Lineare Algebra

mit einem Ausblick auf die Geometrie

Robert Wisbauer

3. Mai 2005

Inhaltsverzeichnis

Vorwort

Die Lineare Algebra ist für alle Gebiete der Mathematik von großer Bedeutung. Sie steht daher aus gutem Grund mit am Anfang jeder Ausbildung in Mathematik. An dieser Stelle fällt es ihr aber auch zu, die Studierenden mit Überlegungen zu den Grundlagen der Mathematik vertraut zu machen. Dazu erscheint es mir unerläßlich, sich etwas mit den Grundideen der Mengenlehre zu befassen.

Zentraler Gegenstand der Untersuchungen in der Linearen Algebra sind die *Moduln* (*lineare Räume*) über Ringen. Moduln über Körpern werden *Vektorräume* genannt. Die verbreitete Gepflogenheit, die Einführung in die Lineare Algebra auf die Behandlung von Vektorräumen zu beschränken, halte ich weder sachlich noch didaktisch für vorteilhaft. Das bei Bedarf dann nachgeschobene Argument, daß einige der Sätze über Vektorräume auch für Moduln über Ringen gelten würden, ist irritierend und insbesondere für Anfänger eine Zumutung.

Der Anwendungsbereich der zu entwickelnden Theorie ist so weit (z.B. Biologie, Wirtschaft, Informatik), daß eine zu starke Ausrichtung der Grundlagen auf geometrische Vorstellbarkeit hin für das Verständnis der Allgemeinheit der Begriffe durchaus hinderlich sein kann. Als *zusätzliche* Quelle für Beispiele ist die Geometrie der Ebene und des Raumes jedoch gelegentlich hilfreich.

In der vorliegenden Einführung werden am Anfang die Grundlagen der Mengenlehre vorgestellt und darauf aufbauend die algebraischen Strukturen eingeführt. Moduln und Matrizenrechnung werden zunächst über beliebigen Ringen betrachtet. Zielsetzung bleibt dabei allerdings die Darstellung der allgemeinen Grundlagen und Zusammenhänge. Der Versuchung, auf dieser Basis *Modultheorie* im eigentlichen Sinne zu betreiben (z.B. die Erforschung der Beziehung zwischen Eigenschaften eines Ringes und seiner Moduln), wird nicht nachgegeben.

Das Fundament, das gelegt wird, soll solide genug sein, um auch bei späterer Anwendung als verläßliche Bezugsquelle zu dienen.

Ich möchte allen, die an der Fertigstellung dieses Buches mitgewirkt haben, ganz herzlich dafür danken.

Düsseldorf, im Juli 1994

Robert Wisbauer

Vorbetrachtung

Aufgabe und Ziel dieses Kurses ist es, eine Einführung in die mathematische Denk- und Arbeitsweise zu geben und zugleich mit den Methoden der linearen Algebra vertraut zu machen.

Dabei soll der Aufbau einer mathematischen Theorie auch exemplarisch geübt werden. Das dabei zu errichtende Begriffsgebäude, nämlich die lineare Algebra, wird sich als äußerst nützlich in vielen Teilen der Mathematik und deren Anwendungen erweisen. Deshalb wird das Fundament in der später nötigen Allgemeinheit gelegt.

Diese Zielsetzung ist nicht ganz problemlos. Eine (zu) große Allgemeinheit und Abstraktheit macht Anfängern erfahrungsgemäß Schwierigkeiten. Andererseits hat die übermäßige Fixierung auf ein Beispiel auch ihre Tücken. Die Besonderheiten des speziellen Falles können den Blick von der Allgemeingültigkeit eines Konzepts ablenken.

Im folgenden wird dem dadurch Rechnung getragen, daß die Grundbegriffe abstrakt formuliert und dann an mehreren Beispielen veranschaulicht werden.

Um eine gewisse Vorstellung von dem zu geben, was später gemacht wird, betrachten wir einige algebraische Strukturen, die allen gut bekannt sind. Dies gibt uns Gelegenheit, gleich einige Notationen festzulegen:

$\mathbb{N}$	die natürlichen Zahlen $\{0, 1, 2, 3, \ldots\}$
$\mathbb{Z}$	die ganzen Zahlen $\{0, \pm 1, \pm 2, \ldots\}$
$\mathbb{Q}$	die rationalen Zahlen $\{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$
$\mathbb{R}$	die reellen Zahlen
$\mathbb{C}$	die komplexen Zahlen

In $\mathbb{N}$ haben wir zum Beispiel die Verknüpfungen $+$ und $\cdot$, das sind Abbildungen

$$\begin{aligned} + &: \mathbb{N} \times \mathbb{N} \to \mathbb{N}, \quad (a, b) \mapsto a + b\,, \\ \cdot &: \mathbb{N} \times \mathbb{N} \to \mathbb{N}, \quad (a, b) \mapsto a \cdot b\,, \end{aligned}$$

mit den Gesetzmäßigkeiten (für alle a, b, $c \in \mathbb{N}$):

$$\begin{aligned} (a + b) + c &= a + (b + c) \quad \text{Assoziativgesetze} \\ (a \cdot b) \cdot c &= a \cdot (b \cdot c) \\ a + b &= b + a \quad \text{Kommutativgesetze} \\ a \cdot b &= b \cdot a \end{aligned}$$

Im Zusammenwirken der beiden Verknüpfungen gilt:

$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c \quad \text{Distributivgesetze} \\ (a + b) \cdot c &= a \cdot c + b \cdot c \end{aligned}$$

Bezüglich Multiplikation und Addition gibt es jeweils ein Element, das die anderen Elemente nicht verändert:

$$\begin{aligned} 1 \cdot a = a &\quad \text{neutrales Element (bzgl. } \cdot \text{)} \\ 0 + a = a &\quad \text{neutrales Element (bzgl. } + \text{)} \end{aligned}$$

In $\mathbb{Z}$ können wir zu jedem $a \in \mathbb{Z}$ ein Element b finden mit

$$a + b = 0 \qquad \text{inverses Element bzgl. } + \ .$$

In $\mathbb{Q}$ gibt es zu jedem Element $a \neq 0$ ein b mit

$$a \cdot b = 1 \qquad \text{inverses Element bzgl. } \cdot \ .$$

Diese Eigenschaften haben auch $\mathbb{R}$ und $\mathbb{C}$. Man beachte, daß bei den oben herausgestellten Beziehungen z.B. die Anordnung der Elemente der betrachteten Mengen keine Rolle spielt. Die Unterschiede zwischen $\mathbb{Q}$ und $\mathbb{R}$ bzw. $\mathbb{C}$ beziehen sich nicht auf die bisher erwähnten Gesetzmäßigkeiten, sondern sind anderer Art. So ist etwa in $\mathbb{R}$ die Gleichung $x^2 = a$ für $a \in \mathbb{R}$ mit $a > 0$ lösbar. In $\mathbb{C}$ ist dies sogar für beliebige $a \in \mathbb{C}$ der Fall.

Als weitere Besonderheit sei auf das Zusammenwirken von verschiedenen Bereichen hingewiesen. Wenn man zum Beispiel die Elemente aus $\mathbb{Q}$ mit Elementen aus $\mathbb{Z}$ multipliziert, so ergibt sich wieder ein Element aus $\mathbb{Q}$, d.h. wir haben eine Abbildung

$$(\mathbb{Z}, \mathbb{Q}) \to \mathbb{Q},\ (z, g) \mapsto z \cdot g,$$

mit den Eigenschaften $(z, z_i \in \mathbb{Z},\ q, q_i \in \mathbb{Q},\ i = 1, 2)$

$$\begin{aligned} (z_1 + z_2)q &= z_1 q + z_2 q \\ z_1(z_2 q) &= (z_1 z_2)q \\ z(q_1 + q_2) &= z q_1 + z q_2 \\ 1z &= z\,. \end{aligned} \qquad (*)$$

Ähnliches läßt sich auch für die Paare $(\mathbb{Z}, \mathbb{R})$, $(\mathbb{Q}, \mathbb{R})$, $(\mathbb{Q}, \mathbb{C})$ und viele andere beobachten.

Die Algebra beschäftigt sich ganz allgemein mit Verknüpfungen auf irgendwelchen Mengen M, also Abbildungen

$$\tau : M \times M \to M, \quad (m, n) \mapsto m \tau n,$$

die gewissen Gesetzmäßigkeiten genügen. Die oben formulierten spielen dabei eine herausragende Rolle. So werden wir (M, τ) eine *Halbgruppe* nennen, wenn für τ das Assoziativgesetz gilt. Gilt zudem das Kommutativgesetz, so spricht man von einer *kommutativen Halbgruppe*. $\mathbb{N}$ ist also sowohl für $+$, als auch für $\cdot$ eine kommutative Halbgruppe.

Gibt es für $\tau : M \times M \to M$ ein neutrales Element und zu jedem $m \in M$ ein Inverses, so nennt man (M, τ) eine *Gruppe*.

Mengen M mit zwei Verknüpfungen $+$ und $\cdot$ nennt man *Ringe*, wenn $(M, +)$ eine kommutative Gruppe und $(M, \cdot)$ eine Halbgruppe mit neutralem Element ist und zudem die Distributivgesetze gelten. Somit sind $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ Ringe.

Ist in einem Ring jedes von Null veschiedene Element invertierbar, so spricht man von einem *Divisionsring* oder *Schiefkörper*. Ist zudem die Multiplikation kommutativ, so hat man einen *Körper*. $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ sind Körper.

Schließlich noch ein Blick auf die Situation, in der zwei algebraische Systeme verbunden sind.

Sei $(R, +, \cdot)$ ein Ring und $(M, +)$ eine kommutative Gruppe. Gibt es eine Abbildung

$$(R, M) \to M,\ (r, m) \mapsto rm,$$

welche die Bedingungen $(*)$ erfüllt, so nennt man das Paar (R, M) einen *R-Modul*. Ist R ein Körper, so werden R-Moduln auch als *Vektorräume* (oder *lineare Räume*) bezeichnet.

Damit haben wir die wichtigsten Grundbegriffe der linearen Algebra angesprochen. Wir haben uns dabei der Schreibweisen bedient, die aus der Schule geläufig sind.

Um damit in allgemeinen Situationen arbeiten zu können, müssen wir Begriffe wie „Paar von Mengen“ oder „Abbildungen“ präziser fassen. Dies gehört zur Problemstellung der Mengenlehre und soll im ersten Kapitel geschehen.

Die algebraischen Strukturen wie Gruppen, Ringe und Körper werden im zweiten Kapitel eingeführt und ihre grundlegenden Eigenschaften herausgestellt, soweit sie für die hier vorgesehene Anwendung relevant sind.

Im dritten Kapitel kommen wir zu den Moduln, deren Untersuchung Gegenstand der Linearen Algebra ist. Die weiteren Kapitel befassen sich dann mit zusätzlichen Eigenschaften von Moduln und Vektorräumen sowie mit Anwendungen der aufgebauten Theorie.

Als Beispiele für Moduln und Vektorräume werden wir schon früh den zwei- und dreidimensionalen (euklidischen) Raum über $\mathbb{R}$ heranziehen. Damit lassen sich einige abstrakte Begriffe geometrisch veranschaulichen. Genaueres über die Wechselbeziehung zwischen Geometrie und Linearer Algebra werden wir erst im letzten Kapitel erfahren.

Kapitel 1

Mengentheoretische Grundlagen

Ziel dieses Kapitels ist es, die mengentheoretischen Begriffe und Techniken bereitzustellen, die wir beim Aufbau der (Linearen) Algebra benötigen.

Auch wenn wir eine gewisse Vertrautheit im formalen Umgang mit Mengen voraussetzen, so sollen doch die Grundtatsachen der Mengenlehre festgehalten werden, die wir als gegeben ansehen wollen (Axiome). An der Entwicklung dieser Theorie zu Beginn unseres Jahrhunderts waren zum Beispiel die Mathematiker G. Cantor, E. Zermelo und A. Fraenkel maßgeblich beteiligt.

Das Überprüfen von eventuellen logischen Abhängigkeiten oder der Vollständigkeit eines solchen Axiomensystems ist ein nicht-triviales Problem der Mengenlehre, auf das wir hier nicht eingehen können.

Beim ersten Durchlesen wird vielleicht der tiefere Sinn oder die Zweckmäßigkeit der Formulierungen nicht gleich erkennbar sein. Mit zunehmender Erfahrung im Umgang mit diesen Begriffen wird jedoch das Verständnis dafür wachsen.

1 Axiome der Mengenlehre

Eine *Menge* setzt sich aus ihren Elementen zusammen. Der grundlegende Begriff der Mengenlehre ist die Element-Beziehung: Für jedes Objekt a muß sich feststellen lassen, ob es zu einer gegebenen Menge A gehört oder nicht. Es gilt also

a ist Element von A (a ist enthalten in A, $a \in A$), oder

a ist nicht Element von A ($a \notin A$).

Mit dieser Beziehung soll festgestellt werden können, ob zwei Mengen gleich sind. Dies geschieht in Form des folgenden Postulats:

1.1 Extensionalitätsaxiom
Zwei Mengen A, B sind genau dann gleich, wenn sie dieselben Elemente haben, also $A = B$ genau dann, wenn

$$x \in A \Rightarrow x \in B \quad \text{und} \quad x \in B \Rightarrow x \in A.$$

Damit sind Mengen eindeutig durch ihre Elemente bestimmt.

Der Pfeil $P \Rightarrow Q$ zwischen zwei Aussagen P, Q bedeutet dabei die logische Implikation *falls P gilt, dann gilt auch Q*; man sagt dazu auch *aus P folgt Q* oder *P impliziert Q*.

Die logische Äquivalenz von P und Q wird mit $P \Leftrightarrow Q$ bezeichnet. Die obige Aussage könnte also auch so formuliert werden:

$$A = B \text{ genau dann, wenn } [x \in A \Leftrightarrow x \in B].$$

Da wir die Zugehörigkeit zu einer Menge als entscheidbar annehmen, können wir nun festlegen:

1.2 Definition
Eine Menge B heißt *Teilmenge* einer Menge A, wenn jedes Element aus B auch Element von A ist, d.h.

$$B \subset A \text{ genau dann, wenn } x \in B \Rightarrow x \in A.$$

Man beachte, daß in dieser Notation $B \subset A$ auch für $B = A$ gilt.

Als nächste Grundforderung wird verlangt, daß man durch Eigenschaften von Elementen Teilmengen aussondern kann.

1.3 Aussonderungsaxiom
Zu jeder Menge A und jeder Eigenschaft P (die ein Element von A haben kann) gibt es eine Teilmenge B von A, die gerade aus den Elementen von A mit dieser Eigenschaft besteht:

$$B = \{x \in A \mid P(x)\} \subset A.$$

Bislang haben wir zwar von Eigenschaften von Mengen gesprochen, doch wissen wir noch nicht, ob es überhaupt Mengen gibt. Dies wollen wir natürlich haben und fordern daher als Axiom, daß es (mindestens) eine Menge gibt.

1.4 Existenz der leeren Menge
Es gibt eine Menge, die keine Elemente enthält.
Man nennt diese die leere Menge *und bezeichnet sie mit $\emptyset$.*

Falls es überhaupt eine Menge A gibt, so folgt aus dem Aussonderungsaxiom die Existenz der leeren Menge als

$$\emptyset = \{x \in A \mid x \neq x\}.$$

Nach Defintion ist $\emptyset$ Teilmenge jeder Menge A, also $\emptyset \subset A$. So gilt auch $\emptyset \subset \emptyset$, aber $\emptyset \notin \emptyset$.

Eine strengeres Fundament für die Mengenlehre war notwendig geworden, als man um die Jahrhundertwende erkannte, daß man mit den bis dahin als zulässig angesehenen Bildungen zu widersprüchlichen Ergebnissen gelangen konnte. So war damals die Annahme zulässig, daß es eine Menge gibt, die alle anderen Mengen enthält (*Allmenge*). Durch Eigenschaften von Elementen sollten dann auch Teilmengen davon festgelegt sein. Der britische Philosoph und Mathematiker B. Russell machte (im Jahre 1902) darauf aufmerksam, daß die (erlaubte) Bildung der Menge A aller Mengen, die nicht Element von sich selbst sind, also

$$A := \{x \mid x \notin x\},$$

nicht sinnvoll ist (*Russellsche Paradoxie*). Man sieht leicht, daß weder die Aussage $A \in A$ noch $A \notin A$ gelten kann.

Es ist eine Konsequenz des Aussonderungsaxioms, daß die Bildung einer Allmenge in unserem Rahmen nicht zulässig ist:

1.5 Satz
Es gibt keine Menge von Mengen, die jede Menge als Element enthält.

Beweis: Es ist zu zeigen, daß es zu jeder Menge A von Mengen eine Menge B gibt, die nicht Element von A ist. Dazu betrachten wir

$$B = \{x \mid x \in A, x \notin x\}.$$

Angenommen $B \in A$. Es gilt $B \in B$ oder $B \notin B$.
- Aus $B \in B$ folgt $B \notin B$, nach Definition von B.
- Aus $B \notin B$ folgt $B \in B$, ebenfalls nach Definition von B.

Dies sind Widersprüche, und somit muß $B \notin A$ gelten. □

Das nächste Axiom fordert, daß man aus *vorgegebenen* Mengen eine Menge bilden kann, die jede dieser Mengen als Teilmenge enthält. Aus sprachlichen Gründen nennen wir eine *Menge von Mengen* auch ein *Mengensystem*.

1.6 Vereinigungsaxiom
Zu jedem Mengensystem $\mathcal{M}$ gibt es eine Menge, welche genau alle Elemente enthält, die zu mindestens einer Menge des gegebenen Systems gehören:

$$\bigcup_{A \in \mathcal{M}} = \{x \mid \textit{es gibt ein } A \in \mathcal{M} \textit{ mit } x \in A\}.$$

Speziell ergibt dies für zwei Mengen A und B die Vereinigung

$$A \cup B = \{x \mid x \in A \text{ oder } x \in B\}.$$

Damit kann man zu zwei Elementen x, y die *Paarmenge* bilden:

$$\{x, y\} = \{x\} \cup \{y\}$$

Aus der Kommutativität von $\cup$ (d.h. $A \cup B = B \cup A$) folgt, daß

$$\{x, y\} = \{y, x\} \qquad \text{(ungeordnetes Paar)}.$$

Will man die Reihenfolge von zwei Elementen berücksichtigen, so kann man dies mit der von K. Kuratowski vorgeschlagenen

1.7 Definition
Als *geordnetes Paar* von Elementen $x, y \in A$ bezeichnet man

$$(x, y) = \{\{x\}, \{x, y\}\}.$$

Für die Menge aller solcher Paare schreibt man

$$A \times A = \{(a, b) \mid a, b \in A\}.$$

Dabei kommt es wirklich auf die Reihenfolge der Elemente an:

1.8 Satz
Seien A eine Menge und $x, y, u, v \in A$. Dann gilt $(x, y) = (u, v)$ genau dann, wenn $x = u$ und $y = v$.

Beweis: $\Leftarrow$ ist klar.
$\Rightarrow$: Nehmen wir an, daß $\{\{x\}, \{x, y\}\} = \{\{u\}, \{u, v\}\}$. Dann gilt

$$\{x\} \in \{\{u\}, \{u, v\}\}, \quad \{x, y\} \in \{\{u\}, \{u, v\}\},$$

also $\{x\} = \{u\}$ oder $\{x\} = \{u, v\}$ und $\{x, y\} = \{u\}$ oder $\{x, y\} = \{u, v\}$.

Angenommen $\{x\} = \{u, v\}$. Dann gilt $x = u = v$ und damit auch $x = y$. Damit ist dann auch die Behauptung des Satzes gezeigt.

Angenommen $\{x\} = \{u\}$, also $x = u$.
$\{x, y\} = \{u\}$ ergibt wieder $x = u = y = v$, also obigen Fall.
$\{x, y\} = \{u\}$ hat $y = v$ zur Folge. Also ist die Behauptung des Satzes ebenfalls erfüllt. □

Obige Bildung kann auch auf die Vereinigung von zwei Mengen angewendet werden. Dies ermöglicht die Formulierung eines Begriffs, der sich als sehr nützlich erweisen wird:

1.9 Definition
Als *geordnetes Paar* von Mengen A, B bezeichnet man

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Als weitere Folgerung aus dem Vereinigungsaxiom ergibt sich mit Hilfe des Aussonderungsaxioms die Existenz des *Durchschnitts* von Mengen:

1.10 Satz
Zu jedem Mengensystem $\mathcal{M}$ gibt es genau eine Menge, welche genau diejenigen Elemente enthält, die in jeder Menge aus $\mathcal{M}$ enthalten sind:

$$\bigcap_{M\in\mathcal{M}} M := \{x \in \bigcup_{M\in\mathcal{M}} M \mid x \in M \text{ für alle } M \in \mathcal{M}\}.$$

Man bezeichnet diese Menge als den Durchschnitt *der Mengen aus $\mathcal{M}$.*
Speziell für zwei Mengen A und B bedeutet dies

$$A \cap B = \{x \mid x \in A \text{ und } x \in B\}.$$

Wir wollen einige Eigenschaften von Durchschnitt und Vereinigung zusammenstellen, die sich unmittelbar aus den Definitionen ergeben:

Eigenschaften
A, B und C seien Teilmengen einer Menge I. Dann gilt:

(1) $A \cap A = A$, $A \cup A = A$ (Idempotenz);

(2) $A \cap B = B \cap A$, $A \cup B = B \cup A$ (Kommutativität);

(3) $A \cap (B \cap C) = (A \cap B) \cap C$,
$A \cup (B \cup C) = (A \cup B) \cup C$ (Assoziativität);

(4) $A \cap (A \cup B) = A$, $A \cup (A \cap B) = A$ (Absorption);

(5) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (Distributivität);

(6) $A \cap B = A \Leftrightarrow A \subset B \Leftrightarrow A \cup B = B$ (Konsistenz).

Eine weitere mengentheoretische Bildung, die durch das Aussonderungsaxiom ermöglicht wird, ist die *Differenzmenge*:

1.11 Definition
Sind A und B Mengen, so heißt

$$A \setminus B = \{x \in A \mid x \notin B\}$$

die *Differenzmenge* zwischen A und B.
Gilt $B \subset A$, so nennt man $A \setminus B$ das *Komplement* von B in A.

Als Zusammenhang zwischen $\cap$, $\cup$ und $\setminus$ läßt sich etwa für Mengen A, B, C als leichte Übung zeigen:

$$\begin{aligned} A \setminus (A \setminus B) &= A \cap B \\ A \cap (B \setminus C) &= (A \cap B) \setminus (A \cap C) \\ A \setminus (B \cup C) &= (A \setminus B) \cap (A \setminus C). \end{aligned}$$

Im nächsten Axiom wird verlangt, daß die Gesamtheit der Untermengen einer Menge wieder eine Menge (ein Mengensystem) bildet:

1.12 Potenzmengenaxiom
Zu jeder Menge A gibt es eine Menge, deren Elemente die Teilmengen von A sind. Man nennt sie die Potenzmenge von A, also

$$\mathcal{P}(A) = \{U \mid U \subset A\}.$$

Formale Folgerungen aus dieser Festlegung sind für Mengen A, B:

(i) $B \subset A \Leftrightarrow B \in \mathcal{P}(A)$

(ii) $B \subset A \Leftrightarrow \mathcal{P}(B) \subset \mathcal{P}(A)$

(iii) $\emptyset \in \mathcal{P}(A)$, $A \in \mathcal{P}(A)$.

Wir haben zwar schon die Existenz von Mengen gefordert, wissen aber noch nicht, ob es Mengen mit unendlich vielen Elementen gibt. Dies müssen wir durch weitere Forderungen sicherstellen.

1.13 Definition
Als *Nachfolger* einer Menge A bezeichnen wir die Menge

$$A^+ = A \cup \{A\}.$$

Eine Menge von Mengen A heißt *induktiv*, wenn $\emptyset \in A$ und zu jedem Element aus A auch sein Nachfolger zu A gehört.

1.14 Unendlichkeitsaxiom
Es gibt eine induktive Menge.

Mit den bisher festgelegten Axiomen haben wir die Möglichkeit, ein Modell für die natürlichen Zahlen $\mathbb{N}$ anzugeben. Deren Existenz haben wir zwar ohnehin geglaubt, wenn wir aber unsere weitere Theorie nur auf vorgegebene Sachverhalte stützen wollen, so muß auch $\mathbb{N}$ seinen Platz in diesem System haben.

Man kann die Menge der natürlichen Zahlen nun als minimale induktive Menge definieren und etwa durch eine Menge von Mengen repräsentieren:

$$\begin{aligned} 0 &= \emptyset \\ 1 &= \emptyset \cup \{\emptyset\} = \{\emptyset\} \\ 2 &= \{0,1\} = \{\emptyset, \{\emptyset\}\} \\ 3 &= \{0,1,2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \\ &\vdots \\ n+1 &= n^+ . \end{aligned}$$

Basierend auf den Festlegungen

$$m + 0 = m, \quad m + n^+ = (m + n)^+$$

kann dann auch die Arithmetik von $\mathbb{N}$ gewonnen werden. Eine genauere Ausführung dazu findet man z.B. in dem Buch von Enderton zur Mengenlehre.

Da hierbei die natürlichen Zahlen als minimale induktive Menge festgelegt wurden, ergibt sich die

Induktionseigenschaft: Ist $A \subset \mathbb{N}$ eine Teilmenge mit

(i) $0 \in A$ und

(ii) $n \in A \Rightarrow n + 1 \in A$,

dann ist $A = \mathbb{N}$.

Dies sind natürlich nur Andeutungen zur Einführung der natürlichen Zahlen, um klar zu machen, in welchem Kontext man sie realisieren kann. Wir werden die allgemein vertrauten Eigenschaften von $\mathbb{N}$ ohne weitere Rechtfertigung benutzen.

Ausgehend von $\mathbb{N}$, können mit algebraischen Methoden $\mathbb{Z}$ und $\mathbb{Q}$ konstruiert werden. Um die rationalen Zahlen $\mathbb{Q}$ zu den reellen Zahlen $\mathbb{R}$ zu erweitern, benötigt man topologische Überlegungen.

Nun kommen wir zu einer Forderung, deren Gültigkeit man ohne weiteres akzeptiert, die jedoch nicht aus unseren bisherigen Axiomen gefolgert werden kann.

Ist ein System nicht-leerer Mengen gegeben, so soll es möglich sein, aus jeder Menge genau ein Element herauszugreifen, und diese herausgenommenen Elemente zu einer neuen Menge zusammenzufassen. Setzen wir voraus, daß die Mengen des gegebenen Mengensystems paarweise disjunkt sind, so läßt sich dies folgenderweise formulieren:

1.15 Auswahlaxiom
Sei $\mathcal{M}$ ein Mengensystem nicht-leerer Mengen und $A \cap B = \emptyset$ für alle $A, B \in \mathcal{M}$ mit $A \neq B$. Dann gibt es eine Menge M, so daß für jedes $A \in \mathcal{M}$ die Menge $A \cap M$ genau ein Element enthält.

Für die Bedeutung des Auswahlaxioms werden wir später mehr Verständnis gewinnen. Man kann auch *ohne* das Auswahlaxiom Mengenlehre und Mathematik betreiben, doch werden wir es heranziehen, wenn wir es brauchen (etwa zum Beweis der Existenz einer Basis in einem Vektorraum).

Mit den angeführten Axiomen und den ersten Folgerungen daraus können wir die mengentheoretischen Begriffsbildungen begründen, die wir für die Algebra benötigen. Sie sind jedoch nicht für alle Probleme der Mengenlehre und Mathematik ausreichend. Für spezielle Konstruktionen können und müssen weitere Axiome dazugenommen werden.

1.16 Aufgaben

(1) Für die Teilmengen der reellen Zahlen
$A_1 := \{x \in \mathbb{R} \mid 0 \leq x < 2\}$,
$A_2 := \{x \in \mathbb{R} \mid 0 < x \leq 2\}$,
$A_3 := \{x \in \mathbb{R} \mid -1 \leq x < 1\}$
bestimme man

(a) $A_1 \cup A_2$, *(b)* $A_1 \cup A_3$, *(c)* $\bigcup_{i=1}^{3} A_i$,

(d) $A_1 \cap A_2$, *(e)* $A_1 \cap A_3$, *(f)* $\bigcap_{i=1}^{3} A_i$,

(g) $A_1 \setminus A_2$, *(h)* $A_2 \setminus A_1$, *(i)* $A_1 \setminus A_3$.

(2) Es seien $A := \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ und $B := \{x \in \mathbb{R} \mid 0 \leq x \leq 2\}$.
Wie kann man $A \times B$ und $B \times A$ geometrisch deuten?
Gilt $A \times B = B \times A$?

(3) A, B, C seien Mengen mit $A \cap B = A \cap C$ und $A \cup B = A \cup C$.
Man folgere $B = C$.

(4) I sei eine beliebige Menge, A und B seien Teilmengen von I. Man nennt $\mathcal{C}(A) := I \setminus A$ das Komplement von A (bzgl. I). *Man zeige:*

(a) $A \cap \mathcal{C}(A) = \emptyset$, $A \cup \mathcal{C}(A) = I$;

(b) $\mathcal{C}(A \cap B) = \mathcal{C}(A) \cup \mathcal{C}(B)$;

(c) $\mathcal{C}(A \cup B) = \mathcal{C}(A) \cap \mathcal{C}(B)$.

(5) Zeigen Sie, daß für zwei Mengen A, B gilt:

(i) $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$

(ii) $\mathcal{P}(A \cup B) \supset \mathcal{P}(A) \cup \mathcal{P}(B)$. *Wann gilt dabei die Gleichheit?*

(6) Zeigen Sie, daß für Mengen A, B, C, D gilt:

(i) $A \times B \cap C \times D = (A \cap C) \times (B \cap D)$;

(ii) $A \times B \cup C \times D \subset (A \cup C) \times (B \cup D)$;

(iii) *Sind A und B nicht leer, so gilt:*
$A \times B \subset C \times D \Rightarrow A \subset C$ *und* $B \subset D$.
Was passiert, wenn A oder B leer sind?

2 Relationen

Der Begriff einer allgemeinen *Relation* zwischen zwei Mengen A, B geht vom geordneten Paar $A \times B$ aus. Die dazu angegebenen Bildungen mögen auf den ersten Blick etwas ungewohnt erscheinen. Ihre Bedeutung wird jedoch bei der Behandlung spezieller Relationen in den nachfolgenden Abschnitten klar.

2.1 Definition
Seien A und B Mengen. Eine Teilmenge $R \subset A \times B$ nennen wir eine *Relation zwischen A und B*.

Speziell heißt eine Teilmenge $R \subset A \times A$ eine *Relation auf A*.

Man sagt $x \in A$ und $y \in B$ *stehen in Relation* R, wenn $(x, y) \in R$, und schreibt dafür xRy.

Als *Definitionsbereich* bzw. *Wertebereich* von R bezeichnen wir

$$\begin{aligned} \mathbb{D}(R) &= \{x \in A \mid \text{es gibt ein } y \in B \text{ mit } (x, y) \in R\}, \\ \mathbb{W}(R) &= \{y \in B \mid \text{es gibt ein } x \in A \text{ mit } (x, y) \in R\}. \end{aligned}$$

$\mathbb{D}(R)$ ist die Menge aller ersten Komponenten der Elemente in R, $\mathbb{W}(R)$ die Menge der zweiten Komponenten.

Nach diesen Definitionen gilt $R \subset \mathbb{D}(R) \times \mathbb{W}(R)$, d.h. R ist auch eine Relation zwischen $\mathbb{D}(R)$ und $\mathbb{W}(R)$. Außerdem ist R auch eine Relation auf $\mathbb{D}(R) \cup \mathbb{W}(R)$, denn

$$R \subset \mathbb{D}(R) \times \mathbb{W}(R) \subset (\mathbb{D}(R) \cup \mathbb{W}(R)) \times (\mathbb{D}(R) \cup \mathbb{W}(R)).$$

Man beachte, daß im allgemeinen nicht $R = \mathbb{D}(R) \times \mathbb{W}(R)$ gelten wird.

Zu zwei Relationen R, S zwischen A und B sind offensichtlich auch $R \cap S$ und $R \cup S$ Relationen zwischen A und B.

2.2 Beispiele. A und B seien Mengen.

(i) Die *leere Relation*: $R = \emptyset \subset A \times B$
Es gibt kein Paar (x, y), das diese Relation erfüllt.
$\mathbb{D}(R) = \mathbb{W}(R) = \emptyset$.

(ii) Die *Allrelation*: $R = A \times B \subset A \times B$
Alle $(x, y) \in A \times B$ erfüllen diese Relation.
$\mathbb{D}(R) = A$, $\mathbb{W}(R) = B$, $R = \mathbb{D}(R) \times \mathbb{W}(R) = A \times B$.

(iii) Die *Gleichheitsrelation*: $R = \Delta_A = \{(x, x) \mid x \in A\} \subset A \times A$
$(x, y) \in R$ genau dann, wenn $x = y$.
$\mathbb{D}(R) = \mathbb{W}(R) = A$.

(iv) Eine bekannte Relation auf $\mathbb{N}$ ist die $\leq$-Beziehung, gegeben durch

$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid y - x \in \mathbb{N}\}.$$

$\mathbb{D}(R) = \mathbb{W}(R) = \mathbb{N}$.

(v) Die Relation auf $\mathbb{N}$ *x ist Nachbar von y*, d.h. x und y unterscheiden sich um 1, ist bestimmt durch

$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x - y = 1 \text{ oder } y - x = 1\}.$$

$\mathbb{D}(R) = \mathbb{W}(R) = \mathbb{N}$.

(vi) Die Relation auf $\mathbb{R}$,

$$R = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x^2 + y^2 = 1\},$$

wird durch den Einheitskreis in der Ebene dargestellt.
$\mathbb{R} \neq \mathbb{W}(R) = \mathbb{D}(R) = \{x \in \mathbb{R} \mid -1 \leq x \leq 1\}$.

(vii) Auf der Potenzmenge $\mathcal{P}(A)$ einer Menge A ist durch die Teilmengenbeziehung (Inklusion) eine Relation gegeben:

$$R = \{(U, V) \in \mathcal{P}(A) \times \mathcal{P}(A) \mid U \subset V\}.$$

$\mathbb{D}(R) = \mathbb{W}(R) = \mathcal{P}(A)$.

Durch bloßes Vertauschen der Argumente läßt sich aus einer gegebenen Relation eine neue gewinnen:

2.3 Definition
Ist R eine Relation zwischen den Mengen A und B, so nennen wir die Relation zwischen B und A

$$R^{-1} = \{(y, x) \in B \times A \mid (x, y) \in R\}$$

die *Umkehrrelation von R* (*konverse* Relation).

Zu jeder Relation R kann R^{-1} gebildet werden, und es gilt $(R^{-1})^{-1} = R$.

Eine wichtige Bildung ist die *Verknüpfung* oder *Komposition* von zwei Relationen zwischen passenden Mengen:

2.4 Definition
A, B, C seien Mengen, R eine Relation zwischen A und B, S eine Relation zwischen B und C. Dann heißt die Relation

$$S \circ R = \{(x, z) \in A \times C \mid \text{es gibt ein } y \in B \text{ mit } (x, y) \in R \text{ und } (y, z) \in S\}$$

die *Verknüpfung* (*Komposition*) von R und S.

Damit kann man auch mehrere Relationen R, S, T zwischen geeigneten Mengen verknüpfen, und es gilt das *Assoziativgesetz* (Aufgabe 3)

$$T \circ (S \circ R) = (T \circ S) \circ R.$$

Zwischen Verknüpfung und Umkehrrelation haben wir folgende Beziehung:

2.5 Hilfssatz
Seien A, B, C Mengen, R eine Relation zwischen A und B und S eine Relation zwischen B und C. Dann gilt

$$(S \circ R)^{-1} = R^{-1} \circ S^{-1} \subset C \times A$$

Beweis:

$$\begin{aligned}
(z,x) \in (S \circ R)^{-1} &\Leftrightarrow (x,z) \in S \circ R \\
&\Leftrightarrow \text{es gibt ein } y \in B \text{ mit } (x,y) \in R,\ (y,z) \in S \\
&\Leftrightarrow \text{es gibt ein } y \in B \text{ mit } (y,x) \in R^{-1},\ (z,y) \in S^{-1} \\
&\Leftrightarrow (z,x) \in R^{-1} \circ S^{-1}\,.
\end{aligned}$$

□

Speziell kann man jede Relation mit ihrer Umkehrrelation verknüpfen:

2.6 Hilfssatz
Sei R eine Relation zwischen den Mengen A und B. Dann gilt:

(a) *$R^{-1} \circ R$ ist eine Relation auf A mit*
  (1) $R^{-1} \circ R = \{(x,z) \in A \times A \mid$ *es gibt* $y \in B$ *mit* $(x,y),\ (z,y) \in R\}$
  (2) $\Delta_{\mathbb{D}(R)} \subset R^{-1} \circ R$
  (3) $(R^{-1} \circ R)^{-1} = R^{-1} \circ R$

(b) *$R \circ R^{-1}$ ist eine Relation auf B mit*
  (1) $R \circ R^{-1} = \{(x,z) \in B \times B \mid$ *es gibt* $y \in A$ *mit* $(y,x),\ (y,z) \in R\}$
  (2) $\Delta_{\mathbb{W}(R)} \subset R \circ R^{-1}$
  (3) $(R \circ R^{-1})^{-1} = R \circ R^{-1}$

Beweis: (a.1) Nach Definition der Komposition von R mit R^{-1} gilt

$$\begin{aligned}
(x,z) \in R^{-1} \circ R &\Leftrightarrow \text{es gibt } y \in B \text{ mit } (x,y) \in R,\ (y,z) \in R^{-1} \\
&\Leftrightarrow \text{es gibt } y \in B \text{ mit } (x,y) \in R,\ (z,y) \in R.
\end{aligned}$$

(a.2) Für jedes $x \in \mathbb{D}(R)$ gibt es - nach Definition von $\mathbb{D}(R)$ - ein $y \in B$ mit $(x,y) \in R$. Damit ist $(x,x) \in R^{-1} \circ R$.

(a.3) Dies ergibt sich mit 2.5:

$$(R^{-1} \circ R)^{-1} = R^{-1} \circ (R^{-1})^{-1} = R^{-1} \circ R.$$

(b) folgt aus Teil (a). Man beachte, daß der Definitionsbereich von R gleich dem Wertebereich von R^{-1} ist: $\mathbb{D}(R) = \mathbb{W}(R^{-1})$. □

Es ist klar, daß im allgemeinen $R \circ R^{-1}$ und $R^{-1} \circ R$ verschieden sind.

2.7 Definition
Sei R eine Relation zwischen den Mengen A, B und U eine Teilmenge von A. Dann schreiben wir

$$R(U) = \{b \in B \mid \text{es gibt ein } u \in U \text{ mit } (u, b) \in R\} \subset B\,.$$

Speziell für $U = \{x\}$, $x \in A$, setzt man

$$R(x) = \{b \in B \mid (x, b) \in R\}.$$

Damit gilt $R(A) = \mathbb{W}(R)$, $R^{-1}(B) = \mathbb{D}(R)$ und

$$\begin{aligned} (x, R(x)) &= \{(x, y) \mid y \in R(x)\} \subset R \text{ für alle } x \in A, \\ R &= \bigcup \{(x, R(x)) \mid x \in A\}, \\ R^{-1}(y) &= \{x \in A \mid (x, y) \in R\} \text{ für alle } y \in B\,. \end{aligned}$$

Für zwei Relationen R, S zwischen geeigneten Mengen haben wir:

$$\begin{aligned} S(R(U)) &= S \circ R(U), \\ S(R(x)) &= S \circ R(x), \\ R^{-1} \circ R(x) &= \{z \in A \mid \text{ es gibt ein } y \in R(x) \text{ mit } (z, y) \in R\}\,. \end{aligned}$$

2.8 Aufgaben

(1) Bestimmen Sie die Umkehrrelationen zu den in 2.2 gegebenen Beispielen.

(2) Seien A, B Mengen und $R \subset A \times B$ eine Relation. Zeigen Sie:

$$R \circ \Delta_A = R \text{ und } \Delta_B \circ R = R\,.$$

(3) Es seien M, N, P, Q Mengen und $R \subset M \times N$, $S \subset N \times P$, $T \subset P \times Q$ Relationen. Zeigen Sie:

$$T \circ (S \circ R) = (T \circ S) \circ R \quad \textit{(Assoziativgesetz).}$$

(4) Seien R und S Relationen auf $\mathbb{N}$ mit

$$\begin{aligned} R &= \{(x, y) \in (\mathbb{N}, \mathbb{N}) \mid y = x^2\}, \\ S &= \{(a, b) \in (\mathbb{N}, \mathbb{N}) \mid b = a + 1\}. \end{aligned}$$

Man bestimme $R \circ S$ und $S \circ R$ und vergleiche diese Relationen.

3 Abbildungen

In diesem und den nächsten beiden Paragraphen wollen wir spezielle Relationen betrachten. Dazu gehört auch der Begriff der „Abbildung“ zwischen zwei Mengen, den Sie wahrscheinlich nicht als Relation kennengelernt haben, sondern als *Zuordnung* oder *Zuordnungsvorschrift*.

Die Darstellung von Abbildungen als Relationen ist zwar nicht anschaulicher als die Beschreibung als „Zuordnung“, sie ist jedoch – mit der Mengenlehre als Grundlage – logisch exakt und elementar formulierbar. Zudem erlaubt uns dies, die in §2 beobachteten Gesetzmäßigkeiten für Relationen auch für Abbildungen zu benutzen.

3.1 Definition
Eine Relation F zwischen zwei Mengen A, B heißt *Abbildung* oder *Funktion*, wenn

(1) $\mathbb{D}(F) = A$;

(2) gilt für $x \in A$ und $y, z \in B$, daß $(x, y) \in F$ und $(x, z) \in F$, so ist $y = z$.

F nennt man dann *Abbildung (Funktion) von A nach B*.

Eine Abbildung wird bestimmt durch das Tripel $(A, B; F)$. Also sind zwei Abbildungen $(A, B; F)$ und $(C, D; G)$ gleich, wenn $A = C$, $B = D$ und $F = G$. Man nennt A die *Quelle* und B das *Ziel* der Abbildung F. Nach Definition gilt

$$\begin{array}{rcccl} A & = & \text{Quelle } F & = & \mathbb{D}(F), \\ B & = & \text{Ziel } F & \supset & \mathbb{W}(F). \end{array}$$

Relationen mit der Eigenschaft (2) in 3.1 heißen *eindeutige Relationen*. Solche Relationen werden durch Einschränkung auf $\mathbb{D}(F)$ zu Abbildungen.

Äquivalent zu 3.1 können wir sagen:

Eine Relation F zwischen den Mengen A und B ist eine *Abbildung*, wenn für jedes $x \in A$ die Menge $F(x)$ aus *genau einem* Element besteht.

Somit ermöglicht eine Abbildung F eine eindeutige Zuordnung

$$f : A \to B, x \mapsto F(x) \text{ für alle } x \in A.$$

Man schreibt $f(x) := F(x) = \{b \in B \mid (x, b) \in F\}$, und für $U \subset A$ setzt man

$$f(U) := F(U) = \{b \in B \mid \text{es gibt ein } u \in U \text{ mit } (u, b) \in F\}.$$

Man nennt F auch den *Graphen der Abbildung f*. Diese Bezeichnung ist von dem Spezialfall $f : \mathbb{R} \to \mathbb{R}$ abgeleitet, in dem der Graph von f gerade die Kurve in der reellen Ebene ergibt, die zu f gehört.

Ist eine Abbildung als Zuordnung $f : A \to B$ gegeben, dann wird durch

$$F = \{(x, f(x)) \mid x \in A\} \subset A \times B$$

die zugehörige Relation beschrieben.

Von den in 2.2 gegebenen Relationen ist nur die Gleichheit eine Abbildung. $\Delta_A \subset A \times A$ bestimmt die *identische Abbildung*

$$id_A : A \to A, \quad x \mapsto x \text{ für alle } x \in A.$$

Die in 2.2(vi) betrachtete Relation auf $\mathbb{R}$, $R = \{(x, y) \mid x^2 + y^2 = 1\}$, kann durch Einschränkung der Quelle auf $\mathbb{D}(R)$ und geeignete Beschränkung des Zielbereichs zur Abbildung werden:

$$\mathbb{D}(R) = \{x \in \mathbb{R} \mid -1 \leq x \leq 1\} \to \mathbb{R}^+ = \{y \in \mathbb{R} \mid y > 0\}\,.$$

Sind zwei Abbildungen F und G mit geeigneten Quellen und Zielen gegeben, so kann man diese zu einer neuen Relation $G \circ F$ verknüpfen. Dies ist wieder eine Abbildung:

3.2 Hilfssatz
Seien A, B, C Mengen und $f : A \to B$, $g : B \to C$ Abbildungen, bestimmt durch $F \subset A \times B$ und $G \subset B \times C$. Dann ist auch die durch die Verknüpfung $G \circ F$ gebildete Relation eine Abbildung.
Diese bezeichnen wir mit $g \circ f : A \to C$.

Beweis: Es ist zu zeigen, daß für alle $x \in A$ die Menge $G \circ F(x) = G(F(x))$ aus genau einem Element besteht:
Da F Abbildung ist, besteht $F(x)$ aus genau einem Element von B.
Da G Abbildung ist, besteht $G(F(x))$ aus genau einem Element von C. □

Im allgemeinen braucht die Umkehrrelation F^{-1} einer Abbildung keine Abbildung zu sein. Dies kann man sich etwa an der Funktion $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$, klar machen. Es gilt jedoch:

3.3 Hilfssatz. *Sei $F \subset A \times B$ eine Abbildung.*

(1) *Dann gilt für $F^{-1} \circ F \subset A \times A$:*
   (i) $\Delta_A \subset F^{-1} \circ F$
   (ii) $(F^{-1} \circ F)^{-1} = F^{-1} \circ F$
   (iii) $(F^{-1} \circ F) \circ (F^{-1} \circ F) \subset F^{-1} \circ F$.
   Es gilt $(x, z) \in F^{-1} \circ F$ genau dann, wenn $F(x) = F(z)$.

(2) $F \circ F^{-1} = \Delta_{\mathbb{W}(F)} \subset B \times B$.

Beweis: (i) und (ii) folgen sofort aus Hilfssatz 2.6.

(iii) Für $(x,z) \in (F^{-1} \circ F) \circ (F^{-1} \circ F)$ gilt: Es gibt $y \in A$ mit

$$(x,y) \in F^{-1} \circ F, \quad (y,z) \in F^{-1} \circ F,$$

d.h. es gibt ein $u \in B$ mit $(x,u) \in F$, $(y,u) \in F$ und ein $v \in B$ mit $(y,v) \in F$, $(z,v) \in F$.

Da F Abbildung ist, folgt aus $(y,u) \in F$ und $(y,v) \in F$, daß $u = v$. Dann ist $(x,u) \in F$ und $(z,u) \in F$, also $(x,z) \in F^{-1} \circ F$ □

Der Fall (2) in 3.3 beschreibt eine Situation, in der die Verknüpfung einer Relation mit einer Abbildung wieder eine Abbildung ergibt.

Daß F^{-1} keine Abbildung ist, kann als Ursache haben:

(1) $\mathbb{D}(F^{-1}) = \mathbb{W}(F) \subset B$, es muß nicht $\mathbb{W}(F) = B$ gelten.

(2) Für ein $y \in \mathbb{W}(F)$ kann $F^{-1}(y)$ mehr als nur ein Element enthalten.

Abbildungen, in denen solche „Defekte“ nicht auftreten, verdienen besonderes Interesse:

3.4 Definition

Sei $f : A \to B$ eine Abbildung.

(1) f heißt *surjektiv*, wenn $f(A) = B$,
d.h. zu jedem $z \in B$ gibt es ein $x \in A$ mit $f(x) = z$.

(2) f heißt *injektiv* (oder *eineindeutig*), wenn für $x \neq y \in A$ auch $f(x) \neq f(y)$,
d.h. aus $f(x) = f(y)$ folgt $x = y$.

(3) f heißt *bijektiv*, wenn es injektiv und surjektiv ist.

Folgerungen

(1) f ist genau dann surjektiv, wenn $\mathbb{W}(F) = B$.

(2) Eine Abbildung $f = (A, B; F)$ ist genau dann *injektiv*, wenn die Relation $F^{-1} \subset \mathbb{W}(F) \times A$ eine Abbildung ist.

(3) $f = (A, B; F)$ ist genau dann *bijektiv*, wenn die Relation $F^{-1} \subset B \times A$ eine Abbildung ist.
Dann heißt $f^{-1} = (B, A; F^{-1})$ die *Umkehrabbildung* zu f, und es gilt:

$$\begin{aligned} f^{-1} \circ f &= id_A \text{ (da } F^{-1} \circ F = \Delta_A) \\ f \circ f^{-1} &= id_B \text{ (da } F \circ F^{-1} = \Delta_B, \text{ vgl. 3.3}). \end{aligned}$$

Beweis: (1) Dies folgt unmittelbar aus den Definitionen.

(2) ⇒ Sei f injektiv.

$$\begin{aligned}(z,x) \in F^{-1},\, (z,y) \in F^{-1} &\Leftrightarrow (x,z) \in F,\, (y,z) \in F\\ &\Rightarrow f(x) = f(y)\\ &\Rightarrow x = y \text{ wegen } F \text{ injektiv}\\ &\Rightarrow F^{-1} \text{ ist Abbildung.}\end{aligned}$$

$\Leftarrow$ F^{-1} sei Abbildung, $f(x) = f(y) =: z$.

$$\begin{aligned}(x,z) \in F,\, (y,z) \in F &\Leftrightarrow (z,x) \in F^{-1},\, (z,y) \in F^{-1}\\ &\Rightarrow x = y\\ &\Rightarrow f \text{ ist injektiv.}\end{aligned}$$

(3) Die Behauptung ergibt sich aus (1) und (2). □

Für die Komposition von Abbildungen haben wir die Beziehungen:

3.5 Satz
Seien $f : A \to B$, $g : B \to C$ Abbildungen. Dann gilt:

(1) *Sind f und g injektiv (surjektiv, bijektiv), so ist auch $g \circ f$ injektiv (surjektiv, bijektiv).*

(2) *Ist $g \circ f$ injektiv, so ist auch f injektiv.*
Ist $g \circ f$ surjektiv, so ist auch g surjektiv.
Ist $g \circ f$ bijektiv, so ist f injektiv und g surjektiv.

Der einfache Beweis dazu sei dem Leser zur Übung belassen. Damit können wir folgende Kennzeichnung von bijektiven Abbildungen angeben:

3.6 Korollar
Eine Abbildung $f : A \to B$ ist genau dann bijektiv, wenn es eine Abbildung $g : B \to A$ gibt mit

$$g \circ f = id_A, \quad f \circ g = id_B$$

Es gilt dann $g = f^{-1}$ (die inverse Abbildung ist somit eindeutig bestimmt).

Beweis: Nach 3.5 folgt aus $g \circ f = id_A$, daß f injektiv ist und aus $f \circ g = id_B$, daß f surjektiv ist. Damit existiert f^{-1}, und aus der ersten Gleichung folgt $f^{-1} = g \circ (f \circ f^{-1}) = g$. □

Man beachte, daß die Gültigkeit von nur einer der beiden Gleichungen, z.B. $f \circ g = id_B$, *nicht* die Bijektivität von f zur Folge hat.

Es folgt aus dem Auswahlaxiom, daß es zu *jeder* surjektiven Abbildung $f : A \to B$ eine Abbildung $g : B \to A$ gibt mit $f \circ g = id_B$. Dazu ist folgende Formulierung des Auswahlaxioms 1.15 von Nutzen:

3.7 Auswahlaxiom II
Zu jeder Relation $R \subset A \times B$ gibt es eine Abbildung $F \subset \mathbb{D}(R) \times B$ mit $F \subset R$.

Beweis: Die Relation R ordnet jedem $x \in \mathbb{D}(R)$ eine Menge $R(x) \subset B$ zu. Nach dem Auswahlaxiom 1.15 können wir aus jeder Menge $R(x)$ ein y_x herausnehmen. Die so gebildeten Paare $F = \{(x, y_x) \mid x \in \mathbb{D}(R)\}$ bilden eine Abbildung $\mathbb{D}(R) \to B$.

Nehmen wir nun an, das Auswahlaxiom II gilt. Sei $\mathcal{M}$ eine Menge von elementfremden Mengen, und setze $B := \bigcup_{A \in \mathcal{M}} A$. Betrachte die Relation

$$R := \{(A, a) \mid A \in \mathcal{M} \text{ und } a \in A\} \subset \mathcal{M} \times B\,.$$

Dann gibt es eine Abbildung $F \subset R \subset \mathcal{M} \times B$, und für die Menge $F(\mathcal{M})$ gilt $F(A) = F(\mathcal{M}) \cap A$ für jedes $A \in \mathcal{M}$. Damit sind die Bedingungen des Auswahlaxioms erfüllt. □

Aus Hilfssatz 3.3 kann man dann ableiten:

3.8 Satz
Sei $f = (A, B; F)$ eine surjektive Abbildung. Dann gibt es eine Abbildung $g = (B, A; G)$ mit $f \circ g = id_B$.

Beweis: Nach 3.3 gilt $F \circ F^{-1} = \Delta_B$. Zu der Relation F^{-1} können wir wegen 3.7 eine Abbildung $g = (B, A; G)$ finden mit $G \subset F^{-1}$.
Dafür gilt $F \circ G \subset F \circ F^{-1} = \Delta_B$. Außerdem ist $\Delta_B \subset F \circ G$, denn zu jedem $x \in B$ gibt es

$$y \in A \text{ mit } (x, y) \in G \subset F^{-1} \subset B \times A,$$

also $(y, x) \in F$, und somit ist $(x, x) \in F \circ G$. □

Wir haben bei den Folgerungen zu 3.4 gesehen, daß für eine injektive Abbildung $f = (A, B; F)$ die Umkehrrelation F^{-1} eine Abbildung von $\mathbb{W}(F)$ in A ist mit $F^{-1} \circ F = id_A$ (beachte $\mathbb{W}(F^{-1}) = A$).

Man kann die Abbildung $(\mathbb{W}(F), A; F^{-1})$ so zu einer Funktion $g = (B, A; G)$ fortsetzen (erweitern), daß die Beziehung $G \circ F = id_A$ erhalten bleibt. Damit gilt analog zu 3.8:

3.9 Satz
Sei $f = (A, B; F)$ eine injektive Abbildung. Dann gibt es eine Abbildung $g = (B, A; G)$ mit $g \circ f = id_A$.

Beweis: Wir wählen ein beliebiges (festes) $x_0 \in A$ und definieren

$$g(y) = \begin{cases} F^{-1}(y) & \text{für } y \in \mathbb{W}(F) \\ x_0 & \text{für } y \notin \mathbb{W}(F) \end{cases}$$

Die zugehörige Menge $G \subset B \times A$ ist dabei

$$G = \{(y, F^{-1}(y)) \mid y \in \mathbb{W}(F)\} \cup \{(y, x_0) \mid y \in B \setminus \mathbb{W}(F)\}$$

Man beachte, daß für $g \circ f$ der Wert von g an den Stellen $y \notin \mathbb{W}(F)$ keine Rolle spielt. Es ist leicht nachzuprüfen, daß $g(f(x)) = x$ für alle $x \in A$. □

Abbildungen $f : A \to B$, die jedem Element $x \in A$ den gleichen Wert $y \in B$ zuordnen, heißen *konstante Abbildungen*. Solche Abbildungen lassen sich zwischen zwei beliebigen nicht-leeren Mengen A, B angeben. Eine Abbildung $f = (A, B; F)$ ist genau dann konstant, wenn $F = \{(x, y) \mid x \in A\}$ für ein (festes) $y \in B$.

In der nachstehenden Liste wird angegeben, wie sich Abbildungen gegenüber mengentheoretischen Bildungen wie Durchschnitt, Vereinigung und Komplement verhalten.

3.10 Abbildungen und Mengenoperationen

Seien $f : A \to B$ eine Abbildung und $U, U' \subset A$. Dann gilt:

(1) $U \subset U' \Rightarrow f(U) \subset f(U'); \quad f(U \cup U') = f(U) \cup f(U').$

(2) $f(U) \setminus f(U') \subset f(U \setminus U'); \quad f(U \cap U') \subset f(U) \cap f(U').$

Für jede Menge $\mathcal{U}$ von Teilmengen von A gilt:

(3) $f(\bigcup_{U \in \mathcal{U}} U) = \bigcup_{U \in \mathcal{U}} f(U); \quad f(\bigcap_{U \in \mathcal{U}} U) \subset \bigcap_{U \in \mathcal{U}} f(U).$

Für Teilmengen V, V' von B gilt:

(4) $V \subset V' \Rightarrow f^{-1}(V) \subset f^{-1}(V'); \quad f^{-1}(V \cup V') = f^{-1}(V) \cup f^{-1}(V').$

(5) $f^{-1}(V) \setminus f^{-1}(V') = f^{-1}(V \setminus V'); \quad f^{-1}(V \cap V') = f^{-1}(V) \cap f^{-1}(V').$

Für jede Menge $\mathcal{V}$ von Teilmengen von B gilt:

(6) $f^{-1}(\bigcup_{V \in \mathcal{V}} V) = \bigcup_{V \in \mathcal{V}} f^{-1}(V); \quad f^{-1}(\bigcap_{V \in \mathcal{V}} V) = \bigcap_{V \in \mathcal{V}} f^{-1}(V).$

Der Beweis dazu ist nicht schwierig und sei dem Leser überlassen.

Hat man mit einer Menge von Elementen oder Mengen zu arbeiten, so kann man häufig die Ausdrucksweise dadurch vereinfachen, daß man diese mit einem Index versieht (indiziert). Wählt man etwa zwei Elemente aus einer Menge B,

so schreibt man $b_1, b_2 \in B$. In diesem Fall ist die Indexmenge $I = \{1, 2\}$, und die Elemente $b_1, b_2 \in B$ sind bestimmt als Bilder einer Abbildung

$$f : I \to B, \quad 1 \mapsto b_1,\ 2 \mapsto b_2.$$

Allgemeiner formulieren wir:

3.11 Familie von Elementen
Seien I und M Mengen, $f : I \to M$ eine Abbildung. Man nennt dann f auch eine *(I-indizierte) Familie von Elementen* und schreibt dafür

$$(f(i) \,|\, i \in I) \ \text{ oder } (f(i))_{i \in I}.$$

Setzt man $f(i) := b_i$, so beschreibt auch $(b_i)_{i \in I}$ die Abbildung f.

Zwei Abbildungen $f, g : I \to M$ sind genau dann gleich, wenn die Familien $(f(i))_{i \in I}$ und $(g(i))_{i \in I}$ gleich sind, d.h. wenn $f(i) = g(i)$ für alle $i \in I$.

Man achte darauf, die Familie $(f(i))_{i \in I}$ von der Menge der Bildelemente $\{f(i) \mid i \in I\} \subset B$ zu unterscheiden. Die Schreibweise $(f(i) \,|\, i \in I)$ legt die Abbildung fest, die Bildmenge dagegen nicht.

Spezialfälle

(1) $I = \{1\}$. $\{1\} \to B$ wird durch ein $b \in B$ bestimmt.

(2) $I = \{1, 2\}$. $\{1, 2\} \to B$ ergibt Paare $b_1, b_2 \in B$ ($b_1 = b_2$ möglich).

(3) $I = \{1, \ldots, n\}$. $\{1, \ldots, n\} \to B$ nennt man *n-Tupel* $(b_1, \ldots, b_n)$.

(4) $I = \mathbb{N}$. $\mathbb{N} \to B$ nennt man *Folgen*.

(5) Ist $B = \mathcal{B}$ eine Menge von Mengen, so ergibt eine Abbildung $I \to \mathcal{B}$ eine *Familie von Mengen* oder ein *indiziertes Mengensystem*.

Nach Definition der Abbildungen zwischen Mengen A und B bildet die Gesamtheit dieser Abbildungen eine Menge, die wir mit $\mathrm{Abb}(A, B)$ bezeichnen wollen ($\mathrm{Abb}(A, B) \subset \mathcal{P}(A \times B)$). Die oben betrachteten Fälle ergeben folgende Entsprechungen (mit $\widehat{=}$ bezeichnet):

(1) $\mathrm{Abb}(\{1\}, B) \ \widehat{=}\ B$.
(2) $\mathrm{Abb}(\{1, 2\}, B) \ \widehat{=}\ B \times B$.
(3) $\mathrm{Abb}(\{1, \ldots, n\}, B) \ \widehat{=}$ Menge der n-Tupel mit Elementen aus B.
(4) $\mathrm{Abb}(\mathbb{N}, B) \ \widehat{=}$ Menge der Folgen mit Elementen aus B.
(5) $\mathrm{Abb}(I, \mathcal{M}) \ \widehat{=}\ \{(M_i)_{i \in I} \mid M_i \in \mathcal{M}\}$.

In (2) haben wir das geordnete Paar $B \times B$ als $\mathrm{Abb}(\{1,2\}, B)$ wiederentdeckt. Auch das geordnete Paar $A \times B$ läßt sich ähnlich darstellen:

$$\begin{aligned} A \times B &\widehat{=} \{f \in \mathrm{Abb}(\{1,2\}, A \cup B) \mid f(1) \in A, f(2) \in B\}, \\ A_1 \times A_2 &\widehat{=} \{f \in \mathrm{Abb}(\{1,2\}, A_1 \cup A_2) \mid f(1) \in A_1, f(2) \in A_2\}. \end{aligned}$$

Man nennt dies auch das *Produkt* der Mengen A_1, A_2. Ausgehend davon kann man schrittweise das *Produkt* von mehreren Mengen bilden:

$$A_1 \times A_2 \times A_3 \times \cdots = ((A_1 \times A_2) \times A_3) \times \cdots.$$

Man kann dies auch beschreiben durch

$$A_1 \times \cdots \times A_n = \{f \in \mathrm{Abb}(\{1,\ldots,n\}, \bigcup_{i=1}^{n} A_i) \mid f(i) \in A_i,\ i = 1,\ldots,n\}.$$

Allgemein definieren wir daher:

3.12 Definition
Sei I eine Menge, $(A_i)_{i\in I}$ eine Familie von Mengen. Dann nennt man

$$\textstyle\prod_{i\in I} A_i = \{f \in \mathrm{Abb}(I, \bigcup_{i\in I} A_i) \mid f(i) \in A_i \text{ für jedes } i \in I\}$$

das (*kartesische*) *Produkt* der Mengen A_i. Kurzschreibweise: $\prod_I A_i$.
Die Elemente von $\prod A_i$ sind die Familien $(a_i)_{i\in I}$ mit $a_i \in A_i$.
Dabei ist $(a_i)_{i\in I} = (b_i)_{i\in I}$ genau dann, wenn $a_i = b_i$ für alle $i \in I$.

Ist $A_i = A$ für alle $i \in I$, so schreibt man $\prod_{i\in I} A_i = A^I$, d.h. es gilt

$$A^I = \textstyle\prod_{i\in I} A_i = \mathrm{Abb}(I, A).$$

Für $I = \{1,\ldots,n\}$ haben wir damit

$$A^n = A \times \ldots \times A = \mathrm{Abb}(\{1,\ldots,n\}, A).$$

In 3.12 ist nichts darüber gesagt, ob das Produkt einer beliebigen Familie von Mengen überhaupt existiert. Diese Existenz wird uns durch eine weitere Version des Auswahlaxioms gesichert.

3.13 Auswahlaxiom III
Zu jeder Familie $(A_i)_{i\in I}$ von nicht-leeren Mengen A_i existiert das kartesische Produkt $\prod_{i\in I} A_i$.

Beweis: Wählen wir gemäß dem Auswahlaxiom für jedes $i \in I$ ein $a_i \in A_i$ und definieren

$$f : I \to \bigcup_{i\in I} A_i, \quad i \mapsto a_i,$$

dann ist $f \in \prod_{i\in I} A_i$.

Gilt andererseits $\prod_{i\in I} A_i \neq \emptyset$ für nicht-leere Mengen A_i, dann gibt es $f \in \prod_{i\in I} A_i$ und $f(i) = a_i \in A_i$ für alle $i \in A$. □

Wir haben $\prod A_i$ als Abbildungen $I \to \bigcup A_i$ definiert. Durch Anwenden dieser Abbildungen auf ein Element $k \in I$ ergibt sich eine Abbildung

$$\{f \in \mathrm{Abb}(I, \bigcup_i I A_i) \mid f(i) \in A_i\} \to A_k, \quad f \mapsto f(k)\,.$$

In etwas handlicherer Notation läßt sich das so ausdrücken:

3.14 Definition
Sei $(A_i)_{i\in I}$ eine Familie von Mengen. Die zu $k \in I$ definierte Abbildung

$$\pi_k : \prod_{i\in I} A_i \to A_k, \quad (a_i)_{i\in I} \mapsto a_k,$$

nennt man die *k-te Projektion* von $\prod_I A_i$ auf A_k.

3.15 Hilfssatz
Ist $(A_i)_{i\in I}$ eine Familie von nicht-leeren Mengen, so ist für jedes $k \in I$ die Projektion $\pi_k : \prod_I A_i \to A_k$ surjektiv.

Beweis: Sei $a_k \in A_k$. Dann wählen wir beliebige $a_i \in A_i$ für $i \in I \setminus \{k\}$ (Auswahlaxiom), und wir erhalten

$$\pi_k((a_i)_{i\in I}) = a_k.$$

Also ist π_k surjektiv. □

Folgende Eigenschaft des Produkts von Mengen ist von Bedeutung:

3.16 Satz *(Universelle Eigenschaft des Produkts von Mengen)*
Zur Indexmenge I und einer Familie von Mengen $(A_i)_{i\in I}$ seien

$\prod_I A_i$ das kartesische Produkt dieser Mengen und

$(\pi_k : \prod_I A_i \to A_k)_{k\in I}$ die Familie der Projektionen.

Dann gibt es zu jeder Menge B und jeder Familie $(f_k : B \to A_k)_{k\in I}$ von Abbildungen genau eine Abbildung $f : B \to \prod_I A_i$ mit $\pi_k \circ f = f_k$,

Man sagt, für jedes $k \in I$ ist folgendes Diagramm kommutativ:

$$\begin{array}{ccc} B & \xrightarrow{f_k} & A_k \\ f \searrow & & \nearrow \pi_k \\ & \prod_I A_i & \end{array}$$

Beweis: Wir definieren $f : B \to \prod_I A_i$, $b \mapsto (f_i(b))_{i\in I}$ und erhalten

$$\pi_k \circ f(b) = \pi_k((f_i(b))_{i\in I}) = f_k(b) \text{ für jedes } b \in B.$$

Es bleibt noch die Eindeutigkeit von f zu zeigen. Angenommen, es gibt ein $g : B \to \prod_I A_i$ mit $\pi_k \circ g = f_k$, also $\pi_k \circ g(b) = f_k(b)$ für alle $b \in B$.

Angenommen $f \neq g$, d.h. es gibt ein $b \in B$ mit $f(b) \neq g(b) \in \prod_I A_i$. Dann gibt es ein $k \in I$ mit

$$\pi_k \circ f(b) \neq \pi_k \circ g(b).$$

Dies bedeutet aber $f_k(b) \neq f_k(b)$, ein Widerspruch. □

Zu einer Menge B und einer Familie $(A_i)_{i\in I}$ von Mengen kann man die Menge $\mathrm{Abb}(B, \prod_I A_i)$ und das Produkt der Mengen $\mathrm{Abb}(B, A_i)$, bezeichnet mit $\prod_I \mathrm{Abb}(B, A_i)$, bilden. Die universelle Eigenschaft des Produktes besagt gerade, daß wir diese beiden Mengen identifizieren können:

3.17 Korollar
Sei B eine Menge und $(A_i)_{i\in I}$ eine Familie von Mengen. Dann ist folgende Abbildung bijektiv:

$$\mathrm{Abb}(B, \textstyle\prod_I A_i) \to \prod_I \mathrm{Abb}(B, A_i), \quad f \mapsto (\pi_i \circ f)_{i\in I}\,.$$

Beweis: Zunächst zeigen wir die Injektivität.

$$\begin{array}{rcl} f \neq g & \Rightarrow & \text{es gibt ein } b \in B \text{ mit } f(b) \neq g(b) \\ & \Rightarrow & \text{es gibt ein } k \in I \text{ mit } \pi_k \circ f(b) \neq \pi_k \circ g(b) \\ & \Rightarrow & \pi_k \circ f \neq \pi_k \circ g\,. \end{array}$$

Nun zur Surjektivität.

Sei $(h_i)_{i\in I} \in \prod_I \mathrm{Abb}(B, A_i)$. Dann gibt es nach 3.16 ein $f : B \to \prod_I A_i$ mit $\pi_k \circ f = h_k$ für alle $k \in I$, also $(\pi_i \circ f)_{i\in I} = (h_i)_{i\in I}$. □

3.18 Aufgaben

In den ersten vier Aufgaben sei $f = (A, B; F)$ eine Abbildung.

(1) Zeigen Sie:

(i) *Für alle Teilmengen $U \subset A$, $V \subset B$ gilt $U \subset F^{-1} \circ F(U)$ und $F \circ F^{-1}(V) \subset V$.*

(ii) *f ist genau dann surjektiv, wenn für alle $V \subset B$ gilt: $F \circ F^{-1}(V) = V$.*

(iii) *f ist genau dann injektiv, wenn für alle $U \subset A$ gilt: $F^{-1} \circ F(U) = U$.*

(2) Für Teilmengen $U_1, U_2 \subset A$ und $V_1, V_2 \subset B$ gilt:

(i) $F(U_1 \cup U_2) = F(U_1) \cup F(U_2), \quad F^{-1}(V_1 \cup V_2) = F^{-1}(V_1) \cup F^{-1}(V_2)$;

(ii) $F(U_1 \cap U_2) \subset F(U_1) \cap F(U_2), \quad F^{-1}(V_1 \cap V_2) = F^{-1}(V_1) \cap F^{-1}(V_2)$;

(iii) $F(U_1) \setminus F(U_2) \subset F(U_1 \setminus U_2), \quad F^{-1}(V_1) \setminus F^{-1}(V_2) = F^{-1}(V_1 \setminus V_2)$.

(3) Man zeige, daß folgende Aussagen äquivalent sind:

(a) *f ist surjektiv;*

(b) *für jede Menge C und alle Abbildungen g, h von B nach C gilt $g \circ f = h \circ f \Rightarrow g = h$;*

(c) *es gibt eine Abbildung $g = (B, A; G)$ mit $f \circ g = id_B$.*

(4) Folgende Aussagen sind äquivalent:

(a) *f ist injektiv;*

(b) *für jede Menge C und alle Abbildungen g, h von C nach A gilt $f \circ g = f \circ h \Rightarrow g = h$;*

(c) *es gibt eine Abbildung $g = (B, A; G)$ mit $g \circ f = id_A$.*

(5) Seien $f : A \to B$, $g : B \to C$ und $h : C \to D$ Abbildungen. Zeigen Sie: Sind $g \circ f$ und $h \circ g$ bijektiv, so sind auch f, g und h bijektiv.

(6) A, B und C seien nicht-leere Mengen, $g : B \to C$ eine Abbildung. Betrachten Sie die Abbildung

$$\widetilde{g} : \mathrm{Abb}(A, B) \to \mathrm{Abb}(A, C), \quad f \mapsto g \circ f.$$

Man beweise: $\widetilde{g}$ ist genau dann injektiv, wenn g injektiv ist.

4 Äquivalenzrelationen

Wir haben in §3 *Abbildungen* als Relationen mit speziellen Eigenschaften kennengelernt. Auch in diesem Paragraphen werden wir Relationen auf einer Menge mit besonderen Eigenschaften untersuchen.

4.1 Definition
Sei R eine Relation auf der Menge A, also $R \subset A \times A$.

(1) R heißt *reflexiv*, wenn $\Delta_A \subset R$, d.h. $(a, a) \in R$ für alle $a \in A$.

(2) R heißt *symmetrisch*, wenn $R = R^{-1}$, d.h. $(a, b) \in R \Rightarrow (b, a) \in R$.

(3) R heißt *transitiv*, wenn $R \circ R \subset R$,
d.h. $(x, y) \in R$ und $(y, z) \in R \Rightarrow (x, z) \in R$.

(4) R heißt *Äquivalenzrelation*, wenn R reflexiv, symmetrisch und transitiv ist.

Beispiele

(1) Die *gröbste* Äquivalenzrelation ist die Allrelation $(R = A \times B)$.

(2) Die *kleinste* oder *feinste* Äquivalenzrelation ist die Gleichheit $(R = \Delta_A)$: Dabei steht jedes Element nur mit sich selbst in Relation.

(3) Eine Äquivalenzrelation R auf $\mathbb{N}$ ist gegeben durch
$R = \{(n, m) \in \mathbb{N} \times \mathbb{N} \mid n$ und m haben die gleiche Quersumme$\}$
(in Dezimaldarstellung).

Sei R eine Relation auf A. Zu jedem Element $a \in A$ haben wir die Menge $R(a)$ gebildet, also die Menge der Elemente $b \in A$ mit $(a, b) \in R$. Falls R transitiv ist, gilt dafür $R(R(a)) = R(a)$.

Ist R eine Äquivalenzrelation, so nennt man die Elemente in $R(a)$ *äquivalent* zu a (bezüglich R). Wir halten fest:

4.2 Definition
Sei R eine Äquivalenzrelation auf A. Dann schreiben wir für $a \in A$

$$[a] := R(a) = \{b \in A \mid (a, b) \in R\}$$

und nennen dies die *Äquivalenzklasse* von a (bezüglich R).

Die Gesamtheit der Äquivalenzklassen bildet eine Menge, nämlich eine Teilmenge der Potenzmenge $\mathcal{P}(A)$ von A, die wir mit A/R bezeichnen, also

$$A/R = \{[a] \mid a \in A\} = \{R(a) \mid a \in A\}.$$

Man beachte, daß bei dieser Beschreibung für verschiedene $a, b \in A$ durchaus $[a] = [b]$ sein kann. Mit obigen Bezeichnungen gilt:

4.3 Hilfssatz
Sei R eine Äquivalenzrelation auf A. Für Elemente $a, b \in A$ sind folgende Aussagen äquivalent:

(a) $[a] = [b]$;

(b) $[a] \cap [b] \neq \emptyset$;

(c) $(a, b) \in R$.

Beweis: (a)⇒(b): $a \in [a] \cap [b] \neq \emptyset$.

(b)⇒(c): Ist $[a] \cap [b] \neq \emptyset$, dann gibt es ein $c \in A$ mit $(a, c) \in R$ und $(c, b) \in R$, also $(a, b) \in R$ (wegen Transitivität).

(c)⇒(a): Gilt $(a, b) \in R$, so ist $a \in [b]$ und $b \in [a]$, also $[a] \subset [b]$ und $[b] \subset [a]$. □

Die Bildung der Äquivalenzklassen definiert eine Abbildung von A in die Menge A/R der Äquivalenzklassen. Wie bezeichnen sie so:

4.4 Definition
Sei R eine Äquivalenzrelation auf A. Dann heißt die Abbildung

$$p_R : A \to A/R, \quad a \mapsto [a] \text{ für } a \in A,$$

die (zu R gehörende) *kanonische Abbildung (Projektion).*

Es ist klar, daß p_R surjektiv ist.

Erinnern wir uns nun an eine Äquivalenzrelation, die wir früher schon kennengelernt haben:

Ist $f = (A, B; F)$ eine Abbildung, so wird $F^{-1} \circ F$ eine Relation auf A. Wir haben in 3.3 gesehen, daß dies eine reflexive, symmetrische und transitive Relation ist. Dabei bedeutet $(a, b) \in F^{-1} \circ F$, daß es ein $c \in B$ gibt mit $(a, c) \in F$ und $(c, b) \in F^{-1}$, also $(b, c) \in F$.

Somit gilt $(a, b) \in F^{-1} \circ F$ genau dann, wenn $f(a) = f(b)$, also:

4.5 Satz
Sei $f = (A, B; F)$ eine Abbildung. Dann ist

$$R_f = F^{-1} \circ F = \{(a, b) \in A \times A \mid f(a) = f(b)\}$$

eine Äquivalenzrelation auf A.

Die dazu gehörende kanonische Projektion auf die Menge der Äquivalenzklassen bezeichnen wir mit $p_f : A \to A/R_f$.

Damit haben wir das Diagramm

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle p_f}\searrow & & \\ & A/R_f & . \end{array}$$

Wir suchen nun eine Abbildung $\overline{f} : A/R_f \to B$, die dieses Diagramm kommutativ ergänzt, also mit $f = \overline{f} \circ p_f$. Falls es ein solches $\overline{f}$ gibt, muß somit $f(a) = \overline{f} \circ p_f(a) = \overline{f}([a])$ gelten. Wir schlagen daher die Zuordnung vor:

$$\overline{f} : A/R_f \to B, \quad \overline{f}([a]) = f(a).$$

Diese Festlegung erscheint zunächst von der Auswahl des Repräsentanten a aus $[a]$ abhängig. Betrachten wir also $a, b \in A$ mit $[a] = [b]$. Dann gilt nach 4.3 $(a, b) \in R_f$, was – nach Definition von R_f – gerade $f(a) = f(b)$ bedeutet. Unsere Zuordnung ist somit nicht von der Auswahl des Repräsentanten abhängig, d.h. $\overline{f}$ ist eine Abbildung.

$\overline{f}$ ist sogar injektiv. Gilt nämlich $[a] \neq [b]$, also $f(a) \neq f(b)$, so ist auch

$$\overline{f}([a]) = f(a) \neq f(b) = \overline{f}([b]).$$

Fassen wir zusammen:

4.6 Satz
Jede Abbildung $f : A \to B$ läßt sich darstellen als Komposition der

surjektiven Abbildung $p_f : A \to A/R_f$ und der

injektiven Abbildung $\overline{f} : A/R_f \to B$,

d.h. folgendes Diagramm ist kommutativ:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle p_f}\searrow & & \nearrow{\scriptstyle \overline{f}} \\ & A/R_f & \end{array}$$

Als Folgerung daraus halten wir fest:

4.7 Korollar
Zu jeder Abbildung $f : A \to B$ gibt es eine bijektive Abbildung zwischen A/R_f und $\mathbb{W}(f)$ (= Bild von f).

Beweis: Nach Konstruktion gilt $\mathbb{W}(f) = \mathbb{W}(\overline{f})$.
Damit ist $\overline{f} : A/R_f \to \mathbb{W}(f)$ injektiv und surjektiv, also bijektiv. □

Wir haben sowohl Abbildungen als auch Äquivalenzrelationen als Relationen kennengelernt. Es ist leicht zu sehen, daß eine Relation, die Abbildung und Äquivalenzrelation ist, schon die Identität sein muß. Weitere Eigenschaften von Relationen werden im nächsten Abschnitt untersucht.

4.8 Aufgaben.

(1) Prüfen Sie, ob die folgenden Relationen auf $\mathbb{N}$ jeweils reflexiv, symmetrisch oder transitiv sind:
$R_1 = \{(n,m) \in \mathbb{N} \times \mathbb{N} \mid n < m\}$
$R_2 = \{(n,m) \in \mathbb{N} \times \mathbb{N} \mid n \leq m\}$
$R_3 = \{(n,m) \in \mathbb{N} \times \mathbb{N} \mid n + m = 26\}$
$R_4 = \{(n,m) \in \mathbb{N} \times \mathbb{N} \mid n \text{ und } m \text{ sind Primzahlen}\}$
$R_5 = \{(n,m) \in \mathbb{N} \times \mathbb{N} \mid n = m \text{ oder } (n,m) = (5,7)\}$
$R_6 = \{(n,m) \in \mathbb{N} \times \mathbb{N} \mid n = m \text{ oder } n \cdot m = 72\}$.

(2) R und S seien Äquivalenzrelationen auf der Menge A. Untersuchen Sie, ob dann $R \cap S$ und $R \cup S$ wieder Äquivalenzrelationen auf A sind.

(3) Auf den ganzen Zahlen $\mathbb{Z}$ sei folgende Relation gegeben:

$$R_7 = \{(a,b) \in \mathbb{Z} \times \mathbb{Z} \mid a - b \text{ ist durch 7 teilbar}\}.$$

(i) *Zeigen Sie, daß R_7 eine Äquivalenzrelation ist.*

(ii) *Bestimmen Sie die Äquivalenzklasse von 0 bezüglich R_7.*

(iii) *Wie groß ist die Anzahl der Äquivalenzklassen?*

(4) Es seien $f : A \to B$ eine Abbildung und R eine Äquivalenzrelation auf B. Man zeige, daß $Q = \{(a,b) \in A \times A \mid (f(a), f(b)) \in R\}$ eine Äquivalenzrelation auf A ist.

(5) Sei A eine Menge und $(A_i)_{i \in I}$ eine Familie von Teilmengen $A_i \subset A$, die eine Partition *von A bilden, d.h. es gilt:*

$$\bigcup_{i \in I} A_i = A \quad \text{und} \quad A_i \cap A_j = \emptyset \text{ für } i \neq j.$$

Man zeige, daß

$$R = \{(a,b) \in A \times A \mid \text{es gibt ein } i \in I \text{ mit } a \in A_i \text{ und } b \in A_i\}$$

eine Äquivalenzrelation auf A ist.
Läßt sich jede Äquivalenzrelation auf A in dieser Form darstellen?

5 Ordnungsrelationen

Auch in diesem Paragraphen wollen wir uns mit Relationen mit besonderen Eigenschaften befassen, den *Ordnungsrelationen*. Diese sind für die Analysis von größerer Bedeutung als für die lineare Algebra. Wir wollen hier hauptsächlich jene Punkte herausarbeiten, die für uns von Interesse sein werden. Zunächst die Definition:

5.1 Definition
Sei R eine Relation auf A (d.h. $R \subset A \times A$).

(1) R heißt *antisymmetrisch*, wenn $R \cap R^{-1} \subset \Delta_A$ (vgl. 2.2(iii)), d.h. wenn für alle $a, b \in A$ gilt: $(a, b) \in R$ und $(b, a) \in R \Rightarrow a = b$.

(2) R heißt *Ordnungsrelation*, wenn R reflexiv, antisymmetrisch und transitiv ist. Man nennt dann A auch eine (durch R) *(teilweise) geordnete Menge*.

Schreibweise: $(a, b) \in R \Leftrightarrow a \leq_R b$ oder auch $a \leq b$.

Eine Ordnungsrelation R, die zugleich Äquivalenzrelation ist, kann nur die Identität sein, denn

$$\left.\begin{array}{ll} \text{symmetrisch} & R = R^{-1} \\ \text{antisymmetrisch} & R \cap R^{-1} \subset \Delta_A \\ \text{reflexiv} & \Delta_A \subset R \end{array}\right\} \Rightarrow \Delta_A = R.$$

Bekannte Beispiele von Ordnungsrelationen sind:

(1) $\leq$ auf $\mathbb{N}$, $\quad R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid y - x \in \mathbb{N} \cup \{0\}\}$.

(2) $\subset$ auf der Potenzmenge einer Menge A:

$$R = \{(U, V) \in \mathcal{P}(A) \times \mathcal{P}(B) \mid U \subset V\}.$$

Im ersten Beispiel beobachten wir als zusätzliche Eigenschaft, daß zwei Elemente $x, y \in \mathbb{N}$ immer vergleichbar sind: $x \leq y$ oder $y \leq x$. Relationen mit dieser Eigenschaft spielen eine besondere Rolle:

5.2 Definition
Eine Ordnungsrelation R auf A heißt *lineare* (oder *totale*) *Ordnung*, wenn

$$R \cup R^{-1} = A \times A,$$

d.h. für je zwei Elemente $a, b \in A$ gilt $(a, b) \in R$ oder $(b, a) \in R$.

Eine symmetrische Relation mit dieser Eigenschaft wäre die Allrelation.

Lineare Ordnungen werden auch *vollständige* oder *totale Ordnungen* genannt. Bei gewöhnlichen Ordnungsrelationen spricht man auch von *teilweisen Ordnungen* oder *Halbordnungen*.

$(\mathbb{N}, \leq)$, $(\mathbb{Q}, \leq)$ und $(\mathbb{Z}, \leq)$ sind Beispiele für lineare Ordnungen,
$(\mathcal{P}(A), \subset)$ ist eine teilweise Ordnung.

Elemente in geordneten Mengen können sich durch folgende Eigenschaften auszeichnen:

5.3 Definition
Sei $(A, \leq)$ eine geordnete Menge, $B \subset A$ eine Teilmenge.

$b \in B$ heißt *größtes Element* von B, wenn für alle $b' \in B$ stets $b' \leq b$.

$b \in B$ heißt *maximales Element* von B, wenn für alle $b' \in B$ gilt:
$b \leq b' \Rightarrow b' = b$, d.h. es gibt kein Element in B, das größer als b ist.

$a \in A$ heißt *obere Schranke* von B, wenn $b \leq a$ für alle $b \in B$.

$a \in A$ heißt *Supremum* oder *obere Grenze* von B in A, wenn a kleinste obere Schranke von B ist, d.h. für alle oberen Schranken a' von B gilt $a \leq a'$.

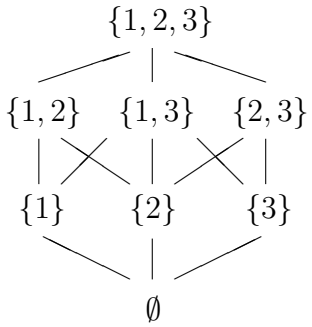
Entsprechend werden *kleinste* und *minimale Elemente*, *untere Schranke* und *Infimum* definiert.

Elemente mit diesen Eigenschaften braucht es nicht zu geben.

Jedes größte Element in B ist auch maximal. Dagegen kann es mehrere maximale Elemente in B geben, die nicht vergleichbar sind.

Wir wollen diese Begriffe an einem einfachen Beispiel erläutern, bei dem wir die Teilmengenbeziehung angeben:

$A =$ Potenzmenge von $\{1, 2, 3\}$:



$\{1, 2, 3\}$ ist größtes Element in A, $\emptyset$ ist kleinstes Element in A.

In $B = A \setminus \{\{1, 2, 3\}, \emptyset\}$ gibt es weder größtes noch kleinstes Element. $\{1, 2\}, \{1, 3\}$ und $\{2, 3\}$ sind maximale Elemente darin, $\{1, 2, 3\}$ ist Supremum davon.

Die Teilmenge $\{\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}\}$ von A ist bezüglich $\subset$ linear geordnet, ebenso die Teilmenge $\{\{2\}, \{2, 3\}\}$. Beide Mengen besitzen ein Supremum (sogar ein größtes Element).

5.4 Definition
Eine nicht-leere, geordnete Menge $(A, \leq)$ heißt *induktiv geordnet*, wenn jede nicht-leere (bezüglich $\leq$) linear geordnete Teilmenge eine obere Schranke besitzt.

Man beachte, daß Teilmengen von induktiv geordneten Mengen nicht wieder induktiv geordnet sein müssen.

Beispiel
Für jede Menge $A \neq \emptyset$ ist die Potenzmenge $(\mathcal{P}(A), \subset)$ induktiv geordnet:
Sei $\mathcal{U} \subset \mathcal{P}(A)$, $\mathcal{U}$ linear geordnet. Dann ist

$$S = \bigcup_{U \in \mathcal{U}} U \in \mathcal{P}(A)$$

Supremum von $\mathcal{U}$: Da $U \subset S$ für alle $U \in \mathcal{U}$, ist S obere Schranke. Für jedes $T \in \mathcal{P}(A)$ mit $U \subset T$ für alle $U \in \mathcal{U}$ gilt $S = \bigcup_{U \in \mathcal{U}} U \subset T$.

Die Bedeutung der induktiv geordneten Mengen liegt darin, daß man in ihnen - mit Hilfe des Auswahlaxioms - die Existenz von maximalen Elementen zeigen kann. Dies führt zu einer vierten Variation des Auswahlaxioms.

5.5 Auswahlaxiom IV: Zornsches Lemma
Jede induktiv geordnete Menge besitzt maximale Elemente.

Der Beweis ist zwar im wesentlichen mit den uns bereits bekannten Begriffen zu führen, erfordert aber doch eine gewisse Vertiefung in die Denkweise der Mengenlehre (vgl. Enderton, Theorem 6M).

Fragen wir uns, was das Zornsche Lemma für die induktiv geordnete Potenzmenge einer Menge A bringt. Nun, für diesen Fall ergibt sich nichts neues, da wir in $\mathcal{P}(A)$ bereits ein maximales Element kennen: A selbst ist sogar größtes Element in $\mathcal{P}(A)$.
Die hauptsächlichen Anwendungen werden sich für uns auf *Teilmengen* von $\mathcal{P}(A)$ beziehen.

Nach diesen Ausführungen zur (abstrakten) Mengenlehre wollen wir uns nun den eigentlichen Objekten unseres Interesses, den algebraischen Strukturen zuwenden.
Eine ausführlichere Darstellung der bislang angesprochenen Problemkreise findet man zum Beipiel in Enderton [4], Halmos [6], oder FU Hagen [5].

5.6 Aufgaben

Seien R, S Ordnungsrelationen auf den Mengen A bzw. B. Zeigen Sie, daß folgende Teilmengen von $(A \times B) \times (A \times B)$ Ordnungsrelationen definieren:

(i) $Q = \{((a_1, b_1), (a_2, b_2)) \,|(a_1 = a_2 \text{ und } (b_1, b_2) \in S)$
$\text{oder } (a_1 \neq a_2 \text{ und } (a_1, a_2) \in R)\}$
(lexikographische Ordnung).

(ii) $P = \{((a_1, b_1), (a_2, b_2)) \mid (a_1, a_2) \in R \text{ und } (b_1, b_2) \in S\}$.

Kapitel 2

Algebraische Grundstrukturen

Bei der Addition oder Multiplikation von Zahlen wird zwei Zahlen a, b eine neue Zahl $a+b$ oder $a \cdot b$ zugeordnet. Man nennt dies eine *Verknüpfung* auf der Menge der (natürlichen, reellen) Zahlen.

Wir werden im nächsten Paragraphen zunächst Mengen mit *einer* Verknüpfung untersuchen (Gruppen, Halbgruppen). Im darauffolgenden Paragraphen wird dann das Zusammenspiel von zwei Verknüpfungen behandelt, wie wir es ebenfalls von den Zahlen her kennen (Ringe, Körper).

6 Halbgruppen und Gruppen

Zur Festlegung der algebraischen Grundbegriffe bedienen wir uns der Sprache und der Methoden der Mengenlehre.

6.1 Definition
Sei A eine nicht-leere Menge. Eine *Verknüpfung* τ auf A ist eine Abbildung von $A \times A$ in A:

$$\tau : A \times A \to A, \quad (a, b) \mapsto a \,\tau\, b.$$

Als Beispiel dazu haben wir schon $+$ und $\cdot$ auf $\mathbb{N}$ kennengelernt.

Ist τ eine Verknüpfung auf A, so ist für $c \in A$ auch $(a \,\tau\, b, c) \in A \times A$ und $(a \,\tau\, b) \,\tau\, c \in A$. Analog läßt sich auch $a \,\tau\, (b \,\tau\, c) \in A$ bilden. Für beliebiges τ muß sich dabei keineswegs das gleiche Element ergeben. Schauen wir uns ein einfaches Beispiel dazu an.

Verknüpfungen auf endlichen Mengen A können wir durch *Verknüpfungstafeln* angeben. Dabei schreibt man die Elemente von A einmal als erste Zeile und einmal als erste Spalte einer quadratischen Tafel und setzt das Element $a \,\tau\, b$ in den Schnitt der Zeile a mit der Spalte b.

Sei z. B. $A = \{a, b\}$ und $\tau : A \times A \to A$ gegeben durch die

Verknüpfungstafel

τ	a	b
a	b	b
b	a	b

$$\begin{aligned}(a\,\tau\,b)\,\tau\,a &= b\,\tau\,a = a\\ a\,\tau\,(b\,\tau\,a) &= a\,\tau\,a = b\end{aligned}$$

Dabei führen die beiden angesprochenen Bildungen zu verschiedenen Ergebnissen. Wir werden uns hier jedoch für solche Verknüpfungen interessieren, bei denen das Ergebnis unabhängig von der Reihenfolge der Ausführung ist.

6.2 Definition
Sei H eine nicht-leere Menge.

Eine Verknüpfung $\tau : H \times H \to H$ heißt *assoziativ*, wenn

$$(a\,\tau\,b)\,\tau\,c = a\,\tau\,(b\,\tau\,c) \text{ für alle } a,\, b,\, c \in H.$$

(H, τ) nennt man dann eine *Halbgruppe*.

(H, τ) heißt *kommutative Halbgruppe*, wenn zudem $a\,\tau\,b = b\,\tau\,a$ für alle $a,\, b \in H$ gilt.

Ist H endlich, so nennt man die Zahl der Elemente von H die *Ordnung von* H.

Ist aus dem Zusammenhang klar, welche Verknüpfung gemeint ist, so schreibt man für (H, τ) nur H.

