern $\widehat{\mathfrak{A}} \subseteq \widehat{\mathfrak{A}}^{-1}(b)$.

$\subseteq$: Für die umgekehrte Inklusion nehmen wir $\mathfrak{y} \in \widehat{\mathfrak{A}}^{-1}(b)$ an. Dann gilt $\widehat{\mathfrak{A}}(\mathfrak{y} - \mathfrak{x}) = \widehat{\mathfrak{A}}(\mathfrak{y}) - \widehat{\mathfrak{A}}(\mathfrak{x}) = \mathfrak{b} - \mathfrak{b} = 0$ und damit $\mathfrak{y} - \mathfrak{x} \in \text{Kern } \widehat{\mathfrak{A}}$. Damit ist $\mathfrak{y} = \mathfrak{x} + (\mathfrak{y} - \mathfrak{x}) \in \mathfrak{x} + \text{Kern } \widehat{\mathfrak{A}}$. □

Affine Räume und lineare Mannigfaltigkeiten

Um die Lösungsmengen linearer Gleichungssysteme genauer studieren zu können, führen wir affine Räume und lineare Mannigfaltigkeiten ein.

3.6.8 Definition Sei $M \neq \emptyset$, V ein K-Vektorraum und $\phi\colon M \times M \longrightarrow V$ eine Surjektion. Anstelle von $\phi(P,Q)$ schreiben wir $\overrightarrow{PQ}$ aus Gründen, die gleich näher erläutert werden sollen.
Das Tripel (M, ϕ, V) heißt ein *affiner Raum*, wenn die folgenden Axiome erfüllt sind

(A1) $(\forall P \in M)(\forall a \in V)(\exists Q \in M)[a = \overrightarrow{PQ}]$,

(A2) $(\forall P, Q, R \in M)[\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}]$,

(A3) $(\forall P, Q \in M)[\overrightarrow{PQ} = 0 \Rightarrow P = Q]$.

Wir nennen M die Menge der *Punkte* des affinen Raumes, $\phi(P,Q) =: \overrightarrow{PQ} \in V$ ist die *Strecke*, die P und Q verbindet. Man nennt V den *Differenzenraum* des affinen Raumes $\mathfrak{A} = (M, \phi, V)$.

In einem affinen Raum gelten weiterhin

(A4) $\overrightarrow{PP} = 0$,

denn es es ist nach (A2) $\overrightarrow{PP} = \overrightarrow{PP} + \overrightarrow{PP}$, woraus $\overrightarrow{PP} = 0$ folgt.

(A5) $\overrightarrow{PQ} = \overrightarrow{PR} \Rightarrow Q = R$,

denn es gilt nach (A2) $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR} = \overrightarrow{PQ}$, woraus $\overrightarrow{QR} = 0$ und daraus nach (A3) $Q = R$ folgt.
Es folgt aus (A1) und (A5), daß es zu $P \in M$ und $a \in V$ einen eindeutig bestimmten Punkt $Q \in M$ gibt mit $\overrightarrow{PQ} = a$. Wir schreiben $Q =: P + a$, d.h. $\overrightarrow{P, P+a} = a$. Man sagt: "$Q$ entsteht durch Abtragen des Vektors a vom Punkte P aus."

(A6) $(P + a) + b = P + (a + b)$,

denn setzen wir $R := (P + a) + b$, so erhalten wir $\overrightarrow{PR} = \overrightarrow{P P + a} + \overrightarrow{P + a R} = a + b = \overrightarrow{P P + (a + b)}$, d.h. $R = P + (a + b)$.

(A7) $\overrightarrow{PQ} = \overrightarrow{RS} \Leftrightarrow \overrightarrow{PR} = \overrightarrow{QS}$, denn nach (A2) kommutiert das folgende Diagramm:

$$\begin{array}{ccc} P & \xrightarrow{\overrightarrow{PQ}} & Q \\ {\scriptstyle\overrightarrow{PR}}\big\downarrow & & \big\downarrow{\scriptstyle\overrightarrow{QS}} \\ R & \xrightarrow{\overrightarrow{RS}} & S \end{array}$$

3.6.9 Satz *Ist V ein K-Vektorraum, U ein Teilraum von V und $v \in V$, so ist das Tripel $(v+U, \phi, U)$ mit $\phi(v+u_1, v+u_2) := u_2 - u_1$ ein affiner Raum mit Differenzenraum U.*

Beweis: Für $u \in U$ ist $\overrightarrow{v, v+u} = \phi(v, v+u) = u - 0 = u$. Damit ist ϕ surjektiv. Wir prüfen nun die einzelnen Axiome nach.
(A1) Aus $P = v + u_1 \in v + U$ und $a \in U$ folgt $\overrightarrow{v+u_1, v+u_1+a} = u_1 + a - u_1 = a$. D.h. wir können $Q = v + (u_1 + a) \in v + U$ setzen.
(A2) Sei $P = v + u_1$, $Q = v + u_2$ und $R = v + u_3$. Dann ist $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{v+u_1, v+u_2} + \overrightarrow{v+u_2, v+u_3} = u_2 - u_1 + u_3 - u_2 = u_3 - u_1 = \overrightarrow{v+u_1, v+u_3} = \overrightarrow{PR}$.
(A3) Aus $\overrightarrow{v+u_1, v+u_2} = 0$ erhalten wir $u_2 - u_1 = 0$, woraus $u_1 = u_2$ und damit auch $v + u_1 = v + u_2$ folgt. □

Es folgt aus Satz 3.6.9, daß sich jeder Vektorraum V als affiner Raum (V, ϕ, V) mit $\phi(u, v) := v - u$ auffassen läßt. Der Begriff des affinen Raumes ist also der allgemeinere. In Abbildung 3.1 haben wir den $\mathbb{R}^2$ als affinen Raum dargestellt.
Wir wollen nun zeigen, daß sich jeder affine Raum A bis auf Isomorphie als ein "um einen Vektor v parallel verschobener" Vektorraum $A \cong (v + U, \phi, U)$ darstellen läßt. Dazu müssen wir den Begriff der Isomorphie auf affine Räume ausweiten. Seien dazu $A = (M, \phi, U)$ und $A' = (M', \phi', U)$ zwei affine Räume mit gleichem Differenzenraum U. Eine Bijektion $f : M \longrightarrow M'$ heißt ein *Isomorphismus der affinen Räume*, wenn gilt:

$$\begin{aligned} &(\forall P, Q \in M)[\phi(P,Q) = \phi'(f(P), f(Q))], \quad \text{d.h.} \\ &(\forall P, Q \in M)[\overrightarrow{PQ} = \overrightarrow{f(P)f(Q)}]. \end{aligned} \tag{3.10}$$

Bemerkung *Eine Bijektion $f : M \longrightarrow M'$ ist genau dann ein Isomorphismus der affinen Räume (M, ϕ, U) und (M', ϕ', U), wenn*

$$(\forall P \in M)(\forall a \in U)[f(P + a) = f(P) + a] \tag{3.11}$$

gilt.

Beweis: $\Rightarrow$: Sei $P \in M$ und $a \in U$. Dann gilt für $Q := P + a$ bereits $a = \overrightarrow{PQ} = \overrightarrow{f(P)f(Q)}$. Also ist $f(P) + a = f(Q) = f(P + a)$.

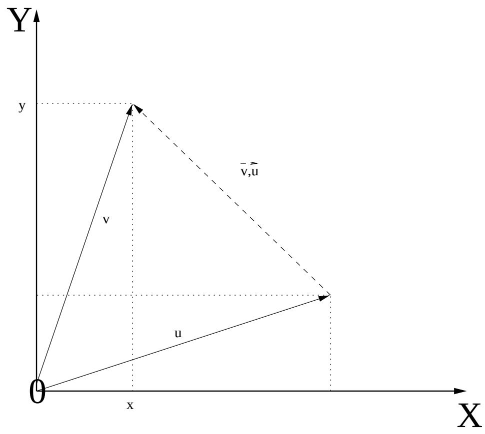

Abbildung 3.1: Der $\mathbb{R}^2$ als affiner Raum. Die Punkte des Raumes sind die Vektoren $\{(x,y) \mid x, y \in \mathbb{R}\}$, die wir hier als Endpunkte der "Zeiger" dargestellt haben. Die Strecken sind die Verbindungen dieser Endpunkte.

$\Leftarrow$: Seien $P, Q \in M$ und $a := \overrightarrow{PQ}$. Dann ist $Q = P + a$ und damit $f(Q) = f(P + a) = f(P) + a$. Damit folgt $\overrightarrow{PQ} = a = \overrightarrow{f(P)f(Q)}$. □

3.6.10 Lemma *Ist f ein Isomorphismus der affinen Räume $\mathfrak{A} = (M, \phi, U)$ und $\mathfrak{A}' = (M', \phi', U)$ und ist $S \in M$, so ist f durch das Bild $f(S)$ von S bereits eindeutig bestimmt. Umgekehrt definiert für $S \in M$ und $S' \in M'$ die Abbildung $f\colon M \longrightarrow M'$, $P \longmapsto S' + \overrightarrow{SP}$ einen Isomorphismus der affinen Räume.*

Beweis: Sei $f\colon M \longrightarrow M'$ ein Isomorphismus der affinen Räume. Zu $P \in M$ finden wir ein $a \in U$ mit $P = S + a$, und es folgt $f(P) = f(S) + a = f(S) + \overrightarrow{SP}$. Ist $g\colon M \longrightarrow M'$ ein weiterer Isomorphismus mit $f(S) = g(S)$, so folgt $g(P) = g(S) + \overrightarrow{SP} = f(S) + \overrightarrow{SP} = f(P)$. Damit ist $f = g$.
Seien nun $S \in M$ und $S' \in M'$ gegeben. Wir definieren $f(P) := S' + \overrightarrow{SP}$. Ist $P' \in M'$, so sei $a := \overrightarrow{S'P'}$. Nach (A1) finden wir ein $P \in M$ mit $\overrightarrow{SP} = a$, und es ist $f(P) = S' + a = P'$. Damit ist f surjektiv. Gilt $f(P) = f(Q)$ für $P, Q \in M$, so ist $S' + \overrightarrow{SP} = f(P) = f(Q) = S' + \overrightarrow{SQ}$. Es folgt $\overrightarrow{SP} = \overrightarrow{S'f(P)} = \overrightarrow{S'f(Q)} = \overrightarrow{SQ}$ und daraus nach (A5) $P = Q$. Also ist f auch injektiv. Schließlich folgt mit $a := \overrightarrow{PQ}$ $f(P + a) = f(Q) = S' + \overrightarrow{SQ} = S' + \overrightarrow{SP} + \overrightarrow{PQ} = f(P) + \overrightarrow{PQ} = f(P) + a$. Damit ist f ein Isomorphismus. □

3.6.11 Satz *Ist A ein affiner Raum mit Differenzenraum U, so ist $A \cong (U, \vec{\ }, U)$ mit $\overrightarrow{vu} = u - v$. Insbesondere sind affine Räume mit gleichem Differenzenraum isomorph.*

Beweis: Sei $A = (M, \phi, U)$. Wähle $S \in M$. Dann ist nach Lemma 3.6.10 die Abbildung $f\colon M \longrightarrow U$ mit $f(P) := \overrightarrow{SP}$ ein Isomorphismus. □

3.6.12 Definition Sei V ein K-Vektorraum und $M \subseteq V$, $M \neq \emptyset$. Wir nennen M eine *lineare Mannigfaltigkeit* in V, wenn gilt

$$(\forall u_1, \dots, u_m \in M)(\forall \alpha_1, \dots, \alpha_m \in K)[\sum_{i=1}^{m} \alpha_i = 1 \;\Rightarrow\; \sum_{i=1}^{m} \alpha_i u_i \in M].$$

Zu einer linearen Mannigfaltigkeit M definieren wir den Differenzenraum

$$\Delta M := \{u - v \mid u, v \in M\}.$$

3.6.13 Lemma *Ist M eine lineare Mannigfaltigkeit in V, so ist $\Delta M \neq \emptyset$, $M = v + \Delta M$ für jedes $v \in M$ und ΔM ein Teilraum von V.*

Beweis: Da $M \neq \emptyset$ ist $0 \in \Delta M$ und damit auch $\Delta M \neq \emptyset$. Sei $v \in M$. Dann läßt sich jedes $u \in M$ darstellen als $u = v + (u - v) \in v + \Delta M$. Ist umgekehrt $x \in v + \Delta M$, so gibt es $u, w \in M$ mit $x = v + (u - w)$. Da $1 + 1 - 1 = 1$ folgt daraus $x \in M$. Es bleibt zu zeigen, daß ΔM ein Teilraum von V ist. Ist $a \in \Delta M$ und $\alpha \in K$, so finden wir $u, w \in M$ mit $\alpha a = \alpha(u - w) = \alpha u - \alpha w + w - w = \alpha u + (1 - \alpha)w - w \in \Delta M$, denn wegen $\alpha + 1 - \alpha = 1$ ist $\alpha u - (1 - \alpha)w \in M$. Damit haben wir gezeigt, daß ΔM unter skalarer Multiplikation abgeschlossen ist. Etwas schwieriger gestaltet sich der Nachweis der Abgeschlossenheit gegenüber der Addition. Hier müssen wir uns den Grundkörper genauer ansehen. Unter der *Charakteristik eines Körpers K* versteht man die Anzahl der 1en, die man benötigt um 0 darzustellen, d.h.

$$\mathrm{Char}K := \begin{cases} \min\{n \in \mathbb{N} \mid 1 \cdot n = 0\} & \text{falls es so ein } n \in \mathbb{N} \text{ gibt} \\ 0 & \text{sonst,} \end{cases}$$

wobei $1 \cdot n$ als Abkürzung $\underbrace{1 + \ldots + 1}_{n-\text{fach}}$ steht. Ohne daß wir hier näher darauf eingehen wollen, sei ohne Beweis bemerkt, daß die Charakteristik eines Körpers immer eine Primzahl oder 0 ist. (So ist z.B. Char($\mathbb{Q}$) $= 0$). Alles, was wir hier benötigen, ist die Tatsache, daß ein Körper einer Charakteristik $\neq 2$ mehr als zwei Elemente besitzt. Das folgt daraus, daß ein Körper immer mindestens zwei Elemente besitzen muß, 0 und 1, die neutralen Elemente der additiven und der multiplikativen Gruppe. Wären dies die einzigen, so müßte 1 bezüglich der Addition zu sich selbst

invers sein, d.h. es müßte $1 + 1 = 0$ gelten, was im Widerspruch zu Char$K \neq 2$ stünde.
Wir nehmen nun an, daß der Grundkörper des Vektorraumes eine Charakteristik $\neq 2$ besitzt. Wir wählen nun ein $\alpha \in K$, das von 0 und 1 verschieden ist, und erhalten für $a, b \in \Delta M$ sowohl $\alpha^{-1}a \in \Delta M$ als auch $(1-\alpha)^{-1}b \in \Delta M$. Wegen $M = v + \Delta M$ erhalten wir $v + \alpha^{-1}a \in M$ und $v + (1-\alpha)^{-1}b \in M$. Damit folgt $v + a + b = \alpha(v + \alpha^{-1}a) + (1-\alpha)[v + (1-\alpha)^{-1}b] \in M$ und somit $a + b \in \Delta M$.
Ist Char$K = 2$, so ist $1 + 1 = 0$ und wegen $1 + 1 + 1 = 1$ folgt aus $v + a \in M$ und $v + b \in M$ auch $1(v+0) + 1(v+a) + 1(v+b) = (1+1)v + v + a + b = v + a + b \in M$ und damit wieder $a + b \in \Delta M$. □

3.6.14 Satz *Die linearen Mannigfaltigkeiten von V sind genau die affinen Räume der Gestalt $M = v + U$, wobei $v \in V$ und U ein Unterraum von V ist.*

Beweis: Wir haben in Lemma 3.6.13 gesehen, daß sich jede lineare Mannigfaltigkeit M auf V in der Form $M = v + \Delta M$ darstellen läßt, wobei $v \in M \subseteq V$ und ΔM ein linearer Teilraum von V ist. Nach Satz 3.6.9 ist $(M, \phi, \Delta M)$ mit $\phi(u, v) = v - u$ ein affiner Raum.
Ist umgekehrt $M = v + U$ für ein $v \in V$ und einen Teilraum $U \subseteq V$ und sind $w_1, \ldots, w_n \in U$, so erhalten wir für $\alpha_1, \ldots, \alpha_n \in K$ mit $\sum_{i=1}^n \alpha_i = 1$

$$\sum_{i=1}^n \alpha_i(v + w_i) = (\sum_{i=1}^n \alpha_i)v + \sum_{i=1}^n \alpha_i w_i = v + \sum_{i=1}^n \alpha_i w_i \in v + U = M.$$

Damit ist M eine lineare Mannigfaltigkeit. □

Ein Beispiel

Als Beispiele für lineare Mannigfaltigkeiten wollen wir *Geraden* einführen. Wir definieren

Definition Sei V ein Vektorraum und $u, v \in V$, so heißt die Menge $G_{uv} := \{v + \alpha(u - v) \mid \alpha \in K\}$ die *Gerade* durch u und v.

Zuerst wollen wir uns davon überzeugen, daß G_{uv} eine lineare Mannigfaltigkeit ist. Sind $a_1, \ldots, a_n \in G_{u,v}$, so hat jedes a_i die Gestalt $a_i = v + \beta_i(u - v)$. Sind $\alpha_1, \ldots, \alpha_n \in K$ mit $\sum_{i=1}^n \alpha_i = 1$, so folgt $\sum_{i=1}^n \alpha_i a_i = \sum_{i=1}^n \alpha_i v + (\sum_{i=1}^n \alpha_i \beta_i)(u - v) = v + \beta(u - v) \in G_{u,v}$ mit $\beta = \sum_{i=1}^n \alpha_i \beta_i$.
Andererseits können wir einsehen, daß eine lineare Mannigfaltigkeit M mit zwei Punkten u, v auch immer die Gerade $G_{u,v}$ durch diese Punkte enthält. Sind nämlich $u, v \in M$ und ist $x \in G_{u,v}$, so folgt $x = v + \alpha(u - v) = v + \alpha u - \alpha v \in M$, da $1 + \alpha - \alpha = 1$.

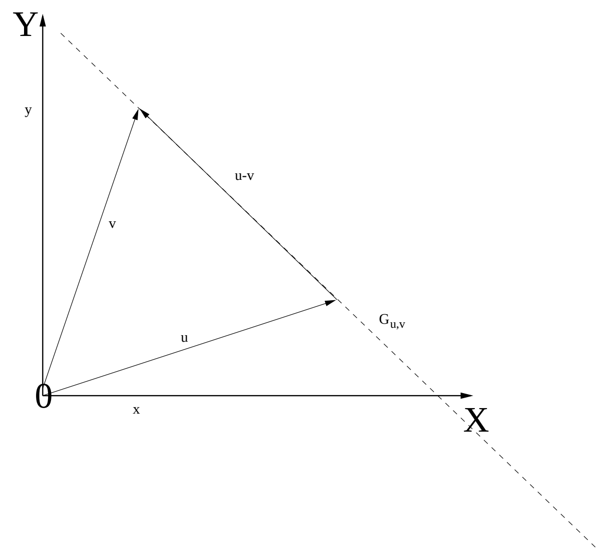

Abbildung 3.2: Eine Gerade im $\mathbb{R}^2$. Ein Vergleich mit Abbildung 3.1 zeigt, daß der $\mathbb{R}^2$ als affiner Raum aufgefaßt alle Geraden enthält.

Die Sätze 3.6.7 und 3.6.14 liefern uns nun den Zusammenhang zwischen linearen Gleichungssystemen und linearen Mannigfaltigkeiten. Fassen wir beide zusammen, so erhalten wir den folgenden Satz.

3.6.15 Satz *Ist* $\mathfrak{A} \in K^{m\times n}$ *und* $\mathfrak{b} \in K_m$*, so ist die Lösungsmenge des Gleichungssystems* $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ *eine lineare Mannigfaltigkeit im* K_n *der Gestalt* $\mathfrak{a} + \mathrm{Kern}\,(\widehat{\mathfrak{A}})$*, wobei* $\mathfrak{a}$ *eine partikuläre Lösung ist, d.h. es gilt* $\mathfrak{A} \cdot \mathfrak{a} = \mathfrak{b}$.

3.6.16 Bemerkung Wir wollen diesen Abschnitt mit einigen Bemerkungen zum praktischen Lösen linearer Gleichungssysteme abschließen. Dabei werden wir uns jedoch auf grundsätzliche Verfahren beschränken. Wirklich effiziente Verfahren, die auch mit großen Gleichungssystemen zu Rande kommen, werden Sie in der numerischen Mathematik kennenlernen.

Nach Satz 3.6.7 hat jede Lösung der Gleichung $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$, so sie existiert, die Gestalt $\mathfrak{a} + \mathfrak{u}$ mit $\mathfrak{A} \cdot \mathfrak{a} = \mathfrak{b}$, d.h. $\mathfrak{a}$ ist partikuläre Lösung und $\mathfrak{u} \in \mathrm{Kern}\,(\widehat{\mathfrak{A}})$. Um ein $\mathfrak{u} \in \mathrm{Kern}\,(\widehat{\mathfrak{A}})$ zu erhalten, müssen wir das Gleichungssystem $\widehat{\mathfrak{A}}(\mathfrak{x}) = 0$, d.h. $\mathfrak{A} \cdot \mathfrak{x} = 0$ lösen. Das Gleichungssystem $\mathfrak{A} \cdot \mathfrak{x} = 0$ wird als das zum Gleichungssystem $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ gehörende *homogene Gleichungssystem* bezeichnet. Das homogene

Gleichungssystem hat also die Form:

$$\begin{array}{ccccccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \cdots & + & \alpha_{1n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + & \cdots & + & \alpha_{mn}x_n & = & 0 \end{array}$$

Nun ist $\widehat{\mathfrak{A}}\colon K_n \longrightarrow K_m$ und damit $n = \dim K_n =$ Rang $\widehat{\mathfrak{A}} + \dim(\text{Kern } \widehat{\mathfrak{A}})$. Also ist $\dim(\text{Kern } \widehat{\mathfrak{A}}) = n - r$. Um eine Basis von Kern $\widehat{\mathfrak{A}}$ zu erhalten, müssen wir daher $n - r$ linear unabhängige Vektoren finden, die das homogene Gleichungssystem lösen. Haben wir so eine Basis $(b_1, \ldots, b_{n-r})$ gefunden, erhalten wir die allgemeine Lösung in der Form

$$\mathfrak{a} + \sum_{i=1}^{n-r} \alpha_i b_i, \tag{3.12}$$

wobei α_i alle Körperelemente durchlaufen und $\mathfrak{a}$ eine partikuläre Lösung ist.
Um einen Algorithmus zur Lösung von Gleichungssystemen zu erhalten, erinnern wir uns an die Elementaroperationen. Nach Satz 3.5.14 gibt es ein Produkt $\mathfrak{V}$ von Elementarmatrizen, so daß $\mathfrak{V} \cdot \mathfrak{A}$ Zeilenstufenform hat. Das Gleichungssystem lautet dann

$$\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{x} = \mathfrak{V} \cdot \mathfrak{b}, \tag{3.13}$$

woraus folgt, daß jede Lösung des Gleichungssystems (3.13) auch eine Lösung des ursprünglichen Gleichungssystems ist. In der Matrix $\mathfrak{A}' := \mathfrak{V} \cdot \mathfrak{A}$, die sich laut Satz 3.5.14 aus $\mathfrak{A}$ durch das Gaußsche Eliminationsverfahren ergibt, sind die Zeilen $\mathfrak{a}'_{r+1}, \ldots, \mathfrak{a}'_m$ alle 0. Daher müssen die Komponenten $\beta'_{r+1}, \ldots, \beta'_m$ des Vektors $\mathfrak{b}' := \mathfrak{V} \cdot \mathfrak{b}$ alle verschwinden. Den Vektor $\mathfrak{b}'$ erhält man am besten, indem man das Gaußsche Eliminationsverfahren auf die geänderte Matrix $(\mathfrak{A}, \mathfrak{b})$ anwendet. Der Vektor in der $n + 1$-ten Spalte ist dann $\mathfrak{b}'$. Das Gleichungssystem $\mathfrak{A}' \cdot \mathfrak{x} = \mathfrak{b}'$ hat nun die folgende Gestalt

$$\begin{array}{ccccccc} x_{i_1} + & \alpha'_{1(i_1+1)}x_{i_1+1} & + & \cdots\cdots\cdots\cdots & + & \alpha'_{1n}x_n & = & \beta'_1 \\ & x_{i_2} & + & \alpha'_{2i_2+1}x_{i_2+1} + \cdots\cdots & + & \alpha'_{2n}x_n & = & \beta'_2 \\ & & \ddots & & & \vdots & & \vdots \\ & & & x_{i_r} + \alpha'_{r(i_r+1)}x_{i_r+1} + \cdots & + & \alpha'_{rn}x_n & = & \beta'_r \end{array} . \tag{3.14}$$

Wir wählen für die Komponenten x_j mit $j \notin \{i_1, \ldots, i_r\}$ beliebige Elemente $\xi_j \in K$ und setzen diese in die letzte Zeile von (3.14) ein und lösen diese nach x_{i_r} auf. Dies liefert einen Wert ξ_{i_r}, und wir setzen die gewählten Werte ξ_j und ξ_{i_r} in die vorletzte Zeile von (3.14) ein und bekommen so einen Wert $\xi_{i_{r-1}}$ für $x_{i_{r-1}}$. Dieses Verfahren setzen wir fort und erhalten nach r-vielen Schritten einen Vektor

$(\xi_1, \dots, \xi_n)$, der eine partikuläre Lösung des Gleichungssystems (3.14) und damit auch des ursprünglichen Gleichungssystems ist. Eine Lösung des homogenen Gleichungssystems

$$\begin{array}{llllllll} x_{i_1}+ & \alpha'_{1(i_1+1)}x_{i_1+1} & + & \cdots\cdots\cdots & + & \alpha'_{1n}x_n & = & 0 \\ & x_{i_2} & + & \alpha'_{2i_2+1}x_{i_2+1} + \cdots\cdots & + & \alpha'_{2n}x_n & = & 0 \\ & & \ddots & & & \vdots & & \vdots \\ & & & x_{i_r} + \alpha'_{r(i_r+1)}x_{i_r+1} + \cdots & + & \alpha'_{rn}x_n & = & 0 \end{array} \tag{3.15}$$

erhalten wir durch das gleiche Verfahren. Da wir $n-r$ viele Komponenten des Lösungsvektors frei wählen können, erreichen wir, beispielsweise indem wir die kanonischen Basisvektoren des K_{n-r} auswählen, daß wir $n-r$ linear unabhängige Lösungen erhalten und können damit die allgemeine Lösung des Gleichungssystems darstellen. Zur weiteren Erläuterung des Verfahrens wollen wir ein einfaches Beispiel rechnen.
Wir betrachten das Gleichungssystem

$$\begin{array}{rcrcrcrcrcr} x_1 & + & x_2 & + & x_3 & - & x_4 & - & x_5 & = & 2 \\ -x_1 & + & 3x_2 & + & x_3 & - & x_4 & + & x_5 & = & 0 \\ 2x_1 & + & 6x_2 & + & x_3 & - & 2x_4 & - & x_5 & = & 3. \end{array} \tag{3.16}$$

Die geränderte Matrix lautet dann

$$\begin{pmatrix} 1 & 1 & 1 & -1 & -1 & 2 \\ -1 & 3 & 1 & -1 & 1 & 0 \\ 2 & 6 & 1 & -2 & -1 & 3 \end{pmatrix}, \tag{3.17}$$

die sich zunächst auf

$$\begin{pmatrix} 1 & 1 & 1 & -1 & -1 & 2 \\ 0 & 4 & 2 & -2 & 0 & 2 \\ 0 & 4 & -1 & 0 & 1 & -1 \end{pmatrix} \tag{3.18}$$

und dann auf die Form

$$\begin{pmatrix} 1 & 1 & 1 & -1 & -1 & 2 \\ 0 & 4 & 2 & -2 & 0 & 2 \\ 0 & 0 & -3 & 2 & 1 & -3 \end{pmatrix} \tag{3.19}$$

bringen läßt. Dies ist bereits die Zeilenstufenform. Die geränderte Matrix sowie die Koeffizientenmatrix haben den Rang 3. Damit ist das Gleichungssystem lösbar.
Lösen wir die unterste Zeile nach x_3 auf, so erhalten wir

$$3x_3 = 3 + 2x_4 + x_5. \tag{3.20}$$

Nun wählen wir $x_4 = x_5 = 1$ und erhalten aus (3.20) $x_3 = 2$. Damit ergibt sich aus der mittleren Zeile von (3.19)

$$2 = 4x_2 + 2x_3 - 2x_4 = 4x_2 + 2$$

und damit $x_2 = 0$. Schließlich erhalten wir aus der ersten Zeile von (3.19)

$$2 = x_1 + x_2 + x_3 - x_4 - x_5 = x_1 + 0 + 2 - 1 - 1 = x_1.$$

Damit haben wir eine partikuläre Lösung $\mathfrak{x} = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 1 \\ 1 \end{pmatrix}$.

Um die allgemeine Lösung zu erhalten, lösen wir das homogene Gleichungssystem

$$\begin{pmatrix} 1 & 1 & 1 & -1 & -1 \\ 0 & 4 & 2 & -2 & 0 \\ 0 & 0 & -3 & 2 & 1 \end{pmatrix} \cdot \mathfrak{x} = \begin{pmatrix} 0 \\ 0 \\ 0. \end{pmatrix} \tag{3.21}$$

Wir beginnen wieder mit der letzten Zeile von (3.21) und haben

$$-3x_3 + 2x_4 + x_5 = 0 \tag{3.22}$$

zu lösen. Wir setzen $x_5 = 0$ und $x_4 = 3$ und erhalten daraus $x_3 = 2$. Aus der mittleren Zeile von (3.21) folgt

$$4x_2 + 2x_3 - 2x_4 = 0, \quad \text{d.h.} \tag{3.23}$$

$$4x_2 + 4 - 6 = 0.$$

Damit ist $x_2 = \frac{1}{2}$. Aus der ersten Zeile von (3.21) erhalten wir

$$x_1 + x_2 + x_3 - x_4 - x_5 = 0, \quad \text{d.h.} \tag{3.24}$$

$$x_1 + \frac{1}{2} + 2 - 3 = 0,$$

woraus sich $x_1 = \frac{1}{2}$ ergibt. Damit haben wir einen Lösungsvektor $\mathfrak{y} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 2 \\ 3 \\ 0 \end{pmatrix}$.

Nun ist $\dim(\text{Kern } \widehat{\mathfrak{A}}) = n - r = 5 - 3 = 2$, und wir haben eine weitere linear unabhängige Lösung von (3.21) zu finden. Dazu setzen wir in (3.22) $x_4 = 0$ und $x_5 = 3$. Dann ergibt sich $x_3 = 1$. Setzen wir dies in (3.23) ein, so erhalten wir $x_2 = -\frac{1}{2}$. Dies in (3.24) eingesetzt ergibt $x_1 = \frac{5}{2}$. Damit erhalten wir einen weiteren

Lösungsvektor $\mathfrak{z} = \begin{pmatrix} \frac{5}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \\ 3 \end{pmatrix}$, der offensichtlich von $\mathfrak{y}$ linear unabhängig ist. Damit ergibt sich die Lösungsmenge des Gleichungssystems (3.16)

$$\widehat{\mathfrak{A}}^{-1}(b) = \{(2,0,2,1,1)^t + \alpha(\frac{5}{2}, -\frac{1}{2}, 1, 0, 3)^t + \beta(\frac{1}{2}, \frac{1}{2}, 2, 3, 0)^t \mid \alpha, \beta \in \mathbb{R}\}.$$

Ist die Koeffizientenmatrix $\mathfrak{A}$ des Gleichungssystems $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ invertierbar, so läßt sich nach Korollar 3.5.16 $\mathfrak{A}$ schon durch Zeilenumformungen allein in die Einheitsmatrix überführen. Die Zeilenstufenform von A ist also bereits die Einheitsmatrix. Die Umformung der geänderten Matrix liefert also eine Matrix der Form

$$\begin{pmatrix} 1 & 0 \cdots\cdots & 0 & \beta_1' \\ 0 & 1 \cdots\cdots & 0 & \beta_2' \\ \vdots & \ddots & \vdots & \vdots \\ 0 \cdots\cdots & 1 & 0 & \beta_{n-1}' \\ 0 \cdots\cdots & 0 & 1 & \beta_n'. \end{pmatrix} \tag{3.25}$$

Wir können nun den eindeutig bestimmten Lösungsvektor $\mathfrak{x} = (\beta_1', \ldots, \beta_n')^t$ direkt aus (3.25) ablesen.

Läßt es sich im allgemeinen Falle einer Matrix $\mathfrak{A} \in K^{m \times n}$ erreichen, daß die Umformung der geänderten Matrix eine Matrix der Gestalt

$$(\mathfrak{A}', \mathfrak{b}') := \begin{pmatrix} 1 & 0 \cdots\cdots & 0 & \alpha_{1,r+1}' & \cdots & \alpha_{1,n}' & \beta_1' \\ 0 & 1 \cdots\cdots & 0 & \alpha_{2,r+1}' & \cdots & \alpha_{2,n}' & \beta_2' \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \\ 0 \cdots\cdots & 1 & 0 & \alpha_{r-1,r+1}' & \cdots & \alpha_{r-1,n}' & \beta_{r-1}' \\ 0 \cdots\cdots & 0 & 1 & \alpha_{r,r+1}' & \cdots & \alpha_{r,n}' & \beta_r' \\ 0 \cdots\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 \cdots\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix} \tag{3.26}$$

ergibt, so können wir sofort die allgemeine Lösung aus (3.26) ablesen. Setzen wir nämlich

$$\mathfrak{k}_k := \begin{pmatrix} -\alpha'_{1,r+k} \\ \vdots \\ -\alpha'_{r,r+k} \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K_n \qquad \text{mit der 1 in der } r+k\text{-ten Zeile,}$$

so folgt $\mathfrak{A}' \cdot \mathfrak{k}_k = \begin{pmatrix} -\alpha_{1,r+k} + \alpha_{1,r+k} \\ \vdots \\ -\alpha_{r,r+k} + \alpha_{r,r+k} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$, d.h. die Vektoren $\mathfrak{k}_1, \ldots, \mathfrak{k}_{n-r}$ spannen Kern $\widehat{\mathfrak{A}}$ auf. Wir sehen aber auch, daß $\mathfrak{a} = (\beta'_1, \ldots, \beta'_r, 0, \ldots, 0)^t \in K_n$ eine partikuläre Lösung ist. Bezeichnen wir nämlich das Produkt von Elementarmatrizen, das den Zeilenumformungen entspricht, die $\mathfrak{A}$ auf die Gestalt $\mathfrak{A}'$ gebracht haben, mit $\mathfrak{V}$, so folgt $\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{a} = \mathfrak{A}' \cdot \mathfrak{a} = \mathfrak{b}' = \mathfrak{V} \cdot \mathfrak{b}$ und damit $\mathfrak{A} \cdot \mathfrak{a} = \mathfrak{b}$. Damit erhalten wir die Lösungsmenge als $\{\mathfrak{a} + \sum_{i=1}^{n-r} \mu_k \mathfrak{k}_k \mid \mu_k \in \mathbb{R}\}$.

Leider läßt sich im allgemeinen Fall die Form in (3.26) nicht durch Zeilenumformungen allein erreichen. Was sich aber erreichen läßt, ist eine Transformation der Form $\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{W}$, die die Gestalt in (3.26) bewirkt. Dabei folgt aus Satz 3.5.14, daß das Produkt von Elementarmatrizen $\mathfrak{W}$ nur Spaltenvertauschungen bewirken muß. Durch Spaltenvertauschungen der Zeilenstufenmatrix können wir nämlich erreichen, daß alle Pivots in der Hauptdiagonalen stehen. Aus dieser Form erreichen wir aber die Gestalt in (3.26) durch Zeilenumformungen allein. Das Gleichungssystem $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ wird damit umgeformt zu

$$\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{W} \cdot \mathfrak{W}^{-1} \cdot \mathfrak{x} = \mathfrak{V} \cdot \mathfrak{b}. \tag{3.27}$$

Die Lösungen des Gleichungssystems

$$\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{W} \cdot \mathfrak{y} = \mathfrak{V} \cdot \mathfrak{b} \tag{3.28}$$

lassen sich wie eben beschrieben aus der Matrix $\mathfrak{V} \cdot \mathfrak{A} \cdot \mathfrak{W}$ ablesen. Dann gilt aber $\mathfrak{y} = \mathfrak{W}^{-1} \cdot \mathfrak{x}$ und damit $\mathfrak{x} = \mathfrak{W} \cdot \mathfrak{y}$. D.h. die Spaltenvertauschungen in $\mathfrak{A}$ werden

zu Zeilenvertauschungen in $\mathfrak{y}$ und Zeilenvertauschungen in $\mathfrak{y}$ bedeuten nur Umbenennung der Unbekannten. Wir erhalten also die Lösungen $\mathfrak{x}$, wenn wir in den abgelesenen Lösungen $\mathfrak{y}$ die Komponenten entsprechend den in $\mathfrak{A}$ vorgenommenen Spaltenvertauschungen vertauschen. Auch dies wollen wir an einem Beispiel klarmachen.
Sei

$$\mathfrak{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 3 \\ 4 & 3 & 2 & 1 \end{pmatrix} \quad \text{und} \quad \mathfrak{b} = \begin{pmatrix} 6 \\ 3 \\ 2 \\ 5 \\ 4 \end{pmatrix}.$$

Wir wollen das Gleichungssystem $\mathfrak{A} \cdot x = \mathfrak{b}$ lösen. Die geränderte Matrix ist dann

$$\mathfrak{A} = \begin{pmatrix} 1 & 2 & 3 & 4 & 6 \\ 0 & 1 & 2 & 2 & 3 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 3 & 5 \\ 4 & 3 & 2 & 1 & 4 \end{pmatrix}.$$

Durch Gaußelimination erhalten wir die Zeilenstufenform

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 6 \\ 0 & 1 & 2 & 2 & 3 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Wir vertauschen nun die 3te und die 4te Spalte und erhalten

$$\begin{pmatrix} 1 & 2 & 4 & 3 & 6 \\ 0 & 1 & 2 & 2 & 3 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Durch Zeilenumformungen erreichen wir die Diagonalgestalt

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

Wir lesen nun die Lösungsmenge aus dieser Matrix ab, wobei wir beachten, daß wir die 3te und 4te Komponente der Lösungsvektoren vertauschen müssen, da wir ja die 3te und 4te Spalte zum Erreichen der Diagonalgestalt vertauschen mußten. Damit erhalten wir die Lösungsmenge als

$$\{(0,1,0,1) + \mu(1,-2,1,0) \mid \mu \in \mathbb{R}\} = \{(\mu, 1-2\mu, \mu, 1) \mid \mu \in \mathbb{R}\}.$$

4. Determinanten

4.1 Permutationen

4.1.1 Definition Sei $\mathbb{N}_n := \{1, \ldots, n\}$. Eine Bijektion $\pi : \mathbb{N}_n \longrightarrow \mathbb{N}_n$ heißt eine *Permutation* der Zahlen $1, \ldots, n$. Die Hintereinanderausführung von Permutationen bezeichnen wir als Produkt der Permutationen. Wir notieren dies auch als Produkt $\pi \cdot \sigma$. Wir definieren

$$S_n := \{\pi \mid \pi : \mathbb{N}_n \longrightarrow \mathbb{N}_n \text{ ist bijektiv}\}.$$

4.1.2 Lemma *Die Menge S_n bildet mit der Komposition von Abbildungen eine Gruppe. Man nennt sie die* symmetrische Gruppe.

Zum Beweis vergleiche man die Beispiele in 1.1.12.

4.1.3 Definition Ein Element $\tau \in S_n$ heißt eine *Transposition*, wenn τ zwei Zahlen vertauscht und die übrigen unverändert läßt, d.h. wenn es $i, j \in \mathbb{N}_n$ gibt mit $i \neq j$ und für alle $k \in \mathbb{N}_n$ gilt

$$\tau(k) = \begin{cases} i & \text{für } k = j \\ j & \text{für } k = i \\ k & \text{sonst.} \end{cases}$$

Es ist aus der Definition sofort klar, daß $\tau \cdot \tau = id$, d.h. $\tau^{-1} = \tau$ ist. Insbesondere ist das Inverse einer Transposition wieder eine Transposition.

4.1.4 Satz *Jede Permutation $\pi \in S_n$ $(n \geq 2)$ läßt sich als ein Produkt endlich vieler Transpositionen darstellen.*

Beweis: Wir wollen vorausschicken, daß nach Definition 4.1.3 die Identität nicht als Transposition betrachtet wird. Für $\pi \in S_n$ definieren wir

$$k_\pi := \begin{cases} \min\{j \in \mathbb{N}_n \mid \pi(j) \neq j\} & \text{falls dies existiert} \\ n & \text{sonst.} \end{cases}$$

Wir zeigen nun durch Induktion nach $n - k_\pi$, daß es ein Produkt $\prod_{i=1}^{l} \tau_i$ von Transpositionen gibt, mit

$$\prod_{i=1}^{l} \tau_i \cdot \pi = id\,. \tag{i}$$

Ist $k_\pi = n$, so ist $\pi = id$ und es ist $id = \tau \cdot \tau^{-1}$ für jede Transposition τ. Sei nun $k_\pi < n$ und τ_k die Transposition, die k_π und $\pi(k_\pi)$ vertauscht. Dann ist $(\tau_k \cdot \pi)(k_\pi) = \tau_k(\pi(k_\pi)) = k_\pi$ und $(\tau_k \cdot \pi)(m) = \pi(m)$ für $m < k_\pi$. Damit ist $k_\pi < k_{\tau_k \cdot \pi}$. Nach Induktionsvoraussetzung gibt es daher ein Produkt $id = \prod_{i=1}^{l} \tau_i \cdot (\tau_k \pi) = (\prod_{i=1}^{l} \tau_i \cdot \tau_k)\pi$. Damit ist (i) gezeigt. Aus (i) folgt aber

$$\pi = \prod_{i=0}^{l-1} {\tau_{l-i}}^{-1}, \tag{ii}$$

und dies ist die Behauptung, da die Inversen von Transpositionen wieder Transpositionen sind. □

4.1.5 Definition Für $\pi \in S_n$ definieren wir dessen *Signatur*

$$\epsilon(\pi) := \prod_{1 \le i < j \le n} \frac{\pi(i) - \pi(j)}{i - j}.$$

4.1.6 Lemma *Ist τ eine Transposition, so ist $\epsilon(\tau) = -1$.*

Beweis: Vertauscht τ die Zahlen j und k, wobei wir ohne Einschränkung der Allgemeinheit $j < k$ annehmen dürfen, so erhalten wir

$$\begin{aligned}
\epsilon(\tau) = \frac{\tau(j) - \tau(k)}{j - k} \cdot \prod_{\substack{1 \le l < m \le n \\ l \ne j; m \ne k}} \frac{\tau(l) - \tau(m)}{l - m} \cdot \prod_{1 \le m < j} \frac{\tau(m) - \tau(j)}{m - j} \cdot \\
\prod_{\substack{j < m \le n \\ m \ne k}} \frac{\tau(j) - \tau(m)}{j - m} \cdot \prod_{\substack{1 \le m < k \\ m \ne j}} \frac{\tau(m) - \tau(k)}{m - k} \cdot \prod_{k < m \le n} \frac{\tau(k) - \tau(m)}{k - m} \\
= (-1) \cdot 1 \cdot \prod_{1 \le m < j} \frac{m - k}{m - j} \cdot \prod_{\substack{j < m \le n \\ m \ne k}} \frac{m - k}{m - j} \cdot \prod_{\substack{1 \le m < k \\ m \ne j}} \frac{m - j}{m - k} \cdot \prod_{k < m \le n} \frac{m - j}{m - k} \\
= -1.
\end{aligned}$$

□

4.1.7 Lemma *Für $\pi, \rho \in S_n$ gilt $\epsilon(\pi\rho) = \epsilon(\pi)\epsilon(\rho)$.*

Beweis: Es ist $\epsilon(\pi\rho) = \prod_{1\le i<j\le n} \frac{\pi\rho(i)-\pi\rho(j)}{i-j} = \prod_{1\le i<j\le n} \frac{\pi\rho(i)-\pi\rho(j)}{\rho(i)-\rho(j)} \cdot \frac{\rho(i)-\rho(j)}{i-j} = \prod_{1\le i<j\le n} \frac{\pi\rho(i)-\pi\rho(j)}{\rho(i)-\rho(j)} \cdot \epsilon(\rho) = \epsilon(\pi)\epsilon(\rho)$, da wir im ersten Produkt, wann immer $\rho(j) < \rho(i)$ ist, den Faktor $\frac{\pi\rho(i)-\pi\rho(j)}{\rho(i)-\rho(j)}$ durch $\frac{\pi\rho(j)-\pi\rho(i)}{\rho(j)-\rho(i)}$ ersetzen können. □

4.1.8 Satz *Die Abbildung $\epsilon\colon S_n \longrightarrow \{1,-1\}$ ist ein Homomorphismus der Gruppen.*

Beweis: Es folgt aus Lemma 4.1.7, daß ϵ die Multiplikation respektiert. Aus Satz 4.1.4 und Lemma 4.1.6 erhalten wir $\epsilon\colon S_n \longrightarrow \{1,-1\}$. □

4.1.9 Korollar *Ist $\pi \in S_n$ und gilt $\pi = \prod_{i=1}^k \tau_i$ und $\pi = \prod_{i=1}^l \tau_i'$ für Transpositionen τ_i und τ_i', so ist $k - l \equiv 0 \mod 2$, d.h. entweder sind k und l beide gerade oder beide ungerade.*

Beweis: Aus $\pi = \prod_{i=1}^k \tau_i = \prod_{i=1}^l \tau_i'$ folgt mit Satz 4.1.8 $\epsilon(\pi) = (-1)^l = (-1)^k$. Also ist $k - l \equiv 0 \mod 2$. □

4.1.10 Definition Es sei $A_n := \text{Kern}\,(\epsilon) = \{\pi \in S_n \mid \epsilon(\pi) = 1\}$. Damit ist A_n eine Untergruppe der S_n. Man nennt A_n die *alternierende Gruppe*. Sie besteht aus denjenigen Permutationen, die sich durch ein gerade Anzahl von Transpositionen darstellen lassen. Daher nennt man die Elemente der A_n auch *gerade Permutationen*. Analog nennt man Permutationen π mit $\epsilon(\pi) = -1$ *ungerade Permutationen*.

Nach dem Homomorphiesatz für Gruppen gilt

$$S_n/A_n \cong \{1,-1\}. \tag{4.1}$$

Damit ist $[S_n : A_n] = 2$ und nach Lemma 2.2.8 gilt für jede ungerade Permutation π

$$S_n = A_n \cup A_n \cdot \pi \;\; und \;\; A_n \cap A_n \cdot \pi = \emptyset. \tag{4.2}$$

Insbesondere gilt (4.2) für Transpositionen.

4.2 Determinantenfunktionen

Vorbemerkung *Im folgenden setzen wir* **immer** *voraus, daß der Grundkörper K eine Charakteristik $\neq 2$ hat.*

4.2.1 Definition Sei V ein K-Vektorraum der Dimension n. Eine Abbildung $\Delta: V^n \longrightarrow K$ heißt eine *Determinantenfunktion* auf V, wenn gilt

(Det 1) Δ *ist multilinear, d.h. linear in jedem Argument,*

(Det 2) Δ *ist alternierend, d.h.* $\Delta(a_1, \ldots, a_n) = \epsilon(\pi)\Delta(a_{\pi(1)}, \ldots, a_{\pi(n)})$.

Ist Δ nicht die Nullabbildung, so heißt Δ *nicht trivial*.

4.2.2 Folgerungen *(i) Gilt $a_i = a_j$ für $1 \leq i < j \leq n$, so ist $\Delta(a_1, \ldots, a_n) = 0$.*
(ii) Für $i, k \in \{1, \ldots, n\}$ mit $i \neq k$ gilt $\Delta(a_1, \ldots, a_i, \ldots, a_n) = \Delta(a_1, \ldots, \alpha a_k + a_i, \ldots, a_n)$.

