Skript zur Vorlesung „Mathematik für die Informatik 1“ im Wintersemester 2022

Wolfgang Merkle

Universität Heidelberg, Institut für Informatik

Version 20. Februar 2023

Inhaltsverzeichnis

1 Grundlagen

1.1 Einleitung

1.1.1 Informatik und Mathematik

Mathematik in der Informatik Wer Informatik betreiben will kommt nicht darum herum, sich auch mit Mathematik zu beschäftigen. Bereits die grundlegenden Herangehensweisen und Methoden der Informatik und Mathematik sind sich sehr ähnlich. Es wird von Irrelevantem abstrahiert um dadurch Sachverhalte exakt darstellen zu können. Entsprechend sind Aussagen wahr oder falsch und bieten keinen Raum für subjektive Interpretationen. Die Frage, ob eine bestimmte natürliche Zahl gerade ist, oder ob ein Programm bei einer bestimmte Eingabe terminiert hat eine eindeutige Antwort. Weiter wird beispielsweise in einem mathematischen Beweis durch eine logische Argumentationskette zweifelsfrei nachgewiesen, dass eine bestimmte Aussage wahr ist, falls bestimmte Voraussetzungen wahr sind. Auf ganz ähnliche Weise kann in der Informatik nachgewiesen werden, dass ein bestimmtes Programm in dem Sinne korrekt ist, dass es für alle Eingaben terminiert und das richtige Ergebnis liefert. Allgemein wird in der Informatik häufig in der Sprache der Mathematik formuliert. Dadurch können Sachverhalte exakt und unmissverständlich dargestellt werden, so dass sie für alle mathematisch Geschulten gut verständlich und klar sind.

Teile der Forschung der Mathematik einerseits und der Informatik andererseits sind eng miteinander verzahnt. So lassen sich viele Probleme der Informatik unter Verwendung von mathematischen Verfahren lösen, beispielsweise basieren einige der wichtigsten in der Praxis eingesetzten kryptographischen Verfahren auf Methoden aus der Zahlentheorie, einem Teilgebiet der Mathematik. Umgekehrt erfordert die praktische Anwendung von mathematischen Ergebnissen regelmäßig die Übersetzung dieser Ergebnisse in Algorithmen und Programme. Es ist ein Unterschied, für einen bestimmte Typ von Gleichungssystem ein mathematisches Verfahren zur Bestimmung einer reellwertigen Lösung zu kennen, oder eine solche Lösungen durch ein Programm bis auf eine gegebene Genauigkeit zu berechnen. Es gibt allerdings auch Teilgebiete der Mathematik in denen in sehr hohem Maße abstrakte und komplexe Strukturen untersucht werden, wobei die praktische Anwendbarkeit der Ergebnisse weder als relevant angesehen wird noch zu erwarten ist.

Neben einer Einführung in grundlegende Begriffe, Ergebnisse und Methoden der linearen Algebra ist diese Vorlesung auch ganz allgemein als Einführung in die Arbeits- und Herangehensweise der Mathematik gedacht. Das mathematische Handwerkszeug wie mathematische Notation, exaktes Formulieren und Schließen oder Beweistechniken soll vermittelt werden. Dies alles mit dem Ziel der Anwendung in der Informatik und ohne die in der Mathematik

übliche umfassende Behandlung abstrakter Strukturen, diese werde im Folgenden nur am Beispiel einiger zentraler Begriffe wie dem einer Gruppe oder eines Körperss behandelt.

1.1.2 Definition, Satz, Beweis

Die mathematische Methodologie, das ist die Vorgehensweise in der mathematische Forschung, lässt sich so beschreiben, dass zunächst Begriffe exakt definiert und dann wahre Aussagen, die als Sätze bezeichnet werden, über diese Begriffe formal bewiesen werden. In diesem Sinne besteht die Mathematik und die mathematische Literatur zu einem großen Teil aus Definitionen, Sätzen und Beweisen.

Wir wollen uns dies am Beispiel eines Satzes über Primzahlen verdeutlichen. Dabei setzen wir ein intuitives Verständnis der folgenden Begriffe voraus: die natürlichen Zahlen $0, 1, 2, \ldots$ werden durch die Relation $<$ wie üblich geordnet, und auch die Addition $+$ und Multiplikation $\cdot$ natürlicher Zahlen sind wie üblich definiert. Wir werden später als Beispiel für den mengentheoretischen Aufbau der Mathematik sehen, wie sich die natürlichen Zahlen, Relationen wie $<$ und Funktionen wie die Addition formal definieren lassen.

1 Definition. *Eine natürliche Zahl t ungleich 0* TEILT *eine natürliche Zahl n, falls es eine natürliche Zahl k mit $n = k \cdot t$ gibt, die Zahl 0 teilt keine natürliche Zahl. Dass eine Zahl t eine Zahl n teilt, lässt sich auch so ausdrücken, dass t* TEILER *von n ist oder dass n* VIELFACHES *von t ist.*

Eine natürliche Zahl ist GERADE*, wenn sie von 2 geteilt wird, sonst* UNGERADE.

Die TRIVIALEN TEILER *einer natürlichen Zahl n ungleich 0 sind 1 und n, alle anderen Teiler von n heißen* NICHTTRIVIAL. *Eine* PRIMZAHL *ist eine von 0 und 1 verschiedene natürliche Zahl, die nur triviale Teiler hat. Primzahlen werde auch als* PRIM *bezeichnet, alle andere natürlichen Zahlen ungleich 0 und 1 als* ZUSAMMENGESETZT.

Beachte, dass die trivialen Teiler einer Zahl tatsächlich Teiler dieser Zahl sind und somit jede natürliche Zahl mindestens einen Teiler hat. Nach Definition teilt die Zahl 0 keine natürliche Zahl, auch nicht sich selbst, jedoch teilt jede natürliche Zahl ungleich 0 die Zahl 0, insbesondere ist 0 gerade. Die fünf kleinsten Primzahlen sind 2, 3, 5, 7 und 11.

2 Bemerkung. *Der Vollständigkeit halber sei angemerkt, dass sich der Begriff Teiler und damit auch der Begriff Vielfaches mit fast wörtlich der gleichen Definition auf die ganzen Zahlen $\ldots, -2, -1, 0, 1, 2, \ldots$ erweitern lassen. Zum Beispiel teilt jede ganze Zahl ungleich 0 die Zahl 0, und die Zahlen 2, 3, -2 und -3 teilen sowohl 6 als auch -6.*

Entsprechend heißen auch die ganzen Zahlen $-2, -4, \ldots$ gerade und die ganzen Zahlen $-1, -3, \ldots$ ungerade. Den Begriff Primzahl verwenden wir dagegen nur für natürliche Zahlen.

Als Beispiel für einen mathematischen Beweis beweisen wir den Satz von Euklid. Im Beweis verwenden wir die beiden nachfolgende Lemmas und im Beweis der Lemmas die folgende Bemerkung.

3 Bemerkung. *Sei n eine von 0 und 1 verschiedene natürliche Zahl. Nach Definition hat n die trivialen Teiler 1 und n, und n hat genau dann auch nichttriviale Teiler, wenn n nicht prim ist.*

Für jeden nichttrivialen Teiler t von n gilt $1 < t < n$. Dies folgt, da ein nichttrivialer Teiler t von n nach Definition verschieden von 0, 1 und n ist. Es kann auch nicht $n < t$ gelten, da in diesem Fall für alle natürlichen Zahlen k das Produkt $k \cdot t$ verschieden von n ist. Letztere Aussage werden wir erst später mathematisch exakt beweisen, nachdem wir die natürlichen Zahlen formal eingeführt haben. Die Aussage ist aber anschaulich klar, im Fall $n < t$ sind alle Produkte der Form $k \cdot t$ gleich 0 oder echt größer n.

Die Zahlen 0 und 1 haben beide den trivialen Teiler 1, die Zahl 1 hat keine weiteren Teiler, die Zahl 0 hat noch die nichttrivialen Teiler $2, 3, \ldots$.

4 Lemma. *Die Teiler-Relation ist* TRANSITIV*, das heißt, für alle natürliche Zahlen n_1, n_2 und n_3 gilt: Falls n_1 Teiler von n_2 ist und n_2 Teiler von n_3, dann ist n_1 Teiler von n_3.*

Beweis. Seien n_1, n_2 und n_3 natürliche Zahlen, so dass n_1 Teiler von n_2 ist und n_2 Teiler von n_3. Es gibt also natürlich Zahlen k_1 und k_2 mit $n_2 = k_1 \cdot n_1$ und $n_3 = k_2 \cdot n_2$, woraus folgt

$$n_3 = k_2 \cdot n_2 = k_2 \cdot (k_1 \cdot n_1) = (k_2 \cdot k_1) \cdot n_1.$$

Die natürliche Zahl $k_2 \cdot k_1$ bezeugt also, dass n_1 Teiler von n_3 ist. □

5 Lemma. *Jede von 1 verschiedene natürliche Zahl n wird von einer Primzahl geteilt.*

Beweis. Sei n eine von 1 verschieden natürliche Zahl. Im Fall $n = 0$ ist die Behauptung offensichtlich richtig, ebenso, falls n einer Primzahl ist. Wir können also im Folgenden annehmen, dass n von 0 und 1 verschieden und keine Primzahl ist. Nach Bemerkung 3 hat n dann auch nichttriviale Teiler. Sei t der kleinste nichttriviale Teiler von n. Wir zeigen, dass t prim ist, woraus dann sofort das Lemma folgt.

Für einen Beweis durch Widerspruch nehmen wir an, dass t nicht prim ist. Dann hat t nichttriviale Teiler. Sei t' der kleinste nichttriviale Teiler von t.

Nach Lemma 4 ist t' auch Teiler von n. Nach Bemerkung 3 gilt

$$1 < t' < t < n.$$

Der Teiler t' von n ist verschieden von 1 und n und folglich nichttrivialer Teiler von n. Da t' echt kleiner als t ist, widerspricht dies der Wahl von t. Dieser Widerspruch folgt aus unserer Annahme, dass t nicht prim ist, die Annahme muss also falsch sein, folglich ist t prim. □

6 Satz von Euklid. *Es gibt unendlich viele Primzahlen.*

Beweis. Wir nehmen an, dass es nur die endlich vielen Primzahlen $p_1, \ldots, p_t$ gibt und setzen $n = p_1 \cdot p_2 \cdot \cdots \cdot p_t$. Nach Konstruktion ist dann n Vielfaches aller p_i, zum Beispiel können wir n in der Form $n = (p_2 \cdot \cdots \cdot p_t) \cdot p_1$ schreiben, die natürliche Zahl $p_2 \cdot \cdots \cdot p_t$ bezeugt also, dass n Vielfaches von p_1 ist.

Da sich zwei verschiedene Vielfache einer Zahl d mindestens um den Betrag d unterscheiden, ist die Zahl $n + 1$ nicht Vielfaches einer der Zahlen p_1 bis p_t, wird also von keiner dieser Zahlen geteilt. Da diese Zahlen aber nach Annahme alle Primzahlen umfassen, wird $n + 1$ nicht von einer Primzahl geteilt. Dies widerspricht Lemma 5. Unsere Annahme, dass es nur endlich viele Primzahlen gibt, muss somit falsch sein. □

7 Bemerkung. *In den Beweisen von Lemma 4 und des Satzes von Euklid haben wir unter anderem folgende Eigenschaften der natürlichen Zahlen verwendet, die wir erst später beweisen werden, nachdem wir die natürlichen Zahlen formal eingeführt sind.*

Die Relation $<$ ist transitiv. Die Multiplikation natürlicher Zahlen ist kommutativ, das heißt, in einem Produkt kommt es nicht auf die Reihefolge der Faktoren ankommt. Die Multiplikation ist auch assoziativ, das heißt, in einem Produkt aus mehreren Faktoren spielt die durch die Klammerung festgelegte Reihenfolge der Ausführung der Multiplikationen keine Rolle, der Wert des Produkts ändert sich nicht, wenn die Klammern anders gesetzt werden.

8 Bemerkung. *Übliche Bezeichnungen für bereits bewiesene und damit als wahr erkannte mathematische Aussagen sind* Theorem, Satz, Proposition, Beobachtung, Lemma *und* Korollar. *Theorem und Satz haben in etwa die gleiche Bedeutung und bezeichnen wichtigere und zentralere Aussagen eines mathematischen Gebiets, meist wird nur einer der beiden Begriffe verwendet. Proposition wird für Aussagen verwendet, die weniger relevant oder leicht zu zeigen sind. Eine Beobachtung ist eine Aussage, so dass die Aussage selbst oder ihr Beweis offensichtlich ist und deshalb auf eine Darstellung des Beweises verzichtet werden kann. Ein Lemma ist eine Aussage, die für den Beweis einer anderen Aussage, etwa eines Satzes, wichtig ist. Das Lemma selbst ist dabei weniger interessant, zum Beispiel weil sein Inhalt sehr technisch oder aus anderen Gründen nicht intuitiv ist. Ein Lemma*

kann einfach oder auch sehr schwierig zu beweisen sein. Ein Korollar ist eine relativ direkte Folgerung aus einer anderen Aussage, zum Beispiel aus einem Satz. Die Verwendung dieser Begriffe in der Literatur ist nicht ganz einheitlich. Zum Beispiel wird nicht immer zwischen Sätzen und Propositionen unterschieden und stattdessen für alle entsprechenden Aussagen der Begriff Satz verwendet.

9 Bemerkung. *Mathematische Aussagen, von denen nicht bekannt ist, ob sie wahr oder falsch sind, werden manchmal als* Vermutung *oder* ungelöstes Problem *bezeichet. Die* Goldbachsche Vermutung *besagt, dass jede von* 0 *und* 2 *verschiedene gerade Zahl die Summe von zwei Primzahlen ist. Beim* Problem der Primzahlzwillinge *wird danach gefragt, ob es unendlich viele Primzahlen p gibt, so dass auch $p+2$ prim ist.*

Die Goldbachsche Vermutung könnte prinzipiell durch die Angabe eines Gegenbeispiels widerlegt werden, also durch eine gerade Zahl echt größer zwei, die nicht die Summe zweier Primzahlen ist, was sich ja für eine gegebene Zahl nachrechnen lässt. Computergestütze Suchen nach einem Gegenbeispiel haben aber bisher nur für immer größere Schranken t gezeigt, dass jede von 0 *und* 2 *verschiedene gerade Zahl kleiner t tatsächlich die Summe von zwei Primzahlen ist.*

Die Aussage zum Problem der Primzahlzwillinge kann mit Gegenbeispielen weder bewiesen noch widerlegt werden, trotzdem gibt es Bestrebungen, computergestützt und unter Ausnutzung geeigneter mathematischer Zusammenhänge möglichst große p zu finden, so dass p und $p+2$ prim sind.

Im Netz finden sich Verweise auf Untersuchungen aus den 2010er Jahren, bei denen Werte für t größer als 10^{18} und Werte für p größer als 2^{151618} erreicht wurden.

1.2 Mengenlehre

1.2.1 Naive Vewendung von Mengen

Die Definition des Begriffs Primzahl, der Satz von Euklid und dessen Beweis sind in einer Sprache formuliert, die wir als mathematische Umgangssprache bezeichnen wollen. Diese ist der üblichen Umgangssprache sehr ähnlich, unterscheidet sich aber von letzterer insbesondere insofern als gewisse Konventionen eingehalten werden und die darin auftretenden Begriffe alle exakt definiert sind. Die mathematische Umgangsprache verwendet ganz wesentlich die mengentheoretische Notation, zum einen um die verwendeten mathematischen Begriffe zu definieren, zum anderen, um über diese Begriffe zu reden.

Tatsächlich kann die Mathematik vollständig mengentheoretisch aufgebaut werden. Der Einfachheit halber werden wir zu diesem Aufbau nur einigen Beispiele behandeln, etwa die Einführung der natürlichen, der rationalen

und der reellen Zahlen oder die formale Beschreibung von Relationen und Funktionen durch kartesische Produkte. Für andere im Folgenden verwendete mathematische Objekte werden wir einfach annehmen, dass diese in geeigneter Weise mengentheoretisch eingeführt werden können.

In der mathematischen Umgangssprache werden mathematische Objekte ohne Rückgriff auf deren mengentheoretisch Definition in der üblichen Weise verwendet: Sätze über Primzahlen beispielsweise lassen sich meist einfacher beweisen, wenn die mengentheoretische Definition der natürlichen Zahlen außen vor bleibt. Prinzipiell lässt sich die mathematische Umgangssprache aber vollständig in die formale mengentheoretische Notation übersetzen, da auch beim Sprechen über mathematische Objekte mengentheoretischen Begriffen und Methoden verwendet werden. Wir werden uns in diesem Abschnitt mit der mengentheoretischen Begriffen, insbesondere mit der zugehörigen Notation, soweit vertraut machen, wie diese benötig wird, um in der mathematischen Umgangssprache über mathematische Objekte zu reden.

In einem Exkurs in Abschnitt 1.2.2 werden wir uns kurz mit grundlegenden Problemen beim Aufbau der Mengenlehre beschäftigen und mit dem heute favorisierten Ansatz zu deren Lösung: einem axiomatischen Aufbau der Mengenlehre. Ansonsten werden wir die Mengenlehre in dem Sinne NAIV verwenden, dass wir uns nicht weiter um deren axiomatischen Aufbau kümmern. Unsere naive Verwendung wird dabei immer in Einklang mit der axiomatischen Mengenlehre sein, da wir uns auf bestimmte Mengen und mengentheoretische Konstruktionen beschränken, die sozusagen einen unproblematischen Teil der axiomatische Mengenlehre darstellen.

Element und Mengen Anschaulich gesprochen ist eine MENGE eine Zusammenfassung von Objekten, die als ELEMENTE der Menge bezeichnet werden, die Elemente einer Menge sind in der Menge ENTHALTEN, alle anderen Objekte sind nicht in der Menge enthalten. Mengen können durch Terme der Form $\{\ldots\}$ beschrieben werden. Wenn wir die natürlichen Zahlen mit den üblichen Bezeichnungen voraussetzen, wird zum Beispiel durch

$$\mathbb{N} = \{0, 1, 2, 3, \ldots\}, \qquad U = \{1, 3, 5, \ldots\}, \quad \text{und} \quad A = \{2, 3, 5, 7, 11\}$$

die Menge $\mathbb{N}$ der natürlichen Zahlen und die Menge U der ungeraden natürlichen Zahlen definiert, sowie eine Menge A, deren Elemente die ersten 5 Primzahlen sind. Die Elementbeziehung wird als $\in$ geschrieben, ihre Negation als $\notin$, es gilt zum Beispiel

$$1 \in \mathbb{N}, \qquad 1 \in U, \qquad 1 \notin A \quad \text{und} \quad 2 \notin U.$$

Zugehörige Sprechweisen sind: die Zahl 1 IST ELEMENT von U, ist in U ENTHALTEN oder ist IN U.

Gleichheit von Mengen gemäß dem Extensionalitätsprinzip Zwei Mengen A und B sind gleich, auch: identisch, kurz $A = B$, falls beide Mengen dieselben Elemente enthalten, also jedes Element von A auch Element von B ist und umgekehrt. Eine Menge ist also bereits durch ihre Elemente festgelegt, dies wird als EXTENSIONALITÄTSPRINZIP bezeichnet. Dieses lässt sich auch unter Verwendung der TEILMENGENBEZIEHUNG ausdrücken, wobei A eine TEILMENGE von B ist, kurz: $A \subseteq B$, falls jedes Element von A auch Element von B ist. Damit gilt dann für alle Mengen A und B,

$$A = B \text{ ist genau dann wahr, wenn } A \subseteq B \text{ und } B \subseteq A \text{ wahr ist.}$$

Mit dieser Charakterisierung lässt sich häufig die Gleichheit zweier Mengen beweisen, indem gezeigt wird, dass die Mengen wechselseitig Teilmengen voneinander sind. Eine Menge A ist eine ECHTE TEILMENGE von B, kurz: $A \subset B$, falls A Teilmenge von B, aber von B verschieden ist.

Aus dem Extensionalitätsprinzip folgt auch, dass es für eine Menge, anschaulich gesprochen, keinen Unterschied macht, ob sie ein Element ein mal oder mehrfach enthält, zum Beispiel sind nach dem Extensionalitätsprinzip die Mengen $\{1, 2, 3\}$ und $\{1, 2, 2, 3, 3, 3\}$ identisch. Dies lässt sich auch so ausdrücken, dass ein Element in einer Menge höchstes ein mal enthalten sein kann. Auch die Reihenfolge, in der die Elemente einer Menge angegeben werden, spielt keine Rolle, zum Beispiel bezeichnen die Terme $\{1, 2, 3\}$ und $\{3, 1, 2\}$ dieselbe Menge.

Leere Menge Es gibt eine leere Menge, also eine Menge ohne Elemente, geschrieben als $\{\}$ oder $\emptyset$. Nach dem Extensionalitätsprinzip kann es keine zwei verschiedenen Mengen ohne Elemente geben, die leere Menge ist somit eindeutig. Nach Definition der Teilmengenbeziehung ist die leere Menge Teilmenge jeder Menge.

Aussonderung Werden zu einer gegebenen Menge A alle Elemente von A zusammengefasst, die eine bestimmte Eigenschaft besitzen, so ergibt sich eine Menge, falls diese Eigenschaft in der mathematischen Umgangssprache und damit auch in mengentheoretischer Notation formulierbar ist. Zum Beispiel sind die Mengen

$$\{x \in \mathbb{N} \colon x > 1 \text{ und } x \text{ hat keine echten Teiler}\} \quad \text{und} \quad \{x \in \mathbb{N} \colon x \text{ ist prim}\}$$

beide gleich der Menge aller Primzahlen. Ein solcher Übergang von einer Menge zur Teilmenge aller Elemente mit einer bestimmten Eigenschaft wird als AUSSONDERUNG bezeichnet.

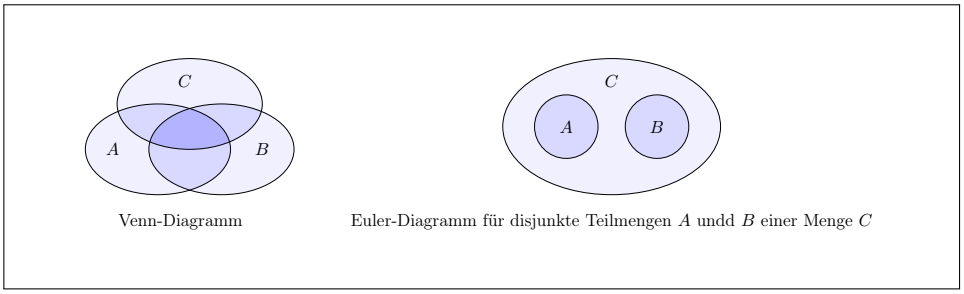

Abbildung 1.1: Venn- und Euler-Diagramme

Durchschnitt und Differenz Für zwei Mengen A und B sind deren DURCHSCHNITT, auch: SCHNITT, sowie deren DIFFERENZ definiert als

$$A \cap B = \{x \in A \colon x \in B\} \qquad \text{beziehungsweise} \qquad A \setminus B = \{x \in A \colon x \notin B\},$$

diese sind wieder Mengen, da beide aus der Menge A durch Aussonderung entstehen.

Zwei Mengen heißen DISJUNKT, falls ihr Schnitt gleich der leeren Menge ist, sie also keine gemeinsamen Elemente haben.

Für eine Teilmenge B von A wird $A \setminus B$ auch als RELATIVES KOMPLEMENT von B bezüglich A bezeichnet. Ist aus dem Zusammenhang klar, dass eine feste Menge G als Grundmenge angesehen wird, so wird für eine Teilmenge X von G das relative Komplement $G \setminus X$ auch als KOMPLEMENT von X, kurz $\overline{X}$, bezeichnet.

Vereinigung und symmetrische Differenz Für zwei Mengen A und B ist deren VEREINIGUNG $A \cup B$ eine Menge, die dadurch definiert ist, dass für alle x gilt

$$x \in A \cup B \qquad \text{genau dann, wenn} \qquad x \in A \text{ oder } x \in B,$$

die Vereinigung von A und B ist also die eindeutig definierte Menge, welche genau die Elemente von A und von B enthält.

Mit Aussonderung ist dann auch die SYMMETRISCHE DIFFERENZ

$$A \triangle B \;=\; (A \setminus B) \cup (B \setminus A) \;=\; \{x \in A \cup B \colon x \text{ ist nicht in } A \text{ und in } B\}$$

zweier Mengen A und B eine Menge. Die Menge $A \triangle B$ enthält also genau die x, die entweder Element von A oder Element von B sind, die also genau in einer der beiden Mengen enthalten sind.

10 Bemerkung. *Mengen und ihre Beziehungen durcheinander können graphisch dargestellt werden, indem jede einzelne Menge als Teilmenge der Ebene dargestellt wird, in Form einer geometrischen Figur wie eines Kreise oder*

einer Ellipsen, und die Überlappungen dieser Figuren für die Schnittmengen der entsprechenden Mengen stehen. In einem VENN-DIAGRAMM *entsprechen die verschiedenen Schnittmenge der beteiligten Mengen jeweils einem nichtleeren Überlappungsbereich, in einem* EULER-DIAGRAMM *können Überlappungsbereiche zu leeren Schnittmengen weggelassen werden, siehe Abbildung 1.1.*

11 Bemerkung. *Zu zwei Mengen A und B gibt es acht Teilmengen, die als Vereinigung der drei Mengen $A \setminus B$, $A \cap B$ und $B \setminus A$ dargestellt werden können. Sechs dieser Teilmengen sind in Abbildung 1.2 als Venn-Diagramme dargestellt, es fehlen die Teilmengen A und B.*

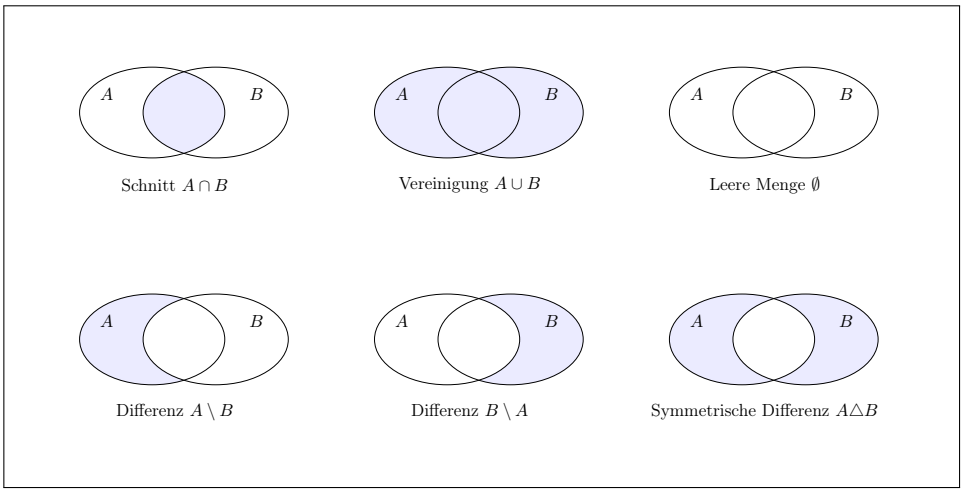

Abbildung 1.2: Mengentheoretische Operationen im Venn-Diagramm

Potenzmenge Mengen können andere Mengen als Elemente enthalten. Zu jeder Menge A gibt es eine Menge $\mathrm{Pot}(A)$, die POTENZMENGE von A, deren Elemente die Teilmengen von A sind. Nach dem Extensionalitätsprinzip kann es nur eine Menge mit letzterer Eigenschaft geben, die Potenzmenge einer Menge ist also eindeutig. Zum Beispiel gilt

$$\mathrm{Pot}(\{1,2,3\}) = \{\emptyset,\ \{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}\}.$$

Paarmenge und geordnetes Paar Zu je zwei Mengen A und B gibt es eine Menge $\{A,B\}$, die PAARMENGE von A und B, die diesen beiden Mengen als einzige Elemente enthält. Falls A ungleich B ist, kurz: $A \neq B$, hat die Paarmenge zwei Elemente, sonst nur eins.

Für zwei Mengen A und B sind dann auch $\{A\} = \{A,A\}$ und $\{A,B\}$ Mengen, und damit auch das wie folgt definierte GEORDNETE PAAR (A,B) der

Mengen A und B

$$(A, B) \;=\; \{\{A\}, \{A, B\}\}.$$

Im Gegensatz zur Paarmenge kommt es für die Gleichheit zweier geordneten Paare auf die Reihenfolge an: die Paarmengen $\{A, B\}$ und $\{B, A\}$ sind immer identisch, während im Fall A ungleich B die geordneten Paare (A, B) und (B, A) verschieden sind. Entsprechend wird A als ERSTE und B als ZWEITE KOMPONENTE des geordneten Paars (A, B) bezeichnet.

12 Lemma. *Seien A_0, B_0, A_1 und B_1 Mengen. Die geordneten Paare (A_0, B_0) und (A_1, B_1) sind genau dann identisch, wenn A_0 gleich A_1 und B_0 gleich B_1 ist.*

Beweis. Falls A_0 gleich A_1 und B_0 gleich B_1 ist, sind die Paare (A_0, B_0) und (A_1, B_1) offensichtlich identisch, denn es folgt

$$\{\{A_0\}, \{A_0, B_0\}\} \;=\; \{\{A_1\}, \{A_1, B_1\}. \tag{1.1}$$

indem im Ausdruck links A_0 durch A_1 und B_0 durch B_1 ersetzt wird.

Für den Beweis der umgekehrte Richtung der Äquivalenz sei angenommen, dass die geordneten Paare (A_0, B_0) und (A_1, B_1) identisch sind, was nach Definition gerade bedeutet, dass (1.1) wahr ist. Wir unterscheiden die Fälle $A_0 = B_0$ und $A_0 \neq B_0$. Im ersten Fall steht auf der linken Seite von (1.1) die Menge $\{\{A_0\}\}$, also eine Menge die nur das Element $\{A_0\}$ enthält und somit nur dann gleich der Menge auf der rechten Seite sein kann, wenn diese ebenfalls nur ein Element enthält. Es muss also $A_1 = B_1$ gelten und weiter $\{A_0\} = \{A_1\}$, folglich sind alle vier betrachteten Mengen identisch. Im zweiten Fall enthält die Menge auf der linken Seite von (1.1) eine ein- und eine zweielementige Menge, dies muss dann auch für die Menge auf der rechten Seite gelten, insbesondere folgt $A_1 \neq B_1$. Gleichung (1.1) kann also nur erfüllt sein, falls $\{A_0\} = \{A_1\}$ und $\{A_0, B_0\} = \{A_1, B_1\}$ gilt. Es folgt $A_0 = A_1$ und damit $B_0 = B_1$. □

Kartesisches Produkt Das KARTESISCHE PRODUKT $A \times B$ zweier Menge A und B ist die durch

$$A \times B \;=\; \{(a, b) \colon a \in A \text{ und } b \in B\}$$

definierte Menge aller geordneten Paare mit erster Komponente in A und zweiter Komponente in B.

Endliche und unendliche Durchschnitte, Vereinigungen und Produkte Es sei X eine nichtleere Menge. Der DURCHSCHNITT DER MENGEN IN X, auch: DURCHSCHNITT VON oder ÜBER X, ist definiert als

$$\bigcap X = \{x \in A \colon \text{ für alle } z \in X \text{ gilt } x \in z\}.$$

wobei A eine beliebige Menge in X ist. Der Durchschnitt über X ist eine Menge, da er durch Aussonderung aus der Menge A gebildet wird. Es lässt sich zeigen, dass der Durchschnitt über X insofern wohldefiniert ist, dass er nicht von der Wahl der Menge A in X ankommt, aus der ausgesondert wird.

Die Vereinigung einer Menge X, kurz: $\bigcup X$, auch: VEREINIGUNG VON oder ÜBER X, ist dadurch definiert, dass für alle x gilt

$$x \in \bigcup X \qquad \text{genau dann, wenn} \qquad \text{es ein } z \in X \text{ mit } x \in z \text{ gibt.}$$

Ist beispielsweise $X = \{\{1,4\},\{1,2,4\},\{1,2,3,4\}$, so gilt

$$\bigcap X = \{1,4\} \qquad \text{und} \qquad \bigcup X = \{1,2,3,4\}.$$

Die Vereinigung einer Menge X ist wieder eine Menge, dies wird durch ein entsprechendes Axiom sichergestellt. Die Schreibweisen $\bigcap X$ und $\bigcup X$ sind in der Mengenlehre üblich, in anderen Gebieten werden stattdessen meist die folgenden äquivalenten Schreibweisen verwendet. Für eine endliche Menge der Form $X = \{X_1, \ldots, X_n\}$ definieren wir

$$\bigcap_{i=1,\ldots,n} X_i \;=\; \bigcap_{i=1}^{n} X_i \;=\; \bigcap X,$$

und für eine unendliche Menge der Form $X = \{X_1, X_2, \ldots\}$

$$\bigcap_{i=1,2,\ldots} X_i \;=\; \bigcap_{i=1}^{\infty} X_i \;=\; \bigcap X$$

und analog für Vereinigungen. Zum Beispiel gilt für $X = \{X_1, X_2, \ldots\}$ mit $X_i = \{1, \ldots, i\}$

$$\bigcap_{i=1,2,\ldots} X_i \;=\; \{1\} \qquad \text{und} \qquad \bigcup_{i=1,2,\ldots} X_i \;=\; \{1,2,\ldots\}.$$

13 Bemerkung. *Endlich viele Mengen $A_1, \ldots, A_n$ beziehungsweise unendlich viele Mengen $A_1, A_2, \ldots$ werden als* PAARWEISE DISJUNKT *bezeichnet, falls je zwei verschiedene dieser Mengen disjunkt sind, das heißt, für alle Indizes i und j mit $i \neq j$ gilt $A_i \cap A_j = \emptyset$. Der Schnitt von paarweise disjunkte Mengen ist gleich der leeren Menge, die Umkehrung gilt nicht, drei oder mehr Mengen mit leerem Schnitt sind im Allgemeinen nicht paarweise disjunkt.*

Auch der Begriff kartesisches Produkt lässt sich auf den Fall endlich oder unendlich vieler Faktoren erweitern. Die Elemente des kartesischen Produktes von Mengen $X_1, \ldots, X_n$ werden als n-Tupel bezeichnet und in der Form $(x_1, \ldots, x_n)$ geschrieben, das kartesische Produkt selbst ist

$$\bigtimes_{i=1,\ldots,n} X_i \;= X_1 \times \cdots \times X_n = \{(x_1, \ldots, x_n) \colon x_i \in X_i \text{ für } i = 1, \ldots, n\}.$$

Endliche kartesische Produkte werden wieder so eingeführt, dass je zwei Tupel $(x_1, \ldots, x_n)$ und $(x'_1, \ldots, x'_n)$ genau dann identisch sind, wenn $x_i = x'_i$ für $i = 1, \ldots, n$ gilt, es kommt also auch bei n-Tupeln auf die Reihenfolge der Komponenten an. Im Fall $n = 1$ wird ein Tupel der Form (x) mit dem Element x von X_1 identifiziert und entsprechend das kartesische Produkt gleich X_1 gesetzt. Falls die Mengen $X_1, \ldots, X_n$ alle gleich einer Menge X sind, wird das kartesische Produkt der X_i auch als X^n geschrieben. In diesem Fall werden die n-Tupel im kartesischen Produkt der X auch als ENDLICHE FOLGE DER LÄNGE n über X bezeichnet und in der Form $x_1, \ldots, x_n$ geschrieben.

Das UNENDLICHE KARTESISCHE PRODUKT von Mengen $X_1, X_2, \ldots$ ist

$$\bigtimes_{i=1,2,\ldots} X_i \;=\; X_1 \times X_2 \times \cdots \;=\; \{x_1, x_2, \ldots : x_i \in X_i \text{ für } i = 1, 2 \ldots\},$$

die Elemente $x_1, x_2, \ldots$ eines solchen Produkts werden als UNENDLICHE FOLGEN bezeichnet. Zwei Folgen $x_1, x_2, \ldots$ und $x'_1, x'_2, \ldots$ sind genau dann gleich, falls für alle i gilt $x_i = x'_i$. Im Fall, dass die X_i alle gleich einer Menge X sind, wird das kartesische Produkt der X_i mit X^∞ bezeichnet, seine Elemente als UNENDLICHE FOLGEN VON ELEMENTEN AUS X oder ÜBER X.

Limes inferior und Limes superior von Mengenfolgen Zu einer Folge $A_1, A_2, \ldots$ von Mengen sei deren LIMES INFERIOR und LIMES SUPERIOR definiert als

$$\liminf_{i\to\infty} A_i \;=\; \bigcup_{m=1}^{\infty} \bigcap_{i=m}^{\infty} A_i \quad \text{beziehungsweise} \quad \limsup_{i\to\infty} A_i \;=\; \bigcap_{m=1}^{\infty} \bigcup_{i=m}^{\infty} A_i.$$

14 Bemerkung. *Der Limes inferior einer Folge $A_1, A_2, \ldots$ von Mengen ist gleich der Menge aller x, die in dem Sinne in* FAST ALLEN *A_i enthalten sind, dass sie nur in endlich vielen der Mengen A_i nicht enthalten sind.*

Der Limes superior einer solchen Folge ist gleich der Menge aller x, die in unendlich vielen A_i enthalten sind, also aller x, so dass es beliebig große Indizes i mit $x \in A_i$ gibt.

Wir beweisen die erste der beiden Aussagen und verzichten auf den ähnlichen Beweis der zweiten Aussage. Sei also A_1, $A_2, \ldots$ eine Folge von Mengen mit Limes inferior A, und sei $D_m = \bigcap_{i=m}^{\infty} A_i$. Weiter sei B gleich der Menge aller x, die in fast allen A_i enthalten sind. Es genügt, die Inklusionen $A \subseteq B$ und $B \subseteq A$ zu zeigen.

Wir zeigen zunächst $A \subseteq B$. Sei dazu x ein beliebiges Element von A. Da A gleich der Vereinigung der Mengen $D_1, D_2, \ldots$ ist, gibt es einen Index m mit $x \in D_m$. Nach Definition von D_m ist x in allen Mengen A_i mit $i \geq m$ enthalten und ist somit in B.

Für einen Beweis von $B \subseteq A$ sei x ein beliebiges Element von B, es gibt also einen Index m, so dass x in allen Mengen A_i mit $i \geq m$ enthalten ist. Für diesen Index m gilt dann $x \in D_m$. Da D_m Teilmenge von A ist, folgt $x \in A$.

1.2.2 Exkurs über axiomatische Mengenlehre

Ende des 19. Jahrhunderts wurden, unter anderem von Cantor, Mengen im Sinne von Zusammenfassungen von Elementen eingeführt. Dabei wurde zunächst postuliert, dass es zu jeder mathematisch formulierbaren Eigenschaft eine Menge aller Objekte mit dieser Eigenschaft gibt, also zum Beispiel die Menge aller Primzahlen oder die Menge aller Kreise in einer Ebene. Wenige Jahre später konnte Russell zeigen, dass dies für bestimmte Eigenschaften zu Widersprüchen führt.

Russell betrachtete die Eigenschaft einer Menge, sich nicht selbst zu enthalten. Diese Eigenschaft trifft zum Beispiel auf die leere Menge zu. Russell zeigte, dass es keine Menge R geben kann, die genau die Mengen mit dieser Eigenschaft enthält, für die also gilt

$$R = \{x \colon x \notin x\}.$$

Angenommen es gäbe eine solche Menge R. Dann gibt es zwei mögliche Fälle: R kann sich selbst enthalten oder nicht. Falls sich R selbst enthält, folgt nach Definition von R, dass R kein Element von R ist. Falls sich R nicht selbst enthält, folgt nach Definition von R, dass R Element von R ist. In beiden Fällen ergibt sich ein Widerspruch. Dies kann auch so ausgedrückt werden, dass für die Menge R gelten müsste

$$R \in R \quad \text{genau dann, wenn} \quad R \notin R.$$

Es kann also keine solche Menge R geben. Die auf den ersten Blick plausibel erscheinenden Annahme, dass R eine Menge ist, führt zu einem Widerspruch, und wird deshalb als Russellsche Antinomie oder Russellsches Paradoxon bezeichnet.

In der axiomatischen Mengenlehre wird versucht, durch die Wahl geeigneter Axiome zu einer widerspruchsfreien Mengenlehre zu kommen. Zum Beispiel wird durch ein spezielles Axiom ausgeschlossen, dass sich eine Menge selbst enthält. Daraus folgt zunächst, dass es keine Menge geben kann, die alle Mengen enthält, diese müsste sich ja auch selbst enthalten. Damit ist aber auch R keine Menge, da sich nach unserer Annahme keine Menge selbst enthält, R also alle Mengen enthält.

Das Kernproblem der Mengenlehre ist, dass sich Widersprüche ergeben, wenn beliebige Zusammenfassungen von Objekten zu Mengen erlaubt sind. Die Bildung von Mengen muss somit eingeschränkt werden. Andererseits sollen alle

Mengen verfügbar sein, die zum Aufbau der Mathematik benötigt werden, zum Beispiel Mengen von geordneten Paaren um Relationen und Funktionen darstellen zu können. Die Axiome der Mengenlehre stellen dies sicher, indem sie garantieren, dass bestimme Zusammenfassungen von Objekten tatsächlich Mengen sind und weiter bestimmte Konstruktionsmöglichkeiten erzwingen, um aus gegebenen Mengen neue Mengen zu erhalten. Einige dieser Axiome wurden bereits im Abschnittt 1.2.1 über die naive Verwendung von Mengen angesprochen, zum Beispiel wird jeweils durch ein Axiom sichergestellt, dass zu jeder Menge deren Potenzmenge und zu je zwei Mengen die zugehörige Paarmenge existiert. Im Folgenden stellen wir kurz einige weitere Axiome vor. Diese Axiome stellen eher technische Eigenschaften von Mengen sicher, die im hier verwendeten Teil der Mengenlehre keine wesentliche Rolle spielen. Für eine ausführlichere Darstellung der axiomatische Mengenlehre sei auf das gut lesbare Lehrbuch *Einführung in die Mengenlehre* von Ebbinghaus (Springer Spektrum, 2021) verwiesen.

Unendliche Menge Eines der Axiome stellt sicher, dass es eine INDUKTIVE MENGE gibt, das heißt, die Menge enthält die leere Menge $\emptyset$ und mit jeder Menge z auch die Menge $z \cup \{z\}$. Daraus folgt insbesondere dass $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots\}$ eine Menge ist, mit dieser werden wir später die natürlichen Zahlen formal einführen.

Fundierung Die Elementbeziehung ist FUNDIERT, insbesondere gibt es keine sich schließenden Ketten von Elementbeziehungen wie $A \in B \in C \in A$.

Auswahl Für bestimmte Anwendungen in der Mathematik wird gefordert, dass es zu jeder nichtleeren Menge A, deren Elemente paarweise disjunkte Menge sind, eine AUSWAHLMENGE gibt, die aus jeder Menge in A genau ein Element enthält. Wir werden später Funktionen formal als Mengen von geordneten Paaren einführen. Äquivalent zur Existenz von Auswahlmengen wie gerade beschrieben ist dann die Aussage, dass es zu jeder nichtleeren Menge, deren Elemente alle nichtleere Mengen sind, eine Funktion gibt, die jedem solchen Element X ein Element von X zuweist. Das Auswahlaxiom wird zwar in der Mathematik allgemein verwendet, ist aber auch umstritten, da es paradox erscheinende Konsequenzen wie das Banach-Tarski-Paradoxon hat. Dieses Paradoxon besagt, dass unter Verwendung des Auswahlaxioms eine Kugel so zerlegt werden kann, dass aus den sich ergebenden Teilen zwei disjunkte Kugeln zusammengesetzt werden können, die beide dieselbe Größe wie die ursprüngliche Kugel haben.

Zermelo-Fraenkel Ein weit verbreitetes Axiomensystem ist das von Zermelo und Fraenkel. Dieses enthält neben den bereits genannten noch weitere

Axiome, die zum Beispiel sicherstelllen, dass das Bild einer Menge unter einer Funktion wieder eine Menge ist. Das Auswahlaxiom, englisch: axiom of choice, ist nicht Teil des Axiomensystems, wird es hinzugenommen, erhält man das Axiomensystem von Zermelo und Fraenkel plus Auswahl, kurz: ZFC, das als eine Art Standardaxiomensystem der Mengenlehre angesehen werden kann.

1.3 Eigenschaften und Relationen

Eigenschaften von mathematischen Objekten und Beziehungen zwischen solchen Objekten lassen sich formal als Mengen darstellen, zum Beispiel

> die Eigenschaft eine Primzahl zu sein, durch die Menge der Primzahlen,
>
> die übliche strikte Ordnung $<$ auf den natürlichen Zahlen durch die Menge aller geordneten Paare (x, y) von natürlichen Zahlen mit $x < y$,
>
> die Summation von zwei natürlichen Zahlen zu einer dritten durch die Menge aller Tripel (a, b, s) von natürlichen Zahlen mit $a + b = s$.

15 Definition. *Es seien n eine natürliche Zahl und A eine Menge. Eine* n-STELLIGE RELATION ÜBER A *ist eine Teilmenge des n-fachen kartesischen Produkts A^n. Für eine solche Relation R wird auch $R(x_1, \ldots, x_n)$ geschrieben um auszudrücken, dass das Tupel $(x_1, \ldots, x_n)$ Element von R ist.*

Einstellige Relationen über A werden als EIGENSCHAFTEN *bezeichnet und mit Teilmengen von A identifiziert. Für eine solche Eigenschaft E bedeutet somit $E(x)$, dass x Element der Teilmenge E von A ist.*

Beziehungen zwischen Elementen verschiedener Mengen können in ähnlicher Weise als Mengen von n-Tupeln dargestellt werden, zum Beispiel die Beziehung zwischer einer natürlichen Zahl und ihrer Quadratwurzel aus der Menge $\mathbb{R}$ der reellen Zahl durch die Menge

$$\{(n, r) \in \mathbb{N} \times \mathbb{R} \colon \ r \text{ ist die Quadratwurzel von } n\}$$

16 Bemerkung. *Die Teilmengenrelation $\subseteq$ zwischen den Teilmengen einer gegebenen Menge A wird durch die Menge*

$$\{(x, y) \colon x \text{ und } y \text{ sind Teilmengen von } A \text{ mit } x \subseteq y\}$$

dargestelllt. Die Teilmengenrelation als Ganzes, das heißt als Relation zwischen beliebigen Mengen, kann dagegen nicht als Menge von Paaren dargestellt werden. Eine solche Menge würde auch alle Paare der Form (x, x) für eine beliebige Menge x enthalten. Aus den mengentheoretischen Axiomen würde dann folgen, dass es eine Menge aller Mengen gibt, und durch Aussonderung würde sich die Menge aller Mengen, die sich nicht selbst enthalten, ergeben, was gemäß des Russellschen Paradoxons ein Widerspruch ist. Aus ähnlichen Gründen kann auch die Elementrelation als Ganzes nicht als Menge von Paaren dargestellt werden.

1.3.1 Äquivalenzklassen und Äquivalenzrelationen

Im Rest dieses Unterkapitels werden wir uns einige wichtige Typen von Relationen anschauen: Äquivalenzrelationen und verschiedene Varianten von

Ordnungsrelationen. Diese Relationen sind alle zweistellig und der Einfachheit halber sei vereinbart, dass die hier betrachteten Relationen zweistellig sind, falls die Stelligkeit nicht explizit angegeben ist. Die betrachteten Relationen werden meist in INFIXNOTATION geschrieben, zum Beispiel steht xRy für $R(x, y)$, also dafür, dass das Paar (x, y) Element von R ist.

Die betrachteten Relationen sind alle reflexiv und transitiv, ausgenommen die am Ende behandelten strikten Varianten von partiellen Ordnungen.

17 Definition. *Es sei R eine zweistellige Relation auf einer Menge A.*

Die Relation R ist REFLEXIV*, wenn für alle x in A gilt xRx.*

Die Relation R ist TRANSITIV*, wenn für alle x_1, x_2 und x_3 in A mit x_1Rx_2 und x_2Rx_3 auch x_1Rx_3 gilt.*

18 Definition. *Für eine reflexive und transitive Relation R auf einer Menge A ist die* ÄQUIVALENZKLASSE *eines Elements x von A gleich*

$$[x]_R \;=\; \{y \in A\colon xRy \text{ und } yRx\}.$$

Für eine reflexive und transitive Relation R auf einer Menge A ist sjedes x in A in seiner Äquivalenzklasse $[x]_R$, aber in keiner anderen Äquivalenzklasse enthalten. Dies bedeutet, dass A in dem Sinne in Äquivalenzklassen zerfällt, dass jedes x in A in genau einer Äquivalenzklasse enthalten ist.

