Lineare Algebra für Informatiker
Universität Leipzig
WS 2008 / 09

Claus Diem

Inhaltsverzeichnis

Kapitel 1

Grundlagen

1.1 Vorbemerkungen

Ziel dieses Abschnitts ist, einige logische Grundlagen zu klären und die sogenannte *Aussagenlogik* zu motivieren. Die Aussagenlogik selbst wird in der Vorlesung "Logik" von Herrn Prof. Brewka behandelt.

Wir beginnen mit einigen mathematischen Aussagen $A, B, C, \ldots$. Zum Beispiel könnten dies diese Aussagen sein:

- 2 *ist gerade.*
- *Jede durch 4 teilbare natürliche Zahl ist durch 2 teilbar.*
- *Jede durch 2 teilbare natürlich Zahl ist durch 4 teilbar.*
- *Für je drei ganze Zahlen* x, y, z *mit* $x^3 + y^3 = z^3$ *gilt:* $x \cdot y \cdot z = 0$.
- *Für je drei ganze Zahlen* x, y, z *und jede natürliche Zahl* $n \geq 3$ *mit* $x^n + y^n = z^n$ *gilt:* $x \cdot y \cdot z = 0$.
- *Jede gerade natürliche Zahl* ≥ 4 *ist eine Summe von zwei Primzahlen.*

Jede dieser Aussagen ist entweder wahr oder falsch.

Wir können nun diese oder andere (wahre oder falsche) Aussagen verwenden, um komplexere Aussagen zu betrachten. Ein Beispiel ist

$$A \textit{ und } B \,,$$

was man mit

$$A \wedge B$$

abkürzt. Es ist klar, dass das Wort *und* bedeutet, dass diese Aussage genau dann wahr ist, wenn beide Aussagen A und B wahr sind. Dies kann man mittels einer *Wahrheitstabelle* ausdrücken.

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

Eine oft benutzte Form ist auch

$A \wedge B$	w	f
w	w	f
f	f	f

.

Andere "Operatoren", die wir auf Aussagen anwenden können, sind *nicht*, *oder* und *entweder oder*, mit den folgenden offensichtlichen Wahrheitstabellen. (Das Wort *oder* wird in der Mathematik immer als *und/oder* benutzt.)

$\neg A$	w	f
	f	w

$A \vee B$	w	f
w	w	w
f	w	f

$A \dot{\vee} B$	w	f
w	f	w
f	w	f

Implikationen

Wir betrachten nun die Aussage A *impliziert* B, abgekürzt $A \to B$. Wiederum wollen wir in Abhängigkeit davon, ob A wahr oder falsch ist, diese Aussage als wahr oder falsch betrachten. Wir legen fest, dass $A \to B$ genau dann *falsch* ist, wenn A wahr und B falsch ist. Wir haben also die folgende Wahrheitstabelle:

$A \to B$	w	f
w	w	f
f	w	w

Dies bedeutet also insbesondere, dass $A \to B$ automatisch wahr ist, wenn A falsch ist. Wir können nun die fünfte Beispiel-Aussage oben wie folgt umformulieren:

Für je drei ganze Zahlen x, y, z und jede natürliche Zahl $n \geq 3$ gilt:

$$x^n + y^n = z^n \longrightarrow x \cdot y \cdot z = 0$$

Die Aussage A *impliziert* B wird auch mit

Wenn A [gilt], dann [gilt] B

umschrieben. Z.B.

Für je drei ganze Zahlen x, y, z und jede natürliche Zahl $n \geq 3$ gilt:

$$\textit{Wenn } x^n + y^n = z^n, \textit{ dann gilt } x \cdot y \cdot z = 0 \ .$$

Ich möchte darauf hinweisen, dass die Bedeutung von A *impliziert* B bzw. *Wenn* A, *dann [gilt]* B nicht unbedingt dem allgemeinen Sprachgebrauch entspricht. Insbesondere könnte man geneigt sein, Aussagen der Form A *impliziert* B weder als wahr oder falsch sondern einfach als *unsinnig* zu betrachten, falls es keinen (offensichtlichen) engen Zusammenhang zwischen A und B gibt.

Mögliche Beispiele hierfür sind die folgenden beiden wahren Aussagen über ganze Zahlen:

$$3 > 4 \longrightarrow 100 < 0$$

$$3 = 2 + 1 \longrightarrow 3^2 + 4^2 = 5^2$$

Ich erwähne noch den Operator *genau dann wenn*, der durch die folgende Wahrheitstabelle definiert ist.

$A \leftrightarrow B$	w	f
w	w	f
f	f	w

Die Aussage $A \longleftrightarrow B$ liest man auch so: A *ist äquivalent zu* B.

Komplexere Zusammensetzungen

Die Operatoren *und*, *oder*, ... kann man selbstverständlich mehrfach anwenden. Man sollte Klammern setzen, um die Interpretation einer Aussage genau festzulegen.

Einige Bespiele:

Seien A, B, C drei sinnvolle (d.h. wahre oder falsche) mathematische Aussagen. Dann sind die folgenden beiden Aussagen äquivalent:

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

genauso:

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

(Diese beiden Äquivalenzen sind unter dem Namen *De Morgan'sche Gesetze* bekannt.)

Es sind auch äquivalent:

- $A \to B$
- $\neg(A \wedge \neg B)$
- $\neg A \vee B$
- $\neg B \to \neg A$

sowie:

- $A \wedge (B \vee C)$
- $(A \wedge B) \vee (A \wedge C)$

sowie:

- $A \vee (B \wedge C)$
- $(A \vee B) \wedge (A \vee C)$

und:

- $(A \to B) \to C$
- $(A \wedge B) \to C$

Dies kann man z.B. leicht mittels Wahrheitstabellen einsehen.

Beweisschemata

Nehmen wir an, wir wollen beweisen dass B wahr ist. Falls wir wissen, dass A wahr ist und $A \to B$ wahr ist, folgt dass B wahr ist. Dies kann man formal so beschreiben:

$$\frac{\begin{array}{l} A \\ A \to B \end{array}}{B}$$

Dies ist ein Beispiel eines *direkten Beweises*.

Wir nehmen nun an, dass wir wissen, dass A wahr ist und $\neg B \to \neg A$ gilt. In diesem Fall können wir auch schließen, dass B wahr ist.

$$\frac{\begin{array}{l} A \\ \neg B \to \neg A \end{array}}{B}$$

Dies ist ein Beispiel eines *Beweises durch Widerspruch*.

Noch einige Bemerkungen:

- Statt $A \wedge B$ schreibt man oft A, B.
- Statt des Implikationspfeils $\to$ wird oft ein Doppelpfeil $\Longrightarrow$ geschrieben.
- Ein Beweis der Form *Es gilt A, und es gilt $A \to B$, folglich gilt also auch B.* wird oft (insbesondere in Vorlesungen) in der Form *Es gilt A.* $\Longrightarrow B$ abgekürzt.
- Die Aussagen $A \leftrightarrow B$ und $B \leftrightarrow C$ (und folglich auch $A \leftrightarrow C$) werden oft zu $A \leftrightarrow B \leftrightarrow C$ zusammengefasst (und statt $\leftrightarrow$ wird meist $\Longleftrightarrow$ geschrieben).

Um Missverständnisse zu vermeiden sei noch angemerkt, dass mathematische Aussagen nicht nur wahr oder falsch sondern auch nicht interpretierbar bzw. sinnlos sein können.

Betrachten Sie z.B. die folgende Aussage:

x ist gerade.

Diese Aussage ist so nicht interpretierbar, weil nicht klar ist, was x ist. Wenn wir vorher x als die ganze Zahl 2 definieren (d.h. x und 2 bezeichnen nun dasselbe mathematische Objekt), dann ist die Aussage wahr. Wenn wir aber x als die rationale Zahl $3/2$ definieren, dann ist die Aussage wieder nicht interpretierbar (weil die Eigenschaft *gerade* nur für ganze Zahlen definiert ist).

1.2 Mengen

Was ist eine Menge? Wir stellen uns auf den folgenden Standpunkt Georg Cantors (1845-1918), der als Begründer der Mengenlehre gelten kann:

Eine Menge ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.

Aus der Schule kennen Sie die Mengen der *natürlichen Zahlen*, der *ganzen Zahlen*, der *rationalen Zahlen* oder der *reellen Zahlen*. Die Bezeichnungen sind $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$. Wir legen hier fest, dass 0 keine natürliche Zahl ist und definieren $\mathbb{N}_0$ als die Menge der ganzen Zahlen ≥ 0.

Die Frage, was genau die reellen Zahlen sind, behandeln wir später. Mengen werden oft durch (möglicherweise unvollständige) Aufzählung ihrer Elemente beschrieben. Beispiele sind

- $\emptyset = \{\}$, die leere Menge
- $\mathbb{N} = \{1, 2, 3, \ldots\}$
- $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$
- $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

Ein anderes Beispiel ist: Sei $a \in \mathbb{N}$, sei $X := \{1, 2, \ldots, a\}$. Nun ist X eine Menge mit a Elementen.

Die Schreibweise $X = \{x_1, \ldots, x_n\}$ bedeutet hingegen nicht, dass alle x_i verschieden sind; X hat *höchstens* n aber nicht notwendigerweise genau n Elemente.

Beispielsweise bedeutet die Aussage $X = \{a, b, c\}$, dass X höchstens drei Elemente enthält, welche mit a, b und c bezeichnet werden.

In der Informatik spricht man jedoch oft von *Mengen von Symbolen* oder von *Alphabeten*. In der Aussage "Sei Σ die Menge der Symbole $\{a, b, c\}$" ist dann nicht eine Menge mit höchstens drei Elementen gemeint, die mit a, b, c bezeichnet werden, sondern die Menge, die genau aus den drei (unterschiedlichen) *Symbolen* a, b, c besteht.

Diskussion Es stellt sich die Frage, inwieweit Symbole *Objekte unserer Anschauung oder unseres Denkens* sind. Vom mathematischen Gesichtspunkt wäre es genauer, unter einer *Menge von Symbolen* eine Menge zu verstehen, die zu jedem der angegeben Symbole ein Objekt enthält, welche alle unterschiedlich sind.

Teilmengen

Notation Sei A eine Teilmenge von X. Dann schreiben wir $A \subseteq X$. Wenn es sich um eine echte Teilmenge handelt (d.h. es gibt ein $x \in X$ so dass $x \notin A$), schreiben wir $A \subsetneq X$.

Bemerkung In Analogie zu den Relationen "kleiner-gleich" und "kleiner" für Zahlen wäre es folgerichtig, statt $A \subsetneq X$ einfach nur $Y \subset X$ zu schreiben. Allerdings schreiben viele Autoren $A \subset X$, wenn sie $A \subseteq X$ meinen. Ich vermeide die Bezeichnung $A \subset X$ ganz.

Definition Sei $A \subseteq X$ eine Teilmenge. Dann ist

$$A^c := \{x \in X \mid x \notin A\}$$

das *Komplement* von A in X.

Definition Seien A und B zwei Teilmengen von X. Dann ist X eine *disjunkte Vereinigung* von A und B, wenn gilt:

$$\text{Für alle } x \in X: \ x \in A \,\dot\vee\, x \in B \ .$$

In diesem Fall schreiben wir

$$X = A \,\dot\cup\, B \ .$$

Für $A \subseteq X$ gilt also offensichtlich $A \dot\cup A^c = X$. Die folgenden Aussagen für Teilmengen A und B von X folgen sofort aus den De Morgan'schen Gesetzen für Aussagen:

$$(A \cap B)^c = A^c \cup B^c \qquad (A \cup B)^c = A^c \cap B^c \; .$$

Auch die folgende Aussage ist offensichtlich:

$$A \subseteq B \longleftrightarrow B^c \subseteq A^c \; .$$

Quantoren

Mittels Mengen kann man mathematischen Aussagen wesentlich einfacher formulieren. Beispielsweise ist das Folgende eine Umformulierung der fünften Aussage zu Beginn.

Für alle $x, y, z \in \mathbb{Z}$, *für alle* $n \in \mathbb{N} : (n \geq 3 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0$.

Das kann man elegant mittels des All-Quantors $\forall$ aufschreiben:[1]

$$\forall x, y, z \in \mathbb{Z}, \forall n \in \mathbb{N} : (n \geq 3 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0 \; .$$

Diese Aussage ist (nach Auskunft der Experten) wahr. (Aber der Beweis ist sehr lang und schwierig.) Anderseits ist die folgende Aussage falsch:

$$\forall x, y, z \in \mathbb{Z}, \forall n \in \mathbb{N} : (n \geq 2 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0 \; .$$

(Ein Gegenbeispiel ist $x = 3, y = 4, z = 5, n = 2$.) Die Existenz so eines Gegenbeispiels man man mittels des Existenz-Quantors $\exists$ ausdrücken:

$$\exists x, y, z \in \mathbb{Z}, \exists n \in \mathbb{N} : \neg((n \geq 2 \wedge x^n + y^n = z^n) \longrightarrow x \cdot y \cdot z = 0)$$

bzw.

$$\exists x, y, z \in \mathbb{Z}, \exists n \in \mathbb{N} : (n \geq 2 \wedge x^n + y^n = z^n) \wedge x \cdot y \cdot z \neq 0 \; .$$

(Die Klammern kann man weglassen.)

[1] Den All-Quantor schreibe ich mit $\forall$, den Existenzquantor mit $\exists$. Eine ältere Schreibweise ist $\bigwedge$ für den All-Quantor und $\bigvee$ für den Existenzquantor. Da letzteres leicht mit dem All-Quantor $\forall$ verwechselt werden kann, bitte ich Sie, diese Schreibweise nicht zu benutzen.

Paradoxien

Wenn man den Mengenbegriff also generös auslegt, stößt leicht auf Paradoxien. Ein Beispiel ist die folgende sogenannte *Russelsche Antinomie*:

Sei $\mathcal{M}$ die Menge aller Mengen, die nicht ein Element von sich selbst sind. D.h.

$$\mathcal{M} := \{M \text{ Menge} \mid M \notin M\}$$

Für jede Menge M gilt also

$$M \in \mathcal{M} \longleftrightarrow M \notin M \ .$$

Somit gilt insbesondere $\mathcal{M} \in \mathcal{M} \longleftrightarrow \mathcal{M} \notin \mathcal{M}$. Das ist offensichtlich absurd.

Aufgrund solcher Paradoxien (das Fachwort ist *Antinomien*), sollte man aufpassen, wenn man Mengen bildet. Als Regeln sollte man sich merken:

- Wenn X eine Menge ist und E eine Eigenschaft, die für Elemente aus X definiert ist (und jeweils wahr oder falsch sein kann), dann gibt es die Teilmenge
$$\{x \in X \mid E \text{ trifft auf } x \text{ zu } \} \ .$$
- Zu jeder Menge X gibt es die *Potenzmenge* $\mathcal{P}(X)$. Die Elemente von $\mathcal{P}(X)$ sind genau die Teilmengen von X.
- Man kann Vereinigungen von Mengen bilden.

Tupel

Seien X und Y Mengen. Dann besteht die Menge $X \times Y$ aus *geordneten Paaren* (x, y) von Elementen $x \in X, y \in Y$. Solche geordneten Paare heißen auch *Tupel* (oder *Zweiertupel*).

Die Menge $X \times Y$ heißt *kartesisches Produkt* von X und Y. Etwas allgemeiner erhält man zu Mengen $X_1, \ldots, X_n$ das kartesische Produkt $X_1 \times \cdots \times X_n$, das aus den so genannten n-Tupeln $(x_1, \ldots, x_n)$ mit $x_i \in X_i$ besteht. Wenn $X_1 = X_2 = \cdots = X_n$, schreibt man auch X^n für das kartesische Produkt.

Das Prinzip der vollständigen Induktion

Das *Prinzip der vollständigen Induktion* ist eine Beweismethode für Aussagen, die für alle natürlichen Zahlen gelten. Es handelt sich also um Aussagen der Form

$$\forall n \in \mathbb{N} : A(n) \ ,$$

wobei für $n \in \mathbb{N}$ $A(n)$ eine Aussage über die Zahl n ist.

Aussagen dieser Form kann man wie folgt beweisen:

1. Man beweist, dass $A(1)$ gilt.

2. Man beweist für alle $n \in \mathbb{N}$, dass die Implikation $A(n) \longrightarrow A(n+1)$ gilt.

Wenn nun $n \in \mathbb{N}$ beliebig ist, haben wir die folgenden (bewiesenen) Aussagen:

$$A(1), A(1) \longrightarrow A(2), A(2) \longrightarrow A(3), \ldots, A(n-1) \longrightarrow A(n) \; .$$

Hieraus folgt $A(n)$.

Es gibt zwei oft benutzte Varianten der vollständigen Induktion:

- Man beginnt die Induktion nicht bei 1 sondern bei $n_0 \in \mathbb{Z}$ und zeigt $\forall n \in \mathbb{Z}, n \geq n_0 : A(n)$.

- Man setzt im Induktionsschritt nicht nur $A(n)$ sondern alle vorherigen Aussagen voraus. D.h. man zeigt $A(n_0) \wedge A(n_0+1) \wedge \cdots \wedge A(n) \longrightarrow A(n+1)$.

Wir geben zwei Beispiele für Beweise mittels "vollständiger Induktion".

Beispiel 1.1 Wir wollen beweisen:

$$\forall n \in \mathbb{N} \; : \; 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \; .$$

Die Aussage $A(n)$ ist also $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.

Beweis von $A(1)$

Es ist $1 = \frac{1 \cdot 2}{2}$.

Beweis der Implikation $A(n) \longrightarrow A(n+1)$

Es gelte $A(n)$, also $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$.

Dann ist $1 + + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+2)(n+1)}{2}$. □

Beispiel 1.2 Wir wollen beweisen:

Für alle $n \in \mathbb{N}, m \in \mathbb{N}_0$ gibt es eindeutig bestimmte $p \in \mathbb{N}_0$ und $r \in \{0, \ldots, n-1\}$ mit

$$m = pn + r \; .$$

("Division mir Rest")

Hierzu *fixieren* wir $n \in \mathbb{N}$ und betrachten die folgende Aussage:[2]

$$\forall m \in \mathbb{N}_0 \,:\, \exists!\, (p, r) \in \mathbb{N}_0 \times \{0, \ldots, n-1\} : m = pn + r$$

Wir zeigen dies nun mittels vollständiger Induktion nach m.

Beweis von $A(0)$
Es ist $0 = 0 \cdot n + 0$, und dies ist offensichtlich die einzige Möglichkeit, 0 in der Form $pn + r$ mit $p \in \mathbb{N}_0, r \in \{0, \ldots, n-1\}$ zu schreiben.

Beweis der Implikation $A(m) \longrightarrow A(m+1)$
Wir setzen voraus: Es gibt eindeutig bestimmte $p \in \mathbb{N}_0$ und $r \in \{0, \ldots, n-1\}$ mit $m = pn + r$.

Zuerst zur *Existenz* der Darstellung von $m + 1$. Seien $p_0 \in \mathbb{N}_0, r_0 \in \{0, \ldots, n-1\}$ mit

$$m = p_0 n + r_0 \,.$$

Dann ist also $m+1 = p_0 n + r + 1$. Es gibt nun zwei Fälle: Wenn $r < n-1$, dann ist $m + 1 = p_0 n + (r + 1)$ eine Darstellung wie gewünscht. Wenn andererseits $r = n - 1$, dann ist $m + 1 = (p_0 + 1)n + 0$ eine Darstellung wie gewünscht.

Nun zur *Eindeutigkeit*. Seien $m + 1 = p_1 n + r_1$ und $m + 1 = p_2 n + r_2$ zwei Darstellungen mit $p_1, p_2 \in \mathbb{N}_0, r_1, r_2 \in \{0, \ldots, n-1\}$.

Wir machen eine Fallunterscheidung in vier Fälle.

$r_1 \geq 1, r_2 \geq 1$
In diesem Fall ist $m = p_1 n + (r_1 - 1)$ und $m = p_2 n + (r_2 - 1)$ mit $r_1 - 1, r_2 - 1 \in \{0, \ldots, n-1\}$. Damit ist nach der Eindeutigkeit der Darstellung von m $p_1 = p_2$ und $r_1 = r_2$.

$r_1 = 0, r_2 \geq 1$
In diesem Fall ist $m = (p_1 - 1)n + (n - 1)$ und $m = p_2 n + (r_2 - 1)$ mit $p_1 - 1, p_2 \in \mathbb{N}_0, r_1 - 1 \in \{0, \ldots, n-1\}$. Nach der Eindeutigkeit der Darstellung von m kann dieser Fall nicht auftreten.

$r_1 \geq 1, r_2 = 0$
Analog zum zweiten Fall kann dieser Fall nicht auftreten.

$r_1 = 0, r_2 = 0$
In diesem Fall ist $m = (p_1 - 1)n + (n - 1)$ und $m = (p_2 - 1)n + (n - 1)$ mit mit $p_1 - 1, p_2 - 1 \in \mathbb{N}_0$. Damit ist nach der Eindeutigkeit der Darstellung von m $p_1 = p_2$. □

[2] Das Ausrufezeichen hinter "$\exists$" deutet an, dass es genau ein Element mit der angegeben Eigenschaft gibt.

1.3 Abbildungen

Seien X, Y Mengen. Intuitiv ist ein *Abbildung* von X nach Y ist eine Vorschrift, die jedem Element aus X in eindeutiger Weise ein Element aus Y zuordnet. Wenn f eine Abbildung von X nach Y ist, schreibt man $f : X \longrightarrow Y$, X heißt dann *Definitionsbereich* und Y heißt *Bildbereich* oder *Wertebereich*.

Zwei Abbildungen $f : X \longrightarrow Y, g : X \longrightarrow Y$ sind genau dann gleich, wenn für alle $x \in X$ $f(x) = g(x)$ ist.

Ein Beispiel ist

$$f : \mathbb{Z} \longrightarrow \mathbb{Z} \ , \ x \mapsto 2x \ .$$

Hier gilt also $f(x) = 2x$ für alle $x \in \mathbb{Z}$.

Wenn allgemein $f : X \longrightarrow Y$ eine Abbildung ist, schreibt man $f(x)$ für den *Wert* von f an x, d.h. für dasjenige Element aus Y, welches x zugeordnet ist, bzw. auf welches x abgebildet wird.

Zwei Abbildungen $f : X \longrightarrow Y, g : X \longrightarrow Y$ sind genau dann gleich, wenn für alle $x \in X$ $f(x) = g(x)$ ist.

Aus der Schule kennen Sie den Begriff der *Funktion*. "Funktion" und "Abbildung" kann man synonym benutzen, allerdings spricht man in der Regel eher dann von Funktionen, wenn der Wertebereich aus Zahlen besteht.

Die Abbildungen von X nach Y bilden auch eine Menge. Wir definieren:

Definition Die Menge der Abbildungen von X nach Y wird mit $\mathrm{Abb}(X, Y)$ bezeichnet.

Einige Beispiele:

Beispiel 1.3 Sei $X = \{x\}$. Dann besteht $\mathrm{Abb}(X, Y)$ genau aus den Abbildungen $x \mapsto a$ mit $a \in Y$.

Beispiel 1.4 Sei andererseits $Y = \{y\}$. Dann besteht $\mathrm{Abb}(X, Y)$ aus genau einem Element.

Eine kleine Spitzfindigkeit ist das folgende Beispiel:

Beispiel 1.5 Sei Y eine beliebige Menge. Natürlich können wir dann *jedem* Element der *leeren Menge* ein Element von Y zuordnen. Somit besteht $\mathrm{Abb}(\emptyset, Y)$ aus genau einem Element. Andererseits ist $\mathrm{Abb}(X, \emptyset)$ leer, wenn $X \neq \emptyset$.

Definition Die *identische Abbildung* id_X auf einer Menge X ist durch $x \mapsto x$ gegeben.

Definition Eine Abbildung $f : X \longrightarrow Y$ heißt

- *injektiv*, wenn für alle $x, x' \in X$ gilt: $f(x) = f(x') \longrightarrow x = x'$
- *surjektiv*, wenn für alle $y \in Y$ gilt: $\exists x \in X : f(x) = y$
- *bijektiv*, wenn sie injektiv und surjektiv ist.

Notation Wenn $f : X \longrightarrow Y$ injektiv ist, schreibt man auch $X \hookrightarrow Y$. Wenn $f : X \longrightarrow Y$ surjektiv ist, schreibt man auch $X \twoheadrightarrow Z$.

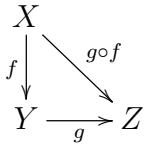
Einige offensichtliche Bemerkungen:

- f ist genau dann injektiv, wenn für alle $x, x' \in X$ gilt: $x \neq x' \longrightarrow f(x) \neq f(x')$.
- f ist genau dann bijektiv, wenn es je jedem $y \in Y$ genau ein $x \in X$ mit $f(x) = y$ gibt. Dies kann man auch so beschreiben: $\forall y \in Y \exists! x \in X : f(x) = y$.
- Wenn f bijektiv ist, dann kann man wie folgt eine Abbildung $g : Y \longrightarrow X$ definieren: Jedem $y \in Y$ wird das eindeutig bestimmte $x \in X$ mit $f(x) = y$ zugeordnet. Diese Abbildung erfüllt $g \circ f = \mathrm{id}_X, f \circ g = \mathrm{id}_Y$. Sie heißt die *Umkehrabbildung* zu f und wird mit $f^{-1} : Y \longrightarrow X$ bezeichnet.

Ich erinnere noch daran, dass man Abbildungen verknüpfen kann: Gegeben zwei Abbildungen $f : X \longrightarrow Y, g : Y \longrightarrow Z$, hat man die Abbildung

$$g \circ f : X \longrightarrow Z,\ x \mapsto g(f(x)) \ .$$

Diese Definition kann man anhand eines *kommutativen Diagramms* veranschaulichen:



Allgemein ist ein *kommutatives Diagramm* ein Diagramm von Mengen und Abbildungen so dass gilt: Wenn immer man von einer Menge zu einer anderen "auf mehreren Wegen gehen kann", erhält man dieselbe Abbildung.

Beispiel 1.6 Seien X, Y, Z, W Mengen, und sei $f : X \longrightarrow Y, g : Y \longrightarrow W, h : X \longrightarrow Z, k : Z \longrightarrow W$. Die Aussage, dass das Diagramm

$$\begin{array}{ccc} X & \xrightarrow{h} & Z \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle k} \\ Y & \xrightarrow[g]{} & W \end{array}$$

kommutiert, heißt, dass $g \circ f = k \circ h : X \longrightarrow W$.

Definition Sei $U \subseteq X$ eine Teilmenge. Durch "Einschränkung" erhalten wir dann eine Abbildung

$$f|_U : U \longrightarrow Y\ ,\ x \mapsto f(x)\ .$$

Die Menge

$$f(U) := \{y \in Y \mid \exists x \in U : f(x) = y\}$$

heißt *Bild* von U unter f. Die Menge $f(X)$ wird auch mit $\mathrm{Bild}(f)$ bezeichnet. Es ist also $f(U) = \mathrm{Bild}(f|_U)$.

Eine ungenauere aber oft vorkommende Beschreibung ist

$$f(U) = \{f(x) \mid x \in U\}\ .$$

Bemerkung Beachten Sie den Unterschied zwischen dem *Bildbereich* (=Wertebereich) von f und dem *Bild* von f!

Definition Sei nun $V \subseteq Y$ eine Teilmenge. Die Menge

$$f^{-1}(V) := \{x \in X \mid f(x) \in V\}$$

heißt die *Urbildmenge* von V unter f.

Familien

Wir diskutieren noch eine weitere *Sichtweise* auf Abbildungen.

Seien zwei beliebige Mengen X und I gegeben.

Für jedes $i \in I$ sei genau ein Element aus X gegeben. Wir erhalten somit eine *Familie* von Elementen von X, die wir z.B. mit $(x_i)_{i \in I}$ bezeichnen können. Die Menge I heißt hierbei auch *Indexmenge*. Wenn $I = \{1, \ldots, n\}$, erhalten wir so die n-Tupel von Elementen aus X, d.h. die Elemente $(x_1, \ldots, x_n) \in X^n$.

So eine Familie $(x_i)_{i\in X}$ ist nichts weiter als eine Abbildung (d.h. Zuordnung): Jedem $i \in I$ wird genau das Element x_i zugeordnet, oder formaler: Die Famlie $(x_i)_{i\in I}$ ist per Definition identisch mit der Abbildung $I \longrightarrow X, i \mapsto x_i$.

Es gibt also keinen inhaltlichen Unterschied zwischen einer Familie und einer Abbildung. Es ist eine Frage der Sichtweise bzw. der Notation.

Beispiel 1.7 Man kann die Menge X^n als die Menge der Abbildungen $\{1, \ldots, n\} \longrightarrow X$ *definieren*. In diesem Sinne definieren wir für jede beliebige Menge I:

$$X^I := \mathrm{Abb}(I, X) \; .$$

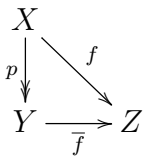
Beispiel 1.8 Eine *Folge* ist eine Familie von reellen Zahlen über der Indexmenge $\mathbb{N}$, d.h. ein Element $(a_n)_{n\in\mathbb{N}}$ mit $a_n \in \mathbb{R}$. Mit anderen Worten: Eine Folge ist per Definition eine Abbidlung $\mathbb{N} \longrightarrow \mathbb{R}$. Die Menge der Folgen ist also $\mathbb{R}^{\mathbb{N}}$.

Beispiel 1.9 Seien $X_1, \ldots, X_n$ Mengen. Oben haben wir von Tupeln $(x_1, \ldots, x_n)$ (mit $x_i \in X_i$) sowie vom kartesischen Produkt $X_1 \times \cdots \times X_n$ gesprochen. Dieses Produkt kann man mittels Familien (d.h. mittels Abbidlungen) so definieren: Oben haben wir bereits X^n für eine Menge X definiert. Wir definieren nun $X_1 \times \cdots \times X_n$ als die Menge der Familien $(x_1, \ldots, x_n) \in X_1 \cup \cdots \cup X_n$ mit $x_i \in X_i$.

Wohldefiniertheit

Es kommt häufig vor, dass die folgende Situation gegeben ist:

Gegeben sind drei Mengen X, Y, Z, eine *surjektive* Abbildung $p : X \twoheadrightarrow Y$ sowie eine weitere Abbildung $f : X \longrightarrow Z$. Man fragt sich nun, ob es eine Abbildung $\overline{f} : Y \longrightarrow Z$ mit $\overline{f} \circ p = f : X \longrightarrow Z$ gibt. Mit anderen Worten: Wir wollen, dass das Diagramm



kommutiert. Wiederum mit anderen Worten: Wir wollen, dass für alle $x \in X$ $\overline{f}(p(x)) = f(x)$ gilt. Da p surjektiv ist, ist $\overline{f}$ hierdurch – wenn es existiert – eindeutig bestimmt. Wir erhalten auch sofort eine *notwendige Bedingung* damit $\overline{f}$ existieren kann: Es muss gelten:

$$\forall x, x' \in X : p(x) = p(x') \longrightarrow f(x) = f(x') \tag{1.1}$$

Denn, wenn $\overline{f}$ existiert und $x, x' \in X$ mit $p(x) = p(x')$ gegeben sind, dann ist $f(x) = \overline{f}(p(x)) = \overline{f}(p(x')) = f(x')$. Wenn andererseits (1.1) gilt, dann können wir mittels

$$y \mapsto f(x) \text{ für } \textit{irgendein } x \in X \text{ mit } p(x) = y$$

eine Abbildung $\overline{f} : Y \longrightarrow Z$ definieren. Der entscheidende Punkt ist, dass $\overline{f}(y)$ nun nicht von der Wahl von x abhängt, man also eine *eindeutige Zuordnung*, eben eine Abbildung erhält. Die Unabhängigkeit von der Wahl von x nennt man *Wohldefiniertheit*. Wir fassen dies zusammen:

Aussage 1.10 *Seien X, Y, Z drei Mengen, $p : X \longrightarrow Y$ eine surjektive Abbildung, $f : X \longrightarrow Z$ irgendeine Abbildung. Dann gibt es höchstens eine Abbildung $\overline{f} : Y \longrightarrow Z$ mit $\overline{f} \circ p = f$. So eine Abbildung $\overline{f}$ gibt es genau dann, wenn die Bedingung (1.1) erfüllt ist.*

Aussage 1.11 *Seien die Notationen wie oben. Dann existiert $\overline{f}$ genau dann und ist injektiv, wenn gilt:*

$$\forall x, x' \in X : p(x) = p(x') \longleftrightarrow f(x) = f(x') \tag{1.2}$$

Der Beweis ist leicht.

Kardinalität

Definition Eine Menge X heißt *endlich*, wenn es eine natürlich Zahl n und eine Bijektion $\{1, \ldots, n\} \longrightarrow X$ gibt, andernfalls *unendlich*. Eine Menge heißt *abzählbar*, wenn es eine Surjektion $\mathbb{N} \longrightarrow X$ gibt.

Bemerkung Beachten Sie, dass eine "abzählbare Menge" endlich sein kann!

Man kann (mittels vollständiger Induktion) zeigen:

Lemma 1.12 *Sei X eine endliche Menge, und seien n und m zwei natürliche Zahlen so dass es Bijektionen $\{1, \ldots, n\} \longrightarrow X$ und $\{1, \ldots, m\} \longrightarrow X$ gibt. Dann ist $n = m$.*

Damit können wir definieren:

Definition Wenn es eine Bijektion $\{1, \ldots, n\} \longrightarrow X$ gibt, dann heißt n die *Kardinalität* von X und wird mit $\#X$ oder $|X|$ bezeichnet. Wenn X unendlich ist, so schreibt man $\#X = |X| = \infty$.

1.4 Relationen

Sei X eine Menge. Eine *Relation* auf X ist intuitiv eine Eigenschaft, die für je zwei Elemente aus X gelten kann oder nicht. Wenn die Eigenschaft für $(x, y) \in X \times X$ gilt, dann sagt man auch, dass x und y in *Relation zueinander stehen* (bez. der gegebenen Eigenschaft).

Einfache Beispiele für die ganzen Zahlen sind die Beziehungen $<, \leq, >, \geq$ und natürlich auch $=$.

Ein anderes Beispiel wiederum für $\mathbb{Z}$ ist: "$x - y$ ist gerade" (wobei $x, y \in \mathbb{Z}$). Hier stehen also je zwei gerade Zahlen zueinander in Relation und je zwei ungerade Zahlen stehen zueinander in Relation. Hingegen stehen jeweils eine gerade und einer ungerade Zahl (oder umgekehrt) nicht zueinander in Relation.

Eine formale Definition einer Relation erhält man, indem man von der oben erwähnten Eigenschaft zu der Teilmenge von $X \times X$ übergeht, die durch die Eigenschaft definiert wird.

Definition Eine Relation auf einer Menge X ist eine Teilmenge von $X \times X$.

Sei nun R eine Relation, und seien $x, y \in X$. Dann sagen wir, dass x (bezüglich R) *in Relation zu* y *steht*, wenn $(x, y) \in R$. In diesem Fall schreiben wir $x \sim_R y$.[3] (Andere Autoren schreiben auch xRy.)

Andernfalls schreiben wir $x \nsim_R y$.

Oft wird $\sim_R$ durch ein anderes Symbol wie z.B. die obigen $(<, \leq, \ldots)$ oder auch einfach $\sim$ ersetzt.

Beispiel 1.13 Die (übliche) Relation $\geq$ auf $\mathbb{Z}$ ist durch $x \geq y :\Longleftrightarrow x - y \in \mathbb{N}_0$ definiert. Mit anderen Worten: Sie ist durch die Menge

$$\{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid x - y \in \mathbb{N}_0\}$$

gegeben.

Äquivalenzrelationen

Sei X eine Menge und R eine Relation auf X.

[3] Immer wenn ich in einer *Definition* schreibe *"Wenn ..., dann sagen wir ..."*, meine ich *"Genau dann wenn ..."*.

Definition Die Relation R heißt *Äquivalenzrelation*, falls gilt:

$$\begin{aligned}
&(R) \quad \forall x \in X : x \sim_R x \\
&(S) \quad \forall x, y \in X : x \sim_R y \longleftrightarrow y \sim_R x \\
&(T) \quad \forall x, y, z \in X : x \sim_R y \wedge y \sim_R z \longrightarrow x \sim_R z
\end{aligned}$$

Die Bedingungen (R), (S), (T) heißen *Reflexivität*, resp. *Symmetrie*, resp. *Transitivität*.

Wenn R eine Äquivalenzrelation ist, schreibt man meist $x \sim y$ anstatt $x \sim_R y$. Man sagt dann auch, dass $\sim$ eine Äquivalenzrelation ist. Aufgrund der Symmetrie sagt man auch, dass x und y *zueinander in Relation stehen*, wenn $x \sim_R y$.

Beispiel 1.14 Zu Beginn dieses Abschnitts haben wir die folgende Relation auf $\mathbb{Z}$ erwähnt:

$$x \sim y :\Longleftrightarrow x - y \text{ ist gerade} \quad .$$

Die Eigenschaften (R), (S), (T) sind offensichtlich, es handelt sich also um eine Äquivalenzrelation.

Man kann dieses Beispiel leicht wie folgt verallgemeinern: Sei n eine natürliche Zahl > 1, und seien $x, y \in \mathbb{Z}$. Dann heißen x und y *kongruent* zueinander modulo n, wenn $x - y$ durch n teilbar ist (d.h. falls ein $a \in \mathbb{Z}$ mit $y = x + an$ existiert). Wenn x und y (modulo n) kongruent zueinander sind, schreibt man

$$x \equiv y \mod n \; .$$

Diese "Kongruenz modulo n" ist offensichtlich auch eine Äquivalenzrelation.

Beispiel 1.15 Sei X eine beliebige Menge. Dann ist "=" eine Äquivalenzrelation.

Sei nun eine Äquivalenzrelation $\sim$ gegeben.

Definition Zu $x \in X$ definieren wir

$$[x]_\sim := \{y \in X \mid x \sim y\} \; ,$$

die *Äquivalenzklasse* zu x. Wir schreiben auch $[x]$, wenn die Relation offensichtlich ist.

Aussage 1.16 *Seien $x, y \in X$ und (und somit $[x]_\sim$ und $[y]_\sim$ zwei Äquivalenzklassen). Dann gilt: Entweder ist $[x]_\sim = [y]_\sim$ oder es ist $[x]_\sim \cap [y]_\sim = \emptyset$.*

Beweis. Es sei zunächst $[x]_\sim = [y]_\sim$. Dann ist $x \in [x]_\sim = [y]_\sim$ und somit $[x]_\sim \cap [y]_\sim \neq \emptyset$.

Wir müssen nun zeigen:

$$[x]_\sim \cap [y]_\sim \neq \emptyset \longrightarrow [x]_\sim = [y]_\sim \,. \tag{1.3}$$

Bevor wir diese Implikation beweisen, legen wir noch eine Notation fest:

Wenn $x_1, \ldots, x_n \in X$ gegeben sind und $x_1 \sim x_2, x_2 \sim x_3, \ldots, x_{n-1} \sim x_n$, dann stehen nach der Transitivität alle x_i zueinander in Relation. Dies deuten wir mit $x_1 \sim x_2 \sim \cdots \sim x_n$ an.

Wir zeigen nun die Implikation. Sei also $[x]_\sim \cap [y]_\sim \neq \emptyset$. Dann gibt es ein $z \in [x]_\sim \cap [y]_\sim$. Es gilt also $x \sim z$, $y \sim z$. Nach der Symmetrie gilt $z \sim y$. Somit gilt $x \sim z \sim y$.

Sei nun $x' \in [x]_\sim$ beliebig. Dann ist

$$y \sim x \sim x' \,.$$

Somit ist also $x' \in [y]_\sim$.

Soeben haben wir gezeigt, dass $[x]_\sim \subseteq [y]_\sim$. Analog zeigt man $[y]_\sim \subseteq [x]_\sim$. Damit sind die beiden Mengen gleich. □

Alle Äquivalenzklassen sind Teilmengen von X, und somit sind sie *Elemente* der Potenzmenge $\mathcal{P}(X)$. Wir können somit die *Menge der Äquivalenzklassen* (in X bez. $\sim$) betrachten.

Notation Die Menge der Äquivalenzklassen bezeichnen wir mit $X_{/\sim}$.

Wir haben also

$$X_{/\sim} = \{M \in \mathcal{P}(X) \mid \exists x \in X : M = [x]_\sim\}$$

oder (etwas ungenauer)

$$X_{/\sim} = \{[x]_\sim \mid x \in X\} \,.$$

Beispiel 1.17 Wir kommen auf die in Beispiel 1.14 diskutierte Äquivalenzrelation "Kongruenz modulo n" zurück. Für $x \in \mathbb{Z}$ bezeichnen wir die Äquivalenzklasse "modulo n" mit $[x]_n$. Es ist also

$$[x]_n = \{x + an | a \in \mathbb{Z}\} \,.$$

Wir haben n Äquivalenzklassen, nämlich $[0]_n, [1]_n, \ldots, [n-1]_n$.

Im letzten Abschnitt haben wir den Konzept der Wohldefiniertheit einer Abbildung diskutiert. Wir wenden dies nun auf Äquivalenzrelationen an.

Wir nehmen an, dass wir eine Abbildung $f : X \longrightarrow Y$ gegeben haben. Wir fragen uns, ob es eine Abbildung $\overline{f} : X_{/\sim} \longrightarrow Y$ mit $\overline{f}([x]_\sim) = f(x)$ für alle $x \in X$ gibt. Hierzu betrachten wir die surjektive Abbildung $p : X \longrightarrow X_{/\sim}$, $x \mapsto [x]_\sim$. Aussage 1.10 liefert nun:

Aussage 1.18 *Sei $f : X \longrightarrow Y$ gegeben. Dann gibt es genau dann eine Abbildung $\overline{f} : X_{/\sim} \longrightarrow Y$ mit $\overline{f}([x]_\sim) = f(x)$ für alle $x \in X$, wenn gilt:*

$$\forall x, x' \in X : x \sim x' \longrightarrow f(x) = f(x')$$

(und in diesem Fall ist $\overline{f}$ eindeutig bestimmt).

Definition Eine *Partition* einer Menge X ist eine Teilmenge $\mathcal{M}$ von $\mathcal{P}(X)$ (d.h. eine Menge von Teilmengen von X) mit der folgenden Eigenschaften:

- Alle Mengen in $\mathcal{M}$ sind nicht-leer.
- Für alle $x \in X$ existiert genau eine Menge $M \in \mathcal{M}$ mit $x \in M$.

Wenn $\mathcal{M}$ eine Menge von Teilmengen von X mit der zweiten obigen Eigenschaft ist, sagt man auch, dass X die *disjunkte Vereinigung* der Mengen in $\mathcal{M}$ ist. (Dies ist eine Verallgemeinerung der entsprechenden Definition in Abschnitt 1.2.) Hierfür schreibt man:

$$X = \dot{\bigcup}_{M \in \mathcal{M}} M$$

Aus der Reflexivität und Lemma 1.16 folgt:

Aussage 1.19 *Sei $\sim$ eine Äquivalenzrelation auf X. Dann bildet die Menge der Äquivalenzklassen $X_{/\sim}$ eine Partition von X.*

Umgekehrt kann man auch jeder Partition eine Äquivalenzrelation zuordnen. Sei hierzu $\mathcal{M}$ eine Partition von X. Zuerst ordnen wir jedem x die eindeutig bestimmte Menge $M \in \mathcal{M}$ mit $x \in M$ zu. Diese bezeichnen wir mit $[x]_\mathcal{M}$. Dann definieren wir wie folgt eine Relation:

$$x \sim y :\Longleftrightarrow [x]_\mathcal{M} = [y]_\mathcal{M}$$

Die Eigenschaften (R), (S), (T) folgen sofort, und $[x]_\mathcal{M}$ ist nun die Äquivalenzklasse zu x.

Definition Sei $\mathcal{M}$ eine Partition von X. Ein *Repräsentantensystem* zu $\mathcal{M}$ ist eine Teilmenge A von X mit der folgenden Eigenschaft: Für alle $M \in \mathcal{M}$ existiert genau ein $a \in A$ mit $[a]_{\mathcal{M}} = M$.

Ein Repräsentantensystem einer Äquivalenzrelation ist der Definition ein Repräsentantensystem der zugehörigen Partition. D.h. ein Repräsentantensystem zu einer Äquivalenzrelation $\sim$ auf X ist eine Teilmenge A von X mit:

$$\forall x \in X \, \exists! \, a \in A : x \sim a \; .$$

Beispiel 1.20 Wir betrachten wieder die Relation "Kongruenz modulo n" auf $\mathbb{Z}$. Ein Repräsentantensystem dieser Relation ist z.B. die Menge $\{0, 1, \ldots, n-1\}$.

