Lineare Algebra, WS 97/98, SS 98
Algebra, WS 98/99

Adalbert Kerber

16. Februar 1999

Inhaltsverzeichnis

Kapitel 1

Grundlagen

In diesem Kapitel werden die grundlegenden Begriffe der Aussagenlogik und der Mengenlehre eingeführt, die wir im folgenden benötigen. Danach werden Relationen, Funktionen und Äquivalenzrelationen definiert. Zwei zentrale Sätze über die Faktorisierung von Abbildungen schließen sich an.

1.1 Aussagen

Mathematische Resultate sind in der Regel als *Aussagen* formuliert, also als Sprachgebilde, die je nach Inhalt entweder wahr (w) oder falsch (f) sind. Hier ist ein Beispiel:

> Die Anzahl der positiven Teiler einer positiven natürlichen Zahl ist genau dann ungerade, wenn die Zahl ein Quadrat ist.

Wir werden später zeigen, daß dies eine *wahre* Aussage ist.

Aussagen können durch *Junktoren*

$\neg$ ("nicht ..."), $\wedge$ ("... und ..."), $\vee$ ("... oder ..."), $\Rightarrow$ ("wenn ..., dann ...")

zu neuen Aussagen verbunden werden. Den Gebrauch dieser Junktoren regelt die *Wahrheitswertetafel*:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$
w	w	f	w	w	w
w	f	f	f	w	f
f	w	w	f	w	w
f	f	w	f	f	w

Hier ist besonders die letzte Spalte zu beachten, weil diese Zuordnung von Wahrheitswerten zu $A \Rightarrow B$ vielleicht auf den ersten Blick ungewohnt erscheinen mag — hier wird nämlich festgelegt, daß “aus Falschem alles geschlossen werden kann”, “wenn der Mond aus grünem Käse besteht, dann haben alle Katzen Flügel” wird als wahre Aussage angesehen.

Die vierte Spalte zeigt, daß dieses “oder” nicht im ausschließenden Sinne zu verstehen ist, also nicht mit “entweder ... oder” verwechselt werden darf!

Darüberhinaus werden wir auch noch den folgenden Junktor

$$\Leftrightarrow \quad (\text{“genau dann, wenn \ldots”})$$

verwenden, der als Abkürzung für

$$(A \Rightarrow B) \wedge (B \Rightarrow A)$$

steht, dem also der folgende Wahrheitswerteverlauf zugeordnet wird:

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

.

Aussagen können durch werteverlaufsgleiche, also durch äquivalente Aussagen, ersetzt werden, beispielsweise sind $A \wedge B$ und $B \wedge A$ äquivalent bzw. werteverlaufsgleich. Anhand des Werteverlaufs kann man leicht nachprüfen, daß die umgangssprachliche Formulierung “entweder A oder B” in folgender Weise formalisiert werden kann:

$$(A \vee B) \wedge \neg(A \wedge B),$$

denn

A	B	$A \vee B$	$\neg(A \wedge B)$	$(A \vee B) \wedge \neg(A \wedge B)$
w	w	w	f	f
w	f	w	w	w
f	w	w	w	w
f	f	f	w	f

zeigt, daß diese Aussage genau dann den Wahrheitswert w bekommt, wenn genau einer der beiden Bestandteile A und B diesen Wert hat. Als weitere Anwendung lassen sich Beweisverfahren begründen:

1.1.1 Anwendungen (Beweisverfahren)

i) Man kann leicht nachweisen, daß $A \Rightarrow B$ und $\neg B \Rightarrow \neg A$ werteverlaufsgleich sind, anstelle eines Beweises von $A \Rightarrow B$ kann man demnach genausogut die Implikation $\neg B \Rightarrow \neg A$ (die sogenannte *Kontraposition* von $A \Rightarrow B$) herleiten, manchmal geht das einfacher.

ii) $\neg(A \Rightarrow B)$ und $A \wedge \neg B$ sind werteverlaufsgleich, woraus sich das Verfahren des *indirekten* Beweisens ergibt: Unter der Annahme, daß $\neg B$ wahr sei, leitet man, unter Zuhilfenahme von A, einen Widerspruch her und hat damit $\neg(\neg(A \Rightarrow B))$, also $A \Rightarrow B$ bewiesen.

$\diamond$

Es gibt auch Sprachgebilde, die *Variable* $x, y, \ldots$ enthalten und durch Einsetzen oder Quantifizieren von Werten für die Variablen in Aussagen übergehen. Solche Sprachgebilde werden als *Aussageformen* bezeichnet. Die Quantifizierung von Variablen erfolgt durch Einsetzen — beispielsweise wird $s < r$ durch Einsetzen von $s = 2$ und $r = 3$ zu einer wahren Aussage — oder mit Hilfe der *Quantoren,*

$$\forall \text{ (“für alle”) und } \exists \text{ (“es gibt”)}.$$

Bei Quantoren muß man unbedingt *die Reihenfolge beachten*, das zeigt folgendes Beispiel:

1.1.2
$$\begin{array}{lllll} \forall\, r \in \mathbb{R} & \exists\, s \in \mathbb{R}: & s < r & w \\ \exists\, s \in \mathbb{R} & \forall\, r \in \mathbb{R}: & s < r & f \end{array}$$

1.1.3 Bemerkung Wichtig ist festzuhalten, daß die Negation einer Aussage mit Quantoren dadurch vorgenommen werden kann, daß man *die Quantoren umdreht und die Aussage negiert:*

$$\neg\,(\forall\ \ldots\ \exists\ \ldots\ : A) \iff \exists\ \ldots\ \forall\ \ldots\ :\ \neg A.$$

1.2 Klassen und Mengen

Als undefinierten Grundbegriff verwenden wir den Begriff der *Klasse.* Dieser ist allgemeiner als der Mengenbegriff und wird in der Algebra zur Definition sogenannter *Kategorien* benötigt. (Genauer: Man kann diesen Klassenbegriff als undefinierten Grundbegriff in einem axiomatischen Aufbau der Mengenlehre verwenden und Mengen definieren, so daß das sogenannte Russellsche Paradoxon kein Paradoxon mehr ist!) Klassen bezeichnen wir mit kleinen Buchstaben: $a, b, c, \ldots$

Gegeben sei eine *Enthaltenseinsbeziehung* $\in$ (“... liegt in ...”), und für je zwei Klassen a und b gelte entweder $a \in b$ oder $\neg(a \in b)$ (hierfür schreiben wir auch kurz: $a \notin b$). Mit ihrer Hilfe definieren wir jetzt den Mengenbegriff:

1.2.1 Definition (Menge, Unmenge)

- *Die Klassen x mit der folgenden Eigenschaft heißen* Mengen:

  $$\exists\ a\colon\ x \in a.$$

  Hierfür schreiben wir auch kurz $Mg(x)$.

- *Klassen, die keine Mengen sind, heißen* Unmengen. *Dies sind die Klassen x mit $\neg Mg(x)$.*

Aussageformen der speziellen Gestalt $a \in b$ heißen *einfache* Aussageformen. Mittels Junktoren oder Quantoren erhält man aus diesen die sogenannten *einschlägigen* Aussageformen. Mit diesen Begriffen können wir jetzt einige *Axiome* formulieren, deren Gültigkeit wir im folgenden unterstellen wollen.

1.2.2 Das Klassenbildungsaxiom *Zu jeder einschlägigen Aussageform $A(x)$ gibt es eine Klasse a, die genau die Mengen als Elemente enthält, die dieser Aussageform genügen:*

$$\exists\ a\colon\ \forall\ x:\ (x \in a \iff Mg(x) \wedge A(x)).$$

$\square$

Diese Klasse ist durch $A(x)$ vollständig bestimmt, denn wir fordern auch die Gültigkeit von

1.2.3 Das Umfangsbestimmtheitsaxiom *Klassen sind durch ihre Elemente eindeutig bestimmt:*

$$\forall\ a, b\colon\ (a = b \iff \forall\ x:\ (x \in a \Leftrightarrow x \in b)).$$

$\square$

Die durch $A(x)$ bestimmte Klasse können wir deshalb auch so beschreiben:

$$a := \{x \mid Mg(x) \wedge A(x)\}.$$

1.2.4 Bemerkung Die durch das Umfangsbestimmtheitsaxiom definierte *Gleichheit* von Klassen hat die folgenden bemerkenswerten Eigenschaften:

- Die Gleichheit $=$ ist *reflexiv:* $\forall\ a : a = a$.
- Die Gleichheit $=$ ist *symmetrisch:* $\forall\ a, b : (a = b \Rightarrow b = a)$.
- Die Gleichheit $=$ ist *transitiv:* $\forall\ a, b, c : ((a = b \wedge b = c) \Rightarrow a = c)$.

1.2.5 Beispiele

i) $\emptyset := \{x \mid Mg(x) \wedge \neg Mg(x)\}$, die *leere Klasse,* sie enthält kein einziges Element. Schon im nächsten Axiom werden wir fordern, daß $\emptyset$ eine Menge ist.

ii) $\mathcal{R} := \{x \mid Mg(x) \wedge x \notin x\}$, die *Russell'sche Klasse,* sie ist eine Unmenge (Übungsaufgabe!).

iii) $\mathcal{A} := \{x \mid Mg(x)\}$, die *Allklasse,* sie enthält alle Mengen als Elemente und ist ebenfalls eine Unmenge (s.u.).

$\diamond$

1.2.6 Das Axiom der leeren Menge *Die leere Klasse ist eine Menge:*

$$Mg(\emptyset).$$

$\square$

Gilt für zwei Klassen a und b

$$\forall x: \ (x \in a \Longrightarrow x \in b),$$

dann heißt a *Teilklasse* von b und wir schreiben

$$a \subseteq b.$$

Gilt darüber hinaus noch $a \neq b$, dann schreiben wir

$$a \subset b.$$

(Leider ist die Verwendung dieser beiden Bezeichnungen in der Literatur nicht einheitlich, $\subset$ wird oft auch geschrieben, wenn Gleichheit zugelassen wird!)

1.2.7 Bemerkung Die *Teilklassenbeziehung* hat die folgenden bemerkenswerten Eigenschaften:

- Sie ist *reflexiv:* $\forall\ a\colon a \subseteq a$.
- Sie ist *antisymmetrisch:* $\forall\ a, b\colon ((a \subseteq b \wedge b \subseteq a) \Rightarrow a = b)$.
- Sie ist *transitiv:* $\forall\ a, b, c\colon ((a \subseteq b \wedge b \subseteq c) \Rightarrow a = c)$.
- Die leere Klasse liegt in jeder anderen, und jede Klasse liegt in der Allklasse: $\forall\ a\colon \emptyset \subseteq a \subseteq \mathcal{A}$.

1.2.8 Das Axiom der Teilmenge *Teilklassen von Mengen sind Mengen:*

$$Mg(a) \wedge b \subseteq a \Longrightarrow Mg(b).$$

□

Hiermit kann man zeigen, daß die Allklasse eine Unmenge ist (Übungsaufgabe!). Für die *Vereinigung* und den *Durchschnitt* zweier Klassen verwenden wir die üblichen Symbole

$$a \cup b := \{x \mid Mg(x) \wedge (x \in a \vee x \in b)\}$$

und

$$a \cap b := \{x \mid Mg(x) \wedge x \in a \wedge x \in b\}.$$

1.2.9 Das Axiom der Vereinigungsmenge *Die Vereinigung zweier Mengen ist wieder eine Menge:*

$$Mg(a) \wedge Mg(b) \Longrightarrow Mg(a \cup b).$$

□

Zusammen mit 1.2.8 ergibt sich daraus

1.2.10 Folgerung *Der Schnitt zweier Mengen ist ebenfalls eine Menge:*

$$Mg(a) \wedge Mg(b) \Longrightarrow Mg(a \cap b).$$

□

Für diese *Verknüpfungen* $\cap$ und $\cup$ von Klassen gelten die für Mengen bekannten Gesetze:

1.2.11 Eigenschaften von $\cap$ und $\cup$ *Für alle Klassen a, b, c gilt:*

$$a \cap (b \cap c) = (a \cap b) \cap c, \;\; a \cup (b \cup c) = (a \cup b) \cup c,$$

d. h. $\cap$ und $\cup$ sind assoziativ. *Es gilt auch*

$$a \cap b = b \cap a,$$

$\cap$ und $\cup$ sind demnach kommutativ. *Sie sind auch* idempotent:

$$a \cap a = a, \;\; a \cup a = a,$$

und es gelten die beiden Distributivgesetze:

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c), \;\; a \cup (b \cap c) = (a \cup b) \cap (a \cup c).$$

Diese Gleichungen beweist man mit Hilfe der Aussagenlogik, die Klasse a wird durch die Aussage $x \in a$ ersetzt usw.

Als *Komplementärklasse* bezeichnen wir

$$\overline{a} := \{x \mid Mg(x) \wedge x \notin a\},$$

als *relatives Komplement* bzgl. der Klasse b :

$$a \backslash b := \{x \mid Mg(x) \wedge x \in b \wedge x \notin a\}.$$

Hierfür gelten

1.2.12 Die de Morganschen Regeln *Für Klassen a, b und deren absolute Komplemente gilt:*

$$\overline{a \cup b} = \overline{a} \cap \overline{b}, \;\; \overline{a \cap b} = \overline{a} \cup \overline{b}.$$

Entsprechendes gilt auch für relative Komplemente.

1.2.13 Definition (Vereinigung und Schnitt) *Vereinigung und Schnitt zweier Klassen können noch deutlich verallgemeinert werden: Zu einer einschlägigen Aussage A definieren wir*

$$\bigcup_{A(x)} x := \{y \mid Mg(y) \wedge \exists x(A(x) \wedge y \in x)\}$$

und

$$\bigcap_{A(x)} x := \{y \mid Mg(y) \wedge \forall x(A(x) \wedge y \in x)\}.$$

•

1.2.14 Beispiele Spezialfälle einschlägiger Aussagen sind die einfachen Aussagen wie $x \in a$, wir erhalten also die Beispiele

$$V(a) := \bigcup_{x \in a} x, \;\; S(a) := \bigcap_{x \in a} x.$$

1.2.15 Das große Axiom der Vereinigungsmenge *Einerklassen von Mengen sind Mengen:*

$$Mg(a) \iff Mg(V(a)).$$

1.2.16 Definition (Einerklasse) *Wir definieren zunächst zu jeder Klasse a die sogenannte* Einerklasse

$$\{a\} := \{x \mid Mg(x) \wedge (Mg(a) \Longrightarrow x = a)\}.$$

Ist a eine Menge, dann ist $\{a\}$ eine Klasse, die ein einziges Element enthält, nämlich a. Ist dagegen a keine Menge, dann ist $\{a\}$ die Allklasse. Die Bildung der Einerklasse ist also eigentlich nur für Mengen interessant. Wir verwenden diese Klassenbildung auch zur Einführung der Klassen

$$\{a, b\} := \{a\} \cup \{b\}, \{a, b, c\} := \{a\} \cup \{b\} \cup \{c\}, \ldots$$

•

Als nächstes wird gefordert, daß die Einerklasse einer Menge ebenfalls Menge ist:

1.2.17 Das Axiom der Einermenge: $Mg(a) \Longrightarrow Mg(\{a\})$.

Diese Mengen werden natürlich als *Einermengen* bezeichnet.

1.2.18 Beispiele (die natürlichen Zahlen) Wir haben bereits durch ein Axiom gefordert, daß die leere Klasse eine Menge sei. Diese definieren wir als *die natürliche Zahl* 0:

$$0 := \emptyset,$$

und wir erhalten daraus $Mg(0)$, die natürliche Zahl Null *ist* die leere Menge. Daraus ergibt sich die natürliche Zahl 1 als die vielleicht wichtigste Einermenge:

$$1 := \{\emptyset\} = \{0\}.$$

Nach dem Axiom der Einer ist also auch diese natürliche Zahl eine Menge! Ganz entsprechend wird dann

$$2 := \{0, 1\}$$

gesetzt, die natürliche Zahl n ist

$$n := \{0, 1, \ldots, n - 1\}.$$

Nach dem gleich folgenden Prinzip der vollständigen Induktion sind damit alle natürlichen Zahlen definiert.

$\diamond$

1.2.19 Definition (die Klasse $\mathbb{N}$ der natürlichen Zahlen) *Wir definieren zunächst zu jeder Klasse a deren* Nachfolger:

$$a^+ := a \cup \{a\}.$$

Die Klasse $\mathbb{N}$ der natürlichen Zahlen wird dann als die kleinste Klasse definiert, die die Null und zu jedem ihrer Elemente auch deren Nachfolger enthält:

$$\mathbb{N} := \bigcap_{0 \in a, \forall x (x \in a \Rightarrow x^+ \in a)} a.$$

$\bullet$

Wir fordern schließlich, daß $\mathbb{N}$ eine Menge ist:

1.2.20 Axiom der Menge der natürlichen Zahlen: $Mg(\mathbb{N})$.

Eine unmittelbare Folgerung hieraus ist der nachstehende Satz, der ein weiteres sehr wichtiges Beweisprinzip beschreibt, die sogenannte *vollständige Induktion.* Sie besagt, daß jede Aussage $A(x)$ für *alle* natürlichen Zahlen x wahr ist, wenn man zeigen kann, daß $A(0)$ wahr ist und auch (für beliebiges n) die Implikation $A(n) \Rightarrow A(n^+)$. (Beim Beweis dieser Implikation kann man voraussetzen, daß $A(0), A(1), \ldots, A(n-1)$ wahr sind.)

1.2.21 Satz *(Prinzip der vollständigen Induktion)*

$$[M \subseteq \mathbb{N} \wedge 0 \in M \wedge \forall x (x \in M \Rightarrow x^+ \in M)] \Longrightarrow M = \mathbb{N}$$

Beweis: Wegen der Voraussetzung $0 \in M \wedge \forall x(x \in M \Rightarrow x^+ \in M)]$ gilt $\mathbb{N} \subseteq M$, da auch $M \subseteq \mathbb{N}$ gelten soll, folgt mit der Antisymmetrie von $\subseteq$ die Gleichheit $M = \mathbb{N}$. $\square$

Schließlich definieren wir noch die *Potenzklasse* von a :

$$P(a) := \{x \mid Mg(x) \wedge x \subseteq a\}.$$

Aufgabe 1.2.1

Zeigen Sie, daß sowohl die Russelsche Klasse $\mathcal{R}$ als auch die Allklasse $\mathcal{A}$ Unmengen sind.

Aufgabe 1.2.2

Zeigen Sie ausführlich, daß gilt $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$, und beweisen Sie die de Morgansche Regel $\overline{a \cup b} = \overline{a} \cap \overline{b}$.

1.3 Relationen und Funktionen

Es gibt eine Konstruktion (Übungsaufgabe!) einer Klasse (a, b) mit der Eigenschaft

1.3.1 $$(a, b) = (c, d) \iff a = c \wedge b = d.$$

Diese Klasse (a, b) heißt auch das *(geordnete) Paar* aus a und b. Aus der Konstruktion folgt

1.3.2 $$(Mg(a) \wedge Mg(b)) \iff Mg((a, b)).$$

Entsprechend kann man *Tripel, Quadrupel,* ... *n-tupel* definieren:

$$(a, b, c) := ((a, b), c), \ldots$$

etc. Als *cartesisches Produkt* zweier Klassen definiert man die Klasse aus den geordneten Paaren ihrer Elemente:

$$a \times b := \{(x, y) \mid x \in a \wedge y \in b\}.$$

Hier gilt entsprechend

1.3.3 $$(Mg(a) \wedge Mg(b)) \iff Mg(a \times b).$$

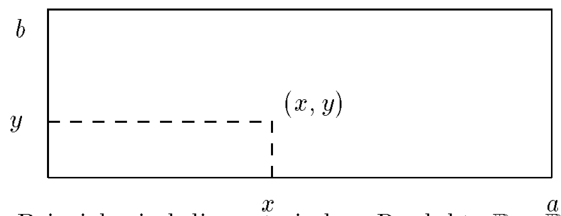
Diese Konstruktion kann natürlich iteriert werden, z.B. sei

$$a \times b \times c := (a \times b) \times c,$$

usw. Wir schreiben auch

$$a^n := \underbrace{a \times \ldots \times a}_{n \text{ Faktoren}}.$$

Das cartesische Produkt $a \times b$ zweier Mengen veranschaulicht man sich gerne als Rechteck, wobei der linke Rand die Menge b, der untere Rand die Menge a bedeutet:



Prominenteste Beispiele sind die cartesischen Produkte $\mathbb{R} \times \mathbb{R}$ und $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$, die reelle Zahlenebene und der dreidimensionale reelle "Anschauungsraum".
Aus schreibtechnischen Gründen erweist es sich später (bei der Verwendung von Vektoren) als zweckmäßig, geordnete Paare, Tripel, ... n-Tupel, nicht nur in *Zeilenschreibweise,* als $(a, b, \ldots)$ einzuführen, sondern auch in *Spaltenschreibweise,* d. h. in der Form

$$\begin{pmatrix} a \\ b \\ \vdots \end{pmatrix}.$$

1.3.4 Definition (Relation, Funktion)

- *Unter einer* (zweistelligen) Relation *verstehen wir eine Teilklasse R der Allklasse, also $R \subseteq \mathcal{A} \times \mathcal{A}$.*
- *Mit einer (zweistelligen) Relation* zwischen *zwei Klassen a und b meinen wir eine Teilklasse des cartesischen Produkts dieser beiden Klassen:*
$$R \subseteq a \times b.$$
Für $(x, y) \in R$ schreiben wir auch kurz: xRy.
- *Als* Definitionsbereich *von $R \subseteq a \times b$ bezeichnen wir*
$$\mathrm{Def}(R) := \{x \in a \mid \exists\ y \in b \colon xRy\}.$$
- *Das* Bild *von $R \subseteq a \times b$ ist*
$$\mathrm{Bild}(R) := \{y \in b \mid \exists\ x \in a \colon xRy\}.$$
- *Die* Umkehrrelation *zu $R \subseteq a \times b$ sei*
$$R^{-1} := \{(y, x) \mid xRy\} \subseteq b \times a.$$
- *Die* Komposition *zweier Relationen $R, S \subseteq \mathcal{A} \times \mathcal{A}$ wird definiert als*
$$S \circ R := \{(x, z) \mid \exists\ y \colon (xRy \wedge ySz)\}.$$
Gelesen wird die Komposition als "erst R, dann S", also von rechts nach links.
- *Eine Relation R heißt* eindeutig, *wenn gilt*
$$(xRy \wedge xRz) \Rightarrow y = z.$$
- *Eindeutige Relationen heißen* Funktionen.

•

Unter einer Funktion f *von a nach b,* also einer eindeutigen Relation $f \subseteq a \times b$, kurz: $f \colon a \to b$, versteht man demnach genau genommen das Tripel
$$(a, \{(x, y) \mid x \in a, y \in b, (x, y) \in f\}, b).$$
Ist $(x, y) \in f$, dann schreibt man auch $f(x)$ anstelle von y und nennt dies den *Wert von f an der Stelle x.* Man läßt jedoch bei der Beschreibung von Funktionen die beiden Klassen a und b gerne weg und betrachtet nur
$$\mathrm{Graph}(f) := \{(x, y) \mid x \in a, y \in b, (x, y) \in f\} \subseteq a \times b,$$
den sogenannten *Graphen* von f. Die Funktion wird also oft einfachheitshalber mit ihrem Graphen identifiziert.
Unter den Funktionen heben wir spezielle Typen durch Attribute hervor:

1.3.5 Definition (injektiv, surjektiv, bijektiv) *Sei f eine Funktion zwischen a und b, wir sagen*

- *f sei* injektiv, *falls $f^{-1}: b \to a$ gilt. Diese Forderung ist äquivalent zu*
$$(f(x) = f(x')) \Rightarrow x = x'.$$
Wir werden dies auch so abkürzen:
$$f: a \rightarrowtail b.$$
- *f sei* surjektiv, *wenn f eine Abbildung* auf *(=sur) b ist, also wenn*
$$\forall y \in b \;\exists\; x \in a : \; f(x) = y.$$
Diese Forderung ist äquivalent zu
$$\mathrm{Bild}(f) = b.$$
Wir werden dies auch so abkürzen:
$$f: a \twoheadrightarrow b.$$
- *f sei* bijektiv, *wenn f sowohl injektiv als auch surjektiv ist. Als Abkürzung verwenden wir*
$$f: a \rightarrowtail\!\!\!\twoheadrightarrow b.$$
- *f heiße* Einbettung, *wenn f injektiv ist und* $\mathrm{Def}(f) = a$ *gilt.*
- *sei die* identische *Abbildung oder* Identität, *wenn $b = a$ gilt und $f(a) = a$.*

•

Mit Hilfe bijektiver Abbildungen werden wir jetzt den Begriff der *Ordnung* einer Klasse einführen.
Man kann zeigen, daß es zu jeder Klasse *a höchstens ein* $n \in \mathbb{N}$ gibt mit $a \rightarrowtail\!\!\!\twoheadrightarrow n$. Dieses n nennt man (gegebenenfalls) die *Kardinalzahl, Mächtigkeit* oder *Ordnung* von a. Man nennt a dann *endlich* und kürzt dies mit $|a| = n$ ab, d. h. man definiert

1.3.6 $$(|a| = n) :\iff (a \rightarrowtail\!\!\!\twoheadrightarrow n).$$

Andernfalls schreibt man $|a| = \infty$ und nennt a *unendliche* Klasse. Im Fall $a \rightarrowtail\!\!\!\twoheadrightarrow \mathbb{N}$ sagt man auch, die Klasse a sei *abzählbar* unendlich.

Funktionen mit geeigneten Eigenschaften sind wichtige *Werkzeuge* in der linearen Algebra und der Algebra. Betrachten wir ein ganz einfaches Beispiel hierfür, nämlich den Beweis der Aussage, die ganz am Anfang des ersten Paragraphen erwähnt wurde:

> Die Anzahl der positiven Teiler einer positiven natürlichen Zahl ist genau dann ungerade, wenn die Zahl ein Quadrat ist.

Betrachten wir dazu eine positive natürliche Zahl n und die Menge $T(n)$ ihrer positiven Teiler:

$$T(n) := \{t \in \mathbb{N} \mid t|n\}.$$

Auf dieser Menge betrachten wir die Abbildung f, die dem Teiler t den Teiler $\frac{n}{t}$ zuordnet. Sie gruppiert die Elemente von $T(n)$ in Teilmengen $\{t, \frac{n}{t}\}$, die offenbar aus 1 oder 2 Elementen Elementen bestehen. Maximal *eine* solche Teilmenge besteht aus nur einem Element, sie existiert genau dann, wenn es einen Teiler t gibt mit $t = \frac{n}{t}$, also genau dann, wenn n ein Quadrat ist: $n = t^2$. Das beweist die Behauptung!

Sehr hilfreich sind die folgenden Charakterisierungen von injektiv, surjektiv und bijektiv:

1.3.7 Satz *Sei f eine Abbildung zwischen einer nicht leeren Klasse a und einer (wegen $a \neq \emptyset$ ebenfalls nicht leeren) Klasse b. Dann gilt:*

- *f ist genau dann i̲njektiv, wenn es eine L̲inksinverse gibt:*

$$\exists\, g: b \to a :\ g \circ f = \mathrm{id}_a.$$

- *f ist genau dann sur̲jektiv, wenn es eine R̲echtsinverse gibt:*

$$\exists\, g: b \to a :\ f \circ g = \mathrm{id}_b.$$

- *f ist genau dann bijektiv, wenn es sowohl eine Rechts- als auch eine Linksinverse gibt. Diese sind gleich.*
- *f ist genau dann i̲njektiv, wenn f l̲inkskürzbar ist, d. h. wenn gilt:*

$$f \circ g = f \circ h \Longrightarrow g = h.$$

- *f ist genau dann sur̲jektiv, wenn f r̲echtskürzbar ist:*

$$g \circ f = h \circ f \Longrightarrow g = h.$$

Beweis: Wir beweisen die erste Behauptung, die restlichen werden analog bewiesen (Übungsaufgabe!):

a) Wir setzen zunächst voraus, f sei injektiv, und wir konstruieren eine Abbildung $g: b \to a$ mit den geforderten Eigenschaften.

   a ist nicht leer, es gibt also mindestens ein Element $x_0 \in a$. Hiermit setzen wir

   1.3.8 $$g: b \to a, y \mapsto \begin{cases} x, & \text{falls } y = f(x), \\ x_0, & \text{falls } y \notin \mathrm{Bild}(f). \end{cases}$$

   Hier ist es wichtig zu beachten, daß dies *noch genau begründet werden muß, denn so wie es da steht ist das keine Definition!* g ist nämlich nur dann eine Funktion, wenn jedem $y \in b$ auf diese Weise *eindeutig* ein Element aus

a zugeordnet wird. Die Zuordnung "x, falls $y = f(x)$" ist aber nicht immer eindeutig, da es ja — ohne weitere Voraussetzungen an f — *mehrere* x geben kann mit $f(x) = y$! Hier ist das glücklicherweise nicht der Fall, denn f ist ja als injektiv vorausgesetzt. Man formuliert dies kurz und knapp so: *Wegen der Injektivität von f ist g durch 1.3.8 wohldefiniert.*

Es bleibt jetzt nur noch $g \circ f = \mathrm{id}_a$ nachzuprüfen, was sehr leicht ist. Aus der Injektivität folgt also tatsächlich die Existenz einer Linksinversen.

b) Existiert umgekehrt eine Linksinverse g, dann ist f injektiv, denn eine Anwendung von g auf beide Seiten einer Gleichung $f(x) = f(x')$ ergibt

$$x = (g \circ f)(x) = (g \circ f)(x') = x',$$

f ist also injektiv.

Damit ist die erste Behauptung bewiesen.

□

Das nächste Axiom ermöglicht den Nachweis, daß Bilder von Mengen ebenfalls Mengen sind:

1.3.9 Das Ersetzungsaxiom: *Ist f eine Funktion mit einer Menge als Definitionsbereich, dann ist auch das Bild eine Menge:* $Mg(\mathrm{Bild}(f))$.

1.3.10 Anwendungen ($\mathbb{N}^n$ und $\mathbb{Q}$ sind Mengen)

- Das cartesische Produkt zweier Mengen ist ebenfalls eine Menge (Übungsaufgabe!) Daraus folgt unter anderem die Tatsache, daß die Menge der n-Tupel natürlicher Zahlen eine Menge ist:

$$Mg(\mathbb{N}^n)$$

- Wir können $\mathbb{N}$ auf die Klasse $\mathbb{Q}_{\geq 0}$ der nicht negativen rationalen Zahlen abbilden, etwa mit Hilfe der im folgenden skizzierten Abbildung (die Pfeile deuten die Reihenfolge an, in der die rationalen Zahlen $\frac{i}{k}$ den natürlichen Zahlen $0, 1, 2, \ldots$ zugeordnet werden sollen, die ersten 10 Funktionswerte sind angegeben, daß einige Werte mehrfach auftreten macht nichts, es ist wegen der besseren Überschaubarkeit in Kauf genommen):

$$\begin{array}{ccccccc}
0 & \to & 1 & & 2 & \to & 3 \\
 & \swarrow & & \nearrow & & \swarrow & \\
\frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & \\
\downarrow & \nearrow & & \swarrow & & & \\
\frac{1}{3} & & \frac{2}{3} & & & & \\
 & \swarrow & & & & & \\
\frac{1}{4} & & & & & & \\
\vdots & & & & & &
\end{array}$$

Mit dem Ersetzungsaxiom ergibt sich hieraus $Mg(\mathbb{Q}_{\geq 0})$, ganz analog folgt $Mg(\mathbb{Q}_{\leq 0})$, insgesamt also mit dem Axiom der Vereinigungsmenge:

$$Mg(\mathbb{Q}).$$

◇

Die n-te cartesische Potenz a^n einer Klasse a kann als Menge von Abbildungen aufgefaßt werden: $(x_0, \ldots, x_{n-1}) \in a^n$ wird dazu einfach als die folgende Abbildung interpretiert:

$$f\colon n \to b, i \mapsto x_i.$$

(Bei dieser Schreibweise wird die Angabe $f\colon a \to b$ noch durch die Definition der Abbildung mit Hilfe der Angabe der Bilder ergänzt!) Die Bezeichnung a^n stimmt mit folgender Notation für die Klasse aller Abbildungen von X nach Y überein:

$$Y^X := \{f \mid Mg(f) \wedge f\colon X \to Y\}.$$

(Abbildungen $(a, \mathrm{Graph}(f), b)$ sind genau dann Mengen, wenn a und b Mengen sind.) Diese Bezeichnung Y^X werden wir meistens verwenden.
Eine Verallgemeinerung des oben eingeführten cartesischen Produkts zweier Klassen und der cartesischen Potenz einer Klasse ist, für eine Menge I (eine sogenannte *Indexmenge*), und die Klassen a_i, für alle $i \in I$, die Bildung von

$$\times_{i\in I} a_i := \{f\colon I \to \bigcup_{i\in I} a_i \mid f(i) \in a_i\}$$

Schließlich führen wir noch ein Axiom an, das die *Existenz* gewisser Funktionen, sogenannter *Auswahlen* fordert:

1.3.11 Das Auswahlaxiom: *Ist I eine nicht leere Indexmenge, $I \neq \emptyset$, und ist zu jedem Index $i \in I$ eine Klasse $a_i \neq \emptyset$ vorgegeben (kurz: eine* Folge, Familie *oder auch* indiziertes Klassensystem *$(a_i)_{i\in I}$ nicht leerer Klassen), dann gilt*

$$\times_{i\in I} a_i \neq \emptyset,$$

d. h. es gibt Auswahlen!

Dieses Axiom wird selten herangezogen werden, es ist aber wichtig und kann auch scheinbar ganz anders formuliert werden. Die prominentesten, zum Auswahlaxiom äquivalenten, Forderungen sind die Gültigkeit des *Wohlordnungssatzes* und das *Zornsche Lemma.* Wir werden es beim Beweis des Satzes verwenden, daß alle Vektorräume Basen besitzen.

Nachdem nun die axiomatische Grundlage gelegt ist, können wir uns — bis zur Einführung der Kategorien — das Leben dadurch erleichtern, daß wir bis dorthin, also bis auf weiteres, voraussetzen, alle betrachteten Klassen seien Mengen. Um dies deutlich zu machen, verwenden wir für die betrachteten Mengen Großbuchstaben wie $L, M, N, \ldots$ Darüberhinaus unterstellen wir, daß $f\colon M \to N$ stets bedeute, M sei der Definitionsbereich von f!

1.3.12 Definition (Äquivalenzrelation) *Sei $R \subseteq M \times M$, also eine "Relation auf M". Sie heißt genau dann* Äquivalenzrelation, *wenn sie reflexiv, symmetrisch und transitiv ist.*

Die zu einem Element m von M äquivalenten Elemente bilden die sogenannte Äquivalenzklasse von m :

$$[m]_R := \{m' \mid mRm'\}.$$

•

Die wichtigste Eigenschaft von Äquivalenzrelationen ist die folgende:

1.3.13 Satz *Jede Äquivalenzrelation R auf einer Menge M (das steht kurz für $R \subseteq M \times M$) ergibt eine* Partition *von M, das heißt eine vollständige Zerlegung von M in disjunkte Teilmengen:*

$$\bigcup_{m \in M} [m]_R = M.$$

(Beachten Sie, daß nirgends behauptet wird, zu verschiedenen Elementen $m \neq m'$ würden auch verschiedene Klassen $[m]_R$ und $[m']_R$ gehören!)

Beweis: Die Reflexivitätsbedingung garantiert, daß jedes m in mindestens einer Äquivalenzklasse liegt, z. B. in $[m]_R$. Mit Transitivität und Symmetrie folgt, daß verschiedene Klassen disjunkt sind:

$$[m]_R \neq [m']_R \Longrightarrow [m]_R \cap [m']_R = \emptyset.$$

Wir beweisen die Kontraposition hiervon: Sei $x \in [m]_R \cap [m']_R$. Es gilt hierfür nach der Definition der Äquivalenzklasse: $mRx \wedge m'Rx$. Mit der Symmetrieeigenschaft von R folgt $mRx \wedge xRm'$, was mit der Transitivität die Äquivalenz von m und m' liefert. Hieraus ergibt nochmalige Anwendung der Transitivität die beiden Inklusionen $[m]_R \subseteq [m']_R$ und $[m]_R \supseteq [m']_R$, woraus die Gleichheit der beiden Äquivalenzklassen folgt.

□

1.3.14 Beispiele

- Die Gleichheit, also $R :==$, ist eine Äquivalenzrelation (auf jeder Menge M).
- Jede vollständige Zerlegung in disjunkte Teilmengen $M = \dot{\bigcup}_{i \in I} M_i$ liefert eine Äquivalenzrelation R :

  $$mRm' :\Longleftrightarrow \exists\ i\colon\ m, m' \in M_i.$$

- Jede Funktion $f\colon M \to N$ (bzw. kurz: $f \in N^M$) induziert auf ihrem Definitionsbereich M die folgende Relation R_f :

  $$mR_f m' :\Longleftrightarrow f(m) = f(m').$$

  Diese Äquivalenzrelation heißt auch die von f auf M *induzierte* Äquivalenzrelation.

◇

Zu einer gegebenen Äquivalenzrelation R auf M bezeichnen wir mit M_R die Menge ihrer Äquivalenzklassen:

$$M_R := \{[m]_R \mid m \in M\}.$$

Viele Strukturen der Mathematik können als solche Äquivalenzklassenmengen definiert werden, d. h. durch Wahl geeigneter Mengen M und geeigneter Äquivalenzrelationen auf diesen. Besonders prominente Beispiele sind die folgenden

1.3.15 Beispiele (Konstruktion von $\mathbb{Z}$ und $\mathbb{Q}$)

- Die Konstruktion von $\mathbb{Z}$ aus $\mathbb{N}$: Ihr liegt die Idee zugrunde, daß man jede ganze Zahl als Differenz zweier natürlicher Zahlen schreiben kann: $z = n_1 - n_2$, wobei allerdings zu beachten ist, daß diese Darstellung nicht eindeutig ist, man faßt deshalb die Paare (n_1, n_2), mit deren Hilfe man z darstellen kann *zu einer Äquivalenzklasse zusammen:*

  Wir betrachten $\mathbb{N}$ und setzen die Addition natürlicher Zahlen als bekannt voraus (die man auch als mengentheoretische Operation formulieren kann: $n + m := (n + (m - 1))^+$). Auf $\mathbb{N} \times \mathbb{N}$ betrachten wir die folgende Äquivalenzrelation:

  $$(n_1, n_2)R(n_3, n_4) :\Longleftrightarrow n_1 + n_4 = n_2 + n_3.$$

  Wir setzen dann

  $$\mathbb{Z} := (\mathbb{N} \times \mathbb{N})_R,$$

  und anstelle von $[(n_1, n_2)]_R$ schreiben wir $n_1 - n_2$, falls $n_1 > n_2$ bzw. $-(n_2 - n_1)$, falls $n_1 < n_2$, und 0, *sonst*.

- Die Konstruktion von $\mathbb{Q}$ aus $\mathbb{Z}$: Hier liegt — dem vorigen Punkt ganz entsprechend — die Idee zugrunde, daß jede rationale Zahl ja ein Quotient $r = \frac{z_1}{z_2}$ aus ganzen Zahlen ist, wobei der Nenner natürlich von Null verschieden sein muß. Diese Darstellung ist ebenfalls nicht eindeutig, so daß die entsprechende Äquivalenzrelation betrachtet wird:

  Gegeben sei $\mathbb{Z}$ und die Multiplikation ganzer Zahlen (auch diese läßt sich mengentheoretisch einführen, ähnlich wie die Addition!). Die Menge der von Null verschiedenen ganzen Zahlen sei mit $\mathbb{Z}^*$ bezeichnet, auf $\mathbb{Z} \times \mathbb{Z}^*$ wird die folgende Äquivalenzrelation betrachtet:

  $$(z_1, z_2)R(z_3, z_4) :\Longleftrightarrow z_1 \cdot z_4 = z_2 \cdot z_3.$$

  Wir setzen

  $$\mathbb{Q} := (\mathbb{Z} \times \mathbb{Z}^*)_R$$

  und bezeichnen die Äquivalenzklassen wie üblich in der folgenden Weise:

  $$\frac{z_1}{z_2} := [(z_1, z_2)]_R.$$

Für diese *Brüche* gilt infolgedessen:

$$\frac{z_1}{z_2} = \frac{z_3}{z_4} \iff z_1 \cdot z_4 = z_2 \cdot z_3.$$

- Viele Strukturen in Mathematik und Naturwissenschaften kann man bequem mit Hilfe von Äquivalenzrelationen auf Mengen von Abbildungen definieren. Ein interessantes Beispiel bilden die (unnumerierten) Graphen, die man als Äquivalenzklassen numerierter Graphen definiert, wie wir gleich sehen werden. Numerierte Graphen kann man nicht vermeiden, denn beispielsweise sind Computer "nur" zur Verarbeitung von Mengen in der Lage, deren Elemente Nummern tragen.

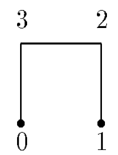
  Die *numerierten* Graphen mit n Punkten kann man wie folgt als Menge Y^X von Abbildungen definieren. Man numeriert zunächst die Punkte von 0 bis $n-1$ und identifiziert die Menge $n = \{0, \ldots, n-1\}$ der Nummern der Punkte mit der Punktemenge selbst. Dann betrachtet man die Menge der *Punktepaare,*

$$\binom{n}{2} := \{\{i,j\} \mid i,j \in n, i \neq j\}.$$

  Für X nimmt man diese Menge, für Y die Menge $2 = \{0,1\}$ und betrachtet die Menge von Abbildungen

$$Y^X := 2^{\binom{n}{2}}$$

  von der Menge der Punktepaare in die Menge $\{0,1\}$. Ist γ in dieser Menge, dann interpretiert man den Wert $\gamma(\{i,j\}) = 1$ als "die Punkte i und j sind durch eine Kante verbunden". Hier ist ein Beispiel:

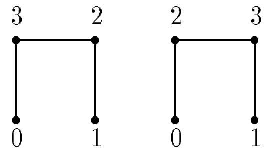


  Dieser Graph wird auf diese Weise mit der folgenden Abbildung identifiziert:

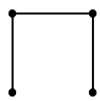
$$f\colon \binom{4}{2} \to \{0,1\}, \begin{cases} \{0,1\} \mapsto 0, \\ \{0,2\} \mapsto 0, \\ \{0,3\} \mapsto 1, \\ \{1,2\} \mapsto 1, \\ \{1,3\} \mapsto 0, \\ \{2,3\} \mapsto 1. \end{cases}$$

  Demnach kann die Menge der numerierten Graphen mit n Punkten mit dieser Menge von Abbildungen $\gamma\colon \binom{n}{2} \to \{0,1\}$ gleichgesetzt werden.

  Die *unnumerierten* Graphen mit n Punkten definiert man — weil es bei ihnen auf die Nummern der Punkte nicht mehr ankommt — als die Äquivalenzklassen der folgenden Relation R : Zwei numerierte Graphen heißen *äquivalent* oder *isomorph,* wie man auch sagt, wenn sie durch Umnumerierung auseinander hervorgehen. Hier sind zwei äquivalente Graphen mit 4 Punkten:



Da beliebiges Umnumerieren zugelassen wird, kann man eine solche Äquivalenzklasse numerierter Graphen durch einen Graphen repräsentieren, den man durch Weglassen der Nummern aus irgendeinem Element der Klasse bekommt. Hier ist der unserem Beispiel auf diese Weise entsprechende *unnumerierte Graph:*



Solche unnumerierten Graphen dienen als *Wechselwirkungsmodelle* in Chemie, Physik und Wirtschaftswissenschaften, zum Beispiel zur Beschreibung von Molekülen (ein bekanntes Beispiel ist der Benzolring) oder von Schaltkreisen oder von Betriebsabläufen (Netzplantechnik) usw.

◇

Es ist bereits erwähnt worden, daß Abbildungen ein wichtiges Werkzeug zur Untersuchung von algebraischen Strukturen sind. Oft sind diese Funktionen recht kompliziert, und man sucht nach Wegen, ihre Anwendung in kleine, möglichst übersichtliche Schritte zu zerlegen. Beispielsweise kann man oft eine gegebene Funktion $f: L \to N$ in zwei Faktoren $g: L \to M$ und $h: M \to N$ zerlegen, sie *faktorisieren:* $f = h \circ g$. Man drückt das gerne in Form eines *Diagramms* aus:

$$\begin{tikzcd} L \arrow[r, "g"] \arrow[dr, "f"'] & M \arrow[d, "h"] \\ & N \end{tikzcd}$$

Man sagt, dieses Diagramm sei *kommutativ,* wenn $f = h \circ g$ gilt. Betrachten wir als erstes Beispiel die natürliche Faktorisierung einer Abbildung mit Hilfe der von ihr induzierten Äquivalenzrelation.
Die Abbildung

$$\nu_R: M \to M_R, m \mapsto [m]_R,$$

die M auf die Äquivalenzklassenmenge M_R abbildet heißt auch die *natürliche* Abbildung auf die Äquivalenzklassenmenge, weil sie sich sozusagen "von selbst" ergibt. Für die von einer Abbildung $f\colon M \twoheadrightarrow N$ induzierte Äquivalenzrelation R_f ist folgendes Diagramm mit der natürlichen Abbildung kommutativ:

$$\begin{tikzcd} M \arrow[r, "\nu_{R_f}"] \arrow[dr, "f"'] & M_{R_f} \arrow[d, "{[m]_{R_f} \mapsto f(m)}"] \\ & N \end{tikzcd}$$

Interessant und oft hilfreich ist die Bemerkung, daß sich *jede Abbildung $f\colon M \to N$ mit Hilfe der natürlichen Abbildung ν_{R_f} zerlegen läßt:*

1.3.16 $$\begin{array}{ccccccc} M & \overset{\nu_{R_f}}{\twoheadrightarrow} & M_{R_f} & \rightarrowtail\!\!\!\twoheadrightarrow & \mathrm{Bild}(f) & \hookrightarrow & N \\ m & \longmapsto & [m]_{R_f} & \longmapsto & f(m) & \longmapsto & f(m) \end{array}.$$

(Dabei wird mit $\hookrightarrow$ wie meist üblich die sogenannte *Einbettung* bezeichnet, das ist, für $M \subseteq N$, die Abbildung $m \mapsto m$.)

Wichtiger ist aber

1.3.17 Der Abbildungssatz *Ist L eine nicht leere Menge, f eine Abbildung von L nach N, g eine Abbildung von L nach M, dann ist f genau dann* über *g* faktorisierbar, *d. h. es gibt ein $h\colon M \to N$ mit $f = h \circ g$, bzw. folgendes Diagramm ist kommutativ:*

$$\begin{tikzcd} L \arrow[r, "g"] \arrow[dr, "f"'] & M \arrow[d, "h"] \\ & N \end{tikzcd},$$

wenn die induzierten Äquivalenzrelationen so ineinander liegen:

$$R_g \subseteq R_f.$$

Ist g surjektiv, dann ist h eindeutig bestimmt. Ist f surjektiv, dann ist auch h surjektiv, und ist g surjektiv sowie $R_f = R_g$, dann ist h injektiv.

Beweis: Beweisen wir zunächst die Faktorisierbarkeit von g. Wegen $L \neq \emptyset$ sind M und N ebenfalls nicht leer. Wir können deshalb ein $n_0 \in N$ auswählen und damit die folgende Zuordnung treffen:

$$h\colon M \to N\colon m \mapsto \begin{cases} n_0, & \text{falls } m \notin \mathrm{Bild}(g) \\ f(l), & \text{falls } m = g(l). \end{cases}.$$

Sie definiert eine Funktion, denn die Voraussetzung $R_g \subseteq R_f$ liefert die Implikation

$$g(l) = g(l') \Longrightarrow m := f(l) = f(l').$$

Für diese Abbildung gilt ganz offensichtlich $f = h \circ g$, sie faktorisiert also f. Sei umgekehrt $f = h \circ g$. Die Gleichheit $g(l) = g(l')$ impliziert dann

$$f(l) = h(g(l)) = h(g(l')) = f(l'),$$

also die Implikation $R_f \subseteq R_g$.
Die übrigen Behauptungen sind ebenso leicht nachzuprüfen: Daß die Surjektivität von g die Abbildung h festlegt, ist klar (vgl. die Definition von h), ebenso, daß die Surjektivität von f die von h impliziert. Falls g surjektiv ist und $R_f = R_g$, dann ergibt sich die Injektivität von h wie folgt: Sei $h(m) = h(m')$. Wegen der Surjektivität von g gibt es l, l' mit $m = g(l)$ und $m' = g(l')$. Die Faktorisierung $f = h \circ g$ liefert $f(l) = f(l')$, was wegen der Gleichheit der Relationen die Identität $g(l) = g(l')$, also auch $m = m'$ impliziert, wie behauptet.

$\square$

Ein Beispiel hierfür ist gegeben, wenn R eine Äquivalenzrelation ist mit $R_f \subseteq R$, denn dann läßt sich f über ν_R faktorisieren:

$$\begin{tikzcd} L \arrow[r, "\nu_R"] \arrow[dr, "f"'] & M \arrow[d, "h"] \\ & N \end{tikzcd} .$$

1.3.18 Der Satz über induzierte Abbildungen *Ist f eine Abbildung von M nach N und sind R und S zwei Äquivalenzrelationen auf M bzw. N, die mit f wie folgt verträglich sind:*

$$mRm' \Longrightarrow f(m)Sf(m'),$$

dann gibt es genau eine Abbildung $h: M_R \to N_S$, die das folgende Diagramm kommutativ ergänzt:

$$\begin{tikzcd} M \arrow[r, "f"] \arrow[d, "\nu_R"'] & N \arrow[d, "\nu_S"] \\ M_R \arrow[r] & N_S \end{tikzcd} .$$

Beweis: Übungsaufgabe!

Weitere Abbildungssätze werden wir bei Bedarf kennenlernen. Damit sind die grundlegenden Definitionen, Axiome und Resultate der Mengenlehre zusammengestellt, so daß wir uns jetzt den algebraischen Strukturen zuwenden können. Dabei werden wir über die bereits erwähnten Resultate hinaus die Grundeigenschaften von $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ als bekannt voraussetzen. $\mathbb{N}, \mathbb{Z}$ und $\mathbb{Q}$ wurden definiert, Konstruktionen von $\mathbb{R}$ und $\mathbb{C}$ werden voraussichtlich im dritten Semester in der Vorlesung “Algebra I” angegeben. Über die bereits erwähnten

Resultate hinaus gehen isnbesondere die Definitionen von Addition und Multiplikation auf $\mathbb{N}$. Es sei hier nur erwähnt, daß man dazu Tatsachen über rekursive Definierbarkeit benötigt. Ein erstes Resultat in dieser Richtung ist, daß für jede Menge W ("Wertemenge"), und alle $w \in W$ ("Anfangswert") die Nachfolgerfunktion

$$-^+ : \mathbb{N} \to \mathbb{N}, n \mapsto n \cup \{n\} = n^+$$

und jedes $\rho: W \to W$ ("Rekursionsfunktion") durch die Forderung, folgendes Diagramm sei kommutativ, eine Funktion $f: \mathbb{N} \to W$ eindeutig bestimmt wird:

$$\begin{tikzcd}
\{0\} \arrow[r, "0 \mapsto 0"] \arrow[d, "f"'] & \mathbb{N} \arrow[r, "-^+"] \arrow[d, "f"] & \mathbb{N} \arrow[d, "f"] \\
\{w\} \arrow[r, "w \mapsto w"] & W \arrow[r, "\rho"] & W
\end{tikzcd}$$

Die Funktion f heißt *die durch ρ rekursiv definierte Funktion bzw. Folge.* Für sie gilt $f(0) = w$ und, für alle $n \in \mathbb{N}$, die *Rekursionsformel*

$$f(n+1) = \rho(f(n)).$$

Hiermit kann man rekursiv die Addition und die Multiplikation natürlicher Zahlen definieren.

Kapitel 2

Algebraische Strukturen

In diesem Kapitel werden die grundlegenden algebraischen Strukturen eingeführt. Als Aufhänger für ihre Verwendung dient die Behandlung linearer Gleichungssyteme

$$\begin{array}{ccccccccc}
a_{00}x_0 & + \ldots + & a_{0k}x_k & + \ldots + & a_{0,n-1}x_n & = & b_0 & & \\
\vdots & & & & & & & & \\
a_{i0}x_0 & + \ldots + & a_{ik}x_k & + \ldots + & a_{i,n-1}x_n & = & b_i & . & \\
\vdots & & & & & & & & \\
a_{m-1,0}x_0 & + \ldots + & a_{m-1,k}x_k & + \ldots + & a_{m-1,,n-1}x_n & = & b_{m-1}, & &
\end{array}$$

Unter Benutzung des Summenzeichens $\sum$ können wir das auch kurz so schreiben:

2.0.19
$$\sum_{k\in n} a_{ik}x_k = b_i,\ i \in m.$$

Dabei sind die $a_{ik}, b_i \in \mathbb{R}$ gegeben, und gesucht sind *Lösungen*

$$x = \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix}, x_i \in \mathbb{R},$$

die 2.0.19 erfüllen. Um dieses Problem anpacken zu können, stellen wir die folgende Hilfsüberlegung an. Gleichzeitig mit 2.0.19 betrachten wir das zugehörige *homogene Gleichungssystem* (2.0.19 heißt *inhomogenes* System, da möglicherweise ein $b_i \neq 0$) :

2.0.20
$$\sum_{k} a_{ik}x_k = 0,\ i \in m.$$

Die zugehörigen *Lösungsgesamtheiten*, die wir ja suchen, bezeichnen wir so:

$$L_I := \{x \mid x \text{ erfüllt } 2.0.19\},\quad L_H := \{x \mid x \text{ erfüllt } 2.0.20\}.$$

Ganz wichtig sind jetzt die folgenden Beobachtungen: Definieren wir

$$x+y=\begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} + \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix} := \begin{pmatrix} x_0+y_0 \\ \vdots \\ x_{n-1}+y_{n-1} \end{pmatrix}, rx := \begin{pmatrix} rx_0 \\ \vdots \\ rx_{n-1} \end{pmatrix},$$

dann gilt (nachrechnen!):

- $x, y \in L_H, r \in \mathbb{R} \Longrightarrow x+y, rx \in L_H$
- $x, y \in L_I \Longrightarrow x-y \in L_H$

Hieraus ergibt sich die

2.0.21 Folgerung *Ist $x \in L_I$, dann ist jede Lösung z von L_I von der Form $z = x + y$, mit einem geeigneten $y \in L_H$, kurz: $L_I = x + L_H$.*

Zur Bestimmung von L_I genügt also die Ermittlung *irgendeiner speziellen* Lösung des inhomogenen Systems und die Angabe der *Lösungsgesamtheit des homogenen Systems.*

2.1 Gruppen

Wir führen jetzt eine Hierarchie von algebraischen Strukturen ein, die für die weiteren Überlegungen sehr wichtig sind. Dabei betrachten wir zunächst diejenigen, die aus einer Menge zusammen mit lediglich einer einzigen Verknüpfung bestehen:

2.1.1 Definition (Gruppoide, Halbgruppen, Monoide, Gruppen)

i) Unter einem Gruppoid *verstehen wir eine Menge G zusammen mit einer* Verknüpfung *"$*$", d. h. mit einer Abbildung*

$$*\colon G \times G \to G\colon (g, g') \mapsto g * g'.$$

ii) Sei $(G, *)$ ein Gruppoid. Ein Element $e \in G$ heißt

$$\begin{array}{rcl} \text{linksneutral} & :\Longleftrightarrow & \forall\, g \in G\colon e * g = g, \\ \text{rechtsneutral} & :\Longleftrightarrow & \forall\, g \in G\colon g * e = g, \\ \text{neutral} & :\Longleftrightarrow & \forall\, g \in G\colon g * e = e * g = g. \end{array}$$

iii) Ist $(G, *)$ ein Gruppoid mit neutralem Element e, $g \in G$, dann heißt $g' \in G$

$$\begin{array}{rcl} \text{linksinvers}\ \textit{zu}\ g & :\Longleftrightarrow & g' * g = e, \\ \text{rechtsinvers}\ \textit{zu}\ g & :\Longleftrightarrow & g * g' = e, \\ \text{invers}\ \textit{zu}\ g & :\Longleftrightarrow & g' * g = g * g' = e. \end{array}$$

iv) Halbgruppen *heißen die Gruppoide mit* assoziativer *Verknüpfung:*

$$\forall\ g, g', g'' \in G: \quad g * (g' * g'') = (g * g') * g''.$$

v) Monoid *heißt jede Halbgruppe mit neutralem Element.*

vi) Gruppe *nennt man ein Monoid mit Inversen:*

$$\forall\ g \in G \quad \exists g' \in G: \quad g * g' = g' * g = e.$$

vii) Abelsche *Gruppen sind die Gruppen mit kommutativer Verknüpfung:*

$$\forall\ g, g' \in G: \quad g * g' = g' * g.$$

2.1.2 Beispiele

- $(\mathbb{N}, +)$ ist ein Monoid, aber keine Gruppe, $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +)$ sind abelsche Gruppen, ebenso $(\mathbb{Q}^*, \cdot), (\mathbb{R}^*, \cdot), (\mathbb{C}^*, \cdot)$, wenn $\mathbb{Q}^* := \mathbb{Q}\backslash\{0\}$ etc.

- Für Mengen X ist $(X^X, \circ)$ ein Monoid.

- Die wichtigste Klasse von Beispielen ist die folgende. Sei X eine nicht leere Menge, $(G, *)$ Gruppoid, Halbgruppe, Monoid, oder Gruppe. Dann wird G^X zu Gruppoid, Halbgruppe, Monoid, Gruppe durch Einführung der *punktweisen* Verknüpfung $*'$ wie folgt:

  $$\forall\ f, f' \in G^X: \quad (f *' f')(x) := f(x) * f'(x).$$

  Das Standardbeispiel ist hier $X := n = \{0, \ldots, n-1\}, G := \mathbb{R}$, also die Menge
  $$\mathbb{R}^n = \{(x_0, \ldots, x_{n-1}) \mid x_i \in \mathbb{R}\}$$
  bei Zeilenschreibweise, bzw.

  $$\mathbb{R}^n = \{ \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} \mid x_i \in \mathbb{R}\}$$

  bei Spaltenschreibweise, zusammen mit der punktweisen Addition

  $$(x_0, \ldots, x_{n-1}) + (y_0, \ldots, y_{n-1}) = (x_0 + y_0, \ldots, x_{n-1} + y_{n-1}),$$

  bzw.
  $$\begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} + \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} x_0 + y_0 \\ \vdots \\ x_{n-1} + y_{n-1} \end{pmatrix}.$$

- Die *symmetrische Gruppe* auf einer Menge X wird definiert als die Menge

  $$S_X := \{\pi : X \rightarrowtail\!\!\twoheadrightarrow X\},$$

  zusammen mit der Hintereinanderausführung als Verknüpfung:

  $$(S_X, \circ).$$

  Z. B. ist — in naheliegender Schreibweise, bei der die Bilder unter die Urbilder geschrieben werden —

  $$S_3 = \{\begin{pmatrix} 012 \\ 012 \end{pmatrix}, \begin{pmatrix} 012 \\ 102 \end{pmatrix}, \begin{pmatrix} 012 \\ 210 \end{pmatrix}, \begin{pmatrix} 012 \\ 021 \end{pmatrix}, \begin{pmatrix} 012 \\ 120 \end{pmatrix}, \begin{pmatrix} 012 \\ 201 \end{pmatrix}\}.$$

  S_X ist leicht als Gruppe nachzuweisen, falls $|X| > 2$ ist diese Gruppe *nicht abelsch,* denn es gilt z. B.

  $$\begin{pmatrix} 012 \\ 210 \end{pmatrix} \circ \begin{pmatrix} 012 \\ 021 \end{pmatrix} = \begin{pmatrix} 012 \\ 201 \end{pmatrix}, \text{aber } \begin{pmatrix} 012 \\ 021 \end{pmatrix} \circ \begin{pmatrix} 012 \\ 210 \end{pmatrix} = \begin{pmatrix} 012 \\ 120 \end{pmatrix},$$

  und diese Nichtvertauschbarkeit ist natürlich nicht von der Numerierung der Elemente von X sondern nur von deren Anzahl abhängig. S_3 ist also eine nicht abelsche Gruppe.

Es gelten die (echten!) Inklusionen

$$\text{Gruppoide} \supset \text{Halbgruppen} \supset \text{Monoide} \supset \text{Gruppen} \supset \text{abelsche Gruppen}.$$

In der Regel werden wir die *multiplikative Schreibweise* verwenden, also $*$ durch $\cdot$ ersetzen (und diesen Multiplikationspunkt oft einfachheitshalber sogar weglassen). Höchstens im kommutativen Fall werden wir die *additive Schreibweise* benutzen, also $+$ anstelle von $*$.
Es ist leicht zu zeigen, daß es in einem Gruppoid *höchstens ein neutrales Element* gibt: Für neutrale Elemente $e, e' \in G$ gilt nämlich $e = e * e'$, weil e' neutral ist, und die Neutralität von e ergibt daraus $e * e' = e'$. Bei multiplikativer Schreibweise werden wir dieses Element ggf. als 1 (oder genauer als 1_G) bezeichnen, bei additiver Schreibweise als 0 oder 0_G.
Ebenso mühelos zeigt man, daß Elemente von Monoiden *höchstens ein Inverses* besitzt, denn für zwei Inverse g' und g'' von g gilt offensichtlich

$$g' = e * g' = (g'' * g) * g' = g'' * (g * g') = g'' * e = g''.$$

Das Inverse wird ggf. mit g^{-1} bzw. $-g$ bezeichnet.

Hilfreich ist noch, daß man zum Nachweis der Gruppeneigenschaft von $(G, \cdot)$ manche Rechnung einsparen kann, wenn man folgendes Resultat berücksichtigt:

2.1.3 Satz *Halbgruppen $(G, \cdot)$ mit linksneutralem Element und diesbezüglichen Linksinversen sind Gruppen.*

Beweis: Sei e ein linksneutrales Element von G, g' linksinvers zu g bzgl. e.
i) Wir zeigen zunächst, daß g' auch rechtsinvers ist bzgl. e : Es gibt auch zu g' ein Linksinverses bzgl. e, es sei mit g'' bezeichnet. Hierfür gilt:

$$gg' = e(gg') = (g''g')(gg') = g''(g'g)g' = g''eg' = g''g' = e.$$

ii) Das Element e ist auch rechtsneutral:

$$ge = g(g'g) = (gg')g =_{i)} eg = g.$$

$\square$

Die oben eingeführte *Verknüpfung* $*$ ist, genauer gesagt, eine *zweistellige* Verknüpfung. Allgemeiner kann man *n-stellige* Verknüpfungen als Abbildungen

$$*\colon G^n \to G$$

einführen und auf diese Weise dann die Existenz von Inversen und neutralen Elementen als ein- bzw. zweistellige Verknüpfungen verstehen. Davon werden wir gleich in der nächsten Definition Gebrauch machen:

2.1.4 Definition (Unterstruktur) *Ist G (genauer: die Menge G mit den vorgegebenen Verknüpfungen, der zweistelligen Verknüpfung $*$ und (gegebenenfalls) den ein- oder nullstelligen) ein Gruppoid, eine Halbgruppe ..., dann versteht man unter einem* Untergruppoid, *einer* Unterhalbgruppe, ... *eine Teilmenge $U \subseteq G$, so daß die Einschränkung der Verknüpfungen auf U^2 bzw. auf U^1 oder U^0 eine ebensolche Struktur bilden, also ein Gruppoid, eine Halbgruppe .... Wir kürzen das mit*

$$U \leq G$$

ab, wenn klar ist, welche Art algebraischer Struktur gemeint ist.

•

Daraus folgt beispielsweise, daß Untergruppen ebenfalls nicht leer sind und dasselbe neutrale Element besitzen wie die Obergruppe!

2.1.5 Folgerung *Jeder Durchschnitt von Untergruppoiden, -halbgruppen, -monoiden oder -gruppen ist ebenfalls Untergruppoid, -halbgruppe, -monoid, -gruppe.*

$\square$

Von besonderer Bedeutung ist es, ökonomische Verfahren zu entwickeln, mit denen eine Teilmenge $M \subseteq G$ daraufhin überprüft werden kann, ob sie eine Unterstruktur ist oder nicht. Zunächst einmal genügt natürlich die Überprüfung, ob die Einschränkungen der Verknüpfungen — von G^2, G^1, G^0 auf M^2, M^1, M^0 — aus M hinausführen oder nicht. Man kann dies bei Gruppen aber auch "in einem Aufwasch" verifizieren:

2.1.6 Satz *Sei $(G, \cdot)$ eine Gruppe, $\emptyset \neq U \subseteq G$. Dann gilt:*

i) $U \leq G \iff [u, u' \in U \implies u \cdot (u')^{-1} \in U]$.

ii) Ist U endlich, dann gilt $U \leq G \Longleftrightarrow [u, u' \in U \Longrightarrow u \cdot u' \in U]$.

iii) $U, U' \leq G \Longrightarrow [U \cdot U' \leq G \Longleftrightarrow U \cdot U' = U' \cdot U]$.
Dabei sei $U \cdot U' := \{u \cdot u' \mid u \in U, u' \in U'\}$, *das sogenannte* Komplexprodukt *von U mit U'.*

Beweis: i) ist leicht nachzuprüfen:

a) Ist U eine Untergruppe, dann liegt mit u' auch dessen Inverses u'^{-1} in U und natürlich auch, zu jedem $u \in U$, das Produkt uu'^{-1}.

b) Liegt, umgekehrt, mit $u, u' \in U$ auch das Produkt uu'^{-1} in U, dann gilt das auch für $u' = u$, so daß $1_G = uu^{-1} \in U$ richtig ist. Setzen wir jetzt $u := 1_G$, dann folgt auch $u'^{-1} \in U$, die Inversen liegen also ebenfalls in U, diese Teilmenge ist also eine Untergruppe.

Für ii) beachten wir, daß es, wegen der Endlichkeit von G, für $u' \in U$ natürliche Zahlen $m \neq n$, etwa $m > n$, gibt mit $u'^n = u'^m$, also, wenn etwa $m > n$: $u'^{m-n} = e$, so daß eine Potenz von u' gleich dem Inversen von u' ist. Jetzt folgt aus i) die Behauptung.

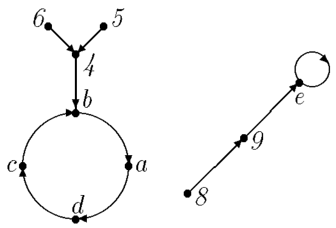
iii) ist Übungsaufgabe.

□

Sehen wir uns jetzt noch ein Beispiel etwas genauer an, die symmetrischen Gruppen auf endlichen Mengen, die bereits eingeführt worden sind. Es sollen einige Bezeichnungsweisen eingeführt werden, die den Umgang mit den Elementen symmetrischer Gruppen erleichtern.

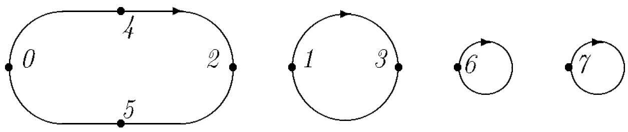
2.1.7 Definition (Endofunktionen, Zyklen, disjunkte Zyklen)

- *Bevor wir* bijektive *Abbildungen einer endlichen Menge X, also Elemente der symmetrischen Gruppe S_X, betrachten, wollen wir uns einmal* beliebige *Abbildungen $f \in X^X$ ansehen. Zunächst ein Beispiel, eine Abbildung $f \in X^X$, wobei $X := \{4, 5, 6, 8, 9, a, b, c, d, e\}$, solche Funktionen heißen* Endofunktionen *auf X. Wir skizzieren uns diese Funktion durch einen* gerichteten Graphen, *die Pfeile geben an, welches Element von X worauf abgebildet wird. Hier ist ein Beispiel:*



Es ist klar, daß eine Endofunktion einer endlichen *Menge diese in zwei Teilmengen einteilt, von denen die eine Teilmenge aus den Elementen x besteht, die von einer geeigneten Potenz f^n auf sich selbst abgebildet werden: $f^n(x) = x$, für geeignetes n. Diese Punkte liegen also auf (gerichteten)* Kreisen *oder auch* Zyklen.

- *Ist die Abbildung dagegen* bijektiv, *dann müssen* alle $x \in X$ *auf solchen Kreisen liegen, hier ist ein Beispiel, ein Element* π *der symmetrischen Gruppe auf* $X := \{0,1,2,3,4,5,6,7\}$:



- *Elemente* $\pi \in S_X$, *die höchstens einen Kreis aus mehr als einem Element enthalten heißen* zyklisch, Zyklus *oder genauer* $r-$Zyklus. *Das ist genau dann der Fall, wenn es — bei geeigneter Numerierung der Elemente von* X *—* $x_0, \ldots, x_{r-1} \in X$ *gibt mit*

$$\pi = \begin{pmatrix} x_0 & x_1 & \ldots & x_{r-2} & x_{r-1} & x_r & \ldots & x_{|X|-1} \\ x_1 & x_2 & \ldots & x_{r-1} & x_0 & x_r & \ldots & x_{|X|-1} \end{pmatrix}.$$

Wir schreiben dann auch kurz

$$\pi = (x_0 \ldots x_{r-1})(x_r) \ldots (x_{|X|-1}),$$

oder gar nur

$$\pi = (x_0 \ldots x_{r-1}),$$

wenn klar ist, auf welcher Gesamtmenge π *wirkt.*

- *Zwei Zyklen* $\pi = (x_{i_0} \ldots x_{i_{r-1}})$ *und* $\rho = (x_{j_0} \ldots x_{j_{s-1}})$ *in* S_X *heißen* disjunkt, *wenn* r *oder* s *gleich 1 ist oder, bei* $r, s \neq 1$,

$$\{x_{i_0}, \ldots, x_{i_{r-1}}\} \cap \{x_{j_0}, \ldots, x_{j_{s-1}}\} = \emptyset.$$

Schreiben wir als Beispiel erneut die Elemente der symmetrischen Gruppe auf der Menge 3 hin, diesmal unter Verwendung der gerade eingeführten Notation für Zyklen:

$$S_3 = \{1, (01), (02), (12), (012), (021)\}.$$

2.1.8 Hilfssatz

i) Disjunkte Zyklen sind vertauschbare Abbildungen.

ii) Ist X *eine endliche Menge, dann ist jedes* $\pi \in S_X$ *darstellbar als Komposition verschiedener und bis auf die Reihenfolge eindeutig bestimmter disjunkter Zyklen.*

iii) $(x_{i_0} \ldots x_{i_{r-1}}) = (x_{i_1} \ldots x_{i_{r-1}} x_{i_0}) = (x_{i_2} \ldots x_{i_{r-1}} x_{i_0} x_{i_1}) = \ldots.$

iv) $(x_{i_0} \ldots x_{i_{r-1}})^{-1} = (x_{i_{r-1}} x_{i_{r-1}} \ldots x_{i_0}),$

v) $(x_{i_0} \ldots x_{i_{r-1}}) = (x_{i_0} x_{i_1})(x_{i_1} x_{i_2}) \ldots (x_{i_{r-1}} x_{i_{r-1}}).$

Beweis: trivial.

□

2.1.9 Definition (Zyklenschreibweise, Listenschreibweise) *Die Darstellung von $\pi \in S_X$ als Produkt disjunkter Zyklen nennt man die* Zyklenschreibweise, *bei Weglassen der 1-Zyklen auch die* verkürzte Zyklenschreibweise. *Ist $\pi \in S_n$ und*

$$\pi = \prod_{\nu=0}^{z(\pi)-1} (j_\nu \pi j_\nu \dots \pi^{l_\nu - 1} j_\nu) \qquad (*)$$

die Zyklenschreibweise für π ($z(\pi)$ steht also für die Anzahl der zyklischen Faktoren, l_ν für deren Längen), dann wird diese eindeutig, wenn wir noch fordern

a) $\forall\ n \in \mathbb{N}$: $j_\nu \leq \pi^n j_\nu$,

b) $j_0 < j_1 < \dots$.

In diesem Fall heißt $(*)$ auch die Standardzyklenschreibweise.
Neben dieser Zyklenschreibweise wird noch die Listenschreibweise *benutzt, insbesondere in Computerprogrammen. Hier schreibt man einfach für $\pi \in S_n$ die Folge der Funktionswerte:*

$$\pi = [\pi 0, \pi 1, \dots, \pi(n-1)].$$

Bijektionen von Mengen auf sich selbst heißen Permutationen, *Gruppen, die aus Permutationen einer Menge X bestehen, also Untergruppen von S_X, X geeignet, heißen entsprechend* Permutationsgruppen, *$|X|$ heißt dabei der* Grad *der Permutation $\pi \in S_X$ bzw. der Untergruppe $U \leq S_X$.*

•

Eine weiter Definition von grundlegender Bedeutung ist das sogenannte Erzeugnis einer Teilmenge:

2.1.10 Definition (Erzeugnis) *Als* Erzeugnis *$\langle T \rangle$ einer Teilmenge T eines Gruppoids, einer Halbgruppe, eines Monoids, einer Gruppe G bezeichnet man die kleinste Unterstruktur von G, die T umfaßt, also, wegen 2.1.5, den Durchschnitt aller Untergruppoide, Unterhalbgruppen, Untermonoide, Untergruppen, die T enthalten:*

$$\langle T \rangle := \bigcap_{U : T \subseteq U \leq G} U.$$

•

Weil das Erzeugnis von T als *kleinste* Unterstruktur definiert ist, die T umfaßt, geht man beim Nachweis, daß eine Teil*menge* U die von T erzeugte Unterstruktur ist, am besten wie folgt vor:

2.1.11 Hilfssatz *Ist G Gruppoid, Halbgruppe, Monoid oder Gruppe, $T \subseteq G$ und U eine Teilmenge von G, dann genügt zum Beweis von $\langle T \rangle = U$ der Nachweis von*

1. *U ist Unterstruktur von G,*

2. *T liegt in U,*

3. *U liegt in jeder Unterstruktur, die T umfaßt.*

□

2.1.12 Beispiele

- Das Erzeugnis von z in der abelschen Gruppe $(\mathbb{Z}, +)$ wird kurz mit $\langle z \rangle$ (anstelle von $\langle \{z\} \rangle$) bezeichnet. Es gilt

$$\langle z \rangle = \{0, \pm z, \pm 2z, \ldots\}.$$

Das Erzeugnis von $\emptyset$ in dieser Gruppe ist

$$\langle \emptyset \rangle = \{0\}.$$

- Ganz allgemein bezeichnet man das Erzeugnis eines Elements g in einer Gruppe G mit $\langle g \rangle$, es gilt offenbar

$$\langle g \rangle = \{1, g, g^2, g^3, \ldots\}.$$

Diese Untergruppe heißt *die von g erzeugte zyklische Untergruppe von G.*

◇

Zur Untersuchung und Anwendung von Gruppen ist sehr oft die Kenntnis eines *möglichst kleinen* Erzeugendensystems hilfreich, wir wollen ein solches für die symmetrischen Gruppen S_n angeben.

2.1.13 Hilfssatz *Die symmetrische Gruppe S_n hat die Ordnung*

$$|S_n| = n! := 1 \cdot 2 \cdots n,$$

(dabei ist zu beachten, daß das leere Produkt *als 1 definiert wird, also $|S_0| = 0! := 1$ behauptet wird) und die folgenden Erzeugendensysteme:*

$$S_n = \langle (01), (12), \ldots, (n-2, n-1) \rangle = \langle (01), (0, \ldots, n-1) \rangle.$$

Beweis: Die Behauptung über die Ordnung ergibt sich durch Induktion nach n :
i) Die Induktionsbasis ist $|S_0| = 1 = 0!$, was offensichtlich richtig ist, denn die einzige Bijektion auf der leeren Menge ist die leere Abbildung.
ii) Zum Vollzug des Induktionsschlusses von n auf $n+1$ betrachten wir die möglichen Bilder des Punktes n in der Menge $n+1 = \{0, \ldots, n\}$ unter Bijektionen auf der Menge $n+1$. Er kann die $n+1$ verschiedenen Werte $0, \ldots, n$ annehmen, wir erhalten also

$$|S_{n+1}| = (n+1)|S_n| = (n+1) \cdot |S_n| = n!,$$

wie behauptet.
Es bleiben noch die Behauptungen über die Erzeugendensysteme nachzuweisen. Zunächst zeigen wir, daß alle zyklischen Permutationen als Produkte von *Transpositionen* (das sind die Bijektionen, die genau zwei Punkte vertauschen und alle anderen auf sich selbst abbilden) schreiben lassen:

$$(i_0 \dots i_{r-1}) = (i_0 i_1)(i_1 i_2) \dots (i_{r-2} i_{r-1}).$$

Jetzt braucht nur noch gezeigt zu werden, daß man von den Transpositionen eigentlich nur die von *benachbarten Punkten* benötigt, was sich aus folgender Gleichung ergibt:

$$(j, k+1) = (k, k+1)(j, k)(k, k+1).$$

Damit ist bewiesen, daß die symmetrische Gruppe von den Transpositionen $(i, i+1)$ benachbarter Ziffern erzeugt wird.
Es bleibt schließlich noch zu verifizieren, daß auch das angegeben System aus nur *zwei* Elementen genügt. Hierzu bemerken wir, daß

$$(i, i+1) = (0, \dots, n-1)^i (01)(0, \dots, n-1)^{-i}$$

gilt, also jede Transposition benachbarter Ziffern aus (01) mit Hilfe der Permutation $(0, \dots, n-1)$ konstruiert werden kann.