19 Definition. *Eine Menge Z ist eine* ZERLEGUNG*, auch:* PARTITION*, einer Menge A, falls Z die leere Menge nicht enthält, die Menge A gleich der Vereinigung der Mengen in Z ist und die Mengen in Z paarweise disjunkt sind, falls also gilt $\emptyset \notin Z$,*

$$A \;=\; \bigcup Z \quad \text{und} \quad X \cap X' \;=\; \emptyset \text{ für alle Mengen } X, X' \text{ in } Z \text{ mit } X \neq X'.$$

Nach Definition sind die Mengen in einer Zerlegung einer Menge A alle nichtleere Teilmengen von A und jedes Element von A ist in genau einer dieser Teilmengen enthalten. Dass eine endliche Menge $\{X_1, \ldots, X_n\}$ eine Zerlegung von A ist, kann auch so gedrückt werden, dass die Mengen $X_1, \ldots, X_n$ eine Zerlegung von A bilden und entsprechend für eine unendlich Zerlegung $\{X_1, X_2, \ldots\}$.[1]

20 Beispiel. *Die Menge* G *und* U *der geraden beziehungsweise ungeraden Zahlen bilden eine Zerlegung $\{G, U\}$ der Menge $\mathbb{N}$ der natürlichen Zahlen.*

[1] Eine Zerlegung $\{X, Y\}$ einer Menge A kann für $Z = Y$ auch als $\{X, Y, Z\}$ geschrieben werden. Um dies zu vermeiden, vereinbaren wir, dass falls eine Zerlegung in der Form $\{X_1, \ldots, X_n\}$ oder $\{X_1, X_2, \ldots\}$ angegeben wird, im Folgenden immer vorausgesetzt sei, dass für zwei verschiedene Indizes i und j die Mengen X_i und X_j verschieden sind.

Die Menge $\{\{1,5\},\{2\},\{3,4\}\}$ *ist eine Zerlegung der Menge* $\{1,2,3,4,5\}$. *Eine Zerlegung von* $\mathbb{N}$ *in unendlich viele Mengen ist*

$$\{\{2^n,\ldots,2^{n+1}-1\}\colon n\in\mathbb{N}\}\cup\{\{0\}\} \;=\; \{\{0\},\{1\},\{2,3\},\{4,5,6,7\},\ldots\}.$$

21 Lemma. *Es sei* R *eine reflexive und transitive Relation auf einer Menge* A. *Dann bilden die Äquivalenzklassen von* R *eine Zerlegung von* A.

Beweis. Wir zeigen, dass $Z=\{[x]_R\colon x\in A\}$ eine Zerlegung von A ist.

Da R reflexiv ist, ist jedes Element von A Element seiner Äquivalenzklasse $[x]_R$ enthalten, die Äquivalenzklassen sind somit alle nichtleer und ihre Vereinigung ist gleich A. Es bleibt zu zeigen, dass die Äquivalenzklassen paarweise disjunkt sind. Seien also x und y Elemente von A mit $[x]_R\neq[y]_R$. Für einen Beweis durch Widerspruch nehmen wir an, dass der Schnitt der beiden Klassen ein Element z von A enthält und folgern daraus im Widerspruch zur Wahl von x und y, dass $[x]_R=[y]_R$ gilt. Dabei zeigen wir nur die Inklusion $[x]_R\subseteq[y]_R$ und verzichten auf den im Wesentlichen identischen Beweis der symmetrischen Inklusion $[y]_R\subseteq[x]_R$. Sei also v ein beliebiges Element von $[x]_R$. Dann gilt nach Annahme über z

$$vRx \quad\text{und}\quad xRz \quad\text{und}\quad zRy.$$

Da R transitiv ist, folgt zunächst vRz und dann vRy. Ganz ähnlich folgt yRv und somit ist v Element von $[y]_R$. Es folgt $[x]_R\subseteq[y]_R$, da v in $[x]_R$ beliebig gewählt war. □

Für eine reflexive und transitive Relation auf einer Menge A stehen genau die Elemente von A wechselseitig zueinander in Relation, die in derselben Äquivalenzklasse liegen.

22 Korollar. *Es sei* R *eine reflexive und transitive Relation auf einer Menge* A, *und* x *und* y *seien Elemente von* A. *Dann sind die folgenden Aussagen äquivalent.*

(i) *Es gilt* xRy *und* yRx.

(ii) *Das Element* x *liegt in der Äquivalenzklasse* $[y]_R$.

(iii) *Das Element* y *liegt in der Äquivalenzklasse* $[x]_R$.

(iv) *Die Äquivalenzklassen* $[x]_R$ *und* $[y]_R$ *sind identisch.*

(v) *Die Elemente* x *und* y *liegen in derselben Äquivalenzklasse.*

Beweis. Aus (i) folgt nach Definition von Äquivalenzklasse sowohl (ii) als auch (iii). Im Rest des Beweises wird mehrfach benutzt, dass die Äquivalenzklassen eine Zerlegung von A bilden und somit jedes z in A in genau einer

Äquivalenzklasse liegt, die wegen der Reflexivität gleich $[z]_R$ ist. Liegt x in der Äquivalenzklasse $[y]_R$, müssen $[x]_R$ und $[y]_R$ identisch sein, aus (ii) folgt somit (iv), und ein symmetrische Argument zeigt die Implikation von (iii) nach (iv). Aus (iv) folgt dann (v), weil $[x]_R$ und $[y]_R$ die eindeutig bestimmten Äquivalenzklassen sind, in denen x beziehungsweise y liegen. Falls (v) gilt, liegt y in der Äquivalenzklasse $[x]_R$ von x und (i) folgt nach Definition von Äquivalenzklasse. □

Es sei R eine reflexive und transitive Relation auf einer Menge A und x und y seien zwei Elemente von A. Nach Korollar 22 stehen x und y genau dann wechselseitig zueinander in Relation, wenn beide in derselben Äquivalenzklasse liegen. Sind x und y dagegen in verschiedenen Äquivalenzklassen, kann höchstens eine der Aussagen xRy oder yRx gelten. Ein Spezialfall davon ist, dass Elemente in verschiedenen Äquivalenzklassen nie zueinander in Relation stehen. In diesem Fall ist die Relation durch ihre Äquivalenzklassen bereits vollständig festgelegt, die Relation wird durch die Äquivalenzklassen im Sinne der folgenden Definition induziert.

23 Definition. *Es sei Z eine Zerlegung einer Menge A. Die durch Z auf A* INDUZIERTE *Relation ist die wie folgt definierte zweistellige Relation $\sim$ auf A*

$$x \sim y \quad \textit{genau dann, wenn} \quad \textit{es eine Menge } X \textit{ in } Z \textit{ mit } x, y \in X \textit{ gibt.}$$

Eine ÄQUIVALENZRELATION *auf einer Menge ist eine Relation, die durch eine Zerlegung der Menge induziert wird.*

Wir werden gleich sehen, dass Äquivalenzrelationen immer reflexiv und transitiv sind und durch die Äquivalenzklassen der Relation induziert werden, woraus insbesondere folgt, dass Element in verschiedenen Äquivalenzklassen nie zueinander in Relation stehen.

Eine Zerlegung einer Menge ist, anschaulich gesprochen, eine Klassifizierung der Elemente dieser Menge gemäß einer Eigenschaft oder gemäß dem Wert eines Parameters, wobei die Klassen gleich den Mengen der Zerlegung sind. Wird eine Äquivalenzrelation auf einer Menge durch eine Zerlegung der Menge induziert, so stehen je zwei Elemente der Menge genau dann in Relation zueinander, wenn sie in derselben Klasse liegen. Die Relation zwischen Personen, dasselbe Geburtsjahr zu haben ist eine Äquivalenzrelation, ebenso die Relation zwischen natürlichen Zahlen, beim Teilen durch 5 denselben Rest aus $\{0, \ldots, 4\}$ zu ergeben.

Äquivalenzrelationen lassen sich äquivalent durch die Eigenschaft symmetrisch zu sein beschreiben. Die Idee dabei ist, dass für eine reflexive und transitive Relation nach Korollar 22 Elemente aus verschiedenen Äquivalenzklassen nicht wechselseitig in Relation zueinander stehen können. Für eine symmetrische Relation können solche Elemente dann auch nicht nur

in einer Richtung in Relation stehen, da daraus die wechselseitige Relation folgen würde.

24 Definition. *Eine Relation R auf einer Menge A ist* SYMMETRISCH*, wenn für alle x und y in A aus xRy folgt yRx.*

25 Satz. *Eine Relation ist genau dann eine Äquivalenzrelation, wenn die Relation reflexiv, transitiv und symmetrisch ist.*

Beweis. Zunächst sei R eine Äquivalenzrelation auf einer Menge A. Dann können wir eine Zerlegung Z von A wählen, welche die Relation R induziert. Jedes x in A ist in einer Menge der Zerlegung Z enthalten und es gilt somit xRx, folglich ist die Relation R reflexiv. Falls xRy gilt, liegen x und y in derselben Menge der Zerlegung, es gilt also auch yRx, folglich ist die Relation R symmetrisch. Falls xRy und yRz gilt, liegen x und y sowie y und z jeweils beide in einer Menge in Z. Da y aber genau in einer Menge der Zerlegung liegt, müssen beide Mengen identisch sein und somit gilt auch xRz. Die Relation R ist also transitiv.

Sei nun umgekehrt R eine reflexive, transitive und symmetrische Relation auf einer Menge A. Nach Lemma 21 bilden die Äquivalenzklassen von R eine Zerlegung von A. Nach Definition von Äquivalenzrelation genügt es somit zu zeigen, dass diese Zerlegung die Relation R induziert, dass also für alle x und y in A genau dann xRy gilt, wenn x und y in derselben Äquivalenzklasse liegen. Dies ist aber gerade die Äquivalenz der Aussagen (i) und (iv) in Korollar 22, wenn berücksichtigt wird, dass für eine symmetrische Relation genau dann xRy gilt, wenn sowohl xRy als auch yRx gelten. □

26 Bemerkung. *Die durch eine Zerlegung einer Menge A induzierte Äquivalenzrelation auf A hängt nur von der Zerlegung ab und ist somit eindeutig. Umgekehrt gibt es zu einer Äquivalenzrelation auf einer Menge A genau eine Zerlegung, welche diese Relation induziert, und diese ist gerade die Zerlegung, die von den Äquivalenzklassen der Relation gebildet wird.*

Für einen Beweis sei R eine Äquivalenzrelation auf einer Menge A. Gemäß der Vorwärtrichtung der Äquivalenz in Satz 25 ist R dann reflexiv, transitiv und symmetrisch. Weiter wird im Beweis der anderen Richtung der Äquivalenz gezeigt, dass jede Relation mit diesen drei Eigenschaften von der von den Äquivalenzklassen der Relation gebildeten Zerlegung induziert wird.

Nach Korollar 22 gilt für alle x und y in A genau dann xRy und yRx, wenn x und y beide in derselben Äquivalenzklasse von R liegen. Für jede Zerlegung Z von A, welche die Relation R induziert, muss deshalbe jede Äquivalenzklasse Teilmenge einer Menge in Z sein, wohingegen keine Menge in Z Elemente aus verschiedenen Äquivalenzklassen enthalten darf. Dies ist nur möglich, wenn die Mengen in Z gerade die Äquivalenzklassen sind.

27 Bemerkung. *Der Begriff Äquivalenzklasse wird üblicherweise nur im Zusammenhang mit Äquivalenzrelationen verwendet. Wir weichen davon ab und verwenden den Begriff Äquivalenzklasse auch für reflexive und transitive, aber nicht notwendigerweise symmetrische Relationen. Für eine solche Relation bilden die Äquivalenzrelationen nach Lemma 21 eine Zerlegung der betrachteten Mengen, im Fall einer nichtsymmetrischen Relation wird die Relation aber nicht durch diese Zerlegung induziert.*

28 Definition. *Für eine Äquivalenzrelation heißen die Elemente einer Äquivalenzklasse* Repräsentanten*, auch:* Vertreter*, dieser Äquivalenzklassen.*

29 Bemerkung. *Eigenschaften von Äquivalenzklassen müssen in dem Sinne* wohldefiniert *sein, dass es nicht vom betrachteten Repräsentanten abhängt, ob die Eigenschaft zutrifft. Entsprechendes gilt später für Funktionen: der Funktionswert einer Äquivalenzklasse darf nicht vom gewählten Repräsentanten abhängen.*

Als Beispiel betrachten wir die Äquivalenzrelation $\sim$ auf den natürlichen Zahlen, bei Division durch 4 denselben Rest zu ergeben. Diese Äquivalenzrelation hat die vier Äquivalenzklassen

$$[1]_\sim = \{1, 5, \ldots\},\ \ [2]_\sim = \{2, 6, \ldots\},\ \ [3]_\sim = \{3, 7, \ldots\},\ \ [4]_\sim = \{4, 8, \ldots\},$$

und es gilt zum Beispiel $[1]_\sim = [5]_\sim$. Wird nun eine Äquivalenzklasse der Form $[n]_\sim$ als gerade definiert, falls n gerade ist, so hängt es nicht vom gewählten Repräsentanten n ab, ob eine Äquivalenzklasse gerade ist, da alle vier Äquivalenzklassen jeweils nur gerade oder nur ungerade Zahlen enthalten.

Für die Äquivalenzklassen bezüglich des Rests bei Division durch 5 wäre eine entsprechende Definition nicht wohldefiniert, da dann zum Beispiel die ungerade Zahl 1 und die gerade Zahl 6 in derselben Äquivalenzklasse liegen.

1.3.2 Ordnungsrelationen

Nichtstrikte Ordnungsrelatione In diesem Abschnitt werden verschiedene Arten von Ordnungsrelationen eingeführt. Solche Ordnungsrelationen sind immer transitiv, sie können reflexiv sein oder auch nicht. Die dabei betrachteten nicht reflexiven Ordnungsrelationen sind sogar irreflexiv, das heißt, kein Element der betrachteten Menge steht zu sich selbst in Relation. Solche Ordnungsrelationen werden auch als strikte Ordnungsrelationen bezeichnet, die reflexiven als nichtstrikt. Wir behandeln zunächst reflexive Ordnungsrelationen und führen dann die irreflexiven Ordnungsrelationen auf diese zurück indem wir zu einer reflexiven Ordnungsrelation $\leq$ eine zugehörige irreflexive Ordnungsrelation $<$ defnieren, so dass $x < y$ genau dann gelten soll, wenn $x \leq y$, aber nicht $x = y$ gilt.

Reflexive Ordnungsrelationen werden meist mit dem Symbol $\leq$ bezeichnet, eventuell mit zusätzlichen unteren oder oberen Indizes, und in Infixnotation geschrieben. Bei den zugehörigen Äquivalenzklassen lassen wir der besseren Lesbarkeit wegen den untere Index $\leq$ weg, wenn die Relation $\leq$ aus dem Kontext klar ist.

30 Definition. *Eine* PARTIELLE PRÄORDNUNG *auf einer Menge A ist eine reflexive und transitive Relation auf A.*

Gemäß Lemma 21 bilden die Äquivalenzklassen einer partiellen Präordnung auf einer Mengen A eine Zerlegung von A. Eine symmetrische partielle Präoordnug ist eine Äquivalenzrelation, diese wird nach Bemerkung 26 durch diese Zerlegung induziert und insbesondere können Element in verschiedenen Äquivalenzklassen nicht zueinander in Relation stehen. Letzteres gilt nicht für eine partielle Präordnung $\leq$, die nicht symmetrisch ist, diese ist nach Satz 25 keine Äquivalenzrelation und wird deshalb insbesondere nicht durch die Zerlegung in Äquivalenklassen induziert. Letzteres gilt genau dann, wenn es x und y aus verschiedenen Äquivalenzklassen mit $x \leq y$ gibt, und genau in diesem Fall ist eine partielle Ordnung keine Äquivalenzrelation.

Die übliche Ordnung $\leq$ auf den natürlichen Zahlen ist eine partielle Ordnung, hat aber zwei Eigenschaften, die nicht für alle partiellen Ordnungen gelten:

(i) Alle Äquivalenzklassen enthalten genau ein Element.

(ii) Je zwei natürliche Zahlen sind vergleichbar, für alle natürlichen Zahlen x und y gilt $x \leq y$ oder $y \leq x$.

Diese beiden Eigenschaften sind für eine reflexive Relation äquivalent dazu, dass die Relation antisymmetrisch beziehungsweise konnex im Sinne der folgenden Definition ist.

31 Definition. *Es sei R eine zweistellige Relation R auf einer Menge A.*

(i) Die Relation R ist ANTISYMMETRISCH*, falls für alle x und y in A aus xRy und yRx immer $x = y$ folgt.*

(ii) Die Relation R ist KONNEX*, falls für alle x und y in A mit x ungleich y gilt xRy oder yRx. Dies ist äquivalent dazu, dass für alle x und y in A gilt $x = y$ oder xRy oder yRx.*

32 Bemerkung. *Eine reflexive Relation R ist genau dann konnex, wenn für alle x und y gilt xRy oder yRx. Die kompliziertere Bedingung in Definition 31 hat den Vorteil, dass der Begriff konnex damit auch auf irreflexive Relationen sinnvoll angewendet werden kann. Für eine solche Relation R gilt ja nie xRx, es lässt sich also nur im Fall x ungleich y fordern, dass xRy oder yRx gilt.*

33 Definition. *Eine* PARTIELLE ORDNUNG *auf einer Menge A ist eine reflexive, transitive und antisymmetrische Relation auf A.*

Eine PRÄORDNUNG *auf einer Menge A ist eine reflexive, transitive und konnexe Relation auf A.*

Eine ORDNUNG*, auch:* TOTALE ORDNUNG, LINEARE ORDNUNG*, auf einer Menge A ist eine reflexive, transitive, antisymmetrische und konnexe Relation auf A.*

In den Definitionen 30 und 33 wurden vier verschiedene Typen von reflexiven Ordnungsrelationen eingeführt. Für eine solche Ordnungsrelation wird dabei

(i) durch den Begriff Präordnung ausgedrückt, dass die Äquivalenzklassen der Relation mehr als ein Element enthalten können, die Relation also nicht notwendigerweise asymmetrisch ist,

(ii) durch das Attribut partiell ausgedrückt, dass nicht alle Elemente der betrachteten Menge vergleichbar sind, dass also die Relation nicht notwendigerweise konnex ist.

Zwischen den oben eingeführten vier Begriffen von Ordnungsrelationen bestehen die in Abbilddung 1.3 dargestellten Implikationen. Über diese Implikationen hinaus gelten keine weiteren, insbesondere sind alle dargestellten Implikationen echt, das heißt, die Implikation in die umgekehrte Richtung gilt jeweils nicht, und alle vier Begriffe sind paarweise verschieden.

Dass der Begriff partielle Ordnung von den anderen drei Begriffen verschieden ist, wird durch Äquivalenzrelationen bezeugt. Jede Äquivalenzrelation ist reflexiv und transitiv, ist also eine partielle Präordnung. Falls eine Äquivalenzrelation mindestens zwei Äquivalenzklassen hat, ist sie nicht konnex und somit auch keine Präordnung. Falls es Äquivalenzklassen mit mehr als einem Element gibt, ist die Äquivalenzordnung nicht antisymmetrisch und somit keine partielle Ordnung. Die drei folgenden Beispiele zeigen, dass auch die anderen drei Begriffe von Ordnungsrelationen paarweise verschieden sind.

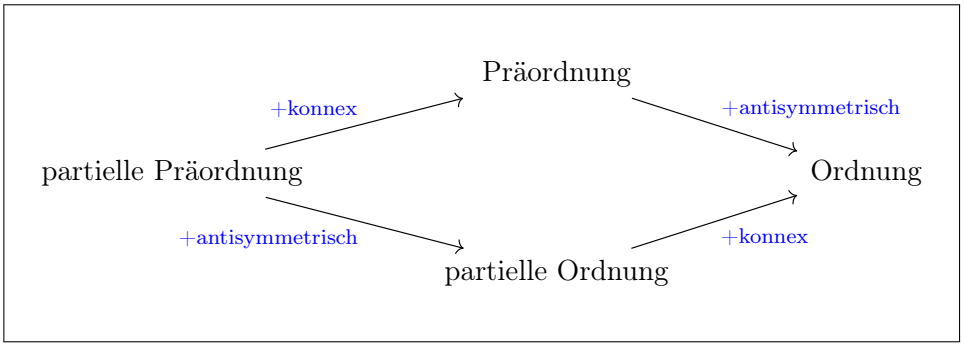

Abbildung 1.3: Vier Typen von Ordnungsrelationen

34 Beispiel. *Für jede Menge A ist die Inklusionsrelation $\subseteq$ eine partielle Ordnung auf der Potenzmenge von A, die aber, falls A mindestens zwei Elemente A, nicht konnex und somit auch keine Ordnung ist.*

Die Relation $\subseteq$ ist offensichtlich reflexiv und transitiv und auch antisymmetrisch, da aus $X \subseteq Y$ und $Y \subseteq X$ folgt $X = Y$. Falls A zwei verschiedene Elemente a und b hat, ist die Relation nicht konnex, da weder $\{a\} \subseteq \{b\}$ noch $\{b\} \subseteq \{a\}$ gilt.

Im Folgenden Beispiel setzen wir die Begriffe endliche Menge und Maximum einer endlichen Menge von natürlichen Zahlen in ihrer übliche Bedeutung als gegeben voraus.

35 Beispiel. *Zu einer nichtleeren endlichen Teilmenge X der natürlichen Zahlen sei deren Maximum $\max X$ gleich der größten Zahl in X. Die durch*

$$X \leq_{\max} Y \qquad g.d.w. \qquad \max X \leq \max Y$$

definierte Relation $\leq_{\max}$ ist dann eine Präordnung auf der Menge der nichtleeren endlichen Teilmengen der natürlichen Zahlen, die aber nicht antisymmetrisch und somit auch keine Ordnung ist.

Die Relation $\leq_{\max}$ ist offensichtlich reflexiv und transitiv und auch konnex, da zwei endliche Teilmengen der endlichen Zahlen immer vergleichbar sind. Die Relation ist nicht antisymmetrisch und somit keine Ordnung, die einzelnen Äquivalenzklassen bestehen gerade aus allen nichtleeren endlichen Teilmengen der natürlichen Zahlen mit demselben Maximum.

36 Beispiel. *Die übliche Kleiner-oder-gleich-Relation $\leq$ auf den natürlichen Zahlen ist eine Ordnung. Die Relation $\leq$ ist offensichtlich reflexiv und transitiv, aber auch antisymmetrisch, da aus $x \leq y$ und $y \leq x$ immer $x = y$ folgt, sowie konnex, da zwei natürlichhe Zahlen immer vergleichbar sind.*

37 Bemerkung. *Eine partielle Präordnung $\leq$ auf einer Menge A induziert auf ihren Äquivalenzklasse eine partielle Ordnung $\leq_{\mathrm{K}}$ gemäß der Definition*

$$[x] \leq_{\mathrm{K}} [y] \quad g.d.w. \quad x \leq y.$$

Die induzierte Relation ist wohldefiniert: da die Relation $\leq$ transitiv ist, gilt im Fall $x \leq y$ auch $x' \leq y'$ für alle x' in $[x]$ und y' in $[y]$. Die induzierte Relation ist eine partielle Ordnung: sie ist offensichtlich reflexiv und sie erbt, anschaulich gesprochen, die Transitivität von der Relation $\leq$. Die induzierte Relation ist auch antisymmetrisch, da aus $[x] \leq_{\mathrm{K}} [y]$ und $[y] \leq_{\mathrm{K}} [x]$ folgt, dass $x \leq y$ beziehungsweise $y \leq x$ gilt und somit die Äquivalenzklassen $[x]$ und $[y]$ identisch sind.

Eine beliebige zweistellige Relation kann dadurch reflexiv gemacht werden, dass alle Paare der Form (x, x) hinzugefügt werden, dies wird als Übergang zur reflexiven Hülle bezeichnet. Ganz ähnliche lassen sich auch Symmetrie und Transitivität erzwingen.

38 Definition. *Es sei R eine zweistellige Relation auf einer Menge A.*

Die REFLEXIVE HÜLLE *der Relation R ist die Relation $R \cup \{(x, x) \colon x \in A\}$.*

Die SYMMETRISCHE HÜLLE *von R ist die Relation $R \cup \{(y, x) \colon (x, y) \in R\}$.*

Die TRANSITIVE HÜLLE *von R ist die Relation*

$$\begin{aligned} R \cup \{(x, y) \colon & \text{ es gibt eine natürliche Zahl } t \text{ und } x_1, \ldots, x_t \text{ in } A \\ & \text{mit } x = x_1 \text{ und } y = x_t, \text{ so dass die Paare} \\ & (x_i, x_{i+1}) \text{ für } i = 1, \ldots, t-1 \text{ in } R \text{ sind}\}. \end{aligned}$$

In den folgenden Bemerkungen werden ohne Beweis einige Eigenschaften der reflexiven, der symmetrische und der transitiven Hülle zusammengestellt.

39 Bemerkung. *Die reflexive Hülle einer Relation ist eindeutig bestimmt. Die reflexive Hülle einer Relation ist reflexiv und die reflexive Hülle einer reflexiven Relation ist die Relation selbst.*

Die reflexive Hülle einer Relation R kann äquivalent dadurch definiert werden, dass sie die bezüglich der mengentheoretischen Inklusion kleinste Relation ist, welche reflexiv ist und R enthält. Eine derartige kleinste Relation existiert, sie ist gleich dem Schnitt über alle zweistelligen Relationen auf derselben Menge wie R die reflexiv sind und R enthalten.

Die Aussagen dieses Bemerkung gelten entsprechend auch für die symmetrische und für die transitive Hülle: alle Aussagen bleiben richtig, wenn überall reflexiv durch symmetrisch ersetzt wird, und ebenso, wenn überall reflexiv durch transitiv ersetzt wird.

40 Bemerkung. *Wird ausgehend von einer beliebigen Relation R zunächst die reflexive Hülle gebildet, dann aus dem Ergebnis die symmetrische und schließlich die transitive Hülle, so ergibt sich die reflexive, symmetrische und transitive Hülle der Relation R, wobei durch diese Bezeichnung auch die Reihenfolge der Hüllenoperationen ausgedrückt wird. Entsprechend für andere Kombinationen von Hüllenoperationen.*

Die Reihenfolge kann relevant sein, zum Beispiel ist für die Relation

$$R = \{(2i-1, 2i) \colon i > 0\} \cup \{(2i+1, 2i) \colon i \in \mathbb{N}\}$$

die symmetrische und transitive Hülle gleich $\mathbb{N} \times \mathbb{N}$, die transitive und symmetrische Hülle ist dagegen gleich $\{(i, i+1), (i+1, i) \colon i \in \mathbb{N}\}$.

Wird nacheinander die reflexive und die transitive Hülle gebildet, spielt die Reihenfolge keine Rolle, ebenso für die reflexive und die symmetrische Hülle.

41 Bemerkung. *Für jede Relation ist die reflexive und transitive Hülle eine partielle Präordnung, und die reflexive, symmetrische und transitive Hülle ist eine Äquivalenzrelation.*

Für einen gerichteten Graphen ist die reflexive und transitive Hülle der Kantenrelation gleich der Erreichbarkeitsrelation des Graphen. Die Äquivalenzklassen der reflexiven, symmetrische und transitive Hülle der Kantenrelation sind gerade die Zusammenhangskomponente des zugehörigen gerichteten Graphen, dieser ergibt sich, wenn alle Kanten als ungerichtet angesehen werden. Dies bedeutet, dass genau die Knoten in derselben Äquivalenzklasse liegen, die im Graphen durch einen ungerichteten Weg verbunden sind.

Strikte Ordnungsrelationen Als nächstes betrachten wir Ordnungsrelationen, bei denen die Elemente der geordneten Menge nie zu sich selbst in Relation stehen. Solche Ordnungen werden als STRIKTE ORDNUNGEN, auch: STRENGE ORDNUNGEN, bezeichnet und meist als $<$ geschrieben. Um den Unterschied zu solchen strengen Ordnungen hervorzuheben, wird für reflexive Ordnungen gelegentlich auch die Bezeichnung NICHTSTRIKTE ORDNUNG verwendet.

42 Definition. *Es sei R eine zweistellige Relation auf einer Menge A.*

Die Relation R ist IRREFLLEXIV*, falls xRx für kein x in A gilt.*

Die Relation R ist ASYMMETRISCH*, falls es keine Elemente x und y aus A gibt, für die sowohl xRy als auch yRx gilt.*

43 Bemerkung. *Nach Definition ist jede asymmetrische Relation R auch irreflexiv: falls es ein x mit xRx gibt, kann die Definition von aysmmetrisch nicht erfüllt sein.*

Ist R transitiv, gilt auch die umgekehrte Implikation: ist R irreflexiv, dann auch asymmetrisch. Wäre R nicht asymmetrisch, gäbe es x und y mit xRy

und yRx, und da R transitiv ist, würde auch xRx gelten, im Widerspruch zur Annahme, dass R irreflexiv ist.

44 Definition. *Eine* STRIKTE PARTIELLE ORDNUNG *auf einer Menge A ist eine irreflexive und transitive Relation auf A.*

Eine STRIKTE ORDNUNG, *auch:* STRIKTE TOTALE ORDNUNG ODER STRIKTE LINEARE ORDNUNG*auf einer Menge A ist eine irreflexive, transitive und konnexe Relation auf A.*

45 Beispiel. *Für jede Menge ist die strikte Inklusion $\subset$ eine strikte partielle Ordnung auf der Potenzmenge dieser Menge. Die übliche Echt-kleiner-gleich-Relation $<$ auf den natürlichen Zahlen ist eine strikte Ordnung.*

46 Bemerkung. *Es sei $<$ eine strikte Ordnung $<$ auf einer Menge A und x und y seien in A mit $x \neq y$. Dann gilt entweder $x < y$ oder $y < x$.*

Da die Relation $<$ konnex ist, gilt $x < y$ oder $y < x$, es können aber nicht beide Aussagen gelten, da $<$ nach Bemerkung 43 asymmetrisch ist.

47 Bemerkung. *Es lässt sich zeigen, dass für eine Ordnung $\leq$ die durch*

$$x < y \qquad g.d.w. \qquad (x \leq y \ \ und \ \ x \neq y)$$

defnierte Relation $<$ eine strikte Ordnung ist, diese wird als die der Ordnung $\leq$ ZUGEHÖRIGE STRIKTE ORDNUNG *bezeichnet. Umgekehrt wird für jede strikte Ordnung $<$ durch*

$$x \leq y \qquad g.d.w. \qquad (x < y \ \ oder \ \ x = y)$$

eine Ordnung $\leq$ definiert, die der strikten Ordnung $<$ ZUGEHÖRIGE ORDNUNG*. Dies lässt sich auch so ausdrücken, dass die reflexive Hülle einer strikten Ordnung eine Ordnung ist. Für strikte und nichtstrikte partielle Ordnungen gelten entsprechende Aussagen.*

1.4 Funktionen

Einstellige Funktionen Anschaulich gesprochen ordnet eine Funktion jedem Objekt genau ein anderes Objekte zu, zum Beispiel einer Person ihr Geburtsjahr oder einer natürlichen Zahl ihr Quadrat. Eine Funktion die jedem Element x aus einer Menge A ein Element y aus einer Menge B zuordnet kann als die Menge aller entsprechender Paare (x, y) dargestellt werden, diese ist eine Teilmenge des kartesischen Produkts $A \times B$.

48 Definition. *Für Mengen A und B ist eine* FUNKTION VON A NACH B, *auch:* FUNKTION VON A IN B, *eine Teilmenge f des kartesischen Produkts $A \times B$ mit den beiden folgenden Eigenschaften*

(i) zu jedem x in A enthält f ein Paar der Form (x, y),

(ii) zu jedem x in A enthält f höchstens ein Paar der Form (x, y),

Für eine solche Funktion f ist der FUNKTIONSWERT $f(x)$ *eines Elements x von A der eindeutig bestimmte Wert y mit (x, y) in f, die Funktion f* ORDNET DEM ARGUMENT x DEN FUNKTIONSWERT $f(x)$ *zu.*

Ein Funktionswert der Form $f(x)$ wird auch als BILD VON x UNTER f *bezeichnet, für jedes Paar (x, y) in f ist x ein* URBILD DES WERTS y UNTER f.

Die Menge A wird als DEFINITIONSMENGE, *kurz:* $\mathrm{def}(f)$, *und B als* ZIELMENGE, *kurz:* $\mathrm{ziel}(f)$, *von f bezeichnet.*

Anstelle von Funktion wird auch der Ausdruch ABBILDUNG verwendet, eine Funktion f BILDET DAS ARGUMENT x AUF $f(x)$ AB. Bisher hatten wir n-stellige Relationen nur in der Form von Teilmengen des n-fachen kartesischen Produkts A^n einer Menge A mit sich selbst eingeführt, analog wird auch eine Teilmenge von zum Beispiel $A \times B$ als zweistellige RELATION AUF A UND B bezeichnet. Eine Funktion von A nach B ist dann formal eine Relation auf A und B welche LINKSTOTAL ist, dies ist gerade Eigenschaft (i), und RECHTSEINDEUTIG, dies ist gerade Eigenschaft (ii).

Eine Funktion wird durch Angabe ihres Definitionsbereichs A und Zielbereichs B sowie der Funktion selbst als Teilmenge des kartesischen Produkts $A \times B$ festgelegt, die Schreibweise dafür ist:

$$\begin{aligned} f\colon A &\longrightarrow B, \\ x &\longmapsto f(x), \end{aligned}$$

wobei $f(x)$ auch durch einen definierenden Term ersetzt werden kann, zum Beispiel lässt sich die Quadratfunktion auf den natürlichen Zahlen wie folgt schreiben

$$\begin{aligned} f\colon \mathbb{N} &\longrightarrow \mathbb{N}, \\ x &\longmapsto x^2. \end{aligned}$$

Wenn Definitions- und Zielmenge klar sind, wird auch nur $f\colon x \mapsto f(x)$ geschrieben und beispielsweise von der Funktion $x \mapsto x^2$ gesprochen.

Bilder und Urbilder

49 Definition. *Es seien A und B nichtleere Mengen und f sei eine Funktion von A nach B. Das* BILD EINER TEILMENGE A_0 *von A unter f ist*

$$f(A_0) \;=\; \{f(x)\colon x \text{ in } A_0\},$$

das URBILD EINER TEILMENGE B_0 *von* B *unter* f *ist*

$$f^{-1}(B_0) \;=\; \{x \in A\colon f(x) \text{ in } B_0\},$$

Die BILDMENGE*, auch:* WERTEMENGE*, von* f *ist gleich dem Bild von* A*.*

In der Literatur finden sich auch andere Bezeichnungen: es wird Definitions- und Wertebereich anstelle von Definitions- und Wertemenge verwendet oder der Begriff Wertemenge und insbesondere dessen englische Form *range* bezeichnen nicht die Bild-, sondern die Zielmenge der betrachteten Funktion.

Hintereinanderausführung von Funktionen

50 Definition. *Es seien* f *eine Funktion von einer Menge* A *in eine Menge* B *und* g *eine Funktion von* B *in eine Menge* C*. Die* KOMPOSITION VON f UND g*, auch:* HINTEREINANDERAUSFÜHRUNG VON f UND g *oder* DIE FUNKTION g NACH f*, ist die Funktion*

$$\begin{aligned} g \circ f\colon A &\longrightarrow C, \\ x &\longmapsto g(f(x)). \end{aligned}$$

Die Komposition von mehr als zwei Funktionen wird entsprechend definiert, zum Beispiel gilt für Funktionen f, g und h mit passenden Definitions- und Zielmengen $h \circ g \circ f(x) = h(g(f(x)))$.

Injektive, surjektive und bijektive Funktionen

51 Definition. *Eine Funktion* f *ist* INJEKTIV*, wenn für alle* x *und* x' *in ihrer Definitionsmenge mit* $x \neq x'$ *gilt* $f(x) \neq f(x')$*.*

Eine Funktion ist SURJEKTIV*, falls ihre Bildmenge gleich der Zielmenge ist.*

Eine Funktion ist BIJEKTIV*, wenn sie injektiv und surjektiv ist.*

Eine INJEKTION *ist eine injektive Funktion, entsprechend für die Begriffe* SURJEKTION *und* BIJEKTION*.*

Eine Funktion ist eine Abbildung von ihrem Wertebereich in ihre Zielmenge. Die Sprechweise, dass eine Funktion eine Abbildung AUF ihren Zielbereich ist, bedeutet gerade, dass die Funktion surjektiv ist.

Für eine injektive Funktion f sind die Urbilder der Funktionswerte eindeutig. Somit kann jedes Element der Bildmenge von f für genau ein x in der Form $f(x)$ geschrieben werden, das heißt, $f(x) \mapsto x$ ist eine wohldefinierte Funktion von der Bildmenge von f in die Definitionsmenge von f.

52 Definition. *Es sei f sei eine injektive Funktion von A nach B. Dann ist die* Umkehrfunktion *von f gleich*

$$\begin{aligned} f^{-1}\colon f(A) &\longrightarrow A, \\ f(x) &\longmapsto x. \end{aligned}$$

53 Bemerkung. *Für jede Funktion f von einer Menge A in eine Menge B bezeichnet f^{-1} die* Urbildfunktion *von f, das heißt die Funktion von $\mathrm{Pot}(B)$ nach $\mathrm{Pot}(A)$, die einer Teilmenge von B ihr Urbild unter f zuordnet. Falls f injektiv ist, bezeichnet f^{-1} auch die Umkehrfunktion von f, das heißt die Funktion von $f(A)$ nach A, die jedem Element y von B, das in der Form $f(x)$ geschrieben werden kann, den Wert x zuordnet. Bei der Verwendung der Schreibweise f^{-1} muss immer klar sein, zum Beispiel aus dem Zusammenhang, welche der beiden Funktionen gemeint ist.*

54 Bemerkung. *Es sei f eine Funktion von A nach B. Dann gilt für alle Teilmengen A_0 von A und B_0 von B*

$$\text{(i)}\ \ A_0 \subseteq f^{-1}(f(A_0)) \qquad \textit{und} \qquad \text{(ii)}\ \ f(f^{-1}(B_0)) \subseteq B_0.$$

Diese Aussagen gelten sogar mit Gleichheit anstelle der Inklusionsrelation, falls f injektiv beziehungsweise surjektiv ist: falls f injektiv ist, gilt für alle Teilmengen A_0 von A die folgende Aussage (i′)*, und falls f surjektiv ist, gilt für alle Teilmengen B_0 von B die folgende Aussage* (ii′)

$$\text{(i}'\text{)}\ \ A_0 = f^{-1}(f(A_0)) \qquad \textit{und} \qquad \text{(ii}'\text{)}\ \ f(f^{-1}(B_0)) = B_0.$$

Falls f nicht injektiv ist, gibt es immer eine Teilmenge A_0 von A, für die (i′) *falsch ist. Falls f nicht surjektiv ist, gibt es immer eine Teilmenge B_0 von B, für die* (ii′) *falsch ist. Wir lassen die Beweise für die Aussagen dieser Bemerkung hier aus, diese werden in den Übungen behandelt.*

55 Bemerkung. *Für jede injektive Funktion f von A nach B ist*

$$\begin{aligned} f\colon A &\longrightarrow f(A), \\ x &\longmapsto f(x). \end{aligned}$$

eine Bijektion von A in die Bildmenge von f.

56 Bemerkung. *Eine Funktion von A nach B ist eine Menge von Paaren aus $A \times B$. Eine solche Funktion ist genau dann bijektiv, wenn unter den Paaren in der Funktion jedes x in A genau ein mal als erste Komponente und jedes y in B genau ein mal als zweite Komponente vorkommt. Dies kann anschaulich als* Matching *oder* Paarung *der Mengen A und B beschrieben werden, ein im Wesentlichen äquivalenter formaler Begriff von Matching wird in der Graphentheorie eingeführt.*

Wachsende und fallende Funktionen Ist auf der Definitions- und Zielmenge einer Funktion jeweils eine Ordnung definiert, liegt es nahe zu fragen, ob mit den Argumenten auch die Funktionswerte größer werden. Der Einfachheit halber führen wir die entsprechenden Monotoniebegriffe nur für den Fall einer Funktion von einer Menge in sich selbst ein. Für den Fall einer Funktion von einer Menge A in eine Menge B ändern sich die Definitionen nur insofern, als die Ordnungnen auf A und auf B zu unterscheiden sind.

57 Definition. *Auf einer Menge A sei eine Ordnung $\leq$ definiert, die zugehörige strikte Ordnung sei $<$. Es sei f eine Funktion von A nach A.*

Die Funktion f ist MONOTON WACHSEND*, wenn für alle x_1 und x_2 in A mit $x_1 < x_2$ gilt $f(x_1) \leq f(x_2)$.*

Die Funktion f ist STRENG MONOTON WACHSEND*, wenn für alle x_1 und x_2 in A mit $x_1 < x_2$ gilt $f(x_1) < f(x_2)$.*

Die Funktion f ist MONOTON FALLEND*, wenn für alle x_1 und x_2 in A mit $x_1 < x_2$ gilt $f(x_2) \leq f(x_1)$.*

Die Funktion f ist STRENG MONOTON FALLEND*, wenn für alle x_1 und x_2 in A mit $x_1 < x_2$ gilt $f(x_2) < f(x_1)$.*

Der Zusatz monoton kann auch weggelassen werden, und anstelle der Ausdrücke wachsend und fallend werden auch die Ausdrücke ZUNEHMEND beziehungsweise ABNEHMEND verwendet. Eine streng monoton wachsende Funktion wird somit auch als streng wachsend oder als streng zunehmend bezeichnet. Alle in Definition 57 eingeführte Begriffe könnnen verwendet werden, ohne die Relationen $\leq$ und $<$ aus der Definition explizit zu nennen, falls aus dem Zusammenhang klar ist, welche Relationen gemeint sind oder dies nicht relevant ist.

58 Beispiel. *Die Funktion, welche einer natürlichen Zahl n die kleinste Quadratzahl größer oder gleich n zuordnet, ist monoton wachsend. Die Funktion $n \mapsto n^2$ auf den natürlichen Zahlen ist streng monoton wachsend.*

Für jede Menge A mit einer Ordnung ist die Identitätsfunktion $x \mapsto x$ auf A streng monoton wachsend.

Die Funktion, welche einer natürlichen Zahl ihren Rest bei Division durch 5 *zuordnet ist weder wachsend noch fallend.*

59 Bemerkung. *Streng monoton wachsende Funktion sind immer injektiv, ebenso streng monoton fallende Funktionen.*

Undefinierte Funktionswerte Für eine Funktion, welche einer natürlichen Zahl ihren kleinsten Primteiler zuordnet, ist nicht klar, welcher Funktionswert dem Argument 1 zugeordnet werden soll. Diese Problem kann dadurch behoben werden, dass ein Wert festgelegt wird oder dadurch, dass die

Funktion an dieser Stelle UNDEFINIERT bleibt. Auch die Funktion $x \mapsto \frac{1}{x}$ von der Menge der reellen Zahlen $\mathbb{R}$ in sich selbst ist an der Stelle 0 undefiniert.

Die Stellen, an denen eine Funktion undefiniert ist, sind formal nicht Elemente des Definitionsbereichs der Funktion, in den Beispielen oben sind die Defnitionsbereiche also $\mathbb{N} \setminus \{1\}$ und $\mathbb{R} \setminus \{0\}$. In der Literatur wird der Begriff Definitionsbereich auch so verwendet, dass diese Funktionen als Funktionen mit Definitionsbereich $\mathbb{N}$ beziehungsweise $\mathbb{R}$ angesehen werden, die an einzelnen Stellen undefiniert sind.

1.5 Zahlbereiche

1.5.1 Die natürlichen und die ganzen Zahlen

Natürliche Zahlen In diesem Abschnitt werden die Mengen der natürlichen, der ganzen, der rationalen und der reellen Zahlen eingeführt. Als Beispiel für den mengentheoretischen Aufbau der Mathematik werden wir dabei skizzieren, wie die natürlichen und die reellen Zahlen formal definiert werden, eine ausführlichere Darstellung findet sich im bereits genannten Lehrbuch zur Mengenlehre von Ebbinghaus und in etwas kompaktere Form bei Friedrichsdorf und Prestel, *Mengenlehre für den Mathematiker*, Vieweg 1985. Später werden wir die verschiedenen Typen von Zahlen und die für diese Typen definierten Relationen und Funktionen in der üblichen Weise verwenden, ohne uns um die formalen Grundlagen zu kümmern. Beispielsweise werden wir die natürlichen Zahlen und die Ordnung sowie die Addition und die Multiplikation auf diesen formal definieren, im Weiteren aber nicht mit diesen formalen Definitionen arbeiten. Stattdessen werden wir bestimmte Eigenschaften der natürlichen Zahlen als gegeben voraussetzen, wobei diese Eigenschaften prinzipiell unter Rückgriff auf die formalen Definitionen bewiesen werden könnsen.

Die mengentheoretischen Axiome stellen sicher, dass es eine Menge gibt, welche genau die folgenden Mengen enthält

$$\emptyset, \qquad \{\emptyset\}, \qquad \{\emptyset, \{\emptyset\}\}, \qquad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \quad \dots \quad . \tag{1.2}$$

Formal ist diese Menge die kleinste Menge, welche die leere Menge enthält und für jedes Element z auch die Menge $z \cup \{z\}$. Im Rahmen des mengentheoretischen Aufbaus der Mathematik wird die Menge $\mathbb{N}$ der NATÜRLICHEN ZAHLEN gleich dieser Menge gesetzt. Die üblichen Bezeichnungen $0, 1, 2, \dots$ für die natürlichen Zahlen werden als Abkürzungen für die Elemente dieser Mengen eingeführt, zum Beispiel steht 0 für die leere Menge und 1 für die Menge $\{\emptyset\}$. Damit können wir die Mengen aus (1.2) auch in der Form

$$\emptyset, \qquad \{0\}, \qquad \{0, 1\}, \qquad \{0, 1, 2\}, \quad \dots$$

schreiben. Wenn wir unter Absehung der gerade eingeführten Formalisierung wie üblich über die natürlichen Zahlen reden, wird die natürliche Zahl n formal gleich der Menge $\{0, \dots, n-1\}$ gesetzt, und ist damit gleich einer Menge mit n Elementen.

Die NACHFOLGERFUNKTION s auf der Menge der natürlichen Zahlen wird wie folgt definiert

$$\mathrm{s}(z) \;=\; z \cup \{z\},$$

damit gilt dann

$$\mathrm{s}(0) = 1, \quad \mathrm{s}(\mathrm{s}(\emptyset)) = \mathrm{s}(1) = 2, \quad \mathrm{s}(\mathrm{s}(\mathrm{s}(\emptyset))) = \mathrm{s}(2) = 3, \quad \dots \ .$$

In der üblichen Sprechweise ist der Nachfolger $\mathrm{s}(n)$ einer natürlichen Zahl n gleich $n+1$ und n ist gleich dem Funktionswert der n-fach iterierten Nachfolgerfunktion bei Argument 0.

Die übliche Kleiner-gleich-Relation $\leq$ und Echt-kleiner-gleich-Relation $\leq$ auf den natürlichen Zahlen lässt sich in einfacher Weise formal definieren, es gelte

$$s \leq t \quad \text{g.d.w.} \quad s \subseteq t \qquad \text{undm} \qquad s < t \quad \text{g.d.w.} \quad s \subset t.$$

Peano-Axiome Für diese Formalisierung der natürlichen Zahlen lassen sich dann die folgenden Aussagen beweisen, die als PEANO-AXIOME bezeichnet werden.

P1 Die Nachfolgerfunktion ist injektiv: für alle m und n in $\mathbb{N}$ mit $m \neq n$ gilt $\mathrm{s}(m) \neq \mathrm{s}(n)$.

P2 Die Zahl 0 ist nicht im Bild der Nachfolgerfunktion: für alle n in $\mathbb{N}$ gilt $\mathrm{s}(n) \neq 0$.

P3 Für alle Teilmengen E von $\mathbb{N}$ gilt: wenn E die Zahl 0 enthält und mit einer Zahl n aus $\mathbb{N}$ immer auch schon deren Nachfolger $\mathrm{s}(n)$, dann ist E gleich $\mathbb{N}$.

Durch die Peano-Axiome wird eine Menge mit Nachfolgerfunktion bereits bis auf ISOMORPHIE festgelegt. Dies bedeutet anschaulich gesprochen, dass jede Menge X mit einer Nachfolgerfunktion s_X und einem Element 0_X, welche die Peano-Axiome erfüllt, die gleiche Struktur oder Gestalt hat wie die natürlichen Zahlen: das Element 0_X ist in dem Sinne das erste oder kleinste Element ist, dass es kein Nachfolger eines anderen Elements ist, alle Elemente ergeben sich als iterierte Nachfolger des Elements 0_X, und die reflexive und transitive Hülle der Nachfolgerrelation ist eine lineare Ordnung auf X. Insbesondere kann sich die Folge der iterierten Nachfolger nicht zu einem Kreis schließen und enthält auch keine Elemente, an der sich zwei oder mehr Abfolgen von iterierten Nachfolgern treffen.

Das dritte Peano-Axiom gilt für $\mathbb{N}$ auch in folgender Form: jede mengentheoretische Eigenschaft, und somit insbesondere jede mathematische Eigenschaft, welche die 0 enthält und mit einer Zahl auch deren Nachfolger, gilt für alle Elemente von $\mathbb{N}$. Tatsächlich ist diese Form äquivalent zum dritten Peano-Axiom, da sich für jede solche Eigenschaft die zugehörige Teilmenge von $\mathbb{N}$ aus der Menge $\mathbb{N}$ aussondern lässt und andererseits in einer bestimmten Menge zu sein eine mengentheoretisch Eigenschaft ist.

