LINEARE ALGEBRA I

ao.Univ.Prof. Dr. Michael Drmota

Wintersemester 2005/2006

Inhaltsverzeichnis

Kapitel 1

Mengen und Relationen

1.1 Logische Grundbegriffe

1.1.1 Verknüpfungen von Aussagen

Unter einer **(mathematischen) Aussage** versteht man einen sprachlichen Ausdruck, dem eindeutig der **Wahrheitswert wahr** (= **w**) oder **falsch** (= **f**) zugeordnet werden kann. Üblicherweise werden *Aussagen* durch einen *Aussagesatz* formuliert, wie z.B.:

> 5 *ist eine Primzahl.*
> 17 *ist eine gerade Zahl.*
> 4 *ist kleiner als* 7*.*

Man beachte, daß man hier von dem Prinzip ausgeht, daß eine Aussage nur entweder wahr oder falsch sein kann. Dieses Prinzip heißt **Prinzip vom ausgeschlossenen Dritten** (Terium non datur). Man spricht auch von einer **zweiwertigen Logik**.

Aussagen können auf verschiedene Arten und Weisen miteinander verknüpft werden.

1. **Konjunktion**: Aus zwei Aussagen wird durch Einfügen des Wortes **und** eine neue Aussage gewonnen, die Konjunktion der beiden Aussagen. Z.B. ist

   > 4 *is kleiner als* 7 **und** 7 *ist eine Primzahl.*

   die Konjunktion der beiden Aussagen

   > *4 ist kleiner als 7.* *7 ist eine Primzahl.*

   Die Konjunktion zweier Aussagen erhält genau dann den Wahrheitswert **w**, wenn die beiden ursprünglichen Aussagen den Wert **w** haben. (In allen anderen Fällen erhält die Konjunktion den Wert **f**.)

2. **Disjunktion**: Aus zwei Aussagen wird durch Einfügen des Wortes **oder** eine neue Aussage gewonnen, die Disjunktion der beiden Aussagen. Z.B. ist

   > 4 *ist eine gerade Zahl* **oder** 4 *ist eine ungerade Zahl.*

die Disjunktion der beiden Aussagen

4 *ist eine gerade Zahl.* 4 *ist eine ungerade Zahl.*

Die Disjunktion zweier Aussagen erhält genau dann den Wahrheitswert **w**, wenn wenigstens eine der ursprünglichen Aussagen den Wert **w** hat. (Nur wenn beide Aussagen den Wert **f** haben, hat auch die Disjunktion den Wert **f**.)

3. **Implikation**: Aus zwei Aussagen wird durch Einfügen der Worte **wenn - dann** eine neue Aussage gewonnen, die Implikation der beiden Aussagen. Z.B. ist

   Wenn 5 *eine gerade Zahl ist,* **dann** *ist* 5 *keine Primzahl.*

   die Implikation der beiden Aussagen

   5 *ist eine gerade Zahl.* 5 *ist keine Primzahl.*

   Die Implikation zweier Aussagen hat genau dann den Wert **f**, wenn die erste Aussage den Wert **w**, aber die zweite den Wert **f** hat. Man beachte, daß bei der Implikation die Reihenfolge wesentlich ist. Weiters hat eine Implikation immer den Wert **w**, wenn die erste Aussage den Wert **f** hat, unabhängig davon, wie die zweite Aussage lautet (ex falso quodlibet) und wenn die zweite Aussage den Wert **w** hat (verum ex quodlibet).
   Statt der Worte *wenn - dann* kann man auch *aus - folgt* oder *impliziert* verwenden.

4. **Äquivalenz**: Aus zwei Aussagen wird durch Einfügen der Worte **genau dann - wenn** eine neue Aussage gewonnen, die Äquivalenz der beiden Aussagen. Z.B. ist

   521 *ist* **genau dann** *durch* 3 *teilbar,* **wenn** $5+2+1$ *durch* 3 *teilbar ist.*

   die Äquivalenz der beiden Aussagen

   521 *ist durch* 3 *teilbar.* $5+2+1$ *ist durch* 3 *teilbar.*

   Die Äquivalenz zweier Aussagen hat genau dann den Wert **w**, wenn den beiden ursprünglichen Aussagen dieselben Wahrheitswerte zugeordnet sind.
   Statt der Worte *genau dann - wenn* werden auch die Worte *dann und nur dann - wenn* oder *äquivalent zu* verwendet.

5. **Negation**: Fügt man in einer Aussage (an geeigneter Stelle) das Wort **nicht** ein, so entsteht eine neue Aussage, die Negation der ursprünglichen Aussage. Z.B. ist

   17 *ist* **nicht** *durch* 3 *teilbar.*

   die Negation der Aussage

   17 *ist durch* 3 *teilbar.*

   Eine negierte Aussage hat genau dann den Wert **w**, wenn die unnegierte (ursprüngliche) Aussage der Wert **f** hat.

1.1.2 Junktoren

Zur Vereinfachung der Notation ist es üblich, Aussagen durch *Symbole* $p_1, p_2, \ldots$ zu bezeichnen und anstelle von *und* das Symbol $\wedge$, anstelle von *oder* das Symbol $\vee$, anstelle von *wenn − dann* das Symbol $\rightarrow$, anstelle von *genau dann, wenn* das Symbol $\leftrightarrow$ und anstelle von *nicht* das Symbol $\neg$ zu verwenden. Die Symbole $\wedge, \vee, \rightarrow, \leftrightarrow, \neg$ werden in diesem Zusammenhang als **Junktoren** bezeichnet. Ist z.B.

$p_1 = 3 + 5$ *ist gerade.*
$p_2 = 3$ *ist gerade.*
$p_3 = 5$ *ist gerade.*

so kann die Aussage

$3 + 5$ *ist genau dann gerade, wenn* 3 *und* 5 *gerade sind oder* 3 *und* 5 *nicht gerade sind.*

durch

$$p_1 \leftrightarrow ((p_2 \wedge p_3) \vee (\neg p_2 \wedge \neg p_3))$$

formalisiert werden.

Die Operationssymbole $\wedge, \vee, \rightarrow, \leftrightarrow, \neg$ können sinnvollerweise auch auf die Wahrheitswerte **w** und **f** angewandt werden.

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

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

$\rightarrow$	**f**	**w**
f	**w**	**w**
w	**f**	**w**

$\leftrightarrow$	**f**	**w**
f	**w**	**f**
w	**f**	**w**

$\neg$	
f	**w**
w	**f**

.

Durch diese Wahl ist sichergestellt, daß der Wahrheitswert einer durch Verknüpfungen von Einzelaussagen $p_1, p_2, \ldots$ gewonnenen Aussage dadurch bestimmt werden kann, daß man $p_1, p_2, \ldots$ durch ihre Wahrheitswerte ersetzt und die Junktoren als Operationssymbole für **w** und **f** interpretiert.

1.1.3 Quantoren und Prädikate

In der Mathematik haben sogenannte *atomare* (d.h. unzerlegbare) Aussagen eine *Subjekt-Prädikatstruktur*, z.B. kann in

4 *ist eine gerade Zahl.*

4 als Subjekt und *gerade Zahl* als Prädikat interpretiert werden. Die Aussage

17 *ist eine gerade Zahl.*

unterscheidet sich von der ersten nur im Subjekt. Es ist daher naheliegend, das Prädikat *gerade Zahl* zu einem Symbol P zu abstrahieren und die **Aussageform** $P(x)$

x ist eine gerade Zahl.

zu betrachten, wobei jetzt x als **Gegenstandsvariable** fungiert. Setzt man für x einen Wert ein, so entsteht aus der Aussageform $P(x)$ wieder eine Aussage, z.B. $P(4)$, $P(17)$, der wieder ein Wahrheitswert zugeordnet werden kann. Andererseits kann man Aussagen der Form

Für alle ... gilt ... und *Es gibt ein ... so daß ...*

mit Hilfe von Aussageformen bilden:

Für alle x gilt $P(x)$. Es gibt ein x so daß $P(x)$.

Fomalisiert wird dies durch die **Quantoren**, den **Allquantor** $\forall$ und den **Existenzquantor** $\exists$. So bedeutet

$$\forall x\ P(x),$$

daß *alle möglichen x die Eigenschaft P haben* und

$$\exists x\ P(x),$$

daß *es wenigstens ein x gibt, das die Eigenschaft P hat.*

Der Wahrheitswert von $\forall x\ P(x)$ ist genau dann **w**, wenn die Aussage $P(x)$ für alle möglichen x der Wert **w** hat, entsprechend hat $\exists x\ P(x)$ genau dann den Wert **w**, wenn es wenigstens ein x gibt, für das $P(x)$ den Wert **w** hat.

Eine Aussageform $P(x)$ heißt auch **einstelliges Prädikat**. Entsprechend werden auch **mehrstellige Prädikate** $Q(x_1, x_2, \ldots, x_n)$ verwendet. Beispielsweise ist

x ist größer als y

ein zweistelliges Prädikat.

Selbstverständlich kann man Prädikate mit Junktoren untereinander bzw. mit gewöhnlichen Aussagen verbinden und, solange noch **freie Gegenstandsvariable** vorhanden sind, Quantoren anwenden. Führt man dies mit Aussagensymbolen $(p_1, p_2, \ldots)$ und Prädikatsymbolen $(P_1, P_2, \ldots)$ durch, so erhält man eine sogenannte **Formel**.

Beispielsweise ist

$$\forall x_1\ ((P_1(x_1) \wedge p_1) \rightarrow (\exists x_2\ (P_2(x_2) \wedge p_2) \rightarrow P_3(x_1, x_2)))$$

so eine Formel, in der die beiden auftretenden Gegenstandsvariablen gebunden sind.

Üblicherweise werden Quantoren in mathematischen Ausagen verwendet, um über *Elemente* einer *Menge* (siehe Kapitel 1.2) eine Aussage zu treffen. Man verwendet

$$\forall x \in E : P(x)$$

als Kurzschreibweise für $\forall x\ ((x \in E) \rightarrow P(x))$ und

$$\exists x \in E : P(x)$$

als Kurzschreibweise für $\exists x\ ((x \in E) \wedge P(x))$.

1.1.4 Äquivalente Formeln

Es ist klar, daß die Negation der Aussage

8 *ist eine Primzahl* **und** 8 *ist größer als* 5.

gleichbedeutend mit

8 *ist keine Primzahl* **oder** 8 *ist nicht größer als* 5.

ist, d.h. anstelle von $\neg(p \wedge q)$ kann auch $(\neg p) \vee (\neg q)$ verwendet werden, ohne daß irgend eine Änderung der Aussage eintritt. Genauer bedeutet dies, daß $\neg(p \wedge q)$ immer denselben Wahrheitswert wie $(\neg p) \vee (\neg q)$ hat, und zwar für alle möglichen Wahrheitswertbelegungen von p und q. Das heißt, egal welche *inhaltliche Bedeutung* p und q haben mögen, die Aussage $\neg(p \wedge q)$ ist *immer* gleichbedeutend mit der Aussage $(\neg p) \vee (\neg q)$. In diesem Fall sagt man, daß die beiden Formeln $\neg(p \wedge q)$ und $(\neg p) \vee (\neg q)$ **äquivalent** sind und schreibt dafür

$$\neg(p \wedge q) \Longleftrightarrow (\neg p) \vee (\neg q).^1$$

In ähnlicher Weise läßt sich auch die Äquivalenz $F_1 \Longleftrightarrow F_2$ von allgemeinen Formeln F_1, F_2 definieren.

Im folgenden werden einige dieser Äquivalenzen, die man auch als *logische Regeln* bezeichnen kann, angegeben werden.

Dabei ist zu bemerken, daß bei Formeln, die keine Prädikate enthalten, immer in endlich vielen Schritten mit einer **Wahrheitstafel** überprüft werden kann, ob sie äquivalent sind oder nicht. Bei allgemeinen Formeln ist dies i.a. nicht entscheidbar.

Zur Illustration soll die Wahrheitstafel der Formel $F = p \leftrightarrow (q \rightarrow (\neg q \vee p))$ angegeben werden:

q	p	p	$\leftrightarrow$	$(q$	$\rightarrow$	$(\neg$	q	$\vee$	$p))$
w	**w**		**w**		**w**	**f**		**w**	
w	**f**		**w**		**f**	**f**		**f**	
f	**w**		**w**		**w**	**w**		**w**	
f	**f**		**f**		**w**	**w**		**w**	

In der zweiten Zeile wird etwa jener Fall diskutiert, wo q den Wert **w** und b den Wert **f** annimmt. $\neg a$ hat dann den Wert **f**, $\neg q \vee p$ den Wert **f**, $q \rightarrow (\neg q \vee p)$ den Wert **f** und schließlich $p \leftrightarrow (q \rightarrow (\neg q \vee p))$ den Wert **w**.

[1] Man beachte, daß hier Aussagen über Aussagen gemacht werden, d.h. $\neg(p \wedge q) \leftrightarrow (\neg p) \vee (\neg q)$ wäre eine Aussage, $\neg(p \wedge q) \Longleftrightarrow (\neg p) \vee (\neg q)$ ist aber eine Aussage über Aussagen.

Die folgende Liste enthält einige der wichtigsten Äquivalenzen.

1.	$a \wedge b$	$\iff$	$b \wedge a$,	$a \vee b$	$\iff$	$b \vee a$,
2.	$a \wedge (b \wedge c)$	$\iff$	$(a \wedge b) \wedge c$,	$a \vee (b \vee c)$	$\iff$	$(a \vee b) \vee c$,
3.	$a \wedge (a \vee b)$	$\iff$	a,	$a \vee (a \wedge b)$	$\iff$	a,
4.	$a \wedge (b \vee c)$	$\iff$	$(a \wedge b) \vee (a \wedge c)$,	$a \vee (b \wedge c)$	$\iff$	$(a \vee b) \wedge (a \vee c)$,
5.	$\neg(a \wedge b)$	$\iff$	$\neg a \vee \neg b$,	$\neg(a \vee b)$	$\iff$	$\neg a \wedge \neg b$,
6.	$a \wedge b$	$\iff$	$\neg(a \to \neg b)$,	$a \vee b$	$\iff$	$\neg a \to b$,
7.	$a \leftrightarrow b$	$\iff$	$(a \to b) \wedge (b \to a)$			
		$\iff$	$\neg((a \to b) \to \neg(b \to a))$,			
8.	$a \to b$	$\iff$	$\neg a \vee b$,	$a \leftrightarrow b$	$\iff$	$(a \wedge b) \vee (\neg a \wedge \neg b)$,
9.	$a \to b$	$\iff$	$\neg b \to \neg a$	$\neg(a \to b)$	$\iff$	$a \wedge \neg b$,
10.	$\forall x \forall y\ P(x,y)$	$\iff$	$\forall y \forall x\ P(x,y)$,	$\exists x \exists y\ P(x,y)$	$\iff$	$\exists y \exists x\ P(x,y)$,
11.	$a \wedge \forall x\ P(x)$	$\iff$	$\forall x\ (a \wedge P(x))$,	$a \vee \exists x\ P(x)$	$\iff$	$\exists x\ (a \vee P(x))$,
12.	$a \wedge \exists x\ P(x)$	$\iff$	$\exists x\ (a \wedge P(x))$,	$a \vee \forall x\ P(x)$	$\iff$	$\forall x\ (a \vee P(x))$,
13.	$\neg\,(\forall x\ P(x))$	$\iff$	$\exists x\ \neg P(x)$,	$\neg\,(\exists x\ P(x))$	$\iff$	$\forall x\ \neg P(x)$,

Besonders beachtenswert sind die Regeln 5. und 9. und 13.

5. wird auch DeMorgansche Regel bezeichnet: Die Negation eine Disjunktion ist die Konjunktion der Negationen und umgekehrt (siehe das einleitende Beispiel).

9. ist die logische Grundlage des **indirekten Beweises**. Anstelle der Aussage

Wenn $a^2 \neq b^2$, dann ist $a \neq b$.

kann auch

Wenn $a = b$, dann ist $a^2 = b^2$.

bewiesen werden, d.h. um "*Wenn $a^2 \neq b^2$, dann ist $a \neq b$.*" zu beweisen, nimmt man an, die Aussage "$a \neq b$" sei negiert, also "$a = b$", und folgert daraus die negierte Aussage von "$a^2 \neq b^2$", also "$a^2 = b^2$".

Mit Hilfe von 13. können auch Formeln mit Quantoren negiert werden. Beispielsweise ist die Negation der Formel

$$\forall x \exists y\ (P(x,y) \to Q(x,y))$$

die Formel

$$\exists x \forall y\ (P(x,y) \wedge \neg Q(x,y)).$$

Neben der Äquivalenz von Formeln gibt es auch die **Implikation** $F_1 \implies F_2$,[2] von Formeln F_1, F_2, d.h. F_2 ist sicher wahr, wenn F_1 wahr ist.

1.	$a \wedge b$	$\implies$	a,	a	$\implies$	$a \vee b$,
2.	$a \wedge (a \to b)$	$\implies$	b,			
3.	$\forall x\ P(x)$	$\implies$	$P(x_0)$,	$P(x_0)$	$\implies$	$\exists x\ P(x)$,
4.	$\forall x\ (P(x) \to Q(x)) \wedge P(x_0)$	$\implies$	$Q(x_0)$,			
5.	$\exists x \forall y\ P(x,y)$	$\implies$	$\forall y \exists x\ P(x,y)$,			

[2] Die Implikation von Formeln ist wieder eine Aussage über Aussagen.

2. ist die Abrennungsregel, der sogenannte *modus ponens*.

Das klassische Beispiel zu 4. ist das folgende:

$P(x)$: x ist ein Mensch.
$Q(x)$: x ist sterblich.
x_0: Sokrates.

Weiters beachte man, daß das Analogon zu 5. $\forall y \exists x\ P(x,y) \Longrightarrow \exists x \forall y\ P(x,y)$ *nicht* gilt. (Man betrachte etwa das Beispiel $P(x,y)$: y ist die Mutter von x.)

1.2 Mengen

1.2.1 Der Mengenbegriff

Die Mengenlehre wurde vor etwa 100 Jahren von Georg Cantor begründet. Er benutzte damals die folgende Definition, die zwar streng formal widersprüchlich ist, sich für unsere Anwendungen aber durchaus als zweckmäßig und ausreichend erweist.

Definition 1.1 (Cantor) *Eine* **Menge** *ist eine Zusammenfassung von wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.*

Beispielsweise stellt sich heraus, daß die *Menge aller Mengen, die sich nicht selbst enthalten* $\{M \mid M \notin M\}$, die nach der obigen Definition eine Menge sein müßte, ein widersprüchlicher Begriff ist. Formal wurde dieser Widerspruch dadurch gelöst, daß die Mengenlehre streng axiomatisch aufgebaut wurde (Axiomensystem von Zermelo und Fraenkel mit oder ohne Auswahlaxiom, siehe unten). Noch einfacher ist es, eine genügend große Menge (von Mengen), ein **Universum** E, vorauszusetzen und nur Elemente des Universums zu betrachten. Dadurch können keine Widersprüche wie eben erwähnt entstehen.

Beispiel 1.2 Die Zahlen $1, 2, 3$ bilden eine *endliche* Menge $A = \{1, 2, 3\}$, die Menge der ganzen Zahlen $\mathbb{Z}$ eine *unendliche*.

Definition 1.3 *Die Objekte x, die in einer Menge A zusammengefaßt werden, bezeichnet man als* **Elemente** *der Menge A. Man sagt auch, daß x in A enthalten ist und schreibt $x \in A$. Ist x in A nicht enthalten, so schreibt man dafür $x \notin A$.*

Die Menge $\emptyset$, die keine Elemente enthält, heißt **leere Menge**.

Definition 1.4 *Eine Menge A heißt* **Teilmenge** *einer Menge B, i.Z. $A \subseteq B$, wenn jedes Element x aus A auch in B enthalten ist.*

Definition 1.5 *Zwei Mengen A, B sind gleich, i.Z. $A = B$, wenn sie dieselben Elemente enthalten.*

Satz 1.6 *Zwei Mengen A, B sind genau dann gleich, wenn sowohl A Teilmenge von B als auch B Teilmenge von A ist, d.h.*

$$A = B \Longleftrightarrow A \subseteq B \text{ und } B \subseteq A.$$

Es gibt verschiedene Möglichkeiten, eine Menge anzugeben. Die einfachste Möglichkeit ist die **aufzählende Darstellung**, z.B. $A = \{1, 2, 3\}$, die sich aber nur für endliche (und gelegentlich für abzählbare - siehe Abschnitt 1.5) Mengen eignet. Die häufigste Form ist die **beschreibende Darstellung**

$$A = \{x \,|\, P(x)\},$$

wobei $P(x)$ ein Prädikat ist. Die Menge A enthält nun jene a, für die $P(a)$ den Wert **w** hat, d.h. $a \in A \iff P(a)$. Beispielsweise ist $\{x \,|\, x \in \mathbb{Z} \wedge 1 \leq x \leq 3\}$ die Menge $\{1, 2, 3\}$. Verlangt man, daß die zu beschreibende Menge A Teilmenge einer Menge E sein soll, d.h. $P(x)$ hat die Form $(x \in E) \wedge Q(x)$, so wird anstelle $A = \{x \,|\, x \in E \wedge Q(x)\}$ einfach

$$A = \{x \in E \,|\, Q(x)\}$$

geschrieben.

Die Definition einer Menge impliziert, daß ein Element x nur *einmal* in eine Menge A aufgenommen werden kann, x ist entweder in A enthalten oder nicht in A enthalten. Es ist aber oft sinnvoll *verallgemeinerte Mengen* zu betrachten, wo die Elemente mit einer gewissen Vielfachheit auftreten, z.B. $A = \{1, 1, 2, 2, 2, 3, 4, 4\}$. Solche Objekte werden als **Multimengen** bezeichnet.

Aus Gründen der Vollständigkeit wird auch eine Version des Axiomensystems von Zermelo und Fraenkel mit Auswahlaxiom angegeben:

1. **Extensionalitätsaxiom** Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente enthalten.
2. **Paarmengenaxiom** Zu je zwei Mengen x und y gibt es eine Menge, die genau diese beiden Elemente enthält: $\{x, y\}$.
3. **Aussonderungsschema** Aus jeder Menge kann man die Teilmenge jener Elemente bilden, die eine vorgebene Eigenschaft besitzen.
4. **Vereinigungsmengenaxiom** Zu jeder Menge von Mengen kann man die Vereinigungsmenge bilden.
5. **Potenzmengenaxiom** Zu jeder Menge existiert die Potenzmenge, die Menge aller Teilmengen.
6. **Unendlichkeitsaxiom** Es gibt eine Menge M, welche die leere Menge und mit jeder Menge x auch den sogenannten Nachfolger $x \cup \{x\}$ enthält.
7. **Ersetzungsschema** Ist A eine Menge und ist E eine zweistellige Eigenschaft derart, daß es zu jedem $a \in A$ höchstens ein b mit $E(a, b)$ gibt, dann bilden alle solchen b wieder eine Menge.
8. **Regularitätsaxiom** Jede nichtleere Menge X besitzt eine $x \in X$ mit leerem Schnitt $X \cap x = \emptyset$.
9. **Auswahlaxiom** Zu jeder Menge X nichtleerer Mengen gibt es eine Funktion f, welche jedem $x \in X$ ein $f(x) \in x$ zuordnet.

1.2.2 Operationen mit Mengen

Definition 1.7 *Die* **Vereinigung** *$A \cup B$ zweier Mengen A, B enthält genau jene Elemente x, die in A oder in B enthalten sind, d.h.*

$$A \cup B := \{x \,|\, x \in A \vee x \in B\}.$$

Definition 1.8 *Der* **Durchschnitt** *$A \cap B$ zweier Mengen A, B enthält genau jene Elemente x, die sowohl in A als auch in B enthalten sind, d.h.*

$$A \cap B := \{x \,|\, x \in A \wedge x \in B\}.$$

Es ist oft notwendig, nicht nur zwei oder endlich viele Mengen zu vereinigen, sondern ein ganzes System von Mengen zu vereinigen. Sei I eine Menge (**Indexmenge**, und für jedes $i \in I$ sei A_i eine Menge. Dann bezeichnet man $(A_i)_{i\in I}$ bzw. $(A_i \mid i \in I)$ als **Mengensystem** bzw. als **Mengenfamilie**.[3]

Definition 1.9 *Sei I eine Menge (Indexmenge) und für alle $i \in I$ sei A_i eine Menge. Dann ist duch*

$$\bigcup_{i\in I} A_i = \bigcup (A_i \mid i \in I) := \{x \mid \exists i \in I : x \in A_i\}$$

die **Vereinigung** *aller A_i, $i \in I$, und*

$$\bigcap_{i\in I} A_i = \bigcap (A_i \mid i \in I) := \{x \mid \forall i \in I : x \in A_i\}$$

der **Durchschnitt** *aller A_i, $i \in I$.*

Ist I die Menge $\{0, 1, 2, \ldots, n\}$ bzw. $\{1, 2, \ldots, n\}$ oder die Menge der natürlichen Zahlen $\{0, 1, 2, 3, \ldots\}$, so schreibt man anstelle von $\bigcup_{i\in I} A_i$ resp. $\bigcap_{i\in I} A_i$

$$\bigcup_{i=0}^{n} A_i, \quad \bigcup_{i=1}^{n} A_i, \quad \bigcup_{i=0}^{\infty} A_i, \qquad \text{resp.} \qquad \bigcap_{i=0}^{n} A_i, \quad \bigcap_{i=1}^{n} A_i, \quad \bigcap_{i=0}^{\infty} A_i.$$

Gelegentlich wird auch die Vereinigung bzw. der Durchschnitt von Mengen A gebildet, die eine Eigenschaft $Q(A)$ besitzen ($Q(x)$ ist ein Prädikat):

$$\bigcup \{A \mid Q(A)\} := \{x \mid \exists A \; x \in A \wedge Q(A)\},$$

$$\bigcap \{A \mid Q(A)\} := \{x \mid \forall A \; x \in A \to Q(A)\}.$$

Definition 1.10 *Die* **Mengendifferenz** *$A \setminus B$ zweier Mengen A, B enthält genau jene Elemente x, die in A, aber nicht in B enthalten sind, d.h.*

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

Die **symmetrische Differenz** *$A \triangle B$ zweier Mengen A, B ist durch*

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

gegeben.

Definition 1.11 *Sei $A \subseteq E$. Dann bezeichnet*

$$A' := \{x \in E \mid x \notin A\}$$

das **Komplement** *von A (bezüglich E).*

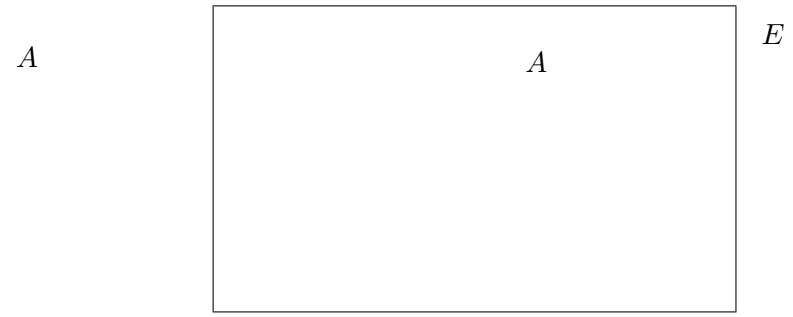
[3] Genaugenommen ist eine Mengenfamilie eine Funktion $A : I \to M$, wobei M eine Menge von Mengen bezeichnet, siehe Abschnitt 1.4.

Satz 1.12 *Seien A, B, C und A_i, $i \in I$, Teilmengen einer Menge E. Dann gelten die folgenden Rechenregeln.*

1.	$A \cup B$	$=$	$B \cup A$,	$A \cap B$	$=$	$B \cap A$,
2.	$A \cup (B \cup C)$	$=$	$(A \cup B) \cup C$,	$A \cap (B \cap C)$	$=$	$(A \cap B) \cap C$,
3.	$A \cap (B \cup C)$	$=$	$(A \cap B) \cup (A \cap C)$,	$A \cup (B \cap C)$	$=$	$(A \cup B) \cap (A \cup C)$,
3.′	$A \cap \bigcup_{i \in I} A_i$	$=$	$\bigcup_{i \in I}(A \cap A_i)$,	$A \cup \bigcap_{i \in I} A_i$	$=$	$\bigcap_{i \in I}(A \cup A_i)$,
4.	$A \cup \emptyset$	$=$	$\emptyset \cup A = A$,	$A \cap E$	$=$	$E \cap A = A$
5.	$A \cup A'$	$=$	E,	$A \cap A'$	$=$	$\emptyset$
6.	$A \cup A$	$=$	A,	$A \cap A$	$=$	A
7.	$A \cup E$	$=$	E,	$A \cap \emptyset$	$=$	$\emptyset$
8.	$A \cup (A \cap B)$	$=$	A,	$A \cap (A \cup B)$	$=$	A
9.	$(A \cup B)'$	$=$	$A' \cap B'$,	$(A \cap B)'$	$=$	$A' \cup B'$
9.′	$\left(\bigcup_{i \in I} A_i\right)'$	$=$	$\bigcap_{i \in I} A_i'$,	$\left(\bigcap_{i \in I} A_i\right)'$	$=$	$\bigcup_{i \in I} A_i'$

1. nennt man auch **Kommutativgesetz**, *2.* **Assoziativgesetz**, *3.* **Distributivgesetz**, *8.* **Verschmelzungsgesetz** *und 9.* **DeMorgansche Regel**.

Es ist oft nützlich, eine Menge A bildlich durch ein sogenanntes **Venndiagramm** darzustellen.



Auf diesem Weg lassen sich die Mengenoperationen Vereinigung, Durchschnitt, Komplement, Mengendifferenz und symmetrische Differenz auf einfache Art graphisch verdeutlichen (siehe p. 11).

1.2.3 Elementtabelle

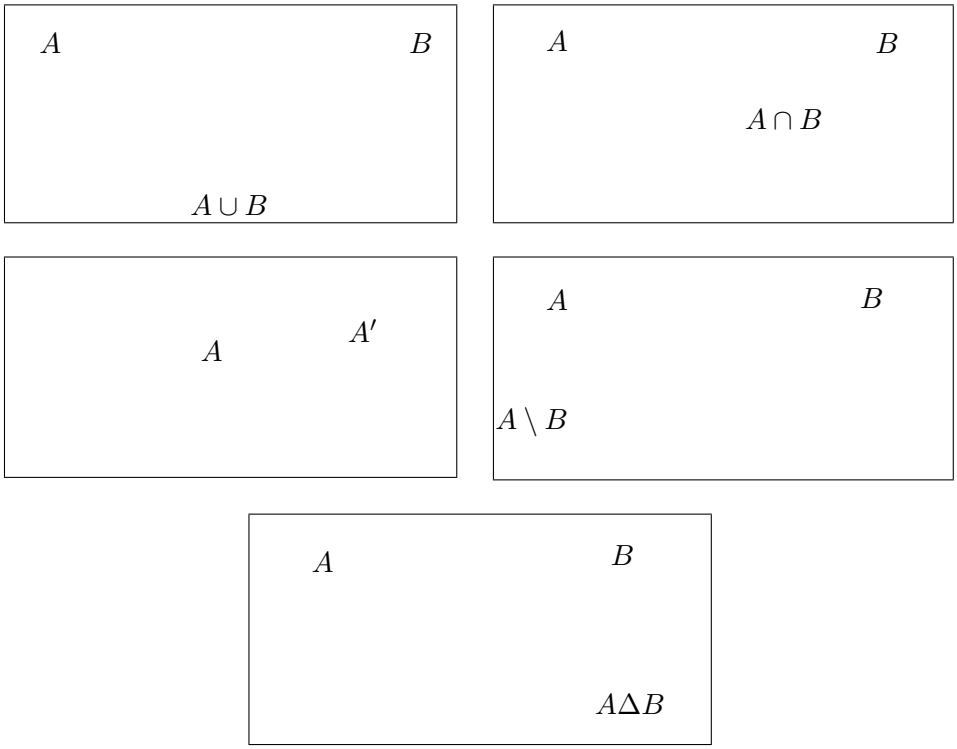
Das einfachste Verfahren zum Überprüfen von Mengenidentitäten ist eine **Elementtabelle**. Dieses Verfahren soll an der folgenden Identität

$$A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$$

demonstriert werden.

Da zwei Mengen nach Definition 1.5 genau dann gleich sind, wenn jedes (potentielle) Element x[4] dann und nur dann in der einen Menge enthalten ist, wenn es in der anderen enthalten ist, genuegt es, bei drei involvierten Mengen A, B, C acht Fälle zu unterscheiden, entsprechend ob ein x in A resp. B resp. C enthalten ist oder nicht.

[4] aus einer Obermenge von $A \cup B \cup C$



Ist z.B. ein x in A und B, aber nicht in C enthalten, so ist es in A und in $B \triangle C$ und damit auch in $A \cap (B \triangle C)$ enthalten. Andererseits ist es in $A \cap B$, aber nicht in $A \cap C$ enthalten, womit es allerdings in $(A \cap B) \triangle (A \cap C)$ enthalten ist. Ein x, das in A und B, aber nicht in C enthalten ist, ist daher sowohl in $A \cap (B \triangle C)$ als auch in $(A \cap B) \triangle (A \cap C)$ enthalten. Die anderen sieben Fälle können ähnlich behandelt werden, und in jedem Fall ist x entweder (wie im eben behandelten Fall) Element von beiden Teilen oder kein Element. Da mit den acht Fällen alle möglichen Situationen abgedeckt sind, müssen $A \cap (B \triangle C)$ und $(A \cap B) \triangle (A \cap C)$ nach Definition 1.5 gleich sein. In einer **Elementtabelle** können alle Fälle übersichtlich dargestellt und so der Nachweis von Mengenidentitäten erbracht werden. Die oben angeführte Überlegung entspricht übrigens der zweiten Zeile.

A	B	C	$B \triangle C$	$A \cap (B \triangle C)$	$A \cap B$	$A \cap C$	$(A \cap B) \triangle (A \cap C)$
$\in$	$\in$	$\in$	$\notin$	$\notin$	$\in$	$\in$	$\notin$
$\in$	$\in$	$\notin$	$\in$	$\in$	$\in$	$\notin$	$\in$
$\in$	$\notin$	$\in$	$\in$	$\in$	$\notin$	$\in$	$\in$
$\in$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$
$\notin$	$\in$	$\in$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$
$\notin$	$\in$	$\notin$	$\in$	$\notin$	$\notin$	$\notin$	$\notin$
$\notin$	$\notin$	$\in$	$\in$	$\notin$	$\notin$	$\notin$	$\notin$
$\notin$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$	$\notin$

Man kann auf einem Blick erkennen, ob Mengengleichheit besteht, die fünfte und die achte Spalte müssen gleich sein. Entsteht in wenigstens einem Fall ein unterschiedliches Bild (d.h. x ist in einem Teil enthalten, im anderen aber nicht), so sind die betrachteten Mengenausdrücke nicht gleich. Es gibt dann Mengen $A, B, C, \ldots$, für die die Identität nicht gilt. (Solche können

auch leicht konstruiert werden.)

Durch eine einfache Modifikation können auch Enthaltenseinsrelationen ($\subseteq$) von Mengenausdrücken überprüft werden. (Ist links ein $\in$–Zeichen, so muß auch rechts eines sein.)

1.2.4 Potenzmenge

Definition 1.13 *Die* **Potenzmenge** $\mathbf{P}(A)$ *einer Menge A ist die Menge aller Teilmengen von A, d.h.*

$$\mathbf{P}(A) = \{C \mid C \subseteq A\}.$$

Beispiel 1.14

$$\begin{aligned} \mathbf{P}(\{1,2\}) &= \{\emptyset, \{1\}, \{2\}, \{1,2\}\}, \\ \mathbf{P}(\emptyset) &= \{\emptyset\} \end{aligned}$$

Manchmal bezeichnet man die Potenzmenge einer Menge A auch als 2^A. Ein Grund dafür ist der folgende Satz. ($|A|$ bezeichnet die Anzahl der Elemente von einer endlichen Menge A - siehe Abschnitt1.5.)

Satz 1.15 *Für eine endliche Menge A gilt*

$$|\mathbf{P}(A)| = 2^{|A|},$$

d.h. eine Menge mit n Elementen hat genau 2^n Teilmengen.

Definition 1.16 *Seien $0 \leq k \leq n$ ganze Zahlen. Der* **Binomialkoeffizient** $\binom{n}{k}$ *(sprich: n über k) ist definiert durch*

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

($0! = 1$ und $n! = 1 \cdot 2 \cdots (n-1) \cdot n$ für $n > 0$.)

Für ganze Zahlen n, k mit $n \geq 0$ und $k < 0$ oder $k > n$ setzt man auch $\binom{n}{k} = 0$. Mit Hilfe dieser Zusatzdefinition gilt der folgende Satz uneingeschränkt und ist Grundlage des **Pascalschen Dreiecks**.

Satz 1.17 *Die Binomialkoeffizienten erfüllen die Rekursion*

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

Satz 1.18 *Eine endliche Menge mit n Elementen hat genau $\binom{n}{k}$ Teilmengen mit k Elementen.*

Man beachte, daß daraus folgt, daß $\binom{n}{k}$ immer eine natürliche Zahl ist, was aus der Definition nicht unmittelbar ersichtlich ist.

Beispiel 1.19 In einer Liga von n Mannschaften müssen genau $\binom{n}{2} = \frac{1}{2}n(n-1)$ Spiele ausgetragen werden, damit jeder gegen jeden gespielt hat.

Korollar 1.20 *Für $n \geq 0$ gilt*

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Eine Erweiterung dieses Korollars ist der sogenannte **Binomische Lehrsatz**

Satz 1.21 *Für $n \geq 0$ und beliebige $x, y \in \mathbb{C}$ gilt*

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n.$$

Korollar 1.20 ergibt sich aus dem Spezialfall $x = y = 1$.

1.2.5 Kartesisches Produkt

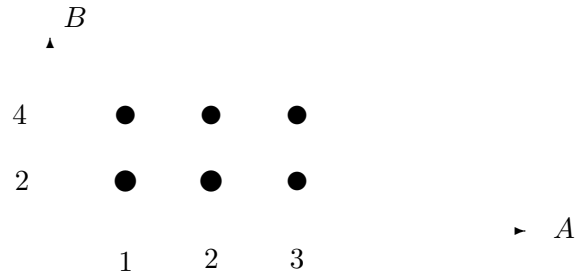
Definition 1.22 *Seien A, B zwei Mengen und $a \in A$, $b \in B$. Dann bezeichnet man durch (a, b) das* **geordnete Paar** *von a und b, wobei zwei geordnete Paare (a, b), (a', b') nur dann als gleich angesehen werden, wenn $a = a'$ und $b = b'$, d.h. die Reihenfolge der Eintragungen ist wesentlich.*[5]

Das **kartesische Produkt** *$A \times B$ zweier Mengen A, B ist die Menge aller geordneter Paare (a, b) mit $a \in A$ und $b \in B$, d.h.*

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

In ähnlicher Weise definiert man auch das Produkt $A_1 \times A_2 \times \cdots \times A_n$ von endlich vielen Mengen $A_1, A_2, \ldots, A_n$ als die Menge aller n–tupel $(a_1, a_2, \ldots, a_n)$ mit $a_j \in A_j$ $(1 \leq j \leq n)$. Sind alle Mengen A_j gleich einer Menge A, so schreibt man statt $A \times A \times \cdots \times A$ auch A^n.

Beispiel 1.23
Sei $A = \{1, 2, 3\}$ und $B = \{2, 4\}$. Dann ist $A \times B = \{(1,2), (1,4), (2,2), (2,4), (3,2), (3,4)\}$. Dies läßt sich auch folgendermaßen *kartesisch* darstellen:



[5] Formal kann dies etwa durch die Festlegung $(a, b) = \{\{a\}, \{a, b\}\}$ geschehen.

Es ist auch möglich das kartesische Produkt von einer Mengenfamilie $(A_i \mid i \in I)$ zu betrachten. Dazu benötigt man aber den Begriff einer Abbildung (siehe Abschnitt 1.4).

Definition 1.24 *Das kartesische Produkt*

$$\prod_{i \in I} A_i := \{(a_i)_{i \in I} \mid \forall i \in I : a_i \in A_i\}$$

einer Mengenfamilie $(A_i \mid i \in I)$ ist das System aller Abbildungen Tupel $(a_i)_{i \in I}$[6] mit der Eigenschaft, daß $a_i \in A_i$ für alle $i \in I$ gilt.

Ist $I = \{1, 2\}$ und $A_1 = A, A_2 = B$, so kann man die Gesamtheit aller Tupel $(a_i)_{i \in \{1,2\}}$ mit $a_1 = a \in A$ und $a_2 = b \in B$ mit der Menge aller geordneter Paare identifizieren.

1.3 Relationen

1.3.1 Grundlegende Begriffe

Um den mengentheoretischen Begriff einer **Relation** zu motivieren, sollen zunächst einige Beispiele angegeben werden.

Beispiel 1.25 Man betrachte eine Gruppe von Personen. Sind a und b zwei Personen dieser Gruppe, so können folgende Situationen eintreten:

- a und b kennen einander,
- a kennt b, aber b kennt nicht a,
- b kennt a, aber a kennt nicht b,
- a und b kennen einander nicht.

Beispiel 1.26 Gewisse Städte können durch Direktflüge voneinander erreicht werden, andere nicht.

Beispiel 1.27 Eine Schachtel (alter) Schrauben kann so sortiert werde, daß jeweils Schrauben gleicher Länge in ein eigenes Fach kommen.

Abstrahiert man von den angegebenen Beispielen, so steht man vor folgender Situation. Zwei Elemente a, b einer Menge stehen miteinander in einer gewissen *Relation* (wobei die Reihenfolge eine Rolle spielen kann) oder eben nicht. Um diesen *Relationsbegriff* mathematisch zu fassen, verwendet man den Begriff des kartesischen Produkts.

[6] Genaugenommen ist ein Tupel $(a_i)_{i \in I}$ eine Abbildung $a : I \to \bigcup(A_i \mid i \in I)$ mit der Eigenschaft, daß $a_i := a(i) \in A_i$ für alle $i \in I$ gilt.

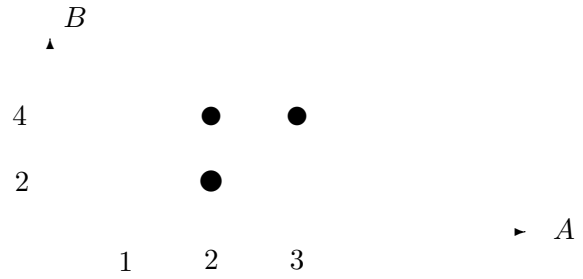
Definition 1.28 *Eine* **Relation** R *zwischen zwei Mengen* A *und* B *ist eine Teilmenge des kartesischen Produkts* $A \times B$.

