Einführung in das mathematische Arbeiten

Hermann Schichl
Roland Steinbauer

Fakultät für Mathematik
Universität Wien
Sommersemester 2007

Inhaltsverzeichnis

KAPITEL 1

Einleitung

Im Vergleich mit vielen anderen Studienrichtungen, selbst mit den naturwissenschaftlichen, hat das Mathematikstudium eine höhere Drop-Out–Rate, und viele Studierende geben bereits bald nach Studienbeginn auf.

Ein Hauptgrund für dieses Faktum liegt darin, dass die Art wie Mathematik an der Universität betrieben wird sich grundlegend von dem aus der Schule gewohnten unterscheidet. Während in der Schule das Hauptaugenmerk auf das Lösen von Beispielen gerichtet ist und für die meisten LehrerInnen das Algorithmische im Vordergrund steht (das Erlernen von Schemata zur Behandlung von Standardproblemen), tritt dies an der Universität merklich in den Hintergrund. Es ist in Wahrheit so, dass selbst die besten Fähigkeiten in diesem Gebiet nicht ausreichen, ein Mathematikstudium, sei es zum Lehramt oder zum Diplom, erfolgreich abzuschließen.

In der Vergangenheit hat die Erfahrung gezeigt, dass bereits in der Studieneingangsphase (in den ersten wenigen Wochen) zwei Fakten zu einer Fehleinschätzung des Studiums durch die Studierenden führen.

(1) *Die scheinbare Einfachheit des zu Beginn gelehrten Stoffes* — der Stoff, der in den Vorlesungen zu Beginn vorgetragen wird, scheint den meisten wohlbekannt und leicht verständlich. Dies verführt dazu, sich zu Beginn auf dem in der Schule gelernten „auszuruhen“ und den Punkt zu verschlafen, an dem der sichere Hafen des bereits Erlernten verlassen wird. Der Stoff sieht nämlich nur auf den ersten Blick einfach aus, denn **die wahre Schwierigkeit liegt nicht darin *was* behandelt wird sondern *wie* es behandelt wird**. Es wird Ihnen also helfen die scheinbare Einfachheit am Beginn dazu zu nützen zu verstehen, *wie* der Stoff präsentiert wird und warum das gerade *so* geschieht.

(2) Der *Abstraktionsschock* hängt unmittelbar mit dem zuvor gesagten zusammen. Während in der Schule die meisten LehrerInnen Mathematik an Hand von Beispielen erklären und weiterentwickeln, ja der gesamte Unterricht meist darauf fokussiert wird, dienen in der höheren Mathematik Beispiele vor allem dazu Sachverhalte zu illustrieren. Die wahre Entwicklung erfolgt innerhalb abstrakter Strukturen; diese werden durch möglichst wenige grundlegende Attribute **definiert**, und weitere gültige Eigenschaften sowie Querbeziehungen zu anderen Strukturen werden in **Beweisen** mittels logischer Schlussfolgerungen aus diesen Grundlagen und bereits bekannten Tatsachen abgeleitet.

Einer der häufigsten Fehler von StudienanfängerInnen liegt darin, den Beweisen nicht die nötige Aufmerksamkeit zukommen zu lassen. Das heißt den wahren Geist der Mathematik zu verkennen und die wahren Schwierigkeiten, besonders am Anfang, zu übersehen. Zusätzlich führt es dazu, dass bereits nach wenigen Wochen des Studiums die geschaffenen Strukturen einen Umfang und ein Abstraktionsniveau erreicht haben, das sich mit Schulwissen und Beispielen allein nicht mehr überblicken lässt. Mitlernen und Hinterfragen des Gehörten bereits zu Beginn des Studiums helfen, den Schock zu verringern oder gar zu verhindern.

Dieses Buch ist aus Skripten zur Lehrveranstaltung „Einführung in das mathematische Arbeiten“ entstanden, die an der Universität Wien im Studienplan Mathematik mit der Absicht eingeführt wurde, eine Brücke zwischen der Mittel- und der Hochschulmathematik zu schlagen. Buch und Lehrveranstaltung sollen also dazu dienen, die StudienanfängerInnen an die abstrakte Art und Weise, in der Mathematik an Universitäten gelehrt wird, zu gewöhnen und außerdem die Studierenden auf ein annähernd einheitliches Wissensniveau zu führen, das auf Grund unterschiedlicher Lehrpläne in den verschiedenen Schultypen zu Studienbeginn nicht gegeben ist.

1.1. Hürden zu Studienbeginn

Das Mathematikstudium bietet den meisten Studierenden zu Beginn einige grundlegende Hürden, die in diesem Kapitel angesprochen werden sollen.

1.1.1. „Buchstabenrechnen“ versus „Zahlenrechnen“ — Abstraktion. Zahlen spielen im Mathematikstudium eine gegenüber der Schule untergeordnete Bedeutung. Reines Rechnen ist kein grundlegender Bestandteil des Lehrstoffs, es ist allerdings Voraussetzung und wird nicht wiederholt. Im Rahmen von *Beispielen* wird das Rechnen mit Zahlen dazu herangezogen, die abgeleiteten Theoreme zu illustrieren.
ACHTUNG: Das bedeutet nicht, dass richtiges Rechnen im Mathematikstudium zweitrangig ist! Es ist unverzichtbare Grundlage.

Ein großer Teil der mathematischen Theorie wird durch abstrakteres Ableiten gewonnen. Dabei spielen mitunter auch Rechenvorgänge eine wichtige Rolle, diese Ableitungen zielen jedoch meist darauf ab, möglichste Allgemeinheit in den Aussagen zu erzielen. Das „Buchstabenrechnen“ steht also im Mathematikstudium im Vordergrund.

1.1.2. „Ich habe genau einen Bruder“ — Sprache. Die Sprache dient in der Mathematik, wie auch im täglichen Leben, der Informationsübermittlung. Die Aufgabe des Sprechenden ist es dabei, durch geeignete Sprachwahl dem Hörenden möglichst wenig Mühe beim Verstehen zu verursachen. Der Beruf des/der MathematikerIn prägt die verwendete Sprache, wie das bei jedem Beruf der Fall ist.

Genauso wie von ÄrztInnen in der Regel anstelle des Wortes „Ellenbogenbruch“ meist „Olekranonfraktur“ verwendet wird, kann man von MathematikerInnen mitunter „ich habe genau einen Bruder“ hören. Während jedoch angehende MedizinerInnen einige Monate Zeit haben, ihre Sprache an das Berufsbild anzupassen, ist es für Mathematikstudierende notwendig, die grundlegenden Sprechweisen äußerst rasch zu erlernen. Ohne diese Fähigkeit gehen viel wesentliche Information und das Grundverständnis der mathematischen Aussagen verloren.

Nachdem die Mathematik ein Gebiet ist, in dem es auf Exaktheit ankommt, ist die mathematische Sprache Regeln unterworfen, die über jene hinausgehen, die für Umgangssprache (Hochsprache) und Literatur gelten.

In diesem Buch werden sprachliche Regeln durch grau hinterlegte Schrift hervorgehoben. Viele der hier zitierten Regeln sind ebenso wie viele dazu gehörende Beispiele dem Buch [**Beutelspacher 1999**] entnommen.

Beachten Sie, dass mathematische Sprache als Grundlage die Hochsprache bzw. die Literatur hat. Grundsätzlich kann man daher davon ausgehen, dass mathematische Texte zwar Gebrauchsliteratur aber immerhin Literatur sind. Wenn Sie also die Lösungen von Übungsaufgaben, oder Seminararbeitenverfassen, so halten Sie die folgenden *literarischen Grundregeln* zusätzlich zu den in dieser Vorlesung behandelten mathematischen Konventionen ein.

Schreiben Sie in vollständigen Sätzen und formulieren Sie überschaubar und klar: Bedenken Sie, dass ein Satz zumindest Subjekt und Prädikat enthalten sollte. Lange, verschachtelte Sätze sind schwer verständlich und lassen weder den Verfasser intelligenter wirken noch den Text glaubwürdiger werden.

Jeder Satz, den Sie schreiben, muss (zumindest für Sie) einen Sinn haben: Vermeiden Sie, durch übertriebene Symbolsetzung und logische Formalismen Ihre Aussagen so zu verschlüsseln, dass am Ende nicht einmal Sie selbst auf Anhieb ihren Inhalt verstehen.

Schließlich die wichtigste Regel: Brechen Sie ruhig alle in diesem Skriptum vorgestellten Regeln, wenn Sie sich durch sie eingeengt fühlen, und wenn Sie wissen, was Sie tun.

1.1.3. „Q.E.D.“ — Beweise. Seit Euklid im dritten Jahrhundert vor Christus seine *Elemente* geschaffen hat, in der er die gesamte damals bekannte Mathematik zusammengefasst hat, ist die logische Struktur, das Fundament der Mathematik, auf Beweisen errichtet.

Auf diese Weise wird sichergestellt, dass in der mathematischen Welt die gemachten Aussagen rein logisch nachgewiesen oder widerlegt werden können. Sie müssen nicht durch „Experimente“ oder „Expertengutachten“ gestützt werden. Auch der in vielen Wissenschaften wohlbekannte philosophische Kampf zwischen verschiedenen Schulen und Lehrmeinungen findet in der Mathematik nicht statt, oder beschränkt sich zumindest darauf, ob ein bestimmtes Gebiet interessant bzw. modern ist oder eben nicht.

Das Beweisen ist für StudienanfängerInnen ungewohnt, die aus der Schule gewöhnt sind, die Aussagen der LehrerInnen aufzunehmen und die vorgestellten Methoden nachzuvollziehen. Es ist in der Schule unökonomisch, alle Aussagen der Lehrenden zu hinterfragen. Auf der Universität wird dies anders. Grundsätzlich sollte man scheinbar sein gesamtes Vorwissen hinter sich lassen und sich von neuem von den bisher geglaubten Tatsachen überzeugen (lassen).

Ein großer Fehler von StudienanfängerInnen besteht darin, bei Übungsaufgaben von bis dahin unbewiesenen Tatsachen auszugehen und Beispiele oder Beweise dadurch fälschlicherweise abzukürzen oder gar zu verderben. Darum

> **Unterscheiden Sie im Rahmen eines Beweises oder einer Übungsaufgabe immer genau zwischen den Resultaten, die sie verwenden dürfen und denen die Sie kennen, oder zu kennen glauben.**

Das scheint nur auf den ersten Blick sinnlos. In Wahrheit wird damit ein zweifacher Zweck verfolgt. Zum einen wird der Blick dafür geschult, keine „Lücken im mathematischen Gebäude“ zu hinterlassen; oft ist das der Sinn hinter einer scheinbar einfachen Übungsaufgabe. Zum anderen wird darauf vorbereitet, auch Beweise in mathematischen Strukturen zu finden, die ärmer an Eigenschaften sind und für die manche Resultate nicht gelten.

Zuletzt noch einige sprachliche Hinweise:

Stellen Sie ihre Beweise sorgfältig dar. Dadurch vermeiden Sie es, Lücken in der Kette logischer Schlüsse zu übersehen. Wesentlich bei der Erstellung von Beweisen ist eine sinnvolle Gliederung und sinnvolle Untergliederungen.

Beachten Sie beim Beweisen zu Beginn die folgenden Prinzipien:

Sagen Sie, was Sie beweisen. Stellen Sie an jeder Stelle im Beweis sicher, dass die Hörerin oder der Leser genau weiß, welche Teilbehauptung Sie gerade untersuchen. Folgen Sie dem folgenden Grundprinzip:

> **Sagen Sie immer, was Sie als nächstes vorhaben, führen Sie es durch, und sagen Sie danach, dass Sie es getan haben.**

Es empfiehlt sich auch, zu Beginn die zu beweisende Aussage in mathematische Form zu übersetzen.

Gliedern Sie ihren Beweis: Alle Beweise, die länger als etwa eine halbe Seite sind, sollten in Teilabschnitte unterteilt werden. Zerlegen Sie den Beweis in eine Reihe von Teilbehauptungen oder Fällen. Kennzeichnen Sie diese mit Einschüben wie *Schritt 1:*, *Schritt 2:*, bzw. *Fall 1:*, *Fall 2:*, etc. Achten Sie besonders bei der Unterteilung in Fälle, dass Sie keinen Fall vergessen und führen Sie keine Fälle ein, die nicht gesondert behandelt werden müssen.

Kennzeichnen Sie den Schluss eines Beweises: Es ist äußerst ermüdend für einen Leser, wenn er sich nie sicher sein kann, wo ein Beweis beginnt und wo er genau endet. Als Kennzeichen für das Ende eines Beweises dienen manchmal Phrasen wie

- *Damit ist alles gezeigt.* oder
- *... was wir behauptet hatten.*

und ähnliche Sätze. Das zwingt den Leser dazu, den Beweis bis zum Ende zu lesen und erschwert es, sich einen schnellen Überblick zu verschaffen, speziell wenn mehrere Resultate und Zwischentexte aufeinander folgen. Übersichtlicher sind die Standardabkürzungen

- *w.z.z.w* — was zu zeigen war — oder die lateinische Variante
- *Q.E.D.* (auch *q.e.d.* oder *qed.*) — quod erat demonstrandum.

In modernen Büchern hat sich das ökonomische Beweisabschlusszeichen, das meist am Ende der letzten Beweiszeile steht,

$$\ldots \qquad\qquad \square$$

durchgesetzt.

Achten Sie im Verlauf der Vorlesung auf die Struktur der vorgetragenen Beweise, nehmen Sie diese als Beispiele und achten Sie auf die grau hinterlegten Textstellen, mit denen typische Redewendungen und die Struktur hervorgehoben werden.

1.2. Schulstoff

Parallel zu dieser Vorlesung werden wichtige Aspekte des AHS–Schulstoffs im Rahmen von Workshops wiederholt. Ein Großteil dieses Stoffes wird in nicht exakter Form vorgetragen. Die Darstellung orientiert sich am Lehrstoff, der für Realgymnasien vorgesehen ist.

Die Wiederholung des Schulstoffs soll hauptsächlich dazu dienen, die Studierenden auf vorhandene Wissenslücken hinzuweisen und die grundlegenden *algorithmischen Fertigkeiten* zu Beginn des Studiums nochmals darzustellen.

> **Es seien alle Studierenden dazu angehalten, den Schulstoff erneut zu lernen, denn die vollständige Beherrschung der dort vermittelten Fakten und Fertigkeiten wird im gesamten folgenden Studium kommentarlos vorausgesetzt werden.**
> **Fehler, auch Rechenfehler, deren Grundlage der Schulstoff ist, sind keine Kavaliersdelikte. Sie zählen bei Übungen und Prüfungen grundsätzlich als *schwere Fehler* und entwerten ein Beispiel vollständig.**

Arbeiten Sie also bei Prüfungen und Übungen sorgfältig und üben Sie den Schulstoff gut ein.

Einige abschreckende Beispiele aus Prüfungen der jüngeren Vergangenheit, die im Mathematikstudium nicht toleriert werden.

- $\frac{a}{b} + \frac{c}{d} = \frac{a+b}{c+d}$.
- $\frac{3x+1}{3y+1} = \frac{x+1}{y+1}$.
- $(e^x)' = x\,e^{x-1}$ bei Ableitung nach x.
- $\int_0^1 e^x\,dx = e$.

- Werden zwei Würfeln geworfen, dann errechnet sich die Wahrscheinlichkeit, dass dabei eine 6 geworfen wird zu $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.
- $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.
- $\log ab = \log a \log b$, $\log 0 = 0$.

1.3. Aufbaustoff

Einige Teile des Schulstoffs und die darüber hinaus gehenden Fakten werden in der Vorlesung selbst mit „voller mathematischer Exaktheit" vorgetragen. Sie bilden gemeinsame Grundlage der nachfolgenden Vorlesungen *Analysis 1* und *Lineare Algebra und Geometrie 1*. Im Rahmen dieses Erweiterungsstoffs werden außerdem weitere Sprachregeln und Sprechweisen erklärt, sowie das Beweisprinzip illustriert.

Einige Teile des Erweiterungsstoffs sind nicht Gegenstand der Prüfung und nur gedacht als Hinweise und Informationen für die besonders Interessierten. Die Teile des Skriptums, die gekennzeichnet sind wie dieser Absatz, bilden diesen Zusatzstoff.

Abschließend noch ein Wort zur Exaktheit: Im Prinzip wäre es möglich, den Stoff der Vorlesung zu entwickeln, ohne irgendwelche mathematische Vorkenntnisse vorauszusetzen. Dann würde allerdings Ihr Vorwissen über weite Strecken brach liegen und viele der vorgestellten Konstruktionen würden blutleer und gekünstelt wirken. Um Sie nicht so in eine Motivationskrise zu treiben, werden wir während wir die Inhalte der Vorlesung exakt aufbauen auf den reichen Fundus der Schulmathematik zurückgreifen, um so den Stoff anhand von Beispielen zu motivieren und zu untermalen. Das macht es allerdings nötig, einige mathematische Begriffe und Objekte zu verwenden, bevor sie im Rahmen der Vorlesung exakt definiert wurden, d.h. sie *naiv* zu verwenden und an Ihre bisherige mathematische Erfahrung zu appellieren.

Denken Sie z.B. an Ihre Kenntnisse über Zahlen. Schon seit der Volksschule „wissen" Sie, dass $1 + 1 = 2$ ist. Allerdings, ohne genau sagen zu können, was die einzelnen Symbole 1, $+$ und $=$ eigentlich „sind". Die hier vorkommenden Zahlen $1, 2, \dots$, also jene Zahlen die man zum Zählen verwendet, heißen die *natürlichen Zahlen*. Per Konvention wollen wir auch die Null zu den natürlichen Zahlen zählen und diese mit $\mathbb{N}$ bezeichnen. Nehmen wir zu den natürlichen Zahlen auch die negative Zahlen hinzu, so gelangen wir zu den *ganzen* Zahlen, die mit $\mathbb{Z}$ bezeichnet werden. Die Quotienten (Brüche) ganzer Zahlen mit von Null verschiedenem Nenner heißen die *rationalen Zahlen*; diese sind bereits ausreichend, um die meisten praktisch auftretenden Rechnungen zu bewältigen und werden mit $\mathbb{Q}$ bezeichnet. Allerdings wurden schon früh in der Geschichte der Mathematik Zahlen „entdeckt", die nicht rational sind, z.B. die Länge der Diagonale im Einheitsquadrat. Nimmt man diese Zahlen hinzu, so landet man schließlich bei den *reellen Zahlen*, die durch die Zahlengerade veranschaulicht werden können. In der Schulmathematik wird oft die folgende Vorstellung bemüht: Man fügt zu den rationalen Zahlen, die man sich als abbrechende oder periodische Dezimalzahlen vorstellen kann, alle — also auch die nicht abbrechenden — Dezimalzahlen hinzu.

Haben Sie in der Schule sogar die *komplexen Zahlen* kennengelernt, so werden Sie sich sicher freuen, diese am Ende dieses Skriptums wiederzufinden. Bis Sie dorthin gelangt sind werden Sie allerdings auch schon genau wissen, wie die natürlichen, ganzen, rationalen und sogar die reellen Zahlen *mathematisch exakt definiert* sind. Nun aber nocheinmal als Überblick und zum „naiven" Einstieg:

natürliche Zahlen	$\mathbb{N}$:	$0, 1, 2, 3, \ldots$
ganze Zahlen	$\mathbb{Z}$:	$0, \pm 1, \pm 2, \pm 3, \ldots$
rationale Zahlen	$\mathbb{Q}$:	Brüche ganzer Zahlen mit nicht verschwindendem Nenner.
reelle Zahlen	$\mathbb{R}$:	Zahlen auf der Zahlengeraden, *alle* Dezimalzahlen.

KAPITEL 2

Grundlagen

Bevor wir uns auf den Ozean der Mathematik hinauswagen, müssen wir als ersten Schritt einiges an Grundlagenwissen ansammeln, einfache Schreibweisen und Ideen, ohne die wir unser Ziel, das Wesen der „höheren“ Mathematik zu erforschen, nicht erreichen können.

2.1. Beweise

Wie wir schon in der Einleitung (Abschnitt 1.1.3) erwähnt haben, bilden *Beweise* die Grundlage des mathematischen Gebäudes. Während wir in den weiteren Abschnitten tiefer auf die Art und Weise eingehen werden, wie Beweise aufgebaut und geführt werden, wollen wir zunächst mit ein paar einfach verständlichen Beispielen beginnen.

Zu Beginn erinnern wir uns, dass eine gerade Zahl eine ganze Zahl ist, die durch 2 teilbar ist.

Proposition 2.1.1 (Quadrate gerader Zahlen)**.** *Das Quadrat einer geraden Zahl ist gerade.*

Man kann sich die gesamte Mathematik denken als eine Ansammlung von Aussagen, die aus gewissen Grundaussagen (den **Axiomen**) durch logische Schlussfolgerungen abgeleitet werden. Dieser Vorgang heißt **beweisen**. Gilt eine Aussage A als bewiesen, und kann man eine weitere Aussage B logisch aus A ableiten, so gilt auch B als bewiesen.

Die solcherart bewiesenen Aussagen nennt man **Sätze** oder auch **Theoreme**. Üblich in der Literatur ist, zuerst die Aussage des Satzes aufzuschreiben und danach den Beweis anzuschließen, in dem die Aussage des Satzes aus bekannten Resultaten hergeleitet wird. Mit diesem Prinzip steht und fällt die Mathematik, daran lässt sich nicht deuteln.

Anstelle von **Satz** bzw. **Theorem** werden auch zuweilen andere Ausdrücke verwendet, die den Stellenwert der Aussagen untereinander im Rahmen einer Theorie andeuten. Ob und wie man diese Begriffe verwendet ist teilweise auch Geschmackssache.

Satz, Theorem: Dies ist das typische Resultat einer Theorie.

Hauptsatz: So wird ein besonders wichtiger Satz in einem Teilgebiet der Mathematik genannt. Ein Beispiel ist etwa der Hauptsatz der Differential- und Integralrechnung, den Sie im Rahmen der Analysis Vorlesungen kennen lernen werden.

Lemma: Dieses Wort stammt aus dem Griechischen (die Mehrzahl ist daher **Lemmata**) und bedeutet „Stichwort“ oder „Hauptgedanke“. Es wird in zwei verschiedenen Zusammenhängen verwendet. Zum einen bezeichnet es ein kleines, meist technisches Resultat, einen **Hilfssatz**, der im Rahmen des Beweises eines wichtigen Satzes verwendet wird aber selbst meist uninteressant ist. Zum anderen handelt es sich dabei um besonders wichtige **Schlüsselgedanken**, die in vielen Situationen nützlich sind. Solche genialen Erkenntnisse tragen meist den Namen des Erfinders (Lemma von Zorn, Lemma von Urysohn,. . .).

Proposition: Dies ist die lateinische Bezeichnung für Satz und wird manchmal an dessen Stelle verwendet, meist aber um ein Resultat zu bezeichnen, dessen Wichtigkeit zwischen der eines Hilfssatzes und der eines Theorems liegt.

Korollar, Folgerung: Dies ist ein Satz, der aus einem anderen Satz durch triviale oder sehr einfache Schlussweise folgt. Manchmal ist es ein Spezialfall einer bereits bewiesenen allgemeineren Aussage. Das Wort Korollar stammt übrigens vom lateinischen Wort *corollarium* ab, welches ein Kränzchen bezeichnet, das der Gastgeber dem Gast „einfach so“ schenkt.

Beweis. Sei n eine beliebige gerade Zahl. Nachdem n durch 2 teilbar ist, existiert eine ganze Zahl m mit $n = 2m$.

Wir können also nun das Quadrat von n durch m ausdrücken und erhalten $n^2 = (2m)^2 = 4m^2$. Natürlich ist $4m^2$ durch 2 teilbar, und daher ist n^2 gerade. □

Falls Ihnen der Beweis (zu) einfach erscheint, dann beherzigen Sie bitte nochmals Punkt (1) auf Seite 1 und lesen aufmerksam weiter.

Im obigen Beweis haben wir mit der Voraussetzung (die ursprüngliche Zahl ist gerade) begonnen, sie ein wenig umgeformt und daraus die Behauptung (ihr Quadrat ist gerade) hergeleitet. Beweise, die auf diese Art vorgehen, nennen wir **direkte Beweise**.

Definition 2.1.2 (Primzahl). *Eine* Primzahl *ist eine natürliche Zahl* $p > 1$*, die nur die trivialen Teiler besitzt, d.h. deren einzige Teiler* 1 *und sie selbst sind.*

Definitionen dienen zur Vergabe von *Namen*. Sie sind weder richtig noch falsch (außer bei der Reproduktion schon vorhandener Definitionen im Rahmen einer Prüfung); sie können allerdings sinnvoll oder unsinnig sein.

Eine Definition verändert nicht das mathematische Gebäude, bloß die Sprache darüber wird um ein weiteres Vokabel ergänzt.

Theorem 2.1.3 (Satz von Euklid). *Es gibt unendlich viele Primzahlen.*

Beweis. Nehmen wir einmal an, es gäbe nur endlich viele Primzahlen. Wenn dem so ist, können wir sie mit $p_1, \dots, p_n$ bezeichnen (wobei n die endliche Anzahl von Primzahlen ist).

Nun bilden wir die Zahl

$$m = p_1 p_2 \cdots p_n + 1.$$

Weiters verwenden wir, dass jede natürliche Zahl größer 1 in Primfaktoren zerlegt, d.h. als Produkt von Primzahlen geschrieben werden kann (siehe unten). Insbesondere ist jede solche Zahl durch mindestens eine Primzahl teilbar.

Daher muß es eine Primzahl geben, die m teilt. Dies ist jedoch nicht möglich, da m durch keine der Primzahlen p_i teilbar ist. (Offensichtlich bleibt immer ein Rest von 1.) So endet unsere logische Schlusskette in einem Widerspruch.

Wir müssen also unsere zu Beginn des Beweises getroffene Annahme verwerfen, und daher existieren tatsächlich unendlich viele Primzahlen. □

In diesem Beweis sind wir anders herum vorgegangen. Wir haben mit einer Annahme begonnen, deren Aussage gerade das Gegenteil unserer Behauptung war. Danach haben wir eine logische Schlusskette bis zu einem Widerspruch verfolgt. Die Annahme konnte also nicht richtig gewesen sein, und daher musste zwangsläufig ihr Gegenteil stimmen („tertium non datur“), also unsere Behauptung wahr sein.

Beweise dieser Struktur nennen wir **indirekte Beweise**. Wir werden ihre logische Struktur im Abschnitt 3.2.2.1 noch genauer kennenlernen.

Obiger Beweis des Satzes von Euklid (ca. 325–265 v.Chr.) wird vielfach als *das* Musterbeispiel eines einfachen, klaren und *eleganten* Beweises angesehen. Er findet sich erstmals in

Euklids Werk „Die Elemente“ in dem er den Großteils des damals bekannten mathematischen Wissens zusammentrug.

Schließlich beweisen wir noch die oben oben verwendetet Tatsache, dass jede natürliche Zahl größer 1 in Primfaktoren zerlegt werden kann. Wieder werden wir einen indirekten Beweis führen.

Lemma 2.1.4. *(Existenz der Primfaktorzerlegung) Sei $a > 1$ eine natürliche Zahl. Dann gibt es Primzahlen $p_1, \ldots, p_k$, sodass*

$$a = p_1 p_2 \cdots p_k.$$

Beweis. Wir nehmen indirekt an, dass die Aussage des Lemmas falsch ist, also nicht alle Zahlen größer 1 in Primfaktoren zerlegt werden können. Sei a die kleinste solche Zahl.

Zunächst kann a keine Primzahl sein, denn sonst könnten wir $a = p$ schreiben und a wäre ein (triviales) Produkt von Primzahlen. Also muß es eine Zahl b geben, die a teilt. D.h. wir können schreiben $a = b\,c$ für ein geeignetes c und $b, c < a$.

Nun können wir b und c in Primfaktoren zerlegen (sie sind ja kleiner als a); es gilt also $b = p_1 \cdots p_s$ und $c = p_{s+1} \cdots p_k$ für gewisse Primzahlen $p_1, \ldots, p_k$.

Zusammengefasst gilt also

$$a = b\,c = p_1 \cdots p_s \cdot p_{s+1} \cdots p_k,$$

was einen Widerspruch zu unserer Annahme darstellt. □

2.2. Indizes

Im Beweis von Theorem 2.1.3 sind Ausdrücke der Form $p_1, \ldots, p_n$ und p_i vorgekommen. Die unter das p tiefer gestellten Zahlen und Buchstaben nennt man **Indizes**.

Indizes dienen dazu, miteinander verwandte Objekte weitgehend einheitlich zu bezeichnen. Darum keine Angst vor Indizes. In vielen Fällen sind sie einfacher und klarer als alle anderen Darstellungsmöglichkeiten. Besonders im Zusammenhang mit Summen und Produkten (siehe Abschnitt 2.3) treten sie häufig auf.

Eine wichtige Eigenschaft eines Index ist, dass er verschiedene Werte annehmen kann, ganz wie eine Variable. So kann der Index i im Ausdruck p_i im Beweis zu Theorem 2.1.3 als Wert alle natürlichen Zahlen von 1 bis n annehmen.

Die Einzahl von Indizes ist übrigens *Index* und nicht Indiz, deren Mehrzahl lautet Indizien, und diese haben in Gerichtssälen nicht aber in Mathematiktexten ihren Platz.

Es ist z.B. offensichtlich, dass die Argumente der Funktion h im folgenden Beispiel allesamt Variable sein sollen, und dass h genau n Argumente benötigt.

$$h(x_1, \ldots, x_n)$$

Vergleichen Sie das mit der viel unklareren Schreibweise

$$h(x, y, \ldots, z)$$

Besonders in der linearen Algebra werden Indizes von Anfang an auftreten. Auch Doppel- (A_{12}, a_{kl}, $b_{i,j+1}$) und sogar Mehrfachindizes (r_{12345}, p_{ijkm}, $Y_{i,i+1,\ldots,i+n}$), ja selbst indizierte Indizes ($Y_{i_1,\ldots,i_n}$) sind möglich und sinnvoll. Folgender Rat:

> Machen Sie sich immer klar, was welcher Index bedeutet. Falls Buchstaben als Index auftreten, behalten sie immer im Auge, welche Werte der Index annehmen kann.

Beispiel 2.2.1 (Matrizen)**.** *Wir ordnen die Zahlen* $1, 2, \ldots, 20$ *in einer* Matrix*, also einem rechteckigen Schema von Zahlen, wie folgt an. Dabei bezeichnen wir die Matrix mit* A*.*

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \end{pmatrix}$$

Mit Hilfe eines Doppelindex können wir die einzelnen Einträge der Matrix bezeichnen, wobei allgemein die Konvention gilt, dass der erste Index die Zeile bezeichnet, der zweite die Spalte (Merkregel: Zeile zuerst; Spalte später). So haben wir z.B. für den Eintrag in der 2. Zeile in der 3. Spalte $A_{23} = 8$ *und für den 1. Eintrag in der 3. Zeile* $A_{31} = 11$*.*

Wir können sogar die gesamte Matrix A *über ihre Elemente mit Hilfe der Indizes definieren, indem wir schreiben*

$$A_{ij} = 5i + j - 5, \quad i = 1, \ldots, 4,\ j = 1, \ldots, 5.$$

Oft werden die Einträge von Matrizen auch mit kleinen Buchstaben bezeichnet, also die Einträge der Matrix A *mit* a_{ij}*.*

Bei Umformungen von Ausdrücken sind Indizes „in Schachteln verpackt“. Das bedeutet, dass man sie nicht „wegkürzen“ oder ähnliches kann. Zusätzlich werden Doppelindizes durch Beistriche getrennt, wenn immer das zur leichteren Lesbarkeit beiträgt. Zur Illustration seien einige richtige und einige falsche Beispiele angegeben.

$$\begin{aligned} A_{i+1+3\cdot5,j} &= A_{i+16,j} \\ f_i - 1 &\neq f_{i-1} \\ B_s B_s &= B_s^2 \neq B_{s^2} \\ \frac{B_s}{s} &\neq B \end{aligned}$$

2.3. Summen, Produkte — Zeichen

In der Mathematik untersucht man häufig Summen, in denen die Anzahl der Terme nicht a priori fest steht. So hat etwa ein allgemeines **Polynom** n–ten Grades (für ein beliebiges n in $\mathbb{N}$) die Form

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

mit $n+1$ Termen, die aufsummiert werden. Um die Schreibweise von den Punkten $(+\ldots+)$ zu befreien, verwendet man das Summenzeichen. Dieses erlaubt es eine mehrfache Addition ähnlicher Ausdrücke vereinfacht darzustellen. So kann man mit Hilfe des Summenzeichens Σ das Polynom im oberen Beispiel schreiben als

$$p(x) = \sum_{i=0}^{n} a_i x^i. \tag{2.1}$$

Genauer betrachtet besteht der allgemeine Summenausdruck mit dem Summenzeichen aus vier verschiedenen Termen.

- Es gibt es eine **Laufvariable**, den **Summationsindex**, in unserem Beispiel i.
- Diese Variable nimmt *alle ganzen* Zahlen beginnend mit der **unteren Grenze**, im Beispiel 0,
- bis zur **oberen Grenze**, in Gleichung (2.1) ist sie n, in Einserschritten an.

- Der Gesamtausdruck entspricht dann einer Summe von Termen, die aussehen wie der **allgemeine Summand**, hier $a_i x^i$, in dem der Summationsindex jeweils durch alle Werte ersetzt wird. **In der dadurch gebildeten Summe kommt der Summationsindex also nicht mehr vor!**

Betrachtet man eine Summe, so kann man sofort erkennen, aus wievielen Termen die Summe besteht

$$\text{Anzahl der Summanden} = \text{obere Grenze} - \text{untere Grenze} + 1.$$

Dies ist auch der erste Schritt in der Analyse eines allgemeinen Summenausdrucks.

Man kann das Summenzeichen dazu verwenden, die Verknüpfung einer bestimmten Anzahl von Ausdrücken darzustellen. Ein einfaches Beispiel dazu ist

$$\sum_{i=1}^{4} \frac{1}{i+1} = \frac{1}{1+1} + \frac{1}{2+1} + \frac{1}{3+1} + \frac{1}{4+1} = \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}$$

Die wahre Stärke besteht allerdings, wie erwähnt, darin, dass man eine unbestimmte Anzahl von Termen summieren kann:

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n$$

In der Analysis wird gezeigt werden, dass selbst die Unendlichkeit hier **keine** Grenze bildet! Man kann zum Beispiel eine **unendliche Reihe** (hier an einem Beispiel) bilden, und schreiben:

$$\sum_{i=1}^{\infty} \frac{1}{i} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots$$

Den tieferen mathematischen Sinn dieses Ausdrucks wollen wir an dieser Stelle allerdings nicht untersuchen.

Die Laufvariable kann an die jeweiligen Bedürfnissen des Problems anpasst werden. Sie kann beliebig umbenannt und sogar weiteren Transformationen unterworfen werden (ähnlich der Substitutionsregel für Integrale), wenn dabei beachtet wird, dass sich das Ergebnis nicht ändert. So kann etwa eine **Indexverschiebung** durchgeführt werden: Setze zum Beispiel $i = j + 2$ so gilt:

$$\sum_{i=3}^{9} a_i = \sum_{j=1}^{7} a_{j+2}$$

Wir haben dabei die neuen Grenzen für j durch Einsetzen berechnet

$$\begin{aligned} \text{untere Grenze:} \quad & 3 = i = j + 2 \Rightarrow j = 1 \\ \text{obere Grenze:} \quad & 9 = i = j + 2 \Rightarrow j = 7 \end{aligned}$$

und im allgemeinen Summanden i durch $j + 2$ ersetzt.

Nach Definition ist übrigens das Ergebnis einer allgemeinen Summe gleich 0, falls die untere Grenze größer als die obere Grenze ist.

Es treten in der Mathematik natürlich nicht nur Summen variierender Länge auf, auch für andere Operationen, etwa Produkte, benötigt man ein ähnliches Prinzip, und daher hat man viele dem Summenzeichen entsprechende Zeichen eingeführt. So gibt es etwa das bereits in der Analysis wichtige Produktzeichen ($\prod$) und noch weitere, etwa $\bigcup$, $\bigcap$, $\bigodot$, $\bigoplus$, usw., die in anderen Bereichen der Mathematik ihre Rolle spielen.

Die Anwendung dieser Zeichen folgt demselben Schema wie die des Summenzeichens. So ist etwa

$$\prod_{i=1}^{5} b_i = b_1 b_2 b_3 b_4 b_5,$$

$$\prod_{i=1}^{0} x_i = 1,$$

Das „leere Produkt“ (obere Grenze ist kleiner als untere Grenze) wird also als 1 festgelegt.

Oft lassen sich Teile der verknüpften Ausdrücke vor das Verknüpfungszeichen ziehen, wobei man stets darauf achten muss, dass dies nach den Rechenregeln für die jeweilige Operation geschieht—beim Summenzeichen also das Herausheben, d.h.:

$$\sum_{i=1}^{n} 7x_i = 7 \sum_{i=1}^{n} x_i.$$

ACHTUNG: Nur Konstante können herausgehoben werden! Also nicht:

$$\sum_{i=1}^{n} i x_i \neq i \sum_{i=1}^{n} x_i.$$

Beim Produktzeichen ist zu beachten, dass solche Konstanten ja multipliziert werden! Daher:

$$\prod_{i=1}^{n} 7x_i = 7^n \prod_{i=1}^{n} x_i.$$

Man kann das Produktzeichen auch verwenden um **Fakultäten** anzuschreiben. Zunächst definieren wir.

Definition 2.3.1 (Fakultät)**.** *Die* Fakultät $n!$ *(sprich: n Fakultät, oder: n faktorielle) einer natürlichen Zahl n ist rekursiv definiert durch:*

$$0! := 1$$

$$(n+1)! := (n+1)n!$$

Das Zeichen := (definitorisches Gleichheitszeichen) bedeutet, dass die linke Seite (hier $0!$ resp. $(n+1)!$) durch die rechte Seite **definiert** wird.

Unter einer **rekursiven Darstellung oder Definition** eines von n in $\mathbb{N}$ abhängigen Ausdrucks $A(n)$ verstehen wir allgemein die Angabe einer Vorschrift zur Berechnung von $A(n)$ mit Hilfe einiger oder aller seiner bereits erklärten Vorgänger $A(n-1)$, $A(n-2), \ldots$. Genau das haben wir oben für $A(n) = n!$ getan, indem wir zunächst $A(0) = 0! := 1$ definiert und dann angegeben haben, wie $A(n+1) = (n+1)!$ aus $A(n) = n!$ zu berechnen ist (eben gemäß $A(n+1) := (n+1)\,A(n)$).

Im Unterschied dazu sprechen wir von einer **geschlossenen Darstellung**, falls $A(n)$ ohne Bezugnahme auf seine Vorgänger angegeben wird.

Nun geben wir wie versprochen eine Darstellung der Fakultäten mit Hilfe des Produktzeichens an. Dabei handelt es sich um die unmittelbar einsichtige, geschlossene Darstellung

$$n! = \prod_{i=1}^{n} i$$

Dieser Ausdruck wird besonders für kombinatorische Probleme benötigt. So gibt $n!$ die Anzahl der Möglichkeiten an, n verschiedene Dinge hintereinander aufzureihen.

Eine wesentliche Vereinfachung ist bei Summanden spezieller Gestalt möglich, nämlich für sogenannte **Teleskopsummen**:

$$\sum_{i=1}^{n}(a_i - a_{i-1}) = \cancel{a_1} - a_0 + \cancel{a_2} - \cancel{a_1} + \cancel{a_3} - \cancel{a_2} + \cdots + \cancel{a_{n-1}} - \cancel{a_{n-2}} + a_n - \cancel{a_{n-1}} = a_n - a_0$$

Analog ergeben sich **Teleskopprodukte**:

$$\prod_{i=1}^{n} \frac{a_i}{a_{i-1}} = \frac{a_n}{a_0}$$

Zum Abschluss stellen wir noch eine weitere Verwendung des Summenzeichens vor (Analoges gilt natürlich auch für die verwandten Zeichen). Der Ausdruck

$$\sum_{i \in I} a_i$$

definiert eine Summe, die für jedes Element der Menge I einen Term enthält. (Wir verwenden hier den Mengenbegriff naiv, er wird im Kapitel 4 präzisiert.) Ähnlich wie zuvor wird im allgemeinen Summanden die Laufvariable i jeweils durch das ausgewählte Element ersetzt. Diese Notation hat vor allem zwei Vorteile. Zum einen können auch „unregelmäßige" Indexmengen verwendet werden, und zum anderen bleibt die Anzahl der Indizes nicht auf endlich (oder abzählbar; vgl. 4.4 unten) viele beschränkt.

Beispiel 2.3.2. *Es gilt*

$$\sum_{i \in \{1,4,7,21\}} a_i^2 = a_1^2 + a_4^2 + a_7^2 + a_{21}^2.$$

2.4. Gleichungsumformungen in Beweisen — Stil und Fallen

2.4.1. Elementare Umformungen. Zunächst zur Schreib- und Sprechweise:

Werden Ketten von Gleichungen untereinander geschrieben, so bedeutet das, dass die *untere* Gleichung *aus der oberen* folgt, d.h.: Wenn die obere Gleichung gilt, dann gilt auch die untere.

Beispiel 2.4.1. *Betrachten wir die Ableitung*

$$\begin{array}{rcll}
3r^2 + 4r + 5 & = & -r^3 + r + 4 & \mid + r^3 - r - 4 \\
r^3 + 3r^2 + 3r + 1 & = & 0 & \\
(r+1)^3 & = & 0 & \mid \sqrt[3]{\ } \\
r + 1 & = & 0 & \mid -1 \\
r & = & -1 &
\end{array}$$

Sie ist, wie in der Mathematik üblich, von oben nach unten gültig. Das bedeutet, wenn wir Folgerungspfeile einführen, können wir die Implikationen hervorheben

$$\begin{array}{rclll}
3r^2 + 4r + 5 & = & -r^3 + r + 4 & \mid + r^3 - r - 4 & \Rightarrow \\
r^3 + 3r^2 + 3r + 1 & = & 0 & & \Rightarrow \\
(r+1)^3 & = & 0 & \mid \sqrt[3]{\ } & \Rightarrow \\
r + 1 & = & 0 & \mid -1 & \Rightarrow \\
r & = & -1 & &
\end{array}$$

und wenn wir alle Zwischenschritte weglassen, ergibt sich der logische Schluss

$$3r^2 + 4r + 5 = -r^3 + r + 4 \Rightarrow r = -1.$$

Werden Umformungen durchführt und soll ausdrückt werden, dass sie in beide Richtungen stimmen, so **muss** dies durch explizites Setzen von Äquivalenzpfeilen ($\Leftrightarrow$) angezeigt werden.

Beispiel 2.4.2. *In Beispiel 2.4.1 folgen in Wahrheit die oberen Gleichungen auch aus den unteren, d.h. sie sind wirklich alle äquivalent. Um das zu unterstreichen, wollen wir daher*

$$\begin{array}{rcll} 3r^2+4r+5 &=& -r^3+r+4 & \mid +r^3-r-4 \quad \Leftrightarrow \\ r^3+3r^2+3r+1 &=& 0 & \qquad\qquad\qquad\quad \Leftrightarrow \\ (r+1)^3 &=& 0 & \mid \sqrt[3]{} \qquad\qquad\quad \Leftrightarrow \\ r+1 &=& 0 & \mid -1 \qquad\qquad\quad \Leftrightarrow \\ r &=& -1 & \end{array}$$

schreiben.

Auch bei Schlüssen von unten nach oben in einer Umformung müsste die Implikationsrichtung durch Setzen des entsprechenden Pfeils ($\Leftarrow$) angegeben werden. **Schlüsse von unten nach oben gelten nicht als guter mathematischer Stil und sollten daher unbedingt vermieden werden.** Machen Sie sich daher immer klar, womit eine Umformung beginnt und was Sie abzuleiten gedenken. Wenn Sie die Rechnung vom Ergebnis zum Ausgangspunkt hin durchführen, so kehren sie die Schlussweise in der Reinschrift um!

Welche Umformungen sind eigentlich erlaubt? Auf beiden Seiten darf dasselbe addiert (subtrahiert) werden. Es dürfen auch beide Seiten mit derselben Zahl multipliziert werden; Wie steht es mit der Division?

Theorem 2.4.3 (Sinnlosigkeit der Zahlen)**.** *Alle Zahlen sind gleich.*

Beweis. O.B.d.A. werden wir den Spezialfall $1 = 2$ beweisen. Wir werden nur elementare Umformungen benutzen. Wir beginnen mit reellen Zahlen a und b mit $a = b$.

Die Abkürzung **O.B.d.A.** steht für *ohne Beschränkung der Allgemeinheit*. Korrekt wird sie zu Beginn eines Beweises oder Beweisteils verwendet. Damit wird der Leser auf zwei Dinge aufmerksam gemacht. Einerseits soll nur ein Teil der Aussage bewiesen werden, und andererseits ist der Autor des Beweises der Meinung, dass die Gesamtaussage einfach aus dem Bewiesenen folgt. Es steckt also hinter o.B.d.A. ein weiterer mathematischer Satz („aus dem tatsächlich Bewiesenen folgt die Aussage des Satzes"), und o.B.d.A. bedeutet dann, dass diese Implikation nach Meinung des Autors *trivial*, also besonders einfach herzuleiten ist.

Zusätzlich zur Beschränkung auf einen Sonderfall, aus dem schon die gesamte Aussage folgt, kann man O.B.d.A. auch noch zur Vereinfachung der Bezeichnung oder zum Ausschließen trivialer Sonderfälle verwenden. Beispiele zu diesen Verwendungen werden Sie in späteren Beweisen finden.

$$\begin{array}{rcll} a &=& b & \\ a^2 &=& ab & \text{nach Multiplikation mit } a \\ a^2+a^2 &=& a^2+ab & \text{nach Addition von } a^2 \\ 2a^2 &=& a^2+ab & \\ 2a^2-2ab &=& a^2+ab-2ab & \text{nach Subtraktion von } 2ab \\ 2a^2-2ab &=& a^2-ab & \\ 2(a^2-ab) &=& 1(a^2-ab) & \\ 2 &=& 1 & \text{nach Division durch } a^2-ab, \end{array}$$

woraus unsere Behauptung folgt. $\square$

Natürlich haben wir in diesem Beweis einen Fehler gemacht. Können Sie ihn entdecken?

An diesem Beispiel sieht man schön die Falle, in die man bei Verwendung der Division als Äquivalenzumformung tappen kann. Es muss immer sichergestellt werden, dass nicht durch 0 dividiert wird wie im obigen Beweis, und 0 kann sich hinter komplizierten Ausdrücken verbergen.

2.4.2. Anwendung von Funktionen. Man kann nicht nur auf beiden Seiten der Gleichung elementare arithmetische Operationen ausführen, sondern man kann auch versuchen, geeignete Funktionen anzuwenden um zu vereinfachen. Besonders beliebt sind Umkehrfunktionen von Funktionen, die auf beiden Seiten der Gleichung auftauchen.

Ein einfaches Beispiel bietet die nächste Umformungskette, in der wir im ersten Schritt die Logarithmusfunktion (log), die Umkehrfunktion der Exponentialfunktion angewendet haben.

$$
\begin{aligned}
e^{3x+4} &= e^{x-2} \quad &|\log\_ \\
3x+4 &= x-2 \\
2x &= -6 \\
x &= -3
\end{aligned}
$$

In der Mathematik wird der natürliche Logarithmus oft mit log und nicht mit ln bezeichnet. Der Logarithmus zur Basis a wird $\log_a$ geschrieben und für den Logarithmus zur Basis 10 müssen wir bei Verwendung obiger Konvention $\log_{10}$ schreiben und nicht bloß log.

Theorem 2.4.4 (Sinnlosigkeit der Zahlen — 2. Versuch)**.** *Alle Zahlen sind gleich.*

Beweis. O.B.d.A werden wir den Spezialfall $4 = 5$ beweisen:

$$
\begin{aligned}
-20 &= -20 \\
16-36 &= 25-45 \\
16-36+\tfrac{81}{4} &= 25-45+\tfrac{81}{4} \\
4^2-2\cdot 4\cdot \tfrac{9}{2}+\left(\tfrac{9}{2}\right)^2 &= 5^2-2\cdot 5\cdot\tfrac{9}{2}+\left(\tfrac{9}{2}\right)^2 \\
\left(4-\tfrac{9}{2}\right)^2 &= \left(5-\tfrac{9}{2}\right)^2 \qquad\qquad \text{weil } (a-b)^2 = a^2-2ab+b^2 \\
4-\tfrac{9}{2} &= 5-\tfrac{9}{2} \\
4 &= 5,
\end{aligned}
$$

womit die Sinnlosigkeit des Zahlenbegriffs erwiesen ist. □

Offensichtlich steckt in diesem Beweis ein Fehler, denn die Ungültigkeit des Satzes steht wohl außer Zweifel. Können Sie den Fehler entdecken?

Die falsche Umformung steht in der vorletzten Zeile: Das Ziehen der Quadratwurzel ist keine Äquivalenzumformung! Soll eine Gleichung durch Wurzel Ziehen umgeformt werden, so muss man sich zuvor überzeugen, dass die Vorzeichen auf beiden Seiten überein stimmen. Dies ist im obigen Beispiel (die Ausdrücke in den Klammern) nicht der Fall, und daher hätten wir schreiben müssen

$$
\begin{aligned}
\left(4-\tfrac{9}{2}\right)^2 &= \left(5-\tfrac{9}{2}\right)^2 \qquad \Leftarrow \\
4-\tfrac{9}{2} &= 5-\tfrac{9}{2}.
\end{aligned}
$$

Allgemein muss man bei der Anwendung von Umkehrfunktionen f^{-1} darauf achten, dass die Funktion f, die man „entfernen“ möchte, *injektiv* (siehe Abschnitt 4.3) ist, auf den Definitionsbereichen beider Seiten der Gleichung.

Beispiel 2.4.5. *Normalerweise ist das Quadratwurzel Ziehen nicht erlaubt, weil die Funktion $f(x) = x^2$ sowohl x als auch $-x$ auf x^2 abbildet; also das Ziehen der Wurzel nicht eindeutig ist! Schränken wir aber f auf positive reelle Zahlen ein, so vermeiden wir dieses Problem und können gefahrlos Wurzel ziehen.*

Sei $x \geq 0$, und seien a, b reelle Zahlen. Dann gilt

$$\begin{aligned} 4x^2 &= (a^2+b^2)^2 \\ 2x &= a^2+b^2 \\ x &= \tfrac{1}{2}(a^2+b^2), \end{aligned}$$

und diese Umformung ist richtig, da wir schon wissen, dass $x \geq 0$ und $a^2+b^2 \geq 0$ (warum?) gelten.

Ist die Anwendung der Umkehrfunktion zwingend nötig, um eine Rechnung fortsetzen zu können, so müssen wir bei Mehrdeutigkeit Fallunterscheidungen durchführen.

Um wieder zum Beispiel „Quadratwurzel" zurückzukehren, sehen wir uns an, wie der vorletzte Umformungsschritt im falschen Beweis von Theorem 2.4.4 richtigerweise geführt hätte werden müssen.

$$\begin{aligned} \left(4-\tfrac{9}{2}\right)^2 &= \left(5-\tfrac{9}{2}\right)^2 \\ 4-\tfrac{9}{2} &= \pm(5-\tfrac{9}{2}) \\ \text{1. Fall: Vorzeichen } +: & \\ 4-\tfrac{9}{2} &= 5-\tfrac{9}{2} \\ -\tfrac{1}{2} &= \tfrac{1}{2} \quad \text{ist offensichtlich falsch} \\ \text{2. Fall: Vorzeichen } -: & \\ 4-\tfrac{9}{2} &= -\left(5-\tfrac{9}{2}\right) \\ -\tfrac{1}{2} &= -\tfrac{1}{2} \quad \text{was stimmt.} \end{aligned}$$

Der 1. Fall führt offensichtlich zu einem unsinnigen Ergebnis und muss daher verworfen werden. Der 2. Fall hingegen liefert das richtige Resultat. Nur dieser darf im Beweis des Theorems verwendet werden und wir sind daher erwartungsgemäß nicht in der Lage, die Behauptung $4 = 5$ zu beweisen.

2.5. Vollständige Induktion

Wir haben im Abschnitt 2.1 bereits die beiden grundlegenden Beweisprinzipien, den direkten und den indirekten Beweis kennengelernt.

Die erste *Beweisidee*, die wir kennenlernen wollen, wird oft benötigt, wenn wir eine Behauptung *für alle* natürlichen Zahlen beweisen möchten.

Beispiel 2.5.1. *Betrachten wir die folgende Reihe von Ausdrücken.*

$$\begin{aligned} 1 &= 1 = 1^2 \\ 1+3 &= 4 = 2^2 \\ 1+3+5 &= 9 = 3^2 \\ 1+3+5+7 &= 16 = 4^2 \\ 1+3+5+7+9 &= 25 = 5^2 \end{aligned}$$

Nach einem „Intelligenztest" finden wir also heraus, dass die Summe der ersten n ungeraden Zahlen genau das Quadrat von n ergibt.

Nun, besser gesagt hätten wir behaupten sollen, dass wir *vermuten*, dass dem so ist. Die ersten fünf Testbeispiele zu überprüfen ist natürlich nicht genug, um daraus schon auf die allgemeine Aussage schließen zu können, ja nicht einmal das überprüfen der ersten 10 Millionen Fälle würde genügen.

> Was wir benötigen, ist eine Technik, um mit einem Schlag das Resultat *für alle unendlich vielen natürlichen Zahlen auf einmal* zu beweisen.

Machen wir einen Zwischenausflug ins tägliche Leben: Welche Hilfsmittel würden Sie verwenden, um ein Dach zu erklimmen? Wahrscheinlich eine Leiter. Ist es zum Erklimmen einer Leiter wichtig, deren Höhe zu kennen? Nein. Das Wissen um die Technik des Leiterkletterns genügt (abgesehen von Höhenangst und eingeschränkter Kondition — das wollen wir wegabstrahieren).

Was müssen wir wissen, um die Technik des Leiterkletterns zu erlernen. Erstaunlicherweise nur zwei Dinge:

(1) Wie komme ich auf die unterste Leitersprosse? (Leiteranfang)
(2) Wie komme ich von einer Leitersprosse auf die nächst höhere Sprosse? (Leiterschritt)

Finden Sie eine Antwort auf diese beiden Fragen, und kein Dach wird vor Ihnen sicher sein (sofern Sie eine Leiter auftreiben können, die lang genug ist).

Wenn wir nun den Gipfel der Erkenntnis über natürliche Zahlen erklimmen wollen, so gehen wir ganz ähnlich vor. Die mathematische Version des Leiterkletterns heißt **vollständige Induktion**.

Um sie korrekt durchzuführen müssen wir ganz analog zum Leiteranfang erst eine Grundlage, einen Anfang für unsere Behauptung finden. Meist werden wir also unsere für alle natürlichen Zahlen zu beweisende Behauptung erst einmal in einem einfachen Fall überprüfen. Üblicherweise ist das der Fall für $n = 0$ oder $n = 1$ aber jede andere natürliche Zahl kann ebenfalls als **Induktionsanfang** dienen.

Danach müssen wir eine Methode finden, den Leiterschritt zu imitieren. Für so einen Schritt gehen wir davon aus, dass wir uns bereits auf einer Leitersprosse befinden, wir also die Aussage schon bewiesen haben für eine bestimmte natürliche Zahl n. Diese Aussage heißt **Induktionsannahme** oder **Induktionsbehauptung**. Von dieser Sprosse ausgehend müssen wir nun eine Methode finden, die nächst höhere Sprosse zu erklimmen. Im Falle der Leiter ist das ein einfacher Schritt, in der Mathematik ist dazu ein Beweis von Nöten. In diesem **Induktionsschritt** wird aus der Behauptung für n die Aussage für die Zahl $n + 1$ (die nächste Sprosse) hergeleitet.

Ist das geschafft, so ist der **Induktionsbeweis** beendet, und die Behauptung ist tatsächlich für alle natürlichen Zahlen bewiesen (resp. für alle natürlichen Zahlen größer als der Induktionsanfang).

Warum ist das so? Für jede natürliche Zahl können wir die „Induktionsleiter“ so lange hinaufklettern bis die Behauptung auch für diese Zahl bewiesen ist — die Höhe des Daches ist nicht wichtig, so lange wir nur die Technik des Kletterns beherrschen.

Verwenden wir also nun unsere neue Technik, um die Behauptung über die Summe ungerader Zahlen aus Beispiel 2.5.1 zu beweisen.

Proposition 2.5.2 (Summen ungerader Zahlen)**.** *Es gilt für* $n \geq 1$[1]

$$\sum_{k=1}^{n}(2k-1) = n^2.$$

Beweis. Wir beweisen die Aussage mit vollständiger Induktion.

[1]Die Aussage ist zwar für $n = 0$ ebenfalls richtig (und zwar trivialerweise), aber in diesem Zusammenhang weniger natürlich, sodass wir darauf verzichten, sie in die Proposition mit einzubeziehen.

Induktionsanfang: Sei $n = 1$. Es gilt $\sum_{k=1}^{1}(2k-1) = 1 = 1^2$. (Wie gesagt, der Induktionsanfang ist meist leicht.)

Induktionsannahme: Es sei die Behauptung für n bereits bewiesen, also

$$\sum_{k=1}^{n}(2k-1) = n^2.$$

Induktionsschritt: Wir müssen nun die Behauptung für $n+1$ zeigen, also

$$\sum_{k=1}^{n+1}(2k-1) = (n+1)^2$$

beweisen. Beginnen wir den Beweis mit der linken Seite

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

Für diese Umformung haben wir einfach die Definition des Summensymbols Σ verwendet und den letzten Term explizit aufgeschrieben. Durch diesen Trick (ein Standardtrick in Induktionsbeweisen) haben wir auf der rechten Seite einen Term (den Summenausdruck) erzeugt, der in der Induktionsannahme vorkommt. Wir können also die Induktionsannahme einsetzen und erhalten

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

Die rechte Seite ist ein vollständiges Quadrat, und daher können wir fertig umformen

$$n^2 + 2n + 1 = (n+1)^2,$$

und wir haben den Induktionsschritt beendet.

Damit ist alles bewiesen — in einem Schritt für unendlich viele, ja für alle, natürlichen Zahlen. □

Als ein komplexeres Beispiel für die Anwendung der vollständigen Induktion zum Beweis einer wichtigen mathematischen Tatsache behandeln wir im folgenden Abschnitt den *binomischen Lehrsatz*.

2.5.1. Der binomische Lehrsatz. Der binomische Lehrsatz dient der Auflösung von Potenzen der Form $(a+b)^n$ in eine Summe von Produkten. Er lautet:

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

Er begründet sich durch folgende Überlegung: Beim Ausmultiplizieren von n gleichen Binomen $(a+b)$ wird für jedes Produkt aus jedem Binom entweder ein a oder ein b verwendet. Somit entstehen Produkte der Formen $a^n b^0, a^{n-1}b^1, \dots, a^1 b^{n-1}, a^0 b^n$. Die entstehenden Produkte werden additiv verknüpft, bleibt also nur noch die Frage, welche Produkte wie oft entstehen. Diese Frage nach dem *Koeffizienten* wird im binomischen Lehrsatz mit $\binom{n}{k}$ beantwortet. Weil er der Koeffizient in der Entwicklung der Potenz eines Binoms $(a+b)$ ist, heißt er **Binomialkoeffizient**.

Die mathematische Disziplin, die sich unter anderem mit dem Abzählen von Objekten beschäftigt, ist die **Kombinatorik**. Dort besteht eine übliche Lösungsmethode darin, ein Problem durch ein äquivalentes Problem zu ersetzen (die Äquivalenz ist oft schwierig zu zeigen), welches leichter zu lösen ist. Ein im Zusammenhang mit Binomialkoeffizienten stets zitiertes äquivalentes Problem ist das *Pascalsche Dreieck*. Es folgt nachstehenden Regeln:

- Die oberste Ebene enthält eine Position.
- Jede Ebene enthält eine Position mehr als die darüberliegende.
- Jeder Position werden in der darunterliegenden Ebene zwei benachbarte Positionen als Linksuntere und Rechtsuntere zugeordnet.
- Die Linksuntere einer Position ist stets gleich der Rechtsunteren ihrer links benachbarten Position und umgekehrt.
- Um einen Weg zu einer Zielposition zu erhalten, startet man von der einzigen Position der obersten Ebene. Dann geht man immer zur Links- oder Rechtsunteren der aktuellen Position, bis man bei der Zielposition angekommen ist.
- An jeder Position notieren wir dann die Anzahl der Wege, die zu ihr führen. Dabei gilt die Position in der obersten Ebene als Weg zu sich selbst, bekommt also eine 1 zugeordnet.