Beweis: (i): Sei τ_{ij} die Transposition, die i und j vertauscht. Dann erhalten wir $\Delta(a_1, \ldots, a_n) = \epsilon(\tau_{ij})\Delta(a_{\tau_{ij}(1)}, \ldots, a_{\tau_{ij}(n)}) = -\Delta(a_1, \ldots, a_n)$. Zusammen mit Char$K \neq 2$ folgt daraus $\Delta(a_1, \ldots, a_n) = 0$. (Man beachte, daß in einem Körper der Charakteristik 2 immer $1 = -1$ ist. Die Voraussetzung Char$K \neq 2$ geht also wirklich ein. Will man darauf verzichten und Determinantenfunktionen auch für Vektorräume über Körpern mit einer Charakteristik 2 erklären, so muß man einen anderen Weg wählen. Eine Möglichkeit besteht darin, die Eigenschaft (Det 2) durch die Forderung zu ersetzen, daß eine Determinatenfunktion auf linear abhängigen n-tupeln verschwinden soll.)
(ii): Es ist nach (i) $\Delta(a_1, \ldots, \alpha a_k + a_i, \ldots, a_n) = \alpha\Delta(a_1, \ldots, a_k, \ldots, a_n) + \Delta(a_1, \ldots, a_i, \ldots, a_n) = \Delta(a_1, \ldots, a_i, \ldots, a_n)$. □

4.2.3 Lemma *Sind $a_1, \ldots, a_n \in V$ linear abhängige Vektoren, so gilt*

$$\Delta(a_1, \ldots, a_n) = 0$$

für jede Determinantenfunktion $\Delta: V^n \longrightarrow K$.

Beweis: Sind $a_1, \ldots, a_n$ linear abhängig, so gilt ohne Einschränkung der Allgemeinheit $a_1 = \sum_{i=2}^n \alpha_i a_i$. Also ist nach Folgerung 4.2.2 (i) $\Delta(a_1, \ldots, a_n) = \sum_{i=2}^n \alpha_i \Delta(a_i, a_2, \ldots, a_n) = 0$. □

Wir wollen nun umgekehrt die Wirkung einer Determinantenfunktion Δ auf linear unabhängige Vektoren $a_1, \ldots, a_n \in V$ studieren. Ist Δ trivial, so ist natürlich

$\Delta(a_1,\ldots,a_n) = 0$. Wir wollen daher annehmen, daß Δ nicht trivial ist. Wegen $\dim V = n$ bilden $a_1,\ldots,a_n$ eine Basis von V. Für ein beliebiges n-tupel $v_1,\ldots,v_n \in V$ mit $v_i = \sum_{j=1}^n \alpha_{ij}a_j$ gilt dann

$$\begin{aligned}\Delta(v_1,\ldots,v_n) &= \Delta(\sum_{j=1}^n \alpha_{1j}a_j,\ldots,\sum_{j=1}^n \alpha_{nj}a_j)\\ &= \sum_{j_1,\ldots,j_n=1}^n \alpha_{1j_1}\cdots\alpha_{nj_n}\Delta(a_{j_1},\ldots,a_{j_n}).\end{aligned} \tag{4.3}$$

Nach Folgerung 4.2.2 liefern nur solche Summanden in (4.3) einen Beitrag, für die die $a_{j_1},\ldots,a_{j_n}$ paarweise verschieden sind. Dies ist genau dann der Fall, wenn die $j_1,\ldots,j_n$ paarweise verschieden und damit eine Permutation der Zahlen $1,\ldots,n$ sind. Damit erhalten wir aus (4.3)

$$\begin{aligned}\Delta(v_1,\ldots,v_n) &= \sum_{\pi\in S_n} \alpha_{1\pi(1)}\cdots\alpha_{n\pi(n)}\Delta(a_{\pi(1)},\ldots,a_{\pi(n)})\\ &= \sum_{\pi\in S_n} \alpha_{1\pi(1)}\cdots\alpha_{n\pi(n)}\cdot\epsilon(\pi)\cdot\Delta(a_1,\ldots,a_n).\end{aligned} \tag{4.4}$$

Da wir Δ als nicht trivial vorausgesetzt haben, gibt es ein Tupel $v_1,\ldots,v_n \in V$ mit $\Delta(v_1,\ldots,v_n) \neq 0$. Aus (4.4) folgt dann aber auch $\Delta(a_1,\ldots,a_n) \neq 0$. Damit haben wir den folgenden Satz bewiesen.

4.2.4 Satz *Ist Δ eine nicht triviale Determinantenfunktion, so gilt $\Delta(a_1,\ldots,a_n) = 0$ genau dann, wenn $a_1,\ldots,a_n$ linear abhängig sind.*

4.2.5 Definition Ist $B = (a_1,\ldots,a_n)$ eine Basis von V und ist $v_i = \sum_{j=1}^n \alpha_{ij}a_j$ für $i = 1,\ldots,n$, so gilt nach (4.4)

$$\Delta(v_1,\ldots,v_n) = \sum_{\pi\in S_n} \epsilon(\pi)\cdot\alpha_{1\pi(1)}\cdots\alpha_{n\pi(n)}\Delta(a_1,\ldots,a_n). \tag{4.5}$$

Wir definieren

$$\Delta_B(v_1,\ldots,v_n) := \sum_{\pi\in S_n} \epsilon(\pi)\cdot\alpha_{1\pi(1)}\cdots\alpha_{n\pi(n)}. \tag{4.6}$$

4.2.6 Satz *Ist $B = (a_1,\ldots,a_n)$ eine Basis des n-dimensionalen K-Vektorraums V, so ist Δ_B eine Determinantenfunktion mit $\Delta_B(a_1,\ldots,a_n) = 1$. Insbesondere ist Δ_B nicht trivial.*

Beweis: Wir zeigen als erstes, daß Δ_B multilinear ist. Sei $v_i = \sum_{j=1}^n \alpha_{ij} a_j$ und $w = \sum_{j=1}^n \beta_j a_j$. Dann ist

$$\Delta_B(v_1, \ldots, v_i + w, \ldots, v_n) = \sum_{\pi \in S_n} \epsilon(\pi) \cdot \alpha_{1\pi(1)} \cdots (\alpha_{i\pi(i)} + \beta_{\pi(i)}) \cdots \alpha_{n\pi(n)}$$

$$= \sum_{\pi \in S_n} \epsilon(\pi) \cdot \alpha_{1\pi(1)} \cdots \alpha_{i\pi(i)} \cdots \alpha_{n\pi(n)} + \sum_{\pi \in S_n} \epsilon(\pi) \cdot \alpha_{1\pi(1)} \cdots \beta_{\pi(i)} \cdots \alpha_{n\pi(n)}$$

$$= \Delta_B(v_1, \ldots, v_n) + \Delta_B(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n).$$

Aus der definierenden Gleichung (4.6) liest man sofort ab, daß

$$\Delta_B(v_1, \ldots, \alpha v_i, \ldots, v_n) = \alpha \Delta_B(v_1, \ldots, v_i, \ldots, v_n)$$

ist. Damit ist Δ_B im i-ten Argument linear und, da wir i beliebig gewählt haben, in jedem Argument linear.
Nun zeigen wir, daß Δ_B alternierend ist. Offenbar genügt es dazu zu zeigen, daß für eine Transposition $\tau \in S_n$

$$\Delta_B(v_{\tau(1)}, \ldots, v_{\tau(n)}) = -\Delta_B(v_1, \ldots, v_n) \tag{i}$$

gilt. Bei der folgenden Rechnung werden wir auf Gleichung (4.2) zurückgreifen, der zufolge sich die S_n in die alternierende Gruppe A_n und $A_n \cdot \tau$ disjunkt zerlegen läßt. Wegen $v_{\tau(i)} = \sum_{j=1}^{n} \alpha_{\tau(i)j} a_j$ erhalten wir

$$\Delta_B(v_{\tau(1)}, \ldots, v_{\tau(n)}) = \sum_{\pi \in S_n} \epsilon(\pi) \prod_{i=1}^n \alpha_{\tau(i)\pi(i)}$$

$$= \sum_{\pi \in A_n} \epsilon(\pi) \prod_{i=1}^n \alpha_{\tau(i)\pi(i)} + \sum_{\sigma \in A_n} \epsilon(\sigma \cdot \tau) \prod_{i=1}^n \alpha_{\tau(i)\sigma(\tau(i))} \tag{ii}$$

$$= \sum_{\pi \in A_n} \epsilon(\pi) \prod_{i=1}^n \alpha_{\tau(i)\pi(i)} + \sum_{\sigma \in A_n} - \epsilon(\sigma) \prod_{j=1}^n \alpha_{j\sigma(j)} \tag{iii}$$

wobei wir $j := \tau(i)$ gesetzt haben. Setzen wir im ersten Summanden $j := \tau(i)$, so ist $i = \tau(j)$, und wir können (iii) folgendermaßen fortsetzen

$$= \sum_{\pi \in A_n} \epsilon(\pi) \prod_{j=1}^n \alpha_{j\pi(\tau(j))} + \sum_{\sigma \in A_n} - \epsilon(\sigma) \prod_{j=1}^n \alpha_{j\sigma(j)}. \tag{iv}$$

Setzen wir $\sigma = \pi \cdot \tau$ im ersten Summanden von (iv), so durchläuft σ alle Elemente von $A_n \cdot \tau$. Damit können wir (iv) wie folgt fortsetzen

$$\begin{aligned} &= \sum_{\sigma \in A_n \cdot \tau} -\epsilon(\sigma) \prod_{j=1}^{n} \alpha_{j\sigma(j)} + \sum_{\sigma \in A_n} -\epsilon(\sigma) \prod_{j=1}^{n} \alpha_{j\sigma(j)} \\ &= \sum_{\sigma \in S_n} -\epsilon(\sigma) \prod_{j=1}^{n} \alpha_{j\sigma(j)} = -\Delta_B(v_1, \ldots, v_n). \end{aligned} \tag{v}$$

Es bleibt zu zeigen, daß $\Delta_B(a_1, \ldots, a_n) = 1$ ist. Wegen $a_i = \sum_{j=1}^{n} \delta_{ij} a_j$ erhalten wir

$$\Delta_B(a_1, \ldots, a_n) = \sum_{\pi \in S_n} \epsilon(\pi) \prod_{i=1}^{n} \delta_{i\pi(i)} = 1,$$

denn alle Summanden mit $\pi \neq id$ verschwinden. □

4.2.7 Satz *Sind Δ und Δ' Determinantenfunktionen auf einem n-dimensionalen K-Vektorraum V und ist Δ' nicht trivial, so gibt es ein $\gamma \in K$ mit*

$$(\forall v_1, \ldots, v_n \in V)[\Delta(v_1, \ldots, v_n) = \gamma \cdot \Delta'(v_1, \ldots, v_n)].$$

Beweis: Wir wählen eine Basis $B = (b_1, \ldots, b_n)$ von V. Nach (4.5) erhalten wir dann

$$\Delta(v_1, \ldots, v_n) = \Delta_B(v_1, \ldots, v_n) \cdot \Delta(b_1, \ldots, b_n) \tag{i}$$

und

$$\Delta'(v_1, \ldots, v_n) = \Delta_B(v_1, \ldots, v_n) \cdot \Delta'(b_1, \ldots, b_n). \tag{ii}$$

Da wir Δ' als nicht trivial vorausgesetzt haben, folgt $\Delta'(b_1, \ldots, b_n) \neq 0$ nach Satz 4.2.6. Wir setzen $\beta := \Delta(b_1, \ldots, b_n)$ und $\beta' := \Delta'(b_1, \ldots, b_n)$. Dann erhalten wir aus (i)und (ii)

$$\Delta(v_1, \ldots, v_n) = \beta \cdot \Delta_B(v_1, \ldots, v_n) = \frac{\beta}{\beta'} \cdot \Delta'(v_1, \ldots, v_n). \tag{iii}$$

Mit $\gamma := \frac{\beta}{\beta'}$ folgt die Behauptung. □

4.3 Die Determinante eines Endomorphismus und einer Matrix

4.3.1 Definition Sei V ein n-dimensionaler K-Vektorraum und Δ eine nicht triviale Determinatenfunktion auf V. Für einen Endomorphismus $f: V \longrightarrow V$ definieren wir

$$\Delta_f(v_1, \ldots, v_n) := \Delta(f(v_1), \ldots, f(v_n)).$$

4.3.2 Lemma *Die Funktion Δ_f ist eine Determinantenfunktion.*

Beweis: Aus der Linearität von f und der Multilinearität von Δ folgt sofort die Multilinearität von Δ_f. Wegen $\Delta_f(v_{\pi(1)}, \ldots, v_{\pi(n)}) = \Delta(f(v_{\pi(1)}), \ldots, f(v_{\pi(n)})) = \epsilon(\pi) \cdot \Delta(f(v_1), \ldots, f(v_n)) = \epsilon(\pi) \cdot \Delta_f(v_1, \ldots, v_n)$ ist Δ_f auch alternierend. □

4.3.3 Lemma *Sei V ein n-dimensionaler K-Vektorraum. Zu jedem $f \in \mathrm{End}(V)$ gibt es ein $\alpha_f \in K$ mit*

$$\Delta_f(v_1, \ldots, v_n) = \alpha_f \cdot \Delta(v_1, \ldots, v_n) \tag{4.7}$$

für alle nicht trivialen Determinantenfunktionen Δ und alle n-tupel $(v_1, \ldots, v_n) \in V^n$.

Beweis: Zu jeder nicht trivialen Determinantenfunktion gibt es nach Lemma 4.3.2 und Satz 4.2.7 ein $\alpha_{f,\Delta}$ mit

$$\Delta_f(v_1, \ldots, v_n) = \alpha_{f,\Delta} \cdot \Delta(v_1, \ldots, v_n). \tag{i}$$

Wir haben zu zeigen, daß $\alpha_{f,\Delta}$ nicht von der Wahl von Δ abhängt. Sei daher Γ eine weitere nicht triviale Determinantenfunktion. Dann gilt

$$\Gamma_f(v_1, \ldots, v_n) = \alpha_{f,\Gamma} \cdot \Gamma(v_1, \ldots, v_n) \tag{ii}$$

und nach Satz 4.2.7 auch

$$\Gamma(v_1, \ldots, v_n) = \gamma \cdot \Delta(v_1, \ldots, v_n). \tag{iii}$$

Damit folgt

$$\begin{aligned}\Gamma_f(v_1, \ldots, v_n) &= \Gamma(f(v_1), \ldots, f(v_n)) = \gamma \cdot \Delta(f(v_1), \ldots, f(v_n)) \\ &= \gamma \cdot \alpha_{f,\Delta}\Delta(v_1, \ldots, v_n) = \gamma \cdot \alpha_{f,\Delta} \cdot \gamma^{-1}\Gamma(v_1, \ldots, v_n).\end{aligned} \tag{iv}$$

Aus (ii) und (iv) folgt aber $\alpha_{f,\Gamma} = \alpha_{f,\Delta}$. □

4.3.4 Definition Zu $f \in \mathrm{End}(V)$ sei $\det(f)$ das nach Lemma 4.3.3 eindeutig bestimmte Körperelement mit

$$\Delta_f(v_1, \ldots, v_n) = \det(f) \cdot \Delta(v_1, \ldots, v_n) \tag{4.8}$$

für jede nicht triviale Determinantenfunktion Δ auf V. Wir nennen $\det(f)$ die *Determinante* des Endomorphismus f.

4.3.5 Lemma *Ist $B = (b_1, \dots, b_n)$ eine Basis des K-Vektorraumes V und $f \in \mathrm{End}(V)$, so ist* $\det f = \Delta_B(f(b_1), \dots, f(b_n))$.

Beweis: Mit Gleichung (4.8) und Satz 4.2.6 erhalten wir $\Delta_B(f(b_1), \dots, f(b_n)) = \det(f) \cdot \Delta_B(b_1, \dots, b_n) = \det(f)$. □

Beachte Mit Lemma 4.3.5 haben wir die Möglichkeit $\det(f)$ zu berechnen, wenn $\mathfrak{D}_{B,B}(f) = (\alpha_{ij})$ bekannt ist. Wegen $f(b_i) = \sum_{j=1}^{n} \alpha_{ij} b_j$ folgt dann

$$\det(f) = \Delta_B(f(b_1), \dots, f(b_n)) = \sum_{\pi \in S_n} \epsilon(\pi) \prod_{i=1}^{n} \alpha_{i\pi(i)}. \tag{4.9}$$

Es folgt aus Lemma 4.3.5, daß (4.9) nicht von der Wahl der Basis $B = (b_1, \dots, b_n)$ abhängt.

4.3.6 Satz *Sei V ein n dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Dann gelten:*

(i) $\det(\mathit{id}) = 1$,

(ii) $\det(\alpha f) = \alpha^n \cdot \det(f)$,

(iii) $\det(g \circ f) = \det(g) \cdot \det(f)$,

(iv) $\det(f) \neq 0 \;\Leftrightarrow\; f$ *ist bijektiv.*

Beweis: (i): Wegen $\Delta_{id} = \Delta$ ist diese Aussage klar.
(ii): Es ist

$$\begin{aligned}\Delta_{\alpha f}(v_1, \dots, v_n) &= \Delta(\alpha f(v_1), \dots, \alpha f(v_n)) = \alpha^n \cdot \Delta(f(v_1), \dots, f(v_n)) \\ &= \alpha^n \cdot \det(f) \cdot \Delta(v_1, \dots, v_n).\end{aligned}$$

Nach (4.8) folgt $\det(\alpha f) = \alpha^n \cdot \det(f)$.
(iii): Durch Nachrechnen erhalten wir

$$\begin{aligned}\Delta_{g \circ f}(v_1, \dots, v_n) &= \Delta(g(f(v_1)), \dots, g(f(v_n))) = \det(g) \cdot \Delta(f(v_1), \dots, f(v_n)) \\ &= \det(g) \cdot \det(f) \cdot \Delta(v_1, \dots, v_n).\end{aligned}$$

Damit ist $\det(g \circ f) = \det(g) \cdot \det(f)$.
(iv): $\Leftarrow$: Ist f bijektiv, so gibt es ein $g \in \mathrm{End}(V)$ mit $g \circ f = \mathit{id}$ und damit nach (i) und (iii) $\det(f) \cdot \det(g) = 1$. Damit ist $\det(f) \neq 0$.
$\Rightarrow$: Sei $\det(f) \neq 0$ und $B = (b_1, \dots, b_n)$ eine Basis von V. Dann ist $\det(f) = \Delta_B(f(b_1), \dots, f(b_n)) \neq 0$. Nach Satz 4.2.6 ist Δ_B nicht trivial und damit folgt

nach Satz 4.2.4 die lineare Unabhängigkeit von $f(b_1), \ldots, f(b_n)$. Also ist Rang $f = n$, und nach den Sätzen 3.4.5 und 2.3.9 ist f damit bijektiv. □

4.3.7 Satz *Die Abbildung* $\det\colon \mathrm{GL}_K(V) \longrightarrow K^*$ *ist ein Homomorphismus der Gruppen. Insbesondere ist* $\det(f^{-1}) = \det(f)^{-1}$.

Beweis: Nach Satz 4.3.6 (iv) ist $\det\colon \mathrm{GL}_K(V) \longrightarrow K^*$ und nach (iii) ein Homomorphismus. □

4.3.8 Definition Für $\mathfrak{A} \in K^{n\times n}$ definieren wir $\det(\mathfrak{A}) := \det(\overline{\mathfrak{A}})$. Man nennt $\det(\mathfrak{A})$ die *Determinante der Matrix* $\mathfrak{A}$.

4.3.9 Satz *Ist* $\mathfrak{A} = (\alpha_{ij}) \in K^{n\times n}$ *so ist*

$$\det(\mathfrak{A}) = \sum_{\pi\in S_n} \epsilon(\pi) \prod_{i=1}^{n} \alpha_{i\pi(i)}. \tag{4.10}$$

Beweis: Wegen $\mathfrak{D}_{\mathbb{E}^n,\mathbb{E}^n}(\overline{\mathfrak{A}}) = \mathfrak{A}$ folgt dies sofort aus (4.9). □

4.3.10 Lemma *Sind* V *und* W *Vektorräume endlicher Dimension und* $g\colon V \longrightarrow W$ *ein Isomorphismus, so ist* $\det(f) = \det(g \circ f \circ g^{-1})$ *für jeden Endomorphismus* $f \in \mathrm{End}(V)$.

Beweis: Ist Δ eine Determinatenfunktion auf W, so definiert die Abbildung $\Delta^g(v_1, \ldots, v_n) := \Delta(g(v_1), \ldots, g(v_n))$ eine Determinantenfunktion auf V, die nicht trivial ist, wenn Δ nicht trivial ist. Damit folgt für $f \in \mathrm{End}(V)$

$$\begin{aligned}\Delta((g\circ f\circ g^{-1})(v_1), \ldots, (g\circ f\circ g^{-1})(v_n)) &= \Delta^g((f\circ g^{-1})(v_1), \ldots, (f\circ g^{-1})(v_n))\\ &= \det(f)\cdot \Delta^g(g^{-1}(v_1), \ldots, g^{-1}(v_n))\\ &= \det(f)\cdot \Delta(v_1, \ldots, v_n)\end{aligned}$$

und damit $\det(g\circ f\circ g^{-1}) = \det(f)$. □

4.3.11 Satz *Ist* V *ein endlich dimensionaler* K*-Vektorraum und ist* B *eine Basis von* V*, so gilt* $\det(f) = \det(\mathfrak{D}_{B,B}(f))$ *für jeden Endomorphismus* $f \in \mathrm{End}(V)$.

Beweis: Es gilt $f = h_B{}^{-1} \circ \overline{\mathfrak{D}_{B,B}(f)} \circ h_B$. Mit Lemma 4.3.10 folgt daraus $\det(f) = \det(\overline{\mathfrak{D}_{B,B}(f)}) = \det(\mathfrak{D}_{B,B}(f))$.
Wir hätten die Aussage von Satz 4.3.11 auch aus (4.9) und (4.10) folgern können. □

4.3.12 Satz *Sei V ein endlich dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$. Dann gilt* $\det(f) = \det(f^{**})$.

Beweis: Nach Satz 3.2.12 ist $f^{**}(v^{**}) = (f(v))^{**}$. Damit ist $f^{**} =^{**} \circ f \circ {}^{**-1}$ und ${}^{**}\colon V \longrightarrow V^{**}$ ein Isomorphismus. Mit Lemma 4.3.10 folgt dann die Behauptung. □

4.3.13 Satz *Für* $\mathfrak{A} \in K^{n\times n}$ *gilt* $\det \mathfrak{A} = \det(\mathfrak{A}^t)$.

Beweis: Sei $\mathfrak{A} = (\alpha_{ij})$. Nach (4.10) ist dann

$$\begin{aligned}\det \mathfrak{A} &= \sum_{\pi\in S_n} \epsilon(\pi) \cdot \prod_{i=1}^{n} \alpha_{i\pi(i)} = \sum_{\pi\in S_n} \epsilon(\pi) \cdot \prod_{i=1}^{n} \alpha_{\pi^{-1}(\pi i)\pi i} \\ &= \sum_{\pi^{-1}\in S_n} \epsilon(\pi^{-1}) \prod_{j=1}^{n} \alpha_{\pi^{-1}(j)j} = \det(\mathfrak{A}^t),\end{aligned}$$

denn mit π durchläuft auch π^{-1} alle Elemente von S_n, und offensichtlich ist $\epsilon(\pi) = \epsilon(\pi^{-1})$. □

4.3.14 Korollar *Ist V ein endlich dimensionaler Vektorraum und $f \in \mathrm{End}(V)$, so ist* $\det(f) = \det(f^*)$.

Beweis: Nach Satz 3.2.8 gilt für eine Basis B von V stets $\mathfrak{D}_{B^*,B^*}(f^*) = \mathfrak{D}_{B,B}(f)^t$. Nach Satz 4.3.11 und Satz 4.3.13 gilt daher $\det(f^*) = \det(\mathfrak{D}_{B^*,B^*}(f^*)) = \det(\mathfrak{D}_{B,B}(f)) = \det(f)$. □

4.3.15 Satz *Die Funktionen* $\det\colon (K_n)^n \longrightarrow K$ *und* $\det\colon (K^n)_n \longrightarrow K$*, die definiert sind durch* $\det(\mathfrak{a}_1, \dots, \mathfrak{a}_n) = \det(\mathfrak{A})$ *mit* $\mathfrak{A} := (\mathfrak{a}_1, \dots, \mathfrak{a}_n)$ *für* $\mathfrak{a}_i \in K_n$ *bzw.* $\mathfrak{A} = \begin{pmatrix}\mathfrak{a}_1\\ \vdots \\ \mathfrak{a}_n\end{pmatrix}$ *für* $\mathfrak{a}_i \in K^n$*, sind nicht triviale Determinantenfunktionen.*

Beweis: Ist $\mathfrak{a}_i = (\alpha_{i1}, \dots, \alpha_{in})$, so ist $\mathfrak{a}_i = \sum_{j=1}^{n} \alpha_{ij} e_j$. Nach Satz 4.3.9 gilt dann $\det(\mathfrak{A}) = \sum_{\pi\in S_n} \epsilon(\pi) \prod_{i=1}^{n} \alpha_{i\pi(i)} = \Delta_{\mathbb{E}^n}(\mathfrak{a}_1, \dots, \mathfrak{a}_n)$. Nach Satz 4.2.6 ist $\Delta_{\mathbb{E}^n}$ aber eine nicht triviale Determinatenfunktion. □

4.3.16 Folgerungen

(i) $\det(\mathfrak{E}) = 1$.

(ii) *Für* $\alpha \in K$ *und* $\mathfrak{A} \in K^{n\times n}$ *ist* $\det(\alpha\mathfrak{A}) = \alpha^n \cdot \det(\mathfrak{A})$.

(iii) *Für* $\mathfrak{A}, \mathfrak{B} \in K^{n\times n}$ *ist* $\det(\mathfrak{A}\mathfrak{B}) = \det(\mathfrak{A}) \cdot \det(\mathfrak{B})$.

(iv) *Für* $\mathfrak{A} \in K^{n\times n}$ *gilt* $\mathfrak{A} \in \mathrm{GL}(n, K) \;\Leftrightarrow\; \det(\mathfrak{A}) \neq 0$.

(v) *Für* $\mathfrak{A} \in \mathrm{GL}(n, K)$ *ist* $\det(\mathfrak{A}^{-1}) = (\det(\mathfrak{A}))^{-1}$.

(vi) *Für* $\mathfrak{A} \in K^{n\times n}$ *gilt* $\det(\mathfrak{A}) = 0$, *wenn* Rang $(\mathfrak{A}) < n$ *ist.*

(vii) *Entsteht* $\mathfrak{A}'$ *aus* $\mathfrak{A} \in K^{n\times n}$ *durch das Vertauschen zweier Spalten bzw. Zeilen, so gilt* $\det(\mathfrak{A}') = -\det(\mathfrak{A})$.

(viii) *Entsteht* $\mathfrak{A}'$ *aus* $\mathfrak{A} \in K^{n\times n}$*, indem man das Vielfache einer Spalte (bzw. einer Zeile) zu einer* **anderen** *Spalte (bzw. Zeile) hinzuaddiert, so ist* $\det(\mathfrak{A}') = \det(\mathfrak{A})$.

Beweis: Die Behauptungen in (i)–(iv) folgen sofort aus Satz 4.3.6 (i) – (iv). Die Behauptung (v) ergibt sich aus Satz 4.3.7 und der Tatsache, daß $\overline{\mathfrak{A}^{-1}} = \overline{\mathfrak{A}}^{-1}$ ist (vgl. den Beweis zu Satz 3.3.13). Die Behauptung (vi) ergibt sich aus der Tatsache, daß $\det$ eine Determinantenfunktion ist (Satz 4.3.15), Lemma 4.2.3 und der Tatsache, daß für Rang $(\mathfrak{A}) < n$ nach Satz 3.5.9 die Zeilen- und Spaltenvektoren in $\mathfrak{A}$ linear abhängig sein müssen. Die Aussagen (vii) und (viii) folgen sofort aus Satz 4.3.15 und den Eigenschaften von Determinantenfunktionen (vgl. Folgerung 4.2.2 (ii) für (viii)). □

4.3.17 Einige Rechenregeln für Determinanten

(i) $\det \begin{pmatrix} \mathfrak{A} & 0 \\ 0 & \mathfrak{B} \end{pmatrix} = \det(\mathfrak{A}) \cdot \det(\mathfrak{B})$, *wobei* $\mathfrak{A}$ *und* $\mathfrak{B}$ *quadratische Matrizen sind.*

Beweis: Schreiben wir $\mathfrak{A} = (\mathfrak{a}_1, \dots, \mathfrak{a}_r)$ und $\mathfrak{B} = (\mathfrak{b}_{r+1}, \dots, \mathfrak{b}_n)$, so erhalten wir $\det(\mathfrak{a}_1, \dots, \mathfrak{a}_r)$ und $\det(\mathfrak{b}_{r+1}, \dots, \mathfrak{b}_n)$ als nicht triviale Determinantenfunktionen auf K_r bzw. K_{n-r}. Ebenso sind $\Delta_{\mathfrak{B}}(\mathfrak{a}_1, \dots, \mathfrak{a}_r) := \det \begin{pmatrix} \mathfrak{A} & 0 \\ 0 & \mathfrak{B} \end{pmatrix}$ und $\Delta_{\mathfrak{A}}(\mathfrak{b}_{r+1}, \dots, \mathfrak{b}_n) := \det \begin{pmatrix} \mathfrak{A} & 0 \\ 0 & \mathfrak{B} \end{pmatrix}$ Determinantenfunktionen auf K_r und K_{n-r}. Nach Satz 4.2.7 gilt dann

$$\Delta_{\mathfrak{B}}(\mathfrak{a}_1, \dots, \mathfrak{a}_r) = \alpha_{\mathfrak{B}} \cdot \det(\mathfrak{a}_1, \dots, \mathfrak{a}_r) \tag{i}$$

und

$$\Delta_{\mathfrak{A}}(\mathfrak{b}_{r+1}, \dots, \mathfrak{b}_n) = \alpha_{\mathfrak{A}} \cdot \det(\mathfrak{b}_{r+1}, \dots, \mathfrak{b}_n). \tag{ii}$$

Aus (i) und (ii) folgt

$$\det\begin{pmatrix}\mathfrak{A} & 0\\ 0 & \mathfrak{B}\end{pmatrix} = \alpha_{\mathfrak{B}} \cdot \det(\mathfrak{A}) = \alpha_{\mathfrak{A}} \cdot \det(\mathfrak{B}). \tag{iii}$$

Wählen wir in (iii) speziell $\mathfrak{B} = \mathfrak{E}$, so erhalten wir

$$\alpha_{\mathfrak{E}} \cdot \det(\mathfrak{A}) = \alpha_{\mathfrak{A}}. \tag{iv}$$

Damit erhalten wir

$$\det\begin{pmatrix}\mathfrak{A} & 0\\ 0 & \mathfrak{B}\end{pmatrix} = \frac{\alpha_{\mathfrak{A}}}{\det(\mathfrak{A})} \cdot \det(\mathfrak{A}) \cdot \det(\mathfrak{B}) = \alpha_{\mathfrak{E}} \cdot \det(\mathfrak{A}) \cdot \det(\mathfrak{B}). \tag{v}$$

Wählen wir $\mathfrak{A} = \mathfrak{B} = \mathfrak{E}$ in (v), so folgt $\alpha_{\mathfrak{E}} = 1$, und wir erhalten die Behauptung aus (v). □

(ii) $$\det\begin{pmatrix}1 & * & \cdots & *\\ 0 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & *\\ 0 & \cdots & 0 & 1\end{pmatrix} = 1$$

Beweis: Wir bezeichnen die Matrix mit $\mathfrak{A} \in K^{n\times n}$. Durch Addition von geeigneten Vielfachen der ersten Spalte läßt sich die Matrix auf die Form

$$\mathfrak{A}' = \begin{pmatrix}1 & 0 & \cdots & \cdots & 0\\ 0 & 1 & \star & \cdots & \star\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ \vdots & & \ddots & \ddots & \star\\ 0 & \cdots & \cdots & 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & \mathfrak{B}\end{pmatrix}$$

bringen, wobei $\mathfrak{B} \in K^{(n-1)\times(n-1)}$ die gleiche Gestalt wie $\mathfrak{A}$ hat. Nach Folgerung 4.3.16 (viii) ist $\det \mathfrak{A} = \det \mathfrak{A}'$, und nach Folgerung 4.3.17 (i) ist $\det \mathfrak{A}' = 1 \cdot \det B$. Damit erhalten wir $\det \mathfrak{A} = 1$ durch Induktion nach n. □

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

Beweis: Ist $\mathfrak{A} \notin \mathrm{GL}(n, K)$, so sind die Spalten in $\mathfrak{A}$ linear abhängig und damit auch in der Matrix $\mathfrak{D} := \begin{pmatrix}\mathfrak{A} & \mathfrak{B}\\ 0 & \mathfrak{C}\end{pmatrix}$. Damit ist $\det \mathfrak{D} = 0 = \det \mathfrak{A} \cdot \det \mathfrak{C}$. Nehmen wir $\mathfrak{A} \in \mathrm{GL}(n, K)$ an, so folgt

$$\det\begin{pmatrix}\mathfrak{A} & \mathfrak{B}\\ 0 & \mathfrak{C}\end{pmatrix} = \det\begin{pmatrix}\mathfrak{A} & 0\\ 0 & \mathfrak{C}\end{pmatrix} \cdot \det\begin{pmatrix}\mathfrak{E} & \mathfrak{A}^{-1}\mathfrak{B}\\ 0 & \mathfrak{E}\end{pmatrix} = \det \mathfrak{A} \cdot \det \mathfrak{C},$$

wobei wir im letzten Schritt (i) und (ii) benutzt haben. □

4.4 Die inverse Matrix und die Cramersche Regel

4.4.1 Definition Für $\mathfrak{A} \in K^{n \times n}$ sei $\mathfrak{A}_{ij}$ die Matrix, die aus $\mathfrak{A}$ entsteht, wenn man die i-te Zeile und die j-te Spalte durch Nullen ersetzt und das Element $\alpha_{ij} := 1$ setzt, d.h.

$$\mathfrak{A}_{ij} := (\mathfrak{E} - \mathfrak{E}_{ii}) \cdot \mathfrak{A} \cdot (\mathfrak{E} - \mathfrak{E}_{jj}) + \mathfrak{E}_{ij}.$$

Das Körperelement $\tilde{\alpha}_{ij} := \det(\mathfrak{A}_{ji})$ heißt der *Cofaktor* oder das *algebraische Komplement* der Matrix $\mathfrak{A}$ bezüglich i, j. (Man beachte die Vertauschung der Indizes i und j in der Definition von $\tilde{\alpha}_{ij}$.)
Schließlich sei $\mathfrak{A}^{ij} \in K^{(n-1)\times(n-1)}$ die Matrix, die aus $\mathfrak{A}$ durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

4.4.2 Lemma *Es ist* $\tilde{\alpha}_{ij} = \det(\mathfrak{A}_{ji}) = (-1)^{i+j} \det(\mathfrak{A}^{ji})$.

Beweis: Durch $j-1$-faches Vertauschen von Zeilen bringt man die j-te Zeile von $\mathfrak{A}_{ji}$ in die erste Zeile. Dann bringt man durch $i-1$-faches Vertauschen von Spalten, die i-te Spalte in die erste Spalte und erhält so die Matrix

$$\mathfrak{B} = \begin{pmatrix} 1 & 0 \\ 0 & \mathfrak{A}^{ji} \end{pmatrix}.$$

Da die Determinante bei jeder Vertauschung ihr Vorzeichen wechselt folgt mit 4.3.17 (i)

$$\det(\mathfrak{A}_{ji}) = (-1)^{i+j-2} \cdot \det(\mathfrak{B}) = (-1)^{i+j} \cdot \det(\mathfrak{A}^{ji}).$$ □

4.4.3 Lemma *Schreiben wir die Matrix* $\mathfrak{A} \in K^{n \times n}$ *als Matrix von Spaltenvektoren* $\mathfrak{A} = (\mathfrak{a}_1, \ldots, \mathfrak{a}_n)$ *mit* $\mathfrak{a}_k \in K_n$ *für* $k = 1, \ldots, n$*, so ist*

$$\det(\mathfrak{A}_{ij}) = \det(\mathfrak{a}_1, \ldots, \mathfrak{a}_{j-1}, \mathbf{e}_i^t, \mathfrak{a}_{j+1}, \ldots, \mathfrak{a}_n).$$

Beweis: Man erhält die Matrix $\mathfrak{A}_{ij}$ aus $\mathfrak{B} := (\mathfrak{a}_1, \ldots, \mathfrak{a}_{j-1}, \mathbf{e}_i^t, \mathfrak{a}_{j+1}, \ldots, \mathfrak{a}_n)$ durch Subtraktion des α_{ik}-fachen der j-ten Spalte $\mathbf{e}_i^t$ von den Spalten $\mathfrak{a}_k$ für $k \in \{1, \ldots, n\} \setminus \{j\}$. Wegen 4.3.16 (viii) gilt dann $\det(\mathfrak{A}_{ij}) = \det(\mathfrak{B})$. □

4.4.4 Definition Die Matrix $\widetilde{\mathfrak{A}} := (\tilde{\alpha}_{ij})$ heißt die zu $\mathfrak{A}$ *konjungierte Matrix* oder die *Komplementärmatrix* von $\mathfrak{A}$.

4.4.5 Lemma *Es ist* $\widetilde{\mathfrak{A}}^t = \widetilde{\mathfrak{A}^t}$.

Beweis: Sei $\mathfrak{A}^t = (\beta_{ij}) =: \mathfrak{B}$, d.h. $\beta_{ij} = \alpha_{ji}$ für $i, j = 1, \dots, n$. Dann gilt $\mathfrak{B}_{ij} = (\mathfrak{A}_{ij})^t$ und damit $\det(\mathfrak{A}^t_{ji}) = \tilde{\beta}_{ij} = \det(\mathfrak{B}_{ji}) = \det(\mathfrak{A}_{ji})^t = \det(\mathfrak{A}_{ij}) = \tilde{\alpha}_{ji}$. Also $\widetilde{\mathfrak{A}^t} = (\tilde{\beta}_{ij}) = \widetilde{\mathfrak{B}} = (\tilde{\alpha}_{ji}) = (\widetilde{\mathfrak{A}})^t$. □

4.4.6 Satz *Für $\mathfrak{A} \in K^{n\times n}$ gilt $\widetilde{\mathfrak{A}} \cdot \mathfrak{A} = \mathfrak{A} \cdot \widetilde{\mathfrak{A}} = \det(\mathfrak{A}) \cdot \mathfrak{E}$.*

Beweis: Sei $\mathfrak{A} = (\alpha_{ij})$. Dann ist

$$\sum_{k=1}^{n} \tilde{\alpha}_{ik}\alpha_{kj} = \delta_{ij} \cdot \det(\mathfrak{A}), \tag{i}$$

denn

$$\begin{aligned}
\sum_{k=1}^{n} \tilde{\alpha}_{ik}\alpha_{kj} &= \sum_{k=1}^{n} \alpha_{kj} \cdot \det(\mathfrak{A}_{ki}) \\
&= \sum_{k=1}^{n} \alpha_{kj} \cdot \det(\mathfrak{a}_1, \dots, \mathfrak{a}_{i-1}, \mathbf{e}_k^t, \mathfrak{a}_{i+1}, \dots, \mathfrak{a}_n) \\
&= \det(\mathfrak{a}_1, \dots, \mathfrak{a}_{i-1}, \sum_{k=1}^{n} \alpha_{kj}\mathbf{e}_k^t, \mathfrak{a}_{i+1}, \dots, \mathfrak{a}_n) \\
&= \det(\mathfrak{a}_1, \dots, \mathfrak{a}_{i-1}, \mathfrak{a}_j, \mathfrak{a}_{i+1}, \dots, \mathfrak{a}_n) \\
&= \delta_{ij} \cdot \det(\mathfrak{A}).
\end{aligned}$$

Aus (i) erhalten wir sofort

$$\widetilde{\mathfrak{A}} \cdot \mathfrak{A} = \det(\mathfrak{A}) \cdot \mathfrak{E} \tag{ii}$$

und aus (ii) mit Lemma 4.4.5

$$\mathfrak{A} \cdot \widetilde{\mathfrak{A}} = (\widetilde{\mathfrak{A}}^t \mathfrak{A}^t)^t = (\widetilde{\mathfrak{A}^t} \mathfrak{A}^t)^t = (\det(\mathfrak{A}^t) \cdot \mathfrak{E})^t = \mathfrak{E} \cdot \det(\mathfrak{A}). \tag{iii}$$

□

4.4.7 Satz *(Entwicklungssatz für Determinanten) Für $\mathfrak{A} = (\alpha_{ij}) \in K^{n\times n}$ gilt*

$$\det(\mathfrak{A}) = \sum_{k=1}^{n} \alpha_{ki} \det(\mathfrak{A}_{ki}) = \sum_{k=1}^{n} (-1)^{k+i} \alpha_{ki} \det(\mathfrak{A}^{ki}) \tag{4.11}$$

und

$$\det(\mathfrak{A}) = \sum_{k=1}^{n} \alpha_{ik} \det(\mathfrak{A}_{ik}) = \sum_{k=1}^{n} (-1)^{k+i} \alpha_{ik} \det(\mathfrak{A}^{ik}). \tag{4.12}$$

Man nennt (4.11) die Entwicklung der Determinante nach der i-ten Spalte und (4.12) die Entwicklung nach der i-ten Zeile.

Beweis: Setzen wir $j := i$ in Gleichung (i) des Beweises von Satz 4.4.6, so erhalten wir

$$\det(\mathfrak{A}) = \sum_{k=1}^{n} \tilde{\alpha}_{ik}\alpha_{ki} = \sum_{k=1}^{n} \alpha_{ki} \cdot \det(\mathfrak{A}_{ki}). \tag{i}$$

Aus (i) und Lemma 4.4.2 erhalten wir die erste Behauptung. Aus Gleichung (iii) des Beweises von Satz 4.4.6 folgt

$$\delta_{ij} \cdot \det(\mathfrak{A}) = \sum_{k=1}^{n} \alpha_{ik}\tilde{\alpha}_{kj}. \tag{ii}$$

Setzen wir $j := i$ in (ii), so erhalten wir die zweite Behauptung analog wie die erste. □

Beachte Da $\mathfrak{A}^{ki} \in K^{(n-1)\times(n-1)}$ liefert Satz 4.4.7 ein rekursives Verfahren zur Berechnung von Determinanten (das jedoch nicht sehr effizient ist).

Eine unmittelbare Folgerung aus Satz 4.4.6 ist ein (wieder nur mäßig effizientes) Berechnungsverfahren für die inverse Matrix.

4.4.8 Satz *Ist* $\mathfrak{A} \in \mathrm{GL}(n, K)$*, so gilt* $\mathfrak{A}^{-1} = \frac{1}{\det(\mathfrak{A})} \cdot \widetilde{\mathfrak{A}}$.

Wenden wir diese Verfahren auf lineare Gleichungssysteme an, so erhalten wir die

4.4.9 Cramersche Regel *Ist* $\mathfrak{A} \in \mathrm{GL}(n, K)$ *und* $\mathfrak{b} = (\beta_1, \ldots, \beta_n)^t \in K_n$*, so erhalten wir die Lösung des Gleichungssystems* $\mathfrak{A} \cdot \mathfrak{x} = \mathfrak{b}$ *als* $\mathfrak{x} = (\xi_1, \ldots, \xi_n)^t$ *mit*

$$\xi_i = \frac{1}{\det(\mathfrak{A})} \sum_{j=1}^{n} \beta_j \cdot \det(\mathfrak{A}_{ji}).$$

Beweis: Die Lösung des Gleichungssystems ist

$$\begin{aligned} \mathfrak{x}^t = \mathfrak{A}^{-1} \cdot \mathfrak{b} = \frac{1}{\det(\mathfrak{A})} \cdot \widetilde{\mathfrak{A}} \cdot \mathfrak{b} &= \frac{1}{\det(\mathfrak{A})} \cdot (\sum_{j=1}^{n} \beta_j \tilde{\alpha}_{1j}, \ldots, \sum_{j=1}^{n} \beta_j \tilde{\alpha}_{nj})^t \\ &= \frac{1}{\det(\mathfrak{A})} \cdot (\sum_{j=1}^{n} \beta_j \det(\mathfrak{A}_{j1}), \ldots, \sum_{j=1}^{n} \beta_j \det(\mathfrak{A}_{jn}))^t. \end{aligned}$$

□

4.4.10 Bemerkung Einprägsamer läßt sich die Cramersche Regel folgendermaßen formulieren. Sei $\mathfrak{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{pmatrix} \in \mathrm{GL}(n, K)$ die Koeffizientenmatrix des Gleichungssystems

$$\begin{array}{ccccccc} \alpha_{11}\xi_1 & + & \cdots & + & \alpha_{1n}\xi_n & = & \beta_1 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{n1}\xi_1 & + & \cdots & + & \alpha_{nn}\xi_n & = & \beta_n \end{array},$$

so erhalten wir die "eingesetzte Matrix" $\mathfrak{A}_i(\mathfrak{b})$, indem wir die i-te Spalte durch den Vektor $\mathfrak{b} := \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ ersetzen, d.h.

$$\mathfrak{A}_i(\mathfrak{b}) := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,i-1} & \beta_1 & \alpha_{1,i+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,i-1} & \beta_n & \alpha_{n,i+1} & \cdots & \alpha_{n,n} \end{pmatrix}.$$

Dann gilt nach der Cramerschen Regel

$$\xi_i = \frac{\det(\mathfrak{A}_i(\mathfrak{b}))}{\det \mathfrak{A}}.$$

5. Normalformdarstellungen von Endomorphismen

Im Normalformensatz (Satz 3.5.1) haben wir gesehen, daß sich jeder Homomorphismus $f \in \mathrm{Hom}(V, W)$ bezüglich geeignet gewählter Basen durch eine Matrix in Diagonalgestalt darstellen läßt. Für Endomorphismen ist diese Fragestellung deutlich schwieriger zu beantworten. Natürlich können wir den Normalformensatz auf $f \in \mathrm{End}(V) = \mathrm{Hom}(V, V)$ anwenden, aber dann erhalten wir zwei verschiedene Basen von V, bezüglich derer wir f darzustellen haben. Unser Ziel sollte es aber sein, **eine** Basis zu finden. In diesem Kapitel wollen wir studieren, unter welchen Bedingungen dies möglich sein wird und wie die Normalformdarstellungen aussehen können. Dazu benötigen wir allerdings noch einige algebraische Hilfsmittel, die wir zunächst bereitstellen wollen.

5.1 Teilbarkeit in Ringen

5.1.1 Definition Es sei R ein Ring mit Einselement 1 (vgl. Definition 1.1.13). Ein Element $u \in R$ heißt *invertierbar* oder eine *Einheit* in R, wenn es ein $w \in R$ mit $u \cdot w = w \cdot u = 1$ gibt.

5.1.2 Folgerung *(i) Ist $u \in R$ invertierbar, so ist dessen inverses Element eindeutig bestimmt. Wir bezeichnen es mit u^{-1}. Es gilt insbesondere $1^{-1} = 1$.*
(ii) Sind $u, v \in R$ beide invertierbar, so ist auch $u \cdot v$ invertierbar, und es gilt $(u \cdot v)^{-1} = v^{-1} \cdot u^{-1}$.
(iii) Ist $u \in R$ invertierbar, so ist auch u^{-1} invertierbar, und es gilt $(u^{-1})^{-1} = u$.

Beweis: (i) Nehmen wir $uv = vu = 1$ und $uw = wu = 1$ an, so erhalten wir $v = v \cdot 1 = v(uw) = (vu)w = w$. Offensichtlich ist $1 \cdot 1 = 1$ und damit $1^{-1} = 1$.
(ii) Dies folgt durch Nachrechnen.
(iii) Dies ist klar. □

Es folgt unmittelbar aus 5.1.2, daß die invertierbaren Elemente eines Ringes eine Gruppe bilden. Dies wollen wir in einem Satz festhalten.