Das Auswahlaxiom (Diskussion)

Betrachten wir nun die folgende Aussage.

Jede Partition einer Menge hat ein Repräsentantensystem.

Ein Beweis dieser Aussage ist scheinbar sehr leicht: Gegeben eine Partition $\mathcal{M}$ auf X wählt man aus jeder Menge $M \in \mathcal{M}$ ein (beliebiges) Element $m \in M$ aus. Nehmen wir an, wir haben so eine Auswahl getroffen. Dann haben wir also eine Zuordnung (d.h. Abbildung) $f : \mathcal{M} \longrightarrow X$ mit $f(M) \in M$ für alle $M \in \mathcal{M}$.

Aber gibt es so eine Abbildung immer? Hier fragen wir nach einem Argument, dass die Existenz einer solchen Abbildung auf "offensichtlichere" Aussagen zurückführt. Eine überraschende Erkenntnis der Mengenlehre ist, dass dies in einer gewissen Hinsicht nicht möglich ist. Man sollte die obige Aussage als ein zusätzliches Axiom der Mengenlehre akzeptieren. Es ist das sogenannte *Auswahlaxiom*.

Es sei noch bemerkt, dass die Tatsache, dass das Auswahlaxiom nicht so naheliegend wie die anderen Axiome der Mengenlehre ist, auch Kritiker auf den Plan ruft, die die Verwendung des Auswahlaxioms ablehnen. Diese Kritiker nehmen jedoch eine Außenseiterrolle in der Gemeinschaft der Mathematiker ein.

Ordnungsrelationen

Sei wiederum X eine Menge und R eine Relation auf X.

Definition Die Relation R heißt *Ordnungsrelation*, falls gilt:

$$(R) \quad \forall x \in X : x \sim_R x$$
$$(A) \quad \forall x, y \in X : x \sim_R y \wedge y \sim_R x \longrightarrow x = y$$
$$(T) \quad \forall x, y, z \in X : x \sim_R y \wedge y \sim_R z \longrightarrow x \sim_R z$$

Eine Ordnungsrelation heißt *vollständig* oder eine *lineare Relation*, falls gilt:

$$(V) \quad \forall x, y \in X : x \sim_R y \vee y \sim_R x \ .$$

Die Bedingungen (R) und (T) sind die schon bekannte Reflexivität und Transitivität, die Bedingung (A) heißt *Antisymmetrie*. Wenn R eine vollständige Relation auf X ist, heißt X (bezüglich R) *linear geordnet*.

Beispiel 1.21 Die Relationen kleiner-gleich bzw. größer-gleich auf $\mathbb{Z}, \mathbb{Q}$ oder $\mathbb{R}$ sind lineare Relationen.

Beispiel 1.22 Sei X eine beliebige Menge. Dann ist "$\subseteq$" eine Ordnungsrelation auf der Potenzmenge $\mathcal{P}(X)$. Diese Relation ist aber nicht linear, wenn X mehr als ein Element besitzt. (Wenn x und y zwei verschiedene Elemente sind, gilt weder $\{x\} \subseteq \{y\}$ noch $\{y\} \subseteq \{x\}$.)

Lemma 1.23 *Sei X eine Menge, und sei $\leq$ eine Ordnungsrelation auf X. Dann gibt es höchstens ein $x \in X$ mit $\forall y \in X : y \leq x$.*

Beweis. Seien x_1, x_2 zwei solche Elemente. Dann ist insbesondere $x_2 \leq x_1$ und $x_1 \leq x_2$. Damit ist $x_1 = x_2$. □

Definition Sei eine Ordnungsrelation $\leq$ auf der Menge X gegeben. Ein Element $x \in X$ wie im letzten Lemma heißt *größtes Element* von X.

Ein Element $x \in X$ so dass

$$\forall y \in X : x \leq y \longrightarrow y = x$$

heißt ein *maximales Element* von X.

Wenn X ein größtes Element hat, dann ist dies (offensichtlich) auch ein maximales Element, und auch das einzige maximale Element.

In vielen wichtigen Beispielen gibt es jedoch mehrere maximale Elemente und kein größtes Element. Hier ist ein instruktives Beispiel.

Beispiel 1.24 Sei $X := \{1, 2, \ldots, 100\}$, und sei Y die Teilmenge von $\mathcal{P}(X)$, die aus den Teilmengen von X mit höchstens 10 Elementen besteht. Wir

betrachten die partielle Ordnung "$\subseteq$" auf Y. Nun ist jede Teilmenge mit genau 10 Elementen ein maximales Element von Y, und es gibt kein größtes Element (es gibt keine Teilmenge von X mit höchstens 10 Elementen, die alle Teilmengen mit höchstens 10 Elementen umfasst).

Bemerkung Analog zu den obigen Definitionen kann man *kleinste Elemente* und *minimale Elemente* definieren.

Relationen zwischen zwei Mengen

Man kann den Begriff der Relation erweitern und allgemeiner Relationen zwischen zwei Mengen betrachten. Seien dazu zwei Mengen X und Y gegeben.

Definition Eine *Relation zwischen X und Y* ist eine Teilmenge von $X \times Y$.

Notation Wenn R eine Relation zwischen X und Y ist und $x \in X, y \in Y$, dann schreiben wir $x \sim_R y$ falls $(x, y) \in R$.

Wir werden diese Verallgemeinerung nicht weiter verfolgen und bemerken nur, wie man mittels dieses Begriffs der Relation definieren kann, was eine Abbildung ist. Wir erinnern uns, dass eine Abbildung von X nach Y jedem Element von X genau ein Element von Y zuordnet. Wir erhalten somit die folgende formale Definition.

Definition Eine Abbildung $f : X \longrightarrow Y$ ist eine Relation zwischen X und Y so dass für jedes $x \in X$ genau ein $y \in Y$ mit $x \sim_f y$ existiert.

Bemerkung Aus der Schule kennen Sie den Begriff des *Graphen* einer Funktion. Dies kann man wie folgt für beliebige Abbildungen definieren: Sei $f : X \longrightarrow Y$ eine Abbildung. Dann ist der *Graph* von f die Menge $\{(x, y) \in X \times Y \mid y = f(x)\}$. Nach der obigen Definition sind allerdings eine Abbildung und ihr Graph identisch! Es ist jedoch üblich und sinnvoll, zwischen einer Abbildung und ihrem Graphen zu unterscheiden und z.B. den Graphen einer Abbildung f mit Γ_f zu bezeichen. (Dann ist also $f(x) = y \longleftrightarrow x \sim_f y \longleftrightarrow (x, y) \in \Gamma_f$.)

Diskussion In Beispiel 1.9 haben wir diskutiert, wie man Tupel mittels Abbildungen definieren kann, und oben haben wir Abbildungen mittels Tupel definiert. Diese zirkuläre Definition sollte natürlich aufgehoben werden. Der übliche Weg ist, rein mengentheoretisch zu definieren, was unter einem

Zweiertupel (x, y) für $x, y \in X \times Y$ zu verstehen ist. Man definiert z.B.: $(x, y) := \{x, \{x, y\}\}$.

Eine alternative Möglichkeit wäre, den Abbildungsbegriff axiomatisch vorauszusetzen und die Mengenlehre darauf aufzubauen.

1.5 Halbgruppen, Monoide und Gruppen

Sei im Folgenden X eine Menge.

Definition Eine *Abbildung* $X \times X \longrightarrow X$ heißt eine *Verknüpfung* auf X.

Beispiele für Mengen mit Verknüpfungen sind die Zahlbereiche $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ jeweils mit der Addition und der Multiplikation. Die Subtraktion ist eine Verknüpfung auf $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, nicht aber auf $\mathbb{N}_0$. Die Division ist eine Verknüpfung auf $\mathbb{Q} - \{0\}$ sowie $\mathbb{R} - \{0\}$, nicht jedoch auf $\mathbb{Z} - \{0\}$, $\mathbb{Q}$ oder $\mathbb{R}$.

An diesen Beispielen fällt auf: Man schreibt z.B. $2 + 3 = 5$ und nicht $+(2, 3) = 5$. Eine analoge Schreibweise ist ganz allgemein bei Verknüpfungen üblich. Übliche Symbole für allgemeine Verknüpfungen sind "$\circ$", "$\cdot$", "$*$".

Verknüpfungen auf endlichen Mengen können auch durch eine *Verknüpfungstabelle* angegeben werden.

Beispiel 1.25 Die folgende Tabelle definiert eine Verknüpfung auf der 5-elementigen Menge $\{1, 2, 3, 4, 5\}$.

$\circ$	1	2	3	4	5
1	1	2	3	4	5
2	2	4	1	5	3
3	3	5	4	2	1
4	4	1	5	3	2
5	5	3	2	1	4

Zum Beispiel ist $2 \circ 3 = 1$ und $3 \circ 2 = 5$.

Definition Sei $\circ : X \times X \longrightarrow X$ eine Verknüpfung. Dann heißt $\circ$

- *assoziativ*, falls $\forall x, y, z \in X : x \circ (y \circ z) = (x \circ y) \circ z$.
- *kommutativ*, falls $\forall x, y \in X : x \circ y = y \circ x$.

Ein Element $e \in X$ heißt *neutrales Element*, wenn gilt: $\forall x \in X : x \circ e = e \circ x = x$.

Lemma 1.26 *Jede Verknüpfung hat höchstens ein neutrales Element.*

Beweis. Seien $e, e' \in X$ neutrale Elemente.[4] Dann gilt

$$e' = e \circ e' = e \ .$$

Bei der ersten Gleichung haben wir benutzt, dass e ein neutrales Element ist, und bei der zweiten Gleichung haben wir benutzt, dass e' ein neutrales Element ist. □

Definition Sei $\circ : X \times X \longrightarrow X$ eine Verknüpfung mit einem neutralen Element e, und sei $x \in X$. Ein Element $y \in Y$ mit

- $y \circ x = e$ heißt ein *Links-Inverses* zu x
- $x \circ y = e$ heißt ein *Rechts-Inverses* zu x
- $y \circ x = e$ und $x \circ y = e$ heißt ein *(beidseites) Inverses* zu x.

Beispiel 1.27 Die Verknüpfung aus Beispiel 1.25 hat ein neutrales Element (die 1), und jedes Element hat (eindeutig bestimmte) Rechts- und Links-Inverse (die aber nicht identisch sind). Die Verknüpfung ist aber nicht assoziativ. Z.B. ist $(2 \circ 2) \circ 3 = 4 \circ 3 = 5$ und $2 \circ (2 \circ 3) = 2 \circ 1 = 2$. Sie ist nicht kommutativ, was man leicht daran sieht, dass die Tabelle nicht symmetrisch bez. Spiegelung an der Diagonalen (von oben links nach unten rechts) ist.

Beispiel 1.28 Sei X eine beliebige Menge. Wenn $f : X \longrightarrow X$ und $g : X \longrightarrow X$ zwei Abbildungen sind, dann können wir die Verknüpfung $f \circ g$ der beiden Abbildungen betrachten. Die Zuordnung $(f, g) \mapsto f \circ g$ ist eine Verknüpfung auf der Menge der Abbildungen $\mathrm{Abb}(X, X)$ im Sinne der obigen Definition. Diese Verknüpfung ist offensichtlich assoziativ. Außerdem gibt es ein neutrales Element, nämlich die identische Abbildung $\mathrm{id}_X : X \longrightarrow X,\ x \mapsto x$. Die Elemente mit beidseitigem Inversen sind genau die bijektiven Abbildungen.

Frage Welche Elemente in $\mathrm{Abb}(X, X)$ haben Links-, welche Rechts-Inverse? (Hierbei muss man das Auswahlaxiom benutzen.)

Im Folgenden betrachten wir ausschließlich assoziative Verknüpfungen. Dies bedeutet, dass Klammern grundsätzlich weggelassen werden können.

[4] Mit dieser Formulierung meine ich, dass e und e' nicht notwendigerweise verschieden sein müssen.

Definition

- Eine *Halbgruppe* ist eine Menge mit einer assoziativen Verknüpfung.
- Ein *Monoid* ist eine Halbgruppe mit einem neutralen Element.
- Eine *Gruppe* ist ein Monoid so dass jedes Element ein Inverses hat.

Beispiele für Halbgruppen die bereits erwähnten Zahlbereiche $\mathbb{N}_0, \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ jeweils mit der Addition oder der Multiplikation.

Die Zahlbereiche $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ bilden bezüglich der Addition auch Monoide (das neutrale Element ist die 0). Allerdings ist $\mathbb{N}$ bezüglich der Addition kein Monoid. Bezüglich der Multiplikation sind $\mathbb{N}_0, \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ Monoide (das neutrale Element ist die 1). Eine andere Beispielklasse ist in Beispiel 1.28 gegeben.

Beispiele für Gruppen sind die Zahlbereiche $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ bezüglich der Addition. Allerdings ist $\mathbb{N}_0$ bezüglich der Addition keine Gruppe. Bezüglich der Multiplikation sind $\mathbb{Q} - \{0\}$ und $\mathbb{R} - \{0\}$ Gruppen. Andererseits sind $\mathbb{Q}$ und $\mathbb{R}$ bezüglich der Multiplikation keine Gruppen (0 hat kein Inverses!).

Die ein-elementige Menge $\{e\}$ ist mit der Verknüpfung $e \circ e = e$ eine Gruppe (genannt die *triviale Gruppe*).

Ein (zugegebenerweise merkwürdiges) Beispiel für eine Halbgruppe ist die leere Menge.

Lemma 1.29 *Sei M ein Monoid. Zu $x \in M$ gibt es höchstens ein beidseitig Inverses.*

Beweis. Seien y, z zwei beidseitige Inverse zu x. Dann ist

$$y = y \circ e = y \circ (x \circ z) = (y \circ x) \circ z = e \circ z = z \ .$$

□

Es gilt sogar die folgende stärkere Aussage:

Lemma 1.30 *Sei $x \in M$ invertierbar[5] mit Inversem y. Dann ist y das einzige Rechts- und das einzige Links-Inverse von x.*

Beweis. Sei z ein Rechts-Inverses zu x. Dann ergibt die obige Rechnung $y = z$. Die Aussage über Links-Inverse zeigt man analog. □

[5]Mit *invertierbar* meinen wir immer, dass es ein beidseitiges Inverses gibt.

Notation Sei $x \in M$. Wenn x ein Inverses hat,[6] wird dies oft mit x^{-1} bezeichnet.

Lemma 1.31 *Sei M ein Monoid, und seien $x, y \in M$ invertierbar. Dann ist auch $x \circ y$ invertierbar, und es ist $(x \circ y)^{-1} = y^{-1} \circ x^{-1}$.*

Beweis. Seien $x, y \in M$. Dann ist $(y^{-1} \circ x^{-1}) \circ (x \circ y) = y^{-1} \circ (x^{-1} \circ x) \circ y = y^{-1} \circ y = e = x \circ x^{-1} = x \circ y \circ y^{-1} \circ x^{-1} = (x \circ y) \circ (y^{-1} \circ x^{-1})$. Damit ist per Definition $y^{-1} \circ x^{-1}$ ein Inverses (das Inverse) von $x \circ y$. □

Notation Wenn eine beliebige Verknüpfung $\circ$ auf einer Menge X gegeben ist, wird für $x \in X$ und $n \in \mathbb{N}$ das Element $\overbrace{x \circ \cdots \circ x}^{n \text{ mal}}$ mit x^n bezeichnet.

Sei nun M wieder ein Monoid. Wenn x ein Inverses hat, so ist (für $n \in \mathbb{N}$) $\overbrace{x^{-1} \circ \cdots \circ x^{-1}}^{n \text{ mal}} = (\overbrace{x \circ \cdots \circ x}^{n \text{ mal}})^{-1}$, was man leicht sieht (siehe auch das obige Lemma). Dieses Element wird mit x^{-n} bezeichnet. Man setzt $x^0 := e$. Mit dieser Notation gilt

$$x^n \circ x^m = x^{n+m} \text{ und } (x^n)^m = x^{nm} .$$

Das Verknüpfungssymbol selbst wird oft weggelassen. Also ist $xy = x \circ y$.

Definition Eine Halbgruppe resp. ein Monoid resp. eine Gruppe mit einer kommutativen Verknüpfung heißt *kommutative* oder *abelsche* Halbgruppe etc.

Notation Wenn eine abelsche Halbgruppe gegeben ist, benutzt man oft die folgende "additive Notation": Man benutzt das Symbol "+" für die Verknüpfung, und für ein Element x und $n \in \mathbb{N}$ setzt man $nx := \overbrace{x + \cdots + x}^{n \text{ mal}}$. Wenn Elemente $x_1, \ldots, x_k$ gegeben sind, setzt man $\sum_{i=1}^{k} x_i := x_1 + \cdots + x_k$.

Wenn ein abelsches Monoid M gegeben ist, schreibt man dann 0_M (oder 0) für das neutrale Element, und man setzt $0 \cdot x := 0_M$ für $x \in M$. Wenn x ein inverses Element hat, bezeichnet man dies mit $-x$.

Beachten Sie, dass die additive Notation der üblichen Notation bezüglich der Addition von Zahlen entspricht.

Beispiel 1.32 Sei Σ ein beliebige nicht-leere Menge. (Die Notation Σ deutet an, dass wir Σ als Alphabet betrachten, aber wie gesagt muss Σ nicht

[6] Mit einem *Inversen* meinen wir immer ein beidseitiges Inverses.

notwendigerweise endlich und auch nicht abzählbar sein.) Wir betrachten die Menge aller Tupel beliebiger Länge ≥ 1 von Elementen von Σ zusammen mit einem weiteren Element $\square$. [7] Diese Menge heißt die Menge der *Worte* in Σ und wird mit Σ^* bezeichnet, $\square$ wird *leeres Wort* genannt. Statt $(a_1, \dots a_n)$ schreibt man $a_1 \cdots a_n$.

Durch "Hintereinanderschreiben" vw kann man zwei Wörtern v, w ein neues zuordnen (wobei $v\square := v, \square w := w$). Man erhält also eine Verknüpfung auf Σ^*. Mit dieser Verknüpfung ist Σ^* ein Monoid (mit neutralem Element $\square$). Wenn Σ mehr als ein Element enthält, ist dieses Monoid nicht abelsch.

Beispiel 1.33 Sei X eine beliebige Menge. Dann ist $\mathrm{Abb}(X, X)$ ein Monoid (siehe Beispiel 1.28). Wenn allerdings X mindestens zwei Elemente enthält[8] ist $\mathrm{Abb}(X, X)$ keine Gruppe und auch nicht abelsch.

Beweis. Seien $a, b \in X$ zwei verschiedene Elemente. Dann ist die Abbildung $f : x \mapsto a$ nicht injektiv (da insbesondere a und b auf a abgebildet werden), besitzt also kein Inverses (keine Umkehrabbildung). Damit ist X keine Gruppe. Sei $g : x \mapsto b$. Dann ist $g \circ f$ ist durch $x \mapsto b$ gegeben und $f \circ g$ durch $x \mapsto a$ gegeben. Damit ist X nicht abelsch. $\square$

Definition Sei wiederum X eine beliebige Menge. Eine *Permutation* auf X ist eine bijektive Abbildung $X \longrightarrow X$. Die Menge der Permutationen von X wird mit $\mathrm{Perm}(X)$ oder mit $S(X)$ bezeichnet. Für $n \in \mathbb{N}$ ist $S_n := S(\{1, \dots, n\})$ die Menge der Permutationen auf $\{1, \dots, n\}$.

Beispiel 1.34 $\mathrm{Perm}(X)$ ist eine Gruppe. Diese Gruppe ist genau dann abelsch, wenn X höchstens zwei Elemente besitzt.

Diese Gruppe heißt die *symmetrische Gruppe* auf X; die Gruppe S_n heißt die *symmetrische Gruppe auf n Elementen.*

Beweis. Da jede bijektive Abbildung eine Umkehrabbildung hat, ist $\mathrm{Perm}(X)$ eine Gruppe.

Wenn X kein oder ein Element besitzt, besteht $\mathrm{Perm}(X)$ nur aus der identischen Abbildung. Wenn X zwei Elemente a, b besitzt, besteht $\mathrm{Perm}(X)$ aus id_X und τ mit $\tau : a \mapsto b, b \mapsto a$. Damit ist $\mathrm{Perm}(X)$ kommutativ.

Seinen nun a, b, c drei verschiedene Elemente von X. Betrachte die Abbildungen

$$f : a \mapsto b,\ b \mapsto a,\ x \mapsto x \text{ für } x \neq a, b$$

[7] Da man für $n \geq 1$ die Menge der n-Tupel Σ^n als die Menge $\Sigma^{\{1,\dots,n\}}$ auffassen kann (siehe Beispiel 1.9), ist es naheliegend, $\square$ als die eindeutige Abbildung $\emptyset \longrightarrow \Sigma$ zu definieren.

[8] Mit dieser Redewendung meine ich "mindesteins zwei *verschiedene* Elemente".

sowie

$$g : a \mapsto b,\ b \mapsto c,\ c \mapsto a,\ x \mapsto x \text{ für } x \neq a, b, c\ .$$

Dann ist $(g \circ f)(a) = g(b) = c$, $(f \circ g)(a) = f(b) = a$, also insbesondere $f \circ g \neq g \circ f$. □

Produkte

Seien X, Y zwei Halbgruppen. Wir definieren wie folgt eine Verknüpfung auf $X \times Y$:

$$(x, y) \circ (x', y') := (x \circ x', y \circ y')\ .$$

Diese "komponentenweise definierte" Verknüpfung ist offensichtlich assoziativ. Damit ist auch $X \times Y$ eine Halbgruppe.

Wenn X und Y Monoide mit neutralen Elementen e_X, e_Y sind, dann ist auch $X \times Y$ ein Monoid mit neutralem Element (e_X, e_Y).

Wenn nun x und y invertierbar sind, ist offensichtlich auch (x, y) invertierbar mit Inversem (x^{-1}, y^{-1}). Insbesondere ist $X \times Y$ eine Gruppe, wenn X und Y Gruppen sind.

Außerdem erhält man abelsche Halbgruppen, abelsche Monoide oder abelsche Gruppen, wenn X und Y abelsche Halbgruppen, abelsche Monoide oder abelsche Gruppen sind.

Selbstverständlich gelten diese Aussagen auch für mehr als zwei Faktoren X, Y. Damit ist also insbesondere X^n (für $n \in \mathbb{N}$) in natürlicher Weise eine Halbgruppe, ein Monoid oder eine Gruppe, wenn X eine Halbgruppe, ein Monoid oder eine Gruppe ist.

Sei nun X weiterhin eine Halbgruppe, und sei I eine Menge. Auch auf X^I können wir in natürlicher Weise eine Verknüpfung definieren. Wir erinnern daran, dass X^I aus den Abbildungen $I \longrightarrow X$ besteht, und diese Abbildungen werden oft in der Form $(x_i)_{i \in I}$ geschrieben. Wir folgen dieser Schreibweise. Wir haben die folgende Verknüpfung auf X^I:

Gegeben $(x_i)_{i \in I}, (x'_i)_{i \in I}$, definieren wir

$$(x_i)_{i \in I} \circ (x'_i)_{i \in I} := (x_i \circ x'_i)_{i \in I}\ .$$

Damit ist X^I wiederum eine Halbgruppe. Wenn X ein Monoid mit neutralem Element e ist, dann ist X^I ein Monoid mit neutralem Element $(e)_{i \in I}$ (dies ist die Abbildung, die jedem $i \in I$ das neutrale Element e von X zuordnet). Die Halbgruppe X^I ist eine Gruppe, wenn X eine Gruppe ist.

Unterstrukturen

Definition Sei X eine Menge mit einer Verknüpfung "$\circ$", und sei Y eine Teilmenge von X. Dann heißt Y *abgeschlossen* bezüglich "$\circ$", wenn gilt:

$$\forall y, y' \in Y : y \circ y' \in Y$$

In diesem Fall definiert $\circ$ durch Einschränkung auf $Y \times Y$ eine Verknüpfung auf Y, man spricht auch von der *induzierten Verknüpfung*.

Beachten Sie: Wenn die Verknüpfung auf X assoziativ (resp. kommutativ) ist, so ist auch die induzierte Verknüpfung assoziativ (resp. kommutativ). Wir haben somit:

- Sei H eine Halbgruppe, und sei $U \subseteq H$ abgeschlossen (bezüglich der Verknüpfung auf H). Dann ist U mit der induzierten Verknüpfung eine Halbgruppe; man spricht von einer *Unterhalbgruppe*.
- Sei M ein Monoid mit neutralem Element e, und sei $U \subseteq M$ abgeschlossen mit $e \in U$. Dann ist U mit der induzierten Verknüpfung ein Monoid mit neutralem Element e; man spricht von einem *Untermonoid*.
- Sei G eine Gruppe mit neutralem Element e, und sei $U \subseteq G$ abgeschlossen mit $e \in U$ so dass für jedes $x \in U$ auch das inverse Element x^{-1} in U liegt. Dann ist U mit der induzierten Verknüpfung eine Gruppe mit neutralem Element e; man spricht von einer *Untergruppe*.

Beispiel 1.35 Sei $(G, +)$ eine additiv geschriebene abelsche Gruppe, und seien $g_1, \ldots, g_k \in G$. Dann ist

$$\langle g_1, \ldots, g_k \rangle := \{z_1 g_1 + \cdots + z_k g_k \mid z_i \in \mathbb{Z}\}$$

eine Untergruppe von G (nachrechnen!). Es gilt: Wenn $U \subseteq G$ irgendeine Untergruppe mit $g_1, \ldots, g_k \in U$ ist, dann ist $\langle g_1, \ldots, g_k \rangle \subseteq U$ (warum?). $\langle g_1, \ldots, g_k \rangle$ ist also die *kleinste* Untergruppe von G, die $g_1, \ldots, g_k$ umfasst (siehe den Unterabschnitt über Ordnungsrelationen im vorherigen Abschnitt).

Die Untergruppe $\langle g_1, \ldots, g_k \rangle$ von G heißt die von $g_1, \ldots, g_k$ *erzeugte* Untergruppe bzw. das *Erzeugnis* von $g_1, \ldots, g_k$.

Frage Natürlich kann man eine abelsche Gruppe auch "multiplikativ" schreiben. Sei $(G, \circ)$ so eine Gruppe. Wie lautet dann das Erzeugnis von $g_1, \ldots, g_k \in G$? Wie lautet z.B. das Erzeugnis von $2, 3, 5 \in \mathbb{Q}^*$? Sind die Untergruppen $\langle 2 \rangle$ und $\langle -2 \rangle$ von $\mathbb{Q}^*$ identisch?

Beispiel 1.36 Sei M ein Monoid, und sei $G \subseteq M$ die Menge der invertierbaren Elemente von M. Dann ist (nach Lemma 1.31) G abgeschlossen und somit ein Untermonoid. Es ist per Definition auch eine Gruppe, genannt die *Gruppe der invertierbaren Elemente* von M.

Beispiel 1.37 Die Gruppe der invertierbaren Elemente von $(\mathbb{Z}, \cdot)$ ist $\{1, -1\}$.

Beispiel 1.38 Sei X eine Menge. Dann ist die Gruppe der invertierbaren Elemente von $\mathrm{Abb}(X, X)$ gleich $\mathrm{Perm}(X)$.

Aussage 1.39 (Untergruppenkriterium) *Sei G eine Gruppe, und sei $U \subseteq G$ eine Teilmenge. Dann ist U (mit der induzierten Verknüpfung) genau dann eine Untergruppe von G, wenn gilt:*

- *U ist nicht-leer*
- $\forall x, y \in U : x \circ y^{-1} \in U$.

Beweis. Wenn U eine Untergruppe ist, gelten die Eigenschaften offensichtlich. Seien also nun die beiden Eigenschaften erfüllt. Da U nicht-leer ist, gibt es ein $x_0 \in U$. Damit ist nach der zweiten Eigenschaft $e = x_0 \circ x_0^{-1} \in U$. Für $x \in U$ beliebig ist dann auch $x^{-1} = e \circ x^{-1} \in U$. Seien nun $x, y \in U$ beliebig. Dann ist $y^{-1} \in U$, wie wir gerade gesehen haben. Somit ist auch $x \circ y = x \circ (y^{-1})^{-1} \in U$. Dies ist die Abgeschlossenheit. Die beiden anderen Eigenschaften wurden zuvor gezeigt. □

Faktorgruppen

Sei nun $(G, +)$ eine *abelsche* Gruppe, und sei U eine Untergruppe. Wir führen wie folgt eine Relation auf G ein:

$$x \sim_U y :\Longleftrightarrow x - y \in U \ .$$

Man sieht leicht, dass dies eine Äquivalenzrelation ist; die Äquivalenzklasse zu $x \in G$ bezeichnen wir mit $[x]_U$. Damit gilt also: Für $x \in G$ ist

$$[x]_U = \{x + u \mid u \in U\} \ .$$

Diese Menge wird auch mit $x + U$ bezeichnet.

Wir wollen auf $G_{/\sim_U}$ eine Verknüpfung definieren durch

$$[x]_U + [y]_U := [x + y]_U \ .$$

Hierzu müssen wir die Wohldefiniertheit nachweisen, d.h. wir müssen nachweisen, dass gilt:

$$\forall x, y, x', y' \in G : x \sim_U x' \wedge y \sim_U y' \longrightarrow x + y \sim_U x' + y'$$

Seien also $x, y, x', y' \in G$ mit $x \sim_U x'$ und $y \sim_U y'$. Dann ist also $x - x' \in U$ und $y - y' \in U$. Nun ist $(x + y) - (x' + y') = (x - x') + (y - y') \in U$, d.h. $x + y \sim_U x' + y'$. □

Man rechnet leicht nach, dass die Verknüpfung auf $G_{/\sim_U}$ assoziativ und abelsch ist. Damit ist also $G_{/\sim_U}$ mit der soeben definieren Verknüpfung eine Halbgruppe. Außerdem ist $[0_G]$ ein neutrales Element von $G_{/\sim_U}$, und für $x \in G$ ist $[-x]$ ein inverses Element von $[x]$.

Damit ist $G_{/\sim_U}$ sogar eine abelsche Gruppe.

Definition Die soeben definierte Gruppe wird mit G/U bezeichnet und heißt die *Faktorgruppe* von G *nach* U (oder G *modulo* U), die Verknüpfung heißt wiederum die *induzierte Verknüpfung*.

Zur Verdeutlichung: Es ist $0_{G/U} = [0_G]_U$ und $-[x]_U = [-x]_U$ für alle $x \in G$.

Beispiel 1.40 Sei n eine natürliche Zahl > 1, und sei $n\mathbb{Z} := \{z \in \mathbb{Z} \mid n \text{ teilt } z\} = \{na \mid a \in \mathbb{Z}\}$. Dann ist $n\mathbb{Z}$ eine Untergruppe von $(\mathbb{Z}, +)$. Die Äquivalenzrelation $\sim_{n\mathbb{Z}}$ ist (per Definition) durch

$$x \sim_{n\mathbb{Z}} y \longleftrightarrow x - y \in n\mathbb{Z} \longleftrightarrow n \text{ teilt } x - y \longleftrightarrow x \equiv y \mod n$$

gegeben. Mit anderen Worten, es ist die Relation "Kongruenz modulo n", die wir in den Beispielen 1.14 und 1.17 diskutiert haben. Ich erinnere daran, dass wir für $x \in \mathbb{Z}$ die Äquivalenzklasse mit $[x]_n$ bezeichnen, und es ist $[x]_n = \{x + na \mid a \in \mathbb{Z}\}$. Die Gruppe $\mathbb{Z}/n\mathbb{Z}$ besteht folglich aus den n Elementen $[0]_n, [1]_n, \ldots, [n-1]_n$.

1.6 Ringe und Körper

Definition Ein *Ring* ist eine Menge R mit zwei Verknüpfungen "+" und "·" so dass

- $(R, +)$ eine abelsche Gruppe ist,
- $(R, \cdot)$ ein Monoid ist,

- die *Distributivgesetze* gelten, d.h.

$$\forall a, b, c \in R : (a + b)c = ac + bc \, , \quad c(a + b) = ca + cb \, .$$

Ein Ring heiß *kommutativ*, wenn die Multiplikation eine kommutative Verknüpfung ist.

Notation Bei den Distributivgesetzen haben wir die übliche Rechenregel "mal vor plus" benutzt und die Klammern auf der rechten Seite der Gleichung entsprechend weglassen. So verfahren wir auch im Folgenden.

Bemerkung Beachten Sie, sich das Adjektiv *kommutativ* hier nur auf die Multiplikation bezieht; die Addition eines Rings ist per Definition immer kommutativ.

Beispiel 1.41 Die ganzen, die rationalen sowie die reellen Zahlen bilden jeweils kommutative Ringe mit den Verknüpfungen "+" und "$\cdot$". Die neutralen Elemente sind 0 und 1.

Beispiele für nicht-kommutative Ringe werden wir später kennenlernen.

Notation Das neutrale Element bezüglich "+" wird mit 0 (oder genauer mit 0_R) bezeichnet, und das neutrale Element bezüglich "$\cdot$" wird mit 1 (oder genauer mit 1_R) bezeichnet. Es gilt also ("wie üblich") $0 + r = r + 0 = r$ und $1 \cdot r = r \cdot 1 = r$ für alle $r \in R$.

Aussage 1.42 *Sei R ein Ring.*

a) Für alle $r \in R$ gilt $0 \cdot r = 0$.

b) Für alle $r \in R$ gilt $(-1) \cdot r = -r$.

c) Wenn $0 = 1$ (in R), dann ist $R = \{0\}$.

Beweis.
zu a) Sei $r \in R$ beliebig. Dann ist $0 \cdot r = (0 + 0) \cdot r = 0 \cdot r + 0 \cdot r$. Hieraus folgt $0 \cdot r = 0$.
zu b) Sei wiederum $r \in R$ beliebig. Dann ist $0 = 0 \cdot r = (1 - 1) \cdot r = 1 \cdot r + (-1) \cdot r = r + (-1) \cdot r$. Daraus folgt die Behauptung.
zu c) Sei $0 = 1$ und sei $r \in R$ beliebig. Dann ist $r = 1 \cdot r = 0 \cdot r = 0$. (Andererseits ist die Menge $\{0\}$ mit den Verknüpfungen $0 + 0 = 0$ und $0 \cdot 0 = 0$ ein Ring.) □

Definition Ein *Körper* ist ein kommutativer Ring mit $0 \neq 1$, in dem jedes Element $\neq 0$ ein Inverses bezüglich der Multiplikation hat.

Beispiel 1.43 Beispiele für Körper sind die rationalen und die reellen Zahlen.

Notation Körper werden oft mit K bezeichnet.

Definition Sei R ein Ring. Die Gruppe der invertierbaren Elemente von $(R, \cdot)$ wird mit R^* bezeichnet (siehe Beispiel 1.36).

Aussage 1.44 *Sei R ein Ring mit $0 \neq 1$. Dann ist $0 \notin R^*$. Wenn K ein Körper ist, dann ist $K^* = K - \{0\}$ (mit der induzierten Verknüpfung).*

Beweis. Sei also R ein Ring mit $0 \neq 1$. Angenommen $0 \in R^*$. Dann gibt es ein $r \in R$ mit $0 \cdot r = 1$. Dies ist ein Widerspruch, da nach Aussage 1.40 $0 \cdot r = 0$ ist.

Sei nun K ein Körper. Soeben haben wir gesehen, dass $K^* \subseteq K - \{0\}$. Es gilt aber auch $K - \{0\} \subseteq K^*$ nach Definition eines Körpers. □

Die folgende Definition ist analog zur Definition der Unterstrukturen in Abschnitt 1.5.

Bemerkung / Definition Sei R ein Ring, und sei $U \subseteq R$ abgeschlossen bezüglich Addition und Multiplikation so dass U eine Untergruppe von R bezüglich der Addition und ein Untermonoid von R bezüglich der Multiplikation ist (insbesondere ist also $0, 1 \in U$). Dann ist U ein Ring; man spricht von einem *Unterring*.

Sei nun R ein Körper. Dann haben insbesondere alle Elemente von $U - \{0\}$ ein multiplikatives Inverses in K. Falls diese Inversen alle in U liegen, ist U ein Körper; man spricht von einem *Unterkörper* oder einem *Teilkörper*.

Faktorringe

Sei von nun an R ein *kommutativer* Ring und sei U eine Untergruppe (!) von $(R, +)$. Wie in Abschnitt 1.5 erhalten wir dann die Faktorgruppe R/U, deren Operation von der Addition induziert wird (die Multiplikation von R vernachlässigen wir im Moment).

Wir wollen nun R/U auch zu einem Ring machen. Genauer wollen wir eine Operation "$\cdot$" auf R/U definieren mit $[r]_U \cdot [s]_U = [r \cdot s]_U$ für alle $r, s \in R$. Es stellt sich nun die Frage, unter welchen Bedingungen an U dies möglich ist.

Wir nehmen zunächst an, dass wir eine solche Operation auf R/U gegeben haben. Dann gilt für alle $u \in U$: $0_{R/U} = [u]_U$, und somit gilt für alle $u \in U$ und $r \in R$: $0_{R/U} = [r]_U \cdot [u]_U = [ru]_U$, d.h. $ru \in U$.

Sogleich werden wir sehen, dass diese notwendige Bedingung auch hinreichend ist. Doch zunächst halten wir das Kriterium in einer Definition fest.

Definition Sei R ein kommutativer Ring. Ein *Ideal* von R ist eine Teilmenge $I \subseteq R$ so dass

- I eine Untergruppe von $(R, +)$ ist
- $\forall a \in I \; \forall r \in R : ra \in I$.

Beispiel 1.45 Sei R ein Ring, und seien $r_1, \ldots, r_k \in R$. Dann ist

$$(r_1, \ldots, r_k) := \{x_1 r_1 + \cdots + x_k r_k \mid x_i \in R\}$$

ein Ideal von R (nachrechnen!).[9]

Es gilt: Wenn $I \subseteq R$ ein Ideal mit $r_1, \ldots, r_k \in I$ ist, dann ist $(r_1, \ldots, r_k) \subseteq I$. Das Ideal $(r_1, \ldots, r_k)$ ist also das *kleinste* Ideal, das $r_1, \ldots, r_k$ umfasst (siehe den Unterabschnitt über Ordnungsrelationen).

Das Ideal $(r_1, \ldots, r_k)$ heißt das von $r_1, \ldots, r_k$ *erzeugte* Ideal in R.

Aussage 1.46 *Sei R ein kommutativer Ring, und sei $I \subseteq R$ ein Ideal. Dann gibt es eine eindeutig bestimmte Verknüpfung "$\cdot$" auf R/I so dass*

$$[r]_I \cdot [s]_I = [rs]_I \; .$$

Mit der bereits definierten Addition und dieser Multiplikation ist R/I ein kommutativer Ring.

Beweis. Die Eindeutigkeit ist klar. Für die Existenz müssen wir das Folgende nachweisen:

$$\forall r, s, r', s' \in R : r \sim_I r', s \sim_I s' \longrightarrow rs \sim_I r's'$$

Seien also $r, s, r', s' \in R$ mit $r \sim_I r', s \sim_I s'$, d.h. $r - r' \in I, s - s' \in I$. Dann ist $rs - r's' = rs - r's + r's - r's' = (r - r')s + r'(s - s')$. Nun sind $(r - r')s = s(r - r') \in I$ und $r'(s - s') \in I$ aufgrund der zweiten Eigenschaft eines Ideals. Damit ist auch $rs - r's' \in I$.

Wir müssen nun noch zeigen, dass mit der bereits definierten Addition und der soeben definierten Multiplikation R/I ein kommutativer Ring ist. Dies ist leicht (Übungsaufgabe). □

[9] Vorsicht! Die Notation $(r_1, \ldots, r_k)$ hat nun zwei Bedeutungen: Einerseits steht sie für das Tupel $(r_1, \ldots, r_k)$, andererseits für das Ideal $(r_1, \ldots, r_k)$.

Definition Der soeben definierte Ring R/I heißt der *Faktorring* von R *nach* I (oder R *modulo* I).

Beispiel 1.47 Sei $n > 1$, und wie in Beispiel 1.40 sei $n\mathbb{Z} := \{na | a \in \mathbb{Z}\}$. Diese Untergruppe von $\mathbb{Z}$ ist das von n erzeugte Ideal (also $n\mathbb{Z} = (n)$). Damit ist $\mathbb{Z}/n\mathbb{Z}$ mit der induzierten Addition und Multiplikation ein Ring.

Wenn n keine Primzahl ist, dann ist $\mathbb{Z}/n\mathbb{Z}$ aber kein Körper. Denn: Sei $n = n_1 n_2$ mit $n_1, n_2 \neq 1$. Dann ist $[n_1]_n, [n_2]_n \neq [0]_n$, aber es ist $[n_1]_n \cdot [n_2]_n = [n]_n = [0]_n$. In Abschnitt 1.9 werden wir sehen, dass umgekehrt $\mathbb{Z}/n\mathbb{Z}$ immer ein Körper ist, wenn n prim ist.

1.7 Die ganzen und die rationalen Zahlen

Bisher sind wir stillschweigend davon ausgegangen, dass die ganzen und die rationalen Zahlen existieren und einige offensichtliche Eigenschaften haben. Hier wollen wir nun zeigen, wie man – ausgehend von den natürlichen Zahlen – explizit definieren kann, was der Ring der ganzen Zahlen und der Körper der rationalen Zahlen ist, bzw. eine explizite Konstruktion der ganzen Zahlen und der rationalen Zahlen angeben kann.

Wir stellen uns auf den Standpunkt, dass wir die die Menge $\mathbb{N}_0$ der natürlichen Zahlen einschließlich der Null existiert, und die Addition und Multiplikation in dieser Menge die üblichen Gesetze wir Assoziativität, Kommutativität und Distributivität erfüllen.

Nun zuerst zu den ganzen Zahlen. Es gibt zwei naheliegende Möglichkeiten, von den natürlichen Zahlen zu den ganzen Zahlen zu gelangen.

Die erste Möglichkeit ist, $\mathbb{Z}$ als die Vereinigung von positiven, negativen Zahlen und der Null zu definieren. Man fixiert eine Menge M, die bijektiv zu $\mathbb{N}$ ist aber von $\mathbb{N}$ verschieden ist. Wir haben also eine Bijektion $m : \mathbb{N} \longrightarrow M$. Dann setzt $Z := \mathbb{N}_0 \cup M$; dies ist nun eine disjunkte Vereinigung, und für jedes Element $z \in Z$ gilt nun: Entweder es ist $z \in \mathbb{N}_0$ oder es gibt ein (eindeutig bestimmtes $n \in \mathbb{N}$ mit $z = m(n)$.

Nun kann man die Operationen "+" und "$\cdot$" per Fallunterscheidung auf von $\mathbb{N}_0$ auf Z ausdehnen. Für "+" sieht das dann so aus: Wir definieren $0 + z := z$ und $z + 0 = z$ für alle $z \in Z$, sowie für $a, b \in \mathbb{N}$:

$$\begin{array}{lll} a + b & := a + b & \\ a + m(b) & := a - b & \text{wenn } a \geq b \\ a + m(b) & := m(b - a) & \text{wenn } b > a \\ m(a) + b & := b - a & \text{wenn } b \geq a \\ m(a) + b & := m(a - b) & \text{wenn } a > b \end{array}$$

Man beachte, dass sich die Operationen "+" und "-" auf der rechten Seite auf die schon bekannten natürlichen Zahlen beziehen.

Jetzt muss man noch die Multiplikation definieren und dann nachweisen, dass Assoziativität, Kommutativität und Distributivität immer noch gelten. Der Ring $\mathbb{Z}$ ist dann per Definition $(Z, +, \cdot)$.

Wir schildern nun eine andere Möglichkeit, die auf einer allgemeinen Methode beruht, die auch in anderem Kontext Anwendung findet.

Die Idee ist, dass sich jede ganze Zahl in der Form $a - b$ mit $a, b \in \mathbb{N}_0$ schreiben lässt, aber diese Darstellung ist nicht eindeutig. Genauer: Es gilt für $a, c, b, d \in \mathbb{N}_0$: $a - b = c - d \longleftrightarrow a + d = b + c$. Diese Überlegung nimmt man nun zum Anlass, die ganzen Zahlen mittels Äquivalenzklassen von Tupeln in $\mathbb{N}_0 \times \mathbb{N}_0$ zu definieren.