□

2.2 Operationen von Gruppen

In diesem Paragraphen wollen wir zeigen, wie Gruppen zur Definition, Abzählung und Konstruktion vieler Strukturen aus Mathematik und Naturwissenschaften benutzt werden können, wenn diese Strukturen als Äquivalenzklassen auf endlichen Mengen definiert sind. Ein Beispiel ist oben bereits erwähnt worden, die unnumerierten Graphen wurden als Äquivalenzklassen numerierter eingeführt. Die vorgesehene Anwendung von Gruppen auf Probleme dieser Art ermöglicht beispielsweise die *Abzählung* von Äquivalenzklassen, also u. a. die Bestimmung der Anzahl unnumerierter Graphen mit vorgegebener Punktezahl.

2.2.1 Definition (Operationen von Gruppen auf Mengen) *Sei G eine (multiplikativ geschriebene) Gruppe, M eine nicht leere Menge. Man sagt, G* operiere *auf M oder M sei eine $G-$*Menge, *wenn eine Abbildung*

$$G \times M \to M : (g, m) \mapsto gm$$

gegeben ist mit

$$g(g'm) = (gg')m, 1m = m.$$

Dies, also die Vorgabe von G, M und einer solchen Abbildung, wird auch mit

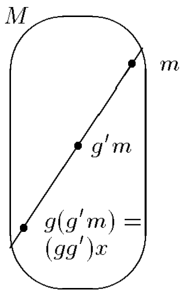
$${}_G M$$

abgekürzt, weil hier G von links *auf M operiert. Ganz entsprechend kann man natürlich Operationen* von rechts *definieren.*

•

Wir bemerken zunächst, daß jede Operation ${}_GM$ eine Äquivalenzrelation auf M definiert:

$$m \sim_G m' :\Longleftrightarrow \ \exists\, g \in G\colon\ gm = m'.$$



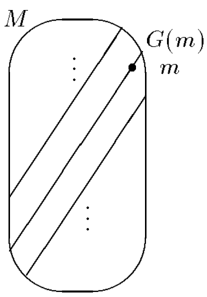
Die Klassen dieser Relation, also die Mengen

$$G(m) := \{gm \mid g \in G\}$$

heißen die *Bahnen* von G auf M. Als Äquivalenzklassen sind zwei Bahnen $G(m)$ und $G(m')$ entweder gleich oder disjunkt. Die Menge *aller* Bahnen wollen wir mit

$$G\backslash\backslash M := \{G(m) \mid m \in M\}$$

bezeichnen.



Ein Repräsentantensystem T von $G\backslash\backslash M$, eine sogenannte *Transversale* von $G\backslash\backslash M$, liefert also eine vollständige Zerlegung von M in disjunkte Teilmengen, eine *Partition* von M :

$$M = \bigcup_{m\in T} G(m).$$

Neben diesen Bahnen der Elemente von M definiert die vorgegebene Operation noch *Untergruppen* von G, und zwar zu jedem $m \in M$:

$$G_m := \{g \in G \mid gm = m\}$$

heißt der *Stabilisator* von m. Betrachten wir einige Beispiele wichtiger Strukturen aus der Algebra, die sich als Bahnen bzw. als Stabilisatoren erweisen:

2.2.2 Anwendungen (Nebenklassen, Konjugiertenklassen, Zentralisatoren)

- Ist U eine Gruppe von G, dann operiert diese per Linksmultiplikation auf G :

  $$U \times G \to G, (u, g) \mapsto ug.$$

  Bahnen sind hier die Mengen

  $$U(g) = Ug = \{ug \mid u \in U\},$$

  die sogenannten *Rechtsnebenklassen* von U in G. Die Menge aller Bahnen bzw. Rechtsnebenklassen bezeichnen wir mit

  $$U\backslash G := U\backslash\backslash G = \{Ug \mid g \in G\}.$$

Bei analoger Operation von U auf G per Rechtsmultiplikation ergeben sich als Bahnen die *Linksnebenklassen*

$$gU = \{gu \mid u \in U\}.$$

Entsprechend die Bezeichnung:

$$G/U := \{gU \mid g \in G\}$$

für die Menge aller Linksnebenklassen von U in G. Bei der Linksmultiplikation ist der Stabilisator von g offenbar die Untergruppe $\{1\}$, ebenso bei der Rechtsmultiplikation.

Als erstes Ergebnis erhalten wir demnach, daß sowohl die Rechtsnebenklassen, als auch die Linksnebenklassen einer Untergruppe U von G eine Partition von G bilden.
Ein konkretes Beispiel ist übrigens die oben bereits erwähnte Beschreibung der Lösungsgesamtheit eines inhomogenen Systems linearer Gleichungen. Setzt man nämlich $G := \mathbb{R}^n$ und $U := L_H$, dann operiert diese (Untergruppe) L_H auf $\mathbb{R}^n$, und die Bahn (irgendeiner) speziellen Lösung x des inhomogenen Systems ist die Nebenklasse (Links- oder Rechts- braucht hier wegen der Kommutativität nicht unterschieden zu werden)

$$L_I = x + L_H.$$

- Eine interessante Operation von G auf sich selbst ist die sogenannte *Konjugation:*

$$G \times G \to G \colon (g', g) \mapsto g'gg'^{-1}.$$

Die Bahnen sind die *Konjugiertenklassen* (der Elemente von G),

$$C^G(g) := \{g'gg'^{-1} \mid g' \in G\}.$$

Der Stabilisator von g wird meist mit

$$C_G(g) := \{g' \in G \mid g'gg'^{-1} = g\}$$

bezeichnet und heißt der *Zentralisator* von g.

Wir erhalten daraus als weiteres Ergebnis, daß auch die Konjugiertenklassen von Elementen der Gruppe eine Partition der Gruppe bilden, und daß die Zentralisatoren Untergruppen sind.

$\diamond$

Mit Hilfe der — gerade als Bahnen beschriebenen — Linksnebenklassen von Untergruppen können wir jetzt ein grundlegend wichtiges Resultat formulieren und beweisen:

2.2.3 Das Fundamentallemma *Ist*

$$G \times M \to M, (g, m) \mapsto gm$$

eine Operation von G auf M, dann ist, für jedes Element m von M, die Bahn von m auf natürliche Weise bijektiv zur Menge der Linksnebenklassen von G_m :

$$G(m) \rightarrowtail\!\!\!\twoheadrightarrow G/G_m, gm \mapsto gG_m.$$

Insbesondere gilt also

$$|G(m)| = |G/G_m|,$$

wenn, wie üblich, G/G_m die Menge der Linksnebenklassen von G_m in G bezeichnet.

Beweis: Wir betrachten folgende Kette von Äquivalenzen:

$$gm = g'm \iff g^{-1}g'm = m \iff g^{-1}g' \in G_m \iff gG_m = g'G_m.$$

Liest man dies von links nach rechts, so ergibt sich die Wohldefiniertheit von $gm \mapsto gG_m$. Liest man von rechts nach links, so ergibt sich die Injektivität, die Surjektivität ist trivial.
Die einzelnen Äquivalenzen ergeben sich wie folgt:

i) Die erste folgt mit den beiden Bedingungen aus der Definition einer Operation.

ii) Die zweite Äquivalenz folgt aus der Definition des Stabilisators.

iii) Die dritte Äquivalenz benutzt eine einfache, aber sehr wichtige Überlegung, die im folgenden in vielen Fällen benutzt werden wird. Sie ist unmittelbar aus der Definition von Links- bzw. Rechtsnebenklassen ersichtlich: $g^{-1}g' \in G_m$ bedeutet die Existenz eines $g'' \in G_m$ mit $g^{-1}g' = g''$, was dasselbe ist wie $g' = gg''$ bzw. wie $g' \in gG_m$. Da verschiedene Linksnebenklassen disjunkt sind, ist das aber äquivalent zu $g'G_m = gG_m$.

□

Hieraus und aus den vorangegangenen Beispielen können wir viele wichtige Folgerungen ziehen: Bei der Operation von U auf G per Linksmultiplikation sind alle Stabilisatoren trivial, $U_g = \{1\}$, mit dem Fundamentallemma ergibt sich

2.2.4 Der Satz von Lagrange *Ist U eine Untergruppe von G, dann haben alle Links- und alle Rechtsnebenklassen von U in G dieselbe Ordnung:*

$$|U| = |Ug| = |gU|.$$

Ist G eine endliche *Gruppe, dann ist $|U|$ ein Teiler von $|G|$ und wir haben*

$$|G/U| = \frac{|G|}{|U|}.$$

□

Aus diesem Satz folgt zum Beispiel, daß Untergruppen von S_3 höchstens die Ordnungen 1,2,3 oder 6 haben können.
Betrachten wir noch ein weiteres Beispiel einer Operation einer Gruppe mit einer interessanten Anwendung auf die Kombinatorik:

2.2.5 Anwendung (Binomialkoeffizienten) Die *natürliche* Operation der symmetrischen Gruppe S_n auf der Menge n, das ist die Operation

$$S_n \times n \to n, (\pi, i) \mapsto \pi(i),$$

induziert, zu jedem $k \leq n$, folgende Operation von S_n auf der Menge $\binom{n}{k}$ aller $k-$ *Teilmengen* von n :

$$S_n \times \binom{n}{k} \to \binom{n}{k}, (\pi, K) \mapsto \pi(K) = \{\pi(x) \mid x \in K\}.$$

Es ist leicht einzusehen, daß jede k-Teilmenge K von n in jede andere k-Teilmenge K' von n übergeführt werden kann, durch Anwendung eines geeigneten $\pi \in S_n$. Mit dem Fundamentallemma und irgendeinem $K \in \binom{n}{k}$ bekommen wir also:

$$|\binom{n}{k}| = |S_n/(S_n)_K|.$$

Diese *Anzahl der k-Teilmengen* von n kürzt man einfachheitshalber mit demselben Symbol ab:

$$\binom{n}{k} := |\binom{n}{k}|,$$

diese Zahlen heißen bekanntlich *Binomialkoeffizienten.* Um sie genauer angeben zu können, brauchen wir uns jetzt nur noch zu überlegen, welche Ordnung der Stabilisator $(S_n)_K$ hat. Er besteht offenbar aus genau den $\pi \in S_n$, die sowohl K als auch den Rest $n\backslash K$ fest lassen, das sind die Permutationen der Form $\pi = \rho \cdot \sigma$, wobei ρ höchstens Elemente von K, σ höchstens Elemente der Restmenge vertauscht. Die Anzahl dieser Produkte ist aber $k! \cdot (n-k)!$, und wir erhalten deshalb das folgende interessante Ergebnis:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

◇

Das für die allgemeine Abzählung von Bahnen von Gruppenoperationen grundlegende Resultat ist

2.2.6 Das Lemma von Cauchy-Frobenius *Ist G eine endliche Gruppe, M eine endliche G-Menge, dann ist die Anzahl der Bahnen von G auf M gleich der mittleren Fixpunktzahl der Elemente von G :*

$$|G\backslash\backslash M| = \frac{1}{|G|}\sum_g |M_g|.$$

2.2.7 Beispiele

- Ist X eine endliche Menge, auf der die endliche Gruppe G von links operiert, Y irgendeine nicht leere endliche Menge, dann induziert die gegebene Operation ${}_G X$ von G auf X eine Operation ${}_G(Y^X)$ von G auf Y^X auf ganz natürliche Weise:
$$G \times Y^X \to Y^X, (g, f) \mapsto \tilde{f},$$
mit $\tilde{f}(x) := f(g^{-1}x)$. Ist jetzt $\bar{g}$ die Permutation von X, die durch $g \in G$ induziert wird: $\bar{g}: x \mapsto gx$, dann ist ein $f \in Y^X$ genau dann Fixpunkt, wenn f konstant auf den Bahnen (= Punktemengen in den zyklischen Faktoren von $\bar{g}$) ist. Das schließt man direkt aus der folgenden Bedingung:
$$f \in (Y^X)_g \iff \forall\ x \in X: \ f(x) = f(g^{-1}x) = f(g^{-1}x) = \dots$$
Wird die Anzahl dieser zyklischen Faktoren mit $z(\bar{g})$ bezeichnet, dann gilt also
$$|(Y^X)_g| = |Y|^{|\langle \bar{g} \rangle \backslash\!\backslash X|} = |Y|^{z(\bar{g})}.$$
Mit dem Lemma von Cauchy–Frobenius erhalten wir demnach
$$|G \backslash\!\backslash Y^X| = \frac{1}{|G|} \sum_{g \in G} |Y|^{z(\bar{g})}.$$

- Eine konkrete Anwendung hiervon ist die Ermittlung der Anzahl unnumerierter Graphen mit vorgegebener Punktezahl n, denn diese entsprechen ja, wie oben bereits beschrieben, der Bahnenmenge
$$S_n \backslash\!\backslash 2^{\binom{n}{2}}.$$
Mit dem Lemma von Cauchy–Frobenius ergibt sich demnach für die Anzahl aller unnumerierten Graphen mit n Punkten die Zahl
$$|S_n \backslash\!\backslash 2^{\binom{n}{2}}| = \frac{1}{n!} \sum_{\pi \in S_n} 2^{z(\bar{pi})}.$$
Bei der Berechnung konkreter Anzahlen

Übungen:

Ü 2.2.1 Beweisen Sie **??** iii).

Ü 2.2.2 Sei X eine endliche Menge und $f \in X^X$. Zeigen Sie, daß äquivalent sind: i) f injektiv, ii) f surjektiv, iii) f bijektiv.

Ü 2.2.3 Zeigen Sie, daß in jeder Halbgruppe H das Produkt einer beliebigen Folge von Elementen $h_1, \dots, h_n, n \geq 1$, unabhängig von der Beklammerung ist.

Ü 2.2.4 Sei G eine Gruppe, sowie $g \in G$. Zeigen Sie, daß die Linkstranslation
$$l_g: G \to G, \quad x \mapsto g * x,$$
bijektiv ist.

Ü 2.2.5 Geben Sie alle Untergruppen von $S_{\underline{3}}$ an.

2.3 Homomorphismen

Algebraische Strukturen werden mit Hilfe *strukturverträglicher* Abbildungen untersucht, die wie folgt definiert werden:

2.3.1 Definition (Homomorphismus) *G und H seien zwei multiplikativ geschriebene Gruppoide, Halbgruppen, Monoide oder Gruppen. Eine Abbildung $f : G \to H$ heißt dann* Homomorphismus, *kurz:*

$$f : G \tilde{\to} H, \text{ oder } f \in Hom(G, H),$$

wenn sie mit allen Verknüpfungen vertauschbar ist. Injektive (surjektive) Homomorphismen heißen Monomorphismen (Epimorphismen)*, bijektive heißen* Isomorphismen. *In diesen Fällen schreibt man entsprechend: $f : G \overset{\sim}{\rightarrowtail} H$ ($f : G \overset{\sim}{\twoheadrightarrow} H$) bzw. $f : G \simeq H$.*

•

Diese Vertauschbarkeit mit den Verknüpfungen bedeutet für die zweistellige Verknüpfung, daß bei multiplikativer Schreibweise gilt

$$f(gg') = f(g)f(g'),$$

in Worten: "Das Bild des Produkts ist das Produkt der Bilder". Gegebenenfalls gilt für die einstellige Verknpfung "Inversenbildung"

$$f(g^{-1}) = f(g)^{-1},$$

und für die nullstellige Verknüpfung "Existenz eines neutralen Elements" gilt (gegebenenfalls, also bei Monoiden):

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

Man braucht dies jedoch meist nicht in allen Einzelheiten nachzuprüfen, denn es gilt

2.3.2 Hilfssatz *Sind G und H Gruppen, dann ist jede mit der Multiplikation verträgliche Abbildung $f: G \to H$ ein Homomorphismus.*

Beweis: Ist $f: G \to H$ eine mit der Multiplikation vertauschbare Abbildung, $g \in G$, dann gilt

$$f(g) = f(g1_G) = f(g)f(1_G),$$

also $f(1_G) = 1_H$, da H ja eine Gruppe ist (so daß $f(g)$ ein Inverses besitzt). Bild des neutralen Elements ist demnach das neutrale Element. Ganz entsprechend weist man nach, daß das Bild des Inversen das Inverse des Bildes ist.

□

2.3.3 Beispiele

- $exp : (\mathbb{R}, +) \to (\mathbb{R}^*, \cdot) : x \mapsto \exp(x)$ ist Isomorphismus,

- $ln : (\mathbb{R}_{>0}, \cdot) \to (\mathbb{R}, +) : x \mapsto \ln(x)$.

Diese Isomorphismen — sie sind zueinader invers — ermöglichen die Konstruktion des Rechenschiebers, der das Multiplizieren reeller Zahlen, etwa x und y, durch das Addieren der Strecken $\ln(x)$ und $\ln(y)$ ersetzt:

$$\exp(\ln(x) + \ln(y)) = x \cdot y.$$

$\diamond$

Ist $f \in \mathrm{Hom}(G, H)$ ein Homomorphismus zwischen zwei multiplikativ geschriebenen Gruppoiden, Halbgruppen, Monoiden oder Gruppen, dann bilden die Klassen $[g]_{R_f}$ der induzierten Äquivalenzrelation R_f mit der Multiplikation

$$[g]_{R_f} \cdot [g']_{R_f} := [gg']_{R_f}$$

das Gruppoid G_{R_f}, denn es ist leicht nachzuprüfen, daß diese Verknüpfung wohldefiniert ist. Die natürliche Abbildung von G hierauf bezeichnen wir mit

$$\nu_f \colon G \to G_{R_f}, g \mapsto [g]_{R_f}.$$

Aus dem Abbildungssatz folgt die Existenz einer Abbildung $\bar{f}$, die das Diagramm

$$\begin{tikzcd} G \arrow[r, "\nu_f"] \arrow[dr, "f"'] & G_{R_f} \arrow[d, "\bar{f}"] \\ & H \end{tikzcd}$$

kommutativ ergänzt, eindeutig bestimmt sowie injektiv ist und die folgende Form hat:

$$\bar{f} \colon G_{R_f} \to H, [g]_{R_f} \mapsto f(g).$$

Man prüft zudem leicht nach, daß es sich bei $\bar{f}$ um einen Homomorphismus handelt:

$$\bar{f}([g]_{R_f}[g']_{R_f}) = \bar{f}([gg']_{R_f}) = f(gg') = f(g)f(g') = \bar{f}([g]_{R_f})\bar{f}([g']_{R_f}).$$

Wenn wir ihren Wertebereich auf das *Bild* von f, das ist

$$\mathrm{Bild}(f) := \{f(g) \mid g \in G\},$$

einschränken, erhalten wir demnach einen Isomorphismus zwischen den beiden Gruppoiden G_{R_f} und $\mathrm{Bild}(f)$. Dieses wichtige Resultat heißt

2.3.4 Der Homomorphiesatz für Gruppoide *Sind G und H zwei Gruppoide, dann gilt für jeden Homomorphismus $f: G \to H$, die davon induzierte Äquivalenzrelation R_f und das Gruppoid aus deren Äquivalenzklassen die folgende Isomorphie (zwischen Gruppoiden)*

$$G_{R_f} \simeq \mathrm{Bild}(f).$$

Das Bild eines Homomorphismus $f: G \tilde{\to} H$ ist also im wesentlichen dasselbe wie G_{R_f}! Wir wollen deshalb dieses Gruppoid genauer beschreiben, insbesondere für den Fall, daß G und H Gruppen sind. (Eigentlich genügt sogar die Forderung, G sei eine Gruppe, weil man H ja durch $\mathrm{Bild}(f)$, d. h. durch Einschränkung des Blickwinkels auf das Bild von f, ersetzen kann, und dieses Bild ist leicht als Gruppe nachgewiesen, s. u.)
Seien also G und H Gruppen, $f \in \mathrm{Hom}(G, H)$. Wir bezeichnen die folgende Teilmenge von G als den *Kern* von f :

$$\mathrm{Kern}(f) := \{g \mid f(g) = f(1_G)\}.$$

$\mathrm{Kern}(f)$ und $\mathrm{Bild}(f)$ sind leicht als Untergruppen nachgewiesen. Darüberhinaus ist der Kern eine Untergruppe mit speziellen Eigenschaften. Solche Untergruppen wollen wir jetzt einführen:

2.3.5 Definition *Ist G eine Gruppe, dann heißt $T \subseteq G$* normale *Teilmenge von G, wenn gilt:*

$$\forall\, g \in G\colon\ gT = Tg,$$

oder, was dasselbe ist

$$\forall\, g \in G\colon\ gTg^{-1} = T.$$

Normale Untergruppen U heißen Normalteiler*, hierfür schreiben wir kurz:*

$$U \trianglelefteq G.$$

•

Wichtige Beispiele normaler Teilmengen sind die bereits erwähnten Konjugiertenklassen von Elementen:

$$C^G(g') = \{gg'g^{-1} \mid g \in G\},$$

es sind die *kleinsten* normalen Teilmengen von G, d. h. jede normale Teilmenge $T \subseteq G$ ist Vereinigung von Konjugiertenklassen $C^G(g')$. Die Umkehrung gilt ebenfalls: Jede Vereinigung von Konjugiertenklassen ist eine normale Teilmenge. Betrachten wir die normalen Untergruppen etwas genauer, es wird sich nämlich gleich zeigen, daß diese *genau die Kerne von Homomorphismen sind!*
Eine Untergruppe $U \leq G$ ist genau dann ein Normalteiler, wenn die Linksnebenklassen auch Rechtsnebenklassen sind (Übungsaufgabe!):

$$G/U = U\backslash G.$$

Genau bei Normalteilern braucht man demnach Links- und Rechtsnebenklassen nicht zu unterscheiden, man spricht dann einfachheitshalber von *Nebenklassen.*

2.3.6 Beispiele

- Als *triviale* Normalteiler bezeichnet man $\{1\}$ und G. Gruppen, die nur diese beiden Normalteiler besitzen, heißen *einfach*.
- Untergruppen vom Index 2 sind stets normal.
- In abelschen Gruppen sind sämtliche Untergruppen normal.

$\diamond$

2.3.7 Satz *Sind G und H Gruppen und $f \in Hom(G, H)$, dann gilt:*

i) Der Kern von f ist Normalteiler:

$$\mathrm{Kern}(f) \trianglelefteq G.$$

ii) Die von f auf G induzierte Äquivalenzrelation besteht gerade aus den Nebenklassen des Kerns:

$$R_f = G/\mathrm{Kern}(f).$$

iii) Die Nebenklassenmenge $G/\mathrm{Kern}(f)$ ist, zusammen mit der oben bereits eingeführten Multiplikation

$$(g\mathrm{Kern}(f)) \cdot (g'\mathrm{Kern}(f)) := (gg')\mathrm{Kern}(f),$$

eine Gruppe, die sogenannte Faktorgruppe *von G nach* $\mathrm{Kern}(f)$.

iv) Ist N ein Normalteiler in G, dann ist die Abbildung

$$\nu_N : G \to G/N, g \mapsto gN$$

ein Epimorphismus von G auf G/N, die Faktorgruppe von G nach N.

Insgesamt bedeutet das, daß genau die Normalteiler von G die Kerne von Homomorphismen auf G sind und die induzierten Äquivalenzrelationen gerade die jeweiligen Nebenklassenmengen.

Beweis: Wir beweisen zunächst i) und ii) gemeinsam:

$$[g]_{R_f} = [g']_{R_f} \iff f(g) = f(g') \iff f(gg'^{-1}) = 1_H \iff gg'^{-1} \in \mathrm{Kern}(f)$$

$$\iff g \in \mathrm{Kern}(f)g' \iff \mathrm{Kern}(f)g = \mathrm{Kern}(f)g'.$$

Da f auf der Nebenklassen $\mathrm{Kern}(f)g$ den Wert $f(g)$ hat, folgt damit: Die Klassen von R_f sind gerade die Rechtsnebenklassen von $\mathrm{Kern}(f)$. Aus Symmetriegründen sind diese Klassen aber auch die Linksnebenklassen des Kerns. Links- und Rechtsnebenklassenmengen sind also gleich, nach obigem Hilfssatz folgt daraus die Normalteilereigenschaft.
Die Nebenklassen des Kerns bilden demnach das Gruppoid $G_{R_f} = G/\mathrm{Kern}(f)$, dieses ist ganz offensichtlich eine Gruppe.

Ist umgekehrt N ein Normalteiler in G, dann ist ν_N eine Abbildung auf die Menge der Nebenklassen von N, und diese bilden eine Gruppe wie man leicht nachrechnet. Die Multiplikation in G/N zeigt sofort, daß diese Abbildung darüberhinaus ein Homomorphismus ist, insgesamt ist es also ein Epimorphismus. □

2.3.8 Anwendung Das Problem der Lösbarkeit des linearen Gleichungssystems

$$\sum_{k=0}^{n-1} a_{ik} x_k = b_i, 0 \leq i \leq m-1,$$

kann als Frage nach dem Urbild von $b \in \mathbb{R}^m$ unter dem Homomorphismus

$$f: \mathbb{R}^n \to \mathbb{R}^m, x \mapsto \sum_k a_{ik} x_k$$

angesehen werden. Das Gleichungssystem ist nämlich ganz offenbar genau dann lösbar, wenn die rechte Seite, der Vektor b, in $\mathrm{Bild}(f)$ liegt. *Der Homomorphiesatz besagt, daß, bei Lösbarkeit, die Lösungsgesamtheit gerade eine Nebenklasse des Kerns des Homomorphismus ist:*

$$L_I = x + \mathrm{Kern}(f).$$

Und der Kern dieses Homomorphismus ist gerade die Lösungsgesamtheit L_H des homogenen Systems, insgesamt gilt also, falls $b \in \mathrm{Bild}(f)$:

$$L_I = x + L_H,$$

mit irgendeinem $x \in f^{-1}(b)$.
Das bestätigt die obigen Vorüberlegungen und zeigt, daß sich diese Lösungsmenge (gegebenenfalls, d. h. wenn sie nicht leer ist) als Nebenklasse erweist. ◇

Aus all diesen Überlegungen halten wir insbesondere folgendes fest, was aus dem letzten Satz und dem allgemeinen Homomorphiesatz folgt:

2.3.9 Der Homomorphisatz für Gruppen *Sind G und H zwei Gruppen und ist $f \in \mathrm{Hom}(G, H)$, dann besteht folgender enge Zusammenhang zwischen Kern und Bild von f :*

$$G/\mathrm{Kern}(f) \simeq \mathrm{Bild}(f).$$

Das Bild von G unter f ist also im wesentlichen dasselbe wie die Faktorgruppe von G nach dem Kern von f.

□

2.3.10 Beispiel Ein nichttrivialer Homomorphismus ist die Abbildung von Permutationen auf das sogenannte *Vorzeichen.* Sie spielt später, bei der Einführung der Determinanten, eine sehr wichtige Rolle:

i) für $n \in \mathbb{N}$ sei das *Differenzenprodukt* Δ_n wie folgt definiert:

$$\Delta_n := \prod_{0 \le i < j \le n-1} (j - i) \in \mathbb{Z}$$

(beachten Sie, daß das “leere Produkt”gleich 1 gesetzt wird, so daß $\Delta_0 = \Delta_1 = 1$ gilt).
ii) Für $\pi \in S_0$ sei $\pi\Delta_0 := 1$ und, für $\pi \in S_1, \pi\Delta_1 := 1$ gesetzt. Dagegen sei, für $n > 1$:

$$\pi\Delta_n := \prod_{0 \le i < j \le n-1} (\pi(j) - \pi(i)).$$

iii) Das *Signum* von $\pi \in S_n, n \in \mathbb{N}$, sei definiert als

$$\mathrm{sgn}(\pi) := \frac{\pi\Delta_n}{\Delta_n}.$$

2.3.11 Hilfssatz *Die Abbildung sgn:* $\pi \mapsto sgn(\pi)$ *ist ein Homomorphismus von* S_n *in die Gruppe* $(\{1, -1\}, \cdot)$, *und für* $n \geq 2$ *ist dieser ein Epimorphismus.*

Beweis: i) Um die Homomorphie zu zeigen, bemerken wir

$$\mathrm{sgn}(\pi\rho) = \prod_{i<j} \frac{\pi\rho(j) - \pi\rho(i)}{\rho(j) - \rho(i)} \cdot \frac{\rho(j) - \rho(i)}{j - i} = \left(\prod_{i<j} \frac{\pi\rho(j) - \pi\rho(i)}{\rho(j) - \rho(i)} \right) \cdot \mathrm{sgn}(\rho)$$

$$= \left(\prod_{\rho(j)<\rho(i)} \frac{\pi\rho(j) - \pi\rho(i)}{\rho(j) - \rho(i)} \right) \cdot \mathrm{sgn}(\rho) = \left(\prod_{i<j} \frac{\pi(j) - \pi(i)}{j - i} \right) \cdot \mathrm{sgn}(\rho)$$

$$= \mathrm{sgn}(\pi) \cdot \mathrm{sgn}(\rho).$$

ii) Ist π Transposition benachbarter Ziffern, dann gilt offenbar $\mathrm{sgn}(\pi) = -1$. $\square$

2.3.12 Definition *Der Kern* $A_n = \{\pi \in S_n \mid sgn(\pi) = 1\}$ *dieses Homomorphismus heißt die* alternierende Gruppe *(in* S_n*).* $\pi \in S_n$ *heißt* gerade*, wenn* $\pi \in A_n$*, andernfalls heißt* π ungerade.

2.3.13 Folgerungen

i) $A_0 = S_0 = \{1\}$, *ebenso* $A_1 = S_1 = \{1\}$, *dagegen ist, für* $n \geq 2$, A_n *Normalteiler vom Index 2 in* S_n.

ii) Ein r*–Zyklus ist genau dann gerade, wenn* r *ungerade ist.*

iii) $\pi \in S_n$ *liegt genau dann in* A_n, *wenn* π *(bezogen auf die Standardzyklenschreibweise oder jede andere Darstellung von* π *als Produkt von Zyklen) geradzahlig viele Zyklen gerader Länge enthält.*

Um dieses Signum auf symmetrische Gruppen beliebiger endlicher Mengen zu verallgemeinern, überlegen wir uns noch:

2.3.14 Hilfssatz

i) Für $\pi \in S_n$ vom Zykeltyp $a(\pi) = (a_1(\pi), \ldots, a_n(\pi))$ gilt

$$sgn(\pi) = (-1)^{n-c(\pi)}, c(\pi) := \sum_i a_i(\pi).$$

ii) sgn ist, für $n \geq 2$, der einzige Homomorphismus von S_n in $(\mathbb{C}^*, \cdot)$, der nicht trivial ist, d.h. verschieden von $\pi \mapsto 1$.

iii) Für endliche Mengen X, Y gilt

$$S_X \simeq S_Y \iff |X| = |Y|.$$

Beweis:

i) $\mathrm{sgn}(\pi) = (-1)^{\sum(i-1)a_i(\pi)} = (-1)^{n-c(\pi)}$.

ii) Sei $f \in Hom(S_n, \mathbb{C}^*)$. Da die Transpositionen S_n erzeugen, diese konjugiert sind, und ± 1 die einzigen Quadratwurzeln von 1 in $\mathbb{C}^*$ sind, ist f genau dann nicht trivial, wenn für jede Transposition τ gilt $f(\tau) = -1$.

iii) $|X| = |Y|$ heißt, daß es $f: X \mapsto Y$ gibt. Für ein solches f sei $\varphi: S_X \to Y^Y: \pi \mapsto f \circ \pi \circ f^{-1}$. Man sieht leicht, daß $\varphi(\pi) \in S_Y$ und daß φ Homomorphismus ist.

$S_X \simeq S_Y$ ergibt $|X|! = |Y|!$, also wegen der Endlichkeit auch $|X| = |Y|$.

$$|X|! = |S_X| = |S_Y| = |Y|!$$

2.3.15 Folgerung *Ist X eine endliche Menge, dann ist, für $|X| \geq 2$, $sgn: \pi \mapsto (-1)^{|X|-c(\pi)}$ der einzige nicht triviale Homomorphismus, sein Kern $A_X = \{\pi \in S_X \mid sgn(\pi) = 1\}$ also die einzige Untergruppe vom Index 2 in S_X.*

Übungen:

Ü 2.3.1 Sei G eine Gruppe. Zeigen Sie:

i) Für jedes $g \in G$ ist die Abbildung $\kappa_g: G \to G, \quad h \mapsto ghg^{-1}$ ein Gruppenisomorphismus.

ii) Bezeichnet man für jedes $g \in G$ mit $l_g: G \to G, \quad x \mapsto gx$, die Linkstranslation, dann ist $\phi: G \to S_G, g \mapsto l_g$, ein Gruppenmonomorphismus.

iii) Ist G abelsch, dann ist $\psi: G \to G, \quad g \mapsto g^{-1}$ ein Gruppenisomorphismus.

Ü 2.3.2 Betrachten Sie $\mathbb{Z}/3\mathbb{Z}$. Beschreiben Sie die Elemente von $\mathbb{Z}/3\mathbb{Z}$, und geben Sie die Multiplikationstafel des Gruppoids $(\mathbb{Z}/3\mathbb{Z}, \cdot)$ mit $a \cdot b := ab$ an.

Ü 2.3.3 Sei nun $n \in \mathbb{N}$· Zeigen Sie, daß $\mathbb{Z}/n\mathbb{Z}$ durch $''\cdot'': \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}/n\mathbb{Z}, (a, b) \mapsto ab$ zu einer Halbgruppe wird. (Ist $''\cdot''$ wohldefiniert?)

Ü 2.3.4 Es sei $m \geq 3$ und $a := (1m) \circ (2m-1) \circ \ldots, \quad b := (1 \ldots m) \in S_m$. Setze $H :=< a >, \quad N :=< b >, \quad D_{2m} :=< a, b >$.

a) Zeigen Sie: $H^N = D_{2m}$, und bestimmen Sie die Mächtigkeit von D_{2m}.

b) Sei nun $m := 4$. Bestimmen Sie zwei Untergruppen U und V von $D_{2 \cdot 4}$ mit $U \lhd V, \quad V \lhd D_{2 \cdot 4}$, aber $U \not\lhd D_{2 \cdot 4}$.

2.4 Strukturen mit Operatorenbereichen

Zu den Strukturen mit einer inneren Verknüpfung $: G \times G \to G$ kommen jetzt Strukturen hinzu, die außerdem noch eine *äußere* Verknüpfung besitzen:

2.4.1 Definition *Sei $(G, *)$ Gruppoid, Ω eine Menge, $\Delta : \Omega \times G \to G$.*

i) G (genauer:$(G, *, \Delta)$) heißt Gruppoid mit Operatorenbereich Ω *oder auch* $\Omega-$Gruppoid, *wenn gilt*

$$\omega\Delta(g * g') = (\omega\Delta g) * (\omega\Delta g'),$$

d. h. wenn l_ω, die Linksmultiplikation mit $\omega \in \Omega$, in $\mathrm{End}(G) := \mathrm{Hom}(G, G)$ *liegt, also ein Endomorphismus von G ist.*

ii) Sind $(G, *, \Delta)$ und $(G', *', \Delta')$ $\Omega-$Gruppoide, $f \in Hom(G, G')$, dann heißt f ein $\Omega-$Homomorphismus *(kurz: $f \in Hom_\Omega(G, H)$), wenn gilt*

$$f(\omega\Delta g) = \omega\Delta' f(g).$$

iii) Ist $(G, *, \Delta)$ ein $\Omega-$Gruppoid, dann heißt ein Untergruppoid $U \subseteq G$ zulässiges Untergruppoid, *wenn gilt*

$$\omega\Delta u \in U.$$

Entsprechend sind zulässige Unterhalbgruppe, ..., zulässige Untergruppe und zulässiger Normalteiler *definiert.*

•

Man kann jeweils leicht nachprüfen, daß sich aus den voraufgegangenen Sätzen über Homomorphismen Sätze über $\Omega-$Gruppen und $\Omega-$Homomorphismen ergeben, wenn man jeweils "Homomorphismus" durch "$\Omega-$Homomorphismus", "Untergruppoid, ..." durch "zulässiges Untergruppoid, ..." und "Normalteiler" durch "zulässigen Normalteiler" ersetzt.

2.4.2 Beispiele

- $\mathbb{R}^n$ hat $\mathbb{R}$ als Operatorenbereich:

$$r\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}.$$

- Die Lösungsgesamtheit L_H eines homogenen linearen Gleichungssystems

$$\sum_{k=0}^{n-1} a_{ik} x_k = 0,\ 0 \leq i \leq m-1,$$

ist zulässige Untergruppe.

- $\mathbb{R}^n, \mathbb{R}^m$ sind abelsche $\mathbb{R}$−Gruppen, und die Abbildung

$$f\colon \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \sum_k a_{ik} x_k \\ \vdots \\ \sum_k a_{mk} x_k \end{pmatrix}$$

  ist ein $\mathbb{R}$−Homomorphismus.

$\diamond$

2.5 Ringe und Körper

Wir betrachten jetzt algebraische Strukturen mit *zwei* inneren Verknüpfungen.

2.5.1 Definition (Ring) *Ist R eine Menge mit zwei inneren Verknüpfungen*

$$+\colon R \times R \to R \text{ und } \cdot\colon R \times R \to R,$$

dann heißt R Ring, *(mit Einselement) wenn gilt*

- $(R,+)$ *ist abelsche Gruppe,*
- $(R,\cdot)$ *ist Halbgruppe (Monoid)*
- l_r, r_t, *die Linksmultiplikation mit r bzw. die Rechtsmultiplikation mit t, liegen in* $\mathrm{End}(\mathbb{R}) := \mathrm{Hom}((\mathbb{R},+),(\mathbb{R},+))$, *d.h. es gelten die sogenannten* Distributivgesetze

$$r(s+t) = rs + rt, (r+s)t = rt + st.$$

•

Ein Ring $(R,+,\cdot)$ heißt *kommutativer Ring,* wenn $(R,\cdot)$ kommutativ ist. Das neutrale Element von $(R,+)$ wird mit 0_R oder kürzer mit 0 bezeichnet und *Null(element)* genannt. Das neutrale Element von $(R,\cdot)$ wird gegebenenfalls mit 1_R bzw. 1 bezeichnet und *Eins(element)* genannt. Das Inverse zu r in $(R,+)$ wird mit $-r$ bezeichnet, und statt $r+(-s)$ schreiben wir $r-s$.

2.5.2 Beispiele

- $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ sind kommutative Ringe, ebenso die *Restklassenringe*

$$\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$$

  mit Addition und Multiplikation *modulo n.*

- Ist R ein Ring, M eine Menge, dann ist $(R^M,+,\cdot)$ ein Ring, wenn wir $+$ und $\cdot$ wieder *punktweise* definieren:

$$(f+g)(m) := f(m) + g(m), (f \cdot g)(m) := f(m) \cdot g(m).$$

- Es gibt auch interessante *nicht punktweise* Verknüpfungen auf solchen Mengen R^M von Abbildungen, sogenannte *Faltungen,* bei denen eine Struktur auf M benutzt wird, z. B. beim Ring $\mathbb{R}^{\mathbb{N}}$ der *Potenzreihen* über $\mathbb{R}$: Hier ist die *Addition punktweise:*

$$(f+g)(n) := f(n) + g(n),$$

  die *Multiplikation die Faltung,* bei der über alle Paare summiert wird, die eine Gleichung lösen:

$$(f \cdot g)(n) := \sum_{(i,j): i+j=n} f(i) \cdot f(j).$$

Eine Unterring hiervon ist der *Polynomring,* der aus den $f \in \mathbb{R}^{\mathbb{N}}$ besteht, die fast überall den Wert 0 haben. Die Potenzreihen und die Polynome schreibt man üblicherweise nicht als Abbildungen sondern als sogenannte *formale Summen:*

$$f = \sum_{n \in \mathbb{N}} a_n x^n.$$

Die Faltung ist dann gerade das bekannte "Ausmultiplizieren".

- Ist $(G, *)$ eine abelsche Gruppe, dann ist $(\mathrm{End}(G), +, \circ)$ mit

  $$(f + f')(g) := f(g) * f'(g)$$

  ein (i.a. nicht kommutativer!) Ring, der *Endomorphismenring* von G.

$\diamond$

Weil Links- und Rechtsmultiplikationen mit Ringelementen Endomorphismen sind, gilt (da Endomorphismen das neutrale Element auf sich selbst und allgemeine Elemente auf ihr Inverses abbilden:

2.5.3 $$r \cdot 0 = 0 \cdot r = 0_R,\ (-r)s = r(-s) = -(rs),\ (-r)(-s) = rs.$$

Ist R ein vom *Nullring* $\{0\}$ verschiedener Ring mit Einselement, dann ist dieses von Null verschieden:

2.5.4 $$[R \neq \{0_R\} \wedge 1_R \in R] \Longrightarrow 0_R \neq 1_R.$$

Das beweisen wir indirekt:

$$r \in R^* := R \backslash \{0\} \Longrightarrow 0 \neq r = r \cdot 1 = r \cdot 0 =_{2.5.3} 0,$$

ein Widerspruch. Per Induktion zeigt man schließlich noch die Gültigkeit der verallgemeinerten Distributivgesetze

2.5.5 $$r \sum_{i=1}^{n} s_i = \sum_{i=1}^{n} r s_i, (\sum_{i=1}^{n} s_i)\, r = \sum_{i=1}^{n} s_i r.$$

Eine Abbildung $f\colon R \to R'$ zwischen zwei Ringen heißt — gemäß der allgemeinen Homomorphismusdefinition — *Ringhomomorphismus,* wenn sie mit allen Verknüpfungen vertauschbar ist, also mit der Addition und der Multiplikation, sowie — bei Ringen mit Einselement — mit der Auswahl des Einselements. (Da sie mit der Addition vertauschbar ist, gilt dies ja dann auch automatisch für die Auswahl des Nullelements und für die Inversenbildung!) Bei Ringen, für die eine Existenz eines Einselements nicht verlangt ist, genügt also die Forderung

$$f(r + r') = f(r) + f(r') \ \wedge f(rr') = f(r)f(r').$$

Entsprechend sind *Ringmonomorphismen, Ringepimorphismen, Ringisomorphismen, Ringendomorphismen* etc. definiert. Kerne von Ringhomomorphismus sind genau die sogenannten *Ideale,* das sind zulässige Untergruppen (mit R als Links- und Rechtsoperatorenbereich). Für Ideale benutzt man dasselbe Zeichen $\trianglelefteq$ wie für Normalteiler:

2.5.6 $$I \trianglelefteq R \iff \forall\, i,i' \in I, r,r' \in R\colon\ i-i', ri, ir' \in I.$$

Die Nebenklassen
$$r+I = \{r+i \mid i \in I\} \subseteq R$$
eines solchen Ideals bilden den *Faktor-, Quotienten-* oder auch *Restklassenring* R/I von R nach I bzgl. den folgenden Verknüpfungen:

$$(r+I)+(r'+I) := (r+r')+I,\ (r+I)\cdot(r'+I) := (r\cdot r')+I.$$

Zusammenfassend ergibt sich

2.5.7 Der Homomorphiesatz für Ringe *Sind R und R' Ringe, und ist f ein Homomorphismus zwischen diesen, dann gilt*

$$\mathrm{Kern}(f) = \{r \in R \mid f(r) = f(0_r) = 0_{R'}\} \trianglelefteq R,$$

und zwischen Bild und Kern besteht die folgende Beziehung:

$$\varphi\colon R/Kernf \simeq \mathrm{Bild}(f)\colon r+Kernf \mapsto f(r).$$

Ist umgekehrt I ein Ideal in R, dann ist dieses der Kern des folgenden Homomorphismus:
$$f\colon R \to R/I, r \mapsto r+I.$$

2.5.8 Beispiele

- Besonders interessant sind die von einem einzigen Element erzeugten Ideale, die sogenannten *Hauptideale* (vgl. Übungsaufgabe!):

  2.5.9 $$(r) = \mathbb{Z}r + Rr + rR + \{\sum_{i=0}^{n} r_i r r_i' \mid n \in \mathbb{N}, r_i, r_i' \in R\}.$$

  Ist R ein Ring mit Einselement, dann vereinfacht sich diese Darstellung zu $(r) = \{\sum_{i=0}^{n} r_i r r_i' \mid n \in \mathbb{N}, r_i, r_i' \in R\}$.

- Der oben bereits erwähnte Ring $\mathbb{Z}_n$ kann auch als Faktorring beschrieben werden:
  $$n\mathbb{Z} := (n) = \{nz \mid z \in \mathbb{Z}\} \trianglelefteq \mathbb{Z}$$
  ist ein Hauptideal in $\mathbb{Z}$, für jedes $n \in \mathbb{N}$. Der Faktorring
  $$\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$$
  entsteht aus $\mathbb{Z}$ durch die Abbildung von $z \in \mathbb{Z}$ auf den Rest $\bar{z}$ von z modulo n. Addition bzw. Multiplikation sind die *Addition modulo n* bzw. die

Multiplikation modulo n. So sind die Verknüpfungstafeln von $\mathbb{Z}_4$ beispielsweise

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

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

- Sei $\mathbb{R}[x]$ der Ring der Polynome in der Unbestimmten x und mit Koeffizienten in $\mathbb{R}$. Das vom Polynom $1 + x^2$ erzeugte Ideal ist

  $$I = \{p \cdot (1 + x^2) \mid p \in \mathbb{R}[x]\},$$

  es besteht also gerade aus den Vielfachen von $1+x^2$. Der Restklassenring $\mathbb{R}[x]/I$ ist isomorph zum Körper (s. u.)

  $$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$$

  der *komplexen Zahlen,* in dem wie folgt addiert und multipliziert wird:

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

  Die Isomorphie zeigt sich durch Anwenden der Abbildung

  $$\mathbb{C} \ni a + bi \mapsto a + bx + I \in \mathbb{R}[x]/I,$$

  denn sie ist offenbar ein Homomorphismus bzgl. Addition, aber auch bzgl. Multiplikation, da

  $$\begin{aligned} (a + bx + I)(c + dx + I) &= (a + bx)(c + dx) + I \\ &= (a + bx)(c + dx) + \underbrace{(-bd)(1 + x^2) + I}_{=0+I} \\ &= (ac - bd) + (ad + bc)x + I. \end{aligned}$$

  Diese Abbildung ist surjektiv, da es für jedes $p \in \mathbb{R}[x]$ Zahlen $a, b \in \mathbb{R}$ gibt mit $p + I = a + bx + I$. Sie ist zudem offensichtlich injektiv, und $1 + I$ ist das neutrale Element.

$\diamond$

Ideale sind Unterstrukturen allgemeiner Ringe (verlangt man von vornherein die Existenz eines Einselements, dann gilt dies allerdings nicht für alle Ideale, sondern nur für das triviale Ideal $R = (1)$). Die Umkehrung gilt aber nicht, da die Idealdefinition noch mehr verlangt, soll doch $ri \in I$ gelten, für *alle* $r \in R$. Ist andererseits R ein Ring mit Einselement 1_R, dann ist die von dem Einselement erzeugte Untergruppe $\langle 1_R \rangle \leq (R, +)$ abgeschlossen gegenüber allen Verknüpfungen, also ein Unterring, die spitzen Klammern entsprechen also sowohl dem Usus, damit die erzeugte *Untergruppe* (diesmal in $(R, +)$) anzugeben, als auch mit der Konvention damit die erzeugte Unterstruktur, in unserem Fall also den von 1_R erzeugten *Unterring,* zu bezeichnen. Die Ordnung dieses Teilrings hat große Bedeutung, wir geben ihr deshalb einen eigenen Namen:

2.5.10 Definition (Charakteristik) *Ist* $(R, +, \cdot)$ *ein Ring mit Einselement, dann heißt*

$$\mathrm{Char}(R) := |\langle 1_R \rangle|$$

die Charakteristik *von* R. *(Ist diese nicht endlich, dann spricht man oft, anstelle von Charakteristik* ∞, *von Charakteristik 0, denn die Charakteristik wird oft auch als die kleinste natürliche Zahl* n *definiert mit* $n1 := 1 + \ldots + 1 = 0$.*)*

2.5.11 Beispiele

- $\mathrm{Char}(\mathbb{Z}) = \mathrm{Char}(\mathbb{Q}) = \mathrm{Char}(\mathbb{R}) = \mathrm{Char}(\mathbb{C}) = \infty$,
- $\mathrm{Char}(\mathbb{Z}_n) = n$.

$\langle 1_R \rangle$ ist Teilring, der *Primring* von R.

$\diamond$

2.5.12 Definition *Ist* $(R, +, \cdot)$ *ein Ring, dann heißt* R Divisionsring *oder auch* Schiefkörper, *wenn* $(\mathbb{R}^*, \cdot)$ *Gruppe ist, im abelschen Fall heißt* R *dann* Körper.

2.5.13 Beispiel

- $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ sind Körper
- Ist R ein kommutativer Ring mit Einselement, I ein Ideal in R, dann ist R/I genau dann Körper, wenn I *maximal* ist in R, d.h. $I \neq R$ und für jedes Ideal J in R mit $I \subset J$ gilt $J = R$. Um das zu beweisen, rechnet man zuerst nach, daß für jedes $x \in R - I$ das von $I \cup \{x\}$ erzeugte Ideal J gerade die folgende Menge ist:

  $$J = \{i + rx \mid i \in I, r \in R\}.$$

  Wegen der Maximalität von I gilt also $J = R$, d. h.

  $$\forall x \in R - I \quad \exists r \in R \colon rx \in 1 + I.$$

  $((R/I)^*, \cdot)$ ist demzufolge abelsche Gruppe.

$\diamond$

Kapitel 3

Lineare Algebra

3.1 Moduln und Vektorräume

Im Zentrum der Untersuchungen in der Linearen Algebra stehen die Vektorräume, das sind abelsche Gruppen mit *Körpern* als Operatorenbereichen. In vielen Überlegungen und Anwendungen stehen aber auch abelsche Gruppen mit *Ringen* als Operatorenbereichen im Mittelpunkt, wir beginnen deshalb mit dieser etwas allgemeineren Definition. Hier müssen wir allerdings — wegen des Verzichts auf Kommutativität des Operatorenbereichs — genau unterscheiden, ob der Ring von rechts oder von links operiert:

3.1.1 Definition (Moduln) *M sei eine abelsche Gruppe, R ein Ring. Dann heißt M* R-Linksmodul, *wenn R Linksoperatorenbereich ist:*

$$r(m+m') = rm + rm',$$

und noch zusätzlich folgendes gilt;

$$(r+r')m = rm + r'm, \ \ r(r'm) = (rr')m, \ \ \textit{sowie } 1_R m = m, \ \textit{falls } 1_R \in R.$$

Hierfür schreiben wir kurz

$${}_RM.$$

*Entsprechend sind $R-$*Rechtsmoduln *M_R definiert. Ist M $R-$Linksmodul und $R'-$Rechtsmodul, dann nennt man M einen $(R,R')-$*Bimodul, *wenn gilt:*

$$(rm)r' = r(mr').$$

Wir schreiben dann

$${}_RM_{R'}.$$

•

3.1.2 Beispiele

- Jede abelsche Gruppe ist auf natürliche Weise $\mathbb{Z}-$Links-, $\mathbb{Z}-$Rechts- und $(\mathbb{Z},\mathbb{Z})-$Bimodul vermöge

$$z \cdot a := a \cdot z := \begin{cases} a + \ldots + a \ \ (z\text{-mal}), & \text{falls } z \geq 0 \\ -a - \ldots - a \ (|z|\text{-mal}), & \text{falls } z < 0. \end{cases}$$

- Jeder Ring R ist $(R,R)-$Bimodul.

- $\mathbb{R}^n$ ist $\mathbb{R}-$Linksmodul vermöge

$$r \cdot \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} := \begin{pmatrix} rx_0 \\ \vdots \\ rx_{n-1} \end{pmatrix}.$$

  $\mathbb{R}^n$ ist natürlich auch $\mathbb{R}$-Rechtsmodul vermöge

$$\begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} \cdot r := \begin{pmatrix} x_0 r \\ \vdots \\ x_{n-1} r \end{pmatrix}.$$

  Das ergibt, wegen der *Assoziativität* der Multiplikation in $\mathbb{R}$ eine $(\mathbb{R},\mathbb{R})$-Bimodulstruktur auf $\mathbb{R}^n$. Wegen der *Kommutativität* der Multiplikation in $\mathbb{R}$ braucht man hier jedoch Links- und Rechts- und Bimoduln nicht zu unterscheiden, man kann einfach von einem $\mathbb{R}$-Modul sprechen.

  Das gilt natürlich ganz allgemein für jeden *Körper* $\mathbb{K}$ und die abelsche Gruppe $\mathbb{K}^n$ und noch allgemeiner für beliebige $\mathbb{K}$-Moduln M, wenn $\kappa \cdot m = m \cdot \kappa$ gesetzt wird. Man kann also auch hier auf die Unterscheidung von Rechts- und Linksmoduln verzichten, spricht von $\mathbb{K}$-Moduln oder $\mathbb{K}$-*Vektorräumen.*

$\diamond$

Es sei zunächst bemerkt, daß unmittelbar aus der Definition folgt

$$r0_M = 0_R m = 0_M, (-r)m = r(-m) = -(rm).$$

Auch hier stellen sich — wie bei den anderen bisher eingeführten algebraischen Grundstrukturen, den Gruppen und Ringen — die Standardfragen nach dem Aussehen von Unterstrukturen, Erzeugnissen und Homomorphismen.
Eine Teilmenge $\emptyset \neq U \subseteq {}_RM$ ist genau dann Untermodul eines R-Linksmoduls M, kurz: $U \leq {}_RM$, wenn sie Untergruppe von $(M,+)$ ist und abgeschlossen gegenüber Linksmultiplikation mit Elementen von R :

$$\forall\ u, u' \in U, r \in R\colon\ u - u', ru \in U.$$

Ist R ein Ring mit Einselement, dann kann man diese Forderung zu *einer* Inklusion vereinfachen:

$$\forall\ u, u' \in U, r, r\text{"} \in R\colon\ ru - r'u' \in U.$$

Klar ist ebenfalls, daß $f\colon {}_RM \to {}_RN$ genau dann *R-(Linksmodul-)Homomorphismus* heißt, wenn

$$f(m+m') = f(m) + f(m'), \text{ und } f(rm) = rf(m)$$

gelten. Solche Abbildungen nennen wir kurz auch *R-lineare* Abbildungen, und wir schreiben dafür $f \in \mathrm{Hom}_R(M,N)$, wenn klar ist, daß M und N als R-Linksmoduln verstanden werden sollen. $f \in \mathrm{Hom}_R(M,N)$ impliziert natürlich

$$f(0_M) = 0_N, f(-m) = -f(m).$$

Entsprechendes gilt für Rechts- und für Bimoduln, und es ergeben sich bei Existenz eines Einselements Vereinfachungen: Offenbar ist $f\colon_R M \to {}_RN$ *in diesem Fall* genau dann $R-$lineare Abbildung, wenn gilt

$$f(rm + r'm') = rf(m) + r'f(m').$$

Kerne und Bilder von R-Homomorphismen sind Unterstrukturen, also Links-, Rechts- oder Bimoduln, und es gilt die Isomorphie

$$M/\mathrm{Kern}(f) \;\simeq_R\; \mathrm{Bild}(f).$$

(Durch den Index R an $\simeq$ soll betont werden, daß über die Homomorphie der abelschen Gruppen hinaus noch Vertauschbarkeit mit der Multiplikation mit Elementen von R, mit der sogenannten *Skalarmultiplikation,* vorliegt.) Hierbei ist die Bildung des *Faktormoduls* natürlich die folgende: Ist $U \le {}_RM$ ein Untermodul, dann wird die Faktorgruppe

$$M/U = \{m + U \mid m \in M\}$$

Zu einem R-Linksmodul vermöge $r(m+U) := rm+U$. Wir bezeichnen ihn kurz mit ${}_R(M/U)$.

3.1.3 Beispiel Erinnern wir erneut an unser Standardproblem, das lineare Gleichungssystem mit der *Koeffizientenmatrix*

$$A = \begin{pmatrix} a_{00} & \dots & a_{0,n-1} \\ \dots & \dots & \dots \\ a_{m-1,0} & \dots & a_{m-1,n-1} \end{pmatrix} =: (a_{ik}) \in \mathbb{R}^{m\times n}.$$

(Ganz allgemein heißen solche Tafeln aus m Zeilen und n Spalten mit Elementen aus einem Ring R, die man ja auch als Abbildungen von $m \times n$ nach R auffassen kann, *$m \times n$-Matrizen* über R.) Die bereits erwähnte Abbildung

$$f\colon \mathbb{R}^n \to \mathbb{R}^m, \begin{pmatrix} x_0 \\ \vdots \\ x_{n-1} \end{pmatrix} \mapsto \begin{pmatrix} \sum_k a_{0k}x_k \\ \vdots \\ \sum_k a_{m-1,k}x_k \end{pmatrix}$$

ist eine $\mathbb{R}$-lineare Abbildung zwischen den Vektorräumen $\mathbb{R}^n$ *und* $\mathbb{R}^m$. (Moduln über Körpern heißen *Vektorräume.*)

◇

Wir müssen jetzt noch das Erzeugnis einer Teilmenge $T \subseteq {}_RM$ diskutieren, es heißt auch die $R-$*lineare Hülle* von T : ${}_R\langle T\rangle$ besteht aus den endlichen Summen der Form

$$m = \sum_{i=0}^{n-1} r_i t_i + \sum_{j=0}^{n'-1} z_j t'_j,$$

mit $n, n' \in \mathbb{N}, r_i \in R, z_i \in \mathbb{Z}, t_i, t'_j \in T$. Enthält R ein Einselement, dann vereinfacht sich dies entsprechend zu

$${}_R\langle T\rangle = \{\sum_{i=0}^{n-1} r_i t_i \mid n \in \mathbb{N}, t_i \in T, r_i \in R\}.$$

Wir nehmen deshalb bis auf Weiteres an, R enthalte ein Einselement!

3.1.4 Beispiele

- ${}_R\langle\emptyset\rangle = \{0_M\}$.
- ${}_R\langle\begin{pmatrix}1\\0\\0\end{pmatrix}\rangle = \{\begin{pmatrix}r\\0\\0\end{pmatrix} \mid r \in \mathbb{R}\}$.
- ${}_R\langle e_0 := \begin{pmatrix}1\\0\\0\end{pmatrix}, e_1 := \begin{pmatrix}0\\1\\0\end{pmatrix}, e_2 := \begin{pmatrix}0\\0\\1\end{pmatrix}\rangle = \mathbb{R}^3$
- ${}_R\langle e_i \mid i \in n\rangle = \mathbb{R}^n$.

Die Elemente $m \in M$ von der Form $\sum_{i=0}^{n-1} r_i t_i$, mit $n \in \mathbb{N}, r_i \in R, t_i \in T, i \in n$, heißen die *Linearkombinationen* aus Elementen von $T \subseteq {}_RM$. Das Erzeugnis von T besteht also (da wir ein Einselement voraussetzen!) genau aus den Linearkombinationen der Elemente von T. ${}_RM$ heißt *endlich erzeugt* oder *endlich erzeugbar*, wenn es eine endliche Teilmenge $T \subseteq M$ gibt mit ${}_R\langle T\rangle = {}_RM$. ${}_RM$ heißt *zyklisch*, wenn es $m \in M$ gibt mit ${}_RM = {}_R\langle m\rangle := {}_R\langle\{m\}\rangle$, d. h. wenn

$${}_RM = \{rm \mid r \in R\} = Rm.$$

Beispielsweise ist $\mathbb{R}^n$ endlich erzeugbar, ${}_{\mathbb{R}}\mathbb{R}[x]$ jedoch nicht, denn im Erzeugnis ${}_{\mathbb{R}}\mathbb{R}[x]$ ist offensichtlich nicht zyklisch, aber ${}_{\mathbb{R}[x]}\mathbb{R}[x] = \mathbb{R}[x] \cdot 1$. Jeder Ring R mit Einselement ist zyklischer $R-$Linksmodul: ${}_RR = R \cdot 1$.
Wir kommen jetzt zur Definition der wichtigsten Begriffe aus der Linearen Algebra:

3.1.5 Definition (Lineare Unabhängigkeit, Basis)

i) Linearkombinationen der Form $0 = \sum_{i=0}^{n-1} r_i t_i$ *heißen* lineare Beziehungen *zwischen den Elementen t von T (es gibt stets die* triviale *lineare Beziehung* $0 = \sum_{t \in T} 0 \cdot t$*).*

ii) T heißt linear unabhängig *oder* frei, *wenn es nur die triviale lineare Beziehung gibt. Andernfalls heißt T* linear abhängig.

iii) Freie Erzeugendensysteme heißen Basen. *Moduln mit Basen heißen* freie *Moduln.*

Lineare Unabhängigkeit einer Teilmenge T bedeutet also, daß für irgendwelche endlich vielen $t_i \in T$ aus einer Gleichung $0_M = \sum_i r_i t_i$ $r_i = 0_R$ folgt, für alle i. Und umgekehrt bedeutet lineare Abhängigkeit von T, daß es endlich viele $t_i \in T$ und $r_i \in R$ gibt, so daß $0_M = \sum_i r_i t_i$, aber nicht alle r_i sind gleich 0_R.

3.1.6 Beispiele

- Liegt 0_M in T, dann ist T linear abhängig, falls $R \neq \{0\}$.
- $\mathbb{R}^n$ ist frei: $T := \{e_i \mid i \in n\}$ ist ein linear unabhängiges Erzeugendensystem. Analoges gilt für $\mathbb{K}^n$, $\mathbb{K}$ ein Körper.
- ${}_{\mathbb{R}}\mathbb{R}[x]$ ist frei, denn $\{1, x, x^2, \ldots\}$ ist ein linear unabhängiges Erzeugendensystem, da jedes Polynom Linearkombination aus *endlich vielen* Potenzen von x ist. Dieser Modul ist aber nicht endlich erzeugbar, denn im Erzeugnis von endlich vielen Polynomen kommen nur Polynome mit beschränktem Grad vor.
- Ist R ein kommutativer Ring mit $1, x \in R$ Nullteiler (d. h. es gibt $r \neq 0$ mit $rx = 0$) dann ist Rx ein Untermodul von ${}_RR$, der nicht frei ist. Dazu verwendet man, daß $rM = \{0\}$, für einen freien Modul M impliziert $r = 0 \vee M = \{0\}$. (Indirekter Beweis: $r \neq 0 \wedge M \neq \{0\}$ ergäbe, für jedes Element b aus einer Basis, den Widerspruch $rb = 0$.) Wir können also wie folgt argumentieren: $Rx \neq 0$, aber $rRx = Rrx = 0$, implizieren — wenn wir annehmen, Rx sei frei — den Widerspruch $r = 0$.
- $\{e_i \mid 0 \leq i \leq n-1\} \subseteq \mathbb{K}^n, \mathbb{K}$ ein Körper, ist Basis.
- $\{\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix}\} \subseteq \mathbb{R}^3$ ist linear abhängig: Setzen wir $r_0 := 1$, $r_1 := -2$, $r_2 := 1$, so erhalten wir eine nicht triviale lineare Beziehung.

3.1.7 Definition *$U_i \leq {}_RM$, dann heißt*

$$\sum_{i \in I} U_i := {}_R\langle \bigcup_{i \in I} U_i \rangle$$

die Summe *der Untermoduln U_i.*

Offensichtlich ist

3.1.8 $$\sum_{i \in I} U_i = \{m \mid \exists\, n \in \mathbb{N}\ \forall\, \nu \in n\ \exists\, u_{i_\nu} \in U_{i_\nu} \colon\ m = \sum_{\nu=1}^{n} u_{i_\nu}\}.$$

□

3.1.9 Folgerungen *Ist I eine endliche Indexmenge und $m_i \in {}_RM, i \in I$, dann sind äquivalent:*

i) ${}_RM = {}_R\langle m_i \mid i \in I\rangle$,

ii) $\forall\ m \in M\ \exists\ r_i \in R$: $m = \sum_{i \in I} r_i m_i$,

iii) ${}_RM = \sum_{i \in I} Rm_i$.

3.1.10 Definition *Die Summe $\sum_{i\in I} U_i$ von Untermoduln $U_i \leq {}_RM$ heißt* direkt, *wenn für alle i gilt: $U_i \cap \sum_{j\neq i} U_j = \{0\}$. Wir schreiben dann für diese direkte Summe auch kurz:*

$$\bigoplus_{i\in I} U_i.$$

3.1.11 Beispiele

- $\mathbb{R}^3 = \mathbb{R}\begin{pmatrix}1\\0\\0\end{pmatrix} \oplus \mathbb{R}\begin{pmatrix}0\\1\\0\end{pmatrix} \oplus \mathbb{R}\begin{pmatrix}0\\0\\1\end{pmatrix} = \mathbb{R}e_0 \oplus \mathbb{R}e_1 \oplus \mathbb{R}e_2.$
- $\mathbb{R}[x] = \mathbb{R}\cdot 1 \oplus \mathbb{R}\cdot x \oplus \mathbb{R}x^2 \oplus \dots$

3.1.12 Satz ${}_RM \geq U_i, i \in I$, *dann sind äquivalent:*

i) $\sum_{i\in I} U_i = \bigoplus_{i\in I} U_i$,

ii) Die Null ist nur trivial kombinierbar *aus Elementen $u_i \in U_i$: Sind $u_i \in U_i, i \in I$, fast alle $u_i = 0$, dann gilt:*

$$\sum_{i\in I} u_i = 0 \Longrightarrow u_i = 0.$$

iii) Jedes $m \in \sum_{i\in I} U_i$ ist eindeutig darstellbar *in der Form*

$$m = \sum_{i\in I} u_i, u_i \in U_i,\ fast\ alle\ u_i = 0.$$

Beweis:
$i) \Longrightarrow ii)$: Ist $\sum u_i = 0, i_o \in I$, dann gilt

$$-u_{i_o} = \sum_{j\neq i_o} u_j \in U_{i_o} \cap \sum_{j \neq i_o} U_j = \{0\},$$

also $u_{i_o} = 0$.
$ii) \Longrightarrow iii)$: Indirekt. $m = \sum u_i = \sum u_i'$ ergibt $\sum(u_i - u_i') = 0$, also $u_i = u_i', i \in I$, wegen ii).
$iii) \Longrightarrow i)$: $m \in U_i \cap \sum_{j\neq i} U_j$ ergibt zwei Darstellungen von m :

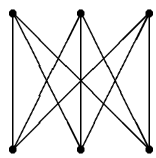
$$m = u_i = \sum_{j\neq i} u_j,$$

was $u_i = 0$ impliziert.

□

3.2 Unabhängigkeitsstrukturen

Unser Ziel ist der Nachweis, daß in Vektorräumen, also in Moduln über Schiefkörpern, Basen existieren und zwei endliche Basen gegebenenfalls von derselben Ordnung sind. (Basen sind sehr wichtig, weil jeder Vektor auf *eindeutige* Weise als Linearkombination der Elemente einer vorgegebenen Basis geschrieben werden kann. Beispielsweise besitzt auch L_H, die Lösungsgesamtheit eines homogenen Gleichungssystems, eine Basis, L_H besteht demnach *genau* aus den Linearkombinationen ihrer Elemente. L_H — eine Menge, die oft unendlich ist — kann deshalb mit Hilfe der endlich vielen Elemente einer solchen Basis vollständig beschrieben werden!)
Zum Beweis dieser Tatsache machen wir einen kurzen Ausflug in die sogenannte *Matroidtheorie,* das ist eine moderne kombinatorische Theorie, die sich als sehr anwendungsrelevant erwiesen hat und die aus der Fragestellung entstanden ist, weshalb diese Gleichheit der Basenordnungen gilt. Mit Hilfe dieser Theorie kann man beispielsweise zeigen, daß das bekannte Spiel, drei Punkte auf einem Blatt Papier jeweils mit drei weiteren Punkten kreuzungsfrei zu verbinden, nicht zum Ziele führen kann:



(In Worten der Graphentheorie: Mit Hilfe der Matroidtheorie kann man zeigen, daß der Graph $K_{3,3}$ *nicht planar* ist!)
Basen sind linear unabhängige Erzeugendensysteme, und man kann sie auch mit Hilfe nur einer dieser beiden Bedingungen definieren, linear unabhängig bzw. Erzeugendensystem zu sein: Basen sind maximale linear unabhängige Mengen, man kann sie aber auch als minimale Erzeugendensysteme charakterisieren. Wir führen deshalb zunächst den Begriff der Halbordnung ein.
Eine Menge H zusammen mit einer Relation $\leq$ heißt *halbgeordnet* bzw. das Paar $(H, \leq)$ heißt *Halbordnung,* wenn $\leq$ reflexiv, antisymmetrisch und transitiv ist. Ein Beispiel hierfür ist in jedem R-Linksmodul M die Menge $\mathcal{U}$ der unabhängigen Teilmengen, zusammen mit der Inklusion:

$$(\mathcal{U}, \subseteq).$$

In einer Halbordnung $(H, \leq)$ heißen diejenigen Elemente h *maximal,* für die $h' \geq h$ die Gleichheit $h = h'$ impliziert, und analog definiert man *minimal.* Ein Beispiel einer Halbordnung ist $(\mathbb{N} \backslash \{0,1\}), |)$, also die Menge der natürlichen Zahlen > 1, mit der Teilbarkeit als Halbordnung. Diese hat genau die Primzahlen als minimale Elemente, maximale Elemente gibt es nicht.
Es ist klar, daß Basen ggf. genau die maximalen Elemente von $(\mathcal{U}, \subseteq)$ sind. Dies erleichtert das Verständnis der folgenden Definition:

3.2.1 Definition (Matroide) *Ist M eine Menge und $\mathcal{U} \neq \emptyset$ ein System von*

Teilmengen von M, dann heißt $(M,\mathcal{U})$ eine Unabhängigkeitsstruktur, *ein* Matroid *oder eine* Prägeometrie *auf M, wenn gilt:*

i) $\mathcal{U}$ ist erblich *oder* hereditär: $S \subseteq T \in U \Longrightarrow S \in U$.

ii) Es gilt der (endliche) Ergänzungssatz:

$$S,T \in \mathcal{U}, |S| = |T|+1 \Longrightarrow \exists\ s \in S\backslash T: T \cup \{s\} \in U.$$

Die Elemente $T \in U$ heißen dann *unabhängige* Mengen, die $S \subseteq M$ mit $S \notin \mathcal{U}$ heißen *abhängige* Mengen.

3.2.2 Beispiele

i) Triviale Unabhängigkeitsstrukturen auf M sind $(M, \{\emptyset\})$ und

$$(M, P(M)),$$

die *universelle* Unabhängigkeitsstruktur, bei der $\mathcal{U}$ der *Potenzmenge* von M gleicht:

$$P(M) := \{T \mid T \subseteq M\}.$$

ii) $M \subseteq {}_{\mathbb{K}}V, \mathbb{K}$ ein Schiefkörper, $\mathcal{U} := \{T \subseteq M \mid T \text{ linear unabhängig}\}$.

a) Die Erblichkeit ist trivial.

b) Der Ergänzungssatz wird indirekt bewiesen. Seien $S,T \in U, |S| = |T|+1$, etwa $S = \{s_0,\ldots,s_m\}_{\neq}, T = \{t_0,\ldots,t_{m-1}\}_{\neq}$, und nehmen wir an, für $0 \le i \le m$ gelte entweder $s_i \in \{t_0,\ldots,t_{m-1}\}$ oder $\{t_0,\ldots,t_{m-1},s_i\}_{\neq}$ ist linear abhängig, Es gibt also jedenfalls $\kappa_{ki} \in \mathbb{K}$ mit

$$s_i = \sum_{k=0}^{m-1} \kappa_{ki} t_k,\ \ 0 \le i \le m.$$

Mit diesen κ_{ki} bilden wir nun folgendes Gleichungssystem:

$$\begin{array}{cccc} x_0\kappa_{00} & +\ldots+ & x_m\kappa_{0,m} & = & 0 \\ & & & \vdots & \\ x_0\kappa_{m-1,0} & +\ldots+ & x_m\kappa_{m-1,m} & = & 0 \end{array} \qquad (1)$$

Da $\mathbb{K}$ Schiefkörper ist, können wir dieses System durch Eliminieren lösen (z.B. können wir, falls $\kappa_{00} \neq 0$, die oberste Gleichung nach x_0 auflösen, das Ergebnis in die nächste Gleichung einsetzen, nach x_1 auflösen, die Ausdrücke für x_0 und x_1 in die dritte Gleichung einsetzen usw.). Vor allem aber gibt es, da es sich bei (1) um $m+1$ *Gleichungen* in nur m *Unbestimmten* handelt, *nicht triviale* Lösungen $(x_0,\ldots,x_m)$.

Für irgendeine dieser nicht trivialen Lösungen gilt aber

$$\sum_{i=0}^{m} x_i s_i = \sum_{i=0}^{m} x_i \sum_{k=0}^{m-1} \kappa_{ki} t_k = \sum_{k=0}^{m-1} (\underbrace{\sum_{i=0}^{m} x_i \kappa_{ki}}_{=0}) t_k = 0,$$

im Widerspruch zur linearen Unabhängigkeit von T.

iii) Ist $(M,\mathcal{U})$ Unabhängigkeitsstruktur, $T \subseteq M$, dann ist auch $(T,\mathcal{U}|_T)$ Unabhängigkeitsstruktur, wenn $\mathcal{U}|_T := \{S \in \mathcal{U} \mid S \subseteq T\}$.

3.2.3 Definition *Ist $(M,\mathcal{U})$ eine Unabhängigkeitsstruktur, dann heißt jedes in der Halbordnung $(\mathcal{U}, \subseteq)$ maximale $T \in \mathcal{U}$ eine* Basis *von $(M,\mathcal{U})$.*

3.2.4 Satz *Ist $(M,\mathcal{U})$ eine Unabhängigkeitsstruktur, dann gilt*

i) Keine echte Teilmenge einer Basis ist ebenfalls Basis.

ii) Gibt es endliche Basen, dann haben alle Basen dieselbe Ordnung.

iii) Gibt es endliche Basen, dann gilt der Austauschsatz: *Sind B, B' Basen, $T \subseteq B$, dann gibt es $T' \subseteq B'$, so daß $|T'| = |T|$ und $(B\backslash T) \cup T'$ ist Basis. D. h. jede Teilmenge T einer Basis B kann gegen eine Teilmenge T' von B' ausgetauscht werden, ohne daß die Basiseigenschaft verloren geht.*

Beweis:

i) ist trivial, ii) folgt aus dem Ergänzungssatz,

iii) folgt aus dem Ergänzungssatz mit vollständiger Induktion. □

3.2.5 Definition *Gegebenenfalls heißt die endliche Ordnung der Basen der* lineare Rang *der Unabhängigkeitsstruktur (kurz: $r(M)$ oder $r(M,\mathcal{U})$. Andernfalls setzt man $r(M) := \infty$ und nennt $(M,\mathcal{U})$* unendlichdimensional.

3.2.6 Satz *Die Rangfunktion $r(-)$ hat auf Unabhängigkeitsstrukturen von endlichem Rang die folgenden Eigenschaften:*

i) $\forall\ S \subseteq T \subseteq M\colon 0 \le r(S) \le r(T) \le r(M)$ (Monotonie),

ii) $\forall\ S,T \subseteq M\colon r(S) + r(T) \ge r(S \cup T) + r(S \cap T)$ *(Submodularität).*

Beweis:

i) ist klar.

ii) $r(M) \in \mathbb{N}$ impliziert $r(S), r(T), r(S\cap T), r(S\cup T) \in \mathbb{N}$. Nach dem Ergänzungssatz kann man eine Basis B von $S \cap T$ ergänzen zu einer Basis $B \cup C$ von S und weiter zu einer Basis $B \cup C \cup D$ von $S \cup T$. Dabei gilt $C \subseteq S - T$ und $D \subseteq T - S$. Offensichtlich gilt

$$r(S \cap T) = |B|, r(S) = |B| + |C|, r(S \cup T) = |B| + |C| + |D|.$$

Wegen $B \cup D \in U$ und $B \cup D \subseteq T$ gilt demnach

$$r(T) \geq |B| + |D| = r(S \cap T) + r(S \cup T) - r(S).$$

$\square$

3.2.7 Beispiel Sei $M := \{1, 2, 3\}$ und $\mathcal{U} := \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}\}$. $(M, \mathcal{U})$ ist offenbar eine Unabhängigkeitsstruktur, und es gilt:

$$2 = r(S) + r(T) > \underbrace{r(S \cup T)}_{=1} + \underbrace{r(S \cap T)}_{=0} = 1.$$

In 3.2.6 ii) kann " $\geq$ " also nicht durch " $=$ " ersetzt werden. $\diamond$

In gewisser Weise dual zu den Basen, den maximalen unabhängigen Mengen, sind die *Kreise,* die minimalen abhängigen Mengen, also die $T \subseteq M$ mit

i) $T \notin \mathcal{U}$,

ii) $T \backslash \{t\} \in \mathcal{U}$, für alle $t \in T$.

Für Vektorräume ergibt sich jetzt:

3.2.8 Satz *Ist* $_{\mathbb{K}}V$ *im Vektorraum,* $B \subseteq V$, *dann sind die folgenden Bedingungen äquivalent:*

i) $V = \bigoplus_{b \in B} \mathbb{K}b$, *und alle* $b \neq 0_v$.

ii) B ist linear unabhängiges Erzeugendensystem.

iii) B ist als Erzeugendensystem minimal.

iv) B ist als linear unabhängige Teilmenge maximal.