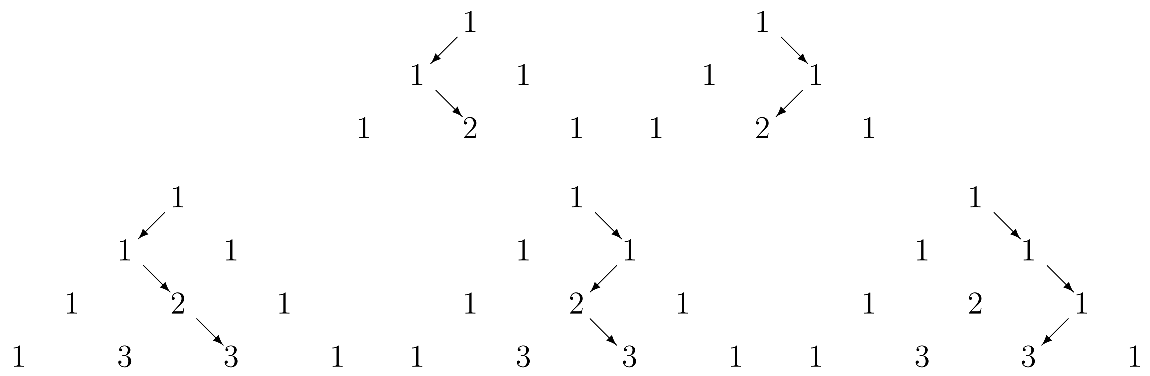


Abbildung 2.1. Pascalsches Dreieck

Der Zusammenhang zwischen dem Pascalschen Dreieck und der Frage, wie oft die einzelnen Produkte beim Ausmultiplizieren auftreten, ist folgender:

- Auf der einen Seite steht beim Finden eines Weges auf jeder Ebene die Entscheidung an, ob man entweder zum Links- oder Rechtsunteren weitergeht.
- Auf der anderen Seite muss man beim Ausmultiplizieren aus jedem Binom entweder ein a oder ein b entnehmen.
- Der an einer Position notierte Wert wird also zum Binomialkoeffizienten des entsprechenden Produktes gleich sein (Dies ist hier noch unbewiesen und wird im Weiteren gezeigt werden.), wobei die Ebene der Potenz entsprechend gewählt werden muss; die Koeffizienten $\binom{n}{k}$ von $(a+b)^n$ findet man also in der n–ten Ebene (wobei wir bei $n = 0$ zu zählen beginnen).

Der Ausdruck $\binom{n}{k}$ beansprucht also, als Ergebnis den Wert der *k–ten Position der n–ten Ebene* des Pascalschen Dreiecks zu haben, wobei die Nummerierung sowohl für n als auch für k mit 0 beginnt. Überlegen wir uns, dass eine Position im Pascalschen Dreieck nur über ihre maximal zwei Oberen zu erreichen ist und alle Wege, zu den beiden Oberen verschieden sind, so ist klarer Weise der Wert einer Position gleich der Summe der Werte ihrer (höchstens zwei) Oberen. Aus dieser Überlegung definieren wir rekursiv.

Definition 2.5.3 (Binomialkoeffizient)**.** *Der* Binomialkoeffizient $\binom{n}{k}$ *für n, k in $\mathbb{N}$ ist (rekursiv) definiert durch*

(i) $\binom{0}{0} := 1$

(ii) $\binom{n}{k} := 0$ *für n in $\mathbb{N}$ und $k < 0$ oder $k > n$,*

(iii) $\binom{n}{k} := \binom{n-1}{k-1} + \binom{n-1}{k}$.

In vielen Situationen (auch im Falle des Binomischen Lehrsatzes) ist die rekursive Definition des Binomialkoeffizienten etwas unhandlich. Wir beweisen daher zunächst eine geschlossene Darstellung desselben.

Proposition 2.5.4 (Geschlossene Darstellung des Binomialkoeffizienten)**.** *Für alle n in $\mathbb{N}$ und alle natürlichen k mit $k \leq n$ gilt:*

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

Beweis. Zu beweisen ist:

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

Dafür müssen wir zeigen, dass die Formel

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

der Definition 2.5.3 von $\binom{n}{k}$ genügt. Dabei haben wir zu beachten, dass die Formel nur für $n \geq 0, 0 \leq k \leq n$ gilt.

Zuerst untersuchen wir die Ränder des Pascalschen Dreiecks und zeigen, dass sie ausschließlich aus Einsen bestehen.

Beginnen wir mit dem linken Rand also dem Fall $k = 0$, d.h. den Binomialkoeffizienten der Form $\binom{n}{0}$, n in $\mathbb{N}$. Aus der rechten Seite der Behauptung ergibt sich tatsächlich

$$\frac{n!}{(n-0)!0!} = \frac{n!}{n!} = 1.$$

Wir müssen nun auch *beweisen*, dass dasselbe aus der rekursiven Definition für $\binom{n}{0}$ folgt. Dazu verwenden wir das Prinzip der vollständigen Induktion:

Behauptung:

$$\text{für alle natürlichen } n \text{ gilt: } \binom{n}{0} = 1$$

Induktionsanfang: $n = 0$

$$\binom{0}{0} = 1 \quad \text{nach Definition 2.5.3(i).}$$

Induktionsannahme: Es gelte

$$\binom{n}{0} = 1.$$

Induktionsschritt:

$$\begin{array}{rcccll} \binom{n+1}{0} & = & \binom{n}{-1} & + & \binom{n}{0} & \text{nach Definition 2.5.3(iii)} \\ & = & 0 & + & 1 & \text{nach Def. 2.5.3(ii) bzw. Induktionsannahme} \\ & = & 1 & & & \end{array}$$

Das beweist die Behauptung über den linken Rand des Pascalschen Dreiecks.

Ganz analog behandeln wir den rechten Rand, also die Binomialkoeffizienten der Form $\binom{n}{n}$. Aus der rechten Seite der Aussage des Satzes berechnen wir

$$\frac{n!}{(n-n)!n!} = \frac{n!}{n!} = 1.$$

Behauptung:

$$\text{für alle natürlichen } n \text{ gilt: } \binom{n}{n} = 1$$

Induktionsanfang: $n = 0$

$$\binom{0}{0} = 1 \quad \text{nach Definition 2.5.3(i).}$$

Induktionsannahme: Es gelte

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

Induktionsschritt:

$$\begin{array}{rcccll} \binom{n+1}{n+1} & = & \binom{n}{n} & + & \binom{n}{n+1} & \text{nach Definition 2.5.3(iii)} \\ & = & 1 & + & 0 & \text{nach Induktionsann. bzw. Def. 2.5.3(ii)} \\ & = & 1 & & & \end{array}$$

Das zeigt die Behauptung über den rechten Rand.

Nun beweisen wir die Formel für alle (restlichen) n und k. Dafür müssen wir nachweisen, dass für alle n in $\mathbb{N}$, $2 \leq n$ und $1 \leq k \leq n-1$

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

gilt. Wir verwenden ein weiteres Mal vollständige Induktion.
Induktionsanfang: $n = 2$, daher $k = 1$

$$\begin{array}{rcccll} \binom{2}{1} & = & \binom{1}{0} & + & \binom{1}{1} & \text{nach Definition 2.5.3(iii)} \\ & = & 1 & + & 1 & \text{nach dem bereits bewiesenen} \\ & = & 2 & & & \end{array}$$

Andererseits gilt

$$\frac{2!}{(2-1)!\,1!} = \frac{2}{1 \cdot 1} = 2$$

Induktionsannahme: Es gelte

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} \quad \text{für } 1 \leq k \leq n-1.$$

Induktionsschritt: Sei $1 \leq k \leq n$ dann gilt:

$$\begin{aligned}
\binom{n+1}{k} &= \binom{n}{k} + \binom{n}{k-1} && \text{Definition 2.5.3(iii)} \\
&= \frac{n!}{(n-k)!k!} + \frac{n!}{(n-k+1)!(k-1)!} && \text{Induktionsann. bzw. obige Beh.} \\
&= \frac{n!(n-k+1)}{(n-k+1)(n-k)!k!} \\
&\quad + \frac{n!k}{(n-k+1)!(k-1)!k} && \text{Erweitern} \\
&= \frac{n!(n-k+1)}{(n-k+1)!k!} + \frac{n!k}{(n-k+1)!k!} && \text{Definition der Fakultät} \\
&= \frac{n!(n-k+1)+n!k}{(n-k+1)!k!} && \text{Zusammenfassen der Brüche} \\
&= \frac{n!(n-k+k+1)}{(n-k+1)!k!} && \text{Herausheben} \\
&= \frac{n!(n+1)}{(n+1-k)!k!} && \text{Addieren} \\
&= \frac{(n+1)!}{(n+1-k)!k!}. && \text{Definition der Fakultät}
\end{aligned}$$

Das beweist, dass die Formel der rekursiven Darstellung von $\binom{n}{k}$ genügt. □

Zum Rechnen mit der Formel aus Proposition 2.5.4 empfiehlt es sich, zu kürzen:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k!} = \frac{\prod_{i=0}^{k-1}(n-i)}{k!}. \tag{2.2}$$

Mit Hilfe der mittels Proposition 2.5.4 nachgewiesenen Formel (2.2) können wir die Definition des Binomialkoeffizienten erweitern: Zunächst haben wir—motiviert durch unsere Betrachtung des Pascalschen Dreieck—in Definition 2.5.3 den Binomialkoeffizienten $\binom{n}{k}$ ja lediglich für n, k in $\mathbb{N}$ definiert. Auf der rechten Seite von (2.2) besteht allerdings keine Notwendigkeit mehr, n auf natürliche Zahlen zu beschränken. So definieren wir wie folgt:

Definition 2.5.5 (Binomialkeffizent — Erweiterung)**.** *Der* Binomialkoeffizient *ist für α in $\mathbb{R}$ und k in $\mathbb{N}$ definiert durch:*

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!} = \frac{\prod_{i=0}^{k-1}(\alpha-i)}{k!}.$$

In der Mathematik ergeben sich häufig Situationen, wo ein Begriff (wie etwa oben der Binomialkoeffizient) zunächst in einer spezielle(re)n Situation definiert wird, es sich nach gründlichem Studium allerdings zeigt, dass der Begriff auch in allgemeineren Situationen Sinn hat. Dann *erweitert* man die Definition. In diesem Sinne ist Definition 2.5.5 eine Erweiterung der Definition 2.5.3. Dabei müssen wir allerdings sorgfältig darauf achten, dass im spezielleren Fall (hier n in $\mathbb{N}$) die erweiterte Definition mit der ursprünglichen übereinstimmt.

Kehren wir nun nach diesem Ausflug in die Kombinatorik zum Binomischen Lehrsatz zurück, den wir zum Abschluss dieses Kapitels beweisen.

Proposition 2.5.6. *(Binomischer Lehrsatz) Es gilt für alle a, b reell und n in $\mathbb{N}$:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Beweis. Zu zeigen ist also

$$\text{für alle } n \text{ in } \mathbb{N} \text{ und alle } a, b \text{ in } \mathbb{R}\text{: } (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Wir beweisen das mittels vollständiger Induktion.
Induktionsanfang: $n = 0$
Klarerweise gilt $(a+b)^0 = 1$. Andererseits

$$\sum_{k=0}^{0} \binom{0}{k} a^k b^{0-k} \;=\; \binom{0}{0} a^0 b^0 \;=\; 1 \cdot 1 \cdot 1 \;=\; 1$$

Induktionsannahme: Es gelte

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}$$

Induktionsschritt:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n && \\
&= (a+b)\sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j} && \text{Induktionsannahme} \\
&= \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}(a+b) && \text{Ausmultiplizieren} \\
&= \sum_{j=0}^{n} \binom{n}{j} (a^{j+1} b^{n-j} + a^j b^{n-j+1}) && \text{Ausmultiplizieren} \\
&= \sum_{j=0}^{n} \left(\binom{n}{j} a^{j+1} b^{n-j} + \binom{n}{j} a^j b^{n-j+1} \right) && \text{Ausmultiplizieren} \\
&= \sum_{j=0}^{n} \binom{n}{j} a^{j+1} b^{n-j} + \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j+1} && \text{Aufspalten der Summe} \\
&= \sum_{i=1}^{n+1} \binom{n}{i-1} a^i b^{n-i+1} + \sum_{j=0}^{n+1} \binom{n}{j} a^j b^{n-j+1} && \begin{array}{l}\text{Indexverschiebung} \\ j+1=i \text{ und } \binom{n}{n+1}=0\end{array} \\
&= \sum_{k=0}^{n+1} \binom{n}{k-1} a^k b^{n-k+1} + \sum_{k=0}^{n+1} \binom{n}{k} a^k b^{n-k+1} && \begin{array}{l}\binom{n}{-1}=0 \text{ und} \\ \text{Laufvariablen umbenannt}\end{array} \\
&= \sum_{k=0}^{n+1} \left(\binom{n}{k-1} a^k b^{n-k+1} + \binom{n}{k} a^k b^{n-k+1} \right) && \text{Vereinigen der Summen} \\
&= \sum_{k=0}^{n+1} a^k b^{n-k+1} \left(\binom{n}{k-1} + \binom{n}{k} \right) && \text{Herausheben} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k} && \text{rekursive Definition von } \binom{n}{k}
\end{aligned}$$

Das beweist den binomischen Lehrsatz. □

KAPITEL 3

Logik

Dieses Kapitel handelt von *den* Grundlagen der Mathematik. Grundlegendes über Booleschen Algebren sollte schon aus der Schule bekannt sein. Versteht man erst das Prinzip von Booleschen Algebren, so ist damit schon der erste Schritt zum Verständnis der Aussagenlogik getan. Wir erklären die grundlegenden Operationen Und, Oder und die Negation sowie die Implikation und die Äquivalenz. Schließlich befassen wir uns mit *Quantoren* und legen somit den Grundstein zur Einführung unserer ersten mathematische Struktur — der *Mengen* — in Kapitel 4.

3.1. Boolesche Algebren

In diesem Abschnitt wollen wir das Kapitel über Boolesche Algebren aus der Schule aufarbeiten. Es soll uns nicht dazu dienen, daraus die Grundlagen der Mathematik zu bauen (was im Prinzip möglich wäre), sondern lediglich die Grundoperation der Aussagenlogik motivieren. Wir beschränken uns dabei auf die *Schaltalgebra*, ein Konzept, das auch für das Verständnis der Informatik von großer Bedeutung ist.

Elektronische (auch elektrische) Schaltungen bestehen aus elektrischen Leitungen und aus Schaltern. Jede Leitung kann sich in zwei Zuständen befinden (Strom führend bzw. nicht Strom führend), so wie jeder Schalter zwei Zustände (Stellungen) hat: „Ein“ und „Aus“.

Mathematisch kann man sowohl den Zustand einer Leitung als auch die Stellung eines Schalters mit Hilfe einer Variable beschreiben, die zwei Werte annehmen kann: 0 oder 1. Eine solche Variable heißt *binäre Variable*.

Mit Schaltern kann man steuern, ob Strom durch eine bestimmte Leitung fließt oder nicht. Das heißt, die Schalterzustände steuern die Zustände von Leitungen. Schaltet man den Schalter ein, so lässt er den Strom passieren, und ergibt sich ein geschlossener Stromkreis, so fließt Strom durch die Leitung. In der Computertechnik wurden mit Hilfe von Transitoren Schaltungen entwickelt, die wie elektronische Schalter funktionieren. Führt dort eine bestimmte Leitung A Strom, so verhält sie sich wie ein Schalter im Zustand „Ein“ für eine andere Leitung B. Fließt kein Strom durch Leitung A, so verhält sie sich wie ein Schalter im „Aus“-Zustand für Leitung B.

Baut man eine komplizierte Schaltung aus mehreren Schaltern, die durch Leitungen verbunden sind, so ist meist auf den ersten Blick nicht zu erkennen, welche Leitungen bei welchen Schalterstellungen Strom führen und welche nicht. Man kann sich dann einen Überblick verschaffen, indem man so genannte Schaltwerttabellen aufstellt. An einigen einfachen Schaltungen sei das Prinzip demonstriert.

(1) Setzt man in einem Stromkreis wie in Abbildung 3.1 zwei Schalter hintereinander, bildet man also eine *Serienschaltung*, und untersucht, wann die Leitung Strom führt, erhält man folgende Schaltwerttabelle. Die Bedeutung der Tabelle ist rechts daneben noch einmal explizit erläutert.

a	b	$a \wedge b$	
0	0	0	$0 \wedge 0 = 0$
0	1	0	$0 \wedge 1 = 0$
1	0	0	$1 \wedge 0 = 0$
1	1	1	$1 \wedge 1 = 1$

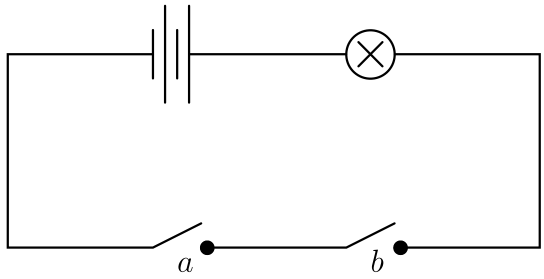


Abbildung 3.1. Serienschaltung — Und-Verknüpfung

Der Strom fließt also, wenn Schalter a **und** Schalter b eingeschaltet sind. Mathematisch schreibt man kurz $a \wedge b$ und spricht a **und** b bzw. von der **Und-Verknüpfung** oder AND-Verknüpfung.

(2) Setzt man in einem Stromkreis wie in Abbildung 3.2 zwei Schalter nebeneinander, so wird man folgendes feststellen: Damit die Leitung Strom führt, reicht es

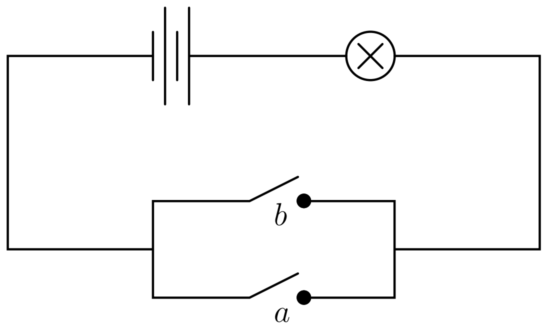


Abbildung 3.2. Parallelschaltung — Oder-Verknüpfung

Schalter a **oder** Schalter b einzuschalten. Eine Schaltung dieser Art nennt man *Parallelschaltung* und die entsprechende mathematische Verknüpfung heißt **Oder-Verknüpfung** bzw. OR-Verknüpfung. Man schreibt $a \vee b$ und spricht a **oder** b. Die Schaltwerttabelle ist

a	b	$a \vee b$
0	0	0
0	1	1
1	0	1
1	1	1

ACHTUNG: Beachten Sie, dass „oder“ im Gegensatz zum umgangssprachlichen Gebrauch bedeutet, dass a oder b **oder beide** eingeschaltet sein müssen (vgl. Beispiel 3.1.2 und Abschnitt 3.2.1 unten). Wir sprechen vom *einschließenden Oder*.

(3) Beschriftet man einen Schalter „verkehrt“, so erhält man die einfachste Schaltung, die **Negation** $\neg a$ mit der Schaltwerttabelle

a	$\neg a$
0	1
1	0

Bemerkung 3.1.1. *Mit* elektrischen *Leitungen und echten Schaltern kann man nicht so leicht komplizierte Schaltungen bauen. Mit* elektronischen *Schaltern hingegen kann man auch Schaltungen bauen, in denen eine Leitung den Strom in mehreren anderen Leitungen schaltet. Mit dieser Technik kann man aus den drei Grundschaltungen Serienschaltung ($\wedge$), Parallelschaltung ($\vee$) und Negation ($\neg$) jede beliebige Schaltung bauen.*

Zum besseren Verständnis der Grundschaltungen bringen wir noch einen Vergleich aus dem „wirklichen Leben“. Wenn Sie als Abenteurer in einem Fantasy-Spiel in ein Haus eindringen müssen, dann werden Sie zuerst die Türen untersuchen. Besitzt eine Tür zwei Schlösser A und B, so müssen Sie A **und** B öffnen, um die Tür zu überwinden. Hat das Haus aber zwei Türen a und b, so müssen Sie a **oder** b öffnen, um einzudringen. Dies ist ein **einschließendes Oder**, denn wenn sie beide Türen aufbekommen, ist das bestimmt kein Hindernis für das Durchsuchen des Hauses — und falls Sie an der logischen Aufgabe mit den Türen und Schlössern scheitern, können Sie immer noch mit Hilfe der vollständigen Induktion ein Fenster im zweiten Stock einschlagen.

Es existieren vier einstellige Operatoren (wie $\neg$) und 16 mögliche binäre Operatoren (wie $\wedge$ oder $\vee$).Über zwei dieser binären Operatoren wollen wir im Folgenden sprechen.

Beispiel 3.1.2 (XOR)**.** *Betrachten wir zunächst die Schaltwerttabelle*

a	b	$a \veebar b$
0	0	0
0	1	1
1	0	1
1	1	0

Diese zweistellige Operation heißt XOR (exklusives Oder, **ausschließendes Oder***). Sie entspricht der Bedeutung von „oder“ in der Umgangssprache:* Entweder a oder b *sind eingeschaltet — keinesfalls beide.*

WICHTIG: In der Mathematik ist es unbedingt notwendig, das Ausschließende der XOR-Operation zu betonen, wie etwa durch Einführen des Wortes „entweder“, um Verwechslungen mit der OR-Operation zu vermeiden, die ja als einschließende Oder einen Einser in der letzten Zeile der Schaltwerttabelle aufweist (vgl. auch Abschnitt 3.2.1).

Beispiel 3.1.3 (NAND)**.** *Interessanterweise gibt es eine Operation — übrigens sehr billig mittels Transistoren herstellbar — die allein ausreicht, um alle anderen Operationen und damit alle möglichen Schaltungen zu erzeugen. Diese binäre Operation hat die Schaltwerttabelle*

a	b	$a \bar{\wedge} b$
0	0	1
0	1	1
1	0	1
1	1	0

und trägt den Namen NAND *(negated AND, also negiertes Und). Der Zusammenhang mit den bereits definierten Operationen ist $a \bar{\wedge} b = \neg(a \wedge b)$, wie leicht aus den Schaltwertabellen zu sehen ist.*

Bemerkung 3.1.4. *Wie können die bereits bekannten Grundoperationen mit Hilfe der NAND Operation zusammensetzt werden?*

(1) *Es gilt $\neg a = a \bar{\wedge} a$, wie wir an Hand der Schaltwerttabelle leicht überprüfen können:*

a	$a \bar{\wedge} a$	$\neg a$
0	1	1
1	0	0

(2) *Für die Oder-Verknüpfung erhalten wir* $a \vee b = (a \bar{\wedge} a) \bar{\wedge} (b \bar{\wedge} b)$*:*

a	b	$a \bar{\wedge} a$	$b \bar{\wedge} b$	$(a \bar{\wedge} a) \bar{\wedge} (b \bar{\wedge} b)$	$a \vee b$
0	0	1	1	0	0
0	1	1	0	1	1
1	0	0	1	1	1
1	1	0	0	1	1

(3) *Zuletzt stellen wir die Und-Verknüpfung ebenfalls durch drei NAND Operationen dar als* $a \wedge b = (a \bar{\wedge} b) \bar{\wedge} (a \bar{\wedge} b)$. *Überprüfen wir die Richtigkeit wieder mit Hilfe der Schaltwerttabelle:*

a	b	$a \bar{\wedge} b$	$(a \bar{\wedge} b) \bar{\wedge} (a \bar{\wedge} b)$	$a \wedge b$
0	0	1	0	0
0	1	1	0	0
1	0	1	0	0
1	1	0	1	1

Eine wichtige Frage bei der technischen Herstellung von Schaltungen ist die folgende: Es sei festgelegt, bei welchen Schalterstellungen welche Leitungen Strom führen sollen und welche nicht; es sei also die Schalttafel gegeben. Was ist die einfachste Schaltung, die genau diese Schalttafel besitzt?

Diese Frage zu beantworten ist nicht ganz einfach. Es ist sicher, dass es eine Schaltung gibt, die der Schalttafel entspricht. Man kann sie auch immer konstruieren mit Hilfe der sogenannten **disjunktiven Normalform**. Es sei also eine Funktion f gegeben, deren Wert 0 oder 1 ist und von den binären Variablen $a_1, \ldots, a_n$ abhängt. Möchte man eine Schaltung konstruieren mit n Schaltern, die den Variablen entsprechen, die immer den Wert $f(a_1, \ldots, a_n)$ ergibt, so folgt man dem folgenden *Algorithmus*:

(1) Stelle die Schaltwerttabelle mit den Variablen links und dem gewünschten Funktionswert rechts auf.
(2) Streiche alle Zeilen, in denen $f(a_1, \ldots, a_n)$ den Wert 0 hat.
(3) Ordne jeder der verbliebenen Zeilen eine Und-Verknüpfung von allen Variablen a_i zu, die in dieser Zeile den Wert 1 haben und von den Negationen $\neg a_j$ aller Variablen, die in dieser Zeile den Wert 0 haben.
(4) Verknüpfe alle gerade konstruierten Und-Glieder durch Oder-Verknüpfungen.

Beispiel 3.1.5. *Konstruieren wir die disjunktive Normalform zur Schaltwerttabelle*

a	b	c	$f(a,b,c)$	*Und-Verknüpfung*
0	0	0	1	$\neg a \wedge \neg b \wedge \neg c$
0	0	1	0	
0	1	0	1	$\neg a \wedge b \wedge \neg c$
0	1	1	1	$\neg a \wedge b \wedge c$
1	0	0	1	$a \wedge \neg b \wedge \neg c$
1	0	1	0	
1	1	0	0	
1	1	1	1	$a \wedge b \wedge c$

Die disjunktive Normalform ist dann

$$f(a,b,c) = (\neg a \wedge \neg b \wedge \neg c) \vee (\neg a \wedge b \wedge \neg c) \vee (\neg a \wedge b \wedge c) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge b \wedge c).$$

Die disjunktive Normalform ist übrigens nicht die einzige Möglichkeit, zu einer gegebenen Schaltwerttabelle eine Schaltung zu konstruieren. Es existiert zum Beispiel auch noch die **konjunktive Normalform**, die sich grob gesprochen dadurch auszeichnet, dass sie eine Und-Verknüpfung von Oder-Ausdrücken ist. Konstruiert wird sie mit einem analogen (dualen) Algorithmus:

(1) Stelle die Schaltwerttabelle mit den Variablen links und dem gewünschten Funktionswert rechts auf.
(2) Streiche alle Zeilen, in denen $f(a_1, \dots, a_n)$ den Wert 1 hat.
(3) Ordne jeder der verbliebenen Zeilen eine Oder-Verknüpfung von allen Variablen a_i zu, die in dieser Zeile den Wert 0 haben und von den Negationen $\neg a_j$ aller Variablen, die in dieser Zeile den Wert 1 haben.
(4) Verknüpfe alle gerade konstruierten Oder-Glieder durch Und-Verknüpfungen.

Die Normalformen zu einem Ausdruck sind üblicherweise sehr kompliziert, und die Frage ist, ob man eine einfachere Schaltung konstruieren kann, die dieselbe Schaltwerttabelle ergibt. Tatsächlich können komplizierte Verknüpfungen mit Hilfe der folgenden Rechenregeln vereinfacht werden (vgl. Beispiel 3.1.8), die man leicht mit Hilfe der jeweiligen Schaltwerttabellen überprüfen kann.

Theorem 3.1.6 (Rechenregeln für logische Operatoren)**.** *Für die Operationen $\wedge$, $\vee$ und $\neg$ gelten die folgenden Rechenregeln.*

Kommutativgesetze:	$a \vee b = b \vee a$	$a \wedge b = b \wedge a$
Assoziativgesetze:	$a \vee (b \vee c) = (a \vee b) \vee c$	$a \wedge (b \wedge c) = (a \wedge b) \wedge c$
Distributivgesetze:	$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$	$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$
Verschmelzungsgesetze:	$a \vee (b \wedge a) = a$	$a \wedge (b \vee a) = a$
Idempotenzgesetze:	$a \vee a = a$	$a \wedge a = a$
Neutralitätsgesetze:	$a \vee 0 = a$	$a \wedge 1 = a$
Absorptionsgesetze:	$a \vee 1 = 1$	$a \wedge 0 = 0$
Komplementaritätsgesetze:	$a \vee \neg a = 1$	$a \wedge \neg a = 0$
	$\neg 0 = 1$	
	$\neg 1 = 0$	
Gesetz der doppelten Verneinung:	$\neg(\neg a) = a$	
Gesetze von De Morgan:	$\neg(a \vee b) = \neg a \wedge \neg b$	$\neg(a \wedge b) = \neg a \vee \neg b$

Beweis. Aufstellen der Schaltwerttabellen. □

Bemerkung 3.1.7 (Boolesche Algebren)**.** *Eine mathematische Struktur mit* 0, 1 *und drei Operationen $\wedge$, $\vee$ und $\neg$, die die Rechengesetze*

(1) *Kommutativgesetze*
(2) *Distributivgesetze*
(3) *Neutralitätsgesetze*
(4) *Komplementaritätsgesetze*

erfüllt, heißt Boolesche Algebra. *Alle anderen Rechengesetze aus Theorem 3.1.6 lassen sich aus diesen acht herleiten.*

Beispiel 3.1.8. *Mit Hilfe der Rechengesetze aus Theorem 3.1.6 können wir versuchen, die disjunktive Normalform aus Beispiel 3.1.5 zu vereinfachen.*

$$\begin{aligned}
f(a,b,c) &= (\neg a \wedge \neg b \wedge \neg c) \vee (\neg a \wedge b \wedge \neg c) \vee (\neg a \wedge b \wedge c) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge b \wedge c) = \\
&= \Big(\neg a \wedge \big((\neg b \wedge \neg c) \vee (b \wedge \neg c) \vee (b \wedge c)\big)\Big) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge b \wedge c) = \\
&= \Big(\neg a \wedge \big(((\neg b \vee b) \wedge \neg c) \vee (b \wedge c)\big)\Big) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge b \wedge c) = \\
&= \Big(\neg a \wedge \big((1 \wedge \neg c) \vee (b \wedge c)\big)\Big) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge b \wedge c) = \\
&= \Big(\neg a \wedge \big(\neg c \vee (b \wedge c)\big)\Big) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge b \wedge c) = \\
&= (\neg a \wedge \neg c) \vee (\neg a \wedge b \wedge c) \vee (a \wedge \neg b \wedge \neg c) \vee (a \wedge b \wedge c) = \\
&= (\neg a \wedge \neg c) \vee (a \wedge \neg b \wedge \neg c) \vee (\neg a \wedge b \wedge c) \vee (a \wedge b \wedge c) = \\
&= (\neg a \wedge \neg c) \vee (a \wedge \neg b \wedge \neg c) \vee \big((\neg a \vee a) \wedge (b \wedge c)\big) = \\
&= (\neg a \wedge \neg c) \vee (a \wedge \neg b \wedge \neg c) \vee \big(1 \wedge (b \wedge c)\big) = \\
&= \big((\neg a \vee (a \wedge \neg b)) \wedge \neg c\big) \vee (b \wedge c) = \\
&= \big((\neg a \vee a) \wedge (\neg a \vee \neg b) \wedge \neg c\big) \vee (b \wedge c) = \\
&= \big(1 \wedge (\neg a \vee \neg b) \wedge \neg c\big) \vee (b \wedge c) = \\
&= \big((\neg a \vee \neg b) \wedge \neg c\big) \vee (b \wedge c) = \\
&= \big(\neg(a \wedge b) \wedge \neg c\big) \vee (b \wedge c) = \\
&= \neg\big((a \wedge b) \vee c\big) \vee (b \wedge c)
\end{aligned}$$

dies ist schon eine wesentlich kompaktere Formel, und an Hand der Schaltwerttabelle kann man leicht überprüfen, dass diese Formel eine äquivalente Schaltung beschreibt.

Beispiel 3.1.9 (Implikation und Äquivalenz)**.** *Zwei weitere Beispiele binärer Operationen, die im folgenden noch wichtig sein werden sind die* Implikation *und die* Äquivalenz

a	b	$a \Rightarrow b$
0	0	1
0	1	1
1	0	0
1	1	1

und

a	b	$a \Leftrightarrow b$
0	0	1
0	1	0
1	0	0
1	1	1

In elementaren Operationen ausgedrückt finden wir die konjunktive Normalform

$$a \Rightarrow b = \neg a \vee b$$

sowie die disjunktive Normalform

$$a \Leftrightarrow b = (\neg a \wedge \neg b) \vee (a \wedge b).$$

3.2. Aussagen, Logik

In der Mathematik werden Begriffe und Regeln der Logik verwendet, um das Theoriegebäude zu erbauen. Die Mathematik arbeitet dabei mit Aussagen. Das hervorstechende Merkmal einer Aussage ist dabei:

Eine **Aussage** ist entweder **wahr** oder **falsch**.

Beispiel 3.2.1 (Aussagen)**.** *Beispiele für Aussagen sind etwa:*

- 7 *ist größer als* 5*, oder in Zeichen* $7 > 5$.
- *Es gibt unendlich viele Primzahlen.*
- *Wale sind Säugetiere.*

Die folgenden Sätze sind keine *Aussagen:*

- *Wer geht heute ins Clubbing?*
- $5 + 8$

Eine Besonderheit der Mathematik besteht darin, dass zu Beginn als Fundament der gesamten Wissenschaft eine Reihe von Aussagen, die **Axiome** als *wahr angenommen* werden. Danach werden ausgehend von diesen Aussagen weitere **wahre** Aussagen abgeleitet. Gewissermaßen könnte man also sagen, dass sich die Mathematiker eine eigene streng logisch aufgebaute „Welt“ erschaffen, in der sie niemals lügen (d.h. sie machen nur wahre Aussagen). Die Gültigkeit dieser Aussagen wird dadurch sicher gestellt, dass sie durch definierte logische Umformungsschritte aus bereits als wahr erkannten Aussagen abgeleitet werden (auch was ableiten bedeutet, kann man exakt definieren — das ist aber Gegenstand der Vorlesungen aus dem Gebiet „Logik“). Diesen Vorgang nennt man **beweisen**.

3.2.1. Und oder oder, oder nicht? Nachdem Aussagen *zwei* mögliche „Werte“ haben können, kann man sie mit den gleichen Augen betrachten wie Schalter oder Stromleitungen, und man kann genau dieselben Verknüpfungen von Aussagen machen wie man aus Schaltern und Leitungen Schaltungen bauen kann. Man beachte, dass bei der Untersuchung von Aussagen an Stelle von Schaltungen die Schaltwerttabellen als **Wahrheitstafeln** bezeichnen werden.

Setzen wir in den Tabellen für **wahr** den Wert **1** und für **falsch** den Wert **0** und werfen wir noch einmal einen Blick auf die drei Grundoperationen, und versuchen wir zu klären, was sie im Zusammenhang mit Aussagen bedeuten.

3.2.1.1. Oder ($\vee$). Bei der Definition der Oder-Verknüpfung muss man aufmerksam sein, und daher wollen wir sie zu Beginn behandeln.

Die Aussage

Peter ist Professor **oder** Student.

bedeutet, dass Peter Professor oder Student *oder beides* ist. Das Oder in der Mathematik ist (wie wir schon aus Abschnitt 3.1 wissen) ein *einschließendes Oder* — im Gegensatz zum umgangssprachlichen Gebrauch. Das entspricht auch der Tabelle zur Verknüpfung $\vee$.

Ein Oder in einer mathematische Aussage ist immer als einschließenden Oder zu verstehen. Möchte man in einer mathematischen Aussage ein Oder so verstanden wissen, dass es, ähnlich zur Umgangssprache, das „oder beides“ ausschließt, möchte man also statt einem einschließenden Oder ein ausschließendes Oder verwenden, so muss man das explizit machen, indem man beispielsweise formuliert:

Peter ist **entweder** Professor **oder** Student.

und eventuell sogar hinzufügt:

Aber nicht beides.

Merke: Hat man zwei Aussagen p und q, dann ist $p \vee q$ (in Sprache p *oder* q) wahr, wenn p *oder* q *oder beide* wahr sind.

So ist den meisten SchülernInnen und Studierenden die Aussage, „Um eine Prüfung zu bestehen, muss man viel lernen *oder* gut schummeln" allzu gut bekannt.

3.2.1.2. Und ($\wedge$). Während die Oder-Verknüpfung einigen Erklärungsbedarf nach sich gezogen hat, ist die **Und-Verknüpfung** aus der Umgangssprache intuitiv klar.

Was bedeutet die folgende Aussage?

Die Zahl 6 ist durch 3 teilbar **und** die Zahl 6 ist durch 2 teilbar.

Klarerweise ist diese Aussage eine Und-Verknüpfung ($\wedge$) der beiden Aussagen „6 ist durch 3 teilbar" und „6 ist durch 2 teilbar". Beide diese Aussagen sind wahr, also ist auch die Und-Verknüpfung der beiden Aussagen wahr, und damit ist auch die Aussage von oben.

Merke: Hat man zwei Aussagen p und q, dann ist $p \wedge q$ (in Sprache p *und* q) wahr, wenn p und q *beide* wahr sind.

Im Gegensatz zu beliebigen Prüfungen seien die StudentInnen aber gewarnt, dass für die Prüfung zur Einführungsvorlesung gilt: „Zum Bestehen der Prüfung müssen die Studierenden viel lernen *und* gut schummeln."

3.2.1.3. Negation ($\neg$). Die Negation einer Aussage ist klarerweise deren Verneinung. Wenn wir etwa die Negation der Aussage

Der Fußboden ist blau.

bilden, so erhalten wir natürlich

Der Fußboden ist **nicht** blau.

ACHTUNG: „Der Fußboden ist gelb" ist **keine** Verneinung der obigen Aussage!

Interessant wird es, wenn wir Aussagen verneinen, in denen bereits Verknüpfungen $\vee$ oder $\wedge$ vorkommen. Dann müssen wir achtgeben. Hier helfen uns die Untersuchungen aus Abschnitt 3.1 weiter, denn in Theorem 3.1.6 haben wir die Regeln von De Morgan kennen gelernt, die uns Aufschluss darüber geben, was passiert, wenn man Und- und Oder-Verknüpfungen negiert. Betrachten wir einige Beispiele:

- Verneint man

  Der Fußboden ist blau und die Decke ist grün.

  so erhält man

  Der Fußboden ist nicht blau **oder** die Decke ist nicht grün.
- Will man dagegen die Aussage

  Die Zahl 3 ist eine Primzahl oder die Zahl 4 ist eine Primzahl.

  negieren, so muss man folgendermaßen formulieren.

  Die Zahl 3 ist keine Primzahl **und** die Zahl 4 ist keine Primzahl.

Merke: Will man $\wedge$- oder $\vee$-Verknüpfungen von Aussagen verneinen, so verneint man die Einzelaussagen und tauscht dann $\wedge$ gegen $\vee$ aus. Es gelten also die Regeln von De Morgan

$$\neg(p \wedge q) = \neg p \vee \neg q \qquad \neg(p \vee q) = \neg p \wedge \neg q.$$

Die letzte wichtige Regel für Negationen ist schließlich, dass doppelte Verneinungen wegfallen:

Wale sind nicht keine Säugetiere.

bedeutet dasselbe wie

Wale sind Säugetiere.

Merke: Doppelte Verneinungen fallen weg. Es gilt $\neg(\neg p) = p$.

Beispiel 3.2.2. *Trifft ein Informatiker seinen Freund, der mit rauchendem Kopf verzweifelt vor dem Computer sitzt. Weil er aus ihm kein vernünftiges Wort herausbringt, blickt er kurz auf den Monitor und liest:* `Nicht alle Dateien nicht löschen? (J/N)`.

Schließlich sei an dieser Stelle bemerkt, dass wir — obwohl wir in der Mathematik natürlich hauptsächlich an wahren Aussagen interessiert sind — oft mit Aussagen zu tun haben, über deren Wahrheitswert wir nicht unmittelbar Bescheid wissen (etwa auf dem Weg zu oder der Suche nach einer wahren Aussage) oder auch mit falschen Aussagen (etwa im Rahmen indirekter Beweise) konfrontiert sind. Daher ist es wichtig mit Aussagen unabhängig von ihrem Wahrheitswert hantieren (etwa die Negation bilden) zu lernen/können. Dies bereitet den AnfängerInnen manchmal Schwierigkeiten: Unterscheiden Sie also sorgfältig zwischen dem Wahrheitsgehalt einer Aussage und einer Manipulation dieser Aussage. So haben wir oben z.B. die Aussage „Die Zahl 3 ist eine Primzahl oder die Zahl 4 ist eine Primzahl.“ zu „Die Zahl 3 ist keine Primzahl und die Zahl 4 ist keine Primzahl.“ verneint. Die „Funktionsweise“ der Verneinung ist unabhängig von der Tatsache, dass die erste Aussage wahr ist und die zweite falsch. So haben wir ja auch Aussagen über Fußboden und Decke verneint, denen unmittelbar kein Wahrheitswert zugeordnet werden kann.

3.2.2. Implikation und Äquivalenz. Wie versprochen tauchen die in Beispiel 3.1.9 eingeführten binäre Operationen hier an wichtiger Stelle wieder auf.

3.2.2.1. Die Implikation ($\Rightarrow$). Wir haben schon diskutiert, dass in der Mathematik neue Aussagen aus bereits bekannten Resultaten *abgeleitet* werden. Werfen wir nun einen genaueren Blick auf diesen Vorgang. **Alle** mathematischen Sätze haben bei genauer Betrachtung das folgende Aussehen:

Theorem 3.2.3. *Aus* den Voraussetzungen *folgt* das Resultat.

Genauer: Ein Theorem ist **eine Aussage** der Form: Voraussetzungen $\Rightarrow$ Resultat. Der Beweis stellt sicher dass diese Aussage **wahr** ist.

Was das bedeutet, können wir erst erkennen, wenn wir die Wahrheitstafel der $\Rightarrow$-Operation noch einmal betrachten.

p	q	$p \Rightarrow q$
0	0	1
0	1	1
1	0	0
1	1	1

Wir erkennen, dass es nur *einen Fall* gibt, in dem die Aussage einer Implikation *falsch* ist, nämlich wenn die Voraussetzung wahr aber die Folgerung falsch ist. Das entspricht durchaus unserer Intuition. Ebenso steht wohl die letzte Zeile der Wahrheitstabelle außer Diskussion: Wenn die Voraussetzung und Resultat wahr ist, dann ist auch die Implikation wahr.

Eine spezielle Betrachtung verdienen die beiden Fälle, in denen p, also die Voraussetzung, falsch ist. In diesen Fällen ist die Aussage der Implikation nämlich wahr unabhängig vom Wahrheitswert des Resultats („ex falso quodlibet“ — lat. aus falschem wie es beliebt). Diese mathematische Definition widerspricht ein wenig der sprachlichen Intuition und es hat sich gezeigt, dass diese Tatsache zu Beginn meist (philosophische) Probleme bereitet.

Überlegen wir: Der Ergebniswert kann in beiden Fällen nur 0 oder 1 sein, denn eine dritte Möglichkeit kennt die formale (zweiwertige) Logik nicht („tertium non datur!“). Ein pragmatischer Standpunkt wäre nun zu sagen: „Wir wollen möglichst viele wahre Aussagen in unseren Theorien haben, und daher setzen wir an beide Stellen 1.“ Das macht Sinn, denn wir wollen mit dem Theorem nur Aussagen machen über Fälle, in denen die Voraussetzungen erfüllt sind, und alle anderen Fälle wollen wir nicht betrachten. Dann soll das Theorem immer noch wahr sein, auch wenn die Voraussetzungen einmal nicht erfüllt sein sollten.

Schließlich wollen wir ein Beispiel betrachten, das aufzeigt, dass die Wahrheitstabelle der Implikation im täglichen Leben durchaus eine Entsprechung findet. Wir betrachten die folgende Aussage.

$(*)$ Es wird ein Stein durch die Glasscheibe geworfen,
und daher zerbricht sie.

Diese Aussage steht, denke ich, außer Zweifel. Sie ist also wahr. Analysieren wir die Sache genauer. Wir haben die folgenden Aussagen:

p: Ein Stein wird durch die Glasscheibe geworfen.
q: Die Glasscheibe zerbricht.
$p \Rightarrow q$: Ein Stein wird durch die Glasscheibe geworfen, und daraus folgt, dass die Glasscheibe zerbricht.

Die Aussage $p \Rightarrow q$ ist eine etwas deutlichere Formulierung unserer Beispielaussage $(*)$ von oben, deren Wahrheit wir akzeptiert haben. Nun gehen wir alle Fälle unserer Wahrheitstabelle durch:

$p = 0, q = 0$: *Kein Stein wird durch die Glasscheibe geworfen. Die Glasscheibe zerbricht nicht.* Dies ist mit der Wirklichkeit durchaus verträglich, und widerspricht nicht im Mindesten unserer Beispielbehauptung $(*)$.
$p = 1, q = 1$: *Ein Stein wird durch die Glasscheibe geworfen. Die Glasscheibe zerbricht.* Auch das ist ein üblicher Vorgang (nicht das Werfen aber das darauf folgende Zerbrechen). Auch in diesem Fall entsteht kein Zweifel an $(*)$.
$p = 0, q = 1$: *Kein Stein wird durch die Glasscheibe geworfen. Die Glasscheibe zerbricht.* Dieser Fall bereitet üblicherweise Schwierigkeiten. Doch bei genauerer Betrachtung verblasst das Problem schnell. Vielleicht haben wir die Glasscheibe etwa mit einem Eisenträger durchstoßen. Die Scheibe ist kaputt ohne dass ein Stein geflogen wäre. Was der Scheibe auch immer passiert ist, genau können wir das aus dem Wahrheitsgehalt der Aussagen p und q nicht ableiten, die Tatsache, dass $(*)$ wahr ist, wird davon nicht berührt.
$p = 1, q = 0$: *Ein Stein wird durch die Glasscheibe geworfen. Die Glasscheibe zerbricht nicht.* Für einen solchen Fall fände ich keine Erklärung — Magie vielleicht? In der wirklichen Welt tendieren Scheiben zu zerbrechen, wenn man Steine durchwirft. Sollte aber tatsächlich der Fall eintreten, dass ein Stein geworfen wird, er durch die Scheibe fliegt und dann die Scheibe noch ganz ist, dann haben wir ein Problem. In diesem einen Fall müssten wir unsere Überzeugung aufgeben, dass $(*)$ gilt. Die Aussage $(*)$ wäre also tatsächlich *falsch*.

Wir haben also die Wahrheitswerte der Tabelle für $\Rightarrow$ in unserem Beispiel auf natürliche Weise wiedergefunden.

Alternativ dazu könnten wir versuchen herauszufinden, was es bedeutet, wenn wir die Ergebniswerte in den ersten beiden Zeilen anders setzen. Betrachten wir die anderen möglichen Fälle:

p	q	$p \wedge q$
0	0	0
0	1	0
1	0	0
1	1	1

p	q	$p \Leftrightarrow q$
0	0	1
0	1	0
1	0	0
1	1	1

p	q	q
0	0	0
0	1	1
1	0	0
1	1	1

.

Der erste Fall ist die Und-Verknüpfung der Aussagen p und q. Wir hätten also nur dann eine gültige Folgerung, wenn p und q beide wahr sind. Der Satz: „Das Quadrat einer geraden Zahl ist gerade.“ wäre also nicht wahr sondern hätte keinen zuordenbaren Wahrheitswert — das wäre zumindest unpraktisch.

Der zweite Fall ist die Äquivalenz. Auch das ist ein wenig zu restriktiv. In diesem Fall wäre der Satz „Sind zwei Zahlen gleich, dann sind auch ihre Quadrate gleich.“ nicht wahr, denn $2 \neq -2$ aber $2^2 = (-2)^2$.

Im letzten Fall stimmen die Wahrheitswerte des Theorems mit denen von q, also denen des Resultates überein, und der Wahrheitsgehalt der Voraussetzung wird gar nicht betrachtet. Das ist ebenfalls unpraktisch; auch in diesem Fall wäre der Satz über die Quadrate gerader Zahlen nicht wahr.

Wir sehen also, dass vieles dafür spricht, die Implikation so und nicht anders zu definieren. Alle die jetzt noch nicht überzeugt sind, seien dazu angehalten, die Tatsache einfach zu akzeptieren und sich daran zu gewöhnen.

Nachdem Schlussfolgerungen *das* Instrument der Mathematik sind, kommen sie in mathematischen Texten ausgesprochen oft vor. Deshalb haben sich auch eine Reihe von Standardformulierungen ausgebildet, die an Stelle der Formulierung „daraus folgt“ angewendet werden können.

- **also**; **auf Grund** von; das **bedeutet**, dass; unter **Berücksichtigung** von; **daher**; **damit**; es **ergibt** sich; daraus **erhalten** wir; dies hat zur **Folge**; man kann **folgern**; wir **folgern**; **folglich**; genauer **gesagt**; dies **impliziert**; **insbesondere**; dies hat zur **Konsequenz**; **mithin**; dies lässt sich **schreiben** als; wir **sehen**; **somit**; ein Spezialfall hiervon ist; nach **Umformung** ergibt sich; mit anderen **Worten**; es **zeigt** sich, dass,...

Es haben zwar nicht alle diese Formulierungen dieselbe Bedeutung, doch wenn Sie ein bisschen überlegen, wird es Ihnen nicht schwer fallen, vielleicht mit ein wenig Erfahrung, die feinen Unterschiede herauszuarbeiten.

Gut ist auch, wenn Sie Ihre Leserin oder Ihren Hörer darauf hinweisen, warum eine Folgerung richtig ist.

- nach **Annahme**; **auf Grund** von Satz 4.29; unter **Berücksichtigung** der Theorie der...; **da** V endlich dimensional ist; aus der **Definition** ergibt sich; **per definitionem** ist; nach **Voraussetzung**; **wegen** Lemma 12.2; **weil** f stetig ist...

Zuletzt können Sie noch den Aufwand verdeutlichen, der benötigt wird, um ein Resultat nach zu vollziehen.

- durch **einfaches Ausrechnen**; durch **genaues Hinsehen**; wie man **leicht sieht**; **offenbar**; **offensichtlich**; durch **technische und uninteressante Abschätzungen**; durch **triviale und langweilige Rechnung**; **trivialerweise**; durch **mühsame Umformungen**; durch **Überprüfen der Wahrheitstabellen**;...

Verjuxen Sie nicht den Vertrauensvorschuss Ihrer LeserIn durch falsche Angaben über den Aufwand. Behaupten Sie grundsätzlich nicht, dass etwas *leicht* einzusehen ist, wenn Sie mehr als 15 Minuten gebraucht haben, um es selbst einzusehen.

Zum Gebrauch des Wortes **trivial** ist zu sagen, dass die wenigsten Schritte in der Mathematik tatsächlich trivial sind. Trivial ist ein Beweisschritt *nur* dann, wenn er unmittelbar folgt (etwa direkt durch Anwendung einer Definition). Steckt ein, wenn auch noch so leicht einzusehender Beweis hinter dem Schritt, so ist er schon nicht mehr trivial.

Übrigens existiert noch eine zweite, technische Bedeutung des Wortes trivial in der Mathematik, nämlich als Adjektiv wie in

> Die **trivialen** Teiler einer natürlichen Zahl n sind 1 und n.

oder

> Ein homogenes lineares Gleichungssystem hat immer zumindest eine Lösung, nämlich die **triviale**.

Hier bedeutet trivial ein oder mehrere ausgezeichnete Objekte, die nach Definition immer existieren aber meist uninteressant sind.

Wollen wir einen Satz beweisen, so müssen wir sicher stellen, dass seine Aussage wahr ist. Die Wahrheitstabelle gibt uns dazu zwei Möglichkeiten (vgl. auch Abschnitt 2.1).

(1) Wir können annehmen, dass die Voraussetzungen (dies sind selbst Aussagen) gelten, dass also p wahr ist, und zeigen, dass dann das Resultat (die Aussage q) ebenfalls wahr ist. Beweise dieser Art heißen **direkte Beweise**.

(2) Alternativ können wir annehmen, dass das Resultat (q) *falsch* ist und dann daraus folgern, dass die Voraussetzungen (die Aussage p) ebenfalls falsch sind bzw. einen Widerspruch zu diesen herleiten (vgl. Abschnitt 2.1). Beweise dieser Art nennt man **indirekte Beweise**. Dieses Beweisprinzip funktioniert, da die Aussage des Satzes bei falschem q nur dann wahr ist, wenn auch p falsch ist. Ist jedoch q wahr, so kann p beliebig sein.

Nachdem schon einige direkte Beweise (z.B. Induktionsbeweise) vorgekommen sind, betrachten wir hier nun ein weiteres — man könnte sagen „klassisches“ — Beispiel eines indirekten Beweises und zeigen die Irrationalität von $\sqrt{2}$.

Theorem 3.2.4. *Die Zahl $\sqrt{2}$ ist irrational.*

Beweis. Die Aussage des Satzes als Implikation aufgeschrieben lautet:

Ist q eine rationale Zahl, so gilt $q \neq \sqrt{2}$.

Wir führen einen indirekten Beweis. Davor schreiben wir noch einmal alle Voraussetzungen an, die wir verwenden wollen.

Für jede rationale Zahl q gibt es teilerfremde ganze Zahlen m und $n \neq 0$ mit $q = \frac{m}{n}$, und jede Bruchzahl ist rational. Daher ist q in $\mathbb{Q}$ gleichbedeutend damit, dass q als Bruch zweier teilerfremder ganzer Zahlen mit nicht verschwindendem Nenner darstellbar ist.

Wir können die Aussage des Satzes also auch folgendermaßen formulieren: Sind m und $n \neq 0$ zwei teilerfremde ganze Zahlen, so gilt $\frac{m}{n} \neq \sqrt{2}$.

Für den indirekten Beweis müssen wir das Resultat verneinen, also nehmen wir an, dass $\frac{m}{n} = \sqrt{2}$. Daraus reicht es zu folgern, dass es solche m und n nicht gibt.

Beweisen wir dies. Sei

$$\begin{aligned} \frac{m}{n} &= \sqrt{2} \\ \frac{m^2}{n^2} &= 2 \\ m^2 &= 2n^2. \end{aligned}$$

Dies bedeutet aber, dass m^2 gerade ist, und da das Quadrat einer ungeraden Zahl ungerade ist, muss folglich m selbst gerade sein. Damit können wir m aber schreiben als $m = 2k$ und einsetzen,

$$\begin{aligned} (2k)^2 &= 2n^2 \\ 4k^2 &= 2n^2 \\ 2k^2 &= n^2. \end{aligned}$$

Wir sehen, dass auch n^2 und damit n gerade ist. Nachdem wir jetzt bewiesen haben, dass n und m beide gerade sind, können sie nicht länger teilerfremd sein (sie sind als gerade Zahlen beide durch 2 teilbar). Dies widerlegt unsere Voraussetzung, und der indirekte Beweis ist geglückt. $\square$

Im Zusammenhang mit Implikationen tauchen in mathematischen Texten oft die die Wörter **notwendig** und **hinreichend** auf. Wenn A und B Aussagen sind und $A \Rightarrow B$ gilt, so heißt A *hinreichend* für B, und B heißt *notwendig* für A. Lernen Sie das auswendig und versuchen Sie nicht die Bedeutung zu hinterfragen. Beispiele sind:

- Notwendig dafür, dass eine Zahl $n > 2$ eine Primzahl ist, ist, dass sie ungerade ist.
- Hinreichend für die Stetigkeit einer Funktion ist ihre Differenzierbarkeit.

Nun zu **dann, wenn** und **nur dann, wenn**:

- „A gilt dann, wenn B gilt“ bedeutet: $A \Leftarrow B$.
- „A gilt nur dann, wenn B gilt“ heißt hingegen $A \Rightarrow B$.

Um ein Beispiel für letztere Formulierung zu geben, betrachten wir die Aussagen: A sei „Ein neuer Papst wird gewählt.“, B sei „Der alte Papst ist gestorben.“. Die Formulierung „Ein neuer Papst wird **nur dann** gewählt, **wenn** der alte gestorben ist“ entspricht dann der Folgerung $A \Rightarrow B$. Wenn wir den Satz umdrehen, so ergibt das die Aussage „Wenn ein neuer Papst gewählt wird, dann ist der alte jedenfalls gestorben.“ Seien Sie in jedem Fall vorsichtig, wenn Sie die Formulierungen mit dann und wenn benutzen.

3.2.2.2. Die Äquivalenz ($\Leftrightarrow$). Eine zweite Klasse von Sätzen der Mathematik hat die logische Äquivalenz (die Operation $\Leftrightarrow$) als Grundlage. Eine leichte Rechnung mit den Wahrheitstabellen ergibt $a \Leftrightarrow b = (a \Rightarrow b) \wedge (b \Rightarrow a)$.

Die typische Aussage eines Äquivalenzsatzes sieht so aus

Theorem 3.2.5. *Resultat 1 gilt* ***genau dann, wenn*** *Resultat 2 gilt.*

Auch an Stelle der Standardaussage „das gilt genau dann, wenn“ haben sich einige andere Formulierungen eingebürgert.

- das ist **äquivalent** zu; dies ist **gleichbedeutend** mit; dies ist **gleichwertig** mit; die beiden Aussagen **gehen auseinander hervor**; dies ist **notwendig und hinreichend** für; **dann und nur dann**...

Die übrigen Hinweise, wie Aufwandsangabe und Erwähnung der Begründung, die wir bei den Implikationen schon besprochen haben, gelten natürlich auch für Äquivalenzen.

Äquivalenzen kommen in der Mathematik sehr häufig vor. Die Äquivalenz zweier Aussagen A und B beweist man dabei so wie es von der obigen Formel suggeriert wird. Man weist die Gültigkeit von $A \Rightarrow B$ nach *und* auch die der *umgekehrten Richtung* $B \Rightarrow A$.
WICHTIG: Der Beweis einer Äquivalenz ist erst dann vollendet, wenn *beide* Implikationsrichtungen gezeigt sind. Um dies zu verdeutlichen betrachten wir die folgende Aussage.

Proposition 3.2.6 (Quadrate gerader Zahlen — die Zweite)**.** *Eine Zahl ist genau dann gerade, wenn ihr Quadrat gerade ist.*

Beweis. Umformuliert bedeutet die Aussage für eine beliebige ganze Zahl n

$$n \text{ gerade } \Leftrightarrow n^2 \text{ gerade }.$$

Wir müssen also beide Implikationen beweisen und beginnen mit der „Hinrichtung“.
$\Rightarrow$: Diese Implikation ist aber genau die Aussage von Proposition 2.1.1, sodass wir nichts mehr zu beweisen haben.

Es bleibt uns die „Rückrichtung“ zu zeigen.
$\Leftarrow$: Genauer ist zu zeigen n^2 gerade $\Rightarrow n$ gerade.
Das beweisen wir indirekt. Sei also n ungerade, d.h. $n = 2k + 1$ für eine ganze Zahl k. Dann ist $n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ und n^2 somit ungerade.

Nachdem wir beide Implikationen bewiesen haben, gilt die im Satz behauptete Äquivalenz. □

Ganz nebenbei schließt die „Rückrichtung“ der Proposition mittels der Äquivalenz $(\neg p \Rightarrow \neg q) \Leftrightarrow (q \Rightarrow p)$ (Diese ist ganz einfach mittels Wahrheitstabellen nach zu weisen.) auch die (kleine) Lücke im Beweis von Theorem 3.2.4 (Ganz ehrlich: Haben Sie diese bemerkt?).

Hat man mehr als zwei Aussagen, von denen man die Äquivalenz zeigen möchte, etwa A, B und C, so kann man einen sogenannten **Zirkelschluss** $A \Rightarrow B$, $B \Rightarrow C$, $C \Rightarrow A$ durchführen, um die Äquivalenz der Aussagen sicher zu stellen. Vorsicht: Solche Zirkelschlüsse beweisen nur die Äquivalenz von Aussagen. Über deren Wahrheitswert wird durch solch einen Beweis nichts bekannt.

Interessant ist noch die Verneinung einer Äquivalenz. Mit Hilfe der Wahrheitstabelle sehen wir nämlich $\neg(p \Leftrightarrow q) = p \veebar q$, also „$p$ ist nicht äquivalent zu q“ ist gleichbedeutend mit „entweder p oder q“. Umgekehrt ist natürlich die Verneinung einer Entweder-Oder-Aussage eine Äquivalenz.

3.2.3. Quantoren. Viele mathematische Aussagen gelten für bestimmte oder auch alle Objekte einer „Gattung“; diesen Formulierung wollen wir uns nun zuwenden.

3.2.3.1. Der Allquantor ($\forall$). Ein Großteil der mathematischen Theorien handelt von Strukturen und Regeln. Ein Beispiel für Regeln sind Rechengesetze, die etwa **für alle** Objekte eine bestimmten Menge gelten. In diesem Fall verwenden wir das Zeichen $\forall$, den **Allquantor**.