Man definiert eine Äquivalenzrelation auf $\mathbb{N}_0 \times \mathbb{N}_0$ wie folgt: $(a, b) \sim (c, d) :\Longleftrightarrow a + d = b + c$, und man setzt $Z := (\mathbb{N}_0 \times \mathbb{N}_0)_{/\sim}$. Wir wollen nun eine Verknüpfung "+" ("Addition") auf $\mathbb{Z}$ definieren durch

$$[(a, b)] + [(c, d)] := [(a + c, b + d)] \; .$$

Hierzu müssen wir nachprüfen, dass dies *wohldefiniert* ist. Sei hierzu $(a, b) \sim (a', b')$ und $(c, d) \sim (c', d')$ mit $a, a', b, b', c, c', d, d' \in \mathbb{N}_0$. Dann ist also $a + b' = a' + b$ und $c + d' = c' + d$. Somit ist

$$(a + c) + (b' + d') = (a + b') + (c + d') = (a' + b) + (c' + d) = (a' + c') + (b + d)$$

und somit $(a + c, b + d) \sim (a' + c', b' + d')$. □

Um die Multiplikation zu definieren, erinnern wir uns, dass $(a-b)\cdot(c-d) = ac + bd - (ad + bc)$. In diesen Sinne wollen wir eine weitere Verknüpfung "$\cdot$" ("Multiplikation") auf Z definieren durch

$$[(a, b)] \cdot [(c, d)] := [(ac + bd, ad + bc)] \; .$$

Wieder rechnet man nach, dass dies wohldefiniert ist. Dann muss man noch zeigen, dass die Assoziativ-, Kommutativ- und Distributivgesetze gelten.

Der Ring $\mathbb{Z}$ ist nun per Definition wieder $(Z, +, \cdot)$.

Wir kommen nun zu den rationalen Zahlen. Nun gehen wir von den ganzen Zahlen $\mathbb{Z}$ aus. Diesmal ist der Ausgangspunkt, dass sich jede rationale Zahl als Bruch $\frac{a}{b}$ mit $a \in \mathbb{Z}$ und $b \in \mathbb{N}$ schreiben lässt, und es gilt für $a, c \in \mathbb{Z}$ und $b, d \in \mathbb{N}$: $\frac{a}{b} = \frac{c}{d} \longleftrightarrow ad = cb$.

Somit definieren wir eine Äquivalenzrelation auf $\mathbb{Z}\times\mathbb{N}$: $(a, b) \sim (c, d) :\Longleftrightarrow ad = cb$, und wir setzen $Q := (\mathbb{Z} \times \mathbb{N})_{/\sim}$.

Wir wollen nun zwei Verknüpfungen "+" und "·" ("Addition" und "Multiplikation") auf Q wie folgt definieren:

$$[(a,b)] + [(c,d)] := [(ad+bc, bd)] \;, \quad [(a,b)] \cdot [(c,d)] := [(ac, bd)]$$

Beachten Sie, dass diese Operationen die üblichen Operation "+" und "·" auf den rationalen Zahlen nachempfinden, wenn man Äquivalenzklassen als Brüche auffasst.

Wiederum zeigt man, dass diese Operationen wohldefiniert sind und die Assoziativ-, Kommutativ- und Distributivgesetze gelten.

Nun kann man noch $\frac{a}{b} := [(a,b)]$ setzen und erhält die rationalen Zahlen in bekannter Darstellung. Der Körper der rationalen Zahlen ist dann $\mathbb{Q} := (Q, +, \cdot)$.

1.8 Morphismen

Oftmals will man "Rechnungen" von einem Monoid, einer Gruppe, einem Ring usw. in anderes Objekt "der gleichen Art" transferieren. Hierzu kann man *Homomorphismen* (strukturerhaltende Abbildungen) benutzen. Eine andere Frage, die hiermit in engem Zusammenhang steht, ist, wann man zwei mathematische Strukturen als "strukturgleich" ansehen sollte.

Homomorphismen von Halbgruppen, Monoiden und Gruppen

Definition

- Seien H und H' zwei Halbgruppen. Ein *Homomorphismus von Halbgruppen* von H nach H' ist eine Abbildung $\varphi : H \longrightarrow H'$ mit $\varphi(a \circ_H b) = \varphi(a) \circ_{H'} \varphi(b)$ für alle $a, b \in H$.
- Seien M und M' Monoide mit neutralen Elementen e und e'. Ein *Homomorphismus von Monoiden* von M nach M' ist eine Abbildung $\varphi : M \longrightarrow M'$ mit $\varphi(a \circ_H b) = \varphi(a) \circ_{H'} \varphi(b)$ für alle $a, b \in M$ und $\varphi(e) = e'$.
- Seien G und G' Gruppen. Dann sind G und G' insbesondere Monoide, und ein Homomorphismus von Gruppen von G nach G' ist ein Homomorphismus von G nach G' als Monoide.

Beispiel 1.48 Die Abbildung $\mathbb{N}_0 \longrightarrow \mathbb{N}_0,\ x \mapsto 2x$ ist ein Homomorphismus von Monoiden von $(\mathbb{N}_0, +)$ nach $(\mathbb{N}_0, +)$. Ebenso ist die "Null-Abbildung" $\mathbb{N}_0 \longrightarrow \mathbb{N}_0,\ x \mapsto 0$ ein Homomorphismus von Monoiden von $(\mathbb{N}_0, +)$ nach

$(\mathbb{N}_0, +)$. Es ist auch ein Homomorphismus von Halbgruppen von $(\mathbb{N}_0, \cdot)$ nach $(\mathbb{N}_0, \cdot)$, aber es ist kein Homomorphismus von Monoiden von $(\mathbb{N}_0, \cdot)$ nach $(\mathbb{N}_0, \cdot)$.

Beispiel 1.49 Sei H eine Halbgruppe (resp. ein Monoid, resp. eine Gruppe) und $U \subseteq H$ eine Unterhalbgruppe (resp. ein Untermonoid, resp. eine Untergruppe. Dann ist die Inklusion $U \hookrightarrow H$ ein Homomorphismus von Halbgruppen (resp. von Monoiden, resp. von Gruppen).

Beispiel 1.50 Sei G eine (additiv geschriebene) abelsche Gruppe, und sei $U \subseteq G$ eine Untergruppe. In Abschnitt 1.5 haben wir die Faktorgruppe G/U eingeführt. Ich wiederhole, dass die Verknüpfung "+" auf G/U $[a]_U + [b]_U = [a+b]_U$ für alle $a, b \in G$ erfüllt. Damit ist die Abbildung $G \twoheadrightarrow G/U,\ a \mapsto [a]_U$ ein Homomorphismus von Gruppen.

Lemma 1.51 *Seien M und M' Monoide und sei $\varphi : M \longrightarrow M'$ eine Abbildung mit $\varphi(a \circ_M b) = \varphi(a) \circ_{M'} \varphi(b)$ für alle $a, b \in M$ so dass $\varphi(e)$ ein Rechts- oder ein Links-Inverses hat. Dann ist $\varphi(e) = e'$, und folglich ist φ ein Homomorphismus von Monoiden.*

Beweis. Es ist $\varphi(e) = \varphi(e \circ_M e) = \varphi(e) \circ_{M'} \varphi(e)$. Multiplikation mit einem Rechts- bzw. Links-Inversen (von Rechts bzw. Links) liefert $e' = \varphi(e)$. □

Da in einer Gruppe jedes Element invertierbar ist, ergibt sich:

Aussage 1.52 *Seien G und G' Gruppen und sei $\varphi : G \longrightarrow G'$ mit $\varphi(a \circ_G b) = \varphi(a) \circ_{G'} \varphi(b)$ für alle $a, b \in G$ (d.h. φ ist ein Homomorphismus von Halbgruppen von G nach G'). Dann ist φ ein Homomorphismus von Gruppen.*

Aussage 1.53 *Seien M und M' Monoide, $\varphi : M \longrightarrow M'$ ein Homomorphismus von Monoiden, und sei $a \in M$ invertierbar. Dann ist $\varphi(a)$ invertierbar, und es ist $\varphi(a)^{-1} = \varphi(a^{-1})$.*

Beweis Es ist $e' = \varphi(e) = \varphi(a \circ_M a^{-1}) = \varphi(a) \circ_{M'} \varphi(a^{-1})$ sowie $e' = \varphi(e) = \varphi(a^{-1} \circ_M a) = \varphi(a^{-1}) \circ_{M'} \varphi(a)$. Damit folgt die Behauptung. □

Aussage 1.54 *Seien M und M' Monoide, $\varphi : M \longrightarrow M'$ ein Homomorphismus von Monoiden. Dann ist* Bild(φ) *ein Untermonoid von M'. Wenn M eine Gruppe ist, dann ist auch* Bild(φ) *eine Gruppe.*

Beweis. Bild(φ) ist offensichtlich abgeschlossen. Da es auch e' enthält, ist es ein Untermonid von M'. Die zweite Behauptung folgt aus der obigen Aussage. □

Bemerkung Sei $\iota : M \longrightarrow M'$ ein injektiver Homomorphismus von Monoiden. Oftmals "identifiziert" man dann M mit seinem Bild in M'. Dies bedeutet, dass man nicht zwischen $a \in M$ und $\iota(a) \in M'$ unterscheidet. Ein typisches Beispiel hierfür ist das folgende: Im vorherigen Abschnitt haben wir die Menge $Z := (\mathbb{N}_0 \times \mathbb{N}_0)_{/\sim}$ definiert; die Elemente in dieser Menge sind per Definition die ganzen Zahlen. Wir haben einen injektiven Homomorphismus $\iota : \mathbb{N}_0 \longrightarrow Z, a \mapsto [(a,0)]_\sim$. Dies ist ein Homomorphismus von Monoiden bezüglich der Addition und der Multiplikation. Sicher macht es Sinn, natürliche Zahlen mit ihrem Bild (der entsprechenden ganzen Zahl) zu identifizieren.

Man muss aber aufpassen, mittels dieser "Identifizierungen" keine unsinnigen "Identitäten" abzuleiten. Ein Beispiel hierzu: Die injektive Abbildung $\iota : \mathbb{Z} \longrightarrow \mathbb{Z}\,,\ x \mapsto 2x$ ist ein Homomorphismus von Monoiden bezüglich der Addition. Wenn wir nun x mit $\iota(x)$ "identifizieren", "erhalten" wir $x = 2x$ für alle $x \in \mathbb{Z}$!

Definition Sei $\varphi : M \longrightarrow M'$ ein Homomorphismus von Monoiden. Dann ist $\mathrm{Kern}(\varphi) := \{a \in M \mid \varphi(a) = e'\}$.

Aussage 1.55 $\mathrm{Kern}(\varphi)$ *ist ein Untermonoid von* M*.*

Beweis. Es ist $e \in \mathrm{Kern}(\varphi)$, denn $\varphi(e) = e'$ nach Definition. Sei $a, b \in \mathrm{Kern}(\varphi)$. Dann ist $\varphi(a \circ_M b) = \varphi(a) \circ_{M'} \varphi(b) = e' \circ_{M'} e' = e'$, also $a \circ_M b \in \mathrm{Kern}(\varphi)$. □

Aussage 1.56 *Sei* M *eine Gruppe. Dann ist* $\mathrm{Kern}(\varphi)$ *eine Untergruppe von* M*. Die Abbildung* φ *genau dann injektiv, wenn* $\mathrm{Kern}(\varphi) = \{e\}$*.*

Beweis. Sei $a \in \mathrm{Kern}(\varphi)$. Dann ist $\varphi(a^{-1}) = \varphi(a)^{-1} = e'$ nach Aussage 1.53.

Offensichtlich ist $e \in \mathrm{Kern}(\varphi)$. Wenn nun φ injektiv ist, gilt insbesondere $\#\,\mathrm{Kern}(\varphi) = \#\varphi^{-1}(\{e'\}) \leq 1$, also $\mathrm{Kern}(\varphi) = \{e\}$.

Sei andererseits $\mathrm{Kern}(\varphi) = \{e\}$, und seien $a, b \in M$ mit $\varphi(a) = \varphi(b)$. Dann ist also $\varphi(a \circ_M b^{-1}) = \varphi(a) \circ_{M'} \varphi(b^{-1}) = \varphi(a) \circ_{M'} \varphi(b)^{-1} = e'$, also $a \circ_M b^{-1} \in \mathrm{Kern}(\varphi)$ und somit $a \circ_M b^{-1} = e$. Damit ist $a = b$. □

Homomorphismen kann man verknüpfen, und man erhält wieder einen Homomorphismus:

Aussage 1.57 *Seien* A, A', A'' *Halbgruppen (resp. Monoide), und seien* $\varphi : A \longrightarrow A'$ *und* $\psi : A' \longrightarrow A''$ *Homomorphismen von Halbgruppen (resp. Monoiden). Dann ist* $\psi \circ \varphi : A \longrightarrow A''$ *ein Homomorphismus von Halbgruppen (resp. Monoiden).*

Beweis. Seien $a, B \in A$ beliebig. Dann ist $(\psi \circ \varphi)(a \circ_A b) = \psi(\varphi(a \circ_A b)) = \psi(\varphi(a) \circ_{A'} \varphi(b)) = \psi(\varphi(a)) \circ_{A''} \psi(\varphi(b)) = (\psi \circ \varphi)(a) \circ_{A''} (\psi \circ \varphi)(b)$.

Wenn es sich um Monoide handelt, gilt zusätzlich $(\psi \circ \varphi)(e) = \psi(\varphi(e)) = \psi(e') = e''$, wobei e, e' und e'' jeweils die neutralen Elemente in A, A' und A'' sind. □

Bemerkung Diese Aussage gilt selbstverständlich auch für Gruppen. Denn Homomorphismen von Gruppen sind ja per Definition Homomorphismen von Monoiden.

Homomorphismen von Ringen und Körpern

Definition

- Seien R und R' Ringe. Ein *Homomorphismus von Ringen* von R nach R' ist eine Abbildung $\varphi : R \longrightarrow R'$ mit $\varphi(a +_R b) = \varphi(a) +_{R'} \varphi(b)$ und $\varphi(a \cdot_R b) = \varphi(a) \cdot_{R'} \varphi(b)$ für alle $a, b \in R$ sowie $\varphi(1_R) = 1_{R'}$.

- Seien K und K' Körper. Dann sind K und K' insbesondere Ringe. Ein *Homomorphismus von Körpern* von K nach K' ist ein Homomorphismus von Ringen von K nach K'.

Bemerkung Ein Homomorphismus von Ringen von R nach R' ist also eine Abbildung $R \longrightarrow R'$, welche ein Homomorphismus der abelschen Gruppe $(R, +)$ sowie des Monoids $(R, \cdot)$ ist.

Beispiel 1.58 Analog zu Beispiel 1.49 gilt: Sei R ein Ring und $U \subseteq R$ ein Unterring. Dann ist die Inklusion $U \hookrightarrow R$ ein Homomorphismus von Ringen.

Beispiel 1.59 Sei R ein kommutativer Ring, und sei $I \subseteq R$ ein Ideal. In Abschnitt 1.6 haben wir den Faktorring R/I eingeführt. Nun ist $R \twoheadrightarrow R/I$ ein Homomorphismus von Ringen. Beachten Sie, dass dies analog zu Beispiel 1.50 ist.

Wiederum gilt:

Aussage 1.60 *Seien A, A', A'' Ringe (resp. Körper), und seien $\varphi : A \longrightarrow A'$ und $\psi : A' \longrightarrow A''$ Homomorphismen von Ringen (resp. Körpern). Dann ist $\psi \circ \varphi : A \longrightarrow A''$ ein Homomorphismus von Ringen (resp. Körpern).*

Der *Beweis* ist analog zum Beweis von Aussage 1.57.

Definition Sei $\varphi : R \longrightarrow R'$ ein Homomorphismus von Ringen. Dann ist $\mathrm{Kern}(\varphi) := \{r \in R \mid \varphi(r) = 0_{R'}\}$.

Aussage 1.61 *Sei $\varphi : R \longrightarrow R'$ ein Homomorphismus von kommutativen Ringen. Dann ist* $\mathrm{Kern}(\varphi)$ *ist ein Ideal von R.*

Beweis. Wir haben schon gesehen, dass $\mathrm{Kern}(\varphi)$ eine Untergruppe von $(R, +)$ ist. Sei also $a \in \mathrm{Kern}(\varphi)$ und $r \in R$. Dann ist $\varphi(r \cdot_R a) = \varphi(r) \cdot_{R'} \varphi(a) = \varphi(r) \cdot 0_{R'} = 0_{R'}$, also $r \cdot_R a \in \mathrm{Kern}(\varphi)$. □

Isomorphismen, Endomorphismen und Automorphismen

Definition Sei $\varphi : A \longrightarrow A'$ ein Homomorphismus von Halbgruppen (resp. Monoiden, resp. Gruppen, resp. Ringen, resp. Körpern). Dann heißt φ *Isomorphismus* wenn es einen Homomorphismus $\psi : A' \longrightarrow A$ mit $\psi \circ \varphi = \mathrm{id}_A, \varphi \circ \psi = \mathrm{id}_{A'}$ von Halbgruppen (resp. Monoiden, resp. Gruppen, resp. Ringen, resp. Körpern) gibt. Wenn es einen Isomorphismus $A \longrightarrow A'$ von Halbgruppen (resp. Monoiden, resp. Gruppen, resp. Ringen. resp. Körpern) gibt, heißen A und A' *isomorph* (als Halbgruppen (resp. Monoide, resp. Gruppen, resp. Ringe, resp. Körper)).

Aussage 1.62 *Ein Homomorphismus (von Halbgruppen, Monoiden, Gruppen, Ringen oder Körpern) ist genau dann ein Isomorphismus, wenn er bijektiv ist.*

Beweis. Ein Isomorphismus ist offensichtlich bijektiv (es gibt eine Umkehrabbildung). Sei andererseits $\varphi : A \longrightarrow A'$ ein bijektiver Homomorphismus, und sei $\varphi^{-1} : A' \longrightarrow A$ die Umkehrabbildung. Zu zeigen ist nun, dass φ^{-1} auch ein Homomorphismus ist.

Wir betrachten zuerst den Fall von Halbgruppen. Seien $a', b' \in A'$ mit $a' = \varphi(a), b' = \varphi(b)$. Dann ist $\varphi^{-1}(a' \circ_{A'} b') = \varphi^{-1}(\varphi(a) \circ_{A'} \varphi(b)) = \varphi^{-1}(\varphi(a \circ_A b)) = a \circ_A b = \varphi^{-1}(a') \circ_A \varphi^{-1}(b')$, was zu zeigen war.

Sei nun φ ein Homomorphismus von Monoiden. Dann ist φ^{-1} ein Homomorphismus von Halbgruppen, und es gilt $\varphi^{-1}(e') = \varphi^{-1}(\varphi(e)) = e$, d.h. φ^{-1} ist ein Homomorphismus von Monoiden.

Für Homomorphismen von Gruppen ist nun nichts mehr zu zeigen, und die Aussagen für Ringe und Körper folgen sofort aus den obigen Argumenten (angewandt auf Addition und Multiplikation). □

Definition Ein Homomorphismus $A \longrightarrow A$ heißt *Endomorphismus* von A; die Menge der Endomorphismen von A wird mit $\mathrm{End}(A)$ bezeichnet. Ein Endomorphismus, der zusätzlich ein Isomorphismus ist, heißt *Automorphismus*; die Menge der Automorphismen von A wird mit $\mathrm{Aut}(A)$ bezeichnet.

Bemerkung Wenn A ein Monoid ist, ist A auch eine Halbgruppe. Nun ist nicht notwendigerweise jeder Endomorphismus von A als Halbgruppe ein Endomorphismus von A als Monoid. (Beispiel: Die Nullabbildung auf $(\mathbb{N}_0, \cdot)$.) Insbesondere sollte man also aufpassen, wenn man die Notation $\mathrm{End}(A)$ benutzt und explizit angeben, ob man A als Monoid oder als Halbgruppe "betrachtet".

Ein ähnliches Problem tritt bei Ringen auf: Wenn man die Multiplikation "vergißt", "wird" jeder Ring eine abelsche Gruppe. Wenn R ein Ring ist, ist aber nicht notwendigerweise jeder Homomorphismus der abelsche Gruppe $(R, +)$ auch ein Endomorphismus von R als Ring.

Aussage 1.63 *Sei A eine Halbgruppe, ein Monoid, eine Gruppe, ein Ring oder ein Körper. Dann ist* $\mathrm{End}(A)$ *bezüglich der normalen Verknüpfung von Abbildungen ein Monoid, und* $\mathrm{Aut}(A)$ *ist die Gruppe der invertierbaren Elemente in* $\mathrm{End}(A)$ *(siehe Beispiel 1.36).*

Bemerkung Beachten Sie, dass diese Aussage analog zu den Aussagen, dass $\mathrm{Abb}(X, X)$ ein Monoid und $\mathrm{Perm}(X)$ eine Gruppe ist, ist (siehe Beispiel 1.38).

Seien G und G' (additiv geschriebene) abelsche Gruppen. In Abschnitt 1.5 (Unterabschnitt über Produkte) haben wir eine Verknüpfung auf $\mathrm{Abb}(G, G')$ definiert. Da G' eine abelsche Gruppe ist, ist $\mathrm{Abb}(G, G')$ mit dieser Verknüpfung auch eine abelsche Gruppe. Ich wiederhole, dass die Verknüpfung wie folgt definiert ist: Für $\varphi, \psi \in \mathrm{Abb}(G, G')$ ist $\varphi + \psi : G \longrightarrow G'$ durch

$$(\varphi + \psi)(a) := \varphi(a) + \psi(a)$$

definiert. (Diese Definition ist identisch zu der in Abschnitt 1.5, nur die Notation ist anders). Beachten Sie, dass wir bisher nur die Gruppenstruktur von G' und nicht die von G ausgenutzt haben.

Sei $\mathrm{Hom}(G, G')$ die Menge der Homomorphismen von G nach G'. Dies ist eine Teilmenge von $\mathrm{Abb}(G, G')$ und sogar eine Untergruppe (nachrechnen!). Damit ist $\mathrm{Hom}(G, G')$ also auch eine abelsche Gruppe.

Insbesondere ist also $\mathrm{End}(G)$ eine abelsche Gruppe mittels der soeben definierten Addition von Homomorphismen. Wir wissen auch schon, dass $\mathrm{End}(G)$ ein Monoid bezüglich der Verknüpfung von Abbildungen ist. Außerdem gelten die Distributivgesetze (für $\varphi, \chi, \psi \in \mathrm{End}(G)$)

$$\varphi \circ (\chi + \psi) = \varphi \circ \chi + \varphi \circ \psi \qquad (\chi + \psi) \circ \varphi = \chi \circ \varphi + \psi \circ \varphi \; .$$

(Nachrechnen!) Damit ist $(\mathrm{End}(G), +, \circ)$ ein Ring, genannt der *Endomorphismenring* von G. Beachten Sie, dass dieser Ring nicht notwendigerweise kommutativ ist.

Strukturtransport

Sei nun A eine Halbgruppe, sei X eine Menge, und sei $f : A \longrightarrow X$ eine *bijektive* Abbildung. Wir definieren wie folgt auf X eine Verknüpfung $*$: Für $x, y \in X$ definieren wir

$$x * y := f(f^{-1}(x) \circ f^{-1}(y)) \; .$$

Damit gilt für alle $a, b \in A$:

$$f(a) * f(b) = f(f^{-1}(f(a)) \circ f^{-1}(f(b))) = f(a \circ b) \; . \tag{1.4}$$

Man sieht leicht, dass die Verknüpfung $*$ auf X assoziativ ist (nachrechnen!), d.h. X ist mit $*$ eine Halbgruppe. Aus (1.4) folgt nun, dass f ein Homomorphismus ist. Da f auch bijektiv ist, ist f somit ein Isomorphismus, und A und $(X, *)$ sind isomorph als Halbgruppen.

Man sieht nun leicht: Wenn A ein Monoid ist, dann ist auch $(X, *)$ ein Monoid (mit neutralem Element $f(e)$), und wenn A eine Gruppe ist, so auch $(X, *)$.

Analoge Aussagen gelten, wenn A ein Ring oder ein Körper ist.

Die soeben angewandte Methode, eine Verknüpfung auf X zu definieren und f zu einem Isomorphismus zu machen, nennt man *Strukturtransport*.

Beispiel 1.64 Sei $n > 1$. Die Zahlen $0, 1, \ldots, n-1$ bilden ein Repräsentantensystem bezüglich der Äquivalenzrelation "Kongruenz modulo n". Wir erhalten somit mittels "Strukturtransport" zwei Verknüpfungen "$\oplus_n$" und "$\odot_n$" auf $\{0, 1, \ldots, n\}$ so dass für alle $a, b \in \{0, 1, \ldots, n-1\}$ gilt: $[a \odot_n b]_n = [a]_n + [b]_n = [a + b]_n$ und $[a \odot_n b]_n = [a]_n \cdot [b]_n = [a \cdot b]_n$. Da $\mathbb{Z}/n\mathbb{Z}$ ein Ring ist, ist somit $(\{0, 1, \ldots, n-1\}, \oplus_n, \odot_n)$ auch ein Ring.

Wahrscheinlich kennen Sie die Verknüpfungen "$\oplus_n$" und "$\odot_n$" aus der Schule: Sie beschreiben die Arithmetik "modulo n": $a \oplus_n b$ ist der Rest von $a + b$ bei der Division mit n, analog ist $a \odot_n b$ der Rest von $a \cdot b$ bei der Division mit n.

Sei für eine natürliche Zahl $n > 1$ und $a \in \mathbb{Z}$ $\mathrm{mod}(a, n)$ die kleinste Zahl in $\mathbb{N}_0$, die kongruent zu a modulo n ist. Dann ist für $a \in \mathbb{Z}$ $\mathrm{mod}(a, n)$ der Repräsentant von $[a]_n$ in $0, 1, \ldots, n-1$. Es ist also $a \oplus_n b = \mathrm{mod}(a + b, n)$ und $a \odot_n b = \mathrm{mod}(a \cdot b, n)$.

Kategorien (Diskussion)

Oben haben wir zu einer Vielzahl verschiedener mathematischer Objekte entsprechende Homomorphismen zugeordnet. Dabei fällt auf, dass es in diesem

Kontext einige Aussagen gibt, die gelten egal ob man nun von Homomorphismen von Halbgruppen, Monoiden, Gruppen, Ringen oder Körpern redet. Eine Beispiel hierfür sind Aussageen 1.57 und 1.60. Diese kann man wie folgt knapp zusammenfassen:

Seien A, A', A'' "Objekte von gleichem Typ", und seien $\varphi : A \longrightarrow A'$ und $A' \longrightarrow A''$ Homomorphismen dieser Objekte. Dann ist auch $\psi \circ \varphi : A \longrightarrow A''$ ein Homomorphismus dieser Objekte.

Dies ist natürlich keine mathematisch rigorose Aussage, da nicht klar ist, was "Objekte von gleichem Typ" und "Homomorphismen dieser Objekte" sein sollen.

Dies kann man jedoch präzise machen, indem man den Begriff einer *Kategorie* einführt. Der Begriff der Kategorie ist ziemlich abstrakt, ich versuche eine intuitive Annäherung zu geben.

Betrachten Sie als Beispiel alle Mengen zusammen mit allen Abbildungen. Wie wir zu Beginn gesehen haben, macht es keinen Sinn, von der *Menge aller Menge* zu reden. Um trotzdem alle Mengen zusammenfassen zu können, führt man den Begriff einer *Klasse* ein und spricht z.B. von der *Klasse* aller Menge. Die Elemente der Klasse (in diesem Fall die Mengen) heißen nun *Objekte*.

Eine Kategorie besteht nun aus einer Klasse zusammen mit Folgendem: Zu je zwei Objekten A, A' der Klasse gibt es eine Menge $\mathrm{Mor}(A, A')$ von sogenannten *Morphismen*. Diese Morphismen schreibt man suggestiv wie Abbildungen (d.h. wenn $\varphi \in \mathrm{Mor}(A, A')$, dann schreibt man $\varphi : A \longrightarrow A'$), obwohl es nicht notwendigerweise Abbildungen sein müssen.

Dabei sollen gewisse Eigenschaften gelten, die für Abbildungen offensichtlich sind. Z.B. soll man Morphismen "hintereinanderschalten" können.

Wir haben nun schon einige Kategorien gesehen, nämlich die Kategorien der Mengen, der Halbgruppen, der Monoide, der Gruppen, der Ringe und der Körper. Im Fall der Mengen sind die Morphismen per Definition die Abbildungen, und ansonsten sind die Morphismen die oben definierten Homomorphismen. Außerdem macht es noch Sinn, von den Kategorien der abelschen Halbgruppen, der abelschen Monoide, der abelschen Gruppen und der kommutativen Ringe zu sprechen.

1.9 Der Euklidische Algorithmus und Moduloarithmetik

Der größte gemeinsame Teiler

Notation Seien $a, b \in \mathbb{Z}$. Wir schreiben $a|b$, wenn a die Zahl b teilt, d.h. wenn es eine ganze Zahl c mit $b = ac$ gibt. Wenn dies nicht der Fall ist,

schreiben wir $a \nmid b$.

Bemerkungen

- Die Relation "teilt" ist offensichtlich reflexiv und transitiv. Auf Zahlen in $\mathbb{N}_0$ eingeschränkt ist sie auch antisymmetrisch und somit eine Ordnungsrelation. Für ganze Zahlen ist sie aber nicht antisymmetrisch und somit keine Ordnungsrelation (wenn $a|b$ und $b|a$, dann ist $a = \pm b$, und beide Vorzeichen sind möglich).

- Seien $a, b \in \mathbb{Z}$, und seien $(a) = \{za | z \in \mathbb{Z}\}$ und $(b) = \{zb | z \in \mathbb{Z}\}$ die von a bzw. b erzeugten Ideale. Dann gilt

$$a|b \longleftrightarrow b \in (a) \longleftrightarrow (b) \subseteq (a) \ .$$

Definition Seien $a_1, a_2, \ldots, a_k$ ganze Zahlen, die nicht alle $= 0$ sind. Die größte natürliche Zahl g mit $g|a_1, g|a_2, \ldots, g|a_k$ heißt der *größte gemeinsame Teiler* von $a_1, \ldots, a_k$, Bezeichung: $\mathrm{ggT}(a_1, \ldots, a_k)$.

Satz 1.1 *Seien $a_1, a_2, \ldots, a_k \in \mathbb{Z}$, nicht alle $= 0$, und sei $b \in \mathbb{Z}$ mit $b|a_1, \ldots, b|a_k$. Dann gilt $b|\,\mathrm{ggT}(a_1, \ldots, a_k)$.*

Der *Beweis* dieses Satzes ist recht einfach, wenn man die eindeutige Primfaktorzerlegung von natürlichen Zahlen voraussetzt.

Ich gebe nun einen Beweis, der auf dem Studium von Idealen in $\mathbb{Z}$ beruht. Dabei werden noch einige andere interessante Ergebnisse abfallen. Mit Methoden wie den folgenden kann man übrigens auch die eindeutige Primfaktorzerlegung beweisen.

Lemma 1.65 *Sei I ein Ideal von $\mathbb{Z}$. Dann gibt es ein (eindeutig bestimmtes) $n \in \mathbb{N}_0$ mit $I = (n)$.*

Beweis. Falls $I = \{0\}$ ist, ist die Behauptung offensichtlich richtig. Sei also $I \neq \{0\}$. Dann enthält I natürliche Zahlen (wenn $z \in I$, dann ist auch $-z \in I$). Sei n die kleinste natürliche Zahl in I. Ich behaupte, dass $I = (n)$.

Denn: Sei $z \in I$. Dann gibt es ein $r \in \{0, \ldots, n-1\}$ und ein $p \in \mathbb{Z}$ mit $z = pn + r$. Da $pn \in I$, ist auch $r = z - pn \in I$. Da n (per Definition) die kleinste natürliche Zahl in I ist, ist $r = 0$. Damit ist also $z = pn \in (n)$.

Die Zahl $n \in \mathbb{N}_0$ ist offensichtlich eindeutig bestimmt: Wenn $I = (n)$ und $I = (m)$ mit $n, m \in \mathbb{N}_0$, dann ist $n|m$ und $m|n$, also $n = m$. □

Lemma 1.66 *Seien $a_1, \ldots, a_k \in \mathbb{Z}$, nicht alle $= 0$, und sei $n \in \mathbb{N}$ mit*

$$(a_1, \ldots, a_k) = (n) \ .$$

Dann gilt $n|a_i$ für alle $i = 1, \ldots, k$.

Für $b \in \mathbb{N}_0$ mit $b|a_i$ für alle $i = 1, \ldots, k$ gilt $b|n$.

Insbesondere ist $n = \mathrm{ggT}(a_1, \ldots, a_k)$, und der obige Satz ist richtig.

Beweis. Da $a_i \in (n)$, gilt $n|a_i$.

Sei nun $b \in \mathbb{N}_0$ mit $b|a_i$ für alle $i = 1, \ldots, k$. Dann gilt also $a_i \in (b)$ für alle i. Daraus folgt $(a_1, \ldots, a_k) \subseteq (b)$ (weil $(a_1, \ldots, a_k)$ das kleinste Ideal ist, das alle a_i enthält). Es gilt also $(n) \subseteq (b)$, also $b|n$.

Wir haben gerade gesehen, dass gilt: Für alle $b \in \mathbb{N}_0$ mit $b|a_i$ für $i = 1, \ldots, k$ gilt $b|n$, und also insbesondere auch $b \leq n$. Damit erfüllt n die Definition des ggT von $a_1, \ldots, a_k$, und es gilt $n = \mathrm{ggT}(a_1, \ldots, a_k)$. □

Eine Umformulierung des obigen Lemmas ist: Seien $a_1, \ldots, a_k \in \mathbb{Z}$. Dann ist $\mathrm{ggT}(a_1, \ldots, a_k)$ die eindeutig bestimmte Zahl in $\mathbb{N}_0$ mit

$$(a_1, \ldots, a_k) = (\mathrm{ggT}(a_1, \ldots, a_k)) \; . \tag{1.5}$$

Aus diesem Grund schreibt man auch oft $(a_1, \ldots, a_k)$ anstelle von $\mathrm{ggT}(a_1, \ldots, a_k)$. Beachten Sie, dass es aufgrund von (1.5) ganze Zahlen $z_1, \ldots, z_k$ mit

$$\mathrm{ggT}(a_1, \ldots, a_k) = z_1 a_1 + \cdots + z_k a_k \tag{1.6}$$

gibt.

Beachten Sie, dass so eine Darstellung des ggT nicht eindeutig ist. Ich verdeutliche dies am Beispiel von zwei ganzen Zahlen a, b. Sei $\mathrm{ggT}(a, b) = xa + yb$. Dann ist z.B. auch $\mathrm{ggT}(a, b) = (x + b)a + (y - a)b$.

Bemerkung Für alle $z_1, \ldots, z_k \in \mathbb{Z}$, nicht alle $= 0$, gilt $\mathrm{ggT}(a_1, \ldots, a_k) \mid z_1 a_1 + \cdots + z_k a_k$, wie man leicht sieht. Insbesondere gilt also: Wenn es ganze Zahlen $z_1, \ldots, z_k$ mit $z_1 a_1 + \cdots + z_k a_k = 1$ gibt, dann ist $\mathrm{ggT}(a_1, \ldots, a_k) = 1$.

Wie Sie wissen, ist eine Primzahl der Definition eine natürliche Zahl > 1, die nur unter den natürlichen Zahlen nur von sich selbst und 1 geteilt wird. Als eine Anwendung der Darstellung (1.6) erhalten wir den folgenden Satz:

Satz 1.2 *Sei p eine Primzahl. Dann gilt:*

$$\forall a, b \in \mathbb{N} : p|ab \longrightarrow p|a \vee p|b \; .$$

Beweis. Seien $a, b \in \mathbb{N}$. Wir zeigen: Wenn p sowohl a also auch b nicht teilt, dann teilt p auch ab nicht.

Es gelte also $p \nmid a$ und $p \nmid b$. Dann ist also $\mathrm{ggT}(a, p) = 1$ und $\mathrm{ggT}(b, p) = 1$. Damit gibt es $w, x, y, z \in \mathbb{Z}$ mit $wa + xp = 1$ und $yb + zp = 1$. Es folgt:

$$1 = (wa + xp) \cdot (yb + zp) = wy \cdot ab + (waz + xyb + xzp) \cdot p \; .$$

Damit ist $\mathrm{ggT}(ab, p) = 1$ nach der obigen Bemerkung. Insbesondere ist also p kein Teiler von ab. □

Aussage 1.67 *Sei $n > 1$ eine natürliche Zahl, und sei $a \in \mathbb{Z}$. Dann ist $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ genau dann invertierbar, wenn $\mathrm{ggT}(a, n) = 1$.*

Beweis. Die folgenden Aussagen sind äquivalent:
- $[a]_n$ ist invertierbar.
- Es gibt ein $x \in \mathbb{Z}$ so dass $[a]_n \cdot [x]_n = [1]_n$.
- Es gibt ein $x \in \mathbb{Z}$ so dass $ax \equiv 1 \mod n$.
- Es gibt $x, y \in \mathbb{Z}$ so dass $ax + ny = 1$.
- $\mathrm{ggT}(a, n) = 1$. □

Insbesondere gilt:

Aussage 1.68 *Sei p eine Primzahl. Dann ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper.*

Notation Der Körper $\mathbb{Z}/p\mathbb{Z}$ wird auch mit $\mathbb{F}_p$ bezeichnet.

Der Euklidische Algorithmus

Der *Euklidische Algorithmus* ist ein Algorithmus, um den ggT zweier ganzer Zahlen explizit auszurechnen. Man kann dann auch leicht eine Darstellung der Form (1.6) zu finden. Ausgangspunkt ist die folgende Beobachtung:

Wir wollen den ggT von $a, b \in \mathbb{Z}$ berechnen. Offensichtlich können wir uns auf $a, b \in \mathbb{N}_0$ und $a > b$ beschränken. Es gibt nun eindeutig bestimmte Zahlen $p_1 \in \mathbb{N}_0$ und $r_1 \in \{0, \ldots, b-1\}$ mit

$$a = p_1 b + r_1 \ .$$

(Teilen mit Rest). Nun gilt

$$\mathrm{ggT}(a, b) = \mathrm{ggT}(r_1, b) \ . \tag{1.7}$$

Ich gebe zwei Beweise dieser Aussage an:

1. Beweis. Es gilt $\mathrm{ggT}(a, b) | a, b$ und somit gilt auch $\mathrm{ggT}(a, b) | r_1$ (das ist leicht). Damit gilt nach der Definition von $\mathrm{ggT}(r_1, b)$: $\mathrm{ggT}(a, b) \leq \mathrm{ggT}(r_1, b)$. Analog sieht man, dass umgekehrt $\mathrm{ggT}(r_1, b) \leq \mathrm{ggT}(a, b)$. Damit sind beide ggT's gleich.

2. Beweis Die beiden Ideale (a, b) und (b, r_1) sind gleich. Damit folgt die Aussage über den ggT nach Lemma 1.66. □

Wenn nun $r_1 = 0$ ist, dann gilt $b|a$, d.h. $\mathrm{ggT}(a,b) = b$. Wenn anderseits $r_1 \neq 0$, dann gibt es eindeutig bestimmte Zahlen $p_2 \in \mathbb{N}_0$ und $r_2 \in \{, \ldots, r_1 - 1\}$ mit

$$b = p_2 r_1 + r_2 \ ,$$

und wiederum ist

$$\mathrm{ggT}(b, r_1) = \mathrm{ggT}(r_2, r_1) \ .$$

Wenn man so fortfährt, erhält man

$$\begin{array}{rcl} a & = & p_1 b + r_1 \\ b & = & p_2 r_1 + r_2 \\ r_1 & = & p_3 r_2 + r_3 \\ & \vdots & \\ r_i & = & p_{i+2} r_{i+1} + r_{i+2} \\ & \vdots & \end{array}$$

mit $r_{i+2} \in \{0, \ldots, r_{i+1} - 1\}$. Diese Prozedur terminiert immer (da die Reste immer kleiner werden). Sagen wir, dass $r_k = 0$ (und $r_{k-1} \neq 0$). Also:

$$\begin{array}{rcl} a & = & p_1 b + r_1 \\ b & = & p_2 r_1 + r_2 \\ r_1 & = & p_3 r_2 + r_3 \\ & \vdots & \\ r_{k-3} & = & p_{k-1} r_{k-2} + r_{k-1} \\ r_{k-2} & = & p_k r_{k-1} + 0 \end{array}$$

Dann gilt $p_k r_{k-1} = r_{k-2}$, und es ist $\mathrm{ggT}(a,b) = \mathrm{ggT}(b, r_1) = \mathrm{ggT}(r_1, r_2) = \cdots = \mathrm{ggT}(r_{k-2}, r_{k-1}) = r_{k-1}$.

Wir kommen nun zur Berechnung einer Darstellung des ggT in der Form (1.6). Die Idee ist, von "unten nach oben" zurückzurechnen.

Wir haben $\mathrm{ggT}(a,b) = r_{k-1} = r_{k-3} - p_{k-1} r_{k-2}$. Nun substituieren wir r_{k-2} mittels der Zeile "darüber". Wir erhalten eine Darstellung der Form $\mathrm{ggT}(a,b) = x_{k-4} r_{k-4} + y_{k-4} r_{k-3}$ mit gewissen ganzen Zahlen x_{k-4}, y_{k-4}. Wenn wir so fortfahren, erhalten wir Darstellungen

$$\mathrm{ggT}(a,b) = x_i r_i + y_i r_{i+1}$$

mit gewissen $x_i, y_i \in \mathbb{Z}$ für alle $i \geq 1$. Ausgehend von

$$\mathrm{ggT}(a,b) = x_1 r_1 + y_1 r_2$$

liefert eine Substitution von r_2 mittels der zweiten Zeile eine Darstellung

$$\mathrm{ggT}(a,b) = x_0 b + y_0 r_1 \ ,$$

und eine weitere Substitution mittels der ersten Zeile liefert eine Darstellung

$$\mathrm{ggT}(a,b) = xa + yb \ ,$$

wie gewünscht.

Der gesamte soeben beschriebene Algorithmus heißt *erweiterter Euklidischer Algorithmus.*

Der erweiterte Euklidische Algorithmus kann auch benutzt werden, um inverse Elemente "modulo n" zu bestimmen. Die Standardaufgabe ist diese: Geben sei $n > 1$ und $a \in \{1, \ldots, n-1\}$. Bestimme nun, ob $[a]_n$ invertierbar ist, und wenn dies der Fall ist, bestimme ein $z \in \{1, \ldots, n-1\}$ mit $[a]_n \cdot [z]_n = [1]_n$!

Hierzu bestimmt man mittels des Euklidischen Algorithmus zuerst $\mathrm{ggT}(a,n)$. Nach Aussage 1.68 ist $[a]_n$ genau dann invertierbar, wenn dieser ggT gleich 1 ist. Falls dies der Fall ist, bestimmt man wie oben beschrieben $x, y \in \mathbb{Z}$ mit $xa + yn = 1$. Dann ist also $[x]_n \cdot [a]_n = [1]_n$. Das gewünschte $z \in \{1, \ldots, n-1\}$ ist die kleinste natürliche Zahl kongruent zu x modulo n. (Da man nur an $\mathrm{mod}(x,n)$ interessiert ist, können beim "Rückrechnen" alle Rechnungen "modulo n" erfolgen.)

Sei nun $E(a,b)$ die Anzahl der Schritte, die man mit dem Euklidischen Algorithmus benötigt, um den größten gemeinsamen Teiler von a und b auszurechnen. D.h. es gilt (per Definition): $E(a,b) = 1$, falls $b|a$ und $E(a,b) = E(b, \mathrm{mod}(a,b)) + 1$, falls $b \nmid a$.

Lemma 1.69 *Für $a > b$ gilt $E(a,b) \leq 2\lfloor \log_2(a) \rfloor$. In jedem Fall gilt $E(a,b) \leq 2\lfloor \log_2(b) \rfloor + 1$.*

Beweis. Wir zeigen zunächst die erste Aussgage und führen den Beweis per Induktion nach a.

Beachten Sie, dass nach Vor. $a \geq 2$ ist.

Sei nun $a \geq 2$. Wir setzen voraus, dass die Behauptung für $a' < a$ richtig ist (d.h. für alle $a' < a$ und alle $b < a'$ gilt $E(a',b) \leq 2\lfloor \log_2(a') \rfloor$). Wir zeigen die Behauptung für a.

Sei $a > b$. Seien nun p_1, r_1 mit $a = p_1 b + r_1$ und $r_1 < b$ gegeben. Wenn $r_1 = 0$, ist $E(a,b) = 1$, und die Behauptung ist richtig. Sei also $r_1 \neq 0$.