5.1.3 Satz *Sei R ein Ring mit 1. Die Menge $R^* := \{u \in R \mid u$ ist invertierbar $\}$ bildet mit der Multiplikation des Ringes eine Gruppe.*

5.1.4 Beispiel Wir wissen aus Satz 2.5.4, daß die Menge $\mathrm{End}(V)$ der Endomorphismen eines Vektorraumes einen Ring bildet. Die Menge der invertierbaren Elemente ist offensichtlich die Gruppe $\mathrm{GL}(V)$, d.h. wir haben $\mathrm{End}(V)^* = \mathrm{GL}(V)$.

5.1.5 Definition Sei R ein kommutativer Ring mit 1. Eine Menge $I \subseteq R$ heißt ein *Ideal*, wenn gilt:

(I.1) *I ist eine Untergruppe der additiven abelschen Gruppe von R.*

(I.2) *Für $u \in I$ und $r \in R$ ist $u \cdot r \in I$.*

Beachte Überzeugen Sie sich davon, daß jedes Ideal von R ein Teilring von R ist. Umgekehrt muß **nicht** jeder Teilring auch ein Ideal sein.

5.1.6 Bemerkung Verzichtet man auf die Voraussetzung der Kommutativität, so hat man in der naheliegenden Weise *linke*, *rechte* und *beidseitige* Ideale zu definieren, indem man $(\forall u \in I)(\forall r \in R)[r \cdot u \in I]$, $(\forall u \in I)(\forall r \in R)[u \cdot r \in I]$ und $(\forall u \in I)(\forall r \in R)[r \cdot u \in I \wedge u \cdot r \in I]$ fordert.

5.1.7 Lemma *Ist R ein kommutativer Ring mit 1 und sind $a_1, \ldots, a_n \in R$, so ist die Menge*

$$(a_1, \ldots, a_n)_R := \{\sum_{i=1}^{n} r_i a_i \mid r_i \in R\}$$

ein Ideal in R. Sie heißt das durch $a_1, \ldots, a_n$ erzeugte *Ideal. Wann immer der Zusammenhang es erlaubt, lassen wir den unteren Index R weg.*

Beweis: Offensichtlich ist $\sum_{i=1}^{n} r_i a_i - \sum_{i=1}^{n} r_i' a_i = \sum_{i=1}^{n}(r_i - r_i')a_i \in (a_1, \ldots, a_n)_R$ und $r \cdot \sum_{i=1}^{n} r_i a_i = \sum_{i=1}^{n} r r_i a_i \in (a_1, \ldots, a_n)_R$, womit ein Ideal vorliegt. □

5.1.8 Definition Ein Ideal $I \subseteq R$ heißt ein *Hauptideal*, wenn es ein $a \in R$ gibt mit $I = (a)_R = \{r \cdot a \mid r \in R\}$.

5.1.9 Beispiele Ist R ein kommutativer Ring, so haben wir das *Nullideal* $\mathbf{0} := \{0\}$ als ein (triviales) Hauptideal. Ebenso ist $R = (1)_R$ ein Hauptideal. Für eine Einheit $u \in R^*$ ist $1 = u^{-1}u \in (u)_R$ und damit für $r \in R$ auch $r = r \cdot 1 \in (u)_R$. Damit gilt $(u)_R = R$ für jede Einheit $u \in R^*$.

Wir erinnern an die Definition eines Nullteilers (Definition 1.1.15). Dort haben wir auch den Begriff des Integritätsring eingeführt. Wir wollen diesen jetzt noch etwas enger fassen und sagen, daß ein *Integritätsbereich* ein **kommutativer**, nullteilerfreier Ring mit 1 ist.

5.1.10 Definition Es sei R ein Integritätsbereich. Ein Element $a \in R$ heißt *Teiler* eines Elements $b \in R$, wenn es ein $r \in R$ gibt mit $b = a \cdot r$. In Formelschreibweise bedeutet dies

$$a \mid b \;\; :\Leftrightarrow \;\; (\exists r \in R)[b = a \cdot r].$$

Wir schreiben $a \nmid b$ um auszudrücken, daß a kein Teiler von b ist.

5.1.11 Folgerungen *Für einen Integritätsbereich R gelten:*

(i) $(\forall a \in R)[1 \mid a \wedge a \mid a]$.

(ii) $(\forall a, b \in R)[a \mid b \wedge b \mid c \;\Rightarrow\; a \mid c]$.

(iii) $a \mid b_1 \wedge \ldots \wedge a \mid b_n \;\Rightarrow\; a \mid \sum_{i=1}^{n} r_i b_i$ *für alle* $r_1, \ldots, r_n \in R$.

(iv) $a \mid 1 \;\Leftrightarrow\; a \in R^*$.

(v) $a \mid b \wedge b \mid a \;\Leftrightarrow\; (\exists u \in R^*)[a = bu]$.

(vi) $a \mid b \;\Leftrightarrow\; (b)_R \subseteq (a)_R$.

(vii) $(a)_R = (b)_R \;\Leftrightarrow\; (\exists u \in R^*)[a = bu]$.

Beweis: Die Behauptung (i) ist klar. Wir zeigen (ii) durch die Rechnung $b = ar_1 \wedge c = br_2 \;\Rightarrow\; c = a(r_1 r_2)$ und (iii) durch die Rechnung $b_i = ac_i$ für $i = 1, \ldots n$, welche $\sum_{i=1}^{n} r_i b_i = \sum_{i=1}^{n} r_i a c_i = a \cdot \sum_{i=1}^{n} r_i c_i$ impliziert. Die Behauptung (iv) ist klar. Bei (v) zeigen wir zunächst $\Rightarrow$: Aus $a = br_1$ und $b = ar_2$ folgt $b = b(r_1 r_2)$ und damit $r_1 r_2 = 1$, d.h. $r_i \in R^*$ für $i = 1, 2$. Die Gegenrichtung ist wegen $a = bu$ und $b = au^{-1}$ trivial. Für (vi) zeigen wir zunächst $\Rightarrow$: Sei dazu $x \in (b)_R$ und $b = a \cdot r_1$. Dann ist $x = b \cdot r = a \cdot r_1 \cdot r \in (a)_R$. Für die Gegenrichtung $\Leftarrow$ erhalten wir $b = b \cdot 1 \in (b)_R \subseteq (a)_R$ und damit $b = a \cdot r$ für ein $r \in R$. D.h. $a \mid b$. Die Eigenschaft (vii) folgt nun unmittelbar wegen $(a)_R = (b)_R \;\Leftrightarrow\; (a)_R \subseteq (b)_R \wedge (b)_R \subseteq (a)_R \;\Leftrightarrow\; a \mid b \wedge b \mid a \;\Leftrightarrow\; (\exists u \in R^*)[a = b \cdot u]$. ☐

5.1.12 Satz *Die Relation $a \sim_R b \;\; :\Leftrightarrow \;\; (\exists u \in R^*)[a = b \cdot u]$ definiert eine Äquivalenzrelation auf R. Gilt $a \sim_R b$, so heißen a und b in R* assoziiert.

Beweis: Da R^* eine Gruppe ist, erhalten wir $a \sim_R a$ wegen $a = 1 \cdot a$. Aus $a \sim_R b$ mit $a = b \cdot u$ erhalten wir $b = a \cdot u^{-1}$ und damit $b \sim_R a$. Schließlich folgt aus $a \sim_R b \sim_R c$ mit $a = b \cdot u_1$ und $b = c \cdot u_2$ auch $a = c \cdot u_1 \cdot u_2$, d.h. $a \sim_R c$. □

Beachte *Nach dem Hilfsatz im Beweis von Lemma 2.2.8 bilden die Äquivalenzklassen von $\sim_R$ eine Partition von R.*
Aus 5.1.11 (v) erhalten wir

$$a \sim_R b \quad \Leftrightarrow \quad a \mid b \wedge b \mid a \quad \Leftrightarrow \quad (a)_R = (b)_R. \tag{5.1}$$

5.1.13 Definition Ein Integritätsbereich R heißt ein *Hauptidealring*, wenn jedes Ideal von R schon ein Hauptideal ist.

Sind $a_1, \ldots, a_n \in R$, so heißt $r \in R$ ein *gemeinsamer Teiler* , wenn $r \mid a_i$ für $i = 1, \ldots, n$ gilt. Gilt $r \in R^*$ für jeden gemeinsamen Teiler von $a_1, \ldots, a_n$, so heißen $a_1, \ldots, a_n$ *teilerfremd* oder *relativ prim*.

5.1.14 Satz *Sei R ein Hauptidealring. Dann sind äquivalent:*

(i) $a_1, \ldots, a_n \in R$ *sind teilerfremd.*

(ii) $(a_1, \ldots, a_n)_R = R$.

(iii) $(\exists r_1, \ldots, r_n \in R)[\sum_{i=1}^n r_i a_i = 1]$.

Beweis: (i) $\Rightarrow$ (ii): Da R ein Hauptidealring ist, gibt es ein Element $r \in R$ mit $(a_1, \ldots, a_n)_R = (r)_R$. Wegen $a_i \in (r)_R$ folgt $a_i = r \cdot r_i$ für $i = 1, \ldots, n$. Damit ist r ein gemeinsamer Teiler von $a_1, \ldots, a_n$ und, da diese teilerfremd sind, somit $r \in R^*$. Nach 5.1.9 folgt $(a_1, \ldots, a_n)_R = (r)_R = R$.
(ii) $\Rightarrow$ (iii): Da $1 \in (a_1, \ldots, a_n)_R$ gibt es $r_1, \ldots, r_n \in R$ mit $1 = \sum_{i=1}^n r_i a_i$.
(iii) $\Rightarrow$ (i): Gilt $r \mid a_i$ für $i = 1, \ldots, n$, so folgt nach 5.1.11 (iii) $r \mid (\sum_{i=1}^n r_i a_i)$ für alle $r_1, \ldots, r_n \in R$ und damit $r \mid 1$. Also ist $r \in R^*$. □

5.1.15 Satz *Ist R ein Hauptidealring und sind $a, b \in R$ teilerfremd, so folgt aus $a \mid b \cdot r$ bereits $a \mid r$.*

Beweis: Da a und b teilerfremd sind, gibt es nach Satz 5.1.14 Elemente $r_1, r_2 \in R$ mit $1 = a \cdot r_1 + b \cdot r_2$. Damit ist $r = a \cdot r_1 \cdot r + b \cdot r_2 \cdot r$, und wegen $a \mid b \cdot r$ folgt nach Satz 5.1.11 (iii) $a \mid r$. □

Wir nennen $p \in R$ ein *Primelement*, wenn

$$p \neq 0 \wedge p \notin R^* \wedge (\forall d \in R)[d \mid p \quad \Rightarrow \quad d \sim_R p \vee d \sim_R 1] \tag{5.2}$$

gilt, d.h. wenn jeder Teiler von p entweder eine Einheit oder mit p assoziiert ist.

5.1.16 Satz *Es sei R ein Hauptidealring und p ein Primelement von R. Dann gelten:*

(i) *Für $a \in R$ gilt entweder $p \mid a$ oder p und a sind teilerfremd.*

(ii) *Aus $p \mid a \cdot b$ folgt $p \mid a$ oder $p \mid b$.*

Beweis: (i): Sind p und a nicht teilerfremd, so gibt es ein $r \in R \setminus R^*$ mit $r \mid a$ und $r \mid p$. Dann ist aber $r \sim_R p$, und wegen $r \mid a$ erhalten wir damit auch $p \mid a$.
(ii): Nach (i) gilt: $p \mid a$ oder p und a sind teilerfremd. In diesem Falle folgt aber aus $p \mid ab$ nach Satz 5.1.15 $p \mid b$. $\square$

5.1.17 Korollar *Zwei Primelemente in einem Hauptidealring sind entweder teilerfremd oder assoziiert.*

Beweis: Sind p und q Primelemente, die nicht teilerfremd sind, so folgt nach Satz 5.1.16 $p \mid q$ und $q \mid p$. Nach (5.1) folgt daraus $p \sim_R q$. $\square$

5.1.18 Bemerkung Wir erhalten aus Satz 5.1.15, daß für paarweise teilerfremde Elemente $a_1, \ldots, a_n$ mit $a_i \mid b$ für $i = 1, \ldots, n$ auch $\prod_{i=1}^n a_i \mid b$ gilt. Aus $a_1 \mid b$ folgt nämlich $b = a_1 \cdot r_1$ und aus $a_2 \mid a_1 \cdot r_1$ wegen der Teilerfremdheit von a_2 und a_1 dann $a_2 \mid r_1$, d.h. $b = a_1 \cdot a_2 \cdot r_2$. Nun erhalten wir in analoger Weise $a_3 \mid r_2$ und durch Iteration schließlich $b = \prod_{i=1}^n a_i \cdot r_n$.

5.2 Euklidische Ringe

Das Beispiel für einen Integritätsbereich ist $\mathbb{Z}$, der Ring der ganzen Zahlen. In $\mathbb{Z}$ gibt es nur zwei Einheiten $\{-1, 1\}$. Ein Ideal in $\mathbb{Z}$ ist beispielsweise die Menge $k \cdot \mathbb{Z} := \{k \cdot r \,|\; r \in \mathbb{Z}\}$. Wir werden in Kürze sehen, daß $\mathbb{Z}$ ein Hauptidealring ist. Damit haben alle Ideale die Gestalt $k \cdot \mathbb{Z}$. Der Ring $\mathbb{Z}$ hat aber noch zusätzliche Eigenschaften. Eine davon ist die Division mit Rest, die den euklidischen Algorithmus zum Auffinden des größten gemeinsamem Teilers zweier Zahlen ermöglicht. Wie Sie sich erinnern, besagt die Division mit Rest, daß zu zwei gegebenen Zahlen a und b eine Zahl r so existiert, daß der Rest $b - ar$ kleiner als der Teiler a wird. Da wir auf beliebigen Ringen im allgemeinen keine “kleiner” Relation haben, müssen wir uns anderweitig behelfen. Dies ist der Hintergrund der folgenden Definition.

5.2.1 Definition Es sei R ein Integritätsbereich. Eine Abbildung $N \colon R \longrightarrow \mathbb{Z}$ heißt eine *euklidische Norm*, wenn sie den folgenden Bedingungen genügt:

(EN.0) $(\forall a \in R)[N(a) = 0 \;\Leftrightarrow\; a = 0]$.

(EN.1) $(\forall a \in R)[a \neq 0 \;\Rightarrow\; N(a) > 0]$.

(EN.2) $(\forall a, b \in R)[N(ab) = N(a)N(b)]$.

(EN.3) $(\forall a, b \in R \setminus \{0\})(\exists r \in R)[N(b - ar) < N(a)]$.

Ein Integritätsbereich, auf dem sich eine euklidische Norm definieren läßt, heißt ein *euklidischer Ring*.

5.2.2 Folgerung *Ist R ein euklidischer Ring mit euklidischer Norm N, so gelten:*

(i) $b \mid a \;\Rightarrow\; N(b) \mid N(a)$,

denn aus $a = br$ folgt $N(a) = N(b)N(r)$.

(ii) $a \sim_R b \;\Rightarrow\; N(a) = N(b)$,

denn aus $a \sim_R b$ folgt $a \mid b$ und $b \mid a$ und daraus $N(a) \mid N(b)$ und $N(b) \mid N(a)$. Damit gilt $N(a) \sim_{\mathbb{Z}} N(b)$ und, da $N(a)$ und $N(b)$ beide nicht negativ und $\{-1, 1\}$ die einzigen Einheiten in $\mathbb{Z}$ sind, schließlich $N(a) = N(b)$.

(iii) $N(1) = 1$,

denn nach (EN.2) ist $N(r) = N(1 \cdot r) = N(1)N(r)$ und damit $N(1) = 1$.

(iv) $u \in R^* \;\Leftrightarrow\; N(u) = 1$,

denn $u \in R^*$ impliziert $u \mid 1$ und damit $N(u) \mid 1$ nach (iii). Da $0 < N(u)$ folgt $N(u) = 1$. Ist umgekehrt $N(u) = 1$, so finden wir nach (EN.3) ein $r \in R$ mit $N(1 - ur) < N(1) = 1$. Dann ist $1 - ur = 0$ und damit $ur = 1$, d.h. $u \in R^*$.

(v) $a \mid b \wedge a \not\sim_R b \wedge b \neq 0 \;\Rightarrow\; N(a) < N(b)$,

denn aus $b = ar$ mit $r \notin R^* \cup \{0\}$ folgt $N(b) = N(a)N(r)$ mit $N(r) > 1$. Also ist $N(a) < N(b)$.

5.2.3 Satz *Jeder euklidische Ring ist ein Hauptidealring.*

Beweis: Sei R ein euklidischer Ring und $I \subseteq R$ ein Ideal. Da das Nullideal ein Hauptideal ist, dürfen wir $I \neq \mathbf{0}$ annehmen. Sei $n := \min\{N(a) \mid a \in I \wedge a \neq 0\}$ und $a_0 \in I$ so gewählt, daß $N(a_0) = n$ ist. Wegen $n \neq 0$ folgt $a_0 \neq 0$. Zu jedem $b \in I$ gibt es nach (EN.3) ein $r \in R$ mit $N(b - a_0 r) < N(a_0)$. Da $b - a_0 r \in I$ folgt daraus $b = a_0 r$, d.h. $b \in (a_0)_R$. Damit ist $I \subseteq (a_0)_R$ und die entgegengesetzte Inklusion folgt sofort wegen $a_0 \in I$. $\square$

5.2.4 Satz *Ist R ein euklidischer Ring, so läßt sich jedes von 0 verschiedene $r \in R \setminus R^*$ bis auf Assoziiertheit eindeutig als Produkt von Primelementen darstellen. Ringe, deren Elemente sich bis auf Assoziiertheit eindeutig als Primzahlprodukte darstellen lassen, heißen* faktoriell. *Jeder euklidische Ring ist somit faktoriell.*

Beweis: Wir haben zwei Dinge zu zeigen:

$$\text{Zu } r \in R \setminus (R^* \cup \{0\}) \text{ gibt es Primelemente } p_1, \ldots, p_n \text{ mit } r = \prod_{i=1}^{n} p_i \tag{i}$$

und

$$\text{Ist } \prod_{i=1}^{n} p_i \sim_R \prod_{i=1}^{m} q_i \text{ für Primelemente } p_1, \ldots, p_n \text{ und } q_1, \ldots, q_m, \text{ so ist } m = n \text{ und } p_i \sim_R q_{\pi(i)} \text{ für eine Permutation } \pi \in S_n. \tag{ii}$$

Beginnen wir mit (i). Wir führen Induktion nach $N(r)$. Wegen $r \notin R^*$ und 5.2.2 (iv) ist $N(r) \neq 1$. Ist r bereits ein Primelement, so wählen wir $n = 1$. Anderenfalls gibt es $a, b \in R \setminus R^*$ mit $r = ab$. Dann ist nach 5.2.2 (v) $N(a) < N(r)$ und $N(b) < N(r)$. Nach Induktionsvoraussetzung sind a und b als Primzahlprodukte darstellbar. Damit ist aber auch ab ein Primzahlprodukt.
Nun zeigen wir (ii). Da R euklidisch ist, ist R ein Hauptidealring. Aus $p_1 \mid \prod_{i=1}^{m} q_i$ folgt nach Satz 5.1.16 $p_1 \mid q_{j_1}$ für ein $j_1 \in \{1, \ldots, m\}$. Wir setzen $\pi(1) := j1$ und erhalten $p_1 \sim_R q_{\pi(1)}$ nach Korollar 5.1.17 und $\prod_{i=1}^{n} p_i \sim_R \prod_{\substack{i=1 \\ i \neq j_1}}^{m} q_i$. Nun erhalten wir $p_2 \mid q_{j_2}$ für ein $j_2 \in \{1, \ldots, m\} \setminus \{j_1\}$. Wir setzen $\pi(2) := j_2$, erhalten $p_2 \sim_R q_{\pi(2)}$ und durch Iteration schließlich $n = m$ und $p_i \sim_R q_{\pi(i)}$ für $i = 1, \ldots, n$. □

5.2.5 Satz *Der Ring $\mathbb{Z}$ der ganzen Zahlen ist euklidisch und damit faktoriell und ein Hauptidealring.*

Beweis: Man prüft sofort nach, daß die Abbildung $z \longmapsto |z|$ eine euklidische Norm definiert. □

5.2.6 Bemerkung *In der Algebra-Vorlesung wird gezeigt werden, daß bereits jeder Hauptidealring faktoriell ist. Das ist jedoch deutlich aufwendiger zu zeigen und wird in dieser Vorlesung nicht gebraucht.*

5.3 Der Polynomring $K[x]$

Wir erinnern an den Abschnitt 1.3.3, in dem wir den Vektorraum $\mathrm{Abb}[\mathbb{N}, K] =: K[x]$ der Polynome über einem Körper K eingeführt haben. Wir hatten

$$G(f) := \begin{cases} \min\{n \in \mathbb{N} \mid (\forall m \in \mathbb{N})[m > n \Rightarrow f(m) = 0]\} & \text{für } f \neq 0 \\ -\infty & \text{für } f = 0 \end{cases}$$

definiert und $G(f)$ den *Grad* von f genannt. Für das zusätzliche Zeichen $-\infty$ vereinbaren wir die folgenden Rechenregeln:

$$\begin{aligned} &-\infty + (-\infty) = -\infty \\ &-\infty + a = a + (-\infty) = -\infty. \end{aligned}$$

Jedes Polynom vom Grad $n \geq 0$ läßt sich in der Form $f = \sum_{i=0}^{n} f(i)x^i$ mit einer *Unbestimmten* x schreiben, was den Vorteil hat, daß man mit diesen formalen Summen wie gewohnt rechnen kann. Dies kommt insbesondere zum Tragen, wenn wir das Produkt zweier Polynome

$$f \cdot g := \sum_{k=0}^{m+n} \Big(\sum_{i+j=k} f(i)g(j) \Big) x^k \tag{5.3}$$

einführen, denn dieses Produkt erhält man, wenn man die Ausdrücke $\sum_{i=0}^{m} f(i)x^i$ und $\sum_{j=0}^{n} g(j)x^j$ formal miteinander multipliziert und nach Potenzen von x ordnet. Offensichtlich hat man dann

$$G(f \cdot g) = G(f) + G(g). \tag{5.4}$$

Man rechnet sofort nach, daß die in (5.3) definierte Multiplikation assoziativ und kommutativ ist und den Distributivgesetzen gehorcht. Damit ist $K[x]$ ein Ring. Wir nennen ihn den *Ring der Polynome mit Koeffizienten in* K oder kurz den *Polynomring* über K.

Für jedes $\alpha \in K$ läßt sich eine Abbildung

$$f\colon \mathbb{N} \longrightarrow K, \quad p_\alpha(n) := \begin{cases} \alpha & \text{für } n = 0 \\ 0 & \text{sonst} \end{cases}$$

erklären, die fast überall verschwindet und somit ein Polynom ist. Offensichtlich ist dann $G(p_\alpha) = 0$. Die Abbildung $\alpha \longmapsto p_\alpha$ ist offensichtlich injektiv. Wir können daher K als Teilmenge von $K[x]$ auffassen, indem wir α und p_α identifizieren. Dies wird noch deutlicher, wenn wir $p_\alpha = \sum_{i=0}^{0} \alpha x^i = \alpha$ notieren.

5.3.1 Definition Jedes Polynom $f = \sum_{i=0}^{n} \alpha_i x^i \in K[x]$ definiert eine Abbildung $K \ni \xi \mapsto f(\xi) := \sum_{i=0}^{n} \alpha_i \xi^i \in K$, die wir ebenfalls mit f bezeichnen wollen.

Dies sollte nicht zu Verwirrungen mit der Abbildung $f\colon \mathbb{N} \longrightarrow K$ führen, da beide unterschiedliche Argumente haben. Gilt $f(\xi) = 0$ für ein $\xi \in K$, so heißt ξ eine *Nullstelle* oder *Wurzel* von f in K.

5.3.2 Bemerkung *Fassen wir $\alpha \in K$ vermöge der Einbettung von K in $K[x]$ als Polynom auf, so gilt $\alpha(\xi) = \alpha \cdot \xi^0 = \alpha$ für alle $\xi \in K$. Daher bezeichnet man die $\alpha \in K$ auch als* konstante Polynome. *Offenbar sind die konstanten Polynome genau die Polynome vom Grad* 0.

5.3.3 Lemma *Ist $f \in K[x]$ und ist ξ eine Wurzel von f in K, so gibt es ein Polynom $g \in K[x]$ mit $f = (x - \xi) \cdot g$. Insbesondere ist $G(g) = G(f) - 1$.*

Beweis: Wir erhalten $f = f - f(\xi) = \sum_{i=0}^n \alpha_i x^i - \sum_{i=0}^n \alpha_i \xi^i = \sum_{i=0}^n \alpha_i \cdot (x^i - \xi^i) = (x - \xi) \cdot \sum_{i=0}^n \alpha_i (\sum_{k=1}^i x^{i-k}\xi^{k-1})$, und wir setzen $g := \sum_{i=0}^n \alpha_i (\sum_{k=1}^i x^{i-k}\xi^{k-1})$. Damit ist $G(f) = G(x - \xi) + G(g) = 1 + G(g)$. □

5.3.4 Satz *Ist $f \in K[x], f \neq 0$, und sind $\xi_1, \ldots, \xi_n$ alle Wurzeln von f in K, so gibt es ein Polynom $r \in K[x]$, das keine Wurzeln in K hat, und $k_1, \ldots, k_n \in \mathbb{N}$ mit*

$$f = (x - \xi_1)^{k_1} \cdots (x - \xi_n)^{k_n} \cdot r, \tag{5.5}$$

und es ist $G(f) = \sum_{i=1}^n k_i + G(r)$.

Beweis: Wir führen Induktion nach $G(f)$. Besitzt f keine Wurzeln in K, so sei $r := f$ und $n = 0$. Besitzt f eine Wurzel ξ, so gilt nach Lemma 5.3.3 $f = (x-\xi)\cdot g$ und $G(g) < G(f)$. Nach Induktionsvoraussetzung ist dann $g = (x - \xi_1)^{k_1} \cdots (x - \xi_m)^{k_m} \cdot r$, wobei $\xi_1, \ldots, \xi_m$ alle Wurzeln von g in K sind und r keine Wurzel in K besitzt. Dann ist $f = (x - \xi) \cdot (x - \xi_1)^{k_1} \cdots (x - \xi_m)^{k_m} \cdot r$, und $\{\xi\} \cup \{\xi_1, \ldots, \xi_m\}$ sind alle Wurzeln von f in K. Ist $\xi = \xi_i$ für ein $i \in \{1, \ldots, m\}$, so ist $n = m$, und der Exponent k_i erhöht sich um 1. Anderenfalls ist $n = m + 1$, und $(x - \xi)$ kommt als neuer Faktor hinzu. Schließlich erhalten wir $G(f) = G((x - \xi_1)^{k_1}) + \cdots + G((x - \xi_n)^{k_n}) + G(r) = \sum_{i=1}^n k_i + G(r)$. □

Wir werden in Kürze feststellen, daß $K[x]$ ein faktorieller Ring ist und jedes Polynom der Form $x - \xi$ ein Primelement in $K[x]$ ist. Damit sind die k_i in (5.5) eindeutig bestimmt. Die Zahl k_i heißt dann die *Vielfachheit* der Wurzel ξ_i in K. Ist r in (5.5) ein konstantes Polynom, so sagen wir, daß f über K in *Linearfaktoren* zerfällt. Ein Körper K, über dem jedes Polynom $f \in K[x]$ in Linearfaktoren zerfällt, heißt *algebraisch abgeschlossen*. In der Vorlesung über Algebra wird gezeigt, daß jeder Körper einen algebraisch abgeschlossenen Oberkörper besitzt. Ein Beispiel für einen algebraisch abgeschlossenen Körper ist der Körper $\mathbb{C}$ der komplexen Zahlen.

5.3.5 Korollar *Ist $f \in K[x], f \neq 0$, und $G(f) = n$, so besitzt f über K höchstens n viele Wurzeln, wobei jede Wurzel in ihrer Vielfachheit gezählt wird. Zerfällt f über K, so hat f genau n Wurzeln in K.*

Beweis: Nach Satz 5.3.4 ist die Anzahl der Wurzeln von f in K in ihrer Vielfachheit gezählt $\sum_{i=1}^{n} k_i = G(f) - G(r)$. Damit ist diese Anzahl kleiner als $G(f)$ und gleich $G(f)$ genau dann, wenn $G(r) = 0$ ist, d.h. wenn f in Linearfaktoren zerfällt. □

5.3.6 Lemma *Es ist $K[x]^* = \{f \mid G(f) = 0\} = K^*$.*

Beweis: Es folgt aus Definition (5.3), daß 1 das 1-Element in $K[x]$ ist. Aus $f \cdot g = 1$ erhalten wir wegen (5.4) $G(f) + G(g) = G(1) = 0$ und daher $G(f) = G(g) = 0$. Damit sind f und g konstant und somit in K. Da K ein Körper ist, folgt aus $fg \neq 0$ auch $f \neq 0$ und $g \neq 0$. Also ist $f, g \in K^*$. Die umgekehrte Inklusion gilt aber wegen $\alpha^{-1}\alpha = 1$ für $\alpha \in K^*$ trivialerweise. □

5.3.7 Satz *Der Ring $K[x]$ über einem Körper K ist euklidisch.*

Beweis: Bis jetzt haben wir eingesehen, daß $K[x]$ ein kommutativer Ring mit 1 ist. Aus (5.3) erhalten wir aber auch sofort $f \cdot g \neq 0$, wenn $f \neq 0$ und $g \neq 0$ ist. Wir definieren nun

$$N(f) := \begin{cases} 0 & \text{falls } f = 0 \\ 2^{G(f)} & \text{sonst.} \end{cases}$$

Dann gilt $N(f) = 0 \iff f = 0$ und $f \neq 0 \Rightarrow N(f) > 0$. Ebenso erhalten wir $N(fg) = 2^{G(f)+G(g)} = 2^{G(f)} \cdot 2^{G(g)} = N(f) \cdot N(g)$ und haben damit die Eigenschaften *(EN.0) – (EN.2)* einer euklidischen Norm für N nachgeprüft. Es bleibt zu zeigen, daß wir zu $f, g \in K[x]$ ein Polynom p so finden können, daß $N(f - gp) < N(g)$ ist. Seien also f und g gegeben. Ist $G(f) < G(g)$, so setzen wir $p = 0$ und sind fertig. Daher nehmen wir $f = \sum_{i=0}^{n} \alpha_i x^i$ und $g = \sum_{i=0}^{m} \beta_i x^i$ mit $n \geq m$ und $\beta_m \neq 0$ an. Sei $r_1 := f - g \cdot \frac{\alpha_n}{\beta_m} x^{n-m}$. Dann ist $G(r_1) < G(f)$. Gilt auch $G(r_1) < G(g)$, so setzen wir $p := \frac{\alpha_n}{\beta_m} x^{n-m}$ und sind fertig. Anderenfalls ist $r_1 = \sum_{i=0}^{n-1} \alpha_i' x^i$ mit $n - 1 \geq m$ und wir setzen $r_2 := r_1 - g \cdot \frac{\alpha_{n-1}'}{\beta_m} \cdot x^{n-1-m}$. Dann ist $G(r_2) = n - 2 < G(r_1)$. Ist $G(r_2) < G(g)$, so haben wir $r_2 = f - g(\frac{\alpha_n}{\beta_m} x^{n-m} + \frac{\alpha_{n-1}'}{\beta_m} \cdot x^{n-1-m})$ und setzen $p := \frac{\alpha_n}{\beta_m} x^{n-m} + \frac{\alpha_{n-1}'}{\beta_m} \cdot x^{n-1-m}$. Anderenfalls setzen wir das Verfahren fort, definieren r_3 usf. Da die Grade der Polynome r_i strikt abnehmen, erhalten wir schließlich ein $r_k = f - g(\gamma_1 x^{n-m} + \gamma_2 x^{n-1-m} + \cdots + \gamma_k x^{n-m-k+1})$ mit $G(r_k) < G(g)$. Wir setzen $p := \gamma_1 x^{n-m} + \gamma_2 x^{n-1-m} + \cdots + \gamma_k x^{n-m-k+1}$ und sind fertig. □

5.3.8 Satz *Der Polynomring $K[x]$ über einem Körper K ist ein Hauptidealring und faktoriell.*

Beweis: Dies folgt mit den Sätzen 5.2.3 und 5.2.4 aus Satz 5.3.7. $\square$

5.3.9 Lemma *Ein Polynom $f = \sum_{i=0}^{n} \alpha_i x^i$ vom Grad n heißt normiert, wenn $\alpha_n = 1$ ist. Jedes Polynom $f \neq 0$ ist zu einem normierten Polynom assoziiert.*

Beweis: Ist $\alpha_n \neq 1$, so gilt nach Lemma 5.3.6: $f \sim {\alpha_n}^{-1} f$ und ${\alpha_n}^{-1} f$ ist normiert. $\square$

5.3.10 Lemma *Zwei normierte Polynome sind genau dann assoziiert, wenn sie gleich sind.*

Beweis: Sei $x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_0 = f \sim g = x^n + \beta_{n-1}x^{n-1} + \cdots + \beta_0$. Dann gibt es nach Lemma 5.3.6 ein $\alpha \in K^*$ mit $\alpha \cdot f = g$. Dann ist $\alpha \cdot x^n = x^n$ und damit $\alpha = 1$. Damit folgt $\beta_i = \alpha\alpha_i = \alpha_i$ für $i = 0, \ldots, n-1$ und damit $f = g$. Die Gegenrichtung ist klar. $\square$

5.3.11 Satz *Ist $I \subseteq K[x]$ ein Ideal, das nicht das Nullideal ist, so gibt es ein eindeutig bestimmtes normiertes Polynom p mit $I = (p)_{K[x]}$.*

Beweis: Da $K[x]$ ein Hauptidealring und $I \neq (0)$ ist, gibt es ein normiertes Polynom p mit $I = (p)_{K[x]}$. Gilt $I = (q)_{K[x]}$ für ein weiteres normiertes Polynom, so folgt $p \sim q$ nach (5.1). Nach Lemma 5.3.10 folgt damit $p = q$. $\square$

5.3.12 Definition Die Primelemente in $K[x]$ nennt man *irreduzible Polynome*. Der Name rührt daher, daß sich ein irreduzibles Polynom nicht als Produkt von Polynomen kleinerer Grade darstellen läßt. Ist nämlich f ein Primelement und $f = p \cdot q$, so folgt $f \sim p$ oder $p \sim 1$ und damit entweder $G(p) = G(f)$ oder $G(p) = 0$ und damit $G(q) = G(f)$.

5.3.13 Satz *Ist $f \in K[x]$ irreduzibel, so ist $G(f) = 1$ oder f hat keine Wurzel in K.*

Beweis: Dies folgt sofort aus Satz 5.3.3. $\square$

5.3.14 Satz *Zu jedem Polynom f gibt es eindeutig bestimmte irreduzible normierte Polynome $p_1, \ldots, p_n$, natürliche Zahlen $k_1, \ldots, k_n$ größer als 0 und ein $\alpha \in K$ mit $f = \alpha \cdot \prod_{i=1}^{n} p_i^{k_i}$.*

Beweis: Da $K[x]$ euklidisch und damit faktoriell ist, erhalten wir $f = \prod_{i=1}^{n} {p'}_i^{k_i}$ für Primelemente $p'_1, \ldots, p'_n$, die bis auf Assoziiertheit eindeutig bestimmt sind. Normieren wir diese Polynome, so erhalten wir mit Lemma 5.3.10 die Behauptung. □

5.3.15 Bemerkung Jedes Polynom vom Grad 1 ist offensichtlich irreduzibel. Damit erhalten wir, daß die Darstellung in (5.5) tatsächlich eindeutig ist. Insbesondere ist die Vielfachheit einer Wurzel ξ eines Polynoms $f \in K[x]$ in K wohldefiniert. Wir wollen sie in dieser Vorlesung mit $\nu_K(\xi, f)$ bezeichnen, wobei wir im Regelfall den unteren Index K weglassen. Ist $\nu_K(\xi, f) \neq 0$, so gilt nämlich $\nu_{K'}(\xi, f) = \nu_K(\xi, f)$ für alle Oberkörper K' von K.
Ist K ein algebraisch abgeschlossener Körper, so sind die irreduziblen Polynome genau die Polynome vom Grad 1. Damit sind beispielsweise alle Polynome eines Grades > 1 mit reellen (oder auch komplexen) Koeffizienten über $\mathbb{C}$ reduzibel. Über $\mathbb{R}$ gibt es jedoch irreduzible Polynome des Grades 2, z. B. $x^2 + 1$. Über $\mathbb{Q}$ ist sogar das Polynom $x^3 - 2$ irreduzibel (obwohl alle Polynome mit Graden größer als 2 über $\mathbb{R}$ reduzibel sind).

5.4 Eigenwerte und Eigenvektoren von Endomorphismen

5.4.1 Definition Sei V ein Vektorraum über einem Körper K. Ein Element $\lambda \in K$ heißt ein *Eigenwert* eines Endomorphismus $f \in \operatorname{End}(V)$, wenn es einen Vektor $a \neq 0$ in V gibt mit $f(a) = \lambda \cdot a$. Der Vektor a heißt dann ein *Eigenvektor* von f zum Eigenwert λ.

5.4.2 Lemma *Sei V ein K-Vektorraum, $\lambda \in K$ und $f \in \operatorname{End}(V)$. Dann gilt* $\operatorname{Eig}(\lambda, f) := \{v \in V \mid f(v) = \lambda \cdot v\} = \operatorname{Kern}(f - \lambda \cdot id_V)$. *Somit ist* $\operatorname{Eig}(\lambda, f)$ *ein Unterraum von V. Er heißt der* Eigenraum *zum Eigenwert λ von f.*

Beweis: Klar wegen $f(v) = \lambda \cdot v \;\Leftrightarrow\; (f - \lambda \cdot id_v)(v) = 0$. □

5.4.3 Lemma *Sei V ein K-Vektorraum und $f \in \operatorname{End}(V)$. Ein Element $\lambda \in K$ ist genau dann ein Eigenwert von f, wenn die Abbildung $f - \lambda \cdot id_V$ nicht injektiv ist.*

Beweis: Nach Definition 5.4.1 ist λ genau dann ein Eigenwert, wenn $\mathrm{Eig}(\lambda, f) \neq \{0\}$ ist. Nach Lemma 5.4.2 ist dies aber genau dann der Fall, wenn der Kern von $f - \lambda \cdot \mathit{id}_V$ nicht der Nullraum ist, d.h. wenn $f - \lambda \cdot \mathit{id}_V$ nicht injektiv ist. □

5.4.4 Satz *Sei V ein endlich dimensionaler Vektorraum über K. Ein Element $\lambda \in K$ ist genau dann ein Eigenwert eines Endomorphismus $f \in \mathrm{End}(V)$, wenn $\det(f - \lambda \cdot \mathit{id}_V) = 0$ ist.*

Beweis: Nach Lemma 5.4.3 ist λ genau dann ein Eigenwert von $f \in \mathrm{End}(V)$,wenn $f - \lambda \cdot \mathit{id}_V$ nicht injektiv und damit, da V endlich dimensional ist, nicht bijektiv ist. Nach Satz 4.3.6 ist dies aber genau dann der Fall, wenn $\det(f - \lambda \cdot \mathit{id}_V) = 0$ ist. □

Um einen Eigenwert für einen Endomorphismus f zu berechnen, muß man nach Satz 5.4.4 also die Gleichung

$$\det(f - x \cdot \mathit{id}_V) = 0 \tag{5.6}$$

für eine Unbestimmte x lösen. Man nennt daher (5.6) die *charakteristische Gleichung* des Endomorphismus f. Welche Form diese Gleichung hat, wollen wir näher untersuchen.

5.4.5 Satz *Ist V ein n-dimensionaler Vektorraum über einem Körper K, so sind die Lösungen der charakteristischen Gleichung eines Endomorphismus f die Wurzeln eines Polynoms in $K[x]$ vom Grad n.*

Beweis: Wählen wir eine Basis $B = (b_1, \ldots, b_n)$ von V, so gilt nach Lemma 4.3.5

$$\det(f - x \cdot \mathit{id}_V) = \Delta_B((f - x \cdot \mathit{id}_V)(b_1), \ldots, (f - x \cdot \mathit{id}_V)(b_n)). \tag{i}$$

Um die rechte Seite in (i) darstellen zu können, definieren wir für eine Menge $I \subseteq \{1, \ldots, n\}$

$$\Delta_B^I(b_1, \ldots, b_n) := \Delta_B(b_1', \ldots, b_n') \text{ mit } b_i' := \begin{cases} b_i & \text{falls } i \in I \\ f(b_i) & \text{sonst.} \end{cases} \tag{ii}$$

Wegen der Multilinearität von Δ_B folgt dann

$$\Delta_B((f - x \cdot \mathit{id}_V)(b_1), \ldots, (f - x \cdot \mathit{id}_V)(b_n)) = \sum_{k=0}^{n} (-1)^k \cdot \Big(\sum_{\substack{I \subseteq \{1,\ldots,n\} \\ |I| = k}} \Delta_B^I(b_1, \ldots, b_n) \Big) \cdot x^k. \tag{iii}$$

Setzen wir

$$\alpha_k := \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ |I|=k}} \Delta_B^I(b_1, \ldots, b_n), \tag{iv}$$

so erhalten wir nach Lemma 4.3.5

$$\alpha_0 = \Delta_B(f(b_1), \ldots, f(b_n)) = \det f, \tag{v}$$

nach Satz 4.2.6

$$\alpha_n = \Delta_B(b_1, \ldots, b_n) = 1 \tag{vi}$$

und

$$\alpha_{n-1} = \sum_{j=1}^{n} \Delta_B(b_1, \ldots, f(b_j), \ldots, b_n). \tag{vii}$$

Man bezeichnet α_{n-1} als die *Spur* des Endomorphismus f (siehe Bemerkung 5.4.6). Zusammenfassend erhalten wir

$$\det(f - x \cdot id_V) = (-1)^n \cdot x^n + (-1)^{n-1} \operatorname{spur} f \cdot x^{n-1} + \cdots + \det f. \tag{viii}$$

Man nennt das Polynom in (viii) das *charakteristische Polynom* des Endomorphismus f. Wir wollen es mit $Charpol_f$ bezeichnen. Ein Element $\lambda \in K$ ist genau dann Eigenwert eines Endomorphismus f, wenn es Wurzel des charakteristischen Polynoms $Charpol_f$ ist. □

5.4.6 Anmerkung Ist Δ eine Determinantenfunktion und f ein Endomorphismus, so ist die Abbildung

$$(v_1, \ldots, v_n) \longmapsto \sum_{i=1}^{n} \Delta(v_1, \ldots, f(v_i), \ldots, v_n)$$

wieder eine Determinantenfunktion. Daher erhalten wir nach Satz 4.2.7 für eine nichttriviale Determinantenfunktion Δ ein $\alpha \in K$ mit

$$\sum_{i=1}^{n} \Delta(v_1, \ldots, f(v_i), \ldots, v_n) = \alpha \cdot \Delta(v_1, \ldots, v_n).$$

Ähnlich wie bei der Definition der Determinante eines Endomorphimus zeigt man, daß α unabhängig von der Wahl von Δ ist, und definiert

$$\operatorname{spur} f := \frac{\sum_{i=1}^{n} \Delta(v_1, \ldots, f(v_i), \ldots, v_n)}{\Delta(v_1, \ldots, v_n)}. \tag{5.7}$$

Man nennt $\operatorname{spur} f$ die *Spur des Endomorphismus* f. Es ist leicht zu sehen, daß $\operatorname{spur}\colon \operatorname{End}(V) \longrightarrow K$ linear ist und mit etwas mehr Mühe, daß $\operatorname{spur}(f \circ g) = \operatorname{spur}(g \circ f)$ gilt. Aus (5.7) und Satz 4.2.6 folgt, daß für eine Basis $B = (b_1, \ldots, b_n)$

$$\operatorname{spur} f = \sum_{i=1}^{n} \Delta_B(b_1, \ldots, f(b_i), \ldots, b_n) \tag{5.8}$$

ist. Für $\mathfrak{D}_{B,B}(f) = (\alpha_{ij})$ erhalten wir aus (5.8)

$$\operatorname{spur} f = \sum_{i=1}^{n} \alpha_{ii}. \tag{5.9}$$

Man definiert daher die Spur einer $n \times n$-Matrix als die Summe der Elemente in der Hauptdiagonalen.

Als Folgerung aus den Sätzen 5.4.5 und 5.3.4 erhalten wir

5.4.7 Korollar *Ist f ein Endomorphismus eines n dimensionalen Vektorraumes V, so besitzt f höchstens n Eigenwerte. Zerfällt das charakteristische Polynom in Linearfaktoren, so ist $Charpol_f = (x - \lambda_1)^{k_1} \cdots (x - \lambda_m)^{k_m}$, wobei $\lambda_1, \ldots, \lambda_m$ alle verschiedenen Eigenwerte von f sind. Der Exponent k_i heißt dann die* algebraische Vielfachheit *des Eigenwertes λ_i. Weiter ist dann $\det f = \prod_{i=1}^{m} \lambda_i^{k_i}$ und $\operatorname{spur} f = \sum_{i=1}^{m} k_i \cdot \lambda_i$.*

5.5 Diagonalisierung

5.5.1 Anmerkung Gehen wir nun davon aus, daß wir eine Basis $B = (b_1, \ldots, b_n)$ von Eigenvektoren eines Endomorphismus f vorliegen haben, wobei b_i ein Eigenvektor zum Eigenwert λ_i ist, so erhalten wir für die Matrixdarstellung von f

$$\lambda_i \cdot b_i = f(b_i) = \sum_{j=1}^{n} \alpha_{ij} b_j \tag{5.10}$$

und damit

$$\mathfrak{D}_{B,B}(f) = (\lambda_i \delta_{ij}) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}. \tag{5.11}$$

In der Darstellung in (5.11) haben wir gleiche Eigenwerte in ihrer algebraischen Vielfachheit gezählt, d.h. wir fordern nicht, daß alle λ_i verschieden sind.

Unser Ziel einer Normalformdarstellung eines Endomorphismus ist nach (5.11) also erreicht, wenn wir eine Basis aus Eigenvektoren des Endomorphismus finden können. Es geht also nun darum zu untersuchen, unter welchen Bedingungen dies gelingt.

5.5.2 Lemma *Sind $a_1, \ldots, a_m$ Eigenvektoren eines Endomorphismus f, deren zugehörige Eigenwerte paarweise verschieden sind, so ist die Menge $\{a_1, \ldots, a_m\}$ linear unabhängig.*

Beweis: Sei $\sum_{i=1}^m \alpha_i a_i = 0$. Allgemein gilt

$$(f - \lambda_l \cdot \mathit{id}\,)(\sum_{i=1}^m \beta_i a_i) = \sum_{i=1}^m \beta_i(\lambda_i - \lambda_l)a_i = \sum_{\substack{i=1 \\ i \neq l}}^m \beta_i(\lambda_i - \lambda_l)a_i. \tag{i}$$

Durch sukzessive Anwendung von (i) erhalten wir

$$0 = \prod_{\substack{l=1 \\ l \neq k}}^m (f - \lambda_l \cdot \mathit{id}\,)(\sum_{i=1}^m \alpha_i a_i) = \alpha_k \cdot \left(\prod_{\substack{l=1 \\ l \neq k}}^m (\lambda_k - \lambda_l) \right) a_k. \tag{ii}$$

Da wegen der Verschiedenheit der Eigenwerte keiner der Faktoren $\lambda_k - \lambda_l$ verschwindet, folgt aus (ii) $\alpha_k = 0$. Da wir dies für beliebiges $k \in \{1, \ldots, m\}$ erhalten, folgt $\alpha_1 = \cdots = \alpha_m = 0$. □

Nach Lemma 5.5.2 haben wir uns also nur noch um die Eigenvektoren zum gleichen Eigenwert zu kümmern. Zerfällt das charakteristische Polynom eines Endomorphismus f und sind alle Eigenwerte verschieden, so erhalten wir n Eigenvektoren zu verschiedenen Eigenwerten, die nach Lemma 5.5.2 dann bereits eine Basis bilden. Bezüglich dieser Basis läßt sich f, wie in (5.11) angegeben, durch eine Diagonalmatrix darstellen.