Es läßt sich durch Induktion zeigen, daß sich in einer Halbgruppe bei jedem endlichen Produkt die Klammern beliebig setzen lassen, also z.B.

$$(a_1\,\tau\,(a_2\,\tau\,a_3))\,\tau\,a_4 = a_1\,\tau\,((a_2\,\tau\,a_3)\,\tau\,a_4).$$

6.3 Beispiele von Halbgruppen

(1) $\mathbb{N}$ bildet mit $\tau = +$ und $\tau = \cdot$ je eine kommutative Halbgruppe.

(2) Für die Abbildung $\tau : A \times A \to A$, $(a, b) \mapsto a$, gilt die Assoziativität, nicht aber die Kommutativität, falls es $a \neq b$ in der Menge A gibt, denn

$$a\,\tau\,b = a \neq b = b\,\tau\,a.$$

(3) Die Potenzmenge $\mathcal{P}(A)$ einer Menge A ist eine Halbgruppe mit $\cap$ (oder $\cup$).

(4) Sei $A \neq \emptyset$ eine Menge, $\mathrm{Abb}(A) := \mathrm{Abb}(A, A)$ die Menge der Abbildungen von A in A. Dann ist $(\mathrm{Abb}(A), \circ)$ eine Halbgruppe mit

$$\mathrm{Abb}(A) \times \mathrm{Abb}(A) \to \mathrm{Abb}(A),\ (f, g) \mapsto g \circ f.$$

Sie ist nicht kommutativ, wenn A mehr als ein Element enthält:
Seien $c \neq b \in A$, $f : A \to A$, $a \mapsto b$ für alle $a \in A$,
$g : A \to A$, $a \mapsto c$ für alle $a \in A$.
Dann gilt $g \circ f(a) = c \neq b = f \circ g(a)$, also $g \circ f \neq f \circ g$.

(5) Ist $A \neq \emptyset$ eine Menge und (H, τ) eine Halbgruppe, dann wird $\mathrm{Abb}(A, H)$ zu einer Halbgruppe durch

$$f\ \tau'\ g : A \to H,\ a \mapsto f(a)\ \tau\ g(a) \text{ für alle } a \in A.$$

(6) Ist $(H_i, \tau_i)_{i \in I}$ eine Familie von Halbgruppen, so läßt sich auch auf dem kartesischen Produkt $\prod_I H_i$ eine Verknüpfung τ definieren, die $(\prod_I H_i, \tau)$ zu einer Halbgruppe macht:

$$(a_i)_I\ \tau\ (b_i)_I := (a_i\ \tau_i\ b_i)_I.$$

Sehen wir uns an, welche besonderen Eigenschaften Elemente einer Halbgruppe haben können:

6.4 Definition
Sei (H, τ) eine Halbgruppe.

(1) Ein Element $e \in H$ heißt

rechtsneutrales Element, wenn $a\ \tau\ e = a$ für alle $a \in H$,

linksneutrales Element, wenn $e\ \tau\ a = a$ für alle $a \in H$,

neutrales Element, wenn $a\ \tau\ e = e\ \tau\ a = a$ für alle $a \in H$.

(2) Hat H ein neutrales Element e, dann heißt ein Element $b \in H$

rechtsinvers zu $a \in H$, wenn $a\ \tau\ b = e$,

linksinvers zu $a \in H$, wenn $b\ \tau\ a = e$,

invers zu $a \in H$, wenn $a\ \tau\ b = b\ \tau\ a = e$. Man schreibt dafür $b =: a^{-1}$.

Wir wollen dazu gleich einige Beobachtungen festhalten:

(1) Eine Halbgruppe hat höchstens ein neutrales Element: Sind e und e' neutrale Elemente, dann gilt $e = e\ \tau\ e' = e'$.

(2) Zu einem $a \in H$ gibt es höchstens ein inverses Element: Sind b und b' invers zu a, dann gilt: $b = b\ \tau\ e = b\ \tau\ (a\ \tau\ b') = (b\ \tau\ a)\ \tau\ b' = e\ \tau\ b' = b'$.

(3) Ist b invers zu a und b' invers zu a', dann ist $b'\ \tau\ b$ invers zu $a\ \tau\ a'$.

Prüfen wir, ob es in den Beispielen von 6.3 solche Elemente gibt:

(i) $(\mathbb{N}, +)$: 0 ist neutrales Element; $(\mathbb{N}, \cdot)$: 1 ist neutrales Element.
Nur zu 0 bzw. 1 gibt es inverse Elemente.

(ii) $(a, b) \mapsto a$: Jedes Element ist rechtsneutral; es gibt kein linksneutrales Element, wenn A mehr als ein Element enthält.

(iii) $(\mathcal{P}(A), \cap)$: A ist neutrales Element;
$(\mathcal{P}(A), \cup)$: $\emptyset$ ist neutrales Element.

(iv) $(\mathrm{Abb}(A), \circ)$: id_A ist neutrales Element. Zu $f \in \mathrm{Abb}(A)$ gibt es ein Inverses, wenn f invertierbar ist (Umkehrabbildung).

(v) $(\mathrm{Abb}(A, H), \tau')$: Hat H ein neutrales Element e, so ist das neutrale Element von $\mathrm{Abb}(A, H)$ die konstante Abbildung $\widetilde{e} : A \to H$, $a \mapsto e$.

(vi) $(\prod_I H_i, \tau)$: Hat jedes H_i ein neutrales Element e_i, so ist $(e_i)_{i \in I}$ neutrales Element in $\prod_I H_i$.

Wie an den Beispielen zu sehen ist, braucht es in Halbgruppen weder neutrale noch inverse Elemente zu geben. Halbgruppen, in denen es diese Elemente immer gibt, sind von besonderer Bedeutung:

6.5 Definition
Eine Halbgruppe (H, τ) heißt *Gruppe*, wenn gilt:

(1) Es gibt ein neutrales Element $e \in H$;

(2) Zu jedem $a \in H$ gibt es ein Inverses.

Ist die Verknüpfung τ kommutativ, so nennt man (H, τ) eine *kommutative* oder auch *abelsche Gruppe* (in Erinnerung an den norwegischen Mathematiker N.H. Abel).

Die in 6.5 gestellten Forderungen (1) und (2) lassen sich durch die Lösbarkeit bestimmter Gleichungen ausdrücken. Es gilt:

6.6 Satz
Für eine Halbgruppe H sind folgende Eigenschaften äquivalent:

(a) *(H, τ) ist eine Gruppe;*

(b) *Zu je zwei Elementen a, $b \in H$ gibt es genau ein $x \in H$ mit $a \tau x = b$ und genau ein $y \in H$ mit $y \tau a = b$.*

Beweis: $(a) \Rightarrow (b)$ Sei (H, τ) eine Gruppe, a, $b \in H$. Für $x = a^{-1} \tau b$ gilt

$$a \tau x = a \tau (a^{-1} \tau b) = (a \tau a^{-1}) \tau b = b.$$

Zur Eindeutigkeit von x: Gelte $a \tau x = b = a \tau x'$ für x, $x' \in H$. Dann folgt $x = a^{-1} \tau (a \tau x) = a^{-1} \tau b$ und $x' = a^{-1} \tau (a \tau x') = a^{-1} \tau b$.
Analog erhält man die Lösung von $y \tau a = b$.

$(b) \Rightarrow (a)$ Es gelte (b). Dann hat $a \tau x = a$ eine Lösung $e \in H$, also $a \tau e = a$. Wir zeigen, daß dieses e neutrales Element ist.

Sei $b \in H$ beliebig und $c \in H$ mit $c \tau a = b$. Dann gilt

$$b \tau e = (c \tau a) \tau e = c \tau (a \tau e) = c \tau a = b,$$

also ist e rechts neutral. Betrachte nun ein $b' \in H$ mit $e \tau b' = b$. Dann gilt

$$e \tau b = e \tau (e \tau b') = (e \tau e) \tau b' = e \tau b' = b.$$

Somit ist e auch links neutral.

Es bleibt noch, ein Inverses zu $a \in H$ zu finden. Nach Voraussetzung gibt es zu a, $e \in H$ Elemente b, $b' \in H$ mit $a \tau b = e$, $b' \tau a = e$. Daraus ergibt sich

$$b' = b' \tau e = b' \tau (a \tau b) = (b' \tau a) \tau b = e \tau b = b.$$

Also ist b invers zu a. □

Als Beispiel für Gruppen kennen wir etwa die ganzen Zahlen $(\mathbb{Z}, +)$ und die rationalen Zahlen $(\mathbb{Q}, +)$ bzw. $(\mathbb{Q} \setminus \{0\}, \cdot)$.

Die in 6.3 angegebenen Beispiele für Halbgruppen sind allesamt keine Gruppen. Einige von ihnen lassen sich in Gruppen einbetten, etwa $(\mathbb{N}, +)$ und $(\mathbb{N} \setminus \{0\}, \cdot)$. Die meisten davon enthalten (wenn auch manchmal triviale) Gruppen. Solche Unterstrukturen sind von großem Interesse:

6.7 Definition
Sei (H, τ) eine Halbgruppe.

Eine Teilmenge $U \subset H$ heißt *Unterhalbgruppe*, wenn für a, $b \in U$ auch $a \tau b \in U$ gilt. Man sagt dazu, U *ist abgeschlossen gegenüber* τ. (U, τ) ist Halbgruppe.

Besitzt (H, τ) ein neutrales Element, dann heißt $U \subset H$ *Untergruppe*, wenn es Unterhalbgruppe ist und zu jedem $a \in U$ auch $a^{-1} \in U$ gilt.

(U, τ) ist dann eine Gruppe mit neutralem Element $e \in U$.

In jeder Halbgruppe (H, τ) mit neutralem Element e ist die Menge $H^\times$ der invertierbaren Elemente in H eine Untergruppe, denn
$e \in H^\times$, $a \in H^\times$ (also invertierbar) $\Rightarrow a^{-1} \in H^\times$ und
a, $b \in H^\times \Rightarrow a \tau b \in H^\times$.

Sehen wir uns wieder die in 6.3 gegebenen Beispiele daraufhin an:

(i) In $(\mathbb{N}, +)$ und $(\mathbb{N}, \cdot)$ ist $2\mathbb{N}$ (= gerade Zahlen) eine Unterhalbgruppe. $\{1\}$ ist Untergruppe von $(\mathbb{N}, \cdot)$.

(ii) $(a, b) \mapsto a$: $\{a\} \subset A$ ist wohl Gruppe, aber nicht Untergruppe, da A kein neutrales Element hat.

(iii) $\{A\}$ ist Untergruppe von $(\mathcal{P}(A), \cap)$, $\{\emptyset\}$ Untergruppe von $(\mathcal{P}(A), \cup)$. $\{\emptyset, A\}$ ist Unterhalbgruppe von $(\mathcal{P}(A), \cap)$ und $(\mathcal{P}(A), \cup)$, aber in keinem Fall Untergruppe ($\emptyset \cap ? = A$, $A \cup ? = \emptyset$).

(iv) $(\mathrm{Abb}(A), \circ)$: Die surjektiven, die injektiven sowie die konstanten Abbildungen bilden jeweils Unterhalbgruppen. Die bijektiven Abbildungen sind eine Untergruppe.

(v) $(\mathrm{Abb}(A, H), \tau')$: Ist U Untergruppe von H, dann ist $\mathrm{Abb}(A, U)$ Untergruppe von $\mathrm{Abb}(A, H)$.

Als Abbildungen zwischen Halbgruppen interessieren uns vor allem solche, die mit den Verknüpfungen verträglich sind:

6.8 Definition
Seien (H_1, τ_1) und (H_2, τ_2) Halbgruppen. Eine Abbildung $f : H_1 \to H_2$ heißt (Halbgruppen-)*Homomorphismus*, wenn für alle $a, b \in H_1$ gilt

$$f(a\,\tau_1\,b) = f(a)\,\tau_2\,f(b)$$

Man nennt dann $f : H_1 \to H_2$ einen

Monomorphismus,	wenn f injektiv ist,
Epimorphismus,	wenn f surjektiv ist,
Isomorphismus,	wenn f bijektiv ist,
Endomorphismus,	wenn $H_1 = H_2$,
Automorphismus,	wenn $H_1 = H_2$ und f Isomorphismus ist.

Sind H_1 und H_2 Gruppen, so nennt man einen Halbgruppen-Homomorphismus $f : H_1 \to H_2$ auch *Gruppen-Homomorphismus*.

Die Mengen der Homo-, Endo- bzw. Automorphismen bezeichnet man mit $\mathrm{Hom}(H_1, H_2)$, $\mathrm{End}(H_1)$ bzw. $\mathrm{Aut}(H_1)$. $\mathrm{End}(H_1)$ und $\mathrm{Aut}(H_1)$ sind offensichtlich Unterhalbgruppen von $\mathrm{Abb}(H_1, H_1)$.

Schauen wir uns einige Beispiele zu diesen Begriffen an.

6.9 Beispiele von Homomorphismen

(1) Sei $(G, \cdot)$ eine Gruppe. Zu $a \in G$ definieren wir

$$I_a : G \to G, \quad x \mapsto a \cdot x \cdot a^{-1}.$$

I_a ist ein Homomorphismus, denn

$$\begin{aligned} I_a(x \cdot y) &= a \cdot (x \cdot y) \cdot a^{-1} \\ &= a \cdot x \cdot (a^{-1} \cdot a) \cdot y \cdot a^{-1} \quad = \quad I_a(x) \cdot I_a(y). \end{aligned}$$

I_a ist injektiv, denn aus $a \cdot x \cdot a^{-1} = a \cdot y \cdot a^{-1}$ folgt $x = y$.
I_a ist surjektiv, denn für $b \in G$ gilt

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

Man nennt I_a einen *inneren Automorphismus*.

(2) Seien H_1, H_2 Halbgruppen und e_2 neutrales Element in H_2. Dann ist

$$H_1 \to H_2, \quad x \mapsto e_2 \text{ für alle } x \in H_1,$$

ein (trivialer) Homomorphismus.

(3) Die *Exponentialfunktion*

$$\exp : (\mathbb{R}, +) \to (\mathbb{R}_{>0}, \cdot), \quad x \mapsto e^x,$$

ist ein Homomorphismus, da $e^{x+y} = e^x e^y$.

Auch der *Logarithmus*

$$\log : (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +), \quad x \mapsto \log(x),$$

ist ein Homomorphismus, denn $\log(xy) = \log(x) + \log(y)$.

(4) Sei $(H, \cdot)$ eine Halbgruppe. Dann ist die Linksmultiplikation mit einem $a \in H$ eine Abbildung von H in sich,

$$L_a : H \to H, \quad x \mapsto a \cdot x.$$

Damit erhalten wir einen Homomorphismus

$$\lambda : (H, \cdot) \to (\mathrm{Abb}(H, H), \circ), \quad a \mapsto L_a.$$

Nach Definition gilt nämlich

$$L_{a \cdot b}(x) = (a \cdot b) \cdot x = a \cdot (b \cdot x) = L_a(L_b(x)) = L_a \circ L_b(x),$$

also $a \cdot b \mapsto L_a \circ L_b$. Ist a invertierbar, dann ist L_a bijektiv.

Ist H eine Gruppe, so ist λ monomorph, und H ist isomorph zu einer Untergruppe der bijektiven Abbildungen von H in sich.

Ähnliche Bildungen kann man natürlich auch mit Rechtsmultiplikationen durchführen. Man beachte, daß sich dann bei λ die Reihenfolge der Multiplikation im Bild umdreht.

Halten wir weitere Eigenschaften von Homomorphismen fest:

6.10 Satz
Sei $f : (H_1, \tau_1) \to (H_2, \tau_2)$ ein Homomorphismus von Halbgruppen.

(1) *Ist U_1 eine Unterhalbgruppe von H_1, dann ist $f(U_1)$ eine Unterhalbgruppe von H_2.*

(2) *Ist U_2 eine Unterhalbgruppe von H_2, dann ist $f^{-1}(U_2) = \emptyset$ oder eine Unterhalbgruppe von H_1.*

Beweis: (1) Sei $a_2, b_2 \in f(U_1) \subset H_2$. Dann gibt es $a_1, b_1 \in H_1$ mit $f(a_1) = a_2$, $f(b_1) = b_2$, und es gilt

$$a_2 \,\tau_2\, b_2 = f(a_1) \,\tau_2\, f(b_1) = f(a_1 \,\tau_1\, b_1) \in f(U_1).$$

(2) Seien $f^{-1}(U_2) \neq \emptyset$ und $a_1, b_1 \in f^{-1}(U_2)$, d.h. $f(a_1), f(b_1) \in U_2$, und

$$f(a_1 \,\tau_1\, b_1) = f(a_1) \,\tau_2\, f(b_1) \in U_2,$$

also $a_1 \,\tau_1\, b_1 \in f^{-1}(U_2)$. □

Wenn auch die Kennzeichnung von Homomorphismen zwischen Halbgruppen die gleiche ist wie zwischen Gruppen, so erzwingt die Gruppenstruktur doch zusätzliche Eigenschaften der Homomorphismen:

6.11 Satz
Seien $f : (G_1, \tau_1) \to (G_2, \tau_2)$ ein Gruppen-Homomorphismus, $e_1 \in G_1$ und $e_2 \in G_2$ die neutralen Elemente. Dann gilt:

(1) *$f(e_1) = e_2$ und $f(a^{-1}) = (f(a))^{-1}$ für alle $a \in G_1$.*

(2) *Ist U_1 Untergruppe von G_1, dann ist $f(U_1)$ Untergruppe von G_2.*

(3) *Ist U_2 Untergruppe von G_2, dann ist $f^{-1}(U_2)$ Untergruppe von G_1.*

Beweis: (1) Aus $e_1 \,\tau_1\, e_1 = e_1$ folgt $f(e_1) = f(e_1) \,\tau_2\, f(e_1)$ und

$$e_2 = f(e_1) \,\tau_2\, (f(e_1))^{-1} = f(e_1).$$

(2) und (3) folgen damit aus Satz 6.10. □

Eine wichtige Eigenschaft von Homomorphismen ist die Tatsache, daß ihre Komposition (als Abbildungen) wieder strukturverträglich ist:

6.12 Satz
Seien $f : H_1 \to H_2$, $g : H_2 \to H_3$ Homomorphismen von Halbgruppen.

(1) *Dann ist auch $g \circ f : H_1 \to H_3$ ein Homomorphismus.*

(2) *Ist f ein Isomorphismus, dann ist auch die Umkehrabbildung $f^{-1} : H_2 \to H_1$ ein Isomorphismus.*

Beweis: (1) ist leicht nachzuprüfen.

(2) Wir haben $f(f^{-1}(x\ \tau_2\ y)) = x\ \tau_2\ y = f(f^{-1}(x)\ \tau_1\ f^{-1}(y))$. Wegen der Injektivität von f gilt damit auch

$$f^{-1}(x\ \tau_2\ y) = (f^{-1}(x)\ \tau_1\ f^{-1}(y)).$$

□

Dies impliziert insbesondere, daß die Komposition von Endomorphismen wieder Endomorphismen ergibt, also

6.13 Korollar
Sei H eine Halbgruppe. Dann ist $\mathrm{End}(H)$ *eine Unterhalbgruppe von* $\mathrm{Abb}(H,H)$*, und die Automorphismen* $\mathrm{Aut}(H)$ *bilden die Gruppe der invertierbaren Elemente in* $\mathrm{End}(H)$*.*

Ist $f : G_1 \to G_2$ ein Gruppen-Homomorphismus, so ist das Urbild **jeder** Untergruppe von G_2 eine Untergruppe von G_1. Speziell ist also auch das Urbild von $\{e_2\} \subset G_2$ eine Untergruppe in G_1. Diese ist für den Homomorphismus f von großer Bedeutung. Man gibt ihr daher einen besonderen Namen:

6.14 Definition
Sei $f : G_1 \to G_2$ ein Gruppen-Homomorphismus, $e_2 \in G_2$ das neutrale Element. Dann nennt man die Untergruppe $f^{-1}(e_2)$ von G_1 den *Kern von f*, also

$$\mathrm{Kern}\, f = \{a \in G_1 \mid f(a) = e_2\}$$

Die wichtigsten Eigenschaften davon fassen wir zusammen in:

6.15 Satz
Seien $f : G_1 \to G_2$ ein Homomorphismus von Gruppen und e_1, e_2 die neutralen Elemente von G_1 bzw. G_2. Für den Kern von f gilt:

(1) *Für $a, b \in G_1$ ist genau dann $f(a) = f(b)$, wenn $ab^{-1} \in \mathrm{Kern}\, f$ (oder $a^{-1}b \in \mathrm{Kern}\, f$).*

(2) *Für $c \in \mathrm{Kern}\, f$ ist $aca^{-1} \in \mathrm{Kern}\, f$, also $a(\mathrm{Kern}\, f)a^{-1} \subset \mathrm{Kern}\, f$ für alle $a \in G_1$.*

(3) *Für alle $a \in G_1$ gilt $a\,\mathrm{Kern}\, f = (\mathrm{Kern}\, f)a$.*

(4) *f ist genau dann injektiv (= monomorph), wenn $\mathrm{Kern}\, f = \{e_1\}$ ist.*

Beweis: (1) Ist $f(a) = f(b)$, dann ist

$$e_2 = f(a)(f(b)^{-1}) = f(a)f(b^{-1}) = f(ab^{-1}),$$

also $ab^{-1} \in \operatorname{Kern} f$. Umgekehrt folgt aus $f(ab^{-1}) = e_2$ auch

$$f(b) = e_2 f(b) = f(a)f(b^{-1})f(b) = f(a).$$

(2) Für $f(c) = e_2$ gilt $f(aca^{-1}) = f(a)f(c)f(a^{-1}) = f(aa^{-1}) = e_2$.

(3) Betrachte $b = ak$, $k \in \operatorname{Kern} f$. Dann ist $ba^{-1} = aka^{-1} \in \operatorname{Kern} f$ und $b \in (\operatorname{Kern} f)a$ nach (2).

(4) Gilt $\operatorname{Kern} f = \{e_1\}$, dann gilt in (1) $ab^{-1} = e_1$, also $a = b$. □

6.16 Definition
Eine Untergruppe U einer Gruppe (G, τ) heißt *Normalteiler* oder *normale Untergruppe*, wenn

$$a \,\tau\, U \,\tau\, a^{-1} \subset U \text{ für alle } a \in G.$$

Es ist leicht zu sehen, daß eine Untergruppe $U \subset G$ genau dann Normalteiler ist, wenn

$$a \,\tau\, U = U \,\tau\, a \text{ für alle } a \in G.$$

In abelschen Gruppen ist natürlich jede Untergruppe Normalteiler.

Zu jeder Untergruppe U einer Gruppe (G, τ) wird eine Äquivalenzrelation auf G definiert durch

$$R_U = \{(a, b) \in G \times G \mid b^{-1} \,\tau\, a \in U\}.$$

Die Äquivalenzklasse zu einem $a \in G$ ist dann $R_U(a) = [a] = a \,\tau\, U$.
Wir bezeichnen die Menge der Äquivalenzklassen mit G/U.

Die Normalteiler einer Gruppe zeichnen sich nun dadurch unter den gewöhnlichen Untergruppen aus, daß auf G/U wieder eine Gruppenstruktur eingeführt werden kann.

6.17 Faktorgruppen
Sei (G, τ) eine Gruppe mit neutralem Element e. Sei U Normalteiler in G, so wird die Menge der Äquivalenzklassen G/U zu einer Gruppe durch die Verknüpfung

$$[a] \,\tau'\, [b] := [a \,\tau\, b] \text{ für } a, b \in G.$$

Dabei ist $[e]$ das neutrale Element in $(G/U, \tau')$, und $[a^{-1}]$ ist das Inverse zu $[a]$.

Beweis: Es ist zu zeigen, daß diese Festlegung unabhängig ist von der Auswahl der Repräsentanten. Dazu betrachten wir $a' \in [a]$ und $b' \in [b]$, also $a' = a \,\tau\, k$ und $b' = a \,\tau\, h$ für geeignete $k, h \in U$. Hierfür gilt

$$\begin{aligned}[a' \,\tau\, b'] \;=\; [a \,\tau\, k \,\tau\, b \,\tau\, h] \;&=\; a \,\tau\, k \,\tau\, b \,\tau\, h \,\tau\, U \;=\; a \,\tau\, k \,\tau\, (b \,\tau\, U) \\ &=\; a \,\tau\, k \,\tau\, (U \,\tau\, b) \;=\; a \,\tau\, U \,\tau\, b \\ &=\; a \,\tau\, b \,\tau\, U \;=\; [a \,\tau\, b].\end{aligned}$$

Also definiert τ' tatsächlich eine Verknüpfung auf G/U.
Es ist klar, daß $[e]$ neutrales Element ist.
Aus $[a]\,\tau'\,[a^{-1}] = [a \,\tau\, a^{-1}] = [e]$ folgt $[a]^{-1} = [a^{-1}]$. □

Nach 6.15 ist der Kern eines Gruppen-Homomorphismus ein Normalteiler in seiner Quelle.

In 4.5 haben wir zu einer Abbildung $f : G_1 \to G_2$ eine Äquivalenzrelation auf G_1 definiert. Ist f ein Gruppen-Homomorphismus, so haben wir nach 6.15

$$\begin{aligned}R_f \;&=\; \{(a,b) \in G_1 \times G_1 \mid f(a) = f(b)\} \\ &=\; \{(a,b) \in G_1 \times G_1 \mid a^{-1}b \in \operatorname{Kern} f\}.\end{aligned}$$

Damit ist $b \in [a]$ (also $f(b) = f(a)$) genau dann, wenn $b \in a \operatorname{Kern} f$, d.h.

$$a \operatorname{Kern} f = [a] = (\operatorname{Kern} f)a$$

ist die Äquivalenzklasse von $a \in G_1$ bzgl. R_f. Für die Menge der Äquivalenzklassen bedeutet dies

$$G_1/R_f = \{a \operatorname{Kern} f \mid a \in G_1\} =: G_1/\operatorname{Kern} f.$$

Wie wir oben gezeigt haben, haben wir auf $G_1/\operatorname{Kern} f$ eine Gruppenstruktur. Damit kommen wir zu einem der wichtigen Sätze der Gruppentheorie:

6.18 Homomorphiesatz für Gruppen
Sei $f : G_1 \to G_2$ ein Homomorphismus von Gruppen. Dann ist die kanonische Projektion

$$p_f : G_1 \to G_1/\operatorname{Kern} f, \quad a \mapsto [a],$$

ein Gruppenepimorphismus, und es gibt genau einen Monomorphismus

$$\overline{f} : G_1/\operatorname{Kern} f \to G_2$$

mit $f = \overline{f} \circ p_f$, d.h. wir haben das kommutative Diagramm

$$\begin{array}{ccc} G_1 & \xrightarrow{\;f\;} & G_2 \\ {\scriptstyle p_f}\searrow & & \nearrow{\scriptstyle \overline{f}} \\ & G_1/\operatorname{Kern} f\,. & \end{array}$$

Beweis: Das kommutative Diagramm kennen wir bereits aus 4.6, und wir wissen von dort, daß p_f surjektiv und $\overline{f}$ injektiv ist. Es bleibt zu zeigen, daß p_f und $\overline{f}$ Homomorphismen sind. Dies sieht man aus den Gleichungen

$$\begin{aligned} p_f(a\,\tau_1\,b) &= [a\,\tau_1\,b] = [a]\,\tau_1'\,[b] = p_f(a)\,\tau_1'\,p_f(b), \\ \overline{f}([a]\,\tau_1'\,[b]) &= \overline{f}([a\,\tau_1\,b]) = f(a\,\tau_1\,b) = f(a)\,\tau_2\,f(b) = \overline{f}([a])\,\tau_2\,\overline{f}([b]). \end{aligned}$$

□

Neu gegenüber Satz 4.6 ist im Homomorphiesatz, daß man bei einem Gruppen-Homomorphismus $f : G_1 \to G_2$ auf $G_1/\operatorname{Kern} f$ eine Gruppenstruktur hat und die auftretenden Abbildungen Homomorphismen sind.

Wie schon angemerkt, ist in abelschen Gruppen jede Untergruppe Normalteiler. Sehen wir uns dazu einen einfachen Fall an.

6.19 Beispiel
Betrachte die Äquivalenzrelation auf $\mathbb{Z}$, die durch die Untergruppe $7\mathbb{Z}$ bestimmt ist, also

$$\begin{aligned} R_7 &= \{(a,b) \in \mathbb{Z} \times \mathbb{Z} \mid a - b \text{ ist durch 7 teilbar}\} \\ &= \{(a,b) \in \mathbb{Z} \times \mathbb{Z} \mid (-b) + a \in 7\mathbb{Z}\}. \end{aligned}$$

Dann ist $\mathbb{Z}/7\mathbb{Z}$ eine additive Gruppe mit $[a] + [b] = [a + b]$.
Die Projektion $p_7 : \mathbb{Z} \to \mathbb{Z}/7\mathbb{Z}$, $z \mapsto [z]$ ist ein Epimorphismus.
Wie nach Satz 3.8 zu erwarten ist, gibt es eine Abbildung von Mengen, etwa

$$q : \mathbb{Z}/7\mathbb{Z} \to \mathbb{Z}, \quad [z] \mapsto z_0 \in [z], \text{ mit } z_0 < 7,$$

für die $p_7 \circ q = \mathrm{id}_{\mathbb{Z}/7\mathbb{Z}}$ gilt. Man beachte aber, daß q *kein* Homomorphismus (bzgl. $+$) ist, denn es gilt z.B. $q([5]) = 5$, $q([6]) = 6$ und

$$q([5] + [6]) = q([5 + 6]) = 4 \neq 11 = 5 + 6 = q([5]) + q([6]).$$

In der Tat gibt es keinen Homomorphismus $g : \mathbb{Z}/7\mathbb{Z} \to \mathbb{Z}$ mit $p_7 \circ g = \mathrm{id}_{\mathbb{Z}/7\mathbb{Z}}$. Dies zeigt uns, daß nicht alle Sätze über Abbildungen von Mengen auch für Homomorphismen von Gruppen gelten.

Analog zur Situation bei Abbildungen haben wir auch hier:

6.20 Satz *(Universelle Eigenschaft des Produkts von (Halb-) Gruppen)*
Sei $(H_i, \tau_i)_I$ eine Familie von Halbgruppen und $(\prod_I H_i, \tau')$ das Produkt der H_i mit der Verknüpfung

$$(a_i)_I\,\tau'\,(b_i)_I := (a_i\,\tau_i\,b_i)_I.$$

(1) *Dann sind die Projektionen*

$$\pi_k : \prod_I H_i \to H_k, \quad (a_i)_{i\in I} \mapsto a_k,$$

(Halbgruppen-)Epimorphismen.

(2) *Ist H' eine Halbgruppe und $(f_i : H' \to H_i)_{i\in I}$ eine Familie von Homomorphismen, so gibt es genau einen Homomorphismus $f : H' \to \prod_I H_i$ mit*

$$\pi_k \circ f = f_k \quad \text{für alle } k \in I,$$

d.h. folgende Diagramme sind kommutativ:

$$\begin{array}{ccc} H' & \xrightarrow{f_k} & H_k \\ {\scriptstyle f}\searrow & & \nearrow{\scriptstyle \pi_k} \\ & \prod_I H_i. & \end{array}$$

□

Wir wollen schließlich die neu erlernten Begriffe anhand einer wichtigen Gruppe ansehen, auf die wir später zurückgreifen werden.

6.21 Permutationsgruppen

Die bijektiven Abbildungen einer (endlichen) Menge A in sich nennt man *Permutationen* von A.

Die Gruppe aller Permutationen von A heißt *Permutationsgruppe* oder *symmetrische Gruppe* von A. Sie hat die Identität auf A als neutrales Element.

Bei diesen Betrachtungen können wir ohne Einschränkung der Allgemeinheit eine endliche Menge mit n Elementen durch

$$\{1, 2, \ldots, n\}$$

vorgeben. Die symmetrische Gruppe dazu bezeichnet man mit $\mathcal{S}_n$.

Eine Permutation σ davon kann man beschreiben durch

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}.$$

Permutationen, die lediglich ein Paar von Elementen vertauschen und den Rest festlassen, nennt man *Transpositionen.*

Für jede Transposition τ gilt $\tau^2 = \mathrm{id}$. Ein Beispiel dafür ist etwa

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}, \quad \tau^2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \mathrm{id}\,.$$

Transpositionen sind sozusagen die Bausteine der Permutationen, denn:

6.22 Satz
Jede Permutation ist als Produkt von Transpositionen darstellbar.

Beweis: Es ist klar, daß eine Permutation genau dann die Identität ist, wenn sie n Elemente unverändert läßt.

Sei nun σ eine Permutation, die $r < n$ Elemente festläßt. Für ein $q < n$ mit $\sigma(q) \neq q$ definieren wir eine Transposition τ_1, die q mit $\sigma(q)$ vertauscht und den Rest festläßt. Die Komposition $\tau_1 \circ \sigma$ hat dann $r+1$ Fixelemente.

Durch weitere Wahl von Transpositionen $\tau_2, \tau_3, ..., \tau_k$ erhält man schließlich n Fixpunkte, also

$$\tau_k \circ \tau_{k-1} \cdots \tau_1 \circ \sigma = \mathrm{id} \qquad \text{und } \sigma^{-1} = \tau_k \circ \cdots \circ \tau_1.$$

Da jede Permutation invertierbar ist, folgt daraus die Behauptung. □

6.23 Definition
Für eine Permutation $\sigma \in \mathcal{S}_n$ definieren wir das *Signum* durch

$$\operatorname{sgn}\sigma = \prod_{i<j}[\sigma(j) - \sigma(i)] \Big/ \prod_{i<j}[j-i]\,.$$

Wie leicht zu sehen ist, stehen über und unter dem Bruchstrich bis auf das Vorzeichen die gleichen Faktoren. Also ist $\operatorname{sgn}\sigma$ gleich $+1$ oder -1.

Die Permutation σ heißt $\begin{cases} \textit{gerade}, & \text{wenn } \operatorname{sgn}\sigma = 1 \\ \textit{ungerade}, & \text{wenn } \operatorname{sgn}\sigma = -1 \end{cases}$.

Für eine Transposition $\tau \in \mathcal{S}_n$ gilt offensichtlich $\operatorname{sgn}\tau = -1$.

Als wichtige Eigenschaft von Signum stellen wir fest:

6.24 Satz
Für je zwei Permutationen σ, $\tau \in \mathcal{S}_n$ gilt

$$\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}\sigma \cdot \operatorname{sgn}\tau.$$

Beweis: Nehmen wir zunächst an, τ sei eine Transposition, welche die Zahlen $k < l$ vertauscht. Ist $i < j$, so haben wir

$$\tau(j) < \tau(i) \text{ genau dann, wenn } i = k,\ j = l,$$

oder – äquivalent dazu –

$$\tau(i) < \tau(j) \text{ genau dann, wenn } (i,j) \neq (k,l).$$

Für jedes $\sigma \in \mathcal{S}_n$ ergibt sich damit

$$\begin{aligned}
&\prod_{i<j}[\sigma \circ \tau(j) - \sigma \circ \tau(i)] \\
&\quad = [\sigma \circ \tau(l) - \sigma \circ \tau(k)] \prod_{i<j,\tau(i)<\tau(j)} [\sigma \circ \tau(j) - \sigma \circ \tau(i)] \\
&\quad = [\sigma \circ \tau(l) - \sigma \circ \tau(k)] \prod_{\tau(i)<\tau(j),i<j} [\sigma(j) - \sigma(i)] \\
&\quad = -\prod_{i<j}[\sigma(j) - \sigma(i)].
\end{aligned}$$

Daraus ersieht man $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}\sigma \cdot \operatorname{sgn}\tau$.

Sei nun τ eine beliebige Permutation. Nach 6.22 können wir

$$\tau = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$$

mit lauter Transpositionen τ_i schreiben. Damit ist nun leicht zu sehen, daß auch für beliebige τ die gewünschte Beziehung gilt. □

Es gibt noch andere Möglichkeiten, das Signum einer Permutation zu interpretieren. Für $\sigma \in \mathcal{S}_n$ bezeichne $s(\sigma)$ die Anzahl der Paare (i, j) mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$. Diese Paare bezeichnet man als *Fehlstände* von σ. Man kann sich überlegen, daß

$$\operatorname{sgn}\sigma = (-1)^{s(\sigma)}.$$

Eine weitere Beschreibung von Signum erhält man durch die Zerlegung von $\sigma \in \mathcal{S}_n$ in ein Produkt von Transpositionen τ_i, also

$$\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k,$$

mit einer ganzen Zahl k. Solche Zerlegungen sind zwar keineswegs eindeutig, und somit ist auch k nicht eindeutig bestimmt, dennoch folgt aus 6.24

$$\operatorname{sgn}\sigma = (-1)^k.$$

Damit können wir zusammenfassen:

6.25 Alternierende Gruppe
Die geraden Permutationen in $\mathcal{S}_n$ sind solche, die eine gerade Anzahl von Fehlständen haben, oder – gleichbedeutend – die als Produkt einer geraden Anzahl von Transpositionen darstellbar sind. Sie bilden eine Untergruppe und – nach 6.24 – sogar einen Normalteiler in $\mathcal{S}_n$.

Man nennt sie die *alternierende Gruppe* $\mathcal{A}_n$.

6.26 Beispiele

(1) Berechnung von Signum in $\mathcal{S}_5$.

$$\begin{array}{rcl} \sigma_1 & = & \begin{pmatrix} 1\,2\,3\,4\,5 \\ 3\,5\,4\,1\,2 \end{pmatrix}, \qquad s(\sigma_1) = 7, \qquad \operatorname{sgn}\sigma_1 = -1; \\ \sigma_2 & = & \begin{pmatrix} 1\,2\,3\,4\,5 \\ 2\,5\,3\,1\,4 \end{pmatrix}, \qquad s(\sigma_2) = 5, \qquad \operatorname{sgn}\sigma_2 = -1; \\ \sigma_1 \circ \sigma_2 & = & \begin{pmatrix} 1\,2\,3\,4\,5 \\ 3\,4\,1\,2\,5 \end{pmatrix}, \quad s(\sigma_1 \circ \sigma_2) = 4, \quad \operatorname{sgn}\,\sigma_1 \circ \sigma_2 = 1. \end{array}$$

(2) Eigenschaften der $\mathcal{S}_3$. Setze

$\alpha = \begin{pmatrix} 1\,2\,3 \\ 2\,1\,3 \end{pmatrix}$ und $\beta = \begin{pmatrix} 1\,2\,3 \\ 1\,3\,2 \end{pmatrix}$.

Dann gilt $\alpha^2 = \mathrm{id}$, $\beta^2 = \mathrm{id}$ und

$$\beta \circ \alpha = \begin{pmatrix} 1\,2\,3 \\ 3\,2\,1 \end{pmatrix} \neq \begin{pmatrix} 1\,2\,3 \\ 2\,3\,1 \end{pmatrix} = \alpha \circ \beta.$$

Dies bestätigt, daß $\mathcal{S}_3$ eine nicht-kommutative Gruppe ist.

$U := \{\mathrm{id}, \alpha\}$ ist Untergruppe, da $\alpha^{-1} = \alpha$, aber kein Normalteiler, denn

$$\beta^{-1} \circ \alpha \circ \beta = \begin{pmatrix} 1\,2\,3 \\ 3\,2\,1 \end{pmatrix} \notin U.$$

Die Komposition $\gamma := \alpha \circ \beta$ ist eine gerade Permutation. Die Menge $\{\mathrm{id}, \gamma, \gamma^2\}$ enthält alle geraden Permutationen und ist Normalteiler ($= \mathcal{A}_3$).

6.27 Aufgaben

(1) Auf $\mathbb{Z}$ wird folgende Verknüpfung definiert :

$$a \,\tau\, b := a + b + 1 \text{ für } a, b \in \mathbb{Z}.$$

Man zeige:

(i) $a \,\tau\, b = b \,\tau\, a$ *für alle* $a, b \in \mathbb{Z}$;

(ii) $(a \,\tau\, b) \,\tau\, c = a \,\tau\, (b \,\tau\, c)$ *für alle* $a, b, c \in \mathbb{Z}$;

(iii) *Es gibt ein* $e \in \mathbb{Z}$ *mit* $e \,\tau\, a = a \,\tau\, e = a$ *für alle* $a \in \mathbb{Z}$.

(2) Finden Sie alle Möglichkeiten, die folgende Verknüpfungstafel so zu vervollständigen, daß man eine Halbgruppe mit neutralem Element erhält:

τ	a	b	c
a	c	b	a
b			b
c	a		

(3) Sei (G, τ) eine Gruppe. Zeigen Sie:

(i) *Jedes linksinverse Element aus G ist auch rechtsinvers.*

(ii) *Jedes linksneutrale Element aus G ist auch rechtsneutral.*

(iii) *Gilt für alle $a, b, c \in G$ daß $a \tau c = b \tau c$, so ist $a = b$.*

(iv) *G ist genau dann kommutativ, wenn für alle $a, b \in G$ und für alle $n \in \mathbb{N}$ gilt: $(a \tau b)^n = a^n \tau b^n$.*

(v) *$Z(G) := \{a \in G \mid$ für alle $b \in G$ gilt $a \tau b = b \tau a\}$, das* Zentrum von G, *ist eine kommutative Untergruppe von G.*

(vi) *U und V seien Untergruppen von G.*
Dann ist auch $U \cap V$ eine Untergruppe von G.
$U \cup V$ ist genau dann Untergruppe von G, wenn $U \subset V$ oder $V \subset U$.

(4) Sei $(G_i, \tau_i)_I$ eine Familie von Halbgruppen. Man zeige:

(i) *Das Produkt $\prod_I G_i$ ist dann mit der komponentenweisen Verknüpfung ebenfalls eine Halbgruppe.*

(ii) *Ist G_i eine Gruppe für alle $i \in I$, so ist auch $\prod_I G_i$ eine Gruppe.*

(iii) *Ist G_i kommutativ für alle $i \in I$, so ist auch $\prod_I G_i$ kommutativ.*

(5) Sei (G, τ) eine Gruppe und U eine Untergruppe von G.

(i) *Zeigen Sie, daß für alle $a, b \in G$ gilt:*

$$b^{-1} \tau a \in U \Leftrightarrow a \tau U = b \tau U \Leftrightarrow (a \tau U) \cap (b \tau U) \neq \emptyset.$$

(ii) *Zeigen Sie, daß auf G eine Äquivalenzrelation gegeben ist durch*

$$R_U = \{(a, b) \in G \times G \mid b^{-1} \tau a \in U\}.$$

(iii) *G sei eine endliche Gruppe der Ordnung $n \in \mathbb{N}$.*
Bestimmen Sie die Anzahl der verschiedenen Äquivalenzklassen bezüglich R_U, und beweisen Sie so den Satz von Lagrange*:*
Ist U eine Untergruppe einer endlichen Gruppe G, so ist die Ordnung von U ein Teiler der Ordnung von G.

(6) Sei (G, τ) eine Gruppe und U eine Untergruppe von G. Zeigen Sie:

(i) *Folgende Aussagen sind äquivalent:*

(a) *$aUa^{-1} \subset U$ für alle $a \in G$ (U ist* Normalteiler*);*

(b) *$aUa^{-1} = U$ für alle $a \in G$;*

(c) *$aU = Ua$ für alle $a \in G$.*

(ii) *Ist U Normalteiler, dann läßt sich auf der Menge $G/U := G/R_U$ der Äquivalenzklassen bezüglich R_U (s.o.) durch $[a]\overline{\tau}[b] := [a \tau b]$ eine Verknüpfung definieren, die $(G/U, \overline{\tau})$ zu einer Gruppe macht.*

(iii) *Die Abbildung* $p_U : (G, \tau) \to (G/U, \overline{\tau})$ *mit* $P_{R_U}(a) = [a]$ *ist ein (Gruppen-)Epimorphismus.*

(iv) *Jeder Normalteiler ist Kern eines (Gruppen-)Homomorphismus.*

(7) Sei (G, τ) *eine Gruppe mit Untergruppe* U *und Normalteiler* N*. Zeigen Sie:*

(i) $U \tau N := \{u \tau n \mid u \in U \text{ und } n \in N\}$ *ist Untergruppe in* G*.*

(ii) $U \cap N$ *ist Normalteiler in* U*.*

(iii) $U/(U \cap N)$ *ist isomorph zu* $(U \tau N)/N$*.*

(8) Seien G_1*,* G_2*,* G_3 *Gruppen,* $f : G_1 \to G_2$ *ein Epimorphismus und* $g : G_2 \to G_3$ *eine Abbildung. Zeigen Sie:*
Ist $g \circ f : G_1 \to G_3$ *ein Homomorphismus, so ist auch* g *ein Homomorphismus.*

(9) Man bestimme alle Endomorphismen der Gruppe $(\mathbb{Q}, +)$*.*

(10) Zeigen Sie, daß jede Gruppe der Ordnung n *isomorph ist zu einer Untergruppe von* $\mathcal{S}_n$ *(vgl. 6.9,(4)).*

(11) Sei $\prod_I H_i$ *das Produkt einer Familie von Halbgruppen mit neutralen Elementen* $e_i \in H_i$ *(vgl. 6.20). Zeigen Sie, daß es zu jedem* $k \in I$ *einen Homomorphismus*

$$\varepsilon : H_k \to \prod\nolimits_I H_i$$

gibt mit $\pi \circ \varepsilon = \mathrm{id}_{H_k}$*.*

Man folgere, daß jedes H_k *isomorph ist zu einer Untergruppe von* $\prod_I H_i$*.*

7 Ringe und Körper

In diesem Abschnitt untersuchen wir algebraische Strukturen mit zwei Verknüpfungen. Da diese sehr weitgehend den Eigenschaften von $+$ und $\cdot$ in den ganzen Zahlen entsprechen sollen, belegt man sie meist mit diesen Symbolen:

7.1 Definition
Eine Menge R mit zwei Verknüpfungen $+$, $\cdot$, also ein Tripel $(R, +, \cdot)$, heißt ein *Ring*, wenn gilt:

(i) $(R, +)$ ist eine abelsche Gruppe,

(ii) $(R, \cdot)$ ist eine Halbgruppe mit neutralem Element,

(iii) für alle $a, b, c \in A$ gelten die *Distributivgesetze*
$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c, \\ (a + b) \cdot c &= a \cdot c + b \cdot c. \end{aligned}$$
Dabei folgen wir der Konvention, daß die *Multiplikation* $\cdot$ stärker bindet als die *Addition* $+$.

Man nennt den Ring $(R, +, \cdot)$ *kommutativ*, wenn zusätzlich gilt

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

wenn also $(R, \cdot)$ eine kommutative Halbgruppe ist.

Das neutrale Element von $(R, +)$ nennt man die *Null* in R und schreibt dafür 0. Das neutrale Element von $(R, \cdot)$ heißt die *Eins* von R, auch *Einselement*, und man bezeichnet es meist mit 1 (oder e). Statt $a \cdot b$ schreibt man häufig nur ab.

Bemerkung: Es macht auch Sinn, Ringe *ohne Einselemente* zu betrachten. Wir wollen hier aber in Ringen im allgemeinen die Existenz einer Eins voraussetzen.

Als elementare Folgerungen notieren wir:

In einem Ring $(R, +, \cdot)$ gelten für alle $a, b \in R$:

$$\begin{aligned} &0 \cdot a = a \cdot 0 = 0, \\ &(-a)b = a(-b) = -ab \text{ und} \\ &(-a)(-b) = ab. \end{aligned}$$

Beweis: Aus $a = (a + 0)$ folgt $aa = aa + a0$ und somit $0 = 0a$.
Analog sieht man $0 = a0$.
Aus $0 = (a + (-a))b = ab + (-a)b$ folgt $-(ab) = (-a)b$. □

7.2 Definition
Ein Ring $(R, +, \cdot)$ heißt *Divisionsring* (auch *Schiefkörper*), wenn $(R \setminus \{0\}, \cdot)$ eine Gruppe ist, wenn es also zu jedem $0 \neq r \in R$ ein Inverses bzgl. der Multiplikation gibt. Dies bezeichnet man mit r^{-1}.

Ein Divisionsring heißt *Körper*, wenn $(R, \cdot)$ kommutativ ist, also $a \cdot b = b \cdot a$ für alle $a, b \in R$.

In einem Divisionsring R ist das Produkt von zwei Elementen a, b nur dann 0, wenn eines der beiden Elemente schon 0 ist: Aus $ab = 0$ und $b \neq 0$ folgt nämlich $0 = (ab)b^{-1} = a$.

7.3 Definition
Sei R ein Ring. Ein Element $a \in R$ heißt *Nullteiler*, wenn es ein $0 \neq b \in R$ gibt mit $ab = 0$.
Ringe, in denen es keine Nullteiler $\neq 0$ gibt, heißen *nullteilerfrei*.
Kommutative Ringe, die nullteilerfrei sind, nennt man *Integritätsringe*.

Zum Beispiel ist jeder Divisionsring nullteilerfrei, und jeder Körper ist Integritätsring.

Wie bei den Gruppen, so interessieren uns auch hier die Abbildungen, welche die Ringstruktur berücksichtigen:

7.4 Definition
Sei $f : R \to S$ eine Abbildung zwischen zwei Ringen R und S. f heißt *(Ring-)Homomorphismus*, wenn für alle $a, b \in R$ gilt:

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

Man nennt f einen *Anti-Homomorphismus*, wenn an Stelle der zweiten Bedingung gilt

$$f(a \cdot b) = f(b) \cdot f(a).$$

Als Kern f bezeichnet man das Urbild der $0 \in S$:

$$\operatorname{Kern} f = \{a \in R \mid f(a) = 0\}$$

Ein Ringhomomorphismus ist also ein Halbgruppen-Homomorphismus zwischen $(R, +)$ und $(S, +)$ sowie zwischen $(R, \cdot)$ und $(S, \cdot)$.

7.5 Hilfssatz
Ist $f : R \to S$ ein Ringhomomorphismus, so gilt:

(1) $\operatorname{Kern} f$ *ist eine Untergruppe von* $(R,+)$*, und für alle* $b \in R$ *gilt*

$$b \operatorname{Kern} f \subset \operatorname{Kern} f \quad \textit{und} \quad (\operatorname{Kern} f)b \subset \operatorname{Kern} f.$$

(2) f *ist genau dann ein Monomorphismus, wenn* $\operatorname{Kern} f = \{0\}$ *ist.*

(3) *Ist* R *ein Divisionsring, so ist* f *ein Monomorphismus oder die Nullabbildung.*

Beweis: (1) $\operatorname{Kern} f$ ist Untergruppe nach 6.14.
Sei $k \in \operatorname{Kern} f$ und $b \in R$. Dann ist $f(kb) = f(k)f(b) = 0$, also $kb \in \operatorname{Kern} f$.

(2) wurde in Satz 6.15 gezeigt.

(3) Angenommen $0 \neq c \in \operatorname{Kern} f$. Nach (1) gilt dann $1 = c \cdot c^{-1} \in \operatorname{Kern} f$, und für jedes $a \in R$ ergibt sich $f(a) = f(1a) = f(1)f(a) = 0$. □

Wie in 6.17 ausgeführt, haben wir auf $R/\operatorname{Kern} f$ eine Addition. Wegen der in 7.5 angegebenen Eigenschaften der Kerne läßt sich darauf auch eine Multiplikation definieren durch

$$(a + \operatorname{Kern} f) \cdot (b + \operatorname{Kern} f) = ab + \operatorname{Kern} f.$$

Es ist nicht schwierig zu zeigen, daß diese Festlegung unabhängig ist von der Auswahl der Repräsentanten.

Dafür gilt nun wieder

7.6 Homomorphiesatz für Ringe
Sei $f : R \to S$ *ein Ringhomomorphismus. Dann gilt:*
Die kanonische Projektion $p_f : R \to R/\operatorname{Kern} f$ *ist ein (surjektiver) Ringhomomorphismus, und die Abbildung*

$$\overline{f} : R/\operatorname{Kern} f \to S, \quad [a] \mapsto f(a),$$

ist ein injektiver Ringhomomorphismus. Wir haben somit folgendes kommutative Diagramm von Ringhomomorphismen:

$$\begin{array}{ccc} R & \xrightarrow{f} & S \\ {\scriptstyle p_f}\searrow & & \nearrow{\scriptstyle \overline{f}} \\ & R/\operatorname{Kern} f. & \end{array}$$

7.7 Definition
Sei R ein Ring. Eine Untergruppe $U \subset (R,+)$ heißt

Linksideal in R, falls $aU \subset U$ für alle $a \in R$,

Rechtsideal in R, falls $Ua \subset U$ für alle $a \in R$,

Ideal in R, falls U Links- und Rechtsideal ist.

Diese Forderungen implizieren, daß jedes einseitige Ideal U selbst bzgl. $\cdot$ abgeschlossen und damit ein Unterring (ohne Eins) ist. Andererseits braucht jedoch ein Unterring kein Links- oder Rechtsideal in R zu sein.

In jedem Ring R sind $\{0\}$ und R Ideale. Man nennt sie die *trivialen Ideale*. Ideale, die ungleich R sind, nennt man *echte* Ideale.

Da R ein Einselement hat, gilt für ein Linksideal $U \subset R$ genau dann $U = R$, wenn $1 \in U$. Letzteres impliziert nämlich $a = a \cdot 1 \in U$ für alle $a \in R$.

Ist $U \subset R$ ein Ideal, so läßt sich auf der Faktorgruppe R/U durch

$$(a + U) \cdot (b + U) := ab + U$$

eine Multiplikation einführen, die $(R/U, +, \cdot)$ zu einem Ring macht.

Die kanonische Projektion $p : R \to R/U$ ist dann ein Ringhomomorphismus mit $\operatorname{Kern} p = U$.

Die Ideale sind bestimmend für die Struktur eines Ringes. Man sagt:

7.8 Definition
Ein Ring R heißt *einfach*, wenn er keine Ideale $\neq \{0\}, R$ enthält.

7.9 Satz. *Sei R ein Ring.*

(1) *R ist genau dann Divisionsring, wenn er keine nicht-trivialen Linksideale enthält.*

(2) *Sei R kommutativ. Dann ist R genau dann ein Körper, wenn er einfach ist.*

Beweis: (1) Sei R ein Ring ohne nicht-triviale Linksideale und $0 \neq a \in R$. Dann ist $0 \neq Ra$ ein Linksideal, also $Ra = R$. Somit gibt es $b \in R$ mit $ba = 1$, d.h. R ist Divisionsring.

Sei R Divisionsring und $U \subset R$ ein Linksideal $\neq 0$. Für ein $0 \neq u \in U$ gilt dann $1 = u^{-1}u \in U$ und damit $U = R$.

(2) Dies ist ein Spezialfall von (1). □

Sehen wir einige Beispiele zu den gegebenen Definitionen an.

7.10 Beispiele

(1) $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Integritätsring.

Für jedes $n \in \mathbb{N}$ bildet $n\mathbb{Z}$ ein Ideal in $\mathbb{Z}$, und wir können dazu den Faktorring $\mathbb{Z}/n\mathbb{Z}$ bilden. Dieser besteht aus n Äquivalenzklassen, die wir mit

$\{0, 1, \ldots, n-1\}$ markieren können. Addition und Multiplikation darin ergeben sich durch Rechnen *modulo n*.

(2) Über jedem Ring R kann man *Matrizenringe* definieren.

Wir werden später noch allgemeinere Bildungen kennenlernen. Hier wollen wir den einfachsten nicht-trivialen Fall angeben:

Matrizenring $R^{(2,2)}$: Auf den $(2,2)$-Matrizen mit Elementen in R können Addition und Multiplikation definiert werden durch:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix},$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} := \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}.$$

Dieser Ring enthält immer Nullteiler, denn

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Er ist nicht kommutativ, da z.B.

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

$\begin{pmatrix} R & 0 \\ R & 0 \end{pmatrix}$ und $\begin{pmatrix} 0 & R \\ 0 & R \end{pmatrix}$ sind Linksideale in $R^{(2,2)}$;

$\begin{pmatrix} R & R \\ 0 & 0 \end{pmatrix}$ und $\begin{pmatrix} 0 & 0 \\ R & R \end{pmatrix}$ sind Rechtsideale in $R^{(2,2)}$;

$\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}$ ist Unterring, aber kein Ideal.

7.11 Ringstruktur auf Mengen von Abbildungen

Sei $(R, +, \cdot)$ ein Ring und I eine Menge. Dann kann auf $\mathrm{Abb}(I, R) = R^I$, der Menge der Abbildungen von I in R (vgl. 3.12), eine Ringstruktur festgelegt werden durch

$$f \oplus g(x) := f(x) + g(x), \quad f \odot g(x) := f(x) \cdot g(x).$$

Die Bedingungen an die Verknüpfungen $\oplus$ und $\odot$ können durch Rückführung auf die entsprechenden Eigenschaften von $+$ und $\cdot$ nachgewiesen werden.

Auch dieser Ring enthält immer Nullteiler, falls I mindestens zwei Elemente hat: Sei R ein Ring und $I = \{i, j\}$ mit $i \neq j$. Definiere Abbildungen $f, g : I \to R$ durch

$$f(i) := 0,\ f(j) := 1; \quad g(i) := 1,\ g(j) := 0.$$

Diese Abbildungen sind nicht Null, aber $f \odot g$ ist Null.

Obige Konstruktion macht (für jeden Ring R) speziell $R \times R$ zu einem Ring (mit Nullteiler) durch komponentenweise Festlegung der Operationen.

Mit den bereitgestellten Mitteln können wir aus gegebenen Ringen weitere Ringe konstruieren. Mit einem ähnlichen Ansatz wie in Beispiel (3) erhalten wir auch den Ring der *Polynome*. Das Rechnen mit Polynomen ist sicher von der Schule her vertraut. Bei genauerem Hinsehen muß man aber doch nachfragen, was es mit der dabei auftretenden *Unbestimmten* auf sich hat. Die nachfolgende Beschreibung liefert den Nachweis, daß wir im Rahmen der von uns aufgebauten Grundlagen mit Polynomen so umgehen dürfen, wie wir es gewohnt sind.

7.12 Polynomring
Sei R ein kommutativer Ring. Auf $\mathrm{Abb}(\mathbb{N}, R) = R^{\mathbb{N}}$, der Menge der Abbildungen $\mathbb{N} \to R$, führen wir eine Addition wie im vorangehenden Beispiel ein, aber eine andere Multiplikation. Bemerkenswert daran ist, daß die Multiplikation auch die Eigenschaften von $\mathbb{N}$ heranzieht:

$$\begin{aligned}(a_i)_{i\in\mathbb{N}} \oplus (b_i)_{i\in\mathbb{N}} &= (a_i + b_i)_{i\in\mathbb{N}},\\ (a_i)_{i\in\mathbb{N}} \odot (b_i)_{i\in\mathbb{N}} &= (c_i)_{i\in\mathbb{N}},\ \text{mit } c_i = \sum_{k=0}^{i} a_k b_{i-k}.\end{aligned}$$

Damit wird $\mathrm{Abb}(\mathbb{N}, R)$ ein kommutativer Ring. Es ist leicht zu sehen, daß das Einselement dargestellt werden kann durch

$$e = (1, 0, 0, \ldots).$$

Sehen wir uns das folgende Element genauer an:

$$X := (0, 1, 0, \ldots), \quad \text{also } X(n) = \begin{cases} 1 & \text{für } n = 1 \\ 0 & \text{für } n \neq 1 \end{cases}.$$

Multiplizieren wir X mit sich, so erhalten wir

$$X^2 = (0, 0, 1, 0, \ldots); \quad X^2(n) = \begin{cases} 1 & \text{für } n = 2 \\ 0 & \text{für } n \neq 2 \end{cases}.$$

Allgemein ergibt sich für $k \in \mathbb{N}$,

$$X^k(n) = \begin{cases} 1 & \text{für } n = k \\ 0 & \text{für } n \neq k \end{cases}.$$

Bezeichne $\mathrm{Abb}_e(\mathbb{N}, R)$ die Abbildungen $\mathbb{N} \to R$, die nur auf endlich vielen $n \in \mathbb{N}$ ungleich Null sind, also darstellbar sind durch

$$(a_i)_{i\in\mathbb{N}} \ \text{ mit } a_i = 0 \text{ für fast alle } i \in \mathbb{N}.$$

Dies ist offensichtlich ein Unterring von $\mathrm{Abb}(\mathbb{N}, R)$. Jedes Element daraus hat die Form

$$(a_0, a_1, a_2, \ldots, a_n, 0, 0, \ldots) = a_0 + a_1 X + \ldots + a_n X^n.$$

Man nennt $R[X] := (\mathrm{Abb}_e(\mathbb{N}, R), \oplus, \odot)$ den *Polynomring über R in einer Unbestimmten X.*

Die Elemente daraus sind die uns vertrauten Polynome, und wir kehren wieder zur üblichen Notation zurück. Insbesodere haben wir für Addition und Multiplikation von zwei Polynomen (wie erwartet)

$$\begin{aligned} \sum_{i=0}^{n} a_i X^i + \sum_{j=0}^{m} b_j X^j &= \sum_{i=0}^{max(n,m)} (a_i + b_i) X^i, \\ \sum_{i=0}^{n} a_i X^i \cdot \sum_{j=0}^{m} b_j X^j &= \sum_{i=0}^{n+m} (\sum_{k=0}^{i} a_k b_{i-k}) X^i. \end{aligned}$$

Ohne Einschränkung können wir annehmen, daß in obiger Darstellung von Polynomen $a_n \neq 0$ und $b_m \neq 0$. Ist R Integritätsring, so ist auch $a_n b_m \neq 0$, und damit ist das rechts stehende Polynom nicht Null, das heißt:

Ist R ein Integritätsring, so ist auch $R[X]$ ein Integritätsring.

Für jeden kommutativen Ring R ist die Abbildung

$$\varepsilon : R \to Re = (R, 0, 0, \ldots)$$

ein Homomorphismus, der jedem Polynom sein konstantes Glied zuordnet.

Als nützliche Eigenschaft von $R[X]$, die sich leicht nachprüfen läßt, halten wir fest:

Universelle Abbildungseigenschaft von $R[X]$
Zu jedem Ring S, $s \in S$ und jedem Homomorphismus $\varphi : R \to S$ gibt es genau einen Homomorphismus $\Phi : R[X] \to S$ mit $\Phi(X) = s$ und dem kommutativen Diagramm

$$\begin{array}{ccc} R & \xrightarrow{\varphi} & S \\ {\scriptstyle\varepsilon}\searrow & & \nearrow{\scriptstyle\Phi} \\ & R[X]. & \end{array}$$

Die Wirkung von Φ auf ein Polynom ist beschrieben durch

$$\Phi : a_0 + a_1 X + \ldots + a_n X^n \mapsto \varphi(a_0) + \varphi(a_1) s + \ldots + \varphi(a_n) s^n.$$

Ein wichtiger Schritt beim Aufbau des Zahlensystems ist die Konstruktion der rationalen Zahlen aus den ganzen Zahlen. Auf gleiche Weise erhält man zu jedem Integritätsring den Quotientenkörper, und wir wollen dies skizzieren:

7.13 Quotientenkörper eines Integritätsrings

Sei R ein Integritätsring. Definiere eine Relation auf $R \times (R \setminus \{0\})$:

$$(a, b) \sim (a', b') \text{ genau dann, wenn } ab' = ba'.$$

Dies ist eine Äquivalenzrelation. Wir schreiben für die Äquivalenzklassen $[(a, b)] = [a, b]$ und bezeichnen ihre Gesamtheit mit

$$Q(R) := \{[a, b] \mid a \in R,\, b \in R \setminus \{0\}\}.$$

Darauf betrachten wir die Verknüpfungen

$$\begin{aligned} [a, b] + [a', b'] &:= [ab' + ba', bb'] \\ [a, b] \cdot [a', b'] &:= [aa', bb']. \end{aligned}$$

Es ist nachzuprüfen, daß diese Festlegungen unabhängig sind von der Auswahl der Repräsentanten, und daß damit $(Q(R), +, \cdot)$ zu einem kommutativen Ring wird. Dabei wird $[1, 1]$ das Einselement und $[0, 1]$ das Nullelement.
Das Inverse zu $[a, b]$ mit $a \neq 0$ ist $[b, a]$.

$Q(R)$ ist also ein Körper, der *Quotientenkörper* von R. Die Abbildung

$$R \to Q(R), \quad a \mapsto [a, 1],$$

ist eine *Einbettung* (injektiver Homomorphismus) von R in $Q(R)$.

Setzen wir $\frac{a}{b} := [a, b]$, so erkennen wir, daß die obigen Definitionen gerade die üblichen Rechenregeln für Brüche ergeben.

Es sei angemerkt, daß bei dieser Konstruktion die Kommutativität von R wesentlich mitbenutzt wird. Für nicht-kommutative Ringe sind vergleichbare Bildungen wesentlich aufwendiger.

Als Spezialfall erhält man für $R = \mathbb{Z}$ wie erwartet $Q(\mathbb{Z}) = \mathbb{Q}$.

Für einen Integritätsring R ist auch der Polynomring $R[X]$ Integritätsring (vgl. 7.12), und wir erhalten mit der angegebenen Konstruktion den Quotientenkörper $Q(R[X])$ dazu. Man nennt diesen den Körper der *rationalen Funktionen*. Seine Elemente lassen sich als Brüche von zwei Polynomen beschreiben.

Mit den vorangegangenen Betrachtungen haben wir gesehen, daß wir mit den bereitgestellten (mengentheoretischen) Grundlagen auch die rationalen Zahlen in unsere Theorie einfügen können. Prinzipiell gilt dies auch für die reellen Zahlen, doch sind dazu auch nicht-algebraische Überlegungen notwendig, mit denen wir uns hier nicht befassen: Man muß $\mathbb{Q}$ zu einem Körper erweitern, in dem alle *Cauchy-Folgen* von Elementen aus $\mathbb{Q}$ einen Grenzwert haben.

Der Schritt von den reellen zu den komplexen Zahlen ist allerdings wieder rein algebraischer Natur. Daher wollen wir ihn hier angeben.

7.14 Die komplexen Zahlen

Auf dem kartesischen Produkt $\mathbb{R} \times \mathbb{R}$ führen wir Addition und Multiplikation ein durch

$$\begin{aligned}(a,b) + (a',b') &:= (a+a', b+b'),\\ (a,b) \cdot (a',b') &:= (aa' - bb', ab' + ba').\end{aligned}$$

Mit diesen Verknüpfungen wird $\mathbb{R} \times \mathbb{R}$ ein Körper, den man den Körper der *komplexen Zahlen* $\mathbb{C}$ nennt.

$(1,0)$ ist das Einselement, $(0,0)$ das Nullelement darin. Das Inverse zu $(0,0) \neq (a,b) \in \mathbb{C}$ bekommt man durch

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

Wir haben eine Einbettung

$$\mathbb{R} \to \mathbb{C}, \quad a \mapsto a(1,0) = (a,0),$$

und dürfen daher die Elemente $a \in \mathbb{R}$ mit $(a,0) \in \mathbb{C}$ identifizieren. Das Element $i := (0,1)$ hat die Eigenschaft $i^2 = -(1,0)$, und mit obiger Festlegung gilt dann für die Elemente in $\mathbb{C}$

$$(a,b) = a + bi.$$

Dies gibt uns die vertraute und übliche Schreibweise der komplexen Zahlen.

Auf eine weitere algebraische Darstellung von $\mathbb{C}$ wird bei den Aufgaben zu diesem Abschnitt hingewiesen.

Die wohl wichtigste Eigenschaft von $\mathbb{C}$ ist für uns, daß jedes nicht konstante Polynom aus $\mathbb{C}[X]$ eine Nullstelle in $\mathbb{C}$ hat. Dies wird manchmal der *Fundamentalsatz der Algebra* genannt, obwohl er nicht mit algebraischen Methoden alleine bewiesen werden kann. Der erste Beweis dazu stammt von C.F. Gauß (1799).

Eine weitere Besonderheit von $\mathbb{C}$ sei noch erwähnt: Die Abbildung

$$\gamma : \mathbb{C} \to \mathbb{C}, \quad a + ib \mapsto a - ib,$$

ist ein Automorphismus von $\mathbb{C}$ mit $\gamma^2 = \mathrm{id}_{\mathbb{C}}$. Dies kann man einfach nachrechnen.

Einen solchen Automorphismus gibt es für die reellen Zahlen nicht, und natürlich hat auch nicht jedes nicht-konstante Polynom aus $\mathbb{R}[X]$ eine Nullstelle in $\mathbb{R}$.

7.15 Die Quaternionen

Auf dem kartesischen Produkt $\mathbb{R}^4$ führen wir Addition und Multiplikation ein durch

$$\begin{aligned}(a,b,c,d)+(a',b',c',d') &:= (a+a',b+b',c+c',d+d'),\\ (a,b,c,d)\cdot(a',b',c',d') &:= (aa'-bb'-cc'-dd', ab'+ba'+cd'-dc',\\ &\quad ac'-bd'+ca'+db', ad'+bc'-cb'+da').\end{aligned}$$

Mit diesen Verknüpfungen wird $\mathbb{R}^4$ ein Divisionsring, den man den Ring der *Quaternionen* $\mathbb{H}$ nennt. Das Symbol $\mathbb{H}$ soll an *W.R. Hamilton* erinnern, der als einer der ersten auf die Bedeutung dieses Divisionsrings (in Geometrie und Zahlentheorie) aufmerksam machte.

$1_{\mathbb{H}} := (1,0,0,0)$ ist das Einselement, $(0,0,0,0)$ das Nullelement darin. Das Inverse bekommt man durch

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

Wir haben eine Einbettung

$$\mathbb{R} \to \mathbb{H}, \quad a \mapsto a(1,0,0,0) = (a,0,0,0),$$

und dürfen daher die Elemente $a \in \mathbb{R}$ mit $(a,0,0,0) \in \mathbb{H}$ identifizieren.

Setzt man die letzten beiden Komponenten von Elementen aus $\mathbb{H}$ gleich Null, so erhält man gerade die komplexen Zahlen $\mathbb{C}$. Dies ergibt auch eine Einbettung

$$\mathbb{C} \to \mathbb{H}, \quad (a,b) \mapsto (a,b,0,0).$$

Die angegebene Multiplikation der Quaternionen läßt sich handlicher angeben, wenn man bedenkt, daß sie schon durch die Produkte von gewissen *Basiselementen* bestimmt ist. Folgende Notation hat sich dabei eingebürgert:

$$i := (0,1,0,0), \quad j := (0,0,1,0), \quad k := (0,0,0,1).$$

Damit können wir die Elemente von $\mathbb{H}$ schreiben als

$$(a,b,c,d) = a + bi + cj + dk.$$

Die Multiplikation der Basiselemente ergibt

$$\begin{gathered}i^2 = j^2 = k^2 = -1_{\mathbb{H}}\\ ij = -ji = k, \quad jk = -jk = i, \quad ki = -ik = j.\end{gathered}$$

Diese Beziehungen legen die Multiplikation in $\mathbb{H}$ fest. Sie zeigen, daß $\mathbb{H}$ nicht kommutativ ist.

Man sieht daraus auch, daß über $\mathbb{H}$ das Polynom X^2+1 drei verschiedene Nullstellen hat. Über einem kommutativen Körper kann ein Polynom zweiten Grades dagegen höchstens zwei Nullstellen haben.

Ähnlich wie in $\mathbb{C}$ haben wir auch in $\mathbb{H}$ einen Antiautomorphismus

$$\gamma : \mathbb{H} \to \mathbb{H}, \quad a + bi + cj + dk \mapsto a - bi - cj - dk,$$

mit $\gamma^2 = \mathrm{id}_{\mathbb{H}}$. Dies kann man einfach nachrechnen. Damit läßt sich die *Norm* eines Elements definieren durch

$$N(a+bi+cj+dk) := (a+bi+cj+dk)(a-bi-cj-dk) = (a^2+b^2+c^2+d^2)1_{\mathbb{H}},$$

mit der man das Inverse eines Elements $z \in \mathbb{H}$ ausdrücken kann als

$$z^{-1} = \frac{1}{N(z)}\gamma(z).$$

Eine Darstellung von $\mathbb{H}$ als Matrizenring über $\mathbb{C}$ wird in den Aufgaben angegeben.

Wir haben in 7.7 angegeben, daß zu jedem Ideal U in einem Ring R auf R/U eine Ringstruktur definiert werden kann. Dabei stehen Eigenschaften des Ideals U in Wechselbeziehung zu Eigenschaften des Ringes R/U. Folgende Eigenschaft von Idealen ist in diesem Zusammenhang von Bedeutung:

7.16 Definition
Ein echtes Linksideal $U \subset R$ heißt *maximal*, wenn es maximales Element in der Menge der echten Linksideale bezüglich der Inklusion $\subset$ ist, d.h.
für jedes Linksideal $V \subset R$ mit $U \subset V$ ist $U = V$ oder $V = R$.

Die Bedeutung dieser Ideale für die Struktur der zugehörigen Faktorringe liegt in folgender Beobachtung:

7.17 Satz
Sei R ein kommutativer Ring. Ein Ideal $U \subset R$ ist genau dann maximal, wenn der Faktorring R/U ein Körper ist.

Beweis: $\Rightarrow$ Sei U ein maximales Ideal, $p : R \to R/U$ die kanonische Projektion. Ist I ein Ideal in R/U, dann ist $p^{-1}(I)$ ein Ideal in R, das U enthält. Also gilt entweder $p^{-1}(I) = U$ und damit

$$I = p(p^{-1}(I)) = p(U) = [0],$$

oder es ist $p^{-1}(I) = R$ und damit $I = p(R) = R/U$.

Somit gibt es in R/U nur die trivialen Ideale. Nach Satz 7.9 ist dann R/U ein Körper.

$\Leftarrow$ Sei R/U ein Körper, $V \subset R$ ein Ideal mit $U \subset V$. Dann ist $p(V)$ Ideal in R/U, also $p(V) = [0]$ und damit $U = V$, oder $p(V) = R/U$, woraus $V = V + U = R$ folgt. Damit ist U maximales Ideal. $\square$

Die Frage nach der Existenz von maximalen Idealen läßt sich mit Hilfe des Zornschen Lemmas beantworten. Dazu zeigen wir zunächst:

7.18 Hilfssatz
Sei R ein Ring und $K \subset R$ ein Linksideal. Dann ist die Menge $\mathcal{U}$ aller von R verschiedenen Linksideale, die K enthalten, also

$$\mathcal{U} = \{I \subset R \mid I \text{ Linksideal}, K \subset I \text{ und } 1 \notin I\},$$

bezüglich der Inklusion $\subset$ induktiv geordnet.

Beweis: $\mathcal{U}$ ist nicht leer, da $K \in \mathcal{U}$. Sei $\mathcal{V}$ eine linear geordnete Teilmenge von $\mathcal{U}$. Wir zeigen, daß $\overline{U} = \bigcup_{U \in \mathcal{V}} U$ eine kleinste obere Schranke von $\mathcal{V}$ in $\mathcal{U}$ ist. Offensichtlich gilt $1 \notin \overline{U} \supset K$. Auch die gewünschte Minimaleigenschaft von $\overline{U}$ ist klar.

Es bleibt also nur nachzuweisen, daß $\overline{U}$ ein Linksideal ist. Betrachte dazu $a, b \in \overline{U}$. Dann gibt es ein $U \in \mathcal{U}$ mit $a, b \in U$, und damit gilt auch

$$a + b \in U \subset \overline{U} \text{ und } Ra \subset U \subset \overline{U}.$$

Also ist $\overline{U}$ ein Linksideal. □

Als Folgerung daraus erhalten wir den wichtigen

7.19 Satz von Krull
Jedes echte Linksideal in einem Ring R ist in einem maximalen Linksideal enthalten.

Beweis: Sei $K \neq R$ Linksideal in R. Nach 7.18 ist die Menge der echten Linksideale von R, die K enthalten, induktiv geordnet. Das Zornsche Lemma 5.5 garantiert die Existenz von maximalen Elementen in solchen Mengen.

Es ist leicht zu sehen, daß dies maximale Linksideale sind. □

Da in jedem Ring $\{0\}$ ein Ideal ist, ergibt sich aus 7.19 die Existenz von maximalen Linksidealen in allen Ringen (mit Eins!).

Mit einem analogen Beweis kann man auch zeigen, daß in einem Ring jedes Linksideal (Rechtsideal) in einem *maximalen Linksideal (Rechtsideal)* enthalten ist.

Für kommutative Ringe haben wir das

Korollar
Zu jedem kommutativen Ring gibt es einen Epimorphismus $R \to K$ auf einen Körper K.

Beweis: Sei U ein maximales Ideal in R. Dann ist R/U nach 7.17 ein Körper, und die kanonische Projektion $R \to R/U$ ist der gewünschte Epimorphismus. □

7.20 Aufgaben

(1) Sei $(G,+)$ eine abelsche Gruppe, $\mathrm{End}(G)$ die Menge der Gruppenhomomorphismen. Für $f,g \in \mathrm{End}(G)$ definieren wir

$$f+g : G \to G, \quad a \mapsto f(a)+g(a).$$

Zeigen Sie, daß mit dieser Addition und der Komposition von Abbildungen als Multiplikation $(\mathrm{End}(G),+,\circ)$ ein Ring mit Eins ist.

(2) Sei $(R,+,\cdot)$ ein Ring mit Eins. Das Zentrum *von R ist definiert als*

$$Z(R) := \{c \in R \mid a \cdot c = c \cdot a \text{ für alle } a \in R\}.$$

Zeigen Sie, daß $(Z(R),+,\cdot)$ ein Unterring von R ist.

(3) Sei $(R,+,\cdot)$ ein Ring mit $a \cdot a = a$ für alle $a \in R$. Man nennt dann R einen Booleschen Ring. Zeigen Sie:

(i) *R ist kommutativ.*

(ii) *$a + a = 0$ für alle $a \in R$.*

(iii) *Alle Unterringe und Faktorringe von R sind Boolesche Ringe.*

(iv) *Für jede Indexmenge I ist R^I ein Boolescher Ring.*

(v) *R ist genau dann ein einfacher Ring, wenn R höchstens zwei Elemente hat.*

(4) Seien A eine nicht-leere Menge und $\mathbb{Z}_2$ der Ring (Körper), der nur aus zwei Elementen $\{0,1\}$ besteht. Zeigen Sie, daß mit den in 7.10(3) definierten Verknüpfungen $(\mathrm{Abb}(A,\mathbb{Z}_2),\oplus,\odot)$ ein Boolescher Ring ist.

Die Bedeutung dieses Ringes liegt in folgendem Sachverhalt: Zu jeder Teilmenge $B \subset A$ gibt es genau ein $g \in \mathrm{Abb}(A,\mathbb{Z}_2)$ mit $g^{-1}(1) = B$ (charakteristische Funktion von B). Damit hat man eine Bijektion zwischen $\mathrm{Abb}(A,\mathbb{Z}_2)$ und der Potenzmenge $\mathcal{P}(A)$.

(5) Sei $K^{(2,2)}$ der Ring der $(2,2)$-Matrizen über einem Körper K. Zeigen Sie:

(i) *$K^{(2,2)}$ ist ein einfacher Ring.*
Bestimmen Sie das Zentrum (vgl. Aufgabe (2)) von $K^{(2,2)}$.

(ii) *Folgende Menge ist ein Unterring von $K^{(2,2)}$:*

$$C(K) := \{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a,b \in K\}.$$

(iii) *Für $K = \mathbb{R}$ ist $C(\mathbb{R})$ ein Körper.*

(iv) *$f : \mathbb{C} \to C(\mathbb{R})$ mit $a + ib \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ ist ein Isomorphismus.*

(6) Betrachten Sie die Normabbildung *auf den komplexen Zahlen $\mathbb{C}$,*

$$| \quad | : \mathbb{C} \to \mathbb{R}_{\geq 0}, \quad a + ib \mapsto \sqrt{a^2 + b^2}.$$

Zeigen Sie, daß dies ein Halbgruppen-Homomorphismus $(\mathbb{C}, \cdot) \to (\mathbb{R}_{\geq 0}, \cdot)$ ist, und daß die Dreiecksungleichung *gilt:*

$$|z_1 + z_2| \leq |z_1| + |z_2| \text{ für alle } z_1, z_2 \in \mathbb{C}.$$

(7) Ähnlich wie $\mathbb{C}$ aus $\mathbb{R}$ lassen sich die Quaternionen $\mathbb{H}$ aus $\mathbb{C}$ gewinnen. $\mathbb{H}$ wurde in 7.15 durch eine Multiplikation auf $\mathbb{R}^4$ eingeführt.
Sei $\mathbb{C}^{(2,2)}$ der Ring der $(2,2)$-Matrizen über $\mathbb{C}$. Für $z = a + ib \in \mathbb{C}$ bezeichne $\bar{z} = a - ib$ (konjugiert komplexe Zahl). Zeigen Sie:

(i) *Folgende Menge ist ein Unterring von $\mathbb{C}^{(2,2)}$:*

$$H(\mathbb{C}) := \{\begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix} \mid u, v \in \mathbb{C}\}.$$

(ii) *$f : \mathbb{H} \to H(\mathbb{C})$ mit $(a, b, c, d) \mapsto \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix}$*
ist ein (Ring-) Isomorphismus.

(iii) *$f(\mathbb{R}, \mathbb{R}, 0, 0) \simeq \mathbb{C}$ liegt nicht im Zentrum von $H(\mathbb{C})$.*

(8) Zeigen Sie, daß für einen Ring $R \neq \{0\}$ folgende Aussagen äquivalent sind:

(a) *R ist ein Divisionsring;*

(b) *$\{0\}$ ist maximales Linksideal in R;*

(c) *$\{0\}$ ist maximales Rechtsideal in R;*

(d) *in R gibt es keine nicht-trivialen Linksideale (Rechtsideale).*

Kapitel 3

Moduln und Vektorräume

8 Definitionen und Beispiele

Wir haben bei Gruppen und Ringen Verknüpfungen untersucht, die auf einer einzelnen Menge definiert waren. Nun werden wir uns mit Verbindungen von zwei Mengen befassen, die jeweils eine algebraische Struktur tragen. Dabei wird eine gewisse Verträglichkeit dieser Strukturen gefordert:

8.1 Definition
Sei R ein Ring mit 1 und $(M, +)$ eine abelsche Gruppe. Dann heißt M ein *R-Linksmodul*, wenn eine Abbildung

$$\cdot : R \times M \to M,\ (a, m) \mapsto a \cdot m,$$

mit folgenden Eigenschaften gegeben ist (für alle $a, b \in R$; $m, n \in M$):

M1) $a \cdot (m + n) = a \cdot m + a \cdot n$

M2) $(a + b) \cdot m = a \cdot m + b \cdot m$

M3) $a \cdot (b \cdot m) = (ab) \cdot m$

M4) $1 \cdot m = m$.

Man nennt diese Abbildung auch *Skalarmultiplikation*.
Statt $a \cdot m$ schreibt man meist am.

Analog dazu sind *R-Rechtsmoduln* definiert, d.h. die Skalare aus R werden von rechts an die Elemente von M multipliziert. Die Notation ${}_RM$ bzw. M_R soll klarstellen, ob wir einen Links- oder Rechtsmodul meinen.

Ist R ein kommutativer Ring und M ein Linksmodul darüber, so kann man durch die Festlegung $m \cdot a := a \cdot m$ die Gruppe M auch als R-Rechtsmodul auffassen. Man sieht leicht, daß dafür analog M1)–M4) gelten.

Für nicht-kommutative R gilt dies nicht mehr, da dann die Forderung M3) von rechts nicht mehr erfüllt zu sein braucht.

Unter einem *R-Modul* wollen wir im allgemeinen einen *R-Linksmodul* verstehen. Über kommutativen Ringen R braucht man zwischen *Links-* und *Rechts-* ohnehin nicht zu unterscheiden.

Ist R ein Divisionsring, so nennt man einen R-Linksmodul auch *Linksvektorraum* und einen R-Rechtsmodul entsprechend *Rechtsvektorraum*.

Ist klar, welche Seite gemeint ist, so sprechen wir einfach von einem *Vektorraum*. Über einem Körper braucht man da sowieso keinen Unterschied zu machen.

Elemente eines Moduls oder Vektorraums bezeichnet man als *Vektoren*.

Aus den Bedingungen in der Definition ergeben sich sofort elementare

Eigenschaften
Ist M ein R-Modul, dann gilt für alle $a \in R$ und $m \in M$:

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

(ii) $a \cdot (-m) = -(a \cdot m) = (-a) \cdot m$, speziell $(-1) \cdot m = -m$

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

(iv) Ist R Körper und $a \cdot m = 0$, dann gilt $m = 0$ oder $a = 0$.

Beweis: Man beachte, daß wir hier zwischen dem neutralen Element in M und dem neutralen Element von $(R, +)$ unterscheiden müssen. Es führt aber nicht zu Unklarheiten, wenn wir beide mit 0 bezeichnen.

(i) $m = (1 + 0) \cdot m = 1 \cdot m + 0 \cdot m = m + 0 \cdot m$, also $0 \cdot m = 0$.
$a \cdot m = a \cdot (m + 0) = a \cdot m + a \cdot 0$ und damit $a \cdot 0 = 0$.

(ii) $0 = a \cdot (m + (-m)) = a \cdot m + a \cdot (-m)$, d.h. $a(-m) = -(a \cdot m)$.

(iii) $(-a)(-m) = (-1)(-a) \cdot m = a \cdot m$.

(iv) Ist $a \cdot m = 0$ und $a \neq 0$, dann ist $a^{-1}(a \cdot m) = a^{-1}0 = 0$.

□

In vielen der bislang betrachteten algebraischen Strukturen sind auch Moduln enthalten. Sehen wir uns einige davon an.

8.2 Beispiele

(1) **Der Ring selbst.** Jeder Ring R ist Links- und Rechtsmodul über sich selbst, denn aus den Rechengesetzen in R folgt, daß die Abbildung

$$\cdot : R \times R \to R,\ (r, r') \mapsto r \cdot r',$$

alle für Moduln geforderten Bedingungen erfüllt.

So ist etwa ${}_{\mathbb{Z}}\mathbb{Z}$ ein Modul, und ${}_{\mathbb{Q}}\mathbb{Q}$, ${}_{\mathbb{R}}\mathbb{R}$, ${}_{\mathbb{C}}\mathbb{C}$ sind Vektorräume.

(2) **Modul über Unterring.** Ist R Unterring eines Ringes S, so ist ${}_RS$ ein R-Modul mit

$$R \times S \to S,\ (r,s) \mapsto r \cdot s.$$

Analog kann man S als R-Rechtsmodul auffassen.

Insbesondere sind also $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ sowohl $\mathbb{Z}$-Moduln als auch $\mathbb{Q}$-Vektorräume. $\mathbb{C}$ ist $\mathbb{R}$-Vektorraum und $\mathbb{Q}$-Vektorraum.

(3) **Ringhomomorphismen.** Ist $\varphi : R \to S$ ein Ringhomomorphismus, so wird S ein R-Modul durch

$$R \times S \to S,\ (r,s) \mapsto \varphi(r) \cdot s.$$

Auch hier kann man analog S als R-Rechtsmodul auffassen.

So macht der kanonische Homomorphismus

$$\varphi : R \to R^{(2,2)}, \quad r \mapsto \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} = r \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

den Matrizenring $R^{(2,2)}$ zu einem Links- und Rechtsmodul über R.

Für kommutative Ringe R haben wir die Einbettung in den Polynomring,

$$\varepsilon : R \to R[X], \quad r \mapsto (r, 0, \ldots).$$

Damit wird nach obigem Rezept $R[X]$ zu einem R-Modul:

$$R \times R[X] \to R[X], \quad (r, \sum a_i X^i) \mapsto \sum (ra_i) X^i.$$

Ist R ein Körper, so ist $R[X]$ ein Vektorraum über R.

(4) **Produkt und direkte Summe.** Für jede Menge I und jeden Ring R bilden die Abbildungen $\mathrm{Abb}(I,R) = R^I$ eine abelsche Gruppe mit (vgl. 6.3)

$$(a_i)_{i\in I} + (b_i)_{i\in I} = (a_i + b_i)_{i\in I}.$$

Diese wird zu einem R-Linksmodul (*Produkt*) vermöge

$$R \times R^I \to R^I, \quad (r, (a_i)_{i\in I}) \mapsto (ra_i)_{i\in I}.$$

Analog kann das Produkt auch als R-Rechtsmodul aufgefaßt werden.

Die Menge der Abbildungen $I \to R$, die nur auf endlich vielen Elementen von I ungleich Null sind, bilden offensichtlich eine Untergruppe von $\mathrm{Abb}(I,R) = R^I$,

die wir mit $\mathrm{Abb}_e(I,R) = R^{(I)}$ bezeichnen. Sie erlaubt die gleiche Skalarmultiplikation wie oben und bildet daher einen R-Modul.

Man nennt $R^{(I)}$ die *(äußere) direkte Summe* von (Kopien von) R.

Für endliche Mengen I gilt natürlich $R^I = R^{(I)}$.

Ist R ein Körper, so bilden R^I und $R^{(I)}$ beide R-Vektorräume.

(5) **Ebene und Raum.** Als Spezialfälle erhalten wir für $I = \{1,2\}$ und $I = \{1,2,3\}$ die R-Moduln R^2 und R^3,

$$\begin{aligned} R \times R^2 \to R^2, &\quad (r,(a_1,a_2)) \mapsto (ra_1, ra_2), \\ R \times R^3 \to R^3, &\quad (r,(a_1,a_2,a_3)) \mapsto (ra_1, ra_2, ra_3). \end{aligned}$$

Insbesondere für $R = \mathbb{R}$ sind diese Vektorräume als *reelle Ebene* und *reeller (dreidimensionaler) Raum* unserer Anschauung vertraut. In der sie beschreibenden *euklidischen Geometrie* stecken aber wesentlich mehr Strukturen, als wir mit der Eigenschaft *Vektorraum* zunächst erfassen. Man beachte, daß wir bislang weder Längen noch Winkel in unseren Bildungen berücksichtigen.

Anschaulich klar ist auch die Interpretation der $\mathbb{Z}$-Moduln $\mathbb{Z}^2$ und $\mathbb{Z}^3$. Es sind die Punkte mit ganzzahligen Koordinaten in der Ebene bzw. im Raum, sogenannte *Gitter*.

Auch die Vektorräume $\mathbb{Q}^2$ und $\mathbb{Q}^3$ finden sich im $\mathbb{R}^2$ bzw. $\mathbb{R}^3$. Zeichnerisch sind sie davon aber nicht zu unterscheiden, da jedes Element im $\mathbb{R}^3$ mit Elementen aus $\mathbb{Q}^3$ beliebig genau dargestellt werden kann.

(6) **Abelsche Gruppen.** Jede abelsche Gruppe G ist ein $\mathbb{Z}$-Modul mit

$$\mathbb{Z} \times G \to G, \quad n \cdot g = \begin{cases} g + \ldots + g \ (n - \text{mal}) & \text{für } n > 0, \\ -(|n|g) & \text{für } n < 0, \\ 0 & \text{für } n = 0. \end{cases}$$

Da andererseits auch jeder $\mathbb{Z}$-Modul per definitionem eine abelsche Gruppe ist, fällt die Theorie der abelschen Gruppen mit der Theorie der $\mathbb{Z}$-Moduln zusammen. Die allgemeine Modultheorie ist somit eine Verallgemeinerung der Theorie abelscher Gruppen.

(7) **Moduln über Endomorphismenring.** Abelsche Gruppen G kann man noch auf andere Weise als Moduln auffassen. Die Endomorphismen von G bilden einen Ring, bezeichnet mit $\mathrm{End}(G)$, und G ist ein Modul darüber durch

$$\mathrm{End}(G) \times G \to G, \quad (h,g) \mapsto h(g).$$

Wie bei anderen algebraischen Strukturen sind auch bei Moduln die Unterstrukturen von Interesse:

8.3 Definition
Sei M ein R-Modul. Eine additive Untergruppe $U \subset M$ heißt *R-Untermodul von M*, wenn

$$aU \subset U \quad \text{für alle } a \in R.$$

Diese Eigenschaft kann man suggestiv als $RU = U$ schreiben.

Ist R ein Körper, so nennt man einen Untermodul des Vektorraums ${}_RM$ einen *Untervektorraum* oder kurz *Unterraum*.

Wir haben folgende Möglichkeiten, Untermoduln zu kennzeichnen:

8.4 Hilfssatz
Für eine nicht-leere Teilmenge U eines R-Moduls M sind folgende Eigenschaften äquivalent:

(a) *U ist R-Untermodul von M;*

(b) *für alle $a, b \in R$ und $u, v \in U$ gilt $au + bv \in U$;*

(c) *für alle $a \in R$ und $u, v \in U$ gilt $u + v \in U$ und $au \in U$.*

Beweis: (a)⇒(b) Seien $a, b \in R$, $u, v \in U$. Dann sind nach Definition 8.3 auch $au, bv \in U$ und $au + bv \in U$.

(b)⇒(c) Gilt (b), und setzt man $a = 1$ und $b = 1$, so erhält man $u + v \in U$. Für $b = 0$ und $a \in R$ ergibt sich $au \in U$.

(c)⇒(a) Nach (c) ist U zunächst Unterhalbgruppe von M. Für $a = -1$ folgt aus (c): $(-1)u = -u \in U$, d.h. U ist eine Untergruppe. □

8.5 Beispiele von Untermoduln
(1) Zu jedem Modul M sind $\{0\}$ und M Untermoduln.

(2) In einem Ring R sind die Untermoduln von ${}_RR$ gerade die Linksideale und die Untermoduln von R_R die Rechtsideale.