Die Formulierung „$\forall x \in M$:“ bedeutet „Für alle x in M gilt…“.

Andere Formulierung für dieselbe Zeichenfolge sind etwa

- Für jedes x in M gilt…
- Sei $m \in M$ beliebig. Dann gilt…
- Für eine beliebiges Element von M gilt…
- Ist $m \in M$, dann gilt…
- Jedes Element aus M erfüllt…
- Die Elemente von M erfüllen…
- $\bigwedge m \in M$.

Bezieht sich ein $\forall$ auf mehrere Variable auf einmal, so verwendet man auch oft „je zwei“, „je drei“, …

- Durch je zwei verschiedene Punkte P und Q geht genau eine Gerade.

bedeutet „Für jeden Punkt P und jeden Punkt $Q \neq P$ gibt es genau eine…“

Der Unterschied zwischen „alle“ und „jedes“ besteht meist darin, dass „für alle“ auf die Gesamtheit aller Objekte abzielt, während „jedes“ ein beliebig herausgegriffenes Objekt meint:

- Alle bijektiven Funktionen sind invertierbar.
- Für jede bijektive Funktion f existiert die Umkehrfunktion, welche wir mit f^{-1} bezeichnen.

Merke: Um eine Allaussage zu widerlegen genügt die Angabe *eines* Gegenbeispieles.
Behauptung: Alle ungeraden Zahlen sind Primzahlen. Dies ist natürlich falsch, denn die Zahl $9 = 3 \cdot 3$ ist eine ungerade Zahl, die keine Primzahl ist.

3.2.3.2. Existenz ($\exists$ und $\exists!$). Oftmals wird eine mathematische Aussage nicht über alle Elemente einer Menge getroffen, sondern es wird nur die **Existenz** eines bestimmten Objektes behauptet.

Für ein homogenes lineares Gleichungssystem existiert eine Lösung.

Die Formulierung in Zeichen mit Hilfe des **Existenzquantors** ist „$\exists x \in M$:“ und in Worten: „Es existiert ein x in M mit...“. Diese Aussage bedeutet, dass es **mindestens ein** Element in M gibt mit...

Möchte man in Zeichen ausdrücken dass es **genau ein** Element in M gibt mit..., so schreibt man „$\exists! x \in M$:“.

Auch für die Existenzaussage gibt es viele Formulierungen.

- Es gibt ein $x \in M$ mit...
- Jede monotone beschränkte Folge reeller Zahlen hat einen Häufungspunkt (d.h. es existiert ein Häufungspunkt)
- Für ein geeignetes x ist $\log x \leq x$. Das bedeutet nichts anderes als, dass solch ein x existiert.
- Im allgemeinen gilt nicht, dass $x^2 + x + 41$ eine Primzahl ist. (Das wiederum heißt, dass ein x existiert, sodass $x^2 + x + 41$ keine Primzahl ist.)
- $\bigvee x \in M$:

Merke: Die Verneinung einer Existenzaussage ist eine Allaussage und umgekehrt.

- Die Verneinung von „Alle Kinder hassen die Schule“ ist „Es gibt ein Kind, das die Schule nicht hasst“.
- Die Verneinung von „Es gibt einen klugen Assistenten“ ist „Alle Assistenten sind dumm.“

In Zeichen ausgedrückt, gilt für die Verneinungen:

$$\neg(\forall x \in M : A(x)) \quad \text{entspricht} \quad \exists x \in M : \neg A(x),$$

wenn A eine Aussage über Elemente von M ist, etwa $A(x) = (x < 7)$. Für den Existenzquantor gilt analoges:

$$\neg(\exists x \in M : A(x)) \quad \text{entspricht} \quad \forall x \in M : \neg A(x).$$

ACHTUNG: Die Verneinung einer Existiert-Genau-Ein-Aussage ist *keine* Allaussage! Man muss komplizierter formulieren. Die Verneinung von „Ich habe genau einen Bruder.“ ist am kürzesten formuliert „Ich habe nicht genau einen Bruder.“ Möchte man das „*nicht*“ zur Aussage befördern, dann muss man mit einer Fallunterscheidung operieren: „Ich habe keinen Bruder oder mehr als einen Bruder.“

3.2.3.3. Reihenfolge von Quantoren ($\forall\exists$ oder $\exists\forall$?). Seien Sie vorsichtig, wenn mehr als ein Quantor $\forall$ oder $\exists$ in einem Satz vorkommt. Dabei kommt es nämlich wesentlich auf die Reihenfolge an.

Beispiel 3.2.7. *Sei M die Menge aller Männer und F die Menge aller Frauen. Die Aussage $h(m, f)$ sei „m ist verliebt in f“. Unter diesen Voraussetzungen machen Sie sich die Bedeutung der beiden Aussagen klar. Danach werden Sie immer auf die Reihenfolge der Quantoren achten.*

(1) $\forall m \in M : \exists f \in F : h(m, f)$.
(2) $\exists f \in F : \forall m \in M : h(m, f)$.

Mitunter ist es aus der Formulierung nur schwer zu erkennen, dass ein $\exists\forall$ oder ein $\forall\exists$ versteckt ist. Dann ist es besonders wichtig, die Formulierung sehr lange zu prüfen und eventuell auch formalisiert noch einmal aufzuschreiben.

- „Der Wert von $y = f(x)$ ist unabhängig von der Wahl von x“ ist gleichbedeutend mit $\exists y : \forall x : f(x) = y$.

KAPITEL 4

Mengenlehre

In diesem Kapitel führen wir nun unsere erste mathematische Struktur, die der *Menge* ein. An diesem Punkt stoßen wir also zum ersten Mal auf das in der Einleitung erwähnte Grundprinzip der Mathematik: Definition und Untersuchung von *Strukturen*.

Ein Großteil der mathematischen Theorien ist darauf aufgebaut, Objekte mit bestimmten Eigenschaften und deren Beziehungen untereinander zu untersuchen. Strukturen können neben einander existieren oder aber auf einander aufbauen, d.h. sie sind Spezialisierungen oder Kombinationen von bereits bestehenden Strukturen.

Die Basisstruktur für die meisten Gebiete der Mathematik ist diejenige der *Menge*, die wir auf naive Weise im Abschnitt 4.1 einführen. Nach einer Darstellung der elementaren Mengenoperationen untersuchen wir *Relationen* als Struktur, die auf dem Begriff der Menge aufbaut und *Abbildungen* zwischen Mengen. Zum Schluss des Kapitels geben wir eine kurze Einführung in die *axiomatische Mengenlehre* (Aufbaustoff).

4.1. Naive Mengenlehre

Bevor wir in Abschnitt 4.5 kurz einen logisch exakten Zugang zur Mengenlehre skizzieren, wollen wir uns hier, aus Gründen der Motivation und des besseren Verständnisses, auf den Zugang von Georg Cantor (1845–1918) zurückziehen, den dieser gegen Ende des 19. Jahrhunderts erstmals formuliert hat:

> Unter einer **Menge** verstehen wir jede Zusammenfassung S von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die **Elemente** von S genannt werden) zu einem Ganzen.

Vorstellen kann man sich eine Menge gewissermaßen als einen Sack. Die Elemente sind die Gegenstände, die sich in dem Sack befinden. Natürlich können Mengen andere Mengen enthalten so, wie sich auch weitere Säcke innerhalb eines Sackes befinden können.

Beispiel 4.1.1 (Mengen)**.** *Bilden kann man etwa die folgenden Mengen.*

- *Die Menge aller Studierenden im Hörsaal.*
- *Die Menge der natürlichen Zahlen.*
- *Die Menge der Lösungen einer Ungleichung.*
- *Die* ***leere Menge*** *(„ein leerer Sack“).*

Der Gebrauch des bestimmten Artikels ist in der Mathematik äußerst eingeschränkt. Es gibt eine feste Regel, die nie gebrochen werden darf.

> **Ein bestimmter Artikel darf nur dann verwendet werden, wenn es klar ist, dass das fragliche Objekt eindeutig bestimmt ist.**

So ist es unzulässig zu formulieren

- … ~~der~~ Teiler von 6 (denn es gibt 1, 2, 3 und 6).
- … ~~die~~ Matrix, die einer lineare Abbildung f entspricht… (denn sie ist nicht eindeutig).
- … ~~die~~ Basis des $\mathbb{R}^3$.

Richtig wäre es dagegen zu sagen:

- Sei n **die** kleinste natürliche Zahl, die ...
- ... **die** leere Menge.
- ... **die** Menge der natürlichen/ganzen/rationalen/reellen Zahlen.

Bevor wir den Begriff „Menge“ weiter studieren, machen wir einen kurzen historischen Exkurs, denn die Geschichte der Mengenlehre unterscheidet sich grundlegend von der fast aller anderen Gebiete der Mathematik, wie etwa in [**O'Connor, Robertson 1996**] dargestellt.

Üblicherweise geht die mathematische Entwicklung verschlungene Wege. Theorien werden über Jahrhunderte hinweg von mitunter konkurrierenden Schulen von Mathematikern gepflegt und weiterentwickelt. Plötzlich ist die Theorie an einem Punkt angelangt, an dem oftmals mehrere Mathematiker gleichzeitig einen Geistesblitz haben und ein bedeutendes Resultat entdeckt wird. Die Mengenlehre steht dem vollständig entgegen. Bis auf wenige zusätzliche Arbeiten ist sie die Entwicklung eines einzigen Mannes, Georg Cantor.

Abbildung 4.1. Georg Cantor (1845–1918)

Die „Unendlichkeit“ hat die Philosophie (und damit die Mathematik) jedenfalls seit Zeno von Elea (ca. 490-425 v.Chr), also seit etwa 450 v.Chr. beschäftigt. Später haben sich bedeutende Philosophen, unter anderen Aristoteles (384–322 v.Chr.), René Descartes (1596–1650), Georeg Berkeley (1685–1753), Gottfried W. Leibniz (1646–1716), aber auch Albert von Sachsen (1316-1390), der die Volumina unendlicher Mengen (Strahlen, Raum,...) verglichen hat, mit diesem Problem auseinandergesetzt.

Anders die Idee der Menge. Diese begann erst Mitte des 19. Jahrhunderts langsam in die Köpfe der Mathematiker Einzug zu halten. So hat etwa der tschechische Mathematiker Bernard Bolzano (1781–1848) ein Jahr vor seinem Tod folgendermaßen formuliert:

> ... eine Verkörperung der Idee oder des Konzeptes, das wir erhalten, wenn wir die Anordnung seiner Teile für gleichgültig erachten.

Der wirkliche Durchbruch der Mengenlehre kam aber erst mit Georg Cantor, der nach einem Besuch bei Richard Dedekind (1831–1916) und darauf folgender Korrespondenz im Jahr 1874 eine wissenschaftliche Arbeit im Crelle-Journal publizierte, in der er Mengen und das Konzept verschiedener Klassen von Unendlichkeit einführte.

Im Jahr 1878 versuchte er eine weitere Publikation im Crelle-Journal, doch stieß er auf heftigen Widerstand der damals Ton angebenden mathematischen Schule der *Konstruktivisten* mit ihrem führenden Kopf Leopold Kronecker (1823–1891), die keine mathematischen Sachverhalte akzeptieren wollten, die sich nicht in endlich vielen Schritten aus den natürlichen Zahlen konstruieren ließen. Erst nach massiver Intervention von Karl Weierstrass (1815–1897) wurde die Arbeit schließlich akzeptiert. Das war der Beginn eines langen Kampfes innerhalb der Mathematik um ihre philosophischen und später auch logischen Grundlagen, der z.B. durch folgende Zitate schön belegt werden kann:

> „Aus dem Paradies [die Mengenlehre], das Cantor uns geschaffen hat, soll uns niemand mehr vertreiben können.“ David Hilbert (1862–1943)

> „Spätere Generationen werden die Mengenlehre als Krankheit ansehen, die man überwunden hat.“ Jules Henri Poincare (1854–1912)

Dieser Kampf wurde nicht nur auf mathematischer sondern auch auf menschlicher Ebene ausgetragen, so blockierten etwa Kronecker und Hermann Schwarz (1843–1921) Cantors Stellenbewerbungen.

Von 1879 bis 1884 veröffentlichte Cantor in den *Mathematischen Annalen* eine sechsteilige Abhandlung über die Mengenlehre, die zu großen Kontroversen in der mathematischen Welt führte. Einige Mathematiker hielten sich an Kronecker, doch andere folgten Cantors Weg. So führte etwa Giuseppe Peano (1858–1932), nach seinem berühmten Satz über Differentialgleichungen (1886) und der ersten Definition eines Vektorraumes (1888) und vor seinen berühmten Peano-Kurven (1890), neben der Axiomatisierung der natürlichen Zahlen 1889 auch das Zeichen $\in$ in die Mengenlehre ein.

Im Jahr 1897 fand Cesare Burali-Forti (1861-1931) das erste Paradoxon in der Mengenlehre, obwohl es durch eine fehlerhaft verstandene Definition des Begriffes „wohlgeordnete Menge“ entwertet, wenn auch nicht ausgelöscht wurde. Interessanter Weise ereignete sich der erste persönliche Erfolg für Cantor im selben Jahr auf einem Mathematiker-Kongress in Zürich, wo sein Werk zum ersten Mal positiv aufgenommen, ja von manchen in höchsten Tönen gepriesen wurde.

Nachdem Cantor selbst 1899 ein weiteres Paradoxon gefunden hatte, entdeckte schließlich Bertrand Russell (1872-1970) im Jahre 1902 das ultimative Paradoxon (heute Russellsche Antinomie), das insbesondere wegen seiner Einfachheit die neuen Grundlagen der Mathematik in ihren Grundfesten erschütterte. Russel betrachtete die Menge R aller Mengen, die sich nicht selbst enthalten; nennen wir solche Mengen „vernünftig“. Die Frage ob R sich selbst enthält, führt nun zu einem Widerspruch: Nehmen wir nämlich an, dass R ein Element von R ist, so ist R keine „vernünftige“ Menge und kann daher nicht Element von R sein. Ist andererseits R nicht Element von R, so ist R eine „vernünftige“ Menge und es muss gelten, dass R Element von R ist.

Zu diesem Zeitpunkt hatte sich die Mengenlehre allerdings schon durchgesetzt. Sowohl die Analysis baute darauf auf als auch Teile der Algebra. Die Maßtheorie und das mengentheoretische Integral waren 1901 bzw. 1902 von Henri Lebesgue (1875–1941) erfunden worden. Daher wurde die Mengenlehre nicht gleich wieder verworfen, sondern es begann eine fieberhafte Suche nach einer „Rettung“ der Mengenlehre ohne ihre wichtigsten Eigenschaften aufgeben zu müssen.

Russell selbst versuchte, sein Paradoxon aus der Mathematik „wegzudefinieren“. In seinem sehr einflussreichen Werk *Principia Mathematica*, das er gemeinsam mit Alfred Whitehead (1861–1947) in den Jahren 1910–1913 veröffentlichte, stellte er eine Lösung mit Hilfe der *Theory of types* vor, doch diese wurde von den meisten nicht als befriedigend erachtet.

Der erste, der eine Lösung für das Paradoxien-Problem fand war Ernst Zermelo (1871–1953), der im Jahr 1908 das erste befriedigende Axiomensystem für die Mengenlehre publizierte, das im Jahr 1922 von Adolf Fraenkel (1891–1965) nochmals verbessert wurde, und das heute aus zehn Axiomen bestehend für viele die Grundlage der Mathematik darstellt (siehe Abschnitt 4.5). Auch andere berühmte Mathematiker wie Kurt Gödel (1906–1978), Paul Bernays (1888–1977) und John von Neumann (1903–1957) axiomatisierten die Mengenlehre auf unterschiedliche Weisen, und welches der Axiomensysteme die Grundlage bilden soll, wird in der heutigen Zeit von den meisten Mathematikern als „reine Geschmackssache" angesehen.

Um die Jahrhundertwende strebten noch viele Mathematiker allen voran David Hilbert und Gottlob Frege (1848–1925) danach die Mathematik (und auch die Physik) vollständig auf die formale Logik zu reduzieren. Hilbert — wahrscheinlich der einflussreichste Mathematiker seiner Zeit — erwähnte dies noch 1900 in seiner berühmten Rede auf dem Internationalen Mathematiker-Kongresses in Paris. Für dieses Ziel war eine möglichst umfassende und widerspruchsfreie Axiomatisierung der Mengenlehre wesentlich. Nach Gödels ω-Unvollständigkeitssatz im Jahr 1931, der die Grenzen jedes axiomatischen Systems aufzeigte, wurden alle diese Versuche zerschlagen und weitere Ansätze bereits im Keim erstickt.

Geblieben von dieser Entwicklung ist das heutige Bestreben der Mathematiker nach exaktem und logischem Vorgehen beim Entwickeln und Beweisen von mathematischen Theorien, beim Aufbau des mathematischen Theoriegebäudes. Jetzt ist es wichtig, Grundlagen zu *haben*, die die mathematisch exakte Behandlung der Theorie erlauben. Nachdem alle heute gängigen Axiomensysteme das bieten, ist die genaue Auswahl eines bestimmten Systems den meisten Mathematikern nicht mehr so wichtig.

Nach diesem historischen Überblick wollen wir die Mengenlehre genauer kennenlernen und zunächst wie Cantor naiv beginnen.

Wollen wir über Mengen sprechen, so müssen wir zuerst erklären, wie wir sie beschreiben können. Grundsätzlich stehen uns zwei Methoden zur Verfügung.

(1) **Aufzählen:** Wir können **alle** Elemente einer *endlichen* Menge angeben, um die Menge zu definieren. So könnten wir etwa durch

$$M := \{0, 2, 5, 9\}$$

die Menge M einführen. Sie enthält als Elemente die vier Zahlen 0, 2, 5 und 9.

Diese in der Mengenlehre fundamentale Beziehung zwischen den Mengen und ihren Elementen wird durch das Symbol $\in$ ausgedrückt. Hier also: $0 \in M$, $2 \in M$, …sprich „0 in M", oder „2 Element M",…

Auf die Reihenfolge kommt es bei der Aufzählung übrigens nicht an: $\{0, 5, 2, 9\}$ ist ebenso wie $\{9, 0, 2, 5\}$ die gleiche Menge wie M.

Das Zeichen := (vgl. auch die Box auf Seite 12) bedeutet, dass wir gerade etwas **definieren**, in diesem Fall geben wir der Menge bestehend aus den Zahlen 0, 2, 5 und 9 den Namen M. Merke: Der Doppelpunkt im Zeichen := (oder =:) steht immer auf der Seite des Gleichheitszeichens, auf der der zu definierende Begriff steht.

Grundsätzlich dienen Definitionen dazu, neue *Abkürzungen* einzuführen. Der definierte Begriff kann jederzeit durch die definierende Beschreibung ersetzt werden, und manchmal muss das auch geschehen, speziell in Beweisen.

Den Sinn von Definitionen rein darauf zu reduzieren, dass sie Abkürzungen einführen, heißt aber, die Bedeutung von Definitionen stark unter zu bewerten. Eine Definition ist ein schöpferischer Akt! Es ist einer der bedeutendsten Schritte in der Entwicklung einer mathematischen Theorie, die wichtigen Objekte zu erkennen und ihnen Namen zu geben.

Dadurch rücken sie ins Zentrum des Interesses, es werden neue Begriffe geschaffen, und man kann beginnen, sich mit diesen neuen Begriffen auseinander zu setzen.

In diesem Zusammenhang ist noch einmal wichtig heraus zu streichen, dass eine Definition niemals *falsch* sein kann (abgesehen von Prüfungen, wenn bereits bestehende Definitionen falsch rezitiert werden), sie kann allerdings *sinnlos* oder *wenig hilfreich* sein.

Scheuen Sie nicht davor zurück, bei der Lösung Ihrer Aufgaben, wichtigen Objekten eigene Namen zu geben z.B. „starke“ Matrizen, „coole“ Elemente,. . . .

(2) **Beschreiben:** Gemäß einer Idee von Cantor können wir eine Menge auch dadurch definieren, dass wir *Eigenschaften ihrer Elemente* angeben. Dies läßt sich auch auf *unendliche* Mengen anwenden. Die Menge P aller Primzahlen ließe sich etwa definieren durch

$$P := \{p \in \mathbb{N} \mid p > 1 \wedge \forall m \in \mathbb{N} : (m|p \Rightarrow (m = 1 \vee m = p))\}.$$

Genauer bedeutet das, dass man P als die Menge all jener Elemente p von $\mathbb{N}$ definiert, die größer 1 sind und folgende Eigenschaft besitzen: Jedes Element m von $\mathbb{N}$, das p teilt, ist entweder 1 oder p selbst. Das bedeutet aber wiederum, dass p größer 1 ist und nur die trivialen Teiler 1 und p hat — genau unsere Definition 2.1.2.

Oft wird statt des senkrechten Strichs ($|$) auch ein Doppelpunkt verwendet, also $P := \{p \in \mathbb{N} :\ \dots\}$.

Viele Definitionen verwenden nicht nur verbale Ausdrücke sondern auch mathematische Symbolik, wie obige Definition von P. Man muss aber nicht rein symbolisch formulieren. Eine ähnlich gute Definition von P wäre

$$P := \{p \in \mathbb{N} \mid p > 1 \text{ und } p \text{ besitzt nur die Teiler } 1 \text{ und } p\}.$$

Symbole im Text erhöhen zwar oft dessen Präzision, doch im selben Maße verringern sie seine Lesbarkeit. Geht man zu sorglos mit ihnen um, so kann der Text sogar mehrdeutig werden. Beherzigt man allerdings einige Regeln, so verbessert das die Lage sofort.

- **Ein Satz sollte nicht mit einem Symbol beginnen**. Man formuliert den Satz

  $\cancel{\mathbb{R}}$ bezeichnet die Menge der reellen Zahlen.

  besser um

  Die Menge der reellen Zahlen bezeichnen wir mit $\mathbb{R}$.
- **Axiom von Siegel** (nach dem Mathematiker Carl Siegel (1896–1981)): **Zwei mathematische Symbole** (die sich nicht zu einem größeren Symbolkomplex ergänzen) **müssen stets durch mindestens ein Wort getrennt sein!**

  Eine 10-elementige Menge hat genau $\cancel{45\ 2}$-elementige Teilmengen.

  könnte bei engerem Druck fehlinterpretiert werden. Besser wäre etwa die Formulierung

  Die Anzahl der 2-elementigen Teilmengen einer 10-elementigen Menge ist 45.
- **Verwenden Sie niemals mathematische Symbole als Abkürzungen für Worte im Text.**

  Sei V ein Vektorraum $\cancel{+}$ endlich dimensional.

Verwenden Sie die Symbole sorgfältig und behalten Sie ihre mathematische Bedeutung stets im Auge. Konzentrieren Sie die Symbolik nicht zu sehr. Eine gute Mischung aus Symbolik und Text garantiert einerseits die Präzision und erhöht andererseits die Lesbarkeit.

Unverzichtbar ist, dass Sie stets in der Lage sind, zwischen verbaler und formaler Beschreibung hin und her zu schalten. Es ist wichtig, schon zu Beginn die Fähigkeit zu trainieren, die eine Beschreibung in die andere zu verwandeln.

Beispiel 4.1.2 (Elementbeziehung)**.**

- *Es gilt* $2 \in \{2, 4, 7, 9\}$*,*
- *weiters haben wir* $42 \in \mathbb{N}$*.*
- *Steht die Menge links vom Element, so dreht man das Zeichen* $\in$ *einfach um:* $\mathbb{R} \ni \pi$*.*
- *Wollen wir ausdrücken, dass ein Objekt* nicht *Element einer bestimmten Menge ist, so streichen wir das Zeichen* $\in$ *einfach durch, wie in* $\frac{1}{2} \notin \mathbb{N}$*.*

Definition 4.1.3 (Gleicheit von Mengen)**.** *Zwei Mengen gelten genau dann als* gleich*, wenn sie dieselben Elemente haben; in Symbolen notiert*

$$A = B \quad \textit{genau dann wenn} \quad \forall x : (x \in A \Leftrightarrow x \in B).$$

Definition 4.1.4 (Leere Menge)**.** *Die* leere Menge $\emptyset$ *ist definiert als die Menge, die kein Element enthält. Formal kann das z.B. so ausgedrückt werden*

$$\emptyset := \{x \mid x \neq x\}.$$

In der Mathematik ist das Symbol $\emptyset$ *üblich, auch wenn mitunter* $\{\}$ *als Bezeichnung für die leere Menge verwendet wird.*

WICHTIG: Beachten Sie, dass ein Element in einer Menge enthalten ist, oder eben nicht. Es steht immer eindeutig fest, welche Elemente zu einer Menge gehören.

Ein und dasselbe Element kann nicht mehrfach in einer Menge auftreten. Eine Menge ist eine Ansammlung *verschiedener* Objekte! Allerdings ist es nicht verboten, einige Elemente mehrfach anzuführen. $\{1, 2\}$ ist die gleiche Menge wie $\{1, 1, 2\}$. Zugegeben, dieses Beispiel ist gekünstelt — der Sinn dieser Vereinbarung wird erst dann deutlich wenn man z.B. eine Menge $\{a, b, c\}$ untersucht, wobei a, b und c erst später festgesetzt oder näher bestimmt werden. Dann ist es sehr praktisch, $\{a, b, c\}$ schreiben zu können, selbst wenn sich später herausstellen sollte, dass $a = b$ gilt.

4.1.1. Teilmengen. Bevor wir untersuchen, wie wir Mengen mit einander verknüpfen können, betrachten wir das einfachste Konzept, das von *Teilmengen*.

Definition 4.1.5 (Teilmenge, Obermenge)**.** *Eine Menge* B *heißt* Teilmenge *der Menge* A*, wenn* B *nur Elemente von* A *enthält. In der Sprache der Logik formuliert, bedeutet das*

$$\forall x : x \in B \Rightarrow x \in A,$$

oder kürzer und etwas salopper

$$\forall x \in B : x \in A.$$

Ist B *Teilmenge von* A*, so schreiben wir*

$$B \subseteq A \quad \textit{oder} \quad A \supseteq B.$$

A *heißt dann* Obermenge *von* B*.*

Beispiel 4.1.6 (Teilmengen)**.** *Wir finden etwa:*

- *Die leere Menge ist Teilmenge* jeder *Menge.*
- *Jede Menge* M *ist ihre eigene Teilmenge. Die Menge* M *und* $\emptyset$ *heißen die* ***trivialen Teilmengen*** *von* M*.*
- *Alle Teilmengen, die ungleich der Menge selbst sind, nennt man auch* ***echte Teilmengen****. Möchte man betonen, dass* B *echte Teilmenge von* A *ist, so schreibt man meist*

$$B \subset A \quad \textit{oder expliziter} \quad B \subsetneqq A.$$

- *Alle Teilmengen von* $\{1, 2, 3\}$ *sind* $\emptyset$*,* $\{1\}$*,* $\{2\}$*,* $\{3\}$*,* $\{1, 2\}$*,* $\{1, 3\}$*,* $\{2, 3\}$ *und* $\{1, 2, 3\}$*.*

Leider wird in manchen mathematischen Texten das Symbol $\subset$ für Teilmenge verwendet und nicht für *echte* Teilmenge, wie wir das oben definiert haben. Daher unser Tipp: verwenden Sie in eigenen Texten immer $\subseteq$ um Teilmengen und $\subsetneqq$ um echte Teilmengen zu kennzeichnen. Finden Sie in einem Text $\subset$, so ist es ratsam nach der Teilmengendefinition zu suchen, um festzustellen, ob $\subset$ für echte Teilmenge oder nur Teilmenge steht.

Die Teilmengenrelation entspricht, wie schon in der Definition explizit gemacht wurde, der logischen Implikation ($\Rightarrow$) der Elementbeziehung. Daraus läßt sich auch sofort ableiten, wie man Gleichheit von Mengen überprüfen kann.

Proposition 4.1.7 (Gleichheit von Mengen). *Zwei Mengen A und B sind genau dann gleich, wenn $A \subseteq B$ und $B \subseteq A$ gilt; formal*

$$A = B \Leftrightarrow A \subseteq B \ \wedge \ B \subseteq A.$$

Beweis. Dieser Satz behauptet eine Äquivalenz und wir müssen wiederum beide Implikationsrichtungen beweisen.
$\Rightarrow$: Zu zeigen ist, dass wenn $A = B$ gilt, auch die beiden Enthalten-Relationen $A \subseteq B$ und $B \subseteq A$ gelten. Dies ist aber trivial, da $A \subseteq A$ für jede Menge stimmt.
$\Leftarrow$: Wir müssen zeigen, dass aus beiden Enthalten-Relationen schon die Gleichheit folgt. Es gelte also $A \subseteq B$ und $B \subseteq A$. Wegen $A \subseteq B$ gilt $x \in A \Rightarrow x \in B$. Andererseits folgt aus $B \subseteq A$, dass $x \in B \Rightarrow x \in A$ gilt. Fassen wir die beiden Implikationen zusammen, erhalten wir für beliebiges x den Zusammenhang $x \in A \Leftrightarrow x \in B$. Das wiederum bedeutet laut Definition 4.1.3, dass $A = B$ gilt. $\square$

4.1.2. Mengenoperationen. Wenn man mehr als eine Menge betrachtet, so kann man aus diesen Mengen weitere Mengen erzeugen. Die folgenden Mengenoperationen werden dabei standardmäßig verwendet.

Definition 4.1.8 (Vereinigung).

(i) *Seien zwei Mengen A und B gegeben. Wir konstruieren eine neue Menge aus allen Elementen von A und B. Diese Menge heißt* Vereinigungsmenge $A \cup B$ *von A und B, und in formalerer Schreibweise ist sie definiert als*

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

Etwas anders ausgedrückt haben wir eine Operation $\cup$ *für Paare von Mengen definiert — die* Vereinigung *— die jedem Paar (A, B) von Mengen deren Vereinigungsmenge $A \cup B$ zuordnet.*

(ii) *Man kann auch mehr als zwei Mengen vereinigen, sogar beliebig viele. Sei A_i, $i \in I$ eine Familie von Mengen. Dann ist*

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

die Vereinigung aller A_i. *Das bedeutet, wir nehmen alle x auf, die in **wenigstens einer** der Mengen A_i liegen. Die **Indexmenge** I kann dabei beliebig (groß) sein.*

Beispiel 4.1.9. *Es gelten:*

- $\{1, 3, 6\} \cup \{2, 6\} = \{1, 2, 3, 6\}$,
- $M \cup \emptyset = M$,
- $\bigcup_{n \in \mathbb{N}} \{-n, n\} = \mathbb{Z}$.

Definition 4.1.10 (Durchschnitt)**.**

(i) *Seien wieder zwei Mengen A und B gegeben. Wir bezeichnen die Menge, die alle Elemente von A enthält, die auch in B enthalten sind, mit $A \cap B$ und nennen sie* Durchschnittsmenge *von A und B. Formal ist sie definiert durch*

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

Analog zu Definition 4.1.8 (i) haben wir somit die Operation $\cap$ für Paare von Mengen definiert, die je zwei Mengen A, B ihre Durchschnittsmenge $A \cap B$ zuordnet.

(ii) *Genau wie die Vereinigung kann man auch den Durchschnitt beliebig vieler Mengen definieren. Sei wieder A_i, $i \in I$ eine Familie von Mengen. Dann ist*

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

der Durchschnitt aller *A_i. Wir nehmen also alle jene Elemente auf, die in* ***allen*** *Mengen A_i liegen.*

(iii) *Haben zwei Mengen A und B leeren Durchschnitt ($A \cap B = \emptyset$), so sagen wir A und B sind* disjunkt.

Sind alle von uns betrachteten Mengen Teilmengen eines (sog.) Universums U, so können wir eine weitere Definition hinzufügen.

Definition 4.1.11 (Komplement)**.** *Sei A eine Teilmenge der Menge U. Dann definieren wir das* Komplement $\complement A$ *von A (in U) durch die Beziehung*

$$\complement A := \{x \in U \mid x \notin A\}.$$

Oft werden auch die Bezeichnungen A' und A^c für das Komplement von A verwendet.

Hinweis: Beachten Sie, dass wir die Universalmenge *nur* zur Bildung des Komplements einführen und verwenden. Alle Rechenoperationen und Rechenregeln, in denen kein Komplement vorkommt, gelten unabhängig von der Existenz solch einer Universalmenge. Ohne Universalmenge muss man auf die Bildung des Komplements verzichten. Man kann es in den meisten Fällen durch die Mengendifferenz (siehe Definition 4.1.13) ersetzen. In diesem Fall müssen aber die Rechenregeln geeignet anpasst werden.

Vergleichen wir die Definitionen mit den logischen Operatoren, die wir in Abschnitt 3.1 eingeführt haben, so erkennen wir rasch die Zusammenhänge. Die Vereinigung ($\cup$) wird gewonnen durch logische Oder-Verknüpfung ($\vee$) der Elementbeziehung zu den zu vereinigenden Mengen. Der Durchschnitt ($\cap$) entspricht der Und-Verknüpfung ($\wedge$), und die Bildung des Komplements ($\complement$) der Negation ($\neg$) der Elementbeziehungen. Diese enge Verwandtschaft zwischen den logischen Verknüpfungen und den Mengenoperationen hat als Konsequenz, dass die Mengenoperationen analoge Rechengesetze erfüllen.

Theorem 4.1.12 (Rechenregeln für Mengenoperationen)**.** *Die mengentheoretischen Operationen $\cup$, $\cap$ und $\complement$ erfüllen die folgenden Operationen, wobei U das für die Komplementbildung notwendige Universum bezeichne.*

Kommutativgesetze:	$A \cup B = B \cup A$	$A \cap B = B \cap A$
Assoziativgesetze:	$A \cup (B \cup C) = (A \cup B) \cup C$	$A \cap (B \cap C) = (A \cap B) \cap C$
Distributivgesetze:	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
Verschmelzungsgesetze:	$A \cup (B \cap A) = A$	$A \cap (B \cup A) = A$
Idempotenzgesetze:	$A \cup A = A$	$A \cap A = A$
Neutralitätsgesetze:	$A \cup \emptyset = A$	$A \cap U = A$
Absorptionsgesetze:	$A \cup U = U$	$A \cap \emptyset = \emptyset$
Komplementaritätsgesetze:	$A \cup \complement A = U$	$A \cap \complement A = \emptyset$
	$\complement\emptyset = U$	
	$\complement U = \emptyset$	
Gesetz des doppelten Komplements:	$\complement(\complement A) = A$	
Gesetze von De Morgan:	$\complement(A \cup B) = \complement A \cap \complement B$	$\complement(A \cap B) = \complement A \cup \complement B$

Beweis. Alle Aussagen des Theorems können entweder durch Aufstellen von Mengentafeln oder aber durch zurückführen auf Theorem 3.1.6 bewiesen werden. Wir beweisen ein Verschmelzungsgesetz nach der 1. und eine Distributivgesetz nach der 2. Methode. Alle anderen Behauptungen folgen analog.

Zu zeigen ist $A \cup (B \cap A) = A$.

A	B	$B \cap A$	$A \cup (B \cap A)$
$\in$	$\in$	$\in$	$\in$
$\in$	$\notin$	$\notin$	$\in$
$\notin$	$\in$	$\notin$	$\notin$
$\notin$	$\notin$	$\notin$	$\notin$

Da die erste und die letzte Spalte übereinstimmen, ist der Beweis des Verschmelzungsgesetz geglückt.

Zu zeigen ist: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. Wir wissen,

$$\begin{aligned}
x \in A \cap (B \cup C) &\Leftrightarrow x \in A \wedge x \in B \cup C \\
&\Leftrightarrow x \in A \wedge (x \in B \vee x \in C) \\
&\Leftrightarrow (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C) \quad \text{(wegen Theorem 3.1.6)} \\
&\Leftrightarrow x \in A \cap B \vee x \in A \cap C \\
&\Leftrightarrow x \in (A \cap B) \cup (A \cap C).
\end{aligned}$$

Außer der explizit angegebenen Äquivalenz gelten alle anderen Zeilen wegen der Definitionen von $\cup$ und $\cap$. Die behauptete Aussage folgt schließlich aus Definition 4.1.3. □

Eine weitere Mengenoperation, die mit der Komplementbildung „verwandt“ ist, ist die bereits erwähnte *Differenz von Mengen.*

Definition 4.1.13 (Mengendifferenz)**.** *Seien A und B zwei Mengen. Die Menge $A \setminus B$ (sprich: A ohne B) ist die Menge aller Elemente von A, die nicht in B sind, also*

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

Bemerkung 4.1.14. *Die Komplementbildung* $\complement A$ *kann mit Hilfe der Mengendifferenz und dem Universum* U *kurz beschreiben als*

$$\complement A = U \setminus A.$$

Beispiel 4.1.15. *Seien* $A = \{2, 3, 6\}$ *und* $B = \{2, 5, 7\}$. *Dann ist* $A \setminus B = \{3, 6\}$.

Die *symmetrische Mengendifferenz* ist die letzte Grundoperation, die wir für Mengen einführen wollen.

Definition 4.1.16 (Symmetrische Mengendifferenz)**.** *Es seien wieder zwei Mengen* A *und* B *gegeben. Wir definieren die* symmetrische Differenz $A \triangle B$ *von* A *und* B *als die Menge derjenigen Elemente von* A *und* B, *die nicht in beiden Mengen liegen; formal*

$$A \triangle B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

Diese Definition beinhaltet genau genommen eine kleine Behauptung, nämlich dass die beiden Ausrücke rechts des definierenden Gleichheitszeichens (:=), d.h. $(A \setminus B) \cup (B \setminus A)$ und $(A \cup B) \setminus (A \cap B)$ ihrerseits gleich sind. Wird in einem mathematischen Text eine solche „Mini-Behauptung“ nicht weiter begründet (wie etwa oben), so bedeutet das, dass die AutorInnen der Ansicht sind, dass die Gültigkeit der Behauptung klar auf der Hand liegt. Die Aufgabe der LeserInnen und besonders der AnfängerInnen ist es, solche kleinen Behauptungen in mathematischen Texten aufzuspüren (d.h. diese nicht zu übersehen) und sich ihre Gültigkeit klar zumachen. Ist der Text gut geschrieben (d.h. wird Ihr Wissensstand von den AutorInnen richtig eingeschätzt), so sollten Sie auch keine Probleme haben, diese „Mini-Behauptungen“ mit „Mini-Beweisen“ zu belegen. Also, falls Ihnen die Aussage $(A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$ nicht klar ist, so nehmen Sie jetzt Papier und Bleistift zur Hand und beweisen diese etwa mit einer der Methoden, die im Beweis von Theorem 4.1.12 verwendet wurden...

Beispiel 4.1.17. *Seien wiederum* $A = \{2, 3, 6\}$ *und* $B = \{2, 5, 7\}$. *Dann ist* $A \triangle B = \{3, 6, 5, 7\}$.

4.1.3. Potenzmenge, Produktmenge.

Kommen wir nun, nachdem wir Operationen definiert haben, um aus bestehenden Mengen neue Mengen zu definieren, zum nächsten Schritt. Zunächst verwenden wir die Tatsache, dass Mengen wieder Mengen enthalten dürfen, um die Potenzmenge einer Menge zu definieren.

Definition 4.1.18 (Potenzmenge)**.** *Sei* M *eine Menge. Die* Potenzmenge $\mathbb{P}M$ *von* M *ist definiert als die Menge aller Teilmengen von* M.

Beispiel 4.1.19 (Potenzmengen)**.**

- *Die Potenzmenge von* $\{1, 2, 3\}$ *ist*

$$\mathbb{P}\{1, 2, 3\} = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\big\}.$$

- *Die Potenzmenge der leeren Menge ist* nicht *die leere Menge sondern eine einelementige Menge, die nur die leere Menge* enthält. *(Also ein Sack, der nur einen leeren Sack enthält!)*

$$\mathbb{P}\emptyset = \{\emptyset\}.$$

Allgemein bezeichnet man eine Menge, die wieder Mengen enthält als **Mengensystem**.

Sind schließlich zwei Mengen A und B gegeben, so kann man die Produktmenge $A \times B$ bilden. Zu diesem Zweck formen wir aus den Elementen a von A und b von B **geordnete Paare** (a, b). In diesen Paaren schreiben wir die Elemente von A an erster und die Elemente von B an zweiter Stelle. Zwei dieser geordneten Paare wollen wir nur dann als gleich betrachten, wenn *beide* Komponenten übereinstimmen.

Definition 4.1.20 (Produkt)**.**

(i) *Seien A und B Mengen. Die* Produktmenge $A \times B$*, auch genannt das* Cartesische Produkt *von A und B, ist die Menge aller geordneten Paare (a, b) aus Elementen von A und B, formal*

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

(ii) *Sind k-stück Mengen $M_1, \ldots, M_k$ gegeben, so können wir analog die geordneten k-tupel bilden $(m_1, \ldots, m_k)$ mit $m_i \in M_i$ für $i = 1, \ldots, k$. Das* Cartesische Produkt $\bigtimes_{i=1}^k M_i = M_1 \times \ldots \times M_k$ *der M_i ist dann die Menge aller geordneten k-tupel dieser Form, d.h.*

$$\bigtimes_{i=1}^{k} M_i := \{(m_1, \ldots, m_k) \mid \forall i : m_i \in M_i\}.$$

(iii) *Ist $A = B$ bzw. $M_i = A$ für alle i, so schreiben wir statt $A \times A$ und $\underbrace{A \times \ldots \times A}_{k \text{ mal}}$ kurz A^2 bzw. A^k.*

Beispiel 4.1.21 (Produkte)**.** *Seien $A = \{1, 2, 3\}$ und $B = \{a, b\}$, dann ist*

$$A \times B = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$$

und

$$A^2 = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}.$$

Man kann auch das Cartesische Produkt beliebig vieler Mengen M_i, $i \in I$ bilden; die Definition ist allerdings ein wenig komplizierter und benötigt Funktionen. Daher wird sie erst in Abschnitt 4.3 nachgeholt werden.

4.2. Relationen

In diesem Abschnitt geht es darum einen Formalismus zu entwickeln, der es gestattet, Elemente von Mengen miteinander in Beziehung zu setzen.

Beispiel 4.2.1. *Sei etwa M die Menge aller HörerInnen in einem Hörsaal. Betrachten wir die Beziehung „ist verwandt mit“. Wir können dann zu je zwei Personen A und B im Hörsaal eine Aussage darüber machen, ob A mit B verwandt ist.*

Eine andere Beziehung, die wir auf M betrachten könnten ist „ist Bruder von“. Natürlich ist jeder Bruder auch ein Verwandter. Umgekehrt muss das nicht der Fall sein.

Schließlich ist eine dritte mögliche Beziehung „wohnt im selben Bezirk wie“.

Beziehungen in der Art von Beispiel 4.2.1 zwischen Elementen von Mengen nennt man Relationen. Im folgenden wollen wir eine mathematische Definition dafür geben.

Definition 4.2.2 (Relation)**.** *Seien M und N Mengen. Eine* Relation auf $M \times N$ *ist eine Teilmenge R des Cartesischen Produkts, d.h. $R \subseteq M \times N$.*

Für zwei Elemente $a \in M$ und $b \in N$ sagen wir: a steht in Relation mit b*, falls $(a, b) \in R$ gilt. Wir schreiben dann in Symbolen*

$$a\,R\,b.$$

Stehen a und b nicht miteinander in Relation, so schreiben wir $a \not{R}\, b$.

Meist werden Relationen nicht mit R sondern mit Symbolen bezeichnet. Typische Relationssymbole sind $<$, $\subset$, $\sim$, $\cong$, $\ll$, $\equiv$, $\simeq$, $\sqsubset$, $\frown$, $\preccurlyeq$ und viele andere mehr. Gerichtete Symbole wie $<$ werden üblicherweise für Ordnungsrelationen (siehe Abschnitt 4.2.2) verwendet, während symmetrische Symbole wie $\simeq$ meist für Äquivalenzrelationen (siehe Abschnitt 4.2.1) stehen.

Ist R eine Relation auf $M \times N$ und gilt $M = N$, so sprechen wir von einer *Relation auf* M. Im Folgenden befassen wir uns (fast) ausschließlich mit diesem Fall.

Beispiel 4.2.3 (Relationen)**.** *Die Beziehungen aus Beispiel 4.2.1 sind natürlich Relationen auf der Menge M der HörerInnen im Hörsaal. Haben wir etwa ein Geschwisterpaar S und B im Hörsaal, so müssen wir in unsere Relation V für „verwandt“ die beiden Paare (S, B) und (B, S) aufnehmen. Ist S weiblich und B männlich, so darf in der „Bruder“-Relation R nur das Paar (B, S) vorkommen (es gilt ja „B ist Bruder von S“ aber nicht „S ist Bruder von B“).*

Zwei wichtige Hauptgruppen von Relationen wollen wir in den folgenden Abschnitten untersuchen. Zuvor definieren wir jedoch noch zwei Eigenschaft für Relationen, die in beiden Abschnitten wichtig sein werden.

Definition 4.2.4 (Transitivität, Reflexivität)**.** *Sei R eine Relation auf einer Menge M.*

(i) *R heißt* transitiv*, wenn für alle $a, b, c \in M$ gilt, dass*

$$a\,R\,b \wedge b\,R\,c \Rightarrow a\,R\,c.$$

(ii) *R heißt* reflexiv*, wenn für alle $a \in M$ gilt, dass $a\,R\,a$.*

Beispiel 4.2.5 (Transitivität, Reflexivität)**.** *Kehren wir noch einmal — aber nicht zum letzten Mal — zu den Relationen aus Beispiel 4.2.1 zurück. Nicht alle sind transitiv, denn wenn A mit B und B mit C verwandt sind, so ist noch lange nicht A mit C verwandt. Anderes gilt für Brüder. Ist A Bruder von B und B Bruder von C, so ist auch A Bruder von C. Auch das Wohnen im gleichen Bezirk ist eine transitive Relation.*

Man könnte sagen, die Verwandtschaftsrelation ist reflexiv, wenn man festlegt, dass jeder Mensch mit sich selbst verwandt ist. Die Bruderbeziehung ist jedoch nicht reflexiv.

Auch ohne weitere Definition ist „wohnt im selben Bezirk wie“ eine reflexive Relation.

4.2.1. Äquivalenzrelationen. In diesem Abschnitt befassen wir uns mit einer speziellen Klasse von Relationen, nämlich jenen die zusätzlich zu den beiden oben definierten Eigenschaften der Reflexivität und Transitivität die Eigenschaft der Symmetrie besitzen. Formal definieren wir:

Definition 4.2.6 (Äquivalenzrelation)**.** *Eine reflexive und transitive Relation $\sim$ auf einer Menge M heißt* Äquivalenzrelation*, falls sie folgende weitere Eigenschaft erfüllt:*

Symmetrie: $\forall x, y \in M : (x \sim y \Rightarrow y \sim x)$.

Gilt $x \sim y$, so nennen wir x und y äquivalent*.*

Beispiel 4.2.7 (Äquivalenzrelationen)**.** *Wenn wir ein weiteres Mal die Relationen aus Beispiel 4.2.1 bemühen, so erkennen wir schnell, dass „wohnt im selben Bezirk wie“ eine Äquivalenzrelation ist. Die Symmetrie ist erfüllt, denn wenn A und B im selben Bezirk wohnen, wohnen auch B und A im selben Bezirk.*

Die zweite Relation „ist Bruder von“ ist keine *Äquivalenzrelation, da weder Reflexivität noch Symmetrie gelten.*

„Ist verwandt mit“ ist zwar symmetrisch, aber da die Transitivität falsch ist, ist es keine Äquivalenzrelation.

Ist eine Äquivalenzrelation $\sim$ auf einer Menge definiert, so können wir die Relation dafür verwenden, miteinander äquivalente Elemente von M in Gruppen zusammenzufassen. Dieses Prinzip ist wohlbekannt, denn in Telefonbüchern werden etwa jene Ärzte in eine Gruppe zusammengefasst, die im selben Bezirk praktizieren.

Definition 4.2.8 (Äquivalenzklasse). *Sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Wir definieren die* Äquivalenzklasse von $a \in M$ *durch*

$$C_a := \{b \in M \mid b \sim a\}.$$

Alternative Bezeichnungen für C_a sind auch $[a]$ und $\bar{a}$.

Aus der Definition sehen wir unmittelbar, dass für jedes $a \in M$ die Äquivalenzklasse $C_a \subseteq M$ erfüllt. Wegen der Reflexivität von $\sim$ gilt $a \in C_a$ (Äquivalenzklassen sind also niemals leer!) und somit $\bigcup_{a \in M} C_a = M$.

Eine zweite wichtige Eigenschaft der Äquivalenzklassen wollen wir in der nachfolgenden Proposition fest halten.

Proposition 4.2.9 (Eigenschaften von Äquivalenzklassen). *Sei M eine Menge und $\sim$ eine Äquivalenzrelation auf M. Dann sind zwei Äquivalenzklassen C_a und C_b entweder disjunkt oder gleich. In Symbolen*

$$C_a \cap C_b \neq \emptyset \Leftrightarrow C_a = C_b.$$

Wie in Definition 4.1.16 (vgl. den dieser Definition folgenden Abschnitt) ist in Proposition 4.2.9 eine kleine (zusätzliche) Behauptung versteckt, nämlich dass für zwei Mengen C_a und C_b die Aussage C_a und C_b sind entweder disjunkt oder gleich gleichbedeutend mit der Aussage $C_a \cap C_b \neq \emptyset \Leftrightarrow C_a = C_b$ ist. Ehrlich, haben Sie das bemerkt? Gut, und falls Ihnen diese „Mini-Behauptungen“ nicht klar ist, dann …(Tipp: Wahrheitstabelle)

Beweis. Da es sich um eine Äquivalenz handelt …
$\Leftarrow$: Ist $C_a = C_b$, so ist auch $C_a \cap C_b = C_a \neq \emptyset$, weil Äquivalenzklassen niemals leer sind.
$\Rightarrow$: Ist umgekehrt $C_a \cap C_b \neq \emptyset$. Dann existiert ein $y \in C_a \cap C_b$, und somit gelten $y \sim a$ und $y \sim b$. Aus Symmetrie und Transitivität folgt $a \sim b$. Es bleibt zu zeigen, dass $a \sim b \Rightarrow C_a = C_b$. Sei dazu $x \in C_a$. Dann wissen wir $x \sim a$ und wegen der Transitivität auch $x \sim b$ und damit $x \in C_b$. Also gilt $C_a \subseteq C_b$. Nachdem wir analog durch Vertauschen von a und b in obiger Argumentation $C_b \subseteq C_a$ beweisen können, folgt $C_a = C_b$, was wir behauptet hatten. □

Wir finden also für jede Äquivalenzrelation $\sim$ auf einer Menge M eine Familie von Teilmengen von M, die Äquivalenzklassen C_a, die

(i) $\bigcup_{a \in M} C_a = M$ (Man sagt: Die C_a **überdecken** M) und
(ii) $C_a \cap C_b \neq \emptyset \Leftrightarrow C_a = C_b$

erfüllen.

Eine Familie von Teilmengen einer gegebenen Menge M, die die obigen Eigenschaften (i) und (ii) erfüllt ist ein mathematisch interessantes Objekt — und zwar unabhängig davon, ob diese Familie aus den Äquivalenzklassen einer Äquivalenzrelation auf M besteht oder *irgendwelchen* Teilmengen von M. Wir definieren daher:

Definition 4.2.10 (Partition). *Eine Familie disjunkter Teilmengen einer Menge, die die gesamte Menge überdecken, nennt man* Partition.

Etwas formaler können wir auch schreiben: Eine Familie U_i, $i \in I$ (I bezeichnet hier eine beliebige Indexmenge) von Teilmengen einer Menge M heißt Partition *von M, falls die folgenden beiden Eigenschaften gelten:*

(i) $U_i \cap U_j \neq \emptyset \Leftrightarrow U_i = U_j$
(ii) $\bigcup_{i \in I} U_i = M$.

Anders ausgedrückt ist eine Familie von Teilmengen einer Menge M genau dann eine Partition, wenn jedes Element von M in genau einer der Teilmengen liegt. (Eigenschaft (ii) sorgt dafür, dass jedes Element von M in *mindestens* einer der Teilmengen liegt, Eigenschaft (i) dafür, dass es *höchstens* eine ist.)

Theorem 4.2.11 (Äquivalenzklassen als Partition)**.** *Jede Äquivalenzrelation $\sim$ auf einer Menge M definiert eine Partition von M, und umgekehrt kann man aus jeder Partition U_i, $i \in I$ einer Menge M eine Äquivalenzrelation $\sim$ gewinnen durch*

$$a \sim b :\Leftrightarrow \exists i \in I : a, b \in U_i.$$

Beweis. Wir wissen bereits, dass eine Äquivalenzrelation auf M eine Partition definiert (Proposition 4.2.9), nämlich die Partition in Äquivalenzklassen.

Sei umgekehrt eine Partition U_i, $i \in I$ gegeben, und sei die Relation $\sim$ wie in der Aussage des Theorems definiert. Es bleibt zu zeigen, dass $\sim$ eine Äquivalenzrelation ist.

Reflexivität: Für alle $a \in M$ gilt $a \sim a$, da wegen $\bigcup_{i\in I} U_i = M$ ein $j \in I$ existieren muss mit $a \in U_j$.

Symmetrie: Das folgt ganz offensichtlich aus der Definition von $\sim$.

Transitivität: Gelten $a \sim b$ und $b \sim c$, so wissen wir, dass ein $j \in I$ existiert mit $a, b \in U_j$ und ein $k \in I$ mit $b, c \in U_k$. Es ist somit $b \in U_j \cap U_k$, und daher ist wegen (i) in Definition 4.2.10 $U_j = U_k$. Daraus wiederum folgt, dass $a, b, c \in U_j$ und daher $a \sim c$ gilt.

Also ist $\sim$ tatsächlich eine Äquivalenzrelation. □

Von Äquivalenzrelationen induzierte Partitionen von Mengen (d.h. Partitionen von Mengen in Äquivalenzklassen einer gegebenen Äquivalenzrelation) sind in der Mathematik äußerst wichtig. Aus diesem Grund hat man der Menge aller Äquivalenzklassen einen eigenen Namen gegeben.

Definition 4.2.12 (Quotient)**.** *Sei $\sim$ eine Äquivalenzrelation auf M. Wir definieren die* Quotientenmenge *oder auch* Faktormenge *$M/_\sim$ (sprich: M modulo Tilde) als die Menge aller Äquivalenzklassen bezüglich $\sim$.*

Beispiel 4.2.13 (Restklassen)**.** *Sei $\mathbb{N} \ni p \geq 2$ gegeben. Wir definieren auf $\mathbb{Z}$ die Relation*

$$n \sim_p m :\Leftrightarrow \exists k \in \mathbb{Z} \text{ mit } n = m + kp.$$

Gilt $n \sim_p m$ so sagen wir n ist kongruent m modulo p und schreiben manchmal auch $n \equiv m(p)$. Die Relation $\sim_p$ ist eine Äquivalenzrelation:

Reflexivität: $n \sim_p n$, weil $n = n + 0p$,

Symmetrie: Ist $n \sim_p m$, so finden wir ein $k \in \mathbb{Z}$ mit $n = m + kp$, und durch Umformen finden wir $m = n + (-k)p$. Da mit k auch $-k$ in $\mathbb{Z}$ ist, folgt $m \sim_p n$.

Transitivität: Gelten $n_1 \sim_p n_2$ und $n_2 \sim_p n_3$, so finden wir k_1 und k_2 mit $n_1 = n_2 + k_1p$ und $n_2 = n_3 + k_2p$. Setzen wir die Gleichungen zusammen, finden wir $n_1 = n_3 + (k_1 + k_2)p$, und $k_1 + k_2$ ist als Summe ganzer Zahlen eine ganze Zahl. Deshalb folgt $n_1 \sim_p n_3$.

Diese Äquivalenzrelation erzeugt genau p Äquivalenzklassen

$$\begin{aligned}
\overline{0} &= \{0, \pm p, \pm 2p, \pm 3p, \dots\} \\
\overline{1} &= \{1, 1 \pm p, 1 \pm 2p, 1 \pm 3p, \dots\} \\
&\vdots \quad \vdots \\
\overline{p-1} &= \{-1, -1 \pm p, -1 \pm 2p, -1 \pm 3p, \dots\}.
\end{aligned}$$

Die p-elementige Faktormenge $\mathbb{Z}/_{\sim_p}$ wird in der Mathematik üblicherweise mit $\mathbb{Z}_p$ bezeichnet, und man nennt sie die ***Restklassen modulo*** *p. Dieser Name begründet sich aus* *der Tatsache, dass zwei Zahlen aus der selben Restklasse modulo p bei Division durch p denselben Rest aufweisen.*

4.2.2. Ordnungsrelationen. Die zweite große Klasse von Relationen, die neben der Reflexivität und Transitivität die Eigenschaft der Antisymmetrie besitzt, dient dazu, Mengen zu ordnen.

Definition 4.2.14 (Ordnung)**.**

(i) *Eine reflexive und transitive Relation $\preceq$ (sprich: vor oder gleich) auf M heißt* Ordnungsrelation *oder* Halbordnung*, falls sie die folgende zusätzliche Eigenschaft erfüllt:*

Antisymmetrie: *Die Beziehungen $a \preceq b$ und $b \preceq a$ implizieren schon Gleichheit $a = b$. In Symbolen*

$$a \preceq b \wedge b \preceq a \Rightarrow a = b.$$

(ii) *Gilt für zwei Elemente von M weder $x \preceq y$ noch $y \preceq x$, so sagt man x und y sind* nicht vergleichbar *(bezüglich $\preceq$). Andernfalls nennt man die beiden Elemente* vergleichbar*.*

(iii) *Sind je zwei Elemente von M vergleichbar, gilt also für je zwei Elemente $x, y \in M$ wenigstens eine der Relationen $x \preceq y$ oder $y \preceq x$, so nennt man die Relation eine* Totalordnung *oder* lineare Ordnung *auf M.*

(iv) *Betrachten wir eine Menge M zusammen mit einer (Total)-Ordnung $\preceq$, so nennen wir das Paar $(M, \preceq)$* (total) geordnete Menge*.*

Um mit Ordnungsrelationen leichter hantieren zu können, müssen wir einige Schreibweisen vereinbaren. Gilt $x \preceq y$, so schreiben wir auch manchmal $y \succeq x$. Haben wir $x \preceq y$ und gilt $x \neq y$, so kürzen wir ab zu $x \prec y$ (sprich: x vor y). Analog definieren wir $y \succ x$. Gilt andererseits $x = y$ oder $x \prec y$, so schreiben wir $x \preceq y$.

Beispiel 4.2.15 (Ordnungen)**.**

- *Das bekannteste Beispiel für eine Ordnungsrelation (eine Totalordnung) ist die Beziehung $\leq$ auf den reellen Zahlen $\mathbb{R}$. Wir nennen die geordnete Menge $(\mathbb{R}, \leq)$ oft $\mathbb{R}$ mit der natürlichen Ordnung.*
- *Sei M die Menge aller Menschen. Wir definieren die Relation $\prec$ durch $A \prec B$, wenn A ein Vorfahre von B ist. Die entstehende Relation $\preceq$ ist klarerweise reflexiv und transitiv. Die Antisymmetrie folgt aus der Tatsache, dass kein Mensch Vorfahre von sich selbst sein kann. Es gibt aber Paare von Menschen, die nicht miteinander vergleichbar sind, für die also weder $A \preceq B$ noch $A \succeq B$ gelten. Die Relation „Ist Vorfahre von" ist also eine Halbordnung auf M.*

So wie eine Äquivalenzrelation auf einer Menge M eine Struktur definiert, die wichtige Folgestrukturen entstehen lässt, erzeugt auch eine Ordnungsrelation auf M Folgebegriffe.