Seien nun $p_2, r_2 \in \mathbb{N}_0$ mit $b = p_2 r_1 + r_2$ und $r_2 < r_1$ gegeben. Wenn $r_2 = 0$, ist $E(a,b) = 2$, und die Behauptung ist wiederum richtig. Sei also $r_2 \neq 0$.

Nun ist $E(a,b) = E(b,r_1)+1 = E(r_1,r_2)+2$, und es ist $a \geq b+r_1 > 2r_1$. (Beachten Sie, dass $p_1 > 0$.) Nach I.V. ist nun $E(r_1,r_2) \leq 2\lfloor \log_2(r_1) \rfloor \leq 2\lfloor \log_2(\frac{a}{2}) \rfloor = 2\lfloor \log_2(a) - 1 \rfloor = 2\lfloor \log_2(a) \rfloor - 2$. Damit ist $E(a,b) \leq 2\lfloor \log_2(a) \rfloor$.

Die zweite Behauptung folgt nun aus der ersten Behauptung aufgrund von $E(a,b) = E(b, \mathrm{mod}(a,b)) + 1$ für $a \nmid b$. □

Bemerkung / Frage Im obigen Induktionsbeweis scheint die Indunktionsbasis "$a = 2$" zu fehlen. Der Beweis ist aber richtig, und die Beh. für $a = 2$ ist insbesondere auch bewiesen. Warum ist das so?

Komplexität arithmetischer Operationen

Wir kommen zur Frage, wieviel "Zeit" man (auf einem Computer) benötigt, um elementare Rechenoperationen auszuführen. Insbesondere geht es um die Frage, wie schnell man "modulo n" rechnen kann.

Das Wort *Zeit* im vorherigen Absatz steht in Anführungszeichen, da es nicht um eine konkrete Zeitangabe geht sondern darum, wieviele elementare Operationen man auf einem idealisierten Modell eines Computers ausführen müßte.

Um diese Frage rigoros zu beantworten, müßten wir zuerst so einen idealisierten Computer formal definieren. Das würde uns aber zu weit weg führen.

Wir setzen voraus, dass unser idealisierter Computer Wörter über dem Alphabet $\{0,1\}$ verarbeitet; solche Wörter nennen wir *Bit-Stings*. (Zusätzlich sollte man noch ein Stoppzeichen haben, das das Ende der Eingabe anzeigt.). Wenn wir nun eine natürliche Zahl n gebeben haben, stellen wir diese in "2-adischer Entwicklung" dar:

$$n = n_0 2^0 + n_1 2^1 + \cdots + n_{k-1} 2^{k-1}$$

mit $n_i \in \{0,1\}$ für alle $i = 0, \ldots, k-1$ und $n_{k-1} = 1$. Bei den folgenden Rechenoperationen benutzen wir dann auch "intern" die 2-adische Darstellung.

Beachten Sie, dass $k = \lfloor \log_2(n) \rfloor + 1$.

Die Zahl k ist die Länge des Bit-Strings, der die Zahl repräsentiert und wird *Bit-Länge* genannt. Mit einem zusätzlichen Bit kann man auch ganze Zahlen darstellen vom Absolutbetrag $\leq 2^k - 1$ darstellen.

Wir legen fest:

Speicherzugriffszeiten werden bei der Laufzeitanalyse vernachlässigt.

Die Laufzeit messen wir in *Bit-Operationen*. Wie gesagt, gehen wir nicht vollkommen rigoros vor. Die wesentliche Idee ist, dass es sich bei einer Bit-Operation um eine Anwendung eines Operators $\neg, \vee, \wedge$ auf einzelne Bits handelt.

Wir geben uns zwei natürliche Zahlen mit höchstens k Bit vor und wollen diese addieren, multiplizieren sowie den Rest bei der Division bestimmen.

Nun lassen sich zwei Zahlen mit höchstens k Bit in $\mathcal{O}(k)$ Bit-Operationen addieren, wobei die hier verwendete "$\mathcal{O}$-Notation" folgendes bedeutet: Konstante Faktoren vernachlässigt, und die Aussagen werden nur für *genügend große Eingabewerte* getroffen. Z.B. bedeutet die soeben getätigte Aussage:

Es gibt Konstanten $C > 0$ und $K > 0$ so dass für alle $k \geq K$ zwei Zahlen mit höchstens k Bit sich mit $\leq C \cdot k$ Bit-Operationen addieren lassen.

Die Anwendung der $\mathcal{O}$-Notation ist das übliche Vorgehen bei der Laufzeitanalyse. In der impliziten Konstante C verstecken sich alle "Feinheiten" des Rechnermodels.

Mit der "Schulbuchmethode" (Sie erinnern sich?) lassen sich zwei solche Zahlen multiplizieren, indem man höchstens k Additionen von Zahlen mit höchstens $2k$ Bit durchführt. Dies führt zu $\mathcal{O}(k^2)$ Bit-Operationen für die Multiplikation.

Seien nun $m, n \in \mathbb{N}_0$. Wir wollen $p \in \mathbb{N}_0$ und $r \in \{0, 1, \ldots, n-1\}$ mit

$$m = pn + r$$

bestimmen. Die "Schulbuchmethode" terminiert diesmal nach höchstens k Iterationen, wobei jede Iteration aus einer Subtraktion von natürlichen Zahlen mit höchstens k Bit besteht. Dies führt wieder zu $\mathcal{O}(k^2)$ Bit-Operationen.

Wir geben uns nun eine natürliche Zahl n vor und wollen "modulo n" rechnen. Gegeben sind nun $a, b \in \{0, 1, \ldots, n-1\}$, und die Aufgabe ist, die Zahlen $\mathrm{mod}(a + b, n) = a \oplus_n b$ und $\mathrm{mod}(a \cdot b, n) = a \odot_n b$ zu bestimmen. Außerdem wollen wir entscheiden, ob a "modulo n" invertierbar ist und gegebenenfalls ein "Inverses modulo n" bestimmen. D.h. wir wollen überprüfen, ob $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ invertierbar ist und gegebenfalls ein $x \in \{1, \ldots, n-1\}$ mit $[x]_n \cdot [a]_n = [1]_n$ (d.h. $x \odot_n a = 1$) bestimmen.

Die Laufzeitangaben wollen wir dabei für variables n angeben.

Die Laufzeit für die Addition ist dann, wie soeben erläutert, $\mathcal{O}(\log(n))$, und die Laufzeit für die Multiplikation ist $\mathcal{O}(\log(n)^2)$ (zuerst multipliziert man die Zahlen und dann bestimmt man den Rest bei der Division mit n).

Der Test, ob $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ invertierbar ist, kann mit den Euklidischen Algorithmus durchgeführt werden (Berechnung von $\mathrm{ggT}(a, n)$). Jeder Schritt im Euklidischen Algorithmus besteht aus einer Division mit Rest, er benötigt also eine Laufzeit von $\mathcal{O}(\log(n)^2)$. Der Algorithmus terminiert nach $\mathcal{O}(\log(n))$ Schritten (siehe Lemma 1.69). Dies führt zu einer Laufzeit von $\mathcal{O}(\log(n)^3)$.

Das Zurückrechnen besteht auch aus $\mathcal{O}(\log(n)^2)$ Schritten. Es kann "modulo n" erfolgen. Damit benötigen wir hierfür auch nur $\mathcal{O}(\log(n)^3)$ Bit-Operationen.

1.10 Polynome

Sei im Folgenden R ein kommutativer Ring.

Intuitiv ist ein Polynom über R in der Unbestimmten X ein "formaler Ausdruck" der Form $\sum_{i=0}^{d} a_i X^i = a_d X^d + a_{d-1} X^{d-1} \cdots + a_1 X + a_0$ mit $a_i \in R$.

Solche "formalen Ausdrücke" kann man dann addieren und multiplizieren, wobei die üblichen Rechenregeln (Assoziativität, Distributivität, Kommutativität) gelten und X als Unbestimmte aufgefasst wird.

Beispielsweise ist $(1 + X + X^2) \cdot (1 - X) = 1 - X^3$.

Aber inwiefern kann ein "formaler Ausdruck" der angegeben Form ein wohldefiniertes mathematisches Objekt sein? Sicher sollte man zwischen den mathematischen Objekten und seiner symbolischen Darstellung unterscheiden.

Sie erinnern sich, wie wir bei den ganzen und den rationalen Zahlen vorgegangen sind: Wir haben nicht definiert, was eine ganze oder einer rationale Zahl "ist", sondern wir haben die Menge der ganzen bzw. der rationalen Zahlen definiert, und eine ganze bzw. rationale Zahl ist dann ein Element aus dieser Menge. So gehen wir auch hier vor.

Wir haben die folgende Wunschliste: Wir wollen einen kommutativen Ring, genannt $R[X]$, definieren, der ein bestimmtes Element X enthält, so dass

- R ein Unterring von $R[X]$ ist
- sich jedes Element $f \in R[X]$ $(f \neq 0)$ in eindeutiger Weise in der Form $f = \sum_{i=0}^{d} a_i X^i$ mit $a_i \in R$ und $a_d \neq 0$ schreiben läßt.

Wir wollen nun einen solchen Ring definieren.

Beachten Sie, dass die zweite Eigenschaft besagt, dass jedes Polynom eindeutig durch das Tupel $(a_0, \ldots, a_d)$ gegeben ist. Da Polynome beliebig lang werden können, liegt es nahe nicht Tupel sondern "Folgen" mit Werten in R zu betrachten,[10] allerdings mit zwei Modifikationen: Erstens sollte das erste Folgenglied nicht a_1 sondern a_0 heißen. Zweitens sollten wir nur solche Folgen betrachten, für die es nur endlich viele Folgenglieder gibt, die $\neq 0$ sind. (Weil Polynome auch nur endliche viele Terme enthalten.)

Definition und einfache Eigenschaften

Wir beginnen mit der Menge $R^{\mathbb{N}_0}$. Die Elemente hierin sind die Abbildungen $\mathbb{N}_0 \longrightarrow R$. Diese Abbildungen schreiben wir wie bei Folgen üblich in der

[10] In Beispiel 1.8 haben wir Folgen als Abbildungen $\mathbb{N} \longrightarrow \mathbb{R}$ definiert. Mit einer *Folge mit Werten in einem Ring* R meinen wir eine Abbildung $\mathbb{N} \longrightarrow R$.

Form $a = (a_n)_{n\in\mathbb{N}_0}$. Dies ist mit der "komponentenweisen Verknüpfung" eine abelsche Gruppe (siehe den Unterabschnitt über Produkte in Abschnitt 1.5).

Ferner definieren wir für $r \in R$ und $a = (a_n)_{n\in\mathbb{N}_0} \in R^{\mathbb{N}_0}$: $r{\cdot}a := (r{\cdot}a_n)_{n\in\mathbb{N}_0}$.

Für $i, j \in \mathbb{N}_0$ definieren wir das so genannte *Kronecker-Delta* durch

$$\delta_{i,j} := \begin{cases} 1 \text{ falls } i = j \\ 0 \text{ falls } i \neq j \end{cases} ,$$

und wir definieren $e_j := (\delta_{i,j})_{i\in\mathbb{N}_0} \in R^{\mathbb{N}_0}$. Wir betrachten nun die folgende Teilmenge von $R^{\mathbb{N}_0}$:

$$P := \{a = (a_n)_{n\in\mathbb{N}_0} \in R^{\mathbb{N}_0} \mid \text{ es gibt nur endlich viele } n \in \mathbb{N}_0 \text{ mit } a_n \neq 0\}$$

Dies ist eine Untergruppe von $R^{\mathbb{N}_0}$. (Warum?) Auf dieser Gruppe P wollen wir eine Multiplikation definieren und ein Element $X \in P$ identifizieren, so dass P zusammen mit X die gewünschten Eigenschaften hat.

Beachten Sie: Wenn $a = (a_n)_{n\in\mathbb{N}} \in P$, dann gibt es ein $d \in \mathbb{N}_0$ so dass $a_n = 0$ für $n > d$. Mit so einem d gilt dann

$$a = \sum_{n=0}^{d} a_n e_n \ . \tag{1.8}$$

Dies schreiben wir auch in der Form

$$a = \sum_{n\in\mathbb{N}_0} a_n e_n \ , \tag{1.9}$$

wobei zu beachten ist, dass in dieser Summe immer nur endlich viele Terme $\neq 0$ sind.

Wir definieren nun eine Multiplikation auf P durch

$$(a_n)_{n\in\mathbb{N}_0} \cdot (b_n)_{n\in\mathbb{N}_0} := (\sum_{i=0}^{n} a_i b_{n-i})_{n\in\mathbb{N}_0} \ . \tag{1.10}$$

Beachten Sie, dass das Ergebnis wieder in P liegt, da nur endlich viele a_n und b_n von 0 verschieden sind.

Beachten Sie weiter: Wenn wir (1.8) in (1.10) einsetzen, erhalten wir

$$(\sum_{n\in\mathbb{N}_0} a_n e_n) \cdot (\sum_{n\in\mathbb{N}_0} b_n e_n) = \sum_{n\in\mathbb{N}_0} (\sum_{i=0}^{n} a_i b_{n-i}) e_n \tag{1.11}$$

$$= \sum_{k\in\mathbb{N}_0} \sum_{\ell\in\mathbb{N}_0} a_k b_\ell \, e_{k+\ell} \tag{1.12}$$

Insbesondere ist

$$e_k \cdot e_\ell = e_{k+\ell} \ , \tag{1.13}$$

für alle $i, j \in \mathbb{N}$, und hieraus folgt insbesondere

$$e_1^n = e_n$$

für alle $n \in \mathbb{N}$. Außerdem sieht man, dass e_0 ein neutrales Element bezüglich der Multiplikation ist. Wie üblich setzen wir also $1 = 1_P := e_0$.

Wir setzen nun $X := e_1$. Dann ist also $e_n = X^n$. Wenn wir dies in (1.9) einsetzen, erhalten wir

$$a = \sum_{n \in \mathbb{N}_0} a_n X^n$$

für alle $a \in P$.

Seien nun $a, b \in P$. Dann ist

$$a \cdot b = (\sum_{n \in \mathbb{N}_0} a_n X^n) \cdot (\sum_{n \in \mathbb{N}_0} a_n X^n) =$$

$$\sum_{n \in \mathbb{N}_0} (\sum_{i=0}^{n} a_i b_{n-i}) X^n = \sum_{k \in \mathbb{N}_0} \sum_{\ell \in \mathbb{N}_0} a_k b_k X^{k+\ell}$$

nach (1.11). Man sieht leicht, dass die Distributivgesetze gelten. Damit ist P ein Ring. Außerdem ist $\iota : R \hookrightarrow P, r \mapsto re_0$ ein injektiver Ringhomomorphismus.

Damit ist P zusammen mit X ein Ring wie gewünscht; wir setzen also $R[X] := P$. □

Sei nun $f = f(X) = \sum_{i=0}^{d} a_i X^i \in R[X]$ ein Polynom. Sei $r \in R$. Dann können wir r in p "einsetzen" bzw. – was das selbe ist – p an r "auswerten". Wir erhalten

$$f(r) := \sum_{i=0}^{d} a_i r^i \in R \ .$$

Wenn wir nun r variieren, erhalten wir eine Abbildung

$$R \longrightarrow R \ , \ r \mapsto f(r) = \sum_{i=0}^{d} a_i r^i \ .$$

Man kann zeigen, dass diese Abbildung ein Ringhomomorphismus ist (Übungsaufgabe). Diese Abbildung heißt die *Polynomfunktion* zu R. Aus der Schule kennen Sie die Polynomfunktionen $\mathbb{R} \longrightarrow \mathbb{R}$ (die wahrscheinlich "Polynome" genannt wurden).

Wenn wir nun die Polynome variieren, erhalten wir eine Abbildung

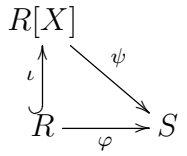
$$R \longrightarrow \mathrm{Abb}(R, R)\ ,\ f(X) \mapsto (r \mapsto f(r))\ .$$

Wenn z.B. $R = \mathbb{R}$ (aber auch $R = \mathbb{Z}$), ist diese Abbildung injektiv, d.h. ein Polynom ist durch die entsprechende Polynomfunktion eindeutig bestimmt (ohne Beweis). Dies ist aber *nicht* immer der Fall. Sei z.B. $R = \mathbb{Z}/2\mathbb{Z} = \mathbb{F}_2$, und sei $f_1(X) := 0, f_2(X) := X^2 + X$. Nun ist sowohl die Polynomfunktion zu $f_1(X)$ als auch die Polynomfunktion zu $f_2(X)$ die Null-Abbildung, aber es ist $f_1(X) \neq f_2(X)$.

Begründung: Es ist klar, dass die Polynomfunktion zu $f_1(X)$ durch $r \mapsto 0$ gegeben ist. Zu $p_2(X)$: Der Körper $\mathbb{F}_2$ hat nur zwei Elemente, $0 = [0]_2$ und $1 = [1]_2$, und es ist $f_2(0) = 0 + 0 = 0$ und $f(1) = 1 + 1 = 0$.

Sei weiterhin $f(X) = \sum_{i=0}^{d} a_i X^i \in R[X]$, sei S ein kommutativer Ring, und sei $\varphi : R \longrightarrow S$ ein Homomorphismus von Ringen und $s \in S$. Wir definieren $f(s) := \sum_{i=0}^{d} \varphi(a_i) s^i \in S$. Man kann zeigen:

Aussage 1.70 *Die Abbildung $\psi : R[X] \longrightarrow S, : f(X) \mapsto f(s)$ ist ein Homomorphismus von Ringen, und es ist der einzige Homomorphismus $\psi : R[X] \longrightarrow S$ für den $\psi(X) = s$ gilt und das Diagramm*



kommutativ ist.

(Übungsaufgabe)

Definition Sei $f(X) = \sum_{i=0}^{d} a_i X^i$ mit $a_d \neq 0$. Dann heißt d der *Grad* von $f(X)$, Bezeichnung $\mathrm{Grad}(f(X))$. Der Grad des Nullpolynoms wird mit $-\infty$ definiert.

Ein Polynom $f(X) = \sum_{i=0}^{d} a_i X^i$ mit $a_d = 1$ heißt *normiert*.

Polynome über Körpern

Sei ab nun $R = K$ ein Körper.

Definition Seien $a(X), b(X) \in K[X]$. Dann sagen wir, dass $a(X)$ das Polynom $b(X)$ *teilt* und schreiben $a(X)|b(X)$, wenn es ein Polynom $c(X) \in K[X]$ mit $b(X) = a(X) \cdot c(X)$ gibt.

Definition Ein Polynom $f(X) \in K[X]$ mit $\mathrm{Grad}(f(X)) \geq 1$ heißt *irreduzibel*, wenn es nicht in der Form $f(X) = f_1(X) \cdot f_2(X)$ mit $f_1(X), f_2(X) \in K[X]$ und $\mathrm{Grad}(f(X)) \geq 1$ geschrieben werden kann.

Aus der Schule kennen sie die *Polynomdivision*, die größe Ähnlichkeit mit der Division mit Rest von ganzen Zahlen hat. Auf Grundlage der Polynomdivision kann man beweisen:

Aussage 1.71 *Seien $a(X), b(X) \in K[X]$ zwei Polynome mit $b(X) \neq 0$. Dann gibt es eindeutig bestimmte Polynome $p(X), r(X) \in K[X]$ mit $\mathrm{Grad}(r(X)) < \mathrm{Grad}(b(X))$ und*

$$a(X) = p(X) \cdot b(X) + r(X) \; .$$

Mittels dieser Aussage kann man viele Aussagen über den Ring $\mathbb{Z}$ auf den Ring $K[X]$ "übertragen" (wobei die normierten Polynome die Rolle der natürlichen Zahlen und die irreduziblen Polynome die Rolle der Primzahlen einnehmen).

Wir definieren:

Definition Seien $a_1(X), a_2(X), \ldots, a_k(X) \in K[X]$, nicht alle $= 0$. Das normierte Polynom mit maximalem Grad $g(X)$ mit $g(X)|a_1(X)$, $g(X)|a_2(X), \ldots, g(X)|a_k(X)$ heißt der *größte gemeinsame Teiler* von $a_1(X), \ldots, a_k(X)$, Bezeichnung: $\mathrm{ggT}(a_1(X), \ldots, a_k(X))$.

Analog zu den ganzen Zahlen haben wir den Satz:

Satz 1.3 *Seien $a_1(X), a_2(X), \ldots, a_k(X) \in K[X]$, nicht alle $= 0$, und sei $b(X) \in K[X]$ mit $b(X) \,|\, a_1(X), \ldots, b(X) \,|\, a_k(X)$. Dann gilt $b(X) \,|\, \mathrm{ggT}(a_1(X), \ldots, a_k(X))$.*

Analog zu Satz 1.1 folgt dieser Satz aus dem folgenden Lemmata:

Lemma 1.72 *Sei I ein Ideal in $K[X]$. Dann gilt entweder $I = \{0\}$ oder es gibt ein (eindeutig bestimmtes) normiertes Polynom $f(X) \in K[X]$ mit $I = (f(X))$.*

Lemma 1.73 *Seien $a_1(X), \ldots, a_k(X) \in K[X]$, nicht alle $= 0$, und sei $f(X) \in K[X]$ normiert mit*

$$(a_1(X), \ldots, a_k(X)) = (f(X)) \; .$$

Dann gilt $f(X)|a_i(X)$ für alle $i = 1, \ldots, k$.

Für $b(X) \in K[X]$ mit $b|a_i$ für alle $i = 1, \ldots, k$ gilt $b(X)|f(X)$.

Insbesondere ist $f(X) = \mathrm{ggT}(a_1(X), \ldots, a_k(X))$, und Satz 1.3 ist richtig.

Der Beweis dieser beiden Lemmata ist analog zum Beweis von Lemmata 1.65 und 1.66 (Übungsaufgaben!).

Analog zur Situation in $\mathbb{Z}$ haben wir die folgende Umformulierung dieses Lemmas: Seien $a_1(X), \ldots, a_k(X) \in \mathbb{Z}$. Dann ist $\mathrm{ggT}(a_1(X), \ldots, a_k(X))$ das eindeutig bestimmte normierte Polynom mit

$$(a_1(X), \ldots, a_k(X)) = (\mathrm{ggT}(a_1(X), \ldots, a_k(X))) \; .$$

Insbesondere gibt es also Polynome $z_1(X), \ldots, z_k(X) \in K[X]$ mit

$$\mathrm{ggT}(a_1, \ldots, a_k) = z_1(X)a_1(X) + \cdots + z_k(X)a_k(X) \; .$$

Sei nun $f(X) = \sum_{i=0}^{d} a_i X^i \in K[X]$ mit $d \geq 1$ und $a_d \neq 0$. Analog zur Relation "mod n auf $\mathbb{Z}$ für ein $n \in \mathbb{Z}$ betrachten wir nun die Relation "mod $f(X)$" auf $K[X]$. D.h. für $a(X), b(X) \in K[X]$ soll gelten:

$$a(X) \equiv b(X) \mod f(X) :\Longleftrightarrow f(X) \mid (a(X) - b(X)) \; .$$

Es sei $f(X))$ das von $f(X)$ erzeugte Ideal in $K[X]$. Dann sind quivalent:

- $a(X) \equiv b(X)$
- $a(X) - b(X) \in (f(X))$
- $a(X) \sim_{(f(X))} b(X)$
- $[a(X)]_{(f(X))} = [b(X)]_{(f(X))} \in K[X]/(f(X))$

Die Relationen "mod $f(X)$" und $\sim_{(f(X))}$ sind also identisch. In der letzten Äquivalenz ist $K[X]/(f(X))$ der in Abschnitt 1.6 definierte Faktorring von $K[X]$ nach $(f(X))$. Man beachte insgesamt die Analogie zu Beispiel 1.47!

Wir haben den Homomorphismus von Ringen

$$K \hookrightarrow K[X]/(f(X)) \, , \; c \mapsto [c]_{(f(X))} \; ,$$

der offensichlich injektiv ist. Wir "identifizieren" K mit seinem Bild, d.h. wir schreiben c statt $[c]_{(f(X))}$. Dann ist also $[\sum_{i=0}^{e} c_i X^i]_{(f(X))} = \sum_{i=0}^{e} c_i [X]^i_{(f(X))}$ (mit $c_i \in K$).

Beachten Sie, dass nach Lemma 1.71 die Menge der Polynome in $K[X]$ vom Grad $< d$ ein Repräsentantensystem bezüglich der Relation $\sim_{(f(X))}$ bildet. Mit anderen Worten: $K[X]/(f(X))$ besteht genau aus den Restklassen $[a(X)]_{(f(X))}$ mit $\mathrm{Grad}(a(X)) < d$, und alle diese Klassen sind verschieden. Oder wiederum anders ausgedrückt:

Lemma 1.74 *Jedes Element in $K[X]/f(X)$ lässt sich in eindeutiger Weise in der Form*

$$\sum_{i=0}^{d-1} c_i [X]^i_{(f(X))}$$

mit $c_i \in K$ darstellen.

Aussage 1.68 hat nun das folgende Analogon:

Aussage 1.75 *Sei $f(X) \in K[X]$ mit $\mathrm{Grad}(f(X)) \geq 1$, und sei $a(X) \in K[X]$. Dann ist $[a(X)]_{(f(X))} \in K[X]/(f(X))$ genau dann invertierbar, wenn $\mathrm{ggT}(a(X), f(X)) = 1$.*

Wir erhalten den folgenden wichtigen Satz:

Satz 1.4 *Sei $f(X) \in K[X]$ irreduzibel. Dann ist $K[X]/(f(X))$ ein Körper.*

Dieser Satz hat eine ganz zentrale Bedeutung in der Algebra, und zwar insbesondere aufgrund des Folgenden:

Sei $f(X) = \sum_{i=0}^{d} a_i X^i \in K[X]$ irreduzibel. Dann hat $f(X)$ insbesondere keine Nullstellen in K, d.h. es gibt kein $\alpha \in K$ mit $f(\alpha) = 0$ (sonst würde $(X - \alpha) | f(X)$ gelten (Übungsaufgabe)).

Wir wollen nun einen größeren Körper schaffen, in dem so eine Nullstelle α vorhanden ist. Außerdem sollen alle Elemente des größeren Körpers die Form $\sum_{i=0}^{e} c_i \alpha^i$ mit $e \in \mathbb{N}$ und $c_i \in K$ haben.

Wir haben so einen Körper, nämlich $K[X]/(f(X))$!

Sei hierzu $\alpha := [X]_{(f(X))}$. Dann ist

$$f(\alpha) = f([X]_{(f(X))}) = \sum_{i=0}^{d} a_i [X]^i_{(f(X))} =$$

$$[\sum_{i=0}^{d} a_i X]_{(f(X))} = [f(X)]_{(f(X))} = [0]_{(f(X))} = 0 \in K[X]/(f(X)) \; .$$

In der Tat kann man jedes Element in $K[X]/(f(X))$ in eindeutiger Weise in der Form $\sum_{i=0}^{d-1} c_i \alpha^i$ mit $c_i \in K$ darstellen (siehe Lemma 1.74).

Es folgen einige illustrative Beispiele. Doch zunächst:

Lemma 1.76 *Sei $f(X) \in K[X]$ vom Grad 2 oder 3 (diese Voraussetzung ist wichtig!). Dann ist $f(X)$ genau dann irreduzibel, wenn es keine Nullstelle in K hat.*

Der *Beweis* ist einfach (Übungsaufgabe).

Satz 1.5 *Es gibt keine rationale Zahl α mit $\alpha^2 = 2$.*

Kennen Sie einen Beweis dieser Aussage? Hier ist einer (durch Widerspruch):

Nehmen wir an, es gäbe so ein a. Wir schreiben a in der Form $\alpha = \frac{z}{n}$ mit $z, n \in \mathbb{Z}$, wobei $\mathrm{ggT}(z, n) = 1$. Dann ist also $(\frac{z}{n})^2 = 2$, also $z^2 = 2n^2$. Damit gilt $2|z^2$. Somit gilt auch $2|z$. (Dies folgt aus Satz 1.2, man kann es aber auch "elementar" zeigen.) Wir haben also $z = 2b$ für ein $b \in \mathbb{Z}$. Daraus folgt $4b^2 = 2n^2$. Und hieraus folgt $2b^2 = n^2$, also $2|n^2$. Hieraus folgt wiederum $2|n$. Also sind z und n nicht teilerfremd. □

Beispiel 1.77 Sei $f(X) := X^2 - 2 \in \mathbb{Q}[X]$. Dann hat $f(X)$ keine Nullstellen in $\mathbb{Q}$ nach obigem Lemma, und da es den Grad 2 hat, ist es somit irreduzibel.

Sei wie oben $\alpha := [X]_{(f(X))}$. Dann ist $\alpha^2 - 2 = 0$, also $\alpha^2 = 2$, und jedes Element von $\mathbb{Q}[X]/(X^2 - 2)$ hat eine eindeutige Darstellung der Form $c + d\alpha$ mit $c, d \in \mathbb{Q}$.

Den Körper $\mathbb{Q}[X]/(X^2 - 2)$ bezeichnet man mit $\mathbb{Q}[\sqrt{2}]$, man sagt "$\mathbb{Q}$ adjungiert $\sqrt{2}$".

Beispiel 1.78 Sei $f(X) := X^2 + 1 \in \mathbb{R}[X]$. Dann hat $f(X)$ wiederum keine Nullstellen (es gibt keine reelle Zahl α mit $\alpha^2 = -1$), ist also wiederum irreduzibel.

Sei nun $i := [X]_{(f(X))}$. Dann ist $i^2 + 1 = 0$, also $i^2 = -1$, und jedes Element von $\mathbb{R}[X]/(X^2 + 1)$ hat eine eindeutige Darstellung der Form $r + is$ mit $r, s \in \mathbb{R}$.

Der Körper $\mathbb{R}[X]/(X^2 + 1)$ heißt der Körper der *komplexen Zahlen* und wird mit $\mathbb{C}$ bezeichnet. Das Element i heißt *imaginäre Einheit*.

Sicher haben Sie von den komplexen Zahlen schon gehört, und vielleicht kennen Sie auch den

Fundamentalsatz der Algebra Jedes nicht-konstante Polynom $f(X) \in \mathbb{C}[X]$ hat eine Nullstelle (in $\mathbb{C}$).

Dieser Satz wurde zuerst von Carl Friedrich Gauß (1777-1855) bewiesen. Ein Beweis dieses Satzes würde uns zu weit wegführen. Wir geben noch das folgende einfache Korollar des Fundamentalsatzes an:

Korollar 1.79 *Jedes nicht-konstante Polynom $f(X) \in \mathbb{C}[X]$ zerfällt in Linearfaktoren, d.h. es gibt $\alpha_1, \ldots, \alpha_d \in \mathbb{C}$ (mit $d = \mathrm{Grad}(f(X))$ und $c \in \mathbb{C}$ mit*

$$f(X) = c \cdot (X - \alpha_1) \cdots (X - \alpha_d) \; .$$

Beispiel 1.80 Sei $f(X) := X^2 + X + 1 \in \mathbb{F}_2[X]$. Wiederum ist $f(X)$ irreduzibel (warum?). Sei $\alpha := [X]_{(f(X))}$. Dann kann man jedes Element in $\mathbb{F}_2[X]/(f(X))$ in eindeutiger Weise in der Form $c + d\alpha$ mit $c, d \in \mathbb{F}_2$ darstellen.

Wir haben also einen Körper mit 4 Elementen.

Beachten Sie, dass hingegen $\mathbb{Z}/4\mathbb{Z}$ kein Körper ist!

Man kann zeigen, dass für jede Primzahl p und jede natürliche Zahl d ein irreduzibles Polynom $f(X) \in \mathbb{F}_p[X]$ vom Grad d existiert. Der Körper $\mathbb{F}_p[X]/(f(X))$ hat dann p^d Elemente.

Außerdem kann man zeigen: Wenn K ein endlicher Körper ist (d.h. $\#K$ ist endlich), dann ist $\#K$ eine Potenz einer Primzahl (siehe Aussage 2.50), und wenn zwei endliche Körper gleich viele Elemente enthalten, dann sind sie isomorph (das beweisen wir nicht).

Es gibt also zu jeder Primpotenz (=Potenz einer Primzahl) "im wesentlichen einen" endlichen Körper, und weitere endliche Körper gibt es nicht. "Der" endliche Körper mit q Elementen (q eine Primpotenz) wird mit $\mathbb{F}_q$ bezeichnet.

Die endlichen Körper, insbesondere die Körper der Form $\mathbb{F}_{2^d}$, spielen in der Informatik eine wichtige Rolle.

Auch den Euklidischen Algorithmus kann man leicht von ganzen Zahlen auf Polynome übertragen. Gegeben $a(X), b(X) \in K[X]$ kann man also $\mathrm{ggT}(a(X), b(X))$ sowie zwei Polynome $c(X), d(X)$ mit $c(X)a(X)+d(X)b(X) = \mathrm{ggT}(a(X), b(X))$ ausrechnen. Hierbei setzen wir natürlich voraus, dass man in K rechnen kann!

Komplexitäten

Wir überlegen uns nun, wie schnell man mit Polynomen rechnen kann. Da K beliebig ist, ist es natürlich, die Komplexitäten (Laufzeiten) in Körperoperationen (d.h. in der Anzahl der benötigten Additionen, Subtraktionen, Multiplikationen oder Divisionen in K) anzugeben.

Gegeben seien zwei Polynome vom Grad $\leq d$. Die Komplexitäten werden in Abhängigkeit von d angegeben.

Nun kann man in $\mathcal{O}(d)$ Körperoperationen die Polynome addieren und in $\mathcal{O}(d^2)$ Körperoperationen die Polynome multiplizieren. Auch die Polynomdivision benötigt $\mathcal{O}(d^2)$ Körperoperationen.

Der Euklidische Algorithmus terminiert nach höchstens d Schritten, da bei jedem Schritt der Grad des zweiten Polynoms abnimmt. Jeder dieser Schritte benötigt $\mathcal{O}(d^2)$ Körperoperationen. Also kann man in $\mathcal{O}(d^3)$ Körperoperationen den ggT der zwei Polynome berechnen.

Gegeben sei nun ein Polynom $f(X) \in K[X]$ vom Grad d. Wir fragen uns, wie schnell man (in Abhängigkeit von d) in dem Ring $K[X]/(f(X))$ rechnen kann. Beachten Sie, dass diese Frage analog zur Frage ist, wie schnell man im Ring $\mathbb{Z}/n\mathbb{Z}$ rechnen kann.

Die Elemente von $K[X]/(f(X))$ werden hierbei durch Polynome mit Grad $< d$ dargestellt.

Zum Beispiel bedeutet Addition dann das Folgende: Gegeben zwei Polynome $a(X), b(X) \in K[X]$ mit $\mathrm{Grad}(a(X)), \mathrm{Grad}(b(X)) < d$, bestimme das eindeutig bestimmte Polynom $c(X)$ vom Grad $< d$ mit $(a(X)+b(X)) \equiv c(X)$ mod $f(X)$, d.h. $f(X)|(a(X)+b(X)-c(X))$. Analoges gilt für die Multiplikation.

Zur Addition in $K[X]/(f(X))$ werden die Polynome addiert. (Es muss keine weitere Operation ausgeführt werden, da $a(X) + b(X)$ wieder Grad $< d$ hat.) Wir erhalten also eine Komplexität von $\mathcal{O}(d)$ Körperoperationen.

Zur Multiplikation kann man so vorgehen: Es werden die beiden Polynome multipliziert, und dann wird die Polynomdivision mit $f(X)$ ausgeführt. Beachten Sie, dass $\mathrm{Grad}(a(X) \cdot b(X)) \leq 2d - 2$ ist. Damit benötigt man sowohl für die Multiplikation als auch für die anschließenden Polynomdivision $\mathcal{O}(d^2)$ Körperoperationen.

Besser ist dies: Es sei $a(X) = \sum_{i=0}^{d-1} a_i X^i$. Dann wird der Reihe nach $[X]^i_{(f(X))} \cdot [b(X)]_{(f(X))}$ berechnet (d.h. es werden die eindeutigen Repräsentaten mit Grad $\leq d-1$ berechnet, wobei immer vom schon Berechneten ausgegangen wird), und dann wird daraus $[a(X) \cdot b(X)]_{(f(X))}$ als $\sum_{i=0}^{d-1} a_i [X]^i_{(f(X))} \cdot [b(X)]_{(f(X))}$ berechnet.

Analog zu $\mathbb{Z}/n\mathbb{Z}$ kann man auch den Euklidischen Algorithmus verwenden, um zu überprüfen, ob Elemente in $K[X]/(f(X))$ invertierbar sind und ggf. das Inverse berechnen.

Hierzu sei ein $a(X) \in K[X]$ vom Grad $< d$ gegeben. Wir wollen wissen, ob $[a(X)]_{(f(X))} \in K[X]/(f(X))$ invertierbar ist und ggf. ein Polynom $c(X) \in K[X]$ vom Grad $< d$ mit $[a(X)]_{(f(X))} \cdot [c(X)]_{(f(X))} = [1]_{(f(X))}$ berechnen.

Hierzu gehen wir wie in $\mathbb{Z}/n\mathbb{Z}$ vor: Zunächst bestimmen wir mit dem Euklidischen Algorithmus den ggT von $a(X)$ und $f(X)$. Nun ist $[a(X)]_{(f(X))} \in K[X]/(f(X))$ genau dann invertierbar, wenn dieser ggT gleich 1 ist (warum?). Dann berechnen wir mit dem erweiterten Euklidischen Algorithmus (das eindeutig bestimmte) $c(X)$ so dass es ein $d(X)$ mit $c(X)a(X) + d(X)f(X) = 1$ gibt. (Dieses $c(X)$ ist das gesuchte Polynom.) Während dieser Rechnung stellen wir sicher, dass nie Polynome vom Grad $\geq 2d$ vorkommen. (Dies geht, weil wir "modulo $f(X)$" rechnen können.) Diese Berechnungen benötigen dann $\mathcal{O}(d^3)$ Körperoperationen.

Wir überlegen uns nun, wie schnell man mit Polynomen über $\mathbb{F}_2$ rechnen

kann. Beachten Sie, dass ein Polynom vom Grad d über $\mathbb{F}_2$ durch einen Bit-String der Länge $d+1$ dargestellt werden kann.

Über $\mathbb{F}_2$ ist eine Körperoperation das Gleiche wie eine Bit-Operation. Deshalb erhalten wir:

Gegeben zwei Polynome vom Grad $\leq d$ kann man diese in einer Zeit von $\mathcal{O}(d)$ addieren und in einer Zeit von $\mathcal{O}(d^2)$ multiplizieren. Die Polynomdivision benötigt $\mathcal{O}(d^2)$ Bit-Operationen. Der ggT kann in $\mathcal{O}(d^3)$ Schritten ausgerechnet werden.

Wie schon angedeutet spielen die Körper $\mathbb{F}_{2^d}$ in der Informatik eine besondere Rolle, und wir diskutieren deshalb, wie schnell man in diesen Körpern rechnen kann. Sei dazu $f(X) \in \mathbb{F}_2[X]$ irreduzibel vom Grad d. Die Aufgabe besteht genauer darin, in $\mathbb{F}_2[X]/(f(X))$ zu rechnen, wobei man mit Polynomen über $\mathbb{F}_2$ vom Grad $\leq d$, die die Klassen in $\mathbb{F}_2[X]/(f(X))$ repräsentieren, rechnet.

Nach den obigen Ergebnissen erhalten wir:

Für die Addition benötigen wir $\mathcal{O}(d)$ Bit-Operationen. Die Multiplikation von zwei Polynomen über $\mathbb{F}_2$ vom Grad $\leq d$ kann in $\mathcal{O}(d^2)$ Bit-Operationen erfolgen. Mit dem Euklidischen Algorithmus kann man in $\mathcal{O}(d^3)$ Bit-Operationen das Inverse eines Elements ausrechnen.

Beachten Sie hier, dass wir für das Rechnen in $\mathbb{F}_{2^d}$ und $\mathbb{F}_p$ (p eine Primzahl) (bis auf Konstanten) die gleichen Laufzeiten erhalten, wenn wir die Laufzeiten als Funktion der Bit-Länge betrachten (vgl. Abschnitt 1.9). (Genauer gesagt: Man wir haben bis auf Konstanten die gleichen oberen Schranken für die Laufzeiten erhalten.)

Kapitel 2

Lineare Algebra

2.1 Lineare Gleichungssysteme und lineare Unterräume

Wir legen im Folgenden immer einen festen Körper zugrunde, den wir mit K bezeichnen. Die Elemente von K werden auch *Skalare* genannt.

Die Elemente aus K^n (die "n-Tupel") werden von nun ab als "Spaltenvektoren" geschrieben, beispielsweise so:

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Dementsprechend nennen wir die Elemente von K^n auch *Vektoren*.

Man definiert eine Addition von Vektoren in K^n komponentenweise:

$$+ : K^n \times K^n \longrightarrow K^n , \quad (\underline{x}, \underline{y}) \mapsto \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Außerdem definiert man eine so genannte *Skalarmultiplikation*:

$$\cdot : K \times K^n \longrightarrow K^n , \quad (a, \underline{x}) \mapsto \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}$$

Definition Wir setzen $\underline{0} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$; dies ist der *Nullvektor*. Außerdem setzen wir

$$\underline{e}_i := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i - \text{te Zeile ;}$$

dies ist der i-te *Standardvektor*.

Bemerkung Die Addition ist also eine Verknüpfung auf K^n. Die Skalarmultiplikation ist hingegen keine Verknüpfung auf K^n. Wir wissen bereits, dass K^n mit der Addition eine abelsche Gruppe mit neutralem Element $\underline{0}$ ist (siehe Unterabschnitt über Produkte in Abschnitt 1.5).

Ein *lineares Gleichungssystem (LGS)* über K ist ein Gleichungssystem der Form

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & b_1 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & b_2 \\ \vdots \qquad\qquad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & b_m \end{array} \tag{2.1}$$

mit $a_{i,j} \in K$ und den "Unbestimmten" $X_1, \ldots, X_n$.

Wir wollen die Lösungsmenge so eines Gleichungssystems explizit bestimmen.

Zunächst eine Definition.

Definition Ein lineares Gleichungssystem der Form

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & 0 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & 0 \\ \vdots \qquad\qquad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & 0 \end{array} \tag{2.2}$$

(d.h. mit "trivialer rechter Seite") heißt *homogenes lineares Gleichungssystem*. Ein lineares Gleichungssystem wie (2.1) heißt auch *inhomogenes lineares Gleichungssystem*.

Wenn man ein konkretes inhomogenes System (2.1) zugrundelegt und die rechte Seite durch Nullen ersetzt, erhält man das *zugehörige homogene lineare Gleichungssystem.*

Bemerkung Ein "inhomogenes System" kann auch eine "triviale rechte Seite" haben, d.h. ein "homogenes System" ist ein Spezialfall eines inhomogenen Systems.

Notation Gegeben ein LGS, wird die Lösungsmenge mit $\mathbb{L}$ bezeichnet. Die Lösungsmenge des zugehörigen homogenen LGS wird mit $\mathbb{L}_h$ bezeichnet.

Notation Sei $A \subseteq K^n$ eine Teilmenge und $\underline{x} \in K^n$. Dann definieren wir

$$\underline{x} + A := \{\underline{x} + \underline{a} \mid \underline{a} \in A\} .$$

Bemerkung Man kann (insbesondere für $K = \mathbb{R}$ und $n = 2$ oder $n = 3$) die Menge $\underline{x} + A$ als eine Parallelverschiebung von A interpretieren.

Aussage 2.1 *Sei ein LGS (2.1) gegeben. Wir nehmen an, dass das LGS lösbar ist und fixieren eine Lösung $\underline{x}_0$. Dann gilt*

$$\mathbb{L} = \underline{x}_0 + \mathbb{L}_h .$$

Der *Beweis* ist einfach.

Bemerkung Eine Lösung $\underline{x}_0$ wie in der Aussage heißt auch "spezielle Lösung". Man sagt: Man erhält alle Lösungen eines inhomogenen LGS, indem man eine spezielle Lösung sowie alle Lösungen des zugehörigen homogenen LGS berechnet.

Definition Ein *linearer Unterraum* von K^n ist eine Teilmenge U von K^n mit den folgenden Eigenschaften.