In einem kommutativen Ring sind die Untermoduln von R die Ideale.

So sind etwa für $n \in \mathbb{N}$ die Teilmengen $n\mathbb{Z}$ Ideale in $\mathbb{Z}$. Übrigens sind alle Ideale in $\mathbb{Z}$ von dieser Form.

(3) In jedem R-Modul M und für jedes $m \in M$ ist die Teilmenge

$$Rm = \{rm \mid r \in R\}$$

ein Untermodul von M. Es ist der kleinste Untermodul von M, der m enthält. Man nennt ihn den *von m erzeugten Untermodul*.

(4) Für jeden Ring R und jede Menge I ist die direkte Summe $R^{(I)}$ Untermodul vom Produkt R^I (vgl. 8.2).

(5) Anschaulich vorstellen kann man sich wieder die Untervektorräume in den reellen Räumen $\mathbb{R}^2$ und $\mathbb{R}^3$: Es sind die Geraden und Ebenen, die durch den Punkt $(0,0)$ bzw. $(0,0,0)$ gehen.

Man kann den $\mathbb{R}^2$ auf verschiedene Weise als Unterraum von $\mathbb{R}^3$ auffassen, etwa $(\mathbb{R},\mathbb{R},0) \subset \mathbb{R}^3$ oder $(0,\mathbb{R},\mathbb{R}) \subset \mathbb{R}^3$.

Aus gegebenen Untermoduln lassen sich auf vielfältige Weise neue bilden:

8.6 Summe und Durchschnitt von Untermoduln
Sind U_1 und U_2 Untermoduln des R-Moduls M, so ist auch

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1,\, u_2 \in U_2\}$$

ein Untermodul von M, denn

$$a(u_1+u_2)+b(v_1+v_2) = (au_1+bv_1)+(au_2+bv_2) \in U_1+U_2.$$

Man nennt $U_1 + U_2$ *die Summe von* U_1 *und* U_2.

Es ist leicht zu sehen, daß auch $U_1 \cap U_2$ ein Untermodul ist. Dagegen ist $U_1 \cup U_2$ im allgemeinen kein Untermodul.

Diese Bildungen lassen sich in offensichtlicher Weise auf endlich viele Untermoduln $U_1, \ldots, U_n$ von M erweitern, und man erhält dabei die Untermoduln $U_1 \cap \ldots \cap U_n$ und

$$U_1 + \cdots + U_n = \{u_1 + \cdots + u_n \mid u_i \in U_i \text{ für } i = 1, \ldots, n\}.$$

Auch für unendliche Familien $(U_i)_{i\in I}$ von Untermoduln von M ist $\bigcap\limits_{i\in I} U_i$ ein Untermodul von M, der größte Untermodul, der in allen U_i enthalten ist.

Die *Summe* aller U_i kann man als die Menge aller endlichen Linearkombinationen der Elemente aus den U_i definieren, also

$$\sum_{i\in I} U_i = \{\sum_{j\in J} u_j \,|\, u_j \in U_j,\ J \subset I,\ J \text{ endlich }\}.$$

Dies ist der kleinste Untermodul, der alle U_i enthält.

Die oben gemachten Beobachtungen führen zu der

8.7 Definition
Sei N eine nicht-leere Teilmenge des R-Moduls M. Dann heißt

$$\langle N \rangle := \sum_{n\in N} Rn$$

die *lineare Hülle* von N, oder der *von* N *erzeugte Untermodul* von M.

Daß $\langle N\rangle$ ein Untermodul ist, folgt aus 8.6. Er ist der kleinste Untermodul, der N enthält.

Für $N = \{m\} \subset M$ ist $\langle m\rangle = Rm$.

Für eine endliche Teilmenge $N = \{n_1, \ldots, n_k\} \subset M$ gilt

$$\langle n_1, \ldots, n_k\rangle = Rn_1 + \cdots + Rn_k = \sum_{i=1}^{k} Rn_i.$$

Für unendliches N kann man schreiben

$$\langle N\rangle = \{\sum_{i=1}^{k} r_i n_i \mid r_i \in R,\, n_i \in N,\, k \in \mathbb{N}\}.$$

Dies ist die Menge der endlichen Linearkombinationen von Elementen in N.

8.8 Definition
Sei M ein R-Modul. Eine Teilmenge N von M heißt *Erzeugendensystem* von M, wenn $\langle N\rangle = M$, wenn also jedes $m \in M$ Linearkombination von Elementen $n_1, \ldots, n_k \in N$ ist,

$$m = \sum_{i=1}^{k} r_i n_i \text{ für geeignete } r_i \in R.$$

Der R-Modul M heißt *endlich erzeugt* (auch *endlich erzeugbar*), wenn es eine endliche Teilmenge gibt, von der M erzeugt wird.

Jeder R-Modul M besitzt ein Erzeugendensystem N, etwa $N = M$.

Ein Erzeugendensystem wird häufig als indizierte Teilmenge (= Familie) $(n_i)_{i\in I}$ von M gegeben.

Beispiele

(1) ${}_{\mathbb{Z}}\mathbb{Z}$ ist endlich erzeugt, ${}_{\mathbb{Z}}\mathbb{Q}$ ist nicht endlich erzeugt.

(2) Für jeden Ring R ist R^3 endlich erzeugt, da etwa $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ ein Erzeugendensystem ist.

Natürlich ist für jedes $n \in \mathbb{N}$ der Modul R^n endlich erzeugt.

(3) Sei R ein kommutativer Ring. Dann ist der Polynomring $R[X]$ ein R-Modul, und 1, X, $X^2, \ldots$ ist ein (unendliches) Erzeugendensystem davon. Man kann sich leicht überlegen, daß es hier kein endliches Erzeugendensystem geben kann.

8.9 Definition
Sei $(U_i)_{i\in I}$ eine Familie von Untermoduln des R-Moduls M. Die Summe $\sum_{i\in I} U_i$ heißt *(innere) direkte Summe* der U_i, falls gilt

$$U_j \cap \sum_{i\in I\setminus\{j\}} U_i = \{0\} \text{ für alle } j \in I.$$

Man nennt dann $(U_i)_I$ eine *unabhängige* Familie von Untermoduln und schreibt

$$\sum_{i\in I} U_i = \bigoplus_{i\in I} U_i = \bigoplus_I U_i.$$

Bei einer endlichen Familie $(U_i)_{i\leq n}$ schreibt man $U_1 \oplus \cdots \oplus U_n$.

Speziell ist $U_1 + U_2$ genau dann direkte Summe, wenn $U_1 \cap U_2 = \{0\}$.

Ist $M = U_1 \oplus \cdots \oplus U_n$, so spricht man von einer *(endlichen) Zerlegung von M in direkte Summanden $U_1, \ldots, U_n$.*

Für beliebige Ringe R haben wir zum Beispiel die Zerlegungen

$$R^3 = (R, R, 0) \oplus (0, 0, R) = (R, 0, 0) \oplus (0, R, 0) \oplus (0, 0, R).$$

Die Bedeutung der unabhängigen Familien von Untermoduln liegt in folgenden Eigenschaften:

8.10 Hilfssatz
Für eine Familie $(U_i)_{i\in I}$ von Untermoduln eines R-Moduls M sind folgende Aussagen äquivalent:

(a) *Die Summe $\sum_I U_i$ ist direkt (also ist $(U_i)_I$ eine unabhängige Familie);*

(b) *jedes Element $u \in \sum_I U_i$ läßt sich eindeutig darstellen als*

$$u = u_{i_1} + \ldots + u_{i_k}, \text{ mit } u_{i_r} \in U_{i_r} \text{ und verschiedenen } i_r \in I;$$

(c) *aus $v_{i_1} + \ldots + v_{i_k} = 0$, mit $v_{i_r} \in U_{i_r}$ und verschiedenen $i_r \in I$, folgt $v_{i_1} = \ldots = v_{i_k} = 0$.*

Beweis: (a)⇒(b) Sei $u_{i_1} + \ldots + u_{i_k} = u'_{i_1} + \ldots + u'_{i_k}$. Dann gilt

$$U_{i_1} \ni u'_{i_1} - u_{i_1} = (u_{i_2} - u'_{i_2}) + \ldots + (u_{i_k} - u'_{i_k}) \in \sum_{j\in I\setminus\{i_1\}} U_j,$$

also $u'_{i_1} - u_{i_1} = 0$. Analog schließt man für die anderen i_r.

(b)⇒(c) ist offensichtlich.

(c)⇒(a) Sei $u_j \in U_j \cap \sum_{i\in I\setminus\{j\}} U_i$, also $u_j = \sum_{i\neq j} u_i$ und $0 = \sum_{i\neq j} u_i + (-u_j)$. Wegen (c) folgt daraus $u_j = 0$. □

Eng verwandt mit der Unabhängigkeit von Untermoduln ist der entsprechende Begriff für Elemente:

8.11 Definition
Eine endliche Menge $\{n_1, \ldots, n_k\}$ von Elementen eines R-Moduls M heißt *linear unabhängig*, wenn gilt:

$$\text{Ist } \sum_{i=1}^{k} r_i n_i = 0 \text{ für } r_i \in R, \text{ so folgt } r_1 = r_2 = \ldots = r_k = 0.$$

Eine (unendliche) Menge $N \subset M$ heißt *linear unabhängig*, wenn jede endliche Teilmenge linear unabhängig ist. Man nennt N *linear abhängig*, wenn N nicht linear unabhängig ist.

Nach Definition ist ein einzelnes Element $m \in M$ linear unabhängig, wenn aus $rm = 0$ auch $r = 0$ folgt. In einem Vektorraum gilt dies bekanntlich für alle Elemente $\neq 0$.

8.12 Hilfssatz
Für eine Familie $(n_i)_{i\in I}$ von Elementen eines R-Moduls M sind folgende Eigenschaften äquivalent:

(a) *$(n_i)_{i\in I}$ ist linear unabhängig;*

(b) *$(Rn_i)_{i\in I}$ ist eine unabhängige Familie von Untermoduln, und jedes n_i ist linear unabhängig;*

(c) *$\sum Rn_i = \bigoplus Rn_i$, und jedes n_i ist linear unabhängig.*

Beweis: Die Äquivalenz von (b) und (c) kennen wir aus 8.10.

(a)⇒(b) Ist $r_1 n_1 + \ldots + r_k n_k = 0$, dann folgt $r_i = 0$, also $r_i n_i = 0$ und $\sum Rn_i$ ist direkt nach 8.10.

(b)⇒(a) Ist $r_1 n_1 + \ldots + r_k n_k = 0$, so folgt $r_i n_i = 0$, da $\sum Rn_i$ direkt ist. Daraus ergibt sich nun $r_i = 0$, da n_i linear unabhängig ist. □

8.13 Definition
Eine Familie $(n_i)_{i\in I}$ von Elementen aus M heißt *Basis von M*, wenn gilt

(i) $(n_i)_I$ ist linear unabhängig;

(ii) $(n_i)_I$ ist Erzeugendensystem von M.

Ein Modul M heißt *freier Modul*, wenn er eine Basis besitzt.

Beispiele

(1) Für jeden Ring R ist $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ eine Basis in R^3. Man nennt sie die *kanonische Basis*.

Jedes Element aus R^3 ist auf genau eine Weise als Linearkombination der e_i darstellbar.

Analog erhält man für jedes $n \in \mathbb{N}$ eine kanonische Basis von R^n.

(2) Die entsprechende Bildung für eine unendliche Indexmenge I ist die (äußere) direkte Summe von R über I,

$$R^{(I)} = \{(a_i)_{i\in I} \in R^I \mid a_i \neq 0 \text{ für nur endlich viele } i \in I\},$$

die wir in Beispiel 8.2, (4) betrachtet haben. Die kanonische Basis davon ist die Familie

$$e_i = (\delta_{ij})_{j\in I} \text{ mit } \begin{cases} \delta_{ii} = 1 \\ \delta_{ij} = 0 \quad \text{für } i \neq j \end{cases}.$$

δ_{ij} nennt man das *Kronecker-Symbol*. Wir werden in 10.8 sehen, daß im wesentlichen alle freien Moduln von dieser Form sind.

Man beachte, daß ein beliebiger R-Modul *keine* Basis zu haben braucht. Zum Beispiel gibt es für $0 \neq n \in \mathbb{N}$ im $\mathbb{Z}$-Modul $\mathbb{Z}/n\mathbb{Z}$ kein (über $\mathbb{Z}$) linear unabhängiges Element.

Auch ${}_{\mathbb{Z}}\mathbb{Q}$ hat keine Basis, denn je zwei Elemente sind linear abhängig.

Wir können aber zeigen, daß es in jedem Modul eine maximale linear unabhängige Teilmenge gibt. Dies ist für $\mathbb{Z}/n\mathbb{Z}$ dann eben die leere Menge.

8.14 Hilfssatz
Sei M ein Modul über einem Ring R. Dann gilt:

Die Menge $\mathcal{U}$ aller linear unabhängigen Teilmengen von M ist bezüglich der Inklusion induktiv geordnet.

Beweis: Gibt es keine linear unabhängigen Elemente in M, dann ist $\mathcal{U} = \{\emptyset\}$. Sei $\mathcal{V}$ eine nicht-leere, linear geordnete Teilmenge von $\mathcal{U}$. Dann ist

$$W = \bigcup_{V\in\mathcal{V}} V$$

auch eine linear unabhängige Teilmenge von $\mathcal{V}$: Jede endliche Teilmenge von W ist in einem $V \in \mathcal{V}$ enthalten und damit linear unabhängig. Man sieht leicht, daß W Supremum von $\mathcal{V}$ ist. Somit ist $\mathcal{U}$ induktiv geordnet. □

Mit Hilfe des Zornschen Lemmas folgt daraus:

8.15 Korollar
Jeder R-Modul besitzt eine maximale linear unabhängige Teilmenge.

Über beliebigen Ringen braucht eine maximale linear unabhängige Teilmenge eines Moduls keine Basis zu sein. So bildet etwa in $\mathbb{Z}$ jedes Element $a \neq 0$ eine maximale linear unabhängige Teilmenge, die keine Basis ist falls, $a \neq 1, -1$.

Wir werden sehen, daß in Moduln über Divisionsringen (Vektorräumen) maximale linear unabhängige Teilmengen schon Basen sind. Daraus ergeben sich einige zusätzliche Eigenschaften von Vektorräumen, die wir im nächsten Abschnitt zusammenstellen wollen.

8.16 Aufgaben

(1) Man zeige, daß man in der Definition eines R-Moduls M das Axiom (M4) $1 \cdot m = m$ für alle $m \in M$ durch das folgende Axiom ersetzen kann:

$$(M4^*)\ 1 \cdot m \neq 0 \text{ für alle } m \in M \setminus \{0\}.$$

(2) Sei M ein R-Modul und U und V Untermoduln von M. Zeigen Sie: $U \cup V$ ist genau dann Untermodul von M, wenn $U \subset V$ oder $V \subset U$.

(3) Prüfen Sie, welche der Teilmengen $\mathbb{Z}$-Untermoduln von $\mathbb{Z}^3$ sind:

$$\begin{aligned} U_1 &:= \{(a,b,c) \in \mathbb{Z}^3 \mid a+b+c = 0\} \\ U_2 &:= \{(a,b,c) \in \mathbb{Z}^3 \mid a+b+c = 1\} \\ U_3 &:= \{(a,b,c) \in \mathbb{Z}^3 \mid a = 0\} \\ U_4 &:= \{(a,b,c) \in \mathbb{Z}^3 \mid a > 0\} \\ U_5 &:= \{(a,b,c) \in \mathbb{Z}^3 \mid a,b,c \in \mathbb{N}\} \end{aligned}$$

(4) Seien $n_1, \ldots, n_k$ Elemente des R-Moduls M. Zeigen Sie:

(i) $\langle n_1, \ldots, n_k \rangle = \bigcap \{U \subset M \mid U \text{ ist Untermodul mit } \{n_1, \ldots, n_k\} \subset U\}$.

(ii) *Für beliebige $a_1, \ldots, a_{k-1} \in R$ gilt:*

$$\langle n_1, \ldots, n_k \rangle = \langle n_1, n_2 + a_1 n_1, n_3 + a_2 n_2, \ldots, n_k + a_{k-1} n_{k-1} \rangle .$$

(5) Seien U und V Teilmengen des R-Moduls M. Man zeige:

(i) $\langle U \cup V \rangle = \langle U \rangle + \langle V \rangle$.

(ii) $\langle U \cap V \rangle \subset \langle U \rangle \cap \langle V \rangle$.

(iii) *Finden Sie Teilmengen $U_0, V_0 \subset \mathbb{Z}^2$ mit $\langle U_0 \cap V_0 \rangle \neq \langle U_0 \rangle \cap \langle V_0 \rangle$.*

(6) Sei M ein endlich erzeugbarer R-Modul. Man zeige:
Ist U ein (beliebiges) Erzeugendensystem von M, so gibt es eine endliche Teilmenge von U, die M erzeugt.

(7) R sei ein Ring und M ein R-Modul. Der Annullator *(oder* Annihilator*) von M in R ist definiert durch* $I := \{r \in R \mid rM = 0\}$.
Man zeige:

(i) *I ist ein Ideal in R;*

(ii) *M besitzt eine R/I-Modulstruktur.*

(8) R sei ein Ring und M ein R-Modul mit Untermoduln H, K, L. Zeige:

(i) $(H \cap K) + (H \cap L) \subset H \cap (K + L)$;

(ii) *ist* $K \subset H$, *so gilt das* Modulargesetz, *d.h.*

$$(H \cap K) + (H \cap L) = K + (H \cap L) = H \cap (K + L).$$

(iii) *in* $M = \mathbb{Z}^2$ *gibt es Untermoduln* H, K *und* L *mit*

$$(H \cap K) + (H \cap L) \neq H \cap (K + L).$$

9 Basis in Vektorräumen

Wir beginnen mit dem im letzten Abschnitt angekündigten Ergebnis.

9.1 Satz
Sei V ein Vektorraum über einem Divisionsring K. Dann besitzt V eine Basis.

Beweis: Nach 8.15 gibt es in V eine maximale linear unabhängige Teilmenge $(v_i)_{i\in I}$. Wir zeigen

$$\sum_{i\in I} Kv_i = \bigoplus_{i\in I} Kv_i = V.$$

Wegen der linearen Unabhängigkeit der gegebenen Untermoduln ist die Summe direkt. Es bleibt zu zeigen, daß sie gleich V ist.

Angenommen, es gibt ein $v \in V \setminus \bigoplus Kv_i$. Wegen der Maximalitätsforderung ist dann v linear abhängig von $(v_i)_{i\in I}$, also $Kv \cap \sum Kv_i \neq 0$ und

$$kv = k_1v_1 + \ldots + k_rv_r \text{ für geeignete } k, k_i \in K, k \neq 0.$$

Daraus folgt aber nun

$$v = \frac{k_1}{k}v_{i_1} + \ldots + \frac{k_r}{k}v_{i_r} \in \sum Kv_i,$$

ein Widerspruch. Somit gilt $V = \bigoplus Kv_i$, und $(v_i)_I$ ist eine Basis (vgl. 8.13). □

Über Divisionsringen können wir daraus verschiedene Kennzeichnungen von Basen ableiten:

9.2 Satz
Sei V ein Vektorraum über dem Divisionsring K. Für eine Familie $(v_i)_{i\in I}$ von Elementen aus V sind folgende Aussagen äquivalent:

(a) *$(v_i)_I$ ist eine Basis von V;*

(b) *$(v_i)_I$ ist ein minimales Erzeugendensystem von V;*

(c) *$(v_i)_I$ ist eine maximale Familie von linear unabhängigen Elementen.*

Beweis: (a)⇒(b) $(v_i)_I$ ist ein Erzeugendensystem. Da es linear unabhängig ist, kann eine echte Teilmenge kein Erzeugendensystem mehr sein.

(a)⇒(c) Jeder Vektor $w \in V$ ist Linearkombination von Elementen aus $(v_i)_I$. Also ist die Menge $\{v_i \mid i \in I\} \cup \{w\}$ linear abhängig.

(b)⇒(a) Es ist zu zeigen, daß jede endliche Teilmenge $\{v_1, \ldots, v_m\}$ von $(v_i)_I$ linear unabhängig ist. Angenommen $\sum_{i=0}^{m} k_i v_i = 0$ für $k_i \in K$. Nehmen wir ohne Einschränkung $k_m \neq 0$ an, dann gilt

$$v_m = -\frac{1}{k_m} \sum_{i=0}^{m-1} k_i v_i.$$

Damit wäre $(v_i)_{i \in I \setminus \{m\}}$ ein Erzeugendensystem von V, was der Minimalität von $(v_i)_{i \in I}$ widerspricht.

(c)⇒(a) Dies wurde im Beweis von Satz 9.1 gezeigt. □

9.3 Korollar
Sei V ein endlich erzeugbarer K-Vektorraum. Dann gilt:
Jedes Erzeugendensystem von V enthält eine Basis.

Beweis: Da V endlich erzeugbar ist, enthält jedes Erzeugendensystem von V eine endliche Teilmenge, die ebenfalls Erzeugendensystem ist. Durch Weglassen von überflüssigen Vektoren bekommt man daraus ein minimales Erzeugendensystem. Nach 9.2 ist dies eine Basis. □

Andererseits läßt sich jede linear unabhängige Teilmenge zu einer Basis ergänzen. Dies folgt aus dem

9.4 Austauschsatz von Steinitz
Sei V ein endlich erzeugbarer K-Vektorraum und $V_0 = \{v_1, \ldots, v_m\}$ ein Erzeugendensystem von V. Ist $\{w_1, \ldots, w_r\}$ eine linear unabhängige Teilmenge von V, so gilt:

(1) *$r \leq m$;*

(2) *man kann geeignete $V_\alpha := \{v_{\alpha_1}, \ldots, v_{\alpha_r}\} \subset V_0$ so auswählen, daß $\{w_1, \ldots, w_r\} \cup (V_0 \setminus V_\alpha)$ ein Erzeugendensystem von V bildet.*

Beweis: w_1 läßt sich schreiben als

$$w_1 = \sum_{i=1}^{m} r_i v_i \text{ mit geeigneten } r_i \in K.$$

Angenommen $r_{\alpha_1} \neq 0$. Dann ist

$$v_{\alpha_1} = \frac{1}{r_{\alpha_1}} w_1 + \sum_{i \neq \alpha_i} r_i' v_i.$$

Somit ist $\{w_1\} \cup \{V_0 \setminus \{v_{\alpha_1}\}\}$ ein Erzeugendensystem. Nun ist

$$w_2 = sw_1 + \sum_{i \neq \alpha_i} s_i v_i \text{ mit } s, s_i \in K,$$

wobei nicht alle $s_i = 0$ sein können, da w_1 und w_2 linear unabhängig sind.

Durch Induktion erhalten wir jeweils ein Erzeugendensystem

$$\{w_1, \ldots, w_t\} \cup \{V_0 \setminus \{v_{\alpha_1}, \ldots, v_{\alpha_t}\}\} \text{ für alle } t \leq \min(r, m).$$

Angenommen $r > m$. Dann ergibt sich für $t = m$, daß $\{w_1, \ldots, w_m\}$ ein Erzeugendensystem und somit Basis von V ist.

Damit ist aber die Menge $\{w_1, \ldots, w_m, w_{m+1}\}$ linear abhängig, im Widerspruch zur Voraussetzung. Also gilt $r \leq m$. □

Als einfache Folgerung erhalten wir den

9.5 Basisergänzungssatz
Jede linear unabhängige Teilmenge eines endlich erzeugbaren Vektorraums läßt sich zu einer Basis ergänzen.

Beweis: Ist $\{w_1, \ldots, w_r\}$ eine linear unabhängige Teilmenge von V, so ergänzt man diese zu einer maximalen linear unabhängigen Teilmenge. Diese ist nach 9.4 endlich und nach 9.2 eine Basis. □

Man beachte, daß die dadurch garantierte Ergänzung einer linear unabhängigen Menge zu einer Basis keineswegs eindeutig bestimmt ist. So kann man zum Beispiel die linear unabhängigen Vektoren $(1, 0, 0)$ und $(0, 1, 0)$ im K^3 sowohl durch $(0, 0, 1)$ als auch durch $(0, 1, 1)$ oder $(1, 1, 1)$ zu einer Basis ergänzen.

Es sei angemerkt, daß 9.5 auch für nicht notwendig endlich erzeugte Vektorräume richtig bleibt. Zum Beweis muß man dann das Zornsche Lemma heranziehen.

Eine weitere wichtige Folgerung aus 9.4 ist die Beobachtung:

9.6 Länge von Basen
In einem endlich erzeugbaren K-Vektorraum besteht jede Basis aus gleich vielen Elementen.

Beweis: Wird der Vektorraum V von m Elementen erzeugt, dann hat jede Basis höchstens m Elemente, d.h. jede Basis ist endlich.

Sind zwei Basen mit r bzw. s Elementen gegeben, so folgt aus 9.4 sowohl $r \leq s$ als auch $s \leq r$, also $r = s$. □

Damit wird folgende Bezeichnung sinnvoll:

9.7 Definition
Sei V ein endlich erzeugbarer Vektorraum über einem Divisionsring K. Dann nennt man die Anzahl der Elemente einer Basis von V die *Dimension von* V und schreibt dafür $\dim_K V$.

Ist V nicht endlich erzeugbar, so nennt man V *unendlich-dimensional* und setzt $\dim V = \infty$.

Auch für unendlich-dimensionale Vektorräume kann man eine *Dimension* definieren als die *Mächtigkeit* einer Familie von Basiselementen. Wir werden allerdings den Dimensionsbegriff nur für endlich-dimensionale Vektorräume gebrauchen.

Sehen wir uns die Dimensionen einiger vertrauter Moduln über einem Divisionsring K an:

Beispiele

(1) $\dim_K K = 1$;

(2) $\dim_K 0 = 0$, da 0 keine linear unabhängigen Elemente enthält;

(3) $\dim K^n = n$, $n \in \mathbb{N}$;

(4) $\dim K[X] = \infty$;

(5) $\dim_{\mathbb{Q}} \mathbb{R} = \infty$.

Als wichtige Eigenschaft stellen wir fest:

9.8 Dimension von Unterräumen
Sei K ein Divisionsring und V ein endlich-dimensionaler K-Vektorraum. Dann gilt für einen Unterraum $U \subset V$:

$$\begin{aligned} \dim_K U &\leq \dim_K V \text{ und} \\ \dim_K U &= \dim V \text{ genau dann, wenn } U = V. \end{aligned}$$

Beweis: In U kann es nicht mehr linear unabhängige Elemente geben als in V. Ist $\dim U = \dim V$, so enthält U eine maximale linear unabhängige Teilmenge, also eine Basis von V. □

Wir haben gesehen, daß es über jedem Ring R Moduln mit einer Basis gibt, z.B. R^n mit $n \in \mathbb{N}$. Die obige Beobachtung für Vektorräume wird jedoch für Untermoduln von R^n nicht mehr richtig sein. Ja wir können nicht einmal erwarten, daß Untermoduln und Faktormoduln von R^n wieder eine Basis haben.

Es ist bemerkenswert, daß die Forderung nach der Existenz von Basen in allen R-Moduln impliziert, daß R ein Divisionsring ist. Wir werden dies in 10.9 zeigen.

9.9 Aufgaben

(1) Sei V die Menge aller Abbildungen $f : \mathbb{R} \to \mathbb{R}$. Für $f, g \in V$ und $r \in \mathbb{R}$ definiert man

$$\begin{aligned} f+g: &\quad a \mapsto f(a)+g(a) \quad \text{für alle } a \in \mathbb{R}, \\ r \cdot f: &\quad a \mapsto r \cdot f(a) \quad \text{für alle } a \in \mathbb{R}. \end{aligned}$$

(i) Zeigen Sie, daß V ein $\mathbb{R}$-Vektorraum ist.

(ii) Welche der folgenden Teilmengen sind $\mathbb{R}$-Untervektorräume von V?

$$\begin{aligned} W_1 &:= \{f \in V \mid f(3) = 0\} \\ W_2 &:= \{f \in V \mid f(7) = f(1)\} \\ W_3 &:= \{f \in V \mid f(7) = 2 + f(1)\} \\ W_4 &:= \{f \in V \mid f(-a) = -f(a) \text{ für alle } a \in \mathbb{R}\} \\ W_5 &:= \{f \in V \mid f \text{ ist stetig}\}. \end{aligned}$$

(2) Sei (v_1, v_2, v_3) eine linear unbhängige Familie von Vektoren eines $\mathbb{R}$-Vektorraumes V. Zeigen Sie:

(i) $(v_1 + v_2, v_2 + v_3, v_3 + v_1)$ *ist linear unabhängig.*

(ii) $(v_1 + v_3, v_1 + v_2 + v_3, v_1 - 2v_2 + v_3)$ *ist linear abhängig.*

(3) Untersuchen Sie, welche der folgenden Familien von Vektoren des $\mathbb{R}^3$ linear unabhängig, Erzeugendensysteme, Basen des $\mathbb{R}^3$ sind:

$$F_1 = \left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \right), \qquad F_2 = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right)$$

$$F_3 = \left(\begin{pmatrix} 1 \\ 5 \\ -4 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ -6 \\ -2 \end{pmatrix}, \begin{pmatrix} 4 \\ -5 \\ -1 \end{pmatrix} \right), \; F_4 = \left(\begin{pmatrix} \sqrt{2} \\ \sqrt{8} \\ \sqrt{18} \end{pmatrix}, \begin{pmatrix} \frac{1}{3}\sqrt{3} \\ \frac{2}{3}\sqrt{3} \\ \sqrt{18} \end{pmatrix} \right).$$

(4) K sei ein Körper, V ein endlich-dimensionaler K-Vektorraum und U ein Untervektorraum von V. Man zeige, daß es einen Untervektorraum W von V mit $V = U \oplus W$ gibt.
Ist der Untervektorraum W eindeutig bestimmt?

(5) Es seien folgende Teilmengen von $\mathbb{R}^3$ gegeben:

$$\begin{aligned} A_1 &:= \{(1,1,1), (1,0,-1)\} \\ A_2 &:= \{(1,1,1), (1,0,-1), (0,\pi,-\pi)\} \\ A_3 &:= \{(1,1,1), (1,0,-1), (0,\sqrt{2},\pi/2), (44,44,4444)\} \\ A_4 &:= \{(1,1,1), (1,0,-1), (4,3,2), (1,-1,-3)\}. \end{aligned}$$

(a) *Man begründe die Antworten auf folgende Fragen:*

(i) *Ist* A_i, $i = 1, 2, 3, 4$, *linear abhängig oder unabhängig?*

(ii) *Ist* A_i, $i = 1, 2, 3$, *ein Erzeugendensystem für* $\mathbb{R}^3$*?*

(iii) *Ist* A_i, $i = 1, 2, 3, 4$, *eine Basis von* $\mathbb{R}^3$*?*

(b) *Man zeige*

$$\langle A_2 \rangle = \langle A_3 \rangle \neq \langle A_2 \cap A_3 \rangle = \langle A_1 \rangle .$$

(c) *Man zeige* $A_4 \subset \langle A_1 \rangle$. *Hieraus bestimme man* $\dim \langle A_4 \rangle$ *und beantworte die Frage (a.ii) für* $i = 4$.
Außerdem beweise man die Gleichheit

$$\langle A_4 \rangle = \{(\alpha, \beta, 2\beta - \alpha) \mid \alpha, \beta \in \mathbb{R}\}.$$

(6) Es seien folgende Unterräume von $\mathbb{R}^4$ *gegeben:*

$$\begin{array}{rcll} U & := & \langle (-5,0,5,5), (1,5,2,1), (-3,-5,0,1) \rangle & \textit{und} \\ V & := & \langle (2,1,1,0), (-1,1,4,3), (9,3,0,-3) \rangle . & \end{array}$$

(a) *Man bestimme* $\dim U$, $\dim V$, $\dim(U + V)$ *und* $\dim(U \cap V)$.

(b) *Man gebe jeweils eine Basis von* $U + V$ *und von* $U \cap V$ *an.*

(7) Gegeben seien die Vektoren

$$\begin{array}{rcllrcl} v_1 & := & (1,0,2,-4,0), & & v_2 & := & (0,-1,5,6,1), \\ v_3 & := & (3,0,6,-12,0), & & v_4 & := & (2,-1,3,0,-2) \textit{ und} \\ v_5 & := & (-6,1,1,12,8) \textit{ aus } \mathbb{R}^5. & & & & \end{array}$$

Es sei $A := \{v_1, v_2, v_3, v_4, v_5\}$ *und* $U := \langle A \rangle$.

(i) *Man berechne* $\dim U$.

(ii) *Bestimmen Sie alle maximalen linear unabhängigen Teilmengen von* A.

(8) Gegeben seien die Vektoren aus $\mathbb{C}^3$*:*
$u := (2, 1 + i, 2 - i)$, $v := (1 - i, 2 - 2i, 3i)$ *und* $w := (0, -3 + i, 5 - 4i)$.
Man beweise oder widerlege:

(i) u, v, w *sind linear unabhängig im* $\mathbb{C}$*–Vektorraum* $\mathbb{C}^3$.

(ii) u, v, w *sind linear unabhängig im* $\mathbb{R}$*–Vektorraum* $\mathbb{C}^3$.

10 Homomorphismen von Moduln

Wie bei Gruppen und Ringen sind auch bei Moduln die strukturverträglichen Abbildungen von großer Bedeutung. Sie werden ebenfalls als *Homomorphismen* bezeichnet. Das Besondere daran ist, daß man sie meist mit Hilfe von Matrizen darstellen kann. Dadurch werden sie rechnerisch leicht zugänglich. Wir werden dies in einem späteren Abschnitt ausführlich behandeln. Bevor wir uns mit diesen Rechnungen befassen, wollen wir die theoretischen Grundlagen verstehen.

10.1 Definition
Eine Abbildung $f : M \to N$ von R-Moduln heißt *(Modul-)Homomorphismus*, wenn für alle $x, y \in M$ und $r \in R$ gilt

$$f(x+y) = f(x) + f(y) \quad \text{und} \quad f(rx) = rf(x).$$

Man nennt diese Abbildungen auch *R-Homomorphismen* oder *R-lineare Abbildungen*.

Mit $\mathrm{Hom}_R(M, N)$ bezeichnen wir die Menge der Homomorphismen von M nach N. Die Begriffe *Mono-*, *Epi-*, *Iso-*, *Endo-* und *Automorphismus* werden analog zur Definition 6.8 gebildet.

Das Urbild der $0 \in N$ unter $f : M \to N$ nennt man wieder den *Kern von f*,

$$\mathrm{Kern}\, f = \{m \in M \mid f(m) = 0\}.$$

Da f auch Gruppenhomomorphismus $(M, +) \to (N, +)$ ist, sind die bereits gezeigten Eigenschaften von $\mathrm{Kern}\, f$ auch hier gültig (vgl. 6.15).

Zudem stellen wir fest, daß $\mathrm{Kern}\, f$ ein Untermodul von M ist: Für $k \in \mathrm{Kern}\, f$ und $r \in R$ gilt nämlich $f(rk) = rf(k) = 0$, also $rk \in \mathrm{Kern}\, f$.

Man bestätigt leicht folgende elementare

Eigenschaften
Sei $f : M \to N$ ein Homomorphismus von R-Moduln. Dann gilt

(1) $f(0_M) = 0_N$.

(2) Für alle $x \in M$ ist $f(-x) = -f(x)$.

(3) Ist $g : N \to L$ ein R-Homomorphismus, so ist auch $g \circ f : M \to L$ ein R-Homomorphismus.

(4) f ist genau dann Homomorphismus, wenn

$$f(rx + sy) = rf(x) + sf(y) \quad \text{für alle } r, s \in R,\ x, y \in M.$$

Beispiele von Homomorphismen

(1) Für jeden R-Modul M ist $\mathrm{id}_M : M \to M$ ein R-Homomorphismus.

(2) Sind M und N R-Moduln, dann ist die *Nullabbildung*

$$M \to N, \quad m \mapsto 0 \text{ für alle } m \in M,$$

ein R-Homomorphismus.

(3) Sei R ein kommutativer Ring und M ein R-Modul. Für alle $r \in R$ ist

$$M \to M,\ m \mapsto rm \text{ für alle } m \in M,$$

ein R-Homomorphismus (*Homothetie*).
Für nicht-kommutative Ringe R sind die Homothetien nur $Z(R)$-Homomorphismen (mit $Z(R)$= Zentrum von R).

(4) Sind G und H abelsche Gruppen, dann ist jeder Gruppenhomomorphismus $f : G \to H$ auch $\mathbb{Z}$-Homomorphismus.

Die Menge $\mathrm{Hom}_R(M, N)$ von Homomorphismen zwischen den Moduln M und N ist eine Teilmenge von $\mathrm{Abb}(M, N)$. Wie wir in 8.2,(4) gesehen haben, trägt $\mathrm{Abb}(M, N)$ eine Modulstruktur, die bestimmt ist durch die Modulstruktur von N. Für $f, g \in \mathrm{Abb}(M, N)$ und $r \in R$ haben wir für jedes $m \in M$:

$$\begin{aligned} (f+g)(m) &:= f(m) + g(m) \\ (rf)(m) &:= r(f(m)). \end{aligned}$$

Es stellt sich die Frage, ob sich diese Struktur auch auf $\mathrm{Hom}_R(M, N)$ übertragen läßt. Die Antwort darauf lautet:

10.2 Satz
Seien M und N Moduln über einem Ring R. Dann gilt:

(1) $\mathrm{Hom}_R(M, N)$ *ist eine Untergruppe von* $\mathrm{Abb}(M, N)$.

(2) *Ist R ein kommutativer Ring, so ist* $\mathrm{Hom}_R(M, N)$ *ein R-Untermodul von* $\mathrm{Abb}(M, N)$.

Beweis: (1) Es ist zu zeigen, daß für $f, g \in \mathrm{Hom}_R(M, N)$ auch $f + g$ und $-f$ R-Homomorphismen sind. Dies ersieht man aus ($m_i \in M$, $a \in R$)

$$\begin{aligned} (f+g)(m_1+m_2) &= f(m_1+m_2) + g(m_1+m_2) \\ &= (f+g)(m_1) + (f+g)(m_2), \text{ und} \\ (f+g)(am) &= f(am) + g(am) \\ &= af(m) + ag(m) \\ &= a[(f+g)(m)]. \end{aligned}$$

(2) Sei nun R kommutativ. Wegen (1) bleibt nur noch nachzuweisen, daß für $f \in \mathrm{Hom}_R(M,N)$ und $r \in R$ auch rf ein R-Homomorphismus ist. Dazu betrachtet man

$$\begin{aligned}(rf)(m_1+m_2) &= r(f(m_1+m_2)) = r(f(m_1)) + r(f(m_2)) \\ &= (rf)(m_1) + (rf)(m_2) \quad \text{und} \\ (rf)(am) &= r(f(am)) = (ra)f(m) = (ar)f(m) \\ &= a(rf)(m).\end{aligned}$$

Für die zweite Beziehung haben wir tatsächlich die Kommutativität von R benötigt. □

10.3 Faktormoduln

Sei U Untermodul des R-Moduls M. Dann wird die Faktorgruppe

$$M/U = \{m+U \mid m \in M\}$$

ein R-Modul durch die Skalarmultiplikation

$$r(m+U) := rm+U \text{ für } r \in R, m \in M,$$

und die kanonische Projektion $p_U : M \to M/U$ ist ein R-Homomorphismus.

Beweis: Da U auch Untergruppe von $(M,+)$ ist, wird die Menge der Restklassen M/U eine additive Gruppe mit (vgl. 6.16)

$$(m+U)+(n+U) := (m+n)+U.$$

Wir müssen uns noch überlegen, ob die Definition der Skalarmultiplikation unabhängig von der Auswahl der Repräsentanten ist.

Sei $m+U = m'+U$ für $m, m' \in M$, also $m'-m \in U$. Dann gilt

$$\begin{aligned}r(m'+U) &= rm'+U \\ &= r[m+(m'-m)]+U = rm+U \\ &= r(m+U).\end{aligned}$$

Wir wissen schon, daß $p_U : M \to M/U$ ein Gruppenhomomorphismus ist. Außerdem gilt für $m \in M$, $r \in R$:

$$p_U(rm) = rm+U = r(m+U) = rp_U(m).$$

Somit ist p_U ein R-Homomorphismus. □

Mit diesen Komponenten können wir nun den Homomorphiesatz für Gruppen auch auf Moduln übertragen.

10.4 Homomorphiesatz für Moduln
Sei $f : M_1 \to M_2$ ein R-Modulhomomorphismus. Dann ist die kanonische Projektion

$$p_f : M_1 \to M_1/\operatorname{Kern} f,\ a \mapsto [a] = a + \operatorname{Kern} f,$$

ein (Modul-)Epimorphismus, und es gibt genau einen Monomorphismus

$$\overline{f} : M_1/\operatorname{Kern} f \to M_2$$

mit $f = \overline{f} \circ p_f$, d.h. wir haben das kommutative Diagramm

$$\begin{array}{ccc} M_1 & \xrightarrow{f} & M_2 \\ {\scriptstyle p_f}\searrow & & \nearrow{\scriptstyle \overline{f}} \\ & M_1/\operatorname{Kern} f & \end{array}$$

Der nächste Satz bringt die Bedeutung von Erzeugendensystem und Basis für lineare Abbildungen zum Ausdruck:

10.5 Satz. *M und N seien R-Moduln.*

(1) *Ist $U \subset M$ ein Erzeugendensystem, und sind $f, g \in \operatorname{Hom}_R(M, N)$ mit*

$$f(u) = g(u) \quad \text{für alle } u \in U,$$

dann ist $f = g$.

(2) *Ist $(m_i)_{i\in I}$ eine Basis von M und $(n_i)_{i\in I}$ eine Familie von Elementen aus N, dann gibt es genau ein $h \in \operatorname{Hom}_R(M, N)$ mit*

$$h(m_i) = n_i \quad \text{für alle } i \in I.$$

Die erste Aussage bedeutet, daß zwei lineare Abbildungen schon dann gleich sind, wenn sie auf einem Erzeugendensystem übereinstimmen. Die zweite Behauptung besagt, daß man jede Abbildung einer Basis von M in N eindeutig zu einer linearen Abbildung von M in N fortsetzen kann.

Beweis: (1) Da U Erzeugendensystem ist, hat jedes $m \in M$ die Form

$$m = r_1u_1 + \ldots + r_ku_k \ \text{ mit } r_i \in R, u_i \in U.$$

Dann gilt:

$$\begin{aligned} f(m) = f(r_1u_1 + \ldots + r_ku_k) &= r_1f(u_1) + \ldots + r_kf(u_k) \\ &= r_1g(u_1) + \ldots + r_kg(u_k) \\ &= g(r_1u_1 + \ldots + r_ku_k) \\ &= g(m). \end{aligned}$$

Dies bedeutet gerade $f = g$.

(2) Für jede endliche Summe $m = \sum r_i m_i \in M$ mit $r_i \in R$ setze man

$$h(m) = h(\sum r_i m_i) := \sum r_i n_i.$$

Da zu jedem $m \in M$ diese r_i eindeutig festgelegt sind, ist h eine Abbildung von M in N. Wir zeigen, daß h ein R-Homomorphismus ist:

Sei $m' = \sum r_i' m_i$ ein weiteres Element aus M. Dann gilt (alle auftretenden Summen sind endlich):

$$\begin{aligned} h(m+m') &= h[\textstyle\sum(r_i + r_i')m_i] &&= \textstyle\sum(r_i + r_i')n_i \\ &= \textstyle\sum r_i n_i + \sum r_i' n_i &&= h(m) + h(m'). \end{aligned}$$

Also ist h ein Gruppenhomomorphismus $(M,+) \to (N,+)$. Für $s \in R$ gilt:

$$h(am) = \sum ar_i n_i = a \sum r_i n_i = ah(m).$$

Somit ist h ein R-Homomorphismus.

Die Eindeutigkeit von h folgt aus (1), da eine Basis insbesondere Erzeugendensystem ist. □

Man beachte, daß die $(n_i)_{i\in I}$ nicht linear unabhängig sein müssen. Im allgemeinen ist auch das Bild einer linear unabhängigen Teilmenge nicht linear unabhängig. Es gilt:

10.6 Satz
Sei $f : M \to N$ ein Homomorphismus von R-Moduln.

(1) *Ist $U \subset M$ ein Erzeugendensystem von M, dann ist $f(U)$ ein Erzeugendensystem von $f(M)$.*

(2) *Ist f injektiv und $U \subset M$ eine linear unabhängige Teilmenge von M, so ist $f(U)$ linear unabhängig.*

(3) *Ist $U \subset M$ eine Basis von M, und ist*

(i) *f injektiv, dann ist $f(U)$ Basis von $f(M)$;*

(ii) *f surjektiv, dann ist $f(U)$ Erzeugendensystem von N;*

(iii) *f Isomorphismus, dann ist $f(U)$ Basis von N.*

Beweis: (1) Sei $n \in f(M)$ und $m \in M$ mit $f(m) = n$. Schreiben wir

$$m = r_1u_1 + \ldots + r_ku_k, \text{ mit } u_i \in U, r_i \in R,$$

so erhalten wir

$$n = f(m) = r_1f(u_1) + \ldots + r_kf(u_k) \in \langle f(U) \rangle ,$$

d.h. $f(M)$ wird von $f(U)$ erzeugt.

(2) Seien $f(u_1), \ldots, f(u_k) \in f(U)$ und $r_1f(u_1) + \ldots + r_kf(u_k) = 0$.
Dann ist $f(r_1u_1 + \ldots + r_ku_k) = 0$, und wegen der Injektivität von f gilt auch $r_1u_1 + \ldots + r_ku_k = 0$. Da die u_i linear unabhängig sind, bedeutet dies $r_i = 0$ für $i = 1, \ldots, k$. Also sind die $f(u_i)$ linear unabhängig.

(3) Diese Aussagen sind einfache Konsequenzen aus (1) und (2). □

Als unmittelbare Folgerung halten wir fest:

10.7 Korollar
Sei $f : M \to N$ ein Epimorphismus.

(1) *Ist M endlich erzeugbar, dann ist auch N endlich erzeugbar.*

(2) *Ist f Isomorphismus und M ein freier Modul, dann ist auch N frei.*

(3) *Ist f Isomorphismus und M endlich erzeugbar und frei, dann ist N endlich erzeugbar und frei.*

Damit können wir nun die Bedeutung der freien Moduln $R^{(I)}$ zeigen:

10.8 Korollar

(1) *Ein R-Modul M ist genau dann frei, wenn $M \simeq R^{(I)}$ für eine geeignete Indexmenge I ist.*

(2) *M ist endlich erzeugbar und frei, wenn $M \simeq R^k$ für ein $k \in \mathbb{N}$ ist.*

(3) *Jeder R-Modul M ist homomorphes Bild eines Moduls $R^{(I)}$.*
Jeder endlich erzeugbare R-Modul ist Bild von R^k mit $k \in \mathbb{N}$.

Beweis: (1) Sei $(n_i)_I$ eine Basis von M und bezeichne $(e_i)_I$ die kanonische Basis von $R^{(I)}$ (siehe Beispiel nach 8.13). Nach 10.5 erhalten wir dann einen Homomorphismus

$$f : R^{(I)} \to M,\ e_i \mapsto n_i \text{ für } i \in I.$$

Da $(n_i)_I$ Erzeugendensystem ist, ist f epimorph. f ist auch monomorph, da die $f(e_i) = n_i$ linear unabhängig sind.

Aus 10.6 folgt, daß jeder zu $R^{(I)}$ isomorphe Modul ebenfalls frei ist.

(2) Dies ergibt sich aus (1) mit $I = \{1, \ldots, k\}$.

(3) Die Abbildung von (1) ergibt für jedes Erzeugendensystem $(n_i)_I$ von M einen Epimorphismus. □

Mit den nun bereitgestellten Mitteln können wir leicht zeigen, was es bedeutet, wenn jeder R-Modul frei ist:

10.9 Satz

Für einen Ring R sind folgende Aussagen äquivalent:

(a) *Jeder R-Modul besitzt eine Basis;*

(b) *R ist ein Divisionsring.*

Beweis: (b)⇒(a) haben wir bereits in Satz 9.1 gezeigt.

(a)⇒(b) Nach dem Satz von Krull (7.19) gibt es ein maximales Linksideal $U \subset R$. Dann ist R/U ein R-Modul mit den einzigen Untermoduln 0 und R/U. Da alle R-Moduln eine Basis haben sollen, ist nach 10.8 $R/U \simeq R^k$. Dabei muß natürlich $k = 1$ sein, also $R/U \simeq R$.

Dann hat aber auch R keine Linksideale (Untermoduln) $\neq \{0\}, R$, und somit ist R Divisionsring (Satz 7.9). □

10.10 Aufgaben

Beweisen Sie, daß U und W Unterräume von $\mathbb{R}^4$ sind, mit

$$U := \{\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_4 \end{pmatrix} \in \mathbb{R}^4 \mid \sum_{i=1}^{4} \alpha_i = 0\} \text{ und}$$

$$W := \{\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_4 \end{pmatrix} \in \mathbb{R}^4 \mid \alpha_2 = 2\alpha_1 \text{ und } \alpha_3 = 3\alpha_4\}.$$

Bestimmen Sie jeweils eine Basis von U, W, $U \cap W$, $\mathbb{R}^4/U$ und $\mathbb{R}^4/W$.

11 Dimension und lineare Abbildungen

In diesem Abschnitt betrachten wir lineare Abbildungen von Vektorräumen über einem Divisionsring K.

Kerne und Bilder von solchen Abbildungen sind wiederum Vektorräume. Deren Dimensionen (vgl. 9.7) geben Aufschluß über Eigenschaften der Abbildung. Grundlage dafür ist die

11.1 Dimensionsformel
Sei $f : V \to W$ eine lineare Abbildung von Vektorräumen. Dann gilt

$$\dim V = \dim(\text{Bild } f) + \dim(\text{Kern } f).$$

Beweis: Formal gilt diese Beziehung auch für unendlich-dimensionale Vektorräume. Interessant ist sie jedoch nur für endlich-dimensionale V.

Sei also $\dim V = n < \infty$. Wähle $v_1, \ldots, v_p$ als Basis des Unterraums Kern f von V, wobei $p \leq n$. Wir ergänzen diese durch $v_{p+1}, \ldots, v_n$ zu einer Basis von V. Dann ist $f(v_{p+1}), \ldots, f(v_n)$ (und 0) das Bild der Basis $v_1, \ldots, v_n$ und somit ein Erzeugendensystem von Bild f.

Wir zeigen, daß dies linear unabhängige Vektoren sind und daher eine Basis von Bild f bilden. Angenommen

$$0 = \sum_{i=p+1}^{n} r_i f(v_i) = f(\sum_{i=p+1}^{n} r_i v_i),$$

also $\sum\limits_{i=p+1}^{n} r_i v_i \in \text{Kern } f$ und folglich

$$\sum_{i=p+1}^{n} r_i v_i = \sum_{i=1}^{p} s_i v_i \text{ für geeignete } s_i \in R.$$

Wegen der linearen Unabhängigkeit der v_i folgt daraus

$$r_i = 0 \text{ für alle } i = p+1, \ldots, n.$$

Somit ist $f(v_{p+1}), \ldots, f(v_n)$ eine Basis von Bild f und $\dim(\text{Bild } f) = n - p$. Wir haben also

$$\dim(\text{Kern } f) + \dim(\text{Bild } f) = p + n - p = n = \dim V.$$

Man beachte, daß hierbei die Dimension von W keine Rolle spielt. □

11.2 Korollar
Für jeden Unterraum U eines endlich-dimensionalen Vektorraums V gilt

$$\dim V/U = \dim V - \dim U\,.$$

Beweis: U ist Kern der linearen Abbildung $p : V \to V/U$. Nach 11.1 gilt daher

$$\dim V = \dim V/U + \dim U\,.$$

□

11.3 Korollar
Für Vektorräume V_1, V_2 gilt

$$\dim V_1 \oplus V_2 = \dim V_1 + \dim V_2\,.$$

Beweis: Hat einer der Vektorräume unendliche Dimension, so ist die Gleichung richtig. Betrachten wir also endlich-dimensionale Räume. Die Projektion

$$p_2 : V_1 \oplus V_2 \to V_2,\ (v_1, v_2) \mapsto v_2,$$

ist eine lineare Abbildung mit $\operatorname{Bild} p_2 = V_2$ und $\operatorname{Kern} p_2 = V_1 \oplus 0 \simeq V_1$. Nun liefert wieder die Dimensionsformel das gewünschte Ergebnis. □

Die Kombination der vorangehenden Beobachtungen ergibt ein weiteres

11.4 Korollar
Für Unterräume U, V eines Vektorraums W gilt

$$\dim U + \dim V = \dim(U + V) + \dim(U \cap V)\,.$$

Beweis: Wiederum können wir uns auf endliche Dimensionen beschränken. Wir bilden die äußere direkte Summe $U \oplus V$ $(= U \otimes V)$ und betrachten die Abbildung

$$\mu : U \oplus V \to U + V,\ (u, v) \mapsto u + v.$$

Dann ist $\operatorname{Bild} \mu = U + V$ und

$$\operatorname{Kern} \mu = \{(a, -a) \mid a \in U \cap V\} \simeq U \cap V.$$

Somit gilt $\begin{aligned} \dim(U \oplus V) &= \dim U + \dim V \quad \text{(nach 11.3)} \\ &= \dim(U + V) + \dim(U \cap V) \quad \text{(nach 11.3).} \end{aligned}$

□

Wegen der Bedeutung der Dimension des Bildes wurde dafür eine kürzere Bezeichnung eingeführt:

11.5 Definition
Sei $f : V \to W$ ein Homomorphismus von Vektorräumen. Die Dimension von $\operatorname{Bild} f$ nennt man den *Rang von f* und schreibt dafür $\operatorname{Rang} f$.

Nach der Dimensionsformel ist $\operatorname{Rang} f = \dim V - \dim \operatorname{Kern} f$.
Ist also $\dim V$ endlich, so ist auch $\operatorname{Rang} f$ endlich.
Ist $\operatorname{Bild} f$ nicht endlich-dimensional, so schreibt man $\operatorname{Rang} f = \infty$.

Wir notieren folgende Eigenschaften von $\operatorname{Rang} f$:

11.6 Satz
Für einen Homomorphismus $f : V \to W$ von endlich-dimensionalen Vektorräumen gilt:

(1) $\operatorname{Rang} f \leq \min(\dim V, \dim W)$.

(2) *f ist genau dann injektiv, wenn* $\operatorname{Rang} f = \dim V$.

(3) *f ist genau dann surjektiv, wenn* $\operatorname{Rang} f = \dim W$.

(4) *f ist genau dann bijektiv, wenn* $\dim V = \operatorname{Rang} f = \dim W$.

Beweis: (1) Da $\operatorname{Bild} f$ Unterraum von W ist, gilt nach 9.8

$$\operatorname{Rang} f = \dim \operatorname{Bild} f \leq \dim W.$$

Nach der Dimensionsformel ist natürlich auch $\operatorname{Rang} f \leq \dim V$.

(2) Falls $\operatorname{Rang} f = \dim V$ gilt, dann folgt aus der Dimensionsformel $\dim \operatorname{Kern} f = 0$, also $\operatorname{Kern} f = 0$, d.h. f ist injektiv.

(3) Gilt $\operatorname{Rang} f = \dim W$, dann ist $\operatorname{Bild} f$ ein Unterraum von W mit gleicher Dimension, also $\operatorname{Bild} f = W$ (vgl. 9.8), d.h. f ist surjektiv.

(4) ist eine Folgerung aus (2) und (3). □

Angewendet auf Endomorphismen, ergeben diese Betrachtungen:

11.7 Korollar
Für einen Endomorphismus $f : V \to V$ eines endlich-dimensionalen Vektorraums sind folgende Aussagen äquivalent:

(a) *f ist injektiv;*

(b) *f ist surjektiv;*

(c) *f ist bijektiv;*

(d) $\operatorname{Rang} f = \dim V$.

Beweis: Die Aussagen folgen direkt aus 11.6 mit $W = V$. □

11.8 Aufgaben

(1) Betrachten Sie die Unterräume in $\mathbb{R}^4$

$$\begin{aligned} U &= \langle (2,1,4,-1),(1,1,3,-1),(-1,-1,-3,1)\rangle \text{ und} \\ V &= \langle (3,0,3,0),(1,-1,2,-2),(1,1,0,2)\rangle\,. \end{aligned}$$

(i) *Bestimmen Sie* $\dim U$, $\dim V$, $\dim(U+V)$ *und* $\dim(U\cap V)$.

(ii) *Geben Sie eine Basis von* $U+V$ *und von* $U\cap V$ *an.*

(2) Man finde alle Untervektorräume von $\mathbb{R}^2$ und $\mathbb{R}^3$ und interpretiere sie geometrisch.

(3) R sei ein Ring mit Eins. Man zeige, daß

$$U := \{(a_n)_{n\in\mathbb{N}} \in R^{\mathbb{N}} \mid \text{für alle } n\in\mathbb{N} : a_{n+2} = a_{n+1} + a_n\}$$

ein zu R^2 isomorpher R-Untermodul von $R^{\mathbb{N}}$ ist.

(4) V sei ein dreidimensionaler $\mathbb{R}$-Vektorraum mit Basis (b_1,b_2,b_3) und W ein zweidimensionaler $\mathbb{R}$-Vektorraum mit Basis (c_1,c_2).
Die lineare Abbildung $f: V\to W$ sei gegeben durch ($\alpha_i\in\mathbb{R}$)

$$f(\alpha_1 b_1+\alpha_2 b_2+\alpha_3 b_3) = (\alpha_1 - 2\alpha_2 + 3\alpha_3)c_1 + (2\alpha_1+\alpha_2-\alpha_3)c_2.$$

(i) *Ist f injektiv?*

(ii) *Ist f surjektiv?*

(iii) *Bestimmen Sie die Dimensionen von Bild und Kern von f.*

Kapitel 4

Homomorphismen und Matrizen

Lineare Abbildungen von *freien* Moduln lassen sich mit Hilfe von *Matrizen* beschreiben. Dies gibt die Möglichkeit zur rechnerischen Bestimmung der Eigenschaften solcher Abbildungen.

Da über einem Divisionsring K alle K-Moduln frei sind, sind diese Methoden insbesondere für alle (endlich-dimensionalen) Vektorräume anwendbar. Darüberhinaus lassen Matrizen über Divisionsringen weitere Bildungen zu, die im nachfolgenden Abschnitt angegeben werden.

Die dabei entwickelten Methoden erlauben eine übersichtliche Behandlung von linearen Gleichungssystemen.

12 Homomorphismen und Matrizen

Zunächst wollen wir freie Moduln über beliebigen Ringen R betrachten.

12.1 Beschreibung linearer Abbildungen

Sei $f : M \to N$ eine lineare Abbildung von R-Moduln. Wir haben in Satz 10.5 gesehen, daß f durch die Werte auf einem Erzeugendensystem von M eindeutig bestimmt ist.

Ist also $x_1, \ldots, x_m$ ein endliches Erzeugendensystem von M, so ist f durch die Bilder $f(x_1), \ldots, f(x_m)$ festgelegt.

Ist nun $y_1, \ldots, y_n$ ein Erzeugendensystem von N, so lassen sich die $f(x_i)$ als Linearkombinationen der $y_1, \ldots, y_n$ darstellen, etwa durch

$$\begin{aligned} f(x_1) &= a_{11}\, y_1 + a_{12}\, y_2 + \ldots + a_{1n}\, y_n, \\ f(x_2) &= a_{21}\, y_1 + a_{22}\, y_2 + \ldots + a_{2n}\, y_n, \\ &\vdots \\ f(x_m) &= a_{m1}\, y_1 + a_{m2}\, y_2 + \ldots + a_{mn}\, y_n, \end{aligned}$$

mit geeigneten Koeffizienten $a_{ij} \in R$.

Sind $x_1, \ldots, x_m \subset M$ und $y_1, \ldots, y_n \subset N$ sogar *Basen*, so sind die $m \cdot n$ Elemente $a_{ij} \in R$ eindeutig bestimmt, und f ist durch diese Daten eindeutig festgelegt.

Man schreibt diese Elemente in ein Schema, das man (m,n)-*Matrix* nennt:

$$\text{Matrix}(a_{ij}) = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

Das n-Tupel $(a_{i1}, \ldots, a_{in})$ bezeichnet man als i-te *Zeile* und das m-Tupel $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ als j-te *Spalte* der Matrix.

m ist die *Zeilenzahl* und n die *Spaltenzahl* von (a_{ij}).

Bemerkung: Die Zuordnung der Matrix (a_{ij}) zu der linearen Abbildung f in 12.1 ist in gewissem Sinne willkürlich. Man hätte auch eine Matrix wählen können, in der Zeilen und Spalten vertauscht sind. Dies macht keinen Unterschied im mathematischen Gehalt, führt aber gelegentlich zu anderen Formeln.

Eine Matrix läßt sich als Abbildung

$$A : \{1, \cdots, m\} \times \{1, \cdots n\} \to R, \quad (i,j) \mapsto a_{ij},$$

auffassen. Damit haben wir für diese Matrizen Operationen zur Verfügung, die wir bei Abbildungen von Mengen in Moduln definiert haben (vgl. 8.2).

Bezeichnen wir mit $R^{(m,n)}$ die Menge der (m,n)-Matrizen bei gegebenen $n, m \in \mathbb{N}$. Für (a_{ij}), $(b_{ij}) \in R^{(m,n)}$ und $r \in R$ haben wir dann

$$\begin{aligned} \text{Matrizenaddition}: \quad (a_{ij}) + (b_{ij}) &= (a_{ij} + b_{ij}) \\ \text{Skalarmultiplikation}: \quad r(a_{ij}) &= (ra_{ij}). \end{aligned}$$

Das neutrale Element bezüglich der Addition ist dabei die (m,n)-Matrix, die aus lauter Nullen besteht, die sogenannte *Nullmatrix*.

12.2 Satz
Für jeden Ring R und $n, m \in \mathbb{N}$ bildet $R^{(m,n)}$ einen freien R-Modul.

Beweis: Die Modulstruktur von $R^{(m,n)}$ haben wir oben angegeben. Um zu zeigen, daß es ein freier Modul ist, geben wir eine Basis (mit $m \cdot n$ Elementen) an. Dazu definieren wir (mit dem Kronecker-Symbol δ) die (m,n)-Matrizen

$$E_{ij} := (\delta_{ik} \cdot \delta_{jl}),$$

also die (m,n)-Matrizen mit 1 am Schnitt der i-ten Zeile mit der j-ten Spalte und 0 überall sonst.

Es ist leicht zu sehen, daß dies $m \cdot n$ linear unabhängige Elemente in $R^{(m,n)}$ sind. Sie bilden auch ein Erzeugendensystem, da für jede Matrix $(a_{ij}) \in R^{(m,n)}$ gilt

$$(a_{ij}) = \sum_{i=1}^{m}(\sum_{j=1}^{n} a_{ij}E_{ij}).$$

Man bezeichnet E_{ij} als die *kanonische Basis* von $R^{(m,n)}$. □

Spezialfälle
Eine $(1,n)$-Matrix heißt *Zeilenvektor* $(\in R^n)$,
eine $(m,1)$-Matrix heißt *Spaltenvektor* $(\in R^m)$.

Die am Beginn des Paragraphen betrachtete Zuordnung einer Matrix zu einem Homomorphismus können wir nun genauer beschreiben.

12.3 Satz
Seien R ein Ring, M ein R-Modul mit Basis $X = (x_1, \ldots, x_m)$ und N ein R-Modul mit Basis $Y = (y_1, \ldots, y_n)$.

(1) *Die durch 12.1 gegebene Zuordnung*

$$\mathrm{Mat}_{XY} : \mathrm{Hom}_R(M,N) \to R^{(m,n)}, \quad f \mapsto (a_{ij}),$$

ist ein Gruppenisomorphismus.

(2) *Ist R ein kommutativer Ring, so ist Mat_{XY} sogar ein Isomorphismus von R-Moduln.*

Beweis: Man beachte, daß wir für $\mathrm{Hom}_R(M,N)$ nur dann eine R-Modulstruktur haben, wenn R kommutativ ist (vgl. 10.2).

(1) Aus den Anmerkungen zu 12.1 ergibt sich, daß Mat_{XY} eine injektive Abbildung ist.

Sei nun irgendeine Matrix $(a_{ij}) \in R^{(m,n)}$ gegeben. Dann stellt die rechte Seite von 12.1 n Elemente b_i aus N dar. Da $x_1, \ldots, x_m$ Basis von M ist, gibt es nach Satz 10.5 genau eine lineare Abbildung $f' : M \to N$ mit

$$f'(x_i) = b_i = a_{i1}y_1 + \ldots + a_{in}y_n.$$

Nach Konstruktion gilt $\mathrm{Mat}_{XY}(f') = (a_{ij})$. Also ist Mat_{XY} surjektiv.

Für $f, g \in \mathrm{Hom}(M,N)$ seien $\mathrm{Mat}_{XY}(f) := (a_{ij})$, $\mathrm{Mat}_{XY}(g) := (b_{ij})$, also

$$f(x_i) = \sum_{j=1}^{n} a_{ij}y_j \quad \text{und} \quad g(x_i) = \sum_{j=1}^{n} b_{ij}y_j.$$

Dann gilt $(f+g)(x_i) = f(x_i) + g(x_i) = \sum_{j=1}^{n}(a_{ij}+b_{ij})y_j$, und damit

$$\mathrm{Mat}_{XY}(f+g) = \mathrm{Mat}_{XY}(f) + \mathrm{Mat}_{XY}(g).$$

(2) Bei kommutativem R ist für $f \in \mathrm{Hom}_R(M,N)$ und $r \in R$ auch rf ein R-Homomorphismus und – wie leicht zu bestätigen –

$$\mathrm{Mat}_{XY}(rf) = r\mathrm{Mat}_{XY}(f).$$

Somit ist Mat_{XY} ein R-Isomorphismus. □

Wir hatten die Matrizen $E_{ij} = (\delta_{ik}\delta_{jl})$ als Basis von $R^{(m,n)}$ kennengelernt. Welchen Abbildungen $M \to N$ entsprechen diese – bei vorgegebenen Basen – unter dem Isomorphismus Mat? Bei kommutativem R werden die Bilder davon eine Basis von $\mathrm{Hom}_R(M,N)$ ergeben.

Nach 12.1 haben wir dafür die Zuordnung

$$x_k \longmapsto \sum_{l=1}^{n} \delta_{ik}\delta_{jl}\, y_l = \begin{cases} y_j & \text{für } k = i \\ 0 & \text{für } k \neq i \end{cases}.$$

Dabei wird also dem i-ten Basisvektor von M der j-te Basisvektor von N und den anderen Basisvektoren von M die 0 zugeordnet.

Bezeichnen wir diese Abbildungen, die von der Wahl der Basen abhängig sind, mit $\widehat{E}_{ij}$, also

$$\widehat{E}_{ij} : M \to N, \quad x_k \mapsto \begin{cases} y_j & \text{für } k = i \\ 0 & \text{für } k \neq i \end{cases},$$

so folgt aus Satz 12.3:

12.4 Korollar
Sei R ein kommutativer Ring. Sind M und N freie R-Moduln mit Basen $\{x_1,\ldots,x_m\} \subset M$ und $\{y_1,\ldots,y_n\} \subset N$, dann ist $\mathrm{Hom}_R(M,N)$ ein freier R-Modul mit Basis $\widehat{E}_{ij}$, $i \leq m$, $j \leq n$.

Nachdem wir die Beziehung zwischen Homomorphismen und Matrizen kennengelernt haben, stellt sich die Frage, ob auch die Komposition von Abbildungen eine entsprechende Bildung bei Matrizen erlaubt. Dies führt zu einer *Multiplikation* von Matrizen. Sehen wir uns dazu die Matrizen von zusammengesetzten Abbildungen an:

Seien M, N und P freie R-Moduln mit den Basen

$$X = (x_1,\ldots,x_m),\ Y = (y_1,\ldots,y_n) \text{ und } Z = (z_1,\ldots,z_p).$$

Betrachten wir Homomorphismen $f : M \to N$ und $g : N \to P$. Dann ist auch $g \circ f : M \to P$ ein Homomorphismus, und f, g und $g \circ f$ bestimmen nach 12.3 drei Matrizen

$$\begin{array}{rcll} \mathrm{Mat}_{XY}(f) & =: & (a_{ij}) & \in \; R^{(m,n)} \\ \mathrm{Mat}_{YZ}(g) & =: & (b_{ij}) & \in \; R^{(n,p)} \\ \mathrm{Mat}_{XZ}(g \circ f) & =: & (c_{ij}) & \in \; R^{(m,p)} \end{array}$$

Aus den Definitionen erhalten wir

$$\begin{aligned} g \circ f(x_i) &= g(\sum_{k=1}^{n} a_{ik}\, y_k) \;=\; \sum_{k=1}^{n} a_{ik}\, g(y_k) \\ &= \sum_{k=1}^{n} a_{ik}(\sum_{j=1}^{p} b_{kj}\, z_j) \;=\; \sum_{j=1}^{p}(\sum_{k=1}^{n} a_{ik} b_{kj}) z_j \\ &= \sum_{j=1}^{p} c_{ij}\, z_j. \end{aligned}$$

Somit gilt $c_{ij} = \sum\limits_{k=1}^{n} a_{ik} b_{kj}$, und wir nehmen dies zur Definition von

12.5 Produkt von Matrizen
Seien R ein Ring und $A = (a_{ij}) \in R^{(m,n)}$, $B = (b_{ij}) \in R^{(n,p)}$ Matrizen über R. Als *Produkt* von A und B bezeichnet man die Matrix

$$A \cdot B := C = (c_{ij}) \in R^{(m,p)} \;\text{ mit }\; c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Man beachte, daß dieses Produkt nur gebildet werden kann, wenn
Spaltenzahl von A = Zeilenzahl von B.

Merkregel:
Für die Berechnung des Eintrags c_{ij} im Produkt muß man die j-te Spalte von B über die i-te Zeile von A legen, die zusammentreffenden Skalare multiplizieren und dann die Summe der Produkte bilden:

$$i \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \begin{matrix} \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \qquad\quad j \\ \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \dots & b_{nj} & \dots & b_{np} \end{pmatrix} \end{matrix} = \begin{matrix} \qquad j \\ i \begin{pmatrix} c_{11} & \dots & c_{1p} \\ \vdots & c_{ij} & \vdots \\ c_{m1} & \dots & c_{mp} \end{pmatrix} \end{matrix}$$

Nach Definition dieser Produktbildung gilt die wichtige Beziehung:

12.6 Satz
Seien R ein Ring und M, N, P drei R-Moduln mit endlichen Basen X, Y und Z. Dann gilt

$$\mathrm{Mat}_{XZ}(g \circ f) = \mathrm{Mat}_{XY}(f) \cdot \mathrm{Mat}_{YZ}(g).$$

Aufgrund dieses Zusammenhangs lassen sich sofort folgende Rechenregeln für die Produktbildung von Matrizen angeben, die man natürlich auch direkt bestätigen kann:

12.7 Eigenschaften der Multiplikation
Seien R ein Ring und A, A_1, $A_2 \in R^{(m,n)}$, B, B_1, $B_2 \in R^{(n,p)}$, $C \in R^{(p,q)}$. Dann gilt:

(1) $(rA)B = r(AB)$ *für alle* $r \in R$.
Ist R kommutativ, so gilt auch $A(rB) = r(AB)$ *für alle* $r \in R$.

(2) $A(B_1 + B_2) = AB_1 + AB_2$.

(3) $(A_1 + A_2)B = A_1B + A_2B$.

(4) $(AB)C = A(BC)$.

Von besonderem Interesse ist die Multiplikation mit den in 12.2 eingeführten Matrizen E_{ij}. Wir stellen daher einige dieser Beziehungen zusammen, die sich leicht bestätigen lassen:

12.8 Multiplikation mit E_{ij}
Seien R ein Ring und $A = (a_{ij}), E_{ij} \in R^{(n,n)}$.

(1) $E_{ij}E_{rs} = \begin{cases} E_{is} & \text{falls } j = r \\ 0 & \text{falls } j \neq r. \end{cases}$

(2) $E_{rk}A = E_{rk}(\sum_{i,j} a_{ij}E_{ij}) = \sum_j a_{kj}E_{rj}$;
$AE_{rk} = (\sum_{i,j} a_{ij}E_{ij})E_{rk} = \sum_i a_{ir}E_{ik}$.

(3) $E_{rk}AE_{ls} = E_{rk}(\sum_{i,j} a_{ij}E_{ij})E_{ls} = a_{kl}E_{rs}$.

(4) $AB = BA$ *gilt genau dann für alle* $B \in R^{(n,n)}$, *wenn* $A = rE_n$ *für ein* $r \in Z(R)$ *(Zentrum von R) und E_n die Einheitsmatrix in* $R^{(n,n)}$ *ist.*

Beweis: Die ersten Aussagen folgen aus der Matrizenmultiplikation.
(4) ergibt sich aus dem Vergleich der beiden Ausdrücke in (2). □

Das Vertauschen von Zeilen und Spalten einer Matrix ist eine wichtige Operation, die wir formal festlegen wollen:

12.9 Definition
Sei R ein Ring und $A = (a_{ij}) \in R^{(m,n)}$ eine Matrix. Als *Transponierte* von A bezeichnet man die Matrix

$$A^t := (a_{ji}) \in R^{(n,m)}.$$

Eine Matrix heißt *symmetrisch*, wenn $A = A^t$ gilt.

Bei quadratischen Matrizen bewirkt das Transponieren gerade eine Spiegelung an der Hauptdiagonalen.

Halten wir einige allgemeine Eigenschaften dieser Bildung fest.

12.10 Transponieren von Matrizen. *Sei R ein Ring.*

(1) *Für jede Matrix A über R gilt $(A^t)^t = A$.*

(2) *Für Matrizen $A, B \in R^{(m,n)}$ und $r \in R$ gilt $(A + B)^t = A^t + B^t$ und $(rA)^t = rA^t$.*

(3) *$R^{(m,n)} \to R^{(n,m)}, A \mapsto A^t$, ist ein Isomorphismus von R-Moduln.*

(4) *Ist R kommutativ, $A \in R^{(m,n)}$ und $C \in R^{(n,k)}$, so gilt $(AC)^t = C^t A^t$.*

(5) *Ist R kommutativ, so ist $R^{(m,m)} \to R^{(m,m)}, A \mapsto A^t$, ein Antiautomorphismus, und für invertierbare A gilt $(A^{-1})^t = (A^t)^{-1}$.*

Beweis: (1) Dies folgt unmittelbar aus der Definition.

(2) Diese Aussagen lassen sich leicht bestätigen.

(3) Die Behauptung folgt aus (1). Dabei können die Matrizen sowohl als Linksmoduln als auch als Rechtsmoduln über R aufgefaßt werden.

(4) Auch dies kann man rechnerisch nachprüfen. Die angegebene Beziehung gilt nicht mehr über nicht-kommutativen Ringen. (In diesem Fall müßte man auf der rechten Seite die Multiplikation in R durch die *opposite* Multiplikation $a * b := ba$ ersetzen.)

(5) Dies ergibt sich aus (3) und (4). □

Betrachten wir nun $\mathrm{Hom}_R(M, N)$ für den Fall $M = R^m$ und $N = R^n$.

Fassen wir den R^m als Zeilenraum $R^{(1,m)}$ auf, d.h. wir behandeln die Elemente $(r_1, \ldots, r_m) \in R^m$ als $(1, m)$-Matrizen. Diese kann man dann von rechts mit (m, n)-Matrizen multiplizieren. Für $A = (a_{ij}) \in R^{(m,n)}$ bekommen wir damit die Zuordnung

$$\begin{array}{rccl} \widehat{A}: & R^{(1,m)} & \to & R^{(1,n)}, \\ & (r_1, \ldots, r_m) & \mapsto & (r_1, \ldots, r_m)(a_{ij}) = (\sum\limits_{i=1}^{m} r_i a_{i1}, \ldots, \sum\limits_{i=1}^{m} r_i a_{in}). \end{array}$$

Aus den Eigenschaften der Matrizenmultiplikation folgt, daß dies eine lineare Abbildung von R-Linksmoduln ist, also $\widehat{A} \in \mathrm{Hom}_R(R^m, R^n)$.

Für die kanonischen Basisvektoren

$$e_i = (\delta_{il}) = (0, \ldots, 1_i, \ldots, 0) \in R^{(1,m)},$$
$$e'_j = (\delta_{jl}) = (0, \ldots, 1_j, \ldots, 0) \in R^{(1,n)},$$

bedeutet dies die Zuordnung (für alle $i \leq m$)

$$e_i \mapsto \sum_{j=1}^{n} a_{ij} e'_j = (a_{i1}, \ldots, a_{in}).$$

Somit spannen die Zeilenvektoren von A das Bild des Homomorphismus' $\widehat{A}$ auf.

Bezüglich der kanonischen Basen in R^m, R^n erhält man $\mathrm{Mat}(\widehat{A}) = A$.

Andererseits ergibt sich für $f \in \mathrm{Hom}_R(R^m, R^n)$ mit dieser Notation $\widehat{\mathrm{Mat}(f)} = f$. Also ist

$$R^{(m,n)} \to \mathrm{Hom}_R(R^m, R^n), \quad A \mapsto \widehat{A},$$

die inverse Abbildung zu $\mathrm{Hom}_R(R^m, R^n) \to R^{(m,n)}$ aus 12.3.

Fassen wir nun den R^m als Spaltenraum $R^{(m,1)}$ auf. Dann kann man seine Elemente von links mit (n, m)-Matrizen multiplizieren, und für $B = (b_{ij}) \in R^{(n,m)}$ bekommen wir die lineare Abbildung von R-Rechtsmoduln

$$\begin{array}{rccl} \widehat{B}: & R^{(m,1)} & \to & R^{(n,1)}, \\ & \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} & \mapsto & (b_{ij}) \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} = \begin{pmatrix} \sum_{l=1}^{m} b_{1l} r_l \\ \vdots \\ \sum_{l=1}^{m} b_{nl} r_l \end{pmatrix}. \end{array}$$

Stellen wir diese Sachverhalte zusammen:

12.11 Satz

(1) Sei $A = (a_{ij}) \in R^{(m,n)}$. Dann gilt für die lineare Abbildung von R-Linksmoduln

$$\widehat{A} : R^{(1,m)} \to R^{(1,n)}, \quad (r_1, \ldots, r_m) \mapsto (r_1, \ldots, r_m) \cdot (a_{ij}),$$

(i) *Bild des i-ten Basisvektors von R^m unter $\widehat{A}$ ist die i-te Zeile von A.*

(ii) *Bild $\widehat{A}$ wird von den Zeilen von A erzeugt.*

(2) Sei $B = (b_{ij}) \in R^{(n,m)}$. Dann gilt für die lineare Abbildung von R-Rechtsmoduln

$$\widehat{B} : R^{(m,1)} \to R^{(n,1)}, \quad \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} \mapsto (b_{ij}) \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} :$$

(i) *Bild des i-ten Basisvektors von R^m unter $\widehat{B}$ ist die i-te Spalte von B.*

(ii) $\operatorname{Bild}\widehat{B}$ *wird von den Spalten von B erzeugt.*

(3) Für $f \in \operatorname{Hom}_R(R^{(1,m)}, R^{(1,n)})$ und $g \in \operatorname{Hom}_R(R^{(m,1)}, R^{(n,1)})$ gilt

$$\begin{aligned} f(r_1,\dots,r_m) &= (r_1,\dots,r_m)\cdot \operatorname{Mat}(f), \text{ und} \\ g\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} &= \operatorname{Mat}(g)\cdot \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}, \end{aligned}$$

wobei Mat *jeweils bezüglich der kanonischen Basis zu bilden ist.*

Man beachte, daß in unserem Ansatz die Matrizenmultiplikation bei Zeilenräumen Homomorphismen von Linksmoduln, bei Spaltenräumen aber Homomorphismen von Rechtsmoduln über R bewirken. Bei kommutativen Ringen brauchen wir die Seiten nicht zu unterscheiden:

12.12 Korollar
Sei R ein kommutativer Ring und $A = (a_{ij}) \in R^{(m,n)}$. Dann haben wir (mit den oben betrachteten Abbildungen) das kommutative Diagramm

$$\begin{array}{ccccc} R^{(1,m)} & \xrightarrow{\widehat{A}} & R^{(1,n)}, & (r_1,\dots,r_m) & \mapsto & (r_1,\dots,r_m)\cdot A, \\ \downarrow & & \downarrow & \downarrow & & \downarrow \\ R^{(m,1)} & \xrightarrow{\widehat{A^t}} & R^{(n,1)}, & \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix} & \mapsto & A^t\cdot\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}, \end{array}$$

in dem die senkrechten Isomorphismen durch Transponieren gegeben sind.

Die Bedeutung der oben betrachteten Situation liegt darin, daß sich jeder Homomorphismus zwischen freien Moduln damit in Beziehung bringen läßt.

12.13 Koordinatenisomorphismus
Sei M ein freier R-Modul mit Basis $X = (x_1,\dots,x_m)$. Für die Standardbasis $e_1,\dots,e_m$ von R^m bezeichnen wir den durch

$$\kappa_X : M \to R^m, \quad x_i \mapsto e_i,$$

festgelegten Isomorphismus als *Koordinatenisomorphismus*. Dabei ist es zunächst gleich, ob wir R^m als Zeilen- oder Spaltenraum betrachten.

Sei N ein freier R-Modul mit Basis $Y = (y_1, \ldots, y_n)$ und entsprechendem Koordinatenisomorphismus

$$\kappa_Y : N \to R^n,\ y_i \mapsto e_i.$$

Zu jedem Homomorphismus $f : M \to N$ haben wir das kommutative Diagramm

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ \kappa_X \downarrow & & \downarrow \kappa_Y \\ R^m & \xrightarrow{\widetilde{f}} & R^n \end{array} \qquad \text{mit } \widetilde{f} = \kappa_Y \circ f \circ \kappa_X^{-1}.$$

Aus Satz 12.6 läßt sich daraus ableiten:

12.14 Satz
In obiger Situation gilt

$$\mathrm{Mat}_{XY}(f) = \mathrm{Mat}(\widetilde{f}),$$

wobei die rechte Seite bezüglich der kanonischen Basen zu bilden ist.

Beweis: Mit $\mathrm{Mat}_{XY}(f) =: (a_{ij})$ haben wir

$$\begin{array}{rclcrcl} \widetilde{f}(e_i) & = & \kappa_Y \circ f \circ \kappa_X^{-1}(e_i) & = & \kappa_Y(f(x_i)) & & \\ & = & \kappa_Y(\sum\limits_{j=1}^{n} a_{ij} y_j) & = & \sum\limits_{j=1}^{n} a_{ij}\kappa_Y(y_j) & = & \sum\limits_{j=1}^{n} a_{ij} e_j \end{array}.$$

Daraus ersieht man $\mathrm{Mat}(\widetilde{f}) = \mathrm{Mat}_{XY}(f)$. □

12.15 Matrizenringe

(1) Sei R ein Ring. Für jedes $m \in \mathbb{N}$ bilden die quadratischen Matrizen $R^{(m,m)}$ mit der Matrizenaddition und Matrizenmultiplikation einen Ring mit Einselement (Einheitsmatrix)

$$E_m = (\delta_{ij}) = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

(2) Sei M ein freier R-Modul mit Basis $X = \{x_1, \ldots, x_m\}$.

(i) *Die in 12.3 betrachtete Abbildung*

$$\mathrm{Mat}_{XX} : \mathrm{End}_R(M) \to R^{(m,m)}$$

ist ein Ringisomorphismus mit $\mathrm{Mat}_{X,X}(\mathrm{id}_M) = E_m$.

(ii) *$f \in \mathrm{End}_R(M)$ ist genau dann ein Isomorphismus, wenn $\mathrm{Mat}_{XX}(f)$ eine invertierbare Matrix ist.*

Beweis: (1) Aus den in 12.7 festgehaltenen Eigenschaften der Matrizenmultiplikation folgt, daß $R^{(m,m)}$ einen Ring bildet.

(2) (i) Nach 12.3 ist die betrachtete Abbildung

$$\mathrm{Mat}_{XX} : \mathrm{End}_R(M) \to R^{(m,m)}$$

ein Isomorphismus von additiven Gruppen.

Nach 12.6 gilt für $f, g \in \mathrm{End}_R(M)$ auch $\mathrm{Mat}(g \circ f) = \mathrm{Mat}(g) \cdot \mathrm{Mat}(f)$. Somit ist Mat ein Ringisomorphismus.

(ii) Sei $f : M \to M$ invertierbar. Dann gibt es $g : M \to M$ mit $g \circ f = \mathrm{id}_M$ und $f \circ g = \mathrm{id}_M$. Setzen wir $\mathrm{Mat}_{XX} = \mathrm{Mat}$, so folgt damit

$$\mathrm{Mat}(g \circ f) = E_m = \mathrm{Mat}(g) \cdot \mathrm{Mat}(f), \quad \mathrm{Mat}(f \circ g) = E_m = \mathrm{Mat}(f) \cdot \mathrm{Mat}(g).$$

□

Bezüglich einer festen Basis X eines freien Moduls M wurde der Identität durch Mat_{XX} die m-te Einheitsmatrix zugeordnet, $\mathrm{Mat}_{XX}(\mathrm{id}_M) = E_m$.

Hat man verschiedene Basen $X = (x_1, \ldots, x_m)$ und $X' = (x'_1, \ldots, x'_m)$ in M, so wird der Identität durch $\mathrm{Mat}_{XX'}$ nicht die Einheitsmatrix zugeordnet, sondern die Matrix (t_{ij}) mit

$$\begin{array}{ccccc} x_1 & = & \mathrm{id}_M(x_1) & = & t_{11}x'_1 + \ldots + t_{1n}x'_n \\ \vdots & & \vdots & & \vdots \\ x_m & = & \mathrm{id}_M(x_m) & = & t_{m1}x'_1 + \ldots + t_{mn}x'_n \end{array}$$

Nach 12.6 erhalten wir damit (wegen $\mathrm{id}_M \circ \mathrm{id}_M = \mathrm{id}_M$)

$$E_m = \mathrm{Mat}_{X'X'}(\mathrm{id}_M) = \mathrm{Mat}_{X'X}(\mathrm{id}_M) \cdot \mathrm{Mat}_{XX'}(\mathrm{id}_M).$$

Daraus folgt, daß die Matrix $\mathrm{Mat}_{XX'}(\mathrm{id}) = (t_{ij})$ invertierbar ist.

Für $f : M \to N$ erhält man nun für verschiedene Basen in M bzw. N folgenden Zusammenhang:

12.16 Satz (Basiswechsel)
Über einem Ring R seien
M ein freier Modul mit Basen $X = (x_1, \ldots, x_m)$ und $X' = (x'_1, \ldots, x'_m)$ und
N ein freier Modul mit Basen $Y = (y_1, \ldots, y_n)$ und $Y' = (y'_1, \ldots, y'_n)$.

(1) *$T = \mathrm{Mat}_{XX'}(\mathrm{id}_M) \in R^{(m,m)}$ ist invertierbar mit $T^{-1} = \mathrm{Mat}_{X'X}(\mathrm{id}_M)$.*
$S = \mathrm{Mat}_{YY'}(\mathrm{id}_N) \in R^{(n,n)}$ ist invertierbar mit $S^{-1} = \mathrm{Mat}_{Y'Y}(\mathrm{id}_N)$.

(2) *Für einen Homomorphismus $f : M \to N$ gilt:*

$$\mathrm{Mat}_{XY}(f) = T \cdot \mathrm{Mat}_{X'Y'}(f) \cdot S^{-1}.$$

Beweis: (1) Dies wurde in unserer Vorbetrachtung ausgeführt.

(2) Die Behauptung ergibt sich mit 12.6 aus dem kommutativen Diagramm

$$\begin{array}{ccc} {}_X M & \xrightarrow{f} & N_Y \\ \mathrm{id}\ \downarrow & & \downarrow\ \mathrm{id}\,. \\ {}_{X'} M & \longrightarrow & N_{Y'} \end{array}$$

□

Matrizen, die bezüglich verschiedener Basen den gleichen Homomorphismus darstellen, faßt man in Klassen zusammen. Bei Homomorphismen zwischen unterschiedlichen Moduln kann man die Basis in Quelle und Ziel ändern. Die Matrix eines Endomorphismus $f : M \to M$ hat man gerne bezüglich der gleichen Basis in M als Quelle und Ziel. Damit haben wir folgende Einteilung der Matrizen:

12.17 Definition
Zwei Matrizen $A, B \in R^{(m,n)}$ nennt man *äquivalent*, wenn es invertierbare Matrizen $T \in R^{(m,m)}$, $S \in R^{(n,n)}$ gibt mit

$$A = T \cdot B \cdot S^{-1}.$$

Man sagt, $A, B \in R^{(m,m)}$ sind *ähnlich* (oder *konjugiert*), wenn es eine invertierbare Matrix $T \in R^{(m,m)}$ gibt mit

$$A = T \cdot B \cdot T^{-1}.$$

Es ist eine leichte Übung, zu zeigen, daß beide Beziehungen Äquivalenzrelationen auf den entsprechenden Klassen von Matrizen bestimmen.

Da die invertierbaren Matrizen Basiswechsel in entsprechenden freien Moduln bewirken, ergibt sich aus 12.16:

12.18 Folgerung
Seien M und N freie R-Moduln mit Basen der Länge m bzw. n.

(1) *Zwei Matrizen $A, B \in R^{(m,n)}$ sind genau dann* äquivalent, *wenn es eine lineare Abbildung $f : M \to N$ und Basen X, X' von M und Y, Y' von N gibt mit*

$$A = \mathrm{Mat}_{XY}(f) \quad \textit{und} \quad B = \mathrm{Mat}_{X'Y'}(f).$$

(2) *Zwei Matrizen $A, B \in R^{(m,m)}$ sind genau dann* ähnlich, *wenn es ein $f \in \mathrm{End}_R(M)$ und Basen X, X' von M gibt mit*

$$A = \mathrm{Mat}_{XX}(f) \quad \textit{und} \quad B = \mathrm{Mat}_{X'X'}(f).$$

12.19 Aufgaben

(1) R sei ein kommutativer Ring mit Eins,
$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in R^{(2,2)}$ *und* $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in R^{(2,2)}$.

Man berechne $A^2 - (a+d)A + (ad-bc)E$.

(2) Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{Z}^{(2,3)}$ *und* $B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ -1 & 0 \end{pmatrix} \in \mathbb{Z}^{(3,2)}$.

Man berechne die Matrizenprodukte $A \cdot B$, $B \cdot A$, $A \cdot B \cdot A$ *und* $B \cdot A \cdot B$.

(3) Seien $f : \mathbb{Q}^2 \to \mathbb{Q}^3$ *und* $g : \mathbb{Q}^3 \to \mathbb{Q}^2$ *gegeben durch*

$$\begin{aligned} f(x,y) &= (x+2y, x-y, 2x+y) \\ g(x,y,z) &= (x-2y+3z, 2y-3z). \end{aligned}$$

(i) *Man berechne* $\mathrm{Mat}(f)$, $\mathrm{Mat}(g)$, $\mathrm{Mat}(g \circ f)$ *und* $\mathrm{Mat}(f \circ g)$ *bezüglich der kanonischen Basen* $E_2 \subset \mathbb{Q}^2$ *und* $E_3 \subset \mathbb{Q}^3$.

(ii) *Man zeige, daß* $g \circ f$ *ein Isomorphismus ist, und bestimme* $\mathrm{Mat}((g \circ f)^{-1})$ *bezüglich* E_2.

(4) $B = ((1,1,0),(1,-1,1),(1,1,1)) \subset \mathbb{Q}^3$ *und* $C = ((1,1),(1,-1)) \subset \mathbb{Q}^2$ *sind Basen in* $\mathbb{Q}^3$ *bzw.* $\mathbb{Q}^2$.

(i) *Man bestimme die Matrizen* $\mathrm{Mat}_{E_3,B}(\mathrm{id}_{\mathbb{Q}^3})$, $\mathrm{Mat}_{B,E_3}(\mathrm{id}_{\mathbb{Q}^3})$, $\mathrm{Mat}_{E_2,C}(\mathrm{id}_{\mathbb{Q}^2})$ *und* $\mathrm{Mat}_{C,E_2}(\mathrm{id}_{\mathbb{Q}^2})$.

(ii) *Für die Homomorphismen aus der vorigen Aufgabe berechne man die Matrizen* $\mathrm{Mat}_{C,B}(f)$, $\mathrm{Mat}_{B,C}(g)$, $\mathrm{Mat}_{C,C}(g \circ f)$ *und* $\mathrm{Mat}_{B,B}(f \circ g)$.

13 Matrizen über Divisionsringen

Wie wir in Satz 11.6 gesehen haben, lassen sich Eigenschaften von Homomorphismen von Vektorräumen mit Hilfe des Dimensionsbegriffs beschreiben. Da andererseits diese Homomorphismen mittels Matrizen dargestellt werden können, ist zu erwarten, daß auch Matrizen über Divisionsringen zusätzliche Eigenschaften haben. Solche wollen wir in diesem Abschnitt behandeln.