Definition 4.2.16 (Schranken)**.** *Sei $(M, \preceq)$ eine geordnete Menge, und sei $E \subseteq M$ eine Teilmenge.*

(i) *Gibt es ein $\beta \in M$ mit der Eigenschaft*

$$x \preceq \beta \text{ für jedes Element } x \in E,$$

so nennen wir β eine obere Schranke *von E.*

(ii) Untere Schranken *sind analog durch Ersetzen von $\preceq$ durch $\succeq$ definiert.*

(iii) *Die Teilmenge E heißt* nach oben (unten) beschränkt*, falls sie eine obere (untere) Schranke besitzt. Sie heißt* beschränkt*, falls sie nach oben und unten beschränkt ist.*

Um Beispiele zur Illustration von Definition 4.2.16 geben zu können, wiederholen wir aus dem Schulstoff den Begriff des **Intervalls**. Seien $a < b \in (\mathbb{R}, \leq)$. Das *offene Intervall* $]a, b[$ (oft auch (a, b) geschrieben), das *abgeschlossene Intervall* $[a, b]$ bzw. die *halboffenen Intervalle* $[a, b[$ und $]a, b]$ sind definiert als die folgenden Teilmengen der reellen Zahlen:

$$\begin{aligned}]a,b[&:= \{x \in \mathbb{R} \mid a < x < b\}, & [a,b] &:= \{x \in \mathbb{R} \mid a \leq x \leq b\} \\ [a,b[&:= \{x \in \mathbb{R} \mid a \leq x < b\}, &]a,b] &:= \{x \in \mathbb{R} \mid a < x \leq b\}. \end{aligned}$$

Die einseitig und zweiseitig *unendlichen Intervalle* sind definiert als

$$\begin{aligned}]-\infty, a[&:= \{x \in \mathbb{R} \mid x < a\}, &]-\infty, a] &:= \{x \in \mathbb{R} \mid x \leq a\} \\ ]a, \infty[&:= \{x \in \mathbb{R} \mid x > a\}, & [a, \infty[&:= \{x \in \mathbb{R} \mid x \geq a\} \\ ]-\infty, +\infty[&:= \mathbb{R}. \end{aligned}$$

Beispiel 4.2.17 (Schranken)**.**

- *Betrachten wir die geordnete Menge* $(\mathbb{R}, \leq)$. *Das Intervall* $E = [0, 1]$ *ist eine Teilmenge von* $\mathbb{R}$. *Jede Zahl im Intervall* $[1, \infty[$ *ist obere Schranke von* E, *und jede Zahl im Intervall* $]-\infty, 0]$ *ist untere Schranke von* E. E *ist also beschränkt.*
- *Die Menge* $M := \{1/n \mid n \in \mathbb{N} \setminus \{0\}\}$ *ist nach oben und unten beschränkt.* M *hat dieselben unteren und oberen Schranken wie* E.
- *Die Menge aller Primzahlen ist als Teilmenge von* $\mathbb{R}$ *nach unten beschränkt, sie besitzt aber keine obere Schranke.*
- *Die Menge* $\mathbb{Z} \subset \mathbb{R}$ *ist weder nach oben noch nach unten beschränkt.*

Wir sehen aus dem vorigen Beispiel, dass obere und untere Schranken bei weitem nicht eindeutig sind. Die interessante Frage ist, ob es eine ausgezeichnete obere bzw. untere Schranke gibt. Die Beantwortung dieser Frage für die geordnete Menge $(\mathbb{Q}, \leq)$ wird uns in Kapitel 6 zu den reellen Zahlen führen. Hier wollen wir uns mit einer Definition begnügen.

Definition 4.2.18 (Supremum und Infimum, Maximum und Minimum)**.** *Sei* $(M, \preceq)$ *eine geordnete Menge.*

(i) *Sei* E *eine nach oben beschränkte Teilmenge von* M. *Existiert ein* $\alpha \in M$ *mit den Eigenschaften*
 (1) α *ist eine obere Schranke von* E,
 (2) *Ist* $M \ni \gamma \prec \alpha$, *so ist* γ *keine obere Schranke von* E,
 so nennen wir α *die kleinste obere Schranke oder das* Supremum *von* E, *und wir schreiben*

$$\alpha = \sup E$$

(ii) *Analog definieren wir die größte untere Schranke, das* Infimum

$$\alpha = \inf E$$

 einer nach unten beschränkten Teilmenge.

(iii) *Besitzt eine Teilmenge* E *ein Supremum (Infimum)* α, *und gilt* $\alpha \in E$, *dann nennen wir* α *auch das* Maximum *(*Minimum*) von* E, *in Zeichen* $\max E$ *(*$\min E$*).*

Beispiel 4.2.19 (sup, max, inf, min)**.** *Seien* E *und* M *wie in Beispiel 4.2.17. Es gilt* $0 = \inf E = \inf M$ *und* $1 = \sup E = \sup M$. *Für* E *sind* 1 *und* 0 *sogar Maximum bzw. Minimum. Für* M *ist* 1 *ein Maximum, aber* 0 *ist kein Minimum, da* $0 \notin M$.

4.3. Abbildungen

Wie bereits früher erwähnt, besteht ein großer Teil der modernen Mathematik in der Analyse von Strukturen. Diese Strukturen bestehen aus Objekten und den Beziehungen zwischen diesen Objekten. Wir haben schon erwähnt, dass *Mengen* für die meisten Strukturen die Basis bilden. Die in diesem Abschnitt behandelten Abbildungen sind neben den Relationen *die* Basis für die Beziehungen zwischen Objekten.

Wegen der zentralen Bedeutung des Begriffs Abbildung bzw. Funktion (diese beiden Bezeichnungen werden synonym verwendet; siehe auch unten) werden wir zwei Definitionen geben. Wir beginnen mit der weniger abstrakten, die Ihnen sicherlich bekannt vorkommen wird.

Definition 4.3.1 (Funktion)**.** *Seien A und B Mengen.*

(i) *Unter einer* Funktion *oder* Abbildung f von A nach B *verstehen wir eine Vorschrift, die jedem $a \in A$* ***genau ein*** *$b \in B$ zuordnet.*
(ii) *Das dem Element a zugeordnete Element b bezeichnen wir mit $f(a)$ und nennen es* den Wert der Funktion f an der Stelle a *oder das* das Bild von a unter f*; a wird als* Urbild von b unter f *bezeichnet.*
(iii) *Weiters wird A als* Definitionsmenge *oder* -bereich *von f bezeichnet und B als* Zielmenge *oder* -bereich *von f.*

WICHTIG: Zur Festlegung einer Funktion **muß** man ausdrücklich Definitions- und Zielmenge angeben. Meist werden die Schreibweisen

$$f : A \to B \text{ oder } A \xrightarrow{f} B$$

verwendet. Die Angabe der Zuordnungsvorschrift alleine ist keinesfalls ausreichend (siehe auch Beispiel 4.3.10 unten)!

Das Symbol

$$a \mapsto f(a)$$

drückt aus, dass die Funktion f dem Element a des Definitionsbereichs das Bild $f(a)$ im Zielbereich zuordnet. Oft wird dieses Symbol auch zur Bezeichnung der Funktion selbst verwendet und man spricht von „der Funktion $a \mapsto f(a)$“ (lies: a geht über (in) $f(a)$). Dann muss allerdings Definitions- und Zielbereich gesondert angegeben werden.

Die ausführlichste und genaueste Darstellung einer Funktion erfolgt durch die Notation

$$\begin{array}{rcl} f : A & \to & B \\ a & \mapsto & f(a) \end{array} \qquad \text{bzw.} \qquad \begin{array}{rcl} f : A & \to & B \\ f(a) & = & \dots \end{array}$$

Beispiel 4.3.2 (Funktionsschreibweise)**.** *Die Funktion, die jeder nichtnegativen reellen Zahl ihre Quadratwurzel zuordnet schreiben wir*

$$\begin{array}{rcl} f : [0, \infty[& \to & \mathbb{R} \\ x & \mapsto & \sqrt{x} \end{array} \qquad \textit{bzw.} \qquad \begin{array}{rcl} f : [0, \infty[& \to & \mathbb{R} \\ f(x) & = & \sqrt{x}. \end{array}$$

Einen Schönheitsfehler hat unserer Definition allerdings: sie beruht auf dem Wort „Zuordnung“, das wir strenggenommen gar nicht definiert haben — obwohl in diesem Fall die alltagssprachliche Bedeutung den mathematsichen Inhalt sehr genau trifft. Dieses Problem können wir umgehen, indem wir eine rein mengentheoretische Definition des Funktionsbegriffs geben. Das ist unser nächstes Ziel.

Nach unserer obigen Definition setzt eine Funktion $f : A \to B$ die Elemente a von A mit gewissen Elementen $b = f(a)$ von B in Verbindung; fassen wir diese zu geordneten Paaren (a, b) zusammen, so folgt aus der **Eindeutigkeit** der Zuordnung, dass Paare mit gleichen ersten Komponenten gleiche zweite Komponenten besitzen, also bereits (als geordnetes Paar) gleich sind. Die durch f gegebene „Zuordnung“ kann deshalb als *spezielle* Teilmenge des Produkts $A \times B$ beschrieben werden. Genau das macht sich die folgende Definition zunutze.

Definition 4.3.3 (Mengentheoretische Definition des Funktionsbegriffs)**.** *Eine* Funktion *ist ein Tripel $f = (A, B, G)$ bestehend aus einer Menge A, genannt* Definitionsbereich, *einer Menge B, genannt* Zielbereich *und einer Teilmenge G des Produkts $A \times B$ mit den Eigenschaften*

(1) $\forall a \in A\ \exists b \in B : (a, b) \in G$
(D.h. jedes $a \in A$ tritt als erste Komponente eines Paares in G auf.)
(2) $\forall a \in A\ \forall b_1, b_2 \in B :\ (a, b_1) \in G \wedge (a, b_2) \in G\ \Rightarrow b_1 = b_2$
(D.h. stimmen die ersten Komponenten eines Paares in G überein, dann auch die zweiten.)

Die Menge G heißt Graph *der Funktion f und wird oft auch mit $G(f)$ bezeichnet. Gilt $(a, b) \in G$, so schreiben wir $f(a) = b$ und wir können den Graphen schreiben als*

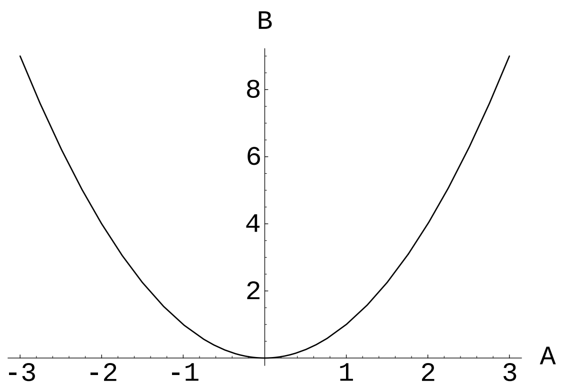
$$G(f) = \{(a, f(a)) | a \in A\}.$$

Die Paarungen $(a, f(a))$ sind gewissermaßen das abstrakte Analogon zur Zusammenstellung der a-Werte und der zugehörigen Funktionswerte $f(a)$ in einer Wertetabelle oder der graphischen Darstellung der Funktion als “Kurve“ in einem kartesischen Koordinatensystem (falls A und B Teilmengen von $\mathbb{R}$ sind); zwei Konzepte, die Ihnen sicherlich aus der Schule ein Begriff sind.

Vielleicht haben Sie bemerkt, dass wir noch immer nicht gesagt haben, *was* eine „Zuordnung“ denn eigentlich *ist*. Die moderne Mathematik zieht sich in dieser und ähnlichen Situationen aus der Affäre, indem Sie Ihnen die Objekte „aufzählt“, die einander „zugeordnet“ sind; in unserem Fall die geordneten Paare $(a, f(a))$.

Nach dieser „philosophischen“ Bemerkung kommen wir zu einem einfachen, konkreten Beispiel.

Beispiel 4.3.4 (Funktion)**.** *Sei $A = \mathbb{R} = B$. Wir betrachten die Funktion $f : x \mapsto x^2$. Dann gilt $G = \{(x, x^2) | x \in \mathbb{R}\} \subseteq \mathbb{R} \times \mathbb{R}$, d.h. $(0, 0) \in G$, $(1, 1) \in G$, $(-1, 1) \in G$, $(2, 4) \in G$, ...*



$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^2$$

Bemerkung 4.3.5 (Funktionen als Relationen)**.** *Aus Definition 4.3.3 wird auch ersichtlich, dass jede Funktion $f : A \to B$ als spezielle Relation auf $A \times B$ aufgefasst werden kann: für jedes $a \in A$ existiert* ***genau ein*** *$b \in B$, mit dem es in Relation steht. Oft wird eine Funktion auch als derartig „eindeutige“ Relation* definiert.

Obwohl der Begriff der Abbildung zentral für die moderne Mathematik ist, wurde er erst sehr spät (im zwanzigsten Jahrhundert!) formalisiert. Daher existieren abhängig vom betrachteten Gebiet viele verschiedene Ausdrücke für Abbildung.

Der Terminus **Abbildung** ist der allgemeinste, doch der Begriff **Funktion** ist ein Synonym, auch wenn er meist dann verwendet wird, wenn B ein Körper (siehe Kapitel 5) ist.

Eine **Transformation** ist eine Abbildung einer Menge in sich (also für $A = B$). Eine bijektive (siehe Definition 4.3.9 (iii) unten) Transformation einer endlichen Menge heißt auch **Permutation**.

Ein **Operator** ist eine Abbildung zwischen Mengen von Abbildungen. So bildet etwa der *Ableitungsoperator* jede differenzierbare Funktion auf ihre Ableitungsfunktion ab.

Schließlich taucht besonders in der Linearen Algebra und der Funktionalanalysis der Begriff **Form** auf. Dieser beschreibt eine multilineare Abbildung in den Grundkörper eines Vektorraums (siehe Lineare Algebra!).

Es ist wichtig, in Texten zwischen der Funktion f und den Werten $f(x)$ einer Funktion zu unterscheiden.

Die Abbildung $\cancel{f(x)}$ ...

Dafür hat man die $\mapsto$-Notation.

Die Abbildung $f : x \mapsto f(x)$ ...

wäre in Ordnung.

Wir können mit Hilfe einer Abbildung ganze Teilmengen von A nach B abbilden.

Definition 4.3.6 (Bild)**.** *Sei $f : A \to B$ eine Abbildung, und sei $M \subseteq A$ eine Teilmenge. Wir nennen die Menge*

$$f(M) := \{b \in B \mid \exists a \in M : f(a) = b\}$$

das Bild der Menge M unter f*.*

Umgekehrt können wir für eine Teilmenge $N \subseteq B$ des Bildbereichs alle Elemente in A suchen, deren Bilder in N liegen.

Definition 4.3.7 (Urbild)**.** *Sei wieder $f : A \to B$ eine Abbildung.*

(i) *Sei $N \subseteq B$ eine Teilmenge des Bildbereichs. Wir definieren die Menge*

$$f^{-1}(N) := \{a \in A \mid f(a) \in N\}$$

und nennen sie das Urbild der Menge N unter f*.*

(ii) *Für ein Element $b \in B$ definieren wir das* Urbild von b *durch $f^{-1}(b) := f^{-1}(\{b\})$. Beachten Sie dabei, dass das Urbild von b **eine Menge** ist!*

Beispiel 4.3.8 (Bild und Urbild)**.** *Betrachten wir nochmals die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$. Das Bild der Menge $M = [-1, 1]$ ist die Menge $f(M) = [0, 1]$. Das Urbild von $N = [-4, 4]$ ist die Menge $f^{-1}(N) = [-2, 2]$, und das Urbild des Punktes* 9 *ist* die Menge *$f^{-1}(9) = \{-3, 3\}$.*

Kommen wir jetzt zu *den* drei grundlegenden Eigenschaften von Abbildungen.

Definition 4.3.9 (Injektiv, surjektiv, bijektiv)**.** *Sei $f : A \to B$ eine Abbildung.*

(i) *Wir sagen f ist* injektiv*, wenn jedes Element $b \in B$ von f **höchstens einmal** getroffen wird, d.h. **höchstens ein** Urbild hat. Anders ausgedrückt verlangen wir, dass verschiedene Urbilder auch verschiedene Bilder haben. In Symbolen können wir schreiben*

$$x \neq y \in A \Rightarrow f(x) \neq f(y) \quad \textit{oder} \quad f(x) = f(y) \Rightarrow x = y.$$

(ii) *f heißt* surjektiv*, wenn **jedes** Element $b \in B$ von f **getroffen** wird, also **mindestens ein** Urbild besitzt. In Symbolen:*

$$\forall b \in B : \exists a \in A : f(a) = b.$$

(iii) *Wir nennen f* bijektiv*, wenn f injektiv und surjektiv ist. Das ist der Fall, wenn **jedes** Element in der Bildmenge B **genau ein** Urbild besitzt.*

Mitunter werden für die Begriffe *injektiv* und *bijektiv* auch die alten Begriffe *eindeutig* und *eineindeutig* verwendet. Das wäre ja leicht zu merken, doch unglücklicherweise verwenden manche Autoren den Begriff „eineindeutig“ statt für bijektiv für injektiv. Daher raten wir dringend zur Verwendung der lateinischen Bezeichnungen.

Ist $f : A \to B$ surjektiv, so sagt man auch f ist eine Abbildung von A **auf** B.

Wenn man Injektivität und Surjektivität von Abbildungen untersucht, ist es wichtig, nicht zu vergessen, Definitions- und Zielmenge genau zu beachten. Wenn wir etwa die Zuordnungsvorschrift $x \mapsto x^2$ untersuchen, dann können wir abhängig von Definitions- und Bildbereich alle Varianten finden, wie uns das folgende Beispiel zeigt.

Beispiel 4.3.10 (Injektiv, surjektiv, bijektiv)**.** *Wir bezeichnen mit $\mathbb{R}_0^+$ die Menge der nichtnegativen reellen Zahlen, d.h. $\mathbb{R}_0^+ := \{x \in \mathbb{R} \mid x \geq 0\}$.*

- *$x \mapsto x^2 : \mathbb{R} \to \mathbb{R}$ ist weder injektiv noch surjektiv, weil $f(-1) = 1 = f(1)$, was der Injektivität widerspricht und -1 nicht von f getroffen wird, was die Surjektivität ausschließt.*
- *$x \mapsto x^2 : \mathbb{R}_0^+ \to \mathbb{R}$ ist injektiv aber nicht surjektiv.*
- *$x \mapsto x^2 : \mathbb{R} \to \mathbb{R}_0^+$ ist surjektiv aber nicht injektiv.*
- *$x \mapsto x^2 : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ ist bijektiv.*

Ein einfaches aber wichtiges Beispiel für eine bijektive Abbildung ist für jede Menge M die **Identität** (auch **identische Abbildung** genannt), die jedem $m \in M$ wieder m zuordnet, d.h.

$$\begin{aligned} \mathbb{1}_M : M &\to M \\ \mathbb{1}_M(m) &= m. \end{aligned}$$

In vielen Texten wird die identische Abbildung auch mit id_M bezeichnet. Ist aus dem Zusammenhang klar, was die Menge M ist, so wird sie gerne notationell unterdrückt, d.h. die Identität wird mit $\mathbb{1}$ oder id bezeichnet.

Sind $f : A \to B$ und $g : B \to C$ zwei Abbildungen, so können wir diese hinter einander ausführen, indem wir das Ergebnis von f in g einsetzen: $g(f(a))$. Dies ist ein wichtiges Konzept.

Definition 4.3.11 (Verknüpfung von Abbildungen)**.** *Seien $f : A \to B$ und $g : B \to C$ zwei Abbildungen. Wir definieren die* Verknüpfung von f mit g *(oder die* Hintereinanderausführung von f und g *bzw. die* Komposition von f mit g*) in Zeichen, $g \circ f : A \to C$ durch*

$$(g \circ f)(a) := g(f(a)). \tag{4.3}$$

Achten Sie sorgfältig auf die Reihenfolge bei Verknüpfungen! Die Notation ist so gewählt, dass auf beiden Seiten von Gleichung (4.3) die Reihenfolge der Symbole g und f übereinstimmt. Das Symbol $g \circ f$ wird auch „g nach f“ gelesen, um klar zu stellen, dass zuerst die Funktion f und dann die Funktion g ausgeführt wird.

Sind $f : A \to B$, $g : B \to C$ und $h : C \to D$ drei Abbildungen, so gilt das **Assoziativgesetz** $(h \circ g) \circ f = h \circ (g \circ f)$ (dies folgt leicht aus der Definition). Man darf also beim Zusammensetzen von Abbildungen die Klammern weglassen.

Ist $f : A \to B$ bijektiv, so gibt es zu jedem Bild $b \in B$ genau ein Urbild $a \in A$ mit $f(a) = b$. Diese Tatsache können wir benutzen, um eine neue Funktion, die **Umkehrfunktion** von f zu konstruieren, die jedem $b = f(a)$ *das* Urbild a zuordnet. Die Surjektivität von f garantiert dabei, dass *jedem* b ein a zugeordnet werden kann, die Injektivität von f hingegen, dass die Zuordnung eindeutig ist.

Definition 4.3.12 (Umkehrfunktion)**.** *Sei $f : A \to B$ bijektiv. Die* inverse Abbildung *von f oder die* Umkehrfunktion *von f ist definiert durch*

$$\begin{aligned} f^{-1} : B &\to A \\ f(a) &\mapsto a. \end{aligned}$$

Verwechseln Sie nicht die Umkehrfunktion f^{-1}, die wir nur für bijektive Abbildungen f definiert haben mit der "Mengenabbildung" aus 4.3.7, die für beliebiges f existiert! Um die latent vorhandene Verwechslungsgefahr zu entschärfen wird das alleinstehende Symbol f^{-1} nur für die inverse Abbildung verwendet; für die "Mengenabbildung", die einer Teilmenge N des Bildbereichs ihr Urbild zuordnet wird das Symbol f^{-1} ausschließlich zusammen mit der Menge verwendet, also $f^{-1}(N)$. Ist $f : N \to N$ bijektiv, so könnte $f^{-1}(N)$ allerdings auch das Bild von N unter f^{-1} bezeichnen. Das ist aber kein Problem, da dieses mit dem Urbild von N unter f übereinstimmt. (Achtung: hier versteckt sich wieder eine „Mini-Behauptung"!) Bleibt alleine das Problem, dass $f^{-1}(b)$ für ein $b \in B$ einerseits das Bild von $b \in B$ unter der inversen Funktion f^{-1} und andererseits das Urbild von $\{b\} \subseteq B$ unter f (vgl. Definition 4.3.7(ii)) bezeichnet. Da f injektiv ist, ist letzteres die einelementige Menge $\{f^{-1}(b)\} \subseteq A$, sodass $f^{-1}(b)$ im ersten Fall ein Element von A bezeichnet im zweiten eine Teilmenge von A, die als einziges Element eben $f^{-1}(b)$ enthält; welche der beiden Möglichkeiten gemeint ist, wird aber immer aus dem Zusammenhang klar sein. Auch wenn Ihnen obige Bemerkung spitzfindig erscheint: wenn Sie sie verstehen, können Sie sicher sein, die Begriffe Bild, Urbild und Umkehrfunktion *gut* verstanden zu haben!

Bemerkung 4.3.13 (Umkehrabbildung)**.** *Die Zusammensetzung einer bijektiven Funktion f mit ihrer Umkehrabbildung f^{-1} ergibt für alle $a \in A$ und alle $b \in B$, wie man leicht einsehen kann*

$$f(f^{-1}(b)) = b, \qquad f^{-1}(f(a)) = a$$

oder in Funktionsnotation

$$f \circ f^{-1} = \mathbb{1}_B, \qquad f^{-1} \circ f = \mathbb{1}_A.$$

Ein erstes Beispiel für eine mathematische Struktur war diejenige einer Menge. Die zugehörigen Beziehungen sind die Abbildungen. Wir haben aber im letzten Abschnitt eine weitere, etwas spezialisierte Struktur definiert, die *geordnete Menge*. Was sind die Beziehungen zwischen geordneten Mengen? Ganz einfach: Diejenigen Abbildungen, die die Ordnungsstruktur erhalten, also die monotonen Abbildungen.

Definition 4.3.14 (Monotonie)**.** *Seien $(A, \preceq)$ und $(B, \trianglelefteq)$ zwei geordnete Mengen. Eine Abbildung $f : A \to B$ heißt* monoton wachsend*, falls aus $a \preceq b$ schon $f(a) \trianglelefteq f(b)$ folgt. Sie heißt* monoton fallend*, falls sich aus $a \preceq b$ die Relation $f(a) \trianglerighteq f(b)$ ergibt.*

Beispiel 4.3.15 (Monotonie)**.** *Die Funktion $f : \mathbb{R}_0^+ \to \mathbb{R}$, $x \mapsto x^2$ ist monoton wachsend.*

Wir haben also bereits zwei Beispiele für typische mathematische Strukturen kennengelernt: *Mengen und Abbildungen* und *geordnete Mengen und monotone Abbildungen*.

Hat man zwei Mengen A und B, so können wir alle Abbildungen von A nach B wieder zu einer Menge zusammenfassen, der **Menge aller Abbildungen von A nach B**, die oft mit B^A bezeichnet wird. Warum das so ist, können wir gleich sehen.

Zuletzt, sei nämlich wie versprochen noch die Definition des Cartesischen Produktes von endlich vielen Mengen auf beliebig viele Mengen verallgemeinert.

Definition 4.3.16 (Unendliches Produkt)**.** *Seien M_i, $i \in I$ Mengen. Wir definieren*

$$\prod_{i\in I} M_i := \bigtimes_{i\in I} M_i := \{f : I \to \bigcup_{i\in I} M_i \mid \forall i \in I : f(i) \in M_i\}$$

das Cartesische Produkt *der M_i.*

Sind alle Mengen $M_i = M$ gleich, dann schreiben wir statt $\prod_{i\in I} M$ auch M^I, und das stimmt mit der oberen Bezeichnung von M^I als Menge aller Abbildungen von I nach M überein!

Man beachte, dass diese Definition für endliche Indexmengen I äquivalent ist zur Definition mit k-tupeln. Haben wir etwa die Mengen M_0 und M_1, dann ist unsere Indexmenge $I = \{0, 1\}$. Setzen wir in die Definition 4.3.16 ein, so erhalten wir

$$M_0 \times M_1 = \{f : \{0, 1\} \to M_0 \cup M_1 \mid f(0) \in M_0 \wedge f(1) \in M_1\}. \tag{4.4}$$

Eine Abbildung f von $\{0, 1\}$ aus in irgendeine Menge ist schon eindeutig bestimmt durch die Werte bei 0 und 1. Die einzige Forderung an f ist, dass $f(0) \in M_0$ und $f(1) \in M_1$ liegen müssen. Verstehen wir nun die Indexmenge I als Positionsangaben und schreiben wir die Abbildung f ein wenig anders auf, dann sehen wir

$$\begin{array}{cccccc} I & 0 & & 1 & \\ (& f(0) & , & f(1) &), \\ & \in M_0 & & \in M_1 & \end{array}$$

dass jede Funktion einem geordneten Paar entspricht, dessen erster Eintrag in M_0 liegt, und dessen zweiter Eintrag Element von M_1 sein muss. Alle möglichen Funktionen die die Form von (4.4) haben, findet man, indem man $f(0)$ alle möglichen Elemente von M_0 durchlaufen lässt und für $f(1)$ jedes Element von M_1 einsetzt. Man konstruiert also wirklich alle geordneten Paare von M_0 und M_1.

Zum weiteren Verständnis wollen wir die Konstruktion für $I = \{0, 1, 2\}$ und $M_0 = \{a, b\}$, $M_1 = \{1, 2, 3\}$ und $M_2 = \{\alpha, \beta\}$ genau vorrechnen:

$$\begin{aligned} M_0 \times M_1 \times M_2 = \{&(a, 1, \alpha), (a, 1, \beta), (a, 2, \alpha), (a, 2, \beta), (a, 3, \alpha), (a, 3, \beta), \\ &(b, 1, \alpha), (b, 1, \beta), (b, 2, \alpha), (b, 2, \beta), (b, 3, \alpha), (b, 3, \beta)\} \end{aligned}$$

entspricht unserer ursprünglichen Definition durch Tripel (3-tupel).

Untersuchen wir die Menge aller Abbildungen

$$\begin{aligned} X := \{f : \{0, 1, 2\} \to M_0 \cup M_1 \cup M_2 = \{1, 2, 3, a, b, \alpha, \beta\} \mid& \\ f(0) \in \{a, b\} \wedge f(1) \in \{1, 2, 3\} \wedge f(2) \in \{\alpha, \beta\}\}.& \end{aligned} \tag{4.5}$$

Es gibt zwölf verschiedene Abbildungen in dieser Menge X:

$$\begin{array}{llll} f_0 : 0 \mapsto a & f_1 : 0 \mapsto a & f_2 : 0 \mapsto a & f_3 : 0 \mapsto a \\ \quad 1 \mapsto 1 & \quad 1 \mapsto 1 & \quad 1 \mapsto 2 & \quad 1 \mapsto 2 \\ \quad 2 \mapsto \alpha & \quad 2 \mapsto \beta & \quad 2 \mapsto \alpha & \quad 2 \mapsto \beta \\ \\ f_4 : 0 \mapsto a & f_5 : 0 \mapsto a & f_6 : 0 \mapsto b & f_7 : 0 \mapsto b \\ \quad 1 \mapsto 3 & \quad 1 \mapsto 3 & \quad 1 \mapsto 1 & \quad 1 \mapsto 1 \\ \quad 2 \mapsto \alpha & \quad 2 \mapsto \beta & \quad 2 \mapsto \alpha & \quad 2 \mapsto \beta \\ \\ f_8 : 0 \mapsto b & f_9 : 0 \mapsto b & f_{10} : 0 \mapsto b & f_{11} : 0 \mapsto b \\ \quad 1 \mapsto 2 & \quad 1 \mapsto 2 & \quad 1 \mapsto 3 & \quad 1 \mapsto 3 \\ \quad 2 \mapsto \alpha & \quad 2 \mapsto \beta & \quad 2 \mapsto \alpha & \quad 2 \mapsto \beta \end{array}$$

Sorgfältiger Vergleich zwischen den Mengen X und $M_0 \times M_1 \times M_2$ zeigt, dass in der Tat beide Mengen dasselbe beschreiben.

Beispiel 4.3.17. *Die Menge aller Abbildungen von $\mathbb{N}$ nach $\mathbb{R}$ (oder das Cartesische Produkt von „$\mathbb{N}$-vielen Kopien von $\mathbb{R}$“) ist die Menge aller reellen Zahlenfolgen*

$$(x_0, x_1, x_2, x_3, \ldots), \qquad \text{mit } x_i \in \mathbb{R} \text{ für } i \in \mathbb{N}.$$

4.4. Mächtigkeit

Eine interessante Eigenschaft von Mengen, die diesen intrinsisch ist, ist ihre **Mächtigkeit**. Wie fast alles in der Mengenlehre geht auch dieses Konzept auf Georg Cantor zurück. Für unendliche Mengen hat er als erster definiert, wann es legitim ist zu sagen, dass zwei Mengen A und B *gleich mächtig* („gleich groß“) sind.

Für endliche Mengen M, also Mengen, die n Stück Elemente (für ein $n \in \mathbb{N}$) enthalten ist die Mächtigkeit $|M|$ einfach die Anzahl ihrer Elemente. Für unendliche Mengen (also Mengen, die nicht endlich viele Elemente besitzen) müssen wir aber etwas trickreicher vorgehen.

Definition 4.4.1 (Gleichmächtigkeit)**.** *Zwei Mengen A und B heißen* gleich mächtig, *wenn eine bijektive Abbildung (eine Bijektion) von A auf B existiert. In diesem Fall sagt man auch A und B haben gleiche* Kardinalität *oder die gleiche* Kardinalzahl *und schreibt* $\mathrm{card}(A) = \mathrm{card}(B)$.

Diese einfache Definition hat weit reichende Konsequenzen. Es wird unter anderem möglich, dass eine Menge zu einer ihrer echten Teilmengen gleich mächtig ist.

Beispiel 4.4.2 (Gleichmächtigkeit von $\mathbb{N}$ und $\mathbb{N}_g$)**.** *Betrachten wir die Menge $\mathbb{N}$ und die Menge $\mathbb{N}_g$ aller geraden Zahlen. Es gilt $\mathbb{N}_g \subsetneqq \mathbb{N}$, doch die Abbildung $f : \mathbb{N} \to \mathbb{N}_g$ mit $f : x \mapsto 2x$ ist eine Bijektion. Die Mengen $\mathbb{N}$ und $\mathbb{N}_g$ sind also gleich mächtig.*

Es stellt sich heraus, dass nur die endlichen Mengen die Eigenschaft haben, eine größere Mächtigkeit zu besitzen als alle ihre echten Teilmengen.

Proposition 4.4.3 (Charakterisierung unendlicher Mengen)**.** *Eine Menge ist unendlich genau dann, wenn sie eine gleich mächtige echte Teilmenge besitzt.*

Beweis. Ohne Beweis. □

Schon Cantor hat gezeigt, dass aus der Mächtigkeitsdefinition gefolgert werden kann, dass unendlich große Mengen nicht gleich groß zu sein brauchen. **Es gibt auch bei unendlichen Menge Größenunterschiede**. In der Mengentheorie ist also „unendlich nicht gleich unendlich“.

Das Wort *unendlich* ist in der Mathematik allgegenwärtig. Die meisten vom Mathematiker behandelten Gegenstände sind unendlich (z.B. $\mathbb{N}$, $\mathbb{R}^n$, …), die meisten Aussagen in mathematischen Theorien handeln von unendlich vielen Objekten.

Das Symbol für den Ausdruck *unendlich* ist ∞. Dass es ein (und nur ein) Symbol für „unendlich“ gibt, führt leider oft zu Missverständnissen, wird doch von vielen daraus geschlossen, dass man mit unendlich so umgehen kann wie mit den reellen oder komplexen Zahlen.

Eine Menge M hat unendlich viele Elemente

Diese Aussage bedeutet, dass es keine natürliche Zahl n gibt mit $|M| = n$. Man schreibt abkürzend manchmal $|M| = \infty$. Es bezeichnet $|M|$ die Mächtigkeit (**Kardinalität**) von M, doch ∞ ist keine Kardinalzahl. Daher ist obige Formulierung *keine* mathematisch exakte Aussage.

Man verwendet ∞ bei der Beschreibung von Grenzübergängen wie etwa in

$$\lim_{n\to\infty} a_n$$

oder in

Für $n \to \infty$ strebt die Folge $(x_n)_n$ gegen x.

Auch hier ist ∞ nur eine *Abkürzung* für die ε-δ-Definition, die Sie in der Analysis noch sehr genau studieren werden. Dasselbe gilt für die Notation in unendlichen Reihen, z.B.

$$\sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Eine wirkliche mathematische Bedeutung hat das Symbol ∞ etwa in der Maßtheorie, in der die Menge $\bar{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ eingeführt wird. In diesem Fall bezeichnet ∞ ein bestimmtes von allen reellen Zahlen wohlunterschiedenes Element von $\bar{\mathbb{R}}$ mit genau definierten Eigenschaften. Auch in der projektiven Geometrie kommt das Symbol ∞ vor, und auch dort hat es eine genau festgelegte Bedeutung. In diesen Fällen ist ∞ keine Abkürzung mehr; dort hat es aber auch eine fixe Bedeutung frei von Mythen.

Ihnen wird vielleicht aufgefallen sein, dass wir nur definiert haben, was es für zwei Mengen bedeutet, die gleiche Kardinalzahl zu haben, aber nicht, was eine Kardinalzahl ist; dies geht über den Rahmen dieser Vorlesung hinaus. Wir befinden uns hier in etwa in der Situation eines Cowboys, der die Größe zweier Rinderherden vergleichen soll, ohne weit genug zählen zu können. Um festzustellen, ob die beiden Herden gleich viele Tiere enthalten braucht man diese aber nur paarweise durch ein Tor laufen zu lassen!

Was ist eigentlich die „kleinste“ unendliche Menge? Diese Frage läßt sich beantworten. Es kann relativ leicht gezeigt werden, dass jede unendliche Menge mindestens so groß wie $\mathbb{N}$ sein muss.

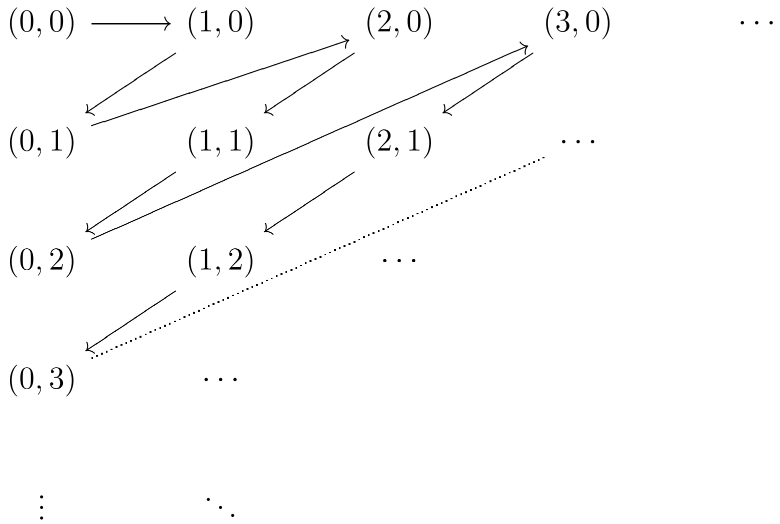
Heuristisch lässt sich das so begründen: Wenn wir $\mathbb{N}$ genauer untersuchen, dann erkennen wir folgende Eigenheit: In der natürlichen Ordnung von $\mathbb{N}$ besitzt jede Teilmenge $T \subset \mathbb{N}$ ein kleinstes Element (ein Minimum). Man sagt, die Menge $\mathbb{N}$ ist **wohlgeordnet**. Nun finden wir, dass für Teilmengen T von $\mathbb{N}$ nur zwei Möglichkeiten in Betracht kommen.

(1) Die Menge T ist nach oben beschränkt. Dann ist T endlich. Ist nämlich α eine obere Schranke von T, so ist T Teilmenge der endlichen Menge $\{0, 1, \ldots, \alpha\}$.
(2) Die Menge T ist nicht nach oben beschränkt. Dann kann man zu jedem Element t in T das nächst größere Element t' in T finden. Auf diese Weise kann man die Elemente von T durchnummerieren und eine Bijektion auf $\mathbb{N}$ konstruieren.

Also ist jede Teilmenge von $\mathbb{N}$ entweder endlich oder unendlich und genauso groß wie $\mathbb{N}$ selbst. „Zwischen“ den endlichen Mengen und $\mathbb{N}$ gibt es also keine Größenordnung mehr.

Um über die Mächtigkeit von $\mathbb{N}$ reden zu können, müssen wir ein neues Symbol einführen. Wir schreiben $|\mathbb{N}| =: \aleph_0$ (dieser Buchstabe stammt aus dem hebräischen Alphabet und heißt *Aleph*; die Mächtigkeit von $\mathbb{N}$ ist also Aleph-Null) und nennen jede Menge, die gleich mächtig mit $\mathbb{N}$ ist, also Kardinalität $\aleph_0$ hat, **abzählbar**.

Cantor hat bereits knapp vor der Jahrhundertwende bewiesen, dass $\mathbb{N} \times \mathbb{N}$ abzählbar ist. Dazu hat er das nach ihm benannte Diagonalverfahren verwendet, das unten dargestellt ist. Alle geordneten Paare in $\mathbb{N} \times \mathbb{N}$ werden in einem Rechteckschema angeschrieben und entlang der Pfeile der Reihe nach durchlaufen; so wird eine Bijektion auf $\mathbb{N}$ hergestellt.



Eine Formel für die Zuordnung ist

$$f : (i, j) \mapsto \tfrac{1}{2}(i + j)(i + j + 1) + j.$$

Nachdem die positiven resp. die negativen rationalen Zahlen $\mathbb{Q}$ als Teilmenge von $\mathbb{N} \times \mathbb{N}$ aufgefasst werden können ($q = \pm\frac{m}{n}$ für zwei teilerfremde natürliche Zahlen m und $n \neq 0$), ist also auch $\mathbb{Q}$ abzählbar. Auch die Vereinigung von abzählbar vielen abzählbaren Mengen ist wieder abzählbar. Das kann man mit Hilfe des gleichen Prinzips beweisen (schreibe in Gedanken alle Mengen untereinander auf und konstruiere die Bijektion analog zur Diagonalabzählung).

Cantor hat aber auch bewiesen, dass es verschiedene Größenklassen von Mengen gibt. So ist etwa die Potenzmenge $\mathbb{P}M$ einer Menge M **immer** mächtiger als M selbst (die Mächtigkeit der Potenzmenge einer Menge erfüllt $|\mathbb{P}M| = 2^{|M|}$).

Interessant ist, dass die reellen Zahlen $\mathbb{R}$ mächtiger sind als $\mathbb{N}$. Man sagt, die reellen Zahlen sind **überabzählbar**. Das hat ebenfalls Cantor gezeigt.

Genauer hat Cantor bewiesen, dass $]\,0, 1\,[$ überabzählbar ist. Die Tatsache, dass $]\,0, 1\,[$ die gleiche Mächtigkeit wie $\mathbb{R}$ hat, ist einfach zu zeigen. So bildet etwa die Funktion $\frac{1}{\pi}(\arctan(x) + \frac{\pi}{2})$ ganz $\mathbb{R}$ bijektiv auf $]\,0, 1\,[$ ab. Zum Beweis der Überabzählbarkeit von $]\,0, 1\,[$ verwenden wir die Tatsache, dass sich jede reelle Zahl r als Dezimalentwicklung aufschreiben lässt, und dann gehen wir indirekt vor. Angenommen, es gäbe eine Bijektion b von $\mathbb{N}$ auf $]\,0, 1\,[$. Dann stellen wir uns vor, dass wir alle Zahlen in der Reihenfolge untereinander schreiben wie sie durch die Bijektion auf $\mathbb{N}$ gegeben ist. Im nachfolgenden Diagramm mögen die a_{ij} für Dezimalziffern stehen. Die oberste Reihe repräsentiere die Dezimalentwicklung der ersten Zahl, die nächste Zeile die der zweiten, usw. Wäre $\mathbb{R}$ abzählbar, so müsste in diesem Schema *jede* reelle Zahl aus $]\,0, 1\,[$ irgendwo auftauchen.

$$\begin{array}{rcccccccc}
0: & 0, & \mathbf{a_{01}} & a_{02} & a_{03} & a_{04} & a_{05} & a_{06} & \cdots \\
1: & 0, & a_{11} & \mathbf{a_{12}} & a_{13} & a_{14} & a_{15} & a_{16} & \cdots \\
2: & 0, & a_{21} & a_{22} & \mathbf{a_{23}} & a_{24} & a_{25} & a_{26} & \cdots \\
3: & 0, & a_{31} & a_{32} & a_{33} & \mathbf{a_{34}} & a_{35} & a_{36} & \cdots \\
4: & 0, & a_{41} & a_{42} & a_{43} & a_{44} & \mathbf{a_{45}} & a_{46} & \cdots \\
5: & 0, & a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & \mathbf{a_{56}} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

Betrachten wir jetzt die reelle Zahl r mit der Dezimalentwicklung

$$r = 0, \widehat{a_{01}}\ \widehat{a_{12}}\ \widehat{a_{23}}\ \widehat{a_{34}}\ \widehat{a_{45}}\ \widehat{a_{56}} \dots \widehat{a_{n\,n+1}} \dots,$$

wobei wir $\widehat{a_{ij}}$ definieren durch

$$\widehat{a_{ij}} := \begin{cases} a_{ij} + 2 & \text{falls } a_{ij} \leq 4 \\ a_{ij} - 2 & \text{falls } a_{ij} \geq 5 \end{cases}$$

Versuchen wir nun herauszufinden, an welcher Stelle r in der Liste eingetragen ist, so müssen wir feststellen, dass r gar nicht in der Aufzählung enthalten sein kann. Sei nämlich n diejenige natürliche Zahl mit $b(n) = r$. Dann gilt aber

$$\begin{aligned} b(n) &= 0,\ a_{n1}\ a_{n2}\ a_{n3}\ a_{n4}\ a_{n5}\ a_{n6} \dots \\ r &= 0,\ \widehat{a_{01}}\ \widehat{a_{12}}\ \widehat{a_{23}}\ \widehat{a_{34}}\ \widehat{a_{45}}\ \widehat{a_{56}} \dots \end{aligned}$$

Damit wirklich $b(n) = r$ gilt, müssen die Dezimalentwicklungen von $b(n)$ und r überein stimmen. Es gilt aber $\widehat{a_{n,n+1}} \neq a_{n,n+1}$. Daher sind $b(n)$ und r verschieden, und r war tatsächlich nicht in der Liste enthalten.

Genauer untersuchend sieht man, dass $\mathbb{R}$ gleich mächtig ist mit der Potenzmenge von $\mathbb{N}$. Man könnte nun vermuten, dass $\mathbb{R}$ die nächst höhere Mächtigkeit nach $\aleph_0$ besitzt, also $\aleph_1$.

Trotzdem bezeichnet man aus gutem Grund die Mächtigkeit von $\mathbb{R}$ mit $|\mathbb{R}| = c$, der sogenannten Mächtigkeit des Kontinuums. Es lässt sich nämlich nicht $c = \aleph_1$ beweisen (**man kann beweisen, dass sich das nicht beweisen lässt** — das hat Kurt Gödel 1938 getan), es lässt sich übrigens auch nicht widerlegen (das hat Paul J. Cohen (geb. 1934) 1963 **bewiesen**). Die sogenannte **Kontinuumshypothese** von Georg Cantor, dass $c = \aleph_1$ ist, ist **unabhängig von den Axiomen der Mengenlehre**. Das heißt, es gibt *Modelle* der axiomatischen Mengenlehre, in denen $c = \aleph_1$ gilt und andere *Modelle*, in denen $c \neq \aleph_1$ zutrifft. Die Axiomatisierung des Mengenbegriffs bringt solche unangenehme Fakten mit sich, die zeigen, dass es noch nicht geschafft wurde, den naiven Mengenbegriff so gut zu axiomatisieren, dass die Axiome unsere Vorstellungswelt ganz einzufangen im Stande sind.

4.5. Axiomatische Mengenlehre

Eine Möglichkeit, die Mathematik auf ein festes Fundament zu stellen, ist die Axiomatisierung der Mengenlehre nach Zermelo und Fraenkel. Mit der Festlegung dieser *Axiome* gibt man ihr einen Satz von Grundaussagen. Aus diesen werden dann die mathematischen Theoreme abgeleitet, auf diesen Fundamenten wird das Gebäude der Mathematik entwickelt — theoretisch jedenfalls.

Der Ursprung der axiomatischen Mengenlehre liegt in den Paradoxien, die die naive Mengenlehre um die Jahrhundertwende geplagt haben, wie etwa die Russellsche Antinomie (siehe Seite 43). Sie wurde 1908 von Zermelo erfunden und hat bald eine große Bedeutung gewonnen. Die Mengenlehre ist die Basis für beinahe die gesamte Mathematik, und ihre Axiomatisierung erlaubt es, diese Basis einwandfrei zu legen.

Es gibt mehrere verschiedene Axiomensysteme, die alle die naive Mengenlehre präzisieren aber untereinander fundamentale Unterschiede aufweisen. Wir präsentieren hier die Axiome von Zermelo und Fraenkel (ZFC), etwa im Gegensatz zu den Systemen von Neumann-Bernays-Gödel oder Morse-Kelley, auch weil die Einführung von *Klassen* dadurch vermieden werden kann.

Grundlage für die Axiomatisierung der Mengenlehre ist die Logik, und obwohl man auch die Theorie der Aussagen (Aussagenlogik, Prädikatenlogik) formal exakt machen könnte, werden wir hier stoppen und die logischen Grundlagen naiv verwenden. Es sei nur festgehalten, dass *alle* auftretenden Zeichen Bedeutung in der Logik haben (auch =) mit der einzigen

Ausnahme $\in$, und dass φ und ψ beliebige Formeln bezeichnen, deren Variable in Klammern angegeben werden.

Mit Hilfe der ersten sechs ZFC Axiome kann die gesamte *endliche* Mathematik konstruiert werden. Sie lauten wie folgt:

(ZF1)	$\exists x : (x = x)$	(Existenz)
(ZF2)	$\forall x : \forall y : \forall z : ((z \in x \Leftrightarrow z \in y) \Rightarrow x = y)$	(Extensionalität)
(ZF3)	$\forall U : \forall p : \exists Z : \forall x : (x \in Z \Leftrightarrow (x \in U \wedge \varphi(x,p)))$	(Separation)
(ZF4)	$\forall x : \forall y : \exists Z : (x \in Z \wedge y \in Z)$	(Paare)
(ZF5)	$\forall \mathcal{F} : \exists Z : \forall F : \forall x : ((x \in F \wedge F \in \mathcal{F}) \Rightarrow x \in Z)$	(Vereinigung)
(ZF6)	$\forall U : \exists \mathcal{Z} : \forall Y : (\forall x : (x \in Y \Rightarrow x \in U) \Rightarrow Y \in \mathcal{Z})$	(Potenzmenge)

Wegen des hohen Abstraktionsgrades dieser Axiome ist ein wenig Erklärung von Nöten, und außerdem müssen wir einige Abkürzungen einführen. Das Axiom ZF1 stellt sicher, dass Mengen existieren, und ZF2 erklärt, dass zwei Mengen genau dann gleich sind, wenn sie dieselben Elemente haben. Mit Hilfe von ZF3 wird das erste Konstruktionsprizip für neue Mengen eingeführt, die Auswahl einer Teilmenge Z aus einer *gegebenen* Menge U mit Hilfe einer „Auswahlregel" φ. Für diese Menge Z führen wir die Abkürzung $\{x \in U \mid \varphi(x)\}$ ein. Weitere Abkürzungen seien die Formulierungen $\forall x \in U$, die für $\forall x : x \in U$ stehe, und $\exists x \in U$ für $\exists x : x \in U$. ZF3 besagt in gewisser Art und Weise, dass man für *jedes* Element einer Menge überprüfen kann, ob es eine bestimmte Eigenschaft φ aufweist oder nicht. Das ist natürlich nur theoretisch möglich, weshalb dies von E. Bishop in [**Bishop 1967**] als *Prinzip der Allwissenheit* bezeichnet wurde.

Aus ZF4 definieren wir $\{x,y\} := \{z \in Z \mid z = x \vee z = y\}$ und $\{x\} := \{x,x\}$. Das Vereinigungs-Axiom ZF5 ermöglicht es uns mit Hilfe von $\mathcal{F} = \{X,Y\}$ zu definieren

$$X \cup Y := \{z \in Z \mid z \in Z \vee z \in Y\}.$$

Drei weitere Symbole müssen wir einführen, um die weiteren Axiome formulieren zu können. Es sind dies das *Leere Menge-Symbol* $\emptyset := \{z \in Z \mid \neg(z = z)\}$ für eine fixe Menge Z und $S(x) := x \cup \{x\}$. Schließlich erklären wir das (uns bereits naiv bekannte) Symbol $\exists!$ durch folgende Abkürzungsvereinbarung

$$\exists! y : \varphi(y) \text{ entspreche } \exists y : \varphi(y) \wedge (\forall y : \forall x : (\varphi(y) \wedge \varphi(x)) \Rightarrow x = y).$$

Die drei nächsten Axiome sind dann:

(ZF7)	$\exists Z : \forall X : (\emptyset \in Z \wedge (X \in Z \Rightarrow S(X) \in Z))$	(Unendlichkeit)
(ZF8)	$\forall U : \forall p : \big(\forall x \in U : \exists! z : \varphi(x,z,U,p) \Rightarrow$ $\exists Z : \forall x \in U : \exists z \in Z : \varphi(x,z,U,p)\big)$	(Ersetzung)
(ZF9)	$\forall x : \big(\neg(x = \emptyset) \Rightarrow \exists y : (y \in x \wedge \neg \exists z : (z \in x \wedge z \in y))\big)$	(Fundierung)

Hier ist wieder einiges an Erläuterungen von Nöten. ZF7 garantiert die Existenz einer Menge mit den Elementen $\emptyset, S(\emptyset), S(S(\emptyset)), \ldots$. Diese scheinbar schräge Konstruktion wird aber sofort verständlicher, wenn man die Bezeichungen $0 := \emptyset$, $1 := S(\emptyset)$, $2 := S(S(\emptyset))$, und allgemein $n + 1 := S(n)$ einführt.

ZF8 hat die komplexeste Formel, doch dieses Axiom stellt nichts anderes sicher als dass man aus einer Menge U und einer Zuordnung f, die jeder Menge $x \in U$ eine Menge y zuordnet, eine weitere Menge als Bild von U unter f konstruieren kann. Dieses Axiom rechtfertigt auch die Abkürzung $\{f(x) \mid x \in U\}$ für die Definition einer Menge.

Das Fundierungsaxiom ZF9 zu guter Letzt schließt unter anderem die Russellsche Antinomie aus zusammen mit allen Mengen, die in gewissem Sinne „zu groß" sind. Es werden alle Mengen verboten, die sich selbst enthalten oder aber Mengen enthalten, die wiederum andere Mengen enthalten, und so weiter ad infinitum.

Das letzte Axiom von ZFC hat in der Vergangenheit viele Kontroversen verursacht, da es dem Mathematiker gestattet, auf nicht konstruktivem Weg neue Mengen zu definieren.

Analog zum Prinzip der Allwissenheit könnte man das Axiom auch wie J. Cigler und H.C. Reichel in [**Cigler, Reichel 1987**] als das *Prinzip der Allmächtigkeit* bezeichnen. Heute akzeptiert ein überwiegender Teil der Mathematiker dieses Axiom auf Grund seiner Verwendbarkeit und der Vielfalt praktischer Theoreme, die zu diesem Axiom äquivalent sind. Zuvor wir das Axiom aber anführen benötigen wir eine weitere Abkürzung

$$F \cap G := \{z \in F \cup G \mid z \in F \wedge z \in G\}.$$

Das zehnte Axiom, das Auswahlaxiom, ist

$$\textbf{(ZF10)} \quad \forall \mathcal{F} : \big(\forall H \in \mathcal{F} : \neg(H = \emptyset) \wedge \forall F \in \mathcal{F} : \forall G \in \mathcal{F} : (F = G \vee F \cap G = \emptyset) \\ \Rightarrow \exists S : \forall F \in \mathcal{F} : \exists! s(s \in S \wedge s \in F)\big) \qquad \text{(Auswahl)}$$

Es besagt, dass es zu jeder gegebenen Familie von nichtleeren, paarweise disjunkten Mengen M_i, $i \in I$ eine weitere Menge gibt, die aus jedem M_i genau ein Element enthält.

Diese axiomatische Einführung der Mengen ist natürlich nicht umfassend; im Rahmen dieses Buchs können wir nur einen kurzen Einblick in die Fundamente der Mathematik geben. Weiterführende Informationen finden Sie in Vorlesungen oder Büchern zu den Themenkreisen „Grundbegriffe der Mathematik“ und „Axiomatische Mengenlehre“.

KAPITEL 5

Grundlegende Algebra

In diesem Kapitel widmen wir uns dem Ausbau mathematischer Strukturen. Die hier definierten Gruppen, Ringe und besonders die Körper bilden die Grundlage für die Theorien in Lineare Algebra und Analysis.

Einige für AnfängerInnen zunächst weniger wichtige Begriffe sind als Erweiterungsstoff gekennzeichnet; der Text kann sowohl mit als auch ohne den doppelt gestrichenen Teilen *fortlaufend* gelesen werden.

Schon in der Zeit der Antike haben in Griechenland berühmte Mathematiker gewirkt. Euklid (ca. 325–265 v.Chr.) (*euklidische* Geometrie, *euklidische* Räume) ist heute vor allem bekannt für sein Werk „Die Elemente" (13 Bücher), das das erste bekannte mathematische Lehrwerk ist, in dem Axiome, Theoreme und Beweise in klarer Abfolge vorkommen und das auf rigorosen Umgang mit der Mathematik abzielt. Es enthält unter anderem Aussagen über ebene und räumliche Geometrie (etwa die Platonischen Körper), Zahlentheorie (z.B. den euklidischen Algorithmus), rationale und irrationale Zahlen. Etwa fünfhundert Jahre später schrieb Diophantus von Alexandria (ca. 200–284) (*diophantische* Gleichung, *diophantische* Approximation) neben anderen Büchern sein 13-bändiges Werk „Arithmetica", von dessen Name unser heutiges „Arithmetik" abgeleitet ist. In diesem machte er als erster einen Schritt in Richtung moderner Algebra. Er studierte lineare und quadratische Gleichungen sowie zahlentheoretische Probleme. Da aber zu seiner Zeit die Null noch nicht erfunden war und das Konzept negativer Zahlen noch in weiter Ferne lag, war die Behandlung dieser Gleichungen noch auf Fallunterscheidungen angewiesen. Darüber hinaus erschienen ihm einige dieser Gleichungen als sinnlos, etwa $4 = 4x + 20$, weil sie keine (d.h. negative) Lösungen hatten. Auch das „Buchstabenrechnen" hatte er noch nicht eingeführt, und es gab noch kein praktisches Zahlensystem. Alle Theoreme und Rechnungen wurden in Worten präsentiert.

Weitere fünfhundert Jahre später verfasste der arabische Mathematiker Abu Abd Allah Mohammed Ibn Musa Al-Khwarizmi (ca. 780–850), Hofmathematiker in Bagdad, sein Hauptwerk „al-kitab almukhtamar fi hisab **al-jabr** wa'l-muqabala", zu deutsch „Kurzgefasstes Buch über das Rechnen durch Vervollständigen und Ausgleichen". Ein weiterer Meilenstein in der Mathematik (nicht primär im Inhalt aber bestimmt in der Wirkung), beschreibt dieses Buch die vollständige Behandlung der linearen und quadratischen Gleichungen, auch die negativen Fälle, beide Lösungen, aber noch immer ohne die Verwendung von Null und negativen Zahlen. Auch das Rechnen mit Buchstaben wurde zu dieser Zeit noch nicht erfunden. Allerdings wurden zum ersten Mal detaillierte Rechenschritte zur Lösung mathematischer Probleme angegeben. Außerdem wurde das hinduistische Zahlensystem (die heutigen arabischen Zahlen) mit den Ziffern 0 bis 9 und den Dezimalstellen zum ersten Mal ausführlich erklärt. Im zwölften Jahrhundert wurde es in Latein übersetzt und beginnt dort mit den Worten „Dixit Algoritmi" (Al-Khwarizmi hat gesagt). Aus dieser Lateinisierung des Herkunftsnamens von Al-Khwarizmi (Khwarizm, das heutige Khiva südlich des Aralsees in Usbekistan und Turkmenistan) wird übrigens das Wort *Algorithmus* für das schrittweise Lösen mathematischer Probleme abgeleitet. Teile des arabischen Titels, besonders das **al-jabr**, wurden auch in späteren Büchern arabischer Mathematiker verwendet, und so wurde über viele Zwischenstufen aus dem arabischen al-jabr (Auffüllen, Vervollständigen) das moderne Wort *Algebra*.

Heute wird unter Algebra vor allem die mathematische Theorie von Strukturen verstanden, und was das genau ist, wollen wir uns in den nächsten Abschnitten genauer ansehen.

5.1. Motivation

Alle hier zu besprechenden Strukturen basieren auf dem Mengenkonzept. Es sind **Mengen zusammen mit Abbildungen**, die bestimmte Eigenschaften aufweisen. Wir beginnen mit einigen Beispielen.

Beispiel 5.1.1.

Hauptwörter: *Sei W die Menge aller Hauptwörter der deutschen Sprache. Wählt man zwei Wörter aus W, dann kann man (meist) durch (fast bloßes) Hintereinandersetzen ein weiteres Wort aus W erzeugen. Wir können etwa aus „Leiter“ und „Sprosse“ das Wort „Leitersprosse“ bilden. Auch „Dampf“ und „Schiff“ lassen sich zu „Dampfschiff“ verbinden, „Schiff“ und „Kapitän“ ergeben „Schiffskapitän“.*

Strichblöcke: *Sei S die Menge aller Strichblöcke. Ein Strichblock ist einfach eine Ansammlung hintereinander geschriebener gleich langer Striche:*

$$s = |||||||||||||||$$

Fügen wir zwei Strichblöcke aneinander, dann erhalten wir wieder einen (längeren) Strichblock.