- $\underline{0} \in U$
- $\forall \underline{x}, \underline{y} \in U : \underline{x} + \underline{y} \in U$
- $\forall \underline{x} \in U \, \forall a \in K : a\underline{x} \in U$

Bemerkung Ein linearer Unterraum ist also insbesondere eine Untergruppe von $(K^n, +)$.

Aussage 2.2 *Die Lösungsmenge eines homogenen LGS ist ein linearer Unterraum.*

Der *Beweis* ist wiederum einfach.

Definition Sei A eine Teilmenge von K^n. Dann ist A ein *affiner Unterraum* von K^n falls es ein $\underline{x}_0 \in K^n$ und einen linearen Unterraum U von K^n mit $A = \underline{x}_0 + U$ gibt.

Lemma 2.3 *Seien $\underline{x}_0, \underline{y}_0 \in K^n$ und $U, V \subseteq K^n$ lineare Unterräume mit $\underline{x}_0 + U = \underline{y}_0 + V$. Dann ist $U = V$ und $\underline{x}_0 - \underline{y}_0 \in U$.*

Beweis. Es ist $(\underline{y}_0 - \underline{x}_0) + V = U$. Da $\underline{0} \in V$ ist $\underline{y}_0 - \underline{x}_0 \in U$ (und somit auch $\underline{x}_0 - \underline{y}_0 \in U$). Analog zeigt man, dass auch $\underline{x}_0 - \underline{y}_0 \in V$.

Sei nun $\underline{u} \in U$. Dann ist $\underline{x}_0 - \underline{y}_0 + \underline{u} \in V$. Somit ist auch $\underline{u} = \underline{x}_0 - \underline{y}_0 + \underline{u} - (\underline{x}_0 - \underline{y}_0) \in V$. Wir haben gesehen, dass $U \subseteq V$. Analog zeigt man, dass $V \subseteq U$. □

Definition Sei A ein affiner Raum mit $A = \underline{x}_0 + U$. Dann heißt U der *zu A gehörige lineare Unterraum*. Dieser Unterraum wird mit U_A bezeichnet.

Beachten Sie, dass nach dem obigen Lemma dieser lineare Unterraum nur von A abhängt.

Der Zusammenhang mit linearen Gleichungssystemen ist durch die folgende Aussage gegeben.

Aussage 2.4 *Wenn die Lösungsmenge $\mathbb{L}$ eines inhomogenen LGS nicht-leer ist, dann ist sie ein affiner Unterraum. In diesem Fall ist die Lösungsmenge $\mathbb{L}_h$ des zugehörigen homogenen LGS der zu diesem affinen Unterraum gehörige lineare Unterraum.*

Dies folgt aus Aussage 2.2 und Aussage 2.1. □

Definition Ein *(endliches) System von Vektoren* in K^n ist ein Tupel von Vektoren in K^n, einschließlich des *leeren Tupels*.

Ich erinnere hier an die Definition von *Tupel von Elementen einer Menge*. Sei X eine beliebige Menge, und sei $r \in \mathbb{N}$. Dann ist ein *r-Tupel* auf X eine Abbildung $\{1, \ldots, r\} \longrightarrow X$. In diesem Sinne definieren wir das *leere Tupel* (0-Tupel) als die eindeutige Abbildung $\emptyset \longrightarrow X$ (dies wurde schon in Fußnote 7 angeschnitten).

Ein System von Vektoren schreiben wir auch als $\underline{x}_1, \ldots, \underline{x}_r$ anstatt $(\underline{x}_1, \ldots, \underline{x}_r)$. Ich betone, dass dann $r = 0$ *immer zugelassen* ist, womit gemeint ist, dass keine Vektoren gegeben sind.

Wenn wir alle Elemente eines solchen Systems von Vektoren $\underline{x}_1, \ldots, \underline{x}_r$ $(r \geq 1)$ aufaddieren, erhalten wir den Vektor $\sum_{i=1}^{r} \underline{x}_i$. Wir definieren die Summe über das *leere System* als $\underline{0}$.

(Dies ist u.a. durch das Folgende gerechtfertigt: Wenn nun $\underline{x}_1, \ldots, \underline{x}_r$ ein System ist und $\underline{x}$ ein beliebiger Vektor ist, ist immer (auch für $r = 0$) $\underline{x}_1 + \cdots + \underline{x}_r + \underline{x}$ die Summe der Vektoren des Systems $\underline{x}_1, \ldots, \underline{x}_r, \underline{x}$.

Definition Sei $U \subseteq K^n$ ein linearer Unterraum, und seien $\underline{x}_1, \ldots, \underline{x}_r \in U$ $(r \geq 0)$. Dann bilden $\underline{x}_1, \ldots, \underline{x}_r$ eine *Basis* von U, wenn gilt: Für alle $\underline{x} \in U$ gibt es eindeutig bestimmte $a_1, \ldots, a_r \in K$ mit

$$\underline{x} = a_1\underline{x}_1 + \cdots + a_r\underline{x}_r \ .$$

Bemerkung Der lineare Unterraum $\underline{0}$ hat das leere System als Basis.

Die Standardaufgaben zu linearen Gleichungssystemen lauten nun wie folgt:

- Gegeben ein homogenes lineares Gleichungssystem, finde eine Basis des Lösungsraums $\mathbb{L} = \mathbb{L}_h$!
- Gegeben ein inhomogenes lineares Gleichungssystem, finde eine Basis von $\mathbb{L}_h$ sowie eine "spezielle Lösung" $\underline{x}_0 \in \mathbb{L}_h$!

Diese Aufgaben kann man mit dem *Gauß-Algorithmus* lösen, den wir im nächsten Abschnitt behandeln.

Doch zunächst noch weitere Begriffe.

Sei im Folgenden U immer ein linearer Unterraum, und seien $\underline{x}_1, \ldots, \underline{x}_r \in U$ $(r \geq 0)$.

Definition Die Vektoren $\underline{x}_1, \ldots, \underline{x}_r$ bilden ein *Erzeugendensystem* von U, wenn gilt: Für alle $\underline{x} \in U$ gibt es $a_1, \ldots, a_r \in K$ so dass $\underline{x} = a_1\underline{x}_1 + \cdots + a_r\underline{x}_r$.

Man sagt dann auch, dass U von $\underline{x}_1, \ldots, \underline{x}_r$ *erzeugt* wird.

Definition Der von $\underline{x}_1, \ldots, \underline{x}_r$ *erzeugte lineare Unterraum* ist

$$\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K := \{a_1\underline{x}_1 + \cdots + a_r\underline{x}_r \mid a_1, \ldots, a_r \in K\} \ .^1$$

[1] Man schreibt auch $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle$. Vorsicht: Hier besteht u.U. Verwechslungsgefahr mit der von $\underline{x}_1, \ldots, \underline{x}_r$ erzeugten abelschen Untergruppe von K^n (Siehe Beispiel 1.35.)

Bemerkung Der vom leeren System erzeuge lineare Unterraum ist $\{0\}$.

Man sieht leicht, dass dies in der Tat ein linearer Unterraum ist. Genauer ist dies der kleinste lineare Unterraum von K^n, der $\underline{x}_1, \ldots, \underline{x}_r$ enthält. (Vergleichen Sie dies mit Beispiel 1.35!) Offensichtlich wird er von $\underline{x}_1, \ldots, \underline{x}_r$ erzeugt.

Wir haben das folgende offensichtliche Lemma.

Lemma 2.5 *$\underline{x}_1, \ldots, \underline{x}_r$ ist genau dann ein Erzeugendensystem von U, wenn $U = \langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$.*

Definition Seien $\underline{x}_1, \ldots, \underline{x}_r \in K^n$, und sei $\underline{x} \in K^n$. Dann heißt $\underline{x}$ *linear abhängig* von $\underline{x}_1, \ldots, \underline{x}_r$, wenn es $a_1, \ldots, a_r \in K$ mit

$$\underline{x} = a_1\underline{x}_1 + \cdots + a_r\underline{x}_r$$

gibt. Wenn dies nicht der Fall ist, heißt $\underline{x}$ linear unabhängig von $\underline{x}_1, \ldots, \underline{x}_r$.

Bemerkung Der Nullvektor ist von jedem System linear abhängig, auch vom leeren System.

Lemma 2.6 *Die folgenden Aussagen sind äquivalent:*

- *$\underline{x}$ ist linear unabhängig von $\underline{x}_1, \ldots, \underline{x}_r$.*
- *$\underline{x} \notin \langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$.*
- *$\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K \subsetneq \langle \underline{x}_1, \ldots, \underline{x}_r, \underline{x} \rangle_K$.*

Beweis. Die ersten beiden Aussagen sind offensichtlich äquivalent, und die zweite impliziert die dritte.

Es ist zu zeigen, dass die dritte Aussage die zweite impliziert bzw. dass das Gegenteil der zweiten Aussage das Gegenteil der dritten Aussage impliziert.

Es gelte also $x \in \langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$. Dann ist $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$ ein linearer Unterraum, der $\underline{x}_1, \ldots, \underline{x}_r, \underline{x}$ enthält, und es ist auch der kleinste solche lineare Unterraum. Denn: Sei V ein weiterer Unterraum mit dieser Eigenschaft. Dann enthält V auch $\underline{x}_1, \ldots, \underline{x}_r$, und somit enthält V auch $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$.

Der kleinste lineare Unterraum, der $\underline{x}_1, \ldots, \underline{x}_r, \underline{x}$ enthält, ist jedoch $\langle \underline{x}_1, \ldots, \underline{x}_r, \underline{x} \rangle_K$. Damit ist $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K = \langle \underline{x}_1, \ldots, \underline{x}_r, \underline{x} \rangle_K$. □

Definition Die Vektoren $\underline{x}_1, \ldots, \underline{x}_r$ heißen *linear unabhängig*, wenn keiner der Vektoren von den anderen linear abhängig ist, mit anderen Worten, falls für alle $i = 1, \ldots, r$ gilt: $\underline{x}_i$ ist linear unabhängig von $\underline{x}_1, \ldots, \underline{x}_{i-1}, \underline{x}_{i+1}, \ldots, \underline{x}_r$.

Lemma 2.7 *Seien $\underline{x}_1, \ldots, \underline{x}_r \in K^n$. Dann sind diese Vektoren genau dann linear unabhängig, wenn gilt:*

$$\forall a_1, \ldots, a_r \in K : a_1\underline{x}_1 + \cdots + a_r\underline{x}_r = 0 \longrightarrow a_1 = \cdots = a_r = 0 \; .$$

Beweis. Wir zeigen, dass die Vektoren genau dann linear abhängig sind, wenn das Kriterium im Lemma nicht gilt.

Wenn die Vektoren linear abhängig sind, ist das Kriterium im Lemma offensichtlich falsch.

Sei nun das Kriterium im Lemma falsch. Dann gibt es $a_1, \ldots, a_r$, nicht alle $= 0$, so dass $a_1\underline{x}_1 + \cdots + a_r\underline{x}_r = 0$. Sei $a_i \neq 0$. Dann ist also

$$\underline{x}_i = -\frac{a_1}{a_i}\underline{x}_1 - \cdots - \frac{a_{i-1}}{a_i}\underline{x}_{i-1} - \frac{a_{i+1}}{a_i}\underline{x}_{i+1} - \cdots - \frac{a_r}{a_i}\underline{x}_r \; ,$$

d.h. $\underline{x}_i$ ist linear abhängig von den anderen Vektoren. □

Bemerkung Man sagt: "Das System $\underline{x}_1, \ldots, \underline{x}_r$ ist genau dann linear unabhängig, wenn sich der Nullvektor nur auf triviale Weise als Linearkombination der $\underline{x}_i$ darstellen lässt."

Bemerkung Man sollte immer versuchen, das Kriterium im obigen Lemma anzuwenden, wenn man zeigen will, dass ein System von Vektoren linear unabhängig ist.

Definition

- ein *maximales linear unabhängiges System* ist ein linear unabhängiges System so dass für alle Vektoren $\underline{x}$ gilt: Das System $\underline{x}_1, \ldots, \underline{x}_r, \underline{x}$ ist linear abhängig.

- ein *minimales Erzeugendensystem* ist ein Erzeugendensystem so dass für alle $i = 1, \ldots, r$ gilt: $\underline{x}_1, \ldots, \underline{x}_{i-1}, \underline{x}_{i+1}, \ldots, \underline{x}_r$ ist kein Erzeugendensystem.

Übungsaufgabe Es gibt einen Zusammenhang zwischen den obigen Definitionen und den Begriffen "maximales Element" / "minimales Element" bei Ordnungsrelationen. Welche Menge und welche Ordnungsrelation sollte man betrachten, damit sich die obige Definition aus den allgemeinen Definitionen von Ordnungsrelationen ergeben?

Aussage 2.8 *Sei U ein linearer Unterraum, und seien $\underline{x}_1, \ldots, \underline{x}_r \in U$. Dann sind die folgenden Aussagen äquivalent:*

a) $\underline{x}_1, \ldots, \underline{x}_r$ ist eine Basis von U.

b) $\underline{x}_1, \ldots, \underline{x}_r$ ist ein linear unabhängiges Erzeugendensystem von U.

c) $\underline{x}_1, \ldots, \underline{x}_r$ ist ein maximales linear unabhängiges System von U.

d) $\underline{x}_1, \ldots, \underline{x}_r$ ist ein minimales Erzeugendensystem von U.

Beweis. Wir zeigen, dass alle Aussagen äquivalent zur Aussage b) sind. Zunächst zeigen wir, dass jede der Aussagen a), c), d) (für sich) die Aussage b) impliziert.

Es gelte a). Dann haben wir insbesondere ein Erzeugendensystem. Es ist zu zeigen, dass das System auch linear unabhängig ist. Seien dazu $a_1, \cdots, a_r \in K$ mit $a_1\underline{x}_1 + \ldots + a_r\underline{x}_r = \underline{0}$. Wir haben auch $0\,\underline{x}_1 + \cdots + 0\,\underline{x}_r = \underline{0}$. Hieraus folgt $a_1 = 0, \ldots, a_r = 0$ aufgrund der Eindeutigkeit der "Darstellung von $\underline{0}$". Damit gilt b).

Es gelte nun c). Dann haben wir also ein linear unabhängiges System. Es ist zu zeigen, dass wir auch ein Erzeugendensystem haben. Sei dazu $\underline{x} \in U$. Dann ist nach Voraussetzung das System $\underline{x}_1, \ldots, \underline{x}_r, \underline{x}$ linear abhängig. Es gibt also $a_1, \ldots, a_r, a \in K$, nicht alle $= 0$, mit $a_1\underline{x}_1 + \cdots + a_r\underline{x}_r + a\underline{x} = \underline{0}$. Wenn nun $a = 0$ wäre, dann wären die Vektoren $\underline{x}_1, \ldots, \underline{x}_r$ linear abhängig, was aber nach Voraussetzung nicht der Fall ist. Also ist $a \neq 0$. Damit gilt $\underline{x} = -\frac{a_1}{a}\underline{x}_1 - \cdots - \frac{a_r}{a}\underline{x}_r$. Damit gilt wiederum b).

Es gelte nun d). Dann haben wir also ein Erzeugendensystem. Es ist zu zeigen, dass wir auch ein linear unabhängiges System haben. Sei dazu $i = 1, \ldots, r$. Nach Voraussetzung ist $\langle \underline{x}_1, \ldots, \underline{x}_{i-1}, \underline{x}_{i+1}, \underline{x}_r \rangle_K \subsetneq U = \langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$ (siehe auch Lemma 2.5). Damit ist nach Lemma 2.6 $\underline{x}_i$ linear unabhängig von den anderen Vektoren.

Es gelte nun b).

Wir zeigen zuerst a). Offensichtlich können wir jeden Vektor in der gewünschten Weise darstellen. Wir müssen die Eindeutigkeit der Darstellung zeigen. Sei also $\underline{x} \in U$ und seien $a_1, \ldots, a_r \in K$ und $a'_1, \ldots, a'_r \in K$ mit $\underline{x} = a_1\underline{x}_1 + \cdots + a_r\underline{x}_r = a'_1\underline{x}_1 + \cdots + a'_r\underline{x}_r$. Dann ist

$$\underline{0} = (a_1 - a'_1)\underline{x}_1 + \cdots + (a_r - a'_r)\underline{x}_r \ .$$

Nach Voraussetzung ist nun $a_1 - a'_1 = \cdots = a_r - a'_r = 0$, d.h. $a_1 = a'_1, \ldots, a_r = a'_r$.

Nun zu c). Da $\underline{x}_1, \ldots, \underline{x}_r$ ein Erzeugendensystem von U ist, ist jeder Vektor in U linear abhängig von $\underline{x}_1, \ldots, \underline{x}_r$. Damit gilt c).

Zu d). Sei $i = 1, \ldots, r$. Nach Voraussetzung ist $\underline{x}_i$ linear unabhängig von $\underline{x}_1, \ldots, \underline{x}_{i-1}, \underline{x}_{x+1}, \underline{x}_r$, also gilt $\underline{x}_i \notin \langle \underline{x}_1, \ldots, \underline{x}_{i-1}, \underline{x}_{i+1}, \ldots, \underline{x}_r \rangle_K$. Damit ist insbesondere $\underline{x}_1, \ldots, \underline{x}_{i-1}, \underline{x}_{x+1}, \ldots, \underline{x}_r$ kein Erzeugendensystem von U. □

2.2 Der Gauß-Algorithmus

In diesem Abschnitt diskutieren wir, wie man lineare Gleichungssysteme explizit löst.

Gegeben sei also ein lineares Gleichungssystem

$$\begin{array}{rcl} a_{1,1}X_1 + \cdots + a_{1,n}X_n & = & b_1 \\ a_{2,1}X_1 + \cdots + a_{2,n}X_n & = & b_2 \\ \vdots \qquad\qquad\quad \vdots & & \vdots \\ a_{m,1}X_1 + \cdots + a_{m,n}X_n & = & b_m \end{array} \tag{2.3}$$

über dem Körper K. Die Aufgabe besteht darin, zu entscheiden, ob das System lösbar ist, und gegebenenfalls eine Lösung $\underline{x}_0$ (genannt "spezielle Lösung") sowie eine Basis $\underline{x}_1, \ldots, \underline{x}_r$ des zugehörigen homogenen Systems zu finden. Der Lösungsraum des Systems ist dann der affine Raum

$$\underline{x}_0 + \langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K \ .$$

Wir sagen, dass zwei lineare Gleichungssysteme *äquivalent* sind, wenn ihre Lösungsmengen gleich sind. Der *vollständige Gauß-Algorithmus* (auch *Gauß-Jordan-Algorithmus* genannt) besteht nun darin, das System solange mittels bestimmter (einfacher) Operationen umzuformen, bis man eine "spezielle Lösung" sowie eine Basis des zugehörigen homogenen Systems ablesen kann.

Wir betrachten hierzu die folgenden drei Operationen.

(I) Multiplikation einer Gleichung mit einem Körperelement $\neq 0$.
(II) Vertauschen von zwei Gleichungen.
(III) Addition von c-mal Gleichung i zu Gleichung j (wobei $i \neq j$ und $c \in K$).

Zur Verdeutlichung: Die Operationen (I) und (III) sind konkret wie folgt gegeben:

(I) Sei $c \in K - \{0\}$ und $i = 1, \ldots, m$. Dann wird die i-te Gleichung aus (2.3) durch die Gleichung

$$ca_{i,1}X_1 + \cdots + ca_{i,n}X_n = cb_i$$

ersetzt.
(III) Sei $c \in K$ (c muss nicht $\neq 0$ sein, aber wenn es 0 ist, passiert nichts), und seien $i, j = 1, \ldots, m$ mit $i \neq j$. Dann wird die j-te Gleichung aus (2.3) durch die Gleichung

$$(a_{j,1} + ca_{i,1})X_1 + \cdots + (a_{j,n} + ca_{i,n})X_n = (b_j + cb_i)$$

ersetzt.

Lemma 2.9 *Operationen (I), (II), (III) überführen ein lineares Gleichungssystem in ein äquivalentes lineares Gleichungssystem.*

Beweis. Die Aussage ist offensichtlich für Operationen (I) und (II), wie betrachten also Operation (III).

Seien hierzu $i, j = 1, \ldots, k (i \neq j)$ und $c \in K$. Sei nun $\underline{x}$ eine Lösung von (2.3). Dann gilt insbesondere

$$a_{i,1}x_1 + \cdots + a_{i,n}x_n = b_i$$

und

$$a_{j,1}x_1 + \cdots + a_{j,n}x_n = b_j \ .$$

Hieraus folgt:

$$(a_{j,1} + ca_{i,1})x_1 + \cdots + (a_{j,n} + ca_{i,n})x_n = (b_j + cb_i) \ ,$$

und somit erfüllt $\underline{x}$ auch das umgeformte System.

Sei nun umgekehrt $\underline{x}$ eine Lösung des umgeformten Systems. Dann gilt insbesondere

$$a_{i,1}x_1 + \cdots + a_{i,n}x_n = b_i$$

und

$$(a_{j,1} + ca_{i,1})x_1 + \cdots + (a_{j,n} + ca_{i,n})x_n = (b_j + cb_i) \ .$$

Damit gilt auch

$$a_{j,1}x_1 + \cdots + a_{j,n}x_n = b_j \ .$$

Außerdem erfüllt $\underline{x}$ auch die Gleichungen $1, \ldots, j-1, j+1, \ldots, m$ des ursprünglichen Systems, weil diese nicht verändert werden. □

Definition Die Operationen (I), (II), (III) heißen *elementare Umformungen.*

Der Gauß-Algorithmus besteht nun darin, mittels der Operationen (I), (II), (III) ein System (2.3) in ein System in "Treppenform" umzuformen. Bei

Gauß-Jordan-Algorithmus rechnet man noch ein wenig weiter, bis man ein System in einer bestimmten, sehr einfachen, "Treppenform" hat. Neben den elementaren Umformungen ist es noch zweckmäßig, Gleichungen der Form $0X_1 + \cdots + 0X_n = 0$ ("Nullzeilen") wegzulassen.

Ich verdeutliche die Algorithmen an einem Beispiel.

Beispiel 2.10 Die Lösungsmenge des folgenden Systems über $\mathbb{Q}$ sei gesucht.

$$\begin{array}{rrrrrr} X_1 & -X_2 & & +X_4 & +X_5 & = 1 \\ -X_1 & & +X_3 & -2X_4 & -X_5 & = 0 \\ & 2X_2 & +X_3 & -X_4 & +3X_5 & = 1 \\ -2X_1 & +3X_2 & +X_3 & -3X_4 & & = -1 \end{array}$$

Für die Rechnungen ist es zweckmäßig, das System in der folgenden symbolischen Form hinzuschreiben.

$$\begin{array}{rrrrr|r} X_1 & X_2 & X_3 & X_4 & X_5 & \\ 1 & -1 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -2 & -1 & 0 \\ 0 & 2 & 1 & -1 & 3 & 1 \\ -2 & 3 & 1 & -3 & 0 & -1 \end{array} \tag{2.4}$$

Die Variablen in der ersten Zeile werden wir im Folgenden auch weglassen. Wir wenden Operation (III) an: Wir addieren die erste Zeile zur zweiten, und wir addieren das 2-fache der ersten Zeile zur vierten. Wir erhalten:

$$\begin{array}{rrrrr|r} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & -1 & 0 & 1 \\ 0 & 2 & 1 & -1 & 3 & 1 \\ 0 & 1 & 1 & -1 & 2 & 1 \end{array}$$

Wir addieren nun $2\times$ die zweite Zeile zur dritten Zeile sowie die zweite Zeile zur vierten und erhalten:

$$\begin{array}{rrrrr|r} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & -1 & 0 & 1 \\ 0 & 0 & 3 & -3 & 3 & 3 \\ 0 & 0 & 2 & -2 & 2 & 2 \end{array}$$

Wir multiplizieren die zweite Zeile mit -1, die dritte mit $\frac{1}{3}$ und die vierte mit $\frac{1}{2}$ und erhalten:

$$\begin{array}{rrrrr|r} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}$$

Nun steht in der vierten Zeile das Gleiche wie in der dritten, und wir können die vierte Zeile weglassen. (Mittels der elementaren Umformungen geht das so: Wir addieren $(-1)\times$ die dritte Zeile zur vierten. Dann erhalten wir eine "Null-Zeile", und diese können wir weglassen.) Wir erhalten:

$$\left.\begin{array}{rrrrr|r} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right|$$

Die ist ein lineares Gleichungssystem in so genannter *Treppenform* oder *(Zeilen-)Stufenform* (siehe unten für Definition).

Die Lösungsmenge kann nun durch "Auflösen" bestimmt werden. Hierzu geben wir uns einen belieben Vektor $\underline{x} \in \mathbb{Q}^n$ vor und setzen $\lambda := x_4, \mu := x_5$. Dann ist $\underline{x}$ genau dann eine Lösung, wenn gilt:

$$\begin{aligned} x_3 &= \lambda - \mu + 1 \\ x_2 &= x_3 - x_4 - 1 = (\lambda - \mu + 1) - \lambda - 1 = -\mu \\ x_1 &= x_2 - x_4 - x_5 + 1 = -\mu - \lambda - \mu + 1 = -\lambda - 2\mu + 1 \ . \end{aligned}$$

Damit ist die Lösungsmenge

$$\left\{ \begin{pmatrix} -\lambda - 2\mu + 1 \\ -\mu \\ \lambda - \mu + 1 \\ \lambda \\ \mu \end{pmatrix} | \lambda, \mu \in \mathbb{Q} \right\} =$$

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} -2 \\ -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \mid \lambda, \mu \in \mathbb{Q} \right\}$$

Der soeben durchgeführte Algorithmus heißt *Gauß-Algorithmus*.

Das "Auflöseverfahren" ist jedoch recht unübersichtlich und fehleranfällig. Besser ist es, noch ein wenig mit elementaren Umformungen weiterzurechnen. Das Ziel ist, dass auf allen "Treppenstufen" eine 1 steht (das sind hier die Elemente mit Indices (1,1), (2,2), (3,3) und das ist hier schon der Fall), und dass über all diesen "Treppenstufen" nur Nullen stehen. Das bedeutet hier, dass die Elemente mit Indices $(1,2)$, $(1,3)$ und $(2,3)$ Null sein sollen.

Hierzu addieren wir zuerst die dritte Zeile zur zweiten und erhalten:

$$\left.\begin{array}{rrrrr|r} 1 & -1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right|$$

Jetzt addieren wir die zweite zur ersten und erhalten:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right)$$

Dies ist ein System in so genannter *reduzierter (Zeilen-)Stufenform* oder *reduzierter Treppenform* (siehe unten für Definition).

Hieraus kann man direkt ablesen, dass $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ eine "spezielle Lösung" ist.

Gesucht ist nunnoch eine Basis des Lösungsraums des zugehörigen homogenen Systems:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array}\right)$$

(Die Spalte rechts des Strichs kann man nun weglassen.)

Nun gibt es einen "Ablesetrick", der wie folgt funktioniert: Man fügt neue *Zeilen* ein, und zwar so: Die Zeilen enthalten nur Nullen und genau eine -1. Sie werden so eingefügt, dass man ein System mit gleich viel Spalten wie Zeilen erhält, das die folgenden Eigenschaften hat: Unter der Diagonalen stehen nur Nullen und jeder Eintrag auf der Diagonalen ist entweder eine 1 oder eine -1. In unserem Fall sieht das dann so aus:

$$\begin{array}{rccccc} & 1 & 0 & 0 & 1 & 2 \\ & 0 & 1 & 0 & 0 & 1 \\ & 0 & 0 & 1 & -1 & 1 \\ \text{eingefügt} \rightarrow & 0 & 0 & 0 & -1 & 0 \\ \text{eingefügt} \rightarrow & 0 & 0 & 0 & 0 & -1 \\ & & & & \uparrow & \uparrow \end{array}$$

(Wenn wir die Null-Zeile nicht gestrichen hätten, wäre diese nun weggefallen.)

Diejenigen *Spalten*, in denen eine -1 eingefügt wurde (mit ↑ gekennzeichnet), bilden nun eine Basis von $\mathbb{L}_h$. In unserem Fall sind dies die Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix}.$$

Somit ist die Lösungsmenge des ursprünglichen Systems gleich

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \langle \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ -1 \end{pmatrix} \rangle_{\mathbb{Q}} .$$

Vergleichen Sie dies mit der zuerst berechneten Darstellung der Lösungsmenge!

Überlegen Sie sich auch, warum der "Ablesetrick" allgemein funktioniert!

Der gesamte Algorithmus bis zum "Ablesetrick" heißt *vollständiger Gauß-Algorithmus* oder *Gauß-Jordan-Algorithmus*.

Eine große Bitte. Der "Ablesetrick" (einfügen der "-1-Zeilen") ist eine grundsätzlich andere Operation als die vorangegangenen elementaren Umformungen. Deshalb sollte man das Einfügen auch entsprechend kenntlich machen, z.B. in dem man schreibt "Ich wende nun den 'Ablesetrick' an." und / oder indem man die eingefügten "-1-Zeilen" mit Bleistift schreibt.

Bevor wir das Lösen linearer Gleichungssysteme genauer betrachten zunächst eine Definition.

Definition Seien $m, n \in \mathbb{N}$. Eine $m \times n$-*Matrix* über K ist eine Abbildung $\{1, \ldots, m\} \times \{1, \ldots, n\} \longrightarrow K,\ (i, j) \mapsto a_{i,j}$.

So eine Matrix schreibt man oft als rechteckiges Schema mit "Klammern darum":

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{2,2} & \cdots & a_{m,n} \end{pmatrix}$$

Eine andere übliche Schreibweise ist $((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n}$. Matrizen werden meist mit großen Buchstaben bezeichnet. Wir schreiben $K^{m\times n}$ für $K^{\{1,\ldots,m\}\times\{1,\ldots,n\}}$, die Menge der $m\times n$-Matrizen. Andere oft benutzte Schreibweisen sind $\mathcal{M}_{m,n}(K)$ oder $\mathcal{M}_{m\times n}(K)$.

Einem homogenen linearen Gleichungssystem (2.2) kann man auf offensichtliche Weise die so genannte *Koeffizientenmatrix* $A = ((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n} \in K^{m\times n}$ zuordnen. Einem beliebigen (inhomogenen) linearen Gleichungssystem (2.3) ordnet man die so genannte *erweiterte Koeffizientenmatrix* zu. Diese entsteht, indem man rechts an die Koeffizientenmatrix A des zugehörigen

homogenen Systems den Vektor $\underline{b}$ für die rechte Seite "anhängt". Die entstehende $m \times (n+1)$-Matrix bezeichen wir mit $(A|\underline{b})$. (Der Strich dient nur der Abgrenzung und hat keine Bedeutung.)

Die elementaren Umformungen lassen sich nun als Operationen auf Matrizen auffassen. In diesem Kontext spricht man von *elementaren Zeilentransformationen*.

Umgekehrt kann man die Begriffe "(reduzierte) (Zeilen-)Stufenform" auch für Gleichungssysteme statt Matrizen anwenden (wie im Beispiel oben geschehen).

Definition Wir definieren die folgende Multiplikation von $m \times n$-Matrizen über K mit Vektoren in K^n:

$$\cdot : K^{m\times n} \times K^n \longrightarrow K^n , \quad (A, \underline{x}) \mapsto \begin{pmatrix} \sum_{j=1}^n a_{1,j}x_j \\ \sum_{j=1}^n a_{2,j}x_j \\ \vdots \\ \sum_{j=1}^n a_{m,j}x_j \end{pmatrix}$$

Seien $\underline{a}_1, \ldots, \underline{a}_n$ die Spalten von A, d.h. $A = (\underline{a}_1, \ldots, \underline{a}_k)$. Dann gilt also

$$A \cdot \underline{x} = x_1\underline{a}_1 + \cdots + x_n\underline{a}_n .$$

Nun ist $\underline{x} \in K^n$ genau dann eine Lösung des linearen Gleichungssystems (2.1), wenn

$$A \cdot \underline{x} = \underline{b}$$

gilt. Das LGS selbst können wir in der Form

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$$

schreiben, wobei wie oben die X_i die Unbestimmten sind.

Definition Eine Matrix in *(Zeilen-)Stufenform* bzw. in *Treppenform* ist eine Matrix der Gestalt

$$\begin{pmatrix}
* & \circ & \cdots & \circ & \circ & \circ & \cdots & \circ & \circ & \circ & \cdots & \circ & \cdots & \circ & \circ & \cdots & \circ \\
 & & & & * & \circ & \cdots & \circ & \circ & \circ & \cdots & \circ & \cdots & \circ & \circ & \cdots & \circ \\
 & & & & & & & & * & \circ & \cdots & \circ & \cdots & \circ & \circ & \cdots & \circ \\
 & & & & & & & & & & & & \ddots & \vdots & \vdots & \cdot & \vdots \\
 & & & & & & & & & & & & & * & \circ & \cdots & \circ \\
 & & & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & & &
\end{pmatrix},$$

wobei die mit $*$ bezeichneten Einträge alle $\neq 0$ sind, die mit $\circ$ bezeichneten Einträge beliebig sind, und die Einträge ohne Bezeichnung (d.h. die Einträge unter der "Treppe" 0 sind.

Mit anderen Worten: So eine Matrix hat eine "Treppe", wobei jede Treppenstufe ein Eintrag $\neq 0$ ist. Jede Treppenstufe hat Höhe 1, und neben jeder Stufe dürfen beliebige Einträge stehen. Unter der Treppe stehen nur Nullen.

Beachten Sie, dass links 0-Spalten und unten 0-Zeilen stehen dürfen (aber nicht müssen) (diese Spalten bzw. Zeilen sind durch den freien Platz links und unten angedeutet.)

Definition Eine Matrix in *reduzierter (Zeilen-)Stufenform* bzw. in *reduzierter Treppenform* ist eine Matrix der Gestalt

$$\begin{pmatrix} & 1 & \circ & \cdots & \circ & 0 & \circ & \cdots & \circ & 0 & \circ & \cdots & \circ & \cdots & 0 & \circ & \cdots & \circ \\ & & & & & 1 & \circ & \cdots & \circ & 0 & \circ & \cdots & \circ & \cdots & 0 & \circ & \cdots & \circ \\ & & & & & & & & & 1 & \circ & \cdots & \circ & \cdots & 0 & \circ & \cdots & \circ \\ & & & & & & & & & & & & & \ddots & \vdots & \vdots & \cdot & \vdots \\ & & & & & & & & & & & & & & 1 & \circ & \cdots & \circ \\ & & & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & & & & & & \end{pmatrix},$$

wobei die $\circ$'s wieder beliebige Einträge repräsentieren und unter der "Treppe" wiederum nur Nullen stehen. Mit anderen Worten: Eine Matrix in reduzierter (Zeilen-)Stufenform ist eine Matrix in (Zeilen-)Stufenform so dass
- die "Treppenstufen" alle gleich 1 sind,
- über den "Treppenstufen" nur Nullen stehen.

(Diese Bedingungen beziehen sich wirklich nur auf die Stufen, nicht auf die Elemente, die rechts daneben stehen!)

Definition Die $n \times n$ *Einheitsmatrix* ist die Matrix

$$I = I_n = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} = (\underline{e}_1 | \cdots | \underline{e}_n) \,,$$

wobei "wie immer" die nicht gekennzeichneten Einträge gleich Null sind. Beachten Sie, dass dies ein Spezialfall einer Matrix in reduzierter (Zeilen-) Stufenform ist.

Bemerkung Mittels des "Kronecker-Deltas" (siehe 1.10) kann man auch $I_n = ((\delta_{i,j}))_{i,j=1,\ldots,n}$ schreiben.

Ich beschreibe nun den Gauß-Algorithmus für Matrizen zur Transformation in eine Matrix in (Zeilen-)Stufenform. Ich wähle eine rekursive Beschreibung. Man würde den Algorithmus allerdings eher wohl mit Schleifen implementieren. In dem folgenden Algorithmus wird die Matrix ohne Kopien anzufertigen fortlaufend transformiert. (D.h. bei den rekursiven Aufrufen werden keine Kopien der Matrix (oder Teile der Matrix) angefertigt.)

Der Gauß-Algorithmus

Eingabe. Eine Matrix $A \in K^{m\times n}$.
Ausgabe. Eine Matrix $\tilde{A}$, die aus A durch elementare Zeilentransformationen hervorgeht und in (Zeilen-)Stufenform ist.

Wenn die erste Spalte eine Nullspalte ist,
  wenn die Matrix mehr als eine Spalte hat,
    wende den Algorithmus auf die Matrix an, die entsteht,
      indem man die erste Spalte streicht.
Ansonsten
  Wähle ein i so dass $a_{i,1} \neq 0$.
  Vertausche die Zeilen i und j (Transformation (II)).
  (Für die Matrix gilt nun, dass $a_{1,1} \neq 0$.)
  Multipliziere eventuell die erste Zeile mit einer Konstanten (z.B. mit $a_{1,1}^{-1}$) (Transformation (I)).
  Für $i = 2, \ldots, m$: Addiere jeweils $-\frac{a_{i,1}}{a_{1,1}}$-mal die erste Zeile zur i-ten Zeile (Transformation (III)).
  Wenn die Matrix mehr als eine Zeile und mehr als eine Spalte hat,
    Wende den Algorithmus auf die Matrix an, die entsteht,
      indem man die erste Zeile und die erste Spalte streicht.

Wenn die erste Spalte keine Nullspalte ist, sieht die Matrix nach dem vorletzten Schritt so aus:

$$\begin{pmatrix} * & \circ & \cdots & \circ \\ 0 & \circ & \cdots & \circ \\ \vdots & \vdots & \cdot & \vdots \\ 0 & \circ & \cdots & \circ \end{pmatrix} \tag{2.5}$$

(Mit $* \neq 0$ und $\circ$ beliebig.) Im letzten Schritt wird dann der Algorithmus mit der Matrix aufgerufen, die entsteht, wenn man in dieser Matrix die erste

Zeile und die erste Spalte weglässt. Das Endergebnis ist das Ergebnis dieser Berechnung zusammen mit der ersten Zeile und der ersten Spalte der Matrix (2.5).

Sicher terminiert der Algorithmus, und der Algorithmus ist offensichtlich korrekt: Das Ergebnis des Algorithmus ist offensichtlich eine Matrix in (Zeilen-)Stufenform, die aus der ursprünglichen Matrix durch elementare Zeilentransformationen hervorgegangen ist.

Um eine Matrix in reduzierter (Zeilen-)Stufenform zu erhalten, geht man ausgehend von einer Matrix in (Zeilen-)Stufenform wie folgt vor:

Zuerst teilt man alle nicht-Nullzeilen durch ihren ersten Eintrag $\neq 0$. Dann sind also die Einträge auf allen Stufen gleich 1.

Dann "räumt" man mittels Operationen von Typ (III) die Einträge oberhalb der Stufen aus (man erzeugt Nullen). Dabei kann man die "Stufenspalten" in beliebiger Reihenfolge behandeln.

Das Ergebnis ist eine Matrix in reduzierter (Zeilen-)Stufenform. Der gesamte soeben beschriebene Algorithmus heißt *vollständiger Gauß-Algorithmus* oder *Gauß-Jordan-Algorithmus*.

Sei nun das LGS

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$$

gegeben. Um die Lösungsmenge zu bestimmen (genauer um eine spezielle Lösung und eine Basis des zugehörigen homogenen Systems zu bestimmen), kann man nun wie folgt vorgehen:

- Man wendet den Gauß-Algorithmus auf die Matrix $(A|\underline{b})$ an. Sei $(\tilde{A}|\tilde{b})$ das Ergebnis. Dann ist das System genau dann lösbar, wenn in der letzten Spalte keine Treppenstufe ist.

Wenn das System lösbar ist,
- erzeugt man eine Matrix in reduzierter (Zeilen-)Stufenform.
- liest man die Lösung mittels des "Ablesetricks" ab.

Bemerkung Oftmals wird im Gauß-Algorithmus noch eine weitere Operation erlaubt: Man vertauscht die Spalten und merkt sich dabei, wie man die Variablen geändert hat. (Die Variablen sollte man dann beim Rechnen "von Hand" über das System schreiben, so wie in (2.4).) Ich empfehle allerdings dringend, diese Operation nur in Ausnahmefällen anzuwenden, da sie sehr fehlerbehaftet ist. Dies gilt insbesondere für die Klausur!

Diskussion Das Wort "Algorithmus" besagt, dass zu jedem Zeitpunkt genau festgelegt sein muss, welche Operation als nächstes ausgeführt wird. In dem oben beschriebenen Gauß-Algorithmus sind allerdings wesentliche Unbestimmtheiten vorhanden:[2]
- Es ist nicht festgelegt, welche Zeile man "nach oben holt". - Es ist nicht festgelegt, unter welchen Bedingungen man eine Transformation (I) durchführen soll, und wenn, mit welcher Konstante.
Somit handelt es sich streng genommen nicht um einen Algorithmus sondern eher um ein *Algorithmenschema*.

Man erhält einen Algorithmus, indem man eine konkrete Regel vorgibt, welche Zeile ausgewählt werden soll und nach dem Vertauschen die erste Zeile mit $a_{1,1}^{-1}$ multipliziert.

Die Auswahl so einer Zeile heißt *Pivotwahl*, und dasjenige Element so einer Zeile, dass für die Elimination benutzt wird (in der obigen Darstellung das erste Element), wird *Piviotelement* genannt. Eine Möglichkeit für die Pivotwahl (in der obigen Darstellung des Gauß-Algorithmus) ist, das kleinste i mit $a_{1,i} \neq 0$ zu bestimmen und dann die 1-ste und die i-te Zeile zu vertauschen. (Dies bezieht sich wirklich nur auf die obige rekursive Darstellung des Algorithmus. Natürlich wählt man keine Zeile derjenigen Zeilen aus, die man schon behandelt hat.)

Für Rechnungen von Hand wird man jedoch eher eine Zeile mit einem "möglichst einfach aussehenden" ersten Eintrag nach oben holen.

Für Berechnungen mit Fließkommazahlen (die reelle Zahlen darstellen) mittels eines Computers sollte man darauf achten, dass nur möglichst kleine Rundungsfehler auftreten. Hierzu ist es geschickt, ein Pivotelement mit möglichst großem Absolutbetrag zu wählen. (Dies hat etwas damit zu tun, dass man durch das Pivotelement teilen muss). Es ist also optimal, das größte Element der gesamten (restlichen) Matrix zu wählen und entsprechend Zeilen und Spalten (!) umzunummerieren (nicht umzukopieren!) (kein Problem auf dem Computer).

Komplexität

Die *Komplexität* des Gauß-Jordan-Algorithmus in Körperoperationen ist leicht berechnet.

Sei eine $m \times n$-Matrix über K gegeben.

[2]Immer wenn man einen Algorithmus in Pseudocode angibt, handelt man sich bestimmte Unbestimmtheiten ein, oder anders ausgedrückt, man läßt gewisse Freiheiten bei der Implementierung. Somit ist die Abgrenzung, ob man nun einen Algorithmus oder nur ein Algorihmenschema hat, etwas ungenau.

Offensichtlich benötigen wir höchstens $m-1$ Operationen vom Typ (II) und höchstens m Operationen vom Typ (I). Wir benötigen höchstens $(m-1)+(m-2)+\cdots+3+2+1=\frac{m\cdot(m-1)}{2}$ Operationen vom Typ (III), um eine Matrix in (Zeilen)-Stufenform zu erzeugen. Danach benötigen wir noch höchstens $\frac{m\cdot(m-1)}{2}$ Operationen vom Typ (III), um eine Matrix in reduzierter (Zeilen-)Stufenform zu erzeugen.

Für jede dieser Operationen benötigen wir höchstens n Körperoperationen.

Damit erhalten wir:

Aussage 2.11 *Gegeben eine $m \times n$-Matrix $A \in K^{m\times n}$, kann man in $\mathcal{O}(m^2 \cdot n)$ Körperoperationen eine reduzierte (Zeilen-)Stufenform berechnen. Insbesondere kann man die Lösungsmenge eines lineares Gleichungssystem mit n Unbestimmten und m Gleichungen in $\mathcal{O}(m^2 \cdot n)$ Körperoperationen bestimmen.*

Ich erinnere noch einmal daran, dass "Bestimmen der Lösungsmenge" bedeutet, eine "spezielle Lösung" und eine Basis der Lösungsmenge des zugehörigen homogenen Systems anzugeben.

Bemerkung Durch die Angabe der Komplexität in Körperoperationen wird wenig über das Problem ausgesagt, wie man nun möglichst schnell und möglichst exakt die Lösungsmenge eines Gleichungssystems über den reellen Zahlen mit Fließkommazahlen berechnet. Dieses Problem haben wir oben kurz angedeutet. Hierzu und zu verwandten Fragen gibt es eine umfangreiche Theorie, die in die *Numerische Mathematik* fällt.