Setzen wir also voraus, daß K ein Divisionsring ist.

In 12.11 wurde gezeigt, daß das Bild eines Homomorphismus $f : K^m \to K^n$ von den Zeilen(vektoren) einer Matrix erzeugt wird. Nun hat $\operatorname{Bild} f$ eine Dimension (den *Rang* von f), die man somit aus den Zeilen der zugehörigen Matrix ablesen kann.

13.1 Definition
Die Dimension des von den Zeilen von $A \in K^{(m,n)}$ erzeugten Unterraums von K^n heißt der *Zeilenrang* von A.

Die Dimension des von den Spalten von A erzeugten Unterraums von K^m nennt man den *Spaltenrang* von A.

Unserer Gepflogenheit folgend, fassen wir dabei sowohl den Zeilenraum als auch den Spaltenraum als Linksvektorräume auf. Man kann beide auch als Rechtsvektorräume über K betrachten und muß bei nicht-kommutativem K diese Fälle unterscheiden.

Man sieht sofort, daß

$$\begin{aligned} \text{Zeilenrang } A &\leq \min(m,n), \\ \text{Spaltenrang } A &\leq \min(m,n). \end{aligned}$$

Man sagt, A hat *maximalen* Zeilenrang (Spaltenrang), wenn Zeilenrang $A = \min(m,n)$ (Spaltenrang $A = \min(m,n)$).

Im allgemeinen können Zeilen- und Spaltenrang verschieden sein (vgl. 13.9, Aufgabe 10). Wir werden später sehen, daß für Matrizen über *kommutativen* Divisionsringen Zeilenrang = Spaltenrang gilt. Aus 12.11 und 12.14 ergibt sich:

13.2 Satz
Sei $f : V \to W$ ein Homomorphismus endlich-dimensionaler K-Vektorräume. Für beliebige Basen $X \subset V$, $Y \subset W$ und $A := \operatorname{Mat}_{XY}(f)$ gilt:

(1) *Zeilenrang $A = \operatorname{Rang} f$.*

(2) *f ist genau dann injektiv, wenn Zeilenrang $A = \dim V$.*

(3) *f ist genau dann surjektiv, wenn Zeilenrang $A = \dim W$.*

(4) *f ist genau dann bijektiv, wenn Zeilenrang $A = \dim W = \dim V$.*

Beweis: (1) Setze $\dim V = m$, $\dim W = n$. Mit den Bezeichnungen von 12.14 hat man $\operatorname{Rang} f = \operatorname{Rang} \widetilde{f}$ mit $\widetilde{f} : K^m \to K^n$. Nach 12.11 wird $\operatorname{Bild} \widetilde{f}$ von den Zeilen von $\operatorname{Mat}(\widetilde{f}) = \operatorname{Mat}_{XY}(f)$ erzeugt, also

$$\operatorname{Rang} f = \dim \operatorname{Bild} \widetilde{f} = \text{Zeilenrang } A.$$

(2)–(4) In 11.6 haben wir die Bedeutung von $\operatorname{Rang} f$ für den Homomorphismus f kennengelernt. Damit folgen die Behauptungen aus (1). □

Für quadratische Matrizen haben wir damit folgende Kennzeichnung der Invertierbarkeit:

13.3 Satz
Sei V ein m-dimensionaler K-Vektorraum mit Basis X. Für $f \in \operatorname{End}_K(V)$ mit $A := \operatorname{Mat}_{XX}(f) \in K^{(m,m)}$ sind folgende Aussagen äquivalent:

(a) *f ist Isomorphismus (injektiv, surjektiv, vgl. 11.7);*

(b) *A ist invertierbare Matrix;*

(c) *es gibt eine Matrix $B \in K^{(m,m)}$ mit $AB = E$;*

(d) *es gibt eine Matrix $B \in K^{(m,m)}$ mit $BA = E$;*

(e) *Zeilenrang $A = m$.*

Beweis: (a)⇔(b) $\operatorname{End}(V) \simeq K^{(m,m)}$, und Isomorphismen führen invertierbare Elemente wieder in invertierbare über (vgl. 12.15).

(b)⇒(c),(d) Diese Implikationen sind klar.

(c)⇒(e) Aus $AB = E$ folgt, daß f injektiv ist, also $\operatorname{Rang} f =$ Zeilenrang $A = m$.

(d)⇒(e) Aus $BA = E$ folgt, daß f surjektiv ist, also $\operatorname{Rang} f =$ Zeilenrang $A = m$.

(e)⇒(a) Nach Korollar 11.7 folgt aus $\operatorname{Rang} f =$ Zeilenrang $A = m$, daß f Isomorphismus ist. □

Bemerkung: Die Äquivalenzen von (a) bis (d) in 12.1 gelten auch für endlich erzeugte freie Moduln über kommutativen Ringen (mit anderem Beweis).

Wir wollen nun den Rang einer Matrix $A \in K^{(m,n)}$ ermitteln, also die Dimension des von den Zeilen von A erzeugten Unterraums von $K^{(1,m)}$.

Diese ändert sich nicht, wenn man etwa statt der Zeilen a_1, a_2, a_3 die Zeilen a_1, $a_2 + a_1$, a_3 betrachtet. Wir stellen daher fest:

Der Rang einer Matrix ändert sich nicht, wenn man

(1) zu einer Zeile eine andere addiert,

(2) eine Zeile mit einem Skalar $0 \neq r \in K$ (von links) multipliziert.

Daraus lassen sich weitere *erlaubte* Zeilenoperationen ableiten, also Umformungen, die den Zeilenrang der Matrix nicht verändern. Man nennt diese auch *elementare* Zeilenumformungen:

(3) Addition des r-fachen eines Zeilenvektors zu einem anderen,

(4) Vertauschen von Zeilenvektoren.

(1) und (2) lassen sich durch Multiplikation mit Matrizen ausdrücken. Dies geschieht mit den im Beweis von 12.2 eingeführten Matrizen

$$E_{ij} := (\delta_{ik} \cdot \delta_{jl}) \in K^{(m,m)}.$$

Mit der Einheitsmatrix $E_m \in K^{(m,m)}$ bewirkt die Multiplikation von A von links

- mit $(E_m + E_{ij})$ die Addition der j-ten Zeile zur i-ten Zeile und
- mit $(E_m + (r-1)E_{ii})$ die Multiplikation der i-ten Zeile mit $r \neq 0$.

Da diese Operationen umkehrbar sind, sind auch die zugehörigen Matrizen (*Elementarmatrizen*) invertierbar.

13.4 Satz
Jede Matrix $A \in K^{(m,n)}$ läßt sich durch elementare Zeilenoperationen in eine (m,n)-Matrix der Form

$$\begin{pmatrix}
0 & \dots & 0 & c_{1,r_1} & * & \dots & \dots & * \\
0 & \dots & \dots & 0 & c_{2,r_2} & * & \dots & * \\
\vdots & & & & \ddots & & & \vdots \\
0 & \dots & \dots & \dots & 0 & c_{k,r_k} & * \dots & * \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
\vdots & & & & & & & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & 0
\end{pmatrix}$$

überführen (Zeilenstufenform), wobei

$$1 \leq r_1 < r_2 < \ldots < r_k \leq m,\ k = \operatorname{Rang} A,\ \textit{und}$$
$$c_{1,r_1} \neq 0,\ c_{2,r_2} \neq 0, \ldots, c_{k,r_k} \neq 0.$$

Man kann immer $c_{i,r_i} = 1$ für $i = 1, \ldots, k$ erreichen.

Die ersten k Zeilen sind eine Basis des von den Zeilen von A aufgespannten Unterraums von K^n.

Beweis: Zunächst sucht man eine Zeile i mit $a_{i1} \neq 0$ oder $a_{ir_1} \neq 0$ für ein möglichst kleines $r_1 \leq n$. Diese Zeile bringt man (durch Vertauschen) nach oben.

Nun kann man durch erlaubte Zeilenoperationen erreichen, daß alle anderen Elemente in der r_1-ten Spalte Null werden.

Durch Wiederholen dieser Prozedur erreicht man die gewünschte Form. □

Hat die Matrix A den Rang $r = n < m$, so ergibt die obige Darstellung die Form

$$m \begin{pmatrix} 1 & * & \dots\dots\dots & * \\ 0 & 1 & * & \dots & * \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & * \\ 0 & \dots\dots\dots & & 0 & 1 \\ 0 & \dots\dots\dots\dots & & & 0 \\ \vdots & & & & \vdots \\ 0 & \dots\dots\dots\dots & & & 0 \end{pmatrix}$$

Es ist leicht zu sehen, daß für $m = n$ daraus durch weitere elementare Zeilenumformungen sogar die Einheitsmatrix gewonnen werden kann. Dies eröffnet eine vorteilhafte Methode zur Ermittlung der Inversen:

13.5 Satz
Sei $A \in K^{(m,m)}$ eine Matrix mit $\operatorname{Rang} A = m$. *Dann gilt:*

(1) *A kann durch elementare Zeilenumformungen zur Einheitsmatrix E umgeformt werden.*

(2) *Führt man diese Umformungen gleichzeitig mit der Einheitsmatrix E durch, so ergibt sich dabei die zu A inverse Matrix A^{-1}.*

(3) *Jede invertierbare Matrix ist Produkt von Elementarmatrizen.*

Beweis: (1) Dies wurde in der Vorbetrachtung erläutert.

(2) Nach (1) gibt es Elementarmatrizen $B_1, \dots, B_n$ mit der Eigenschaft

$$B_n \cdots B_2 \cdot B_1 \cdot A = E.$$

Also gilt $B_n \cdots B_2 \cdot B_1 = B_n \cdots B_2 \cdot B_1 \cdot E = A^{-1}$.

(3) ist eine Folge von (2). □

Beispiel
Sehen wir uns an, wie nach der oben angegebenen Methode die Inverse zur Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{Q}^{(3,3)}$$

berechnet werden kann. Wir bezeichnen die Zeilen mit römischen Ziffern.

$$\begin{array}{r|l}
 & \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \\
\text{II} \leftrightarrow \text{III} & \\
 & \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & 0 & 1 & 0 \end{array}\right) \\
\text{I} - 2 \cdot \text{II} & \\
 & \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & -2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 2 & 0 & 1 & 0 \end{array}\right) \\
\text{III} - \text{I} & \\
 & \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & -2 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 & 2 \end{array}\right) \\
\text{I} + 3 \cdot \text{III} & \\
 & \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 3 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 & 2 \end{array}\right) \\
(-1) \cdot \text{III} & \\
 & \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & 3 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 & -2 \end{array}\right)
\end{array}$$

Daraus lesen wir ab

$$A^{-1} = \begin{pmatrix} -2 & 3 & 4 \\ 0 & 0 & 1 \\ 1 & -1 & -2 \end{pmatrix}.$$

Fassen wir eine Matrix $A \in K^{(m,n)}$ als Homomorphismus $K^m \to K^n$ auf (vgl. 12.11), dann entspricht nach Satz 12.16 (Basiswechsel) die Umformung von A in Zeilenstufenform einer Basistransformation in K^m.

Führen wir auch Basistransformationen in K^n durch, so läßt sich die Form von A noch weiter vereinfachen. Wir wollen uns dies ganz allgemein überlegen.

Sei $f : V \to W$ ein Homomorphismus von K-Vektorräumen mit $\operatorname{Rang} f = r$, $\dim V = m$ und $\dim W = n$.

Die Wahl von Basen $X \subset V$ und $Y \subset W$ beeinflußt die Gestalt der Matrix $\operatorname{Mat}_{XY}(f)$. Wie kann man eine besonders handliche Gestalt dieser Matrix erreichen?

Dazu wählen wir folgende Basis X in V:

$v_{r+1}, \ldots, v_m$ sei eine Basis von $\operatorname{Kern} f$,
$v_1, \ldots, v_r$ soll dies zu einer Basis von V ergänzen.

Dann ist $w_1 := f(v_1), \ldots, w_r := f(v_r)$ eine Basis von Bild f. Diese ergänzen wir durch $w_{r+1}, \ldots, w_n$ zu einer Basis Y von W.

Dafür gilt nun $f(v_i) = \begin{cases} w_i & \text{für } i = 1, \ldots, r \\ 0 & i > r \end{cases}$, und damit

$$\mathrm{Mat}_{XY}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

wobei E_r die (r, r)-Einheitsmatrix bezeichnet.

Hat man nun andere Basen X', Y' in V bzw. W gegeben mit Transformationsmatrizen $T = \mathrm{Mat}_{XX'}(id_V)$ und $S = \mathrm{Mat}_{YY'}(id_W)$, so erhält man aus 12.16:

$$\begin{array}{ccccccc} \mathrm{Mat}_{XX'}(\mathrm{id}_V) & \cdot & \mathrm{Mat}_{X'Y'}(f) & \cdot & \mathrm{Mat}_{Y'Y}(\mathrm{id}_W) & = & \\ T & \cdot & \mathrm{Mat}_{X'Y'}(f) & \cdot & S^{-1} & = & \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}. \end{array}$$

Fassen wir diese Beobachtungen zusammen:

13.6 Satz
Sei K ein Divisionsring.

(1) *Ist $f : V \to W$ ein Homomorphismus von K-Vektorräumen mit Rang $f = r$, so gilt für geeignete Basen X, Y von V bzw. W,*

$$\mathrm{Mat}_{XY}(f) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

(2) *Ist $A \in K^{(m,n)}$ eine Matrix mit Zeilenrang r, dann gibt es invertierbare Matrizen $T_1 \in K^{(m,m)}$ und $S_1 \in K^{(n,n)}$ mit*

$$T_1 A S_1 = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Man nennt dies die Normalform *von A.*

(3) *Ist $B \in K^{(m,n)}$ eine weitere Matrix mit Zeilenrang r, dann gibt es invertierbare $T_2 \in K^{(m,m)}$, $S_2 \in K^{(n,n)}$ mit*

$$T_2 A S_2 = B.$$

Beweis: (1) und (2) wurden in den vorangehenden Ausführungen gezeigt.

(3) läßt sich leicht aus (2) ableiten. □

Aus den gewonnenen Beziehungen können wir Angaben über den Rang von Matrizen ableiten:

13.7 Satz
Für jede Matrix $A \in K^{(m,n)}$ gilt:

(1) *Sind $T \in K^{(m,m)}$ und $S \in K^{(n,n)}$ invertierbar, so ist*

$$\textit{Zeilenrang } A = \textit{Zeilenrang } TAS.$$

(2) *Ist K kommutativ und T, S wie oben, dann gilt*

$$\textit{Spaltenrang } A = \textit{Spaltenrang } TAS.$$

In diesem Fall ist Zeilenrang A = Spaltenrang A. Man spricht dann vom *Rang von A.*

Beweis: (1) folgt unmittelbar aus Satz 13.6.

(2) Wenden wir (2) aus Satz 13.6 auf die Matrix $A^t \in K^{(n,m)}$ an, so finden wir invertierbare Matrizen $C \in K^{(n,n)}$ und $D \in K^{(m,m)}$ mit

$$CA^tD = \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} \in K^{(n,m)},$$

mit $s =$ Zeilenrang $A^t =$ Spaltenrang A.

Durch Transponieren erhalten wir

$$D^tAC^t = \begin{pmatrix} E_s & 0 \\ 0 & 0 \end{pmatrix} \in K^{(m,n)}.$$

Nach (1) bedeutet dies $s =$ Zeilenrang $A =$ Spaltenrang A. □

Bemerkung: Für Matrizen A über nicht-kommutativen Divisionsringen K gilt (allgemeiner als (2) in 13.7): Der Rang des Zeilenraums als Linksvektorraum ist gleich dem Rang des Spaltenraums als Rechtsvektorraum über K.

In einem endlich-dimensionalen Vektorraum kann es keine endlos aufsteigenden Unterräume geben. Dies impliziert viele nützliche Eigenschaften des Endomorphismenrings. Eine davon wollen wir herausstellen:

13.8 Fitting'sches Lemma
Sei V ein endlich-dimensionaler Vektorraum über einem Divisionsring K und $g \in \mathrm{End}_K(V)$.

(1) *Es gibt ein $k \in \mathbb{N}$ mit* $\operatorname{Kern} g^k = \operatorname{Kern} g^{k+l}$ *für alle $l \in \mathbb{N}$.*

(2) *Für das k aus (1) gilt:* $V = \operatorname{Kern} g^k \oplus \operatorname{Bild} g^k$.
Man nennt dies eine *Fittingzerlegung von V bezüglich g.*

Beweis: (1) Betrachten wir zunächst die aufsteigende Kette von Unterräumen

$$\operatorname{Kern} g \subset \operatorname{Kern} g^2 \subset \operatorname{Kern} g^3 \subset \dots .$$

Dies kann keine echt aufsteigende unendliche Kette von Unterräumen sein, da sie durch die Dimension von V beschränkt ist. Also gibt es ein $k \in \mathbb{N}$ mit $\operatorname{Kern} g^k = \operatorname{Kern} g^{k+1}$.

Wir schließen nun durch Induktion nach l. $l = 1$ ist klar. Nehmen wir an $\operatorname{Kern} g^k = \operatorname{Kern} g^{k+l}$.

Es gilt immer $\operatorname{Kern} g^k \subset \operatorname{Kern} g^{k+l+1}$. Zeige $\operatorname{Kern} g^k \supset \operatorname{Kern} g^{k+l+1}$.
Für $u \in \operatorname{Kern} g^{k+l+1}$ ist

$$0 = g^{k+l+1}(u) = g^{k+l}(g(u)),$$

also – nach Induktionsannahme – $g(u) \in \operatorname{Kern} g^{k+l} = \operatorname{Kern} g^k$. Das bedeutet aber $g^{k+l}(g(u)) = g^{k+1}(u) = 0$ und $u \in \operatorname{Kern} g^{k+1} = \operatorname{Kern} g^k$.

(2) Sei k wie oben. Zeigen wir zunächst $\operatorname{Kern} g^k \cap \operatorname{Bild} g^k = 0$.
Für $v \in \operatorname{Kern} g^k \cap \operatorname{Bild} g^k$ gilt

$$g^k(v) = 0 \quad \text{und} \quad v = g^k(u) \text{ für ein } u \in V.$$

Das bedeutet $g^{2k}(u) = 0$ und – nach (1) – $u \in \operatorname{Kern} g^k$, also $v = g^k(u) = 0$.

Nun folgt aus dem Dimensionssatz

$$\dim V = \dim \operatorname{Kern} g^k + \dim \operatorname{Bild} g^k,$$

und somit $V = \operatorname{Kern} g^k \oplus \operatorname{Bild} g^k$. □

13.9 Aufgaben

(1) Sei V ein endlich-dimensionaler K-Vektorraum. Man zeige, daß für $f \in \operatorname{End}_K(V)$ folgende Aussagen äquivalent sind:

(a) *Es existiert ein $k \in K$ mit $f = k \cdot \mathrm{id}_V$;*

(b) *Für alle eindimensionalen Unterräume U von V gilt $f(U) \subset U$;*

(c) *Für alle Unterräume U von V gilt $f(U) \subset U$;*

(d) *für alle $g \in \operatorname{End}_K(V)$ gilt $f \circ g = g \circ f$.*

(2) V sei ein $\mathbb{R}$-Vektorraum mit Basis $B = (b_1, b_2, b_3, b_4)$.

(i) *Zeigen Sie, daß es genau eine lineare Abbildung $f : V \to V$ gibt mit*
$f \circ f = f$,
$f(b_1 + 2b_3 + b_4) = b_1 + b_2 - b_3 - b_4$, *und*
$f(2b_1 + b_2 - b_3) = b_1 + b_3$.

(ii) *Man bestimme* $\mathrm{Mat}(f)$ *bezüglich der Basis* B.

(iii) *Wie groß ist der Rang von* f*?*

(3) V *und* W *seien* $\mathbb{R}$*-Vektorräume mit Basen* $B = (b_1, b_2, b_3) \subset V$ *und* $C = (c_1, c_2, c_3, c_4) \subset W$*. Gegeben seien* $\mathbb{R}$*-Homomorphismen* $f : V \to W$ *und* $g : W \to V$ *durch*

$$f(r_1b_1 + r_2b_2 + r_3b_3) = (2r_2 - r_1 + r_3)c_1 + (r_3 - 3r_1)c_2 \text{ für } r_i \in \mathbb{R}$$

$$\begin{array}{lll} \textit{und} & g(c_1) = 2b_1 - b_2 + 7b_3, & g(c_2) = 5b_2 - 3b_1, \\ & g(c_3) = -b_1 - 2b_3, & g(c_4) = b_1 + b_2 + b_3. \end{array}$$

(i) *Bezüglich der Basen* $B \subset V$ *und* $C \subset W$ *bestimme man die Matrizen* $\mathrm{Mat}(f)$, $\mathrm{Mat}(g)$, $\mathrm{Mat}(g \circ f)$ *und* $\mathrm{Mat}(f \circ g)$.

(ii) *Man bestimme den Rang von* f, g, $g \circ f$ *und* $f \circ g$.

(iii) *Man gebe das Bild von* $-b_1 - b_2 + 3b_3 \in V$ *unter* $g \circ f$ *und das Bild von* $c_1 - c_3 \in W$ *unter* $f \circ g$ *an.*

(4) Sei $n \in \mathbb{N}$ *und* $\mathrm{Pol}_n(\mathbb{R}) := \{p \in \mathbb{R}[x] \mid p = 0 \text{ oder } \mathrm{grad}\, p \leq n\}$.
$f \in \mathrm{Hom}(\mathrm{Pol}_3(\mathbb{R}), \mathrm{Pol}_1(\mathbb{R}))$ *sei definiert durch*

$$f(\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3) := \alpha_0 - \alpha_2 + \alpha_3 + (\alpha_1 - \alpha_3)x.$$

(i) *Man verifiziere, daß* $B := (1, x, 2x^2 - 1, 4x^3 - 3x)$ *bzw.* $C := (1 + x, 1 - x)$ *eine Basis von* $\mathrm{Pol}_3(\mathbb{R})$ *bzw.* $\mathrm{Pol}_1(\mathbb{R})$ *ist.*

(ii) *Man berechne* $\mathrm{Mat}_{B,C}(f)$.

(iii) *Bestimmen Sie den Kern und das Bild (durch Angabe einer Basis) von* f *sowie den Rang von* $\mathrm{Mat}_{B,C}(f)$.

(5) Sei $B = (b_1, b_2, b_3, b_4)$ *eine Basis des* $\mathbb{R}$*-Vektorraums* V *und* $C = (c_1, c_2, c_3)$ *eine Basis des* $\mathbb{R}$*-Vektorraums* W*. Die lineare Abbildung* $f : V \to W$ *sei gegeben durch die Matrix*

$$\mathrm{Mat}_{B,C}(f) = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & -1 \\ 3 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{(4,3)}.$$

(i) *Man berechne* $\mathrm{Rang}\, f$ *und* $\dim \mathrm{Kern}\, f$.

(ii) *Man gebe jeweils eine Basis von* $\mathrm{Bild}\, f$ *und von* $\mathrm{Kern}\, f$ *an.*

(iii) *Besitzt* $8c_1 + 3c_2 + 2c_3 \in W$ *ein Urbild unter* f*?*

(iv) *Man bestimme das Bild von* $b_1 + b_2 + b_3 + b_4 \in V$ *unter* f.

(v) *Finden Sie eine Basis* B' *von* V, *so daß* $\mathrm{Mat}_{B'C}(f)$ *Zeilenstufenform hat.*

(vi) *Finden Sie eine Basis* C' *von* W, *so daß* $\mathrm{Mat}_{B'C'}(f)$ *Normalform hat.*

(6) Seien $f : \mathbb{R}^3 \to \mathbb{R}^3$ *und* $g : \mathbb{R}^3 \to \mathbb{R}^2$ *lineare Abbildungen. Bezüglich der kanonischen Basen* $E_2 \subset \mathbb{R}^2$ *und* $E_3 \subset \mathbb{R}^3$ *gelte*

$$\mathrm{Mat}_{E_3,E_3}(f) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ -10 & 1 & 4 \end{pmatrix} \text{ und } \mathrm{Mat}_{E_3,E_2}(g \circ f) = \begin{pmatrix} -1 & 1 \\ -2 & 4 \\ 3 & 7 \end{pmatrix}.$$

(a) *Man zeige, daß* f *ein Isomorphismus ist und berechne* $\mathrm{Mat}_{E_3,E_3}(f^{-1})$.

(b) *Man bestimme* $\mathrm{Mat}_{E_3,E_2}(g)$.

(7) Es sei

$$A := \begin{pmatrix} -1 & -2 & -3 \\ 2 & 3 & 5 \\ -2 & 3 & 1 \\ -2 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{(4,3)}.$$

(a) *Man ermittle den Rang von* A.

(b) *Man bestimme invertierbare Matrizen* $S \in \mathbb{R}^{(4,4)}$ *und* $T \in \mathbb{R}^{(3,3)}$ *mit*

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

(c) *Gibt es invertierbare Matrizen* $\widetilde{S} \in \mathbb{R}^{(4,4)}$ *und* $\widetilde{T} \in \mathbb{R}^{(3,3)}$ *mit*

$$\widetilde{S}A\widetilde{T}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}?$$

Begründung!

Hinweis: Man betrachte den Homomorphismus $f : \mathbb{R}^4 \to \mathbb{R}^3, x \mapsto xA$.

(8) Seien V *ein zweidimensionaler* $\mathbb{R}$*-Vektorraum mit Basis* $B = (b_1, b_2)$ *und* W *ein dreidimensionaler* $\mathbb{R}$*-Vektorraum mit Basis* $C = (c_1, c_2, c_3)$.

(i) *Man zeige, daß es genau einen* $\mathbb{R}$*-Homomorphismus* $f : V \to V$ *mit* $f \circ f = f + \mathrm{id}_V$ *und* $f(b_1 - 3b_2) = 2b_1 + b_2$ *gibt.*

(ii) *Man bestimme* $\mathrm{Mat}_{B,B}(f)$ *sowie* $\mathrm{Rang}\, f$.

(iii) *Für* $\operatorname{Kern} f$ *und* $\operatorname{Bild} f$ *finde man jeweils eine Basis.*

(iv) *Sei* $g : V \to W$ *als* $\mathbb{R}$*-Homomorphismus durch*
$g(\alpha_1 b_1 + \alpha_2 b_2) = (\alpha_1 - \alpha_2)c_1 + (2\alpha_1 - \lambda\alpha_2)c_2 + (3\alpha_1 - \mu\alpha_2)c_3$
für $\alpha_1, \alpha_2, \alpha_3, \lambda, \mu \in \mathbb{R}$ *definiert. Man bestimme* $\operatorname{Mat}_{B,C}(g \circ f)$ *und ermittle, welche Werte der Rang von* $g \circ f$ *in Abhängigkeit von* $\lambda, \mu \in \mathbb{R}$ *annehmen kann.*

(9) Für $\alpha \in \mathbb{R}$ *sei* $T_\alpha \in \mathbb{R}^{(2,3)}$ *die Matrix*

$$T_\alpha := \begin{pmatrix} 1 & \alpha & \alpha^2 \\ \alpha & 4 & 8 \end{pmatrix}.$$

Folgende lineare Abbildungen seien gegeben :

$$\begin{aligned} \varphi_\alpha &: \mathbb{R}^2 \to \mathbb{R}^3, \qquad (x_1, x_2) \mapsto (x_1, x_2)T_\alpha, \qquad \text{und} \\ \varphi_\alpha^t &: \mathbb{R}^3 \to \mathbb{R}^2, \qquad (x_1, x_2, x_3) \mapsto (x_1, x_2, x_3)T_\alpha^t. \end{aligned}$$

(a) *Man berechne (in Abhängigkeit von* α*)* $\operatorname{Rang} \varphi_\alpha$.

(b) *Daraus bestimme man die folgenden Dimensionen:*
$\dim \operatorname{Kern} \varphi_\alpha$, $\dim \operatorname{Bild} \varphi_\alpha$, $\dim \operatorname{Kern} \varphi_\alpha^t$ *und* $\dim \operatorname{Bild} \varphi_\alpha^t$.

(c) *Man bestimme jeweils eine Basis von* $\operatorname{Kern} \varphi_\alpha$, $\operatorname{Bild} \varphi_\alpha$, $\operatorname{Kern} \varphi_\alpha^t$ *und* $\operatorname{Bild} \varphi_\alpha^t$.

(10) Betrachten Sie die Matrix über den Quaternionen $\mathbb{H}$,

$$A := \begin{pmatrix} i & -1 \\ j & k \end{pmatrix},$$

wobei i, j, k *die kanonischen Basiselemente von* $\mathbb{H}$ *sind. Zeigen Sie:*

(i) *Der Spaltenrang von* A *als Linksvektorraum über* $\mathbb{H}$ *ist* 1.

(ii) *Der Zeilenrang von* A *als Linksvektorraum über* $\mathbb{H}$ *ist* 2.

14 Lineare Gleichungen

In diesem Abschnitt wollen wir zeigen, wie die bislang entwickelten Methoden zur Behandlung von *linearen Gleichungssystemen* genutzt werden können. Wir beginnen mit der Betrachtung von allgemeinen *Gleichungen*.

14.1 Definitionen
Seien X, Y Mengen und $f : X \to Y$ eine Abbildung.

(1) Zu jedem $b \in Y$ nennt man das Paar (f, b) eine *Gleichung*.

(2) Ein Element $x \in X$ mit $f(x) = b$ heißt *Lösung* dieser Gleichung.

(3) Das Urbild $f^{-1}(b) \subset X$ von b unter f heißt *Lösungsmenge* von (f, b). Sie ist also gleich $\{x \in X \mid f(x) = b\}$.

(4) Ist die Lösungsmenge nicht leer, so heißt die Gleichung *lösbar*.

Sei nun R ein Ring und $f : M \to N$ eine *lineare Abbildung* von R-Moduln. Dann nennt man (f, b) für $b \in N$ eine *lineare Gleichung*.
Es ist klar, daß (f, b) genau dann lösbar ist, wenn $b \in \operatorname{Bild} f$.
Für $b = 0$ kennen wir die Lösungsmenge bereits, es ist $\operatorname{Kern} f$.
Man nennt $(f, 0)$ eine *homogene Gleichung* und (f, b) eine *inhomogene Gleichung*, wenn $b \neq 0$.

Eine Übersicht über die Gesamtheit der Lösungen wird gegeben durch

14.2 Satz
Sei $f : M \to N$ eine lineare Abbildung von R-Moduln und $b \in N$. Ist (f, b) lösbar, so ist die Lösungsmenge gleich $x_0 + \operatorname{Kern} f$, wobei x_0 eine (beliebige) Lösung der Gleichung ist.

Beweis: Ist $x_0 \in M$ eine Lösung von (f, b), dann gilt für jede Lösung $x \in M$

$$f(x - x_0) = f(x) - f(x_0) = b - b = 0,$$

also $x - x_0 \in \operatorname{Kern} f$ und $x \in x_0 + \operatorname{Kern} f$.

Andererseits ist für jedes $z \in \operatorname{Kern} f$

$$f(x_0 + z) = f(x_0) + f(z) = f(x_0) = b,$$

und damit ist $x_0 + z$ eine Lösung von (f, b). □

Für die weitere Behandlung von Gleichungen wollen wir uns auf Vektorräume über Körpern beschränken. Ein Teil der Aussagen gilt auch noch für Divisionsringe.

Sei also K ein Körper. Von *linearen Gleichungssystemen* spricht man, wenn man speziell die K-Homomorphismen $f : K^{(n,1)} \to K^{(m,1)}$ betrachtet. Sie werden durch Multiplikation mit einer Matrix $A = (a_{ij}) \in K^{(m,n)}$ dargestellt:

$$f : K^{(n,1)} \longrightarrow K^{(m,1)}, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Für ein $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^{(m,1)}$ hat dann die Gleichung (f, b) die Gestalt:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array} \qquad (*)$$

Man nennt dies ein *lineares Gleichungssystem.*

Setzen wir $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, so haben wir $A \cdot x = b$.

Wir wissen, daß in der gewählten Darstellung von f das Bild f von den Spalten von A erzeugt wird. Somit ist das System genau dann lösbar, wenn b von den Spalten linear abhängig ist. Wir können also festhalten:

14.3 Satz

(1) *Das Gleichungssystem* $(*)$ *ist genau dann lösbar, wenn*

$$\operatorname{Rang} A = \operatorname{Rang}(A, b),$$

wobei $(A, b) \in K^{(m,n+1)}$ *die* erweiterte Matrix *heißt.*

(2) *Die Lösungsgesamtheit des homogenen Systems* $A \cdot x = 0$ *ist ein Unterraum* $U \subset K^{(n,1)}$ *mit Dimension* $= n - \operatorname{Rang} A$.

(3) *Die Lösungsgesamtheit von* $(*)$ *ist leer oder von der Form* $y + U$, *wobei* $y = (y_1, \dots, y_n)^t \in K^{(n,1)}$ *eine spezielle Lösung von* $(*)$ *ist und* U *wie in (2).*

(4) *Die Gleichung ist genau dann eindeutig lösbar, wenn* $\operatorname{Rang} A = n$.

Beweis: (1) Wie oben festgestellt, ist $(*)$ genau dann lösbar, wenn b von den Spalten von A linear abhängig ist. Dies ist äquivalent zu der Forderung

$\operatorname{Rang} A = \text{Spaltenrang } A = \text{Spaltenrang } (A, b) = \operatorname{Rang}(A, b)$.

(2) Die Lösungsmenge des homogenen Systems ist gerade der Kern der von A bestimmten Abbildung $f : K^{(n,1)} \to K^{(m,1)}$. Somit ist sie ein K-Unterraum mit $\dim U = n - \operatorname{Rang} f$ (Dimensionsformel 11.1).

(3) Die Behauptung aus 14.2.

(4) Ist die Gleichung eindeutig lösbar, so muß $\operatorname{Kern} f = 0$ gelten, und damit $\operatorname{Rang} f = \operatorname{Rang} A = n$.

Gilt andererseits $\operatorname{Rang} A = n$, dann ist auch $\operatorname{Rang}(A, b) = n$, und nach (1) ist das System lösbar. Außerdem gilt in diesem Fall $\operatorname{Kern} f = 0$, und somit ist die Lösung eindeutig. □

Mit Hilfe von Matrizenumformungen (vgl. 12.16) erhalten wir folgende Möglichkeit zum Lösen eines Gleichungssystems:

Multipliziert man die Gleichung $Ax = b$ von links mit einer invertierbaren (m, m)-Matrix B, so ist offenbar die Lösungsmenge der Gleichung

$$BAx = Bb$$

die gleiche wie für die Ausgangsgleichung. Speziell heißt dies, daß man die gleiche Lösungsmenge erhält, wenn man auf die Matrizen

$$A \in K^{(m,n)},\ b \in K^{(m,1)}$$

die gleichen Zeilenumformungen anwendet. Dies ist natürlich nicht verwunderlich, weil dabei eben nur Linearkombinationen der gegebenen Gleichungen gebildet werden.

Bringt man A auf diese Weise in die Zeilenstufenform von 13.4, so hat man ein Gleichungssystem der Form (mit $k = \operatorname{Rang} A$)

$$\begin{array}{ccccccccccc}
c_{1,r_1}x_{r_1} & + & \dots\dots\dots\dots\dots\dots\dots\dots & + & c_{1,n}x_n & = & b'_1 \\
& c_{2,r_2}x_{r_2} & + & \dots\dots\dots\dots\dots & + & c_{2,n}x_n & = & b'_2 \\
& & \ddots & & & \vdots & & \vdots \\
& & & c_{k,r_k}x_{r_k} & + \dots + & c_{k,n}x_n & = & b'_k \\
& & & & & 0 & = & b'_{k+1}
\end{array}$$

Ist das System lösbar, so muß sich dabei $b_{k+1} = 0$ ergeben. $b_{k+1} \neq 0$ bedeutet, daß das System nicht lösbar ist.

Die oben beschriebene Umformung des Gleichungssystems bezeichnet man als *Gauß'sches Eliminationsverfahren*:

Man sucht die Gleichung, in der ein $a_{i,r_1}x_{r_i} \neq 0$ mit möglichst kleinem r_1 vorkommt. Diese setzt man in die erste Zeile und *eliminiert* x_{r_1} aus den anderen Gleichungen durch Addition eines geeigneten Vielfachen der ersten Zeile. Diese Prozedur wiederholt man, bis die gewünschte Form des Systems erreicht ist.

Bemerkung: Um $\mathrm{Rang}(A, b)$ zu bestimmen, könnte man auch *Spaltenumformungen* darauf anwenden. Bei der gewählten Schreibweise würden diese jedoch eine Veränderung der Unbekannten x_i bewirken. Zur Bestimmung einer Lösung sind sie daher hier nicht geeignet.

Sehen wir uns die oben angegebene Methode in einem konkreten Fall an.

Beispiel
Gegeben sei das Gleichungssystem mit Koeffizienten aus $\mathbb{Q}$:

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 2 \\ 2x_1 & + & 4x_2 & + & 3x_3 & = & -1 \\ 4x_1 & + & 6x_2 & + & 5x_3 & = & 3 \end{array}$$

Wir nehmen die Koeffizienten als Matrix heraus und führen dann die angegebenen Zeilenoperationen durch:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & 4 & 3 & -1 \\ 4 & 6 & 5 & 3 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ \mathrm{II} - 2 \cdot \mathrm{I} & 0 & 2 & 1 & -5 \\ \mathrm{III} - 4 \cdot \mathrm{I} & 0 & 2 & 1 & -5 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & \frac{1}{2} & -\frac{5}{2} \end{array}\right)$$

Um eine spezielle Lösung zu finden, kann man nun x_3 beliebig wählen, nehmen wir $x_3 = 0$. Aus der letzten Zeile ergibt sich dann $x_2 = -\frac{5}{2}$.

Durch Einsetzen dieser Werte in die erste Zeile findet man

$$x_1 = -x_2 - x_3 + 2 = \frac{5}{2} + \frac{4}{2} = \frac{9}{2}.$$

Somit haben wir eine spezielle Lösung: $x_1 = \frac{9}{2}$, $x_2 = -\frac{5}{2}$, $x_3 = 0$.

Nun zur Lösung der homogenen Gleichung. Auch diese kann man aus der letzten Matrix ablesen. Da die Matrix den Rang 2 hat, hat die Lösungsmenge Dimension 1. Jede Lösung $\neq 0$ ist daher Basis des Lösungsraums. Wie leicht zu sehen ist, ist $x_1 = -1$, $x_2 = -1$, $x_3 = 2$ eine nicht-triviale Lösung, und $\mathbb{Q} \cdot (-1, -1, 2)$ ist der Lösungsraum des homogenen Teils.

Als Lösungsgesamtheit der Ausgangsgleichnung haben wir dann

$$\left(\frac{9}{2}, -\frac{5}{2}, 0\right) + \mathbb{Q} \cdot (-1, -1, 2).$$

Durch Lösen eines homogenen Gleichungssystems läßt sich auch der Kern eines Homomorphismus bestimmen. Man benutzt dabei die zugeordnete Matrix:

14.4 Kern eines Homomorphismus
V und W seien endlich-dimensionale K-Vektorräume mit Basen X bzw. Y. Ist $f : V \to W$ ein Homomorphismus und $A = \mathrm{Mat}_{XY}(f)$, dann sind die Koordinaten der Elemente von $\mathrm{Kern}\, f$ *gerade die Lösungsmenge des Gleichungssystems $A^t \cdot x^t = 0$.*

Beweis: Wegen Satz 12.14 können wir ohne Einschränkung $V = K^m$ und $W = K^n$ annehmen. Dann ist $A \in K^{(m,n)}$, und die Bilder unter f sind Linearkombinationen der Zeilen von A. Damit sich Null ergibt, müssen die Koeffizienten $x = (x_1, \ldots, x_m)$ der Gleichung $x \cdot A = 0$ genügen. Durch Transponieren erhält man die gewünschte Form $A^t \cdot x^t = 0$. □

Sehen wir auch dazu einen konkreten Fall an.

Beispiel
Sei $f : \mathbb{Q}^{(1,3)} \to \mathbb{Q}^{(1,4)}$ ein Homomorphismus, dessen Matrix bezüglich der kanonischen Basen folgende Gestalt hat:

$$\mathrm{Mat}(f) = \begin{pmatrix} 1 & -2 & 1 & 2 \\ 1 & 1 & -1 & 1 \\ 1 & 7 & -5 & -1 \end{pmatrix}$$

Durch Zeilenumformungen erhalten wir eine Basis von Bild f:

$$\left.\begin{array}{rrrr} 1 & -2 & 1 & 2 \\ 0 & 3 & -2 & -1 \\ 0 & 9 & -6 & -3 \end{array}\right| \quad \begin{array}{rrrr} 1 & -2 & 1 & 2 \\ & 3 & -2 & -1 \end{array}$$

Daraus ersehen wir, daß $\mathrm{Rang}\, f = 2$ und $\dim \mathrm{Kern}\, f = 3 - 2 = 1$.

Den Kern erhalten wir als Lösungsmenge des durch $\mathrm{Mat}(f)^t$ bestimmten Gleichungssystems

$$\begin{pmatrix} 1 & 1 & 1 \\ -2 & 1 & 7 \\ 1 & -1 & -5 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

Zeilenumformungen ergeben die Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix},$$

aus der sich eine Lösung $(x_1, x_2, x_3) = (2, -3, 1)$ ablesen läßt.

Insgesamt gilt damit $\mathrm{Kern}\, f = \mathbb{Q} \cdot (2, -3, 1)$.

14.5 Aufgaben

(1) Für welche $\lambda \in \mathbb{R}$ ist das folgende Gleichungssystem lösbar? Man bestimme dazu die Lösungsmenge.

$$\begin{array}{rcrcrcrcl} 5\,x_1 & - & 3\,x_2 & + & 2\,x_3 & + & 4\,x_4 & = & 3 \\ 4\,x_1 & - & 2\,x_2 & + & 3\,x_3 & + & 7\,x_4 & = & 1 \\ 8\,x_1 & - & 6\,x_2 & - & x_3 & - & 5\,x_4 & = & 9 \\ 7\,x_1 & - & 3\,x_2 & + & 7\,x_3 & + & 17\,x_4 & = & \lambda \end{array}$$

(2) Man untersuche, für welche $a, b \in \mathbb{R}$ das folgende Gleichungssystem über $\mathbb{R}$ lösbar ist und bestimme dann für diese $a, b \in \mathbb{R}$ die Lösungsmenge.

$$\begin{array}{rcrcrcl} a^2\,x & + & 5\,y & + & z & = & b \\ a\,x & + & (a+3)\,y & + & 3\,z & = & 0 \\ x & + & 2\,y & + & z & = & 0 \end{array}$$

(3) Man bestimme in Abhängigkeit von $t \in \mathbb{R}$ die Lösungsmenge des folgenden Gleichungssystems über $\mathbb{R}$:

$$\begin{array}{rcrcrcrcr} x_1 & + & x_2 & - & 2\,x_3 & = & 1 & & \\ 3\,x_1 & + & 4\,x_2 & - & 4\,x_3 & = & 4 & + & t \\ x_1 & - & x_2 & - & 6\,x_3 & = & 1 & - & 4t \end{array}$$

(4) Man bestimme in Abhängigkeit von $t \in \mathbb{R}$ die Lösungsmenge des folgenden Gleichungssystems über $\mathbb{R}$:

$$\begin{array}{rcrcrcl} x & + & 2\,y & - & z & = & 1 \\ 2\,x & + & 5\,y & + & t\,z & = & 3 \\ x & + & (t+1)\,y & + & 3\,z & = & 2 \end{array}$$

Kapitel 5

Determinante und Spur

15 Determinanten

Über kommutativen Ringen können Eigenschaften von Endomorphismen von Moduln und Matrizen durch eine numerische Größe beschrieben werden, die *Determinante*.

Nach einigen theoretischen Überlegungen dazu werden wir Berechnungsmöglichkeiten dafür angeben. Die Kenntnis der abstrakten Grundlagen erleichtern sowohl das Verständnis als auch die Herleitung von Eigenschaften der Determinante.

Insbesondere bei der Ermittlung des Wertes einer Determinante ist die Kommutativität des Grundringes sehr wichtig. Wir wollen daher in diesem Abschnitt voraussetzen, daß R ein *kommutativer* Ring ist.

Beginnen wir mit einer Verallgemeinerung von linearen Abbildungen.

15.1 Multilineare Abbildungen

M, N seien R-Moduln, und für $n \in \mathbb{N}$ bezeichne M^n das n-fache Produkt $M \times \ldots \times M$. Eine Abbildung

$$d : M^n \to N, \quad (a_1, \ldots, a_n) \mapsto d(a_1, \ldots, a_n),$$

nennt man *multilinear* oder *n-linear*, wenn sie in den einzelnen a_i R-linear ist, wenn also für jedes $i \leq n$ und $r, s \in R$ gilt

$$(a_1, \ldots, ra_i + sb_i, \ldots, a_n) = r(a_1, \ldots, a_i, \ldots, a_n) + s(a_1, \ldots, b_i, \ldots, a_n).$$

Eine multilineare Abbildung d heißt *alternierend*, wenn $d(a_1, \ldots, a_n) = 0$, falls $a_i = a_j$ für ein Paar $i \neq j$ gilt.

Für $n = 1$ ist *multilinear* offenbar gleichbedeutend mit *linear*.

Man beachte, daß eine multilineare Abbildung *keine* lineare Abbildung auf dem R-Modul M^n ist.

Folgerung
Ist d eine alternierende Abbildung, so ändert sich bei Vertauschung von zwei Komponenten das Vorzeichen, also z.B.

$$d(a_1, a_2, a_3, \ldots, a_n) = -d(a_2, a_1, a_3, \ldots, a_n).$$

Beweis: Dies sieht man aus den Gleichungen

$$\begin{aligned} 0 &= d(a_1 + a_2, a_1 + a_2, a_3, \ldots, a_n) \\ &= d(a_2, a_1, a_3, \ldots, a_n) + d(a_1, a_2, a_3, \ldots, a_n) \\ &\quad + d(a_1, a_1, a_3, \ldots, a_n) + d(a_2, a_2, a_3, \ldots, a_n) \\ &= d(a_2, a_1, a_3, \ldots, a_n) + d(a_1, a_2, a_3, \ldots, a_n), \end{aligned}$$

da die Funktionswerte bei zwei gleichen Komponenten Null werden. □

Man kann daraus ableiten, wie sich die Abbildung bei beliebiger Vertauschung der Komponenten verhält.

Erinnern wir uns, daß man eine Permutation σ *gerade* nennt, wenn sie Produkt einer geraden Anzahl von Transpositionen ist, wenn also $\operatorname{sgn}\sigma = 1$ (vgl. 6.21).

15.2 Hilfssatz
Sei M ein R-Modul und $d : M^n \to N$ eine multilineare, alternierende Abbildung. Dann gilt für alle $a_1, \ldots, a_n \in M$ und $\sigma \in \mathcal{S}_n$

$$d(a_1, \ldots, a_n) = \operatorname{sgn}\sigma \cdot d(a_{\sigma(1)}, \ldots, a_{\sigma(n)}).$$

Beweis: Für eine Permutation $\sigma = \tau_1 \cdots \tau_k$ mit Transpositionen τ_i, gilt $\operatorname{sgn}\sigma = (-1)^k$ und

$$d(a_1, \ldots, a_n) = (-1)^k d(a_{\sigma(1)}, \ldots, a_{\sigma(n)}).$$

□

Damit kann man nun angeben, wie eine multilineare, alternierende Abbildung $M^n \to R$ aussehen muß, falls sie existiert:

15.3 Hilfssatz
Sei M ein R-Modul mit Basis $(b_1, \ldots, b_n)$. Ist $d : M^n \to R$ multilinear und alternierend, so gilt für $a_i = \sum_{j=1}^{n} a_{ij} b_j \in M$, $(a_{ij}) \in R^{(n,n)}$,

$$d(a_1, \ldots, a_n) = \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn}\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \cdot d(b_1, \ldots, b_n).$$

Beweis: Durch wiederholtes Ausnutzen der Linearität in den einzelnen Komponenten ergibt sich

$$\begin{aligned}
d(a_1,\ldots,a_n) &= d(\sum_{j_1=1}^{n} a_{1j_1}b_{j_1}, a_2,\ldots,a_n)\\
&= \sum_{j_1=1}^{n} a_{1j_1} d(b_{j_1}, a_2,\ldots,a_n)\\
&= \sum_{j_1=1}^{n} a_{1j_1}\left(\sum_{j_2=1}^{n} a_{2j_2} d(b_{j_1}, b_{j_2}, a_3,\ldots,a_n)\right)\\
&\quad\cdots\\
&= \sum_{(j_1,\ldots,j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n} d(b_{j_1},\ldots,b_{j_n})\\
\text{(alternierend)} &= \sum_{\sigma\in\mathcal{S}_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)}\cdot d(b_{\sigma(1)},\ldots,b_{\sigma(n)})\\
&= \sum_{\sigma\in\mathcal{S}_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)}\cdot (\operatorname{sgn}\sigma) d(b_1,\ldots,b_n)\\
&= \sum_{\sigma\in\mathcal{S}_n} (\operatorname{sgn}\sigma)\, a_{1\sigma(1)}\cdots a_{n\sigma(n)}\cdot d(b_1,\ldots,b_n).
\end{aligned}$$

□

Die angegebenen Eigenschaften von multilinearen, alternierenden Abbildungen weisen auch einen Weg zu deren Existenz und Eindeutigkeit:

15.4 Satz
Sei M ein R-Modul mit Basis $(b_1,\ldots,b_n)$ und $\beta\in R$ invertierbar.
Dann gibt es genau eine alternierende n-lineare Abbildung

$$d: M^n \to R, \quad \textit{mit } d(b_1,\ldots,b_n)=\beta.$$

Beweis: Sei d eine solche Abbildung und $d(b_1,\ldots,b_n)=\beta$ vorgegeben. Dann ist der Wert von d auf

$$(a_1,\ldots,a_n) \text{ mit } a_i=\sum_{j=1}^{n} a_{ij}b_j \in M, \quad (a_{ij})\in R^{(n,n)},$$

durch den in 15.3 gefundenen Ausdruck festgelegt.

Andererseits läßt sich dieser Ausdruck als Definition verwenden. Setzen wir

$$d(a_1,\ldots,a_n) := \beta\cdot\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\cdot a_{1\sigma(1)}\cdots a_{n\sigma(n)},$$

und zeigen, daß dadurch eine Abbildung mit den gewünschten Eigenschaften definiert ist.

Nehmen wir für (a_{ij}) die Einheitsmatrix, so erhalten wir $d(b_1, \ldots, b_n) = \beta$.

Prüfen wir die Linearität an der ersten Stelle. Dazu sei $a'_1 = \sum_{j=1}^{n} a'_{1j} b_j$. Wir haben

$$\begin{aligned} d(a_1 + a'_1, a_2, \ldots, a_n) &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}\sigma \cdot (a_{1\sigma(1)} + a'_{1\sigma(1)}) a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}\sigma \cdot [a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ &\qquad + a'_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}] \\ &= d(a_1, a_2, \ldots, a_n) + d(a'_1, a_2, \ldots, a_n). \end{aligned}$$

Für $r \in R$ bestätigt man leicht

$$d(r\, a_1, a_2, \ldots, a_n) = r\, d(a_1, a_2 \ldots, a_n).$$

Somit ist d linear in a_1. Durch die gleiche Rechnung an den anderen Stellen sieht man, daß d in der Tat multilinear ist.

Um zu sehen, ob d alternierend ist, nehmen wir ohne Einschränkung $a_1 = a_2$ an, also $a_{1j} = a_{2j}$. Dann ist

$$\begin{aligned} d(a_1, a_1, a_3, \ldots, a_n) &= \beta \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}\sigma\, a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)} \cdots a_{n\sigma(n)} \\ &= \beta \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}\sigma\, a_{1\sigma(1)} a_{1\sigma(2)} a_{3\sigma(3)} \cdots a_{n\sigma(n)} \\ &= \beta [\sum_{\sigma(1)<\sigma(2)} \operatorname{sgn}\sigma\, a_{1\sigma(1)} a_{1\sigma(2)} a_{3\sigma(3)} \cdots a_{n\sigma(n)} \\ &\qquad + \sum_{\sigma(2)<\sigma(1)} \operatorname{sgn}\sigma\, a_{1\sigma(2)} a_{1\sigma(1)} a_{3\sigma(3)} \cdots a_{n\sigma(n)}]. \end{aligned}$$

Sei τ die Vertauschung von 1 mit 2. Dann können wir oben schreiben:

$$\sum_{\sigma(2)<\sigma(1)} \operatorname{sgn}\sigma = \sum_{\sigma\tau(1)<\sigma\tau(2)} \operatorname{sgn}\sigma\tau = -\sum_{\sigma(1)<\sigma(2)} \operatorname{sgn}\sigma$$

Damit haben wir $d(a_1, a_1, a_3, \ldots, a_n) = 0$. Mit entsprechender Rechnung sieht man, daß $d(a_1, \ldots, a_n) = 0$ wird, wenn zwei beliebige Komponenten übereinstimmen.

Also ist d alternierend. □

Nunmehr sind wir in der Lage, die wichtigsten multilinearen, alternierenden Abbildungen zu kennzeichnen:

15.5 Satz (Determinantenformen)
Sei M ein freier R-Modul mit einer Basis der Länge n und $d : M^n \to R$ eine alternierende, n-lineare Abbildung. Dann sind folgende Aussagen äquivalent:

(a) *Es gibt eine Basis $(b_1, \ldots, b_n)$ von M mit $d(b_1, \ldots, b_n) = 1$;*

(b) *für jede Basis $(c_1, \ldots, c_n)$ von M ist $d(c_1, \ldots, c_n)$ invertierbar in R;*

(c) *zu jeder alternierenden, n-linearen Abbildung $h : M^n \to R$ gibt es ein $\delta \in R$ mit*

$$h(a_1, \ldots, a_n) = \delta \cdot d(a_1, \ldots, a_n).$$

Eine solche Abbildung $d : M^n \to R$ nennt man *Determinantenform.*
Ist R ein Körper, so ist *jedes* nicht-triviale d Determinantenform.

Beweis: (a)⇒(b) Stellt man die Basis $b_1, \ldots, b_n$ durch die $c_1, \ldots, c_n$ dar, so gilt nach 15.3

$$1 = d(b_1, \ldots, b_n) = d(c_1, \ldots, c_n) \cdot (\sum \operatorname{sgn} \ldots).$$

Somit ist $d(c_1, \ldots, c_n)$ invertierbar.

(b)⇒(a) Ist $d(c_1, \ldots, c_n) = r$ invertierbar, so ist $(\frac{1}{r}c_1, c_2, \ldots, c_n)$ eine Basis von M mit der gewünschten Eigenschaft.

(a)⇒(c) Mit $\delta := h(b_1, \ldots, b_n)$ erhalten wir das gewünschte Ergebnis. Nach 15.3 folgt nämlich aus $h(a_1, \ldots, a_n) = \delta \cdot d(a_1, \ldots, a_n)$, daß h und $\delta \cdot d$ die gleiche Abbildung beschreiben.

(c)⇒(b) Zu einer Basis $(c_1, \ldots, c_n)$ wählen wir die alternierende, n-lineare Abbildung $h : M^n \to R$ mit $h(c_1, \ldots, c_n) = 1$. Nach (c) gibt es $\delta \in R$ mit

$$h(c_1, \ldots, c_n) = \delta \cdot d(c_1, \ldots, c_n).$$

Somit ist $d(c_1, \ldots, c_n)$ invertierbar. □

Es ist leicht zu verifizieren, daß die n-linearen, alternierenden Abbildungen $M^n \to R$ einen R-Modul bilden, nämlich einen Untermodul von $\mathrm{Abb}(M^n, R)$. Die Aussage (c) in 15.5 läßt sich damit so interpretieren:

Der R-Modul der n-linearen, alternierenden Abbildungen $M^n \to R$ besitzt eine Basis aus einem Element.
Jede Determinantenform kann als Basis genommen werden.

Eine weitere wichtige Beobachtung wollen wir herausstellen:

15.6 Korollar
Sei M ein freier R-Modul und $d : M^n \to R$ sei n-linear und alternierend.

(1) *Sind $a_1, \ldots, a_n \in M$ linear abhängig, so ist $d(a_1, \ldots, a_n)$ Nullteiler in R.*

(2) *Sei R Körper und d eine Determinantenform. Dann sind für Elemente $a_1, \ldots, a_n \in M$ folgende Aussagen äquivalent:*
(a) $d(a_1, \ldots, a_n) = 0$;
(b) die $a_1, \ldots, a_n$ sind linear abhängig.

Beweis: (1) Seien die $a_1, \ldots, a_n$ linear abhängig, also $\sum_{i=1}^{n} r_i a_i = 0$ mit nicht allen $r_i \in R$ gleich Null. Ohne Einschränkung können wir $r_1 \neq 0$ annehmen. Dann gilt

$$r_1\, d(a_1, \ldots, a_n) = d(r_1 a_1, \ldots, a_n) = -d(\sum_{i=2}^{n} r_i a_i, a_2, \ldots, a_n) = 0.$$

Also ist $d(a_1, \ldots, a_n)$ Nullteiler in R.

(2) (a)⇒(b) Sind die $a_1, \ldots, a_n$ linear unabhängig, so bilden sie eine Basis von M, und nach 15.5 ist $d(a_1, \ldots, a_n) \neq 0$.

(b)⇒(a) ergibt sich aus (1). □

Wir benutzen die Determinantenformen, um die *Determinante* eines Endomorphismus $f : M \to M$ eines freien R-Moduls zu definieren.

Für jede n-lineare, alternierende Abbildung $d : M^n \to R$ ist offensichtlich auch folgende Komposition n-linear und alternierend:

$$\begin{array}{rccccc} d \circ f^n : & M^n & \to & M^n & \to & R \\ & (a_1, \ldots, a_n) & \mapsto & (f(a_1), \ldots, f(a_n)) & \mapsto & d(f(a_1), \ldots, f(a_n)). \end{array}$$

Wählt man nun für d eine Determinantenform, so braucht $d \circ f^n$ keine Determinantenform zu sein. Nach 15.5 gibt es jedoch ein $\delta \in R$ mit

$$d \circ f^n = \delta \cdot d.$$

Da sich zwei Determinantenformen d, d' nur um ein invertierbares Element $\gamma \in R$ unterscheiden, also $d' = \gamma d$, ist δ nur von f, nicht aber von der Auswahl von d abhängig.

15.7 Definition *(Determinante von Endomorphismen)*
Sei M ein freier R-Modul mit einer Basis der Länge n, und $d : M^n \to R$ eine Determinantenform.

Ist $f : M \to M$ ein Endomorphismus, so bezeichnet man das Element $\det(f) \in R$ mit

$$d \circ f^n = \det(f) \cdot d$$

als *Determinante von f*.

Bei Wahl einer Basis $(b_1, \ldots, b_n)$ von M haben wir

$$d(f(b_1), \ldots, f(b_n)) = \det(f) \cdot d(b_1, \ldots, b_n).$$

Da $d(b_1, \ldots, b_n)$ invertierbar ist (vgl. 15.5), erhält man daraus

$$\det(f) = \delta = \frac{d(f(b_1), \ldots, f(b_n))}{d(b_1, \ldots, b_n)}.$$

Diese Beziehung gibt uns die Möglichkeit zur Berechnung von $\det(f)$. Es ist leicht nachzuprüfen, daß das Ergebnis sowohl von der Wahl der Determinantenform als auch der Wahl der Basis $(b_1, \ldots, b_n)$ von M unabhängig ist.

Ohne eine genaue Rechenprozedur zu kennen, lassen sich aus der Definition schon wichtige Aussagen über die Determinante ableiten.

15.8 Eigenschaften der Determinante
Sei M ein freier R-Modul mit einer Basis der Länge n. Dann gilt:

(1) $\det(0) = 0$, $\det(\mathrm{id}) = 1$ *und* $\det(r \cdot \mathrm{id}) = r^n$ *für* $r \in R$.

(2) $f \in \mathrm{End}(M)$ *ist injektiv, wenn* $\det(f)$ *kein Nullteiler in R ist.*

(3) *Für* $f, g \in \mathrm{End}(M)$ *ist*

$$\det(g \circ f) = \det(f) \cdot \det(g) = \det(f \circ g).$$

Beweis: Sei d eine Determinantenform und $(b_1, \ldots, b_n)$ eine Basis von M.

(1) Diese Beziehungen sind leicht zu bestätigen.

(2) Ist f nicht injektiv, so sind die $f(b_1), \ldots, f(b_n)$ linear abhängig, und

$$d(f(b_1), \ldots, f(b_n)) = \det(f) \cdot d(b_1, \ldots, b_n)$$

ist nach 15.6 ein Nullteiler. Damit ist $\det(f)$ Nullteiler.

(3) Nach Definition der Determinante haben wir

$$\begin{aligned} \det(g \circ f) \cdot d(b_1, \ldots, b_n) &= d(g \circ f(b_1), \ldots, g \circ f(b_n)) \\ &= \det(g) \cdot d(f(b_1), \ldots, f(b_n)) \\ &= \det(g) \cdot \det(f) \cdot d(b_1, \ldots, b_n). \end{aligned}$$

Wir dürfen $d(b_1, \ldots, b_n)$ kürzen und bekommen dann (R kommutativ)

$$\det(g \circ f) = \det(g) \cdot \det(f) = \det(f \circ g).$$

□

Auch folgende interessante Beziehung können wir herleiten, ohne die Determinante explizit berechnen zu müssen:

15.9 Adjunkte Matrix eines Endomorphismus
Sei M ein R-Modul mit Basis $B = (b_1, \ldots, b_n)$ und $d : M^n \to R$ eine Determinantenform mit $d(b_1, \ldots, b_n) = 1$.

Zu $f \in \mathrm{End}_R(M)$ bilden wir die Matrix $\mathrm{Ad}(f) = (d_{kl}) \in R^{(n,n)}$ mit

$$d_{kl} = d(f(b_1), \ldots, b_k, \ldots, f(b_n)), \quad b_k \text{ an } l\text{-ter Stelle},$$

welche wir die zu f adjunkte *Matrix nennen. Dafür gilt*

$$\mathrm{Mat}_{BB}(f) \cdot \mathrm{Ad}(f) = \det(f) E_n,$$

wobei E_n die n-te Einheitsmatrix bezeichnet.

Beweis: Setzen wir $\mathrm{Mat}_{BB}(f) = (a_{ij}) \in R^{(n,n)}$, also $f(b_1) = \sum_{k=1}^n a_{1k} b_k$. Damit gilt

$$\det(f) = \det(f) \cdot d(b_1, \ldots, b_n) = d(\sum_{k=1}^n a_{1k} b_k, f(b_2), \ldots, f(b_n)) = \sum_{k=1}^n a_{1k}\, d_{k1}.$$

Außerdem gilt

$$0 = d(f(b_2), f(b_2), \cdots f(b_n)) = d(\sum_{k=1}^n a_{2k} b_k, f(b_2), \ldots f(b_n)) = \sum_{k=1}^n a_{2k}\, d_{k1}.$$

Mit analogen Einsetzungen an den verschiedenen Stellen erhalten wir

$$\sum_{k=1}^n a_{ik}\, d_{kj} = \delta_{ij} \det(f).$$

Dies ergibt die gewünschte Relation. □

Damit bekommen wir verschiedene Beschreibungen für Isomorphismen:

15.10 Kennzeichnung von Isomorphismen
Sei M ein freier R-Modul mit einer Basis der Länge n.

Für $f \in \mathrm{End}_R(M)$ sind folgende Aussagen äquivalent:

(a) *f ist ein Isomorphismus;*

(b) *f ist surjektiv;*

(c) *es gibt ein $g \in \mathrm{End}_R(M)$ mit $f \circ g = id_M$;*

(d) *es gibt ein $h \in \mathrm{End}_R(M)$ mit $h \circ f = id_M$;*

(e) $\det(f)$ *ist invertierbar in* R.

Es gilt dann $\det(f^{-1}) = \det(f)^{-1}$.
Ist R *ein Körper, so ist* f *genau dann injektiv, wenn* $\det(f) \neq 0$.

Beweis: Sei $B = (b_1, \ldots, b_n)$ eine Basis von M.

Offensichtlich impliziert (a) die Aussagen (b), (c) und (d).

(b)⇒(c) Sei f surjektiv. Wir definieren eine R-lineare Abbildung

$$g : M \to M, \quad b_i \mapsto c_i \text{ mit } c_i \in f^{-1}(b_i).$$

Damit gilt $f \circ g = id_M$.

(c)⇒(e) Nach (3) in 15.8 folgt aus $f \circ g = id_M$

$$1 = \det(id_M) = \det(f \circ g) = \det(f) \cdot \det(g).$$

Somit ist $\det(f)$ invertierbar.

(b)⇒(a) Ist f surjektiv, so ist – wie eben gezeigt – $\det(f)$ invertierbar. Nach 15.8,(2) ist dann f auch injektiv.

(d)⇒(e) sieht man analog zu (c)⇒(e).

(e)⇒(a) Ist $\det(f)$ invertierbar, so ist f injektiv (siehe 15.8,(2)). Es bleibt noch zu zeigen, daß f auch surjektiv ist. Ist R ein Körper, so folgt dies bereits aus 15.6, da dann $f(b_1), \ldots, f(b_n)$ eine Basis von M ist.

Über Ringen hilft uns 15.9 weiter. Danach gilt für die zu f adjunkte Matrix D, daß $\mathrm{Mat}(f) \cdot D = \det(f)E$. Somit ist $\overline{D} := \frac{1}{\det(f)} D$ eine rechtsinverse Matrix zu $\mathrm{Mat}_{BB}(f)$.

Zur gegebenen Basis B definiert $\bar{D}$ einen Endomorphismus $h : M \to M$ (vgl. 12.3), und nach Satz 12.6 gilt

$$\mathrm{Mat}_{BB}(h \circ f) = \mathrm{Mat}_{BB}(f) \cdot \mathrm{Mat}_{BB}(h) = E.$$

Also ist $h \circ f = \mathrm{id}_M$, und h ist surjektiv. Wie in (b)⇒(a) gezeigt, ist damit h bijektiv. Daraus folgt wiederum, daß auch f bijektiv ist.

Ist f invertierbar, so gilt

$$1 = \det(f^{-1} \circ f) = \det(f) \cdot \det(f^{-1}).$$

□

Nachdem uns die Bedeutung der Determinante klar geworden ist, wollen wir uns auch überlegen, wie wir ihren Wert feststellen können. Dies wird uns durch die Matrix eines Endomorphismus ermöglicht:

15.11 Berechnung der Determinante

Sei M ein freier R-Modul mit Basis $B = (b_1, \ldots, b_n)$. Ist $f \in \mathrm{End}_R(M)$ und $\mathrm{Mat}_{BB}(f) = (a_{ij})$, so gilt

$$\det(f) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn} \sigma \, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

Beweis: Sei $d : M^n \to R$ eine Determinantenform. Es gilt $f(b_i) = \sum_{j=1}^{n} a_{ij} b_j$, und mit der Berechnung in 15.3 folgt

$$\begin{aligned} \det(f) \cdot d(b_1, \ldots, b_n) &= d(f(b_1), \ldots, f(b_n)) \\ &= \sum_{\sigma \in \mathcal{S}_n} (\operatorname{sgn} \sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \cdot d(b_1, \ldots, b_n). \end{aligned}$$

Durch Kürzen von $d(b_1, \ldots, b_n)$ ergibt sich die gewünschte Beziehung.

Wie früher schon angemerkt, ist $\mathrm{Mat}_{BB}(f)$ von der Wahl der Basis abhängig, $\det(f)$ jedoch nicht. □

16 Die Determinante einer Matrix

In 15.11 haben wir gezeigt, wie die Determinante eines Endomorphismus' aus der zugehörigen Matrix (bezüglich irgendeiner Basis) berechnet werden kann. Wir nehmen dies zur Definition der Determinante einer beliebigen quadratischen Matrix.

Auch hier setzen wir wieder voraus, daß R ein kommutativer Ring ist.

16.1 Definition
Zu einer Matrix $A = (a_{ij}) \in R^{(n,n)}$ definieren wir als *Determinante*

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn} \sigma \; a_{1\sigma(1)} \dots a_{n\sigma(n)}.$$

Man benutzt auch die Notation $\det(A) = |A|$.

Fassen wir den R^n als Zeilenraum auf mit den kanonischen Basisvektoren $e_i = (0, \dots, 1, \dots, 0)$ (1 an i-ter Stelle), und betrachten die Determinantenform (vgl. 15.4)

$$d : (R^n)^n \to R \quad \text{mit} \;\; d(e_1, \dots, e_n) = 1.$$

Jede Matrix $A = (a_{ij}) \in R^{(n,n)}$ bestimmt eine R-lineare Abbildung

$$\widehat{A} : R^{(1,n)} \to R^{(1,n)}, \quad (r_1, \dots, r_n) \mapsto (r_1, \dots, r_n) \cdot (a_{ij}),$$

und $\operatorname{Mat}(\widehat{A}) = A$ (bezogen auf die kanonische Basis, vgl. 12.11).

Nach 15.11 gilt dann $\det(\widehat{A}) = \det(A)$, wobei die erste Determinante für Endomorphismen, die zweite Determinante für Matrizen (nach 16.1) gemeint ist. Wegen dieser Gleichheit brauchen wir die Notationen dafür nicht zu unterscheiden.

Die Bilder der e_i unter $\widehat{A}$ sind die Zeilen von A.

Durch die vorangegangenen Interpretationen von A können wir unmittelbar angeben:

16.2 Eigenschaften der Determinante von Matrizen
Sei $A = (a_{ij}) \in R^{(n,n)}$ eine Matrix. Dann gilt

(1) $\det(A) = \det(A^t)$.

(2) $\det(A)$ *ist Nullteiler, wenn die Zeilen (Spalten) von A linear abhängig sind.*

(3) *Für jede Matrix $B \in R^{(n,n)}$ ist*

$$\det(AB) = \det(A) \det(B).$$

Beweis: (1) Da R kommutativ ist, gilt für jedes $\tau \in \mathcal{S}_n$

$$a_{1\tau(1)} \cdots a_{n\tau(n)} = a_{1\tau^{-1}(1)} \cdots a_{n\tau^{-1}(n)}.$$

Damit erhalten wir nun

$$\begin{aligned} \det(A^t) &= \sum_{\tau \in \mathcal{S}_n} (\operatorname{sgn}\tau) a_{\tau(1),1} a_{\tau(2),2} \cdots a_{\tau(n),n} \\ &= \sum_{\tau^{-1} \in \mathcal{S}_n} (\operatorname{sgn}\tau^{-1}) a_{1\tau^{-1}(1)} \cdots a_{n\tau^{-1}(n)} \;=\; \det(A). \end{aligned}$$

(2) Dies folgt aus 15.6 (und (1)).

(3) Unter Benutzung von 15.8 und 12.6 haben wir

$$\begin{aligned} \det(AB) &= \det(\widehat{AB}) = \det(\widehat{A}) \det(\widehat{B}) \\ &= \det(\widehat{B} \circ \widehat{A}) = \det(\widehat{A}) \det(\widehat{B}) \\ &= \det(A) \det(B). \end{aligned}$$

□

Auch die Kennzeichnung von Isomorphismen in 15.10 können wir auf Matrizen übertragen:

16.3 Invertierbare Matrizen

Für jede Matrix $A \in R^{(n,n)}$ sind äquivalent:

(a) *A ist invertierbar;*

(b) *es gibt ein $B \in R^{(n,n)}$ mit $AB = E$;*

(c) *es gibt ein $B \in R^{(n,n)}$ mit $BA = E$;*

(d) $\det(A)$ *ist invertierbar in R.*

Es gilt dann $\det(A^{-1}) = \det(A)^{-1}$.

16.4 Berechnung von Determinanten

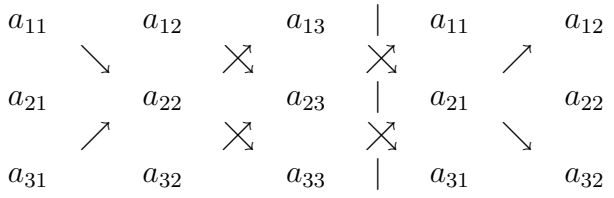
Die Definition der Determinante einer Matrix gibt auch die Möglichkeit ihrer Berechnung durch Permutationen und Summen von Koeffizienten.

Für $n = 2$ kann man dies leicht aufschreiben:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Regel von Sarrus

Für $n = 3$ kann man die Berechnung durch geschickte Notation recht übersichtlich gestalten. Man fügt Kopien der beiden ersten Spalten von rechts an die Matrix an, bildet die Summe der Produkte der Elemente in der Pfeilrichtung $\searrow$ und subtrahiert die Summe der entsprechenden Produkte in der Richtung $\nearrow$:



Schreiben wir das auch ausführlich auf:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$$

Bei $n \geq 4$ wird die Auswertung der definierenden Gleichung der Determinante deutlich mühsamer. Zum Glück gibt es verschiedene Methoden, sich die Berechnung zu erleichtern.

Berechnung nach Zeilenumformungen

Die Determinante von $A = (a_{ij})$ bleibt bei gewissen Zeilenumformungen unverändert. Da $\det : (R^n)^n \to R$ eine alternierende, n-lineare Abbildung in den Zeilen ist, gilt:

(i) Vertauschen von zwei Zeilen ändert das Vorzeichen der Determinante.

(ii) Addition einer Vielfachen einer Zeile zu einer anderen verändert den Wert der Determinante nicht.

Ist R ein Körper, so kann man auf diese Weise (a_{ij}) in eine Dreiecksmatrix überführen,

$$A' = \begin{pmatrix} a'_{11} & a'_{12} & \dots & * \\ 0 & a'_{22} & & \vdots \\ & & \ddots & \vdots \\ 0 & & & a'_{nn} \end{pmatrix},$$

die – eventuell bis auf das Vorzeichen – die gleiche Determinante hat:

$$\pm \det(A) = \det(A') = a'_{11} \cdots a'_{nn}.$$

Berechnung aus Teilmatrizen
Seien R ein Körper und $A \in R^{(n,n)}$ eine Matrix der Form

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

mit $B \in R^{(p,p)}$, $D \in R^{(p,n-p)}$ und $C \in R^{(n-p,n-p)}$. Dann gilt

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

Beweis: Ist die Determinante ungleich Null, so sind die Zeilen von B und C linear unabhängig. Somit lassen sich B und C durch Zeilenumformungen von A auf obere Dreiecksgestalt bringen. Dann ist $\det(A)$ das Produkt der Diagonalelemente dieser Matrix, und $\det(B)$, $\det(C)$ erhält man als Produkt der Diagonalelemente der entsprechenden Teilmatrizen. Damit ist die angegebene Formel klar. □

Allgemein kann die Berechnung der Determinante einer (n,n)-Matrix auf die Berechnung der Determinanten von $(n-1,n-1)$-Matrizen zurückgeführt werden. Erinnern wir uns dazu an die Definition der adjunkten Matrix $\mathrm{Ad}(f) = (d_{kl})$ zu einem Endomorphismus f eines freien R-Moduls M mit Basis $B = (b_1, \ldots, b_n)$ und einer Determinantenform $d : M^n \to R$ (siehe 15.9):

$$d_{kl} = d(f(b_1), \ldots, b_k, \ldots, f(b_n)), \quad b_k \text{ an } l\text{-ter Stelle.}$$

Für $M = R^n$ bestimmt eine Matrix $A = (a_{ij}) \in R^{(n,n)}$ den Endomorphismus

$$\widehat{A} : R^n \to R^n, \quad (a_1, \ldots, a_n) \mapsto (a_1, \ldots, a_n) \cdot A.$$

Wähle die Determinatenform $d : R^n \to R$ mit $d(e_1, \ldots, e_n) = 1$, wobei $(e_1, \ldots, e_n)$ die kanonische Basis von R^n bezeichne.

16.5 Definition
Als *Adjunkte* $\mathrm{Ad}(A) = (A_{kl})$ von A definieren wir die Adjunkte von $\widehat{A}$ (mit einer Determinantenform d), also

$$\begin{aligned} A_{kl} &= d(e_1 A, \ldots, e_k, \ldots, e_n A), \quad e_k \text{ an } l\text{-ter Stelle} \\ &= \det \begin{pmatrix} a_{1,1} & \ldots & a_{1,k-1} & a_{1,k} & a_{1,k+1} & \ldots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{l-1,1} & \ldots & a_{l-1,k-1} & a_{l-1,k} & a_{l-1,k+1} & \ldots & a_{l-1,n} \\ 0 & 0 & 0 & 1 & 0 & \ldots & 0 \\ a_{l+1,1} & \ldots & a_{l+1,k-1} & a_{l+1,k} & a_{l+1,k+1} & \ldots & a_{l+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \ldots & a_{n,k-1} & a_{n,k} & a_{n,k+1} & \ldots & a_{n,n} \end{pmatrix} \end{aligned}$$

$$= (-1)^{k+l} \det \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & a_{1,1} & \dots & a_{1,k-1} & a_{1,k+1} & \dots & a_{1,n} \\ & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{l-1,1} & \dots & a_{l-1,k-1} & a_{l-1,k+1} & \dots & a_{l-1,n} \\ 0 & a_{l+1,1} & \dots & a_{l+1,k-1} & a_{l+1,k+1} & \dots & a_{l+1,n} \\ & \vdots & & \vdots & \vdots & & \vdots \\ 0 & a_{n,1} & \dots & a_{n,k-1} & a_{n,k+1} & \dots & a_{n,n} \end{pmatrix}$$

Bis auf das Vorzeichen berechnen sich also die A_{kl} als Determinanten der $(n-1, n-1)$-Matrizen, die sich aus A durch Streichen der k-ten Spalte und l-ten Zeile ergeben.

Die Determinante von A läßt sich auf verschiedene Weise als Linearkombination dieser Größen angeben:

16.6 Laplacescher Entwicklungssatz
Sei $A = (a_{ij}) \in R^{(n,n)}$. Für jedes $l \leq n$ gilt

$$\det(A) = \sum_{k=1}^{n} a_{lk} \cdot A_{kl}.$$

Man nennt dies die *Entwicklung von* $\det(A)$ *nach der l-ten Zeile.*

Beweis: Die entsprechende Beziehung wurde im Beweis 15.9 für Endomorphismen gezeigt. Schreiben wir das Argument mit $l = 1$ nochmals in unserer Situation auf:

$$\det(A) = \det(A) \cdot d(e_1, \dots, e_n) = d(\sum_{k=1}^{n} a_{1k} e_k, e_2 A, \dots, e_n A) = \sum_{k=1}^{n} a_{1k}\, A_{k1}.$$

□

Mit den eingeführten Bezeichnungen folgt nun aus 15.9:

16.7 Satz
Für jede Matrix $A = (a_{ij}) \in R^{(n,n)}$ gilt

$$A \cdot \mathrm{Ad}(A) = \det(A) \cdot E_n,$$

wobei E_n die n-te Einheitsmatrix bezeichne.
Ist $\det(A)$ *invertierbar, dann gilt*

$$A^{-1} = \frac{1}{\det(A)} \cdot \mathrm{Ad}(A).$$

Für eine Matrix mit invertierbarer Determinante ergibt die oben angegebene Beziehung einen Weg zur Berechnung der inversen Matrix. Meist ist dies jedoch nicht die einfachste Möglichkeit, die Inverse zu finden.

Beispiel
Sehen wir uns die allgemeinen Ausführungen für $n = 3$ an. Sei

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \in R^{(3,3)}.$$

Die Entwicklung nach der zweiten Zeile ergibt:

$$\det(A) = -d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + e \begin{vmatrix} a & c \\ g & i \end{vmatrix} - f \begin{vmatrix} a & b \\ g & h \end{vmatrix}.$$

Entwickeln wir nach der dritten Zeile, so haben wir:

$$\det(A) = g \begin{vmatrix} b & c \\ e & f \end{vmatrix} - h \begin{vmatrix} a & c \\ d & f \end{vmatrix} + i \begin{vmatrix} a & b \\ d & e \end{vmatrix}.$$

Die Adjunkte zu A sieht so aus:

$$\mathrm{Ad}(A) = \begin{pmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ \begin{vmatrix} d & e \\ g & h \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{pmatrix}$$

Die Komponente $(2, 2)$ von $A\,\mathrm{Ad}(A)$ ergibt sich als

$$(A\,\mathrm{Ad}(A))_{22} = -d \begin{vmatrix} b & c \\ h & i \end{vmatrix} + e \begin{vmatrix} a & c \\ g & i \end{vmatrix} - f \begin{vmatrix} a & b \\ g & h \end{vmatrix} = \det(A).$$

Mit Hilfe der Determinante läßt sich auch die Lösung von *eindeutig* lösbaren Gleichungssystemen angeben. Wir wissen, daß dann die Matrix des Systems invertierbar, also deren Determinante ebenfalls invertierbar sein muß.

16.8 Cramersche Regel

Sei $A = (a_{ij}) \in R^{(n,n)}$ eine invertierbare Matrix und $b \in R^n$. Bezeichnen wir mit a_i die i-te Spalte von A, dann gilt für die Lösung $x = (x_1, \ldots, x_n)^t$ des Gleichungssystems $Ax = b$:

$$x_i = \frac{1}{\det(A)} \cdot \det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n).$$

Beweis: Ist $x = (x_1, \ldots, x_n)$ die Lösung des Systems, also

$$b = x_1 a_1 + \ldots + x_n a_n,$$

dann folgt aus den Eigenschaften der Determinante

$$\begin{aligned}
&\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n) \\
&\quad = \textstyle\sum_{k=1}^{n} x_k \det(a_1, \ldots, a_{i-1}, a_k, a_{i+1}, \ldots, a_n) \\
&\quad = x_i \cdot \det(A).
\end{aligned}$$

□

Sehen wir uns diese Lösungsmethode in einem konkreten Fall an.

Beispiel
Gegeben sei das lineare Gleichungssystem mit Koeffizienten aus $\mathbb{Q}$:

$$\begin{array}{rcrcrcr}
x_1 & + & 2x_2 & - & 3x_3 & = & 14 \\
2x_1 & - & x_2 & + & x_3 & = & -3 \\
-x_1 & + & 7x_2 & - & 2x_3 & = & 19
\end{array}$$

Nehmen wir die bestimmende Matrix heraus, und berechnen die Determinante:

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & -1 & 1 \\ -1 & 7 & -2 \end{pmatrix}, \qquad \det(A) = -38.$$

Somit ist das System eindeutig lösbar mit den Koeffizienten

$$\begin{aligned}
x_1 &= -\frac{1}{38} \begin{vmatrix} 14 & 2 & -3 \\ -3 & -1 & 1 \\ 19 & 7 & -2 \end{vmatrix} = 1; \\
x_2 &= -\frac{1}{38} \begin{vmatrix} 1 & 14 & -3 \\ 2 & -3 & 1 \\ -1 & 19 & -2 \end{vmatrix} = 2; \\
x_3 &= -\frac{1}{38} \begin{vmatrix} 1 & 2 & 14 \\ 2 & -1 & -3 \\ -1 & 7 & 19 \end{vmatrix} = -3.
\end{aligned}$$

Also haben wir die Lösung $(x_1, x_2, x_3) = (1, 2, -3)$.

16.9 Aufgaben

(1) Über einem kommutativen Ring R seien $A, B \in R^{(n,n)}$ gegeben. Zeigen Sie:

$$\det \begin{pmatrix} A & B \\ B & A \end{pmatrix} = \det(A+B) \cdot \det(A-B).$$

(2) (i) Berechnen Sie $\det \begin{pmatrix} 4 & -6 & 10 \\ 11 & 1 & \sqrt{7} \\ -1 & 1,5 & -3 \end{pmatrix}$ *mit der Regel von Sarrus.*

(ii) Berechnen Sie $\det \begin{pmatrix} 1 & 5 & 14 & 15 \\ 2 & 11 & 30 & 34 \\ 3 & 12 & 37 & 35 \\ 4 & 13 & 45 & 36 \end{pmatrix}$.

(3) Für $(x, y, z) \in \mathbb{R}^3$ sei definiert

$$A_{(x,y,z)} := \begin{pmatrix} 1 & 5 & 6 \\ x & y & z \\ 4 & 2 & 3 \end{pmatrix} \in \mathbb{R}^{(3,3)}.$$

Weiter sei

$$U := \{(x, y, z) \in \mathbb{R}^3 \mid A_{(x,y,z)} \text{ ist } \textbf{nicht} \text{ invertierbar}\}.$$

(a) *Man zeige, daß U ein Unterraum von $\mathbb{R}^3$ ist und gebe eine Basis von U an.*

(b) *Für $(x, y, z) \in \mathbb{R}^3 \setminus U$ berechne man $A^{-1}_{(x,y,z)}$.*

(4) Sei R ein kommutativer Ring, $n \in \mathbb{N}$. Zu $b_1, \ldots, b_n \in R$ bilden wir die (n, n)-Matrizen

$$B_n := \begin{pmatrix} 1 & b_1 & b_1^2 & \ldots & b_1^{n-1} \\ 1 & b_2 & b_2^2 & \ldots & b_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & b_n & b_n^2 & \ldots & b_n^{n-1} \end{pmatrix}.$$

*Man zeige (*Vandermondesche Determinante*):*

$$\det B_n = \prod_{1 \le i < k \le n} (b_k - b_i).$$

17 Die Spur von Endomorphismen und Matrizen

Ähnlich wie die Determinante, lassen sich den Endomorphismen von freien Moduln mit Hilfe von Determinantenformen noch andere Größen zuordnen. Eine davon wollen wir in diesem Abschnitt einführen.

Wir werden wieder voraussetzen, daß R ein kommutativer Ring ist.

M sei ein freier R-Modul mit Basis $(b_1, \ldots, b_n)$ und $d : M^n \to R$ eine Determinantenform mit $d(b_1, \ldots, b_n) = 1$.

Zu einem Endomorphismus $f : M \to M$ haben wir die alternierende, multilineare Abbildung

$$M^n \xrightarrow{f^n} M^n \xrightarrow{d} R$$

gebildet. Nach Satz 15.5 ist dies ein Vielfaches der Determinantenform d, und $d \circ f^n = \det(f) \cdot d$ war die definierende Gleichung für die Determinante.

Wir geben nun zu f eine andere alternierende, n-lineare Abbildung $M^n \to R$ an. Auch diese ist nach Satz 15.5 ein Vielfaches der Determinantenform d, und wir werden uns für den auftretenden Faktor interessieren.

17.1 Hilfssatz
Mit den oben festgelegten Notationen ist die Zuordnung

$$\sigma_{f,d} : M^n \to R, \quad (a_1, \ldots, a_n) \longmapsto \sum_{i=1}^{n} d\Big(a_1, \ldots, f(a_i), \ldots, a_n\Big),$$

eine multilineare und alternierende Abbildung.

Beweis: Sehen wir uns das Verhalten beim Vertauschen der ersten beiden Argumente an:

$$\begin{aligned}
&\sigma_{f,d}(a_1, a_2, \ldots, a_n) + \sigma_{f,d}(a_2, a_1, \ldots, a_n) \\
&\quad = \sum_{i=1}^{n} \Big(d(a_1, a_2, \ldots, f(a_i), \ldots, a_n) + d(a_2, a_1, \ldots, f(a_i), \ldots, a_n)\Big) \\
&\quad = d(f(a_1), a_2, \ldots, a_n) + d(f(a_2), a_1, \ldots, a_n) \\
&\qquad + d(a_1, f(a_2), \ldots, a_n) + d(a_2, f(a_1), \ldots, a_n) \;=\; 0.
\end{aligned}$$

Die gleichen Umformungen lassen sich an den anderen Stellen durchführen. Auch ist leicht zu sehen, daß man Skalare aus jeder Komponente nach vorne ziehen kann. □

Als alternierende Multilinearform ist $\sigma_{f,d}$ ein skalares Vielfaches von d. Damit legen wir fest:

17.2 Definition *(Spur von Endomorphismen)*
Seien M ein freier R-Modul mit einer Basis der Länge n, $d : M^n \to R$ eine Determinantenform und $f : M \to M$ ein Endomorphismus.

Das Element $\mathrm{Sp}(f) \in R$ mit der Eigenschaft

$$\sigma_{f,d} = \mathrm{Sp}(f) \cdot d$$

bezeichnet man als *Spur von* f.

Für eine Basis $b_1, \ldots, b_n$ von M gilt somit

$$\mathrm{Sp}(f) = \frac{\sum\limits_{i=1}^{n} d\big(b_1, \ldots, f(b_i), \ldots, b_n\big)}{d(b_1, \ldots, b_n)} \qquad (*)$$

Jede andere Determinantenform hat die Gestalt $d' = \gamma \cdot d$ mit invertierbarem $\gamma \in R$. Daher ist $\mathrm{Sp}(f)$ unabhängig von der Wahl der Form d.

Aus der angegebenen Beziehung kann man direkt ablesen:

17.3 Eigenschaften der Spur
Sei M ein freier R-Modul mit Basis $B = (b_1, \ldots, b_n)$. Dann gilt:

(1) $\mathrm{Sp}(0) = 0$ *und* $\mathrm{Sp}(\mathrm{id}_M) = n \cdot 1$ $(1 \in R)$.

(2) *Für $f, g \in \mathrm{End}(M)$ und $r \in R$ gilt*

$$\mathrm{Sp}(rf) = r\,\mathrm{Sp}(f) \quad \textit{und} \quad \mathrm{Sp}(f+g) = \mathrm{Sp}(f) + \mathrm{Sp}(g).$$

Somit ist $\mathrm{Sp} : \mathrm{End}_R(M) \to R$ *ein R-Modulhomomorphismus.*

(3) *Für $(a_{ij}) := \mathrm{Mat}_{BB}(f)$ gilt*

$$\mathrm{Sp}(f) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Beweis: (1) und (2) lassen sich direkt aus $(*)$ ablesen.

(3) Nach Definition von (a_{ij}) gilt $f(b_i) = \sum_{j=1}^{n} a_{ij} b_j$ und damit

$$\sum_{i=1}^{n} d(b_1, \ldots, \sum_{j=1}^{n} a_{ij} b_j, \ldots, b_n) = (a_{11} + a_{22} + \cdots + a_{nn}) d(b_1, \ldots, b_n).$$

Damit folgt die Behauptung ebenfalls aus $(*)$. □

Ähnlich wie bei den Determinanten nehmen wir die Beziehung (3) zur

17.4 Definition *(Spur einer Matrix)*
Als *Spur einer Matrix* $A = (a_{ij}) \in R^{(n,n)}$ bezeichnen wir

$$\mathrm{Sp}(A) = \sum_{i=1}^{n} a_{ii}.$$

Die Spur ist also die Summe der Elemente der Hauptdiagonale.

Mit diesen Begriffsbildungen stellen wir fest:

17.5 Folgerung
Sei M ein freier R-Modul mit Basis $B = \{b_1, \ldots, b_n\}$ und $f : M \to M$ ein Endomorphismus mit $A := \mathrm{Mat}_{B,B}(f) \in R^{(n,n)}$, dann gilt

$$\mathrm{Sp}(f) = \mathrm{Sp}(A).$$

Die Spur einer Matrix läßt sich also sehr leicht berechnen. Daher lassen sich auch Eigenschaften der Spur von Endomorphismen einfach durch entsprechende Rechnungen in Matrizen zeigen.

17.6 Eigenschaften der Spur einer Matrix
Für $A, B, C \in R^{(n,n)}$, $r \in R$ und $E_n \in R^{(n,n)}$ (Einheitsmatrix) gilt:

(1) $\mathrm{Sp}(0) = 0$ *und* $\mathrm{Sp}(E_n) = n \cdot 1_R$.

(2) $\mathrm{Sp}(rA) = r \cdot \mathrm{Sp}(A)$ *und* $\mathrm{Sp}(A + B) = \mathrm{Sp}(A) + \mathrm{Sp}(B)$.
Also ist $\mathrm{Sp} : R^{(n,n)} \to R$ *ein R-Modulhomomorphismus.*

(3) $\mathrm{Sp}(AB) = \mathrm{Sp}(BA)$ *und* $\mathrm{Sp}(ABC) = \mathrm{Sp}(CAB) = \mathrm{Sp}(BCA)$.

(4) *Ist B invertierbar, so ist* $\mathrm{Sp}(B^{-1}AB) = \mathrm{Sp}(A)$.

(5) $\mathrm{Sp}(A) = \mathrm{Sp}(A^t)$.

Beweis: Die in (1) und (2) angegebenen Beziehungen kennen wir schon für Endomorphismen. Sie lassen sich ohne Mühe direkt nachrechnen.

(3) Mit $A = (a_{ij})$, $B = (b_{ij})$ ergibt sich die erste Gleichung aus

$$\mathrm{Sp}(AB) = \sum_{i=1}^{n}(\sum_{k=1}^{n} a_{ik}\, b_{ki}) = \sum_{k=1}^{n}\left(\sum_{i=1}^{n} b_{ki}\, a_{ik}\right) = \mathrm{Sp}(BA).$$

Die zweite Gleichung erhält man aus der Assoziativität der Multiplikation von Matrizen und wiederholte Anwendung der ersten Identität.

(4) ist direkte Folge von (3).