Aus dem dritten Peano-Axiom folgt, dass INDUKTIONSBEWEISE der folgenden Form möglich sind: um zu zeigen, dass alle Elemente von $\mathbb{N}$, also alle natürlichen Zahlen, eine bestimmte Eigenschaft haben, genügt es nachzuweisen, dass die Eigenschaft auf 0 zutrifft und, falls sie auf eine natürliche Zahl n

zutrifft, dann auch auf deren Nachfolger $n + 1$. Weiter können Funktionen auf $\mathbb{N}$ durch eine INDUKTIVE DEFINITION definiert werden: eine Funktion ist auf ganz $\mathbb{N}$ definiert, falls die Funktion auf 0 definiert ist und immer, wenn sie auf einer Zahl n definiert ist, dann auch auf deren Nachfolger $n+1$. Falls dabei die Definition der Funktion an einer Stelle $n+1$ jeweils nur vom Funktionswert an der Stelle n abhängt, ist die Funktion auf ganz $\mathbb{N}$ eindeutig definiert. Als Beispiel betrachten wir die induktive Definition der Addition und Multiplikation, Induktionsbeweise werden später im Abschnitt über Beweistechniken behandelt.

Induktive Definition von Addition und Multiplikation Die Addition lässt sich auf den natürliche Zahlen induktiv definieren, indem für ein gegebenes n in $\mathbb{N}$ die Werte von $n + i$ sukzessive für $i = 0, 1, 2, \ldots$ festgelegt werden:

$$\begin{aligned} &\text{Induktionsanfang } i = 0: && n + 0 = n, \\ &\text{Induktionsschritt } i \curvearrowright i + 1: && n + \mathrm{s}(i) = \mathrm{s}(n + i). \end{aligned}$$

Unter Verwendung der Addition lässt sich ganz ähnlich auch die Multiplikation induktiv definieren:

$$\begin{aligned} &\text{Induktionsanfang } i = 0: && n \cdot 0 = 0, \\ &\text{Induktiontsschritt } i \curvearrowright i + 1: && n \cdot \mathrm{s}(i) = (n \cdot i) + n. \end{aligned}$$

Ganze Zahlen Unter Verwendung der natürlichen Zahlen lässt sich die Menge $\mathbb{Z}$ DER GANZEN ZAHLEN auf verschiedene Weise einführen. Zum Beispiel kann eine ganze Zahl mit einem Paar von natürlichen Zahlen gleichgesetzt werden, so dass ein Paar (a, b) für die Differenz $a - b$ steht. Alternativ kann eine ganze Zahl als Paar aus einem geeignet kodierten Vorzeichen und einer natürlichen Zahl dargestellt werden. In beiden Fällen lassen sich Relationen wie die strikte Ordnung $<$ und Operationen wie die Addition und Multiplikation von $\mathbb{N}$ auf $\mathbb{Z}$ übertragen, beispielsweise gilt $a_0 - b_0 < a_1 - b_1$ genau dann, wenn $a_0 + b_1 < a_1 + b_0$ gilt. Wir verzichten auf die Ausarbeitung dieser Konstruktionen und setzen im Folgenden das Verständnis der Menge $\mathbb{N}$ der natürlichen Zahlen und der Menge $\mathbb{Z}$ der ganzen Zahlen zusammen mit den üblichen Relationen und Operationen auf diesen Mengen voraus.

1.5.2 Die rationalen und die reellen Zahlen

Um Gleichungen wie $5x = 3$ lösen zu können, werden die ganzen Zahlen zur Menge $\mathbb{Q}$ der RATIONALEN ZAHLEN erweitert, diese enthalten anschaulich

gesprochen auch die Werte aller BRÜCHE, also alle Zahlen, die in der Form $\frac{a}{b}$, auch: a/b, $a : b$, geschrieben werden können, so dass der ZÄHLER a und der Nenner b ganze Zahlen sind und b ungleich 0 ist. Entsprechend definieren wir die Menge der RATIONALEN ZAHLEN als

$$\mathbb{Q} = \{(a, b) \colon a,\ b \text{ in } \mathbb{Z} \text{ und } b \text{ ungleich } 0\}.$$

Dabei stellen zum Beispiel die Paare $(3, 5)$ und $(-6, -10)$ zwei Brüche mit demselben Wert dar. Dies kann durch die wie folgt defnierte Äquivalenzrelation $\sim$ formalisiert werden

$$(a, b) \sim (c, d) \quad \text{g.d.w.} \quad \frac{a}{b} = \frac{c}{d} \quad \text{g.d.w.} \quad ad = cb.$$

Für eine Äquivalenzklasse der Relation $\sim$, welche das Paar $(0, 1)$, welches die 0 repräsentiert, nicht enthält, bietet sich als Repräsentant das eindeutig bestimmte Paar (a, b) in der Äquivalenzklasse an, so dass b eine natürliche Zahl ist und a und b TEILERFREMD sind, das heißt, keine gemeinsamen Teiler ungleich 1 und -1 haben. Dass a und b teilerfremd sind, wird auch so formuliert, dass der Bruch $\frac{a}{b}$ in GEKÜRZTER FORM vorliegt.

Relationen wie die strikte Ordnung $<$ und Operationen wie die Addition und Multiplikation auf den ganzen Zahlen lassen sich auf die rationalen Zahlen erweitern, zum Beispiel durch die Definitionen

$$(a, b) < (c, d) \ \text{ g.d.w. } \ ad < cb \qquad \text{und} \qquad (a, b) + (c, d) = (ad + cb, bd),$$

wobei die Definition der Relation $<$ für den Fall angegeben ist, dass die beteiligten Nenner b und d nichtnegativ sind.

Die rationalen Zahlen sind DICHT, für je zwei rationale Zahlen (a, b) und (c, d) mit $(a, b) < (c, d)$ gibt es eine rationale Zahl (e, f), die ECHT ZWISCHEN diesen beiden liegt, für die also $(a, b) < (e, f) < (c, d)$ gilt. Eine solche Zahl ergibt sich zum Beispiel, indem zur kleineren Zahl (a, b) die Hälfte der Differenz $(c, d) - (a, b)$ addiert wird.

Es gibt rationale Zahlen r, so dass sich r^2 beliebig wenig von 2 unterscheidet, aber die Gleichung $r^2 = 2$ besitzt keine Lösung in der Menge der rationalen Zahle. In der üblichen Sprechweise lässt sich dies so formulieren: die Wurzel aus 2 oder kurz $\sqrt{2}$, das heißt, die nichtnegative reelle Zahl, deren Quadrat gleich 2 ist, lässt sich durch rationale Zahlen beliebig gut annähern oder approximieren, aber $\sqrt{2}$ selbst ist nicht rational.

60 Satz. *Die Zahl $\sqrt{2}$ ist nicht rational.*

Beweis. Zunächst halten wir fest, dass das Quadrat einer natürlichen Zahl genau dann gerade ist, wenn die Zahl selbst gerade ist: für eine gerade Zahl der Form $2n$ ist das Quadrat gleich $4n^2$, also gerade, für eine ungerade Zahl der Form $2n + 1$ ist das Quadrat gleich $4n^2 + 4n + 1$ und somit ungerade.

Wir führen einen Beweis durch Widerspruch und nehmen an, dass es eine rationale Zahl r mit $2 = r^2$ gibt. Wir können r als positiv annehmen, es gibt also teilerfremde natürliche Zahlen p und q, so dass sich r in gekürzter Form als $r = p/q$ schreiben lässt. Es folgt

$$2 = \left(\frac{p}{q}\right)^2 = \frac{p^2}{q^2} \quad \text{und somit} \quad 2q^2 = p^2.$$

Nach der Vorüberlegung muss p gerade sein, etwa von der Form $p = 2n$ und wir erhalten

$$2q^2 = 4n^2 \quad \text{und somit} \quad q^2 = 2n^2,$$

somit ist auch q gerade. Im Widerspruch zur Wahl der Zahlen p und q haben diese also den gemeisamen Teiler 2. □

Nichtrationale reelle Zahlen wie die Wurzel aus 2 lassen sich beliebig gut durch rationale Zahlen annähern. Eine reelle Zahl kann entsprechend durch das Paar aus der Menge aller echt kleineren rationalen Zahlen und der Menge aller echt größeren rationalen Zahlen beschrieben werden. Da die Abbildung $x \mapsto x^2$ auf den positiven rationalen Zahlen streng monoton wachsend ist, lässt sich somit Wurzel 2 durch die beiden folgenden Mengen darstellen

$$\{x \in \mathbb{Q} \colon x^2 < 2\} \qquad \text{und} \qquad \{x \in \mathbb{Q} \colon 2 \leq x^2\}.$$

Ein solches Paar wird als Dedekindscher Schnitt bezeichnet, seine Komponenten als Untermenge und Obermenge des Schnitts. Formal ist ein Dedekindscher Schnitt definiert als ein Paar (A, B) von Mengen für das gilt

(i) die Mengen A und B bilden eine Zerlegung von $\mathbb{Q}$: die beiden Mengen sind disjunkt und ihre Vereinigung ist gleich $\mathbb{Q}$,

(ii) jede Zahl in A ist kleiner als jede Zahl in B,

(iii) die Menge A besitzt kein größtes Element: für alle x in A gibt es ein x' in A mit $x < x'$.

Die Menge $\mathbb{R}$ der reellen Zahlen wird dann gleich der Menge der Dedekindschen Schnitte gesetzt Entsprechend definieren wir die Menge der rationalen Zahlen als

$$\mathbb{R} = \{(A, B) \colon (A, B) \text{ ist Dedekindscher Schnitt}\}.$$

Werden die reellen Zahlen als bekannt vorausgesetzt, so bedeutet dies anschaulich, dass eine reelle Zahl r durch den Dedekindschen Schnitt mit Untermenge $\{x \in \mathbb{Q} \colon x < r\}$ und Obermenge $\{x \in \mathbb{Q} \colon r \leq x\}$ dargestellt wird.

Beide Mengen enthalten rationale Zahlen die sich von r beliebig wenig unterscheiden, für rationales r ist r das kleinste Element der Obermenge, falls r nicht rational ist, ist r weder in der Unter- noch in der Obermenge enthalten.

Die übliche Echt-Kleiner-Relation $<$ auf den reellen Zahlen lässt sich wie folgt auf deren Darstellungen durch Dedekindsche Schitte definieren

$$(A, B) \ < \ (A', B') \qquad \text{g.d.w.} \qquad A \subset A'.$$

Da Unter- und Obermenge eines Dedekindschen Schnitte eine Zerlegung von $\mathbb{Q}$ bilden, folgt aus $A \subset A'$, dass auch $B' \subset B$ gilt, letzteres muss also in dieser Definition nicht zusätzlich gefordert werden. Allgemein ist die Obermenge eines Dedekindschen Schnitts bereits durch die Untermenge festgelegt, so dass reelle Zahlen auch nur durch die Untermengen des jeweiligen Dedekindschen Schnitts dargestellt werden könnten.

Mit der Darstellung durch Dedekindsche Schnitte lassen sich Addition und Multiplikation von reellen Zahlen in einfacher Weise auf die Addition und Multiplikation rationaler Zahlen zurückführen. Beispielsweise wird die Addition durch

$$(A, B) \ + \ (C, D) \ = \ (A + C, B + D),$$

definiert, wobei $A + C$ für die Menge $\{s + t \colon s \text{ in } A \text{ und } t \text{ in } C\}$ steht und entsprechend für die Menge $B + D$.

61 Bemerkung. *Ein* ABGESCHLOSSENES INTERVALL *der reellen Zahlen ist eine Menge der Form* $\{x \in \mathbb{R} \colon a \leq x \leq b\}$ *für reelle Zahlen* a *und* b *mit* $a < b$*, ein solches Intervall wird als* $[a, b]$ *geschrieben und hat* LÄNGE $b - a$*. Eine* INTERVALLSCHACHTELUNG *ist eine Folge* $I_1, I_2, \ldots$ *von abgeschlossenen Intervallen der Form* $I_j = [a_i, b_i]$ *für reelle Zahlen* a_i *und* b_i *für die gilt*

$$a_1 \leq a_2 \leq \ldots, \qquad b_1 \geq b_2 \geq \ldots \qquad \textit{und} \qquad a_i < b_j \textit{ für alle } i \textit{ und } j$$

gelten und weiter die Intervalllängen $b_i - a_i$ GEGEN NULL GEHEN*, das heißt, für jede natürliche Zahl* k *sind fast alle Intervalllängen kleiner als* $\frac{1}{k}$*. Insbesondere gilt also* $I_{j+1} \subseteq I_j$ *für* $j = 1, 2, \ldots$*.*

Zu jeder Intervallschachtelung $I_1, I_2, \ldots$ *gibt es genau eine reelle Zahl, die in allen Intervallen* I_j *und damit im Durchschnitt* $\cap_{j=1,2,\ldots} I_j$ *über alle Intervalle liegt. Gäbe es zwei solche Zahlen* x *und* $x' < x$*, so wären beide in allen Intervallen enthalten, diese hätten somit alle mindestens Länge* $x - x' > 0$*. Die Menge der rationalen Zahlen ist nun in dem Sinne* UNVOLLSTÄNDIG*, dass es Intervallschachtelungen wie oben gibt, bei denen die Intervallendpunkte* $a_1, a_2, \ldots$ *und* $b_1, b_2, \ldots$ *alle rational sind, der Durchschnitt über die Intervalle aber nicht. Im Gegensatz dazu sind nach dem oben Gesagten die reellen Zahlen in diesem Sinne* VOLLSTÄNDIG*: der Durchschnitt über die Intervalle einer Intervallschachtelung enthält immer eine reelle Zahl.*

Falls für eine Intervallschachtelung wie oben die Intervallendpunkte alle rational sind, wird die reelle Zahl r im Schnitt der Intervalle durch den Dedekindschen Schnittt (A, B) mit

$$A = \{q \in \mathbb{Q} \colon \textit{es gibt einen Index } i \textit{ mit } a_i < q\},$$
$$B = \{q \in \mathbb{Q} \colon \textit{es gibt einen Index } i \textit{ mit } q < b_i\}.$$

dargestellt, wobei zu B noch r hinzugfügt werden muss, falls r rational ist.

1.6 Endliche und unendliche Mengen

Vergleich von Mengen der Größe nach Wollen wir zwei Mengen A und B, anschaulich gesprochen, hinsichtlich ihrer Größe vergleichen, so lässt sich wie folgt argumentieren. Gibt es eine bijektive Abbildung f zwischen A und B, so gibt es eine Paarung der Elemente von A und B bei der sich jeweils ein x aus A und dessen Funktionswert $f(x)$ aus B entsprechen, die Mengen sind sozusagen wechselseitig Kopien voneinander, in diesem Sinne haben die beiden Mengen dieselbe Größe. Gibt es eine injektive Abbildung f von A nach B, so gibt es zu jedem Element x von A ein Element $f(x)$ von B, wobei verschiedenen Elementen von A auch verschiedene Elemente von B zugeordnet werden, die Menge B enthält sozusagen eine Kopie von A und ist damit mindestens so groß wie A.

62 Definition. *Es seien A und B zwei Mengen. Die Mengen A und B sind* GLEICHMÄCHTIG*, kurz: $A \sim B$, falls es eine bijektive Funktion von A nach B gibt. Die Menge A ist* HÖCHSTENS SO MÄCHTIG *wie die Menge B, kurz: $A \preceq B$, falls es eine injektive Funktion von A nach B gibt. Die Menge B ist* MÄCHTIGER *als A, kurz: $A \prec B$, falls $A \preceq B$ gilt, aber nicht $B \sim A$.*

Für zwei gleichmächtige Mengen A und B gilt $A \preceq B$ und $B \preceq A$. Für eine Bijektion $f\colon A \to B$ ist ja f selbst eine injektive Abbildung von A nach B und die Umkehrabbildung f^{-1} von f ist eine injektive Abbildung von B nach A. Tatsächlich gilt auch die Umkehrung: falls es injektive Funktionen von A nach B und von B nach A gibt, dann auch eine bijektive Funktion von A nach B.

63 Äquivalenzsatz von Bernstein, Cantor und Schröder. *Zwei Mengen A und B sind genau dann gleichmächtig, wenn A höchstens so mächtig ist wie B und umgekehrt, das heißt, es gilt*

$$A \sim B \qquad \text{genau dann, wenn} \qquad A \preceq B \quad \text{und} \quad B \preceq A.$$

Der Beweis dieses Satzes ist zu verwickelt, um hier dargestellt zu werden, er findet sich in Kapitel 9 des angegebenen Lehrbuchs von Ebbinghaus.

64 Bemerkung. *Es seien A und B Mengen. Aus dem Äquivalenzsatz folgt unmittelbar, dass folgende Äwuivalenzen gelten*

$$A \prec B \quad \text{genau dann, wenn} \quad A \preceq B \text{ und nicht } A \sim B,$$
$$A \preceq B \quad \text{genau dann, wenn} \quad A \prec B \quad \text{oder} \quad A \sim B.$$

65 Bemerkung. *Die durch $\prec$, $\preceq$ und $\sim$ bezeichneten Beziehungen zwischen Mengen werden in der mathematischen Umgangssprache als Relationen bezeichnet, sie sind allerdings keine Relationen im formalen Sinn einer Menge von Paaren. Entsprechend werden wir zum Beispiel die Relation $\sim$ nicht als*

Äquivalenzrelation ansehen, obwohl sie die für eine Äquivalenzrelation geforderten Eigenschaften hat.

Wäre zum Beispiel $\sim$ als formal Relation darstellbar, müsste diese auch alle Paare der Form (x, x) für beliebige Mengen x enthalten. Aus den mengentheoretischen Axiomen würde damit folgen, dass es eine Menge gibt, die alle Mengen enthält, und mit Aussonderung würde sich wie in der Russelschen Antinomie als Widerspruch ergeben, dass es eine Menge gibt, die alle Mengen enthält, die sich nicht selbst enthalten.

66 Satz. *Die Relation $\sim$ hat die definierenden Eigenschaften einer Äquivalenzrelation, das heißt, sie ist reflexiv, symmetrisch und transitiv.*

Die Relationen $\prec$ und $\preceq$ sind beide transitiv, $\prec$ ist irreflexiv und $\preceq$ ist reflexiv. Unter Verwendung des Auswahlaxioms lässt sich zeigen, dass beide Relationen in dem Sinne konnex sind, dass je zwei verschiedene Mengen immer vergleichbar sind. Beide Relationen besitzen somit die definierenden Eigenschaften einer strikten beziehungsweise nichtstrikten Präordnung.

Endliche Mengen Die natürlichen Zahlen hatten wir formal als Mengen $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, ... eingeführt, die natürliche Zahl i entspricht dabei einer Menge mit i Elementen. Auf den so formalisierten natürlichen Zahlen lassen sich dann in einfacher Weise die übliche strikte und nichtstrikte Ordnung $<$ beziehungsweise $\leq$ definieren: für zwei natürliche Zahlen i und j gelte $i < j$, wenn i eine echte Teilmenge von j ist, und es gelte $i \leq j$, wenn i Teilmenge von j ist.

67 Definition. *Eine Menge ist* ENDLICH*, wenn sie gleichmächtig zu einer natürlichen Zahl ist. Eine endliche Menge die gleichmächtig zur natürlichen Zahl i ist, hat die* GRÖSSE *i, wir schreiben $|\mathrm{A}|$ für die Größe der endlichen Menge A.*

Die Größe einer Menge endlichen Menge A ist eindeutig. Für zwei verschiedene natürliche Zahlen i und j, wobei $i < j$ gelten soll, kann es keine Bijektionen $f_i \colon A \to i$ und $f_j \colon A \to j$ geben. Andernfalls wäre $f_i \circ f_j^{-1}$ eine bijektive Funktion von j in die echte Teilmenge i von j, ein Widerspruch zu Lemma 68, nach dem es keine injektive Funktion von einer endlichen Menge A in eine echte Teilmenge von A gibt. Diese Aussage ist anschaulich klar: das Bild eine injektiven Funktion auf einer endlichen Menge A kann nicht echt weniger Elemente haben als A. Als Beispiel für den mengentheoretischen Aufbau der Mathematik beweisen wir diese Aussage und einige ihrer Folgerungen ausgehend von der Formalisierung der natürlichen Zahlen als Mengen und der formalen Definitionen der Begriffe endliche Menge und Größe einer endlichen Menge.[2]

[2]Dass die Größe einer endlichen Menge eindeutig ist, folgt aus Lemma 68. Im Beweis

68 Lemma. *Für eine endliche Menge gibt es keine injektive Funktion von der Menge in eine ihrer echten Teilmengen.*

Beweis. Im Folgenden schreiben wir $\{0, 1, \ldots, n-1\}$ für die als Menge aufgefasste natürliche Zahl n. Dies ist gerechtfertigt, da gemäß der formalen Definition der natürllichen Zahlen als Mengen die natürliche Zahl n tatsächlich genau die den Zahlen 0 bis $n-1$ entsprechenden Mengen enthält.

Wir zeigen das Lemma für alle endlichen Mengen A durch Induktion über die Größe n von A. Im Induktionsanfang ist $n = 0$, das heißt, die Menge A ist leer und hat somit keine echten Teilmengen, folglich ist die Aussage des Lemmas für A richtig.

Im Induktionsschritt nehmen wir für eine beliebige natürliche Zahl n an, dass die Aussage des Lemmas für alle endlichen Mengen der Größe n gilt und zeigen, dass die Aussage auch für alle Mengen der Größe$n + 1$ gilt. Sei also A eine beliebige Menge der Größe $n + 1$. Für einen Beweis durch Widerspruch nehmen wir an, dass es eine injektive Funktion f von A in eine echte Teilmenge von A gibt. Da A die Größe $n+1$ hat, gibt es eine Bijektion

$$f_A \colon A \to n + 1 = \{0, 1, \ldots, n\}.$$

Sei a das Element von A mit $f(a) = n$ und sei $A_0 = A \setminus \{a\}$. Dann ist die Einschränkung von f_A auf A_0 eine Bijektion von A_0 auf $n-1$, die Menge A_0 hat also Größe $n - 1$.

Wir zeigen, zunächst, dass wir die Funktion f so abändern können, dass nach der Änderung $f(a) = a$ gilt und f immer noch eine injektive Funktion f von A in eine echte Teilmenge von A ist. Wir unterscheiden bei der Änderung drei Fälle. Sei dazu $f(a) = y$. Falls $y = a$ gilt, sind wir fertig. Falls a nicht im Bild von a ist, setzen wir $f(a) = a$, und erhalten eine injektive Funktion von A nach $A \setminus \{y\}$. Falls es ein x ungleich a mit $f(x) = a$ gibt, vertauschen wir die Funktionswerte von x und a, wir setzen $f(x) = y$ und $f(a) = a$. Auch in diesem Fall bleibt f injektiv, und da sich das Bild von f nicht ändert, bleibt dieses eine echte Teilmenge von A.

Im Folgenden steht f für die so geänderte Funktion, diese ist immer noch eine injektive Funktion von A in eine echte Teilmenge von A, es existiert somit ein z in A, dass nicht im Bild von f ist. Wegen $f(a) = a$ ist z ungleich a, somit ist z in A_0. Wir betrachten die Einschränkung von f auf die Menge A_0, das ist die Funktion mit Definitionsbereich A_0, die dort mit f übereinstimmt. Diese Einschränkung ist wieder injektiv und eine Funktion von A_0 in die echte Teilmenge $A_0 \setminus \{z\}$ von A_0, im Widerspruch zur Induktionsannahme

des Lemmas darf daher die Eindeutigkeit der Größe noch nicht vorausgesetzt werden. Der Begriff Größe wird dort nur so benutzt, dass eine Menge A eben Größe i hat, wenn es eine Bijektion von A in i gibt, ob es zusätzlich noch Bijektionen von A in naturliche Zahlen j ungleich i gibt, spielt keine Rolle.

für Mengen der Größe n. Die Widerspruchsannahme, dass es eine injektive Funktion von A in eine echte Teilmenge von A gibt, muss also falsch sein. Dies beendet den Induktionsschritt. □

69 Satz. *Es sei A eine endliche Mengen und f sei eine Funktion von A nach A. Dann sind die folgenden Aussagen äquivalent.*

(i) Die Funktion f ist injektiv.

(ii) Die Funktion f ist surjektiv.

(iii) Die Funktion f ist bijektiv.

Beweis. Falls f injektiv ist, kann das Bild der endlichen Menge A nach Lemma 68 keine echte Teilmenge von A sein, folglich ist f surjektiv

Sei nun f surjektiv. Zu y in A sei $U_y = f^{-1}(\{y\})$ die Menge der Urbilder von y. Da f surjektiv ist, sind alle Urbilder U_y nichtleer. Es gibt somit eine Funktion, die jedem y in A ein Element x_y in U_y zuordnet. Diese Funktion ist injektiv, da nach Definition von Urbild für $y \neq y'$ die Urbilder U_y und $U_{y'}$ disjunkt und folglich die Funktionswerte x_y und $x_{y'}$ verschieden sind. Die Funktion ist somit nach Lemma 68 surjektiv. Dies ist aber nur möglich, wenn alle Urbilder der Form U_y nur ein Element enthalten, was äquivalent dazu ist, dass f injektiv ist.

Die Aussagen (i) und (ii) sind also äquivalent. Ist f injektiv, dann auch surjektiv und umgekehrt, aus jeder der ersten beiden Aussagen folgt also, dass f bijektiv ist. Umgekehrt ist jede bijektive Funktion nach Definition injektiv und surjektiv. □

Als weitere Folgerung aus Lemma 68 erhalten wir das Schubfachprinzip, englisch: pigeonhole principle. Dieses besagt, dass bei einer Abbildung von einer endlichen Menge A in eine echt kleinere endliche Menge B zwei Elemente von A auf dasselbe Element von B abgebildet werden.

70 Schubfachprinzip. *Es seien A und B endliche Mengen mit $|A| < |B|$. Dann gibt es keine injektive Abbildung von B nach A.*

Beweis. Die Größen $|A|$ und $|B|$ sind natürliche Zahlen und damit formal endliche Mengen. Nach Definition der Relation $<$ auf den natürlichen Zahlen gilt $|A| < |B|$ genau dann, wenn $|A|$ eine echte Teilmenge von $|B|$ ist. Nach Lemma 68 kann es also keine injektive Abbildung von $|B|$ nach $|A|$ geben. Wir zeigen, dass es auch keine injektive Abbildung von B nach A geben kann. Für einen Beweis durch Widerspruch nehmem wir an, dass f eine solche Abbildung sei und wählen Bijektionen $f_A \colon A \to |A|$ und $f_B \colon B \to |B|$. Dann ist die Abbildung $f_A \circ f \circ f_B^{-1}$ als Komposition injektiver Abbildungen eine

injektive Abbildung von $|B|$ nach $|A|$, im Widerspruch zur gerade bewiesenen Aussage. □

71 Bemerkung. *Endliche Mengen können einerseits hinsichtlicher ihrer Mächtigkeit verglichen werden, also hinsichtlich der Relationen $\prec$, $\preceq$ und $\sim$, andererseits hinsichtlich der auf den natürlichen Zahlen, und damit auf den Größen von endlichen Mengen definierten Relationen $<$, $\leq$ und $=$.*

Beide Arten zu vergleichen sind äquivalent: es lässt sich zeigen, dass für alle endlichen Mengen A und B die folgenden Äquivalenzen gelten

(i) $A \prec B$ *genau dann, wenn* $|A| < |B|$,

(ii) $A \preceq B$ *genau dann, wenn* $|A| \leq |B|$,

(iii) $A \sim B$ *genau dann, wenn* $|A| = |B|$.

Wir verzichten auf den Beweis dieser Aussagen. Ähnlich wie der Beweis des Schubfachprinzips beruht dieser darauf, dass es nach Definition der Größe einer endlichen Menge Bijektionen von A nach $|A|$ und von B nach $|B|$ gibt, die Mengen A also in diesem Sinn eine Kopie der Mengen $|A|$ darstellt, und ebenso für B und $|B|$.

Unendliche Mengen Der Größenbegriff für endliche Mengen kann auf unendliche Mengen erweitert werden, die Größe einer Menge wird dann als KARDINALITÄT oder MÄCHTIGKEIT der Menge bezeichnet und formal gleich einer speziellen gleichmächtigen Menge gesetzt. Entsprechend wird auch die Größe von endlichen Mengen gelegentlich als deren Kardinalität bezeichnet.

Für die unendliche Menge $\mathbb{N}$ ist $n \mapsto n + 1$ eine injektive Funktion, deren Bild $\mathbb{N}$ eine echte Teilmenge von $\mathbb{N}$ ist.

72 Lemma. *Für jede unendliche Menge gibt es eine injektive Funktion von der Menge in eine ihrer echten Teilmengen.*

Beweisskizze. Sei A eine unendliche Menge. Es genügt zu zeigen, dass A eine Teilmenge A_0 der Form $\{x_0, x_1, \ldots\}$ enthält, so dass die x_i paarweise verschiedenen sind. Damit erhalten wir dann eine injektive Funktion von A auf eine echte Teilmenge von A, indem wir die Funktion auf $A \setminus A_0$ gleich der Identität und auf A_0 das Bild von x_i gleich x_{i+1} setzen, analog zur Abbildung $i \mapsto i+1$ auf $\mathbb{N}$. Die x_i werden induktiv definiert, wobei wir ausnutzen, dass die Menge A nach Voraussetzung nicht endlich ist. Da A insbesondere nichtleer ist, können wir ein Element x_0 von A wählen. Sind $x_0, \ldots, x_i$ schon definiert, ist A verschieden von seiner endlichen Teilmenge $x_0, \ldots, x_i$, wir können also ein Element x_{i+1} in A wählen, das von $x_0, \ldots, x_i$ verschieden ist. Dass diese Konstruktion auch formal möglich ist und eine Teilmengen A_0

von A der gewünschten Form ergibt, wird durch die mengentheoretischen Axiome sichergestellt. □

Aus den Lemmas 68 und 72 folgt nun unmittelbar der folgende Satz.

73 Satz. *Eine Menge ist genau dann endlich, wenn es keine injektive Funktion von der Menge in eine ihrer echten Teilmengen gibt.*

74 Definition. *Eine Menge ist* ABZÄHLBAR UNENDLICH*, wenn sie gleichmächtig mit der Menge der natürliche Zahlen ist. Eine Menge ist* ABZÄHLBAR *wenn sie endlich oder abzählbar unendlich ist.*

Eine Menge ist ÜBERABZÄHLBAR*, wenn sie mächtiger als die Menge der natürlichen Zahlen ist.*

Das Standardbeispiel für eine überabzählbare Menge ist die Menge der reellen Zahlen, es gibt aber auch noch mächtigere Mengen. Tatsächlich gibt es zu jeder Menge eine mächtigere Menge.

75 Satz von Cantor. *Für jede Menge A gilt $A \prec \mathrm{Pot}(A)$.*

Beweis. Es sei A eine Menge. Die Abbildung $x \mapsto \{x\}$ ist eine injektive Abbildung von A in die Potenzmenge von A, somit gilt $A \preceq \mathrm{Pot}(A)$. Es bleibt zu zeigen, dass $\mathrm{Pot}(A) \preceq A$ falsch ist. Für einen Beweis durch Widerspruch nehmen wir an, dass letztere Aussage wahr ist, es also eine injektive Abbildung $g\colon \mathrm{Pot}(A) \to A$ gibt. Die Umkehrabbildung g^{-1} von g ordnet dann jedem x im Bild von g eine Menge $A_x = g^{-1}(x)$ zu. Dann ist

$$D \;=\; \{x \in \mathrm{bild}(g)\colon x \notin A_x\}$$

eine Teilmenge von A, die aber nach Konstruktion von allen Mengen in der Potenzmenge von A verschieden ist. Genauer gilt: nach Konstruktion ist D Teilmenge von A und für $x = g(D)$ ist D gleich A_x. Nach Definition von D ist folglich x genau dann in D, wenn x nicht in D ist. Dies ist ein Widerspruch, also muss die Widerspruchsannahme falsch sein. □

76 Satz. *Die Menge der reellen Zahlen ist gleichmächtig zur Potenzmenge der natürlichen Zahlen.*

Beweisskizze. Die nichtnegativen reellen Zahlen sind gleichmächtig mit dem Einheitsintervall $[0, 1] = \{x \in \mathbb{R}\colon 0 \leq x \leq 1\}$, siehe die Übungen. Ganz ähnlich lässt sich zeigen, dass auch die reellen Zahlen als Ganzes gleichmächtig mit dem Einheitsintervall sind.

Eine Teilmenge A von $\mathbb{N}$ kann mit ihrer charakteristische Folge $A(1)A(2)\ldots$ identifiziert werden, dabei ist $A(n)$ gleich 1, wenn n in A ist und sonst gleich 0. Eine Teilmenge A der natürlichen Zahlen entspricht so auch der

reellen Zahl im Einheitsintervall mit Binärdarstellung $0.A(1)A(2)\ldots$. Dies ist fast schon eine Bijektion, es ist allerdings noch zu berücksichtigen, dass zum Beispiel die reellen Zahlen $0,01$ und $0,001111\ldots$ identisch sind. □

Das folgendes Korollar zu Satz 76 ergibt sich sofort mit dem Satz von Cantor und der Definition von überabzählbar.

77 Korollar. *Die Menge der reellen Zahlen ist überabzählbar.*

1.7 Aussagenlogische Funktionen und Quantifikation

Funktionen auf Wahrheitswerten Aussagenlogische Funktionen bilden Wahrheitswerte auf Wahrheitswerte ab. Die Wahrheitswerte wahr und falsch werden dabei als 1 beziehungsweise 0 geschrieben. Zum Beispiel ist die Funktion $\neg$, die NEGATION, wie folgt definiert

$$\begin{aligned} \neg\colon \{0,1\} &\longrightarrow \{0,1\}, \\ 0 &\longmapsto \neg(0) = 1, \\ 1 &\longmapsto \neg(1) = 0. \end{aligned}$$

Aussagenlogische Funktionen, deren Argument zwei Wahrheitswerte umfasst, sind formal Funktionen von $\{0,1\}^2$ nach $\{0,1\}$, werden aber meist als aussagenlogische Operatoren aufgefasst und in Infixnotation geschrieben. Solche Funktionen können durch ihre WAHRHEITSTAFEL beschrieben werden. Es folgen die wichtigsten aussagenlogische Funktionen und deren Wahrheitstafeln.

- $\wedge$ bezeichnet die UND-VERKNÜPFUNG, auch: KONJUNKTION,
- $\vee$ bezeichnet die ODER-VERKNÜPFUNG, auch: DISJUNKTION, genauer: INKLUSIVE ODER-VERKNÜPFUNG, INLUSIVE DISJUNKTION
- $\oplus$ bezeichnet die EXKLUSIVES ODER-VERKNÜPFUNG, auch: EXKLUSIVE DISJUNKTION, PARITYFUNKTION,
- $\leftrightarrow$ bezeichnet die ÄQUIVALENZ,
- $\rightarrow$ bezeichnet die IMPLIKATION.

x	y	$x \wedge y$	$x \vee y$	$x \oplus y$	$x \leftrightarrow y$	$x \rightarrow y$
0	0	0	0	0	1	1
0	1	0	1	1	0	1
1	0	0	1	1	0	0
1	1	1	1	0	1	1

Werden in mathematischen Texten Aussagen durch das Wort „und“ verküpft, so ergibt sich der Wahrheitswert der Gesamtaussage entsprechend der Wahrheitstafel der Konjunktion. Eine Aussage der Form „A und B“ ist also wahr, falls die Aussagen A und B beide wahr sind und ist sonst falsch. Entsprechendes gilt für andere Formulierungen und aussagenlogische Funktionen.

Konjunktion Die Und-Verknüpfung $\wedge$ wird durch Formulierungen wie „und“, „sowie“, „sowohl ... als auch“ ausgedrückt, auch das Wort „aber“ bezeichnet eine Und-Verküpfung wobei zusätzlich ein Gegensatz zwischen den verknüpften Aussagen ausgedrückt werden soll.

Disjunktion Das inklusive Oder $\vee$ entspricht dem Wort „oder“.

Exklusive Disjunktion Das exklusive Oder $\oplus$ entspricht der Formulierung „entweder ... oder“.

Äquivalenz Die Äquivalenz $\leftrightarrow$ kann durch „äquivalent“ und „genau dann, wenn“ ausgedrückt werden, als abkürzende Schreibweise auch durch $\Leftrightarrow$.

Implikation Die Implikation $\rightarrow$ kann durch „impliziert“ und „wenn, ... dann“ ausgedrückt werden, als abkürzende Schreibweise auch durch $\Rightarrow$. Weiter durch Formulierungen mit den Wörtern folgern oder schließen: dass die Aussage A die Aussage B impliziert, kann so ausgedrückt werden, dass Aussage B aus A gefolgert oder geschlossen werden kann.

78 Bemerkung. *Falls die Implikation $A \Rightarrow B$ gilt, heißt A* HINREICHENDE BEDINGUNG *für B und B heißt* NOTWENDIGE BEDINGUNG *für A:*

- *wenn A gilt, gilt auch B,*
- *damit A gelten kann, muss B gelten.*

Damit lässt sich zum Beispiel zeigen, dass A nicht gilt, indem gezeigt wird, dass B nicht gilt. Tatsächlich sind die Implikation $A \Rightarrow B$ und ihre KONTRAPOSITION $\neg B \Rightarrow \neg A$ *logisch äquivalent.*

Quantifikation Eine EXISTENZIELL QUANTIFIZIERTE Aussage wie

> Es gibt eine natürliche Zahl, die Nachfolger von 7 ist.

besagt, dass es in einer Menge ein oder mehrere Elemente mit einer bestimmten Eigenschaft gibt. Eine UNIVERSELL QUANTIFIZIERTE Aussage wie

> Für alle natürlichen Zahlen n ist der Nachfolger von n echt größer als n.

besagt, dass alle Elemente einer Menge eine bestimmte Eigenschaft haben. Solche Aussagen können unter Verwendung des EXISTENZQUANTORS $\exists$ und des ALLQUANTORS $\forall$ wie folgt geschrieben werden

$$\exists x \in \mathbb{N}\,(\mathrm{s}(7) = n), \qquad \forall x \in \mathbb{N}\,(x < \mathrm{s}(x)).$$

Diese Schreibweisen sind ähnlich zu prädikatenlogischen Formeln und als abkürzende Schreibweisen zu verstehen, wenn sie wie hier nicht im Zusammenhang mit Betrachtungen zur Prädikatenlogik verwendet werden.

Bei existenziell oder universell quantifizierten Aussagen muss klar sein, über welche Menge quantifiziert wird. Die beiden folgenden Aussagen unterscheiden sich nur in der Menge, über die quantifiziert wird, die erste Aussage ist falsch, die zweite ist richtig.

$$\exists x \in \mathbb{N}\ (5 + x = 2), \qquad \exists x \in \mathbb{Z}\ (5 + x = 2).$$

Wir werden im Folgenden die Bedeutung von Aussagen mit Quantifikation anhand von Beispielen verdeutlichen, verzichten aber sowohl auf die Betrachtung des allgemeinen Falls als auch auf eine formale Definition.

Die Bedeutung einer Aussage mit zwei oder mehr gleichartige Quantoren ist naheliegend, zum Beispiel ist eine Aussage der Form $\forall x \in \mathbb{N}\ \forall y \in \mathbb{N}(\ldots)$ eben genau dann wahr, wenn für alle Paare x und y von natürlichen Zahlen die Aussage in Klammern wahr ist. Letztere Aussage muss auch im Fall $x = y$ wahr sein, dies ist nur für die zweite der beiden folgenden Aussagen der Fall.

$$\forall x \in \mathbb{N}\ \forall y \in \mathbb{N}\ (x < y \vee y < x), \qquad \forall x \in \mathbb{N}\ \forall y \in \mathbb{N}\ (x \leq y \vee y \leq x).$$

Durch alternierende existenzielle und universelle Quantifikation können komplizierte Aussagen ausgedrückt werden, zum Beispiel, dass es keine größte natürliche Zahl gibt

$$\forall x \in \mathbb{N}\ \exists y \in \mathbb{N}\ (x < y).$$

oder dass eine Folge $x_1, x_2, \ldots$ von reellen Zahlen in dem Sinne gegen 0 konvergiert, dass für jede noch so kleine positive Zahl sich fast alle x_i höchstens um diese Zahl von 0 unterscheiden

$$\forall k \in \mathbb{N}\ \exists i_0 \in \mathbb{N}\ \forall i \in \mathbb{N}\ (\text{aus } i_0 \leq i \text{ folgt } -\tfrac{1}{k} < x_i < \tfrac{1}{k}). \tag{1.3}$$

Die Reihenfolge der Quantifikation ist wichtig, die erste der beiden folgenden Aussagen ist richtig, die zweite ist falsch

$$\forall x \in \mathbb{N}\ \exists y \in \mathbb{N}\ (x < y) \qquad \text{und} \qquad \exists y \in \mathbb{N}\ \forall x \in \mathbb{N}\ (x < y).$$

Allgemein darf in einem Block von Quantoren eine existenziell quantifizierte Variablen nur von den vorangehenden quantifizierten Variablen abhängen, nicht von den nachfolgenden. Zum Beispiel darf in (1.3) die Wahl von i_0 nur von k abhängen, aber nicht von i: für jedes k muss es ein i_0 geben, dass die Implikation in Klammern für alle i wahr macht. Falls i_0 von i abhängen dürfte, würde es genügen, zu gegebenen i immer $i_0 = i + 1$ setzen, die Implikation in Klammern in (1.3) wäre dann immer richtig, unabhängig davon, wie groß die Werte x_i sind.

1.8 Beweismethoden

Beweise und Implikationen In einem mathematische Beweis wird aus einer VORAUSSETZUNG eine FOLGERUNG abgeleitet, wobei neben der Voraussetzung auch bekannte Zusammenhänge verwendet werden dürfen, zum Beispiel die Aussage eines bereits bewiesenen Satzes oder die definierenden Bedingungen eines Begriffs.

Dies kann auch so formuliert werden, dass in einem Beweis gezeigt wird, dass die Voraussetzung A die Folgerung B impliziert, es wird also die Implikation $A \Rightarrow B$ bewiesen. Ist dann die Voraussetzung wahr, so auch die Folgerung. Die Implikation $A \Rightarrow B$ wiederum wird in mehreren Beweisschritten gezeigt, die jeweils ebenfalls einer Implikation entsprechen und in der jeweils ein Hilfsaussagen C_i bewiesen wird. Der Beweis hat als Voraussetzung $A = C_1$ und für ein natürliche Zahl t und $i = 1, \ldots, t-1$ wird sukzessive gezeigt, dass aus C_1 bis C_i die Ausage C_{i+1} folgt, wobei C_t gleich B sei. Damit ist dann die Implikation $A \Rightarrow B$ bewiesen, aus A folgt B. Falls die Aussage A wahr, sind auch die Aussage C_1 bis C_t und damit auch B wahr. Ist A dagegen falsch, sind die Implikationen in den Beweisschritten trivialerweise wahr, so dass auch in diesem Fall die Aussage B aus Aussage A folgt.

Die in einem solchen Beweis gezeigten einzelnen Implikationen müssen nachvollziehbar sein. Gemäß der Wahrheitstafel der Implikation wird jede wahre Aussage von jeder anderen wahren Aussage impliziert, zum Beispiel impliziert die Aussage $2 + 2 = 4$ über die natürlichen Zahlen jeden wahren mathematischen Satz. Eine solche Implikation wäre aber als Beweisschritt nicht zulässig, wenn nicht nachvollziehbar ist, warum die Folgerung gilt.

Beweis durch Kontraposition Die KONTRAPOSITION einer Implikation $x \to y$ ist die Implikation $\neg y \to \neg x$. Eine Implikation und ihre Kontraposition sind logisch äquivalent, wie sich an den zugehörigen Wahrheitstafeln ablesen lässt.

x	y	$x \to y$	$\neg x$	$\neg y$	$\neg y \to \neg x$
0	0	1	1	1	1
0	1	1	1	0	1
1	0	0	0	1	0
1	1	1	0	0	1

Dass die Voraussetzung A die Folgerung B impliziert, kann mit einem BEWEIS DURCH KONTRAPOSITION gezeigt werden, indem bewiesen wird, dass die Voraussetzung, dass B nicht gilt impliziert, dass A nicht gilt. Statt der eigentlich zu beweisenden Implikation $A \Rightarrow B$ wird also deren Kontraposition $\neg B \Rightarrow \neg A$ bewiesen.

79 Beispiel. *Für alle natürlichen Zahlen n gilt die Implikation:*

$$\text{Wenn } n^2 \text{ ungerade ist, dann ist auch } n \text{ ungerade.} \tag{1.4}$$

Für einen Beweis dieser Aussage genügt es, für eine beliebig gewählte natürliche Zahl n die Implikation (1.4) *zu beweisen. Wir führen diesen Beweis durch Kontraposition und zeigen*

$$\text{Wenn } n \text{ gerade ist, dann ist auch } n^2 \text{ gerade.}$$

Dies folgt leicht: jede gerade natürliche Zahl n kann in der Form $n = 2i$ für eine natürliche Zahl i geschrieben werden, also ist auch $n^2 = 4i^2$ gerade. Bei der Formulierung der Kontraposition haben wir verwendet, dass eine natürliche Zahl genau dann gerade ist, wenn sie nicht ungerade ist.

Beweis durch Widerspruch Dass aus einer Vorausetzung A eine Folgerung B abgeleitet werden kann, wird in einem BEWEIS DURCH WIDERSPRUCH bewiesen, indem aus der Voraussetzung $A \wedge \neg B$ eine Aussage abgeleitet wird, die im Widerspruch zur Voraussetzug oder zu andern als wahr bekannten Aussagen steht. Dies zeigt dann, dass A und die Negation von B nicht gleichzeitig wahr sein können, wenn also die Voraussetzung A wahr ist, dann muss auch B wahr sein. Ein Beweis durch Widerspruch wird auch als INDIREKTER BEWEIS oder als REDUCTIO AD ABSURDUM bezeichnet, letzterer Begriff wird auch außerhalb der Mathematik verwendet.

Wir hatten bereits Beweise durch Widerspruch gesehen. Im Beweis des Satzes von Euklid wurde aus der Annahme, dass es nur endlich viele Primzahlen gibt ein Widerspruch zu einem zuvor gezeigten Lemma abgeleitet, nach dem jede natürliche Zahl ungleich 1 einen Primteiler hat. Um zu beweisen, dass die Wurzel aus 2 nicht rational ist, hatten wir angenommen, dass die Wurzel aus 2 durch einen Bruch $\frac{p}{q}$ in gekürzter Form dargestellt werden kann und daraus den Widerspruch abgeleitet, dass dann p und q beide gerade sind.

Beweise durch vollständige Induktion Mit einem Beweis durch vollständige Induktion kann eine Aussage für alle natürlichen Zahlen bewiesen werden, indem gezeigt wird, dass die Aussage auf 0 zutrifft und dass wenn die Aussage für eine natürliche Zahl n gilt, dann auch für $n+1$. Diese beiden Teile des Beweises werden als INDUKTIONANFANG und INDUKTIONSSCHRITT bezeichnet. Im Induktionsschritt von n nach $n+1$ darf die INDUKTIONSANNAHME, auch: INDUKTIONSVORAUSSETZUNG benutzt werden, das heißt, dass die zu beweisende Aussage für n bereits bewiesen ist.

Im Zusammenhang mit Beweisen durch vollständige Induktion gibt es die Sprechweise, dass für eine natürliche Zahl n die INDUKTIONSBEHAUPTUNG AUF n ZUTRIFFT. Dies bedeutet, dass die Aussage, die für alle natürlichen

Zahlen zu zeigen ist, für die Zahl n richtig ist. Damit lassen sich die Teile eines Induktionsbeweises wie folgt beschreiben.

> Im *Induktionsanfang* wird gezeigt, dass die Induktionsbehauptung für $n = 0$ gilt.
>
> Im *Induktionsschritt* von n nach $+1$ wird für eine beliebige natürliche Zahl n gezeigt, dass die Induktionsbehauptung für $n + 1$ richtig ist, falls sie für n richtig ist.

Im Induktionsanfang kann statt mit $n = 0$ auch mit $n = 1$ oder sogar mit $n > 1$ begonnen werden, wenn die zu beweisenden Aussage für kleinere Werte von n uninteressant ist oder keinen Sinn ergibt.

80 Beispiel. *Wir beweisen die Summenformel* $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$ *durch vollständige Induktion. Diese Formel ist auch für* $n = 0$ *richtig, da die Summe einer Summation von* 1 *bis* 0 *nach Definition gleich* 0 *ist. Dieser Fall ist aber weniger anschaulich und auch nicht besonders interessant, wir zeigen die Summenformel deswegen nur für* $n \geq 1$*, entsprechend wird im Induktionsanfang der Fall* $n = 1$ *betrachtet.*

Induktionsanfang $n = 1$*:* *Es gilt* $\sum_{i=1}^{1} i = 1 = \frac{1\cdot(1+1)}{2}$.