Ist $A = B$ *so spricht man von einer* **binären Relation** *und bezeichnet sie auch durch* $\langle A, R \rangle$.

Anstelle von $(a, b) \in R$ *schreibt man auch* aRb, *anstelle von* $(a, b) \notin R$ *auch* $a \not R b$.

Die drei einleitenden Beispiele sind übrigens alle binäre Relationen.

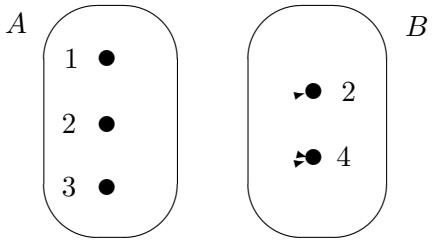
Beispiel 1.29 Sei $A = \{1, 2, 3\}$ und $B = \{2, 4\}$. Dann ist $R = \{(2, 2), (2, 4), (3, 4)\}$ eine Relation zwischen A und B, d.h. $2R2$, $2R4$ und $3R4$, aber $1 \not R 2$, $1 \not R 4$ und $3 \not R 2$. Dies kann natürlich auch graphisch ausgedrückt werden:



Es gibt verschiedene Möglichkeiten, eine Relation $R \subseteq A \times B$ bildlich darzustellen:

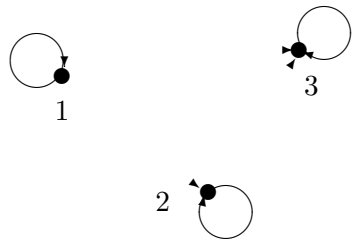
1. **Kartesische Darstellung**: In der kartesischen Darstellung von $A \times B$ (siehe Beispiel 1.23) werden nur nur jene Elemente (= Punkte) von $A \times B$ *markiert*, die der Relation $R \subseteq A \times B$ angehören (siehe Beispiel 1.29).

2. **Pfeildiagramm**: Die Mengen A, B werden durch Venndiagramme dargestellt und Paare $(a, b) \in R$ durch einen Pfeil verbunden:



3. **Graph einer binären Relation**: Da bei einer binären Relation $A = B$ gilt (d.h. $R \subseteq A \times A = A^2$), reicht es, (im Gegensatz zum Pfeildiagramm) die Menge A nur einmal zu repräsentieren. Paare $(a, b) \in A \times A$, die in Relation stehen, werden nun ähnlich wie beim Pfeildiagramm miteinander verbunden. Der **Graph** $G(R)$ besteht daher aus einer Menge von *Punkten (Knoten)* entsprechend den Elementen aus A und einer Mengen von *gerichteten Kanten*, die genau jene Punkte miteinander verbinden, die miteinander in Relation stehen.

   Ist beispielsweise $A = \{1, 2, 3\}$ und $R = \{(1, 1), (1, 2), (1, 3), (2, 2), (2, 3), (3, 3)\}$, so entsteht folgendes Bild:



Man beachte, daß in einem Graphen einer Relation auch sogenannte *Schlingen* auftreten, das sind Kanten, die von $a \in A$ wieder auf a zeigen (a steht mit sich selbst in Relation: aRa).

1.3.2 Äquivalenzrelation

Definition 1.30 *Eine binäre Relation $\langle A, R\rangle$ heißt* **Äquivalenzrelation***, wenn folgende drei Eigenschaften erfüllt sind:*

1. $\forall a \in A : aRa$ **(Reflexivität)**,
2. $\forall a, b \in A : aRb \rightarrow bRa$ **(Symmetrie)**,
3. $\forall a, b, c \in A : (aRb \wedge bRc) \rightarrow aRc$ **(Transitivität)**.

Eine Relation mit der Eigenschaft 1. heißt **reflexiv**, eine mit der Eigenschaft 2. **symmetrisch** und eine mit der Eigenschaft 3. **transitiv**. Eine Äquivalenzrelation ist also reflexiv, symmetrisch und transitiv.

Beispiel 1.31 Gleichheitsrelation: $\langle A, =\rangle$, d.h. jedes Element $a \in A$ steht nur mit sich selbst in Relation.

Beispiel 1.32 Allrelation: $\langle A, A^2\rangle$, d.h. jedes Element $a \in A$ steht mit allen anderen Elementen $b \in A$ in Relation.

Beispiel 1.33 Sei $A = \mathbb{Z}$ und $aRb \Longleftrightarrow a \equiv b \bmod 2$.

In der *kartesischen Darstellung* einer Relation R äußert sich die Reflexivität dadurch, daß die 1. Hauptdiagonale (1. Mediane) in der Relation enthalten ist. Ist die Relation symmetrisch, so ist die kartesische Darstellung symmetrisch zur 1. Hauptdiagonale, d.h. R geht durch Spieglung an der 1. Hauptdiagonale in sich über. Die Transitivität hat in der kartesischen Darstellung keine offensichtliche Entsprechung.

Stellt man eine Relation R als Graph $G(R)$ dar, so entspricht einer reflexiven Relation ein Graph, bei dem jeder Punkt (Knoten) mit einer Schlinge ausgestattet ist. Bei einer symmetrischen Relation treten die Kanten (mit Ausnahme der Schlingen) gepaart auf. Zu einer Kante von a nach b gibt es immer auch eine Gegenkante von b nach a. Ebenso läßt sich die Transitivität sofort übersetzen.

Im Beispiel 1.33 fällt auf, daß durch die Äquivalenzrelation die ganzen Zahlen $\mathbb{Z}$ in zwei Teilmengen zerlegt wird, in die geraden Zahlen und in die ungeraden Zahlen. Alle geraden Zahlen stehen miteinander in Relation, und entsprechend alle ungeraden Zahlen. Sie sind jeweils kongruent modulo 2. Aber eine gerade Zahl steht mit keiner ungeraden Zahl in Relation.

Ein entsprechender Sachverhalt gilt ganz allgemein. Äquivalenzrelationen zerlegen die Grundmenge in sogenannte Äquivalenzklassen. Dies soll nun präzisiert werden.

Definition 1.34 *Ein System von nichtleeren Teilmengen $(A_i \mid i \in I)$ einer Menge A heißt* **Partition** *oder* **Zerlegung** *von A, wenn die A_i $(i \in I)$ paarweise disjunkt sind, d.h.*

$$A_i \cap A_j = \emptyset \quad \text{für } i \neq j,$$

und A die Vereinigung

$$A = \bigcup_{i \in I} A_i$$

ist.

Definition 1.35 *Sei R eine Äquivalenzrelation auf A. Für $a \in A$ heißt die Menge*

$$K(a) := \{b \in A \mid aRb\}$$

die von a erzeugte **Äquivalenzklasse**.

Man beachte, daß wegen der Reflexivität immer $a \in K(a)$ gilt. Weiters gilt die folgende Eigenschaft.

Lemma 1.36 *Sei R eine Äquivalenzrelation auf A. Dann gilt*

$$aRb \Longleftrightarrow K(a) = K(b).$$

Daraus ergibt sich leicht der folgende Zusammenhang zwischen Äquivalenzrelationen und Partitionen.

Satz 1.37 *Sei R eine Äquivalenzrelation auf A. Dann bilden die (verschiedenen) Äquivalenzklassen der Elemente von A eine Partition von A.*

Sei umgekehrt A_i $(i \in I)$ eine Partition von A und bezeichne $C(a)$ $(a \in A)$ jene Teilmenge A_i der Partition mit $a \in A_i$. Definiert man aRb genau für jene $a, b \in A$, für die $C(a) = C(b)$ gilt, so ist R eine Äquivalenzrelation.

1.3.3 Halbordnung

Definition 1.38 *Eine binäre Relation $\langle A, R\rangle$ heißt* **Halbordnung** *oder* **partielle Ordnung**, *wenn folgende drei Eigenschaften erfüllt sind:*

1. $\forall a \in A : aRa$ **(Reflexivität)**,

2. $\forall a, b \in A : (aRb \wedge bRa) \rightarrow a = b$ **(Antisymmetrie)**,

3. $\forall a, b, c \in A : (aRb \wedge bRc) \rightarrow aRc$ **(Transitivität)**.

Eine Relation mit der Eigenschaft 2. heißt antisymmetrisch. Eine Halbordnung ist daher eine reflexive, antisymmetrische und transitive Relation. (Im Zusammenhang mit Halbordnungen wird anstelle von R oft das Symbol $\leq$ verwendet.)

Definition 1.39 *Eine Halbordnung $\langle A, R\rangle$ heißt* **Totalordnung** *oder* **Kette** *oder* **lineare Ordnung**, *wenn für je zwei Elemente $a, b \in A$ entweder aRb oder bRa gilt, d.h. je zwei Elemente sind vergleichbar.*

Beispiel 1.40 $\langle \mathbb{R}, \leq\rangle$ bildet eine Totalordnung (Ordnung der reellen Zahlen).

Beispiel 1.41 $A = \mathbb{N}$ mit $mRn :\Longleftrightarrow m$ teilt n ist eine Halbordnung, aber keine Totalordnung.

$A = \mathbb{Z}$ mit $mRn :\Longleftrightarrow m$ teilt n ist keine Halbordnung, da R auf $\mathbb{Z}$ nicht mehr antisymmetrisch ist (z.B. -2 teilt 2 und umgekehrt, aber $-2 \neq 2$).

Beispiel 1.42 $A = \mathbf{P}(M)$ (Potenzmenge einer Menge M) mit $BRC :\Longleftrightarrow B \subseteq C$ bildet eine Halbordnung, aber für $|M| > 1$ keine Totalordnung.

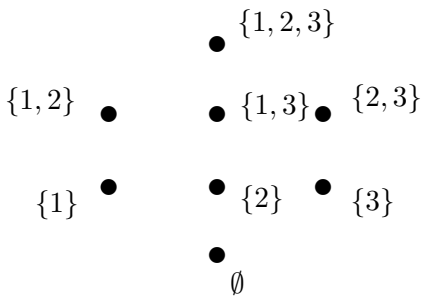
Insbesondere sind alle damit alle Relationen $R \subseteq A \times B$ zwischen zwei (festen) Mengen A, B in natürlicher Weise geordnet.

Wie jede binäre Relation kann man natürlich auch Halbordnungen durch einen Graphen darstellen. Viele der darzustellenden Kanten sind allerdings redundant, sie lassen sich aus den den definierenden Eigenschaften leicht wieder rekonstruieren. Führt man die folgenden drei Schritte durch, so erhält man aus dem Graphen $G(R)$ einer Halbordnung R das **Hassediagramm** von R:

- Weglassen aller Schlingen.
- Weglassen aller Kanten, die sich aufgrund der Transitivitätsbedingung rekonstruieren lassen, d.h. ist aRb, aber gibt es kein c mit aRc und cRb, so bleibt die Kante von a nach b erhalten, allen anderen Kanten werden *gestrichen*. Mit anderen Worten: nur die *unmittelbaren Nachfolger* von a werden von a mit einer Kante verbunden.
- Weglassen aller Orientierungen. Wegen der Antisymmetrie kann für $a \neq b$ entweder aRb oder bRa gelten aber nie beides zugleich. Zur Übersicht zeichnet man bei aRb $(a \neq b)$ b oberhalb von a und kann die Orientierung der Kante weglassen.

Beispiel 1.43 Sei $A = \mathbf{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$ die Potenzmenge von $M = \{1,2,3\}$ mit der Inklusion $(\subseteq)$ als Relation (siehe Beispiel 1.42). Dann hat das Hassediagramm dieser Halbordnung die folgende Gestalt:

Oft wird eine Halbordnung auf einer endlichen Menge nur durch Angabe des Hassediagramms definiert.



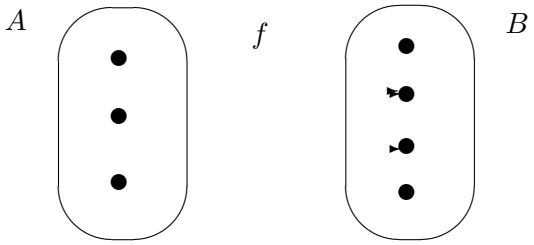
1.4 Funktionen

1.4.1 Begriffsbildung

Der Begriff der **Funktion** oder **Abbildung** ist sicher einer der wichtigsten in der Mathematik. Formal ist eine Funktion Spezialfall einer Relation.

Definition 1.44 *Seien A, B zwei nichtleere Mengen.*

Eine **Funktion** *oder* **Abbildung** *$f : A \to B$ ist eine Relation $R_f \subseteq A \times B$ mit der Eigenschaft, daß zu jedem $a \in A$* genau ein *$b \in B$ mit aR_fb existiert. Man schreibt dafür auch $b = f(a)$. Die Menge A heißt auch* **Wertemenge** *oder* **Definitionsmenge** *und die Menge B* **Zielmenge**



oder **Bildmenge**.

Ist $C \subseteq A$, so wird mit $f(C) = \{f(a) \,|\, a \in C\}$ das **Bild** *von C unter f bezeichnet.*

Entsprechend heißt für $D \subseteq B$ die Menge $f^{-1}(D) = \{a \in A \,|\, f(a) \in D\}$ das **Urbild** *von D unter f.*

Die Menge aller Funktionen $f : A \to B$ wird durch B^A bezeichnet.[7]

Man kann eine Funktion $f : A \to B$ auch als *Zuordnung* oder als *Automat* interpretieren. Einem $a \in A$ wird ein (und nur ein) $b = f(a) \in B$ *zugeordnet* bzw. bei Eingabe von $a \in A$ wird ein $b = f(a) \in B$ *ausgegeben*. Üblicherweise wird eine Funktion $f : A \to B$ nicht als Teilmenge von $A \times B$ definiert, sondern durch Angabe des Bildes $f(a)$ für $a \in A$. Man schreibt dafür auch

$$\begin{aligned} f :\ & A \to B \\ & a \mapsto f(a). \end{aligned}$$

Die Menge $\{(a, f(a)) \,|\, a \in A\} \subseteq A \times B$ wird auch als **Graph** von f bezeichnet.

[7] Bei endlichen Mengen A, B gilt $|A^B| = |A|^{|B|}$.

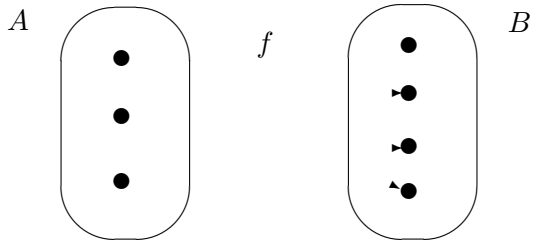
Man beachte, daß zwei Funktionen $f : A \to B$, $g : C \to D$ nur dann gleich sind, wenn $A = C$, $B = D$ und $f(a) = g(a)$ für alle $a \in A$.

Definition 1.45 *Sei $f : A \to B$ eine Funktion und $C \subseteq A$ eine nichtleere Teilmenge von A. Dann bezeichnet man die durch $f|_C(a) = f(a)$ $(a \in C)$ definierte Funktion $f|_C : C \to B$ als* **Einschränkung** *von $f : A \to B$ auf C.*

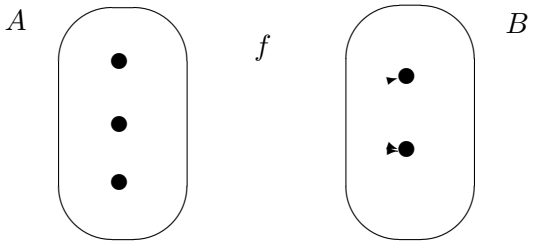
Ist andererseits $A \subseteq D$ und $g : D \to B$ eine Funktion mit $g|_A = f$, so heißt $g : D \to B$ **Fortsetzung** *von $f : A \to B$ auf D.*

1.4.2 Injektive, surjektive und bijektive Funktionen

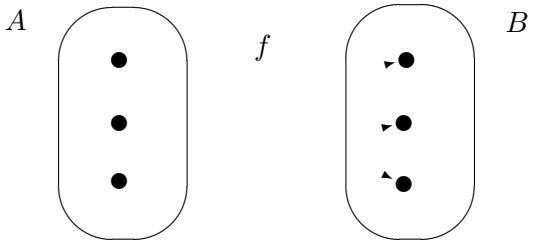
Definition 1.46 *Eine Funktion $f : A \to B$ heißt* **injektiv** *oder* **Injektion**, *wenn es zu jedem $b \in B$* höchstens *ein $a \in A$ mit $b = f(a)$ gibt, d.h. aus $f(a) = f(a')$ folgt $a = a'$.*



Eine Funktion $f : A \to B$ heißt **surjektiv** *oder* **Surjektion**, *wenn es zu jedem $b \in B$* mindestens *ein $a \in A$ mit $b = f(a)$ gibt, d.h. $f(A) = B$.*



Eine Funktion $f : A \to B$ heißt **bijektiv** *oder* **Bijektion**, *wenn es zu jedem $b \in B$* genau *ein $a \in A$ mit $b = f(a)$ gibt.*



Satz 1.47 *Eine Funktion $f : A \to B$ ist genau dann bijektiv, wenn sie injektiv und surjektiv ist.*

Man beachte, daß die Einschränkung $f|_{f^{-1}(B)}$ einer injektive Funktion $f : A \to B$ bijektiv ist.

Definition 1.48 *Sind $f : A \to B$ und $g : B \to C$ Funktionen, so wird die Zusammensetzung $g \circ f : A \to C$ durch $(g \circ f)(a) = g(f(a))$ definiert. $g \circ f$ ist dann wieder eine Funktion.*

Satz 1.49 *Sind die Funktionen $f : A \to B$ und $g : B \to C$ beide injektiv (bzw. surjektiv bzw. bijektiv), so ist auch $g \circ f : A \to C$ injektiv (bzw. surjektiv bzw. bijektiv).*

Definition 1.50 *Eine Funktion $f^{-1} : B \to A$ heißt die zu der Funktion $f : A \to B$* **inverse Funktion***, wenn $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$ sind.*

Dabei bezeichnet id_A die **identische Funktion** *auf einer Menge A, d.h. $\mathrm{id}_A(a) = a$ für alle $a \in A$.*

Satz 1.51 *Eine Funktion $f : A \to B$ besitzt genau dann eine inverse Funktion $f^{-1} : B \to A$, wenn f bijektiv ist. Die inverse Funktion f^{-1} ist dann auch bijektiv.*

1.4.3 Mengenfamilien und kartesische Produkte

Teilweise wurde der Funktionsbegriff in den vorangestellten Abschnitten vorweggenommen, um an passender Stelle eine entsprechend allgemeine Definition angeben zu können.

Im Abschnitt 1.2 wurden Mengenfamilien $(A_i \mid i \in I)$ betrachtet. Genaugenommen ist eine **Mengenfamilie** eine Funktion $A : I \to M$, wobei M eine Menge von Mengen bezeichnet.

Entsprechend benötigt die Definition des **kartesischen Produkts** (siehe Abschnitt 1.2) einer Mengenfamile $(A_i \mid i \in I)$

$$\prod_{i \in I} A_i := \{(a_i)_{i \in I} \mid \forall i \in I : a_i \in A_i\}$$

den Begriff der Funktion, da ein Tupel $(a_i)_{i \in I}$ genaugenommen als Abbildung $a : I \to \bigcup(A_i \mid i \in I)$ zu interpretieren ist. Man beachte, daß das Auswahlaxiom (siehe Abschnitt 1.5) sichert, daß das kartesische Produkt einer Mengenfamilie nichtleer ist.

1.5 Unendliche Mengen

1.5.1 Die natürlichen Zahlen

Die Zahlen $0, 1, 2, 3, \ldots$ heißen **natürliche Zahlen**.[8] Die wesentlichen Eigenschaft der natürlichen Zahlen ist, daß es zu jeder natürlichen Zahl n einen Nachfolger $n' = n + 1$ gibt. Das entspricht dem intuitiven *Immerweiterzählen*. Streng genommen können die natürlichen Zahlen etwa durch die **Peanoaxiome** charakterisiert werden.

1. 0 (Null) ist eine natürliche Zahl.
2. Jede natürliche Zahl n hat genau einen Nachfolger.

[8] Nach ÖNORM ist 0 auch eine natürliche Zahl.

3. 0 ist nicht Nachfolger einer natürlichen Zahl.
4. Verschiedene natürliche Zahlen besitzen verschiedene Nachfolger.
5. Jede Eigenschaft, welche 0 zukommt und sich von jeder natürlichen Zahl auf den Nachfolger überträgt, kommt bereits allen natürlichen Zahlen zu.

Das letzte Axiom heißt auch **Induktionsaxiom**.

Wir wollen die natürlichen Zahlen als Menge $\mathbb{N}$ wiederfinden. Dazu bedient man sich folgender Konstruktion.

$$\begin{aligned} 0 &:= \emptyset, \\ 1 &:= 0 \cup \{0\}, \\ 2 &:= 1 \cup \{1\} = \{0, 1\}, \\ 3 &:= 2 \cup \{2\} = \{0, 1, 2\}, \\ &\vdots \end{aligned}$$

und definiert $\mathbb{N}$ durch

$$\mathbb{N} := \{0, 1, 2, 3, \ldots\}.$$

Genaugenommen muß man $\mathbb{N}$ durch

$$\mathbb{N} := \bigcap \{U \mid (0 \in U) \wedge (\forall x \ (x \in U \to x' \in U))\}$$

definieren, wobei $x' = x \cup \{x\}$ den *Nachfolger* der Menge x bezeichnet. Das Unendlichkeitsaxiom der Mengenlehre garantiert, daß es eine Menge U gibt, die $0 = \emptyset$ enthält und mit jedem x auch den Nachfolger x' enhält. Bildet man den Durchschnitt aller Mengen mit dieser Eigenschaft, so erhält man wieder eine Menge, die die *kleinste* Menge mit dieser Eigenschaft ist, die als $\mathbb{N}$ bezeichnet wird. Es ist leicht nachzuweisen, daß diese Menge $\mathbb{N}$ die Peanoaxiome erfüllt.

Satz 1.52 *Die Menge $\mathbb{N}$ hat folgende Eigenschaften:*

1. $0 \in \mathbb{N}$.
2. $x \in \mathbb{N} \Longrightarrow x' \in \mathbb{N}$.
3. $x' \neq 0$ *für alle* $x \in \mathbb{N}$.
4. $x' = y' \Longrightarrow x = y$ *für alle* $x, y \in \mathbb{N}$.
5. *Ist* $T \subseteq \mathbb{N}$ *mit den Eigenschaften, daß* $0 \in T$ *und* $\forall x \ (x \in T \to x' \in T)$, *so gilt* $T = \mathbb{N}$.

Aus der *Auflistung* $0, 1, 2, 3, \ldots$ der natürlichen Zahlen ergibt sich eine *natürlich Ordnung*. Man sagt m ist kleiner als n, i.Z. $m < n$, wenn m in der *Liste* vor n gereiht ist. Man sagt auch, m ist kleiner oder gleich n, wenn $m < n$ oder $m = n$. Weiters schreibt man auch $m > n$ anstelle von $n < m$ und $m \geq n$ anstelle von $n \leq m$.

Mengentheoretische bedeutet $m \leq n$ nichts anderes als $m \subseteq n$ und $m < n$ kann auch durch $m \in n$ charakterisiert werden.

Mit Hilfe dieser Ordnungsstruktur kann das Induktionsaxiom auch umformuliert werden:

5'. Ist $T \subseteq \mathbb{N}$ mit der Eigenschaft, daß für all $x \in \mathbb{N}$

$$\{y \in \mathbb{N} \mid y < x\} \subseteq T \Longrightarrow x \in T^9$$

gilt, so ist $T = \mathbb{N}$.

[9] Man beachte, daß aus dieser Bedingung sofort folgt, daß $0 \in T$. Man setze $x = 0$.

Das Induktionsaxiom hat das Beweisprinzip der **vollständigen Induktion** zur Folge.

Es sei $P(x)$ ein Prädikat und es ist zu untersuchen, ob $P(n)$ für alle $n \in \mathbb{N}$ wahr ist. Bezeichnet man mit $T \subseteq \mathbb{N}$ jene Teilmenge von $\mathbb{N}$, für die $P(n)$ wahr ist, so erhält man direkt aus den beiden Formulierungen *5.* und *5'.* der mengentheoretischen Formulierungen des Induktionsaxioms die folgenden Schlußregeln:

$$P(0) \wedge \forall n \in \mathbb{N} : (P(n) \to P(n+1)) \Longrightarrow \forall n \in \mathbb{N} : P(n)$$

und

$$\forall n \in \mathbb{N} : ((\forall k < n : P(k)) \to P(n)) \Longrightarrow \forall n \in \mathbb{N} : P(n).$$

Beispiel 1.53 Es sei $P(n)$ die Aussage

$$\sum_{k=0}^{n} k = \frac{n(n+1)}{2}.$$

Offensichtlich ist $P(0)$ wahr, da $\sum_{k=0}^{0} k = 0$. Ist nun $P(n)$ wahr, dann gilt

$$\begin{aligned} \sum_{k=0}^{n+1} k &= \sum_{k=0}^{n} k + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) \\ &= \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Also ist auch $P(n+1)$ wahr. Damit ist gezeigt, daß $P(n)$ für alle $n \in \mathbb{N}$ wahr ist.

Beispiel 1.54 Eine natürliche Zahl $n > 1$ heißt prim oder unzerlegbar, wenn sie nicht als Produkt $n = r \cdot s$ zweier natürlicher Zahlen r, s darstellbar ist, die beide kleiner sind als n.

Für $n > 1$ sei $P(n)$ die Aussage, daß n entweder selbst prim ist oder als Produkt endlich vieler primer Zahlen darstellbar ist. Für den Beweis nehme man an, daß $P(k)$ für alle $k < n$ wahr ist. Wenn n nicht prim ist, dann gibt es natürliche Zahlen $r < n$ und $s < n$ mit $n = r \cdot s$. Unter der eben angeführten Annahme sind $P(r)$ und $P(s)$ wahr. (Man beachte, daß $r > 1$ und $s > 1$ sein müssen.) Folglich kann n auch als Produkt von endlich vielen primen Zahlen dargestellt werden. Daher ist $P(n)$ wahr.

Es wurde damit gezeigt, daß jede natürliche Zahl $n > 1$ eine Primfaktorenzerlegung besitzt. Um zu zeigen, daß diese Zerlegung bis auf die Reihenfolge der auftretenden Primzahlen eindeutig ist, benötigt man noch zusätzliche Überlegungen.

Weitere Beispiele von Mengen sind:

$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ (Menge der **ganzen Zahlen**),
$\mathbb{Q} = \{\frac{m}{n} \mid m \in \mathbb{Z} \wedge n \in \mathbb{Z} \wedge n \neq 0\}$ (Menge der **rationalen Zahlen**),
$\mathbb{R}$ (Menge der **reellen Zahlen**),
$\mathbb{C}$ (Menge der **komplexen Zahlen**).

Bekanntlich gilt

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

All diese Zahlmengen können mit Hilfe der natürlichen Zahlen schrittweise (mengentheoretisch) aufgebaut werden.[10]

An einigen Stellen wird es notwendig sein, aus einer Teilmenge der komplexen Zahlen 0 zu entfernen. Wir definieren daher

$$T^\times := T \setminus \{0\}.$$

Wir werden z.B. $\mathbb{N}^\times, \mathbb{Z}^\times, \mathbb{Q}^\times, \mathbb{R}^\times, \mathbb{C}^\times$ gebrauchen.

1.5.2 Wohlordnungen und transfinite Induktion

Definition 1.55 *Eine Totalordnung $\langle H, \leq\rangle$ heißt* **Wohlordnung**, *wenn jede nichtleere Teilmenge $T \subseteq H$ ein minimales Element besitzt, d.h. es gibt ein $x \in T$ mit $x \leq y$ für alle $y \in T$.*

Beispiel 1.56 $\langle \mathbb{N}, \leq\rangle$ ist eine Wohlordnung,[11] $\langle \mathbb{Z}, \leq\rangle$ ist keine Wohlordnung. Es ist allerdings möglich, $\mathbb{Z}$ auf andere Weise, etwa durch $0 < 1 < 2 < \cdots < -1 < -2 < \cdots$, zu ordnen, so daß $\mathbb{Z}$ wohlgeordnet ist.

Für Wohlordnungen gilt das Prinzip der **transfiniten Induktion**, welches eine Verallgemeinerung der vollständigen Induktion der natürlichen Zahlen ist.

Satz 1.57 *Sei $\langle H, \leq\rangle$ eine Wohlordnung und T eine Teilmenge von H mit der Eigenschaft, daß für alle $x \in H$*

$$\{y \in H \mid y < x\} \subseteq T \Longrightarrow x \in T$$

gilt. Dann ist $T = H$.

1.5.3 Auswahlaxiom

Wie oben schon erwähnt, ist die formale Grundlage der Mengenlehre ein Axiomensystem, das im wesentlichen von Zermelo und Fraenkel stammt. Eine Sonderrolle spielt das **Auswahlaxiom**. Einerseits wurde zunächst gar nicht bemerkt, daß man es eigenlich schon benützt. Es wurde erst nachträglich ergänzt. Andererseits ist es von den übrigen Axiomen unabhängig. Es ist daher durchaus sinnvoll, anstelle des Auwahlaxioms ein anderes zu setzen (auch eines das dem Auswahlaxiom widerspricht) und daraus eine *andere Mathematik* zu entwickeln. Das Auswahlaxiom erscheint zwar in seiner Formulierung einsichtig und naheliegend, es hat aber bei konsequenter Anwendung überraschende und teilweise der Anschauung widersprechende Konsequenzen.[12]

[10] Die genaue Konstruktion der Menge der reellen Zahlen $\mathbb{R}$ gehört zur Vorlesung Analysis. Die komplexen Zahlen können beispielsweise als Paare $z = (a, b)$ reeller Zahlen aufgefaßt werden. Man definiert $(a, b)+(c, d) := (a+c, b+d)$ $(a, b) \cdot (c, d) := (ad + bc, ab - cd)$ und schreibt anstelle von $(a, 0)$ einfach a und anstelle von $(0, 1)$ die *imaginäre Einheit* i. Dann läßt sich $z = (a, b)$ auch durch $z = a + bi$ darstellen.

[11] Ein strenger Beweis dieser Beobachtung folgt aus dem Induktionsaxiom. Man betrachte zu einer beliebigen Teilmenge $T \subseteq \mathbb{N}$ das Prädikat $P(x) := (T \cap \{k \in \mathbb{N} \mid k < x\}$ hat ein minimales Element$) \vee T \cap \{k \in \mathbb{N} \mid k < x\} = \emptyset$. Es folgt sofort, daß$P(n)$ für alle $n \in \mathbb{N}$ wahr ist. Insbesondere ist $P(n_0 + 1)$ wahr für ein $n_0 \in T$, womit $T \cap \{k \in \mathbb{N} \mid k \leq n_0\}$ (und damit auch T) ein minimales Element hat.

[12] Beispielsweise folgt aus dem Auswahlaxiom, daß es möglich ist, einen kleinen Würfel in endlich viele Teile zu zerlegen, die (nach etwaiger Verschiebung und Drehung) die Erdkugel vollständig zusammenzusetzen.

Wir werden folgende Formulierung des Auswahlaxioms verwenden:

Auswahlaxiom *Zu jeder Menge M nichtleerer Mengen gibt es eine Funktion (***Auswahlfunktion***) $f : M \to \bigcup\{A \mid A \in M\}$ mit $f(A) \in A$ für alle $A \in M$.*

Insbesondere ist $f(M) = \{f(A) \mid A \in M\}$ eine Menge, die zu jedem $A \in M$ ein Element $x = f(A) \in A$ enthält. Diese Menge wird auch als **Auswahlmenge** bezeichnet.

Ist $(A_i \mid i \in I)$ eine Mengenfamilie (d.h. eine Abbildung $A : I \to M$, wobei M eine Menge von Mengen bezeichnet), so gibt es auch eine Abbildung $F : I \to M$ mit $F(i) \in A_i$ für alle $i \in I$.[13]

Insbesondere ergibt sich aus dem Auswahlaxiom, daß das kartesische Produkt $\prod_{i \in I} A_i$ einer Mengenfamilie $(A_i \mid i \in I)$ nichtleer ist.

1.5.4 Das Hausdorffsche Maximalitätsprinzip und das Lemma von Zorn

Definition 1.58 *Sei $\langle H, \leq\rangle$ eine Halbordnung.*

Ein Element $x \in H$ heißt **maximal** *(bzw.* **maximales Element***), wenn es kein $y \in H$ mit $x \leq y$ gibt, das von x verschieden ist.*

Ein Element $x \in H$ heißt **minimal** *(bzw.* **minimales Element***), wenn es kein $y \in H$ mit $y \leq x$ gibt, das von x verschieden ist.*

Definition 1.59 *Sei $\langle H, \leq\rangle$ eine Halbordnung und $T \subseteq H$.*

Ein Element $x \in H$ heißt **obere Schranke** *von T, wenn für alle $y \in T$ $y \leq x$ gilt.*

Ein Element $x \in H$ heißt **untere Schranke** *von T, wenn für alle $y \in T$ $x \leq y$ gilt.*

Satz 1.60 (Hausdorffsches Maximalitätsprinzip) *Jede Halbordnung $\langle H, \leq\rangle$ besitzt eine maximale Teilkette, d.h. es gibt eine nichtleere Teilmenge $T \subseteq H$, so daß $\langle T, \leq\rangle$ eine Totalordnung ist, und kein Element $x \in H \setminus T$ ist mit allen Elementen aus T vergleichbar.*

Satz 1.61 (Lemma von Zorn) *Sei $\langle H, \leq\rangle$ eine Halbordnung mit der Eigenschaft, daß jede Teilkette $T \subseteq H$ eine obere Schranke (in H) besitzt. Dann gibt es ein maximales Element $x \in H$.*

Satz 1.62 (Wohlordnungssatz) *Für jede Menge A gibt es eine Relation $\leq$, so daß $\langle A, \leq\rangle$ eine Wohlordnung ist.*

1.5.5 Mächtigkeit von Mengen

Es soll zunächst untersucht werden, wann zwei Mengen als *gleich groß* bezeichnet werden können.

Definition 1.63 *Zwei Mengen A, B heißen* **gleichmächtig**, *wenn es eine bijektive Abbildung $f : A \to B$ gibt. Man schreibt dafür auch $|A| = |B|$ und bezeichnet $|A|$ als die* **Kardinalität** *von A.*

[13] Dazu bezeichne man mit $M' = \{A_i \mid i \in I\}$ und betrachte $F := f \circ A$, wobei $f : M' \to \bigcup\{A_i \mid i \in I\}$ eine Auswahlfunktion bezeichnet.

Insbesondere heißt eine Menge A **endlich**, *wenn es eine natürliche Zahl $n \in \mathbb{N}$ gibt, so daß A mit $\{k \in \mathbb{N} \mid k < n\}$ gleichmächtig ist. In diesem Fall schreibt man auch $|A| = \#A = n$ und bezeichnet $n \in \mathbb{N}$ als die Anzahl der Elemente von A.*

Jede nichtendliche Menge heißt **unendlich**.

Insbesondere ist die Menge $\mathbb{N}$ unendlich.

Definition 1.64 *Eine Menge A heißt* **abzählbar**, *wenn sie gleichmächtig zu den natürlichen Zahlen $\mathbb{N}$ ist.*

Die Kardinalität einer abzählbaren Menge wird mit $\aleph_0$ bezeichnet. ($\aleph$ – sprich "aleph" – ist ein hebräischer Buchstabe.)

Für abzählbare Mengen A gibt es daher eine (bijektive) Funktion $f : \mathbb{N} \to A$, mit anderen Worten, man kann die Elemente von A duch

$$a_0 = f(0), a_1 = f(1), a_2 = f(2), \ldots$$

tatsächlich *abzählen*. Jedes Element aus A wird in dieser Liste genau einmal aufgenommen.

Manchmal werden endliche Mengen auch als abzählbar bezeichnet. In diesem Fall bezeichnet man unendlichen Mengen, die abzählbar sind, auch als **abzählbar unendlich**.

Satz 1.65 *Jede unendliche Teilmenge einer abzählbaren Menge ist abzählbar.*

Satz 1.66 *Die Mengen $\mathbb{Z}, \mathbb{Q}$ sind abzählbar.*

Als nächstes soll untersucht werden, wann eine Menge als größer (oder kleiner) bezeichnet werden kann als eine andere.

Definition 1.67 *Seien A, B zwei Mengen.*

Gibt es eine injektive Abbildung $f : A \to B$, so schreibt man für die Kardinalitäten $|A| \leq |B|$.
Gibt es eine surjektive Abbildung $f : A \to B$, so bezeichnet man dies durch $|A| \geq |B|$.

Im folgenden wird auch $|A| < |B|$ als Abkürzung für $(|A| \leq |B|) \wedge \neg(|A| = |B|)$ geschrieben.

Satz 1.68 *Für zwei Mengen A, B gelten die folgenden Eigenschaften:*

1. $A \subseteq B \Longrightarrow |A| \leq |B|$.
2. $|A| \leq |B| \wedge |B| \leq |C| \Longrightarrow |A| \leq |C|$.
3. $|A| \leq |B| \Longleftrightarrow |B| \geq |A|$.
4. $|A| \leq |B| \vee |B| \leq |A|$.
5. $|A| \leq |B| \wedge |B| \leq |A| \Longleftrightarrow |A| = |B|$.

Satz 1.69 *Für eine unendliche Menge A gilt immer $|A| \geq \aleph_0$.*

Definition 1.70 *Eine unendlichen Menge B, die nicht abzählbar ist, heißt* **überabzählbar**, *d.h. jede injektive Funktion $f : \mathbb{N} \to B$ ist nicht surjektiv.*

Satz 1.71 *Für eine beliebige Menge A gilt*

$$|A| < |\mathbf{P}(A)|.$$

Startet man beispielsweise mit den natürlichen Zahlen $\mathbb{N}$, so folgt aus Satz 1.71, daß $\mathbf{P}(\mathbb{N})$ (die Menge aller Teilmengen von $\mathbb{N}$) überabzählbar ist. (Übrigens sind $\mathbf{P}(\mathbb{N})$ und $\mathbb{R}$ gleichmächtig, also $\mathbb{R}$ ist überabzählbar.[14]) Weiters ist $\mathbf{P}(\mathbf{P}(\mathbb{N}))$ mächtiger als $\mathbf{P}(\mathbb{N})$ usw. Es gibt also unendlich viele *Unendlichkeitsstufen*.

Anders als bei der Potzenzmenge gibt es bei kartesischen Produkten und Vereinigungen keinen Sprung in der Mächtigkeit.

Satz 1.72 *Seien A, B zwei Mengen mit $|A| \leq |B|$, wobei B unendlich ist. Dann gelten die folgenden Eigenschaften:*

1. $|A \cup B| = |B|$.
2. $|A \times B| = |B|$.
3. $\left|\bigcup_{n \geq 1} B^n\right| = |B|$.
4. $|\{T \subseteq B \,|\, |T| < \aleph_0\}| = |B|$.

Abschließend sei noch eine einfache, aber sehr nützliche Eigenschaft für endliche Mengen angeführt.

Satz 1.73 *Haben zwei endliche Mengen A, B gleich viele Elemente, d.h. $|A| = |B|$, dann ist eine injektive Funktion $f : A \to B$ auch surjektiv und damit bijektiv. Entsprechend ist eine surjektive Funktion $f : A \to B$ auch injektiv und ebenfalls bijektiv.*

[14] Die Kontinuumshypothese besagt, daß es keine Menge A mit $|\mathbb{N}| < |A| < \mathbb{R}$ gibt. Diese Aussage kann aber im Rahmen der Mengenlehre weder bewiesen noch widerlegt werden. Sie ist davon unabhängig und könnte als zusätzliches Axiom aufgenommen werden.

Kapitel 2

Algebraische Grundlagen

2.1 Gruppen

2.1.1 Binäre Operationen und Gruppen

Definition 2.1 *Sei A eine nichtleere Menge. Eine* **binäre Operation** $\circ$ *auf A ist eine Abbildung $A \times A \to A$, d.h. je zwei Elementen $a, b \in A$ wird ein Element $a \circ b$ zugeordnet.*

Ein Paar $\langle A, \circ \rangle$ heißt **algebraische Struktur** *oder* **Gruppoid**, *wenn A eine nichtleere Menge ist und $\circ$ eine binäre Operation auf A ist.*

Definition 2.2 *Sei $\langle A, \circ \rangle$ eine algebraische Struktur. Dabei werden folgende Gesetze von $\langle A, \circ \rangle$ definiert.*

(1) **Assoziativgesetz***:*

$$\forall a, b, c \in A : \ (a \circ b) \circ c = a \circ (b \circ c).$$

(2) Existenz eines **neutrales Elements** *e:*

$$\exists e \in A \, \forall a \in A : \ e \circ a = a \circ e = a$$

(3) Existenz **inverser Elemente** *a':*

$$\forall a \in A \, \exists a' \in A : \ a \circ a' = a' \circ a = e,$$

(e bezeichnet das neutrale Element aus (2).)

(4) **Kommutativgesetz***:*

$$\forall a, b \in A : \ a \circ b = b \circ a.$$

Benützt man für das binäre Operationssymbol das *Malzeichen* $\cdot$, so schreibt man für das inverse Element von a auch a^{-1}, verwendet man hingegen das *Pluszeichen* $+$, so bezeichnet man das inverse Element von a durch $-a$.

Satz 2.3 *In einer algebraischen Struktur $\langle A, \circ\rangle$ gibt es höchstens ein neutrales Element, und zu jedem $a \in A$ gibt es höchstens ein inverses Element.*

Es wird daher im folgenden nur mehr vom neutralen Element bzw. vom inversen Element gesprochen werden, sofern diese existieren.

Definition 2.4 *Eine algebraische Struktur $\langle A, \circ\rangle$ heißt*

- **Halbgruppe**, *wenn sie (1) erfüllt,*
- **Monoid**, *wenn sie (1) und (2) erfüllt, und*
- **Gruppe**, *wenn sie (1), (2) und (3) erfüllt.*[1]

Erfüllt eine der Strukturen Gruppoid, Halbgruppe, Monoid bzw. Gruppe auch (4), so heißen sie auch **kommutative(s)** *Gruppoid, Halbgruppe, Monoid bzw. Gruppe.*

Kommutative Gruppen werden auch als **abelsche Gruppen** (im Andenken an den Mathematiker Niels Henrik Abel) bezeichnet

Beispiel 2.5 $A = \mathbb{N}$ mit $a \circ b = a^b$ ist nur ein Gruppoid.

Beispiel 2.6 $\langle \mathbb{N}^\times, +\rangle$ ist eine Halbgruppe, $\langle \mathbb{N}, +\rangle$ und $\langle \mathbb{N}, \cdot\rangle$ sind Monoide, aber keine Gruppen.