(5) ist klar, da die Diagonalen von A und A^t übereinstimmen. □

Diese Beobachtungen können wir auch für Endomorphismen festhalten:

17.7 Eigenschaften der Spur von Endomorphismen
Sei M ein freier R-Modul mit endlicher Basis. Für $f, g, h \in \mathrm{End}_R(M)$ gilt:

(1) $\mathrm{Sp}(g \circ f) = \mathrm{Sp}(f \circ g)$.

(2) $\mathrm{Sp}(h \circ g \circ f) = \mathrm{Sp}(g \circ f \circ h) = \mathrm{Sp}(f \circ h \circ g)$.

(3) *Ist g Isomorphismus, so ist* $\mathrm{Sp}(g \circ f \circ g^{-1}) = \mathrm{Sp}(f)$.

Bemerkungen

(1) $\mathrm{Sp}(A) = 0$ kann auch für invertierbare Matrizen gelten, z.B. für $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ oder $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(2) Zeilenumformungen, die die Determinante unverändert lassen, können sehr wohl die Spur verändern.

(3) Für einen endlich dimensionalen $\mathbb{Q}$-Vektorraum V und $f, g \in \mathrm{End}_{\mathbb{Q}}(V)$, kann *nicht*

$$g \circ f - f \circ g = r \cdot \mathrm{id}$$

gelten, da

$$0 = \mathrm{Sp}(g \circ f - f \circ g) \neq r \cdot \mathrm{Sp}(\mathrm{id}) = r \cdot n.$$

Dies ist bemerkenswert, da in anderen Räumen solche Beziehungen vorkommen (z.B. Quantentheorie, Heisenberg'sche Unschärferelation).

Wir haben die Spur als lineare Abbildung $\mathrm{End}(M) \to R$ (bzw. $R^{(n,n)} \to R$) kennengelernt. Lineare Abbildungen in den Grundring nennt man *Linearformen*. Näheres dazu werden wir später ausführen. Hier wollen wir festhalten, daß die Spur dabei eine Sonderrolle einnimmt:

17.8 Spur und Linearformen auf $\mathrm{End}_R(M)$
Sei M ein freier R-Modul mit Basis der Länge n.

(1) *Zu jeder Linearform $\alpha : \mathrm{End}_R(M) \to R$ gibt es ein $h \in \mathrm{End}_R(M)$ mit*

$$\alpha(g) = \mathrm{Sp}(h \circ g) = 0 \text{ für alle } g \in \mathrm{End}_R(M).$$

(2) *Sei $f \in \mathrm{End}_R(M)$. Ist $\mathrm{Sp}(f \circ g) = 0$ für alle $g \in \mathrm{End}_R(M)$, so ist $f = 0$.*

(3) *Ist $\alpha : \mathrm{End}_R(M) \to R$ eine Linearform mit $\alpha(f \circ g) = \alpha(g \circ f)$ für alle $f, g \in \mathrm{End}_R(M)$, dann gibt es (genau) ein*

$$r \in R \text{ mit } \alpha(f) = r \cdot \mathrm{Sp}(f) \text{ für alle } f \in \mathrm{End}_R(M).$$

Die obigen Zusammenhänge können wir auch für Matrizen formulieren, und wir werden den Beweis dafür angeben:

17.9 Spur und Linearformen auf $R^{(n,n)}$

(1) *Zu jeder Linearform* $\alpha : R^{(n,n)} \to R$ *gibt es ein* $H \in R^{(n,n)}$ *mit*

$$\alpha(A) = \mathrm{Sp}(H \cdot A) = 0 \text{ für alle } A \in R^{(n,n)}.$$

(2) *Sei* $A \in R^{(n,n)}$. *Ist* $\mathrm{Sp}(AB) = 0$ *für alle* $B \in R^{(n,n)}$, *so ist* $A = 0$.

(3) *Ist* $\alpha : R^{(n,n)} \to R$ *eine Linearform mit* $\alpha(AB) = \alpha(BA)$ *für alle* $A, B \in R^{(n,n)}$, *dann gibt es (genau) ein*

$$r \in R \text{ mit } \alpha(A) = r \cdot \mathrm{Sp}(A) \text{ für alle } A \in R^{(n,n)}.$$

Beweis: (1) Da α eine lineare Abbildung ist, genügt es, die Aussage auf der Basis E_{ij} von $R^{(n,n)}$ zu verifizieren. Für die gewünschte Matrix $H = (h_{ij})$ muß also gelten

$$\alpha(E_{ij}) = \mathrm{Sp}(HE_{ij}) = \mathrm{Sp}(\sum_{k,l} h_{kl}E_{kl}E_{ij}) = h_{ji},$$

da in der Diagonale der Matrix HE_{ij} nur das Element h_{ji} von Null verschieden sein kann (vgl. 12.8).

Somit erfüllt (nur) die Matrix $H = (\alpha(E_{ij}))$ die gewünschte Bedingung.

(2) Die Behauptung ergibt sich aus (1), da nur die Nullmatrix der Nullform $R^{(n,n)} \to R$ entspricht.

(3) Sei $\alpha(AB) = \alpha(BA)$ für alle $A, B \in R^{(n,n)}$ und $H \in R^{(n,n)}$ wie in (1). Aus den Eigenschaften der Spur ergibt sich damit

$$\mathrm{Sp}(BHA) = \mathrm{Sp}(HAB) = \alpha(AB) = \alpha(BA) = \mathrm{Sp}(HBA),$$

und damit $\mathrm{Sp}((HB - BH)A) = 0$ für alle $A, B \in R^{(n,n)}$.

Nach (2) bedeutet das $HB - BH = 0$ für alle $B \in R^{(n,n)}$. Nach 12.8 ist dann $B = rE$ für ein geeignetes $r \in R$.

Aus (1) ergibt sich nun $\alpha(A) = r \cdot \mathrm{Sp}(A)$ für alle $A \in R^{(n,n)}$. □

17.10 Aufgaben

(1) Sei $A = (a_{ij}) \in R^{(n,n)}$. *Zeigen Sie* $\mathrm{Sp}(AA^t) = \sum_{i=1}^{m} \sum_{j=i}^{n} (a_{ij})^2$.

(2) Zeigen Sie für $A \in R^{(2,2)}$: $A^2 = 0$ *genau dann, wenn* $\mathrm{Sp}(A) = \mathrm{Sp}(A^2) = 0$.

(3) Betrachten Sie die Linearform

$$\alpha : R^{(4,4)} \to R, \quad (a_{ij}) \mapsto a_{11} + \cdots + a_{14}.$$

Bestimmen Sie eine Matrix $H \in R^{(4,4)}$ *mit* $\alpha = \mathrm{Sp}(H-)$.

Kapitel 6

Eigenwerte und Jordansche Normalform

In diesem Kapitel werden wir uns mit der Feinstruktur von Endomorphismen und Matrizen befassen. Es wird unter anderem darum gehen, durch Wahl einer geeigneten Basis eine vorteilhafte Form der einem Endomorphismus zugeordneten Matrix zu finden (etwa Diagonalform, Dreiecksform). Dazu suchen wir Unterräume, die von einem gegebenen Endomorphismus in sich übergeführt werden.

18 Eigenwerte und Eigenvektoren

Sei K ein Körper. Beschäftigen wir uns zunächst mit den Elementen, die von einem Endomorphismus eines K-Vektorraums V in ein Vielfaches von sich selbst abgebildet werden.

18.1 Definition
Sei V ein K-Vektorraum und $f : V \to V$ ein Endomorphismus.
Ein Element $0 \neq m \in V$ heißt *Eigenvektor* von f, wenn

$$f(m) = rm \quad \text{für ein } r \in R.$$

Man nennt dann r einen *Eigenwert* von f.

Für jedes $s \in R$ bezeichnet man den Unterraum

$$E(s) = \{n \in V \mid f(n) = sn\} = \operatorname{Kern}(f - s\,\mathrm{id}) \subset V$$

als *Eigenraum* von f zu s.

Man beachte, daß man den Nullvektor $0 \in V$ nicht Eigenvektor nennt, obwohl er formal die definierende Gleichung erfüllt. Dagegen kann $0 \in K$ durchaus ein Eigenwert sein. So ist $s \in K$ genau dann Eigenwert von f, wenn $E(s) \neq 0$.

$E(0) = \operatorname{Kern} f$ und $E(1) = \{n \in V \mid f(n) = n\}$ (*Fixelemente*).

18.2 Definition
Sei $f \in \mathrm{End}_K(V)$. Ein Untervektorraum $U \subset V$ heißt *f-invariant* oder *f-stabil*, wenn $f(U) \subset U$.

Es ist offensichtlich, daß für $f \in \mathrm{End}_K(V)$ alle Eigenräume dazu f-invariant sind. Sie haben noch weitere wichtige Eigenschaften:

18.3 Satz
Sind $r_1, \ldots, r_n \in K$ paarweise verschiedene Eigenwerte von $f \in \mathrm{End}_K(V)$, dann bilden die Eigenräume $E(r_1) \ldots, E(r_n)$ von f eine unabhängige Familie von Untermoduln (vgl. 8.9), also

$$\sum_{i=1}^{n} E(r_i) = E(r_1) \oplus \cdots \oplus E(r_n).$$

Beweis: Zu zeigen ist, daß aus

$$(*) \quad 0 = \sum_{i=1}^{n} z_i, \quad z_i \in E(r_i),$$

stets $z_i = 0$ folgt. Wir beweisen dies durch vollständige Induktion über n. Dabei ist $n = 1$ klar.

Nehmen wir an, die Behauptung sei für $n-1$ richtig. Aus $(*)$ folgt

$$0 = f(0) = \sum_{i=1}^{n} f(z_i) = \sum_{i=1}^{n} r_i z_i$$

und damit

$$0 = \sum_{i=1}^{n-1} (r_i - r_n) z_i.$$

Wegen $(r_i - r_n)z_i \in E(r_i)$ folgt aus der Induktionsannahme $(r_i - r_n)z_i = 0$ für $i \leq n-1$, also $z_i = 0$, da nach Voraussetzung $r_i - r_n \neq 0$. □

Aus Dimensionsbetrachtungen erhalten wir das

18.4 Korollar
Ist V ein endlich dimensionaler Vektorraum, so hat jedes $f \in \mathrm{End}_K(V)$ höchstens $\mathrm{Rang}\, f$ *($\leq \dim V$) verschiedene Eigenwerte.*

Beispiel
Zu einem Endomorphismus braucht es keine Eigenwerte zu geben. Betrachten wir dazu

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad (1,0) \mapsto (0,1),\ (0,1) \mapsto (-1,0).$$

Angenommen, $k \in \mathbb{R}$ ist ein Eigenwert mit Eigenvektor (r_1, r_2), also

$$k(r_1, r_2) = f(r_1, r_2) = (-r_2, r_1).$$

Dies bedeutet $r_2 = -kr_1$, $r_1 = kr_2$ und damit $r_1^2 + r_2^2 = 0$. Also gilt $E(k) = 0$ und k ist kein Eigenwert.

Endomorphismen mit möglichst vielen Eigenvektoren lassen sich folgendermaßen kennzeichnen:

18.5 Satz
Sei V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}(V)$.
Gibt es zu f genau n verschiedene Eigenwerte $r_1, \ldots, r_n$, und bezeichnet b_i jeweils den Eigenvektor zu r_i, dann gilt:

(1) *$E(r_i) = Kb_i$ (also hat jeder Eigenraum Dimension 1).*

(2) *$V = \bigoplus_{i=1}^{n} Kb_i$, d.h. $B = (b_1, \ldots, b_n)$ ist eine Basis.*

(3) $\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} r_1 & & 0 \\ & r_2 & \\ & & \ddots \\ 0 & & r_n \end{pmatrix}$.

Beweis: (1), (2) Nach Voraussetzung hat jeder Eigenraum $E(r_i)$ mindestens Dimension 1, und somit gilt $\dim(\bigoplus_{i=1}^{n} E(r_i)) \geq n$.

Nach 18.3 bedeutet dies $\bigoplus_{i=1}^{n} E(r_i) = V$ und $\dim E(r_i) = 1$.

(3) Da $f(b_i) = r_i b_i$ für alle $i \leq n$, folgt die Behauptung aus der Definition von $\mathrm{Mat}_{B,B}(f)$. □

18.6 Definition
Ein Endomorphismus $f \in \mathrm{End}_K(V)$ heißt *diagonalisierbar*, wenn bezüglich einer geeigneten Basis B von V die Matrix $\mathrm{Mat}_{B,B}(f)$ Diagonalform hat.

Eine Matrix $A \in R^{(n,n)}$ nennt man *diagonalisierbar*, wenn es eine invertierbare Matrix $C \in K^{(n,n)}$ gibt, so daß $C^{-1}AC$ Diagonalform hat (also wenn A zu einer Diagonalmatrix ähnlich ist).

Als Beispiele für diagonalisierbare Endomorphismen haben wir in 18.5 die Endomorphismen mit genügend vielen verschiedenen Eigenwerten kennengelernt.

Nach den gegebenen Definitionen ist klar, daß $f \in \mathrm{End}_K(K)$ genau dann diagonalisierbar ist, wenn dies für $\mathrm{Mat}_{B,B}(f)$ (mit beliebiger Basis B) gilt.

Bemerkung: Die Definitionen dieses Abschnitts sind offensichtlich auch sinnvoll für Endomorphismen von freien Moduln und Matrizen über kommutativen Ringen. Jedoch ist etwa die Aussage in 18.5 (2) in dieser allgemeinen Situation nicht mehr gültig, da der Schluß mit der Dimension nicht mehr möglich ist.

Sehen wir uns zum Beispiel folgenden Homomorphismus an:

$$f : \mathbb{Z}^2 \to \mathbb{Z}^2, \quad (1,0) \mapsto (1,0),\ (0,1) \mapsto (1,-1).$$

Es gilt $f^2 = \mathrm{id}$, und f ist ein Isomorphismus. Als Eigenwerte findet man 1 und -1 mit den Eigenräumen

$$E(1) = \mathbb{Z}(1,0), \quad E(-1) = \mathbb{Z}(1,-2).$$

Nun kann aber $(0,1)$ nicht als $\mathbb{Z}$-Linearkombination von $(1,0)$ und $(1,-2)$ dargestellt werden. Daher gilt $E(1) \oplus E(-1) \neq \mathbb{Z}^2$.

Wir haben gesehen, daß $r \in K$ genau dann Eigenwert für $f \in \mathrm{End}_K(V)$ ist, wenn $\mathrm{Kern}(f - r\,\mathrm{id}) \neq 0$, d.h. wenn $f - r\,\mathrm{id}$ kein Isomorphismus ist. Wir suchen also nach solchen $r \in K$, für die $\det(f - r\,\mathrm{id}) = 0$ gilt. Dieses Problem wird im nächsten Abschnitt angegangen.

19 Das charakteristische Polynom

Sei R ein kommutativer Ring, M ein freier R-Modul mit Basis $(b_1, \ldots, b_n)$ und $f \in \mathrm{End}_R(M)$. Wie können wir die $r \in R$ finden mit $\det(f - r\,\mathrm{id}) = 0$?

Betrachten wir die Matrix $A = (a_{ij}) := \mathrm{Mat}_{B,B}(f) \in R^{(n,n)}$.

Dann ist $\det(r\,\mathrm{id} - f) = 0$ genau dann, wenn $\det(rE - A) = 0$.

Setzen wir für r eine Unbestimmte X und bilden

$$XE - A = \begin{pmatrix} X - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & X - a_{22} & \ldots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \ldots & X - a_{nn} \end{pmatrix} \in R[X]^{(n,n)}.$$

Da $R[X]$ ein kommutativer Ring ist, können wir davon die Determinante als ein Element aus $R[X]$ bestimmen.

19.1 Definition
Zu einer Matrix $A = (a_{ij}) \in R^{(n,n)}$ bilden wir

$$\mathrm{ch}(A) = \det(XE - A) \in R[X]$$

und nennen dies das *charakteristische Polynom* von A.

Schreiben wir $\mathrm{ch}(A)$ *als Linearkombination der Potenzen von* X*,*

$$\mathrm{ch}(A) = X^n + c_1 X^{n-1} + \ldots + c_{n-1}X + c_n,$$

so haben wir als zweiten und letzten Koeffizienten

$$c_1 = -\,\mathrm{Sp}(A), \quad c_n = (-1)^n \det(A).$$

Beweis: Nach der Determinantenformel gilt

$$\mathrm{ch}(A) = \prod_{i=1}^{n}(X - a_{ii}) + \underbrace{\sum_{\sigma \in \mathcal{S}_n \setminus \mathrm{id}} (\ldots)}_{Q}.$$

Dabei ist Q vom Grad $\leq n-2$, da für $\sigma \neq \mathrm{id}$ mindestens für *zwei* verschiedene i gilt, daß $i \neq \sigma(i)$. Also sind mindestens zwei der Faktoren nicht aus der Diagonale und

$$\mathrm{ch}(A) = X^n + (-1)^n(a_{11} + \ldots + a_{nn})X^{n-1} + Q',$$

mit $\deg Q' \leq n-2$. Ersetzt man X durch 0, so ergibt sich $c_n = (-1)^n \det(A)$. □

Eigenschaften von $\mathrm{ch}(\boldsymbol{A})$
Für eine Matrix $A \in R^{(n,n)}$ *gilt:*

(1) $\mathrm{ch}(A) = \mathrm{ch}(A^t)$.

(2) *Ist* $T \in R^{(n,n)}$ *invertierbar, so ist* $\mathrm{ch}(A) = \mathrm{ch}(TAT^{-1})$. *Somit haben ähnliche Matrizen das gleiche charakteristische Polynom.*

Beweis: (1) Da transponierte Matrizen die gleiche Determinante haben, gilt

$$\begin{aligned} \mathrm{ch}(A^t) &= \det(XE - A^t) = \det(XE - A)^t \\ &= \det(XE - A) = \mathrm{ch}(A). \end{aligned}$$

(2) Für invertierbares T erhalten wir $XE - TAT^{-1} = T(XE - A)T^{-1}$ und somit

$$\mathrm{ch}(TAT^{-1}) = \det(XE - TAT^{-1}) = \det(XE - A) = \mathrm{ch}(A).$$

□

Die Bildung des charakteristischen Polynoms ist also invariant gegenüber Basistransformationen und somit durch einen Endomorphismus eindeutig bestimmt.

19.2 Definition
Sei M ein freier R-Modul mit Basis $B = (b_1, \ldots, b_n)$.

Für einen Endomorphismus $f \in \mathrm{End}_R(M)$ nennen wir

$$\mathrm{ch}(f) = \mathrm{ch}(\mathrm{Mat}_{B,B}(f))$$

das *charakteristische Polynom von* f.

Nach den obigen Betrachtungen ist diese Definition unabhängig von der Wahl der Basis B. Auf die Bedeutung von $\mathrm{ch}(f)$ für Bestimmung und Eigenschaften der Eigenwerte und Eigenräume werden wir später zurückkommen.

Eine bemerkenswerte Eigenschaft von $\mathrm{ch}(A)$ ist der

19.3 Satz von Cayley-Hamilton
Jede Matrix $A \in R^{(n,n)}$ *ist Nullstelle ihres charakteristischen Polynoms,*

$$\mathrm{ch}(A)|_{X=A} = A^n - \mathrm{Sp}(A)A^{n-1} + \ldots (-1)^n \det(A) = 0.$$

Jeder Endomorphismus f *eines* R*-Moduls mit endlicher Basis ist Nullstelle seines charakteristischen Polynoms,*

$$\mathrm{ch}(f)|_{X=f} = f^n - \mathrm{Sp}(f)f^{n-1} + \ldots (-1)^n \det(f) = 0.$$

Beweis: Es genügt, die erste Behauptung zu beweisen.

Man könnte dazu verleitet sein, das Problem durch Einsetzen

$$\mathrm{ch}(A)|_{X=A} = \det(XE - A)|_{X=A} = \det(AE - A) = \det(0) = 0$$

zu lösen. Dies ist nicht zulässig, da die Matrix A ja nicht im Grundring R liegt. Wir wollen daher das Problem über einem Ring behandeln, der A enthält.

Bezeichne E die (n,n)-Einheitsmatrix, und betrachte den Untermodul

$$R[A] = \Big\{ \sum_{i=0}^{k} r_i A^i \mid k \in \mathbb{N}, r_i \in R \Big\} \subset R^{(n,n)}.$$

Dies ist ein kommutativer Unterring von $R^{(n,n)}$ mit Einselement $A^0 = E$. Vermöge der Einbettung $R \to R[A]$, $r \mapsto rE$, fassen wir R als Unterring von $R[A]$ auf und betrachten die Matrix

$$\widetilde{A} = (a_{ij}E) \in R^{(n^2,n^2)}.$$

Damit wird $\mathrm{ch}(\widetilde{A})$ ein Polynom aus $R[A][X]$, das wir mit $\mathrm{ch}(A) \in R[X]$ identifizieren dürfen (genauer: $\mathrm{ch}(\widetilde{A}) = \mathrm{ch}(A)E$). Es hat den Wert

$$\det(XE - \widetilde{A}) = \det\Big(\begin{pmatrix} X & & 0 \\ & \ddots & \\ 0 & & X \end{pmatrix} - \widetilde{A} \Big) = \det \begin{pmatrix} X - a_{11}E & \cdots & -a_{1n}E \\ -a_{21}E & \ddots & -a_{2n}E \\ \vdots & & \\ -a_{n1}E & \cdots & X - a_{nn}E \end{pmatrix}.$$

Da nun A Element im Grundring $R[A]$ ist, dürfen wir $X = A$ setzen, und die Determinante der entstehenden Matrix – oder der transponierten davon – soll gleich Null sein, also etwa der Matrix

$$C := (c_{ij}) = \begin{pmatrix} A - a_{11}E & -a_{21}E & \cdots & -a_{n1}E \\ -a_{12}E & A - a_{22}E & & -a_{n2}E \\ \vdots & & \ddots & \\ -a_{1n}E & \cdots & & A - a_{nn}E \end{pmatrix}.$$

Mit den kanonischen Einheitsvektoren (Zeilen) $e_1, \ldots, e_n$ gilt für $i \le n$,

$$e_i A = \sum_{j=1}^{n} e_j a_{ij}, \quad \text{also} \quad e_i(A - a_{ii}) - \sum_{j \neq i} e_j a_{ij} = (0, \ldots, 0).$$

Für die oben definierte Matrix bedeutet dies

$$\sum_{j=1}^{n} e_j c_{ji} = (0, \ldots, 0) \text{ für } i = 1, \ldots, n.$$

Von der zu C adjunkten Matrix $\mathrm{Ad}(C) = (C_{ij})$ wissen wir (vgl. 15.9, 16.7)

$$\det(C)\delta_{jk} = \sum_{i=1}^{n} c_{ji}C_{ik} \quad \text{für } j,k \leq n.$$

Durch Multiplikation der oben stehenden Beziehung mit C_{ik} und Summieren über i erhalten wir für $k \leq n$

$$(0,\ldots,0) = \sum_{i=1}^{n}\Big(\sum_{j=1}^{n} e_j c_{ji}\Big)C_{ik} = \sum_{j=1}^{n} e_j\Big(\sum_{i=1}^{n} c_{ji}C_{ik}\Big) = \sum_{j=1}^{n} e_j \det(C)\delta_{jk}.$$

Damit ist $e_k \det(C) = (0,\ldots,0)$ für alle $k \leq n$ und auch $\det(C) = 0$. □

Beispiele

(1) Für $A = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$ ist $\mathrm{ch}(A,X) = X^n$ und somit $A^n = 0$.

(2) Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist $A^2 - (a+d)A + (ad-bc)E = 0$.
(vgl. 12.19, Aufgabe 1).

19.4 Aufgaben

Für eine Matrix $A \in R^{(n,n)}$ sei das charakteristische Polynom

$$\mathrm{ch}(A) = X^n + c_1X^{n-1} + \ldots + c_n \in R[X].$$

Zeigen Sie für invertierbare A, daß

$$\mathrm{Ad}(A) = (-1)^{n-1}(A^{n-1} + c_1A^{n-2} + \ldots + c_{n-2}A + c_{n-1}E).$$

Damit ist die Adjunkte (und die Inverse) Linearkombination von Potenzen $\leq n-1$ von A.

20 Dreiecksform von Matrizen

Wir waren bei der Suche nach Eigenwerten einer Matrix auf das charakteristische Polynom aufmerksam geworden. Diese Wechselbeziehung wollen wir noch weiter analysieren. Da es dabei um Nullstellen von Polynomen geht, werden wir die Betrachtungen über einem Körper K durchführen.

Zunächst sei an den Beginn von Abschnitt 19 erinnert, wo wir festgestellt hatten:

20.1 Eigenwerte – charakteristisches Polynom
Sei V ein endlich dimensionaler K-Vektorraum. Dann sind die Eigenwerte von $f \in \mathrm{End}_K(V)$ genau die Nullstellen des charakteristischen Polynoms $\mathrm{ch}(f)$.

Die Existenz von Eigenwerten zu f hängt also auch von Eigenschaften des Körpers K ab: Gibt es in K Nullstellen von $\mathrm{ch}(f)$ oder nicht? Wir wissen zum Beispiel von den komplexen Zahlen, daß jedes Polynom aus $\mathbb{C}[X]$ Nullstellen in $\mathbb{C}$ hat. Entsprechendes gilt aber nicht von $\mathbb{R}$.

Wie schon in Abschnitt 18 angesprochen, können wir in gewissen Fällen die Matrix eines Endomorphismus' durch geeignete Basistransformation in Diagonalform bringen. Sehen wir uns das in einem konkreten Fall an.

Beispiel
Bezeichne E die kanonische Basis von $\mathbb{R}^2$, und sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch $\mathrm{Mat}_{E,E}(f) = \begin{pmatrix} 8 & 18 \\ -3 & -7 \end{pmatrix}$. Das charakteristische Polynom von f ist dann

$$\mathrm{ch}(f) = \det \begin{pmatrix} X-8 & -18 \\ 3 & X+7 \end{pmatrix} = X^2 - X - 2.$$

Die Nullstellen ergeben die Eigenwerte $r_1 = -1$ und $r_2 = 2$.

Wir haben lauter verschiedene Eigenwerte und wissen daher, daß f diagonalisierbar ist. Die Matrix von f hat genau dann Diagonalform, wenn man sie bezüglich einer Basis von Eigenvektoren bildet. Wir bekommen die Eigenvektoren als Lösung der Gleichungssysteme (transponierte Form)

$$\begin{pmatrix} -9 & 3 \\ -18 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \quad \text{und} \quad \begin{pmatrix} -6 & 3 \\ -18 & 9 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

Die Lösungen sind $(\,2 \quad -1\,)$ und $(\,3 \quad -1\,)$. Somit ist

$$B = \{(\,2 \quad -1\,), (\,3 \quad -1\,)\}$$

eine Basis von Eigenvektoren, und es gilt

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Die bisher betrachteten diagonalisierbaren Endomorphismen f von n-dimensionalen Vektorräumen hatten n verschiedene Eigenwerte. Dies ist jedoch keine notwendige Bedingung für Diagonalisierbarkeit.

Ist r_i ein Eigenwert von f, so läßt sich $\mathrm{ch}(f)$ durch $(X - r_i)$ teilen. Zudem gibt es eine natürliche Zahl k_i und ein Polynom $P(X) \in K[X]$ mit

$$\mathrm{ch}(f) = (X - r_i)^{k_i} \cdot P(X), \quad P(r_i) \neq 0.$$

k_i heißt dann die *Vielfachheit* von r_i in $\mathrm{ch}(f)$.

Als entscheidendes Kriterium für Diagonalisierbarkeit wird sich die Beziehung zwischen Vielfachheit eines Eigenwertes und der Dimension des zugehörigen Eigenraums herausstellen. Dazu eine erste Abschätzung:

20.2 Lemma
Seien V ein n-dimensionaler K-Vektorraum und $f \in \mathrm{End}_K(V)$. Für jeden Eigenwert $r \in K$ von f gilt

$$\textit{Vielfachheit von } r \geq \dim E(r).$$

Beweis: Sei $v_1, \ldots, v_l$ eine Basis des Eigenraums $E(r) \subset V$. Ergänzt man diese (beliebig) zu einer Basis $B = \{v_1, \ldots, v_l, v_{l+1}, \ldots, v_n\}$ von V, so erhält man

$$A = \mathrm{Mat}_{B,B}(f) = \begin{pmatrix} r & & 0 & \\ & \ddots & & 0 \\ 0 & & r & \\ & 0 & & A' \end{pmatrix}, \quad A' \in K^{(n-l,n-l)}.$$

Das charakteristische Polynom hat dann die Gestalt

$$\mathrm{ch}(f) = \det(XE - A) = (X - r)^l \cdot \det(XE - A'),$$

woraus sich mit $l = \dim E(r)$ die Behauptung ergibt. □

Nun lassen sich die diagonalisierbaren Endomorphismen allgemein kennzeichnen:

20.3 Satz
Sei V ein n-dimensionaler K-Vektorraum. Für ein $f \in \mathrm{End}_K(V)$ sind folgende Aussagen äquivalent:

(a) *f ist diagonalisierbar;*

(b) $\mathrm{ch}(f)$ *zerfällt in Linearfaktoren, und*
für jeden Eigenwert r von f ist die Vielfachheit k gleich der Dimension des Eigenraums $E(r)$;

(c) *für die verschiedenen Eigenwerte* $r_1, \ldots, r_l$ *von* f *gilt*

$$V = E(r_1) \oplus \cdots \oplus E(r_l).$$

Beweis: Seien $r_1, \ldots, r_l$ die verschiedenen Eigenwerte von f.

(a)⇒(b) Da f diagonalisierbar ist, muß $\mathrm{ch}(f)$ gleich dem charakteristischen Polynom einer Diagonalmatrix sein, also

$$\mathrm{ch}(f) = \prod_{i=1}^{l} (X - r_i)^{k_i}.$$

Damit haben wir die Gleichung

$$k_1 + \ldots + k_l = n = \dim E(r_1) + \ldots + \dim E(r_l).$$

Wegen $k_i \geq \dim E(r_i)$ folgt daraus $k_i = \dim E(r_i)$.

(b)⇒(c) Aus den beiden Bedingungen ergibt sich

$$k_1 + \ldots + k_l = n \quad \text{ und } \dim E(r_1) + \cdots + E(r_l) = n.$$

Mit 18.3 haben wir dann $V = \bigoplus_{i=1}^{l} E(r_i)$.

(c)⇒(a) Ist B eine Basis, die aus Eigenvektoren besteht (in richtiger Reihenfolge), so hat $\mathrm{Mat}_{B,B}(f)$ Diagonalgestalt. □

Auch die oben beschriebene Situation wollen wir uns in einem konkreten Fall ansehen.

Beispiel
Mit der kanonischen Basis E sei die lineare Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch

$$\mathrm{Mat}_{E,E}(f) = \begin{pmatrix} 0 & -3 & -2 \\ -1 & -2 & -2 \\ 1 & 3 & 3 \end{pmatrix}.$$

Das charakteristische Polynom ist

$$\mathrm{ch}(f) = \det \begin{pmatrix} X & 3 & 2 \\ 1 & X+2 & 2 \\ -1 & -3 & X-3 \end{pmatrix} = (X-1)^2(X+1).$$

Also ist 1 ein Eigenwert mit Vielfachheit 2, und -1 ist einfacher Eigenwert. Die Eigenvektoren zu 1 erhalten wir als Lösung des Systems

$$(x_1, x_2, x_3) \begin{pmatrix} 1 & 3 & 2 \\ 1 & 3 & 2 \\ -1 & -3 & -2 \end{pmatrix} = 0.$$

Der Rang dieses Systems ist 1, und daher ist die Dimension des Lösungsraumes gleich 2, mit der Basis $(1,0,1)$, $(0,1,1)$.

Die Eigenvektoren zu -1 sind die Lösungen von

$$(x_1,x_2,x_3)\begin{pmatrix}-1 & 3 & 2\\ 1 & 1 & 2\\ -1 & -3 & -4\end{pmatrix} \stackrel{!}{=} 0,$$

welche von $(1,3,2)$ erzeugt werden. Somit ist

$$B = \Big((1,0,1),(0,1,1),(1,3,2)\Big)$$

eine Basis von Eigenvektoren und

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}, \quad \mathrm{Mat}_{B,E}(id) = \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & 3 & 2\end{pmatrix}.$$

Das folgende Diagramm stellt die Zusammenhänge dar:

$$\begin{array}{ccc} {}_E\,\mathbb{R}^3 & \xrightarrow{f} & \mathbb{R}^3\,{}_E \\ \mathrm{id}\downarrow & & \downarrow\mathrm{id} \\ {}_B\,\mathbb{R}^3 & \xrightarrow{f} & \mathbb{R}^3\,{}_B \end{array}$$

Daraus lesen wir für die beteiligten Matrizen ab:

$$\begin{aligned}\mathrm{Mat}_{E,E}(f) &= \mathrm{Mat}_{E,B}(\mathrm{id})\cdot\mathrm{Mat}_{B,B}(f)\cdot\mathrm{Mat}_{B,E}(\mathrm{id}),\\ \mathrm{Mat}_{B,B}(f) &= \mathrm{Mat}_{B,E}(\mathrm{id})\cdot\mathrm{Mat}_{E,E}(f)\cdot\mathrm{Mat}_{E,B}(\mathrm{id}).\end{aligned}$$

Allgemeiner als in 20.2 wollen wir nun fragen, welche Matrixform man für Endomorphismen erreichen kann, deren charakteristisches Polynom in Linearfaktoren (ohne Zusatzbedingungen) zerfällt. Dazu eine Definition:

20.4 Definition
Ein Endomorphismus f eines endlich dimensionalen K-Vektorraums V heißt *trigonalisierbar*, wenn es eine Basis B von V gibt, für die $\mathrm{Mat}_{B,B}(f)$ untere Dreiecksform hat.

Eine Matrix $A \in K^{(n,n)}$ heißt *trigonalisierbar*, wenn sie zu einer unteren Dreiecksmatrix ähnlich ist, wenn es also ein invertierbares $T \in K^{(n,n)}$ gibt, so daß TAT^{-1} untere Dreiecksmatrix ist.

Wir werden sehen, daß wir für solche Endomorphismen eine Familie von Unterräumen in V mit besonderen Eigenschaften finden können.

20.5 Definition
Sei V ein n-dimensionaler K-Vektorraum. Eine Kette von Unterräumen

$$V_1 \subset V_2 \subset \ldots \subset V_n = V$$

nennt man *Fahne in* V, wenn $\dim V_i = i$ für alle $i = 1, \ldots, n$.

Sie heißt f*-stabil* bzgl. $f \in \mathrm{End}_K(V)$, wenn $f(V_i) \subset V_i$ für $i = 1, \ldots, n$.

Die Existenz solcher Fahnen hängt von den Eigenheiten des Endomorphismus f ab. Wir haben dazu folgende Kennzeichnung:

20.6 Trigonalisierbare Endomorphismen
Sei V ein n-dimensionaler K-Vektorraum. Für $f \in \mathrm{End}_K(V)$ sind folgende Aussagen äquivalent:

(a) *Es gibt eine f-stabile Fahne $V_1 \subset V_2 \subset \cdots \subset V_n$ in V;*

(b) *f ist trigonalisierbar;*

(c) *für jede Basis C von V ist $\mathrm{Mat}_{C,C}(f)$ trigonalisierbar;*

(d) $\mathrm{ch}(f)$ *zerfällt in (nicht notwendig verschiedene) Linearfaktoren.*

Beweis: (a)⇒(b) Ist $V_1, \ldots, V_n$ eine f-stabile Fahne, so wähle man eine Basis $B = \{v_1, \ldots, v_n\}$ mit $v_i \in V_i$. Wegen $f(V_i) \subset V_i$ gilt

$$f(v_i) = a_{i1}v_1 + a_{i2}v_2 + \ldots + a_{ii}v_i,$$

also

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & 0 & \\ \vdots & & \ddots & \\ a_{n1} & \ldots\ldots\ldots & & a_{nn} \end{pmatrix}.$$

(b)⇔(c) Dies folgt aus dem Verhalten der f zugeordneten Matrizen bei Basistransformationen.

(b)⇒(d) Offensichtlich zerfällt die Determinante einer Dreiecksmatrix in Linearfaktoren.

(d)⇒(a) Wir zeigen dies durch Induktion nach $n = \dim V$. Für $n = 1$ ist die Aussage trivial. Nehmen wir an, die Behauptung sei richtig für alle $n-1$-dimensionalen K-Vektorräume.

Sei nun $\dim V = n$. Wir suchen eine f-stabile Fahne in V.
Zur Nullstelle r_1 von $\mathrm{ch}(f)$ wählen wir einen Eigenvektor $v_1 \in V$ und ergänzen ihn zu einer Basis $B = (v_1, w_2, \ldots, w_n)$ von V. Setzen wir

$$V_1 := Kv_1 \quad \text{und} \quad W := \sum_{i=2}^{n} Kw_i \subset V,$$

so gilt $f(V_1) \subset V_1$, aber nicht notwendig $f(W) \subset W$.

Für die Bilder der $w_i \in W$ gilt

$$f(w_i) = k_i v_1 + c_{i2} w_2 + \ldots + c_{in} w_n, \quad k_i, c_{ij} \in K.$$

Wir definieren ein $f' : W \to W$ durch Weglassen der v_1,

$$f'(w_i) := c_{i2} w_2 + \ldots + c_{in} w_n \quad \text{für } i = 2, \ldots, n.$$

Mit dieser Festlegung gilt $[f - f'](W) \subset V_1$ und

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} r_1 & 0 & \ldots & 0 \\ * & & & \\ \vdots & & (b_{ij}) & \\ * & & & \end{pmatrix},$$

woraus wir $\mathrm{ch}(f) = (X - r_1)\,\mathrm{ch}(f')$ ablesen.

Somit ist auch $\mathrm{ch}(f')$ Produkt von Linearfaktoren, und nach Induktionsannahme gibt es eine f'-stabile Fahne $W_1 \subset \cdots \subset W_{n-1}$ in W. Dann ist

$$V_i = \begin{cases} Kv_1 & \text{für } i = 1 \\ V_1 + W_{i-1} & \text{für } i = 2, \ldots, n \end{cases}$$

eine f-stabile Fahne in V, denn aus $W_i \subset W_{i+1}$ folgt $V_i \subset V_{i+1}$ und

$$\dim V_i = \dim V_1 + \dim W_{i-1} = 1 + (i-1) = i.$$

Außerdem ist $f(V_1) \subset V_1$ (v_1 Eigenvektor), und für $w \in W_{i-1}$ gilt

$$f(w) = (f - f')(w) + f'(w) \in V_1 + W_{i-1} = V_i.$$

Somit erhalten wir die gewünschte Beziehung

$$f(V_i) = f(V_1 + W_{i-1}) \subset V_1 + W_{i-1} = V_i.$$

□

Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes Polynom aus $K[X]$ eine Nullstelle hat, wenn also jedes Polynom aus $K[X]$ in Linearfaktoren zerfällt. Wir hatten bereits angemerkt, daß die komplexen Zahlen diese Eigenschaft haben.

20.7 Korollar

Sei V ein endlich dimensionaler Vektorraum über einem algebraisch abgeschlossenen Körper K. Dann gilt:

(1) *Zu jedem Endomorphismus $f : V \to V$ gibt es eine f-stabile Fahne.*

(2) *Jede Matrix $A \in K^{(n,n)}$ ist trigonalisierbar.*

20.8 Aufgaben

(1) Seien $f : \mathbb{R}^3 \to \mathbb{R}^3$ und $g : \mathbb{R}^3 \to \mathbb{R}^3$ lineare Abbildungen mit

$$\mathrm{Mat}_{E_3,E_3}(f) = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix} \text{ und } \mathrm{Mat}_{E_3,E_3}(g) = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ -1 & 1 & 4 \end{pmatrix}.$$

(i) *Man bestimme die Eigenwerte und Eigenräume von f und g.*

(ii) *Sind f und g diagonalisierbar?*

(2) Es sei

$$A := \begin{pmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{pmatrix} \in \mathbb{R}^{(3,3)}$$

gegeben. Man untersuche, ob A zu einer reellen oder komplexen Diagonalmatrix D ähnlich ist. Ist dies der Fall, so bestimme man eine invertierbare Matrix T mit $D = TAT^{-1}$.

(3) Es sei

$$A := \begin{pmatrix} 5 & 2 & 16 \\ 4 & -2 & 8 \\ -4 & -1 & -11 \end{pmatrix} \in \mathbb{R}^{(3,3)}$$

gegeben. Man finde eine zu A ähnliche Diagonalmatrix $D \in \mathbb{R}^{(3,3)}$. Außerdem bestimme man eine invertierbare Matrix $T \in \mathbb{R}^{(3,3)}$ mit $D = TAT^{-1}$.

21 Nilpotente Endomorphismen

Bei den Betrachtungen zur Darstellbarkeit gewisser Endomorphismen wollen wir uns in diesem Abschnitt mit *nilpotenten* Endomorphismen f befassen, für die es also ein $k \in \mathbb{N}$ gibt mit $f^k = 0$.

Welche Matrixdarstellung kann man durch geschickte Basiswahl für solche Abbildungen erreichen? Wir wollen dies über Körpern K untersuchen.

21.1 Satz
Seien V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$, und es gebe ein $k \in \mathbb{N}$ mit $f^{k-1} \neq 0$ und $f^k = 0$. Dann gilt:

(1) *Für $v \in V \setminus \operatorname{Kern} f^{k-1}$ sind die Vektoren $v, f(v), \ldots, f^{k-1}(v)$ linear unabhängig und spannen einen f-stabilen Unterraum auf.*

(2) *Ist $k = n$, so gilt für die Basis $B = (v, f(v), \ldots, f^{k-1}(v))$ von V*

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ & 0 & 1 & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ & 0 & & 0 & 1 \\ & & & & 0 \end{pmatrix} =: N(n) \in K^{(n,n)}.$$

Man nennt $N(n)$ einen *elementaren Nilpotenzblock*.
In diesem Fall erlaubt V keine echte Zerlegung in f-invariante Unterräume.

Beweis: (1) Zur Untersuchung der linearen Unabhängigkeit betrachten wir eine Linearkombination

$$\sum_{i=0}^{k-1} a_i f^i(v) = 0 \quad \text{mit } a_i \in K.$$

Durch wiederholte Anwendung von Potenzen von f erhalten wir

$$a_0 f^{k-1}(v) = 0, \quad a_1 f^{k-1}(v) = 0, \ldots$$

also $a_0 = a_1 = \ldots = a_{k-1} = 0$.

Jede Linearkombination $\sum_{i=1}^{k-1} a_i f^i(v)$ geht unter f wieder in eine Linearkombination gleichen Typs über. Somit spannen die $f^i(v)$ einen f-invarianten Unterraum auf.

(2) Die angegebene Gestalt der Matrix folgt unmittelbar aus der Definition von $\mathrm{Mat}_{B,B}(f)$.

Angenommen, $V = U_1 \oplus U_2$ mit nicht-trivialen f-invarianten Unterräumen U_1, U_2. Mit Dimensionsbetrachtungen folgt dann aus (1), daß

$$f^{k_i}|_{U_i} = 0 \quad \text{für } k_i \leq \dim U_i < n, \quad i = 1, 2,$$

und somit $f^k = 0$ für $k = \max(k_1, k_2) < n$.

Dies steht im Widerspruch zur Annahme $k = n$. □

Im allgemeinen kann für $f \in \mathrm{End}_K(V)$ natürlich $f^k = 0$ für $k < n$ gelten. Durch geeignete Zerlegung von V kann man dann auf den oben beschriebenen Fall zurückgreifen:

21.2 Zerlegungssatz für nilpotente Endomorphismen

Seien V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$ und $k < n$ mit $f^{k-1} \neq 0$ und $f^k = 0$. Dann gilt:

(1) *Es gibt eine f-invariante Zerlegung $V = Z \oplus D$ mit*

$$f^{k-1}|_Z \neq 0, \quad f^k|_Z = 0 \quad \text{und } \dim Z = k.$$

(2) *Es gibt eine f-invariante Zerlegung $V = Z_1 \oplus \ldots \oplus Z_l$ mit*

$$f^{k_i-1}|_{Z_i} \neq 0, \quad f^{k_i}|_{Z_i} = 0 \quad \text{und } \dim Z_i = k_i.$$

(3) *Für eine geeignete Basis B von V ist*

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} N(k_1) & & 0 \\ & \ddots & \\ 0 & & N(k_l) \end{pmatrix},$$

wobei die $N(k_i)$ Nilpotenzblöcke der Größe k_i sind.

Beweis: (1) Für $v \in V \setminus \mathrm{Kern}\, f^{k-1}$ ist $Z = \bigoplus_{i=0}^{k-1} K f^i(v)$ nach 21.1 ein f-invarianter Unterraum mit Dimension k.

Wir ergänzen $f^{k-1}(v)$ durch $w_2, \ldots, w_n$ zu einer Basis von V und setzen

$$W := \bigoplus_{i=2}^{n} K w_i, \quad D := W \cap f^{-1}(W) \cap \ldots \cap (f^{k-1})^{-1}(W).$$

Wir zeigen zunächst, daß D f-invariant und $Z \cap D = 0$ ist.

Für $u \in D$ sind alle $f^i(u) \in W$. Somit sind auch für $f(u)$ alle $f^i(f(u)) \in W$ (beachte $f^k = 0$), und wir folgern daraus $f(u) \in D$. Also ist D f-invariant.

Für $a \in Z \cap D$ gilt $f^i(a) \in W$ für alle $i \in \mathbb{N}$ und

$$a = a_0 v + a_1 f(v) + \cdots + a_{k-1} f^{k-1}(v).$$

Bedenkt man $f^k = 0$, so erhalten wir daraus

$$W \ni f^{k-1}(a) = a_0 f^{k-1}(v) \notin W, \quad \text{falls } a_0 \neq 0.$$

Damit muß $a_0 = 0$ sein und – mit ähnlichem Schluß – $a_1 = \ldots = a_{k-1} = 0$.

Dies bedeutet $a = 0$ und $Z \cap D = 0$.

Es genügt, noch zu zeigen, daß $\dim D \geq n - k$.

Nun ist $\dim W = n - 1$, und aus $V \xrightarrow{f} V \xrightarrow{p} V/W$ folgern wir

$$\dim f^{-1}(W) = \dim \operatorname{Kern} p \circ f = n - \dim \operatorname{Bild} p \circ f \geq n - 1.$$

Aus Korollar 11.4 folgt $\dim(W \cap f^{-1}(W)) \geq n - 2$ und schließlich

$$\dim D = \dim(W \cap \ldots \cap (f^{k-1})^{-1}(W)) \geq= n - k.$$

(2) Nach (1) gibt es eine Zerlegung $V = Z \oplus D$, mit f-invarianten Z und D. Dann ist $f' := f|_D$ ein nilpotenter Endomorphismus von D und $\dim D < \dim V$.

Nach der in (1) aufgezeigten Methode kann man nun in D einen f'-invarianten Unterraum abspalten. Dies führt nach endlich vielen Schritten zu der angegebenen Zerlegung.

(3) Ist $V = Z_1 \oplus \ldots \oplus Z_l$ (nach (2)), so ist

$$\operatorname{Mat}(f) = \begin{pmatrix} \operatorname{Mat}_{B_1,B_1}(f_1) & & 0 \\ & \ddots & \\ 0 & & \operatorname{Mat}_{B_l,B_l}(f_l) \end{pmatrix},$$

und nach 21.1 gilt

$$\operatorname{Mat}_{B_i,B_i}(f_i) = N(k_i)$$

bezüglich geeigneter Basen B_i von Z_i. □

Praktische Bestimmung von *D*

Um nach dem Beweis von 21.2 konstruktiv eine Zerlegung $V = Z \oplus D$ zu finden, braucht man eine Möglichkeit, den Unterraum D zu ermitteln. Dies läßt sich folgendermaßen durchführen:

Jedes W ist als $(n-1)$-dimensionaler Unterraum von V Kern einer geeigneten linearen Abbildung $h : V \to K$, wobei $h(f^{k-1}(V)) \neq 0$.

Dann ist

$$\begin{aligned} f^{-1}(W) &= \{v \in V \mid f(v) \in \operatorname{Kern} h\} \\ &= \{v \in V \mid h \circ f(v) = 0\} \quad = \quad \operatorname{Kern} h \circ f, \\ (f^i)^{-1}(W) &= \operatorname{Kern} h \circ f^i, \end{aligned}$$

und man erhält $D = W \cap f^{-1}(W) \cap \ldots \cap (f^{k-1})^{-1}(W)$ als Lösungsmenge des linearen Gleichungssystems

$$\begin{aligned} h(x) &= 0 \\ h \circ f(x) &= 0 \\ &\vdots \\ h \circ f^{k-1}(x) &= 0. \end{aligned}$$

Wird $f : K^n \to K^n$ durch die Matrix $A \in K^{(n,n)}$ (bzgl. der kanonischen Basis) und $h \in \mathrm{Hom}_K(V, K)$ durch $H = (h_1, \ldots, h_n)^t \in K^{(n,1)}$ dargestellt, so ergibt dies das lineare Gleichungssystem (x ist Zeilenvektor)

$$\begin{aligned} x \cdot H &= 0 \\ x \cdot AH &= 0 \\ &\vdots \\ x \cdot A^{k-1}H &= 0 \end{aligned}$$

Diese theoretischen Vorgaben sollen an einem praktischen Fall nachvollzogen werden.

Beispiel
Sei $f : \mathbb{R}^5 \to \mathbb{R}^5$ eine lineare Abbildung, die bezüglich der kanonischen Basis E gegeben ist durch

$$\mathrm{Mat}_{EE}(f) = \begin{pmatrix} -1 & -2 & 1 & 2 & -1 \\ -1 & -2 & 2 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -4 & 3 & 4 & -2 \\ -1 & -2 & 1 & 2 & -1 \end{pmatrix}.$$

Wir suchen eine Basis im $\mathbb{R}^5$, bezüglich welcher die Matrix von f aus Nilpotenzblöcken besteht. Das charakteristische Polynom ergibt sich als

$$\mathrm{ch}(f) = -\det \begin{pmatrix} -1-X & -2 & 1 & 2 & -1 \\ -1 & -2-X & 2 & 2 & -1 \\ 0 & 0 & -X & 0 & 0 \\ -2 & -4 & 3 & 4-X & -2 \\ -1 & -2 & 1 & 2 & -1-X \end{pmatrix} = X^5.$$

Aus dem Satz von Cayley-Hamilton folgt damit $f^5 = 0$. Somit wissen wir, daß es eine Basis mit den gewünschten Eigenschaften gibt. Man kann direkt nachrechnen, daß bereits $f^2 = 0$.

Um $\mathrm{Kern}\, g$ zu bestimmen, ist folgendes Gleichungssystem zu lösen (Matrix transponieren)

$$\begin{pmatrix} -1 & -1 & 0 & -2 & -1 \\ -2 & -2 & 0 & -4 & -2 \\ 1 & 2 & 0 & 3 & 1 \\ 2 & 2 & 0 & 4 & 2 \\ -1 & -1 & 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Durch elementare Zeilenumformungen erhalten wir die Matrix

$$\begin{pmatrix} 1 & 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

und lesen daraus als Basis für den Lösungsraum ab

$$(0,0,1,0,0), (-1,-1,0,1,0), (-1,0,0,0,1).$$

Wegen $f^2 = 0$ ist $\operatorname{Kern} f^2 = \mathbb{R}^5$. Als erste Basiselemente nehmen wir

$$\begin{aligned} b_1 &:= (1,0,0,0,0) \in \mathbb{R}^5 \setminus \operatorname{Kern} f, \\ b_2 &:= f(b_1) = (-1,-2,1,2,-1). \end{aligned}$$

Nun suchen wir ein $h_1 : \mathbb{R}^5 \to \mathbb{R}$ mit $b_2 \notin \operatorname{Kern} h_1$, etwa $\operatorname{Mat}(h_1) = (1,0,0,0,0)^t$. Dann ist $\operatorname{Mat}(h_1 \circ f) = (-1,-1,0,-2,-1)$, und wir bekommen D_1 als Lösung des homogenen Systems mit der Matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -2 & -1 \end{pmatrix},$$

für die wir als Basis erhalten

$$(0,0,1,0,0), (0,-2,0,1,0), (0,-1,0,0,1).$$

Als weitere Basisvektoren für $\mathbb{R}^5$ wählen wir

$$\begin{aligned} b_3 &:= (0,-2,0,1,0) \in D_1 \setminus \operatorname{Kern} f, \\ b_4 &:= f(b_3) = (0,0,1,0,0). \end{aligned}$$

Um den letzten Basisvektor zu finden, brauchen wir ein $h_2 : \mathbb{R}^3 \to \mathbb{R}$ mit $b_4 \notin \operatorname{Kern} h_2$. Dies wird mit $\operatorname{Mat}(h_2) = (0,0,1,0,0)^t$ erfüllt, und $\operatorname{Mat}(h_2 \circ f) = (1,2,0,3,1)$.
Zusammen mit obigen Gleichungen bekommen wir das System mit Matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 3 & 1 \end{pmatrix}.$$

Die Lösung dazu liefert als fünften Basisvektor von $\mathbb{R}^5$

$$b_5 := (0,1,0,-1,1).$$

Bezüglich der so gewählten Basis $B = (b_1, b_2, b_3, b_4, b_5)$ gilt nun

$$\mathrm{Mat}_{BB}(f) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Transformationsmatrix setzt sich aus den Basisvektoren zusammen,

$$\mathrm{Mat}_{BE}(id) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & -2 & 1 & 2 & -1 \\ 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 \end{pmatrix}.$$

21.3 Aufgaben

Sei $g : \mathbb{R}^4 \to \mathbb{R}^4$ eine lineare Abbildung mit

$$\mathrm{Mat}_{EE}(g) = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & -3 & 3 \end{pmatrix}.$$

Man zeige, daß g nilpotent ist und bestimme eine Basis B im $\mathbb{R}^4$, bezüglich welcher $\mathrm{Mat}_{BB}(g)$ aus Nilpotenzblöcken besteht.

22 Haupträume und Jordansche Normalform

Sozusagen als Defekt bei der Suche nach der Diagonaldarstellung einer linearen Abbildung $f : V \to V$ war die Tatsache anzusehen, daß für einen Eigenwert $r \in K$ die Dimension von $E(r) = \text{Kern}(r\,\text{id} - f)$ echt kleiner als die Vielfachheit des Eigenwertes r im charakteristischen Polynom sein konnte. War dies der Fall, so konnten wir keine Diagonalmatrix für f finden, sondern nur eine Dreiecksgestalt, wenn $\text{ch}(f)$ in Linearfaktoren zerfiel.

Durch verfeinertes Studium dieser Situation können wir dies etwas verbessern. Dazu betrachten wir nicht nur $\text{Kern}(r\,\text{id} - f)$, sondern auch die Kerne der Potenzen $(r\,\text{id} - f)^i$ mit $i \in \mathbb{N}$.

Wir stellen diese Untersuchungen über einem Körper K an.

22.1 Definition
Seien V ein K-Vektorraum und $f \in \text{End}_K(V)$. Ist $r \in K$ ein Eigenwert von f, so nennt man

$$\begin{aligned} H(r) &= \{v \in V \mid (r\,\text{id} - f)^k(v) = 0 \text{ für ein } k \in \mathbb{N}\} \\ &= \bigcup_{k \in \mathbb{N}} \text{Kern}\,(r\,\text{id} - f)^k \end{aligned}$$

den *Hauptraum* (*verallgemeinerten Eigenraum*) von f zum Eigenwert r.

Natürlich ist für jeden Eigenwert $r \in K$ von $f \in \text{End}_K(V)$ der Eigenraum $E(r)$ im Hauptraum $H(r)$ enthalten. Wir werden sehen, daß bei zerfallendem charakteristischem Polynom die (direkte) Summe der Haupträume schon ganz V ergibt.

Zunächst stellen wir einige allgemeine Beziehungen für Haupträume fest, die denen von Eigenräumen recht ähnlich sind (vgl. 18.3).

22.2 Eigenschaften von Haupträumen
Seien V ein n-dimensionaler K-Vektorraum, $f \in \text{End}_K(V)$ und $r, s \in K$ Eigenwerte von f. Dann gilt:

(1) *Der Hauptraum $H(r)$ ist ein f-stabiler Unterraum von V.*

(2) *$H(r) = \text{Kern}(r\,\text{id} - f)^{k_0}$ für ein geeignetes $k_0 \leq n$.*

(3) *Gilt $r \neq s$, dann ist $H(r) \cap H(s) = 0$.*

(4) *Sind $r_1, \ldots, r_k$ verschiedene Eigenwerte von f, so bilden die Haupträume $H(r_1) \ldots, H(r_k)$ eine unabhängige Familie von Untermoduln (vgl. 8.9), also*

$$\sum_{i=1}^{k} H(r_i) = H(r_1) \oplus \cdots \oplus H(r_k).$$

Beweis: (1) Sei $u \in H(r)$, also $(r\,\mathrm{id} - f)^k(u) = 0$ für ein $k \in \mathbb{N}$. Dann gilt auch

$$(r\,\mathrm{id} - f)^k[f(u)] = f[(r\,\mathrm{id} - f)^k(u)] = 0,$$

also $f(u) \in H(r)$, und somit ist $H(r)$ f-invariant.

Mit (2) wird sich ergeben, daß $H(r)$ ein K-Unterraum ist.

(2) Aus dem Fittingschen Lemma 13.8 folgt, daß die aufsteigende Kette

$$\mathrm{Kern}(r\,\mathrm{id} - f) \subset \mathrm{Kern}(r\,\mathrm{id} - f)^2 \subset \mathrm{Kern}(r\,\mathrm{id} - f)^3 \subset \ldots$$

stationär wird, d.h. es gibt ein $k_0 \leq n$ mit

$$\mathrm{Kern}(r\,\mathrm{id} - f)^{k_0} = \mathrm{Kern}(r\,\mathrm{id} - f)^{k_0+l} \quad \text{für alle } l \in \mathbb{N}.$$

Dies impliziert $H(r) = \mathrm{Kern}(r\,\mathrm{id} - f)^{k_0}$.

(3) Nach (2) gibt es $k_r, k_s \in \mathbb{N}$ mit

$$H(r) = \mathrm{Kern}(r\,\mathrm{id} - f)^{k_r} \quad \text{und } H(s) = \mathrm{Kern}(s\,\mathrm{id} - f)^{k_s}.$$

Nehmen wir an, daß $D := H(r) \cap H(s)$ nicht Null ist. Nach (1) ist D ein f-invarianter Unterraum und somit auch invariant gegenüber $(r\,\mathrm{id} - f)$ und $(s\,\mathrm{id} - f)$. Damit können wir folgende Endomorphismen von D bilden:

$$g := (r\,\mathrm{id} - f)|_D \quad \text{und } (s\,\mathrm{id} - f)|_D = (g + t\,\mathrm{id})|_D \text{ mit } t := s - r,$$

mit den Eigenschaften (nach Definition von D)

$$g^{k_r} = 0 \quad \text{und } (g + t\,\mathrm{id}_D)^{k_s} = 0.$$

Die rechte der beiden Gleichungen ergibt die Beziehung

$$t^{k_s}\,\mathrm{id}_D = -g(t^{k_s-1}\,\mathrm{id}_D + \ldots),$$

bei der die linke Seite nicht nilpotent (da $t \neq 0$), die rechte aber nilpotent ist (da g nilpotent). Dies ist ein Widerspruch. Also ist die Annahme $D \neq 0$ falsch.

(4) Den Beweis führen wir durch Induktion nach k (vgl. 18.3). Für $k = 1$ ist die Behauptung klar, für $k = 2$ wurde sie in (3) gezeigt.

Nehmen wir an, die Behauptung sei für $k-1$ richtig. Es ist zu zeigen, daß in einer Summe

$$(*) \qquad z_1 + \cdots + z_k = 0 \quad \text{mit } z_i \in H(r_i)$$

alle $z_i = 0$ sein müssen. Ist $H(r_k) = \mathrm{Kern}(r_k\,\mathrm{id} - f)^{k_0}$, so folgt aus $(*)$

$$(r_k\,\mathrm{id} - f)^{k_0}(z_1) + \cdots + (r_k\,\mathrm{id} - f)^{k_0}(z_{k-1}) = 0.$$

Wegen der f-Invarianz der $H(r_i)$, gilt auch $(r_k \,\mathrm{id} - f)^{k_0}(z_i) \in H(r_i)$, und nach Induktionsannahme folgt daraus

$$(r_k \,\mathrm{id} - f)^{k_0}(z_i) = 0 \quad \text{für } i < k.$$

Dies wiederum bedeutet $z_i \in H(r_i) \cap H(r_k) = 0$ (wegen (3)), also $z_i = 0$ für $i < k$. Nach $(*)$ ist dann auch $z_k = 0$. □

Wir können nun für Endomorphismen mit zerfallendem charakteristischem Polynom zeigen:

22.3 Zerlegungssatz
Seien V ein n-dimensionaler Vektorraum und $f \in \mathrm{End}_K(V)$. Zerfällt $\mathrm{ch}(f)$ in Linearfaktoren, und sind $r_1, \ldots, r_k$ die verschiedenen Eigenwerte von f, so ist

$$V = H(r_1) \oplus \cdots \oplus H(r_k).$$

Beweis: Wegen 22.2 bleibt nur noch zu zeigen, daß V von den Haupträumen erzeugt wird. Dies beweisen wir durch Induktion nach der Dimension von V. Für $n = 1$ ist die Aussage trivial. Nehmen wir an, die Behauptung gelte für alle Vektorräume mit Dimension $< n$.

Sei nun $\dim V = n$ und $t \in \mathbb{N}$ mit $H(r_k) = \mathrm{Kern}(r_k \,\mathrm{id} - f)^t$. Setzen wir $U := \mathrm{Bild}(r_k \,\mathrm{id} - f)^t$, so bekommen wir mit dem Fitting-Lemma eine Zerlegung

$$V = H(r_k) \oplus U$$

in offensichtlich f-invariante Unterräume, und $\dim U < n$.

Ist $U = 0$, so bleibt nichts mehr zu zeigen. Für $U \neq 0$ wollen wir die Induktionsannahme auf die Restriktion

$$f' := f|_U \in \mathrm{End}_K(U)$$

anwenden. Dazu müssen wir uns überlegen, ob auch $\mathrm{ch}(f')$ in Linearfaktoren zerfällt. Wählen wir eine Basis B' in U und ergänzen sie (irgendwie) zu einer Basis B von V. Wegen der f-Invarianz von U gilt dann

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} \mathrm{Mat}_{B',B'}(f') & 0 \\ * & C \end{pmatrix}.$$

Daraus sieht man, daß $\mathrm{ch}(f')$ ebenfalls in Linearfaktoren zerfällt.

Die Eigenwerte $r_1, \ldots, r_m$ von f' sind auch Eigenwerte von f, und nach Induktionsannahme ist

$$U = \bigoplus_{i=1}^{m} H'(r_i).$$

Außerdem ist $H'(r_i) \subset H(r_i)$, denn für $u \in H'(r_i) \subset U$ gilt

$$0 = (r_i \,\mathrm{id} - f')(u) = (r_i \,\mathrm{id} - f)(u).$$

Somit haben wir $V = H(r_1) \oplus \cdots \oplus H(r_k)$. □

Bemerkung: Durch die Zerlegung von $V = \bigoplus H(r_i)$ mit f-invarianten $H(r_i)$ können wir f durch die

$$f_i := f|_{H(r_i)} \in \mathrm{End}(H(r_i))$$

darstellen. Für diese gilt $(r_i\,\mathrm{id} - f_i)^{k_i} = 0$.

Wir wollen nun die Zerlegung von V in Haupträume und die früher gewonnene Darstellung von nilpotenten Endomorphismen benutzen, um eine spezielle Darstellung eines Endomorphismus mit zerfallendem charakteristischem Polynom zu gewinnen.

22.4 Jordansche Normalform

Sei V ein n-dimensionaler K-Vektorraum, $f \in \mathrm{End}_K(V)$, und das charakteristische Polynom von f zerfalle in Linearfaktoren. Dann gibt es eine Zerlegung

$$V = V_1 \oplus \cdots \oplus V_t$$

in f-invariante Unterräume V_i mit Basen B_i, so daß für die damit gebildete Basis $B = (B_1, \ldots, B_t)$ von V gilt

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} J(k_1, r_1) & & 0 \\ & \ddots & \\ 0 & & J(k_t, r_t) \end{pmatrix}$$

mit elementaren Jordanblöcken

$$J(k_i, r_i) = r_i E_{k_i} + N(k_i) = \begin{pmatrix} r_i & 1 & 0 \\ & \ddots & 1 \\ 0 & & r_i \end{pmatrix} \in K^{(k_i,k_i)},$$

wobei die $N(k_i)$ elementare Nilpotenzblöcke und die r_i (nicht notwendig verschiedene) Eigenwerte von f sind.

Eine solche Basis nennt man Jordanbasis *zu f.*

Beweis: Sind $r_1, \ldots, r_k$ die verschiedenen Eigenwerte von f, so haben wir nach 22.3 eine f-invariante Zerlegung

$$V = H(r_1) \oplus \ldots \oplus H(r_k) \quad \text{mit } H(r_i) = \mathrm{Kern}(r_i\,\mathrm{id} - f)^{k_i}.$$

Betrachten wir die Endomorphismen von $H(r_i)$,

$$g_i := (f - r_i\,\mathrm{id})|_{H(r_i)} \quad \text{und } f_i := f|_{H(r_i)}.$$

Die g_i sind nilpotent vom Grad $\leq k_i$.

Konzentrieren wir uns zunächst auf g_1. Nach dem Zerlegungssatz für nilpotente Endomorphismen 21.2 gibt es eine Zerlegung

$$H(r_1) = V_1 \oplus \cdots \oplus V_r$$

in g_1-invariante Teilräume V_s mit Basen B_s, für die $\mathrm{Mat}_{B_s,B_s}(g_1)$ elementare Nilpotenzblöcke der Größe $k_s = \dim V_s$ sind.

Wegen $f_1 = r_1 \,\mathrm{id} + g_1$ haben wir damit

$$\mathrm{Mat}_{B_s,B_s}(f_1) = r_1 \mathrm{Mat}_{B_s,B_s}(\mathrm{id}) + \mathrm{Mat}_{B_s,B_s}(g_1).$$

Die g_1-invarianten Teilräume V_s sind auch invariant gegenüber f_1 und f. Daher setzt sich die Matrix von f_1 bezüglich der Basis $B_1, \ldots, B_r$ von $H(r_1)$ gerade aus den $\mathrm{Mat}_{B_s,B_s}(f_1)$ in der Diagonalen zusammen.

Die entsprechenden Zerlegungen lassen sich natürlich auch für die anderen Eigenwerte durchführen. Setzt man die dabei gefundenen Basen der $H(r_i)$ zu einer Basis B von V zusammen, so hat $\mathrm{Mat}_{BB}(f)$ die angegebene Gestalt.

Nach Konstruktion ist klar, daß zu einem Eigenwert mehrere elementare Jordanblöcke auftreten können. □

Formulieren wir den vorangehenden Satz auch für Matrizen:

22.5 Jordansche Normalform von Matrizen
Sei $A \in K^{(n,n)}$ eine Matrix, für die das charakteristische Polynom in Linearfaktoren zerfällt. Dann gibt es eine invertierbare Matrix $T \in K^{(n,n)}$ mit

$$TAT^{-1} = \begin{pmatrix} J(k_1, r_1) & & 0 \\ & \ddots & \\ 0 & & J(k_t, r_t) \end{pmatrix}$$

mit elementaren Jordanblöcken

$$J(k_i, r_i) = r_i E_{k_i} + N(k_i) = \begin{pmatrix} r_i & 1 & 0 \\ & \ddots & 1 \\ 0 & & r_i \end{pmatrix} \in K^{(k_i,k_i)},$$

wobei die r_i (nicht notwendig verschiedene) Eigenwerte von f sind.

Zum Auffinden der angegebenen Matrix T kann man sich am Beweis von 22.4 orientieren. Sehen wir uns das an einem konkreten Fall an.

Beispiel
Gegeben sei die Matrix

$$A = \begin{pmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{pmatrix} \in \mathbb{R}^{(4,4)},$$

und wir betrachten die dadurch bestimmte lineare Abbildung

$$f : \mathbb{R}^4 \to \mathbb{R}^4, \quad (x_1, x_2, x_3, x_4) \mapsto (x_1, x_2, x_3, x_4).$$

Berechnen wir das charakteristische Polynom,

$$\det(XE - A) = \det \begin{pmatrix} X-3 & 1 & 0 & 0 \\ -1 & X-1 & 0 & 0 \\ -3 & 0 & X-5 & 3 \\ -4 & 1 & -3 & X+1 \end{pmatrix} = (X-2)^4.$$

Zu $g := f - 2\,\mathrm{id}$ erhalten wir die Matrix

$$B := \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 0 & 3 & -3 \\ 4 & -1 & 3 & -3 \end{pmatrix},$$

für die wir $B^2 = 0$ ermitteln, also auch $g^2 = 0$. Die Bestimmung der Jordanbasis geht nun von dem nilpotenten Endomorphismus g aus (vgl. 21.2). Wegen $g^2 = 0$ ist $\operatorname{Kern} g^2 = \mathbb{R}^4$, und wir nehmen als erste Basiselemente (von Z)

$$\begin{aligned} b_1 &:= (1,0,0,0) \in \mathbb{R}^4 \setminus \operatorname{Kern} g, \\ b_2 &:= g(b_1) = (1,-1,0,0). \end{aligned}$$

Zur Bestimmung von D brauchen wir ein $h : \mathbb{R}^4 \to \mathbb{R}$ mit $h(b_2) \neq 0$. Wählen wir $\operatorname{Mat}(h) = (0,1,0,0)^t$, so ergibt sich für $h \circ g$ die Matrix

$$\operatorname{Mat}(h \circ g) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 0 & 3 & -3 \\ 4 & -1 & 3 & -3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \\ -1 \end{pmatrix},$$

und für $D = \operatorname{Kern} h \cap \operatorname{Kern} h \circ g$ haben wir das Gleichungssystem mit der Matrix

$$\begin{pmatrix} \operatorname{Mat}(h)^t \\ \operatorname{Mat}(h \circ g)^t \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & -1 \end{pmatrix}.$$

Als Basis für D erhalten wir $(0,0,1,0)$ und $(1,0,0,-1)$. Somit setzen wir als nächste Basisvektoren für $\mathbb{R}^4$

$$\begin{aligned} b_3 &:= (0,0,1,0) \in D \setminus \operatorname{Kern} g, \\ b_4 &:= g(b_3) = (3,0,3,-3). \end{aligned}$$

Aus der Jordanbasis b_1, b_2, b_3, b_4 ergibt sich die Transformationsmatrix

$$T = \mathrm{Mat}_{BE}(id) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 3 & 0 & 3 & -3 \end{pmatrix},$$

für die man (mit entsprechender Rechnung) bestätigt:

$$TAT^{-1} = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

22.6 Aufgaben

(1) Sei V ein $\mathbb{R}$-Vektorraum mit Basis $B = (b_1, b_2, b_3, b_4, b_5)$, und es sei $f \in \mathrm{End}_{\mathbb{R}}(V)$ gegeben mit

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} 0 & -1 & -4 & -3 & -7 \\ 0 & -1 & -2 & -2 & -4 \\ 1 & -1 & 0 & 2 & -5 \\ 0 & 1 & 2 & 1 & 6 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

(i) *Bestimmem Sie die Eigenwerte und Haupträume von f.*

(ii) *Bestimmem Sie die Jordansche Normalform und eine Jordanbasis für f.*

(2) Sei V ein $\mathbb{R}$-Vektorraum mit Basis $B = (b_1, b_2, b_3, b_4, b_5, b_6)$, und es sei $f \in \mathrm{End}_{\mathbb{R}}(V)$ gegeben mit

$$\mathrm{Mat}_{B,B}(f) = \begin{pmatrix} 4 & 2 & 2 & 3 & 0 & 1 \\ 0 & 5 & 0 & 1 & 3 & 2 \\ 0 & -2 & 4 & -2 & -1 & -4 \\ 0 & 1 & 0 & 5 & -1 & 2 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & -1 & 0 & -1 & -1 & 2 \end{pmatrix}.$$

(i) *Bestimmen Sie die Eigenwerte und Haupträume von f.*

(ii) *Bestimmen Sie die Jordansche Normalform und eine Jordanbasis von f.*

(3) Gegeben sei die Matrix

$$A = \begin{pmatrix} -5 & -9 & -7 & 12 & 17 \\ -3 & -11 & -7 & 12 & 19 \\ 0 & 0 & -2 & 0 & 6 \\ -3 & -9 & -7 & 10 & 22 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix} \in \mathbb{R}^{(5,5)}.$$

(a) *Zeigen Sie, daß sich A durch eine geeignete Basistransformation über $\mathbb{R}$ in Jordansche Normalform bringen läßt.*

(b) *Berechnen Sie die Haupträume von A.*

(c) *Bestimmen Sie eine invertierbare Matrix $T \in \mathbb{R}^{(5,5)}$, so daß TAT^{-1} die Jordansche Normalform von A ist.*

(4) Sei folgende Matrix gegeben:

$$B = \begin{pmatrix} -4 & 1 & 0 & 0 \\ -6 & 0 & 1 & 0 \\ -4 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{(4,4)}$$

(a) *Berechnen Sie die Eigenwerte von B.*

(b) *Geben Sie die Jordansche Normalform von B an (ohne Angabe einer Jordanbasis), und begründen Sie deren Gestalt.*

Kapitel 7

Bilinearformen

23 Linearformen und der Dualraum

Wir haben bereits in Abschnitt 10 festgestellt, daß die Homomorphismen zwischen zwei R-Linksmoduln M und N eine Gruppe $\mathrm{Hom}_R(M,N)$ bilden. Für kommutatives R ist dies sogar ein R-Modul.

Von besonderer Bedeutung sind die Homomorphismen von M in den Grundring R. Nun ist R selbst Links- und Rechtsmodul über sich selbst. Deswegen kann auch bei nicht kommutativem R die Gruppe $\mathrm{Hom}_R(M,R)$ zu einem Rechtsmodul gemacht werden, nämlich durch die Festlegung der Skalarmultiplikation von $s \in R$ mit $f \in \mathrm{Hom}_R(M,R)$

$$fs(m) = f(m)s \text{ für alle } m \in M.$$

Damit gilt für jedes $r \in R$ und $m \in M$

$$fs(rm) = f(rm)s = rf(m)s = r(fs)(m),$$

das heißt $fs \in \mathrm{Hom}_R(M,R)$. Es ist leicht nachzuprüfen, daß $\mathrm{Hom}_R(M,R)$ damit zu einem R-Rechtsmodul wird.

23.1 Definition
Für einen R-Linksmodul M bezeichnet man den R-Rechtsmodul

$$M^* = \mathrm{Hom}_R(M,R)$$

als den *Dualraum* von M, oder auch als den zu M *dualen Modul*. Die Elemente aus M^* nennt man *Linearformen* auf M.

Zu einem Morphismus $M \to N$ von R-Linksmoduln läßt sich ein Morphismus zwischen den dualen Moduln definieren – allerdings in der anderen Richtung:

23.2 Transponierter Homomorphismus

Ist $h : M \to N$ ein Homomorphismus von R-Linksmoduln, so nennt man den Homomorphismus von R-Rechtsmoduln

$$h^t : N^* \to M^*,\ f \mapsto f \circ h,$$

den zu h transponierten *Homomorphismus. Er hat die Eigenschaften*

(1) *Für $\mathrm{id}_M : M \to M$ gilt $(\mathrm{id}_M)^t = \mathrm{id}_{M^*}$.*

(2) *Für $k : M \to N$ gilt $(h+k)^t = h^t + k^t$.*

(3) *Für $p : N \to P$ gilt $(p \circ h)^t = h^t \circ p^t$.*

(4) *Ist h surjektiv, so ist h^t injektiv.*

(5) *Ist h ein Isomorphismus, so ist auch h^t ein Isomorphismus, und $(h^t)^{-1} = (h^{-1})^t$.*

Beweis: h^t ist ein Homomorphismus von R-Rechtsmoduln, denn aus den Definitionen folgt für $f, g \in N^*$ und $s \in R$

$$\begin{array}{ccccccc} h^t(f+g) & = & (f+g)\circ h & = & f\circ h + g \circ h & = & h^t(f) + h^t(g), \\ h^t(fs) & = & (fs)\circ h & = & (f\circ h)s & = & h^t(f)s. \end{array}$$

(1) folgt unmittelbar aus der Definition.

(2) Nach Definition gilt für jedes $f \in N^*$:

$$(h+k)^t(f) = f \circ (h+k) = f \circ h + f \circ k = h^t(f) + k^t(f).$$

(3) Für jedes $g \in P^*$ gilt

$$(p \circ h)^t(g) = g \circ p \circ h = h^t(g \circ p) = h^t \circ p^t(g).$$

(4) Angenommen, für $f \in N^*$ gelte $0 = h^t(f) = f \circ h$. Ist h surjektiv, so bedeutet dies $f = 0$, und somit ist h^t injektiv.

(5) Sei $h : M \to N$ invertierbar. Dann ist $h^{-1} \circ h = \mathrm{id}_M$, und aus (1) und (3) folgt

$$\mathrm{id}_{M^*} = (\mathrm{id}_M)^t = (h^{-1} \circ h)^t = h^t \circ (h^{-1})^t.$$

□

In 12.4 wurde gezeigt, daß für endlich erzeugte freie Moduln M, N über einem kommutativen Ring R auch $\mathrm{Hom}_R(M, N)$ einen freien R-Modul bildet. Dies gilt entsprechend für den Dualraum von freien Moduln. Wir können sogar folgendes zeigen:

23.3 Dualraum von freien Moduln

Sei M ein freier R-Modul mit Basis $\{m_\lambda \mid \lambda \in \Lambda\}$. Für jedes $\lambda \in \Lambda$ definieren wir eine Linearform durch Vorgabe der Werte auf den Basiselementen:

$$m_\lambda^* : M \to R, \quad m_\lambda^*(m_\mu) = \begin{cases} 0 & \text{falls } \lambda \neq \mu \\ 1 & \text{falls } \lambda = \mu. \end{cases}$$

Diese haben folgende Eigenschaften:

(1) $\{m_\lambda^* \mid \lambda \in \Lambda\}$ *ist eine linear unabhängige Teilmenge von* M^*.

(2) *Für* $m = \sum_{\lambda \in \Lambda_0} r_\lambda m_\lambda$, $\Lambda \supset \Lambda_0$ *endlich, gilt* $m_\mu^*(m) = r_\mu$.

(3) *Ist* Λ *endlich, dann ist* $\{m_\lambda^* \mid \lambda \in \Lambda\}$ *eine Basis von* M^*.
Man nennt sie die duale Basis *zu* $\{m_\lambda \mid \lambda \in \Lambda\}$.

(4) *Ist* Λ *endlich und* R *kommutativ, dann ist* $M \simeq M^*$ *als* R*-Modul.*

Beweis: (1) Angenommen, $\sum_{\lambda \in \Lambda_0} m_\lambda^* s_\lambda = 0$, $\Lambda \supset \Lambda_0$ endlich. Dann gilt für jedes $\mu \in \Lambda_0$

$$0 = \sum_{\lambda \in \Lambda_0} m_\lambda^* s_\lambda(m_\mu) = \sum_{\lambda \in \Lambda_0} m_\lambda^*(m_\mu) s_\lambda = s_\mu.$$

Also sind die $\{m_\lambda^* \mid \lambda \in \Lambda\}$ linear unabhängig.

(2) Nach Definition von m_μ^* gilt

$$m_\mu^*(m) = \sum_{\lambda \in \Lambda_0} r_\lambda m_\mu^*(m_\lambda) = r_\mu.$$

(3) Für beliebiges $f \in M^*$ und $m = \sum_{\lambda \in \Lambda} r_\lambda m_\lambda$ haben wir nach (2)

$$\begin{aligned} \left(\sum_{\lambda \in \Lambda} m_\lambda^* f(m_\lambda) \right)(m) &= \sum_{\lambda \in \Lambda} m_\lambda^*(m) f(m_\lambda) \\ &= \sum_{\lambda \in \Lambda} r_\lambda f(m_\lambda) \quad = \quad f(m). \end{aligned}$$

Dies bedeutet $f = \sum_{\lambda \in \Lambda} m_\lambda^* f(m_\lambda)$, und somit ist $\{m_\lambda^* \mid \lambda \in \Lambda\}$ ein Erzeugendensystem von M^*.

(4) Durch die Zuordnung der Basiselemente $m_\lambda \mapsto m_\lambda^*$ ist ein $\mathbb{Z}$-Homomorphismus $M \to M^*$ bestimmt. Dies ist offensichtlich ein Isomorphismus.

Man beachte, daß die Abbildung von der gewählten Basis von M abhängig ist. □

Ein Morphismus zwischen freien R-Moduln kann durch eine Matrix beschrieben werden. Aus dieser bekommt man auch die Matrix für den transponierten Morphismus:

23.4 Matrix der Transponierten

Seien M und N freie R-Moduln mit Basen $X = \{x_1, \ldots, x_m\} \subset M$ und $Y = \{y_1, \ldots, y_n\} \subset N$. Bezeichne $X^* = \{x_1^*, \ldots, x_m^*\} \subset M^*$ und $Y^* = \{y_1^*, \ldots, y_n^*\} \subset N^*$ die dualen Basen.

Ist $f : M \to N$ ein R-Homomorphismus, so gilt für $f^t : N^* \to M^*$

$$\mathrm{Mat}_{Y^*X^*}(f^t) = \mathrm{Mat}_{XY}(f)^t,$$

d.h. bezüglich der gewählten Basen ist die Matrix von f^t gleich der Transponierten der Matrix von f.

Beweis: Setzen wir $f(x_i) = \sum\limits_{j=1}^{n} a_{ij} y_j$ und $f^t(y_k^*) = \sum\limits_{l=1}^{m} b_{kl} x_l^*$, also

$$\mathrm{Mat}_{XY}(f) = (a_{ij}) \in R^{(m,n)}, \quad \mathrm{Mat}_{Y^*X^*}(f^t) = (b_{kl}) \in R^{(n,m)}.$$

Nach Definition gilt $f^t(g) = g \circ f$ für alle $g \in N^*$, also

$$f^t(g)(x_i) = g(f(x_i)) \quad \text{für} \quad i = 1, \ldots, m.$$

Für $g = y_k^* \in N^*$ erhalten wir aus der ersten Gleichung

$$y_k^* f(x_i) = \sum_{j=1}^{n} a_{ij} y_k^*(y_j) = a_{ik}.$$

Setzen wir die x_i in die zweite Gleichung ein, so ergibt sich

$$f^t(y_k^*)(x_i) = \sum_{l=1}^{m} b_{kl}\, x_l^*(x_i) = b_{ki}.$$

Also ist $\mathrm{Mat}_{Y^*X^*}(f^t)$ die Transponierte von $\mathrm{Mat}_{XY}(f)$. □

Beispiel

(1) Sei R ein kommutativer Ring und $M = R^m$ mit kanonischer Basis $e_i = (0, \ldots, 1_i, \ldots, 0)$, $i \leq n$.

Die duale Basis ist festgelegt durch

$$e_i^*(e_j) = \delta_{ij} \quad \text{(Kronecker-Symbol)},$$

und die Matrix von e_i^* bezüglich der kanonischen Basen ist

$$\mathrm{Mat}_{E,1}(e_i^*) = e_i^t = \begin{pmatrix} 0 \\ \vdots \\ 1_i \\ \vdots \\ 0 \end{pmatrix}.$$

Damit haben wir einen R-Homomorphismus

$$\begin{array}{rcccl} R^m \to (R^m)^*, & & e_i & \mapsto & e_i^*, \\ & & (r_1,\ldots,r_m) & \mapsto & \sum r_i e_i^* = \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}. \end{array}$$

(2) Über einem Ring R sei $f : R^m \to R^n$ gegeben durch $\mathrm{Mat}_E E(f) = A \in R^{(m,n)}$, also

$$f : R^m \to R^n, \quad (r_1,\ldots,r_m) \mapsto (r_1,\ldots,r_m)A.$$

Dann wird die duale Abbildung beschrieben durch

$$f^t : (R^n)^* \to (R^m)^*, \quad (s_1,\ldots,s_n) \mapsto (s_1,\ldots,s_n)A^t.$$

Symmetrisch zu den Bildungen für Linksmoduln wird auch der Dualraum zu einem R-Rechtsmodul K als der R-Linksmodul

$$K^* = \mathrm{Hom}_R(K, R)$$

definiert.

23.5 Bidualraum
Ist M ein Linksmodul, so können wir also zu dem R-Rechtsmodul M^* wiederum den Dualraum bilden,

$$M^{**} = (M^*)^* = \mathrm{Hom}_R\Big(\mathrm{Hom}_R(M, R), R\Big),$$

den wir als den *Bidualraum* von M bezeichnen.

Für nicht kommutatives R haben M und M^* wenig miteinander zu tun, da M ein Links- und M^* ein Rechtsmodul über R ist.

M^{**} dagegen ist wieder ein R-Linksmodul, der durch einen Homomorphismus eng mit M verbunden ist. Die Elemente von M^{**} sind ja Linearformen $M^* \to R$, und für jedes $m \in M$ ist die Auswertung von $f \in M^*$, $f \mapsto f(M)$ ein Element aus M^{**}. Somit haben wir den Auswertungshomomorphismus

$$\Phi_M : M \to M^{**}, \quad m \mapsto \Big[f \mapsto f(m)\Big].$$

Es ist leicht nachzuprüfen, daß Φ_M ein Homomorphismus von R-Linksmoduln ist. Bei freien Moduln werden wir später noch mehr dazu sagen.

Sei $h : M \to N$ ein Homomorphismus von R-Linksmoduln und dazu $h^t : N^* \to M^*$ die Transponierte. Auch zu h^t können wir die Transponierte bilden und erhalten

$$h^{tt} = (h^t)^t : M^{**} \to N^{**}, \quad g \mapsto g \circ h^t.$$

Dies ist ein Homomorphismus von R-Linksmoduln, für den folgende bemerkenswerte Beziehung gilt:

23.6 Eigenschaft des Auswertungsmorphismus
Zu jedem Homomorphismus $h : M \to N$ von R-Linksmoduln ist folgendes Diagramm kommutativ:

$$\begin{array}{ccc} M & \xrightarrow{h} & N \\ \Phi_M \downarrow & & \downarrow \Phi_N \\ M^{**} & \xrightarrow{h^{tt}} & N^{**} \end{array}$$

Beweis: Für $m \in M$ ergibt $\Phi_N \circ h(m)$ die Linearform

$$N^* \to R, \quad \alpha \mapsto \alpha(h(m)).$$

Andererseits entspricht $h^{tt} \circ \Phi_M(m)$ der Abbildung

$$N^* \xrightarrow{h^t} M^* \xrightarrow{\Phi_M(m)} R, \quad \alpha \mapsto \alpha \circ h(m) = \alpha(h(m)).$$

Also ist das Diagramm kommutativ. □

Für freie Moduln M hat Φ_M besonders schöne Eigenschaften:

23.7 Der Bidualraum von freien Moduln
Für einen freien R-Modul M gilt:

(1) *$\Phi_M : M \to M^{**}$ ist injektiv.*

(2) *Ist M zudem endlich erzeugbar, so ist Φ_M ein Isomorphismus.*

Beweis: (1) Sei $\{m_\lambda \mid \lambda \in \Lambda\}$ Basis von M und $\{m_\lambda^* \mid \lambda \in \Lambda\} \subset M^*$ die duale Basis. Nehmen wir an, daß $m = \sum\limits_{\lambda \in \Lambda_0} r_\lambda m_\lambda$ im Kern von Φ_M liegt, also insbesondere

$$0 = m_\lambda^*(m) = r_\lambda \text{ für alle } \lambda \in \Lambda_0.$$

Damit sind alle $r_\lambda = 0$ und auch $m = 0$, d.h. Φ_M ist injektiv.

(2) Sei nun Λ endlich. Aus den Beziehungen

$$\Phi_M(m_\lambda)(m_\mu^*) = m_\mu^*(m_\lambda) = \begin{cases} 0 & \text{falls } \lambda \neq \mu \\ 1 & \text{falls } \lambda = \mu, \end{cases}$$

ersehen wir, daß $\{\Phi_M(m_\lambda) \mid \lambda \in \Lambda\}$ die zu $\{m_\lambda^* \mid \lambda \in \Lambda\}$ duale Basis von M^{**} bildet. Somit ist Φ_M ein Isomorphismus. □

23.8 Aufgaben

(1) Seien $f_1, \ldots, f_m$ Linearformen eines n-dimensionalen Vektorraumes. Zeigen Sie: Für $U := \operatorname{Kern} f_1 \cap \ldots \cap \operatorname{Kern} f_m$ gilt $\dim U \geq n - m$.

(2) f und g seien Linearformen auf einem Vektorraum V über einem Körper K mit $f \neq 0$. Zeigen Sie, daß folgende Aussagen äquivalent sind:

(a) *es gibt ein $c \in K \setminus \{0\}$ mit $g = cf$;*

(b) $\operatorname{Kern} f = \operatorname{Kern} g$.

(3) Seien $f : M \to N$ ein Homomorphismus von Vektorräumen über einem Körper K und $f^t : N^* \to M^*$ die zu f transponierte Abbildung.
Beweisen Sie, daß f genau dann surjektiv ist, wenn f^t injektiv ist.

(4) Zu den folgenden Basen von $\operatorname{Pol}_3(\mathbb{R}) = \{p \in \mathbb{R}[X] \mid grad(p) \leq 3\}$ bestimme man jeweils die duale Basis von $\operatorname{Pol}_3(\mathbb{R})^*$:

(i) $B := (1, x, x^2, x^3)$

(ii) $C := (1, 1 + x, 2x^2 - 1, 4x^3 - 3x)$

(5) Seien V ein $\mathbb{R}$-Vektorraum und $f, g \in V^*$ Linearformen. Zeigen Sie:
Ist $h : V \to \mathbb{R}$ mit $h(v) := f(v) \cdot g(v)$ eine Linearform auf V, so ist $f = 0$ oder $g = 0$.

24 Tensorprodukt

In diesem Abschnitt betrachten wir einige fundamentale Bildungen über (nichtkommutativen) Ringen, die von allgemeinem Interesse sind. Als Anwendung werden wir dann Aussagen und Sätze für *bilineare Abbildungen* von Moduln über kommutativen Ringen ableiten.

24.1 Definition
Seien M_R ein Rechtsmodul, ${}_RN$ ein Linksmodul über dem Ring R und G eine abelsche Gruppe ($\mathbb{Z}$-Modul). Eine Abbildung

$$\beta : M \times N \to G$$

heißt *R-balanciert*, wenn gilt

$$\begin{aligned} \beta(m+m',n) &= \beta(m,n)+\beta(m',n), \\ \beta(m,n+n') &= \beta(m,n)+\beta(m,n'), \\ \beta(mr,n) &= \beta(m,rn), \end{aligned}$$

für alle $m, m' \in M$, $n, n' \in N$, $r \in R$. β ist also $\mathbb{Z}$-linear in jeder der beiden Komponenten, und es gilt zusätzlich eine Vertauschungseigenschaft für Elemente aus R.

Die Menge aller balancierten Abbildungen $M \times N \to G$ bezeichnen wir mit $\mathrm{Bal}_R(M \times N, G)$.

24.2 Beispiele

(1) Die Multiplikation in R ergibt eine balancierte Abbildung

$$\mu : R \times R \to R, \quad (r,s) \mapsto rs.$$

Die Rechengesetze in R entsprechen gerade den Bedingungen, die μ balanciert machen.

(2) Für einen R-Modul M mit Dualraum M^* ist

$$\mu : M^* \times M \to R, \quad (f,m) \mapsto f(m),$$

eine balancierte Abbildung (*kanonische Abbildung*).

Die nachfolgende Begriffsbildung dient dazu, die balancierten Abbildungen durch gewisse lineare Abbildungen darzustellen. Wir definieren das *Tensorprodukt* von M und N als abelsche Gruppe mit einer universellen Eigenschaft bezüglich aller R-balancierten Abbildungen auf $M \times N$.

24.3 Definition
Seien M_R ein Rechtsmodul, ${}_RN$ ein Linksmodul über dem Ring R. Eine abelsche Gruppe T zusammen mit einer R-balancierten Abbildung

$$\tau : M \times N \to T$$

heißt *Tensorprodukt von M und N*, wenn sich jede R-balancierte Abbildung $\beta : M \times N \to G$, G abelsche Gruppe, eindeutig über T faktorisieren läßt, d.h. es gibt genau einen Gruppenhomomorphismus $\gamma : T \to G$, der folgendes Diagramm kommutativ macht:

$$\begin{array}{ccc} M \times N & \xrightarrow{\tau} & T \\ & \beta \searrow \quad \swarrow \gamma & \\ & G & \end{array}$$

Noch wissen wir nicht, ob es eine Gruppe T mit diesen Eigenschaften gibt. Dennoch können wir uns schon überlegen, daß es bis auf Isomorphie höchstens eine solche Gruppe geben kann.