Induktionsschritt $n \curvearrowright n + 1$*:*

$$\sum_{i=1}^{n+1} i = \left(\sum_{i=1}^{n} i\right) + (n+1) \overset{IV}{=} \frac{n(n+1)}{2} + \frac{2(n+1)}{2} = \frac{(n+1)(n+2)}{2}.$$

81 Bemerkung. *In einem Beweis durch vollständige Induktion darf im Induktionsschritt von* n *nach* $n+1$ *benutzt werden, dass die Induktionsannahme für alle natürlichen Zahlen* $i \leq n$ *gilt. Diese Variante der vollständigen Induktion wird als* Verlaufsinduktion *bezeichnet, weil sie im Spezialfall einer Behauptung über die Werte einer Funktion* f *im Induktionsschritt die Induktionsannahme nicht nur auf einen einzelnen Funktionswert* $f(n)$*, sondern auf das Anfangsstück* $f(1)f(2)\cdots f(n)$ *des Verlaufs der Werte der Funktion verwendet: es wird vorausgesetzt.*

Die Verlaufsinduktion lässt sich auf die übliche Induktion zurückführen. Um mit Verlaufsinduktion eine Aussage $P(n)$ *für alle* n *zu zeigen, zeigen wir mit der üblichen Induktion die Aussage* $P(1) \wedge \cdots \wedge P(n)$ *für alle* n*. Es wird sozusagen ein Beweis von* $P(n)$ *für alle* n *mit Verlaufsinduktion durch die übliche Induktion simuliert, wobei in der Simulation die Induktionsannahme nicht darin besteht, dass* $P(n)$ *für eine einzelne natürliche Zahl* n *wahr ist, sondern darin, dass* $P(i)$ *für alle* $i \leq n$ *wahr ist.*

Der Induktionsanfang besteht dann aus dem Beweis von $P(1)$*. Im Induktionsschritt von* n *nach* $n+1$ *ist* $P(1) \wedge \cdots \wedge P(n) \wedge P(n+1)$ *zu beweisen, was*

auf einen Beweis von $P(n+1)$ hinausläuft, da nach Induktionsannahme die Aussagen $P(1)$ bis $P(n)$ richtig sind. Diese Aussagen dürfen dann auch im Beweis von $P(n+1)$ verwendet werden.

82 Beispiel. *Wir zeigen mit Verlaufsinduktion über n, dass jede natürliche Zahl $n \geq 2$ als Produkt von Primzahlen dargestellt werden kann. Im Induktionsanfang wird $n = 2$ betrachtet, da n Primzahl ist, ist nichts zu zeigen. Im Induktionsschritt von n nach $n+1$ werden zwei Fälle unterschieden. Falls $n+1$ Primzahl ist, ist wie im Induktionsanfang nichts zu zeigen. Falls $n+1$ keine Primzahl ist, sei t ein nichttrivialer Teiler von $n+1$ und es sei $n+1 = k \cdot t$. Es gilt $2 \leq t < n+1$ und damit auch $2 \leq k < n+1$. Nach Induktionsannahme der Verlaufsinduktion könnnen somit k und t als Produkte von Primzahlen dargestellt werden, das Produkt dieser beiden Darstellungen ist eine Darstellung von $n+1$ als Produkt von Primzahlen.*

Im folgenden Beispiel für einen Beweis mit Verlaufsinduktion wird im Induktionsanfang die Induktionsbehauptung für $n = 1$ und $n = 2$ gezeigt, der Induktionsschritt geht von $n+1$ nach $n+2$ für beliebiges $n \geq 1$.

83 Beispiel. *Die* FIBONACCI-ZAHLEN $\mathrm{fib}(1), \mathrm{fib}(2), \ldots$ *sind definiert durch* $\mathrm{fib}(1) = \mathrm{fib}(2) = 1$ *und, für $n \geq 1$, durch*

$$\mathrm{fib}(n+2) = \mathrm{fib}(n) + \mathrm{fib}(n+1).$$

Viele natürliche Prozesse lassen sich durch die Fibonacci-Zahlen beschreiben, in der Informatik sind sie das Standardbeispiel für die Umwandlung eines rekursiven und ausgesprochen ineffizienten Berechnungsverfahren in ein effizientes iteratives Berechnungsverfahren.

Erstaunlicherweise gilt die folgende geschlossene Darstellung der Fibonacci-Zahlen

$$\mathrm{fib}(n) = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right). \tag{1.5}$$

Wir beweisen die Korrektheit dieser Darstellung durch Verlaufsinduktion, die Induktionsbehauptung ist also gerade (1.5)*. Wir definieren*

$$a = \frac{1+\sqrt{5}}{2} \quad \textit{und} \quad b = \frac{1-\sqrt{5}}{2}, \qquad (1.5) \textit{ wird damit zu} \qquad \mathrm{fib}(n) = \frac{a^n - b^n}{\sqrt{5}}.$$

Zur weiteren Verwendung im Beweis halten wir fest

$$\begin{aligned} a^2 &= \frac{1+2\sqrt{5}+5}{4} = \frac{3+\sqrt{5}}{2} = 1 + \frac{1+\sqrt{5}}{2} = 1 + a, \\ b^2 &= \frac{1-2\sqrt{5}+5}{4} = \frac{3-\sqrt{5}}{2} = 1 + \frac{1-\sqrt{5}}{2} = 1 + b, \end{aligned} \tag{1.6}$$

Induktionsanfang $n = 1$ und $n = 2$:

Für $n = 1$ beziehungsweise $n = 2$ ist die Induktionsbehauptung (1.5) *wahr, da gilt*

$$\frac{a^1-b^1}{\sqrt{5}} = \frac{1}{\sqrt{5}}\frac{2\sqrt{5}}{2} = 1 = \mathrm{fib}(1) \qquad \textit{und} \qquad \frac{a^2-b^2}{\sqrt{5}} \overset{(1.6)}{=} \frac{a-b}{\sqrt{5}} = 1 = \mathrm{fib}(2).$$

Induktionsschritt $n+1 \curvearrowright n+2$ für $n \geq 1$: Im Induktionsschritt verwenden wir zwei Gleichungen, die sich direkt aus (1.6) *ergeben*

$$\begin{aligned} a^{n+2} &= a^n \cdot a^2 = a^n(1+a) = a^n + a^{n+1}, \\ b^{n+2} &= b^n \cdot b^2 \ = b^n(1+b) \ = b^n + b^{n+1}. \end{aligned}$$

Damit folgt dann die Induktionsbehauptung für $n+2$

$$\begin{aligned} \frac{a^{n+2}-b^{n+2}}{\sqrt{5}} &= \frac{a^n + a^{n+1} - b^n - b^{n+1}}{\sqrt{5}} \\ &= \frac{a^n - b^n}{\sqrt{5}} \ + \ \frac{a^{n+1} - b^{n+1}}{\sqrt{5}} \overset{IV}{=} \mathrm{fib}(n) + \mathrm{fib}(n+1) = \mathrm{fib}(n+2). \end{aligned}$$

84 Bemerkung. *Eine* Wohlordnung *ist eine Ordnung auf einer Menge A, so dass jede nichtleere Teilmenge von A ein kleinstes Element besitzt. Zum Beispiel ist die übliche Kleiner-oder-gleich-Relation $\leq$ ein Wohlordnung auf den natürlichen, nicht aber auf den ganzen Zahlen.*

Die vollständige Induktion entspricht einem Spezialfall des folgenden Induktionsprinzips für Wohlordnungen. Um zu beweisen, dass eine Eigenschaft P für alle Elemente einer wohlgeordnete Menge A gilt, genügt es zu zeigen, dass für alle x in A

$$P(x) \textit{ gilt, falls } P(y) \textit{ für alle } y \textit{ in } A \textit{ gilt, die echt kleiner als } x \textit{ sind.} \quad (1.7)$$

Der Beweis von 1.7 entspricht dem Induktionsschritt der Verlaufsinduktion: es wird gezeigt, dass die Eigenschaft P auf x zutrifft, falls die Eigenschaft auf alle echt kleineren y zutrifft. Der Beweis von 1.7 umfasst aber auch ein Äquivalent zum Induktionsanfang. Für das kleinste Element von A gilt trivialerweise, dass P auf alle echt kleineren Elemente von A zutrifft, einfach weil es keine echt kleineren Elemente gibt. Um 1.7 zu beweisen, muss somit insbesondere gezeigt werden, dass P auf das kleinste Element in A zutrifft.

Aus (1.7) *folgt $P(x)$ für alle x in A. Für einen Beweis sei F die Menge aller x in A, für die $P(x)$ falsch ist. Falls F leer ist, sind wir fertig. Anderfalls hat F ein kleinstes Element x, das heißt, alle y in A, die echt kleiner sind als x sind nicht in F, für alle solche y gilt also $P(x)$. Damit folgt aus* (1.7)*, dass $P(x)$ gilt, im Widerspruch zur Wahl von x in F.*

Wir beschließen den Abschnitt über Induktionsbeweise mit einem Beispiel für die strukturelle Induktion über den Aufbau von Ausdrücken.

85 Beispiel. *Im Abschnitt 1.7 über aussagenlogische Funktionen wurde die Parity-Funktion $\oplus$ eingeführt, die auch als exklusives Oder bezeichnet wird. Im Folgenden definieren wir induktuktiv die Menge der Funktionsausdrücke, die mit den Konstanten 0 und 1 und dem Symbol $\oplus$ gebildet werden können, wobei 0 für falsch und 1 für wahr steht. Wir beweisen dann durch* STRUKTURELLE INDUKTION *über den induktiven Aufbau von Parity-Ausdrücken, dass der Wahrheitswert eines solchen Ausdrucks nur von der Anzahl der im Ausruck vorkommenden Einsen abhängt.*

Parity-Ausdrücke über den Konstanten 0 und 1, im Rest dieses Beispiels kurz als Parity-Ausdrücke bezeichnet, seien induktiv wie folgt definiert

(i) *0 und 1 sind Parity-Ausdrücke.*

(ii) *Sind α und β Parity-Ausdrücke, dann ist auch $(\alpha \oplus \beta)$ ein Parity-Ausdruck.*

Diese Definition ist so zu verstehen, dass es keine weiteren Parity-Ausdrücke gibt. Formal bedeutet dies, dass alle Mengen von Ausdrücken betrachtet werden, die 0 und 1 enthalten und unter der Bildung von neuen Ausdrücken gemäß (ii) abgeschlossen sind. Der Durchschnitt über alle solchen Menge gehört zu diesen Mengen und ist damit unter all diesen Mengen die kleinste bezüglich der mengentheoretischen Inklusion. Die Menge der Parity-Ausddrücke ist dann gleich diesem Durchschnitt, also gleich der kleinsten Menge, welche (i) und (ii) erfüllt.

Entsprechend der induktiven Definition von Parity-Ausdruck lassen sich Funktionen auf Parity-Ausdrücken induktiv definieren, zum Beispiel für einen Parity-Ausdruck α die Anzahlen der Vorkommen von 0 und von 1 in α, diese seien mit $\#_0(\alpha)$ beziehungsweise $\#_1(\alpha)$ bezeichnet und wie folgt definiert

(i) $\#_0(0) = \#_1(1) = 1$ *und* $\#_0(1) = \#_1(0) = 0$.

(ii) $\#_i(\ (\alpha \oplus \beta)\ ) = \#_i(\alpha) + \#_i(\beta)$ *für* $i = 0, 1$.

Der besseren Lesbarkeit wegen lassen wir bei der Anwendung der Funktionen $\#_0$ und $\#_1$ die äußeren Klammern des Arguments weg. Nach Definition gilt beispielsweise

$$\#_1(0 \oplus 1) = 1, \qquad \#_1((0 \oplus 1) \oplus 1) = 2, \qquad \#_1((1 \oplus 1) \oplus (1 \oplus 0)) = 3.$$

Von den Parity-Ausdrücken in diesen Beispielen hat der mittlere den Wahrheitswert wahr, die beiden anderen den Wert falsch. Allgemein gilt

Ein Parity-Ausdruck γ ist genau dann wahr, wenn $\#_1(\gamma)$ ungerade ist.

Wir beweisen diese Aussage durch STRUKTURELLE INDUKTION ÜBER DEN AUFBAU VON PARITY-AUSDRÜCKEN.

(i) *Im Induktionsanfang ist der betrachtete Parity-Ausdruck γ gleich einem der Ausdrücke 0 oder 1. Durch Nachrechnen ergibt sich, dass die zu beweisende Aussage für beide Ausdrücke wahr ist.*

(ii) *Im Induktionsschritt betrachten wir einen Parity-Ausdruck γ der Form $(\alpha \oplus \beta)$. Wir unterscheiden vier Fälle, je nachdem, ob die Werte $\#_1(\alpha)$ und $\#_1(\beta)$ jeweils gerade oder ungerade sind.*

Sind beide Werte gerade oder beide ungerade, so ist $\#_1(\alpha \oplus \beta)$ gerade. In diesem Fall sind α und β nach Induktionsannahme beide falsch oder beide wahr, folglich ist $\alpha \oplus \beta$ falsch.

Ist einer der Werte $\#_1(\alpha)$ und $\#_1(\beta)$ gerade und der andere ungerade, so ist $\#_1(\alpha \oplus \beta)$ ungerade. In diesem Fall ist nach Induktionsannahme genau einer der Parity-Ausdrücke α und β wahr, folglich ist $\alpha \oplus \beta$ wahr.

Beim Wahrheitswert eines Partiy-Ausdrucks kommt es nur auf die Anzahl der Einsen an und inbesondere ist die Klammerung des Ausdrucks irrelevant. Die Klammern in Parity-Ausdrücke können also weggelassen werden. Dies bedeutet gerade, dass für die Parity-Operation das Assoziativgesetz gilt.

Die strukturelle Induktion über den induktiven Aufbau von Parity-Ausdrücken kann auch als Verlaufsinduktion über die, geeignet definierte, Länge von Parity-Ausdrücke formuliert werden. Der Beweis ändert sich dadurch nicht wesentlich, im Induktionsanfang werden weiter die Parity-Ausdrücke 0 und 1 behandelt, diese haben Länge 1, und im Induktionsschritt wird weiter die Aussage für $(\alpha \oplus \beta)$ auf die Induktionsannahme für die kürzeren Ausdrücke α und β zurückgeführt.

2 Gruppen und Körper

2.1 Gruppen

Die Gruppenaxiome Eine Gruppe ist ein Paar $(G, \cdot)$ aus einer Menge G und einer Verknüpfung $\cdot$ auf G, welche die GRUPPENAXIOME G1, G2 und G3 erfüllen. Der Einfachheit halber wird manchmal auch die Menge G selbst als Gruppe bezeichnet, wenn die GRUPPENOPERATION $\cdot$ klar ist.

86 Definition. *Eine* GRUPPE *ist ein Paar $(G, \cdot)$ aus einer Menge G und einer zweistelligen Verknüpfung $\cdot$ auf G mit folgenden Eigenschaften.*

(G1) ***Assoziativität*** *Für alle a, b und c in G gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.*

(G2) ***Neutrales Element*** *Es gibt ein Element e in G, so dass $e \cdot a = a$ für alle a in G gilt.*

(G3) ***Inverse Elemente*** *Für ein Element e wie in G2 gibt es für jedes Element a in G ein Element a^{-1} so dass $a^{-1}a = e$ gilt.*

Eine Gruppe $(G, \cdot)$ ist ABELSCH*, auch:* KOMMUTATIV*, falls folgende Eigenschaft gilt.*

(G4) ***Kommutativität*** *Für alle a und b in G gilt $a \cdot b = b \cdot a$.*

Dass die Gruppenoperation $\cdot$ einer Gruppes G eine zweistellige Verknüpfung auf G ist bedeutet insbesondere, dass für alle x und y in G auch $x \cdot y$ in G liegt, mit anderen Worten: die Menge G ist UNTER DER GRUPPENOPERATION $\cdot$ ABGESCHLOSSEN. Weiter sei vereinbart, dass es auf der leeren Menge keine Verküpfung geben kann, eine Menge mit Verknüpfung somit nichtleer ist.

Standardbeispiele für kommutative Gruppen sind $(\mathbb{Z}, +)$ und $(\mathbb{Q} \setminus \{0\}, \cdot)$, wobei die Verknüpfungen $+$ und $\cdot$ für die übliche Addition beziehungsweise Multiplikation stehen. Ein Beispiel für eine nichtkommutative Gruppe ist die Menge der bijektiven Abbildungen auf einer Menge mit drei oder mehr Elementen mit der Hintereinanderausführung als Gruppenoperation, siehe Beispiel 92.

87 Bemerkung. *Es lässt sich zeigen, dass für eine assoziative Verknüpfung auch bei komplizierteren Ausdrücken als denen im Assoziativgesetz die Klammerung irrelevant für den Wert des Ausdrucks ist und somit weggelassen werden kann. Wir verzichten auf den einfachen, aber technischen und wenig instruktiven Beweis.*

88 Vereinbarung. *Gruppen könne in* MULTIPLIKATIVER FORM *und in* ADDITIVER FORM *geschrieben werden. In der multiplikativen Form wird die*

Verknüpfung als $\cdot$ geschrieben, das neutrale Element als e oder 1 und inverse Elemente als a^{-1}. In der additiven Form wird die Verknüpfung als $+$ geschrieben, das neutralem Element als e oder 0 und inverse Elemente als $-a$. In der multiplikativen Form kann das Symbol für die Verknüpfung weggelassen werden, beispielsweise steht dann abc für $a \cdot b \cdot c$.

Ein Element e wie in G2 heißt ein LINKSNEUTRALES ELEMENT DER GRUPPE, ein Element x^{-1} wie in G3 heißt ein LINKSINVERSES ELEMENT ZU x BEZÜGLICH e, RECHTSNEUTRALE und RECHTSINVERSE Element sind analog definiert, zum Beispiel ist e rechtsneutral, wenn für alle a in der Gruppe $a \cdot e = a$ gilt.

Als nächstes zeigen wir, dass es in einer Gruppe nur ein linksneutrales Element gibt und dieses gleich dem einzigen rechtsneutralen Element der Gruppe ist. Links- und rechtsinverse Elemente sind dann immer invers bezüglich des eindeutig bestimmten linksneutralen Elements, so dass dieser Bezug weggelassen werden kann. Weiter gibt es zu gegebenem a nur ein Linksinverses, dieses ist gleich dem einzigen Rechtsinversen zu a und wird als INVERSES zu a bezeichnet.

89 Satz. *Es sei G eine Menge und $\cdot$ sei eine assoziative Verknüpfung auf G. Das Paar $(G, \cdot)$ ist genau dann eine Gruppe, wenn folgende Eigenschaften gelten.*

(G2′) Es gibt ein eindeutig bestimmtes Element e in G, das NEUTRALE ELEMENT *von G, so dass $ea = ae = a$ für alle a in G gilt.*

(G3′) Für das neutrale Element e von G gibt es zu jedem Element a in G ein eindeutig bestimmtes Element a^{-1}, das INVERSE ELEMENT *zu a, so dass $a^{-1}a = aa^{-1} = e$ gilt.*

Satz 89 besagt, dass das neutrale Element und die inversen Elemente zu den verschiedene Gruppenelementen jeweils eindeutig sind. Aus dem Beweis des Satzes folgt zusätzlich, dass es neben dem neutralen Element keine weiteren links- oder rechtneutralen Elemente und für jedes Gruppenelmennt a neben dem Inversen zu a keine weiteren Links- oder Rsechtsinversen zu a geben kann.

Beweis von Satz 89. Wenn die Eigenschaften G2′ und G3′ gelten, sind die Eigenschaften G2 und G3 aus der Definition von Gruppe offensichtlich erfüllt, weiter ist die Gruppenoperation als assoziativ vorausgesetzt, folglich ist $(G, \cdot)$ eine Gruppe.

Für den Beweis der umgekehrten Implikation nehmen wir an, dass $(G, \cdot)$ eine Gruppe ist. Wir wählen ein linksneutrales Element e wie in G2 für das G3 gilt, das heißt, zu jedem a in G gibt es ein Linksinverses a^{-1} zu a bezüglich e.

Wir zeigen als erstes, dass a^{-1} auch rechtsinvers zu a bezüglich e ist, dies gilt wegen

$$aa^{-1} = e(aa^{-1}) = ((a^{-1})^{-1}a^{-1})(aa^{-1}) = (a^{-1})^{-1}(a^{-1}a)a^{-1} = e. \quad (2.1)$$

Daraus folgt, dass e auch rechtsneutral ist, da für alle a in G gilt

$$ae = a(a^{-1}a) = (aa^{-1})a \overset{(2.1)}{=} ea = a.$$

Somit ist e links- und rechtsneutral. Für ein beliebiges linksneutrales Element e_1 und rechtsneutrales Element e_2 gilt aber $e_1 = e_1e_2 = e_2$, folglich sind alle links- und rechtsneutralen Elemente gleich e. Daraus folgt weiter, dass es zu jedem a nur ein linksinverses Element geben kann, denn für jedes linksinverse Element x zu a gilt, da wie schon gezeigt a^{-1} auch rechtsinvers zu a ist,

$$xa = e \;\Leftrightarrow\; (xa)a^{-1} = ea^{-1} \;\Leftrightarrow\; x(aa^{-1}) = a^{-1} \;\Leftrightarrow\; x = a^{-1}.$$

Alle linksinversen Element zu a sind also gleich a^{-1}. Ganz ähnlich lässt sich zeigen, dass es zu jedem a nur ein rsechtinverses Element gibt. Da das eindeutig bestimmte Linksinverse zu a auch rechtsinvers zu a ist, sind das Linksinverse und das Rechtsinverse zu a gleich. □

90 Bemerkung. *In einer Gruppe $(G, \cdot)$ sind Gleichungen der Form $ax = b$ und $xa = b$ immer eindeutig lösbar: zu gegebenen Elementen a und b von G gilt die erste Gleichung genau für $x = a^{-1}b$ und die zweite Gleichung genau für $x = ba^{-1}$.*

Es folgt, dass Abbildungen von G nach G der Form $x \mapsto ax$ und $x \mapsto xa$ mit a in G immer bijektiv sind. Die Surjektivität gilt wegen der der Lösbarkeit der zugehörigen Gleichungen $ax = b$ beziehungsweise $xa = b$, die Injektivität wegen der Eindeutigkeit der Lösungen.

Beispiele von Gruppen Die natürlichen Zahlen mit der Addition als Verknüpfung sind keine Gruppe, da die inversen Elemente fehlen. Die Mengen $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ der ganzen, der rationalen und der reelleen Zahlen mit der üblichen Addition als Verknüpfung sind Gruppen mit neutralem Element 0. Mit der Multiplikation als Verknüpfung bilden diese Mengen keine Gruppe, es gibt zwar ein neutrales Element 1, aber die 0 hat kein multiplikatives Inverses, für die ganzen Zahlen gilt letzteres auch für alle anderen Elemente außer 1 und -1. Die Mengen $\mathbb{Q} \setminus \{0\}$ und $\mathbb{R} \setminus \{0\}$ dagegen sind Gruppen bezüglich der Multiplikation.

Standardbeispiele für endliche Gruppen ergeben sich aus den im folgenden Abschnitt ausführlicher behandelten Restklassen: für jede natürliche Zahl $n \geq 2$ bilden die Restklassen zu n eine Gruppe bezüglich der Addition, und für jede Primzahl n bilden die von der Restklasse von 0 verschiedenen Restklassen eine Gruppe bezüglich der Multiplikation.

91 Beispiel. *Es sei X eine nichtleere Menge X und $(G,+)$ sei eine Gruppe mit neutralem Element 0. Dann ist die Menge der Funktionen von X nach G eine Gruppe unter der wie folgt definierten Summe von Funktionen.*

$$(f+g)(x) = f(x) + g(x),$$

Das neutrale Element dieser Gruppe ist die konstante Funktion mit Wert 0, das Inverse einer Funktion f ist die Funktion $x \mapsto -f(x)$, die x auf das Inverse von $f(x)$ abbildet.

92 Beispiel. *Für eine nichtleere Menge X bildet die Menge der bijektiven Abbildungen von X nach X mit der Hintereinanderausführung $\circ$ als Verküpfung eine Gruppe. Das neutrale Element ist die Identitätsabbildung, das Inverse einer Abbildung ist deren Umkehrabbildung. Falls X mindestens drei Elemente hat, ist diese Gruppe nichtkommutativ, siehe die Übungen.*

Untergruppen

93 Definition. *Eine Teilmengen U einer Gruppe G ist eine* Untergruppe *von G, falls U mit der auf U eingeschränkten Verknüpfung der Gruppe G eine Gruppe ist.*

Eine Untergruppe ist nie leer, das sie als Gruppe ein neutrales Element enthalten muss. Jede Gruppe G mit neutralem Element e hat die trivialen Untergruppen G und $\{e\}$. Die Menge $\{-1,1\}$ ist eine nichttriviale Untergruppe der Gruppe $(\mathbb{R}\setminus\{0\},\cdot)$. Die Menge $d\mathbb{Z} = \{dn\colon n \in \mathbb{Z}\}$ der Vielfachen einer natürlichen Zahl d ist eine Untergruppe von $(\mathbb{Z},+)$.

94 Bemerkung. *Eine Untergruppe U einer Gruppe G ist eine Gruppe, enthält also nach Satz 89 ein eindeutiges neutrales Element und für jedes a in U ein eindeutiges inverse Elemente. Diese sind identisch mit dem neutralen Element beziehungsweise den jeweiligen inversen Elementen in G.*

Für einen Beweis seien e und e_U die neutralen Elemente von G beziehungsweise U und für jedes a in U seien a^{-1} und a_U^{-1} das inverse Element zu a in G beziehungsweise in U. Für ein beliebiges a in U gilt dann

$$e_U \cdot a = a \;\Leftrightarrow\; (e_U \cdot a)a^{-1} = a \cdot a^{-1} \;\Leftrightarrow\; e_U(a \cdot a^{-1}) = e \;\Leftrightarrow\; e_U = e,$$

das heißt, e_U ist gleich e. Damit folgt aber für alle a in U

$$a_U^{-1} \cdot a = e_U \;\Leftrightarrow\; (a_U^{-1} \cdot a)a^{-1} = ea^{-1} \;\Leftrightarrow\; a_U^{-1}(aa^{-1}) = a^{-1} \;\Leftrightarrow\; a_U^{-1} = a^{-1},$$

das heißt, auch die Inversen in U stimmen mit den Inversen in G überein.

Eine Untergruppe einer Gruppe G ist eine Gruppe und daher nichtleer und unter der Gruppenoperation von G abgeschlossen, nach Bemerkung 94 muss

weiter die Untergruppe zu jedem ihrer Element das zugehörige Inverse aus G enthalten. Nach dem folgenden Untergruppenkriterium sind diese notwendigen Bedingungen zusammen auch hinreichend dafür, dass eine Teilmenge einer Gruppe eine Untergruppe ist.

95 Satz (Untergruppenkriterium)**.** *Es sei $(G, \cdot)$ eine Gruppe. Eine Teilmenge U von G ist eine Untergruppe von G, falls U nichtleer und unter der Verknüpfung $\cdot$ abgeschlossen ist und zu einem Element auch immer dessen inverses Element in G enthält, das heißt, falls gilt*

$$\text{(i) } U \neq \emptyset, \qquad \text{(ii) } \forall a, b \in U\ (ab \in U), \qquad \text{(iii) } \forall a \in U\ (a^{-1} \in U).$$

Beweis. Die Teilmenge U von G erfülle (i) bis (iii). Wir zeigen, dass U bezüglich der Verknüpfung $\cdot$ eine Gruppe ist. Zunächst folgt aus (ii), dass die Gruppenoperation $\cdot$ tatsächliche eine Verknüpfung auf U ist. Das Assoziativgesetz für $\cdot$ überträgt sich dann von G auf U.

Da U nicht leer ist, können wir ein Element a aus U wählen, nach (iii) sind dann auch a^{-1} und damit nach (ii) das neutrale Element e von G in U, welches dann auch in U die Eigenschaften eines neutralen Elements hat. Die Existenz der inversen Elemente in U ergibt sich damit direkt aus (iii). □

96 Definition. *Es sei U eine Untergruppe der Gruppe $(G, \cdot)$. Zu gegebenem a aus G sei*

$$a \cdot U \ = \ \{a \cdot u \colon u \in U\}.$$

Mengen dieser Form werden als Nebenklasse *von G nach U bezeichnet.*

97 Lemma. *Es sei U Untergruppe einer Gruppe G. Alle Nebenklassen zu einer Untergruppe U sind gleichmächtig zu U. Die Nebenklassen von U bilden eine Zerlegung von G.*

Beweis. Jede Nebenklasse von U hat die Form $a \cdot U$ für ein Element a von G. Die Abbilddung $a \mapsto a \cdot x$ ist nach Bemerkung 90 eine Bijektion von G nach G. Die Einschränkung der Abbildung auf U ist somit eine Bijektion von U auf $\text{bild}(U) = a \cdot U$, die Untergruppe ist also gleichmächtig zur Nebenklasse $a \cdot U$.

Da die Untergruppe U das neutrale Element von G enthält, ist jedes Element a von G in der Nebenklasse $a \cdot U$ enthalten. Die Vereinigung aller Nebenklassen ist folglich gleich G. Es bleibt zu zeigen, dass die Nebenklassen paarweise disjunkt sind. Seien dazu a und b beliebige Element von G, so dass die zugehörigen Nebenklassen verschieden sind. Für einen Beweis durch Widerspruch nehmen wir an, dass der Schnitt der beiden Nebenklassen nicht leer ist. Dann gibt es u_a und u_b in U mit $a \cdot u_a = b \cdot u_b$. Damit gilt dann für jedes u in U, dass

$$a \cdot u = a \cdot e \cdot u = a \cdot (u_a \cdot u_a^{-1}) \cdot u = (a \cdot u_a) \cdot u_a^{-1} \cdot u = b \cdot \underbrace{u_b \cdot u_a^{-1} \cdot u}_{\text{in } U}$$

Element von $b \cdot U$ ist, die Nebenklasse $a \cdot U$ ist folglich Teilmenge der Nebenklasse $b \cdot U$. Mit einem symmetrischen Argument folgt auch die umgekehrte Inklusion, die beiden Nebenklassen sind somit identisch, im Widerspruch zur Annahme. □

98 Definition. *Für eine endliche Gruppe G ist die* Ordnung *von G gleich der Anzahl $|G|$ der Elemente von G.*

Für endliche Gruppen folgt aus Lemma 97 sofort, dass die Ordnung jeder Untergruppe einer Gruppe die Ordnung der Gruppe teilt.

99 Satz von Lagrange. *Es sei U eine Untergruppe der endlichen Gruppe G. Dann gilt $|G| = k|U|$, dabei ist k gleich der Anzahl der Nebenklassen zu U in G.*

Restklassen Weitere Beispiele für endliche Gruppen und Untergruppen ergeben sich aus den Restklassen zu einer natürlichen Zahl d ungleich 0. Eine Restklasse enthält dabei alle ganzen Zahlen, die bei Division durch d denselben Rest ergeben. Wir werden sehen, dass sich die Addition und Multiplikation auf den ganzen Zahlen in natürlicher Weise auf die Restklassen übetragen.

100 Lemma. *Es sei d eine natürliche Zahl ungleich 0. Dann gibt es für jede ganze Zahl n eindeutig bestimmte ganze Zahlen t und r für die gilt*

$$n = td + r \qquad \text{und} \qquad 0 \leq r \leq d - 1. \tag{2.2}$$

Beweis. Zu einer gegebenen ganzen Zahl n, sei t die größte ganze Zahl mit $td \leq n$. Nach Wahl von t gilt $n < (t+1)d$, das heißt $n \leq td + (d-1)$ und somit $n = td + r$ für ein r mit $0 \leq r \leq d-1$. Diese Darstellung ist eindeutig. Für jede weiter Darstellung $n = t'd + r'$ mit t' in $\mathbb{Z}$ und $0 \leq r' \leq d-1$ folgt $(t - t')d = r - r'$. Im Fall $t \neq t'$ würden sich die Reste r und r' um mindestens d unterscheiden, was ihrer Wahl in $\{0, \ldots, d-1\}$ widerspricht. Also gilt $t = t'$ und damit auch $r = r'$. □

101 Definition. *Es sei d eine natürliche Zahl ungleich 0 und n sei eine ganze Zahll. Die eindeutig bestimmte Zahl r, so dass (2.2) für ein t in $\mathbb{Z}$ gilt, heißt* Rest *von n bei Division durch d. Ein solcher Rest wird auch als* Rest von n modulo d*, als n* modulo d *oder kurz als $n \bmod d$, bezeichnet. Zwei ganze Zahlen sind* kongruent modulo d*, kurz*

$$n \equiv m \mod d,$$

falls sie den gleichen Rest bei Division durch n haben.

Zu $d = 5$ haben zum Beispiel -10, 0, 5 und 25 den Rest 0 sowie -28, -3, 2 und 42 den Rest 2 und es gilt zum Beispiel $-3 \equiv 2 \mod 5$. Für jede natürliche Zahl d ungleich 0 wird durch

$$s \sim_d t \quad \text{genau dann, wenn} \quad s \equiv t \mod d$$

eine Äquivalenzrelation $\sim_d$ auf $\mathbb{Z}$ definiert. Die Äquivalenzklassen der Relation $\sim_d$ werden als RESTKLASSEN MODULO d bezeichnet und wir schreiben $[n]_d$ für die Äquivalenzklasse von n. Die Menge der Äquivalenzklassen von $\sim_d$ ist

$$\mathbb{Z}_d \;=\; \{[0]_d, [1]_d, \ldots, [d-1]_d\},$$

da ja zum Beispiel $[-d+2]_d = [2]_d = [d+2]_d = [2d+2]_d$ gilt.

102 Bemerkung. *Für zwei ganze Zahlen m und n gilt genau dann $m \sim_d n$, wenn d die Differenz $m - n$ teilt.*

Nach Definition gilt $m \sim_d n$ genau dann, wenn die eindeutigen Darstellungen von m und n gemäß Lemma 100 denselben Rest r haben, also genau dann, wenn es ganze Zahlen t_m und t_n sowie ein r mit $0 \leq r \leq d-1$ gibt mit

$$m = t_m \cdot d + r \qquad \textit{und} \qquad n = t_n \cdot d + r.$$

Dies ist genau dann der Fall, wenn die Differenz $m - n$ durch d teilbar ist, auf den einfachen formalen Beweis verzichten wir.

Die Addition und Multiplikation auf den ganzen Zahlen induzieren auf den Restklasen modulo d gemäß dem folgenden Satz kanonisch zwei Verküpfungen $+$ und $\cdot$ die als ADDITION und MULTIPLIKATION von Restklassen bezeichnet werden.

103 Satz. *Es sei d eine natürliche Zahl. Dann werden auf $\mathbb{Z}_d$, der Menge der Restklassen modulo d, durch*

$$[m]_d + [n]_d \;=\; [m+n]_d \qquad \textit{und} \qquad [m]_d \cdot [n]_d \;=\; [m \cdot n]_d$$

zwei Verknüpfungen $+$ und $\cdot$ definiert.

Beweis. Wir müssen zeigen, dass die Verknüpfung $+$ und $\cdot$ auf den Restklassen wohldefiniert sind. Seien dazu m, m', n und n' beliebige ganze Zahlen mit $[m] = [m']$ und $[n] = [n']$. Um zu beweisen, dass die Restklassenaddition und -multiplikation wohldefiniert sind, genügt es dann zu zeigen, dass

$$\text{(i)} \quad [m+n] = [m'+n'] \qquad \text{und} \qquad \text{(ii)} \quad [m \cdot n] = [m' \cdot n']$$

gilt. Nach Voraussetzung und Bemerkung 102 sind $m - m'$ und $n - n'$ Vielfache von d. Damit ist aber auch

$$(m+n) - (m'+n') = (m-m') + (n-n') \tag{2.3}$$

ein Vielfaches von d und mit Bemerkung 102 folgt (i). Im Fall der Multiplikation folgt, dass

$$mn - m'n' = mn + mn' - mn' - m'n' = (m - m')n' + m(n - n')$$

ein Vielfaches von d ist, und mit Bemerkung 102 folgt (ii). □

104 Satz. *Für jede natürliche Zahl d ist $(\mathbb{Z}_d, +)$ eine kommutative Gruppe.*

Beweis. Wir müssen zeigen, dass $(\mathbb{Z}_d, +)$ die Bedingungen G_1 bis G_4 erfüllt. Dabei gelten G1 und G4, weil die Addition auf den ganzen Zahlen assoziativ und kommutativ ist und sich dies wie folgt auf die Addition der Restklassen überträgt. Für alle ganzen Zahlen m, n und q gilt

$$\begin{aligned}([m]_d + [n]_d) + [q]_d &= [m + n]_d + [q]_d \\ &= [(m + n) + q]_d \\ &= [m + (n + q)]_d \\ &= [m]_d + [(n + q)]_d = [m]_d + ([n]_d + [q]_d)\end{aligned}$$

und weiter $[m]_d + [n]_d = [m + n]_d = [n + m]_d = [n]_d + [m]_d$.

Die Restklasse $[0]_d$ ist ein neutrales Element der Addition, somit gilt G2. Es gilt auch G3, das inverse Element zu einer Restklasse $[m]$ mit $0 \leq m \leq d-1$ ist die Restklasse von $d - m$: es gilt $[d - m]_d + [m]_d = [d]_d = [0]_d$. □

Für die Multiplikation von Restklassen modulo d ist die Restklasse $[1]_d$ ein neutrales Element. Zur Restklasse $[0]_d$ gibt es aber kein Inverses, so dass sich auf diese Weise keine Gruppe ergibt. Wird die Restklasse $[0]_d$ weggelassen, ergibt sich im Allgemeinen das Problem, dass die Multiplikation aus der verbliebenen Menge hinausführt, zum Beispiel ist für $d = 6$ das Produkt der Restklassen $[2]_d$ und $[3]_d$ gleich $[0]_d$. Dieses Problem tritt nicht auf, falls d eine Primzahl ist, und genau in diesem Fall bilden die von $[0]_d$ verschiedenen Restklassen mit der Multiplikation als Verküpfung einer Gruppe.

105 Satz. *Für jede Primzahl p ist $(\mathbb{Z}_p \setminus \{[0]_p\}, \cdot)$ eine kommutative Gruppe.*

Beweis. Wir werden im Folgenden ohne Beweis benutzen, dass eine Primzahl das Produkt ab von ganzen Zahlen a und b genau dann teilt, wenn die Primzahl a oder b teilt. Weiter verwenden wir, dass die Menge $\mathbb{Z}_p \setminus \{[0]_p\}$ nach Definition aus den $p - 1$ Restklassen $[1]_p, \ldots, [p-1]_p$ besteht.

Wir zeigen zunächst, dass die Multiplikation eine Verknüpfung auf $\mathbb{Z}_p \setminus \{[0]_p\}$ ist, also nicht aus dieser Menge herausführt. Seien $[m]_p$ und $[n]_p$ Restklassen, die beide ungleich $[0]_p$ sind, das heißt, weder m noch n ist ein Vielfaches der Primzahl p. Dann ist auch mn kein Vielfaches von p, folglich ist das Produkt $[m]_p[n]_p = [mn]_p$ ungleich $[0]_p$.

Die Restklasse $[1]_d$ ist ein neutrales Element der Multiplikation. Somit bleibt noch die Existenz der inversen Elemente nachzuweisen. Wir zeigen dazu, dass für jede Restklasse $[a]_p$ ungleich $[0]_p$ die Abbildung $[x]_p \mapsto [x]_p[a]_p$ injektiv und damit, als Abbildung einer endlichen Menge in sich selbst, auch surjektiv ist. Folglich gibt es eine Restklasse $[x]_p$ mit $[x]_p[a]_p = [1]_p$.

Für einen Beweis durch Widerspruch nehmen wir an, dass es eine Restklasse $[a]_p$ ungleich $[0]_p$ gibt, so dass die Abbildung $[x]_p \mapsto [x]_p[a]_p$ nicht injektiv ist. Dann gibt es zwei voneinander verschiedene Restklassen mit demselben Funktionswert, für die wir Repräsentanten m und n mit $1 \leq n < m \leq p-1$ wählen können. Es gilt also $[m]_p[a]_p = [n]_p[a]_p$ und somit

$$[ma]_p = [na]_p, \quad \text{das heißt,} \quad ma - na = (m-n)a \text{ ist durch } p \text{ teilbar.}$$

Dies ist ein Widerspruch: nach Wahl von m und n liegt $m-n$ zwischen 1 und $p-2$, wird also nicht durch p geteilt, und letztere gilt auch für a, da die Restklasse $[a]_p$ nach Annahmne ungleich $[0]_d$ ist. □

106 Bemerkung. *Für jede natürliche Zahl d hat $(\mathbb{Z}_d, \cdot)$ alle Eigenschaften einer kommutativen Gruppe ausgenommen die Existenz von inversen Elementen, insbesondere hat die Restklasse von 0 kein Inverses. Der Beweis dieser Aussage lässt sich ganz ähnlich wie beim Nachweis dieser Eigenschaften im Beweis von Satz 105 führen.*

Exkurs über den Chinesischen Restsatz und verteilte Datensicherung Der Chinesische Restsatz sagt etwas darüber aus, inwieweit eine Zahl d bereits dadurch festgelegt ist, dass für gegebene Zahlen n_1 bis n_k jeweils der Rest von d beim Teilen durch n_j vorgegeben ist. Dazu sei eine RESTEKOMBINATION zu natürliche Zahlen $n_1, \ldots, n_k$ ein Vektor

$$(r_1, \ldots, r_k) \in \bigtimes_{j=1,\ldots,k} \{0, \ldots, n_j - 1\}.$$

Eine ganze Zahl d ERFÜLLT eine solche Restekombination, wenn gilt

$$d \equiv r_j \mod n_j \qquad \text{für } j = 1, \ldots, k.$$

Chinesischer Restsatz. *Es seien $n_1, \ldots, n_k$ paarweise teilerfremde natürliche Zahlen und es sei $n = n_1 \cdot \cdots \cdot n_k$. Dann wird jede Restekombination zu diesen Zahlen von genau einer natürlichen Zahl d in $\{0, \ldots, n-1\}$ erfüllt.*

Beweis. Wir zeigen eine etwas allgemeinere Aussage als die des Chinesischen Restsatzes. Sei dazu I eine zusammenhängede Teilmenge der ganzen Zahlen der Größe n, das heißt, I hat die Form $\{z, z+1, \ldots, z+n-1\}$ für eine ganze Zahl z. Weiter sei g die Abbildung von I in die Menge der Restekombinationen zu $n_1, \ldots, n_k$, welche einer Zahl d die Restekombination zuordnet,

die von d erfüllt wird. Wir zeigen, dass g bijektiv ist und somit jede Restekombination von genau einer Zahl in I erfüllt wird. Der Satz folgt dann als Spezialfall für $I = \{0, \ldots, n-1\}$.

Die Menge der Restekombinationen ist das kartesische Produkt von Mengen der Größe n_1 bis n_k und hat folglich dieselbe Größe n wie die Menge I. Als Abbildung zwischen endlichen Mengen derselben Größe ist g genau dann bijektiv, wenn g injektiv ist, es genügt also, Letzteres zu zeigen. Dazu sei $g(d) = g(d')$ für zwei beliebige Zahlen $d < d'$ in I. Für $j = 1, \ldots, k$ gilt dann

$$d \equiv d' \mod n_j, \qquad \text{somit ist } n_j \text{ Teiler von } d' - d$$

gemäß Bemerkung 102. Da die n_j paarweise teilerfremd sind, ist auch deren Produkt n Teiler von $d' - d$. Letztere Differenz muss dann gleich 0 und d gleich d' sein, da der Abstand von je zwei Zahlen in I echt kleiner als n ist. Folglich is g injektiv. □

Bemerkung (Verteilte Datenspeicherung)**.** *Eine Datei soll so auf k Festplatten verteilt werden, dass es für ein festes $r \leq k$ genügt, auf eine beliebige Auswahl von r dieser Festplatten zuzugreifen, um die ursprüngliche Datei rekonstruieren zu können.*

Die Datei bestehe aus einer Folge von Bits, das heißt, von Werten 0 oder 1, die in zusammenhängende Blöcke von je l Bits aufgeteilt ist. Die Aufteilung auf die k Festplatten soll für jeden Block separat durchgeführt werden. Für einen Block gibt es $n = 2^l$ Möglichkeiten seine l Bits zu wählen, entsprechend können wir einen aufzuteilenden Block in natürlicher Weise mit einer Zahl d im Bereich von 0 bis $n-1$ identifizieren.

Für die Aufteilung wählen wir paarweise teilerfremde Zahlen $n_1 < \cdots < n_k$ mit $n \leq n_1 \cdot \cdots \cdot n_r$. Für $j = 1, \ldots, k$ setzen wir $d_j = d \mod n_j$ und speichern d_j auf der j-ten Festplatte. Sind dann r paarweise verschiedene Werte $d_{i_1}, \ldots, d_{i_r}$ gegeben, so wird nach dem Chinesischen Restsatz die zugehörige Restekombination von genau einer natürlichen Zahl echt kleiner $n_{i_1} \cdot \cdots \cdot n_{i_r}$ erfüllt. Diese Zahl ist gleich d nach Wahl der d_j und wegen

$$d < n \leq n_1 \cdot \cdots \cdot n_r \leq n_{i_1} \cdot \cdots \cdot n_{i_r}.$$

Die oben beschriebene Aufteilung führt für $r < k$ zu einer redudanten Speicherung. Zum Beispiel ist für $k = 2r$ die Information über einen aufgeteilten Block sowohl in den ersten r wie auch in den zweiten r Werten d_j enthalten. Werden die Zahlen n_j etwa gleich groß und dabei möglichst klein gewählt, so ist n_j ungefähr gleich der r-ten Wurzel aus n, es gilt somit näherungsweise

$$\log n_j = \log \sqrt[r]{n} = \log n^{\frac{1}{r}} = \frac{\log n}{r} = \frac{l}{r},$$

dabei steht log *für den Logarithmus zur Basis 2. Es sind also etwa $\frac{l}{r}$ bits nötig, um einen der Werte d_j abzuspeichern. Dies ist näherungsweise optimal,*

da ja jeweils r der Werte d_j einen Wert d und damit einen Block aus l Bits darstellen.

Es folgt ein Beispiel für die verteilte Datenspeicherung wie in Bemerkung 22.1.

Beispiel. *Wir wollen einen Block von $l = 8$ Bits so auf $k = 3$ Festplatten verteilen, dass es genügt, auf $r = 2$ beliebige dieser Festplatten zuzugreifen, um den Block zu rekonstruieren. Wir wählen k paarweise teilerfremde Zahlen $n_1 = 16$, $n_2 = 17$ und $n_3 = 19$, für $n = 2^l = 256$ gilt $n \leq n_1 \cdot n_2$. Der aufzuteilende Blockinhalt wird mit einer Zahl d im Bereich 0 bis $n - 1$ identifiziert, für diese berechnen wir $d_j = d \mod n_j$ für $j = 1, 2, 3$ und speichern d_j auf der j-ten Festplatte. Ist zum Beispiel d gleich 100, so gilt $d_1 = 4$, $d_2 = 15$, und $d_3 = 5$. Gemäß dem Chinesischen Restsatz lässt sich aus je zwei dieser Werte d konstruieren, zum Beispiel ist d die eindeutige natürliche Zahl echt kleiner 256, die beim Teilen durch n_1 und n_3 den Rest d_1 beziehungsweise d_3 ergibt.*

2.2 Körper

107 Definition. *Ein* KÖRPER *ist ein Tripel $(K, +, \cdot)$ aus einer Menge K und zwei zweistelligen Funktionen $+$ und $\cdot$ auf K, so dass die drei folgenden Bedingungen erfüllt sind.*

> *(Additive Gruppe) Die Struktur $(K, +)$ ist eine kommutative Gruppe.*
>
> *(Multiplikative Gruppe) Die Struktur $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe, dabei ist 0 das neutrale Element der Gruppe $(K, +)$.*
>
> *(Distributivgesetz) Es gilt $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$ für alle x, y, und z in K.*

Für einen gegebenen Körper $(K, +, \cdot)$ wird die Verknüpfung $+$ als KÖRPERADDITION *oder kurz als Addition bezeichnet und $(K, +)$ als* ADDITIVE GRUPPE. *Die Verknüpfung $+$ wird als* KÖRPERMULTIPLIKATION *oder kurz als Multiplikation bezeichnet und $(K \setminus \{0\}, \cdot)$ als* MULTIPLIKATIVE GRUPPE.

108 Beispiel. *Die Strukturen $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper. Die Gruppeneigenschaften, dass also beispielsweise $(\mathbb{Q}, +)$ und $(\mathbb{Q}\setminus\{0\}, \cdot)$ kommutative Gruppen sind, hatten wir schon gesehen. Das Distributivgesetz gilt in $\mathbb{R}$ und damit auch in $\mathbb{Q}$, auf den Beweis verzichten wir.*

109 Vereinbarung. *Für einen Körper $(K, +, \cdot)$ werden wir auch die Menge K als Körper bezeichnen und einfach von der Addition und Multiplikation in K sprechen, wenn aus dem Zusammenhang klar oder irrelevant ist, welche Verknüpfungen gemeint sind. Im Zusammenhang mit einem Körper K*

sind mit $+$ und $\cdot$ immer die Körperaddition beziehungsweise -multiplikation gemeint, mit 0 und 1 das neutrale Element und mit $-a$ und a^{-1} das inverse Element zu a der additiven beziehungsweise multiplikativen Gruppe.

Das Verknüpfungsymbol für die Körpermultiplikation kann weggelassen werden, für $x \cdot y$ schreiben wir auch xy. Die Körpermultiplikation soll stärker binden als die Addition, so dass sich zum Beispiel das Distributivgesetz auch als $x(y+z) = xy + xz$ schreiben lässt.