Beispiel 2.7 Sei Σ eine endliche Menge, das *Alphabet* und bezeichne Σ^* die Menge aller endlichen *Wörter* über Σ, das sind alle endlichen Folgen $x_1x_2\ldots x_k$ mit $x_j \in \Sigma$ $(1 \leq j \leq k)$, wobei auch das *leere Wort* ε darin enthalten ist. Sind $w_1 = x_{11}x_{12}\ldots x_{1k}$ und $w_2 = x_{21}x_{22}\ldots x_{2l}$ zwei Wörter in Σ^* so definiert man

$$w_1 \circ w_2 = x_{11}x_{12}\ldots x_{1k}x_{21}x_{22}\ldots x_{2l} \in \Sigma^*.$$

$\langle \Sigma^*, \circ\rangle$ ist damit ein Monoid mit neutralem Element ε. Man bezeichnet Σ^* auch als **freies Monoid** über dem Alphabet Σ.

Beispiel 2.8 $\langle \mathbb{Z}, +\rangle$, $\langle \mathbb{Q}, +\rangle$, $\langle \mathbb{Q}^\times, \cdot\rangle$,$\langle \mathbb{R}, +\rangle$, $\langle \mathbb{R}^\times, \cdot\rangle$ etc. sind abelsche Gruppen.

Beispiel 2.9 Sei M eine beliebige Menge. Dann bildet $\langle \mathbf{P}(M), \Delta\rangle$, d.h. die Teilmengen von M mit der symmetrischen Mengendifferenz eine Gruppe. Das neutrale Element ist $\emptyset$ und jedes Element ist selbstinvers.

Beispiel 2.10 Die Menge aller bijektiven Abbildungen $\sigma : M \to M$ auf einer nichtleeren Menge M nennt man Permutationen $\mathbf{S}(M)$ von M. Sie bilden bezüglich der Hintereinanderausführung die sogenannte **symmetrische Gruppe** von M.

[1] Die Gruppenaxiome können dahingehend abgeschwächt werden, daß neben der Assoziativität nur verlangt wird, daß es ein linksneutrales Element e_l gibt, d.h. $e_l \circ a = a$ für alle $a \in A$, und für jedes Element $a \in A$ ein linksinverses Element a'_l, das $a'_l \circ a = e_l$ erfüllt. Aus diesen Eigenschaften kann leicht abgeleitet werden, daß das linksneutrale Element auch rechtsneutral, also neutral, ist und daß das linksinverse Element auch rechtsinvers, also invers, ist.

Beispiel 2.11 Die **Symmetriegruppe** eines gleichseitigen Dreiecks besteht aus allen Isometrien der Ebene, die ein gleichseitiges Dreieck auf sich selbst abbilden. Da ein Dreieck durch seine Eckpunkte eindeutig gegeben ist, reicht es aus, die Auswirkung solcher Isometrien auf die Eckpunkte zu betrachten. Es entstehen gewisse Permuationen der Eckpunkte $\{1, 2, 3\}$. Bei den Drehungen um 0°, 120° und 240° werden die Eckpunkte zyklisch vertauscht, und bei den Spiegelungen an den drei Höhen werden jeweils zwei Eckpunkte miteinander vertauscht. Insgesamt erhält man also sechs verschiedene Symmetrien, die bezüglich Hintereinanderausführung eine Gruppe bilden. In diesem speziellen Fall eines gleichseitigen Dreicks ist die Symmetriegruppe nichts anderes als die symmetrische Gruppe auf den drei Eckpunkten.

Beispiel 2.12 Kleine algebraische Strukturen kann man auch durch sogenannte **Operationstafeln** definieren. Um dies zu demonstrieren, werden alle Möglichkeiten von *kleinen Gruppen* mit ≤ 5 Elementen aufgelistet. (e bezeichnet immer das neutrale Element.)

$\circ$	e
e	e

$\circ$	e	a
e	e	a
a	a	e

$\circ$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

$\circ$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

$\circ$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

$\circ$	e	a	b	c	d
e	e	a	b	c	d
a	a	b	c	d	e
b	b	c	d	e	a
c	c	d	e	a	b
d	d	e	a	b	c

2.1.2 Untergruppen

Definition 2.13 *Eine (nichtleere) Teilmenge $U \subseteq G$ einer Gruppe $\langle G, \circ\rangle$ heißt* **Untergruppe** *von G, wenn $\langle U, \circ\rangle$ selbst ein Gruppe ist,*[2] *i.Z. $\langle U, \circ\rangle \leq \langle G, \circ\rangle$ oder nur $U \leq G$.*

Die Teilmengen $U_1 = \{e\}$ und $U_2 = G$ von G bilden immer Untergruppen, die sogenannten **trivialen Untergruppen** von G.

Satz 2.14 *Sei $\langle G, \circ\rangle$ Gruppe und U nichtleere Teilmenge von G. Dann sind die folgenden drei Bedingungen äquivalent:*

(i) $U \leq G$.

(ii) $\forall a, b \in U : a \circ b \in U \wedge a' \in U$.

(iii) $\forall a, b \in U : a \circ b' \in U$.

[2] Genaugenommen müßte die Operation $\circ : G \times G \to G$ auf $H \times H$ eingeschränkt werden. Es ist aber üblich (und praktisch), das gleiche Operationssymbol auch für diese Einschränkung zu verwenden.

Beispiel 2.15 Für $m \in \mathbb{N}$ bilden die Mengen $m\mathbb{Z} = \{0, \pm m, \pm 2m, \pm 3m, \ldots\}$ Untergruppen bezüglich der Addition.

Definition 2.16 *Sei $\langle G, \circ \rangle$ Gruppe, U Untergruppe von G und $a \in G$. Dann heißt*

$$a \circ U = \{a \circ u \,|\, u \in U\}$$

Linksnebenklasse *von U in G und*

$$U \circ a = \{u \circ a \,|\, u \in U\}$$

Rechtsnebenklasse *von U in G.*

Lemma 2.17 *Sei $\langle G, \circ \rangle$ Gruppe und $U \leq G$. Dann gilt*

$$b \in a \circ U \Longleftrightarrow b \circ U = a \circ U.$$

Satz 2.18 *Sei $\langle G, \circ \rangle$ Gruppe und $U \leq G$. Dann bildet die Menge der Linksnebenklassen $a \circ U$ ($a \in G$) eine Partition von G. Die Relation $a \sim b :\Longleftrightarrow a \circ U = b \circ U$ ist die entsprechende Äquivalenzrelation.*

Eine entsprechende Aussage gilt für die Rechtsnebenklassen $U \circ a$.

Satz 2.19 *Sei $\langle G, \circ \rangle$ Gruppe und $U \leq G$. Dann sind alle Links- und Rechtsnebenklassen gleichmächtig, d.h. für alle $a \in G$ gilt $|a \circ U| = |U \circ a| = |U|$*

Korollar 2.20 *Sei $\langle G, \circ \rangle$ endliche Gruppe und $U \leq G$. Dann stimmt die Anzahl der Linksnebenklassen von U in G mit der Anzahl der Rechtsnebenklassen überein. Diese Anzahl ist durch $|G|/|U|$ gegeben.*

Definition 2.21 *Sei $\langle G, \circ \rangle$ endliche Gruppe und $U \leq G$. Die Anzahl der Links- bzw. Rechtsnebenklassen von wird als* **Index** *$|G : U| = |G|/|U|$ von G nach U bezeichnet.*

Die Anzahl der Elemente $|G|$ einer Gruppe wird als **Ordnung** *von G bzw. als* **Gruppenordnung** *bezeichnet.*

Satz 2.22 (Satz von Lagrange) *Ist $\langle G, \circ \rangle$ endliche Gruppe so ist die Ordnung $|U|$ einer Untergruppe $U \leq G$ stets Teiler der Gruppenordnung $|G|$.*

Definition 2.23 *Sei $\langle G, \circ \rangle$ Gruppe mit neutralem Element e. Für $a \in G$ werden die* **Potenzen** *a^n von a mit $n \in \mathbb{Z}$ folgendermaßen definiert:*

$$a^n = \begin{cases} e & \text{für } n = 0, \\ a & \text{für } n = 1, \\ a^{n-1} \circ a & \text{rekursiv für } n > 1, \\ (a')^{-n} & \text{für } n < 0. \end{cases}$$

Ist das Operationssymbol $+$, so schreibt man statt a^n auch na, z.B. $3a$ für $a + a + a$.

Lemma 2.24 *Sei $\langle G, \circ \rangle$ Gruppe und $a \in G$. Dann gilt für alle $n, m \in \mathbb{Z}$*

$$a^{n+m} = a^n \circ a^m.$$

Weiters sind enweder alle Potenzen a^n $(n \in \mathbb{Z})$ voneinander verschieden oder es gibt ein $n \in \mathbb{N}$ mit $a^n = e$.

Definition 2.25 *Sei $\langle G, \circ \rangle$ Gruppe und $a \in G$. Sind alle Potenzen a^n $(n \in \mathbb{Z})$ voneinander verschieden, so hat a* **unendliche Ordnung** $\mathrm{ord}_G(a) = \infty$. *Andernfalls bezeichnet man*

$$\mathrm{ord}_G(a) = \min\{n \in \mathbb{N}^\times \mid a^n = e\}$$

als **Ordnung** *von a. a hat dann* **endliche Ordnung**.

Satz 2.26 *Ist H_i $(i \in I)$ ein System von Untergruppen einer Gruppe G, so ist*

$$H = \bigcap_{i \in I} H_i$$

wieder eine Untergruppe von G.

Definition 2.27 *Sei G eine Gruppe und $K \subseteq G$ eine nichtleere Teilmenge von G. Die von K* **erzeugte Untergruppe** $[K]$ *ist der Durchschnitt aller Unterguppen $H \leq G$, die K enthalten:*

$$[K] := \bigcap \{U \leq G \mid K \subseteq U\}.$$

$[K]$ ist wegen Satz 2.26 immer eine Untergruppe von G. Bei einelementigen Mengen $K = \{a\}$ ergibt sich folgendes Bild:

Satz 2.28 *Hat $a \in G$ unendliche Ordnung, so ist*

$$[a] := [\{a\}] = \{a^n \mid n \in \mathbb{Z}\}$$

die von a erzeugte Untergruppe, bei endlicher Ordnung $\mathrm{ord}_G(a)$ *ist*

$$[a] := [\{a\}] = \{a^n \mid 0 \leq n < \mathrm{ord}_G(a)\}$$

die von a erzeugt Untergruppe.

Man beachte, daß $[a]$ in jedem Fall eine Untergruppe von G bildet und daß $|[a]| = \mathrm{ord}_G(a)$ gilt.

Satz 2.29 *Sei $\langle G, \circ \rangle$ endliche Gruppe und $a \in G$. Dann hat a endliche Ordnung und es* $\mathrm{ord}_G(a)$ *ist ein Teiler der Gruppenordnung $|G|$.*

Satz 2.30 (Allgemeine Version des Kleinen Fermatschen Satzes) *Für jedes Element $a \in G$ einer endlichen Gruppe $\langle G, \circ \rangle$ gilt $a^{|G|} = e$.*

Definition 2.31 *Eine Gruppe $\langle G, \circ \rangle$ heißt* **zyklisch**, *wenn es ein $a \in G$ mit $G = [a]$ gibt.*

Unendliche zyklische Gruppen haben daher die Form $G = \{a^n \mid n \in \mathbb{Z}\}$ und endliche die Gestalt $G = \{a^n \mid 0 \leq n < |G|\}$, d.h. es gibt ein $a \in G$ mit $\mathrm{ord}_G(a) = |G|$. Weiters beachte man, daß eine zyklische Gruppe immer abelsch ist.

Satz 2.32 *Ist $\langle G, \circ \rangle$ eine endliche Gruppe mit Primzahlordnung, d.h. $|G|$ ist eine Primzahl, so ist G zyklisch (und daher abelsch), und es gibt keine nichttrivialen Untergruppen.*

2.1.3 Produkte von Gruppen

Sind $\langle G_1, \circ\rangle$ und $\langle G_2, \star\rangle$ zwei Gruppen, so kann auch das kartesische Produkt $G_1 \times G_2$ in natürlicher Weise zu einer Gruppe gemacht werden. Definiert man

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 \circ a_2, b_1 \star b_2) \quad (a_1, a_2 \in G_1,\ b_1, b_2 \in G_2),$$

so ist $\langle G_1 \times G_2, \cdot\rangle$ wieder eine Gruppe, das **Produkt** der Gruppen G_1, G_2. Dies motiviert die folgende allgemeine Definition.

Definition 2.33 *Ist $\langle G_i, \circ_i\rangle$ $(i \in I)$ ein System von Gruppen, dann wird das kartesische Produkt $\prod_{i\in I} G_i$ mit der Operation*

$$(a_i)_{i\in I} \cdot (b_i)_{i\in I} := (a_i \circ_i b_i)_{i\in I}$$

zu einer Gruppe, dem **direkten Produkt** *der Gruppen G_i $(i \in I)$.*

Die Teilmenge

$$\prod_{i\in I}^{*} G_i := \{(a_i)_{i\in I} \,|\, |\{i \in I \,|\, a_i \neq e_i\}| < \aleph_0\}$$

bildet eine Untergruppe von $\prod_{i\in I} G_i$ und wird **semidirektes Produkt** *der Gruppen G_i $(i \in I)$ genannt. ($e_i \in G_i$ bezeichnet das neutrale Element von G_i.)*

Das semidirekte Produkt stimmt für endliche Indexmengen I mit dem direkten Produkt überein.

2.1.4 Normalteiler

Definition 2.34 *Eine Untergruppe H einer Gruppe G heißt* **Normalteiler**, *i.Z. $H \trianglelefteq G$, wenn die Links- und Rechtsnebenklassen übereinstimmen.*

Offensichtlich ist jede Untergruppe H einer kommutativen Gruppe G ein Normalteiler.

Weiters ist jede Untergruppe H mit Index $|G : H| = 2$ Normalteiler, da es in diesem Fall nur zwei Links- bzw. Rechtnebenklassen gibt. Die eine ist $e \circ H = H \circ e = H$ und die andere $G \setminus H$.

Satz 2.35 *Für eine Untergruppe H einer Gruppe G sind folgende drei Eigenschaften äquivalent:*

(i) $H \trianglelefteq G$.

(ii) $\forall a \in G : a \circ H = H \circ a$.

(iii) $\forall a \in G : a \circ H \circ a' \subseteq H$.

Lemma 2.36 *Sei H Normalteiler einer Gruppe G. Dann folgt aus $a_1 \circ H = a_2 \circ H$ und $b_1 \circ H = b_2 \circ H$ auch $(a_1 \circ b_1) \circ H = (a_2 \circ b_2) \circ H$.*

Mit Hilfe dieser Eigenschaft von Normalteilern kann auch auf der Menge der Nebenklassen eine Gruppenoperation definiert werden.

Definition 2.37 *Sei H Normalteiler einer Gruppe G und bezeichne G/H die Menge der Nebenklassen von G nach H. Dann wird duch die Operation*

$$(a \circ H) \circ (b \circ H) := (a \circ b) \circ H$$

eine Gruppenoperation auf G/H definiert. Die Gruppe $\langle G/H, \circ \rangle$ heißt **Faktorgruppe** *von G nach H.*

Es ist bei Faktorgruppen G/H üblich, dasselbe Operationszeichen (hier $\circ$) zu verwenden wie bei der ursprünglichen Gruppe G, da bei der Deutung als *Komplexprodukt*

$$(a \circ H) \circ (b \circ H) = \{c \circ d \,|\, c \in a \circ H, d \in b \circ H\}$$

tatsächlich (wegen der Normalteilereigenschaft)

$$\begin{aligned}(a \circ H) \circ (b \circ H) &= (H \circ a) \circ (b \circ H) \\ &= (H \circ (a \circ b)) \circ H \\ &= (a \circ b) \circ (H \circ H) \\ &= (a \circ b) \circ H\end{aligned}$$

dasselbe Ergebnis erhalten wird. Die Gruppenstruktur von G/H ist daher natürlich. Ist H kein Normalteiler von G, so ist das Komplexprodukt von $a \circ H$ und $b \circ H$ i.a. keine Linksnebenklasse.

Beispiel 2.38 Sei $G = \mathbb{Z}$ (mit der Addition $+$) und $H = m\mathbb{Z}$ (mit $m \in \mathbb{N}$). Dann besteht $\mathbb{Z}/m\mathbb{Z} = \mathbb{Z}_m$ aus m Nebenklassen $\overline{0} = 0 + m\mathbb{Z} = m\mathbb{Z}$, $\overline{1} = 1 + m\mathbb{Z}$, …, $\overline{m-1} = (m-1) + m\mathbb{Z}$, den sogenannten **Restklassen modulo** m. $\mathbb{Z}_m$ ist übrigens eine zyklische Gruppe, sie wird etwa von $\overline{1} = 1 + m\mathbb{Z}$ erzeugt.

2.1.5 Gruppenhomomorphismen

Definition 2.39 *Eine Abbildung $\varphi : G \to H$ zwischen zwei Gruppen $\langle G, \circ \rangle$ und $\langle H, \star \rangle$ heißt* **Homomorphismus** *oder* **Gruppenhomomorphismus**, *wenn für alle $a, b \in G$*

$$\varphi(a \circ b) = \varphi(a) \star \varphi(b)$$

gilt. Die Menge aller Gruppenhomomorphismen $\varphi : G \to H$ wird durch $\mathrm{Hom}(G, H)$ bezeichnet.

Ist ein Gruppenhomomorphismus φ injektiv, so heißt φ auch **Monomorphismus**, *und ist φ surjektiv, so nennt man ihn* **Epimorphismus**.

Ist φ bijektiv, so heißt er **Isomorphismus**. *Die inverse Abbildung $\varphi^{-1} : H \to G$ ist dann auch ein Isomorphismus. Existiert zwischen zwei Gruppen G, H ein Isomorphismus, so heißen G und H* **isomoph** *und man schreibt dafür $G \cong H$.*

Ein Homomorphismus $\varphi : G \to G$ heißt **Endomorphismus** *und ein Isomorphismus $\varphi : G \to G$* **Automorphismus**. *Die entsprechenden Mengen von Abbildungen werden mit $\mathrm{End}(G)$ und $\mathrm{Aut}(G)$ bezeichnet.*

Die Automorphisem $\mathrm{Aut}(G)$ bilden bezüglich der Hintereinanderausführung eine Gruppe, die sogenannte **Automorphismengruppe**.

Lemma 2.40 *Ist $\varphi : G \to H$ ein Gruppenhomomorphismus, so wird das neutrale Element e_G von G auf das neutrale Element e_H von H abgebildet, d.h. $\varphi(e_G) = e_H$. Weiters gilt $\varphi(a') = \varphi(a)'$ für alle $a \in G$.*

Definition 2.41 *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Das Urbild $\varphi^{-1}(\{e_H\})$ des neutralen Elements e_H wird als* **Kern** *von φ*

$$\text{kern}(\varphi) := \{a \in G \,|\, \varphi(a) = e_H\}$$

bezeichnet.

Weiters nennt man

$$\text{im}(\varphi) = \varphi(G) := \{b \in H \,|\, \exists a \in A : \varphi(a) = b\}$$

Bild *von G unter φ.*

Satz 2.42 *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann ist kern(φ) ein Normalteiler von G und $\varphi(G)$ eine Untergruppe von H.*

Einer der wichtigsten Sätze der Gruppentheorie ist der **Homomorphiesatz**.

Satz 2.43 *Sei $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann ist die Faktorgruppe $G/\text{kern}(\varphi)$ mit $\varphi(G)$ isomorph:*

$$G/\text{kern}(\varphi) \cong \varphi(G)$$

Die Nebenklasse $a \circ \text{kern}(\varphi) \in G/\text{kern}(\varphi)$ entspricht dem Element $\varphi(a) \in \varphi(G)$.

Man beachte, daß für jeden Normalteiler N von G die Abbildung $\varphi_N : G \to G/N$, $a \mapsto a \circ N$ ein Gruppenhomomorphismus ist. Die Faktorgruppen geben daher (bis auf Isomorphie) einen *Überblick* über die möglichen homomorphen Bilder von G. Übrigens heißt eine Gruppe **einfach**, wenn es bis auf die trivialen Untergruppen keine weiteren Normalteiler gibt, d.h. ein Homomorphismus $\varphi : G \to H$ ist entweder injektiv oder $\varphi(G) = \{e_H\}$.

2.2 Ringe

2.2.1 Halbringe und Ringe

In den ganzen Zahlen $\mathbb{Z}$ verwendet man (wenigstens) zwei verschiedene binäre Operationen, die *Addition* und die *Multiplikation*. Bezüglich der Addition ist $\mathbb{Z}$ eine Gruppe und bezüglich der Multiplikation ein Monoid. Man erfaßt aber die Struktur der ganzen Zahlen $\mathbb{Z}$ (bezüglich Addition und Multiplikation) nicht vollständig, wenn man nur die Stukturen $\langle \mathbb{Z}, + \rangle$ und $\langle \mathbb{Z}, \cdot \rangle$ betrachtet. Es gelten auch Rechenregeln, wie das *Distributivgesetz*

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

wo Addition und Multiplikation gemeinsam auftreten.

Im folgenden werden daher algebraische Strukturen $\langle A, +, \cdot \rangle$ mit zwei binären Operation behandelt, die notationstechnischen Gründen mit + (plus) und $\cdot$ (mal) bezeichnet werden (auch wenn

sie mit der *gewöhnlichen* Addition und Multiplikation nichts zu tun haben). Entsprechend bezeichnet man das neutrale Element von $+$, sofern eines existiert, mit 0 (Null) und das von $\cdot$ mit 1 (Eins). Das additiv inverse Element von a ist dann $-a$ und das multiplikative a^{-1}. Schließlich wird, um Klammern zu sparen, wie üblich die Multiplikation vor der Addition ausgeführt.

Definition 2.44 *Eine algebraische Struktur $\langle R, *, \cdot \rangle$ (mit zwei binären Operation) heißt* **Halbring**, *wenn die folgenden vier Eigenschaften erfüllt sind:*

1. *$\langle R, + \rangle$ ist ein kommutatives Monoid mit neutralem Element 0.*
2. *$\langle R, \cdot \rangle$ ist ein Monoid mit neutralem Element 1, das von 0 verschieden ist.*
3. *Es gelten die* **Distributivgesetze**:

$$\forall a, b, c \in R: \ a \cdot (b + c) = a \cdot b + a \cdot c \ \wedge \ (a + b) \cdot c = a \cdot c + b \cdot c$$

4. *$\forall a \in R : a \cdot 0 = 0 \cdot a = 0$.*

Beispiel 2.45 $\langle \mathbb{N}, +, \cdot \rangle$ ist ein Halbring.

Beispiel 2.46 $R = \{0, 1\}$ mit den Operationen

$+$	0	1
0	0	1
1	1	1

$\cdot$	0	1
0	0	0
1	0	1

ist der sogenannte **Boolesche Halbring**.

Definition 2.47 *Eine algebraische Struktur $\langle R, *, \cdot \rangle$ (mit zwei binären Operation) heißt* **Ring**, *wenn die folgenden drei Eigenschaften erfüllt sind:*

1. *$\langle R, + \rangle$ ist ein kommutative Gruppe (mit neutralem Element 0).*
2. *$\langle R, \cdot \rangle$ ist eine Halbgruppe.*
3. *Es gelten die Distributivgesetze.*

Besitzt R bezüglich $\cdot$ ein neutrales Element ($= 1$), so nennt man R **Ring mit Einselement**, *und ist R bezüglich $\cdot$ kommutativ, so nennt man R* **kommutativen Ring**.

Insbesondere ist jeder Ring mit Einselement 1 ($\neq 0$) auch ein Halbring.

Beispiel 2.48 $\langle \mathbb{Z}, +, \cdot \rangle$ und $\langle \mathbb{Z}_m, +, \cdot \rangle$ ($m \geq 1$) sind kommutative Ringe mit Einselement. (Ähnlich wie die Addition auf $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ definiert man auch die Multiplikation zweier Nebenklassen durch $\overline{a} \cdot \overline{b} = (a + m\mathbb{Z}) \cdot (b + m\mathbb{Z}) := (a \cdot b) + m\mathbb{Z} = \overline{a \cdot b}$.)

Beispiel 2.49 Die Menge $R[x]$ der **Polynome** mit Koeffizienten aus R bildet mit der üblichen Polynomaddition und -multiplikation wieder eine Ring, den **Polynomring über** R.

Sind $p(x) = \sum_{k=0}^{\infty} a_k x^k$, $q(x) = \sum_{k=0}^{\infty} b_k x^k$ zwei Polynome über R, d.h. $a_k, b_k \in R$ und nur endliche viele a_k bzw. b_k sind von 0 verschieden, so berechnen sich Summe $p(x) + q(x)$ und Produkt $p(x) \cdot q(x)$ durch

$$\begin{aligned} p(x) + q(x) &= \sum_{k=0}^{\infty} (a_k + b_k) x^k, \\ p(x) \cdot q(x) &= \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} a_j \cdot b_{k-j} \right) x^k. \end{aligned}$$

Man bezeichnet das maximale k mit $a_k \neq 0$ eine Polynoms $p(x) = \sum_{k=0}^{\infty} a_k x^k$ als den **Grad** von $p(x)$, i.Z. $\operatorname{grad}(p(x))$. Man beachte, daß immer

$$\begin{aligned} \operatorname{grad}(p(x) + q(x)) &\leq \max\left(\operatorname{grad}(p(x)), \operatorname{grad}(q(x))\right), \\ \operatorname{grad}(p(x) \cdot q(x)) &\leq \operatorname{grad}(p(x)) + \operatorname{grad}(q(x)) \end{aligned}$$

gelten.

Beispiel 2.50 Die Menge $R[[x]]$ der **(formalen) Potenzreihen** $\sum_{k=0}^{\infty} a_k x^k$ mit Koeffizienten aus $a_k \in R$ bildet mit formal denselben Operation $+, \cdot$ wie bei den Polynomen aus Beispiel 2.49 wieder eine Ring, den **Ring der formalen Potenzreihen über** R.

2.2.2 Nullteiler und Integritätsbereiche

Lemma 2.51 *In jedem Ring R gilt $a \cdot 0 = 0 \cdot a = 0$.*

In einem Halbring kann diese Eigenschaft nicht aus den anderen Axiomen abgeleitet werden.

In einem Ring kann das Produkt zweier Elemente von 0 verschiedener Elemente 0 sein, in $\mathbb{Z}_6$ gilt z.B. $\overline{2} \cdot \overline{3} = \overline{6} = \overline{0}$.

Definition 2.52 *Eine Element $a \neq 0$ eines Ringes R heißt* **Nullteiler***, wenn es ein $b \neq 0$ aus R gibt, so daß entweder $a \cdot b = 0$ ist oder $b \cdot a = 0$ ist.*

Dieses b ist damit natürlich auch ein Nullteiler.

Definition 2.53 *Ein kommutativer Ring mit Einselement ohne Nullteiler heißt* **Integritätsbereich** *oder* **Integritätsring**.

Beispiel 2.54 $\langle \mathbb{Z}, +, \cdot \rangle$ ist ein Integritätsbereich.

Beispiel 2.55 $\langle \mathbb{Z}_m, +, \cdot \rangle$ $(m \geq 1)$ ist nur dann ein Integritätsbereich, wenn m eine Primzahl ist. Ist m zusammengesetzt, so besitzt $\mathbb{Z}_m$ sicherlich Nullteiler.

Satz 2.56 *Der Polynomring $R[x]$ und der Ring der formalen Potentreihen $R[[x]]$ über einem Ingegritätsbereich R sind wieder Integritätsbereiche.*

Definition 2.57 *Ein Element a eines Ringes (mit Einselement 1) heißt* **Einheit**, *wenn es bezüglich $\cdot$ ein inverses Element a^{-1} gibt.*

Die Menge aller Einheiten R^* von R bildet eine Gruppe, die sogenannte **Einheitengruppe**.

Beispiel 2.58 $\mathbb{Z}^* = \{-1, 1\}$.

Beispiel 2.59 $\mathbb{Z}_m^* = \{\overline{a} = a + m\mathbb{Z} \,|\, \mathrm{ggT}(a, m) = 1\}$. Dies Einheitengruppe heißt auch **Gruppe der primen Restklassen modulo** m. Ihre Ordnung $|\mathbb{Z}_m^*|$ wird auch als **Eulersche Phifunktion** $\varphi(m)$ bezeichnet. Kennt man die Primfaktorenzerlegung von $m = p_1^{e_1} p_2^{e_2} \cdots p_n^{e_n}$, so berechnet man $\varphi(m)$ durch

$$\varphi(m) = \prod_{j=1}^{n} p_j^{e_j - 1}(p_j - 1) = m \prod_{j=1}^{n} \left(1 - \frac{1}{p_j}\right).$$

Wendet man die allgemeine Beziehung $a^{|G|} = e$ auf die Einheitengruppe $\mathbb{Z}_m^*$ an, so erhält man den **kleinen Fermatschen Satz**

$$a^{\varphi(m)} \equiv 1 \bmod m \qquad \text{für } a \text{ mit } \mathrm{ggT}(a, m) = 1.$$

Insbesondere gilt für eine Primzahl p die Formel $\varphi(p) = p - 1$, d.h. alle Elemente aus $\mathbb{Z}_p$ außer $\overline{0}$ sind Einheiten.

2.3 Körper

2.3.1 Integritätsbereiche und Körper

Definition 2.60 *Ein kommutativer Ring $\langle K, +, \cdot\rangle$ mit Einselement $1 \neq 0$, in dem jedes Element $a \neq 0$ eine Einheit ist, heißt* **Körper**.

Eine algebraische Struktur $\langle K, +, \cdot\rangle$ ist also genau dann ein Körper, wenn $\langle K, +\rangle$ und $\langle K \setminus \{0\}, \cdot\rangle$ kommutative Gruppen sind und die Distributivgesetze gelten.

Beispiel 2.61 $\langle \mathbb{Q}, +, \cdot\rangle$, $\langle \mathbb{R}, +, \cdot\rangle$ und $\langle \mathbb{C}, +, \cdot\rangle$ sind Körper.

Satz 2.62 *Jeder Körper ist ein Integritätsbereich.*

Die Umkehrung dieses Satzes gilt im allgemeinen nicht, wie das Beispiel $\mathbb{Z}$ zeigt. Es gilt allerdings:

Satz 2.63 *Jeder endliche Integritätsbereich ist ein Körper.*

Satz 2.64 $\langle \mathbb{Z}_m, +, \cdot\rangle$ *ist genau dann ein Körper, wenn $m = p$ ein Primzahl ist. $\langle \mathbb{Z}_p, +, \cdot\rangle$ ist dann ein* **endlicher Körper** *der Ordnung p.*[3]

[3] Allgemein läßt sich zeigen, daß es nur für Primzahlpotenzen p^m ($m \geq 1$) endliche Körper mit p^m Elementen gibt. Es gibt daher keinen Körper mit 6 Elementen, aber einen mit 8 und einen mit 9.

2.3.2 Euklidischer Algorithmus

Im Beispiel 2.59 wird die Einheitengruppe $\mathbb{Z}_m^* = \{\overline{a} = a + m\mathbb{Z} \,|\, \mathrm{ggT}(a,m) = 1\}$ betrachtet. Ein Element $\overline{a}$ besitzt daher ein inverses Element $\overline{a}'$ genau dann, wenn $\mathrm{ggT}(a,m) = 1$. Abschließend soll nun der **Euklidische Algorithmus** vorgestellt werden, mit dessen Hilfe man nicht nur den größten gemeinsamen Teiler zweier ganzer Zahlen, sonder auch das inverse Element $\overline{a}'$ sehr effektiv berechnet werden kann. Er beruht auf folgender einfachen Eigenschaft.

Lemma 2.65 (Division mit Rest) *Zu je zwei ganzen Zahlen a, b mit $b \neq 0$ gibt es ganze Zahlen q, r mit*

$$a = bq + r \quad \textit{und} \quad 0 \leq r < b.$$

q heißt Quotient und r Rest.

Satz 2.66 *Führt man zu zwei ganzen Zahlen a, b mit $b \neq 0$ die Divisionskette*

$$\begin{aligned}
a &= bq_0 + r_0, & 0 < r_0 < b \\
b &= r_0 q_1 + r_1 & 0 < r_1 < r_0 \\
r_0 &= r_1 q_2 + r_2 & 0 < r_2 < r_1 \\
&\vdots & \\
r_{k-2} &= r_{k-1} q_k + r_k & 0 < r_k < r_{k-1} \\
r_{k-1} &= r_k q_{k+1} + 0 & (r_{k+1} = 0)
\end{aligned}$$

durch, so muß diese wegen $b > r_0 > r_1 > r_2 > \cdots \geq 0$ einmal abbrechen, d.h. es gibt irgendeinmal einen verschwindenden Rest. Der letzte Rest $r_k \neq 0$ ist dann der größte gemeinsame Teiler $\mathrm{ggT}(a,b)$.

Beispiel 2.67 Zur Bestimmung des $\mathrm{ggT}(59, 11)$ ermittelt man die Divisionskette

$$\begin{aligned}
59 &= 11 \cdot 5 + 4 \\
11 &= 4 \cdot 2 + 3 \\
4 &= 3 \cdot 1 + 1 \\
3 &= 1 \cdot 3 + 0.
\end{aligned}$$

Es ist also $\mathrm{ggT}(59, 11) = 1$. Umgekehrt kann man mit Hilfe dieser Divisionskette auch den ggT zweier Zahlen als ganzzahlige Linearkombination von a, b darstellen, indem man ausgehend von der Gleichung $\mathrm{ggT}(59, 11) = 4 - 3 \cdot 1$ sukzessive die weiteren Reste 3 und 4 rückeinsetzt:

$$\begin{aligned}
1 &= 4 - 3 \cdot 1 \\
&= 4 - (11 - 4 \cdot 2) \cdot 1 \\
&= 3 \cdot 4 - 1 \cdot 11 \\
&= 3 \cdot (59 - 5 \cdot 11) - 1 \cdot 11 \\
&= 3 \cdot 59 - 15 \cdot 11 - 1 \cdot 11 \\
&= 3 \cdot 59 - 16 \cdot 11.
\end{aligned}$$

Das ergibt z.B. $11^{-1} = -16 \equiv 43 \bmod 59$.

Satz 2.68 *Ist d der größte gemeinsame Teiler der von Null verschiedenen ganzen Zahlen a, b, so gibt es ganze Zahlen k, l mit*

$$ak + bl = d,$$

die mit Hilfe der Divisionskette von a und b effektiv berechnet werden können.

Die Division mit Rest (Lemma 2.65) funktioniert nicht nur für den Bereich der ganzen Zahlen.

Lemma 2.69 *Seien $a(x)$ und $q(x)$ zwei nichtverschwindende Polynome mit Koeffizienten aus einem Körper K. Dann gibt es Polynome $q(x), r(x) \in K[x]$ mit*

$$a(x) = b(x)q(x) + r(x),$$

wobei $r(x)$ entweder das Nullpolynom ist oder $\operatorname{grad}(r(x)) < \operatorname{grad}(b(x))$.

Genauso wie in den ganzen Zahlen kann jetzt mit einer Divisionskette der größte gemeinsame Teiler zweier Polynome über einem Körper bestimmt werden. Dabei heißt ein Polynom $d(x) \in K[x]$ größter gemeinsamer Teiler der Polynome $a(x), b(x) \in K[x]$, wenn $d(x)$ ein gemeinsamer Teiler von $a(x)$ und $b(x)$ ist und jeder gemeinsame Teiler $t(x) \in K[x]$ von $a(x)$ und $b(x)$ ein Teiler von $d(x)$ ist.

Satz 2.70 *Zu je zwei nichtverschwindenden Polynomen $a(x), b(x)$ mit Koeffizienten aus einem Körper K gibt es immer einen größten gemeinsamen Teiler $d(x) \in K[x]$. Weiters gibt es Polynome $k(x), l(x) \in K[x]$ mit*

$$a(x)k(x) + b(x)l(x) = d(x).$$

Ähnlich wie den Ring $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ kann man auch einen Faktorpolynomring über einem Körper K aufbauen. Sei $p(x) \in K[x]$ ein festes Polynom (ungleich dem Nullpolynom). Die polynomiellen Vielfachen $p(x)K[x]$ dieses Polynoms bilden eine (additiven) Normalteiler von $K[x]$. Auf der Faktorgruppe $K[x]/p(x)K[x]$ kann aber auch (wie in $\mathbb{Z}/m\mathbb{Z}$) in natürlicher Weise eine Multpikation definiert werden, und $K[x]/p(x)K[x]$ wird dadurch wieder zu einem Ring.

Satz 2.71 *Sei K ein Körper und $p(x) \in K[x]$ ein Polynom mit Koeffizienten aus K. Dann ist der Faktorring $\langle K[x]/p(x)K[x], +, \cdot \rangle$ genau dann ein Körper, wenn das Polynom $p(x)$* **irreduzibel** *über K ist, d.h. $p(x)$ kann nicht als Produkt zweier Polynome $a(x), b(x) \in K[x]$ mit kleinerem Grad (als $p(x)$) dargestellt werden.*

Beispiel 2.72 Das Polynom $p(x) = x^2 + 1$ ist irreduzibel über $\mathbb{R}$. Daher ist $\mathbb{R}[x]/(x^2 + 1)\mathbb{R}[x]$ ein Körper. Die Nebenklasse $x+(x^2+1)\mathbb{R}[x]$ hat die Eigenschaft $(x+(x^2+1)\mathbb{R}[x])^2 = -1+(x^2+1)\mathbb{R}[x]$. Es ist leicht einzusehen, daß $\mathbb{R}[x]/(x^2 + 1)\mathbb{R}[x]$ nichts anderes als die komplexen Zahlen $\mathbb{C}$ repräsentiert. Die imaginäre Einheit i muß nur mit $x + (x^2 + 1)\mathbb{R}[x]$ identifiziert werden.

Beispiel 2.73 Ist $q(x)$ ein irreduzibles Polynom über $\mathbb{Z}_p$ (mit einer Primzahl p) vom Grad k, so ist $\mathbb{Z}_p[x]/q(x)\mathbb{Z}_p[x]$ ein endlicher Körper (auch Galoisfeld genannt) mit p^k Elementen.

2.3.3 Charakteristik eines Körpers

Die additive Ordnung von 1 in einem Körper $\langle K, +, \cdot \rangle$ kann endlich oder unendlich sein. Im endliche Fall gilt folgende Eigenschaft.

Lemma 2.74 *Sei $\langle K, +, \cdot \rangle$ ein Körper und sei die additive Ordnung von 1 endlich. Dann ist diese Ordnung eine Primzahl.*

Dies motiviert die folgende Definition.

Definition 2.75 *Die* **Charakteristik** $\mathrm{char}(K)$ *eines Körpers $\langle K, +, \cdot \rangle$ ist 0, wenn die additive Ordnung von 1 unendlich ist und $\mathrm{char}(K) := p$, wenn die additive Ordnung von 1 gleich einer Primzahl p ist.*

Beispiel 2.76 Für eine Primzahl p ist $\mathrm{char}(\mathbb{Z}_p) = p$. Andererseits gilt $\mathrm{char}(\mathbb{Q}) = \mathrm{char}(\mathbb{R}) = \mathrm{char}(\mathbb{C}) = 0$.

Eine andere Möglichkeit, die Charakterisik eines Körpers zu definieren ist den kleinsten Unterkörper

$$K' := \bigcap \{L \mid L \leq K\}$$

von K zu betrachten. Im Fall $\mathrm{char}(K) = 0$ ist K' eine *Kopie* von $\mathbb{Q}$ und im Fall $\mathrm{char}(K) = p$ ist K' eine Kopie von $\mathbb{Z}_p$.

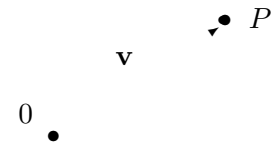
Kapitel 3

Vektorräume

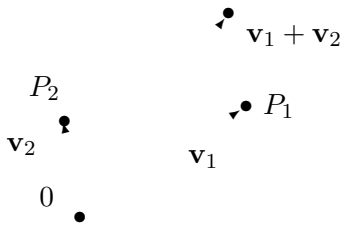
3.1 Vektoren in der Ebene

Es soll zunächst versucht werden, einen Überblick über aller Verschiebungen $\mathbf{V}$ der euklidischen Ebene $\mathbf{E}$ (in sich) zu bekommen.

Dazu zeichnet man einen *Nullpunkt* 0 der Ebene aus. Bei jeder Verschiebung $\mathbf{v} \in \mathbf{V}$ wird der Nullpunkt in einen Punkt P der Ebene übergeführt. Dabei fällt auf, daß dieser Punkt P (in den 0 übergeführt wird) die Verschiebung $\mathbf{v}$ auch schon eindeutig charakterisiert. Man kann daher eine Verschiebung der Ebene etwa durch einen Pfeil von 0 zu P repräsentieren, der nun auch durch $\mathbf{v}$ bezeichnet werden soll.

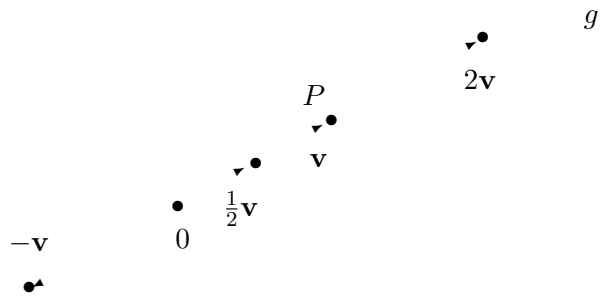


Die Menge $\mathbf{V}$ all dieser Pfeile (Verschiebungen) soll nun als Menge von *Vektoren* $\mathbf{v}$ bezeichnet werden. (Die *Nullverschiebung* wird mit mit einem "*Pfeil*" mit Länge 0 und ohne Richtung, dem *Nullvektor* $\mathbf{0}$ identifiziert.) Führt man nun zwei Verschiebungen $\mathbf{v}_1, \mathbf{v}_2$ hintereinander aus, so entsteht wieder eine Verschiebung $\mathbf{w}$, die wir mit $\mathbf{v}_1 + \mathbf{v}_2$ bezeichnen wollen. Der Vektor $\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2$ kann mit Hilfe der sogenannten *Paralellogrammregel* bestimmt werden



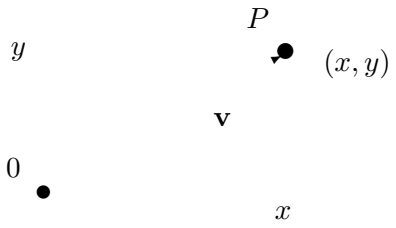
Es ist leicht zu überlegen, daß $\langle \mathbf{V}, + \rangle$ eine abelsche Gruppe ist. Das neutrale Element ist der Nullvektor $\mathbf{0}$, und der negative Vektor $-\mathbf{v}$ entspricht einfach der Verschiebung in die entgegengesetzte Richtung.