24.4 Eindeutigkeit des Tensorprodukts
Sind $\tau_1 : M \times N \to T_1$ und $\tau_2 : M \times N \to T_2$ zwei Tensorprodukte von M_R und ${}_RN$, dann gibt es einen Isomorphismus $\gamma : T_1 \to T_2$ mit $\gamma\tau_1 = \tau_2$.

Beweis: Wir haben das Diagramm

$$\begin{array}{ccc} M \times N & \xrightarrow{\tau_1} & T_1 \\ & \tau_2 \searrow & \\ & & T_2. \end{array}$$

Wegen der universellen Eigenschaft von T_1 gibt es ein $\gamma : T_1 \to T_2$ mit $\gamma\tau_1 = \tau_2$. Aus Symmetriegründen können wir auch ein $\delta : T_2 \to T_1$ mit $\delta\tau_2 = \tau_1$ finden. Also haben wir

$$\tau_1 = \delta\tau_2 = \delta\gamma\tau_1.$$

Das Diagramm

$$\begin{array}{ccc} M \times N & \xrightarrow{\tau_1} & T_1 \\ & \tau_1 \searrow & \\ & & T_1 \end{array}$$

kann nach Definition von T_1 nur durch *einen* Homomorphismus $T_1 \to T_2$ kommutativ ergänzt werden. Nun erfüllen aber sowohl id_{T_1} als auch $\delta\gamma$ diese Bedingung, also gilt $\mathrm{id}_{T_1} = \delta\gamma$.

Analog folgert man $\mathrm{id}_{T_2} = \gamma\delta$, und somit ist γ (und δ) ein Isomorphismus. □

Zeigen wir nun, daß es tatsächlich eine Gruppe mit den in 24.3 geforderten Eigenschaften gibt:

24.5 Existenz des Tensorprodukts
In den R-Moduln M_R, ${}_RN$ nehmen wir $M \times N$ als Indexmenge und bilden den freien $\mathbb{Z}$-Modul

$$F = \mathbb{Z}^{(M\times N)}.$$

Zu $(m, n) \in M \times N$ bezeichnen wir das entsprechende Basiselement in F mit $[m, n]$. Damit gilt

$$F = \bigoplus_{M\times N} \mathbb{Z}[m, n].$$

In F bilden wir die Elemente

$$\begin{aligned}&[m_1 + m_2, n] - [m_1, n] - [m_2, n],\\ &[m, n_1 + n_2] - [m, n_1] - [m, n_2],\\ &[mr, n] - [m, rn].\end{aligned}$$

Bezeichne K den Untermodul von F, der von allen diesen Elementen mit $m_1, m_2, m \in M$, $n_1, n_2, n \in N$ und $r \in R$ erzeugt wird.

Wir setzen nun $M \otimes_R N = F/K$ und betrachten die Komposition der kanonischen Abbildungen $M \to F \to F/K$, nämlich

$$\tau : M \times N \to M \otimes_R N, \quad (m, n) \mapsto m \otimes n := [m, n] + K.$$

Nach der Definition von K ist τ offensichtlich R-balanciert.

Sei $\beta : M \times N \to G$ eine R-balancierte Abbildung. Durch Vorgabe der Werte auf der Basis definieren wir einen $\mathbb{Z}$-Homomorphismus

$$\tilde{\gamma} : F \to G, \quad [m, n] \mapsto \beta(m, n).$$

Dabei gilt etwa für die Bilder von $[mr, n]$ und $[m, rn]$

$$\tilde{\gamma}([mr, n]) = \beta(mr, n) = \beta(m, rn) = \tilde{\gamma}([m, rn]).$$

Ähnlich sieht man, daß in der Tat $K \subset \operatorname{Kern} \tilde{\gamma}$ gilt. Somit faktorisiert $\tilde{\gamma}$ über τ, d.h. wir haben das kommutative Diagramm

$$\begin{array}{ccccc} M \times N & \longrightarrow & F & \longrightarrow & M \otimes_R N \\ & \beta \searrow & \downarrow \tilde{\gamma} & \swarrow \gamma & \\ & & G & & \end{array}$$

Die Abbildung $\tau : M \times N \to M \otimes_R N$ ist zwar nicht surjektiv, ihr Bild

$$\operatorname{Bild} \tau = \{m \otimes n \mid m \in M,\, n \in N\}$$

ist jedoch ein Erzeugendensystem von $M \otimes_R N$ als $\mathbb{Z}$-Modul.

Da im obigen Diagramm die $\gamma(m \otimes n) = \beta(m, n)$ eindeutig bestimmt sind, ist auch γ eindeutig festgelegt. Somit ist $\tau : M \times N \to M \otimes_R N$ ein Tensorprodukt von M und N.

Man beachte, daß jedes Element aus $M \otimes_R N$ als endliche Summe

$$m_1 \otimes n_1 + \ldots + m_k \otimes n_k$$

dargestellt werden kann. Die Darstellung ist jedoch nicht eindeutig.

Sind $\beta_1, \beta_2 : M \times N \to G$ zwei balancierte Abbildungen, so ist mit der üblichen Addition auch

$$\beta_1 + \beta_2 : \quad M \times N \quad \to \quad G, \quad (m, n) \quad \mapsto \quad \beta_1(m, n) + \beta_2(m, n),$$

eine balancierte Abbildung. Also ist $\mathrm{Bal}_R(M \times N, G)$ ein $\mathbb{Z}$-Modul.

Andererseits ist für jedes $\beta \in \mathrm{Bal}_R(M \times N, G)$ und $n \in N$ die Zuordnung

$$\beta(-, n) : M \to G, \quad m \mapsto \beta(m, n),$$

ein $\mathbb{Z}$-Homomorphismus, also aus $\mathrm{Hom}_{\mathbb{Z}}(M, G)$.

Wir betrachten $\mathrm{Hom}_{\mathbb{Z}}(M, G)$ als R-Linksmodul mit der Operation

$$rf(m) = f(rm) \text{ für } r \in R,\ m \in M.$$

Damit ergeben sich aus den obigen Definitionen folgende Zusammenhänge zwischen balancierten Abbildungen und Homomorphismen, die sich als sehr nützlich erweisen werden:

24.6 Hom-Tensor-Relation

Seien M_R ein Rechtsmodul, ${}_RN$ ein Linksmodul über dem Ring R, G eine abelsche Gruppe und $\tau : M \times N \to M \otimes_R N$ die kanonische Abbildung.

Dann sind folgende Abbildungen Isomorphismen von $\mathbb{Z}$-Moduln:

$$\begin{array}{rcl}
\mathrm{Hom}_{\mathbb{Z}}(M \otimes_R N, G) & \xrightarrow{\psi_1} & \mathrm{Bal}_R(M \times N, G), \\
\alpha & \longmapsto & \alpha\tau. \\
\mathrm{Bal}_R(M \times N, G) & \xrightarrow{\psi_2} & \mathrm{Hom}_R(N, \mathrm{Hom}_{\mathbb{Z}}(M, G)), \\
\beta & \longmapsto & [n \to \beta(-, n)]. \\
\mathrm{Hom}_{\mathbb{Z}}(M \otimes_R N, G) & \xrightarrow{\psi_M} & \mathrm{Hom}_R(N, \mathrm{Hom}_{\mathbb{Z}}(M, G)), \\
\delta & \longmapsto & [n \mapsto \delta(- \otimes n)].
\end{array}$$

Beweis: Es ist leicht nachzuprüfen, daß die angegebenen Abbildungen in dem angegebenen Sinn definiert und $\mathbb{Z}$-linear sind.

Die Bijektivität von ψ_1 folgt aus der definierenden Eigenschaft des Tensorprodukts.

Für jedes $\varphi \in \mathrm{Hom}_R(N, \mathrm{Hom}_{\mathbb{Z}}(M, G))$ ist

$$\widetilde{\varphi} : M \times N \to G, \quad (m, n) \mapsto \varphi(n)[n],$$

eine R-balancierte Abbildung, und die Zuordnung $\varphi \mapsto \widetilde{\varphi}$ ist invers zu ψ_2.

ψ_M ist gerade die Komposition von ψ_1 und ψ_2. □

Wir kommen nun zu weiteren Konstruktionen mit dem Tensorprodukt.

24.7 Tensorprodukt von Homomorphismen
Sind $f : M_R \to M'_R$, $g : {}_RN \to {}_RN'$ R-Homomorphismen, so gibt es genau eine $\mathbb{Z}$-lineare Abbildung

$$f \underline{\otimes} g : M \otimes_R N \to M' \otimes_R N'$$

mit $f \underline{\otimes} g(m \otimes n) = f(m) \otimes g(n)$, $m \in M$, $n \in N$.

Man nennt $f \underline{\otimes} g$ das Tensorprodukt *von f und g. Dafür gilt:*

(1) $\mathrm{id}_M \underline{\otimes} \mathrm{id}_N = \mathrm{id}_{M \otimes N}$.

(2) $f \underline{\otimes}(g_1 + g_2) = f \underline{\otimes} g_1 + f \underline{\otimes} g_2$, $(f_1 + f_2) \underline{\otimes} g = f_1 \underline{\otimes} g + f_2 \underline{\otimes} g$.

(3) *Für R-Homomorphismen $h : M'_R \to M''_R$, $k : {}_RN' \to {}_RN''$ ist*

$$(h \underline{\otimes} k) \circ (f \underline{\otimes} g) = (h \circ f) \underline{\otimes}(k \circ g).$$

Beweis: Wir definieren die Abbildung

$$f \times g : M \times N \to M' \otimes_R N', \quad (m, n) \mapsto f(m) \otimes g(n).$$

Diese ist $\mathbb{Z}$-bilinear und R-balanciert, da

$$f(mr) \otimes g(n) = f(m) \otimes g(rn).$$

Also faktorisiert $f \times g$ über $M \otimes_R N$, und wir erhalten die gewünschte Abbildung $f \underline{\otimes} g$. □

Bemerkung: Die Unterscheidung zwischen $\underline{\otimes}$ und $\otimes$ ist bei der Definition hilfreich. In der Praxis schreibt man aber meist auch $\otimes$ an Stelle von $\underline{\otimes}$.

24.8 Tensorprodukt und direkte Summen
Seien M_R und ${}_RN$ R-Moduln. Ist $N = N_1 \oplus \ldots \oplus N_k$, dann gilt

$$M \otimes_R N \simeq \bigoplus_{i=1}^{k} M \otimes_R N_i.$$

Man sagt, *das Tensorprodukt ist mit direkten Summen vertauschbar.*

Beweis: Wir zeigen die Behauptung für $N = N_1 \oplus N_2$. Der allgemeine Fall ergibt sich daraus durch Induktion.

$$\begin{array}{ll} \text{Bezeichne} & \varepsilon_1 : N_1 \to N, \quad \varepsilon_2 : N_2 \to N \\ & \pi_1 : N \to N_1, \quad \pi_2 : N \to N_2 \end{array}$$

die kanonischen Injektionen und Projektionen. Dann ist

$$\varepsilon_1\pi_1 + \varepsilon_2\pi_2 = \mathrm{id}_N \text{ und}$$
$$\mathrm{id}_{M\otimes N} = \mathrm{id}_M \otimes \mathrm{id}_N = \mathrm{id}_M \otimes(\varepsilon_1\pi_1 + \varepsilon_2\pi_2) = \mathrm{id}_M \otimes\varepsilon_1\pi_1 + \mathrm{id}_M\, \varepsilon_2\pi_2.$$

$\mathrm{id}_M \otimes\varepsilon_1\pi_1$ und $\mathrm{id}_M \otimes\varepsilon_2\pi_2$ sind orthogonale Idempotente im Endomorphismenring $\mathrm{End}_{\mathbb{Z}}(M \otimes_R N)$, und somit gilt

$$M \otimes_R N = [\mathrm{id}_M \otimes\varepsilon_1\pi_1 + \mathrm{id}_M \otimes\varepsilon_2\pi_2](M \otimes N) \simeq M \otimes_R N_1 \oplus M \otimes_R N_2.$$

□

Bemerkung: Es ist lediglich eine schreibtechnische Übung, zu zeigen, daß die Aussage in 24.8 auch für unendliche direkte Summen gilt.

24.9 Beispiele

(1) Für einen Linksmodul M ist die Abbildung

$$\mu : M^* \times M \to R, \quad (f, m) \mapsto f(m),$$

balanciert, sie faktorisiert also über einen $\mathbb{Z}$-Homomorphismus

$$\overline{\mu} : M^* \otimes_R M \to R.$$

(2) Für Linksmoduln M, N ist die Abbildung

$$\vartheta_{M,N} : M^* \times N \to \mathrm{Hom}_R(M, N), \quad (f, n) \mapsto [m \mapsto f(m) \cdot n]$$

R-balanciert, denn nach Definition der R-Modulstruktur von M^* gilt für $r \in R$

$$(fr)(m) \cdot n = f(rm) \cdot n.$$

Somit faktorisiert sie über einen Gruppenhomomorphismus

$$\overline{\vartheta}_{M,N} : M^* \otimes_R N \longrightarrow \mathrm{Hom}_R(M, N).$$

Ist M endlich erzeugt und frei, so ist $\overline{\vartheta}_{M,N}$ ein Isomorphismus.

Beweis: Seien $x_1, \ldots, x_m$ eine Basis von M und $x_1^*, \ldots, x_m^*$ die dazu duale Basis von M^*. Wir definieren eine Abbildung

$$\gamma_{M,N} : \mathrm{Hom}_R(M,N) \to M^* \otimes_R N, \quad f \mapsto \sum_{i=1}^{m} x_i^* \otimes f(x_i),$$

und zeigen, daß diese invers zu $\overline{\vartheta}_{M,N}$ ist.
Für $x = \sum\limits_{i=1}^{m} r_i\, x_i \in M$ $(r_i \in R)$ gilt nach Definition der dualen Basis:

$$\overline{\vartheta}_{M,N}\left(\sum_{i=1}^{m} x_i^* \otimes f(x_i)\right)(x) = \sum_{i=1}^{m} x_i^*(x) \cdot f(x_i) = \sum_{i=1}^{m} r_i f(x_i) = f(x).$$

Also gilt $\overline{\vartheta}_{M,N} \circ \gamma_{M,N}(f) = f$ und $\overline{\vartheta}_{M,N} \circ \gamma_{M,N} = \mathrm{id}_{M^* \otimes_R N}$.

Mit ähnlichen Argumenten erhält man auch $\gamma_{M,N}\overline{\vartheta}_{M,N} = \mathrm{id}_{\mathrm{Hom}_R(M,N)}$. □

(3) Sei M ein endlich erzeugter, freier R-Linksmodul. Dann können wir die in (1) und (2) gewonnenen Abbildungen hintereinander ausführen und erhalten

$$\mathrm{Sp} : \mathrm{End}_R(M) \xrightarrow{\overline{\gamma}_{M,M}} M^* \otimes M \xrightarrow{\mu} R,$$

die sogenannte *Spurform* auf $\mathrm{End}_R(M)$.

Ist $f \in \mathrm{End}_R(M)$, $X = \{x_1, \ldots, x_n\}$ eine Basis von M und $\mathrm{Mat}_{XX}(f) = (a_{ij})$, so gilt

$$\begin{aligned} \mathrm{Sp}(f) \;=\; \mu(\gamma_{M,M}(f)) \;&=\; \overline{\mu}\left(\sum_{i=1}^{n} x_i^* \otimes f(x_i)\right) \;=\; \overline{\mu}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} x_i^* \otimes a_{ij}x_j\right) \\ &=\; \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i^*(x_j) \;=\; \sum_{i=1}^{n} a_{ii}. \end{aligned}$$

Somit ist $\mathrm{Sp}(f)$ gerade die Spur der Matrix $\mathrm{Mat}_{XX}(f)$, unabhängig von der gewählten Basis X von M.

Über kommutativen Ringen R ist das Tensorprodukt nicht nur ein $\mathbb{Z}$-Modul (abelsche Gruppe), sondern erlaubt zudem eine R-Modulstruktur:

24.10 Tensorprodukt über kommutativen Ringen
Sei R ein kommutativer Ring, und M, N, L seien R-Moduln. Dann gilt:

(1) *$M \otimes_R N$ ist ein R-Modul.*

(2) *$M \otimes_R N \simeq N \otimes_R M$ und $(M \otimes_R N) \otimes_R L \simeq M \otimes_R (N \otimes_R L)$.*

Beweis: (1) Für $r \in R$ ist die Abbildung $r : M \to M$, $m \mapsto rm$, ein R-Homomorphismus. Nach 24.7 gibt es daher einen Homomorphismus

$$r \otimes \mathrm{id} : M \otimes N \to M \otimes N, \quad m \otimes n \mapsto (rm) \otimes n.$$

Setzen wir $r(m \otimes n) := r \otimes \mathrm{id}(m \otimes n) = (rm) \otimes n$, so wird $M \otimes_R N$ zu einem R-Modul.

(2) Die Abbildung $M \times N \to N \otimes_R M$, $(m, n) \mapsto n \otimes m$, ist balanciert und faktorisiert daher über $M \otimes_R N$. Entsprechend erhalten wir eine Abbildung in der umgekehrten Richtung, und die Komposition der beiden ergibt die Identität.

Zur Assoziativität: Zu $l \in L$ definieren wir zunächst

$$f_l : N \to N \otimes_R L, \quad n \mapsto n \oplus l.$$

Dazu bilden wir $\mathrm{id}_M \otimes f_l : M \otimes_R N \to M \otimes_R (N \otimes_R L)$
und dann

$$\beta : (M \otimes_R N) \times L \to M \otimes_R (N \otimes_R L), \quad (m \otimes n, l) \mapsto \mathrm{id}_M \otimes f_l(m \otimes n).$$

Man bestätigt leicht, daß β R-balanciert ist, und wir erhalten eine Abbildung

$$\bar{\beta} : (M \otimes_R N) \otimes_R L \to M \otimes_R (N \otimes_R L), \quad (m \otimes n) \otimes l \mapsto m \otimes (n \otimes l).$$

Aus Symmetriegründen erhalten wir auch eine Abbildung in der umgekehrten Richtung, die zu $\bar{\beta}$ invers ist. □

Wir wollen nun das Tensorprodukt von freien Moduln genauer ansehen. Sei M ein R-Modul und N ein freier R-Modul mit einem Basiselement $y \in N$, also $N = Ry$. Dann läßt sich jedes Element in $M \otimes_R Ry$ als Summe der Form $m \otimes ry$ mit $m \in M$ und $r \in R$ schreiben. Wegen $m \otimes ry = mr \otimes y$ läßt sich also jedes Element aus $M \otimes_R Ry$ darstellen durch

$$\sum_{i=1}^{k} (m_i \otimes n) = \left(\sum_{i=1}^{k} m_i \right) \otimes n, \quad m_i \in M.$$

Somit ist jedes Element vom Typ $m \otimes y$ für ein geeignetes $m \in M$. Die Abbildung

$$M \times Ry \to M, \quad (m, ry) \mapsto rm,$$

ist R-balanciert, induziert also einen $\mathbb{Z}$-Homomorphismus

$$M \otimes_R Ry \to M, \quad m \otimes ry \mapsto rm.$$

Andererseits haben wir einen $\mathbb{Z}$-Homomorphismus

$$M \to M \otimes_R Ry, \quad m \mapsto m \otimes y.$$

Diese Abbildungen sind zueinander invers, und daher gilt

$$M \otimes_R Ry \simeq M.$$

Jedes Element aus $M \otimes_R Ry$ hat eine eindeutige Darstellung $m \otimes y$, $m \in M$.

Ausgehend von dieser Beobachtung, können wir zeigen:

24.11 Tensorprodukt von freien Moduln

Seien M ein Rechtsmodul und N ein freier Linksmodul über dem Ring R mit Basis $\{y_1, \ldots, y_n\}$. Dann gilt:

(1) *Jedes Element von $M \otimes_R N$ läßt sich eindeutig darstellen als*

$$\sum_{i=1}^{n} m_i \otimes y_i, \quad m_i \in M.$$

(2) *Ist R kommutativ und $\{x_1, \ldots, x_m\}$ eine Basis von M, dann ist*

$$\{x_i \otimes y_j \mid i \leq m,\, j \leq n\}$$

eine Basis von $M \otimes_R N$.

(3) *Ist R ein Körper, dann ist*

$$\dim(M \otimes_R N) = \dim(M) \cdot \dim(N).$$

Beweis: (1) Da das Tensorprodukt mit direkten Summen vertauschbar ist (siehe 24.8), folgt aus $N = Ry_1 \oplus \ldots \oplus Ry_n$,

$$M \otimes_R N = (M \otimes_R Ry_1) \oplus \ldots \oplus (M \otimes_R Ry_n).$$

In der Vorbereitung haben wir $M \otimes Ry \simeq M$ gezeigt und dafür die gewünschte Darstellung angegeben.

(2) Wende (1) analog auf $M \otimes Ry_i$ an.

(3) Ist $\dim(M)$ unendlich, so gilt dies auch für $\dim(M \otimes_R N)$. Damit folgt die Behauptung aus (2). □

Beispiel

Seien $f : \mathbb{R}^3 \to \mathbb{R}^4$ und $g : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch

$$A = Mat_{E_3,E_4}(f) = \begin{pmatrix} 1 & 0 & 3 & 2 \\ 4 & 1 & 1 & 0 \\ 2 & -1 & 4 & 1 \end{pmatrix}, \quad B = Mat_{E_2,E_2}(g) = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Wir wollen die Matrix zu $f \otimes g : \mathbb{R}^3 \otimes \mathbb{R}^2 \to \mathbb{R}^4 \otimes \mathbb{R}^2$ bezüglich der kanonischen Basen berechnen.

Als Basis von $\mathbb{R}^3 \otimes \mathbb{R}^2$ erhalten wir die Vektoren

$$\begin{aligned} v_1 &= (1,0,0) \otimes (1,0), & v_2 &= (1,0,0) \otimes (0,1), \\ v_3 &= (0,1,0) \otimes (1,0), & v_4 &= (0,1,0) \otimes (0,1), \\ v_5 &= (0,0,1) \otimes (1,0), & v_6 &= (0,0,1) \otimes (0,1), \end{aligned}$$

und als Basis von $\mathbb{R}^4 \otimes \mathbb{R}^2$:

$$\begin{aligned} w_1 &= (1,0,0,0) \otimes (1,0), & w_2 &= (1,0,0,0) \otimes (0,1), \\ & & \vdots & \\ w_7 &= (0,0,0,1) \otimes (1,0), & w_8 &= (0,0,0,1) \otimes (0,1). \end{aligned}$$

Damit berechnen wir nun

$$\begin{aligned} f \underline{\otimes} g(v_1) &= (1,0,3,2) \otimes (1,-1) \\ &= \Big(1(1,0,0,0) + 3(0,0,1,0) + 2(0,0,0,1)\Big) \otimes (1,-1) \\ &= (1,0,0,0) \otimes (1,-1) + 3 \cdot (0,0,1,0) \otimes (1,-1) \\ &\quad + 2 \cdot (0,0,0,1) \otimes (1,-1) \\ &= \quad (1,0,0,0) \otimes \Big((1,0) - (0,1))\Big) \\ &\quad + 3 \cdot (0,0,1,0) \otimes \Big((1,0) - (0,1))\Big) \\ &\quad + 2 \cdot (0,0,0,1) \otimes \Big((1,0) - (0,1))\Big) \\ &= \quad 1 \cdot (1,0,0,0) \otimes (1,0) - 1 \cdot (1,0,0,0) \otimes (0,1) \\ &\quad + 3 \cdot (0,0,1,0) \otimes (1,0) - 3 \cdot (0,0,1,0) \otimes (0,1) \\ &\quad + 2 \cdot (0,0,0,1) \otimes (1,0) - 2 \cdot (0,0,0,1) \otimes (0,1) \\ &= \quad 1 \cdot w_1 - 1 \cdot w_2 + 0 \cdot w_3 + 0 \cdot w_4 \\ &\quad + 3 \cdot w_5 - 3 \cdot w_6 + 2 \cdot w_7 - 2 \cdot w_8 \end{aligned}$$

Durch analoge Rechnung bekommt man

$$\begin{aligned} f \underline{\otimes} g(v_2) &= (1,0,3,2) \otimes (2,1) \\ &= \quad 2 \cdot w_1 + 1 \cdot w_2 + 0 \cdot w_3 + 0 \cdot w_4 \\ &\quad + 6 \cdot w_5 + 3 \cdot w_6 + 4 \cdot w_7 + 2 \cdot w_8 \end{aligned}$$

Damit finden wir schließlich bezüglich der kanonischen Basen

$$\begin{aligned}\mathrm{Mat}(f \underline{\otimes} g) &= \begin{pmatrix} 1 & -1 & 0 & 0 & 3 & -3 & 2 & -2 \\ 2 & 1 & 0 & 0 & 6 & 3 & 4 & 2 \\ 4 & -4 & 1 & -1 & 1 & -1 & 0 & 0 \\ 8 & 4 & 2 & 1 & 2 & 1 & 0 & 0 \\ 2 & -2 & -1 & 1 & 4 & -4 & 1 & -1 \\ 4 & 2 & -2 & -1 & 8 & 4 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a_{11} \cdot B & a_{12} \cdot B & a_{13} \cdot B & a_{14} \cdot B \\ a_{21} \cdot B & a_{22} \cdot B & a_{23} \cdot B & a_{24} \cdot B \\ a_{31} \cdot B & a_{32} \cdot B & a_{33} \cdot B & a_{34} \cdot B \end{pmatrix}.\end{aligned}$$

Diese Matrix bezeichnet man auch als das *Kroneckerprodukt* von A und B.

24.12 Aufgaben

(1) Seien R ein Ring, M_R ein Rechts- und ${}_RN$ ein Linksmodul. Man zeige: Sind $f : M_R \to M_R$, $g : {}_RN \to {}_RN$ epimorph, so ist auch $f \underline{\otimes} g$ epimorph.

(2) R sei ein Unterring eines Ringes S und ${}_RM$ ein Linksmodul über R. Beweisen Sie, daß es eine S-Linksmodulstruktur auf $S \otimes_R M$ gibt mit

$$s(t \otimes m) = (st) \otimes m \text{ für alle } s, t \in S,\ m \in M.$$

(3) R sei ein Ring. Zeigen Sie: Für jeden R-Linksmodul ${}_RM$ gibt es einen R-Isomorphismus $\mu_M : R \otimes_R M \to M$.

(4) R sei ein kommutativer Integritätsring. M und N seien freie R-Moduln mit Basen $(m_1, \ldots, m_k)$ bzw. $(n_1, \ldots, n_l)$, $k, l \in \mathbb{N}$, und $\tau : M \times N \to M \otimes_R N$ die kanonische Abbildung. Man zeige:

(i) *Ist $k \geq 2$ und $l \geq 2$, so ist $m_1 \otimes n_2 + m_2 \otimes n_1 \notin \mathrm{Bild}\,\tau$.*

(ii) *τ ist surjektiv $\Leftrightarrow k = 1$ oder $l = 1$.*

(5) V und W seien endlich-dimensionale Vektorräume über einem Körper K. Man beweise:

(i) *Für alle $v \in V$, $w \in W$ gilt: $v \otimes_K w = 0 \Leftrightarrow v = 0$ oder $w = 0$.*

(ii) *$V \otimes_K W = 0 \Leftrightarrow V = 0$ oder $W = 0$.*

(6) $f : V \to V$ und $g : W \to W$ seien Endomorphismen von Moduln über einem kommutativen Ring R. Beweisen Sie:

(i) *Ist f nilpotent, so ist $f \underline{\otimes} g$ nilpotent.*

Ist R ein Körper, und sind V, W endlich dimensional, so gilt:

(ii) *Ist $f \underline{\otimes} g$ nilpotent, so ist f nilpotent oder g nilpotent.*

(iii) *Ist $\lambda \in R$ ein Eigenwert von f und $\mu \in R$ ein Eigenwert von g, so ist $\lambda \cdot \mu$ ein Eigenwert von $f \underline{\otimes} g$.*

25 Bilinearformen

Wie wir am Ende des letzten Abschnitts schon festgestellt haben, erhalten wir über kommutativen Ringen noch weitere Bildungen mit dem Tensorprodukt. Diese werden im folgenden eine Rolle spielen.

Wir setzen in diesem Abschnitt voraus, daß R ein kommutativer Ring ist.

25.1 Definition
M, N und L seien R-Moduln. Eine Abbildung

$$\beta : M \times N \to L$$

heißt *bilinear*, wenn β R-balanciert ist und zudem für alle $m \in M$, $n \in N$ und $r \in R$ gilt

$$\beta(rm, n) = r\beta(m, n) \quad (= \beta(m, rn)).$$

Eine bilineare Abbildung auf $M \times N$ ist also R-linear in jeder der beiden Komponenten.

Die Menge der bilinearen Abbildungen $M \times N \to L$ bezeichnen wir mit $\mathrm{Bil}_R(M \times N, L)$.

Die bilinearen Abbildungen in den Grundring, $M \times N \to R$, nennt man *Bilinearformen* auf $M \times N$ und bezeichnet sie mit $\mathrm{Bil}_R(M, N)$.

25.2 Beispiele (beachte R kommutativ)
(1) Die Multiplikation in R ergibt eine Bilinearform (vgl. 24.2)

$$\mu : R \times R \to R, \quad (r, s) \mapsto rs.$$

(2) Für einen R-Modul M mit Dualraum M^* ist (vgl. 24.2)

$$\mu : M^* \times M \to R, \quad (f, m) \mapsto f(m),$$

eine Bilinearform (*kanonische Bilinearform* auf $M^* \times M$).

(3) Für jede Matrix $B \in R^{(m,n)}$ ist

$$R^m \times R^n \to R, \quad (x, y) \mapsto xBy^t,$$

eine Bilinearform. Wir werden sehen, daß auf endlich erzeugten, freien Moduln alle Bilinearformen von diesem Typ sind.

25.3 Bilineare Abbildungen und Tensorprodukt
M, N und L seien R-Moduln.

(1) *Die kanonische Abbildung $\tau : M \times N \to M \otimes_R N$ ist R-bilinear.*

(2) *Jede R-bilineare Abbildung $M \times N \to L$ läßt sich über eine R-lineare Abbildung $M \otimes_R N \to L$ faktorisieren.*

Beweis: (1) Dies ergibt sich aus der Definition der Skalarmultiplikation auf $M \otimes_R N$ (vgl. 24.10).

(2) Eine R-bilineare Abbildung $\beta : M \times N \to L$ ist insbesondere balanciert, d.h. wir haben einen $\mathbb{Z}$-Homomorphismus

$$\gamma : M \otimes_R N \to L, \quad m \otimes n \mapsto \beta(m, n).$$

Dafür gilt nun

$$\gamma(r(m \otimes n)) = \gamma((rm) \otimes n) = \beta(rm, n) = r\beta(m, n) = r\gamma(m \otimes n),$$

d.h. γ ist ein R-Homomorphismus. □

Sind $\beta_1, \beta_2 : M \times N \to L$ bilinear und $r \in R$, so sind auch mit der üblichen Addition und Skalarmultiplikation die Abbildungen

$$\begin{array}{rrcl} \beta_1 + \beta_2 : & M \times N & \to & R, \quad (m,n) \mapsto \beta_1(m,n) + \beta_2(m,n), \text{ und} \\ r\beta_1 : & M \times N & \to & R, \quad (m,n) \mapsto r\beta_1(m,n), \end{array}$$

bilinear. Also ist $\mathrm{Bil}_R(M \times N, L)$ ein R-Modul.

Andererseits sind für $\beta \in \mathrm{Bil}_R(M \times N, L)$ und $m \in M$ bzw. $n \in N$ die Abbildungen

$$\begin{array}{l} \beta(m, -) : N \to L, \quad x \mapsto \beta(m, x), \\ \beta(-, n) : M \to L, \quad y \mapsto \beta(y, n), \end{array}$$

R-Homomorphismen, also aus $\mathrm{Hom}_R(N, L)$ bzw. $\mathrm{Hom}_R(M, L)$.

Es ist leicht nachzuprüfen, daß über kommutativen Ringen die in 24.6 angegebenen Isomorphismen von abelschen Gruppen zu Isomorphismen von R-Moduln werden:

25.4 Homomorphismen und Tensorprodukt
M, N und L seien R-Moduln und $\tau : M \times N \to M \otimes_R N$ die kanonische Abbildung. Dann sind folgende Abbildungen Isomorphismen von R-Moduln:

$$\begin{array}{rcl} \mathrm{Hom}_R(M \otimes_R N, L) & \xrightarrow{\psi_1} & \mathrm{Bil}_R(M \times N, L), \\ \alpha & \longmapsto & \alpha\tau. \\ \mathrm{Bil}_R(M \times N, L) & \xrightarrow{\psi_2} & \mathrm{Hom}_R(N, \mathrm{Hom}_R(M, L)), \\ \beta & \longmapsto & [n \to \beta(-, n)]. \\ \mathrm{Hom}_R(M \otimes_R N, L) & \xrightarrow{\psi_M} & \mathrm{Hom}_R(N, \mathrm{Hom}_R(M, L)), \\ \delta & \longmapsto & [n \mapsto \delta(- \otimes n)]. \end{array}$$

Speziell für Bilinearformen leiten wir daraus ab:

25.5 Bilinearformen und Tensorprodukte

(1) *Für R-Moduln M, N haben wir die R-Isomorphismen*

$$\begin{aligned}\mathrm{Bil}_R(M,N) &\simeq (M \otimes_R N)^* \\ &\simeq \mathrm{Hom}_R(M, N^*) \\ &\simeq \mathrm{Hom}_R(N, M^*).\end{aligned}$$

(2) *Ist M oder N endlich erzeugt und frei, so gilt für die Dualräume*

$$(M \otimes_R N)^* \simeq M^* \otimes_R N^*.$$

(3) *Sind M und N endlich erzeugt und frei, so ist*

$$\mathrm{End}_R(M) \otimes_R \mathrm{End}_R(N) \to \mathrm{End}_R(M \otimes_R N), \quad f \otimes g \mapsto f \underline{\otimes} g,$$

ein Isomorphismus von Ringen und R-Moduln.

Beweis: (1) Dies sind die Isomorphismen aus 25.4.

(2) Ist M endlich erzeugt und frei, so ist nach 24.9,(2)

$$M^* \otimes_R N^* \simeq \mathrm{Hom}(M, N^*).$$

Die angegebene Beziehung folgt damit aus (1).

Ist N endlich erzeugt und frei, so läßt sich analog argumentieren.

(3) Mit Hilfe von 24.9 und Eigenschaften des Tensorprodukts (aus 24.10) haben wir die Isomorphismen

$$\begin{aligned}\mathrm{End}_R(M) \otimes_R \mathrm{End}_R(N) &\simeq (M^* \otimes M) \otimes (N^* \otimes N) \\ &\simeq (M^* \otimes N^*) \otimes (M \otimes N) \\ &\simeq (M \otimes N)^* \otimes (M \otimes_R N) \\ &\simeq \mathrm{End}_R(M \otimes_R N).\end{aligned}$$

Die zugehörigen Abbildungen sind

$$\begin{aligned}f \otimes g &\mapsto \left(\sum_i x_i^* \otimes f(x_i)\right) \otimes \left(\sum_j y_j^* \otimes f(y_j)\right) \\ &\mapsto \sum_{i,j} (x_i^* \otimes y_j^*) \otimes f(x_i) \otimes g(y_j) \\ &\mapsto \sum_{i,j} (x_i \otimes y_j)^* \otimes f \underline{\otimes} g(x_i \otimes y_j) \\ &\mapsto f \underline{\otimes} g.\end{aligned}$$

□

Wie wir gesehen haben, sind mit einer Bilinearform $\beta : M \times N \to R$ folgende R-Homomorphismen festgelegt:

$$\beta_\ell : M \to N^*, \quad m \mapsto \beta(m, -),$$
$$\beta_r : N \to M^*, \quad n \mapsto \beta(-, n).$$

Eigenschaften dieser Abbildungen bestimmen auch Eigenschaften von β:

25.6 Definition
Eine Bilinearform β auf $M \times N \to R$ nennt man

links nicht-ausgeartet, wenn β_ℓ injektiv ist;

rechts nicht-ausgeartet, wenn β_r injektiv ist;

links nicht-singulär, wenn β_ℓ Isomorphismus ist;

rechts nicht-singulär, wenn β_r Isomorphismus ist.

Da Homomorphismen genau dann injektiv sind, wenn ihr Kern trivial ist, haben wir folgendes Kriterium:

25.7 Nicht-ausgeartete Bilinearformen
Eine Bilinearform β auf $M \times N$ ist genau dann links nicht-ausgeartet, wenn

$$\operatorname{Kern} \beta_\ell = \{m \in M \mid \beta(m, n) = 0 \text{ für alle } n \in N\} = 0.$$

β ist genau dann rechts nicht-ausgeartet, wenn

$$\operatorname{Kern} \beta_r = \{n \in N \mid \beta(m, n) = 0 \text{ für alle } m \in M\} = 0.$$

Wir wollen uns nun die Begriffe an einem speziellen Fall veranschaulichen.

25.8 Kanonische Bilinearform auf $M^* \times M$
Für die kanonische Bilinearform $\mu : M^* \times M \to R$ (vgl. 25.2(2)) gilt:

(1) *μ ist links nicht-singulär.*

(2) *Ist M freier Modul, so ist μ rechts nicht-ausgeartet.*

(3) *Ist M endlich erzeugt und frei, dann ist μ rechts nicht-singulär.*

Beweis: (1) Die Abbildung

$$\mu_\ell : M^* \to M^*, \quad f \mapsto [\mu(f, -) : m \mapsto f(m)],$$

ist die Identität.

(2),(3) Die Abbildung

$$\mu_r : M \to M^{**}, \quad m \mapsto [\mu(-, m) : g \mapsto g(m)],$$

ist gerade der Auswertungsmorphismus Φ_M. Wie in 23.7 gezeigt, ist Φ_M für freies M injektiv und für endlich erzeugtes, freies M sogar bijektiv. □

Betrachten wir ein $\beta \in \mathrm{Bil}_R(M,N)$, das links nicht-singulär ist, also $M \simeq N^*$. Dann ist jede Linearform auf N von der Gestalt $\beta(m,-)$ mit geeignetem $m \in M$.

Liegt nun ein weiteres $\gamma \in \mathrm{Bil}_R(M,N)$ vor, so ist für jedes $x \in M$ die Abbildung $\gamma(x,-) \in N^*$ und daher

$$\gamma(x,-) = \beta(m,-) \text{ für geeignetes } m \in M.$$

Zu jedem $x \in M$ ist dieses $m \in M$ eindeutig bestimmt, und wir haben somit eine Abbildung

$$f : M \to M, \quad x \mapsto f(x) \text{ mit } \gamma(x,-) = \beta(f(x),-).$$

Aus den Beziehungen

$$\begin{array}{rcl} \gamma(x_1+x_2,-) & = & \beta(f(x_1+x_2),-) \quad = \\ \gamma(x_1,-)+\gamma(x_2,-) & = & \beta(f(x_1),-)+\beta(f(x_2),-) \\ & = & \beta(f(x_1)+f(x_2),-) \end{array}$$

folgern wir

$$f(x_1+x_2) = f(x_1)+f(x_2)$$

und – mit einem ähnlichen Argument –

$$f(rx) = rf(x).$$

Somit ist $f : M \to M$ ein Endomorphismus. Diese Beobachtung fassen wir zusammen:

25.9 Adjungierte Abbildungen bezüglich β

M, N seien R-Moduln und $\beta \in \mathrm{Bil}_R(M,N)$ links nicht-singulär. Dann gilt:

(1) *Zu jeder Bilinearform $\gamma : M \times N \to R$ gibt es ein $f \in \mathrm{End}(M)$ mit*

$$\gamma(x,-) = \beta(f(x),-) \text{ für jedes } x \in M.$$

(2) *Zu jedem $h \in \mathrm{End}(N)$ gibt es ein $\widetilde{h} \in \mathrm{End}(M)$ mit*

$$\beta(x,h(y)) = \beta(\widetilde{h}(x),y) \text{ für alle } x \in M,\ y \in N.$$

Man nennt $\widetilde{h}$ die *Adjungierte* von h (bezüglich β).

Beweis: (1) Dies wurde in den Vorbetrachtungen gezeigt.

(2) Es ist leicht nachzuprüfen, daß

$$\beta(-,h(-)) : M \times N \to R, \quad (m,n) \mapsto \beta(m,h(n)),$$

bilinear ist. Wendet man darauf (1) an, so erhält man die Behauptung. □

Analoge Bildungen lassen sich natürlich auch für rechts nicht-singuläre Bilinearformen durchführen.

Sehen wir uns wieder einen Spezialfall an:

25.10 Adjungierte und Transponierte
Die kanonische Bilinearform $\mu : M^* \times M \to R$ ist nicht-singulär (vgl. 25.8).

Für $h \in \mathrm{End}_R(M)$ ist $\widetilde{h} : M^* \to M^*$ bestimmt durch

$$g(h(m)) = \mu(g, h(m)) = \mu(\widetilde{h}(g), m) = \widetilde{h}(g)(m)$$

für alle $g \in M^*$, $m \in M$, also $\widetilde{h}(g) = g \circ h$.

Somit ist $\widetilde{h}$ gleich der Transponierten h^t (vgl. 23.2).

Wir befassen uns weiterhin mit Elementen aus $M \times N$, auf denen eine Bilinearform verschwindet. Dazu eine weitere

25.11 Definition *(Orthogonalität)*
M und N seine R-Moduln und $\beta \in \mathrm{Bil}_R(M, N)$.

Elemente $x \in M$, $y \in N$ heißen *orthogonal* bzgl. β, wenn $\beta(x, y) = 0$.

Für Teilmengen $K \subset M$, $U \subset N$ bezeichnen wir die Untermoduln

$$\begin{aligned} K^\perp &= \{y \in N \mid \beta(k, y) = 0 \text{ für alle } k \in K\} \subset N \\ U^\perp &= \{x \in M \mid \beta(x, u) = 0 \text{ für alle } u \in U\} \subset M \end{aligned}$$

als das *orthogonale Komplement* von K bzw. U.

Damit ist β genau dann links (rechts) nicht-ausgeartet, wenn $N^\perp = 0$ (bzw. $M^\perp = 0$).

Folgende Beziehungen sind leicht zu verifizieren:

(1) $K \subset K^{\perp\perp}$, $U \subset U^{\perp\perp}$;

(2) Für $K_1 \subset K_2$ gilt $K_2^\perp \subset K_1^\perp$;

(3) $K^{\perp\perp\perp} = K^\perp$.

Bei endlich-dimensionalen Vektorräumen lassen sich für die Komplemente weitergehende Aussagen machen:

25.12 Dimensionsformel
Seien V, W endlich-dimensionale Vektorräume über dem Körper K und $\beta \in (V, W)$. Dann gilt für Unterräume $X \subset V$, $Y \subset W$:

(1) $\dim X + \dim X^\perp = \dim W + \dim(W^\perp \cap X)$,
$\dim Y + \dim Y^\perp = \dim V + \dim(V^\perp \cap Y)$.

(2) *Ist β nicht-singulär, also insbesondere* $\dim V = \dim W$*, so gilt*

$$\dim X + \dim X^\perp = \dim V = \dim Y + \dim Y^\perp \quad \textit{und}$$
$$X = X^{\perp\perp}, \quad Y = Y^{\perp\perp}.$$

Beweis: (1) Betrachte die Abbildung $V \to W^*$, $v \mapsto \beta(v, -)$. Die Einschränkung auf $X \subset V$ hat als Kern gerade $X \cap W^\perp$ und als Bild $\beta(X, -) = \{\beta(x, -) \mid x \in X\}$. Aus der Dimensionsformel für Morphismen haben wir damit

$$\dim \beta(X, -) = \dim X - \dim(X \cap W^\perp).$$

Wählen wir $x_1, \ldots, x_r \in X$ so, daß $\beta(x_1, -), \ldots, \beta(x_r, -)$ eine Basis von $\beta(X, -)$ bilden. Dann ist

$$X^\perp = \bigcap_{i=1}^{r} \operatorname{Kern} \beta(x_i, -).$$

Nach Wahl der $\beta(x_i, -)$ folgt daraus (durch Induktion)

$$\dim X^\perp = \dim W - r = \dim W - \dim \beta(X, -).$$

Zusammen mit der oberen Gleichung ergibt dies die Behauptung.

(2) Ist β nicht-singulär, so müssen V und W ($\simeq W^*$) gleiche Dimension haben, und $W^\perp = 0$, $V^\perp = 0$. Damit folgt die Dimensionsgleichung aus (1). Diese wiederum ergibt wegen $X \subset X^{\perp\perp}$ und $X^\perp \subset X^{\perp\perp\perp}$ die Gleichheit $X = X^{\perp\perp}$.

Analog sieht man $Y = Y^{\perp\perp}$. □

Für endlich-dimensionale Vektorräume ist die kanonische Bilinearform nicht-singulär. Somit haben wir als Folgerung:

25.13 Orthogonalität bezüglich $V^* \times V \to K$
Sei V ein endlich-dimensionaler Vektorraum über dem Körper K und

$$\mu : V^* \times V \to K, \quad (f, y) \mapsto f(y).$$

Dann ist für jeden Unterraum $U \subset V$

$$U^\perp = \{f \in V^* \mid f(U) = 0\},$$

und es gilt:

(1) $\dim U^\perp = \dim V - \dim U$ *und* $U = U^{\perp\perp}$.

(2) *Die Zuordnung $U \mapsto U^\perp$ ergibt eine ordnungsumkehrende Bijektion zwischen den Unterräumen von V und V^*.*

Beweis: (1) erhält man unmittelbar aus 25.12.

(2) Aus $U = U^{\perp\perp}$ sehen wir, daß die Zuordnung umkehrbar ist. □

25.14 Aufgaben

(1) Man zeige: Sind f und g Linearformen auf dem R-Modul M, so wird durch

$$\psi(m,n) := f(m) \cdot g(n) \text{ für alle } m,n \in M$$

eine Bilinearform auf M definiert.

(2) Untersuchen Sie, welche der folgenden Abbildungen $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ bilinear sind:

(i) $f_1((x_1,x_2),(y_1,y_2)) = 3x_1 + 4y_1$

(ii) $f_2((x_1,x_2),(y_1,y_2)) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$

(iii) $f_3((x_1,x_2),(y_1,y_2)) = \pi$

(iv) $f_4((x_1,x_2),(y_1,y_2)) = (x_2 - x_1)(3y_1 + y_2)$

(3) Sei V ein endlich-dimensionaler Vektorraum über einem Körper K, und $\beta : V \times V \to K$ sei eine nicht-singuläre Bilinearform. Für Untervektorräume $K, L \subset V$ zeige man:

(i) $(K + L)^\perp = K^\perp \cap L^\perp$

(ii) $(K \cap L)^\perp = K^\perp + L^\perp$.

26 Bilinearformen auf freien Moduln

Ähnlich wie die linearen Abbildungen, sind auch die bilinearen Abbildungen auf freien Moduln durch die Werte auf den Basiselementen bestimmt, die durch Matrizen dargestellt werden.

R sei wiederum ein kommutativer Ring.

26.1 Matrix einer Bilinearform

Seien M und N freie R-Moduln mit Basen $\{x_1, \ldots, x_m\} \subset M$ und $\{y_1, \ldots, y_n\} \subset N$.

(1) *Eine Bilinearform $\beta : M \times N \to R$ ist eindeutig bestimmt durch die (m,n)-Matrix*

$$B_\beta = \Big(\beta(x_i, y_j)\Big).$$

Man nennt B_β die *Matrix von β* (bezüglich der gewählten Basen).

(2) *Für eine Matrix $B \in R^{(m,n)}$, $x = \sum\limits_{i=1}^{m} r_i x_i \in M$, $y = \sum\limits_{j=1}^{n} s_j y_j \in N$, wird durch*

$$\beta_B(x,y) = \sum_{i=1}^{m}\sum_{j=1}^{n} r_i b_{ij} s_j = (r_1, \ldots, r_m) B (s_1, \ldots, s_n)^t$$

eine Bilinearform $\beta_B : M \times N \to R$ definiert.

(3) *Ist $B = B_\beta$, so gilt $\beta = \beta_B$. Für fest gewählte Basen haben wir daher eine Bijektion*

$$\mathrm{Bil}(M,N) \to R^{(m,n)}, \quad \beta \mapsto B_\beta.$$

Beweis: (1) Für beliebige Elemente $x = \sum\limits_{i=1}^{m} r_i x_i \in M$, $y = \sum\limits_{j=1}^{n} s_j y_j \in N$, $r_i, s_j \in R$, gilt

$$\beta(x,y) = \sum_{i=1}^{m}\sum_{j=1}^{n} r_i \beta(x_i, y_j) s_j.$$

Damit ist die Behauptung klar.

(2) Die Bedingungen für eine Bilinearform sind leicht nachzuprüfen.

(3) Dies folgt aus (1) und (2). □

Wie schon angemerkt, ist die Matrix einer Bilinearform von den gewählten Basen abhängig. Schauen wir, wie sie sich bei einem Basiswechsel verändert.

26.2 Matrizen bei Basiswechsel
Seien M und N freie R-Moduln mit je zwei Basen

$$X = \{x_1, \ldots, x_m\} \subset M \text{ bzw. } Y = \{y_1, \ldots, y_n\} \subset N,$$
$$X' = \{x'_1, \ldots, x'_m\} \subset M \text{ bzw. } Y' = \{y'_1, \ldots, y'_n\} \subset N$$

mit den Transformationsmatrizen

$$S = (s_{ij}) = \mathrm{Mat}_{X'X}(\mathrm{id}_M), \quad T = (t_{ij}) = \mathrm{Mat}_{Y'Y}(\mathrm{id}_N).$$

Ist $\beta : M \times N \to R$ eine Bilinearform mit den Matrizen

$$B = \Big(\beta(x_i, y_j)\Big) \text{ und } B' = \Big(\beta(x'_i, y'_j)\Big),$$

so gilt

$$B' = SBT^t.$$

Beweis: Durch Einsetzen erhalten wir für jedes Paar (i, k)

$$\begin{aligned} \beta(x'_i, y'_k) &= \beta(\sum_{j=1}^{m} s_{ij}x_j, \sum_{l=1}^{n} t_{kl}y_l) = \sum_{j=1}^{m}\sum_{l=1}^{n} s_{ij}\beta(x_j, y_k)t_{kl} \\ &= (s_{i1}, \ldots, s_{im})B(t_{k1}, \ldots, t_{kn})^t. \end{aligned}$$

Daraus folgt nun $B' = SBT^t$. □

Die Matrix einer Bilinearform beschreibt auch die durch β bestimmten Morphismen:

26.3 Matrizen von Bilinearformen und Morphismen
M und N seien R-Moduln mit Basen

$$X = \{x_1, \ldots, x_m\} \subset M \text{ und } Y = \{y_1, \ldots, y_n\} \subset N.$$

Die dazu dualen Basen bezeichnen wir mit

$$X^* = \{x_1^*, \ldots, x_m^*\} \subset M^* \text{ und } Y^* = \{y_1^*, \ldots, y_n^*\} \subset N^*.$$

Sei $\beta : M \times N \to R$ eine Bilinearform mit Matrix $B = \Big(\beta(x_i, y_j)\Big)$ und

$$\beta_\ell : M \to N^*,\ m \mapsto \beta(m, -), \quad \beta_r : N \to M^*,\ n \mapsto \beta(-, n).$$

Dann gilt

$$\mathrm{Mat}_{XY^*}(\beta_\ell) = B, \quad \mathrm{Mat}_{YX^*}(\beta_r) = B^t.$$

Beweis: Bezüglich der Basis $Y^* \subset N^*$ läßt sich das Bild des Basiselements $x_i \in M$ als

$$\beta(x_i, -) = \sum_{j=1}^{n} \beta(x_i, y_j)\, y_j^* = \sum_{j=1}^{n} b_{ij}\, y_j^*$$

schreiben. Durch Einsetzen der Basiselemente $y_1, \ldots, y_n \in N$ kann man diese Gleichung bestätigen. Entsprechend erhält man

$$\beta(-, y_j) = \sum_{i=1}^{m} \beta(x_i, y_j)\, x_i^* = \sum_{i=1}^{m} b_{ij}\, x_i^*.$$

Also haben die Matrizen B und B^t die angegebenen Eigenschaften. □

Als Folgerung daraus können wir der Matrix einer Bilinearform ansehen, ob β nicht-ausgeartet ist:

26.4 Matrix einer nicht-ausgearteten Bilinearform
Seien M und N freie R-Moduln mit Basen $\{x_1, \ldots, x_m\} \subset M$, $\{y_1, \ldots, y_m\} \subset N$. Ist $\beta : M \times N \to R$ eine Bilinearform und $B = \big(\beta(x_i, y_j)\big)$, so gilt:

(1) *Folgende Aussagen sind äquivalent:*
   (a) *β ist links nicht-ausgeartet;*
   (b) $\det B$ *ist kein Nullteiler;*
   (c) *β ist rechts nicht-ausgeartet.*

(2) *Auch folgende Aussagen sind äquivalent:*
   (a) *β ist links nicht-singulär;*
   (b) $\det B$ *ist invertierbar;*
   (c) *β ist rechts nicht-singulär.*

Man nennt $\det B$ die *Diskriminante* von β.

Beweis: (1) (a)⇔(b) $\beta_\ell : M \to N^*$ ist genau dann injektiv, wenn $\det B$ kein Nullteiler ist.

(b)⇔(c) Wegen $\det B = \det B^t$ sieht man dies analog zu (a)⇔(b).

(2) β_ℓ ist genau dann invertierbar, wenn $\det B$ invertierbar ist. □

Die obigen Ausführungen zeigen, daß man in dem betrachteten Fall nicht mehr zwischen links oder rechts nicht-ausgearteten (bzw. nicht-singulären) Bilinearformen zu unterscheiden braucht.

Es ist klar, daß über einem Körper K eine Bilinearform $\beta : V \times W \to K$ mit endlich-dimensionalen Vektorräumen $V \simeq W$ genau dann nicht-ausgeartet ist, wenn sie nicht-singulär ist oder – gleichbedeutend – die Diskriminante von Null verschieden ist.

Beispiel
Sei M ein freier R-Modul mit Basis $\{x_1, \ldots, x_m\}$. Bezüglich der Basis $\{x_1^*, \ldots, x_m^*\}$ von M^* hat die kanonische Bilinearform

$$\mu : M^* \times M \to R$$

die Matrix

$$B = \Big(\mu(x_i^*, x_j)\Big) = \Big(x_i^*(x_j)\Big) = (\delta_{ij}),$$

also die (m, m)-Einheitsmatrix. Die Determinante von B ist gleich 1, was bestätigt, daß μ nicht-singulär ist.

26.5 Aufgaben

(1) Auf $\mathbb{R}^2$ sei folgende Bilinearform f definiert:

$$f((x_1, x_2), (y_1, y_2)) := 5x_1y_1 - 3x_1y_2 - x_2y_1 + 2x_2y_2$$

(i) *Bestimmen Sie die Matrix $A \in \mathbb{R}^{(2,2)}$ von f bezüglich der Basis $U := ((1, 3), (1, 1))$.*

(ii) *Bestimmen Sie die Matrix $B \in \mathbb{R}^{(2,2)}$ von f bezüglich der Basis $V := ((2, 1), (-1, 1))$.*

(iii) *Finden Sie eine invertierbare Matrix $S \in \mathbb{R}^{(2,2)}$ mit $B = S \cdot A \cdot S^t$.*

(2) Bezüglich der Basen $X := ((1, 0, 2), (1, 3, 1), (2, 2, 0))$ von $\mathbb{R}^3$ und $Y := ((1, -1), (1, -2))$ von $\mathbb{R}^2$ sei eine Bilinearform $f : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}$ gegeben durch die Matrix

$$B_f = \begin{pmatrix} 5 & 0 \\ -2 & 2 \\ 1 & 7 \end{pmatrix}.$$

(i) *Für $a, b, c, d, e \in \mathbb{R}$ berechne man $f((a, b, c), (d, e))$.*

(ii) *Bestimmen Sie die Matrix von f bezüglich der Basen $C := ((1, 1, 0), (1, 0, 1), (0, 1, 1))$ und $D := ((1, 4), (4, 1))$.*

27 Bilinearformen auf M

Seien R ein kommutativer Ring und M ein R-Modul. Wir wollen hier Bilinearformen des Typs $M \times M \to R$ betrachten, die wir als *Bilinearformen auf M* bezeichnen.

Natürlich gilt alles für Bilinearformen $M \times N \to R$ Gezeigte auch für den speziellen Fall $M = N$. Darüber hinaus können noch zusätzliche Eigenschaften auftreten. Schauen wir zunächst, was bei Basiswechsel passiert. Aus den allgemeineren Betrachtungen in §26 ergibt sich:

27.1 Basiswechsel
Sei M ein freier R-Modul mit Basen $X = \{x_1, \ldots, x_m\}$, $X' = \{x'_1, \ldots, x'_m\}$ und $S = \mathrm{Mat}_{X',X}(\mathrm{id}_M)$.

Ist $\beta : M \times M \to R$ eine Bilinearform mit $B = \big(\beta(x_i, x_j)\big)$, $B' = \big(\beta(x'_i, x'_j)\big) \in R^{(m,m)}$, so gilt

$$B' = SBS^t.$$

27.2 Definition
Zwei Matrizen B, $B' \in R^{(m,m)}$ heißen *kongruent*, wenn es eine invertierbare Matrix $S \in R^{(m,m)}$ gibt mit

$$B' = SBS^t.$$

Dies ist eine Äquivalenzrelation auf $R^{(m,m)}$.

27.3 Definition
Eine Bilinearform $\beta : M \times M \to R$ heißt *symmetrisch*, wenn

$$\beta(m, n) = \beta(n, m) \text{ für alle } m, n \in M.$$

27.4 Die Matrix symmetrischer Bilinearformen
Sei M ein freier R-Modul mit Basis $X = \{x_1, \ldots, x_m\}$.

Eine Bilinearform $\beta : M \times M \to R$ ist genau dann symmetrisch, wenn $B = \big(\beta(x_i, x_j)\big)$ eine symmetrische Matrix ist, d.h. $B = B^t$.

Beweis: Man überlegt sich leicht, daß $\beta(x, y) = \beta(y, x)$, wenn für alle Basiselemente $\beta(x_i, x_j) = \beta(x_j, x_i)$ gilt. □

27.5 Beispiele
(1) Auf R^n definiert man für $x = (r_1, \ldots, r_n)$, $y = (s_1, \ldots, s_n) \in R^n$

$$\beta_E(x, y) = \sum_{i=1}^{n} r_i s_i.$$

Dies ergibt eine symmetrische Bilinearform, die *Standardform*. Für die Standardbasis $\{e_1, \ldots, e_n\}$ gilt $\Big(\beta(e_i, e_j)\Big) = E$.

Für jede Matrix $B = (b_{ij})$ definiert

$$\beta_B(x, y) = \sum_{i,j} r_i b_{ij} s_j = (r_1, \ldots, r_n) B (s_1, \ldots, s_n)^t$$

eine Bilinearform. Diese ist genau dann symmetrisch, wenn B eine symmetrische Matrix ist.

(2) $V = \mathbb{R}^4$, $M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$,

$\beta_M(x, y) = r_1 s_1 + r_2 s_2 + r_3 s_3 - r_4 s_4$.
Diese Bilinearform nennt man *Minkowski-Form*.

(3) $R = \mathbb{R}$, $V = \{\text{integrierbare Funktionen auf } [0, 1]\}$.
Für $\varphi, \psi \in V$ ergibt

$$\beta(\varphi, \psi) = \int_0^1 \varphi(x)\psi(x)dx$$

eine symmetrische Bilinearform auf V.

27.6 Definition
Sei $\beta : M \times M \to R$ eine Bilinearform. Ein Endomorphismus $f : M \to M$ heißt *Isometrie* (bzgl. β), wenn

$$\beta(m, n) = \beta(f(m), f(n))$$

für alle $m, n \in M$.

Bezüglich nicht-singulärer Bilinearformen lassen sich Isometrien durch Eigenschaften der adjungierten Abbildung kennzeichnen.

27.7 Isometrien und nicht-singuläre Bilinearformen
Sei M ein R-Modul und $\beta : M \times M \to R$ eine links nicht-singuläre Bilinearform. Dann sind für $f \in \mathrm{End}(M)$ folgende Aussagen äquivalent:

(a) *f ist eine Isometrie bezüglich β;*

(b) *für die Adjungierte $\widetilde{f}$ bezüglich β gilt $\widetilde{f} f = \mathrm{id}$.*

Hat M eine Basis $X = \{x_1, \ldots, x_m\}$, dann sind zudem äquivalent:

(c) *f ist invertierbar und $f^{-1} = \widetilde{f}$;*

(d) *für $B = \Big(\beta(x_i, x_j)\Big)$ gilt*

$$B = \mathrm{Mat}_X(f) B \mathrm{Mat}_X(f)^t.$$

Beweis: (a)$\Rightarrow$ (b) Für eine Isometrie $f \in \mathrm{End}(M)$ gilt nach Definition der Adjungierten $\widetilde{f}$:

$$\beta(x,y) = \beta(f(x), f(y)) = \beta(\widetilde{f}f(x), y)$$

für alle $x, y \in M$ und damit $\widetilde{f} \circ f = \mathrm{id}_M$.

(b)$\Rightarrow$ (a) Aus $\widetilde{f} \circ f = \mathrm{id}_M$ folgt aus obiger Gleichung, daß f eine Isometrie bezüglich β ist.

(c)$\Leftrightarrow$ (b) Da M endlich erzeugt und frei ist, folgt aus $\widetilde{f} \circ f = \mathrm{id}_M$ schon, daß f invertierbar ist, und $\widetilde{f} = f^{-1}$.

(c)$\Rightarrow$ (d) Für $\mathrm{Mat}_X(f) = (a_{ij}) \in R^{(m,m)}$ gilt:

$$\begin{aligned} \beta(x_i, x_k) \quad = \quad \beta(f(x_i), f(x_k)) \quad &= \quad \beta(\sum_j a_{ij}x_j, \sum_l a_{lk}x_l) \\ &= \quad \sum_j \sum_l a_{ij}\beta(x_j, x_l)a_{lk} \\ &= \quad (a_{i1}, \ldots, a_{im})B(a_{1k}, \ldots, a_{mk})^t \end{aligned}$$

und somit $B = \mathrm{Mat}(f)B\mathrm{Mat}(f)^t$. □

Zu einer Bilinearform $\beta : M \times M \to R$ ergibt die Orthogonalität eine Abbildung zwischen den Untermoduln U von M:

$$U \mapsto U^\perp = \{m \in M \mid \beta(m, U) = 0\}.$$

Man nennt ein Element $x \in M$ *isotrop*, wenn es zu sich selbst orthogonal ist, d.h. wenn $\beta(x,x) = 0$. Auch zu nicht-singulären Bilinearformen kann es isotrope Elemente geben.

Allgemeiner kann man für einen Untermodul $U \subset M$ nach den Beziehungen zu $U^\perp$ fragen. Für endlich-dimensionale Vektorräume erhalten wir dabei weitergehende Aussagen.

27.8 Orthogonale Komplemente

Seien V ein endlich-dimensionaler Vektorraum über dem Körper K und $\beta : V \times V \to K$ eine nicht-ausgeartete Bilinearform. Für einen Unterraum $U \subset V$ sind dann folgende Aussagen äquivalent:

(a) $U \cap U^\perp = 0$*;*

(b) $\beta|_U$ *ist nicht-ausgeartet;*

(c) $V = U \oplus U^\perp$.

Beweis: (a)$\Leftrightarrow$(b) Es ist klar, daß die Einschränkung von β auf U eine Bilinearform auf U ergibt.

Angenommen, für ein $x \in U$ gilt $\beta(x, y) = 0$ für alle $y \in U$, so bedeutet das $x \in U \cap U^\perp = 0$.

Also ist $\beta|_U$ genau dann nicht-ausgeartet, wenn $U \cap U^\perp = 0$.

(a)$\Leftrightarrow$(c) Nach der Dimensionsformel 25.12 gilt $\dim U + \dim U^\perp = \dim V$. Damit folgen die Behauptungen aus Dimensionsbetrachtungen. □

Als Folgerung daraus halten wir fest:

27.9 Korollar
Sei β eine nicht-ausgeartete Bilinearform auf V. Dann sind folgende Aussagen äquivalent:

(a) *Es gibt keine isotropen Elemente in V, d.h. aus $\beta(x, x) = 0$ folgt $x = 0$;*

(b) *Für jeden Unterraum $U \subset V$ gilt $U \oplus U^\perp = V$.*

Wie bei linearen Abbildungen von freien Moduln kann man auch hier fragen, wie man durch geeignete Basiswahl die Matrix einer Bilinearform β auf M besonders vorteilhaft – etwa in Diagonalform – bekommt. Dazu zeigen wir:

27.10 Die Orthogonalbasis
Sei K ein Körper mit $\operatorname{Char} K \neq 2$ und V ein endlich-dimensionaler K-Vektorraum. Ist β eine symmetrische Bilinearform, so gilt:

(1) *Es gibt eine Basis $v_1, \ldots, v_n$ von V mit*

$$\beta(v_i, v_j) = 0 \text{ für } i \neq j.$$

Eine solche Basis heißt *Orthogonalbasis* von V.

(2) *Folgende Aussagen sind äquivalent:*

(a) *β ist nicht-ausgeartet;*

(b) *für eine Orthogonalbasis $v_1, \ldots, v_n$ von V ist $\beta(v_i, v_i) \neq 0$ für $i \leq n$.*

Beweis: (1) Wir zeigen dies durch Induktion nach $n = \dim V$. Für $n = 1$ ist die Aussage klar.

Angenommen, die Behauptung ist richtig für alle Vektorräume mit Dimension $\leq n - 1$. Ist β identisch Null, so ist nichts zu zeigen.

Ist β nicht trivial, so gibt es ein $v_1 \in V$ mit $\beta(v_1, v_1) \neq 0$. Wir werden dies in 28.15 für einen etwas allgemeineren Fall zeigen. Dafür gilt $Kv_1 \cap (Kv_1)^\perp = 0$.

$(Kv_1)^\perp$ ist Lösungsmenge der linearen Gleichung

$$\beta(v_1, x) = 0.$$

Also gilt $\dim(Kv_1)^\perp = n-1$ und somit

$$V = Kv_1 \oplus (Kv_1)^\perp.$$

Nach Annahme gibt es für die Einschränkung von β auf den $(n-1)$-dimensionalen Vektorraum $(Kv_1)^\perp$ eine Orthogonalbasis. Zusammen mit v_1 haben wir dann eine Orthogonalbasis für V.

Man beachte, daß eine Orthogonalbasis isotrope Vektoren enthalten kann.

(2) Ist $v_1, \ldots, v_n$ eine Orthogonalbasis von V, so ist die zugehörige Matrix von β eine Diagonalmatrix:

$$\Big(\beta(v_i, v_j)\Big) = \begin{pmatrix} \beta_{11} & & 0 \\ & \ddots & \\ 0 & & \beta_{nn} \end{pmatrix}.$$

Nach 26.4 ist β genau dann nicht-ausgeartet, wenn

$$\det\Big(\beta(v_i, v_i)\Big) = \prod_{i=1}^{n} \beta(v_i, v_i) \neq 0.$$

Daraus folgt die Behauptung. □

Es hängt von Eigenschaften des Körpers K ab, ob man die Matrix einer Bilinearform noch weiter standardisieren kann. Die Möglichkeiten, die sich über den reellen bzw. komplexen Zahlen ergeben, werden wir im nächsten Kapitel untersuchen.

Über den komplexen Zahlen wird dabei noch die *Konjugation* zu berücksichtigen sein, also der Isomorphismus $\mathbb{C} \to \mathbb{C}$, $z \to \overline{z}$. Die dabei auftretenden Abweichungen von linearen und bilinearen Abbildungen wollen wir zuvor im nachfolgenden Abschnitt zusammenstellen.

27.11 Aufgaben

Bezüglich der kanonischen Basis des $\mathbb{Q}$-Vektorraumes $\mathbb{Q}^4$ sei eine Bilinearform $\beta : \mathbb{Q}^4 \times \mathbb{Q}^4 \to \mathbb{Q}$ durch die folgende Matrix gegeben:

$$C = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & -3 & 1 & 2 \\ -1 & 1 & 5 & 4 \\ 3 & 2 & 4 & 12 \end{pmatrix} \in \mathbb{Q}^{(4,4)}.$$

(i) *Begründen Sie, warum es eine Orthogonalbasis von $\mathbb{Q}^4$ bezüglich β gibt.*

(ii) *Berechnen Sie eine Orthogonalbasis, und geben Sie die Matrix von β bezüglich dieser Basis an.*

(iii) *Für den Unterraum $U := \mathbb{Q} \cdot (1,1,-1,-1) + \mathbb{Q} \cdot (2,1,1,-1)$ berechne man $U^\perp$ bezüglich β.*

28 Semi- und Sesquilinearformen

Neben den linearen Abbildungen spielen auch additive Abbildungen von R-Moduln eine wichtige Rolle, die einen Ringhomomorphismus $R \to R$ berücksichtigen. Auch bei bilinearen Abbildungen gibt es entsprechende Variationen, die wir uns später ansehen werden.

In diesem Abschnitt seien R ein kommutativer Ring und $\varphi : R \to R$ ein Ringhomomorphismus.

28.1 Semilineare Abbildungen
M und N seien R-Moduln. Eine $\mathbb{Z}$-lineare Abbildung $f : M \to N$ heißt *semilinear* (bezüglich φ) oder *φ-semilinear*, wenn

$$f(rm) = \varphi(r)f(m) \text{ für alle } m \in M,\ r \in R.$$

Der R-Modul ${}_RN$ kann durch

$$r \cdot n := \varphi(r)n \text{ für } r \in R,\ n \in N$$

mit einer neuen Skalarmultiplikation versehen werden. Den so definierten R-Modul bezeichnen wir mit ${}_{\varphi(R)}N$.

Damit entsprechen die φ-linearen Abbildungen ${}_RM \to {}_RN$ genau den R-Modulhomomorphismen ${}_RM \to {}_{\varphi(R)}N$.

Die Eigenschaften solcher Abbildungen hängen auch von den Eigenschaften von φ ab. Für $\mathrm{id} : R \to R$ ergeben sich daraus die linearen Abbildungen.

Beispiel
Sei M ein R-Modul. Die Linksmultiplikation

$$M \to M, \quad m \mapsto rm,$$

mit $r \in R$ ist für nicht kommutatives R kein R-Homomorphismus. Für invertierbares r ist sie jedoch eine semilineare Abbildung bezüglich des Homomorphismus

$$\varphi : R \to R, \quad s \mapsto rsr^{-1}.$$

Dies folgt aus der Gleichung $r(sm) = (rsr^{-1})rm$.

Die Abbildungen in den Grundring bekommen wieder einen eigenen Namen:

28.2 Semilinearformen
φ-semilineare Abbildungen $M \to R$ nennt man *Semilinearformen* (bzgl. φ) und bezeichnet sie mit $\mathrm{Hom}_\varphi(M, R)$.

Für jedes $f \in \mathrm{Hom}_R(M,R)$ gilt

$$\varphi \circ f(rm) = \varphi(rf(m)) = \varphi(r) \cdot \varphi \circ f(m).$$

Damit haben wir eine semilineare Abbildung

$$\mathrm{Hom}_R(M,R) \to \mathrm{Hom}_\varphi(M,R), \quad f \mapsto \varphi f.$$

Bei speziellen Homomorphismen wird diese Beziehung sogar bijektiv:

28.3 Definition
Eine Ringhomomorphismus $\varphi : R \to R$ heißt *Involution*, wenn $\varphi^2 = id_R$, wenn also

$$\varphi^2(a) = a \text{ für alle } a \in R.$$

Es ist klar, daß φ ein Isomorphismus ist mit $\varphi = \varphi^{-1}$. Für jeden Ring R ist die Identität eine Involution.

Auf den komplexen Zahlen $\mathbb{C}$ ist die Konjugation

$$^- : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \overline{z},$$

eine Involution, das wichtigste Beispiel für die späteren Anwendungen.

28.4 Hilfssatz
Ist $\varphi : R \to R$ eine Involution, dann ist

$$\mathrm{Hom}_R(M,R) \to \mathrm{Hom}_\varphi(M,R), \quad f \mapsto \varphi f,$$

eine bijektive semilineare Abbildung.

Beweis: Wegen $\varphi^2 = id_R$ ist $\mathrm{Hom}_\varphi(M,R) \to \mathrm{Hom}_R(M,R)$, $g \mapsto \varphi \circ g$, eine semilineare Abbildung, die zu der oben betrachteten Zuordnung invers ist. □

Für freie Moduln hat auch $\mathrm{Hom}_\varphi(M,R)$ besondere Eigenschaften:

28.5 φ-Dualraum von freien Moduln
Sei M ein freier R-Modul mit Basis $\{m_\lambda \mid \lambda \in \Lambda\}$.

Zu $\lambda \in \Lambda$ definieren wir Semilinearformen durch Vorgabe der Werte auf den Basiselementen

$$m_\lambda^* : M \to R, \quad m_\lambda^*(m_\mu) = \begin{cases} 0 & \text{falls } \lambda \neq \mu \\ 1 & \text{falls } \lambda = \mu, \end{cases}$$

und der Festlegung

$$m^*(rm_\lambda) = \varphi(r)m^*(m_\lambda) \text{ für } \lambda \in \Lambda,\ r \in R.$$

Diese haben die Eigenschaften:

(1) $\{m_\lambda^* \mid \lambda \in \Lambda\}$ *ist eine* R*-linear unabhängige Teilmenge von* $\mathrm{Hom}_\varphi(M, R)$.

(2) *Für* $m = \sum r_\lambda m_\lambda$ *gilt* $m_\mu^*(m) = \varphi(r_\mu)$.

(3) *Ist* Λ *endlich, dann ist* $\{m_\lambda^* \mid \lambda \in \Lambda\}$ *eine Basis von* $\mathrm{Hom}_\varphi(M, R)$.

Beweis: Die Aussagen ergeben sich aus dem Beweis von 23.3. □

Für den Rest des Abschnitts setzen wir voraus, daß $\varphi : R \to R$ eine Involution ist.

28.6 Definition
M und N seien R-Moduln. Wir nennen $\beta : M \times N \to R$ eine *Sesquilinearform* (d.h. $1\frac{1}{2}$*-Linearform*) (bezüglich φ), wenn β R-linear in der ersten und φ-semilinear in der zweiten Komponente ist.

Die Bedingung an Bilinearformen ist also ersetzt durch die Forderung

$$\beta(rm, n) = r\beta(m, n) = \beta(m, \varphi(r)n).$$

Die Menge der Sesquilinearformen bezüglich φ auf $M \times N$ bezeichnen wir mit $\mathrm{Bil}_\varphi(M, N)$.

Für $\varphi = \mathrm{id}_R$ gilt natürlich $\mathrm{Bil}_R(M, N) = \mathrm{Bil}_\varphi(M, N)$.

28.7 Beispiele (R kommutativ, $\varphi : R \to R$ Involution)

(1) Die Abbildung

$$\nu : R \times R \to R, \quad (r, s) \mapsto r\varphi(s),$$

ist eine Sesquilinearform, denn es gilt

$$\nu(r, ts) = r\varphi(ts) = \varphi(t)r\varphi(s) = \varphi(t)\nu(r, s).$$

(2) Ähnlich sieht man, daß

$$\mathrm{Hom}_\varphi(M, R) \times M \to R, \quad (g, m) \mapsto g(m),$$

eine Sesquilinearform ist.

(3) Für eine Matrix $B \in R^{(m,n)}$ ist

$$\begin{array}{rcl} R^m \times R^n & \longrightarrow & R, \\ \Big((r_1, \ldots, r_m), (s_1, \ldots, s_n)\Big) & \longmapsto & (r_1, \ldots, r_m)B(\varphi(s_1), \ldots, \varphi(s_n))^t, \end{array}$$

eine Sesquilinearform.

Ist $\beta : M \times N \to R$ eine Sesquilinearform, dann ist

$$\beta(m, -) : N \to L, \quad x \mapsto \beta(m, x),$$

semilinear, also $\beta(m, -) \in \mathrm{Hom}_\varphi(N, R)$, und

$$\beta(-, n) : M \to L, \quad y \mapsto \beta(y, n),$$

ist Linearform. Es sind somit folgende Abbildungen festgelegt:

$$\begin{aligned} \beta_\ell &: M \to \mathrm{Hom}_\varphi(N, R), \quad m \mapsto \beta(m, -), \\ \beta_r &: N \to \mathrm{Hom}_R(M, R), \quad n \mapsto \beta(-, n). \end{aligned}$$

wobei β_ℓ R-linear und β_r R-semilinear ist.

28.8 Definition
Entsprechend den Festlegungen bei Bilinearformen (vgl. 25.6), nennt man β

links (rechts) nicht-ausgeartet, wenn β_ℓ (bzw. β_r) injektiv ist, und

links (rechts) nicht-singulär, wenn β_ℓ (bzw. β_r) Isomorphismus ist.

Ähnlich wie lineare und bilineare Abbildungen sind auch semilineare und sesquilineare Abbildungen auf freien Moduln durch die Werte auf den Basiselementen bestimmt.

28.9 Matrix einer Sesquilinearform
Seien M und N freie R-Moduln mit Basen $\{x_1, \ldots, x_m\} \subset M$ und $\{y_1, \ldots, y_n\} \subset N$.

(1) *Eine Sesquilinearform $\beta : M \times N \to R$ ist eindeutig bestimmt durch die* Matrix von β *(bezüglich der gewählten Basen)*

$$B_\beta = \big(\beta(x_i, y_j)\big) \in R^{(m,n)}.$$

(2) *Für $B \in R^{(m,n)}$, $x = \sum_{i=1}^{m} r_i x_i \in M$, $y = \sum_{j=1}^{n} s_j y_j \in N$, wird durch*

$$\beta_B(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} r_i b_{ij} \varphi(s_j) = (r_1, \ldots, r_m) B (\varphi(s_1), \ldots, \varphi(s_n))^t$$

eine Sesquilinearform $\beta_B : M \times N \to R$ definiert.

Beweis: Die Aussagen aus 26.1 zu Bilinearformen übertragen sich problemlos auf Sesquilinearformen.

Die von β bestimmte Matrix B_β ergibt mit (2) wieder β. □

Auch das Verhalten bei Basiswechsel ist fast wie bei Bilinearformen:

28.10 Matrizen bei Basiswechsel
Seien M und N freie R-Moduln mit je zwei Basen

$$X = \{x_1, \ldots, x_m\} \subset M \text{ bzw. } Y = \{y_1, \ldots, y_n\} \subset N,$$
$$X' = \{x'_1, \ldots, x'_m\} \subset M \text{ bzw. } Y' = \{y'_1, \ldots, y'_n\} \subset N,$$

mit den Transformationsmatrizen

$$S = (s_{ij}) = \mathrm{Mat}_{X',X}(\mathrm{id}_M), \quad T = (t_{ij}) = \mathrm{Mat}_{Y',Y}(\mathrm{id}_N).$$

Ist $\beta : M \times N \to R$ eine Sesquilinearform mit den Matrizen

$$B = \Big(\beta(x_i, y_j)\Big) \text{ und } B' = \Big(\beta(x'_i, y'_j)\Big),$$

so gilt mit $\varphi(T) := \Big(\varphi(t_{ij})\Big)$

$$B' = SB\varphi(T)^t.$$

Beweis: Man folgt den gleichen Argumenten wie in 26.2. □

28.11 Definition
Eine Sesquilinearform $\beta : M \times M \to R$ heißt *φ-symmetrisch*, wenn

$$\beta(m, n) = \varphi(\beta(n, m)) \text{ für alle } m, n \in M.$$

Speziell für $\varphi : \mathbb{C} \to \mathbb{C}$, $z \mapsto \overline{z}$, nennt man φ-symmetrische Sesquilinearformen meist *hermitesch*. Die Bedingung dafür ist dann

$$\beta(m, n) = \overline{\beta(n, m)},$$

und für $m = n$ bedeutet das $\overline{\beta(m, m)} = \beta(m, m)$, also $\beta(m, m) \in \mathbb{R}$.

Bei freien Moduln kann man die φ-Symmetrie wieder von der Matrix ablesen. Dazu legen wir fest:

28.12 Definition
Eine Matrix $A = (a_{ij}) \in R^{(n,n)}$ heißt *φ-symmetrisch*, wenn

$$a_{ij} = \varphi(a_{ji}) \text{ für } i, j \leq n, \text{ also } A = \varphi(A)^t.$$

Eine Matrix $A = (a_{ij}) \in \mathbb{C}^{(n,n)}$ heißt *hermitesch*, wenn

$$a_{ij} = \overline{a_{ji}} \text{ für } i, j \leq n, \text{ also } A = \overline{A}^t.$$

Damit können wir nun feststellen:

28.13 Die Matrix φ-symmetrischer Bilinearformen
Sei M ein freier R-Modul mit Basis $X = \{x_1, \ldots, x_n\}$ und $\beta : M \times M \to R$ eine Sesquilinearform.

(1) *β ist genau dann φ-symmetrisch, wenn*

$$\beta(x_i, x_j) = \varphi(\beta(x_j, x_i)),$$

wenn also die Matrix $\big(\beta(x_i, x_j)\big)$ φ-symmetrisch ist.

(2) *Sei $R = \mathbb{C}$. Dann ist β genau dann hermitesch, wenn*

$$\beta(x_i, x_j) = \overline{\beta(x_j, x_i)},$$

wenn also $\big(\beta(x_i, x_j)\big)$ eine hermitesche *Matrix ist.*

Beweis: (1) Es ist klar, daß die angegebene Beziehung notwendig ist. Ist sie andererseits gegeben, so gilt für $x = \sum r_i x_i$, $y = \sum s_j x_j$,

$$\begin{aligned}
\beta(x, y) &= \beta(\sum_i r_i x_i, \sum_j s_j x_j) \\
&= \sum_i \sum_j r_i \beta(x_i, x_j) \varphi(s_j) \\
&= \sum_i \sum_j r_i \varphi(\beta(x_j, x_i)) \varphi(s_j) \\
&= \varphi\Big(\sum_i \sum_j \varphi(r_i) \beta(x_j, x_i) s_j\Big) \\
&= \varphi\Big(\beta(\sum_j s_j x_j, \sum_i r_i x_i)\Big) \quad = \quad \varphi(\beta(y, x)).
\end{aligned}$$

(2) ist ein Spezialfall von (2). □

28.14 Beispiele
Sei $V = \mathbb{C}^n$. Für $x = (r_1, \ldots, r_n)$, $y = (s_1, \ldots, s_n) \in \mathbb{C}^n$ wird durch

$$\beta_E(x, y) = \sum_{i=1}^{n} r_i \overline{s}_i$$

eine hermitesche Sesquilinearform definiert.

Für jede Matrix $B \in \mathbb{C}^{(n,n)}$ definiert

$$\beta_B(x, y) = \sum_{i,j} r_i b_{ij} \overline{s}_j = (r_1, \ldots, r_n) B (\overline{s}_1, \ldots, \overline{s}_n)^t$$

eine Sesquilinearform auf $\mathbb{C}^n$. Diese ist genau dann hermitesch, wenn dies auch für B gilt.

Wie bei linearen und bilinearen Abbildungen von freien Moduln wollen wir durch geeignete Basiswahl die Matrix einer Sesquilinearform in eine vorteilhafte Gestalt bringen. Dazu zeigen wir zunächst die Existenz nicht-isotroper Vektoren.

28.15 Die Existenz nicht-isotroper Elemente
Sei K ein Körper mit Involution $\varphi : K \to K$ und $\operatorname{Char} K \neq 2$.

Ist $\beta : V \times V \to K$ eine nicht-triviale φ-symmetrische Bilinearform auf dem K-Vektorraum V, so gibt es nicht-isotrope Vektoren in V.

Beweis: Angenommen, $\beta(x, x) = 0$ für alle $x \in V$. Dann ist für $k \in K$, $y \in V$,

$$\begin{aligned} 0 \;=\; \beta(kx + y, kx + y) &= k^2\beta(x, x) + \beta(kx, y) + \beta(y, kx) + \beta(y, y) \\ &= \beta(kx, y) + \varphi\beta(kx, y). \end{aligned}$$

Nach Voraussetzung gibt es ein Paar $x_0, y_0 \in V$ mit $\beta(x_0, y_0) \neq 0$. Setzen wir $k_0 = \beta(x_0, y_0)^{-1}$, so gilt

$$1 = \beta(k_0x_0, y_0) = \varphi\beta(k_0x_0, y_0),$$

und aus obiger Gleichung folgt $1 + 1 = 0$. Dies steht im Widerspruch zur Voraussetzung $\operatorname{Char} K \neq 2$. □

28.16 Die Orthogonalbasis
Sei K ein Körper mit Involution $\varphi : K \to K$ und $\operatorname{Char} K \neq 2$, *und sei V ein endlich-dimensionaler K-Vektorraum.*

Ist β eine φ-symmetrische Sesquilinearform auf V, so gilt:

(1) *Es gibt eine Basis $v_1, \ldots, v_n$ von V (*Orthogonalbasis*) mit*

$$\beta(v_i, v_j) = 0 \text{ für } i \neq j.$$

(2) *Folgende Aussagen sind äquivalent:*

   (a) *β ist nicht-ausgeartet;*

   (b) *für eine Orthogonalbasis $v_1, \ldots, v_n$ von V ist $\beta(v_i, v_i) \neq 0$ für alle $i \leq n$.*

Beweis: Man kann den Beweis von 27.10 übernehmen. □

28.17 Aufgaben

(1) Man zeige: Sind f eine Linearform auf V und g eine Semilinearform auf dem R-Modul M, so wird durch

$$\psi(m, n) := f(m) \cdot g(n) \text{ für alle } m, n \in M$$

eine Sesquilinearform auf M definiert.

(2) Seien R ein kommutativer Ring und $\varphi : R \to R$ eine Involution.

(i) *Zeigen Sie: Ist $A \in R^{(n,n)}$ φ-symmetrisch, so ist für jedes $S \in R^{(n,n)}$ auch $SA(\varphi(S))^t$ φ-symmetrisch.*

(ii) *Formulieren Sie entsprechende Aussagen für den Fall, daß A symmetrisch bzw. $R = \mathbb{C}$ und A hermitesch ist.*

Kapitel 8

Vektorräume mit Skalarprodukt

Gegenüber allgemeinen Körpern haben die Körper der reellen und komplexen Zahlen einige besondere Eigenschaften. Zum Beispiel gibt es auf $\mathbb{R}$ eine lineare Ordnung, jedes Quadrat ist größer Null, und die Summe von Quadraten $\neq 0$ in $\mathbb{R}$ ist nicht Null.

Über $\mathbb{C}$ zerfällt jedes nicht-konstante Polynom in Linearfaktoren.

Diese Eigenschaften ermöglichen es, Bilinearformen und Sesquilinearformen über $\mathbb{R}$ bzw. $\mathbb{C}$ noch genauer zu beschreiben als über beliebigen Körpern. Dies werden wir im nächsten Abschnitt tun.

29 Skalarprodukte

Sehen wir uns zunächst an, wie sich Bilinearformen über $\mathbb{R}$ darstellen lassen.

29.1 Bilinearformen über $\mathbb{R}$

Sei β eine symmetrische Bilinearform auf dem n-dimensionalen $\mathbb{R}$-Vektorraum V. Dann gilt:

Es gibt eine Basis $v_1, \ldots, v_n$ von V und ganze Zahlen p, q mit $p+q \leq n$, so daß für $x = \sum x_i v_i$, $y = \sum y_i v_i$ gilt

$$\beta(x,y) = x_1y_1 + x_2y_2 + \cdots + x_py_p - x_{p+1}y_{p+1} - \cdots - x_{p+q}y_{p+q}.$$

β ist genau dann nicht-ausgeartet, wenn $p+q = n$.

Bezüglich einer solchen Basis hat β die Matrix

$$\begin{pmatrix} \overbrace{\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}}^{p} & & & \\ & \overbrace{\begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix}}^{q} & & 0 \\ & & \overbrace{\begin{matrix}0 & & \\ & \ddots & \\ & & 0\end{matrix}}^{n-(p+q)} & \\ 0 & & & \end{pmatrix}$$

Beweis: Wir wählen eine Orthogonalbasis $\{v'_1, \ldots, v'_n\}$ und können o.E. annehmen

$$\beta(v'_i, v'_i) \begin{cases} > 0 & \text{für } 1 \leq i \leq p \\ < 0 & \text{für } p < i \leq p+q \\ = 0 & \text{für } p+q < i \leq n. \end{cases}$$

Setzen wir nun

$$v_i = \begin{cases} \frac{1}{\sqrt{\beta(v'_i, v'_i)}} v'_i & \text{für } 1 \leq i \leq p \\ \frac{1}{\sqrt{-\beta(v'_i, v'_i)}} v'_i & \text{für } p < i \leq p+q \\ v'_i & \text{für } p+q < i \leq n, \end{cases}$$

so hat die Basis $v_1, \ldots, v_n$ die gewünschten Eigenschaften.

Nach 28.16 ist β genau dann nicht-ausgeartet, wenn alle $\beta(v_i, v_i) \neq 0$, d.h. wenn $p + q = n$. □

Mit dem gleichen Beweis erhalten wir

29.2 Sesquilinearformen über $\mathbb{C}$

Sei β eine hermitesche Sesquilinearform auf dem n-dimensionalen $\mathbb{C}$-Vektorraum V. Dann gibt es eine Basis $v_1, \ldots, v_n$ von V und $p, q \in \mathbb{N}$, so daß gilt

$$\beta(x, y) = x_1\overline{y}_1 + x_2\overline{y}_2 + \cdots + x_p\overline{y}_p - x_{p+1}\overline{y}_{p+1} - \cdots - x_{p+q}\overline{y}_{p+q}.$$

β ist genau dann nicht-ausgeartet, wenn $p + q = n$.

Man kann natürlich verschiedene Basen finden, die den in 29.1 oder 29.2 angegebenen Bedingungen genügen. Interessant ist jedoch die Beobachtung, daß die Zahlen p und q dabei immer die gleichen sind.

29.3 Der Sylvestersche Trägheitssatz
Seien V ein n-dimensionaler Vektorraum über $\mathbb{R}$ (bzw. $\mathbb{C}$) und β eine symmetrische Bilinearform (bzw. hermitesche Sesquilinearform) auf V. Sind $v_1, \ldots, v_n$ und $u_1, \ldots, u_n$ Basen von V mit

$$\beta(v_i, v_i) \begin{cases} 1 & \text{für } i \leq p_v \\ -1 & \text{für } p_v < i \leq p_v + q_v \\ 0 & \text{für } p_v + q_v < i \leq n \end{cases}$$

und

$$\beta(u_i, u_i) \begin{cases} 1 & \text{für } i \leq p_u \\ -1 & \text{für } p_u < i \leq p_u + q_u \\ 0 & \text{für } p_u + q_u < i \leq n, \end{cases}$$

dann gilt $p_v = p_u$ und $q_v = q_u$.

Beweis: Sei $x = \sum x_i v_i = \sum y_i u_i$. Setzen wir

$$V^+ = \bigoplus_{i=1}^{p_v} Rv_i, \qquad V^- = \bigoplus_{i=p_v+1}^{p_v+q_v} Rv_i,$$
$$U^+ = \bigoplus_{i=1}^{p_u} Ru_i, \qquad U^- = \bigoplus_{i=p_u+1}^{p_u+q_u} Ru_i,$$

dann gilt

$$V = V^+ \oplus V^- \oplus V^\perp = U^+ \oplus U^- \oplus V^\perp.$$

Mit $r = \dim V^\perp$ haben wir $p_v + q_v + r = n = p_u + q_u + r$.
Nun gilt für $x \in U^+ \cap (V^- \oplus V^\perp)$

$$\beta(x, x) = \sum y_i^2 \geq 0 \text{ und } \beta(x, x) = \sum -x_i^2 \leq 0,$$

d.h. $U^+ \cap (V^- \oplus V^\perp) = 0$. Daraus folgt $p_u + (q_v + r) \leq n$ und somit $p_u \leq p_v$.
Analog erhält man $p_v \leq p_u$, also $p_v = p_u$ und $q_u = q_v$. □

Die Zahlen p und q bestimmen das Verhalten von β. Dazu die

29.4 Definition
Sei β eine hermitesche Sesquilinearform auf dem n-dimensionalen Vektorraum V über $\mathbb{C}$ (bzw. $\mathbb{R}$). Man nennt β

positiv definit, wenn $\beta(x, x) > 0$ für alle $0 \neq x \in V$;

negativ definit, wenn $\beta(x, x) < 0$ für alle $0 \neq x \in V$;

positiv semidefinit, wenn $\beta(x, x) \geq 0$ für alle $0 \neq x \in V$;

negativ semidefinit, wenn $\beta(x,x) \leq 0$ für alle $0 \neq x \in V$;

indefinit, wenn es $x, y \in V$ mit $\beta(x,x) > 0$ und $\beta(y,y) < 0$ gibt.