Beweis:

i) $\Longrightarrow$ *ii)* :

a) Wäre B linear abhängig, dann gäbe es eine nicht triviale lineare Beziehung $0 = \sum_{b \in B} \kappa_b b$. Daraus folgte, wenn etwa $\kappa_{b_o} \neq 0$, da $\mathbb{K}$ Schiefkörper ist:

$$b_o = \sum_{b \neq b_o} -\kappa_{b_o}^{-1} \kappa_b b \in \mathbb{K}b_o \cap (\sum_{b \neq b_o} \mathbb{K}b)$$

im Widerspruch dazu, daß $\sum \mathbb{K}b$ als direkt vorausgesetzt ist.

b) Daß zudem noch B Erzeugendensystem sein muß , ist trivial.

ii) $\Longrightarrow$ *iii)* : Wäre B als Erzeugendensystem nicht minimal, könnte man ein Element $b_o \in B$ weglassen, dieses wäre dann aber $\neq 0$ und Linearkombination aus Elementen von $B\backslash\{b_o\}$, B also nicht linear unabhängig.

iii) $\Longrightarrow$ *iv)* : Minimale Erzeugendensysteme B sind linear unabhängig, das ist klar, man könnte ja sonst Elemente weglassen. Wären sie als linear unabhängige Teilmengen nicht maximal, dann könnte man ein von B unabhängiges $v \in V - B$ hinzufügen, B wäre dann aber kein Erzeugendensystem.

iv) $\Longrightarrow$ *i)* : Ist B als linear unabhängige Menge maximal, dann gilt wegen der Maximalität, daß $V = \sum_{b \in B} \mathbb{K}b$, und wegen der linearen Unabhängigkeit von B ist diese Summe direkt. □

3.2.9 Folgerungen

i) Die Basen eines $\mathbb{K}-$ Vektorraumes V (gemäß Def. 2.3.6) sind genau die Basen der Unabhängigkeitsstruktur

$$(V,\mathcal{U}),\ \mathcal{U} := \{T \subseteq V \mid T \text{ linear unabhängig}\}.$$

ii) Besitzt V eine endliche Basis, dann haben alle Basen von V dieselbe (endliche) Ordnung

$$dim_{\mathbb{K}}(V) := r(V,\mathcal{U}),$$

die auch als $\mathbb{K}$-Dimension von V bezeichnet wird.

iii) In solchen endlichdimensionalen Vektorräumen gilt der Austauschsatz von Steinitz: *Ist $T \subseteq V$ linear unabhängig, $B \subseteq V$ eine Basis, dann kann man T durch Elemente aus B zu einer Basis von V ergänzen.*

3.2.10 Satz *Jeder Vektorraum besitzt eine Basis.*

Beweis: Wir unterscheiden zwei Fälle:

i) Ist E ein endliches Erzeugendensystem, dann gibt es in E eine Teilmenge T, die als Erzeugendensystem minimal ist, T ist nach 3.2.8 Basis.

ii) Falls dagegen V nicht endlich erzeugbar ist, dann gibt es Ketten

$$T \subset T' \subset T'' \subset \dots$$

linear unabhängiger Teilmengen $T^{(i)}$, die nicht abbrechen, da sonst V endlich erzeugbar wäre. Die Vereinigung

$$S := \bigcup_{i \in \mathbb{N}} T^{(i)}$$

aller Kettenglieder ist offensichtlich linear unabhängig. Also besitzt jede Kette in U eine obere Grenze in $\mathcal{U}$, die Menge der unabhängigen Teilmengen ist, wie man sagt, *strikt induktiv geordnet.* Nach dem *Lemma von Zorn,* das man aus dem Auswahlaxiom herleiten kann (vgl. Scheja/Storch: Lehrbuch der Algebra I, zweite Auflage), besitzen strikt induktiv geordnete Mengen maximale Elemente. □

Ist $B \subseteq V$ eine Basis von V, dann schreiben wir

$$V =_{\mathbb{K}} \ll B \gg .$$

Aus dem Beweis ergibt sich noch (wenn man anstelle *aller* unabhängigen Mengen in V nur die betrachtet, die eine vorgegebene unabhängige Teilmenge $T \subseteq V$ umfassen):

3.2.11 Folgerungen

i) Jede unabhängige Teilmenge $T \subseteq V$ läßt sich zu einer Basis von V ergänzen.

ii) Ist $W \leq_{\mathbb{K}} V$, dann gibt es $W' \leq_{\mathbb{K}} V$ mit $V = W \oplus W'$.

□

Wegen 3.2.11 gilt

3.2.12 Satz *Sind $W, W' \leq_{\mathbb{K}} V$ und ist $dim_{\mathbb{K}}(V) \in \mathbb{N}$, dann gilt*

i) $dim_{\mathbb{K}}(W) < dim_{\mathbb{K}}(V) \Longleftrightarrow W \subset V$.

ii) $dim_{\mathbb{K}}(W + W') + dim_{\mathbb{K}}(W \cap W') = dim_{\mathbb{K}}(W) + dim_{\mathbb{K}}(W')$.

Beweis:

i) Ist B_W eine Basis von W, dann läßt sich B_W zu Basis B von V ergänzen: $B = B_W \bigcup (B \backslash B_W)$, daraus folgt i).

ii) Wir erweitern eine Basis von $W \cap W'$ zu einer Basis von W und einer von W' :

$$B_W = B_{W \cap W'} \bigcup (B_W \backslash B_{W \cap W'}),\ B_{W'} = B_{W \cap W'} \bigcup (B_{W'} \backslash B_{W \cap W'}).$$

Diese Mengen erzeugen jedenfalls $W + W'$:

$$W + W' =_{\mathbb{K}} < \underbrace{B_{W \cap W'} \bigcup (B_W \backslash B_{W \cap W'})}_{=:B_1} \bigcup \underbrace{(B_{W'} \backslash B_{W \cap W'})}_{=:B_2} >$$

B_1 und B_2 sind sogar disjunkt, und ihre Vereinigung ist linear unabhängig, denn eine lineare Relation

$$0 = \sum_{b_1 \in B_1} \kappa_{b_1} b_1 + \sum_{b_2 \in B_2} \kappa_{b_2} b_2$$

ergibt

$$\sum \kappa_{b_1} b_1 = -\sum \kappa_{b_2} b_2 \in W \cap W',$$

also $\kappa_{b_1} = 0$, falls $b_1 \notin B_{W \cap W'}$ und alle $\kappa_{b_2} = 0$. Da $B_{W \cap W'}$ aber Basis ist, folgt daraus sogar, daß *alle* $\kappa_{b_1} = 0$ sind. Die angesetzte lineare Beziehung muß also trivial sein, die beiden Mengen sind demnach disjunkt und ihre Vereinigung ist eine linear unabhängige Menge.

□

3.2.13 Beispiele

i) $\dim_{\mathbb{K}}(\{0_V\}) = \dim_{\mathbb{K}}(\langle\emptyset\rangle) = 0$.

ii) $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$.

iii) $\dim_{\mathbb{K}}(\mathbb{K}^{m\times n}) = m \cdot n$.

iv) $\dim_{\mathbb{K}}(\mathbb{K}[x]) = \infty$.

v) $dim_{\mathbb{C}}(\mathbb{C}) = 1,\ aber\ dim_{\mathbb{R}}(\mathbb{C}) = 2.$

3.3 Lineare Abbildungen und Matrizen

Wir wollen jetzt die *numerische Behandlung* linearer Abbildungen zwischen Vektorräumen beschreiben, bei der vorgegebene Basen die Hauptrolle spielen. Dazu sei daran erinnert, daß $f : {}_{\mathbb{K}}V \to {}_{\mathbb{K}}V'$ genau dann linear, also ein $\mathbb{K}$-Homomorphismus ist, wenn $f(u+v) = f(u) + f(v)$ und $f(\kappa v) = \kappa f(v)$ gelten, für alle $u, v \in V$ und alle $\kappa \in \mathbb{K}$. Es war bereits erwähnt worden, daß man diese beiden Bedingungen auch durch die einzige Bedingung

$$f(\kappa u + \lambda v) = \kappa f(u) + \lambda f(v)$$

ersetzen kann. Lineare Abbildungen sind also Abbildungen zwischen Vektorräumen mit demselben Grundkörper $\mathbb{K}$, wir können deshalb den Index $\mathbb{K}$ meistens weglassen. Hier kommt zunächst eine Liste weiterer sehr wichtiger Eigenschaften solcher linearen Abbildungen:

3.3.1 Satz *Für $\mathbb{K}$-Vektorräume V, V' und eine Teilmenge T von V gilt:*

i) Genau diejenigen $f : V \to V'$ sind linear, die lineare Beziehungen erhalten (d. h. $0 = \sum \kappa_i v_i$ impliziert $0 = \sum \kappa_i f(v_i)$).

ii) Bilder linear unabhängiger Teilmengen unter injektiven linearen Abbildungen sind linear unabhängig.

iii) Ist $V = \mathbb{K}\langle T\rangle$, $\bar{f} : T \to V'$, dann kann $\bar{f}$ genau dann zu einer linearen Abbildung $f : V \to V'$ fortgesetzt werden, wenn $\bar{f}$ lineare Beziehungen erhält. Eine solche lineare Abbildung ist gegebenenfalls eindeutig bestimmt.

iv) Auf linear unabhängigen Teilmengen T definierte Abbildungen sind linear fortsetzbar, auf Basen B definierte sogar eindeutig.

v) Ist $f \in Hom_{\mathbb{K}}(V, V'), T \subseteq V$, dann gilt $f(\langle T\rangle) = \langle f(T)\rangle$.

vi) Ist $T \subseteq V_{\mathbb{K}}$, dann gilt:

a) $f \in Epi_{\mathbb{K}}(V, V'),\ V = \langle T\rangle \Rightarrow V' = \langle f(T)\rangle$,

b) $f \in Mon_{\mathbb{K}}(V, V'),\ V = \ll B \gg \Rightarrow f(V) = \ll f(B) \gg$,

c) $f \in Iso_{\mathbb{K}}(V, V'),\ V = \ll B \gg \Rightarrow V' = \ll f(B) \gg$.

Beweis: i) und ii) sind leicht nachzuprüfen. Zum Beweis von iii) betrachten wir das Diagramm

$$\begin{tikzcd} T \arrow[r, "\iota"] \arrow[dr, "\bar{f}"'] & V \arrow[d, "f"] \\ & V' \end{tikzcd}$$

wobei $\iota\colon T \to V\colon t \mapsto t$ ist, die *Einbettung* von T in V. Wegen der Injektivität von ι gibt es — nach dem Abbildungssatz — eine Abbildung f, für welche dieses Diagramm kommutativ ist, die also $\bar{f}$ auf V fortsetzt. Darüberhinaus ist *jede* Fortsetzung von $\bar{f}$, d. h. jede Abbildung von V nach V' mit $f(t) = \bar{f}(t)$, eine kommutative Ergänzung des Diagramms, also auch die Abbildung

$$f\colon V \to V'\colon \sum_t \kappa_t t \mapsto \sum_t \kappa_f \bar{f}(t),$$

wenn diese wohldefiniert ist, d. h. wenn sie lineare Beziehungen auf T erhält. Und dieses f ist ganz offensichtlich eine lineare Abbildung. Weil T ganz V erzeugt, ist sie sogar *eindeutig bestimmt.* Das beweist iii), iv) und v) und vi) folgen hieraus unmittelbar. □

Sind V und V' endlichdimensional und B, B' Basen, dann wird nach 3.3.1 jedes $f \in \mathrm{Hom}_{\mathbb{K}}(V, V')$ durch Angabe der Bilder $f(b), b \in B$, vollständig bestimmt. Zur systematischen numerischen Beschreibung von f ordnen wir deshalb B und B' zu *Basisfolgen* an:

$$\mathcal{B} = (b_0, \ldots, b_{n-1}), \mathcal{B}' = (b'_0, \ldots, b'_{m-1}).$$

f ist dann vollständig bestimmt durch die Koeffizienten a_{ik} aus den Gleichungen

3.3.2
$$f(b_k) = \sum_{i=0}^{m-1} a_{ik} b'_i.$$

Diese Koeffizienten a_{ik} (die von f, $\mathcal{B}$ und $\mathcal{B}'$ abhängen !) füllen die $m \times n$–Matrix

3.3.3
$$M(\mathcal{B}', f, \mathcal{B}) := (a_{ik}).$$

(Die dabei beachtete Konvention, *das Bild des k–ten Basisvektors in die k–te Spalte* zu schreiben (genauer: dessen Komponenten bzgl. $\mathcal{B}'$), heißt die *Spaltenkonvention.* Gelegentlich findet man auch Bücher, in denen die *Zeilenkonvention* benutzt wird.)

3.3.4 Beispiele Im folgenden bezeichnen wir mit $\mathcal{E}$ bzw. $\mathcal{E}_n$ stets die *Standardbasisfolge* $(e_0, \ldots, e_{n-1})$ von $\mathbb{K}^n$, wobei

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Mit den Basisfolgen

$$\mathcal{B}_1 := (e_2, e_1), \mathcal{B}'_1 := (\begin{pmatrix} 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}), \mathcal{B}_2 := (\begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix}), \mathcal{B}'_2 := (\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 7 \end{pmatrix})$$

von ${}_{\mathbb{R}}V := {}_{\mathbb{R}}\mathbb{R}^2 =: V'$ gilt dann für die lineare Abbildung

$$f: V \to V': \begin{pmatrix} \alpha \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ y \end{pmatrix}$$

folgendes:

$$M(\mathcal{E}, f, \mathcal{E}) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, M(\mathcal{B}_1', f, \mathcal{B}_1) = \begin{pmatrix} 1/3 & 0 \\ 0 & 0 \end{pmatrix}$$

$$M(\mathcal{B}_2', f, \mathcal{B}_2) = \begin{pmatrix} -9/5 & -3 \\ 6/5 & 2 \end{pmatrix}.$$

Die $m \times n$-Matrizen beschreiben also die linearen Abbildungen, aber es gilt noch mehr, denn auch die *Komposition zweier linearen Abbildungen* wird durch eine Matrix beschrieben, die das *Produkt der Matrizen der Kompositionsfaktoren* ist. Hierzu müssen wir aber erst einmal ein Produkt definieren. Zusammen mit der (kanonischen) Addition ergibt sich dabei sogar eine Ringstruktur auf den quadratischen Matrizen vorgegebener Zeilen und Spaltenzahl!

3.3.5 Definition *Ist R ein Ring, $m, n \in \mathbb{N}^*$, dann heißen die Elemente von*

$$R^{m \times n} = \{(a_{ik}) \mid a_{ik} \in R, i \in m, k \in n\}$$

*die $m \times n-$*Matrizen *über R. Die* Elemente *oder* Einträge *a_{ik} von $A = (a_{ik}) \in R^{m \times n}$ werden wie folgt in Zeilen und Spalten angeordnet:*

$$A = \begin{pmatrix} a_{00} & a_{12} & \dots & a_{0,n-1} \\ a_{10} & a_{11} & \dots & a_{1,n-1} \\ \dots & \dots & & \\ a_{m-1,0} & a_{m-1,1} & \dots & a_{m-1,n-1} \end{pmatrix},$$

i ist also der Zeilenindex, *k der* Spaltenindex.

3.3.6 Hilfssatz

i) $R^{m \times n}$ ist $R-$Linksmodul mit der punktweisen Addition und Skalarmultiplikation:

$$(a_{ik}) + (b_{ik}) := (a_{ik} + b_{ik}), r(a_{ik}) := (ra_{ik}).$$

ii) $R^{m \times m}$ ist Ring mit der punktweisen Addition und folgender Multiplikation:

$$(a_{ik})(b_{ik}) := (\sum_{j=0}^{m-1} a_{ij} b_{jk}).$$

Beweis: Nachrechnen. □

Diese Produktbildung läßt sich auf *nicht quadratische* Matrizen verallgemeinern, vorausgesetzt, *die Spaltenzahl des linken Faktors gleicht der Zeilenzahl des rechten Faktors:*

3.3.7 Hilfssatz *Für $A \in R^{m \times n}, B \in R^{n \times p}$ definiert man als Produkt die Matrix*

$$AB := (\sum_{j=1}^{n} a_{ij} b_{jk}).$$

Die Produktbildung ist überall dort, wo sie definiert ist, assoziativ und distributiv:

$$A(BC) = (AB)C, \; A(B+C) = AB + AC, \; (A+B)C = AC + BC.$$

Beweis: Nachrechnen. □

Für $m > 1$ ist der Ring $R^{m \times m}$ nicht kommutativ, wenn R ein Ring mit 1 ist, denn

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ist R ein Ring mit Eins, dann ist demnach auch $R^{m \times m}$ ein Ring mit Einselement

$$I_m := \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Diese Matrix heißt *Einheitsmatrix.* Eine Matrix $A \in R^{m \times m}$ heißt dann *invertierbar* oder *regulär,* wenn es $C \in R^{m \times m}$ gibt mit

$$AC = CA = E_m.$$

Die Matrix $D = (d_{ik}) \in R^{n \times m}$ mit $d_{ik} := a_{ki}$ heißt die zu $A \in R^{m \times n}$ *transponierte* Matrix und wird mit

$${}^{t}A$$

bezeichnet. Ist R kommutativ, dann gilt, für $A \in R^{m \times n}, B \in R^{n \times p}$:

$${}^{t}(AB) = {}^{t}B\,{}^{t}A.$$

Sind $A, B \in R^{m \times m}$ invertierbar, dann gilt

$$(AB)^{-1} = B^{-1}A^{-1}.$$

3.3.8 Satz *Gilt $\dim V_{\mathbb{K}} = n, \dim V'_{\mathbb{K}} = m$, so ist, für Basisfolgen $\mathcal{B}$ von V, $\mathcal{B}'$ von V' :*

i) $\varphi_{\mathcal{B},\mathcal{B}'}\colon Hom_{\mathbb{K}}(V, V') \simeq_{\mathbb{K}} \mathbb{K}^{m \times n} \colon f \mapsto M(\mathcal{B}', f, \mathcal{B})$.

ii) Ist $\mathbb{K}$ kommutativ, also ein Körper, *dann ist die Abbildung*

$$\varphi_{\mathcal{B},\mathcal{B}}\colon End_{\mathbb{K}}(V) \to \mathbb{K}^{n \times n} \colon f \mapsto M(\mathcal{B}, f, \mathcal{B})$$

ein Isomorphismus zwischen Ringen mit Eins.

Beweis:

i) Unmittelbar aus der Definition von $M(\mathcal{B}', f, \mathcal{B})$ und den Definitionen der Verknüpfungen in $\mathbb{K}^{m\times n}$ folgt, daß $\varphi_{\mathcal{B},\mathcal{B}'}$ Homomorphismus, linear und injektiv ist. Die Surjektivität ist ebenfalls offensichtlich: Jede $m\times n-$Matrix beschreibt eine lineare Abbildung.

ii) Die Homomorphie von $\varphi_{\mathcal{B},\mathcal{B}}$ bzgl. Addition folgt aus i). Für die Multiplikation haben wir zu zeigen, daß

$$M(\mathcal{B}, g\circ f, \mathcal{B}) = M(\mathcal{B}, g, \mathcal{B})\cdot M(\mathcal{B}, f, \mathcal{B}).$$

Dies ergibt sich so:

$$\begin{aligned}(g\circ f)(b_k) &= g\sum_i a_{ik}b_i = \sum_i a_{ik}g(b_i)\\ &= \sum_{i,j} a_{ik}b_{ji}b_j = \sum_j(\sum_i b_{ji}a_{ik})b_j.\end{aligned}$$

Die Tatsache $\varphi_{\mathcal{B},\mathcal{B}}(id_V) = E_m$ ist trivial. □

Hätten wir konsequent die Skalarfaktoren *rechts* von den Vektoren geschrieben, dann würde die entsprechende Formel auch für Vektorräume über Schiefkörpern gelten. Wir bleiben aber dabei, die Skalare *links* von den Vektoren zu schreiben, wie meist üblich, und *setzen von jetzt an bis auf weiteres voraus,* $\mathbb{K}$ *sei ein Körper,* also kommutativ.
Ganz analog folgt, für endlichdimensionale V, V', V'', lineare Abbildungen $f \in Hom_{\mathbb{K}}(V, V')$, $g \in Hom_{\mathbb{K}}(V', V'')$ und Basisfolgen $\mathcal{B}, \mathcal{B}'$ und $\mathcal{B}''$:

3.3.9 $$M(\mathcal{B}'', g, \mathcal{B}')M(\mathcal{B}', f, \mathcal{B}) = M(\mathcal{B}'', g\circ f, \mathcal{B}),$$

d. h. die Matrixmultiplikation ist gerade so eingerichtet, daß das Produkt der Matrizen die Komposition der entsprechenden linearen Abbildungen beschreibt!

Ist $V_{\mathbb{K}} = \ll B \gg_{\mathbb{K}}$ und $v = \sum_{b\in B}\kappa_b b$, dann heißen die κ_b die *Komponenten* von v bzgl. B. Im endlichdimensionalen Fall (wenn $\mathcal{B} = (b_0, \ldots, b_{n-1})$ Basisfolge ist), ergibt sich der Isomorphismus

$$\varphi\colon V \simeq_{\mathbb{K}} \mathbb{K}^n : v \mapsto \begin{pmatrix} v_0\\ \vdots\\ v_{n-1}\end{pmatrix},\ falls\ v = \sum_{i=0}^{n-1} v_i b_i.$$

Je zwei n-dimensionale $\mathbb{K}-$Vektorräume sind also zueinander isomorph.
Sehr wichtig ist noch, wie man mit Hilfe der $f \in Hom_{\mathbb{K}}(V, V')$ darstellenden Matrix $A = M(\mathcal{B}', f, \mathcal{B})$ die *Komponenten des Bildes* eines Vektors ermittelt. Ist $v = \sum_i v_i b_i$, dann gilt nämlich, wenn $f(v) = v' = \sum v_i' b_i'$: $v_i' = \sum_k a_{ik}v_k$, oder, in Matrixschreibweise:

3.3.10 $$\begin{pmatrix} v_0'\\ \vdots\\ v_{n-1}'\end{pmatrix} = A\cdot\begin{pmatrix} v_0\\ \vdots\\ v_{m-1}\end{pmatrix}.$$

Sehen wir uns noch an, wie sich die eine lineare Abbildung beschreibende Matrix bei einem Basiswechsel verhält:

3.3.11 Satz *Sind $\mathcal{B}_1, \mathcal{B}_2$ bzw. $\mathcal{B}'_1, \mathcal{B}'_2$ Basisfolgen endlichdimensionaler Vektorräume, und sind $C = (c_{ik}), D = (d_{ik})$ die jeweiligen* Übergangsmatrizen, *d. h. gilt*

$$b_k^1 = \sum_i c_{ik} b_i^2, \; b_k'^1 = \sum_i d_{ik} b_i''^2,$$

dann folgt:

$$D \cdot M(\mathcal{B}'_1, f, \mathcal{B}_1) = M(\mathcal{B}'_2, f, \mathcal{B}_2) \cdot C.$$

Beweis: Man kann das nachrechnen oder sich so klarmachen:

$$\begin{tikzcd}
(V, \mathcal{B}_1) \arrow[r, "C"] \arrow[d, "M(\mathcal{B}'_1, f, \mathcal{B}_1)"'] & (V, \mathcal{B}_2) \arrow[d, "M(\mathcal{B}'_2, f, \mathcal{B}_2)"] \\
(V', \mathcal{B}'_1) \arrow[r, "D"'] & (V', \mathcal{B}'_2)
\end{tikzcd}$$

3.4 Der Gaußsche Algorithmus

Wir kommen jetzt zur expliziten numerischen Lösung des eingangs als eine Motivierung für die Lineare Algebra angegebenen Gleichungssystems

3.4.1 $$\sum_{k=0}^{n-1} a_{ik} x_k = b_i, \quad 0 \leq i \leq m-1.$$

Wie bereits mehrfach erwähnt, ist dazu herauszufinden, ob es überhaupt Lösungen gibt, gegebenenfalls muß dann eine spezielle Lösung $x' \in L_I$ gefunden und eine Basis von L_H angegeben werden, denn $L_I = x + L_H$. All dies leistet der *Gaußsche Algorithmus,* der jetzt vorgeführt werden wird.
Wir gehen dazu von der *Matrixversion* des linearen Gleichungssystems aus:

3.4.2 $$A \cdot x = b, A \in \mathbb{K}^{m \times n}, b \in \mathbb{K}^m.$$

Außerdem benutzen wir, daß diese *Koeffizientenmatrix* A auch als lineare Abbildung f verstanden werden kann, so daß wir die Lösbarkeit des Gleichungssystems als Problem formulieren können festzustellen, ob die *rechte Seite* b in Bild(f) liegt. Dies ist aber genau dann der Fall, wenn die Dimension des *Spaltenraums* von A, also des von den Spalten von A erzeugten Unterraums von $\mathbb{K}^m$, dieselbe ist wie die Dimension des von den Spalten *zuzüglich rechter Seite* erzeugten Unterraums. Dieser letztere, von den Spalten und der rechten Seite erzeugte Unterraum ist der Spaltenraum der um b *erweiterten Matrix:*

$$(A \mid b),$$

die es deshalb zu untersuchen gilt.
Hierzu kann man invertierbare Matrizen C auf beide Seiten der Gleichung $A \cdot x = b$ anwenden, das Gleichungssystem $CAx = Cb$ hat *dieselbe* Lösungsgesamtheit L_I. Wir listen deshalb zunächst diejenigen invertierbaren linearen Abbildungen und die ihnen entsprechenden Matrizen auf, die im Gaußschen Algorithmus Verwendung finden:

3.4.3 Hilfssatz *Ist E die Standardbasis von $\mathbb{K}^m$, dann sind die linearen Fortsetzungen f_i der folgenden Abbildungen $\bar{f}_i: E \to \mathbb{K}^m_{\mathbb{K}}$, $i = 1, 2, 3$, invertierbar:*

a) $\bar{f}_1$ vertauscht e_i mit e_j und läßt alle anderen e_k fest.

b) $\bar{f}_2$ multipliziert e_i mit $\kappa \neq 0$ von rechts und läßt alle anderen e_k fest.

c) $\bar{f}_3$ ersetzt e_i durch $e_i + e_j \cdot \kappa$ mit einem $j \neq i$ und läßt alle $e_k, k \neq i$, fest.

Die Matrizen $C_i := M(\mathcal{E}_m, f_i, \mathcal{E}_m)$ unterscheiden sich von der Einheitsmatrix E_m nur geringfügig:

α) C_1 geht aus E_m durch Vertauschen von $i-$ter Zeile und $j-$ter Spalte hervor.

β) C_2 unterscheidet sich von E_m nur durch den Eintrag κ (statt 1) in $i-$ter Zeile und $i-$ter Spalte.

γ) C_3 unterscheidet sich von E_m nur durch einen Eintrag k (statt 0) in $i-$ter Spalte, $j-$ter Zeile.

Diese Matrizen C_i können wir also von links an A sowie b multiplizieren ohne daß sich die Lösungsgesamtheit L_I ändert:

3.4.4 Folgerung *An der Koeffizientenmatrix A und der rechten Seite b des Gleichungssystems können wir folgende* Zeilenumformungen *anwenden, ohne die Lösungsgesamtheit zu verändern.:*

i) Vertauschen zweier Zeilen,

ii) Multiplizieren einer Zeile mit $\kappa \neq 0$,

iii) Addieren des $\kappa-$fachen der $j-$ten Zeile zur $i-$ten, wobei $i \neq j$.

□

Diese Art von Umformungen nennen wir *elementare Zeilemumformungen* und wenden sie in dem nun folgenden *Gauß–Algorithmus* zur Lösung linearer Gleichungssysteme an!

1. Schritt
(Aus bezeichnungstechnischen Gründen beschränken wir uns auf den "Spezialfall", bei dem die erste Spalte der Ausgangsmatrix A nicht aus lauter Nullen besteht. Andernfalls läßt sich die Unbestimmte x_1 beliebig wählen.) Mit Hilfe von elementaren Zeilenumformungen bringt man das Ursprungsschema, die erweiterte Matrix,

$$(A \mid b) = \left(\begin{array}{ccc|c} a_{00} & \dots & a_{0,n-1} & b_0 \\ \vdots & & \vdots & \vdots \\ a_{m-1,0} & \dots & a_{m-1,n-1} & b_{m-1} \end{array}\right),$$

auf die Form

3.4.5
$$\left(\begin{array}{cccccccccc|c} 1 & * & \dots & * & & & & & & * & * \\ & & & & 1 & * & \dots & * & & & * \\ & & & & & & & & \ddots & & \vdots \\ & 0 & & & & & & & & & * \end{array}\right)$$

Dabei hat man *spaltenweise* von *links* nach *rechts* vorzugehen! (Sterne im Schema 6.2.27 sollen nicht näher spezifizierte Elemente andeuten.)

3.4.6 Satz *Das lineare Gleichungssystem $A \cdot x = b$ ist genau dann* nicht lösbar, *wenn nach Umformung auf die Form 6.2.27 die letzte Spalte eine neue "Stufe" darstellt, also die folgende Form hat:*

$$\left(\begin{array}{ccccccccccccc|c} 1 & * & \dots & * & & & & & & & & & * & * \\ & & & & 1 & * & \dots & * & & & & & & \vdots \\ & & & & & & & & \ddots & & & & & \vdots \\ & & & & & & & & & 1 & * & \dots & * & * \\ & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & 0 \\ & & & & & & & & & & & & & \vdots \\ 0 & & & & & & & & & & & & & 0 \end{array}\right).$$

Genau dann, wenn $A \cdot x = b$ lösbar *ist, geht bei geeigneter Umformung die erweiterte Matrix in eine Matrix der folgenden Form über:*

3.4.7
$$\left(\begin{array}{ccccccccccccc|c} 1 & * & \dots & * & & & & & & & & & * & * \\ & & & & 1 & * & \dots & * & & & & & & \vdots \\ & & & & & & & & \ddots & & & & & \vdots \\ & & & & & & & & & 1 & * & \dots & * & * \\ & & & & & & & & & & & & & 0 \\ & & & & & & & & & & & & & \vdots \\ 0 & & & & & & & & & & & & & 0 \end{array}\right).$$

□

Nun sei das lineare Gleichungssystem als lösbar vorausgesetzt. Wir verabreden folgende Sprechweisen: Die Spalten im Schema 3.4.7, bei denen eine tiefer liegende "Treppenstufe" beginnt, nennen wir *Basisspalten,* die restlichen Spalten (ohne die letzte zusätzliche Spalte) heißen *Fehlspalten.* Basisspalten kennzeichnen wir durch Pfeile $\downarrow$, Fehlspalten durch Pfeile $\uparrow$:

$$\begin{array}{c}
B\;a\;s\;i\;s\;s\;p\;a\;l\;t\;e\;n \\[1em]
\begin{array}{ccccccccccccc}
\downarrow & & & & \downarrow & & & & & \downarrow & & & \\
\end{array} \\
\left(\begin{array}{ccccccccccccc|c}
1 & * & \dots & * & & & & & & & & & * & * \\
 & & & & 1 & * & \dots & * & & & & & & \vdots \\
 & & & & & & & & \ddots & & & & & \vdots \\
 & & & & & & & & & 1 & * & \dots & * & * \\
 & & & & & & & & & & & & & 0 \\
 & & & & & & & & & & & & & \vdots \\
0 & & & & & & & & & & & & & 0
\end{array}\right) \\
\begin{array}{ccccccccccccc}
 & \uparrow & \dots & \uparrow & & \uparrow & \dots & \uparrow & & & \uparrow & \dots & \uparrow
\end{array} \\[1em]
F\;e\;h\;l\;s\;p\;a\;l\;t\;e\;n
\end{array}$$

2. Schritt

Wieder nur mit elementaren Zeilenumformungen machen wir die in den Basisspalten von 3.4.7 oberhalb der 1 stehenden Elemente zu Null. Dabei fangen wir - im Gegensatz zum ersten Schritt - möglichst weit *rechts,* also mit der letzten Basisspalte an und arbeiten uns von rechts nach links spalten-, genauer: *basisspaltenweise* vor. Resultat ist ein Matrixschema der folgenden Form, bei dem wir in die erste Spalte zur Verdeutlichung die *Nummern* j_i der Basisspalten geschrieben haben:

$$\begin{array}{c}
\begin{array}{ccccccccccccccccc}
j_0 & & & & j_1 & & & & j_2 & & & & & j_{r-1} & & &
\end{array} \\
\left(\begin{array}{c|c}
\underbrace{\begin{array}{ccccccccccccccccc}
1 & * & \dots & * & 0 & * & \dots & * & 0 & * & \dots & * & \dots & 0 & * & \dots & * \\
 & & & & 1 & * & \dots & * & 0 & * & \dots & * & \dots & 0 & * & \dots & * \\
 & & & & & & & & 1 & * & \dots & * & \dots & 0 & * & \dots & * \\
 & & & & & & & & & & & & \ddots & & & & \\
 & & & & & & & & & & & & & 1 & * & \dots & * \\
 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & & \\
0 & & & & & & & & & & & & & & & &
\end{array}}_{=:C=(c_{ij})}
&
\begin{array}{c}
t_0 \\ t_1 \\ t_2 \\ \vdots \\ t_{r-1} \\ 0 \\ \vdots \\ 0
\end{array}
\end{array}\right)
\end{array}$$

3.4.8 Satz *Ist die Spalte mit der Nummer p in der Matrix $C = (c_{ij})$ gleich*

$$\begin{pmatrix} c_{0p} \\ \vdots \\ c_{mp} \end{pmatrix}$$

und gleich der Fehl*spalte von C mit der Nummer k, so setze*

$$y^k := \left(\begin{array}{c} 0 \\ \vdots \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{array}\right)\begin{array}{l} \\ \\ \\ \\ p \\ \\ \\ \\ \end{array} + \left(\begin{array}{c} c_{0p} \\ 0 \\ \vdots \\ 0 \\ c_{1p} \\ 0 \\ \vdots \\ 0 \\ c_{2p} \\ \vdots \\ \vdots \end{array}\right)\begin{array}{l} j_0 \\ \\ \\ \\ j_1 \\ \\ \\ \\ j_2 \\ \\ \\ \end{array} \quad , \quad x' := \left(\begin{array}{c} t_0 \\ 0 \\ \vdots \\ 0 \\ t_1 \\ 0 \\ \vdots \\ 0 \\ t_2 \\ \vdots \\ \vdots \end{array}\right)\begin{array}{l} j_0 - te\ Zeile \\ \\ \\ \\ j_1 - te\ Zeile \\ \\ \\ \\ j_2 - te\ Zeile \\ \\ \\ \end{array}$$

Dann gilt:

1. *x' ist eine spezielle Lösung von $A \cdot x = b$.*
2. *Die $d = n - r$ Vektoren $y^0, \ldots, y^{d-1}$ bilden eine Basis des Vektorraumes $\{y \in \mathbb{R}^n \mid A \cdot y = 0\}$.*
3. *Die Lösungsgesamtheit von $A \cdot x = b$ ist (vgl. ??):*

$$\{x' + \kappa_0 y^0 + \ldots + \kappa_{d-1} y^{d-1} \mid \kappa_i \in \mathbb{K}\}.$$

□

3.4.9 Beispiel Gegeben sei folgendes lineare Gleichungssystem über $\mathbb{R}$:

$$\begin{array}{rcrcrcrcrcrcrcrcr}
-4x_1 &+& 4x_2 &-& 8x_3 &-& 24x_4 &-& 44x_5 &+& 4x_6 &-& 56x_7 &-& 44x_8 &=& -24 \\
3x_1 &-& 3x_2 &+& 6x_3 &+& 18x_4 &+& 30x_5 &-& 9x_6 &+& 42x_7 &+& 24x_8 &=& 15 \\
2x_1 &-& 2x_2 &+& 4x_3 &+& 10x_4 &+& 16x_5 &-& 4x_6 &+& 20x_7 &+& 12x_8 &=& 8 \\
-2x_1 &+& 2x_2 &-& 4x_3 &-& 12x_4 &-& 18x_5 &+& 10x_6 &-& 28x_7 &-& 10x_8 &=& -8 \\
2x_1 &-& 2x_2 &+& 4x_3 &+& 10x_4 &+& 18x_5 && &+& 20x_7 &+& 18x_8 &=& 10
\end{array}.$$

Die dazugehörige erweiterte Matrix $(A \mid b)$ lautet:

$$\left(\begin{array}{rrrrrrrr|r}
-4 & 4 & -8 & -24 & -44 & 4 & -56 & -44 & -24 \\
3 & -3 & 6 & 18 & 30 & -9 & 42 & 24 & 15 \\
2 & -2 & 4 & 10 & 16 & -4 & 20 & 12 & 8 \\
-2 & 2 & -4 & -12 & -18 & 10 & -28 & -10 & -8 \\
2 & -2 & 4 & 10 & 18 & 0 & 20 & 18 & 10
\end{array}\right).$$

1. Wir bringen zunächst eine Eins an die Position $(0,0)$, etwa durch die Division der nullten Zeile durch -4, kurz: $(Z_0 \| Z_0/(-4))$:

$$\left(\begin{array}{rrrrrrrr|r} \boxed{1} & -1 & 2 & 6 & 11 & -1 & 14 & 11 & 6 \\ 3 & -3 & 6 & 18 & 30 & -9 & 42 & 24 & 15 \\ 2 & -2 & 4 & 10 & 16 & -4 & 20 & 12 & 8 \\ -2 & 2 & -4 & -12 & -18 & 10 & -28 & -10 & -8 \\ 2 & -2 & 4 & 10 & 18 & 0 & 20 & 18 & 10 \end{array}\right).$$

2. Nun räumen wir die *unterhalb* von $\boxed{1}$ stehenden Elemente aus durch $(Z_1 \| Z_1 - 3Z_0;\ Z_2 \| Z_2 - 2Z_0;\ Z_3 \| Z_3 + 2Z_0;\ Z_4 \| Z_4 - 2Z_0)$:

$$\left(\begin{array}{rrrrrrrr|r} 1 & -1 & 2 & 6 & 11 & -1 & 14 & 11 & 6 \\ 0 & 0 & 0 & 0 & -3 & -6 & 0 & -9 & -3 \\ 0 & 0 & 0 & -2 & -6 & -2 & -8 & -10 & -4 \\ 0 & 0 & 0 & 0 & 4 & 8 & 0 & 12 & 4 \\ 0 & 0 & 0 & -2 & -4 & 2 & -8 & -4 & -2 \end{array}\right).$$

3. Nachdem wir die nullte Spalte von A weitgehend ausgeräumt haben, suchen wir in dem neuen Zahlenschema die nächste Spalte, die *unterhalb* der nullten Zeile vom Nullvektor verschieden ist. (Hier ist das also die Spalte mit der Nummer 3.) Durch Zeilenvertauschungen bringen wir eine von 0 verschiedene Zahl an die oberste Position des Rechtecks (etwa durch die Vertauschung $(Z_1 \leftrightarrow Z_2)$):

$$\left(\begin{array}{rrrrrrrr|r} 1 & -1 & 2 & 6 & 11 & -1 & 14 & 11 & 6 \\ 0 & 0 & 0 & -2 & -6 & -2 & -8 & -10 & -4 \\ 0 & 0 & 0 & 0 & -3 & -6 & 0 & -9 & -3 \\ 0 & 0 & 0 & 0 & 4 & 8 & 0 & 12 & 4 \\ 0 & 0 & 0 & -2 & -4 & 2 & -8 & -4 & -2 \end{array}\right).$$

4. Durch Anwendung von $(Z_1 \| Z_1/(-2))$ erreichen wir, daß eine Eins an der Position $(1,3)$ steht. Um bequemer rechnen zu können, wenden wir noch $(Z_2 \| Z_2/(-3);\ Z_3 \| Z_3/4;\ Z_4 \| Z_4/(-2))$ an und erhalten:

$$\left(\begin{array}{rrrrrrrr|r} 1 & -1 & 2 & 6 & 11 & -1 & 14 & 11 & 6 \\ 0 & 0 & 0 & \boxed{1} & 3 & 1 & 4 & 5 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & 2 & -1 & 4 & 2 & 1 \end{array}\right).$$

5. Jetzt sind die unterhalb von $\boxed{1}$ stehenden Elemente zu 0 zu machen

$(Z_4 \| Z_4 - Z_1)$:

$$\left(\begin{array}{cccccccc|c}
1 & -1 & 2 & 6 & 11 & -1 & 14 & 11 & 6 \\
0 & 0 & 0 & 1 & 3 & 1 & 4 & 5 & 2 \\
0 & 0 & 0 & 0 & \boxed{1} & 2 & 0 & 3 & 1 \\
0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 1 \\
0 & 0 & 0 & 0 & -1 & -2 & 0 & -3 & -1
\end{array}\right).$$

Nun ist diejenige Spalte zu suchen, die unterhalb der Zeile mit der Nummer 1 vom Nullvektor verschieden ist.

6. Da in unserem Beispiel bereits eine $\boxed{1}$ an der Position $(2,4)$ steht, müssen nur noch die Einträge unterhalb der Eins zu 0 gemacht werden ($Z_3 \| Z_3 - Z_2$; $Z_4 \| Z_4 + Z_2$):

$$\begin{array}{c}
\left(\begin{array}{cccccccc|c}
1 & -1 & 2 & 6 & 11 & -1 & 14 & 11 & 6 \\
0 & 0 & 0 & 1 & 3 & 1 & 4 & 5 & 2 \\
0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right) \\
\begin{array}{cccccccc}
\uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\
\mathrm{B} & \mathrm{F} & \mathrm{F} & \mathrm{B} & \mathrm{B} & \mathrm{F} & \mathrm{F} & \mathrm{F}
\end{array}
\end{array}$$

(B= Basisspalte, F= Fehlspalte).

7. Damit ist der erste Schritt durchgeführt, die Matrix besitzt Treppengestalt. Wir können folgende Eigenschaften direkt ablesen:

   - Das Gleichungssystem $Ax = b$ ist lösbar.
   - $\mathrm{Rang}(A) = 3$ (Anzahl der Basisspalten).
   - Der Vektorraum $\{y \in \mathbb{R}^8 \mid Ay = 0\}$ ist fünfdimensional (Anzahl der Fehlspalten).

8. Als nächstes sind die in den Basisspalten *oberhalb* der unteren 1 stehenden Elemente zu eliminieren. Dabei gehen wir von rechts nach links vor, beginnen also mit der Basisspalte mit der Nummer 2. Die Umformungen ($Z_1 \| Z_1 - 3Z_2$; $Z_0 \| Z_0 - 11Z_2$) liefern:

$$\begin{array}{c}
\left(\begin{array}{cccccccc|c}
1 & -1 & 2 & 6 & 0 & -23 & 14 & -22 & -5 \\
0 & 0 & 0 & 1 & 0 & -5 & 4 & -4 & -1 \\
0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right). \\
\begin{array}{cccccccc}
 & & & & \uparrow & & &
\end{array}
\end{array}$$

9. Mit $(Z_0 \| Z_0 - 6Z_1)$ erreichen wir das entsprechende für die zweite Basisspalte:

$$\begin{array}{c}\left(\begin{array}{rrrrrrrr|r} 1 & -1 & 2 & 0 & 0 & 7 & -10 & 2 & 1 \\ 0 & 0 & 0 & 1 & 0 & -5 & 4 & -4 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) . \\ \begin{array}{rrrrrrrrr} & & & \uparrow & & & & & \end{array}\end{array}$$

Die nullte Basisspalte hat bereits die gewünschte Gestalt.

10. Erinnern wir uns, daß die Spalten von C den Komponenten $x_0, \ldots, x_7$ des Vektors x entsprechen. x_0, x_3, x_4 sind also Basisspaltenvariable, während x_1, x_2, x_5, x_6, x_7 Fehlspaltenvariable sind. Im Schema liest sich dies so:

$$x = \begin{pmatrix} \boxed{B} \\ \boxed{F} \\ \boxed{F} \\ \boxed{B} \\ \boxed{B} \\ \boxed{F} \\ \boxed{F} \\ \boxed{F} \end{pmatrix} .$$

Eine spezielle Lösung erhalten wir nun dadurch, daß wir die Einträge der rechten Spalte von oben nach unten auf die durch ein B gekennzeichneten Plätze verteilen und die F-Plätze mit 0 besetzen:

$$X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} .$$

Eine Basis $y^0, \ldots, y^4$ von $\{y \in \mathbb{R}^8 \mid Ay = 0\}$ erhalten wir so: Für die nullte Fehlspalte betrachten wir den nullten F-Platz im obigen Schema. Wir setzen an diese Stelle eine -1. Alle anderen F-Plätze werden mit 0 vorbesetzt. Nun verteilen wir die Einträge der nullten Fehlspalte von oben nach unten auf die B-Plätze und erhalten y^0. Entsprechend ergeben sich aus den anderen Fehlspalten die Vektoren $y^1, \ldots, y^4$.

$$y^0 = \begin{pmatrix} -1 \\ \boxed{-1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} y^1 = \begin{pmatrix} 2 \\ 0 \\ \boxed{-1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} y^2 = \begin{pmatrix} 7 \\ 0 \\ 0 \\ -5 \\ 2 \\ \boxed{-1} \\ 0 \\ 0 \end{pmatrix} y^3 = \begin{pmatrix} -10 \\ 0 \\ 0 \\ 4 \\ 0 \\ 0 \\ \boxed{-1} \\ 0 \end{pmatrix} y^4 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -4 \\ 3 \\ 0 \\ 0 \\ \boxed{-1} \end{pmatrix}$$

Wir bemerken vor allem, daß diese y^i linear unabhängig sind, das homogene System lösen und deshalb eine Basis von L_H besitzen, denn ihre Anzahl entspricht gerade der Dimension von L_H. Dieser Raum ist ja gerade der Kern der von A beschriebenen Abbildung, so daß gilt:

$$\dim(L_H) = \dim(\text{Kern}(f)) = n - \dim(\text{Bild}(f)) = n - \dim(\text{Spaltenraum von} A)$$

$$= n - \text{Anzahl der Basisspalten} = \text{Anzahl der Fehlspalten.}$$

Die *Lösungsgesamtheit* des linearen Gleichungssystems ist demnach:

$$\{x \in \mathbb{R}^8 \mid x = x' + \kappa_0 y^0 + \kappa_1 y^1 + \kappa_2 y^2 + \kappa_3 y^3 + \kappa_4 y^4 \text{ mit } \kappa_i \in \mathbb{R}\}.$$

3.5 Duale Vektorräume und Abbildungen

Wir wollen im Folgenden auch geometrische Zusammenhänge mathematisch beschreiben und beginnen deshalb jetzt mit der Einführung hierfür geeigneter Begriffe.
Betrachten wir zunächst eine einzelne homogene lineare Gleichung

$$\sum_k a_{ik} x_k = 0$$

unter geometrischen Aspekten. Wenn nicht sämtliche Koeffizienten verschwinden, dann hat die (einzeilige) Koeffizientenmatrix den Rang 1, der Lösungsraum des einzeiligen homogenen Systems also die Dimension $n-1$, solche Unterräume heißen *Hyperebenen,* in Verallgemeinerung der Definition von Ebene als Raum der Dimension $2 = 3-1$, also als Hyperebene im Dreidimensionalen. Ein homogenes lineares Gleichungssystem hat also als Lösungsgesamtheit einen Schnitt von Hyperebenen. Wir werden bald zeigen können, daß auch umgekehrt zu vorgegebenen Hyperebenen leicht ein Gleichungssystem aufgestellt werden kann, das den Schnitt dieser Hyperebenen als Lösungsgesamtheit besitzt.
Wir verweisen deshalb zunächst auf die abbildungstheoretische Interpretation einer einzelnen linearen Gleichung: Zu ${}_{\mathbb{K}}V$ ergibt sich auf natürliche Weise der Vektorraum

$$L(V) := \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$$

der *Linearformen* λ auf V. Beispiele von Linearformen sind linke Seiten einzelner homogener linearer Gleichungen:

$$\lambda: \mathbb{K}^n \to \mathbb{K}: x \mapsto \sum_k a_{ik} x_k .$$

Eine Verallgemeinerung des Begriffs der Linearform ist der der *Bilinearform,* worunter man Abbildungen $\beta: V^* \times V \to \mathbb{K}$ versteht, für $\mathbb{K}$-Vektorräume V^*, V, versteht, die in beiden Komponenten linear sind, d. h. für alle $v^* \in V^*$ bzw. $v \in V$ ist $\beta(v^*, -): v \mapsto \beta(v^*, v)$ bzw. $\beta(-, v): v^* \mapsto \beta(v^*, v)$ Linearform. Die Menge *aller* solchen Bilinearformen sei mit

$$BLF(V^*, V)$$

bezeichnet.

3.5.1 Definition (Skalarprodukte) *Sei $\beta \in BLF(V^*, V)$.*

i) Als Nullräume *von β bezeichnet man den Unterraum*

$$N_{V^*} := \{v^* \mid \forall\ v\colon\ \beta(v^*, v) = 0\}$$

von V^* und den analog definierten Unterraum N_V von V.

ii) β heißt nicht ausgeartet, *wenn die Nullräume trivial sind: $N_{V^*} = \{0_{V^*}\}$ und $N_V = \{0_V\}$. Solche nicht ausgearteten Bilinearformen bezeichnen wir auch oft mit $\langle - \mid - \rangle$, schreiben also $\langle v^* \mid v\rangle$ statt $\beta(v^*, v)$.*

iii) Ist $\langle - \mid - \rangle \in BLF(V^*, V)$ *(also nicht ausgeartet), dann heißen* V^* *und* V dual *bzgl.* $\langle - \mid - \rangle$, *die Bilinearform heißt dann* Skalarprodukt *und* $\langle v^* \mid v \rangle$ *das* skalare Produkt *von* v^* *und* v.

•

3.5.2 Beispiele

i) $\langle - \mid - \rangle \colon \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K} \colon (v^*, v) \mapsto \sum_{i=0}^{n-1} v_i^* v_i$ heißt das *Standardskalarprodukt* auf $\mathbb{K}^n$, $\mathbb{K}^n$ ist also diesbezüglich zu sich selbst dual.

ii) $\langle - \mid - \rangle \colon L(V) \times V \to \mathbb{K} \colon (\lambda, v) \mapsto \lambda(v)$ wird ebenfalls häufig herangezogen. Es zeigt, daß $L(V)$ dual ist zu V. Der Raum der Linearformen ist geradezu der *Prototyp* der zu V dualen Vektrorräume!

iii) Sind $\mathcal{B}^*, \mathcal{B}$ Basisfolgen von $V^*, V, m = \dim_{\mathbb{K}}(V^*)$, $n = \dim_{\mathbb{K}}(V)$, und $A \in \mathbb{K}^{m \times n}$, dann ist

$$\beta \colon V^* \times V \to \mathbb{K} \colon (v^*, v) \mapsto (v_0^*, \ldots, v_{m-1}^*) \cdot A \cdot \begin{pmatrix} v_0 \\ \vdots \\ v_{n-1} \end{pmatrix}$$

eine Bilinearform. Für $A := I_m$, die Einheitsmatrix, ergibt sich so das Standardskalarprodukt. Umgekehrt gilt für $\beta \in BLF(V^*, V)$ mit $B := (\beta(b_i^*, b_k))$:

$$\beta(v^*, v) = (v_0^*, \ldots, v_{m-1}^*) \cdot B \cdot \begin{pmatrix} v_0 \\ \vdots \\ v_{n-1} \end{pmatrix}$$

iv) $\beta \in BLF(V^*, V)$ ergibt, für $U^* \leq V^*, U \leq V$, die Einschränkung $\beta \downarrow U^* \times U \in BLF(U^*, U)$. Umgekehrt liefert $\beta' \in BLF(U^*, U)$, mit $f_1 \in \mathrm{Hom}_{\mathbb{K}}(V^*, U^*)$, $f_2 \in \mathrm{Hom}_{\mathbb{K}}(V, U)$ die Bilinearform

$$\beta \colon V^* \times V \to \mathbb{K} \colon (v^*, v) \mapsto \beta'(f_1(v^*), f_2(v)).$$

v) Ist $\beta \in BLF(V^*, V)$, dann ergibt sich mit deren Nullräumen wie folgt eine nicht ausgeartete Bilinearform:

$$\bar{\beta} \colon (V^*/N_{V^*}) \times (V/N_V) \to \mathbb{K} \colon (v^* + N_{V^*}, v + N_V) \mapsto \beta(v^*, v).$$

◇

3.5.3 Definition (orthogonal) *Sei* $\langle - \mid - \rangle \in BLF(V^*, V)$, $v^* \in V, v \in V$.

i) v^*, v *heißen zueinander* orthogonal, *kurz:* $v^* \perp v$, *wenn gilt* $\langle v^* \mid v \rangle = 0$.

ii) Als orthogonales Komplement *von* $U^* \leq V^*$ *bezeichnet man den Unterraum*

$$U^{*\perp} := \{v \in V \mid \forall\ u^* \in U^* \colon \langle u^* \mid v \rangle = 0\}$$

von V. *Analog ist das orthogonale Komplement von* $U \leq V$ *definiert, ein Unterraum von* V^*.

3.5.4 Beispiele

i) $V^{*\perp} = \{0_V\}, V^\perp = \{0_{V^*}\}$, da $\langle - \mid - \rangle$ nicht ausgeartet ist.

ii) Offenbar gilt $U^* \leq U^{*\perp\perp}$.

iii) $U_i \leq V \Rightarrow (U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$.

$\diamond$

3.5.5 Definition (duale Abbildungen) *Sind V^*, V und W^*, W Paare dualer $\mathbb{K}-$ Vektorräume, dann heißen $f \in Hom_{\mathbb{K}}(V, W)$ und $f^* \in Hom_{\mathbb{K}}(W^*, V^*)$ zueinander* dual, *wenn gilt:*

$$\langle f^*(w^*) \mid v \rangle = \langle w^* \mid f(v) \rangle.$$

3.5.6 Beispiele

i) $V^* := L(V)$ ist dual zu V. Ist außerdem W^* dual zu W bzgl. $\langle - \mid - \rangle$ sowie $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$, dann ist zu dieser Abbildung dual:

$$f^* : W^* \to L(V) : w^* \mapsto \langle w^* \mid f(-) \rangle.$$

ii) Sind V^*, V zueinander dual und ist $U \leq V_{\mathbb{K}}$, dann sind $V^*/U^\perp$ und U zueinander dual bzgl. (vgl. 4.1.3 v))

$$\langle - \mid - \rangle' : (V^*/U^\perp) \times U \to \mathbb{K} : (v^* + U^\perp, u) \mapsto \langle v^* \mid u \rangle.$$

Dual zur Einbettung $\iota_U : U \to V : u \mapsto u$ ist die Abbildung

$$\nu_{U^\perp} : V^* \to V^*/U^\perp : v^* \mapsto v^* + U^\perp.$$

3.5.7 Satz

i) Duale Abbildungen sind (gegebenenfalls) eindeutig bestimmt.

ii) Sind f^*, g^* zu $f, g \in Hom_{\mathbb{K}}(V, W)$, dann gilt

$$(f + g)^* = f^* + g^*, (\kappa \cdot f)^* = \kappa \cdot f^*.$$

iii) Sind f^*, g^* dual zu $f \in Hom_{\mathbb{K}}(U, V), g \in Hom_{\mathbb{K}}(V, W)$, dann ist

$$(g \circ f)^* = f^* \circ g^*.$$

iv) Sind f, f^* zueinander dual, dann gilt

$$\mathrm{Kern}(f^*) = \mathrm{Bild}(f)^\perp, \mathrm{Kern}(f) = \mathrm{Bild}(f^*)^\perp.$$

Beweis: i) gilt, da Skalarprodukte nicht ausgeartet sind, ii) und iii) sind leicht nachzurechnen. Zum Beweis von iv) bemerken wir:

$$w^* \in \mathrm{Kern}(f^*) \iff \forall\, v\colon\ 0 = \langle f^*(w^*) \mid v\rangle = \langle w^* \mid f(v)\rangle \iff w^* \in \mathrm{Bild}(f)^\perp.$$

Die zweite Identität $\mathrm{Kern}(f) = \mathrm{Bild}(f^*)^\perp$ ergibt sich ganz analog. □

3.5.8 Folgerungen *V^*, V und W^*, W seien Paare zueinander dualer $\mathbb{K}-$ Vektorräume, und zu $f \in Hom_{\mathbb{K}}(V, W)$ sei $f^* \in Hom_{\mathbb{K}}(W^*, V^*)$ dual. Dann gilt:*

i) $W^*/\mathrm{Bild}(f)^\perp = W^*/\mathrm{Kern}(f^*)$ *und* $\mathrm{Bild}(f)$ *sind dual bzgl.*

$$\langle w^* + \mathrm{Bild}(f)^\perp \mid f(v)\rangle := \langle w^* \mid f(v)\rangle.$$

ii) $\mathrm{Bild}(f^*)$ *und* $V/\mathrm{Bild}(f^*)^\perp = V/\mathrm{Kern}(f)$ *sind dual bzgl.*

$$\langle f^*(w^*) \mid v + \mathrm{Kern}(f)\rangle := \langle f^*(w^*) \mid v\rangle.$$

Beweis: 3.5.6 ii).

□

Das Paradebeispiel für ein Paar dualer Vektorräume ist $L(V), V$ wir fassen zusammen, was wir darüber bereits wissen und zeigen noch einiges darüber hinaus:

3.5.9 Satz *Zu V ist $L(V)$ dual bzgl.*

$$\langle - \mid - \rangle\colon L(V) \times V \to \mathbb{K}\colon (\lambda, v) \mapsto \lambda(v).$$

Zu einer linearen Abbildung $f \in Hom_{\mathbb{K}}(V, W)$ dual ist

$$f^*\colon L(W) \to L(V)\colon \mu \mapsto \mu \circ f.$$

Für die Kerne und Bilder dieser Abbildungen gilt:

$$\mathrm{Kern}(f) = \mathrm{Bild}(f^*)^\perp, \mathrm{Kern}(f^*) = \mathrm{Bild}(f)^\perp,$$

und analog natürlich auch

$$\mathrm{Bild}(f^*) = \mathrm{Kern}(f)^\perp, \mathrm{Bild}(f) = \mathrm{Kern}(f^*)^\perp.$$

Schließlich ist noch für orthogonale Komplemente folgendes richtig:

$$U^{\perp\perp} = U.$$

Beweis: Die Aussagen über die Dualität von V und $L(V)$ wie auch die Aussage über die Form der dualen Abbildung sind bereits bekannt. Wir wissen auch schon, daß

$$\mathrm{Kern}(f^*) = \mathrm{Bild}(f)^\perp,\ \mathrm{Kern}(f) = \mathrm{Bild}(f^*)^\perp.$$

Es bleibt also noch zu zeigen, daß

$$\mathrm{Bild}(f^*) = \mathrm{Kern}(f)^\perp,\ \mathrm{Bild}(f) = \mathrm{Kern}(f^*)^\perp.$$

Zum Beweis der ersten Gleichung bemerken wir, daß

$$\lambda \in \mathrm{Kern}(f)^{\perp} \iff \mathrm{Kern}(f) \subseteq \mathrm{Kern}(\lambda).$$

Nach dem Abbildungssatz ist dies äquivalent zur Existenz einer kommutativen Ergänzung $\mu \in L(W)$ des folgenden Diagramms:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "\lambda"'] & W \arrow[d, "\mu"] \\ & \mathbb{K} \end{tikzcd}$$

Hierfür gilt also $\lambda = \mu \circ f$ und damit $\lambda \in \mathrm{Bild}(f^*)$, was zu zeigen war.
Vor dem Beweis von $\mathrm{Bild}(f) = \mathrm{Kern}(f^*)^{\perp}$ beweisen wir $U^{\perp\perp} = U$, für $U \leq V$. Dazu betrachten wir die kanonische Projektion

$$\pi: V \to V/U : v \mapsto v + U.$$

Mit Hilfe des bereits Bewiesenen ergibt sich:

$$U = \mathrm{Kern}(\pi) = \mathrm{Bild}(\pi^*)^{\perp} = \mathrm{Kern}(\pi)^{\perp\perp} = U^{\perp\perp}.$$

Hiermit können wir jetzt den Beweis vervollständigen:

$$\mathrm{Kern}(f^*)^{\perp} = \mathrm{Bild}(f)^{\perp\perp} = \mathrm{Bild}(f).$$

□

Wir fragen nun nach Dimension und Basen von zueinander dualen Vektorräumen. Sind $\mathcal{B}, \mathcal{B}'$, endliche Basisfolgen für V und V', dann gilt offenbar

3.5.10 $$\mathrm{Hom}_{\mathbb{K}}(V, V') = \ll f_{ik} : b_j \mapsto \delta_{jk} b_i' \mid i \in m, j \in n \gg,$$

also insbesondere

3.5.11 $$\dim_{\mathbb{K}}(\mathrm{Hom}_{\mathbb{K}}(V, V')) = \dim_{\mathbb{K}}(V) \cdot \dim_{\mathbb{K}}(V').$$

3.5.12 Folgerung

i) $L(V) = \ll \lambda_i : b_j \mapsto \delta_{ij} \mid i \in m \gg$,

ii) $dim_{\mathbb{K}}(L(V)) = dim_{\mathbb{K}}(V)$.

Das legt die Vermutung nahe, daß endlichdimensionale zueinander duale Vektorräume V^*, V dieselbe Dimension haben und damit $V^* \simeq_{\mathbb{K}} L(V)$ gilt. Das ist tatsächlich richtig, weshalb $L(V)$ auch als *der* zu V duale Vektorraum bezeichnet werden kann. Genauer gilt:

3.5.13 Satz *Sind V^*, V dual bzgl. $\langle - \mid - \rangle$ und ist $dim_{\mathbb{K}}(V) \in \mathbb{N}$, dann gilt:*

i) $\varphi: V^* \simeq_{\mathbb{K}} L(V): v^* \mapsto \langle v^* \mid - \rangle$,

ii) $dim_{\mathbb{K}}(V^*) = dim_{\mathbb{K}}(V)$.

Beweis:

a) Offensichtlich ist $\varphi(v^*) \in L(V)$, und da $\langle - \mid - \rangle$ nicht ausgeartet ist, ist φ zudem injektiv.

b) Die Injektivität von φ ergibt $\dim_{\mathbb{K}}(V^*) \leq \dim_{\mathbb{K}}(V)$.

c) Betrachten von $\psi: V \to L(V^*): v \mapsto \langle - \mid v \rangle$ ergibt ganz analog $\dim_{\mathbb{K}}(V) \leq \dim_{\mathbb{K}}(V^*)$. Mit b) folgt also die Behauptung.

□

3.5.14 Definition (duale Basen) *Sind V^*, V dual bzgl. $\langle - \mid - \rangle$, ist $dim_{\mathbb{K}}(V) \in \mathbb{N}$ und sind $\mathcal{B}^*, \mathcal{B}$ Basisfolgen, dann heißen diese zueinander* dual, *wenn gilt*

$$\langle b_i^* \mid b_k \rangle = \delta_{ik} = \begin{cases} 1, & i{=}k, \\ 0, & \text{sonst.} \end{cases}$$

3.5.15 Beispiel
Die zur Basisfolge $\mathcal{B} = (b_1, \ldots, b_m)$ von V oben bereits angegebene Basisfolge

$$\mathcal{L} := (\lambda_1, \ldots, \lambda_m)$$

von $L(V)$ ist die zu $\mathcal{B}$ duale Basisfolge.

3.5.16 Satz *Ist $V_{\mathbb{K}}$ endlichdimensional, $\mathcal{B}$ Basisfolge, dann gibt es in jedem zu V dualen Vektorraum V^* eine zu $\mathcal{B}$ duale Basisfolge $\mathcal{B}^*$.*

Beweis: Ist $\mathcal{L}$ die zu $\mathcal{B}$ duale Basisfolge in ${}_{\mathbb{K}}L(V)$, so ergibt die Abbildung $\varphi: V^* \simeq_{\mathbb{K}} L(V)$ vermöge

$$b_i^* \in \varphi^{-1}(\{\lambda_i\})$$

eine zu $\mathcal{B}$ duale Basisfolge $\mathcal{B}^* = (b_1^*, \ldots, b_m^*)$:

$$\langle b_i^* \mid b_k \rangle = \langle b_i^* \mid - \rangle(b_k) = \varphi(b_i^*)(b_k) = \lambda_i(b_k) = \delta_{ik}.$$

□

3.5.17 Satz *Zueinander duale lineare Abbildungen zwischen endlichdimensionalen $\mathbb{K}-$Vektorräumen werden bzgl. dualer Basisfolgen durch zueinander transponierte Matrizen dargestellt.*

Beweis: V^*, V und W^*, W seien Paare endlichdimensionaler dualer $\mathbb{K}-$ Vektorräume mit zueinander dualen Basisfolgen $\mathcal{B}^*, \mathcal{B}$ und $\mathcal{C}^*, \mathcal{C}$. Weiter sei $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ mit $A = (a_{ik}) := M(\mathcal{C}, f, \mathcal{B})$, und dazu dual $f^* \in \mathrm{Hom}_{\mathbb{K}}(W^*, V^*)$ mit $B = (b_{ik}) := M(\mathcal{B}^*, f^*, \mathcal{C}^*)$. Dann gilt:

$$\langle f^*(c_k^*) \mid b_j\rangle = \sum_i b_{ik}\langle b_i^* \mid b_j\rangle = b_{jk}.$$

Wegen der Dualität ergibt sich andererseits:

$$\langle f^*(c_k^*) \mid b_j\rangle = \langle c_k^* \mid f(b_j)\rangle = \sum_i \langle c_k^* \mid c_k\rangle a_{ij} = a_{kj},$$

also folgt insgesamt: $b_{jk} = a_{kj}$, wie behauptet. □

3.5.18 Satz *Ist $U \leq V$, $dim_{\mathbb{K}}(V) \in \mathbb{N}$, dann gilt*

$$dim_{\mathbb{K}}(V) = dim_{\mathbb{K}}(U) + dim_{\mathbb{K}}(U^{\perp}).$$

Beweis: $L(V)$ und V sind zueinander dual, also auch $L(V)/U^{\perp}, U$, weshalb gilt

$$\dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}}(L(V)) - \dim_{\mathbb{K}}(U^{\perp}) = \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U^{\perp}).$$

□

3.5.19 Satz *Sind V^*, V dual bzgl. $\langle - \mid - \rangle$, und ist $\beta \in BLF(V^*, V)$ sowie $dim_{\mathbb{K}}(V) \in \mathbb{N}$, dann gibt es genau ein $f_\beta \in End_{\mathbb{K}}(V)$ mit*

$$\beta(v^*, v) = \langle v^* \mid f_\beta(v)\rangle.$$

Beweis:

i) Die Existenz von f_β :
Sei $\psi\colon V \simeq L(V^*)\colon v \mapsto \langle - \mid v\rangle$. Wegen $\beta(-, v) \in L(V^*)$ gibt es genau ein $v' \in V$ mit $\psi(v') = \beta(-, v)$, also ist

$$f_\beta\colon V \to V\colon v \mapsto v',\ \mathit{mit}\ \psi(v') = \beta(-, v),$$

wohldefiniert.

ii) f_β ist $\mathbb{K}-$ Endomorphismus:

$$f_\beta(v_1\kappa_1 + v_2\kappa_2) = (v_1\kappa_1 + v_2\kappa_2)' = v_1'\kappa_1 + v_2'\kappa_2,$$

letzteres wegen

$$\begin{aligned}\langle - \mid (v_1\kappa_1 + v_2\kappa_2)'\rangle &= \beta(-, v_1\kappa_1 + v_2\kappa_2) = \beta(-, v_1) + \beta(-, v_2)\kappa_2\\ &= \langle - \mid v_1'\rangle\kappa_1 + \langle - \mid v_2'\rangle\kappa_2 = \langle - \mid v_1'\kappa_1 + v_2'\kappa_2\rangle,\end{aligned}$$

woraus sich $(v_1\kappa_1+v_2\kappa_2)' = v_1'\kappa_1+v_2'\kappa_2$ ergibt, da $\langle - \mid - \rangle$ nicht ausgeartet ist.

iii) Die Eindeutigkeit:

$$\langle v^* \mid f_\beta(v)\rangle = \langle v^* \mid f(v)\rangle \Rightarrow \langle v^* \mid (f_\beta - f)(v)\rangle$$

ergibt $f_\beta = f$, da $\langle - \mid - \rangle$ nicht ausgeartet ist.

3.5.20 Definition *Die Menge $BLF(V^*, V)$ aller Bilinearformen β ist auf natürliche Weise sowohl $\mathbb{K}-$ Vektorraum.*

3.5.21 Folgerung $BLF(V^*, V) \simeq End_{\mathbb{K}}(V)$.

3.5.22 Anwendungen

i) Ist $f \in \mathrm{Hom}_{\mathbb{K}}(V, V')$, und sind V, V' endlichdimensional, dann heißt der Spaltenrang $Sr(A)$, für irgendeine Matrix A mit $M(\mathcal{B}', f, \mathcal{B})$, der *Rang* von f. Kurz: $Rg(f) := Sr(A)$. Wir wollen (nochmals) zeigen, daß der Spaltenrang dem Zeilenrang gleicht:

$$\begin{aligned}
Rg(f) &= \dim_{\mathbb{K}}(\mathrm{Bild}(f)) \\
&= \dim_{\mathbb{K}}(V/\mathrm{Kern}(f)) \\
&= \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(\mathrm{Kern}(f)) \\
&= \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(\mathrm{Bild}(f^*)^{\perp}) \\
&= \dim_{\mathbb{K}}(V) - (\dim_{\mathbb{K}}(V^*) - \dim_{\mathbb{K}}(\mathrm{Bild}(f^*)) \\
&= \dim_{\mathbb{K}}(\mathrm{Bild}(f^*)) \\
&= Rg(f^*).
\end{aligned}$$

Dies beweist, erneut, daß für Matrizen A gilt $Sr(A) = Zr(A)$, denn wir wissen ja, daß die duale Abbildung bzgl. der dualen Basis durch die Transponierte von A dargestellt wird.

ii) $Ax = b$ ist genau dann lösbar, wenn jede Lösung des *transponierten* homogenen Systems ${}^tA \cdot y = 0$ auf b senkrecht steht: $\sum y_i b_i = 0$.

$$Ax = b \text{ lösbar } \Leftrightarrow b \in \mathrm{Bild}(f) = \mathrm{Kern}(f^*)^{\perp}.$$

iii) Ein numerisches Beispiel zu ii):

$$A := \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 3 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 3}, b := \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} \in \mathbb{R}^3, also$$

$${}^tA = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 3 \\ 1 & -1 & 3 \end{pmatrix}.$$

Dann ist ${}_{\mathbb{R}}\langle \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \rangle = L_H({}^tA)$. Wegen

$$\langle \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \mid \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} \rangle = -6 + 4 + 1 \neq 0$$

ist das System also nicht lösbar.

3.6 Determinanten

Wir verallgemeinern jetzt den Begriff bilinear zu *multilinear.* Unser Ziel ist dabei insbesondere die Einführung der sogenannten Determinante. Der Grundkörper $\mathbb{K}$ wird weiterhin als kommutativ vorausgesetzt.

3.6.1 Definition (alternierend, symmetrisch, schiefsymmetrisch) *Seien V und V' $\mathbb{K}-$ Vektorräume, $n \in \mathbb{N}^*$, $f\colon V^n \to V'$, dann heißt f*

- multilinear: $\iff$ $f(v_0, \ldots, v_{i-1}, -, v_{i+1}, \ldots, v_{n-1}) \in Hom_{\mathbb{K}}(V, V')$,
- alternierend: $\iff$ $[\exists\ i \neq k\colon\ v_i = v_k \Rightarrow f(v_0, \ldots, v_{n-1}) = 0]$,
- symmetrisch: $\iff$ $\forall\ \pi \in S_n\colon\ f(v_0, \ldots, v_{n-1}) = f(v_{\pi(0)}, \ldots, v_{\pi(n-1)})$,
- schiefsymmetrisch: $\iff$ $f(v_0, \ldots, v_{n-1}) = sgn(\pi) f(v_{\pi(0)}, \ldots, v_{\pi(n-1)})$.

3.6.2 Beispiele

i) Für $n = 2$ ist z.B.

$$f\colon \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}\colon (\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}) \mapsto x_0 y_1 - x_1 y_0$$

bilinear, alternierend und schiefsymmetrisch. Falls char($\mathbb{K}$) $\neq 2$ ist, ist f nicht symmetrisch.

ii) Bei char($\mathbb{K}$) $= 2$ fallen die Begriffe symmetrisch und schiefsymmetrisch zusammen.

iii) Ein für die Anwendungen in der euklidischen Geometrie und Physik sehr wichtiges Beispiel ist (für $n = 2$) das sogenannte *Kreuzprodukt* im $\mathbb{R}^3$:

$$\times\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3\colon (\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \\ y_2 \end{pmatrix}) \mapsto \begin{pmatrix} x_1 y_2 - x_2 y_1 \\ x_2 y_0 - x_0 y_2 \\ x_0 y_1 - x_1 y_0 \end{pmatrix} =: x \times y,$$

es ist bilinear und schiefsymmetrisch. $\diamond$

3.6.3 Hilfssatz

i) Jede alternierende multilineare Abbildung ist schiefsymmetrisch.

ii) Bei char($\mathbb{K}$) $\neq 2$ sind die multilinearen schiefsymmetrischen Abbildungen alternierend.

Beweis:

i) $f: V^n \to V'$ sei alternierend und multilinear, $i \in n$, dann gilt:

$$\begin{aligned} 0 &= f(\ldots, v_{i-1}, v_i + v_{i+1}, v_i + v_{i+1}, v_{i+2}, \ldots) \\ &= f(v_0, \ldots, v_{n-1}) + f(\ldots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \ldots). \end{aligned}$$

Für die Transposition $\tau := (i, i+1)$ gilt also

$$f(v_0, \ldots, v_{n-1}) = \mathrm{sgn}(\tau) f(v_{\tau(0)}, \ldots, v_{\tau(n-1)}).$$

Da diese Transpositionen aber S_n erzeugen, und sgn ein Homormophismus ist, folgt daraus die Behauptung.

ii) ist klar. □

Von besonderer Bedeutung sind für uns nun gewisse *Multilinearformen,* also der Fall $V' = \mathbb{K}$, und dabei wiederum der Fall $n := \dim_{\mathbb{K}}(V)$.

3.6.4 Definition (Determinantenformen) *Ist* $n := dim_{\mathbb{K}} V \in \mathbb{N}^*$, *dann heißt* $\Delta: V^n \to \mathbb{K}$ Determinantenform *auf* V, *kurz:* $\Delta \in DF(V)$, *wenn* Δ *multilinear und alternierend ist.*

3.6.5 Beispiele $\Delta: V^n \to \mathbb{K}: (v_0, \ldots, v_{n-1}) \mapsto 0_{\mathbb{K}}$ heißt die *triviale* Determinantenform auf V. 3.6.2 i) ist eine nicht triviale Determinantenform, 3.6.2 iii) ist multilinear, alternierend und schiefsymmetrisch, aber keine Determinantenform, da weder $n = 3$ noch der Bildraum gleich $\mathbb{K}$ ist. ◇

3.6.6 Hilfssatz *Ist* $\Delta \in DF(V)$ *und* $V = \ll b_0, \ldots, b_{n-1} \gg$, *dann gilt für* $v_i = \sum_{\nu=0}^{n-1} v_{\nu,i} b_\nu \in V$:

$$\begin{aligned} \Delta(v_0, \ldots, v_{n-1}) &= \Delta(b_0, \ldots, b_{n-1}) \sum_{\pi \in S_n} sgn(\pi) v_{\pi(0),0} \cdots v_{\pi(n-1),n-1} \\ &= \Delta(b_0, \ldots, b_{n-1}) \sum_{\pi \in S_n} sgn(\pi) v_{0,\pi(0)} \cdots v_{n-1,\pi(n-1)}. \end{aligned}$$

Beweis: Wegen der Multilinearität von Δ folgt zunächst:

$$\Delta(v_0, \ldots, v_{n-1}) = \sum_{\varphi: n \to n} v_{\varphi(0),0} \cdots v_{\varphi(n-1),n-1} \Delta(b_{\varphi(0)}, \ldots, b_{\varphi(n-1)}).$$

Da Δ alternierend ist, ergibt das einmal

$$\Delta(v_0, \ldots, v_{n-1}) = \sum_{\pi \in S_n} v_{\pi(0),0} \cdots v_{\pi(n-1),n-1} \Delta(b_{\pi(0)}, \ldots, b_{\pi(n-1)}),$$

und — weil alternierende Multilinearformen schiefsymmetrisch sind — auch noch

$$= \Delta(b_0, \ldots, b_{n-1}) \sum_{\pi \in S_n} \mathrm{sgn}(\pi) v_{\pi(0),0} \cdots v_{\pi(n-1),n-1},$$

insgesamt also die erste der behaupteten Gleichungen. Schließlich gilt, wegen der Kommutativität von $\mathbb{K}$,

$$v_{\pi(0),0}\cdots v_{\pi(n-1),n-1} = v_{0,\pi^{-1}(0)}\cdots v_{n-1,\pi^{-1}(n-1)},$$

so daß mit $\mathrm{sgn}(\pi) = \mathrm{sgn}(\pi^{-1})$ auch die zweite der behaupteten Gleichungen folgt, denn $\pi \mapsto \pi^{-1}$ ist eine Bijektion auf jeder Gruppe, insbesondere also auch auf S_n. □

3.6.7 Folgerungen

i) Jede Determinantenform $\Delta \in DF(V)$ ist durch den Wert $\Delta(b_0,\ldots,b_{n-1})$ auf irgendeiner Basisfolge $\mathcal{B}$ von V vollständig bestimmt.

ii) Nichttriviale Determinantenformen sind Vielfache voneinander.

iii) Zu jeder Basisfolge $\mathcal{B}$ von V gibt es höchstens ein $\Delta \in DF(V)$ mit $\Delta(b_0,\ldots,b_{n-1}) = 1$.