Die additive und die multiplikative Gruppe eines Körpers sind durch das Distributivgesetz, anschaulich gesprochen, gekoppelt. Der folgende Satz zeigt, dass das neutrale Element 0 der additiven Gruppe besondere Eigenschaften bezüglich der Multiplikation hat.

110 Satz. *In einem Körper K ist ein Produkt xy genau dann gleich 0, wenn x oder y gleich 0 sind. Insbesondere ist 1 nicht nur in der multiplikativen Gruppe $K\setminus\{0\}$ neutrales Element, sondern auf ganz K, da $1\cdot 0 = 0\cdot 1 = 0$ gilt.*

Beweis. Sei zunächst einer der Faktoren x und y gleich 0. Da die Multiplikation kommutativ ist, gilt $xy = yx$, es genügt also den Fall y gleich 0 zu betrachten und $x \cdot 0 = 0$ zu zeigen. Es gilt

$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0,$$

und durch Addition des additiven Inversen $-(x \cdot 0)$ von $x \cdot 0$ zum ersten und letzten Term der Gleichungskette ergibt sich $0 = x \cdot 0$.

Für die Implikation in die andere Richtung zeigen wir deren Kontraposition: falls x und y beide ungleich 0 sind, dann gilt dies auch für das Produkt xy. Nach Definition des Begriffs Körper ist $(K\setminus\{0\}, \cdot)$ eine Gruppe, insbesondere ist die Multiplikation eine Verküpfung auf der Menge $K\setminus\{0\}$, somit ist diese Menge unter Multiplikation abgeschlossen. □

111 Beispiel. *Für jede Primzahl p bilden die Restklassen modulo p mit der Restklassenaddition $+$ und -multiplikation $\cdot$ einen Körper $(\mathbb{Z}_p, +, \cdot)$ mit p Elementen.*

Wir hatten schon gesehen, dass für p prim $(\mathbb{Z}_p, +)$ und $(\mathbb{Z}_p \setminus \{0\}, \cdot)$ kommutative Gruppen sind. Es bleibt zu zeigen, dass das Distributivgesetz gilt. Dieses überträgt sich für alle natürlichen Zahlen d ungleich 0 von $\mathbb{Z}$ auf $\mathbb{Z}_d$ da gemäß der Definition der Restklassenaddition und -multiplikation für alle

ganzen Zahlen m, n und q gilt

$$\begin{aligned}[m]_d \cdot ([n]_d + [q]_d) &= [m]_d \cdot [n+q]_d \\ &= [m \cdot (n+q)]_d \\ &= [m \cdot n + m \cdot q]_d \\ &= [m \cdot n]_d + [m \cdot q]_d \\ &= [m]_d \cdot [n]_d + [m]_d \cdot [q]_d.\end{aligned}$$

112 Bemerkung. *Für jede Primzahl p und jede natürliche Zahl $k \geq 1$ gibt es einen Körper mit p^k Elementen, dieser ist eindeutig bestimmt und wird als $\mathbb{F}_{p^k}$ bezeichnet. Im Fall $k = 1$ ist dieser Körper gleich $(\mathbb{Z}_p, +, .)$.*

Im Fall $k > 1$ ist der Körper gleich $(\mathbb{Z}_p^k, +, \cdot)$ für geeignet gewählte Verküpfungen $+$ und $\cdot$. Die Elemente des Körpers sind also k-Tupel über $\mathbb{Z}_p$, die Komponenten eines solche Tupels stehen dabei für die Koeffizienten eines Polynoms vom Grad echt kleiner k und die Addition ist gleich der komponentenweise Addition in $\mathbb{Z}_p$. Die Multiplikation entspricht der Multiplikation solcher Polynome modulo eines geeigneten Polynoms vom Grad k. Die komponentenweise Multiplikation in $\mathbb{Z}_p$ als Multiplikation würde dagegen keinen Körper ergeben, da für Elemente mit Komponenten der Form $[0]_p$ kein Inverses bezüglich der Multiplikativion existiert.

113 Beispiel. *Im Körper $\mathbb{F}_2$ gilt $1 + 1 = 0$. Für $x_1, \ldots, x_t$ in $\mathbb{F}_2$ ist die Summe $x_1 + x_2 + \cdots + x_t$ genau dann gleich 0, wenn eine gerade Anzahl der Summanden x_1 bis x_t gleich 1 ist und das Produkt $x_1 \cdot x_2 \cdot \cdots \cdot x_t$ ist genau dann gleich 0, wenn einer der Faktoren x_1 bis x_t gleich 0 ist.*

3 Vektorräume

3.1 Vektorräume

In diesem Abschnitt werden Vektorräume behandelt. Ein Vektorraum besteht, grob gesagt, aus einer Menge von Vektoren, die bezüglich der Addition von Vektoren eine Gruppe bilden und einem Körper, mit dessen Elementen die Vektoren multipliziert werden können. Für einen gegebenen Vektorraum wird dabei das Symbol + sowohl für die Körper- wie auch für die Vektoraddition verwendet, und das Symbol $\cdot$ sowohl für die Körpermultiplikation wie auch für die Multiplikation von Vektoren mit einem Körperelement. Es wird dabei immer aus dem Kontext klar sein, welche Verküpfung gemeint ist.

114 Definition. *Es sei $(K, +, \cdot)$ ein Körper. Ein* Vektorraum über dem Körper K*, kurz: ein K-*Vektorraum*, ist ein Paar aus einer kommutativen Gruppe $(V, +)$ und einer in Infixnotation als Verknüpfung geschriebenen Abbildung $\cdot\colon K \times V \to V$, die als* skalare Multiplikation *bezeichnet wird, so dass für alle a und b in K und alle u und v in V die folgenden Bedingungen, die* Vektorraumgesetze*, gelten*

$$a \cdot (b \cdot u) = (a \cdot b) \cdot u \qquad \textit{(Assoziativgesetz der skalaren Multiplikation)},$$

$$(a + b) \cdot u = (a \cdot u) + (b \cdot u) \qquad \textit{(Distributivgesetz der Skalaraddition)},$$

$$a \cdot (u + v) = (a \cdot u) + (a \cdot v) \qquad \textit{(Distributivgesetz der Vektoraddition)},$$

$$1 \cdot u = u \qquad \textit{(Neutrales Element der skalaren Multiplikation)},$$

dabei ist 1 das neutrale Element der Multiplikation im Körper K.

Die Elemente des Körpers K werden als Skalare *bezeichnet, die Elemente der Menge V als* Vektoren*, die Verküpfung $+$ auf V als* Vektoraddition*. Ein* reeller Vektorraum *ist ein Vektorraum über dem Körper $\mathbb{R}$.*

115 Vereinbarung. *Für einen Vektorraum $(V, +)$ wird auch die Menge V selbst als Vektorraum bezeichnet. Auch ohne dass für einen Vektorraum zum Beispiel der zugehörige Körper K explizit eingeführt wurde, sprechen wir einfach von der Körperaddition $+$, der Körpermultiplikation $\cdot$, der Vektoraddition $+$ und der skalaren Multiplikation $\cdot$ wenn aus dem Zusammenhang klar oder irrelevant ist, welche Verknüpfungen gemeint sind.*

Im Zusammenhang mit einem K-Vektorraum steht V immer für die Menge der Vektoren und o für das neutrale Element der Vektoraddition. Für die additive Gruppe von K steht 0 für das neutrale Element und $-a$ für das inverse Element zu einem Körperelement a. Für die multiplikative Gruppe steht 1 für das neutrale Element sowie a^{-1} oder $\frac{1}{a}$ für das inverse Element von a.

Das Verknüpfungsymbol $\cdot$ kann immer weggelassen werden, unabhängig davon, ob es für die Körpermultiplikation oder die skalare Multiplikation steht. Die Körpermultiplikation und die skalare Multiplikation sollen stärker binden als die Körper- und die Vektoraddition, so dass sich zum Beispiel das Distributivgesetz der skalaren Multiplikation auch als $(a + b)u = au + bu$ schreiben lässt.

Nach Satz 110 sind die additive und multiplikative Gruppe eines Köpers über das Distributivgesetz gekoppelt. Ganz ähnlich sind in einem K-Vektorraum der Vektorraum selbst und der Körper K über die skalare Multiplikation und die Vektorraumgesetze gekoppelt.

116 Satz. *Es sei K ein Körper und $(V, +)$ ein K-Vektorraum.*

(i) Es sei a in K und u in V. Das Produkt $a \cdot u$ ist genau dann gleich dem Nullvektor o, wenn $a = 0$ oder $u = o$ ist.

(ii) Für alle u in V ist das Inverse $-u$ von u bezüglich der Vektoraddition gleich $(-1) \cdot u$.

Dabei sind wie üblich 0 und 1 das neutralen Element der Köperaddition beziehungsweise -multiplikation und -1 ist das Inverse von 1 bezüglich der Köperaddition.

Beweis. Wir zeigen die Äquivalenz (i) und wählen dazu a in K und u in V beliebig. Um die Rückwärtsrichtung der Äquivalenz zu zeigen, beobachten wir zunächst, dass ein Vektor v für den die Gleichung $v = v + v$ gilt, gleich dem Nullvektor sein muss. Dies folgt indem auf beiden Seiten der Gleichung v subtrahiert wird, das heißt, der inverse Vektor $-v$ addiert wird. Die Gleichung selbst gilt sowohl für $v = 0 \cdot u$ wie auch für $v = a \cdot o$, da aus den Vektorraumgesetzen folgt

$$0 \cdot u = (0 + 0)u = 0 \cdot u + 0 \cdot u \qquad \text{und} \qquad a \cdot o = a \cdot (o + o) = a \cdot o + a \cdot o,$$

das heißt, falls $a = 0$ oder $u = o$ ist, folgt $a \cdot u = o$.

Um die Vorwärtsrichtung der Äquivalenz (i) zu zeigen, nehmen wir nun an, dass $a \cdot u$ gleich dem Nullvektor ist. Falls a gleich 0 ist, sind wir fertig. Wir können also annehmen, dass a ungleich 0 ist und somit bezüglich der Körpermultiplikation ein Inverses a^{-1} besitzt. Aus den Vektorraumgesetzen und der schon bewiesenen Vorwärtsrichtung von (i) folgt dann

$$u = 1 \cdot u = (a^{-1} \cdot a) \cdot u = a^{-1} \cdot (a \cdot u) = a^{-1} \cdot o = o.$$

Aussage (ii) folgt aus den Vektorraumgesetzen und (i): für alle u in V gilt

$$u + (-1)u = 1u + (-1)u = (1 + (-1))u = 0u = o. \qquad \square$$

117 Beispiel (Die Vektorräume $\mathbb{R}^2$ und $\mathbb{R}^3$)**.** *Das Standarbeispiel eines reellen Vektorraums ist die Menge $V = \mathbb{R}^3$ mit komponentenweise ausgeführter Vektoraddition und skalarer Multiplikation, das heißt, für alle Vektoren (x, y, z) und (x', y', z') in V und alle reellen Zahl a gilt*

$$(x, y, z)+(x', y', z') \;=\; (x+x', y+y', z+z') \quad \textit{und} \quad a(x, y, z) = (ax, ay, az).$$

Die Menge V wird auch als EUKLIDISCHER RAUM *bezeichnet, ihre Elemente als Punkte. Der Punkt $(0, 0, 0)$ heißt* URSPRUNG *und wird als o geschrieben, da er gleich dem neutralen Element der Vektoraddition ist.*

Eine analoge Konstruktionen für $V = \mathbb{R}^2$ ergibt einen Vektorraum, der als EUKLIDISCHE EBENE *bezeichnet wird.*

118 Beispiel (Verschiebungen)**.** *Eine* VERSCHIEBUNG*, genauer: eine Verschiebung des euklidischen Raums, ist eine Abbildung des euklidischen Raums, bei der anschaulich gesprochen jeder Punkt parallel in dieselbe Richtung um dieselbe Entfernung verschoben wird, eine solche Abbildung wird auch als Translation bezeichnet. Eine Verschiebung t ist bereits festgelegt, wenn für einen beliebigen Punkt p dessen Bild $t(p)$ bekannt ist, es genügt also, ein Paar der Form $(p, t(p))$ anzugeben. Für Punkte p_1 und p_2 des euklidischen Raums wird das Paar (p_1, p_2) auch als* PFEIL *bezeichnet und als $\overrightarrow{p_1\, p_2}$ geschrieben.*

Die Menge V der Verschiebungen wird zu einem reellen Vektorraum, indem Vektoraddition und skalare Multiplikation wie folgt erklärt werden. Die Summe zweier Verschiebungen ist gleich deren Hintereinanderausführung. Für die skalare Multiplikation einer Verschiebung mit einer reellen Zahl a wird, anschaulich gesprochen, die Länge der Pfeile, die eine Verschiebung darstellen, mit dem Faktor a multipliziert, falls a negativ ist, wird zusätzlich die Richtung der Pfeile „umgedreht“.

119 Bemerkung. *Eine Verschiebung t ist insbesonder durch den Punkt $t(o)$ oder das Paar $(o, t(o))$ festgelegt, dabei steht o für den Ursprung, also den Punkt $(0, 0, 0)$, ein Paar der Form (o, p) wird auch als* ORTSPFEIL *bezeichnet. Die Abbildung $\pi\colon t \mapsto t(o)$ ist eine Bijektion von der Menge der Verschiebungen nach $\mathbb{R}^3$. Bezeichnen wir die Operationen im Vektorraum der Verschiebungen wie gewohnt mit $+$ und $\cdot$ und die Operationen in $\mathbb{R}^3$ mit $+^{\mathbb{R}^3}$ und $\cdot^{\mathbb{R}^3}$, so gilt für alle Verschiebungen t und t' und alle reellen Zahlen a*

$$(t + t')(o) = t(o) +^{\mathbb{R}^3} t'(o) \qquad \textit{und} \qquad (a \cdot t)(o) = a \cdot^{\mathbb{R}^3} t(o),$$

Mit der Abbildung π lässt sich dies auch als

$$\pi(t + t') = \pi(t) +^{\mathbb{R}^3} \pi(t') \qquad \textit{und} \qquad \pi(a \cdot t) = a \cdot^{\mathbb{R}^3} \pi(t),$$

schreiben. Das heißt, die Zuordnung einer Verschiebung t zur zugehörigen reellen Zahl $\pi(t)$ ist mit der Vektoraddition beziehungsweise skalaren Multiplikation vertauschbar. Die Abbildung π heißt ISOMORPHIE *zwischen den beiden*

Vektorräumen, die Vektorräume selbst heißen isomorph. Anschaulich gesprochen bedeutet dies, dass die beiden Vektorräume bis auf die Bezeichung ihrer Elemente identisch sind: zum Beispiel gilt im Vektorraum der Verschiebungen genau dann $t+t'=t''$, wenn im Vektorraum $\mathbb{R}^3$ gilt $\pi(t)+\pi(t')=\pi(t'')$.

Aus der Isomorphie der beiden Vektorräume folgt auch, dass die Vektoraddition im Vektorraum der Verschiebungen kommutativ ist, weil dies in $\mathbb{R}^3$ der Fall ist, während die Hintereinanderausführung von beliebigen Funktionen im Allgemeinen nicht kommutativ ist. Da in beiden Vektorräumen das neutrale Element und die inversen Elemente bereits durch die Vektoraddition festgelegt sind, folgt weiter

$$\pi(o)=0 \qquad \textit{und} \qquad \pi(-t)=-\pi(t),$$

das heißt, das Bild unter π des neutralen Elements ist das neutrale Element und das Bild des Inversen der Verschiebung t ist das inverse der Verschiebung $\pi(t)$.

120 Bemerkung. *Zwei Pfeile im euklidischen Raum heißen* ÄQUIVALENT*, wenn sie parallel sind und dieselbe Richtung und Länge haben. Die Äquivalenz von Pfeilen ist eine Äquivalenzrelation, die Äquivalenzklassen heißen* VERSCHIEBUNGSVEKTOREN*. Wir schreiben $[(p_1,p_2)]$ für die Äquivalenzklasse eines Pfeils (p_1,p_2). Die Abbildung*

$$\pi\colon t\mapsto[(o,t(o)]=\{(x,t(x))\colon x\in\mathbb{R}^3\}$$

ist eine Bijektion von der Menge der Verschiebungen in die Menge der Verschiebungsvektoren.

Die Menge V der Verschiebungsvektoren wird zu einem reellen Vektorraum und die Bijektion π wird zu einem Isomorphismus zwischen Vektorräumen, indem die Vektoraddition und skalare Multiplikation im Vektorraum der Verschiebungen in der natürlichen Weise auf Verschiebungsvektoren übertragen werden. Das heißt, für alle Verschiebungen t und t' und alle reellen Zahlen a soll gelten

$$[(o,t(o))]+[(o,t'(o))]=[(o,t(o)+t'(o))] \quad \textit{und} \quad a\cdot[(o,t(o))]=[(o,a\cdot t(o))].$$

Wir verzichten auf den Beweis, dass die so definierte Addition und skalare Multiplikation wohldefiniert ist, also Pfeile im selben Verschiebungsvektor auf Pfeile in ein und demselben Verschiebungsvektor abgebildet werden.

121 Beispiel (Der Vektorraum K^n)**.** *Die Konstruktion des Vektorraums $\mathbb{R}^3$ lässt sich auf beliebige Körper K und auf von 3 verschiedene Dimensionen übertragen. Dazu sei für einen Körper K und eine natürliche Zahl n die Menge V gleich dem n-fachen kartesischen Produkt K^n von K mit sich selbst.*

Wir erhalten einen K-Vektorraum, indem wir die Vektoraddition und die skalare Multiplikation als komponentenweise Verknüpfungen definieren:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),$$
$$a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n).$$

122 Beispiel (Der Vektorraum der unendlichen Folgen über K)**.** *Zu einem Körper K sei V die Menge aller unendlichen Folgen über K, wobei wir eine solche Folge im Zusammenhang mit Vektorräumen als $(a_1, a_2, \ldots)$ schreiben, es ist also*

$$V = \{(a_1, a_2, \ldots) \colon a_i \in K \text{ für alle } i \geq 1\}.$$

Die Menge V wird durch komponentenweise Definition der Addition und skalaren Multiplikation zu einem Vektorraum, wir definieren also

$$(a_1, a_2, \ldots) + (b_1, b_2, \ldots) = (a_1 + b_1, a_2 + b_2, \ldots)$$
$$a(a_1, a_2, \ldots) = (aa_1, aa_2, \ldots).$$

Die Menge U aller unendlichen Folgen über K, bei denen nur endlich viele Folgenglieder ungleich 0 sind, ist nach dem Untervektorraumkriterium ein Untervektorraum von V und damit insbesondere ein Vektorraum.

Die Folgen in U können mit Polynomen über K identifiziert werden, das sind Abbildungen von K nach K der Form $x \mapsto a_n \cdot x^n + \cdot + a_1 \cdot x + a_0$ mit a_i in K und a_n ungleich 0. Die a_i heißen Koeffizienten des Polynoms, n heißt Grad des Polynoms. Zum Beispiel entspricht das Polynom $ax^2 - b$ der Folge $(-b, 0, a, 0, 0, \ldots)$.

Untervektorräume

123 Definition (Untervektorraum)**.** *Es sei K ein Köper und V sei ein K-Vektorraum. Eine Teilmenge U von V ist ein* UNTERVEKTORRAUM *von V, falls U mit der auf U eingeschränkten Vektoraddition und skalaren Multipikation ein K-Vektorraum ist.*

Ein Vektorraum V ist nach Definition eine Gruppe bezüglich der Vektoraddition und enthält somit den Nullvektor o, das neutrale Element der Vektoraddition. Jeder Vektorraum enthält als TRIVIALE UNTERVEKTORRÄUME sich selbst und den NULLVEKTORRAUM $\{o\}$.

124 Satz (Untervektorraumkriterium)**.** *Eine Teilmenge U eines Vektorraums V ist genau dann ein Untervektorraum von V, falls U nicht leer und unter Vektoraddition und skalarer Multiplikation abgeschlossen ist.*

Beweis. Es sei K ein Körper und $(V, +)$ sei ein K-Vektorraum. Ist U Untervektorraum von V, so ist U insbesondere ein Vektorraum und somit nichtleer,

außerdem ist U unter Vektoraddition und skalarer Multiplikation abgeschlossen. Um die Implikation in die umgekehrte Richtung zu zeigen, sei U eine nichtleere Teilmenge von V für welche die beiden letzteren Abschlusseigenschaften gelten. Wir zeigen mit dem Untergruppenkriterium aus Satz 95, das $(U, +)$ eine Untergruppe von $(V, +)$ und damit eine Gruppe ist. Da U nach Voraussetzung nichtleer und gegen die Gruppenoperation Vektoraddition abgeschlossen ist, muss dazu nur noch gezeigt werden, dass U mit einem Vektor u auch immer dessen Inverses $-u$ unter der Vektoraddition enthält. Dies folgt sofort aus Satz 116, da U nach Voraussetzung unter skalarer Multiplikation abgeschlossen ist und folglich mit u immer auch $-u = (-1)u$ enthält. Es bleiben nur noch die Vektorraumgesetze zu zeigen. Diese übertragen sich von V auf U. Zum Beispiel gilt das Distributivgesetz der skalaren Multiplikation auch in U: für alle Skalare a und b und alle u in U gilt $(a + b)u = au + bu$, weil diese Gleichung sogar für alle u in V gilt. □

3.2 Linearkombinationen und Erzeugnis

Linearkombinationen

125 Definition. *Sei V ein K-Vektorraum. Für eine endliche Folge $v_1, \ldots, v_n$ mit $n \geq 1$ von Vektoren aus V heißt jeder Vektor der Form*

$$a_1 v_1 + \cdots + a_n v_n \qquad \text{mit } a_i \text{ in } K \tag{3.1}$$

LINEARKOMBINATION *der Vektoren v_1 bis v_n*

Für eine Menge X von Vektoren heißt ein Vektor LINEARKOMBINATION ÜBER X*, falls der Vektor Linearkombination einer endlichen Folge von Vektoren aus X ist.*

Nach Definition ist für eine Linearkombination wie in (3.1) nicht gefordert, dass die Vektoren v_1 bis v_n paarweise verschieden sind. Letzteres kann immer erreicht werden, indem jeweils alle Summanden zum selben Vektoren zusammengefasst werden, also zum Beispiel $au + bv + cu$ in $(a + c)u + bv$ umgeformt wird.

126 Bemerkung. *Der Einfachheit halber wird der Begriff Linearkombination im Folgenden in zwei verschiedenen Bedeutungen verwendet, wobei sich jeweils aus dem Zusammenhang ergeben wird, welche der beiden Bedeutungen gemeint ist. Zum einen ist eine Linearkombination gemäß Definition 125 ein Vektor. Zum anderen werden wir, ohne dies formal zu definieren, den Begriff Linearkombination auch im Sinne einer Darstellung eines Vektors verwenden, eine solche Darstellung ist dann durch die endliche Folge $v_1, \ldots, v_n$ der beteiligten Vektoren und die zugehörigen Koeffizienten $a_1, \ldots, a_n$ gegeben. Damit lässt sich dann zum Beispiel formulieren, dass solche Darstellungen*

unter bestimmten Voraussetzungen in dem Sinne eindeutig sind, dass je zwei Linearkombinationen desselben Vektors in den beteiligten Vektoren und deren Koeffizienten übereinstimmen.

Erzeugnis

127 Definition (Erzeugnis)**.** *Es sei V ein K-Vektorraum. Das* Erzeugnis *einer nichtleeren Teilmenge V_0 von V, kurz: $\langle V_0 \rangle$, ist die Menge alle Vektoren, die sich als Linearkombination von Vektoren aus V_0 darstellen lassen, das heißt*

$$\langle V_0 \rangle \;=\; \{a_1 v_1 + \cdots + a_n v_n \colon n \in \mathbb{N} \setminus \{0\} \text{ und } a_1, \ldots, a_n \in K, v_1, \ldots, v_n \in V_0\}.$$

Das Erzeugnis einer endlichen Menge $\{v_1, \ldots, v_n\}$ wird als $\langle v_1, \ldots, v_n \rangle$ geschrieben, es gilt dann

$$\langle v_1, \ldots, v_n \rangle \;=\; \{a_1 v_1 + \cdots + a_n v_n \colon a_1, \ldots, a_n \in K\}.$$

Das Erzeungis der leeren Menge wird gleich der Einermenge des Nullvektors gesetzt, das heißt, es gilt $\langle \emptyset \rangle = \{o\}$.

Das Erzeugnis einer Menge enthält nach Definition immer den Nullvektor, ist also nicht leer. Tatsächlich ist das Erzeugnis einer Teilmenge eines Vektorraums V immer ein Untervektorraum von V.

128 Satz. *Es sei V_0 eine Teilmenge eines K-Vektorraums V. Dann ist das Erzeugnis $\langle V_0 \rangle$ von V_0 ein Untervektorraum von V.*

Beweisskizze. Falls V_0 leer ist, ist das Erzeugnis von V_0 gleich dem Nullvektorraum und somit ein Untervektorraum von V. Sei also V_0 nichtleer, das Erzeugnis von V_0 enthält dann gerade die Linearkombinationen von Vektoren aus V_0. Es genügt zu zeigen, dass die Bedingungen des Untervektorraumkriteriums erfüllt sind. Das Erzeugnis enthält den Nullvektor, ist also nicht leer. Weiter ist das Erzeugnis $\langle V_0 \rangle$ unter Vektoraddition und unter skalarer Multiplikation abgeschlossen: die Summe von Linearkombinationen über V_0 und ebenso das Produkt eines Skalars mit einer Linearkombination über V_0 ist wieder eine Linearkombination über V_0. Sind beispielsweise $a_1 u_1 + \cdots a_m u_m$ und $b_1 v_1 + \cdots b_n v_n$ Linearkombinationen, so sind deren Summe und das Produkt eines Skalars a mit der ersteren Linearkombination gleich

$$a_1 u_1 + \cdots + a_m u_m \;+\; b_1 v_1 + \cdots + b_n v_n \quad \text{beziehungsweise} \quad a a_1 u_1 + \cdots a a_m u_m$$

und damit wieder Linearkombinationen über V_0. Im Fall der Summe folgt dies, da die in einer Linearkombination auftretenden Vektoren nicht paarweise verschieden sein müssen. □

3.3 Lineare Unabhängigkeit und Basen

Linear abhängige und linear unabhängige Vektoren Der Nullvektor kann als Linearkombination beliebiger Vektoren dargestellt werden, indem deren Koeffizienten alle gleich 0 gesetzt werden.

129 Definition. *Wird der Nullvektor als Linearkombination von Vektoren dargestellt, so heißt diese Linearkombination* TRIVIAL*, wenn die Koeffizienten der beteiligten Vektoren alle gleich* 0 *sind, sonst* NICHTTRIVIAL*.*

130 Definition. *Es sei V ein K-Vektorraum. Die Vektoren $v_1, \ldots, v_n$ aus V mit $n \geq 1$ heißen*

> LINEAR ABHÄNGIG*, falls sich der Nullvektor als nichttriviale Linearkombination dieser Vektoren darstellen lässt, das heißt, es gibt Koeffizienten $a_1, \ldots, a_n$, die nicht alle gleich* 0 *sind, mit $a_1v_1 + \cdots + a_nv_n = o$,*
>
> LINEAR UNABHÄNGIG*, falls diese Vektoren nicht linear abhängig sind, das heißt, falls sich der Nullvektor mit diesen Vektoren nur als triviale Linearkombination darstellen lässt: aus $a_1v_1 + \cdots + a_nv_n = o$ mit $a_1, \ldots, a_n$ in K folgt immer, dass alle a_i gleich* 0 *sind.*

Eine Teilmenge X von V ist

> LINEAR ABHÄNGIG*, falls es in X endlich viele paarweise verschiedene Vektoren gibt, die linear abhängig sind,*
>
> LINEAR UNABHÄNGIG*, falls die Menge X nicht linear abhängig ist, das heißt, endlich viele paarweise verschiedene Vektoren in X sind immer linear unabhängig.*

Es ist sinnvoll, in der Definition von linear abhängiger Menge nur paarweise verschiedene Vektoren zu betrachten, da jeder Vektor v die folgende nichttriviale Linearkombination des Nullvektors erlaubt

$$1 \cdot v + (-1) \cdot v \;=\; (1 + (-1))v = 0 \cdot v = o.$$

Nach Definition ist die leere Menge linear unabhängig. Eine Menge, die genau einen Vektor enthält ist genau dann linear abhängig, wenn dieser Vektor gleich dem Nullvektor ist: nach Definition ist ein einzelner Vektor v linear abhängig, falls $av = o$ für einen Skalar a ungleich 0 gilt, dies ist nach Satz 116 genau dann der Fall, wenn v der Nullvektor ist.

131 Satz (Charakterisierung der linearen Abhängigkeit)**.** *Für $n \geq 2$ sind Vektoren $v_1, \ldots, v_n$ genau dann linear abhängig, wenn es eine Index Index i*

gibt, so dass sich v_i als Linearkombination der restlichen Vektoren schreiben lässt, das heißt, wenn es Koeffizienten a_j gibt mit

$$v_i \;=\; \sum_{j\neq i} a_j v_j. \tag{3.2}$$

Beweis. Zunächst gelte Gleichung (3.2). Es ergibt sich eine Linearkombination des Nullvektors indem auf beiden Seiten der Gleichung der Vektor v_i abgezogen wird, diese ist nichttrivial, da der Vektor v_i den Koeffizienten -1 hat, die Vektoren $v_1, \ldots, v_n$ sind also linear abhängig.

Sind umgekehrt die Vektoren $v_1, \ldots, v_n$ linear abhängig, so gibt es nach Definition der linearen Abhängigkeit Koeffizienten $a_1, \ldots, a_n$, die nicht alle gleich 0 sind, so dass gilt

$$a_1 v_1 + \cdots + a_n v_n \;=\; o.$$

Wir wählen einen Index i mit a_i ungleich 0, insbesondere hat der Skalar a_i ein multiplikatives Inverses $\frac{1}{a_i}$. Es gilt dann

$$a_i v_i \;=\; \sum_{j\neq i} -a_j v_j \qquad \text{und somit} \qquad v_i \;=\; \sum_{j\neq i} \frac{-a_j}{a_i} v_j. \qquad \square$$

132 Satz (Koeffizientenvergleich)**.** *Für $n \geq 1$ seien $v_1, \ldots, v_n$ linear unabhängige Vektoren in einem K-Vektorraum. Gilt dann für Koeffizienten $a_1, \ldots, a_n, b_1, \ldots, b_n$ aus K die Gleichung*

$$a_1 v_1 + \cdots + a_n v_n \;=\; b_1 v_1 + \cdots + b_n v_n, \tag{3.3}$$

so müssen die Koeffizienten der Vektoren v_i auf beiden Seiten jeweils gleich sein, das heißt, es gilt, $a_i = b_i$ für $i = 1, \ldots, n$.

Beweis. Gemäß den Vektorraumgesetzen ist (3.3) äquivalent zu

$$(a_1 - b_1) v_1 + \cdots + (a_n - b_n) v_n \;=\; 0.$$

Da die Vektoren $v_1, \ldots, v_n$ linear unabhänging sind, müssen die Koeffizienten $a_i - b_i$ alle gleich 0 sein, es folgt $a_i = b_i$ für $i = 1, \ldots, n$. $\square$

Erzeugendensystem und Basis

133 Definition (Erzeugendensystem und Basis)**.** *Es sei V ein Vektorraum. Eine Teilmenge von V ist ein* Erzeugendensystem *von V, falls das Erzeugnis der Menge gleich V ist. Eine* Basis *von V ist ein linear unabhängiges Erzeugendensystem von V.*

Nach Definition ist die leere Menge linear unabhängig und ihr Erzeugnis ist gleich dem Nullvektorraum $\{o\}$, die leere Menge ist somit eine Basis des Nullvektorraums.

134 Beispiel. *Zu einem Körper K betrachten wir den K-Vektorraum K^n. Die kanonische Basis dieses Vektorraums ist die Menge der* EINHEITSVEKTOREN

$$e_1 = (1, 0, \ldots, 0), \qquad e_2 = (0, 1, \ldots, 0), \qquad \ldots, \qquad e_n = (0, 0, \ldots, 1).$$

135 Beispiel. *Zu einem Körper K betrachten wir den K-Vektorraum der unendlichen Folgen über K, bei denen nur endlich viele Folgenglieder ungleich 0 sind. Die kanonische Basis dieses Vektorraums ist die abzählbar unendliche Menge der Vektoren*

$$e_1 = 1, 0, 0, \ldots, \qquad e_2 = 0, 1, 0, \ldots, \qquad e_3 = 0, 0, 1, \ldots, \qquad \ldots \quad .$$

136 Satz (Eindeutig der Darstellung durch eine endliche Basis)**.** *Sei V ein K-Vektorraum mit einer endlichen Basis $B = \{v_1, \ldots, v_n\}$ der Größe $n \geq 1$. Dann kann jeder Vektor in V in der Form $a_1 v_1 + \ldots + a_n v_n$ als Linearkombination der Basisvektoren dargestellt werden, die Koeffizienten a_1 bis a_n aus K sind dabei eindeutig bestimmt.*

Beweis. Die Basis B ist nach Definition ein Erzeugendensystem von V, folglich lässt sich jeder Vektor in V als Linearkombination der Basisvektoren darstellen. Die Koeffizienten a_1 bis a_n sind dabei eindeutig, dies ergibt sich unmittelbar aus dem Satz über den Koeffizientvergleich, da die Vektoren der Basis B nach Definition linear unabhängig sind. □

Für unendliche Basen sind Darstellung durch eine Linearkombination aus paarweise verschiedenen Basisvektoren insofern eindeutig, als sich zwei solche Linearkombinationen desselben Vektors nur in der Reihenfolge der Summanden und in Summanden mit Koeffizienten 0 unterscheiden können. Durch diesen schwächeren Begriff von Eindeutigkeit wird zum Beispiel berücksichtigt, dass für eine Linearkombination $a_1 v_1 + \cdots + a_n v_n$ der dargestellte Vektor für jeden von den Basisvektoren v_1 bis v_n verschiedenen Basisvektor v auch durch die Linearkombination $a_1 v_1 + \cdots + a_n v_n + 0v$ dargestellt wird.

137 Satz (Eindeutig der Darstellung durch eine Basis)**.** *Es sei B eine Basis eines Vektorraums V. Jeder Vektor in V lässt sich als Linearkombination paarweise verschiedenener Vektoren in B darstellen. Eine solche Darstellung ist eindeutig bis auf die Reihenfolge der Summanden und bis auf Summanden mit Koeffizienten 0.*

Beweisskizze. Da die Basis B nach Definition ein Erzeugendensystem von V ist, lässt sich jeder Vektor als Linearkombination von Vektoren aus B darstellen. Es sei nun angenommen, dass sich ein Vektor w in V auf die beiden

folgenden Weisen als Linearkombinationen von jeweils paarweise verschiedenen Basisvektoren darstellen lässt

$$w \;=\; a_1u_1 + \ldots + a_mu_m \;=\; b_1v_1 + \ldots + b_nv_n. \tag{3.4}$$

Für $E_u = \{u_1, \ldots, u_m\}$ und $E_v = \{v_1, \ldots, v_n\}$ ergeben sich daraus die beiden folgenden Linearkombination von w über $E_u \cup E_v$ aus jeweils paarweise verschiedenen Vektoren

$$\begin{aligned} w \;&=\; a_1u_1 + \ldots + a_mu_m + \sum_{v \in E_v \setminus E_u} 0v \\ &=\; b_1v_1 + \ldots + b_nv_n \;+ \sum_{u \in E_u \setminus E_v} 0u. \end{aligned}$$

Die Menge $E_u \cup E_v$ ist als Teilmenge der Basis B linear unabhängig. Nach dem Satz über Koeffizientenvergleich sind die Koeffizienten der Vektoren in $E_v \setminus E_u$ und in $E_u \setminus E_v$ in beiden Linearkombinationen identisch und somit gleich 0. In beiden Linearkombinationen von w in (3.4) können somit nur Vektoren in $E_u \cap E_v$ Koeffizienten ungleich 0 haben. Durch Streichen von Summanden mit Koeffizienten 0 ergeben sich zwei Linearkombinationen von w über der Menge $E_u \cap E_v$, so dass in beiden Linearkombinationen jeder Vektor in $E_u \cap E_v$ jeweils genau einmal vorkommt. Nach dem Satz über Koeffizientenvergleich müssen die Koeffizienten von jeweils identischen Vektoren in diesen beiden Darstellungen übereinstimmen. Dies bedeutet aber gerade, dass in den beiden Linearkombinationen in (3.4), bis auf eventuell einige Vektoren aus $E_v \setminus E_u$ und $E_u \setminus E_v$ mit Koeffizienten 0, dieselben Vektoren mit jeweils denselben Koeffizienten vorkommen, das sind gerade die Vektoren in $E_u \cap E_v$. Die beiden Linearkombinationen in (3.4) unterscheiden sich also höchstens in der Reihenfolge der Summanden und in Summanden mit Koeffizienten 0. □

138 Definition. *Es sei X eine Teilmenge eines Vektorraums.*

Die Menge X ist MAXIMAL LINEAR UNABHÄNGIG, *falls X, aber keine echte Obermenge von X linear unabhängig ist.*

Die Menge X ist ein MININMALES ERZEUGENDENSYSTEM, *falls X, aber keine echte Teilmenge von X ein Erzeugendensystem ist.*

Die in Definition 138 eingeführten Begriffe sind Spezialfälle von zwei in der folgenden Bemerkung behandelten Sprechweisen.

139 Bemerkung. *Zu einer Menge A sei eine Eigenschaft von Teilmengen von A gegeben, formal ist diese Eigenschaft also eine Teilmenge der Potenzmenge von A. Eine Teilmenge X von A ist* MAXIMAL MIT DIESER EIGENSCHAFT, *falls X, aber keine echte Obermenge von X die Eigenschaft hat. Eine Teilmenge X von A ist* MINIMAL MIT DIESER EIGENSCHAFT, *falls X, aber keine echte Teilmenge von X die Eigenschaft hat.*

140 Bemerkung. *Sei X eine Teilmenge eines Vektorraums V. Dann ist X maximal linear unabhängig, falls X linear unabhängig ist, aber letztere Eigenschaft für keine Menge der Form $X \cup \{v\}$ für ein $v \notin X$ gilt. Weiter ist X ein minimales Erzeugendensystem von V, falls X ein Erzeugendensystem von V ist, letztere Eigenschaft aber verlorengeht, falls aus X ein beliebiges Element entfernt wird.*

Die linear abhängigen Teilmengen eines Vektorraums sind unter Inklusion nach oben abgeschlossen: für eine linear abhängig Menge ist auch jede Obermenge dieser Menge linear abhängig. Sind also für eine linear unabhängige Menge X alle Mengen der Form $X \cup \{v\}$ mit $v \notin X$ linear abhängig, dann sind schon alle Obermenge von X linear abhängig, somit ist X maximal linear unabhängig. Auch die die Eigenschaft, Erzeugendensysteme eines Vektorraums zu sein, ist unter Inklusion nach oben abgeschlossen. Daraus folgt, dass die Eigenschaft, kein Erzeugendensystem zu sein unter Inklusion nach unten abgeschlossen ist. Sind also für ein Erzeugendensystem X alle Teilmengen von X der Form $X \setminus \{v\}$ keine Erzeugendensysteme, dann gilt Letzteres schon für alle echten Teilmengen von X.

141 Lemma. *Es sei V ein Vektorraum und X sei eine linear unabhängige Teilmenge von V. Die Menge X ist genau dann maximal linear unabhängig, wenn X ein Erzeugendensystem von V ist.*

Beweis. Falls X leer ist, gilt die Behauptung, da dann die Eigenschaften, dass X maximal linear unabhängig ist und dass X ein Erzeugendensystem ist beide dazu äquivalent sind, dass V der Nullvektorraum ist. Wir können also ab jetzt annehmen, dass X nichtleer ist.

Sei zunächst X maximal linear unabhängig und sei v ein beliebiger Vektor aus V. Falls v in X ist, liegt v im Erzeugnis von X. Im Fall $v \notin X$ ist die Menge $X \cup \{v\}$ linear abhängig, dies bedeutet nach Definition, dass es paarweise und von v verschiedene Vektoren $v_1, \ldots, v_n$ in X gibt, so dass

$$a_1 v_1 + \cdots + a_n v_n + a_{n+1} v \;=\; o.$$

eine nichttriviale Linearkombination des Nullvektors ist. Da die Menge X linear unabhängig ist, muss der Koeffizient a_{n+1} ungleich 0 sein, folglich gilt

$$v \;=\; \tfrac{-a_1}{a_{n+1}} v_n + \cdots + \tfrac{-a_n}{a_{n+1}} v_n.$$

Somit ist X ein Erzeugendensystem von V, da v ein beliebiger Vektor aus v ist. Für den Beweis der Implikation in die anderer Richtung sei X ein Erzeugendensystem von V und v sei ein beliebiger Vektor v in V, der nicht in X ist. Dann lässt sich v als Linearkombination von paarweise verschiedenen Vektoren $v_1, \ldots, v_n$ aus X darstellen, es gilt

$$v = a_1 v_1 + \cdots + a_n v_n, \quad \text{und somit} \quad o = a_1 v_1 + \cdots + a_n v_n - v,$$

wobei nach Wahl von v auch die Vektoren $v_1, \ldots, v_n, v$ paarweise verschieden sind. Folglich ist die Menge $X \cup \{v\}$ linear abhängig. Da v als beliebiger Vektor gewählt war, der nicht in X ist, folgt mit Bemerkung 140, dass X maximal linear unabhängig ist. □

142 Lemma. *Es sei V ein Vektorraum und die Teilmenge X von V sei ein Erzeugendensystem von V. Die Mengen X ist genau dann ein minimales Erzeugendensystem von V, wenn X linear unabhängig ist.*

Beweis. Der Nullvektorraum hat genau ein minimales Erzeugendensystem und genau eine linear unabhängige Teilmenge, diese sind jeweils gleich der leeren Menge, die behauptete Äquivalenz gilt also in diesem Fall. Wir können somit im Weiteren annehmen, dass V nicht der Nullvektorraum ist und insbesondere alle Erzeugendensysteme von V nichtleer sind.

Die zu beweisende Aussage ist logisch äquivalent zur Aussage, dass das Erzeugendensystem X von V genau dann nicht minimal ist, wenn X linear abhängig ist. Um letztere Äquivalenz zu beweisen, nehmen wir zunächst an, dass X nicht minimal ist, das heißt, für einen Vektor v aus X ist auch die Menge $X \setminus \{v\}$ ein Erzeugendensystem und nach Annahme insbesondere nichtleer. Der Vektor v kann dann als Linearkombination der anderen Vektoren in X dargestellt werden, die Menge X ist also linear abhängig.

Für den Beweis der anderen Richtung der Äquivalenz sei nun X linear abhängig, es gibt also eine nichttriviale Linearkombination

$$a_1 v_n + \cdots + a_n v_n \;=\; o.$$

des Nullvektors mit paarweise verschiedenen Vektoren $v_1, \ldots, v_n$ aus X. Im Fall, dass einer der Vektoren v_i gleich dem Nullvektor ist, ist auch $X \setminus \{v_i\}$ ein Erzeugendensystem von V, das Erzeugendensystem X ist also nicht minimal. Andernfalls muss $n \geq 2$ gelten und wir können für einen beliebig gewählten Index i mit a_i ungleich 0 den Vektor v_i in der Form

$$v_i \;=\; \sum_{j \neq i} \frac{-a_j}{a_i} v_j$$

schreiben. Damit ist auch $X \setminus \{v_i\}$ ein Erzeugendensystem von V, das Erzeugendensystem X ist folglich nicht minimal. □

143 Satz (Basisergänzungssatz und Basisauswahlsatz)**.** *Für eine Teilmenge X eines Vektorraums V sind die folgenden Aussagen äquivalent.*

(i) Die Menge X ist eine Basis von V.

(ii) Die Menge X ist maximal linear unabhängig.

(iii) Die Menge X ist ein minimales Erzeugendensystem.

Beweis. Wir zeigen zunächst, dass die Aussagen (i) und (ii) äquivalent sind. Jeder der beiden Aussagen impliziert, dass X linear unabhängig ist. Es genügt also zu zeigen, dass eine linear unabhängige Menge genau dann eine Basis ist, wenn sie maximal linear unabhängig ist. Dies folgt sofort aus Lemma 141, da eine lineare unabhängige Menge genau dann eine Basis ist, wenn sie ein Erzeugendensystem ist.

Wir zeigen weiter, dass die Aussagen (i) und (iii) äquivalent sind. Jeder der beiden Aussagen impliziert, dass X ein Erzeugendensystem von V ist. Mit dieser Voraussetzung folgt aus Lemma 142, dass X genau dann ein minimales Erzeugendensystem ist, wenn X linear unabhängig und damit als Erzeugendensystem auch eine Basis ist. □

In Satz 143 wird die Äquivalenz von (i) und (ii) als Basisergänzungssatz bezeichnet, die von (ii) und (iii) als Basisauswahlsatz. Nach dem Basisergänzungssatz lassen sich – unter Erhaltung der linearen Unabhänigkeit – zu einer gegebenenen linear unabhänigen Menge solange sukzessive Vektoren hinzufügen, bis eine maximal linear unabhängige Menge und damit eine Basis erreicht wird. Nach dem Basisauswahlsatz lassen sich – unter Erhaltung der Eigenschaft Erzeugendensystem von V zu sein – aus einem gegebenenen Erzeugendensystem von V solange sukzessive Vektoren streichen, bis eine minimales Erzeugendensystem von V und damit eine Basis erreicht wird.

Aus dem Basisergänzungssatz lässt sich folgern, dass jeder Vektorraum eine Basis hat. Im allgemeine Fall ist letztere Aussage äquivalent zum Auswahlaxiom der Mengenlehre und in ihrem Beweis wird das ebenfalls zum Auswahlaxiom äquivalente Zornsche Lemma verwendet.[3] Für ENDLICH ERZEUGTE Vektorräume, das sind Vektorräume mit einem endlichen Erzeugendensystem, folgt die Existenz von Basen dagegen in einfacher Weise aus dem Basisauswahlsatz.

144 Satz. *Jeder Vektorraum hat eine Basis. Jeder endlich erzeugte Vektorraum hat eine endliche Basis.*

Beweis. Wir verzichten auf den aufwändigen Beweis des allgemeinen Falls und zeigen nur die Aussage für endlich erzeugte Vektorräume. Für einen solchen Vektorraum V können wir ein endliches Erzeugendensystem E wählen. Falls dieses Erzeugendensystem nicht minimal ist, können wir einen Vektor aus E entfernen und erhalten ein echt kleineres Erzeugendensystem von V. Wir wiederholen das Entfernen eines Elements solange das aktuell betrachtete Erzeugendensystem noch nicht minimal ist. Da wir mit einem endlichen

[3] Das Zornsche Lemma besagt, dass es in einer partiell geordneten Menge ein maximales Element gibt, falls jede Kette, also jede total geordnete Teilmenge, eine obere Schranke hat. Ein Element ist dabei maximal, falls kein anderes Element echt größer ist.

Erzeugendensystem beginnen, erreichen wir nach endlich vielen Wiederholungen ein minimales Erzeugendensystem, welches nach dem Basisauswahlsatz eine Basis ist. □

Da im Wesentlichen alle hier relevanten Vektorräume endlich erzeugt sind, werden wir uns im Folgenden auf diesen Spezialfall beschränken.