5.5.3 Lemma *Sind $\lambda_1, \ldots, \lambda_m$ paarweise verschiedene Eigenwerte eines Endomorphismus f, so gilt*

$$\sum_{i=1}^m \mathrm{Eig}(\lambda_i, f) = \bigoplus_{i=1}^m \mathrm{Eig}(\lambda_i, f).$$

Beweis: Sei $u_i \in \mathrm{Eig}(\lambda_i, f)$ für $i = 1, \ldots, m$. Dann sind nach Lemma 5.5.2 alle von 0 verschiedenen Vektoren u_i linear unabhängig. Aus $\sum_{i=1}^m u_i = 0$ folgt daher $u_i = 0$ für $i = 1, \ldots, m$, denn wären einige der u_i von 0 verschieden, erhielten wir eine nicht triviale Darstellung der 0 aus linear unabhängigen Vektoren, was unmöglich ist. Damit ist die Summe direkt. □

5.5.4 Satz *Ist V ein endlich dimensionaler Vektorraum und ist λ ein Eigenwert eines Endomorphismus $f \in \mathrm{End}(V)$, so gilt*

$$\dim(\mathrm{Eig}(\lambda, f)) \leq \nu(\lambda, \mathit{Charpol}_f). \tag{5.12}$$

Man nennt $\dim(\mathrm{Eig}(\lambda, f))$ *die* geometrische Vielfachheit *des Eigenwerts* λ.

Beweis: Es sei $\{b_1, \ldots, b_s\}$ eine Basis von $\mathrm{Eig}(\lambda, f)$. Ergänzen wir $\{b_1, \ldots, b_s\}$ zu einer Basis B von V, erhalten wir für $\mathfrak{D}_{B,B}(f) = (\alpha_{ij})$ schon

$$f(b_i) = \sum_{j=1}^{n} \alpha_{ij} b_j = \lambda b_i \tag{i}$$

und damit $\alpha_{ij} = \delta_{ij} \cdot \lambda$ für $i = 1, \ldots, s$. Somit ist

$$\mathfrak{D}_{B,B}(f) = \begin{pmatrix} \lambda & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda & 0 \\ * & \cdots & \cdots & * & \mathfrak{B} \end{pmatrix} \tag{ii}$$

und daher

$$\mathfrak{C} := \mathfrak{D}_{B,B}(f - x \cdot id) = \begin{pmatrix} \lambda - x & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda - x & 0 \\ * & \cdots & \cdots & * & \mathfrak{B}' \end{pmatrix}. \tag{iii}$$

Aus Satz 4.3.11 und 4.3.17 (iii) folgt nun

$$\mathit{Charpol}_f = \det \mathfrak{C} = (\lambda - x)^s \cdot \det \begin{pmatrix} \mathfrak{C} & 0 \\ * & \mathfrak{B}' \end{pmatrix} = (\lambda - x)^s \cdot \det \mathfrak{B}'. \tag{iv}$$

Aus (iv) folgt aber $s \leq \nu(\lambda, \mathit{Charpol}_f)$. □

5.5.5 Satz *Es sei V ein endlich dimensionaler Vektorraum über einem Körper K und $f \in \mathrm{End}(V)$. Dann sind äquivalent:*

(i) Es gibt eine Basis B, bezüglich der die Matrix $\mathfrak{D}_{B,B}(f)$ Diagonalgestalt hat.

(ii) Das charakteristische Polynom $\mathit{Charpol}_f$ zerfällt über K in Linearfaktoren, und es gilt $\dim(\mathrm{Eig}(\lambda_i, f)) = \nu(\lambda_i, \mathit{Charpol}_f)$ für jeden Eigenwert λ_i von f.

(iii) Es ist $V = \bigoplus_{i=1}^{m} \mathrm{Eig}(\lambda_i, f)$, wobei $\lambda_1, \ldots, \lambda_m$ alle paarweise verschiedenen Eigenwerte von f sind.

Trifft eine der Eigenschaften auf $f \in \mathrm{End}(V)$ zu, so nennen wir f diagonalisierbar

Beweis: (i) $\Rightarrow$ (ii): Seien $\mu_1, \dots, \mu_n$ die Diagonalelemente in $\mathfrak{D}_{B,B}(f)$. Dann gilt $f(b_i) = \mu_i \cdot b_i$ für $i = 1, \dots, n$, d.h. jedes μ_i ist ein Eigenwert. Damit hat nach Satz 5.4.5 $Charpol_f$ n-Wurzeln und zerfällt daher nach Satz 5.3.4 in Linearfaktoren. Seien nun $\lambda_1, \dots, \lambda_m$ die paarweise verschiedenen Eigenwerte von f. Für $i = 1, \dots, m$ definieren wir

$$I_i := \{j \mid 1 \leq j \leq n \wedge \mu_j = \lambda_i\}. \tag{i}$$

Wir behaupten, daß dann

$$V_i := \langle\{b_j \mid j \in I_i\}\rangle = \mathrm{Eig}(\lambda_i, f) \tag{ii}$$

gilt. Ist nämlich $a \in V_i$, so gilt $a = \sum_{j \in I_i} \alpha_j b_j$ und damit $f(a) = f(\sum_{j \in I_i} \alpha_j b_j) = \sum_{j \in I_i} \alpha_j \lambda_i b_j = \lambda_i \cdot \sum_{j \in I_i} \alpha_j b_j = \lambda_i a$. Damit ist $V_i \subseteq \mathrm{Eig}(\lambda_i, f)$. Ist umgekehrt $a \in \mathrm{Eig}(\lambda_i, f)$, so gilt $a = \sum_{j=1}^{n} \alpha_j b_j$ und damit

$$\sum_{j=1}^{n} \alpha_j \lambda_i b_j = \lambda_i \cdot \sum_{j=1}^{n} \alpha_j b_j = f(a) = \sum_{j=1}^{n} \alpha_j \mu_j b_j. \tag{iii}$$

Aus (iii) erhalten wir aber $\alpha_j \mu_j = \alpha_j \lambda_i$ für $j = 1, \dots, n$. Also ist $\alpha_j = 0$ für $\mu_j \neq \lambda_i$, d.h. für $j \notin I_i$, und wir erhalten $a = \sum_{j \in I_i} \alpha_j b_j \in V_i$.
Aus (ii) folgt $\dim(\mathrm{Eig}(\lambda_i, f)) = |I_i| =: s_i$. Nun ist

$$\sum_{i=1}^{m} s_i = n = \sum_{i=1}^{m} \nu(\lambda_i, Charpol_f). \tag{iv}$$

Da nach Lemma 5.5.4 $1 \leq s_i \leq \nu(\lambda_i, Charpol_f)$ ist, impliziert (iv) schon $s_i = \nu(\lambda_i, Charpol_f)$ für $i = 1, \dots, m$.
(ii) $\Rightarrow$ (iii): Seien $\lambda_1, \dots, \lambda_m$ die paarweise verschiedenen Eigenwerte von f. Sei $W := \sum_{i=1}^{m} \mathrm{Eig}(\lambda_i, f,)$. Dann gilt $W = \bigoplus_{i=1}^{m} \mathrm{Eig}(\lambda_i, f) \subseteq V$, und es ist $\dim W = \sum_{i=1}^{m} \dim(\mathrm{Eig}(\lambda_i, f)) = \sum_{i=1}^{m} \nu(\lambda_i, Charpol_f) = n$. Damit ist $W = V$.
(iii) $\Rightarrow$ (i): Sind $B_1, \dots, B_m$ die Basen von $\mathrm{Eig}(\lambda_1, f), \dots, \mathrm{Eig}(\lambda_m, f)$, so ist $B = B_1 \cup \dots \cup B_m$ eine Basis von V aus Eigenvektoren von f. Wie wir in 5.5.1 gesehen haben, hat dann $\mathfrak{D}_{B,B}(f)$ Diagonalgestalt. □

Wir wollen noch untersuchen, wann und unter welchen Bedingungen sich zwei Endomorphismen bezüglich einer Basis simultan durch eine Diagonalmatrix darstellen lassen.

5.5.6 Satz *Sind $f, g \in \mathrm{End}(V)$ und sind $\mathfrak{D}_{B,B}(f)$ als auch $\mathfrak{D}_{B,B}(g)$ Diagonalmatrizen, so sind f und g* vertauschbar, *d.h. es gilt $f \circ g = g \circ f$.*

Beweis: Es folgt aus der Definition der Matrixmultiplikation, daß $\mathfrak{A}\mathfrak{B} = \mathfrak{B}\mathfrak{A}$ für Diagonalmatrizen $\mathfrak{A}$ und $\mathfrak{B}$ gilt. Demnach erhalten wir

$$\begin{aligned} f \circ g &= ({h_B}^{-1} \circ \overline{\mathfrak{D}_{B,B}(f)} \circ h_B) \circ ({h_B}^{-1} \circ \overline{\mathfrak{D}_{B,B}(g)} \circ h_B) \\ &= {h_B}^{-1} \circ \overline{\mathfrak{D}_{B,B}(g) \cdot \mathfrak{D}_{B,B}(f)} \circ h_B \\ &= {h_B}^{-1} \circ \overline{\mathfrak{D}_{B,B}(f) \cdot \mathfrak{D}_{B,B}(g)} \circ h_B \\ &= ({h_B}^{-1} \circ \overline{\mathfrak{D}_{B,B}(g)} \circ h_B) \circ ({h_B}^{-1} \circ \overline{\mathfrak{D}_{B,B}(f)} \circ h_B) = g \circ f. \end{aligned}$$

□

5.5.7 Satz *Seien $f, g \in \mathrm{End}(V)$ diagonalisierbar und vertauschbar. Dann gibt es eine Basis B, so daß sowohl $\mathfrak{D}_{B,B}(f)$ als auch $\mathfrak{D}_{B,B}(g)$ Diagonalmatrizen sind.*

Beweis: Nach Satz 5.5.5 gilt $V = \bigoplus_{i=1}^{k} \mathrm{Eig}(\lambda_i, f)$. Nun ist jeder der Eigenräume $W_i := \mathrm{Eig}(\lambda_i, f)$ invariant unter g, d.h. es gilt $g[W_i] \subseteq W_i$, denn für $w \in W_i$ gilt $f(g(w)) = g(f(w)) = g(\lambda_i \cdot w) = \lambda_i \cdot g(w)$ und damit $g(w) \in W_i$. Um den Satz zu beweisen, genügt es nach Satz 5.5.5 daher, jeden der Eigenräume W_i von f in eine direkte Summe von Eigenräumen von g zu zerlegen. Sei daher $W_{ij} := W_i \cap \mathrm{Eig}(\mu_j, g)$. Zu $w \in W_i$ erhalten wir wegen $V = \bigoplus_{j=1}^{m} \mathrm{Eig}(\mu_j, g)$ Vektoren $w_j \in \mathrm{Eig}(\mu_j, g)$ mit $w = w_1 + \cdots + w_m$. Dann folgt aber $f(w_1) + \cdots + f(w_m) = f(w) = \lambda_i \cdot w = \lambda_i w_1 + \cdots + \lambda_i w_m$ und wegen der Eindeutigkeit der Summendarstellung damit $f(w_j) = \lambda_i w_j$, d.h. $w_j \in W_{ij}$. Also ist $W_i = \sum_{j=1}^{m} W_{ij}$ und, da $\mu_1, \ldots, \mu_m$ paarweise verschiedene Eigenwerte von g sind, nach Lemma 5.5.3 schon $W_i = \bigoplus_{j=1}^{m} W_{ij}$. □

Man überlegt sich leicht, daß sich die eben gemachten Überlegungen sofort auf die simultane Darstellung endlicher Mengen von Endomorphismen übertragen lassen. Hier hat man paarweise Vertauschbarkeit zu fordern. Für unendliche Mengen wird die Beweisführung verwickelter.

5.6 Algebren

5.6.1 Definition Es sei K ein Körper. Eine Menge A mit den Verknüpfungen $\cdot\colon K \times A \longrightarrow A$, $+\colon A \times A \longrightarrow A$ und $\cdot\colon A \times A \longrightarrow A$ heißt eine *Algebra*, wenn die folgenden Bedingungen erfüllt sind:

(A.1) *Mit den Verknüpfungen $\cdot\colon K \times A \longrightarrow A$ und $+\colon A \times A \longrightarrow A$ ist die Menge A ein K-Vektorraum.*

(A.2) *Für $\alpha, \beta, \gamma \in K$ und $a, b, c \in A$ gilt*

$$a(\beta b + \gamma c) = \beta ab + \gamma ac \quad \textit{und} \quad (\alpha a + \beta b)c = \alpha(ac) + \beta(bc). \tag{5.13}$$

Gilt $a(bc) = (ab)c$ für alle $a, b, c \in A$, so heißt die Algebra A assoziativ.

Bemerkung Es gibt durchaus nicht assoziative Algebren, die in der Mathematik eine wichtige Rolle spielen. So spricht man von einer *Lie–Algebra*, wenn $ab = -ba$ und $a(bc) + b(ca) + c(ab) = 0$ ist. Der Vektorraum $\mathbb{R}^3$ zusammen mit dem sogenannten *Vektorprodukt*

$$(x_1, x_2, x_3) \times (y_1, y_2, y_3) := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$$

ist eine Lie–Algebra.
Jede assoziative Algebra ist mit den Verknüpfungen $+$ und $\cdot\colon A \times A \longrightarrow A$ ein Ring. Aus der Tatsache, daß $\operatorname{End}(V)$ mit $+$ und $\circ$ ein Ring ist, und (i) von Bemerkung 3.1.10 folgt, daß $\operatorname{End}(V)$ eine assoziative Algebra ist. Für einen n-dimensionalen Vektorraum V sind $K^{n\times n}$ und $\operatorname{End}(V)$ als Algebren isomorph (s.u). Daher ist auch $K^{n\times n}$ eine Algebra.
Ebenso überzeugt man sich sofort, daß der Polynomring $K[x]$ eine assoziative Algebra über K ist.
Ist 0_A das Nullelement des Vektorraumes, so erhalten wir wieder $0_A \cdot a = (b - b) \cdot a = ba - ba = 0_A$ und analog $a \cdot 0_A = 0_A$ für alle $a \in A$.

5.6.2 Definition Unter der *Dimension einer Algebra A* über einem Körper K verstehen wir die Dimension des Vektorraumes A über K.
Wir sagen, daß A eine Algebra mit 1-Element ist, wenn es ein Element $1_A \in A$ mit $1_A \cdot a = a \cdot 1_A = a$ gibt. Wegen $1'_A = 1'_A \cdot 1_A = 1_A$ ist das 1-Element einer Algebra eindeutig bestimmt.
Sind A und A' zwei Algebren, so nennen wir $f\colon A \longrightarrow A'$ einen *Homomorphismus der Algebren*, wenn f ein Homomorphismus der Vektorräume ist und $f(ab) = f(a)f(b)$ für alle $a, b \in A$ gilt.
Das Bild und der Kern eines Homomorphismus ist wie bei Vektorräumen erklärt. Man überzeugt sich sofort, daß für einen Homomorphismus $f\colon A \longrightarrow B$ der Algebren $\operatorname{Im}(f)$ eine Teilalgebra von B und $\operatorname{Kern}(f)$ eine Teilalgebra von A ist.

5.6.3 Definition Wie in einem Ring definieren wir für eine assoziative Algebra mit 1-Element und $a \in A$ die Potenzen von a rekursiv durch $a^0 = 1_A$ und $a^{n+1} := a^n \cdot a$. Den durch die Potenzen eines Elementes $a \in A$ aufgespannten Vektorraum bezeichnen wir mit $K[a]$. Es ist also

$$K[a] := \{\sum_{i=0}^{n} \alpha_i a^i \mid n \in \mathbb{N} \wedge \alpha_i \in K\}. \tag{5.14}$$

5.6.4 Lemma *Sei A eine assoziative Algebra über K und $a \in A$. Die Abbildung*

$$\begin{aligned} \rho_a\colon K[x] &\longrightarrow A \\ \sum_{i=0}^n \alpha_i x^i &\longmapsto \sum_{i=0}^n \alpha_i a^i \end{aligned} \tag{5.15}$$

ist ein Homomorphismus der Algebren. Man nennt sie den Einsetzungshomomorphismus. *Insbesondere ist* Im $\rho_a = K[a]$ *und damit* $K[a]$ *eine Teilalgebra von* A. *Weiter ist* Kern ρ_a *ein Ideal in* $K[x]$.

Beweis: Man rechnet sofort nach, daß ρ_a linear ist und $(\sum_{i=0}^m \alpha_i a^i)\cdot(\sum_{j=0}^n \beta_i a^i) = \sum_{k=0}^{m+n}(\sum_{i+j=k} \alpha_i\beta_j)a^k$ gilt. Damit ist ρ_a ein Homomorphismus der Algebren. Wir haben Im $\rho_a = K[a]$ nach Definition des Einsetzungshomomorphismus. Ist $f(x) \in$ Kern ρ_a und $g(x) \in K[x]$, so ist $f(a)g(a) = 0_A \cdot g(a) = 0_A = g(a) \cdot 0_A = g(a)f(a)$, d.h. $f(x)g(x) \in$ Kern ρ_a und $g(x)f(x) \in$ Kern ρ_a gelten für alle $g(x) \in K[x]$. Damit liegt ein Ideal vor. $\square$

5.6.5 Lemma *Sei* A *eine assoziative Algebra mit 1. Ist für* $a \in A$ *der Einsetzungshomomorphismus* ρ_a *nicht injektiv, so gelten:*

(i) Kern ρ_a *ist ein vom Nullideal verschiedenes Ideal in* $K[x]$.

(ii) Es gibt ein eindeutig bestimmtes normiertes Polynom $\mathit{Minpol}_a(x)$ *mit* Kern $\rho_a = (\mathit{Minpol}_a(x))_{K[x]}$. *Man nennt* $\mathit{Minpol}_a(x)$ *das* Minimalpolynom *von* a.

(iii) Jedes Polynom $f(x) \in K[x]$ *mit* $f(a) = 0$ *wird von* $\mathit{Minpol}_a(x)$ *geteilt.*

(iv) $\dim(K[a]) = G(\mathit{Minpol}_a(x))$.

Beweis: Nach Lemma 5.6.4 ist Kern ρ_a ein Ideal in $K[x]$. Da ρ_a nicht injektiv ist, ist es nicht das Nullideal. Nach Satz 5.3.11 gibt es daher ein eindeutig bestimmtes, normiertes Polynom $\mathit{Minpol}_a(x)$, das Kern ρ_a erzeugt. Gilt $f(a) = 0$ so ist $f(x) \in$ Kern $\rho_a = (\mathit{Minpol}_a(x))$. Daraus folgt $\mathit{Minpol}_a(x) \mid f(x)$. Es bleibt also nur

$$\dim K[a] = G(\mathit{Minpol}_a(x)) \tag{i}$$

zu zeigen. Ist $\mathit{Minpol}_a(x) = \sum_{i=0}^m \alpha_i x^i$, so folgt

$$a^m = \sum_{i=0}^{m-1} -\alpha_i a^i. \tag{ii}$$

Aus (ii) folgt aber, daß sich jede Potenz a^k mit $k \geq m$ als Linearkombination von $\{1, a, \ldots, a^{m-1}\}$ darstellen läßt. Damit ist

$$G(\mathit{Minpol}_a(x)) = m \geq \dim K[a]. \tag{iii}$$

Nehmen wir an, daß $\{1, a, \ldots, a^{m-1}\}$ linear abhängig sind, so erhalten wir $\sum_{i=0}^{m-1} \beta_i a^i = 0$, wobei nicht alle β_i, $i = 0, \ldots, m-1$ verschwinden. Damit ist $0 \neq \sum_{i=0}^{m-1} \beta_i x^i \in (Minpol_a(x))$, woraus $Minpol_a(x) \mid \sum_{i=0}^{m-1} \beta_i x^i$ folgt. Wegen $G(\sum_{i=0}^{m-1} \beta_i x^i) = m-1 < m = G(Minpol_a(x))$ und (5.4) ist dies unmöglich. Also sind $1, a, \ldots, a^{m-1}$ linear unabhängig und es folgt

$$\dim K[a] \geq m. \tag{iv}$$

Aus (iii) und (iv) erhalten wir aber $\dim K[a] = G(Minpol_a(x))$. □

Fassen wir diese Ergebnisse zusammen, so erhalten wir den folgenden Satz.

5.6.6 Satz *Es sei A eine assoziative Algebra mit 1 über einem Körper K. Für ein Element $a \in A$ liegt dann einer der beiden Fälle vor:*

(i) *Der Einsetzungshomomorphismus $\rho_a\colon K[x] \longrightarrow K[a]$ ist ein Isomorphismus der Algebren. Dann ist $\dim K[a] = \dim K[x] = \infty$ und die Menge $\{a^i \mid i \in \mathbb{N}\}$ ist linear unabhängig. In diesem Falle heißt $a \in A$* transzendent über K.

(ii) *Die Algebra $K[a]$ hat endliche Dimension n über K. In diesem Fall sind $1, a, \ldots, a^n$ linear abhängig, und es gibt ein eindeutig bestimmtes, normiertes Polynom $Minpol_a(x) \in K[x]$ vom Grad n mit $Minpol_a(a) = 0$. In diesem Fall heißt $a \in A$* algebraisch über K *und $Minpol_a(x)$ das* Minimalpolynom *von a.*

5.6.7 Korollar *Ist A eine endlich dimensionale Algebra über einem Körper K, so ist jedes Element $a \in A$ algebraisch über K.*

5.6.8 Satz *Sei A eine assoziative Algebra mit 1 über einem Körper K. Für ein algebraisches Element $u \in K$ sind dann äquivalent:*

(i) *$u \in A^*$, d.h. u ist eine Einheit in A.*

(ii) *$u \in K[u]^*$, d.h. u ist eine Einheit von $K[u]$.*

(iii) *0 ist keine Wurzel des Minimalpolynoms von u.*

(iv) *u ist kein Nullteiler von $K[u]$.*

Beweis: (i) $\Rightarrow$ (ii): Wir definieren die Menge

$$K_1[u] := \{\sum_{i=1}^{n} \alpha_i u^i \mid n \in \mathbb{N} \wedge \alpha_i \in K\}. \tag{i}$$

Man rechnet sofort nach, daß $K_1[u]$ eine Teilalgebra von A ist, und wir definieren eine Abbildung $\phi\colon K_1[u] \longrightarrow K[u]$ durch $\phi(b) := b \cdot u^{-1}$. Durch Nachrech-

nen verifizieren wir, daß ϕ ein Homomorphismus der Vektorräume ist. Nach Definition ist ϕ surjektiv. Wir erhalten die Elemente in $K_1[u]$ als Bilder des Einsetzungshomomorphismus ρ_u angewandt auf Polynome der Form $x \cdot f(x)$. Ist daher $0 = \phi(b) = \phi(uf(u)) = uf(u)u^{-1} = f(u)$, so folgt auch $b = uf(u) = 0$. Damit ist ϕ injektiv, und es folgt $\dim(K_1[u]) = \dim(K[u])$. Da $K_1[u]$ ein Teilraum von $K[u]$ ist, folgt demnach $K_1[u] = K[u]$. Damit ist $1 \in K_1[u]$ und folglich $\phi(1) = u^{-1} \in K[u]$. Damit ist u eine Einheit in $K[u]$.
(ii) $\Rightarrow$ (iii): Ist u eine Einheit von $K[u]$, so ist $u^{-1} = f(u)$ für ein $f(x) \in K[x]$. Sei $g(x) := 1 - x \cdot f(x)$. Dann ist $g(u) = 1 - uu^{-1} = 0$ und damit das Minimalpolynom $\mathit{Minpol}_u(x)$ von u ein Teiler von $g(x)$. Wegen $g(0) \neq 0$ folgt dann aber $\mathit{Minpol}_u(0) \neq 0$.
(iii) $\Rightarrow$ (iv): Da 0 genau dann eine Wurzel des Minimalpolynoms ist, wenn x ein Teiler von $\mathit{Minpol}_u(x)$ ist und 0 nach Voraussetzung keine Wurzel des Minimalpolynoms von u ist, sind $\mathit{Minpol}_u(x)$ und x teilerfremd in $K[x]$. Sei $a = f(u) \in K[u]$ mit $a \cdot u = 0$. Dann ist $uf(u) = 0$ und daher $\mathit{Minpol}_u(x)$ ein Teiler von $xf(x)$. Nach Satz 5.1.15 folgt daher $\mathit{Minpol}_u(x) \mid f(x)$. Damit folgt aber $a = f(u) = 0$.
(iv) $\Rightarrow$ (i): Da u kein Nullteiler in $K[u]$ ist, ist die Multiplikation von links $l_u\colon K[u] \longrightarrow K[u] \quad a \longmapsto u \cdot a$ injektiv und nach Satz 2.3.9 auch bijektiv. Damit existiert ein Urbild der 1, d.h., es gibt ein $v \in K[u]$ mit $uv = 1$. Also ist u eine Einheit in $K[u]$ und da $K[u] \subseteq A$ auch schon in A. □

5.6.9 Korollar *Ist A ein assoziative Algebra mit 1 über einem Körper K, so gilt für jede Einheit $u \in A$ schon $K[u] = K[u^{-1}]$.*

Beweis: Ist $u \in A^*$, so folgt nach Satz 5.6.8 $u^{-1} \in K[u]$. Also ist auch $\sum_{i=0}^{n} \alpha_i (u^{-1})^i \in K[u]$ und damit $K[u^{-1}] \subseteq K[u]$. Die Gegenrichtung erhalten wir in analoger Weise, wenn wir von u^{-1} ausgehen. □

5.7 Die Minimalzerlegung eines algebraischen Elements

5.7.1 Lemma *Ein Element $a \in A$ einer assoziativen Algebra A heißt* nilpotent, *wenn es ein $n \in \mathbb{N}$ gibt mit $a^n = 0$. Ist $n_0 := \min\{n \in \mathbb{N} \mid a^n = 0\}$, so ist $\mathit{Minpol}_a(x) = x^{n_0}$.*

Beweis: Es ist $a^{n_0} = 0$ und x^{n_0} ein normiertes Polynom. Es genügt daher zu zeigen, daß das Minimalpolynom von a den Grad n_0 haben muß. Nach Lemma 5.6.5 genügt es dazu, $\dim(K[a]) \geq n_0$ zu zeigen. Wir behaupten, daß $1, a, a^2, \ldots, a^{n_0-1}$

linear unabhängig sind. Ist nämlich $\sum_{i=0}^{n_0-1} \alpha_i a^i = 0$, so erhalten wir durch Multiplikation mit a^{n_o-1} zunächst $\alpha_0 = 0$ und somit $\sum_{i=1}^{n_0-1} \alpha_i a^i = 0$. Durch Multiplikation mit a^{n_0-2} folgt daraus $\alpha_1 = 0$ und durch Iteration schließlich $\alpha_0 = \cdots = \alpha_{n_0-1} = 0$. □

5.7.2 Satz *Ist A eine assoziative und kommutative Algebra über einem Körper K, so ist die Menge der nilpotenten Elemente von A ein Ideal in A.*

Beweis: Sind a, b nilpotent in A, so gibt es $r, s \in \mathbb{N}$ mit $a^r = b^s = 0$. Dann folgt

$$(a-b)^{r+s} = \sum_{i=0}^{r+s} (-1)^i \cdot \binom{r+s}{i} a^{r+s-i} b^i = 0$$

und

$$(\alpha a)^r = 0 \quad \text{sowie} \quad (ca)^r = c^r a^r = 0.$$ □

5.7.3 Definition Sei A eine assoziative Algebra mit 1 über einem Körper K. Ein Element $a \in A$ heißt *split über K*, wenn

(i) *a algebraisch über K ist*

und

(ii) *das Minimalpolynom $\mathit{Minpol}_a(x)$ von a über K in Linearfaktoren zerfällt.*

Hat das Minimalpolynom von a nur einfache Wurzeln, so nennen wir a ein *halbeinfaches Element* von A.

Gilt $a^2 = a$ für ein von 0 verschiedenes Element der Algebra A, so nennen wir a *idempotent*.
Endlich viele idempotente Elemente $c_1, \ldots, c_n$ einer assoziativen Algebra A mit 1 heißen ein *vollständiges Orthogonalsystem* wenn gilt:

(i) $1 = c_1 + \cdots + c_n$.

(ii) $c_i c_j = \delta_{ij} c_i$ *für* $i, j = 1, \ldots, n$.

Beachte Ist $c_1, \ldots, c_n$ ein vollständiges Orthogonalsystem, so ist die Menge $\{c_1, \ldots, c_n\}$ linear unabhängig, denn aus $\sum_{i=1}^{n} \alpha_i c_i = 0$ erhalten wir durch Multiplikation mit c_j für $j = 1, \ldots, n$ sofort $\alpha_j = 0$.

5.7.4 Satz (Minimalzerlegung eines algebraischen Elements) *Sei A eine assoziative Algebra mit 1 über einem Körper K. Ist $a \in A$ split über K und sind $\lambda_1, \ldots, \lambda_m$*

die verschiedenen Wurzeln des Minimalpolynoms von a, so gibt es ein vollständiges Orthogonalsystem $c_1, \ldots, c_m \in K[a]$ und ein nilpotentes Element $v \in K[a]$ mit

$$a = \sum_{i=1}^{m} \lambda_i c_i + v. \tag{5.16}$$

Das Element $a \in A$ ist genau dann halbeinfach, wenn der nilpotente Anteil v verschwindet. Für halbeinfaches a gilt

$$f(a) = \sum_{i=1}^{m} f(\lambda_i) c_i \tag{5.17}$$

für jedes Polynom $f(x) \in K[x]$.

Beweis: Da a split über K ist, gilt für das Minimalpolynom von a

$$Minpol_a(x) = \prod_{i=1}^{m} (x - \lambda_i)^{k_i}. \tag{i}$$

Setzen wir

$$q_j(x) := \prod_{\substack{i=1 \\ i \neq j}}^{m} (x - \lambda_i)^{k_i} \tag{ii}$$

und

$$p_j(x) := (x - \lambda_j)^{k_j}, \tag{iii}$$

so erhalten wir sofort

$$q_j(x) \cdot p_j(x) = Minpol_a(x) \quad \text{für } \; j = 1, \ldots, m, \tag{iv}$$

$$Minpol_a(x) \mid q_i(x) \cdot q_j(x) \quad \text{für } \quad i \neq j \tag{v}$$

und

$$q_j(x) \;\text{ und }\; p_j(x) \quad \text{sind teilerfremd für } \; j = 1, \ldots, m, \tag{vi}$$

denn die Körperelemente $\lambda_1, \ldots, \lambda_m$ sind paarweise verschieden. Da $K[x]$ ein Hauptidealring ist, erhalten wir aus (vi) nach Satz 5.1.14 Elemente $f_j(x), g_j(x) \in K[x]$ mit

$$1 = q_j(x) f_j(x) + p_j(x) g_j(x). \tag{vii}$$

Sei nun

$$h_i(x) := q_i(x)f_i(x) \tag{viii}$$

und

$$c_i := h_i(a) \quad \text{für } i = 1, \ldots, m. \tag{ix}$$

Wir wollen zeigen, daß $c_1, \ldots, c_m$ ein vollständiges Orthogonalsystem bilden und $a - \sum_{j=1}^{m} \lambda_j c_j$ nilpotent ist.
Nach (vii) gilt $h_i(x)^2 - h_i(x) = h_i(x) \cdot (h_i(x) - 1) = -q_i(x)f_i(x)p_i(x)g_i(x)$, und aus (iv) folgt daher

$$Minpol_a(x) \mid (h_i(x)^2 - h_i(x)). \tag{x}$$

Damit ist $c_i^2 - c_i = h_i(a)^2 - h_i(a) = 0$ und somit

$$c_i^2 = c_i \quad \text{für } i = 1, \ldots, n. \tag{xi}$$

Aus (v) folgt

$$Minpol_a(x) \mid h_i(x)h_j(x) \quad \text{für } i \neq j \tag{xii}$$

und damit

$$c_i \cdot c_j = 0 \quad \text{für } i \neq j. \tag{xiii}$$

Setzen wir $g(x) := (\sum_{i=1}^{m} h_i(x)) - 1 = (\sum_{\substack{i=1 \\ i \neq j}}^{m} h_i(x)) + h_j(x) - 1$, so erhalten wir $p_j(x) \mid g(x)$ für $j = 1, \ldots, m$, denn $p_j(x) \mid q_i(x)$ für $j \neq i$ nach Konstruktion und $p_j(x) \mid (h_j(x) - 1)$ nach (vii). Da die $p_j(x)$ paarweise teilerfremd sind, erhalten wir nach Bemerkung 5.1.18 wegen $Minpol_a(x) = \prod_{j=1}^{m} p_j(x)$

$$Minpol_a(x) \mid (\sum_{i=1}^{m} h_i(x) - 1). \tag{xiv}$$

Aus (xiv) folgt nun $\sum_{i=1}^{m} h_i(a) - 1 = 0$ und damit

$$\sum_{i=1}^{m} c_i = 1. \tag{xv}$$

Aus (xi), (xiii) und (xv) folgt, daß $c_1, \ldots, c_m$ ein vollständiges Orthogonalsystem bilden.
Um nachzuweisen, daß $a - \sum_{i=1}^{m} \lambda_i c_i$ nilpotent ist, untersuchen wir das Polynom $f(x) := x - \sum_{j=1}^{m} \lambda_j h_j(x)$. Wegen $h_j(\lambda_i) = q_j(\lambda_i)f_j(\lambda_i) = 0 \cdot f_j(\lambda_i) = 0$ für $i \neq j$ und $h_i(\lambda_i) = 1 - p_i(\lambda_i)g_i(\lambda_i) = 1$ folgt, daß $f(\lambda_i) = \lambda_i - \lambda_i = 0$ für $i = 1, \ldots, m$ ist. Also ist jedes λ_i eine Wurzel von $f(x)$. Damit ist $f(x)$ durch $x - \lambda_i$ für $i = 1, \ldots, m$ teilbar, und für $k := \max\{k_1, \ldots, k_m\}$ folgt damit

$$Minpol_a(x) \mid (x - \sum_{i=1}^{m} \lambda_i h_i(x))^k. \tag{xvi}$$

Für $v := a - \sum_{i=1}^{m} \lambda_i h_i(a)$ erhalten wir daher

$$v^k = (a - \sum_{i=1}^{m} \lambda_i h_i(a))^k = 0 \tag{xvii}$$

und

$$a = \sum_{i=1}^{m} \lambda_i c_i + v. \tag{xviii}$$

Es bleibt also nur noch die Aussage über die halbeinfachen Elemente zu zeigen. Ist a halbeinfach, so hat $Minpol_a(x)$ nur einfache Wurzeln. Damit ist $k = 1$ und damit nach (xvii) schon $v = 0$, d.h. der nilpotente Anteil verschwindet. Ist umgekehrt $v = 0$, so erhalten wir aus (xviii), (xiii) und (xi)

$$a^r = \sum_{i=1}^{m} \lambda_i^r c_i \quad \text{für } r \geq 0. \tag{xix}$$

Aus (xix) folgt aber für jedes $f(x) \in K[x]$ wegen $f(a) = \sum_{j=0}^{n} \alpha_j a^j = \sum_{j=0}^{n} \alpha_j (\sum_{i=1}^{m} \lambda_i^j c_i) = \sum_{i=1}^{m} (\sum_{j=0}^{n} \alpha_j \lambda_i^j) c_i$

$$f(a) = \sum_{i=1}^{m} f(\lambda_i) c_i. \tag{xx}$$

Wählen wir $f(x) = \prod_{i=1}^{m} (x - \lambda_i)$ in (xx), so ist $G(f(x)) \leq G(Minpol_a(x))$ und $f(a) = 0$. Damit haben wir $Minpol_a(x) \mid f(x)$ und, da beide Polynome normiert sind, schließlich $f(x) = Minpol_a(x)$. Damit hat $Minpol_a(x)$ nur einfache Wurzeln, d.h. a ist halbeinfach. $\square$

Die Zerlegung $a = h + v$ mit $h = \sum_{i=1}^{m} \lambda_i c_i$ für ein vollständiges Orthogonalsystem $c_1, \ldots, c_m \in K[a]$ und ein nilpotentes $v \in K[a]$ nennt man die *Minimalzerlegung* von a. Dabei heißt h der *halbeinfache Anteil* und v der *nilpotente Anteil* des Elementes a. Man beachte, daß die Algebra $K[a]$ als homomorphes Bild der kommutativen Algebra $K[x]$ kommutativ ist. Da h und v beide in $K[a]$ liegen, gilt $h \cdot v = v \cdot h$. Das heißt h und v sind *vertauschbar*. Ebenso sind h und a und v und a vertauschbar.

Aus der Tatsache, daß die Koeffizienten des halbeinfachen Anteils genau die Wurzeln des Minimalpolynoms sind, und Satz 5.6.8 erhalten wir den folgenden Satz.

5.7.5 Satz *Ein Element $a \in A$ mit der Minimalzerlegung $a = \sum_{i=1}^{m} \lambda_i c_i + v$ ist genau dann invertierbar, wenn alle λ_i ungleich Null sind.*

5.8 Die Algebra der Endomorphismen eines endlich dimensionalen Vektorraumes

Wir wollen nun die abstrakt gewonnenen Ergebnisse des Abschnittes 5.7 auf die Algebra der Endomorphismen anwenden. Wir haben bereits eingesehen, daß $\mathrm{End}(V)$ für einen n-dimensionalen K-Vektorraum eine assoziative Algebra mit 1 der Dimension n^2 über K ist. Wir wollen, um Verwechslungen mit Polynomen zu vermeiden, im folgenden die Elemente von $\mathrm{End}(V)$ mit großen lateinischen Buchstaben bezeichnen.

5.8.1 Satz *Ist V ein endlich dimensionaler K-Vektorraum und $F \in \mathrm{End}(V)$, so stimmen die in K liegenden Wurzeln des Minimalpolynoms von F mit den Eigenwerten von F überein.*

Beweis: Sei λ ein Eigenwert von F und v ein Eigenvektor zum Eigenwert λ. Dann gilt $F^n(v) = \lambda^n \cdot v$ und damit $f(F)(v) = f(\lambda) \cdot v$ für jedes Polynom $f(x) \in K[x]$. Insbesondere erhalten wir $Minpol_F(\lambda) \cdot v = Minpol_F(F)(v) = 0$ und damit $Minpol_F(\lambda) = 0$.
Ist λ eine Wurzel des Minimalpolynoms, so gilt $Minpol_F(x) = (\lambda - x) \cdot g(x)$. Damit folgt

$$(\lambda \cdot id - F) \cdot g(F) = Minpol_F(F) = 0. \qquad \text{(i)}$$

Da aus $g(F) = 0$ schon $Minpol_F(x) \mid g(x)$ folgen müßte, was aber wegen $G(g(x)) < G(Minpol_F(x))$ unmöglich ist, haben wir $g(F) \neq 0$. Damit gibt es ein $v \in V$ mit $g(F)(v) \neq 0$, und wir erhalten aus (i) $\lambda \cdot g(F)(v) = F(g(F)(v))$. Damit ist λ ein Eigenwert von F. □

5.8.2 Korollar *Die Wurzeln des Minimalpolynoms und des charakteristischen Polynoms eines Endomorphismus stimmen bis auf Vielfachheiten überein. Insbesondere ist ein Endomorphismus genau dann split über dem Grundkörper, wenn alle Wurzeln des charakteristischen Polynoms im Grundkörper liegen.*

Ist nun $F \in \mathrm{End}(V)$ ein Endomorphismus, dessen Eigenwerte alle im Grundkörper von K liegen, so ist F split über K, und nach Satz 5.7.4 erhalten wir eine Minimalzerlegung von F in der Form

$$F = \sum_{i=1}^{m} \lambda_i C_i + V, \qquad (5.18)$$

wobei $\lambda_1, \ldots, \lambda_m$ alle verschiedenen Eigenwerte von F sind, $C_1, \ldots, C_m$ ein vollständiges Orthogonalsystem bilden und V nilpotent ist. Weiter wissen wir, daß alle C_i und V in $K[F]$ liegen und daher insbesondere vertauschbar sind. Unser nächstes Ziel muß daher darin bestehen zu klären, welche geometrische Bedeutung vollständige Orthogonalsysteme und nilpotente Endomorphismen haben.

5.8.3 Lemma *Sind F und N vertauschbare Endomorphismen, wobei N nilpotent ist, so haben F und $F + N$ die (möglicherweise bis auf Vielfachheiten) gleichen Eigenwerte.*

Beweis: Sei λ ein Eigenwert von $F + N$. Für einen Eigenvektor v zu λ gilt dann $\lambda v = (F + N)(v) = F(v) + N(v)$ und damit

$$(\lambda \cdot id - F)(v) = N(v). \tag{i}$$

Es ist $\min\{n \in \mathbb{N} \mid N^n(v) = 0\} = m + 1$ für eine natürliche Zahl m. Aus der Vertauschbarkeit von N und F folgt aber sofort auch die Vertauschbarkeit von $\lambda \cdot id - F$ mit allen Potenzen von N. Damit erhalten wir aus (i)

$$(\lambda \cdot id - F)N^m(v) = N^m(\lambda \cdot id - F)(v) = N^{m+1}(v) = 0 \tag{ii}$$

und erhalten damit $N^m(v)$ als einen Eigenvektor zum Eigenwert λ von F. Somit ist jeder Eigenwert von $F+N$ auch ein Eigenwert von F. Setzen wir $F' := F+N$ und $N' := -N$, so bleibt N' nilpotent und F' und N' vertauschbar. Damit ist, wie wir gerade gezeigt haben, jeder Eigenwert von $F' + N' = F$ auch ein Eigenwert von $F' = F + N$. □

5.8.4 Satz *Ist V ein endlich dimensionaler K-Vektorraum und $C_1, \ldots, C_m \in \mathrm{End}(V)$ ein vollständiges Orthogonalsystem in $\mathrm{End}(V)$, so ist*

$$V = C_1[V] \oplus \cdots \oplus C_m[V].$$

Beweis: Wegen $C_1 + \cdots + C_m = 1$ erhalten wir $v = (C_1 + \cdots + C_m)(v) = C_1(v) + \cdots + C_m(v)$. Also ist $V = C_1[V] + \cdots + C_m[V]$. Für $0 = v_1 + \cdots + v_m$ mit $v_i \in C_i[V]$ erhalten wir $v_i = C_i(u_i)$ für ein $u_i \in V$. Dann ist $0 = C_j(v_1 + \cdots + v_m) = C_jC_1(u_1) + \cdots + C_jC_m(u_m) = C_j^2(u_j) = C_j(u_j) = v_j$ für $j = 1, \ldots, m$. Somit folgt $v_1 = \cdots = v_m = 0$, und die Summe ist direkt. □

Wir haben jedoch auch eine Art Umkehrung von Satz 5.8.4.

5.8.5 Satz *Gilt $V = U_1 \oplus \cdots \oplus U_m$ für einen endlich dimensionalen Vektorraum V, so bilden die Projektionen $P_i: V \longrightarrow U_i$ mit $P_i(v_1 + \cdots + v_m) := v_i$ ein vollständiges Orthogonalsystem in $\mathrm{End}(V)$.*

Beweis: Nach Definition gilt $P_iP_j = \delta_{ij}P_i$. Für $v = v_1 + \cdots + v_m$ ist $(P_1 + \cdots + P_m)v = v_1 + \cdots + v_m = v$ und damit $P_1 + \cdots + P_m = id$. □

Fassen wir die Sätze 5.8.4 und 5.8.5 zusammen, so erhalten wir den folgenden Satz.

5.8.6 Satz *Es sei V ein endlich dimensionaler Vektorraum. Zwischen der Zerlegung von V in direkte Summen von Teilräumen und den vollständigen Orthogonalsystemen in* $\mathrm{End}(V)$ *besteht eine bijektive Zuordnung.*

Wir müssen uns nun näher mit den nilpotenten Elementen befassen. Für einen Endomorphismus $F \in \mathrm{End}(V)$ und $a \in V$ definieren wir

$$\mu(F,a) := \min\{n \in \mathbb{N} \mid F^n(a) = 0\} \tag{5.19}$$

und

$$\mu(F) := \min\{n \in \mathbb{N} \mid F^n = 0\}. \tag{5.20}$$

In beiden Fällen handelt es sich um partielle Funktionen, die nicht für jedes Paar (F,a) bzw. für jedes F definiert sein müssen.

5.8.7 Lemma *Sei V ein endlich dimensionaler Vektorraum und $F \in \mathrm{End}(V)$. Gibt es ein von Null verschiedenes $a \in V$, so daß $\mu(F,a)$ definiert ist, so ist $U := \langle\{F^i(a) \mid i \in N\}\rangle$ ein Teilraum der Dimension $m := \mu(F,a)$ von V und $\{F^0(a), \ldots, F^{m-1}(a)\}$ ist eine Basis von U. Man nennt U dann einen bezüglich F zyklischen Unterraum von V.*

Beweis: Da $\{F^i(a) \mid i \in N\}$ ein Erzeugendensystem ist, genügt es zu zeigen, daß $a, F(a), \ldots, F^{m-1}(a)$ linear unabhängig sind. Sei dazu $\sum_{i=0}^{m-1} \alpha_i F^i(a) = 0$. Dann erhalten wir zunächst $0 = F^{m-1}(\sum_{i=0}^{m-1} \alpha_i F^i(a)) = \alpha_0 F^{m-1}(a)$. Aus $F^{m-1}(a) \neq 0$ folgt $\alpha_0 = 0$ und damit $0 = F^{m-2}(0) = F^{m-2}(\sum_{i=1}^{m-1} \alpha_i F^i(a)) = \alpha_1 F^{m-1}(a)$. Damit folgt $\alpha_1 = 0$. Durch Iteration erhalten wir somit $\alpha_0 = \cdots = \alpha_{m-1} = 0$. □

5.8.8 Lemma *Ist U ein bezüglich $F \in \mathrm{End}(V)$ zyklischer Unterraum von V mit $\dim U = m$, so gilt $F^m(v) = 0$ für alle $v \in U$.*

Beweis: Sei $a \in V$ so gewählt, daß $a \neq 0$ und $a, F(a), \ldots, F^{m-1}(a)$ eine Basis von U bilden. Für $v = \sum_{i=0}^{m-1} \alpha_i F^i(a)$ folgt dann $F^m(v) = \sum_{i=0}^{m-1} \alpha_i F^{m+i}(a) = 0$. □

Das folgende Korollar ist eine unmittelbare Folgerung aus Lemma 5.8.8.

5.8.9 Korollar *Ist V ein zyklischer Raum bezüglich F, so ist F nilpotent. In diesem Falle heißt der Endomorphismus F* nilzyklisch.

5.8.10 Satz *Es sei V ein endlich dimensionaler K-Vektorraum und $F \in \mathrm{End}(V)$ nilpotent. Dann gibt es Teilräume $U_1, \ldots, U_r$ von V mit*

(i) *Jeder Raum U_i ist zyklisch bezüglich F und F-invariant, d.h. es gilt $F[U_i] \subseteq U_i$ für $i = 1, \ldots, r$.*

(ii) $V = U_1 \oplus \cdots \oplus U_r$.