Man betrachtet nun eine Gerade $g \in \mathbf{E}$, die den Nullpunkt 0 enthält, und einen Vektor (Pfeil) $\mathbf{v} \neq \mathbf{0}$, der 0 mit einem Punkt $P \neq 0$ auf g verbindet. Dann liegen alle (Spitzen der) ganzzahligen Vielfachen $2\mathbf{v} := \mathbf{v} + \mathbf{v}$, $3\mathbf{v} := \mathbf{v} + \mathbf{v} + \mathbf{v}$, ..., $-1\mathbf{v} := -\mathbf{v}$, $-2\mathbf{v} := (-\mathbf{v}) + (-\mathbf{v})$, ... auf dieser Geraden g. Weiters gibt es einen eindeutig bestimmten Vektor $\mathbf{v}'$ (desses Spitze wirder auf g liegt) mit $2\mathbf{v}' = \mathbf{v}' + \mathbf{v}' = \mathbf{v}$. Wir bezeichnen $\mathbf{v}'$ daher durch $\frac{1}{2}\mathbf{v}$, usw.



Schließlich stellt sich heraus, daß man jeden Vektor $\mathbf{w}$, dessen Spitze auf g liegt, als *Vielfaches* von $\mathbf{v}$ interpretieren kann, d.h. es gibt genau eine reelle Zahl $x \in \mathbb{R}$ mit $\mathbf{w} = x\mathbf{v}$.

Zeichnet man auf $\mathbf{E}$ neben einem Nullpunkt 0 auch zwei Koordinatenrichtungen aus, so kann jeder Punkt $P \in \mathbf{E}$ durch ein geordneten Paar (x, y) von reellen Zahlen *koordinatisiert* werden.



Entsprechend kann jeder Vektor $\mathbf{v} \in \mathbf{V}$ durch so ein Paar (x, y) von reellen Zahlen charakterisiert werden. Entsprechen etwa die Paare (x_1, y_1) und (x_2, y_2) den Vektoren $\mathbf{v}_1, \mathbf{v}_2$, so folgt aus der Parallelogrammregel, daß das Paar $(x_1 + x_2, y_1 + y_2)$ dem Summenvektor $\mathbf{v}_1 + \mathbf{v}_2$ entspricht. In ähnlicher Weise erkennt man, daß das Paar (xx_1, xy_1) dem Vektor $x\mathbf{v}_1$ entspricht. Das Rechnen mit Verschiebungen (Vektoren) in $\mathbf{V}$ ist daher im wesentlichen dasselbe wie das Rechnen in $\mathbb{R}^2$.

Man kann noch eine weitere Beobachung machen. Bezeichnet man den Vektor, der dem Paar $(1, 0)$ entspricht, mit $\mathbf{b}_1$ und der, dem das Paar $(0, 1)$ entspricht, mit $\mathbf{b}_2$, so hat jeder beliebige Vektor $\mathbf{v} \in \mathbf{V}$ (dem etwa das Paar (x, y) entspricht) eine eindeutige Darstellung der Form

$$\mathbf{v} = x\mathbf{b}_1 + y\mathbf{b}_2.$$

Alle Vektoren können eindeutig als *Linearkombination* von zwei *Basisvektoren* dargestellt werden.

Dieses einfache Beispiel soll als Motivation für die allgemeine Definition eines *Vektorraums* (siehe Abschnitt 3.2) dienen. Es ist wahrscheinlich auch günstig, bei den weiteren Begriffen (wie *lineare*

Unabhängigkeit, Basis ect.) dieses Beispiel *vor Augen zu haben*, um eine bessere Intuition zu gewinnen.

3.2 Vektorräume

3.2.1 Definition und Beispiele

Im folgenden wird mit $\langle K, +, \cdot \rangle$ immer ein Körper bezeichnet. Weiters ist 0 immer das neutrale Element der Addition $+$ und 1 das neutrale Element der Multiplikation $\cdot$. Zur Vereinfachung der Schreibweise wird auch das Multiplikationszeichen $\cdot$ weggelassen, d.h. xy bedeutet $x \cdot y$.

Definition 3.1 *Sei K ein Körper und $\langle \mathbf{V}, + \rangle$ eine abelsche Gruppe. Weiters wird jedem $x \in K$ und $\mathbf{a} \in V$ ein "Produkt" $x\mathbf{a} \in V$ zugeordnet.*[1]

V *heißt* **Vektorraum** *über K, wenn die folgenden Eigenschaften (für alle $x, y \in K$ und $\mathbf{a}, \mathbf{b} \in \mathbf{V}$) erfüllt sind:*

1. $x(\mathbf{a} + \mathbf{b}) = x\mathbf{a} + x\mathbf{b}$,
2. $(x + y)\mathbf{a} = x\mathbf{a} + y\mathbf{a}$,
3. $(xy)\mathbf{a} = x(y\mathbf{a})$,
4. $1\mathbf{a} = \mathbf{a}$.

Man beachte, daß das $+$-Zeichen hier für zwei verschiedene Operationen verwenden wird, für die Addition in K und für die (abelsche) Gruppenoperation in $\mathbf{V}$. Dies ist formal nicht präzise, aber erleichtert das Arbeiten in Vektorräumen.

Der Körper K wird im Zusammenhang mit Vektorräumen über K als **Skalarkörper** und die Elemente als **Skalare**. Die Elemente eines Vektorraums $\mathbf{V}$ werden **Vektoren** genannt. Das neutrale Element $\mathbf{0}$ von $\mathbf{V}$ heißt **Nullvektor**.

Zur Vereinfachung der Lesbarkeit werden Skalare *kursiv* geschrieben, wobei bevorzugt Kleinbuchstaben aus dem hinteren Teil des lateinischen Alphabets verwendet werden. Vektoren werden hingeben mit **fett** gedruckten Kleinbuchstaben aus dem vorderen Teil des Alphabets bezeichnet.

Satz 3.2 *In einem Vektorraum $\mathbf{V}$ gelten folgende Rechenregeln:* $(x \in K, \mathbf{a} \in \mathbf{V})$

1. $x\mathbf{0} = \mathbf{0}$,
2. $0\mathbf{a} = \mathbf{0}$,
3. $(-x)\mathbf{a} = x(-\mathbf{a}) = -(x\mathbf{a})$,
4. $x\mathbf{a} = \mathbf{0} \Longrightarrow x = 0 \vee \mathbf{a} = \mathbf{0}$.

[1] Formal ist diese Verknüpfung $x\mathbf{a}$ von $x \in K$ und $\mathbf{a} \in V$ eine Abbildung $K \times V \to V$.

Beispiel 3.3 Sei $n > 0$ eine natürliche Zahl und K^n das direkte Produkt der Gruppen $\langle K, + \rangle$, d.h.

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n).$$

Definiert man nun das Produkt $x(x_1, x_2, \ldots, x_n)$ durch

$$x(x_1, x_2, \ldots, x_n) := (xx_1, xx_2, \ldots, xx_n),$$

so wird K^n zu einem Vektorraum.

Beispiel 3.4 Unter einer $m \times n$-**Matrix A** versteht man ein $m \cdot n$-Tupel $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$. Üblicherweise stellt man **A** durch ein rechteckiges Schema aus m Zeilen und n Spalten dar. Beispielsweise ist

$$\begin{pmatrix} 1 & 2 & 5 \\ 2 & 2 & 1 \end{pmatrix}$$

eine 2×3-Matrix.

Bezeichnet man mit $K^{m \times n}$ die Menge aller $m \times n$-Matrizen mit Eintragungen aus K, so wird $K^{m \times n}$ mit den Operationen

$$(x_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} + (y_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} := (x_{ij} + y_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$$

und

$$x(x_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} := (xx_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$$

zu einem Vektorraum.

Beispiel 3.5 Sei M eine nichtleere Menge und $\mathbf{V} = K^M$ die Menge aller Funktionen $f : M \to K$. Definiert man $f + g$ durch $(f + g)(a) := f(a) + g(a)$ und xf durch $(xf)(a) := xf(a)$, dann ist **V** ein Vektorraum. (Die Spezialfälle $M = \{1, 2, \ldots, n\}$ und $M = \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$ wurden in den vorigen Beispielen behandelt.)

Beispiel 3.6 Sei R ein Ring und $K \subseteq R$ (mit denselben Operationen wie in R) ein Körper. Dann ist R auch Vektorraum über K. Beispielsweise bilden die Polynome $K[x]$ und die formalen Potenzreihen $K[[x]]$ Vektorräume über K. Insbesondere bildet auch jeder Oberkörper L von K einen Vektorraum über K.

3.2.2 Unterräume

Definition 3.7 *Eine Teilmenge* **U** *eines Vektorraums* **V** *über* K *heißt* **Unterraum** *oder* **Teilraum** *von* **V**, *i.Z.* $\mathbf{U} \leq \mathbf{V}$, *wenn* **U** *(mit den Operationen aus* **V***) wieder einen Vektorraum über* K *bildet.*

Beispiel 3.8 **V** und der sogenannte **Nullraum** $\{\mathbf{0}\}$ sind immer Unterräume von **V**. Sie sind die sogenannte trivialen Unterräume von **V**.

Satz 3.9 *Sei* **V** *Vektorraum über* K *und* **U** *nichtleere Teilmenge von* **V**. *Dann sind folgende drei Bedingungen äquivalent:*

(i) $\mathbf{U} \leq \mathbf{V}$.

(ii) $\forall \mathbf{a}, \mathbf{b} \in \mathbf{U}\ \forall x \in K : \mathbf{a} + \mathbf{b} \in \mathbf{U} \wedge x\mathbf{a} \in \mathbf{U}$.

(iii) $\forall \mathbf{a}, \mathbf{b} \in \mathbf{U}\ \forall x \in K : \mathbf{a} + x\mathbf{b} \in \mathbf{U}$.

Um sicherzustellen, daß eine Teilmenge $\mathbf{U} \subseteq \mathbf{V}$ Unterraum ist, muß also nur $\mathbf{U} \neq \emptyset$ und *(ii)* oder *(iii)* überprüft werden.

Beispiel 3.10 $\mathbf{U} = \{(x_1, x_2, x_3) \in K^3 \,|\, x_1 + x_2 + x_3 = 0\}$ ist Unterraum von $\mathbf{V} = K^3$.

Satz 3.11 *Sei* $(\mathbf{U}_i, i \in I)$ *ein System von Unterräumen eines Vektorraums* $\mathbf{V}$. *Dann ist auch*

$$\mathbf{U} := \bigcap_{i \in I} \mathbf{U}_i$$

eine Unterraum von $\mathbf{V}$.

Definition 3.12 *Sei* $\mathbf{V}$ *ein Vektorraum und* $\mathbf{M}$ *eine Teilmenge von* $\mathbf{V}$. *Der von* $\mathbf{M}$ **erzeugte Unterraum** $[\mathbf{M}]$ *ist der Durchschnitt aller Unterräume* $\mathbf{U}$ *von* $\mathbf{V}$, *die* $\mathbf{M}$ *enthalten:*

$$[\mathbf{M}] := \bigcap \{\mathbf{U} \leq \mathbf{V} \,|\, \mathbf{M} \subseteq \mathbf{U}\}.$$

Eine nichtleere Teilmenge $\mathbf{M}$ *von* $\mathbf{V}$ *heißt* **Erzeugendensystem** *von* $\mathbf{V}$, *wenn* $[\mathbf{M}] = \mathbf{V}$.

Definition 3.13 *Seien* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ *Vektoren eines Vektorraums* $\mathbf{V}$ *(über* K*) und* $x_1, x_2, \ldots, x_n \in K$. *Dann heißt der Vektor*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \sum_{i=1}^{n} x_i\mathbf{a}_i$$

Linearkombination *der Vektoren* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. *Die Skalare* $x_1, x_2, \ldots, x_n \in K$ *heißen* **Koeffizienten** *der Linearkombination. Eine Linearkombination heißt* **trivial**, *wenn alle Koeffizienten* $x_i = 0$ *sind* $(1 \leq i \leq n)$. *Andernfalls heißt sie* **nichttrivial**.

Die **lineare Hülle** $H(\mathbf{M})$ *einer nichtleeren Teilmenge* $\mathbf{M}$ *von* $\mathbf{V}$ *ist die Menge aller Linearkombinationen (von endlichen Auswahlen) von Vektoren aus* $\mathbf{M}$.

Satz 3.14 *Für jede nichtleere Teilmenge* $\mathbf{M}$ *eines Vektorraums gilt*

$$H(\mathbf{M}) = [\mathbf{M}].$$

Beispiel 3.15 Der von einem Vektor $\mathbf{a} \in \mathbf{V} \setminus \mathbf{0}$ erzeuge Unterraum ist durch $[\mathbf{a}] = \{x\mathbf{a} \,|\, x \in K\}$ gegeben.

3.2.3 Faktorräume

In Analogie zur Gruppentheorie werden mit Hilfe von Unterräumen von Vektorräumen auch Faktorräume betrachtet.

Definition 3.16 *Sei* $\mathbf{U}$ *Teilraum eines Vektorraums* $\mathbf{V}$ *(über* K*). Eine Nebenklasse* $\mathbf{a}+\mathbf{U}$ *der Untergruppe* $\langle \mathbf{U}, + \rangle$ *von* $\langle \mathbf{V}, + \rangle$ *heißt* **Nebenraum**.

Die Menge $\mathbf{V}/\mathbf{U}$ *aller Nebenräume* $\mathbf{a}+\mathbf{U}$ $(\mathbf{a} \in \mathbf{V})$ *mit den Operationen*

$$(\mathbf{a}+\mathbf{U})+(\mathbf{b}+\mathbf{U}) := (\mathbf{a}+\mathbf{b})+\mathbf{U}$$

und

$$x(\mathbf{a}+\mathbf{U}) := (x\mathbf{a})+\mathbf{U}$$

ist ein Vektorraum über K*, der* **Faktoraum** $\mathbf{V}/\mathbf{U}$.

3.2.4 Summe von Unterräumen

Definition 3.17 *Ist* $\mathbf{U}_i$ $(i \in I)$ *ein System von Unterräumen eines Vektorraums* $\mathbf{V}$*, so bezeichnet*

$$\sum_{i\in I} \mathbf{U}_i := \left[\bigcup (\mathbf{U}_i \mid i \in I)\right]$$

die **Summe** *der Unterräume* $\mathbf{U}_i$ $(i \in I)$*, d.h. den kleinsten Unterraum, der alle* $\mathbf{U}_i$ $(i \in I)$ *enthält.*

Gilt für alle $j \in I$

$$\left(\sum_{i\in I\setminus\{j\}} \mathbf{U}_i\right) \cap \mathbf{U}_j = \{\mathbf{0}\},$$

so bezeichnet man $\sum\limits_{i\in I} \mathbf{U}_i$ *als* **direkte Summe** *der Unterräume* $\mathbf{U}_i$ $(i \in I)$ *und schreibt dafür*

$$\bigoplus_{i\in I} \mathbf{U}_i.$$

Ist $I = \{1, 2, \ldots, n\}$, so schreibt man für die Summe von $\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_n$ auch

$$\mathbf{U}_1 + \mathbf{U}_2 + \cdots + \mathbf{U}_n$$

und für die direkte Summe

$$\mathbf{U}_1 \oplus \mathbf{U}_2 \oplus \cdots \oplus \mathbf{U}_n.$$

Satz 3.18 *Sei* $\mathbf{U}_i$ $(i \in I)$ *ein System von Unterräumen eines Vektorraums* $\mathbf{V}$*. Dann gilt*

$$\sum_{i\in I} \mathbf{U}_i = \left\{ \sum_{i\in I} \mathbf{a}_i \;\middle|\; (\forall i \in I : \mathbf{a}_i \in \mathbf{U}_i) \wedge |\{i \in I \mid \mathbf{a}_i \neq \mathbf{0}\}| < \aleph_0 \right\},$$

d.h. $\sum\limits_{i\in I} \mathbf{U}_i$ *besteht auch allen endlichen Summen von Vektoren aus* $\bigcup\{\mathbf{U}_i \mid i \in I\}$.

Ist die Summe der $\mathbf{U}_i$ $(i \in I)$ *überdies direkt, so besitzt jeder Vektor aus* $\bigoplus\limits_{i\in I} \mathbf{U}_i$ *eine eindeutige Darstellung als endliche Summe von Vektoren* $\mathbf{a}_i$ $(i \in I)$.

Die direkte Summe von Unterräumen $\bigoplus_{i \in I} \mathbf{U}_i$ kann daher mit den semidirekten Produkt der Vektorräume $(\mathbf{U}_i \,|\, i \in I)$ identifiziert werden.

Definition 3.19 *Sei $\mathbf{U} \leq \mathbf{V}$ Unterraum eines Vektorrausm $\mathbf{V}$. Ein Unterraum $\mathbf{W} \leq \mathbf{V}$ heißt* **Komplementärraum***, wenn*

$$\mathbf{U} \oplus \mathbf{W} = \mathbf{V}.$$

Satz 3.20 *Jeder Unterraum $\mathbf{U} \leq \mathbf{V}$ eines Vektorraums $\mathbf{V}$ besitzt einen Komplementärraum.*

3.3 Dimension und Basis

3.3.1 Linear unabhängige und linear abhängige Vektoren

Nächstes Ziel ist es, minimale Erzeugendensysteme, sogenannte Basen, zu charakterisieren. Die beiden folgenden Begriffe, der der linearen Unabhängigkeit und der der linearen Abhängigkeit, sind nicht nur für die Definition einer Basis nützlich, sondern spielen in der gesamten Linearen Algebra eine zentrale Rolle.

Definition 3.21 *Sei $n > 0$ eine natürliche Zahl und seien $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ Vektoren aus einem Vektorraum $\mathbf{V}$ (über K).*

Die Vektoren $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ heißen **linear unabhängig** *über K, wenn aus*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \sum_{i=1}^{n} x_i\mathbf{a}_i = \mathbf{0}$$

folgt, daß alle Koeffizienten $x_i = 0$ sind $(1 \leq i \leq n)$, d.h. jede nichttriviale Linearkombination ist $\neq \mathbf{0}$. Eine nichtleere Teilmenge $\mathbf{M}$ von $\mathbf{V}$ heißt **linear unabhängig***, wenn jede endliche Auswahl (von paarweise verschiedenen) Vektoren von $\mathbf{M}$ linear unabhängig ist.*

Die Vektoren $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ heißen **linear abhängig** *über K, wenn sie nicht linear unabhängig sind, d.h. es gibt eine nichttriviale Linearkombination mit*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \sum_{i=1}^{n} x_i\mathbf{a}_i = \mathbf{0}.$$

Eine nichtleere Teilmenge $\mathbf{M}$ von $\mathbf{V}$ heißt **linear abhängig***, wenn es eine endliche Auswahl (von paarweise verschiedenen) Vektoren von $\mathbf{M}$ gibt, die linear abhängig sind.*

Man beachte, daß linear unabhängige Vektoren $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ immer paarweise verschieden sein müssen.

Beispiel 3.22 $\mathbf{0}$ ist linear abhängig, $\mathbf{a} \neq \mathbf{0}$ ist linear unabhängig.
$\emptyset$ ist linear unabhängig, $\mathbf{V}$ ist linear abhängig.

Lemma 3.23 *n Vektoren* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ *eines Vektorraums* $\mathbf{V}$ *sind genau dann linear abhängig, wenn es ein i gibt, so daß der Vektor* $\mathbf{a}_i$ *als Linearkombination der restlichen* $n-1$ *Vektoren*

$$\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n$$

dargestellt werden kann.

Satz 3.24 *Sei* $\mathbf{T}$ *eine nichtleere Teilmenge eines Vektorraums* $\mathbf{V}$*. Dann gelten die folgenden Eigenschaften:*

$$\begin{aligned} \mathbf{T} \text{ linear unabhängig} &\iff \forall\, \mathbf{a} \in \mathbf{T} : \mathbf{a} \notin [\mathbf{T} \setminus \{\mathbf{a}\}], \\ \mathbf{T} \text{ linear abhängig} &\iff \exists\, \mathbf{a} \in \mathbf{T} : \mathbf{a} \in [\mathbf{T} \setminus \{\mathbf{a}\}]. \end{aligned}$$

3.3.2 Basis eines Vektorraums

Definition 3.25 *Eine Teilmenge* $\mathbf{B}$ *eines Vektorraums* $\mathbf{V}$ *heißt* **Basis** *von* $\mathbf{V}$ *wenn* $\mathbf{B}$ *ein linear unabhängiges Erzeugendensystem von* $\mathbf{V}$ *ist.*

Beispiel 3.26 Die Vektoren $\mathbf{e}_1 := (1, 0, 0, \ldots, 0)$, $\mathbf{e}_2 := (0, 1, 0, \ldots, 0)$,..., $\mathbf{e}_n := (0, 0, \ldots, 0, 1)$ bilden eine Basis des Vektorraums $\mathbf{V} = K^n$.

Beispiel 3.27 Die Monome $1 = x^0, x = x^1, x^2, x^3, \ldots$ bilden eine Basis von $K[x]$.

Offensichtliche können Basen auf verschiedene Arten charakterisiert werden.

Satz 3.28 *Für eine Teilmenge* $\mathbf{B}$ *eines Vektorraums* $\mathbf{V}$ *sind folgende Aussagen äquivalent:*

(i) $\mathbf{B}$ *is eine Basis von* $\mathbf{V}$*.*

(ii) $\mathbf{B}$ *ist ein minimales Erzeugendensystem von* $\mathbf{V}$*, d.h. jede echte Teilmenge von* $\mathbf{B}$ *ist kein Erzeugendensystem mehr.*

(iii) $\mathbf{B}$ *ist eine maximale linear unabhängige Teilmenge von* $\mathbf{V}$*, d.h. jede Teilmenge von* $\mathbf{V}$*, die* $\mathbf{B}$ *echt umfaßt, ist linear abhängig.*

(iv) Jeder Vektor aus $\mathbf{V}$ *besitzt eine eindeutige Darstellung als Linearkombination von Vektoren aus* $\mathbf{B}$

Mit Hilfe dieser Charakterisierungen und des Lemmas von Zorn läßt sich der folgende Satz beweisen.

Satz 3.29 *Ist* $\mathbf{M}$ *eine linear unabhängige Teilmenge eines Vektorraums* $\mathbf{V}$*, dann gibt es eine Basis* $\mathbf{B}$ *von* $\mathbf{V}$*, die* $\mathbf{M}$ *umfaßt. Insbesondere hat jeder Vektorraum eine Basis.*

3.3.3 Koordinaten

Definition 3.30 *Sei* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *eine Basis eines Vektorraums* $\mathbf{V}$ *(über* K*). Dann heißen die eindeutig bestimmten Koeffizienten* $x_i \in K$ $(i \in I)$ *der Linearkombination, die einen Vektor* $\mathbf{a} \in \mathbf{V}$ *darstellen,*

$$\mathbf{a} = \sum_{i \in I} x_i \mathbf{b}_i,$$

Koordinaten *des Vektors* $\mathbf{a}$ *bezüglich der Basis* $\mathbf{B}$*.*

Die Abbildung $\Phi_{\mathbf{B}} : \mathbf{V} \to K^I, \mathbf{a} \mapsto (x_i)_{i \in I}$ *wird als* **Koordinatisierung** *oder* **Koordinatenabbildung** *bezeichnet.*

Man beachte, daß ganau jene I-Tupel $(x_i)_{i \in I}$ als Koordinaten auftreten, wo nur endlich viele von 0 verschienden sind. Jeder Vektorraum $\mathbf{V}$ kann daher mit dem semidirekten Produkt von $|I|$ Kopien von K identifiziert werden.

Dieses semidirekte Produkt kann auch folgendermaßen beschrieben werden. Sei $\mathbf{e}_j = (e_{j,i})_{i \in I} \in K^I$ definiert durch $e_{j,j} := 1$ und $e_{j,i} := 0$ für $i \in I \backslash \{j\}$. Dann ist die lineare Hülle $K^{I*} := [\{\mathbf{e}_j \mid j \in I\}]$ gerade dieses semidirekte Produkt. K^{I*} ist wieder ein Vektorraum und $\mathbf{E} := \{\mathbf{e}_j \mid j \in I\}$ ist eine Basis, die **kanonische Basis** von K^{I*}. $\Phi_{\mathbf{B}}$ kann dann auch als bijektive Abbildung $\mathbf{V} \to K^{I*}$ interpretiert werden.[2]

Ist $\mathbf{V}$ endlichdimensional, d.h. wir können für I ohne Beschränkung der Allgemeinheit $I = \{1, 2, \ldots, n\}$ verwenden, dann ist $\Phi_{\mathbf{B}}$ eine bijektive Abbildung $\mathbf{V} \to K^n$. (Das semidirekte Produkt K^{I*} fällt mit dem direkten Produkt K^n zusammen.)

Satz 3.31 *Sei* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *eine Basis eines Vektorraums* $\mathbf{V}$ *(über* K*). Dann ist die Abbildung* $\Phi_{\mathbf{B}} : \mathbf{V} \to K^{I*}$ *bijektiv und erfüllt die Eigenschaften:*

$$\begin{aligned} \Phi_{\mathbf{B}}(\mathbf{a} + \mathbf{b}) &= \Phi_{\mathbf{B}}(\mathbf{a}) + \Phi_{\mathbf{B}}(\mathbf{b}), \\ \Phi_{\mathbf{B}}(x\mathbf{a}) &= x\Phi_{\mathbf{B}}(\mathbf{a}). \end{aligned}$$

In anderen Worten, $\Phi_{\mathbf{B}} : \mathbf{V} \to K^{I*}$ ist ein Vektorraumisomorphismus (siehe Definition 4.9).

3.3.4 Der Austauschsatz von Steinitz

Aus der Definition einer Basis (und auch aus den vorigen beiden Sätzen) ist nicht ersichtlich, ob zwei verschiedene Basen eines Vektorraums gleichmächtig sein müssen oder nicht. Tatsächlich sind alle Basen eines Vektorraums gleichmächtig. Für den Beweis benötigt man den Austauschsatz von Steinitz.

Lemma 3.32 (Austauschlemma) *Sei* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *eine Basis eines Vektorraums* $\mathbf{V}$ *(über* K*) und*

$$\mathbf{a} = \sum_{i \in I} x_i \mathbf{b}_i.$$

[2] $\Phi_{\mathbf{B}}$ ist ein Vektorraumisomorphismus, siehe Abschnitt 4.1.

Dann ist für jedes $j \in I$ mit $x_j \neq 0$ die Menge

$$\mathbf{B}' = (\mathbf{B} \setminus \{\mathbf{b}_j\}) \cup \{\mathbf{a}\}$$

wieder eine Basis von $\mathbf{V}$.

Satz 3.33 (Austauschsatz von Steinitz) *Sei* $\mathbf{V}$ *ein Vektorraum (über* K*),* $\mathbf{B}$ *ein endliche Basis von* $\mathbf{V}$ *und* $\mathbf{A}$ *eine linear unabhängige Teilmenge von* $\mathbf{V}$*. Dann ist* $\mathbf{A}$ *endlich, und es gibt eine Teilmenge* $\mathbf{T}$ *von* $\mathbf{B}$ *mit* $|\mathbf{T}| = |\mathbf{A}| \leq |\mathbf{B}|$ *so, daß die Menge*

$$\mathbf{B}' = (\mathbf{B} \setminus \mathbf{T}) \cup \mathbf{A}$$

wieder eine Basis von $\mathbf{V}$ *ist.*

Dieser Satz bleibt übrigens richtig, wenn man *endliche Basis* $\mathbf{B}$ durch *endliches Erzeugendensystem* $\mathbf{M}$ ersetzt, da es zu jedem endlichen Erzeugendensystem $\mathbf{M}$ immer eine (endliche) Basis $\mathbf{B} \subseteq \mathbf{M}$ gibt.

3.3.5 Dimension eines Vektorraums

Satz 3.34 *Je zwei Basen eines Vektorraums sind gleichmächtig.*

Definition 3.35 *Ein Vektorraum* $\mathbf{V}$ *heißt* **endlichdimensional***, wenn es eine endliche Basis* $\mathbf{B}$ *gibt. Die* **Dimension** $\dim(\mathbf{V})$ *von* $\mathbf{V}$ *ist die Anzahl der Element von* $\mathbf{B}$*, d.h. die Kardinalität von* $\mathbf{B}$*:*

$$\dim(\mathbf{V}) := |\mathbf{B}|.$$

Besitzt ein Vektorraum keine endliche Basis, so heißt er **unendlichdimensional** *und man schreibt*

$$\dim(\mathbf{V}) := \infty.$$

Beispiel 3.36 $\dim(K^n) = n$, $\dim(K[x]) = \infty$.

Satz 3.37 *Für jeden Unterraum* $\mathbf{U}$ *eines Vektorraums* $\mathbf{V}$ *gilt* $\dim(\mathbf{U}) \leq \dim(\mathbf{V})$.

Satz 3.38 *Ist* $\mathbf{U}$ *Unterraum eines endlichdimensionalen Vektorraums* $\mathbf{V}$ *mit* $\dim(\mathbf{U}) = \dim(\mathbf{V})$*, dann ist* $\mathbf{U} = \mathbf{V}$.

3.3.6 Dimensionsformel

Satz 3.39 *Für je zwei Unterräume* $\mathbf{U}_1, \mathbf{U}_2$ *eines Vektorraums* $\mathbf{V}$ *gilt*

$$\dim(\mathbf{U}_1) + \dim(\mathbf{U}_2) = \dim(\mathbf{U}_1 \cap \mathbf{U}_2) + \dim(\mathbf{U}_1 + \mathbf{U}_2).$$

Insbesondere gilt für die direkte Summe $\mathbf{U}_1 \oplus \mathbf{U}_2$ zweier Unterräume

$$\dim(\mathbf{U}_1 \oplus \mathbf{U}_2) = \dim(\mathbf{U}_1) + \dim(\mathbf{U}_2).$$

Sind $\mathbf{U}_1, \mathbf{U}_2$ Komplementärräume von $\mathbf{V}$, d.h. $\mathbf{U}_1 \oplus \mathbf{U}_2 = \mathbf{V}$, dann gilt auch

$$\dim(\mathbf{U}_1) + \dim(\mathbf{U}_2) = \dim(\mathbf{V}).$$

Man bezeichet $\dim(\mathbf{V}) - \dim(\mathbf{U}_1)$ auch als **Kodimension** des Unterraums $\mathbf{U}_1$. Die Dimension eines Komplementärraums von $\mathbf{U}_1$ ist daher genau die Kodimension von $\mathbf{U}_1$.

Kapitel 4

Lineare Abbildungen

4.1 Der Vektorraum der linearen Abbildungen

4.1.1 Lineare Abbildungen

Definition 4.1 *Eine Abbildung $f : \mathbf{V} \to \mathbf{W}$ zwischen zwei Vektorräumen über demselben Körper K heißt* **linear**, *wenn folgende zwei Eigenschaften gelten:*

(i) $\forall\, \mathbf{a}, \mathbf{b} \in \mathbf{V} : f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b})$,

(ii) $\forall\, \mathbf{a} \in \mathbf{V}\ \forall x \in K : f(x\mathbf{a}) = xf(\mathbf{a})$.

Die Menge aller linearen Abblildungen $f : \mathbf{V} \to \mathbf{W}$ wird mit $\mathrm{L}(\mathbf{V}, \mathbf{W})$ *bezeichnet.*

Die Menge $f(\mathbf{V}) \subseteq \mathbf{W}$ heißt **Bild** *von $\mathbf{V}$ unter f und* $\operatorname{kern}(f) := f^{-1}(\{\mathbf{0}\})$ **Kern** *von f.*

Eine lineare Abbildung ist daher ein Gruppenhomomorphismus $\langle \mathbf{V}, + \rangle \to \langle \mathbf{W}, + \rangle$, der zusätzlich mit der Skalarmultiplikation verträglich ist.

Beispiel 4.2 Sei $\mathbf{V} = \mathbf{U} \oplus \mathbf{W}$ direkte Summe der Unterräume $\mathbf{U}$ und $\mathbf{W}$, d.h. jeder Vektor $\mathbf{a} \in \mathbf{V}$ hat eine eindeutig Darstellung in der Form $\mathbf{a} = \mathbf{a_U} + \mathbf{a_W}$ mit $\mathbf{a_U} \in \mathbf{U}$ und $\mathbf{a_W} \in \mathbf{W}$. Die Abbildung

$$\begin{aligned} p: \quad & \mathbf{V} \to \mathbf{U} \\ & \mathbf{a} \mapsto \mathbf{a_U} \end{aligned}$$

heißt **Projektion** von $\mathbf{V}$ auf $\mathbf{U}$ in Richtung $\mathbf{W}$ und ist eine lineare Abbildung.

So ist etwa die Projektion des *Anschauungsraums* auf die *x-y-Ebene* $\mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (x, y)$ ist eine lineare Abbildung.

Beispiel 4.3 Eine Drehung um den Ursprung ist eine lineare Abbildung $\mathbb{R}^3 \to \mathbb{R}^3$.

Beispiel 4.4 Sei $\mathbf{V}$ der Vektorraum aller Polynome mit Koeffizieten aus K. Dann ist die Abbildung $D : \mathbf{V} \to \mathbf{V}$,

$$p(x) = \sum_{n=0}^{\infty} a_n x^n \mapsto p'(x) = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$$

linear.

Beispiel 4.5 Sei $\mathbf{V} = C[0,1]$ der Vektoraum aller stetigen Funktionen $f : [0,1] \to \mathbb{R}$ und $\mathbf{W} = \mathbb{R}$. Dann ist die Abbildung $I : \mathbf{V} \to \mathbf{W}$,

$$f \mapsto \int_0^1 f(x)\, dx$$

eine lineare Abbildung.

Beispiel 4.6 Sei $\mathbf{V} = L^1(\mathbb{R})$ der Vektorraum aller Funktionen $f : \mathbb{R} \to \mathbb{C}$ mit $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$ und $\mathbf{W}$ der Vektorraum aller Funktionen $g : \mathbb{R} \to \mathbb{C}$ mit $\lim_{|x|\to\infty} g(x) = 0$. Dann ist Abbildung $F : \mathbf{V} \to \mathbf{W}$,

$$f \mapsto F(f), \quad (F(f))(t) := \int_{-\infty}^{\infty} f(x) e^{-ixt}\, dx$$

linear. (F wird auch als Fouriertransformation bezeichnet.)

Lemma 4.7 *Sei $f : \mathbf{V} \to \mathbf{W}$ eine lineare Abbildung zwischen zwei Vektorräumen $\mathbf{V}, \mathbf{W}$ über demselben Körper K, und seien $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbf{V}$ und $x_1, x_2, \ldots, x_n \in K$. Dann gilt*

$$f\left(\sum_{i=1}^{n} x_i \mathbf{a}_i\right) = \sum_{i=1}^{n} x_i f(\mathbf{a}_i).$$

Satz 4.8 *Sei $f : \mathbf{V} \to \mathbf{W}$ eine lineare Abbildung zwischen zwei Vektorräumen über demselben Körper K. Dann gelten folgende Eigenschaften:*

1. $\mathbf{U} \leq \mathbf{V} \Longrightarrow f(\mathbf{U}) \leq \mathbf{W}$.
2. $\mathbf{T} \leq \mathbf{W} \Longrightarrow f^{-1}(\mathbf{T}) \leq \mathbf{V}$.
3. $\mathbf{M} \subseteq \mathbf{V} \Longrightarrow f([\mathbf{M}]) = [f(\mathbf{M})]$.
4. *Ist $f : \mathbf{V} \to \mathbf{W}$ bijektiv, so ist auch die inverse Abbildung $f^{-1} : \mathbf{W} \to \mathbf{V}$ linear.*

Insbesondere sind der Kern $\ker(f)$ und das Bild $f(\mathbf{V})$ Unterräume von $\mathbf{V}$ bzw. $\mathbf{W}$.

Definition 4.9 *Zwei Vektorräume $\mathbf{V}$, $\mathbf{W}$ über demselben Körper K heißen* **isomorph**, *i.Z. $\mathbf{V} \cong \mathbf{W}$, wenn es eine bijektive lineare Abbildung $f : \mathbf{V} \to \mathbf{W}$ gibt. So eine lineare Abbildung heißt auch* **Vektorraumisomorphismus**.

Wegen Satz 4.8.3 ist die Relation $\cong$ symmetrisch.

Satz 4.10 *Seien* $\mathbf{V}, \mathbf{W}$ *zwei Vektorräume über demselben Körper* K *und* $\mathbf{B}$ *eine Basis von* $\mathbf{V}$. *Dann gelten die folgenden Eigenschaften.*

1. *Ist* $\tilde{f} : \mathbf{B} \to \mathbf{W}$ *eine (beliebige Abbildung), dann gibt es genau eine lineare Abbildung* $f : \mathbf{V} \to \mathbf{W}$ *mit* $f|_{\mathbf{B}} = \tilde{f}$.

2. *Eine lineare Abbildung* $f : \mathbf{V} \to \mathbf{W}$ *ist genau dann injektiv, wenn* $f|_{\mathbf{B}}$ *injektiv und* $f(\mathbf{B})$ *linear unabhängig ist. Das ist auch genau dann der Fall, wenn* $\operatorname{kern} f = \{\mathbf{0}\}$ *ist*

3. *Eine lineare Abbildung* $f : \mathbf{V} \to \mathbf{W}$ *ist genau dann surjektiv, wenn* $[f(\mathbf{B})] = \mathbf{W}$ *gilt, d.h. wenn die Bilder der Basis ein Erzeugendensystem von* $\mathbf{W}$ *bilden.*

4. *Eine lineare Abbildung* $f : \mathbf{V} \to \mathbf{W}$ *ist genau dann bijektiv, wenn* $f|_{\mathbf{B}}$ *injektiv und* $f(\mathbf{B})$ *eine Basis von* $\mathbf{W}$ *ist.*

Satz 4.11 *Zwei Vektorräume (über demselben Körper* K*) sind genau dann isomorph, wenn sie gleichmächtige Basen besitzen insbesondere haben isomorphe Vektorräume gleiche Dimension.*

Definition 4.12 *Die Menge* $\mathrm{L}(\mathbf{V}, \mathbf{W})$ *der linearen Abbildungen zwischen zwei Vektorräumen* $\mathbf{V}, \mathbf{W}$ *(über demselben Körper* K*) bilden mit den Operationen*

$$(f + g)(\mathbf{a}) := f(\mathbf{a}) + g(\mathbf{a})$$

und

$$(xf)(\mathbf{a}) := xf(\mathbf{a})$$

den **Vektorraum der linearen Abbildungen**.

Satz 4.13 *Für endlichdimensionale Vektorräume* $\mathbf{V}, \mathbf{W}$ *gilt*

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

4.1.2 Rang und Defekt einer linearen Abbildung

Definition 4.14 *Seien* $\mathbf{V}, \mathbf{W}$ *zwei Vektorräume über demselben Körper* K*, und sei* $\mathbf{V}$ *endlichdimensional. Die Dimension der Kernes* $\dim(\operatorname{kern}(f)$ *einer linearen Abbildung* $f : \mathbf{V} \to \mathbf{W}$ *heißt* **Defekt** *von* f*, i.Z.* $\operatorname{def}(f)$*. Die Dimension des Bildes* $\dim(f(\mathbf{V}))$ *heißt* **Rang** *von* f*, i.Z.* $\operatorname{rg}(f)$.

Satz 4.15 (Rangformel) *Sei* $f : \mathbf{V} \to \mathbf{W}$ *eine lineare Abbildung zwischen zwei Vektorräumen über demselben Körper* K*, wobei* $\mathbf{V}$ *endlichdimenional ist. Dann gilt*

$$\operatorname{def}(f) + \operatorname{rg}(f) = \dim(\mathbf{V}).$$

4.1.3 Faktorräume und lineare Abbildungen

Satz 4.16 *Sei $f : \mathbf{V} \to \mathbf{W}$ eine lineare Abbildung zwischen zwei Vektorräumen $\mathbf{V}, \mathbf{W}$ (über demselben Körper K). Dann ist der Faktorraum $\mathbf{V}/\mathrm{kern}(f)$ mit $f(\mathbf{V})$ isomorph:*

$$\mathbf{V}/\mathrm{kern}(f) \cong f(\mathbf{V})$$

Der Nebenraum $\mathbf{a} + \mathrm{kern}(f) \in \mathbf{V}/\mathrm{kern}(f)$ entspricht dem Element $f(\mathbf{a}) \in f(\mathbf{V})$.

Ist andererseits $\mathbf{U} \leq \mathbf{V}$ ein Unterraum von $\mathbf{V}$, so ist die Abbildung $f_{\mathbf{U}} : \mathbf{V} \to \mathbf{V}/\mathbf{U}, \mathbf{a} \mapsto \mathbf{a} + \mathbf{U}$ eine lineare Abbildung mit $\mathrm{kern}(f_{\mathbf{U}}) = \mathbf{U}$.

Alle möglichen Faktorräume $\mathbf{V}/\mathbf{U}$ geben daher (bis auf Isomorphie) einen *Überblick* über alle möglichen Bilder von linearen Abbildungen.

Aus der Rangformel (Satz 4.15) ergibt sich auch

$$\dim(\mathbf{V}/\mathrm{kern}(f)) + \dim(\mathrm{kern}(f)) = \dim(\mathbf{V})$$

und

$$\dim \mathbf{V}/\mathbf{U} = \dim \mathbf{V} - \dim \mathbf{U}.$$

4.2 Matrizen

4.2.1 Addieren und Multiplizieren von Matrizen

Definition 4.17 *Unter einer $m \times n$-**Matrix** A mit Koeffizienten aus einem Körper K versteht man ein $m \cdot n$-Tupel $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ mit $a_{ij} \in K$, $1 \leq i \leq m, 1 \leq j \leq n$. Man stellt eine Matrix durch ein rechteckiges Schema*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

aus m Zeilen und n Spalten dar.

Die Menge aller $m \times n$-Matrizen mit Eintragungen aus K wird mit $K^{m \times n}$ bezeichnet.

Die einspaltigen Matrizen aus $K^{m \times 1}$ nennt man auch **Spaltenvektoren** *und die einzeiligen aus $K^{1 \times n}$ auch* **Zeilenvektoren**.

Sei $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in K^{m \times n}$. Dann heißen die Spaltenvektoren $\mathbf{a}_j := (a_{ij})_{1 \leq i \leq m} \in K^{m \times 1}$, $1 \leq j \leq n$, **Spalten** der Matrix A. Man schreibt auch

$$A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{pmatrix}.$$

Ensprechend heißen die Zeilenvektoren $\tilde{\mathbf{a}}_i := (a_{ij})_{1 \leq j \leq n} \in K^{1 \times n}$, $1 \leq i \leq m$, **Zeilen** der Matrix A:

$$A = \begin{pmatrix} \tilde{\mathbf{a}}_1 \\ \vdots \\ \tilde{\mathbf{a}}_m \end{pmatrix}.$$

Definition 4.18 *Eine Matrix $A \in K^{n\times n}$, bei der die Anzahl der Spalten gleich der Anzahl der Zeilen ist, heißt* **quadratisch**.