□

Daß es zu jeder Basisfolge $\mathcal{B} := (b_0,\ldots,b_{n-1})$ von V *genau eine* Determinantenform Δ mit $\Delta(b_0,\ldots,b_{n-1}) = 1$ gibt, folgt damit aus

3.6.8 Hilfssatz *Die Gleichung*

$$\Delta_{\mathcal{B}}(\ldots,\sum_{\nu=0}^{n-1} v_{\nu,i}b_\nu,\ldots,) := \sum_{\pi\in S_n} sgn(\pi)v_{\pi(0),0}\ldots v_{\pi(n-1),n-1}$$

definiert eine nicht triviale Determinantenform $\Delta_{\mathcal{B}}$ auf V. Für diese gilt

$$\Delta_{\mathcal{B}}(b_0,\ldots,b_{n-1}) = 1.$$

Beweis: Die Abbildung $\Delta_{\mathcal{B}}$ ist offensichtlich multilinear, sie ist auch alternierend: Sei $i \neq k$ und $v_i = v_k$. Für die Transposition $\tau := (ik) \in S_n$ gilt dann zunächst einmal die folgende Zerlegung von S_n in Rechtsnebenklassen von $A_n = \{\pi \in S_n \mid \mathrm{sgn}(\pi) = 1\}$:

$$S_n = A_n \cup A_n\tau.$$

Damit ergibt sich, wegen $v_i = v_k$,

$$\begin{aligned}\Delta_{\mathcal{B}}(v_0,\ldots,v_{n-1}) &= \Delta_{\mathcal{B}}(v_{\tau(0)},\ldots,v_{\tau(n-1)})\\ &= \sum_{\pi\in S_n}\mathrm{sgn}(\pi)v_{\pi(0),\tau(0)}\ldots v_{\pi(n-1),\tau(n-1)}\end{aligned}$$

$$= \sum_{\rho\in A_n}(\mathrm{sgn}(\rho)v_{\rho(0),\tau(0)}\ldots v_{\rho(n-1),\tau(n-1)} + \mathrm{sgn}(\rho\tau)v_{\rho\tau(0),\tau(0)}\ldots v_{\rho\tau(n-1),\tau(n-1)})$$

$$= \sum_{\rho\in A_n}(\mathrm{sgn}(\rho)v_{\rho(0),0}\ldots v_{\rho(n-1),n-1} - \mathrm{sgn}(\rho)v_{\rho(0),0}\ldots v_{\rho(n-1),n-1}).$$

Jeder dieser Summanden ist aber, ganz unabhängig von $\mathrm{char}(\mathbb{K})$, gleich Null. □

Insgesamt ist damit folgendes bewiesen:

3.6.9 Folgerung *Ist $\mathcal{B} = (b_0, \ldots, b_{n-1})$ Basisfolge für V, dann gilt*

i) Die eindeutig bestimmte Determinantenform $\Delta_{\mathcal{B}}$ mit $\Delta_{\mathcal{B}}(b_0, \ldots, b_{n-1}) = 1$ wird definiert durch

$$\Delta_{\mathcal{B}}(\ldots, \sum_{\nu=0}^{n-1} v_{\nu,i} b_\nu, \ldots) := \sum_{\pi \in S_n} sgn(\pi) v_{\pi(0),0} \ldots v_{\pi(n-1),n-1} =$$

$$\sum_{\pi \in S_n} sgn(\pi) v_{0,\pi(0)} \ldots v_{n-1,\pi(n-1)}.$$

ii) Jede andere Determinantenform $\Delta \in DF(V)$ ist Vielfache von $\Delta_{\mathcal{B}}$:

$$\Delta(v_0, \ldots, v_{n-1}) = \Delta(b_0, \ldots, b_{n-1}) \cdot \Delta_{\mathcal{B}}(v_0, \ldots, v_{n-1}).$$

□

Es soll nun gezeigt werden, wie mit Hilfe einer Determinantenform $\Delta \in DF(V)$ einem Endomorphismus $f \in End_{\mathbb{K}}(V)$ ein Skalar, die sogenannte Determinante $\det(f)$ von f zugeordnet werden kann, die sehr nützliche Eigenschaften hat.

3.6.10 Hilfssatz *$\Delta \in DF(V)$ sei nicht trivial, $f \in End_{\mathbb{K}}(V)$, dann gilt:*

i) $\Delta_f(v_0, \ldots, v_{n-1}) := \Delta(f(v_0), \ldots, f(v_{n-1}))$ definiert eine Determinantenform Δ_f.

ii) Der Faktor κ_f aus $\Delta_f = \kappa_f \cdot \Delta$ ist unabhängig von der Wahl der (nicht trivialen) Determinantenform Δ.

Beweis: i) ist klar.

ii) Sei $\Delta' = \lambda \cdot \Delta, \lambda \neq 0$, dann folgt aus

$$\Delta'(b_0, \ldots, b_{n-1}) = \lambda \cdot \Delta(b_0, \ldots, b_{n-1})$$

die linke Seite der folgenden Sequenz von Gleichungen:

$$\begin{aligned} \Delta'_f(b_0, \ldots, b_{n-1}) &= \lambda \cdot \Delta_f(b_0, \ldots, b_{n-1}) = \lambda \cdot \kappa_f \cdot \Delta(b_0, \ldots, b_{n-1}) \\ &= \kappa_f \cdot \Delta'(b_0, \ldots, b_{n-1}), \end{aligned}$$

also gilt auch für Δ': $\Delta'_f = \kappa_f \cdot \Delta'$. □

3.6.11 Definition (die Determinante eines Endomorphismus) *Der Faktor*

$$\det(f) := \frac{\Delta_f(b_0, \ldots, b_{n-1})}{\Delta(b_0, \ldots, b_{n-1})},$$

$\mathcal{B}$ irgendeine Basisfolge von V, Δ irgendeine nicht triviale Determinantenfolge auf V, $f \in End_{\mathbb{K}}(V)$, heißt die Determinante *von $f \in End_{\mathbb{K}}(V)$.*

Für diese Zahl gilt also die Gleichung:

3.6.12 $$\det(f)\cdot\Delta = \Delta_f.$$

3.6.13 Satz *Für $f, g \in End_{\mathbb{K}}(V), n := dim_{\mathbb{K}}(V), \kappa \in \mathbb{K}$ gilt:*

i) $f = \kappa \cdot id_V \Longrightarrow \det(f) = \kappa^n$,

ii) $f \in Aut_{\mathbb{K}}(V) \iff \det(f) \neq 0$,

iii) $\det(g \circ f) = \det(g)\cdot\det(f) = \det(f \circ g)$,

iv) $f \in Aut_{\mathbb{K}}(V) \Longrightarrow \det(f^{-1}) = \det(f)^{-1}$,

v) $V = V_1 \oplus V_2, f_i := f \downarrow V_i, f_i(V_i) \subseteq V_i, i = 1,2 \Longrightarrow \det(f) = \det(f_1)\cdot\det(f_2)$.

Beweis: Sei $\Delta \in DF(V)$ nicht trivial, $\mathcal{B} = (b_0, \ldots, b_{n-1})$ eine Basisfolge.

i) Wegen $f = f = \kappa \cdot id_V$ gilt

$$\Delta(f(b_0), \ldots, f(b_{n-1})) = \Delta(\kappa b_0, \ldots, \kappa b_{n-1}) = \kappa^n \Delta(b_0, \ldots, b_{n-1}).$$

ii) Es ist

$$\begin{aligned} f \in Aut_{\mathbb{K}}(V) &\iff (f(b_0,), \ldots, f(b_{n-1})) \text{ Basisfolge} \\ &\iff 0 \neq \Delta(f(b_0), \ldots, f(b_{n-1})) = \det(f)\cdot\Delta(b_0, \ldots, b_{n-1}) \\ &\iff \det(f) \neq 0. \end{aligned}$$

iii) Wir haben

$$\begin{aligned} \det(g\circ f)\Delta(b_0, \ldots, b_{n-1}) &= \Delta(g(f(b_0)), \ldots, g(f(b_{n-1}))) \\ &= \det(g)\Delta(f(b_0), \ldots, f(b_{n-1})) \\ &= \det(g)\cdot\det(f)\cdot\Delta(b_0, \ldots, b_{n-1}) \end{aligned}$$

iv) $f \in Aut_{\mathbb{K}}(V)$ impliziert $f^{-1} \in Aut_{\mathbb{K}}(V)$, die Determinanten dieser beiden Abbildungen sind also, nach ii), von Null verschieden, und wir können wie folgt schließen:

$$1 =_{i)} \det(f \circ f^{-1}) =_{iii)} \det(f)\cdot\det(f^{-1})^{-1}.$$

v) Für $g_1 := f_1 \oplus id_{v_2} : v = v_1 + v_2 \mapsto f_1(v_1) + v_2$ und die entsprechend definierte Abbildung $g_2 := id_{v_1} \oplus f_2$ gilt $f = g_1 \circ g_2$, also, nach iii): $\det(f) = \det(g_1)\cdot\det(g_2)$, so daß $\det(f_i) = \det(g_i)$ zu zeigen bleibt. Ist deshalb $V_2 = \ll b_r, \ldots, b_{n-1} \gg$, so definiert (falls $V_2 \neq 0$, andernfalls ist nichts zu zeigen)

$$\Delta_1(v_0, \ldots, v_{r-1}) := \Delta(v_0, \ldots, v_{r-1}, b_r, \ldots, b_{n-1})$$

offenbar eine nicht triviale Determinantenform auf V_1. Für sie gilt

$$\begin{aligned}
\det(f_1)\Delta_1(b_0,\dots,b_{r-1}) &= \Delta_1(f(b_0),\dots,f(b_{r-1}))\\
&= \Delta(f(b_0),\dots,f(b_{r-1}),b_r,\dots,b_{n-1})\\
&= \Delta(g_1(b_0),\dots,g(b_{n-1})) = \det(g_1)\cdot\Delta(b_0,\dots,b_{n-1})\\
&= \det(g_1)\cdot\Delta_1(b_0,\dots,b_{r-1}),
\end{aligned}$$

also $\det(f_1)=\det(g_1)$. Analog folgt $\det(f_2)=\det(g_2)$. □

Mit Hilfe der Determinante eines Endomorphismus kann nun eine Determinante für jede quadratische Matrix definiert werden. Dies geschieht so, daß für jede Basisfolge $\mathcal{B}$ von V gilt:

$$\det(A)=\det(f), \text{ falls } A=M(\mathcal{B},f,\mathcal{B}).$$

Sei deshalb $A=(a_{ik})=M(\mathcal{B},f,\mathcal{B})\in\mathbb{K}^{n\times n}$. Dann gilt

$$\det(f)\cdot\Delta_{\mathcal{B}}(b_0,\dots,b_{n-1})=\Delta_{\mathcal{B}}(f(b_0),\dots,f(b_{n-1}))$$

$$=\Delta_{\mathcal{B}}(\dots,\sum_{i=0}^{n-1}a_{ik}b_i,\dots)=\sum_{\pi\in S_n}\operatorname{sgn}(\pi)a_{\pi(0),0}\dots a_{\pi(n-1),n-1}.$$

Wir sehen, daß dieser Ausdruck nicht von $\mathcal{B}$ und der hierzu passend zu wählenden Abbildung f abhängt, definieren deshalb wie folgt:

3.6.14 Definition (die Determinante einer Matrix) *Ist* $A=(a_{ik})\in\mathbb{K}^{n\times n}$, *dann sei*

$$\det(A):=\sum_{\pi\in S_n}sgn(\pi)\prod_{i=0}^{n-1}a_{\pi(i),i}=\sum_{\pi\in S_n}sgn(\pi)\prod_{i=0}^{n-1}a_{i,\pi(i)}$$

die Determinante *der Matrix* A.

3.6.15 Beispiele

i) $\det(a_{00})=a_{00}$,

ii) $\det\begin{pmatrix}a_{00}&a_{01}\\a_{10}&a_{11}\end{pmatrix}=a_{00}a_{11}-a_{01}a_{10}$,

iii) $\det\begin{pmatrix}a_{00}&a_{01}&a_{02}\\a_{10}&a_{11}&a_{12}\\a_{20}&a_{21}&a_{22}\end{pmatrix}=a_{00}a_{11}a_{22}-a_{10}a_{01}a_{22}+a_{20}a_{11}a_{02}$

$$-a_{00}a_{21}a_{12}+a_{10}a_{21}a_{02}+a_{20}a_{01}a_{12}.$$

3.6.16 Folgerungen *Sind* $A,B\in\mathbb{K}^{n\times n}$, $\kappa\in\mathbb{K}$, *dann gilt:*

i) *Ist $\mathcal{B}$ Basisfolge, $f \in End_{\mathbb{K}}(V), A = M(\mathcal{B}, f, \mathcal{B})$, so ist*

$$\det(f) = \det(A).$$

ii) $\det(\kappa \cdot I_n) = \kappa^n$,

iii) *A ist genau dann invertierbar, wenn* $\det(A) \neq 0$,

iv) *Die Abbildung einer Matrix auf ihre Determinante ist ein* Homomorphismus *(ein Ringhomomorphismus auf dem Ring $\mathbb{K}^{n\times n}$ der n-reihigen Matrizen, ein Gruppenhomomorphismus auf der Gruppe $GL_n(\mathbb{K})$ der invertierbaren n-reihigen Matrizen):*

$$\det(A \cdot B) = \det(A) \cdot \det(B) = \det(B \cdot A),$$

v) *Ist B invertierbar, dann gilt*

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

d. h. die Determinante ist invariant *unter der Transformation $A \mapsto BAB^{-1}$.*

vi) *A ist invertierbar* $\Longrightarrow \det(A^{-1}) = \det(A)^{-1}$,

vii) $\det(A) = \det({}^tA)$.

(Die Punkte i) bis v) folgen umnittelbar aus 3.6.13, der sechste Punkt ergibt sich aus $\prod_\nu a_{\pi(\nu),\nu} = \prod_\nu a_{\nu,\pi^{-1}(\nu)}$.) Für die konkrete Berechnung der Determinante einer Matrix A verwendet man zweckmäßigerweise nicht die oben angegebene explizite Formel, da diese aus $n!$ Summanden besteht. Vielmehr geht man in der Regel zu einer Matrix A' über, die dieselbe Determinante besitzt, bei der aber sichtlich weniger Summanden zu berechnen sind, da die anderen verschwinden. Die möglichen Umformungen ergeben sich aus 3.6.16 mit der Tatsache, daß Determinantenformen multilinear und alternierend sind:

3.6.17 Satz *Die Determinante* $\det(A)$ *ist multilinear, alternierend und schiefsymmetrisch in den Zeilen und in den Spalten von A. Es gilt also insbesondere:*

i) *Das Vertauschen zweier Zeilen oder Spalten ändert am Wert der Determinante nur das Vorzeichen.*

ii) *Entsteht A' aus A durch Multiplizieren einer Zeile oder einer Spalte mit κ, dann gilt* $\det(A') = \kappa \cdot \det(A)$.

iii) *Addition eines Vielfachen einer Zeile (Spalte) zu einer anderen (!) Zeile (Spalte) ändert den Wert der Determinante nicht.*

iv) *Ist der Rang von A kleiner als die Zeilenzahl, dann gilt* $\det(A) = 0$.

3.6.18 Beispiele
i)

$$\det\begin{pmatrix} a_{00} & & * \\ & \ddots & \\ 0 & & a_{n-1,n-1}\end{pmatrix} = \det\begin{pmatrix} a_{00} & & 0 \\ & \ddots & \\ * & & a_{n-1,n-1}\end{pmatrix} = a_{00}\cdots a_{n-1,n-1}$$

folgt unmittelbar aus der (wichtigen) Bemerkung, daß jeder Summand $\prod_i a_{\pi(i),i}$ aus jeder Zeile und Spalte von A genau ein Element enthält, bei oberen oder unteren Dreiecksmatrizen ist also maximal ein einziger Summand (der zu $\pi = \mathrm{id}$) von Null verschieden.
ii) Die *Vandermondesche Determinante:* Für paarweise verschiedene $\kappa_i \in \mathbb{K}^*$ gilt

$$\det\begin{pmatrix} 1 & 1 & \dots & 1 \\ \kappa_0 & \kappa_1 & \dots & \kappa_{n-1} \\ \kappa_0^2 & \kappa_1^2 & \dots & \kappa_{n-1}^2 \\ \vdots & \vdots & & \vdots \\ \kappa_0^{n-1} & \kappa_1^{n-1} & \dots & \kappa_{n-1}^{n-1}\end{pmatrix} = \det\begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & \kappa_1-\kappa_0 & \dots & \kappa_{n-1}-\kappa_0 \\ 0 & \kappa_1^2-\kappa_0\kappa_1 & \dots & \kappa_{n-1}^2-\kappa_0\kappa_{n-1} \\ \vdots & \vdots & & \vdots \\ 0 & \kappa_1^{n-1}-\kappa_0\kappa_1^{n-1} & \dots & \kappa_{n-1}^{n-1}-\kappa_0\kappa_{n-1}^{n-1}\end{pmatrix},$$

da man von der $(i+1)-$ten Zeile das κ_0-fache der $i-$ten Zeile abziehen kann, ohne daß sich die Determinante ändert. Benutzt man jetzt die explizite Formel, so sieht man, daß (wegen der Nullen in der ersten Spalte) gilt

$$= \det\begin{pmatrix} \kappa_1-\kappa_0 & \kappa_2-\kappa_0 & \dots & \kappa_{n-1}-\kappa_0 \\ \kappa_1^2-\kappa_0\kappa_1 & \kappa_2^2-\kappa_0\kappa_2 & \dots & \kappa_n^2-\kappa_0\kappa_{n-1} \\ \vdots & & & \\ \kappa_1^{n-1}-\kappa_0\kappa_1^{n-1} & \kappa_2^{n-1}-\kappa_0\kappa_2^{n-1} & \dots & \kappa_{n-1}^{n-1}-\kappa_0\kappa_{n-1}^{n-1}\end{pmatrix}$$

$$= (\kappa_1-\kappa_0)(\kappa_2-\kappa_0)\dots(\kappa_{n-1}-\kappa_0)\det\begin{pmatrix} 1 & 1 & \dots & 1 \\ \kappa_1 & \kappa_2 & \dots & \kappa_{n-1} \\ \vdots & \vdots & & \vdots \\ \kappa_1^{n-2} & \kappa_2^{n-2} & \dots & \kappa_{n-1}^{n-2}\end{pmatrix},$$

so daß per Induktion nach n folgt:

$$\det\begin{pmatrix} 1 & 1 & \dots & 1 \\ \kappa_0 & \kappa_1 & \dots & \kappa_{n-1} \\ \vdots & \vdots & & \vdots \\ \kappa_0^{n-1} & \kappa_1^{n-1} & \dots & \kappa_{n-1}^{n-1}\end{pmatrix} = \prod_{0\le i<j\le n-1}(\kappa_j-\kappa_i),$$

denn

$$\det(1) = 1.$$

$\diamond$

Zur Vorbereitung eines Satzes über die rekursive Berechnung von Determinanten definieren wir

3.6.19 Definition (Cofaktoren) *Sei $A = (a_{ik}) \in \mathbb{K}^{n\times n}$ und a_i der i-te Zeilenvektor, a'_k der k-te Spaltenvektor dieser Matrix. Wir definieren mit Hilfe der Basisfolge $\mathcal{E} = (e_0, \ldots, e_{n-1})$ aus den Einheitsvektoren den* Cofaktor *von a_{ik} durch*

$$\mathrm{Cof}(a_{ik}) := \det(A_{ik}),$$

wobei A_{ik} aus A durch Ersetzen der $i-$ten Zeile durch den Einheitsvektor e_k und der k-ten Spalte durch den Einheitsvektor e_i entsteht:

$$A_{ik} = \begin{pmatrix} & & & 0 & & & \\ & * & & \vdots & & * & \\ & & & 0 & & & \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ & & & 0 & & & \\ & * & & \vdots & & * & \\ & & & 0 & & & \end{pmatrix}.$$

Es ist leicht einzusehen, daß diese Zahl auch mit Hilfe der Basisfolge $\mathcal{E}$ aus den Einheitsvektoren beschrieben werden kann, d.h. daß gilt

3.6.20 Bemerkung *Die Cofaktoren lassen sich wie folgt auch mit Hilfe der Determinantenform $\Delta_{\mathcal{E}}$ ausdrücken:*

$$\begin{aligned} \mathrm{Cof}(a_{ik}) &= \det(A_{ik}) \\ &= \Delta_{\mathcal{E}}(a_0, \ldots, a_{i-1}, e_k, a_{i+1}, \ldots, a_{n-1}) \\ &= \Delta_{\mathcal{E}}(a'_0, \ldots, a'_{k-1}, e_i, a'_{k+1}, \ldots, a'_{n-1}). \end{aligned}$$

Weiterhin sei die adjungierte Matrix *zu A definiert durch*

$$\mathrm{Ad}(A) := {}^t(\mathrm{Cof}(a_{ik})).$$

3.6.21 Satz $A \cdot \mathrm{Ad}(A) = \det(A) \cdot I_n$.

Beweis:

$$\begin{aligned} (A \cdot \mathrm{Ad}(A))_{ik} &= \sum_\nu a_{i\nu}\mathrm{Cof}(a_{k\nu}) \\ &= \Delta_{\mathcal{E}}(a_0, \ldots, a_{k-1}, \sum_\nu a_{i\nu}e_\nu, a_{k+1}, \ldots, a_{n-1})) \\ &= \Delta_{\mathcal{E}}(a_0, \ldots, a_{k-1}, a_i, a_{k+1}, \ldots, a_{n-1})) \\ &= \delta_{ik} \det(A). \end{aligned}$$

□

3.6.22 Folgerung *Ist A invertierbar, dann gilt*

$$A^{-1} = \frac{\mathrm{Ad}(A)}{\det(A)}.$$

3.6.23 Beispiel

$$A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \Rightarrow \mathrm{Ad}(A) = {}^t(\mathrm{Cof}(a_{ik})) = {}^t\begin{pmatrix} \det\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} & \det\begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix} \\ \det\begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} & \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix}$$

$$= {}^t\begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} \Rightarrow A^{-1} = \frac{1}{4-6}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}.$$

Als Anwendung auf lineare Gleichungssysteme mit *regulärer,* d.h. invertierbarer Koeffizientenmatrix A ergibt sich

3.6.24 Die Cramersche Regel *Ist $A \in \mathbb{K}^{n\times n}$ invertierbar, dann gilt für die Komponenten x_i der eindeutig bestimmten Lösung von $Ax = b$:*

$$x_i = \frac{1}{\det(A)} \cdot \det\begin{pmatrix} a_{00} & \dots & a_{0,i-1} & b_1 & a_{0,i+1} & \dots & a_{0,n-1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n-1,0} & \dots & a_{n-1,i-1} & b_n & a_{n-1,i+1} & \dots & a_{n-1,n-1} \end{pmatrix}.$$

Beweis:
Aus $x = A^{-1}b = \frac{1}{\det(A)}\mathrm{Ad}(A)\cdot b$ ergibt sich

$$x_i = \frac{1}{\det(A)} \cdot (\sum_\nu \mathrm{Cof}(a_{\nu i})b_\nu)).$$

□

Weiter erhalten wir jetzt die angekündigte Rekursionsformel

3.6.25 Der Laplacesche Entwicklungssatz *Ist $A \in \mathbb{K}^{n\times n}$, dann gilt, für jedes $i, k \in n$:*

$$\det(A) = \sum_\nu a_{i\nu}\det(A_{i\nu}) = \sum_\nu a_{\nu k}\det(A_{\nu k}),$$

d.h. man kann $\det(A)$ *sowohl* nach der $i-$ten Zeile *als auch* nach der $k-$ten Spalte entwickeln.

Beweis: Es gilt offenbar

$$\begin{aligned} \det(A) &= \Delta_{\mathcal{E}}(a'_0,\dots,a'_{n-1}) = \sum_\nu a_{\nu k}\Delta_{\mathcal{E}}(a'_0,\dots,a'_{k-1},e_\nu,a'_{k+1},\dots,a'_{n-1}) \\ & \sum_\nu a_{\nu k}\mathrm{Cof}(a_{k\nu}) = \sum_\nu a_{\nu k}\det(A_{\nu k}). \end{aligned}$$

Entsprechendes gilt für die Entwicklung nach der $i-$ten Zeile. □
Dies ist ein Rekursionssatz, denn es gilt

3.6.26 $$\det(A_{ik}) = (-1)^{i+k}\det(A'_{ik}),$$

wenn A'_{ik} die Matrix ist, die aus A durch *Streichen der* $i-$*ten Zeile und der* $k-$*ten Spalte* entsteht. Diese Gleichung folgt durch Anwenden elementarer Zeilen- und Spaltenumformungen.
Anschließend sei noch als Verallgemeinerung angeführt:

3.6.27 *Die vorstehenden Überlegungen lassen sich auf endlich erzeugbare freie Moduln (das sind Moduln mit endlichen Basen, die alle dieselbe Ordnung haben) über kommutativen Ringen übertragen.*

3.7 Eigenwerte und Eigenvektoren

Wir wollen jetzt lineare Endomorphismen durch Matrizen besonders übersichtlicher Gestalt (u. a. mit möglichst vielen Nullen) beschreiben, was durch die Auswahl geeigneter Basen geschieht. Da wir dazu den Determinantenbegriff benutzen werden, bleibt nach wie vor $\mathbb{K}$ als *Körper,* also als kommutativ vorausgesetzt.
Zu einer besonders einfachen Form der $f \in \mathrm{End}_{\mathbb{K}}(V)$ darstellenden Matrix verhelfen gegebenenfalls natürlich solche Basisvektoren, die von f auf ein Vielfaches von sich selbst abgebildet werden. Wir bemerken zunächst, daß für solche Vektoren folgendes gilt:

3.7.1 $f(v) = \kappa \cdot v \iff (f - \kappa \cdot \mathrm{id}_V)(v) = 0 \iff v \in \mathrm{Kern}(f - \kappa \cdot \mathrm{id}_V)$.

Wir führen deshalb die folgenden Bezeichnungen ein:

3.7.2 Definition (Eigenwerte, Eigenvektoren, Eigenräume) *Ist* $f \in \mathrm{End}_{\mathbb{K}}(V), v \in V, \kappa \in \mathbb{K}$, *dann heißt* κ Eigenwert *von* f *genau dann, wenn*

$$\mathrm{Kern}(f - \kappa \cdot \mathrm{id}_V) \neq \{0_V\}.$$

Die von Null verschiedenen Elemente des Kerns von $f - \kappa \cdot \mathrm{id}_V$ *heißen ggf.* Eigenvektoren *zum Eigenwert* κ, *und als* Eigenraum *von* f *zum Eigenwert* κ *bezeichnen wir den gesamten Kern:*

$$E(\kappa) := \{v \in V \mid f(v) = \kappa v\} = \mathrm{Kern}(f - \kappa \cdot \mathrm{id}_V).$$

•

Falls die Dimension von V endlich ist, gibt es weitere dazu äquivalente Formulierungen, denn wir können die Determinante von f heranziehen:

3.7.3 Hilfssatz *Ist* $f \in \mathrm{End}_{\mathbb{K}}(V)$, $dim_{\mathbb{K}}(V) \in \mathbb{N}^*, \kappa \in \mathbb{K}$, *dann sind äquivalent:*

1. κ *ist Eigenwert von* f,
2. $\mathrm{Kern}(f - \kappa \cdot \mathrm{id}_V) \neq \{0_v\}$,
3. $\det(f - \kappa \cdot \mathrm{id}_V) = 0$,
4. $f - \kappa \cdot \mathrm{id}_V$ *ist nicht injektiv,*
5. $f - \kappa \cdot \mathrm{id}_V$ *ist nicht surjektiv,*
6. $\exists\ v \in V \backslash \{0_V\}\colon\ f(v) = \kappa \cdot v$.

□

Die Gleichung $\det(f - \kappa \cdot \mathrm{id}_V) = 0$ zeigt, daß die Eigenwerte von f als Nullstellen eines Polynoms aufgefaßt werden können. Wir wollen dies präzisieren und beginnen dazu mit

3.7.4 Definition (Formale Potenzreihen, Polynome, Polynomabbildungen) $(R, +, \cdot)$ *sei ein kommutativer Ring mit Einselement.*

- *Wir erinnern zunächst an den Ring* $R^{\mathbb{N}}$ *der* formalen Potenzreihen *über* R *in einer* Unbestimmten x. *Die Elemente* $p \in R^{\mathbb{N}}$ *schreibt man gern in der Form*
$$p = \sum_{n=0}^{\infty} p_n x^n, \text{ wobei } p_n := p(n).$$
Dieser Ring wird meist mit $R[\![x]\!]$ *bezeichnet, um die Bezeichnung der* Unbestimmten x *hervorzuheben. Er enthält — auch das wurde bereits erwähnt — den Teilring*
$$R[x] := \{p \in R^{\mathbb{N}} \mid p(n) = 0, \text{ für fast alle } n\}$$
der Polynome *über* R *in einer Unbestimmten* x.

- *Ist* $p \in R[x]^*$ *und* $n \in \mathbb{N}$ *maximal mit* $p(n) \neq 0$, *dann heißt* $r_n = p(n)$ *der* Leitkoeffizient *von* p *und* n *der* Grad *von* p. *Dem Nullpolynom wird kein Grad zugeordnet.*

- *Zu* $p := \sum r_n x^n \in R[x]$ *gehört die* Polynomabbildung
$$P \colon R \to R \colon \rho \mapsto \sum r_n \rho^n.$$
Umgekehrt heißt $f \in R^R$ polynomiale Abbildung, *wenn es* $p \in R[x]$ *gibt mit* $P = f$, *d.h.* $P(\rho) = f(\rho)$, *für alle* $\rho \in R$. *Falls* $\mathrm{Char}(R) \neq 0$ *ist* P *nicht notwendig eindeutig bestimmt, z.B. ist* $P = 0$, *falls* $p_i = x + x^2 \in \mathbb{Z}_2[x]$.

- $\rho \in R$ *heißt* Wurzel *von* p, *wenn* ρ Nullstelle *von* P *ist, d. h. wenn* $P(\rho) = 0$ *gilt.*

•

Ist jetzt $A \in R^{n\times n}$, dann ist $A - x \cdot I_n \in R[x]^{n\times n}$, und nach 3.6.16, Punkt v), gilt für jede invertierbare Matrix T über $R[x]$:

3.7.5 $$\det(A - xI_n) = \det(TAT^{-1} - xI_n).$$

Das ermöglicht die folgende Definition:

3.7.6 Definition (charakteristisches Polynom, charakteristische Koeffizienten) *Ist* $f \in \mathrm{End}_{\mathbb{K}}(V)$, $n := dim_{\mathbb{K}}(V) \in \mathbb{N}^*$, $\mathcal{B}$ *eine Basisfolge für* V, $A := M(\mathcal{B}, f, \mathcal{B})$, *dann heißt*
$$p_f := p_A := \det(A - xI_n) = \sum \alpha_i x^i$$
das charakteristische Polynom *von* f *bzw. von* A, *die* a_i *heißen die* charakteristischen Koeffizienten *von* f *bzw.* A. •

3.7.7 Beispiel Für

$$A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

gilt

$$p_A = \det \begin{pmatrix} 1-x & 2 \\ 3 & 4-x \end{pmatrix} = x^2 - 5x - 2.$$

◇

3.7.8 Hilfssatz *Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$, $n := dim_{\mathbb{K}}(V) \in \mathbb{N}^* A = M(\mathcal{B}, f, \mathcal{B})$, dann gilt für die charakteristischen Koeffizienten von f bzw. A :*

$$\alpha_0 = \det(A),\ \alpha_{n-1} = (-1)^{n-1} \sum_i \alpha_{ii},\ \alpha_n = (-1)^n.$$

□

Aus der Definition des charakteristischen Polynoms ergibt sich direkt:

3.7.9 Folgerung *Eigenwerte von $f \in \mathrm{End}_{\mathbb{K}}(V)$ bzw. $A = M(\mathcal{B}, f, \mathcal{B})$ sind — bei endlichdimensionalem V — gegebenenfalls (!) genau die Wurzeln (in $\mathbb{K}$!) des charakteristischen Polynoms $p_f = p_A$, also die Nullstellen der zugehörigen Polynomabbildung $P_f = P_A$.* □

Deshalb liefern u.a. auch Hilfsmittel aus der Analysis, nämlich Aussagen über Nullstellen von Polynomabbildungen, Existenzsätze für Eigenwerte. So sagt z.B. der *Fundamentalsatz der Algebra* — der meist in der Vorlesung über Funktionentheorie bewiesen wird — aus, daß jede Polynomabbildung $P: \mathbb{C} \to \mathbb{C}, p$ vom Grad n, n Nullstellen (die aber nicht paarweise verschieden sein müssen) in $\mathbb{C}$ hat. Das liefert, mit 4.3.9:

3.7.10 Satz *Jeder lineare Endomorphismus eines endlichdimensionalen Vektorraums $V \neq \{0_V\}$ über $\mathbb{C}$ besitzt Eigenwerte (und damit auch Eigenvektoren).*

□

Das gilt für $\mathbb{K} := \mathbb{R}$ dagegen nicht, z.B. hat das Polynom $p := 1 + x^2$ keine Wurzeln in $\mathbb{R}$, die lineare Abbildung $f \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^2)$ mit

$$M(\mathcal{E}, f, \mathcal{E}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

hat also *keine* Eigenwerte. Immerhin gilt noch

3.7.11 Satz *Ist $n := dim_{\mathbb{R}}(V) \in \mathbb{N}^*, f \in \mathrm{End}_{\mathbb{R}}(V)$, dann gilt:*

1. *Ist n ungerade, dann besitzt $\mathbb{R}$ Eigenwerte, und zwar mindestens einen positiven (negativen), falls $\det(f)$ positiv (negativ) ist.*

2. *Ist n gerade und $\det(f)$ negativ, dann besitzt f mindestens einen positiven und einen negativen Eigenwert.*

□

Beweis: Bei ungeradem n ist $a_n = (-1)^n = -1$, und nach Sätzen aus der Analysis ergibt sich, daß $P_f(x)$ mit $x \to -\infty$ gegen ∞ geht, mit $x \to \infty$ dagegen nach $-\infty$. Wegen der Stetigkeit von P_f folgt daraus mit dem Zwischenwertsatz die Behauptung. Analog folgt für gerade n die zweite Behauptung, da dann $\alpha_0 = P_f(0) = \det(f)$. □

Kehren wir wieder zu dem allgemeinen Fall zurück.

3.7.12 Satz *Jede Folge $(v_0, \ldots, v_{m-1})$ von Eigenvektoren zu paarweise verschiedenen Eigenwerten $\kappa_0, \ldots, \kappa_{m-1}$ von $f \in \mathrm{End}_{\mathbb{K}}(V)$ ist linear unabhängig.*

Beweis per Induktion nach m :

I $m = 1$: Die Folge (v_0) ist linear unabhängig, da $v_0 \neq 0$.

II $m \to m+1$: Wir schließen indirekt. Wäre $(v_0, \ldots, v_m)$ linear abhängig, dann gäbe es eine nicht triviale lineare Relation zwischen den Elementen der Folge. Ohne Einschränkung der Allgemeinheit können wir annehmen, sie sei von der Form $v_m = \sum_0^{m-1} \lambda_i v_i$, wobei nicht alle λ_i gleich Null sind. Daraus würde sich aber folgendes ergeben:

$$\begin{aligned} 0 &= f(v_m) - \kappa_m \cdot v_m \\ &= \sum_0^{m-1} \lambda_i f(v_i) - \kappa_m \sum_0^{m-1} \lambda_i v_i \\ &= \sum_0^{m-1} \lambda_i \underbrace{(\kappa_i - \kappa_m)}_{\neq 0} v_i, \end{aligned}$$

ein Widerspruch zur Induktionsannahme, der linearen Unabhängigkeit von $(v_0, \ldots, v_{m-1})$. □

3.7.13 Folgerung *Ist f ein linearer Endomorphismus von V, $dim_{\mathbb{K}}(V) \in \mathbb{N}^*$, mit $dim_{\mathbb{K}}(V)$ paarweise verschiedenen Eigenwerten, dann besitzt V eine Basisfolge $\mathcal{B}$ aus Eigenvektoren. f wird also bezüglich dieser Basisfolge durch eine Diagonalmatrix $M(\mathcal{B}, f, \mathcal{B})$ beschrieben, in deren Hauptdiagonale die Eigenwerte von f stehen.*

□

3.7.14 Satz *Ist $V = U \oplus W, n = dim_{\mathbb{K}}(V) \in \mathbb{N}^*, f \in \mathrm{End}_{\mathbb{K}}(V), f(U) \subseteq U, \pi_W \colon V \to W \colon u + w \mapsto w$ (die kanonische Projektion von V auf W), dann gilt: Für $g := \pi_W \circ f \downarrow W$ haben wir*

$$p_f = p_{f \downarrow U} \cdot p_g.$$

□

Beweis: Wir wählen eine Basisfolge $\mathcal{B}$ von V, die an die Zerlegung "angepaßt" ist, d.h. $\mathcal{B} = (b_0, \ldots, b_{n-1})$, mit

$$U = \ll b_0, \ldots, b_{m-1} \gg, W = \ll b_m, \ldots, b_{n-1} \gg .$$

Dann gilt für $A := M(\mathcal{B}, f, \mathcal{B})$ und $\mathcal{B}_1 = (b_1, \ldots, b_m), \mathcal{B}_2 = (b_{m+1}, \ldots, b_n)$:

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}, \ mit\ B = M(\mathcal{B}, f \downarrow U, \mathcal{B}_1),$$

sowie (mit einer Nullmatrix 0 aus n Zeilen und m Spalten):

$$\begin{aligned} M(\mathcal{B}_2, g, \mathcal{B}_2) &= M(\mathcal{B}_2, \pi_w, \mathcal{B}) \cdot M(\mathcal{B}, f \downarrow W, \mathcal{B}_2) \\ &= (0 \mid I_{n-m}) \cdot \begin{pmatrix} C \\ D \end{pmatrix} = D. \end{aligned}$$

Damit erhalten wir:

$$p_f = \det(A - xI_n) = \det(B - xI_m) \cdot \det(D - xI_{n-m}) = p_{f|U} \cdot p_g.$$

□

3.7.15 Folgerung *Ist U Unterraum von V mit $f(U) \subseteq U$, dann teilt $p_{f\downarrow U}$ das Polynom p_f.* □

Heißt $\kappa \in \mathbb{K}$ ein *r−facher* Eigenwert von f, wenn $(x-\kappa)^r$ Teiler von p_f ist, und verstehen wir unter der *Vielfachheit* von κ das maximale dieser r, dann ergibt sich:

3.7.16 Satz *Die Dimension des Eigenraums von $\kappa \in \mathbb{K}$ ist höchstens gleich der Vielfachheit von κ als Eigenwert.*

Beweis: Man sieht sehr leicht, daß $f \downarrow E(\kappa) = \kappa \cdot \mathrm{id}_{E(\kappa)}$ gilt, also

$$p_{f\downarrow E(\kappa)} = (\kappa - x)^{dim(E(\kappa))}.$$

Dieses Polynom teilt aber p_f, in dem der Faktor $\kappa - x$ genau mit dem Exponenten r steckt. □

Der Grad des Produkts zweier Polynome $\neq 0$ ist die Summe der Grade der Faktoren. Also ist die Summe der Vielfachheiten der Eigenwerte von f höchstens $\dim_{\mathbb{K}}(V)$. Das, zusammen mit 4.3.16, ergibt:

3.7.17 Satz *Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $n := dim_{\mathbb{K}}(V) \in \mathbb{N}^*$, dann sind äquivalent:*

- *f läßt sich durch eine Diagonalmatrix beschreiben.*
- *V hat eine Basis aus Eigenvektoren.*
- *V ist direkte Summe der Eigenräume von f.*

- *p_f ist Produkt von Linearfaktoren (:=Polynome vom Grad 1) und die Vielfachheiten der Eigenwerte gleichen den Dimensionen der zugehörigen Eigenräume.*

□

Will man eine Matrix auf Diagonalisierbarkeit untersuchen, hat man sich also zuerst das charakteristische Polynom vorzunehmen. Läßt es sich in Linearfaktoren zerlegen, hat man noch die Dimensionen $n - Rg(A - \kappa E)$ der Eigenräume zu ermitteln und mit den Vielfachheiten zu vergleichen.

3.7.18 Beispiel

1. Die Matrix
$$A := \begin{pmatrix} -1 & 4 & 6 \\ 0 & -1 & 0 \\ 0 & 5 & 2 \end{pmatrix}$$
hat das charakteristische Polynom
$$\det(A - xE) = \det \begin{pmatrix} -1-x & 4 & 6 \\ 0 & -1-x & 0 \\ 0 & 5 & 2-x \end{pmatrix} = -(1+x)^2(2-x),$$
also die Eigenwerte $\kappa_0 = -1$ und $\kappa_1 = 2$. Wegen
$$A - \kappa_0 I_3 = \begin{pmatrix} 0 & 4 & 6 \\ 0 & 0 & 0 \\ 0 & 5 & 3 \end{pmatrix},$$
ist die Dimension des Eigenraums, also $3 - Rg(A - \kappa_1 E)$, gleich 1 und damit kleiner als die Vielfachheit des Eigenwerts κ_0.. Die Matrix A besitzt demnach keine Basis aus Eigenvektoren.

2. Für die Matrix
$$A := \begin{pmatrix} -1 & 4 & 6 \\ 0 & 2 & 0 \\ 0 & 5 & 3 \end{pmatrix}$$
dagegen gilt $p_A = -(1+x)(2-x)(3-x)$. Die Eigenwerte sind also $\kappa_0 = -1$, $\kappa_1 = 2$ und $\kappa_2 = 3$. Nach 4.3.12 existiert also eine Basis aus Eigenvektoren.

3. Einen Eigenvektor b_i zum Eigenwert κ_i erhält man durch Lösen des Gleichungssystems $(A - \kappa_i I_n)b_i = 0$, für unser Beispiel erhält man beispielsweise (die Eigenvektoren sind nur bis auf Zahlenfaktoren eindeutig bestimmt!):
$$b_0 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, b_1 = \begin{pmatrix} 26/3 \\ -1 \\ 5 \end{pmatrix}, b_2 = \begin{pmatrix} -9 \\ 0 \\ -6 \end{pmatrix}.$$

◇

3.8 Vektorräume mit innerem Produkt

In diesem Paragraphen sei stets V ein $\mathbb{R}$−Vektorraum, ohne daß dies noch gesondert erwähnt wird. Wir betrachten nun gewisse Skalarprodukte auf V, die die Einführung von Begriffen wie Länge, Winkel usw. erlauben.

3.8.1 Definition (inneres Produkt) *$\langle - \mid - \rangle: V^2 \to \mathbb{R}$ sei Skalarprodukt (d.h. nicht ausgeartete Bilinearform, vgl. 3.5.1). Dann heißt $\langle - \mid - \rangle$* inneres Produkt *wenn gilt:*

i) *$\langle - \mid - \rangle$ ist* symmetrisch:

$$\langle v \mid w \rangle = \langle w \mid v \rangle,$$

ii) *$\langle - \mid - \rangle$ ist* positiv definit:

$$\langle v \mid v \rangle \geq 0 \ \wedge \ [\langle v \mid v \rangle = 0 \Rightarrow v = 0_V].$$

In diesem Fall heißt das Paar $(V, \langle - \mid - \rangle)$ ein Raum mit innerem Produkt. *Genau die endlichdimensionalen Vektorräume mit innerem Produkt heißen* euklidische Vektorräume.

iii) Die entsprechende Funktion

$$\| - \|: V \to \mathbb{R}: v \mapsto \sqrt{\langle v \mid v \rangle}$$

heißt die Norm. *Ist $\|v\| = 1$, dann heißt v* Einheitsvektor.

iv) u und w aus V heißen orthogonal, *kurz: $u \perp w$, wenn*

$$\langle u \mid w \rangle = 0.$$

v) Unterräume $U, W \leq$ von V heißen orthogonal, *kurz: $U \perp W$, wenn gilt:*

$$\forall \ u \in U, w \in W: \langle u \mid w \rangle = 0.$$

•

3.8.2 Beispiele

1. $\mathbb{R}^n$ mit $\langle v \mid w \rangle := \sum_{i=1}^n v_i w_i$.
2. Der Raum aller stetigen $f: [0,1] \to \mathbb{R}$ ist

$$\langle f \mid g \rangle := \int_0^1 f(t) g(t) dt.$$

◇

3.8.3 Hilfssatz *In einem Raum mit innerem Produkt sind Teilmengen paarweise orthogonaler Vektoren $\neq 0_V$ linear unabhängig.* □

Beweis: Ist T eine solche Teilmenge, $v_0, \ldots, v_{r-1} \in T$, mit $\sum_0^{r-1} \kappa_i v_i = 0$, dann gilt:

$$\forall\ j \in r\colon\ 0_{\mathbb{R}} = \langle 0_V \mid v_j \rangle = \sum_{i=0}^{r-1} \kappa_i \langle v_i \mid v_j \rangle = \kappa_j \langle v_j \mid v_j \rangle.$$

(Die erste Gleichung folgt aus der Tatsache, daß $v \mapsto \langle v \mid w \rangle$ eine lineare Abbildung ist.) Wegen $\langle v_j \mid v_j \rangle \neq 0$ folgt daraus $\kappa_j = 0$. □

Zwischen dem inneren Produkt und der zugehörigen Norm bestehen u.a. diese Zusammenhänge:

3.8.4 Hilfssatz *Ist* $(V, \langle - \mid - \rangle)$ *Raum mit innerem Produkt,* $v, w \in V$, *dann gilt:*

i) $\langle v \mid w \rangle = \frac{1}{2}(\|v+w\|^2 - \|v\|^2 - \|w\|^2)$

ii) $\langle v \mid w \rangle^2 \leq \|v\|^2 \|w\|^2$ (Cauchy- Schwarzsche Ungleichung),
und dabei gilt Gleichheit genau dann, wenn $\dim_{\mathbb{R}} \langle v, w \rangle \leq 1$.

iii) $\|v+w\| \leq \|v\| + \|w\|$ (Dreiecksungleichung)

Beweis:

i) $\|v+w\|^2 = \langle v+w \mid v+w \rangle = 2\langle v \mid w \rangle + \langle v \mid v \rangle + \langle w \mid w \rangle$ ergibt die erste Behauptung.

ii) Zum Beweis der Ungleichung verwenden wir die Tatsache, daß für beliebiges $r \in \mathbb{R}$ gilt

$$0 \leq \|v + rw\|^2 = \|v\|^2 + r^2 \|w\|^2 + 2r\langle v \mid w \rangle.$$

Für $w := 0_V$ gilt die Behauptung, im anderen Fall können wir r wie folgt wählen:

$$r := -\frac{\langle v \mid w \rangle}{\|w\|^2},$$

was folgendes ergibt:

$$0 \leq \|v\|^2 + \frac{\langle v \mid w \rangle^2}{\|w\|^2} - 2\frac{\langle v \mid w \rangle^2}{\|w\|^2} = \frac{\|v\|^2 \|w\|^2 - \langle v \mid w \rangle^2}{\|w\|^2},$$

also $0 \leq \|v\|^2 \|w\|^2 - \langle v \mid w \rangle^2$.

Ist dabei $\dim_{\mathbb{R}}(\langle v, w \rangle) \leq 1$, so gilt o.B.d.A.: $v = rw$, mit geeignetem $r \in \mathbb{R}$, und damit $\langle v \mid w \rangle^2 = \|w\|^4 = \|v\|^2 \|w\|^2$, es gilt also die behauptete Gleichheit.

Ist umgekehrt $\langle v \mid w \rangle^2 = \|v\|^2 \|w\|^2$, so gilt in der obigen Ungleichung das Gleichheitszeichen, also

$$0 = \|v + rw\|^2 = \|v\|^2 + r^2 \|w\|^2 + 2r\langle v \mid w \rangle,$$

zumindest für das angegebene $r \in \mathbb{R}$. Da $\langle - \mid - \rangle$ positiv definit ist, gilt $v + rw = 0$, also $\dim_{\mathbb{R}}(\langle v, w \rangle) \leq 1$.

iii) $\|v+w\| = \sqrt{\|v\|^2 + \|w\|^2 + 2\langle v \mid w\rangle} \leq_{ii)} \sqrt{\|v\|^2 + \|w\|^2 + 2\|v\|\,\|w\|} = \|v\| + \|w\|$.

□

3.8.5 Folgerung *Ist $(V, \langle - \mid - \rangle)$ Raum mit innerem Produkt, $v, w \in V \backslash \{0_V\}$, dann gilt:*

$$-1 \leq \frac{\langle v \mid w\rangle}{\|v\|\|w\|} \leq 1,$$

es gibt also genau ein $\omega \in [0, \pi] \subseteq \mathbb{R}$ mit:

$$\cos(\omega) = \frac{\langle v \mid w\rangle}{\|v\|\|w\|}.$$

□

Wir sind demnach zur Einführung der folgenden Bezeichnung berechtigt:

3.8.6 Definition (Winkel) *Das ω aus 3.8.5 heiße der* Winkel *zwischen v und w. (Offenbar ist $\omega \in \{0, \pi\}$ bei linearer Abhängigkeit von $\{v, w\}_{\neq}$, und gleich $\pi/2$ bei $v \perp w$.)* •

Aus 3.8.5 folgt mit 3.8.4 i) noch:

3.8.7 Der Cosinussatz *Es gilt*

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos(w).$$

□

Der Winkelbegriff der anschaulichen Geometrie ist allerdings komplizierter, da es sich um *orientierte* Winkel handelt. Wir führen deshalb den Begriff der Orientierung ein. Ist $n := \dim_{\mathbb{R}} V \in \mathbb{N}^*$, so gibt es nichttriviale $\Delta \in DF(V)$ und auf diesen wegen 3.6.7 ii) folgende Äquivalenzrelation:

3.8.8 $$\Delta \sim \Delta' :\Longleftrightarrow \exists\, r \in \mathbb{R}_{>0}\colon\ \Delta = r\Delta'.$$

3.8.9 Definition (orientierter Vektorraum)

i) Die beiden Klassen von "$\sim$" heißen die Orientierungen *von V. V zusammen mit einer dieser beiden Klassen $[\Delta]_\sim$ heißt* orientierter Vektorraum.

ii) Ist $(V, [\Delta]_\sim)$ orientierter Vektorraum $\mathcal{B} = (b_1, \ldots, b_n)$ eine Basisfolge von V, dann heißt $\mathcal{B}$ positive Basisfolge *genau dann, wenn $\Delta(b_1, \ldots, b_n) > 0$ gilt.*

•

3.8.10 Beispiel

$$\begin{aligned}\mathbb{R}^4 &= \ll e_0+e_1, e_0+e_1+e_2, e_0+e_1+e_2+e_3, e_0-e_1+e_3 \gg \\ &= \ll e_0-3e_2, e_1+e_3, e_1-e_0-e_3, e_1 \gg .\end{aligned}$$

Diese Basen sind stets verschieden orientiert, da die beiden Determinanten mit diesen Basisfolgen als Spalten, also

$$\det\begin{pmatrix}1&1&1&1\\1&1&1&-1\\0&1&1&0\\0&0&1&1\end{pmatrix}=-2,\ \det\begin{pmatrix}1&0&-1&0\\0&1&1&1\\-3&0&0&0\\0&1&-1&0\end{pmatrix}=3$$

verschiedene Vorzeichen haben. ◇

Jeder $\mathbb{R}$−Isomorphismus f induziert eine Orientierung im Bildraum, nämlich diejenige, bzgl. der f orientierungserhaltend ist:

3.8.11 Definition (orientierungserhaltend) *Sind* $(V,[\Delta]_\sim)$ *und* $(V',[\Delta']_\sim)$ *orientierte Vektorräume,* $f \in Iso_{\mathbb{R}}(V,V')$, *dann heißt* f orientierungserhaltend, *wenn gilt*

$$(\Delta')_f \in [\Delta]_\sim.$$

•

3.8.12 Satz *Ein* $\mathbb{R}$−*Automorphismus* f *eines orientierten* $\mathbb{R}$−*Vektorraums ist genau dann orientierungserhaltend, wenn gilt* $\det(f) > 0$ *ist.*

Beweis: Ist $[\Delta]_\sim$ die Orientierung, dann gilt $\Delta_f = \det(f)\cdot\Delta$, also $\Delta_f \in [\Delta]_\sim$ genau dann, wenn $\det(f) > 0$. □

3.8.13 Definition (Orthonormalbasis) *Ist wieder* $(V,\langle - \mid - \rangle)$ *ein Raum mit innerem Produkt, dann heißt eine Basis* B Orthonormalbasis *(kurz: ON-Basis) von* V, *wenn sie aus paarweise orthogonalen Vektoren der Norm 1 besteht.* •

3.8.14 Satz *Jeder Vektorraum mit innerem Produkt und abzählbarer Basis besitzt ON-Basen.*

Beweis: Das *Orthonormalisierungsverfahrens von E. Schmidt:*

Aus einer beliebigen vorgegebenen Basisfolge $\mathcal{B} = (b_0, b_1 \ldots, b_{n-1}, \ldots)$ wird eine ON-Basisfolge $(e_0, e_1, \ldots, e_{n-1}, \ldots)$ konstruiert. Dazu zeigen wir per Induktion: Für $n \in \mathbb{N}$ und $U := \ll b_0, \ldots, b_{n-1} \gg$ existiert eine ON-Basisfolge $\mathcal{E}' = (e_0, \ldots, e_{n-1})$.

I $n = 0$: Die leere Menge erzeugt den Nullraum, und die leere Folge ist offenbar eine ON-Basisfolge.

II $n \to n+1$: Entsprechend der Induktionsannahme sei $(e_0, \ldots, e_{n-1})$ eine ON-Basisfolge für $\ll b_0, \ldots, b_{n-1} \gg$. Wir setzen jetzt

$$a_n := b_n - \sum_{j=0}^{n-1} \langle b_n \mid e_j \rangle e_j.$$

Offenbar gilt damit

$$\ll b_0, \ldots, b_n \gg = \ll e_0, \ldots, e_{n-1}, a_n \gg,$$

zudem sind die Vektoren $e_0, \ldots, e_{n-1}, a_n$ paarweise orthogonal:

$$\langle a_n \mid e_i \rangle = \langle b_n \mid e_i \rangle - \sum_{1}^{n} \langle b_n \mid e_j \rangle \underbrace{\langle e_j \mid e_i \rangle}_{=\delta_{ij}} = 0.$$

Also ist $\mathcal{E}' := (e_0, \ldots, e_n)$, mit

$$e_n := \frac{a_n}{||a_n||}$$

eine ON-Basisfolge von $U = \ll b_0, \ldots, b_n \gg$.

$\square$

3.8.15 Beispiel V sei der Raum der Polynomabbildungen $f: \mathbb{R} \to \mathbb{R}$ vom Grad ≤ 2 mit dem inneren Produkt

$$\langle f \mid g \rangle := \int_{-1}^{1} f(x) g(x) dx.$$

Aus der Basisfolge $(1, x, x^2)$ ergibt sich die ON-Folge

$$e_0 = \frac{1}{\sqrt{2}}, e_1 = \sqrt{\frac{3}{2}} x, e_2 = \sqrt{\frac{45}{8}} (x^2 - \frac{1}{3}).$$

$e_0, e_1, \ldots$ sind, bis auf Normierungsfaktoren, die sogenannten *Legendre-Polynome.*
$\diamond$

3.8.16 Definition (orthogonale Matrizen) $A \in \mathbb{R}^{n \times n}$ *heißt* orthogonal, *wenn* ${}^tA = A^{-1}$. $\bullet$

3.8.17 Satz *Ist V endlichdimensional, dann wird der Übergang von einer ON-Basis zu einer anderen stets durch eine orthogonale Matrix beschrieben. Umgekehrt ergibt ein Basiswechsel, der durch eine orthogonale Matrix beschrieben wird, aus einer ON-Basis wieder eine solche.*

Beweis: Nachrechnen. □

Ist V euklidischer Vektorraum, $U \leq V$, dann kann man eine an U angepaßte ON-Basisfolge finden, es gilt also

3.8.18 $$V = U \oplus U^{\perp}.$$

Demgemäß kann $v \in V$ zerlegt werden in $v = u + w, u \in U, w \in U^{\perp}$. Ist $(u_0, \ldots, u_{m-1})$ eine ON-Basis von U, dann gilt

3.8.19 $$u = \sum_0^{m-1} \langle u \mid u_i \rangle u_i.$$

u heißt die *Orthogonalprojektion* von v auf U, $||w||$ die *Distanz* von v zum Unterraum U. Wegen $||v||^2 = ||u||^2 + ||w||^2$ gilt

3.8.20 $$||v|| \geq ||u|| \text{ (Besselsche Ungleichung)}.$$

Ist $\Delta \in DF(V)$ nicht trivial, $v_i, w_i \in V$, $(e_0, \ldots, e_{n-1})$ eine ON-Basisfolge, so ist zunächst $\det(\langle v_i \mid w_k \rangle)$ multilinear und alternierend sowohl in den v_i als auch in den w_k. Es gibt deshalb $r \in \mathbb{R}$ mit

3.8.21 $$r \cdot \det(\langle v_i \mid w_k \rangle) = \Delta(v_0, \ldots, v_{n-1})\Delta(w_0, \ldots, w_{n-1}).$$

Wegen $\det(\langle e_i \mid e_k \rangle) = \det(I_n) = 1$ gilt insbesondere $r = \Delta(e_0, \ldots, e_{n-1})^2 > 0$. Für $\Delta' := \pm\Delta/\sqrt{r}$ gilt also

3.8.22 $$\Delta'(v_0, \ldots, v_{n-1})\Delta'(w_0, \ldots, w_{n-1}) = \det(\langle v_i \mid w_k \rangle).$$

3.8.23 Definition (normierte Determinantenform) *Determinantenformen auf Räumen mit inneren Produkt, die 3.8.22 genügen, heißen* normiert.

•

$\pm\Delta/\sqrt{r}$ sind die beiden einzigen normierten Determinantenformen, genau eine von ihnen repräsentiert die Orientierung. Auf jedem euklidischen Vektorraum mit Orientierung gibt es also genau eine normierte Determinantenform, die die Orientierung repräsentiert.

Mit dieser die Orientierung repräsentierenden normierten Determinantenform Δ_0 können wir auch Winkel orientieren: In einem orientierten 2-dimensionalen Vektorraum V mit innerem Produkt $\langle - \mid - \rangle$ gilt nach 3.8.22:

3.8.24 $$\Delta_0(v, w)^2 = det\begin{pmatrix} \langle v \mid v \rangle & \langle v \mid w \rangle \\ \langle w \mid v \rangle & \langle w \mid w \rangle \end{pmatrix} = ||v||^2||w||^2 - \langle v \mid w \rangle^2.$$

Für $v \neq 0_V \neq w$ ergibt dies

$$\frac{\langle v \mid w \rangle^2}{||v||^2||w||^2} + \frac{\Delta_0(v, w)^2}{||v||^2||w||^2} = 1.$$

Es gibt also genau im $\omega \in (-\pi, \pi] \subseteq \mathbb{R}$ mit

3.8.25 $$\cos(\omega) = \frac{\langle v \mid w \rangle}{||v||||w||} \text{ und } \sin(\omega) = \frac{\Delta_0(v, w)}{||v||||w||}.$$

3.8.26 Definition (orientierter Winkel) *Das ω aus 3.8.25 heißt der* orientierte Winkel *zwischen v und w. Er wechselt das Vorzeichen mit der Orientierung $[\Delta_0]_\sim$ und auch bei Vertauschen von v und w.* •

3.8.27 Definition (die Gramsche Determinante) *Ist $(V, \langle - \mid - \rangle)$ Vektorraum mit innerem Produkt und sind $v_0, \ldots, v_{m-1} \in V$, dann heißt*

$$G(v_0, \ldots, v_{m-1}) := \det(\langle v_i \mid v_k \rangle)$$

die Gramsche Determinante *der Vektorenfolge $(v_0, \ldots, v_{m-1})$.* •

3.8.28 Hilfssatz *Für die Gramsche Determinante gilt*

i) $G(v_0, \ldots, v_{m-1}) \geq 0$

ii) $G(v_0, \ldots, v_{m-1}) = 0 \iff dim_{\mathbb{R}}(\langle v_0, \ldots, v_{m-1} \rangle) < m.$

Beweis: Die erste Behauptung folgt aus

$$G(v_0, \ldots, v_{m-1}) = \Delta_0(v_0, \ldots, v_{m-1})^2 \geq 0.$$

Zum Beweis der zweiten Behauptung argumentieren wir wie folgt:
$\dim(\langle v_0, \ldots, v_{m-1} \rangle) < m$ ergibt offensichtlich $G(v_0, \ldots, v_{m-1}) = 0$. Ist andererseits $\dim(\langle v_0, \ldots, v_{m-1} \rangle) = m$, dann ist $\ll v_0, \ldots, v_{m-1} \gg$ Unterraum mit innerem Produkt, also auch mit normierten Determinantenformen Δ. Die Ungleichung

$$G(v_0, \ldots, v_{m-1}) = \Delta(v_0, \ldots, v_{m-1})^2 > 0$$

ergibt dann die Behauptung. □

3.8.29 Definition (Parallelpiped,Volumen) *Ist $(V, \langle - \mid - \rangle)$ Raum mit innerem Produkt, $(v_0, \ldots, v_{r-1})$ linear unabhängig, Δ eine normierte Determinantenform auf $\ll v_0, \ldots, v_{r-1} \gg$, dann heißt*

$$\{v \mid v = \sum_0^{r-1} \rho_i v_i, 0 \leq \rho_i \leq 1\}$$

das von den v_i aufgespannte $r-$dimensionale Parallelpiped, *und*

$$V(v_0, \ldots, v_{r-1}) := |\Delta(v_0, \ldots, v_{r-1})|$$

heißt dessen Volumen. •

Offenbar gilt dafür:

3.8.30 Satz

1. $V(v_0, \ldots, v_{r-1})^2 = \det(\langle v_i \mid v_k \rangle)$,

2. *Bei $r = 2$ haben wir $V(v_0, v_1)^2 = ||v_0||^2 ||v_1||^2 \sin^2(\omega)$, also*

$$V(v_0, v_1) = ||v_0|| ||v_1|| |\sin(\omega)|.$$

□

Auf euklidischen Vektorräumen definierte lineare Funktionale lassen sich mit Hilfe des inneren Produkts beschreiben:

3.8.31 Satz *Der Rieszsche Darstellungssatz: Ist* $(V, \langle - \mid - \rangle)$ *ein euklidischer Vektorraum,* $f \in L(V)$*, dann gibt es genau ein* $v \in V$ *mit*

$$f = \langle v \mid - \rangle.$$

□

Beweis: Die Abbildung $\varphi\colon v \mapsto \langle v \mid - \rangle$ ist injektiv, und, wegen der Endlichkeit der Dimension $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{R}}(L(V))$ sogar bijektiv, also ein Isomorphismus.

□

Das benutzen wir jetzt zu folgender Definition:

3.8.32 Definition (Kreuzprodukt) *Sei* $(V, \langle - \mid - \rangle)$ *ein dreidimensionaler, euklidischer Vektorraum,* Δ *die normierte, eine Orientierung repräsentierende Determinantenform,* $v, w \in V$*. Dann ist* $\Delta(v, w, -) \in L(V)$*, nach dem Rieszschen Darstellungssatz gibt es also genau ein* $u \in V$ *mit*

$$\Delta(v, w, -) = \langle u \mid - \rangle.$$

Dieses $u \in V$ *heißt das* Kreuzprodukt *von* v *und* w :

$$v \times w := u.$$

•

3.8.33 Eigenschaften des Kreuzprodukts *Für das Kreuzprodukt gilt:*

i) $(v, w) \mapsto v \times w$ *ist schiefsymmetrische Bilinearform,*

ii) $\langle v \times w \mid v \rangle = \langle v \times w \mid w \rangle = 0_{\mathbb{R}}$,

iii) $v \times w \neq 0 \iff dim_{\mathbb{R}}(\langle v, w \rangle) = 2$,

iv) $\Delta(v, w, v \times w) = \|v \times w\|^2$,

v) $dim_{\mathbb{R}}(\langle v, w \rangle) = 2 \Longrightarrow (v, w, v \times w)$ *ist positiv orientierte Basisfolge,*

vi) $\langle u \times v \mid w \times x \rangle = \langle u \mid w \rangle \langle v \mid x \rangle - \langle u \mid x \rangle \langle v \mid w \rangle$.

vii) Ist $v \neq 0 \neq w$ *und* ω *der Winkel zwischen* v *und* w*, dann gilt*

$$\|v \times w\| = \|v\| \cdot \|w\| \cdot |\sin(\omega)|,$$

viii) Für $u, v, w \in V$ *gilt:*

$$u \times (v \times w) = \langle u \mid w \rangle v - \langle u \mid v \rangle w.$$

ix) Es gilt die Jacobische Gleichung

$$u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0.$$

x) Ist $\mathcal{B} = (b_0, b_1, b_2)$ eine positiv orientierte Orthonormalbasisfolge, dann gilt:

- $b_i \times b_j = sgn([ijk])b_k$, *falls* $k \neq i \neq j \neq k$,
- $v \times w = (v_1w_2 - v_2w_1)b_0 + (v_2w_0 - v_0w_2)b_1 + (v_0w_1 - v_1w_0)b_2$.

□

Beweis: Aus der Gleichung

$$\Delta(v, w, -) = \langle v \times w \mid - \rangle$$

folgen Bilinearität und Schiefsymmetrie des Kreuzprodukts und auch die Tatsache, daß das Kreuzprodukt auf jedem seiner Faktoren senkrecht steht; i) und ii) sind damit bewiesen.
Zum Nachweis von iii) bemerken wir, daßs

$$v \times w \neq 0 \iff \Delta(v, w, -) \neq 0 \iff (v, w) \text{ linear unabhängig.}$$

Behauptung iv) ist ebenfalls offensichtlich richtig:

$$\Delta(v, w, v \times w) = \langle v \times w \mid v \times w \rangle = \|v \times w\|^2.$$

Ist, wie bei v) vorausgesetzt, $\dim(\langle v, w \rangle) = 2$, dann ist (v, w) linear unabhängig, nach iii) also $v \times w \neq 0$. Letzteres ergibt

$$0 < \|v \times w\|^2 = \Delta(v, w, v \times w).$$

Diese Ungleichung impliziert die positive Orientierung, weil Δ die vorgegebene Orientierung repräsentiert.
Zum Beweis von vi) betrachten wir, für später geeignet zu wählende u_i und v_i aus V, den Ausdruck

$$\begin{aligned} \langle u_0 \times u_1 \mid u_2 \rangle \langle v_0 \times v_1 \mid v_2 \rangle &= \Delta(u_0, u_1, u_2)\Delta(v_0, v_1, v_2) \\ &= \det(\langle u_i \mid v_k \rangle). \end{aligned}$$

Setzt man hier $v_2 := u_0 \times u_1$ und entwickelt nach der letzten Zeile, so bleibt von den drei Summanden nur der letzte übrig, und es ergibt sich die Gleichung

$$\langle u_0 \times u_1 \mid u_2 \rangle \langle v_0 \times v_1 \mid u_0 \times u_1 \rangle = \langle u_0 \times u_1 \mid u_2 \rangle (\langle u_0 \mid v_0 \rangle \langle u_1 \mid v_1 \rangle - \langle u_0 \mid v_1 \rangle \langle u_1 \mid v_0 \rangle).$$

Für $u_0 \times u_1 \neq 0$ gibt es u_2 mit $\langle u_0 \times u_1 \mid u_2 \rangle \neq 0$, durch diesen Faktor darf also dividiert werden, und es verbleibt

$$\langle u_0 \times u_1 \mid v_0 \times v_1 \rangle = \langle u_0 \mid v_0 \rangle \langle u_1 \mid v_1 \rangle - \langle u_0 \mid v_1 \rangle \langle u_1 \mid v_0 \rangle,$$

woraus die Behauptung vi) folgt.
Setzt man in dieser Gleichung vi) $u_0 = v_0 := v$ und $u_1 = v_1 := w$, dann erhält man die Gleichung

$$||v \times w||^2 = ||v||^2 ||w||^2 - \langle v \mid w \rangle^2,$$

was mit der Definition des Winkels zwischen v und w die Behauptung vii) liefert.
Zum Beweis von viii) verwenden wir

$$\begin{aligned} \langle u \times (v \times w) \mid x \rangle &= \Delta(u, v \times w, x) \\ &= -\Delta(u, x, v \times w) \\ &= -\langle u \times x \mid v \times w \rangle \\ &= -\langle u \mid v \rangle \langle x \mid w \rangle + \langle u \mid w \rangle \langle x \mid v \rangle \end{aligned}$$

Es gilt demnach, für alle $x \in V$,

$$\langle u \times (v \times w) \mid - \rangle 0 \langle \langle u \mid w \rangle \mid - \rangle - \langle \langle u \mid v \rangle w \mid - \rangle,$$

was die Behauptung beweist.
ix) folgt direkt aus viii).

□

3.9 Normierte Räume

Ist $(V, \langle - \mid - \rangle)$ ein Raum mit innerem Produkt, dann haben wir gemäß ?? eine Abbildung $\| - \| \colon V \to \mathbb{R}$ mit allen Eigenschaften einer Norm gemäß folgender

3.9.1 Definition (Norm) *Eine Abbildung* $\| - \| \colon V \to \mathbb{R}$ *heißt* Norm, *wenn folgende Bedingungen erfüllt sind:*

i) $\|v\| \geq 0$, *mit Gleichheit genau dann, wenn* $v = 0$.

ii) Es gilt die Dreiecksungleichung $\|v + w\| \leq \|v\| + \|w\|$.

iii) $\|\rho \cdot v\| = |\rho| \cdot \|v\|$.

Das Paar $(V, \| - \|)$ *heißt dann* normierter Raum. •

Jeder Raum $(V, \langle - \mid - \rangle)$ mit innerem Produkt ist demnach auf kanonische Weise auch ein normierter Raum. Die Umkehrung gilt aber nicht, denn nicht jede Norm ergibt anhand der Gleichung

$$\langle v \mid w \rangle = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2)$$

ein inneres Produkt. Als Beispiel hierfür betrachten wir den Raum C der auf $[0, 1]$ stetigen reellwertigen Funktionen mit der bekannten und wichtigen *Maximumsnorm*

$$\|f\| := \max\{|f(x)| \mid x \in [0, 1]\}.$$

Die durch

$$\langle f \mid g \rangle := \frac{1}{2}(\|f + g|^2 - \|f\|^2 - \|g\|^2)$$

definierte Abbildung ist nämlich nicht bilinear: $\langle x \mid x \rangle = 1$, $\langle 1 - x \mid x \rangle = -1/2$, ber $\langle 1 \mid x \rangle = 1$. (Mit zusätzlichen Forderungen an die Norm ergibt sich allerdings ein inneres Produkt (vgl. Übungsblatt).

Mittels $\| - \|$ läßt sich auch eine *Distanzfunktion* $d_{\|-\|}$ (kurz auch: d) definieren durch

3.9.2 $$d_{\|-\|} \colon V \times V \to \mathbb{R} \colon (v, w) \mapsto \|v - w\|.$$

Sie hat offenbar die folgenden Eigenschaften (vgl. 3.8.4):

3.9.3 Eigenschaften der Distanzfunktion *Für alle* $u, v, w \in V$ *gilt:*

1. *Distanzen sind nicht negativ:*

$$d(v, w) \geq 0,$$

mit Gleichheit genau dann, wenn $v = w$.

2. *Es gilt die* Dreiecksungleichung

$$d(v, w) \leq d(v, u) + d(u, w).$$

3. *Die Distanz ist* translationsinvariant:

$$d(u+v, u+w) = d(v, w).$$

□

$d = d_{||-||}$ ist also eine *Metrik* und $(V, d_{||-||})$ ein *metrischer Raum* im Sinne der Topologie. Man beachte, daß jede Funktion $d\colon V \times V \to \mathbb{R}$ mit 3.9.3 umgekehrt auch ein inneres Produkt liefert. Es ist also im wesentlichen dasselbe, ob wir von einem Raum mit innerem Produkt, einem normierten Raum oder einem metrischen Vektorraum reden.

3.9.4 Definition *Sind V, W normierte Räume, dann heißt $f\colon V \to W$*

1. beschränkt, *wenn gilt*

$$\exists\ \rho \in \mathbb{R}\ \forall\ v \in V\colon\ \|f(v)\| \leq \rho \cdot \|v\|.$$

2. stetig an der Stelle $v \in V$, *falls*

$$\forall\ \epsilon > 0\ \exists\ \delta > 0\ \forall\ v' \in V\colon\ \|v - v'\| < \delta\ \Rightarrow\ \|f(v) - f(v')\| < \epsilon.$$

3. stetig, *wenn f an jeder Stelle stetig ist.*

•

Schreibt man diese Bedingungen zur Verdeutlichung mit Hilfe der Distanzfunktion d um, dann ist zunächst die Beschränktheit von f äquivalent zu

$$\exists\ \rho \in \mathbb{R}\ \forall\ v \in V\colon\ d(f(0), f(v)) \leq \rho \cdot d(0, v).$$

Dies *zeigt zunächst, daß Beschränktheit die Stetigkeit an der Stelle 0_V impliziert.* Die Bedingung für Stetigkeit an der Stelle v sieht umformuliert so aus:

$$\forall\ \epsilon > 0\ \exists\ \delta > 0\ \forall\ v' \in V\colon\ d(v, v') < \delta\ \Rightarrow\ d(f(v), f(v')) < \epsilon.$$

Dies wiederum zeigt, daß sich — wegen der Translationsinvarianz — ggf. die Stetigkeit an der Stelle 0_V auf jede andere Stelle überträgt. Wir erhalten also den

3.9.5 Satz *Sind V, W normiert, $f \in Hom_{\mathbb{R}}(V, W)$, dann sind äquivalent:*

1. *f ist stetig,*

2. *f ist an der Stelle 0_V stetig,*

3. *es gibt eine Stelle $v' \in V$, an der f stetig ist.*

□

3.9.6 Satz *Unter den linearen Abbildungen zwischen normierten Räumen sind genau die beschränkten stetig.* □

Beweis: Ist f beschränkt, dann ist f an 0_V stetig, also stetig.
Ist umgekehrt f stetig, dann insbesondere an der Stelle 0_V, es gibt also, zu $\epsilon := 1$, ein $\delta > 0$ mit:

$$\|v\| < \delta \;\Rightarrow\; \|f(v)\| < 1.$$

Zu $v \neq 0_V$ sei jetzt ein Vektor w definiert durch $w := \frac{\delta}{2}\frac{v}{\|v\|}$, dann ist $\|w\| = \frac{\delta_0}{2}$, also

$$1 > \|f(w)\| = \frac{\delta}{2}\frac{\|f(v)\|}{\|v\|},$$

und damit folgt

$$\|f(v)\| < \frac{2}{\delta}\|v\|. \tag{$*$}$$

Da letzteres aber auch für $v := 0_V$ richtig ist, haben wir für *alle* $v \in V$ die Ungleichung $(*)$ bewiesen, f ist demnach beschränkt (mit $\rho = 2/\delta$).

□

3.9.7 Satz *Jede lineare Abbildung zwischen euklidischen Vektorräumen ist beschränkt und damit auch stetig.*

Beweis: Sei $\mathcal{B}$ eine Orthonormalbasisfolge von V, $\mathcal{C}$ eine Basisfolge von W, $f \in Hom_{\mathbb{R}}(V,W)$, V und W euklidisch, $A := M(\mathcal{C}, f, \mathcal{B})$. Dann gilt:

$$\begin{aligned}
&\|f(v)\| = \|A \cdot v\| = \|A \sum_i \langle v \mid b_i\rangle b_i\| = \| \sum_i \langle v \mid b_i\rangle A b_i \\
&\leq \sum_i |\langle v \mid b_i\rangle| \cdot \|Ab_i\| \leq_{3.8.4} \sum_i \|v\| \cdot \|b_i\| \cdot \|Ab_i\| \\
&\leq \|v\| \cdot \underbrace{n \cdot \max\{\|Ab_i\| \mid i \in n\}}_{=:\rho}.
\end{aligned}$$

□

Es gibt aber auch andere interessante Abbildungen, die unter gewissen Bedingungen stetig sind:

3.9.8 Satz *Ist $(V, \langle - \mid - \rangle)$ euklidisch, $f \in End_{\mathbb{R}}$, dann ist $\varphi\colon v \mapsto \langle v \mid f(v)\rangle$ stetig.*

Beweis:

$$\begin{aligned}
&\|\varphi(v)\| = |\langle v \mid f(v)\rangle| =_{3.8.4} \frac{1}{2}\left|\|v + f(v)\|^2 - \|v\|^2 - \|f(v)\|^2\right| \\
&\leq \frac{1}{2}\|v + f(v)\|^2 + \frac{1}{2}\|v\|^2 + \frac{1}{2}\|f(v)\|^2 \\
&\leq \|v\|^2 + \|f(v)\|^2 + \|v\|\|f(v)\|.
\end{aligned}$$

Da f beschränkt ist (vgl. 3.9.7), etwa $\|f(v)\| \leq \rho\|v\|$, für alle $v \in V$, gilt weiter

$$\leq (1 + \rho^2 + \rho)\|v\|^2.$$

Für $\rho := \epsilon/(1 + \rho^2 + \rho)$ gilt also

$$\|\varphi(v - v')\| < \epsilon,$$

sobald $\|v - v'\| < \delta$.