Beweis: Ist $F = 0$, so wählen wir eine Basis $b_1, \ldots, b_n$ von V und setzen $r := n$ und $U_i := \langle b_i \rangle$ für $i = 1, \ldots, n$. Sei also $F \neq 0$. Wir beweisen den Satz durch Induktion nach $n := \dim(V)$.
Ist $\dim V = 1$, so ist V selbst zyklisch bezüglich F und F-invariant.
Sei also $\dim V > 1$ und $\mu(F) =: m+1$. Dann gibt es ein $a \in V$ mit $\mu(F, a) = m+1$, und nach Lemma 5.8.7 ist $U := \langle F^0(a), \ldots, F^m(a) \rangle$ ein bezüglich F zyklischer Teilraum der Dimension $m+1$. Da $F^0(a), \ldots, F^m(a)$ eine Basis von U bilden, ist U auch F-abgeschlossen. Ist nun $U = V$, so sind wir fertig. Anderenfalls erhalten wir $\dim(V/U) < \dim V$ wegen

$$\dim V = \dim U + \dim(V/U). \tag{i}$$

Die Zuordnung

$$\tilde{F}(a + U) := F(a) + U \tag{ii}$$

ist wohldefiniert. Ist nämlich $a + U = b + U$, so ist $b = a + u$ für ein $u \in U$ und wir erhalten $\tilde{F}(b + U) = \tilde{F}(a + u + U) = F(a + u) + U = F(a) + F(u) + U = F(a) + U$, denn wegen der F-Abgeschlossenheit von U ist $F(u) \in U$. Damit haben wir eine Abbildung $\tilde{F}\colon V/U \longrightarrow V/U$, und aus der Definition von $\tilde{F}$ folgt, daß das Diagramm

$$\begin{tikzcd} V \arrow[r, "F"] \arrow[d, "\pi_U"'] & V \arrow[d, "\pi_U"] \\ V/U \arrow[r, "\tilde{F}"'] & V/U \end{tikzcd}$$

kommutiert. Weiter ist

$$\tilde{F}^{m+1}(a + U) = F^{m+1}(a) + U = 0 + U. \tag{iii}$$

Damit ist $\tilde{F}$ nilpotent auf V/U, und wir erhalten nach Induktionsvoraussetzung

$$V/U = \tilde{U}_1 \oplus \cdots \oplus \tilde{U}_s, \tag{iv}$$

wobei jeder der Räume $\tilde{U}_i$ zyklisch bezüglich $\tilde{F}$ und $\tilde{F}$-abgeschlossen ist. Setzen wir $m_i + 1 := \dim \tilde{U}_i$, so erhalten wir aus (i)

$$n = (m+1) + \sum_{i=1}^{s} (m_i + 1). \tag{v}$$

Da jeder der Räume $\tilde{U}_i$ zyklisch bezüglich $\tilde{F}$ ist, gibt es ein $\tilde{a}_i \in \tilde{U}_i$, so daß $\mu(\tilde{F}, \tilde{a}_i) = m_i + 1$ und $\{\tilde{F}^j(\tilde{a}_i) \mid 0 \le j \le m_i\}$ eine Basis von $\tilde{U}_i$ ist. Da der kanonische Epimorphismus π_U eine Surjektion ist, finden wir zu jedem $\tilde{a}_i$ ein Urbild $a_i \in V$. Bei der Wahl von a_i haben wir einige Freiheit. Wir wollen a_i so wählen, daß

$$\mu(F, a_i) = m_i + 1 \tag{vi}$$

ist. Um einzusehen, daß dies geht, benötigen wir die folgende Aussage

$$0 \le k \le m \wedge F^k(b) \in U \quad \Rightarrow \quad (\exists c \in V)[F^k(c) = 0 \wedge b - c \in U]. \tag{vii}$$

Bevor wir (vii) beweisen, überlegen wir uns, daß (vi) mit (vii) zu bewerkstelligen ist. Wegen $\mu(\tilde{F}, \tilde{a}_i) = m_i + 1$ ist $F^{m_i+1}(a_i) \in U$ für jedes Urbild a_i von $\tilde{a}_i$. Ist $F^{m_i+1}(a_i) = 0$, so sind wir fertig. Anderenfalls ist $m_i + 1 \le m$ und nach (vii) finden wir ein $b_i \in V$ mit $F^{m_i+1}(b_i) = 0$. Wegen $a_i - b_i \in U$ ist dann $\pi_U(b_i) = \pi_U(a_i) = \tilde{a}_i$.

Nun zum Beweis von (vii). Da $F^k(b) \in U$ ist $F^k(b) = \sum_{i=0}^{m} \alpha_i F^i(a)$. Damit erhalten wir $0 = F^{m+1}(b) = \sum_{i=0}^{m} \alpha_i F^{m+1-k+i}(a) = \sum_{i=0}^{k-1} \alpha_i F^{m+1-k+i}(a)$. Da die $F^i(a)$ linear unabhängig sind, folgt daraus $\alpha_0 = \cdots = \alpha_{k-1} = 0$. Damit ist $F^k(b) = \sum_{i=k}^{m} \alpha_i F^i(a) = F^k(\sum_{i=0}^{m-k} \alpha_{k+i} F^i(a))$. Setzen wir nun $b' := \sum_{i=0}^{m-k} \alpha_{k+i} F^i(a)$ und $c := b - b'$, so ist $F^k(c) = F^k(b) - F^k(b') = 0$ und $b - c = b' \in U$.

Sei nun

$$B := \{F^0(a), \ldots, F^m(a), F^0(a_1), \ldots, F^{m_1}(a_1), \ldots, F^0(a_s), \ldots, F^{m_s}(a_s)\}. \tag{viii}$$

Dann enthält B nach (v) genau n Vektoren. Wir zeigen, daß B eine linear unabhängige Menge ist. Sei dazu $0 = \sum_{i=1}^{s}(\sum_{j=0}^{m_i} \alpha_{ij} F^j(a_i)) + \sum_{i=0}^{m} \alpha_i F^i(a)$. Dann ist $\pi_U(0) = \sum_{i=1}^{s}(\sum_{j=0}^{m_i} \alpha_{ij} \tilde{F}^j(\tilde{a}_i))$. Aus (iv) folgt, daß jeder der Summanden $\sum_{j=0}^{m_i} \alpha_{ij} F^j(a_i)$ in V/U verschwindet und, da $F^0(\tilde{a}_i), \ldots, F^{m_i}(\tilde{a}_i)$ linear unabhängig sind, daß alle α_{ij} für $i = 1, \ldots, s$ und $j = 0, \ldots, m_i$ verschwinden. Damit ist auch $\sum_{i=0}^{m} \alpha_i F^i(a) = 0$ und damit auch $\alpha_0 = \cdots = \alpha_m = 0$. Also ist B linear unabhängig und damit eine Basis von V. Setzen wir nun

$$U_i := \langle F^0(a_i), \ldots, F^{m_i}(a_i) \rangle, \tag{ix}$$

so ist jeder der Räume U_i F-abgeschlossen und zyklisch bezüglich F. Da B eine Basis von V ist, folgt

$$V = U \oplus U_1 \oplus \cdots \oplus U_s,$$

und der Satz ist bewiesen. □

Ist nun F ein nilzyklischer Endomorphismus eines endlich dimensionalen K-Vektorraumes V und ist $B = (b_1, \ldots, b_n) = (F^0(a), \ldots, F^{n-1}(a))$ eine zyklische Basis von V, so erhalten wir aus $F(b_i) = b_{i+1} = \sum_{i=1}^{n} \alpha_{ij} b_j$, daß $\alpha_{ij} = \delta_{(i+1)j}$ und damit

$$\mathfrak{D}_{B,B}(F) = (\delta_{(i+1)j}) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} \tag{5.21}$$

ist. Wir nennen daher Matrizen, die die in (5.21) dargestellte Gestalt haben, *nilzyklisch*.

Ist nun $F \in \mathrm{End}(V)$ ein nilpotenter Endomorphismus und zerlegen wir $V = U_1 \oplus \cdots \oplus U_r$ gemäß Satz 5.8.10, so ist $F{\restriction}U_i$, d.h. die Abbildung $F\colon U_i \longrightarrow U_i$, nilzyklisch auf jedem der U_i. Stellen wir F bezüglich der im Beweis des Satzes konstruierten Basis dar, so erhalten wir als Korollar zu Satz 5.8.10:

5.8.11 Korollar *Ist V ein endlich dimensionaler Vektorraum und $F \in \mathrm{End}(V)$ nilpotent, so gibt es eine Basis B mit*

$$\mathfrak{D}_{B,B}(F) = \begin{pmatrix} \mathfrak{A}_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{A}_r \end{pmatrix},$$

wobei alle Matrizen $\mathfrak{A}_i$, $i = 1, \ldots, r$, nilzyklisch sind.

Wir nennen eine Matrix $\mathfrak{N} \in K^{n\times n}$ *nilpotent*, wenn die dazugehörige Abbildung $\overline{\mathfrak{N}}$ nilpotent ist. Das ist äquivalent zu der Forderung, daß es ein $n \in \mathbb{N}$ gibt mit $\mathfrak{N}^n = 0$. Da der Übergang von der kanonischen Basis des K^n zu der in Satz 5.8.10 konstruierten Basis durch eine Transformationsmatrix aus $\mathrm{GL}(n, K)$ beschrieben wird, können wir Korollar 5.8.11 wie folgt umformulieren.

5.8.12 Satz *Ist $\mathfrak{N} \in K^{n\times n}$ eine nilpotente Matrix, so gibt es eine umkehrbare Matrix $\mathfrak{S} \in \mathrm{GL}(n, K)$, so daß*

$$\mathfrak{S}\mathfrak{N}\mathfrak{S}^{-1} = \begin{pmatrix} \mathfrak{A}_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{A}_r \end{pmatrix}$$

ist, wobei alle Matrizen $\mathfrak{A}_i \quad i = 1, \ldots, r$ *nilzyklisch sind.*

Ist nun

$$F = \sum_{i=1}^{m} \lambda_i C_i + N \tag{5.22}$$

die Minimalzerlegung eines Endomorphismus $F \in \text{End}(V)$, so ist nach Satz 5.8.4

$$V = U_1 \oplus \cdots \oplus U_m \quad \text{mit} \quad U_i = C_i[V]. \tag{5.23}$$

Dann folgt aber

$$C_i \restriction U_i = \textit{id} \quad \text{für} \quad i = 1, \ldots, m \quad \text{und} \quad C_j \restriction U_i = 0 \quad \text{für} \quad i \neq j, \tag{5.24}$$

denn zu $u \in U_i$ gibt es ein $v \in V$ mit $u = C_i(v)$, und damit ist $C_i(u) = C_i^2(v) = C_i(v) = u$ und $C_j(u) = C_j(C_i(v)) = 0$ für $i \neq j$.
Als Folgerung von (5.24) erhalten wir mit $n_i := \dim U_i$

$$\mathfrak{D}_{B,B}(C_i \restriction U_i) = \mathfrak{E}^{(n_i,n_i)} \tag{5.25}$$

und damit

$$\mathfrak{D}_{B,B}(\lambda_i \cdot C_i \restriction U_i) = \lambda_i \cdot \mathfrak{E}^{(n_i,n_i)} \tag{5.26}$$

für $i = 1, \ldots, m$ und jede Basis B von U_i.
Ist nun F ein halbeinfacher Endomorphismus, so folgt aus (5.26), daß F diagonalisierbar ist. Sei umgekehrt F diagonalisierbar und

$$\mathfrak{D}_{B,B}(F) = \begin{pmatrix} \lambda_1 \mathfrak{E}^{(n_1,n_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_m \mathfrak{E}^{(n_m,n_m)} \end{pmatrix}, \tag{5.27}$$

wobei $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von F und n_i die dazugehörigen Dimensionen der Eigenräume sind. Bezeichnen wir mit $\mathfrak{C}_i$ die Matrix, die aus der Matrix in (5.27) entsteht, indem man $\lambda_i \mathfrak{E}^{(n_i,n_i)}$ durch $\mathfrak{E}^{(n_i,n_i)}$ und alle übrigen Kästchen durch Nullmatrizen ersetzt, so rechnet man leicht nach, daß die Endomorphismen

$$C_i := {h_B}^{-1} \circ \overline{\mathfrak{C}_\mathrm{i}} \circ h_B$$

ein Orthogonalsystem bilden und

$$F = \sum_{i=1}^{m} \lambda_i C_i$$

ist. Nach Satz 5.7.4 ist F dann halbeinfach. Damit haben wir den folgenden Satz.

5.8.13 Satz *Sei V ein endlich dimensionaler Vektorraum über einem Körper K und $F \in \mathrm{End}(V)$ ein Endomorphismus, der split über K ist mit Minimalzerlegung $F = \sum_{i=1}^{m} \lambda_i C_i + N$. Dann sind äquivalent:*

(1) F ist diagonalisierbar.

(2) F ist halbeinfach, d.h. das Minimalpolynom von F besitzt nur einfache Wurzeln.

(3) Der nilpotente Anteil in der Minimalzerlegung von F verschwindet.

(4) Der Vektorraum V besitzt eine Basis aus Eigenvektoren von F.

(5) Es gilt $\mathrm{Eig}(\lambda_i, F) = C_i[V]$ für $i = 1, \ldots, m$.

(6) Es ist $\nu(\lambda_i, Charpol_F) = \dim(\mathrm{Eig}(\lambda_i, F)) = \dim(C_i[V])$.

Beweis: Wir haben uns gerade überlegt, daß (1) und (3) äquivalent sind. Die Äquivalenz von (2) und (3) ist in Satz 5.7.4 gezeigt worden. Die Äquivalenz von (1) zu (4) folgt aus Satz 5.5.5. Wir zeigen die Äquivalenz von (5) zu (1). Gilt (5), so erhalten wir mit Satz 5.8.4, daß V die direkte Summe der Eigenräume von F zu dessen verschiedenen Eigenwerten ist. Nach Satz 5.5.5 ist F dann diagonalisierbar. Ist umgekehrt F diagonalisierbar, so ist nach (3) $F = \sum_{i=1}^{m} \lambda_i C_i$. Wir zeigen

$$\mathrm{Eig}(\lambda_i, F) = C_i[V] =: U_i. \tag{i}$$

Für $u \in U_i$ gilt $F(u) = \sum_{i=1}^{m} \lambda_i C_i(u) = \lambda_i C_i(u) = \lambda_i u$ nach (5.24). Damit haben wir die Inklusion von rechts nach links in (i). Nach Satz 5.8.4 ist $V = U_1 \oplus \cdots \oplus U_m$. Für $u \in \mathrm{Eig}(\lambda_i, F)$ erhalten wir

$$\lambda_i(u_1 + \cdots + u_m) = \lambda_i u = F(u) = \sum_{j=1}^{m} \lambda_j C_j(u_1 + \cdots + u_m) = \sum_{j=1}^{m} \lambda_j u_j. \tag{ii}$$

Damit ist $\sum_{j=1}^{m} (\lambda_i - \lambda_j) u_j = 0$. Also haben wir $u_j = 0$ für $j \neq i$ und folglich $u = u_i \in U_i$. □

Kehren wir zum allgemeinen Fall zurück. Da die Endomorphismen C_i und N von (5.22) in $K[F]$ liegen, sind sie vertauschbar. Als Konsequenz davon sind alle Teilräume U_i in (5.23) N-invariant. Aus $u \in U_i$ erhalten wir nämlich wieder $u =$

$C_i(v)$ für ein $v \in V$ und damit $N(u) = N(C_i(v)) = C_i(N(v)) \in C_i[V] = U_i$. Da die nilpotenten Endomorphismen nach Satz 5.7.2 in der kommutativen Teilalgebra $K[F]$ ein Ideal bilden, sind die Endomorphismen

$$N_i := NC_i \tag{5.28}$$

wieder nilpotent. Weiter gilt

$$N = N(C_1 + \cdots + C_m) = N_1 + \cdots + N_m. \tag{5.29}$$

und

$$N_i \upharpoonright U_j = 0 \quad \text{für} \quad i \neq j, \tag{5.30}$$

denn für $u = C_j(v) \in U_j$ ist $N_i(u) = N(C_i(C_j(v))) = 0$. Damit folgt für $F_i := \lambda_i C_i + N_i$, daß F_i auf U_j für $j \neq i$ verschwindet, jeder Teilraum U_i somit F_i-invariant ist und nach (5.22)

$$F = \sum_{i=1}^{m} (\lambda_i C_i + N_i) = \sum_{i=1}^{m} F_i \tag{5.31}$$

gilt.
Wählen wir nun Basen $B_i = (b_{i1}, \ldots, b_{in_i})$ von U_i, so ist $B = \bigcup_{i=1}^m B_i$ eine Basis von V, und wir erhalten

$$\mathfrak{D}_{B,B}(F) = \begin{pmatrix} \mathfrak{A}_1 & 0 & \cdots & 0 \\ 0 & \mathfrak{A}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{A}_m \end{pmatrix} \tag{5.32}$$

mit $\mathfrak{A}_i = \lambda_i \cdot \mathfrak{E}^{(n_i,n_i)} + \mathfrak{N}_i$, wobei die $\mathfrak{N}_i = \mathfrak{D}_{B_i,B_i}(N_i)$ nilpotente Matrizen sind. Da die Endomorphismen N_i nilpotent sind, können wir jeden der Teilräume U_i gemäß Satz 5.8.10 in eine direkte Summe von Unterräumen $U_i = U_{i1} \oplus \cdots \oplus U_{ir_i}$ mit $n_i \geq \mu(N_i)$ zerlegen, und bezüglich der zyklischen Basen dieser Zerlegung erhalten wir nach Korollar 5.8.11 eine Darstellung

$$\mathfrak{N}_i = \begin{pmatrix} \mathfrak{N}_{i1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{N}_{ir_i} \end{pmatrix}, \tag{5.33}$$

in der die $\mathfrak{N}_{ij}$ alle nilzyklische Matrizen (vgl. (5.21)) sind. Beachten wir noch die Tatsache, daß nach Korollar 5.8.2 ein Endomorphismus genau dann split über K ist, wenn alle seine verschiedenen Eigenwerte in K liegen, so haben wir den folgenden Satz bewiesen.

5.8.14 Satz (Jordansche Normalform eines Endomorphismus) *Sei V ein endlich dimensionaler Vektorraum über einem Körper K und F ein Endomorphismus von V, dessen verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_m$ alle in K liegen. Dann gibt es eine Basis B von V derart, daß*

$$\mathfrak{D}_{B,B}(F) = \begin{pmatrix} \mathfrak{A}_1 & 0 & \cdots & 0 \\ 0 & \mathfrak{A}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{A}_m \end{pmatrix}$$

ist, wobei

$$\mathfrak{A}_i = \lambda_i \cdot \mathfrak{E}^{(n_i,n_i)} + \mathfrak{N}_i,$$

$$\mathfrak{N}_i = \begin{pmatrix} \mathfrak{N}_{i1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{N}_{ir_i} \end{pmatrix}$$

und alle $\mathfrak{N}_{ij}$ nilzyklische Matrizen sind.

5.8.15 Korollar (Jordansche Normalform einer quadratischen Matrix) *Ist $\mathfrak{A} \in K^{n\times n}$ eine quadratische Matrix, deren charakteristisches Polynom über K in Linearfaktoren zerfällt, so gibt es eine invertierbare Matrix $\mathfrak{S} \in \mathrm{GL}(n, K)$ mit*

$$\mathfrak{S}\mathfrak{A}\mathfrak{S}^{-1} = \begin{pmatrix} \mathfrak{A}_1 & 0 & \cdots & 0 \\ 0 & \mathfrak{A}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{A}_m \end{pmatrix},$$

wobei

$$\mathfrak{A}_i = \lambda_i \cdot \mathfrak{E}^{(n_i,n_i)} + \mathfrak{N}_i,$$

$$\mathfrak{N}_i = \begin{pmatrix} \mathfrak{N}_{i1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathfrak{N}_{ir_i} \end{pmatrix}$$

und alle $\mathfrak{N}_{ij}$ nilzyklische Matrizen sind.

Beweis: Wir wenden Satz 5.8.14 auf den Endomorphismus $\overline{\mathfrak{A}}$ an. Da die Wurzeln des Minimalpolynoms mit denen des charakteristischen Polynoms (bis auf Vielfachheiten) übereinstimmen, ist $\overline{\mathfrak{A}}$ split über K. Dann hat $\overline{\mathfrak{A}}$ bezüglich der nach Satz 5.8.13 existierenden Basis B die behauptete Darstellung. Definieren wir $\mathfrak{S}$ als die Transformationsmatrix $\mathfrak{T}_{B,\mathbb{E}^n}$, so folgt die Behauptung mit Satz 3.3.13. $\square$

5.8.16 Korollar *Hat $F \in \mathrm{End}(V)$ bzw. $\mathfrak{A} \in K^{n\times n}$ die in Satz 5.8.14 oder Korollar 5.8.15 dargestellte Jordan-Normalform, so ist $Charpol_F = \prod_{i=1}^m (\lambda_i - x)^{n_i}$ bzw. $Charpol_{\mathfrak{A}} = \prod_{i=1}^m (\lambda_i - x)^{n_i}$.*

Beweis: Kürzen wir mit $\mathfrak{J}$ die Jordan-Normalform ab, so gilt $Charpol_F = \det(\mathfrak{J} - x \cdot \mathfrak{E})$ bzw. $Charpol_{\mathfrak{A}} = \det(\mathfrak{J} - x \cdot \mathfrak{E})$. Nach 4.3.17 (ii) erhalten wir daraus die Behauptung. $\square$

Für ein halbeinfaches $F \in \mathrm{End}(V)$ erhalten wir nach Satz 5.7.4 $Charpol_F(F) = \sum_{i=1}^m Charpol_F(\lambda_i)C_i = 0$, da die Eigenwerte alle Wurzeln des charakteristischen Polynoms sind. Wir wollen dies auf Endomorphismen verallgemeinern, die split sind.

5.8.17 Satz (Hamilton–Cayley) *Ist V ein endlich dimensionaler Vektorraum über einem Körper K und liegen alle Eigenwerte des Endomorphismus $F \in \mathrm{End}(V)$ in K, so gilt $Charpol_F(F) = 0$. Insbesondere teilt das Minimalpolynom von F stets dessen charakteristisches Polynom.*

Beweis: Nach Korollar 5.8.16 gilt $Charpol_F = \prod_{i=1}^m (\lambda_i - x)^{n_i}$. Damit ist nach (5.31) $Charpol_F(F) = \prod_{i=1}^m (\lambda_i \cdot \mathit{id} - F)^{n_i} = \prod_{i=1}^m (\lambda_i \cdot \mathit{id} - \sum_{j=1}^m \lambda_j C_j - N_j)^{n_i}$. Für $v \in V$ erhalten wir $v = u_1 + \cdots + u_m$ mit $u_i \in U_i := C_i[V]$. Also ist mit (5.24) und (5.30) $(\lambda_i \cdot \mathit{id} - F)(u_i) = \lambda_i u_i - \sum_{j=1}^m \lambda_j C_j u_i - N(u_i) = -N(u_i)$. Da U_i invariant unter N_i ist, folgt $(\lambda_i \cdot \mathit{id} - F)^k(u_i) = (-1)^k N_i^k(u_i)$ und wegen $n_i = \mu(N_i)$ schließlich $(\lambda_i \cdot \mathit{id} - F)^{n_i}(u_i) = (-1)^k N_i^{n_i}(u_i) = 0$. Damit ist $U_i \subseteq$ Kern $(\lambda_i \cdot \mathit{id} - F)^{n_i}$ und folglich $V = U_1 \oplus \cdots \oplus U_m \subseteq$ Kern $(\prod_{i=1}^m (\lambda_i \cdot \mathit{id} - F)^{n_i})$, d.h. $\prod_{i=1}^m (\lambda_i \cdot \mathit{id} - F)^{n_i} = 0$. Aus Lemma 5.6.5 (iii) folgt nun, daß $Charpol_F$ von $Minpol_F$ geteilt wird. $\square$

In der Sprache der Matrizen lautet der Satz von Hamilton–Cayley wie folgt.

5.8.18 Satz *Ist $\mathfrak{A} \in K^{n\times n}$ eine quadratische Matrix deren Eigenwerte alle in K liegen, so gilt $Charpol_{\mathfrak{A}}(\mathfrak{A}) = 0$. Insbesondere ist das Minimalpolynom von $\mathfrak{A}$ ein Teiler des charakteristischen Polynoms von $\mathfrak{A}$.*

6. Euklidische und unitäre Vektorräume

6.1 Bilinearformen

6.1.1 Definition Seien V und W Vektorräume über einem gemeinsamen Grundkörper K. Eine Abbildung $\sigma\colon V \times W \longrightarrow K$ heißt eine *Bilinearform*, wenn σ in beiden Argumenten linear ist, d.h. wenn

$$\sigma(\alpha u + \beta v, w) = \alpha \cdot \sigma(u, w) + \beta \cdot \sigma(v, w)$$

und

$$\sigma(u, \alpha v + \beta w) = \alpha \cdot \sigma(u, v) + \beta \cdot \sigma(u, w)$$

gelten. Wir definieren den linken und rechten *Bilinearkern* einer Bilinearform σ als die Mengen

$$\text{Bkern}\,(\sigma, V) := \{v \in V \mid (\forall w \in W)[\sigma(v, w) = 0]\}$$

und

$$\text{Bkern}\,(\sigma, W) := \{w \in W \mid (\forall v \in V)[\sigma(v, w) = 0]\}.$$

6.1.2 Lemma *Ist $\sigma\colon V \times W \longrightarrow K$ eine Bilinearform, so ist* Bkern (σ, V) *ein Teilraum von V und* Bkern (σ, W) *ein Teilraum von W.*

Beweis: Klar. □

6.1.3 Lemma *Seien V und W K-Vektorräume. Die Menge* BL (V, W) *der Linearformen auf $V \times W$ bildet mit den Verknüpfungen $(\sigma + \tau)(u, v) := \sigma(u, v) + \tau(u, v)$ und $(\alpha \cdot \sigma)(u, v) := \alpha \cdot \sigma(u, v)$ einen K-Vektorraum. Ist $V = W$, so notieren wir* BL (V, V) *als* BL (V).

Beweis: Klar durch Nachrechnen. □

6.1.4 Definition Eine Bilinearform $\sigma: V \times V \longrightarrow K$ heißt

symmetrisch, wenn $\quad (\forall x \in V)(\forall y \in V)[\sigma(x,y) = \sigma(y,x)]$

und

alternierend oder *antisymmetrisch* , wenn $\quad (\forall x \in V)(\forall y \in V)[\sigma(x,y) = -\sigma(y,x)]$

gilt.
Für symmetrische und alternierende Bilinearformen stimmen die rechten und linken Bilinearkerne überein. Wir notieren beide daher als

$$\text{Bkern}\,(\sigma) = \{v \in V \mid (\forall x \in V)[\sigma(x,v) = 0]\} = \{v \in V \mid (\forall x \in V)[\sigma(v,x) = 0]\}.$$

Ein symmetrische oder alternierende Bilinearform heißt *nicht ausgeartet,* wenn Bkern $(\sigma) = \{0\}$ ist. Wenn wir von einer nicht ausgearteten Bilinearform σ sprechen, antizipieren wir immer, daß σ entweder symmetrisch oder alternierend ist.

6.1.5 Bemerkung Ist $\sigma: V \times V \longrightarrow K$ eine Bilinearform und $a \in V$, so erhalten wir eine Abbildung $\sigma_a: V \longrightarrow K$ durch $\sigma_a(x) := \sigma(a,x)$. Dann ist $\sigma_a \in V^*$, und wegen $\sigma_{\alpha a+\beta b}(x) = \sigma(\alpha a + \beta b, x) = \alpha\sigma(a,x) + \beta\sigma(b,x) = (\alpha\sigma_a + \beta\sigma_b)(x)$ ist die Abbildung $a \longmapsto \sigma_a$ ein Homomorphismus der Vektorräume V und V^*.

6.1.6 Satz *Ist σ eine nicht ausgeartete Bilinearform auf V, so ist die Abbildung $\sigma^*: V \longrightarrow V^* \quad \sigma^*(a) := \sigma_a$ ein Monomorphismus der Vektorräume. Ist V endlich dimensional, so ist σ^* ein Isomorphismus.*

Beweis: Es gilt

$$\begin{aligned} \sigma^*(a) = 0 \quad &\Leftrightarrow \quad \sigma_a = 0 \\ &\Leftrightarrow \quad (\forall x \in V)[\sigma(a,x) = 0] \\ &\Leftrightarrow \quad a \in \text{Bkern } \sigma. \end{aligned}$$

Damit ist

$$\text{Kern}\,(\sigma^*) = \text{Bkern}\,(\sigma). \tag{6.1}$$

Ist σ nicht ausgeartet, so ist Bkern $(\sigma) = \{0\}$ und damit σ^* injektiv. Ist $\dim V$ endlich, so ist $\dim V = \dim V^*$, und nach Satz 2.3.9 ist σ^* somit bijektiv. □

6.1.7 Satz *Sei V ein endlich dimensionaler K-Vektorraum, σ eine nicht ausgeartete Bilinearform auf V und $F \in$ End(V). Dann gibt es einen Endomorphismus $F^\sigma \in$ End(V) mit $\sigma(a, Fb) = \sigma(F^\sigma a, b)$ für alle $a, b \in V$.*

Beweis: Definieren wir $\lambda(x) := \sigma(a, Fx)$, so ist $\lambda\colon V \longrightarrow K$ linear, d.h. $\lambda \in V^*$. Nach Satz 6.1.6 gibt es daher ein eindeutig bestimmtes $c \in V$ mit $\lambda = \sigma^*(c)$. Wir definieren $F^\sigma a := c$, d.h. $\sigma^*(F^\sigma a)(x) = \sigma(a, Fx)$. Damit folgt

$$\begin{aligned}\sigma^*(F^\sigma(\alpha a + \beta b))(x) &= \sigma(\alpha a + \beta b, Fx) = \alpha\sigma(a, Fx) + \beta\sigma(b, Fx)\\ &= \alpha\sigma^*(F^\sigma a)(x) + \beta\sigma^*(F^\sigma b)(x)\\ &= (\alpha\sigma^*(F^\sigma a) + \beta\sigma^*(F^\sigma b))(x) = \sigma^*(\alpha F^\sigma a + \beta F^\sigma b)(x)\end{aligned}$$

und, da σ^* bijektiv ist, damit $F^\sigma(\alpha a + \beta b) = \alpha F^\sigma a + \beta F^\sigma b$. Also ist $F^\sigma \in \operatorname{End}(V)$. □

Man nennt den Endomorphismus F^σ den bezüglich σ zu F *adjungierten* Endomorphismus. Gilt $A^\sigma = A$, so nennt man A bezüglich σ *selbstadjungiert*.

6.1.8 Bemerkung Die adjungierte und die duale Abbildung sind eng miteinander verknüpft. Wir erhalten nämlich $\sigma^*(F^\sigma a)(x) = \sigma(a, Fx) = \sigma_a(Fx) = F^*(\sigma_a)(x) = F^*(\sigma^*(a))(x)$, d.h. $\sigma^*(F^\sigma(a)) = F^*(\sigma^*(a))$ und damit

$$\sigma^* \circ F^\sigma = F^* \circ \sigma^*. \tag{6.2}$$

Gilt $\sigma^* \circ F = \sigma^* \circ G$, so folgt $(\forall v \in V)[\sigma^*(F(v)) = \sigma^*(G(v))]$ und wegen der Bijektivität von σ^* daher $(\forall v \in V)[F(v) = G(v)]$, d.h. $F = G$.

6.1.9 Satz *Sei V ein endlich dimensionaler K-Vektorraum und σ eine nicht ausgeartete Bilinearform auf V. Die Abbildung ${}^\sigma\colon \operatorname{End}(V) \longrightarrow \operatorname{End}(V)$ ist dann ein Isomorphismus der Vektorräume und ein involutorischer Antiisomorphismus der Ringe.*

Beweis: Wir zeigen zunächst, daß ${}^\sigma$ eine lineare Abbildung ist. Wir erhalten

$$\begin{aligned}&\sigma^* \circ (\alpha F + \beta G)^\sigma = (\alpha F + \beta G)^* \circ \sigma^* = (\alpha F^* + \beta G^*) \circ \sigma^* =\\ &= \alpha F^* \circ \sigma^* + \beta G^* \circ \sigma^* = \sigma^* \circ \alpha F^\sigma + \sigma^* \circ \beta G^\sigma = \sigma^* \circ (\alpha F^\sigma + \beta G^\sigma)\end{aligned}$$

und damit $(\alpha F + \beta G)^\sigma = \alpha F^\sigma + \beta G^\sigma$, d.h. ${}^\sigma$ ist linear. Nun gilt

$$\begin{aligned}F^\sigma = 0 \;&\Leftrightarrow\; (\forall a \in V)[F^\sigma(a) = 0]\\ &\Leftrightarrow\; (\forall x \in V)(\forall a \in V)[\sigma(a, Fx) = 0]\\ &\Leftrightarrow\; (\forall x \in V)[Fx \in \text{Bkern } \sigma].\end{aligned}$$

Da σ nicht ausgeartet ist, folgt $Fx = 0$ für alle $x \in V$ und damit $F = 0$. Damit ist ${}^\sigma$ injektiv und als Endomorphismus des endlich dimensionalen Vektorraumes $\operatorname{End}(V)$ damit auch bijektiv. Es bleibt nur noch zu zeigen, daß ein involutorischer Antihomomorphismus der Ringe vorliegt. Dies rechnen wir einfach nach.

Zunächst ist $\sigma((FG)^\sigma a, b) = \sigma(a, FGb) = \sigma(F^\sigma a, Gb) = \sigma(G^\sigma F^\sigma a, b)$. Also folgt $(FG)^\sigma = G^\sigma F^\sigma$ und es liegt ein Antihomomorphismus vor. Aus der Symmetrie von σ erhalten wir weiter $\sigma(u, F^{\sigma\sigma}v) = \sigma(F^{\sigma\sigma}v, u) = \sigma(v, F^\sigma u) = \sigma(F^\sigma u, v) = \sigma(u, Fv)$. Da σ nicht ausgeartet ist, folgt $F^{\sigma\sigma}v = Fv$ für alle $v \in V$, d.h. $F^{\sigma\sigma} = F$ und $^\sigma$ ist involutorisch. Ist σ alternierend, so führen wir die gleiche Rechnung durch, wobei wir beim Vertauschen der Argumente jedes Mal das Vorzeichen zu wechseln haben. Da wir die Argumente genau zweimal vertauschen, ändert dies am Ergebnis der Rechnung nichts. □

6.1.10 Satz *Sei V ein endlich dimensionaler Vektorraum und σ eine nicht ausgeartete Bilinearform auf V. Dann gibt es einen Isomorphismus $f\colon \mathrm{BL}(V) \longrightarrow \mathrm{End}(V)$ mit $\tau(x,y) = \sigma(f(\tau)x, y)$ für alle $\tau \in \mathrm{BL}(V)$, d.h. daß sich jede Bilinearform auf V in der Form $\tau(x,y) = \sigma(Ax, y)$ für ein $A \in \mathrm{End}(V)$ darstellen läßt.*

Beweis: Ist $g(x_1, \ldots, x_n)$ eine Funktion, so bezeichnen wir mit $\lambda x_k \,.\, g(x_1, \ldots, x_n)$ die Funktion, die sich aus $g(x_1, \ldots, x_n)$ ergibt, wenn wir nur das Argument x_k als variabel betrachten und die übrigen konstant lassen. Für $a \in V$ und $\tau \in \mathrm{BL}(V)$ ist dann $\lambda x \,.\, \tau(a,x)$ eine Linearform. Nach Satz 6.1.6 gibt es daher ein $a' \in V$ mit $\lambda x \,.\, \tau(a,x) = \sigma^*(a')$. Wir rechnen nach, daß die Abbildung $a \longmapsto a'$ linear ist. Es ist $\lambda x \,.\, \tau(\alpha a + \beta b, x) = \alpha \cdot \lambda x \,.\, \tau(a,x) + \beta \cdot \lambda x \,.\, \tau(b,x) = \alpha\sigma^*(a') + \beta\sigma^*(b') = \sigma^*(\alpha a' + \beta b')$ und damit $(\alpha a + \beta b)' = \alpha a' + \beta b'$. Nun definieren wir $f(\tau)(a) := a'$. Damit ist $f(\tau) \in \mathrm{End}(V)$, und es gilt $\sigma^*(f(\tau)a) = \lambda x \,.\, \tau(a,x)$. Wir haben nachzuprüfen, daß f ein Homomorphismus ist. Dazu berechnen wir $\sigma^*(f(\alpha\tau + \beta\rho)a) = \lambda x \,.\, (\alpha\tau + \beta\rho)(a,x) = \alpha \cdot \lambda x \,.\, \tau(a,x) + \beta \cdot \lambda x \,.\, \rho(a.x) = \alpha \cdot \sigma^*(f(\tau)a) + \beta \cdot \sigma^*(f(\rho)a) = \sigma^*(\alpha f(\tau)a + \beta f(\rho)a)$. Also ist $f(\alpha\tau + \beta\rho) = \alpha f(\tau) + \beta f(\rho)$ und damit f ein Homomorphismus, für den $\lambda x \,.\, \tau(a,x) = \sigma^*(f(\tau)a) = \lambda x \,.\, \sigma(f(\tau)a, x)$ gilt. Ist $f(\tau) = 0$, so ist $\sigma^*(f(\tau)a) = 0$ und damit $\lambda x \,.\, \tau(a,x) = 0$ für alle $a \in V$. Das bedeutet aber, daß $\tau = 0$ ist. Also ist Kern $f = \{0\}$ und f injektiv. Umgekehrt gilt $\tau := \lambda xy \,.\, \sigma(Ax, y) \in \mathrm{BL}(V)$ und $f(\tau) = A$ für jedes $A \in \mathrm{End}(V)$. Damit ist f auch surjektiv, und wir haben einen Isomorphismus der Vektorräume vorliegen.

□

6.1.11 Lemma *Ist σ eine nicht ausgeartete, symmertische Bilinearform und bezeichnet $f\colon \mathrm{BL}(V) \longrightarrow \mathrm{End}(V)$ den in Satz 6.1.10 definierten Isomorphismus, so ist*

(i) die Bilinearform τ genau dann symmetrisch, wenn $f(\tau)$ bezüglich σ selbstadjungiert ist

und

(ii) $\operatorname{Bkern}(\tau) = \operatorname{Kern}(f(\tau))$ *und damit ist* τ *genau dann nicht ausgeartet, wenn* $f(\tau) \in \mathrm{GL}(V)$ *ist.*

Beweis: (i): Ist τ symmetrisch, so folgt $\sigma(f(\tau)y, x) = \tau(y, x) = \tau(x, y) = \sigma(f(\tau)x, y) = \sigma(y, f(\tau)x) = \sigma(f(\tau)^\sigma y, x)$ und damit $f(\tau)^\sigma = f(\tau)$, d.h. $f(\tau)$ ist selbstadjungiert bezüglich σ. Ist umgekehrt $f(\tau)$ selbstadjungiert bezüglich σ, so folgt $\tau(x, y) = \sigma(f(\tau)x, y) = \sigma(f(\tau)^\sigma x, y) = \sigma(x, f(\tau)y) = \sigma(f(\tau)y, x) = \tau(y, x)$, und τ ist symmetrisch.
(ii): Da σ als nicht ausgeartet vorausgesetzt ist, gilt

$$\begin{aligned} x \in \operatorname{Bkern}(\tau) &\Leftrightarrow (\forall y \in V)[\tau(x, y) = 0] \\ &\Leftrightarrow (\forall y \in V)[\sigma(f(\tau)x, y) = 0] \\ &\Leftrightarrow f(\tau)x \in \operatorname{Bkern}(\sigma) \\ &\Leftrightarrow f(\tau)x = 0 \quad \Leftrightarrow \quad x \in \operatorname{Kern}(f(\tau)). \end{aligned}$$

Also ist $\operatorname{Bkern}(\tau) = \operatorname{Kern}(f(\tau))$, und der Rest ist klar. □

6.1.12 Definition Für eine Bilinearform $\sigma \in \mathrm{BL}(V)$ und eine Basis $B = (b_1, \ldots, b_n)$ definieren wir die Matrixdarstellung von σ bezüglich B durch

$$\mathfrak{D}_B(\sigma) := (\sigma_{ij})_{n,n} := (\sigma(b_i, b_j))_{n,n}.$$

6.1.13 Satz *Sei* V *ein* K*-Vektorraum mit* $\dim V = n$*. Dann ist die Abbildung* $\mathfrak{D}_B \colon \mathrm{BL}(V) \longrightarrow K^{n\times n}$ *ein Isomorphismus der Vektorräume. Für* $u, v \in V$ *gilt*

$$\sigma(u, v) = h_B(u) \cdot \mathfrak{D}_B(\sigma) \cdot \hat{h}_B(v) = h_B(u) \cdot \mathfrak{D}_B(\sigma) \cdot (h_B(v))^t. \tag{6.3}$$

Beweis: Zunächst berechnen wir

$$\begin{aligned} \mathfrak{D}_B(\alpha\sigma + \beta\tau) &= ((\alpha\sigma + \beta\tau)(b_i, b_j))_{n,n} = (\alpha\sigma(b_i, b_j) + \beta\tau(b_i, b_j))_{n,n} \\ &= \alpha \cdot (\sigma(b_i, b_j))_{n,n} + \beta \cdot (\tau(b_i, b_j))_{n,n} = \alpha\mathfrak{D}_B(\sigma) + \beta\mathfrak{D}_B(\tau). \end{aligned}$$

Damit liegt ein Homomorphismus vor. Für $u = \sum_{i=1}^n \alpha_i b_i$ und $v = \sum_{j=1}^n \beta_j b_j$ erhalten wir

$$\begin{aligned} \sigma(u, v) &= \sum_{j=1}^n \beta_j (\sum_{i=1}^n \alpha_i \sigma(b_i, b_j)) = \left(\sum_{i=1}^n \alpha_i \sigma(b_i, b_1), \ldots, \sum_{i=1}^n \alpha_i \sigma(b_i, b_n)\right) \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \\ &= (\alpha_1, \ldots, \alpha_n)(\sigma(b_i, b_j))_{n,n} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = h_B(u)\mathfrak{D}_B(\sigma)\hat{h}_B(v). \end{aligned}$$

Damit haben wir (6.3) gezeigt. Ist $\mathfrak{D}_B(\sigma) = 0$, so erhalten wir mit (6.3) aber auch $\sigma(u, v) = 0$ für alle $u, v \in V$ und damit $\sigma = 0$. Also ist $\mathfrak{D}_B$ injektiv und, da nach Satz 6.1.10 $\dim(\mathrm{BL}\,(V)) = \dim(\mathrm{End}(V)) = n^2 = \dim(K^{n\times n})$ ist, auch bijektiv. □

6.1.14 Korollar *Eine Bilinearform σ ist genau dann symmetrisch, wenn $\mathfrak{D}_B(\sigma) = (\mathfrak{D}_B(\sigma))^t$ ist. Eine symmetrische Bilinearform σ ist genau dann nicht ausgeartet, wenn $\mathfrak{D}_B(\sigma)$ invertierbar ist.*

Beweis: Die erste Behauptung erhalten wir direkt aus (6.3). Für die zweite Behauptung lesen wir aus (6.3) ab, daß Bkern $\sigma = \{u \in V \mid \mathfrak{D}_B(\sigma)\hat{h}_B(u) = 0\}$ ist. Damit ist σ genau dann nicht ausgeartet, wenn Kern $(\widehat{\mathfrak{D}_B(\sigma)} \circ \hat{h}_B) = \{0\}$ ist, und dies ist genau dann der Fall, wenn $\widehat{\mathfrak{D}_B(\sigma)}$ bijektiv und damit $\mathfrak{D}_B(\sigma)$ invertierbar ist. □

Wir nennen Matrizen $\mathfrak{S} \in K^{n\times n}$, für die $\mathfrak{S} = \mathfrak{S}^t$ gilt, *symmetrisch*.

6.1.15 Satz *Sei V ein endlich dimensionaler K-Vektorraum und σ eine nicht ausgeartete Bilinearform auf V. Ist $B = (b_1, \ldots, b_n)$ eine Basis von V und $B^* = (b_1^*, \ldots, b_n^*)$ die dazu duale Basis von V^*, so gilt*

$$\mathfrak{D}_B(\sigma) = \mathfrak{D}_{B,B^*}(\sigma^*). \tag{6.4}$$

Beweis: Sei $\mathfrak{D}_{B,B^*}(\sigma^*) = (\alpha_{ij})_{n,n}$. Dann erhalten wir aus $\lambda x\,.\,\sigma(b_i, x) = \sigma^*(b_i) = \sum_{j=1}^n \alpha_{ij} b_j^*$ sofort $\sigma(b_i, b_k) = (\sum_{j=1}^n \alpha_{ij} b_j^*)(b_k) = \alpha_{ik}$ und damit $\mathfrak{D}_B(\sigma) = \mathfrak{D}_{B,B^*}(\sigma^*)$. □

Kombinieren wir (6.2) und das Diagramm in Satz 3.3.6, so erhalten wir die Kommutativität des Diagramms in Abbildung 6.1. Diesem Diagramm entnehmen wir nun sofort

$$\overline{\mathfrak{D}_{B,B}(A^\sigma)} = \overline{\mathfrak{D}_{B,B^*}(\sigma^*)}^{-1} \circ \overline{\mathfrak{D}_{B^*,B^*}(A^*)} \circ \overline{\mathfrak{D}_{B,B^*}(\sigma^*)}.$$

Daraus erhalten wir zusammen mit Satz 6.1.15

$$\mathfrak{D}_{B,B}(A^\sigma) = \mathfrak{D}_B(\sigma) \cdot \mathfrak{D}_{B,B}(A)^t \cdot \mathfrak{D}_B(\sigma)^{-1}.$$

Wir haben also den folgenden Satz bewiesen.

6.1.16 Satz *Sei V ein endlich dimensionaler Vektorraum und σ eine nicht ausgeartete Bilinearform auf V. Dann gilt*

$$\mathfrak{D}_{B,B}(A^\sigma) = \mathfrak{D}_B(\sigma) \cdot \mathfrak{D}_{B,B}(A)^t \cdot \mathfrak{D}_B(\sigma)^{-1}$$

für jeden Endomorphismus $A \in \mathrm{End}(V)$.

$$\begin{tikzcd}
K^n \arrow[rrr, "\overline{\mathfrak{D}_{B,B}(A^\sigma)}"] \arrow[ddd, "\overline{\mathfrak{D}_{B,B^*}(\sigma^*)}"'] & & & K^n \arrow[ddd, "\overline{\mathfrak{D}_{B,B^*}(\sigma^*)}"] \\
& V \arrow[ul, "h_B"'] \arrow[r, "A^\sigma"] \arrow[d, "\sigma^*"] & V \arrow[ur, "h_B"] \arrow[d, "\sigma^*"'] & \\
& V^* \arrow[r, "A^*"'] \arrow[dl, "h_B{}^*"] & V^* \arrow[dr, "h_B{}^*"'] & \\
K^n \arrow[rrr, "\overline{\mathfrak{D}_{B^*,B^*}(A^*)}"'] & & & K^n
\end{tikzcd}$$

Abbildung 6.1: Zusammenhang zwischen adjungierter und dualer Abbildung

Zum Ende dieses Abschnittes untersuchen wir noch das Transformationsverhalten der darstellenden Matrix einer nicht ausgearteten Bilinearform bei Basiswechsel.

6.1.17 Satz *Es sei V ein endlich dimensionaler Vektorraum und σ eine nicht ausgeartete Bilinearform auf V. Sind B und C Basen von V, so gilt $\mathfrak{D}_C(\sigma) = \mathfrak{T}_{C,B} \cdot \mathfrak{D}_B(\sigma) \cdot \mathfrak{T}^t_{C,B}$.*

Beweis: Nach (6.4) ist $\mathfrak{D}_B(\sigma) = \mathfrak{D}_{B,B^*}(\sigma^*)$ für jede Basis B. Damit erhalten wir mit dem Transformationssatz (Satz 3.3.13)

$$\mathfrak{D}_C(\sigma) = \mathfrak{D}_{C,C^*}(\sigma^*) = \mathfrak{T}_{C,B} \cdot \mathfrak{D}_{B,B^*}(\sigma^*) \cdot \mathfrak{T}_{B^*,C^*} = \mathfrak{T}_{C,B} \cdot \mathfrak{D}_B(\sigma) \cdot \mathfrak{T}_{B^*,C^*}. \quad \text{(i)}$$

Es ist $\mathfrak{T}_{B^*,C^*} = \mathfrak{D}_{B^*,C^*}(id_{V^*}) = \mathfrak{D}_{B^*,C^*}(id^*_V) = (\mathfrak{D}_{C,B}(id_V))^t = \mathfrak{T}^t_{C,B}$. Setzen wir dies in (i) ein, so folgt die Behauptung. □

6.2 Positiv definite Bilinearformen

6.2.1 Definition Sei V ein Vektorraum über dem Körper $\mathbb{R}$ der reellen Zahlen. Wir nennen in Zukunft solche Räume *reelle Vektorräume* . Eine symmetrische Bilinearform $\sigma \colon V \times V \longrightarrow \mathbb{R}$ heißt

positiv definit, wenn $(\forall x \in V)[x \neq 0 \;\Rightarrow\; \sigma(x,x) > 0]$ gilt,

positiv semi-definit, wenn $(\forall x \in V)[\sigma(x,x) \geq 0]$ gilt, *negativ definit,* wenn $(\forall x \in V)[x \neq 0 \;\Rightarrow\; \sigma(x,x) < 0]$ gilt,

indefinit, wenn $(\exists x \in V)(\exists y \in V)[\sigma(x,x) > 0 \wedge \sigma(y,y) < 0]$ gilt.