Definition 4.19 *Sind $A = (a_{ij})_{1\le i\le m,1\le j\le n}$ und $B = (b_{ij})_{1\le i\le m,1\le j\le n}$ zwei Matrizen aus $K^{m\times n}$, so bezeichnet man durch*

$$A + B := (a_{ij} + b_{ij})_{1\le i\le m,1\le j\le n} \in K^{m\times n}$$

die **Summe** *von A und B. Entsprechend definiert man für $x \in K$ und $A = (a_{ij})_{1\le i\le m,1\le j\le n} \in K^{m\times n}$*

$$xA := (xa_{ij})_{1\le i\le m,1\le j\le n} \in K^{m\times n}.$$

Satz 4.20 *Sei K ein Körper. Dann ist $K^{m\times n}$ ein Vektorraum über K der Dimension*

$$\dim(K^{m\times n}) = mn.$$

Insbesondere ist der Vektorraum $K^{m\times 1}$ der einspaltigen Matrizen (Spaltenvektoren) m-dimensional, und der Vektorraum $K^{1\times n}$ der einzeiligen Matrizen (Zeilenvektoren) n-dimensional.

Definition 4.21 *Sind $A = (a_{ij})_{1\le i\le m,1\le j\le n} \in K^{m\times n}$ und $B = (b_{jl})_{1\le j\le n,1\le l\le q} \in K^{n\times q}$ zwei Matrizen, wobei die Anzahl der Spalten der ersten gleich der Anzahl der Zeilen der zweiten ist, so wird durch*

$$(c_{il})_{1\le i\le m,1\le l\le q}$$

mit

$$c_{il} = \sum_{j=1}^{n} a_{ij}b_{jl} \qquad (1 \le i \le m, 1 \le l \le q),$$

eine Matrix in $K^{m\times q}$ definiert, die als **Produkt**

$$AB := (c_{il})_{1\le i\le m,1\le l\le q}$$

der Matrizen A und B bezeichnet wird.

Man beachte, daß das Element c_{il} in der i-ten Zeile von und l-ten Spalte von AB durch ein Matrizenprodukt[1] der i-ten Zeile von A und der l-ten Spalte von B gebildet wird:

$$c_{il} = \tilde{\mathbf{a}}_i\, \mathbf{b}_l = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1l} \\ b_{2l} \\ \vdots \\ b_{nl} \end{pmatrix}.$$

Satz 4.22 *Die Matrizenmultiplikation ist assoziativ, d.h. für $A \in K^{m\times n}$, $B \in K^{n\times q}$ und $C \in K^{q\times r}$ gilt*

$$(AB)C = A(BC).$$

[1] Das kann als Spezialfall eines *Skalarprodukts* gesehen werden.

Weiters ist die Matrizenmultiplikation bezüglich der Addition distributiv, d.h. für $A, B \in K^{m\times n}$ und $C \in K^{m\times q}$ gilt

$$(A+B)C = AC + AC$$

und für $A \in K^{m\times n}$ und $B, C \in K^{n\times q}$ gilt

$$A(B+C) = AB + AC.$$

Definition 4.23 *Sei $n \geq 1$ eine ganze Zahl. Unter der n-dimensionalen* **Einheitsmatrix** *$E_n \in K^{n\times n}$ versteht man die Matrix*

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

wobei δ_{ij} das sogenannte **Kroneckerdelta** *bezeichnet:*

$$\delta_{ij} := \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Die Spalten $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ der Einheitsmatrix bilden die **kanonische Basis**

$$E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$$

von $K^{n\times 1}$.

Lemma 4.24 *Für $A \in K^{m\times n}$ gilt*

$$AE_n = E_m A = A.$$

Satz 4.25 *Sei K ein Körper und $n \geq 1$ eine ganze Zahl. Dann ist $\langle K^{n\times n}, +, \cdot\rangle$ ein Ring mit 1. Für $n > 1$ ist $K^{n\times n}$ weder kommutativ noch nullteilerfrei.*

$K^{n\times n}$ ist daher sowohl ein Vektorraum (über K) als auch ein Ring. Man beachte, daß die Skalarmultiplikation und die Ringmultiplikation im folgenden Sinn verträglich sind:

$$(xA)B = A(xB) = x(AB)$$

Ein Vektorraum $\mathbf{V}$ (über einem Körper K), auf dem auch eine Multiplikation definiert ist, so daß $\langle \mathbf{V}, +, \cdot\rangle$ ein Ring ist und die beiden Multiplikationen im gerade erwähnten Sinn verträglich sind, heißt K-**Algebra**. $K^{n\times n}$ ist daher eine K-**Algebra**.

Genauso wie in allgemeinen Ringen mit 1 definiert man auch im Ring $K^{n\times n}$ die Einheitengruppe $(K^{n\times n})^*$ als jene Matrizen A, die invertierbar sind:

Definition 4.26 *Eine Matrix $A \in K^{n\times n}$ heißt* **invertierbar**, *wenn es eine Matrix A' mit $AA' = A'A = E_n$ gibt. $AA' = A'A = E_n$. Die Matrix A' wird* **inverse Matrix** *genannt und durch A^{-1} bezeichnet.*

4.2.2 Transponierte Matrix

Definition 4.27 *Für $A = (a_{ij})_{1\le i\le m, 1\le j\le n} \in K^{m\times n}$ bezeichnet man die Matrix*

$$A^T = (a'_{ji})_{1\le j\le n, 1\le i\le m} \in K^{n\times m}$$

mit

$$a'_{ji} = a_{ij} \qquad (1 \le i \le m, 1 \le j \le n)$$

als **transponierte Matrix**.

A^T entsteht aus A durch Vertauschen der Spalten und Zeilen. Die Zeilen von A^T sind die Spalten von A und umgekehrt.

Definition 4.28 *Eine quadratische Matrix $A = (a_{ij})_{1\le i,j\le n} \in K^{n\times n}$ heißt* **symmetrisch**, *wenn*

$$A = A^T,$$

d.h. $a_{ij} = a_{ji}$ $(1 \le i, j \le n)$.

Satz 4.29 *Für $A, B \in K^{m\times n}$, $C \in K^{n\times q}$ und $x \in K$ gelten:*

1. $(xA)^T = xA^T$.
2. $(A + B)^T = A^T + B^T$.
3. $(AC)^T = C^T A^T$.
4. $(A^{-1})^T = (A^T)^{-1}$, *falls $m = n$ und A invertierbar.*

4.2.3 Elementare Operationen auf Matrizen

Definition 4.30 *Sei $A = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix} \in K^{m\times n}$ mit den Spalten $\mathbf{a}_1, \ldots, \mathbf{a}_n \in K^{m\times 1}$. Die drei Operationen*

1. *Multiplikation einer Spalte $\mathbf{a}_j$ $(1 \le j \le n)$ mit einem Skalar $x \in K^\times$.*
2. *Addieren eines Vielfachen einer Spalte $\mathbf{a}_i$ $(1 \le i \le n)$ zu einer Spalte $\mathbf{a}_j$ $(1 \le j \le n,$ $i \ne j)$, d.h. Ersetzen der Spalte $\mathbf{a}_j$ durch $x\mathbf{a}_i + \mathbf{a}_j$ mit $x \in K$ und $i \ne j$.*
3. *Vertauschen zweier Spalten $\mathbf{a}_i, \mathbf{a}_j$ $(1 \le i \ne j \le n)$.*

heißen **elementare Spaltenumformungen** *der Matrix A.*

Sind $\tilde{\mathbf{a}}_1, \ldots, \mathbf{a}_m \in K^{1\times n}$ die Zeilen einer Matrix $A \in K^{m\times n}$, dann heißen die drei Operationen

1. *Multiplikation einer Zeile $\tilde{\mathbf{a}}_j$ $(1 \le j \le m)$ mit einem Skalar $x \in K^\times$.*
2. *Addieren eines Vielfachen einer Zeile $\tilde{\mathbf{a}}_i$ $(1 \le i \le m)$ zu einer Zeile $\tilde{\mathbf{a}}_j$ $(1 \le j \le m,$ $i \ne j)$, d.h. Ersetzen der Zeile $\tilde{\mathbf{a}}_j$ durch $x\tilde{\mathbf{a}}_i + \tilde{\mathbf{a}}_j$ mit $x \in K$ und $i \ne j$.*

3. Vertauschen zweier Zeilen $\tilde{\mathbf{a}}_i, \tilde{\mathbf{a}}_j$ *(*$1 \leq i \neq j \leq m$*).*

elementare Zeilenumformungen *der Matrix A.*

Satz 4.31 *Seien* $A = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix}, A' = \begin{pmatrix} \mathbf{a}'_1 & \cdots & \mathbf{a}'_n \end{pmatrix} \in K^{m\times n}$ *zwei Matrizen, wobei* A' *durch elementare Spaltenumformungen aus A gewonnen wird, so gilt*

$$[\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}] = [\{\mathbf{a}'_1, \ldots, \mathbf{a}'_n\}].$$

Satz 4.32 *Jede Matrix* $A \in K^{m\times n}$ *kann durch elementare Spaltenumformungen in eine Matrix umgeformt werden, die dieselben Zeilen hat wie die Matrix*

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ a'_{r+1,1} & a'_{r+1,2} & a'_{r+1,3} & \cdots & a'_{r+1,r} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a'_{m,1} & a'_{m,2} & a'_{m,3} & \cdots & a'_{m,r} & 0 & \cdots & 0 \end{pmatrix},$$

mit $a'_{ij} \in K$ $(r < i \leq m, 1 \leq j \leq r)$.

Zus Illustration dieses Satzes soll ein Beispiel vorgerechnet werden.

Beispiel 4.33 Man betrachte die Matrix

$$A = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 2 & 5 & 1 & 8 \\ -1 & -2 & 4 & 1 \\ 4 & 0 & 2 & 6 \end{pmatrix}$$

und bezeiche ihre Spalten durch $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$. Zunächst ersetzt man die zweite Spalte $\mathbf{a}_2$ durch $\mathbf{a}_2 - 2\mathbf{a}_1$ und die dritte durch $\mathbf{a}_3 + 3\mathbf{a}_1$:

$$A' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 7 & 8 \\ -1 & 0 & 1 & 1 \\ 4 & -8 & 14 & 6 \end{pmatrix}.$$

In dieser Matrix ersetzt man die erste Spalte $\mathbf{a}'_1$ durch $\mathbf{a}'_1 - 2\mathbf{a}_2$, die dritte durch $\mathbf{a}'_3 - 7\mathbf{a}'_2$ und die vierte durch $\mathbf{a}'_4 - 8\mathbf{a}'_2$:

$$A'' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 \\ 20 & -8 & 70 & 70 \end{pmatrix}.$$

Im letzten Schritt ersetzt man noch die erste Spalte $\mathbf{a}_1''$ durch $\mathbf{a}_1''+\mathbf{a}_3''$ und die vierte durch $\mathbf{a}_4''-\mathbf{a}_3''$ und erhält schließlich eine Matrix des gewünschten Typs:

$$A''' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 90 & -8 & 70 & 0 \end{pmatrix}.$$

Die Spalten der ursprünglichen Matrix A sind daher linear abhängig und spannen einen dreidimensionalen Unterraum von $K^{4\times 1}$ auf.

In diesem Beispiel war es nicht notwendig, Spalten zu vertauschen oder sie mit einem Skalar $\neq 0$ zu multiplizieren.

Um im folgenden die Notation zu vereinfachen, wird anstelle der in Satz 4.32 angegebenen Matrix

$$\begin{pmatrix} E_r & \mathbf{0}^{r\times(n-r)} \\ A'' & \mathbf{0}^{(m-r)\times(n-r)} \end{pmatrix}$$

geschrieben werden, wobei $\mathbf{0}^{k\times l}$ die Nullmatrix in $K^{k\times l}$ bezeichnet.

Die ersten r Spalten dieser Matrix sind linear unabhängig. Es gilt daher

$$r = \dim[\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}],$$

wobei $\mathbf{a}_1, \ldots, \mathbf{a}_n$ die Spalten von A bezeichnen.

Elementare Spaltenumformungen können daher verwendet werden, die Dimension r der linearen Hülle einer endlichen Auswahl von Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_n$ zu bestimmen. Aus dem Ergebnis kann man auch ablesen, ob die Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_n$ linear unabhängig sind oder nicht. Bei $r = n$ sind sie linear unabhängig und bei $r < n$ linear abhängig.

Elementare Spaltenumformungen können auch durch das Multiplizieren geeigneter Transformationsmatrizen realisiert werden:

1. $\begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & x\mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}$
   $= \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{j-1} & x\mathbf{a}_j & \mathbf{a}_{j+1} & \cdots & \mathbf{a}_n \end{pmatrix}$.
2. $\begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & x\mathbf{e}_i+\mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}$
   $= \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{j-1} & x\mathbf{a}_i+\mathbf{a}_j & \mathbf{a}_{j+1} & \cdots & \mathbf{a}_n \end{pmatrix}$.
3. $\begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix} \begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{i-1} & \mathbf{e}_j & \mathbf{e}_{i+1} & \cdots & \mathbf{e}_{j-1} & \mathbf{e}_i & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}$
   $= \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{a}_j & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_{j-1} & \mathbf{a}_i & \mathbf{a}_{j+1} & \cdots & \mathbf{a}_n \end{pmatrix}$.

Die rechtsstehenden Matrizen

1. $\begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & x\mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix} \quad (x \in K^\times)$,
2. $\begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & x\mathbf{e}_i+\mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix} \quad (i \neq j,\ x \in K)$,
3. $\begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{i-1} & \mathbf{e}_j & \mathbf{e}_{i+1} & \cdots & \mathbf{e}_{j-1} & \mathbf{e}_i & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix} \quad (i < j)$

werden auch als **Elementarmatrizen** bezeichnet. Sie entstehen übrigens aus der Einheitsmatrix E_n durch entsprechende elementare Spaltenumformugen. Anders ausgedrückt, bedeuten die obigen Bezieungen auch,

$$\mathcal{T}(A) = A \cdot \mathcal{T}(E_n),$$

wobei $\mathcal{T}$ eine elementare Spaltenumformung notiert.

Man beachte insbesondere, daß Elementarmatrizen invertierbar sind.

1. $\begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & x\mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}^{-1}$
   $= \begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & x^{-1}\mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}$.
2. $\begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & x\mathbf{e}_i + \mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}^{-1}$
   $= \begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{j-1} & -x\mathbf{e}_i + \mathbf{e}_j & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}$.
3. $\begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{i-1} & \mathbf{e}_j & \mathbf{e}_{i+1} & \cdots & \mathbf{e}_{j-1} & \mathbf{e}_i & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}^{-1}$
   $= \begin{pmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_{i-1} & \mathbf{e}_j & \mathbf{e}_{i+1} & \cdots & \mathbf{e}_{j-1} & \mathbf{e}_i & \mathbf{e}_{j+1} & \cdots & \mathbf{e}_n \end{pmatrix}$.

Das Umformen einer Matrix A in eine Matrix A' mittels elementarer Spaltenumformungen entpricht also der Multiplikation mit einer invertierbaren Matrix T,

$$A' = AT,$$

wobei T Produkt geeigneter Transformationsmatrizen ist. Satz 4.32 kann daher folgendermaßen *umformuliert* werden.

Satz 4.34 *Zu jeder Matrix $A \in K^{m\times n}$ gibt es eine invertierbare Matrix $T \in K^{n\times n}$, die als Produkt von Elementarmatrizen, die elementaren Spaltenumformungen entsprechen, darstellbar ist, so daß AT dieselben Zeilen wie die Matrix*

$$\begin{pmatrix} E_r & \mathbf{0}^{r\times(n-r)} \\ A'' & \mathbf{0}^{(m-r)\times(n-r)} \end{pmatrix},$$

hat, wobei r die Dimension des von den Spalten von A aufgespannten Raums ist und A'' eine Matrix in $K^{(m-r)\times r}$ bezeichnet.

Für quadratische Matrizen, deren Spalten linear unabhängig sind, läßt sich dieser Satz etwas verschärfen.

Satz 4.35 *Sei $A \in K^{n\times n}$ eine quadratische Matrix, deren Spalten linear unabhängig sind. Dann gibt es eine invertierbare Matrix $T \in K^{n\times n}$, die als Produkt von Elementarmatrizen, die elementaren Spaltenumformungen entsprechen, darstellbar ist, mit*

$$AT = E_n.$$

Anders ausgedrückt erhält man auch das folgende Kriterium:

Satz 4.36 *Eine Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn sie als Produkt von Elementarmatrizen darstellbar ist.*

Alle Eigenschaften über elementare Spaltenoperationen gelten natürlich sinngemäß auch für Zeilenoperationen. Besonders einfach wird die Situation, wenn man Spalten- und Zeilenoperationen zuläßt.

Satz 4.37 *Zu jeder Matrix $A \in K^{m\times n}$ gibt es eine invertierbare Matrizen $U \in K^{m\times m}$, $V \in K^{n\times n}$ mit*

$$UAV = \begin{pmatrix} E_r & \mathbf{0}^{r\times(n-r)} \\ \mathbf{0}^{(m-r)\times r} & \mathbf{0}^{(m-r)\times(n-r)} \end{pmatrix},$$

wobei r die Dimension des von den Spalten von A aufgespannten Raums ist.

4.2.4 Matrizen und $L(K^{n\times 1}, K^{m\times 1})$

Satz 4.38 *Sei $A \in K^{m\times n}$. Dann ist die Abbildung*

$$f_A : K^{n\times 1} \to K^{m\times 1}, \quad \mathbf{v} \mapsto A\mathbf{v}$$

eine lineare Abbildung. Umgekehrt gibt es zu jeder linearen Abbildung $f \in L(K^{n\times 1}, K^{m\times 1})$ eine eindeutig bestimmte Matrix $A \in K^{m\times n}$ mit $f = f_A$. Die Spalten von A sind die Bilder der kanonischen Basis $\mathbf{e}_1, \ldots, \mathbf{e}_n \in K^{n\times 1}$ unter f:

$$A = \begin{pmatrix} f(\mathbf{e}_1) & \cdots & f(\mathbf{e}_n) \end{pmatrix}.$$

Korollar 4.39 *Die Vektorräume $K^{m\times n}$ und $L(K^{n\times 1}, K^{m\times 1})$ sind isomorph. Die Abbildung*

$$\Phi : K^{m\times n} \to L(K^{n\times 1}, K^{m\times 1}), \quad A \mapsto f_A$$

ist ein Vektorraumisomorphismus. Für die Dimensionen gilt

$$\dim L(K^{n\times 1}, K^{m\times 1}) = \dim K^{m\times n} = mn.$$

Damit ist auch Satz 4.13 bewiesen.

Satz 4.40 *Seien $A \in K^{m\times n}$ und $B \in K^{n\times q}$. Dann gilt*

$$f_{AB} = f_A \circ f_B.$$

Satz 4.41 *Eine lineare Abbildung $f \in L(K^{n\times 1}, K^{n\times 1})$ ist genau dann bijektiv, wenn die f entsprechende Matrix $A \in K^{n\times n}$ invertierbar ist.*

Definition 4.42 *Eine Matrix $A \in K^{n\times n}$ heißt* **regulär**, *wenn es eine Matrix $A' \in K^{n\times n}$ mit*

$$AA' = E_n$$

gibt. Eine nicht-reguläre Matrix $A \in K^{n\times n}$ heißt **singulär**.

Satz 4.43 *Eine Matrix $A \in K^{n\times n}$ ist genau dann regulär, wenn sie invertierbar ist, d.h. aus $AA' = E_n$ folgt $A'A = E_n$ und $A' = A^{-1}$.*

Kombiniert man diesen Satz mit Satz 4.35, so ergibt sich auch:

Satz 4.44 *Jede invertierbare Matrix $A \in K^{n\times n}$ kann als Produkt von Elementarmatrizen, die elementaren Spaltenumformungen entsprechen, dargestellt werden.*

Ebenso läßt sich jede jede invertierbare Matrix auch als Produkt von Elementarmatrizen, die elementaren Zeilenumformungen entsprechen, darstellen.

Man kann diesen Satz auch dafür ausnützen, explizit die inverse Matrix A^{-1} einer regulären Matrix A zu ermitteln. Das Spaltenumformen von A zur Einheitsmatrix E_n ist nichts anderes als das Multiplizien mit A^{-1}: $AA^{-1} = E_n$. Wendet man daher dieselben Spaltenumformungen auf die Einheitsmatrix E_n an, so gewinnt man $E_nA^{-1} = A^{-1}$. Diese beiden Vorgänge können gleichzeitig durchgeführt werden, indem man etwa A und E_n untereinander schreibt und zu einer $2n \times n$-Matrix zusammenfaßt.

Beispiel 4.45 Es soll die inverse Matrix A^{-1} von

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 3 & 5 & 0 \end{pmatrix}$$

bestimmt werden. Durch elementare Spaltenumformungen erhält man

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 3 & 5 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 3 & -1 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 3 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & 0 & 2 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -5 & 0 & 2 \\ 3 & 0 & -1 \\ -6 & 1 & 2 \end{pmatrix}.$$

Die inverse Matrix A^{-1} ist also

$$A^{-1} = \begin{pmatrix} -5 & 0 & 2 \\ 3 & 0 & -1 \\ -6 & 1 & 2 \end{pmatrix}.$$

4.2.5 Der Rang einer Matrix

Definition 4.46 *Der* **Rang** *(oder* **Spaltenrang***)* $\mathrm{rg}(A)$ *einer Matrix $A \in K^{m\times n}$ ist die maximale Anzahl von linear unabhängigen Spalten von A.*

Der **Zeilenrang** *einer Matrix $A \in K^{m\times n}$ ist die maximale Anzahl von linear unabhängigen Zeilen von A.*

Satz 4.47 *Für eine Matrix* $A = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix} \in K^{m\times n}$ *gilt*

$$\mathrm{rg}(A) = \dim[\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}] = \mathrm{rg}(f_A),$$

wobei $f_A \in L(K^{n\times 1}, K^{n\times 1})$ *die* A *entsprechende Abbildung ist.*

Satz 4.48 *Sei* $A \in K^{m\times n}$*. Dann ist die* A *entsprechende Abbildung* $f_A \in L(K^{n\times 1}, K^{m\times n})$ *genau dann surjektiv, wenn* $\mathrm{rg}(A) = m$*, und* f_A *ist genau dann injektiv, wenn* $\mathrm{rg}(A) = n$*.*

Satz 4.49 *Eine quadratische Matrix* $A \in K^{n\times n}$ *ist genau dann regulär, wenn* $\mathrm{rg}(A) = n$*.*

Satz 4.50 *Sei* $f \in L(V, W)$ *eine lineare Abbildung und seien* $g \in L(V_1, V)$ *und* $h \in L(W, W_1)$ *Isomorphismen. Dann gilt*

$$\mathrm{rg}(h \circ f \circ g) = \mathrm{rg}(f).$$

Der Satz gilt auch, wenn g surjektiv und h injektiv ist.

Satz 4.51 *Sei* $A \in K^{m\times n}$ *und seinen* $U \in K^{m\times m}$ *und* $V \in K^{n\times n}$ *reguläre Matrizen. Dann gilt*

$$\mathrm{rg}(UAV) = \mathrm{rg}(A).$$

Mit Hilfe des Begriffes des Rangs einer Matrix gewinnen die Sätze aus Abschnitt 4.2.3 eine neue Interpretation.

Satz 4.52 *Eine Matrix* $A \in K^{m\times n}$ *hat genau dann* $\mathrm{rg}(A) = r$*, wenn es eine invertierbare Matrix* $T \in K^{n\times n}$ *gibt, so daß* AT *dieselben Zeilen wie die Matrix*

$$\begin{pmatrix} E_r & \mathbf{0}^{r\times(n-r)} \\ A'' & \mathbf{0}^{(m-r)\times(n-r)} \end{pmatrix},$$

hat, wobei A'' *eine Matrix in* $K^{(m-r)\times r}$ *bezeichnet.*

Satz 4.53 *Eine Matrix* $A \in K^{m\times n}$ *hat genau dann* $\mathrm{rg}(A) = r$*, wenn es invertierbare Matrizen* $U \in K^{m\times m}$*,* $V \in K^{n\times n}$ *mit*

$$UAV = \begin{pmatrix} E_r & \mathbf{0}^{r\times(n-r)} \\ \mathbf{0}^{(m-r)\times r} & \mathbf{0}^{(m-r)\times(n-r)} \end{pmatrix}$$

gibt.

Wendet man diesen Satz für die transponierte Matrix A^T, so erhält man folgendes überraschende Resultat.

Satz 4.54 *Für jede Matrix* $A \in K^{m\times n}$ *stimmen Spalten- und Zeilenrang überein.*

Man kann daher immer vom *Rang* einer Matrix sprechen und muß nicht zwischen Spalten- und Zeilenrang unterscheiden. Der tiefere Grund für diesen Sachverhalt wird in der Theorie der dualen Vektorräume (siehe Abschnitt 6.3) erläutert.

4.2.6 Äquivalente Matrizen

Definition 4.55 *Zwei Matrizen $A, B \in K^{m\times n}$ heißen* **äquivalent**, *wenn es invertierbare Matrizen $U \in K^{m\times m}$, $V \in K^{n\times n}$ mit*

$$B = UAV$$

gibt.

Aus Satz 4.51 ergibt sich sofort das folgende Kriterium.

Satz 4.56 *Zwei Matrizen $A, B \in K^{m\times n}$ sind genau dann äquivalent, wenn sie denselben Rang haben.*

4.3 Matrix einer linearen Abbildung

4.3.1 Lineare Abbildungen zwischen endlichdimensionalen Vektorräumen

Definition 4.57 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$. Die Abbildung*

$$\begin{array}{rcl} \Phi_{\mathbf{B}} : \mathbf{V} & \to & K^{n\times 1} \\ \mathbf{a} = \sum_{i=1}^{n} x_i \mathbf{b}_i & \mapsto & \sum_{i=1}^{n} x_i \mathbf{e}_i = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{array}$$

bezeichet man als **Koordinatisierung** *von $\mathbf{V}$ bezüglich (der Basis) $\mathbf{B}$. $x_1, x_2, \ldots, x_n \in K$ heißen* **Koordinaten** *von $\mathbf{a}$ bezüglich (der Basis) $\mathbf{B}$.*

Die folgende wichtige Eigenschaft von $\Phi_{\mathbf{B}}$ ist schon im Satz 3.31 enthalten.

Satz 4.58 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$. Dann ist die Abbildung $\Phi_{\mathbf{B}} : \mathbf{V} \to K^{n\times 1}$ ein Vektorraumisomorphismus.*

Definition 4.59 *Sei $f : \mathbf{V} \to \mathbf{W}$ eine lineare Abbildung zwischen zwei endlichdimensionalen Vektorräumen, und sei $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$ und $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m\}$ eine Basis von $\mathbf{W}$. Sind die Skalare a_{ij} $(1 \le i \le m, 1 \le j \le n)$ durch*

$$f(\mathbf{b}_j) = \sum_{i=1}^{m} a_{ij} \mathbf{c}_i \qquad (1 \le j \le n)$$

gegeben, so bezeichnet man mit

$$\Phi_{\mathbf{BC}}(f) := (a_{ij})_{1\le i\le m, 1\le j\le n}$$

die **Koordinatenmatrix** *der linearen Abbildung $f : \mathbf{V} \to \mathbf{W}$ bezüglich der Basen $\mathbf{B}$ und $\mathbf{C}$.*

Lemma 4.60 *Die Koordinatenmatrix $\Phi_{\mathbf{BC}}$ ist jene Matrix aus $K^{m\times n}$, die der Abbildung $\Phi_{\mathbf{C}} \circ f \circ \Phi_{\mathbf{B}}^{-1} : K^{n\times 1} \to K^{m\times 1}$ entspricht, d.h. für $\mathbf{a} \in \mathbf{V}$ gilt*

$$\Phi_{\mathbf{BC}}(f)\, \Phi_{\mathbf{B}}(\mathbf{a}) = \Phi_{\mathbf{C}}(f(\mathbf{a})).$$

Die j-te Spalte von $\Phi_{\mathbf{BC}}(f)$ enthält genau die Koordinaten von $f(\mathbf{b}_j)$ bezüglich $\mathbf{C}$.

$$\begin{array}{ccc} \mathbf{V} & \xrightarrow{f} & \mathbf{W} \\ {\scriptstyle \Phi_{\mathbf{B}}}\downarrow & & \downarrow{\scriptstyle \Phi_{\mathbf{C}}} \\ K^{n\times 1} & \xrightarrow{\Phi_{\mathbf{BC}}(f)} & K^{m\times 1} \end{array}$$

Satz 4.61 *Seien $\mathbf{V}, \mathbf{W}, \mathbf{X}$ endlichdimensionale Vektorräume, und sei $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$, $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m\}$ eine Basis von $\mathbf{W}$ und $\mathbf{D} = \{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_q\}$ eine Basis von $\mathbf{X}$. Dann gilt für alle linearen Abbildungen $f \in L(\mathbf{V}, \mathbf{W})$ und $g \in L(\mathbf{W}, \mathbf{X})$*

$$\Phi_{\mathbf{BD}}(g \circ f) = \Phi_{\mathbf{CD}}(g)\, \Phi_{\mathbf{BC}}(f).$$

Man beachte, daß die Abbildung

$$\Phi_{\mathbf{BC}} : L(\mathbf{V}, \mathbf{W}) \to K^{m\times n}$$

ein Vektorraumisomorphismus ist (was einen erneuten Beweis von Satz 4.13 liefert). Insbesondere ist $\Phi_{\mathbf{BC}}$ mit der Addition verträglich. Für Matrizen gelten die Distributivgesetze

$$(A + B)C = AB + BC \quad \text{und} \quad A(B + C) = AB + AC.$$

Deshalb ist zu erwarten, daß für lineare Abbildungen auch Distributivgesetze erfüllt sind:

$$(f + g) \circ h = f \circ h + g \circ h \quad \text{und} \quad f \circ (g + h) = f \circ g + f \circ h.$$

Dies ist tatsächlich der Fall. Während die erste Beziehung $(f + g) \circ h = f \circ h + g \circ h$ für alle (auch nichtlinearen) Abbildung erfüllt ist, gilt die zweite Beziehung $f \circ (g + h) = f \circ g + f \circ h$ nicht allgemein. Bei linearem f ist sie aber auch erfüllt.

Die linearen Abbildungen aus $L(\mathbf{V}, \mathbf{V})$ bilden daher (ähnlich wie die quadratischen Matrizen $K^{n\times n}$) eine K-Algebra. Für $\dim \mathbf{V} = n$ und $\dim \mathbf{W} = m$ sind $L(\mathbf{V}, \mathbf{W})$ und $K^{m\times n}$ auch als K-Algebren isomorph.

4.3.2 Basiswechsel

Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum und seien $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, $\tilde{\mathbf{B}} = \{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_n\}$ zwei Basen von $\mathbf{V}$. Es soll nun die Frage untersucht werden, wie sich die Koordinaten eines Vektors $\mathbf{a} \in \mathbf{V}$ verändern, wenn man von der Basis $\mathbf{B}$ zur Basis $\tilde{\mathbf{B}}$ wechselt. Formal ist dieser **Basiswechsel** durch den Isomorphismus $\Phi_{\tilde{\mathbf{B}}} \circ \Phi_{\mathbf{B}}^{-1} : K^{n\times 1} \to K^{n\times 1}$ gegeben.

Definition 4.62 *Seinen $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ und $\tilde{\mathbf{B}} = \{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_n\}$ zwei Basen eines n-dimensionalen Vektorraums $\mathbf{V}$ und sei*

$$\mathbf{b}_j = \sum_{i=1}^{n} c_{ij} \tilde{\mathbf{b}}_i \quad (1 \le j \le n),$$

so bezeichnet man mit

$$\mathbf{T}_{\mathbf{B}\tilde{\mathbf{B}}} = (c_{ij})_{1\leq i,j\leq n} \in K^{n\times n}$$

die Basistransformationsmatrix zwischen den Matrizen $\mathbf{B}$ *und* $\tilde{\mathbf{B}}$*.*

Satz 4.63 *Sind* $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ *und* $\tilde{\mathbf{B}} = \{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2, \ldots, \tilde{\mathbf{b}}_n\}$ *zwei Basen eines n-dimensionalen Vektorraums* $\mathbf{V}$*, dann gilt*

$$\mathbf{T}_{\mathbf{B}\tilde{\mathbf{B}}} = \Phi_{\mathbf{B}\tilde{\mathbf{B}}}(\mathrm{id}),$$

d.h. die Basistransformationsmatrix $\mathbf{T}_{\mathbf{B}\tilde{\mathbf{B}}}$ *ist die Koordinatenmatrix der identischen Abbildung bezüglich der Basen* $\mathbf{B}$ *und* $\tilde{\mathbf{B}}$*. Weiters gilt*

$$\Phi_{\tilde{\mathbf{B}}}(\mathbf{a}) = \mathbf{T}_{\mathbf{B}\tilde{\mathbf{B}}} \cdot \Phi_{\mathbf{B}}(\mathbf{a}),$$

d.h. $\mathbf{T}_{\mathbf{B}\tilde{\mathbf{B}}}$ *führt den Koordinatenwechsel von* $\mathbf{B}$ *nach* $\tilde{\mathbf{B}}$ *mittels Matrizenmultiplikation durch.*

$$\begin{array}{ccc} \mathbf{V} & \xrightarrow{\mathrm{id}} & \mathbf{V} \\ \Phi_{\mathbf{B}} \downarrow & & \downarrow \Phi_{\tilde{\mathbf{B}}} \\ K^{n\times 1} & \xrightarrow{\mathbf{T}_{\mathbf{B}\tilde{\mathbf{B}}}} & K^{n\times 1} \end{array}$$

Satz 4.64 *Jede Basistransformationsmatrix* $\mathbf{T}_{\mathbf{B}\tilde{\mathbf{B}}}$ *ist regulär. Es gilt*

$$\mathbf{T}^{-1}_{\mathbf{B}\tilde{\mathbf{B}}} = \mathbf{T}_{\tilde{\mathbf{B}}\mathbf{B}}.$$

Weiters gibt ist zu jeder regulären Matrix $\mathbf{C} \in K^{n\times n}$ *und zu jeder Basis* $\mathbf{B}$ *eines n-dimensionalen Vektorraums* $\mathbf{V}$ *eine Basis* $\tilde{\mathbf{B}}$ *mit* $\mathbf{T}_{\tilde{\mathbf{B}}\mathbf{B}} = \mathbf{C}$*.*

Satz 4.65 *Sei* $f \in L(\mathbf{V}, \mathbf{W})$ *eine lineare Abbildung zwischen zwei endlich dimensionalen Vektoräumen* $\mathbf{V}, \mathbf{W}$*. Sind* $\mathbf{B}, \tilde{\mathbf{B}}$ *zwei Basen von* $\mathbf{V}$ *und* $\mathbf{C}, \tilde{\mathbf{C}}$ *zwei Basen von* $\mathbf{W}$*, so gilt*

$$\Phi_{\tilde{\mathbf{B}}\tilde{\mathbf{C}}}(f) = \mathbf{T}_{\mathbf{C}\tilde{\mathbf{C}}}\, \Phi_{\mathbf{B}\mathbf{C}}(f)\, \mathbf{T}_{\tilde{\mathbf{B}}\mathbf{B}}.$$

$$\begin{array}{ccccccc} \mathbf{V} & \xrightarrow{\mathrm{id}} & \mathbf{V} & \xrightarrow{f} & \mathbf{W} & \xrightarrow{\mathrm{id}} & \mathbf{W} \\ \Phi_{\tilde{\mathbf{B}}} \downarrow & & \downarrow \Phi_{\mathbf{B}} & & \Phi_{\mathbf{C}} \downarrow & & \downarrow \Phi_{\tilde{\mathbf{C}}} \\ K^{n\times 1} & \xrightarrow{\mathbf{T}_{\tilde{\mathbf{B}}\mathbf{B}}} & K^{n\times 1} & \xrightarrow{\Phi_{\mathbf{B}\mathbf{C}}(f)} & K^{m\times 1} & \xrightarrow{\mathbf{T}_{\mathbf{C}\tilde{\mathbf{C}}}} & K^{m\times 1} \end{array}$$

Multipliziert man eine Matrix $A \in K^{m\times n}$ mit regulären Matrizen $U \in K^{m\times m}$, $V \in K^{n\times n}$:

$$UAV,$$

so kann dies als als Koordinatenmatrix von f_A bezüglich anderer Basen in $K^{m\times 1}$ und $K^{n\times 1}$ interpretiert werden. Äquivalente Matrizen können daher auch als Koordinatenmatrizen ein und derselben Abbildung aufgefaßt werden.

4.4 Lineare Gleichungssysteme

4.4.1 Lineare Gleichungssysteme und lineare Abbildungen

Definition 4.66 *Seien $m, n \geq 1$ ganze Zahlen und K ein Körper. Weiters seien $a_{ij} \in K$ $(1 \leq i \leq m, 1 \leq j \leq n)$ und $b_i \in K$ $(1 \leq i \leq m)$ Elemente in K Dann heißt ein System der Form*

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{12}x_1 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & \vdots & & & \vdots \\ a_{1m}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

Lineares Gleichungssystem *in den Unbekannten $x_1, x_2 \ldots, x_n \in K$.*

Sind alle $b_1 = b_2 = \cdots = b_m = 0$, so heißt das lineare Gleichungssystem **homogen**, *sonst* **inhomogen**.

Es besteht nun die Aufgabe, ein lineares Gleichungssystem vollständig zu lösen, d.h. alle n-Tupel $(x_1, x_2, \ldots, x_n) \in K^n$ anzugeben, die das obige Gleichungssystem erfüllen.

Faßt man die **Koeffizienten** a_{ij} zu einer Matrix $A = (a_{ij}) \in K^{m \times n}$ zusammen, und entsprechend auch die **rechte Seite** $b_1, b_2, \ldots, b_m$ zu einer Spaltenmatrix $\mathbf{b} \in K^{m \times 1}$ und die Unbekannten $x_1, x_2, \ldots, x_n$ zu einer Spaltenmatrix $\mathbf{x} \in K^{n \times 1}$, so läßt sich ein lineares Gleichungssystem auch mit Hilfe des Matrizenprodukts durch

$$A\mathbf{x} = \mathbf{b}$$

darstellen, bzw. auch mit der A entsprechenden linearen Abbildung $f_A \in L(K^{n \times 1}, K^{m \times 1})$ durch

$$f_A(\mathbf{x}) = \mathbf{b}.$$

Das heißt, die Lösungsmenge ist durch

$$f_A^{-1}(\{\mathbf{b}\})$$

gegeben.

Lemma 4.67 *Sei $f \in L(\mathbf{V}, \mathbf{W})$ eine lineare Abbildung und $\mathbf{x}_0 \in \mathbf{V}$ und $\mathbf{b} \in \mathbf{W}$ mit $f(\mathbf{x}_0) = \mathbf{b}$. Dann ist*

$$f^{-1}(\{\mathbf{b}\}) = \mathbf{x}_0 + \operatorname{kern}(f)$$

ein Nebenraum des Kerns von f.

4.4.2 Lösbarkeitskriterien

Im folgenden wird ein lineares Gleichunssystem kurz durch $A\mathbf{x} = \mathbf{b}$ dargestellt, und $f_A : K^{n \times 1} \to K^{m \times 1}$ bezeichnet die A entsprechende lineare Abbildung (mit $f_A(\mathbf{x}) = A\mathbf{x}$).

Satz 4.68 *Sei $A \in K^{m\times n}$ und $\mathbf{b} \in K^{m\times 1}$.*

Das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ ist genau dann lösbar, wenn $\mathbf{b} \in f_A(K^{n\times 1})$.

Ist f_A surjektiv, d.h. $\mathrm{rg}(f_A) = \mathrm{rg}(A) = m$, so ist das lineare Gleichungssystem für alle rechten Seiten $\mathbf{b} \in K^{m\times 1}$ lösbar.

Ist f_A injektiv, d.h. $\mathrm{rg}(f_A) = \mathrm{rg}(A) = n$, so hat das lineare Gleichungssystem höchstens eine Lösung.

Satz 4.69 (Kronecker-Capelli) *Sei $A \in K^{m\times n}$ und $\mathbf{b} \in K^{m\times 1}$. Dann ist das lineares Gleichungssystem $A\mathbf{x} = \mathbf{b}$ genau dann lösbar, wenn*

$$\mathrm{rg}(A) = \mathrm{rg}\begin{pmatrix} A & \mathbf{b} \end{pmatrix}.$$

Korollar 4.70 *Ist der Rang einer Matrix $A \in K^{m\times n}$ gleich der Anzahl der Zeilen, d.h. $\mathrm{rg}(A) = m$, so ist das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ für jede rechte Seite $\mathbf{b} \in K^{m\times 1}$ lösbar.*

Korollar 4.71 *Ist der Rang einer Matrix $A \in K^{m\times n}$ gleich der Anzahl der Spalten, d.h. $\mathrm{rg}(A) = n$, so hat das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ für jede rechte Seite $\mathbf{b} \in K^{m\times 1}$ höchstens eine Lösung.*

Korollar 4.72 *Ist $A \in K^{n\times n}$ regulär, so ist das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ für jede rechte Seite $\mathbf{b} \in K^{n\times 1}$ eindeutig lösbar. Die Lösung ist durch*

$$\mathbf{x} = A^{-1}\mathbf{b}$$

gegeben.

Satz 4.73 *Sei $A \in K^{m\times n}$ und $\mathbf{b} \in K^{m\times 1}$. Ist das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ lösbar, so gibt es $r' := n - \mathrm{rg}(A)$ linear unabhängige Vektoren $\mathbf{x}_1, \ldots, \mathbf{x}_{r'} \in \mathrm{kern}(f_A) \subseteq K^{n\times 1}$, d.h. Lösungen des homogenen linearen Gleichungssystems $A\mathbf{x} = \mathbf{0}$, so daß alle Lösungen von $A\mathbf{x} = \mathbf{b}$ durch die Menge*

$$\{\mathbf{x}_0 + t_1\mathbf{x}_1 + \cdots + t_{r'}\mathbf{x}_{r'} \mid t_1, \ldots, t_r \in K\}$$

gegeben sind, wobei $\mathbf{x}_0$ eine beliebige, aber fest gewählte Lösung von $A\mathbf{x} = \mathbf{b}$ ist.