Translationen: *Sei T die Menge aller Möglichkeiten, ein Objekt im dreidimensionalen Raum geradlinig zu verschieben, also die Menge der* Translationen. *Bei der Betrachtung solcher Verschiebungen können wir uns auf deren Richtung und Länge beschränken. Zusammen mit der Position des Objekts vor der Translation ist es uns dann leicht möglich, seine Endposition zu bestimmen. Verschieben wir ein Objekt zweimal, so hätten wir dieselbe Endposition auch mit einer einzigen Translation erreichen können. Das Hintereinander-Ausführen von Translationen ist also wieder eine Translation.*

Drehungen: *Betrachten wir wieder einen Gegenstand im dreidimensionalen Raum. Wir wählen eine beliebige Gerade g, die durch seinen Schwerpunkt geht. Dann geben wir uns einen Winkel φ vor und verdrehen das Objekt bezüglich der Drehachse g um den Winkel φ. Die Menge aller dieser Drehungen sei D. Wie bei den Translationen ergibt das Hintereinander-Ausführen zweier Drehungen wieder eine Drehung.*

Abbildungen: *Sei $M^M = \mathrm{Abb}(M)$ die Menge aller Abbildungen von M nach M. Die Hintereinander-Ausführung $\circ$ von Abbildungen (vgl. 4.3.11) ist eine Verknüpfung auf $\mathrm{Abb}(M)$.*

Zahlenmengen:

- **$\mathbb{N}$:** *Wenn wir zwei natürliche Zahlen addieren oder multiplizieren, erhalten wir wieder eine natürliche Zahl.*
- **$\mathbb{Z}$:** *Auch das Produkt und die Summe zweier ganzer Zahlen ist eine ganze Zahl.*
- **$\mathbb{R}$:** *Auch reelle Zahlen können wir addieren und multiplizieren, um eine neue reelle Zahl zu berechnen.*

Matrizen:

Addition: *Sei $M_2(\mathbb{R})$ die Menge aller 2×2–Matrizen reeller Zahlen. Eine 2×2–Matrix ist dabei ein kleines Zahlenquadrat der Form (vgl. 2.2.1)*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

aus Zahlen $a_{ij} \in \mathbb{R}$. *Wir definieren die Summe zweier Matrizen komponentenweise, d.h.*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} := \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix}$$

und erhalten wieder eine 2×2*–Matrix.*

Multiplikation: *Auf* $M_2(\mathbb{R})$ *kann man auch ein Produkt einführen, das zwei Matrizen eine weitere Matrix zuordnet. Die Definition ist nicht-trivial und lautet*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} := \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}.$$

Visualisieren lässt sich die Verknüpfung an Hand der grauen Pfeile in Ab-

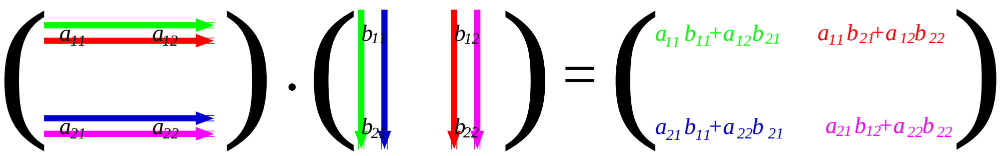


Abbildung 5.1. Multiplikation von Matrizen

bildung 5.1. Um etwa den hellsten Eintrag in der Ergebnismatrix zu erhalten, wandert man die hellsten Pfeile in den beiden Faktoren entlang. Dann berechnet man das Produkt der ersten Zahl links mit der ersten Zahl im Pfeil rechts, das der zweiten Zahl im linken Pfeil mit der zweiten Zahl im rechten Pfeil und summiert die Ergebnisse.

Gleitkommazahlen: *Sei* FP_2 *die Menge aller rationalen Zahlen, die sich schreiben lassen als* $\pm 0.z_1z_2 \cdot 10^n$ *mit Ziffern* z_1 *und* z_2 *und ganzzahligem Exponenten* n*. Diese Zahlen heißen auch dezimale Gleitkommazahlen mit zwei signifikanten Stellen. Addieren wir zwei solche Zahlen, erhalten wir wieder eine rationale Zahl. Diese Zahl lässt sich aber meist nicht in der obigen Form schreiben:*

$$0.23 + 4.5 = 4.73.$$

Wir zwingen das Ergebnis nun in die Gleitkommaform, indem wir runden*. Dann wird*

$$0.23 + 4.5 = 4.73 \approx 4.7$$

und wir definieren eine „veränderte“ Addition $\oplus$*, die* Addition mit Runden*, sodass*

$$0.23 \oplus 4.5 = 4.7.$$

Damit ergibt die $\oplus$*-Summe zweier Elemente von* FP_2 *wieder eine Gleitkommazahl mit zwei signifikanten Stellen.*

Alle diese Beispiele haben eines gemeinsam. Wir starten mit einer Menge M und einer Methode, wie wir aus zwei Elementen von M ein weiteres Element von M erzeugen. In der nächsten Definition schälen wir diese Struktur heraus.

Definition 5.1.2 (Gruppoid, algebraische Struktur)**.** *Sei G eine nichtleere Menge.*

(i) *Eine* Verknüpfung *auf G ist eine Abbildung*

$$\circ : G \times G \to G.$$

An Stelle von $\circ(g,h)$ für zwei Elemente $g, h \in M$ schreiben wir $g \circ h$, und wir nennen das Bild von (g,h) das Ergebnis *der Verknüpfung.*

(ii) *Wenn wir die Menge G zusammen mit der Verknüpfung $\circ$ untersuchen, so schreiben wir meist $(G, \circ)$ und nennen sie* Gruppoid *oder* algebraische Struktur *(oder* Magma*). In diesem Zusammenhang nennen wir G auch* Grundmenge*.*

Es sind also alle im Beispiele 5.1.1 betrachteten Mengen mit den entsprechenden Abbildungen Gruppoide.

Die Stärke, die in dieser und ähnlichen Definitionen von Strukturen liegt, ist dass man die Eigenschaften der Struktur und Konsequenzen aus diesen Eigenschaften unabhängig vom tatsächlichen Beispiel untersuchen kann. Die Ergebnisse dieser Untersuchung lassen sich dann auf alle zu dieser Struktur passenden Beispiele anwenden und erlauben es dadurch auf sehr elegantem Wege neue Erkenntnissen über die Beispiele zu gewinnen.

Verknüpfungen von Elementen werden meist mit Symbolen bezeichnet. Typische Symbole sind $\circ$, $+$, $\cdot$, $*$, $\oplus$, $\otimes$, $\boxdot$, $\circledast$,...

Betrachten wir Mengen mit mehr als einer Verknüpfung, so nehmen wir auch die anderen Verknüpfungssymbole in die Bezeichnung auf, z.B. $(B, \wedge, \vee)$.

Wird die Verknüpfung mit $\circ$ oder mit $\cdot$ bezeichnet, so lässt man das Verknüpfungssymbol meist weg, sofern keine Mehrdeutigkeiten bestehen. Man schreibt dann statt $g \circ h$ einfach gh. Kommen $\circ$ und $\cdot$ vor, so lässt man (meist) $\cdot$ weg. Z.B. schreibt man $(g \circ h)k$ statt $(g \circ h) \cdot k$. Falsch wäre $(gh) \cdot k$.

Alle Strukturen, die wir in diesem Abschnitt kennen lernen werden bauen aufeinander und insbesondere auf Definition 5.1.2 auf.

Je mehr Eigenschaften eine Struktur aufweist, um so spezieller ist sie. Umgekehrt kann man aus einer spezielleren Struktur immer eine allgemeinere machen, indem man die Eigenschaften, die „zuviel" sind, einfach *vergisst*. So ist etwa jede Gruppe (siehe Definition 5.2.18) auch eine Halbgruppe (siehe Definition 5.2.2). Die Abbildung 5.2 gibt ein grobes Diagramm der Strukturhierarchie wie wir sie in diesem Abschnitt kennen lernen werden. In dieser Abbildung sehen wir, dass zusätzlich geforderte Eigenschaften, jeweils angedeutet durch ein Rechteck, die Menge der passenden Strukturen einschränken. Es gilt aber immer, dass speziellere Strukturen eben speziellere Varianten von weniger speziellen (d.h. allgemeineren) Strukturen sind. So ist, wie in diesem Bild zu sehen ist, jeder Körper auch ein Ring und jeder Ring auch eine kommutative Gruppe und erst recht ein Gruppoid.

5.2. Gruppen

In diesem Abschnitt wollen wir uns zunächst auf Mengen zusammen mit einer Verknüpfung beschränken.

Beispiel 5.2.1 (Assoziativität)**.**

$(W, \circ)$**:** *Sei $(W, \circ)$ die Menge aller Hauptwörter der deutschen Sprache mit dem Hintereinandersetzen als Verknüpfung. Man kann natürlich auch zusammengesetzte Hauptwörter mit weiteren Wörtern verknüpfen und dadurch längere (mehrfach) zusammengesetzte Hauptwörter konstruieren. „Dampf" und „Schiffskapitän" liefern etwa „Dampfschiffskapitän". Wenig überraschend setzen sich auch „Dampfschiff" und „Kapitän" zu „Dampfschiffskapitän" zusammen. Wir sehen also, dass das Ergebnis beim Hintereinandersetzen von „Dampf", „Schiff" und „Kapitän" das Wort*

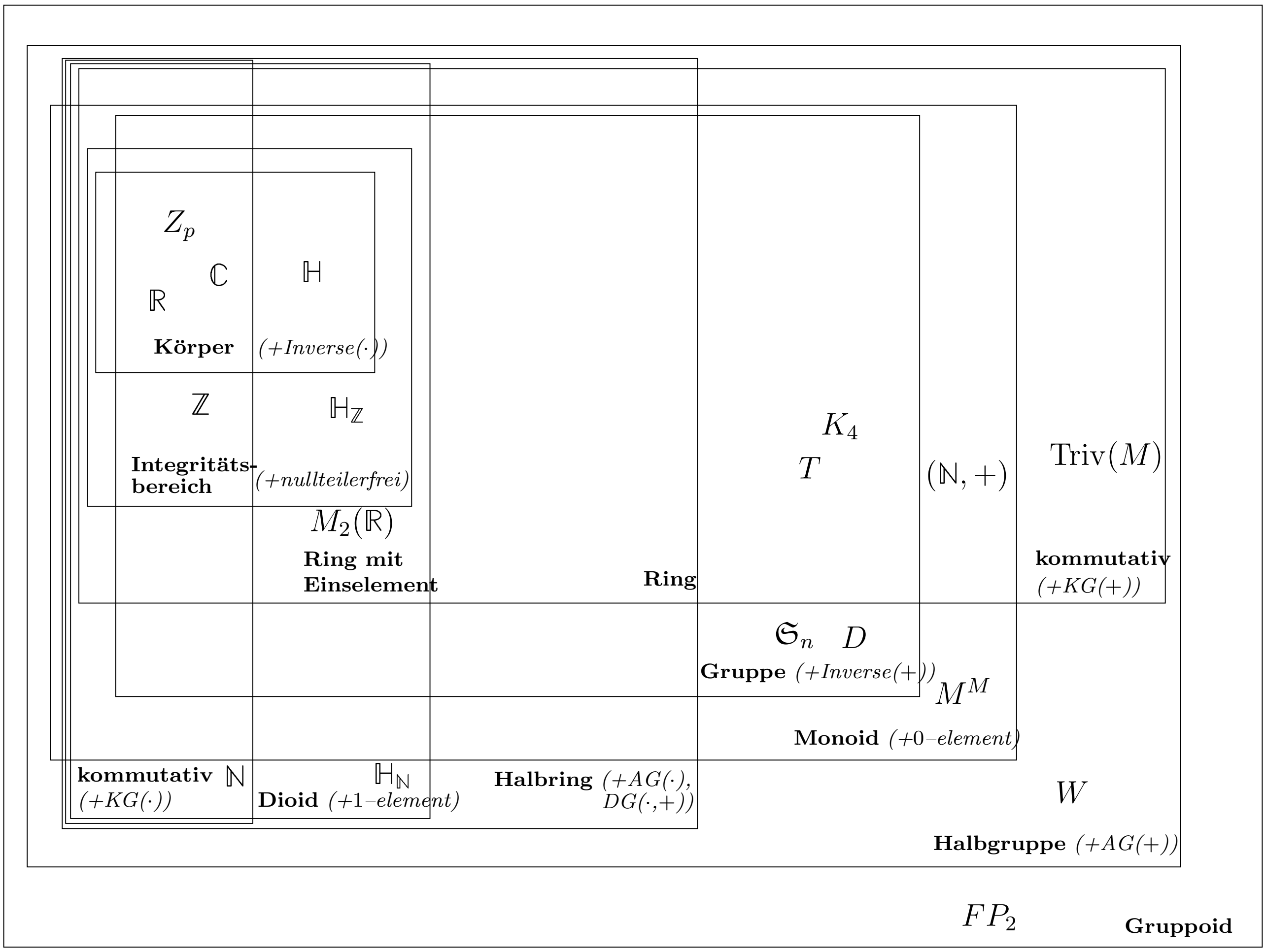


Abbildung 5.2. Hierarchie einiger algebraischer Strukturen

„Dampfschiffskapitän“ ergibt und das unabhängig von der Reihenfolge des Zusammensetzens.

$(S, \circ)$**:** *Bezeichnen wir mit* $(S, \circ)$ *das Gruppoid der Strichblöcke mit der Zusammensetzung als Verknüpfung. Beim Zusammensetzen von drei Strichblöcken kommt es nicht darauf an, ob zuerst die ersten beiden zusammengefasst werden und danach der dritte hinzugefügt wird, oder ob zuerst die beiden hinteren verknüpft werden und danach der erste Strichblock daran gehängt wird.*

$(T, \circ)$, $(D, \circ)$**:** *Ebenso verhält sich die Verknüpfung von drei Translationen oder Drehungen.*

$(\mathrm{Abb}(M), \circ)$**:** *Allgemein ist das Hintereinander-Ausführen von Abbildungen assoziativ (das haben wir schon in Abschnitt 4.3 beobachtet).*

$(\mathbb{N}, +)$, $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{N}, \cdot)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{R}, \cdot)$**:** *Auch bei der Addition natürlicher, ganzer und reeller Zahlen sowie bei der Multiplikation dieser macht es keinen Unterschied, welche einer Reihe von Verknüpfungen zuerst ausgeführt wird.*

$(M_2(\mathbb{R}),+)$: *Für 2×2–Matrizen überprüfen wir, ob $+$ diese Eigenschaft auch besitzt. Nehmen wir Elemente A, B und C aus $(M_2(\mathbb{R}),+)$. Dann finden wir*

$$\left(\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} + \begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix}\right) + \begin{pmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{pmatrix} =$$

$$\begin{pmatrix}a_{11}+b_{11} & a_{12}+b_{12}\\ a_{21}+b_{21} & a_{22}+b_{22}\end{pmatrix} + \begin{pmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{pmatrix} = \begin{pmatrix}a_{11}+b_{11}+c_{11} & a_{12}+b_{12}+c_{12}\\ a_{21}+b_{21}+c_{21} & a_{22}+b_{22}+c_{22}\end{pmatrix} =$$

$$\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} + \begin{pmatrix}b_{11}+c_{11} & b_{12}+c_{12}\\ b_{21}+c_{21} & b_{22}+c_{22}\end{pmatrix} = \begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix} + \left(\begin{pmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{pmatrix} + \begin{pmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{pmatrix}\right)$$

$(M_2(\mathbb{R}),\cdot)$: *Auch in $(M_2(\mathbb{R}),\cdot)$ verhält es sich ähnlich.*

$(\mathbb{FP}_2,\oplus)$: *Nun zum letzten Beispiel. Wir betrachten die Menge $(\mathbb{FP}_2,\oplus)$ der Gleitkommazahlen mit zwei signifikanten Stellen und der Addition mit Runden. In diesem Fall kommt es sehr wohl auf die Reihenfolge der Verknüpfungen an, denn*

$$(0.47 \oplus 0.57) \oplus 0.88 = 1.0 \oplus 0.88 = 1.9$$
$$0.47 \oplus (0.57 \oplus 0.88) = 0.47 \oplus 1.5 = 2.0$$

liefert verschiedene Resultate.

Wir erkennen also: Die Eigenschaft, dass man auf die genaue Festlegung der Verknüpfungsreihenfolge verzichten kann, ist zwar (sehr) oft aber nicht immer erfüllt. Darum führen wir für solche speziellere Strukturen einen neuen Begriff ein.

Definition 5.2.2 (Assoziativgesetz, Halbgruppe). *Ein Gruppoid $(G,\circ)$ heißt* Halbgruppe*, falls die Verknüpfung* assoziativ *ist, also das* Assoziativgesetz

$$\textbf{(AG)} \quad \forall g,h,k \in G : (g \circ h) \circ k = g \circ (h \circ k)$$

gilt. In diesem Fall ist das Setzen von Klammern nicht notwendig, und wir dürfen an Stelle von $(g \circ h) \circ k$ einfach $g \circ h \circ k$ schreiben.

Beispiel 5.2.3 (Halbgruppen).

(i) *Wie schon erwartet bilden die Mengen W bis $M_2(\mathbb{R})$ mit den in Beispiel 5.1.1 definierten Verknüpfungen Halbgruppen.*

(ii) Keine *Halbgruppe ist etwa die Menge $(FP_2,\oplus)$ der Gleitkommazahlen mit zwei signifikanten Stellen mit der Addition mit Runden.*

(iii) *Immer nach der Einführung einer Struktur kann man untersuchen, welche Objekte diese Struktur beschreibt. Meist kann man schnell sehr einfach gebaute Objekte finden, die dazu passen. Abgesehen von der Halbgruppe, die nur ein Element besitzt, gibt es auch noch eine andere „triviale“ Halbgruppe. Sei nämlich M eine beliebige Menge und $m \in M$ ein Element, dann definiert $m_1 \circ m_2 := m$ für alle $m_1, m_2 \in M$ eine assoziative Verknüpfung auf M, also eine Halbgruppe, die wir hier mit* $\mathrm{Triv}(M)$ *bezeichnen wollen.*

Wenn wir die mathematischen Beispiele $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{R}$ betrachten, dann wissen wir aus unserer Erfahrung, dass es die speziellen Elemente 0 und 1 gibt, die bei Addition bzw. Multiplikation ein besonders einfaches Verhalten zeigen; dieses motiviert die folgende Definition.

Definition 5.2.4 (Links-/ Rechtseinselement). *Sei $(G,\circ)$ ein Gruppoid.*

(i) *Ein Element $e_L \in G$ heißt* Linkseinselement (linksneutrales Element)*, falls die Beziehung*

$$\forall g \in G : e_L \circ g = g$$

stimmt.

(ii) *Analog heißt ein Element* $e_R \in G$ Rechtseinselement (rechtsneutrales Element), *wenn sich bei Verknüpfung von rechts „nichts ändert“:*

$$\forall g \in G : g \circ e_R = g.$$

Gibt es ein Element $e \in G$, das sowohl Links- als auch Rechtseinselement ist, dann heißt e Einselement. Formal definieren wir.

Definition 5.2.5 (Einselement)**.** *Ein Element e eines Gruppoids G heißt* Einselement *oder* neutrales Element*, falls*

$$\forall g \in G : \ g \circ e = e \circ g = g$$

gilt. Wird die Verknüpfung mit + bezeichnet (additiv geschrieben), so bezeichnet man e oft mit 0 *oder* $\mathbb{0}$ *und nennt es* Nullelement*. Einselemente bezüglich multiplikativ geschriebener Verknüpfungen erhalten auch oft die Bezeichnung* 1 *oder* $\mathbb{1}$*.*

Beispiel 5.2.6 (Neutralitätseigenschaft)**.**

$(\mathbb{N}, +)$, $\ldots$, $(\mathbb{R}, \cdot)$**:** *Für die Addition von natürlichen, ganzen und reellen Zahlen ist klarerweise* 0 *das Nullelement, und für die Multiplikation ist* 1 *das Einselement.*

$(T, \circ)$**:** *Die Menge T enthält die Translation der Länge* 0*, welche das Objekt nicht von der Stelle bewegt. (Die Richtung ist hierbei egal!) Sie ist das Einselement von T.*

$(D, \circ)$**:** *Die Drehung um* 0 *Grad (die Achse ist dabei unerheblich) ist das Einselement der Halbgruppe D.*

$(\mathrm{Abb}(M), \circ)$**:** *In der Menge der Abbildungen* $\mathrm{Abb}(M)$ *bildet die Identität* $\mathbb{1}_M$ *(vgl. Seite 60) auf M das Einselement.*

$(W, \circ)$, $(S, \circ)$**:** *Führt man nicht künstlich leere Hauptwörter oder leere Strichblöcke ein, so enthalten W und S keine neutralen Elemente.*

$(M_2(\mathbb{R}), +)$**:** *Die Nullmatrix* $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ *ist das Nullelement von* $(M_2(\mathbb{R}), +)$*.*

$(M_2(\mathbb{R}), \cdot)$**:** *Auch* $(M_2(\mathbb{R}), \cdot)$ *hat ein Einselement, nämlich die Einheitsmatrix* $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$*.*

$(FP_2, \oplus)$**:** *Die Menge FP_2 hat ebenfalls ein Nullelement. Die Zahl* 0 *ist in FP_2 enthalten und besitzt alle Eigenschaften eines neutralen Elements.*

Nach Definition 5.2.5 können wir uns schon einmal fragen, welche Konsequenzen die Existenz eines Einselements hat. Die ersten beiden Ergebnisse finden wir in den folgenden Propositionen. Beim Beweis derselben, sowie bei den übrigen Beweisen in diesem Abschnitt müssen wir genauestens auf die Eigenschaften achten, die wir verwenden dürfen. Einer der beliebtesten Fehler in der Algebra ist, in Beweisen ohne zu zögern Eigenschaften der Verknüpfung zu verwenden, die gar nicht erfüllt sind — also Achtung!

Die Stärke der mathematischen Strukturtheorie gilt es auszunutzen. Wir wollen zum Beispiel die interessante Frage beantworten, ob in all unseren Beispielen das angegebene Einselement das einzige Element der Grundmenge ist, das die Neutralitätseigenschaft aufweist. Um nicht jedes Beispiel einzeln untersuchen zu müssen, verwenden wir *nur* die Struktureigenschaften für den Beweis.

Proposition 5.2.7 (Eindeutigkeit des neutralen Elements)**.**

(i) *Besitzt ein Gruppoid $(G, \circ)$ ein Einselement e, d.h. $\exists e \in G$:*

$$\forall g \in G : \ g \circ e = e \circ g = g,$$

so ist e eindeutig bestimmt (und wir können also von dem *Einselement in G sprechen).*

(ii) *Ist* $(G, \circ)$ *ein Gruppoid mit Linkseinselement* e_L *und Rechtseinselement* e_R*, so gilt*

$$e_L = e_R =: e$$

und e *ist Einselement in* G

Beweis.

(i) Sei $\tilde{e}$ ein weiteres neutrales Element in $(G, \circ)$, d.h. es existiert $\tilde{e}$ mit

$$\forall g \in G: \; g \circ \tilde{e} = \tilde{e} \circ g = g.$$

Wir zeigen, dass dann $e = \tilde{e}$ gilt. Zunächst gilt (setze $g = e$ in obiger Formel)

$$\tilde{e} \circ e = e.$$

Analog folgt (setze $g = \tilde{e}$ in der Formel in der Proposition)

$$\tilde{e} \circ e = \tilde{e}.$$

Insgesamt folgt also $\tilde{e} = e$.

(ii) Es gilt $e_L = e_L e_R$, da e_R ein Rechtseinselement ist, und weil e_L linksneutral ist, haben wir $e_L e_R = e_R$. Aus diesen Gleichungen sieht man aber sofort $e_L = e_R$. Setzen wir $e = e_L = e_R$, so erhalten wir das gewünschte Einselement.

□

Ein Element g in einem Gruppoid $(G, \circ)$ heißt *idempotent*, falls

$$g \circ g = g$$

gilt. Damit haben wir.

Proposition 5.2.8. *Ein (Links-, Rechts-) Einselement* e *eines Gruppoids* $(G, \circ)$ *ist immer idempotent.*

Beweis. Es gilt $e \circ e = e$, weil e (Links-, Rechts-) Einselement ist. □

Nachdem Einselemente häufig anzutreffen sind, hat man *Halbgruppen*, die ein solches enthalten, einen eigenen Namen gegeben.

Definition 5.2.9 (Monoid)**.** *Ist* $(G, \circ)$ *eine Halbgruppe und existiert ein Einselement* $e \in G$*, so nennt man* G *auch* Monoid *und schreibt oft* $(G, \circ, e)$*.*

Beispiel 5.2.10 (Monoide)**.**

(i) *Sowohl* $(\mathbb{N}, +)$ *als auch* $(\mathbb{N}, \cdot)$ *sind Monoide. Auch* $(\mathbb{Z}, +)$*,* $(\mathbb{Z}, \cdot)$*,* $(\mathbb{R}, +)$*,* $(\mathbb{R}, \cdot)$ *sind Monoide, so wie* $(\mathrm{Abb}(M), \circ)$ *und* $(T, \circ)$ *bzw.* $(D, \circ)$*.*

(ii) *Die Menge* $(FP_2, \oplus)$ *ist kein Monoid. Sie besitzt zwar ein neutrales Element aber die Verknüpfung ist nicht assoziativ (*FP_2 *ist ja nicht einmal eine Halbgruppe!).*

(iii) $(W, \circ)$ *und* $(S, \circ)$ *sind ebenfalls keine Monoide, weil sie kein neutrales Element besitzen. Wir könnten aber durch Hinzufügen des* leeren Hauptwortes *bzw. des* leeren Strichblockes *Einselemente in* W *und* S *definieren.*
*Auf diese Weise kann man übrigens aus jeder Halbgruppe durch Hinzufügen (*Adjungieren*) eines neutralen Elements ein Monoid machen.*

Fahren wir fort, die verschiedenen Beispiele miteinander zu vergleichen. Vielleicht können wir noch weitere Eigenschaften der Verknüpfungen isolieren.

Da stoßen wir übrigens auf ein wichtiges mathematisches Prinzip. Wir spüren eine Eigenschaft auf, geben ihr einen Namen und machen sie so reif für eine Untersuchung. Kreative Namensgebung ist bereits der erste Schritt zur erfolgreichen Behandlung einer Theorie. Die Kreativität liegt dabei natürlich mehr darauf, *was* und nicht darauf *wie* etwas benannt wird

— meist jedenfalls. Hätte nämlich der amerikanische Physiker George Zweig die kleinen Teilchen, aus denen die Elementarteilchen aufgebaut sind, nicht *Aces* genannt, so wäre heute sein Name berühmt und nicht der Name Murray Gell-Mann, der zur selben Zeit wie Zweig die Theorie der *Quarks* entdeckt aber den erfolgreicheren Namen gewählt hat.

Beispiel 5.2.11 (Kommutativität). *Wenn wir die Verknüpfungen untersuchen, die wir seit Beispiel 5.1.1 betrachten, dann fällt an manchen eine weitere Besonderheit auf.*

$(\mathbb{N}, +), \ldots, (\mathbb{R}, \cdot)$**:** *Am ehesten offensichtlich ist es bei den Zahlenmengen. In allen Beispielen von* $(\mathbb{N}, +)$ *bis* $(\mathbb{R}, \cdot)$ *kann man erkennen, dass es beim Addieren und Multiplizieren auf die Reihenfolge der Operanden nicht ankommt. Jeder „weiß“, dass etwa* $4 + 5 = 5 + 4$ *und* $3 \cdot 6 = 6 \cdot 3$ *gelten.*

$(T, \circ)$**:** *Die Translationen* T *haben ebenfalls diese Eigenschaft. Egal welche von zwei Translationen zuerst durchgeführt wird, das verschobene Objekt wird am selben Platz landen.*

$(D, \circ)$**:** *Drehungen sind allerdings anders: Legen wir das Koordinatenkreuz so, dass Ursprung und Schwerpunkt des zu drehenden Objektes zusammen fallen. Drehen wir zuerst um* $90°$ *um die* x_1*–Achse und danach um* $90°$ *um die* x_3*–Achse, so ergibt das eine Gesamtdrehung um die Achse, die durch den Punkt* $(1, -1, 1)$ *geht, um den Winkel* $120°$*. Vertauscht man die beiden Drehungen, dann ergibt sich eine Gesamtdrehung um die Achse durch den Punkt* $(1, 1, 1)$ *wieder um den Winkel* $120°$*. Die Reihenfolge, in der Drehungen ausgeführt werden, ist also wesentlich.*

$(\mathrm{Abb}(M), \circ)$**:** *Auch bei der Verknüpfung (allgemeiner) Abbildungen darf man nicht einfach die Reihenfolge vertauschen. Sind etwa* $f : \mathbb{R} \to \mathbb{R}$*,* $f : x \mapsto x^2$ *und* $g : \mathbb{R} \to \mathbb{R}$*,* $g : x \mapsto -x$ *gegeben. Dann gilt* $f \circ g : x \mapsto x^2$*, aber* $g \circ f : x \mapsto -x^2$*.*

$(M_2(\mathbb{R}), +)$**:** *Bei der Addition von* 2×2*–Matrizen darf man die Terme vertauschen. Das folgt trivialerweise aus der Tatsache, dass die Addition komponentenweise definiert ist.*

$(M_2(\mathbb{R}), \cdot)$**:** *Die Multiplikation in* $M_2(\mathbb{R})$ *ist da schon problematischer. Es gilt etwa*

$$A := \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \quad B := \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}$$
$$AB = \begin{pmatrix} -1 & 5 \\ -3 & 6 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & 5 \\ -1 & 4 \end{pmatrix}$$

Das Ergebnis der Multiplikation reeller 2×2*–Matrizen hängt also von der Reihenfolge der beiden Faktoren ab.*

$(S, \circ)$**:** *Das Ergebnis der Verknüpfung von Strichblöcken* S *ist wieder unabhängig von der Reihenfolgen der Operanden.*

$(W, \circ)$**:** *Bei Worten macht es dagegen einen Unterschied. „Dampfschiff“ hat eine gänzlich andere Bedeutung als „Schiffsdampf“.*

Wir sehen also, dass manchmal die Operanden einer Verknüpfung vertauscht werden dürfen ohne das Ergebnis zu ändern, manchmal aber aber nicht. Jetzt fehlt nur noch der Name für diese Eigenschaft.

Definition 5.2.12 (Kommutativgesetz). *Eine Verknüpfung in einem Gruppoid* $(G, \circ)$ *heißt* kommutativ*, falls das* Kommutativgesetz *erfüllt ist, d.h.*

$$\textbf{(KG)} \quad \forall g, h \in G : \; g \circ h = h \circ g.$$

Beispiel 5.2.13 (Kommutativgesetz).

(i) *Aus unserer Beispielliste erfüllen die Zahlenmengen, die Translationen,* $(M_2(\mathbb{R},+)$ *und die Strichblöcke mit den jeweiligen Verknüpfungen das Kommutativgesetz.*

(ii) *Nicht das Kommutativgesetz erfüllen hingegen die Drehungen, die Abbildungen (beide bzgl. der Hintereinanderausführung von Abbildungen),* $(M_2(\mathbb{R}),\cdot)$ *und die Menge der Hauptwörter (mit der Zusammensetzung).*

Und weiter führt uns unsere Entdeckungsreise durch die verschiedenen Verknüpfungseigenschaften. Wir fragen uns, ob eine einmal erfolgte Verknüpfungen wieder rückgängig gemacht werden kann. Bei den Translationen T kann man etwa nach jeder Verschiebung die Translation gleicher Länge aber entgegen gesetzter Richtung ausführen und damit das Objekt wieder an seinen ursprünglichen Platz zurückschieben. Translationen kann man also wieder ungeschehen machen. Wie das bei den anderen Verknüpfungen aussieht, wollen wir uns nach der folgenden Definitionen ansehen.

Definition 5.2.14 (Links-/ Rechtsinverses). *Sei* $(G,\circ,e)$ *ein Gruppoid mit Einselement.*

(i) *Ist* $a \in G$*, so nennen wir* $a' \in G$ *ein zu* a linksinverses Element*, falls*

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

(ii) *Ein Element* $a' \in G$ *heißt zu* a rechtsinvers*, wenn die umgekehrte Beziehung gilt, d.h.*

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

Analog zur Situation mit den einseitigen neutralen Elementen nennen wir ein a', das sowohl links- als auch rechtsinvers zu a ist, inverses Element von a. Formal definieren wir.

Definition 5.2.15 (Inverses Element). *Sei* $(G,\circ,e)$ *ein Gruppoid mit Einselement und sei* $a \in G$*. Wir nennen* $a' \in G$ inverses Element von a*, falls*

$$a' \circ a = a \circ a' = e$$

gilt. Wir nennen a' *auch* Inverses zu a *und schreiben meist* a^{-1}*. Ist das Verknüpfungszeichen ein* $+$*, schreiben wir die Operation also additiv, dann bezeichnen wir das Inverse von* a *üblicherweise mit* $-a$*.*

Beispiel 5.2.16 (Inverses).

$(\mathbb{N},+),\dots,(\mathbb{R},\cdot)$**:** *Bis zu diesem Zeitpunkt sind die Zahlenmengen brav neben einander marschiert und haben jeweils die gleichen Eigenschaften gehabt. Doch nun trennt sich die Verknüpfungsspreu vom Weizen.*

- *In* $(\mathbb{N},+,0)$ *gibt es außer für* 0 *zu keinem Element ein Inverses.*
- *In* $(\mathbb{Z},+,0)$ *und* $(\mathbb{R},+,0)$*, andererseits, hat jedes Element* $n \in \mathbb{Z}$ *bzw.* $n \in \mathbb{R}$ *ein inverses Element, nämlich* $-n$*.*
- *In* $(\mathbb{N},\cdot,1)$ *resp.* $(\mathbb{Z},\cdot,1)$ *besitzt außer* 1 *resp.* 1 *und* -1 *kein Element ein Inverses.*
- *In* $(\mathbb{R},\cdot,1)$ *hat jedes Element außer* 0 *ein Inverses.*

$(T,\circ)$**:** *Wir haben schon gesehen, dass die Translationen aus* T *Inverse besitzen, einfach die Verschiebung um dieselbe Länge in die Gegenrichtung.*

$(D,\circ)$**:** *Auch alle Drehungen in* D *haben Inverse, die Drehungen um dieselbe Achse um den negativen Winkel.*

$(\mathrm{Abb}(M),\circ)$**:** *In der Menge der Abbildungen* $\mathrm{Abb}(M)$ *haben nur die* bijektiven *Abbildungen Inverse (vgl. 4.3.12). Alle anderen können nicht rückgängig gemacht werden.*

$(M_2(\mathbb{R}),+)$**:** *In* $(M_2(\mathbb{R}),+)$ *hat jede Matrix* A *ein Inverses, nämlich diejenige Matrix* $-A$*, bei der man bei jedem Element von* A *das Vorzeichen gewechselt hat.*

$(M_2(\mathbb{R}),\cdot)$**:** *Für* $(M_2(\mathbb{R}),\cdot)$ *kann man beweisen (und das wird in der Linearen Algebra auch getan!) dass eine Matrix* A *genau dann ein Inverses hat, wenn* $a_{11}a_{22}-a_{12}a_{21} \neq 0$ *gilt; diese* A *heißen invertierbar.*

Wieder stehen wir vor der Frage, ob das Inverse zu einem Element, falls es überhaupt existiert, eindeutig bestimmt ist oder ob mehr als ein (Links-, Rechts-) Inverses existieren kann. Wieder beantwortet uns die Untersuchung der Struktureigenschaften die Frage für alle Beispiele auf einmal.

Proposition 5.2.17 (Eindeutigkeit des Inversen)**.** *Sei* $(G,\circ,e)$ *ein Monoid und* $g \in G$*.*

(i) *Existiert zu* g *ein Inverses* g^{-1}*, d.h.* $\exists g^{-1} \in G$*:*

$$g \circ g^{-1} = g^{-1} \circ g = e,$$

so ist g^{-1} *eindeutig bestimmt.*

(ii) *Ist* g_L^{-1} *ein Linksinverses von* g *und* g_R^{-1} *ein Rechtsinverses von* g*, d.h. gilt*

$$g_L^{-1} g = e = g g_R^{-1},$$

so ist $g_L^{-1} = g_R^{-1} =: g^{-1}$ *und* g^{-1} *ist Inverses zu* g*.*

Beweis.

(i) Sei $\tilde{g}$ ein weiteres Inverses zu g, d.h. es gelte $g \circ \tilde{g} = \tilde{g} \circ g = e$. Wir zeigen, dass dann schon $g^{-1} = \tilde{g}$ gilt. Tatsächlich ist

$$\tilde{g} = \tilde{g} \circ e = \tilde{g} \circ (g \circ g^{-1}) = (\tilde{g} \circ g) \circ g^{-1} = e \circ g^{-1} = g^{-1}.$$

(ii) Wir haben

$$g_L^{-1} = g_L^{-1} e = g_L^{-1}(g g_R^{-1}) = (g_L^{-1} g) g_R^{-1} = e g_R^{-1} = g_R^{-1}.$$

Daher sind g_L^{-1} und g_R^{-1} gleich und klarerweise ist $g^{-1} := g_L^{-1} = g_R^{-1}$ Inverses zu g.

□

Beachten Sie, dass wir im Beweis das Assoziativgesetz verwendet haben. Das ist in Ordnung, da wir ja $(G,\circ,e)$ als Monoid (also insbesondere als Halbgruppe) vorausgesetzt haben und nicht bloß als Gruppoid mit Einselement!

Jetzt haben wir alle Eigenschaften zusammen gesammelt und benannt und können endlich die Struktur definieren, auf die wir schon die ganze Zeit hinarbeiten.

Definition 5.2.18 (Gruppe)**.** *Sei* $(G,\circ)$ *ein Gruppoid. Gelten die folgende Eigenschaften*

(G1) *Assoziativgesetz:*

$$\forall g,h,k \in G : (g \circ h) \circ k = g \circ (h \circ k),$$

(G2) *Einselement:*

$$\exists e \in G : \forall g \in G : e \circ g = g \circ e = g,$$

(G3) *Inverse:*

$$\forall g \in G : \exists g^{-1} \in G : g^{-1} \circ g = g \circ g^{-1} = e,$$

dann heißt $(G,\circ)$ Gruppe. *Gilt außerdem noch*

(G4) *Kommutativgesetz:*

$$\forall g,h \in G : g \circ h = h \circ g,$$

dann heißt G kommutative *oder* abelsche Gruppe *(nach Nils Henrik Abel (1802–1829)).*

Die Eigenschaften (G1) bis (G3) nennt man auch oft die *Gruppenaxiome*. Sie besagen — anders ausgedrückt — dass eine Gruppe ein Monoid mit der zusätzlichen Eigenschaft ist, dass jedes $g \in G$ ein Inverses besitzt; siehe dazu auch Abbildung 5.2.

Gruppen werden in weiten Teilen der Mathematik benötigt. Sie beschreiben nicht nur Bewegungen sondern auch Symmetrien. Sie spielen ihre Rolle bei der Untersuchung von Differentialgleichungen genauso wie bei der Lösung von Optimierungsaufgaben oder der Lösung kombinatorischer Probleme. Zweifellos gehören Gruppen zu den zentralen Begriffen der Mathematik.

Auch im nächsten Abschnitt und in der linearen Algebra werden Gruppen gebraucht werden. Es ist also unerlässlich, diesen Begriff sorgfältig mit Fleisch (also mit Beispielen) zu füllen.

Beispiel 5.2.19 (Gruppen)**.**

(i) *Aus unserer Beispielliste bilden die ganzen Zahlen* $(\mathbb{Z}, +)$ *und die reellen Zahlen* $(\mathbb{R}, +)$ *eine abelsche Gruppe, so wie auch* $(\mathbb{R} \setminus \{0\}, \cdot)$. *Das selbe gilt für die Translationen und* $(M_2(\mathbb{R}), +)$

(ii) *Nicht kommutative Gruppen sind die Drehungen und die invertierbaren Matrizen in* $(M_2(\mathbb{R}), \cdot)$.

(iii) *Die einelementige Menge* $M = \{e\}$ *ist eine abelsche Gruppe mit der einzig möglichen Verknüpfung* $e \circ e = e$, *sie heißt Permutationsgruppe von einem Element* $\mathfrak{S}^1$ *oder* triviale Gruppe.

Wir wollen nun zeigen, dass Definition 5.2.18 teilweise redundant ist. Tatsächlich genügt es, folgende Abschwächung der Gruppenaxiome zu fordern.

Proposition 5.2.20. *Sei* $(G, \circ)$ *ein Gruppoid. Sind folgende Eigenschaften erfüllt, dann ist* G *eine Gruppe.*

(G1) *Assoziativgesetz:*

$$\forall g, h, k \in G : (g \circ h) \circ k = g \circ (h \circ k),$$

(G2') *Linkseinselement:*

$$\exists e \in G : \forall g \in G : e \circ g = g,$$

(G3') *Linksinverse:*

$$\forall g \in G : \exists g^{-1} \in G : g^{-1} \circ g = e.$$

Gilt außerdem noch

(G4) *Kommutativgesetz:*

$$\forall g, h \in G : g \circ h = h \circ g,$$

dann ist G *eine abelsche Gruppe.*

Beweis. Wir haben nicht alles vorausgesetzt, was wir vorher von einer Gruppe verlangt hatten. Eigenschaft G1, das Assoziativgesetz macht $(G, \circ)$ zu einer Halbgruppe, doch wir haben nur *Links*einselement und *Links*inverse vorausgesetzt. Wir müssen also zeigen, dass das Linkseinselement auch Rechtseinselement ist und dass alle Linksinversen auch Rechtsinverse sind.

Schritt 1: Wir beginnen mit einer Teilbehauptung. Ist $g \in G$ idempotent, so gilt schon $g = e$. Wir haben nämlich

$$\begin{aligned} gg &= g \\ g^{-1}(gg) &= g^{-1}g && \text{das Linksinverse } g^{-1} \text{ existiert immer} \\ (g^{-1}g)g &= g^{-1}g && \text{Assoziativität} \\ eg &= e && \text{weil } g^{-1} \text{ Linksinverses ist} \\ g &= e && \text{weil } e \text{ Linkseinselement ist} \end{aligned}$$

Das beweist unsere Teilbehauptung.
Schritt 2: Jetzt beweisen wir, dass das Linksinverse g^{-1} auch $gg^{-1} = e$ erfüllt, also Rechtsinverses ist.

$$\begin{aligned} gg^{-1} &= g(eg^{-1}) && \text{weil } e \text{ Linkseinselement ist} \\ &= g((g^{-1}g)g^{-1}) && \text{weil } g^{-1} \text{ Linksinverses ist} \\ &= (gg^{-1})(gg^{-1}) && \text{Assoziativität.} \end{aligned}$$

Aus obiger Beziehung folgt, dass gg^{-1} idempotent ist. Wir haben aber in Schritt 1 bewiesen, dass dann schon $gg^{-1} = e$ gilt.
Schritt 3: Es bleibt noch zu zeigen, dass für alle $g \in G$ auch $ge = g$ gilt, e also Rechtseinselement ist.

$$\begin{aligned} ge &= g(g^{-1}g) && \text{weil } g^{-1} \text{ Linksinverses ist} \\ &= (gg^{-1})g && \text{Assoziativität} \\ &= eg && \text{das haben wir in Schritt 2 gezeigt} \\ &= g && e \text{ ist Linkseinselement} \end{aligned}$$

Wir haben also gezeigt, dass e Einselement ist. Darum ist $(G, \circ, e)$ ein Monoid, und jedes Element besitzt ein Inverses wegen Schritt 2. Daher ist G eine Gruppe.

Die Aussage über die Kommutativität ist offensichtlich. □

Bemerkung 5.2.21.

(i) *Es existiert nur eine zweielementige Gruppe, nämlich* $\mathbb{Z}_2 := (\{0, 1\}, +)$ *mit* $0+0 = 0$, $1 + 0 = 0 + 1 = 1$ *und* $1 + 1 = 0$.

(ii) *Ist die Menge M endlich, so kann man jede Verknüpfung direkt angeben, indem man den Wert jedes Elements von $M \times M$ in einer Tabelle, der **Verknüpfungstabelle** auch **Cayley–Tafel**, anschreibt.*

Für $\mathbb{Z}_2$ würde das die Tabelle

+	0	1
0	0	1
1	1	0

ergeben. Sie drückt aus, was wir über das Addieren gerader und ungerader Zahlen wissen (0 ist die Äquivalenzklasse der geraden Zahlen und 1 diejenige der ungeraden Zahlen). Gerade plus gerade ist gerade, ungerade plus ungerade ist gerade, gerade plus ungerade ist ungerade.

Beispiel 5.2.22. *Betrachten wir ein ebenes gleichseitiges Dreieck und alle Abbildungen, die das Dreieck auf sich selbst abbilden (solche Abbildungen nennt man* Deckabbildungen*). Es gibt sechs verschiedene solche Abbildungen:*

(1) *Die Identität I,*
(2) *Drehung um $\frac{2}{3}\pi$ (120°) D_1,*
(3) *Drehung um $\frac{4}{3}\pi$ (240°) D_2,*
(4) *Spiegelung S_a an der Höhe auf a,*

(5) *Spiegelung S_b an der Höhe auf b,*
(6) *Spiegelung S_c an der Höhe auf c.*

Die Menge dieser Abbildungen bildet eine Gruppe bezüglich Verknüpfung von Abbildungen. Man kann die Wirkung der Abbildung am einfachsten veranschaulichen, indem man beobachtet, wohin die Eckpunkte abgebildet werden. Die Abbildung D_1 etwa bildet die Ecken ABC auf die Ecken BCA (in der Reihenfolge) ab. Die Spiegelung S_a bildet ABC auf ACB ab. Man sieht also, dass die Deckabbildungen des gleichseitigen Dreiecks genau die Permutationen (siehe die grauen Box auf Seite 59) der Eckpunkte sind. Die dabei entstehende Gruppe heißt $\mathfrak{S}^3$, und ihre Verknüpfungstabelle ist

$\circ$	I	S_a	S_b	S_c	D_1	D_2
I	I	S_a	S_b	S_c	D_1	D_2
S_a	S_a	I	D_2	D_1	S_c	S_b
S_b	S_b	D_1	I	D_2	S_a	S_c
S_c	S_c	D_2	D_1	I	S_b	S_a
D_1	D_1	S_b	S_c	S_a	D_2	I
D_2	D_2	S_c	S_a	S_b	I	D_1

Diese Gruppe ist die ***Permutationsgruppe*** *von drei Elementen oder auch Diedergruppe D_3 der Ordnung 3, eine* nicht abelsche *Gruppe. Sie ist sogar* die kleinste nicht abelsche Gruppe*.*

Beispiel 5.2.23. *Die Kleinsche Vierergruppe (V_4), auch Diedergruppe D_2 der Ordnung 2 genannt ist definiert durch die Verknüpfungstabelle*

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

Sie ist übrigens die kleinste nicht-zyklische Gruppe (wobei eine Gruppe zyklisch heißt, wenn sich alle Elemente als Potenzen eines einzigen Elements schreiben lassen).

Nun wenden wir uns wieder dem Studium der abstrakten Struktur einer Gruppe zu. Zunächst zeigen wir, dass das Gesetz der doppelten Inversion auch in Gruppen gilt.

Proposition 5.2.24 (Doppelte Inversion)**.** *Ist $(G, \circ)$ eine Gruppe, so haben wir für jedes $g \in G$*

$$(g^{-1})^{-1} = g.$$

Beweis. Das Element $(g^{-1})^{-1}$ ist das Inverse von g^{-1}. Wir wissen aber, dass $gg^{-1} = e$ gilt. Daher ist auch g das Inverse von g^{-1}. Wegen der Eindeutigkeit der Inversen (Proposition 5.2.17) folgt $g = (g^{-1})^{-1}$. □

Achtgeben muss man, wenn man das Verhältnis von Gruppenoperation und Inversion untersucht.

Proposition 5.2.25 (Rechenregeln für Gruppen)**.** *Ist $(G, \circ)$ eine Gruppe und $g, h, k \in G$. Dann gelten die Rechenregeln*

(1) $(g \circ h)^{-1} = h^{-1} \circ g^{-1}$ *(die Verknüpfung dreht sich um!),*
(2) $k \circ g = k \circ h \Rightarrow g = h$ *(es gilt die* Kürzungsregel*),*
(3) *Die Gleichung $g \circ x = h$ hat in G die eindeutige Lösung $x = g^{-1} \circ h$.*

Beweis.

(1) Es gilt

$$(g \circ h) \circ (h^{-1} \circ g^{-1}) = g \circ (h \circ h^{-1}) \circ g^{-1} = g \circ g^{-1} = e.$$

Die Aussage folgt nun aus der Eindeutigkeit der Inversen.

(2) Wir haben

$$\begin{aligned} k \circ g &= k \circ h \\ k^{-1} \circ (k \circ g) &= k^{-1} \circ (k \circ h) \\ (k^{-1} \circ k) \circ g &= (k^{-1} \circ k) \circ h \\ e \circ g &= e \circ h \\ g &= h. \end{aligned}$$

(3) Zunächst ist $x = g^{-1} \circ h$ eine Lösung, da $g \circ x = g \circ (g^{-1} \circ h) = (g \circ g^{-1}) \circ h = e \circ h = h$. Angenommen x' ist eine weitere Lösung so gilt $g \circ x = h = g \circ x'$ und aus der Kürzungsregel folgt nun $x = x'$.

□

Im Folgenden werden wir Teilmengen von Gruppen studieren und dabei unser erstes Beispiel einer *Teilstruktur* kennen lernen. Wenn $H \subseteq G$ gilt und $(G, \circ, e)$ eine Gruppe ist, dann ist automatisch eine Abbildung

$$\circ : H \times H \to G$$

definiert; jedes $h \in H$ ist ja auch Element in G und daher ist für jedes Paar $(g, h) \in H \times H$ die Verknüpfung $g \circ h$ definiert. Man spricht von der von G *ererbten* oder auch *induzierten* Operationen auf H.

Besonders interessant sind nun solche Teilmengen von Gruppen, die mit der ererbten Operation dieselbe Struktur aufweisen wie ihre Obermenge, also selbst Gruppen sind.

Definition 5.2.26 (Untergruppe). *Sei $(G, \circ, e)$ eine Gruppe. Eine Teilmenge $H \subseteq G$ heißt* Untergruppe*, falls $(H, \circ, e)$ eine Gruppe ist.*

Beispiel 5.2.27.

- *Jede Gruppe G besitzt die beiden trivialen Untergruppen $\{e\}$ und G.*
- *Die Gruppe $(\mathbb{Z}, +)$ ist eine Untergruppe von $(\mathbb{R}, +)$.*
- *Die Gruppe $(\mathbb{Z}, +)$ besitzt etwa die Untergruppe $\mathbb{Z}_g$ aller geraden ganzen Zahlen.*

Man bezeichnet Teilstrukturen (die gleiche Struktur auf einer Teilmenge) meist mit Unter... oder mit Teil...

In der Algebra kommen etwa *Untergruppen*, *Unterringe* und *Unterkörper* vor. In der linearen Algebra spricht man von *Teilräumen*, *Teilalgebren*,...

Sei $H \subseteq G$ Teilmenge der Gruppe $(G, \circ, e)$. Damit $(H, \circ, e)$ eine Gruppe ist, ist es notwendig, dass

$$\forall g, h \in H : \; g \circ h \in H$$

gilt also alle Verknüpfungen von Elementen aus H wiederum in H liegen (und nicht bloß in G). Die Verknüpfung darf demnach nicht aus H herausführen. Diese Eigenschaft nennt man *Abgeschlossenheit*; genauer sagt man, dass die Verknüpfung $\circ$ auf der Teilmenge $H \subseteq G$ **abgeschlossen** ist.

Nun stellt sich die Frage, ob und welche weiteren Eigenschaften gelten müssen, damit $H \subseteq G$ zur Untergruppe wird. Die Antwort gibt die folgende Proposition.

Proposition 5.2.28 (Charakterisierung von Untergruppen)**.** *Eine Teilmenge $H \subseteq G$ einer Gruppe $(G, \circ, e)$ ist genau dann eine Untergruppe, wenn eine der beiden äquivalenten Bedingungen gilt:*

(1) *Für alle $g, h \in H$ ist auch $g \circ h^{-1} \in H$*
(2) *Für alle $g, h \in H$ liegt die Verknüpfung $g \circ h \in H$ und zusätzlich liegt zu jedem Element $h \in H$ auch das Inverse $h^{-1} \in H$.*

Ist G abelsch, dann auch H.

Beweis. Zuerst beweisen wir die Äquivalenz der Eigenschaften.
(1) $\Rightarrow$ (2): Ist für je zwei Elemente $g, h \in H$ auch $g \circ h^{-1} \in H$, so sehen wir sofort, dass $e = g \circ g^{-1} \in H$ liegt. Damit ist aber auch zu jedem $g \in H$ das Element $e \circ g^{-1} = g^{-1} \in H$. Ferner muss dann aber für $g, h^{-1} \in H$ das Element $g \circ (h^{-1})^{-1} = g \circ h \in H$ liegen.
(2) $\Rightarrow$ (1): Seien $g, h \in H$. Dann erhalten wir $h^{-1} \in H$, und daher ist auch $g \circ h^{-1} \in H$.
Das beweist die behauptete Äquivalenz.

Nun zeigen wir, dass die Bedingung (2) impliziert, dass H eine Gruppe ist. Der erste Schritt dabei ist zu zeigen, dass $(H, \circ)$ ein Gruppoid bildet, dass also $\circ$ eine Verknüpfung auf H ist. Das ist aber tatsächlich der Fall, weil wir schon wissen, dass für je zwei Elemente $g, h \in H$ auch $g \circ h \in H$ liegt. Damit ist aber H bereits eine Halbgruppe, denn das Assoziativgesetz gilt, weil es sogar für alle Elemente in G erfüllt ist.

Das Einselement e von G liegt ebenfalls in H, da für jedes Element $g \in H$ auch $g \circ g^{-1} = e \in H$ sein muss. Schließlich besitzt jedes Element $g \in H$ ein Inverses in G, nämlich g^{-1}, von dem wir bereits wissen, dass es in H liegt. Das beweist alle Gruppeneigenschaften für $(H, \circ, e)$, und daher ist H eine Untergruppe von G.

Die umgekehrte Richtung, d.h. dass für eine Untergruppe U die Eigenschaft (2) gilt, ist klar.

Nun fehlt nurmehr die Aussage über die Kommutativität, die aber ebenfalls leicht einzusehen ist. Wenn G abelsch ist, dann erfüllen alle Elemente in G das Kommutativgesetz, also erst recht alle in H. $\square$

Ein wichtiger Begriff der Algebra fehlt noch. Wir haben jetzt aus zuvor unbedarften Mengen neue mathematische Strukturen geschaffen, indem wir auf ihnen eine Verknüpfung eingeführt haben. Dann haben wir die Eigenschaften dieser Verknüpfungen untersucht und sind so schließlich zur Definition der Gruppe gekommen. Wo sind aber die versprochenen Verbindungen zwischen unseren Gruppenobjekten? Bei den Mengen hatten wir die Abbildungen. Was sollen wir bei den Gruppen verwenden.

Die Lösung ist einfach. Gruppen sind Mengen, also können wir mit Abbildungen anfangen. Um allerdings die Gruppenstruktur nicht zu vergessen, müssen wir von den Abbildungen verlangen, dass sie die Gruppenstruktur nicht zerstören. Das führt zur folgenden Definition.

Definition 5.2.29 (Gruppenhomomorphismus)**.** *Seien $(G, \circ)$ und $(H, \Box)$ Gruppen.*

(i) *Ein* Gruppenhomomorphismus *von G nach H ist eine Abbildung $f : G \to H$ mit*

$$\forall g_1, g_2 \in G : f(g_1 \circ g_2) = f(g_1) \Box f(g_2).$$

Das bedeutet also, dass es unerheblich ist, ob man zuerst in G verknüpft und dann nach H abbildet oder zuerst nach H abbildet und dann dort verknüpft.

(ii) *Ist die Abbildung f bijektiv, dann heißt sie* Gruppenisomorphismus*; in diesem Fall sagen wir die beiden Gruppen sind* isomorph*.*

(iii) *Sind G und H nur Halbgruppen, so heißt f* Halbgruppenhomomorphismus. *Ist f zusätzlich bijektiv, so heißt f* Halbgruppenisomorphismus.

(iv) *Sind G und H lediglich Gruppoide, so heißt f* Gruppoidhomomorphismus. *Ist f zusätzlich bijektiv, so heißt f* Gruppoidisomorphismus.

Wie aus der obigen Definition bereits erahnt werden kann, werden in der Mathematik Abbildungen zwischen Mengen mit zusätzlicher Struktur, die diese Struktur erhalten **Homomorphismen** genannt und der Name der Struktur vorangestellt.

Bijektive Homomorphismen heißen **Isomorphismen**, wobei ebenfalls der Name der Struktur vorangestellt wird.

Homomorphismen von Mengen auf sich heißen **Endomorphismen** und Isomorphismen von Mengen auf sich **Automorphismen**, wobei wiederum jeweils der Name der Struktur vorangestellt wird.

Ein Gruppenisomorphismus (wie jeder andere Isomorphismus in der Mathematik auch) ist im wesentlichen nichts anderes als eine *Umbenennung* der Gruppenelemente. Dass solche Umbenennungen mitunter sehr praktisch sein können, muss nicht extra erwähnt werden. Zwei isomorphe Strukturen sind vom Standpunkt der Strukturtheorie aus ununterscheidbar. Oftmals kann man sich bei der Untersuchung der Eigenschaften eines bestimmten Objektes damit wesentlich weiter helfen, einen Isomorphismus zu einem bereits bekannten Objekt zu konstruieren.

Beispiel 5.2.30 (Gruppenhomomorphismen)**.**

- *Die Abbildung, die jedem $z \in \mathbb{Z}$ die reelle Zahl $z \in \mathbb{R}$ zuordnet, ist ein Gruppenhomomorphismus von $(\mathbb{Z}, +)$ in $(\mathbb{R}, +)$.*
- *Die Abbildung f von S nach $\mathbb{N}$, die jedem Strichblock die Anzahl der enthaltenen Striche zuordnet, ist ein Halbgruppenhomomorphismus von S nach $\mathbb{N}$. Haben wir zu S den leeren Strichblock hinzugefügt, dann ist $f : S \to \mathbb{N}$ bijektiv, also ein Halbgruppenisomorphismus. Die Menge der Strichblöcke ist also von den natürlichen Zahlen nicht unterscheidbar vom Standpunkt der Halbgruppentheorie aus. Die Menge S ist eine Möglichkeit, $\mathbb{N}$ zu konstruieren. Eine andere Variante, in der $\mathbb{N}$ aus den Mengenaxiomen hergeleitet wird, findet sich in Abschnitt 6.1.1.*
- *Seien G und H Gruppen, so kann trivialerweise immer der folgende Gruppenhomomorphismus definiert werden: $f(g) = e' \ \forall g \in G$, wobei e' das Einselement in H ist.*
- *Ein äußerst wichtiges Beispiel ist die Abbildung*

$$\begin{aligned} f : (\mathbb{R}, +) &\to (\mathbb{R} \setminus \{0\}, \cdot) \\ f(x) &= e^x \qquad (e,\ \textit{die Eulersche Zahl}). \end{aligned}$$

f ist ein Gruppenhomomorphismus, da für alle $x, y \in \mathbb{R}$ gilt, dass $f(x+y) = e^{x+y} = e^x e^y$. Schränkt man den Zielbereich von f auf $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$ ein ($(R^+, \cdot)$ ist Untergruppe von $(\mathbb{R} \setminus \{0\}, \cdot)$!), so ist f sogar ein Gruppenisomorphismus mit Umkehrabbildung $f^{-1}(z) = \log z$.

Zum Abschluss des Abschnitts zeigen wir, dass Definition 5.2.29 tatsächlich ausreicht, um die gesamte Gruppenstruktur zu respektieren. Es ist nicht nur unerheblich ob vor oder nach der der Anwendung von f verknüpft wird, sondern das Einselement wird auf das Einselement abgebildet und so ist es auch egal, ob vor oder nach der Abbildung invertiert wird.

Proposition 5.2.31 (Eigenschaften von Gruppenhomomorphismen)**.** *Sei f ein Gruppenhomomorphismus von $(G, \circ, e)$ nach $(H, \square, e')$. Dann gilt*

(1) *$f(e) = e'$ und*

(2) *$\forall g \in G :\ f(a^{-1}) = (f(a))^{-1}$.*

Beweis.

(1) Es gilt $f(e) = f(e \circ e) = f(e)\Box f(e)$, daher $e'\Box f(e) = f(e) = f(e)\Box f(e)$ und mit der Kürzungsregel (in H) folgt die Aussage.

(2) Es gilt $f(e) = f(a \circ a^{-1}) = f(a)\Box f(a^{-1})$ und nach (1) $e' = f(a)\Box f(a^{-1})$. Analog folgt $e' = f(a^{-1})\Box f(a)$ und wegen der Eindeutigkeit der Inversen ist $f(a^{-1})$ das Inverse (in H) zu $f(a)$, d.h. $(f(a))^{-1} = f(a^{-1})$.

□

5.3. Ringe

Um uns den „Schmuckstücken" der Mathematik zu nähern, kehren wir zurück zu unseren Gruppoiden aus Beispiel 5.1.1. Einige dort betrachtete Mengen haben als doppeltes Beispiel gedient. So etwa $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{R}$ aber auch die 2×2–Matrizen $M_2(\mathbb{R})$. Für alle diese Mengen haben wir Summen und Produkte definiert. Alle diese Mengen sind also Gruppoide bezüglich zweier Verknüpfungen.

Beispiel 5.3.1 (Distributivität)**.** *Wichtig an den oben erwähnten Mengen mit Gruppoid–Strukturen bezüglich zweier Verknüpfungen ist die Eigenschaft, dass „Ausmultiplizieren" und „Herausheben" („Ausklammern") gültige Rechenregeln sind. Wir alle wissen ja, dass etwa* $(3 + 4) \cdot 5 = 3 \cdot 5 + 4 \cdot 5$ *gilt.*

Von nun an werden wir Mengen betrachten, auf denen *zwei* Verknüpfungen definiert sind. Wir schreiben die beiden Verknüpfungen $+$ und $\cdot$, vereinbaren, dass $\cdot$ stärker bindet als $+$ („Punktrechnung vor Strichrechung"), und lassen, wie schon angekündigt, den Punkt weg, wenn immer angebracht.