Die Abbildung

$$\begin{aligned} q_\sigma : V &\longrightarrow \mathbb{R} \\ x &\longmapsto \sigma(x,x) \end{aligned}$$

heißt die von der symmetrischen Bilinearform σ induzierte *quadratische Form.*

6.2.2 Satz *Sei V ein reeller Vektorraum. Jede positiv definite Bilinearform auf V ist nicht ausgeartet.*

Beweis: Ist σ positiv definit und $v \in$ Bkern σ, so folgt $(\forall x \in V)[\sigma(v,x) = 0]$ und damit insbesondere $\sigma(v,v) = 0$. Daraus ergibt sich aber $v = 0$, und es ist Bkern $\sigma = \{0\}$, d.h. σ ist nicht ausgeartet. □

6.2.3 Satz *Es sei V ein reeller Vektorraum und σ eine positiv semi-definite Bilinearform auf V. Dann gilt die* Cauchy-Schwarzsche Ungleichung

$$(\forall u,v \in V)[\sigma(u,v)^2 \leq \sigma(u,u)\sigma(v,v)]. \tag{6.5}$$

Für positiv definite Bilinearformen σ steht in Gleichung (6.5) genau dann das Gleichheitszeichen, wenn u und v linear abhängig sind.

Beweis: Für $\alpha, \beta \in \mathbb{R}$ erhalten wir

$$0 \leq \sigma(\alpha u + \beta v, \alpha u + \beta v) = \alpha^2\sigma(u,u) + 2\alpha\beta\sigma(u,v) + \beta^2\sigma(v,v). \tag{i}$$

Ist $\sigma(u,u) = 0$, so wählen wir $\beta = 1$ und erhalten $2\alpha\sigma(u,v) + \sigma(v,v) \geq 0$ für alle $\alpha \in \mathbb{R}$, was nur für $\sigma(u,v) = 0$ möglich ist. Ist $\sigma(u,u) \neq 0$, so wählen wir $\beta := \sigma(u,u)$ und teilen Ungleichung (i) durch β. Das ergibt

$$0 \leq \alpha^2 + 2\alpha\sigma(u,v) + \sigma(u,u)\sigma(v,v). \tag{ii}$$

Nun wählen wir $\alpha := -\sigma(u,v)$ in (ii) und erhalten

$$0 \leq -\sigma(u,v)^2 + \sigma(u,u)\sigma(v,v), \tag{iii}$$

d.h. $\sigma(u,v)^2 \leq \sigma(u,u)\sigma(v,v)$.
Sind u und v linear abhängig, so ist $u = \lambda v$ für ein $\lambda \in \mathbb{R}$, und es folgt $\sigma(u,v)^2 = \lambda^2\sigma(v,v)^2 = \sigma(\lambda v, \lambda v)\sigma(v,v)$. Sind u und v linear unabhängig und ist σ positiv definit, so steht für $\alpha \neq 0$ in (i) das $<$-Zeichen. Dann ist aber auch $\sigma(u,u) > 0$, womit folgt, daß auch in (iii) das $<$-Zeichen steht. Damit folgt aber $\sigma(u,v)^2 < \sigma(u,u)\sigma(v,v)$. □

6.2.4 Definition Sei V ein reeller Vektorraum. Eine Abbildung $\|\|: V \longrightarrow \mathbb{R}$ heißt eine *Normfunktion* oder kurz eine *Norm* auf V, wenn sie den folgenden Bedingungen genügt:

(No.0) $\|0\| = 0$,

(No.1) $(\forall x \in V)[x \neq 0 \ \Rightarrow \ \|x\| > 0]$,

(No.2) $(\forall \alpha \in \mathbb{R})(\forall x \in V)[\|\alpha x\| = |\alpha|\|x\|]$,

(No.3) $(\forall x, y \in V)[\|x + y\| \leq \|x\| + \|y\|]$.

Ein reeller Vektorraum, auf dem sich eine Norm definieren läßt, heißt eine *normierter Raum*.

In normierten Räumen lassen sich Längen messen. Um dies genauer zu erläutern, benötigen wir den Begriff eines metrischen Raumes.

6.2.5 Definition Sei M eine nicht leere Menge. Eine Abbildung $d: M \times M \longrightarrow \mathbb{R}$, die die Bedingungen

(D.0) $(\forall x \in M)[d(x, x) = 0]$,

(D.1) $(\forall x, y \in M)[x \neq y \ \Rightarrow \ d(x, y) > 0]$,

(D.2) $(\forall x, y \in M)[d(x, y) = d(y, x)]$,

(D.3) $(\forall x, y, z \in M)[d(x, y) \leq d(x, z) + d(z, y)]$

erfüllt, heißt eine *Metrik* oder *Abstandsfunktion* auf M. Das Paar (M, d) heißt ein *metrischer Raum*. Die Bedingung (D.3) nennt man die *Dreiecksungleichung* der Metrik.

6.2.6 Satz *Ist V ein normierter Raum, so definiert die Abbildung $d: V \times V \longrightarrow \mathbb{R}$ mit $d(u, v) := \|u - v\|$ eine Metrik auf V.*

Beweis: Dies folgt durch einfaches Nachrechnen. □

6.2.7 Satz *Ist V ein reeller Vektorraum und σ eine positiv definite Bilinearform auf V, so definiert $\|v\|_\sigma := \sqrt{\sigma(v, v)}$ eine Norm auf V. Insbesondere trägt V dann eine Metrik.*

Beweis: Die Eigenschaften (No.0) - (No.2) rechnet man einfach nach. Um (No.3) zu erhalten, folgern wir aus der Cauchy-Schwarzschen Ungleichung (6.5)

$$(\forall x, y \in V)[\sigma(x, y) \leq \|x\|_\sigma \|y\|_\sigma]. \tag{6.6}$$

Damit berechnen wir

$$\begin{aligned}\sigma(x+y,x+y) &= \sigma(x,x)+2\sigma(x,y)+\sigma(y,y)\\ &\le \sigma(x,x)+2\|x\|_\sigma\|y\|_\sigma+\sigma(y,y) = (\|x\|_\sigma+\|y\|_\sigma)^2\end{aligned} \tag{i}$$

, und aus (i) folgt

$$\|x+y\|_\sigma \le \|x\|_\sigma + \|y\|_\sigma. \qquad \square$$

6.2.8 Beispiele 1. Sei $V=\mathbb{R}^n$. Für $\mathfrak{x},\mathfrak{y}\in\mathbb{R}^n$ definieren wir $\langle\mathfrak{x},\mathfrak{y}\rangle := \mathfrak{x}\cdot\mathfrak{y}^t$. Offensichtlich definiert dies eine Bilinearform. Für $\mathfrak{x}=(\xi_1,\ldots,\xi_n)$ und $\mathfrak{y}=(\eta_1,\ldots,\eta_n)$ erhalten wir dann $\langle\mathfrak{x},\mathfrak{y}\rangle=\sum_{i=1}^n\xi_i\eta_i$. Insbesondere ist $\|\mathfrak{x}\|=\sqrt{\langle\mathfrak{x},\mathfrak{x}\rangle}=\sqrt{\sum_{i=1}^n\xi_i^2}$. Daraus ersehen wir, daß $\langle\ ,\ \rangle$ positiv definit ist. Man nennt $\langle\mathfrak{x},\mathfrak{y}\rangle$ das *kanonische Skalarprodukt* auf $\mathbb{R}^n$. Offensichtlich erhalten wir $\mathfrak{D}_{\mathbb{E}^n}(\lambda\mathfrak{x}\mathfrak{y}\,.\,\langle\mathfrak{x},\mathfrak{y}\rangle)=\mathfrak{E}^{n,n}$.
Es folgt aus (6.3), daß die symmetrischen Bilinearformen auf $\mathbb{R}^n$ alle die Form $\sigma_{\mathfrak{S}}(\mathfrak{x},\mathfrak{y})=\mathfrak{x}\cdot\mathfrak{S}\cdot\mathfrak{y}^t$ mit einer symmetrischen Matrix $\mathfrak{S}$ haben. Hier erhalten wir $\mathfrak{D}_{\mathbb{E}^n}(\sigma_{\mathfrak{S}})=\mathfrak{S}$.
Ist die von der symmetrischen Matrix $\mathfrak{S}$ induzierte Bilinearform $\lambda\mathfrak{x}\mathfrak{y}\,.\,\sigma_{\mathfrak{S}}(\mathfrak{x},\mathfrak{y})$ positiv definit, so nennen wir die Matrix $\mathfrak{S}$ *positiv definit*. Wegen $\sigma_{\mathfrak{S}}(\mathbf{e}_i,\mathbf{e}_i)=\sigma_{ii}$ für $\mathfrak{S}=(\sigma_{ij})_{n,n}$ folgt, daß die Diagonalelemente einer positiv definiten Matrix stets positiv sind.

2. Es seien $\alpha,\beta\in\mathbb{R}$, $\alpha\le\beta$ und $V=C[\alpha,\beta]:=\{f\,|\ f\colon[\alpha,\beta]\longrightarrow\mathbb{R}\wedge\ f \text{ ist stetig}\}$. Man prüft mühelos nach, daß $C[\alpha,\beta]$ ein reeller Vektorraum ist. Nun definieren wir eine Bilinearform auf $C[a,b]$ durch $\sigma(f,g):=\int_\alpha^\beta f(\xi)g(\xi)d\xi$. Es folgt aus den Rechenregeln für Integrale, daß eine symmetrische Bilinearform vorliegt. Ist $f\ne 0$, so finden wir wegen der Stetigkeit von f ein $\gamma\in[\alpha,\beta]$, ein $\epsilon>0$ und ein $\delta>0$, so daß $f(\xi)^2\ge\epsilon>0$ für alle $\xi\in[\gamma-\delta,\gamma+\delta]$ gilt. Damit erhalten wir $\sigma(f,f)=\int_\alpha^\beta f^2(\xi)d\xi\ge\int_{\gamma-\delta}^{\gamma+\delta}f^2(\xi)d\xi\ge 2\delta\cdot\epsilon>0$. Die Bilinearform ist somit positiv definit. Diese Bilinearform spielt in der Funktionalanalysis eine Rolle.

6.3 Euklidische Vektorräume

6.3.1 Definition Ein Paar (V,σ) heißt ein *euklidischer Vektorraum*, wenn gilt:

(EVR.1) *V ist ein reeller Vektorraum.*

(EVR.2) *Die Abbildung $\sigma\colon V\times V\longrightarrow\mathbb{R}$ ist eine positiv definite Bilinearform. Man nennt σ das* Skalarprodukt *des euklidischen Raumes.*

6.3.2 Definition Ist (V,σ) ein euklidischer Vektorraum, so ist

$$\|x\|_\sigma := \sqrt{\sigma(x,x)}$$

eine Norm auf V. Man nennt $\|x\|_\sigma$ die *Länge* des Vektors x in (V,σ). Ein Vektor $v \in V$ heißt *normiert*, wenn $\|v\|_\sigma = 1$ ist. Wegen $\|\|v\|_\sigma{}^{-1} v\|_\sigma = 1$ läßt sich jeder von Null verschiedene Vektor in V normieren.
Nach Satz 6.2.6 definiert

$$d_\sigma(x,y) := \|x-y\|_\sigma$$

eine Metrik auf V.
Gilt $\sigma(x,y) = 0$, so heißen die Vektoren x und y *orthogonal* in (V,σ).

6.3.3 Winkelmessung In euklidischen Räumen haben wir nach Definition 6.3.2 den Begriff der Orthogonalität. Wir können also das “Senkrecht Stehen” zweier Vektoren beschreiben. Einen ähnlichen Begriff haben wir in Definition 3.2.13 schon zwischen den Vektoren eines beliebigen Vektorraums V und den Vektoren des dualen Raumes V^* eingeführt. Da für endlich dimensionale Vektorräume V und V^* isomorph sind, können wir nach Wahl einer Basis von B den Begriff des “Senkrecht Stehens” bezüglich einer Basis durch $v \perp w \quad :\Leftrightarrow \quad v^*(w) = 0$ definieren, wobei * der in Satz 3.2.7 eingeführte Isomorphismus ist. Daß dies eng mit der in euklidischen Räumen eingeführten Orthogonalität verknüpft ist, folgt aus Satz 6.1.6. Allerdings können wir in euklidischen Räumen die Winkelmessung verfeinern. Aus der Ungleichung von Cauchy-Schwarz erhalten wir $|\sigma(x,y)| \le \|x\|_\sigma \|y\|_\sigma$ und daher

$$(\forall x, y \in V \setminus \{0\}) \left[-1 \le \frac{\sigma(x,y)}{\|x\|_\sigma \|y\|_\sigma} \le 1 \right],$$

und wir definieren

$$\Theta(x,y) := \arccos \frac{\sigma(x,y)}{\|x\|_\sigma \|y\|_\sigma} \in [0,\pi]. \tag{6.7}$$

Wir nennen $\Theta(x,y)$ den von x und y eingeschlossenen *Winkel*. Nach Definition gilt dann

$$\sigma(x,y) = \|x\|_\sigma \|y\|_\sigma \cos\Theta(x,y). \tag{6.8}$$

6.3.4 Satz *Sei (V,σ) ein euklidischer Raum und $x, y \in V \setminus \{0\}$. Für den von x und y eingeschlossenen Winkel gelten dann*

(1) $\Theta(x,y) = \Theta(y,x)$,

(2) $\Theta(x,-y) = \pi - \Theta(x,y)$,

(3) $\alpha \ne 0 \land \beta \ne 0 \;\Rightarrow\; \Theta(\alpha x, \beta y) = \Theta(x,y)$,

(4) *x und y sind genau dann orthogonal, wenn $\Theta(x, y) = \frac{\pi}{2}$ ist,*

(5) *x und y sind genau dann linear abhängig, wenn $\Theta(x, y) = 0$ oder $\Theta(x, y) = \pi$ ist,*

(6) $\|x - y\|_\sigma^2 = \|x\|_\sigma^2 + \|y\|_\sigma^2 - 2\|x\|_\sigma\|y\|_\sigma \cos(\Theta(x, y))$.

Die Aussage (6) in Satz 6.3.4 ist als *Cosinus-Satz* bekannt. Stehen x und y senkrecht, so erhalten wir den *Satz des Pythagoras*.
Beweis: Die Behauptungen (1) und (3) folgen sofort aus der Definition von Θ, die Behauptungen (2) und (4) aus den Eigenschaften des Cosinus. Aus Satz 6.2.3 erhalten wir $\frac{|\sigma(x,y)|}{\|x\|_\sigma\|y\|_\sigma} = 1$ für linear abhängige Vektoren x und y. Also ist $\cos(\Theta(x, y)) \in \{-1, 1\}$, d.h. $\Theta(x, y) = \pi$ oder $\Theta(x, y) = 0$.
Um (6) zu zeigen, berechnen wir $\|x-y\|_\sigma^2 = \sigma(x-y, x-y) = \sigma(x, x)+\sigma(y, y)-2\sigma(x, y) = \|x\|_\sigma^2 + \|y\|_\sigma^2 - 2\|x\|_\sigma\|y\|_\sigma \cos(\Theta(x, y))$ nach (6.8). $\square$

6.3.5 Definition Ist $U \subseteq V$, so definieren wir

$$U^{\perp_\sigma} := \{x \in V \mid (\forall y \in U)[\sigma(x, y) = 0]\}$$

und nennen $U^{\perp_\sigma}$ das *orthogonale Komplement* von U in (V, σ).

6.3.6 Lemma *Für einen endlich dimensionalen euklidischen Vektorraum (V, σ) und eine Teilmenge $U \subseteq V$ gilt $\sigma^*[U^{\perp_\sigma}] = U^\perp$.*

Beweis: Wir haben

$$\begin{aligned}
x \in \sigma^*[U^{\perp_\sigma}] &\Leftrightarrow (\exists v \in U^{\perp_\sigma})[x = \sigma^*(v)]\\
&\Leftrightarrow (\exists v)[(\forall y \in U)(\sigma(v, y) = 0) \wedge x = \sigma^*(v)]\\
&\Leftrightarrow (\exists v)[(\forall y \in U)(\sigma^*(v)(y) = 0) \wedge x = \sigma^*(v)]\\
&\Leftrightarrow (\forall y \in U)(x(y) = 0) \quad \Leftrightarrow \quad x \in U^\perp.
\end{aligned}$$

$\square$

6.3.7 Satz *Ist (V, σ) ein endlich dimensionaler euklidischer Vektorraum und U ein Teilraum von V, so gilt $\dim V = \dim U + \dim U^{\perp_\sigma}$ und $V = U \oplus U^{\perp_\sigma}$.*

Beweis: Nach Satz 3.2.17 gilt $\dim V = \dim U + \dim U^\perp$ für endlich dimensionale Vektorräume V und Teilräume $U \subseteq V$. Da $\sigma^* : V \longrightarrow V^*$ ein Isomorphismus ist, erhalten wir aus Lemma 6.3.6 $\dim V = \dim U + \dim U^{\perp_\sigma}$. Ist $u+v = 0$ für $u \in U$ und $v \in U^{\perp_\sigma}$, so folgt $0 = \sigma(u + v, u + v) = \sigma(u, u) + \sigma(v, v) + 2\sigma(u, v) = \sigma(u, u) + \sigma(v, v)$ und damit $u = v = 0$. Damit ist $U + U^{\perp_\sigma} = U \oplus U^{\perp_\sigma} \subseteq V$ und wegen $\dim(U \oplus U^{\perp_\sigma}) = \dim U + \dim(U^{\perp_\sigma}) = \dim V$ schließlich $V = U \oplus U^{\perp_\sigma}$.
$\square$

6.3.8 Korollar *Ist (V, σ) ein endlich dimensionaler euklidischer Vektorraum und U ein Teilraum von V, so gilt $(U^{\perp_\sigma})^{\perp_\sigma} = U$.*

Beweis: Wir haben $U \subseteq (U^{\perp_\sigma})^{\perp_\sigma}$. Wegen $\dim U = \dim((U^{\perp_\sigma})^{\perp_\sigma})$ gilt Gleichheit. □

In Abschnitt 3.6 haben wir die *Gerade* als Beispiel einer linearen Mannigfaltigkeit eingeführt. Eine Gerade hatte die Form $G = \{x + \alpha y \mid \alpha \in K\}$ mit $x, y \in V$, d.h. die Geraden sind gerade die affinen Teilräume der Dimension 1. Ist $\dim V = n$, so nennen wir einen affinen Teilraum von V, dessen Dimension $n - 1$ ist, eine *Hyperebene* in V. Die Hyperebenen in V sind also genau die affinen Teilräume der Gestalt

$$H = u + U \quad \text{mit} \quad \dim U = \dim V - 1. \tag{6.9}$$

Die Hyperebenen durch 0 sind dann genau die $n-1$-dimensionalen Teilräume von V. Ist (V, σ) nun ein euklidischer Vektorraum und H eine Hyperebene in V, so ist nach Satz 6.3.7 $\dim H^{\perp_\sigma} = 1$ und damit $H^{\perp_\sigma} = \langle a \rangle$ für einen Vektor $a \in V$. Dies läßt sich folgendermaßen verallgemeinern.

6.3.9 Lemma *Ist (V, σ) ein euklidischer Vektorraum, so gibt es zu jeder Hyperebene H in V einen normierten Vektor a und ein $\alpha \in \mathbb{R}$ mit*

$$H = H_{a,\alpha} := \{x \in V \mid \sigma(a, x) = \alpha\}. \tag{6.10}$$

Man nennt $H_{a,\alpha}$ die Hessesche Normalform *der Hyperebene H.*

Beweis: (vgl. Abbildung 6.2) Sei $H = v + U$ mit $\dim U = \dim V - 1$. Dann gibt es ein $a \in V$ mit $\langle a \rangle = U^{\perp_\sigma}$, und damit ist $U = \langle a \rangle^{\perp_\sigma}$. Ohne Einschränkung der Allgemeinheit dürfen wir $\|a\|_\sigma = 1$ annehmen. Dann gilt $x \in H \;\Leftrightarrow\; (\exists u \in U)[x = v + u] \;\Leftrightarrow\; (\exists u)[x = v + u \wedge \sigma(a, u) = 0] \;\Leftrightarrow\; (\exists u)[\sigma(a, x) = \sigma(a, v + u) = \sigma(a, v) + \sigma(a, u) = \sigma(a, v)] \;\Leftrightarrow\; \sigma(a, x) = \sigma(a, v) \;\Leftrightarrow\; x \in H_{a,\sigma(a,v)}$. □

Ist H eine Hyperebene in V und $v \in V$, so definieren wir

$$d_\sigma(v, H) := \inf\{d_\sigma(v, x) \mid x \in H\} \tag{6.11}$$

und nennen $d_\sigma(v, H)$ den *Abstand* des Vektors v von der Hyperebene H.

6.3.10 Satz *Ist (V, σ) ein euklidischer Raum und $H_{a,\alpha}$ die Hessesche Normalform der Hyperebene H, so gilt für $v \in V$ $d_\sigma(v, H) = |\sigma(a, v) - \alpha|$.*

Beweis: (vgl. Abbildung 6.3) Wir fällen das “Lot” von v auf H. Dies ergibt einen Vektor $v_0 := v + (\alpha - \sigma(a, v))a$. Wegen $\sigma(a, v_0) = \sigma(a, v) + (\alpha - \sigma(a, v))\sigma(a, a) = \alpha$ ist $v_0 \in H$. Weiter ist $v_0 - v \in \langle a \rangle$ und daher für $x \in H$

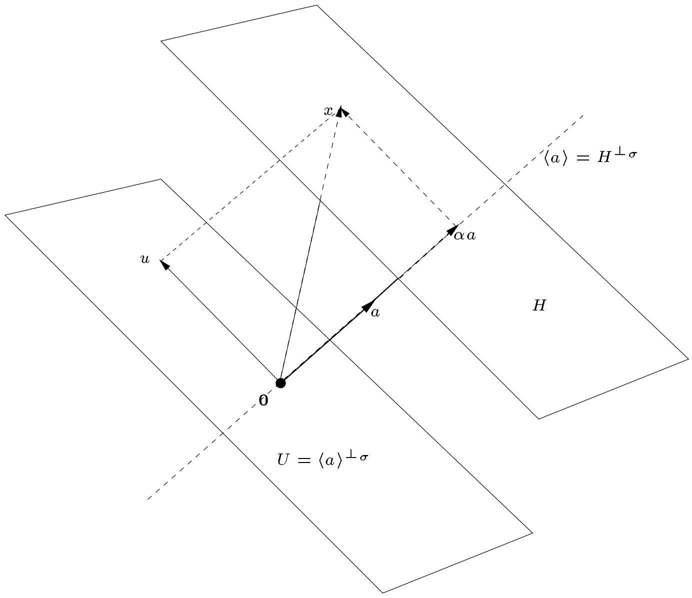

Abbildung 6.2: Hessesche Normalform im $\mathbb{R}^3$. Es gilt $\sigma(x, a) = \sigma(\alpha a + u, a) = \alpha\sigma(a, a) + \sigma(a, u) = \alpha$.

stets $\sigma(x - v_0, v_0 - v) = \sigma(x - v_0, \beta a) = \beta\sigma(x, a) - \beta\sigma(v_0, a) = 0$. Daher folgt für $x \in H$ schon $\|v - x\|_\sigma^2 = \|x - v_0 + v_0 - v\|_\sigma^2 = \|x - v_0\|_\sigma^2 + \|v_0 - v\|_\sigma^2 + 2\sigma(x - v_0, v_0 - v) = \|x - v_0\|_\sigma^2 + \|v_0 - v\|_\sigma^2 \geq \|v_0 - v\|_\sigma^2$. Damit ist $\|v_0 - v\|_\sigma = \inf\{\|v - x\|_\sigma \mid x \in H\} = d_\sigma(v, H)$, und wir erhalten $\|v_0 - v\|_\sigma = \|(\alpha - \sigma(a, v))a\|_\sigma = |\alpha - \sigma(a, v)|\|a\|_\sigma = |\alpha - \sigma(a, v)|$. $\square$

6.3.11 Definition Sei (V, σ) ein euklidischer Vektorraum und $U_1, \ldots, U_n$ Unterräume von V. Wir sagen, daß V die *orthogonale Summe* der Vektorräume $U_1, \ldots, U_n$ ist, wenn gelten

(OS.1) $V = U_1 + \cdots + U_n$,

(OS.2) $U_i \subseteq U_j{}^{\perp_\sigma}$ *für* $i \neq j$.

Wir notieren mit $V = U_1 \bigoplus\!\!\!\!\perp \cdots \bigoplus\!\!\!\!\perp U_n$, daß V die orthogonale Summe der Unterräume $U_1, \ldots, U_n$ ist.

6.3.12 Lemma *Aus* $V = U_1 \bigoplus\!\!\!\!\perp \cdots \bigoplus\!\!\!\!\perp U_n$ *folgt* $V = U_1 \oplus \cdots \oplus U_n$.

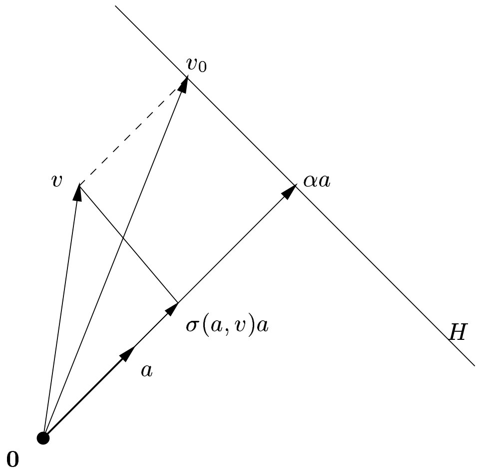

Abbildung 6.3: Zum Beweis von Satz 6.3.10

Beweis: Ist $V = U_1 \bigoplus \cdots \bigoplus U_n$ und $0 = v = u_1 + \cdots + u_n$ mit $u_i \in U_i$, so erhalten wir sofort $0 = \sigma(u_k, v) = \sigma(u_k, u_k)$ und damit $u_k = 0$ für $k = 1, \ldots, n$. □

Es ist eine einfache Beobachtung, daß für einen Unterraum U eines euklidischen Vektorraumes (V, σ) der Raum $(U, \sigma\restriction(U \times U))$ wieder ein euklidischer Raum ist. Dies benötigen wir zum Beweis des folgenden Satzes.

6.3.13 Satz *Ist (V, σ) ein endlich dimensionaler euklidischer Vektorraum, so gibt es eine Basis $B = (b_1, \ldots, b_n)$ von V mit $V = \langle b_1 \rangle \bigoplus \cdots \bigoplus \langle b_n \rangle$, d.h. V ist die orthogonale Summe von Geraden. Eine derartige Basis heißt eine* Orthogonalbasis.

Beweis: Wir führen Induktion nach $\dim V$. Für $\dim V = 0$ ist nichts zu beweisen. Ist $\dim V > 0$, so wählen wir $a \in V$ mit $a \neq 0$ und betrachten $U := \langle a \rangle^{\perp_\sigma}$. Da $\dim U < \dim V$ und $(U, \sigma\restriction(U \times U))$ wieder ein euklidischer Raum ist, erhalten wir nach Induktionsvoraussetzung eine Orthogonalbasis $(b_1, \ldots, b_{n-1})$ von $(U, \sigma\restriction(U \times U))$. Wegen $V = \langle a \rangle \bigoplus U$ ist dann $(a, b_1, \ldots, b_{n-1})$ eine Basis von V, und wegen $\langle b_i \rangle \subseteq U$ gilt $\sigma(a, b_i) = 0$ für $i = 1, \ldots, n-1$. Damit ist $\langle a \rangle \subseteq \langle b_i \rangle^{\perp_\sigma}$ und $\langle b_i \rangle \subseteq \langle a \rangle^{\perp_\sigma}$. Dann ist $(a, b_1, \ldots, b_{n-1})$ eine Orthogonalbasis von V. □

6.3.14 Definition Sei (V, σ) ein euklidischer Vektorraum. Wir nennen eine Basis $B = \{b_i \mid i \in I\}$ eine *Orthonormalbasis* von (V, σ), wenn $\sigma(b_i, b_j) = \delta_{ij}$ für alle $i, j \in I$ gilt.

6.3.15 Satz *Jeder euklidische Vektorraum besitzt eine Orthonormalbasis.*

Beweis: Nach Satz 6.3.13 besitzt V eine Orthogonalbasis B. Für $b_i, b_j \in B$ gilt daher $\sigma(b_i, b_j) = \|b_i\|_\sigma^2 \cdot \delta_{ij}$. Definieren wir $b' := \frac{1}{\|b\|_\sigma} b$, so ist $B' := \{b' \mid b \in B\}$ eine Orthonormalbasis. □

6.3.16 Das Schmidtsche Orthonormalisierungsverfahren Oft steht man vor dem Problem, eine Basis eines euklidischen Vektorraumes in eine Orthonormalbasis zu überführen. Dazu gibt es einen Algorithmus, den wir im folgenden beschreiben wollen.

Sei also eine Basis $(c_1, \ldots, c_n)$ gegeben. Wir definieren eine Folge $d_1, \ldots, d_n$, so daß gilt:

$$d_k \neq 0 \quad \text{und für} \quad b_k := \frac{1}{\|d_k\|_\sigma} d_k \tag{6.12}$$

$$(\forall i, j \leq k)[\sigma(b_i, b_j) = \delta_{ij}] \tag{6.13}$$

und

$$b_1, \ldots, b_k, c_{k+1}, \ldots, c_n \quad \text{bilden eine Basis.} \tag{6.14}$$

Wir setzen

$$d_1 := c_1 \quad \text{und} \quad d_{k+1} := c_{k+1} - \sum_{i=1}^{k} \sigma(c_{k+1}, b_i) b_i.$$

Nun zeigen wir (6.13) und (6.14) durch Induktion nach k. Für $k = 1$ erhalten wir $d_1 = c_1 \neq 0$ und $\sigma(b_1, b_1) = \frac{1}{\|c_1\|_\sigma^2} \sigma(c_1, c_1) = 1$. Im Induktionsschritt nehmen wir (6.13) für k an. Für $j \leq k$ erhalten wir $\sigma(d_{k+1}, b_j) = \sigma(c_{k+1} - \sum_{i=1}^{k} \sigma(c_{k+1}, b_i) b_i, b_j) = \sigma(c_{k+1}, b_j) - \sum_{i=1}^{k} \sigma(c_{k+1}, b_i) \cdot \sigma(b_i, b_j) = \sigma(c_{k+1}, b_j) - \sigma(c_{k+1}, b_j) = 0$, da nach Induktionsvoraussetzung $\sigma(b_i, b_j) = \delta_{ij}$ für $i = 1, \ldots, k$ gilt. Nach Induktionsvoraussetzung bilden $b_1, \ldots, b_k, c_{k+1}, \ldots, c_n$ eine Basis. Damit sind insbesondere $b_1, \ldots, b_k, c_{k+1}$ linear unabhängig und somit $d_{k+1} \neq 0$, und wir erhalten $\sigma(b_{k+1}, b_{k+1}) = \frac{1}{\|d_{k+1}\|_\sigma^2} \sigma(d_{k+1}, d_{k+1}) = 1$. Damit ist (6.13) bewiesen. Insbesondere sind $b_1, \ldots, b_{k+1}$ linear unabhängig, und da b_{k+1} eine Linearkombination der $b_1, \ldots, b_k$ und c_{k+1} ist, bilden $b_1, \ldots, b_{k+1}, c_{k+2}, \ldots, c_n$ eine Basis. Für $k = n$ folgt aus (6.13), daß $b_1, \ldots, b_n$ eine Orthonormalbasis bilden.

Als Korollar zu Satz 6.3.15 erhalten wir eine Charakterisierung positiv definiter Matrizen.

6.3.17 Korollar *Eine Matrix* $\mathfrak{S} \in \mathbb{R}^{n \times n}$ *ist genau dann positiv definit, wenn es eine invertierbare Matrix* $\mathfrak{A} \in \mathrm{GL}(n, \mathbb{R})$ *gibt mit* $\mathfrak{S} = \mathfrak{A}\mathfrak{A}^t$.

Beweis: $\Rightarrow$: Ist $\mathfrak{S}$ positiv definit, so definiert $\sigma(\mathfrak{x},\mathfrak{y}) := \mathfrak{x}\mathfrak{S}\mathfrak{y}^t$ eine positiv definite Bilinearform auf $\mathbb{R}^n$. Nach Satz 6.3.15 gibt es eine Orthonormalbasis B von (V,σ). Dann gilt $\mathfrak{D}_B(\sigma) = (\sigma(b_i,b_j))_{n,n} = \mathfrak{E}$. Für $\mathfrak{A} := \mathfrak{T}_{\mathbb{E}^n,B}$ folgt aus Satz 6.1.17 $\mathfrak{S} = \mathfrak{D}_{\mathbb{E}^n}(\sigma) = \mathfrak{A}\mathfrak{A}^t$.
$\Leftarrow$: Ist $\mathfrak{S} = \mathfrak{A}\mathfrak{A}^t$ mit $\mathfrak{A} \in \mathrm{GL}(n,\mathbb{R})$, so folgt $\sigma_{\mathfrak{S}}(\mathfrak{x},\mathfrak{x}) = \mathfrak{x}\mathfrak{S}\mathfrak{x}^t = \mathfrak{x}\mathfrak{A}\mathfrak{A}^t\mathfrak{x}^t = (\mathfrak{x}\mathfrak{A})(\mathfrak{x}\mathfrak{A})^t = \langle \mathfrak{x}\mathfrak{A},\mathfrak{x}\mathfrak{A}\rangle$. Wie wir in Beispiel 6.2.8 gesehen haben, ist das kanonische Skalarprodukt auf $\mathbb{R}^n$ positiv definit. Damit ist $\sigma_{\mathfrak{S}}(\mathfrak{x},\mathfrak{x}) = 0$ genau dann, wenn $\mathfrak{x}\mathfrak{A} = 0$ ist. Da $\mathfrak{A} \in \mathrm{GL}(n,\mathbb{R})$ ist, ist dies genau dann der Fall, wenn $\mathfrak{x} = 0$. Also ist $\sigma_{\mathfrak{S}}$ positiv definit. □

6.4 Isometrien euklidischer Räume

6.4.1 Definition Es seien (V,σ) und (W,τ) zwei euklidische Räume. Eine Abbildung $f\colon V \longrightarrow W$ heißt eine *Isometrie* von (V,σ) in (W,τ), wenn

$$(\forall x,y \in V)[\sigma(x,y)) = \tau(f(x),f(y))]$$

gilt.

6.4.2 Satz *Ist $f\colon V \longrightarrow W$ eine Isometrie, so ist f ein Monomorphismus der Vektorräume.*

Beweis: Da f eine Isometrie ist, folgt $\|f(x)\|_\tau = \|x\|_\sigma$. Wir erhalten $\|f(x+y) - f(x) - f(y)\|_\tau^2 = \|f(x+y)\|_\tau^2 + \|f(x)\|_\tau^2 + \|f(y)\|_\tau^2 - 2\tau(f(x+y),f(x)) - 2\tau(f(x+y),f(y)) + 2\tau(f(x),f(y)) = \|x+y\|_\sigma^2 + \|x\|_\sigma^2 + \|y\|_\sigma^2 - 2\sigma(x+y,x) - 2\sigma(x+y,y) + 2\sigma(x,y) = \|(x+y) - x - y\|_\sigma^2 = 0$. Da τ positiv definit ist, folgt daraus $f(x+y) - f(x) - f(y) = 0$, d.h. $f(x+y) = f(x) + f(y)$. Analog erhalten wir $\|f(\alpha x) - \alpha f(x)\|_\tau^2 = \|f(\alpha x)\|_\tau^2 + \alpha^2\|f(x)\|_\tau^2 - 2\alpha\tau(f(\alpha x),f(x)) = \|\alpha x\|_\sigma^2 + \alpha^2\|x\|_\sigma^2 - 2\alpha\sigma(\alpha x,x) = \|\alpha x - \alpha x\|_\sigma^2 = 0$. Damit folgt $f(\alpha x) = \alpha f(x)$, und f ist linear. Ist $f(x) = 0$, so ist $\|x\|_\sigma = \|f(x)\|_\tau = 0$ und damit $x = 0$. Damit ist f injektiv, und wir haben einen Monomorphismus der Vektorräume. □

6.4.3 Definition Sei (V,σ) ein euklidischer Raum. Wir setzen

$$O(V,\sigma) := \{f \mid\ f\colon V \longrightarrow V \ \text{ ist Isometrie von } (V,\sigma) \text{ in } (V,\sigma)\}.$$

6.4.4 Lemma *Sei (V,σ) ein euklidischer Raum. Die Menge $O(V,\sigma)$ ist dann eine Untergruppe der* $\mathrm{GL}(V)$. *Man nennt sie die* allgemeine orthogonale Gruppe *von* (V,σ).

Beweis: Zunächst ist $id_V \in O(V,\sigma)$ und damit $O(V,\sigma) \neq \emptyset$. Es folgt aus Satz 6.4.2, daß $O(V,\sigma) \subseteq \mathrm{GL}(V)$ ist. Für $f,g \in O(V,\sigma)$ ist $\sigma(f(g(x)), f(g(y))) = \sigma(g(x), g(y)) = \sigma(x,y)$. Damit ist $f \circ g \in O(V,\sigma)$. Weiter folgt $\sigma(f^{-1}(x), f^{-1}(y)) = \sigma(f(f^{-1}(x)), f(f^{-1}(y))) = \sigma(x,y)$ und damit $f^{-1} \in O(V,\sigma)$. □

6.4.5 Satz *Sei (V,σ) ein euklidischer Raum. Ein Endomorphismus $F \in \mathrm{End}(V)$ ist genau dann eine Isometrie von (V,σ) in sich, wenn $F^{-1} = F^{\sigma}$ gilt, d.h. wir haben*

$$F \in \mathrm{End}(V) \quad \Rightarrow \quad [F \in O(V,\sigma) \quad \Leftrightarrow \quad F^{-1} = F^{\sigma}].$$

Beweis: Ist F eine Isometrie, so gilt $\sigma(x,y) = \sigma(Fx, Fy) = \sigma(F^{\sigma}Fx, y)$. Daraus folgt, da σ nicht ausgeartet ist, $F^{\sigma}F = id$. Also ist $F^{-1} = F^{\sigma}$. Ist umgekehrt $F^{-1} = F^{\sigma}$, so folgt $\sigma(Fx, Fy) = \sigma(F^{\sigma}Fx, y) = \sigma(x,y)$, und F ist eine Isometrie. □

6.4.6 Korollar *Ist $F \in O(V,\sigma)$ eine Isometrie des euklidischen Raumes (V,σ), so gilt* $\det F \in \{1, -1\}$.

Beweis: Nach Bemerkung 6.1.8 ist $F^{\sigma} = \sigma^{*\,-1} \circ F^* \circ \sigma^*$, und nach Lemma 4.3.10 und Korollar 4.3.14 erhalten wir daraus $\det(F^{\sigma}) = \det(F^*) = \det F$. Zusammen mit Satz 6.4.5 ergibt dies $(\det F)^2 = \det(F^{\sigma}) \det F = 1$, woraus die Behauptung sofort folgt. □

6.4.7 Spiegelungen Sei V ein endlich dimensionaler euklidischer Raum und H eine Hyperebene in V durch den Nullvektor. Dann gibt es einen normierten Vektor a mit $H^{\perp_\sigma} = \langle a \rangle$ und zu jedem Vektor $v \in V$ ein eindeutig bestimmtes $\alpha \in \mathbb{R}$ und $u \in H$ mit $v = \alpha a + u$. Die Abbildung

$$v = \alpha a + u \longmapsto -\alpha a + u$$

heißt *Spiegelung* an der Hyperebene H.

6.4.8 Bemerkung *Ist (V,σ) ein euklidischer Raum und $a \in V$, so hat die Spiegelung an $\langle a \rangle^{\perp_\sigma}$ die Form*

$$v \longmapsto v - 2\frac{\sigma(a,v)}{\sigma(a,a)}a =: S_a(v).$$

Beweis: Sei $a_N := \frac{1}{\|a\|_\sigma}a$ die Normierung von a. Für $v \in V$ gilt dann $v = \sigma(a_N, v)a_N + u$ für ein (eindeutig bestimmtes) $u \in \langle a \rangle^{\perp_\sigma} =: H$. Damit gilt für das

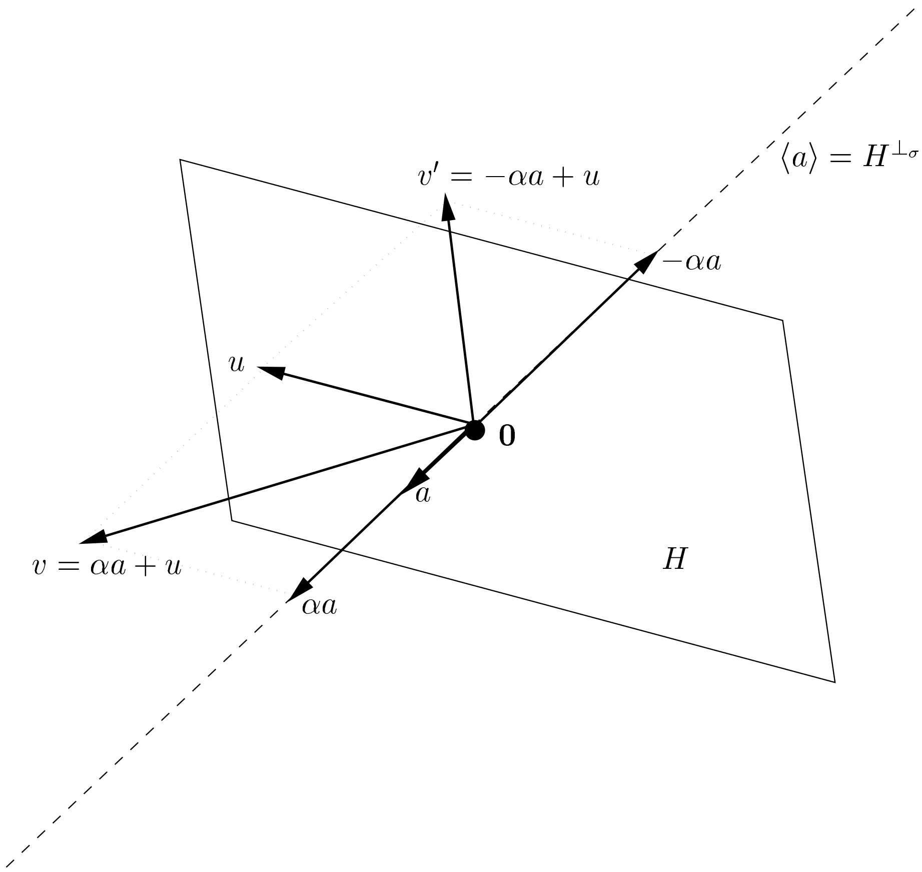

Abbildung 6.4: Spiegelung an einer Hyperebene

Spiegelbild $v' = -\sigma(a_N, v)a_N + u = v - 2\sigma(a_N, v)a_N$. Wegen $\sigma(a_N, v)a_N = \frac{\sigma(a,v)}{\sigma(a,a)}a$ ist dies die Behauptung. □

6.4.9 Satz *Jede Spiegelung ist eine Isometrie.*

Beweis: Seien $x, y \in V$ und S_a die Spiegelung an der Hyperebene $\langle a\rangle^{\perp_\sigma}$ für einen normierten Vektor a. Dann ist $x = \alpha a + u_x$ und $y = \beta a + u_y$ mit $u_x, u_y \in \langle a\rangle^{\perp_\sigma}$, und wir erhalten $\sigma(x, y) = \alpha\beta + \sigma(u_x, u_y) = (-\alpha)(-\beta) + \sigma(u_x, u_y) = \sigma(-\alpha a + u_x, -\beta a + u_y) = \sigma(S_a(x), S_a(y))$. □

6.4.10 Lemma *Ist S eine Spiegelung, so gilt* $\det(S) = -1$.

Beweis: Sei S die Spiegelung an der Hyperebene H durch 0. Dann ist $H^{\perp_\sigma} = \langle a \rangle$ für einen normierten Vektor $a \in V$, und wir erhalten eine Basis $B := \{a, b_1, \ldots, b_{n-1}\}$ von V, so daß $\{b_1, \ldots, b_{n-1}\}$ eine Basis von H ist. Dann gilt nach Lemma 4.3.5 $\det S = \Delta_B(S(a), S(b_1), \ldots, S(b_{n-1})) = \Delta_B(-a, b_1, \ldots, b_{n-1}) = -\Delta_B(a, b_1, \ldots, b_{n-1}) = -1$. □

6.4.11 Bemerkung Man bezeichnet Isometrien eines euklidischen Raumes auch als *Drehungen* oder *Rotationen*. Dabei heißt eine Rotation R *eigentlich*, wenn $\det R = 1$ ist. Anderenfalls heißt sie *uneigentlich*. Ist λ ein Eigenwert einer Rotation R, so muß für einen Eigenvektor a schon $\|a\|_\sigma = \|Ra\|_\sigma = \|\lambda a\|_\sigma = |\lambda| \|a\|_\sigma$ und damit $\lambda \in \{1, -1\}$ gelten.

6.4.12 Lemma *Die eigentlichen Drehungen eines euklidischen Vektorraumes* (V, σ) *bilden eine Untergruppe der allgemeinen orthogonalen Gruppe* $O(V, \sigma)$. *Man nennt sie die* spezielle orthogonale Gruppe *und bezeichnet sie mit* $SO(V, \sigma)$.