Man beachte, daß die Vektoren $\mathbf{x}_1, \ldots, \mathbf{x}_{r'}$ eine Basis des Kerns von f_A sind, d.h. sie lösen alle das homogene lineare Gleichungssystem $A\mathbf{x} = \mathbf{0}$. Es reicht daher, eine Lösung $\mathbf{x}_0$ des inhomogenen linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ und $r' = n - \mathrm{rg}(A)$ linear unabhängige Lösugen des homogenen Gleichungssystems $A\mathbf{x} = \mathbf{0}$ zu finden.

4.4.3 Gaußsches Eliminationsverfahren

Das Lösen eines linearen Gleichungssystems ist besonders einfach, wenn die Matrix A von spezieller Gestalt ist.

Satz 4.74 *Sei $A \in K^{m\times n}$ $(n \geq m)$ eine Matrix der Gestalt*

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 & a_{1,m+1} & \cdots & a_{1,n} \\ 0 & 1 & \cdots & 0 & a_{2,m+1} & \cdots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & a_{m,m+1} & \cdots & a_{n,m} \end{pmatrix} = \begin{pmatrix} E_m & A' \end{pmatrix}$$

(mit $A' \in K^{m\times(n-m)}$), so sind alle Lösungen $\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^T \in K^{n\times 1}$ des Gleichungssystems $A\mathbf{x} = \mathbf{b}$ (mit $\mathbf{b} \in K^{m\times 1}$) durch

$$\mathbf{x} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} -A' \\ E_{n-m} \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_{n-m} \end{pmatrix}$$

resp. durch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \\ x_{m+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} + t_1 \begin{pmatrix} -\mathbf{a}_{m+1} \\ \mathbf{e}_1 \end{pmatrix} + t_2 \begin{pmatrix} -\mathbf{a}_{m+2} \\ \mathbf{e}_2 \end{pmatrix} + \cdots + t_{n-m} \begin{pmatrix} -\mathbf{a}_n \\ \mathbf{e}_{n-m} \end{pmatrix}$$

mit $t_1, t_2, \ldots, t_{n-m} \in K$ gegeben. Dabei bezeichnen $\mathbf{a}_{m+1}, \ldots, \mathbf{a}_n$ die Spalten von A', $\mathbf{0}$ den Nullvektor in $K^{(n-m)\times 1}$ und $\mathbf{e}_1, \ldots, \mathbf{e}_{n-m}$ die Vektoren der kanonischen Basis von $K^{(n-m)\times 1}$.

In diesem Fall können sozusagen die Komponenten $x_{m+1}, \ldots, x_n$ als *Parameter* $t_1, \ldots, t_{n-m}$ verwendet werden, also beliebig und unabhängig voneinander gewählt werden. Die übrigen Komponenten $x_1, \ldots, x_n$ lassen sich darauf in einfacher Weise aus $x_{m+1} = t_1, \ldots, x_n = t_{n-m} \in K$ berechnen.

Dieses Verfahren funktioniert auch, wenn die Matrix A Dreicksgestalt mit nichtverschwindender Diagonale hat.

Satz 4.75 *Sei $A \in K^{m\times n}$ $(n \geq m)$ eine Matrix der Gestalt*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} & a_{1,m+1} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,m} & a_{2,m+1} & \cdots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{m,m} & a_{m,m+1} & \cdots & a_{n,m} \end{pmatrix}$$

mit $a_{1,1} \neq 0, a_{2,2} \neq 0, \ldots, a_{m,m} \neq 0$. Dann gibt es eindeutig bestimmte Spaltenvektoren $\mathbf{b}', \mathbf{a}'_1, \mathbf{a}'_2, \ldots, \mathbf{a}'_{n-m} \in K^{m\times 1}$, so daß alle Lösungen $\mathbf{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^T \in K^{n\times 1}$ des Glei-

chungssystems $A\mathbf{x} = \mathbf{b}$ *(mit* $\mathbf{b} \in K^{m\times 1}$ *durch*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_m \\ x_{m+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \mathbf{b}' \\ \mathbf{0} \end{pmatrix} - t_1 \begin{pmatrix} \mathbf{a}'_1 \\ \mathbf{e}_1 \end{pmatrix} - t_2 \begin{pmatrix} \mathbf{a}'_2 \\ \mathbf{e}_2 \end{pmatrix} - \cdots - t_{n-m} \begin{pmatrix} \mathbf{a}'_{n-m} \\ \mathbf{e}_{n-m} \end{pmatrix}$$

mit $t_1, t_2, \ldots, t_{n-m} \in K$ *gegeben sind.* $\mathbf{0}$ *bezeichnet wieder den Nullvektor in* $K^{(n-m)\times 1}$ *und* $\mathbf{e}_1, \ldots, \mathbf{e}_{n-m}$ *die Vektoren der kanonischen Basis von* $K^{(n-m)\times 1}$*.*

Wieder Fall können die Komponenten $x_{m+1}, \ldots, x_n$ als *Parameter* $t_1, \ldots, t_{n-m}$ verwendet werden. Damit läßt sich zunächst

$$\begin{aligned} x_m &= a_{m,m}^{-1}(b_m - t_1 a_{m,m+1} - \cdots t_{n-m} a_{m,n}) \\ &= b'_m + t_1 a'_{1,m} + \cdots + t_{n-m} a'_{n-m,m} \end{aligned}$$

direkt berechnen. Mit dieser Kenntnis ermittelt man

$$\begin{aligned} x_{m-1} &= a_{m-1,m-1}^{-1}(b_{m-1} - a_{m-1,m} x_m - t_1 a_{m-1,m+1} - \cdots t_{n-m} a_{m-1,n}) \\ &= b'_{m-1} + t_1 a'_{1,m-1} + \cdots + t_{n-m} a'_{n-m,m-1} \end{aligned}$$

und danach rekursiv $x_{m-2}, x_{m-3}, \ldots, x_1$.

Man beachte, daß dasselbe Verfahren auch dann funktioniert, wenn die Matrix A bis auf die Reihenfolge der Spalten von der Gestalt aus Satz 4.75 ist, d.h. $A = A'T$, wobei $T \in K^{n\times n}$ eine Spaltentransformationsmatrix ist, die nur Spalten vertauscht, d.h. $T^{-1} = T^T$, und $A' \in K^{m\times n}$ hat die Gestalt aus Satz 4.75. Das Gleichungssystem

$$A'\mathbf{x}' = \mathbf{b}$$

läßt sich nun wie zuvor lösen, und die Lösungen von $A\mathbf{x} = \mathbf{b}$ sind dann durch

$$\mathbf{x} = T^{-1}\mathbf{x}' = T^T\mathbf{x}'$$

gegeben, d.h. $\mathbf{x}$ ergibt sich aus $\mathbf{x}'$, indem gewisse Komponenten vertauscht werden. T^T ist in diesem Fall als Zeilentransformationsmatrix zu interpretieren, die genau dieselben Zeilen vertauscht wie T Spalten.

Das nächste Ziel ist es, ein beliebiges Gleichungssystem in ein System überzuführen, das bis auf die Reihenfolge der Spalten von der Form aus Satz 4.75 ist.

Definition 4.76 *Zwei lineare Gleichungssysteme heißen* **äquivalent***, wenn sie dieselben Lösungen haben.*

Satz 4.77 *Sei* $A \in K^{m\times n}$*,* $\mathbf{b} \in K^{m\times 1}$ *und* $U \in K^{m\times m}$ *eine reguläre Matrix. Dann sind die linearen Gleichungssysteme*

$$A\mathbf{x} = \mathbf{b} \quad \textit{und} \quad (UA)\mathbf{x} = (U\mathbf{b})$$

äquivalent.

Mit anderen Worten, elementare Zeilenumformungen der erweiterten Matrix $\left(\begin{array}{cc} A & \mathbf{b} \end{array}\right)$ ändern nichts an den Lösungen. Außerdem darf man die Spalten von A vertauschen. (Diese Vertauschungen müssen bei der Lösung mitberücksichtigt werden.)

Damit erhält man direkt ein Verfahren zum Lösen eines linearen Gleichungssystems, das **Gaußsche Eliminationsverfahren** zur Lösung eines linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ ($A \in K^{m\times n}, \mathbf{b} \in K^{m\times 1}$). Dabei startet man mit der erweiterten Matrix $\left(\begin{array}{cc} A & \mathbf{b} \end{array}\right)$ und führt folgenden Algorithmus durch, wobei man voraussetzt, daß A nicht die Nullmatrix ist.[2]

1. Durch etwaiges Zeilenvertauschen in $\left(\begin{array}{cc} A & \mathbf{b} \end{array}\right)$ bzw. Spaltenvertauschen in A erreicht man, daß $a_{11} \neq 0$ ist. Danach ersetzt man die j-te Zeile $\left(\begin{array}{cc} \tilde{\mathbf{a}}_j & b_j \end{array}\right)$ von $\left(\begin{array}{cc} A & \mathbf{b} \end{array}\right)$ $(2 \leq j \leq m)$ durch $\left(\begin{array}{cc} \tilde{\mathbf{a}}_j & b_j \end{array}\right) - a_{11}^{-1}a_{j1}\left(\begin{array}{cc} \tilde{\mathbf{a}}_1 & b_j \end{array}\right)$ und erhält eine Matrix der Form[3]

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix},$$

   d.h. in der ersten Spalte ist nur $a_{11} \neq 0$.

2. Daraufhin betrachtet man die Untermatrix

$$\begin{pmatrix} a_{22} & \cdots & a_{2n} & b_2 \\ a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & & \vdots & \vdots \\ a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} = \left(\begin{array}{cc} A' & \mathbf{b}' \end{array}\right)$$

   und wendet darauf dasselbe Verfahren an wie in 1) auf $\left(\begin{array}{cc} A & \mathbf{b} \end{array}\right)$.[4] Insgesamt erhält man dabei eine Matrix der Gestalt

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} & b_2 \\ 0 & 0 & a_{33} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & a_{m3} & \cdots & a_{mn} & b_m \end{pmatrix}$$

   mit $a_{11} \neq 0$ und $a_{22} \neq 0$.

3. Das soeben beschriebene Verfahren wird solange wie möglich iterativ fortgesetzt. Man

[2] In diesem Fall ist bei $\mathbf{b} \neq \mathbf{0}$ die Lösungsmenge leer und bei $\mathbf{b} = \mathbf{0}$ die Lösungsmenge ganz $K^{n\times 1}$.

[3] Zur Vereinfachung der Notation werden die Elemente der transformierten Matrix mit denselben Buchstaben bezeichnet.

[4] Dies funktioniert natürlich nur dann, wenn es ein Element $a_{ij} \neq 0$, $2 \leq i \leq m$, $2 \leq j \leq n$, gibt. Ist es dabei nötig, zwei Spalten von A' zu vertauschen, so müssen die entsprechenden Elemente der ersten Zeile von A auch vertauscht werden.

gewinnt schließlich eine Matrix der Form

$$\left(\begin{array}{ccccccccc} a_{11} & a_{12} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} & \cdots & a_{2r} & a_{2,r+1} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & a_{rr} & a_{r,r+1} & \cdots & a_{rn} & b_r \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & b_{r+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ & \cdots & 0 & 0 & 0 & \cdots & 0 & b_n \end{array}\right) = \left(\begin{array}{cc} \overline{A} & \overline{\mathbf{b}} \end{array}\right)$$

mit $a_{11} \neq 0, a_{22} \neq 0, \ldots, a_{rr} \neq 0$. r ist dabei der Rang der Matrix A. Diese Transformation wurde durch sukzessive elementare Zeilenumformungen der ursprünglichen erweiterten Matrix $\left(\begin{array}{cc} A & \mathbf{b} \end{array}\right)$ (und gegebenenfalls duch Spaltenvertauschungen von A) gewonnen. Es gibt daher eine reguläre Matrix $U \in K^{m\times m}$ (und eine Spaltentransformationsmatrix $T \in K^{n\times n}$, die nur Spalten vertauscht), so daß

$$\overline{A} = UAT \quad \text{und} \quad \overline{\mathbf{b}} = U\mathbf{b}.$$

4. Gibt es ein Element $b_j \neq 0$, $r < j \leq m$, so ist $\operatorname{rg}\left(\begin{array}{cc} \overline{A} & \overline{\mathbf{b}} \end{array}\right) > \operatorname{rg}(\overline{A})$ und somit das ursprünglich lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ **unlösbar**.

5. Sind alle Elemente $b_j = 0$, $r < j \leq m$, so ist $\operatorname{rg}\left(\begin{array}{cc} \overline{A} & \overline{\mathbf{b}} \end{array}\right) = \operatorname{rg}(\overline{A})$ und damit das ursprünglich lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ **lösbar**, und hat (bis auf die etwaige Komponentenvertauschungen, die in T kodiert sind), dieselben Lösungen wie $\overline{A}\mathbf{x} = \overline{b}$. Das lineare Gleichungssystem $\overline{A}\mathbf{x} = \overline{b}$ kann aber mit den Methoden aus Satz 4.75 vollständig gelöst werden.

Man beachte, daß die Transformationen

$$\overline{A} = UAT \quad \text{und} \quad \overline{\mathbf{b}} = U\mathbf{b}.$$

mit regulären Matrizen als Basiswechsel interpretiert werden können. Das ursprüngliche lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ ist ja die Koordinatendarstellung von $f(\mathbf{x}) = \mathbf{b}$ einer linearen Abbildung $f \in L(K^{n\times 1}, K^{m\times 1})$, wobei jeweils die kanonischen Basen zugrundegelegt werden. Die Matrix T entspricht nun einem Basiswechsel in $K^{n\times 1}$, wobei aber nur die Reihenfolge der Basisvektoren vertauscht wird. Die Matrix U vermittelt einen Basiswechsel in $K^{m\times 1}$, der bewirkt, daß die Koordinatendarstellung von f bezüglich dieser neuen Basis in $K^{m\times 1}$ einer Matrix entspricht, aus der die Lösung von $f(\mathbf{x}) = \mathbf{b}$, d.h. das Finden des Urbilds $f^{-1}(\{\mathbf{b}\})$, einfacher abgelesen werden kann.

Beispiel 4.78 Es soll das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & - & 2x_3 & + & 3x_4 & = & 3 \\ 2x_1 & + & 5x_2 & & & + & x_4 & = & 4 \\ 3x_1 & + & 8x_2 & + & 2x_3 & - & x_4 & = & 5 \\ x_1 & + & 4x_2 & + & 6x_3 & - & 7x_4 & = & -1 \end{array}$$

über einem Körper K vollständig gelöst werden. Die Koeffizientenmatrix A und die rechte Seite $\mathbf{b}$ sind

$$A = \begin{pmatrix} 1 & 2 & -2 & 3 \\ 2 & 5 & 0 & 1 \\ 3 & 8 & 2 & -1 \\ 1 & 4 & 6 & -7 \end{pmatrix} \quad \text{und} \quad \mathbf{b} = \begin{pmatrix} 3 \\ 4 \\ 5 \\ -1 \end{pmatrix}.$$

Durch elementare Zeilenumformungen der erweiterten Matrix $\begin{pmatrix} A & \mathbf{b} \end{pmatrix}$ erhält man

$$\begin{pmatrix} 1 & 2 & -2 & 3 & 3 \\ 2 & 5 & 0 & 1 & 4 \\ 3 & 8 & 2 & -1 & 5 \\ 1 & 4 & 6 & -7 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & -2 & 3 & 3 \\ 0 & 1 & 4 & -5 & -2 \\ 0 & 2 & 8 & -10 & -4 \\ 0 & 2 & 8 & -10 & -4 \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & 2 & -2 & 3 & 3 \\ 0 & 1 & 4 & -5 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Das ursprüngliche lineare Gleichungssystem ist daher lösbar und äquivalent zu

$$\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & - & 2x_3 & + & 3x_4 & = & 3 \\ & & x_2 & + & 4x_3 & - & 5x_4 & = & -2 \end{array}$$

Setzt man $x_3 = t_1$ und $x_4 = t_2$, so errechnet man

$$\begin{aligned} x_2 &= -2 - 4x_3 + 5x_4 \\ &= -2 - 4t_1 + 5t_2, \\ x_1 &= 3 - 2x_2 + 2x_3 - 3x_4 \\ &= 7 + 10t_1 - 13t_2. \end{aligned}$$

Alle Lösungen sind daher durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ -2 \\ 0 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} 10 \\ -4 \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -13 \\ 5 \\ 0 \\ 1 \end{pmatrix}$$

mit $t_1, t_2 \in K$ gegeben.

Mit Hilfe einer weiteren Zeilenumformungen hätte man die erweiterte Matrix $\begin{pmatrix} A & \mathbf{b} \end{pmatrix}$ auch in die Form

$$\cdots \to \begin{pmatrix} 1 & 2 & -2 & 3 & 3 \\ 0 & 1 & 4 & -5 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -10 & 13 & 7 \\ 0 & 1 & 4 & -5 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

bringen können und hätte diese allgemeine Lösung auch direkt ablesen können.

Kapitel 5

Determinanten

5.1 Permutationen

5.1.1 Die symmetrische Gruppe

Definition 5.1 *Sei A eine nichtleere Menge. Eine bijektive Abbildung $\pi : A \to A$ auf einer Menge A heißt* **Permutation** *auf A.*

Die Menge aller Permutationen $\mathbf{S}(A)$ auf einer Menge A heißt **symmetrische Gruppe** *auf A.*[1]

Für $\pi, \sigma \in \mathbf{S}(A)$ wird $\pi\sigma = \sigma \circ \pi$ als das **Produkt** *der Permutationen π und σ bezeichnet, d.h. zuerst wird π ausgeführt und danach σ. Weiters bezeichnet π^{-1} die zu π* **inverse Permutation**.

Ist $A = \{a_1, a_2, \ldots, a_n\}$ endlich, so identifiziert man die Elemente $a_1, a_2, \ldots, a_n$ mit den Zahlen $1, 2, \ldots, n$ und schreibt für $\mathbf{S}(A)$ auch $\mathbf{S}_n$.

Satz 5.2 *Es gibt genau $n!$ verschiedene Permutationen auf den Zahlen $\{1, 2, \ldots, n\}$, d.h. $|\mathbf{S}_n| = n!$.*

Permutationen auf endlichen Mengen besitzen verschiedene **Darstellungsarten**:

1. **Zweizeilige Darstellung:** $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ wird durch

$$\pi = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ \pi(1) & \pi(2) & \pi(3) & \cdots & \pi(n-1) & \pi(n) \end{pmatrix}$$

   dargestellt, z.B.

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}.$$

   Man kann daher eine Permutation auch als **Umordnung** interpretieren.

[1] Die Permuationen aus $\mathbf{S}(A)$ bilden mit der Hintereinaderausführung $\circ$ ein Gruppe.

Die inverse Permutation π^{-1} erhält man am einfachsten dadurch, daß man die beiden Zeilen (in der zweizeiligen Darstellung von π) vertauscht und dann spaltenweise ordnet:

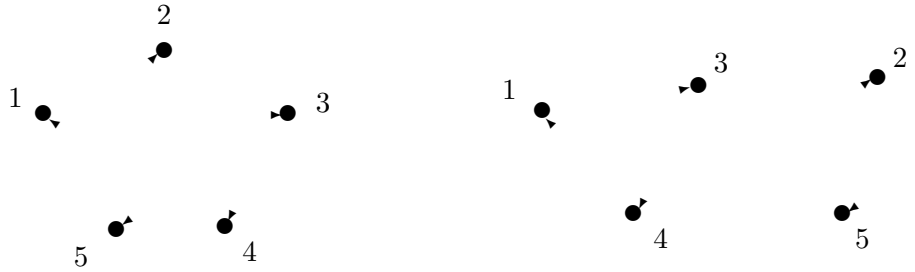
$$\pi^{-1} = \begin{pmatrix} 3 & 5 & 4 & 1 & 2 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{pmatrix}.$$

Ist $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix}$ eine weiter Permutation, so kann auch das Produkt $\pi\sigma$ in einfacher Weise bestimmt werden:

$$\pi\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 4 & 2 \end{pmatrix},$$

z.B. ist $\pi(1) = 3$ und $\sigma(3) = 5$, also $\pi\sigma(1) = 5$.

2. **Graphische Darstellung:** Die Zahlen $\{1, 2, \ldots, n\}$ werden als Punkte (Knoten) dargestellt, und ist $j = \pi(i)$, so verläuft ein Kante von i nach j. (Es gibt daher genau n Kanten.)



Man beachte, daß von jedem Punkt genau eine Kante wegführt und zu jedem Punkt genau eine Kante hinführt. Der Graph muß daher in **Zyklen** zerfallen. Im obigen Beispiel sind dies die Zyklen $1, \pi(1) = 3, \pi(3) = 4, \pi(3) = 1$ und $2, \pi(2) = 5, \pi(5) = 2$. Die Graphen von Permutationen haben daher ein sehr einfache Struktur. Sie bilden eine eine Menge von Zyklen, wobei natürlich auch **Schlingen** auftreten können, und zwar genau dann, wenn ein $j \in \{1, 2, \ldots, n\}$ auf sich selbst abgebildet wird, d.h. $\pi(j) = j$. Solche Punkte heißen auch **Fixpunkte**.

3. **Zyklendarstellung:** Da jede Permuation $\pi \in \mathbf{S}_n$ in eine Menge von Zyklen zerfällt genügt es, einfach diese anzugeben. Ist etwa $1, \pi(1), \pi(\pi(1)), \pi(\pi(\pi(1))) = \pi^3(1), \ldots, \pi^{k-1}(1), \pi^k(1) = 1$ der erste Zyklus, so stellt man diesen durch

$$(1\, \pi(1)\, \pi^2(1) \cdots \pi^{k-1}(1))$$

dar. Im obigen Beispiel gibt es also die Zyklen $(1\,3\,4)$ und $(2\,5)$. Schreibt man nun alle Zyklen von π hintereinander an, so erhält man die Zyklendarstellung von π, z.B.

$$\pi = (1\,3\,4)\,(2\,5).$$

Schreibt man die Zyklen in einer anderen Reihenfolge an, bzw. vertauscht man innerhalb eines Zyklus die Elemente zyklisch, so erhält man auch eine Zyklendarstellung dieser Permutation, z.B.

$$\pi = (5\,2)\,(3\,4\,1),$$

die Menge der Zyklen wird dadurch ja nicht verändert. Man beachte auch, daß damit $\pi = (5\,2) \cdot (3\,4\,1)$ auch als Produkt der Permutationen $(5\,2)$ und $(3\,4\,1)$ dargestellt wird.

Definition 5.3 *Eine Permutation $\pi \in \mathbf{S}_n$ heißt* **Transposition**, *wenn die Zyklendarstellung (in der Fixpunkte nicht aufgenommen werden) nur aus einem Zweierzyklus besteht. Ist $j \neq k$, so bezeichnet man die Transposition $(j\,k)$ durch π_{jk}.*

Bei einer Transposition π_{jk} werden also nur die Zahlen j und k vertauscht.

Satz 5.4 *Ein Zyklus $(b_1 b_2 \cdots b_k) \in \mathbf{S}_n$ der Länge k kann als Produkt von $k-1$ Transpositionen*

$$(b_1 b_2 \cdots b_k) = (b_1 b_2)\,(b_1 b_3) \cdots (b_1 b_k)$$

dargestellt werden.

Daher ist auch jede Permutation $\pi \in \mathbf{S}_n$ als Produkt von Transpositionen darstellbar.

5.1.2 Signum einer Permutation

Definition 5.5 *Das* **Signum** $\operatorname{sgn}(\pi)$ *einer Permuation $\pi \in \mathbf{S}_n$ ist durch*

$$\operatorname{sgn}(\pi) = \prod_{1 \leq i < j \leq n} \frac{\pi(j) - \pi(i)}{j - i}$$

gegeben.

Man beachte, daß $|\operatorname{sgn}(\pi)| = 1$ ist. Es gibt nur die Möglichkeiten $\operatorname{sgn}(\pi) = 1$ und $\operatorname{sgn}(\pi) = -1$.

Definition 5.6 *Ein Paar (i, j) $(1 \leq i, j \leq n)$ heißt* **Inversion** *$\pi \in \mathbf{S}_n$, wenn*

$$i < j \quad \textit{und} \quad \pi(i) > \pi(j).$$

Eine Permutation $\pi \in \mathbf{S}_n$ heißt **gerade**, *wenn die Anzahl der Inversionen von π gerade ist, und sie heißt* **ungerade**, *wenn die Anzahl der Inversionen von π ungerade ist.*

Aus der Definition des Signum ergibt sich daher der folgende einfache Sachverhalt.

Satz 5.7 *Eine Permuation $\pi \in \mathbf{S}_n$ ist genau dann gerade, wenn* $\operatorname{sgn}(\pi) = 1$.

Satz 5.8 *Für jede Transposition $(j\,k) \in \mathbf{S}_n$ $(j \neq k)$ gilt*

$$\operatorname{sgn}(j\,k) = -1.$$

Satz 5.9 *Die Abbildung* $\operatorname{sgn} : \mathbf{S}_n \to \{-1, 1\}$ *ist ein Gruppenhomomorphismus, d.h. für alle $\pi\sigma \in \mathbf{S}_n$ gilt*

$$\operatorname{sgn}(\pi\sigma) = \operatorname{sgn}(\pi)\operatorname{sgn}(\sigma).$$

Das Produkt zweier gerader Permutation ist gerade. Ebenso ist das Produkt zweier ungerader Permutationen gerade. Andererseits ist das Produkt einer geraden mit einer ungeraden Permutation ungerade.

Definition 5.10 *Der Menge der geraden Permutation in* $\mathbf{S}_n$ *wird durch* $\mathbf{A}_n$ *bezeichnet und wird auch* **alternierende Gruppe** *genannt.*

$\mathbf{A}_n$ ist der Kern von sgn. Für $n \geq 2$ ist sgn surjektiv.

Satz 5.11 $\mathbf{A}_n$ *ist ein Normalteiler von* $\mathbf{S}_n$*. Für* $n \geq 2$ *ist der Index 2, d.h.* $|\mathbf{A}_n| = n!/2$ *bzw. es gibt in* $\mathbf{S}_n$ *genauso viele gerade wie ungerade Permutationen.*

Satz 5.12 *Eine Permutation* $\pi \in \mathbf{S}_n$ *ist genau dann gerade, wenn man* π *als Produkt einer geraden Anzahl von Transposition darstellen kann. Entsprechend ist eine Permutation* $\pi \in \mathbf{S}_n$ *ungerade, wenn sie als Produkt einer ungeraden Anzahl von Transposition darstellbar ist.*

5.2 Determinatenformen

5.2.1 Definition

Definition 5.13 *Sei* $\mathbf{V}$ *ein n-dimensionaler Vektorraum über einem Körper* K*. Dann heißt eine Abbildung*

$$\Delta : \mathbf{V}^n \to K$$

Determinantenform, *wenn sie folgende drei Eigenschaften besitzt.*

1. Δ *ist multilinear, d.h. für alle* j, $1 \leq j \leq n$, *gilt*

$$\begin{aligned}&\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, x\mathbf{a} + y\mathbf{b}, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n)\\ &\quad = x\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{a}, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n) + y\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{b}, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n).\end{aligned}$$

2. *Sind von den Vektoren* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbf{V}$ *zwei gleich, so gilt*

$$\Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = 0.$$

3. *Es gibt eine Basis* $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ *von* $\mathbf{V}$ *mit*

$$\Delta(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n) \neq 0.$$

.

Lemma 5.14 *Sei* Δ *eine Determinantenform auf einem n-dimensionalen Vektroraum* $\mathbf{V}$*. Dann gilt für beliebige Vektoren* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbf{V}$ *und für jede Permutation* $\pi \in \mathbf{S}_n$

$$\Delta(\mathbf{a}_{\pi(1)}, \mathbf{a}_{\pi(2)}, \ldots, \mathbf{a}_{\pi(n)}) = \operatorname{sgn}(\pi)\Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n).$$

Wegen dieser Eigenschaft und der Mulitlinearität ist eine Determinantenform festgelegt, wenn man ihren Wert für eine Basis kennt.

Satz 5.15 *Sei Δ eine Determinantenform auf einem n-dimensionalen Vektroraum $\mathbf{V}$. Weiters seien $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbf{V}$ und $C = (c_{ij}) \in K^{n\times n}$. Dann gilt für die Vektoren*

$$\mathbf{a}'_j := \sum_{i=1}^{n} c_{ij}\mathbf{a}_i \quad (1 \le j \le n)$$

$$\Delta(\mathbf{a}'_1, \ldots, \mathbf{a}'_n) = \left(\sum_{\pi\in\mathbf{S}_n} \operatorname{sgn}(\pi) c_{\pi(1)1} c_{\pi(2)2} \cdots c_{\pi(n)n} \right) \Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n).$$

Satz 5.16 *Sei Δ eine Determinantenform auf einem n-dimensionalen Vektroraum $\mathbf{V}$. Dann gilt für beliebige Vektoren $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbf{V}$*

$$\Delta(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \neq 0 \Longleftrightarrow \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \text{ linear unabhängig.}$$

5.2.2 Existenz von Determinantenformen

Mit Hilfe der vorangestellten Sätze können nun alle Determinantenformen beschrieben werden, womit auch die Existenz von Determinantenformen gesichert ist.

Satz 5.17 *Sei $\mathbf{V}$ ein n-dimensionaler Vektorraum und $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$. Sind weiters*

$$\mathbf{a}_j = \sum_{i=1}^{n} x_{ij}\mathbf{b}_i \quad (1 \le j \le n)$$

n Vektoren in $\mathbf{V}$ (mit den Koordinaten x_{ij}) und $c \in K \setminus \{0\}$, so ist durch

$$\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_n) := c \sum_{\pi\in\mathbf{S}_n} \operatorname{sgn}(\pi) x_{\pi(1)1} x_{\pi(2)2} \cdots x_{\pi(n)n}$$

eine Determinantenform $\Delta : \mathbf{V}^n \to K$ gegeben, und jede Determinantenform wird auf diese Weise erzeugt.

5.3 Determinanten

5.3.1 Determinanten und Determinantenformen

Definition 5.18 *Die* **Determinante** *einer Matrix $A = (a_{ij})_{1\le i,j\le n} \in K^{n\times n}$ ist durch*

$$\det A := \sum_{\pi\in\mathbf{S}_n} \operatorname{sgn}(\pi) a_{\pi(1)1} a_{\pi(2)2} \cdots a_{\pi(n)n}$$

gegeben.

In der Darstellung als quadratisches Schema schreibt man anstelle von $\det(a_{ij})$ auch

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Beispiel 5.19 Für $n = 2$ erhält man daher

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

und für $n = 3$ (**Regel von Sarrus**)

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ -a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \end{array}$$

Determinantenformen in $K^{n\times 1}$ stehen in engem Zusammenhang mit der Determinante.

Satz 5.20 *Sei $A \in K^{n\times n}$. Dann gilt für jede Determinantenform $\Delta : (K^{n\times 1})^n \to K$*

$$\det A = \frac{\Delta(\mathbf{a}_1, \ldots, \mathbf{a}_n)}{\Delta(\mathbf{e}_1, \ldots, \mathbf{e}_n)},$$

wobei $\mathbf{a}_1, \ldots, \mathbf{a}_n \in K^{n\times 1}$ die Spalten von A bezeichnen.

Man beachte, daß die Spalten $\mathbf{a}_j$ $(1 \le j \le n)$ der Matrix A genau die Bilder $f_A(\mathbf{e}_j) = A\mathbf{e}_j$ der kanonischen Basisvektoren $\mathbf{e}_j$ $(1 \le j \le n)$ sind. Es gilt also

$$\det A = \frac{\Delta(f_A(\mathbf{e}_1), \ldots, f_A(\mathbf{e}_n))}{\Delta(\mathbf{e}_1, \ldots, \mathbf{e}_n)}.$$

Aus Satz 5.15 folgt sofort, daß man die kanonische Basis $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ durch eine beliebige andere Basis $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ ersetzen kann, und es gilt noch immer

$$\det A = \frac{\Delta(f_A(\mathbf{b}_1), \ldots, f_A(\mathbf{b}_n))}{\Delta(\mathbf{b}_1, \ldots, \mathbf{b}_n)}.$$

Man definiert daher für eine lineare Abbildung $f \in L(\mathbf{V}, \mathbf{V})$ (wobei $\mathbf{V}$ ein n-dimensionaler Vektorraum ist) die **Determinante** von f durch

$$\det f := \frac{\Delta(f(\mathbf{b}_1), \ldots, f(\mathbf{b}_n))}{\Delta(\mathbf{b}_1, \ldots, \mathbf{b}_n)},$$

wobei $\Delta : \mathbf{V}^n \to K$ eine beliebige Determinatenform und $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eine beliebige Basis von $\mathbf{V}$ ist. Der Wert $\det f$ ist von der Wahl von Δ und $\mathbf{B}$ unabhängig.

Die Determinante hat für lineare Abbildungen $f \in L(\mathbb{R}^n, \mathbb{R}^n)$ eine interessante Interpretation, es gilt nämlich

$$|\det f| = \mathrm{Vol}_n(f([0,1)^n),$$

wobei Vol_n das n-dimensionale Volumen und $[0,1)^k$ den Einheitswürfel bezeichnet.[2] Für Matrizen $A \in \mathbb{R}^{n\times n}$ bedeutet dies, dass das das Volumen des von den Spalten von A aufgespannten *Parallelepipeds*

$$P = \{t_1\mathbf{a}_1 + \cdots + t_n\mathbf{a}_n \,|\, 0 \le t_1, \ldots, t_n < 1\}$$

das Volumen

$$\mathrm{Vol}_n(P) = |\det A|$$

hat.

[2] Für jede (meßbare) Menge $\mathbf{M} \subseteq \mathbb{R}^n$ gilt daher für das Bild $f(\mathbf{M}) = \{f(\mathbf{x}) \,|\, \mathbf{x} \in \mathbf{M}\}$ $\mathrm{Vol}_n(f(\mathbf{M}) = |\det f| \,\mathrm{Vol}_n(\mathbf{M})$.

5.3.2 Eigenschaften der Determinante

Satz 5.21 *Eine Matrix $A \in K^{n\times n}$ ist genau dann regulär, wenn*

$$\det A \neq 0.$$

Satz 5.22

1. *Für beliebige Matrizen $A, B \in K^{n\times n}$ gilt*

$$\det(AB) = \det A \cdot \det B.$$

2. *Ist $A \in K^{n\times n}$ invertierbar, so berechnet sich die Determinante der inversen Matrix durch*

$$\det(A^{-1}) = (\det A)^{-1}.$$

3. *Für $A \in K^{n\times n}$ gilt*

$$\det(A^T) = \det A.$$

Satz 5.23

1. *Multipliziert man eine Spalte/Zeile einer Matrix A mit einem Faktor $x \in K$, so ist die Determinante der neuen Matrix* $\det A' = x \det A$.
2. *Addiert man zu einer Spalte/Zeile einer Matrix das Vielfache einer anderen Spalte/Zeile, so verändert sich der Wert der Determinante nicht.*
3. *Vertauscht man in einer Matrix A zwei Spalten/Zeilen, so ist die Determinante der neuen Matrix* $\det A' = -\det A$.

Lemma 5.24 *Ist $A = (a_{ij})_{1\leq i,j\leq n} \in K^{n\times n}$ eine obere Dreicksmatrix*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix},$$

d.h. $a_{ij} = 0$ für $1 \leq j < i \leq n$. Dann gilt

$$\det A = a_{11}a_{22}\cdots a_{nn}.$$

Dieselbe Eigenschaft gilt auch für untere Dreiecksmatrizen.

Mit elementaren Spalten- und Zeilenumformungen kann jede Matrix in eine obere Dreicksmatrix umgeformt werden. Wegen Satz 5.23 verändert sich dabei der Wert der Determinante in kontrollierter Art und Weise. Damit kann jede Determinante (auch ohne explizite Kenntnis von $\mathbf{S}_n$) berechnet werden.

Beispiel 5.25 Die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

berechnet sich nach geeigneten elementaren Zeilenumformungen zu

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

5.3.3 Laplacescher Entwicklungssatz

Definition 5.26 *Sei $A \in K^{n\times n}$. Unter dem* **Kofaktor** *A_{lk} $(1 \le l, k \le n)$ versteht man die Determinante jener Matrix, die dadurch hervorgeht, daß man die k-te Spalte durch den Vektor $\mathbf{e}_l$ der kanonischen Basis ersetzt.*

Addiert man entsprechende Vielfache von $\mathbf{e}_l$ zu den Spalten $\mathbf{a}_1, \ldots, \mathbf{a}_{k-1}, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_n$ von A, so kann man erreichen, daß in der l-ten Zeile dieser Spalten nur mehr 0 steht. Es gilt daher auch

$$A_{lk} = \begin{vmatrix} a_{11} & \ldots & a_{1,k-1} & 0 & a_{1,k+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{l-1,1} & \ldots & a_{l-1,k-1} & 0 & a_{l-1,k+1} & \ldots & a_{l-1,n} \\ 0 & \ldots & 0 & 1 & 0\ldots & 0 & \\ a_{l+1,1} & \ldots & a_{l+1,k-1} & 0 & a_{l+1,k+1} & \ldots & a_{l+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,k-1} & 0 & a_{n,k+1} & \ldots & a_{nn} \end{vmatrix}.$$

Lemma 5.27 *Sei D_{lk} $(1 \le l, k \le n)$ die Determinante jener Matrix aus $K^{(n-1)\times(n-1)}$, die aus $A \in K^{n\times n}$ dadurch hervorgeht, daß die l-te Zeile und die k-te Spalte gestrichen werden. Dann gilt*

$$A_{lk} = (-1)^{l+k} D_{lk}.$$

Satz 5.28 (Laplacescher Entwicklungssatz) *Sei $A = (a_{ij})_{1\le i,j\le n} \in K^{n\times n}$. Dann gilt für jedes l $(1 \le l \le n)$*

$$\det A = \sum_{j=1}^{n} a_{lj} A_{lj} = \sum_{j=1}^{n} (-1)^{l+j} a_{lj} D_{lj}$$

und für jedes k $(1 \le k \le n)$

$$\det A = \sum_{i=1}^{n} a_{ik} A_{ik} = \sum_{i=1}^{n} (-1)^{i+k} a_{ij} D_{ij}.$$

Mit Hilfe dieses Satzes kann das Berechnen der Determinate einer $n\times n$-Matrix auf das Berechnen von n Determinanten von $(n-1)\times(n-1)$-Matrizen zurückgeführt werden.

Beispiel 5.29 Entwicklet man die Determinante

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

nach der 1. Zeile, so ergibt sich

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} + \\ &= (5 \cdot 9 - 6 \cdot 8) - 2(4 \cdot 9 - 7 \cdot 6) + 3(4 \cdot 8 - 5 \cdot 7) = 0. \end{aligned}$$

Satz 5.30 *Sei $A \in K^{n \times n}$ und bezeichne $\hat{A} = (A_{ij})_{1 \leq i,j \leq n} \in K^{n \times n}$ die Matrix der Kofaktoren von A. Dann gilt*

$$A\hat{A}^T = (\det A) E_n.$$

Insbesondere gilt für reguläre Matrizen $A \in K^{n \times n}$

$$A^{-1} = \frac{1}{\det A} \hat{A}^T = \left(\frac{A_{ji}}{\det A} \right)_{1 \leq i,j \leq n}.$$

5.3.4 Cramersche Regel

Sei $A \in K^{n \times n}$ eine reguläre Matrix und $\mathbf{b} \in K^{n \times 1}$ Dann kann das lineare Gleichungssystem

$$A\mathbf{x} = \mathbf{b}$$

immer eindeutig gelöst werden. Die einzige Lösung ist durch

$$\mathbf{x} = A^{-1}\mathbf{b}$$

gegeben. Da A^{-1} durch Determinantenberechnunen explizit berechnet werden kann, ist damit die Lösung $\mathbf{x} = A^{-1}\mathbf{b}$ auch explizit anzugeben. Dieses Verfahren kann aber noch abgekürzt werden.

Satz 5.31 (Cramersche Regel) *Sei $A \in K^{n \times n}$ eine reguläre Matrix und $\mathbf{b} \in K^{n \times 1}$. Bezeichnet man mit A_j $(1 \leq j \leq n)$ jene Matrix, die aus A dadurch hervorgeht, daß man die j-te Spalte durch $\mathbf{b}$ ersetzt, so ist die einzige Lösung des linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ durch*

$$\mathbf{x} = \frac{1}{\det A} \begin{pmatrix} \det A_1 \\ \vdots \\ \det A_n \end{pmatrix}$$

gegeben.

Man kann also durch $x_i = \dfrac{\det A_j}{\det A}$ jede Koordinate der Lösung einzeln berechnen.

5.3.5 $GL(n, K)$ und $SL(n, K)$

Definition 5.32 *Die Menge der regulären $n \times n$-Matrizen A über einem Körper K wird mit $GL(n, K)$ bezeichnet.*

Die Menge der regulären $n \times n$-Matrizen A über einem Körper K mit $\det A = 1$ wird mit $SL(n, K)$ bezeichnet.

Satz 5.33 *$GL(n, K)$ bildet mit der Matrizenmultiplikation eine Gruppe und $SL(n, K)$ einen Normalteiler von $GL(n, K)$.*

Definition 5.34 *Ein Körper K heißt* **angeordnet**, *wenn in K ein* **Positivbereich** *K^+ ausgezeichnet werden kann, so daß $K^+, \{0\}, -K^+$ eine Partition von K ist und mit $x, y \in K^+$ auch $x + y \in K^+$ und $xy \in K^+$ liegen.*

Die Elemente in K^+ heißen positiv, und die Elemente in $-K^+$ negativ.

Beispiel 5.35 $\mathbb{Q}$ und $\mathbb{R}$ sind angeordnet. $\mathbb{C}$ kann nicht angeordnet werden.

Definition 5.36 *Sei K ein angeordneter Körper. Die Menge der regulären $n \times n$-Matrizen A über K mit $\det A \in K^+$ wird mit $GL^+(n, K)$ bezeichnet. Entsprechend wird die Menge der regulären $n \times n$-Matrizen A über K mit $\det A \in -K^+$ wird mit $GL^-(n, K)$ bezeichnet.*

Satz 5.37 *Sei K ein angeordneter Körper. Dann ist $GL^+(n, K)$ ein Normalteiler von $GL(n, K)$ und $SL(n, K)$ ein Normalteiler von $GL^+(n, K)$.*

Definition 5.38 *Sei $\mathbf{V}$ ein endlichdimensionaler Vektorraum über einem angeordneten Körper K. Zwei Bases $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ und $\mathbf{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ heißen* **gleichorientiert**, *wenn $\mathbf{T_{BC}} \in GL^+(n, K)$.*

Kapitel 6

Duale Vektorräume

6.1 Linearformen

6.1.1 Definition und Beispiele

Definition 6.1 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K*. Eine lineare Abbildung* $f \in L(\mathbf{V}, K)$ *heißt* **Linearform** *oder* **lineares Funktional**.