Definition 5.3.2 (Distributivgesetze)**.** *Sei H ein Gruppoid bezüglich zweier Verknüpfungen $+$ und $\cdot$. Wir schreiben dann auch $(H, +, \cdot)$ und sagen $+$ erfüllt die beiden* Distributivgesetze *bzgl. $\cdot$ falls gilt*

$$\textbf{(DG1)} \qquad \forall a, b, c \in H: \ a(b + c) = ab + ac,$$
$$\textbf{(DG2)} \qquad \forall a, b, c \in H: \ (b + c)a = ba + ca.$$

Definition 5.3.3 (Halbring)**.**

(i) *Eine Menge H, die eine Halbgruppe $(H, +)$ und eine Halbgruppe $(H, \cdot)$ bildet, heißt* Halbring*, falls die beiden Distributivgesetze von $+$ bezüglich $\cdot$ erfüllt sind.*

(ii) *Ist $(H, +)$ eine kommutative Halbgruppe, so sprechen wir von einem* additiv kommutativen *Halbring, ist $(H, \cdot)$ kommutativ, so nennen wir die Struktur einen* multiplikativ kommutativen *Halbring. Sind beide Verknüpfungen kommutativ, so liegt ein* kommutativer *Halbring vor.*

Beispiel 5.3.4 (Halbringe)**.**

(i) *Die natürlichen Zahlen $(\mathbb{N}, +, \cdot)$ bilden einen kommutativen Halbring. Manche nennen das sogar **Dioid**, da beide Halbgruppen $(\mathbb{N}, +)$ und $(\mathbb{N}, \cdot)$ sogar Monoide sind.*

(ii) *Auch $(\mathbb{Z}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ besitzen eine Halbringstruktur. Dies folgt aus Beispiel 5.2.3 und der offensichtlichen Gültigkeit der Distributivgesetze.*

(iii) *Die interessante Frage ist: Ist $M_2(\mathbb{R})$ ebenfalls ein Halbring? Die Antwort ist ja, ein additiv kommutativer Halbring. Das Nachrechnen der Distributivgesetze ist allerdings ein bisschen mühsam, ergibt sich aber aus der Gültigkeit der Distributivgesetze für die reellen Zahlen.*

Das Nullelement der Operation $+$ in einem Halbring bezeichnen wir mit 0 und das Einselement von $\cdot$ mit 1, sofern sie existieren.

Beispiel 5.3.5. *Einige unserer Beispielmengen besitzen aber noch mehr Struktur. So ist zwar* $(\mathbb{N}, +)$ *keine Gruppe, sehr wohl sind aber* $(\mathbb{Z}, +)$ *und* $(\mathbb{R}, +)$ *kommutative Gruppen. Auch* $(M_2(\mathbb{R}), +)$ *ist eine abelsche Gruppe.*

Dies führt uns unmittelbar zum nächsten Begriff.

Definition 5.3.6 (Ring)**.** *Sei* $(R, +, \cdot)$ *ein Gruppoid bezüglich beider Verknüpfungen.*

(i) $(R, +, \cdot)$ *heißt* Ring*, falls folgende Eigenschaften erfüllt sind:*

(R1) $(R, +, 0)$ *ist abelsche Gruppe,*

(R2) $(R, \cdot)$ *erfüllt das Assoziativgesetz,*

(R3) $(R, +, \cdot)$ *erfüllt die Distributivgesetze von* $+$ *bzgl.* $\cdot$*.*

(ii) *Ist zusätzlich* $(R, \cdot, 1)$ *ein Monoid und gilt* $0 \neq 1$*, so sagen wir* R *ist ein* Ring mit Einselement *und schreiben oft auch* $(R, +, \cdot, 0, 1)$*.*

(iii) *Ist die Operation* $\cdot$ *kommutativ, so liegt ein* kommutativer Ring *vor.*

(iv) *Hat man beides, Kommutativität und Einselement, dann nennt man die entstehende Struktur ganz einfach* kommutativer Ring mit Einselement*.*

Obige Definition besagt also, dass ein Halbring $(R, +, \cdot)$ ein Ring ist, falls zusätzlich (R1) gilt, also $(R, +)$ eine abelsche Gruppe ist.

Beispiel 5.3.7 (Ringe)**.**

(i) *Die ganzen Zahlen* $(\mathbb{Z}, +, \cdot)$ *und die reellen Zahlen* $(\mathbb{R}, +, \cdot)$ *sind kommutative Ringe mit Einselement.*

(ii) *Die reellen* 2×2*–Matrizen bilden einen Ring mit Einselement, der aber nicht kommutativ ist (vgl. Beispiel 5.2.11).*

Einige Ringe haben wir jetzt identifiziert in unserer täglichen mathematischen Umgebung. Nun spielen wir wieder die Stärken der Algebra aus und suchen *nur an Hand der geforderten Eigenschaften* nach neuen Gesetzen, die in *allen* Ringen gelten.

Proposition 5.3.8 (Rechenregeln in Ringen)**.** *Ist* $(R, +, \cdot)$ *ein Ring, so gelten die Rechenregeln*

(1) $\forall r \in R : r0 = 0r = 0,$

(2) $\forall r, s \in R : -(rs) = (-r)s = r(-s),$

(3) $\forall r, s \in R : rs = (-r)(-s).$

(4) *Besitzt* R *ein Einselement* $1 \neq 0$*, so gilt* $\forall r \in R : (-1)r = r(-1) = -r.$

Beweis.

(1) Es gilt $r0 = r(0+0) = r0 + r0$ und durch Addition von $-(r0)$ auf beiden Seiten der Gleichung folgt $0 = r0$. Die zweite Gleichung folgt analog.

(2) Wir haben $(-r)s + rs = ((-r) + r)s = 0s = 0$ wegen (1). Aus der Eindeutigkeit des Inversen folgt $-(rs) = (-r)s$. Analog finden wir $r(-s) + rs = r((-s) + s) = r0 = 0$ und damit $-(rs) = r(-s)$.

(3) Mit Proposition 5.2.24 und zweimaliger Verwendung von (2) folgt $rs = -(-(rs)) = -((-r)s) = (-r)(-s)$.

(4) Es gilt $0 = 0r = (1 + (-1))r = 1r + (-1)r = r + (-1)r$ und damit $-r = (-1)r$ wegen der Eindeutigkeit der Inversen. Die zweite Gleichung zeigt man analog.

$\square$

Genau wie für Gruppen können wir auch für Ringe Teilstrukturen definieren.

Definition 5.3.9 (Teilring)**.** *Eine Teilmenge* $S \subseteq R$ *eines Ringes* $(R, +, \cdot)$ *heißt* Teilring *oder* Unterring *von* R*, falls* $(S, +, \cdot)$ *mit den induzierten Verknüpfungen ein Ring ist.*

Zur Überprüfung der Tatsache, ob eine Teilmenge eines Rings ein Unterring ist müssen glücklicherweise nicht alle Ringeigenschaften nachgeprüft werden. Im wesentlichen genügt es nämlich wiederum nur zu zeigen, dass die Verknüpfungen aus der Teilmenge nicht hinausführen.

Proposition 5.3.10 (Charakterisierung von Unterringen)**.** *Eine Teilmenge $S \subseteq R$ eines Ringes $(R, +, \cdot)$ ist ein Unterring genau dann, wenn für alle $r, s \in S$ die Elemente $r - s$ und rs in S liegen.*

Ist R kommutativ, dann auch S.

Beweis. Weil für $r, s \in S$ schon $r - s \in S$ folgt, wissen wir aus Proposition 5.2.28, dass $(S, +)$ eine Gruppe ist und zwar eine Untergruppe von $(R, +)$. Die Verknüpfung $\cdot$ ist in S abgeschlossen, denn das haben wir vorausgesetzt. Weil aber das Assoziativgesetz und die Distributivgesetze für alle Elemente in R gelten, stimmen sie erst recht für alle Elemente von S. Daher ist S ein Ring.

Die Aussage über Kommutativität ist offensichtlich. □

Beispiel 5.3.11. *Für zwei ganze Zahlen p und q wissen wir folgende Eigenschaft: Sind $p \neq 0$ und $q \neq 0$, dann ist auch $pq \neq 0$. Auch die Menge der reellen Zahlen erfüllt das.*

In den 2×2–Matrizen können wir so schnell nicht schließen. Es gilt nämlich

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

In $M_2(\mathbb{R})$ ist also das Produkt von Null verschiedener Elemente nicht notwendigerweise auch von Null verschieden. Das ist eine bemerkenswerte Eigenschaft, die uns zur nächsten Definition führt.

Definition 5.3.12. *Ein kommutativer Ring mit Einselement $(R, +, \cdot, 0, 1)$ heißt* Integritätsbereich*, wenn für je zwei Elemente $r, s \in R$ aus $rs = 0$ schon $r = 0$ oder $s = 0$ folgt.*

Anders ausgedrückt, besitzt ein Integritätsbereich keine so genannten **Nullteiler**, wobei Nullteiler Elemente $r, s \neq 0$ mit $rs = 0$ sind.

Beispiel 5.3.13 (Integritätsbereiche)**.**

(i) *Die ganzen Zahlen $(\mathbb{Z}, +, \cdot, 0, 1)$ sind ein Integritätsbereich, ebenso die reellen Zahlen $(\mathbb{R}, +, \cdot, 0, 1)$.*
(ii) *Die Matrizen $M_2(\mathbb{R})$ sind kein Integritätsbereich, denn die Multiplikation ist nicht kommutativ, und $M_2(\mathbb{R})$ ist nicht nullteilerfrei.*

Wie zu den Gruppen gehören auch zu den Ringen bestimmte Abbildungen, die sich mit der Struktur vertragen. Es ist immer das gleiche Prinzip. Ein Ring ist eine Gruppe mit etwas Zusatzstruktur, also ist ein Ringhomomorphismus ein Gruppenhomomorphismus, der „noch ein bisschen mehr kann".

Definition 5.3.14 (Ringhomomorphismus)**.** *Seien $(R, +, \cdot)$ und $(S, \oplus, \otimes)$ zwei Ringe.*

(i) *Ein* Ringhomomorphismus *ist ein Gruppenhomomorphismus $f : (R, +) \to (S, \oplus)$, für den zusätzlich noch*

$$\forall r, r' \in R : f(rr') = f(r) \otimes f(r')$$

gilt (der also außerdem noch ein Halbgruppenhomomorphismus $(R, \cdot) \to (S, \otimes)$ ist).
(i) *Ist f bijektiv, dann heißt f* Ringisomorphismus *und man sagt, R und S sind* isomorph.

Beispiel 5.3.15. *Die Abbildung $\iota : \mathbb{R} \to M_2(\mathbb{R})$, die jeder reellen Zahl r die Matrix $\begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}$ zuordnet, ist ein Ringhomomorphismus von $(\mathbb{R}, +, \cdot)$ nach $(M_2(\mathbb{R}), +, \cdot)$. Es gilt nämlich*

$$\iota(r_1) + \iota(r_2) = \begin{pmatrix} r_1 & 0 \\ 0 & r_1 \end{pmatrix} + \begin{pmatrix} r_2 & 0 \\ 0 & r_2 \end{pmatrix} = \begin{pmatrix} r_1 + r_2 & 0 \\ 0 & r_1 + r_2 \end{pmatrix} = \iota(r_1 + r_2),$$

sowie

$$\iota(r_1)\, \iota(r_2) = \begin{pmatrix} r_1 & 0 \\ 0 & r_1 \end{pmatrix} \begin{pmatrix} r_2 & 0 \\ 0 & r_2 \end{pmatrix} = \begin{pmatrix} r_1 r_2 & 0 \\ 0 & r_1 r_2 \end{pmatrix} = \iota(r_1 r_2).$$

In diesem Beispiel können wir auch leicht sehen, dass die Inversion mit der Abbildung vertauscht — es also egal ist ob zuerst abgebildet, dann Invertiert wird oder umgekehrt. (Allgemein folgt dies ja aus Proposition 5.2.31.) Es gilt

$$\iota(-r) = \begin{pmatrix} -r & 0 \\ 0 & -r \end{pmatrix} = -\iota(r),$$

und

$$\iota(r^{-1}) = \begin{pmatrix} r^{-1} & 0 \\ 0 & r^{-1} \end{pmatrix} = \iota(r)^{-1}.$$

*Der Ringhomomorphismus ι ist sogar injektiv. Man sagt er **bettet** $\mathbb{R}$ in die Menge der 2×2–Matrizen **ein**. Er ist nicht surjektiv, da z.B. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ nicht im Bild von $\mathbb{R}$ liegt.*

Dass ι nicht bijektiv sein kann, wissen wir schon aufgrund folgender Tatsache: Wäre ι ein Ringisomoprhismus, so wären $\mathbb{R}$ und $M_2(\mathbb{R})$ aus Sicht der Ringtheorie ununterscheidbar. Das kann aber nicht sein, da $\mathbb{R}$ ein Integritätsbereich ist und $M_2(\mathbb{R})$ nicht.

5.4. Körper

Jetzt sind wir beinahe am Ende unseres Weges angelangt. Die folgende speziellste Struktur der Algebra für Mengen mit zwei Verknüpfungen spielt in der Mathematik eine herausragende Rolle. Sie wird sowohl in der Analysis, als auch in der Linearen Algebra ein wesentlicher Begleiter sein, und daher ist es wichtig, sich die Eigenschaften möglichst gut einzuprägen.

Definition 5.4.1 (Körper)**.** *Ein Gruppoid $(K, +, \cdot)$ mit den beiden Verknüpfungen $+$ und $\cdot$ heißt* Körper*, falls die folgenden Eigenschaften (die* Körperaxiome*) gelten.*

- **(K1)** $\forall a, b, c \in K : (a + b) + c = a + (b + c)$ *(Assoziativität von $+$),*
- **(K2)** $\forall a, b \in K : a + b = b + a$ *(Kommutativität von $+$),*
- **(K3)** $\exists 0 \in K : \forall a \in K : a + 0 = a$ *(Nullelement),*
- **(K4)** $\forall a \in K : \exists (-a) \in K : a + (-a) = 0$ *(Inverse bzgl. $+$),*
- **(K5)** $\forall a, b, c \in K : (ab)c = a(bc)$ *(Assoziativität von $\cdot$),*
- **(K6)** $\forall a, b \in K : ab = ba$ *(Kommutativität von $\cdot$),*
- **(K7)** $\exists 1 \in K : 1 \neq 0 \wedge \forall a \in K \setminus \{0\} : a1 = a$ *(Einselement),*
- **(K8)** $\forall a \in K \setminus \{0\} : \exists a^{-1} \in K : aa^{-1} = 1$ *(Inverse bzgl. $\cdot$),*
- **(K9)** $\forall a, b, c \in K : a(b + c) = ab + ac$ *(Distributivität).*

Bemerkung 5.4.2. *Die Bedingungen (K1) bis (K4) machen $(R, +, 0)$ zur abelschen Gruppe und die Bedingungen (K5)–(K8) implizieren, dass $(R \setminus \{0\}, \cdot)$ ebenfalls eine abelsche Gruppe ist. Schließlich sorgt das Distributivgesetz (K9) (wegen der Kommutativität gilt* *das zweite Distributivgesetz (DG2) dann automatisch) für die „Verträglichkeit" der beiden Operationen.*

Anders ausgedrückt ist ein Ring mit Einselement $(K, +, \cdot)$ *ein Körper, wenn zusätzlich* $(K \setminus \{0\}, \cdot)$ *eine abelsche Gruppe ist.*

Beispiel 5.4.3 (Körper)**.** *Die rationalen Zahlen* $(\mathbb{Q}, +, \cdot, 0, 1)$ *bilden ebenso einen Körper wie die reellen oder komplexen Zahlen.*

Um weitere Beispiele zu finden, müssen wir ein wenig arbeiten — was wir im folgenden Beispiel auch tun.

Beispiel 5.4.4 (Restklassenkörper)**.** *Die schon aus Beispiel 4.2.13 bekannten Restklassen* $\mathbb{Z}_p$ *bilden einen kommutativen Ring mit Einselement mit den Verknüpfungen*

$$\begin{aligned} \overline{a} + \overline{b} &:= \overline{a+b} \\ \overline{a} \cdot \overline{b} &:= \overline{ab}. \end{aligned}$$

Ist p *eine Primzahl, so ist* $\mathbb{Z}_p$ *sogar ein Körper.*

Zuerst seien die Eigenschaften für Ringe überprüft: Die Operation $+$ *ist* wohldefiniert, *weil für je zwei verschiedene* Repräsentanten $a, a' \in \overline{a}$ *bzw.* $b, b' \in \overline{b}$ *gilt:* $a = a' + kp$ *und* $b = b' + \ell p$ *für geeignete* $k, \ell \in \mathbb{Z}$*. Dann ist aber* $a + b = a' + b' + (k + \ell)p$*, und damit ist* $\overline{a+b} = \overline{a' + b'}$*.*

*Der Ausdruck **wohldefiniert** bedeutet nicht, dass etwas „schön" definiert ist. Diesen Ausdruck verwendet man, wenn man eine Beziehung, eine Operation, eine Abbildung für eine Klasse von Objekten dadurch definiert, dass man einen **Repräsentanten** aus der Klasse wählt und für diesen die Beziehung, Operation, Abbildung erklärt. Dann muss man nämlich überprüfen, ob diese Definition **unabhängig** von der Wahl des Repräsentanten ist oder ob die Definition etwa auf verschiedenen Elementen der Äquivalenzklasse verschiedenes bedeutet, denn das wäre schlecht.*

Ein Beispiel einer nicht wohldefinierten Operation auf $\mathbb{Z}_3$ *ist* $\sqrt{\overline{a}} := \overline{\sqrt{a}}$*, wenn* a *eine Quadratzahl ist. Wollen wir* $\sqrt{\overline{1}}$ *berechnen, so finden wir* $\sqrt{\overline{1}} = \overline{\sqrt{1}} = \overline{1}$*. Gleichzeitig gilt aber* $\overline{1} = \overline{4}$*, und wir hätten* $\sqrt{\overline{1}} = \sqrt{\overline{4}} = \overline{\sqrt{4}} = \overline{2}$*, was zu einem Widerspruch führt. Die Operation* $\sqrt{}$ *wie oben eingeführt ist also* nicht *wohldefiniert.*

Ebenso gilt für $\cdot$*:* $ab = (a' + kp)(b' + \ell p) = (a'b' + (a'\ell + kb' + k\ell p)p)$*, und daher ist* $\overline{ab} = \overline{a'b'}$*. Auch* $\cdot$ *ist also wohldefiniert.*

Weil für ganze Zahlen (und das sind die Repräsentanten der Nebenklassen ja auch!) Assoziativgesetz, Kommutativgesetz und Distributivgesetz gelten, gelten diese Gesetze auch für $+$ *und* $\cdot$ *auf* $\mathbb{Z}_p$*. Das Nullelement ist* $\overline{0}$*, und das Einselement* $\overline{1}$ *erfüllt für* $p > 1$ *auch* $\overline{0} \neq \overline{1}$*. Das additiv Inverse einer Klasse* $\overline{a}$ *ist leicht gefunden. Es ist* $\overline{-a}$*.*

Um zu überprüfen, dass $\mathbb{Z}_p$ *ein Körper ist, wenn* p *eine Primzahl ist, müssen wir nur noch beweisen, dass jedes Element* $\overline{a} \neq \overline{0}$ *ein Inverses besitzt. Dazu müssen wir eine Restklasse* $\overline{b}$ *finden mit* $\overline{a} \cdot \overline{b} = \overline{1}$*. Ein Satz aus der elementaren Zahlentheorie besagt folgendes:*

> *Sind* $x, y \in \mathbb{N}$ *mit* $\mathrm{ggT}(x, y) = 1$*, so gibt es ganze Zahlen* b, n *mit*
>
> $$1 = bx + ny.$$

Für jede Restklasse $\overline{a}$ *mit* $\overline{a} \neq \overline{0}$ *ist* $\mathrm{ggT}(a, p) = 1$*, da* p *Primzahl ist. Somit folgt die Existenz zweier Zahlen* b, n *mit* $ba + np = 1$*. Daher ist* $\overline{b}$ *das Inverse zu* $\overline{a}$*, und* $\mathbb{Z}_p$ *ist tatsächlich ein (endlicher) Körper.*

In der Zahlentheorie sind die Operationen in den $\mathbb{Z}_m$ sehr wichtig. Dort hat sich eine eigene Schreibweise etabliert. Für $a + b = c$ in $\mathbb{Z}_m$ schreibt man

$$a + b \equiv c \mod m$$

und spricht: „a plus b ***kongruent*** *c* ***modulo*** *m“. Ebenso für das Produkt*

$$a \cdot b \equiv c \mod m.$$

Bemerkung 5.4.5. *Nach Beispiel 4.2.13 hat $\mathbb{Z}_p$ genau p Elemente, sodass $\mathbb{Z}_p$, für p eine Primzahl, ein Körper mit p Elementen ist. Wir sprechen von* ***endlichen Körpern*** *im Gegensatz zu $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$, die ja unendlich (genauer abzählbar unendlich im Falle von $\mathbb{Q}$ bzw. überabzählbar unendlich in den Fällen $\mathbb{R}$ und $\mathbb{C}$; vgl. Abschnitt 4.4) viele Elemente besitzen.*

Im folgenden wenden wir uns wieder dem Studium der abstrakten Struktur zu.

Proposition 5.4.6 (Rechenregeln für Körper)**.** *Ist $(K, +, \cdot)$ ein Körper, $a, b \in K$, so gelten die Rechenregeln*

(1) $(ab)^{-1} = a^{-1}b^{-1}$.
(2) $(-a)^{-1} = -a^{-1}$,
(3) $ab = 0$ *impliziert* $a = 0$ *oder* $b = 0$ *(Nullteilerfreiheit)*
(4) *Die Gleichung* $ax = b$ *hat für* $a \neq 0$ *die eindeutige Lösung* $x = a^{-1}b$.

Beweis.

(1) Die Aussage folgt aus Proposition 5.2.25 (1) und der Kommutativität der Multiplikation oder direkt aus
$$(ab)(a^{-1}b^{-1}) = aba^{-1}b^{-1} = aa^{-1}bb^{-1} = 1 \cdot 1 = 1$$
und der Eindeutigkeit der Inversen.
(2) Zunächst gilt $(-1)^{-1} = -1$, denn
$$-1 = (-1) \cdot 1 = (-1)((-1)(-1)^{-1}) = ((-1)(-1))(-1)^{-1} = (-1)^{-1},$$
wobei wir im letzten Schritt Proposition 5.3.8 (3) verwendet haben. Schließlich erhalten wir unter Verwendung von Proposition 5.3.8 (4) und Eigenschaft (1)
$$(-a)^{-1} = ((-1)a)^{-1} = (-1)^{-1}a^{-1} = (-1)a^{-1} = -a^{-1}.$$
(3) Da $(K \setminus \{0\}, \cdot)$ eine Gruppe bildet, folgt aus $a, b \in K \setminus \{0\}$, dass auch $ab \in K \setminus \{0\}$.
(4) Falls $b \neq 0$ folgt die Aussage aus Proposition 5.2.25 (3) (wiederum weil $(K \setminus \{0\}, \cdot)$ eine Gruppe ist). Ist $b = 0$, so folgt aus (3), dass $x = 0$; also gilt auch hier die Eindeutigkeit.

□

Bemerkung 5.4.7. *Proposition 5.4.6 (3) zeigt, dass Körper nullteilerfrei also Integritätsbereiche sind; also sind Körper die speziellste aller hier vorgestellten Strukturen (vgl. Abbildung 5.2).*

Analog zu Ringen kann man auch wieder Unterkörper definieren:

Definition 5.4.8 (Unterkörper)**.** *Eine Teilmenge $Q \subseteq K$ eines Körpers $(K, +, \cdot)$ heißt* Unterkörper*, wenn $(Q, +, \cdot)$ selbst ein Körper ist.*

Beispiel 5.4.9 (Unterkörper)**.** *Die rationalen Zahlen $\mathbb{Q}$ sind ein Unterkörper der reellen Zahlen $\mathbb{R}$. Diese sind wiederum ein Unterkörper der komplexen Zahlen $\mathbb{C}$.*

Proposition 5.4.10 (Charakterisierung von Unterkörpern)**.** *Eine Teilmenge Q eines Körpers $(K, +, \cdot)$ ist genau dann ein Unterkörper, wenn eine der folgenden äquivalenten Bedingungen gilt.*

(1) *Für je zwei Elemente* $a, b \in Q$ *ist sowohl* $a - b \in Q$ *als auch, sofern* $b \neq 0$, $ab^{-1} \in Q$.
(2) *Für je drei Elemente* $a, b, c \in Q$ *mit* $c \neq 0$ *ist auch* $(a - b)c^{-1} \in Q$.

Beweis. Dies folgt aus Proposition 5.2.28 für $(K, +)$ und $(K, \cdot)$. Ferner beachte man, dass $(a - 0)c^{-1} = ac^{-1}$ und $(a - b)1^{-1} = a - b$ gelten. □

Beispiel 5.4.11 ($\mathbb{Q}[\sqrt{2}]$). *Seien auf*

$$K = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\} \subset \mathbb{R}$$

die folgenden Operationen definiert:

$$\begin{aligned}(a_1 + b_1\sqrt{2}) \oplus (a_2 + b_2\sqrt{2}) &= (a_1 + a_2) + (b_1 + b_2)\sqrt{2}\\ (a_1 + b_1\sqrt{2}) \otimes (a_2 + b_2\sqrt{2}) &= (a_1a_2 + 2b_1b_2) + (a_2b_1 + a_1b_2)\sqrt{2}.\end{aligned}$$

Bei genauerer Betrachtung sehen wir, dass $\oplus$ *und* $\otimes$ *genau die von* $\mathbb{R}$ *ererbten Operationen* $+$ *und* $\cdot$ *sind. Wir untersuchen also:*

$$(a_1 + b_1\sqrt{2}) - (a_2 + b_2\sqrt{2}) = (a_1 - a_2) + (b_1 - b_2)\sqrt{2} \in K,$$

und für $(a_2, b_2) \neq (0, 0)$

$$\begin{aligned}(a_1 + b_1\sqrt{2})(a_2 + b_2\sqrt{2})^{-1} = \frac{a_1 + b_1\sqrt{2}}{a_2 + b_2\sqrt{2}} = \frac{(a_1 + b_1\sqrt{2})(a_2 - b_2\sqrt{2})}{a_2^2 - 2b_2^2} &=\\ = \frac{a_1a_2 - 2b_1b_2}{a_2^2 - 2b_2^2} + \frac{a_2b_1 - a_1b_2}{a_2^2 - 2b_2^2}\sqrt{2}&.\end{aligned}$$

Dieses Ergebnis liegt in K, *sofern* $a_2^2 - 2b_2^2 \neq 0$ *gilt. Dies ist aber wahr, da nicht beide* a_2 *und* b_2 *gleich Null sein dürfen. Darüber hinaus gilt noch, dass* $a_2^2 \neq 2b_2^2$ *sein muss, weil* a_2 *und* b_2 *rational sind,* $\sqrt{2}$ *aber irrational ist. Daher sind die Voraussetzungen von Proposition 5.4.10 erfüllt, und* K *ist in der Tat ein Unterkörper von* $\mathbb{R}$. *Wir schreiben auch* $K = \mathbb{Q}[\sqrt{2}]$.

Nach den Definitionen der Struktur und den Beispielen müssen wir uns ein weiteres Mal um die Abbildungen kümmern. Das Prinzip ist wieder dasselbe wie schon zuvor. Jeder Körper ist ein Ring mit zusätzlichen Eigenschaften, also ist ein Körperhomomorphismus — bitte raten! — genau, ein Ringhomomorphismus, der auch diese zusätzlichen Eigenschaften respektiert.

Definition 5.4.12 (Körperhomomorphismus). *Es seien* $(K, +, \cdot)$ *und* $(K', \oplus, \otimes)$ *zwei Körper.*

(i) *Ein* Körperhomomorphismus *ist ein Gruppenhomomorphismus* $f : (K, +) \to (K', \oplus)$, *der auch noch ein Gruppenhomomorphismus* $f : (K \setminus \{0\}, \cdot) \to (K' \setminus \{0\}, \otimes)$ *ist.*
(ii) *Ist* f *bijektiv, so nennt man die Abbildung* Körperisomorphismus *und sagt, die beiden Körper* K *und* K' *sind* isomorph.

Beispiel 5.4.13. *Definieren wir auf* $\mathbb{Q} \times \mathbb{Q}$ *die Verknüpfungen*

$$\begin{aligned}(a_1, a_2) + (b_1, b_2) &:= (a_1 + b_1, a_2 + b_2)\\ (a_1, a_2) \cdot (b_1, b_2) &:= (a_1b_1 + 2a_2b_2, a_1b_2 + a_2b_1)\end{aligned}$$

dann ist $(\mathbb{Q} \times \mathbb{Q}, +, \cdot)$ *ein Körper.*

Wir überprüfen das, indem wir die Körperaxiome nachrechnen:

(K1) *Seien* $a, b, c \in \mathbb{Q} \times \mathbb{Q}$. *Wir finden*

$$\begin{aligned}(a+b)+c &= \big((a_1,a_2)+(b_1,b_2)\big)+(c_1,c_2) = (a_1+b_1,a_2+b_2)+(c_1,c_2) = \\ &= (a_1+b_1+c_1,a_2+b_2+c_2) = (a_1,a_2)+(b_1+c_1,b_2+c_2) = \\ &= (a_1,a_2)+\big((b_1,b_2)+(c_1,c_2)\big) = a+(b+c).\end{aligned}$$

(K2) *Nehmen wir beliebige* $a, b \in \mathbb{Q} \times \mathbb{Q}$. *Es gilt*

$$\begin{aligned}a+b &= (a_1,a_2)+(b_1,b_2) = (a_1+b_1,a_2+b_2) = \\ &= (b_1+a_1,b_2+a_2) = (b_1,b_2)+(a_1,a_2) = b+a.\end{aligned}$$

(K3) *Für* $0 := (0,0) \in \mathbb{Q} \times \mathbb{Q}$ *gilt*

$$a+0 = (a_1,a_2)+(0,0) = (a_1+0,a_2+0) = (a_1,a_2) = a.$$

(K4) *Sei* $a \in \mathbb{Q} \times \mathbb{Q}$ *gegeben. Wir definieren* $-a := (-a_1,-a_2) \in \mathbb{Q} \times \mathbb{Q}$ *und berechnen*

$$a+(-a) = (a_1,a_2)+(-a_1,-a_2) = (a_1+(-a_1),a_2+(-a_2)) = (0,0) = 0.$$

(K5) *Für alle* $a, b, c \in \mathbb{Q} \times \mathbb{Q}$ *folgt*

$$\begin{aligned}(ab)c &= \big((a_1,a_2)(b_1,b_2)\big)(c_1,c_2) = (a_1b_1+2a_2b_2,a_1b_2+a_2b_1)(c_1,c_2) = \\ &= \big((a_1b_1+2a_2b_2)c_1+2(a_1b_2+a_2b_1)c_2,(a_1b_1+2a_2b_2)c_2+(a_1b_2+a_2b_1)c_1\big) = \\ &= \big(a_1b_1c_1+2a_2b_2c_1+2a_1b_2c_2+2a_2b_1c_2,a_1b_1c_2+a_1b_2c_1+a_2b_1c_1+2a_2b_2c_2\big) = \\ &= \big(a_1(b_1c_1+2b_2c_2)+2a_2(b_1c_2+b_2c_1),a_1(b_1c_2+b_2c_1)+a_2(b_1c_1+2b_2c_2)\big) = \\ &= (a_1,a_2)(b_1c_1+2b_2c_2,b_1c_2+b_2c_1) = \\ &= (a_1,a_2)\big((b_1,b_2)(c_1,c_2)\big) = a(bc).\end{aligned}$$

(K6) *Es seien wieder* $a, b \in \mathbb{Q} \times \mathbb{Q}$. *Wir rechnen nach:*

$$\begin{aligned}ab &= (a_1,a_2)(b_1,b_2) = (a_1b_1+2a_2b_2,a_1b_2+a_2b_1) = \\ &= (b_1a_1+2b_2a_2,b_1a_2+b_2a_1) = (b_1,b_2)(a_1,a_2) = ba.\end{aligned}$$

(K7) *Wir definieren* $1 := (1,0) \in \mathbb{Q} \times \mathbb{Q}$. *Klarerweise gilt* $0 \neq 1$, *und außerdem für* $a \in \mathbb{Q} \times \mathbb{Q}$

$$a1 = (a_1,a_2)(1,0) = (a_11+0,0+a_21) = (a_1,a_2) = a.$$

(K8) *Sei* $0 \neq a \in \mathbb{Q} \times \mathbb{Q}$ *gegeben. Wir definieren* $a^{-1} := \left(\frac{a_1}{a_1^2-2a_2^2}, \frac{-a_2}{a_1^2-2a_2^2}\right)$. *Es gilt* a^{-1} *ist für alle* $a \neq 0$ *definiert. Zu diesem Zweck muss* $a_1^2 - 2a_2^2 \neq 0$ *gelten. Das folgende Argument beweist das: Sei* $a_1^2 = 2a_2^2$. *Dann gilt auch, falls* $a_2 \neq 0$ *stimmt, dass* $(a_1/a_2)^2 = 2$. *Die linke Seite dieser Gleichung ist das Quadrat einer rationale Zahl. Das Quadrat einer rationalen Zahl kann aber niemals gleich* 2 *sein, da andernfalls* $\sqrt{2}$ *rational wäre. Folglich ist* $a_2 = 0$. *Dann haben wir aber auch* $a_1 = 0$ *und damit* $a = 0$, *was wir ausgeschlossen haben.*

Es ist a^{-1} *also für alle* $a \neq 0$ *definiert. Nun können wir rechnen*

$$\begin{aligned}aa^{-1} &= (a_1,a_2)(a_1/(a_1^2-2a_2^2),-a_2/(a_1^2-2a_2^2)) = \\ &= \big((a_1^2-2a_2^2)/(a_1^2-2a_2^2),0\big) = (1,0) = 1.\end{aligned}$$

(K9) *Seien wieder $a, b, c \in \mathbb{Q} \times \mathbb{Q}$. Auch das letzte Axiom ist eine längliche Rechnung:*

$$\begin{aligned} ab + ac &= (a_1, a_2)(b_1, b_2) + (a_1, a_2)(c_1, c_2) = \\ &= (a_1 b_1 + 2a_2 b_2, a_1 b_2 + a_2 b_1) + (a_1 c_1 + 2a_2 c_2, a_1 c_2 + a_2 c_1) = \\ &= (a_1 b_1 + a_1 c_1 + 2a_2 b_2 + 2a_2 c_2, a_1 b_2 + a_1 c_2 + a_2 b_1 + a_2 c_1) = \\ &= \big(a_1(b_1 + c_1) + 2a_2(b_2 + c_2), a_1(b_2 + c_2) + a_2(b_1 + c_1)\big) = \\ &= (a_1, a_2)(b_1 + c_1, b_2 + c_2) = (a_1, a_2)\big((b_1, b_2) + (c_1, c_2)\big) = a(b + c). \end{aligned}$$

Wir haben also alle Eigenschaften nachgeprüft, und daher ist $(\mathbb{Q}\times\mathbb{Q}, +, \cdot)$ wirklich ein Körper.

Als nächstes definieren wir eine Abbildung $f : \mathbb{Q}\times\mathbb{Q} \to \mathbb{Q}[\sqrt{2}]$ durch $(a_1, a_2) \mapsto a_1 + a_2\sqrt{2}$. Die Abbildung ist offensichtlich bijektiv, und es gilt

$$\begin{aligned} f(a + b) &= f((a_1, a_2) + (b_1, b_2)) = f((a_1 + b_1, a_2 + b_2)) = (a_1 + b_1) + (a_2 + b_2)\sqrt{2} = \\ &= (a_1 + a_2\sqrt{2}) \oplus (b_1 + b_2\sqrt{2}) = f(a) \oplus f(b), \\ f(ab) &= f((a_1, a_2)(b_1, b_2)) = f((a_1 b_1 + 2a_2 b_2, a_1 b_2 + a_2 b_1)) = \\ &= (a_1 b_1 + 2a_2 b_2) + (a_1 b_2 + a_2 b_1)\sqrt{2} = (a_1 + a_2\sqrt{2}) \otimes (b_1 + b_2\sqrt{2}) = f(a) \otimes f(b), \end{aligned}$$

Daher ist f ein Körperisomorphismus, deshalb ist $(\mathbb{Q}\times\mathbb{Q}, +, \cdot)$ isomorph zu $\mathbb{Q}[\sqrt{2}]$. Die beiden Strukturen sind also identisch bis auf Umbenennen der Elemente.

Damit schließen wir unsere algebraischen Untersuchungen ab. Aufbauend auf den Körperaxiomen werden in der Linearen Algebra darüber hinaus gehend neue Strukturen erschaffen werden wie die eines **Vektorraumes**. Für die Analysis werden wir genauere Untersuchungen der rationalen, reellen und komplexen Zahlen benötigen. Alle diese Mengen sind mit den bereits bekannten Rechengesetzen ausgestattet und bilden Körper.

In der (höheren) Algebra wird mit der genaueren Untersuchung der Strukturen selbst fortgefahren werden. Man wird Fragen stellen wie: Welche Arten von Gruppen, Ringen, Körpern gibt es? Kann man alle endlichen Gruppen, Ringe, Körper finden? Alle diese Fragen und viele andere werden zum Ausbau der mathematischen Theorie beitragen und teilweise tief gehende Resultate hervorbringen.

KAPITEL 6

Zahlenmengen

Dieses letzte Kapitel führt uns zurück zu den konkreten Dingen. Wir werden uns wieder mit Zahlen beschäftigen. Nach der Wanderung durch die Grundbegriffe der Mathematik wie Logik, Mengenlehre und elementare Algebra, kehren wir zurück zu den Anfängen der Mathematik.

Wir haben im Verlauf der vergangenen Kapitel häufig die verschiedenen Zahlenmengen als Beispiel verwendet. Wir sind durch den täglichen Umgang mit den Zahlen überzeugt, sie zu beherrschen, ihre Eigenschaften zu kennen. Es scheint uns, dass wir mit ihnen völlig vertraut sind.

Doch trügt der Schein nicht? Was ist $\sqrt{2}$ eigentlich? Haben wir diese Zahl wirklich verstanden? Das Hinterfragen dessen, was wir zu wissen glauben, die kritische Analyse, ist eines der Grundprinzipien der modernen Naturwissenschaft. Im Gegensatz zu zuvor wollen wir nun den bereits mathematisch geschulten Blick auf das richten, was wir bereits zu kennen glaubten. Wir werden unser Wissen über Mengenlehre und mathematische Strukturen anzuwenden versuchen und die Zahlen selbst in einem etwas veränderten Licht betrachten.

Das Kapitel ist in zwei Teile geteilt, die munter durcheinander gemischt erscheinen. Nur Randstreifen trennen den vergleichsweise beschreibenden Zugang zu den Zahlenmengen vom axiomatischen Zugang, bei dem die Zahlenmengen direkt aus dem Zermelo–Fraenkelschen Axiomensystem ZFC konstruiert werden.

6.1. Die natürlichen Zahlen $\mathbb{N}$

> „Die natürlichen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk.“ Leopold Kronecker (1823–1891)

> „Die natürlichen Zahlen sind freie Schöpfungen des menschlichen Geistes.“ Richard Dedekind (1831–1916)

Die natürlichen Zahlen sind schon seit langer Zeit bekannt. Sie entstanden historisch gesehen aus dem natürlichen Zählbegriff. Die Null als Zeichen und als eigenständige Zahl wurde aber erst Ende des Mittelalters akzeptiert. Wahrscheinlich stammt das Zeichen aus Indien. Die Null ist Element der natürlichen Zahlen. Wir definieren das so, und auch die DIN Norm 5473 stimmt damit überein. Demnach ist

$$\mathbb{N} := \{0, 1, 2, 3, 4, \dots\}.$$

Mit dieser „Definition“ erhalten Sie die richtige Vorstellung von $\mathbb{N}$, doch mathematisch exakt ist sie nicht. Was sollen die Punkte bedeuten? Wir folgen hier Giuseppe Peano (1858–1932) der die einfachste axiomatische Beschreibung von $\mathbb{N}$, die die Punkte in obiger Beschreibung exakt macht, im 19. Jahrhundert gegeben hat.

Bemerkung 6.1.1 ($\mathbb{N}$)**.** *Die natürlichen Zahlen sind eine Menge $\mathbb{N}$ zusammen mit einer Vorschrift S die die **Peano Axiome** erfüllt:*

(PA1) 0 *ist eine natürliche Zahl, d.h.* $0 \in \mathbb{N}$*,*

(PA2) *Jeder natürlichen Zahl wird genau eine natürliche Zahl $S(n)$ zugeordnet, die ihr* Nachfolger *genannt wird, d.h.*

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

(PA3) 0 *ist kein Nachfolger, i.e.,*

$$\forall n \in \mathbb{N} : \neg(S(n) = 0),$$

(PA4) *Sind zwei natürliche Zahlen verschieden, so sind das auch ihre Nachfolger, d.h.*

$$\forall n \in \mathbb{N} : \forall m \in \mathbb{N} : (S(n) = S(m)) \Rightarrow n = m,$$

(PA5) *Enthält eine Menge M natürlicher Zahlen die Zahl* 0 *und mit jeder Zahl ihren Nachfolger, so ist $M = \mathbb{N}$, genauer: Ist $0 \in A \subseteq \mathbb{N}$ und gilt: $n \in A \Rightarrow S(n) \in A$, so gilt schon $A = \mathbb{N}$.*

Das letzte Axiom postuliert übrigens das Induktionsprinzip. Ganz befriedigend ist diese Beschreibung noch immer nicht, da wir ja gerne hätten, dass die Menge $\mathbb{N}$ existiert und eindeutig bestimmt ist, d.h. dass es genau eine Menge $\mathbb{N}$ gibt, die obigen Axiome erfüllt. Das dem tatsächlich so ist, wird in Abschnitt 6.1.1 aus den ZFC Axiomen der Mengenlehre bewiesen.

Auf der Menge $\mathbb{N}$ sind die beiden Operationen $+$, die Addition und $\cdot$, die Multiplikation definiert, wobei $(N, +)$ eine kommutative Halbgruppe mit Nullelement 0 und $(N, \cdot)$ eine kommutative Halbgruppe mit Einselement 1 ist. (In der Sprache von Definition 5.3.3 ist $\mathbb{N}$ ein kommutativen Halbring mit 0 und 1 (ein Dioid) ohne Nullteiler (siehe Beispiel 5.3.4).)

Ferner ist eine Totalordnung (siehe Definition 4.2.14 (iii)) $\leq$ erklärt, die verträglich mit den Verknüpfungen ist, d.h. es gelten die beiden **Ordnungsaxiome**

(O1) Ist $a \leq b$, so ist für alle $c \in \mathbb{N}$ auch $a + c \leq b + c$,
(O2) Sind $x > 0$ und $y > 0$, so ist $xy > 0$.

Die Menge $\mathbb{N}$ ist also ein geordnetes Dioid bezüglich Addition und Multiplikation. Sie ist die kleinstmächtige unendliche Menge, und es gilt $|\mathbb{N}| = \aleph_0$ (siehe Ende Kapitel 4).

6.1.1. Mengentheoretische Konstruktion von $\mathbb{N}$. Die Konstruktion der natürlichen Zahlen aus ZFC (den Axiomen der Mengenlehre von Zermelo und Fraenkel) funktioniert folgendermaßen.

Wir definieren

$$\begin{aligned} 0 &:= \emptyset \\ 1 &:= S(0) = 0 \cup \{0\} = \{\emptyset\} \\ 2 &:= S(1) = 1 \cup \{1\} = \big\{\emptyset, \{\emptyset\}\big\} \\ 3 &:= S(2) = 2 \cup \{2\} = \Big\{\emptyset, \{\emptyset\}, \big\{\emptyset, \{\emptyset\}\big\}\Big\} \end{aligned}$$

$$n := \begin{cases} \emptyset & n = 0 \\ S(n) = n \cup \{n\} & n \neq 0 \end{cases}$$

Somit erhalten wir in Kurzform $0 = \emptyset, 1 = \{0\}, 2 = \{0, 1\}$ und allgemein $n = \{0, 1, \ldots, n-1\}$. Jede Zahl ist also identifiziert als die Menge, die alle kleineren Zahlen enthält.

So stellen wir uns das jedenfalls vor. Die Konstruktoren, die wir verwendet haben, sind alle bereits definiert, und ZF7 garantiert uns, dass eine Menge existiert, die alle diese Zahlen n enthält. Leider wissen wir zwei Dinge noch nicht, nämlich ob es eine Menge gibt die **genau**

alle diese Zahlen enthält, denn nur dann ist sie eindeutig bestimmt (und das, was wir uns naiv unter $\mathbb{N}$ vorstellen).

Theorem 6.1.2. *Sei die Nachfolgereigenschaft* ψ

$$\psi(Y) := \forall X : (\emptyset \in Y \wedge (X \in Y \Rightarrow S(X) \in Y)).$$

gegeben. Dann gilt

$$\exists! \mathbb{N} : \forall M : (\psi(\mathbb{N}) \wedge (\psi(M) \Rightarrow \mathbb{N} \subseteq M)).$$

Mit anderen Worten, es gibt genau eine Menge der natürlichen Zahlen. Sie ist die kleinste Menge, die die Nachfolgereigenschaft besitzt.

Beweis. Wegen ZF7 gibt es eine Menge Z, die die Eigenschaft $\psi(Z)$ besitzt. Wir definieren $\mathcal{N} := \{M \in \mathbb{P}Z \mid \psi(M)\}$. Sei nun $\mathbb{N} := \bigcap \mathcal{N}$. (Für eine Mengenfamilie $\mathcal{F}$ ist $\bigcap \mathcal{F}$ definiert durch $\bigcap \mathcal{F} := \{x \in \bigcup \mathcal{F} \mid \forall F \in \mathcal{F} : (x \in F)\}$.)

Dann gilt $\forall M \in \mathcal{N} : \psi(M)$, und daher $\forall M \in \mathcal{N} : (\emptyset \in M)$, also auch $\emptyset \in \mathbb{N}$. Ferner wissen wir $X \in \mathbb{N} \Rightarrow (\forall M \in \mathcal{N} : (X \in M))$, deshalb $\forall M \in \mathcal{N} : (S(X) \in M)$, was wiederum $S(X) \in \mathbb{N}$ zur Folge hat. Daher gilt $\psi(\mathbb{N})$.

Um Eindeutigkeit zu zeigen, nehmen wir an, dass $\exists M : \psi(M)$ (etwa ein M, das nicht Teilmenge von Z ist). Mit denselben Argumenten wie oben können wir zeigen, dass $\psi(Z \cap M)$ gilt, sowie $(Z \cap M) \subseteq M$ und $\mathbb{N} \subseteq Z \cap M$, was $\mathbb{N} \subseteq M$ impliziert. □

Korollar 6.1.3. *Es gilt das Induktionsprinzip*

$$\forall M \in \mathbb{P}\mathbb{N} : (\psi(M) \Rightarrow M = \mathbb{N}).$$

Beweis. Sei $M \in \mathbb{P}\mathbb{N}$ beliebig. Gilt $\psi(M)$, so ist $M \subseteq \mathbb{N}$, und nach Voraussetzung gilt $\mathbb{N} \subseteq M$, und daher ist $M = \mathbb{N}$. □

Diese (etwas unintuitive) Version der Konstruktion der natürlichen Zahlen ist viel mächtiger als die Definitionen, die im neunzehnten Jahrhundert gegeben wurden. Das sieht man allein daran, dass man das Induktionsprinzip *beweisen* kann und nicht als Axiom fordern muss. Alle fünf von Peano für die natürlichen Zahlen angegebenen Axiome kann man leicht überprüfen.

Proposition 6.1.4. *Die Menge der natürlichen Zahlen* $\mathbb{N}$ *erfüllt die Peano Axiome.*

Beweis. Die Axiome PA1 und PA2 gelten wegen der Definition von $\mathbb{N}$ und PF5 haben wir in Korollar 6.1.3 gezeigt. Es bleiben also nur noch PA3 und PA4.

PA3 beweisen wir indirekt. Sei also $n \in \mathbb{N}$ gegeben mit $S(n) = 0$. Dann ist $S(n) = n \cup \{n\} = \emptyset$, doch es gilt $n \in S(n)$, und daher $S(n) \neq \emptyset$. Dieser Widerspruch beweist PA3.

Zum Beweis von PA4 nehmen wir an, dass $m, n \in \mathbb{N}$ sind mit $S(n) = S(m)$. Sei $k \in n$. Dann ist auch $k \in n \cup \{n\} = S(n) = S(m) = m \cup \{m\}$, also $k \in m$ oder $k \in \{m\}$ wegen der Eigenschaften von $\cup$. Weil aber die Menge $\{m\}$ nur ein Element, nämlich m enthält, folgt daraus die Tatsache $k \in m \vee k = m$. Ist $k = m$, so gilt $n \in k \vee n = k$, weil $n \in S(n) = S(m) = S(k)$, und daher widerspricht entweder $\{n, k\}$ oder $\{k\}$ dem Fundierungsaxiom ZF9. Daher gilt $k \in m$ und auch $n \subseteq m$. Analog zeigt man durch Vertauschen von m und n die Relation $m \subseteq n$, und es folgt $n = m$. Dies beweist auch PA4, und wir sind fertig. □

Die arithmetischen Operationen $+$ und $\cdot$ definiert man ebenfalls über S. Die Totalordnung $\leq$ ist einfach

$$m \leq n :\Leftrightarrow (m \in n \vee m = n).$$

Proposition 6.1.5. *Die Relation* $\leq$ *ist eine Totalordnung.*

Beweis. Reflexivität und Transitivität sind offensichtlich, und wäre die Antisymmetrie nicht erfüllt, dann existierten zwei natürliche Zahlen $m \neq n \in \mathbb{N}$ mit $n \leq m$ und $m \leq n$, also mit $m \in n$ und $n \in m$. Gäbe es diese Zahlen, dann könnten wir die Menge $\{m, n\}$ bilden, welche ZF9 widerspräche. Daher ist die Antisymmetrie erfüllt, und $\leq$ ist eine Halbordnung.

Um zu beweisen, dass $\leq$ eine Totalordnung ist, müssen wir zeigen, dass für je zwei Zahlen $m, n \in \mathbb{N}$ entweder $m < n$ oder $m = n$ oder $m > n$ gilt.

Beweisen wir zwei Hilfsresultate zuerst:

HB1. $\forall m, n \in \mathbb{N} : (m \in n \Rightarrow S(m) \subseteq n)$.
Sei $M := \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : (m \in n \Rightarrow S(m) \subseteq n)\}$. Die 0 erfüllt die Bedingung trivialerweise, daher ist $0 \in M$. Sei nun $n \in M$. Gilt $m \in S(n) = n \cup \{n\}$, so ist entweder $m = n$ oder $m \in n$. Ist $m = n$, so ist $S(m) = S(n)$ und daher gilt $S(m) \subseteq S(n)$. Ist hingegen $m \in n$, so gilt wegen $n \in M$ auch $S(m) \subseteq n \subseteq S(n)$, und somit gilt immer $S(m) \subseteq S(n)$. Daher ist auch $S(n) \in M$ und wegen Korollar 6.1.3 folgt $M = \mathbb{N}$. Dies beweist HB1.

HB2. $\forall m, n \in \mathbb{N} : ((m \subseteq n \wedge m \neq n) \Rightarrow m \in n)$.
Sei $M := \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : ((m \subseteq n \wedge m \neq n) \Rightarrow m \in n)\}$. Ist $m \subseteq 0$, so ist $m = 0$ und daher $0 \in M$. Sei nun $n \in M$. Wir betrachten $S(n)$, und daher sei $m \in \mathbb{N}$ mit $m \subseteq S(n) \wedge m \neq S(n)$. Ist $k \in m$, so gilt wegen $S(n) = n \cup \{n\}$, dass entweder $k \in n$ oder $k = n$. Ist $k = n$, so ist $n \in m$ und wegen HB1 folgt dann $S(n) \subseteq m$. Dies ist aber ein Widerspruch zu $m \subseteq S(n) \wedge m \neq S(n)$. Daher gilt $\forall k \in m : k \in n$, also $m \subseteq n$. Ist $m = n$, dann haben wir $m \in n \cup \{n\} = S(n)$. Sonst gilt $m \subseteq n \wedge m \neq n$, und weil $n \in M$ vorausgesetzt ist auch $m \in n$. Dies impliziert aber $m \in S(n)$, und $S(n) \in M$. Aus Korollar 6.1.3 folgt $M = \mathbb{N}$, was HB2 beweist.

Sei $M = \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : (m < n \vee m = n \vee n < m)\}$. Betrachten wir zuerst 0. Ist $0 \neq n$, so gilt $0 = \emptyset \subseteq n$, also $0 \in n$ wegen HB2, und daher $0 \in M$. Sei nun $n \in M$. Betrachten wir $S(n)$. Sei $m \in \mathbb{N}$ gegeben. Gelten $m \in n$ oder $m = n$, so haben wir $m \in n \cup \{n\} = S(n)$. Gilt andererseits $n \in m$, so folgt aus HB1, dass $S(n) \subseteq m$. Ist $S(n) \neq m$, so ist $S(n) \in m$ wegen HB2. Es gilt also $m \in S(n) \vee m = S(n) \vee S(n) \in m$, und daher $S(n) \in M$. Verwenden wir ein weiteres Mal Korollar 6.1.3, so sehen wir $M = \mathbb{N}$ und wir sind fertig. □

Die arithmetische Operation $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ sei unser nächstes Opfer. Wir definieren

$$n + 0 = n$$
$$n + S(m) = S(n + m)$$

und finden das folgende Resultat

Proposition 6.1.6. *Es gibt genau eine Abbildung* $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, *die obige rekursive Definition erfüllt.*

Beweis. Beginnen wir mit der Eindeutigkeit. Seien $+$ und $\boxplus$ zwei Funktionen, die die rekursive Definition erfüllen. Setzen wir $M := \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : (m + n = m \boxplus n)\}$. Natürlich ist $0 \in M$ wegen $n + 0 = n = n \boxplus 0$. Sei nun $n \in N$, dann haben wir für $m \in \mathbb{N}$ die Gleichung $m + S(n) = S(m + n) = S(m \boxplus n)$ wegen $n \in N$ und $S(m \boxplus n) = m \boxplus S(n)$, und daher $S(n) \in M$. Aus Korollar 6.1.3 folgt $M = \mathbb{N}$, und daher ist $+ = \boxplus$ als Teilmenge von $(\mathbb{N} \times \mathbb{N}) \times \mathbb{N}$. Wir dürfen noch nicht von Abbildung reden, da wir die Abbildungseigenschaft noch nicht nachgewiesen haben. Dies können wir mit einem ähnlichen Induktionsargument erreichen.

Sei für jedes $m \in \mathbb{N}$ die „Abbildung“ $+_m : \mathbb{N} \to \mathbb{N}$ definiert durch $+_0(n) = n$ und $+_{S(m)}(n) = S(+_m(n))$. Dies macht $+_m$ zu einer Relation, aber wir werden unten die Abbildungseigenschaft nachweisen:
Sei $M := \{m \in \mathbb{N} \mid \forall n \in \mathbb{N} : \forall j \leq m : \exists! k \in \mathbb{N} : (+_j(n) = k)\}$. Wegen $\forall n \in \mathbb{N} : (+_0(n) = n)$ folgt sofort $0 \in M$. Ist $m \in M$, dann ist $+_0(n) = n$ eindeutig. Sei also $j \leq m$. Dann existiert für beliebiges $n \in \mathbb{N}$ genau ein k mit $+_j(n) = k$. Also ist für $S(j)$ die Beziehung

$+_{S(j)}(n) = S(+_j(n)) = S(k)$ erfüllt. Somit ist auch $S(m) \in M$, da für $j \in \mathbb{N}$ mit $j \leq S(m)$ entweder $j = 0$ ist oder ein $j' \in \mathbb{N}$ existiert mit $j = S(j')$ und $j' \leq m$. Somit impliziert Korollar 6.1.3 aber $M = \mathbb{N}$. Daher ist für jedes $m \in \mathbb{N}$ die Relation $+_m$ tatsächlich eine Abbildung, und $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ ist dann als Abbildung definiert durch $n + m = +_m(n)$ für alle $m, n \in \mathbb{N}$. □

Mit ähnlichen Induktionsbeweisen zeigt man noch, dass die arithmetische Operation $\cdot$: $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ rekursiv definiert werden kann durch

$$n \cdot 0 = 0$$
$$n \cdot S(m) = (n \cdot m) + n$$

Theorem 6.1.7. *Die natürlichen Zahlen* $(\mathbb{N}, +, \cdot)$ *bilden einen kommutativen Halbring mit* 0 *und* 1.

Beweis. Zeigen wir zunächst, dass $(\mathbb{N}, +)$ eine kommutative Halbgruppe ist.

BH1: $\forall n \in \mathbb{N} : S(m) + n = m + S(n)$.
Sei $M := \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : S(m) + n = m + S(n)\}$. Es gilt $S(0) + 0 = S(0)$ und $0 + S(0) = S(0 + 0) = S(0)$ und daher $0 \in M$. Sei nun $n \in M$. Wir betrachten $S(n)$ und erhalten für $m \in \mathbb{N}$ die Beziehung $S(m) + S(n) = S(S(m) + n) = S(m + S(n)) = m + S(S(n))$ nach Definition von $+$ und weil $n \in M$. Daher ist auch $S(n) \in M$ und Korollar 6.1.3 liefert uns $M = \mathbb{N}$.

BH2: $\forall n \in \mathbb{N} : 0 + n = n$.
Sei $M := \{n \in \mathbb{N} \mid 0 + n = n\}$. Dann ist $0 \in M$ wegen $0 + 0 = 0$. Sei nun $n \in M$ und betrachten wir $S(n)$. Wir erhalten $0 + S(n) = S(0 + n) = S(n)$ aus der Definition von $+$ und weil $n \in M$. Daraus und aus der Definition folgt, dass 0 ein Nullelement ist.

KG(+): $\forall n, m \in \mathbb{N} : n + m = m + n$.
Diese Beziehung zeigen wir ebenfalls mit Induktion. Sei $M := \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : m + n = n + m\}$. Wegen BH2 und der Definition von $+$ gilt für alle $n \in \mathbb{N}$ die Gleichung $0 + n = n + 0$ und daher $0 \in M$. Sei nun $n \in M$. Dann rechnen wir für beliebiges $m \in \mathbb{N}$ wie folgt: $S(n) + m = n + S(n) = S(n + m) = S(m + n) = m + S(n)$. Zweimal haben wir die Definition von $+$ verwendet und je einmal die Tatsache $n \in M$ und BH1. Daher ist $S(n) \in M$, und wegen Korollar 6.1.3 gilt $M = \mathbb{N}$. Daher ist $+$ kommutativ.

AG(+): $\forall k, m, n \in \mathbb{N} : (k + n) + m = k + (n + m)$.
Ein weiterer Induktionsbeweis wird uns das Assoziativgesetz zeigen. Wir definieren $M := \{m \in \mathbb{N} : \forall k, n \in \mathbb{N} : (k + n) + m = k + (n + m)\}$, und wieder gilt $0 \in M$, diesmal wegen $(k + n) + 0 = k + n = k + (n + 0)$. Ist $m \in M$, dann rechnen wir für beliebige $k, n \in \mathbb{N}$

$$\begin{aligned}(k + n) + S(m) &= S((k + n) + m) = S(k + (n + m)) = \\ &= k + S(n + m) = k + (n + S(m)).\end{aligned}$$

Das beweist $S(m) \in M$ und damit $M = \mathbb{N}$ wegen Korollar 6.1.3. Also ist $+$ assoziativ und $(\mathbb{N}, +)$ ein kommutatives Monoid.

BH3: $\forall n \in \mathbb{N} : 0 \cdot n = 0$.
Induktion mit $M = \{n \in \mathbb{N} \mid 0 \cdot n = 0\}$. $0 \in M$ wegen der Definition $0 \cdot 0 = 0$. Ist $n \in M$, so ist auch $S(n) \in M$ wegen $0 \cdot S(n) = (0 \cdot n) + 0 = 0 + 0 = 0$. Korollar 6.1.3 impliziert wieder $M = \mathbb{N}$.