Beweis: Sei $SO(V, \sigma) := \{R \in O(V, \sigma) \mid \det R = 1\}$. Dann ist für $F, G \in SO(V, \sigma)$ auch $\det(FG^{-1}) = \det F \det(G^{-1}) = 1$ und damit $FG^{-1} \in SO(V, \sigma)$. □

6.4.13 Satz *Sei* (V, σ) *ein* n*-dimensionaler euklidischer Vektorraum. Dann ist jede von der Identität verschiedene Isometrie von* V *in sich ein Produkt von höchstens* n*- vielen Spiegelungen.*

Beweis: Als erstes haben wir uns zu überlegen, daß sich jede Spiegelung eines Teilraumes von V zu einer Spiegelung auf V fortsetzen läßt. Sei dazu U ein Teilraum von V und H_U eine Hyperebene durch 0 in U. Dann wird das Komplement ${H_U}^{\perp_\sigma}$ von H_U in (U, σ) durch einen Vektor $a \in U \subseteq V$ aufgespannt. Eine Spiegelung an H in U hat dann die Form $\alpha a + h \longmapsto -\alpha a + h$ mit $h \in H$. Nun betrachten wir $\langle a \rangle^{\perp_\sigma}$ in (V, σ). Dies ist eine Hyperebene durch 0 in (V, σ). Die Abbildung $\alpha a + h \longmapsto -\alpha a + h$ für $h \in \langle a \rangle^{\perp_\sigma}$ definiert eine Spiegelung an $\langle a \rangle^{\perp_\sigma}$ in V, die die Spiegelung in U fortsetzt. Wir bemerken noch, daß die Fortsetzung der Spiegelung auf $V \setminus U$ die Identität ist.
Wir beweisen den Satz nun durch Induktion nach n. Ist $\dim V = 1$, so hat jeder Endomorphismus von V die Gestalt $\alpha \cdot id$ und wegen Korollar 6.4.6 muß für eine Isometrie $|\alpha| = 1$ sein. Damit sind id und $-id$ die einzigen Isometrien auf V, und wir haben nichts zu beweisen.
Sei nun $n > 1$ und F eine Isometrie von V in sich. Gibt es einen Vektor $a \in V$ mit $a \neq 0$ und $Fa = a$, so betrachten wir die Hyperebene $H := \langle a \rangle^{\perp_\sigma}$. Für $h \in H$ gilt $\sigma(a, Fh) = \sigma(Fa, Fh) = \sigma(a, h) = 0$, d.h. $Fh \in H$. Nach Induktionsvoraussetzung gibt es Spiegelungen $S_1, \ldots, S_r$ auf (H, σ) mit $r < n$ und $F{\restriction}H =$

$S_1 \cdots S_r$. Nach unserer Vorbemerkung läßt sich aber jede Spiegelung S_i auf (H, σ) zu einer Spiegelung S_i' auf (V, σ) so fortsetzen, daß $S_i'(\alpha a) = \alpha a$ ist. Damit ist $F(\alpha a + h) = \alpha a + (F \restriction H)(h) = \alpha a + (S_1 \cdots S_r)(h) = (S_1' \cdots S_r')(\alpha a + h)$, und wir erhalten $F = S_1' \cdots S_r'$. Wir haben also

$$(\exists a \in V) Fa = a \quad \Rightarrow \quad (\exists r < n)(\exists S_1, \ldots, S_r)[S_1, \ldots, S_r \text{ sind Spiegelungen} \\ \wedge\ F = S_1 \cdots S_r] \tag{i}$$

gezeigt.
Erfüllt die Isometrie die Voraussetzung des Satzes, so gibt es wegen $F \neq id$ ein $b \in V$ mit $Fb \neq b$. Sei nun $a := Fb - b$ und $H := \langle a \rangle^{\perp_\sigma}$. Die Spiegelung an H bezeichnen wir mit S_H. Dann ist $S_H(a) = -a$ und damit

$$S_H(Fb) - S_H(b) = -Fb + b. \tag{ii}$$

Es ist

$$\sigma(Fb + b, Fb - b) = \sigma(Fb, Fb) - \sigma(b, b) = 0, \tag{iii}$$

womit $Fb + b$ und a orthogonal sind und daher $Fb + b \in H$ ist. Damit gilt

$$S_H(Fb) + S_H(b) = Fb + b, \tag{iv}$$

und aus (ii) und (iv) folgt $(S_H F)(b) = b$. Aus (i) folgt nun, daß $S_H F$ ein Produkt von höchstens $n - 1$ vielen Spiegelungen ist. Wegen $S_H \cdot S_H = id$ folgt

$$F = S_H \cdot S_1 \cdots S_r \tag{v}$$

für $r < n$, und wir sehen, daß F das Produkt von höchstens n Spiegelungen ist.

□

6.4.14 Bemerkung Betrachten wir eine Rotation R bezüglich des kanonischen Skalarprodukts auf dem $\mathbb{R}^n$, so erhalten wir nach Satz 6.1.16

$$\mathfrak{D}_{\mathbb{E}^n, \mathbb{E}^n}(R^\sigma) = \mathfrak{D}_{\mathbb{E}^n, \mathbb{E}^n}(R)^t,$$

denn die darstellende Matrix des kanonischen Skalarproduktes ist die Einheitsmatrix. Damit ist nach Satz 6.4.5 $\overline{\mathfrak{A}}$ genau dann eine Rotation bezüglich des kanonischen Skalarproduktes im $\mathbb{R}^n$, wenn $\mathfrak{A}^{-1} = \mathfrak{A}^t$ ist. Daher definiert man

$$O(n) := \{\mathfrak{A} \in \mathbb{R}^{n \times n} \mid \mathfrak{A}^{-1} = \mathfrak{A}^t\} \tag{6.15}$$

und nennt $O(n)$ die *allgemeine orthogonale Gruppe*. Die Matrizen der orthogonalen Gruppe heißen *orthogonale Matrizen*. Die Menge

$$SO(n) = \{\mathfrak{A} \in O(n) \mid \det \mathfrak{A} = 1\} \tag{6.16}$$

ist eine Untergruppe der $O(n)$. Sie heißt die *spezielle orthogonale Gruppe*. Man prüft leicht nach, daß für einen n-dimensionalen euklidischen Raum (V, σ) und eine Orthonormalbasis B von V

$$O(n) = \{\mathfrak{D}_{B,B}(R) \mid R \in O(V, \sigma)\} \tag{6.17}$$

und

$$SO(n) = \{\mathfrak{D}_{B,B}(R) \mid R \in SO(V, \sigma)\} \tag{6.18}$$

gelten.

6.4.15 Satz *Eine Matrix $\mathfrak{A} \in \mathbb{R}^{n\times n}$ ist genau dann orthogonal, wenn sowohl ihre Zeilen als auch ihre Spalten eine Orthonormalbasis des $\mathbb{R}^n$ mit dem kanonischen Skalarprodukt bilden.*

Beweis: Sei $\mathfrak{A} = \begin{pmatrix} \mathfrak{a}_1 \\ \vdots \\ \mathfrak{a}_n \end{pmatrix}$ mit $\mathfrak{a}_i \in \mathbb{R}^n$. Dann gilt $\mathfrak{A}^t = \mathfrak{A}^{-1}$ genau dann, wenn $\mathfrak{a}_i\mathfrak{a}_j^t = \delta_{ij}$ für $i, j = 1, \ldots, n$ gilt. Dies ist aber genau dann der Fall, wenn $\langle \mathfrak{a}_i, \mathfrak{a}_j \rangle = \delta_{ij}$ für $i, j = 1, \ldots, n$ gilt, d.h. wenn $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ eine Orthonormalbasis bilden. □

6.5 Sesquilinearformen

Im folgenden werden wir gezwungen sein, auch Vektorräume über dem Körper $\mathbb{C}$ der komplexen Zahlen zu betrachten. Obwohl wir die komplexen Zahlen hier voraussetzen wollen, sind einige Vorbemerkungen angebracht.

6.5.1 Komplexe Zahlen Das Polynom x^2+1 hat in $\mathbb{R}$ keine Wurzel, da Quadrate reeller Zahlen niemals negativ sind. Man “adjungiert” daher eine Wurzel dieses Polynoms, die üblicherweise mit i bezeichnet wird und betrachtet den Einsetzungshomomorphismus $\mathbb{R}[x] \longrightarrow \mathbb{R}[i]$. Dann ist $\mathbb{R}[i]$ eine kommutative $\mathbb{R}$-Algebra der Dimension 2, denn das Minimalpolynom von i ist $x^2 + 1$. Insbesondere ist $\mathbb{R}[i]$ ein kommutativer Ring, in dem jedes Element die Gestalt $\alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$ hat. Nun beobachtet man weiter, daß $(\alpha + i\beta)\cdot(\alpha - i\beta) = \alpha^2 + \beta^2$ und damit $(\alpha + i\beta)\cdot \frac{1}{\alpha^2+\beta^2}(\alpha - i\beta) = 1$ gilt. D.h. $\mathbb{R}[i]$ besteht nur aus Einheiten und ist daher ein Körper, den man mit $\mathbb{C}$ bezeichnet. Die Elemente von $\mathbb{C}$ heißen *komplexe Zahlen*. Jede komplexe Zahl hat die Gestalt $z = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$, wobei α und β eindeutig bestimmt sind. Man nennt $\Re(z) := \alpha$ den *Realteil* und $\Im(z) := \beta$ den *Imaginärteil* der komplexen Zahl $z = \alpha + i\beta$. Die Abbildung

$\overline{}: \mathbb{C} \longrightarrow \mathbb{C}$, die durch $\overline{\alpha + i\beta} := \alpha - i\beta$ definiert ist, heißt *Konjugation*, die Zahl $\alpha - i\beta$ die *konjungiert Komplexe* von $\alpha + i\beta$. Man prüft sofort nach, daß die Konjugation ein involutorischer Automorphismus von $\mathbb{C}$ ist.
Wegen $(\alpha + i\beta)\overline{(\alpha + i\beta)} = \alpha^2 + \beta^2$ erhalten wir eine Abbildung

$$\begin{aligned} |\ |: \mathbb{C} &\longrightarrow \mathbb{R} \\ z &\longmapsto \sqrt{z\overline{z}} \end{aligned}$$

und rechnen leicht nach, daß dies eine Normfunktion auf $\mathbb{C}$ ist, für die $|z_1 z_2| = |z_1||z_2|$ gilt.

Eine wichtige Eigenschaft der komplexen Zahlen wird im folgenden Satz ausgedrückt, den wir ohne Beweis zitieren.

Fundamentalsatz der Algebra *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen, d.h. jedes Polynom $f \in \mathbb{C}[x]$ zerfällt über $\mathbb{C}$ in Linearfaktoren.*

6.5.2 Bemerkung Man rechnet sofort nach, daß die Abbildung

$$\begin{aligned} M: \mathbb{C} &\longrightarrow \mathbb{R}^{2\times 2} \\ \alpha + i\beta &\longmapsto \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \end{aligned}$$

ein injektiver Homomorphismus der Ringe ist. Daher ist die Menge

$$M[\mathbb{C}] = \left\{ \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \middle|\ \alpha, \beta \in \mathbb{R} \right\} \subseteq \mathbb{R}^{2\times 2}$$

als Ring isomorph zu $\mathbb{C}$ und somit ein Körper. Wir erhalten $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}^t \cdot \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = (\alpha^2 + \beta^2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, und es folgt aus Satz 6.4.5, daß die Elemente $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$ mit $\alpha^2 + \beta^2 = 1$ Drehungen des $\mathbb{R}^2$ sind. Wegen

$$(\xi_1, \xi_2) \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} = (\alpha\xi_1 - \beta\xi_2, \beta\xi_1 + \alpha\xi_2) \tag{6.19}$$

und

$$(\xi_1, \xi_2) \cdot \begin{pmatrix} \alpha\xi_1 - \beta\xi_2 \\ \beta\xi_1 + \alpha\xi_2 \end{pmatrix} = \alpha(\xi_1^2 + \xi_2^2)$$

ist der Drehwinkel der Drehung $\Theta = \arccos \alpha$. Wegen $\alpha^2 + \beta^2 = 1$ ist dann $\beta = \sin \Theta$, und die Drehungen haben die Form $\begin{pmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{pmatrix}$. Übersetzen wir dies mit M^{-1} zurück in die komplexen Zahlen und beachten wir, daß $|z| =$

$\sqrt{\alpha^2+\beta^2}$ für $z = \alpha + i\beta \in \mathbb{C}$ gilt, so können wir jede komplexe Zahl $z \in \mathbb{C}$ in der Form

$$z = |z| \cdot (\cos\Theta + i \sin\Theta) \tag{6.20}$$

oder, unter Verwendung der Exponentialfunktion, als

$$z = |z| e^{i\Theta} \tag{6.21}$$

darstellen. Dabei heißt $|z|$ der *Betrag* und Θ das *Argument* der komplexen Zahl z. Betrachten wir die Bijektion $V : \mathbb{C} \longrightarrow \mathbb{R}^2$ mit $V(\alpha + i\beta) := (\alpha, \beta)$, d.h. fassen wir $\mathbb{C}$ als den $\mathbb{R}^2$ auf, so ersehen wir aus (6.19), daß $V(z_1) \cdot M(z_2) = V(z_1 z_2)$ ist. Übersetzen wir dies mit V^{-1} und M^{-1} wieder in die komplexen Zahlen zurück, so sehen wir, daß sich die Multiplikation mit komplexen Zahlen vom Betrag 1 als Drehung auffassen läßt, wobei der Drehwinkel durch das Argument der Zahl gegeben ist. Multiplikationen mit i, das durch $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ dargestellt wird, entsprechen daher Drehungen um $\frac{\pi}{2}$ und Multiplikationen mit -1 Drehungen um π.

6.5.3 Definition Seien V und W Vektorräume über dem Körper $\mathbb{C}$ der komplexen Zahlen. Eine Abbildung $F : V \longrightarrow W$ heißt *semilinear* (d.h. halblinear) , wenn $F(u + v) = Fu + Fv$ und $F(\lambda v) = \overline{\lambda} F(v)$ für alle $u, v \in V$ und $\lambda \in \mathbb{C}$ gilt.

Eine Abbildung $\sigma : V \times W \longrightarrow \mathbb{C}$ heißt eine *Sesquilinearform* (eine eineinhalbfach lineare Form), wenn $\lambda x \,.\, \sigma(x, y)$ linear und $\lambda y \,.\, \sigma(x, y)$ semilinear ist, d.h. wenn $\sigma(\alpha u + \beta v, w) = \alpha\sigma(u, w) + \beta\sigma(v, w)$ und $\sigma(u, \alpha v + \beta w) = \overline{\alpha}\sigma(u, v) + \overline{\beta}\sigma(u, w)$ gelten.

Eine Abbildung $\sigma : V \times V \longrightarrow \mathbb{C}$ heißt *hermitesch*, wenn $(\forall u, v \in V)[\sigma(u, v) = \overline{\sigma(v, u)}]$ gilt.
Man beachte, daß jede hermitesche Form σ, für die $\lambda x \,.\, \sigma(x, y)$ linear ist, bereits sesquilinear ist.

Wir übertragen nun die für Bilinearformen in Abschnitt 6.1 und 6.2 eingeführten Begriffe und Sätze auf hermitesche Sesquilinearformen. Die Beweise übertragen sich in der Regel mühelos.
Zunächst definieren wir den *Sesquilinearkern* einer hermiteschen Sesquilinearform in Analogie zum Bilinearkern als

$$\text{Bkern } \sigma = \{v \in V \mid (\forall x \in V)[\sigma(x, v) = 0]\} = \{v \in V \mid (\forall y \in V)[\sigma(v, y) = 0]\}$$

und nennen eine hermitesche Sesquilinearform σ *nicht ausgeartet*, wenn Bkern $\sigma = \{0\}$ ist.
Ist σ eine hermitesche Sesquilinearform, so definiert

$$\sigma_a := \lambda x \,.\, \sigma(x, a)$$

eine Linearform, und damit erhalten wir eine semilineare Abbildung

$$\sigma^*: V \longrightarrow V^* \quad \text{mit} \quad \sigma^*(v) := \sigma_v = \lambda x \,.\, \sigma(x, v)$$

und

$$\text{Bkern } \sigma = \text{Kern } \sigma^*.$$

Für nicht ausgeartetes σ ist σ^* daher eine injektive semilineare Abbildung von V in V^* und demnach eine semilineare Bijektion für endlich dimensionale V.
Ist V ein endlich dimensionaler Vektorraum, $F \in \text{End}(V)$ ein Endomorphismus und $a \in V$, so ist $\lambda x \,.\, \sigma(Fx, a) \in V^*$, und es gibt daher ein eindeutig bestimmtes $a' \in V$ mit $\lambda x \,.\, \sigma(Fx, a) = \sigma^*(a') = \lambda x \,.\, \sigma(x, a')$. Die Abbildung $a \longmapsto a'$ ist wieder linear und heißt die zu σ adjungierte Abbildung F^σ. Es gilt also

$$\sigma(Fu, v) = \sigma(u, F^\sigma v) \tag{6.22}$$

und

$$\sigma(u, Fv) = \overline{\sigma(Fv, u)} = \overline{\sigma(v, F^\sigma u)} = \sigma(F^\sigma u, v). \tag{6.23}$$

Weiter erhalten wir $(\sigma^* \circ F^\sigma)(a) = \lambda x \,.\, \sigma(x, F^\sigma a) = \lambda x \,.\, \sigma(Fx, a) = \lambda x \,.\, \sigma(x, a) \circ F = \sigma_a \circ F = F^*(\sigma_a) = F^*(\sigma^*(a)) = (F^* \circ \sigma^*)(a)$ und damit

$$\sigma^* \circ F^\sigma = F^* \circ \sigma^*. \tag{6.24}$$

Wie im Falle von Bilinearformen prüft man nach, daß die Abbildung $F \longmapsto F^\sigma$ ein involutorischer Antiisomorphismus des Endomorphismenringes ist. Allerdings ist sie kein Homomorphismus der Algebren, sondern wegen $\sigma^*(\alpha F)^\sigma = (\alpha F)^* \sigma^* = \alpha(F^* \sigma^*) = \alpha(\sigma^* F^\sigma) = \sigma^*(\overline{\alpha} F^\sigma)$ nur semilinear.
Ein Unterschied zur Situation symmetrischer Bilinearformen ergibt sich auch bei der Matrixdarstellung von hermiteschen Sesquilinearformen. Wir definieren wie im Falle der Bilinearformen für eine Sesquilinearform σ auf einem endlich dimensionalen $\mathbb{C}$-Vektorraum V mit Basis $B = \{b_1, \ldots, b_n\}$

$$\mathfrak{D}_B(\sigma) := (\sigma(b_i, b_j))_{n,n}. \tag{6.25}$$

Ist $\mathfrak{A} = (\alpha_{ij}) \in \mathbb{C}^{n \times m}$, so sei $\overline{\mathfrak{A}} := (\overline{\alpha_{ij}})$ die konjungiert komplexe Matrix zu $\mathfrak{A}$. (*Um $\overline{\mathfrak{A}}$ typographisch von der Abbildung $\overline{\mathfrak{A}}$ zu unterscheiden, ist die Überstreichung fetter gesetzt*). Ist $B = (b_1, \ldots, b_n)$ eine Basis von V und $\mathfrak{D}_{B,B^*}(\sigma^*) = (\alpha_{ij})$, so erhalten wir $\sigma^*(b_i)(b_k) = (\sum_{j=1}^n \alpha_{ij} b_j^*)(b_k) = \alpha_{ik}$ und damit $\mathfrak{D}_{B,B^*}(\sigma^*) = (\sigma^*(b_i)(b_j))_{n,n}$. Damit erhalten wir für eine hermitesche Sesquilinearform σ

$$\mathfrak{D}_{B,B^*}(\sigma^*) = (\sigma^*(b_i)(b_j))_{n,n} = (\sigma(b_j, b_i))_{n,n} = \left(\overline{\sigma(b_i, b_j)}\right)_{n,n} = \overline{\mathfrak{D}_B(\sigma)}. \quad (6.26)$$

Um das Diagramm 6.1 zu modifizieren, führen wir die *konjungierte Koordinatendarstellung* bezüglich einer Basis $B = \{b_1, \ldots, b_n\}$ durch

$$\overline{h_B}(\sum_{i=1}^{n} \alpha_i b_i) := (\overline{\alpha}_1, \ldots, \overline{\alpha}_n)$$

ein. Dann modifiziert das Diagramm 6.1 zum Diagramm in Abbildung 6.5. Um

$$\begin{tikzcd}
\mathbb{C}^n \arrow[rrr, "\overline{\overline{\mathfrak{D}_{B,B}(A^\sigma)}}"] \arrow[ddd, "\overline{\mathfrak{D}_{B,B^*}(\sigma^*)}"'] & & & \mathbb{C}^n \arrow[ddd, "\overline{\mathfrak{D}_{B,B^*}(\sigma^*)}"] \\
& V \arrow[ul, "\overline{h_B}"'] \arrow[r, "A^\sigma"] \arrow[d, "\sigma^*"] & V \arrow[ur, "\overline{h_B}"] \arrow[d, "\sigma^*"'] & \\
& V^* \arrow[dl, "h_{B^*}"] \arrow[r, "A^*"'] & V^* \arrow[dr, "h_{B^*}"'] & \\
\mathbb{C}^n \arrow[rrr, "\overline{\mathfrak{D}_{B^*,B^*}(A^*)}"'] & & & \mathbb{C}^n
\end{tikzcd}$$

Abbildung 6.5: Zusammenhang zwischen adjungierter und dualer Abbildung bezüglich einer hermiteschen Sesquilinearform

die Kommutativität des Diagramms zu zeigen, berechnen wir $\sigma^*(\sum_{i=1}^n \beta_i b_i) = \sum_{i=1}^n \overline{\beta_i}\sigma^*(b_i) = \sum_{i=1}^n \overline{\beta_i} \sum_{j=1}^n \alpha_{ij} b_j^* = \sum_{j=1}^n (\sum_{i=1}^n \overline{\beta_i}\alpha_{ij}) b_j^*$ für $\mathfrak{D}_{B,B^*}(\sigma^*) = (\alpha_{ij})$, woraus sich $\overline{\mathfrak{D}_{B,B^*}(\sigma^*)} \circ \overline{h_B} = h_{B^*} \circ \sigma^*$ ergibt. Damit sind das linke und rechte Trapez kommutativ. Weiter gilt $\overline{h_B}(A^\sigma(\sum_{i=1}^n \beta_i b_i)) = (\sum_{i=1}^n \overline{\beta_i \alpha_{i1}}, \ldots, \sum_{i=1}^n \overline{\beta_i \alpha_{in}}) = (\sum_{i=1}^n \overline{\beta_i}\,\overline{\alpha_{i1}}, \ldots, \sum_{i=1}^n \overline{\beta_i}\,\overline{\alpha_{in}}) = \overline{\overline{\mathfrak{D}_{B,B}(A^\sigma)}}(\overline{h_B}(\sum_{i=1}^n \beta_i b_i))$, womit auch die Kommutativität des oberen Trapezes gezeigt ist.
Aus dem Diagramm 6.5 ergibt sich die Transformationsformel

$$\overline{\overline{\mathfrak{D}_{B,B}(A^\sigma)}} = \overline{\mathfrak{D}_B(\sigma)} \cdot \mathfrak{D}_{B,B}(A)^t \cdot \overline{\mathfrak{D}_B(\sigma)}^{-1},$$

woraus sich durch Übergang zum konjugiert komplexen

$$\mathfrak{D}_{B,B}(A^\sigma) = \mathfrak{D}_B(\sigma) \cdot \overline{\mathfrak{D}_{B,B}(A)}^t \cdot \mathfrak{D}_B(\sigma)^{-1} \quad (6.27)$$

ergibt.
Sind B und C zwei Basen eines endlich dimensionalen komplexen Raumes, so kommutiert das Diagramm in Abbildung 6.6. Die beiden Trapeze haben wir gerade nachgerechnet, die Kommutativität des rechten Dreiecks wurde schon früher gezeigt (Satz 3.3.13), und die Kommutativität des linken Dreiecks rechnet man mühelos nach.

$$\begin{tikzcd}
\mathbb{C}^m \arrow[rrr, "\overline{\mathfrak{D}_{B,B^*}(\sigma^*)}"] & & & \mathbb{C}^n \arrow[dd, "\overline{\mathfrak{T}_{B^*,C^*}}"] \\
& V \arrow[ul, "\overline{h_B}"'] \arrow[dl, "\overline{h_C}"] \arrow[r, "\sigma^*"] & V^* \arrow[ur, "h_{B^*}"] \arrow[dr, "h_{C^*}"'] & \\
\mathbb{C}^m \arrow[uu, "\overline{\overline{\mathfrak{T}_{C,B}}}"] \arrow[rrr, "\overline{\mathfrak{D}_{C,C^*}(\sigma^*)}"'] & & & \mathbb{C}^n
\end{tikzcd}$$

Abbildung 6.6: Basistransformation der Matrix einer hermiteschen Sesquilinearform

Aus dem Diagramm in Abbildung 6.6 erhalten wir die Basistransformationsformel

$$\mathfrak{D}_C(\sigma) = \overline{\mathfrak{D}_{C,C^*}(\sigma^*)} = \overline{\overline{\mathfrak{T}_{C,B}}\mathfrak{D}_{B,B^*}(\sigma^*)\mathfrak{T}_{B^*,C^*}} = \mathfrak{T}_{C,B}\mathfrak{D}_B(\sigma)\overline{\mathfrak{T}_{C,B}}^t. \tag{6.28}$$

6.5.4 Lemma *Sei V ein $\mathbb{C}$-Vektorraum und σ eine Sesquilinearform auf V. Die von σ induzierte* quadratische Form $q_\sigma\colon V \longrightarrow \mathbb{C}$ *ist genau dann reellwertig, wenn σ hermitesch ist.*

Beweis: Ist σ eine hermitesche Sesquilinearform auf V, so ist wegen $\sigma(x,x) = \overline{\sigma(x,x)}$ stets $q_\sigma(x) = \sigma(x,x) \in \mathbb{R}$. Für die Gegenrichtung berechnen wir

$$q_\sigma(x+y) - q_\sigma(x) - q_\sigma(y) = \sigma(x,y) + \sigma(y,x) \tag{i}$$

und

$$q_\sigma(x+iy) - q_\sigma(x) - q_\sigma(y) = i\sigma(y,x) - i\sigma(x,y) \tag{ii}$$

aus (i) und (ii) erhalten wir

$$[q_\sigma(x+y) - q_\sigma(x) - q_\sigma(y)] + i[q_\sigma(x+iy) - q_\sigma(x) - q_\sigma(y)] = 2\cdot\sigma(x,y). \tag{iii}$$

Ist q_σ reellwertig, so sind beide Klammern reell. Vertauschen wir x und y in (iii), so verändert sich der der Realteil von (iii) nicht. Für die Summe der Imaginärteile ergibt sich $q_\sigma(x+iy)+q_\sigma(y+ix)-2q_\sigma(x)-2q_\sigma(y) = q_\sigma(x)+q_\sigma(y)+\sigma(x,iy)+\sigma(iy,x)+q_\sigma(y)+q_\sigma(x)+\sigma(y,ix)+\sigma(ix,y)-2q_\sigma(x)-2q_\sigma(y) = -i\sigma(x,y)+$

$i\sigma(y,x)-i\sigma(y,x)+i\sigma(x,y)=0$. Damit erhalten wir $\Im(2\sigma(x,y)+2\sigma(y,x))=0$, woraus sich $\sigma(x,y)=\overline{\sigma(y,x)}$ ergibt. Somit ist σ hermitesch. $\square$

6.5.5 Definition Sei V ein Vektorraum über $\mathbb{C}$. Eine hermitesche Sesquilinearform heißt

positiv definit, wenn $(\forall x\in V)[x\neq 0 \;\Rightarrow\; \sigma(x,x)>0]$ gilt,

positiv semi-definit, wenn $(\forall x\in V)[\sigma(x,x)\geq 0]$ gilt,

negativ definit, wenn $(\forall x\in V)[x\neq 0 \;\Rightarrow\; \sigma(x,x)<0]$ gilt,

indefinit, wenn $(\exists x\in V)(\exists y\in V)[\sigma(x,x)>0\wedge\sigma(y,y)<0]$ gilt.

Analog wie für positiv definite Bilinearformen erhalten wir den folgenden Satz.

6.5.6 Satz *Es sei V ein komplexer Vektorraum (d.h. ein Vektorraum über dem Grundkörper $\mathbb{C}$) und σ eine positiv semi-definite Sesquilinearform auf V. Dann gilt die* Cauchy-Schwarzsche Ungleichung

$$(\forall u,v\in V)[|\sigma(u,v)|^2\leq\sigma(u,u)\sigma(v,v)]. \tag{6.29}$$

Für positiv definite Sesquilinearformen σ steht in Gleichung (6.29) genau dann das Gleichheitszeichen, wenn u und v linear abhängig sind.

Beweis: Der Beweis ist eine leichte Modifikation des Beweises von Satz 6.2.3. Für $\alpha,\beta\in\mathbb{C}$ erhalten wir

$$0\leq\sigma(\alpha u+\beta v,\alpha u+\beta v)=\alpha\overline{\alpha}\sigma(u,u)+\beta\overline{\beta}\sigma(v,v)+\alpha\overline{\beta}\sigma(u,v)+\overline{\alpha}\beta\sigma(v,u). \tag{i}$$

Ist $\sigma(u,u)=0$, so wählen wir $\beta=1$ und erhalten $0\leq\alpha\sigma(u,v)+\overline{\alpha}\sigma(v,u)+\sigma(v,v)=\alpha\sigma(u,v)+\overline{\alpha\sigma(u,v)}+\sigma(v,v)$ für alle $\alpha\in\mathbb{C}$, was nur für $\sigma(u,v)=0$ möglich ist. Ist $\sigma(u,u)\neq 0$, so wählen wir $\beta:=\sigma(u,u)=\overline{\beta}$ und teilen Ungleichung (i) durch β. Das ergibt

$$0\leq\alpha\overline{\alpha}+\sigma(u,u)\sigma(v,v)+\alpha\sigma(u,v)+\overline{\alpha}\sigma(v,u). \tag{ii}$$

Nun wählen wir $\alpha:=-\overline{\sigma(u,v)}$ in (ii) und erhalten

$$\begin{aligned}0&\leq\sigma(u,v)\overline{\sigma(u,v)}+\sigma(u,u)\sigma(v,v)-\overline{\sigma(u,v)}\sigma(u,v)-\sigma(u,v)\sigma(v,u)\\&=-\sigma(u,v)\overline{\sigma(u,v)}+\sigma(u,u)\sigma(v,v),\end{aligned} \tag{iii}$$

d.h. $\sigma(u,v)\overline{\sigma(u,v)}\leq\sigma(u,u)\sigma(v,v)$ und damit $|\sigma(u,v)|^2\leq\sigma(u,u)\sigma(v,v)$.
Sind u und v linear abhängig, so ist $u=\lambda v$ für ein $\lambda\in\mathbb{C}$, und es folgt $|\sigma(u,v)|^2=|\lambda|^2\sigma(v,v)=\sigma(\lambda v,\lambda v)\sigma(v,v)$. Sind u und v linear unabhängig und ist σ positiv

definit, so steht für $\alpha \neq 0$ in (i) das $<$-Zeichen. Dann ist aber auch $\sigma(u,u) > 0$ womit folgt, daß auch in (iii) das $<$-Zeichen steht. Damit folgt aber $|\sigma(u,v)|^2 < \sigma(u,u)\sigma(v,v)$. □

6.5.7 Satz *Ist V ein komplexer Vektorraum und σ eine hermitesche positiv definite Sesquilinearform auf V, so definiert $\|v\|_\sigma := \sqrt{\sigma(v,v)} = \sqrt{q_\sigma(v)}$ eine Norm und $d_\sigma(u,v) := \|u-v\|_\sigma$ eine Metrik auf V.*

Die Beweise erfolgen wörtlich wie die der Sätze 6.2.6 und 6.2.7. □

6.6 Unitäre Vektorräume

6.6.1 Definition Ein Paar (V,σ), wobei V ein komplexer Vektorraum und σ eine positiv definite hermitesche Sesquilinearform auf V ist, heißt ein *unitärer Raum*.

Wir übernehmen die für euklidische Räume getroffenen Redeweisen und Definitionen sinngemäß für unitäre Räume. So nennen wir die positiv definite Sesquilinearform das *Skalarprodukt* des unitären Raumes und bezeichnen die von dem Skalarprodukt σ induzierte Norm mit $\| \ \|_\sigma$. Zwei Vektoren u,v heißen *orthogonal* oder *senkrecht,* wenn $\sigma(u,v) = 0$ ist. Wir übertragen die Definition des *orthogonalen Komplements* einer Teilmenge $U \subseteq V$

$$U^{\perp_\sigma} := \{v \in V \mid (\forall x \in V)[\sigma(v,x) = 0]\}.$$

Die Definitionen, Lemmata, Sätze und Korollare 6.3.6 - 6.3.16 übertragen sich mühelos auf unitäre Räume.
Auf dem $\mathbb{C}^n$ erhalten wir das *kanonische Skalarprodukt* durch

$$\langle \mathfrak{x}, \mathfrak{y} \rangle := \mathfrak{x} \cdot \overline{\mathfrak{y}}^t.$$

Die darstellende Matrix des kanonischen Skalarproduktes bezüglich der kanonischen Basis ist die Einheitsmatrix. Die kanonische Basis bildet eine Orthonormalbasis bezüglich des kanonischen Skalarproduktes. Ist $\mathfrak{S} \in \mathbb{C}^{n\times n}$ und gilt $\mathfrak{S} = \overline{\mathfrak{S}}^t$, so definiert

$$\sigma_\mathfrak{S}(\mathfrak{x}, \mathfrak{y}) := \mathfrak{x}\mathfrak{S}\overline{\mathfrak{y}}^t$$

eine hermitesche Sesquilinearform. Wir nennen $\mathfrak{S} \in \mathbb{C}^{n\times n}$ positiv definit, wenn $\mathfrak{S} = \overline{\mathfrak{S}}^t$ und die von $\mathfrak{S}$ erzeugte hermitesche Sesquilinearform positiv definit ist. In Analogie zu Korollar 6.3.17 erhalten wir mit dem gleichen Beweis aus (6.28), daß $\mathfrak{S}$ genau dann eine positiv definite Matrix ist, wenn es ein $\mathfrak{A} \in \mathrm{GL}(n,\mathbb{C})$ gibt mit $\mathfrak{S} = \mathfrak{A}\overline{\mathfrak{A}}^t$.

6.7 Komplexifizierung reeller Vektorräume

Wir wollen einen reellen Vektorraum in natürlicher Weise zu einem komplexen Vektorraum machen. Die Idee dazu besteht darin, aus Vektoren $x, y \in V$ den "komplexen" Vektor $x + iy$ zu bilden und die skalare Multiplikation durch formales "Ausrechnen" $(\xi + i\eta)(x + iy) = \xi x - \eta y + i(\eta x + \xi y)$ festzusetzen. Die Elemente des komplexifizierten Raumes müssen somit als Paare $(x, y) \in V \times V$ aufgefaßt werden, wobei x den Realteil und y den Imaginärteil repräsentiert. Dies ist der Hintergrund der folgenden Definition.

6.7.1 Lemma *Sei V ein reeller Vektorraum. Die Menge $V \times V$ wird mit der punktweisen Addition und der Multiplikation $(\xi + i\eta) \cdot (x, y) := (\xi x - \eta y, \xi y + \eta x)$ ein $\mathbb{C}$-Vektorraum, den wir mit $\tilde{V}$ bezeichnen wollen. Der Vektorraum $\tilde{V}$ heißt die* Komplexifizierung *von V.*

Beweis: Wir haben bereits früher gesehen, daß $V \times V$ mit der punktweisen Addition eine additive abelsche Gruppe bildet. Die Axiome für die skalare Multiplikation rechnet man einfach nach. □

6.7.2 Lemma *Die Abbildung $\tilde{}\colon V \longrightarrow \tilde{V}$ mit $\tilde{x} := (x, 0)$ ist injektiv, und es gelten*

(1) $(\forall x, y \in V)[(x, y) = \tilde{x} + i\tilde{y}]$,

(2) $(\forall \xi, \eta \in \mathbb{R})(\forall x, y \in V)[\widetilde{(\xi x + \eta y)} = \xi\tilde{x} + \eta\tilde{y}]$,

(3) $(\forall \xi, \eta \in \mathbb{R})(\forall x \in V)[(\xi + i\eta)\tilde{x} = (\xi x, \eta x) = \xi\tilde{x} + i\eta\tilde{x}]$.

Beweis: Offensichtlich ist $\tilde{}$ injektiv. Die Behauptung (2) ist klar, und die erste Gleichung in (3) rechnet man anhand der Definition der skalaren Multiplikation sofort nach. Dann ist insbesondere $i\eta(y, 0) = (0, \eta y)$, woraus sich die zweite Gleichung in (3) und (1) sofort ergeben. □

6.7.3 Satz *Ist V ein reeller Vektorraum und B eine Basis von V, so ist $\tilde{B} := \{\tilde{b} \mid b \in B\}$ eine Basis von $\tilde{V}$. Insbesondere ist $\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} \tilde{V}$.*

Beweis: Ist $(x, y) \in \tilde{V}$, so ist nach Lemma 6.7.2 $(x, y) = \tilde{x} + i\tilde{y}$. Da B eine Basis von V ist gibt es ein n, $\xi_1, \ldots, \xi_n$ und $\eta_1, \ldots, \eta_n$ mit $x = \sum_{k=1}^n \xi_k b_k$ und $y = \sum_{j=1}^n \eta_j b_j$. Dann folgt mit Lemma 6.7.2 $(x, y) = \sum_{k=1}^n \widetilde{\xi_k b_k} + i\sum_{j=1}^n \widetilde{\eta_j b_j} = \sum_{k=1}^n \xi_k \tilde{b}_k + i\sum_{j=1}^n \eta_j \tilde{b}_j = \sum_{k=1}^n (\xi_k + i\eta_k)\tilde{b}_k$, und wir haben $\tilde{B}$ als Erzeugendensystem nachgewiesen. Gilt $\sum_{k=1}^n \alpha_k \tilde{b}_k = 0$ mit $\alpha_k = (\xi_k + i\eta_k)$, so folgt mit Lemma 6.7.2 $(0, 0) = \sum_{k=1}^n (\xi_k b_k, \eta_k b_k) = (\sum_{k=1}^n \xi_k b_k, \sum_{k=1}^n \eta_k b_k)$, woraus sich $\xi_1 = \cdots = \xi_n = \eta_1 = \cdots = \eta_n = 0$ und folglich auch $\alpha_1 = \cdots = \alpha_n = 0$ ergeben. Damit ist $\tilde{B}$ eine linear unabhängige Menge und somit eine Basis. $\square$

Damit haben wir nachgewiesen, daß sich unsere Ausgangsidee verwirklichen läßt. Durch die Abbildung $\tilde{\ }$ wird V in $\tilde{V}$ eingebettet, und wir können daher x und $\tilde{x}$ identifizieren. Nach Lemma 6.7.2 hat dann jedes $z \in \tilde{V}$ die Gestalt $z = x + iy$.

6.7.4 Lemma *Sind V und W reelle Vektorräume und $F \in \mathrm{Hom}(V, W)$ und definieren wir $\tilde{F}: \tilde{V} \longrightarrow \tilde{W}$ durch $\tilde{F}(x, y) := (Fx, Fy)$, d.h. $\tilde{F}(\tilde{x} + i\tilde{y}) := \widetilde{Fx} + i\widetilde{Fy}$, so ist $\tilde{F} \in \mathrm{Hom}(\tilde{V}, \tilde{W})$. Die Abbildung $\tilde{\ }: \mathrm{Hom}(V, W) \longrightarrow \mathrm{Hom}(\tilde{V}, \tilde{W})$ mit $F \longmapsto \tilde{F}$ ist ein Homomorphismus der Ringe und $\mathbb{R}$-linear, d.h. es gilt $\widetilde{\xi F + \eta G} = \xi\tilde{F} + \eta\tilde{G}$ für $\xi, \eta \in \mathbb{R}$.*
Ist B eine Basis von V und C eine Basis von W, so folgt $\mathfrak{D}_{\tilde{B},\tilde{C}}(\tilde{F}) = \mathfrak{D}_{B,C}(F)$.

Beweis: Klar durch Nachrechnen. $\square$

6.7.5 Korollar *Für $F \in \mathrm{End}(V)$ gilt $\det F = \det \tilde{F}$. Insbesondere ist dann $Charpol_F = Charpol_{\tilde{F}}$.*

Beweis: Wegen $\det F = \det \mathfrak{D}_{B,B}(F)$ erhalten wir die erste Aussage aus Lemma 6.7.4. Nun gilt aber $Charpol_{\tilde{F}} = \det(\mathfrak{D}_{\tilde{B},\tilde{B}}(\tilde{F}) - x \cdot \mathfrak{E}) = \det(\mathfrak{D}_{B,B}(F) - x \cdot \mathfrak{E}) = Charpol_F$. $\square$

6.7.6 Lemma *Ist (V, σ) ein euklidischer Raum und definieren wir $\tilde{\sigma}(\tilde{x}_1 + i\tilde{y}_1, \tilde{x}_2 + i\tilde{y}_2) := \sigma(x_1, x_2) + \sigma(y_1, y_2) + i[\sigma(x_1, y_2) - \sigma(x_2, y_1)]$ für $x_1, x_2, y_1, y_2 \in V$, so ist $(\tilde{V}, \tilde{\sigma})$ ein unitärer Raum.*

Beweis: Klar durch Nachrechnen. $\square$

6.7.7 Satz *Ist (V, σ) ein euklidischer Vektorraum und $F \in \mathrm{End}(V)$, so gilt $\widetilde{F^\sigma} = \tilde{F}^{\tilde{\sigma}}$.*

Beweis: Es ist $\tilde{\sigma}(\widetilde{F^\sigma}(\tilde{x}_1 + i\tilde{y}_1), \tilde{x}_2 + i\tilde{y}_2) = \tilde{\sigma}(\widetilde{F^\sigma x_1} + i\widetilde{F^\sigma y_1}, \tilde{x}_2 + i\tilde{y}_2) = \sigma(F^\sigma x_1, x_2) + \sigma(F^\sigma y_1, y_2) + i[\sigma(F^\sigma x_1, y_2) - \sigma(x_2, F^\sigma y_1)] = \sigma(x_1, Fx_2) + \sigma(y_1, Fy_2) + i[\sigma(x_1, Fy_2) - \sigma(Fx_2, y_1)] = \tilde{\sigma}(\tilde{x}_1 + i\tilde{y}_1, \widetilde{Fx_2} + i\widetilde{Fy_2}) = \tilde{\sigma}(\tilde{x}_1 + i\tilde{y}_1, \tilde{F}(\tilde{x}_2 + i\tilde{y}_2))$, woraus sich $\widetilde{F^\sigma} = \tilde{F}^{\tilde{\sigma}}$ ergibt. $\square$

6.8 Isometrien unitärer Räume

6.8.1 Definition Eine lineare Abbildung zwischen zwei unitären Räumen, die das Skalarprodukt respektiert, heißt eine Isometrie.

Mit dem gleichen Beweis wie in Satz 6.4.2 erhalten wir, daß jede Isometrie unitärer Räume schon ein Monomorphismus der Vektorräume und damit im endlich dimensionalen Fall bereits ein Isomorphismus der Vektorräume ist.

6.8.2 Satz *Sei* (V, σ) *ein endlich dimensionaler unitärer Raum. Ist* $F: V \longrightarrow V$ *eine Isometrie von* V *in sich, so nennen wir* F *eine* unitäre *Abbildung. Ein Endomorphismus* $F \in \mathrm{End}(V)$ *ist genau dann unitär, wenn* $F^{-1} = F^\sigma$ *ist.*
Die Menge

$$U(V, \sigma) := \{F \in \mathrm{End}(V) \mid F^{-1} = F^\sigma\}$$

ist mit der Komposition von Abbildungen eine Gruppe. Sie heißt die allgemeine unitäre Gruppe *von* (V, σ).
Für $F \in U(V, \sigma)$ *gilt* $|\det F| = 1$.
Die Menge

$$SU(V, \sigma) := \{F \in U(V, \sigma) \mid \det F = 1\}$$

ist eine Untergruppe der $U(V, \sigma)$. *Sie heißt die* spezielle unitäre Gruppe *von* (V, σ).

Die Beweise erfolgen wie im euklidischen Fall. □

6.8.3 Satz *Ist* (V, σ) *ein euklidischer Raum und* $F \in \mathrm{End}(V)$ *eine Isometrie von* (V, σ), *so ist dessen komplexe Fortsetzung* $\tilde{F}$ *eine Isometrie von* $(\tilde{V}, \tilde{\sigma})$.

Beweis: Wir haben nach Satz 6.8.2 nur nachzuprüfen, daß $\tilde{F}^{\tilde{\sigma}} = \tilde{F}^{-1}$ ist. Das ergibt sich aber aus den Lemmata 6.7.4 und 6.7.7 wegen $\tilde{F}^{-1} = \widetilde{F^{-1}} = \widetilde{F^\sigma} = \tilde{F}^{\tilde{\sigma}}$. □

Ist (V, σ) ein euklidischer oder unitärer Vektorraum, so läßt sich für jeden Endomorphismus $F \in \mathrm{End}(V)$ eine Norm

$$\|F\|_\sigma := \sup\left\{\sqrt{\frac{\sigma(Fx, Fx)}{\sigma(x, x)}} \;\middle|\; x \in V\right\} \tag{6.30}$$

definieren. Für diese Norm erhalten wir

$$\|\alpha F\|_\sigma = \sqrt{\alpha\overline{\alpha}}\|F\|_\sigma = |\alpha| \cdot \|F\|_\sigma. \tag{6.31}$$

Die Isometrien sind dann genau die Endomorphismen mit Norm 1, d.h. wir haben

$$O(V,\sigma) = \{F \in \mathrm{End}(V) \bigm| \|F\|_\sigma = 1\} \tag{6.32}$$

für euklidische bzw.

$$U(V,\sigma) = \{F \in \mathrm{End}(V) \bigm| \|F\|_\sigma = 1\} \tag{6.33}$$

für unitäre Räume. Die Normen von Endomorphismen spielen in der Funktionalanalysis und in der numerischen Mathematik eine wichtige Rolle.
Betrachten wir wieder den Spezialfall des unitären Raumes $\mathbb{C}^n$ mit dem kanonischen Skalarprodukt, so ist dessen darstellende Matrix bezüglich der kanonischen Basis die Einheitsmatrix. Aus (6.27) folgt daher, daß $\overline{\mathfrak{A}}$ genau dann unitär ist, wenn $\mathfrak{A}^{-1} = \overline{\mathfrak{A}}^t$ ist. Daher definiert man

$$U(n) := \{\mathfrak{A} \in \mathbb{C}^{n\times n} \mid \mathfrak{A}^{-1} = \overline{\mathfrak{A}}^t\}$$

und nennt $U(n)$ die *allgemeine unitäre Gruppe*. Ebenso setzt man

$$SU(n) := \{\mathfrak{A} \in U(n) \mid \det \mathfrak{A} = 1\}$$

und nennt $SU(n)$ die *spezielle unitäre Gruppe*. Die Matrizen in $U(n)$ heißen *unitäre Matrizen*. In Analogie zu Satz 6.4.15 erhalten wir

6.8.4 Satz *Eine Matrix ist genau dann unitär, wenn ihre Spalten- bzw. Zeilenvektoren eine Orthonormalbasis des $\mathbb{C}^n$ bezüglich des kanonischen Skalarproduktes bilden.*

6.9 Normale Endomorphismen

Vorbemerkung Unitäre und euklidische Räume haben sehr viele gemeinsame Eigenschaften. Zur Vereinfachung der Notation bedeute daher $\mathbb{K}$ im euklidischen Falle den Körper $\mathbb{R}$ und im unitären Falle den Körper $\mathbb{C}$.

6.9.1 Definition Sei (V,σ) ein euklidischer oder unitärer Vektorraum. Ein Endomorphismus $F \in \mathrm{End}(V)$ heißt normal, wenn er mit seiner adjungierten F^σ vertauschbar ist, d.h. wenn $F^\sigma F = FF^\sigma$ gilt.

6.9.2 Lemma *Ist F ein normaler Endomorphismus, so sind alle Elemente von $\mathbb{K}[F]$ mit den Elementen in $\mathbb{K}[F^\sigma]$ vertauschbar.*

Beweis: Ist $A = \sum_{k=0}^m \alpha_k F^k$ und $B = \sum_{j=0}^n \beta_j F^{\sigma j}$, so erhalten wir $AB = \sum_{l=0}^{m+n} \sum_{k+j=l} \alpha_k\beta_j F^k F^{\sigma j} = \sum_{l=0}^{m+n} \sum_{k+j=l} \alpha_k\beta_j F^{\sigma j} F^k = BA$. □

6.9.3 Lemma *Sind $F_1, \dots, F_n \in \operatorname{End}(V)$ und gilt $\sum_{j=1}^n F^\sigma{}_j F_j = 0$, so folgt $F_1 = \dots = F_n = 0$.*

Beweis: Für $x \in V$ erhalten wir $0 = \sigma(\sum_{j=1}^n F^\sigma{}_j F_j x, x) = \sum_{j=1}^n \sigma(F^\sigma{}_j F_j x, x) = \sum_{j=1}^n \sigma(F_j x, F_j x)$. Das ist nur möglich, wenn $F_j x = 0$ ist. Da wir $x \in V$ beliebig gewählt haben, folgt $F_j = 0$ für $j = 1, \dots, n$. □

6.9.4 Satz *Ist (V, σ) ein euklidischer oder unitärer Vektorraum und F ein normaler Endomorphismus von V. Dann enthält $\mathbb{K}[F]$ keine von 0 verschiedenen nilpotenten Elemente.*

Beweis: Wir zeigen zunächst, daß ein nilpotenter Endomorphismus, für den $F = F^\sigma$ gilt, bereits 0 ist. Ist $F \neq 0$ und nilpotent, so gibt es ein $k \in \mathbb{N}$, so daß $F^{2^k} \neq 0$, aber $F^{2^{k+1}} = 0$ ist. Dann folgt $F^{\sigma 2^k} F^{2^k} = F^{2^{k+1}} = 0$ und damit nach Lemma 6.9.3 $F^{2^k} = 0$. Widerspruch.
Für $A \in \mathbb{K}[F]$ gilt aber $(A^\sigma A)^\sigma = A^\sigma A$, und $(A^\sigma A)$ ist wegen der Vertauschbarkeit von A und A^σ nilpotent, wenn A dies ist. Damit ist $(A^\sigma A) = 0$, woraus sich $\sigma(Ax, Ax) = \sigma(A^\sigma A x, x) = 0$ für alle $x \in V$ und damit $A = 0$ ergibt. □

Zusammen mit Satz 5.7.4 erhalten wir die folgende Aussage.