Der Vektorraum aller Linearformen $\mathbf{V}^* := L(\mathbf{V}, K)$ *heißt* **dualer Vektorraum** *oder* **Dualraum**.

Beispiel 6.2 Sei $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ Basis eines Vektorraums $\mathbf{V}$ und $j \in I$. Dann ist die Abbildung

$$\mathbf{b}_j^* : \mathbf{V} \to K, \qquad \mathbf{a} = \sum_{i \in I} x_i \mathbf{b}_i \mapsto x_j$$

eine Linearform.

Beispiel 6.3 Sei $\mathbf{V} = K^M$ die Menge der Funktionen $f : M \to K$ mit den Operationen $(f+g)(t) := f(t) + g(t)$ und $(xf)(t) := xf(t)$. Dann ist für alle $t_0 \in M$ die *Auswertung*

$$\nu_{t_0} : K^M \to K, \qquad f \mapsto f(t_0)$$

eine Linearform.

Beispiel 6.4 Sei $\mathbf{V} = C[0,1]$ die Menge der stetigen Funktionen $f : [0,1] \to \mathbb{R}$. Dann ist das Integral

$$I : \mathbf{V} \to \mathbb{R}, \qquad f \mapsto I(f) := \int_0^1 f(t)\, dt$$

eine Linearform.

Linearformen werden im folgenden als Vektoren des dualen Vektorraums gesehen. Um sie von den Vektoren $\mathbf{a} \in \mathbf{V}$ optisch zu unterscheiden werden sie immer mit einem $*$ versehen, d.h. $\mathbf{a}^*$ wird immer eine Linearform aus $\mathbf{V}^*$ bezeichnen.

Definition 6.5 *Sei* $\mathbf{V}$ *ein Vektorraum über einem Körper* K *und* $\mathbf{V}^*$ *der duale Vektorraum. Dann beizeichnet man die Abbildung*

$$\mathbf{V}^* \times \mathbf{V} \to K, \quad (\mathbf{a}^*, \mathbf{a}) \mapsto \langle \mathbf{a}^*, \mathbf{a} \rangle := \mathbf{a}^*(\mathbf{a})$$

als **kanonische Paarung**.

Satz 6.6 *Die kanonische Paarung eines Vektorraums* $\mathbf{V}$ *ist bilinear, d.h. für* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$, $\mathbf{a}^*, \mathbf{b}^* \in \mathbf{V}^*$ *und* $x, y \in K$ *gelten*

$$\langle x\mathbf{a}^* + y\mathbf{b}^*, \mathbf{a} \rangle = x\langle \mathbf{a}^*, \mathbf{a} \rangle + y\langle \mathbf{b}^*, \mathbf{a} \rangle$$

und

$$\langle \mathbf{a}^*, x\mathbf{a} + y\mathbf{b} \rangle = x\langle \mathbf{a}^*, \mathbf{a} \rangle + y\langle \mathbf{a}^*, \mathbf{b} \rangle.$$

6.1.2 Duale Basis

Im Beispiel 6.3 wurde zu einer Basis $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ und jedem $j \in I$ eine Linearform $\mathbf{b}_j^* \in \mathbf{V}^*$ zugeordnet. $\mathbf{b}_j^*$ *extrahiert* die j-te Koordinate eines Vektors $\mathbf{a} \in \mathbf{V}$ bezüglich der Basis $\mathbf{B}$. Wegen des Fortsetzungssatzes kann $\mathbf{b}_j^*$ auch durch

$$\langle \mathbf{b}_j^*, \mathbf{b}_i \rangle := \delta_{ij}$$

definiert werden.

Satz 6.7 *Sei* $\mathbf{B} = \{\mathbf{b}_i \mid i \in I\}$ *Basis eines Vektorraums* $\mathbf{V}$. *Dann sind die Linearformen* $\mathbf{b}_j^*$ $(j \in I)$ *alle voneinander verschieden und linear unabhängig in* $\mathbf{V}^*$.

Die Menge $\mathbf{B}^* := \{\mathbf{b}_j^* \mid j \in I\}$ *ist genau dann eine Basis von* $\mathbf{V}^*$, *wenn* $\mathbf{V}$ *endlichdimensional ist.*

Im endlichdimensionalen Fall gilt $\dim \mathbf{V}^* = \dim L(\mathbf{V}, K) = \dim \mathbf{V}$. Es ist also klar, daß $\mathbf{V}^*$ dieselbe Dimension hat wie $\mathbf{V}$.[1] Da alle $\mathbf{b}_j^*$ voneinander verschieden sind und ein linear unabhängiges System bilden, sind sie tatsächlich eine Basis von $\mathbf{V}^*$.

Im unendlichdimensionalen Fall kann man sogar zeigen, daß jede Basis von $\mathbf{V}^*$ echt mächtiger ist als eine von $\mathbf{V}$.

Definition 6.8 *Sei* $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ *Basis eines (n-dimensionalen) Vektorraums* $\mathbf{V}$. *Dann heißt die Menge*

$$\mathbf{B}^* := \{\mathbf{b}_j^* \mid 1 \leq j \leq n\}$$

duale Basis *von* $\mathbf{V}$.

Satz 6.9 *Sei* $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ *Basis eines (n-dimensionalen) Vektorraums* $\mathbf{V}$ *und* $\mathbf{a}^* \in \mathbf{V}^*$. *Dann ist die Koordinatenmatrix von* $\mathbf{a}^*$ *bezüglich* $\mathbf{B} \subseteq \mathbf{V}$ *und* $\{1\} \subseteq K$ *durch*

$$\Phi_{\mathbf{B}\{1\}}(\mathbf{a}^*) = \Phi_{\mathbf{B}^*}(\mathbf{a}^*)^T = \begin{pmatrix} \langle \mathbf{a}^*, \mathbf{b}_1 \rangle & \ldots & \langle \mathbf{a}^*, \mathbf{b}_n \rangle \end{pmatrix} \in K^{1 \times n}$$

gegeben.

[1] $\mathbf{V}$ und $\mathbf{V}^*$ sind in diesem Fall auch isomorph.

Satz 6.10 *Sei $f \in L(\mathbf{V}, \mathbf{W})$ eine lineare Abbildung zwischen zwei endlichdimensionalen Vektorräumen $\mathbf{V}, \mathbf{W}$. Ist $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ eine Basis von $\mathbf{V}$ und $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_m\}$ eine Basis von $\mathbf{W}$ so ist die Koordinatenmatrix von f bezüglich der Basen $\mathbf{B}$ und $\mathbf{C}$ durch*

$$\Phi_{\mathbf{BC}}(f) = (\langle \mathbf{c}_i^*, f(\mathbf{b}_j) \rangle)_{1 \leq i \leq m, 1 \leq j \leq n}$$

gegeben.

6.1.3 Bidualraum

Definition 6.11 *Der Dualraum $\mathbf{V}^{**}$ des Dualraums $\mathbf{V}^*$ eines Vektorraums $\mathbf{V}$ heißt* **Bidualraum** *von $\mathbf{V}$.*

Die Abbildung

$$e : \mathbf{V} \to \mathbf{V}^{**}, \quad \mathbf{a} \mapsto e(\mathbf{a}), \quad e(\mathbf{a})(\mathbf{a}^*) := \mathbf{a}^*(\mathbf{a})$$

heißt **kanaonische Identifizierung von $\mathbf{V}$ und $\mathbf{V}^{**}$.**

Lemma 6.12 *Die kanonische Identifizierung $e : \mathbf{V} \to \mathbf{V}^{**}$ ist eine injektive lineare Abbildung. Für einen endlichdimensionale Vektorraum $\mathbf{V}$ ist e ein Isomorphismus.*

6.2 Annullatorräume

6.2.1 Basis eines Annullatorraums

Definition 6.13 *Sei $\mathbf{V}$ ein Vektorraum über einem Körper K und $\mathbf{V}^*$ der Dualraum. Für eine Teilmenge $\mathbf{M} \subseteq \mathbf{V}$ bezeichnet*

$$\mathbf{M}^\circ := \{\mathbf{a}^* \in \mathbf{V}^* \,|\, \forall \mathbf{m} \in \mathbf{M} : \ \langle \mathbf{a}^*, \mathbf{m} \rangle = 0\} = \bigcap_{\mathbf{m} \in \mathbf{M}} \text{kern } e(\mathbf{m})$$

den **Annullatorraum** *von $\mathbf{M}$.*

Lemma 6.14 *Der Annullatorraum $\mathbf{M}^\circ$ einer Teilmenge $\mathbf{M} \subseteq \mathbf{V}$ eines Vektorraums $\mathbf{V}$ ist ein Teilraum von $\mathbf{V}^*$.*

Lemma 6.15 *Für Teilmengen $\mathbf{M}, \mathbf{N}$ eines Vektorraums $\mathbf{V}$ gelten folgende Eigenschaften:*

1. $\mathbf{M} \subseteq \mathbf{N} \Longrightarrow \mathbf{N}^\circ \subseteq \mathbf{M}^\circ$.
2. $[\mathbf{M}]^\circ = \mathbf{M}^\circ$.
3. $(\mathbf{M} \cup \mathbf{N})^\circ = \mathbf{M}^\circ \cap \mathbf{N}^\circ$.

Satz 6.16 *Sei $\mathbf{U}$ Unterraum eines Vektorraums $\mathbf{V}$ mit endlicher Kodimension m. Dann ist*

$$\dim \mathbf{U}^\circ = m.$$

Ist $\mathbf{B} = \{\mathbf{b}_i \,|\, i \in I\}$ eine Basis von $\mathbf{V}$, so daß es m verschiedene $i_1, i_2, \ldots, i_m \in I$ derart gibt, daß $\{\mathbf{b}_{i_1}, \mathbf{b}_{i_2}, \ldots, \mathbf{b}_{i_m}\}$ Basis eines Komplementärraums von $\mathbf{U}$ ist, dann bilden die Linearformen $\{\mathbf{b}_{i_1}^*, \mathbf{b}_{i_2}^*, \ldots, \mathbf{b}_{i_m}^*\}$ eine Basis von $\mathbf{U}^\circ$.

Satz 6.17 *Sei* $\mathbf{V}$ *ein endlichdimensionaler Vektorraum, der mittels der kanonischen Identifizierung* $e : \mathbf{V} \to \mathbf{V}^{**}$ *mit* $\mathbf{V}^{**}$ *identifiziert wird. Dann gilt für jede Teilmenge* $\mathbf{M} \subseteq \mathbf{V}$

$$(\mathbf{M}^\circ)^\circ = [\mathbf{M}].$$

Insbesondere gilt für Unterräume $\mathbf{U} \leq \mathbf{V}$

$$(\mathbf{U}^\circ)^\circ = \mathbf{U}.$$

Für Unterräume $\mathbf{U}$ in $\mathbf{V} = K^{n\times 1}$ gibt es ein einfaches Verfahren, bei Vorgabe einer Basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$ von $\mathbf{U}$ eine entsprechende Basis von $\mathbf{U}^\circ$ zu bestimmen. Dazu bezeichne $B \in K^{n\times r}$ jene Matrix, die sich aus den Spalten $\mathbf{b}_1, \ldots, \mathbf{b}_r$ zusammensetzt. Nach geeigneten elementaren Spaltenumformungen erhält man eine Matrix, die (bis auf die Reihenfolge der Zeilen) von der Gestalt

$$\begin{pmatrix} E_r \\ A' \end{pmatrix}$$

(mit $A' \in K^{(n-r)\times r}$) ist, deren Spalten natürlich wieder eine Basis von $\mathbf{U}$ bilden. Wegen

$$\begin{pmatrix} -A' & E_{n-r} \end{pmatrix} \begin{pmatrix} E_r \\ A' \end{pmatrix} = -A' + A' = \mathbf{0}^{(n-r)\times r}$$

und $\operatorname{rg}\begin{pmatrix} -A' & E_{n-r} \end{pmatrix} = n - r$ bilden nun die Zeilen von $\begin{pmatrix} -A' & E_{n-r} \end{pmatrix}$ eine Basis von $\mathbf{U}^\circ$.

6.2.2 Summe und Durchschnitt von Annullatorräumen

Satz 6.18 *Sei* $(\mathbf{U}_i \,|\, i \in I)$ *ein System von Unterräumen eines Vektorraums* $\mathbf{V}$*. Dann gilt*

$$\left(\sum_{i\in I} \mathbf{U}_i\right)^\circ = \bigcap_{i\in I} \mathbf{U}_i^\circ$$

und

$$\sum_{i\in I} \mathbf{U}_i^\circ \subseteq \left(\bigcap_{i\in I} \mathbf{U}_i\right)^\circ.$$

Ist I *endlich oder* $\mathbf{V}$ *endlichdimensional, so gilt auch*

$$\sum_{i\in I} \mathbf{U}_i^\circ = \left(\bigcap_{i\in I} \mathbf{U}_i\right)^\circ.$$

6.3 Adjungierte Abbildung

Definition 6.19 *Sei* $f \in L(\mathbf{V}, \mathbf{W})$ *eine lineare Abbildung. Dann ist durch*

$$f^* : \mathbf{W}^* \to \mathbf{V}^*, \qquad \mathbf{c}^* \mapsto \mathbf{c}^* \circ f$$

eine Abbildung definiert, die als **adjungierte Abbildung** *oder* **transponierte Abbildung**[2] *bezeichnet wird.*

[2]Es wird anstelle von f^* auch die Bezeichnung f^T verwendet.

$$\begin{array}{ccc} \mathbf{V} & \xrightarrow{f} & \mathbf{W} \\ {\scriptstyle \mathbf{c}^* \circ f} \downarrow & & \downarrow {\scriptstyle \mathbf{c}^*} \\ K & = & K \end{array}$$

Lemma 6.20 *Die adjungierte Abbildung f^* einer linearen Abbildung $f \in L(\mathbf{V}, \mathbf{W})$ ist linear.*

Satz 6.21 *Für lineare Abbildungen $f \in L(\mathbf{V}, \mathbf{W})$ und $g \in L(\mathbf{W}, \mathbf{X})$ gilt*

$$(g \circ f)^* = f^* \circ g^*.$$

Insbesondere gilt für bijektive lineare Abbildungen $f \in L(\mathbf{V}, \mathbf{W})$

$$(f^{-1})^* = (f^*)^{-1}.$$

Satz 6.22 *Für eine lineare Abbildung $f \in L(\mathbf{V}, \mathbf{W})$ gilt*

$$f(\mathbf{V})^\circ = \operatorname{kern}(f^*) \quad \textit{und} \quad \operatorname{kern}(f)^\circ = f^*(\mathbf{W}^*).$$

Korollar 6.23 *Sei $f \in L(\mathbf{V}, \mathbf{W})$, wobei $\mathbf{V}$ oder $\mathbf{W}$ endlichdimensional ist. Dann gilt*

$$\operatorname{rg}(f^*) = \operatorname{rg}(f).$$

Satz 6.24 *Sei $f \in L(\mathbf{V}, \mathbf{W})$ eine lineare Abbildung zwischen zwei endlichdimensionalen Vektorräumen. Weiters sei $\mathbf{B}$ eine Basis von $\mathbf{V}$ und $\mathbf{C}$ eine Basis von $\mathbf{W}$. Dann ist die Koordinatenmatrix der adjungierten Abbildung $f^* \in L(\mathbf{W}^*, \mathbf{V}^*)$ bezüglich der dualen Basen $\mathbf{C}^*$ und $\mathbf{B}^*$ die transponierte Matrix der Koordinatenmatrix von f bezüglich der Basen $\mathbf{B}$ und $\mathbf{C}$:*

$$\Phi_{\mathbf{C}^*\mathbf{B}^*}(f^*) = \Phi_{\mathbf{B}\mathbf{C}}(f)^T.$$

Man beachte, daß hiermit wieder bewiesen wurde, daß für jede Matrix $A \in K^{m \times n}$

$$\operatorname{rg}(A) = \operatorname{rg}(A^T)$$

gilt, d.h. Spalten- und Zeilenrang von A stimmen überein.

Satz 6.25 *Identifiziert man die Bidualräume $\mathbf{V}^{**}$, $\mathbf{W}^{**}$ zweier endlichdimensionaler Vektroräume $\mathbf{V}$, $\mathbf{W}$ mit $\mathbf{V}$ resp. $\mathbf{W}$, so gilt für jede lineare Abbildung $f \in L(\mathbf{V}, \mathbf{W})$*

$$(f^*)^* = f.$$

Kapitel 7

Lineare Geometrie

7.1 Affine Geometrie

7.1.1 Vorbemerkungen

Grundbausteine jeder Geometrie sind Punkte, Geraden (und eventuell Ebenen etc.) und die Enthaltenseinsrelation dieser Objekte. Je nachdem, welche zusätzlichen Relationen (z.B. die Parallelitätsrelation) zur Verfügung hat und welche Axiome (z.B. daß es zu zwei verschiedenen Punkten genau eine Gerade gibt, die diese beiden Punkte enthält) man fordert, erhält man unterschiedliche Geometrien.

Im folgenden wird ein *Modell* einer *affinen Geometrie* entwickelt, das sich auf die Theorie der Vektorräume stützt.

7.1.2 Nebenräume

Zunächst werden einige Begriffe und Eigenschaften wiederholt und erweitert:

Definition 7.1 *Sei* $\mathbf{V}$ *ein Vektorraum (über einem Köper* K*). Eine Teilmenge* $\mathbf{N} \subseteq \mathbf{V}$ *heißt* **Nebenraum** *in* $\mathbf{V}$*, wenn es einen Unterraum* $\mathbf{U} \leq \mathbf{V}$ *und einen Vektor* $\mathbf{a} \in \mathbf{V}$ *mit*

$$\mathbf{N} = \mathbf{a} + \mathbf{U} = \{\mathbf{a} + \mathbf{c} \,|\, \mathbf{c} \in \mathbf{U}\}$$

gibt.

Lemma 7.2 *Sei* $\mathbf{a} + \mathbf{U}$ *ein Nebenraum in* $\mathbf{V}$ *und* $\mathbf{b} \in \mathbf{V}$*. Dann sind die folgenden drei Bedingungen äquivalent:*

(i) $\mathbf{a} + \mathbf{U} = \mathbf{b} + \mathbf{U}$.

(ii) $\mathbf{b} \in \mathbf{a} + \mathbf{U}$.

(iii) $\mathbf{a} - \mathbf{b} \in \mathbf{U}$.

Lemma 7.3 *Sind* $\mathbf{a}+\mathbf{U}$ *und* $\mathbf{b}+\mathbf{W}$ *zwei Nebenräume in* $\mathbf{V}$*, dann folgt aus* $\mathbf{a}+\mathbf{U}=\mathbf{b}+\mathbf{W}$*, daß* $\mathbf{U}=\mathbf{W}$ *ist.*

Mit Hilfe dieser Eigenschaft kann in eindeutiger Weise die Dimension eines Nebenraums definiert werden.

Definition 7.4 *Ist* $\mathbf{N}=\mathbf{a}+\mathbf{U}$ *ein Neberaum in* $\mathbf{V}$*, so definiert man die* **Dimension** *von* $\mathbf{N}$ *durch*

$$\dim \mathbf{N} := \dim \mathbf{U}.$$

Lemma 7.5 *Es bezeichnen* $\mathbf{N}$ *und* $\mathbf{M}$ *zwei endlichdimensionale Nebenräume in* $\mathbf{V}$*. Dann gelten die folgenden beiden Eigenschaften:*

1. $\mathbf{N}\subseteq\mathbf{M}\Longrightarrow \dim\mathbf{N}\leq\dim\mathbf{M}$.
2. $\mathbf{N}\subseteq\mathbf{M}\wedge\dim\mathbf{N}=\dim\mathbf{M}\Longrightarrow\mathbf{N}=\mathbf{M}$.

Lemma 7.6 *Sei* $(\mathbf{N}_i \,|\, i\in I)$ *ein System von Nebenräumen in* $\mathbf{V}$*. Dann ist der Durchschnitt*

$$\bigcap_{i\in I}\mathbf{N}_i$$

entweder leer oder wieder ein Nebenraum in $\mathbf{V}$*.*

Man beachte, daß in Fall des nichtleeren Durchschnitts alle Nebenräume $\mathbf{N}_i$ einen gemeinsamen Repräsentaten $\mathbf{a}\in\mathbf{V}$ besitzen, d.h.

$$\forall i\in I:\ \mathbf{N}_i=\mathbf{a}+\mathbf{U}_i,$$

wobei $\mathbf{U}_i$ entsprechenden Unterräume bezeichnen.

Definition 7.7 *Sei* $(\mathbf{N}_i \,|\, i\in I)$ *ein System von Nebenräumen in* $\mathbf{V}$*. Der kleinste Nebenraum in* $\mathbf{V}$*, der alle* $\mathbf{N}_i$*,* $i\in I$*, enthält, heißt* **Verbindungsraum**

$$\bigvee_{i\in I}\mathbf{N}_i := \bigcap\{\mathbf{N}\subseteq\mathbf{V}\,|\,(\forall i\in I:\ \mathbf{N}_i\subseteq\mathbf{N})\wedge(\mathbf{N}\ \textit{ist Nebenraum in}\ \mathbf{V})\}$$

von $(\mathbf{N}_i \,|\, i\in I)$*.*

Wegen Lemma 7.6 ist der Verbindungsraum von Nebenräumen wieder ein Nebenraum

7.1.3 Affiner Raum

Es werden nun Punkte, Geraden, etc. einer affinen Geometrie modelliert.

Definition 7.8 *Sei* $\mathbf{N} = \mathbf{a} + \mathbf{U}$ *Nebenraum in* $\mathbf{V}$*. Die Menge*

$$\mathcal{A}(\mathbf{N})$$

aller Nebenräume in $\mathbf{V}$*, die in* $\mathbf{N}$ *enthalten sind, heißt* **affiner Raum** *oder* **affine Geometrie** *von* $\mathbf{N}$*.*

Die nulldimensionalen Elemente von $\mathcal{A}(\mathbf{N})$ *heißen* **Punkte***, die eindimensionalen* **Geraden** *und die zweidimensionalen* **Ebenen***.*

Man beachte den kleinen Unterschied zwischen einem Punkt $p \in \mathcal{A}(\mathbf{N})$ und einem Vektor $\mathbf{p} \in \mathbf{N}$. Ein Punkt p ist ein nulldimensionaler Nebenraum, d.h. er ist eine einelementige Menge $p = \{\mathbf{p}\}$.

Die Enthaltenseinsrelation zwischen den geometrischen Objekten wird einfach durch die mengentheoretische Inklusion induziert. Auch die Parallelitätsrelation kann leicht eingeführt werden.

Definition 7.9 *Zwei Nebenräume* $\mathbf{N}_1 = \mathbf{a}_1 + \mathbf{U}_1$, $\mathbf{N}_1 = \mathbf{a}_1 + \mathbf{U}_1$ *heißen parallel, i.Z.* $\mathbf{N}_1 \| \mathbf{N}_2$*, wennn* $\mathbf{U}_1 \subseteq \mathbf{U}_2$ *oder* $\mathbf{U}_2 \subseteq \mathbf{U}_1$*.*

Der Vorteil der allgemeinen Definition von $\mathcal{A}(\mathbf{N})$ mit einem Nebenraum $\mathbf{N}$ ist, daß man uneingeschränkt affine Unterräume definieren kann.

Definition 7.10 *Sei* $\mathcal{A}(\mathbf{N})$ *ein affiner Raum und* $\mathbf{M} \in \mathcal{A}(\mathbf{N})$*. Dann heißt* $\mathcal{A}(\mathbf{M})$ **affiner Unterraum** *von* $\mathcal{A}(\mathbf{N})$*.*

Definition 7.11 *Sei* $(\mathcal{A}(\mathbf{N}_i) \,|\, i \in I)$ *ein System von affinen Räumen mit Nebenräumen* $\mathbf{N}_i$*,* $i \in I$*, in einem Vektorräumen* $\mathbf{V}$*.*

Ist $\bigcap_{i \in I} \mathbf{N}_i \neq \emptyset$*, so heißt*

$$\bigcap_{i \in I} \mathcal{A}(\mathbf{N}_i) = \mathcal{A}\left(\bigcap_{i \in I} \mathbf{N}_i\right)$$

affiner Durchschnittsraum *von* $(\mathcal{A}(\mathbf{N}_i) \,|\, i \in I)$*.*

Entsprechend heißt

$$\bigvee_{i \in I} \mathcal{A}(\mathbf{N}_i) := \mathcal{A}\left(\bigvee_{i \in I} \mathbf{N}_i\right)$$

affiner Verbindungsraum *von* $(\mathcal{A}(\mathbf{N}_i) \,|\, i \in I)$*.*

7.1.4 Schnitt- und Verbindungraum

Lemma 7.12 *Für zwei Nebenräume* $\mathbf{N}_1 = \mathbf{a}_1 + \mathbf{U}_1$, $\mathbf{N}_2 = \mathbf{a}_2 + \mathbf{U}_2$ *in* $\mathbf{V}$ *gilt*

$$\mathbf{N}_1 \vee \mathbf{N}_2 = \mathbf{a}_1 + ([\mathbf{a}_2 - \mathbf{a}_1] + \mathbf{U}_1 + \mathbf{U}_2).$$

Lemma 7.13 *Für zwei Nebenräume* $\mathbf{N}_1 = \mathbf{a}_1 + \mathbf{U}_1$, $\mathbf{N}_2 = \mathbf{a}_2 + \mathbf{U}_2$ *in* $\mathbf{V}$ *gilt*

$$\mathbf{N}_1 \cap \mathbf{N}_2 \neq \emptyset \Longleftrightarrow \mathbf{a}_2 - \mathbf{a}_1 \in \mathbf{U}_1 + \mathbf{U}_2.$$

Satz 7.14 *Seien* $\mathbf{N}_1 = \mathbf{a}_1 + \mathbf{U}_1$, $\mathbf{N}_2 = \mathbf{a}_2 + \mathbf{U}_2$ *zwei Nebenräume in* $\mathbf{V}$.

1. *Ist* $\mathbf{N}_1 \cap \mathbf{N}_2 \neq \emptyset$, *so gilt*
$$\mathbf{N}_1 \vee \mathbf{N}_2 = \mathbf{a}_1 + (\mathbf{U}_1 + \mathbf{U}_2).$$
2. *Ist* $\mathbf{N}_1 \cap \mathbf{N}_2 = \emptyset$, *so gilt*
$$\mathbf{N}_1 \vee \mathbf{N}_2 = \mathbf{a}_1 + ([\mathbf{a}_2 - \mathbf{a}_1] \oplus (\mathbf{U}_1 + \mathbf{U}_2)).$$

Satz 7.15 *Seien* $\mathbf{N}_1 = \mathbf{a}_1 + \mathbf{U}_1$, $\mathbf{N}_2 = \mathbf{a}_2 + \mathbf{U}_2$ *zwei endlichdimensionale Nebenräume in* $\mathbf{V}$.

1. *Ist* $\mathbf{N}_1 \cap \mathbf{N}_2 \neq \emptyset$, *so gilt*
$$\dim \mathbf{N}_1 + \dim \mathbf{N}_2 = \dim(\mathbf{N}_1 \vee \mathbf{N}_2) + \dim(\mathbf{N}_1 \cap \mathbf{N}_2).$$
2. *Ist* $\mathbf{N}_1 \cap \mathbf{N}_2 = \emptyset$, *so gilt*
$$\dim \mathbf{N}_1 + \dim \mathbf{N}_2 = \dim(\mathbf{N}_1 \vee \mathbf{N}_2) + \dim(\mathbf{U}_1 \cap \mathbf{U}_2) - 1.$$

Satz 7.16 *Seien* $\mathbf{N}_1 = \mathbf{a}_1 + \mathbf{U}_1$, $\mathbf{N}_2 = \mathbf{a}_2 + \mathbf{U}_2$ *zwei endlichdimensionale Nebenräume in* $\mathbf{V}$.

1. *Ist* $\mathbf{N}_1 \cap \mathbf{N}_2 \neq \emptyset$, *so gilt*
$$\mathbf{N}_1 \| \mathbf{N}_2 \Longleftrightarrow \mathbf{N}_1 \subseteq \mathbf{N}_2 \vee \mathbf{N}_2 \subseteq \mathbf{N}_1.$$
2. *Ist* $\mathbf{N}_1 \cap \mathbf{N}_2 = \emptyset$, *so gilt*
$$\mathbf{N}_1 \| \mathbf{N}_2 \Longleftrightarrow \dim(\mathbf{N}_1 \vee \mathbf{N}_2) = \max(\dim \mathbf{N}_1, \dim \mathbf{N}_2) + 1.$$

Allein aus den Sätzen 7.15 und 7.16 lassen sich die folgenden bekannten Eigenschaften der zweidimensionalen und dreidimensionalen Geometrie ableiten.

Satz 7.17 *In einer affinen Ebene* $\mathcal{A}(\mathbf{V})$, *d.h.* $\dim \mathbf{V} = 2$, *gelten die folgenden beiden Eigenschaften:*

1. *Der Verbindungsraum zweier verschiedener Punkte ist eine Gerade.*
2. *Der Durchschnitt zweier nichtparalleler Geraden ist ein Punkt.*

Satz 7.18 *In einem dreidimensionalen affinen Raum* $\mathcal{A}(\mathbf{V})$, *d.h.* $\dim \mathbf{V} = 3$, *gelten die folgenden Eigenschaften:*

1. *Der Verbindungsraum zweier verschiedener Punkte ist eine Gerade.*
2. *Der Durchschnitt zweier nichtparalleler Ebenen ist eine Gerade.*
3. *Der Verbindungsraum zweier Geraden, die genau einen Punkt als Schnitt haben, ist eine Ebene.*
4. *Der Durchschnitt zweier nichtparalleler Geraden, die in einer Ebenen liegen, ist ein Punkt.*
5. *Der Verbindungsraum zweier verschiedener paralleler Geraden ist eine Ebene.*
6. *Der Verbindungsraum einer Geraden mit einem Punkt, der nicht auf dieser Geraden liegt, ist eine Ebene.*
7. *Der Durchschnitt einer Ebenen mit einer Geraden, die nicht parallel liegen, ist ein Punkt.*

7.1.5 Affine Linearkombinationen und affine Koordinaten

Definition 7.19 *Eine Linearkombination*

$$\sum_{i=1}^{n} x_i \mathbf{a}_i$$

von Vektoren $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ *(eines Vektorraums* $\mathbf{V}$*) heißt* **affine Linearkombination**, *wenn*

$$\sum_{i=1}^{n} x_i = 1.$$

Lemma 7.20 *Sei* $\mathbf{N} = \mathbf{a} + \mathbf{U}$ *eine Nebenraum (in einem Vektorraum* $\mathbf{V}$*). Dann ist jede affine Linearkombination von Vektoren aus* $\mathbf{N}$ *wieder in* $\mathbf{N}$ *enthalten.*

Definition 7.21 *Sei* $\mathbf{T}$ *eine Teilmenge eines Vektorraums* $\mathbf{V}$*. Dann heißt die Menge aller affinen Linearkombinationen von Vektoren aus* $\mathbf{T}$ **affine Hülle**

$$H(\mathbf{T})$$

von $\mathbf{T}$*.*

Lemma 7.22 *Die affine Hülle* $H(\mathbf{T})$ *einer Teilmenge* $\mathbf{T}$ *eines Vektorraums* $\mathbf{V}$ *ist ein Nebenraum in* $\mathbf{V}$*:*

$$H(\mathbf{T}) = \mathbf{t}_0 + [\{\mathbf{t}' - \mathbf{t}'' \,|\, \mathbf{t}', \mathbf{t}'' \in \mathbf{T}\}] \qquad (\mathbf{t}_0 \in \mathbf{T}).$$

Satz 7.23 *Sei* $\mathbf{T}$ *Teilmenge eines Vektorraums* $\mathbf{V}$*. Dann ist die affine Hülle* $H(\mathbf{T})$ *der kleinste Nebenraum in* $\mathbf{V}$*, der* $\mathbf{T}$ *enthält:*

$$H(\mathbf{T}) = \bigcap \{\mathbf{N} \in \mathcal{A}(\mathbf{V}) \,|\, \mathbf{T} \subseteq \mathbf{N}\}.$$

Definition 7.24 *Sei* $\mathbf{T}$ *Teilmenge eines Nebenraums* $\mathbf{N} = \mathbf{a} + \mathbf{U}$ *(in einem Vektorraum* $\mathbf{V}$*).*

$\mathbf{T}$ *heißt* **affin unabhängig**, *wenn*

$$\forall \mathbf{t} \in \mathbf{T} : \ \mathbf{t} \notin H(\mathbf{T} \setminus \{\mathbf{t}\}).$$

$\mathbf{T}$ *heißt* **affines Erzeugendensystem** *von* $\mathbf{N}$*, wenn*

$$H(\mathbf{T}) = \mathbf{N}.$$

$\mathbf{T}$ *heißt* **affine Basis** *von* $\mathbf{N}$*, wenn* $\mathbf{T}$ *ein affin unabhängiges Erzeugendensystem von* $\mathbf{N}$ *ist.*

Satz 7.25 *Sei* $\mathbf{T}$ *eine Teilmenge eines Nebenraums* $\mathbf{N} = \mathbf{a} + \mathbf{U}$ *und* $\mathbf{t}_0 \in \mathbf{T}$*.* $\mathbf{T}$ *ist genau dann affin unabhängig (bzw. ein affines Erzeugendensystem von* $\mathbf{N}$ *bzw. eine affine Basis von* $\mathbf{N}$*), wenn*

$$\mathbf{B} = \{\mathbf{t} - \mathbf{t}_0 \,|\, \mathbf{t} \in \mathbf{T} \setminus \{\mathbf{t}_0\}\}$$

linear unabhängig in $\mathbf{U}$ *ist (bzw. Erzeugendensystem von* $\mathbf{U}$ *ist bzw. Basis von* $\mathbf{U}$ *ist).*

Definition 7.26 *Sei* $\mathbf{N} = \mathbf{a} + \mathbf{U}$ *ein Nebenraum mit endlicher Dimension* $\dim \mathbf{N} = n$ *und* $\mathbf{u}, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ *eine affine Basis von* $\mathbf{N}$*. Dann ist* $\mathbf{B} := \{\mathbf{p}_1 - \mathbf{u}, \mathbf{p}_2 - \mathbf{u}, \ldots, \mathbf{p}_n - \mathbf{u}\}$ *eine Basis von* $\mathbf{U}$ *und die Abbildung*

$$\Phi_{\mathbf{u},\mathbf{B}} = \Phi_{\mathbf{B}} \circ \tau_{-\mathbf{u}} : \mathbf{N} \to K^{n\times 1}, \quad \mathbf{c} \mapsto \Phi_{\mathbf{B}}(\mathbf{c} - \mathbf{u})$$

heißt **Koordinatisierung** *zum* **affinen Koordinatensystem** $\mathbf{u}, \mathbf{p}_1, \ldots, \mathbf{p}_n$ *mit* **Ursprung** $\{\mathbf{u}\}$ *und* **Einheitspunkten** $\{\mathbf{p}_1\}, \{\mathbf{p}_2\}, \ldots, \{\mathbf{p}_n\}$*.*

Ein Vektor $\mathbf{c} \in \mathbf{N}$ hat also die affinen Koordination $\Phi_{\mathbf{u},\mathbf{B}}(\mathbf{c}) = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^T$, wenn

$$\mathbf{c} = \mathbf{u} + \sum_{i=1}^{n} x_i(\mathbf{p}_i - \mathbf{u}) = \left(1 - \sum_{i=1}^{n} x_i\right)\mathbf{u} + \sum_{i=1}^{n} x_i \mathbf{p}_i,$$

d.h. $\mathbf{c}$ kann als um $\mathbf{u}$ verschobene Linearkombination der (Basis-)Vektoren $(\mathbf{p}_i - \mathbf{u})$ oder aber als affine Linearkombination der Vektoren $\mathbf{u}$ und $\mathbf{p}_i$ angesehen werden.

Definition 7.27 *Das* **Teilverhältnis** $TV(\mathbf{a}, \mathbf{p}, \mathbf{u})$ *dreier Vektoren (resp. Punkte)* $\mathbf{a}, \mathbf{p}, \mathbf{u}$ *eines affinen Raums, die auf einer Geraden liegen, ist durch*

$$TV(\mathbf{a}, \mathbf{p}, \mathbf{u}) = x \in K$$

gegeben, wenn

$$\mathbf{a} = \mathbf{u} + x(\mathbf{p} - \mathbf{u})$$

ist.

Satz 7.28 *Sei* $\mathbf{u}, \mathbf{p}$ *eine affine Basis einer affinen Geraden* g*. Dann gilt für* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in g$ *mit den Koordinaten* a, b, c *(d.h.* $\mathbf{a} = \mathbf{u} + a(\mathbf{p} - \mathbf{u})$*,* $\mathbf{b} = \mathbf{u} + b(\mathbf{p} - \mathbf{u})$ *und* $\mathbf{c} = \mathbf{u} + c(\mathbf{p} - \mathbf{u})$*)*

$$TV(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \frac{a - c}{b - c}.$$

Beispielweise kann man daraus unmittelbar

$$TV(\mathbf{a}, \mathbf{c}, \mathbf{b}) = 1 - TV(\mathbf{a}, \mathbf{b}, \mathbf{c}) \quad \text{und} \quad TV(\mathbf{b}, \mathbf{a}, \mathbf{c}) = TV(\mathbf{a}, \mathbf{b}, \mathbf{c})^{-1}$$

ablesen.

Für Körper K der Charakteristik $\mathrm{char}(K) \neq 2$ läßt sich mit Hilfe des Teilverhältnisses auch der **Mittelpunkt** $\mathbf{m}$ zweier Vektoren (Punkte) $\mathbf{a}, \mathbf{b}$ durch

$$TV(\mathbf{m}, \mathbf{a}, \mathbf{b}) = 2^{-1}$$

definieren. Eine äquivalente Definition ist

$$\mathbf{m} = 2^{-1}(\mathbf{a} + \mathbf{b}).$$

Üblicherweise stellt man sich den Mittelpunkt von $\mathbf{a}$ und $\mathbf{b}$ als einen Punkt vor, der zwischen $\mathbf{a}$ und $\mathbf{b}$ liegt. Eine präzise Formulierung eines Zwischenbegriffs gelingt (wieder mit Hilfe des Teilverhältnisses), wenn man einen angeordneten Körper zugrudelegt. Man sagt, $\mathbf{m}$ liegt **zwischen** $\mathbf{a}$ und $\mathbf{b}$, wenn

$$TV(\mathbf{m}, \mathbf{a}, \mathbf{b}) \in K^+ \quad \text{und} \quad TV(\mathbf{m}, \mathbf{b}, \mathbf{a}) \in K^+$$

gilt.

7.1.6 Affine Abbildungen

Definition 7.29 *Sei* $\mathbf{V}$ *ein Vektoraum und* $\mathbf{t} \in \mathbf{V}$*. Dann heißt die Abbildung*

$$\tau_{\mathbf{t}} : \mathbf{V} \to \mathbf{V}, \quad \mathbf{a} \mapsto \mathbf{a} + \mathbf{t}$$

Verschiebung

Wegen $\tau_{\mathbf{t}_1} \circ \tau_{\mathbf{t}_2} = \tau_{\mathbf{t}_1+\mathbf{t}_2}$ bilden die Verschiebungen mit der Komposition eine abelsche Gruppe, die zu $\langle \mathbf{V}, + \rangle$ isomorph ist. Der Vektorraum $\mathbf{V}$ (interpretiert als Menge der Verschiebungen) operiert also auf sich selbst. Offensichtlich führt $\tau_{\mathbf{t}}$ auch Nebenräume in Nebenräume über. $\mathbf{V}$ (interpretiert als Menge der Verschiebungen) operiert daher auch auf der affinen Geometrie $\mathcal{A}(\mathbf{V})$. Entprechend operiert $\mathbf{U}$ auf $\mathcal{A}(\mathbf{a} + \mathbf{U})$.

Definition 7.30 *Seien* K, K' *zwei Körper. Eine bijektive Abbildung* $\zeta : K \to K'$ *heißt* **Isomorphismus***, wenn für alle* $x, y \in K$

$$\zeta(x + y) = \zeta(x) + \zeta(y) \quad \textit{und} \quad \zeta(x \cdot y) = \zeta(x) \cdot \zeta(y)$$

gelten.

Ist insbesondere $K = K'$*, so heißt ein Isomorphismus* $\zeta : K \to K$ **Automorphismus***. Die Menge aller Körperautomorphismen eines Körpers* K *wird mit* $Aut(K)$ *bezeichnet und bildet mit der Hintereinanderausführung eine Gruppe, die* **Automorphismengruppe***.*

In vielen Fällen kennt man alle Körperautomorphismen.

Satz 7.31

1. $Aut(\mathbb{Q}) = \{\mathrm{id}\}$.
2. $Aut(\mathbb{R}) = \{\mathrm{id}\}$.
3. *Die Automorphismen* ζ *von* $\mathbb{C}$ *mit* $\zeta|_{\mathbb{R}} = \mathrm{id}_{\mathbb{R}}$ *sind nur die Identität* $\mathrm{id}_{\mathbb{C}}$ *und die komplexe Konjugation* $z \mapsto \overline{z}$*.*
4. *Sei* K *ein endlicher Körper mit* $|K| = q = p^k$ *Elementen (mit einer Primzahl* p*), dann sind alle Automorphismen von der Form* $x \mapsto x^{p^l}$*,* $0 \leq l < k$*. Es gibt also genau* k *Automorphismen.*

Definition 7.32 *Sei* $\mathbf{V}$ *ein Vektorraum über eine Körper* K *und* $\mathbf{V}'$ *ein Vektorraum über dem Körper* K'*. Eine Abbildung* $f : \mathbf{V} \to \mathbf{V}'$ *heißt* **semilinear***, wenn es einen Körperisomorphismus* $\zeta : K \to K'$ *gibt, so daß für alle* $\mathbf{a}, \mathbf{b} \in \mathbf{V}$ *und* $x, y \in K$

$$f(x\mathbf{a} + y\mathbf{b}) = \zeta(x)f(\mathbf{a}) + \zeta(y)f(\mathbf{b})$$

gilt.

Definition 7.33 *Seien $\mathcal{A}(\mathbf{a}+\mathbf{U})$, $\mathcal{A}(\mathbf{a}'+\mathbf{U}')$ zwei affine Räume. Eine Abbildung*

$$\alpha : \mathcal{A}(\mathbf{a}+\mathbf{U}) \to \mathcal{A}(\mathbf{a}'+\mathbf{U}')$$

heißt **affine Abbildung**, *wenn es Vektoren $\mathbf{t} \in \mathbf{a}+\mathbf{U}$, $\mathbf{t} \in \mathbf{a}'+\mathbf{U}'$ und eine semilineare Abbildung $f : \mathbf{U} \to \mathbf{U}'$ gibt, so daß*

$$\alpha = \tau_{\mathbf{t}'} \circ f \circ \tau_{-\mathbf{t}}.$$

Ist f zusätzlich bijektiv, so heißt α **Affinität**.