BH4: $\forall n \in \mathbb{N} : S(0) \cdot n = n \cdot S(0) = n$, also $S(0)$ ist Einselement.
Die erste Gleichung $n \cdot S(0) = n \cdot 0 + n = 0 + n = n$ folgt direkt aus den Definitionen

von $\cdot$ und $+$. Die zweite Gleichung benötigt einen Induktionsbeweis. Sei $M := \{n \in \mathbb{N} \mid S(0) \cdot n = n\}$. Es ist $0 \in M$ nach Definition von $\cdot$, und ist $n \in M$, so können wir rechnen

$$S(0) \cdot S(n) = (S(0) \cdot n) + S(0) = n + S(0) = S(n + 0) = S(n).$$

Daher ist $S(n) \in M$ und $M = \mathbb{N}$ wegen Korollar 6.1.3.

BH5: $\forall n, m \in \mathbb{N} : S(n) \cdot m = n \cdot m + m$.

Dieser erste Schritt zur Kommutativität folgt aus Korollar 6.1.3 nach Definition von $M := \{m \in \mathbb{N} \mid \forall n \in \mathbb{N} : S(n) \cdot m = n \cdot m + m\}$. Es gilt nämlich wegen $S(n) \cdot 0 = 0 = (n \cdot 0) + 0$, dass $0 \in M$ ist. Gilt nun $m \in M$, dann haben wir für beliebiges $n \in \mathbb{N}$

$$\begin{aligned} S(n) \cdot S(m) &= (S(n) \cdot m) + S(n) = (n \cdot m) + m + S(n) = \\ &= (n \cdot m) + S(m) + n = (n \cdot m) + n + S(m) = \\ &= (n \cdot S(m)) + S(m) \end{aligned}$$

und damit $S(m) \in M$.

KG($\cdot$): $\forall m, n \in \mathbb{N} : m \cdot n = n \cdot m$.

Diesmal setzen wir $M := \{n \in \mathbb{N} \mid \forall m \in \mathbb{N} : m \cdot n = n \cdot m\}$. Es ist wegen der Definition von $\cdot$ und BH3 $0 \in M$. Ist $n \in M$, so auch $S(n)$ wegen $m \cdot S(n) = (m \cdot n) + m = (n \cdot m) + m = S(n) \cdot m$. Hier haben wir die Definition und BH5 verwendet. Es ist also $M = \mathbb{N}$ wegen Korollar 6.1.3.

DG: $\forall k, m, n \in \mathbb{N} : k \cdot (m + n) = (k \cdot m) + (k \cdot n)$.

Sei $M = \{k \in \mathbb{N} \mid \forall m, n \in \mathbb{N} : k \cdot (m+n) = (k \cdot m) + (k \cdot n)\}$. Dann ist $0 \in M$ wegen $0 \cdot (m + n) = 0 = 0 + 0 = (0 \cdot m) + (0 \cdot n)$. Haben wir $k \in M$, so ist auch $S(k) \in M$ wegen Definitionen, Eigenschaften von $+$ und BH4

$$\begin{aligned} S(k) \cdot (m + n) &= (k \cdot (m + n)) + (m + n) = (k \cdot m) + (k \cdot n) + m + n = \\ &= (k \cdot m) + m + (k \cdot n) + n = (S(k) \cdot m) + (S(k) \cdot n). \end{aligned}$$

Aus Korollar 6.1.3 erhalten wir $M = \mathbb{N}$.

AG($\cdot$): $\forall k, m, n \in \mathbb{N} : (k \cdot m) \cdot n = k \cdot (m \cdot n)$.

Setzen wir diesmal $M := \{n \in \mathbb{N} \mid \forall k, m \in \mathbb{N} : (k \cdot m) \cdot n = k \cdot (m \cdot n)\}$. Es ist $0 \in M$ erfüllt, weil $(k \cdot m) \cdot 0 = 0 = k \cdot 0 = k \cdot (m \cdot 0)$. Ist nun $n \in M$ und sind $k, m \in \mathbb{N}$ beliebig, so rechnen wir nach dem zuvor bewiesenen

$$\begin{aligned} (k \cdot m) \cdot S(n) &= ((k \cdot m) \cdot n) + (k \cdot m) = (k \cdot (m \cdot n)) + (k \cdot m) = \\ &= k \cdot ((m \cdot n) + m) = k \cdot (m \cdot S(n)). \end{aligned}$$

Verwenden wir ein letztes Mal Korollar 6.1.3, so erhalten wir $M = \mathbb{N}$.

Somit haben wir alle erforderlichen Eigenschaften eines kommutativen Halbrings mit 0 und 1 nachgewiesen. $\square$

Die Vorrangregel $\cdot$ vor $+$ führen wir ein, um uns überflüssige Klammerung zu ersparen. Wir haben nun die natürlichen Zahlen mit ihren Rechenoperationen eingeführt. Wir lassen in Zukunft auch das Multiplikationszeichen weg, wenn dadurch keine Zweideutigkeit entsteht.

Theorem 6.1.8. *Die Ordnungsrelation $\leq$ und die arithmetischen Operationen $+$ und $\cdot$ sind verträglich.*

(1) $\forall k, m, n \in \mathbb{N} : (m \leq n \Rightarrow k + m \leq k + n)$,
(2) $\forall k, \ell, m, n \in \mathbb{N} : ((m \leq n \land k \leq \ell) \Rightarrow k + m \leq \ell + n)$,
(3) $\forall k, m, n \in \mathbb{N} : (n + k \leq n + m \Rightarrow k \leq m)$,
(4) $\forall k, m, n \in \mathbb{N} : (m \leq n \Rightarrow km \leq kn)$,
(5) $\forall k, m, n \in \mathbb{N} : ((n \neq 0 \land nk \leq nm) \Rightarrow k \leq m)$.

Beweis. Im gesamten Beweis definieren wir eine Menge M und beweisen $0 \in M$ und die Implikation $n \in M \implies S(n) \in M$. Dann verwenden wir Korollar 6.1.3, um $M = \mathbb{N}$ zu schließen.

Zu Beginn beweisen wir die Hilfsbehauptung $\forall m, n \in \mathbb{N} : (m \leq n \Leftrightarrow S(m) \leq S(n)))$. Es gelten

$$\begin{aligned} m \leq n \Rightarrow m \in n \vee m = n \Rightarrow S(m) \subseteq n \vee S(m) = S(n) \Rightarrow \\ \Rightarrow (S(m) \subseteq S(n) \wedge S(m) \neq S(n)) \vee S(m) = S(n) \Rightarrow \\ \Rightarrow S(m) \in S(n) \vee S(m) = S(n) \Rightarrow S(m) \leq S(n). \end{aligned}$$

und

$$\begin{aligned} S(m) \leq S(n) \Rightarrow S(m) \in S(n) \vee S(m) = S(n) \Rightarrow S(m) \in n \cup \{n\} \vee m = n \Rightarrow \\ \Rightarrow S(m) \in n \vee S(m) = n \vee m = n \Rightarrow m \in n \vee m = n \Rightarrow m \leq n, \end{aligned}$$

was die Hilfsbehauptung zeigt.

(1) $M := \{k \in \mathbb{N} \mid \forall m, n \in \mathbb{N} : (m \leq n \Rightarrow k + m \leq k + n)\}$. Trivial ist $0 \in M$. Für $k \in M$ wissen wir

$$m \leq n \Rightarrow k + m \leq k + n \Rightarrow S(k + m) \leq S(k + n) \Rightarrow S(k) + m \leq S(k) + n.$$

Daher ist $S(k) \in M$.

(2) Es gilt $k \leq \ell$ und daher ist $k+m \leq \ell+m$. Wegen $m \leq n$ gilt außerdem $\ell+m \leq \ell+n$. Aus der Transitivität von $\leq$ folgt schließlich $k + m \leq \ell + n$.

(3) Sei $M := \{n \in \mathbb{N} \mid \forall k, m \in \mathbb{N} : (n + k \leq n + m \Rightarrow k \leq m)\}$. Es gilt wieder trivialerweise $0 \in M$ und für $n \in M$ finden wir wegen

$$S(n) + k \leq S(n) + m \Rightarrow S(n + k) \leq S(n + m) \Rightarrow n + k \leq n + m \Rightarrow k \leq m$$

und $S(n) \in M$.

(4) $M := \{k \in \mathbb{N} \mid \forall m, n \in \mathbb{N} : (m \leq n \Rightarrow km \leq kn)\}$. Trivial sind $0 \in M$, da $0 \leq 0$, und $S(0) \in M$. Für $k \in M$ wissen wir

$$m \leq n \Rightarrow km \leq kn \Rightarrow km + m \leq kn + n \Rightarrow S(k)m \leq S(k)n.$$

Daher ist $S(k) \in M$.

(5) Sei $M := \{k \in \mathbb{N} \mid \forall n, m \in \mathbb{N} : ((n \neq 0 \wedge nk \leq nm) \Rightarrow k \leq m)\}$. Es gilt trivialerweise $0 \in M$, und für $k \in M$ finden wir

$$nS(k) \leq nm \Rightarrow nk + n \leq nm. \tag{6.6}$$

Nun unterscheiden wir zwei Fälle. Ist $m = 0$, so muss $nk + n = 0$ sein, da die einzige Zahl $z \in \mathbb{N}$ mit $z \leq 0$ die 0 ist. Das ist aber nur möglich, wenn $n = 0$ ist; dies ist aber nicht erlaubt. Also gilt $m \neq 0$, und damit existiert $m' \in \mathbb{N}$ mit $m = S(m')$. Wir folgern in Gleichung (6.6) weiter

$$\begin{aligned} nk + n \leq nS(m') \Rightarrow nk + n \leq nm' + n \Rightarrow nk \leq nm' \Rightarrow \\ \Rightarrow k \leq m' \Rightarrow S(k) \leq S(m') = m. \end{aligned}$$

Daher ist auch $S(k) \in M$ und $M = \mathbb{N}$.

□

Theorem 6.1.9. *Im Halbring $(\mathbb{N}, +, \cdot)$ gelten die folgenden Regeln:*

(1) *Aus $nm = 0$ folgt bereits $n = 0$ oder $m = 0$.*
(2) *Aus $n + m = n + k$ folgt $m = k$.*
(3) *Aus $nm = nk$ für $n \neq 0$ folgt $m = k$.*

Beweis.

(1) Sei $n \neq 0$ und $m \neq 0$. Dann gibt es $m', n' \in \mathbb{N}$ mit $n = S(n')$ und $m = S(m')$ und wir erhalten $mn = S(m')S(n') = m'S(n') + S(n') = m'n' + m' + S(n') = S(m'n' + m' + n') \neq 0$ wegen PA3.

(2) Sei $M := \{n \in \mathbb{N} \mid \forall m, k \in \mathbb{N} : (n + m = n + k \Rightarrow m = k)\}$. Dann ist $0 \in M$ weil aus $0 + m = 0 + k$ trivialerweise $m = k$ folgt. Sei nun $n \in M$. Dann gilt wegen Definitionen und PA4

$$S(n) + m = S(n) + k \Rightarrow S(n + m) = S(n + k) \Rightarrow n + m = n + k \Rightarrow m = k.$$

Daher ist $S(n) \in M$ und $M = \mathbb{N}$ wegen Korollar 6.1.3.

(3) Aus $nm = nk$ können wir $nm \leq nk$ folgern, und daraus wegen Theorem 6.1.8 Punkt (5) auch $m \leq k$. Da wir analog auch $nk \leq nm$ und daraus $k \leq m$ schließen können, folgt der Rest aus der Antisymmetrie der Ordnungsrelation.

Damit hätten wir alle Behauptungen bewiesen. □

6.2. Die ganzen Zahlen $\mathbb{Z}$

Die ganzen Zahlen sind die zweite Zahlenmenge, die in der Schule eingeführt wird. Um keine Probleme mit der Umkehrung der Addition, der Subtraktion $-$ zu erhalten, führt man die *negativen Zahlen* ein, die Ergebnisse, wenn man größere Zahlen von kleineren subtrahiert. Zu jeder natürlichen Zahl n gibt es eine negative Zahl $-n$ mit $n+(-n) = 0$. Auf diese Weise wird $\mathbb{Z}$ zu einer abelschen Gruppe bezüglich der Addition. Wir haben

$$\mathbb{Z} = \{\dots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}.$$

Zusammen mit der Addition $+$ und der Multiplikation $\cdot$ bildet $\mathbb{Z}$ einen Integritätsbereich. Ferner kann man die Totalordnung von $\mathbb{N}$ auf $\mathbb{Z}$ fortsetzen, indem man erklärt

$$-n \leq -m :\Leftrightarrow m \leq n \quad \text{und} \quad -m \leq 0 \qquad \forall m, n \in \mathbb{N}$$

Diese Ordnungsrelation erfüllt dann dieselben Verträglichkeitsbedingungen **(O1)** und **(O2)** wie sie schon in $\mathbb{N}$ gelten.

Die ganzen Zahlen sind gleich mächtig wie $\mathbb{N}$. Es gilt also $|\mathbb{Z}| = \aleph_0$.

6.2.1. Mengentheoretische Konstruktion von $\mathbb{Z}$. Machen wir nun den nächsten Schritt und versuchen wir eine mengentheoretische Konstruktion der ganzen Zahlen.

Gehen wir dazu von $\mathbb{N}$ aus. Bis jetzt ist dies ja die einzige unendliche Zahlenmenge, die wir aus den Axiomen konstruiert haben. Bilden wir $\mathbb{N} \times \mathbb{N}$, die Paare natürlicher Zahlen. Definieren wir eine Relation $\sim$ auf $\mathbb{N} \times \mathbb{N}$ durch

$$(m, n) \sim (m', n') :\Longleftrightarrow m + n' = m' + n$$

Proposition 6.2.1. *Die Relation $\sim$ ist eine Äquivalenzrelation auf $\mathbb{N} \times \mathbb{N}$.*

Beweis. Die Reflexivität ist offensichtlich erfüllt, ebenso wie die Symmetrie. Kommen wir zur Transitivität. Seien $(m, n) \sim (m', n')$ und $(m', n') \sim (m'', n'')$. Dann gelten $m + n' = m' + n$ und $m' + n'' = m'' + n'$. Daher wissen wir $m + n' + m'' = m' + n + m''$ und daraus wiederum folgt $m + m' + n'' = m' + n + m''$. Verwenden wir nun Eigenschaft 2 aus Theorem 6.1.9, so erhalten wir $m + n'' = m'' + n$ und $(m, n) \sim (m'', n'')$. □

Wir definieren $\mathbb{Z} := (\mathbb{N} \times \mathbb{N})/_\sim$ als Faktormenge bezüglich der oben definierten Relation. Nun wollen wir die Operationen $+$ und $\cdot$ und die Relation $\leq$ auch auf $\mathbb{Z}$ definieren.

+: Wir definieren

$$[(m_1, m_2)] + [(n_1, n_2)] := [(m_1 + n_1, m_2 + n_2)].$$

Dies ist wohldefiniert. Seien (m_1, m_2) und (m_1', m_2') zwei verschiedene Repräsentanten von $[(m_1, m_2)]$. Dann gilt $m_1 + m_2' = m_1' + m_2$ und wir erhalten

$$\begin{aligned}(m_1 + n_1) + (m_2' + n_2) &= (m_1 + m_2') + (n_1 + n_2) = \\ &= (m_1' + m_2) + (n_1 + n_2) = \\ &= (m_1' + n_1) + (m_2 + n_2).\end{aligned}$$

Daher ist $(m_1 + n_1, m_2 + n_2) \sim (m_1' + n_1, m_2' + n_2)$. Analog weist man die Wohldefiniertheit im zweiten Term nach.

$\cdot$: Für die Multiplikation setzen wir

$$[(m_1, m_2)] \cdot [(n_1, n_2)] := [(m_1 n_1 + m_2 n_2, m_1 n_2 + m_2 n_1)].$$

Auch das ist wohldefiniert, wie man leicht nachrechnet.

$\leq$: Die Ordnungsrelation führt man auch zurück auf die Relation in $\mathbb{N}$:

$$[(m_1, m_2)] \leq [(n_1, n_2)] :\iff m_1 + n_2 \leq n_1 + m_2.$$

Diese Relation ist wohldefiniert, was man leicht nachrechnet. Sie ist auch offensichtlich reflexiv. Sie ist antisymmetrisch, weil aus $[(m_1, m_2)] \leq [(n_1, n_2)]$ und $[(n_1, n_2)] \leq [(m_1, m_2)]$ und den Eigenschaften von $\leq$ auf $\mathbb{N}$ die Beziehung $m_1 + n_2 = n_1 + m_2$, also $(m_1, m_2) \sim (n_1, n_2)$ und daher $[(m_1, m_2)] = [(n_1, n_2)]$ folgt.

Die Transitivität erhält man so: $[(m_1, m_2)] \leq [(n_1, n_2)]$ impliziert $m_1 + n_2 \leq n_1 + m_2$, und aus $[(n_1, n_2)] \leq [(k_1, k_2)]$ folgt $n_1 + k_2 \leq k_1 + n_2$. Aus Theorem 6.1.8 erhalten wir

$$m_1 + n_2 + k_2 \leq n_1 + m_2 + k_2 \leq k_1 + n_2 + m_2,$$

woraus schließlich $m_1 + k_2 \leq k_1 + m_2$ folgt, also $[(m_1, m_2)] \leq [(k_1, k_2)]$.

Jetzt haben wir die Grundoperationen definiert. Es bleibt noch, ihre Eigenschaften zu beweisen.

Theorem 6.2.2. *Die ganzen Zahlen* $(\mathbb{Z}, +, \cdot)$ *sind ein Integritätsbereich.*

Beweis. Verifizieren wir zuerst, dass $(\mathbb{Z}, +)$ eine abelsche Gruppe ist:

G1: Es gilt $([(m_1, m_2)] + [(n_1, n_2)]) + [(k_1, k_2)] = [(m_1, m_2)] + ([(n_1, n_2)] + [(k_1, k_2)])$, weil die Operation komponentenweise definiert ist und $+$ auf $\mathbb{N}$ assoziativ ist.

G2: Das Element $[(0, 0)]$ ist neutrales Element, wie man sofort einsieht.

G3: Sei $[(m_1, m_2)] \in \mathbb{Z}$ beliebig. Dann ist das Element $[(m_2, m_1)]$ ein Inverses bezüglich der Addition.
Es gilt $[(m_1, m_2)] + [(m_2, m_1)] = [(m_1 + m_2, m_1 + m_2)] = [(0, 0)]$.

G4: Das Kommutativgesetz ist erfüllt, weil es in $(\mathbb{N}, +)$ gilt und die Operation in $\mathbb{Z}$ komponentenweise auf Repräsentanten definiert ist.

Nun müssen wir zeigen, dass $(\mathbb{Z}, \cdot)$ ein kommutatives Monoid ist:

M1: Es gilt $([(m_1, m_2)][(n_1, n_2)])[(k_1, k_2)] = [(m_1, m_2)]([(n_1, n_2)][(k_1, k_2)])$. Das sieht man nach langer aber einfacher Rechnung ein.

M2: Das Element $[(1, 0)]$ ist Einselement. Das ist leicht.

M3: Es gilt das Kommutativgesetz $[(m_1, m_2)][(n_1, n_2)] = [(n_1, n_2)][(m_1, m_2)]$. Das folgt unmittelbar aus der Definition.

D: Ebenso mühsam aber einfach nachzurechnen wie das Assoziativgesetz ist das Distributivgesetz.

Was bleibt, ist die Freiheit von Nullteilern zu zeigen. Seien $[(m_1, m_2)]$ und $[(n_1, n_2)]$ zwei Elemente von $\mathbb{Z}$ mit $[(m_1, m_2)][(n_1, n_2)] = [(0, 0)]$. Aus dieser Beziehung folgt mit Hilfe der Definitionen von $\cdot$ und $\sim$ die Beziehung

$$m_1 n_1 + m_2 n_2 = m_1 n_2 + m_2 n_1. \tag{6.7}$$

Hilfsbehauptung: Wir zeigen nun, dass für je vier Zahlen $m, n, k, \ell \in \mathbb{N}$ aus

$$mk + n\ell = m\ell + nk \quad \wedge \quad m \neq n$$

schon $k = \ell$ folgt. Wie immer beweisen wir das mit vollständiger Induktion. Sei

$$M := \big\{n \in \mathbb{N} \mid \forall k, \ell, m \in \mathbb{N} : \big((mk + n\ell = m\ell + nk \wedge m \neq n) \Rightarrow k = \ell\big)\big\}.$$

Dann gilt $0 \in M$, weil

$$mk + 0\ell = m\ell + 0k \Rightarrow mk = m\ell \Rightarrow k = \ell \quad \text{wegen } m \neq n = 0 \text{ und Theorem 6.1.9.}$$

Sei nun $n \in M$. Dann untersuchen wir

$$mk + S(n)\ell = m\ell + S(n)k$$

Für $m = 0$ haben wir $0k+S(n)\ell = 0\ell+S(n)k$, woraus sofort $\ell = k$ folgt wegen Theorem 6.1.9 (3). Sei also nun $m \neq 0$ und $m \neq S(n)$. Dann existiert $m' \in \mathbb{N}$ mit $S(m') = m$, und wir können unter Verwendung von Theorem 6.1.9 rechnen

$$\begin{aligned}
mk + S(n)\ell &= m\ell + S(n)k \\
mk + n\ell + \ell &= m\ell + nk + k \\
S(m')k + n\ell + \ell &= S(m')\ell + nk + k \\
m'k + k + n\ell + \ell &= m'\ell + \ell + nk + k \\
m'k + n\ell &= m'\ell + nk.
\end{aligned}$$

Falls $n \neq m'$ gilt, dann können wir aus $n \in M$ schon $\ell = k$ folgern. Das ist aber der Fall, weil $S(m') = m \neq S(n)$ vorausgesetzt war. Daher ist auch $S(n) \in M$ und aus Korollar 6.1.3 folgt $M = \mathbb{N}$ und die Hilfsbehauptung.

Kehren wir zurück zu unserer Beziehung (6.7). Aus der Hilfsbehauptung erhalten wir für $m_1 \neq m_2$ die Folgerung $n_1 = n_2$, also $[(n_1, n_2)] = [(0, 0)]$. Gilt andererseits $m_1 = m_2$, so bedeutet das $[(m_1, m_2)] = [(0, 0)]$ und wir schließen die Nichtexistenz von Nullteilern. $\square$

Wir können sehr leicht nachrechnen, dass für die Elemente $[(n, 0)]$ dieselben Rechenregeln gelten wie für natürliche Zahlen n. Außerdem sind alle diese Zahlen verschieden ($n \neq m \Rightarrow [(n, 0)] \neq [(m, 0)]$). Es ist also $\mathbb{N} \subseteq \mathbb{Z}$ mit dieser Identifikation. Wir schreiben in Zukunft auch n für diese Elemente. Es ist nun das Inverse bzgl. $+$ von n die Klasse $[(0, n)]$, und wir schreiben für dieses Element von $\mathbb{Z}$ kurz $-n$. Die Elemente $[(n, 0)]$ und $[(0, n)]$ für $n \in \mathbb{N}$ sind auch schon alle Elemente in $\mathbb{Z}$, da

$$([m_1, m_2]) = m_1 + (-m_2) = \begin{cases} ([m_1 - m_2, 0]) & \text{falls } m_1 \geq m_2 \\ ([0, m_2 - m_1]) & \text{falls } m_1 < m_2. \end{cases}.$$

Damit haben wir endlich die uns vertraute Form der ganzen Zahlen als „$\pm\mathbb{N}$“ erreicht.

Es gilt für alle $n, m \in \mathbb{N}$, dass $[(n, 0)] \leq [(m, 0)]$ genau dann, wenn $n \leq m$. Das folgt direkt aus der Definition. Ebenfalls aus der Definition folgt sogleich $[(0, n)] \leq [(0, m)]$, dann und nur dann wenn $m \leq n$ ist. Schließlich kann man noch aus der Definition ablesen, dass für $\mathbb{N} \ni n \neq 0$ die Ungleichungen $[(0, n)] < [(0, 0)] < [(n, 0)]$ gelten. Die natürlichen Zahlen entsprechen also genau den *positiven* Elementen von $\mathbb{Z}$, und die Elemente $-n$ sind die *negativen* Elemente (die negativen Zahlen).

Theorem 6.2.3. *Für die Ordnungsrelation von $\mathbb{Z}$ finden wir die folgenden Eigenschaften.*

(1) $\forall m, n \in \mathbb{Z} : (m \leq n \implies -m \geq -n)$,
(2) $\forall k, m, n \in \mathbb{Z} : (m \leq n \Rightarrow m + k \leq n + k)$,
(3) $\forall m, n \in \mathbb{Z} : ((m > 0 \wedge n > 0) \Rightarrow mn > 0)$,
(4) $\forall k, m, n \in \mathbb{Z} : ((k > 0 \wedge m \leq n) \Rightarrow km \leq kn)$,
(5) $\forall k, m, n \in \mathbb{Z} : ((k < 0 \wedge m \leq n) \Rightarrow km \geq kn)$,
(6) $\forall k, m, n \in \mathbb{Z} : ((k > 0 \wedge km \leq kn) \Rightarrow m \leq n)$

Beweis.

(1) Sind die Vorzeichen von m und n verschieden, so wissen wir $m \leq 0 \leq n$ und daher $-m \geq 0 \geq -n$. Sind m und n positiv, so sind $-m = [(0, m)]$ und $-n = [(0, n)]$. Wegen $m \leq n$ gilt nach Definition von $\leq$ auf $\mathbb{Z}$ die Beziehung $-m \geq -n$. Haben wir umgekehrt $m \leq n \leq 0$, so impliziert das analog zu oben $-m \geq -n$.
(2) Sind $m = [(m_1, m_2)]$, $n = [(n_1, n_2)]$ und $k = [(k_1, k_2)]$, so erhalten wir wegen Theorem 6.1.8

$$\begin{aligned} m &\leq n \\ [(m_1, m_2)] &\leq [(n_1, n_2)] \\ m_1 + n_2 &\leq m_2 + n_1 \\ m_1 + k_1 + n_2 + k_2 &\leq m_2 + k_2 + n_1 + k_1 \\ [(m_1 + k_1, m_2 + k_2)] &\leq [(n_1 + k_1, n_2 + k_2)] \\ m + k &\leq n + k \end{aligned}$$

(3) Dies folgt aus Theorem 6.1.8.(4) und der Nullteilerfreiheit.
(4) Ist $m \geq 0$, so folgt aus Theorem 6.1.8.(4) sofort $km \geq 0 = k0$. Gilt nun $m \leq n$, so folgt aus (2) $0 \leq n - m$ und aus dem schon bewiesenen $0 \leq k(n - m) = kn - km$ und wir erhalten wieder aus (2) die gesuchte Ungleichung $km \leq kn$.
(5) Für $k \leq 0$ ist $-k \geq 0$ und alles weitere folgt aus (4).
(6) Gilt $km \leq kn$, so erhalten wir aus (2) die Beziehung $0 \leq k(n - m)$. Weil $k > 0$ gilt, können wir aus Theorem 6.1.8.(5) $0 \leq n - m$ und damit wegen (2) $m \leq n$ schließen.

□

Proposition 6.2.4. *Ist $k \neq 0$, so folgt aus $km = kn$ schon $m = n$ für beliebige $m, n \in \mathbb{Z}$.*

Beweis. Es gilt $km = kn \implies 0 = km - kn \implies 0 = k(m - n)$. Weil $k \neq 0$ gilt, muss wegen der Nullteilerfreiheit $m - n = 0$, also $m = n$ gelten. □

6.3. Die rationalen Zahlen $\mathbb{Q}$

Die rationalen Zahlen sind die nächst umfassendere von früher bekannte Zahlenmenge. Ebenso wie man die ganzen Zahlen konstruiert, um die Subtraktion für alle Zahlen durchführen zu können, muss man für die Umkehrung der Multiplikation wieder die Zahlenmenge erweitern.

Man geht von den ganzen Zahlen zu den Bruchzahlen über. Man führt also Ausdrücke der Form

$$q = \frac{m}{n}$$

ein. Hier entdeckt man die ersten beiden Schwierigkeiten, die bei der naiven Einführung der ganzen Zahlen nicht aufgetreten sind. Erstens schafft man es nicht, dem Ausdruck $\frac{m}{0}$ Sinn zu geben, ohne Widersprüche zu verursachen. Zweitens bemerkt man, dass es notwendig ist, Ausdrücke der Form $\frac{m}{n}$ und $\frac{km}{kn}$ für gleich zu erklären ($\frac{1}{2} = \frac{2}{4}$). Mathematisch heißt das, man muss bei der Einführung von $\mathbb{Q}$ *Äquivalenzklassen* bilden und die Null im Nenner verbieten!

Man definiert also $\mathbb{Q}$ als die Äquivalenzklassen von Brüchen der Form $\frac{m}{n}$ ganzer Zahlen mit $n \neq 0$. Man findet, dass es in jeder Äquivalenzklasse einen Bruch gibt, sodass m und n teilerfremd sind und weiters $n > 0$ gilt.

Zusammen mit der Addition $+$ und $\cdot$ bildet $\mathbb{Q}$ einen Körper. Außerdem ist auf $\mathbb{Q}$ eine Ordnungsrelation $\leq$ definiert, für die $\mathbb{Q}$ ein *geordneter* Körper ist, genau bedeutet dies.

Definition 6.3.1 (Geordneter Körper)**.** *Ein Körper $(K, +, \cdot)$, der auch eine totalgeordnete Menge $(K, \leq)$ ist, heißt* geordneter Körper*, falls die beiden Ordnungsaxiome gelten, d.h. für $q, r, s \in K$ gilt*

(O1) $q \leq r \;\Rightarrow\; q + s \leq r + s$,
(O2) $q > 0 \wedge r > 0 \;\Rightarrow\; qr > 0$.

Wir schreiben dann $(K, +, \cdot, \leq)$.

Die Ordnungsrelation muss also mit den Rechenoperationen **verträglich** sein. Aus den Ordnungsaxiomen können wir auch bereits die bekannten Rechengesetze für Ungleichungen herleiten, wie „das Ungleichheitszeichen dreht sich um, wenn man mit einer negativen Zahl multipliziert".

Proposition 6.3.2 (Rechenregeln in geordneten Körpern)**.** *In einem geordneten Körper $(K, +, \cdot, \leq)$ gelten folgende Aussagen ($x, y, z \in K$)*

(1) $x \leq y \;\Leftrightarrow y - x \geq 0$.
(2) $x \leq 0 \Leftrightarrow -x \geq 0$.
(3) *Ist $x \geq 0$ und $y \leq z$, dann folgt $xy \leq xz$.*
(4) *Ist $x < 0$ und $y \leq z$, dann folgt $xy \geq xz$.*
(5) *Für $x \neq 0$ ist $x^2 > 0$ und daher $1 > 0$.*
(6) *Ist $0 < x < y$, dann folgt $0 < y^{-1} < x^{-1}$.*

Beweis.

(1) Aus $x \leq y$ folgt mit (O1) $0 = x + (-x) \leq y + (-x)$ und somit $0 \leq y - x$. Umgekehrt ergibt sich aus $y - x \geq 0$ mit (O1) $y = y - x + x \geq 0 + x = x$.
(2) Folgt aus (1) für y=0.
(3) Für $y = z$ wissen wir $xy = xz$. Für $x = 0$ gilt $0 = xy = xz = 0$. Ist $y < z$, so ist wegen (1) $0 < z - y$. Ist schließlich $x > 0$, dann folgt aus (O2) $0 < x(z - y) = xz - xy$ und somit ist $xy < xz$.
(4) Dies folgt aus (2) und (3).
(5) Ist $x > 0$, so gilt $x^2 = x \cdot x > 0$ wegen (O2). Für $x < 0$ ist $-x > 0$ und $x^2 = (-x)(-x) > 0$. Es ist $1 \neq 0$ und daher $1 = 1^2 > 0$.
(6) Ist $x > 0$, so ist $x^{-1} > 0$. Wäre das nicht so, hätten wir $1 = xx^{-1} < 0$ im Widerspruch zu (5). Gilt $0 < x < y$, so wissen wir $x^{-1}y^{-1} > 0$, und daher folgt

$$\begin{aligned} x &< y \\ x(x^{-1}y^{-1}) &< y(x^{-1}y^{-1}) \\ y^{-1} &< x^{-1}. \end{aligned}$$

□

Proposition 6.3.3. *Die Menge $\mathbb{N}$ ist in $\mathbb{Q}$ nach oben unbeschränkt.*

Beweis. Angenommen, $\mathbb{N}$ sei in $\mathbb{Q}$ beschränkt. Dann existieren positive natürliche Zahlen k und m mit der Eigenschaft, dass $\forall n \in \mathbb{N} : n \leq \frac{m}{k}$. Das ist gleichbedeutend mit der Aussage, dass $\forall n \in \mathbb{N} : nk \leq m$ wegen Proposition 6.3.2(3). Nachdem k positiv ist, muss $nk \geq n$ sein, weil $k \geq 1$ gilt ($k = k' + 1$, daher $nk = nk' + n$ mit $n \geq 0$ und $k' \geq 0$, also $nk' \geq 0$, was $nk \geq n$ impliziert) und daher existiert eine positive natürliche Zahl m so, dass $\forall n \in \mathbb{N} : n \leq m$. Es ist aber $m + 1 > m$, ein Widerspruch. Daher ist $\mathbb{N}$ in $\mathbb{Q}$ unbeschränkt. □

Die Menge Q ist abzählbar; es gilt also $|\mathbb{Q}| = \aleph_0$ (vgl. Abschnitt 4.4). Außerdem besitzt $\mathbb{Q}$ keinen nicht-trivialen Unterkörper.

6.3.1. Mengentheoretische Konstruktion von $\mathbb{Q}$. Wenn wir die ganzen Zahlen konstruiert haben, steht uns nichts im Wege, dieselbe Konstruktion so ähnlich noch einmal durchzuführen. Im folgenden bezeichne $\mathbb{Z}_+ := \{n \in \mathbb{Z} \mid n > 0\}$ die Menge der positiven Elemente in $\mathbb{Z}$, also der natürlichen Zahlen ungleich 0.

Betrachten wir auf der Menge $\mathbb{Z} \times \mathbb{Z}_+$ die Relation

$$(m_1, m_2) \sim (n_1, n_2) :\iff m_1 n_2 = m_2 n_1.$$

Insbesondere gilt für jede positive natürliche Zahl n die Relation $(m_1, m_2) \sim (nm_1, nm_2)$.

Proposition 6.3.4. *Es gilt wieder $\sim$ ist eine Äquivalenzrelation auf $\mathbb{Z} \times \mathbb{Z}_+$.*

Beweis.

Reflexivität: ist offensichtlich,
Symmetrie: erfüllt, weil Definition symmetrisch ist,
Transitivität: Seien $(m_1, m_2) \sim (n_1, n_2)$ und $(n_1, n_2) = (k_1, k_2)$. Dann sind $m_1 n_2 = m_2 n_1$ und $n_1 k_2 = n_2 k_1$. Multiplizieren wir die erste Gleichung mit k_2, so erhalten wir $m_1 n_2 k_2 = m_2 n_1 k_2$. Jetzt können wir die zweite Gleichung einsetzen und erhalten $m_1 n_2 k_2 = m_2 n_2 k_1$. Nachdem $n_2 \neq 0$ gilt und $\mathbb{Z}$ ein Integritätsbereich ist, folgt $m_1 k_2 = m_2 k_1$, also $(m_1, m_2) \sim (k_1, k_2)$.

□

Die Menge der rationalen Zahlen $\mathbb{Q}$ ist definiert als Faktormenge $\mathbb{Q} := \mathbb{Z} \times \mathbb{Z}_+/_\sim$.

Wenn wir die Operationen

$$\begin{aligned}
[(m_1, m_2)] + [(n_1, n_2)] &:= [(m_1 n_2 + m_2 n_1, m_2 n_2)] \\
[(m_1, m_2)] \cdot [(n_1, n_2)] &:= [(m_1 n_1, m_2 n_2)]
\end{aligned}$$

definieren, so sind diese wohldefiniert und es gilt der folgende Satz

Theorem 6.3.5. *Die Menge der rationalen Zahlen $(\mathbb{Q}, +, \cdot)$ ist ein Körper mit Nullelement $[(0,1)]$ und Einselement $[(1,1)]$. Die Menge aller Elemente der Form $[(n,1)]$ für $n \in \mathbb{Z}$ entspricht $\mathbb{Z}$ mit allen seinen Eigenschaften (aka ist **isomorph zu** $\mathbb{Z}$).*

Beweis. Beginnen wir mit der Wohldefiniertheit von $+$. Sei $(m_1', m_2') \in [(m_1, m_2)]$. Dann haben wir $m_1' m_2 = m_1 m_2'$ und

$$\begin{aligned}
[(m_1', m_2')] + [(n_1, n_2)] &= [(m_1' n_2 + m_2' n_1, m_2' n_2)] = [((m_1' n_2 + m_2' n_1) m_2, m_2' n_2 m_2)] = \\
&= [(m_1' n_2 m_2 + m_2' n_1 m_2, m_2' n_2 m_2) = [(m_2' m_1 n_2 + m_2' n_1 m_2, m_2' n_2 m_2)] = \\
&= [(m_2'(m_1 n_2 + n_1 m_2), m_2' n_2 m_2)] = [(m_1 n_2 + n_1 m_2, n_2 m_2)] = \\
&= [(m_1, m_2)] + [(n_1, n_2)].
\end{aligned}$$

Die Wohldefiniertheit im zweiten Term zeigt man analog.

Nun rechnen wir die Gruppenaxiome für $+$ nach

(K1) Seien $q = [(q_1, q_2)]$, $[(r_1, r_2)]$ und $[(s_1, s_2)]$. Wir rechnen

$$\begin{aligned}
(q + r) + s &= [(q_1 r_2 + q_2 r_1, q_2 r_2)] + [(s_1, s_2)] = [((q_1 r_2 + q_2 r_1) s_2 + s_1 q_2 r_2, q_2 r_2 s_2)] = \\
&= [(q_1 r_2 s_2 + q_2 r_1 s_2 + s_1 q_2 r_2, q_2 r_2 s_2)] = [(q_1 r_2 s_2 + q_2(r_1 s_2 + r_2 s_1), q_2 r_2 s_2)] = \\
&= [(q_1, q_2)] + [(r_1 s_2 + r_2 s_1, r_2 s_2)] = q + (r + s)
\end{aligned}$$

(K2) Die Definition von $q + r$ ist symmetrisch in q und r.

(K3) Es gilt $[(q_1, q_2)] + [(0,1)] = [(1q_1 + 0q_2, 1q_2)] = [(q_1, q_2)]$. Daher ist $0 = [(0,1)]$ das neutrale Element.

(K4) Wir rechnen $[(q_1, q_2)] + [(-q_1, q_2)] = [(q_1q_2 - q_1q_2, q_2^2)] = [(0, q_2^2)] = [(0, 1)] = 0$. Das inverse Element von $[(q_1, q_2)]$ ist also $[(-q_1, q_2)]$.

Die Wohldefiniertheit der Multiplikation erkennen wir aus der folgenden Rechnung. Sei $(m_1', m_2') \in [(m_1, m_2)]$ und deshalb $m_1'm_2 = m_1m_2'$. Dann finden wir

$$\begin{aligned}[(m_1', m_2')][(n_1, n_2)] &= [(m_1'n_1, m_2'n_2)] = [(m_1'n_1m_2, m_2'n_2m_2)] = \\ &= [(m_2'm_1n_1, m_2'n_2m_2)] = [(m_1n_1, n_2m_2)] = [(m_1, m_2)][(n_1, n_2)].\end{aligned}$$

Die Wohldefiniertheit im zweiten Faktor zeigt man analog.

Die Gruppenaxiome für $\cdot$ kommen nun.

(K5), (K6) Die Multiplikation ist komponentenweise definiert, und die Multiplikation ganzer Zahlen ist kommutativ und assoziativ.

(K7) Das Element $1 := [(1, 1)] \neq [(0, 1)]$ ist offensichtlich Einselement.

(K8) Ist $q = [(q_1, q_2)] \neq 0$, dann ist $q_1 \neq 0$ und wir finden $q^{-1} = [(q_2, q_1)]$, falls $q_1 > 0$ und $q^{-1} = [(-q_2, -q_1)]$ für $q_1 < 0$. Dass dann q^{-1} das Inverse von q ist, ist einfach einzusehen.

Das Distributivgesetz sieht man so ein.

(K9) Für $q = [(q_1, q_2)]$, $r = [(r_1, r_2)]$ und $s = [(s_1, s_2)]$ rechnen wir

$$\begin{aligned}q(r + s) &= [(q_1, q_2)]\big([(r_1, r_2)] + [s_1, s_2]\big) = [(q_1, q_2)][(r_1s_2 + r_2s_1, r_2s_2)] = \\ &= [(q_1(r_1s_2 + r_2s_1), q_2r_2s_2)] = [(q_1r_1s_2 + q_1r_2s_1, q_2r_2s_2)] = \\ &= [(q_1r_1q_2s_2 + q_2r_2q_1s_1, q_2^2r_2s_2)] = [(q_1r_1, q_2r_2)] + [(q_1s_1, q_2s_2)] = \\ &= [(q_1, q_2)][r_1, r_2] + [(q_1, q_2)][(s_1, s_2)] = qr + qs\end{aligned}$$

Daher ist $\mathbb{Q}$ ein Körper. $\square$

Führen wir darüber hinaus die Relation $\leq$ ein, indem wir fordern

$$[(m_1, m_2)] \leq [(n_1, n_2)] :\iff m_1n_2 \leq n_1m_2,$$

so ist dies wohldefiniert. Hätten wir etwa $(m_1', m_2') \in [(m_1, m_2)]$ gewählt, so ist $m_1m_2' = m_1'm_2$ und wir haben

$$\begin{aligned}m_1n_2 &\leq n_1m_2 \\ m_1m_2'n_2 &\leq n_1m_2m_2' \\ m_1'm_2n_2 &\leq n_1m_2m_2' \\ m_1'n_2 &\leq n_1m_2' \quad \text{wegen } m_2 > 0 \text{ und Theorem 6.2.3.}\end{aligned}$$

Analog zeigen wir die Wohldefiniertheit auf der rechten Seite.

Theorem 6.3.6. *Die Relation $\leq$ macht $\mathbb{Q}$ zu einem geordneten Körper.*

Beweis. Wir müssen die Bedingungen O1 und O2 nachweisen:

(O1) Seien $q = [(q_1, q_2)]$, $r = [(r_1, r_2)]$ und $s = [(s_1, s_2)]$. Dann gilt

$$\begin{aligned}q \leq r &\implies q_1r_2 \leq q_2r_1 \implies q_1s_2r_2 \leq r_1s_2q_2 \implies \\ &\implies (q_1s_2 + s_1q_2)r_2 \leq (r_1s_2 + s_1r_2)q_2 \implies \\ &\implies (q_1s_2 + s_1q_2)r_2s_2 \leq (r_1s_2 + s_1r_2)q_2s_2 \implies \\ &\implies [(q_1s_2 + s_1q_2, q_2s_2)] \leq [(r_1s_2 + s_1r_2, r_2s_2)] \implies q + s \leq r + s.\end{aligned}$$

(O2) Sei $q = [(q_1, q_2)] > 0$, dann folgt $q_1 > 0$. Für $r = [(r_1, r_2)]$ gilt analog $r_1 > 0$. Daher ist $qr = [(q_1r_1, q_2r_2)] > 0$, weil $q_1r_1 > 0$ gilt wegen Theorem 6.2.3.

$\square$

Wenn wir zu guter Letzt die Schreibweise

$$\frac{m}{n} := \begin{cases} [(m,n)] & \text{für } n > 0 \\ [(-m,-n)] & \text{für } n < 0 \end{cases}$$

erklären, dann haben wir die „Bruchzahlen“ wieder eingeführt und die gewohnte Notation von $\mathbb{Q}$ zurückgewonnen.

Auch die ganzen Zahlen $\mathbb{Z}$ können wir in $\mathbb{Q}$ wiederfinden. Wenn wir die Elemente der Form $[(n,1)]$ betrachten, so sehen wir, dass für $m \neq n$ auch $[(m,1)] \neq [(n,1)]$ gilt. Die Rechenoperationen in $\mathbb{Z}$ gelten auch: $[(m,1)] + [(n,1)] = [(m+n,1)]$ und $[(m,1)][(n,1)] = [(mn,1)]$. Die Abbildung $\iota : \mathbb{Z} \to \mathbb{Q}$ mit $\iota : z \mapsto [(z,1)]$ ist ein injektiver Ringhomomorphismus. Wir können also $\mathbb{Z} \cong \iota(\mathbb{Z}) \subseteq \mathbb{Q}$ als Teilring (sogar Teil-Integritätsbereich) sehen. Wir werden Elemente der Form $[(m,1)]$ daher weiterhin mit der ganzen Zahl m identifizieren.

6.4. Die reellen Zahlen ℝ

Die reellen Zahlen sind die vorletzte Zahlenmenge, die wir genauer untersuchen wollen. Weil einige wichtige Beziehungen in $\mathbb{Q}$ nicht berechnet werden können (etwa die Länge der Diagonale des Einheitsquadrates oder die Fläche des Einheitskreises), bleibt uns keine Wahl als die Zahlenmenge ein weiteres Mal zu vergrößern.

Der Körper $\mathbb{Q}$ ist auch „löchrig“ im folgenden Sinn. Betrachten wir die beiden disjunkten Mengen

$$A = \{x \in \mathbb{Q} \mid x > 0 \wedge x^2 < 2\}$$
$$B = \{x \in \mathbb{Q} \mid x > 0 \wedge x^2 > 2\},$$

dann ist deren Vereinigung $A \cup B = \mathbb{Q}_+$. Wir würden aber vom Gefühl erwarten, dass zwischen den beiden Mengen noch eine Zahl sein sollte. Das ist natürlich nicht möglich, da diese Zahl die Gleichung $x^2 = 2$ erfüllen würde, was bekanntermaßen in den rationalen Zahlen nicht möglich ist (Theorem 3.2.4).

Um die Löcher zu „stopfen“, müssen wir zu $\mathbb{Q}$ irrationale Zahlen hinzufügen und erhalten den geordneten Körper $(\mathbb{R}, +, \cdot, \leq)$, den wir auch als **Zahlengerade** repräsentieren. Die rationalen Zahlen sind ein geordneter Unterkörper von $\mathbb{R}$.

Die reellen Zahlen bilden die Grundlage der Analysis, und daher müssen wir einige wichtige Eigenschaften von $\mathbb{R}$ ableiten.

Definition 6.4.1 (Ordnungsvollständigkeit)**.** *Eine geordnete Menge M ist* ordnungsvollständig *(hat die* Supremums–Eigenschaft*), wenn zu jeder nichtleeren nach oben beschränkten Teilmenge $E \subseteq M$ ein Supremum* $\sup E \in M$ *existiert (vgl. Definition 4.2.18).*

Um diese Eigenschaft vernünftig anwenden zu können, müssen wir zuerst einige äquivalente Formulierungen beweisen.

Proposition 6.4.2 (Charakterisierung ordnungvollständiger Mengen)**.** *Sei M eine geordnete Menge. Dann sind äquivalent:*

(1) *M ist ordnungsvollständig.*
(2) *Jede nach unten beschränkte nichtleere Teilmenge $F \subseteq M$ besitzt ein Infimum* $\inf F \in M$.
(3) *Für je zwei nichtleere Teilmengen E und F von M mit*

$$a \leq b \qquad \forall a \in E, \forall b \in F$$

gibt es ein Element $m \in M$ mit

$$a \leq m \leq b \qquad \forall a \in E, \forall b \in F$$

Beweis. Wir beginnen mit (1)⇒(2). Sei $\emptyset \neq F \subseteq M$ und F nach unten beschränkt. Wir definieren

$$E := \{x \in M \mid x \leq f \ \forall f \in F\}.$$

Die Menge E ist nach oben beschränkt, weil jedes Element in F eine obere Schranke für E ist. Außerdem ist E nichtleer, da F als nach unten beschränkt vorausgesetzt war. Nach Voraussetzung existiert daher das Supremum $\alpha = \sup E \in M$. Wir zeigen nun, dass $\alpha = \inf F$ gilt. Nachdem E die Menge aller unteren Schranken von F ist, ist α größer oder gleich allen unteren Schranken von F. Wir müssen also nur zeigen, dass α eine untere Schranke von F ist. Angenommen, das ist nicht der Fall. Dann gäbe es ein $f \in F$ mit $f < \alpha$. Weil E die Menge der unteren Schranken von F ist, gilt $e \leq f \ \forall f \in F$. Daher ist f eine obere Schranke von E, ein Widerspruch zur Supremumseigenschaft von α. Daher ist α tatsächlich eine untere Schranke von F, also $\inf F$.
(2)⇒(3): Seien E und F Mengen wie in der Voraussetzung. Wegen (2) existiert $m := \inf F$. Klarerweise ist $m \leq b$ für alle $b \in F$. Es ist außerdem $\forall a \in E : a \leq m$, denn wäre das nicht der Fall, so gäbe es ein $e \in E$ mit $e > m$. Wegen der Eigenschaften von E und F ist aber e eine untere Schranke von F, was der Infimumseigenschaft von m widerspricht. Daher gilt (3).
(3)⇒(1): Sei E eine nach oben beschränkte Menge. Wir definieren die Menge F aller oberen Schranken von E als

$$F := \{x \in M \mid e \leq x \ \forall e \in E\} \neq \emptyset.$$

Nach Voraussetzung existiert dann ein $m \in M$ mit $e \leq m \leq f$ für alle $e \in E$ und $f \in F$. Daher ist m eine obere Schranke von E. Sei $\alpha < m$. Dann ist $\alpha \notin F$, also keine obere Schranke. Daher ist m das Supremum von E. □

Beispiel 6.4.3. *Die Menge der rationalen Zahlen ist* nicht *ordnungsvollständig. Betrachten wir nämlich die Teilmengen A und B von $\mathbb{Q}$, die definiert sind durch*

$$A = \{x \in \mathbb{Q} \mid x > 0 \wedge x^2 < 2\},$$
$$B = \{x \in \mathbb{Q} \mid x > 0 \wedge x^2 > 2\}.$$

Zunächst sind A und B nichtleer, da $1 \in A$ und $2 \in B$. Weiters gilt $a < b$ für alle $a \in A$ und für alle $b \in B$. Wäre $\mathbb{Q}$ ordnungsvollständig, dann gäbe es ein Element $m \in \mathbb{Q}$ mit

$$a \leq m \leq b \text{ für alle } a \in A \text{ und } b \in B. \tag{6.8}$$

Definieren wir nun

$$c := m - \tfrac{m^2-2}{m+2} = \tfrac{2m+2}{m+2} > 0. \tag{6.9}$$

Damit gilt

$$c^2 - 2 = \frac{2(m^2-2)}{(m+2)^2}. \tag{6.10}$$

Ist nun $m^2 > 2$, dann folgt mit (6.10) dass $c^2 > 2$ und somit $c \in B$. Allerdings ist wegen (6.9) $c < m$, was im Widerspruch zu (6.8) steht.

Andererseits führt aber auch $m^2 < 2$ zu einem Widerspruch. In diesem Fall impliziert nämlich (6.10) $c^2 < 2$ und somit $c \in A$, während (6.9) $c > m$ zur Folge hat. Das widerspricht aber (6.8).

Daher gilt also $m^2 = 2$, was aber in $\mathbb{Q}$ unmöglich ist wegen Theorem 3.2.4.

Die nun dringend benötigte Ordnungsvervollständigung von $\mathbb{Q}$ und damit den Schritt zu den reellen Zahlen liefert uns der folgende Satz.

Theorem 6.4.4 (Richard Dedekind (1831–1916)). *Es existiert bis auf Isomorphie genau ein ordnungsvollständiger geordneter Körper $\mathbb{R}$, der $\mathbb{Q}$ als geordneten Unterkörper besitzt. Wir nennen $\mathbb{R}$ die Menge der reellen Zahlen und die Elemente der Menge $\mathbb{R} \setminus \mathbb{Q}$ die irrationalen Zahlen.*

Beweis. In Abschnitt 6.4.1. □

Will man sich nicht auf die mengentheoretischen Konstruktionen in Abschnitt 6.4.1 einlassen, so kann man sich auf den Standpunkt von Hilbert stellen, der die Einführung der reellen Zahlen als ordnungsvollständigen, geordneten Körper den mengentheoretischen Konstruktionen vorzog, und pragmatisch das Theorem 6.4.4 zur Definition erheben. Das bedeutet also, die reellen Zahlen als ordnungsvollständigen, geordneten Körper zu *definieren*, der $\mathbb{Q}$ als geordneten Unterkörper besitzt; dann garantiert Theorem 6.4.4, dass es einen und nur einen solchen Körper gibt.

Was auch immer man tut, die folgenden Ergebnisse folgen nur aus den Eigenschaften und nicht aus der speziellen mengentheoretischen Konstruktion.

Proposition 6.4.5.

(1) *Zu je zwei reellen Zahlen x, y mit $x > 0$ existiert eine natürliche Zahl n so, dass*
$$nx > y$$
gilt. Das heißt, $\mathbb{R}$ besitzt die ***archimedische Eigenschaft****.*

(2) *Zwischen je zwei reellen Zahlen $x, y \in \mathbb{R}$ mit $x < y$ gibt es eine rationale Zahl $q \in \mathbb{Q}$ und eine irrationale Zahl $r \in \mathbb{R} \setminus \mathbb{Q}$, d.h. es gibt $q \in \mathbb{Q}$ und $r \in \mathbb{R} \setminus \mathbb{Q}$ mit*
$$x < q < y \quad \text{und} \quad x < r < y.$$
Man sagt auch $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ ***liegen dicht*** *in $\mathbb{R}$.*

Beweis. Wir beginnen mit der archimedischen Eigenschaft.

(1) Sei $A := \{nx \mid n \in \mathbb{N}\}$. Wäre die archimedische Eigenschaft nicht erfüllt, dann wäre y eine obere Schranke von A. Damit wäre A nach oben beschränkt und hätte ein Supremum, weil $\mathbb{R}$ die Supremumseigenschaft besitzt. Sei $\alpha := \sup A$. Wegen $x > 0$ ist $\alpha - x < \alpha$, also ist $\alpha - x$ keine obere Schranke von A. Somit existiert nach Definition von A eine natürliche Zahl n mit $\alpha - x < nx$. Dann ist aber $\alpha < (n + 1)x$, ein Widerspruch dazu, dass α obere Schranke von A ist. Also gilt die archimedische Eigenschaft.

(2) Wir beginnen mit der Dichtheit von $\mathbb{Q}$. Sei $x < y$ und damit $y - x > 0$. Wegen der archimedischen Eigenschaft gibt es eine natürliche Zahl n so, dass $n(y - x) > 1$ ist. Wir können auch natürliche Zahlen m_1 und m_2 finden mit $m_1 > nx$ und $m_2 > -nx$. Wir haben jetzt
$$-m_2 < nx < m_1,$$
was die Existenz einer ganzen Zahl m impliziert mit
$$m - 1 < nx < m \quad \text{und} \quad -m_2 \leq m \leq m_1.$$
Die Kombination aller dieser Ungleichungen liefert
$$nx < m < 1 + nx < ny$$
$$x < \tfrac{m}{n} < y,$$
wobei die letzte Ungleichung aus $n > 0$ folgt. Setzen wir $q = \frac{m}{n}$, so haben wir alles bewiesen, was behauptet wurde.

Nun zu $\mathbb{R} \setminus \mathbb{Q}$: Wenden wir obiges Argument zweimal an, so können wir rationale Zahlen q_1 und q_2 an mit $x < q_1 < q_2 < y$. Wir definieren
$$r := q_1 + \frac{q_2 - q_1}{2}\sqrt{2} > q_1.$$

Die Zahl r ist irrational, weil $\sqrt{2}$ irrational ist. Außerdem ist

$$q_2 - r = (q_2 - q_1)(1 + \tfrac{1}{\sqrt{2}}) > 0,$$

und deswegen gilt $x < q_1 < r < q_2 < y$. □

Eine weitere Eigenschaft von $\mathbb{R}$ betrifft das Wurzelziehen. Es folgt nämlich aus der Ordnungsvollständigkeit.

Proposition 6.4.6 (Wurzel)**.** *Für alle $a \in \mathbb{R}$ mit $a > 0$ und alle positiven $n \in \mathbb{N}$ gibt es genau ein $x \in \mathbb{R}$ mit $x > 0$ und $x^n = a$.*

Beweis. Beweisen wir zuerst die Eindeutigkeit: Sind $x \neq y$ zwei Lösungen, so ist o.B.d.A. $x < y$. Mit den Ordnungseigenschaften und vollständiger Induktion folgt dann für jedes $n \in \mathbb{N}$, dass $x^n < y^n$, also $x^n \neq y^n$.

Die Existenzaussagen ist für $n = 1$ oder $a \in \{0, 1\}$ trivial. Seien also zunächst $a > 1$ und $n \geq 2$. Dann definieren wir

$$A := \{x \in \mathbb{R} \mid x > 0 \wedge x^n \leq a\},$$

Weil $1 \in A$ liegt und $\forall x \in A : x < a$ gilt (denn mittels Proposition 6.3.2(3) und Induktion folgt $x \geq a \Rightarrow x^n \geq a^n > a$), wissen wir, dass $s = \sup A$ existiert.

Wir wollen jetzt beweisen, dass $s^n = a$ gilt.

Fall 1: Ist $s^n < a$, so definieren wir $b := (1 + s)^n - s^n > 0$ und wählen $0 < \varepsilon < \min\{1, \frac{a-s^n}{b}\}$. Dann folgt unter Verwendung von $\varepsilon < 1 \Rightarrow \varepsilon^n < \varepsilon$ für alle $1 < n \in \mathbb{N}$ (was ebenfalls mittels Induktion aus Proposition 6.3.2(3) folgt)

$$(s+\varepsilon)^n = \sum_{k=0}^{n-1} \binom{n}{k} s^k \varepsilon^{n-k} + s^n \leq \varepsilon \sum_{k=0}^{n-1} \binom{n}{k} s^k + s^n = \varepsilon b + s^n < a - s^s + s^n = a,$$

ein Widerspruch zur Supremumseigenschaft von s.

Fall 2: Ist $s^n > a$, so definieren bzw. wählen wir

$$c := \sum_{\substack{j \\ 0 \leq 2j-1 \leq n}} \binom{n}{2j-1} s^{n-2j+1} > 0, \text{ und } 0 < \varepsilon < \min\{1, \tfrac{s^n - a}{c}\}.$$

Dann rechnen wir nach (wobei wir wieder verwenden, dass $\varepsilon^n < \varepsilon$ gilt)

$$\begin{aligned}
(s-\varepsilon)^n &= s^n + \sum_{k=1}^{n} \binom{n}{k} s^{n-k}(-\varepsilon)^k \\
&\geq s^n + \sum_{k=1,\ k \text{ ungerade}}^{n} \binom{n}{k} (-1)^k \varepsilon^k s^{n-k} \\
&= s^n + \sum_{\substack{j \\ 0 \leq 2j-1 \leq n}} (-1)^{2j-1} \binom{n}{2j-1} s^{n-2j+1} \varepsilon^{2j-1} \\
&\geq s^n - \varepsilon \sum_{\substack{j \\ 0 \leq 2j-1 \leq n}} \binom{n}{2j-1} s^{n-2j+1} \\
&= s^n - \varepsilon c > s^n - s^n + a = a.
\end{aligned}$$

Dies widerspricht ebenfalls der Tatsache, dass $s = \sup A$ gilt.

Deshalb muss $s^n = a$ gelten, was wir zeigen wollten.

Ist schließlich $a < 1$, dann ist $\frac{1}{a} > 1$. Wir können also ein $y \in \mathbb{R}$ finden mit $y^n = \frac{1}{a}$. Dann aber gilt für $x = \frac{1}{y}$, dass $x^n = a$ ist. $\square$

Es gibt auch bei den irrationalen Zahlen noch gewisse Unterschiede. Die Zahl $\sqrt{2}$ tritt als Nullstelle eines Polynoms mit rationalen Koeffizienten auf. Es ist $\sqrt{2}$ nämlich Nullstelle von $x^2 - 2$.

Definition 6.4.7. *Eine reelle Zahl r heißt* algebraisch*, wenn es $n \in \mathbb{N}$ und rationale Zahlen $a_0, \dots, a_n$ gibt mit*

$$\sum_{i=0}^{n} a_i r^i = 0.$$

Eine bis Cantor ungelöste Frage war, ob alle irrationalen Zahlen algebraisch sind. Er hat diese Frage für die damalige Zeit recht überraschend gelöst, denn jedes rationale Polynom n–ten Grades besitzt höchstens n Nullstellen (siehe Korollar 6.5.11). Ferner gibt es nur abzählbar viele rationale Polynome, die also insgesamt höchstens abzählbar viele Nullstellen besitzen können. Die Mächtigkeit der Menge $\mathbb{R}_a$ der algebraischen Zahlen ist also $\aleph_0$.