Anwendungen

Ich gebe jetzt noch Anwendungen des Gauß-Algorithmus auf die Frage, ob ein System von Vektoren linear unabhängig bzw. ein Erzeugendensystem ist.

Es seien $\underline{a}_1, \ldots, \underline{a}_r \in K^n$ gegeben. Wir betrachten nun die $n \times r$-Matrix $A = (\underline{a}_1 | \cdots | \underline{a}_r)$, die entsteht, wenn man die Vektoren $\underline{a}_1, \ldots \underline{a}_r$ als Spalten einer Matrix auffasst.

Nun sind die Vektoren $\underline{a}_1, \ldots, \underline{a}_r$ genau dann linear unabhängig, wenn das System

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix} = \underline{0} \tag{2.6}$$

ausschließlich die "triviale" Lösung $\underline{0}$ (und keine weitere Lösung) hat. Ob dies der Fall ist, kann man nun mit dem Gauß-Algorithmus überprüfen.

Sei nun $\tilde{A}$ eine Matrix in (Zeilen-)Stufenform, die aus A mittels elementarer Zeilenoperationen entsteht. Dann ist also das System (2.6) äquivalent zum System

$$\tilde{A} \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix} = \underline{0} \,. \tag{2.7}$$

Dieses System hat genau dann eine "nicht-triviale" Lösung (eine Lösung $\neq \underline{0}$), wenn es Spalten von $\tilde{A}$ gibt, die (in $\tilde{A}$) keine "Treppenstufen" sind.

Wenn $r \geq n+1$ ist, gibt es immer eine Spalte von $\tilde{A}$, die keine "Treppenstufe" ist, und folglich gibt es immer eine "nicht-triviale" Lösung. Es gilt also:

Aussage 2.12 *n+1 (oder mehr) Vektoren in K^n sind immer linear abhängig.*

Mit den obigen Überlegungen können wir auch auf die folgende Frage eine algorithmische Antwort geben:

Gebeben ein System $\underline{a}_1, \ldots, \underline{a}_r \in K^n$, finde ein "Teilsystem", das eine Basis von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$ ist!

(Ein "Teilsystem" ist wie folgt definiert. Gegeben Seien $i_1 < i_2 < \ldots < i_k$ für ein $k \leq r$. Dann ist $\underline{a}_{i_1}, \ldots, \underline{a}_{i_k}$ ein Teilsystem von $\underline{a}_1, \ldots, \underline{a}_r$.)

Hierzu definieren wir die Matrix A wie oben und berechnen mittels elementarer Zeilenumformungen eine Matrix $\tilde{A}$. Dann bilden diejenigen Spalten aus A (!), für die in $\tilde{A}$ eine Stufe steht, eine Basis von $\langle \underline{a}_1, \ldots, \underline{a}_r \rangle_K$.

(Denn: Die "Nichtstufenspalten" von $\tilde{A}$ sind von den "Stufenspalten" von $\tilde{A}$ linear abhängig. Und das gilt auch für A, weil die Lösungsmenge eines LGS unter elementaren Zeilentransformationen eben nicht verändert wird. Außerdem sind die Stufenspalten von $\tilde{A}$ linear unabhängig, und das Gleiche gilt dann auch für die Stufenspalten von A.)

Wir kommen nun zu der Frage, wie man entscheiden kann, ob die $\underline{a}_1, \ldots, \underline{a}_r$ ein Erzeugendensystem von K^n bilden. Wieder betrachten wir die Matrix A, die definiert ist wie oben, sowie eine Matrix $\tilde{A}$ in (Zeilen-)Stufenform, die aus A durch elementare Transformationen hervorgeht.

Offensichtlich ist $\underline{a}_1, \ldots, \underline{a}_r$ genau dann ein Erzeugendensystem von K^n,

wenn für alle $\underline{b} \in K^n$ das LGS

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_r \end{pmatrix} = \underline{b} \tag{2.8}$$

lösbar ist. Diese Eigenschaft ist auch invariant unter elementaren Transformationen (warum?) und folglich äquivalent zu der Bedingung, dass das System (2.7) für alle $\underline{b} \in K^n$ lösbar ist. Dies wiederum ist dazu äquivalent, dass $\tilde{A}$ keine Nullzeile enthält. Dies ist jedoch nur dann möglich, wenn $n \leq r$.

Damit erhalten wir:

Aussage 2.13 *Jedes Erzeugendensystem von K^n besteht aus mindestens n Vektoren.*

Im Prinzip kann man mit dem Obigen auch eine algorithmische Antwort auf die Aufgabe geben, das System $\underline{a}_1, \ldots, \underline{a}_r$ zu einem Erzeugendensystem von K^n zu ergänzen:

Sei k die letzte Zeile von $\tilde{A}$, die keine Nullzeile ist. Dann bilden offensichtlich die Spalten von $\tilde{A}$ und die Vektoren $\underline{e}_{k+1}, \ldots, \underline{e}_n$ ein Erzeugendensystem von K^n.

Man betrachtet nun die Matrix $(\tilde{A}|\underline{e}_{k+1}| \cdots |\underline{e}_n)$. Wenn man diese Matrix mittels elementarer Umformungen umformt, bilden die Spalten immer noch ein Erzeugendensystem (siehe oben).

Wenn wir nun die elementaren Umformungen von A auf $\tilde{A}$ "rückgängig" machen (von $\tilde{A}$ nach A "zurückgehen"), und dabei die neuen Spalten auch mit umformen, erhalten wir eine Matrix $(A|N)$. Nach dem eben gesagten bilden die Spalten dieser Matrix ein Erzeugendensystem von K^n. Damit ist die Aufgabe gelöst.

Wir erhalten insbesondere:

Satz 2.1 *Jede Basis von K^n besteht aus genau n Vektoren.*

Genauer haben wir die folgenden Äquivalenzen.

- $\underline{a}_1, \ldots, \underline{a}_r$ ist eine Basis.
- Es gilt $r = n$ und in jeder Spalte von $\tilde{A}$ ist eine Stufe.

Sei nun $\tilde{A}$ eine Matrix, die aus A durch elementare Transformationen hervorgeht und in reduzierter Zeilenstufenform ist. Dann haben wir die Äquivalenzen:

- $\underline{a}_1, \ldots, \underline{a}_r$ ist eine Basis.
- $\tilde{A} = I_n$.

Bemerkung Wir sehen, dass die Matrix $\tilde{A}$ eindeutig ist, wenn $\underline{a}_1, \ldots, \underline{a}_r$ eine Basis ist. Allgemeiner kann man zeigen, dass sich jede Matrix mittels elementarer Zeilenumformungen zu genau einer Matrix in reduzierter (Zeilen-) Stufenform umformen lässt. Diese Matrix heißt dann auch die *Zeilennormalform* von A.

2.3 Lineare Abbildungen und Matrizen

Definition Eine *lineare Abbildung* von K^n nach K^m ist eine Abbildung $\varphi : K^n \longrightarrow K^m$ so dass

- $\forall \underline{x}, \underline{y} \in K^n : \varphi(\underline{x} + \underline{y}) = \varphi(\underline{x}) + \varphi(\underline{y})$
- $\forall a \in K \, \forall \underline{x} \in K^n : \varphi(a\underline{x}) = a\varphi(\underline{x})$.

Bemerkungen

- Eine lineare Abbildung von K^n nach K^m ist also insbesondere ein Gruppenhomomorhismus von $(K^n, +)$ nach $(K^m, +)$.
- Wenn $\varphi : K^r \longrightarrow K^n$ und $\psi : K^n \longrightarrow K^m$ lineare Abbildungen sind, dann ist auch $\psi \circ \varphi : K^r \longrightarrow K^m$ eine lineare Abbildung.
- Wenn $\varphi : K^n \longrightarrow K^m$ eine bijektive lineare Abbildung ist, dann ist auch $\varphi^{-1} : K^m \longrightarrow K^n$ eine lineare Abbildung.[3]

Beispiel 2.14 Sei $A \in K^{m \times n}$. Dann ist die Abbildung

$$\Lambda_A : K^n \longrightarrow K^m \, , \underline{x} \mapsto A\underline{x}$$

linear.

Bemerkung Sei $A = (\underline{a}_1 | \cdots | \underline{a}_n) \in K^{m \times n}$ und $\underline{x} \in K^n$. Dann ist $\Lambda_A(\underline{x}) = A\underline{x} = \sum_{j=1}^{n} x_j \, \underline{a}_j$. Insbesondere ist $\text{Bild}(\Lambda_A) = \langle \underline{a}_1, \ldots, \underline{a}_n \rangle_K$.

Wir haben somit jeder $m \times n$-Matrix eine lineare Abbildung von K^n nach K^m zugeordnet. Umgekehrt ordnen wir wie folgt jeder linearen Abbildung von K^n nach K^m eine Matrix in $K^{m \times n}$ zu:

[3] Wir werden in Satz 2.2 sehen, dass dann auch $n = m$ gilt.

Definition Sei $\varphi : K^n \longrightarrow K^m$ eine lineare Abbildung. Dann ist die zu φ *assoziierte Matrix* $M(\varphi)$ wie folgt definiert:

$$M(\varphi) := (\varphi(\underline{e}_1)|\cdots|\varphi(\underline{e}_n))$$

Hier ist $\underline{e}_j$ der j-te Standardvektor.

Aussage 2.15 *Die Zuordnungen $\varphi \mapsto M(\varphi)$ und $A \mapsto \Lambda_A$ definieren zueinander inverse Bijektionen zwischen der Menge der linearen Abbildungen von K^n nach K^m und der Menge der $m \times n$-Matrizen $K^{m\times n}$.*

Mit anderen Worten: Es gilt

$$\Lambda_{M(\varphi)} = \varphi \quad \textit{und} \quad M(\Lambda_A) = A$$

für alle linearen Abbildungen $\varphi : K^m \longrightarrow K^n$ und alle $m \times n$-Matrizen A.

Beweis. Sei zunächst A eine $m \times n$-Matrix, und sei $A = (\underline{a}_1|\cdots|\underline{a}_n)$. Dann ist $M(\Lambda_A) = (\Lambda_A(\underline{e}_1)|\cdots|\Lambda_A(\underline{e}_n)) = (A\underline{e}_1|\cdots A\underline{e}_n) = (\underline{a}_1|\cdots|\underline{a}_n) = A$, was zu zeigen war.

Sei nun $\varphi : K^n \longrightarrow K^m$ eine lineare Abbildung. Zu zeigen ist, dass für alle $x \in K^n$ gilt: $\Lambda_{M(\varphi)}(\underline{x}) = \varphi(x)$.

Sei hierfür $\underline{x} \in K^n$ beliebig. Dann gilt $\Lambda_{M(\varphi)}(\underline{x}) = M(\varphi) \cdot \underline{x} = (\varphi(\underline{e}_1)|\cdots|\varphi(\underline{e}_n)) \cdot \underline{x} = \sum_{j=1}^n x_i\, \varphi(\underline{e}_j) = \varphi(\sum_{j=1}^n x_j\, \underline{e}_j) = \varphi(\underline{x})$. □

Da lineare Abbildungen auch Gruppenhomomorhismen sind, können wir den Begriffs des *Kerns* auch auf lineare Abbildungen anwenden. Für eine lineare Abbildung $\varphi : K^n \longrightarrow K^m$ haben wir also

$$\mathrm{Kern}(\varphi) = \{\underline{x} \in K^n \,|\, \varphi(\underline{x}) = \underline{0}\}\ .$$

Wir definieren nun für jede Matrix $A \in K^{m\times n}$

$$\mathrm{Kern}(A) := \mathrm{Kern}(\Lambda_A),$$

d.h.

$$\mathrm{Kern}(A) = \{\underline{x} \in K^n \,|\, A\underline{x} = \underline{0}\}\ .$$

Beachten Sie, dass mit dieser Definition insbesondere gilt:

Die Lösungsmenge des homogenen LGS $A \cdot \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \underline{0}$ ist gleich dem Kern von A, also

$$\mathbb{L} = \mathrm{Kern}(A)\ .$$

Andererseits ist $\mathrm{Bild}(\Lambda_A)$ gleich der Menge der Vektoren $\underline{b}$, für die das inhomogene LGS $A \cdot \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$ lösbar ist.

Matrizenaddition und -multiplikation

Wir definieren eine Addition auf der Menge der $m \times n$-Matrizen $K^{m\times n}$ "komponentenweise":

$$+ : K^{m\times n} \times K^{m\times n} \longrightarrow K^{m\times n} ,$$
$$(((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots n}, ((b_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n}) \mapsto ((a_{i,j} + b_{i,j}))_{i=1,\ldots,m,j=1,\ldots n}$$

(Diese Definition ist ein Spezialfall der in 1.5, Unterabschnitt über Produkte, definierten Verknüpfung.)

Mit dieser Addition bildet die Menge der $m \times n$-Matrizen eine abelsche Gruppe.

Beachten Sie, dass man auch die linearen Abbildungen von K^n nach K^n addieren kann (siehe wiederum den erwähnten Unterabschnitt über Produkte oder die Diskussion auf Seite 46): Wenn $\varphi, \psi : K^n \longrightarrow K^n$ lineare Abbildungen sind, dann ist auch die Abbildung $\varphi + \psi : K^n \longrightarrow K^n$ (gegeben durch $\underline{x} \mapsto \varphi(\underline{x}) + \psi(\underline{x})$) linear. Auch die linearen Abbildungenvon von K^n nach K^m bilden eine abelsche Gruppe.

Wir haben nun

$$M(\varphi + \psi) = M(\varphi) + M(\psi) \text{ und } \Lambda_{A+B} = \Lambda_A + \Lambda_B \tag{2.9}$$

für alle linearen Abbildungen $\varphi, \psi : K^n \longrightarrow K^m$ und alle Matrizen $A, B \in K^{m\times n}$.

Mit anderen Worten: Die Zuordungen $\varphi \mapsto M(\varphi)$ und $A \mapsto \Lambda_A$ sind Isomorphismen zwischen der abelschen Gruppe der linearen Abbildungen von K^n nach K^m und der abelschen Gruppe $K^{m\times n}$.

Wir wollen nun eine *Multiplikation* $\cdot : K^{m\times n} \times K^{n\times r} \longrightarrow K^{m\times r}$ definieren, welche der Verknüpfung linearer Abbildungen entspricht. Mit anderen Worten: Wir wollen, dass für alle $A \in K^{m\times n}$ und alle $B \in K^{n\times r}$ gilt:

$$\Lambda_{A\cdot B} = \Lambda_A \circ \Lambda_B , \tag{2.10}$$

bzw.

$$A \cdot B = M(\Lambda_A \circ \Lambda_B) . \tag{2.11}$$

Wir *definieren* die Multiplikation also mittels (2.11). Wir wollen nun wissen, wie man explizit zwei Matrizen multipliziert. Seien dazu $A \in K^{m\times n}$ und $B \in K^{n\times r}$ mit $B = (\underline{b}_1 | \cdots | \underline{b}_r)$, und sei $C := A \cdot B$ mit $C = (\underline{c}_1 | \cdots | \underline{c}_r) = ((c_{i,j}))_{i=1,\ldots,m,j=1,\ldots,r}$. Dann ist

$$\underline{c}_j = C \cdot \underline{e}_j = \Lambda_C(\underline{e}_j) = (\Lambda_A \circ \Lambda_B)(\underline{e}_j) = \Lambda_A(\Lambda_B(\underline{e}_j)) = \Lambda_A(\underline{b}_j) = A \cdot \underline{b}_j .$$

Eine Umformulierung hiervon ist:

$$c_{i,j} = \sum_{\ell=1}^{n} a_{i,\ell} b_{\ell,j}$$

Mit anderen Worten:

$$A \cdot B = (A\underline{b}_1 | \cdots | A\underline{b}_r) \; . \tag{2.12}$$

bzw.

$$((a_{i,j}))_{i,j} \cdot ((b_{i,j}))_{i,j} = ((\sum_{\ell=1}^{n} a_{i,\ell}\, b_{\ell,j}))_{i,j} \; .$$

Insbesondere sieht man, dass die Multiplikation einer Matrix mit einem Spaltenvektor ein Spezialfall der Multiplikation zweier Matrizen ist (setze $r = 1$ in Formel (2.12)).

Da die Verknüpfung zweier (linearer) Abbildungen assoziativ ist, ist die Matrizenmultiplikation auch automatisch assoziativ (warum?). Es gilt also für alle $A \in K^{m\times n}, B \in K^{n\times r}, C \in K^{r\times s} : A(BC) = (AB)C$.

Wir wissen schon, dass für die Addition und Verknüpfung linearer Abbildungen die Distributivgesetze gelten:

$$\varphi \circ (\chi + \psi) = \varphi \circ \chi + \varphi \circ \psi$$

für alle linearen Abbildungen $\varphi : K^n \longrightarrow K^m$ und $\chi, \psi : K^r \longrightarrow K^n$ sowie

$$(\chi + \psi) \circ \varphi = \chi \circ \varphi + \psi \circ \varphi$$

für alle linearen Abbildungen $\chi, \psi : K^n \longrightarrow K^m$ und $\varphi : K^r \longrightarrow K^n$. (Weil lineare Abbildungen Gruppenhomomorphismen sind, siehe S. 46.) Die analoge Aussage gilt nun auch für Matrizen:

$$A(B + C) = AB + AC \qquad (B + C)A = BA + CA \; ,$$

wenn immer die Multiplikation definiert ist. (Dies kann man direkt nachrechnen oder (2.11) anwenden.)

Aussage 2.16 *Die Menge der linearen Abbildungen $K^n \longrightarrow K^n$ bildet mit der Verknüpfung als Multiplikation einen Ring. Genauso bildet die Menge der $n \times n$-Matrizen $K^{n\times n}$ mit der soeben definiertern Addition und Multiplikation einen Ring. Die Zuordnungen $\varphi \mapsto M(\varphi)$ sowie $A \mapsto \Lambda_A$ sind zueinander inverse Isomorphismen zwischen diesen Ringen.*

Hierfür ist schon so gut wie alles gezeigt. Es fehlt nur noch die Angabe der neutralen Elemente bezüglich der Multiplikation. Diese sind die identische Abbildung auf K^n sowie die Einheitsmatrix I_n.

Bemerkungen

- Die Menge der linearen Abbildungen von K^n nach K^n ist ein Unterring des Rings der Endomorphismen der abelschen Gruppe $(K^n, +)$.
- Für $n \geq 2$ sind die Ringe nicht-kommutativ.

Wenn $\varphi : K^n \longrightarrow K^m$ eine lineare Abbildung ist und $c \in K$, dann ist auch die Abbildung $c\varphi : K^n \longrightarrow K^m$, die per Definition durch $\underline{x} \mapsto c\varphi(\underline{x})$ gegeben ist, linear. Es gilt

$$(c\psi) \circ \varphi = c(\psi \circ \varphi) = \psi \circ (c\varphi)$$

für alle $c \in K$ und alle linearen Abbildungen $\varphi : K^r \longrightarrow K^n$ und $\psi : K^n \longrightarrow K^m$.

In Analogie hierzu definieren wir nun noch eine *Skalarmultiplikation*:

$$K \times K^{m\times n} \longrightarrow K^{m\times n} \, , (c, A) \mapsto ((c \cdot a_{i,j}))_{i,j} \, .$$

Damit gilt

$$M(c\varphi) = cM(\varphi) \text{ und } \Lambda_{cA} = c\Lambda_A \, .$$

Außerdem gilt

$$(cA)B = c(AB) = A(cB)$$

für $c \in K, A \in K^{m\times n}$ und $B \in K^{n\times r}$.

2.4 Matrizenmultiplikation und Gauß-Algorithmus

Mittels der Matrizenmultiplikation kann man den Gauß-Algorithmus reinterpretieren:

Sei $A \in K^{m\times n}$. Jede der drei elementaren Zeilentransformationen angewandt auf A entspricht einer Multiplikation mit einer bestimmten invertierbaren Matrix in $K^{m\times m}$ von links. Wir betrachten hierzu die drei elementaren Zeilentransformationen.

(I) Multiplikation der i-ten Zeile mit $c \in K^*$. Dies entspricht der Multiplikation mit der Matrix

$$(\underline{e}_1 \,|\, \cdots \,|\, \underline{e}_{i-1} |\, c\underline{e}_i \,|\, \underline{e}_{i+1} \,|\, \cdots \,|\, \underline{e}_n) \, .$$

(II) Vertauschen der i-ten und der j-ten Zeile. Sei $i < j$. Dann entspricht dies der Multiplikation mit der Matrix

$$(\underline{e}_1 \,|\, \cdots \,|\, \underline{e}_{i-1} \,|\, \underline{e}_j \,|\, \underline{e}_{i+1} \,|\, \cdots \,|\, \underline{e}_{j-1} \,|\, \underline{e}_i \,|\, \underline{e}_{j+1} \,|\, \cdots \,|\, \underline{e}_n) \, .$$

(III) Addition von c-mal Zeile i zu Zeile j (mit $i \neq j$). Dies entspricht der Multiplikation mit der Matrix

$$(\underline{e}_1 \,|\, \cdots \,|\, \underline{e}_{i-1} \,|\, \underline{e}_i + c\underline{e}_j \,|\, \underline{e}_{i+1} \,|\, \cdots \,|\, \underline{e}_n) \; .$$

Definition Die obigen Matrizen heißen *Elementarmatrizen.*

Bemerkung Gegeben eine elementare Zeilentransformation, erhält man die entsprechende Elementarmatrix, indem man die Transformation auf die Einheitsmatrix anwendet.

Bemerkung Beachten Sie die die unintuitive Rolle der Indices i und j in (III)!

Bemerkung /Frage Die Elementarmatrizen sind invertierbar, und die inversen Matrizen sind auch Elementarmatrizen. Wie lauten die inversen Matrizen?

Mittels des Gauß-Algorithmus kann man eine Matrix in eine Matrix in reduzierter (Zeilen-)stufenform transformieren. Dies kann man nun wie folgt ausdrücken:

Aussage 2.17 *Sei $A \in K^{m\times n}$. Dann gibt es Elementarmatrizen $E_1, \ldots, E_k \in K^{m\times m}$ so dass $E_k \cdots E_1 A$ eine Matrix in reduzierter (Zeilen-) stufenform ist.*

So eine Matrix $E_k \cdots E_1$ kann man auch leicht algorithmisch ausrechnen. Beachten Sie, dass

$$E_k \cdots E_1 (A \,|\, I_m) = (E_k \cdots E_1 A \,|\, E_k \cdots E_1)$$

gilt.

Wir können demnach so vorgehen: Wir gehen von der Matrix $(A|I_m)$ aus und wenden elementare Zeilentransformationen an, bis wir eine Matrix der Form $(\tilde{A}|M)$ erhalten, wobei $\tilde{A}$ in reduzierter (Zeilen-)stufenform ist. Dann ist M ein Produkt elementarer Matrizen, und es ist $MA = \tilde{A}$.

Wir können nun auch die Aussage, dass elementare Operationen die Lösungsmenge eines LGS nicht ändern, neu beweisen:

Seien $A \in K^{m\times n}$ und $\underline{b} \in K^n$. Sei $M \in K^{m\times m}$ eine Elementarmatrix oder allgemeiner eine invertierbare Matrix. Dann ist offenbar das LGS

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$$

äquivalent zum LGS

$$MA \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = M\underline{b}\,.$$

Die erweiterte Koeffizientenmatrix des ersten LGS ist $(A|\underline{b})$, und die erweiterte Koeffizientenmatrix des zweiten LGS ist $(MA|M\underline{b}) = M(A|\underline{b})$. Man sieht, dass die Multiplikation der erweiterten Koeffizientenmatrix mit M das erste LGS in das äquivalente zweite LGS überführt.

Lemma 2.18 *Sei $A \in K^{m\times n}$. Dann sind die äquivalent:*

- *Λ_A ist injektiv.*
- *Die Spalten von A sind linear unabhängig.*
- *Das homogene LGS $A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{0}$ hat nur die "triviale Lösung" $\underline{0}$.*

Außerdem sind äquivalent:

- *Λ_A ist surjektiv.*
- *Die Spalten von A bilden ein Erzeugendensystem von K^m.*
- *Für alle $\underline{b} \in K^m$ ist das LGS $A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$ lösbar.*

Und es sind äquivalent:

- *Λ_A ist bijektiv.*

- *Die Spalten von A bilden eine Basis von K^m.*
- *Für alle $\underline{b} \in K^m$ hat das LGS $A \cdot \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$ eine eindeutige Lösung.*

Beweis. Λ_A ist genau dann injektiv, wenn $\mathrm{Kern}(A) = \{\underline{0}\}$, und das bedeutet gerade, dass das angegebene homogene LGS nur die "triviale Lösung" $\underline{0}$ hat. Und dies bedeutet gerade, dass die Spalten von A linear unabhängig sind. (Die Null lässt sich nur auf "triviale Weise" als Linearkombination der Spalten darstellen.) Die zweite Aussage folgt aus der Bemerkung zu Beispiel 2.14, und die dritte folgt aus der ersten und zweiten. □

Lemma 2.19 *Seien W, X, Y, Z Mengen, sei $f : X \longrightarrow Y$ eine Abbildung, und seien $g : Y \longrightarrow Z$ und $h : W \longrightarrow X$ bijektive Abbildungen. Dann ist f genau dann injektiv (resp. surjektiv, resp. bijektiv), wenn $g \circ f \circ h : W \longrightarrow Z$ injektiv (resp. surjektiv, resp. bijektiv) ist.*

Eine Anwendung hiervon ist:

Lemma 2.20 *Sei $\varphi : K^n \longrightarrow K^m$ linear, und seien $\psi : K^m \longrightarrow K^m$ und $\chi : K^n \longrightarrow K^n$ bijektive lineare Abbildungen. Dann ist φ genau dann injektiv (resp. surjektiv, resp. bijektiv), wenn $\psi \circ \varphi \circ \chi : K^n \longrightarrow K^m$ injektiv (resp. surjektiv, resp. bijektiv) ist.*

Wenn man dies mit Lemma 2.18 verbindet, erhält man:

Lemma 2.21 *Sei $A \in K^{m \times n}$, und seien $M \in K^{m \times m}, N \in K^{n \times n}$ invertierbar. Dann sind die Spalten von A genau dann linear unabhängig (resp. ein Erzeugendensystem von K^m, resp. eine Basis von K^m), wenn die Spalten von MAN linear unabhängig (resp. ein Erzeugendensystem von K^m, resp. eine Basis von K^m) sind.*

Hieraus folgt insbesondere:

Aussage 2.22 *Sei $M \in K^{m \times m}$ invertierbar, und sei $\tilde{A} = MA$ in (Zeilen-) Stufenform. Dann gilt:*

- *Die Spalten von A sind genau dann linear unabhängig, wenn $\tilde{A}$ nur "Stufenspalten" hat.*
- *Die Spalten von A bilden genau dann ein Erzeugendensystem, wenn $\tilde{A}$ keine Nullzeilen hat.*

- *Die Spalten von A bilden genau dann eine Basis, wenn $m = n$ und jeder Eintrag auf der Diagonalen von $\tilde{A} \neq 0$ ist. (Wenn $\tilde{A}$ in reduzierter (Zeilen-)Stufenform ist, ist dies äquivalent zu $\tilde{A} = I_m$.)*

Hieraus folgen dann nochmal Aussageen 2.12, 2.13 und Satz 2.1.
Wenn man dies nun mit Aussage 2.17 verbindet, folgt:

Satz 2.2 *Sei $A \in K^{m\times n}$. Dann sind äquivalent:*

a) Die Spalten von A bilden eine Basis von K^m.

b) Λ_A ist bijektiv.

c) Sei $M \in K^{m\times m}$ invertierbar so dass MA in reduzierter (Zeilen-)Stufenform ist. Dann ist $MA = I_m$.

d) $m = n$ und $A \in K^{n\times n}$ ist invertierbar.

e) Es gibt Elementarmatrizen $E_1, \ldots, E_k$ so dass $E_k \cdots E_1 A = I_m$.

f) Es gibt Elementarmatrizen $E_1, \ldots, E_k$ mit $A = E_k \cdots E_1$.

Beweis. Aussagen a) und b) sind äquivalent nach Lemma 2.18. Aussagen a) und c) sind äquivalent nach Aussage 2.22.

Wir zeigen nun, dass sowohl Aussage d) als auch Aussage e) jeweils äquialent zu einer der ersten drei Aussagen sind. Danach zeigen wir, dass Aussagen e) und f) äquivalent zueinander sind.

Wenn Aussage c) gilt, ist offensichtlich $A = M^{-1}$, also gilt Aussage d).

Offensichtlich impliziert Aussage d) Aussage b).

Es gelte nun Aussage a). Nach Aussage 2.17 gibt es Elementarmatrizen $E_1, \ldots, E_k$ so dass $E_k \cdots E_1 A$ in reduzierter (Zeilen-)Stufenform ist. Dies bedeutet nach Aussage 2.22, das $E_k \cdots E_1 A = I_m$ gilt, also Aussage e).

Es gelte nun Aussage e). Dann gilt $A = E_1^{-1} \cdots E_k^{-1}$. Da alle E_i invertierbar ist, ist somit auch A invertierbar, d.h. es gilt d).

Da die Inversen der Elementarmatrizen auch Elementarmatrizen sind, impliziert e) auch f).

Nach dem selben Prinzip impliziert Aussage f) Aussage e). □

Gegeben eine Matrix $A \in K^{n\times n}$, kann man das Verfahren zu Aussage 2.17 benutzen, um zu entscheiden, ob A invertierbar ist und ggf. die inverse Matrix bereichnen:

Man formt mittels elementarer Zeilentransformationen die Matrix $(A|I_n)$ um, bis man eine Matrix $(\tilde{A}|M)$ mit $\tilde{A}$ in reduzierter Treppenform erhält. Wenn nun $\tilde{A} = I_n$, ist $M = A^{-1}$. Ansonsten ist A nicht invertierbar. Natürlich kann man den Algorithmus abbrechen, falls man während der Rechnung eine Nullzeile erhält. In diesem Fall ist A nicht invertierbar.

Transponieren

Gegeben eine Matrix $A \in K^{m\times n}$ definieren wir die *transponierte Matrix* A^t durch

$$A^t := ((a_{j,i}))_{i=1,\dots,n,j=1,\dots,m} \in K^{n\times m} .$$

Die transponierte Matrix zu

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \cdot & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

ist also

$$\begin{pmatrix} a_{1,1} & \cdots & a_{m,1} \\ \vdots & \cdot & \vdots \\ a_{1,n} & \cdots & a_{m,n} \end{pmatrix} .$$

Beachten Sie, dass

$$(A+B)^t = A^t + B^t \quad \text{und} \quad (A^t)^t = A$$

für alle $A, B \in K^{m\times n}$.

Lemma 2.23 *Für $A \in K^{m\times n}$ und $B \in K^{n\times r}$ gilt*

$$(AB)^t = B^t A^t .$$

(Das Produkt auf der rechten Seite ist definiert, da $B^t \in K^{r\times n}$ und $A^t \in K^{n\times m}$.)

Beweis. Der Eintrag an der Stelle (i,j) von $(AB)^t$ ist gleich dem Eintrag an der Stelle (j,i) von AB, also $\sum_{\ell=1}^{n} a_{j,\ell} b_{\ell,i}$.

Andererseits ist der Eintrag an der Stelle (i,j) von $B^t A^t$ per Definition gleich $\sum_{\ell=1}^{n} b_{\ell,i} a_{j,\ell}$.

Damit sind die beiden Einträge gleich. □

Wir erhalten insbesondere:

Aussage 2.24 *Eine Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn A^t invertierbar ist. In diesem Fall ist $(A^t)^{-1} = (A^{-1})^t$.*

Beweis. Sei A invertierbar. Dann gilt $A^t(A^{-1})^t = (A^{-1}A)^t = I_n^t = I_n$ und $(A^{-1})^t A^t = (AA^{-1})^t = I_n^t = I_n$. Damit gilt ist per Definition A invertierbar, und $(A^{-1})^t$ ist das Inverse von A^t.

Die Rückrichtung folgt auch, da $(A^t)^t = A$ ist. □

Spaltentransformationen

Analog zu elementaren Zeilentransformationen kann man eine Matrix auch mittels *elementarer Spaltentransformationen* umformen.

Diese Umformungen sind:

(I) Multiplikation einer Spalte mit einem Skalar $\neq 0$.
(II) Vertauschen von zwei Spalten.
(III) Addition des c-fachen der i-ten Spalte zur j-ten Spalte (für ein $c \in K$ und $i \neq j$).

Jede dieser drei Umformungen kann man erhalten, indem man die Matrix von *rechts* mit einer bestimmten invertierbaren Matrix multipliziert.

Frage Welche Matrizen sind dies genau?

Es gibt einen Zusammenhang zwischen Zeilentransformationen und Spaltentransformationen, der sich durch das Transponieren ergibt: Anstatt eine Spaltentransformation durchzuführen, kann man auch die Matrix transponieren, die "entsprechende" Zeilentransformation durchführen und dann wieder transponieren.

Beispiel 2.25 Sei $A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$. Wenn wir 2-mal die erste Spalte von der zweiten abziehen, erhalten wir $\begin{pmatrix} 1 & 0 & 3 \\ 4 & -3 & 6 \end{pmatrix}$. Wenn wir andererseits A transponieren, dann 2-mal die erste Zeile von der zweiten abziehen und dann wieder transponieren, erhalten wir der Reihe nach

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & 4 \\ 0 & -3 \\ 3 & 6 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 3 \\ 4 & -3 & 6 \end{pmatrix}.$$

Beachten Sie auch: Die Spaltentransformation entspricht der Multplikation mit der Matrix

$$M := \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

von rechts, die Zeilentransformation entspricht der Multpiplikation mit der Matrix

$$M^t = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

von links. (Man kann das Beispiel auch durch die Gleichung $AM = (M^t A^t)^t$ beschreiben; diese folgt aus Lemma 2.23.)

Analog zur Definition einer Matrix in (reduzierter) Zeilenstufenform definiert man, was eine Matrix in (reduzierter) Spaltenstufenform ist. Wir machen es uns einfach und definieren:

Definition Eine Matrix A ist in *Spaltenstufenform* (resp. in *reduzierter Spaltenstufenform*), wenn A^t in Zeilenstufenform (resp. in reduzierter Zeilenstufenform) ist.

Beachten Sie, dass die Stufen einer Matrix in Spaltenstufenform die *Breite* 1 haben.

Wir kommen nun zu einer *Anwendung*.

Seien $\underline{x}_1, \ldots, \underline{x}_r \in K^n$. Wir wollen eine Basis von $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$ bestimmen. Ein Verfahren hierfür haben wir bereits in Abschnitt 2.2 kennen gelernt. Mit dem dort beschriebenen Verfahren kann man eine Basis finden, die ein Teilsystem von $\underline{x}_1, \ldots, \underline{x}_r$ ist.

Mit dem hier beschriebenen Verfahren findet man hingegen eine besonders "schöne" Basis.

Wir betrachten die Matrix $(\underline{x}_1 | \cdots | \underline{x}_r)$. Nehmen wir an, wir haben eine elementare *Spalten*transformation durchgeführt und haben nun die Matrix $(\tilde{\underline{x}}_1 | \cdots | \tilde{\underline{x}}_r)$. Dann liegen (offensichtlich) alle $\tilde{\underline{x}}_i$ in $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$. Umgekert gilt $\underline{x}_i \in \langle \tilde{\underline{x}}_1, \ldots, \tilde{\underline{x}}_r \rangle_K$, da die Transformation (mittels einer anderen elementaren Spaltentransformation) rückgängig gemacht werden kann. Damit gilt:

$$\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K = \langle \tilde{\underline{x}}_1, \ldots, \tilde{\underline{x}}_r \rangle_K \tag{2.13}$$

Mit anderen Worten, der von den Spalten aufgespannte lineare Unterraum ist invariant unter elementaren Spaltentransformationen.

Das Verfahren ist nun wie folgt: Wir transformieren die Matrix $(\underline{x}_1 | \cdots | \underline{x}_r)$ mittels elementarer Spaltentransformationen in eine Matrix $\tilde{A}$ in Spaltenstufenform. Die nicht-Nullspalten von $\tilde{A}$ sind dann offensichtlich linear unabhängig und ein Erzeugendensystem des von den Spalten aufgespannten Raums. Damit bilden diese eine Basis von $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$.

Eine besonders "schöne Basis" erhält man, indem man bis zu einer Matrix in reduzierter Spaltenstufenform weiterrechnet.

Bemerkung Aufgrund des Zusammenhangs zwischen Spalten- und Zeilentransformationen unter Transponieren kann man auch so vorgehen: Man transponiert die Vektoren $\underline{x}_1, \ldots, \underline{x}_r$ und schreibt diese als *Zeilen* in eine Matrix. Dann führt man den Gauß-Algorithmus durch (elementare Zeilenumformungen). Man betrachtet nun die nicht-Nullzeilen. Wenn man diese transponiert, erhält man eine gesuchte Basis von $\langle \underline{x}_1, \ldots, \underline{x}_r \rangle_K$.

Bemerkung Die Invarianz (2.13) kann man mittels Matrizenmultiplikation auch so zeigen: Wie gesagt korrespondiert jede elementare Spaltentransformation zu einer Multiplikation mit einer bestimmten invertierbaren Matrix von rechts.

Sei $A \in K^{n\times r}$ die Ausgangsmatrix und M eine beliebige invertierbare $r \times r$-Matrix. Dann ist

$$\text{Bild}(\Lambda_{AM}) = \text{Bild}(\Lambda_A \circ \Lambda_M) = \text{Bild}(\Lambda_A) \ ,$$

da $\Lambda_M : K^r \longrightarrow K^r$ surjektiv (sogar bijektiv) ist. Es ist aber $\text{Bild}(\Lambda_A)$ der von den Spalten von A aufgespannte Raum und $\text{Bild}(\Lambda_{AM})$ der von den Spalten von AM aufgespannte Raum.

2.5 Vektorräume

Sei nach wie vor K ein Körper. Das Konzept eines *Vektorraums* ist eine Verallgemeinerung des K^n.

Definition Ein K-*Vektorraum* (oder ein *Vektorraum über* K) ist eine abelsche Gruppe $(V, +)$ zusammen mit einer Abbildung (genannt *Skalarmultiplikation* (von K auf V))

$$\cdot : \ K \times V \longrightarrow V, (a, \mathfrak{x}) \mapsto a \cdot \mathfrak{x}$$

so dass

- $\forall a \in K \, \forall \mathfrak{x}, \mathfrak{y} \in V \ : a(\mathfrak{x} + \mathfrak{y}) = a\mathfrak{x} + a\mathfrak{y}$
- $\forall a, b \in K \, \forall \mathfrak{x} \in V \ : a(b\mathfrak{x}) = (ab)\mathfrak{x}$
- $\forall a, b \in K \, \forall \mathfrak{x} \in V \ : (a + b)\mathfrak{x} = a\mathfrak{x} + b\mathfrak{x}$
- $\forall \mathfrak{x} \in V \ : 1_K \cdot \mathfrak{x} = \mathfrak{x}$

Hier haben wir (wie üblich) die Multiplikationspunkte weggelassen und die Regel ("Mal vor Plus") angewandt.

Notation Das Nullelement benzeichnen wir mit $\mathfrak{o}$.

Sprachgebrauch Die Elemente eines Vektorraums heißen *Vektoren*, und die Elemente dem Grundkörper *Skalare*. Beachten Sie, dass der Begriff eines Vektorraums abstrakt ist; insbesondere sind die Elemente eines Vektorraums nicht notwendigerweise in irgendeinem anschaulichen Sinn Vektoren.

Bemerkung / Frage Es gilt immer $-\mathfrak{x} = (-1) \cdot \mathfrak{x}$. Warum?

Einige Beispiele:

- Für alle $n \in \mathbb{N}$ ist K^n mit der üblichen Addition und Skalarmultiplikation ein K-Vektorraum.
- Die "triviale Gruppe" $\{0\}$ ist in offensichtlicher Weise ein K-Vektorraum. (Man setzt $K^0 := \{0\}$.)
- Der Körper K ist mit seiner Addition und Multiplikation ein K-Vektorraum. ("Im Wesentlichen" ist $K = K^1$, siehe unten.)
- Für alle $n, m \in \mathbb{N}$ ist die Menge der $m \times n$-Matrizen $K^{m \times n}$ mit der üblichen Addition und der auf S. 93 definierten Skalarmultiplikation ein K-Vektorraum.

Und noch zwei wichtige Beispiele:

Beispiel 2.26 Sei L ein Körper und $K \subseteq L$ ein Unterkörper. Dann ist L in "natürlicher Weise" ein K-Vektorraum, und zwar wie folgt: $(L, +)$ ist eine abelsche Gruppe, und wenn wir die Multiplikation $\cdot : L \times L \longrightarrow L$ auf $K \times L$ einschränken, erhalten wir eine Skalarmultiplikation von K auf L.

Es ist offensichtlich, dass die vier Vektorraumaxiome gelten.

Beispiel 2.27 Der Polynomring $K[X]$ ist mit der üblichen Addition und der Skalarmultiplikation $K \times K[X] \longrightarrow K[X]\,, (c, \sum_{i=0}^{n} a_i X^i) \mapsto \sum_{i=0}^{n} c a_i X^i$ ein K-Vektorraum.

Aus Abschnitt 1.8 wissen Sie, dass man, immer wenn eine neue Art von mathematischen "Objekten" eingeführt wird, fragen sollte: *Wie lauten die zugehörigen Morphismen?*

Die Morphismen sind hier die *linearen Abbildungen*, die genau so definiert sind wie zuvor:

Definition Seien V und W K-Vektorräume. Eine K-*lineare Abbildung* (kurz: eine *lineare Abbildung*) von V nach W ist eine Abbildung $\varphi : V \longrightarrow W$ so dass

- $\forall \mathfrak{x}, \mathfrak{y} \in V : \varphi(\mathfrak{x} + \mathfrak{y}) = \varphi(\mathfrak{x}) + \varphi(\mathfrak{y})$
- $\forall a \in K\ \forall \mathfrak{x} \in V : \varphi(a\mathfrak{x}) = a\varphi(\mathfrak{x})$.

Bemerkung Seien V, W K-Vektorräume. Dann ist eine Abbildung $\varphi : V \longrightarrow W$ genau dann linear, wenn für alle $a \in K$ und alle $\mathfrak{x}, \mathfrak{y} \in V$ $\varphi(a\mathfrak{x} + \mathfrak{y}) = a\varphi(\mathfrak{x}) + \varphi(\mathfrak{y})$ gilt.

Lemma 2.28

a) Wenn $\varphi : V \longrightarrow W$ und $\psi : W \longrightarrow X$ lineare Abbildungen von K-Vektoräumen sind, dann auch $\psi \circ \varphi : V \longrightarrow X$.

b) Wenn $\varphi, \psi : V \longrightarrow W$ lineare Abbildungen von K-Vektorräumen sind, dann auch $\varphi + \psi : V \longrightarrow W, \mathfrak{x} \mapsto \varphi(\mathfrak{x}) + \psi(\mathfrak{x})$.

c) Wenn $a \in K$ und $\varphi : V \longrightarrow W$ eine lineare Abbildungen von K-Vektorräumen ist, dann auch $a\varphi : V \longrightarrow W, \mathfrak{x} \mapsto a\varphi(\mathfrak{x})$.

d) Wenn $\varphi : V \longrightarrow W$ eine bijektive lineare Abbildung von K-Vektorräumen ist, dann ist auch $\varphi^{-1} : W \longrightarrow V$ linear.

Die K-linearen Abbildungen nennt man auch von *Homomorphismen* von K-Vektorräumen. (Entsprechend der Philosophie von Abschnitt 1.8.

Dementsprechend bezeichnet man die Menge der K-linearen Abbildungen von V nach W mit $\mathrm{Hom}_K(V, W)$ (oder einfach mit $\mathrm{Hom}(V, W)$, wenn es klar ist, dass man "über dem Körper K" arbeitet). Insbesondere ist also $\mathrm{Hom}_K(K^n, K^m)$ die Menge der linearen Abbildungen von K^n nach K^m (die bisher keine Bezeichnung hatte). Beachten Sie, dass $\mathrm{Hom}_K(V, W)$ mit der in Lemma 2.28 b) definierten Addition eine abelsche Gruppe ist.

Ferner ist $\mathrm{Hom}_K(V, W)$ mit dieser Addition und der in Lemma 2.28 c) definierten Skalarmultiplikation ein K-Vektorraum (überprüfen Sie dies!).