145 Vereinbarung. *Im Folgenden seien alle Vektorräume endlich erzeugt, falls nicht ausdrücklich etwas anderes gesagt wird.*

3.4 Dimension eines Vektorraums

Der Austauschsatz von Steinitz

146 Austauschlemma. *Sei B Basis eines Vektorraums V und sei u ein Vektor aus V ungleich dem Nullvektor. Dann gibt es einen Vektor v in B, so dass $(B \setminus \{v\}) \cup \{u\}$ eine Basis von V ist. Für v kann jeder Vektor gewählt werden, der in einer Darstellung von u als Linearkombination von paarweise verschiedenen Vektoren aus B mit einem Koeffizienten ungleich 0 vorkommt.*

Beweis. Wir können im Beweis annehmen, dass u nicht in B ist, da andernfalls die Behauptung für $v = u$ gilt. Die Basis B ist ein Erzeugendensystem, folglich kann u in der Form

$$u \;=\; a_1 v_1 + \cdots + a_n v_n \tag{3.5}$$

als Linearkombination von paarweise verschiedenen Vektoren aus B dargestellt werden. Da u ungleich dem Nullvektor ist, gibt es Indizes j, so dass a_j ungleich 0 ist. Wir wählen einen beliebigen solchen Index t und zeigen, dass die Menge

$$B_t = (B \setminus \{v_t\}) \cup \{u\}$$

eine Basis von V ist. Der Koeffizient a_t ist ungleich 0 und hat somt ein multiplikatives Inverses $\frac{1}{a_t}$. Durch Umstellen von (3.5) erhalten wir

$$v_t \;=\; \tfrac{1}{a_t} u - \sum_{j \neq t} \frac{a_j}{a_t} v_j.$$

Somit ist v_t eine Linearkombination über B_t und folglich sind die Erzeugnisse der Mengen B_t und $B_t \cup \{v_t\}$ gleich. Tatsächlich sind beide Mengen Erzeugendensysteme von V, da die letztere Menge das Erzeugendensystem B enthält, es gilt

$$B = (B \setminus \{v_t\}) \cup \{v_t\} \subseteq (B \setminus \{v_t\}) \cup \{u\} \cup \{v_t\} = B_t \cup \{v_t\}.$$

Um zu zeigen, dass das Erzeugendensystem B_t eine Basis ist, genügt es zu zeigen, dass B_t linear unabhängig ist. Für einen Beweis durch Widerspruch

nehmen wir an, dass B_t linear abhängig ist und es somt eine nichttriviale Linearkombination

$$o = bu + b_1u_1 + \cdots + b_mu_m + bu \tag{3.6}$$

des Nullvektors gibt, so dass die Vektoren $u_1, \ldots, u_m, u$ aus B_t und paarweise verschiedenen sind. Dabei muss der Koeffizient b ungleich 0 sein, da wir sonst eine nichttriviale Linearkombination des Nullvektors aus paarweise verschiedenen Vektoren der Menge $\{u_1, \ldots, u_m\}$ erhalten, die als Teilmenge der Basis B aber linear unabhängig ist. Ersetzen wir nun in (3.6) den Vektor u durch seine Darstellung aus (3.5), so erhalten wir folgende Linearkombination des Nullvektors

$$o = bu + \sum_{j \neq t} b_j u_j = ba_1v_1 + \cdots + ba_tv_t + \cdots + ba_nv_n + \sum_{j \neq t} b_j u_j.$$

Die v_i sind paarweise verschieden, ebenso die u_i. Falls es Indizes i und j mit v_i gleich u_j gibt, so können wir jeweils die zugehörigen Summanden zu $(ba_i + b_j)v_i$ zusammenfassen und erhalten so eine nichttriviale Linearkombination des Nullvektors aus paarweise verschiedenen Vektoren über B, im Widerspruch zur Voraussetzung, dass B eine Basis und somit linear unabhängig ist. Letztere Linearkombination ist nichttrivial, da der Summand zum Vektor v_t den Koeffizienten ba_t hat, dieser ist ungleich 0, da b und a_t beide ungleich 0 sind. In diesem Zusammenhang ist zu beachten, dass v_t ungleich u und damit nach Definition nicht in B_t ist, während die u_i alle in B_t sind, so dass der Summand ba_tv_t nicht mit einem Summanden der Form b_ju_j zusammengefasst wird. □

Im Beweis des Austauschlemmas wurde nirgends verwendet, dass die Basis B endlich ist, das Austauschlemma und der folgende Austauschsatz gelten tatsächlich für beliebige Vektorräume, nicht nur für die hier betrachteten endlich erzeugten.

147 Austauschsatz von Steinitz. *Sei B eine Basis eine Vektorraums V. Für jede endliche linear unabhängige Teilmenge E von V gibt es eine Teilmenge B_0 von B der Größe $|E|$, so dass $(B \setminus B_0) \cup E$ eine Basis von V ist. Insbesondere kann jede endliche linear unabhängige Menge von Vektoren durch Vektoren aus einer gegebenen Basis zu einer Basis ergänzt werden.*

Beweis. Wir beweisen den Austauschsatz durch Induktion über die Größe von E, simultan für alle endlichen Basen des Vektorraums V. Im Induktionsanfang sei $|E| = 0$, das heißt, E ist die leere Menge und die Behauptung ist offensichtlich richtig. Im Induktionsschritt von n nach $n+1$ sei B eine Basis von V und E sei eine linear unabhängige Teilmenge von V der Größe $n+1$. Wir wählen einen Vektor u aus E und setzen $F = E \setminus \{u\}$. Die Menge F

hat dann Größe n und ist linear unabhängig. Nach Induktionsannahme gibt es eine Teilmenge B_F von B der Größe n, so dass $B' = (B \setminus B_F) \cup F$ eine Basis von V ist. Der Vektor u kann dann als Linearkombination über B' dargestellt werden. In dieser Linearkombination muss der Koeffizient eines Vektors v aus $B \setminus B_F$ ungleich 0 sein, sonst wäre u eine Linearkombination der Vektoren in F, im Widerspruch zur Voraussetzung, dass $E = F \cup \{u\}$ linear unabhängig ist. Nach dem Austauschlemma kann also in der Basis B' der Vektor v durch u ersetzt werden. Folglich gilt die Aussage des Austauschsatzes für E und $B_0 = B_F \cup \{v\}$. □

Der Austauschsatz hat eine Reihe interessanter Folgerungen. Es sei daran erinnnert, dass alle betrachteten Vektorräume endlich erzeugt sind.

148 Korollar. *Sei B eine endliche Basis des Vektorraums V. Dann gibt es in V höchstens $|B|$ viele linear unabhängige Vektoren. Insbesondere sind linear unabhängige Mengen immer endlich.*

Beweis. Sei E eine linear unabhängige Teilmenge von V. Nach dem Austauschsatz gibt es eine Teilmenge B_0 von B der Größe $|E|$, die Menge E hat also höchstens so viele Elemente wie die Basis B. Da die hier betrachteten endlich erzeugten Vektorräume nach Satz 144 immer eine endliche Basis haben, folgt dass linear unabhängige Mengen immer endlich sind. □

149 Korollar. *Alle Basen eines Vektorraums haben dieselbe Größe.*

Beweis. Basen sind linear unabhängig und somit nach Korollar 148 endlich. Aus dem Korollar folgt somit, dass für je zwei Basen B_1 und B_2 desselben Vektorraums sowohl $|B_1| \leq |B_2|$ als auch $|B_2| \leq |B_1|$ gilt, beide Basen habe also dieselbe Anzahl von Elementen. □

150 Definition. *Die* DIMENSION *eine Vektorraums V, kurz:* $\dim(V)$*, ist die Anzahl der Elemente der Basen von V.*

Die hier betrachteten endlich erzeugten Vektorräume werden auch als ENDLICH-DIMENSIONALE Vektorräume bezeichnet. Der Nullvektorraum $\{o\}$ hat nach Definition die leere Menge als Basis und entsprechend die Dimension 0.

Auch für nicht endllich erzeugte Vektorräume lässt sich die Dimension als die Kardinalität der Basen einführen, zum Beispiel gibt es Vektorräume, deren Basen alle abzählbar unendllich sind. Entsprechend unsere Einschränkung auf endlich erzeugte Vektorräume sind die im Folgenden auftretenden Dimensionen endlich: hat ein Vektorraum Dimension d, so ist d eine natürliche Zahl.

151 Korollar. *In einem Vektorraum der Dimension d ist jede linear unabhängige Mengen mit d Elementen eine Basis des Vektorraums.*

Beweis. Ein gegebener Vektorraum der Dimension d hat nach Definition eine Basis der Größe d. Nach Korollar 148 hat dann jede linear unabhängige Menge höchstens Größe d, eine linear unabhängige Menge mit d Elementen ist also maximal linear unabhängig und somit nach dem Basisergänzungssatz eine Basis. Alternativ folgt das Korollar aus dem Austauschsatz, indem in einer Basis der Größe d alle Vektoren durch die d Vektoren der betrachteten linear unabhängigen Mengeersetzt werden. □

152 Korollar. *Die Dimension eines Untervektorraums eines Vektorraums V ist höchstens so groß wie die Dimension von V.*

Beweis. Jede Basis eines Untervektorraums von V ist linear unabhängig, enthält also höchstens $\dim(V)$ viele Elemente. □

153 Korollar. *Jede Basis eines Untervektorraums von V kann zu einer Basis von V ergänzt werden.*

Beweis. Folgt direkt aus dem Austauschsatz, da jede Basis eines Untervektorraums endlich und linear unabhängig ist. □

154 Korollar. *Für Untervektorräume U_1 und U_2 eines Vektorraums sei U_1 eine Teilmenge von U_2. Dann gilt $\dim(U_1) \leq \dim(U_2)$ und jede Basis von U_1 kann zu einer Basis von U_2 erweitert werden. Die Dimensionen der beiden Vektorräume sind genau dann gleich, wenn U_1 gleich U_2 ist.*

Beweis. Nach Definition von Untervektorraum ist U_1 Untervektorraum des Vektorraums U_2. Damit folgt sofort $\dim(U_1) \leq \dim(U_2)$ aus Korollar 152 und die Erweiterbarkeit der Basen von U_1 zu Basen von U_2 aus Korollar 153. Falls U_1 und U_2 dieselbe Dimension haben, ist jede Basis von U_1 eine maximal linear unabhängig Teilmenge von U_2 und somit insbesondere eine Basis von U_2, die beiden Untervektorräume sind also identisch. Sind umgekehrt beide Untervektorräume identisch, so haben beide dieselbe Dimension. □

Dimensionsformel für Untervektorräume Wir beschließen diesen Abschnitt mit einer Dimensionsformel für Untervektorräume. Diese ist ähnlich aufgebaut wie die Gleichung

$$|X \cup Y| \;=\; |X| + |Y| - |X \cap Y|,$$

die für die Größen $|X|$ und $|Y|$ beliebiger endlicher Mengen X und Y gilt. Durch den negativen Summanden auf der rechten Seite wird berücksichtigt, dass die Elemente im Schnitt der beiden Mengen X und Y auf der linken Seite einmal, auf der rechten Seite aber zweimal gezählt werden. Als Vorbereitung auf die Dimensionsformel machen wir uns zunächst klar, dass der Schnitt von zwei Untervektorräumen wieder ein Untervektorraum ist.

155 Satz. *Seien U_1 und U_2 Untervektorräume eines Vektorraums V. Dann ist auch der Schnittt $U_1 \cap U_2$ ein Untervektorraum von V.*

Beweis. Nach dem Untervektorraumkriterium genügt es zu zeigen, dass der Schnitt von U_1 und U_2 nichtleer und unter Vektoraddition und skalarer Multiplikation abgeschlossen ist. Zunächst ist der Schnitt nicht leer, da beide Vektorräume den Nullvektor enthalten. Weiter übertragen sich die Asbschlusseigenschaften von U_1 und U_2 auf deren Schnitt. Die Untervektorräume U_1 und U_2 sind jeweils gegen Vektoraddition abgeschlossen, für Vektoren u und v im Schnitt ist deren Summe $u + v$ folglich in U_1 und in U_2 und damit auch im Schnitt dieser beiden Untervektorräume. Der Schnitt ist auch gegen skalare Multiplikation abgeschlosssen ist: für jeden Vektor u im Schnitt ist jedes skalare Vielfache der Form au in U_1 und in U_2 und damit im Schnitt. □

156 Dimensionssatz für Untervektorräume. *Für Untervektorräume U_1 und U_2 eines Vektorraums V gilt*

$$\dim(\langle U_1 \cup U_2 \rangle) \;=\; \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

Beweis. Nach Satz 155 ist $U_1 \cap U_2$ ein Untervektorraum, dessen Dimension m ist höchstens so groß wie die des endlich-dimensionalen Vektorraum V. Sei $B_0 = \{u_1, \ldots, u_m\}$ eine Basis von $U_1 \cap U_2$. Nach Korollar 153 lässt sich B_0 für geeignete $s, t \geq 0$ zu Basen

$$B_1 = \{u_1, \ldots, u_m, v_1, \ldots, v_s\} \text{ und } B_2 = \{u_1, \ldots, u_m, w_1, \ldots, w_t\}$$

von U_1 der Größe $m + s$ beziehungsweise U_2 der Größe $m + t$ ergänzen. Wir setzen

$$B = B_1 \cup B_2 = \{u_1, \ldots, u_m, v_1, \ldots, v_s, w_1, \ldots, w_t\}$$

Die Vektoren $u_1, \ldots, u_m, v_1, \ldots, v_s, w_1, \ldots, w_t$ sind paarweise verschieden. Die $u_1, \ldots, u_m, v_1, \ldots, v_s$ sind paarweise verschieden, da sie die Menge B_1 der Größe $m+s$ bilden, entsprechend für die Vektoren $u_1, \ldots, u_m, w_1, \ldots, w_t$ und B_2. Aber auch alle v_i sind von allen w_j verschieden, da die v_i in U_1 sind, wohingegen die w_j in U_2, aber nicht im Erzeugnis $U_1 \cap U_2$ von B_0 und somit nicht in U_1 sind. Es folgt

$$|B| = (m + s) + (m + t) - m = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

Es genügt also zu zeigen, dass B eine Basis des Vektorraums $\langle U_1 \cup U_2 \rangle$ ist und dieser somit Dimension $|B|$ hat. Zunächst ist B ein Erzeugendensystem dieses Vektorraums, da nach Konstruktion jeder Vektor in U_1 und in U_2 eine Linearkombination über B ist. Es bleibt zu zeigen, dass B linear unabhängig. Für einen Beweis sei

$$a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_sv_s + c_1w_1 + \cdots + c_tw_t \;=\; o \qquad (3.7)$$

eine beliebige Linearkombination des Nullvektors über B. Dann ist der Vektor

$$a_1u_1 + \cdots + a_mu_m + b_1v_1 + \cdots + b_sv_s \;=\; -c_1w_1 + \cdots - c_tw_t$$

in U_1 und in U_2, kann also als Linearkombination über B_0 geschrieben werden. Wegen der Eindeutigkeit der Darstellung über B_0 gemäß Satz 136 sind somit alle b_i gleich 0. Ein symmetrisches Argument mit dem Vektor

$$a_1u_1 + \cdots + a_mu_m + c_1w_1 + \cdots + c_tw_t \;=\; -b_1v_1 + \cdots - b_sv_s$$

zeigt, dass auch alle c_i gleich 0 sind. Es folgt $a_1u_1 + \cdots + a_mu_m = o$, und da die Vektoren $u_1, \ldots, u_m$ linear unabhängig sind, sind auch alle a_i gleich 0. Da die Linearkombination des Nullvektors über B in (3.7) beliebig gewählt war, ist B linear unabhängig. □

3.5 Komplementäre Untervektorräume und Nebenklassen

Komplementäre Untervektorräume

157 Definition. *Zwei Untervektorräume U_1 und U_2 eines Vektorraums V sind* KOMPLEMENTÄR*, falls gilt*

$$\text{(i)}\;\; U_1 \cap U_2 = \{o\} \qquad \textit{und} \qquad \text{(ii)}\;\; \langle U_1 \cup U_2 \rangle = V.$$

Für komplementäre Untervektorräume U_1 und U_2 wird U_2 als Komplement von U_1 bezeichnet und U_1 als Komplement von U_2. Komplemente in diesem Sinn sind nicht eindeutig, wie das nächste Beispiel zeigt.

158 Beispiel. *Es sei V ein d-dimensionaler Vektorraum und U sei ein Untervektorraum von V der Dimension $d-1$, ein Untervektorraum dieser Dimension wird als* HYPEREBENE *von V bezeichnet. Für jeden Vektor v, der nicht in der Hyperebene U liegt, ist dann der von v erzeugte eindimensionale Untervektorraum komplementär zu U.*

Zum Beispiel sind im Vektorraum $\mathbb{R}^2$ je zwei Untervektorräume der Form $\langle u \rangle$ und $\langle v \rangle$ komplementär, falls v nicht Vielfaches von u, also nicht von der Form au für einen Skalar a ist. Ganz ähnlich ist für jeden zweidimensionalen Untervektorraum U des Vektorraums $\mathbb{R}^3$ und für jeden Vektor v, der nicht in der Hyperebene U liegt, der Untervektorraum $\langle v \rangle$ komplementär zu U.

159 Satz. *Zu jedem Untervektorraum gibt es einen komplementären Untervektorraum.*

Beweis. Sei U ein Untervektorraum eines Vektorraums V. Wähle eine Basis $B = \{u_1, \ldots, u_m\}$ von U, die durch eine Menge $B' = \{v_1, \ldots, v_s\}$ zu

einer Basis $B \cup B'$ von V der Größe $m + s$ ergänzt wird. Wir zeigen, dass der von B' erzeugte Untervektorraum U' komplementär zu U ist.

Die Vereinigung $U \cup U'$ enthält die Basis $B \cup B'$ von V als Teilmenge und erzeugt folglich den Vektorraum V. Es bleibt somit nur noch zu zeigen, dass der Schnitt von U_1 und U_2 nur den Nullvektor enthält. Jeder Vektor v in diesem Schnitt kann sowohl als Linearkombination über B als auch über B' in der Form

$$v \;=\; a_1u_1 + \cdots + a_mu_m \;=\; b_1v_1 + \cdots + b_sv_s$$

dargestellt werden, folglich gibt es eine Linearkombination

$$o \;=\; a_1u_1 + \cdots + a_mu_m - b_1v_1 - \cdots - b_sv_s$$

des Nulllvektors aus paarweise verschiedenen Vektoren in der linear unabhängigen Menge $B \cup B'$. Die Koeffizienten in dieser Linearkombination sind dann alle gleich 0 und v ist gleich dem Nullvektor. □

160 Satz. *Es seien U_1 und U_2 komplementäre Untervektorräume eines Vektorraums V. Dann lässt sich jeder Vektor v aus V als Summe der Form $v = u_1 + u_2$ mit eindeutig bestimmten Vektoren u_1 aus U_1 und u_2 aus U_2 schreiben.*

Beweis. Gemäß der Definition von komplementär erzeugt die Vereinigung der Untervektorräume U_1 und U_2 den Vektorraum V, jeder Vektor v aus V lässt sich also für Vektoren $v_1, \ldots, v_s$ aus U_1 und und $w_1, \ldots, w_t$ aus U_2 als Linearkombination der Form

$$v = \underbrace{a_1v_1 + \cdots + a_sv_s}_{=u_1 \in U_1} + \underbrace{b_1w_1 + \cdots + b_tw_t}_{=u_2 \in U_2} \;=\; u_1 + u_2$$

darstellen. Die Vektoren u_1 und u_2 sind dabei eindeutig bestimmt, für u_1' in U_1 und u_2' in U_1 mit

$$u_1 + u_2 = u_1' + u_2' \;\text{ liegt der Vektor }\; u_1 - u_1' \;=\; u_2' - u_2 \text{ in } U_1 \cap U_2$$

und ist somit gleich dem Nullvektor, es gilt also $u_1 = u_1'$ und $u_2 = u_2'$. □

161 Korollar. *Für komplementäre Untervektorräume U_1 und U_2 eines Vektorraums V gilt* $\dim(V) = \dim(U_1) + \dim(U_2)$.

Beweis. Für die komplementären Untervektorräume U_1 und U_2 erzeugt die Vereinigung $U_1 \cup U_2$ den Vektorraum V und der Schnitt $U_1 \cap U_2$ ist gleich dem Nullvektorraum, dieser hat Dimension 0. Die Aussage folgt somit direkt aus dem Dimensionssatz für Untervektorräume. □

Nebenklassen Die im Folgenden definierten Nebenklassen werden später in unterschiedlichen Zusammenhängen vorkommen.

162 Definition. *Es sei U ein Untervektorraum eines Vektorraums V und v sei ein Vektor in V. Die* Nebenklasse *von U durch v ist definiert als*

$$v + U \;=\; \{v + u \colon u \in U\},$$

der Vektor v heißt Repräsentant *der Nebenklasse.*

Jeder Vektor v ist in der Nebenklasse $v + U$ enthalten, da der Untervektorraum U den Nullvektor enthält.

163 Satz. *Es sei U ein Untervektorraum eines Vektorraums V. Dann bilden die Nebenklassen zu U eine Zerlegung von V.*

Beweis. Jeder Vektor v in V ist Element der Nebenklasse $v + U$, die Vereinigung aller Nebenklassen ist also gleich V. Es bleibt zu zeigen, dass je zwei Nebenklassen disjunkt oder identisch sind. Seien dazu $v+U$ und $w+U$ zwei beliebige Nebenklassen zu U. Falls die beiden Nebenklassen disjunkt sind, ist nichts zu zeigen. Andernfalls gib es u_v und u_w in U mit $v + u_v = w + u_w$. Für jeden Vektor u in U gilt dann

$$v + u = w + \underbrace{u_w - u_v + u}_{\in U} \;\in\; w + U.$$

Folglich ist $v + U$ eine Teilmenge von $w + U$, und ein sehr ähnliches symmetrisches Argument zeigt, dass auch $w + U$ Teilmenge von $v + U$ ist, die beiden Nebenklassen sind also gleich. □

Für einen Untervektorraum U eines Vektorraums V ist nach Satz 163 jeder Vektor v in V in genau einer Nebenklasse zu U. Da U den Nullvektor enthält, ist diese Nebenklasse gleich der Nebenklasse $v + U$ die durch v repräsentiert wird, eine Nebenklasse wird also genau durch ihre Elemente repräsentiert.

164 Bemerkung. *Es sei U ein Untervektorraum eines Vektorraums V. Aus Satz 163 folgt, dass die wie folgt definiert Relation $\sim_U$ auf V*

$$v \sim_u w \quad \text{g.d.w.} \quad v \text{ und } w \text{ sind in derselben Nebenklasse zu } U.$$

eine Äquivalenzrelation ist. Nach Definition sind deren Äquivalenzklassen gerade die Nebenklassen zu U, und auch die Begriffe der Repräsentanz von Nebenklassen und von Äquivalenzklassen fallen zusammen: für beide Begriffe wird eine Klasse genau durch ihre Elemente repräsentiert.

165 Bemerkung. *Es sei U ein Untervektorraum eines Vektorraums V. Für Vektoren v und w gilt genau dann $v \sim_u w$, wenn $v - w$ in U liegt. Nach Definition und Satz 163 gilt $v \sim_u w$ genau dann, wenn v in $w + U$ liegt, also genau dann, wenn $v = w + u$ für ein u in U gilt. Dies ist aber äquivalent dazu, dass $v - w$ in U liegt.*

166 Definition. *Es sei U ein Untervektorraum des Vektorraums V. Der* FAKTORRAUM VON V NACH U, *kurz: V/U, ist gleich der Menge der Nebenklassen zu U, es gilt*

$$V/U \;=\; \{v + U \colon v \in V\}.$$

167 Bemerkung. *Die Vektoraddition und skalare Multiplikation in V induzieren auf dem Faktorraum V/U vermöge*

$$[v]_U + [w]_U = [v + w]_U \qquad \textit{und} \qquad a \cdot [v]_U = [av]_U$$

eine Vektoraddition und eine skalare Multiplikation, durch die der Faktorraum zu einem Vektorraum wird. Diese beiden Verküpfungen sind wohldefiniert: das Ergebnis der Verknüpfung hängt nicht von den betrachteten Repräsentanten der beteiligten Äquivalenzklassen ab, den Beweis dieser Tatsache lassen wir aus. Für die Dimension des Faktorraums lässt sich zeigen

$$\dim(V/U) = \dim(V) - \dim(U).$$

168 Bemerkung. *Es sei U ein Untervektorraum des Vektorraums V und der Untervektorraum W von V sei komplementär zu U. Dann sind die Vektorräume V/U und W* ISOMORPH. *Dies bedeutet anschaulich, dass die beiden Vektorräume sozusagen bis auf die Umbenennung ihrer Elemente gleich sind. Formal bedeutet die Isomorphie, dass es eine Bijektion g zwischen den beiden Vektorräumen gibt, welche mit der Vektoraddition und der skalaren Multiplikation in dem Sinne verträglich ist, dass gilt*

$$g(v + w) = g(v) + g(w) \qquad \textit{und} \qquad g(av) = a \cdot g(v).$$

Eine solche Bijektion ist die Abbildung g von W nach V/U, welche einen Vektor v auf seine Äquivalenzklasse $[v]$ bezüglich der Relation $\sim_U$ abbildet, das heißt

$$g \colon v \mapsto [v]$$

Wir verzichte auf den Beweis der Bijektivität. Dass die Bijektion g mit den Verküpfungen in den beiden Vektorräumen verträglich ist, folgt aus der Definition der Verknüpfungen des Faktorraums: für alle v und w in W und alle Skalare a gilt

$$g(v+w) = [v+w] = [v]+[w] = g(v)+g(w) \;\; \textit{und} \;\; g(av) = [av] = a[v] = a{\cdot}g(v).$$

4 Anwendungen von Vektorräumen

4.1 Matrizen

In den nachfolgenden Anwendungen von Vektorräumen werden Matrizen eine wichtiges Hilfsmittel sein.

169 Definition. *Sei K ein Körper und seien m und n beide natürliche Zahlen ungleich 0. Eine (m,n)-*Matrix *über K ist eine Abbildung*

$$\{1,\dots,m\}\times\{1,\dots,n\} \;\to\; K,$$

das Paar (m,n) heißt Dimension *der Matrix. Eine solche Matrix A kann in der Form*

$$A \;=\; \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

geschrieben werden, der Eintrag $a_{i,j}$ *an* Position (i,j) *ist dabei gleich dem Funktionswert $A(i,j)$. Eine solche Matrix A kann auch als (m,n)-Matrix $(a_{i,j})$ oder Matrix $(a_{i,j})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ beschrieben werden.*

Im Zusammenhang mit Matrizen kann die Angabe des betrachteten Körpers weggelassen werden, falls dieser irrelevant ist oder sich aus dem Zusammenhang ergibt. Matrixeinträge 0 und 1 stehen immer für das neutrale Element der Addition beziehungsweise Multiplikation im betrachteten Körper.

170 Definition. *Es sei $A=(a_{i,j})$ eine (m,n)-Matrix. Die* transponierte Matrix *von A, auch: die* Transponierte *von A ist die*

$$(n,m)\text{-}Matrix \quad A^{\mathrm{T}} \;=\; (c_{i,j}) \qquad mit \qquad c_{i,j} \;=\; a_{j,i}.$$

Für jede natürliche Zahl n werden $(n,1)$-Matrizen und $(1,n)$-Matrizen als Spaltenvektor *beziehungsweise* Zeilenvektor *der* Länge n *bezeichnet.*

Beispiele für transponierte Matrizen sind

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Die Transponierte eines Spaltenvektors der Länge n ist ein Zeilenvektor der Länge n und umgekehrt. Wir sprechen auch einfach von einem Vektor, falls sich aus dem Kontext ergibt oder irrelevant ist, ob es sich um einen Spalten- oder Zeilenvektor handelt. In Textzeilen schreiben wir auch Spaltenvektoren in der Form $(a_1,\dots,a_n)$, gelegentlich auch als $(a_1,\dots,a_n)^{\mathrm{T}}$.

171 Definition. *Eine Matrix der Dimension (n, n) heißt quadratisch. In einer solchen quadratischen Matrix bilden die Positionen $(1, 1)$, $(2, 2)$ bis (n, n) die* Hauptdiagonale*, die Positionen $(n, 1)$, $(n-1, 2)$, $\ldots$, $(1, n)$ die* Nebendiagonale.

Die Einheitsmatrix *der Dimension (n, n) ist die (n, n)-Matrix mit Einträgen 1 auf der Hauptdiagonalen und Einträgen 0 an allen anderen Positionen. Die* Nullmatrix *der Dimension (m, n) ist die (m, n)-Matrix, deren Einträge alle gleich 0 sind. Eine* reelle Matrix *ist eine Matrix über dem Körper $\mathbb{R}$.*

Die Transponierte einer quadratischen Matrix ergibt sich durch Spiegelung der Einträge der Matrix an der Hauptdiagonalen.

172 Definition. *Seien $A = (a_{i,j})$ und $B = (b_{i,j})$ zwei (m, n)-Matrizen. Die* Summe *$A + B$ von A und B ist die (m, n)-Matrix*

$$A + B = (c_{i,j}) \quad \textit{mit} \quad c_{i,j} = a_{i,j} + b_{i,j}.$$

Für einen Skalar a ergibt die Multiplikation *von A mit a die (m, n)-Matrix*

$$aA = (c_{i,j}) \quad \textit{mit} \quad c_{i,j} = aa_{i,j}.$$

173 Definition. *Seien m, n und q natürliche Zahlen und sei $A = (a_{i,j})$ eine (m, n)-Matrix und $B = (b_{i,j})$ eine (n, q)-Matrix. Das* Produkt *der beiden Matrizen A und B ist die (m, q)-Matrix*

$$A \cdot B \;=\; (c_{i,j}) \quad \textit{mit} \quad c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}.$$

174 Bemerkung. *Für Matrizen A und B wie in Definition 173 ist der Eintrag an Position (i, j) des Produkts $A \cdot B$ gleich dem Produkt der i-ten Zeile von A und der j-ten Spalte von B, das heißt, gleich dem Produkt des Zeilenvektors $(a_{i,1}, \ldots, a_{i,n})$ und des Spaltenvektors $(b_{1,j}, \ldots, b_{n,j})$.*

175 Beispiel. *Gemäß der Definition der Matrixmultiplikation gilt beispielsweise für das folgende Produkt reeller Matrizen*

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 1 & 2 \end{pmatrix} \;=\; \begin{pmatrix} 1+2+3 & 2+4+6 \\ 4+5+6 & 8+10+12 \end{pmatrix} \;=\; \begin{pmatrix} 6 & 12 \\ 15 & 30 \end{pmatrix}.$$

Auch bei der Multiplikation von Matrizen kann das Verknüpfungssymbol $\cdot$ weggelassen werden, und die Multiplikation soll stärker binden als die Addition, so kann beispielsweise $(A \cdot B) + C$ als $AB + C$ geschrieben werden.

Matrizen können nur multipliziert werden, wenn die Spaltenanzahl des ersten Faktors gleich der Zeilenanzahl des zweiten Faktors ist. Das Produkt einer (m, n)-Matrix mit einer (n, q)-Matrix ergibt eine (m, q)-Matrix, in symbolischer Darstellung

$$\boxed{m}\boxed{n} \cdot \boxed{n}\boxed{q} = \boxed{m}\boxed{q} .$$

176 Bemerkung. *Das Produkt zweier Matrizen, die beide keine Nullmatrix sind, kann eine Nullmatrix sein, zum Beispiel gilt*

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Dieses Beispiel zeigt auch, dass die Matrixmultiplikation nicht kommutativ ist, es gilt

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

177 Satz. *Die Matrixmultiplikation ist in dem Sinne* LINEAR*, auch:* HOMOGEN*, in beiden Argumenten, dass für alle Skalare a, alle (m,n)-Matrizen A und alle (n,q)-Matrizen B gilt*

$$a(AB) = (aA)B = A(aB). \tag{4.1}$$

Beweisskizze. Für $A = (a_{i,j})$ und $B = (b_{i,j})$ sind die drei Matrizen in (4.1) alle (m,q)-Matrizen mit dem Eintrag $a\sum_{k=1}^{n} a_{i,k}b_{k,j}$ an Position (i,j). □

178 Satz. *Die Matrixmultiplikation ist in dem Sinne* ADDITIV *in beiden Argumenten, dass für alle (m,n)-Matrizen A und B und alle (n,q)-Matrizen C und D gilt*

$$A(C+D) = AC + AD \qquad \textit{und} \qquad (A+B)C = AC + BC.$$

Beweis. Wir beweisen die erste der beiden Gleichungen und lassen den sehr ähnlichen Beweis der zweiten aus. Dazu seien eine (m,n)-Matrix $A = (a_{i,j})$ und (n,q)-Matrizen $C = (c_{i,j})$ und $D = (d_{i,j})$ beliebig gewählt. Die Matrizen $A(C+D)$ und $AC + AD$ sind dann beide (m,q)-Matrizen und der Eintrag an Position (i,j) der ersten Matrix ist

$$(A(C+D))(i,j) = \sum_{k=1}^{n} a_{i,k}(c_{k,j} + d_{k,j}).$$

Der Eintrag an Position (i,j) der zweiten Matrix ist

$$(AC+AD)(i,j) = \left(\sum_{k=1}^{n} a_{i,k}c_{k,j}\right) + \left(\sum_{k=1}^{n} a_{i,k}d_{k,j}\right) = \sum_{k=1}^{n} a_{i,k}(c_{k,j}+d_{k,j}),$$

die beiden Matrizen sind also identisch. □

179 Satz. *Die Matrixmultiplikation ist assoziativ: für alle (m,n)-Matrizen A, (n,q)-Matrizen B und (q,r)-Matrizen C gilt $(AB)C = A(BC)$.*

Beweis. Es seien A, B und C Matrizen wie im Satz. Wir setzen $D = (d_{i,j})$ gleich der (m,q)-Matrix AB und $E = (e_{i,j})$ gleich der (n,r)-Matrix BC. Nach Definition gilt dann

$$d_{s,\ell} = \sum_{k=1}^{n} a_{s,k} b_{k,\ell} \qquad \text{und} \qquad e_{k,t} = \sum_{\ell=1}^{q} b_{k,\ell} c_{\ell,t}.$$

Die Matrix $(AB)C = DC$ ist eine (m,r)-Matrix, ihr Eintrag an der Position (s,t) ist gleich

$$((AB)C)(s,t) = \sum_{\ell=1}^{q} d_{s,\ell} c_{\ell,t} = \sum_{\ell=1}^{q} \left(\sum_{k=1}^{n} a_{s,k} b_{k,\ell} \right) c_{\ell,t} = \sum_{\ell=1}^{q} \sum_{k=1}^{n} a_{s,k} b_{k,\ell} c_{\ell,t}.$$

Die Matrix $A(BC) = AE$ ist eine (m,r)-Matrix, ihr Eintrag an der Position (s,t) ist gleich

$$(A(BC))(s,t) = \sum_{k=1}^{n} a_{s,k} e_{k,t} = \sum_{k=1}^{n} a_{s,k} \sum_{\ell=1}^{q} b_{k,\ell} c_{\ell,t} = \sum_{k=1}^{n} \sum_{\ell=1}^{q} a_{s,k} b_{k,\ell} c_{\ell,t}.$$

Die Einträge der Matrizen $(AB)C$ und $A(BC)$ und damit auch die Matrizen selbst sind folglich identisch, da die Vertauschung der beiden endlichen Summationen über k und über ℓ den Wert der Summe nicht ändert. Welche Summation zuerst ausgeführt wird, spielt keine Rolle, da jeweils einfach alle $n \cdot q$ Werte der Form $a_{s,k} b_{k,\ell} c_{\ell,t}$ aufaddiert werden. □

Da die Matrixmultiplikation assoziativ, aber nicht kommutativ ist, kommt es bei einem Produkt mit mehreren Faktoren auf die Reihenfolge der Faktoren an, aber nicht darauf, in welcher Reihenfolge Teilprodukte berechnet werden, in solchen Produkten können somit Klammern immer weggelassen werden.

Rang einer Matrix

180 Definition. *Sei $A = (a_{i,j})$ eine (m,n)-Matrix über einem Körper K. Die* Zeilenvektoren *von A sind die Vektoren der Form $(a_{i,1}, \ldots, a_{i,n})$, die* Spaltenvektoren *von A sind die Vektoren der Form $(a_{1,j}, \ldots, a_{m,j})^{\mathrm{T}}$, dabei läuft i von 1 bis m und j von 1 bis n.*

Der Zeilenrang *von A ist die Dimension des von den Zeilenvektoren von A erzeugten Untervektorraums des K-Vektoraums K^n, das heißt, der Zeilenrang ist gleich*

$$\dim(\langle \{(a_{i,1}, \ldots, a_{i,n}) \colon i = 1, \ldots, m\} \rangle).$$

Entsprechend ist der Spaltenrang *von A gleich der Dimension des von den Spaltenvektoren von A erzeugten Untervektorraums des K-Vektoraums K^m.*

181 Satz. *Der Zeilenrang einer Matrix ist gleich der maximalen Anzahl linear unabhängiger Zeilenvektoren der Matrix.*

Der Spaltenrang einer Matrix ist gleich der maximalen Anzahl linear unabhängiger Spaltenvektoren der Matrix.

Beweis. Wir zeigen die Behauptung über den Zeilenrang und lassen den im Wesentlichen identischen Beweis für den Spaltenrang aus. Es sei A eine Matrix mit m Zeilen und Zeilenrang d, das heißt, der von den Zeilenvektoren erzeugte Untervektorraum U hat die Dimension d. Nach Korollar 148 kann es in U und damit auch unter den Zeilenvektoren von A nicht mehr als d linear unabhängige Vektoren geben. Umgekehrt erhalten wir d linear unabhängige Zeilenvektoren wie folgt. Falls die Menge der Zeilenvektoren kein minimales Erzeugendensystem von U ist, können wir einen geeigneten Zeilenvektor entfernen und erhalten ein echt kleineres Erzeugendensystem von U aus Zeilenvektoren. Wir wiederholen den Übergang zu einem echt kleineren Erzeugendensystems durch das Entfernen eines geeigneten Zeilenvektors solange, wie das aktuell betrachtete Erzeugendensystem von U noch nicht minimal ist. Nach endlich vielen Wiederholungen erreichen wir ein minimales Erzeugendensystem von U. Dieses ist nach Satz 143 eine Basis von U und enthält somit $\dim(U) = d$ viele linear unabhängige Vektoren. Dieser Beweis gilt auch im Fall $d = 0$, alternativ lässt sich argumentieren, dass in diesem Fall A eine Nullmatrix ist und somit keine linear unabhängigen Zeilenvektoren hat, die Behauptung also richtig ist. □

182 Lemma. *Die Matrix A sei das Produkt BC von Matrizen B und C.*

Alle Zeilenvektoren von A können als Linearkombination der Zeilenvektoren von C dargestellt werden. Für den Zeilenvektor i von A können die Koeffizienten dieser Linearkombination gleich den Einträgen in Zeile i von B gewählt werden, in der kanonischen Reihenfolge.

Alle Spaltenvektoren von A lassen sich als Linearkombination der Spaltenvektoren von B darstellen. Für den Spaltenvektor j von A können die Koeffizienten dieser Linearkombination gleich den Einträgen in Spalte j von C gewählt werden, in der kanonischen Reihenfolge.

Beweis. Es sei $A = (a_{i,j})$ eine (m, n)-Matrix. Wegen $A = BC$ muss es eine natürliche Zahl d ungleich 0 geben, so dass $B = (b_{i,j})$ eine (m, d)-Matrix und $C = (c_{i,j})$ eine (d, n)-Matrix ist. Nach Definition der Matrixmultiplikation gilt für den Zeilenvektor i von A

$$(a_{i,1}, \ldots, a_{i,n}) = \left(\sum_{k=1}^{d} b_{i,k} c_{k,1},\ \ldots,\ \sum_{k=1}^{d} b_{i,k} c_{k,n} \right) = \sum_{k=1}^{d} b_{i,k} \left(c_{k,1},\ \ldots,\ c_{k,n} \right).$$

Zeile i von A kann also als Linearkombination der Zeilenvektoren von C dargestellt werden, mit Koeffizienten $b_{i,k}$, die wie gefordert aus Zeile i von B

kommen. Analog gilt für Spalte j von A

$$\begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{d} b_{1,k}c_{k,j} \\ \vdots \\ \sum_{k=1}^{d} b_{m,k}c_{k,j} \end{pmatrix} = \sum_{k=1}^{d} c_{k,j} \begin{pmatrix} b_{1,k} \\ \vdots \\ b_{m,k} \end{pmatrix}.$$

Spalte j von A kann also als Linearkombination der Spaltenvektoren von B dargestellt werden, mit Koeffizienten $c_{k,j}$, die wie gefordert aus Spalte j von C kommen. □

183 Lemma. *Es sei A eine (m, n)-Matrix über einem Körper K und sei d eine natürliche Zahl ungleich 0. Dann sind die folgenden Aussagen äquivalent.*

(i) *Der Zeilenrang von A ist höchstens d.*

(ii) *Der Spaltenrang von A ist höchstens d.*

(iii) *Die Matrix A kann als Produkt BC einer (m, d)-Matrix B über K und einer (d, n)-Matrix C über K geschrieben werden.*

Beweis. Wir nehmen zunächst an, dass die dritte Aussage richtig ist. Gemäß Lemma 182 können dann alle Zeilenvektoren von A als Linearkombination der d Zeilenvektoren von C dargestellt werden. Die Dimension des von den Zeilenvektoren von A erzeugten Untervektorraums und damit der Zeilenrang von A ist somit höchstens gleich d, und es folgt die erste Aussage. Da gemäß Lemma 182 alle Spaltenvektoren von A als Linearkombination der d Spaltenvektoren von B dargestellt werden können, folgt mit einem ähnlichen Argument auch die zweite Aussage.

Es bleibt zu zeigen, dass aus der ersten und zweiten Aussage jeweils die dritte Aussage folgt. Wir zeigen dies für die erste Aussage und verzichten auf den sehr ähnlichen Beweis im Fall der zweiten Aussage. Der Zeilenrang von A sei also höchstens gleich d, das heißt, der von den Zeilenvektoren von A erzeugte Untervektorraum U von K^n hat höchstens Dimension d. Wir können also ein Erzeugendensystem

$$\{(c_{1,1}, \ldots, c_{1,n}), (c_{2,1}, \ldots, c_{2,n}), \ldots, (c_{d,1}, \ldots, c_{d,n})\}$$

von U aus d Vektoren in K^n wählen. Sei $C = (c_{i,j})$ die (d, n)-Matrix, deren Zeilenvektoren gerade gleich diesen d Vektoren sind. Ähnlich wie in Lemma 182 lässt sich nun A in der Form BC für eine (m, d)-Matrix $B = (b_{i,j})$ darstellen, so dass für $i = 1, \ldots, m$ die d Einträge des Zeilenvektors i von B gerade die Koeffizienten einer Linearkombination des Zeilenvektors i von A aus den d Zeilenvektoren von C sind. Eine solche Matrix B existiert, weil die Zeilenvektoren von C ein Erzeugendensystem von U bilden und somit die

Zeilenvektoren von A als Linearkombination der Zeilenvektoren von C dargestellt werden können. Das heißt, für $i = 1, \ldots, m$ gibt es Skalare $b_{i,1}, \ldots, b_{i,d}$ mit

$$(a_{i,1}, \ldots, a_{i,n}) \;=\; \sum_{k=1}^{d} b_{i,k}(c_{k,1}, \ldots, c_{k,n}) \;=\; \left(\sum_{k=1}^{d} b_{i,k}c_{k,1}, \ldots, \sum_{k=1}^{d} b_{i,k}c_{k,n} \right).$$

Nach Definition der Matrixmultiplikation ist also jeder Eintrag $a_{i,j}$ von A gleich dem Eintrag der Matrix BC an der Position (i, j), es folgt $A = BC$. □

184 Satz. *Für jede Matrix ist der Zeilenrang gleich dem Spaltenrang.*

Beweis. Es sei A eine Matrix mit Zeilenrang d_{Z} und Spaltenrang d_{S}. Falls einer der beiden Werte d_{Z} und d_{S} gleich 0 ist, ist A eine Nullmatrix und auch der andere der beiden Werte ist gleich 0. Wir können also annehmen, das Zeilen- und Spaltenrang beide ungleich 0 sind. Somit ist für $d = d_{\mathrm{Z}}$ Lemma 183 auf A und d anwendbar. Nach Wahl von d ist Aussage (i) des Lemmas wahr, somit ist auch Aussage (ii) wahr und es gilt $d_{\mathrm{S}} \leq d = d_{\mathrm{Z}}$. Für $d = d_{\mathrm{S}}$ erhalten wir mit einem symmetrischen Argument $d_{\mathrm{Z}} \leq d_{\mathrm{S}}$, folglich sind Zeilen- und Spaltenrang von A gleich. □

185 Definition. *Der* Rang *einer Matrix A, kurz:* $\operatorname{rang}(A)$*, ist gleich dem Zeilenrang der Matrix.*

4.2 Lineare Gleichungssysteme

186 Definition. *Sei K ein Körper. Eine* lineares Gleichungssystem *in den Unbekannten $x_1, \ldots, x_n$ über K ist eine endliche Folge von Gleichungen der Form*

$$\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &=& b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &=& b_2 \\ \vdots \quad\quad \vdots \quad\quad \vdots \quad\quad \vdots &\vdots& \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &=& b_m \end{array} \tag{4.2}$$

mit $a_{i,j}$ und b_i aus K. Ein solches Gleichungssystem kann für die (m, n)-Matrix $A = (a_{i,j})$ und Spaltenvektoren $x = (x_1, \ldots, x_n)^{\mathrm{T}}$ und $b = (b_1, \ldots, b_m)^{\mathrm{T}}$ auch in der Form $Ax = b$ geschrieben werden.

Eine Lösung *des Gleichungssystems ist ein Spaltenvektor $c = (c_1, \ldots, c_m)^{\mathrm{T}}$ über K, so dass die Gleichung $Ac = b$ gilt, das heißt, alle Gleichungen des Gleichungssystems sind richtig, wenn dort überall x_i durch c_i ersetzt wird.*

187 Definition. *Ein Gleichungssystem der Form $Ax = b$ heißt* HOMOGEN, *falls b ein Nullvektor ist, also alle Einträge von b gleich 0 sind. Im allgemeinen Fall, wenn der Vektor b nicht unbedingt ein Nullvektor sein muss, heißt das Gleichungsystem* INHOMOGEN.

Die LÖSUNGSMENGE *eines linearen Gleichungssystems ist die Menge aller Lösungen des Gleichungssystems. Die Lösungsmenge eines Gleichungssystems $Ax = b$ wird mit $\mathrm{L}(A, b)$ bezeichnet, die Lösungsmenge des zugehörigen homogenen Gleichungssystems $Ax = o$ entspreched mit $\mathrm{L}(A, o)$.*

188 Satz. *Es sei K ein Körper und $A = (a_{i,j})$ sei eine (m, n)-Matrix über K. Die Lösungsmenge des homogenen Gleichungssystems $Ax = o$ über K ist ein Untervektorraum des K-Vektorraums K^n.*

Beweis. Die Elemente der Lösungsmenge $\mathrm{L}(A, o)$ sind Vektoren der Länge n über K, die Lösungsmenge ist somit eine Teilmenge von V. Es genügt also zu zeigen, dass die Bedingungen des Untervektorraumkriteriums für die Lösungsmenge erfüllt sind. Zunächst ist diese nicht leer, da sie den Nullvektor 0^n enthält. Gemäß den Rechenregeln der Matrixmultiplikation gilt für alle Lösungen c_1 und c_2

$$A(c_1 + c_2) \;=\; Ac_1 + Ac_2 \;=\; 0^n + 0^n \;=\; 0^n,$$

und für alle Lösungen c und Skalare a gilt

$$A(ac) = a(Ac) = a0^n = 0^n.$$

Der Lösungsraum $\mathrm{L}(A, o)$ ist also nichtleer und unter Vektoraddition und skalarer Multiplikation abgeschlossen und ist somit ein Untervektorraum von V. □

189 Satz. *Es sei K ein Körper und $Ax = b$ sei ein lineares Gleichungssystem über K, dass eine Lösung besitzt. Dann ist die Lösungsmenge $\mathrm{L}(A, b)$ gleich einer Nebenklasse der Lösungsmenge $\mathrm{L}(A, o)$ des zugehörigen homogenen Gleichungssystems. Genauer gilt für jede Lösung c des Gleichungssystems $Ax = b$*

$$\mathrm{L}(A, b) \;=\; c + \mathrm{L}(A, o).$$

Beweis. Sei c eine beliebig aber fest gewählte Lösung von $Ax = b$. Nach Satz 188 ist $U = \mathrm{L}(A, o)$ ein Untervektorraum von V. Es ist somit zu zeigen, dass $\mathrm{L}(A, b)$ gleich der Nebenklasse $c + U$ ist. Für jede Lösung c' des inhomogenen Gleichungssystems gilt

$$A(c' - c) \;=\; Ac' - Ac \;=\; b - b \;=\; o.$$

Die Differenz $u = c' - c$ ist also in U, damit ist $c' = c + u$ in $c + U$. Da c' als beliebige Lösung des inhomogen Systems gewählt wurde, ist $\mathrm{L}(A, b)$ eine

Teilmenge der Nebenklasse $c+U$. Umgekehrt lässt sich jedes Element in $c+U$ in der Form $c+u$ mit u in U schreiben, es folgt

$$A(c+u) \;=\; Ac+Au \;=\; b+o \;=\; b,$$

das heißt, $c+U$ ist Teilmenge von $\mathrm{L}(A,b)$. Folglich ist $\mathrm{L}(A,b)$ gleich $c+U$. □

190 Definition. *, Für eine (m,n)-Matrix A und einen Spaltenvektor b der Länge n sei $A \mid b$ die $(m,n+1)$-Matrix, die dadurch ensteht, dass b als neuer letzter Spaltenvektor zu A hinzugefügt wird. Für ein lineares Gleichungssystem der Form $Ax = b$ heißt A* Matrix *und $A \mid b$* erweiterte Matrix *des Gleichungssystems.*

191 Satz. *Für ein lineares Gleichungssystem der Form $Ax = b$ sind die folgenden Aussagen äquivalent.*

(i) *Das Gleichungssystem hat eine Lösung.*

(ii) *Die Untervektorräume, die durch die Spaltenvektoren von A beziehungsweise von $A \mid b$ erzeugt werden, sind identisch.*

(iii) *Die Matrizen A und $A \mid b$ haben denselben Rang.*

Beweis. Es seien U_0 und U_b die von den Spaltenvektoren von A beziehungsweise von $A \mid b$ erzeugten Untervektorräume. Nach Definition des Begriffs Rang gilt

$$\mathrm{rang}(A) \;=\; \dim(U_0) \qquad \text{und} \qquad \mathrm{rang}(A \mid b) \;=\; \dim(U_b).$$

Die Aussagen (ii) und (iii) besagen dann gerade, dass die Untervektorräume U_0 und U_b identisch sind beziehungsweise dieselbe Dimension haben. Da U_0 eine Teilmenge von U_b ist, sind beide Aussagen nach Lemma 154 äquivalent. Es bleibt die Äquivalenz der ersten beiden Aussagen zu zeigen.