Die entsprechenden Eigenschaften einer Sesquilinearform lassen sich nun aus den Vorzeichen der Matrix bzgl. einer Orthogonalbasis ablesen:

29.5 Kennzeichnung von Sesquilinearformen
Sei β eine hermitesche Sesquilinearform auf dem n-dimensionalen $\mathbb{C}$-Vektorraum V, p und q die in 29.1 definierten Zahlen. Dann gilt

(1) *$p = n$ genau dann, wenn β positiv definit ist.*

(2) *$q = n$ genau dann, wenn β negativ definit ist.*

(3) *$q = 0$ genau dann, wenn β positiv semidefinit ist.*

(4) *$p = 0$ genau dann, wenn β negativ semidefinit ist.*

(5) *$p \neq 0$, $q \neq 0$ genau dann, wenn β indefinit ist.*

(6) *$p + q = n$ genau dann, wenn β nicht-ausgeartet ist.*

Beweis: Bezüglich einer geeigneten Orthogonalbasis $v_1, \ldots, v_n$ von V gilt für $x = \sum x_i v_i$

$$\beta(x,x) = \sum_{i=1}^{p} x_i \overline{x}_i - \sum_{i=p+1}^{p+q} x_i \overline{x}_i.$$

(1) Ist $p = n$, dann ist $\beta(x,x) = \sum x_i \overline{x}_i > 0$, falls $x \neq 0$.

(2) bis (5) sieht man auf ähnliche Weise.

(6) Falls $p + q = n$, so ist $\dim(V^\perp) = 0$, also $V^\perp = 0$, und somit ist β nicht-ausgeartet.

□

Wir werden uns im folgenden hauptsächlich für die positiv definiten Bi- und Sesquilinearformen interessieren.

29.6 Definition
Sei V ein Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$.

Eine positiv definite, symmetrische Bilinearform $\beta : V \times V \to \mathbb{R}$, bzw.

eine positiv definite hermitesche Form $\beta : V \times V \to \mathbb{C}$

nennt man *Skalarprodukt auf V*.

Ein reeller Vektorraum mit Skalarprodukt heißt *euklidischer Vektorraum*, ein komplexer Vektorraum mit Skalarprodukt wird als *unitärer Vektorraum* bezeichnet.

Wir geben zunächst eine Kennzeichnung von positiv definiten Formen.

29.7 Kennzeichnung von positiv definiten Formen
Sei V ein n-dimensionaler Vektorraum über $\mathbb{R}$ (bzw. $\mathbb{C}$).

Für eine symmetrische Bilinearform (bzw. hermitesche Form) β auf V sind folgende Aussagen äquivalent:

(a) *β ist positiv definit;*

(b) *es gibt eine Basis $v_1, \ldots, v_n$ von V mit*

$$\beta(v_i, v_j) = \delta_{ij} \quad \textit{(Kronecker-Symbol).}$$

Eine solche Basis nennt man *Orthonormalbasis.*

Beweis: Die Behauptung folgt sofort aus 29.5 für $p = n$. □

Wir hatten in 27.10 durch einen Induktionsbeweis gezeigt, daß es zu jeder Bilinearform eine Orthogonalbasis gibt, und daraus dann in gewissen Fällen eine Orthonormalbasis gewonnen (vgl. 29.1 ff.). Es gibt auch ein konstruktives Verfahren, um aus einer beliebigen Basis eine Orthonormalbasis zu gewinnen:

Wir schreiben $\|x\| := \sqrt{\beta(x,x)}$.

29.8 Gram-Schmidtsches Orthonormalisierungsverfahren
Sei V ein n-dimensionaler Vektorraum über $\mathbb{R}$ (bzw. $\mathbb{C}$) und β ein Skalarprodukt auf V. Ist $v_1, \ldots, v_n$ eine Basis von V, so erhalten wir eine Orthonormalbasis $u_1, \ldots, u_n$ auf folgende Weise:

Wir setzen $u_1 := \frac{1}{\|v_1\|}\, v_1$ und bilden schrittweise

$$\begin{aligned} w_2 &:= v_2 - \beta(v_2, u_1)\, u_1 && \longrightarrow \quad u_2 := \tfrac{1}{\|w_2\|}\, w_2 \\ w_{i+1} &:= v_{i+1} - \sum_{l=1}^{i} \beta(v_{i+1}, u_l)\, u_l && \longrightarrow \quad u_{i+1} := \tfrac{1}{\|w_{i+1}\|}\, w_{i+1} \end{aligned}$$

Nach Konstruktion ist w_{i+1} orthogonal zu $u_1, \ldots, u_i$, und $u_1, \ldots, u_n$ ist eine Orthonormalbasis.

Beispiel
Betrachte $\mathbb{R}^3$ mit Standard-Skalarprodukt

$$\beta\Big((r_1, r_2, r_3), (t_1, t_2, t_3)\Big) = r_1 t_1 + r_2 t_2 + r_3 t_3$$

und mit Basis $v_1 = (1,1,1)$, $v_2 = (0,1,1)$, $v_3 = (0,0,1)$.

$$u_1 \;=\; \frac{1}{\sqrt{3}}\,(1,1,1)$$

$$\begin{aligned} w_2 &= v_2 - \beta(v_2, u_1)\, u_1 \\ &= (0,1,1) - \frac{2}{\sqrt{3}\sqrt{3}}\,(1,1,1) \quad = \quad (-\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{1}{3}) \\ u_2 &= \frac{w_2}{\|w_2\|} \quad = \quad \frac{1}{\sqrt{6}}\,(-2,1,1) \\ w_3 &= (0,0,1) - \tfrac{1}{3}(1,1,1) - \tfrac{1}{6}(-2,1,1) \quad = \quad (0, -\tfrac{1}{2}, \tfrac{1}{2}) \\ u_3 &= \frac{1}{\sqrt{2}}\,(0,1,1) \end{aligned}$$

Man kann auch aus der Matrix einer Bilinearform bezüglich einer beliebigen Basis ablesen, ob sie positiv definit ist:

29.9 Satz
Seien β eine symmetrische Bilinearform auf dem n-dimensionalen $\mathbb{R}$-Vektorraum V, $v_1, \ldots, v_n$ eine Basis von V und

$$A = (a_{ij}) := \Big(\beta(v_i, v_j)\Big).$$

β ist genau dann positiv definit, wenn für alle $i = 1, \ldots, n$

$$\det \begin{pmatrix} a_{11} & \ldots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \ldots & a_{ii} \end{pmatrix} > 0$$

Diese Untermatrizen von A bezeichnet man als *Hauptminoren*.

Beweis: Sei β positiv definit. Nach 29.5 ist die Determinante einer positiv definiten Form immer positiv. Für den Unterraum $U_i := \bigoplus_{j=1}^{i} \mathbb{R}v_j$ ist $U_i \cap U_i^{\perp} = 0$, und die Einschränkung von β auf U_i ist positiv definit mit der Matrix

$$\begin{pmatrix} a_{11} & \ldots & a_{1i} \\ \vdots & & \vdots \\ a_{i1} & \ldots & a_{ii} \end{pmatrix},$$

deren Determinante also positiv sein muß.

Nehmen wir nun an, daß die Matrix A die angegebenen Eigenschaften hat. Wir zeigen durch Induktion, daß β auf U_i, für $i = 1, \ldots, n$, positiv definit ist.

Für $i = 1$ wissen wir, daß $a_{11} > 0$ und daher

$$\beta(rv_1, rv_1) = r^2\beta(v_1, v_1) = r^2 a_{11} > 0 \ \text{ für } r \neq 0.$$

Also ist β positiv definit auf U_1.

Sei die Behauptung für $k-1$ richtig, $1 < k \leq n$, d.h. die Einschränkung von β auf U_{k-1} ist positiv definit. Dann gibt es eine Orthonormalbasis in U_{k-1}. Durch Ergänzung dieser Basis zu einer Basis von U_k durch Hinzunahme eines orthogonalen Vektors und Normierung haben wir für β auf U_k eine der (k,k)-Matrizen

$$k-1 \begin{pmatrix} 1 & & 0 & \\ & \ddots & & \\ 0 & & 1 & \\ & & & -1 \end{pmatrix}, \begin{pmatrix} 1 & & 0 & \\ & \ddots & & \\ 0 & & 1 & \\ & & & 0 \end{pmatrix}, \begin{pmatrix} 1 & & 0 & \\ & \ddots & & \\ 0 & & 1 & \\ & & & 1 \end{pmatrix}.$$

Nach Voraussetzung hat die Matrix von β bezüglich der Basis $v_1, \ldots, v_k$ von U_k eine positive Determinante. Da sich bei Basiswechsel das Vorzeichen nicht ändert, kommt nur die dritte Matrix in Frage, d.h. β hat eine Orthonormalbasis in U_k, ist also positiv definit. □

Wir werden sehen, daß das Skalarprodukt eine Längenmessung in Vektorräumen ermöglicht. Wir wollen zunächst formulieren, was wir von einer solchen Länge erwarten:

29.10 Definition
Sei V ein $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum. Eine Abbildung $N : V \to \mathbb{R}$ heißt *Norm auf* V, wenn für alle $v, w \in V$, $r \in \mathbb{R}$ (bzw. $\mathbb{C}$) gilt:

(N1) $N(rv) = |r|\, N(v)$

(N2) $N(v+w) \leq N(v) + N(w)$ (Dreiecksungleichung)

(N3) $N(v) = 0 \Leftrightarrow v = 0$

Aus den Axiomen folgt übrigens, daß $N(v) \geq 0$, denn

$$0 = N(v-v) \leq N(v) + N(-v) = 2N(v).$$

In engem Zusammenhang mit der Längenmessung steht die Bestimmung des Abstandes zwischen zwei Punkten. Eine solche Abstandsfunktion sollte folgende Eigenschaften haben:

29.11 Definition
Sei X eine Menge. Unter einer *Metrik auf* X versteht man eine Abbildung $d : X \times X \to \mathbb{R}$ mit den Eigenschaften $(x, y, z \in X)$

(M1) $d(x,y) = d(y,x)$ (Symmetrie)

(M2) $d(x,z) \leq d(x,y) + d(y,z)$ (Dreiecksungleichung)

(M3) $d(x,y) = 0 \Leftrightarrow x = y$.

Auch aus diesen Axiomen folgt, daß $d(x,y) \geq 0$ für alle $x, y \in X$ gilt, denn

$$0 = d(x,x) \leq d(x,y) + d(y,x) = 2d(x,y).$$

Man beachte, daß die Metriken auf beliebigen Mengen definiert sind, wir haben daher keine Verknüpfungen zwischen den Elementen von X zu berücksichtigen.

Zwischen Norm und Metrik auf einem Vektorraum stellen wir folgenden Zusammenhang fest:

29.12 Lemma
Ist $N : V \to \mathbb{R}$ eine Norm auf dem $\mathbb{R}$- (bzw. $\mathbb{C}$-)Vektorraum V, dann wird durch

$$d(v,w) = N(v-w), \quad v, w \in V,$$

auf V eine Metrik festgelegt.

Beweis: Dies sei dem Leser als Übung belassen. □

Man beachte, daß nicht jede Metrik auf einem Vektorraum von einer Norm abgeleitet werden kann.

Wir wollen nun zeigen, daß wir mit Hilfe des Skalarproduktes auf einem Vektorraum durch

$$N(v) = \sqrt{\beta(v,v)} =: \|v\|$$

eine Norm – und damit eine Metrik – erhalten. Dazu brauchen wir die

29.13 Cauchy-Schwarz-Ungleichung
Sei V ein Vektorraum über $\mathbb{R}$ (bzw. $\mathbb{C}$) mit Skalarprodukt β. Dann gilt für alle $v, w \in V$

$$|\beta(v,w)| \leq \|v\| \cdot \|w\|.$$

Das Gleichheitszeichen gilt genau dann, wenn v und w linear abhängig sind.

Beweis: Zunächst gilt für alle $r \in \mathbb{R}$ (bzw. $\mathbb{C}$)

$$\begin{aligned} 0 &\leq \beta(v - rw, v - rw) \\ &= \beta(v,v) - r\beta(w,v) - \overline{r}\beta(v,w) + r\overline{r}\beta(w,w). \end{aligned}$$

Ist $w \neq 0$, so setzen wir $r = \beta(v,w)\beta(w,w)^{-1}$. Multiplizieren wir nun obige Gleichung mit $\beta(w,w)$:

$$\begin{aligned} 0 &\leq \beta(v,v)\beta(w,w) - \beta(v,w)\beta(w,v) - \beta(w,v)\beta(v,w) + \beta(v,w)\beta(w,v) \\ &= \beta(v,v)\beta(w,w) - |\beta(v,w)|^2 , \end{aligned}$$

und somit

$$\beta(v,w) \leq \|v\| \cdot \|w\| .$$

Gleichheit erhält man nur, falls $v = sw$ für ein $s \in \mathbb{R}$ (bzw. $\mathbb{C}$). □

29.14 Satz
Jeder Vektorraum mit Skalarprodukt wird durch

$$N(v) = \sqrt{\beta(v,v)} = \|v\| \quad \textit{für } v \in V$$

zu einem normierten Vektorraum.

Beweis: Zunächst halten wir fest: $0 \leq \beta(v,v) \in \mathbb{R}$ und $\sqrt{\beta(v,v)} \in \mathbb{R}$.

$$\begin{aligned} \text{(N1)} \quad \|rv\| &= \sqrt{\beta(rv,rv)} = \sqrt{r\overline{r}\beta(v,v)} \\ &= \sqrt{|r|^2 \beta(v,v)} = |r| \sqrt{\beta(v,v)} = |r| \cdot \|v\| . \end{aligned}$$

$$\begin{aligned} \text{(N2)} \qquad \|v+w\|^2 &= \beta(v+w, v+w) \\ &= \beta(v,v) + \beta(v,w) + \beta(w,v) + \beta(w,w) \\ &\leq \|v\|^2 + 2\,|\beta(v,w)| + \|w\|^2 \\ \text{(Cauchy-Schwarz)} \qquad &\leq \|v\|^2 + 2\,\|v\|\,\|w\| + \|w\|^2 \\ &= \big(\|v\| + \|w\|\big)^2 , \end{aligned}$$

also $\|v+w\| \leq \|v\| + \|w\|$.

(N3) β ist positiv definit.

□

Nicht jede Norm auf einem Vektorraum ist aus einem Skalarprodukt abgeleitet.

Die Cauchy-Schwarz-Ungleichung ermöglicht es, auf folgende Weise einen Winkel zwischen zwei Vektoren einzuführen:

29.15 Definition
Sei β ein Skalarprodukt auf dem Vektorraum V über $\mathbb{R}$. Zu zwei Vektoren $u, v \in V$, $u, v \neq 0$ definieren wir den *(unorientierten) Winkel* zwischen u und v durch

$$(u|v) = \sphericalangle(u,v) = \arccos \frac{\beta(u,v)}{\|u\|\,\|v\|},$$

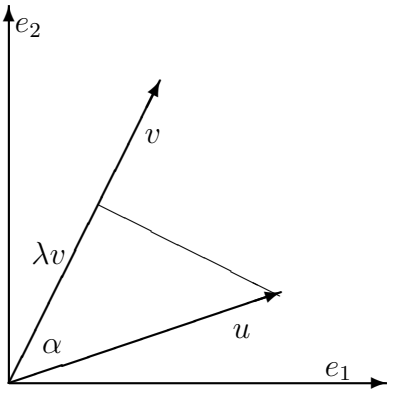
also $\cos(u|v) = \frac{\beta(u,v)}{\|u\|\|v\|}$ oder auch

$$\beta(u,v) = \|u\|\,\|v\|\cos(u|v).$$

Diese Definition ist sinnvoll, da nach Cauchy-Schwarz

$$-1 \le \frac{\beta(u,v)}{\|u\|\,\|v\|} \le 1.$$

Beispiel
Wir wollen nachprüfen, ob dies in der Ebene den gewünschten Winkelbegriff ergibt. Sei also $V = \mathbb{R}^2$ mit Standardbasis $e_1 = (1,0)$, $e_2 = (0,1)$, und die Matrix für das Skalarprodukt sei $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.



$$\cos\alpha = \frac{\|\lambda v\|}{\|u\|} = \frac{\lambda\,\|v\|^2}{\|u\|\cdot\|v\|}$$

Bedingung für λ:

$$0 = \beta(u-\lambda v, v) = \beta(u,v) - \lambda\,\|v\|^2$$

also

$$\cos\alpha = \frac{\beta(u,v)}{\|u\|\cdot\|v\|}.$$

Bemerkung: Sei β ein Skalarprodukt eines reellen Vektorraums und $v_1,\dots,v_n$ eine Orthonormalbasis. Für $u \in V$ treten die Winkel zwischen v_i und u in den Koeffizienten von u bzw. $v_1,\dots,v_n$ auf, d.h. für $u = \sum x_i v_i$

$$x_j = \beta(u,v_j) = \|u\|\,\frac{\beta(u,v_j)}{\|u\|} = \|u\|\cos(u|v_j),$$

also

$$u = \|u\| \cdot \sum_{i=1}^{n} \cos(u|v_i)\, v_i.$$

Aus den Eigenschaften des Skalarprodukts erhalten wir folgende geometrische Eigenschaften:

29.16 Satz
Sei V ein euklidischer (bzw. unitärer) Vektorraum mit Skalarprodukt β und dadurch induzierte Norm $\|\cdot\|$. Dann gilt für alle $v, w \in V$:

(1) $\|v+w\|^2 = \|v\|^2 + \|w\|^2 + \beta(v,w) + \beta(w,v)$ *(Satz von Pythagoras).*

(2) $\|v+w\|^2 + \|v-w\|^2 = 2\Big(\|v\|^2 + \|w\|^2\Big)$ *(Parallelogrammgleichung).*

Beweis: (1) $\beta(v+w, v+w) = \beta(v,v) + \beta(v,w) + \beta(w,v) + \beta(w,w)$.

(2) $\beta(v-w, v-w) = \beta(v,v) - \beta(v,w) - \beta(w,v) + \beta(w,w)$
$\Rightarrow \beta(v+w, v+w) + \beta(v-w, v-w) = 2\beta(v,v) + 2\beta(w,w)$.

Ist v senkrecht zu w, dann gilt $\|v+w\|^2 = \|v\|^2 + \|w\|^2$. □

Bemerkung: Jede Norm, die die Parallelogrammgleichung erfüllt, wird durch ein Skalarprodukt induziert.

29.17 Aufgaben

(1) Sei V ein $\mathbb{C}$-Vektorraum, und $\beta : V \times V \to \mathbb{C}$ sei eine Sesquilinearform bezüglich $\varphi : \mathbb{C} \to \mathbb{C}$ mit $\varphi(z) := \overline{z}$ für alle $z \in \mathbb{C}$. Zeigen Sie:

(i) *Für alle $v, w \in V$ gilt:*

$$\begin{aligned} 4\beta(v,w) &= \beta(v+w, v+w) - \beta(v-w, v-w) + \\ &\quad + i\beta(v+iw, v+iw) - i\beta(v-iw, v-iw). \end{aligned}$$

(ii) *β ist genau dann hermitesch, wenn $\beta(v,v) \in \mathbb{R}$ für alle $v \in V$ gilt.*

(2) Sei (e_1, e_2, e_3) die kanonische Basis des $\mathbb{C}$-Vektorraumes $\mathbb{C}^3$ und $g : \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}$ eine hermitesche Form auf $\mathbb{C}^3$ mit

$$\Big(g(e_k, e_l)\Big) = \begin{pmatrix} 1 & 2 & i \\ 2 & 3 & 2-i \\ -i & 2+i & 4 \end{pmatrix}.$$

Bestimmen Sie eine Basis $B = (b_1, b_2, b_3)$ von $\mathbb{C}^3$, die bezüglich g orthogonal ist, d.h. es gilt $g(b_k, b_l) = 0$ für $k \neq l$.

(3) Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum mit Basis $v_1, \ldots, v_n \in V$ und $x = \sum_{i=1}^{n} x_i v_i \in V$, $x_i \in \mathbb{R}$.

(i) *Zeigen Sie, daß durch*

$$N_1(x) := \sum_{i=1}^{n} |x_i|, \quad N_2(x) := \max_{1 \leq i \leq n} |x_i|, \quad N_3(x) := \sqrt{\sum_{i=1}^{n} x_i^2}$$

Normen auf V definiert sind.

(ii) *Welche dieser Normen werden von einem Skalarprodukt auf V induziert?*

(iii) *Skizzieren Sie für $V = \mathbb{R}^2$ die Teilmengen*
$E_i := \{v \in \mathbb{R}^2 \mid N_i(v) \leq 1\}$ für $i = 1, 2, 3$.

(4) In dem dreidimensionalen $\mathbb{R}$-Vektorraum V mit Basis $v_1, v_2, v_3 \in V$ sei eine Bilinearform $\langle -, - \rangle$ gegeben durch die Matrix

$$\langle v_i, v_j \rangle := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 5 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

(i) *Zeigen Sie, daß $\langle -, - \rangle$ ein Skalarprodukt auf V ist.*

(ii) *Bestimmen Sie eine Orthonormalbasis von V bezüglich $\langle -, - \rangle$.*

(iii) *Bestimmen Sie das orthogonale Komplement bezüglich $\langle -, - \rangle$ des Unterraums $U := \mathbb{R}(v_1 + v_2 + v_3)$.*

(5) Sei $V = \{f : [0,1] \to \mathbb{R} \mid f \text{ ist stetig}\}$ der $\mathbb{R}$-Vektorraum der reellwertigen stetigen Funktionen auf dem Intervall $[0,1] \subset \mathbb{R}$.

(i) *Man zeige, daß durch*

$$\begin{aligned} N_1(f) &:= \int_0^1 |f(x)|\, dx \\ N_2(f) &:= \max_{0 \le x \le 1} |f(x)| \end{aligned}$$

für $f \in V$ jeweils eine Norm auf V definiert ist.

(ii) *Wird eine der beiden Normen durch ein Skalarprodukt induziert?*

(6) Sei $V = \mathbb{R}^{(2,2)}$ der Vektorraum der $(2,2)$-Matrizen über $\mathbb{R}$.

(i) *Man zeige, daß durch $\beta(A, B) := \mathrm{Sp}(B^t \cdot A)$ für $A, B \in V$ ein Skalarprodukt auf V definiert wird.*

(ii) *Man bestimme eine Orthonormalbasis von V bezüglich β.*

(7) Sei V ein $\mathbb{R}$-Vektorraum und $N : V \to \mathbb{R}$ eine Norm auf V.
Beweisen Sie, daß durch $d(v, w) := N(v - w)$ für alle $v, w \in V$ eine Metrik auf V definiert wird.

(8) V sei ein $\mathbb{C}$-Vektorraum und $N : V \to \mathbb{R}$ eine Norm auf V, so daß die Parallelogrammgleichung gilt, d.h.

$$N(v + w)^2 + N(v - w)^2 = 2N(v)^2 + 2N(w)^2 \text{ für alle } v, w \in V.$$

Man definiert $\beta : V \times V \to \mathbb{C}$ durch

$$\beta(v, w) := \frac{1}{4} \left(N(v + w)^2 - N(v - w)^2 + iN(v + iw)^2 - iN(v - iw)^2 \right)$$

für alle $v, w \in V$ (vgl. §27, Aufgabe (2)). Zeigen Sie:

(i) $\beta(v,w) = \overline{\beta(w,v)}$ *für alle* $v, w \in V$.

(ii) β *ist positiv definit.*

(iii) $\beta(u+v,w) = \beta(u,w) + \beta(v,w)$ *für alle* $u, v, w \in V$.
Hinweis: $N(u+v+w)^2 = \frac{1}{2}N(u+(v+w))^2 + \frac{1}{2}N(v+(u+w))^2$. *Formen Sie beide Terme der rechten Seite dieser Gleichung mit Hilfe der Parallelogrammgleichung um, damit vergleichen Sie jeweils den Real- und Imaginärteil der Gleichung in (iii).*

(iv) $\beta(zv,w) = z\beta(v,w)$ *für alle* $z \in \mathbb{Z}$.

(v) $\beta(qv,w) = q\beta(v,w)$ *für alle* $q \in \mathbb{Q}$.

Mit Hilfe eines Stetigkeitsargumentes folgt nun:
$\beta(rv,w) = r\beta(v,w)$ *für alle* $r \in \mathbb{R}$ *(ohne Beweis).*

(vi) $\beta(cv,w) = c\beta(v,w)$ *für alle* $c \in \mathbb{C}$.

(vii) β *ist ein Skalarprodukt auf* V.

30 Homomorphismen und Skalarprodukte

Lineare Abbildungen zwischen Vektorräumen berücksichtigen die Vektorraumstruktur. Bei Vektorräumen mit Skalarprodukt sind diejenigen linearen Abbildungen von Interesse, die zusätzlich auf das Skalarprodukt Rücksicht nehmen.

30.1 Definition
Seien (V, β) und (W, γ) Vektorräume mit Skalarprodukten. Ein Vektorraum-Homomorphismus $f : V \to W$ heißt *Isometrie*, wenn für alle $v_1, v_2 \in V$ gilt

$$\beta(v_1, v_2) = \gamma\Big(f(v_1), f(v_2)\Big).$$

Eine Isometrie, die zugleich Isomorphismus ist, nennt man *isometrischen Isomorphismus.*

Zwei Vektorräume heißen *isometrisch*, wenn es einen isometrischen Isomorphismus zwischen ihnen gibt.

Folgende Aussagen lassen sich leicht bestätigen:

Eigenschaften
Für eine Isometrie $f : V \to W$ gilt:

(1) $\|f(v_1)\| = \|v_1\|$.

(2) $\sphericalangle(f(v_1), f(v_2)) = \sphericalangle(v_1, v_2)$ (f erhält Winkel).

(3) $\beta(v_1, v_2) = 0$ genau dann, wenn $\gamma(f(v_1), f(v_2)) = 0$
(f erhält Orthogonalität).

(4) f ist Monomorphismus.

(5) Eine Orthonormalbasis von V wird in eine Orthonormalbasis von Bild f überführt.

Aus den Eigenschaften folgt, daß für endlich-dimensionales V jede Isometrie $V \to V$ schon isometrischer Automorphismus ist.

Der nächste Satz gibt Aufschluß über die Existenz von Isometrien:

30.2 Satz
Seien (V, β) und (W, γ) Vektorräume mit Skalarprodukt, $v_1, \ldots, v_n$ eine Orthonormalbasis in V, $w_1, \ldots, w_n$ eine Orthonormalbasis in W. Dann gibt es genau eine Isometrie $f : V \to W$ mit $f(v_i) = w_i$ für $i = 1, \ldots, n$.

Beweis: Wir wissen bereits, daß es genau einen Homomorphismus f mit den gewünschten Eigenschaften gibt, und daß f ein Isomorphismus sein muß. Nach Wahl der Basen gilt

$$\beta(v_i, v_j) = \delta_{ij} = \gamma(w_i, w_j) = \gamma(f(v_i), f(v_j)).$$

Daraus folgt für beliebige $v, v' \in V$

$$\gamma(f(v), f(v')) = \beta(v, v'),$$

d.h. f ist eine Isometrie. □

Der vorausgegangene Satz besagt, daß zwei Vektorräume mit Skalarprodukt über $\mathbb{R}$ (bzw. $\mathbb{C}$) genau dann isometrisch sind, wenn sie gleiche Dimension haben.

Einem Homomorphismus $f : V \to W$ von Vektorräumen mit Basen $v_1, \ldots, v_n$ bzw. $w_1, \ldots, w_n$ haben wir eine Matrix zugeordnet durch

$$f(v_i) = \sum_{j=1}^{n} a_{ij} w_j, \quad i = 1, \ldots, n.$$

Es stellt sich die Frage, welche Eigenschaft der Matrix $A = (a_{ij})$ den Homomorphismus f zu einem isometrischen Isomorphismus macht. Da wir in endlichdimensionalen Vektorräumen mit Skalarprodukt immer Orthonormalbasen finden können, wollen wir $\mathrm{Mat}(f)$ bezüglich solcher Basen bestimmen, d.h. wir nehmen an

$$\beta(v_i, v_k) = \delta_{ik}, \quad \gamma(w_i, w_k) = \delta_{ik}.$$

Als Bedingung für Isometrie ergibt sich

$$\begin{aligned} \delta_{ik} = \beta(v_i, v_k) &= \gamma(f(v_i), f(v_k)) \\ &= \gamma\Big(\sum_j a_{ij} w_j, \sum_l a_{kl} w_l\Big) \\ &= \sum_{j,l} a_{ij} \overline{a}_{kl}\, \gamma(w_j, w_l) \\ &= \sum_{j=1}^{n} a_{ij} \overline{a}_{kj}, \end{aligned}$$

also $A\overline{A}^t = E$ (vgl. 27.7). Solchen Matrizen geben wir eigene Namen:

30.3 Definition
Eine Matrix $A \in \mathbb{R}^{(n,n)}$ heißt *orthogonal*, wenn $AA^t = E$.
Eine Matrix $A \in \mathbb{C}^{(n,n)}$ heißt *unitär*, wenn $A\overline{A}^t = E$.

Diese Bedingungen bedeuten, daß die Zeilen von A bezüglich des Standardskalarprodukts in $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) ein Orthonormalsystem bilden. Wir halten nochmals fest:

Folgerung
Eine lineare Abbildung $f : V \to W$ von euklidischen (bzw. unitären) Vektorräumen ist genau dann ein isometrischer Isomorphismus, wenn die Matrix von f bezüglich Orthonormalbasen von V und W eine orthogonale (bzw. unitäre) Matrix ist.

Eigenschaften
von orthogonalen (bzw. unitären) Matrizen. Gilt $A\overline{A}^t = E$, so folgt

$$\overline{A}^t = A^{-1}, \quad \overline{A}A^t = E, \quad \overline{A}^t A = E,$$
$$1 = \det E = \det \overline{A}^t \det A = \overline{\det A} \det A = |\det A|^2 .$$

Somit gilt für orthogonale $A \in \mathbb{R}^{(n,n)}$ mit $\det A = \pm 1$:

$$\begin{array}{ll} \det A = 1: & A \text{ ist eigentlich orthogonal} \\ \det A = -1: & A \text{ ist uneigentlich orthogonal.} \end{array}$$

Speziell für $V = W$ betrachtet man:

30.4 Definition
Sei V ein n-dimensionaler euklidischer Vektorraum. Dann heißt

$$\mathrm{Orth}(V) = \{f : V \to V \mid f \text{ ist Isometrie}\}$$

die *orthogonale Gruppe von V*. Die Isometrien $V \to V$ nennt man auch *Drehungen von V*.

Ist V ein n-dimensionaler unitärer Vektorraum, dann heißt

$$\mathrm{Un}(V) = \{f : V \to V \mid f \text{ ist Isometrie}\}$$

die *unitäre Gruppe von V* (bzw. die Gruppe der unitären Automorphismen).

Bemerkungen: (1) Die orthogonalen Matrizen entsprechen genau den Drehungen. Sie bilden eine Untergruppe der invertierbaren Matrizen in $\mathbb{R}^{(n,n)}$. Die eigentlichen orthogonalen Matrizen (mit Determinante $+1$) bilden darin eine Untergruppe. Man spricht von eigentlichen bzw. uneigentlichen Drehungen.

(2) Ein Wechsel von einer Orthonormalbasis zu einer anderen Orthonormalbasis erfolgt durch orthogonale (bzw. unitäre) Matrizen.

30.5 Drehungen im $\mathbb{R}^2$

Wir wollen feststellen, welche $(2,2)$-Matrizen über $\mathbb{R}$ eigentlich orthogonal sind. Betrachte also

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{(2,2)}$$

mit $AA^t = E$ und $\det A = 1$, das heißt $\det A = a_{11}a_{22} - a_{21}a_{12} = 1$ und

$$\begin{pmatrix} a_{11}^2 + a_{12}^2 & a_{11}a_{21} + a_{12}a_{22} \\ a_{11}a_{21} + a_{12}a_{22} & a_{21}^2 + a_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Aus den Eigenschaften der Winkelfunktionen wissen wir, daß es $\varphi, \psi \in \mathbb{R}$ gibt mit

$$\begin{aligned} a_{11} &= \cos\varphi, \quad a_{12} = \sin\varphi, \\ a_{21} &= \sin\psi, \quad a_{22} = \cos\psi, \end{aligned}$$

und dafür gilt

$$\cos(\varphi + \psi) = \cos\varphi\cos\psi - \sin\varphi\sin\psi = 1.$$

Hieraus ergibt sich $\varphi + \psi = 2k\pi$, $k \in \mathbb{Z}$, und wir erhalten

$$\begin{array}{rcccccl} a_{21} & = & \sin(2k\pi - \varphi) & = & \sin(-\varphi) & = & -\sin\varphi \\ a_{12} & = & \cos(2k\pi - \varphi) & = & \cos(-\varphi) & = & \cos\varphi\,. \end{array}$$

Die vorgegebene Matrix können wir nun so schreiben:

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Sei die Isometrie $f : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch $x = (x_1, x_2) \mapsto (x_1, x_2)A$, dann ist

$$f(x_1, x_2) = (x_1\cos\varphi - x_2\sin\varphi, x_1\sin\varphi + x_2\cos\varphi),$$

$$\begin{aligned} \cos\sphericalangle(x, f(x)) &= \frac{\beta(x, f(x))}{\|x\|\,\|f(x)\|} \\ &= \frac{x_1(x_1\cos\varphi - x_2\sin\varphi) + x_2(x_1\sin\varphi + x_2\cos\varphi)}{x_1^2 + x_2^2} \\ &= \frac{(x_1^2 + x_2^2)\cos\varphi}{x_1^2 + x_2^2} \quad = \quad \cos\varphi. \end{aligned}$$

Die Abbildung f ist also eine Drehung um den Winkel φ.

Wir können $\mathbb{R}^2$ als Gaußsche Zahlenebene $\mathbb{C}$ mit Basis 1, i auffassen.

Sei $z = x_1 + ix_2$. Eine Abbildung $f : \mathbb{C} \to \mathbb{C}$, $z \mapsto f(z)$, ist genau dann eine eigentliche Drehung, wenn gilt

$$\begin{aligned} f(z) &= (x_1\cos\varphi - x_2\sin\varphi) + (x_1\sin\varphi + x_2\cos\varphi)i \\ &= (\cos\varphi + i\sin\varphi)(x_1 + ix_2) \quad = \quad e^{i\varphi}z, \end{aligned}$$

d.h. die eigentlichen Drehungen von $\mathbb{C}$ werden genau durch die Abbildungen $z \mapsto e^{i\varphi}z$ beschrieben.

Analog zeigt man, daß die uneigentlichen Drehungen genau den Abbildungen $z \mapsto e^{i\varphi}\overline{z}$ entsprechen.

30.6 Aufgaben

(1) Sei $A = (a_{ij}) \in \mathbb{R}^{(n,n)}$ *mit Spalten* $a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$. *Zeigen Sie:*

(i) $A^{-1} = A^t \quad \Leftrightarrow \quad a_i^t \cdot a_j = \delta_{ij}$ *für alle* $i, j \leq n$.

(ii) $A^{-1} = A^t \quad \Leftrightarrow \quad \det(A) = \pm 1$.

(iii) *Ist* $A^{-1} = A^t$ *und* λ *ein Eigenwert von* A, *so ist* $\lambda = \pm 1$.

(iv) $\{A \in \mathbb{R}^{(n,n)} \mid A^{-1} = A^t\}$ *ist bezüglich der Multiplikation eine Untergruppe von* $\mathbb{R}^{(n,n)}$.

(2) Sei K *ein Körper mit* $1 + 1 \neq 0$ *in* K. *Für* $A, B \in K^{(n,n)}$ *gelte*

$$(*) \quad AB + A + B = E$$

mit der Einheitsmatrix $E \in K^{(n,n)}$. *Zeigen Sie:*

(i) $\det(A + E) \neq 0$ *und* $\det(B + E) \neq 0$.

(ii) $A^t = A^{-1} \Leftrightarrow B^t = -B$.

(iii) *Zu jeder Matrix* $A_1 \in K^{(n,n)}$ *mit* $A_1^t = A_1^{-1}$ *und* $\det(A_1 + E) \neq 0$ *gibt es eine Matrix* $B_1 \in K^{(n,n)}$ *mit* $B_1^t = -B_1$ *und* $\det(B_1 + E) \neq 0$, *so daß* A_1 *und* B_1 *die Gleichung* $(*)$ *erfüllen.*

(iv) *Ist* $A^{-1} = A^t$ *und* $\det(A + E) \neq 0$, *so ist* $\det(A) = 1$.

Hinweis *zu (ii): Schreiben Sie* $(*)$ *in der Form* $A(B + E) = E - B$ *und multiplizieren Sie mit der transponierten Gleichung.*

(3) V, W *seien euklidische Vektorräume und* $f \in \mathrm{Hom}(V, W)$. *Zeigen Sie:* f *ist genau dann Isometrie, wenn* $\|v\|_V = \|f(v)\|_W$ *für alle* $v \in V$.

(4) Man beweise, daß für die Drehmatrizen

$$D(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \in \mathbb{R}^{(2,2)}$$

mit $\varphi \in \mathbb{R}$ *die folgenden Regeln gelten:*

(i) $D(\varphi_1) \cdot D(\varphi_2) = D(\varphi_1 + \varphi_2)$ *für alle* $\varphi_1, \varphi_2 \in \mathbb{R}$.

(ii) $D(\varphi)^{-1} = D(-\varphi)$ *für alle* $\varphi \in \mathbb{R}$.

Folgern Sie daraus, daß die eigentlich-orthogonalen $(2, 2)$*-Matrizen über* $\mathbb{R}$ *mit der Multiplikation eine* kommutative *Gruppe bilden.*

31 Adjungierte Endomorphismen

Wir haben in §18 untersucht, welche Endomorphismen eines Vektorraums V diagonalisierbar bzw. trigonalisierbar sind. Dazu haben wir geeignete Basen von V gesucht. Ist auf V ein Skalarprodukt gegeben, so fragt man zudem, ob es eine *Orthonormalbasis* gibt, bezüglich welcher die Matrix von f Diagonal- oder Dreiecksgestalt hat. Folgender Satz zeigt, daß diese Forderung für trigonalisierbare Endomorphismen keine Einschränkung bedeutet:

31.1 Satz
Sei V ein Vektorraum mit Skalarprodukt. Dann gilt:
Ein Endomorphismus $f : V \to V$ ist genau dann trigonalisierbar, wenn es eine Orthonormalbasis B von V gibt, für die $\mathrm{Mat}_B(f)$ Dreiecksgestalt hat.

Beweis: Nach dem Fahnensatz 20.6 ist f genau dann trigonalisierbar, wenn es eine f-stabile Fahne $V_1 \subset V_2 \subset \ldots \subset V_n = V$ gibt. Dann hat die Matrix von f Dreiecksgestalt bezüglich der Basis $v_1, \ldots, v_n$. Durch das Gram-Schmidtsche Verfahren erhalten wir eine Orthonormalbasis $u_1, \ldots, u_n$ mit

$$V_1 = Ku_1,\ V_2 = Ku_1 + Ku_2,\ \ldots,\ V_k = \sum_{i=1}^{k} Ku_i.$$

Offensichtlich hat f auch bezüglich dieser Basis Dreiecksgestalt. □

Bemerkungen: (1) Äquivalent zur Trigonalisierbarkeit von $f : V \to V$ ist, daß $\mathrm{ch}(f)$ in Linearfaktoren zerfällt. Da dies über $\mathbb{C}$ immer gilt, wissen wir:

In einem endlich-dimensionalen unitären Vektorraum ist jeder Endomorphismus bezüglich einer Orthonormalbasis trigonalisierbar.

(2) Ist B eine Orthonormalbasis von V, $f : V \to V$ ein Endomorphismus und $A = \mathrm{Mat}_B(f)$, so ist A genau dann trigonalisierbar, wenn es eine orthogonale (bzw. unitäre) Matrix T gibt, so daß TAT^{-1} Dreiecksgestalt hat.

Für Diagonalisierbarkeit kann ein 31.1 entsprechendes Ergebnis nicht gezeigt werden. Nur spezielle Endomorphismen sind orthogonal (bzw. unitär) diagonalisierbar.

Erinnern wir uns, daß wir in 25.9 zu einem Endomorphismus $f : V \to V$ einen adjungierten Homomorphismus $f^{\mathrm{ad}} : V \to V$ bezüglich einer nichtsingulären Bilinearform $\beta : V \times V \to K$ durch die Eigenschaft

$$\beta(x, f(y)) = \beta(f^{\mathrm{ad}}(x), y) \text{ für alle } x, y \in V$$

definiert haben. Wir nennen f

selbstadjungiert, wenn $f = f^{\mathrm{ad}}$, und

normal, wenn $f \circ f^{\mathrm{ad}} = f^{\mathrm{ad}} \circ f$.

Natürlich ist jeder selbstadjungierte Endomorphimus normal. Auch jede Isometrie f ist normal, da für sie $f^{\mathrm{ad}} \circ f = \mathrm{id}$ und damit $f^{\mathrm{ad}} = f^{-1}$ gilt (falls V endlich-dimensional).

Um anzugeben, wie die Matrizen von selbstadjungierten bzw. normalen Endomorphismen aussehen, wollen wir uns zunächst die Matrix der adjungierten Abbildung ansehen:

31.2 Matrix der Adjungierten
Sei V ein endlich-dimensionaler Vektorraum mit Skalarprodukt β und Orthonormalbasis $B = (v_1, \ldots, v_n)$.

Ist $\mathrm{Mat}_B(f) = A \in \mathbb{R}^{(n,n)}$ *(bzw.* $\mathbb{C}^{(n,n)}$*), so ist*

$$\mathrm{Mat}_B(f^{\mathrm{ad}}) = \overline{A}^t.$$

Beweis: Sei $A = (a_{ij})$ und $D = (d_{ij}) = \mathrm{Mat}_B(f^{\mathrm{ad}})$. Dann gilt

$$\begin{aligned} \beta(v_i, f^{\mathrm{ad}}(v_k)) &= \beta(f(v_i), v_k), \\ \beta(v_i, \sum_{l=1}^{n} d_{kl}\, v_l) &= \beta(\sum_{j=1}^{n} a_{ij} v_j, v_k), \text{ und damit} \\ \overline{d}_{ki} = \sum_{l=1}^{n} \overline{d}_{kl} \delta_{il} &= \sum_{j=1}^{n} a_{ij} \delta_{jk} = a_{ik}, \end{aligned}$$

also $D = \overline{A}^t$. □

31.3 Korollar
Sei $f : V \to V$ ein Endomorphismus, B eine Orthonormalbasis von V und $A = \mathrm{Mat}_B(f)$.

(1) *Über $\mathbb{C}$ gilt:*

f ist genau dann selbstadjungiert, wenn $A = \overline{A}^t$ (d.h. A hermitesch).

f ist genau dann normal, wenn $A\overline{A}^t = \overline{A}^t A$ (A heißt dann *normal).*

(2) *Über $\mathbb{R}$ gilt:*

f ist genau dann selbstadjungiert, wenn $A = A^t$ (A symmetrisch).

f ist genau dann normal, wenn $AA^t = A^t A$ (A normal).

Ist f Isometrie, dann gilt $A\overline{A}^t = E = \overline{A}^t A = E$, d.h. A ist normal.

Bezeichnen wir mit $\mathrm{Eig}(f, r) = \mathrm{Kern}(f - r\,\mathrm{id})$ den Eigenraum von f zu $r \in \mathbb{C}$.

31.4 Normale Endomorphismen

Sei V ein endlich-dimensionaler unitärer Vektorraum und $f \in \mathrm{End}_{\mathbb{C}}(V)$ normal. Dann gilt:

(1) $\mathrm{Kern}\, f = \mathrm{Kern}\, f^{\mathrm{ad}}$.

(2) $\mathrm{Eig}(f, r) = \mathrm{Eig}(f^{\mathrm{ad}}, \overline{r})$ *für alle* $r \in \mathbb{C}$.

(3) $\beta(\mathrm{Eig}(f, r), \mathrm{Eig}(f, s)) = 0$ *für* $r \neq s \in \mathbb{C}$.

Beweis: (1) Für alle $v \in V$ gilt

$$\begin{aligned} \beta(f(v), f(v)) &= \beta(v, f^{\mathrm{ad}} \circ f(v)) = \beta(v, f \circ f^{\mathrm{ad}}(v)) \\ &= \beta(f^{\mathrm{ad}}(v), f^{\mathrm{ad}}(v)), \end{aligned}$$

also $f(v) = 0$ genau dann, wenn $f^{\mathrm{ad}}(v) = 0$.

(2) Der Eigenraum von f zu $r \in \mathbb{C}$ ist $\mathrm{Kern}(f - r\,\mathrm{id})$. $f - r\,\mathrm{id}$ ist wiederum normal, und $(f - r\,\mathrm{id})^{\mathrm{ad}} = f^{\mathrm{ad}} - \overline{r}\,\mathrm{id}$. Mit (1) gilt dann

$$\mathrm{Eig}(f, r) = \mathrm{Kern}(f - r\,\mathrm{id}) = \mathrm{Kern}(f^{\mathrm{ad}} - \overline{r}\,\mathrm{id}) = \mathrm{Eig}(f^{\mathrm{ad}}, \overline{r}).$$

(3) Sei $v \in E(r)$, also $f(v) = rv$, und $w \in E(s)$, also $f(w) = sw$. Nach (1) gilt $f^{\mathrm{ad}}(w) = \overline{s}w$ und damit

$$\begin{aligned} r\beta(v, w) &= \beta(rv, w) = \beta(f(v), w) = \beta(v, f^{\mathrm{ad}}(w)) = \beta(v, \overline{s}w) \\ &= s\beta(v, w), \end{aligned}$$

also $(r - s)\beta(v, w) = 0$. □

Aus (3) folgt übrigens für normale $f \in \mathrm{End}_{\mathbb{C}}(V)$ aus der Existenz einer Basis von Eigenvektoren die Existenz einer Orthonormalbasis von Eigenvektoren. Es gilt aber noch mehr:

31.5 Unitär diagonalisierbare Endomorphismen

Sei V ein endlich-dimensionaler unitärer Vektorraum. Für $f \in \mathrm{End}_{\mathbb{C}}(V)$ sind folgende Aussagen äquivalent:

(a) *f ist normal;*

(b) *es gibt eine Orthonormalbasis von V, die aus Eigenvektoren besteht;*

(c) *ist B eine Orthonormalbasis von V und $A = \mathrm{Mat}_B(f)$, dann gibt es eine unitäre Matrix $T \in \mathbb{C}^{(n,n)}$, so daß TAT^{-1} Diagonalmatrix ist (d.h. A ist unitär diagonalisierbar).*

Beweis: (a)⇒(b) Dies zeigt man durch Induktion nach der Dimension von V. Für $n = 1$ ist nichts zu zeigen.

Nehmen wir an, die Aussage ist für unitäre V mit Dimension $\leq n-1$ richtig, und sei V unitärer Vektorraum mit $\dim V = n$.

Zu einem Eigenwert r wähle einen Eigenvektor $v_1 \in V$ mit $\|v_1\| = 1$. Betrachte

$$W := v_1^{\perp} = \{w \in V \mid \beta(w, v_1) = 0\}.$$

Für $w \in W$ gilt

$$\beta(f(w), v_1) = \beta(w, f^{\mathrm{ad}}(v_1)) = \beta(w, \overline{r}v_1) = r\beta(w, v_1) = 0,$$

also $f(w) \in W$ und $f(W) \subset W$. Somit ist $V = \mathbb{C}v_1 \oplus W$ eine Zerlegung in orthogonale, f-invariante Unterräume.

$f|_W$ ist ein normaler Endomorphismus des $(n-1)$-dimensionalen Vektorraums W. Nach Induktionsannahme gibt es in W eine Basis von orthogonalen Eigenvektoren. Zusammen mit v_1 ergibt dies eine Orthonormalbasis aus Eigenvektoren von V.

(b)⇒(a) Sei $B = (v_1, \ldots, v_n)$ eine Basis von orthonormalen Eigenvektoren von f. Dann hat $\mathrm{Mat}_B(f)$ Diagonalgestalt, ist also normal. Nach 31.3 ist damit auch f normal.

(b)⇔(c) Dies ergibt sich aus dem Zusammenhang zwischen Basistransformation und Matrizendarstellung von f. □

Im Beweis von 31.5 haben wir benutzt, daß es immer nicht-triviale Eigenwerte und Eigenvektoren gibt. Dies wurde durch die Eigenschaften von $\mathbb{C}$ (algebraisch abgeschlossen) garantiert. Dieser Schluß kann nicht über $\mathbb{R}$ benutzt werden, d.h. wir bekommen nicht das gleiche Ergebnis für euklidische Vektorräume. Es gilt jedoch ein ähnliches Ergebnis, wenn man die Endomorphismen spezialisiert, d.h. für selbstadjungierte f. Dies liegt im wesentlichen an folgendem Sachverhalt:

31.6 Selbstadjungierte Endomorphismen

Sei V ein n-dimensionaler euklidischer oder unitärer Vektorraum.

Ist $f \in \mathrm{End}(V)$ selbstadjungiert, dann sind alle Eigenwerte von f reell, und $\mathrm{ch}(f)$ *zerfällt in Linearfaktoren, d.h.*

$$\mathrm{ch}(f) = \prod_{i=1}^{n}(X - r_i), \quad r_i \in \mathbb{R}.$$

Beweis: Betrachten wir zunächst unitäre V. Sei $v \in V$ Eigenvektor von f zum Eigenwert $r \in \mathbb{C}$. Dann ist

$$r\beta(v, v) = \beta(rv, v) = \beta(f(v), v) = \beta(v, f(v)) = \beta(w, rv) = \overline{r}\beta(v, v),$$

und somit $r = \overline{r}$ (da $\beta(v,v) \neq 0$).

Sei nun V ein euklidischer Vektorraum mit Orthonormalbasis B. Für $A = \mathrm{Mat}_B(f)$ gilt $A = A^t$. Faßt man A als Matrix über $\mathbb{C}$ auf, so gilt $A = \overline{A}^t$. Also definiert A einen selbstadjungierten Endomorphismus

$$\widetilde{f} : \mathbb{C}^n \to \mathbb{C}^n, \quad (z_1, \ldots, z_n) \mapsto (z_1, \ldots, z_n)A.$$

Nach dem oben Gezeigten besitzt $\widetilde{f}$ nur reelle Eigenwerte (und $\mathrm{ch}(\widetilde{f})$ zerfällt in Linearfaktoren). Somit sind auch die Eigenwerte von A reell, und $\mathrm{ch}(A)$ zerfällt in Linearfaktoren. □

Als Antwort auf unsere Frage erhalten wir damit:

31.7 Orthogonal diagonalisierbare Endomorphismen
Sei V ein n-dimensionaler euklidischer Vektorraum. Für $f \in \mathrm{End}_{\mathbb{R}}(V)$ sind folgende Aussagen äquivalent:

(a) *f ist selbstadjungiert;*

(b) *es gibt eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht;*

(c) *ist B eine Orthonormalbasis von V und $A = \mathrm{Mat}_B(f)$, dann gibt es eine orthogonale Matrix $T \in \mathbb{R}^{(n,n)}$, so daß TAT^{-1} Diagonalmatrix ist (d.h. A ist orthogonal diagonalisierbar).*

Beweis: (a)⇒(b) Unter Berücksichtigung von 31.6 folgt dies mit dem Beweis (a)⇒(b) von 31.5.

(b)⇒(c) Betrachte Basistransformation zwischen Orthonormalbasen.

(c)⇔(a) Sei T eine Matrix mit $T^t = T^{-1}$, für die $D := TAT^t$ Diagonalmatrix ist. Dann gilt

$$A = T^tDT \text{ und } A^t = T^tD^tT^{tt} = T^tDT = A.$$

Also ist A symmetrisch, und nach 31.3 ist f selbstadjungiert. □

31.8 Korollar
Jede symmetrische Matrix $A \in \mathbb{R}^{(n,n)}$ hat ein zerfallendes charakteristisches Polynom und ist orthogonal diagonalisierbar.

Bemerkung: Über die Gestalt von normalen Endomorphismen von euklidischen Räumen (über $\mathbb{R}$) haben wir keine Aussagen gemacht. Ist $A \in \mathbb{R}^{(n,n)}$ mit $A^tA = AA^t$, so kann durch eine geeignete Basiswahl folgende Gestalt erreicht werden:

$$\begin{pmatrix} r_1 & & & & & \\ & \ddots & & & & \\ & & r_m & & & \\ & & & Q_1 & & \\ & & & & \ddots & \\ & & & & & Q_s \end{pmatrix}, \quad Q_i = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad b > 0.$$

31.9 Aufgaben

(1) $A \in \mathbb{R}^{(n,n)}$ sei eine symmetrische Matrix. Zeigen Sie, daß A genau dann positiv definit ist, wenn alle Eigenwerte von A positiv sind.

(2) Sei $f : \mathbb{C}^3 \to \mathbb{C}^3$ der durch

$$\begin{aligned} f(x,y,z) \;:=\; \Big(&4ix + (2+i)z, (5-2i)x + (-1+3i)y - 5iz, \\ &2x + (1-i)y - (1-i)z\Big) \end{aligned}$$

für alle $x, y, z \in \mathbb{C}$ definierte Endomorphismus. Bestimmen Sie die zu f adjungierte Abbildung $f^{\mathrm{ad}}(x,y,z)$.

(3) U, V und W seien endlich-dimensionale unitäre Vektorräume, und es sei $f \in \mathrm{Hom}(U,V)$. Zeigen Sie, daß die zu f adjungierte Abbildung $f^{\mathrm{ad}} : V \to U$ $\mathbb{C}$-linear ist und außerdem folgende Eigenschaften hat:

(i) $(f^{\mathrm{ad}})^{\mathrm{ad}} = f$

(ii) $(f+g)^{\mathrm{ad}} = f^{\mathrm{ad}} + g^{\mathrm{ad}}$ *für alle* $g \in \mathrm{Hom}(U,V)$

(iii) $(g \circ f)^{\mathrm{ad}} = f^{\mathrm{ad}} \circ g^{\mathrm{ad}}$ *für alle* $g \in \mathrm{Hom}(V,W)$

(iv) $(zf)^{\mathrm{ad}} = \overline{z} f^{\mathrm{ad}}$ *für alle* $z \in \mathbb{C}$.

(4) Sei $h : \mathbb{C}^3 \to \mathbb{C}^3$ der Endomorphismus mit

$$\mathrm{Mat}_{E,E}(h) = \begin{pmatrix} 1 & i & -i \\ -i & 2 & 0 \\ i & 0 & 2 \end{pmatrix} \in \mathbb{C}^{(3,3)}$$

bezüglich der kanonischen Basis E von $\mathbb{C}^3$.
Geben Sie eine Orthonormalbasis B des $\mathbb{C}^3$ an, so daß $\mathrm{Mat}_{B,B}(h)$ Diagonalgestalt hat.

(5) Gegeben sei die Matrix

$$A := \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}.$$

(i) *Betrachten Sie $A \in \mathbb{C}^{(3,3)}$ und entscheiden Sie, ob A unitär diagonalisierbar ist.*

(ii) *Betrachten Sie $A \in \mathbb{R}^{(3,3)}$, und entscheiden Sie, ob A*
orthogonal diagonalisierbar,
orthogonal trigonalisierbar oder
diagonalisierbar ist.

(6) Zeigen Sie, daß die Matrix

$$A = \begin{pmatrix} 17 & -8 & 4 \\ -8 & 17 & -4 \\ 4 & -4 & 11 \end{pmatrix} \in I\!R^{(3,3)}$$

orthogonal diagonalisierbar ist, und finden Sie eine orthogonale Matrix $T \in I\!R^{(3,3)}$*, so daß* TAT^{-1} *eine Diagonalmatrix ist.*

(7) Sei

$$B = \begin{pmatrix} 2 & 10i & -2 \\ -10i & 5 & -8i \\ -2 & 8i & 11 \end{pmatrix} \in \mathbb{C}^{(3,3)}$$

gegeben. Prüfen Sie, ob B *unitär diagonalisierbar ist, und finden Sie gegebenenfalls eine unitäre Matrix* $S \in \mathbb{C}^{(3,3)}$*, so daß* SBS^{-1} *eine Diagonalmatrix ist.*

(8) Gegeben sei die folgende Matrix

$$D := \begin{pmatrix} \frac{7}{25}i & 0 & \frac{24}{25}i \\ 0 & i & 0 \\ \frac{24}{25}i & 0 & \frac{-7}{25}i \end{pmatrix} \in \mathbb{C}^{(3,3)}.$$

(i) *Zeigen Sie, daß* D *unitär diagonalisierbar ist.*

(ii) *Man finde eine unitäre Matrix* $S \in \mathbb{C}^{(3,3)}$ *(d.h.* $S^{-1} = \overline{S}^t$*), so daß* $SD\overline{S}^t$ *eine Diagonalmatrix ist.*

32 Vektorprodukt in R^3

Sei R ein kommutativer Ring. Speziell in R^3 läßt sich ein Produkt von zwei Vektoren definieren, das eine handliche Darstellung verschiedener Zusammenhänge ermöglicht. Insbesondere über den reellen Zahlen, also im Fall $R = \mathbb{R}$, erweist es sich bei der Behandlung von geometrischen und physikalischen Fragestellungen als nützlich. Ähnliche Konstruktionen in R^n werden in Aufgabe (4) skizziert.

Wir bezeichnen die kanonische Basis von R^3 mit

$$e_1 = (1,0,0), \quad e_2 = (0,1,0), \quad e_3 = (0,0,1)$$

und schreiben die Standardform von $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3) \in R^3$ (vgl. 27.5) als

$$a \cdot b = a_1b_1 + a_2b_2 + a_3b_3.$$

Für $R = \mathbb{R}$ ist dies gerade das Standard-Skalarprodukt auf $\mathbb{R}^3$.

32.1 Definiton
Als *Vektorprodukt* von $a, b \in R^3$ bezeichnen wir den Vektor

$$a \times b := (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1) \in R^3 .$$

Damit haben wir eine Abbildung

$$R^3 \times R^3 \to R^3, \quad (a,b) \mapsto a \times b,$$

mit den Eigenschaften

$$\begin{aligned} (a+c) \times b &= a \times b + c \times b \\ a \times (b+c) &= a \times b + a \times c \\ (ra) \times b &= r(a \times b) = a \times (rb) \\ a \times b &= -b \times a \\ a \times a &= 0 \end{aligned}$$

für alle $a, b, c \in R^3$ und $r \in R$. Diese Aussagen zeigen, daß die Abbildung R-bilinear und alternierend ist.

Ähnlich wie bei Bilinearformen kann man hier sehen, daß das Vektorprodukt schon durch die Produkte der Basisvektoren bestimmt ist. Diese haben die Werte

$$e_i \times e_i = 0 \ , \quad e_i \times e_j = -\, e_j \times e_i \ , \quad e_i \times e_j = e_k \ ,$$

wobei für i, j, k zyklische Vertauschungen von $1, 2, 3$ einzusetzen sind. Wir können dies in folgender Multiplikationstafel zusammenstellen:

$\times$	e_1	e_2	e_3
e_1	0	e_3	$-e_2$
e_2	$-e_3$	0	e_1
e_3	e_2	$-e_1$	0

Interessante Bildungen ergeben sich in Verbindung mit der Standard(bilinear)-form auf R^3.

32.2 Grassmann Identitäten. *Für alle $a, b, c \in R^3$ gelten*

$$a \times (b \times c) = (a \cdot c)b - (a \cdot b)c \quad \textit{und} \quad (a \times b) \times c = (a \cdot c)b - (b \cdot c)a \ .$$

Beweis: Diese Gleichungen sind in allen Komponenten linear. Daher genügt es, ihre Richtigkeit für die Basisvektoren zu zeigen. Für die erste Identität sind dies die Bedingungen

$$e_i \times (e_j \times e_k) = (e_i \cdot e_k)e_j - (e_i \cdot e_j)e_k \, ,$$

für alle $i, j, k \in \{1, 2, 3\}$. Da $e_i \cdot e_j = 0$ für $i \neq j$, bestätigt man diese anhand der Multiplikationstafel. Die zweite Identität folgt aus der ersten wegen der Antikommutativität von $\times$. □

Aus den Grassmann-Identitäten sieht man leicht, daß für $\times$ nicht das Assoziativgesetz gilt. Zusammen mit der Identität $b \times (c \times a) = (b \cdot a)c - (b \cdot c)a$ erhalten wir durch Differenz- und Summenbildung die

32.3 Jacobi Identität. *Für alle $a, b, c \in R^3$ gilt*

$$a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0 \ .$$

Betrachten wir $(R^3, +, \times)$ als algebraische Struktur, so sehen wir, daß alle Eigenschaften eines Ringes mit Ausnahme der Assoziativität der Multiplikation erfüllt sind. Man nennt einen solchen 'Ring', in dem die Jacobi-Identität gilt und das Quadrat eines jeden Elementes Null ergibt, einen *Lie-Ring* (nach dem norwegischen Mathematiker Sophus Lie). Derartige Strukturen sind in verschiedenen Zweigen der Mathematik und deren Anwendung in der Physik von großem Interesse.

32.4 Vektorprodukt und Determinanten. *Für $a, b, c, d \in R^3$ gelten*

$$a \cdot (b \times c) = \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \quad \textit{und} \ (a \times b) \cdot (c \times d) = \det \begin{pmatrix} a \cdot c & b \cdot c \\ a \cdot d & b \cdot d \end{pmatrix} .$$

Beweis: Die linke Seite folgt aus dem Laplaceschen Entwicklungssatz.

Die rechte Identität ist wiederum linear in allen Argumenten. Es genügt daher, sie durch Einsetzen der Basisvektoren e_1, e_2, e_3 zu bestätigen. □

Wie schon erwähnt, ergibt für $R = \mathbb{R}$ die Standardform ein Skalarprodukt auf $\mathbb{R}^3$, und wir erhalten die Norm eines Vektors und den Winkel $(a|b)$ zwischen zwei Vektoren $a, b \in \mathbb{R}^3$ durch

$$\|a\| = \sqrt{a \cdot a} \quad \text{und} \quad a \cdot b = \|a\| \, \|b\| \cos(a|b).$$

Aus 32.4 erhalten wir in diesem Fall eine Verschärfung der Cauchy-Schwarz-Ungleichung:

32.5 Geometrische Beschreibung. *Seien $a, b \in \mathbb{R}^3$.*

(1) $\|a \times b\|^2 = \|a\|^2 \, \|b\|^2 - (a \cdot b)^2$ und $\|a \times b\| = \|a\| \, \|b\| \, |\sin(a|b)|$.

(2) $a \times b = 0$ gilt genau dann, wenn a und b linear abhängig sind.

Beweis: (1) Die erste Gleichung ist ein Spezialfall von 32.4.

Mit der oben erwähnten Beziehung für $\cos(a|b)$ erhält man

$$\|a \times b\|^2 = \|a\|^2 \, \|b\|^2 - \|a\|^2 \, \|b\|^2 \cos^2(a|b) = \|a\|^2 \, \|b\|^2 \sin^2(a|b).$$

(2) Wir dürfen $a \neq 0 \neq b$ annehmen. Nach (1) ist $a \times b = 0$ gleichbedeutend mit

$$\|a\| \, \|b\| = |a \cdot b| = \|a\| \, \|b\| \, |\cos(a|b)|,$$

also $|\cos(a|b)| = 1$, d.h. a ist ein Vielfaches von b. □

Es ist aus der ebenen Geometrie bekannt, daß damit $\|a \times b\|$ gerade die Fläche des von a und b aufgespannten Parallelogramms ist. Daraus läßt sich folgern, daß die Gleichung

$$|a \cdot (b \times c)| = \|a\| \, \|b \times c\| \, |\cos(a|b \times c)|$$

das Volumen des von a, b und c aufgespannten Parallelotops (Spats) darstellt. Diese Bildung wird daher auch das *Spatprodukt* von a, b, c genannt.

32.6 Aufgaben.

(1) Sei R ein kommutativer Ring. Zeigen Sie für $a, b, c, d \in R^3$:

(i) *Durch formale Entwicklung nach der ersten Spalte ergibt sich*

$$a \times b = \det \begin{pmatrix} e_1 & a_1 & b_1 \\ e_2 & a_2 & b_2 \\ e_3 & a_3 & b_3 \end{pmatrix}.$$

(ii) *Ist $a \times b = 0$ für alle $b \in R^3$, dann ist $a = 0$.*

(iii) $(a \times b) \times (c \times d) = (a \cdot (b \times d))c - (a \cdot (b \times c))d$.

(iv) $a \times (a \times (a \times c)) = -(a \cdot a)\, a \times c$.

(2) Seien $a, b \in \mathbb{R}^3$. Zeigen Sie:

(i) *Die Gleichung $a \times x = b$ ist genau dann lösbar mit $x \in \mathbb{R}^3$, wenn $a \cdot b = 0$.*

(ii) *Finden Sie ein $x \in \mathbb{R}^3$ mit $x + a \times x = b$.*

(3) Für $a \in \mathbb{R}^3$ betrachte man die $\mathbb{R}$-lineare Abbildung

$$L_a : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto a \times x.$$

(i) *Bestimmen Sie die Matrix von L_a bezüglich der Basis (e_1, e_2, e_3).*

(ii) *Finden Sie das Minimalpolynom von L_a.*

(iii) *Geben Sie eine geometrische Interpretation für L_a, insbesondere falls $a \cdot a = 1$ oder $a = e_1, e_2, e_3$.*

(4) Sei R ein kommutativer Ring und $n \in \mathbb{N}$. Zu gegebenen (Spalten-) Vektoren $a_1, \ldots, a_{n-1} \in R^n$ ist die Zuordnung

$$h : R^n \to R, \quad x \mapsto \det(a_1, \ldots, a_{n-1}, x),$$

eine Linearform auf R^n. Nach 26.4 ist $R^n \to (R^n)^*, b \mapsto b \cdot (-)$, ein Isomorphismus. Daher gibt es genau einen Vektor $c \in R^n$ mit

$$h(x) = c \cdot x \ \text{ für alle } \ x \in R^n.$$

(i) *Bestimmen Sie die Koeffizienten von c (aus den Koeffizienten der a_i).* Hinweis: Laplacescher Entwicklungssatz.

(ii) *Vergleichen Sie das Ergebnis für $n = 3$ mit den Aussagen in diesem Abschnitt.*

(iii) *Welche der Aussagen aus diesem Abschnitt sind für $n > 3$ noch sinnvoll?*

Kapitel 9

Affine Geometrie

33 Affine Räume

Wir haben zwar bisher gelegentlich geometrische Anschauung zum Verständnis der Eigenschaften von Vektorräumen (insbesondere im $\mathbb{R}^2$ und $\mathbb{R}^3$) benutzt – Geometrie im eigentlichen Sinne haben wir jedoch nicht betrieben. So hatten wir es zum Beispiel nicht mit Punkten und Geraden zu tun.

Andererseits haben wir in der Linearen Algebra ein Hilfsmittel, mit dem sich Geometrie sehr vorteilhaft definieren und beschreiben läßt. Wir wählen – wie meist in den vorangegangenen Paragraphen – den axiomatischen Zugang, um uns dann mit den „natürlichen“ Situationen zu befassen. Der grundlegende Begriff ist:

33.1 Definition
Sei A eine Menge, V ein K-Vektorraum und $\alpha : A \times V \to A$ eine Abbildung. Das Tripel (A, V, α) nennt man *affinen Raum* über K, wenn gilt

(i) Für alle $P, Q \in A$ gibt es ein $v \in V$ mit $\alpha(P, v) = Q$;

(ii) für alle $P \in A$, $v \in V$ gilt $\alpha(P, v) = P \Rightarrow v = 0$;

(iii) für alle $P \in A$, $v, w \in V$ ist $\alpha(\alpha(P, v), w) = \alpha(P, v + w)$.

Die Elemente von A nennt man die *Punkte* und A selbst die *Punktmenge* des affinen Raums (A, V, α). V heißt der *zugrundeliegende* oder *zugehörige Vektorraum*.

Gilt $\alpha(P, v) = Q$, so bezeichnet man v auch als den *Verbindungsvektor* von P und Q und schreibt $v = \overrightarrow{PQ}$. Als *Dimension* von (A, V, α) bezeichnet man die Dimension des Vektorraums V.

Man sagt, V *operiert transitiv* auf A, da ausgehend von einem beliebigen Punkt $P \in A$ jedes $Q \in A$ durch ein $v \in V$ erreicht werden kann.

Mit der Schreibweise $\alpha(P, v) =: P + v$ sehen die obigen Bedingungen so aus:

(i) Für alle $P, Q \in A$ gibt es ein $v \in V$ mit $P + v = Q$;

(ii) für alle $P \in A$, $v \in V$ gilt: aus $P + v = P$ folgt $v = 0$;

(iii) für alle $P \in A$, $v, w \in V$ gilt $(P + v) + w = P + (v + w)$.

Aus diesen Vorgaben folgt, daß der Verbindungsvektor $v = \overrightarrow{PQ}$ eindeutig bestimmt ist: Angenommen $P + v = P + w$. Dann gilt

$$\begin{aligned} P \overset{\text{(ii)}}{=} P + 0 &= P + (v - v) \overset{\text{(iii)}}{=} (P + v) + (-v) \\ &= (P + w) + (-v) \overset{\text{(iii)}}{=} P + (w - v), \end{aligned}$$

woraus wir, wegen (ii), $w - v = 0$ folgern. Daraus ergibt sich

33.2 Satz
Sei (A, V, α) ein affiner Raum. Dann gibt es eine Abbildung

$$t : A \times A \to V, \quad (P, Q) \mapsto t(P, Q) = \overrightarrow{PQ}$$

mit den Eigenschaften:

(A1) Für jedes $P \in A$ und $v \in V$ gibt es genau ein $Q \in A$ mit $t(P, Q) = v$.

(A2) Für alle $P, Q, R \in A$ gilt

$$t(P, Q) + t(Q, R) = t(P, R).$$

Beweis: (A1) $Q = \alpha(P, v) = P + v$.

(A2) Es gilt $R = P + t(P, Q) + t(Q, R) = P + t(P, R) = R$.
Wegen der Eindeutigkeit des Verbindungsvektors heißt das

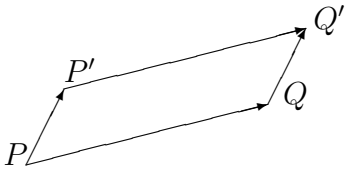
$$t(P, Q) + t(Q, R) = t(P, R).$$

□

Bemerkung: Aus 33.2 sieht man, daß die Abbildung $t : A \times A \to V$ mit den gegebenen Eigenschaften ebenfalls zur Definition eines affinen Raums geeignet ist.

Aus (A2) ergibt sich ein elementarer geometrischer Sachverhalt:

33.3 Parallelogrammregel
Für vier Punkte P, P', Q, Q' eines affinen Raums (A, V, α) gilt:
Ist $\overrightarrow{PQ} = \overrightarrow{P'Q'}$, dann auch $\overrightarrow{PP'} = \overrightarrow{QQ'}$.



Beweis: $\overrightarrow{PP'} = \overrightarrow{PQ} + \overrightarrow{QP'} = \overrightarrow{P'Q'} + \overrightarrow{QP'} = \overrightarrow{QP'} + \overrightarrow{P'Q'} = \overrightarrow{QQ'}$. □

Nehmen wir als Menge A gerade die Elemente des Vektorraums V, so haben wir als wichtiges Beispiel:

33.4 Affiner Standardraum über V
Ist V ein K-Vektorraum, so haben wir für $V = A$ die Abbildungen

$$\alpha : V \times V \to V, \quad (v, w) \mapsto v + w$$

mit den Eigenschaften (i), (ii) und (iii) aus 33.1 und

$$t : V \times V \to V, \quad (u, v) \mapsto v - u$$

mit den Eigenschaften (A1) und (A2) aus 33.2.

Dabei wird also α durch die Addition in V bestimmt. Folgende Beobachtung belegt, daß der Standardraum auch die Situation in allgemeinen affinen Räumen gut reflektiert:

33.5 Lemma
Ist (A, V, α) ein affiner Raum, so ist für jeden Punkt $O \in A$ die Abbildung

$$t_O : A \to V, \quad P \mapsto \overrightarrow{OP} = t(O, P),$$

bijektiv mit Umkehrabbildung

$$t_O^{-1} : V \to A, \quad v \mapsto O + v.$$

Beweis: Dies ergibt sich aus den bereits angesprochenen Eigenschaften affiner Räume. □

33.6 Definition
Sind (A, V, α) und (A', V', α') affine Räume über K, dann heißt (A', V', α') *(affiner) Unterraum* oder *Teilraum* von (A, V, α), wenn gilt:

(U1) $A' \subset A$, $V' \subset V$.

(U2) Für alle $P' \in A'$ und $v' \in V'$ gilt $\alpha(P', v') = \alpha'(P', v')$.

Unterräume kann man auf verschiedene Weisen gewinnen:

33.7 Charakterisierung affiner Unterräume
Sei (A, V, α) ein affiner Raum. Dann gilt:

(1) *Gilt für eine Teilmenge $A' \subset A$, daß für ein $O' \in A'$ die Menge*

$$V' := \{\overrightarrow{O'P'} \mid P' \in A'\} \subset V$$

ein Untervektorraum ist, so ist $(A', V', \alpha|_{(A'\times V')})$ ein affiner Unterraum von (A, V, α).

(2) *Ist $V' \subset V$ ein Untervektorraum, und setzt man für einen Punkt $O' \in A$*

$$A' := O' + V' = \{O' + v' \mid v' \in V'\} \subset A,$$

so ist $(A', V', \alpha|_{(A'\times V')})$ ein affiner Unterraum von (A, V, α).

(3) *Für zwei affine Unterräume (A_1, V_1, α_1), (A_2, V_2, α_2) gilt genau dann $A_1 = A_2$, wenn $A_1 \cap A_2 \neq \emptyset$ und $V_1 = V_2$.*

Beweis: (1) Wir zeigen, daß (A', V', α') ein affiner Raum ist, also $\alpha'(A', V') \subset V'$: Sei $Q' \in A'$ und $v' \in V'$. Dann gibt es ein $P' \in A'$, so daß

$$Q' + v' = O' + (\overrightarrow{O'Q'} + v') = O' + \overrightarrow{O'P'} = P' \in A'.$$

(2) Es gilt $O' = O' + 0 \in A'$ und

$$\{\overrightarrow{O'P'} \mid P' \in A'\} = \{\overrightarrow{O'(O' + v')} \mid v' \in V'\} = V' \subset V,$$

also ist $\{\overrightarrow{O'P'} \mid P' \in A'\}$ ein Vektorraum, und die Behauptung folgt aus (1).

(3) $\Rightarrow$: Mit $O' \in A_1 = A_2$ gilt

$$V_1 = \{\overrightarrow{O'P_1} \mid P_1 \in A_1\} = \{\overrightarrow{O'P_2} \mid P_2 \in A_2\} = V_2.$$

$\Leftarrow$: Mit einem $O' \in A_1 \cap A_2$ folgt

$$A_1 = O' + V_1 = O' + V_2 = A_2.$$

□

Aus dem reellen Raum bekannte Bezeichnungen übernimmt man auch für allgemeinere Situationen. So nennt man Unterräume der Dimension 1 die *Geraden* und Unterräume der Dimension 2 die *Ebenen* eines affinen Raumes (A, V, α). Ist $\dim A = n$, so bezeichnet man Unterräume A' mit $\dim A' = n-1$ als *Hyperebenen* in A.

Nach 33.7 haben wir für die erst genannten Fälle folgende

Parameterdarstellungen
Gerade: $\{O' + rv' \mid r \in K\}$ mit $O' \in A$, $v' \in V$,
Ebene : $\{O' + r_1v_1 + r_2v_2 \mid r_1, r_2 \in K\}$ mit $O' \in A$, $v_1, v_2 \in V$.

Man beachte, daß diese Darstellung eine eineindeutige Zuordnung zwischen den Punkten einer (jeden) Geraden und den Körperelementen ergibt.

Eine weitere Folge aus der Charakterisierung der affinen Unterräume ist:

33.8 Lemma
Sind $(A_i, V_i, \alpha_i)_{i\in I}$ Unterräume des affinen Raums (A, V, α), so gilt

$$\bigcap_{i\in I} A_i = \emptyset \quad \text{oder} \quad \bigcap_{i\in I} A_i \text{ ist affiner Unterraum.}$$

Beweis: Sei $O' \in \bigcap_{i\in I} A_i \neq \emptyset$, dann ist $\bigcap_{i\in I} A_i = O' + \bigcap_{i\in I} V_i$. □

Die nächste Konstruktion gibt den kleinsten affinen Unterraum an, der eine vorgegebene Menge von Punkten enthält.

33.9 Definition
Sei (A, V, α) ein affiner Raum. Für eine Teilmenge $B \subset A$ heißt

$$\langle B\rangle = \bigcap\{A' \mid B \subset A',\ A' \text{ Unterraum von } A\}$$

der *von B erzeugte affine Unterraum.*

Ist $B = \{P_1, \ldots, P_k\}$ eine endliche Teilmenge, so schreibt man auch

$$\langle B\rangle = \langle P_1, \ldots, P_k\rangle .$$

Der von zwei verschiedenen Punkten erzeugte Unterraum ist eine Gerade, drei Punkte erzeugen eine Gerade oder eine Ebene. Dafür haben wir schon Parameterdarstellungen betrachtet. Der allgemeine Fall läßt sich ähnlich darstellen:

33.10 Parameterdarstellung affiner Unterräume
Sei (A, V, α) ein affiner Raum, $\emptyset \neq B \subset A$.

(1) *Für den zugrundeliegenden Vektorraum von $\langle B\rangle$ gilt für jeden (festen) Punkt $P_0 \in B$*

$$V_B = \langle \overrightarrow{P_0P} \mid P \in B\rangle \subset V.$$

Damit ist

$$\langle B\rangle = \{Q \in A \mid \overrightarrow{P_0Q} \in V_B\} = P_0 + V_B.$$

(2) *Ist $B = \{P_0, \ldots, P_k\}$ endlich, so gilt*

$$\langle P_0, \ldots, P_k\rangle = \{Q \in A \mid Q = P_0 + \sum_{i=1}^{k} a_i\overrightarrow{P_0P_i},\ a_i \in K\}.$$

(3) *oder für (festes)* $O \in A$

$$\langle P_0, \ldots, P_k \rangle = \{Q \in A \mid \overrightarrow{OQ} = \sum_{i=0}^{k} a_i \overrightarrow{OP_i} \text{ mit } \sum_{i=0}^{k} a_i = 1\}.$$

Beweis: (1) Zunächst ist V_B in jedem Vektorraum eines affinen Unterraums $A' \subset A$ mit $B \subset A$ enthalten.

Andererseits ist $\{Q \in A \mid \overrightarrow{P_0Q} \in V_B\}$ ein solcher Unterraum.

(2) folgt aus (1).

(3) $Q = P_0 + \sum\limits_{i=1}^{k} a_i \overrightarrow{P_0P_i}$, daraus folgt:

$$\begin{aligned}
\overrightarrow{OQ} &= \overrightarrow{OP_0} + \sum_{i=1}^{k} a_i \overrightarrow{P_0P_i} = \overrightarrow{OP_0} + \sum_{i=1}^{k} a_i \left(\overrightarrow{OP_i} - \overrightarrow{OP_0}\right) \\
&= \left(1 - \sum_{i=1}^{k} a_i\right) \overrightarrow{OP_0} + \sum_{i=1}^{k} a_i \overrightarrow{OP_i} \\
&= \sum_{i=0}^{k} a_i \overrightarrow{OP_i} \quad \text{mit } a_0 = 1 - \sum_{i=1}^{k} a_i \,.
\end{aligned}$$

□

Bezeichnungen

Die in (2) gewählte Schreibweise nennt man *inhomogene Parameterdarstellung*, die in (3) heißt *homogene Parameterdarstellung* von $\langle B \rangle$.

Ist $Q \in A$ und

$$\overrightarrow{OQ} = \sum_{i=0}^{k} a_i \overrightarrow{OP_i} \text{ mit } \sum_{i=0}^{k} a_i = 1,$$

dann heißen die $(a_0, \ldots, a_k) \in K^{n+1}$ *Schwerpunktkoordinaten* oder *konzentrische Koordinaten*.

Diese sind nicht eindeutig, falls es linear abhängige $\overrightarrow{OP_i}$ gibt. Sie sind aber in jedem Fall unabhängig von der Wahl des Punkts $O \in A$, denn für $O' \in A$ gilt

$$\overrightarrow{OQ} = \overrightarrow{OO'} + \overrightarrow{O'Q} = \sum_{i=0}^{k} a_i \left(\overrightarrow{OO'} + \overrightarrow{O'P_i}\right) = \overrightarrow{OO'} + \sum_{i=0}^{k} a_i \overrightarrow{O'P_i},$$

und damit $\overrightarrow{O'Q} = \sum_{i=0}^{k} a_i \overrightarrow{O'P_i}$.

Der Durchschnitt von zwei Unterräumen A_1, A_2 war der größte Unterraum, der in A_1 und A_2 enthalten ist. Wir betrachten nun den kleinsten Unterraum, der A_1 und A_2 enthält:

33.11 Definition
Sind (A_1, V_1, α_1), (A_2, V_2, α_2) affine Unterräume von (A, V, α), dann nennt man

$$A_1 + A_2 = \langle A_1 \cup A_2 \rangle$$

die *Summe* oder den *Verbindungsraum* von A_1 und A_2.

Der zu $A_1 + A_2$ gehörige Vektorraum ist

$$V_1 + V_2 + \langle \overrightarrow{P_1 P_2} \rangle$$

für (irgend) zwei Punkte $P_1 \in A_1$, $P_2 \in A_2$. Daraus ergibt sich die Dimension des Verbindungsraumes:

33.12 Dimensionsformel
Sind (A_1, V_1, α_1) und (A_2, V_2, α_2) affine Unterräume von (A, V, α), so gilt:

$$\dim(A_1 + A_2) = \begin{cases} \dim A_1 + \dim A_2 - \dim A_1 \cap A_2 & \text{falls } A_1 \cap A_2 \neq \emptyset \\ \dim A_1 + \dim A_2 - \dim A_1 \cap A_2 + 1 & \text{falls } A_1 \cap A_2 = \emptyset \end{cases}.$$

Beweis: Falls $A_1 \cap A_2 \neq \emptyset$, so wähle $P_1 = P_2 \in A_1 \cap A_2$. □

Analog zu Begriffen in Vektorräumen nennt man Punkte $P_0, \ldots, P_k \in A$

linear unabhängig, wenn $\{\overrightarrow{P_0 P_i}\}_{i \in \{1,\ldots,k\}}$ linear unabhängig ist,

eine *Basis* oder ein *Koordinatensystem*, wenn $\{\overrightarrow{P_0 P_i}\}_{i \in \{1,\ldots,k\}}$ eine Basis von V ist.

Ohne Mühe erhält man nachstehende

33.13 Kennzeichnung von Basen
Sei (A, V, α) ein affiner Raum. Für eine Teilmenge $\{P_0, \ldots, P_n\} \subset A$ sind folgende Aussagen äquivalent:

(a) *$\{P_0, \ldots, P_n\}$ ist minimales Erzeugendensystem von A;*

(b) *$\{\overrightarrow{P_0 P_i}\}_{i=1,\ldots,n}$ ist eine Basis von V;*

(c) *Für jedes P_k, $k \in \{0, \ldots, n\}$ ist $\{\overrightarrow{P_k P_i}\}_{i \neq k}$ eine Basis von V.*

In der Geometrie interessiert man sich für das Verhalten verschiedener Unterräume zueinander. Einige Fälle haben eine eigene Bezeichnung:

33.14 Definition
Sind (A_1, V_1, α_1) und (A_2, V_2, α_2) affine Unterräume von (A, V, α), dann sagt man A_1 und A_2 *sind zueinander*

parallel, wenn $V_1 \subset V_2$ oder $V_2 \subset V_1$,

teilweise parallel, wenn $V_1 \cap V_2 \neq \{0\} \subset V$,

punktfremd, wenn $A_1 \cap A_2 = \emptyset$;

windschief, wenn $A_1 \cap A_2 = \emptyset$ und $V_1 \cap V_2 = 0$.

Für Geraden und Ebenen im $\mathbb{R}^3$ sind diese Begriffe wohlvertraut.

33.15 Aufgaben

(1) Sei A eine Menge, V ein Vektorraum über einem Körper K und $t : A \times A \to V$ eine Abbildung mit:

(A1) Für jedes $P \in A$ und jedes $v \in V$ gibt es genau ein $Q \in A$ mit $t(P, Q) = v$.

(A2) Für alle $P, Q, R \in A$ gilt $t(P, Q) + t(Q, R) = t(P, R)$.

Zeigen Sie, daß damit ein affiner Raum (A, V, α) definiert werden kann.

(2) Sei V ein K-Vektorraum mit Untervektorräumen U_1 und U_2. Für $a_1, a_2 \in V$ werden durch $A_1 = a_1 + U_1$ und $A_2 = a_2 + U_2$ affine Unterräume des affinen Standardraumes (V, V, α) bestimmt. Beweisen Sie:

(i) *$A_1 \cap A_2 \neq 0$ genau dann, wenn $a_1 - a_2 \in U_1 + U_2$.*

(ii) *Ist $V = U_1 \oplus U_2$, so haben A_1 und A_2 genau einen Punkt gemeinsam.*

(3) Im affinen Standardraum $V = \mathbb{R}^5$ seien die folgenden Punkte gegeben:

$$P_0 = (1, 2, 3, 4, 5), \quad P_1 = (0, 2, 5, 9, 2), \quad P_2 = (1, 3, 5, 7, 9)$$
$$P_3 = (3, -1, 3, 5, 6), \quad P_4 = (1, 0, 9, 18, 4)$$

(i) *Berechnen Sie die Dimension des von $P_0, \ldots, P_4$ aufgespannten affinen Unterraumes U.*

(ii) *Bestimmen Sie den Durchschnitt von U mit der durch die Gleichung $2x_1 + x_2 - x_3 + 2x_4 - 2x_5 + 3 = 0$ gegebenen Hyperebene.*

34 Teilverhältnis und Schließungssätze

Wir wollen in diesem Abschnitt ausführen, welche elementargeometrischen Sätze in unserem Rahmen schon gezeigt werden können. Wichtiges Hilfsmittel dabei ist die folgende

34.1 Definition
Seien (A, V, α) ein affiner Raum und $P, Q, R \in A$ drei Punkte, die auf einer Geraden liegen, wobei $P \neq Q$. Das eindeutig bestimmte $\lambda \in K$ mit

$$\overrightarrow{PR} = \lambda \overrightarrow{PQ}$$

heißt das *Teilverhältnis* von P, Q und R, und man schreibt $\lambda = \mathrm{TV}(P, Q, R)$.

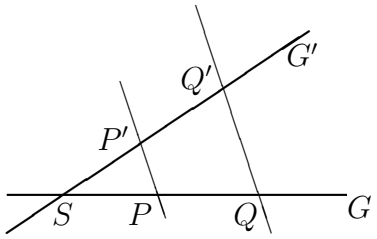
Eine erste Anwendung dieses Begriffs ist der

34.2 Strahlensatz
Seien (A, V, α) ein affiner Raum und G, G' Geraden in A, die sich in einem Punkt $S \in A$ schneiden. Sind P, Q Punkte $\neq S$ auf G und P', Q' Punkte $\neq S$ auf G', so sind folgende Aussagen äquivalent:

(a) *Die Geraden $\langle P, P' \rangle$ und $\langle Q, Q' \rangle$ sind parallel;*

(b) $\mathrm{TV}(S, P, Q) = \mathrm{TV}(S, P', Q')$;

(c) $\overrightarrow{QQ'} = \mathrm{TV}(S, P, Q)\overrightarrow{PP'}$.



Beweis: Setze $\lambda = \mathrm{TV}(S, P, Q)$ und $\lambda' = \mathrm{TV}(S, P', Q')$. Nach Voraussetzung sind $P \neq P'$ und $Q \neq Q'$.

(a)⇒(b) Sind $\langle P, P' \rangle$ und $\langle Q, Q' \rangle$ parallel, so ist $\overrightarrow{QQ'} = \mu \overrightarrow{PP'}, \quad \mu \in K$. Nach Voraussetzung gilt $\overrightarrow{SQ} = \lambda \overrightarrow{SP}$ und $\overrightarrow{SQ'} = \lambda' \overrightarrow{SP'}$ und damit

$$\lambda' \overrightarrow{SP'} - \lambda \overrightarrow{SP} = \overrightarrow{SQ'} - \overrightarrow{SQ} = \overrightarrow{QQ'} \quad = \quad \mu \overrightarrow{PP'} = \mu \overrightarrow{SP'} - \mu \overrightarrow{SP}.$$

Da $\overrightarrow{SP'}$ und $\overrightarrow{SP}$ linear unabhängig sind (denn G ist nicht parallel zu G'), folgt also $\lambda' = \mu = \lambda$.

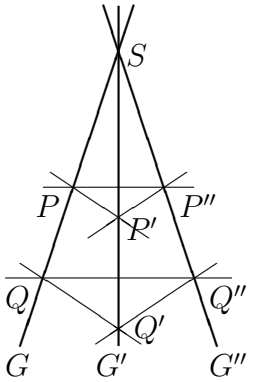
(b)⇒(c) Es gilt $\overrightarrow{QQ'} = \overrightarrow{SQ'} - \overrightarrow{SQ} = \lambda(\overrightarrow{SP'} - \overrightarrow{SP}) = \lambda \overrightarrow{PP'}$.

(c)⇒(a) Aus (c) folgt $\langle \overrightarrow{PP'} \rangle = \langle \overrightarrow{QQ'} \rangle$, d.h. $\langle P, P' \rangle$ ist parallel zu $\langle Q, Q' \rangle$. □

Der Strahlensatz ermöglicht nun den Beweis der folgenden Sätze, die auch bei einem axiomatischen Zugang zur Geometrie von Bedeutung sind:

34.3 Satz von Desargues

Sei (A, V, α) ein affiner Raum, und seien G, G', G'' drei Geraden in einer Ebene, die sich in einem Punkt S schneiden. Außerdem seien von S verschiedene Punkte $P, Q \in G$, $P', Q' \in G'$ und $P'', Q'' \in G''$ gegeben. Dann gilt: Ist $\langle P, P' \rangle \parallel \langle Q, Q' \rangle$ und $\langle P', P'' \rangle \parallel \langle Q', Q'' \rangle$, so ist $\langle P, P'' \rangle \parallel \langle Q, Q'' \rangle$.



Beweis: Aus dem Strahlensatz folgt unter den gegebenen Voraussetzungen

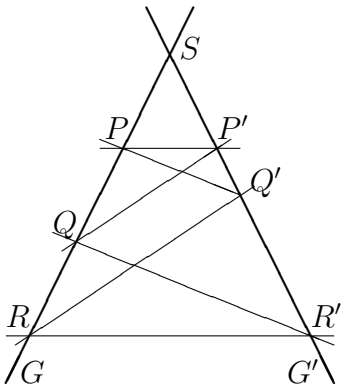
$$\mathrm{TV}(S, P, Q) = \mathrm{TV}(S, P', Q') \quad \text{und} \quad \mathrm{TV}(S, P', Q') = \mathrm{TV}(S, P'', Q''),$$

also $\mathrm{TV}(S, P, Q) = \mathrm{TV}(S, P'', Q'')$.

Wiederum mit dem Strahlensatz ergibt sich daraus $\langle P, P'' \rangle \parallel \langle Q, Q'' \rangle$. □

34.4 Satz von Pappus

Seien (A, V, α) ein affiner Raum, G und G' Geraden in A mit Schnittpunkt S, sowie $P, Q, R \in G$ und $P', Q', R' \in G'$, jeweils von S verschieden. Ist $\langle P, Q' \rangle \parallel \langle Q, R' \rangle$ und $\langle P', Q \rangle \parallel \langle Q', R \rangle$, dann ist $\langle P, P' \rangle \parallel \langle R, R' \rangle$.



Beweis: Aus dem Strahlensatz folgt

$$\lambda := \mathrm{TV}(S, P, Q) = \mathrm{TV}(S, Q', R'), \text{ also } \overrightarrow{SQ} = \lambda \overrightarrow{SP} \text{ und } \overrightarrow{SR'} = \lambda \overrightarrow{SQ'};$$
$$\mu := \mathrm{TV}(S, P', Q') = \mathrm{TV}(S, Q, R), \text{ also } \overrightarrow{SQ'} = \mu \overrightarrow{SP'} \text{ und } \overrightarrow{SR} = \mu \overrightarrow{SQ}.$$

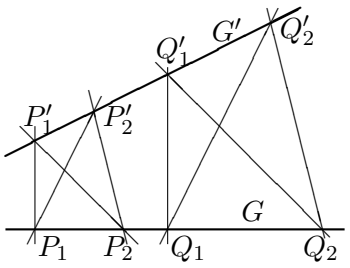
Somit ist $\overrightarrow{SR} = \lambda\mu \overrightarrow{SP}$ und $\overrightarrow{SR'} = \lambda\mu \overrightarrow{SP'}$, also $\mathrm{TV}(S, P, R) = \mathrm{TV}(S, P', R')$.