$\square$

3.10 Die adjungierte Abbildung

Die Vektorräume dieses Paragraphen seien sämtlich euklidisch. Zu $f \in Hom_{\mathbb{R}}(V, W)$ und $w \in W$ ist $\langle f(-) \mid w\rangle \in L(V)$, nach dem Rieszschen Darstellungssatz (3.8.31) gibt es also genau ein $v \in V$ mit

$$\langle f(-) \mid w\rangle = \langle - \mid v\rangle.$$

Die Zuordnung $w \mapsto v$ definiert also eine Abbildung $\tilde{f}$, und diese ist offenbar auch linear. Wir nennen sie die zu f *adjungierte* Abbildung, und sie erfüllt die folgende Gleichung, bzw. ist durch diese definiert:

$$\langle f(v) \mid w\rangle = \langle v \mid \tilde{f}(w)\rangle.$$

3.10.1 Hilfssatz *Die zu $f \in Hom_{\mathbb{R}}(V, W)$ adjungierte Abbildung $\tilde{f}$ hat die folgenden Eigenschaften:*

1. *f ist die adjungierte Abbildung zu $\tilde{f}$,*
2. *$W = \mathrm{Bild}(f) \oplus \mathrm{Kern}(\tilde{f})$,*
3. *Sind $\mathcal{B}, \mathcal{C}$ Orthonormalbasisfolgen für V, W, dann gilt*
$$M(\mathcal{C}, f, \mathcal{B}) = {}^tM(\mathcal{B}, \tilde{f}, \mathcal{C}).$$
4. *Ist $V = W, g \in End_{\mathbb{R}}(V)$, dann ist $\widetilde{f \circ g} = \tilde{g} \circ \tilde{f}$, und irgend zwei Eigenvektoren v von f und $\tilde{v}$ von $\tilde{f}$ zu verschiedenen Eigenwerten κ und $\tilde{\kappa}$ sind orthogonal.*

Beweis:

1. $\langle f(v) \mid w\rangle = \langle v \mid \tilde{f}(w)\rangle = \langle \tilde{\tilde{f}}(v) \mid w\rangle$ ergibt $f = \tilde{\tilde{f}}$, da $\langle - \mid -\rangle$ nicht ausgeartet ist.
2. folgt aus $\mathrm{Bild}(f)^{\perp} = \mathrm{Kern}(\tilde{f})$ mit 4.4.18.
3. Ist $A := M(\mathcal{C}, f, \mathcal{B})$, $\tilde{A} := M(\mathcal{B}, \tilde{f}, \mathcal{C})$, dann gilt
$$a_{jk} = \sum_i a_{ik}\langle c_i \mid c_j\rangle = \langle f(b_k) \mid c_j\rangle = \langle b_k \mid \tilde{f}(c_j)\rangle = \sum_i \tilde{a}_{ij}\langle b_k \mid b_i\rangle = \tilde{a}_{kj}.$$
4. $\langle (f \circ g)(v) \mid w\rangle = \langle g(v) \mid \tilde{f}(w)\rangle = \langle v \mid (\tilde{g} \circ \tilde{f})(w)\rangle$. Außerdem gilt für Eigenvektoren $v, \tilde{v}$ zu Eigenwerten $\kappa, \tilde{\kappa}$:
$$\begin{aligned}(\kappa - \tilde{\kappa})\langle v \mid \tilde{v}\rangle &= \langle \kappa v \mid \tilde{v}\rangle - \langle v \mid \tilde{\kappa}\tilde{v}\rangle \\ &= \langle f(v) \mid \tilde{v}\rangle - \langle v \mid \tilde{f}(\tilde{v})\rangle \\ &= \langle v \mid \tilde{f}(\tilde{v})\rangle - \langle v \mid \tilde{f}(v)\rangle \\ &= 0.\end{aligned}$$

Sind die beiden Eigenwerte verschieden, so folgt daraus $\langle v \mid \tilde{v}\rangle = 0$. □

3.10.2 Satz *Ist V euklidisch, dann gilt*

$$\varphi\colon End_{\mathbb{R}}(V) \simeq_{\mathbb{R}} BLF(V)\colon f \mapsto \langle f(-) \mid -\rangle.$$

Beweis. Die Linearität von φ ist trivial. Die Injektivität folgt aus der Tatsache, daß $\langle f(-) \mid -\rangle\colon v \mapsto 0$ impliziert, daß f die Nullabbildung ist, denn dies bedeutet ja

$$\forall\ u, v \in V\colon\ \langle f(u) \mid v\rangle = 0,$$

also $f(u)$ im Nullraum und demnach $f(u) = 0$, für alle u, d. h. $f = 0$.
Die Surjektivität folgt so: Für $\Phi \in BLF(V)$, $v \in V$, gibt es nach dem Rieszschen Darstellungssatz genau ein $u \in V$ mit

$$\Phi(v, -) = \langle u \mid -\rangle.$$

Damit ist $f\colon v \mapsto u$ wohldefiniert, und es gilt dafür offenbar $\varphi(f) = \Phi$. □

3.10.3 Beispiele
i) Ist $V := \mathbb{R}^n$ und $\langle - \mid -\rangle$ das Standardskalarprodukt, dann kann man die Werte von $\Phi := \langle f(-) \mid -\rangle \in BLF(V)$ mit Hilfe der Matrix $A := M(\mathcal{B}, f, \mathcal{B})$ berechnen:

$$\langle f(u) \mid v\rangle = \sum_i f(u)_i v_i = \sum_{i,k} a_{ik} u_k v_i = {}^t v \cdot A \cdot u.$$

Die Matrix A bezeichnen wir deshalb auch mit

$$M^{\Phi}_{\mathcal{B},\mathcal{B}}.$$

ii) Für $\tilde{\Phi} := \varphi(\tilde{f})$ (nach 3.10.2) gilt

$$\tilde{\Phi}(v, w) = \langle \tilde{f}(v) \mid w\rangle = \langle v \mid f(w)\rangle = \langle f(w) \mid v\rangle = \Phi(w, v).$$

Die zur adjungierten Abbildungen gemäß 4.6.3 gehörende Bilinearform erhält man also durch Vertauschen der Argumente. ◇

3.10.4 Definition (normal) $f \in End_{\mathbb{R}}(V)$ *heißt* normal, *wenn gilt*

$$f \circ \tilde{f} = \tilde{f} \circ f.$$

•

3.10.5 Hilfssatz $f \in End_{\mathbb{R}}(V)$, *dann sind äquivalent.*

1. *f ist normal,*
2. $\langle f(v) \mid f(w)\rangle = \langle \tilde{f}(v) \mid \tilde{f}(w)\rangle$.
3. $\|f(v)\|^2 = \|\tilde{f}(v)\|^2$.

Beweis: Nachrechnen. □

3.10.6 Hilfssatz *Ist $f \in End_{\mathbb{R}}(V)$ normal, dann gilt*

1. $\mathrm{Kern}(f) = \mathrm{Kern}(\tilde{f})$,
2. $V = \mathrm{Bild}(f) \oplus \mathrm{Kern}(f)$,
3. $\forall\ n \in \mathbb{N}^*$: $\mathrm{Rang}(f) = \mathrm{Rang}(f^n)$.

Beweis:

1. Die erste Behauptung folgt unmittelbar aus 3.10.5 iii).
2. Die zweite Behauptung ergibt sich aus 3.10.1 ii) mit der gerade bewiesenen ersten Behauptung.
3. Aus i) folgt, daß die Einschränkung $f \downarrow \mathrm{Bild}(f)$ regulär ist. Es folgt

$$\mathrm{Rang}(f^2) = \dim(f(\mathrm{Bild}(f))) = \dim(\mathrm{Bild}(f)) = \mathrm{Rang}(f),$$

und ganz analog ergibt sich $\mathrm{Rang}(f^3) = \mathrm{Rang}(f)$, usw.

□

3.10.7 Hilfssatz *Ist $f \in End_{\mathbb{R}}(V)$ normal und $\rho \in \mathbb{R}$, dann gilt*

1. $\widetilde{f - \rho \cdot id_V} = \tilde{f} - \rho \cdot \mathrm{id}_V$,
2. *$f - \rho \cdot id_V$ ist normal.*

Beweis: Nachrechnen.

□

3.10.8 Folgerung *Ist f normal, dann haben f und $\tilde{f}$ dieselben Eigenwerte und Eigenvektoren.*

Beweis: Die Gleichungen

$$\mathrm{Kern}(f - \rho \cdot \mathrm{id}_V) = \mathrm{Kern}(\widetilde{f - \rho \cdot \mathrm{id}_V}) = \mathrm{Kern}(\tilde{f} - \rho \cdot \mathrm{id}_V)$$

ergeben die Identität der entsprechenden Eigenräume:

$$E_f(\rho) = E_{\tilde{f}}(\rho).$$

□

3.10.9 Satz *Sei $f \in End_{\mathbb{R}}(V)$ und $V = V_1 \oplus \ldots \oplus V_r$ mit paarweise orthogonalen V_i (d.h. $V_i \subseteq V_j^{\perp}$, falls $i \neq j$), kurz:*

$$V = V_1 \perp \ldots \perp V_r.$$

Sind zudem die V_i invariant unter f, d.h. $f(v_i) \subseteq V_i$, dann ist f genau dann normal, wenn die V_i auch unter den $\tilde{f_i}$ (zu $f_i := f \downarrow V_i$) invariant und die f_i normal sind. □

Beweis:
i) Sei f normal. Wir zeigen zunächst die Invarianz $\tilde{f}_i(V_i) \subseteq V_i$: Ist $v_j \in V_j, j \neq i$, dann gilt
$$\langle v_j \mid \tilde{f}(v_i)\rangle = \langle f(v_j) \mid v_i\rangle = 0,$$
also, für alle $j \neq i$: $\tilde{f}(v_i) = \tilde{f}_i(v_i) \in V_j^{\perp}$, und damit
$$\tilde{f}_i(v_i) \in \bigcap_{j \neq i} V_j^{\perp} = V_i.$$

Zum Nachweis der Normalität von f_i beachten wir, daß
$$\|f_i(v_i)\|^2 = \|f(v_i)\|^2 = \|\tilde{f}(v_i)\|^2 = \|\tilde{f}_i(v_i)\|^2.$$

Daraus folgt die Normalität von f_i nach 3.10.5.
ii) Jetzt seien umgekehrt die f_i normal und die V_i invariant unter den $\tilde{f}_i$. Ist $v = \sum_1^r v_i$, mit $v_i \in V_i$, dann gilt
$$\|f(v)\|^2 = \|\sum_i f_i(v_i)\|^2 = \sum_i \|f_i(v_i)\|^2,$$
letzteres wegen $f_i(v_i) \in V_i \perp V_j, j \neq i$. Da die f_i als normal vorausgesetzt sind, gilt weiter:
$$= \sum_i \|\tilde{f}_i(v_i)\|^2 = \|\tilde{f}(v)\|^2,$$
also ist auch f normal. □

3.10.10 Definition (selbstadjungiert) *$f \in End_{\mathbb{R}}(V)$ heißt* selbstadjungiert, *wenn $\tilde{f} = f$ gilt.* •

Unmittelbar aus 3.10.1 ergibt sich

3.10.11 Folgerung *Ist f selbstadjungiert, $\mathcal{B}$ eine Orthogonalbasisfolge, dann gilt*
$$M(\mathcal{B}, f, \mathcal{B}) = {}^tM(\mathcal{B}, f, \mathcal{B}),$$
d.h. die Matrix, die f beschreibt, ist symmetrisch. □

Eines unserer Ziele ist der Nachweis der Tatsache, daß selbstadjungierte lineare Endomorphismen eines $n-$dim ensionalen euklidischen Raumes n paarweise orthogonale Eigenvektoren besitzen. Dazu betrachten wir zu selbstadjungiertem f die Abbildung
$$\varphi: V\backslash\{0_V\} \to \mathbb{R}: v \mapsto \frac{\langle v \mid f(v)\rangle}{\langle v \mid v\rangle}.$$

3.10.12 Bemerkungen Diese Funktion φ hat die folgenden Eigenschaften:

- $\forall\ \rho \in \mathbb{R}^*, v \in V$: $\varphi(\rho v) = \varphi(v)$.

- Als Quotient stetiger Funktionen nimmt φ auf der *Einheitssphäre,* d.h. auf der Menge $\{v \mid \|v\| = 1\}$, ein Minimum an, etwa bei $b_0 \in V$, d.h. es gilt
$$\|v\| = 1 \Longrightarrow \varphi(v) \geq \varphi(b_1).$$

- Mit dem ersten Punkt folgt dann weiter:
$$\forall\ v \in V\backslash\{0_V\}:\ \varphi(v) \geq \varphi(b_1).$$

3.10.13 Hilfssatz b_0 *ist Eigenvektor von* f *zum Eigenwert* $\rho_1 := \langle b_1 \mid f(b_1)\rangle$.

Beweis: Sei $v \in V$ und
$$\psi\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto \varphi(b_0 + xv).$$

Nach 3.10.12 nimmt ψ an der Stelle $x = 0$ ein Minimum an, es gilt also $\psi'(0) = 0$. Einsetzen der Definition von φ und Differenzieren nach x ergibt
$$\psi'(0) = 2\langle b_0 \mid f(v)\rangle - 2\langle b_0 \mid f(b_0)\rangle\langle b_0 \mid v\rangle.$$

Setzen wir dies gleich Null, so folgt
$$0 = \langle f(b_0) - \langle b_0 \mid f(b_0) b_0 \mid v\rangle,$$

und zwar für *alle* $v \in V$. Das impliziert
$$f(b_0) = \langle b_0 \mid f(b_0)\rangle b_0,$$

wie behauptet.

□

3.10.14 Satz *Zu jedem selbstadjungierten linearen Endomorphismus eines euklidischen Vektorraums gibt es eine Orthonormalbasisfolge aus Eigenvektoren.*

Beweis: Ist $f \in End_{\mathbb{R}}(V), V$ euklidisch, $n := \dim_{\mathbb{R}} V \in \mathbb{N}^*$, dann besitzt f nach 3.10.13 einen Eigenvektor b_0. Wir betrachten die orthogonale Zerlegung
$$V = \langle b_0 \rangle \perp \langle b_0 \rangle^{\perp}$$

und verwenden, daß $f\langle b_0\rangle^{\perp}) \subseteq \langle b_0\rangle^{\perp}$ (da f selbstadjungiert ist). Mit f ist auch die Einschränkung $f \downarrow \langle b_0\rangle^{\perp}$ selbstadjungiert, obige Überlegungen ergeben also einen Eigenvektor der Norm 1 in $\langle b_0\rangle^{\perp}$ usw. Die Behauptung folgt auf diese Weise per Induktion nach der Dimension von V. □

3.10.15 Folgerungen

1. *Reelle symmetrische Matrizen sind diagonalisierbar.*

2. *Ist $A \in \mathbb{R}^{n\times n}$ symmetrisch, dann hat A lauter reelle Eigenwerte λ_i, für diese gilt*

$$p_A = (-1)^n \prod_{i=1}^{n}(x-\lambda_i).$$

□

Mit 3.10.2 erhalten wir noch

3.10.16 Folgerung *Zu jeder symmetrischen Bilinearform auf einem euklidischen Vektorraum V gibt es eine Orthonormalbasisfolge $\mathcal{B}$ in V, bzgl. derer die Form durch eine Diagonalmatrix beschrieben wird.*

□

3.10.17 Definition (Projektion,Orthogonalprojektion) *Ein linearer Endomorphismus f eines Vektorraumes heißt* Projektion, *wenn $f^2 = f$ gilt. Ein Projektion heißt* Orthogonalprojektion, *wenn er selbstadjungiert ist.* •

3.10.18 Hilfssatz *Ist $f \in End_{\mathbb{R}}(V)$ ein Projektion, dann gilt*

$$V = \mathrm{Kern}(f) \oplus \mathrm{Bild}(f).$$

Ist f darüberhinaus eine Orthogonalprojektion, dann ist f normal und

$$f = 0_{\mathrm{Kern}(f)} \oplus \mathrm{id}_{\mathrm{Bild}(f)}.$$

Beweis: Ist f Projektion, d.h. $f^2 = f$, dann ist zunächst die Summe $\mathrm{Kern}(f) + \mathrm{Bild}(f)$ direkt: $v \in \mathrm{Kern}(f) \cap \mathrm{Bild}(f)$, etwa $v = f(w)$, ergibt $0 = f(v) = f^2(w) = f(w) = v$. Diese direkte Summe ist darüberhinaus gleich V : Für $v \in V$ gibt es $w \in V$ mit $v = f(v) + w$, denn V ist eine additive Gruppe. Der Summand w liegt im $\mathrm{Kern}(f)$, da $f(v) = f^2(v) + f(w) = f(v) + f(w)$. Jedes $v \in V$ ist demnach als Summe aus einem Element des Kerns und einem Element des Bildes darstellbar. Damit ist der erste Teil der Behauptung bewiesen.
Der zweite Teil, die Beschreibung von f als Summe aus der Nullabbildung auf dem Kern und der identischen Abbildung auf dem Bild von f folgt direkt aus dem ersten Teil, denn dieser liefert die Zerlegung in Nullabbildung auf dem Kern und der Einschränkung auf das Bild, und die Gleichung $f^2 = f$ zeigt noch, daß diese Einschränkung gleich der identischen Abbildung ist. Die Normalität von f ist trivial. □

3.10.19 Hilfssatz *Ist f normale Projektion, dann auch Orthogonalprojektion.*

Beweis: $v \in V$ und wieder $v = f(v) + w, w \in \mathrm{Kern}(f)$, dann gilt (vgl. den Beweis von 3.10.18)

$$\langle v \mid f(u)\rangle = \langle f(v) + w \mid f(u)\rangle = \langle f(v) \mid f(w)\rangle,$$

letzteres wegen $w \in \mathrm{Kern}(f) = \mathrm{Kern}(\tilde{f})$. Analog ergibt sich

$$\langle f(v) \mid w\rangle = \langle f(v) \mid f(w)\rangle,$$

also insgesamt

$$\langle v \mid f(w)\rangle = \langle f(w) \mid v\rangle.$$

Das liefert $\tilde{f} = f$, wie behauptet.

□

3.10.20 Hilfssatz

1. *Ist U ein Unterraum von V, dann gibt es genau eine Orthogonalprojektion $f \in End_{\mathbb{R}}(V)$ mit $\mathrm{Bild}(f) = U$, nämlich $f := \mathrm{id}_U \oplus 0_{U^\perp}$.*
2. *Sind $U_1, U_2 \leq_{\mathbb{R}} V$ und f_1, f_2 die zugehörigen Orthogonalprojeektionen gemäß i), dann gilt:*
   - $f_2 \circ f_1 = 0 \Longleftrightarrow U_1 \perp U_2$,
   - $f_1 + f_2$ *ist Orthogonalprojektion* $\Longleftrightarrow U_1 \perp U_2$, *item*
   - $f_1 - f_2$ *ist Orthogonalprojektion* $\Longleftrightarrow U_2 \subseteq U_1$,
   - $f_1 \circ f_2$ *ist Orthogonalprojektion* $\Longleftrightarrow f_2 \circ f_1 = f_1 \circ f_2$.

Beweis: Übungsaufgabe. □

3.10.21 Definition (schiefsymmetrisch) $f \in End_{\mathbb{R}}(V)$ *heißt* schiefsymmetrisch, *wenn* $\tilde{f} = -f$ *gilt.*

•

3.10.22 Folgerung *Ist f schiefsymmetrisch und $\mathcal{B}$ eine Orthogonalbasisfolge, dann gilt*

$$M(\mathcal{B}, f, \mathcal{B}) = -\,{}^tM(\mathcal{B}, f, \mathcal{B}),$$

d.h. die f darstellende Matrix ist schiefsymmetrisch.

□

3.10.23 Hilfssatz *Ist $f \in End_{\mathbb{R}}(V)$, dann sind äquivalent:*

1. *f ist schiefsymmetrisch,*
2. $\langle f(v) \mid w\rangle + \langle v \mid f(w)\rangle = 0$,
3. $\langle v \mid f(v)\rangle = 0$.

□

3.10.24 Satz *Ist $f \in End_{\mathbb{R}}(V)$ schiefsymmetrisch, dann gilt:*

1. *0 ist der einzige Eigenwert von f,*

2. *Spur* $f = 0$,

3. $\det(f) = (-1)^{dim(V)} \det(f)$,

4. $dim_{\mathbb{R}}(V)$ *ungerade* $\Longrightarrow \det(f) = 0$,

5. $Rg(f)$ *ist gerade.*

□

Beweis: Die erste Behauptung folgt aus 3.10.23, 3.: Ist v ein Eigenvektor zum Eigenwert κ, dann gilt
$$0 = \langle v \mid f(v)\rangle = \kappa\langle v \mid v\rangle.$$
Zum Nachweis der zweiten Behauptung argumentieren wir so:
$$\mathrm{Spur}(f) = \mathrm{Spur}(\tilde{f}) = -\mathrm{Spur}(f).$$
Entsprechend für die dritte Behauptung:
$$\det(f) = \det(\tilde{f}) = \det(-f) = (-1)^{\dim(V)} \det(f).$$
Die vierte Behauptung folgt jetzt unmittelbar. $\mathrm{Bild}(f) = \mathrm{Kern}(f)^{\perp}$ folgt aus 4.6.7. Daraus ergibt sich, daß $f \downarrow \mathrm{Bild}(f)$ regulär und schiefsymmetrisch ist. Wegen iv) ist deshalb $\dim_{\mathbb{R}}(\mathrm{Bild}(f) = Rg(f)$ gerade.

□

3.10.25 Folgerung *Der Rang jeder schiefsymmetrischen Matrix ist gerade.* □

3.10.26 Satz *Ist $f \in End_{\mathbb{R}}(V)$ schiefsymmetrisch, dann gibt es eine Orthonormalbasisfolge $\mathcal{B}$ mit*
$$M(\mathcal{B}, f, \mathcal{B}) = \begin{pmatrix} 0 & -\kappa_1 & & & & 0 \\ \kappa_1 & 0 & & & & \\ & & 0 & -\kappa_2 & & \\ & & \kappa_2 & 0 & & \\ 0 & & & & \ddots & \end{pmatrix}$$

□

Beweis: Die Abbildung $\varphi := f^2$ ist selbstadjungiert, es gibt also eine 0N-Basisfolge $\mathcal{C} = (c_1, \ldots, c_n)$ aus Eigenvektoren von φ, sei etwa $\varphi(c_i) = \lambda_i c_i$.
i) $\lambda_i \leq 0$: Da die c_i Vektoren mit Norm 1 sind, gilt
$$\lambda_i = \langle c_i \mid \varphi(c_i)\rangle = \langle c_i \mid f^2(c_i)\rangle = -\langle f(c_i) \mid f(c_i)\rangle \leq 0.$$
ii) $\mathrm{Rang}(\varphi)$ ist gerade: $\mathrm{Rang}(\varphi) = \mathrm{Rang}(f^2) = \mathrm{Rang}(f)$, letzterer ist nach 3.10.24 gerade.
iii) Die Anzahl der negativen Eigenwerte von φ gleicht dem Rang von φ, ist also gerade. $c_1, \ldots, c_{2s}$ seien die zugehörigen Eigenvektoren. Setzen wir jetzt
$$b_{2\nu-1} := c_\nu, b_{2\nu} := \sqrt{-\lambda_\nu^{-1}} \cdot f(c_\nu), 1 \leq \nu \leq s,$$

und

$$b_\nu := c_\nu, \nu \geq 2s,$$

dann hat die Basisfolge $\mathcal{B}$ offenbar die gewünschten Eigenschaften, mit $\kappa_i := \sqrt{-\lambda_\nu}$, wie man leicht nachrechnet:

$$f(b_{2\nu-1}) = f(c_\nu) = \sqrt{-\lambda_\nu} \cdot b_{2\nu},$$

$$f(b_{2\nu} = f(\sqrt{-\lambda_\nu^{-1}} \cdot f(c_\nu) = \sqrt{-\lambda_\nu} \cdot f^2(b_{2\nu}),$$

$\square$

Kapitel 4

Anwendungen in der Geometrie

Wir kommen jetzt zu Anwendungen der Linearen Algebra auf die Geometrie.

4.1 Affine Punkträume

Es wird jetzt der Übergang von der linearen Algebra zur analytischen Geometrie beschrieben.

4.1.1 Definition (affiner Punktraum) *Sei V ein $\mathbb{K}$-Vektorraum, A eine nicht leere Menge und*

$$\overrightarrow{} : A \times A \to V : (P,Q) \to \overrightarrow{PQ}$$

eine Abbildung. Dann heißt $(A, \overrightarrow{})$ affiner Punktraum *zu V, wenn gilt:*

1. $\forall\ P \in A, v \in V\ \exists_1\ Q \in A:\ \overrightarrow{PQ} = v.$
   (Kurz: $q = P + v$.)
2. $\forall\ P, Q, R \in A:\ \overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}.$

•

Die Elemente von A heißen *Punkte.* Ist V $n-$dimensional, so heißt auch A *$n-$dimensional:*

$$\dim_{\mathbb{K}}(A) := \dim_{\mathbb{K}}(V).$$

4.1.2 Beispiel Zu V sei $A := V$ und

$$\overrightarrow{} : A \times A = V \times V \to V : (v_1, v_2) \mapsto \overrightarrow{v_1 v_2} := v_2 - v_1.$$

Dieser affine Punktraum zu V wird mit $D(V)$ bezeichnet. □

4.1.3 Hilfssatz *Ist (A, α) affiner Raum zu V und sind $P, Q, R, S \in A$, dann gilt:*

1. $\overrightarrow{PQ} = 0_V \iff P = Q,$
2. $\overrightarrow{PQ} = -\overrightarrow{QP},$
3. $\overrightarrow{PQ} = \overrightarrow{RS} \Longrightarrow \overrightarrow{PR} = \overrightarrow{QS}$ (Parallelogrammregel).

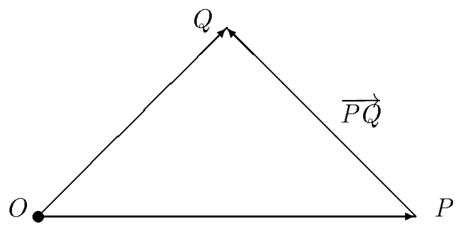
□

Greifen wir uns einen Punkt $0 \in A$ als *Ursprungspunkt* heraus, so ist nach 4.1.1 $P \in A$ eindeutig durch den Vektor $p := \overrightarrow{OP} \in V$ bestimmt, seinen *Ortsvektor.* P kann dann mit p *identifiziert* werden, d.h. wir haben die Bijektion

4.1.4 $$A \to V : P \mapsto p := \overrightarrow{OP}.$$

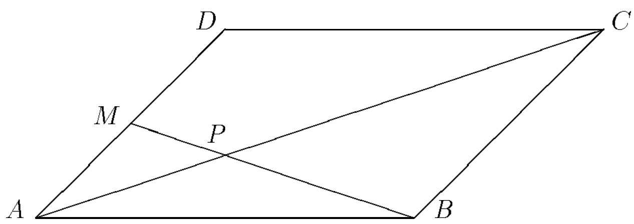
Der Differenzvektor $\overrightarrow{PQ}$ von P und Q genügt dann der Gleichung

4.1.5 $$\overrightarrow{PQ} = q - p.$$



Wie man nun damit die lineare Algebra auf die Elementargeometrie anwenden kann zeigt folgendes Beispiel:

4.1.6 Beispiel Wir wollen beweisen, daß die Diagonale und die Seitenhalbierende sich in nicht ausgearteten Parallelogrammen im Verhältnis 1:2 schneiden.
Beweis: Wir betrachten das Parallelogramm, dessen Punkte wie skizziert bezeichnet seien:



Gesucht sind $\lambda, \mu \in \mathbb{R}$ mit $\overrightarrow{PM} = \lambda \overrightarrow{BM}$ und $AP = \mu \overrightarrow{AC}$. Dazu betrachten wir das Dreieck AMP, für welches gilt:

$$\overrightarrow{AM} + \overrightarrow{MP} + \overrightarrow{PA} = 0,$$

bzw.

$$\frac{1}{2}\overrightarrow{AD} - \overrightarrow{\lambda BM} - \overrightarrow{\mu AC} = 0.$$

Das ist nach 4.1.1 ii) äquivalent zu

$$\frac{1}{2}\overrightarrow{AD} - \lambda(\overrightarrow{BA} + \overrightarrow{AM}) - \mu(\overrightarrow{AB} + \overrightarrow{BC}) = 0.$$

Da es sich um ein Parallelogramm handelt, gilt $\overrightarrow{BC} = \overrightarrow{AD}$, so daß sich hieraus folgendes ergibt:

$$(\frac{1}{2} - \frac{\lambda}{2} - \mu)\overrightarrow{AD} + (\lambda - \mu)\overrightarrow{AB} = 0.$$

Das Parallelogramm ist genau dann nicht ausgeartet, wenn $(\overrightarrow{AD}, \overrightarrow{AB})$ linear unabhängig ist, so daß wir schließlich folgendes Gleichungssystem für die gesuchten Parameter bekommen:

$$\lambda - \mu \quad = \quad 0$$

$$\frac{1}{2} - \frac{\lambda}{2} - \mu \quad = \quad 0.$$

Die einzige Lösung ist $\lambda = \mu = \frac{1}{3}$, wie behauptet. □

4.1.7 Definition (affine Koordinaten, Unterräume und Abbildungen)
A sei ein affiner Unterraum zu V.

i) *Ein Ursprungspunkt $O \in A$ ergibt zusammen mit irgendeiner Basisfolge $\mathcal{B} = (b_0, \ldots, b_{n-1})$ von V ein* affines Koordinatensystem

$$(O; b_0, \ldots, b_{n-1}).$$

Die diesbezüglichen Komponenten p_i des Ortsvektors $p = \overrightarrow{OP}$ von $P \in A$:

$$\overrightarrow{OP} = \sum_i p_i b_i$$

heißen die affinen Koordinaten *von P bezüglich* $(0; b_0, \ldots, b_{n-1})$.

ii) *$B \subseteq A$, heißt B* affiner Unterraum *von A, (kurz: $B \leq A$), wenn gilt:*

$$\exists\ P_0 \in B : \ \{\overrightarrow{P_0Q} \mid Q \in B\} \ \leq \ V.$$

Dieser Vektorraum $\{\overrightarrow{P_0Q} \mid Q \in B\}$ ist offenbar unabhängig von der getroffenen Auswahl von $P_0 \in B$ und kann deshalb mit $V(B)$ bezeichnet werden, er heißt auch die Richtung *von B.*

In Anlehnung an die Elementargeometrie heißen eindimensionale affine Unterräume Geraden, *zweidimensionale heißen* Ebenen. *$(n-1)-$dimensionale Unterräume — n die Dimension von A — heißen* Hyperebenen.

iii) *Sind $B_0, \ldots, B_m$ affine Unterräume von A, dann heißt der kleinste affine Unterraum, der $\cup_i B_i$ umfaßt, der* Verbindungsraum *der B_i. Er wird mit*

$$B_0 \vee \ldots \vee B_m$$

bezeichnet. Beispielsweise sind Punkte $P, Q \in A$ (nulldimensionale) affine Unterräume. Der Verbindungsraum $P \vee Q$ heißt deren Verbindungsgerade, *usw.*

iv) *Zwei Unterräume B_0, B_1 eines affinen Punktraumes A heißen* parallel, *kurz*

$$B_0 || B_1,$$

wenn die Richtung eines von ihnen in der Richtung des anderen enthalten ist:

$$\exists\ i\colon\ V(B_i) \subseteq V(B_j),\ j \neq i.$$

(B_0 und B_1 sind dann offenbar disjunkt oder einer ist im anderen enthalten.) Disjunkte und nicht zueinander parallele Unterräume heißen windschief.

v) Eine Abbildung $f: A \to A: P \mapsto P'$ heißt affine Abbildung, *wenn*

$$\overrightarrow{P_1Q_1} = \overrightarrow{P_2Q_2} \Longrightarrow \overrightarrow{P_1'Q_1'} = \overrightarrow{P_2'Q_2'}$$

und die von f induzierte Abbildung

$$\varphi: V \to V: \overrightarrow{PQ} \mapsto \overrightarrow{P'Q'}$$

linear ist, also die Abbildung φ, die folgendes Diagramm kommutativ ergänzt:

$$\begin{tikzcd}
A \times A \arrow[r, "\overrightarrow{\ }"] \arrow[dr, "\overrightarrow{\ } \circ (f \times f)"'] & V \arrow[d, "\varphi"] \\
 & V
\end{tikzcd}$$

•

4.1.8 Beispiele Für irgend zwei Punkte aus einem affinen Raum A zu einem Vektorraum V über $GF(2)$ gilt:

$$P \vee Q = \{P, Q\}.$$

Demnach enthält jede Teilmenge eines solchen affinen Unterraums mit zwei Punkten auch deren Verbindungsgerade. Es ist jedoch nicht jede Teilmenge eines affinen Raums zu $GF(2)^2$ ein affiner Unterraum. Das liegt an der Charakteristik von $GF(2)$, denn bei $\mathrm{char}(\mathbb{K}) \neq 2$ folgt aus der Tatsache, daß $B \subseteq A$ nicht leer ist und mit je zwei Punkten auch deren Verbindungsgerade enthält, daß B ein affiner Unterraum ist (vgl. Übungsblatt). ◇

4.1.9 Hilfssatz *Zu jedem $\varphi \in End_{\mathbf{K}}(V)$ und jedem Punktepaar $(O, O') \in A^2$ gibt es genau eine affine Abbildung $f: A \to A$, mit $f(O) = O'$, die gemäß 4.1.7 iv) die lineare Abbildung φ induziert, nämlich*

$$f: P \mapsto P', \ \textit{mit}\ \overrightarrow{OP'} = \overrightarrow{OO'} + \varphi(\overrightarrow{OP}).$$

Mit Hilfe von Ortsvektoren $p = \overrightarrow{OP}$ läßt sich diese Abbildung auch wie folgt schreiben:

$$f: x \mapsto x' = \varphi(x) + t, \ \textit{mit}\ t := \overrightarrow{OO'}.$$

Affine Abbildungen $f: x \mapsto x' = \varphi(x) + t$ mit $\varphi = \mathrm{id}_V$ heißen *Translationen.* Offenbar ist die affine Abbildung f genau dann injektiv, surjektiv, bijektiv, wenn dies für die induzierte lineare Abbildung φ gilt. Die Inverse f^{-1} einer affinen Abbildung f induziert die Inverse φ^{-1} der von f induzierten linearen Abbildung

φ. Die bijektiven affinen Abbildungen ergeben, zusammen mit der Komposition als Verknüpfung, eine Gruppe, die *affine Gruppe*

$$\mathrm{Aff}(A)$$

von A.
Unter den Teilmengen affiner Räume sind solche der folgenden Form besonders wichtig:

4.1.10 Definition (allgemeine Lage, Parallelepiped, Simplex) *A sei ein affiner Raum, $P_0, \ldots, P_m \in A$.*

i) Die P_ν heißen Punkte in allgemeiner Lage, *wenn sie in keinem $(m-1)-$dimensionalen Unterraum liegen, d.h. wenn die Folge $(\overrightarrow{P_0P_1}, \ldots, \overrightarrow{P_0P_1})$ linear unabhängig ist, bzw. wenn*

$$dim(P_0 \vee P_1 \vee \ldots \vee P_m) = m.$$

ii) Sind $P_0, \ldots, P_p$, Punkte in allgemeiner Lage und ist $\mathbb{K} = \mathbb{R}$, dann heißt

$$\{P_0 + \sum_{i=1}^{m} \rho_i \overrightarrow{P_0P_i} \mid 0 \leq \rho_i \leq 1\}$$

das von den P_i aufgespannte Parallelepiped. *Die Teilmenge*

$$\{P_0 + \sum_{i=1}^{m} \rho_i \overrightarrow{P_0P_i} \mid 0 \leq \rho_i \leq 1, \sum_i \rho_i = 1\}$$

heißt das von den P_i aufgespannte Simplex.

•

4.2 Isometrien und Rotationen

Affine Räume A mit euklidischen Vektorräumen $V(A)$ nennt man *euklidische Räume.* Die hier definierte Metrik ergibt eine *Distanzfunktion*

$$\rho(P,Q) := \|\overrightarrow{PQ}\|,$$

und auch die anderen für euklidische Vektorräume behandelten metrischen Konzepte erfahren eine geometrische Interpretation. Affine Abbildungen, die die Distanzen erhalten, heißen *starre Bewegungen.* Die zugehörigen linearen φ gehören zu der folgenden Klasse linearer Abbildungen:

4.2.1 Definition (Isometrie) *$f \in Hom_{\mathbb{R}}(V,W)$ heißt* Isometrie, *wenn f das innere Produkt* invariant läßt: $\langle f(v_1) \mid f(v_2)\rangle = \langle v_1 \mid v_2\rangle$.

Aus den vorangegangenen Überlegungen über lineare Abbildungen zwischen euklidischen Vektorräumen ergeben sich leicht die folgenden Eigenschaften von Isometrien:

4.2.2 Satz

i) Genau die die Norm erhaltenden linearen Abbildungen sind Isometrien.

ii) Isometrien sind injektiv.

iii) Isometrien f zwischen euklidischen Vektorräumen derselben Dimension sind Isomorphismen, und für sie gilt $\tilde{f} = f^{-1}$.

iv) Gilt für $f \in Hom_{\mathbb{R}}(V,W)$ die Gleichung $\tilde{f} = f^{-1}$, dann ist f Isometrie.

v) Isometrien zwischen euklidischen Vektorräumen derselben Dimension bilden Orthonormalbasisfolgen auf Orthonormalbasisfolgen ab, und umgekehrt sind Isomorphismen mit dieser Eigenschaft Isometrien.

vi) Ist $f \in \mathrm{End}_{\mathbb{R}}(V)$ eine Isometrie, dann ist

$$\det(f) \in \{1,-1\}$$

und Eigenwerte sind ggf. auch aus der Menge $\{1,-1\}$.

4.2.3 Definition (Rotation) *$f \in End_{\mathbb{R}}(V)$ heißt* Rotation, *wenn f isometrisch ist. Die Rotationen f mit* $\det(f) = 1$ *heißen dabei* eigentliche, *die anderen* uneigentliche *Rotationen.*

Es gibt natürlich auch Rotationen, die überhaupt keine Eigenwerte im vorgegebenen Grundkörper haben, z. B. die Abbildung f mit

$$M(\mathcal{E},f,\mathcal{E}) := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Sie beschreibt eine eigentliche Rotation des euklidischen Vektorraums $\mathbb{R}^2$ mit Standardskalarprodukt $\langle v \mid w\rangle := \sum v_i w_i$ und besitzt keinerlei Eigenwert in $\mathbb{R}$. Aus den Ergebnissen über die Existenz von Eigenwerten im Reellen, die mittels Zwischenwertsatz erzielt worden waren, erhalten wir

4.2.4 Folgerung *Eigentliche (uneigentliche) Rotationen von Räumen ungerader Dimension besitzen einen Eigenwert* $1(-1)$. *Uneigentliche Rotationen auf Räumen gerader Dimension besitzen die beiden Eigenwerte* ± 1.

Eine weitere leicht einzusehende Folgerung formuliert

4.2.5 Hilfssatz *Ist* $U \leq_{\mathbb{R}} V$ *invariant unter der Rotation* f, *dann auch* $U^{\perp}$.

Sind $f, g \in End_{\mathbb{R}}(V)$ Rotationen, dann auch f^{-1} und $g \circ f^{-1}$. Die Rotationen bilden also eine Untergruppe $0(V)$ in der Gruppe $GL(V)$ der regulären linearen Automorphismen von V, sie heißt die *(volle) orthogonale Gruppe* von V, denn die Rotationen sind, wegen $\tilde{f} = f^{-1}$ orthogonale Abbildungen:

$$O(V) \leq GL(V).$$

Den Kern der Abbildungen $\det : GL(V) \to \mathbb{R}^*$ bilden die eigentlichen Rotationen, er heißt die *spezielle orthogonale Gruppe* und wird mit $SO(V)$ bezeichnet. Falls $V \neq 0_V$ ist, gilt

4.2.6

i) $SO(V) := \{f \in O(V) \mid \det(f) = 1\} \triangleleft O(V)$,

ii) $|O(V)/SO(V)| = 2$.

Es soll nun gezeigt werden, daß eine Rotation $f \in End_{\mathbb{R}}(V)$ eine direkte Zerlegung von V ergibt in invariante und orthogonale Unterräume der Dimension 1 oder 2.

4.2.7 Satz *Ist* $f \in Aut_{\mathbb{R}}(V)$ *eine Rotation, dann gibt es eine Orthonormalbasisfolge* $\mathcal{B}$ *von* V *mit*

$$M(\mathcal{B}, f, \mathcal{B}) = \begin{pmatrix} \epsilon_0 & & & & & & & 0 \\ & \ddots & & & & & & \\ & & \epsilon_{m-1} & & & & & \\ & & & \cos(\omega_1) & \sin(\omega_1) & & & \\ & & & -\sin(\omega_1) & \cos(\omega_1) & & & \\ & & & & & \ddots & & \\ & & & & & & \cos(\omega_r) & \sin(\omega_r) \\ 0 & & & & & & -\sin(\omega_r) & \cos(\omega_r) \end{pmatrix},$$

wobei $\epsilon_\nu \in \{1, -1\}, 0 \leq \nu \leq m-1$.

Beweis: i) Wir bemerken zunächst, daß die Summe der Eigenräume $E(1), E(-1)$ direkt ist:

$$U := E(1) + E(-1) = E(1) \oplus E(-1).$$

Nach dem letzten Hilfssatz ist die entsprechende Zerlegung

$$V = U \oplus U^{\perp} = E(1) \oplus E(-1) \oplus (E(1) \oplus E(-1))^{\perp}$$

eine Zerlegung in invariante Unterräume. Die Wahl einer an U angepaßten Basisfolge $\mathcal{B}$ ergibt den ersten diagonalen Teil der Matrix von f mit den Einträgen ϵ_i, wie behauptet:

$$M(\mathcal{B}, f, \mathcal{B}) = \left(\begin{array}{cccccc|c} 1 & & & & & 0 & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & 0 \\ & & & & \ddots & & \\ 0 & & & & & -1 & \\ \hline & & & 0 & & & * \end{array}\right).$$

ii) Da die Einschränkung $f \downarrow U^{\perp} = f \downarrow (E(1) \oplus E(-1))^{\perp}$ keinen Eigenwert besitzt, hat dieser Raum gerade Dimension. Wir zeigen, daß dieser Unterraum eine unter f invariante Ebene enthält:
Die Einschränkung F von $f + \tilde{f} = f + f^{-1}$ auf $(E(1) \oplus E(-1))^{\perp}$ ist selbstadjungiert, besitzt also einen Eigenvektor e, λ sei der zugehörige Eigenwert. Da f in $(E(1) \oplus E(-1))^{\perp}$ keinen Eigenvektor besitzt, sind e und $f(e)$ linear unabhängig:

$$W := \langle e, f(e) \rangle \Longrightarrow \dim_{\mathbb{R}}(W) = 2.$$

Diese Ebene ist invariant unter f, denn

$$F(e) = f(e) + f^{-1}(e) = \lambda \cdot e \Longrightarrow f^2(e) = \lambda \cdot f(e) - e.$$

Die Einschränkung von f auf W ist eine eigentliche Rotation, $W^{\perp}$ also ebenfalls invariant unter f, so daß wir eine verfeinerte Zerlegung von V in invariante Unterräume erhalten:

$$V = U \oplus W \oplus W^{\perp}.$$

Wir haben also von $U^{\perp}$ einen invarianten Unterraum der Dimension 2 abspalten können, so daß die Einschränkung von f auf diesen — wegen der geraden Dimension — eine *eigentliche* Rotation ist. Dies können wir iterieren und erhalten die Behauptung damit per Induktion und mit dem folgenden Hilfssatz. □

4.2.8 Hilfssatz *Ist V zweidimensional euklidisch, Δ die normierte Determinantenform, die eine Orientierung repräsentiert, sowie $f \in End_{\mathbb{R}}(V)$ definiert durch $\langle f(-) \mid - \rangle = \Delta(v, w)$, dann ist f eigentliche Rotation mit $f^2 = -id$.*

Beweis: i) Wir beweisen als erste die Identität $\tilde{f} = -f$:

$$\langle v \mid \tilde{f}(w) \rangle = \langle f(v) \mid w \rangle = \Delta(v, w) = -\Delta(w, v) = -\langle f(w) \mid v \rangle = -\langle v \mid f(w) \rangle.$$

ii) Es zeigt sich jetzt, daß f eine Isometrie ist:

$$\langle v \mid w\rangle^2 + \langle f(v) \mid w\rangle^2 = \langle v \mid w\rangle^2 + \Delta(v,w)^2 =_{3.8.24} ||v||^2||w||^2.$$

Setzen wir jetzt $w := f(v)$, so ergibt sich (mit $\langle v \mid f(v)\rangle = 0$, nach i)):

$$||f(v)||^4 = \langle f(v) \mid f(v)\rangle = ||f(v)||^2||v||^2.$$

Das ergibt $||f(v)|| = ||v||$, f erhält also die Norm und ist demnach Isometrie.
iii) Zum Nachweis von $f^2 = -id$:

$$f^2 = f \circ f =_{i)} -f \circ \tilde{f} =_{ii)} -f \circ f^{-1} = -\mathrm{id}.$$

iv) Mit Hilfe von iii) erhalten wir schließlich

$$\det(f) \cdot \Delta(v,w) = \Delta(f(v), f(w)) = \langle f^2(v) \mid f(w)\rangle = -\langle v \mid f(w)\rangle = \Delta(v,w),$$

was $\det(f) = 1$ ergibt und den Beweis vervollständigt. □
Diese Abbildung f heißt auch die *kanonische Komplexstruktur* der euklidischen Ebene V. Sei nun φ eine eigentliche Drehung von V, $v \in V\backslash\{0_V\}$ und ω der orientierte Winkel zwischen v und $\varphi(v)$. Für ihn gilt

4.2.9 $$\cos(\omega) = \frac{\langle v \mid \varphi(v)\rangle}{||v||||\varphi(v)||},\ \sin(\omega) = \frac{\Delta(v,\varphi(v))}{||v||||\varphi(v)||} = \frac{\langle f(v) \mid \varphi(v)\rangle}{||v||||\varphi(v)||}.$$

Da, für jedes $v \in V$ der Norm 1, $(v, f(v))$ eine ON-Basisfolge ist, gilt also für jede eigentliche Rotation φ und die kanonische Komplexstruktur f zu der Orientierung, bei der diese ON-Basisfolge positiv orientiert ist:

4.2.10 $$\varphi(v) = v \cdot \cos(\omega) - f(v) \cdot \sin(\omega).$$

Mit $f^2 = -\mathrm{id}_V$ ergibt sich daraus

4.2.11 $$\varphi(f(v)) = f(v) \cdot \cos(\omega) + v \cdot \sin(\omega).$$

Das ergibt die behauptete Matrixform:

$$M((v,f(v),\varphi,(v,f(v)) = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix}$$

Tatsächlich sind diese Matrixelemente *unabhängig* von der Wahl von v :

4.2.12 Hilfssatz $\cos(\omega) = \frac{1}{2}\mathrm{Spur}(\varphi)$, $\sin(\omega) = \frac{1}{2}\mathrm{Spur}(f \circ \varphi)$.

Beweis: Die Spur (vgl. Übungsblatt) von φ erfüllt (vgl. Übungsblatt) mit jeder Determinantenform $\Delta \neq 0$ die Gleichung

$$\Delta(\varphi(v), w) + \Delta(v, \varphi(w)) = \mathrm{Spur}(\varphi) \cdot \Delta(v,w).$$

Hieraus folgt (Einsetzen von 4.2.10 und 4.2.11):

$$\mathrm{Spur}(\varphi)\Delta(v, f(v)) = 2\cos(\omega)\Delta(v, f(v)),$$

also der erste Teil der Behauptung, der zweite folgt analog. □
Der Winkel ω kann deshalb als der *Drehwinkel von* φ bezeichnet werden, kurz

$$\omega(\varphi) := \omega.$$

Es gilt also

$$\varphi(v) = \cos(\omega(\varphi)) \cdot v - \sin(\omega(\varphi)) f(v),$$

und

$$\varphi(f(v)) = \sin(\omega(\varphi)) v + \cos(\omega(\varphi)) f(v).$$

Insbesondere haben wir $\omega(\mathrm{id}) = 0$, $\omega(-\mathrm{id}) = \pi$, $\omega(f) = -\pi/2$. Für die Komposition zweier Rotationen gilt

4.7.18 $(\varphi \circ \psi)(v) = (\psi \circ \varphi)(v) = v \cdot cos(\omega(\varphi) + \omega(\psi)) + f(v) \cdot sin(\omega(\varphi) + \omega(\psi)).$

Für die Beschreibung von Rotationen in dreidimensionalen euklidischen Räumen gehen wir vom allgemeinen Satz über die Form von Rotationen φ aus:

$$M(\mathcal{B}, \varphi, \mathcal{B}) = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos(\omega) & \sin(\omega) \\ 0 & -\sin(\omega) & \cos(\omega) \end{pmatrix}.$$

Links oben steht ± 1, je nachdem ob φ eine eigentliche oder eine uneigentliche Rotation beschreibt. Dabei ist $\mathcal{B} = (b_0, b_1, b_2)$ eine ON-Basisfolge, der Unterraum $\langle b_0 \rangle$ heißt *Drehachse* von φ. Zu ihrer Ermittlung betrachtet man die Abbildung

$$\psi := \frac{1}{2}(\varphi - \tilde{\varphi})$$

und argumentiert wie folgt (vgl. Übungsblatt):

i) Zu $v \in V$ ist $F_v \colon w \mapsto v \times w$ schiefsymmetrisch, d. h. $\tilde{F}_v = -F_v$.

ii) Es gilt
$$F_{v \times w} = F_v \circ F_w - F_w \circ F_v.$$

iii) Ist $f \colon V \to V$ schiefsymmetrisch, dann gibt es genau ein $v \in V$ mit $f = F_v$.

iv) Zu obigem ψ gibt es also genau ein $u \in V$ mit $\psi = F_u$. u heißt der *Drehvektor*, und es gilt
$$\psi(v) = u \times v.$$
Diese Gleichung kann zur Bestimmung des Drehvektors verwendet werden.

v) Der Drehvektor liegt auf der Drehachse, ist Vielfaches von b_0.

vi) Der Cosinus des Drehwinkels ω kann aus der Spur von φ ermittelt werden, der Sinus erfüllt die Gleichung
$$\sin(\omega) = \|u\|.$$

Vgl. Übungsblatt.

4.3 Symmetrische Bilinearformen

Alle betrachteten Vektorräume seien nach wie vor Vektorräume über $\mathbb{R}$. Wir fragen jetzt, was auch ohne die Voraussetzung noch gilt, daß $\langle - \mid - \rangle$ positiv definit sei, d.h. wir betrachten Bilinearformen

$$\Phi: V \times V \to \mathbb{R},$$

von denen wir nur voraussetzen, daß sie symmetrisch sind. Ist V insbesondere von endlicher Dimension, $\mathcal{B} = (b_0, \ldots, b_{n-1})$ eine Basisfolge, dann können wir zur Beschreibung von Φ wieder die Matrix

4.3.1 $$A := M_{\mathcal{B}}^{\Phi} := (\Phi(b_i, b_k))$$

benutzen, sie heißt die *Formmatrix.* Wegen der Symmetrie von Φ ist sie symmetrisch, $A = {}^tA$, und es gilt (wenn $v = \sum v_i b_i, w = \sum w_i b_i$) :

4.3.2 $$\Phi(v, w) = {}^tv \cdot M_{\mathcal{B}}^{\Phi} \cdot w = \sum_{i,k} a_{ik} v_i w_k.$$

Zu $\Phi: V \times V \to \mathbb{R}$ sei $\varphi: V \to \mathbb{R}$ definiert durch

4.3.3 $$\varphi(v) := \Phi(v, v).$$

Hierfür gilt, da Φ bilinear und symmetrisch ist:

4.3.4 $$\Phi(v, w) = \frac{1}{2}[\varphi(v + w) - \varphi(v) - \varphi(w)].$$

Φ ist also durch φ eindeutig bestimmt. Aus 4.3.4 folgt die *Parallelogramm-Identität:*

4.3.5 $$\varphi(v + w) + \varphi(v - w) = 2[\varphi(v) + \varphi(w)].$$

Stetige Funktionen $\psi: V \to \mathbb{R}$ mit 4.3.5 heißen *quadratische Formen.* Jede symmetrische Bilinearform liefert also eine quadratische Form, denn φ ist stetig (vgl. den Beweis von 3.9.8), also quadratische Form. Es gibt aber auch die Umkehrung:

4.3.6 Hilfssatz *Jede quadratische Form ergibt sich gemäß 4.3.3 aus einer symmetrischen Bilinearform.*

Beweis: Ist $\varphi: V \to \mathbb{R}$ quadratische Form, so ergibt sich daraus gemäß 4.3.4 eine Abbildung $\Phi: V \times V \to \mathbb{R}$, die als symmetrisch und bilinear nachgewiesen werden muß .

i) Die Symmetrie folgt unmittelbar aus 4.3.4.

ii) $\Phi(v_1 + v_2, w) = \Phi(v_1, w) + \Phi(v_2, w)$ folgt aus 4.3.4, 4.3.5 nach leichten Rechnungen.

iii) $\Phi(-v,w) = -\Phi(v,w)$ ergibt sich aus ii) und liefert

$$\forall\ z \in \mathbb{Z}, v, w\colon \quad \Phi(zv,w) = z\Phi(v,w),$$

woraus sich unmittelbar ergibt:

$$\forall\ q \in \mathbb{Q}, v, w\colon \quad \Phi(qv,w) = q\Phi(v,w).$$

Jetzt ergibt sich mit der rationalen Approximierbarkeit reeller Zahlen aus der Stetigkeit von φ schließlich, daß

$$\forall\ r \in \mathbb{R}, v, w\colon \quad \Phi(rv,w) = r\Phi(v,w),$$

womit alles bewiesen ist, denn $\Phi(v,v) = \varphi(v)$. □

Insgesamt gilt also

4.3.7 Folgerung *4.3.3 und 4.3.4 definieren Bijektionen zwischen den Mengen der symmetrischen Bilinearformen und der quadratischen Form auf V.*

Für φ gilt entsprechend 4.3.2:

4.3.8 $$\varphi(v) = \sum_{i,k} a_{ik} v_i v_k.$$

Es soll nun gezeigt werden, daß V durch die Bilinearform Φ eine direkte Zerlegung $V = V_0 \oplus V_+ \oplus V_-$ erfährt in den Nullraum V_0 von V und die Unterräume $V_\pm$ wo Φ positiv bzw. negativ definit ist. Es war

4.3.9 $$V_0 := N_V = \{w \mid \forall\ v \in V\colon \Phi(v,w) = 0\} \leq V.$$

Die Differenz $\dim_{\mathbb{R}}(V) - \dim_{\mathbb{R}}(V_0)$ heißt auch der *Rang* von Φ. Ist V^* dual zu V bzgl. $\langle - \mid - \rangle\colon V \times V^* \to \mathbb{R}$ und $f\colon V \to V^*$ definiert durch

4.3.10 $$\Phi(v,w) = \langle v \mid f(w)\rangle,$$

dann gilt (nachrechnen) mit der zu $\mathcal{B}$ dualen Basisfolge $\mathcal{B}^*$:

4.3.11 Hilfssatz

i) $V_0 = \mathrm{Kern}(f)$,

ii) $\mathrm{Rang}(\Phi) = \mathrm{Rang}(f) = \mathrm{Rang}(M_{\mathcal{B}}^{\Phi})$,

iii) $M_{\mathcal{B}}^{\Phi} = M(\mathcal{B}^*, f, \mathcal{B})$,

iv) Φ *nicht ausgeartet* $\iff \det(M_{\mathcal{B}}^{\Phi}) \neq 0$.

4.3.12 Definition (definit, semidefinit, indefinit) *Ist $\Phi\colon V \times V \to \mathbb{R}$ symmetrische Bilinearform, dann heißt Φ*

i) positiv definit, *wenn*

$$\forall\ v \neq 0\colon\ \varphi(v) > 0,$$

ii) positiv semidefinit, *falls*

$$\forall\ v \neq 0\colon\ \varphi(v) \geq 0 \wedge\ \exists\ v \neq 0\colon\ \varphi(v) = 0.$$

Entsprechend sind negativ definit *und* negativ semidefinit *zu verstehen. Nimmt* φ *dagegen sowohl positive als auch negative Werte an, dann heißt* Φ indefinit.

•

Ist Φ positiv definit oder positiv semidefinit, dann gilt die Cauchy- Schwarzsche Ungleichung, im letzten Fall jedoch ohne den Zusatz, daß Gleichheit nur bei linearer Abhängigkeit auftritt. Semidefinite Bilinearformen sind stets ausgeartet, indefinite nicht notwendig. Sei nun Φ eine nicht ausgeartete indefinite symmetrische Bilinearform auf einem euklidischen Vektorraum V. Es gibt dann offenbar Unterräume $\neq \{0_V\}$, auf denen Φ positiv definit ist, V_+ sei ein solcher von maximaler Dimension, V_- dessen orthogonales Komplement bzgl. (dem Skalarprodukt!) Φ. Es gilt also

4.3.13 $$V = V_+ \oplus V_-.$$

4.3.14 Hilfssatz *Die Einschränkung* $\Phi \downarrow V_- \times V_-$ *ist negativ definit.*

Beweis: Wir beweisen zunächst indirekt, daß $\varphi(v) \leq 0$, für alle $v \in V_-$. Für $v \in V_-\backslash\{0_V\}$ sei $W := V_+ + \langle v\rangle$. Dann gilt, für alle $w = u + \rho \cdot v \in W$:

$$\varphi(w) = \varphi(u) + \rho^2\varphi(v).$$

Die Annahme $\varphi(v) > 0$ ergäbe also $\varphi(w) > 0$, im Widerspruch zur Maximalität von V_+.
Zum Nachweis der Behauptung $\Phi \downarrow V_- \times V_-$ sei negativ definit benutzen wir, daß für $v, v' \in V_-$ gilt (Cauchy-Schwarz):

$$\Phi(v, v')^2 \leq \varphi(v)\varphi(v').$$

Falls $\varphi(v') = 0$, ist also $\Phi(v, v') = 0$, für alle $v \in V_-$. Wegen $V_- = V_+^\perp$ und 4.3.13 gilt dann sogar $\Phi(v, v') = 0$ für *alle* $v \in V$. Φ war aber als nicht ausgeartet vorausgesetzt, so daß $v' = 0$ folgt. □
Falls Φ ausgeartet ist, haben wir noch den Nullraum V_0 zu berücksichtigen. Φ ist dann nicht ausgeartet auf jedem Komplement zu V_0, so daß wir darauf 4.3.14 anwenden können und insgesamt erhalten:

4.3.15 Satz *Ist* Φ *symmetrische Bilinearform auf dem endlichdimensionalen* $\mathbb{R}-$*Vektorraum* V*, dann gibt es Unterräume* $V_\pm$ *von* V*, so daß* Φ *auf* V_+ *bzw.* V_- *positiv bzw. negativ definit ist und mit dem Nullraum* V_0 *folgendes gilt:*

$$V = V_0 \oplus V_+ \oplus V_-.$$

Wählen wir eine an eine solche Zerlegung von V angepaßte ON-Basisfolge $\mathcal{B} = (b_0, \ldots, b_{n-1})$ so, daß $(b_0, \ldots, b_{s-1})$ Orthonormalbasisfolge von V_+ ist bzgl. Φ, $(b_s, \ldots, b_{s+r-1})$ Orthonormalbasisfolge von V_- bzgl. $-\Phi$ und $(b_{s+r-1}, \ldots, b_{n-1})$ eine Basisfolge für V_0, dann ist

4.3.16
$$M_{\mathcal{B}}^{\Phi} = \begin{pmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ 0 & & & & & & & & 0 \end{pmatrix},$$

Hierbei sind die Anzahlen der Einsen s, der sogenannte *Index* von Φ, und r, die Anzahl der Diagonalelemente -1, eindeutig bestimmt. (Die Zahl $2s - r$ nennt man die *Signatur* von Φ.) Nur scheinbar allgemeiner ist die unmittelbare Folgerung

4.3.17 Der Trägheitssatz von Sylvester *Transformiert man die Formmatrix auf Diagonalform, dann erhält man stets dieselbe Anzahl positiver Elemente auf der Hauptdiagonalen.*

4.3.18 Folgerung *Ist φ quadratische Form auf einem endlichdimensionalen $\mathbb{R}-$Vektorraum, dann ist die zugehörige symmetrische Bilinearform Φ genau dann positiv definit, wenn es eine Formmatrix von φ mit Diagonalgestalt und lauter positiven Diagonalelementen gibt.*

Die Berechnung von Rang und Index von Φ kann wie folgt geschehen. Ist

$$\varphi(v) = \sum_{i,k=1}^{n} a_{ik} v_i v_k$$

und $\Phi \neq 0$, dann ist mindestens ein $a_{ik} \neq 0$. Falls dabei $i \neq k$ gilt (andernfalls liegt ja bereits eine Diagonalmatrix vor), substituiert man

$$v_i = \overline{v}_i + \overline{v}_k, v_k = \overline{v}_i - \overline{v}_k,$$

was

$$\varphi(v) = \sum_{i,k=1}^{n} \overline{a}_{ik} \overline{v}_i \overline{v}_k$$

ergibt mit $\overline{a}_{ii} \neq 0 \neq \overline{a}_{kk}$. Wir können also annehmen, daß mindestens ein Diagonalterm $a_{ii} \neq 0$ ist, sei dies etwa a_{11}. $\varphi(v)$ kann dann geschrieben werden als

$$\begin{aligned}\varphi(v) &= \underbrace{a_{11}\,[v_1^2 + \frac{2}{a_{11}}\sum_{k=2}^{n} a_{1k} v_i v_k]}_{\overline{v}_1^2} + \sum_{i,k=2}^{n} a_{ik} \underbrace{v_i v_k}_{=:\overline{v}_i\overline{v}_k} \\ &= a_{11}\overline{v}_1^2 + \sum_2^n a_{ik}\overline{v}_i\overline{v}_k .\end{aligned}$$

Die Summe auf der rechten Seite beschreibt aber eine symmetrische Bilinearform auf einem $(n-1)-$dimensionalen $\mathbb{R}-$Vektorraum, so daß sich iterativ schließlich eine Diagonalmatrix ergibt.
$\mathbb{R}-$Vektorräume mit nicht ausgearteten indefiniten Bilinearformen Φ treten u.a. in Untersuchungen im Rahmen der Relativitätstheorie auf, sie heißen *pseudo-euklidisch.* Hierbei heißen die Vektoren $v \in V$ mit $\Phi(v,v) > 0$ *raumartig,* solche mit $\Phi(v,v) < 0$ *zeitartig,* die mit $\Phi(v,v) = 0$ heißen *Lichtvektoren* (sie bilden den sogenannten *Lichtkegel.* Ist z.B. $\dim_{\mathbb{R}}(V) = 2$ und 1 der Index von Φ, dann besteht der Lichtkegel aus 2 Geraden. Die Vektoren in $T^{\pm}$ bilden den *Zukunfts-* bzw. den *Vergangenheitskegel.* Diese Konstruktionen dienen also vor allem der Unterscheidung von zeitlichen und räumlichem Abstand zweier Ereignisse.

4.3.19 Definition (Minkowski-Raum, Lorentz-Transformation) *Endlichdimensionale $\mathbb{R}-$ Vektorräume zusammen mit nicht ausgearteten indefiniten Bilinearformen $\langle - \mid - \rangle$ heißen* pseudo-euklidisch, *die Bilinearform heißt dann ebenfalls das* innere Produkt, *ihr Index der* Index *s des pseudo-euklidischen Raums. Eine Basisfolge $\mathcal{B}$ heißt dabei* orthonormal, *wenn gilt*

$$\langle b_i \mid b_k\rangle = \epsilon_i\delta_{ik}\,, \epsilon_i = \begin{cases} 1, & i \in s \\ -1, & i \in n\backslash s \end{cases}$$

Wir haben bereits gesehen, daß solche Orthonormalbasisfolgen stets existieren. Bezüglich einer solchen gilt

$$\langle v \mid w\rangle = \sum_{i=1}^{s} v_i w_i - \sum_{i=s+1}^{n} \forall_i w_i ,$$

und die Gleichung des Lichtkegels ist

$$\sum_{i=1}^{s} v_i^2 - \sum_{i=s+1}^{n} v_i^2 = 0.$$

Interessant sind insbesondere die pseudo-euklidischen Räume V vom Index $n - 1 := dim_{\mathbb{R}}(V) - 1$. Unter diesen heißen die mit $n = 4$ Minkowski-Räume, *deren Rotationen, also die Abbildungen $f \in End_{\mathbb{R}}(V)$ mit*

$$\langle f(v) \mid f(w)\rangle = \langle v \mid w\rangle$$

heißen Lorentz-Transformationen.

4.4 Quadriken

Wir wollen den Begriff des *Kegelschnitts* verallgemeinern und erinnern deshalb zunächst daran, daß Kegelschnitte in der euklidischen Ebene $\mathbb{E}^2 := D(\mathbb{R}^2)$ Mengen von Punkten $x \in \mathbb{E}^2$ sind, deren Koordinaten x_1, x_2 einer Gleichung der Form

4.4.1 $$\sum_{i,k=0}^{1} a_{ik} x_i x_k + 2\sum_{i=0}^{1} b_i x_i = c$$

genügen.

4.4.2 Beispiele

- Die Gleichung
$$\frac{x_0^2}{a^2} + \frac{x_1^2}{b^2} = 1$$
beschreibt eine *Ellipse* bzw. einen *Kreis* (falls $a = b$).
- Die Gleichung
$$\frac{x_0^2}{a^2} - \frac{x_1^2}{b^2} = 1$$
beschreibt eine *Hyperbel.*
- Die Gleichung $x_1^2 = 2px_0$ beschreibt eine *Parabel.*

Diese Kegelschnitte werden also durch quadratische Formen und Linearformen beschrieben. Wir verallgemeinern deshalb wie folgt:

4.4.3 Definition (Quadrik) *Unter einer* Quadrik *im euklidischen Raum* $\mathbb{E}^n := D(\mathbb{R}^n)$ *versteht man eine Menge* Q *von Punkten, die die Lösungsmenge einer Gleichung der Gestalt*
$$\varphi(x) + 2f(x) = \alpha$$
bilden, mit einer quadratischen Form φ, *einer Linearform* f *und einem* $\alpha \in \mathbb{R}$.

4.4.4 Hilfssatz *Quadriken im* $\mathbb{E}^n$ *sind genau die Lösungsmengen von Gleichungen der Form*
$$\langle F(x) \mid x\rangle + 2\langle a \mid x\rangle = \alpha$$
mit selbstadjungiertem $F \in End_{\mathbb{R}}(\mathbb{R}^n), a \in \mathbb{R}^n, \alpha \in \mathbb{R}$.

Beweis: Der Differenzenraum $\mathbb{R}^n$ ist euklidisch, $\langle - \mid -\rangle$ sei das innere Produkt. Dann kann nach 4.3.10 Φ, die symmetrische Bilinearform zu φ, mit einem geeigneten $F \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^n)$ so ausgedrückt werden:
$$\Phi(x,y) = \langle F(x) \mid y\rangle.$$
Nun ist aber Φ symmetrisch, ebenso $\langle - \mid -\rangle$, und wir erhalten
$$\langle F(x) \mid y\rangle = \Phi(x,y) = \Phi(y,x) = \langle F(y) \mid x\rangle = \langle x \mid F(y)\rangle = \langle \tilde{F}(x) \mid y\rangle.$$

Das impliziert $F = \tilde{F}$. Zudem ist jede Linearform $f \in Hom_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ von der Gestalt $f(-) = \langle - \mid a\rangle$, mit geeigneten $a \in \mathbb{R}^n$.

$\square$

Wir unterscheiden jetzt zwei Fälle:

Fall 1: $\exists\ s \in Q$: $F(s) + a = 0$. In diesem Fall kann a durch $-F(s)$ ersetzt werden. Es ergibt sich

$$\langle F(x) \mid x\rangle - 2\langle F(s) \mid x\rangle = \alpha.$$

Einsetzen von $x := s$ ergibt noch $\alpha = -\langle F(s) \mid s\rangle$, so daß wir insgesamt erhalten:

4.4.5 $$0 = \langle F(x) \mid x\rangle - 2\langle F(s) \mid x\rangle + \langle F(s) \mid s\rangle = \varphi(x - s),$$

wobei für die letzte Gleichung die Selbstadjungiertheit von F benutzt wird. Eine solche Quadrik heißt *Kegel* mit der *Spitze* s. Kegel mit Spitzen sind also genau die Quadriken der folgenden Form:

$$Q = \{x \mid \varphi(x - s) = 0\}.$$

Fall 2: $\forall\ x \in Q$: $F(x) + a \neq 0$. Das hat u.a. die Konsequenz, daß für alle Quadrikpunkte $x \in Q$ das orthogonale Komplement von $F(x)+a$, dem sogenannten *Normalenvektor* im Punkt x, eine Hyperebene in $\mathbb{R}^n$ ist, der *Tangentialraum* zum Punkt x :

4.4.6 $$T(x) := \langle F(x) + a\rangle^{\perp} = \{y \mid \langle F(x) + a \mid y\rangle = 0\}.$$

Trägt man die Elemente des Tangentialraums $T(x)$ von x aus ab, dann erhält man die Punkte der *Tangentialhyperebene* in x :

4.4.7 $$H(x) := x + T(x).$$

Mit Φ und f formuliert lesen sich diese Gleichungen so:

4.4.8 $$T(x) = \{y \in \mathbb{R}^n \mid \Phi(x, y) + f(y) = 0\} \leq \mathbb{R}^n,$$

4.4.9 $$H(x) = \{y \in \mathbb{E}^n \mid \Phi(x, y) + f(x + y) = \alpha\} \subseteq \mathbb{E}^n.$$

4.4.10 Satz *Ist x ein Punkt der Quadrik Q, dann ist der Schnitt von Q mit einer Ebene E durch z ein Kegelschnitt.*

Beweis: Sei $\varphi(x) + 2f(x) = \alpha$ wieder die Bestimmungsgleichung von Q und sei die Ebene E definiert durch

$$E := \{x \mid x = z + \xi u + \eta v, \xi, \eta \in \mathbb{R}\}$$

für zwei linear unabhängige $u, v \in \mathbb{R}^n$. Durch Einsetzen in die Bestimmungsgleichung ergibt sich

$$\xi^2\varphi(u) + 2\xi\eta\Phi(u, v) + \eta^2\varphi(v) + 2\xi(\Phi(u, z) + f(u)) + 2\eta(\Phi(v, z) + f(v)) = 0.$$

Die Menge der Vektoren (ξ, η), die dieser Gleichung genügen, ist ein Kegelschnitt.

$\square$

Verschiebt man den Koordinatenursprung im $\mathbb{E}^n$, d. h. ersetzt man x durch $z + x$, dann ergibt sich als neue Bestimmungsgleichung

4.4.11 $$Q = \{z + x \mid \varphi(x) + 2(\Phi(z, x) + f(x)) = \alpha - \varphi(z) - 2f(z)\}.$$

Um diese Bestimmungsgleichung zu vereinfachen, fragen wir, ob es solche z gibt, die die neue Linearform zum Verschwinden bringen, d. h. für die gilt

4.4.12 $$\Phi(z, x) + f(x) = \langle F(z) + a \mid x\rangle = 0,$$

für alle $x \in \mathbb{R}^n$. Solche $z \in \mathbb{E}^n$ heißen gegebenenfalls *Zentren.* Die Menge *aller* Zentren ist demnach

4.4.13 $$Z(Q) := \{z \in \mathbb{E}^n \mid F(z) = -a\}.$$

Zentren sind — als Lösungen der Gleichung $F(z) = -a$ — offenbar bis auf Vektoren aus Kern(F) eindeutig bestimmt. Die Menge der Zentren ist also leer oder ein affiner Unterraum von $\mathbb{E}^n$ mit Kern(F) als zugehörigem Differenzenraum. Ist Q *nicht degeneriert,* d. h. die Q beschreibende Bilinearform nicht ausgeartet (und damit F regulär), dann besitzt Q ggf. genau ein Zentrum. Hat Q ein Zentrum z, dann ist diese Quadrik von der Form

4.4.14 $$Q = \{x \in \mathbb{E}^n \mid \varphi(x - z) = \beta\},$$

und dabei ist β genau dann gleich Null, wenn Q ein Kegel mit Spitze ist (die Spitze ist dann das Zentrum). Wegen $\beta \neq 0$, falls die Quadrik kein Kegel mit Spitze ist, kann (vgl. 4.4.11) — außer bei Kegeln mit Spitze — kein Zentrum auf der Quadrik liegen, es gilt sogar noch mehr:

4.4.15 Hilfssatz *Die Zentren von Quadriken Q, die keine Kegel mit Spitzen sind, liegen in keiner Tangentialhyperebene.*

Beweis: Indirekt. Ist Q eine Quadrik mit Zentrum $z \in Q$, aber kein Kegel mit Spitze, dann ist die Bestimmungsgleichung für die $y \in H(x)$ nach 4.4.9:

$$0 \neq \beta = \Phi(x, y) = \langle F(x) \mid y\rangle.$$

Wegen $a = 0$ ist aber $y \in Z(Q)$ genau dann, wenn $F(y) = 0$. Das ergibt den Widerspruch

$$0 = \langle x \mid F(y)\rangle = \langle F(x) \mid y\rangle.$$

$\square$

Für die Lokalisierung von Zentren benutzt man zweckmäßig die Tatsache, daß jede Quadrik durch Punktspiegelung

$$\tau : x \mapsto 2z - x =: x'$$

an jedem ihrer Zentren in sich selbst übergeht:

$$\varphi(x' - z) = \varphi(2z - x - z) = \varphi(z - x) = \varphi(x - 1).$$

Aus dem Beweis des letzten Hilfssatzes ergibt sich noch mit 4.3.16 und durch Ersetzen von φ durch φ/β :

4.4.16 Folgerung *Eine Quadrik Q mit Zentrum, die kein Kegel mit Spitze ist, genügt einer Bestimmungsgleichung der Form $\varphi(x) = 1$, mit einer geeigneten quadratischen Form φ. Man nennt diese die* Normalform *der Bestimmungsgleichung. Wählt man noch eine Basisfolge $\mathcal{B}$ von $\mathbb{R}^n$ mit*

$$\Phi(b_i, b_k) = \epsilon_i \delta_{ik}, \epsilon_i := \begin{cases} 1, & 0 \leq i \leq s-1 \\ -1, & s \leq i \leq s+r-1 \\ 0, & s+r \leq i \leq n-1 \end{cases}$$

$s :=$ Index von φ, $r :=$ Rang von Φ, dann ist die Normalform der Bestimmungsgleichung von Q gerade diese:

$$\sum_{i=0}^{r-1} \epsilon_i x_i x_i = 1.$$

Beispiele solcher Normalformen von Quadriken mit Zentren in der Ebene sind:

Normalform	Typ
$x_0^2 + x_1^2 = 1$	Ellipse
$x_0^2 - x_1^2 = 1$	Hyperbel
$x_0^2 = 1$	zwei parallele Geraden

Im dreidimensionalen Anschauungsraum haben wir

Normalform	Typ
$x_0^2 + x_1^2 + x_2^2 = 1$	Ellipsoid
$x_0^2 + x_1^2 - x_2^2 = 1$	einschaliges Hyperboloid
$x_0^2 - x_1^2 - x_2^2 = 1$	zweischaliges Hyperboloid
$x_0^2 + x_1^2 = 1$	elliptischer Zylinder
$x_0^2 - x_1^2 = 1$	hyperbolischer Zylinder
$x_0^2 = 1$	zwei parallele Ebenen

Für Quadriken ohne Zentren kann man folgendes beweisen

4.4.17 Satz *Die Normalform einer Quadrik ohne Zentrum ist von der Form $\varphi(y) + 2\xi = 0$, bei Wahl einer Basis wie in 4.4.16 nimmt diese Bestimmungsgleichung die Form*

$$\sum_{i=1}^{r} \epsilon_i y_i y_i + 2\xi = 0$$

an.