(i) $\Rightarrow$ (ii): Es sei c eine Lösung des Gleichungssystems, es gilt also $Ac = b$. Nach Lemma 182 ist dann b eine Linearkombination der Spaltenvektoren von A. Somit ist b in U_0 und U_0 und U_b sind identisch.

(ii) $\Rightarrow$ (i): Nach Voraussetzung sind U_0 und U_b gleich, der Vektor b ist somit in U_0, kann also als Linearkombination der Spaltenvektoren von A dargestellt werden. Aus den Koeffizienten dieser Linearkombination ergibt sich kanonisch ein Vektor c mit $Ac = b$, dass Gleichungssystem hat also die Lösung c. □

4.3 Der Gaußsche Algorithmus

Der Gaußsche Algorithmus ist das Standardverfahrenn zur Lösung von linearen Gleichungssystemen. Er beruht auf elementaren Zeilen- und Spaltenumformungen von Matrizen, diese können auch zur Bestimmung des Rangs einer Matrix und zur Invertierung von Matrizen verwendet werden.

192 Definition. *Für Matrizen über einem Körper K werden die folgenden Umformungen als* ELEMENTARE ZEILENUMFORMUNGEN *bezeichnet.*

- *Vertauschung zweier Zeilen.*
- *Multiplikation einer Zeile mit einem Skalar ungleich* 0.
- *Addition eines skalaren Vielfachen einer Zeile zu einer anderen Zeile.*

Analog werden drei Typen von ELEMENTARE SPALTENUMFORMUNGEN *definiert.*

193 Beispiel. *Im Folgenden Beispiel wird die reelle Ausgangsmatrix durch elementare Zeilen- und Spaltenumformungen in eine Einheitsmatrix überführt, wobei nicht alle Zwischenergebnisse angegeben sind.*

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 5 \\ 2 & 5 & 6 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nach Definition ändern elementare Zeilen- und Spaltenumformungen die Dimension, also die Anzahl der Zeilen und Spalten, einer Matrix nicht. Tatsächlich bleibt auch der Rang einer Matrix erhalten.

194 Satz. *Elementare Zeilen- und Spaltenumformungen ändern den Rang einer Matrix nicht.*

Beweis. Wir zeigen für die drei elementaren Zeilenumformungen, dass diese den Zeilenrang und damit den Rang einer Matrix nicht ändern. Die symmetrischen und nahezu identischen Betrachtungen für die elementaren Spaltenumformungen lassen wir weg. Es genügt zu zeigen, dass sich durch elementare Zeilenumformungen der durch die Zeilenvektoren erzeugte Untervektorraum nicht ändert. Für die Vertauschung zweier Zeilen ist dies offensichtlich. Wird ein Zeilenvektor v durch $v' = av$ für einen Skalar a ungleich 0 ersetzt, so bleibt das Erzeugnis der Zeilenvektoren gleich. Dies folgt, da $v' = av$ als Linearkombination von v, aber auch $v = a^{-1}v'$ als Linearkombination des neuen Zeilenvektors v' dargestellt werden kann. Falls ein Zeilenvektor v durch $v' = v + au$ für einen Skalar a und einen Zeilenvektor u in einer anderen Zeile ersetzt wird, ist v' Linearkombination von v und u, aber auch $v = v' - au$ ist Linearkombination der neuen Zeilenvektoren, da u nach Voraussetzung in einer anderen Zeile als v steht und somit nach der Umformung noch vorhanden ist. □

Bestimmung des Rangs einer Matrix Aus Satz 194 ergibt sich folgender Ansatz um den Rang einer Matrix zu bestimmen: eine gegebene Matrix wird durch Zeilen- und Spaltenumformungen in eine Matrix überführt, deren Rang sich aus der Form dieser Matrix in einfacher Weise ergibt. Da beide Matrizen denselben Rang haben, ist damit dann auch der Rang der gegebenen Matrix bestimmt. Im Folgenden wird ein solches Verfahren vorgestellt, für dessen Beschreibung wir den Begriff s-MATRIX definieren. Eine (m, n)-Matrix A ist für ein $s \geq 0$ eine s-Matrix falls es $t \geq 0$ und $d \geq 0$ mit $m = s+t$ und $n = s+d$ gibt, so dass A in die folgenden vier Teilmatrizen zerlegt werden kann

- die quadratische (s, s)-Teilmatrix oben links, die Einträge $a_1, \ldots, a_s$ auf der Hauptdiagonalen dieser Teilmatrix sind ungleich 0, unter der Hauptdiagonalen sind alle Einträge gleich 0, die restlichen Einträge sind beliebig.
- die (t, s)-Teilmatrix unten links ist eine Nullmatrix,
- die (s, d)-Teilmatrix oben rechts mit beliebigen Einträgen,
- die (t, d)-Teilmatrix unten rechts mit beliebigen Einträgen, dieses wird im Folgenden mit C bezeichnet.

Eine s-Matrix heißt TERMINAL, wenn C eine Nullmatrix ist oder in dem Sinne nicht existiert, dass t oder d gleich 0 ist.

In (4.3) sind die Teilmatrizen einer s-Matrix schematisch dargestellt, dabei steht $*$ für beliebige Einträge. Im Fall $s = 0$ ist C gleich A, die anderen drei Teilmatrizen existieren nicht. Im Fall $t = 0$ existieren die beiden unteren Teilmatrizen nicht, im Fall $d = 0$ die beiden rechten Teilmatrizen.

$$\left(\begin{array}{ccccc|ccc} a_1 & * & \cdots & \cdots & * & * & \cdots & * \\ 0 & a_2 & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & & \ddots & \ddots & * & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & a_s & * & \cdots & * \\ \hline 0 & \cdots & \cdots & \cdots & 0 & & & \\ \vdots & & \vdots & & \vdots & & C & \\ 0 & \cdots & \cdots & \cdots & 0 & & & \end{array}\right) \tag{4.3}$$

Wir beschreiben ein Verfahren, dass eine nichtterminale s-Matrix A durch elementare Zeilen- und Spaltenumformungen in eine $(s + 1)$-Matrix A' umformt.

195 Umformung einer s-Matrix in eine $(s+1)$-Matrix. *Das Verfahren erhält als Eingabe eine s-Matrix A und formt diese in eine $(s+1)$-Matrix A' um. Die s-Matrix A wird als nichtterminal vorausgesetzt, folglich hat die Teilmatrix C von A einen Eintrag ungleich 0.*

Schritt 1 *Unter allen Einträgen ungleich 0 in der Teilmatrix C wähle denjenigen aus, dessen Position die kleinste Spaltennummer und innerhalb dieser Spalte die kleinste Zeilennummer hat.*

Schritt 2 *Bringe den ausgewählten Eintrag an die Position $(s+1, s+1)$ von A, das ist die Position in der ersten Zeile und ersten Spalte von C, indem innerhalb der Zeilen $s+1, \ldots, m$ und der Spalten $s+1, \ldots, n$ von A Zeilen- beziehungsweise Spalten vertauscht werden.*

Schritt 3 *Setze alle anderen Einträge in der ersten Spalte von C auf 0 indem geeignete Vielfache der ersten Zeile von C zu den restlichen Zeilen von C addiert, das heißt, zur jeder Zeile $i = s+2, \ldots, m$ von A wird ein geeignetes skalares Vielfaches der Zeile $s+1$ von A addiert, wodurch der Eintrag an Position $(i, s+1)$ von A zu 0 wird.*

Falls im Verfahren 195 im ersten Schritt der Eintrag in der ersten Spalte und ersten Zeile von C ungleich 0 ist, wird dieser Eintrag ausgewählt und im zweiten Schritt werden keine Vertauschungen vorgenommen.

196 Lemma. *Verfahren 195 wandelt jede nichtterminale s-Matrix in eine $(s+1)$-Matrix derselben Dimension und desselben Rangs um.*

Beweis. Es sei s eine natürliche Zahl und A sei eine nichtterminale s-Matrix. Durch Verfahren 195 werde A in eine Matrix A' umgeformt. Während des Verfahrens werden nur elementare Zeilen und Spaltenumformungen angewendet, die Matrizen A und A' stimmen also in Dimension und Rang überein. Es bleibt zu zeigen, dass A' eine $(s+1)$-Matrix ist. Falls s ungleich 0 ist und somit die quadratische (s,s)-Teilmatrix von A oben links existiert, wird diese durch die vorgenommenen Umformungen nicht verändert, sie bleibt also eine Matrix mit Einträgen ungleich 0 auf und gleich 0 unterhalb der Hauptdiagonalen. Auch die (t,s)-Teilmatrix von A unten links bleibt unverändert, da diese von den vorgenommenen Spaltenvertauschungen nicht betroffen ist und die vorgenommenen Zeilenumformungen nur die letzten t Zeilen von A betreffen und damit die Einträge dieser Nullmatrix nicht verändern. Durch den zweiten Schritt wird erreicht, dass der Eintrag an der Position $(s+1, s+1)$ ungleich 0 ist. Es muss also nur noch gezeigt werden, dass die Einträge an den Positionen $(i, s+1)$ mit $s+2 \leq i \leq m$ von A' gleich 0 sind. Dies wird immer erreicht, da im dritten Schritt des beschriebenen Verfahrens geeignete Faktoren gewählt werden können. Sei dazu a^{-1} das multiplikative Inverse des Skalars a an Position $(s+1, s+1)$ von A' und für $i = s+2, \ldots, m$

sei y_i der Eintrag an Position $(i, s+1)$ von A. Um letzteren Eintrag in 0 zu ändern, genügt es, die Zeile $s+1$ von A mit dem Skalar $-y_i \cdot a^{-1}$ zu multiplizieren und zur Zeile i zu addieren. Der Eintrag wird auf diese Weise durch $y_i - (y_i \cdot a^{-1})a = y_i - y_i(a^{-1}a) = y_i - (y_i \cdot 1) = 0$ ersetzt. □

Das Verfahren zur Bestimmung des Rangs einer gegebenen (m, n)-Matrix A arbeitet nun wie folgt.

197 Bestimmung des Rangs einer Matrix. *Das Verfahren wandelt eine gegebene Matrix A in eine terminale s-Matrix desselben Rangs um und gibt als Rang von A den Wert s aus.*

Beginnend mit $A^0 = A$ wird sukzessive für $s = 0, 1, \ldots$ die s-Matrix A^s durch Verfahren 195 in eine $(s+1)$-Matrix A^{s+1} überführt. Falls dabei A^s eine terminale s-Matrix ist, wird der Übergang zu A^{s+1} nicht mehr ausgeführt und das Verfahren bricht mit Ergebnis s ab.

Unter Verwendung von Lemma 196 folgt mit einem einfachen induktiven Argument, dass für Verfahren 197 die beiden folgenden Invarianten gelten.

Invariante I Für $s = 0, 1, \ldots$ ist jede erzeugte Matrix A^s eine s-Matrix.

Invariante II Alle erzeugten Matrizen $A^0, A^1, \ldots$ haben dieselbe Dimension und denselben Rang wie A.

Spätestens wenn s gleich dem Mininum von m und n ist und folglich die Teilmatrix C nicht existiert, ist eine terminale s-Matrix erreicht. Das Verfahren bricht also nach endlichen vielen Schritte mit einer terminalen s-Matrix ab. Diese hat nach dem folgenden Lemma Rang s, gemäß Invariante II ist dies auch der Rang von A.

198 Lemma. *Eine terminale s-Matrix hat Rang s.*

Beweis. Der Rang einer terminalen s-Matrix kann nicht größer als s sein, da die Matrix im Fall $t = 0$ und im Fall $d = 0$ jeweils nur s Zeilen- beziehungsweise Spaltenvektoren hat, und im Fall, dass C eine Nullmatrix ist, höchstens s linear unabhängige Zeilenvektoren hat. Es bleibt zu zeigen, dass der Rang einer terminale s-Matrix mindestens s ist. Im Fall $s = 0$ ist dies klar, da der Rang nicht negativ sein kann. Im Fall $s > 0$ betrachten wir eine Linearkombination des Nulllvektors aus den ersten s Zeilenvektoren. Es seien $a_1, \ldots, a_s$ die Einträge auf der Hauptdiagonale der oberen linken Teilmatrix in der natürlichen Reihenfolge. Wegen der Dreiecksform der oberen linken Teilmatrix und weil die a_i ungleich 0 sind, folgt zunächst dass der Koeffizient des ersten Zeilenvektors gleich 0 ist. Damit muss aber auch der Koeffizient des zweiten Zeilenvektors gleich 0 sein, und so weiter. Die ersten s Zeilenvektoren der gegebenen Matrix sind also linear unabhängig und der Rang der Matrix ist folglich mindestens s. □

Das Gaußsche Verfahren zur Lösung linearer Gleichungssysteme Ähnlich wie bei der Bestimmung des Rangs einer Matrix, wird im Gaußschen Verfahren ein gegebenes lineares Gleichungssystem der Form $Ax = b$ durch elementare Zeilenumformungen und Spaltenvertauschungen in ein lineares Gleichungssystem $A'x = b'$ überführt, so dass einerseits die Lösungsmengen $\mathrm{L}(A, b)$ und $\mathrm{L}(A', b')$ übereinstimmen und andererseits letztere Lösungsmenge aufgrund der speziellen Form der Matrix A' einfach bestimmt werden kann. Dazu sei vereinbart, dass elementaren Zeilenumformung auf eine erweiterte Matrix der Form $(A \mid b)$ derart angewendet werden, dass sich die Änderungen auch auf b erstrecken, werden also zum Beispiel zwei Zeilen vertauscht, werden die entsprechenden Einträge von b mitvertauscht.

199 Satz. *Wird eine erweiterte Matrix $(A \mid b)$ durch elementare Zeilenumformungen in eine erweiterte Matrix $(A' \mid b')$ überführt, so ändert sich die Lösungsmenge des zugehörigen inhomogenen beziehungsweise homogenen linearen Gleichungssystems nicht, es gilt*

$$\text{(i)} \quad \mathrm{L}(A, b) = \mathrm{L}(A', b') \qquad \textit{und} \qquad \text{(ii)} \quad \mathrm{L}(A, o) = \mathrm{L}(A', o').$$

Beweis. Es genügt, Aussage (i) zu zeigen, da Aussage (ii) gleich dem Spezialfall von Aussage (i) ist, in dem b ein Nullvektor ist. Sei also $(A \mid b)$ eine erweiterte Matrix für eine (m, n)-Matrix $A = (a_{i,j})$ und $b = (b_1, \ldots, b_m)^{\mathrm{T}}$. Nach einer Zeilenvertauschung enthält das zugehörige lineare Gleichungssystem dieselben Gleichungen wie zuvor, die Lösungsmenge ändert sich also nicht. Letzteres gilt auch, wenn Zeile i mit einem Skalar a ungleich 0 multipliziert wird, da für die zugehörige Gleichung gilt

$$\begin{aligned} \overbrace{a_{i,1}x_1 + \cdots + a_{i,n}x_n = b_i}^{(*)} \quad & \text{g.d.w.} \quad a_{i,1}x_1 + \cdots + a_{i,n}x_n - b_i = 0 \\ & \text{g.d.w.} \quad a(a_{i,1}x_1 + \cdots + a_{i,n}x_n - b_i) = 0 \\ & \text{g.d.w.} \quad aa_{i,1}x_1 + \cdots + aa_{i,n}x_n = ab_i. \end{aligned}$$

Wird schließlich für einen Skalar a das a-Fache von Zeile i zu einer anderen Zeile addiert, so bleibt Zeile i erhalten. Jede Lösung des neuen linearen Gleichungssystems muss somit die Gleichung $(*)$ zu Zeile i erfüllen und damit wie gerade gezeigt auch

$$\underbrace{aa_{i,1}x_1 + \cdots + aa_{i,n}x_n}_{=\ell} = \underbrace{ab_i}_{=r}.$$

Wurde zu Zeile j addiert, so hat die Gleichung zu dieser Zeile anschließend die Form

$$a_{j,1}x_1 + \cdots + a_{j,n}x_n + \ell = b_j + r. \tag{4.4}$$

Durch Zeile i wird für alle Lösungen des neuen linearen Gleichungssystems ℓ gleich r erzwungen. Die Summanden ℓ und r können deshalb in Gleichung (4.4)

gestrichen werden, ohne die Lösungsmenge des neuen Gleichungssystems zu ändern. Durch diese Streichung ergibt sich aus (4.4) dann gerade die Gleichung zu Zeile j im alten linearen Gleichungssystem, die Lösungsmengen des neuen und des alten linearen Gleichungssystems sind somit gleich. □

200 Bemerkung. *Im Gaußschen Verfahren werden auf erweiterten Matrizen der Form $(A \mid b)$ auch Vertauschungen der Spalten von A angewendet. Für das zugehörige lineare Gleichungssystem entspricht dies einer Vertauschung von zwei Unbekannten, die Lösungsmenge des Gleichungssystems ändert sich dadurch nur insofern, dass bei den Lösungsvektoren die zwei entsprechenden Positionen ebenfalls vertauscht sind.*

Für die Beschreibung des Gaußschen Verfahrens führen wir den Begriff einer normierten s-Matrix ein und erweitern diesen und die Begriffe s-Matrix und terminal auf erweiterte Matrizen.

201 Definition. *Eine s-Matrix ist* NORMIERT*, falls ihre (s, s)-Teilmatrix oben links eine Einheitsmatrix ist, das heißt, auf der Hauptdiagonale nur Einträge* 1 *und sonst nur Einträge* 0 *hat.*

Eine erweiterte Matrix der Form $(A \mid b)$ ist eine ERWEITERTE s-MATRIX*, falls A eine s-Matrix ist. Eine solche erweiterte Matrix ist* TERMINAL*, wenn A terminal ist und ist* NORMIERT*, wenn A normiert ist.*

Für eine erweiterte s-Matrix $(A \mid b)$ hat A die Form (4.3), b ist beliebig.

Das Gaußsche Verfahren wandelt eine gegebene erweiterte Matrix $(A \mid b)$ in Schritt 1 in eine terminale erweiterte s-Matrix um, und diese wiederum in Schritt 2 in eine normierte, terminale erweiterte s-Matrix, die Ergebnisse der beiden Schritte sind in Abbildung 4.1 dargestellt. Schritt 1 des Gaußschen Verfahrens ist sehr ähnlich zum Verfahren197 zur Bestimmung des Rangs einer Matrix, der Unterschied besteht nur darin, dass im Gaußschen Verfahren anstelle der Matrizen A^{s+1} erweiterte Matrizen $(A^{s+1} \mid b^{s+1})$ betrachtet werden.

202 Gaußsches Verfahren. *Das Gaußsche Verfahren erhält als Eingabe eine (m, n)-Matrix A und einen Vektor b der Länge n, beide über einem Körper K. Das Gaußsche Verfahren wandelt die erweiterte Matrix $(A \mid b)$ durch elementare Zeilen- und Spaltenumformungen zunächst in eine terminale erweiterte s-Matrix $(A^s \mid b^s)$ um, und anschließend in eine mormierte, terminale erweiterte s-Matrix $(A' \mid b')$. Wie vereinbart wirken dabei alle Zeilenumformungen jeweils auch auf die dem Vektor b entsprechende letzte Spalte der umzuformenden erweiterten Matrix.*

Schritt 1 *Umformung in eine terminale erweiterte s-Matrix.*

> *Beginnend mit der erweiterten 0-Matrix $(A^0 \mid b^0) = (A \mid b)$ wird sukzessive für $s = 0, 1, \ldots$ ein Übergang gemäß Verfahren 195 von*

einer erweiterten s-Matrix $(A^s \mid b^s)$ zu einer erweiterten $(s+1)$-Matrix $(A^{s+1} \mid b^{s+1})$ ausgeführt. Falls dabei $(A^s \mid b^s)$ terminal ist, wird der Übergang nicht mehr ausgeführt und das Gaußsche Verfahren wird mit $(A^s \mid b^s)$ und dem aktuellen Wert von s bei Schritt 2 fortgesetzt.

Schritt 2 *Die aus Schritt 1 resultierende terminale erweiterte s-Matrix der Form $(A^s \mid b^s)$ mit $A^s = (a^s_{i,j})$ und $b = (b^s_j)$ wird in eine normierte, terminale erweiterte s-Matrix $(A' \mid b')$ umgewandelt indem für $i = 1, \ldots, s$*

(a) für $i = 1, \ldots, s$ der Eintrag $(a^s_{i,i})$ an der i-ten Position auf der Hauptdiagonalen von A^s durch 1 ersetzt wird, indem Zeile i mit dem Inversen $(a^s_{i,i})^{-1}$ dieses Eintrags multipliziert wird,

(b) sukzessive für $i = 1, \ldots, s$ und $j = 1, \ldots, i-1$ der Eintrag an Position (i, j) durch 0 ersetzt wird, indem ein geeignetes Vielfaches von Zeile i zu Zeile j additiert wird.

Schritt 1 dess Gaußschen Verfahrens ergibt eine terminale erweiterte s-Matrix $(A^s \mid b^s)$, in dieser sind in der quadratischen (s, s)-Teilmatrix oben links alle Einträge unterhalb der Hauptdiagonalen gleich 0. In Schritt 2 wird diese Teilmatrix in eine Einheitsmatrix umgewandelt, indem in dieser Teilmatrix in Teilschritt (a) die Einträge auf der Hauptdiagonalen zu 1 geändert werden und in Teilschritt (b) die Einträge oberhalb der Hauptdiagonalen zu 0.

203 Bemerkung. *In Schritt 2 des Gaußschen Verfahrens spielt es beim Ändern der Einträge auf und oberhalb der Hauptdiagonalen der (s, s)-Teilmatrix oben links keine Rolle, in welcher Reihenfolge diese Änderungen vorgenommen werden, solange nur immer alle zu 0 geänderten Felder in derselben Spalte zusammen geändert werden und dabei die Faktoren geeignet gewählt werden, mit denen die hinzuaddierte Zeile multipliziert wird. Soll der Eintrag x an Position (i, j) oberhalb der Hauptdiagonalen gleich 0 gesetzt werden, und ist y der aktuelle Eintrag an Positionen (i, i), so muss der Faktor gleich $-x \cdot y^{-1}$ gewählt werden, das heißt, gleich $-x$, falls der Eintrag an Position (i, i) bereits gleich 1 ist.*

Wird das Gaußsche Verfahren auf eine erweiterte Matrix $(A \mid b)$ angewendet, so gilt die folgenden Invariante.

Invariante des Gaußschen Verfahrens Für alle im Verfahren erzeugten erweiterten Matrizen ist – bis auf die Vertauschungen von Unbekannten, die sich aus Spaltenvertauschungen in Schritt 1 ergeben – die Lösungsmenge des zugehörigen inhomogenen linearen Gleichungssystems gleich $\mathrm{L}(A, b)$ und die Lösungsmenge des zugehörigen homogenen linearen Gleichungssystems gleich $\mathrm{L}(A, o)$.

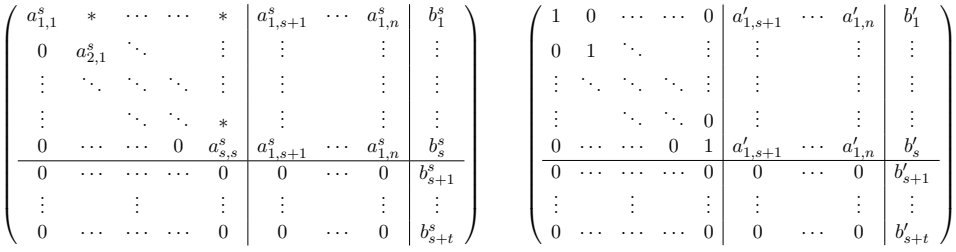

Abbildung 4.1: Ergebnis der Schritte 1 und 2 des Gaußschen Verfahrens: eine terminale erweiterte s-Matrix und deren normierte Variante.

Die Invariante gilt nach Satz 199, da alle erzeugten erweiterten Matrizen aus $(A \mid b)$ durch elementare Zeilenumformungen und eventuell durch zusätzliche Vertauschungen der Spalten der Teilmatrix A in Schritt 1 hervorgehen.

Wird durch das Gaußsche Verfahren die erweiterte Matrix $(A \mid b)$ in eine normierte, terminale erweiterte s-Matrix $(A' \mid b')$ ungeformt, so gilt gemäß der Invariante des Verfahrens, dass für die beiden erweiterten Matrizen die Lösungsmengen der zugehörigen homogenen Gleichungssystems identisch sind, und ebenso für die zugehörigen inhomogenen Systeme, das heißt, es gilt

$$\mathrm{L}(A \mid o) \;=\; \mathrm{L}(A' \mid o) \qquad \text{und} \qquad \mathrm{L}(A \mid b) \;=\; \mathrm{L}(A' \mid b'),$$

jeweils bis auf eventuelle Variablenvertauschungen. Gemäß dem folgenden Satz können diese Lösungsmengen an der durch das Gaußsche Verfahren berechneten erweiterten Matrix $(A' \mid b')$ direkt abgelesen werden.

204 Satz. *Es sei $A' = (a'_{i,j})$ eine (m,n)-Matrix und $b' = (b'_1, \ldots, b'_n)$ ein Vektor der Länge n über einem Körper K, so dass $(A' \mid b')$ eine normierte, terminale erweiterte s-Matrix ist.*

Die Lösungsmenge $\mathrm{L}(A', o)$ des homogenen Gleichungssystems $A'x = o$ ist ein Untervektorraum der Dimension $d = n-s$, eine Basis dieses Vektorraums bilden die Vektoren $v_1, \ldots, v_d$ mit

$$v_j \;=\; (-a_{1,s+j},\; -a_{2,s+j},\; \ldots,\; -a_{s,s+j},\; \underbrace{0,\; \ldots,\; 0}_{j-1 \text{ mal}},\; 1,\; \underbrace{0,\; \ldots,\; 0}_{d-j \text{ mal}}),$$

das heißt, die ersten s Einträge des Vektors v_j sind gleich den ersten s Einträgen in Spalte $s+j$ von A' und die restlichen d Einträge des Vektors v_j sind alle gleich 0, ausgenommen der Eintrag $s+j$, dieser ist gleich 1.

Falls einer der Einträge $b'_{s+1}, \ldots, b'_n$ ungleich 0 ist, so hat das inhomogene lineare Gleichungssystems $A'x = b'$ keine Lösung, anderfalls ist der Vektor $c = (b'_1, \ldots, b'_s, 0, \ldots, 0)$ der Länge n eine Lösung. Im letzteren Fall ist

die Lösungsmenge $L(A', b')$ gleich der Nebenklassen $c + \mathrm{L}(A', o)$ zur Lösungsmenge des homogenen Systems.

Beweis. Satz 191 besagt, dass die Lösungsmenge des inhomogenen Systems genau dann leer ist, wenn die erweiterte Matrix $A' \mid b'$ echt größeren Rang hat als A'. Letzteres ist genau dann der Fall, wenn einer der Einträge $b'_{s+1}, \ldots, b'_n$ ungleich 0 ist: in diesem Fall lässt sich b nicht als Linearkombination der Spaltenvektoren von A' darstellen, der Rang der erweiterten Matrix ist also echt größer. Sind diese Einträge dagegen alle gleich 0, ist b Linearkombination der ersten s Spaltenvektoren von A' und beide Ränge sind gleich. In diesem Fall ist der angegeben Vektor c eine Lösung des inhomogenene Systems, was sich durch Nachrechnen bestätigen lässt. Dass der Lösungsraum des inhomogenen Systems dann wie angegeben als Nebenklasse dargestellt werden kann, folgt aus Satz 189.

Dass die Lösungsmenge $\mathrm{L}(A', o)$ des homogenen Systems ein Untervektorraum ist, gilt nach Satz 188. Es bleibt zu zeigen, dass die Vektoren $v_1, \ldots, v_d$ eine Basis diese Untervektorraums bilden. Wir zeigen zunächst, dass jeder Vektor v_j Lösung des homogenen Systems ist, also $A \cdot v_j = o$ gilt. Dies ist äquivalent dazu, dass das Produkt jedes Zeilenvektors von A' mit v_j gleich 0 ist. Für das Produkt von Zeilenvektor i von A' mit v_j gilt: für $i > s$ ist der Zeilevektor ein Nullvektor und das Produkt ist gleich 0. Für $i < s$ ist das Produkt gleich

$$
\begin{aligned}
&(\underbrace{0, \ldots, 0}_{i-1 \text{ mal}}, \underbrace{1}_{i}, \underbrace{0, \ldots, 0}_{s-i \text{ mal}}, \underbrace{a'_{i,s+1}, -a'_{i,s+2}, \ldots}_{j-1 \text{ Einträge}}, \underbrace{a'_{i,s+j}}_{s+j}, \underbrace{\ldots, a'_{i,s+d}}_{d-j \text{ Einträge}}) \\
&\cdot(\underbrace{-a'_{1,s+j}, \ldots}_{i-1 \text{ Einträge}}, \underbrace{-a'_{i,s+j}}_{i}, \underbrace{\ldots, -a'_{s,s+j}}_{s-i \text{ Einträge}}, \underbrace{0, \ldots, 0}_{j-1 \text{ mal}}, \underbrace{1}_{s+j}, \underbrace{0, \ldots, 0}_{d-j \text{ mal}}) \\
&= 1 \cdot (-a'_{i,s+j}) + a'_{i,s+j} \cdot 1 = 0.
\end{aligned}
$$

Durch die Wahl der jeweils letzten d Einträge der Vektoren v_1 bis v_d wird sichergestellt, dass diese Vektoren linear unabhängig sind, den offensichtlichen Beweis lassen wir aus. Es bleibt somit nur zu zeigen, dass diese Vektoren ein Erzeugendensystem des Untervektorraums $\mathrm{L}(A, o)$ bilden. Da alle v_j Lösungen und damit in diesem Untervektorraum sind, bleibt nur zu zeigen, dass jede Lösung als Linearkombination der v_j dargestellt werden kann. Sei dazu c eine beliebigen Lösung c der Form $(c_1, \ldots, c_n)^{\mathrm{T}}$ in $\mathrm{L}(A, o)$. Wir betrachten die Linearkombination

$$v = c_{s+1}v_1 + \cdots + c_{s+d}v_d$$

bei welcher der Koeffizient von v_j gleich dem Eintrag $s + j$ von c ist, das heißt, die Vektoren v und c stimmen in den letzte d Einträgen überein. Für jede Lösung des homogenen Gleichungssystems sind nun durch die Wahl der

letzten d Einträge alle anderen Einträge bereits festgelegt, da dann für $i = 1, \ldots, s$ die Gleichung zu Zeile i nur durch geeignete Wahl des Eintrags i erfüllt werden kann. Folglich sind die Lösungen v und c gleich. Da c als beliebige Lösung des homogenen Gleichungssystem gewählt war, bilden die Vektoren $v_1, \ldots, v_d$ eine Basis des Untervektorraums $\mathrm{L}(A, o)$, dieser hat somit Dimension d. □

205 Korollar. *Es sei A eine (m, n)-Matrix mit Rang d. Dann hat die Lösungsmenge $\mathrm{L}(A, o)$ des zugehörigen homogenen linearen Gleichungssystems die Dimension $n - d$.*

Beweis. Durch die Anwendung des Gaußschen Verfahrens wird die erweiterte Matrix $(A \mid o)$ in einer normierte, terminale erweiterte s-Matrix $(A' \mid o)$ umgeformt, da der Nullvektor in der letzten Spalte der erweiterten Matrix durch die vorgenommenen Zeilenumformungen nicht verändert wird. Die erweiterte s-Matrix $(A' \mid o)$ hat nach Satz 198 Rang s und das zugehörige homogene Gleichungssystem nach Satz 204 eine Lösungsmenge der Dimension $n - s$. Da die Matrix A und die erweiterten Matrizen $(A \mid o)$ und $A' \mid o)$ alle drei denselben Rang haben, folgt $s = d$ und damit das Korollar. □

Die Berechnung der inversen Matrix Zur Erinnerung: die Einheitsmatrix E_n ist die (n, n)-Matrix mit Einsen auf der Hauptdiagonale und Einträgen 0 sonst.

206 Bemerkung. *Die Einheitsmatrizen sind im folgenden Sinne neutrale Elemente der Matrixmultiplikationn: für jede (m, n)-Matrix A gilt*

$$E_m A = A \qquad \textit{und} \qquad A E_n = A, \tag{4.5}$$

im Fall $m = n$, für eine quadratische Matrix, gilt somit $E_n A = A E_n = A$. Die Gleichungen (4.5) folgen aus der Darstellung der Matrixmultiplikation über Zeilen- oder Spaltenvektoren gemäß Lemma 182. Nach diesem Lemma ist zum Beispiel Zeile i des Produkts $E_m A$ eine Linearkombination der Zeilen von A, deren Koeffizienten gleich den Einträgen in Zeile i von E_m sind. Dort gibt es an Position i einen Eintrag 1 und sonst nur Einträge 0, folglich ist Zeile i von $E_m A$ gleich Zeile i von A.

207 Satz. *Es sei A eine (n, n)-Matrix. Eine Matrix L mit $LA = E_n$ heißt* LINKSINVERS *zu A, eine Matrix R mit $AR = E_n$* RECHTSINVERS. *Die Matrix A ist invertierbar, wenn sie eine linksinverse Matrix besitzt. Die Matrix A hat* VOLLEN RANG, *wenn ihr Rang gleich n.*

Wir zeigen in diesem Abschnitt, dass eine quadratische Matrix genau dann invertierbar ist, wenn sie vollen Rang hat, und dass in diesem Fall jede linksinverse Matrix auch rechtsinvers ist und umgekehrt.

208 Lemma. *Es sei A eine quadratische Matrix. Falls A nicht vollen Rang hat, ist A nicht invertierbar.*

Beweis. Es sei A eine (n,n)-Matrix mit Rang $s < n$. Nach Lemma 205 hat die Lösungmenge $\mathrm{L}(A,o)$ des zugehörige homogene linearen Gleichungssystems Dimension $d = n - s \neq 0$, enthält also einen Vektor v der Länge n, der kein Nullvektor ist. Hätte A eine linkinverse Matrix L, würde gelten

$$L(Av) = Lo = o \qquad \text{und} \qquad (LA)v = E_n v = v,$$

im Widerspruch zur Assoziativität der Matrixmultiplikation. □

209 Lemma. *Es sei A eine (n,n)-Matrix mit vollem Rang. Dann kann A durch elementare Zeilenumformungen in E_n überführt werden.*

Beweis. Das oben beschriebene Verfahren zur Bestimmung des Rangs einer Matrix überführt A in eine n-Matrix der Dimension (n,n), das heißt, eine (n,n)-Matrix mit Einträgen ungleich 0 auf der Hauptdiagonalen und Einträgen gleich 0 unterhalb der Hauptdiagonalen. Wie im Gaußschen Verfahren können dann durch elementare Zeilenumformungen auch die Einträge oberhalb der Hauptdiagonale gleich 0 und die Einträge auf der Hauptdiagonalen gleich 1 gesetzt werden, so dass sich insgesamt E_n ergibt.

Es bleibt zu zeigen, dass für eine Matrix A von vollem Rang gleich n bei der anfänglichen Überführung in eine n-Matrix keine Spaltenvertauschungen auftreten, ddass also bei den sukzessiven Übergängen von der s-Matrix A^s zur $(s+1)$-Matrix A^{s+1} nie die erste Spalte der unteren rechten Teilmatrix C von A^s nur Einträge 0 enthält. Wäre Letzteres für ein s der Fall, wäre der Vektor in Spalte $s+1$ von A^s eine Linearkombination der ersten s Spaltenvektoren dieser Matrix, im Widerspruch dazu, dass A und damit auch alle erzeugten Matrizen A^1, A^2, …, Rang n haben. □

Als nächstes beobachten wir, dass sich elementare Zeilenumformungen auch durch Multiplikation von links mit einer geeigneten Matrix ausdrücken lassen, zum Beispiel gilt für die reellen Matrizen

$$H_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad H_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{pmatrix}$$

und jede reelle $(3,n)$-Matrix A, dass $H_i A$ aus A wie folgt entsteht: für H_1 durch Vertauschen der ersten beiden Zeilen, für H_2 durch Multiplikation der zweiten Zeile mit 5 und für H_3 indem die mit 5 multiplizierte zweite Zeile zur dritten addiert wird.

210 Lemma. *Es sei m eine natürliche Zahl und es sei $I = \{1, \ldots, m\}$*

(i) *Für alle Indizes i_1 und i_2 in I mit $i_1 \neq i_2$ gibt es eine Matrix H, so dass für jede (m,n)-Matrix A die Matrix HA gleich der Matrix ist, die sich durch Vertauschen der Zeilen i_1 und i_2 von A ergibt.*

(ii) *Für alle Indizes i in I und alle Skalare a gibt es eine Matrix H, so dass für jede (m,n)-Matrix A die Matrix HA gleich der Matrix ist, die sich durch Multiplizieren der Zeilen i von A mit a rgibt.*

(iii) *Für alle Indizes i_1 und i_2 in I und alle Skalare a gibt es eine Matrix H, so dass für jede (m,n)-Matrix A die Matrix HA gleich der Matrix ist, die sich aus A ergibt, wenn zu Zeile i_2 die mit a multiplizierte Zeile i_1 addiert wird.*

Beweisskizze. Wir geben ohne weitere Begründung geeignete Matrizen an. Diese ergeben sich in allen drei Fällen durch Abändern der Einheitsmatrix E_m. Für (i) werden in E_m die Zeilen i_1 und i_2 vertauscht, für (ii) wird Zeile i von E_m mit a multipliziert und für (iii) wird der Eintrag 0 an der Position (i_2, i_1) von E_m durch a ersetzt. □

211 Satz. *Jede quadratische Matrix mit vollem Rang hat eine linksinverse Matrix.*

Beweis. Es sei A eine (n,n)-Matrix mit Rang n. Gemäß Lemma 209 kann A durch elementare Zeilenumformungen in die Einheitsmatrix E_n überführt werden. Werden dabei insgesamt t Zeilenumformungen durchgeführt, so gibt es gemäß Lemma 210 eine Folge $H_1, \ldots, H_t$ von Matrizen, so dass für deren Produkt

$$H \;=\; H_t \cdot H_{t-1} \cdot \;\cdots\; \cdot H_1$$

und jede (n,n)-Matrix X gilt: HX ist gleich der Matrix, die sich ergibt, wenn auf X dieselbe Folge von elementaren Zeilenumformungen angewendet wird wie auf A. Insbesondere gilt $HA = E_n$, die Matrix H ist also linksinvers zu A. □

212 Bemerkung. *Die linksinverse Matrix zu einer (n,n)-Matrix A von vollem Rang ergibt sich, indem auf die Einheitsmatrix E_n dieselben Zeilenumformungen angewendet werden wie beim Übergang von A zu E_n im Beweis von Satz 211. Die Anwendung dieser Zeilenumformungen auf eine (n,n)-Matrix entsprich einer Multiplikation von links mit einer Matrix H wie im Beweis, das heißt, die Anwendung der Zeilenumformungen auf E_n ergibt $HE_n = H$ und damit die linksinverse Matrix H zu A. Die Matrix H kann so berechnet werden, dass beim Übergang von A zu E_n die Zeilenumformungen nicht nur auf A, sondern auch auf E_n angewendet werden. Nach Wahl von H und der Definition der Matrixmultiplikation ergibt dies die Matrix*

$$H(A \mid E_n) \;=\; (HA \mid HE_n) \;=\; (E_n \mid H).$$

Dabei ist zum Beispiel $(A \mid E_n)$ die $(n, 2n)$-Matrix, bei der die Teilmatrize den ersten und der letzten n Spalten gleich A beziehungsweise E_n ist.

213 Bemerkung. *Analog zum Beweis von Satz 211 lässt sich zeigen, dass jede quadratische Matrix von vollem Rang eine rechtsinverse Matrix besitzt. Die Argumentation ist im Wesentlichen die gleiche, anstelle von Zeilenumformungen wird aber mit Spaltenumformungen argumentiert und dabei benutzt, dass jede Spaltenumformung durch Multiplikation mit einer geeigneten Matrix von rechts dargestellt werdenn kann. Weiter wird mit einer symmetrische Variante des Gaußschen Verfahren argumentiert, bei der eine quadratische Matrix mit vollem Rang n nur durch Spaltenumformungen sukzessive für $s = 1, \ldots,$ in eine zu einer s-Matrix symmetrischen Form überführt wird. Für $s = n$ hat die symmetrische Form Einträge ungleich 0 auf der Hauptdiagonalen und unterhalb der Hauptdiagonalen nur Einträge 0, ist also auch eine s-Matrix. Es ergibt sich ein Verfahren zur Berechnung der rechtsinverse Matrix zu einer Matrix A, das sehr ähnlich zu dem in Bemerkung 212 beschriebenen Verfahren zur Berechnung einer linksinversen Matrix ist. Dabei werden die Spaltenumformungen, die A in E_n überführen, auf eine $(2n, n)$-Matrix angewendet, bei der die Teilmatrix der ersten n Zeilen gleich A und die Teilmatrix der letzten n Zeilen gleich E_n ist.*

214 Bemerkung. *Nach Satz 211 und Bemerkung 212 besitzt jede quadratische Matrix mit vollem Rank sowohl links- als auch rechtsinverse Matrizen. Diese sind tatsächlich alle gleich und insbesonder ist jede linksinverse Matrix einer quadratischen Matrix auch rechtsinvers und umgekehrt. Statt von einer links- oder rechtsinversen Matrix einer gegebenen Matrix A wird deshalb von der inversen Matrix von A gesprochen.*

Sind L und R eine links- beziehungsweise rechtsinverse Matrix zu einer quadratischen Matrix A, so gilt wegen der Assoziativität der Matrixmultiplikation

$$L \;=\; LE_n \;=\; L(AR) \;=\; (LA)R \;=\; E_nR = R.$$

Daraus folgt auch, dass alle links- und rechtsinversen Matrizen identisch sind. Beispielsweise sind je zwei linksinverse Matrizen L_1 und L_2 identisch, da sie ja beide gleich jeder beliebigen rechtsinversen Matrix sind.

Exkurs über fehlerkorrigierende Codes

215 Definition. *Ein* Alphabet *ist eine endliche Menge, die Elemente des Alphabets werden* Buchstaben, Symbole *oder* Zeichen *genannt. Ein* Wort *über einem Alphabet ist eine endliche Folge von Buchstaben aus dem Alphabet, die Länge dieser Folge wird als* Länge *des Worts bezeichnet. Es sei λ das* leere Wort, *welches einer Folge der Länge 0 entspricht. Das Binäralphabet ist die Menge $\{0,1\}$, ein* Binärwort *ist ein Wort über dem Binäralphabet.*

Falls nicht explizit etwas anderes gesagt wird, soll im Folgenden die Schreibweise $w = w_1 \cdots w_n$ dafür stehen, dass w ein Wort der Länge n mit Buchstaben w_1 bis w_n ist, dabei ist w_i der Buchstabe an der Stelle oder Position i von w.

216 Definition. *Der* Hammingabstand *zweier Wörter $u = u_1 \cdots u_n$ und $v = v_1 \cdots v_n$ gleicher Länge n ist*

$$\mathrm{d_H}(u,v) \;=\; |\{i\colon 1 \le i \le n \text{ und } u_i \neq v_i\}|\,,$$

ist also gleich der Anzahl der Stellen, an denen sich u und v unterscheiden.

Beispielsweise ist $\mathrm{d_H}(abba, aabb) = 2$ und $\mathrm{d_H}(110, 110) = 0$. Im Englischen wird der Hammingabstand auch als *edit distance* bezeichnet, da $\mathrm{d_H}(u,v)$ die Anzahl der Positionen ist, an denen u geändert werden muss, um v zu erhalten.

Der Hammingabstand ist in dem Sinne eine Metrik, dass für alle Wörter u, v und w gleicher Länge gilt

(i) $\mathrm{d_H}(u,v) \ge 0$ und $\mathrm{d_H}(u,v) = 0$ genau für $u = v$ (Positive Definitheit),

(ii) $\mathrm{d_H}(u,v) = \mathrm{d_H}(v,u)$ (Symmetrie),

(iii) $\mathrm{d_H}(u,w) \le \mathrm{d_H}(u,v) + \mathrm{d_H}(v,w)$ (Dreiecksungleichung).

Die Dreiecksungleichung lässt sich wie folgt einsehen: um u in w zu ändern werden höchstens so viele Änderungen benötigt, wie wenn zunächst u in v und dann v in w geändert wird.

Im Folgenden betrachten wir nur das Binäralphabet und Binärwörter, falls nicht explizit etwas anderes gesagt wird, bedeutet Wort immer Binärwort.

217 Definition. *Ein* Code *ist eine Menge von Wörtern gleicher Länge n, diese wird als* Codewortlänge *des Codes bezeichnet, die Elemente des Codes heiße* Codewörter. *Der* Mininmalabstand *eines Codes C ist*

$$\mathrm{d_H}(C) \;=\; \min\{\mathrm{d_H}(u,v)\colon u,v \in C \text{ und } u \neq v\},$$

218 Beispiel. *Die verschiedenen Varianten des weit verbreiteten ASCII-Codes haben alle Codewortlänge* 8*. Wir betrachten eine Variante C, bei der für die Codewörter nur die ersten sieben Buchstaben frei gewählt werden können und an Position* 8 *ein* Prüfzeichen *steht, das so gewählt wird, dass die Anzahl der Einsen im Codeworts gerade ist. Von den $2^8 = 256$ Wörtern der Länge* 8 *enthält C also nur $2^7 = 128$, da für jede Wahl der ersten sieben Buchstaben genau eine der beiden Möglichkeiten für Position* 8 *entfällt.*

In C gibt es Codewörter wie 00000000 *und* 00000011 *mit Abstand* 2*, während die Änderung eines Codeworts an nur einer Stelle immer ein Wort mit einer ungeraden Anzahl von Einsen ergibt, das somit kein Codewort ist. Der Code C hat folglich Minimalabstand* 2*. Dies bedeutet, dass jedes an nur einer Stelle veränderte Codewort als Nicht-Codewort erkannt werde kann, es aber auch Codewörter gibt, die durch Änderungen an nur zwei Stellen in ein anderes Codewort überführt werden.*

Codes werden in der Praxis so verwendet, dass nur Codewörter übertragen oder gespeichert werden. Wird dabei ein Codewort w zum Beispiel durch Übertragungsfehler zu einem Wort u ungleich w geändert, so wird das veränderte Wort u als fehlerhaft bezeichnet, uabhängig davon, ob u selbst ein Codewort ist oder nicht. Beispiel 218 lässt sich dann wie folgt verallgemeinern: für jeden Code mit einem Minimalabstand von mindestens $t + 1$ kann ein an mindestens einer und höchstens t Stellen verändertes Codewort als fehlerhaft erkannt werden, da sich durch eine solche Veränderung kein Codewort ergeben kann.

219 Definition. *Ein t-*fehlererkennender *Code ist ein Code mit einem Minimalabstand von mindestens $t + 1$.*

Wird für den Code C aus Beispiel 218 das Wort $u = 00000001$ empfangen, so kann erkannt werden, dass u kein Codewort ist und somit durch der fehlerhaften Übertragungen eines anderen Worts w entstanden ist. Da es mehrere Codewörte im Hammingabstand 1 zu u gibt, kann auch unter der Annahme, dass nur ein Buchstabe von w fehlerhaft übertragen wurde, aus u nicht auf w geschlossen werden. Eine solche Fehlerkorrektur würde einen Code mit Minimalabstand 3 erfodern.

220 Definition. *Es sei w ein Wort der Länge n und $r \geq 0$ eine ganze Zahl. Die* Kugel um w mit Radius r *ist*

$$\mathrm{B_H}(w, r) \;=\; \{u\colon |u| = n \text{ und } \mathrm{d_H}(w, u) \leq r\},$$

das ist die Menge aller Wörter u gleicher Länge wie w, die sich an höchstens r Stellen von w unterscheiden.

221 Lemma. *Es sei C ein Code mit einem Minimalabstand von mindestens $2t+1$ und v und w seien zwei verschiedene Codewörter. Dann sind die Kugeln $\mathrm{B_H}(v, t)$ und $\mathrm{B_H}(w, t)$ disjunkt.*

Beweis. Für einen Beweis durch Widerspruch nehmen wir an, dass der Schnitt der Kugeln $\mathrm{B_H}(v,t)$ und $\mathrm{B_H}(w,t)$ ein Wort u enthält. Dann gilt $\mathrm{d_H}(v,u) \leq t$ und $\mathrm{d_H}(w,u) \leq t$, und mit der Dreiecksungleichung folgt $\mathrm{d_H}(v,w) \leq 2t$, im Widerspruch zur Annahme, dass der Minimalabstand von C mindestes $2t+1$ ist. □

Für einen Code mit einem Minimalabstand von mindestens $2t+1$ sei angenommen, dass sich ein fehlerhaftes Wort u an höchstens t Stellen von einem Codewort w unterscheiden kann, jedes fehlerhafte Wort befindet sich also in einer Kugel der Form $\mathrm{B_H}(w,t)$ für ein Codewort w. Aus einem solchen Wort u kann dann auf w geschlossen werden, da u in keiner anderen solchen Kugel mit Radius t um ein Codewort ist.