Wir wenden nun auch die anderen Begriffe aus Abschnitt 1.8 auf lineare Abbildungen an:

- Ein *Isomorphismus* von V nach W (beides K-Vektorräume) ist eine lineare Abbildung $\varphi : V \longrightarrow W$ so dass es eine lineare Abbildung $\psi : W \longrightarrow V$ mit $\psi \circ \varphi = \mathrm{id}_V$ und $\varphi \circ \psi = \mathrm{id}_W$ existiert. (Nach Lemma 2.28 d) ist dies äquivalent dazu, dass φ eine bijektive lineare Abbildung ist.) Wenn es einen Isomorphismus zwischen V und W gibt, heißen V und W *isomorph*.

- Ein *Endomorphismus* eines K-Vektorraums V ist eine lineare Abbildung von V nach V.

- Ein *Automorphismus* eines K-Vektorraums V ist ein Isomorphismus von V nach V.

Die Bezeichnungen für die entsprechenden Mengen sind $\mathrm{Iso}_K(V,W), \mathrm{End}_K(V)$ und $\mathrm{Aut}_K(V)$, wobei das K auch weggelassen werden kann. (Beachten Sie, dass die letzeren beiden Bezeichungen analog zu den Bezeichnungen in Abschnitt 1.8 sind.)

Es ist klar, dass $\mathrm{End}_K(V)$ mit der schon diskutierten Addition und der Verknüpfung als Multiplikation ein Ring ist, und $\mathrm{Aut}_K(V)$ ist mit der Verknüpfung eine Gruppe. Dabei ist $\mathrm{Aut}_K(V)$ gleich der Gruppe der (bez. der Multiplikation) invertierbaren Elemente von $\mathrm{End}_K(V)$, also $\mathrm{End}_K(V)^* = \mathrm{Aut}_K(V)$ (vergleiche die letzte Aussage mit Aussage 1.63). Aussage 2.16 besagt, dass die Abbildungen $\mathrm{End}_K(K^n) \longrightarrow K^{n\times n}, \varphi \mapsto M(\varphi)$ und $K^{n\times n} \longrightarrow \mathrm{End}_K(K^n), A \mapsto \Lambda_A$ zueinander inverse Isomorphismen von Ringen sind.

Und nun noch einige Beispiele für Isomorphismen von K-Vektorräumen:

Beispiel 2.29 Die Abbildung $K \longrightarrow K^1, x \mapsto (x)$ (wobei (x) das "1-Tupel" ist, dass x enthält) ist ein Isomorphismus von K-Vektorräumen. (Wir identifizieren diese beiden Vektorräume.)

Beispiel 2.30 Seien $m, n \in \mathbb{N}$. Dann ist das Transponieren $K^{m\times n} \longrightarrow K^{n\times m}, A \mapsto A^t$ ein Isomorphismus von K-Vektorräumen. Insbesondere ist also der Raum der *Zeilenvektoren* der Länge n isomorph zum Raum der Spaltenvektoren der Länge n (über K).

Beispiel 2.31 Die lineare Abbildung

$$K^{m\times n} \longrightarrow K^{m\cdot n}, A = ((a_{i,j}))_{i=1,\ldots,m,j=1,\ldots,n} \mapsto \underline{y} \text{ mit } y_{(j-1)\cdot m+i} := a_{i,j}$$

für alle $i = 1,\ldots,m, j = 1,\ldots,n$ ist ein Isomorphismus von K-Vektorräumen.

Beispiel 2.32 Die Abbildungen $\mathrm{Hom}_K(K^m, K^n) \longrightarrow K^{m\times n}, \varphi \mapsto M(\varphi)$ und $K^{m\times n} \longrightarrow \mathrm{Hom}_K(K^m, K^n), A \mapsto \Lambda_A$ sind zueinander inverse Isomorphismen von K-Vektorräumen.

Definition Sei V ein K-Vektorraum. Dann heißen die linearen Abbildungen $V \longrightarrow K$ *Linearformen* auf V.

Das obige Beispiel besagt insbesondere, dass der Raum $\mathrm{Hom}_K(K^n, K)$ der Linearformen auf K^n isomorph zum Raum der *Zeilenvektoren* der Länge n über K ist.

Bemerkung Wir haben gesehen, dass das Transponieren einen Isomorphismus zwischen dem Raum der Zeielenvektoren der Länge n und dem Raum der Spaltenvektoren der Länge n definiert. Man könnte sagen, dass es sowieso egal ist, ob man einen Vektor als Zeile oder als Spalte schreibt. Man sollte aber *nicht* beliebig zwischen Spalten- und Zeilenvektoren hin- und herwechseln: Viele Argumente sind viel übersichtlicher, wenn man die Elemente von K^n mit Spalten beschreibt und Zeilenvektoren benutzt, um Linearformen anzugeben.

Bemerkung / Definition Sei V ein K-Vektorraum, und sei $U \subseteq V$ mit $\mathfrak{o} \in U$ so dass U abgeschlossen unter Addition und Skalarmultiplikation ist (d.h. $\forall a \in K \, \forall \mathfrak{x}, \mathfrak{y} \in U : \mathfrak{x} + \mathfrak{y} \in U, a\mathfrak{x} \in U$). Dann ist U mit der induzierten Addition und Skalarmultiplikation ein K-Vektorraum; genannt ein *Untervektorraum* von V.

Bemerkung Die Untervektorräume des K^n sind genau die linearen Unterräume.

Lemma 2.33 *Seien V, W K-Vektorräume, $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann ist* Kern(φ) *ein Untervektorraum von V, und* Bild(φ) *ist ein Untervektorraum von W.*

Der *Beweis* ist offensichtlich. □

2.6 Endlich erzeugte Vektorräume

Systeme von Vektoren

Sei V ein K-Vektorraum. Sei $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ ein endliches System von Vektoren aus V. Dann definiert man in offensichtlicher Verallgemeinerung der Definitionen in Abschnitt 2.1:

Definition

- $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ bilden ein *Erzeugendensystem* von V, wenn gilt: Für alle $\mathfrak{x} \in V$ gibt es $a_1, \ldots, a_r \in K$ mit $\mathfrak{x} = a_1\mathfrak{x}_1 + \cdots a_r\mathfrak{x}_r$.
- $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ bilden eine *Basis* von V, wenn gilt: Für alle $\mathfrak{x} \in V$ gibt es eindeutig bestimmte $a_1, \ldots, a_r \in K$ mit $\mathfrak{x} = a_1\mathfrak{x}_1 + \cdots a_r\mathfrak{x}_r$.

Beispiel 2.34 Wie wir schon wissen, bilden die Standardvektoren $\underline{e}_1, \ldots, \underline{e}_r$ eine Basis von K^r. Man spricht von der *Standardbasis* des K^r.

Man definiert den von $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *erzeugten* (oder *aufgespannten*) *Untervektorraum* als

$$\langle \mathfrak{x}_1, \ldots, \mathfrak{x}_r \rangle_K := \{ a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r \mid a_1, \ldots, a_r \in K \} .$$

Dies ist nun der kleinste Untervektorraum von V (bez. der Inklusion), der $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ enthält.

Aussage 2.35 *Sei $\mathfrak{b}_1, \ldots, \mathfrak{b}_r \in V$ eine Basis. Sei W ein weiterer K-Vektorraum, und seien $\mathfrak{x}_1, \ldots, \mathfrak{x}_r \in W$. Dann gibt es genau eine lineare Abbildung $\varphi : V \longrightarrow W$ mit $\varphi(\mathfrak{b}_i) = \mathfrak{x}_i$ für alle $i = 1, \ldots, r$, und diese ist durch $\varphi(a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r) = a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$ für $a_1, \ldots, a_r \in K$ gegeben. Dabei gilt* $\mathrm{Bild}(\varphi) = \langle \mathfrak{x}_1, \ldots, \mathfrak{x}_r \rangle_K$.

Beweis. Wir zeigen zuerst die Eindeutigkeit. Sei $\mathfrak{x} \in V$. Dann gibt es (eindeutig bestimmte) $a_1, \ldots, a_r \in K$ mit $\mathfrak{x} = a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r$. Nun ist $\varphi(\mathfrak{x}) = \varphi(a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r) = a_1\varphi(\mathfrak{b}_1) + \cdots + a_r\varphi(\mathfrak{b}_r) = a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$.

Nun müssen wir noch nachweisen, dass es tatsächlich so eine lineare Abbildung gibt. Sei dazu wiederum $\mathfrak{x} \in V$ beliebig. Wie schon gesagt gibt es eindeutig bestimmte $a_1, \ldots, a_r \in K$ mit $\mathfrak{x} = a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r$. Wir setzen nun $\varphi(\mathfrak{x}) := a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$. Dies ist wohldefiniert, da die "Koeffizienten" $a_1, \ldots, a_r$ eindeutig sind.

Wir müssen noch die Linearität nachweisen.

Seien dazu $\mathfrak{x}, \mathfrak{y} \in V$ und $c \in K$. Sei $\mathfrak{x} = a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r$ und $\mathfrak{y} = b_1\mathfrak{b}_1 + \cdots + b_r\mathfrak{b}_r$.

Dann ist $\varphi(c \cdot \mathfrak{x} + \mathfrak{y}) = \varphi(ca_1\mathfrak{b}_1 + \cdots + ca_r\mathfrak{b}_r + b_1\mathfrak{b}_1 + \cdots + b_r\mathfrak{b}_r) = \varphi((ca_1 + b_1)\mathfrak{b}_1 + \cdots + (ca_r + b_r)\mathfrak{b}_r) \overset{\text{per Def.}}{=} (ca_1 + b_1)\mathfrak{x}_1 + \cdots + (ca_r + b_r)\mathfrak{x}_r = c(a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r) + (b_1\mathfrak{x}_1 + \cdots + b_r\mathfrak{x}_r) \overset{\text{per Def.}}{=} c\varphi(a_1\mathfrak{b}_1 + \cdots + a_r\mathfrak{b}_r) + \varphi(b_1\mathfrak{b}_1 + \cdots + b_r\mathfrak{b}_r) = c\varphi(\mathfrak{x}) + \varphi(\mathfrak{y})$.

Die Aussage über das Bild folgt sofort aus der Definition von φ. □

Als Spezialfall des obigen Aussage erhalten wir: Sei V ein K-Vektorraum, und seien $\mathfrak{x}_1, \ldots, \mathfrak{x}_k \in V$. Dann gibt es genau eine lineare Abbildung $\varphi : K^r \longrightarrow V$ mit $\varphi(\underline{e}_i) = \mathfrak{x}_i$ für $i = 1, \ldots, r$. Explizit ist diese Abbildung durch

$$\varphi\left(\begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}\right) = a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$$

gegeben.

Offensichtlich sind nun die folgenden Aussagen jeweils äquivalent:

- $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ ist ein Erzeugendensystem von V.

- $\langle \mathfrak{x}_1, \ldots, \mathfrak{x}_r \rangle_K = V$
- Die eindeutig bestimmte lineare Abbildung $\varphi : K^r \longrightarrow V$ mit $\varphi(\underline{e}_i) = \mathfrak{x}_i$ ist surjektiv.

Sowie:

- $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ ist eine Basis von V.
- Die eindeutig bestimmte lineare Abbildung $\varphi : K^r \longrightarrow V$ mit $\varphi(\underline{e}_i) = \mathfrak{x}_i$ ist bijektiv (d.h. ein Isomorphismus).

Definition Sei $\mathfrak{x} \in V$. Dann ist $\mathfrak{x}$ *linear abhängig* von $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$, wenn es $a_1, \ldots, a_r \in K$ mit $\mathfrak{x} = a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$ gibt. Wenn dies nicht der Fall ist, heißt $\mathfrak{x}$ *linear unabhängig* von $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$.

Definition Die Vektoren $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ heißen *linear unabhängig*, wenn keiner der Vektoren von den anderen linear abhängig ist.

Wir haben die folgende offensichtliche Verallgemeinerung von Lemma 2.7:

Lemma 2.36 *Die Vektoren $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ sind genau dann linear unabhängig, wenn gilt:*

$$\forall a_1, \ldots, a_r \in K : a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r = 0 \longrightarrow a_1 = \cdots = a_r = 0 \, .$$

Somit sind die folgenden beiden Aussagen äquivalent:

- Die Vektoren $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ sind linear unabhängig.
- Die eindeutig bestimmte lineare Abbildung $\varphi : K^r \longrightarrow V$ mit $\varphi(\underline{e}_i) = \mathfrak{x}_i$ ist injektiv.

Auch die Definitionen von "maximales linear unabhängiges System" und "minimales Erzeugendensystem" verallgemeinern sich sofort. Damit haben wir:

Aussage 2.37 *Die folgenden Aussagen sind äquivalent:*

a) $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *ist eine Basis von* V.

b) $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *ist ein linear unabhängiges Erzeugendensystem.*

c) $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *ist ein maximales linear unabhängiges System.*

d) $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *ist ein minimales Erzeugendensystem.*

e) Die eindeutig bestimmte lineare Abbildung $\varphi : K^r \longrightarrow V$ *mit* $\varphi(\underline{e}_i) = \mathfrak{x}_i$ *für alle* $i = 1, \ldots, r$ *ist ein Isomorphismus.*

Der *Beweis* der Äquivalenzen von Aussage 2.8 gilt wörtlich auch hier. Oben haben wir bereits gesehen, dass a) und e) äquivalent sind. □

Lemma 2.38 *Seien* V, W K*-Vektorräume, seien* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r \in V$, *und sei* $\varphi : V \longrightarrow W$ *eine lineare Abbildung. Dann gilt:*

a) Wenn $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *ein Erzeugendensystem von* V *bilden, dann bilden* $\varphi(\mathfrak{x}_1), \ldots, \varphi(\mathfrak{x}_r)$ *ein Erzeugendensystem von* $\mathrm{Bild}(V)$. *Wenn also zusätzlich* φ *surjektiv ist, bilden sie auch ein Erzeugendensystem von* W.

b) Wenn $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *linear unabhängig sind und* φ *injektiv ist, dann sind auch* $\varphi(\mathfrak{x}_1), \ldots, \varphi(\mathfrak{x}_r)$ *linear unabhängig.*

c) Wenn φ *ein Isomorphismus ist, dann bilden* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *genau dann eine Basis von* V, *wenn* $\varphi(\mathfrak{x}_1), \ldots, \varphi(\mathfrak{x}_r)$ *eine Basis von* W *bilden.*

(Das ist leicht.)

Aussage 2.39 *Sei* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ *ein Ereugendensystem von* V. *Dann gibt es ein Teilsystem von* $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$, *das eine Basis von* V *ist.*

Beweisskizze. Sei $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ ein Erzeugendensystem. Dann gibt es zwei Fälle: Entweder $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ ist linear unahbängig, also eine Basis, oder einer der Vektoren ist linear abhängig von den anderen. Dann lassen wir diesen Vektor weg. Wenn man diese Prozedur iteriert, erhält man irgendwann ein minimales Erzeugendensystem, also eine Basis.

(In einem formalen Beweis sollte man vollständige Induktion benutzen.) □

Definition Der Vektorraum V heißt *endlich erzeugt*, wenn er ein Erzeugendensystem $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ hat.

Wir werden uns im Folgenden fast ausschließlich auf endlich erzeugte Vektorräume beschränken.

Nach Aussage 2.39 gilt insbesondere:

Aussage 2.40 *Sei* V *endlich erzeugt. Dann hat* V *eine Basis.*

Satz 2.3 *Sei* V *endlich erzeugt, und sei* $\mathfrak{x}_1, \ldots, \mathfrak{x}_n$ *eine Basis. Dann gilt:*

- *Jedes Erzeugendensystem von* V *enthält mindestens* n *Vektoren.*

- *Jedes linear unabhängige System von V enthält höchstens n Vektoren.*
- *Jede Basis von V enthält genau n Vektoren.*

Beweis. Sei $\varphi : K^n \longrightarrow V$ die eindeutig bestimmte lineare Abbildung mit $\varphi(\underline{e}_i) = \mathfrak{x}_i$ für alle $i = 1, \ldots, n$. Da $\mathfrak{x}_1, \ldots, \mathfrak{x}_n$ eine Basis ist, ist dies ein Isomorphismus.

Seien $\mathfrak{y}_1, \ldots, \mathfrak{y}_r \in V$.

Wenn die Vektoren $\mathfrak{y}_1, \ldots, \mathfrak{y}_r$ ein Erzeugendensystem von V bilden, dann bilden $\varphi^{-1}(\mathfrak{x}_1), \ldots, \varphi^{-1}(\mathfrak{x}_r)$ ein Erzeugendensystem von K^n. Damit gilt nach Aussage 2.13 $r \geq n$.

Wenn $\mathfrak{y}_1, \ldots, \mathfrak{y}_r$ linear unabhängig sind, dann sind auch $\varphi^{-1}(\mathfrak{x}_1), \ldots, \varphi^{-1}(\mathfrak{x}_r)$ linear unabhängig. Damit gilt nach Aussage 2.12 $r \leq n$.

Aus diesen beiden Aussagen folgt die dritte. (Oder man benutzt Satz 2.1.) □

Definition Sei V ein endlich erzeugter Vektorraum. Dann heißt die Anzahl der Elemente einer Basis (jeder Basis) die *Dimension* von V. Bezeichnung: $\mathrm{Dim}(V)$ oder $\mathrm{Dim}_K(V)$.

Wenn V nicht endlich erzeugt ist, sagt man auch, dass V *unendliche Dimension* hat, und man schreibt $\mathrm{Dim}(V) = \infty$.

Beispiel 2.41 Der Raum K^n hat natürlich die Dimension n; man spricht auch vom *n-dimensionalen Standardvektorraum über K.*

Beispiel 2.42 Sei wie immer K ein Körper, und sei $f(X) \in K[X]$ mit $\mathrm{Grad}(f(X)) = d$. Dann ist K mittels der Inklusion $K \hookrightarrow K[X]/(f(X))$, $c \mapsto [c]_{(f(X))}$ ein Unterring von $K[X]/(f(X))$, und somit ist $K[X]/(f(X))$ in natürlicher Weise ein K-Vektorraum. Lemma 1.74 besagt, dass $1, [X], [X]^2, \ldots, [X]^{d-1}$ eine Basis dieses Vektorraums ist. Also ist dieser Vektorraum d-dimensional.

Beispielsweise ist der Körper $\mathbb{C} := \mathbb{R}[X]/(X^2 + 1)$ ein $\mathbb{R}$-Vektorraum, und $1, i = [X]_{(X^2+1)}$ ist eine Basis dieses Vektorraums. Damit ist $\mathbb{C}$ ein 2-dimensionaler $\mathbb{R}$-Vektorraum.

Aus Lemma 2.38 folgt:

Lemma 2.43 *Seien V und W K-Vektorräume, wobei V endlich erzeugt ist, und sei $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann gilt:*

a) $\mathrm{Bild}(\varphi)$ *ist auch endlich erzeugt, und es ist* $\mathrm{Dim}(\mathrm{Bild}(\varphi)) \leq \mathrm{Dim}(V)$. *Wenn φ darüber hinaus surjektiv ist, ist auch W endlich erzeugt, und es ist* $\mathrm{Dim}(W) \leq \mathrm{Dim}(V)$.

b) Wenn φ injektiv ist, gilt $\mathrm{Dim}(\mathrm{Bild}(V)) = \mathrm{Dim}(V)$. *Wenn φ ein Isomorphismus ist, gilt* $\mathrm{Dim}(V) = \mathrm{Dim}(W)$.

Satz 2.4 *Seien V, W zwei endlich erzeugte K-Vektorräume. Dann gilt: Es gibt genau dann einen Isomorphismus zwischen V und W, wenn* $\mathrm{Dim}(V) = \mathrm{Dim}(W)$.

Beweis. Wir wissen schon, dass $\mathrm{Dim}(V) = \mathrm{Dim}(W)$, wenn φ ein Isomorphismus ist.

Sei also $\mathrm{Dim}(V) = \mathrm{Dim}(W) =: n$. Sei $\mathfrak{x}_1, \ldots, \mathfrak{x}_n$ eine Basis von V, und sei $\mathfrak{y}_1, \ldots, \mathfrak{y}_n$ eine Basis von W. Dann haben wir eine eindeutig bestimmte lineare Abbildung $\varphi : V \longrightarrow W$ mit $\varphi(\mathfrak{x}_i) = \mathfrak{y}_i$ für alle $i = 1, \ldots, n$, und wir haben eine eindeutig bestimmte lineare Abbildung $\psi : W \longrightarrow V$ mit $\psi(\mathfrak{y}_i) = \mathfrak{x}_i$ für alle $i = 1, \ldots, n$. Nun ist $\psi \circ \varphi : V \longrightarrow V$ die eindeutig bestimmte lineare Abbildung mit $(\psi \circ \varphi)(\mathfrak{x}_i) = \mathfrak{x}_i$ für alle $i = 1, \ldots, n$, also $(\psi \circ \varphi) = \mathrm{id}_V$. Andererseits ist $\varphi \circ \psi : W \longrightarrow W$ die eindeutig bestimmte lineare Abbildung mit $(\varphi \circ \psi)(\mathfrak{y}_i) = \mathfrak{y}_i$ für alle $i = 1, \ldots, n$, also $\varphi \circ \psi = \mathrm{id}_W$.

Damit ist also $\varphi : V \longrightarrow W$ ein Isomorphismus. □

Aussage 2.44 *Sei V endlich erzeugt. Dann kann man jedes linear unabhängige System in V zu einer Basis von V ergänzen. D.h.: Seien $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ linear unabhängige Vektoren in V. Dann kann man $\mathfrak{x}_{r+1}, \ldots, \mathfrak{x}_s \in V$ finden, so dass $\mathfrak{x}_1, \ldots, \mathfrak{x}_s$ eine Basis von V bilden.*

Beweisskizze. Man geht wie folgt vor: Wenn das System eine Basis ist, ist man fertig. Wenn dies nicht der Fall ist, gibt es einen von $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ linear unabhängigen Vektor. Man fügt so einem Vektor zu dem System hinzu. Dies iteriert man. Der Prozess terminiert nach genau $\mathrm{Dim}(V) - r$ Schritten. □

Aussage 2.45 *Sei V endlich erzeugt, und sei $U \subseteq V$ ein Untervektorraum. Dann hat ist auch U endlich erzeugt, hat also auch eine Basis.*

Beweis. Angefangen von leeren System geht man analog zum obigen Beweis vor: Wenn immer es einen von dem System linear unabhängigen Vektor in U gibt, ergänzt man das System mit so einem Vektor. Dieser Prozess muss terminieren, da jedes linear unabhängige System in V höchstens $\mathrm{Dim}(V)$ Vektoren entält. Dann hat man ein maximales linear unabhängiges System von U, also eine Basis von U. □

Definition Sei A eine Matrix über K. Die Dimension des von den Spalten von A erzeugten Vektorraums heißt der *Spaltenrang* von A. Die Dimension des von den Zeilen von A erzeugten Vektorraums heißt der *Zeilenrang* von A.

Bemerkung Der Zeilenrang von A ist gleich dem Spaltenrang von A^t (und umgekehrt).

Ziel ist nun der Beweis des folgenden Satzes.

Satz 2.5 *Sei A eine Matrix über K. Dann ist der Spaltenrang von A gleich dem Zeilenrang von A.*

Lemma 2.46 *Sei $A \in K^{m\times n}$ eine beliebige Matrix, und seien $M \in K^{m\times m}$ und $N \in K^{n\times n}$ invertierbare Matrizen. Dann ist der Spaltenrang von A gleich dem Spaltenrang von MAN, und der Zeilenrang von A ist gleich dem Zeilenrang von MAN.*

Beweis. Wir zeigen zunächst die Aussage für den Spaltenrang. Der Spaltenrang von A (resp. MAN) ist per Definition gleich der Dimension von $\mathrm{Bild}(\Lambda_A)$ (resp. $\mathrm{Bild}(\Lambda_{MAN})$). Nun ist $\mathrm{Bild}(\Lambda_{MAN}) = \mathrm{Bild}(\Lambda_M \circ \Lambda_A \circ \Lambda_N) = \mathrm{Bild}(\Lambda_M \circ \Lambda_A)$, da Λ_N surjektiv (sogar bijektiv) ist. Außerdem ist

$$\mathrm{Dim}(\mathrm{Bild}(\Lambda_M \circ \Lambda_A)) = \mathrm{Dim}(\mathrm{Bild}(\Lambda_A)) \ ,$$

da $\Lambda_M|_{\mathrm{Bild}(\Lambda_A)}$ injektiv (sogar bijektiv) ist (siehe Lemma 2.43).

Die Aussage über den Zeilenrang folgt, indem man zu den transponierten Matrizen übergeht und die schon bewiesenen Aussagen über den Spaltenrang anwendet. Genauer ist der Zeilenrang von A gleich dem Spaltenrang von A^t gleich dem Spaltenrang von $N^tA^tM^t = (MAN)^t$, gleich dem Zeilenrang von MAN. □

Lemma 2.47 *Sei $\tilde{A}$ in Zeilenstufenform, und sei r die Anzahl der Nicht-Nullzeilen. Dann gilt: Der Zeilenrang von $\tilde{A}$ ist gleich dem Spaltenrang von $\tilde{A}$ ist gleich r.*

Beweis. Die Nicht-Nullzeilen sind offensichtlich linear unabhängig, d.h. sie bilden eine Basis für die von ihnen aufgespannten Raum.

Andererseits sind die Stufenspalten auch offensichtlich linear unabhängig, und die anderen Spalten sind von diesen Spalten linear abhängig. Damit bilden die Stufenspalten auch eine Basis in dem von ihnen aufgespannten Raum.

Beachten Sie, dass die Anzahl der Stufenspalten gleich der Anzahl der Nicht-Nullzeilen ist. □

Aus diesen beiden Lemmata folgt der Satz. Denn: Sei $M \in K^{m\times n}$ invertierbar so dass MA in Zeilenstufenform ist. Dann ist der Spaltenrang von A gleich dem Spaltenrang von MA gleich dem Zeilenrang von MA gleich dem Zeilenrang von A. □

Definition Da der Zeilenrang immer gleich dem Spaltenrang ist, spricht man einfach vom *Rang* von A. Bezeichnung: $\mathrm{Rang}(A)$.

Faktorräume und Dimensionsformeln

Sei V ein K-Vektorraum, und sei $U \subseteq V$ ein Untervektorraum. Dann ist U insbesondere eine Untergruppe von V, und wir haben die Faktorgruppe V/U mit der induzierten Addition (also $[\mathfrak{x}]_U + [\mathfrak{y}]_U] = [\mathfrak{x} + \mathfrak{y}]_U$ für alle $\mathfrak{x}, \mathfrak{y} \in V$).

Ich behaupte, dass die Skalarmultiplikatition eine Skalarmultiplikation

$$\cdot : K \times V/U \longrightarrow V/U \ (a, [\mathfrak{x}]_U) \longrightarrow [a\mathfrak{x}]_U$$

induziert, und dass V/U mit dieser Skalarmultiplikation ein K-Vektorraum ist.

Wir müssen überprüfen, dass die Multiplikation auf V/U wohldefiniert ist.

Seien dazu $a \in K$ und $\mathfrak{x}, \mathfrak{y} \in V$ mit $[\mathfrak{x}]_U = [\mathfrak{y}]_U$, d.h. $\mathfrak{x} \sim_U \mathfrak{y}$. Dann ist $a\mathfrak{x} - a\mathfrak{y} = a(\mathfrak{x} - \mathfrak{y}) \in U$, also $a\mathfrak{x} \sim_U a\mathfrak{y}$, was zu zeigen war.

Es ist leicht, die Vektorraumaxiome zu überprüfen.

Aussage 2.48 *Sei V endlich erzeugt und $U \subseteq V$ ein Untervektorraum. Dann gilt* $\mathrm{Dim}(V) = \mathrm{Dim}(V/U) + \mathrm{Dim}(U)$.

Beweis. Sei $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ eine Basis von U. Wir ergänzen diese Basis zu einer Basis $\mathfrak{x}_1, \ldots, \mathfrak{x}_s$ von V (siehe Aussage 2.44). Ich behaupte nun, dass $[\mathfrak{x}_{r+1}]_U, \ldots, [\mathfrak{x}_s]_U$ eine Basis von V/U ist. (Dann ist $s = \mathrm{Dim}(V)$, $r = \mathrm{Dim}(U)$ und $s - r = \mathrm{Dim}(V/U)$, und hieraus folgt die Behauptung.)

Offensichtlich bilden $[\mathfrak{x}_1]_U, \ldots, [\mathfrak{x}_s]_U$ ein Erzeugendensystem von V/U. Da aber $[\mathfrak{x}_1]_U = \cdots = [\mathfrak{x}_r]_U = \mathfrak{o}_U$ ist, bilden auch $[\mathfrak{x}_{r+1}]_U, \ldots, [\mathfrak{x}_s]_U$ ein Erzeugendensystem.

Wir zeigen nun die lineare Unabhängigkeit. Seien dazu $a_{r+1}, \ldots, a_s \in K$ mit $a_{r+1}[\mathfrak{x}_{r+1}]_U + \cdots + a_s[\mathfrak{x}_s]_U = \mathfrak{o}$. Dann ist also $a_{r+1}\mathfrak{x}_{r+1} + \cdots + a_s\mathfrak{x}_s \in U$, d.h. es gibt $a_1, \ldots, a_r \in K$ mit $a_{r+1}\mathfrak{x}_{r+1} + \cdots + a_s\mathfrak{x}_s = a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r$, d.h. $-a_1\mathfrak{x}_1 - \cdots - a_r\mathfrak{x}_r + a_{r+1}\mathfrak{x}_{r+1} + \cdots + a_s\mathfrak{x}_s = \mathfrak{o}$. Da $\mathfrak{x}_1, \ldots, \mathfrak{x}_s$ eine Basis von V bilden, folgt $a_1 = \cdots = a_s = 0$. □

Sei nun auch W ein K-Vektorraum und $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann induziert φ einen Isomorphismus von Vektorräumen

$$\overline{\varphi} : V/\operatorname{Kern}(\varphi) \tilde{\longrightarrow} \operatorname{Bild}(\varphi) \ , [\mathfrak{x}]_{\operatorname{Kern}(\varphi)} \mapsto \varphi(\mathfrak{x}) \ . \tag{2.14}$$

Um zu überprüfen, dass die Abbildung wohldefiniert und injektiv ist, nehmen wir uns zwei beliebige Vektoren $\mathfrak{x}, \mathfrak{y} \in V$ vor. Dann gilt:

$$\mathfrak{x} \sim_{\operatorname{Kern}(\varphi)} \mathfrak{y} \longleftrightarrow \mathfrak{x} - \mathfrak{y} \in \operatorname{Kern}(\varphi) \longleftrightarrow \varphi(\mathfrak{x} - \mathfrak{y}) = \mathfrak{o} \longleftrightarrow \varphi(\mathfrak{x}) = \varphi(\mathfrak{y}) \ ,$$

was zu zeigen war.

Es ist leicht zu sehen, dass φ eine lineare Abbildung ist, also ist es ein Isomorphismus.

Satz 2.6 *Sei V endlich erzeugt, und sei $\varphi : V \longrightarrow W$ eine lineare Abbildung. Dann gilt* $\mathrm{Dim}(V) = \mathrm{Dim}(\mathrm{Kern}(\varphi)) + \mathrm{Dim}(\mathrm{Bild}(\varphi))$.

Dies folgt aus Aussage 2.48 und dem Isomorphismus (2.14).

Korollar 2.49 *Sei $A \in K^{m\times n}$. Dann gilt* $\mathrm{Rang}(A) + \mathrm{Dim}(\mathrm{Kern}(A)) = n$.

Definition Seien $U_1, U_2 \subseteq V$ Untervektorräume. Dann definieren wir die *Summe* von U_1 und U_2 als

$$U_1 + U_2 := \{\mathfrak{x}_1 + \mathfrak{x}_2 \mid \mathfrak{x}_1 \in U_1\,, \mathfrak{x}_2 \in U_2\}\ .$$

Dies ist der kleinste Untervektorraum von V, der U enthält (überprüfen!).

Bemerkung Wenn $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ ein Erzeugendensystem von U_1 bilden und $\mathfrak{y}_1, \ldots, \mathfrak{y}_s$ ein Erzeugendensystem von U_2 bilden, dann bilden $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$, $\mathfrak{y}_1, \ldots, \mathfrak{y}_s$ ein Erzeugendensystem von $U_1 + U_2$.

Satz 2.7 *Sei V endlich erzeugt, und seien $U_1, U_2 \subseteq V$ Untervektorräume. Dann gilt* $\mathrm{Dim}(U_1 + U_2) + \mathrm{Dim}(U_1 \cap U_2) = \mathrm{Dim}(U_1) + \mathrm{Dim}(U_2)$.

Beweis. Sei $\mathfrak{x}_1, \ldots, \mathfrak{x}_r$ eine Basis von $U_1 \cap U_2$. Wir ergänzen diese Basis zu einer Basis $\mathfrak{x}_1, \ldots, \mathfrak{x}_s$ von U_1 sowie zu einer Basis $\mathfrak{x}_1, \ldots, \mathfrak{x}_r, \mathfrak{y}_1, \ldots, \mathfrak{y}_t$ von U_2.

Ich behaupte, dass dann $\mathfrak{x}_1, \ldots, \mathfrak{x}_s, \mathfrak{y}_1, \ldots, \mathfrak{y}_t$ eine Basis von $U_1 + U_2$ ist. (Dann ist $s+t = \mathrm{Dim}(U_1+U_2)$, $s = \mathrm{Dim}(U_1)$ und $t = \mathrm{Dim}(U_2) - \mathrm{Dim}(U_1 \cap U_2)$, und hieraus folgt die Behauptung.)

Es ist offensichtlich, dass es sich um ein Erzeugendensystem handelt.

Seien also $a_1, \ldots, a_s, b_1, \ldots, b_t \in K$ mit $a_1\mathfrak{x}_1 + \cdots + a_s\mathfrak{x}_s + b_1\mathfrak{y}_1 + \cdots + b_t\mathfrak{y}_t = \mathfrak{o}$. Dann ist also $a_1\mathfrak{x}_1 + \cdots + a_s\mathfrak{x}_s = -(b_1\mathfrak{y}_1 + \cdots + b_t\mathfrak{y}_t) \in U_1 \cap U_2$. Somit gibt es $c_1, \ldots, c_r \in K$ mit $c_1\mathfrak{x}_1 + \cdots + c_r\mathfrak{x}_r = a_1\mathfrak{x}_1 + \cdots + a_s\mathfrak{x}_s$ bzw. $(a_1 - c_1)\mathfrak{x}_1 + \cdots + (a_r - c_r)\mathfrak{x}_r + a_{r+1}\mathfrak{x}_{r+1} + \cdots + a_s\mathfrak{x}_s = \mathfrak{o}$. Hieraus folgt aufgrund der linearen Unabhängigkeit von $\mathfrak{x}_1, \ldots, \mathfrak{x}_s$: $c_1 = a_1, \ldots, c_r = a_r$ und $a_{r+1} = \cdots = a_s = 0$. Wir haben also $a_1\mathfrak{x}_1 + \cdots + a_r\mathfrak{x}_r + b_1\mathfrak{y}_1 + \cdots + b_t\mathfrak{y}_t = \mathfrak{o}$. Hieraus folgt nun $a_1 = \cdots = a_r = b_1 = \cdots = b_t = 0$ aufgrund der linearen Unabhängigkeit von $\mathfrak{x}_1, \ldots, \mathfrak{x}_r, \mathfrak{y}_1, \ldots, \mathfrak{y}_t$. □

Abbildungsmatrizen und Basiswechsel

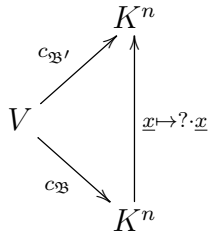
Sei V ein endlich erzeugter K-Vektorraum, und sei $\mathfrak{B} := (\mathfrak{b}_1, \ldots, \mathfrak{b}_n)$ eine Basis von V.

Wir haben den Isomorphismus $c_{\mathfrak{B}} : V \longrightarrow K^n$, der eindeutig durch $\mathfrak{b}_i \mapsto \underline{e}_i$ gegeben ist. Dieser Isomorphismus heißt die *Koordinatenabbildung* zu $\mathfrak{B}$. Für $\mathfrak{x} \in V$, heißt der Vektor $\underline{x} := c_{\mathfrak{B}}(\mathfrak{x})$ der *Koordinatenvektor* von $\mathfrak{x}$ bezüglich $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$.

Wir wählen nun eine zweite Basis $\mathfrak{B}' = (\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n)$ von V. Dann können wir $\mathfrak{b}'_j = \sum_{i=1}^{n} s_{i,j}\mathfrak{b}_i$ mit eindeutig bestimmten $s_{i,j}$ schreiben. Die Matrix $S = ((s_{i,j}))_{i,j}$ nennen wir *Übergangsmatrix* von $\mathfrak{B}$ nach $\mathfrak{B}'$.

Beachten Sie, dass der Vektor $\begin{pmatrix} s_{1,j} \\ \vdots \\ s_{n,j} \end{pmatrix}$ (die j-te Spalte von S) genau der Koordinatenvektor von $\mathfrak{b}'_j$ bez. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ ist.

Wir wollen nun die Koordinatenvektoren von Vektoren von V bez. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ in Koordinatenvektoren bez. $\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n$ umrechnen. D.h. gegeben $\underline{x} \in K^n$ wollen wir $c_{\mathfrak{B}'} \circ c_{\mathfrak{B}}^{-1}(\underline{x})$ berechnen. Beachten Sie, dass $c_{\mathfrak{B}'} \circ c_{\mathfrak{B}}^{-1} : K^n \longrightarrow K^n$ eine lineare Abbildung ist.



Wie muss die Matrix ? lauten, damit das Diagramm kommutativ ist?

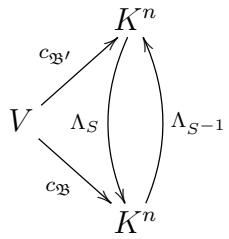
Bevor wir hierzu kommen, betrachten wir ein einfacheres Problem: Wir rechnen die Koordinatenvektoren von $\mathfrak{x} \in V$ bez. $\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n$ in Koordinatenvektoren bez. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ um.

Dies bedeutet: Wir bestimmen die Matrix zur linearen Abbildung $c_{\mathfrak{B}} \circ c_{\mathfrak{B}'}^{-1} : K^n \longrightarrow K^n$. Wir wissen, dass $c_{\mathfrak{B}} \circ c_{\mathfrak{B}'}^{-1}(\underline{e}_j) = c_{\mathfrak{B}}(\mathfrak{b}'_j) = \begin{pmatrix} s_{1,j} \\ \vdots \\ s_{m,j} \end{pmatrix}$. Und das bedeutet: Die gesuchte Matrix ist S.

Mit anderen Worten: Wenn $\underline{x}$ der Koordinatenvektor von $\mathfrak{x}$ bez. $\mathfrak{b}'_1, \ldots, \mathfrak{b}'_n$ ist, dann ist $S\underline{x}$ der Koordinatenvektor von $\mathfrak{x}$ bez. $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$.

Nun können wir auch die ursprüngliche Frage beantworten: Die Abbildung $c_{\mathfrak{B}'} \circ c_{\mathfrak{B}}^{-1} : K^n \longrightarrow K^n$ ist durch $\underline{x} \mapsto S^{-1}\underline{x}$ gegeben. Also: Wenn $\underline{x}$ der

Koordinatenvektor von $\mathfrak{x}$ bez. $\mathfrak{B}$ ist, dann ist $S^{-1}\underline{x}$ der Koordinatenvektor von $\mathfrak{x}$ bez. $\mathfrak{B}'$.



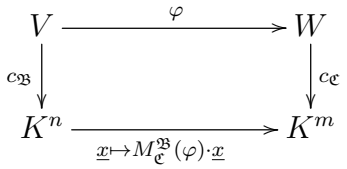
Sei nun W ein zweiter endlich erzeugter K-Vektorraum, sei $\mathfrak{C} := (\mathfrak{c}_1, \ldots, \mathfrak{c}_m)$ eine Basis von W, und sei $\varphi : V \longrightarrow W$ eine lineare Abbildung.

Wir wissen bereits, dass φ durch $\varphi(\mathfrak{b}_1), \ldots \varphi(\mathfrak{b}_n)$ eindeutig bestimmt ist. Wir definieren die *Abbildungsmatrix* von φ bezüglich der Basen $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ und $\mathfrak{c}_1, \ldots, \mathfrak{c}_m$ wie folgt:

Die j-te Spalte der Abbildungsmatrix ist der Koordinatenvektor von $\varphi(\mathfrak{b}_j)$ bezüglich der Basis $\mathfrak{c}_1, \ldots, \mathfrak{c}_m$.

Somit ist die Abbildungsmatrix also eine $m \times n$-Matrix. Die Abbildungsmatrix von φ bezüglich der Basen $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ und $\mathfrak{c}_1, \ldots, \mathfrak{c}_m$ bezeichnen wir mit $M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)$.

Wir haben das kommutative Diagramm:



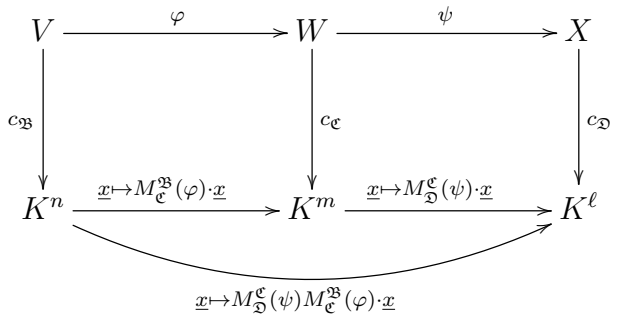
Die Abbildungsmatrix $M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)$ ist endeutig durch die Forderung, dass das Diagramm komutativ sei, bestimmt.

Beachten Sie: Mit den obigen Notationen ist die Abbildungsmatrix von id_V bezüglich $\mathfrak{B}'$ und $\mathfrak{B}$ gleich S, der Übergangsmatrix von $\mathfrak{B}$ auf $\mathfrak{B}'$ (und die Abbildungsmatrix von id_V bezüglich $\mathfrak{B}$ und $\mathfrak{B}'$ ist gleich S^{-1}).

$$M^{\mathfrak{B}'}_{\mathfrak{B}}(\mathrm{id}_V) = S \qquad M^{\mathfrak{B}}_{\mathfrak{B}'}(\mathrm{id}_V) = S^{-1} \tag{2.15}$$

Sei nun X noch ein endlich erzeugter K-Vektorraum mit Basis $\mathfrak{D} := (\mathfrak{d}_1, \ldots, \mathfrak{d}_\ell)$, und sei $\psi : W \longrightarrow X$ eine lineare Abbildung. Dann haben sind das rechte und das linke "Kästchen" sowie der untere Teil des folgenden

Diagramms kommutativ:



Dies bedeutet, dass das gesamte Diagramm kommutativ ist. Insbesondere sieht man, dass die Abbildungsmatrix von $\psi \circ \varphi : V \longrightarrow X$ bez. $\mathfrak{B}$ und $\mathfrak{C}$ gleich $M^{\mathfrak{C}}_{\mathfrak{D}}(\psi)M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)$ ist. Also

$$M^{\mathfrak{B}}_{\mathfrak{D}}(\psi \circ \varphi) = M^{\mathfrak{C}}_{\mathfrak{D}}(\psi)M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi) \; . \tag{2.16}$$

Eselsbrücke: Man kann "$\mathfrak{C}$ kürzen".

Wir betrachten nun, wie sich Abbildungsmatrizen unter Basiswechsel verändern.

Seien dazu $\mathfrak{B}, \mathfrak{B}'$ Basen von V und $\mathfrak{C}, \mathfrak{C}'$ Basen von W. Sei S die Übergangsmatrix von $\mathfrak{B}$ auf $\mathfrak{B}'$ und T die Übergangsmatrix von $\mathfrak{C}$ auf $\mathfrak{C}'$. Dann ist nach (2.16) und (2.15)

$$M^{\mathfrak{B}'}_{\mathfrak{C}'}(\varphi) = M^{\mathfrak{C}}_{\mathfrak{C}'}(\mathrm{id}_W)M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)M^{\mathfrak{B}'}_{\mathfrak{B}}(\mathrm{id}_V) = T^{-1}M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)S \; . \tag{2.17}$$

Wir betrachten noch zwei Spezialfalle:

Sei $\varphi : V \longrightarrow V$ ein Endomorphismus. Die Abbildungsmatrix $M^{\mathfrak{B}}_{\mathfrak{B}}(\varphi)$ heißt dann die *Abbildungsmatrix von φ bezüglich* der Basis $\mathfrak{B}$. Wenn nun $\mathfrak{B}'$ eine weitere Basis ist und S (wie oben) die Übergangsmatrix von $\mathfrak{B}$ zu $\mathfrak{B}'$ ist, dann ist

$$M^{\mathfrak{B}'}_{\mathfrak{B}'}(\varphi) = S^{-1}M^{\mathfrak{B}}_{\mathfrak{C}}(\varphi)S \; . \tag{2.18}$$

Wir geben uns eine Matrix $A \in K^{m\times n}$ vor und betrachten die Abbildung $\Lambda_A : K^n \longrightarrow K^m$. Die Abbildungsmatrix dieser Abbildung bez. den Standardbasen von K^n und K^m ist natürlich A.