Ist f eine lineare Abbildung, d.h. $K = K'$ und $\zeta = \mathrm{id}$, so nennt man α **projektive affine Abbildung** *bzw.* **projektive Affinität**, *wenn f ein Vektorraumisomorphismus ist.*

Man beachte, daß hier aus notationstechnischen Gründen nicht streng zwischen der Abbildung $\alpha : \mathbf{a}+\mathbf{U} \to \mathbf{a}'+\mathbf{U}'$ und der eigentlich davon induzierten Abbildung $\alpha : \mathcal{A}(\mathbf{a}+\mathbf{U}) \to \mathcal{A}(\mathbf{a}'+\mathbf{U}')$ unterschieden wird.

Satz 7.34 *Sei $\alpha = \tau_{\mathbf{t}'} \circ f \circ \tau_{-\mathbf{t}} : \mathcal{A}(\mathbf{a}+\mathbf{U}) \to \mathcal{A}(\mathbf{a}'+\mathbf{U}')$ eine affine Abbildung. Dann gelten die folgenden Eigenschaften:*

1. $\forall \mathbf{p}, \mathbf{q} \in \mathbf{a}+\mathbf{U} : \ f(\mathbf{p}-\mathbf{q}) = \alpha(\mathbf{p}) - \alpha(\mathbf{q})$, *d.h. f ist durch α fixiert.*
2. $\mathbf{t}' = \alpha(\mathbf{t})$.
3. $\forall \mathbf{s} \in \mathbf{a}+\mathbf{U} : \ \alpha = \tau_{\alpha(\mathbf{s})} \circ f \circ \tau_{-\mathbf{s}}$.

Ähnlich wie bei linearen Abbildungen zwischen endlichdimensionalen Vektorräumen kann man auch projektive affine Abbildungen bezüglich affiner Koordinaten betrachten und sie mit Hilfe einer Matrix (und einer Verschiebung) realisieren.

Definition 7.35 *Es seien $\mathbf{u}, \mathbf{p}_1, \ldots, \mathbf{p}_n$ resp. $\mathbf{u}', \mathbf{p}'_1, \ldots, \mathbf{p}'_m$ affine Basen der affinen Räume $\mathcal{A}(\mathbf{a}+\mathbf{U})$ resp. $\mathcal{A}(\mathbf{a}'+\mathbf{U}')$, und es bezeichnen $\mathbf{B} = \{\mathbf{p}_1 - \mathbf{u}, \ldots, \mathbf{p}_n - \mathbf{u}\}$ resp. $\mathbf{B}' = \{\mathbf{p}'_1 - \mathbf{u}', \ldots, \mathbf{p}'_m - \mathbf{u}'\}$ die dazugehörigen Basen von $\mathbf{U}$ resp. $\mathbf{U}'$.*

Ist $\alpha : \mathcal{A}(\mathbf{a}+\mathbf{U}) \to \mathcal{A}(\mathbf{a}'+\mathbf{U}')$ eine affine Abbildung, so bezeichnet man die Abbildung

$$\Phi_{\mathbf{u}',\mathbf{B}'} \circ \alpha \circ \Phi^{-1}_{\mathbf{u},\mathbf{B}} : \mathcal{A}(K^{n\times 1}) \to \mathcal{A}(K^{m\times 1})$$

als **Koordinatendarstellung** *von α bezüglich der affinen Basen $\mathbf{u}, \mathbf{p}_1, \ldots, \mathbf{p}_n$ und $\mathbf{u}', \mathbf{p}'_1, \ldots, \mathbf{p}'_m$.*

Satz 7.36 *Ist α eine projektive affine Abbildung, so wird die Koordinatendarstellung $\Phi_{\mathbf{u}',\mathbf{B}'} \circ \alpha \circ \Phi^{-1}_{\mathbf{u},\mathbf{B}}$ durch die Zuordung*

$$\mathbf{x} \mapsto A\mathbf{x} + \mathbf{b}$$

realisiert, wobei $A \in K^{m\times n}$ und $\mathbf{b} \in K^{m\times 1}$ durch

$$A = \Phi_{\mathbf{B}\mathbf{B}'}(f) \quad \textit{und} \quad \mathbf{b} = \Phi_{\mathbf{B}'}(\alpha(\mathbf{u}) - \mathbf{u}')$$

gegeben sind.

Das Analogon zum Fortsetzungssatz linearer Abbildungen ist die folgende Eigenschaft.

Satz 7.37 *Es seien $\mathcal{A}(\mathbf{N})$ und $\mathcal{A}(\mathbf{N}')$ affine Räume über isomorphen Körpern K und K' und $\mathbf{u}, \mathbf{p}_1, \ldots, \mathbf{p}_n$ eine affine Basis von $\mathcal{A}(\mathbf{N})$. Dann gibt es bei (beliebiger) Vorgabe von $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_n \in \mathbf{N}'$ und einem Insomorphismus $\zeta : K \to K'$ genau eine affine Abbildung $\alpha : \mathcal{A}(\mathbf{N}) \to \mathcal{A}(\mathbf{N}')$ mit*

$$\alpha(\mathbf{u}) = \mathbf{q}_0 \quad \textit{und} \quad \alpha(\mathbf{p}_i) = \mathbf{q}_i \quad (1 \leq i \leq n).$$

Eine weitere Eigenschaft affiner Abbildungen betrifft das Teilverhältnis.

Satz 7.38 *Das Teilverhältnis ist unter affinen Abbildungen $\alpha : \mathcal{A}(\mathbf{N}) \to \mathcal{A}(\mathbf{N}')$ invariant, d.h. für beliebige drei Vektoren (Punkte) $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{N}$, die auf einer Geraden liegen, gilt*

$$TV(\alpha(\mathbf{a}), \alpha(\mathbf{b}), \alpha(\mathbf{c}) = \zeta(TV(\mathbf{a}, \mathbf{b}, \mathbf{c})).$$

7.2 Projektive Geometrie

7.2.1 Vorbemerkungen

Zwei verschiedene affine Geraden (in einer affinen Ebene) haben entweder einen Schnittpunkt oder sie sind parallel. Durch diese Unterscheidung müssen in der affinen Geometrie oft viele Fallunterscheidungen getroffen werden. Es stellt sich jedoch heraus, daß man diese Fallunterscheidungen umgehen kann, wenn man parallelen Geraden einen Schnittpunkt zuordnet. Diese zusätzlichen Punkte liegen natürlich außerhalb der affinen Geometrie und werden als **Fernpunkte**[1] bezeichnet. Erweitert man also eine affine Ebene um dieser Fernpunkte und um die sogenannte **Ferngerade**, die alle Fernpunkte umfaßt, so erhält man eine eine Geometrie, wo (wie bereits in der affinen Geometrie) die Verbindung zweier verschiedener Punkte immer eine Gerade ist und wo (im Gegensatz zur affinen Geometrie) der Schnitt zweier verschiedener Geraden immer ein Punkt ist.

Auf diese Weise gewinnt man eine sogenannte **projektive Geometrie**. Im folgenden wird allerding ein anderer Zugang zu projektiven Geometrien vorgestellt, der sich wieder auf die Theorie der Vektorräume stützt. Daß auch in diesem Fall eine projektive Geometrie als *Fortsetzung* einer affinen Geometie gesehen werden kann, ist Inhalt des Einbettungssatzes 7.46.

7.2.2 Projektiver Raum

Definition 7.39 *Sei $\mathbf{V}$ ein Vektorraum (über einem Körper K). Die Menge aller Unterräume von $\mathbf{V}$*

$$\mathcal{P}(\mathbf{V})$$

heißt **projekiver Raum** *oder* **projektive Geometrie** *auf $\mathbf{V}$.*

Die eindimensionalen Unterräume in $\mathcal{P}(\mathbf{V})$ heißen **projektive Punkte**, *die zweidimensionalen Unterräume von $\mathcal{P}(\mathbf{V})$* **projektive Geraden** *und die dreidimensionalen Unterräume* **projektive Ebenen**.

[1] Diese Bezeichnung rührt von der Vorstellung, daß sich zwei parallele Geraden in einem unendlich fernen Punkt scheiden.

Definition 7.40 *Die* **projektive Dimension** $\operatorname{pdim} \mathbf{U}$ *eines Unterraums* $\mathbf{U} \leq \mathbf{V}$ *wird durch*

$$\operatorname{pdim} \mathbf{U} := \dim \mathbf{U} - 1$$

definiert.

Definition 7.41 *Sei* $\mathcal{P}(\mathbf{V})$ *ein projektiver Raum und* $\mathbf{U} \in \mathcal{P}(\mathbf{V})$*. Dann heißt* $\mathcal{P}(\mathbf{U})$ **projektiver Unterraum** *von* $\mathcal{P}(\mathbf{V})$*.*

Der projektive Unterraum $\mathcal{P}(\{\mathbf{0}\})$ *wird als* **leerer projektiver Raum** *bezeichnet.*

Man beachte, daß der leere projektive Raum zwar im mengentheoretischen Sinn nicht leer ist, aber weder Punkte noch Geraden etc. enthält. Seine projektive Dimension ist

$$\operatorname{pdim}\{\mathbf{0}\} = -1.$$

Definition 7.42 *Sei* $(\mathbf{U}_i \,|\, i \in I)$ *ein System von Unterräumen eines Vektorraums* $\mathbf{V}$*. Dann heißt*

$$\bigcap_{i \in I} \mathcal{P}(\mathbf{U}_i) = \mathcal{P}\left(\bigcap_{i \in I} \mathbf{U}_i\right)$$

projektiver Schnittraum *und*

$$\bigvee_{i \in I} \mathcal{P}(\mathbf{U}_i) := \mathcal{P}\left(\sum_{i \in I} \mathbf{U}_i\right)$$

projektiver Verbindungsraum.

Satz 7.43 *Für zwei Unterräume* $\mathbf{U}_1, \mathbf{U}_2$ *eines Vektorraums* $\mathbf{V}$ *gilt*

$$\operatorname{pdim} \mathbf{U}_1 + \operatorname{pdim} \mathbf{U}_2 = \operatorname{pdim} (\mathbf{U}_1 \vee \mathbf{U}_2) + \operatorname{pdim} (\mathbf{U}_1 \cap \mathbf{U}_2).$$

Wie in der affinen Geometrie können nun allgemeine Eigenschaften der Objekte endlichdimensionaler Geometrien angegeben werden. Beim Vergleich fällt auf, daß die projektiven Versionen einfacher sind. Die Unterscheidung zwischen parallel und nichtparallel entfällt.

Satz 7.44 *In einer projektiven Ebene* $\mathcal{P}(\mathbf{V})$*, d.h.* $\operatorname{pdim} \mathbf{V} = 2$ *gelten die folgenden beiden Eigenschaften:*

1. *Der Verbindungsraum zweier verschiedener Punkte ist eine Gerade.*
2. *Der Durchschnitt zweier verschiedener Geraden ist ein Punkt.*

Satz 7.45 *In einem dreidimensionalen projektiven Raum* $\mathcal{P}(\mathbf{V})$*, d.h.* $\operatorname{pdim} \mathbf{V} = 3$*, gelten die folgenden Eigenschaften:*

1. *Der Verbindungsraum zweier verschiedener Punkte ist eine Gerade.*

2. *Der Durchschnitt zweier verschiedener Ebenen ist eine Gerade.*

3. *Der Verbindungsraum zweier verschiedener sich schneidender Geraden ist eine Ebene.*

4. *Der Durchschnitt zweier verschiedener Geraden, die in einer Ebenen liegen, ist ein Punkt.*

5. *Der Verbindungsraum einer Geraden mit einem Punkt, der nicht auf dieser Geraden liegt, ist eine Ebene.*

6. *Der Durchschnitt einer Ebenen mit einer nicht in dieser Ebenen enthaltenen Geraden ist ein Punkt.*

7.2.3 Einbettungssatz

Satz 7.46 *Sei $\mathcal{P}(\mathbf{V})$ eine endlichdimensionaler projektiver Raum, $\mathbf{H}$ eine Hyperebene von $\mathbf{V}$, d.h. ein Unterraum mit $\dim \mathbf{H} = \dim \mathbf{V} - 1$, und $\mathbf{a} \in \mathbf{V} \setminus \mathbf{H}$. Dann hat die Abbildung*

$$\varphi : \mathcal{A}(\mathbf{a} + \mathbf{H}) \to \mathcal{P}(\mathbf{V}), \quad \mathbf{S} \mapsto [\mathbf{S}]$$

die folgenden Eigenschaften: ($\mathbf{S}, \mathbf{T}$, resp. $\mathbf{S}_i, i \in I$ bezeichnen Nebenräume in $\mathbf{a} + \mathbf{H}$)

1. *φ ist injektiv.*

2. $\varphi(\mathcal{A}(\mathbf{a} + \mathbf{H})) = \{\mathbf{U} \in \mathcal{P}(\mathbf{V}) \,|\, \mathbf{U} \not\subseteq \mathbf{H}\}$.

3. $\mathbf{S} \subseteq \mathbf{T} \iff \varphi(\mathbf{S}) \subseteq \varphi(\mathbf{T})$.

4. $\varphi\left(\bigcap_{i\in I} \mathbf{S}_i\right) = \bigcap_{i\in I} \varphi(\mathbf{S}_i)$, *falls* $\bigcap_{i\in I} \mathbf{S}_i \neq \emptyset$.

5. $\varphi\left(\bigvee_{i\in I} \mathbf{S}_i\right) = \bigvee_{i\in I} \varphi(\mathbf{S}_i) = \sum_{i\in I} \varphi(\mathbf{S}_i)$.

6. $\dim \mathbf{S} = \operatorname{pdim} \varphi(\mathbf{S})$.

7. $\mathbf{S} \| \mathbf{T} \iff (\varphi(\mathbf{S}) \cap \mathbf{H} \subseteq \varphi(\mathbf{T}) \cap \mathbf{H}) \vee (\varphi(\mathbf{T}) \cap \mathbf{H} \subseteq \varphi(\mathbf{S}) \cap \mathbf{H})$.

Dieser Satz zeigt einerseits, wie ein affiner Raum als Teil eines projektiven Raums gesehen werden kann, resp. wie ein affiner Raum zu einem projektiven Raum erweitert werden kann.

Fixiert man in einem projektiven Raum eine Hyperebene $\mathbf{H}$ und betrachtet nur die reduzierte Geometrie

$$\mathcal{A} := \mathcal{P}(\mathbf{V}) \setminus \mathcal{P}(\mathbf{H}) = \{\mathbf{U} \in \mathcal{P}(\mathbf{V}) \,|\, \mathbf{U} \not\subseteq \mathbf{H}\},$$

so stehen die Elemente aus $\mathcal{A}$ in einem eineindeutigen Verhältnis zu den Elementen aus der affinen Geometrie $\mathcal{A}(\mathbf{a} + \mathbf{H})$ ($\mathbf{a} \notin \mathbf{H}$). Inklusionen, Dimensionen, Schnitte und Verbindungen entsprechen einander. Weiters kann mit 7. auch ein Parallelitätsbegriff definiert werden. $\mathcal{A}$ ist nichts anderes als ein Modell für eine affine Geometrie, interpretiert als Teil einer projektiven Geometrie.

Geht man andererseits von einem affinen Raum $\mathcal{A}(\mathbf{a}+\mathbf{H})$ (mit $\mathbf{a} \notin \mathbf{H}$[2]) aus, so kann man mit Hilfe dieses Satzes $\mathcal{A}(\mathbf{a}+\mathbf{H})$ in den projektiven Raum $\mathcal{P}([\mathbf{a}]\oplus\mathbf{H}])$ einbetten. $\mathcal{A}(\mathbf{a}+\mathbf{H})$ wird durch jene Elemente ergänzt, die in $\mathbf{H}$ liegen. Sind z.B. zwei Gerade $g = \mathbf{a}_1 + [\mathbf{v}], h = \mathbf{a}_2 + [\mathbf{v}] \in \mathcal{A}(\mathbf{a} + \mathbf{H})$ parallel, so ist $P = [\mathbf{v}]$ ein projektiver Punkt in $\mathbf{H}$ mit $P = [g] \cap [h]$, d.h. die projektiven Punkte in $\mathcal{P}(\mathbf{H})$ sind gerade die *fehlenden* Schnittpunkte paralleler Geraden in $\mathcal{A}(\mathbf{a} + \mathbf{H})$. Diese werden auch als **Fernpunkte** bezeichnet. Entsprechend haben parallele Ebenen in $\mathcal{A}(\mathbf{a} + \mathbf{H})$ eine (projektive) Schnittgerade in $\mathbf{H}$ etc.

7.2.4 Projektive Basen und homogene Koordinaten

Projektive Punkte $Q = [\mathbf{q}] \in \mathcal{P}(\mathbf{V})$ werden im folgenden auch durch

$$Q = K\mathbf{q} = \{x\mathbf{q} \,|\, x \in K\}$$

bezeichnet.

Definition 7.47 *Sei* $\mathbf{T}$ *eine Menge von projektiven Punkten in* $\mathcal{P}(\mathbf{V})$.

Die **projektive Hülle** $H(\mathbf{T})$ *von* $\mathbf{T}$ *ist durch*

$$H(\mathbf{T}) := \bigvee_{Q \in \mathbf{T}} Q = [\{\mathbf{q} \in \mathbf{V} \,|\, K\mathbf{q} \in \mathbf{T}\}]$$

definiert.

$\mathbf{T}$ *heißt* **projektiv unabhängig**, *wenn*

$$\forall Q \in \mathbf{T} : \; Q \notin H(\mathbf{T} \setminus \{Q\}).$$

$\mathbf{T}$ *heißt* **projektives Erzeugendensystem** *von* $\mathcal{P}(\mathbf{V})$, *wenn*

$$H(\mathbf{T}) = \mathbf{V}.$$

Schließlich heißt $\mathbf{T}$ **projektive Basis**, *wenn* $\mathbf{T}$ *ein projektiv unabhängiges Erzeugendensystem von* $\mathcal{P}(\mathbf{V})$ *ist.*

Satz 7.48 *Sei* $\mathbf{T} = \{Q_i = K\mathbf{q}_i \,|\, i \in I\}$ *eine Menge von Punkten in* $\mathcal{P}(\mathbf{V})$.

1. $\mathbf{T}$ *ist genau dann projektiv unabhängig, wenn die Menge der Vektoren* $\{\mathbf{q}_i \,|\, i \in I\}$ *linear unabhängig in* $\mathbf{V}$ *ist.*

2. $\mathbf{T}$ *ist genau dann projektives Erzeugendensysem von* $\mathcal{P}(\mathbf{V})$, *wenn* $\{\mathbf{q}_i \,|\, i \in I\}$ *ein Erzeugendensystem von* $\mathbf{V}$ *ist.*

3. $\mathbf{T}$ *ist genau eine projektive Basis von* $\mathcal{P}(\mathbf{V})$, *wenn* $\{\mathbf{q}_i \,|\, i \in I\}$ *eine Basis von* $\mathbf{V}$ *ist.*

[2] Dies ist keine Einschränkung und kann notfalls durch eine Verschiebung des affinen Raums erzwungen werden.

Es sein nun $\mathcal{P}(\mathbf{V})$ ein projektiver Raum mit projektiver Dimension n, d.h. $\dim \mathbf{V} = n+1$. Aufgrund des vorigen Satzes könnte man projektive Koordinaten bezüglich einer projektiven Basis $\{Q_i = K\mathbf{q}_i \,|\, 0 \leq i \leq n\}$ dadurch einführen, indem man die Koordinaten eines Punktes $P = K\mathbf{p}$ mit jenen von $\mathbf{p}$ bezüglich $\{\mathbf{q}_i \,|\, 0 \leq i \leq n\}$ gleichsetzt. Hier ergeben sich aber zwei Probleme. Erstens ist $\mathbf{p}$ durch $P = K\mathbf{p}$ nicht eindeutig bestimmt, sondern nur bis auf einen Faktor $x \in K^\times$. Damit beschreiben (verschiedene) Koordinaten, die sich um einen gemeinsamen Faktor $x \in K^\times$ unterscheiden, denselben Punkt P, d.h. projektive Koordinaten sind eigentlich Punkte in $\mathcal{P}(K^{(n+1)\times 1})$. Das zweite Problem ist ein wenig diffiziler. Die soeben beschriebenen Koordinaten hängen nicht von der Basis $\{Q_i = K\mathbf{q}_i \,|\, 0 \leq i \leq n\}$ ab, sonder von den speziell gewählten Vektoren $\{\mathbf{q}_i \,|\, 0 \leq i \leq n\}$. Um dieses Problem zu lösen, muß noch ein zusätzlicher Punkt Q betrachtet werden, dem immer die Koordinaten

$$K \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

zugeordnet werden sollen.

Definition 7.49 *Sei $\mathcal{P}(\mathbf{V})$ ein projektiver Raum der projektiven Dimension n. Eine Menge $\{Q_0, Q_1, \ldots, Q_n, Q\}$ von $n+2$ Punkten in $\mathcal{P}(\mathbf{V})$ heißt* **Bezugssystem** *oder* **projektives Koordinatensystem**, *wenn je $n+1$ Punkte projektiv unabhängig sind.*

Lemma 7.50 *Ist $\{Q_0, Q_1, \ldots, Q_n, Q\}$ ein Bezugssystem eines n-dimensionalen projektiven Raums $\mathcal{P}(\mathbf{V})$, dann gibt es Vektoren $\mathbf{q}_i \in \mathbf{V}$, $0 \leq i \leq n$, mit*

$$Q_i = K\mathbf{q}_i \quad (1 \leq i \leq n) \quad \textit{und} \quad Q = K\left(\sum_{i=0}^{n} \mathbf{q}_i\right).$$

Definition 7.51 *Sei $\{Q_0 = K\mathbf{q}_0, Q_1 = K\mathbf{q}_1, \ldots, Q_n = K\mathbf{q}_n, Q = K(\mathbf{q}_0 + \mathbf{q}_1 + \cdots + \mathbf{q}_n)\}$ ein projektives Koordinatensystem eines n-dimensionalen projektiven Raums $\mathcal{P}(\mathbf{V})$ und bezeichne $\mathbf{B} = \{\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_n\}$. Dann bezeichnet man die Abbildung*

$$\mathcal{P}(\Phi_B) : \mathcal{P}(\mathbf{V}) \to \mathcal{P}(K^{(n+1)\times 1}), \quad K\mathbf{p} \mapsto K\Phi_{\mathbf{B}}(\mathbf{p})$$

als **(projektive) Koordinatenabbildung** *zum projektiven Koordinatensystem $\{Q_0, Q_1, \ldots, Q_n, Q\}$.*

Man beachte, daß durch die Forderung, daß Q die Koordinaten $K\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}^T$ haben soll, die Koordinatisierung $\mathcal{P}(\Phi_B)$ nicht von der speziellen Wahl von $\mathbf{B} = \{\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_n\}$ abhängt.

Die Koordinaten $K\Phi_{\mathbf{B}}(\mathbf{p})$ werden auch als **homogene Koordinaten** bezeichnet und werden manchmal auch durch

$$(x_0 : x_1 : \ldots : x_n) \quad \text{oder} \quad [x_0 : x_1 : \ldots : x_n]$$

dargestellt und man betrachtet zwei solche $(n+1)$-Tupel als gleich, wenn sie sich um einen Faktor $y \in K^\times$ unterscheiden.

Bei richtiger Wahl der Koordinatensysteme besteht auch ein einfacher Zusammenhang der affinen Koordinaten auf den affinen Raum $\mathcal{A}(\mathbf{a}+\mathbf{H})$ und $\mathcal{P}(\mathbf{V})$, wobei $\mathbf{H}$ eine Hyperebene bezeichnet (siehe Satz 7.46). Sei $\{\mathbf{u}, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ eine affine Basis von $\mathcal{A}(\mathbf{a}+\mathbf{H})$, d.h. $\mathbf{u} \in \mathbf{a}+\mathbf{H}$ und $\mathbf{B}' = \{\mathbf{p}_1 - \mathbf{u}, \ldots, \mathbf{p}_n - \mathbf{u}\}$ ist eine Basis von $\mathbf{H}$. Betrachtet man nun das Bezugssystem

$$Q_0 := K\mathbf{u},\ Q_1 := K(\mathbf{p}_1 - \mathbf{u}),\ \ldots,\ Q_n := K(\mathbf{p}_n - \mathbf{u}),\ Q := K\left(\mathbf{u} + \sum_{i=1}^{n}(\mathbf{p}_i - \mathbf{u})\right)$$

und die entsprechende Basis $\mathbf{B} = \{\mathbf{u}, \mathbf{p}_1 - \mathbf{u}, \ldots, \mathbf{p}_n - \mathbf{u}\}$, so entspricht ein affiner Punkt $\{\mathbf{p}\} \in \mathbf{a}+\mathbf{H}$ mit den affinen Koordinaten

$$\Phi_{\mathbf{u},\mathbf{B}'}(\mathbf{p}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

der projektive Punkte $[\mathbf{p}] = K\mathbf{p}$ mit den homogenen Koordinaten

$$\mathcal{P}(\Phi_{\mathbf{B}})(K\mathbf{p}) = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Hat umgekehrt ein projektiver Punkt $P = K\mathbf{p}$ die homogenen Koordinaten

$$\mathcal{P}(\Phi_{\mathbf{B}})(K\mathbf{p}) = \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}$$

mit $x_0 \neq 0$, dann entspricht ihn der affine Punkt $K\mathbf{p} \cap (\mathbf{a}+\mathbf{H})$ mit den affinen Koordinaten

$$\Phi_{\mathbf{u},\mathbf{B}'}(K\mathbf{p} \cap (\mathbf{a}+\mathbf{H})) = \begin{pmatrix} x_1/x_0 \\ \vdots \\ x_n/x_0 \end{pmatrix}.$$

Ist hingegen $x_0 = 0$, so hat $P = K\mathbf{p}$ keine affine Entsprechung und ist ein Fernpunkt.

In Analogie zum Teilverhältnis wird in der projektiven Geometrie das Doppelverhältnis definiert.

Definition 7.52 *Das* **Doppelverhältnis** $DV(X, Q, Q_0, Q_1)$ *von vier Punkten* $X = K\mathbf{x}, Q = K(\mathbf{q}_0 + \mathbf{q}_1), Q_0 = K\mathbf{q}_0, Q_1 = K\mathbf{q}_1$ *eines projektiven Raums, die auf einer Geraden liegen, ist durch*

$$DV(X, Q, Q_0, Q_1) = \frac{x_1}{x_0} \in K$$

gegeben, wenn

$$K\mathbf{x} = K(x_0\mathbf{q}_0 + x_1\mathbf{q}_1)$$

ist.

Beispielsweise ist

$$DV(K\mathbf{a}, K\mathbf{p}, K\mathbf{u}, K(\mathbf{p}-\mathbf{u})) = TV(\mathbf{a}, \mathbf{p}, \mathbf{u}),$$

d.h. das Doppelverhältnis geht in das Teilverhälnis über, wenn der vierte Punkt jener Fernpunkt ist, der von der (affinen) Geraden $\mathbf{a} \vee \mathbf{p} \vee \mathbf{u}$ bestimmt wird.

Satz 7.53 *Sei $Q_0 = K\mathbf{q}_0, Q_1 = K\mathbf{q}_1, Q = K(\mathbf{q}_0 + \mathbf{q}_1)$ ein projektives Koordinatensystem einer projektiven Geraden g. Haben die vier Punkte $A, B, C, D \in g$ bezüglich dieses Koordinatensystems die homogenen Koordinaten*

$$K\begin{pmatrix} a_0 \\ a_1 \end{pmatrix},\ K\begin{pmatrix} b_0 \\ b_1 \end{pmatrix},\ K\begin{pmatrix} c_0 \\ c_1 \end{pmatrix},\ K\begin{pmatrix} d_0 \\ d_1 \end{pmatrix},$$

so bestimmt sich das Doppelverhältnis durch

$$DV(A, B, C, D) = \frac{\begin{vmatrix} a_0 & c_0 \\ a_1 & c_1 \end{vmatrix}}{\begin{vmatrix} b_0 & c_0 \\ b_1 & c_1 \end{vmatrix}} : \frac{\begin{vmatrix} a_0 & d_0 \\ a_1 & d_1 \end{vmatrix}}{\begin{vmatrix} b_0 & d_0 \\ b_1 & d_1 \end{vmatrix}}.$$

Korollar 7.54 *Für kollineare projektive Punkte $K\mathbf{a}, K\mathbf{b}, K\mathbf{c}, K\mathbf{d}$ gilt*

$$DV(K\mathbf{a}, K\mathbf{b}, K\mathbf{c}, K\mathbf{d}) = \frac{TV(\mathbf{a}, \mathbf{b}, \mathbf{c})}{TV(\mathbf{a}, \mathbf{b}, \mathbf{d})}.$$

Definition 7.55 *Vier Punkte A, B, C, D einer projektiven Geraden liegen in harmonische Lage, wenn*

$$DV(A, B, C, D) = -1.$$

7.2.5 Kollineare Abbildungen

Definition 7.56 *Seien $\mathcal{P}(\mathbf{V})$ resp. $\mathcal{P}(\mathbf{V}')$ zwei projektive Räume über isomorphen Körpern K resp. K' und $f : \mathbf{V} \to \mathbf{V}'$ eine injektive semilineare Abbildung. Dann heißt die Abbildung*

$$\pi = \mathcal{P}(f) : \mathcal{P}(\mathbf{V}) \to \mathcal{P}(\mathbf{V}'), \quad K\mathbf{p} \mapsto K'f(\mathbf{p})$$

kollineare Abbildung. *Ist f zusätzlich bijektiv, so heißt π* **Kollinearität**.

Ist $K = K'$ und f eine lineare Abbildung, so heißt π **projektive kollineare Abbildung** *resp.* **projektive Kollinearität**.

Man kann in dieser Definition auch allgemeine (d.h. nicht unbedingt injektive) semilineare bzw. lineare Abbildungen $f : \mathbf{V} \to \mathbf{V}'$ zulassen. π kann dann aber nicht für alle Punkte aus $\mathcal{P}(\mathbf{V})$ definiert werden. In diesem Fall muß der Definitionsbereich eingeschränkt werden.

Lemma 7.57 *Sind $f, g : \mathbf{V} \to \mathbf{V}'$ zwei injektive semilineare Abbildungen mit $\mathcal{P}(f) = \mathcal{P}(g)$, dann gibt es $x \in K^\times$ mit*

$$g = xf.$$

Wieder gibt es einen *Fortsetzungssatz.*

Satz 7.58 *Sei $\mathcal{P}(\mathbf{V})$ ein n-dimensionaler projektiver Raum über einem Körper K und $\mathcal{P}(\mathbf{V}')$ über einem Körper K', der zu K isomorph ist. Ist weiters $\{Q_0, Q_1, \ldots, Q_n, Q\}$ ein Bezugssystem in $\mathcal{P}(\mathbf{V})$ und $\{Q'_0, Q'_1, \ldots, Q'_n, Q'\}$ eine Menge von Punkten in $\mathcal{P}(\mathbf{V}')$, von denen jeweils $n+1$ Punkte projektiv unabhängig sind. Dann gibt es eine eindeutig bestimmte kollineare Abbildung $\pi : \mathcal{P}(\mathbf{V}) \to \mathcal{P}(\mathbf{V}')$ mit*

$$\pi(Q_i) = Q'_i \quad (0 \leq i \leq n) \quad \textit{und} \quad \pi(Q) = Q'.$$

Eine weiters Eigenschaft kollinearer Abbildungen betrifft das Doppelverhältnis.

Satz 7.59 *Das Doppelverhältnis ist unter kollinearen Abbildungen $\pi : \mathcal{P}(\mathbf{V}) \to \mathcal{P}(\mathbf{V}')$ invariant, d.h. für beliebige vier Punkte $A, B, C, D \in \mathcal{P}(\mathbf{V})$ die auf einer Geraden liegen, gilt*

$$DV(\pi(A), \pi(B), \pi(C), \pi(D)) == \zeta(DV(A, B, C, D)).$$

7.2.6 Die Sätze von Desargues und Pappos

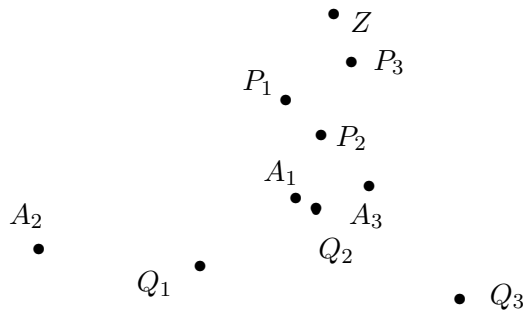
Drei Punkte P, Q, R eines projektiven Raums bilden ein **Dreieck**, wenn sie projektiv unabhängig ist, d.h. P, Q, R sind paarweise verschieden und liegen nicht auf einer Geraden.

Liegt andererseits eine Menge $\mathbf{T}$ von Punkten eines projektiven Raums auf einer Geraden, so bezeichnet man die Punkte in $\mathbf{T}$ als **kollinear**.

Satz 7.60 (Satz von Desargues) *Seien $Z, P_1, P_2, P_2, Q_1, Q_2, Q_3$ sieben verschieden Punkte einer projektiven Ebene $\mathcal{P}(\mathbf{V})$ (d.h.* $\operatorname{pdim} \mathbf{V} = 2$*), so daß P_1, P_2, P_3 und Q_1, Q_2, Q_3 zwei Dreiecke bilden und die Punkte Z, P_i, Q_i, $1 \leq i \leq 3$, jeweils kollinear sind. Dann sind die drei Punkte*

$$A_1 = (P_2 \vee P_3) \cap (Q_2 \vee Q_3), \quad A_2 = (P_1 \vee P_3) \cap (Q_1 \vee Q_3), \quad A_3 = (P_1 \vee P_2) \cap (Q_1 \vee Q_2)$$

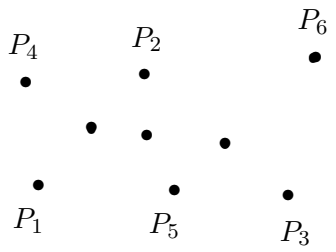
verschieden und kollinear.



Satz 7.61 (Satz von Pappos) *Seien g, h zwei Geraden einer projektiven Ebene $\mathcal{P}(\mathbf{V})$, P_1, P_3, P_5 drei verschiedene Punkte auf $g \setminus h$ und P_2, P_4, P_5 drei verschiedene Punkte auf $h \setminus g$. Dann sind die drei Punkte*

$$A = (P_1 \vee P_2) \cap (P_4 \vee P_5), \quad B = (P_2 \vee P_3) \cap (P_5 \vee P_6), \quad C = (P_3 \vee P_4) \cap (P_6 \vee P_1)$$

kollinear.



Die Sätze von Desargues und Pappos gelten daher in jeder projektiven Geometrie, die mit Hilfe eines Vektorraums (wie hier beschrieben) modelliert werden. Tatsächlich hätte man nur den Satz von Pappos beweisen müssen, da sich zeigen läßt, daß der Satz von Desargues aus dem Satz von Pappos abgeleitet werden kann. Für einen direkten Beweis des Satzes von Desargues können die Vorraussetzungen sogar etwas abgeschwächt werden. Man benötigt im Beweis nicht die Kommutativität der Multiplikation im Körper K. Betrachtet man also einen Schiefkörper (d.h. eine algebraische Strukture mit 2 Operationen, die genauso definiert wird wie ein Körper, mit der Ausnahme, daß die Kommutativität der Multiplikation nicht gefordert wird) und betrachtet *Vektorräume* über diesem Schiefkörper und entsprechende projektive Geometrien, so gilt dort der Satz von Desargues immer noch. Zum Beweis des Satzes von Pappos benötigt man aber die Kommutativität der Multiplikation in K. Es zeigt sich sogar, daß in einer projektiven Geometrie über einem echten Schiefkörper (der also kein Körper ist) der Satz von Pappos nicht allgemein gilt.

Interessant ist auch, daß der Satz von Desargues genau dann gilt, wenn die betrachtete projektive Geometrie isomoph ist zu einer Geometrie, die von einem Vektorraum über einem Schiefkörper oder Körper aufgebaut wird. Entsprechend gilt der Satz von Pappos genau dann, wenn die betrachtete projektive Geometrie isomoph zu einer Geometrie ist, die mit Hilfe eines Vektorraums über einem Körper modelliert wird. (Damit folgt natürlich der Satz von Desagues aus dem Satz von Pappos). Weiters gilt der Satz von Desargues in jedem affinen Raum der Dimension ≥ 3. Solche projektiven Räume sind daher im wesentichen Räume, die aus Vektorräumen aufgebaut werden.

Die projektiven Ebenen sind daher von speziellem Interesse. Tatsächlich gibt es projektive Ebenenen, auf denen der Satz von Desargues nicht gilt (nichtdesarguessche Ebene). Diese können natürlich nicht in einen dreidimensionalen Raum eingebettet werden.

Endlichen projektive Ebenen haben eine weitere interessante Eigenschaft. In ihnen folgt aus dem Satz von Desargues auch der Satz von Pappos. Dies liegt daran, daß es keine echten endlichen Schiefkörper gibt (Satz von Wedderburn).

7.3 Isomorphe Geometrien

7.3.1 Ismorphismen affiner und projektiver Geometrien

Definition 7.62 *Eine bijektive Abbildung $\alpha : \mathcal{A}(\mathbf{N}) \to \mathcal{A}(\mathbf{N}')$ zwischen zwei affinen Räumen heißt* **Isomorphismus** *, wenn für alle Nebenräume $\mathbf{S}, \mathbf{T} \in \mathcal{A}(\mathbf{N})$*

$$\mathbf{S} \subseteq \mathbf{T} \Longleftrightarrow \alpha(\mathbf{S}) \subseteq \alpha(\mathbf{T})$$

gilt.

Beispielsweise sine Affinitäten Isomorphismen.

Überraschenderweise müssen für den Nachweis der folgenden Eigenschaften keine weiteren Voraussetzungen zu verlangt werden.

Satz 7.63 *Sei $\alpha : \mathcal{A}(\mathbf{N}) \to \mathcal{A}(\mathbf{N}')$ ein Isomorphismus zwischen zwei affinen Räumen $\mathcal{A}(\mathbf{N})$ und $\mathcal{A}(\mathbf{N}')$. Dann gelten die folgenden Eigenschaften. ($\mathbf{S}_i, i \in I$, resp. $\mathbf{S}, \mathbf{T}$, bezeichnen Nebenräume in $\mathbf{N}$.)*

1. $\bigcap_{i\in I} \mathbf{S}_i \neq \emptyset \Longrightarrow \alpha\left(\bigcap_{i\in I} \mathbf{S}_i\right) = \bigcap_{i\in I} \alpha(\mathbf{S}_i)$.
2. $\alpha\left(\bigvee_{i\in I} \mathbf{S}_i\right) = \bigvee_{i\in I} \alpha(\mathbf{S}_i)$.
3. $\dim \mathbf{S} = \dim \alpha(\mathbf{S})$.
4. $\mathbf{S} \| \mathbf{T} \Longleftrightarrow \alpha(\mathbf{S}) \| \alpha(\mathbf{T})$.

Für projektive Räume kann ein ähnliches Konzept verwenden werden.

Definition 7.64 *Eine bijektive Abbildung $\tau : \mathcal{P}(\mathbf{V}) \to \mathcal{P}(\mathbf{V}')$ zwischen zwei projektiven Räumen heißt* **Isomorphismus** *, wenn für alle Unterräume $\mathbf{S}, \mathbf{T} \in \mathcal{P}(\mathbf{V})$*

$$\mathbf{S} \subseteq \mathbf{T} \Longleftrightarrow \alpha(\mathbf{S}) \subseteq \alpha(\mathbf{T})$$

gilt.

Entsprechend sind Kollinearitäten in diesem Sinn Isomorphismen.

Satz 7.65 *Sei $\tau : \mathcal{P}(\mathbf{V}) \to \mathcal{P}(\mathbf{V})$ ein Isomorphismus zwischen zwei affinen Räumen $\mathcal{P}(\mathbf{V})$ und $\mathcal{P}(\mathbf{V}')$. Dann gelten die folgenden Eigenschaften. ($\mathbf{S}_i, i \in I$, resp. $\mathbf{S}$, bezeichnen Unterräume in $\mathbf{V}$.)*

1. $\alpha\left(\bigcap_{i\in I} \mathbf{S}_i\right) = \bigcap_{i\in I} \alpha(\mathbf{S}_i)$.
2. $\alpha\left(\bigvee_{i\in I} \mathbf{S}_i\right) = \bigvee_{i\in I} \alpha(\mathbf{S}_i)$.
3. $\operatorname{pdim} \mathbf{S} = \operatorname{pdim} \alpha(\mathbf{S})$.

7.3.2 Charakterisierung isomorpher Räume

Der erste Satz zeigt, daß Isomorphismen zwischen affinen Räumen genau die Affinitäten sind.

Satz 7.66 *Sei $\alpha : \mathcal{A}(\mathbf{a}+\mathbf{U}) \to \mathcal{A}(\mathbf{a}'+\mathbf{U}')$ ein Isomorphismus zwischen zwei affinen Räumen $\mathcal{A}(\mathbf{a}+\mathbf{U}), \mathcal{A}(\mathbf{a}'+\mathbf{U}')$ (über den Körpern K, K') mit*

$$\dim \mathbf{U} = \dim \mathbf{U}' \geq 2.$$

Dann gibt es einen Köperisomorphismus $\zeta : K \to K'$, eine bijektive semilineare Abbiludung $f : \mathbf{U} \to \mathbf{U}'$ und Vektoren $\mathbf{t} \in \mathbf{a}+\mathbf{U}$, $\mathbf{t}' \in \mathbf{a}'+\mathbf{U}'$ mit

$$\alpha = \tau_{\mathbf{t}'} \circ f \circ \tau_{-\mathbf{t}}.$$

Für projektive Räume gilt eine analoge Eigenschaft.

Satz 7.67 *Sei $\pi : \mathcal{P}(\mathbf{V}) \to \mathcal{P}(\mathbf{V}')$ ein Isomorphismus zwischen zwei projektiven Räumen $\mathcal{P}(\mathbf{V}), \mathcal{P}(\mathbf{V}')$ (über den Körpern K, K') mit*

$$\dim \mathbf{V} = \dim \mathbf{V}' \geq 2.$$

Dann gibt es einen Köperisomorphismus $\zeta : K \to K'$ und eine bijektive semilineare Abbiludung $f : \mathbf{V} \to \mathbf{V}'$ mit

$$\pi = \mathcal{P}(f).$$

Index