4.5 Unitäre Räume

Es soll nun untersucht werden, wie man zweckmäßig vorzugehen hat, wenn der Grundkörper $\mathbb{K}$ der Körper

$$\mathbb{C} = \{\zeta := \rho + i\sigma \mid \rho, \sigma \in \mathbb{R}\}$$

der komplexen Zahlen ist, mit den bekannten Rechenregeln und der *Konjugation*

$$\overline{}\colon \mathbb{C} \to \mathbb{C}\colon \zeta \mapsto \overline{\zeta} := \rho - i\sigma.$$

4.5.1 Definition (Semibilinearformen) *Ist V ein $\mathbb{C}$-Vektorraum, dann heißt* $\Phi\colon V \times V \to \mathbb{C}$ Semibilinearform, *wenn gilt*

$$\Phi(\zeta_0 v_0 + \zeta_1 v_1, v_2) = \zeta_0 \Phi(v_0, v_2) + \zeta_1 \Phi(v_1, v_2)$$

und

$$\Phi(v_0, \zeta_1 v_1 + \zeta_2 v_2) = \overline{\zeta_1}\Phi(v_0, v_1) + \overline{\zeta_2}\Phi(v_0, v_2).$$

Die zugehörige quadratische Form φ ist wieder durch $\varphi(v) := \Phi(v, v)$ definiert, und auch hier gilt die Parallelogramm-Identität

$$\varphi(v + w) + \varphi(v - w) = 2[\varphi(v) + \varphi(w)].$$

Weiter haben wir, für $|\zeta| := \sqrt{\zeta\overline{\zeta}} = \sqrt{\rho^2 + \sigma^2}$:

4.5.2 $$\varphi(\zeta v) = |\zeta|^2 \varphi(v),$$

und auch hier im komplexen Fall läßt sich Φ aus der quadratischen Form φ wiedergewinnen:

4.5.3 $$2\Phi(v, w) = [\varphi(v + w) - \varphi(v) - \varphi(w)] + i \cdot [\varphi(v + iw) - \varphi(v) - \varphi(w)].$$

Zu jeder Semibilinearform Φ ergibt sich eine weitere Semibilinearform $\widetilde{\Phi}$ vermöge

4.5.4 $$\widetilde{\Phi}(v, w) := \overline{\Phi(w, v)}.$$

4.5.5 Definition (hermitesche Semibilinearform) *Die Semibilinearform Φ heißt* hermitesch, *wenn $\Phi = \widetilde{\Phi}$ gilt, also wenn*

$$\Phi(v, w) = \overline{\Phi(w, v)}.$$

4.5.6 Hilfssatz *Eine Semibilinearform ist genau dann hermitesch, wenn die zugehörige quadratische Form reellwertig ist.*

Demnach kann man bei hermiteschen Semibilinearformen auch von positiv definiten sprechen, sie eigenen sich also zur Metrisierung!

4.5.7 Definition (hermitesche Matrix) *$A \in \mathbb{C}^{n \times n}$ heißt* hermitesch, *wenn gilt*

$$a_{ik} = \overline{a_{ki}}.$$

Ist z. B. Φ hermitesch, $\mathcal{B} = (b_0, \ldots, b_{n-1})$ eine Basisfolge, dann ist $M_{\mathcal{B}}^{\Phi}$ hermitesch.

4.5.8 Beispiele

- Ist $V := \mathbb{C}^n$, dann ist das *Standardskalarprodukt* Φ, definiert durch

$$\Phi(v, w) := \sum_{i=o}^{n-1} v_i \overline{w}_i,$$

  hermitesch.

- Ein $\mathbb{R}$-Vektorraum V ergibt den $\mathbb{C}$–Vektorraum $V \times V$ vermöge

$$(\rho + i\sigma)(v, w) := (\rho v - \sigma w, \rho w + \sigma v).$$

  Dieser Vektorraum heißt die *Komplexifizierung* von V. Wir bezeichnen ihn mit ${}_{\mathbb{C}}(V \times V)$. Jeder seiner Vektoren (v, w) ist eindeutig darstellbar als Summe

$$(v, w) = (v, 0) + (0, w) = (v, 0) + i(w, 0).$$

  Ist dabei ${}_{\mathbb{R}}V$ euklidisch, $\langle - \mid - \rangle$ das innere Produkt, dann definiert

$$\Phi((v, v'), (w, w')) := \langle v \mid w \rangle + \langle v' \mid w' \rangle + i \cdot [\langle v' \mid w \rangle - \langle v \mid w' \rangle]$$

  eine hermitesche Semibilinearform auf ${}_{\mathbb{C}}(V \times V)$.

4.5.9 Definition (unitärer Raum, hermitesches inneres Produkt) *Ein endlichdimensionaler* $\mathbb{C}$*– Vektorraum zusammen mit einer positiv definiten hermiteschen Semibilinearform*

$$\langle - \mid - \rangle : V \times V \to \mathbb{C}$$

heißt unitärer Raum. $\langle - \mid - \rangle$ *heißt dann auch* hermitesches inneres Produkt.

4.5.10 Definition (unitäre Matrix) $A \in \mathbb{C}^{n \times n}$ *heißt* unitär, *wenn gilt:*

$$A^{-1} = {}^t\overline{A} = (\overline{a_{ki}}).$$

Natürlich ist $\mathbb{C}^n$ mit dem Standardskalarprodukt ein unitärer Raum. Ein nicht ganz so offensichtliches Beispiel ist die oben eingeführte Komplexifizierung ${}_{\mathbb{C}}(V \times V)$ eines reellen Vektorraums V. Für die dort angegebene hermitesche Semibilinearform Φ gilt nämlich:

$$\Phi((v, w), (v, w)) = \langle v \mid v \rangle + \langle w \mid w \rangle + i \cdot [\langle v \mid w \rangle - \langle w \mid v \rangle] = \langle v \mid v \rangle + \langle w \mid w \rangle.$$

Auf unitären Räumen kann, ganz wie für relle Vektorräume mit innerem Produkt, eine *Norm* definiert werden:

$$||v|| := \sqrt{\langle v \mid v \rangle},$$

denn der Wert unter der Wurzel ist ja eine nicht negative reelle Zahl. Auch hier gilt die Cauchy-Schwarzsche Ungleichung

$$|\langle v \mid w\rangle| \leq \|v\|\,\|w\|.$$

mit Gleichheit genau dann, wenn die beiden Vektoren linear abhängig sind. Der Beweis verläuft hier völlig analog zum Beweis im Reellen. Die Dreiecksungleichung

$$\|v+w\| \leq \|v\| + \|w\|$$

gilt mit Gleichheit genau dann, wenn $v = \rho w, \rho \in \mathbb{R}_{\geq 0}$). Um das zu beweisen quadriet man beide Seiten von $\|v+w\| = \|v\| + \|w\|$ und erhält

4.5.11 $$\langle v \mid w\rangle + \overline{\langle v \mid w\rangle} = 2\sqrt{\langle v \mid v\rangle\langle w \mid w\rangle},$$

also für den Realteil von $\langle v \mid w\rangle$ die Gleichung

$$\mathrm{Re}(\langle v \mid w\rangle) =\leq \|v\|\,\|w\|.$$

Aus der Cauchy-Schwarzschen Ungleichung folgt damit die Gleichheit

$$|\langle v \mid w\rangle| = \|v\|\,\|w\|,$$

die Vektoren v und w sind demnach linear abhängig, etwa $v = \zeta w$. Setzen wir dies in 4.5.11 ein, so folgt $\zeta + \overline{\zeta} = 2|\zeta|$, der Koeffizient ζ ist demnach reell.
Ist umgekehrt $v = \zeta w$, mit $\zeta \in \mathbb{R}_{\geq 0}$, dann wird aus der Dreiecksungleichung ganz offensichtlich eine Identität.

Auch im unitären Fall nennen wir zwei Vektoren v, w genau dann *orthogonal,* wenn ihr inneres Produkt verschwindet:

$$v \perp w \;:\Longleftrightarrow\; \langle v \mid w\rangle = 0.$$

Wir haben also einen Orthogonalitätsbegriff, Orthonormalbasisfolgen, orthogonale Komplemente usw.
Eine wichtige Konstruktion ist der zu V *konjugierte* Vektorraum $\overline{V}$. Er hat dieselbe Grundmenge V von Vektoren, aber das Produkt eine $v \in \overline{V}$ mit einem Skalar $\zeta \in \mathbb{C}$ setzt man gleich dem Vektor $\overline{\zeta} \cdot v$ im Vektorraum V, kurz:

$$\zeta \cdot v := \overline{\zeta} \cdot v.$$

(Das sieht vielleicht auf den ersten Blick merkwürdig aus, wird aber dadurch erklärt, daß links ein Vektor aus $\overline{V}$, und rechts einer aus V steht!) Für die Abbildung

$$f\colon V \to \overline{V}\colon v \mapsto v$$

gilt

$$f(i \cdot v) = -i \cdot f(v).$$

Mit ihrer Hilfe definiert man dann zum unitären V (mit $\langle - \mid - \rangle$) die nicht ausgeartete Bilinearform (nachrechnen!)

$$[- \mid -] : V \times \overline{V} \to \mathbb{C} : (v, w) \mapsto \langle v \mid f^{-1}(w) \rangle.$$

V und $\overline{V}$ sind also bzgl. $[- \mid -]$ zueinander *duale* Vektorräume. Hiermit läßt sich auch der *Rieszsche Darstellungssatz* übertragen: Ist $G : V \to \mathbb{C}$ eine Linearform, dann folgt aus einem ganz analogen Beweis, daß $g(v) = [v \mid b]$, für ein geeignetes $b \in V$. Wegen $[v \mid b] = \langle v \mid f^{-1}(b) \rangle =: \langle v \mid a \rangle$ ergibt sich für g die folgende Darstellung mit Hilfe eines geeigneten $a \in V$ und des inneren Produkts auf V :

$$g(-) = \langle - \mid a \rangle.$$

usw.
Dies wird im kommenden Semester in der Vorlesung Algebra I ergänzt und fortgesetzt!
Beachten Sie bitte auch, daß dieses Manuskript noch nicht korrigiert worden ist!

Kapitel 5

Multilineare Algebra

In diesem Kapitel geht es um die Definition von Tensoren, insbesondere um die Konstruktion von Tensorprodukten von Vektorräumen und Moduln.
Tensoren spielen eine wichtige Rolle in den Naturwissenschaften. Darüberhinaus ist das Tensorprodukt wegen seiner Universaleigenschaft von großer Bedeutung in der Algebra. Ein weiteres sehr wichtiges "universelles Element" werden wir mit der sogenannten Freien Gruppe im nächsten Kapitel kennenlernen.

5.1 Universelle Elemente

Analog zur Untersuchung der Lösbarkeit von Gleichungen der Form $b = x \cdot a$ in Halbgruppen, Gruppen, Ringen, Körpern usw., kann man die Auflösbarkeit von Gleichungen der Form $\beta = \xi \circ \alpha$ in vorgegebenen Klassen von *Abbildungen* betrachten. Mit anderen Worten: es geht jetzt um die Ergänzbarkeit von

$$\begin{tikzcd} M \arrow[r, "\alpha"] \arrow[dr, "\beta"'] & A \arrow[d, "\xi"] \\ & B \end{tikzcd}$$

zu einem kommutativen Diagramm bei gegebenen α und β.

5.1.1 Definition *Sei $\mathcal{F}$ eine vorgegebene Klasse von Funktionen mit Definitionsbereich M, sei $\alpha: M \to A$ in $\mathcal{F}$ und $\mathcal{L}$ eine weitere vorgegebene Klasse von Funktionen. $\mathcal{L}$ enthalte (zu den vorkommenden Definitionsbereichen) die Identitäten und sei abgeschlossen gegenüber Komposition (d.h. für $\gamma, \delta \in \mathcal{L}$ mit Bild $\delta \subseteq Def\ \gamma$ gilt $\gamma \circ \delta \in \mathcal{L}$). Dann heißt $\alpha: M \to A$* universell bzgl. $\mathcal{F}$ und *$\mathcal{L}$, wenn gilt:*

$$\forall\ \beta \in \mathcal{F}\ \exists_1\ \xi \in \mathcal{L}\colon\ \beta = \xi \circ \alpha.$$

In abkürzender Diagrammschreibweise:

$$\begin{tikzcd} M \arrow[r, "\alpha"] \arrow[dr, "\beta"'] & A \arrow[d, "\exists_1\ \xi \in \mathcal{L}"] \\ & B \end{tikzcd}$$

Die Klasse $\mathcal{L}$ heißt dabei die Klasse der zulässigen Lösungen*, und A heißt* universelles Element*.*

5.1.2 Beispiel Ist M eine Menge, $\sim$ eine Äquivalenzrelation auf M und

$$\alpha: M \to M/\sim: m \mapsto [m]_\sim$$

die Projektion der Elemente auf ihre Äquivalenzklasse. Dann ist die Abbildung $\alpha: M \to M/\sim$ universell bezüglich der Klasse $\mathcal{F}$ der Funktionen mit Defintionsbereich M, die konstant auf den Äquivalenzklassen sind, und der Klasse $\mathcal{L}$ aller Abbildungen.

◇

Von zentraler Bedeutung ist nun folgender Satz:

5.1.3 Die Eindeutigkeit universeller Elemente *Sind gegebene Abbildungen*

$$\alpha: M \to A \text{ und } \gamma: M \to C$$

universell bzgl. $\mathcal{F}$ und $\mathcal{L}$, dann gibt es eine Bijektion $\lambda \in \mathcal{L}$ mit $\lambda: A \to C$. Mit anderen Worten: Universelle Elemente sind im wesentlichen, d. h. bis auf zulässige Bijektionen, eindeutig bestimmt.

Beweis: Aus der Universalität von α und γ folgen die beiden Existenzaussagen

$$\exists_1\ \xi \in \mathcal{L}: \gamma = \xi \circ \alpha \text{ und } \exists_1\ \eta \in \mathcal{L}: \alpha = \eta \circ \gamma.$$

Daraus ergibt sich unmittelbar durch Einsetzen:

$$\gamma = \underbrace{\xi \circ \eta}_{\in \mathcal{L}} \circ \gamma \text{ und } \alpha = \underbrace{\eta \circ \xi}_{\in \mathcal{L}} \circ \alpha.$$

Nun gilt aber trivialerweise

$$\alpha = id_A \circ \alpha, \text{ und } \gamma = id_C \circ \gamma.$$

Wegen $id_A, id_C \in \mathcal{L}$ und der Eindeutigkeit der Ergänzungen zeigt also ein Vergleich, daß folgendes richtig ist:

$$\xi \circ \eta = id_C \text{ und } \eta \circ \xi = id_A,$$

woraus sich bekanntlich die Bijektivität von $\xi =: \lambda$ ergibt.

□

5.1.4 Anwendungen Ist G eine Gruppe, $\alpha\colon G \to H$ ein Epimorphismus, $\mathcal{F}$ die Klasse der Homomorphismen γ mit Definitionsbereich G und der Eigenschaft $\mathrm{Kern}(\alpha) \subseteq \mathrm{Kern}(\gamma)$, $\mathcal{L}$ die Klasse aller Homomorphismen zwischen Gruppen, dann ist α offenbar universell bzgl. $\mathcal{F}$ und $\mathcal{L}$. Demnach sind, wegen 5.1.3, alle epimorphen Bilder von G mit demselben Kern $Kern\ \alpha$ isomorph.
Ein weiteres, leicht einzusehendes Beispiel ist die Universalität von

$$\alpha\colon n \to \mathbb{K}^n : i \mapsto e_i,$$

bezüglich der Klasse $\mathcal{F}$ der Abbildungen von $n := \{0,\ldots,\text{n-1}\}$ in Vektorräume über $\mathbb{K}$ und der Klasse $\mathcal{L}$ aller $\mathbb{K}$-linearen Abbildungen. Hier ergibt die Eindeutigkeit universeller Elemente die Isomorphie aller n–dimensionalen Vektorräume über $\mathbb{K}$.

$\diamond$

5.2 Tensorprodukte von Moduln

Wir wollen ein Produkt $M \otimes_R N$ aus einem R-Rechtsmodul M und einem R-Linksmodul N bilden, dabei ist R ein Ring mit Einselement. Die Spezialisierung auf Vektorräume soll ein Vektorraum werden, dessen Dimension das *Produkt* der Dimensionen der beiden Faktoren ist.
Dieses Produkt wird als universelles Element eingeführt. Die Klasse $\mathcal{F}$ ist dabei die Klasse der Abbildungen f von $M \times N$ in abelsche Gruppen, die biadditiv sind, und für die gilt:

$$f(mr, n) = f(m, rn).$$

Solche Abbildungen heißen *ausgeglichen,* und wir definieren:

5.2.1 Definition *Ist T eine abelsche Gruppe, $\tau: M \times N \to T$, dann heißt T ein Tensorprodukt von M und N über R, wenn $\tau: M \times N \to T$ universell ist bzgl. der Klasse $\mathcal{F}$ der ausgeglichenen Abbildungen von $M \times N$ in abelsche Gruppen und der Klasse $\mathcal{L}$ aller Homomorphismen zwischen abelschen Gruppen.*

5.2.2 Satz *Es gibt Tensorprodukte von M und N über R.*

Beweis: Wir konstruieren ein Tensorprodukt. Dazu betrachten wir die *freie abelsche Gruppe* (additiv geschrieben) über $M \times N$, das ist

$$(\mathbb{Z}^{M\times N})' := \{f: M \times N \to \mathbb{Z} \mid \text{ fast alle } f(m,n) = 0\}$$

$$:= \{ \sum_{(m,n)\in M\times N} z_{mn}(m,n) \mid z_{mn} \in \mathbb{Z}\}.$$

(Beachte: Bei der letzten Schreibweise handelt es sich um eine *formale Summe*, nicht um Linearkombinationen o.ä.) Diese Gruppe enthält u.a. die folgenden Elemente (für alle m, m_i in M, und alle n, n_i in N, sowie alle $r \in R$):

$$(m_1 + m_2, n) - (m_1, n) - (m_2, n),$$

$$(m, n_1 + n_2) - (m, n_1) - (m, n_2),$$

$$(mr, n) - (m, rn).$$

Die von all diesen Elementen erzeugte Untergruppe heiße U, und wir setzen

$$T := (\mathbb{Z}^{M\times N})'/U, \tau: M \times N \to T: (m,n) \mapsto (m,n) + U.$$

Es bleibt jetzt zu zeigen, daß $\tau: M \times N \to T$ tatsächlich universell ist.

a) τ ist leicht als ausgeglichen nachgewiesen, z.B. gilt

$$\tau(m_1 + m_2, n) = (m_1 + m_2, n) + U$$

$$= (m_1 + m_2, n) \underbrace{-(m_1 + m_2, n) + (m_1, n) + (m_2, n)}_{\in U} + U$$

$$= (m_1, n) + (m_2, n) + U = \tau(m_1, n) + \tau(m_2, n).$$

b) Zum Beweis der Faktorisierbarkeit eines ausgeglichenen $f: M \times N \to B$ über τ (wobei B eine beliebige, i.a. wieder additiv geschriebene, abelsche Gruppe ist) bemerken wir zunächst, daß sich f trivialerweise und eindeutig über die Einbettung

$$\iota: M \times N \hookrightarrow (\mathbb{Z}^{M\times N})': (m,n) \mapsto 1_{\mathbb{Z}}(m,n)$$

faktorisieren läßt:

$$\begin{tikzcd} M \times N \arrow[r, "\iota"] \arrow[dr, "f"'] & (\mathbb{Z}^{M\times N})' \arrow[d, "\exists_1\ \Phi"] \\ & B \end{tikzcd}$$

Die Eindeutigkeit von Φ folgt dabei aus der Tatsache, daß das Bild von $M \times N$ unter ι eine $\mathbb{Z}$-Basis von $(\mathbb{Z}^{M\times N})'$ als $\mathbb{Z}-$Modul ist. Φ ist dabei so zu definieren:

$$\Phi: (\mathbb{Z}^{M\times N})' \to B: \sum z_{mn}(m,n) \mapsto \sum z_{mn} f(m,n).$$

Da f ausgeglichen ist, gilt $U \subseteq Kern\ \Phi$, so daß wir noch folgendes kommutative Diagramm erhalten:

$$\begin{tikzcd} (\mathbb{Z}^{M\times N})' \arrow[r, "\nu"] \arrow[dr, "\Phi"'] & T \arrow[d, "\exists_1\ \varphi"] \\ & B \end{tikzcd}$$

wobei $\nu: (\mathbb{Z}^{M\times N})' \to T$ der natürliche Epimorphismus auf die Faktorgruppe T ist. φ ist ebenfalls Homomorphismus (nachrechnen!). Insgesamt ergibt sich so das kommutative Diagramm

$$\begin{tikzcd} M \times N \arrow[r, "\tau"] \arrow[dr, "f"'] & T \arrow[d, "\exists_1\ \varphi"] \\ & B \end{tikzcd}$$

mit einem $\varphi \in \mathcal{L}$. Da die Elemente $\tau(m,n)$ ganz T erzeugen, ist zudem φ eindeutig bestimmt.

□

5.3 Eigenschaften von Tensorprodukten

Das im Beweis von 5.2.2 konstruierte Tensorprodukt von M und N über R bezeichnen wir so:

5.3.1 $$M \otimes_R N := (\mathbb{Z}^{M \times N})'/U,$$

und die Abbildung τ ersetzen wir durch das Symbol $\otimes$:

5.3.2 $$\otimes : M \times N \to M \otimes_R N : (m, n) \mapsto m \otimes n := (m, n) + U.$$

Aus dem Satz über die Eindeutigkeit universeller Elemente erhalten wir jetzt:

5.3.3 Folgerung *Jedes Tensorprodukt von M und N über R ist, als abelsche Gruppe, isomorph zu $M \otimes_R N$.*

Wir nennen deshalb $M \otimes_R N$ kurzerhand auch *das* Tensorprodukt von M und N über R. Es hat insbesondere die folgenden Eigenschaften:

5.3.4 Eigenschaften des Tensorprodukts *Für Tensorprodukte von Moduln gilt:*

$$M \otimes_R N = \{ \sum_{(m,n) \in M \times N} z_{mn} m \otimes n \mid z_{mn} \in \mathbb{Z}, \textit{ fast alle } z_{mn} = 0 \}$$

$$= \langle m \otimes n \mid (m, n) \in M \times N \rangle,$$

außerdem gelten die Rechenregeln

$$(m_1 + m_2) \otimes n = m_1 \otimes n + m_2 \otimes n,$$

$$m \otimes (n_1 + n_2) = m \otimes n_1 + m \otimes n_2,$$

und

$$mr \otimes n = m \otimes rn.$$

Dabei ist es jedoch sehr wichtig zu beachten, daß

$$\{m \otimes n \mid (m, n) \in M \times N\}$$

keine $\mathbb{Z}$-Basis ist, denn in der Regel ist eine Darstellung

$$x = \sum z_{mn} (m \otimes n)$$

von $x \in M \otimes_R N$ als $\mathbb{Z}$-Linearkombination *nicht* eindeutig, gilt doch beispielsweise

$$3 \otimes 4 = \frac{3}{2} \otimes 8 \in \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}.$$

Bei der Interpretation eines Ausdruckes $m \otimes n$ ist noch weitere Vorsicht geboten. Es muß zweifelsfrei klar sein, *als Element welches Tensorprodukts* dieses $m \otimes n$ verstanden werden soll. Es kann nämlich sehr wohl sein, daß m in einem echten Untermodul M' von M liegt und $m \otimes n$ als Element von $M' \otimes_R N$ verschieden von 0 ist, obwohl $M \otimes_R N = \{0\}$ und damit $m \otimes n = 0$, als Element von $M \otimes_R N$. Z. B. gilt

$$\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}_2 \simeq \mathbb{Z}_2, \ \textit{aber} \ \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Z}_2 = \{0\}.$$

Denn für alle $a \in \mathbb{Z}, b \in \mathbb{Z}_2, q \in \mathbb{Q}$ gilt:

$$a \otimes b = 1 \otimes ab \in \{1 \otimes 0, 1 \otimes 1\} \simeq \mathbb{Z}_2,$$

während

$$q \otimes b = q' \cdot 2 \otimes b = q' \otimes 2b = q' \otimes 0 = q' \cdot 0 \otimes 0 = 0 \otimes 0 = 1 \otimes 0.$$

Seien nun $f\colon M_R \to M'_R$ und $g\colon {}_RN \to {}_RN'$ R-Homomorphismen. Dann gilt:

5.3.5 Satz *Es gibt genau einen Homomorphismus $f \otimes g$ zwischen (den abelschen Gruppen) $M \otimes_R N$ und $M' \otimes_R N'$ mit*

$$(f \otimes g)(m \otimes n) = f(m) \otimes g(n).$$

Beweis: Die Komposition

$$\otimes \circ (f \times g)\colon (m, n) \mapsto f(m) \otimes g(n)$$

ist nach den Rechenregeln 5.3.4 ausgeglichen. Es gibt deshalb genau ein $\varphi \in \mathcal{L}$, der Klasse aller Homomorphismen zwischen abelschen Gruppen, das folgendes Diagramm kommutativ ergänzt:

$$\begin{tikzcd}
M \times N \arrow[r, "\otimes"] \arrow[dr, "\otimes \circ (f \times g)"'] & M \otimes_R N \arrow[d, "\varphi"] \\
& M' \otimes_R N'
\end{tikzcd}$$

Für φ gilt offensichtlich $\varphi(m \otimes n) = f(m) \otimes g(n)$, so daß die Abbildung $f \otimes g := \varphi$ die Behauptung erfüllt.

$\square$

Sei $A \dot{+} B$ die abelsche Gruppe über $A \times B$ mit komponentenweiser Verknüpfung, wobei A und B abelsch sind. Dann gilt der folgende sehr wichtige Satz über die Zerlegung des Tensorprodukts mit Hilfe von Untermoduln:

5.3.6 Satz *Ist $M = M_1 \oplus M_2$ (als R-Rechtsmodul), dann gilt*

$$M \otimes_R N \simeq M_1 \otimes_R N \dot{+} M_2 \otimes_R N.$$

Beweis: π_i sei die Projektion von M auf M_i, $\theta_i := \pi_i \otimes id_N$, und

$$T_i := \theta_i(M \otimes_R N), i = 1, 2.$$

Es gilt: $M \otimes_R N = T_1 \oplus T_2 \simeq T_1 \dot{+} T_2$, so daß zu zeigen bleibt: $T_i \simeq M_i \otimes_R N$, d.h. T_i ist als *Tensorprodukt* von M_i und N, also als universell nachzuweisen. Sei dazu $f_i \colon M_i \times N \to B$ ausgeglichen und g_i definiert durch die Kommutativität von

$$\begin{tikzcd}
M \times N \arrow[rr, "\otimes"] \arrow[dr, "\pi_i \times 1"'] & & M \otimes_R N \arrow[dd, "\exists_1 \, g_i"] \\
& M_i \times N \arrow[dr, "f_i"'] & \\
& & B
\end{tikzcd}$$

Wir setzen $\bar{g}_i := g_i \downarrow T_i$ und $\otimes_i := \otimes \downarrow M_i \times N$. Es soll gezeigt werden, daß folgendes Diagramm kommutativ ist:

$$\begin{tikzcd}
M_i \times N \arrow[r, "\otimes_i"] \arrow[dr, "f_i"'] & T_i \arrow[d, "\bar{g}_i"] \\
& B
\end{tikzcd}$$

Dazu beachten wir, daß für alle $(m_i, n) \in M_i \times N$ gilt:

$$\begin{aligned}
\overline{g}_i \circ \otimes_i(m_i, n) &= \overline{g}_i \circ \otimes(m_i, n) = \overline{g}_i(m_i \otimes n) = g_i(m_i \otimes n) \\
&= f_i \circ (\pi_i \times id_N)(m_i, n) = f_i(m_i, n).
\end{aligned}$$

Das Diagramm ist also tatsächlich kommutativ. Wegen

$$T_i = \langle \theta_i \circ \otimes_i(M_i \times N) \rangle$$

ist $\overline{g}_i$ zudem eindeutig bestimmt, $\otimes_i \colon M_i \times N \to T_i$ also universell.

$\square$

5.4 Tensorprodukte als Moduln

Ist S ein weiterer Ring mit Einselement und M sowohl R-Rechts- als auch S-Linksmodul, dann heißt M ein (S,R)-*Bimodul* (kurz: ${}_SM_R$), wenn gilt

$$(sm)r = s(mr).$$

Zu solchen Bimoduln werden z.B. $\mathbb{K}$-Linksvektorräume, $\mathbb{K}$ ein Körper, wenn man $v\kappa := \kappa v$ setzt, denn dann gilt:

$$(\lambda v)\kappa = \kappa(\lambda v) = (\kappa\lambda)v = (\lambda\kappa)v = \lambda(\kappa v) = \lambda(v\kappa).$$

5.4.1 Satz *Ist M ein (S,R)-Bimodul, dann wird $M \otimes_R N$ zu einem S-Linksmodul vermöge*

$$s\left(\sum z_{mn}(m \otimes n)\right) := \sum z_{mn}(sm \otimes n).$$

Beweis: Für jedes $s \in S$ ist die Abbildung

$$\beta_s \colon M \times N \to M \otimes_R N \colon (m,n) \mapsto sm \otimes n$$

ausgeglichen, so daß genau ein Endomorphismus φ_s existiert mit kommutativem Diagramm

$$\begin{tikzcd}
M \times N \arrow[r, "\otimes"] \arrow[dr, "\beta_s"'] & M \otimes_R N \arrow[d, "\xi_s"] \\
& M \otimes_R N
\end{tikzcd}$$

Für diesen Endomorphismus gilt also

$$\xi_s(m \otimes n) = \beta_s(m,n) = sm \otimes n.$$

Setzt man also, für $x \in M \otimes_R N$, $sx := \xi_s(x)$, dann gilt

$$s(x+y) = sx + sy,$$

$\forall\, x, y \in M \otimes_R N$, da ξ_s Endomorphismus ist. Die anderen Modulgesetze folgen entsprechend aus den Rechenregeln 5.3.4.

□

Ist zudem noch $M = M_1 \oplus M_2$ als (S,R)-Bimodul, dann ist auch

5.4.2 $\quad M \otimes_R N \simeq M_1 \otimes_R N \dot{+} M_2 \otimes_R N$ (als S-Linksmodul).

5.4.3 Satz

$$R \otimes_R N \simeq N \quad \textit{(als R-Linksmodul).}$$

Beweis: Die Abbildung

$$f: R \otimes_R N \to N: r \otimes n \mapsto rn$$

ist leicht als $R-$Isomorphismus zu erkennen, denn mit

$$g: N \to R \otimes_R N: n \mapsto 1 \otimes n$$

genügt sie

$$f \circ g = id_N \wedge g \circ f = id_{R \otimes N}.$$

f ist also bijektiv. Außerdem ist f (R-)Homomorphismus, da

$$\begin{tikzcd} R \times N \arrow[r, "\otimes"] \arrow[dr, "\beta: (r,n) \mapsto r \cdot n"'] & R \otimes_R N \arrow[d, "f"] \\ & N \end{tikzcd}$$

kommutativ ist.

□

5.4.4 Satz *Diese Bildung des Tensorproduktes ist assoziativ, d.h.*

$$L \otimes_R (M \otimes_S N) \simeq (L \otimes_R M) \otimes_S N.$$

Beweis: Wir zeigen dazu, daß durch

$$l \otimes (m \otimes n) \mapsto (l \otimes m) \otimes n$$

ein Isomorphismus definiert wird. Deshalb betrachten wir die Abbildung

$$\gamma: (L \otimes M) \times N \to L \otimes (M \otimes N): ((l \otimes m), n) \mapsto l \otimes (m \otimes n).$$

Man rechnet leicht nach, daß γ ausgeglichen ist. Es gibt also genau einen Homomorphismus Γ mit kommutativem Diagramm

$$\begin{tikzcd} (L \otimes M) \times N \arrow[r, "\otimes"] \arrow[dr, "\gamma"'] & (L \otimes M) \otimes N \arrow[d, "\Gamma"] \\ & L \otimes (M \otimes N) \end{tikzcd}$$

Γ genügt offenbar den Gleichungen

$$\Gamma((l \otimes m) \otimes n) = l \otimes (m \otimes n).$$

Entsprechend zeigt man die Existenz eines Homomorphismus

$$\Delta\colon L \otimes (M \otimes N) \to (L \otimes M) \otimes N$$

mit

$$\Delta(l \otimes (m \otimes n)) = (l \otimes m) \otimes n.$$

Für beide zusammen gilt somit

$$\Delta \circ \Gamma = id_{(L\otimes M)\otimes N}, \Gamma \circ \Delta = id_{L\otimes(M\otimes N)}.$$

Γ ist also Isomorphismus.

□

Statt $L \otimes_R (M \otimes_S N)$ können wir deshalb $L \otimes_R M \otimes_S N$ schreiben, entsprechend auch für endlich viele Faktoren

$$L \otimes_{R_1} M \otimes_{R_2} \ldots \otimes_{R_n} N.$$

Die Elemente solcher Tensorprodukte heißen *Tensoren.* Diejenigen unter ihnen, die von der Form $l \otimes m \otimes \ldots \otimes n$ sind, heißen *zerlegbare* Tensoren. Man nennt ein solches Element auch das *Tensorprodukt* von $l, m, \ldots, n$.

5.5 Tensorprodukte von Vektorräumen

$V_{\mathbb{K}}$ und ${}_{\mathbb{K}}W$ seien jetzt $\mathbb{K}$-Vektorräume, $\mathbb{K}$ ein Körper. Dann ist, wie bereits erwähnt, V auf natürliche Weise ein $\mathbb{K}$-Bimodul ($\kappa v := v\kappa$), und nach 5.4.1 wird deshalb $V \otimes_{\mathbb{K}} W$ zu einem $\mathbb{K}$-Vektorraum vermöge

$$\kappa \sum z_{vw}(v \otimes w) := \sum z_{vw}(\kappa v \otimes w) = \sum z_{vw}(v \otimes \kappa w).$$

5.5.1 Satz *Ist* $V = \ll b_0, \ldots, b_{m-1}\rangle\!\rangle, W = {}_{\mathbb{K}}\langle c_0, \ldots, c_{n-1} \gg$, *dann gilt*

$$V \otimes_{\mathbb{K}} W = \ll b_i \otimes c_k \mid i \in m, k \in n \gg,$$

also ist insbesondere

$$dim_{\mathbb{K}}(V \otimes_{\mathbb{K}} W) = dim_{\mathbb{K}}(V) \cdot dim_{\mathbb{K}}(W).$$

Beweis: Aus $V = \oplus\, \mathbb{K}b_i\,, W = \oplus\, \mathbb{K}c_k$ folgt nach 5.4.2:

$$V \otimes_{\mathbb{K}} W \simeq_{\mathbb{K}} \dot{+}_{i,k}\ (\mathbb{K}b_i \otimes \mathbb{K}c_k).$$

Daraus folgt die Behauptung über die Dimension, denn

$$\mathbb{K}b_i \otimes \mathbb{K}c_k = \{\kappa(b_i \otimes c_k) \mid \kappa \in \mathbb{K}\} = \mathbb{K}(b_i \otimes c_k).$$

Zudem ergibt sich aus den Rechenregeln 5.3.4, daß die $b_i \otimes c_k$ das Tensorprodukt erzeugen, sie müssen also eine Basis bilden.

□

Wir betrachten jetzt lineare Abbildungen auf endlichdimensionalen $\mathbb{K}$-Vektorräumen und die davon auf dem Tensorprodukt induzierte lineare Abbildung. Sei also wieder $\mathcal{B} = (b_0, \ldots, b_{m-1})$ Basisfolge von V, $\mathcal{C} = (c_0, \ldots, c_{n-1})$ Basisfolge von W, und es sei $f \in Hom_{\mathbb{K}}(V, V')$, $g \in Hom_{\mathbb{K}}(W, W')$. Weiter sei $\mathcal{B}' = (b'_0, \ldots, b'_{m'})$ Basisfolge von V', $\mathcal{C}' = (c'_0, \ldots, c'_{n'})$ Basisfolge von W', und wir setzen $A = M(\mathcal{B}', f, \mathcal{B})$, $B = M(\mathcal{C}', g, \mathcal{C})$, die darstellenden Matrizen der linearen Abbildungen. Wegen 5.4.3 gilt

$$(f \otimes g)(b_j \otimes c_l) = \sum_{i,k} a_{ij} b_{kl} (b'_i \otimes c'_k),$$

also folgt für die Matrix $A \otimes B$, die $f \otimes g$ beschreibt bzgl. der doppelt lexikographisch geordneten Basisfolgen $(\ldots, b_i \otimes c_k, \ldots) = (b_0 \otimes c_0, b_0 \otimes c_1, \ldots, b_{m-1} \otimes c_{n-1})$ und die entsprechend angeordnete Folge $(\ldots, b'_i \otimes c'_k, \ldots)$:

$$A \otimes B = \begin{pmatrix} a_{00}B & a_{01}B & \ldots \\ a_{10}B & a_{11}B & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}$$

Diese Matrix heißt ein *Kroneckerprodukt* von A und B (die anderen Kroneckerprodukte von A und B gehören zu den anderen Numerierungen der Basiselemente). Es gilt:

5.5.2 $$\mathrm{Spur}(A \otimes B) = \mathrm{Spur}(A) \cdot \mathrm{Spur}(B),$$

sowie

5.5.3 $$(A \otimes B)(A' \otimes B') = (AA') \otimes (BB'),$$

bei geeigneten Zeilen- und Spaltenzahlen. Für die Determinante schließlich gilt

5.5.4 $$det\ A \otimes B = (det\ A)^n \cdot (det\ B)^m,$$

falls A m-reihig, B n-reihig. Weiterhin haben wir

5.5.5 $$\mathrm{Kern}(f \otimes g) \ = \ \mathrm{Kern}(f) \otimes W + V \otimes \mathrm{Kern}(g),$$

sowie

5.5.6 $$\mathrm{Rang}(f \otimes g) = \mathrm{Rang}(f) \cdot \mathrm{Rang}(g).$$

5.5.7 Satz *Ist $\mathbb{K}$ Teilkörper von $\mathbb{L}$, dann ist $\mathbb{L} \otimes_{\mathbb{K}} V$ ein $\mathbb{L}$-Vektorraum. Hat ${}_{\mathbb{K}}V$ die Basisfolge $\mathcal{B} = (b_0, \ldots, b_{m-1})$, dann gilt*

$$\mathbb{L} \otimes_{\mathbb{K}} V = \ {}_{\mathbb{L}} \ll 1_{\mathbb{L}} \otimes b_i \mid i \in m \gg .$$

Man sagt dazu, $\mathbb{L} \otimes_{\mathbb{K}} V$ entsteht aus V durch *Grundkörpererweiterung*.

5.6 Lineare Symmetrieklassen von Tensoren

Wir betrachten jetzt, zu einem Körper $\mathbb{K}$, einem endlichdimensionalen $\mathbb{K}$-Vektorraum V und einem $n \in \mathbf{N}$, die n-fache Tensorpotenz von V:

$$\overset{n}{\otimes} V := V \otimes_{\mathbb{K}} \ldots \otimes_{\mathbb{K}} V, n \ Faktoren, \ \overset{0}{\otimes} V := \mathbb{K}.$$

Unser Ziel ist die Beschreibung gewisser Unterräume von $\overset{n}{\otimes} V$, über die wir u.a. symmetrische und schiefsymmetrische Abbildungen faktorisieren können. Wir bemerken zunächst, daß folgendes gilt:

5.6.1 Hilfssatz *Zu $\pi \in S_n$ wird durch die Kommutativität des folgenden Diagramms der* Permutationsoperator $P(\pi)$ *definiert (wobei $v^{\otimes} := v_0 \otimes \ldots \otimes v_{n-1}$ und $v^{\otimes}_{\pi^{-1}} := v_{\pi^{-1}(0)} \otimes \ldots \otimes v_{\pi^{-1}(n-1)}$) :*

$$\begin{tikzcd}
\overset{n}{\times} V \arrow[r, "\otimes"] \arrow[dr, "(v_0, \ldots, v_{n-1}) \mapsto v^{\otimes}_{\pi^{-1}}"'] & \overset{n}{\otimes} V \arrow[d, "\exists_1 \ P(\pi)\colon v^{\otimes} \mapsto v^{\otimes}_{\pi^{-1}}"] \\
 & \overset{n}{\otimes} V
\end{tikzcd}$$

Die dadurch definierte Abbildung

$$P\colon S_n \to GL(\overset{n}{\otimes} V)\colon \pi \mapsto P(\pi)$$

ist ein Homomorphismus von S_n in die volle lineare Gruppe $GL(\otimes^n V)$.

P ist also eine Darstellung von S_n auf $\overset{n}{\otimes} V$ im folgenden Sinn:

5.6.2 Definition *Ist G eine Gruppe, $\mathbb{K}$ ein Körper, $W \neq \{0_W\}$ ein $\mathbb{K}$-Vektorraum, dann heißt jeder Homomorphismus $D\colon G \to GL(W)$ eine* (lineare) Darstellung *von G auf W. Dabei nennt man W den* Darstellungsraum*. Die Dimension von W heißt auch die* Dimension *von D, und die Funktion*

$$\chi^D\colon G \to \mathbb{K}\colon g \mapsto \mathrm{Spur}(D(g))$$

heißt der Charakter *der Darstellung.*

5.6.3 Beispiele Die Charaktere eindimensionaler Darstellungen, die sogenannten *eindimensionalen* Charaktere von G, sind genau die Homomorphismen χ von G in die multiplikative Gruppe $\mathbb{K}^* := (\mathbb{K}\backslash\{0\}, \cdot)$ von $\mathbb{K}$. Der *Einscharakter*

$$\iota\colon g \mapsto 1_{\mathbb{K}}$$

ist dabei ein trivialer Spezialfall. Neben ι hat die symmetrische Gruppe S_n, $n \geq 2$, $char\mathbb{K} \neq 2$, noch genau einen weiteren eindimensionalen Charakter, nämlich

$$\epsilon\colon \pi \mapsto sgn(\pi),$$

den sogenannten *alternierenden* Charakter.

5.6.4 Definition *Sei $G \leq S_n$, char$\mathbb{K}$ teile die Ordnung von G nicht, χ sei ein Charakter von G. Dann heißt die lineare Abbildung*

$$P_G^\chi := \frac{\chi(1)}{|G|} \sum_{\pi \in G} \chi(\pi^{-1}) P(\pi) \qquad \in End_{\mathbb{K}}(\overset{n}{\otimes} V)$$

der Symmetrisierungsoperator *zu* (G, χ).

5.6.5 Hilfssatz *Ist P_G^χ Symmetrisierungsoperator zu (G, χ) und zu dem eindimensionalen Charakter χ, dann ist P_G^χ ein Projektionsoperator, d.h. es gilt*

$$(P_G^\chi)^2 = P_G^\chi.$$

Weiterhin gilt für diesen Operator:

$$Rang(P_G^\chi) = \text{Spur}(P_G^\chi) = \frac{1}{|G|} \sum_{\pi \in G} \chi(\pi^{-1}) m^{c(\pi)},$$

wenn $m := dim_{\mathbb{K}}(V)$ ist, und $c(\pi)$ die Anzahl der zyklischen Faktoren von $\pi \in G$ bezeichnet.

Beweis: Die Homomorphieeigenschaft von χ und P ergibt, daß dieser Operator ein Projektionsoperator ist. Hieraus wiederum folgt

$$P_G^\chi = 0_{Kern(P_G^\chi)} \oplus id_{Bild(P_G^\chi)},$$

was Rang$(P_G^\chi) =$ dimBild$(P_G^\chi) =$ Spur(P_G^χ) liefert. Ist jetzt $\mathcal{B} = (b_0, \ldots, b_{m-1})$ eine Basisfolge von V, dann gilt

$$\overset{n}{\otimes} V = \ _{\mathbb{K}} \ll b_\varphi := b_{\varphi(0)} \otimes \ldots \otimes b_{\varphi(n-1)} \mid \varphi \in m^n \gg .$$

Da $P(\pi)$ diese Basiselemente nur permutiert, ist

$$\text{Spur}(P(\pi)) = |\{\varphi \in m^n \mid \varphi = \varphi \circ \pi^{-1}\}|.$$

Nun gilt aber $\varphi = \varphi \circ \pi^{-1}$ genau dann, wenn

$$\forall\ i \in n\colon\ \varphi(i) = \varphi(\pi^{-1}(i)) = \varphi(\pi^{-2}(i)) = \ldots,$$

also genau dann, wenn φ auf den zyklischen Faktoren von π^{-1} konstant ist. Mit $c(\pi^{-1}) = c(\pi))$ ergibt das

$$\text{Spur}(P(\pi)) = m^{c(\pi)},$$

woraus der Rest der Behauptung folgt.

□

Wir wollen nun zeigen, daß die Bilder von Symmetrisierungsoperatoren P_G^χ zu linearen Charakteren χ wichtige Universaleigenschaften haben und *setzen deshalb für den Rest des Kapitels voraus, daß die Charakteristik von $\mathbb{K}$ kein Teiler von $n!$ ist.*

5.6.6 Definition *Sei $f\colon \overset{n}{\times} V \to W$ multilinear, $G \le S_n, \chi\colon G \to \mathbb{K}^*$ ein eindimensionaler Charakter. f heißt dann (G,χ)–symmetrisch, wenn gilt*

$$\chi(\pi) f(v_{\pi(0)}, \dots, v_{\pi(n-1)}) = f(v_0, \dots, v_{n-1}).$$

Dafür schreiben wir kurz:$f \in M_n(V, W, G, \chi)$. Beispielsweise gilt $\det \in M_n(V, \mathbb{K}, S_n, \epsilon)$. Weiter setzen wir

$$V_G^\chi := P_G^\chi(\overset{n}{\otimes} V)$$

und nennen die Unterräume dieser Form von $\overset{n}{\otimes} V$ lineare Symmetrieklassen von Tensoren. Wir führen dazu noch die folgende Bezeichnung ein:

$$v^\Delta := v_0 \Delta \dots \Delta v_{n-1} := P_G^\chi(v_0 \otimes \dots \otimes v_{n-1}) = P_G^\chi(v^\otimes).$$

5.6.7 Satz $\Delta\colon \overset{n}{\times} V \to V_G^\chi\colon (v_0, \dots, v_{n-1}) \mapsto v^\Delta$ *ist universell bzgl.*

$$\mathcal{F} := \{f \mid \exists\, U\colon f \in M_n(V, U, G, \chi)\}$$

und der Klasse $\mathcal{L}$ der linearen Abbildungen zwischen $\mathbb{K}$- Vektorräumen.

Beweis: Δ liegt in $\mathcal{F}$, denn

$$\chi(\pi) v_\pi^\Delta = \chi(\pi) P_G^\chi(v_\pi^\otimes)$$

$$= \frac{1}{|G|} \sum_{\rho \in G} \chi(\pi\rho^{-1}) P(\rho\pi^{-1})(v^\otimes) = P_G^\chi(v^\otimes) = v^\Delta.$$

Die Faktorisierungseigenschaft zeigen wir wie folgt: Jedes $f \in M_n(V, W, G, \chi)$ ist multilinear, so daß genau ein lineares φ existiert mit kommutativem Diagramm

$$\begin{tikzcd}
\overset{n}{\times} V \arrow[r, "\otimes"] \arrow[dr, "f"'] & \overset{n}{\otimes} V \arrow[d, "\varphi"] \\
& W
\end{tikzcd}$$

Dieses φ kann über P_G^χ faktorisiert werden, denn

$$\varphi \circ P_G^\chi(v^\otimes) = \frac{1}{|G|} \sum \chi(\pi^{-1}) \varphi(v_{\pi^{-1}}^\otimes)$$

$$= \frac{1}{|G|} \sum \chi(\pi^{-1}) f(v_{\pi^{-1}(1)}, \dots, v_{\pi^{-1}(n)}) = f(v_0, \dots, v_{n-1}) = \varphi(v^\otimes).$$

Es gilt also $\varphi \circ P_G^\chi = \varphi$, und damit $Kern\ P_G^\chi \subseteq Kern\ \varphi$. Demnach gibt es lineare ψ mit kommutativem Diagramm

$$\begin{tikzcd} \overset{n}{\otimes} V \arrow[r, "P_G^\chi"] \arrow[d, "\varphi"'] & V_G^\chi \arrow[dl, "\psi"] \\ W & \end{tikzcd}$$

Da P_G^χ surjektiv ist, ist zudem ψ eindeutig bestimmt.

□

5.6.8 Bemerkung *Es gibt eine natürliche Verallgemeinerung des Begriffs der linearen Symmetrieklasse auf die Bilder von Projektionen von Symmetrieklassen* P_G^χ *zu nicht notwendig linearen Charakteren (den sogenannten* irreduziblen *Charakteren). Die Beschreibung von diesen setzt jedoch Resultate aus der Darstellungstheorie voraus.*

5.7 Basen und Beispiele von linearen Symmetrieklassen

Wir haben gesehen, daß die lineare Abbildung $P(\pi)$ die Basiselemente $b_\varphi^\otimes$ permutiert. Im Grunde handelt es sich dabei um die folgende Operation (vgl. 2.2.1) von $G \leq S_n$ auf der Indexmenge m^n der Basis $\{b_\varphi^\otimes\}$ von $\overset{n}{\otimes} V$:

5.7.1 $$G \times m^n \to m^n : (\pi, \varphi) \mapsto \varphi \circ \pi^{-1}.$$

Die Untergruppe $G_\varphi := \{\pi \in G \mid \varphi = \varphi \circ \pi^{-1}\}$ heißt der Stabilisator von φ, und die Menge aller Bahnen $G(\varphi) := \{\varphi \circ \pi^{-1} \mid \pi \in G\}$ sei wie folgt bezeichnet:

$$G \backslash\backslash m^n := \{G(\varphi) \mid \varphi \in m^n\}.$$

Sei R_G ein Repräsentantensystem dieser Menge, eine sogenannte *Transversale.* Zu einem eindimensionalen $\chi : G \to \mathbb{K}^*$ sei R_G^χ die folgende Teilmenge:

$$R_G^\chi := \{\varphi \in R_G \mid G_\varphi \subseteq Kern\ \chi\} = \{\varphi \in R_G \mid \forall \pi \in G_\varphi : \chi(\pi) = 1\}.$$

Unser Ziel ist der Nachweis, daß diese Menge eine Basis der Symmetrieklasse zu G und χ ist:

5.7.2 Satz $V_G^\chi = \ll b_\varphi^\Delta \mid \varphi \in R_G^\chi \gg$.

Beweis:
i) Wir bemerken zunächst, daß wir unsere Untersuchung auf die b_φ^Δ beschränken können: Da die $b_\varphi^\otimes$ die Tensorpotenz $\overset{n}{\otimes} V$ erzeugen, spannen die $b_\varphi^\Delta = P_G^\chi(b_\varphi^\otimes), \varphi \in m^n$, deren Bild unter P_G^χ auf.
ii) Da Elemente aus derselben Bahn lineare Vielfache voneinander sind (mit von Null verschiedenen Koeffizienten),

$$b_{\varphi\circ\pi^{-1}}^\Delta = \frac{1}{|G|} \sum_{\rho \in G} \chi(\rho^{-1}) P(\rho\pi) b_\varphi^\otimes = \chi(\pi) b_\varphi^\Delta, \qquad (*)$$

können wir uns sogar auf die Elemente irgendeines Repräsentantensystems Transversale R_G beschränken. V_G^χ wird sogar schon von den $b_\varphi^\Delta, \varphi \in R_G$, erzeugt.
iii) Tatsächlich genügen sogar die $\varphi \in R_G^\chi$, denn die anderen Elemente aus der Transversale sind Null: Für $\varphi \in R_G \backslash R_G^\chi$ gibt es ein $\pi \in G_\varphi$ mit $\chi(\pi^{-1}) \neq 1$, so daß aus $(*)$ folgt:

$$b_\varphi^\Delta = \chi(\pi^{-1}) b_{\varphi\circ\pi^{-1}}^\Delta = \chi(\pi^{-1}) b_\varphi^\Delta = 0.$$

iv) Zum abschließenden Nachweis der linearen Unabhängigkeit der $b_\varphi^\Delta, \varphi \in R_G^\chi$, benutzen wir die zu $(b_0, \ldots, b_{m-1})$ duale Basisfolge $(\lambda_0, \ldots, \lambda_{m-1})$, $\lambda_i \in Hom_\mathbb{K}(V, \mathbb{K})$, also die Linearformen mit der Eigenschaft $\lambda_i(b_j) = \delta_{ij}$. Aus ihnen können wir leicht die zu $\{b_\varphi^\otimes \mid \varphi \in m^n\}$ duale Basis gewinnen, es ist natürlich die Menge

$$\{\lambda_\varphi^\otimes \mid \varphi \in m^n\}.$$

Mit ihrer Hilfe können wir die Koeffizienten κ_φ aus jeder Linearkombination ermitteln. Ist etwa

$$x = \sum_{\psi \in R_G^\chi} \kappa_\psi b_\psi^\Delta,$$

dann gilt, für $\varphi \in R_G^\chi$,

$$\begin{aligned}
\lambda_\varphi^\otimes(x) &= \sum_{\psi \in R_G^\chi} \kappa_\psi \lambda_\varphi^\otimes(b_\psi^\Delta) \\
&= \sum_{\psi \in R_G^\chi} \kappa_\psi \lambda_\varphi^\otimes P_G^\chi(b_\psi^\otimes) \\
&= \sum_{\psi \in R_G^\chi} \kappa_\psi \lambda_\varphi^\otimes \frac{1}{|G|} \sum_{\pi \in G} \chi(\pi^{-1}) P(\pi)(b_\psi^\otimes) \\
&= \sum_{\psi \in R_G^\chi} \kappa_\psi \frac{1}{|G|} \sum_{\pi \in G} \chi(\pi^{-1}) \lambda_\varphi^\otimes(b_{\psi\circ\pi^{-1}}^\otimes) \\
&= \kappa_\varphi \frac{|G_\varphi|}{|G|},
\end{aligned}$$

letzteres wegen

$$\lambda_\varphi^\otimes(b_{\psi\circ\pi^{-1}}^\otimes) = \chi(\pi)\delta_{\varphi,\psi}.$$

Insgesamt erhalten wir demnach die interessante Gleichung

5.7.3 $$\kappa_\varphi = \frac{|G|}{|G_\varphi|} \lambda_\varphi^\otimes(x).$$

Ist beispielsweise

$$0 = \sum_{\psi \in R_G^\chi} \kappa_\psi b_\psi^\Delta,$$

dann gilt für die Koeffizienten in dieser Linearkombination

$$\kappa_\varphi = \frac{|G|}{|G_\varphi|} \lambda_\varphi^\otimes(0) = 0,$$

die Elemente aus R_G^χ sind also tatsächlich linear unabhängig, und der Beweis ist damit abgeschlossen.

□

5.7.4 Beispiele

i) $\bigvee^n V := V_{S_n}^\iota$ Die Bahnen von S_n auf m^n sind offenbar die Teilmengen der Funktionen vom gleichen Gewicht:

5.7.5 $$S_n(\varphi) = S_n(\psi) \iff \forall\ i \in m\colon\ |\varphi^{-1}(i)| = |\psi^{-1}(i)|.$$

Als ein Repräsentantensystem R_{S_n} der Bahnen können wir deshalb die (schwach) monoton wachsenden Funktionen φ wählen. Da wir den Einscharakter ι betrachten, gilt außerdem $R^{\iota}_{S_n} = R_{S_n}$. Es folgt also aus 5.7.2, wenn wir statt b^{Δ}_{φ} zur Verdeutlichung $b^{\vee}_{\varphi} := P^{\iota}_{S_n}(b^{\otimes}_{\varphi})$ schreiben:

5.7.6 $$\bigvee^n V := V^{\iota}_{S_n} = \ll b^{\vee}_{\varphi} \mid \varphi \in m^n \textit{ schwach monoton wachsend} \gg .$$

Die Dimension dieses Raumes erhalten wir aus der Tatsache, daß jedes *schwach* monoton wachsende n-Tupel $(\varphi(1), \ldots, \varphi(n))$ über m ein *streng* monoton wachsendes $n-$Tupel $(\varphi(1), \varphi(2)+1, \ldots, \varphi(n)+n-1)$ über $m+n-1$ ergibt und umgekehrt. Die Anzahl all dieser streng monoton wachsenden $n-$Tupel ist aber

5.7.7 $$\binom{m+n-1}{n};$$

also ist dies die Dimension von $\bigvee^n V$.

<u>*ii)* $\bigwedge^n V := V^{\epsilon}_{S_n}$</u> Ein $\varphi \in R_{S_n}$ liegt genau dann in $R^{\epsilon}_{S_n}$, wenn gilt

$$\forall i\colon\ |\varphi^{-1}[\{i\}]| \leq 1,$$

denn ein solches φ darf ja durch keine Transposition festgelassen werden. In $R^{\epsilon}_{S_n}$ liegen also genau die streng monoton wachsenden φ. Mit $b^{\wedge}_{\varphi} := P^{\epsilon}_{S_n}(b^{\otimes}_{\varphi})$ gilt demnach:

5.7.8 $$\bigwedge^n V = \ll b^{\wedge}_{\varphi} \mid \varphi \in m^n \textit{ streng monoton wachsend} \gg .$$

Für die Dimension ergibt sich also

5.7.9 $$\dim(\bigwedge^n V) = \binom{m}{n}.$$

Zur Berechnung der Komponenten von Elementen x aus $\bigvee^n V$ oder aus $\bigwedge^n V$ können wir 5.7.3 verwenden. Betrachten wir etwa $x := v^{\wedge} := v_0 \wedge \ldots \wedge v_{n-1} \in \bigwedge^n V$. Der Ansatz

$$v^{\wedge} = \sum_{\varphi \in R^{\epsilon}_{S_n}} \kappa_{\varphi} b^{\wedge}_{\varphi}$$

ergibt

$$\kappa_{\varphi} = n! \lambda^{\otimes}_{\varphi}(v^{\wedge}),$$

wobei

$$\lambda^{\otimes}_{\varphi}(v^{\wedge}) = \frac{1}{n!} \sum_{\pi \in S_n} \epsilon(\pi) \lambda^{\otimes}_{\varphi}(v^{\otimes}_{\pi^{-1}}) = \frac{1}{n!} \sum_{\pi} sgn(\pi) \prod_i v_{\pi^{-1}(i),\varphi(i)}$$

$$= \frac{1}{n!} \sum_{\pi} sgn(\pi) \prod_i v_{\pi^{-1}(i),\varphi(i)} = \frac{1}{n!} \underbrace{\det(v_{i,\varphi(k)})}_{:=\det(A[id|\varphi])},$$

wenn $v_i = \sum v_{ij} b_j$ und $A[\mathrm{id} \mid \varphi]$ die Matrix bezeichnet, die in der i-ten Zeile und k-ten Spalte das Element $v_{i,\varphi(k)}$ enthält, für eine Abbildung $\varphi \in m^n$.
Wir erhalten also die folgende Darstellung von $v^\wedge$ als Linearkombination aus den Basiselementen:

5.7.10 $$v^\wedge = \sum_\varphi \det(A[id \mid \varphi]) b^\wedge_\varphi,$$

wobei über die streng monoton wachsenden $\varphi \in m^n$ zu summieren ist.

Wichtige *Anwendungen* von Symmetrieklassen sind beispielsweise Herleitungen von Sätzen über Determinanten, z. B. dera *Cauchy-Binet-Formel,* die abschließend bewiesen werden soll. Der Beweis verwendet Koeffizientenvergleich bei einer Anwendung einer geeigneten Linearform auf einen geeigneten Tensor. Die Anwendung der Linearformen $\mu^\otimes \in Hom_\mathbb{K}(\otimes^n V, \mathbb{K})$ mit den Faktoren

$$\mu_t := \sum_i b_{it} \lambda_i,$$

auf den Tensor $v^\wedge$ mit den Faktoren $v_i := \sum_k a_{i,k} b_k$ ergibt

$$\begin{aligned}
\mu^\otimes(v^\wedge) &= \frac{1}{n!} \sum_\pi sgn(\pi) \mu^\otimes(v^\otimes_{\pi^{-1}}) \\
&= \frac{1}{n!} \sum_\pi sgn(\pi) \prod_t \mu_t(v_{\pi^{-1}(t)}) \\
&= \frac{1}{n!} \sum_\pi sgn(\pi) \prod_t \sum_i b_{i,t} \lambda_i (\sum_k a_{\pi^{-1}(t),k} b_k) \\
&= \frac{1}{n!} \sum_\pi sgn(\pi) \prod_t \sum_i b_{i,t} a_{\pi^{-1}(t),i} \\
&= \frac{1}{n!} \sum_\pi sgn(\pi) \prod_t \underbrace{\sum_i a_{pi^{-1}(t),i} b_{i,t}}_{=(AB)_{\pi^{-1}(t),t}} \\
&= \frac{1}{n!} det(AB)
\end{aligned}$$

Andererseits gilt, mit 5.7.10,

$$\begin{aligned}
\mu^\otimes(v^\wedge) &= \mu^\otimes(\sum_\varphi \kappa_\varphi b^\wedge_\varphi) \\
&= {}_{b)} \sum_\varphi \det(A[id \mid \varphi]) \mu^\otimes(b^\wedge_\varphi),
\end{aligned}$$

und schließlich ist noch

$$\mu^\otimes(b^\wedge_\varphi) = \frac{1}{n!} \sum_\pi sgn(\pi) \prod_t \mu_t(b_{\varphi(\pi^{-1}(t))})$$

$$
\begin{aligned}
&= \frac{1}{n!}\sum_{\pi} sgn(\pi)\prod_{t}\sum_{i} b_{i,t}\lambda_i(b_{\varphi\circ\pi^{-1}(t)}) \\
&= \frac{1}{n!}\sum_{\pi} sgn(\pi)\prod_{t} b_{\varphi\circ\pi^{-1}(t),t} \\
&= \frac{1}{n!}\sum_{\pi} sgn(\pi)\prod_{t} b_{\varphi(t),\pi^{-1}(t)} \\
&= \frac{1}{n!}\det(B[\varphi \mid id])
\end{aligned}
$$

Insgesamt erhalten wir damit die gewünschte Formel zur Berechnung der Determinante eines Produkts von (evtl. rechteckigen) Matrizen, die

5.7.11 Die Formel von Cauchy–Binet

$$
\det(AB) = \sum_{\varphi\in m^n, str.mon.w.} \det(A[id \mid \varphi])\det(B[\varphi \mid id]).
$$

5.8 Induzierte Transformationen

Ist $f \in End_{\mathbb{K}}(V)$ und V_G^χ eine lineare Symmetrieklasse, dann definieren wir die von f auf V_G^χ induzierte Transformation f_G^χ *als kommutative Ergänzung des Diagramms*

$$\begin{tikzcd} \times^n V \arrow[r, "\Delta"] \arrow[dr, "(v_0,\ldots,v_{n-1}) \mapsto f(v_0)\Delta\ldots\Delta\, f(v_{n-1})"'] & V_G^\chi \arrow[d, "f_G^\chi"] \\ & V_G^\chi \end{tikzcd}$$

Es gilt also

5.8.1 $$f_G^\chi(v^\Delta) = f(v_0)\Delta\ldots\Delta f(v_{n-1}).$$

5.8.2 Beispiele

- $f_{\{1\}}^\iota =: \overset{n}{\otimes} f$, *die* $n-$fache Tensorpotenz *von* f,
- $f_{S_n}^\iota =: \bigvee^n f$, *die* n-fache symmetrische Tensorpotenz,
- $f_{S_n}^\epsilon =: \bigwedge^n f$, *die* n-fache schiefsymmetrische Tensorpotenz.

5.8.3 Hilfssatz Ist $f \in End_{\mathbb{K}}(V), G \leq S_n, \chi\colon G \to \mathbb{K}^*$ ein eindimensionaler Charakter, dann gilt:

- $\overset{n}{\otimes} f$ ist mit dem Symmetrisierungsoperator P_G^χ vertauschbar,
- $\overset{n}{\otimes} f$ läßt V_G^χ invariant,
- f_G^χ ist die Einschränkung $\overset{n}{\otimes} f \downarrow V_G^\chi$ auf die Symmetrieklasse,
- $\forall\, g \in End_{\mathbb{K}}(V)$: $f_G^\chi \circ g_G^\chi = (f \circ g)_G^\chi$,
- $V_G^\chi \neq \{0\} \implies [f$ regulär $\iff f_G^\chi$ regulär$]$,
- $\mathrm{Rang}(f) = r \Longrightarrow \mathrm{Rang}(f_G^\chi) = |r^n \cap R_G^\chi|$.

Beweis: Die erste Behauptung ist trivial und impliziert die zweite. Die dritte folgt durch Nachrechnen und ergibt die vierte. Die fünfte Behauptung beweisen wir in zwei Schritten:

a) Ist f regulär, dann auch $\overset{n}{\otimes} f$ und somit f_G^χ, wegen der Gültigkeit der dritten Behauptung.

b) Sei f nicht regulär, $v \in \mathrm{Kern}(f)\backslash\{0_V\}, \varphi \in R_G^\chi, i \in \mathrm{Bild}(\varphi)$, $\mathcal{B} = (b_0, \ldots, b_{m-1})$ eine Basisfolge von V mit $b_i = v$. Dann gilt:

$$f_G^\chi(b_\varphi^\Delta) = \overset{n}{\otimes} f \circ P_G^\chi(b_\varphi^\otimes) = P_G^\chi \circ \overset{n}{\otimes} f(b_\varphi^\otimes) = P_G^\chi(\bigotimes_j f(b_{\varphi(j)})) = 0,$$

da es ein $j \in n$ gibt mit $f(b_{\varphi(j)}) = f(v) = 0$. Das steht aber im Widerspruch zur Regularität von f_G^χ.

Schließlich zur sechsten Behauptung: Sei $\mathrm{Rang}(f) = r$ *und*

$$\mathrm{Kern}(f) = \ll b_r, \ldots, b_{m-1} \gg, \ \mathrm{Bild}(f) = \ll v_0, \ldots, v_{r-1} \gg, \ f(b_i) = v_i, i \in r.$$

Dann ist $V = \ll b_0, \ldots, b_{m-1} \gg$. Weil es für alle $\varphi \in R_G^\chi$, die nicht in r^n liegen, ein i gibt mit $\varphi(i) \geq r$, haben wir für solche φ :

$$f_G^\chi(b_\varphi^\Delta) = 0.$$

Ist dagegen $\varphi \in R_G^\chi \cap r^n$, dann gilt

$$\varphi \in R_G^\chi \cap r^n \Longrightarrow f_G^\chi(b_\varphi^\Delta) = v_\varphi^\Delta,$$

das Bild $f_G^\chi(b_\varphi^\Delta)$ ist also ein Basisvektor. Es folgt deshalb, wie behauptet

$$\mathrm{Rang}(f_\varphi^\Delta) = |R_G^\chi \cap r^n|.$$

□

5.8.4 Satz $\mathcal{B}$ und $\mathcal{C}$ seien Basisfolgen für V, R_G das Repräsentantensystem der lexikographisch kleinsten Elemente aus den Bahnen von G auf m^n, $\mathcal{B}^\Delta := (\ldots, b_\varphi^\Delta, \ldots)$, $\mathcal{C}^\Delta := (\ldots, c_\varphi^\Delta, \ldots)$, $\varphi \in R_G^\chi$, die resultierenden Basisfolgen von V_G^χ (lexikographisch angeordnet). Setzt man $A := M(\mathcal{C}, f, \mathcal{B})$, $D := M(\mathcal{C}^\Delta, f_G^\chi, \mathcal{B}^\Delta)$, so gilt:

- $d_{\varphi\psi} = \frac{1}{|G_\varphi|} \sum_{\pi \in G} \chi(\pi^{-1}) \prod_i a_{\varphi(\pi(i)),\psi(i)}$.
- Ist A eine Dreiecksmatrix, dann auch D; die Hauptdiagonalelemente sind
$$d_{\varphi\varphi} = \prod_i a_{ii}^{|\varphi^{-1}[\{i\}]|}.$$
- Sind $\epsilon_0, \ldots, \epsilon_{m-1}$ die Eigenwerte von A, dann sind die Eigenwerte von D gerade die Zahlen
$$\prod_i \epsilon_i^{|\varphi^{-1}[\{i\}]|}, \ \varphi \in R_G^\chi.$$

Beweis: Ist $f_G^\chi(b_\psi^\Delta) = \sum_\varphi d_{\varphi\psi} c_\varphi^\Delta$, dann gilt nach 5.7.3:

$$d_{\varphi\psi} = \frac{|G|}{|G_\varphi|} \lambda_\varphi^\otimes(f_G^\chi(b_\varphi^\Delta)).$$

Nun ist aber (wegen $f(b_k) = \sum_j a_{jk} c_j$)

$$\lambda_\varphi^\otimes(f_G^\chi(b_\psi^\Delta)) = \frac{1}{|G|}\sum_\pi \chi(\pi^{-1})\lambda_\varphi^\otimes(\bigotimes_i f(b_{\psi(\pi^{-1}(i))}))$$

$$= \frac{1}{|G|}\sum_\pi \chi(\pi^{-1})\prod_i \lambda_{\varphi(i)}(\sum_j a_{j,\psi(\pi^{-1}(i))}c_j)$$

$$= \frac{1}{|G|}\sum_\pi \chi(\pi^{-1})\prod_i a_{\varphi(i),\psi(\pi^{-1}(i))},$$

woraus sich die erste Behauptung ergibt.
Sei A untere *Dreiecksmatrix, d.h. $a_{ik} = 0$, falls $i < k$. Da $\psi \in R_G^\chi$ das kleinste Element seiner Bahn ist, gilt $\psi \circ \pi^{-1} \geq \psi$, für alle $\pi \in G$. Also gilt folgende Implikation:*

$$\varphi < \psi \Longrightarrow \prod_i a_{\varphi(i),\psi(\pi^{-1}(i))} = 0.$$

Entsprechend gilt, wenn A obere *Dreiecksmatrix ist:*

$$\varphi > \psi \Longrightarrow \prod_i a_{\varphi(\pi(i)),\psi(i)} = 0.$$

Für die Hauptdiagonalelemente gilt in letzterem Fall:

$$d_{\varphi\varphi} = \frac{1}{|G_\varphi|}\sum_\pi \chi(\pi^{-1})\prod_i a_{\varphi(\pi(i)),\varphi(i)} = \prod_i a_{\varphi(i),\varphi(i)} = \prod_j a_{jj}^{|\varphi^{-1}[\{j\}]|}.$$

f_G^χ ist unabhängig von $\mathcal{B}$ und $\mathcal{C}$ definiert. $\mathcal{B}$ kann deshalb so gewählt werden, daß A Dreiecksmatrix ist, die in der Hauptdiagonalen ihre Eigenwerte stehen hat. Die dritte Behauptung folgt deshalb aus der zweiten.

□

5.8.5 Definition Sei $n \in \mathbf{N}, G \leq S_n$ und $\chi: G \to \mathbb{K}$ ein (nicht notwendig eindimensionaler) Charakter von G. Die Abbildung

$$d_G^\chi: \mathbb{K}^{n\times n} \to \mathbb{K}: (a_{ik}) \mapsto \sum_{\pi\in G}\chi(\pi^{-1})\prod_i a_{\pi(i),i}$$

heißt dann die zu (G,χ) gehörende *verallgemeinerte Matrixfunktion.*

5.8.6 Beispiele

- $per := d_{S_n}^\iota: (a_{ik}) \mapsto \sum_{\pi\in S_n}\prod_i a_{\pi(i),i}$, die *Permanente.*
- $det = d_{S_n}^\epsilon$, die Determinante.
- $d_{\{1\}}^\iota: (a_{ik}) \mapsto \prod_i a_{ii}$.
- 5.8.4 liest sich jetzt so:

  $$d_{\varphi\psi} = \frac{1}{|G_\varphi|} d_G^\chi(A[\varphi|\psi]),$$

  wenn wir mit $A[\varphi|\psi]$ die Untermatrix $(a_{\varphi(i),\psi(k)})_{ik}$ bezeichnen.

5.9 Die Tensoralgebra

Ist $\mathbb{K}$ jetzt ein *beliebiger Körper*, V ein $\mathbb{K}$-Vektorraum, dann können wir, wie im folgenden beschrieben wird, aus den Vektorräumen $\otimes^n V, n \in \mathbb{N}$, eine wichtige Algebra konstruieren.

5.9.1 Definition $(H,+)$ *sei eine abelsche Halbgruppe. Ein $\mathbb{K}$-Vektorraum W heißt dann ein H-*graduierter Vektorraum*, wenn es zu jedem $h \in H$ einen Unterraum W_h gibt mit $W = \oplus_h W_h$.*

5.9.2 Beispiele

- Ist $(H,+)$ abelsche Halbgruppe und, für jedes $h \in H$, W_h ein $\mathbb{K}$-Vektorraum, dann heißt der $\mathbb{K}$-Vektorraum
$$\prod_{h\in H} W_h := \{f\colon H \to \bigcup W_h \mid f(h) \in W_h\}.$$
das *Produkt* der W_h, und
$$\coprod_{h\in H} W_h := \{f\colon H \to \bigcup W_h \mid f(h) \in W_h,\ \text{fast alle } f(h) = 0\}$$
das *Coprodukt* (oder auch die *äußere direkte Summe* der W_h). Es gilt $\coprod W_h = \oplus \overline{W}_h$, mit
$$\overline{W}_h := \{f \in \coprod W_h \mid \forall h' \neq h\colon f(h') = 0\} \simeq W_h.$$

- Gemäß den vorangegangenen Beispielen haben wir zu V den $\mathbb{N}$-graduierten $\mathbb{K}$-Vektorraum
$$T(V) := \coprod_{n\in\mathbb{N}} \overset{n}{\otimes} V = \bigoplus_{n\in\mathbb{N}} \overline{\overset{n}{\otimes} V} = \overline{\mathbb{K}} \oplus \overline{V} \oplus (\overline{V \otimes V}) \oplus \ldots,$$
das heißt:
$$T(V) = \{f\colon \mathbb{N} \to \bigcup_{n\in\mathbb{N}} \overset{n}{\otimes} V \mid f(n) \in \overset{n}{\otimes} V,\ \text{fast alle } f(n) = 0\}.$$

- Der Polynomring $\mathbb{K}[x]$ enthält die Unterräume $\ll x^n \gg$, für alle $n \in \mathbb{N}$, und es gilt
$$\mathbb{K}[x] = \bigoplus_{n\in\mathbb{N}} \ll x^n \gg,$$
$\mathbb{K}[x]$ ist also ein $\mathbb{N}$-graduierter $\mathbb{K}$-Vektorraum.

5.9.3 Definition $W = \bigoplus W_h$ *sei H-graduierter $\mathbb{K}$-Vektorraum.*

- *$w \in W$ heißt* homogen vom Grad h, *wenn $w \in W_h$. $U \leq W$ heißt* homogen, *wenn U von homogenen Elementen erzeugt wird.*

- *Ist W zudem noch eine $\mathbb{K}$-Algebra, d.h. ist eine Multiplikation definiert, so daß $(W, +, \cdot)$ zu einem Ring wird und*

  $$\kappa(ww') = (\kappa w)w' = w(\kappa w')$$

  gilt, für alle $\kappa \in \mathbb{K}$, $w, w' \in W$, dann heißt W eine H-graduierte $\mathbb{K}$-Algebra*, wenn*

  $$W_h \cdot W_{h'} \subseteq W_{h+h'}.$$

  Ein Ideal I heißt dabei homogenes Ideal*, wenn es (als $\mathbb{K}$-Algebra) von homogenen Elementen erzeugt wird, d.h. wenn es homogene Elemente gibt, so daß jedes $i \in I$ eine Linearkombination aus endlichen Produkten dieser homogenen Elemente ist.*

5.9.4 Beispiele

- $\mathbb{K}[x]$ ist eine $\mathbb{N}$-graduierte Algebra.
- $\{f \in \mathbb{K}[x] \mid \text{Polynomgrad } f \leq 5\}$ ist homogener Unterraum, jedoch kein Ideal.
- Die Polynome mit lauter geraden Koeffizienten bilden ein homogenes Ideal.

5.9.5 Definition *Seien $V = \bigoplus V_h$ und $W = \bigoplus W_h$, beide H-graduiert, $f \in Hom_{\mathbb{K}}(V, W)$. Dann heißt f* homogen vom Grad h'*, wenn gilt*

$$f(V_h) \subseteq W_{h+h'}.$$

5.9.6 Beispiele

- $f: \mathbb{K}[x] \to \mathbb{K}[x]: p \mapsto x^n \cdot p$ ist homogen vom Grad n.
- $\mathbb{K}[x]$ kann auch als $\mathbb{Z}$-graduierte $\mathbb{K}$-Algebra geschrieben werden:

  $$\mathbb{K}[x] := \bigoplus_{z \in \mathbb{Z}} V_z, \ V_z := \ll x^z \gg, \text{ falls } z \geq 0, V_z := \{0\}, \text{ falls } z < 0.$$

  Die Differentiation $\frac{d}{dx}$: $\mathbb{K}[x] \to \mathbb{K}[x]$ ist, als lineare Abbildung auf diesem $\mathbb{Z}$-graduierten $\mathbb{K}$-Vektorraums, homogen vom Grad -1.

5.9.7 Hilfssatz *Ist $W = \bigoplus W_h$ H-graduiert, U ein homogener Unterraum dann gilt:*

- *$U = \bigoplus_h (U \cap W_h)$, ist also ebenfalls H-graduiert.*
- *$W/U = \bigoplus_h \nu_U(W_h) \simeq \coprod_h W_h/U_h$, mit $U_h := U \cap W_h$.*

Graduierung vererbt sich also auf homogene Unterräume und auf Faktorräume nach homogenen Unterräumen.