6.9.5 Korollar *Jeder splitte normale Endomorphismus eines euklidischen und jeder normale Endomorphismus eines unitären Vektorraumes ist halbeinfach und damit diagonalisierbar.*

6.9.6 Lemma *Sei V ein K-Vektorraum und F ein halbeinfacher Endomorphismus von V mit Minimalzerlegung $F = \sum_{j=1}^r \lambda_j C_j$. Gilt dann $F = \sum_{k=1}^s \mu_k B_k$ für ein vollständiges Orthogonalsystem $B_1, \dots, B_s$, so ist $r = s$, und es gibt eine Permutation $\pi \in S_r$ mit $\lambda_j = \mu_{\pi(j)}$ und $C_j = B_{\pi(j)}$.*

Beweis: Sei $F = \sum_{j=1}^r \lambda_j C_j = \sum_{k=1}^s \mu_k B_k$ und sowohl $C_1, \dots, C_r$ als auch $B_1, \dots, B_s$ ein vollständiges Orthogonalsystem. Dann gilt, wie im Beweis von Satz 5.7.4 gezeigt,

$$f(F) = \sum_{j=1}^r f(\lambda_j) C_j = \sum_{k=1}^s f(\mu_k) B_k \tag{i}$$

für jedes Polynom $f \in K[x]$. Wählen wir in (i) für $f_j := \prod_{\substack{k=1 \\ k \neq j}}^s (x - \mu_k)$, so erhalten wir $f_j(\mu_j) B_j = f_j(F) \in K[F]$ und damit

$$B_j \in K[F] \quad \text{für} \quad j = 1, \dots, s. \tag{ii}$$

Demnach gibt es $C_{j_1}, \ldots, C_{j_m} \subseteq \{C_1, \ldots, C_r\}$ und $\gamma_1, \ldots, \gamma_m \in K$ mit

$$B_k = \sum_{i=1}^{m} \gamma_i C_{j_i} \quad \text{für} \quad k = 1, \ldots, s. \tag{iii}$$

Da B_j idempotent ist, folgt aus (iii) $\sum_{i=1}^m \gamma_i^2 C_{j_i} = B_j^2 = B_j = \sum_{i=1}^m \gamma_i C_{j_i}$ und damit $\gamma_i \in \{0,1\}$. Also gilt

$$B_k = \sum_{i=1}^{m_k} C_{j_i} \quad \text{für} \quad k = 1, \ldots, s. \tag{iv}$$

Aus Symmetriegründen erhalten wir auch

$$C_j = \sum_{i=1}^{n_j} B_{ki} \quad \text{für} \quad j = 1, \ldots, r. \tag{v}$$

Da nun sowohl $C_1, \ldots, C_r$ als auch $B_1, \ldots, B_s$ linear unabhängig sind (vgl. 5.7.3), folgt $m_k = 1$ für $k = 1, \ldots, s$ und $n_j = 1$ für $j = 1, \ldots, r$. Also ist $B_j \in \{C_1, \ldots, C_r\}$ für $j = 1, \ldots, s$ und $C_k \in \{B_1, \ldots, B_s\}$. Damit ist $r = s$, und es gibt eine Permutation $\pi \in S_r$ mit $C_j = B_{\pi(j)}$, woraus dann auch $\lambda_j = \mu_{\pi(j)}$ folgt. □

6.9.7 Lemma *Ist (V, σ) ein endlich dimensionaler unitärer oder euklidischer Raum, so gilt* Im $F^\sigma = (\text{Kern } F)^{\perp_\sigma}$ *und* Kern $F^\sigma = (\text{Im } F)^{\perp_\sigma}$.

Beweis: Mit Lemma 6.3.6 erhalten wir $\sigma^*[(\text{Kern } F)^{\perp_\sigma}] = (\text{Kern } F)^\perp = \text{Im } F^* = \sigma^*[\text{Im } F^\sigma]$ und damit $(\text{Kern } F)^{\perp_\sigma} = \text{Im } F^\sigma$. Ähnlich folgt $\sigma^*[(\text{Im } F)^{\perp_\sigma}] = (\text{Im } F)^\perp = \text{Kern } F^* = \sigma^*[\text{Kern } F^\sigma]$. □

6.9.8 Lemma *Sei (V, σ) ein endlich dimensionaler unitärer oder euklidischer Vektorraum. Für jeden normalen Endomorphismus $F \in$* End(V) *gilt dann* Kern $F^\sigma =$ Kern F *und* Im $F^\sigma =$ Im F. *Insbesondere ist dann $V =$* Kern $F \mathbin{\textcircled{\perp}}$ Im F.

Beweis: Es gilt

$$\begin{aligned} x \in \text{Kern } F \quad &\Leftrightarrow \quad \sigma(Fx, Fx) = 0 \\ &\Leftrightarrow \quad \sigma(F^\sigma F x, x) = 0 \\ &\Leftrightarrow \quad \sigma(F F^\sigma x, x) = 0 \\ &\Leftrightarrow \quad \sigma(F^\sigma x, F^\sigma x) = 0 \quad \Leftrightarrow \quad x \in \text{Kern } F^\sigma. \end{aligned}$$

Mit Lemma 6.9.7 folgt dann Im $F^\sigma = (\text{Kern } F)^{\perp_\sigma} = (\text{Kern } F^\sigma)^{\perp_\sigma} =$ Im F. Wegen

$$\begin{aligned} V &= \text{Kern } F \textcircled{\perp} (\text{Kern } F)^{\perp_\sigma} \\ &= \text{Kern } F \textcircled{\perp} \text{Im } F^\sigma \\ &= \text{Kern } F \textcircled{\perp} \text{Im } F \end{aligned}$$

ist alles bewiesen. □

6.9.9 Lemma *Ist (V,σ) ein euklidischer oder unitärer Vektorraum und $F \in \text{End}(V)$ ein normaler Endomorphismus, so ist $\text{Eig}(\lambda, F) = \text{Eig}(\overline{\lambda}, F^\sigma)$ für jeden Eigenwert λ von F.*

Beweis: Wir betrachten den Endomorphismus $G := F - \lambda id$. Dann gilt $G^\sigma = F^\sigma - \overline{\lambda} id$, und wir erhalten

$$\begin{aligned} G^\sigma G &= (F^\sigma - \overline{\lambda} id)(F - \lambda id) \\ &= F^\sigma F - \overline{\lambda} F - \lambda F^\sigma + \lambda\overline{\lambda} id \\ &= F F^\sigma - \lambda F^\sigma - \overline{\lambda} F + \overline{\lambda}\lambda id \\ &= (F - \lambda id)(F^\sigma - \overline{\lambda} id) = G G^\sigma. \end{aligned}$$

Damit ist G normal, und mit Lemma 6.9.8 folgt

$$\text{Eig}(\lambda, F) = \text{Kern } G = \text{Kern } G^\sigma = \text{Eig}(\overline{\lambda}, F^\sigma).$$ □

6.9.10 Satz *Sei (V,σ) ein endlich dimensionaler euklidischer oder unitärer Vektorraum und $F \in \text{End}(V)$ split über $\mathbb{K}$ (im unitären Fall ist dies immer erfüllt). Der Endomorphismus F ist genau dann normal, wenn es eine Orthonormalbasis von Eigenvektoren von F gibt.*

Beweis: $\Leftarrow$: Sei B eine Basis von V aus Eigenvektoren von F. Für $b \in B$ erhalten wir $\sigma(F^\sigma b, b) = \sigma(b, Fb) = \overline{\lambda}\sigma(b,b) = \sigma(\overline{\lambda} b, b)$ und damit $F^\sigma b = \overline{\lambda} b$. Also folgt $F^\sigma F b = F^\sigma \lambda b = \lambda\overline{\lambda} b = \overline{\lambda}\lambda b = \overline{\lambda} F b = F \overline{\lambda} b = F F^\sigma b$. Damit sind F und F^σ auf allen Basisvektoren und damit auf ganz V vertauschbar.
$\Rightarrow$: Ist F normal, so ist F nach Satz 6.9.4 halbeinfach. Ist $F = \sum_{j=1}^r \lambda_j C_j$ die Minimalzerlegung, so gilt nach Satz 5.8.4 $V = U_1 \oplus \cdots \oplus U_r$ mit $U_i := C_i[V]$. Wir behaupten, daß sogar

$$V = U_1 \textcircled{\perp} \cdots \textcircled{\perp} U_r \tag{i}$$

gilt. Sei dazu $u_i \in U_i$ und $u_j \in U_j$ für $i \neq j$. Da $C_i, C_j \in \mathbb{K}[F]$ sind nach Lemma 6.9.2 C_i und C_j^σ vertauschbar. Dann folgt $\sigma(u_i, u_j) = \sigma(C_i u_i, C_j u_j) = \sigma(u_i, C_i^\sigma C_j u_j) = \sigma(u_i, C_j C_i^\sigma u_j) = 0$, denn wegen $i \neq j$ ist $u_j \in \text{Kern } C_i$ und nach Lemma 6.9.8 ist $\text{Kern } C_i = \text{Kern } C_i^\sigma$. Damit ist (i) gezeigt. Nach Satz 5.8.13

gilt nun $U_i = \mathrm{Eig}(\lambda_i, F)$. Die Räume $(U_i, \sigma \restriction U_i)$ sind wieder euklidisch bzw. unitär und besitzen daher Orthonormalbasen B_i. Dann ist aber $B = B_1 \cup \dots \cup B_r$ eine Orthonormalbasis von V, die aus Eigenvektoren von F besteht. □

Zusammen mit Satz 5.8.13 erhalten wir als Korollar des Satzes 6.9.10

6.9.11 Korollar *Jeder splitte normale Automorphismus ist diagonalisierbar.*

6.9.12 Bemerkung Wir betrachten nun den $\mathbb{C}^n$ mit dem kanonischen Skalarprodukt. Für $\mathfrak{A} \in \mathbb{C}^{n\times n}$ ist der Endomorphismus $\overline{\mathfrak{A}}$ genau dann normal, wenn $\overline{\overline{\mathfrak{A}}}^t \cdot \overline{\mathfrak{A}} = \overline{\mathfrak{A}} \cdot \overline{\overline{\mathfrak{A}}}^t$ ist, und dies ist genau dann der Fall, wenn $\overline{\mathfrak{A}}^t \cdot \mathfrak{A} = \mathfrak{A} \cdot \overline{\mathfrak{A}}^t$ ist. Daher nennen wir eine Matrix $\mathfrak{A} \in \mathbb{C}^{n\times n}$ *normal*, wenn $\overline{\mathfrak{A}}^t \cdot \mathfrak{A} = \mathfrak{A} \cdot \overline{\mathfrak{A}}^t$ ist. Da $\mathbb{C}$ algebraisch abgeschlossen ist, ist jede normale Matrix split über $\mathbb{C}$. Damit existiert nach Satz 6.9.10 eine Orthonormalbasis B von Eigenvektoren. Damit ist $\mathfrak{T}_{B,\mathbb{E}^n} \cdot \mathfrak{A} \cdot \mathfrak{T}_{\mathbb{E}^n,B}$ eine Diagonalmatrix, deren Hauptdiagonale alle Eigenwerte in ihrer Vielfachheit gezählt enthält. Da das kanonische Skalarprodukt sowohl bezüglich der kanonischen Basis als auch der Basis B durch die Einheitsmatrix dargestellt wird, folgt $\mathfrak{T}_{B,\mathbb{E}^n}^{-1} = \overline{\mathfrak{T}_{B,\mathbb{E}^n}}^t$ aus (6.28) und die Matrix $\mathfrak{T}_{B,\mathbb{E}^n}$ ist unitär. Damit haben wir den folgenden Satz gezeigt.

6.9.13 Satz *Ist $\mathfrak{A} \in \mathbb{C}^{n\times n}$ eine normale Matrix, so gibt es eine unitäre Matrix $\mathfrak{S}$ mit*

$$\mathfrak{S} \cdot \mathfrak{A} \cdot \overline{\mathfrak{S}}^t = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \dots, \lambda_n$ alle Eigenwerte von $\mathfrak{A}$ in ihrer Vielfachheit gezählt sind.

6.9.14 Bemerkung Um eine normale Matrix $\mathfrak{A} \in \mathbb{C}^{n\times n}$ auf Diagonalgestalt zu transformieren, muß man nicht zuerst deren Minimalzerlegung kennen. Man weiß ja aus Satz 6.9.13, daß eine Orthonormalbasis aus Eigenvektoren existiert. Diese läßt sich finden, indem man zunächst die Wurzeln des charakteristischen Polynoms bestimmt, d.h. die Gleichung

$$\det(\mathfrak{A} - x\mathfrak{E}) = 0$$

löst. Dies liefert alle verschiedenen Eigenwerte $\lambda_1, \dots, \lambda_r$ zusammen mit deren algebraischen Vielfachheiten. Nun bestimmt man die Eigenvektoren durch Lösen der Gleichungssysteme

$$\mathfrak{x}(\mathfrak{A} - \lambda_i\mathfrak{E}) = 0.$$

Die Eigenvektoren zu einem Eigenwert λ_i orthonormalisiert man nach Schmidt zu einer Orthonormalbasis $\mathfrak{b}_{i1}, \ldots, \mathfrak{b}_{in_i}$ von $\mathrm{Eig}(\lambda_i, \mathfrak{A})$. Dadurch erhalten wir eine Orthonormalbasis $(\mathfrak{b}_{11}, \ldots, \mathfrak{b}_{1n_1}, \ldots, \mathfrak{b}_{r1}, \ldots, \mathfrak{b}_{rn_r})$ von $\mathbb{C}^n$, und für

$$\mathfrak{B}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathfrak{b}_{i1} \\ \vdots \\ \mathfrak{b}_{in_i} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{gilt} \quad \mathfrak{B}_i\mathfrak{A} = \lambda_i \mathfrak{B}_i \quad \text{und daher} \quad \mathfrak{B}_i \mathfrak{A} \overline{\mathfrak{B}_i}^t = \lambda_i \mathfrak{C}_i$$

mit

$$\lambda_i \mathfrak{C}_i = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & & & \vdots \\ \vdots & & 0 & & & & \vdots \\ \vdots & & & \lambda_i \mathfrak{E}^{(n_i,n_i)} & & & \vdots \\ \vdots & & & & 0 & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}.$$

Die Matrix $\mathfrak{B} := \begin{pmatrix} \mathfrak{b}_{1n_1} \\ \vdots \\ \mathfrak{b}_{rn_r} \end{pmatrix}$ ist unitär, da ihre Zeilenvektoren eine Orthonormalbasis bilden, und wir erhalten

$$\mathfrak{B}\mathfrak{A}\overline{\mathfrak{B}}^t = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Damit haben wir ein Verfahren zur Berechnung der unitären Transformationsmatrix einer normalen Matrix. Der schwierigste Punkt dabei ist die Berechnung der Eigenwerte, da das charakteristische Polynom den Grad n hat. Hier müssen in der Regel Methoden der numerischen Mathematik in Anspruch genommen werden.

6.10 Unitäre und orthogonale Abbildungen

Wir erinnern daran, daß wir die Isometrien euklidischer Räume als *orthogonale* und die Isometrien unitärer Räume als *unitäre* Abbildungen bezeichnen. Wir wollen in diesem Abschnitt einige Eigenschaften unitärer und orthogonaler Automorphismen genauer studieren.

6.10.1 Satz *Sei (V, σ) ein euklidischer oder unitärer Raum. Sind B und C zwei Orthonormalbasen von (V, σ), so ist die Transformationsmatrix $\mathfrak{T}_{B,C}$ orthogonal bzw. unitär.*

Beweis: Sind B und C Orthonormalbasen, so ist $\mathfrak{D}_B(\sigma) = \mathfrak{D}_C(\sigma) = \mathfrak{E}$. Nach Satz 6.1.17 oder (6.28) folgt daher $\mathfrak{E} = \mathfrak{T}_{C,B} \cdot {\mathfrak{T}_{C,B}}^t$ bzw. $\mathfrak{E} = \mathfrak{T}_{C,B} \cdot \overline{\mathfrak{T}_{C,B}}^t$. Damit ist ${\mathfrak{T}_{C,B}}^{-1} = \mathfrak{T}^t_{C,B}$ bzw. ${\mathfrak{T}_{C,B}}^{-1} = \overline{\mathfrak{T}_{C,B}}^t$. □

6.10.2 Satz *Jeder unitäre und orthogonale Automorphismus ist normal.*

Beweis: Ist F ein orthogonaler oder unitärer Automorphismus, so gilt $F^\sigma F = F^{-1}F = id = FF^{-1} = FF^\sigma$. □

6.10.3 Satz *Ist F ein orthogonaler oder unitärer Automorphismus und λ ein Eigenwert von F, so ist $|\lambda| = 1$. Damit sind 1 und -1 die einzigen möglichen reellen Eigenwerte.*

Beweis: Ist F orthogonal oder unitär, so ist F eine Isometrie, und für einen Eigenvektor v zum Eigenwert λ erhalten wir $\|v\|_\sigma = \|Fv\|_\sigma = \|\lambda v\|_\sigma = |\lambda|\|v\|_\sigma$. Damit ist $|\lambda| = 1$. □

6.10.4 Satz *Jeder unitäre Automorphismus F ist halbeinfach. Damit ist F diagonalisierbar. Ist $F = \sum_{j=1}^r \lambda_j C_j$ die Minimalzerlegung von F, so ist jeder der Endomorphismen C_j selbstadjungiert, d.h. es gilt $C_j^\sigma = C_j$.*

Beweis: Da F insbesondere normal ist, ist F nach Korollar 6.9.11 diagonalisierbar. Sei $F = \sum_{j=1}^r \lambda_j C_j$ die Minimalzerlegung. Nach Satz 6.10.3 ist $\lambda_i^{-1} = \overline{\lambda_i}$, und wir erhalten $(\sum_{j=1}^r \overline{\lambda_j} C_j)(\sum_{j=1}^r \lambda_j C_j) = \sum_{j=1}^r \overline{\lambda_j}\lambda_j C_j = id$. Damit ist $\sum_{j=1}^r \overline{\lambda_j} C_j = F^{-1} = F^\sigma = (\sum_{j=1}^r \lambda_j C_j)^\sigma = \sum_{j=1}^r \overline{\lambda_j} C_j^\sigma$. Da $C_1^\sigma, \ldots, C_r^\sigma$ ebenfalls ein vollständiges Orthogonalsystem bilden, folgt aus der in Lemma 6.9.6 gezeigten Eindeutigkeit der Darstellung $C_j = C_j^\sigma$ für $j = 1, \ldots, r$. □

Wir erinnern an die orthogonale und unitäre Gruppe $O(n) := \{\mathfrak{A} \in \mathbb{R}^{n\times n} \mid \mathfrak{A}^{-1} = \mathfrak{A}^t\}$ bzw. $U(n) := \{\mathfrak{A} \in \mathbb{C}^{n\times n} \mid \mathfrak{A}^{-1} = \overline{\mathfrak{A}}^t\}$.

6.10.5 Satz *Sei (V, σ) ein euklidischer bzw. unitärer Vektorraum. Eine Matrix ist genau dann unitär oder orthogonal, wenn sie die darstellende Matrix eines orthogonalen bzw. unitären Automorphismus bezüglich einer Orthonormalbasis von (V, σ) ist. D.h. wir haben für eine Orthonormalbasis B von (V, σ)*

$$O(n) = \{\mathfrak{D}_{B,B}(F) |\ F \in O(V, \sigma)\}$$

bzw.

$$U(n) = \{\mathfrak{D}_{B,B}(F) |\ F \in U(V, \sigma)\}.$$

Beweis: Für $F \in O(V, \sigma)$ bzw. $F \in U(V, \sigma)$ folgt nach (6.27) bzw. nach Satz 6.1.16 $\mathfrak{D}_{B,B}(F)^{-1} = \mathfrak{D}_{B,B}(F^{-1}) = \mathfrak{D}_{B,B}(F^\sigma) = \mathfrak{D}_B(\sigma) \cdot \overline{\mathfrak{D}_{B,B}(F)}^t \cdot \mathfrak{D}_B(\sigma)^{-1} = \overline{\mathfrak{D}_{B,B}(F)}^t$. Damit ist $\{\mathfrak{D}_{B,B}(F) |\ F \in O(V, \sigma)\} \subseteq O(n)$ im euklidischen und $\{\mathfrak{D}_{B,B}(F) |\ F \in U(V, \sigma)\} \subseteq U(n)$ im unitären Fall. Ist umgekehrt $\mathfrak{A} \in O(n)$ bzw. $\mathfrak{A} \in U(n)$, so erhalten wir für $F := h_B \circ \overline{\mathfrak{A}} \circ h_B{}^{-1}$ sofort $\mathfrak{D}_{B,B}(F) = \mathfrak{A}$ und damit $\mathfrak{D}_{B,B}(F^\sigma) = \mathfrak{D}_B(\sigma)\overline{\mathfrak{A}}^t \mathfrak{D}_B(\sigma)^{-1} = \overline{\mathfrak{A}}^t = \mathfrak{A}^{-1} = \mathfrak{D}_{B,B}(F^{-1})$. Damit ist $F \in O(V, \sigma)$ bzw. $F \in U(V, \sigma)$, und wir haben auch die umgekehrte Inklusion gezeigt. □

6.10.6 Satz *Ist $\mathfrak{A} \in \mathbb{C}^{n \times n}$ eine unitäre Matrix, so gibt es eine unitäre Matrix $\mathfrak{S}$ mit*

$$\mathfrak{S}\mathfrak{A}\overline{\mathfrak{S}}^t = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ alle Eigenwerte von $\mathfrak{A}$ in ihrer Vielfachheit gezählt sind.

Wir wollen nun Normaldarstellungen orthogonaler Automorphismen studieren. Die Schwierigkeit, auf die wir hier unmittelbar stoßen, ist, daß die Endomorphismen euklidischer Räume nicht notwendigerweise split über $\mathbb{R}$ sind. Hier müssen wir die Theorie der Komplexifizierung des Abschnitts 6.7 zu Hilfe nehmen.

6.10.7 Satz *Sei (V, σ) ein euklidischer Vektorraum und F ein orthogonaler Automorphismus von V. Dann gibt es es eine Orthonormalbasis C von (V, σ), so daß*

$$\mathfrak{D}_{C,C}(F) = \begin{pmatrix} \boxed{\mathfrak{E}^{(p,p)}} & 0 & \cdots & \cdots & 0 \\ 0 & \boxed{-\mathfrak{E}^{(q,q)}} & \ddots & & \vdots \\ \vdots & \ddots & \boxed{\begin{matrix} \xi_1 & \eta_1 \\ -\eta_1 & \xi_1 \end{matrix}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \boxed{\begin{matrix} \xi_r & \eta_r \\ -\eta_r & \xi_r \end{matrix}} \end{pmatrix},$$

wobei p die Anzahl der Eigenwerte 1, q die Anzahl der Eigenwerte -1 und mit $\alpha_k := \xi_k + i\eta_k$ für $k = 1, \ldots, r$ die Zahlen $\alpha_1, \overline{\alpha_1}, \ldots, \alpha_r, \overline{\alpha_r}$ alle komplexen Eigenwerte von F in ihrer Vielfachheit gezählt durchlaufen.

Beweis: Wir setzen den Automorphismus F komplex zu $\tilde{F}$ fort. Nach Satz 6.8.3 ist $\tilde{F}$ dann unitär. Die charakteristischen Polynome von $\tilde{F}$ und F stimmen nach Korollar 6.7.5 überein und sind daher Polynome in $\mathbb{R}[x]$. Dann zerfällt $Charpol_{\tilde{F}}$ über $\mathbb{C}$ in die Linearfaktoren

$$Charpol_{\tilde{F}} = (x+1)^p \cdot (x-1)^s \cdot (x-\alpha_1)(x-\overline{\alpha_1}) \cdots (x-\alpha_r)(x-\overline{\alpha_r}).$$

Insbesondere ist mit jedem Eigenwert auch dessen konjungiert Komplexe ein Eigenwert. Sind $\lambda_1, \ldots, \lambda_s$ die verschiedenen reellen und $\lambda_{s+1}, \ldots, \lambda_t$ die verschiedenen komplexen Eigenwerte von $\tilde{F}$, so ist

$$\tilde{F} = \sum_{j=1}^{s} \lambda_j C_j + \sum_{k=s+1}^{t} \lambda_k C_k \qquad \text{(i)}$$

die Minimalzerlegung von $\tilde{F}$. (Hier ist zu bemerken, daß für $j \leq s$ wegen Satz 6.10.3 $\lambda_j \in \{-1, 1\}$ ist und damit $s \leq 2$ ist). Dann erhalten wir

$$\tilde{V} = U_1 \oplus \cdots \oplus U_t \qquad \text{(ii)}$$

mit $U_k := C_k[\tilde{V}] = \mathrm{Eig}(\lambda_k, \tilde{F})$ für $k = 1, \ldots, t$. Zu U_k sei $\overline{U_k} := \{\bar{z} \mid z \in U_k\}$. Zu jedem λ_k ist $\overline{\lambda_k} \in \{\lambda_{s+1}, \ldots, \lambda_t\}$ und damit $\overline{\lambda_k} = \lambda_{k'}$. Dann erhalten wir auch

$$\overline{U_k} = U_{k'}. \qquad \text{(iii)}$$

Definieren wir nämlich für einen Endomorphismus $G \in \mathrm{End}(\tilde{V})$ dessen Konjugation durch $\overline{G}(z) := \overline{G(\bar{z})}$, so gilt für die Fortsetzung eines reellen Endomorphismus $\overline{\tilde{F}}(z) = \tilde{F}(\bar{z})$. Für $u \in \overline{U_j}$ folgt daher $\tilde{F}(u) = \overline{\tilde{F}(\bar{u})} = \overline{\lambda_j \bar{u}} = \overline{\lambda_j} u = \lambda_{j'} u$ und damit $u \in \mathrm{Eig}(\lambda_{j'}, \tilde{F}) = U_{j'}$. Umgekehrt erhalten wir für $u \in U_{j'}$, daß

$\tilde{F}(\overline{u}) = \overline{\tilde{F}(u)} = \overline{\lambda_{j'} u} = \lambda_j \overline{u}$ und damit $u \in \overline{\mathrm{Eig}(\lambda_j, \tilde{F})} = \overline{U_j}$, d.h. $U_{j'} \subseteq \overline{U_j}$ ist. Damit erhalten wir

$$\tilde{V} = U_1 \bigoplus \cdots \bigoplus U_s \bigoplus U_{s+1} \bigoplus \overline{U_{s+1}} \bigoplus \cdots \bigoplus U_l \bigoplus \overline{U_l} \tag{iv}$$

für ein geeignetes l. Nun wählen wir Orthonormalbasen der Räume U_i und $\overline{U_i}$, die zusammen eine Orthonormalbasis $b_1, \dots, b_{p+q}, z_1, \overline{z_1}, \dots, z_r, \overline{z_r}$ von $\tilde{V}$ mit $b_1, \dots, b_{p+q} \in V$ ergeben. Wir wollen B zu einer "reellen" Basis $B' \subseteq V$ transformieren. Dazu betrachten wir eines der Paare $z, \overline{z}$ von konjungierten Eigenvektoren. Für $z = x + iy$ ergibt sich $0 = \tilde{\sigma}(z, \overline{z}) = \tilde{\sigma}(x + iy, x - iy) = \sigma(x,x) - \sigma(y,y) + i[\sigma(x,y) + \sigma(x,y)]$, woraus $\|x\|_\sigma^2 = \|y\|_\sigma^2$ und

$$\sigma(x,y) = 0 \tag{v}$$

folgen. Aus $1 = \tilde{\sigma}(z,z) = \sigma(x,x) + \sigma(y,y) + i[\sigma(x,y) - \sigma(x,y)]$ erhalten wir $\|x\|_\sigma^2 + \|y\|_\sigma^2 = 1$. Damit ist $\|x\|_{\tilde{\sigma}} = \|x\|_\sigma = \|y\|_\sigma = \|y\|_{\tilde{\sigma}} = \frac{1}{2}\sqrt{2}$. Zusammen mit (v) zeigt dies, daß $(\sqrt{2}x, \sqrt{2}y)$ orthonormale Vektoren sind. Wegen

$$2x = z + \overline{z}, \quad -2y = iz - i\overline{z} \tag{vi}$$

ergibt sich als Transformationsmatrix

$$\mathfrak{T}_{(\sqrt{2}x, \sqrt{2}y),(z,\overline{z})} = \frac{1}{2}\sqrt{2}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}. \tag{vii}$$

Ist $\xi + i\eta$ der Eigenwert zu $z = x + iy$ und definieren wir $U_z := \langle z, \overline{z} \rangle$ sowie $C_z := \{\sqrt{2}x, \sqrt{2}y\}$, so ist

$$\mathfrak{D}_{(z,\overline{z}),(z,\overline{z})}(\tilde{F} \restriction U_z) = \begin{pmatrix} \xi + i\eta & 0 \\ 0 & \xi - i\eta \end{pmatrix}, \tag{viii}$$

und wir erhalten

$$\begin{aligned} &\mathfrak{D}_{C_z,C_z}(\tilde{F} \restriction U_z) = \mathfrak{T}_{C_z,(z,\overline{z})} \cdot \mathfrak{D}_{(z,\overline{z}),(z,\overline{z})}(\tilde{F} \restriction U_z) \cdot \overline{\mathfrak{T}_{C_z,(z,\overline{z})}}^t = \\ &\tfrac{1}{2}\sqrt{2}\begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \cdot \begin{pmatrix} \xi + i\eta & 0 \\ 0 & \xi - i\eta \end{pmatrix} \cdot \tfrac{1}{2}\sqrt{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = \\ &\tfrac{1}{2}\sqrt{2}\begin{pmatrix} \xi + i\eta & \xi - i\eta \\ \eta - i\xi & \eta + i\xi \end{pmatrix} \cdot \tfrac{1}{2}\sqrt{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = \\ &\tfrac{1}{2}\sqrt{2}\begin{pmatrix} 2\xi & -2\eta \\ 2\eta & 2\xi \end{pmatrix} \tfrac{1}{2}\sqrt{2} = \begin{pmatrix} \xi & -\eta \\ \eta & \xi \end{pmatrix}. \end{aligned} \tag{ix}$$

Weiter erhalten wir

$$\tilde{V} = U_1 \bigoplus \cdots \bigoplus U_{p+q} \bigoplus U_{z_1} \bigoplus \cdots \bigoplus U_{z_r}, \tag{x}$$

und für $C_z := (\sqrt{2}\Re z, \sqrt{2}\Im z)$ und $C := \{b_1, \ldots, b_{p+q}\} \cup C_{z_1} \cup \cdots \cup C_{z_r}$ ist $\tilde{C}$ eine Orthonormalbasis von $\tilde{V}$ mit

$$\mathfrak{D}_{\tilde{C},\tilde{C}}(\tilde{F}) = \begin{pmatrix} \boxed{\mathfrak{E}^{(p,p)}} & 0 & \cdots\cdots & & 0 \\ 0 & \boxed{-\mathfrak{E}^{(q,q)}} & \ddots & & \vdots \\ \vdots & \ddots & \boxed{\begin{matrix} \xi_1 & \eta_1 \\ -\eta_1 & \xi_1 \end{matrix}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots\cdots & & 0 & \boxed{\begin{matrix} \xi_r & \eta_r \\ -\eta_r & \xi_r \end{matrix}} \end{pmatrix}.$$

Mit Lemma 6.7.4 folgt $\mathfrak{D}_{C,C}(F) = \mathfrak{D}_{\tilde{C},\tilde{C}}(\tilde{F})$ und dies war die Behauptung. $\square$

6.11 Selbstadjungierte Endomorphismen

6.11.1 Definition Es sei (V, σ) ein euklidischer oder unitärer Vektorraum. Ein Endomorphismus $F \in \mathrm{End}(V)$ heißt *selbstadjungiert*, , wenn $F^\sigma = F$ ist. (Wir haben diese Terminologie schon vorher benutzt).
Eine Matrix $\mathfrak{A} \in \mathbb{R}^{n\times n}$ heißt *symmetrisch*, wenn $\mathfrak{A} = \mathfrak{A}^t$ ist.
Eine Matrix $\mathfrak{A} \in \mathbb{C}^{n\times n}$ heißt *hermitesch*, wenn $\mathfrak{A} = \overline{\mathfrak{A}}^t$ ist.

6.11.2 Lemma *Sei (V, σ) ein euklidischer oder unitärer Vektorraum und B eine Orthonormalbasis von (V, σ). Ein Endomorphismus $F \in \mathrm{GL}(V)$ ist genau dann selbstadjungiert in (V, σ), wenn seine darstellende Matrix $\mathfrak{D}_{B,B}(F)$ symmetrisch bzw. hermitesch ist.*

Beweis: Wegen $\mathfrak{D}_{B,B}(F^\sigma) = \mathfrak{D}_B(\sigma) \cdot \overline{\mathfrak{D}_{B,B}(F)}^t \cdot \overline{\mathfrak{D}_B(\sigma)}^t = \overline{\mathfrak{D}_{B,B}(F)}^t$ folgt dies unmittelbar. Im euklidischen Falle entfällt der Übergang zum konjungiert komplexen. $\square$

6.11.3 Satz *Jeder selbstadjungierte Endomorphismus eines unitären oder euklidischen Vektorraumes ist normal.*

Beweis: Dies ist klar wegen $F^\sigma F = FF = FF^\sigma$. $\square$

6.11.4 Satz *Ist F ein selbstadjungierter Endomorphismus eines euklidischen oder unitären Vektorraumes, so besitzt F nur reelle Eigenwerte. Insbesondere ist jeder selbstadjungierte Endomorphismus eines euklidischen Raumes split über $\mathbb{R}$.*

Beweis: Im euklidischen Fall setzen wir F zu $\tilde{F}$ fort. Dann ist $\tilde{F}$ split über $\mathbb{C}$. Da F normal ist, besitzt $\tilde{V}$ nach Satz 6.9.10 eine Orthonormalbasis B aus Eigenvektoren. Dann gilt nach (6.27) $\mathfrak{D}_{B,B}(F) = \mathfrak{D}_{B,B}(F^\sigma) = \overline{\mathfrak{D}_{B,B}(F)}^t$. Da $\mathfrak{D}_{B,B}(F)$ eine Diagonalmatrix mit den Eigenwerten von F in der Hauptdiagonalen ist, folgt $\overline{\lambda} = \lambda$ und damit $\lambda \in \mathbb{R}$ für jeden Eigenwert λ von F. Die gleiche Argumentation gilt auch im unitären Fall. Im euklidischen Fall folgt nun, daß F schon split über $\mathbb{R}$ ist. □

6.11.5 Satz *(Hauptachsentransformation einer symmetrischen Matrix) Ist $\mathfrak{A} \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix, so gibt es eine orthogonale Matrix $\mathfrak{S} \in O(n)$, so daß*

$$\mathfrak{S} \cdot \mathfrak{A} \cdot \mathfrak{S}^t = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \ldots, \lambda_n$ alle Eigenwerte von $\mathfrak{A}$ in ihrer Vielfachheit gezählt sind.

Beweis: Wir betrachten $\mathbb{R}^n$ mit dem kanonischen Skalarprodukt. Der durch $\mathfrak{A}$ induzierte Endomorphismus $\overline{\mathfrak{A}}$ ist dann selbstadjungiert und nach Satz 6.11.4 damit split über $\mathbb{R}$. Aus Satz 6.9.10 erhalten wir die Existenz einer Orthonormalbasis B aus Eigenvektoren von $\mathfrak{A}$. Dann hat $\mathfrak{D}_{B,B}(\overline{\mathfrak{A}})$ die verlangte Diagonalgestalt, und die Transformationsmatrix $\mathfrak{T}_{B,\mathbb{E}^n}$ ist nach Satz 6.10.1 orthogonal. □

6.11.6 Satz *(Hauptachsentransformation symmetrischer Bilinearformen) Es sei $\sigma: \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ eine symmetrische Bilinearform und $\mathfrak{A} = \mathfrak{D}_{\mathbb{E}^n}(\sigma)$. Dann gibt es eine Orthonormalbasis B von $\mathbb{R}^n$ derart, daß*

$$\mathfrak{D}_B(\sigma) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

ist, wobei $\lambda_1, \ldots, \lambda_n$ alle Eigenwerte von $\mathfrak{A}$ in ihrer Vielfachheit gezählt sind. Sind die Eigenwerte so geordnet, daß $\lambda_1, \ldots, \lambda_p > 0$, $\lambda_{p+1}, \ldots, \lambda_{p+q} < 0$ und $\lambda_{p+q+1} = \cdots = \lambda_n = 0$ sind, so gibt es eine Orthogonalbasis C, so daß

$$\mathfrak{D}_C(\sigma) = \begin{pmatrix} \mathfrak{E}^{(p,p)} & 0 & \cdots & \cdots & 0 \\ 0 & -\mathfrak{E}^{(q,q)} & \ddots & & \vdots \\ \vdots & \ddots & 0 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

ist.

Beweis: Die Matrix $\mathfrak{A}$ ist symmetrisch, da σ eine symmetrische Bilinearform ist. Wenden wir Satz 6.11.5 auf $\mathfrak{A}$ an, so transformiert sich die kanonische Basis zu einer Orthonormalbasis B, bezüglich der $\mathfrak{D}_B(\sigma)$ die geforderte Diagonalgestalt hat. Ist $B = \{b_1, \ldots, b_n\}$, so definieren wir

$$c_i := \begin{cases} |\lambda_i|^{-\frac{1}{2}} \cdot b_i & \text{für } i = 1, \ldots, p+q \\ b_i & \text{sonst.} \end{cases}$$

Dann ist $C := \{c_1, \ldots, c_n\}$ eine Orthogonalbasis, und es gilt $\sigma(c_i, c_j) = \frac{1}{|\lambda_i|}\sigma(b_i, b_j) = \frac{\lambda_i}{|\lambda_i|}\delta_{ij}$ für $i \leq p+q$, womit $\mathfrak{D}_C(\sigma)$ die behauptete Gestalt erhält. □

6.11.7 Bemerkung Die Tatsache, daß eine reelle symmetrische Matrix $\mathfrak{A}$ nur reelle Eigenwerte hat, ist zunächst erstaunlich und hat “analytische” Gründe. Wir arbeiten im $\mathbb{R}^n$ mit dem kanonischen Skalarprodukt und betrachten die quadratische Form

$$q_{\mathfrak{A}}(\mathfrak{x}) := \mathfrak{x}\mathfrak{A}\mathfrak{x}^t.$$

Diese definiert eine stetige Funktion $q_{\mathfrak{A}}\colon \mathbb{R}^n \longrightarrow \mathbb{R}$. Die n-Sphäre

$$S_n := \{\mathfrak{x} \in \mathbb{R}^n \mid \|\mathfrak{x}\| = 1\}$$

ist kompakt, und daher gibt es ein $\mathfrak{v} \in S_n$ mit $q_{\mathfrak{A}}(\mathfrak{v}) = \max\{q_{\mathfrak{A}}(\mathfrak{x}) \mid \mathfrak{x} \in S\}$. Wir behaupten, daß $\mathfrak{v}$ ein Eigenvektor von $\mathfrak{A}$ ist. Dazu betrachten wir den Raum $U := \langle\mathfrak{v}\rangle^\perp$. Gilt $\mathfrak{w} \perp \mathfrak{v}\mathfrak{A}$ für alle $\mathfrak{w} \in U$, so ist $\mathfrak{v}\mathfrak{A} \in \langle\mathfrak{v}\rangle$ und damit $\mathfrak{v}$ ein Eigenvektor. Sei also $\mathfrak{w} \in U$. Wir können $\mathfrak{w}$ normieren, so daß wir $\mathfrak{w} \in S_n$ annehmen dürfen. Nun wählen wir ein $\alpha \in [0,1]$ und $\beta := \sqrt{1-\alpha^2}$ und setzen

$$\mathfrak{u} := \alpha\mathfrak{v} + \beta\mathfrak{w}.$$

Da $\mathfrak{v}$ und $\mathfrak{w}$ senkrecht aufeinander stehen, folgt $\mathfrak{u} \in S_n$, und wir erhalten

$$q_{\mathfrak{A}}(\mathfrak{v}) \geq q_{\mathfrak{A}}(\mathfrak{u}) = \alpha^2 q_{\mathfrak{A}}(\mathfrak{v}) + \beta^2 q_{\mathfrak{A}}(\mathfrak{w}) + 2\alpha\beta\mathfrak{v}\mathfrak{A}\mathfrak{w}^t$$

und damit

$$\beta^2 q_{\mathfrak{A}}(\mathfrak{v}) \geq \beta^2 q_{\mathfrak{A}}(\mathfrak{w}) + 2\alpha\beta\mathfrak{v}\mathfrak{A}\mathfrak{w}^t,$$

woraus sich nach Division durch β

$$2\alpha\mathfrak{v}\mathfrak{A}\mathfrak{w}^t \leq \beta(q_{\mathfrak{A}}(\mathfrak{v}) - q_{\mathfrak{A}}(\mathfrak{w}))$$

ergibt. Sind $\mathfrak{v}\mathfrak{A}$ und $\mathfrak{w}$ nicht orthogonal, so dürfen wir, indem wir gegebenfalls $\mathfrak{w}$ durch $-\mathfrak{w}$ ersetzen, $\mathfrak{v}\mathfrak{A}\mathfrak{w}^t > 0$ annehmen. Setzen wir nun $\alpha = 1$ und damit $\beta = 0$, so ergibt sich der Widerspruch $\mathfrak{v}\mathfrak{A}\mathfrak{w}^t = 0$. Also waren $\mathfrak{v}\mathfrak{A}$ und $\mathfrak{w}$ orthogonal und $\mathfrak{v}$ ist ein Eigenvektor. Wegen $\mathfrak{v}\mathfrak{v}^t = 1$ folgt für den dazugehörigen Eigenwert λ schon $\mathfrak{v}\mathfrak{A}\mathfrak{v}^t = \mathfrak{v}\lambda\mathfrak{v}^t = \lambda\mathfrak{v}\mathfrak{v}^t = \lambda$. Damit ist $\mathfrak{v}\mathfrak{A}\mathfrak{v}^t = q_{\mathfrak{A}}(\mathfrak{v})$ der Eigenwert zum Eigenvektor v. Das Verfahren läßt sich nun iterieren, indem man die quadratische Form $q_{\mathfrak{A}}\restriction U$ betrachtet, diese nimmt auf der $(n-1)$-Sphäre wieder ihr Maximum an einem Eigenvektor an, und ein reeller Eigenwert läßt sich wie eben berechnen. So erhält man letztendlich, daß alle Eigenwerte reell sind. Dies liefert auch ein Verfahren zur numerischen Berechnung von Eigenwerten und Eigenvektoren.

6.11.8 Bemerkung Der Name *Hauptachsentransformation* erklärt sich aus der vorherigen Bemerkung. Durch die einer symmetrischen Bilinearform σ zugeordnete quadratische Form q_σ wird eine *Kurve* $\mathfrak{K} = \{x \in V \mid q_\sigma(x) = 1\}$ im euklidischen Raum $\mathbb{R}^n$ erklärt. Im $\mathbb{R}^2$ kann dies ein Ellipse oder Hyperbel, im $\mathbb{R}^3$ ein Ellipsoid oder Hyperboloid sein. Ist $\mathfrak{x}$ ein Vektor in $\mathfrak{K}$, der Eigenvektor zum Eigenwert λ der darstellenden Matrix von σ bezüglich der kanonischen Basis ist, so gilt $\mathfrak{x} = \alpha\mathfrak{v}$ für $\mathfrak{v} \in \mathrm{Eig}(\lambda, \mathfrak{A}) \cap S_n$ und damit $1 = q_\sigma(\mathfrak{x}) = q_\sigma(\alpha\mathfrak{v}) = \alpha^2 q_\sigma(\mathfrak{v}) = \alpha^2\lambda$ und damit $\|\mathfrak{x}\| = |\alpha| = \frac{1}{\sqrt{|\lambda|}}$. Die Eigenvektoren der Matrix, welche die quadratische Form bezüglich der kanonischen Basis darstellt, zeigen dann, wie in 6.11.7 gezeigt, in die Richtung der Hauptachsen der quadratischen Kurven.

Wir wollen dies an einem Beispiel im $\mathbb{R}^2$ klarmachen. Sei $\sigma\colon \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ eine Bilinearform mit darstellender Matrix $\mathfrak{D}_{\mathbb{E}^2}(\sigma) = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$. Für $\mathfrak{x} = (x_1, x_2)$ und $\mathfrak{y} = (y_1, y_2)$ ist dann

$$\sigma(\mathfrak{x}, \mathfrak{y}) = \alpha(x_1y_1 + x_2y_2) + \beta(x_1y_2 + x_2y_1).$$

Das charakteristische Polynom von $\mathfrak{D}_{\mathbb{E}^2}(\sigma)$ erhalten wir zu

$$Charpol_{\mathfrak{D}_{\mathbb{E}^2}(\sigma)} = \det\begin{pmatrix} \alpha - x & \beta \\ \beta & \alpha - x \end{pmatrix} = x^2 - 2\alpha x + (\alpha^2 - \beta^2).$$

Daraus errechnen sich die Eigenwerte

$$\lambda_1 = \alpha + \sqrt{\alpha^2 - \alpha^2 + \beta^2} = \alpha + \beta \quad \text{und} \quad \lambda_2 = \alpha - \beta$$

und die dazugehörigen normierten Eigenvektoren

$$\mathfrak{v}_1 = \frac{1}{2}\sqrt{2}(1,1) \quad \text{und} \quad \mathfrak{v}_2 = \frac{1}{2}\sqrt{2}(-1,1).$$

Die zu σ gehörende quadratische Form $q_\sigma(\mathfrak{x}) := \sigma(\mathfrak{x},\mathfrak{x})$ definiert eine *quadratische Kurve*

$$\mathfrak{K} := \{\mathfrak{x} \in \mathbb{R}^2 \mid q_\sigma(\mathfrak{x}) = 1\}.$$

Ist $\mathfrak{x} = (x_1, x_2)$, so erhalten wir $q_\sigma(\mathfrak{x}) = \alpha({x_1}^2 + {x_2}^2) + 2\beta x_1 x_2$. Ist $B = (\mathfrak{v}_1, \mathfrak{v}_2)$ die Orthonormalbasis aus Eigenvektoren, so erhalten wir

$$\mathfrak{D}_B(\sigma) = \begin{pmatrix} \alpha+\beta & 0 \\ 0 & \alpha-\beta \end{pmatrix}$$

und damit für $h_B(\mathfrak{x}) = (z_1, z_2)$

$$q_\sigma(\mathfrak{x}) = (\alpha+\beta)z_1^2 + (\alpha-\beta)z_2^2 = \lambda_1 z_1^2 + \lambda_2 z_2^2. \tag{6.34}$$

Die gemischten Glieder sind also verschwunden. Um diese Tatsache anschaulich darzustellen, nehmen wir $\lambda_1 > 0$ und $\lambda_2 \neq 0$ an und setzen $\xi := \frac{1}{\sqrt{\lambda_1}}$ und $\eta := \frac{1}{\sqrt{|\lambda_2|}}$. Dann erhält (6.34) die Form

$$q_\sigma(\mathfrak{x}) = \left(\frac{z_1}{\xi}\right)^2 \pm \left(\frac{z_2}{\eta}\right)^2.$$

Im Falle von $+$ ist die Kurve $\mathfrak{K}$ eine Ellipse und im Falle von $-$ eine Hyperbel, deren Hauptachsen $\xi \cdot \mathfrak{v}_1$ und $\eta \cdot \mathfrak{v}_2$ sind. Die Hauptachsentransformation bewirkt also eine Drehung des Koordinatensystems in Richtung der Hauptachsen.

Index

Symbole

Stichworte