Wiederum mit dem Strahlensatz erhalten wir daraus $\langle P, P' \rangle \parallel \langle R, R' \rangle$. □

Auf ähnliche Weise kann man zeigen:

34.5 Scherensatz

G, G' seien Geraden im affinen Raum (A, V, α) mit den Punkten P_1,P_2,Q_1, $Q_2 \in G \setminus G'$ und P_1',P_2',Q_1', $Q_2' \in G' \setminus G$. Ist $\langle P_1, P_1'\rangle \parallel \langle Q_1, Q_1'\rangle$, $\langle P_1, P_2'\rangle \parallel \langle Q_1, Q_2'\rangle$ und $\langle P_2, P_1'\rangle \parallel \langle Q_2, Q_1'\rangle$, dann ist $\langle P_2, P_2'\rangle \parallel \langle Q_2, Q_2'\rangle$.



Sind P, Q, R Punkte auf einer Geraden, so nennt man R den *Mittelpunkt* von $\overline{PQ}$, wenn $\mathrm{TV}(P, Q, R) = \frac{1}{2}$.

Als weitere Aussagen der affinen ebenen Geometrie (über $\mathbb{Q}$) seien genannt:

Die Diagonalen eines Parallelogramms schneiden sich in ihren Mittelpunkten.

Die Seitenhalbierenden eines Dreiecks schneiden sich in einem Punkt und teilen sich im Verhältnis $2 : 1$.

34.6 Aufgaben

(1) Sei (A, V, α) ein affiner Raum über einem Körper K. $P, Q, R \in A$ seien paarweise verschiedene Punkte auf einer Geraden mit Teilverhältnis $\mathrm{TV}(P, Q, R) = \lambda \in K$.

Berechnen Sie alle möglichen $\mathrm{TV}(X, Y, Z)$ mit $\{X, Y, Z\} = \{P, Q, R\}$ in Abhängigkeit von λ.

(2) Beweisen Sie den **Kleinen Satz von Pappus***:*
(A, V, α) sei ein affiner Raum über einem Körper K, und $g \neq h$ seien zwei parallele Geraden aus A mit Punkten $P_1, P_2, P_3 \in g$ und $Q_1, Q_2, Q_3 \in h$. Dann gilt:

Ist $\langle P_1, Q_2\rangle \parallel \langle P_2, Q_3\rangle$ und $\langle P_2, Q_1\rangle \parallel \langle P_3, Q_2\rangle$, so ist $\langle P_1, Q_1\rangle \parallel \langle P_3, Q_3\rangle$.
Fertigen Sie eine Zeichnung an.

(3) (X, V, α) sei ein affiner Raum über dem Körper K mit $\mathrm{Char}\, K \neq 2, 3$. A, B und C seien drei Punkte aus X, die nicht auf einer Geraden liegen. Bezeichne M_a, M_b, M_c die Mittelpunkte der jeweils den Punkten A, B, C gegenüberliegenden Seiten des Dreiecks ABC.

Zeigen Sie, daß sich die drei Seitenhalbierenden des Dreiecks ABC in einem Punkt S schneiden, und daß gilt:

$$\mathrm{TV}(A, S, M_a) = \mathrm{TV}(B, S, M_b) = \mathrm{TV}(C, S, M_c) = \frac{3}{2}.$$

Welche Situation ergibt sich für den Fall $\mathrm{Char}\, K = 3$*?*

(4) (X, V, α) *sei ein affiner Raum über dem Körper* K. A, B, C *seien drei Punkte aus* X, *die nicht auf einer Geraden liegen, und* g *sei eine Gerade, welche die Geraden* $\langle B, C\rangle$, $\langle C, A\rangle$ *und* $\langle A, B\rangle$ *in den Punkten* P, Q *und* R *schneidet.*

Zeigen Sie, daß die Mittelpunkte der Strecken $\overline{AP}$, $\overline{BQ}$ *und* $\overline{CR}$ *auf einer Geraden liegen.*

(5) Sei (A, V, α) *ein zwei-dimensionaler affiner Raum über einem Körper* K. *Zeigen Sie:*

(i) *Zwei Geraden in* A *sind entweder parallel, oder sie besitzen genau einen Schnittpunkt.*

(ii) $g \neq h$ *seien zwei parallele Geraden in* A, *und* $P_1, P_2, P_3 \in g$, $Q_1, Q_2, Q_3 \in h$ *seien verschiedene Punkte.*

Genau dann ist $\mathrm{TV}(P_1, P_2, P_3) = \mathrm{TV}(Q_1, Q_2, Q_3)$, *wenn die Geraden* $\langle P_1, Q_1\rangle$, $\langle P_2, Q_2\rangle$ *und* $\langle P_3, Q_3\rangle$ *sich in einem Punkt schneiden oder alle parallel sind.*

35 Affine Abbildungen

Zu algebraischen Strukturen wie etwa Gruppen und Moduln haben wir jeweils die strukturerhaltenden Abbildungen betrachtet. Wir tun dies auch für affine Räume:

35.1 Definition
Seien (A_1, V_1, α_1) und (A_2, V_2, α_2) affine Räume. Eine Abbildung $f : A_1 \to A_2$ heißt *affin*, *affiner Homomorphismus* oder auch *Affinität*, wenn für jedes $O \in A_1$ die Abbildung

$$f_V : V_1 \to V_2, \quad \overrightarrow{OP} \mapsto \overrightarrow{f(O)f(P)},$$

ein Vektorraumhomomorphismus ist.

f heißt *semiaffin*, wenn f_V eine semilineare Abbildung ist (bzgl. eines Körperautomorphismus φ von K, vgl. 28.1).

Wir stellen einige Beobachtungen zu Affinitäten zusammen. Einige davon gelten auch für Semiaffinitäten.

35.2 Beschreibung von Affinitäten
Sei $f : A_1 \to A_2$ eine Affinität. Dann gilt:

(1) *Für $P, Q \in A_1$ gilt $\overrightarrow{f(P)f(Q)} = f_V(\overrightarrow{PQ})$.*

(2) *Die zugehörige Abbildung $f_V : V_1 \to V_2$, $\overrightarrow{OP} \mapsto \overrightarrow{f(O)f(P)}$ ist unabhängig von der Wahl von $O \in A_1$.*

(3) *Zu einem Vektorraumhomomorphismus $g : V_1 \to V_2$ und $O_1 \in A_1$, $O_2 \in A_2$ gibt es genau eine Affinität $f : A_1 \to A_2$ mit $f(O_1) = O_2$ und $f_V = g$.*

Beweis: (1) Für $P, Q \in A$ gilt:

$$\begin{aligned} \overrightarrow{f(P)f(Q)} &= -\overrightarrow{f(O)f(P)} + \overrightarrow{f(O)f(Q)} \\ &= -f_V(\overrightarrow{OP}) + f_V(\overrightarrow{OQ}) \\ &= f_V(\overrightarrow{PO} + \overrightarrow{OQ}) \;=\; f_V(\overrightarrow{PQ}). \end{aligned}$$

(2) Seien $Q, O' \in A_1$, und bezeichne f'_V die Abbildung

$$f'_V : V_1 \to V_2, \quad \overrightarrow{O'P'} \mapsto \overrightarrow{f(O')f(P')}, \quad P' \in A.$$

Ist $v = \overrightarrow{OP} = \overrightarrow{O'P'} \in V_1$ für $P, P' \in A$, so gilt mit (1)

$$f'_V(v) = \overrightarrow{f(O')f(P')} = f_V(\overrightarrow{O'P'}) = f_V(\overrightarrow{OP}) = f_V(v).$$

(3) Die Affinität $f : A_1 \to A_2$ ist gegeben durch

$$f(P) = O_2 + g(\overrightarrow{O_1P}) \quad \text{für } P \in A_1.$$

□

Beispiel
Nach den Ausführungen zu (3) ist für jeden affinen Raum (A, V, α) der Identität id_V und jedem Punktepaar $O_1, O_2 \in A$ die Affinität

$$\begin{aligned} A \to A, \quad P \mapsto O_2 + \overrightarrow{O_1P} &= O_1 + \overrightarrow{O_1O_2} + \overrightarrow{O_1P} \\ &= P + \overrightarrow{O_1O_2} \end{aligned}$$

zugeordnet. Man nennt diese die *Translation* mit dem Vektor $\overrightarrow{O_1O_2}$.

Eine Translation $t : A \to A$ ist somit schon durch das Bild eines Punktes $O \in A$ bestimmt. Es gilt dann

$$t(P) = P + \overrightarrow{O\,t(O)}.$$

Offensichtlich ergibt die Hintereinanderausführung von zwei Translationen wieder eine Translation.

Die Menge der Translationen bildet eine additive Gruppe ($\simeq (V, +)$).

35.3 Eigenschaften affiner Abbildungen

Sei $f : A_1 \to A_2$ eine Affinität von affinen Räumen. Dann gilt:

(1) *Für $P, Q \in A_1$ gilt $f(P) = f(Q)$ genau dann, wenn $f_V(\overrightarrow{PQ}) = 0$.*

(2) *Für kollineare Punkte $P, Q, R \in A_1$ gilt*

$$\mathrm{TV}(P, Q, R) = \mathrm{TV}(f(P), f(Q), f(R)).$$

(3) *Ist B_1 affiner Unterraum von A_1, so ist $f(B_1)$ affiner Unterraum von A_2.*

(4) *Sind B_1 und C_1 parallele affine Unterräume von A_1, so sind $f(B_1)$ und $f(C_1)$ parallele affine Unterräume von A_2.*

(5) *Ist B_2 affiner Unterraum von A_2, so gilt $f^{-1}(B_2) = \emptyset$, oder $f^{-1}(B_2)$ ist affiner Unterraum von A_1.*

(6) *Sind B_2 und C_2 parallele affine Unterräume von A_2, so sind $f^{-1}(B_2)$ und $f^{-1}(C_2)$ parallele affine Unterräume in A_1, wenn beide Urbilder nicht leer sind.*

Beweis: (1) Ist $f(P) = f(Q)$, so gilt $0 = \overrightarrow{f(P)f(Q)} = f_V(\overrightarrow{PQ})$.

(2) Sei $\mathrm{TV}(P, Q, R) = \lambda \in K$, also $\overrightarrow{PR} = \lambda\overrightarrow{PQ}$. Dann ist

$$\overrightarrow{f(P)f(R)} = f_V(\overrightarrow{PR}) = f_V(\lambda\overrightarrow{PQ}) = \lambda f_V(\overrightarrow{PQ}) = \lambda\overrightarrow{f(P)f(Q)},$$

also $\mathrm{TV}(f(P), f(Q), f(R)) = \lambda$.

(3) Ist $B_1 = P + V_B$, so gilt $\quad f(B_1) = f(P) + f_V(V_B)$.

(4) Sind V_{B_1} bzw. V_{C_1} die Vektorräume zu B_1 bzw. C_1, und gilt $V_{B_1} \subset V_{C_1}$, so folgt $f_V(V_{B_1}) \subset f_V(V_{C_1})$.

(5) Sei $B_2 = p_2 + V_{B_2}$. Ist $f^{-1}(B_1) \neq \emptyset$, so gibt es ein $P_1 \in A_1$ mit $f(P_1) = P_2 \in B_2$, und es gilt

$$f^{-1}(B_2) = P_1 + f_V^{-1}(V_{B_2}).$$

(6) Ist $C_2 = Q_2 + V_{C_2}$ und $V_{B_2} \subset V_{C_2}$, so ist $\quad f_V^{-1}(V_{B_2}) \subset f_V^{-1}(V_{C_2})$. □

35.4 Komposition von Affinitäten
Sind $f : A_1 \to A_2$ und $g : A_2 \to A_3$ affine Abbildungen, so ist auch deren Komposition $g \circ f : A_1 \to A_3$ eine affine Abbildung.

Beweis: Für $g \circ f$ haben wir

$$\overrightarrow{OP} \mapsto \overrightarrow{g \circ f(O)\ g \circ f(P)} = g_V\Big(\overrightarrow{f(O)f(P)}\Big) = g_V \circ f_V(\overrightarrow{OP}).$$

Damit sieht man, daß dies eine affine Abbildung ist. □

Eine affine Abbildung $f : A_1 \to A_2$ ist offensichtlich genau dann bijektiv, wenn der zugehörige Vektorraumhomomorphismus $f_V : V_1 \to V_2$ bijektiv ist.

35.5 Affine Bijektionen
Sei (A, V, α) ein affiner Raum. Die Menge $\mathrm{Aff}(A)$ *der affinen Bijektionen bildet eine Gruppe, die* Gruppe der Affinitäten *von A.*

Die Abbildung

$$\mathrm{Aff}(A) \to \mathrm{End}(V), \quad f \mapsto f_V,$$

ist ein Gruppenhomomorphismus. Sie hat als Bild die linearen Automorphismen von V ($\mathrm{GL}(V)$).

Der Kern der Abbildung besteht aus den Translationen von A.

Beweis: Nach 35.4 ist die Komposition von affinen Abbildungen affin. Damit ist die erste Behauptung klar.

Sind f und g Affinitäten, so gehört nach 35.4 zu $g \circ f$ die lineare Abbildung $g_V \circ f_V$, also ein Homomorphismus.

f_V ist genau dann die Identität, wenn f Translation ist. □

35.6 Kennzeichnung von Affinitäten
Für eine Abbildung $f : A_1 \to A_2$ von affinen Räumen sind folgende Aussagen äquivalent:

(a) *f ist eine Affinität;*

(b) *es gibt zwei Punkte $O_1 \in A_1$, $O_2 \in A_2$ und einen Vektorraumhomomorphismus $f_V : V_1 \to V_2$ mit*

$$f(P) = O_2 + f_V(\overrightarrow{O_1P}) \text{ für } P \in A_1.$$

Ist $\operatorname{Char} K \neq 2$*, so sind dazu äquivalent:*

(c) *f läßt das Teilverhältnis von je drei kollinearen Punkten invariant;*

(d) *für $a \in K$, $P, Q, P', Q' \in A_1$ gilt:*
Ist $\overrightarrow{P'Q'} = a\overrightarrow{PQ}$, so ist $\overrightarrow{f(P')f(Q')} = a\overrightarrow{f(P)f(Q)}$.

Beweis: (a)⇔(b) ergibt sich aus 35.2.

(a)⇒(c) wurde bereits in 35.3 gezeigt.

(a)⇒(d) $\overrightarrow{f(P')f(Q')} = f_V(\overrightarrow{P'Q'}) = f_V(a\overrightarrow{PQ}) = af_V(\overrightarrow{PQ}) = a\overrightarrow{f(P)f(Q)}$.

(d)⇒(a) Für $O \in A_1$ definieren wir $f_V(v) = \overrightarrow{f(O)f(O+v)}$.
Dann gilt für alle $P, Q \in A_1$

$$f_V(\overrightarrow{PQ}) = \overrightarrow{f(O)f(O+\overrightarrow{PQ})} = \overrightarrow{f(P)f(Q)},$$

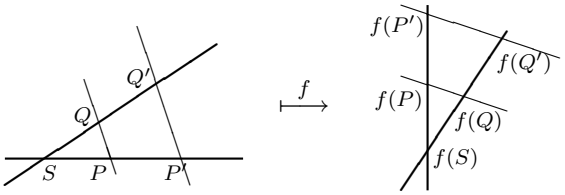
denn für $v = \overrightarrow{PQ}$ folgt aus $\overrightarrow{PQ} = \overrightarrow{OO} + \overrightarrow{PQ}$:

$$f_V(\overrightarrow{PQ}) = \overrightarrow{f(O)f(O+\overrightarrow{PQ})} = \overrightarrow{f(P)f(Q)}.$$

Die Linearität von f_V ersieht man aus

$$\begin{aligned} f_V(v+w) &= \overrightarrow{f(O)f(O+v+w)} \\ &= \overrightarrow{f(O)f(O+v)} + \overrightarrow{f(O+v)f(O+v+w)} = f_V(v) + f_V(w), \\ f_V(rv) &= \overrightarrow{f(O)f(O+rv)} = r\overrightarrow{f(O)f(O+v)} = rf_V(v). \end{aligned}$$

(c)⇒(d) folgt aus dem Strahlensatz (34.2):



□

Wir wissen, daß durch die Bilder einer Basis ein Vektorraumhomomorphismus eindeutig festgelegt ist. Analog haben wir für affine Räume:

35.7 Bestimmung von Affinitäten
Seien (A_1, V_1, α_1) und (A_2, V_2, α_2) affine Räume.
Bilden $P_0, \ldots, P_n$ eine Basis von A_1, so gibt es zu Punkten $Q_0, \ldots, Q_n$ aus A_2 genau eine Affinität $f : A_1 \to A_2$ mit

$$f(P_i) = Q_i \text{ für } i = 0, \ldots, n.$$

Bilden die $Q_0, \ldots, Q_n$ eine Basis von A_2, so ist f eine bijektive Affinität.

Beweis: Nach Voraussetzung bilden die Vektoren $\overrightarrow{P_0P}_1, \ldots, \overrightarrow{P_0P}_n$ eine Basis von V_1, und $\overrightarrow{Q_0Q}_1, \ldots, \overrightarrow{Q_0Q}_n$ sind Vektoren in V_2. Damit haben wir einen Homomorphismus

$$f_V : V_1 \to V_2, \quad \overrightarrow{P_0P_i} \mapsto \overrightarrow{Q_0Q_i}, \quad i = 1, \ldots, n.$$

Für $P = P_0 + \overrightarrow{P_0P}$ setzen wir

$$f : A_1 \to A_2, \quad P \;\mapsto\; Q_0 + f_V(\overrightarrow{P_0P}).$$

Sind die $\overrightarrow{Q_0Q_i}$, $i = 1, \ldots, n$, eine Basis von V_2, so ist f_V ein Isomorphismus, und f ist bijektiv. □

Ein weiteres wichtiges Beispiel für Affinitäten sind die

35.8 Projektionen auf einen Unterraum
Sei (A, V, α) ein affiner Raum. (A_1, V_1, α_1) und (A_2, V_2, α_2) seien affine Unterräume mit $V = V_1 \oplus V_2$. Dann gilt:

Zu jedem $P \in A$ besteht $(P + V_1) \cap A_2$ aus genau einem Punkt, die Zuordnung

$$p : A \to A, \quad P \mapsto (P + V_1) \cap A_2,$$

ist eine Affinität von A mit $p \circ p = p$, $p(A) = A_2$, und für $Q \in A_2$ gilt $p^{-1}(Q) = Q + V_1$.

Man nennt p die Projektion auf A_2 längs A_1.

Beweis: Betrachten wir die Vektorraumprojektion

$$p_V : V \to V, \quad v = v_1 + v_2 \;\mapsto\; v_2.$$

Dafür gilt $p_V(V) = V_2$, $p_V \circ p_V = p_V$ und $\operatorname{Kern} p_V = V_1$. Wir wählen nun einen Punkt $O_2 \in A_2$ und definieren die Affinität

$$p : A \to A, \quad P \mapsto O_2 + p_V(\overrightarrow{O_2P}).$$

Wegen $p_V(\overrightarrow{O_2P}) \in V_2$ ist $p(P) \in A_2$. Es gilt $p(O_2) = O_2$, und für $v = \overrightarrow{OP_2}$ haben wir

$$\overrightarrow{Pp(P)} = \overrightarrow{O_2p(P)} - \overrightarrow{O_2P} = p_V(\overrightarrow{O_2P}) - \overrightarrow{O_2P} = v_2 - v = -v_1 \quad \in V_1.$$

Daraus ergibt sich $p(P) = P + \overrightarrow{P\,p(P)} = P - v_1 \in P + V_1$.

Wegen $V_1 \cap V_2 = 0$ ist $p(P)$ der einzige Punkt in $(P + V_1) \cap A_2$, und p ist die in der Behauptung angesprochene Abbildung. Für sie gilt

$$\begin{aligned} p \circ p(P) &= p\Big(O_2 + p_V(\overrightarrow{O_2P})\Big) \\ &= O_2 + p_V\Big(\overrightarrow{O_2, O_2 + p_V(\overrightarrow{O_2P})}\Big) \\ &= O_2 + p_V \circ p_V(\overrightarrow{O_2P}) \\ &= O_2 + p_V(\overrightarrow{O_2P}) \;=\; p(P). \end{aligned}$$

Außerdem ist

$$P(A) = O_2 + p_V(V) = O_2 + V_2 = A_2.$$

Für alle $Q \in A_2$ ist $\overrightarrow{O_2Q} \in V_2$, und daher

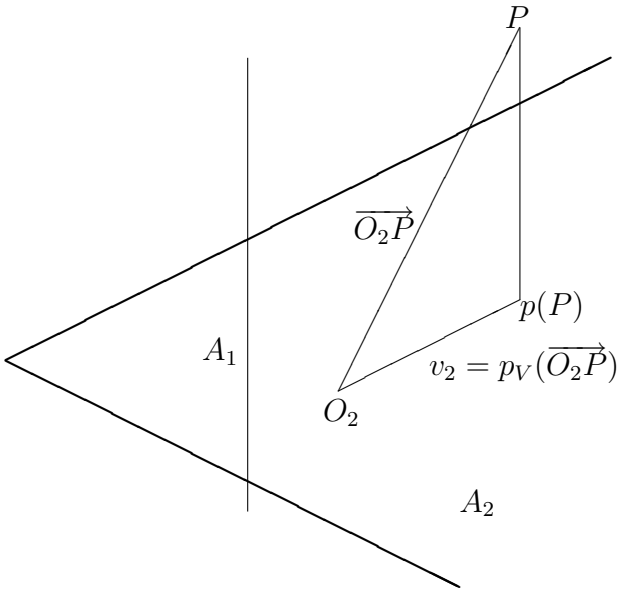
$$p(Q) = O_2 + p_V(\overrightarrow{O_2Q}) = O_2 + \overrightarrow{O_2Q} = Q.$$

Somit ist $Q \in p^{-1}(Q)$ und

$$p^{-1}(Q) = Q + p_V^{-1}(0_V) = Q + \operatorname{Kern} p_V = Q + V_1.$$

□

Beispiel
Wir wollen uns den oben betrachteten Fall im $\mathbb{R}^3$ mit einer Geraden A_1 und einer Ebene A_2 veranschaulichen:



In 33.5 haben wir gesehen, daß in jedem affinen Raum (A, V, α) eine Bijektion zwischen den Mengen A und V besteht. Die dort betrachtete Abbildung ergibt sogar eine Affinität zwischen (A, V, α) und dem Standardraum zu V (vgl. 33.4).

35.9 Affinität zum Standardraum
Sei (A, V, α) ein affiner Raum und $(V, V, +)$ der affine Standardraum über V. Dann bestimmt für ein (beliebiges) $O \in A$ die Abbildung

$$t_O : A \to V, \quad P \mapsto \overrightarrow{OP},$$

eine bijektive Affinität (einen affinen Isomorphismus) $(A, V, \alpha) \to (V, V, +)$.

Beweis: Nach Definition gilt für $P, Q \in A$

$$\overrightarrow{t_O(P)\, t_O(Q)} = t_O(P) - t_O(Q) = \overrightarrow{OQ} - \overrightarrow{OP} = \overrightarrow{PQ} \;=\; \mathrm{id}_V(\overrightarrow{PQ}).$$

Die Zuordnung $\overrightarrow{PQ} \mapsto \overrightarrow{t_O\,(P) t_O(Q)} = \mathrm{id}_V(\overrightarrow{PQ})$ ist also ein Vektorraumhomomorphismus, d.h. t_O ist eine Affinität. □

Für einen K-Vektorraum V mit Basis $v_1, \ldots, v_n$ haben wir *Koordinatenisomorphismen* definiert:

$$\varphi : V \to K^n, \quad v_i \mapsto e_i, \quad e_i \text{ Standardbasis.}$$

Setzen wir diesen Isomorphismus mit dem in 35.9 betrachteten zusammen, so erhalten wir auch Koordinaten für einen affinen Raum:

35.10 Definition
Sei (A, V, α) ein n-dimensionaler affiner Raum über K. Ist $P_0, \ldots, P_n$ eine Basis (Koordinatensystem) von A, so nennt man die Abbildung

$$\begin{array}{rccccc} \varrho_P : A & \longrightarrow & V & \longrightarrow & K^n, \\ P_i & \longmapsto & \overrightarrow{P_0 P_i} & \longmapsto & e_i, \end{array}$$

die *zu* $P_0, \ldots, P_n$ *gehörige Koordinatendarstellung.*

Für $P \in A$ ist $P = P_0 + \sum\limits_{i=1}^{n} r_i \overrightarrow{P_0 P_i}$ mit eindeutig bestimmten $r_i \in K$, und es gilt

$$\varrho_P(P) = (r_1, \ldots, r_n) \in K^n.$$

Faßt man K^n als affinen Raum auf, so ist ϱ_P eine bijektive Affinität (affiner Isomorphismus).

Sei nun $Q_0, \ldots, Q_n$ eine Basis von A mit Koordinatendarstellung

$$\varrho_Q : A \to V \to K^n, \quad Q_i \mapsto \overrightarrow{Q_0 Q_i} \mapsto e_i.$$

Somit haben wir das kommutative Diagramm mit affinen Isomorphismen

$$\begin{array}{ccccc} A & \xrightarrow{t_P} & V & \xrightarrow{\varphi_P} & K^n \\ \| & & \downarrow{\scriptstyle t_Q \circ t_P^{-1}} & & \downarrow{\scriptstyle \varrho_Q \circ \varrho_P^{-1}} \\ A & \xrightarrow{t_Q} & V & \xrightarrow{\varphi_Q} & K^n \end{array}$$

Wir wollen feststellen, wie sich die zugehörigen Koordinaten transformieren.

35.11 Transformation von Koordinaten
Mit den obigen Bezeichnungen seien $\varrho_P(X)$ und $\varrho_Q(X)$ für $X \in A$ die zugehörigen Koordinaten. Dann gibt es genau eine Matrix $T \in \mathrm{GL}(n,K)$ und einen Vektor $u \in K^n$ mit

$$\varrho_Q(X) = u + \varrho_P(X)T \quad \text{für alle } X \in A.$$

Dabei ist $u = \varrho_Q(P_0)$.

Beweis:

$$\begin{aligned}
\varrho_Q(X) &= \varphi_Q(\overrightarrow{Q_0X}) = \varphi_Q(\overrightarrow{Q_0P_0} + \overrightarrow{P_0X}) = \varphi_Q(\overrightarrow{Q_0P_0}) + \varphi_Q(\overrightarrow{P_0X}) \\
&= \varrho_Q(P_0) + \varphi_Q \circ \varphi_P^{-1} \circ \varphi_P(\overrightarrow{P_0X}) \\
&= \varrho_Q(P_0) + \varphi_Q \circ \varphi_P^{-1}\Big(\varrho_P(X)\Big) \\
&= \varrho_Q(P_0) + \varrho_P(X) \cdot \mathrm{Mat}_E(\varphi_Q \circ \varphi_P^{-1}) \\
&= \varrho_Q(P_0) + \varrho_P(X) \cdot \mathrm{Mat}_{Q,P}(\mathrm{id}_V).
\end{aligned}$$

Mit $u = \varrho_Q(P_0)$ und $T = \mathrm{Mat}_{Q,P}(\mathrm{id}_V)$ haben wir die gewünschte Beziehung. □

35.12 Satz
Seien (A,V,α) ein affiner Raum und $(P_0,\ldots,P_n)$ ein Koordinatensystem. Dann gibt es zu jeder invertierbaren Matrix $T \in K^{(n,n)}$ und zu jedem $u \in K^n$ genau ein Koordinatensystem $(Q_0,\ldots,Q_n)$ von A mit

$$\varrho_Q(X) = u + \varrho_P(X)T \text{ für alle } X \in A,$$

wobei ϱ_P, ϱ_Q die entsprechenden Koordinatendarstellungen bezeichnen. Es ist

$$\begin{aligned}
Q_0 &= \varrho_P^{-1}(-uT^{-1}) \quad \text{und} \\
Q_i &= \varrho_P^{-1}(e_iT^{-1} - uT^{-1}) \quad \text{für } i = 1,\ldots,n.
\end{aligned}$$

Beweis: Zu einer Basis $\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_n}$ von V und invertierbarem $T \in K^{(n,n)}$ gibt es genau eine Basis $W = (w_1,\ldots,w_n)$ von V mit

$$\mathrm{Mat}_{W,P}(\mathrm{id}_V) = T.$$

Damit sind die obigen Behauptungen leicht nachzuprüfen. □

35.13 Definition
Seien (A,V,α) ein affiner Raum und $(P_0,\ldots,P_n)$ ein Koordinatensystem. Als *Koordinatentransformation* bezeichnet man jedes Paar (u,T) mit $u \in K^n$ und $T \in \mathrm{GL}(n,K)$ und den *Transformationsgleichungen*

$$\varrho_Q(X) = u + \varrho_P(X)T \qquad \text{bzw. } \varrho_P(X) = \Big(\varrho_Q(X) - u\Big)T^{-1}.$$

Sehen wir uns dazu ein Beispiel in einer affinen Ebene an.

Beispiel
Über einem Körper K betrachten wir den Standardraum (K^2, K^2, α) mit dem kanonischen Koordinatensystem

$$(P_0, P_1, P_2) = \Big((0,0),(1,0),(0,1)\Big).$$

Als neues Koordinatensystem wollen wir

$$(Q_0, Q_1, Q_2) = \Big((1,1),(4,1),(2,3)\Big).$$

Zu dem beliebigen Punkt $X = (x_1, x_2) = P_0 + x_1\overrightarrow{P_0P_1} + x_2\overrightarrow{P_0P_2}$ berechnen wir $\varrho_Q(X) = (y_1, y_2)$. Zunächst suchen wir $\varrho_Q(P_0)$, $\varrho_Q(P_1)$, $\varrho_Q(P_2)$:

$$\begin{array}{rcccl}
\overrightarrow{Q_0P_0} & = & (-1,-1) & = & y_1(3,0) + y_2(1,2) = -\frac{1}{6}(3,0) - \frac{1}{2}(1,2), \\
& \Rightarrow & \varrho_Q(P_0) & = & \left(-\frac{1}{6}, -\frac{1}{2}\right); \\
\overrightarrow{Q_0P_1} & = & (0,-1) & = & \frac{1}{6}(3,0) - \frac{1}{2}(1,2), \\
& \Rightarrow & \varrho_Q(P_1) & = & \left(\frac{1}{6}, -\frac{1}{2}\right); \\
\overrightarrow{Q_0P_2} & = & (-1,0) & = & -\frac{1}{3}(3,0) + 0(1,2), \\
& \Rightarrow & \varrho_Q(P_2) & = & \left(-\frac{1}{3}, 0\right).
\end{array}$$

Damit finden wir nun

$$u = \varrho_Q(P_0) = \left(-\frac{1}{6}, -\frac{1}{2}\right), \quad T = \begin{pmatrix} \frac{1}{3} & 0 \\ -\frac{1}{6} & \frac{1}{2} \end{pmatrix}$$

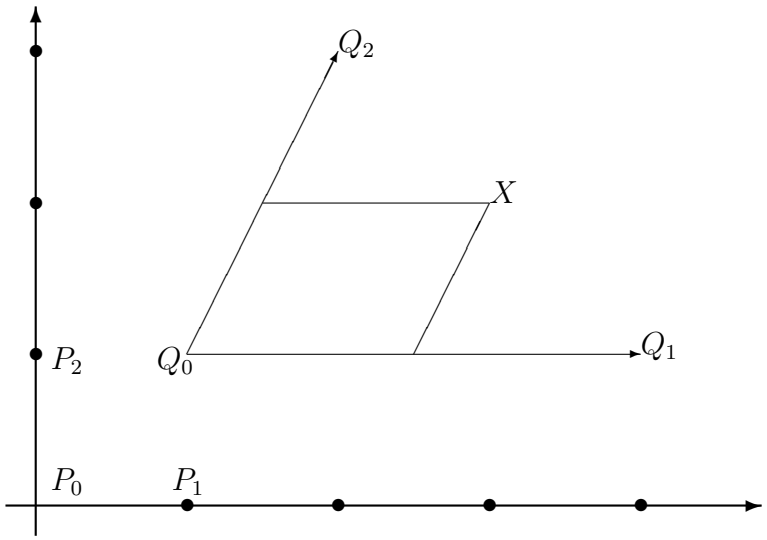
und haben dann die Transformationsgleichnungen

$$(y_1, y_2) = \left(-\frac{1}{6}, -\frac{1}{2}\right) + (x_1, x_2)\begin{pmatrix} \frac{1}{3} & 0 \\ -\frac{1}{6} & \frac{1}{2} \end{pmatrix}.$$

Zum Beispiel finden wir für den Punkt $X = (3,2)$ die neuen Koordinaten

$$\varrho_Q(X) = \left(-\frac{1}{6} + \frac{1}{3}\cdot 3 - \frac{1}{6}\cdot 2, -\frac{1}{2} + \frac{1}{2}\cdot 2\right) = \left(\frac{1}{2}, \frac{1}{2}\right).$$

Wir können diesen Sachverhalt in folgender Skizze veranschaulichen:



In Abschnitt 14 haben wir gesehen, daß sich die Lösungsmenge eines Gleichungssystems

$$(x_1, \ldots, x_n)B = (b_1, \ldots, b_n)$$

mit $B \in K^{(n,n)}$ in der Form $(a_0, \ldots, a_n) + U$ schreiben läßt, wobei $a_0, \ldots, a_n$ eine partikuläre Lösung ist und U der Lösungsraum des zugehörigen homogenen Systems, ein Untervektorraum von K^n.

Mit der inzwischen eingeführten Terminologie können wir sagen, daß die Lösungsmenge ein affiner Unterraum im Standardraum K^n ist. Andererseits ist jeder Unterraum darin von der angegebenen Form, ist also Lösungsmenge eines geeigneten Gleichungssystems.

Mit Hilfe der Koordinatendarstellung können wir zeigen, daß in jedem affinen Raum die affinen Unterräume in analoger Weise beschrieben werden können:

35.14 Darstellung von affinen Unterräumen
Seien (A, V, α) ein n-dimensionaler affiner Raum und $P_0, \ldots, P_n$ ein Koordinatensystem mit Koordinatendarstellung ϱ_P.

Zu jedem p-dimensionalen Unterraum $A' \subset A$ gibt es eine Matrix $B \in K^{(n,m)}$ und einen Vektor $b \in K^m$ mit $\operatorname{Rang} B = n - p = m$*, so daß die Lösungsmenge des linearen Gleichungssystems*

$$(*) \qquad (x_1, \ldots, x_n)B = b$$

affin isomorph zu A' ist. Es gilt

$$A' = \varrho_P^{-1}(\textit{Lösung von } (*)\,) \overset{\text{affin}}{\simeq} (a_0, \ldots, a_n) + \textit{homogene Lösung.}$$

Man nennt $(*)$ eine *Darstellung von A'.*

Beweis: Der zu $\varrho_P(A')$ gehörende Vektorraum $U \subset K^n$ hat die Dimension p, ist also Lösungsmenge eines homogenen Systems

$$xB = 0, \text{ mit } B \in K^{(n,m)}, \operatorname{Rang} B = n - p = m.$$

Für einen Punkt $Q \in A'$ setzen wir $b = \varrho_P(Q)B$. Dann ist $\varrho_P(Q)$ eine Lösung des Gleichungssystems

$$xB = b, \text{ mit der Lösungsmenge } \varrho_P(Q) + U = \varrho_P(A').$$

□

Beispiel
Im $\mathbb{Q}^3$ mit dem kanonischen Koordinatensystem sei der affine Unterraum

$$A' = \langle (0,0,0), (7,3,8), (0,6,1), (14,12,17) \rangle$$

gegeben. Wir wollen dazu ein Gleichungssystem finden. Der zugehörige Vektorraum ist

$$\langle (7,3,8), (0,6,1), (14,12,17) \rangle \;=\; \langle (7,3,8), (0,6,1) \rangle ,$$

also ist $\dim A' = 2$. Da der Nullpunkt (des kanonischen Koordinatensystems) in A' liegt, wird die Darstellung von A' ein homogenes System sein. Das Gleichungssystem

$$\begin{pmatrix} 7 & 3 & 8 \\ 0 & 6 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

hat $(45, -7, 42)^t$ als Lösung, und die Gleichung für A' ist daher

$$(x_1, x_2, x_3) \begin{pmatrix} 45 \\ -7 \\ 42 \end{pmatrix} = 0.$$

Überlegen wir uns nun, wie die Darstellung des Durchschnitts von zwei affinen Unterräumen aussieht.

35.15 Durchschnitt affiner Unterräume

(A, V, α) sei affiner Raum, $P_0, \ldots, P_n$ ein Koordinatensystem und ϱ_P die Koordinatendarstellung. A_1 und A_2 seien affine Unterräume mit den zugehörigen Gleichungen

$$\begin{aligned} &\text{(G1)} \quad xB_1 = b_1, \quad B_1 \in K^{n,m_1}, \quad b_1 \in K^{m_1}, \quad m_1 = n - \dim A_1, \\ &\text{(G2)} \quad xB_2 = b_2, \quad B_2 \in K^{n,m_2}, \quad b_2 \in K^{m_2}, \quad m_2 = n - \dim A_2. \end{aligned}$$

Setze $B = (B_1, B_2) \in K^{m,m_1+m_2}$, $b = (b_1, b_2) \in K^{m_1+m_2}$ und

$$\text{(G)} \qquad xB = b$$

Mit diesen Bezeichnungen gilt:

(1) (G) *ist das zu* $A_1 \cap A_2$ *gehörige Gleichungssystem, also*

$$A_1 \cap A_2 = \varrho_P^{-1}(\textit{Lösung von } (\mathrm{G})).$$

(2) A_1 *und* A_2 *sind genau dann parallel, wenn*

$$\mathrm{Rang}(B) = \max(\mathrm{Rang}\, B_1, \mathrm{Rang}\, B_2).$$

(3) A_1 *und* A_2 *sind teilweise parallel* ($V_1 \cap V_2 \neq 0$)*, wenn* $\mathrm{Rang}\, B < n$.

(4) A_1 *und* A_2 *sind windschief, wenn* $\mathrm{Rang}\, B = n$ *und* (G) *keine Lösung hat.*

Beweis: (1) Da ϱ_P injektiv ist, gilt

$$\begin{aligned} \varrho_P(A_1 \cap A_2) &= \varrho_P(A_1) \cap \varrho_P(A_2) \\ &= \text{Lösung von (G1)} \cap \text{Lösung von (G2)} \\ &= \text{Lösung von (G)}. \end{aligned}$$

(2) bis (4) folgen aus den Beziehungen $\mathrm{Kern}\, B = \mathrm{Kern}\, B_1 \cap \mathrm{Kern}\, B_2$ und $\mathrm{Rang}\, B_i = n - \dim \mathrm{Kern}\, B_i$ für $i = 1, 2$. □

Folgerung
Jeder affine Unterraum ist Durchschnitt von Hyperebenen.

Wir wollen noch eine Berechnungsmöglichkeit für die Gleichung einer Hyperebene angeben, die durch n linear unabhängige Punkte erzeugt wird.

35.16 Gleichung von Hyperebenen
Sei (A, V, α) *ein affiner Raum mit Koordinatensystem* $(P_0, \ldots, P_n)$ *und Koordinatendarstellung* ϱ_P. $\{Q_1, \ldots, Q_n\}$ *seien linear unabhängige Punkte in* A. *Zur Matrix*

$$A = \begin{pmatrix} \varrho_P(Q_1) & 1 \\ \vdots & \vdots \\ \varrho_P(Q_n) & 1 \\ 0 & 1 \end{pmatrix} \in K^{n+1,n+1}$$

bilden wir die adjunkte Matrix $\mathrm{Ad}(A) = (A_{kl})$ *(siehe 16.5) und setzen*

$$\begin{aligned} B &= (A_{n+1,1}, \ldots, A_{n+1,n})^t \in K^n, \\ b &= -A_{n+1,n+1} \in K\,. \end{aligned}$$

Dann ist $xB = b$ *eine Darstellung der von* $Q_1, \ldots, Q_n$ *erzeugten Hyperebene.*

Beweis: Für $Q \in A$ bilden wir die Matrix

$$A(Q) = \begin{pmatrix} \varrho_P(Q_1) & 1 \\ \vdots & \vdots \\ \varrho_P(Q_n) & 1 \\ \varrho_P(Q) & 1 \end{pmatrix} \in K^{n+1,n+1}$$

mit $\varrho_P(Q) = (x_1, \ldots, x_n)$, $x_{n+1} = 1$. Die Entwicklung nach der $(n+1)$-ten Zeile ergibt

$$\det A(Q) = \sum_{i=1}^{n+1} x_i A(Q)_{n+1,i}.$$

Man beachte, daß die Koeffizienten $A(Q)_{n+1,i} = A_{n+1,i}$ unabhängig von Q sind und daher – nach Festlegung von B und b –

$$\det A(Q) = \sum_{i=1}^{n} x_i A_{n+1,i} - (-A_{n+1,n+1}) = xB - b.$$

Offensichtlich ist $\det A(Q_i) = 0$, $i = 1, \ldots, n$, und somit

$$\langle Q_1, \ldots, Q_n \rangle \subset \{Q \in A \mid \varrho_P(Q)B = b\}.$$

Es ist noch $B \neq 0$ zu zeigen, dann gilt Gleichheit aus Dimensionsgründen. Aus $B = 0$ folgt für alle $Q \in A$

$$\det A(Q) = -b = \det A(Q_1) = 0.$$

Nach Voraussetzung kann $Q_1, \ldots, Q_n$ aber durch ein Q_0 zu einem Koordinatensystem ergänzt werden, und dann gilt

$$\det A(Q_0) = \begin{pmatrix} \varrho_P(Q_1) - \varrho_P(Q_0) & 0 \\ \vdots & \vdots \\ \varrho_P(Q_n) - \varrho_P(Q_0) & 0 \\ \varrho_P(Q_0) & 1 \end{pmatrix} \neq 0,$$

da die $\varrho_P(Q_i) - \varrho_P(Q_0)$ linear unabhängig sind. Also ist $B \neq 0$. □

Beispiel
Im $\mathbb{Q}^2$-Standardraum seien die Punkte $P = (p_1, p_2)$, $Q = (q_1, q_2)$ gegeben. Nach obigen Ausführungen bilden wir

$$A = \begin{pmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} p_2 - q_2 \\ p_1 - q_1 \end{pmatrix}, \quad b = p_2 q_1 - p_1 q_2.$$

Als Gleichung für die Hyperebene (Gerade) durch P und Q erhalten wir

$$(x_1, x_2) \begin{pmatrix} p_2 - q_2 \\ -p_1 + q_1 \end{pmatrix} = p_2 q_1 - p_1 q_2,$$

und äquivalent dazu die bekannte Form

$$\frac{x_2 - p_2}{x_1 - p_1} = \frac{q_2 - p_2}{q_1 - p_1}.$$

35.17 Aufgaben

(1) Im affinen Standardraum $V = \mathbb{R}^4$ seien die folgenden affinen Unterräume gegeben:

$$\begin{aligned} A_1 &:= \langle (-1, 2, -3, 4) \rangle, \\ A_2 &:= \langle (0, 1, 2, 3), (2, 5, 7, 3) \rangle, \\ A_3 &:= \langle (1, 2, 3, 0), (1, 2, 0, 1), (0, 1, 7, 2) \rangle \end{aligned}$$

Prüfen Sie für $i, j \in \{1, 2, 3\}$, ob A_i und A_j zueinander parallel, teilweise parallel, punktfremd oder windschief sind.

(2) Im affinen Standardraum $\mathbb{R}^3$ sei folgendes Koordinatensystem gegeben: $P_0 = (0, 0, 0)$, $P_1 = (1, 1, 0)$, $P_2 = (1, 0, 1)$, $P_3 = (0, 1, 1)$.

(i) *Berechnen Sie den Vektor u sowie die Matrix T (siehe §35), die die Koordinatentransformation in das folgende Koordinatensystem bewirken: $Q_0 = (1, 1, 1)$, $Q_1 = (2, 1, 3)$, $Q_2 = (0, 0, 0)$, $Q_3 = (2, 3, 2)$.*

(ii) *Geben Sie die zum Koordinatensystem (Q_0, Q_1, Q_2, Q_3) gehörenden Koordinatendarstellungen der folgenden Punkte aus dem $\mathbb{R}^3$ an: $A = (1, 5, 4)$, $B = (0, 2, -1)$, $C = (1, 4, 0)$.*

(3) Der affine Standardraum $\mathbb{R}^3$ sei mit dem kanonischen Koordinatensystem $((0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1))$ versehen. Auf welches Koordinatensystem führt die durch

$$u := (2, 0, 1) \quad \text{und} \quad T := \begin{pmatrix} 1 & 3 & 4 \\ 2 & 0 & 1 \\ 2 & 7 & 9 \end{pmatrix}$$

gegebene Koordinatentransformation?

(4) Betrachten Sie die affinen Unterräume im affinen Standardraum $\mathbb{R}^4$:

$$\begin{aligned} A_1 &:= \langle (0, 1, 0, -1), (2, 3, 4, 5), (1, -1, 2, -1), (0, 0, 0, -2) \rangle \\ A_2 &:= \langle (1, -3, 1, -3), (2, -4, 1, -2), (-1, -1, 1, -5) \rangle \end{aligned}$$

(a) *Berechnen Sie die Dimensionen von A_1 und A_2.*

(b) *Für A_1 und A_2 gebe man jeweils ein lineares Gleichungssystem an, für das A_1 bzw. A_2 Lösungsmenge ist.*

(c) *Ermitteln Sie eine Parameterdarstellung für den affinen Unterraum $A_1 \cap A_2$.*

36 Affine Quadriken

Sei (A, V, α) ein affiner Raum über einem Körper K.
Für eine affine Abbildung $f : A \to K$ ist die Menge der Urbilder von $0 \in K$

$$\{P \in A \mid f(P) = 0\}$$

eine Hyperebene in A.

Im folgenden wollen wir uns mit der Urbildmenge von 0 bei quadratischen Abbildungen befassen. Obwohl die *quadratischen* Abbildungen nicht unbedingt in die *lineare* Algebra gehören, sind die damit verbundenen Fragestellungen doch weitgehend mit linearen Methoden zu behandeln. Wir werden dabei auf die Theorie der (symmetrischen) Bilinearformen auf Vektorräumen zurückgreifen.

36.1 Definition
Sei (A, V, α) ein affiner Raum. Eine Abbildung $h : A \to K$ nennen wir *quadratisch*, wenn für jedes $P_0 \in A$ gilt:

$$h(P) = h(P_0) + \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}) + 2f_{P_0}(\overrightarrow{P_0P}), \quad P \in A,$$

wobei $\beta : V \times V \to K$ eine nicht-triviale, symmetrische Bilinearform und $f_{P_0} : V \to K$ eine Linearform ist.

Wir werden gleich zeigen, daß die zugehörige Bilinearform β unabhängig vom Ausgangspunkt P_0 ist, während die Linearform f_{P_0} von der Wahl von P_0 abhängt.

36.2 Abhängigkeit vom Ausgangspunkt
Seien (A, V, α) ein endlich-dimensionaler affiner Raum und $h : A \to K$ eine quadratische Abbildung. Dann gilt für $P_0, Q_0 \in A$:

(1)
$$\begin{aligned} h(P) &= \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}) + 2f_{P_0}(\overrightarrow{P_0P}) + h(P_0) \\ &= \beta(\overrightarrow{Q_0P}, \overrightarrow{Q_0P}) + 2f_{Q_0}(\overrightarrow{Q_0P}) + h(Q_0) \end{aligned}$$
mit
$$\begin{aligned} f_{P_0} - f_{Q_0} &= \beta(\overrightarrow{Q_0P_0}, -) \\ h(P_0) - h(Q_0) &= \beta(\overrightarrow{P_0Q_0}, \overrightarrow{P_0Q_0}) + 2f_{Q_0}(\overrightarrow{P_0Q_0}) \end{aligned}$$

(2) *Ist $f_{Q_0} = \beta(\overrightarrow{Q_0X}, -)$ für ein $X \in A$, dann kann man $P_0 \in A$ so wählen, daß $f_{P_0} = 0$, d.h.*

$$h(P) = \beta(\overrightarrow{P_0Q_0}, \overrightarrow{P_0Q_0}) + h(P_0).$$

Dies gilt zum Beispiel, wenn β nicht-singulär ist.

(3) *Ist die Voraussetzung von (2) nicht erfüllt, dann kann man $P_0 \in A$ so wählen, daß $h(P_0) = 0$, also*

$$h(P) = \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}) + 2f_{P_0}(\overrightarrow{P_0P}).$$

Beweis: (1) Mit $\overrightarrow{Q_0P} = \overrightarrow{P_0P} + \overrightarrow{Q_0P_0}$ erhalten wir

$$\begin{aligned}
\beta(\overrightarrow{Q_0P}, \overrightarrow{Q_0P}) &= \beta(\overrightarrow{P_0P} + \overrightarrow{Q_0P_0}, \overrightarrow{P_0P} + \overrightarrow{Q_0P_0}) \\
&= \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}) + 2\beta(\overrightarrow{P_0P}, \overrightarrow{Q_0P_0}) + \beta(\overrightarrow{Q_0P_0}, \overrightarrow{Q_0P_0}), \\
2f_{Q_0}(\overrightarrow{Q_0P}) &= 2f_{Q_0}(\overrightarrow{P_0P}) + 2f_{Q_0}(\overrightarrow{Q_0P_0}), \\
h(Q_0) &= h(P_0) - \beta(\overrightarrow{Q_0P_0}, \overrightarrow{Q_0P_0}) - 2f_{Q_0}(\overrightarrow{Q_0P_0}).
\end{aligned}$$

Somit erhalten wir durch Einsetzen

$$\begin{aligned}
h(P) &= \beta(\overrightarrow{Q_0P}, \overrightarrow{Q_0P}) + 2f_{Q_0}(\overrightarrow{Q_0P}) + h(Q_0) \\
&= \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}) + 2\beta(\overrightarrow{P_0P}, \overrightarrow{Q_0P_0}) + \beta(\overrightarrow{Q_0P_0}, \overrightarrow{Q_0P_0}) \\
&\quad + 2f_{Q_0}(\overrightarrow{Q_0P_0}) + 2f_{Q_0}(\overrightarrow{P_0P}) + h(P_0) - 2f_{Q_0}(\overrightarrow{Q_0P_0}) \\
&\quad - \beta(\overrightarrow{Q_0P_0}, \overrightarrow{Q_0P_0}).
\end{aligned}$$

(2) Wähle $P_0 \in A$ so, daß $f_{Q_0} = -\beta(-, \overrightarrow{Q_0P_0})$. Dann haben wir

$$\begin{aligned}
h(P) &= \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}) + 2\beta(\overrightarrow{P_0P}, \overrightarrow{Q_0P_0}) + 2f_{Q_0}(\overrightarrow{P_0P}) + h(P_0) \\
&= \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}) + h(P_0).
\end{aligned}$$

(3) Sei nun $f_{Q_0} \neq \beta(\overrightarrow{Q_0X}, -)$ für alle $X \in A$.
Dann gibt es ein

$$P_1 \in A \text{ mit } \beta(\overrightarrow{Q_0P_1}, -) = 0 \text{ und } f_{Q_0}(\overrightarrow{Q_0P_1}) \neq 0.$$

Setzen wir $P_0 = Q_0 + \delta\overrightarrow{Q_0P_1}$ mit $\delta = \dfrac{-h(Q_0)}{2f_{Q_0}(\overrightarrow{Q_0P_1})}$, so ergibt sich

$$\begin{aligned}
h(P_0) &= h(Q_0) + \beta(\overrightarrow{Q_0P_0}, \overrightarrow{Q_0P_0}) + 2f_{Q_0}(\overrightarrow{Q_0P_0}) \\
&= h(Q_0) + \underbrace{\delta^2\beta(\overrightarrow{Q_0P_1}, \overrightarrow{Q_0P_1})}_{=0} + 2\frac{-h(Q_0)}{2f_{Q_0}(\overrightarrow{Q_0P_1})} \cdot f_{Q_0}(\overrightarrow{Q_0P_1}) \;=\; 0.
\end{aligned}$$

□

Für Berechnungen zu quadratischen Abbildungen sind natürlich wiederum Koordinatendarstellungen vorteilhaft:

36.3 Die Koordinatenform

Sei (A, V, α) ein n-dimensionaler affiner Raum mit dem Koordinatensystem $(P_0, \ldots, P_n)$.
Ist $h : A \to K$ eine quadratische Abbildung und $X = P_0 + \sum_{i=1}^{n} x_i\overrightarrow{P_0P_i}$, so gilt

$$h(X) = h(P_0) + \sum_{i,j=1}^{n} x_i\,\beta_{ij}x_j + 2\sum_{i=1}^{n} x_i\gamma_i,$$

mit $\beta_{ij} = \beta(\overrightarrow{P_0P_i}, \overrightarrow{P_0P_j})$ *und* $\gamma_i = f_{P_0}(\overrightarrow{P_0P_i})$.

Setzt man $x = (x_1, \ldots, x_n)$, $B = (\beta_{ij})$ *und* $C = (\gamma_1, \ldots, \gamma_n)^t$, *so kann man dies schreiben als*

$$h(X) = h(P_0) + xBx^t + 2xC.$$

Man nennt dies die *Koordinatenform* der quadratischen Abbildung.

Durch geeignete Transformationen konnten die Matrizen von symmetrischen Bilinearformen auf Diagonalform gebracht werden. Mit einem geschickt gewählten Koordinatensystem im affinen Raum A können wir auch die Koordinatenform von quadratischen Abbildungen standardisieren.

36.4 Hauptsatz über quadratische Abbildungen

Seien (A, V, α) *eine affiner Raum über* K, $\operatorname{Char} K \neq 2$ *und* $h : A \to K$ *eine quadratische Abbildung. Dann gibt es ein Koordinatensystem* $P_0, \ldots, P_n$ *in* A, *so daß* h *die folgende Koordinatenform hat:*

(1) *Ist für ein* $Q \in A$ $f_Q = \beta(\overrightarrow{QP_0}, -)$ *für ein geeignetes* $P_0 \in A$, *so gilt*

$$h(X) = \sum_{i=1}^{r} b_i x_i^2 + c, \quad 0 \neq b_i, c \in K \text{ für } i = 1, \ldots, r \leq n.$$

(2) *Andernfalls können wir ein Koordinatensystem so wählen, daß*

$$h(X) = \sum_{i=1}^{r} b_i x_i^2 - 2x_n, \quad 0 \neq b_i \in K \text{ für } i = 1, \ldots, r < n.$$

Dabei ist r *der Rang der symmetrischen Bilinearform* β.

Falls $K = \mathbb{R}$, *so kann* $b_i = 1$ *für* $1 \leq i \leq p$ *und* $b_i = -1$ *für* $p < i \leq r$ *gewählt werden.*

Für $K = \mathbb{C}$ *kann* $b_i = 1$ *für* $1 \leq i \leq r$ *erreicht werden.*

Beweis: (1) Wie wir in 36.2 gesehen haben, fällt bei der vorgeschlagenen Wahl von $P_0 \in A$ das lineare Glied in der Darstellung von h weg.

Nach Satz 27.10 können wir eine Orthogonalbasis von V bezüglich β finden (d.h. $\beta(v_i, v_j) = b_i\delta_{ij}$). Damit haben h und β die angegebene Koordinatenform.

(2) Zunächst wählen wir P_0 wie in 36.2(3) so, daß $h(P_0) = 0$. Da für alle $X \in A$ gilt $f_{P_0} \neq \beta(\overrightarrow{P_0X}, -)$, gibt es ein

$$d_n \in V \text{ mit } \beta(d_n, -) = 0 \text{ und } f_{P_0}(d_n) = -1.$$

Wähle nun eine Basis $d_{r+1}, \ldots, d_n$ von $\operatorname{Kern} \beta_\ell : V \to V^*$, $v \mapsto \beta(v, -)$ (also $\beta(d_i, -) = 0$) mit $f_{P_0} = 0$ für $r < i < n$. Diese Elemente ergänzen wir mit $d'_1, \ldots, d'_r$ zu einer Orthogonalbasis von V bezüglich β. Setzen wir dann

$$d_i = d'_i + f_{P_0}(d'_i)d_n \quad \text{für } i = 1, \ldots, r,$$

so gilt $f_{P_0}(d_i) = 0$ für $i = 1, \ldots, n-1$, und als lineares Glied bleibt in der Koordinatenform von h nur $-2x_n$ übrig.

In 29.1 hatten wir uns überlegt, daß über $\mathbb{R}$ die b_i zu ± 1 normiert werden können. Über $\mathbb{C}$ kann man sogar $b_i = 1$ erreichen (vgl. 29.2). □

Beispiel
Betrachte die quadratische Abbildung

$$h : \mathbb{Q}^2 \to \mathbb{Q}, \quad (x, y) \mapsto x^2 - 19y^2 + 6x - 38y + 15 = (x+3)^2 - 19(y+1)^2 + 25.$$

Betrachten wir diese Zuordnung über den Körpern $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$:

Über $\mathbb{Q}$ setzen wir $x' = x + 3$, $y' = y + 1$ und erhalten

$$h(x', y') = x'^2 - 19y'^2 + 25;$$

über $\mathbb{R}$ können wir $x'' = x'$, $y'' = \sqrt{19}\, y'$ setzen und

$$h(x'', y'') = x''^2 - y''^2 - 13,$$

über $\mathbb{C}$ ist $x''' = x'$, $y''' = \sqrt{19}\, iy'$ möglich und

$$h(x''', y''') = x'''^2 + y'''^2 + 25.$$

36.5 Definition
Sei (A, V, α) ein affiner Raum. Unter einer *affinen Quadrik* in A wollen wir die Nullstellenmenge einer quadratischen Abbildung $h : A \to K$ verstehen, also

$$\{Q \in A \mid h(Q) = 0\}.$$

Anmerkung: Für jedes $a \in K$ ist auch

$$\overline{h} : A \to K, \quad \overline{h}(P) = h(P) - a,$$

eine quadratische Abbildung. Die zugehörige Quadrik ist dann

$$\{Q \in A \mid \overline{h}(Q) = 0\} = \{Q \in A \mid h(Q) = a\}.$$

Zur Unterscheidung verschiedener Quadriken können wir diese zunächst in eine *Normalform* bringen. Die dabei angegebenen Bezeichnungen orientieren sich an der Realisierung der jeweiligen Punktmengen im $\mathbb{Q}^3$ bzw. $\mathbb{R}^3$:

36.6 Normalformen affiner Quadriken
Seien (A, V, α) ein affiner Raum über einem Körper K, $P_0 \in A$ und $h : A \to K$ eine quadratische Abbildung mit

$$h(P) = h(P_0) + f_{P_0}(\overrightarrow{P_0P}) + \beta(\overrightarrow{P_0P}, \overrightarrow{P_0P}).$$

Dann gibt es ein Koordinatensystem $(P_0, \ldots, P_n)$ von A, so daß für $Q \in A$ mit $h(Q) = 0$ die Koordinaten $\varrho_P(Q) = (x_1, \ldots, x_n)$ durch folgende Gleichungen beschrieben werden (mit $0 < r = \operatorname{Rang} B$):

(1) $\sum_{i=1}^{r} b_i x_i^2 = 0, \quad b_i \neq 0$ *Quadratischer Kegel*

(2) $\sum_{i=1}^{r} b_i x_i^2 = 1, \quad b_i \neq 0$ *Echte Mittelpunktsquadrik*

(3) $\sum_{i=1}^{r} b_i x_i^2 - 2x_n = 0, \quad r < n$ *Paraboloid*

Über $K = \mathbb{R}$ kann man erreichen:

(1) $\sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2 = 0,$ *Quadratischer Kegel*

(2) $\sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2 = 1,$ *Echte Mittelpunktsquadrik*

(3) $\sum_{i=1}^{p} x_i^2 - \sum_{i=p+1}^{r} x_i^2 - 2x_n = 0,$ *Paraboloid.*

Über $K = \mathbb{C}$ erhalten wir:

(1) $\sum_{i=1}^{r} x_i^2 = 0, \quad r \leq n$

(2) $\sum_{i=1}^{r} x_i^2 = 1, \quad r \leq n$

(3) $\sum_{i=1}^{r} x_i^2 - 2x_n = 0, \quad r < n.$

Beweis: Die angegebenen Darstellungen können unmittelbar aus 36.4 abgeleitet werden. □

36.7 Definition
Seien (A_1, V_1, α_1), (A_2, V_2, α_2) isomorphe affine Räume über dem Körper K.

Zwei Punktmengen $M_1 \subset A_1$ und $M_2 \subset A_2$ heißen *affin äquivalent*, wenn es einen affinen Isomorphismus $f : A_1 \to A_2$ mit $f(M_1) = M_2$ gibt.

Es ist leicht einzusehen, daß zwei affine Unterräume $M_1 \subset A_1$, $M_2 \subset A_2$ genau dann affin äquivalent sind, wenn sie gleiche Dimension haben.

Es stellt sich die Frage, welche Quadriken affin äquivalent sind. Lassen wir ausgeartete Fälle beiseite, so können wir aus den vorangehenden Ausführungen für reelle Räume folgendes feststellen:

36.8 Klassifikation affiner Quadriken

Im n-dimensionalen reellen affinen Raum A seien zwei Quadriken $M_1, M_2 \subset A$ gegeben, von denen keine in einer Hyperebene von A enthalten ist.

M_1 und M_2 sind genau dann affin äquivalent, wenn sie (bzgl. entsprechender Koordinatensysteme) dieselben Normalformen besitzen.

Als Beispiele geben wir die Möglichkeiten in der reellen Ebene und im reellen Raum an:

Quadriken in $\mathbb{R}^2$:

(1)	$r = 2:$	$x^2 + y^2 = 0$	ein Punkt
		$x^2 - y^2 = 0$	zwei sich schneidende Geraden
	$r = 1:$	$x^2 = 0$	(doppelt zählende) Gerade
(2)	$r = 2:$	$x^2 + y^2 = 1$	Kreis
		$x^2 - y^2 = 1$	Hyperbel
		$-x^2 - y^2 = 1$	$\emptyset$
	$r = 1:$	$x^2 = 1$	2 parallele Geraden
		$-x^2 = 1$	$\emptyset$
(3)		$x^2 - 2y = 0$	Parabel.

Quadriken in $\mathbb{R}^3$:

(1)	$x^2 + y^2 - z^2 = 0$	Kegel
	$x^2 - y^2 = 0$	2 sich schneidende Ebenen
	$x^2 = 0$	Doppelebene
(2)	$x^2 + y^2 + z^2 = 1$	Sphäre
	$x^2 + y^2 - z^2 = 1$	einschaliger Hyperboloid
	$x^2 - y^2 - z^2 = 1$	zweischaliger Hyperboloid
	$x^2 + y^2 = 1$	Kreiszylinder
	$x^2 - y^2 = 1$	Hyperbelzylinder
	$x^2 = 1$	2 parallele Ebenen
(3)	$x^2 + y^2 - 2z = 0$	elliptischer Paraboloid
	$x^2 - y^2 - 2z = 0$	hyperbolischer Paraboloid
	$x^2 - 2z = 0$	Parabelzylinder

37 Euklidische Räume

In den bisher betrachteten affinen Räumen hatten wir weder Längen noch Winkel zur Verfügung. Wie schon an den entsprechenden Stellen angedeutet, kann man solche Größen mit Hilfe des Skalarprodukts erfassen. Dies werden wir im folgenden tun.

37.1 Definition
Ein affiner Raum (A, V, α) über $\mathbb{R}$ (bzw. $\mathbb{C}$) heißt *euklidisch* (bzw. *unitär*), wenn auf V ein Skalarprodukt β definiert ist.

Beispiele

(1) Ist V ein Vektorraum mit Skalarprodukt, so ist der affine Standardraum über V euklidisch (bzw. unitär).

(2) Insbesondere ist für jedes $n \in \mathbb{N}$ der Standardraum $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) mit dem kanonischen Skalarprodukt ein euklidischer (bzw. unitärer) affiner Raum.

In solchen Räumen können wir weitere geometrische Begriffe einführen:

37.2 Definition
Seien (A, V, α) ein euklidischer affiner Raum mit Skalarprodukt $\beta : V \times V \to \mathbb{R}$. Für Punkte $P, Q, R \in A$ betrachten wir

(i) die *Strecke* von P nach Q

$$\overline{PQ} = \{P + \lambda\overrightarrow{PQ} \mid \lambda \in \mathbb{R},\ 0 \leq \lambda \leq 1\},$$

(ii) den *Abstand* von P und Q (die *Länge* von $\overline{PQ}$)

$$d(P, Q) = |P, Q| = \sqrt{\beta(\overrightarrow{PQ}, \overrightarrow{PQ})},$$

(iii) die *Halbgerade* durch Q mit Anfangspunkt P (falls $P \neq Q$)

$$\overrightarrow{PQ} = \{P + \lambda\overrightarrow{PQ} \mid \lambda \in \mathbb{R},\ \lambda > 0\},$$

(iv) den durch Q, P, R gegebenen *Winkel* (falls $P \neq Q$, $P \neq R$; vgl. 29.15)

$$\sphericalangle QPR = \sphericalangle(\overrightarrow{PQ}, \overrightarrow{PR}) = \arccos \frac{\beta(\overrightarrow{PQ}, \overrightarrow{PR})}{\left\|\overrightarrow{PQ}\right\| \left\|\overrightarrow{PR}\right\|}.$$

P heißt der *Scheitelpunkt* dieses Winkels, und die Halbgeraden $\overrightarrow{PQ}$ und $\overrightarrow{PR}$ bezeichnet man als seine *Schenkel* (oder *Halbachsen*).

(v) den *Winkelraum* des Winkels $\sphericalangle QPR$

$$< QPR = \{P + \lambda\overrightarrow{PQ} + \mu\overrightarrow{PR} \mid 0 \leq \lambda, \mu \in \mathbb{R}\}$$

Der Nachweis folgender Zusammenhänge ist eine leichte Übung:

37.3 Eigenschaften
Für Punkte P, Q und R im euklidischen Raum (A, V, α) gilt:

(i) $\overline{PQ} = \overline{QP}$.

(ii) $\overline{PQ} = \overleftarrow{PQ} \cap \overleftarrow{QP}$.

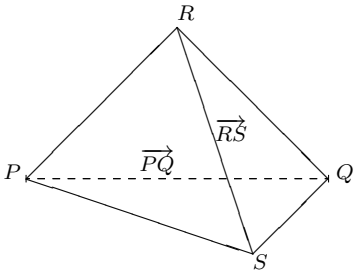
(iii) $\overleftarrow{PQ} = \overleftarrow{PR}$ *genau dann, wenn* $R \in \overleftarrow{PQ}$.

(iv) *Mit* $|-,-| : A \times A \to \mathbb{R}$ *wird* A *zu einem metrischen Raum, denn:*
$|P, Q| \geq 0$, *und* $|P, Q| = 0$ *genau dann, wenn* $P = Q$.
$|P, Q| = |Q, P|$
$|P, R| \leq |P, Q| + |Q, R|$ *(Dreiecksungleichung),*
Gleichheit gilt genau dann, wenn $Q \in \overline{PR}$.

Wir wollen nun zeigen, daß die Berechnung des Skalarprodukts zweier Verbindungsvektoren auf eine Längenberechnung zurückgeführt werden kann (vgl. 29.16).

37.4 Berechnung des Skalarprodukts
Für Punkte P, Q, R und S in einem euklidischen Raum (A, V, α) gilt

$$\begin{aligned} 2\beta(\overrightarrow{PQ}, \overrightarrow{RS}) &= |P, S|^2 + |Q, R|^2 \\ &\quad - |P, R|^2 - |Q, S|^2 . \end{aligned}$$



Beweis: Aus $\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR}$ erhalten wir zunächst

$$\beta(\overrightarrow{PR}, \overrightarrow{PR}) = \beta(\overrightarrow{PQ}, \overrightarrow{PQ}) + 2\beta(\overrightarrow{PQ}, \overrightarrow{QR}) + \beta(\overrightarrow{QR}, \overrightarrow{QR})$$

und bekommen durch Umformen

$$2\beta(\overrightarrow{QP}, \overrightarrow{QR}) = |P, Q|^2 + |Q, R|^2 - |P, R|^2 \quad \text{vgl. (29.16).}$$

Ersetzt man R durch S, so haben wir analog

$$2\beta(\overrightarrow{QP}, \overrightarrow{QS}) = |P, Q|^2 + |Q, S|^2 - |P, S|^2 .$$

Damit folgern wir nun

$$\begin{aligned} 2\beta(\overrightarrow{PQ}, \overrightarrow{RS}) &= 2\beta(\overrightarrow{PQ}, \overrightarrow{RQ} + \overrightarrow{QS}) \\ &= 2\beta(\overrightarrow{PQ}, \overrightarrow{RQ}) + 2\beta(\overrightarrow{PQ}, \overrightarrow{QS}) \\ &= 2\beta(\overrightarrow{QP}, \overrightarrow{QR}) - 2\beta(\overrightarrow{QP}, \overrightarrow{QS}) \\ &= |Q,R|^2 - |Q,S|^2 - |P,R|^2 + |P,S|^2 . \end{aligned}$$

□

Von Abbildungen zwischen euklidischen Räumen erwarten wir, daß sie neben der affinen Struktur auch das Skalarprodukt berücksichtigen:

37.5 Definition
Seien (A_1, V_1, α_1) und (A_2, V_2, α_2) endlich-dimensionale euklidische Räume.
Eine affine Abbildung $f : A_1 \to A_2$ heißt ein *euklidischer Morphismus*, wenn $f_V : V_1 \to V_2$ eine Isometrie ist, d.h. für alle $v, w \in V_1$ gilt

$$\beta_1(v, w) = \beta_2\Big(f_V(v), f_V(w)\Big).$$

Als unmittelbare Folgerung aus den Eigenschaften von Isometrien euklidischer Vektorräume halten wir fest (vgl. 30.1):

37.6 Eigenschaften
Für einen euklidischen Morphismus $f : A_1 \to A_2$ gilt:

(i) $|P,Q| = |f(P), f(Q)|$ *für alle $P, Q \in A_1$ (f ist* längentreu*).*

(ii) $\measuredangle QPR = \measuredangle f(Q)f(P)f(R)$ *für alle $R \neq P \neq Q \in A$ (f ist* winkeltreu*).*

(iii) *f ist* injektiv*.*

(iv) *Jeder euklidische Morphismus $A \to A$ ist ein euklidischer Isomorphismus (*Bewegung*).*

Zum Beispiel ist jede Translation $A \to A$ eine Bewegung.

Bemerkenswert ist, daß jede längentreue Abbildung euklidischer affiner Räume schon ein euklidischer Morphismus ist:

37.7 Kennzeichnung von euklidischen Morphismen
Für eine Abbildung $f : A_1 \to A_2$ euklidischer Räume sind äquivalent:

(a) *f ist ein euklidischer Morphismus;*

(b) *für jedes Paar $P, Q \in A_1$ gilt $|P,Q| = |f(P), f(Q)|$.*

Beweis: (a)⇒(b) haben wir schon in 37.5 festgestellt.

(b)⇒(c) Sei $f : A_1 \to A_2$ eine längentreue Abbildung. Dann gilt nach 37.3 für alle Punkte $P, Q, P', Q' \in A$

$$2\beta_2\Big(\overrightarrow{f(P)f(Q)}, \overrightarrow{f(P')f(Q')}\Big) = 2\beta_1(\overrightarrow{PQ}, \overrightarrow{P'Q'}).$$

Sei nun $\overrightarrow{PQ} = r\overrightarrow{P'Q'}$ für $r \in \mathbb{R}$. Dann folgt aus obiger Gleichung

$$\begin{aligned}
&\|\overrightarrow{f(P)f(Q)} - r\overrightarrow{f(P')f(Q')}\|^2 \\
&\quad = \|\overrightarrow{f(P)f(Q)}\|^2 - 2r\beta\Big(\overrightarrow{f(P)f(Q)}, \overrightarrow{f(P')f(Q')}\Big) + r^2\,\|\overrightarrow{f(P')f(Q')}\|^2 \\
&\quad = \|\overrightarrow{PQ}\|^2 - 2r\beta(\overrightarrow{PQ}, \overrightarrow{P'Q'}) + r^2\,\|\overrightarrow{P'Q'}\|^2 \\
&\quad = \|\overrightarrow{PQ} - r\overrightarrow{P'Q'}\|^2 \;=\; 0,
\end{aligned}$$

also für alle $P, Q, P', Q' \in A$

$$\overrightarrow{f(P)f(Q)} = r\overrightarrow{f(P')f(Q')}.$$

Nach 35.6 ist f damit eine affine Abbildung. Aus dem oben Gezeigten sieht man auch, daß f_V eine Isometrie ist. □

37.8 Definition
Sei (A, V, α) ein n-dimensionaler euklidischer Raum.
Ein Koordinatensystem $(P_0, \ldots, P_n)$ nennt man *kartesisch*, wenn die Koordinatenvektoren $\{\overrightarrow{P_0P_i} \mid i = 1, \ldots, n\}$ eine Orthonormalbasis von V bilden.

Kartesische Koordinatensysteme sind somit durch folgende Bedingungen gekennzeichnet:

$$\begin{aligned}
|P_0, P_i| &= 1 \qquad \text{für } i = 1, \ldots, n \text{ und} \\
\beta(\overrightarrow{P_0P_i}, \overrightarrow{P_0P_j}) &= 0 \qquad \text{für } i \neq j\,.
\end{aligned}$$

In jedem euklidischen Raum gibt es ein kartesisches Koordinatensystem. (Beweis: Ausgehen von einer Orthonormalbasis von V.)

Beispiel
Im euklidischen Standardraum $\mathbb{R}^n$ ist $(0, e_1, \ldots, e_n)$ ein kartesisches Koordinatensystem.

Wir überlegen uns, daß jeder euklidische Raum zu einem solchen Raum euklidisch isomorph ist:

37.9 Koordinatendarstellung euklidischer Räume
Sei (A, V, α) ein n-dimensionaler euklidischer Raum mit einem kartesischen Koordinatensystem $(P_0, \ldots, P_n)$. Dann gilt:

(1) $\varrho_P : A \to V \to \mathbb{R}^n$, $P_i \mapsto \overrightarrow{P_0P_i} \mapsto e_i$, *ist ein euklidischer Isomorphismus.*

(2) *Für alle $P, Q \in A$ gilt $|P, Q| = \|\varrho_P(P) - \varrho_P(Q)\|$.*

(3) *Für alle $P, Q, R \in A$ mit $R \neq P \neq Q$ gilt*

$$\sphericalangle QPR = \sphericalangle\Big(\varrho_P(Q) - \varrho_P(P), \varrho_P(R) - \varrho_P(P)\Big).$$

Beweis: (1) Wir haben schon gesehen, daß ϱ_P ein affiner Isomorphismus ist. Somit bleibt nur noch zu zeigen, daß

$$(\varrho_P)_V : V \to K^n, \quad \overrightarrow{P_0P_i} \mapsto e_i,$$

eine Isometrie ist.

Nach Definition führt $(\varrho_P)_V$ eine Orthonormalbasis in eine Orthonormalbasis über. Wie in §30 gezeigt, sind solche Homomorphismen isometrisch.

(2) und (3) folgen nun aus (1) und den Eigenschaften von euklidischen Isomorphismen (vgl. 37.6). □

37.10 Kartesische Koordinaten und euklidische Isomorphismen
Sei (A, V, α) ein n-dimensionaler euklidischer Raum und $(P_0, \ldots, P_n)$ ein kartesisches Koordinatensystem.

(1) *Ist $(Q_1, \ldots, Q_n)$ ebenfalls ein kartesisches Koordinatensystem von A und (u, T), $u \in \mathbb{R}^n$, $T \in \mathbb{R}^{(n,n)}$, die Koordinatentransformation, die P_i in Q_i überführt, $i = 0, \ldots, n$, dann ist T eine orthogonale Matrix, d.h. $TT^t = E$.*

(2) *Ist T eine orthogonale (n, n)-Matrix und $u \in \mathbb{R}^n$, dann führt die durch (u, T) bestimmte Koordinatentransformation $(P_0, \ldots, P_n)$ wieder in ein kartesisches Koordinatensystem über.*

Beweis: Dies folgt aus der Tatsache, daß die Transformationsmatrizen von Isometrien von Vektorräumen bezüglich Orthonormalbasen gerade die orthogonalen Matrizen sind (vgl. 30.3). □

37.11 Definition
(A_1, V_1, α_1) und (A_2, V_2, α_2) seien isomorphe euklidische Räume.

Zwei Punktmengen $M_1 \subset A_1$ und $M_2 \subset A_2$ heißen *euklidisch äquivalent*, wenn es einen euklidischen Isomorphismus $f : A_1 \to A_2$ mit $f(M_1) = M_2$ gibt.

Bei der *euklidischen* Klassifikation von Quadriken fragt man, welche Quadriken durch einen euklidischen Isomorphismus ineinander übergeführt werden können. Es ist klar, daß dabei eine Ellipse mit verschiedenen Achsen nicht zu einem Kreis werden kann (wie das bei affinen Abbildungen der Fall sein kann). So können etwa in der *euklidischen Normalform* einer Quadrik nicht mehr (wie in 36.6) die Koeffizienten zu ± 1 normiert werden.

Wir wollen uns damit begnügen, die in der reellen Ebene und im reellen Raum auftretenden Fälle aufzulisten. Dabei bezeichnet r wiederum den Rang der zur Quadrik gehörigen Matrix.

Euklidische Klassifikation der Quadriken im $\mathbb{R}^2$

$r=2:$	$x^2+\beta y^2=0,$	$\beta>0,$	Punkt
	$x^2-\beta y^2=0,$	$\beta>0,$	zwei sich schneidende Geraden
$r=1:$	$x^2=0$		Doppelgerade
	$\alpha x^2+\beta y^2=1,$	$\alpha\geq\beta>0,$	Ellipse
	$\alpha x^2-\beta y^2=1,$	$\alpha,\beta>0,$	Hyperbel
	$-\alpha x^2-\beta y^2=1,$	$\alpha\geq\beta>0,$	$\emptyset$
	$\alpha x^2=1,$	$\alpha>0,$	zwei parallele Geraden
	$-\alpha x^2=1,$	$\alpha>0,$	$\emptyset$
	$\alpha x^2-2y=0$	$\alpha>0,$	Parabel.

Euklidische Klassifikation der Quadriken im $\mathbb{R}^3$

$x^2+\beta y^2-\gamma z^2=0,$	$1\geq\beta\geq 0,\ \gamma>0,$	elliptischer Kegel
$x^2-\beta y^2=0,$	$\beta>0,$	zwei sich schneidende Ebenen
$\alpha x^2+\beta y^2-\gamma z^2=1,$	$\alpha\geq\beta>0,\ \gamma>0,$	einschaliger Hyperboloid
$\alpha x^2-\beta y^2-\gamma z^2=1,$	$\alpha>0,\ \beta\geq\gamma>0,$	zweischaliger Hyperboloid
$\alpha x^2+\beta y^2+\gamma z^2=1,$	$\alpha\geq\beta\geq\gamma>0,$	Ellipsoid
$\alpha x^2+\beta y^2=1,$	$\alpha\geq\beta>0,$	elliptischer Zylinder
$\alpha x^2-\beta y^2=1,$	$\alpha,\beta>0,$	hyperbolischer Zylinder
$\alpha x^2=1$		zwei parallele Ebenen
$\alpha x^2+\beta y^2=2z,$	$\alpha\geq\beta>0,$	elliptischer Paraboloid
$\alpha x^2-\beta y^2=2z,$	$\alpha,\beta>0$	hyperbolischer Paraboloid
$\alpha x^2=2z,\ y=c$		Parabelzylinder

38 Axiomatische Geometrie

Die Geometrie war wohl die erste mathematische Disziplin, die einen axiomatischen Aufbau erfuhr. Bekannt ist das Werk *Elemente* des griechischen Mathematikers Euklid, in denen ein erstes Axiomensystem für die Ebene angegeben wurde. Dabei spielte zunächst immer noch etwas die Anschauung mit. Im Laufe der Zeit wurde daran verbessert und ergänzt, und schließlich gab David Hilbert ein logisch einwandfreies Axiomensystem für die euklidische Geometrie.

Man kann verschiedene Axiome wählen und dennoch zur gleichen Geometrie kommen. Eine Möglichkeit ist etwa folgendes

38.1 Axiomensystem einer affinen Inzidenzebene (A, G, ε)
Gegeben seien zwei Mengen A (Punkte) und G (Geraden) und eine Relation ε zwischen A und G (= Teilmenge von $A \times G$).

Gilt $P \,\varepsilon\, g$, so sagt man, P *inzidiert* mit g, oder P *liegt auf* g.

Man sagt zwei Geraden g und h sind *parallel* (und schreibt $g \| h$), wenn $g = h$ oder g und h keinen gemeinsamen Punkt haben.

Inzidenzaxiome

(I) Zu je zwei Punkten $P, Q \in A$ gibt es genau eine Gerade $g \in G$ mit $P \,\varepsilon\, g$ und $Q \,\varepsilon\, g$.

(II) Zu jeder Geraden g und jedem Punkt P gibt es genau eine Gerade h mit $P \,\varepsilon\, h$ und $g \| h$.

(III) Es gibt drei Punkte, die *nicht* auf einer Geraden liegen.

Axiom (II) nennt man das *Parallelenaxiom*. Mit Axiom (III) will man ganz triviale Fälle ausschließen.

Als erste Beobachtungen aus den Axiomen halten wir fest:

38.2 Folgerungen

(1) *Zwei nicht parallele Geraden haben genau einen Schnittpunkt.*

(2) *Die Parallelität ist eine Äquivalenzrelation.*

(3) *Schneidet eine Gerade g eine Gerade h, so schneidet g auch jede Parallele h' zu h.*

(4) *Auf jeder Geraden liegen mindestens zwei Punkte.*

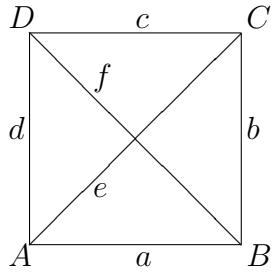
Beweis: (1) Nach Voraussetzung gilt $g \neq h$, und es gilt $P \,\varepsilon\, g$ und $P \,\varepsilon\, h$. Angenommen, es gäbe ein Q mit $Q \,\varepsilon\, g$, $Q \,\varepsilon\, h$ und $Q \neq P$, dann ist nach (I) $g = h$ (eindeutig durch P und Q bestimmt).

(3) Nach Voraussetzung ist g nicht parallel zu h. Wäre $g \| h'$, so folgte mit $h' \| h$ nach (2) auch $g \| h$. Dies ist ein Widerspruch.

(2) und (4) sind ohne weiteres einzusehen. □

Natürlich ist die affine Ebene ein Beispiel, das mit den üblichen Begriffen die obigen Axiome erfüllt. Es gibt aber auch Modelle, die ganz anders aussehen, etwa so:

Minimales Modell einer affinen Inzidenzebene



Mit den angegebenen Axiomen erhalten wir noch keine sehr reichhaltige Geometrie. So können daraus etwa der Satz von Desargues (34.3) und der Satz von Pappus (34.4) nicht abgeleitet werden. Will man solche Zusammenhänge haben, so muß man sie in geeigneter Form als Axiome extra verlangen.

Dabei ist die Forderung nach dem Satz von Desargues gleichbedeutend mit der Forderung nach der Gültigkeit des Scherensatzes (vgl. [9], Theorem 45.A.10). Die Hinzunahme des Satzes von Pappus würde die Gültigkeit des Satzes von Desargues implizieren.

Hat man eine Inzidenzebene gegeben, so interessiert man sich für diejenigen Abbildungen, welche grundlegende Eigenschaften berücksichtigen:

38.3 Definitionen
Sei (A, G, ε) eine Inzidenzebene.

Eine Bijektion $f : A \to A$ nennt man *Kollineation*, wenn sie Geraden auf Geraden abbildet und inzidenzerhaltend ist.

Sei f eine Kollineation.

Eine Gerade g mit $P \,\varepsilon\, g$ und $f(P) \,\varepsilon\, g$ heißt *Spur* von f.

f heißt *Translation*, wenn für alle $g \in G$ gilt $g \| f(g)$, und wenn entweder kein Punkt oder alle Punkte fest bleiben.

f heißt *Streckung*, wenn für alle $g \in G$ gilt $g \| f(g)$ und f mindestens einen Punkt fest läßt.

Wir geben einige Aussagen an, die sich mit diesen Begriffen verbinden:

38.4 Eigenschaften
Sei (A, G, ε) eine Inzidenzebene.

Die Menge der Kollineationen bildet eine Untergruppe der Bijektionen von A.

Kollineationen bilden parallele Geraden in parallele Geraden ab.

Die Spuren einer Translation sind parallel.

Zu zwei Punkten P und Q gibt es höchstens eine Translation $P \mapsto Q$.

Gilt der Satz von Pappus, so ist die Gruppe der Streckungen mit Fixpunkt 0 abelsch.

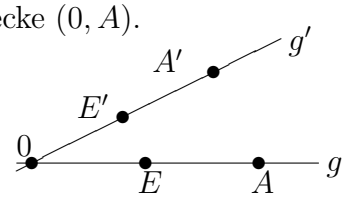
Wir haben schon früher gesehen, daß in einem affinen Raum über einem Körper K jede Gerade bijektiv auf K abgebildet werden kann. Das mag zunächst als eine willkürliche Verbindung zwischen Körpern und Geometrie erschienen sein. Dem ist aber nicht so. Wir können nämlich zeigen, daß sich in jeder Inzidenzebene auf den Geraden algebraische Operationen (Addition und Multiplikation) definieren lassen. Bei genügend starken Voraussetzungen (Satz von Pappus) ergibt sich für die Elemente einer Geraden sogar eine Körperstruktur.

Eine Möglichkeit zur Einführung von Addition und Multiplikation ist die

38.5 Hilbertsche Streckenrechnung *(1899)*

In einer Inzidenzebene wählen wir drei nicht kollineare Punkte $0, E, E' \in A$ und setzen $g = (0, E)$ und $g' = (0, E')$.

Wir bezeichnen $(0, E)$ mit 1, $(0, 0)$ mit 0, und für $A \,\varepsilon\, g$ rechnen wir mit der Strecke $(0, A)$.



Sind zwei Punkte A, B auf der unteren Geraden gegeben, so konstruieren wir die Summe $A + B$ auf folgende Weise:

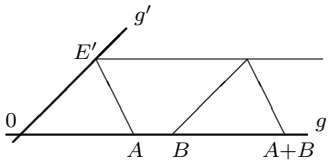
Ziehe eine Parallele zur unteren Geraden durch E';

Ziehe eine Parallele zur oberen Geraden durch B;

durch den Schnittpunkt dieser beiden Geraden ziehe eine Gerade parallel zu $E'A$;

der Schnittpunkt dieser Geraden mit der unteren Gerade ergibt den gewünschten Punkt $A + B$.

Addition:



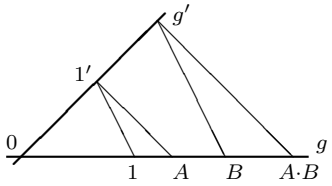
Zur Multiplikation von A und B:

Ziehe Gerade durch B parallel zu EE';

durch Schnittpunkt mit OE' lege Parallele zu $E'A$;

deren Schnittpunkt mit OE ergibt $A \cdot B$.

Multiplikation:



Aus der Konstruktion können wir unmittelbar ablesen:

$$\begin{aligned} A + 0 &= A & A + (-A) &= 0 \\ A \cdot 1 &= A & A \cdot A^{-1} &= 1 \\ A \cdot 0 &= 0 \end{aligned}$$

Durch Hinzunahme weiterer Axiome zu den Inzidenzaxiomen erhalten wir weitere Eigenschaften unserer Operationen, nämlich:

mit Desargues	Assoziativität der Addition
	Assoziativität der Multiplikation
	Kommutativität der Addition
	Distributivität .

In diesem Fall bilden die Elemente einer Geraden einen Schiefkörper.

mit Pappus	Kommutativität der Multiplikation

Damit erhält man also einen Körper.

Durch Hinzunahme von weiteren Forderungen (Anordnungsaxiome, Stetigkeitsaxiome), kann man erreichen, daß die *reellen Zahlen* Koordinatenkörper werden.

Verlangt man als Axiom nicht den allgemeinen Satz von Desargues sondern nur einen Spezialfall davon (den Satz vom vollständigen Vierseit), so spricht man von einer *Moufang-Ebene*. In diesem Fall wird die Multiplikation von beliebigen drei Elementen einer Geraden nicht mehr assoziativ. Allerdings ist jeder von zwei Elementen erzeugte Unterring assoziativ, und jedes Element ungleich Null hat ein Inverses bezüglich der Multiplikation. Man nennt dann K einen *Alternativkörper*. Diese Untersuchungen gehen auf Ruth Moufang [10] zurück.

Wir können (ebene) Geometrien (geometrische Räume) analytisch (als affine Räume, vgl. §33) oder axiomatisch (synthetisch, als Inzidenzebene) einführen. In beiden Fällen erhalten wir ein Koordinatensystem über einem (geeigneten) Körper.

Die interessanten Abbildungen in der affinen Geometrie waren die Affinitäten, in der axiomatischen Geometrie die Kollineationen. Wir wissen bereits, daß Affinitäten auch Kollineationen sind (vgl. 35.3). Auch für Semiaffinitäten läßt sich dies leicht zeigen.

Andererseits wissen wir, daß Kollineationen parallele Geraden in parallele Geraden überführen. Im allgemeinen muß das nicht bedeuten, daß sie auch (Semi-) Affinitäten sind. Es gilt jedoch:

38.6 Hauptsatz der affinen Geometrie
Seien K ein Körper mit mindestens drei Elementen und A eine affine Ebene. Dann gibt es zu jeder Kollineation $f : A \to A$ einen Körperautomorphismus φ von K, so daß f eine Semiaffinität (bzgl. φ) ist.

Für den Körper der reellen (oder rationalen) Zahlen ist die Identität der einzige Automorphismus. Also gilt:

In der reellen affinen Ebene ist jede Kollineation eine Affinität.

Eine genauere Ausführung dieser Zusammenhänge findet man etwa in Artzy [2] oder Klingenberg [9].

Namensverzeichnis

Abel, Niels Hendrik (1802 – 1829), Kristiania, Berlin
Boole, George (1815 – 1864), Waddington, Cork
Cantor, Georg (1848 – 1918), Petersburg, Halle
Cauchy, Augustin Louis (1789 – 1857), Prag, Turin, Paris
Cayley, Arthur (1821 – 1895), London, Cambridge
Cramer, Gabriel (1704 – 1752), Genf
Desargues, Gérard (1591 – 1661), Lyon, Paris
Euklid von Alexandria (ca. 365 – 300 v.Chr.)
Fraenkel, Abraham (1891 – 1965), Marburg, Kiel, Jerusalem
Gauß, Carl Friedrich (1777 – 1855), Braunschweig, Göttingen
Grassmann, Hermann Günther (1809 – 1877), Stettin
Hamilton, William Rowan (1805 – 1865), Dublin
Hermite, Charles (1822 – 1901), Paris
Hilbert, David (1862 – 1943), Königsberg, Göttingen
Jacobi, Carl Gustav (18o4 – 1851), Potsdam, Königsberg, Berlin
Jordan, Camille (1838 – 1922), Lyon
Kronecker, Leopold (1823 – 1891), Liegnitz, Berlin
Krull, Wolfgang (1899 – 1970), Erlangen, Bonn
Lagrange, Joseph Louis (1736 – 1813), Berlin, Paris
Laplace, Pierre Simon (1749 – 1827), Paris
Lie, Marius Sophus (1842 – 1899), Berlin, Kristiania, Leipzig
Moufang, Ruth (1905 – 1977), Frankfurt
Pappus von Alexandria (ca 300 n.Chr)
Russell, Bertrand (1872 – 1970), Cambridge
Sarrus, Pierre-Frédéric (1798 – 1861)
Schmidt, Erhard (1876 – 1959), Berlin
Schwarz, Hermann Amandus (1843 – 1921), Göttingen, Berlin
Sylvester, James Joseph (1814 – 1897), Oxford, London
Vandermond, A.T. (1749 – 1827), Paris

Literaturverzeichnis

[1] Artin, E., *Geometric Algebra*, Interscience Publ. (1966)

[2] Artzy, R., *Linear Geometry*, Addison-Wesley, Reading Mass. (1978)

[3] Blyth, T.S., *Module Theory*, Oxford Science Publ. (1990)

[4] Enderton, H.B., *Elements of Set Theory*, Academic Press, New York (1977)

[5] Fernuniversität Hagen, *Naive Mengenlehre*,
Mathematischer Einführungskurs, FU Hagen (1979)

[6] Halmos, P.R., *Naive Mengenlehre*,
Vandenhoek u. Ruprecht, Göttingen (1968)

[7] Hilbert, D., *Grundlagen der Geometrie*, Leipzig und Berlin (1956)

[8] Koecher, Max, *Lineare Algebra und analytische Geometrie*,
Springer Verlag, Berlin (1985)

[9] Klingenberg, W., *Lineare Algebra und Geometrie*,
Springer Verlag, Berlin (1984)

[10] Moufang, R., *Alternativkörper und der Satz vom vollständigen Vierseit* (D_9),
Abhdlg. Math. Sem. Univ. Hamburg 9, 207-222 (1933)

[11] Oeljeklaus, E., Remmert, R., *Lineare Algebra I*,
Springer Verlag, Berlin (1974)

[12] Wisbauer, R., *Grundlagen der Modul- und Ringtheorie*,
Verlag R. Fischer, München (1988)

Index