Beweis:

i) Ist E ein Erzeugendensystem von U aus lauter homogenen Elementen $\neq 0$, dann ist $E = \bigcup (E \cap W_h)$, also

$$U = \sum_h U \cap W_h = \bigoplus_h (U \cap W_h).$$

ii) Wir betrachten $\alpha_h \colon W_h + U \to W_h/U_h, w_h + U \mapsto w_h + U_h$. Dafür gelten offensichtlich die folgen Äquivalenzen:

$$w_h + U = w'_h + U \Leftrightarrow w_h - w'_h \in U \Leftrightarrow w_h - w'_h \in U_h \Leftrightarrow w_h + U_h = w'_h + U_h.$$

Liest man diese von links nach rechts, so erweist sich α_h als wohldefiniert. Von rechts nach links gelesen ergibt sich die Injektivität. Die Surjektivität ist trivial. Es gilt demnach

$$\nu_U(W_h) \simeq W_h/U_h.$$

Es bleibt also nur noch zu zeigen, daß die Summe $\sum \nu_U(W_h)$ direkt ist. Ein Ansatz $0 = \sum_{h \in} \nu_U(w_h)$ ergibt aber

$$0 = \sum w_h + U \Longrightarrow \sum w_h \in U \Longrightarrow w_h \in U_h \Longrightarrow \nu_U(w_h) = 0.$$

□

Zur Einführung einer Multiplikation auf $T(V)$ definieren wir zunächst Abbildungen

$$\mu \colon \overset{m}{\otimes} V \times \overset{n}{\otimes} V \to \otimes^{m+n} V$$

als bilineare Forsetzungen von

$$(v^{\otimes}, w^{\otimes}) \mapsto v^{\otimes} \otimes w^{\otimes}.$$

Mit Hilfe dieser μ definieren wir die *Multiplikation*, also eine Abbildung

$$\mu \colon T(V) \times T(V) \to T(V) \colon (f,g) \mapsto \mu(f,g),$$

jetzt durch Angabe der Werte von $\nu(f,g)$ an den Stellen $n \in \mathbb{N}$:

$$\mu(f,g)(n) := \sum_{(i,k)\colon\, i+k=n} \mu(f(i), g(k)).$$

(Die Summe auf der rechten Seite dieser definierenden Gleichung ist endlich!) $T(V)$ zusammen mit der durch ν definierten Multiplikation heißt *die Tensoralgebra über* V, oder auch, genauer, die Algebra der *kontravarianten* Tensoren über V. Die Gültigkeit der Rechengesetze kann man leicht nachprüfen. Offenbar ist diese Algebra $\mathbb{N}$-graduiert.

5.9.8 Beispiel

$$(2 + v_0 \otimes v_1 + v'_0 \otimes v'_1 \otimes v'_2) \otimes (w_0 \otimes w_1 \otimes w_2 + w'_0 \otimes w'_1) =$$

$$2(w_0 \otimes w_1 \otimes w_2) + 2(w'_0 \otimes w'_1) + v_0 \otimes v_1 \otimes w_0 \otimes w_1 \otimes w_2$$

$$+ v_0 \otimes v_1 \otimes w'_0 \otimes w'_1 + v'_0 \otimes v'_1 \otimes v'_2 \otimes w_0 \otimes w_1 \otimes w_2 + v'_0 \otimes v'_1 \otimes v'_2 \otimes w'_0 \otimes w'_1.$$

5.10 Die Graßmann-Algebra

In diesem Paragraphen werden wir den Faktorring $T(V)/I$ betrachten, aus der Tensoralgebra und einem interessanten homogenen Ideal I. Zur Vorbereitung dient die folgende Bemerkung:

5.10.1 Satz *Ist $A = \bigoplus A_h$ eine H-graduierte $\mathbb{K}$-Algebra, I ein homogenes Ideal, dann hat A/I die H-Graduierung*

$$A/I = \bigoplus \nu_I(A_h) \simeq \coprod A_h/I_h, \ \mathit{mit}\ I_h := I \cap A_h.$$

Beweis: Nach 5.9.7 gilt jedenfalls

$$A/I = \bigoplus \nu_I(A_h) \text{ (als } \mathbb{K}\text{-Vektorraum)}.$$

Ist $x_h \in \nu_I(A_h), y_{h'} \in \nu_I(A_{h'})$, etwa

$$x_h = \nu_I(a_h), y_{h'} = \nu_I(b_{h'}),$$

dann gilt:

$$x_h \cdot y_{h'} = \nu_I(a_h \cdot b_{h'}) = \nu_I(c_{h+h'}) \in \nu_I(A_{h+h'}).$$

Die angegebene Zerlegung $A/I = \bigoplus \nu_I(A_h)$ ist also auch eine Graduierung für die Ringstruktur A/I.

□

Wir betrachten in der Algebra

$$T(V) = \overline{\mathbb{K}} \oplus \overline{V} \oplus \overline{V \otimes V} \oplus \ldots$$

der kovarianten Tensoren das Ideal I, das von den Tensoren $v \otimes v, v \in V$, erzeugt wird, also

$$I := \{\sum u^{\otimes} \otimes v \otimes v \otimes w^{\otimes} \mid u^{\otimes}, w^{\otimes} \in T(V), v \in V\}.$$

I ist homogenes Ideal, $T(V)/I$ also nach obigem Satz eine $\mathbb{N}-$graduierte $\mathbb{K}-$Algebra, die sogenannte *Graßmann-Algebra* (oder auch: die äußere Algebra) über V:

5.10.2 $$\bigwedge V := T(V)/I = \bigoplus_{n \in \mathbb{N}} \nu_I(\overline{\overset{n}{\otimes} V}).$$

5.10.3 Satz *Der Unterraum $\nu_I(\overline{\otimes^n V})$ von $\bigwedge V$ ist isomorph zu $\bigwedge^n V$. Wir können also schreiben:*

$$\bigwedge V \simeq \coprod_{n \in \mathbb{N}} \bigwedge^n V = \bigoplus_{n \in \mathbb{N}} \overline{\bigwedge^n V}.$$

Beweis: Wir betrachten die Abbildungen $\mu_n := \nu_I \circ \otimes$, also

$$\mu_n : \overset{n}{\times} V \to \nu_I(\overline{\overset{n}{\otimes} V}), (v_0, \ldots, v_{n-1}) \mapsto \nu_I(v^{\otimes}).$$

i) $\mu_n \in M_n(V, \nu_I(\overline{\otimes^n V}, S_n, \epsilon)$: Aus der Definition von I folgt

$$\mu_n(v_0, \ldots, v_{i-1}, v_i + v_{i+1}, v_i + v_{i+1}, v_{i+2}, \ldots, v_{n-1}) = 0,$$

also gilt

$$\mu_n(v_0, \ldots, v_{n-1}) = -\mu_n(v_0, \ldots, v_{i-1}, v_{i+1}, v_i, v_{i+2}, \ldots, v_{n-1}),$$

und das wiederum impliziert

$$\mu(v_0, \ldots, v_{n-1}) = \epsilon(\pi)\mu_n(v_{\pi(0)}, \ldots, v_{\pi(n-1)}).$$

ii) Nach i) kann μ_n über $\bigwedge^n V$ faktorisiert werden:

$$\begin{tikzcd} \times^n V \arrow[r, "P^{\epsilon}_{S_n}"] \arrow[dr, "\mu_n"'] & \bigwedge^n V \arrow[d, "\exists_1 \varphi_n"] \\ & \nu_I(\overline{\otimes^n V}) \end{tikzcd}$$

es bleibt zu zeigen, daß φ_n Isomorphismus ist.

iii) $I \cap \overline{\otimes^n V} = Kern\ P^{\epsilon}_{S_n}$:

a) $Kern\ P^{\epsilon}_{S_n} \subseteq I \cap \overline{\otimes^n V}$: Wir benutzen die Gleichungen

$$Kern\ P^{\epsilon}_{S_n} = Bild(id - P^{\epsilon}_{S_n}) = \langle (id - P^{\epsilon}_{S_n})v^{\otimes} \mid v_i \in V \rangle$$

und beachten, daß

$$(id - P^{\epsilon}_{S_n})v^{\otimes} = \frac{1}{n!}\sum_{\pi}(v^{\otimes} - \epsilon(\pi)v^{\otimes}_{\pi}) \in I \cap \overline{\overset{n}{\otimes} V},$$

denn

$$\nu_I(v^{\otimes} - \epsilon(\pi)v^{\otimes}_{\pi}) = \mu_n(v_0, \ldots, v_{n-1}) - \epsilon(\pi)\mu_n(v_{\pi(0)}, \ldots, v_{\pi(n-1)}) =_{i)} 0.$$

b) $I \cap \overline{\otimes^n V} \subseteq Kern\ P^{\epsilon}_{S_n}$ ist trivial, da jedes $u^{\otimes} \otimes v \otimes v \otimes w^{\otimes}$ aus $\overline{\otimes^n V}$ im Kern von $P^{\epsilon}_{S_n}$ liegt.

iv) Wir zeigen als nächstes, daß die Menge

$$\{\mu_n(b_{\varphi(0)}, \ldots, b_{\varphi(n-1)}) \mid \varphi \textit{ streng monoton wachsend}\}_{\neq}$$

linear unabhängig ist:

$$0 = \sum \kappa_{\varphi}\mu_n(b_{\varphi(0)}, \ldots, b_{\varphi(n-1)}) = \sum \kappa_{\varphi}\nu_I(b^{\otimes}_{\varphi}) = \nu_I(\sum \kappa_{\varphi}b^{\otimes}_{\varphi}).$$

Also gilt

$$\sum \kappa_\varphi b_\varphi^\otimes \in I \cap \overline{\overset{n}{\otimes} V} =_{iii)} Kern\ P_{S_n}^\epsilon,$$

und das wiederum impliziert

$$0 = \sum \kappa_\varphi P_{S_n}^\epsilon (b_\varphi^\otimes) = \sum \kappa_\varphi b_\varphi^\wedge,$$

also $\kappa_\varphi = 0$, für alle streng monoton wachsenden $\varphi \in m^n$.

v) $\langle Bild \mu_n \rangle = \nu_I(\overline{\otimes^n V})$: μ_n ist multilinear, kann also auch über $\overline{\otimes^n V}$ faktorisiert werden, die kommutative Ergänzung des Diagramms ist offenbar die Einschränkung von ν_I:

$$\begin{tikzcd}
\times^n V \arrow[r] \arrow[dr, "\mu_n"'] & \overline{\otimes^n V} \arrow[d, "\exists_1 \nu_I \downarrow \overline{\otimes^n V}"] \\
& \nu_I\left(\overline{\otimes^n V}\right)
\end{tikzcd}$$

Daraus folgt die Behauptung.

vi) Die behauptete Isomorphie $\nu_I(\overline{\otimes^n V}) \simeq \bigwedge^n V$ folgt jetzt aus einem Vergleich der Dimension.

a) Aus iv) und v) ergibt sich die Ungleichung

$$dim_{\mathbb{K}}(\bigwedge^n V) = \binom{m}{n} \leq dim_{\mathbb{K}}(\nu_I(\overline{\overset{n}{\otimes} V})).$$

b) Aus v) und ii) ergibt sich das kommutative Diagramm

$$\begin{tikzcd}
\overline{\otimes^n V} \arrow[r, "P_{S_{\underline{n}}}^\epsilon"] \arrow[dr, "\nu_I \downarrow \overline{\otimes^n V}"'] & \overline{\bigwedge^n V} \arrow[d, "\exists_1\ \varphi_n"] \\
& \nu_I\left(\overline{\otimes^n V}\right)
\end{tikzcd}$$

und wegen der Surjektivität von $\nu_I \downarrow \overline{\otimes^n V}$ ergibt dies

$$dim_{\mathbb{K}}(\bigwedge^n V) = \binom{m}{n} \geq dim_{\mathbb{K}}(\nu_I(\overline{\overset{n}{\otimes} V})).$$

$\square$

5.10.4 Folgerung $dim_{\mathbb{K}}(\bigwedge V) = 2^{dim_{\mathbb{K}}(V)}$.

Beweis:
$$\sum_{n=0}^{dim_{\mathbb{K}}(V)} \binom{dim_{\mathbb{K}}(V)}{n} = (1+1)^{dim_{\mathbb{K}}(V)}.$$

□

Wir benutzen für die Multiplikation in $\bigwedge V$ das Symbol $\wedge$ anstelle von $\cdot$, schreiben also z.B. $V^{\wedge} \wedge w^{\wedge}$ anstelle von $v^{\wedge} \cdot w^{\wedge}$.

5.10.5 Anwendungen

i) Nach 5.7.10 gilt z.B.

5.10.6 $$v^{\wedge} = 0 \iff dim_{\mathbb{K}}(\langle v_0, \ldots, v_{n-1}\rangle) < n.$$

Man kann das jetzt aber auch ohne Benutzung der Determinante herleiten!

ii) Ein weiteres Beispiel für Anwendungen von $\bigwedge V$ ist die Herleitung der Cramerschen Regel. Ist etwa $Ax = b$ ein vorgegebenes lineares Gleichungssystem mit regulärer Koeffizientenmatrix $A \in \mathbb{K}^{n\times n}$ und Spaltenvektoren a_k, dann setzt man

$$v_t := a_0 \wedge \ldots \wedge a_{t-1} \wedge a_{t-1} \wedge \ldots \wedge a_{n-1}, t \in n.$$

Statt $Ax = b$ können wir auch schreiben

$$\sum_k x_k a_k = b.$$

um daraus x_k zu ermitteln multiplizieren wir (in $\bigwedge V$) mit v_t. Es ergibt sich dabei einerseits

$$v_t \wedge b = \sum x_k(v_t \wedge a_k) = x_t(v_t \wedge a_t).$$

Andererseits gilt aber auch

$$v_t \wedge b = (-1)^{n-t} a_0 \wedge \ldots \wedge a_{t-1} \wedge b \wedge a_{t+1} \wedge \ldots \wedge a_{n-1},$$

und das ist, nach 5.7.10,

$$= (-1)^{n-t} det(a_0, \ldots, a_{t-1}, b, a_{t+1}, \ldots, a_{n-1}) b_0 \wedge \ldots b_{n-1},$$

während 5.7.10 auch noch folgendes hergibt:

$$v_t \wedge a_t = (-1)^{n-t} det A b_0 \wedge \ldots \wedge b_{n-1}.$$

Also ergibt sich insgesamt

$$x_t = \frac{det(a_0, \ldots, a_{t-1}, b, a_{t+1}, \ldots, a_{n-1})}{det\ A};$$

das ist die *Cramersche Regel.*

Die wichtigste Eigenschaft von $\bigwedge V$ ist natürlich eine Universaleigenschaft. Wir leiten sie aus der Universaleigenschaft der Tensoralgebra $T(V)$ her.

5.10.7 Hilfssatz *Ist A eine $\mathbb{K}-$Algebra, $f\colon V \to A$ linear, dann kann f eindeutig zu einem Algebrahomomorphismus $\overline{f}\colon T(V) \to A$ fortgesetzt werden.*

Beweis: Sei, für $n > 0$, die Abbildung f_n definiert als kommutative Ergänzung von

$$\begin{tikzcd}
\times^n V \arrow[r, "\otimes"] \arrow[dr, "(v_0,\ldots,v_{n-1}) \mapsto \prod_i f(v_i)"'] & \otimes^n V \arrow[d, "\exists_1\ f_n"] \\
& A
\end{tikzcd}$$

Weiter sei

$$\overline{f}\colon T(V) \to A; \sum x_n \mapsto \sum f_n(x_n),$$

wobei $\sum x_n$ die (eindeutige!) Zerlegung von $x \in T(V)$ in homogene Elemente bedeute. Dann gilt, mit $f_0\colon \kappa \mapsto \kappa 1_A$:

i) $\overline{f} \downarrow \overline{V} = f$, das ist klar.

ii) a) $\overline{f}$ ist multiplikativ auf zerlegbaren Tensoren:

$$\begin{aligned}
\overline{f}(u_1 \otimes \ldots \otimes u_p \otimes v_1 \otimes \ldots \otimes v_q) &= f_{p+q}(u^{\otimes} \otimes v^{\otimes}) \\
&= \prod_i f(u_i) \prod_j f(u_j) \\
&= f_p(u^{\otimes}) f_q(v^{\otimes}) \\
&= \overline{f}(u^{\otimes})\overline{f}(v^{\otimes}).
\end{aligned}$$

$\overline{f}$ ist also Homomorphismus bzgl. Multiplikation.

b) $\overline{f}(1_{T(V)}) = \overline{f}(1_{\mathbb{K}}) = f_0(1_{\mathbb{K}}) = 1_A$.

c) $\overline{f}$ ist eindeutig bestimmt, da $T(V)$ (als Algebra) das Erzeugendensystem $\{v^{\otimes} \mid v_i \in V\} \cup \{1_{\mathbb{K}}\}$ hat.

□

5.10.8 Satz *Die Einbettung $V \simeq \overline{V} \mapsto T(V)$ ist universell bezüglich der Klasse $\mathcal{F}$ aller linearen Abbildungen von V in $\mathbb{K}-$Algebren und der Klasse $\mathcal{L}$ aller Algebrenhomomorphismen.*

5.10.9 Satz $\nu_I \downarrow \overline{V}\colon \overline{V} \to \bigwedge V$ *ist universell bzgl. der Klasse $\mathcal{F}$ der linearen Abbildungen φ von V in $\mathbb{K}-$Algebren mit der Eigenschaft*

$$\varphi(v)^2 = 0.$$

Beweis: Sei $\varphi: V \to A$ mit $\varphi(v)^2 = 0$, $\overline{\varphi}$ die Fortsetzung von φ auf $T(V)$ gemäß 5.10.7. Dann gilt $I \subseteq Kern\ \overline{\varphi}$, so daß eine kommutative Ergänzung h folgenden Diagramms existiert:

$$\begin{tikzcd} T(V) \arrow[r, "\nu_I"] \arrow[dr, "\overline{\varphi}"'] & \bigwedge V \arrow[d, "\exists_1 h"] \\ & B \end{tikzcd}$$

Für $f: \nu_I \downarrow \overline{V} \simeq V$ gilt deshalb:

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "\varphi"'] & \bigwedge V \arrow[d, "\exists_1 h"] \\ & A \end{tikzcd}$$

$\square$

Kapitel 6

Aus der Gruppentheorie

6.1 Präsentation von Gruppen

Es geht jetzt um die Beschreibung von Gruppen durch Erzeugende und Relationen, also z. B. um die genaue Beschreibung dessen, was Zeilen wie die folgende bedeuten:

$$G := \langle x, y \mid x^2 = y^2 = (xy)^3 = 1 \rangle.$$

Aus dieser Zeile erhält man durch naives Rechnen die folgende Multiplikationstafel:

	1	x	y	xy	yx	xyx
1	1	x	y	xy	yx	xyx
x	x	1	xy	y	xyx	yx
y	y	yx	1	xyx	x	xy
xy	xy	xyx	x	yx	1	y
yx	yx	y	xyx	1	xy	x
xyx	xyx	xy	yx	x	y	1

also eine zu S_3 isomorphe Gruppe. Diese ist aber, im Gegensatz zu der Definition der S_3 als Menge von Bijektionen auf der Menge $\{0, 1, 2\}$, eine sogenannte *abstrakte Gruppe*, die gegeben ist durch die Erzeugenden x und y und durch die drei Gleichungen ("Relationen") $x^2 = 1$, $y^2 = 1$ sowie $(xy)^3 = 1$. Ein Isomorphismus zwischen beiden Gruppen ist offenbar die Fortsetzung von $x \mapsto (01), y \mapsto (12)$. Es stellt sich also die Frage, ob man jede Gruppe so beschreiben, und wie man mit abstrakten Gruppen rechnen kann.

6.1.1 Definition *Ist M eine Menge, φ eine Abbildung von M in eine Gruppe G, dann heißt (G, φ) eine von M* frei erzeugte Gruppe, *wenn*

$$\varphi: M \to G$$

universell ist bzgl. der Klasse $\mathcal{F}$ der Abbildungen von M in Gruppen und der Klasse $\mathcal{L}$ der Homomorphismen zwischen Gruppen.

6.1.2 Satz *Zu jeder Menge M existieren von ihr frei erzeugte Gruppen, je zwei von ihnen sind isomorph.*

Beweis: Wir konstruieren eine Gruppe $F(M)$, die die gewünschten Eigenschaften hat.

- Zu M nehmen wir eine bijektive, aber zu M *disjunkte* Menge M^{-1} hinzu, einem $m \in M$ entspreche dabei das Element $m^{-1} \in M^{-1}$. Diese beiden Mengen fassen wir zum *Alphabet* $A := M \cup M^{-1}$ zusammen und bilden die Halbgruppe A^* aus den endlichen Wörtern w über A, einschließlich des *leeren Worts* ϵ, mit dem Anfügen ("Concatenation") als Verknüpfung:
$$ww' = (a_{i_1} \cdots a_{i_k})(a_{j_1} \cdots a_{j_l}) := a_{i_1} \cdots a_{i_k} a_{j_1} \cdots a_{j_l}.$$
Das leere Wort ist hier offenbar das neutrale Element, die Halbgruppe A^* ist also ein *Monoid.*

- Auf A^* definieren wir eine *Reduktion* ρ wie folgt: Aus $w \in A^*$ streichen wir, von links beginnend, alle Teilfolgen der Form mm^{-1} oder $m^{-1}m$, und wiederholen dies, bis keine derartige Teilfolge mehr auftritt. Die Menge $F(M)$ dieser *reduzierten Worte* $\rho(w)$ versehen wir mit der folgenden Verknüpfung:
$$*: F(M) \times F(M) \to F(M): (w, w') \mapsto \rho(ww').$$

- Das Paar $(F(M), *)$ ist eine Gruppe, und die Einbettung
$$M \hookrightarrow F(M): m \mapsto m$$
hat die gewünschte Universaleigenschaft, denn eine Abbildung f von M in eine Gruppe H läßt sich offenbar mittels folgender Abbildung γ von $F(M)$ nach H faktorisieren:
$$\gamma: F(M) \to H: \quad \underbrace{m_{i_1}^{b_1} \cdots m_{i_k}^{b_k}}_{oBdA:\ b_i \in \{-1,+1\}} \quad \mapsto f(m_{i_1})^{b_1} \cdots f(m_{i_k})^{b_k}.$$
Außerdem ist γ offensichtlich Homomorphismus und eindeutig bestimmt, da $F(M)$ von M erzeugt wird.

- Die Isomorphie folgt nach 5.1.3.

$\square$

Wegen der Isomorphie zweier von M frei erzeugter Gruppen können wir $F(M)$ auch als *die* von M frei erzeugte Gruppe bezeichnen. Die Elemente $m \in M$ heißen *freie Erzeugende* dieser Gruppe. Das Attribut *frei* bedeutet dabei insbesondere, daß zwischen den Elementen aus M keine nichttrivialen *Relationen* bestehen, das sind Gleichungen der Form
$$\underbrace{m_{i_1}^{b_1} \cdots m_{i_k}^{b_k}}_{reduziert\ !} = \epsilon.$$

Weil M ganz $F(M)$ erzeugt, heißt $F(M)$ auch die freie Gruppe *über* M.

6.1.3 Beispiele Für die freie Gruppe über einem einzigen Element gilt offenbar

$$F(\{m\}) \simeq \mathbf{Z}: m \mapsto 1,$$

während alle anderen freien Gruppen nicht kommutativ sind:

$$1 < |M| \Longrightarrow F(M) \text{ ist nicht abelsch.}$$

Aus der Existenz und Isomorphie zwischen den freien Gruppen über M ergibt sich noch

6.1.4 Folgerung *Für Gruppen und freie Gruppen gilt:*

- *Jede Gruppe G mit Erzeugendensystem M ist auf eindeutige Weise epimorphes Bild von $F(M)$.*
- *Jede Gruppe G ist homomorphes Bild einer freien Gruppe und damit isomorph zu einer Faktorgruppe einer freien Gruppe.*

6.1.5 Definition *Wird G von M erzeugt und ist γ der (eindeutig bestimmte) Epimorphismus von $F(M)$ auf G, dann heißt der Kern*

$$K_M := \mathrm{Kern}(\gamma)$$

der definierende Kern *von G bzgl. M. Ist E* Normalteilererzeugendensystem *von K_M (d. h. K_M ist der kleinste Normalteiler in $F(M)$, der E enthält), dann heißt das System der Gleichungen*

$$\gamma(e) = 1_G,\ e \in E,$$

ein System definierender Relationen *von G bzgl. M. Gibt es endliche E mit dieser Eigenschaft, dann heißt G* endlich präsentierbar, *und*

$$\langle M \mid \gamma(e) = 1, e \in E\rangle = G$$

heißt eine endliche Präsentation *von G.*

6.1.6 Beispiel Die zyklische Gruppe der Ordnung 5 kann offenbar wie folgt präsentiert werden:

$$G := \langle a \mid a^5 = 1\rangle.$$

$\diamond$

Es stellt sich jetzt natürlich die Frage, wie man eine Präsentation einer vorgegebenen Gruppe berechnen kann. Sei deshalb G eine Gruppe mit einem Erzeugendensystem $M = \{m_1, \ldots, m_n\}$, sowie Relationen

$$R_1(m_1, \ldots, m_n) = 1_G, \ldots, R_s(m_1, \ldots, m_n) = 1_G.$$

Wir fragen, ob dies schon eine Präsentation von G ist, d. h. ob folgendes gilt:

$$G' := \langle M \mid R_1 = \ldots = R_s = 1\rangle \simeq G.$$

Um dies zu entscheiden, kann nun im endlichen Fall wie folgt argumentiert werden:

- Nach 6.1.4 ist G homomorphes Bild von G', es gilt also insbesondere $|G'| \geq |G|$.
- Die Isomorphie ist somit bewiesen, wenn $|G'| = |G|$ verifiziert werden kann. Dazu genügt es, eine Untergruppe $U' \leq G'$ zu finden, für die $|U'|$ und $|G'/U'|$ ermittelt werden können.

6.1.7 Beispiel Betrachten wir als Beispiel den eingangs aufgeführten Fall der Präsentation

$$G := \langle x, y \mid x^2 = y^2 = (xy)^3 = 1 \rangle.$$

Wir wollen zeigen, daß dies tatsächlich eine Präsentation der symmetrischen Gruppe S_3 ist. Dazu schließen wir wie folgt:

- S_3 ist jedenfalls homomorphes Bild von G, also ist $3!$ ein Teiler von $|G|$.
- Die von y erzeugte Untergruppe U hat die Ordnung 2, zum Nachweis der vermuteten Isomorphie genügt also allein die Verifizierung von $|G/U| = 3$.
- Zur Berechnung des gewünschten Index beachten wir, daß U, xU und yxU Linksnebenklassen sind (ob sie verschieden sind brauchen wir hier nicht einmal zu wissen!). Bei Linksmultiplikation mit Erzeugenden ergeben sich daraus *keine weiteren* Linksnebenklassen, der Index ist also tatsächlich ≤ 3.

Zusammenfassend haben wir daraus die folgende Präsentation ermittelt:

$$S_3 \simeq \langle x, y \mid x^2 = y^2 = (xy)^3 = 1 \rangle.$$

6.1.8 Definition *Die Gruppen*

$$D_m := \langle x, y \mid x^2 = y^2 = (xy)^m = 1 \rangle$$

heißen Diedergruppen.

6.1.9 Hilfssatz *D_m hat die Ordnung $2m$ und enthält mit $\langle xy \rangle$ einen Normalteiler der Ordnung m.*

Beweis: Sei $z := xy$. D_m wird erzeugt von $\{x, z\}$, und die Gleichungen

$$yx = y^{-1}x^{-1} = z^{-1}$$

zeigen, daß sich jedes Element von D_m in der Form $x^r z^s$ schreiben läßt, mit $0 \leq r \leq 1$ und $0 \leq s \leq m-1$. Es gilt demnach $|D_m| = 2m$. Die Normalteilereigenschaft von $\langle z \rangle$ ergibt sich aus $xzx = yzy$.

□

6.1.10 Definition *Sind G und H Gruppen und ist $h \mapsto \bar{h}$ ein Homomorphismus von H in die Gruppe $Aut(G)$ der Automorphismen von G, dann wird $G \times H$ zu einer Gruppe durch die Setzung*

$$(g, h)(g', h') := (g \cdot \bar{h}(g'), hh').$$

Diese Gruppe heißt das semidirekte Produkt von G mit H bzgl. $h \mapsto \bar{h}$.

Der Beweis von 6.1.9 zeigt, daß die Diedergruppe semidirektes Produkt einer zyklischen Gruppe der Ordnung m mit einer zyklischen Gruppe der Ordnung 2 ist.

6.2 Endliche Gruppenoperationen

Erinnern wir uns jetzt an die im Paragraphen 2 bereits eingeführten *Gruppenoperationen*

$$G \times M \to M\colon (g,m) \mapsto gm, \text{ mit } g(g'm) = (gg')m, \text{ und } 1m = m.$$

Wir schrieben dafür abkürzend auch ${}_GM$, da G von *links* auf M operiert, und wir hatten festgestellt, daß eine solche Operation wichtige Teilstrukturen auf M und auf G definiert:

- Zu jedem $m \in M$ gehört seine *Bahn*

  $$G(m) := \{gm \mid g \in G\}.$$

  Je zwei Bahnen $G(m)$ und $G(m')$ sind entweder gleich oder disjunkt. Eine Transversale T der Menge

  $$G \backslash\!\backslash M := \{G(m) \mid m \in M\}$$

  aller Bahnen liefert also eine Partition von M:

  $$M = \dot\cup_{t \in T} G(t).$$

- Andererseits gehört zu $m \in M$ sein *Stabilisator*

  $$G_m := \{g \in G \mid gm = m\},$$

  eine Untergruppe von G. Die Abbildung $\varphi\colon gm \mapsto gG_m$ ist eine Bijektion zwischen der Bahn $G(m)$ und der Menge $G/G_m := \{gG_m \mid g \in G\}$ der Linksnebenklassen des Stabilisators von m. Dies zeigt uns auch, daß die Abbildung φ ein *G–Isomorphismus* ist, d.h. eine mit der Operation vertauschbare Bijektion, gilt doch

  $$g'\varphi(gm) = \varphi(g'(gm)).$$

  Wir wollen Operationen ${}_GM$ und ${}_GN$ als *ähnlich* bezeichnen, wenn solche G–Isomorphismen zwischen M und N existieren, und wir wollen diese Situation so abkürzen: ${}_GM \simeq {}_GN$. Es gilt also insbesondere

  6.2.1 $${}_G(G(m)) \simeq {}_G(G/G_m).$$

- Zu jedem $g \in G$ gehört die Menge seiner *Fixpunkte*:

  $$M_g := \{m \in M \mid gm = m\}.$$

Viele Strukturen aus Mathematik und Naturwissenschaften lassen sich als Bahnen, Stabilisatoren oder Fixpunktmengen definieren und werden dadurch einer Untersuchung mit Hilfe der im folgenden beschriebenen Methoden zugänglich.

6.2.2 Anwendungen (bilaterale Klassen) Ist wieder G eine Gruppe, dann operiert jede Untergruppe U von $G \times G$ in folgender Weise kanonisch auf G:

$$U \times G \to G : ((u, u'), g) \mapsto ugu'^{-1}.$$

Die Bahnen dieser Operation heißen *bilaterale Klassen* von U auf G. Beispiele sind:

- Die Konjugiertenklassen von Elementen $g \in G$, sind die bilateralen Klassen der *Diagonalen*

  $$U := \Delta(G \times G) := \{(g, g) \mid g \in G\}.$$

  Wir bezeichnen sie so:

  $$C^G(g') := \Delta(G \times G)(g) = \{gg'g^{-1} \mid g \in G\}.$$

  Für die entsprechenden Stabilisatoren schreiben wir:

  $$C_G(g') := \Delta(G \times G)_g = \{g \in G \mid gg'g^{-1} = g'\}.$$

  Sie heißen auch *Zentralisatoren*.

- Die *Links*- bzw. *Rechtsnebenklassen* von Untergruppen $H \leq G$ sind die Bahnen von $U := 1 \times H$ bzw. $U := H \times 1$:

  $$(1 \times H)(g) = gH \text{ bzw. } (H \times 1)(g) = Hg.$$

- Die (H, K)-Doppelnebenklassen in G sind die Bahnen der Untergruppe $U := H \times K$ von $G \times G$, mit $H, K \leq G$:

  $$(H \times K)(g) = HgK.$$

$\diamond$

Als konkretes Beispiel wollen wir die Konjugiertenklassen von Elementen π der symmetrischen Gruppe S_n beschreiben. Dazu erinnern wir uns an die *Zyklenschreibweise* für Permutationen, π kann als Produkt disjunkter zyklischer Faktoren geschrieben werden:

$$\pi = \prod_{\nu=0}^{z(\pi)-1} (j_\nu \pi j_\nu \ldots \pi^{l_\nu - 1} j_\nu).$$

Wenn wir mit ρ konjugieren, erhalten wir

$$\rho\pi\rho^{-1} = \prod_{\nu=0}^{z(\pi)-1} (\rho j_\nu, \rho\pi j_\nu, \ldots, \rho\pi^{l_\nu - 1} j_\nu),$$

beim Konjugieren bleiben also die Längen der zyklischen Faktoren erhalten.

Umgekehrt sind Permutationen mit zyklischen Faktoren derselben Längen zueinander konjugiert: Für

$$\pi = \prod_{\nu=0}^{z(\pi)-1} (j_\nu, \pi j_\nu, \ldots, \pi^{l_\nu - 1} j_\nu)$$

und

$$\sigma = \prod_{\nu=0}^{z(\pi)-1} (k_\nu, \pi k_\nu, \ldots, \pi^{l_\nu - 1} k_\nu)$$

gilt mit

$$\rho = \prod_{\nu=0}^{z(\pi)-1} \begin{pmatrix} j_\nu, \pi j_\nu, \ldots, \pi^{l_\nu - 1} j_\nu \\ k_\nu, \pi k_\nu, \ldots, \pi^{l_\nu - 1} k_\nu \end{pmatrix}$$

die Gleichung

$$\sigma = \rho \pi \rho^{-1}.$$

6.2.3 Folgerung *Zwei Permutationen aus $\pi, \sigma \in S_n$ sind genau dann konjugiert, wenn die schwach monoton fallende Folge*

$$\alpha(\pi) := (\alpha_0(\pi), \alpha_1(\pi), \ldots)$$

der Längen der zyklischen Faktoren von π gleich der entsprechenden Folge $\alpha(\sigma)$ ist.

Solche schwach monoton fallenden Folgen $\alpha = (\alpha_0, \alpha_1, \ldots)$ natürlicher Zahlen mit $\sum \alpha_i = n$ heißen *Partitionen* der Zahl n, wir kürzen dies wie folgt ab:

$$\alpha \vdash n.$$

Die Konjugiertenklasse der Elemente mit Zykellängenpartition α bezeichnen wir wie folgt:

$$C^\alpha := \{\pi \in S_n \mid \alpha(\pi) = \alpha\}.$$

Sind sowohl G als auch M endlich, dann nennen wir ${}_GM$ eine endliche Operation. Für diese zeigt 6.2.1, daß die Länge der Bahn eines Gruppenelementes dem Index seines Stabilisators gleicht:

6.2.4 $$|G(m)| = \frac{|G|}{|G_m|}.$$

Beispielsweise gilt, für $\pi \in C^\alpha$,

$$|C^\alpha| = \frac{n!}{|C_{S_n}(\pi)|},$$

und man kann sich leicht überlegen, wie man die Ordnung $|C_{S_n}(\pi)|$ dieses Zentralisators erhält: Bezeichnet a_i die Anzahl der i-Zyklen von π, dann heißt $a(\pi) := (a_1(\pi), \ldots, a_n(\pi))$ der *Zykeltyp* von π. Allgemeiner heißen n-Tupel

$a = (a_1, \ldots, a_n)$ mit $a_i \in \mathbb{N}$ und $\sum_i i \cdot a_i = 2$ Zykeltypen von n, und man kann das wie folgt abkürzend bezeichnen:

$$a \vdash\!\dashv n.$$

Der Zykeltyp $a(\pi)$ der Permutationen mit Zykelpartition $\alpha(\pi)$ charakterisiert also ebenfalls die Konjugiertenklasse

$$C^a = C^\alpha = \{\pi \mid a(\pi) = a.$$

Eine Permutation $\rho \in S_n$ liegt genau dann im Zentralisator von π, wenn sie jeden i-Zyklus von π wieder in einen i-Zyklus von π transformiert. Die Anzahl solcher ρ ist

$$|C_{S_n}(\pi)| = \prod_{i=1}^{n} i^{a_i(\pi)} a_i(\pi)!.$$

Insgesamt gilt also

$$|C^a| = \frac{n!}{\prod_{i=1}^n i^{a_i(\pi)} a_i(\pi)!}.$$

Die Gleichung 6.2.4 ist eine sehr einschneidende Bedingung an die Bahnlängen, wie die folgende Anwendung zeigt, H. Wielandts berühmter Beweis (Archiv d. Math. 1959) für die Existenz von Untergruppen von Primzahlpotenzordnung p^r, wenn p^r in $|G|$ aufgeht.

6.2.5 Anwendung(die Existenz von Sylow-Untergruppen) Wir betrachten die natürliche Operation von G auf der Menge ihrer Teilkomplexe K der Ordnung p^r:

$$G \times \binom{G}{p^r} : (g, K) \mapsto gK,$$

definiert s durch $|G| = p^r p^s q$, mit zu p teilerfremdem q, und rechnet nach, daß die Anzahl der p^r–Teilmengen von G nicht durch p^{s+1} teilbar sein kann:

$$|\binom{G}{p^r}| = \frac{|G| \cdot (|G| - 1) \cdots (|G| - (p^r - 1))}{p^r \cdot 1 \cdots (p^r - 1)}$$

zeigt (alle im Nenner stehenden Potenzen von p können gekürzt werden, so daß als einziger durch p noch teilbarer Faktor $|G|/p^r$ verbleibt), Es muß also eine Bahn $G(K)$ geben, deren Länge nicht durch p^{s+1} teilbar ist. Der Stabilisator G_K ist dann eine Untergruppe der Ordnung p^r, wie man leicht durch Abschätzung der Ordnung dieses Stabilisators nach oben und unten beweist:

i) Aus 6.2.4 folgt

$$|G(K)| = \frac{|G|}{|G_K|}.$$

Diese Zahl ist nicht durch p^{s+1} teilber, in $|G_K|$ steckt also mindestens die p-Potenz p^r als Faktor, d. h. $|G_K| \geq p^r$.

ii) Ist $\kappa \in K$, dann liegt $G_K \cdot \kappa$ ganz in K. Dieses Komplexprodukt besteht aber aus genau $|G_K|$ verschiedenen Elementen, es muß also $|G_K| \leq p^r$ gelten.

□

Es gibt also, unter anderem, Untergruppen von G, deren Ordnung der *maximalen* p-Potenz gleicht, die in $|G|$ aufgeht. Diese heißen p-*Sylowuntergruppen.* Man kann zeigen (Übungsblatt), daß diese sämtlich zueinander konjugiert sind und ihre Anzahl kongruent ist zu 1 modulo p. Insgesamt haben wir damit den folgenden sehr wichtigen Existenzsatz bewiesen:

6.2.6 Der Satz von Sylow *Ist G eine endliche Gruppe, p eine Primzahl, dann gibt es zu jeder p-Potenz p^r, die in $|G|$ aufgeht, eine Untergruppe $U \leq G$ von dieser Ordnung.*

◇

Im folgenden seien alle Operationen endlich. Die Gleichheit von Bahnlänge und Stabilisatorindex ermöglicht außerdem (wie bereits in 2 gezeigt) den Beweis der Tatsache, daß die Gesamtanzahl der Bahnen die mittlere Anzahl der Fixpunkte der Elemente von G ist: Das Lemma von Cauchy–Frobenius besagt, daß für endliche Operationen gilt:

$$|G \backslash\backslash M| = \frac{1}{|G|} \sum_{g \in G} |M_g|.$$

Bei Anwendungen dieses Lemmas sollte man beachten, daß die Summation über die ganze Gruppe wesentlich vereinfacht werden kann, weil die Anzahl der Fixpunkte konstant auf den Konjugiertenklassen ist:

6.2.7 $$|M_g| = |M_{g'gg'^{-1}}|,$$

denn $m \mapsto g'm$ ist eine Bijektion zwischen M_g und $M_{g'gg'^{-1}}$. Kurz: $g \mapsto |M_g|$ ist eine *Klassenfunktion.* Man kann sich demnach bei Anwendung des Lemmas von Cauchy-Frobenius auf die Summation über die Elemente einer Transversalen T der Konjugiertenklassen beschränken,

6.2.8 $$|G \backslash\backslash M| = \frac{1}{|G|} \sum_{t \in T} |C^G(t)| \cdot |M_t|.$$

6.2.9 Anwendung(Symmetrieklassen von Abbildungen) Aus einer gegebenen Operation ${}_GX$ erhält man auf natürliche Weise die folgende Operation von G auf Y^X :

$$G \times Y^X \to Y^X : (g, f) \mapsto f \circ g^{-1},$$

d. h. gf ist die Abbildung $\widetilde{f}$ mit

$$\widetilde{f}(x) := f(g^{-1}x).$$

Diese Operation ist uns bereits in der Linearen Algebra I und dann nochmals im Zusammenhang mit Symmetrieklassen von Tensoren begegnet (dort war $Y^X = m^n$). Wir nennen deshalb die Bahnen von G auf Y^X unter dieser Operation *Symmetrieklassen von Abbildungen.* Wir wissen, daß $f \in Y^X$ genau dann ein Fixpunkt von g ist, wenn f auf den Zyklen von $\bar{g}$ (der von g auf X induzierten Permutation) konstant ist. Der Charakter dieser Operation ist demnach

$$\chi(g) = |Y|^{z(\bar{g})},$$

wenn $z(\bar{g})$ die Anzahl der zyklischen Faktoren von $\bar{g}$ ist. Das Lemma von Cauchy–Frobenius, zusammen mit 6.2.8, ergibt demnach für die Anzahl der Symmetrieklassen die Formel

$$|G \backslash\!\backslash Y^X| = \frac{1}{|G|} \sum_{t \in T} |C^G(t)| \cdot |Y|^{z(\bar{t})}.$$

Eine konkrete Anwendung ist die Berechnung der Anzahl unnumerierter Graphen mit p Punkten. Diese gleicht

$$|S_p \backslash\!\backslash 2^{\binom{p}{2}}| = \sum_{a \dashv p} \frac{2^{z(\bar{a})}}{\prod_i i^{a_i} a_i!}.$$

◇

Das Lemma von Cauchy–Frobenius läßt mehrere Verfeinerungen zu und findet zahlreiche Anwendungen. Zum Beispiel folgt sofort, daß für jede endliche Operation ${}_G M$ die folgende Kongruenz richtig ist:

6.2.10
$$\sum_{g \in G} |M_g| \equiv 0 \ (|G|).$$

Hiermit kann man, durch Betrachtung geeigneter Operationen, viele zahlentheoretische Kongruenzen herleiten, zum Beispiel die Kongruenz von Fermat aus

$$|C_p \backslash\!\backslash m^p| = \frac{1}{p}(m^p + (p-1)m),$$

was die Kongruenz

$$m^p \equiv m \ (p)$$

ergibt.

In vielen Fällen ist man mit der Bestimmung der bloßen Anzahl der Bahnen bzw. der als solche definierten Strukturen aber noch nicht zufrieden, sondern man will diese konstruieren. Hat man hierfür einen Algorithmus gefunden, kann man oft mit Hilfe des Lemmas von Cauchy–Frobenius die Gesamtanzahl berechnen und dieses Ergebnis dann bei der Konstruktion als *stopping rule* verwenden, d.h. als Abbruchbedingung.

Definiert man Graphen mit vorgegebener Punkteanzahl p als Bahnen der symmetrischen Gruppe S_p auf $Y^X := 2^{\binom{p}{2}}$, so läßt sich ihre Anzahl demnach mit

Hilfe des Cauchy–Frobenius Lemmas berechnen (vgl. Übungsblatt). In diesem Zusammenhang kann man aber auch die Frage nach der Zahl der Graphen mit p Punkten *und k Kanten* stellen, und tatsächlich läßt sich diese mit folgender Verfeinerung des Lemmas beantworten:

6.2.11 Das Lemma von Cauchy–Frobenius, gewichtete Form: *Ist $_GM$ eine endliche Operation und $w: M \to R$ eine auf den Bahnen konstante Abbildung von M in einen kommutativen Ring R, der den Körper $\mathbb{Q}$ der rationalen Zahlen als Teilring enthält. Dann gilt, für eine Transversale $\mathcal{T}$ der Menge $G \backslash\backslash M$ der Bahnen:*

$$\sum_{t \in \mathcal{T}} w(t) = \frac{1}{|G|} \sum_{g \in G} \sum_{m \in M_g} w(m).$$

Der Beweis dieser Verfeinerung ist im wesentlichen derselbe wie der der ursprünglichen Version, der sogenannten *konstanten Form* des Cauchy-Frobenius Lemmas.

6.2.12 Anwendung(Abzählung von Symmetrieklassen nach Gewicht) Ist $_GX$ wieder eine endliche Operation, $_G(Y^X)$ die davon induzierte Operation auf Y^X, zu einer gegebenen endlichen Menge Y.
Wir können Y als eine Menge von Unbestimmten auffassen und folgende Gewichtsfunktion auf Y^X betrachten, die offenbar konstant auf den Bahnen ist:

$$w: Y^X \to \mathbb{Q}[Y]: f \mapsto \prod_{x \in X} f(x).$$

Weil f genau dann Fixpunkt ist, wenn f auf den zyklischen Faktoren von $\bar{g}$ konstant ist, gilt entsprechend für die Summe aller Fixpunkte von g :

$$\sum_{f \in (Y^X)_g} w(f) = \prod_i \Big(\sum_{y \in Y} y^i\Big)^{a_i(\bar{g})}.$$

Mit der gewichteten Form des Lemmas von Cauchy-Frobenius erhalten wir demnach die

6.2.13 Folgerung *Die Anzahl der Bahnen von G auf Y^X mit vorgegebenem Gewicht*

$$\prod_{y \in Y} y^{b_y}$$

ist der Koeffizient dieses Monoms in dem Polynom

$$\frac{1}{|G|} \sum_{g \in G} \prod_i \Big(\sum_{y \in Y} y^i\Big)^{a_i(\bar{g})}.$$

Wir können dies Abzählung nach Gewicht sofort anwenden. Unser konkretes Standardbeispiel sind wieder die Graphen. Wir wollen die Anzahl der Graphen mit p Punkten *und mit k Kanten* bestimmen.

Hier war $Y = 2 = \{0, 1\}$, wir ersetzen diese Menge zur Verdeutlichung durch

$$Y = \{y_0, y_1\}.$$

Ist jetzt $f \in Y^{\binom{p}{2}}$ wieder ein numerierter Graph, dann ist sein Gewicht bei der oben eingeführten Gewichtsfunktion w gleich

$$w(f) = y_0^{\binom{p}{2}-k} y_1^k,$$

wenn k die Kantenzahl ist. Die erzeugende Funktion für die Abzählung der Graphen mit p Punkten nach Kantenzahl ist als, entsprechend der gewichteten Form des Lemmas von Cauchy-Frobenius, gleich

6.2.14
$$\frac{1}{p!} \sum_{\pi \in S_p} \prod_{i=1}^{\binom{p}{2}} (y_0^i + y_1^i)^{a_i(\bar{\pi})}.$$

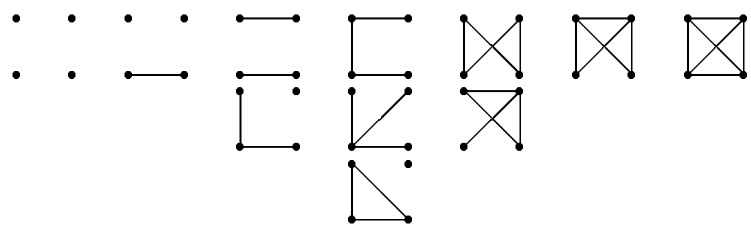
Man kann sich hierbei ganz offenbar das Leben noch dadurch erleichtern, daß man y_0 durch 1 und y_1 durch y ersetzt, es ergibt sich dann als erzeugende Funktion das Polynom

6.2.15
$$\frac{1}{p!} \sum_{\pi \in S_p} \prod_{i=1}^{\binom{p}{2}} (1 + y^i)^{a_i(\bar{\pi})} = \sum_{a \vdash p} \frac{1}{\prod_{i=1}^{p} i^{a_i} a_i!} \prod_{j=1}^{\binom{p}{2}} (1 + y^j)^{\overline{a_i}}.$$

Der Koeffizient von y^k in diesem Polynom ist die Anzahl der Graphen mit p Punkten und k Kanten! Für $p = 4$ erhält man das Polynom

$$1 + y + 2y^2 + 3y^3 + 2y^4 + y^5 + y^6.$$

Das entspricht der folgenden Skizze dieser Graphen:



$\diamond$

Eine andere Verfeinerung der Bahnenabzählung geht von folgender Bemerkung aus:

6.2.16
$$G_{gm} = gG_m g^{-1}.$$

Aus ihr folgt nämlich, daß die Stabilisatoren der Elemente $m' \in G(m)$ die Klasse der zu G_m konjugierten Untergruppen bilden:

6.2.17
$$\widetilde{G_m} := \{gG_m g^{-1} \mid g \in G\} = \{G_{gm} \mid g \in G\}.$$

Die Bahnen, deren Elementstabilisatoren die Konjugiertenklasse $\widetilde{U}$ bilden, heißen Bahnen *vom Typ* $\widetilde{U}$. Man kann also die Frage stellen, wie groß die Anzahl der Bahnen vom Typ $\widetilde{U}$ ist, zu vorgegebener Untergruppe U von G. Diese Menge nennen wir das *Stratum von* U und bezeichnen sie so:

$$G\backslash\!\backslash_{\widetilde{U}} M := \{G(m) \mid G_m \in \widetilde{U}\}.$$

Zur Berechnung ihrer Ordnung bringen wir sie mit der Menge der Fixpunkte von U in Zusammenhang:

$$M_U := \{m \in M \mid \forall g \in U : gm = m\}.$$

Bezeichnet ζ die Zetafunktion auf dem Verband $U(G)$ der Untergruppen von G, also

$$\zeta(U,V) := \begin{cases} 1 & U \leq V \\ 0 & \text{sonst}, \end{cases}$$

dann gilt, wenn wieder T eine Transversale von $G\backslash\!\backslash M$ bezeichnet:

$$\begin{aligned}
|M_U| &= \sum_{m : U \leq G_m} 1 = \sum_{V : U \leq V \leq G} \sum_{m : V = G_m} 1 = \sum_{V : U \leq V \leq G} \frac{1}{|\widetilde{V}|} \sum_{m : G_m \in \widetilde{V}} 1 \\
&= \sum_{V : U \leq V \leq G} \frac{|G/V|}{|\widetilde{V}|} \sum_{t \in T : G_t \in \widetilde{V}} 1 = \sum_{V : U \leq V \leq G} \frac{|G/V|}{|\widetilde{V}|} |G\backslash\!\backslash_{\widetilde{V}} M|, \\
&= \sum_{V \leq G} \zeta(U,V) \frac{|G/V|}{|\widetilde{V}|} |G\backslash\!\backslash_{\widetilde{V}} M|.
\end{aligned}$$

Wir haben damit folgendes bewiesen:

6.2.18
$$|M_U| = \sum_{V \leq G} \zeta(U,V) \frac{|G/V|}{|\widetilde{V}|} |G\backslash\!\backslash_{\widetilde{V}} M|.$$

Wenn wir die Untergruppen U^i von G, also die Elemente der Menge

$$L(G) := \{U^i \mid U^u \leq G\},$$

so numerieren, daß folgende Bedingung gilt (und dies ist insbesondere bei endlichen Gruppen ohne weiteres möglich):

6.2.19
$$U^i \leq U^k \implies i \leq k,$$

dann ist die Matrix $(\zeta(U^i, U^k))$ (i sei der Zeilenindex, k der Spaltenindex) offensichtlich eine obere Dreiecksmatrix mit Einsen in der Hauptdiagonalen, sie ist damit invertierbar, sogar über $\mathbb{Z}$. Die Inverse

6.2.20
$$\mu(G) := \big(\mu(U^i, U^k)\big) := \zeta(G)^{-1}$$

ist ebenfalls obere Dreiecksmatrix, ihre Elemente sind die Werte der *Möbiusfunktion* auf der Halbordnung

$$(L(G), \subseteq).$$

In dieser Menge gibt es zu je zwei Elementen U^i und U^k ein eindeutig bestimmtes Infimum $U^i \wedge U^k$ und ein eindeutig bestimmtes Supremum $U^i \vee U^k$, nämlich

$$U^i \wedge U^k = U^i \cap U^k, \text{ bzw. } U^i \vee U^k = \langle U^i \cup U^k \rangle.$$

Das Quadrupel

$$(L(G), \subseteq, \wedge, \vee)$$

ist also ein *Verband,* der *Untergruppenverband von G*. Die Werte der Möbiusfunktion auf dem Untergruppenverband von G sind nach obiger Argumentation ganze Zahlen. Die Matrix $\mu(G)$ heißt die *Möbiusmatrix* von G, und $\zeta(G) = (\zeta(U^i, U^k))$ heißt die *Zetamatrix* von G.
Betrachten wir als ein einfaches Beispiel die symmetrische Gruppe S_3. Ihren Untergruppenverband zeigt die Abbildung 6.1,

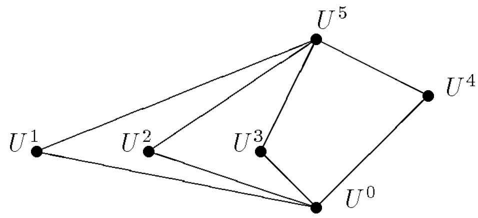


Abbildung 6.1: Der Untergruppenverband von S_3

wobei $U^0 = \langle 1 \rangle$, $U^1 = \langle (01) \rangle$, $U^2 = \langle (02) \rangle$, $U^3 = \langle (12) \rangle$, $U^4 = \langle (012) \rangle$, und $U^5 = S_3$. Es gilt also

$$\zeta(S_3) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 0 & 0 & 1 \\ & & 1 & 0 & 0 & 1 \\ & & & 1 & 0 & 1 \\ & & & & 1 & 1 \\ & & & & & 1 \end{pmatrix},$$

und damit

$$\mu(S_3) = \begin{pmatrix} 1 & -1 & -1 & -1 & -1 & 3 \\ & 1 & 0 & 0 & 0 & -1 \\ & & 1 & 0 & 0 & -1 \\ & & & 1 & 0 & -1 \\ & & & & 1 & -1 \\ & & & & & 1 \end{pmatrix}.$$

Der Untergruppenverband der symmetrischen Gruppe ist im allgemeinen Fall allerdings sehr kompliziert und bisher nicht explizit beschrieben. Dagegen kann man den Untergruppenverband der zyklischen Gruppe

$$C_n = \langle (0, \ldots, n-1) \rangle$$

sehr leicht beschreiben. Dazu verwenden wir (Übungsblatt):

6.2.21 Hilfssatz *Die m-te Potenz $(0, \ldots, n-1)^m$ des Zyklus $(0, \ldots, n-1)$ besteht aus genau $\mathrm{ggT}(m, n)$ disjunkten zyklischen Faktoren, diese haben die Länge $n/\mathrm{ggT}(m, n)$.*

6.2.22 Folgerung *Die Elemente der Ordnung t, t ein Teiler von $|G|$, in C_n sind genau die Potenzen $(0, \ldots, n-1)^m$, bei denen der Exponent m von der Form ist*

$$m = \frac{n}{t} \cdot j,$$

mit

$$j \in t \text{ und } \mathrm{ggT}(j, t) = 1.$$

Beweis: Die Potenz $(0, \ldots, n-1)^m$ hat nach dem Hilfssatz genau dann die Ordnung t, wenn $t = n/\mathrm{ggT}(m, n)$. Gilt dies, dann ist

$$m = \frac{m \cdot n}{n} = \frac{\mathrm{kgV}(m, n) \cdot \mathrm{ggT}(m, n)}{n} = \frac{\mathrm{kgV}(m, n)}{t} = \frac{n}{t} \cdot j,$$

mit $j = \mathrm{kgV}(m, n)/n$. Hierfür gilt

$$\mathrm{ggT}(j, t) = \mathrm{ggT}\Big(\frac{\mathrm{kgV}(m, n)}{n}, \frac{n}{\mathrm{ggT}(m, n)}\Big) = \mathrm{ggT}\Big(\frac{m}{\mathrm{ggT}(m, n)}, \frac{n}{\mathrm{ggT}(m, n)}\Big) = 1.$$

$\square$

Mit Hilfe der Möbiusmatrix ergibt sich durch Invertierung von 6.2.18 jetzt der gesuchte Ausdruck für die Länge des Stratums von U auf M:

6.2.23 $$|G \backslash\!\backslash_{\widetilde{U}} M| = \frac{|\widetilde{U}|}{|G/U|} \sum_V \mu(U, V)|M_V|.$$

Dieses Resultat war im wesentlichen bereits Burnside bekannt, ihm fehlte nur noch der Begriff der Möbius–Inversion. Eine interessante Implikation ist

6.2.24 Folgerung *Die Anzahl der Bahnen der (maximalen) Länge $|G|$ von G auf M gleicht*

$$\frac{1}{|G|} \sum_{U \leq G} \mu(1, U)|M_U|.$$

Die Identität 6.2.23 kann man noch durch Zusammenfassen von Koeffizienten nach Konjugiertenklassen von Untergruppen vereinfachen. Wir betrachten dazu die Menge $\widetilde{U}(G)$ der Konjugiertenklassen von Untergruppen in G:

$$\widetilde{U}(G) := \{\widetilde{U}_0, \ldots, \widetilde{U}_{d-1}\}, \text{ mit Repräsentanten } U_i \in \widetilde{U}_i.$$

Mit ihrer Hilfe definieren wir jetzt die rationalen Zahlen

$$b_{ik} := \frac{|\widetilde{U}_i|}{|G/U_i|} \sum_{V \in \widetilde{U}_k} \mu(U_i, V), \text{ und } B(G) := (b_{ik}).$$

Diese Matrix ist die Inverse der sogenannten *Markentafel*, die bereits Burnside eingeführt hat: $B(G) = M(G)^{-1}$, mit $M(G) := (m_{ik})$, und

$$m_{ik} := \frac{|G/U_k|}{|\widetilde{U}_k|} \sum_{V \in \widetilde{U}_k} \zeta(U_i, V).$$

Mit Hilfe dieser Matrix $B(G)$, die wir als *Burnsidematrix* von G bezeichnen wollen, liest sich jetzt obiges Resultat wie folgt:

6.2.25 Burnsides Lemma: *Ist ${}_GM$ eine endliche Operation, dann gilt:*

$$\begin{pmatrix} \vdots \\ |G \backslash\!\backslash_{\widetilde{U}_i} M| \\ \vdots \end{pmatrix} = B(G) \cdot \begin{pmatrix} \vdots \\ |M_{U_i}| \\ \vdots \end{pmatrix}.$$

Ganz analog ergibt sich auch hier eine Möglichkeit, Bahnen nach Gewicht und Typ abzuzählen:

6.2.26 Burnsides Lemma, gewichtete Form: *Ist ${}_GM$ eine endliche Operation, $w\colon M \to \mathbb{R}[Y]$ konstant auf den Bahnen, dann gilt, wenn $\mathcal{T}_i$ eine Transversale des Stratums vom Typ $\widetilde{U}_i$ ist:*

$$\begin{pmatrix} \vdots \\ \sum_{t \in \mathcal{T}_i} w(t) \\ \vdots \end{pmatrix} = B(G) \cdot \begin{pmatrix} \vdots \\ \sum_{m: U_i \leq G_m} w(m) \\ \vdots \end{pmatrix}.$$

Der Beweis ist im wesentlichen derselbe wie obige Herleitung der konstanten Form.

Viele mathematische Strukturen wie Graphen, Schaltfunktionen, physikalische Zustände usw. können als Bahnen endlicher Gruppen auf endlichen Mengen definiert und damit auch abgezählt werden. Natürlich kann man sie auch auf diese Weise konstruieren, solange die Gruppe und die Menge, auf der die Gruppe operiert, einigermaßen klein sind. Bei Graphen kann man gerade noch einen Katalog aller Graphen mit 10 Punkten herstellen (das sind etwa 12 Millionen Graphen). Man will aber oft auch Graphen mit wesentlich größeren Punktezahlen untersuchen, etwa eine Invariante auf ihre Güte testen. Hierbei hilft die folgende Methode weiter, die eine *gleichverteilte Zufallserzeugung* von Repräsentanten der Bahnen ermöglicht. Sie eröffnet gleichzeitig ein weites Feld für die *experimentelle Mathematik*, kann man hiermit doch immer dann Strukturen gleichverteilt zufallserzeugen, wenn sie als Bahnen endlicher Gruppen auf endlichen Mengen definierbar sind.

6.2.27 Der Algorithmus von Dixon und Wilf: *Ist ${}_GM$ eine endliche Operation, dann erhält man wie folgt Elemente $m \in M$, die über die Bahnen gleichverteilt sind (d.h. für jede Bahn $\omega \in G\backslash\backslash M$ ist die Wahrscheinlichkeit für $m \in \omega$ gleich $|G\backslash\backslash M|^{-1}$):*

- *Zuerst wählt man eine Konjugiertenklasse C von Elementen aus G mit der Wahrscheinlichkeit*
$$p(C) := \frac{|C||M_g|}{|G||G\backslash\backslash M|},$$
für irgendein $g \in C$.
- *Dann entnimmt man C irgendein Element g.*
- *Schließlich wählt man aus der Menge M_g der Fixpunkte von g gleichverteilt ein Element m.*

Beweis: Seien $C_1, \ldots, C_r$ die Konjugiertenklassen von G, mit Repräsentanten $g_i \in C_i$. Dann gilt, nach dem Cauchy-Frobenius Lemma:
$$\sum_{i=1}^{r} p(C_i) = \frac{\sum_i |C_i||M_{g_i}|}{\sum_g |M_g|} = 1,$$
so daß $p(-)$ tatsächlich eine Wahrscheinlichkeitsverteilung definiert. Ist jetzt $\omega \in G\backslash\backslash M$, dann haben wir
$$p(m \in \omega) = \sum_i p(C_i)p(m \in M_{g_i} \cap \omega)$$
$$= \sum_i p(C_i)\frac{|M_{g_i} \cap \omega|}{|M_{g_i}|} = \sum_i \frac{|C_i||M_{g_i}|}{|G||G\backslash\backslash M|}\frac{|M_{g_i} \cap \omega|}{|M_{g_i}|}$$
$$= \frac{1}{|G||G\backslash\backslash M|}\sum_i |C_i||M_{g_i} \cap \omega| = \frac{1}{|G||G\backslash\backslash M|}\sum_g |M_g \cap \omega|.$$
Nun gilt zudem
$$\sum_g |M_g \cap \omega| = \sum_{g\in G}\sum_{m\in M_g\cap\omega} 1 = \sum_{m\in\omega}\sum_{g\in G_m} 1 = \sum_{m\in\omega} |G_m|,$$
das ist aber gleich $|G_y||\omega| = |G|$, für jedes $y \in \omega$, und wir sind fertig.

□

Die Anwendung dieser Methode zur Erzeugung von Repräsentanten von G-Klassen auf Y^X ist recht einfach, da wir die Fixpunkte für diesen Fall sehr genau kennen.

6.2.28 Folgerung *Für endliche ${}_GX$ und Y liefert der folgende Algorithmus Elemente $f \in Y^X$, die gleichverteilt über die G-Klassen sind:*

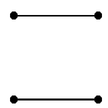


Abbildung 6.2: Der erzeugte Graph

- *Zuerst wählt man eine Konjugiertenklasse C von G mit der Wahrscheinlichkeit*
$$p(C) := \frac{|C||Y|^{c(\bar{g'})}}{\sum_g |Y|^{c(\bar{g})}}, g' \in C.$$
- *Dann nimmt man irgendein $g \in C$, berechnet dessen Zykelindex und konstruiert ein $f \in Y^X$, das Werte $y \in Y$ auf diesen Zykeln annimmt, die zufällig gleichverteilt über Y sind.*

6.2.29 Beispiel Wir wollen nun die Graphen mit vier Punkten zufällig gleichverteilt erzeugen. Die Konjugiertenklassen werden durch die Partitionen (4), $(3,1)$, (2^2), $(2,1^2)$ und (1^4) von 4 charakterisiert. Die Ordnungen dieser Klassen sind 6,8,3,6 und 1. Die Anzahl zyklischer Faktoren der auf der Menge $\binom{4}{2}$ der Punktepaare induzierten Permutationen sind 2,2,4,4 und 6, so daß die Zahl der Fixpunkte auf der Menge $2^{\binom{4}{2}}$ der numerierten Graphen 4,4,16,16 und 64 beträgt. Damit erhält man für die Wahrscheinlichkeiten die Werte (mit $|S_4 \backslash\backslash 2^{\binom{4}{2}}| = 11$)
$$\frac{1}{11}, \frac{4}{33}, \frac{2}{11}, \frac{4}{11}, \frac{8}{33}.$$
Dies liefert uns nach der Multiplikation mit dem Hauptnenner 33 die natürlichen Zahlen 3,4,6,12,8, aus denen sich nach Aufaddieren $3, 7, 13, 25, 33$ ergibt. Ein Zufallsgenerator, der gleichverteilt natürliche Zahlen zwischen 1 und 33 liefert, wird nun verwendet, um C zu wählen. Wird 1, 2 oder 3 erzeugt, so bedeutet dies, daß die erste Konjugiertenklasse gewählt ist. Erhalten wir 4, 5, 6 oder 7, fahren wir mit der zweiten Klasse fort usw. Nehmen wir an, der Generator würde uns 12 ausgeben. Damit ist die dritte Klasse ausgewählt; sie enthält z.B. die Permutation (01)(23). Diese induziert auf der Menge
$$\{a = \{0,1\}, b = \{0,2\}, c = \{0,3\}, d = \{1,2\}, e = \{1,3\}, f = \{2,3\}\}$$
der Punktepaare die Permutation
$$\overline{(01)(23)} = (a)(be)(cd)(f).$$
Ein Zufallsgenerator für Nullen und Einsen wird jetzt eingesetzt, um den zyklischen Faktoren $(a)(be)(cd)(f)$ die Werte 0 oder 1 zuzuordnen. Erzeugt er beispielsweise die Folge 1,0,0,1, so erhalten wir den numerierten Graphen, bei dem nur die Elemente der Paare a und f verbunden sind. Seine Bahn wird durch einen Graphen repräsentiert, wie ihn Abb. 6.2 zeigt.

◇

Falls die symmetrische Gruppe $S_{\underline{n}}$ operiert, entspricht die Wahl einer Konjugiertenklasse der Wahl einer (eigentlichen) Partition von n. Wir sollten uns daher stets vor Augen halten, daß die Anzahl der Partitionen von $n \in \mathbf{N}$ bei zunehmendem n sehr stark wächst.

6.3 Normal- und Kompositionsreihen

Es geht jetzt um die innere Struktur von Gruppen, soweit diese mit Ketten von ineinandergeschachtelten Normalteilern beschrieben werden kann. Erinnern wir uns deshalb zunächst daran, daß *Normalteiler* genau die selbstkonjugierten Untergruppen $N \leq G$ sind, also die Untergruppen mit

$$\forall\, g \in G: \; gNg^{-1} = N.$$

Dies kürzen wir so ab: $N \trianglelefteq G$. Beispiele sind sämtliche Untergruppen vom Index ≤ 2, also z. B. die alternierende Gruppe $A_n \trianglelefteq S_n$. Wir wissen auch, daß genau die Kerne von Homomorphismen Normalteiler sind, und wir haben gesehen, daß die Nebenklassen eines solchen Normalteilers eine Gruppe bilden, die sogenannte *Faktorgruppe*

$$G/N := \{gN \mid g \in G\}.$$

Für jeden Homomorphismus $\varphi: G \to H$ von G in eine Gruppe H ist das Bild isomorph zur Faktorgruppe nach dem Kern:

$$\mathrm{Bild}(\varphi) \simeq G/\mathrm{Kern}(\varphi).$$

Dieses Resultat bezeichnet man auch als den *Homomorphiesatz.* Hierbei ist die Bemerkung wichtig, daß der Homomorphismus φ eine Bijektion zwischen der Menge der Untergruppen von G, die oberhalb $\mathrm{Kern}(\varphi)$ liegen und der Menge der Untergruppen von $\mathrm{Bild}(\varphi)$ stiftet. Diese Bijektion erhält Durchschnitte und Erzeugnisse, ist also ein sogenannter Verbandsisomrophismus

6.3.1 $$(\{U \leq G \mid \mathrm{Kern}(\varphi) \leq U\}, \wedge, \vee) \simeq (\{W \leq \mathrm{Bild}(\varphi)\}, \wedge, \vee).$$

Hat man zwei Normalteiler M und N in G mit $N \subseteq M$, dann gilt

6.3.2 Der erste Isomorphiesatz

$$(G/N)/(M/N) \simeq G/M.$$

Beweis: Wir betrachten das durch einen eindeutig bestimmten Epimorphismus kommutativ ergänzte Diagramm

$$\begin{tikzcd} G \arrow[r, "\nu_N"] \arrow[dr, "\nu_M"'] & G/N \arrow[d, "\exists_1\ \varphi"] \\ & G/M \end{tikzcd}$$

Der Kern von φ ist offenbar M/N, die Behauptung folgt deshalb aus dem Homormorphiesatz. $\square$

Hat man in G einen Normalteiler N und eine Untergruppe U, dann ist UN ebenfalls Untergruppe und es gilt

6.3.3 Der Noethersche Isomorphiesatz

$$U/U\cap N \simeq UN/N.$$

Beweis: Ist ι die Einbettung $U \hookrightarrow UN$ von U in UN, dann hat $\varphi := \nu_N \circ \iota$ als Kern die Untergruppe $U\cap N$, diese ist also ein Normalteiler. Wir haben demnach das eindeutig durch einen Homomorphismus kommutativ ergänzbare Diagramm

$$\begin{tikzcd} U \arrow[r, "\varphi"] \arrow[dr, "\nu_N"'] & UN/N \arrow[d, "\psi"] \\ & U/U\cap N \end{tikzcd}$$

Die Abbildung ψ ist injektiv und surjektiv, also ein Isomorphismus.

□

Wir erinnern in diesem Paragraphen an den Satz von Jordan und Hölder für R-Moduln. Dazu betrachtet man, für eine Gruppe G, sogenannte *Normalketten*, das sind Ketten von Untergruppen, bei denen jedes Glied der Kette Normalteiler des Nachbarn ist:

$$\gamma\colon G = G_0 \trianglerighteq G_1 \trianglerighteq \ldots.$$

Sind hier Nachbarn G_i, G_{i+1} stets verschieden, dann spricht man von einer *eigentlichen* Normalkette. Die Kettenglieder G_i heißen auch die *Elemente* der Kette, die Faktorgruppen G_i/G_{i+1} heißen ihre *Faktoren.* Normalketten der Form

$$\gamma\colon G = G_0 \trianglerighteq \ldots \trianglerighteq G_k = 1$$

heißen *Normalreihen.* Sind γ, γ' Normalketten, dann heißt γ *Verfeinerung* von γ', wenn alle Elemente G_i' von γ' auch in γ vorkommen.

6.3.4 Definition *Zwei Normalketten*

$$\gamma\colon G = G_0 \trianglerighteq \ldots \trianglerighteq G_k$$

und

$$\gamma'\colon G = G_0' \trianglerighteq \ldots \trianglerighteq G_l'$$

heißen äquivalent, *wenn $k = l$ und*

$$\exists\ \pi \in S_k\colon\ G_i/G_{i+1} \simeq G'_{\pi(i)}/G'_{\pi(i)+1}.$$

6.3.5 Beispiele

- $S_4 \triangleright A_4 \triangleright V_4 := \{1,(01)(23),(02)(13),(03)(12)\} \triangleright \{1,(01)(23)\} \triangleright 1.$
- $S_n \triangleright A_n \triangleright 1$, ist für $n \neq 4$ nicht verfeinerbar.

6.3.6 Definition *Eine eigentliche Normalreihe*

$$\gamma\colon G = G_0 \rhd G_1 \rhd \ldots \rhd G_{k-1} \rhd G_k = 1$$

heißt Kompositionsreihe, *wenn alle Faktoren* einfach *sind, d. h. wenn jede G_i maximaler Normalteiler in G_{i-1} ist. Die Zahl k heißt dabei die* Länge *der Kompositionsreihe.*

Kompositionsreihen sind also *nicht mehr verfeinerbare Normalreihen.* Offensichtlich sind eigentliche Normalreihen mit Faktoren von Primzahlordnung Kompositionsreihen, insbesondere also auch die obige Reihe der symmetrischen Gruppe vom Grad 4:

$$S_4 \rhd A_4 \rhd V_4 \rhd C_2 \rhd 1.$$

Jede Gruppe besitzt Normalreihen, z. B. $G \rhd 1$, jedoch nicht immer Kompositionsreihen, was sich an der Gruppe $\mathbb{Z}$ zeigen wird. Endliche Gruppen besitzen Kompositionsreihen.

6.3.7 Hilfssatz *Hat G eine Kompositionsreihe der Länge r, so hat jede eigentliche Normalreihe von G höchstens diese Länge.*

Beweis: Sei

$$\gamma\colon G = G_0 \rhd G_1 \rhd \ldots \rhd G_r = 1$$

eine Kompositionsreihe von G und

$$\gamma'\colon G = G_0 \rhd G_1' \rhd \ldots \rhd G_s' = 1$$

sei eine Normalreihe. Wir führen eine Induktion nach r durch.

$r = 1$: G ist einfach, also auch $s = 1$ und damit die Behauptung richtig.

$r > 1$: Wir unterscheiden drei Fälle:

a. Ist $G_1 = G_1'$, dann gilt nach Induktionsannahme $s - 1 \leq r - 1$, woraus die Behauptung folgt.

b. Ist $G_1' < G_1$, dann ist auch

$$G_1 \rhd G_1' \rhd \ldots \rhd G_s' = 1$$

eine eigentliche Normalreihe, nach der Induktionsannahme gilt also $s \leq r - 1$ und damit auch $s \leq r$.

c. Es bleibt der Fall $G_1' \not\leq G_1$ zu diskutieren. Da G_1 maximal in G ist, ergibt diese Annahme, daß $G = G_1 G_1'$ gilt. Wir betrachten den Schnitt $G_1 \cap G_1'$. Er ist Normalteiler sowohl in G_1 als auch in G_1'. Es ergeben sich also die beiden Normalketten

6.3.8 $$G \rhd G_1 \rhd G_1 \cap G_1' \text{ und } G \rhd G_1' \rhd G_1 \cap G_1'.$$

Nach dem Noetherschen Isomorphiesatz sind sie äquivalent:

6.3.9
$$G/G_1 = G_1G_1'/G_1 \simeq G_1'/G_1 \cap G_1', \quad \text{und}$$
$$G/G_1' = G_1G_1'/G_1' \simeq G_1/G_1 \cap G_1'.$$

Daraus folgt auch, daß $G_1 \cap G_1'$ ein maximaler Normalteiler in G_1' ist. Aus der Induktionsannahme folgt aber, daß jede Normalreihe von G_1 eine Länge $\leq r-1$ hat. Jede Normalreihe von $G_1 \cap G_1'$ ist aber nach 6.3.8 von einer Länge $\leq r-2$, dieser Schnitt besitzt also Kompositionsreihen (so wie jede endliche Gruppe mit eigentlichen Normalreihen von beschränkter Länge), und deren Längen sind zudem $\leq r-2$. Daraus schließen wir, mit Hilfe der zweiten der angegebenen eigentlichen Normalketten, daß auch G_1' eine Kompositionsreihe besitzt und deren Länge $\leq r-1$ ist. Zusammen mit der Induktionsannahme ergibt das die Behauptung.

□

6.3.10 Folgerung *Besitzt G Kompositionsreihen, dann gilt:*

- *Je zwei von diesen sind gleichlang.*
- *Jede eigentliche Normalreihe kann zu einer Kompositionsreihe verfeinert werden.*

Wir können damit den Satz von Jordan und Hölder beweisen.