222 Definition. *Ein t-*FEHLERKORRIGIERENDER *Code ist ein Code mit einem Minimalabstand von mindestens $2t+1$. Ein t-fehlerkorrigierender Code der Codewortlänge n ist* PERFEKT*, falls die Kugeln der Form $\mathrm{B_H}(w,t)$ für ein Codewort w eine Zerlegung der Menge aller Wörter der Länge n bilden.*

Ein perfekter t-fehlererkennender Code hat die maximal mögliche Anzahl von Codewörtern für einen t-fehlererkennenden Codes der gegebenen Codewortlänge. Für den Beweis und die Diskussion des folgenden Satzes identifizieren wir in der natürlichen Weise das Binäralphabet $\{0,1\}$ mit dem Körper $\mathbb{F}_2 = \{0,1\}$ sowie die Binärwörter der Länge n mit Spaltenvektoren der Längen n, beispielsweise entspricht das Binärwort 110 dem Spaltenvektor $(1,1,0)^{\mathrm{T}}$.

223 Satz. *Es sei s eine natürlichen Zahl und es sei $n = 2^s - 1$. Dann gibt es einen perfekten 1-fehlererkennenden Code C_s der Codewortlänge n.*

Beweis. Es seien $v_1, \ldots, v_n$ die $2^s - 1$ vielen Spaltenvektoren über $\mathbb{F}_2$ der Länge s ungleich dem Nullvektor in lexikographischer Ordnung, wobei das Symbol 0 kleiner als das Symbol 1 sein soll. Es ist also

$$\begin{aligned} v_1 &= (0,0,\ldots,0,0,1)^{\mathrm{T}},\\ v_2 &= (0,0,\ldots,0,1,0)^{\mathrm{T}},\\ v_3 &= (0,0,\ldots,0,1,1)^{\mathrm{T}},\\ &\vdots \qquad\qquad \vdots\\ v_n &= (1,1,\ldots,1,1,1)^{\mathrm{T}}, \end{aligned}$$

und es sei $A = A_s$ die (s,n)-Matrix über $\mathbb{F}_2$, deren Spaltenvektor j gleich v_j ist. Wir setzen C_s gleich der Lösungsmenge $\mathrm{L}(A,o)$ des zur Matrix A gehörigen homogenen linearen Gleichungssystems.

Für ein Wort w der Länge n und Indizes j_i mit $1 \leq j_i \leq n$ sei $w[j_1,\ldots,j_t]$ das Wort, das aus w entsteht indem sukzessive die Stellen $j_1,\ldots,j_t$ von w

geändert werden, beispielsweise ist $1000[1,4] = 0001$. Nach Definition der Addition in $\mathbb{F}_2$ gilt

$$Aw[j] \;=\; Aw + v_j \qquad \text{und} \qquad Aw[j_1, \ldots, j_t] \;=\; Aw + v_{j_1} + \cdots + v_{j_t}.$$

Wir zeigen zunächst, dass der Minimalabstand von C_s mindestens gleich 3 ist und somit die Kugeln $\mathrm{B_H}(w, 1)$ mit Radius 1 um die Codewörter paarweise disjunkt sind. Für jedes Codewort w gilt $Aw = o$. Wird w an einer Stelle j oder an zwei verschiedenen Stellen j_1 und j_2 geändert, so gilt für die so erhaltenen Wörter $w[j]$ und $w[j_1, j_2]$

$$Aw[j] = o + v_j = v_j \neq o \qquad \text{und} \qquad Aw[j_1, j_2] = o + v_{j_1} + v_{j_2} = v_{j_1} + v_{j_2} \neq o.$$

Durch Änderung eines Codeworts an einer oder zwei Stellen ergibt sich also kein Codewort, der Minimalabstand ist somit mindestens 3. Der Code C_s ist also 1-fehlerkorrigierend und die Kugeln $\mathrm{B_H}(w, 1)$ mit w in C_s sind paarweise disjunkt. Es bleibt zu zeigen, dass die Vereinigung dieser Kugeln alle Vektoren der Längen n enthält. Sei dazu u ein solcher Vektor. Ist u ein Codewort, so ist u in der Kugel mit Mittelpunkt u. Andernfalls ist Au ungleich o und damit gleich v_j für einen Index j. Für den Vektor $u[j]$ gilt dann

$$Au[j] = Au + v_j = v_j + v_j = o,$$

somit ist $u[j]$ ein Codewort mit Hammingabstand 1 zum Vektor u, dieser ist also Element der Kugel $\mathrm{B_H}(u[j], 1)$. □

224 Bemerkung. *Der Beweis von Theorem 223 zeigt, dass für die dort konstruierten Codes C_s für ein gegebenes Wort u effizient überprüft werden kann, ob u ein Codewort ist, und falls nicht, effizient das eindeutig bestimmte Codeword w mit Hammingabstand 1 zu u bestimmt werden kann. Es genügt, das Wort u mit der Matrix $A = A_s$ zu multiplizieren: ist Au der Nullvektor, so ist u ein Codewort, andernfalls ist Au für einen Index j gleich dem Vektor v_j und es ergibt sich ein Codewort, wenn u an Stelle j geändert wird.*

Die perfekten fehlerkorrigierenden Codes aus Theorem 223 liefern eine überraschende Lösung für das in der folgenden Bemerkung behandelte Hutproblem

225 Bemerkung (Das Hutproblem und fehlerkorrigierende Codes)**.** *Beim Hutproblem wird jeder von $n \geq 1$ Spielerinnen per Wurf einer unverfälschten Münze eine Hutfarbe blau oder rot zugelost. Jede Spielerin sieht alle Hüte und deren Farbe außer ihren eigenen. Jede Spielerin kann entweder einen Tipp zur die Farbe ihres Hutes abgeben oder sich enthalten, dass Team gewinnt als Ganzes, wenn mindestens ein Tipp richtig und keiner falsch ist. Die Spielerinnen sehen nicht, ob und, falls ja, wie die anderen Spielerinnen tippen. Das Team kann sich in einer Vorbereitungsphase vor Beginn des Spiels auf*

eine beliebig komplizierte Strategie einigen, danach können die Spielerinnen nicht mehr miteinander kommunizieren.

Betrachte für $n = 3$ Spielerinnnen folgende Strategie: eine Spielerin gibt genau dann einen Tipp ab, wenn sie zwei gleichfarbige Hüte sieht, in diesem Fall tippt sie auf die Farbe, die sie nicht sieht. Mit dieser Strategie verliert das Team genau dann, wenn alle drei Hüte dieselbe Farbe haben und gewinnt andernfalls, das Team gewinnt also mit Wahrscheinlichkeit $\frac{3}{4}$.

Für das Hutproblem mit n Spielerinnen werden wir eine feste Reihenfolge der Spielerinnen annehmen und eine Zuordnung der Hutfarben mit einem Wort der Länge n identifizieren, dabei entspreche Zeichen i des Wortes der Hutfarbe von Spielerin i.

Zu einer Runde des Spiels sei das wahre Wort die in dieser Runde gewählte Zuordnung. Die Auswahl der Hutfarben durch Münzwürfe ist äquivalent dazu, dass das wahre Wort zufällig und gleichverteilt unter allen 2^n Wörtern der Länge n gewählt wird. Ist das wahre Wort fesgelegt, so gibt es genau zwei Wörter, die konsistent mit der Sicht von Spielerin j sind, das wahre Wort und das Wort, das sich vom wahren Wort genau an der Stelle j unterscheidet, wir nennen diese beiden Wörter die konsistenten Wörter von Spielerin j.

Eine Strategie legt für jede Spielerin j und für jedes mögliche Paar von konsistenten Wörter für diese Spielerin fest, ob die Spielerin einen Tipp abgeben soll und, falls ja, gemäß welchem der beiden konsistenten Wörter. Eine Strategie entspricht somit einem gerichteten Graphen, dem Strategiegraphen, dessen Knotenmenge gleich der Menge der Wörter der Länge n ist, und der genau dann eine Kante von Knoten u nach Knoten v hat, falls u und v ein mögliches Paar von konsistenten Wörtern für eine Spielerin j sind – u und v unterscheiden sich also genau an der Stelle j – und Spielerin j einen Tipp gemäß v abgibt, falls diese beiden Wörter ihre konsistenten Wörter sind.

Für ein gegebenes wahres Wort w gewinnt das Team folglich, falls im Strategiegraphen mindestens eine Kante zu w hin, aber keine Kante von w wegführt. Dies lässt sich anschaulich so beschreiben, dass der Knoten w empfängt, aber nicht sendet. Um einen hohen Anteil solcher Knoten zu erhalten, bietet es sich an, gewisse Knoten als Sender auszuwählen, die dann an alle Knoten im Abstand 1 senden, das heißt, ein solcher Knoten ist der Mittelpunkt einer Kugel mit Radius 1 und es gibt gerichtete Kanten vom Mittelpunkt zu den n Knoten auf der Kugeloberfläche. Wird dann einer der $n+1$ Knoten einer solchen Kugel als wahres Wort gewählt, verliert das Team, falls das wahre Wort gleich dem Mittelpunkt ist und gewinnt sonst. Innerhalb dieser Knoten ist das Team also bei n von $n+1$ Knoten erfolgreich. Für Werte von n der Form $n = 2^s - 1$ können wir eine entsprechende Erfolgswahrscheinlichkeit auch insgesamt erreichen, indem wir für einen perfekten 1-fehlerkorrigierenden Code wie in Theorem 223 die Codewörter zu Sender und alle anderen Wörter zu Empfängern machen. Das Team verliert somit,

falls das wahre Wort ein Codewort ist und gewinnt, falls das wahre Wort Abstand 1 *zu einem Codewort hat. Für den gewählten Code, für den die Einheitskugeln um die Codewörter eine Zerlegung der Menge aller Wörter bilden, bedeutet dies, das in jeder solchen Einheitskugel und dann auch in der Menge aller Wörter das Team nur auf einem Anteil von* $\frac{1}{n+1}$ *aller Wörter verliert und somit mit Wahrscheinlichkeit* $1 - \frac{1}{n+1}$ *gewinnt.*

4.4 Analytische Geometrie im euklidischen Raums

Der euklidische Raum In diesem Abschnitt behandeln wir geometrisch Fragen im euklidischen Raum. Wir verwenden dabei Methoden aus der Theorie der Vektorräume. Der euklidische Raum wird dazu gleich dem üblichen reellen Vektorraum $V = \mathbb{R}^3$ gesetzt, in welchem die Verktoraddition und skalare Multiplikation komponentenweise ausgeführt werden, und Geraden und Ebenen werden als Nebenklassen von ein- beziehungsweise zweidimensionale Unterverktorräumen des Vektorraums $\mathbb{R}^3$ eingeführt. Die folgenden Betrachtungen lassen sich weitgehend von $\mathbb{R}^3$ auf den Vektorraum $\mathbb{R}^n$ für eine beliebige natürliche Zahl n übertragen, zum Teil auch auf beliebige Vektorräume.

Punkte, Geraden, Ebenen

226 Definition. *Eine Nebenklasse von V der Form $v+U$ für einen Vektor v in V heißt*

PUNKT*, falls U Dimension 0 hat,*

GERADE*, falls U Dimension 1 hat und*

EBENE*, falls U Dimension 2 hat.*

Geraden und Ebenen sind Nebenklassen und somit Teilmengen von V, insbesondere sind zwei Geraden oder zwei Ebenen identisch, wenn sie dieselben Elemente enthalten. Punkte sind formal einelementige Teilmengen von V, in Folgenden werden wir aber einen Punkt der Form $v + \{o\}$ mit dem Vektor v identifizieren und zum Beispiel vom Punkt v in V sprechen. Dies entspricht unseren beiden bereits früher eingeführten Betrachtungsweisen des Vektorraums $\mathbb{R}^3$, für den wir ein Element (x, y, z) einerseits als den Punkt mit den Koordinaten x, y und z und andererseits als Vektor vom Ursprung zu diesem Punkt angesehen hatten.

227 Definition. *Für jeden eindimensionalen Untervektorraum U von V heißen Geraden der Form $v_1 + U$ und $v_2 + U$* PARALLEL.

Für jeden zweidimensionalen Untervektorraum U von V heißen Ebenen der Form $v_1 + U$ und $v_2 + U$ PARALLEL.

Für Untervektorräume U_1 der Dimension 1 und U_2 der Dimension 2 heißen eine Gerade der Form $v_1 + U_1$ und eine Ebene der Form $v_2 + U_2$ PARALLEL*, falls U_1 Untervektorraum von U_2 ist.*

228 Bemerkung. *Wir hatten bereits gesehen, dass Nebenklassen zum selben Untervektorraum eine Zerlegung des zugrundeliegenden Vektorraums bilden. Insbesondere sind je zwei Nebenklassen der Form v_1+U und v_2+U entweder*

disjunkt oder identisch. Folglich sind je zwei parallele Geraden und ebenso je zwei parallele Ebenen entweder disjunkt oder gleich.

Eine Gerade v_1+U_1 und eine Ebene v_2+U_2, die parallel sind, sind entweder disjunkt oder die Gerade ist Teilmenge der Ebene. Ist der Schnitt der Geraden und der Ebene nichtleer und enthält somit einen Punkt u, so kann die Gerade auch als $u+U_1$ und die Ebene auch als $u+U_2$ geschrieben werden. Somit ist die Gerade Teilmenge der Ebene, da nach Definition des Begriffs parallel U_1 Untervektorraum und damit insbesondere Teilmenge von U_2 ist.

229 Satz. *Zu je zwei verschiedenen Punkten u und v in V gibt es genau eine Gerade, die beide Punkte enthält, diese kann als $v+\langle u-v\rangle$ geschrieben werden.*

Zu je drei verschiedenen Punkten u, v und w, die nicht alle auf derselben Geraden liegen, gibt es genau eine Ebene, die alle drei Punkte enthält, diese kann als $v+\langle u-v, w-v\rangle$ geschrieben werden.

Beweis. Wir zeigen nur die Aussage über Geraden. Seien zunächst u und v verschiedene Punkte. Dann ist $u-v$ ungleich dem Nullvektor, das Erzeugnis von $u-v$ hat also Dimension 1 und $v+\langle u-v\rangle$ ist eine Gerade. Diese enthält die Punkte $v=v+0(v-u)$ und $u=v+1(u-v)$. Es bleibt zu zeigen, dass es keine weitere Gerade geben kann, die u und v enthält. Für eine Gerade der Form $x+\langle w\rangle$, welche die Punkte u und v enthält, gibt es Skalare a und b mit

$$u = x + aw \quad \text{und} \quad v = x + bw, \quad \text{es gilt somit} \quad u - v = (a-b)w.$$

Wegen u ungleich v ist $a-b$ ungleich 0 und $u-v$ und w erzeugen jeweils denselben Untervektorraum U. Die Geraden $v+U$ und $x+U$ sind also parallel und haben nach Annahme einen nichtleeren Schnitt, sie sind folglich nach Bemerkung 228 gleich. □

Sklararprodukt und euklidische Norm

230 Definition. *Es seien $x=(x_1,x_2,x_3)$ und $y=(y_1,y_2,y_3)$ zwei Vektoren in $\mathbb{R}^3$. Das* Skalarprodukt *von x und y ist*

$$\langle x, y\rangle \;=\; x\cdot y^{\mathrm{T}} \;=\; x_1y_1+x_2y_2+x_3y_3.$$

Die euklidische Norm *oder kurz* Norm *von x ist*

$$|x| \;=\; \sqrt{\langle x,x\rangle} \;=\; \sqrt{x_1^2+x_2^2+x_3^2}.$$

Unter Verwendung des Skalarprodukts und der Norm von Vektoren können Aussagen über Winkel, Längen und Abstände gemacht werden, insbesondere werden wir die folgenden Aussagen beweisen.

Die Norm eines Vektors ist gleich dessen Länge, wird der Vektor als Punkt aufgefasst, ist die Norm des Vektors gleich dem Abstand des Punkts vom Ursprung.

Das Skalarprodukt zweier Vektoren ist genau dann gleich 0, wenn beide Vektoren senkrecht zueinander sind.

Aus den Längen und dem Skalarprodukt zweier Vektoren lässt sich der Winkel zwischen den beiden Vektoren bestimmen.

231 Bemerkung. *Aus der Definition des Skalarprodukts $\langle x, y\rangle$ über die Matrixmultiplikation der Vektoren x und y^{T} ergibt sich sofort, dass das Skalarprodukt in dem Sinn in beiden Argumenten linear ist, dass für alle Skalare a und alle Vektoren x, y und z gilt*

$$\begin{aligned} a \cdot \langle x, y\rangle &= \langle ax, y\rangle = \langle x, ay\rangle, \\ \langle x + y, z\rangle &= \langle x, z\rangle + \langle y, z\rangle, \\ \langle x, y + z\rangle &= \langle x, y\rangle + \langle x, z\rangle. \end{aligned}$$

Insbesondere gilt $\langle ax, ax\rangle = a^2\langle x, x\rangle$, für die Norm folgt daraus $|ax| = a|x|$.

Die Norm eines Vektors ist gleich seiner Länge Im Folgenden sei ein anschauliches Verständnis von Längen und Winkeln in der euklidischen Ebene und im euklidischen Raum und damit auch in den Vektorräumen $\mathbb{R}^2$und $\mathbb{R}^3$ vorausgesetzt. Die folgenen Sätze zeigen, dass sich diese anschaulichen Begriffe durch Skalarprodukt und Norm formal dargestellt werden können.

Als Erstes zeigen wir mit dem Satz des Pythagoras dass die Länge eines Vektors im $\mathbb{R}^3$ gleich dessen Norm ist. In einem rechtwinkligen Dreieck gibt es eine längste Seite, diese wird als Hypotenuse bezeichnet, die beiden kürzeren Seiten als Katheten, der rechte Winkel wird immer durch die beiden Katheten gebildet.

232 Satz des Pythagoras. *In einem rechtwinkligen Dreieck mit Katheten der Länge a und b und einer Hypotenuse der Länge c gilt $c^2 = a^2 + b^2$.*

Beweis. Folgt. □

233 Lemma. *In einem Rechteck mit Seitenlängen a und b hat die Diagonale Länge $\sqrt{a^2 + b^2}$. In einem Quader mit Seitenlängen a, b und c hat die Raumdiagonale Länge $\sqrt{a^2 + b^2 + c^2}$.*

Beweis. Die Diagonale in einem Rechteck ist Hypotenuse eines rechtwinkligen Dreiecks dessen Katheten gleich zwei aneinanderliegenden Seiten des Rechtecks sind. Sind die Seitenlängen des Rechtecks a und b, so hat die

Diagonale nach dem Satz des Pythagoras die Länge $\sqrt{a^2+b^2}$. Die Raumdiagonale eines Quaders ist Hypotenuse eines rechtwinkligen Dreiecks, dessen eine Kathete gleich der Diagonale einer Seitenfläche des Quaders ist und dessen andere Kathete gleich der Kante des Quaders ist, die senkrecht auf dieser Seitenfläche steht. Hat die Seitenfläche Kanten der Länge a und b und ist c die Länge der dazu senkrechten Kante, so sind die quadrierten Längen dieser beiden Katheten gleich a^2+b^2 und c^2, die Raumdiagonale hat folglich Länge $\sqrt{a^2+b^2+c^2}$. □

234 Satz. *Im euklidischen Raum ist der Abstand eines Punkts x vom Ursprung gleich $|x|$. Entsprechend ist die Norm $|x|$ eines Vektors x in $\mathbb{R}^3$ gleich der Länge des Vektors.*

Beweis. Für einen Punkt $x = (x_1, x_2, x_3)$ ist der Abstand von Ursprung gleich der Länge der Raumdiagonalen eines Quaders mit Seitenlängen x_1, x_2 und x_3. Die Behauptung folgt unmittelbar aus Lemma 233 und der Definition der Norm. □

Skalarprodukt und rechte Winkel Wir geben die Größe eines Winkel in Grad an, beispielsweise hat ein rechter Winkel 90 Grad, kurz: 90°, und ein gestreckter Winkel 180 Grad.

235 Definition. *In der euklidische Ebene und entsprechend in $\mathbb{R}^2$ heißt die Menge aller Punkte mit Abstand 1 zum Ursprung* Einheitskreis. *Ein* Einheitsvektor *ist ein Vektor mit Länge 1, das heißt, im $\mathbb{R}^2$ entsprechen die Einheitsvektoren den Punkten auf dem Einheitskreis.*

236 Bemerkung. *In den meisten Gebieten der Mathematik ist es üblich, die Größe von Winkeln nicht in Grad, sondern im Bogenmaß anzugeben. Ist φ der Winkel zwischen dem Einheitvektor $(1,0)$ und einem weitere Einheitsvektor im $\mathbb{R}^2$, so wird der Winkel zwischen diesen beiden Vektoren, die ja gleich Punkten auf dem Einheitskreis sind, durch die Länge des Kreissegments des Einheitskreises zwischen diesen beiden Schnittpunkten angegeben, in mathematisch positive Richtung, also gegen den Uhrzeigersinn vom Punkt $(1,0)$ auf der x-Achse aus. Der Einheitskreis hat eine Umfang von 2π, entsprechend haben beispielsweise Winkel von 45, 90 und 180 Grad ein Bogenmaß von $\frac{\pi}{4}$, $\frac{\pi}{2}$ beziehungweise π. Ein negatives Bogenmaß bedeutet, dass die Länge des Kreissegments in mathematisch negativer Richtung, also im Uhrzeigersinn, gemessen wird. Ein Winkel von 180 Grad hat somit sowohl Bogenmaß π als auch Bogenmaß $-\pi$.*

237 Vereinbarung. *Der Winkel zwischen zwei Vektoren x und y des Vektorraums $\mathbb{R}^3$ ist der Winkel beim Ursprung zwischen den beiden Halbgeraden des euklidischen Raums die im Ursprung beginnen und den Punkt x beziehungsweise y enthalten. Im Rest dieses Kapitels sei dabei immer der kleinere*

der beiden auftretenden und sich zu 360 *Grad addierenden Winkeln gemeint. Durch diese Festlegung liegt der Winkel zwischen zwei Vektoren immer zwischen* 0 *und* 180 *Grad. Beispielsweise bildet der Einheitsvektor* $(1,0,0)$ *mit den Einheitsvektoren* $(0,1,0)$ *und* $(0,-1,0)$ *jeweils einen Winkel von* 90 *Grad und nicht von* 270 *Grad. Zwei Vektoren, die in diesem Sinne einen Winkel von* 90 *Grad bilden, heißen zueinander* SENKRECHT *oder* ORTHOGONAL.

238 Bemerkung. *Es gilt auch folgende Umkehrung des Satzes des Pythagoras: gilt für die Seitenlängen* a, b *und* c *eines Dreiecks* $c^2 = a^2 + b^2$, *so ist das Dreieck ein rechtwinkliges Dreieck mit einer Hypotenuse der Länge* c. *Die Begründung für diese Implikation ist einfach: sind drei solche Zahlen gegeben, so gibt es ein rechtwinkliges Dreieck mit Katheten der Länge* a *und* b. *Nach dem Satz des Pythagoras hat dieses Dreieck eine Hypotenuse der Länge* c. *Da die Form eines Dreiecks durch die Längen der drei Seiten eindeutig bestimmt ist, muss jedes Dreieck mit Seiten der Längen* a, b *und* c *gleich diesem rechtwinkligen Dreieck sein.*

239 Satz. *Zwei Vektoren im* $\mathbb{R}^3$ *sind genau dann senkrecht zueinander, wenn ihr Skalarprodukt gleich* 0 *ist.*

Beweis. Es seien $x = (x_1, x_2, x_3)$ und $y = (y_1, y_2, y_3)$ zwei Vektoren in $\mathbb{R}^3$. Wir betrachten das Dreieck dessen Eckpunkte gleich x, y und dem Ursprung sind. Die Seiten dieses Dreiecks haben die Längen $a = |x|$, $b = |y|$ und $c = |x - y|$ mit $x - y = (x_1 - y_1, x_2 - y_2, x_3 - y_3)$. Es folgt

$$\begin{aligned}
c^2 = |x-y|^2 &= \langle x-y, x-y\rangle \\
&= (x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2 \\
&= x_1^2 - 2x_1y_1 + y_1^2 \;+\; x_2^2 - 2x_2y_2 + y_2^2 \;+\; x_3^2 - 2x_3y_3 + y_3^2 \\
&= \underbrace{x_1^2 + x_2^2 + x_3^2}_{=|x|^2=a^2} \;+\; \underbrace{y_1^2 + y_2^2 + y_3^2}_{=|y|^2=b^2} \;-\; 2\underbrace{(x_1y_1 + x_2y_2 + x_3y_3)}_{=\langle x,y\rangle} \\
&= a^2 + b^2 - 2\langle x, y\rangle.
\end{aligned}$$

Die Vektoren x und y sind genau dann senkrecht zueinander, wenn das betrachtete Dreieck einen rechten Winkel beim Ursprung hat. Letzteres gilt gemäß dem Satz des Pythagoras und Bemerkung 238 genau dann, wenn c^2 gleich $a^2 + b^2$ ist, was wiederum nach der gerade bewiesenen Gleichungskette genau dann der Fall ist, wenn $\langle x, y\rangle$ gleich 0 ist. Der Satz folgt. □

Skalarprodukt und beliebige Winkel. Die Winkelfunktionen Sinus und Kosinus spielen in vielen Gebieten der Mathematik eine wichtige Rolle. Der Einfachheit halber definieren wir im Folgenden nur den Kosinus und diesen auch nur für Winkel zwischen 0 und 180 Grad. Dies genügt für unsere Zwecke, da gemäß unserer Festlegung der Winkel zwischen zwei Vektoren immer in diesem Bereich liegt.

240 Definition. *Der Einheitsvektor* $(1,0)$ *bilde mit einem Einheitsvektor der Form* (x_1,x_2) *mit* $x_2 \geq 0$ *einen Winkel* φ*, dieser liegt gemäß Vereinbarung 237 zwischen* 0 *und* 180 *Grad. Dann wird* x_1 *als* KOSINUS *des Winkels* φ *bezeichnet, kurz:* $\cos\varphi$*.*

241 Bemerkung. *Die in Definition 240 eingeführte Kosinusfunktion* $\varphi \mapsto \cos\varphi$ *ist eine Bijektion zwischen der Menge der Winkel zwischen* 0 *und* 180 *Grad und dem abgeschlossenen reellen Intervall* $[-1,1]$*. Insbesondere ist die Abbildung invertierbar: zu jeder reellen Zahl* a *zwischen* -1 *und* 1 *gibt es einen eindeutig bestimmten Winkel* φ *mit* $a = \cos\varphi$*.*

Definition 240 ist eine Einschränkung der üblichen Definition des Kosinusfunktion als Funktion auf den reellen Zahlen. Diese bildet das reelle Bogenmass eines Winkels zwischen zwei Einheitsvektoren $(1,0)$ *und* (x_1,x_2) *auf den Kosinuswert* x_1 *ab.*

242 Bemerkung. *Es sei* $e = (x_1,x_2)$ *ein Einheitsvektor der mit dem Einheitsvektor* $(1,0)$ *einen Winkel* φ *bildet. Nach Definition ist dann der Kosinus von* φ *gleich der Projektion* x_1 *von* (x_1,x_2) *auf die* x*-Achse, diese ist anschaulich gleich der durch den Vektor* $(1,0)$ *definierte Gerade.*

Für einen beliebigen Vektor v *in, der ein Vielfaches eines Einheitsvektors der Form* (x_1,x_2) *mit* $x_2 \geq 0$ *ist und der folglich mit dem Einheitsvektor* $(0,1)$ *ebenfalls den Winkel* φ *bildet, gilt*

$$v \;=\; |v|(x_1,x_2), \;=\; (|v|\cdot x_1, |v|\cdot x_2),$$

folglich hat die Projektion von v *auf die* x*-Achse die Länge* $|v|\cdot x_1 = |v|\cdot\cos\varphi$*.*

Das bisher Gesagte überträgt sich auf Winkel zwischen Vektoren im euklidischen Raum indem statt diesen zwei Vektoren in der euklidischen Ebene derselben Länge betrachtet werden, die denselben Winkel bilden und von denen einer ein Vielfaches des Einheitsvektors $(1,0)$ *ist.*

243 Lemma. *Es seien* x *und* y *Einheitsvektoren im* $\mathbb{R}^3$*, der Winkel zwischen den beiden Vektoren sei gleich* φ*. Dann gilt* $\langle x,y\rangle = \cos\varphi$*.*

Beweis. Wir zeigen zunächst, dass die Aussage für φ gleich 0, 90 oder 180 Grad wahr ist. Für eine Winkel von 90 Grad ist der Kosinus gleich 0, die Aussagen gilt also gemäß Satz 239. Im Fall 0 Grad sind x und y identisch, das Skalarprodukt der beiden Vektoren ist also gleich $\langle x,y\rangle = |x|^2 = 1$ und damit gleich dem Kosinus von 0 Grad, ganz ähnlich ergibt sich im Fall von 180 Grad, dass beide Werte gleich -1 sind.

Wir können also ab jetzt voraussetzen, dass die Größe von φ ungleich 0, 90 und 180 Grad ist. Wir führen den Beweis zuächst für den Fall $0° < \varphi < 90°$ und betrachten dazu den Vektor

$$\widetilde{y} \;=\; \cos\varphi\cdot y, \qquad \text{für diesen gilt} \quad |\widetilde{y}| \;=\; \cos\varphi|y| \;=\; \cos\varphi.$$

und das durch die Vektoren x und $\widetilde{y}$ aufgespannte Dreieck. Dieses Dreieck ist in dem Sinne nicht entartet, dass alle drei Seiten Länge ungleich 0 haben und alle drei Innenwinkel mehr als 0 Grad betragen. Dies folgt, da nach Fallannahme $\cos\varphi$ echt größer 0 ist, folglich ist mit y auch $\widetilde{y}$ ungleich dem Nullvektor und beide Vektoren bilden mit x denselben Winkel φ, dieser beträgt nach Fallannahme weder 0 noch 180 Grad.

Der Vektor $\widetilde{y}$ hat die Norm $\cos\varphi$ und fällt deshalb mit der Projektion des Vektors x auf die Gerade $o + \langle y\rangle = o + \langle\widetilde{y}\rangle$ zusammen, das Dreieck ist folglich rechtwinklig, mit Hypotenuse x, einer Kathete $\widetilde{y}$ und einer weiteren Kathete $x - \widetilde{y}$. Für die beiden zueinander senkrechten Katheten gilt dann

$$\langle\widetilde{y}, x - \widetilde{y}\rangle = \langle x, \widetilde{y}\rangle - \langle\widetilde{y}, \widetilde{y}\rangle = 0 \quad \text{und somit} \quad \langle x, \widetilde{y}\rangle \;=\; \langle\widetilde{y}, \widetilde{y}\rangle \;=\; |\widetilde{y}|^2. \quad (4.6)$$

Wegen $\widetilde{y} = \cos\varphi \cdot y$ und der Linearität von Skalarprodukt und Norm folgt

$$\cos\varphi \cdot \langle x, y\rangle \;=\; \langle x, \cos\varphi \cdot y\rangle \;=\; \langle x, \widetilde{y}\rangle \overset{(4.6)}{=} |\widetilde{y}|^2 \;=\; \cos\varphi \cdot \cos\varphi.$$

Indem der erste und letzte Term in dieser Gleichungskette mit $\cos\varphi^{-1}$ multipliziert werden, ergibt sich $\langle x, y\rangle = \cos\varphi$ und das Lemma ist für den Fall $0° < \varphi < 90°$ gezeigt. Den noch verbleibenden Fall $90° < \varphi < 180°$ zeigen wir, indem wir für eine solchen Winkel φ und Einheitsvektoren x und y den schon bewiesenen Fall auf die Vektoren x und $y' = -y$ und den zugehörigen Winkel $\varphi' = 180° - \varphi$ mit $\cos\varphi' = -\cos\varphi$ anwenden, es folgt

$$\langle x, y\rangle \;=\; -\langle x, y'\rangle \;=\; -\cos\varphi' \;=\; \cos\varphi.$$

□

244 Satz. *Es seien x und y Vektoren im $\mathbb{R}^3$ die beide vom Nullvektor verschieden sind, und der Winkel zwischen den beiden Vektoren sei gleich φ. Dann gilt*

$$\langle x, y\rangle \;=\; |x| \cdot |y| \cdot \cos\varphi. \quad (4.7)$$

Beweis. Wir betrachten die Einheitsvektoren

$$x' \;=\; \frac{x}{|x|} \qquad \text{und} \qquad y' \;=\; \frac{y}{|y|},$$

diese sind normierte Versionen der Vektoren x und y und bilden insbesondere ebenfalls den Winkel φ. Mit Lemma 243 folgt

$$\langle x, y\rangle \;=\; \left\langle |x|\frac{x}{|x|}, |y|\frac{y}{|y|}\right\rangle \;=\; |x| \cdot |y| \cdot \langle x', y'\rangle \;=\; |x| \cdot |y| \cdot \cos\varphi.$$

□

4.5 Lineare Abbildungen

Lineare Abildungen

245 Definition. *Seien V und W K-Vektorräume. Eine Funktion $g\colon V \to W$ heißt* LINEARE ABBILDUNG *oder* HOMOMORPHISMUS *von V nach W, wenn für alle u und v in V und alle Skalare a gilt*

$$g(u+v) \;=\; g(u)+g(v) \;\; \textit{(Additivität)} \quad \textit{und} \quad g(au) \;=\; ag(u) \;\; \textit{(Linearität).}$$

246 Satz. *Es seien V und W K-Vektorräume und g sei eine lineare Abbildung von V nach W, weiter sei $B = \{v_1, \ldots, v_t\}$ eine Basis von V. Dann ist die Abbildung g bereits durch die Bilder der Basis B unter g eindeutig bestimmt. Dies bedeutet: je zwei lineare Abbildungen von V nach W, die auf B übereinstimmen, sind gleich.*

Beweis. Sei u ein beliebiger Vektor aus V. Der Vektor u kann als Linearkombination $u = a_1v_1 + \cdots + a_tv_t$ über der Basis B geschrieben werden. Für die lineare Abbildung g gilt nach Definition

$$g(u) \;=\; g(a_1v_1 + \cdots + a_tv_t) \;=\; a_1\, g(v_1) + \cdots + a_t\, g(v_t),$$

die Abbildung g ist also bereits durch die Werte $g(v_1)$ bis $g(v_t)$ festgelegt. Für jede andere lineare Abbildung g' bleibt diese Gleichungskette gültig, wenn überall g durch g' ersetzt wird. Falls g und g' auf den Vektoren in B übereinstimmen, steht am Ende der beiden Gleichungsketten jeweils derselbe Vektor, folglich sind auch die Terme $g(u)$ und $g'(u)$ am Anfang der beiden Gleichungsketten identisch, es gilt $g(u) = g'(u)$. Da der Vektor u beliebig aus V gewählt war, sind die Abbildungen g und g' gleich. □

247 Satz. *Es seien V und W K-Vektorräume und $B = \{v_1, \ldots, v_t\}$ eine Basis von V. Für beliebige Vektoren $w_1, \ldots, w_t$ in W gibt es genau eine lineare Abbildung g von V nach W mit*

$$g(v_i) \;=\; w_i \quad \textit{für} \quad i = 1, \ldots, t.$$

Beweis. Das es höchstens eine solche lineare Abbildunge geben kann ist gerade die Aussage von Satz 246. Um zu zeigen, dass es eine solche Abbildung g gibt, genügt es, für einen beliebigen Vektor $u = a_1v_1 + \cdots + a_tv_t$ aus V

$$g(u) \;=\; g(a_1v_1 + \cdots + a_tv_t) \;=\; a_1\, g(v_1) + \cdots + a_t\, g(v_t)$$

zu setzen. Diese Abbildung ist als Abbildung auf V wohldefiniert, da die Darstellung eines Vektors u in V als Linearkombination über B eindeutig ist. Es bleibt noch die Additivität und Linearität nachzurechnen. Wir zeigen die Linearität. Seien u und v beliebige Vektoren aus V, die über der Basis B

als Linearkombinationen $u = a_1v_1 + \cdots + a_tv_t$ und $v = b_1v_1 + \cdots + b_tv_t$ dargestellt werden können. Dann gilt

$$\begin{aligned} g(u+v) &= g((a_1+b_1)v_1 + \cdots + (a_t+b_t)v_t) \\ &= (a_1+b_1)g(v_1) + \cdots + (a_t+b_t)g(v_t) \\ &= a_1g(v_1) + \cdots + a_tg(v_t) \ + \ b_1g(v_1) + \cdots + b_tg(v_t) \\ &= g(u) \ + \ g(v). \end{aligned}$$

□

Darstellung durch Matrizen

248 Definition. *Es seis V ein K-Vektorraum mit Basis $B = \{v_1, \ldots, v_t\}$. Für einen $(1,t)$-Vektor $u = (a_1, \ldots, a_t)$ über K sei u^B gleich dem Vektor aus V in dessen Darstellung als Linearkombination über B der Basisvektor v_i den Koeffizienten a_i hat, das heißt*

$$u^B \ = \ a_1v_1 + \cdots a_tv_t.$$

Ein solcher Vektor u heiße Koeffizientenvektor *zur Basis B des Vektores u^B, während u^B der von u* dargestellte Vektor *ist.*

Falls sich die Basis B aus dem Kontext ergibt oder irrelevant ist, schreiben wir u statt u^B. Dies hatten wir schon für den Vektorraum $\mathbb{R}^3$ so gehandhabt: für die Schreibweise $x = (x_1, x_2, x_3)$ für Vektoren aus $\mathbb{R}^3$ steht x für x^B, wobei die Basis B aus den Einheitsvektoren $e_1 = (1,0,0)$, $e_1 = (0,1,0)$, und $e_1 = (0,0,1)$ besteht.

249 Bemerkung. *Es sei B eine Basis für einen K-Vektorraum. Für alle Koeffizientenvektoren u, u_1 und u_2 zur Basis B und alle Skalare a gilt*

$$(u_1+u_2)^B \ = \ u_1^B + u_2^B \quad \textit{und} \quad (au)^B \ = \ au^B.$$

Mit Koeffizientenvektoren kann eine lineare Abbildung zwischen zwei Vektorräumen in einfacher Weise durch eine Matrix dargestelllt werden. Seien dazu V und W K-Vektorräume mit Basen $B_V = \{v_1, \ldots, v_s\}$ beziehungsweise $B_W = \{w_1, \ldots, w_t\}$. Weiter sei g eine lineare Abbildung von V nach W. Wir konstruieren eine (t,s)-Matrix A, so dass für Koeffizientenvektoren die Multiplikation mit A der Abbildung g in folgendem Sinne entspricht: für jeden Koeffizientenvektor u eines Vektors aus V zur Basis B_V ist der Vektor Au gleich dem Koeffizientenvektor von $g(u^{B_V})$ zur Basis B_W, das heißt

$$(Au)^{B_W} \ = \ g(u^{B_V}). \tag{4.8}$$

Wir setzen für $i = 1, \ldots, s$ Spalte i von A gleich dem Koeffizientenvektor c_i von $g(v_i)$ über B_W. Es sei u_i der Koeffizientenvektor des Vektor v_i über B_V,

dieser hat Länge s, Eintrag i ist gleich 1 und alle anderen Einträge sind gleich 0. Mit der Spaltendarstellung der Matrixmultiplikation folgt somit, dass Au_i gleich Spalte i von A ist, also gleich dem Koeffizientenvektor c_i des Vektors $g(v_i) = g(u_i^{B_V})$. Damit gilt 4.8 für u_i anstelle von u. Da die Abbildung g linear ist, folgt mit den Regeln der Matrixmultiplikation, dass 4.8 für alle u aus V gilt. Für jeden Koeffizientenvektor $u = (a_1, \ldots, a_s)$ über B_V gilt

$$Au \;=\; A(a_1u_1 + \ldots + a_su_s) \;=\; a_1(Au_1) + \ldots + a_s(Au_s) \;=\; a_1c_1 + \ldots + a_sc_s.$$

Da (4.8) für die u_i gilt, folgt mit Bemerkung 249

$$\begin{aligned}(Au)^{B_W} &= (a_1c_1 + \ldots + a_sc_s)^{B_W} \\ &= a_1c_1^{B_W} + \ldots + a_sc_s^{B_W} \\ &= a_1g(v_1) + \ldots + a_sg(v_s) \\ &= g(a_1v_1 + \ldots + a_sv_s) \\ &= g(u^{B_v}).\end{aligned}$$

Lineare Abbildungen im $\mathbb{R}^2$ und $\mathbb{R}^3$ Wir betrachten einige Beispiel für lineare Abbildungen der euklidischen Ebene und des euklidischen Raums. Wie auch bisher schreiben wir dabei zum Beispiel (x_1, x_2, x_3) für einen Punkt des euklidischen Raums und fassen dieses Tripel als Koeffizientenvektor zur Standardbasis auf. Lineare Abbildungen werden dann durch Multiplikation von Koeffizientenvektoren mit einer Matrix dargestellt. Die Spaltenvektoren dieser Matrix sind, von links nach rechts, gerade die Transponierten der Bilder der Vektoren e_1, e_2, e_3 der Standardbasis.

250 Beispiel. *Die drei Matrizen stellen lineare Abbildungen des euklidischen Raums dar: die Matrix A_1 die Identität, die Matrix A_2 die Spiegelung an der x-y-Ebene und die Matrix A_3 eine Punktspiegelung am Ursprung mit zusätzlicher Streckung mit den Faktor* 2*.*

$$A_1 \;=\; \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}, \qquad A_2 \;=\; \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}, \qquad A_3 \;=\; \begin{pmatrix}-2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -2\end{pmatrix}.$$

251 Beispiel. *Die ersten beiden Matrizen stellen lineare Abbildungen der euklidischen Ebene dar: die Matrix A_1 eine Drehung um* 90 *Grad im positiven Sinn, das heißt, gegen den Uhrzeigersinn, die Matrix A_2 eine Drehung um den Winkel φ im positiven Sinn. Die Matrix A_3 stellt eine Drehung um die z-Achse im positiven Sinn um den Winkel φ gegen der Uhrzeigersinn dar, mit zusätzlicher Streckung mit dem Faktor* 3 *vom Ursprung aus in die Richtungen der x-y-Ebene und einer Spiegelung an der x-y-Ebene*

$$A_1 \;=\; \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix},\; A_2 \;=\; \begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix},\; A_3 \;=\; \begin{pmatrix}3\cos\varphi & -3\sin\varphi & 0\\ 3\sin\varphi & 3\cos\varphi & 0\\ 0 & 0 & -1\end{pmatrix}.$$

5 Anhang: Beispielrechnungen

Dieser Anhang enthält einige Beispiele für Probleme im Zusammenhang mit linearen Gleichungssystemen und der analytischen Geometrie des euklidischen Raums. Wir schreiben im Folgenden Vektoren im $\mathbb{R}^3$ als Zeilenvektoren.

252 Beispiel (Schnitt einer Gerade mit einer Ebene)**.** *Gesucht ist der Schnitt der*

Ebene $(0,0,2) + \langle (1,0,0), (0,1,0) \rangle$ *mit der Geraden* $(0,0,0) + \langle (1,1,1) \rangle$.

Wie setzen dazu einen allgemeinen Punkt der Ebene mit einen allgemeinen Punkt der Geraden gleich, die Koeffizientenseien x_1 *und* x_2 *für die Ebene und* x_3 *für die Gerade. Dies ergibt*

$$(0,0,2) + x_1(1,0,0) + x_2(0,1,0) \;=\; (0,0,0) + x_3(1,1,1).$$

Nach einer Umstellung zu

$$x_1(1,0,0) + x_2(0,1,0) - x_3(1,1,1) \;=\; (0,0,0) - (0,0,2)$$

ergibt sich daraus für $x = (x_1, x_2, x_3)^{\mathrm{T}}$ *das inhomogene lineare Gleichungssystem* $Ax = b$ *für die erweiterte Matrix* $(A|b)$

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & -2 \end{array}\right)$$

Die Anwendung des Gaußschen Verfahren auf diese Matrix ergibt

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & -2 \end{array}\right) \quad \to \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & -1 & -2 \end{array}\right) \quad \to \quad \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

Die Matrix A hat somit vollen Rang 3*, die Lösungsmenge* $\mathrm{L}(A, o)$ *des homogenen linearen Gleichungsssystems hat als Untervektorraum Dimension* $d = n - \mathrm{rang}(A) = 0$*, enthält also nur den Nullvektor. Die Lösungsmenge* $\mathrm{L}(A, b)$ *ist somit leer oder enthält genau einen Vektor. Letzteres ist der Fall, da die Matrix A und die erweiterten Matrix* $(A|b)$ *denselben Rang haben, das inhomogone lineare Gleichungssystem also lösbar ist.*

Die einzige Lösung $c = (c_1, c_2, c_3)$ *ergib sich, indem die* c_i *so gewählt werden, dass die zugehörigen Gleichungen wie* $c_1 \cdot 1 + c_2 \cdot 0 + c_3 \cdot 0 = 2$ *erfüllt werden. Dies führt auf* $c = (2, 2, 2)$*.*

253 Beispiel (Abstand eines Punkts zu einer Ebene)**.** *Wir wollen im euklidischen Raum den Abstand des Punkts* $z = (2, -3, 3)$ *zur Ebene*

$$v + \langle u_1, u_2 \rangle \qquad \textit{mit} \qquad v = (0,0,0), u_1 = (1,1,1) \textit{ und } u_2 = (0,1,1)$$

bestimmen. Der Abstand ergibt sich dabei als Länge des Vektors s*, der senkrecht auf der Ebene steht und, anschaulich gesprochen, von einem Punkt* z_0 *der Ebene zu* z *führt. Letzteres bedeutet formal, dass* s *gleich* $z - z_0$ *ist.*

Wir bestimmen zunächst einen Vektor $c = (c_1, c_2, c_3)$*, der senkrecht auf der Ebene steht. Dies ist genau dann der Fall, wenn das Skalarprodukt von* c *mit* u_1 *und* u_2 *jeweils gleich* 0 *ist. Dies führt auf die beiden Gleichungen*

$$\begin{aligned} \langle u_1, c \rangle &= 1 \cdot c_1 + 1 \cdot c_2 + 1 \cdot c_3 = 0 \quad \textit{und} \\ \langle u_2, c \rangle &= 0 \cdot c_1 + 1 \cdot c_2 + 1 \cdot c_3 = 0. \end{aligned}$$

Als senkrechter Vektor c *kann also jeder Vektor ungleich dem Nullvektor in der Lösungsmenge des homogenen linearen Gleichungssystem* $Ac = o$ *oder, äquivalent dazu, des Gleichungssystem* $A'c = o$ *gewählt werden für*

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \textit{und} \qquad A' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Die Matrix A' *ergibt sich dabei durch Anwendung des Gaußschen Verfahrens auf die Matrix* A*, tatsächlich wird dabei nur die zweite Zeile von der ersten abgezogen.*

Die Lösungsmenge des homogenen Gleichungssystem Ac *hat als Untervektorraum Dimension* $d = n - \operatorname{rang}(A) = 1$*. Wir erhalten einen Basisvektor* c *dieses Untervektorraums durch den Ansatz* $c = (c_1, c_2, 1)$*, wobei die* c_i *so gewählt werden, dass die beiden Gleichungen des Systems erfüllt werden. Dies führt auf* $c_1 = 0$ *und* $c_2 = -1$*, wir setzen* $c = (0, -1, 1)$*.*

Zur Bestimmung des Abstands machen wir nun folgenden Ansatz, z_0 *ist dabei des Fußpunkt des Lots von* z *auf die betrachtete Ebene, der Vektor* $z - z_0$ *ist entsprechend ein Vielfaches der Senkrechten* c*.*

$$z_0 = v + x_1 u_1 + x_2 u_2 \qquad \textit{und} \qquad z = z_0 + x_3 c.$$

Durch Einsetzen der ersten Gleichung in die zweite und Umstellen ergibt sich die Gleichung

$$x_1 u_1 + x_2 u_2 + x_3 c = z,$$

diese ist äquivalent zum inhomogenen linearen Gleichungssystem

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 1 & 1 & -1 & -3 \\ 1 & 1 & 1 & 3 \end{array}\right).$$

Durch die Umformungen des Gaußschen Verfahrens werden nacheinander folgende erweiterte Matrizen berechnet, wobei nicht alle Zwischenergebnisse angegeben sind.

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 1 & 1 & -1 & -3 \\ 1 & 1 & 1 & 3 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & -1 & -5 \\ 0 & 1 & 1 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & -1 & -5 \\ 0 & 0 & 2 & 6 \end{array}\right)$$
$$\rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 2 & 6 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Das inhomogene System hat somit nur die Lösung $x_1 = 2,\ x_2 = -2$ *und* $x_3 = 3$. *Der gesuchte Abstand von* z *zur Ebene war als Länge des Vektors* $x_3 \cdot c$ *angesetzt und ist folglich gleich*

$$x_3|c| = 3\sqrt{c_1^2 + c_2^2 + c_3^2} = 3\sqrt{0+1+1} = 3\sqrt{2} = \sqrt{18}.$$

Die Lösung des Gleichungssystems liefert auch den Fußpunkt z_0 *des Lots von* z *auf die Ebene, dieser ist gleich*

$$z_0 = x_1 u_1 + x_2 u_2 = 2 \cdot (1,1,1) - 2 \cdot (0,1,1) = (2,0,0).$$