Seien nun $\mathfrak{B} := (\underline{b}_1, \ldots, \underline{b}_n)$ und $\mathfrak{C} := (\underline{c}_1, \ldots, \underline{c}_m)$ Basen von K^n bzw. K^m, und seien B und C die Matrizen, die man erhält, wenn man die Basisvektoren jeweils als Spalten einer Matrix auffasst. Die Übergangsmatrix von der Standardbasis von K^n zur Basis $\mathfrak{B}$ ist demnach B, und die Übergangsmatrix von der Standardbasis von K^m zur Basis $\mathfrak{C}$ ist C. Demnach ist die Abbildungsmatrix von Λ_A bez. $\mathfrak{B}$ und $\mathfrak{C}$ gleich $C^{-1}AB$. (Überlegen Sie sich anhand der Definition der Abbildungsmatrix, warum dies richtig ist!)

Charakteristik und endliche Körper

Sei K ein Körper. Dann haben wir den Ringhomomorphismus $\varphi : \mathbb{Z} \longrightarrow K\,, z \mapsto z \cdot 1_K$. Wir wissen, dass es ein $n \in \mathbb{N}_0$ mit $\mathrm{Kern}(\varphi) = (n) = \{an, \mid a \in \mathbb{Z}\}$ gibt. Beachten Sie: Entweder ist $n = 0$ oder n ist die kleinste Zahl in $\mathbb{N}$ mit $n \cdot 1_K = 0$.

Sei nun $n > 0$. Dann ist n eine Primzahl. Denn: Angenommen n ist keine Primzahl. Dann gibt es $a, b \in \mathbb{N}$, beide $< n$, so dass $ab = n$. Damit ist dann $a \cdot 1_K \neq 0$ und $b \cdot 1_K \neq 0$ aber $ab1_K = (a1_K) \cdot (b1_K) = 0_K$, ein Widerspruch.

Definition Die Zahl n heißt die *Charakteristik* von K und wird mit $\mathrm{char}(K)$ bezeichnet.

Wenn $\mathrm{char}(K) = 0$, haben wir eine Inklusion $\mathbb{Z} \longrightarrow K\ , z \mapsto z \cdot 1_K$. Man erhält dann auch eine Inklusion $\mathbb{Q} \hookrightarrow K$ von Körpern (warum?). Man "identifiziert" nun $\mathbb{Q}$ mit seinem Bild in K (in Verallgemeinerung des Falls $K = \mathbb{R}$). Dann ist $\mathbb{Q}$ (bez. der Inklusion) der kleinste Unterkörper von K.

Sei nun $\mathrm{char}(K) = p > 0$. Wir haben einen injektiven Homomorphismus $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} \hookrightarrow K$ von Körpern. Das Bild von $\mathbb{F}_p$ ist ein Unterkörper von K; dieser Körper ist nun (bez. der Inklusion) der kleinste Unterkörper von K. Nach Beispiel 2.26 ist K insbesondere ein $\mathbb{F}_p$-Vektorraum.

Sei nun K ein endlicher Körper. Dann gibt es ein $m \in \mathbb{N}$ mit $m \cdot 1_K = 0$, also gilt $\mathrm{char}(K) > 0$. Sei wiederum $p := \mathrm{char}(K)$. Da K nur endliche viele Elemente enthält, ist K insbesondere endlich erzeugt als $\mathbb{F}_p$-Vekorraum. Sei $a_1, \ldots, a_d \in K$ eine Basis von K über $\mathbb{F}_p$. Dann ist K als $\mathbb{F}_p$-Vektorraum isomorph zu K^d. Insbesondere gilt $\#K = p^d$.

Wir haben gezeigt:

Aussage 2.50 *Die Anzahl der Elemente eines endlichen Körpers ist immer eine Primpotenz (d.h. eine Potenz einer Primzahl).*

Wie schon in Abschnitt 1.10, S. 64 erwähnt kann man auch zeigen: Zu jeder Primpotenz q gibt es einen endlichen Körper mit q Elementen, und zwei endliche Körper mit q Elementen sind isomorph.

2.7 Determinanten

Zur Motivation der Definition der Determinante nehmen wir uns die folgende Aufgabe vor:

Gegeben $\underline{x}_1, \ldots, \underline{x}_n \in \mathbb{R}^n$ wollen wir definieren, was das *Volumen* des Spats

$$\{c_1\underline{x}_1 + \cdots + c_n\underline{x}_n \mid 0 \leq c_j \leq 1 \text{ für alle } j = 1, \ldots, n\}$$

ist. Mit anderen Worten: Wir suchen eine Abbildung Vol : $(\mathbb{R}^n)^n \longrightarrow \mathbb{R}$, so dass für alle $\underline{x}_1, \ldots, \underline{x}_n \in \mathbb{R}^n$ $\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_n) \in \mathbb{R}$ unserer intuitiven Vorstellung des Volumen des entsprechenden physikalischen Spats entsricht.

Wir stellen die folgenden naheliegenden Forderungen (wobei $\underline{x}_1, \ldots, \underline{x}_n$ beliebige Vektoren aus $\mathbb{R}^n$ sind, $c \in \mathbb{R}$ und $i, j = 1, \ldots, n$ mit $i \neq j$ ist).

Vol1 $\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_{j-1}, c\underline{x}_j, \underline{x}_{j+1}, \ldots, \underline{x}_n) = |c| \cdot \mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_j, \underline{x}_{j+1}, \ldots, \underline{x}_n)$

Vol2 $\mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_{j-1}, \underline{x}_j + \underline{x}_i, \underline{x}_{j+1}, \ldots, \underline{x}_n) = \mathrm{Vol}(\underline{x}_1, \ldots, \underline{x}_n)$

Vol3 $\mathrm{Vol}(\underline{e}_1, \ldots, \underline{e}_n) = 1$

Man kann zeigen, dass es genau eine solche Abbildung Vol : $\mathbb{R}^n \longrightarrow \mathbb{R}$ gibt.

Eine andere Motivation der Determinante kann man über beliebigen Körpern formulieren: Wir suchen eine Abbildung, die jeder Matriz ein Skalar zuordnet, so dass die Matriz genau dann invertierbar ist, wenn dieses Skalar $\neq 0$ ist. Außerdem soll die Abbildung noch einige weitere "angenehme Eigenschaften" bez. elementaren Spaltentransformationen haben.

Sei also K ein Körper und $n \in \mathbb{N}$. Wir suchen eine Abbildung Det : $K^{n\times n} \longrightarrow K$, die die folgenden Eigenschaften hat:

Det1 Sei $A \in K^{n\times n}$, und sei A' eine Matrix, die aus A durch Multiplikation einer Spalte mit einem Skalar c hervorgeht. Dann gilt $d(A') = c\,\mathrm{Det}(A)$.

Det2 Sei $A \in K^{n\times n}$, und sei A' die Matrix, die aus A durch Addition des c-fachen von Spalte i zu Spalte j (mit $i \neq j$) hervorgeht. Dann gilt $\mathrm{Det}(A) = \mathrm{Det}(A')$.

Det3 $\mathrm{Det}(I_n) = 1$.

(Beachten Sie, dass aber für $c = 0$ die Transformation in Det1 keine elementare Spaltentransformation ist.)

So eine Abbildung $K^{n\times n} \longrightarrow K$ heißt eine *Determinanteabbildung*. Wir werden sogleich sehen, dass es genau eine Determinantenabbildung $K^{n\times n} \longrightarrow K$ gibt. Wir werden auch sehen, dass diese Determinantenabbildung die Eigenschaft hat, dass $\mathrm{Det}(A) = 0$ genau dann wenn $\mathrm{Rang}(A) < n$, was die ursprüngliche Motivation war.

Notation Wenn eine Abbildung $f : K^{m\times n} \longrightarrow K$ gegeben ist, erhält man mittels $(\underline{x}_1, \ldots, \underline{x}_n) \mapsto f(\underline{x}_1|\cdots|\underline{x}_n)$ eine Abbildung $(K^m)^n \longrightarrow K$. Wir bezeichnen diese Abbildung wieder mit f, d.h. wir setzen $f(\underline{x}_1, \ldots, \underline{x}_n) := f(\underline{x}_1|\ldots|\underline{x}_n)$. Umgekehrt identifizieren wir Abbildungen $(K^m)^n \longrightarrow K$ mit Abbildungen $K^{m\times n} \longrightarrow K$.

Bemerkung Wenn wir eine Determinantenabbildung $\mathrm{Det} : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$ haben, dann erfüllt die Abbildung $\mathrm{Vol} : (\mathbb{R}^n)^n \longrightarrow \mathbb{R}$, $(\underline{x}_1, \ldots, \underline{x}_n) \mapsto |\det(\underline{x}_1, \ldots, \underline{x}_n)|$ die obigen Forderungen Vol1, Vol2, Vol3. Die Zahl $\mathrm{Det}(\underline{x}_1, \ldots, \underline{x}_n)$ selbst kann man als *gerichtetes Volumen* auffassen.

Eindeutigkeit und Existenz einer Determinantenabbildung

Wir fixieren nun eine Determinantenabbildung $d : K^{n\times n} \longrightarrow K$ und leiten einige Eigenschaften her. Mittels dieser Eigenschaften werden wir dann insbesondere zeigen, dass es höchstens eine Determintenabbildung $K^{n\times n} \longrightarrow K$ gibt. Danach werden wir eine dieser Eigenschaften als Ansatz für eine Definition benutzen und zeigen, dass die so definierte Abbildung wirklich eine Determinatenabbildung ist.

Es ist leicht zu beschreiben, wie sich $d(A)$ (mit $A \in K^{n\times n}$) unter elementaren Spaltentransformationen angewandt auf A ändert. Wir kennen schon das Transformationsverhalten unter Multiplikation einer Spalte mit einem Skalar $c \neq 0$. Außerdem haben wir:

Lemma 2.51 *Sei $A \in K^{n\times n}$ und sei A' eine Matrix, die aus A durch Anwendung einer elementaren Spaltentransformation hervorgeht. Dann gilt:*

a) Wenn A' aus A durch Vertauschen von zwei Spalten hervorgeht, gilt $d(A') = -d(A)$.

b) Wenn A' aus A durch Addition des c-fachen einer Spalte zu einer anderen Spalte hervorgeht, gilt $d(A') = d(A)$.

Beweis. Wir zeigen zuerst die zweite Aussage. Sei A' die Matrix, die man aus A durch Addition des c-fachen von Spalte i zu Spalte j erhält. Für $c = 0$ ist nichts zu zeigen, sei also $c \neq 0$. Wir haben

$$\begin{aligned} c\,d(A) &= d(\underline{a}_1, \ldots, \underline{a}_{i-1}, c\underline{a}_i, \underline{a}_{i+1}, \ldots, \underline{a}_{j-1}) \\ &\overset{\text{Det1}}{=} d(\underline{a}_1, \ldots, \underline{a}_{i-1}, c\underline{a}_i, \underline{a}_{i+1}, \ldots, \underline{a}_{j-1}, \underline{a}_j + c\underline{a}_i, \underline{a}_{j+1}, \underline{a}_n) \\ &= c\,d(A')\ , \end{aligned}$$

und hieraus folgt die Behauptung. (Wir haben hier implizit angenommen, dass $i < j$, aber das hat nur notationelle Gründe.)

Die erste Behauptung folgt, da das Vertauschen von zwei Spalten durch wiederholte Addition und Subtraktion einer Spalte, gefolgt mit Multiplikation einer Spalte mit -1, dargestellt werden kann (Mit anderen Worten: Man braucht Umformung (II) im Gauß-Algorithmus eigentlich gar nicht.)

In der Tat, die Vertauschung von Spalten i und j kann man so realisieren:
- Man addiert Spalte i zu Spalte j. (Dann steht in Spalte j $\underline{a}_i + \underline{a}_j$.)
- Man subtahiert Spalte j von Spalte i. (Dann steht in Spalte i $\underline{a}_i - (\underline{a}_i + \underline{a}_j) = -\underline{a}_j$.)
- Man addiert Spalte i zu Spalte j. (Dann steht in Spalte j $\underline{a}_i$.)
- Man multipliziert Spalte i mit -1. □

Bemerkung Zusammen mit der Eigenschaft $d(I_n) = 1$ folgt aus dem obigen Lemma insbesondere, dass $d(E)$ für eine Elementarmatrix E durch Axiome Det1, Det2, Det3 eindeutig festgelegt ist und dass immer $d(E) \neq 0$ gilt.

Wenn man beachtet, dass eine elementare Spaltentransformationen zur Multiplikation mit einer Elementarmatrix von rechts korrespondiert, erhält man aus dem obigen Lemma sofort:

Lemma 2.52 *Sei $A \in K^{n\times n}$ eine beliebige Matrix, und sei $E \in K^{n\times n}$ eine Elementarmatrix. Dann gilt $d(AE) = d(A) \cdot d(E)$.*

Per Induktion nach k folgt:

Lemma 2.53 *Sei $A \in K^{n\times n}$ beliebig, und seien $E_1, \ldots, E_k$ $n\times n$-Elementarmatrizen. Dann gilt $d(AE_1 \cdots E_k) = d(A) \cdot d(E_1) \cdots d(E_k)$.*

Aussage 2.54 *Es gibt höchstens eine Determinantenabbildung $d : K^n \longrightarrow K$, und diese erfüllt $d(A) = 0$ genau dann wenn* $\mathrm{Rang}(A) < n$.

Beweis. Sei nach wie vor $d : K^{n\times n} \longrightarrow K$ eine Determinantenabbildung, und sei $A \in K^{n\times n}$.

Sei zunächst $\mathrm{Rang}(A) = n$, also A invertierbar. Dann gibt es Elementarmatrizen $E_1, \ldots, E_k$ mit $A = E_1 \cdots E_k$ (siehe Satz 2.2). Dann gilt nach dem obigen Lemma $d(A) = d(E_1) \cdots d(E_k)$. Nun ist die Determinante einer Elementarmatrix eindeutig festgelegt (s.o.). Damit ist auch die Determinante von A eindeutig festgelegt.

Sei nun $\mathrm{Rang}(A) < n$. Dann gibt es also eine Matrix $\tilde{A}$, die eine Nullspalte enthält (z.B. eine Matrix in Spaltenstufenform) sowie Elementarmatrixen $E_1, \ldots, E_k$ so dass $A = \tilde{A}E_1 \cdots E_k$. Damit ist $d(A) = d(\tilde{A}) \cdot d(E_1) \cdots d(E_k) = 0$. □

Die Beweismethode hat einige recht einfache Konsequenzen:

Aussage 2.55 *Für $A \in K^{n\times n}$ gilt $d(A) = d(A^t)$.*

Beweis. Offensichtlich ist für eine Elementarmatrix E $d(E) = d(E^t)$. Wenn $\text{Rang}(A) = n$, gibt es Elementarmatrizen $E_1, \ldots, E_k$ mit $A = E_1 \cdots E_k$. Damit gilt $d(A^t) = d(E_k^t \cdots E_1^t) = d(E_k^t) \cdots d(E_1^t) = d(E_1) \cdots d(E_k) = d(A)$.

Wenn $\text{Rang}(A) < n$, ist auch $\text{Rang}(A^t) = \text{Rang}(A) < n$. Damit gilt $d(A^t) = 0 = d(A)$. □

Bemerkung Sei wiederum $A \in K^{n\times n}$. Aufgrund der obigen Aussage hat die Determinantenfunktion d die folgenden Eigenschaften bezüglich Transformation von A mittels elementarer Zeilenoperationen:

- Sei A' eine Matrix, die aus A durch Multiplikation einer Zeile mit einem Skalar c hervorgeht. Dann gilt $d(A') = c\, d(A)$.
- Sei A' eine Matrix, die aus A durch Vertauschen von zwei Zeilen hervorgeht. Dann gilt $d(A') = -d(A)$.
- Sei A' die Matrix, die aus A durch Addition des c-fachen einer Zeile zu einer anderen Zeile hervorgeht. Dann gilt $d(A') = d(A)$.

Aussage 2.56 *Für $A, B \in K^{n\times n}$ gilt $d(AB) = d(A)\, d(B)$.*

Beweis. Sei zunächst $\text{Rang}(A) = \text{Rang}(B) = n$. Dann gibt es Elementarmatrizen $E_1, \ldots, E_k$ und $E_{k+1}, \ldots, E_\ell$ mit $A = E_1 \cdots E_k$ und $B = E_{k+1} \cdots E_\ell$. Damit ist $d(AB) = d(E_1 \cdots E_\ell) = d(E_1) \cdots d(E_\ell) = d(A)\, d(B)$.

Sei nun der Rang (mindestes) einer der beiden Matrizen $< n$. Wie das folgende Lemma zeigt, ist dann auch $\text{Rang}(AB) < n$. Damit gilt $d(AB) = 0 = d(A)\, d(B)$. □

Lemma 2.57 *Sei $A \in K^{m\times n}$ und $B \in K^{n\times r}$. Dann ist* $\text{Rang}(AB) \leq \min\{\text{Rang}(A), \text{Rang}(B)\}$.

Beweis. Wir benutzen die Definition des Rangs als Spaltenrang. Nach Definition der Matrizenmultiplikation sind alle Spalten von AB im "Spaltenraum von A" (d.h. im von den Spalten aufgespannten Raum) enthalten. Damit ist der Spaltenraum von AB im Spaltenraum von A enthalten. Dies impliziert $\text{Rang}(AB) \leq \text{Rang}(A)$.

Außerdem ist der Spaltenraum von AB gleich $\text{Bild}(\Lambda_{AB}) = \Lambda_{AB}(K^r) = \Lambda_A(\Lambda_B(K^r))$. Nach Lemma 2.43 a) ist die Dimension dieses Raumes $\leq \text{Dim}(\Lambda_B(K^r)) = \text{Rang}(B)$. □

Eine Aussage von Aussage 2.56 ist wiederum:

Aussage 2.58 *Die Determinantenabbildung d ist ein Homomorphismus von Monoiden von* $(K^{n\times n},\cdot)$ *nach* $(K,\cdot)$*. Insbesondere ist* $(K^{n\times n})^* \longrightarrow K^*$ *ein Homomorphismus von Gruppen.*

Beweis. Die erste Aussage ist eine Zusammenfassung der obigen Aussage und des Axioms Det3: $d(I_n) = 1$. Die zweite Aussage folgt sofort aus der ersten.□

Aussage 2.59 *Es gilt für alle* $j = 1,\ldots,n$ *und alle* $\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{x}_{j+1},\ldots,\underline{x}_n \in K^n$*: Die Abbildung*

$$d : K^n \longrightarrow K\,,\underline{x} \mapsto d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{x},\underline{x}_{j+1},\ldots,\underline{x}_n)$$

ist linear.

Beweis. Wir müssen nur zeigen, dass stets $d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{x}+\underline{y},\underline{x}_{j+1},\ldots,\underline{x}_n) = d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{x},\underline{x}_{j+1},\ldots,\underline{x}_n) + d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{y},\underline{x}_{j+1},\ldots,\underline{x}_n)$ gilt.

Wir wissen, dass $n+1$ Vektoren in K^n immer linear abhängig sind. Wir haben also $c,d \in K$ und $c_1,\ldots,c_{j-1},c_{j+1},\ldots,c_n \in K$, nicht alle $= 0$, mit

$$c\underline{x} + d\underline{y} + \sum_{k\neq j} c_k\underline{x}_k = 0\ .$$

Wenn $c = d = 0$ ist, sind die Vektoren $\underline{x}_k$ $(k \neq j)$ linear abhängig. Damit haben alle drei Matrizen, die hier betrachtet werden, Rang $< n$. Somit gilt die Behauptung.

Sei also $c \neq 0$ oder $d \neq 0$. Wie können o.E. (ohne Einschränkung) annehmen, dass $d \neq 0$ und sogar $d = -1$. Dann ist also $\underline{y} = c\underline{x} + \sum_{k\neq j} c_k\underline{x}_k$, und wir haben nun

$$\begin{array}{l} d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{y},\underline{x}_{j+1},\ldots,\underline{x}_n) = \\ d(\underline{x}_1,\ldots,\underline{x}_{j-1},c\underline{x}+\sum_{k\neq j} c_k\underline{x}_k,\underline{x}_{j+1},\ldots,\underline{x}_n) = \\ d(\underline{x}_1,\ldots,\underline{x}_{j-1},c\underline{x},\underline{x}_{j+1},\ldots,\underline{x}_n) = \\ c\cdot d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{x},\underline{x}_{j+1},\ldots,\underline{x}_n)\ . \end{array}$$

Nach dem selben Argument ist

$$\begin{array}{l} d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{x}+\underline{y},\underline{x}_{j+1},\ldots,\underline{x}_n) = \\ d(\underline{x}_1,\ldots,\underline{x}_{j-1},(1+c)\underline{x}+\sum_{k\neq j} c_k\underline{x}_k,\underline{x}_{j+1},\ldots,\underline{x}_n) = \\ (1+c)\cdot d(\underline{x}_1,\ldots,\underline{x}_{j-1},\underline{x},\underline{x}_{j+1},\ldots,\underline{x}_n)\ . \end{array}$$

Aus der Verbindung dieser beiden Identitäten folgt die Behauptung. □

Bemerkung Die Aussage in der obige Aussage heißt *Multilinearität* von d.

Sei nun $n \geq 2$.

Notation Sei für $A \in K^{n\times n}$ und $i,j = 1,\dots n$ $A_{i,j}$ diejenige $n-1\times n-1$-Matrix, die entsteht, wenn man in A die i-te Zeile und die j-te Spalte streicht.

Betrachten wir die Abbildung $K^{(n-1)\times(n-1)} \longrightarrow K\,, A \mapsto d(\begin{pmatrix} 1 & \\ & A \end{pmatrix})$. Man sieht leicht, dass dies eine Determinantenabbildung auf $K^{(n-1)\times(n-1)}$ ist. Wir bezeichnen diese Abbildung mit d_{n-1}, und die wir setzen $d_n := d$.

Aussage 2.60 (Laplacescher Entwicklungssatz) *Sei $A \in K^{n\times n}$. Dann gilt für alle $j = 1,\dots,n$:*

$$d_n(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j}\, d_{n-1}(A_{i,j}) \qquad \textit{(Entwicklung nach Spalte } j\textit{)}$$

Analog gilt für alle $i = 1,\dots,n$:

$$d_n(A) = \sum_{j=1}^{n} (-1)^{i+j}\, a_{i,j}\, d_{n-1}(A_{i,j}) \qquad \textit{(Entwicklung nach Zeile } i\textit{)}$$

Beweis. Die beiden Aussagen sind äquivalent da $d_n(A) = d_n(A^t)$ und $d_{n-1}(A^t) = d_{n-1}(A^t)$. Wir zeigen die Entwicklung nach Spalte j.

Aufgrund der Multilinearität ist

$$\begin{aligned} d(A) = &\quad d(\underline{a}_1,\dots,\underline{a}_{j-1},\textstyle\sum_{i=1}^n a_{i,j}\,\underline{e}_i,\underline{a}_{j+1},\dots,\underline{a}_n) \\ = &\quad \textstyle\sum_{i=1}^n a_{i,j}\,(\underline{a}_1,\dots,\underline{a}_{j-1},\underline{e}_i,\underline{a}_{j+1},\dots,\underline{a}_n)\,. \end{aligned}$$

Wir müssen also zeigen, dass

$$d_n(\underline{a}_1,\dots,\underline{a}_{j-1},\underline{e}_i,\underline{a}_{j+1},\dots,\underline{a}_n) = (-1)^{i+j} d_n(A_{i,j})\,.$$

Sei zunächst

$$A = (\underline{e}_1|a_2|\cdots|\underline{a}_n) = \begin{pmatrix} 1 & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,n} \end{pmatrix}\,.$$

Dann ist $d_n(A) = d_n(\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,n} \end{pmatrix})$ aufgrund von Spaltentransformationen, und letzteres ist per Definition gleich $d_{n-1}(A_{1,1})$. Somit ist die Formel für solche Matrizen richtig.

Sei nun $A = (\underline{e}_i|a_2|\cdots|\underline{a}_n)$ eine Matrix wie oben, nur das die Eins in in der i-ten Zeile steht. Wieder "räumen wir zunächst die Einträge rechts der 1 aus (Addition von Vielfachen der ersten Spalte zu anderen Spalten). Wir führen hintereinander die Zeilentransformationen "Vertauschen der Zeile i mit Zeile $i-1$, "Vertauschen der Zeile $i-1$ mit Zeile $i-2$", ..., "Vertauschen der Zeile 2 mit Zeile 1" durch und nennen das Ergebnis A'. Dann ist $A'_{1,1} = A_{i,1}$ und $A' = \begin{pmatrix} 1 & 0 \\ 0 & A_{i,1} \end{pmatrix}$. Damit ist $d_n(A) = (-1)^{i-1}\, d_n(A') = (-1)^{i+1}\, d_{n-1}(A_{i,1})$. Die Formel ist also wiederum richtig.

Sei nun $A = (\underline{a}_1|\cdots|\underline{a}_{j-1}|\underline{e}_i|\underline{a}_{j+1}|\cdots|\underline{a}_n)$. Wir gehen analog vor und "räumen zuerst die Einträge rechts und links vom Eintrag mit Index (i,j) aus". Dann vertauschen wir der Reihe nach die Spalte j mit der Spalte $j-1$, ..., die Spalte 2 mit der Spalte 1. Dann vertauschen wir noch die Zeilen wie oben. Wir erhalten die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & A_{i,j} \end{pmatrix}$. Wir haben nun $d_n(A) = (-1)^{i-1+j-1}\, d_{n-1}(A_{i,j}) = (-1)^{i+j}\, d_{n-1}(A_{i,j})$, abermals die richtige Formel. □

Wir kommen nun zum Hauptresultat über Determinantenabbildungen.

Satz 2.8 *Sei K ein Körper und $n \in \mathbb{N}$. Dann gibt es genau eine Determinantenabbildung $K^{n\times n} \longrightarrow K$.*

Beweis. Die Eindeutigkeit haben wir schon gezeigt.

Motiviert durch die obige Entwicklung nach Zeilen (die ja für jede Determinantenabbildung gelten muss) *definieren* wir rekursiv für jeden Körper K:

$$\mathrm{Det}_1(a) := a \qquad \mathrm{Det}_n(A) := \sum_{j=1}^{n} (-1)^{1+j} a_{i,j}\, \mathrm{Det}_{n-1}(A_{i,j})$$

für $n \in \mathbb{N}$ und $A \in K^{n\times n}$. Ich behaupte, dass für alle $n \in \mathbb{N}$ Det_n eine Determinantenabbildung ist.

Wir zeigen dies nach Induktion über n. Der Induktionsanfang $n = 1$ ist trivial. Wir setzen also voraus, dass die Behauptung für n richtig ist und zeigen die Behauptung für $n+1$.

Beachten Sie, dass wir dann insbesondere alle oben bewiesenen Eigenschaften für Determinantenabbildungen für Det_n benutzen dürfen.

Offensichtlich ist $\mathrm{Det}_{n+1}(I_{n+1}) = 1 \cdot \mathrm{Det}(I_1) = 1$.

Sei $A \in K^{(n+1)\times(n+1)}$, $j = 1, \ldots, n+1$ und A' diejenige Matrix, die aus A durch Multiplikation der j-ten Spalte mit $c \in K$ hervorgeht. Dann ist $\mathrm{Det}_n(A'_{1,k}) = c\, \mathrm{Det}_n(A'_{1,k})$ für alle $k \neq j$. Damit gilt $\mathrm{Det}_{n+1}(A') = \sum_{j=1}^{n} (-1)^{1+j}\, c\, a_{1,j}\, \mathrm{Det}_n(A_{i,j}) = \mathrm{Det}_{n+1}(A)$.

Sei nun $A \in K^{(n+1)\times(n+1)}$, und seien $i, j = 1, \ldots, n+1$ mit $i \neq j$. Sei A' diejenige Matrix, die aus A durch Addition der i-ten Spalte zur j-ten Spalte hervorgeht. Für $k \neq i, j$ entsteht dann $A'_{1,i}$ auch aus $A_{1,i}$, indem eine Spalte zu einer anderen addiert wird. Wenn wir nun anwenden, dass Det_n eine Determinantenabbildung ist, erhalten wir, dass $\mathrm{Det}(A'_{1,k}) = \mathrm{Det}(A_{1,k})$ für alle $k \neq i, j$. Außerdem ist dies offensichtlich auch für $k = j$ richtig, denn diese Spalte wird ja gerade gestrichen.

Wir untersuchen nun $\mathrm{Det}_n(A'_{1,i})$. Es ist $\mathrm{Det}_n(A'_{1,i}) = \mathrm{Det}_n(A_{1,i}) + \mathrm{Det}_n(B)$, wobei B aus A hervorgeht, indem man zuerst die Spalten i und j vertauscht und dann die Spalte i streicht. Also hat B bis auf Reihenfolge die selben Spalten wie die Matrix $A_{1,j}$. Genauer geht B aus $A_{j,1}$ durch $|j-i|-1$ Spaltenvertauschungen hervor. Damit ist also $\mathrm{Det}_n(A'_{1,i}) = \mathrm{Det}_n(A_{1,i}) + (-1)^{|j-i|-1}\,\mathrm{Det}_n(A_{1,j}) = \mathrm{Det}_n(A_{1,i}) + (-1)^{j-i+1}\,\mathrm{Det}_n(A_{1,j})$.

Es folgt:

$$
\begin{aligned}
& \mathrm{Det}_{n+1}(A') - \mathrm{Det}_{n+1}(A) \\
= \;& (-1)^{1+i} a_{1,i}\,\mathrm{Det}_{n-1}(A'_{1,i}) + (-1)^{1+j} a'_{1,j}\,\mathrm{Det}_n(A_{1,j}) \\
& -(-1)^{1+i} a_{1,i}\,\mathrm{Det}_n(A_{1,i}) - (-1)^{1+j} a_{1,j}\,\mathrm{Det}_n(A_{1,j}) \\
= \;& (-1)^{1+i} a_{1,i}\,\mathrm{Det}_{n-1}(A_{1,i}) + (-1)^{1+i}(-1)^{j-i+1} a_{1,i}\,\mathrm{Det}_n(A_{1,j}) \\
& \quad +(-1)^{1+j} a_{1,j}\,\mathrm{Det}_n(A_{1,j}) + (-1)^{1+j} a_{1,i}\,\mathrm{Det}_n(A_{1,j}) \\
& -(-1)^{1+i} a_{1,i}\,\mathrm{Det}_n(A_{1,i}) - (-1)^{1+j} a_{1,j}\,\mathrm{Det}\,n(A_{1,j}) \\
= \;& 0
\end{aligned}
$$

□

Notation Im Folgenden schreiben wir Det statt Det_n. Eine andere übliche Schreibweise für $\mathrm{Det}\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & . & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$ ist $\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & . & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}$.

Man kann die Spalten- / Zeilnentwicklungen der Determinante verwenden, um eine nicht-rekursive Formel herzuleiten. Es ergibt sich:

$\underline{n=2}$. Es ist $\mathrm{Det}(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,2}$.

$\underline{n=3}$. Es ist $\mathrm{Det}(A) = a_{1,1}\begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} - a_{2,1}\begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} + a_{3,1}\begin{vmatrix} a_{1,2} & a_{1,2} \\ a_{2,2} & a_{2,3} \end{vmatrix} =$
$a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} - a_{3,1}a_{2,2}a_{1,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3}$.
Dies ist die so genannte *Regel von Sarrus*.

Ich gebe noch ohne Beweis das Ergebnis für beliebies n an.
Zunächst einige Definitionen und einfache Bemerkungen.

Definition / Bemerkung Wir ordnen jeder Permutation $\pi \in S_n$ die entsprechende *Permutationsmatrix* $M_\pi := (\underline{e}_{\pi(1)}, \ldots, \underline{e}_{\pi(n)}) \in K^{n\times n}$ zu. Man sieht leicht, dass diese Matrix immer invertierbar ist, und dass die Abbildung $S_n \longrightarrow K^{n\times n}$ ein Gruppenhomomorphismus ist. Wir betrachten nun Permutationsmatrizen über $\mathbb{Q}$ und definieren nun das *Signum* von $\pi \in S_n$ wie folgt: $\text{sign}(\pi) := \text{Det}(M_\pi) \in \mathbb{Q}^*$. Da auch $\text{Det} : (\mathbb{Q}^{n\times n})^* \longrightarrow \mathbb{Q}^*$ ein Gruppenhomomorphismus ist, ist also $\text{sign} : S_n \longrightarrow \mathbb{Q}^*$ ein Gruppenhomomorphismus.

Eine *Transposition* ist per Definition eine Permutation, die genau zwei Elemente vertauscht und die anderen Elemente fest lässt. Man sieht leicht, dass jede Permutation ein Produkt von Transpositionen ist (= aus Transpositionen durch Verknüpfung hervorgeht). (Beweisen Sie dies per Induktion!)

Wenn π eine Transposition ist, ist $\text{Det}(\pi) = -1$ per Definition. Damit gilt: Sei $\pi = \pi_1 \cdots \pi_k$, wobei die π_i Transpositionen sind. Dann ist $\text{sign}(\pi) = (-1)^k$. Insbesondere ist also $\text{sign}(\pi) = \pm 1$.

Man sieht auch: Wenn $\pi_1 \cdots \pi_k = \sigma_1 \cdots \sigma_\ell$ mit Transpositionen π_i und σ_i, dann sind entweder k und ℓ beide gerade oder beide sind ungerade.

Wir haben nun (ohne Beweis):

Aussage 2.61 (Formel von Leibniz) *Sei $A \in K^{n\times n}$. Dann ist*

$$\text{Det}(A) = \sum_{\pi\in S_n} \text{sign}(\pi) a_{1,\pi(1)} \cdots a_{n,\pi(n)} \ .$$

Bemerkung Diese Formel hat ihren Wert in theoretischen Betrachtungen. Für die algorithmische Berechnung der Determinante ist sie aber katastrophal: Mittels dieser Formel benötigt man $(n-1)\cdot n!$ Körpermultiplikationen. Wenn man mit elementaren Spalten- und Zeilenumformungen rechnet, benötigt man nur $\mathcal{O}(n^3)$ Körperoperationen, genau wie beim Gauß-Algorithmus.

Die adjunkte Matrix und die Cramersche Regel

Definition Für $A \in K^{n\times n}$ definieren wir die *adjunkte Matrix* als

$$A^\# := ((-1)^{i+j} (\text{Det}(A_{j,i})))_{i,j=1,\ldots n} \ .$$

(Beachten Sie die Rolle der Indices!)

Sei nun $A \in K^{n\times n}$ und $\underline{b} \in K^n$, und sei $\underline{c} := A^{\#}\underline{b}$. Dann ist

$$c_i = \sum_{j=1}^{n} (-1)^{i+j} \operatorname{Det}(A_{j,i})\, b_j = \operatorname{Det}(\underline{a}_1, \ldots, \underline{a}_{i-1}, \underline{b}, \underline{a}_{i+1}, \cdots, \underline{a}_n)\ . \qquad (2.19)$$

aufgrund der Formel für die Entwicklung nach Spalten. Also: Man erhält c_i, indem man *die i-te Spalte von A durch $\underline{b}$ ersetzt und die Determinante dieser Matrix bildet.* (Beachten Sie wieder die unkonventionelle Rolle des Index i!)

Somit ist insbesondere $A^{\#}\underline{a}_j = \operatorname{Det}(A)\,\underline{e}_j$. (Wenn man die i-te Spalte von A durch $\underline{a}_j$ ersetzt und dann die Determinante bildet, erhält man $\operatorname{Det}(A)\,\delta_{i,j}$.)

Damit gilt

$$A^{\#}A = \operatorname{Det}(A)\, I_n\ . \qquad (2.20)$$

Da A beliebig war, gilt auch $(A^t)^{\#}A^t = \operatorname{Det}(A^t)\, I_n = \operatorname{Det}(A)\, I_n$. Nun ist $(A^t)^{\#} = ((-1)^{i+j}\,(\operatorname{Det}(A_{i,j})))_{i,j=1,\ldots n} = (A^{\#})^t$.

Damit folgt:

$$(A\,A^{\#})^t = (A^{\#})^t A^t = (A^t)^{\#}A^t = \operatorname{Det}(A)\, I_n$$

Wenn man nochmal transponiert, erhält man

$$A\,A^{\#} = \operatorname{Det}(A)\, I_n\,.$$

Wenn A invertierbar ist, kann man dies natürlich auch durch

$$A^{-1} = \frac{1}{\operatorname{Det}(A)} \cdot A^{\#}$$

ausdrücken.

Wir geben uns nun eine invertierbare Matrix $A \in K^{n\times n}$ und einen Vektor $\underline{b} \in K^n$ vor. Dann wissen wir, dass das LGS $A \cdot \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \underline{b}$ genau eine Lösung hat; sei diese $\underline{x}$. Wenn wir beide Seiten von (2.20) von rechts mit $\underline{x}$ multiplizieren erhalten wir $A^{\#}\,\underline{b} = \operatorname{Det}(A)\,\underline{x}$, also

$$\underline{x} = \frac{1}{\operatorname{Det}(A)} \cdot A^{\#}\,\underline{b}\,.$$

Wenn wir (2.19) beachten, erhalten wir:

$$x_i = \frac{\operatorname{Det}(\underline{a}_i, \cdots, \underline{a}_{i-1}, \underline{b}, \underline{a}_{i+1}, \cdots, \underline{a}_n)}{\operatorname{Det}(A)} \qquad (2.21)$$

Dies ist die so genannte *Cramersche Regel.*

Algorithmisch ist diese Formel aber kein Fortschritt gegenüber dem Gauß-Algorithmus.

Die Determinante eines Endomorphismus

Wir geben uns jetzt einen endlichen erzeugten K-Vektorraum V und einen Endomorphismus $\varphi : V \longrightarrow V$ vor. Wir wollen die *Determinante von* φ definieren.

Hierzu wählen wir uns (irgendwie) eine Basis $\mathfrak{b}_1, \ldots, \mathfrak{b}_n$ von V und betrachten die Determinante der Abbildungsmatrix $M_{\mathfrak{B}}^{\mathfrak{B}}(\varphi)$ von φ bez. $\mathfrak{B}$.

Ich behaupte, dass diese Determinante nicht von der Wahl der Basis abhängt.

Sei hierzu $\mathfrak{B}'$ eine andere Basis von V, und sei S die Übergangsmatrix von $\mathfrak{B}$ nach $\mathfrak{B}'$. Dann gilt nach (2.18) und Aussage 2.58:

$$\begin{aligned}\mathrm{Det}(M_{\mathfrak{B}'}^{\mathfrak{B}'}(\varphi)) = \mathrm{Det}(S^{-1}\, M_{\mathfrak{B}}^{\mathfrak{B}}(\varphi)\, S) = \\ \mathrm{Det}(S^{-1})\ \mathrm{Det}(M_{\mathfrak{B}}^{\mathfrak{B}}(\varphi))\ \mathrm{Det}(S) = \mathrm{Det}(M_{\mathfrak{B}}^{\mathfrak{B}}(\varphi))\end{aligned}$$

Damit können wir definieren: Die *Determinante* von φ ist die Determinante der Abbildungsmatrix von φ bezüglich irgendeiner Basis von V. Bezeichnung: $\mathrm{Det}(\varphi)$.

Langrange-Interpolation und die Vandermondesche Matrix

Ich schließe mit der expliziten Berechnung der Determinante einer bestimmten Art von Matrizen und einer Anwendung.

Sei wie immer K ein Körper, und seien $(x_1, y_1), \ldots, (x_n, y_n) \in K^2$, wobei die x_i paarweise verschieden sind. Wir wollen diese Punkte mittels eines Polynoms *interpolieren*. Dies heißt, wir wollen ein Polynom $f(X) \in K[X]$ mit $p(x_i) = y_i$ für alle $i = 1, \ldots, n$ finden. Wir stellen uns die Aufgabe, ein solches Polynom mit möglichst kleinem Grad zu finden.

Nun hat diese Aufgabe die folgende eindeutige Lösung: *Es gibt ein eindeutig bestimmtes Interpolationspolynom* $f(X) \in K[X]$ *mit* $\mathrm{Grad}(f(X)) \leq n-1$.

Das Auffinden dieses Polynoms heißt *Lagrange Interpolation*.

Dies wollen nun die Behauptung beweisen. Dabei werden wir auch eine Methode kennen lernen, die es erlaubt, dieses Polynom explizit zu berechnen.

Sei zunächst $f(X) = \sum_{j=0}^{d} a_j X^j$ irgendein Polynom. Dann gilt: $p(x_i) = y_i \longleftrightarrow \sum_{j=0}^{d} a_j x_i^j = y_i \longleftrightarrow \begin{pmatrix} 1 & x_i & \cdots & x_i^d \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_d \end{pmatrix} = y_i$. Mit anderen Worten: Das Polynom $f(X)$ interpoliert die vorgegebenen Punkte genau

dann, wenn $\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix}$ eine Lösung des LGS

$$\begin{pmatrix} 1 & x_1 & \cdots & x_1^d \\ \vdots & \vdots & . & \vdots \\ 1 & x_n & \cdots & x_n^d \end{pmatrix} \cdot \begin{pmatrix} A_0 \\ \vdots \\ A_d \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

ist, wobei $\begin{pmatrix} A_0 \\ \vdots \\ A_d \end{pmatrix}$ der Vektor der Unbestimmten ist.

Ich behaupte nun, dass die quadratische Matrix $\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & . & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}$ invertierbar ist. Hieraus folgt die Behauptung, dass es ein eindeutig bestimmtes Interpolationspolynom vom Grad $\leq n-1$ gibt. Außerdem ergibt dies sofort eine Methode, dies zu berechnen: Man löse das LGS! (Es gibt aber noch andere Methoden hierfür.)

Die Matrix $\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & . & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix}$ heißt *Vandermondesche Matrix*. Ich berechne nun die Determinante dieser Matrix und zeige, dass diese $\neq 0$ ist. Beachten Sie, dass wir vorausgesetzt haben, dass die x_i paarweise verschieden sind. Wenn zwei gleich sind, ist die Determinante offensichtlich $= 0$.

Wenn wir der Reihe nach das x_n-fache der Spalte $n-1$ von der Spalte n abziehen, das x_n-fache der Spalte $n-2$ von der Spalte n abziehen, ..., das x_n-fache der Spalte 2 von der Spalte 1 abziehen (in dieser Reihenfolge),

erhalten wir:

$$\mathrm{Det}\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & . & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} = \mathrm{Det}\begin{pmatrix} 1 & x_1 - x_n & \cdots & x_1^{n-1} - x_n x_1^{n-2} \\ \vdots & \vdots & . & \vdots \\ 1 & x_{n-1} - x_n & \cdots & x_{n-1}^{n-1} - x_n x_{n-1}^{n-2} \\ 1 & 0 & \cdots & 0 \end{pmatrix} =$$

$$(-1)^{n+1} \cdot \mathrm{Det}\begin{pmatrix} x_1 - x_n & \cdots & x_1^{n-1} - x_n x_1^{n-2} \\ \vdots & . & \vdots \\ x_{n-1} - x_n & \cdots & x_{n-1}^{n-1} - x_n x_{n-1}^{n-2} \end{pmatrix} =$$

$$(x_2 - x_1) \cdots (x_n - x_1) \cdot \mathrm{Det}\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-2} \\ \vdots & \vdots & . & \vdots \\ 1 & x_{n-1} & \cdots & x_{n-1}^{n-2} \end{pmatrix} .$$

Die Matrix, von der die Determinante genommen wird, ist nun eine $(n-1) \times (n-1)$-Vandermondesche Matrix.

Per Induktion nach n sieht man nun, dass

$$\mathrm{Det}\begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \vdots & \vdots & . & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} = \prod_{i>j} (x_i - x_j) .$$

Insbesondere sieht man, dass die Determinante $\neq 0$ ist, also ist die Matrix invertierbar.