6.3.11 Satz von Jordan und Hölder *Je zwei Kompositionsreihen derselben Gruppe sind äquivalent.*

Beweis: Seien

$$\gamma\colon G = G_0 \rhd G_1 \rhd \ldots \rhd G_r = 1 \text{ und } \gamma'\colon G = G_0 \rhd G_1' \rhd \ldots \rhd G_r' = 1$$

Kompositionsreihen. Wir induzieren nach r. Da der Fall $r = 1$ klar ist (G ist dann ja einfach), können wir gleich $r > 1$ annehmen und die folgenden Fälle unterscheiden:

a. $G_1 = G_1'$, hier folgt die Behauptung direkt aus der Induktionsannahme.

b. Ist $G_1 \neq G_1'$, dann gilt wieder $G = G_1G_1'$. Wir betrachten Reihen, die durch Verfeinerung von $G_1 \rhd G_1 \cap G_1' \rhd 1$ und der entsprechenden Reihe für G_1', zu Kompositionsreihen führen:

$$G_1 \rhd G_1 \cap G_1' \rhd G_3'' \rhd \ldots \rhd G_r'' = 1,$$

$$G_1' \rhd G_1 \cap G_1' \rhd G_3'' \rhd \ldots \rhd G_r'' = 1.$$

Nun folgt, aus der Induktionsannahme, daß

$$G \rhd G_1 \rhd G_1 \cap G_1' \rhd G_3'' \rhd \ldots \rhd G_r'' = 1$$

äquivalent ist zu γ, und daß

$$G \rhd G_1' \rhd G_1 \cap G_1' \rhd G_3'' \rhd \ldots \rhd G_r'' = 1$$

äquivalent ist zu γ'. Das ergibt schon die Behauptung, denn nach 6.3.9 ist

$$G/G_1 \simeq G_1'/G_1 \cap G_1', \text{ und } G/G_1' \simeq G_1/G_1 \cap G_1'.$$

□

6.3.12 Satz *Eine Gruppe $G \neq 1$ besitzt genau dann eine Kompositionsreihe, wenn folgende beiden Bedingungen erfüllt sind:*

- *Jede absteigende Normalkette bricht ab.*
- *Jede aufsteigende Kette von Normalteilern eines Elements G' einer Normalreihe von G bricht ab.*

Beweis: Die Notwendigkeit der beiden angegebenen Bedingungen für die Existenz einer Kompositionsreihe ist klar. Zum Beweis der Umkehrung betrachten wir ein Element $G' \neq 1$ einer Normalreihe von G. Aus der zweiten der angegebenen Bedingungen folgt die Existenz eines maximalen Normalteilers von G', woraus mit der ersten Bedingung die Existenz einer Kompositionsreihe von G' folgt. Setzt man jetzt $G' := G$, dann ergibt sich die Existenz einer Kompositionsreihe auch für die Gruppe G.

□

6.3.13 Folgerung *Eine abelsche Gruppe besitzt genau dann Kompositionsreihen, wenn sie die* aufsteigende–Ketten–Bedingung *und die* absteigende–Ketten–Bedingung *für Untergruppen erfüllt (d. h. daß solche (eigentlichen) Untergruppenketten stets von endlicher Länge sind).*

Wie wir bereits wissen, besitzt $\mathbb{Z}$ nicht diese Eigenschaften, verfügt daher auch nicht über Kompositionsreihen. Eine nicht abbrechende Kette von Untergruppen bilden beispielsweise die von den Potenzen einer Primzahl p erzeugten Ideale:

$$(p) \supset (p^2) \supset (p^3) \supset \ldots$$

6.3.14 Satz *Ist G eine Gruppe mit Kompositionsreihen der Länge r und ist G' ein Normalteiler von G, dann besitzen sowohl G' als auch G/G' Kompositionsreihen, deren Längen sich zu r addieren.*

Beweis: Wir verfeinern $G \rhd G' \rhd 1$ zu einer Kompositionsreihe:

$$G \rhd G_1 \rhd \ldots \rhd G_s = G' \rhd H_1 \rhd \ldots \rhd H_r = 1.$$

Aus ihr ergibt sich die Kompositionsreihe

$$G/G' \rhd G_1/G' \rhd \ldots \rhd G_s/G' = 1.$$

□

Eine wichtige Klasse von Gruppen ist dadurch charakterisiert, daß die Faktoren ihrer Kompositionsreihen von Primzahlordnung sind, wir wollen diese aber anders definieren und später die Äquivalenz der verlangten Eigenschaften nachweisen.

6.3.15 Definition *G sei eine Gruppe.*

Eine Untergruppe $U \leq G$ heißt charakteristische *Untergruppe, wenn für alle Elemente α der Automorphismengruppe $Aut(G)$ von G gilt*

$$\alpha(U) = U.$$

Zu zwei Elementen $g, h \in G$ heißt

$$[g, h] := g^{-1}h^{-1}gh$$

ihr Kommutator*.*

Die von allen Kommutatoren erzeugte Untergruppe

$$G^{(1)} := [G, G] := \langle g^{-1}h^{-1}gh \rangle$$

heißt die Kommutatorgruppe *von G.*

Die höheren Kommutatorgruppen $G^{(i)}$ werden rekursiv so definiert:

$$G^{(i)} := [G^{(i-1)}, G^{(i-1)}].$$

6.3.16 Hilfssatz *Die Kommutatorgruppe ist die kleinste Untergruppe mit abelscher Faktorgruppe in G. Sie ist charakteristische Untergruppe.*

Beweis: Übungsaufgabe (Nr.1, Bl. 1).

6.3.17 Definition *Eine Gruppe heißt* auflösbar*, wenn ihre Kompositionsreihen nur Faktoren von Primzahlordnung besitzen.*

Man kann zeigen, daß eine endliche Gruppe G genau dann auflösbar ist, wenn es ein $k \in \mathbb{N}$ gibt mit $G^{(k)} = 1$.

6.3.18 Satz *Die symmetrische Gruppe S_n ist für $n > 4$ nicht auflösbar, da*

$$S_n \rhd A_n \rhd 1$$

eine Kompositionsreihe ist (d.h. A_n ist für $n \neq 4$ einfach).

Kapitel 7

Aus der Ringtheorie

7.1 Ringe und Ideale

Erinnern wir uns zunächst an die Definition von Ringen, es sind Mengen R mit zwei Verkn üpfungen $+$ und $\cdot$, so daß $(R, +)$ eine abelsche Gruppe, $(R, \cdot)$ eine Halbgruppe ist, und die beiden Distributivgesetze gelten:

$$r(s+t) = rs + rt,\ (r+s)t = rt + st.$$

(Diese Distributivgesetze formuliert man eigentlich besser gleich strukturtheoretisch: Links- und Rechtsmultiplikation mit einem Ringelement ist ein Endomorphismus von $(R, +)$. Das liefert nämlich sofort, daß $0r = r0 = 0$ gilt!)
Ringe, die wir bereits kennengelernt haben sind die Zahlbereiche $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, sowie die Matrixringe über diesen, also z. B. der Ring $\mathbb{Z}^{n\times n}$, der aus den n-reihigen Matrizen über $\mathbb{Z}$ besteht, mit der Matrixaddition und der Matrizenmultiplikation als Verknüpfungen. Allgemeiner gilt: man kann mit einem Ring R und einer nicht leeren Menge M die Menge

$$R^M := \{f \mid f \colon M \to R\}$$

aller Abbildungen von M nach R bilden, die mit punktweiser Addition und Multiplikation,

$$(f+g)(r) := f(r) + g(r),\ (f \cdot g)(r) := f(r) \cdot g(r)$$

offenbar einen Ring bildet.
Daneben bieten die *Endomorphismenringe* abelscher Gruppen eine Fülle an Beispielen: Ist $(A, *)$ abelsche Gruppe, $End(A)$ die Menge ihrer Endomorphismen, dann ist $(End(A), +, \cdot)$ ein Ring mit

$$(f+g)(a) := f(a) * g(a),\ (f \cdot g)(a) := (f \circ g)(a) = f(g(a)).$$

Wegen der Distributivgesetze ist offensichtlich jeder Ring mit Einselement isomorph zu einem Ring von Endomorphismen von sich selbst, denn man braucht ja nur $r \in R$ auf die Linksmultiplikation mit r abzubilden.

Ist R ein Ring mit Einselement — und die meisten in der Vorlesung vorkommenden Ringe besitzen ein solches Element —, dann heißt $r \in R$ *Einheit*, wenn es ein $s \in R$ gibt mit $rs = sr = 1$. Die Menge aller Einheiten wird so bezeichnet:

$$E(R) := \{r \in R \mid r \text{ Einheit}\},$$

sie ist eine (multiplikative) Gruppe, die *Einheitengruppe* von R. Einfache Beispiele sind

$$E(\mathbb{Z}) = \{1, -1\},\ E(\mathbb{Z}^{2\times 2}) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \pm 1\}.$$

Die tragende Rolle, die in der Gruppentheorie die Normalteiler spielen, übernehmen in der Ringtheorie die *zulässigen* Normalteiler, also die zulässigen Untergruppen von $(R, +)$, denn in abelschen Gruppen sind ja alle Untergruppen normal. Diese zulässigen Untergruppen heißen Ideale:

7.1.1 Definition *Sei R ein Ring. Eine nicht-leere Teilmenge L von R heißt* Linksideal*, wenn für alle $l, l' \in L$ und alle $r \in R$ gilt:*

$$l - l' \in L,\ rl \in L.$$

Analog wird der Begriff des Rechtsideals *definiert. Ideale, die sowohl Rechts- als auch Linksideale sind, heißen* (zweiseitige) Ideale, *und wir schreiben hierfür auch*

$$I \trianglelefteq R.$$

Ideale sind somit Untermoduln.

Schnitte von Idealen sind ebenfalls Ideale, so daß wir von dem Ideal sprechen können, das von einer Teilmenge $M \subseteq R$ *erzeugt* wird; wir bezeichnen es mit (M). Besteht M aus einem einzigen Element r, dann schreiben wir kurz (r) statt $(\{r\})$. Solche Ideale (r) heißen *Hauptideale*, Ringe, deren Ideale sämtlich Hauptideale sind, dementsprechend *Hauptidealringe*.

7.1.2 Beispiele Ist R ein Ring, dann enthält

- R die *trivialen* Ideale $0 := \{0\}$ und R.
- Für das von $r \in R$ erzeugte Hauptideal gilt:

  $$(r) = \{x = \sum_{i=1}^{n} s_i r t_i + rt + ur + zr \mid s_i, t_i, u, t \in R, z \in \mathbb{Z}, n \in \mathbb{N}\}.$$

  Enthält R ein Einselement, dann reduziert sich dies zu

  $$(r) = \langle RrR \rangle = \{\sum_{i=1}^{n} s_i r t_i \mid n \in \mathbb{N}, s_i, t_i \in R\}.$$

  Ist R zusätzlich kommutativ, dann gilt entsprechend

  $$(r) = \langle Rr \rangle = \{sr \mid s \in R\}.$$

- Sind $I, I' \trianglelefteq R$, dann ist deren Summe
$$I + I' = \{r \mid r = i + i', i \in I, i' \in I'\}$$
ein Ideal, das *Summenideal.* Analog definieren wir
$$(II') := \{r \mid r = \sum_{\nu=1}^{n} i_\nu i'_\nu, i_\nu \in I, i'_\nu \in I', n \in \mathbb{N}\},$$
das *Produktideal.*

- Summen- und Produktideal genügen den folgenden Ungleichungen:
$$(II') \subseteq I \cap I' \subseteq I + I'.$$

- $\mathbb{Z}$ ist ein Hauptidealring, denn mit Hilfe des Euklidischen Algorithmus, also der Definition mit Rest, sieht man leicht, daß die kleinste positive natürliche Zahl n in einem vom Nullideal verschiedenen Ideal $I \trianglelefteq \mathbb{Z}$ jedes $i \in I$ teilt:
$$i = qn + r, 0 \leq r < n \Rightarrow r = i - qn \in I \Rightarrow r = 0.$$

- Im Hauptidealring $\mathbb{Z}$ gilt
$$(z) \subseteq (z') \iff z' \mid z,$$
und
$$((z)(z')) = (zz') \subseteq (kgV(z, z')) = (z) \cap (z') \subseteq (ggT(z, z')) = (z) + (z').$$

$\diamond$

Die Ideale sind zulässige Normalteiler, für Ringhomomorphismen $f: R \to R'$ gilt demnach:
$$f(r_1 + r_2) = f(r_1) + f(r_2) \quad \text{und} \quad f(r_1 \cdot r_2) = f(r_1) \cdot f(r_2).$$
Ist R ein Ring mit 1, so gilt zudem: $f(1_R) = 1_{R'}$.

7.1.3 Satz *Für Homomorphismen $f: R \to R'$ zwischen Ringen gilt:*

- *Bilder von Ringhomomorphismen sind Ringe, Kerne sind genau die Ideale.*

- *Ist I ein Ideal in R, dann erhält man aus der Faktorgruppe $(R/I, +)$ einen Ring vermöge*
$$[r]_I[s]_I := [rs]_I, \text{ wobei } [r]_I := r + I.$$
Die I entsprechende kanonische Abbildung $\nu_I: r \mapsto [r]_I$ auf die Faktorgruppe R/I ist ein Ringepimorphismus mit I als Kern.

7.1.4 Beispiel Als *Restklassenring modulo n* bezeichnen wir den Faktorring

$$(\mathbb{Z}_n, +, \cdot) := (\mathbb{Z}/(n), +, \cdot) = (\mathbb{Z}/n \cdot \mathbb{Z}, +, \cdot).$$

Ein Ringhomomorphismus zwischen solchen Restklassenringen ist beispielsweise

$$f\colon \mathbb{Z}_{12} \to \mathbb{Z}_3\colon z + (12) \mapsto z + (3).$$

Der Kern ist $(3 + (12))$; da f zusätzlich surjektiv ist, gilt also

$$\mathbb{Z}_{12}/(3 + (12)) \simeq \mathbb{Z}_3.$$

Schreiben wir kurz $n\mathbb{Z}_m$ für das Ideal

$$(n + (m)) = \{0 + (m), n + (m), 2n + (m), \ldots\} \trianglelefteq \mathbb{Z}_m,$$

dann gilt also

$$\mathbb{Z}_{12}/3\mathbb{Z}_{12} \simeq \mathbb{Z}_3.$$

$\diamond$

Eine ganz einfache Anwendung der Abbildung von Ringen auf Restklassenringe ermöglicht eine erhebliche Vereinfachung und Parallelisierung beim Rechnen mit großen ganzen Zahlen. Betrachten wir zunächst einen einfachen Fall, den wir später verallgemeinern und auf $\mathbb{Z}$ anwenden werden:

7.1.5 Hilfssatz *Sei $R \neq 0$ ein Ring mit Einselement und Idealen I, I' mit $R = I + I'$. Dann ist die Abbildung*

$$f\colon R \to R/I \times R/I'\colon r \mapsto ([r]_I, [r]_{I'})$$

ein Ringepimorphismus mit $\mathrm{Kern}(f) = I \cap I'$.

Beweis: Die Homomorphieeigenschaft und die Behauptung über die Form des Kerns von f sind trivial. Zum Beweis der Surjektivität geben wir ein Urbild von $([r]_I, [r']_{I'})$ an: Ist $i + i' = 1$, dann ist

$$f(r'i + ri') = ([ri']_I, [r'i]_{I'}) = (r(1-i) + I, r'(1-i') + I') = ([r]_I, [r']_{I'}).$$

$\square$

Nehmen wir beispielsweise $R := \mathbb{Z}$ und $I := (n)$, $I' := (n')$, mit teilerfremden natürlichen Erzeugenden n, n', dann gilt also

$$\mathbb{Z}_{n \cdot n'} \simeq \mathbb{Z}_n \times \mathbb{Z}_{n'}.$$

Das bedeutet: Anstatt in $\mathbb{Z}_{nn'}$ rechnen zu müssen, können wir demnach auch in dem kartesischen Produkt der beiden Restklassenringe rechnen. Die folgenden Klassen von Idealen werden sich als besonders interessant herausstellen:

7.1.6 Definition *Ist $I \triangleleft R$, dann heißt*

- *I ein* maximales Ideal*, wenn I als Ideal maximal in R ist:*

$$\forall\, I' \trianglelefteq R\colon\ [I \subset I' \Rightarrow I' = R]\,.$$

- *I* Primideal*, wenn $R\backslash I$ nicht-leer und multiplikativ abgeschlossen ist, d.h.:*

$$(R\backslash I, \cdot) \text{ ist Halbgruppe.}$$

Für solche Ideale gilt der wichtige

7.1.7 Satz *Ist R ein kommutativer Ring mit 1, $I \trianglelefteq R$, dann ist I*

- *genau dann Primideal, wenn R/I Integritätsbereich ist, das heißt ein vom Nullring verschiedener kommutativer Ring mit Einselement und ohne Nullteiler.*

- *genau dann maximales Ideal, wenn R/I ein Körper ist.*

Beweis: Es gelten die folgenden Ketten von Äquivalenzen:

$$\begin{array}{rcl} I \text{ prim} & \Longleftrightarrow & (R/I)^* \text{ komm. Halbgruppe} \\ & \Longleftrightarrow & R/I \text{ Integritätsbereich.} \end{array}$$

und

$$\begin{array}{rcl} I \text{ maximal} & \Longleftrightarrow & [R\backslash I \neq \emptyset \wedge \forall x \in R\backslash I : I + (x) = R] \\ & \Longleftrightarrow & [R\backslash I \neq \emptyset \wedge \forall x \in R\backslash I\ \exists i \in I\colon i + rx = 1_R] \\ & \Longleftrightarrow & ((R/I)^*, \cdot) \text{ abelsche Gruppe.} \end{array}$$

□

7.1.8 Folgerung *In kommutativen Ringen mit Einselement sind maximale Ideale auch Primideale.*

Betrachten wir jetzt die Einheitengruppen der Restklassenringe $\mathbb{Z}_n$. Da die *Eulerfunktion* φ wie folgt definiert ist:

$$\varphi(n) := |\{0 \leq i < n \mid ggT(i,n) = 1\}|,$$

haben wir (vgl. 6.2.21):

7.1.9 Folgerung *Für natürliche Zahlen n gilt:*

- *Die* primen Restklassengruppen*, die Einheitengruppen $E(\mathbb{Z}_n)$, haben die Ordnung*

$$|E(\mathbb{Z}_n)| = \varphi(n).$$

- *$\mathbb{Z}_n$ ist genau dann ein Körper, wenn $\varphi(n) = n - 1$ gilt, oder, äquivalent dazu, wenn n Primzahl ist.*

- *(n) ist genau dann Primideal in $\mathbb{Z}$, wenn n Primzahl ist, und dies wiederum ist äquivalent dazu, daß (n) ein maximales Ideal ist.*

Hier fallen also die Begriffe maximales Ideal und Primideal zusammen. In vielen Ringen ist das anders, dort klaffen diese Begriffe auseinander, was wiederum heißt, daß die maximalen Ideale und auch die Primideale als *verschiedene* Verallgemeinerungen des Primzahlbegriffs angesehen werden können.
Für die nächsten Überlegungen ziehen wir ein Axiom der axiomatischen Mengenlehre heran, das schon zum Beweis der Existenz von Basen von Vektorräumen benutzt wurde:

7.1.10 Das Lemma von Zorn *Ist $(M, \leq)$ eine nicht leere Menge mit einer Halbordnung, so daß es zu jeder Kette $(K, \leq)$ von Teilmengen in M eine obere Schranke in M gibt, dann enthält M maximale Elemente.*

Mit dessen Hilfe folgt sofort:

7.1.11 Satz *Jeder Ring $0 \neq R$ mit Eins enthält maximale Ideale.*

Beweis: Man wendet das Lemma von Zorn auf die folgende Halbordnung an:

$$\mathcal{I} := \{I \mid I \lhd R,\ I \neq R\}.$$

$\square$

Ideale in Ringen sind, wie bereits erwähnt, zulässige Untergruppen, also ergibt sich direkt aus dem Homomorphiesatz für Gruppen

7.1.12 Der Homomorphiesatz für Ringe *Ist $f : R \to R'$ ein Homormorphismus zwischen Ringen, und ist I ein Ideal in R, dann gilt:*

- *Liegt I im Kern von f, dann gibt es genau ein $g \colon R/I \to R'$, das folgendes Diagramm kommutativ ergänzt:*

$$\begin{tikzcd} R \arrow[r, "\nu_I"] \arrow[dr, "f"'] & R/I \arrow[d, "g"] \\ & R' \end{tikzcd}$$

  g ist ein Ringhomomorphismus.

- *Ist I gleich dem Kern von f, dann ist g injektiv, also insbesondere*

$$R/I \simeq Bild(f).$$

- *Die Abbildung f stiftet Bijektionen zwischen den Mengen der Links-, Rechts- und zweiseitigen Ideale aus R, die den Kern von f enthalten und den Mengen der Links-, Rechts- und zweiseitigen Ideale aus $Bi(f)$.*

Weiter folgt mit dem Satz über induzierte Abbildungen bzw. Homomorphismen:

7.1.13 Der Satz über induzierte Ringhomomorphismen
Sind R, R', R'', R''' Ringe mit Epimorphismen $f: R \to R'$, $g: R \to R''$, $h: R' \to R'''$, dann gilt

- *Ist $f[Kern(g)] \subseteq Kern(h)$, dann gibt es genau ein $k: R'' \to R'''$ mit kommutativem Diagramm*

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, "g"'] & R' \arrow[d, "h"] \\ R'' \arrow[r, "k"] & R''' \end{tikzcd}$$

  Diese eindeutig bestimmte Abbildung k ist ein Ringepimorphismus.

- *Ist $Kern(g) = f^{-1}[Kern(h)]$, dann ist dieses k sogar ein Ringisomorphismus.*

□

7.2 Ringe von Brüchen, Lokalisierung, Quotientenkörper

Wir verallgemeinern zunächst die Bildung von $\mathbb{Q}$ aus $\mathbb{Z}$ als Menge der Äquivalenzklassen von Brüchen ganzer Zahlen. Dazu bemerken wir, daß folgendes leicht nachgerechnet werden kann:

7.2.1 Satz *Ist* $(0 \neq)R$ *ein kommutativer Ring mit Eins und* $N \subseteq R$ *ein Teilmonoid von* $(R^*, \cdot)$ *(d. h.* N *ist multiplikativ abgeschlossen, also eine Unterhalbgruppe, und* N *enthält das Einselement von* R*), dann gilt:*

- *Folgende Relation auf* $R \times N$ *ist eine Äquivalenzrelation:*

$$(r, n) \sim (r', n') :\iff \exists\, n'' \in N : n''(rn' - r'n) = 0.$$

- *Die Menge* $[R \times N]_\sim$ *der entsprechenden Äquivalenzklassen* $[r, n]_\sim := [(r, n)]_\sim$*, ist zusammen mit den Verknüpfungen*

$$[r, n]_\sim + [r', n']_\sim := [rn' + r'n, nn']_\sim$$

und

$$[r, n]_\sim \cdot [r', n']_\sim := [rr', nn']_\sim$$

ein kommutativer Ring mit dem Einselement $[n, n]_\sim$ *und dem Nullelement* $[0, n]_\sim$*, sowie mit den additiven Inversen*

$$-[r, n]_\sim = [-r, n]_\sim.$$

Die Teilmenge N *heißt dabei die* Nennermenge.

7.2.2 Beispiel Ist $R \neq 0$ ein kommutativer Ring mit Einselement, dann sind R^*, $E(R)$ und $R \backslash I$, für jedes Primideal I geeignete Nennermengen. Beispiele sind also $\mathbb{Z}^*$, $\{1, -1\}$, und $\mathbb{Z}\backslash(p)$, für Primzahlen p.

◇

7.2.3 Hilfssatz *N sei eine Nennermenge gemäß 7.2.1. Es gilt:*

- *Die Abbildung*

$$f : R \to [R \times N]_\sim : r \mapsto [rn, n]_\sim$$

ist ein Ringhomomorphismus mit $f(N) \subseteq E([R \times N]_\sim)$.

- *Ist* R *ein Integritätsbereich, dann ist* f *sogar injektiv, und wir haben eine Einbettung*

$$R \hookrightarrow [R \times N]_\sim : r \mapsto [rn, n]_\sim$$

von R*. Wir können in diesem Fall das Element* $[rn, n]_\sim$ *durch* r *ersetzen. Die oben angegebene Äquivalenzrelation vereinfacht sich zu*

$$(r, n) \sim (r', n') :\iff rn' - r'n = 0,$$

das Element $[r,n]_\sim$ *ist demnach durch den* Bruch $\frac{r}{n}$ *ersetzbar.*

Der so aus $[R\times N]_\sim$ *entstandene Ring heißt* Ring der Brüche *von* R *mit* Nennermenge N. *Wir bezeichnen ihn mit*

$$B(R,N).$$

Das bekannteste Beispiel ist natürlich

$$\mathbb{Q} := B(\mathbb{Z},\mathbb{Z}^*),$$

der *Körper der rationalen Zahlen.* Die herausragende Eigenschaft dieses Ringes der Brüche ist seine Universalität:

7.2.4 Satz *Ist* R *ein Integritätsbereich,* N *eine Nennermenge in* R*, dann ist die Einbettung*

$$R \hookrightarrow B(R,N)$$

universell bzgl. der Klasse $\mathcal{F}$ *der (Ring mit 1)-Homomorphismen von* R *in Ringe* R' *mit Einselement und der Eigenschaft* $f(N)\subseteq E(R')$*, sowie der Klasse* $\mathcal{L}$ *der Homomorphismen zwischen Ringen mit Einselement.*

7.2.5 Definition *Ein Ring* R *heißt* lokaler Ring*, wenn* R *genau ein maximales Ideal besitzt.*

7.2.6 Satz *Ist* $R\neq 0$ *ein Integritätsbereich,* P *ein Primideal in* R*, dann ist der Ring der Brüche*

$$B(R,R\backslash P)$$

ein lokaler Ring.

Beweis: Wir zeigen daß das Ideal (nachprüfen!)

$$I := \{\frac{p}{n} \mid p\in P\}R$$

genau aus den Nichteinheiten von $B(R,R\backslash P)$ besteht:

a. $\frac{a}{n}\notin I \Rightarrow a\notin P \Rightarrow \frac{n}{a}\in B(R,R\backslash P) \Rightarrow \frac{a}{n}$ Einheit.

b. $I\ni \frac{p}{n}$ Einheit, etwa $\frac{p}{n}\cdot\frac{a}{n'} = 1 \Rightarrow \underbrace{pa}_{\in P} = \underbrace{nn'}_{\in R\backslash P}$, ein Widerspruch.

Als Menge aller Nichteinheiten ist dieses Ideal maximal. Umgekehrt besteht jedes maximale Ideal aus Nichteinheiten, liegt also in I und ist deshalb *gleich* I.

□

7.2.7 Beispiel Der Ring der Brüche

$$B(\mathbb{Z},\mathbb{Z}\backslash(p))$$

ist lokal, falls p eine Primzahl ist.

7.2.8 Satz *Ist R Integritätsbereich, dann ist der Ring der Brüche*

$$B(R, R^*)$$

ein Körper, der sogenannte Quotientenkörper *von R.*

Beweis: Das Nullideal ist Primideal, also ist, nach 7.2.6, $B(R, R^*)$ lokal und dessen Nullideal ist maximal, enthält also alle Nichteinheiten. Demnach ist $B(R, R^*)$ Schiefkörper und wegen der Kommutativität auch ein Körper.

$\square$

Das klassische Beispiel hierfür ist natürlich

$$\mathbb{Q} := B(\mathbb{Z}, \mathbb{Z}^*),$$

der *Körper der rationalen Zahlen.*

7.3 Euklidische Bereiche, Hauptideal- und Gaußbereiche

Es geht jetzt um eine Verallgemeinerung der Teilbarkeitseigenschaften von $\mathbb{Z}$ auf andere Ringe und von Begriffen wie den größten gemeinsamen Teiler, Primzahlen usw.

7.3.1 Definition *Ist R Integritätsbereich, dann heißt $\delta: R \to \mathbb{Z}$ genau dann* euklidische Norm, *wenn folgendes gilt:*

- $\forall\ r \in R^*: \delta(r) > 0.$
- $\delta(0) = 0.$
- $\delta(rs) = \delta(r)\delta(s).$
- $\forall\ x \in R, y \in R^*\ \exists\ q, r \in R: x = qy + r,\ 0 \leq \delta(r) < \delta(y).$

Der Ring R (genauer: das Paar (R, δ)) heißt dann euklidischer Bereich.

Bekannte Beispiele hierfür sind der Ring der ganzen Zahlen mit der Betragsfunktion und der Polynomring $\mathbb{K}[x]$, $\mathbb{K}$ ein Körper, mit

$$\delta: \mathbb{K}[x] \to \mathbb{Z}: f \mapsto \begin{cases} 0 & \text{falls } f = 0 \\ 2^{Grad(f)} & \text{sonst,} \end{cases}$$

denn es gilt:

7.3.2 Der Divisionsalgorithmus für $\mathbb{K}[x]$: *Ist $\mathbb{K}$ ein Körper, $f \in \mathbb{K}[x] \ni g \neq 0$, dann gibt es genau ein Paar (q, r) von Polynomen mit der Eigenschaft*

$$f = q \cdot g + r,\ \text{und } [r = 0 \vee Grad(r) < Grad(g)].$$

Beweis: Der Fall $f = 0$ ist trivial, denn $q := r := 0$ erfüllen die Identität $f = qg + r$, und die Annahme $r \neq 0$ ergäbe

$$Grad(r) = Grad(gq) = Grad(g) + Grad(q) \geq Grad(g),$$

also einen Widerspruch.
Sei deshalb $f \neq 0$. Wir beweisen zunächst die Existenz eines Paares (q, r) per Induktion nach $Grad(f)$. Sei

$$f = a_0 + \ldots + a_n x^n,\ a_n \neq 0,\ g = b_0 + \ldots + b_m x^m,\ b_m \neq 0.$$

Wir unterscheiden drei Fälle:
Bei $n = 0 < m$ können wir $q := 0,\ r := f$ setzen.
Falls $n = 0 = m$ wählen wir $q := g^{-1} f = b_0^{-1} f$ und $r := 0$.
Ist dagegen $n > 0$, dann betrachten wir die beiden Fälle

- $n < m$: Wir nehmen $q := 0$ und $r := f$.

- Bei $n \geq m$ betrachten wir $f_1 := f - a_n b_m^{-1} x^{n-m} g$ und spalten erneut in zwei Teilfälle auf:
  - Bei $f_1 = 0$ setzen wir $q := a_n b_m^{-1} x^{n-m},\ r := 0$.
  - Andernfalls ist $Grad(f_1) < Grad(f)$, und die Induktionsannahme liefert die Existenz eines Paare (q_1, r_1) mit

$$f = q_1 g + a_n b_m^{-1} x^{n-m} g + r_1,\ [r_1 = 0 \vee Grad(r_1) < Grad(g)].$$

Zum abschließenden Beweis der Eindeutigkeit betrachten wir zwei Paare (q, r) und (q_1, r_1) mit $f = gq + r = gq_1 + r_1$ und den Ungleichungen für die Grade. Da hierfür gilt $(q - q_1)g = r_1 - r$, folgt aus der Annahme $r \neq r_1$, daß $q \neq q_1$, und damit

$$0 \leq Grad(r - r_1) < Grad(g)$$

gelten müßte, im Widerspruch zu

$$Grad(r - r_1) = Grad(q - q_1)g = Grad(q - q_1) + Grad(g) \geq Grad(g).$$

□

Ganz entsprechend wie für $\mathbb{Z}$ wird jetzt gezeigt:

7.3.3 Satz *Jeder euklidische Bereich ist ein Hauptidealbereich.*

Beweis: Sei $I \trianglelefteq R$ und (R, δ) ein euklidischer Bereich. Das Nullideal ist trivialerweise ein Hauptideal, sei also $I \neq 0$ und $i_0 \in I$ ein Element $\neq 0$, dessen Norm $\delta(i_0)$ minimal ist unter den Normen der von Null verschiedenen Elemente in I (solche Elemente gibt es, denn δ ist eine Abbildung nach $\mathbb{Z}$). Jedes von Null verschiedene $i \in I$ ist dann durch i_0 teilbar:

$$i = q \cdot i_0 + r,\ 0 \leq \delta(r) < \delta(i_0)$$

Dies ergibt wegen der Minimalität der Norm von i_0, daß $\delta(r) = 0$ gilt, und damit $I = (i_0)$.

□

Ein Polynom $f \in \mathbb{K}[x]$ heißt genau dann *normiert*, wenn der Leitkoeffizient gleich 1 ist. Es gilt also: Der Polynomring $\mathbb{K}[x]$ ist Hauptidealbereich, und jedes Ideal $I \neq 0$ in $\mathbb{K}[x]$ wird von einem eindeutig bestimmten normierten Polynom erzeugt.

7.3.4 Definition *Sei R ein Integritätsbereich mit Elementen $r, s \in R$ und einer nicht-leeren Teilmenge T. Wir sagen dann*

- *r* teilt *s, wenn es $t \in R$ gibt mit $rt = s$, kurz:*

$$r \mid s.$$

- *r ist* assoziiert *zu s, wenn es eine Einheit $t \in E(R)$ gibt mit $rt = s$, und das wird wie folgt abgekürzt:*

$$r \sim s.$$

- *r ist ein* echter *Teiler von s, wenn folgendes gilt:*

$$r \mid s \;\wedge\; r \notin E(R) \;\wedge\; r \not\sim s.$$

- *r heißt* unzerlegbar, *wenn r weder 0 noch Einheit ist und keine echten Teiler besitzt.*

- *r heißt* prim, *wenn r weder 0 noch Einheit ist und*

$$r \mid st \Rightarrow [\, r \mid s \;\vee\; r \mid t \,]$$

richtig ist, d.h. r teilt mit einem Produkt auch mindestens einen Faktor.

- *r ist* ein größter gemeinsamer Teiler *von T, wenn folgende beiden Bedingungen erfüllt sind:*
  - $\forall\, t \in T\colon r \mid t.$
  - $\forall\, s \in R\colon [\, \forall\, t \in T\colon\; s \mid t \,] \Rightarrow s \mid r.$

  Dieses wird auch so geschrieben: $r \in ggT(T)$.

- *r ist* ein kleinstes gemeinsames Vielfaches *von T, wenn es die folgenden Eigenschaften hat:*
  - $\forall\, t \in T\colon\; t \mid r.$
  - $\forall\, s \in R\colon\; [\, \forall\, t \in T\colon\; t \mid s \,] \Rightarrow r \mid s,$

  kurz: $r \in kgV(T)$.

- *Die Elemente von T heißen genau dann* teilerfremd, *wenn*

$$ggT(T) = E(R).$$

Teilbarkeitseigenschaften lassen sich sowohl mit Elementen als auch mit den davon erzeugten Idealen formulieren. Wie man leicht nachprüft gilt

7.3.5 Satz *In Integritätsbereichen R gilt für Elemente r, s und nicht-leere Teilmengen T:*

- $r \mid s \iff (s) \subseteq (r),$
- $r \in E(R) \iff (r) = R,$
- $r \sim s \iff [\, r \mid s \;\wedge\; s \mid r \,] \iff (r) = (s),$
- *r ist echter Teiler von s* $\iff (s) \subset (r) \subset R,$
- *r ist prim* $\iff [\, (r) \neq 0 \;\wedge\; (r)$ *ist Primideal* $]$,
- $r \in ggT(T) \Longrightarrow ggT(T) = r \cdot E(R),$
- $r \in kgV(T) \Longrightarrow\in kgV(T) = r \cdot E(R),$

7.3.6 Hilfssatz *In Integritätsbereichen sind Primelemente stets unzerlegbar.*

Beweis: Sei p ein Primelement mit einem Teiler r, etwa $p = rs$. Wir haben zu zeigen, daß dies kein echter Teiler sein kann. Da p Primelement ist, können wir so schließen:

$$p \mid rs \Rightarrow [[p \mid r \vee p \mid s] \wedge [r \mid p \wedge s \mid p]] \Rightarrow p \sim r \vee p \sim s.$$

Falls $p \sim r$ sind wir fertig, andernfalls gilt $s = pt$, mit $t \in E(R)$, also $p = prt$, und damit auch $1 = rt$, also $r \in E(R)$.

□

7.3.7 Satz *Für Elemente $r \neq 0$ in Hauptidealbereichen R sind die folgenden Eigenschaften äquivalent:*

- *r ist unzerlegbar,*
- *(r) ist maximales Ideal,*
- *(r) ist Primideal,*
- *r ist Primelement.*

Beweis: Ist r unzerlegbar, dann ist (r) maximal, denn andernfalls würden echte Teiler existieren, da R Hauptidealbereich ist. (r) ist dann aber auch Primideal, wie wir wissen, r also Primelement, und damit unzerlegbar, nach 7.3.6.

□

7.3.8 Beispiele

- Ist $\mathbb{K}$ ein Körper, dann gilt in $\mathbb{K}[x, y] := \mathbb{K}[x][y]$:
  - x ist unzerlegbar,
  - $(x) \subset (x, y) \subset \mathbb{K}[x, y]$, also ist $\mathbb{K}[x, y]$, nach 7.3.7, kein Hauptidealbereich.
- $R := \{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\} \subset \mathbb{C}$ ist Integritätsbereich, 3 ist darin unzerlegbar, wegen $3 \cdot 3 = (2 + \sqrt{-5})(2 - \sqrt{-5})$ aber nicht prim. R ist also *kein* Hauptidealbereich.

7.3.9 Satz *Jede nicht-leere Teilmenge T eines Hauptidealbereichs besitzt größte gemeinsame Teiler und kleinste gemeinsame Vielfache, nämlich alle Erzeugenden von (T) bzw. von $\cap_{t \in T}(t)$.*

Größte gemeinsame Teiler endlicher Mengen kann man sukzessive mit Hilfe der folgenden Methode ermitteln:

7.3.10 Der Euklidische Algorithmus *Ist (R,δ) ein euklidischer Bereich, dann kann man ein Element aus der Menge $ggT(x_0,x_1)$ der größten gemeinsamen Teiler zweier Elemente $x_0,x_1 \in R$ durch sukzessives Dividieren berechnen (der Algorithmus terminiert, weil δ eine Abbildung in $\mathbb{Z}_{\geq 0}$ ist und die Werte $\delta(x_i)$ der Norm der Reste streng monoton fallen):*

$$\begin{aligned} x_0 &= q_0x_1+x_2 \\ x_1 &= q_1x_2+x_3 \\ &\vdots \\ x_{n-1} &= q_{n-1}x_n+x_{n+1} \\ x_n &= q_nx_{n+1} \end{aligned}$$

Liest man diese Gleichungen von unten nach oben, so ergibt sich, daß x_{n+1} die Elemente X_0,x_1 teilt. Liest man von oben nach unten, so erhält man, daß jeder Teiler von x_0 und x_1 auch in x_{n+1} aufgeht.

Für teilerfremde Elemente gilt insbesondere

7.3.11 Hilfssatz *Ist R ein Hauptidealbereich, dann sind die folgenden Aussagen über Elemente $r_0,\ldots,r_{n-1}$ von R äquivalent:*

- *Die r_i sind teilerfremd,*
- *$\sum Rr_i = R$,*
- *Es gibt $a_i \in R$ mit $\sum a_ir_i = 1$.*

Beweis: Sind die r_i teilerfremd, dann ist das Ideal $\sum Rr_i$ gleich R, denn es wird von jedem der größten gemeinsamen Teiler erzeugt, und diese sind ja Einheiten. Es gibt deshalb auch a_i mit $\sum a_ir_i = 1$. Letzteres wiederum impliziert, daß die r_i teilerfremd sind.

□

7.3.12 Definition *Ist R ein Integritätsbereich mit dem Element r und Nichteinheiten $r_0,\ldots,r_{m-1}$ sowie Nichteinheiten $s_0,\ldots,s_{n-1}$, dann*

- *heißt $\prod_{i\in m} r_i$* äquivalent *zu $\prod_{i\in n} s_i$ (kurz: $\prod_i r_i \approx \prod_i s_i$), wenn gilt:*

$$m=n \ \wedge\ \exists\ \pi\in S_n\colon\ r_i \sim s_{\pi(i)}.$$

- *Weiter heißt r* streng zerlegbar, *wenn gilt:*

$$0\neq r\notin E(R)\wedge\ r \text{ besitzt echte Teiler.}$$

- *Der Integritätsbereich R heißt* Gaußbereich, *wenn jedes streng zerlegbare Element bis auf Äquivalenz eindeutig als Produkt unzerlegbarer Elemente geschrieben werden kann.*

7.3.13 Beispiel Alle Körper sind Gaußbereiche, denn es gibt dort keine streng zerlegbaren Elemente.

Unser Ziel ist der Beweis der Tatsache, daß alle Hauptidealbereiche Gaußbereiche sind.

7.3.14 Definition *Sei $(M, \leq)$ eine angeordnete Menge.*

- $(M, \leq)$ *genügt der* Maximal- *bzw. der* Minimalbedingung, *wenn jede nichtleere Teilmenge maximale bzw. minimale Elemente enthält.*
- $(M, \leq)$ *genügt der* Aufsteigende-Ketten-Bedingung *bzw. der* Absteigende-Ketten-Bedingung, *wenn alle aufsteigenden bzw. alle absteigenden Ketten von Elementen stationär werden.*

7.3.15 Satz *Die Maximalbedingung ist äquivalent zur Aufsteigende-Ketten-Bedingung, die Minimalbedingung ist äquivalent zur Absteigende-Ketten-Bedingung.*

Beweis: Zu einer aufsteigenden Kette

$$m_0 \leq m_1 \leq \ldots$$

betrachten wir $\cup_i\{m_i\}$. Gilt die Maximalbedingung, dann gibt es darin maximale Elemente, die Kette muß also stationär werden. Umgekehrt gibt es, wenn keine maximalen Elemente existieren, offensichtlich aufsteigende Ketten, die nicht stationär werden. Entsprechend argumentiert man bei absteigenden Ketten und der Minimalbedingung.

□

7.3.16 Satz *Ist R ein Hauptidealbereich, dann gilt die Aufsteigende-Ketten-Bedingung für die Menge $(\mathcal{I}, \subseteq)$ seiner Ideale.*

Beweis: Zeige: Ist $I_0 \leq I_1 \leq \ldots$ eine aufsteigende Kette von Idealen, dann ist die Vereinigung $I := \cup I_\nu$ ihrer Elemente ein Ideal. R ist Hauptidealbereich, es gibt also $i \in R$ mit $I = (i)$. Dieses i liegt in einem der I_ν, nach diesem Element bleibt die Kette demnach konstant.

□

7.3.17 Satz *Jeder Hauptidealbereich ist Gaußbereich.*

Beweis: Wir beweisen zunächst die *Existenz* der Zerlegung streng zerlegbarer Elemente in unzerlegbare. Sei $r \in R$ streng zerlegbar, R ein Hauptidealbereich. Es gibt also echte Teiler von r, etwa $r = r_1 r_2$. Wir haben dann die beiden Ketten

$$(r) \subset (r_1) \subset R,\ (r) \subset (r_2) \subset R.$$

Der Teiler r_1 von r ist nun entweder unzerlegbar oder streng zerlegbar, in welchem Falle $r_1 = r_{11} r_{12}$ gilt, mit echten Teilern, also insgesamt

$$(r) \subset (r_1) \subset (r_{11}) \subset R,\ (r) \subset (r_1) \subset (r_{12}) \subset R$$

usw., und entsprechend mit r_2.
Nun ist aber R ein Hauptidealbereich, so daß (nach 7.3.16) all diese Ketten stationär werden. Die jeweils letzten Kettenglieder sind von unzerlegbaren Faktoren von r erzeugt, und das Produkt dieser endlich vielen (!) Faktoren ergibt r.
Um noch die Äquivalenz dieser Faktorisierungen nachzuweisen, betrachten wir zwei Zerlegungen

$$r = p_0 \cdots p_s = q_0 \cdots q_t.$$

Da die unzerlegbaren Faktoren auch prim sind, gibt es j mit $p_0 \mid q_j$. Nun ist aber R kommutativ, demnach können wir ohne Einschränkung annehmen, daß $j = 0$ gilt, also, weil auch q_0 prim ist,

$$p_0 \sim q_0, \text{ etwa } q_0 = e \cdot p_0, \text{ mit } e \in E(R).$$

Es folgt

$$r' = p_1 \cdots p_s = (e \cdot q_1) \cdots q_t.$$

Ist $s = 1$, dann folgt $t = 1$, und wir sind fertig. Andernfalls argumentieren wir mit p_1 wie oben mit p_0 und erhalten schließlich per Induktion, daß $s = t$ gilt und die beiden Faktorisierungen äquivalent sein müssen.

□

7.3.18 Folgerung *Ist $\mathbb{K}$ ein Körper, dann ist $\mathbb{K}[x]$ ein Gaußbereich.*

Wir haben also insgesamt die folgende Kette von Implikationen (mit den Abkürzungen EB für Euklidischer Bereich, HIB für Hauptidealbereich, GB für Gaußbereich, IB für Integritätsbereich):

7.3.19 Satz *Für Integritätsbereiche R gelten die folgenden Implikationen:*

$$EB(R) \Rightarrow HIB(R) \Rightarrow GB(R) \Rightarrow IB(R).$$

Keine dieser Implikationen ist umkehrbar, das sieht man an:

7.3.20 Beispiele

- Die Menge $\{a + b(\frac{1+\sqrt{-19}}{2}) \mid a, b \in \mathbb{Z}\}$ ist ein Hauptidealbereich, aber kein Euklidischer Bereich. (s. T. Motzkin, *Bull. Amer. Math. Soc.* **55** (1949), 1142-1146)

- Der Polynomring $\mathbb{K}[x, y] := \mathbb{K}[x][y]$ ist ein Gaußbereich (s.u.), aber kein Hauptidealbereich (s. 7.3.8).

- $\{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$ ist ein Integritätsbereich, aber kein Gaußbereich:

  $$9 = 3 \cdot 3 = (2 + \sqrt{-5})(2 - \sqrt{-5}),$$

  die Darstellung als Produkt unzerlegbarer Elemente ist also nicht bis auf Äquivalenz eindeutig (s. 7.3.8).

7.3.21 Satz *In Gaußbereichen sind genau die primen Elemente unzerlegbar.*

Beweis: Primelemente sind unzerlegbar, das haben wir bereits gesehen. Die Umkehrung folgt (indirekter Beweis) so: Sei r unzerlegbar und Teiler von $s \cdot t$, etwa $r \cdot p = s \cdot t$. Nehmen wir an, $r \not| s$ und $r \not| t$. Da R Gaußbereich ist, existieren Zerlegungen $s \cdot t = \prod r_i$ und $r \cdot p = r \prod p_i$ in unzerlegbare Elemente, und es gilt: $r \not| r_i$, also $r \not\sim r_i$, und damit $r \prod p_i \not\approx \prod r_i$, ein Widerspruch.

□

7.3.22 Satz *Sind r, s Elemente eines Gaußbereichs R mit Produktdarstellungen*

$$r \sim \prod_{i=1}^{m} p_i^{m_i}, \; s \sim \prod_{i=1}^{m} p_i^{n_i}, \; p_i \textit{ unzerlegbar,}$$

dann besitzen diese Elemente größte gemeinsame Teiler und kleinste gemeinsame Vielfache, es gilt nämlich

$$\prod_{i=1}^{m} p_i^{\min\{m_i,n_i\}} \in ggT(r,s), \; \prod_{i=1}^{m} p_i^{\max\{m_i,n_i\}} \in kgV(r,s).$$

Beweis: Übungsaufgabe.

□

7.3.23 Definition *Ist R Integritätsbereich, so heißt $f \in R[x]$ genau dann* primitiv, *wenn $Grad(f) > 0$ und die Menge der gemeinsamen Teiler* aller *Koeffizienten von f die Einheiten sind.*

7.3.24 Hilfssatz *Ist R ein Integritätsbereich, $f \in R[x]$ und $r \in R$, dann gilt:*

- $r \mid f \iff r$ *teilt jeden Koeffizienten.*
- $f \in E(R[x]) \iff f \in E(R)$.
- *Ist R Gaußbereich, $f \in R[x]$, $Grad(f) > 0$, dann gibt es ein primitives Polynom $g \in R[x]$ und $r \in R$ mit $f = rg$.*
- *Ist $f = r_1 g_1 = r_2 g_2$ mit $r_i \in R$ und primitiven g_i, dann gilt $r_1 \sim r_2$ und $g_1 \sim g_2$.*

Beweis: Übungsaufgabe.

7.3.25 Hilfssatz *Ist R Gaußbereich, dann sind Produkte primitiver Polynome in $R[x]$ primitiv.*

Beweis: Seien

$$f = \sum_{i=1}^{m} a_i x^i, \; g = \sum_{i=1}^{n} b_i x^i, \; a_m \neq 0 \neq b_n,$$

primitive Polynome. Wir führen einen indirekten Beweis, nehmen also an,

$$h := f \cdot g = \sum_i c_i x^i$$

sei nicht primitiv. Es gibt dann ein Primelement $r \in R$, das jedes c_i teilt. Wegen der Primitivität von f und g gibt es minimale Indices i, j mit $r \not| a_i$ und $r \not| b_j$. Das ergibt den Widerspruch

$$r \text{ teilt } [c_{i+j} = \underbrace{a_0 b_{i+j} + a_1 b_{i+j-1} + \ldots}_{\text{teilbar durch } r} + a_i b_j \underbrace{+ \ldots + a_{i+j} b_o}_{\text{teilbar durch } r}].$$

□

7.3.26 Hilfssatz *Ist R ein Gaußbereich, $\mathbb{K} = B(R, R^*)$ der Quotientenkörper, $f \in R[x]$ primitiv, dann gilt:*

$$f \textit{ unzerlegbar in } \mathbb{K}[x] \iff f \textit{ unzerlegbar in } R[x].$$

Beweis:
i) Sei $f \in R[x]$ in $\mathbb{K}[x]$ unzerlegbar, es gebe also keinen Teiler $g \in \mathbb{K}[x]$ von f mit $0 < Grad(g) < Grad(f)$. Jeder echte Teiler t von f in $R[x]$ ist also vom Grad 0: $t \in R$. Da f nach Voraussetzung primitiv ist, und t alle Koeffizienten teilt, gilt $t \in E(R)$, d.h. f ist auch in $R[x]$ unzerlegbar.
ii) Sei jetzt f unzerlegbar in $R[x]$ und Produkt $f = gh$ zweier Polynome in $\mathbb{K}[x]$, die echte Teiler seien. Da $\mathbb{K}$ der Quotientenkörper von R ist, gibt es $a, b \in R^*$ mit $ag, bh \in R[x]$. Nach ?? gibt es primitive $h_1, g_1 \in R[x]$ mit

$$abf = abgh = a_1 b_1 g_1 h_1.$$

Nach 7.3.25 ist deren Produkt $g_1 h_1$ primitiv. Weil f als primitiv vorausgesetzt ist, folgt $f \sim g_1 h_1$, in $R[x]$. Es ergibt sich

$$g_1 \mid f \wedge 0 < Grad(g) = Grad(g_1) < Grad(f).$$

Dies impliziert aber den Widerspruch, g_1 sei ein echter Teiler von f in $R[x]$.

□

7.3.27 Hilfssatz *Ist R Gaußbereich mit Quotientenkörper $\mathbb{K}$ und primitiven Polynomen $f, g \in R[x]$, so gilt*

$$f \sim g \textit{ in } R[x] \iff f \sim g \textit{ in } \mathbb{K}[x].$$

Beweis:
i) Ist $f \sim g$ in $\mathbb{K}[x]$, etwa $f = eg, e \in E(\mathbb{K}[x]) = \mathbb{K}^*$, dann gibt es $s, t \in R^*$ mit $e = s/t$, also $tf = sg$, was nach ?? $f \sim g$ impliziert, in $R[x]$.
ii) Die Umkehrung ist trivial.

□

7.3.28 Satz *Polynomringe über Gaußbereichen sind ebenfalls Gaußbereiche.*

Beweis: Ist R Gaußbereich, dann ist R mindestens Integritätsbereich. Bezeichnen wir mit $\mathbb{K}$ den Quotientenkörper, dann ist $\mathbb{K}[x]$ Hauptidealbereich, also Gaußbereich.
i) Wir beweisen zunächst die *Existenz* von Faktorisierungen streng zerlegbarer Polynome f in unzerlegbare.

1. Sei f ein *primitives* streng zerlegbares Polynom in $R[x]$. Wegen 7.3.26 ist es auch in $\mathbb{K}[x]$ streng zerlegbar, und es gelte dort:
$$f = \overline{p_1} \cdots \overline{p_s},$$
mit unzerlegbaren $\overline{p_i}$. Wir wählen $a_i \in R$ mit $a_i \overline{p_i} \in R[x]$ und dann (vgl. ?? i)) $b_i \in R$ mit
$$a_i \overline{p_i} = b_i p_i \in R[x], p_i \text{ primitiv.}$$
Es gilt dann, für $a := \prod a_i, b := \prod b_i$: $af = bp_1 \cdots p_s$. Nun sind aber auch die p_i unzerlegbar in $\mathbb{K}[x]$. Da sie primitiv sind, sind diese Polynome nach 7.3.26 auch unzerlegbar in $R[x]$. Nach 7.3.25 ist auch deren Produkt primitiv. Da f zu diesem Produkt assoziiert ist (??), ist demnach f selbst Produkt unzerlegbarer Polynome aus $R[x]$.

2. Sei $f \in R[x]$ jetzt *irgendein* streng zerlegbares Polynom in $R[x]$. Wir unterscheiden zwei Fälle:

   a) Ist $f \in R$, dann gibt es, da R Gaußbereich ist, unzerlegbare a_i mit $f = a_1 \cdots a_h$. Eine solche Zerlegung ist in R bis auf Äquivalenz eindeutig, also auch in $R[x]$. Bei $f \in R$ gilt also die Behauptung.

   b) Ist $f \notin R$, dann gibt es $d \in R$ mit $f = dg$ und einem primitiven $g \in R[x]$. Wie unter 1. ergeben sich Zerlegungen $g = p_1 \cdots p_s$, $d = d_1 \cdots d_h$, also auch Faktorisierungen von f:
$$f = d_1 \cdots d_h p_1 \cdots p_s.$$

ii) Zum Beweis der *Äquivalenz* zweier solcher Zerlegungen betrachten wir
$$f = d_1 \cdots d_h p_1 \cdots p_s = e_1 \cdots e_k q_1 \cdots q_t,$$
mit unzerlegbaren $d_i, e_i \in R$, und unzerlegbaren $p_i, q_i \in R[x] \backslash R$, von denen wir ohne Einschränkung annehmen können, sie seien zusätzlich auch primitive Polynome. Nach ?? folgt dann
$$d_1 \cdots d_h \sim e_1 \cdots e_k, \text{ in } R.$$
Da R Gaußbereich ist, impliziert dies die Äquivalenz dieser beiden Produkte. Die Produkte $d_1 \cdots d_h$ und $e_1 \cdots e_k$ sind dann auch in $R[x]$ äquivalent, insbesondere gilt $h = k$.

Auch die polynomialen Faktoren $p_1 \cdots p_s$ und $q_1 \cdots q_t$ sind assoziiert in $R[x]$, also auch in $\mathbb{K}[x]$, dies ist aber ein Gaußbereich, also sind sie dort sogar äquivalent, und insbesondere gilt $s = t$. Geeignete Paare sind also assoziiert in $\mathbb{K}[x]$, nach 7.3.27 auch assoziiert in $R[x]$, so daß diese beiden Produkte sogar in $R[x]$ äquivalent sind. Insgesamt ergibt sich daraus die Äquivalenz der beiden Faktorisierungen von f.

□

Wiederholte Anwendung dieses Satzes ergibt

7.3.29 Folgerung *Ist R ein Gaußbereich, $n \in \mathbf{N}$, dann ist der Polynomring*

$$R[x_1, \ldots, x_n] := R[x_1, \ldots, x_{n-1}][x_n]$$

ein Gaußbereich.

7.3.30 Unzerlegbarkeitskriterien *Sei R ein Integritätsbereich, dann gilt im Polynomring $R[x]$:*

- *Die linearen Polynome $f = r + sx, s \in E(R)$, sind sämtlich unzerlegbar.*
- *Für alle $r \in R$ ist $f(x) \in R[x]$ genau dann unzerlegbar, wenn $f(x + r)$ unzerlegbar ist.*
- *Ist $f \in R[x]$ primitiv und als Polynom über $\mathbb{K} = B(R, R^*)$ unzerlegbar, genau dann ist f auch über R unzerlegbar (7.3.26).*
- *Ist $R = \mathbb{Z}$, $f \in \mathbb{Z}[x]$ unzerlegbar über $\mathbb{Z}_p$ (genauer: $\overline{f}$, das aus f durch Reduktion der Koeffizienten modulo p hervorgeht) und teilt p den Leitkoeffizienten nicht, dann ist f auch in $\mathbb{Z}[x]$ unzerlegbar.*
- Kriterium von Eisenstein: *Ist R Gaußbereich, $f = \sum_i a_i x^i \in R[x]$ primitives Polynom, $p \in R$ ein Primelement, für das in R gilt:*

  $$p \not| a_n, p \mid a_{n-1}, p \mid a_{n-2}, \ldots, p \mid a_0, p^2 \not| a_0,$$

  dann ist f über R unzerlegbar.

Beweis: Das Eisensteinkriterium zeigen wir indirekt: Ist $f = \sum a_i x^i = gh$, mit $g = \sum b_i x^i, h = \sum c_i x^i$, dann teilt p entweder b_0 nicht oder c_0 nicht, o.E. $p \not| c_0$. Da p kein Teiler von a_n ist, gibt es einen kleinsten Index i mit $p \not| b_i$. Nun ist aber

$$a_i = b_i c_0 + (b_{i-1} c_1 + \ldots + b_0 c_i),$$

wobei p den eingeklammerten Summanden auf der rechten Seite teilt, aber nicht $b_i c_0$, also auch nicht a_i. Damit ist $i = n$, und der Grad von h ist Null. Da f als primitiv vorausgesetzt war, ist $h \in E(R)$.

□

Das Eisensteinkriterium veranlaßt uns zu folgender

7.3.31 Definition *Ein Polynom $f \in R[x]$, R Gaußbereich, heißt* irreduzibel, *wenn für $g, h \in R[x]$ gilt:*

$$[f = gh \;\Rightarrow\; Grad(g) = 0 \;\vee\; Grad(h) = 0]$$

Im Falle, daß f primitiv oder R ein Körper ist, sind Unzerlegbarkeit und Irreduzibilität gleichwertig.

7.3.32 Beispiele

- $f(x) := 2 + 4x + 6x^2 + 4x^3 + x^4 = (x+1)^4 + 1$ ist über $\mathbb{Z}$ unzerlegbar nach dem Eisensteinkriterium, damit ist $y^4 + 1$ unzerlegbar über $\mathbb{Q}$.

- $f := -135 + 17x - 8x^2 + x^3$ ergibt, modulo $p = 2$, das Polynom $\overline{f} = 1 + x + x^3 \in \mathbb{Z}_2[x]$, das dort unzerlegbar ist, denn eine Zerlegung müßte einen Linearfaktor enthalten, $1 + x + x^3$ also eine Nullstelle besitzen, was offensichtlich nicht der Fall ist.

- Mit Hilfe einer Kombination des Eisensteinkriteriums mit dem Transformationskriterium ergibt sich auch die Unzerlegbarkeit der wichtigen *Kreisteilungspolynome*

  $$\Phi_p(x) := \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \ldots + 1 \in \mathbb{Z}[X],$$

  für Primzahlen p. Denn

  $$\Phi_p(x+1) = x^{p-1} + \binom{p}{1} x^{p-2} + \binom{p}{2} x^{p-3} + \ldots + \binom{p}{p-1}$$

  ist unzerlegbar nach dem Eisensteinkriterium.

◇

7.4 Zerlegung von Ringen, Moduln und abelschen Gruppen

Die Struktur eines Ringes R wirkt sich auch auf die Struktur von R-Moduln aus, insbesondere gewisse direkte Zerlegung von R in Ideale. Zu deren Beschreibung dienen die sogenannten *Idempotente,* das sind die Elemente $e \in R$ mit

$$0 \neq e = e^2.$$

Solche Idempotente heißen *zentral,* wen sie im *Zentrum* von R liegen, d. h. wenn $er = re$, für alle $r \in R$. Zwei Idempotente $e, f \in R$ heißen *orthogonal,* wenn

$$ef = fe = 0.$$

7.4.1 Beispiele

i) In einem Ring R mit Einselement ist mit einem Idempotent e auch $1-e$ ein Idempotent, und diese beiden Idempotente sind orthogonal. Sie ergeben eine direkte Zerlegung von R in Linksideale, die sogenannte *Piercesche Zerlegung*

$$R = Re \oplus R(1-e).$$

ii) Ist $n > 1$ eine natürliche Zahl mit Primfaktorzerlegung

$$n = \prod_{i=1}^{r} p_i^{a_i}$$

(also $p_i \neq p_j$, für $i \neq j$, und $a_i > 0$), dann sind die Quotienten $n_i := n/p_i^{a_i}$, $i = 1, \ldots, r$ teilerfremd, es gibt also $z_i \in \mathbb{Z}^*$ mit

$$\sum_{i=1}^{r} n_i z_i = 1.$$

Die Restklassen dieser Summanden in $\mathbb{Z}_n$, also die

$$\bar{e_i} := n_i z_i + (n) \in \mathbb{Z}_n$$

sind orthogonale Idempotente: Die Orthogonalität folgt aus der Teilbarkeit von $n_i n_j$ durch n. Die Idempotenz ergibt sich aus der Tatsache, daß die $\bar{e_i}$ eine *Zerlegung der Eins* ergeben:

$$\bar{1} = \sum_i \bar{e_i},$$

denn hieraus folgt durch Multiplikation mit $\bar{e_j}$ und mit der Orthogonalität: $\bar{e_j} = \bar{e_j}^2$.

◇

Das letzte Beispiel läßt sich leicht verallgemeinern: Jede Zerlegung $1 = \sum_i e_i$ der Eins in paarweise orthogonale Idempotente ergibt eine direkte Zerlegung von R in Linksideale:

7.4.2 $$_RR = \oplus_i Re_i.$$

Wenn die Idempotente zusätzlich noch zentral sind ergibt sich daraus ein sehr wichtiges Resultat über Zerlegungen von R-Moduln:

7.4.3 Satz *R sei ein Ring mit 1, $1 = \sum_i e_i$ eine Zerlegung der Eins in zentrale und paarweise orthogonale Idempotente, M ein R-Linksmodul. Dann gilt:*

- *Aus der Zerlegung der Eins des Ringes ergibt sich eine eine direkte Zerlegung von M in R-Linksmoduln:*

  $$_RM = \oplus_i e_i M.$$

  (Entsprechendes gilt natürlich für Rechtsmoduln!)

- *Das zweiseitige Ideal $Re_i = e_i Re_i$ ist ein Unterring von R mit Einselement e_i, es gilt*

  $$R = \oplus_i e_i Re_i.$$

- *$e_i M$ ist ein $e_i Re_i$-Linksmodul. Jede direkte Zerlegung von $e_i M$ in $e_i Re_i$-Untermoduln $e_i M = L_1 \oplus \ldots \oplus L_k$ kann als Zerlegung in R-Untermoduln aufgefaßt werden.*

Beweis:

i) Wegen $re_i m = e_i rm$ ist $e_i M$ ein Untermodul. Die Summe $\sum_i e_i M$ dieser Untermoduln ist $M : m = 1m = \sum_i e_i m$, und sie ist direkt: $0 = \sum_i e_i m_i$ ergibt, durch Multiplikation von links mit e_j, daß $0 = e_j m_j$.
ii) $e_i Re_i$ ist additiv und multiplikativ abgeschlossen, also Unterring. e_i ist offenbar das Einselement, und die angegebene direkte Zerlegung folgt unmittelbar aus 7.4.2.
iii) Die L_j sind, nach Definition, $e_i Re_i$-Untermoduln. Sie sind aber auch R-Untermoduln, bzw. können als solche aufgefaßt werden. Jedes $l \in L_j$ ist ja von der Form $e_i m$, so daß wegen der Orthogonalität folgt

$$rl = re_i m = (\sum_\nu re_\nu) e_i m = (\sum_\nu e_\nu re_\nu) e_i m = e_i re_i e_i m = e_i rm \in L_j.$$

□

Die Isomorphiesätze übertragen sich von Gruppen auf Moduln, wenn man “Normalteiler” durch “zulässige Normalteiler”, d. h. durch “Untermoduln” ersetzt. Es gilt vor allem

7.4.4 Der Satz von Jordan und Hölder *Besitzt ein R-Linksmodul Kompositionsreihen, dann sind diese paarweise isomorph.*

7.4.5 Beispiele

i) Kompositionsreihen n-dimensionaler Vektorräume V sind von der Form

$$V = V_0 > V_1 > \ldots > V_n = 0,$$

ii) Ist R ein Hauptidealbereich, $p \in R$ prim, dann ist

$$R/(p^n) \supset (p)/(p^n) \supset (p^2)/(p^n) \supset \ldots \supset (p^{n-1})/(p^n) \supset (p^n)/(p^n) = 0$$

die einzige Kompositionsreihe des Faktorrings $R/(p^n)$. Um das einzusehen, betrachtet man den kanonischen Epimorphismus $\nu: r \mapsto r + (p^n)$, von R auf den Faktorring. Ein Ideal I im Faktorring ist Bild eines Ideals J aus R, welches (p^n) umfaßt. R ist Hauptidealring, es gibt also $q \in R$ mit $J = (q)$. Ist jetzt p^m die maximale Potenz von p, die in q aufgeht (R ist Gaußbereich!), also $q \sim p^m$, und damit $I = (p^m)/(p^n)$.
iii) Dagegen besitzen Moduln, die direkte Summe $M = M_1 \oplus M_2$ von nicht trivialen Untermoduln M_i sind, mehr als eine Kompositionsreihe, da andernfalls $M \supset M_1 \supset 0$ und $M \supset M_2 \supset 0$ eine gemeinsame Verfeinerung besitzen müßten, was nicht sein kann, da keiner dieser Untermoduln im anderen liegt.

Wir heben Moduln mit eindeutig bestimmter Kompositionsreihe hervor, sie sollen *einreihig* heißen. Moduln, die von einem Element erzeugt werden, also $M = {}_R\langle m\rangle$, nennen wir *zyklisch.*

7.4.6 Satz *Sind die beiden Moduln ${}_RR$ und R_R einreihig, dann ist jeder R-Linksmodul M direkte Summe von zyklischen, also einreihigen, Untermoduln.*

Beweis (W. Müller):

i) Ist $m \in M$, dann ist der hiervon erzeugte Untermodul Rm isomorph zu $R/\mathrm{Kern}(\nu)$, wenn $\nu: r \mapsto rm$ der kanonische Epimorphismus ist. Als Faktormodul hat er demnach Kompositionsreihen endlicher Länge. Für jedes Erzeugendensystem $\{m_0, \ldots, m_{k-1}\}$ von M gilt dann mit den Längen $l(Rm_i)$ der Kompositionsreihen der Rm_i die Abschätzung

$$l(M) \leq \sum_i l(Rm_i).$$

Die rechte Seite dieser Ungleichung ist endlich, wir können also annehmen, daß das Erzeugendensystem so gewählt sei, daß dieser Wert minimal ist. Wir wollen zeigen, daß dann $M = \oplus_i Rm_i$ gilt.
Sei $0 = \sum_i r_i m_i$. Weil R_R einreihig ist, gibt es ein größtes Rechtsideal r_iR mit $r_i m_i \neq 0$, und wir können ohne Einschränkung der Allgemeinheit annehmen, dieses sei r_0R. In r_0R liegen alle $r_i m_i \neq 0$, es gibt also $q_i \in R$ mit $r_0 q_i = r_i$, falls $r_i m_i \neq 0$. Sei $q_0 = 1$ und noch $q_i := 0$ gesetzt, falls $r_i m_i = 0$. Wir haben dann

$$0 = \sum_{i=0}^{k-1} r_i m_i = r_0(m_0 + \sum_{i=1}^{k-1} q_i m_i) =: r_o m_0'.$$

Wir führen jetzt einen Widerspruch herbei durch Nachweis von $l(Rm_0') < l(Rm_0)$. Dazu betrachten wir die kanonischen Epimorphismen ν bzw. ν' von ${}_RR$ auf Rm_0 bzw. auf Rm_0'. Wegen der Einreihigkeit von ${}_RR$ liegt einer der beiden Kerne im anderen. Weil $r_0m_0 \neq 0$, aber $r_0m_0' = 0$, gilt, haben wir demnach $\mathrm{Kern}(\nu) \subset \mathrm{Kern}(\nu')$. Es folgt $l(\mathrm{Kern}(\nu)) < l(\mathrm{Kern}(\nu'))$ und damit

$$l(Rm_0') = l({}_RR) - l(\mathrm{Kern}(\nu')) < l({}_RR) - l(\mathrm{Kern}(\nu)) = l(Rm_0).$$

Da, mit $\{m_0, m_1, \ldots, m_{k-1}\}$ auch $\{m_0', m_1, \ldots, m_{k-1}\}$ Erzeugendensystem ist, steht dies im Widerspruch zur angenommenen Miniamlität der Längensumme $\sum l(Rm_i)$. Es muß also auch $r_im_i = 0$ gelten.

□

Abelsche Gruppen sind $\mathbb{Z}$-Moduln, die Resultate zur Struktur von Ringen und von Moduln über diesen können also auch auf abelsche Gruppen angewandt werden.

7.4.7 der Hauptsatz über abelsche Gruppen *Jede endliche abelsche Gruppe ist direkte Summe zyklischer Untergruppen von Primzahlpotenzordnung.*

Beweis: $(A, +)$ seieine endliche abelsche Gruppe, $n := |A| > 1$. A ist, wie bereits erwähnt, ein $\mathbb{Z}$-Modul,

$$za := \begin{cases} a + \ldots + a, & z\text{-mal, falls } z > 0, \\ -a - \ldots - a, & (-z)\text{-mal, falls } z < 0. \end{cases}$$

Wegen $na = 0$ kann A auch als $\mathbb{Z}_n$-Modula aufgefaßt werden, vermöge

$$\bar{z}a := za, \text{ falls } \bar{z} = z + (n) \in \mathbb{Z}_n.$$

Nun sei wieder $n = \prod_i p_i^{a_i}$ die Primfaktorzerlegung von n, $n_i := n/p_i^{a_i}$, $z_i \in \mathbb{Z}^*$ mit $\sum_i n_iz_i = 1$, $e_i := n_iz_i$, $\bar{e_i} := e_i + (n) \in \mathbb{Z}_n$. Wie oben bereits diskutiert, sind die $\bar{e_i}$ zentrale Idempotente, die eine Zerlegung der Eins in $\mathbb{Z}_n$ bilden, $\sum_i \bar{e_i} = \bar{1}$, und wir erhalten daraus direkte Zerlegungen $\mathbb{Z}_n = \oplus_i \bar{e_i}\mathbb{Z}_n$ sowie

$$A = \oplus_i e_iA = \oplus_i \bar{e_i}A.$$

Die erste der beiden direkten Summen ist eine Zerlegung von A als $\mathbb{Z}$-Modul, die zweite ist eine Zerlegung von A als $\mathbb{Z}_n$-Modul. Die Summanden sind dabei, als Teilmengen von A, gleich: $e_iA = \bar{e_i}A$.
Die e_iA sind sogar $\mathbb{Z}_{p_i^{a_i}}$-Moduln, denn $p_i^{a_i}e_iA = nz_iA = 0$. Die Ringe $\mathbb{Z}_{p_i^{a_i}}$ sind einreihig, nach obigem Satz ist also e_iA eine direkte Summe zyklischer Untermoduln. Diese haben, als Faktormoduln von $\mathbb{Z}_{p_i^{a_i}}$, als Ordnung eine Potenz von p.

□

Analog verläuft der Beweis von

7.4.8 Die Jordansche Normalform einer quadratischen Matrix *Jede quadratische Matrix endlicher Reihenzahl über einem algebraisch abgeschlossenen Zahlkörper kann auf Blockdiagonalgestalt transformiert werden, wobei die*

Blöcke die folgende Form haben

$$\begin{pmatrix} \lambda & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda \end{pmatrix},$$

mit einem Eigenwert λ der betrachteten Matrix.

Beweis: A sei eine $n \times n$-matrix über $\mathbb{K}$, einem algebraisch abgeschlossenen Körper (d. h. jedes Polynom in $\mathbb{K}[x]$ zerfällt in Linearfaktoren.

i) Der Vektorraum $\mathbb{K}^n$ wird zu einem $\mathbb{K}[x]$-Modul, wenn wir die Operation eines Polynoms $\sum_i \kappa_i x^i$ als Anwendung der linearen Abbildung definieren, die durch Einsetzen der Matrix in das Polynom entsteht:

$$(\sum_i \kappa_i x^i)(v) := \sum_i \kappa_i A^i(v).$$

Ist jetzt E_i der i-te Einheitsvektor, $0 \leq i \leq n-1$, dann hat (aus Dimensionsgründen) der Epimorphismus

$$\mathbb{K}[x] \to \mathbb{K}[x]E_i : (\sum_i \kappa_i x^i) \mapsto (\sum_i \kappa_i x^i)(E_i) = \sum_i \kappa_i A^i(E_i)$$

einen nicht trivialen Kern, also ein Ideal, das von einem normierten Polynom vom Grad > 0 erzeugt wird. Es folgt

$$\mathbb{K}[x]E_i \simeq \mathbb{K}[x]/(q_i).$$

Der Vektorraum $\mathbb{K}^n$ wird vom Produkt $q = q_0 \cdots q_{n-1}$ der diese Kerne erzeugenden Polynome q_i annulliert und kann deshalb als $\mathbb{K}[x]/(q)$-Modul aufgefaßt werden. Ist jetzt $q = \prod_i p_i^{a_i}$ die Primfaktorzerlegung von q, mit normierten unzerlegbaren und paarweise verschiedenen Polynomen p_i, dann erhalten wir — ganz wie im Beweis des Hauptsatzes über abelsche Gruppen — aus den teilerfremden $r_i := q/p_i^{a_i}$ eine Zerlegung der Eins:

$$1 = \sum_i r_i s_i,\ s_i \in \mathbb{K}[x]^*.$$

Die Restklassen $\bar{e_i}$ der Summanden e_i, also

$$\bar{e_i} := r_i s_i + (p),$$

sind orthogonale Idempotente mit

$$\mathbb{K}^n = \oplus_i \bar{e_i}\mathbb{K}^n = \oplus_i e_i \mathbb{K}^n,$$

sowie $\bar{e_i}\mathbb{K}^n = e_i \mathbb{K}^n$.

Der Unterraum $e_i\mathbb{K}^n$ ist $\mathbb{K}[x]/(p_i^{a_i})$-Modul also Modul über einem einreihigen Ring und deshalb Summe von Unterräumen U_{ij}, die zyklische $\mathbb{K}[x]/(p_i^{a_i})$-Moduln sind, d. h. isomorph zu Faktormoduln nach den Idealen dieses Rings:

$$U_{ij} = \mathbb{K}[x]u_{ij} \simeq \mathbb{K}[x]/(p_i^{a_i})/(p_i^{b_{ij}})/(p_i^{a_i}) \simeq \mathbb{K}[x]/(p_i^{b_{ij}}).$$

Insgesamt haben wir also die folgende Zerlegung von $\mathbb{K}^n$ als direkte Summe erhalten:

$$\mathbb{K}^n = \oplus_i \oplus_j \mathbb{K}[x]u_{ij}\,.$$

Es bleibt die Wirkung der Matrix A auf Basen dieser Unterräume $\mathbb{K}[x]u_{ij}$ zu berechnen.

Wir verwenden dazu die Voraussetzung, daß $\mathbb{K}$ algebraisch abgeschlossen ist, die unzerlegbaren Polynome, insbesondere auch die p_i, sind also linear: $p_i = x - \lambda_i$, für eingeeignetes $\lambda_i \in \mathbb{K}$. Die Potenzen p_i^k, $0 \leq k \leq b_{ij} - 1$, bilden also eine Basis von $\mathbb{K}[x]/(p_i^{b_{ij}}) \simeq U_{ij}$. Es gilt also

$$U_{ij} = \ll u_{ij}, p_i u_{ij}, p_i^2 u_{ij}, \ldots, p_i^{b_{ij}-1} \gg .$$

Wir brauchen jetzt nur noch zu bemerken, daß

$$Ap_i u_{ij} = xp_i^k u_{ij} = \lambda_i p_i^k u_{ij} + p_i^{k+1} u_{ij},$$

Die Matrix der Einschränkung von A auf den Unterraum U_{ij} hat also wirklich die Form

$$\begin{pmatrix} \lambda_i & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda_i \end{pmatrix}.$$

Dabei zeigt die letzte Spalte dieser Matrix, daß λ_i tatsächlich ein Eigenwert von A ist

□

Will man auf die Voraussetzung algebraischer Abgeschlossenheit von $\mathbb{K}$ verzichten, dann kann man die Linearität der Polynome p_i nicht mehr voraussetzen. Ist etwa $p_i = \sum_i a_i x^i$ normiert und vom Grad m_i, dann ist

$$U_{ij} = \ll u_{ij}, xu_{ij}, \ldots, x^{m_i-1}u_{ij}, p_i u_{ij}, xp_i u_{ij}, \ldots, x^{m_i-1}p_i^{b_{ij}-1} \gg .$$

Die Matrix von A auf U_{ij} hat in diesem Fall die Form

$$\left(\begin{array}{ccccc|ccccc|c}
0 & & & & -a_0 & & & & & & \\
1 & 0 & & & -a_1 & & & & & & \\
 & 1 & 0 & & -a_2 & & & & & & \\
 & & \ddots & \ddots & \vdots & & & & & & \\
 & & & 1 & -a_{m_i-1} & & & & & & \\
\hline
 & & & & 1 & 0 & & & & -a_0 & \\
 & & & & & 1 & 0 & & & -a_1 & \\
 & & & & & & 1 & 0 & & -a_2 & \\
 & & & & & & & \ddots & \ddots & \vdots & \\
 & & & & & & & & 1 & -a_{m_i-1} & \\
\hline
 & & & & & & & & & 1 & \ddots
\end{array}\right).$$