Cantor hat aber auch bewiesen, dass $|\mathbb{R}| = c > \aleph_0$ gilt. Aus diesem Grund ist $\mathbb{R}_t := \mathbb{R} \setminus \mathbb{R}_a \neq \emptyset$, ja es gilt sogar $|\mathbb{R}_t| = c$. Die Elemente von $\mathbb{R}_t$ heißen **transzendente Zahlen**. Z.B. sind π und e transzendent. Ersteres hat übrigens Ferdinand Lindemann (1852–1939) im April 1882 bewiesen.

Zum Abschluss dieses Abschnitts wollen wir noch die Begriffe des Absolutbetrags einer und des Abstands zweier reeller Zahlen diskutieren, die wichtige Werkzeuge der Analysis darstellen.

Definition 6.4.8 (Absolutbetrag)**.** *Seien $x, y \in \mathbb{R}$.*

(i) *Wir definieren den* Absolutbetrag *(oder einfach* Betrag*) von x durch*

$$|x| := \begin{cases} x & \text{falls } x \geq 0 \\ -x & \text{falls } x < 0. \end{cases}$$

(ii) *Unter dem* Abstand *von x und y verstehen wir die Zahl $|x-y|$. (Es ist also $|x-y| = x-y$ oder $|x-y| = y-x$, je nachdem ob $x > y$ oder $y > x$.)*

(iii) *Das* Vorzeichen *oder* Signum *von x ist definiert als*

$$\operatorname{sgn}(x) := \begin{cases} 1 & \text{falls } x > 0 \\ -1 & \text{falls } x < 0 \\ 0 & \text{falls } x = 0. \end{cases}$$

Diese Definition hat die sog. „Spiegelungssymmetrie“

$$|x| = |-x| \text{ und } |x-y| = |y-x|$$

als offensichtliche Konsequenz (Fallunterscheidung!) Weitere gültige Eigenschaften fassen wir in der folgenden Proposition zusammen.

Proposition 6.4.9 (Eigenschaften von Betrag und Abstand)**.** *Für $x, y, z \in \mathbb{R}$ gilt*

(1) $|x| \geq 0$ *und* $|x| = 0 \Leftrightarrow x = 0$ *(Definitheit).*
(2) $|x+y| \leq |x| + |y|$ *(Dreiecksungleichung).*
(3) $|xy| = |x||y|$ *(Multiplikativität).*
(4) $|x-y| \geq 0$ *und* $|x-y| = 0 \Leftrightarrow x = y$ *(Definitheit).*
(5) $|x-z| \leq |x-y| + |y-z|$ *(Dreiecksungleichung).*

Beweis.

(1) Die erste Behauptung folgt sofort aus der Definition des Betrags und Proposition 6.3.2(2). Die „Rückrichtung“ der zweiten Behauptung folgt sofort aus der Definition und die „Hinrichtung“ aus der Tatsache, dass wiederum wegen der Definition des Betrags aus $x \neq 0$ folgt, dass $|x| \neq 0$.
(2) Folgt mittels Fallunterscheidung oder mittels der Aussage
$$a \leq b \text{ und } -a \leq b \Rightarrow |a| \leq b \tag{6.11}$$
(die ebenfalls unmittelbar aus Definition 6.4.8(i) folgt). Denn aus Addition der offensichtlichen Ungleichungen $x \leq |x|$, $-x \leq |x|$ sowie $y \leq |y|$, $-y \leq |y|$ folgt
$$x + y \leq |x| + |y| \text{ und } -x - y \leq |x| + |y|,$$
was mittels (6.11) die Behauptung ergibt.
(3) Folgt ebenfalls mittels Fallunterscheidung.
(4) Ergibt sich sofort aus (1).
(5) Wegen der Dreiecksungleichung für den Betrag gilt
$$|x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z|.$$
□

6.4.1. Die mengentheoretische Konstruktion von $\mathbb{R}$. Das einzige, das uns noch fehlt in unserer Untersuchung über die reellen Zahlen ist der Beweis von Theorem 6.4.4. Wir werden diesen gesamten Abschnitt dafür opfern und $\mathbb{R}$ aus $\mathbb{Q}$ durch mengentheoretische Mechanismen konstruieren. Zu diesem Zweck werden wir die von Dedekind erfundenen Schnitte verwenden. Es gibt viele äquivalente Verfahren zur Konstruktion von $\mathbb{R}$ aus $\mathbb{Q}$. Die Dedekindschen Schnitte sind nicht die einleuchtendste Methode aber jedenfalls diejenige, die nur Mengenoperationen verwendet.

Definition 6.4.10. *Eine nichtleere nach unten beschränkte Teilmenge $S \subseteq \mathbb{Q}$ heißt **Schnitt** (von $\mathbb{Q}$), falls*
$$\textbf{(S1)} \quad \forall q \in \mathbb{Q} \setminus S : \forall s \in S : s \geq q, \quad \text{und}$$
$$\textbf{(S2)} \quad \forall s \in S : \exists s' \in S : s > s'.$$

Motivierend kann man erklären, dass ein Schnitt ein halboffenes Intervall $]\,a, +\infty\,[\cap \mathbb{Q}$ mit $a \in \mathbb{R}$ ist. Noch dürfen wir das allerdings nicht sagen.

Proposition 6.4.11.
(1) *Sei S ein Schnitt. Es gilt*
$$\forall S \in s : \forall q \in \mathbb{Q} : (s \leq q \Rightarrow q \in S).$$
Ist also eine rationale Zahl größer als ein Element des Schnittes, dann liegt sie im Schnitt.
(2) *Zu jeder positiven rationalen Zahl ε gibt es $q, r \in \mathbb{Q}$ mit $q \in S$, $r \in \mathbb{Q} \setminus S$ und $q - r \leq \varepsilon$.*

Beweis.
(1) Seien $s \in S$ und $q \in \mathbb{Q}$ mit $s \leq q$. Ist $q \notin S$, dann liegt natürlich $q \in \mathbb{Q} \setminus S$ und daher gilt $\forall s' \in S : s' \geq q$. Daher ist auch $s \geq q$, und weil $\leq$ eine Ordnungsrelation ist, folgt $s = q$. Das ist ein Widerspruch zu $q \notin S$. Daher ist $q \in S$, und wir sind fertig.
(2) Sei $0 < \varepsilon \in \mathbb{Q}$. Weil S ein Schnitt ist, gibt es $q \in S$ und $r \in \mathbb{Q} \setminus S$. Ist $q - r \leq \varepsilon$, dann sind wir fertig. Andernfalls sei $n \in \mathbb{N}$ so groß, dass $n > \frac{q-r}{\varepsilon}$ gilt. Solch ein n existiert wegen Proposition 6.3.3. Wir bilden nun die Menge
$$M := \{r + k\tfrac{q-r}{n} \mid k \in \{0, \ldots, n\}\} \subseteq \mathbb{Q}.$$

Für $q \in M \cap S$ und $r \in M \cap (\mathbb{Q} \setminus S)$. Es existiert ein kleinstes Element $q_m \in M \cap S$, weil M endlich ist. Dann ist $r_m := q_m - \frac{q-r}{n} \in M \cap (\mathbb{Q} \setminus S)$, und wir haben zwei rationale Zahlen q_m und r_m wie benötigt gefunden, da $q_m - r_m = \frac{q-r}{n} < \varepsilon$ gilt.

□

Definition 6.4.12. *Sei $R \subseteq \mathbb{PQ}$ die Menge aller Schnitte von $\mathbb{Q}$. Wir definieren auf R die Relation $\leq$ durch*

$$S \leq T := S \supseteq T. \tag{6.12}$$

Proposition 6.4.13. *Die Relation $\leq$ macht R zu einer totalgeordneten Menge.*

Beweis. Wir müssen die Ordnungseigenschaften überprüfen. Halbordnung ist eigentlich klar, da $\supseteq$ eine Halbordnung auf $\mathbb{PQ}$ bildet, doch wir schreiben alles noch einmal auf:

Reflexivität: Es für jede Menge $S \supseteq S$.

Symmetrie: Sind $S \supseteq T$ und $T \supseteq S$ erfüllt, so ist $S = T$.

Transitivität: Seien $S \supseteq T$ und $T \supseteq U$. Ist $u \in U$, dann ist $u \in T$, und daher gilt $u \in S$. Das impliziert $S \supseteq U$.

Es bleibt zu zeigen, dass $\leq$ eine Totalordnung ist. Seien S und T zwei Schnitte und $S \neq T$. Ist $S \not\leq T$, dann ist $S \not\supseteq T$, und daher gibt es ein $t \in T$ mit $t \notin S$. In diesem Fall liegt $t \in \mathbb{Q} \setminus S$, also ist für alle $s \in S$ die Ungleichung $s \geq t$ erfüllt. Wegen Proposition 6.4.11.(1) bedeutet das aber $s \in T$, und das impliziert $S \subseteq T$, also $S \geq T$. Damit sind je zwei Schnitte vergleichbar, und $\leq$ ist eine Totalordnung auf R. □

Definition 6.4.14. *Als nächstes führen wir die Abbildung $+ : R \times R \to \mathbb{PQ}$ durch*

$$S + T := \{s + t \mid s \in S \wedge t \in T\} \quad \textit{für } S, T \in R$$

ein.

Proposition 6.4.15. *Diese Abbildung führt sogar wieder nach R. Es ist also $S + T$ wieder ein Schnitt:*

Beweis.

- Sind $s \in S$ und $t \in T$, dann ist $s + t \in S + T$, also ist $S + T \neq \emptyset$.
- Sei σ untere Schranke von S und τ untere Schranke von T. Für beliebiges $x \in S+T$ gibt es $s \in S$ und $t \in T$ mit $x = s + t$. Aus den Eigenschaften von $\leq$ auf $\mathbb{Q}$ folgt ferner $x = s + t \leq \sigma + \tau$. Daher ist $S + T$ nach unten beschränkt.
- Betrachten wir $q \in \mathbb{Q} \setminus (S + T)$. Sei $s \in S$ gegeben, wir wissen $\forall t \in T : s + t \neq q$. Wir formen das um zu $\forall t \in T : t \neq q - s$, und daher ist $q - s \in \mathbb{Q} \setminus T$. Weil T ein Schnitt ist, folgt $\forall t \in T : t \geq q - s$. Bringen wir s zurück auf die linke Seite, ergibt das $\forall t \in T : s+t \geq q$, darum gilt für alle $x \in S+T$, dass $x \geq q$, also ist Eigenschaft S1 erfüllt.
- Sei $x \in S + T$ beliebig. Dann existieren $s \in S$ und $t \in T$ mit $s + t = x$. Weil S und T Schnitte sind, gibt es $s' \in S$ und $t' \in T$ mit $s > s'$ und $t > t'$. Daher ist $x' = s' + t' \in S + T$, und es gilt $x > x'$. Das weist Eigenschaft S2 nach.

□

Das beweist, dass $(R, +)$ ein Gruppoid bildet. Bevor wir die weiteren Eigenschaften nachweisen, betrachten wir noch ein Klasse spezieller Schnitte.

Definition 6.4.16. *Ein Schnitt S heißt **rational**, falls er ein Infimum besitzt.*

Proposition 6.4.17. *Ein Schnitt S ist genau dann rational, wenn es ein $q \in \mathbb{Q}$ gibt mit*

$$S = \mathbb{S}_q := \{q' \in \mathbb{Q} \mid q' > q\}. \tag{6.13}$$

Beweis. Sei S ein Schnitt von der Form (6.13). Nun ist q eine untere Schranke von S, und falls $q' \in \mathbb{Q}$ mit $q' > q$, dann ist q' keine untere Schranke von S. Es ist nämlich $q' > \frac{1}{2}(q'+q) > q$, und daher $\frac{1}{2}(q'+q) \in S$. Daher ist q das Infimum von S und S rational.

Nun sei S ein rationaler Schnitt. Es existiert $q = \inf S$, und wir definieren $\mathbb{S}_q = \{q' \in \mathbb{Q} \mid q' > q\}$. Weil q untere Schranke von S ist, folgt $S \subseteq \mathbb{S}_q$. Sei nun $t \in \mathbb{S}_q$. Falls $t \notin S$ gilt, wissen wir, dass $\forall s \in S : s \geq t$. Daher ist t eine untere Schranke von S mit $t > q$. Das widerspricht der Infimumseigenschaft von q. Daher ist $t \in S$ und $S = \mathbb{S}_q$. □

Auf diese Weise sehen wir, dass für je zwei rationale Zahlen q und r die zugehörigen rationalen Schnitte $\mathbb{S}_q$ und $\mathbb{S}_r$ genau dann gleich sind, wenn $q = r$. Die Abbildung $\iota : \mathbb{Q} \to R$ mit $\iota : q \mapsto \mathbb{S}_q$ ist also injektiv. Auf diese Weise wird $\mathbb{Q}$ in R eingebettet, und wir können in Zukunft die rationale Zahl q mit dem Schnitt $\mathbb{S}_q$ identifizieren.

Proposition 6.4.18. $(G, +)$ *ist eine abelsche Gruppe.*

Beweis. Wir weisen sukzessive alle Eigenschaften nach:

AG: Seien S, T und U Schnitte.

$$\begin{aligned}(S+T)+U &= \{x+u \mid x \in S+T, u \in U\} = \{(s+t)+u \mid s \in S, t \in T, u \in U\} = \\ &= \{s+(t+u) \mid s \in S, t \in T, u \in U\} = \{s+y \mid s \in S, y \in T+U\} = \\ &= S+(T+U).\end{aligned}$$

KG: Für zwei Schnitte S und T sind die Mengen $S+T$ und $T+S$ gleich, weil die Addition in $\mathbb{Q}$ kommutativ ist.

Nullelement: Der rationale Schnitt $0 := \{q \in \mathbb{Q} \mid q > 0\} = \mathbb{S}_0$ ist das Nullelement. Sei nämlich T ein beliebiger Schnitt. Dann erhalten wir

$$0 + T = \{s + t \mid s \in 0, t \in T\}.$$

Wir müssen nachweisen, dass $0 + T = T$ gilt. Sei $x \in 0 + T$, dann gibt es $s \in 0$ und $t \in T$ mit $s + t = x$. Wegen $s > 0$ ist $x > t$ und damit gilt $x \in T$, also $0 + T \subseteq T$. Umgekehrt sei $t \in T$. Weil T ein Schnitt ist, gibt es ein $t' \in T$ mit $t' < t$. Setzen wir nun $s = t - t'$, dann ist $s \in 0$ und $t = s + t' \in S + T$, was wiederum $T \subseteq 0 + T$ beweist.

Inverse: Betrachten wir wieder einen Schnitt S. Wir definieren

$$-S := \{q \in \mathbb{Q} \mid \forall s \in S : q > -s \ \wedge\ \forall t \in \mathbb{Q} : (t = \inf S \Rightarrow q \neq -t)\}$$

den zu S negativen Schnitt. Wir behaupten $S + (-S) = 0$. Zuerst müssen wir aber zeigen, dass $-S$ tatsächlich ein Schnitt ist.

Sei q' eine untere Schranke von S. Dann gilt $\forall s \in S : q = q' - 1 < s$ und deshalb $\forall s \in S : -q > -s$, also ist $-S$ nichtleer. Für ein beliebiges Element $s \in S$ folgt, dass jedes Element $s' \in -S$ die Ungleichung $s' \geq -s$ erfüllen muss, also ist $-s$ eine untere Schranke von $-S$.

Sei nun $q \in \mathbb{Q} \setminus (-S)$. Dann gibt es $s \in S$ mit $q \leq -s$, also $-q \geq s$. Weil S ein Schnitt ist, folgt $-q \in S$. Darum gilt aber $\forall t \in (-S) : t > q$. Das beweist S1.

S2 beweisen wir indirekt. Sei $q \in (-S)$ gegeben mit $\forall t \in (-S) : q \leq t$. Dann ist q eine untere Schranke von $-S$, also ein Minimum und erst recht ein Infimum von $-S$. Das ist aber unmöglich wegen der Definition von $-S$. Falls $\tilde{s} := \inf S$ existiert, dann ist $-\tilde{s}$ das Supremum der Menge $\tilde{S} := \{-s \mid s \in S\}$, des Komplements von $-S$, und damit das Infimum von $-S$. Nach Definition ist $-\tilde{s} \notin (-S)$.

Sei $x \in S + (-S)$, dann existieren $s \in S$ und $t \in -S$ mit $s + t = x$. Dass $t \in -S$ liegt, impliziert $t > -s$ und damit auch $x = s + t > 0$. Daher ist $S + (-S) \subseteq 0$. Nun sei $y > 0$. Wir suchen gemäß Proposition 6.4.11.(2) zwei rationale Zahlen q und r mit $q \in S$, $r \in Q \setminus S$ und $q - r < y$. Es gilt $\forall s \in S : s > r$, und daher ist $-r \in -S$.

Wir definieren $r' := q - r$ und wissen $r' < y$, also $y - r' > 0$. Weil S ein Schnitt ist, bedeutet das $s := y - r' + q \in S$. Setzen wir nun zusammen, so haben wir $-r \in -S$ und $s \in S$ mit

$$-r + s = -r + y - r' + q = -r + y - q + r + q = y.$$

Das impliziert $y \in S + (-S)$ und daher ist $0 = S + (-S)$.

□

Die Verträglichkeit von $+$ und $\leq$, also O1 beweisen wir als nächstes (siehe Definition 6.3.1).

Proposition 6.4.19. *Für je drei Elemente S, T und U von R gilt*

$$S \leq T \implies S + U \leq T + U.$$

Beweis. Seien drei Schnitte S, T und U gegeben mit $S \leq T$. Sei $y \in T + U$. Dann existieren $t \in T$ und $u \in U$ mit $y = t + u$. Weil $s \leq T$ gilt, wissen wir $S \supseteq T$ und damit $t \in S$. Daher ist $t + u = y$ auch in $S + U$, was wiederum $S + U \leq T + U$ bestätigt. □

Ein Schnitt S heißt positiv, falls $S > 0$ gilt. Er heißt nichtnegativ, falls $S \geq 0$ erfüllt ist. Analog führen wir die Bezeichnungen negativ und nichtpositiv ein. Für einen negativen Schnitt S ist $-S$ positiv. Das folgt aus der Verträglichkeit von $+$ und $\leq$ in Proposition 6.4.19.

Es fehlt zum Körper die zweite Operation.

Definition 6.4.20. *Wir definieren die Abbildung $\cdot : R \times R \to \mathbb{P}\mathbb{Q}$ wie folgt: Für zwei nichtnegative Schnitte S und T sei*

$$S \cdot T := \{st \mid s \in S \wedge t \in T\}.$$

Darüber hinaus erklären wir

$$S \cdot T := \begin{cases} -((-S) \cdot T) & \text{falls } S < 0 \text{ und } T \geq 0 \\ -(S \cdot (-T)) & \text{falls } S \geq 0 \text{ und } T < 0 \\ (-S) \cdot (-T) & \text{falls } S < 0 \text{ und } T < 0. \end{cases}$$

Wegen der Bemerkungen vor der Definition ist die Abbildung $\cdot$ wohldefiniert.

Proposition 6.4.21. *Die Abbildung $\cdot$ ist eine Verknüpfung auf R. Es gilt $(R, +, \cdot)$ ist ein Körper.*

Beweis. Zuerst müssen wir beweisen, dass für nichtnegative Schnitte S und T die Menge $S \cdot T$ wieder ein Schnitt ist. Es existiert $s \in S$ und $t \in T$, daher ist $st \in S \cdot T$, welches somit nichtleer ist.

Weil S und T nichtnegativ sind, folgt $0 \supseteq S$ und $0 \supseteq T$, und daher ist $0 \in \mathbb{Q}$ untere Schranke von S und T. Wir erhalten $\forall s \in S : 0 \leq s$ und $\forall t \in T : 0 \leq t$. Wegen Proposition 6.3.2.(3) gilt $\forall s \in S : \forall t \in T : 0 \leq st$, und daher ist $S \cdot T$ nach unten beschränkt.

Sei $q \in \mathbb{Q} \setminus (S \cdot T)$. Ist $s \in S$ beliebig, dann gilt $\forall t \in T : st \neq q$, und daher haben wir wegen $s > 0$ auch $\forall t \in T : t \neq q/s$. Das wiederum bedingt, dass $q/s \in \mathbb{Q} \setminus T$ liegt, weshalb $\forall t \in T : t \geq q/s$. Umgeformt bedeutet das $\forall t \in T : ts \geq q$, was S1 impliziert.

Für $y \in S \cdot T$ existieren $s \in S$ und $t \in T$ mit $y = st$. Ferner gibt es $s' \in S$ mit $s' < s$ und $t' \in T$ mit $t' < t$. Weil alle Zahlen s, s', t, t' größer Null sind, folgt aus Proposition 6.3.2 $s't' < st$, woraus S2 folgt.

Für nichtnegative Schnitte ist das Produkt also wieder ein Schnitt. In den anderen Fällen wird die Definition auf ein Produkt nichtnegativer Schnitte zurückgeführt, und daher ist $\cdot$ tatsächlich eine Verknüpfung auf R.

Wir wissen bereits, dass $(R,+)$ eine abelsche Gruppe ist. Der Rest der Körperaxiome muss noch nachgewiesen werden. Beginnen wir mit den Aussagen über die Multiplikation, doch zuvor wollen wir noch ein Hilfsresultat über positive Schnitte beweisen.

Lemma: Sei S ein positiver Schnitt. Es gibt ein $q > 0$ in $\mathbb{Q}$, das untere Schranke von S ist.

Beweis: Wegen $S > 0$ folgt, dass $0 \supsetneq S$ gilt, und daher existiert ein $q \in 0$ mit $q \notin S$, also $q \in \mathbb{Q} \setminus S$. Es gilt $0 < q$, weil $q \in 0$ und $\forall s \in S : q \leq S$, wegen S1. □

AG: Das Assoziativgesetz für positive Schnitte folgt direkt aus dem Assoziativgesetz für die Multiplikation rationaler Zahlen. Seien S, T und U nichtnegative Schnitte.

$$\begin{aligned}(S \cdot T) \cdot U &= \{xu \mid x \in S \cdot T, u \in U\} = \{(st)u \mid s \in S, t \in T, u \in U\} = \\ &= \{s(tu) \mid s \in S, t \in T, u \in U\} = \{sy \mid s \in S, y \in T \cdot U\} = \\ &= S \cdot (T \cdot U).\end{aligned}$$

Seien nun S, T und U beliebig. Mit einer Anzahl einfacher Fallunterscheidungen kann man das Assoziativgesetz auf den positiven Fall zurückführen. Sei etwa $S < 0$, $T \geq 0$ und $U \geq 0$. Dann folgt

$$\begin{aligned}(S \cdot T) \cdot U &= -((-S) \cdot T) \cdot U = -(((-S) \cdot T) \cdot U) = -((-S) \cdot (T \cdot U)) = \\ &= -(-S) \cdot (T \cdot U) = S \cdot (T \cdot U).\end{aligned}$$

All die anderen sechs Fälle beweist man analog.

KG: Auch die Kommutativität für nichtnegative Schnitte folgt aus der Kommutativität der Multiplikation in $\mathbb{Q}$. Für beliebige Schnitte folgt sie aus der Symmetrie von Definition 6.4.20.

Einselement: Wir definieren $1 := \mathbb{S}_1 = \{x \in \mathbb{Q} \mid x > 1\}$ und behaupten, dass 1 das Einselement bezüglich der Multiplikation ist. Es gilt $1 \neq 0$, und wir betrachten einen nichtnegativen Schnitt S. Für $s \in 1 \cdot S$ gibt es $t \in 1$ und $s' \in S$ mit $s = ts'$. Weil $t > 1$ ist, folgt aus Proposition 6.3.2, dass $s > s'$ und damit auch $s \in S$ gilt. Daher ist $1 \cdot S \subseteq S$.

Sei nun umgekehrt $s \in S$ gegeben. Wir können $s' \in S$ finden mit $0 < s' < s$, weil S ein nichtnegativer Schnitt ist. Aus Proposition 6.3.2 folgt, dass $t = \frac{s}{s'} > 1$ ist, also $t \in 1$, und außerdem wissen wir $ts' = s$. Daher ist auch $S \subseteq 1 \cdot S$.

Für negatives S gilt $1 \cdot S = -(1 \cdot (-S)) = -(-S) = S$.

Inverse: Sei S ein positiver Schnitt. Wir definieren

$$S^{-1} := \{q \in \mathbb{Q} \mid \forall s \in S : q > \tfrac{1}{s} \wedge \forall t \in \mathbb{Q} : (t = \inf S \Rightarrow q \neq \tfrac{1}{t})\}$$

und behaupten das multiplikative Inverse zu S gefunden zu haben.

Zuerst müssen wir beweisen, dass S^{-1} ein Schnitt ist. Wegen des Lemmas existiert eine positiver rationale Zahl q', die untere Schranke von S ist. Die Zahl $q = \frac{q'}{2}$ erfüllt dann für alle $s \in S$, dass $s > q$ und daher $\frac{1}{s} < \frac{1}{q}$, also ist $\frac{1}{q} \in S^{-1}$.

Sei $q \in \mathbb{Q} \setminus S^{-1}$. O.b.d.A. gilt $q > 0$, denn alle Elemente von S^{-1} sind positiv. Es folgt, dass es ein $s \in S$ gibt, für das $q \leq \frac{1}{s}$ erfüllt ist. Aus den Eigenschaften der Ordnungsrelationen folgt aber dann $\frac{1}{q} \geq s$, und daher ist $\frac{1}{q} \in S$. Daher gilt $\forall t \in S^{-1} : t > q$. Das zeigt S1.

S2 folgt wieder aus der Definition von S^{-1}. Ist $q \in S^{-1}$ gegeben mit $\forall s \in S^{-1} : q \leq s$, dann ist q Minimum also Infimum von S^{-1}. Aus der Definition von S^{-1} kann man aber ablesen, dass S^{-1} sein (eventuell existierendes) Infimum nicht enthalten darf.

Nachdem wir jetzt gezeigt haben, dass S^{-1} tatsächlich ein Schnitt ist, müssen wir beweisen, dass S^{-1} das Inverse von S ist. Sei also $q \in S \cdot S^{-1}$. Dann existieren

$s \in S$ und $t \in S^{-1}$ mit $st = q$. Weil $t \in S^{-1}$ folgt, dass $t > \frac{1}{s}$, und daher ist $st > 1$, woraus $S \cdot S^{-1} \subseteq 1$ folgt.

Sei umgekehrt $y \in 1$ gegeben. Wir definieren $\varepsilon = y - 1 > 0$ und wählen uns gemäß dem Lemma eine positive untere Schranke r' von S. Außerdem können wir wegen Proposition 6.4.11.(1) zwei rationale Zahlen $\tilde{r}$ und s mit $r \in \mathbb{Q} \setminus S$ und $s \in S$ und $s - \tilde{r} < r'\varepsilon$ finden. Sei $r = \max\{\tilde{r}, r'\}$. Dann ist immer noch $r \in \mathbb{Q} \setminus S$ und $s - r < r'\varepsilon$. Für r und s gilt darüber hinaus noch

$$\tfrac{s}{r} - 1 < \tfrac{r'\varepsilon}{r} < \varepsilon \quad \text{also} \quad \tfrac{s}{r} < 1 + \varepsilon = y.$$

Wir definieren $t := \frac{yr}{s} > 1$. Dann sind $s < st =: s' \in S$ und $\frac{1}{r} \in S^{-1}$ und weiters

$$s'\tfrac{1}{r} = \tfrac{st}{r} = \tfrac{yrs}{rs} = y,$$

also ist $y \in S \cdot S^{-1}$, und das impliziert $S \cdot S^{-1} = 1$.

Ist S negativ, dann definieren wir $S^{-1} := -((-S)^{-1})$, und wir haben

$$S \cdot S^{-1} = (-S) \cdot (-S^{-1}) = (-S) \cdot (-S)^{-1} = 1.$$

Zu guter letzt fehlt noch das **Distributivgesetz**. Wir beginnen wieder mit nichtnegativen Schnitten S, T und U. Wegen der Distributivität in $\mathbb{Q}$ gilt

$$\begin{aligned}(S+T) \cdot U &= \{xu \mid x \in S+T, u \in U\} = \{(s+t)u \mid s \in S, t \in T, u \in U\} = \\ &= \{su + tu \mid s \in S, t \in T, u \in U\} = \{y + z \mid y \in S \cdot U, z \in T \cdot U\} = \\ &= S \cdot U + T \cdot U.\end{aligned}$$

Für die sieben übrigen Fälle sei als Beispiel einer bewiesen: Mit $U < 0$ und $S \geq 0$ und $T \geq 0$ gilt

$$\begin{aligned}(S+T) \cdot U &= -((S+T) \cdot (-U)) = -(S \cdot (-U) + T \cdot (-U)) = \\ &= -(S \cdot (-U)) + (-(T \cdot (-U))) = S \cdot U + T \cdot U.\end{aligned}$$

Das beweist, dass $(R, +, \cdot)$ ein Körper ist. □

Proposition 6.4.22. *Der Körper $(R, +, \cdot)$ ist geordnet bezüglich $\leq$.*

Beweis. Es genügt O2 zu beweisen, denn O1 haben wir in Proposition 6.4.19 bereits nachgewiesen. Seien also $S > 0$ und $T > 0$. Dann gibt es positive untere Schranken $\underline{s}$ von S und $\underline{t}$ von T, und daher ist $\underline{st} > 0$ eine untere Schranke von $S \cdot T$. Das impliziert $S \cdot T > 0$. □

Nachdem wir nachgewiesen haben, dass die Menge aller Schnitte R einen geordneten Körper bildet, bleibt noch die letzte Eigenschaft nachzuweisen.

Proposition 6.4.23. *Der geordnete Körper $(R, +, \cdot, \leq)$ ist ordnungsvollständig.*

Beweis. Sei E eine nach unten beschränkte Teilmenge von R. Sei $Q \in R$ eine untere Schranke von E, und sei $\alpha \in \mathbb{Q}$ untere Schranke von Q.

Wir betrachten die Menge

$$S := \bigcup_{T \in E} T,$$

die Vereinigung aller Elemente von E.

Wir zeigen zuerst, dass S ein Schnitt ist. Es ist klar, dass S nichtleer ist, denn jedes $T \in E$ ist nichtleer. Außerdem ist α untere Schranke von jedem $T \in E$ (weil $T \subseteq Q$), und daher auch untere Schranke der Vereinigung.

Nun wählen wir ein $q \in \mathbb{Q} \setminus S$. Für dieses Element gilt, dass $\forall T \in E : q \notin T$, also $\forall T \in E : q \in \mathbb{Q} \setminus T$. Für ein beliebiges $s \in S$ gilt nun, dass $\exists T \in E : s \in T$, und daher muss $q \leq s$ sein, was S1 beweist.

Schließlich gilt S2, weil für beliebiges $s \in S$ wieder ein $T \in E$ existiert mit $s \in T$. Da T ein Schnitt ist, gibt es ein $s' \in T$, sodass $s' < s$ gilt. Nun ist aber s' in der Vereinigung aller T, also $s' \in S$.

Wir beschließen den Beweis mit der Behauptung, dass $S = \inf E$ gilt. Offensichtlich ist S untere Schranke von E, da $S \supseteq T$ für alle $T \in E$ erfüllt ist. Sei nun $U \in R$ ein Schnitt mit $U > S$. Dann ist $S \supsetneq U$, und daher existiert ein $s \in S$ mit $s \notin U$. Nun muss es aber ein $T \in E$ geben mit $s \in T$, woraus folgt, dass $U \not\supseteq T$ gilt, also ist U keine untere Schranke von E. Daher stimmt tatsächlich $S = \inf E$ und R ist ordnungsvollständig. □

Lemma 6.4.24. *Sei $(S, +, \cdot, \leq)$ ein ordnungsvollständiger geordneter Körper mit geordnetem Unterkörper $\mathbb{Q}$. Dann ist jedes Element $b \in S$ das Infimum der Menge*

$$\mathbb{T}_b := \{s \in \mathbb{Q} : b < s\}.$$

Beweis. Die Menge $\mathbb{T}_b$ ist durch b nach unten beschränkt, und daher existiert das Infimum $\inf \mathbb{T}_b =: b' \in S$. Angenommen, es gilt $b' > b$. Aus Proposition 6.4.5, für deren Beweis wir nur die Eigenschaften geordneter Körper und Ordnungsvollständigkeit verwendet haben, folgt, dass es ein $q \in \mathbb{Q}$ gibt mit $b < q < b'$. Dann ist aber $q \in \mathbb{T}_b$, und daher ist b' keine untere Schranke von $\mathbb{T}_b$. Das ist ein Widerspruch, also ist $b' = b$. □

Proposition 6.4.25. *Sei $(S, +, \cdot, \leq)$ ein weiterer ordnungsvollständiger geordneter Körper, der $\mathbb{Q}$ als geordneten Unterkörper enthält, dann sind S und R isomorph.*

Beweis. Die Abbildung $f : S \to R$ gegeben durch $f : s \mapsto \mathbb{T}_s$ ist ein monotoner Körperisomorphismus.

Zunächst ist f wohldefiniert, denn jedes $\mathbb{T}_s$ ist ein Schnitt von $\mathbb{Q}$: Dass $\mathbb{T}_s$ nichtleer ist, folgt aus der Unbeschränktheit von $\mathbb{Q}$ in S. Weil s eine untere Schranke von $\mathbb{T}_s$ ist, existiert auch eine rationale Zahl $\tilde{s} < s$, die untere Schranke von $\mathbb{T}_s$ ist. Gilt $q \in \mathbb{Q} \setminus \mathbb{T}_s$, dann muss $q \leq s$ sein wegen der Definition von $\mathbb{T}_s$. Daher ist q ebenfalls untere Schranke von $\mathbb{T}_s$, was S1 beweist. Ist schließlich $r \in \mathbb{T}_s$ eine rationale Zahl, dann können wir wieder Proposition 6.4.5 verwenden, um eine rationale Zahl r' zu erhalten, mit $s < r' < r$, also $r' \in \mathbb{T}_s$, gleichbedeutend mit der Gültigkeit von S2.

Zuerst zeigen wir die Injektivität von f. Seien $s \neq s'$ zwei Elemente von S. O.B.d.A. ist $s > s'$. Dann gilt $\mathbb{T}_s \subsetneq \mathbb{T}_{s'}$, weil es eine rationale Zahl zwischen s und s' gibt (wieder Proposition 6.4.5). Daher ist $f(s) \neq f(s')$.

Die Abbildung f ist surjektiv. Ist T ein beliebiger Schnitt von $\mathbb{Q}$, dann ist $T \subseteq S$ nichtleer und nach unten beschränkt, besitzt also ein Infimum $s \in S$. Sei $t \in T$, dann gilt $t > s$, weil wegen der Schnitteigenschaft S2 die Menge T ihr Infimum nicht enthält. Daher ist $t \in \mathbb{T}_s$. Sei umgekehrt $t \in \mathbb{T}_s$ und damit $t > s$. Ist $t \notin T$, dann folgt aus der Schnitteigenschaft S1, dass $\forall t' \in T : t \leq t'$, also ist t eine untere Schranke von T mit $t > s$, was der Infimumseigenschaft von s widerspricht. Darum gilt $T = \mathbb{T}_s = f(s)$.

Es bleibt zu zeigen, dass f ein Körperhomomorphismus ist.

- Seien $s, t \in S$. Dann ist $f(s) + f(t) = \mathbb{T}_s + \mathbb{T}_t$. Es gilt

$$\begin{aligned}\mathbb{T}_s + \mathbb{T}_t = \{s' + t' \mid s' \in \mathbb{T}_s, t' \in \mathbb{T}_t\} = \{s' + t' \mid s' > s \wedge t' > t\} = \\ = \{s' + t' \mid s' + t' > s + t\} = \mathbb{T}_{s+t} = f(s + t).\end{aligned}$$

- Für $s \in S$ folgt

$$\begin{aligned}-f(s) = -\mathbb{T}_s = \{s' \in \mathbb{Q} \mid \forall t \in \mathbb{T}_s : s' > -t \wedge \forall t' \in \mathbb{Q} : (t' = \inf \mathbb{T}_s \Rightarrow s' \neq -t')\} = \\ = \{s' \in \mathbb{Q} \mid s' > -s\} = \mathbb{T}_{-s} = f(-s).\end{aligned}$$

- Sind wieder $s, t \in S$. Dann folgt für $s \geq 0$ und $t \geq 0$, dass

$$\begin{aligned}\mathbb{T}_s \cdot \mathbb{T}_t &= \{s't' \mid s' \in \mathbb{T}_s, t' \in \mathbb{T}_t\} = \{s't' \mid s' > s \wedge t' > t\} = \\ &= \{s't' \mid s't' > st\} = \mathbb{T}_{st} = f(st).\end{aligned}$$

  Falls $s < 0$ ist und $t \geq 0$ gilt, ist $st = -((-s)t)$ und aus dem bereits Bewiesenen folgt

$$f(st) = f(-((-s)t)) = -f((-s)t) = -(f(-s)f(t)) = -(-(f(s))f(t)) = f(s)f(t).$$

  Der letzte Fall $s < 0$, $t < 0$ ist einfacher:

$$f(st) = f((-s)(-t)) = f(-s)f(-t) = (-f(s))(-f(t)) = f(s)f(t).$$

- Zuletzt sei wieder $s \in S$ mit $s > 0$.

$$\begin{aligned}f(s)^{-1} = \mathbb{T}_s^{-1} &= \{s' \in \mathbb{Q} \mid \forall t \in \mathbb{T}_s : s' > \tfrac{1}{t} \wedge \forall t' \in \mathbb{Q} : (t' = \inf \mathbb{T}_s \Rightarrow s' \neq \tfrac{1}{t'})\} = \\ &= \{s' \in \mathbb{Q} \mid s' > \tfrac{1}{s}\} = \mathbb{T}_{s^{-1}} = f(s^{-1}).\end{aligned}$$

  Ist hingegen $s < 0$, dann erhalten wir

$$f(s^{-1}) = f(-((-s)^{-1})) = -f((-s)^{-1}) = -(f(-s)^{-1}) = (-f(-s))^{-1} = f(s)^{-1}.$$

Daher ist f ein Körperisomorphismus, und tatsächlich sind S und R isomorph. □

Ab nun bezeichnen wir den bis auf Isomorphie eindeutig bestimmten ordnungsvollständigen geordneten Körper R mit $\mathbb{R}$ und nennen ihn die **Menge der reellen Zahlen**.

6.5. Die komplexen Zahlen $\mathbb{C}$

Kommen wir nun zu einer Zahlenmenge, die in der Schule oft vernachlässigt wird, und die darüber hinaus einige philosophische Fragen aufzuwerfen scheint.

Wir erinnern uns, dass die Geschichte der Algebra mit Büchern begonnen hat, in denen unter anderem die Lösung linearer und quadratischer Gleichungen beschrieben wurde. Begonnen hat das in einer Zeit als nur die positiven rationalen Zahlen bekannt waren. Bereits die Lösung der quadratischen Gleichung

$$x^2 = 0 \tag{6.14}$$

bereitet da Schwierigkeiten. Gegen Ende 14. Jahrhunderts setzte sich in Europa die 0 als eigenständige Zahl durch, doch die alte Welt musste ein weiteres Jahrhundert warten bis auch die negativen Zahlen akzeptiert waren. Von diesem Zeitpunkt an war Gleichung (6.14) lösbar, ebenso wie etwa

$$x^2 + 3x + 2 = 0. \tag{6.15}$$

Die Tatsache, dass die rationalen Zahlen nicht genügen, ist seit Hippasos von Metapont (5. Jahrhundert v.Chr.) bekannt, der die Irrationalität von $\sqrt{2}$ als Diagonallänge des Einheitsquadrates erkannte. (Dafür wurde er übrigens, sagt die Geschichte, von einer Gruppe Pythagoreer im Meer ertränkt). Die Polynomgleichung

$$x^2 = 2, \tag{6.16}$$

die aus dem Satz von Pythagoras folgt, ist also — wie wir in Theorem 3.2.4 bewiesen haben — in den rationalen Zahlen nicht lösbar.

Daher wurde in der Mathematik schon früh auf die reellen Zahlen zurückgegriffen, allerdings ohne eine wirkliche Definition als Zahlenmenge anzugeben. Das haben erst Cantor und Dedekind im Jahre 1871 auf äquivalente aber unterschiedliche Weise getan.

Ist nun aber jede quadratische Gleichung lösbar? Die Antwort ist, wie wir alle wissen, nein. Die Gleichung

$$x^2 + 1 = 0 \tag{6.17}$$

hat keine reelle Lösung.

Die komplexen Zahlen wurden in der Mathematik schon einige Zeit verwendet, allerdings ohne richtige Definition. So ist bekannt, dass Girolamo Cardano (1501–1576) während er die Formeln für die Nullstellen von Polynomen dritten und vierten Grades erarbeitete, die komplexen Zahlen vor Augen hatte. Er verwarf sie allerdings wieder als „zu subtil und daher nutzlos“.

Auch Leonhard Euler (1707–1783) kannte bereits die komplexen Zahlen. Er führte 1748 die „Zahl“ i in seiner berühmten Arbeit „Introductio in analysin infinitorum“ als Bezeichnung ein. Dort taucht auch die faszinierende Formel

$$e^{ix} = \cos x + i \sin x,$$

die heute Eulers Namen trägt, das erste Mal auf.

Der erste jedoch, der eine mathematische Arbeit über die komplexen Zahlen verfasst hat, in der eine Definition derselben (die reellen Zahlen vorausgesetzt) vorkommt, war Caspar Wessel (1745–1818). Er hat 1799 in der Königlich Dänischen Akademie seine Arbeit veröffentlicht (übrigens als erstes Nichtmitglied, und es war seine einzige(!) mathematische Arbeit), in der er die geometrische Interpretation der komplexen Zahlen vorstellte. Er entwickelte diese Zahlen übrigens während er Oldenburg trigonometrisch vermaß (triangulierte), und es ist sicher, dass er bereits 1787 die komplexen Zahlen entwickelt hatte (unwissend, dass solche Zahlen bereits in Verwendung waren). Mit Hilfe dieser brillianten mathematischen Idee gelang es ihm als erstem, eine genaue Landkarte Dänemarks herzustellen.

Leider wurde seine Arbeit in Mathematikerkreisen nicht gelesen, und so wurde im Jahr 1806 die geometrische Interpretation von dem Schweizer Jean Robert Argand (1768–1822) wiederentdeckt und erneut neuentwickelt von Johann Carl Friedrich Gauss (1777–1855) im Jahre 1831, der übrigens interessanterweise eine weitere Arbeit von Wessel, nämlich die Triangulierung von Oldenburg im Jahr 1824 wiederholte.

Was sind also diese „mystischen“ komplexen Zahlen, die die Mathematiker so lange in Atem gehalten haben? Als moderne Mathematiker mit geschultem algebraischem Blick können wir den Zahlen den Mythos nehmen. Wir beginnen mit einer Definition.

Definition 6.5.1 ($\mathbb{C}$)**.** *Wir definieren* $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ *und erklären auf dieser Menge die beiden Verknüpfungen* $+$ *und* $\cdot$ *wie folgt*

$$\begin{aligned}(a_1, a_2) + (b_1, b_2) &:= (a_1 + b_1, a_2 + b_2)\\ (a_1, a_2) \cdot (b_1, b_2) &:= (a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1).\end{aligned}$$

Wir nennen $(\mathbb{C}, +, \cdot)$ *die komplexen Zahlen.*

Zuerst untersuchen wir die algebraischen Eigenschaften unseres neuen Konstrukts.

Theorem 6.5.2. $(\mathbb{C}, +, \cdot)$ *ist ein Körper.*

Beweis. Um dieses Theorem zu beweisen, müssen wir die Körperaxiome nachrechnen.

AG(+): Das folgt aus der komponentenweisen Definition von $+$ und der Tatsache, dass $(\mathbb{R}, +)$ eine abelsche Gruppe ist.
KG(+): Hier trifft dasselbe Argument zu wie für das Assoziativgesetz.
Nullelement: Es gilt, dass $(0, 0)$ das neutrale Element bezüglich $+$ ist. $(a_1, a_2) + (0, 0) = (a_1 + 0, a_2 + 0) = (a_1, a_2)$.
Inverse(+): Das Inverse zu (a_1, a_2) ist $(-a_1, -a_2)$, wie man sehr leicht nachrechnet.

AG($\cdot$): Seien (a_1, a_2), (b_1, b_2) und (c_1, c_2) gegeben. Dann gilt

$$\begin{aligned}
(a_1,a_2)\big((b_1,b_2)(c_1,c_2)\big) &= \\
&= (a_1,a_2)(b_1c_1 - b_2c_2, b_1c_2 + b_2c_1) \\
&= (a_1b_1c_1 - a_1b_2c_2 - a_2b_1c_2 - a_2b_2c_1, a_1b_1c_2 + a_1b_2c_1 + a_2b_1c_1 - a_2b_2c_2) \\
&= (a_1b_1 - a_2b_2, a_1b_2 + a_2b_1)(c_1,c_2) \\
&= \big((a_1,a_2)(b_1,b_2)\big)(c_1,c_2).
\end{aligned}$$

KG($\cdot$): Dieses Gesetz folgt aus der Symmetrie der Definition von $\cdot$ und dem Kommutativgesetz in $(\mathbb{R}, \cdot)$.

Einselement: Das Einselement ist $(1, 0)$, eine sehr einfache Rechnung.

Inverse($\cdot$): Ist $(a_1, a_2) \neq (0, 0)$, dann ist das Element

$$\left(\frac{a_1}{a_1^2 + a_2^2}, \frac{-a_2}{a_1^2 + a_2^2}\right)$$

das Inverse zu (a_1, a_2). Beachte, dass für reelle Zahlen a_1 und a_2 der Nenner $a_1^2 + a_2^2$ nur dann verschwinden kann, wenn beide Zahlen gleich 0 sind. Das haben wir aber ausgeschlossen. Es gilt

$$(a_1,a_2)\left(\frac{a_1}{a_1^2 + a_2^2}, \frac{-a_2}{a_1^2 + a_2^2}\right) = \left(\frac{a_1^2 + a_2^2}{a_1^2 + a_2^2}, \frac{-a_1a_2 + a_2a_1}{a_1^2 + a_2^2}\right) = (1,0).$$

Distributivgesetz: Seien wiederum (a_1, a_2), (b_1, b_2) und (c_1, c_2) gegeben. Dann gilt wegen des Distributivgesetzes in $\mathbb{R}$

$$\begin{aligned}
(a_1,a_2)\big((b_1,b_2) + (c_1,c_2)\big) &= \\
&= \big(a_1(b_1+c_1) - a_2(b_2+c_2), a_1(b_2+c_2) + a_2(b_1+c_1)\big) \\
&= (a_1b_1 - a_2b_2, a_1b_2 + a_2b_1) + (a_1c_1 - a_2c_2, a_1c_2 + a_2c_1) = \\
&= (a_1,a_2)(b_1,b_2) + (a_1,a_2)(c_1,c_2).
\end{aligned}$$

$\square$

Die reellen Zahlen sind ein Unterkörper von $\mathbb{C}$, wie man sieht, indem man die Abbildung

$$\begin{aligned}
\iota : \mathbb{R} &\to \mathbb{C} \\
r &\mapsto (r, 0)
\end{aligned}$$

betrachtet. Diese Abbildung erlaubt es uns, $\mathbb{R}$ derart als Teilmenge von $\mathbb{C}$ aufzufassen, dass die von $\mathbb{C}$ ererbten Operationen mit den ursprünglichen $\mathbb{R}$-Operationen übereinstimmen. Bleibt also nur, die Voraussetzungen von Proposition 5.4.10 nachprüfen. Tatsächlich genügen einfache Rechnungen, $(r, 0) + (s, 0) = (r + s, 0)$ und $(r, 0)(s, 0) = (rs, 0)$ nachzuweisen und weiters $-(r, 0) = (-r, 0)$ sowie $(r, 0)^{-1} = (\frac{1}{r}, 0)$ zu zeigen. In Zukunft werden wir also die reellen Zahlen mit den komplexen Elementen $(r, 0)$ identifizieren und im weiteren wieder r für diese Zahlen schreiben. Außerdem sehen wir, dass $(r, 0)(a_1, a_2) = (ra_1, ra_2)$ gilt.

Interessant wird es nun, wenn wir die Eigenschaften anderer Elemente betrachten, z.B.

$$(0,1)(0,1) = (-1,0),$$

und damit finden wir in $\mathbb{C}$ eine Nullstelle des Polynoms $x^2 + 1$. Um die Schreibweise zu vereinfachen, führen wir eine Abkürzung für $(0, 1)$ ein, indem wir sagen

Definition 6.5.3 (Imaginäre Einheit). *Es gelte die Bezeichnung*

$$i := (0,1).$$

Wir nennen i die imaginäre Einheit*.*

Wir haben schon nachgerechnet, dass $i^2 = -1$ gilt, und es folgt aus der Struktur von $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ und der komponentenweisen Definition der Addition, dass sich jedes Element (a_1, a_2) von $\mathbb{C}$ eindeutig schreiben lässt als $(a_1, a_2) = (a_1, 0) + a_2(0, 1)$ oder mit Hilfe der Abkürzungen aus Definition 6.5.3 als $(a_1, a_2) = a_1 + ia_2$.

Damit gewinnen wir Eulers Schreibweise für die komplexen Zahlen zurück. Mythisches oder Philosophisches haben wir dazu nicht benötigt. Wir fassen unsere neue Notation zusammen und definieren.

Definition 6.5.4 (Real- und Imaginärteil)**.** *Sei $z = (x, y) \in \mathbb{C}$. Dann schreiben wir z auch als*

$$z = x + iy$$

und bezeichnen x als den Realteil *und y als den* Imaginärteil *von z und schreiben $x = \operatorname{Re} z = \Re z$ bzw. $y = \operatorname{Im} z = \Im z$.*

Die komplexen Zahlen lassen sich als Elemente von $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ klarerweise auch als Punkte in der Ebene deuten. Das führt auf die Definition von Wessel, Argand und Gauss. Auch die Polarkoordinatenrepräsentation durch Länge und Winkel ist auf diese geometrische Interpretation zurückzuführen, siehe Abbildung 6.1:

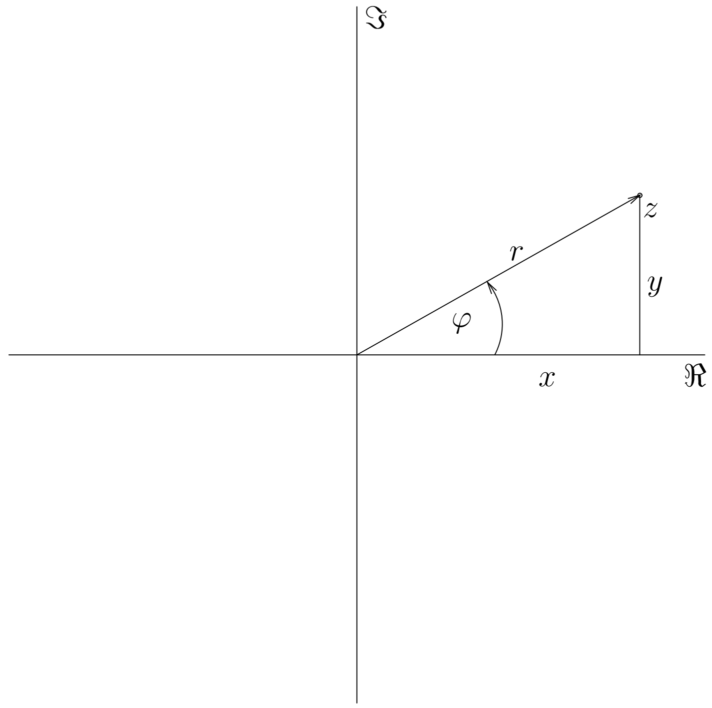


Abbildung 6.1. Die komplexe Zahlenebene

Bemerkung 6.5.5 (Polardarstellung komplexer Zahlen)**.** *Führen wir in der komplexen Ebene* Polarkoordinaten *ein, so läßt sich jede komplexe Zahl $z = x + iy$ als*

$$z = r(\cos\varphi + i\sin\varphi)$$

schreiben (siehe Abbildung 6.1). Hier ist der Radius *r durch den Betrag $|z|$ von z gegeben, d.h. $r = |z| := \sqrt{x^2 + y^2}$. Der Winkel φ wird auch* Argument *von z genannt und es gilt* $\tan\varphi = \frac{y}{x}$.

Die Darstellung von z in Polarschreibweise ist in bezug auf r eindeutig. Der Winkel φ ist zu gegebenem $z \neq 0$ allerdings nur bis auf Addition von ganzzahligen Vielfachen von 2π bestimmt. Ist $z = 0$, dann ist der Winkel gänzlich unbestimmt; $z = 0(\cos\varphi + i\sin\varphi)$ für beliebiges φ.

Das Multiplizieren komplexer Zahlen ist in der Polardarstellung besonders einfach. Es gilt nämlich

$$z_1 z_2 = r_1 r_2\big(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)\big),$$

wobei r_i bzw. φ_i der zu z_i $(i = 1, 2)$ gehörige Radius bzw. Winkel ist. Es werden also die Radien multipliziert und die Winkel addiert. Das Inverse von $z \neq 0$ ist ebenfalls sehr einfach zu bestimmen; es gilt $z^{-1} = \frac{1}{r}(\cos(-\varphi) + i\sin(-\varphi))$.

In der cartesischen Darstellung ist die Division ein wenig mühsamer:

$$\frac{a_1 + ib_2}{a_2 + ib_2} = \frac{(a_1 + ib_1)(a_2 - ib_2)}{(a_2 + ib_2)(a_2 - ib_2)} = \frac{a_1a_2 + b_1b_2 + i(a_2b_1 - a_1b_2)}{a_2^2 + b_2^2} = \frac{a_1a_2 + b_1b_2}{a_2^2 + b_2^2} + i\frac{a_2b_1 - a_1b_2}{a_2^2 + b_2^2}.$$

In diesem Fall haben wir den Bruch oben und unten mit derselben komplexen Zahl multipliziert, nämlich $a_2 - ib_2$. Diese verdient eine besondere Hervorhebung.

Definition 6.5.6 (Konjugiert komplexe Zahl)**.** *Sei* $z = x + iy \in \mathbb{C}$ *Wir definieren die zu* z konjugiert komplexen Zahl *als*

$$\overline{z} := x - iy.$$

Wie wir im Nenner obiger Rechnung sehen können, in dem das Quadrat des Betrags von $a_2 + ib_2$ auftaucht, gilt für jede komplexe Zahl z

$$z\overline{z} = |z|^2,$$

und außerdem

$$\begin{aligned} \overline{\overline{z}} &= z, \\ \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2}, \\ \overline{z_1 z_2} &= \overline{z_1}\,\overline{z_2}. \end{aligned}$$

Wir wollen nun untersuchen, ob es uns gelingt, die Ordnungsrelation von $\mathbb{R}$ auf $\mathbb{C}$ auszudehnen, sodass wieder O1 und O2 gelten. Folgendes (zunächst) erstaunliche Resultat kommt dabei zu Tage.

Theorem 6.5.7. *Es gibt keine Ordnungsrelation auf* $\mathbb{C}$*, mit der* $(\mathbb{C}, +, \cdot)$ *ein geordneter Körper wird.*

Beweis. Angenommen, es gäbe eine Ordnungsrelation $\leq$, die alle notwendigen Eigenschaften aufweist. Dann gilt jedenfalls $-1 < 0 < 1$ wegen Proposition 6.3.2.(2) und (5).

Wegen $i \neq 0$ folgt aber wieder wegen Proposition 6.3.2.(5), dass $-1 = i^2 > 0$, ein Widerspruch. Daher existiert keine solche Ordnungsrelation. □

Nun aber zurück zu den Polynomen. Für ein beliebiges quadratisches Polynom mit komplexen Koeffizienten α_i können wir jetzt jedenfalls die Nullstellen ausrechnen. Sei nämlich

$$p(z) = \alpha_2 z^2 + \alpha_1 z + \alpha_0,$$

dann kann man alle Nullstellen von p mit Hilfe der wohlbekannten Formel

$$z_{1,2} = \frac{-\alpha_1 \pm \sqrt{\alpha_1^2 - 4\alpha_2\alpha_0}}{2\alpha_2}$$

berechnen.

Beispiel 6.5.8. *Sei das Polynom*

$$z^2 - (3 - 8i)z - 13 - 11i$$

gegeben. Die Nullstellen bestimmen wir wie folgt:

$$\begin{aligned} z_{1,2} &= \frac{3-8i \pm \sqrt{(3-8i)^2 + 4(13+11i)}}{2} \\ &= \frac{3-8i \pm \sqrt{-55-48i+52+44i}}{2} \\ &= \frac{3-8i \pm \sqrt{-3-4i}}{2} \end{aligned}$$

Wir müssen nun die Wurzel aus $-3-4i$ ziehen, wofür sich zwei Möglichkeiten anbieten. Zum einen können wir in Polarkoordinaten verwandeln und $\sqrt{(r,\varphi)} = (\sqrt{r}, \frac{\varphi}{2})$ verwenden. Zum anderen ist es möglich, die Wurzel direkt zu ziehen. Dazu verwenden wir einen unbestimmten Ansatz. *Sei $\sqrt{-3-4i} = a + ib$. Dann gilt*

$$\begin{aligned} (a+ib)^2 &= -3-4i \\ a^2 - b^2 + 2iab &= -3-4i. \end{aligned}$$

Das führt zu dem Gleichungssystem

$$\begin{aligned} a^2 - b^2 &= -3 \\ 2ab &= -4. \end{aligned}$$

Mit einem kleinen Trick können wir den Lösungsweg abkürzen. Wir wissen, dass

$$a^2 + b^2 = |a+ib|^2 = |(a+ib)^2| = |-3-4i| = \sqrt{9+16} = 5$$

gilt, und aus dieser erhalten wir durch Addition bzw. Subtraktion mit der oberen Gleichung:

$$\begin{aligned} 2a^2 &= 2 \\ 2b^2 &= 8. \end{aligned}$$

Wir haben also $a = \pm 1$ und $b = \pm 2$, und aus der Beziehung $2ab = -4$ erhalten wir die Lösungen

$$\sqrt{-3-4i} = \pm(1-2i).$$

Setzen wir das in die Lösungsformel ein, dann erhalten wir

$$\begin{aligned} z_{1,2} &= \frac{3-8i \pm (1-2i)}{2} \\ z_1 &= 2-5i \\ z_2 &= 1-3i. \end{aligned}$$

Für quadratische Polynome haben die komplexen Zahlen also das Nullstellenproblem erledigt, doch wir wissen noch immer nicht, ob wir dasselbe für beliebige Polynome tun können. Die Frage ist, ob jedes nichtkonstante komplexe Polynom eine Nullstelle besitzt, ob also $\mathbb{C}$ **algebraisch abgeschlossen** ist. Dieses Problem hat J.C.F. Gauss 1799 in seiner Dissertation gelöst:

Theorem 6.5.9 (Fundamentalsatz der Algebra)**.** *Sei $p(z)$ ein beliebiges nichtkonstantes Polynom mit komplexen Koeffizienten:*

$$p(z) = \sum_{i=0}^{n} a_i z^i \quad \text{mit } a_i \in \mathbb{C},\ n \geq 1,\ a_n \neq 0.$$

Dann existiert ein $\alpha \in \mathbb{C}$ mit $p(\alpha) = 0$, es gibt also immer wenigstens eine (komplexe) Nullstelle.

Beweis. Der Beweis dieses Satzes würde das Lehrziel dieses Skriptums sprengen, und daher lassen wir ihn aus. Übrigens gibt es viele — über 100 — verschiedenen Beweise für diesen Satz. In jedem guten Buch über Funktionentheorie (komplexe Analysis) ist einer zu finden; siehe etwa [**Remmert, Schumacher 2001**]. □

Es lässt sich sogar noch ein klein wenig mehr sagen, denn wenn man eine Nullstelle eines Polynoms gefunden hat, dann kann man mit Hilfe der Polynomdivision folgenden Satz beweisen:

Theorem 6.5.10. *Sei p ein komplexes Polynom n–ten Grades und α eine Nullstelle von p. Dann gibt es ein Polynom q vom Grad $n-1$, und es gilt*

$$p(z) = q(z)(z-\alpha).$$

Man kann also den Linearfaktor $z-\alpha$ *abspalten.*

Beweis. Ebenfalls in guten Funktionentheorie-Büchern nachzulesen. □

Fasst man die beiden Theoreme 6.5.9 und 6.5.10 zusammen, dann kann man die wichtige Folgerung über Polynome und ihre Nullstellen beweisen:

Korollar 6.5.11. *Sei p ein komplexes Polynom vom Grad n. Dann existieren genau n Linearfaktoren $z-\alpha_i$ mit $i=1,\dots,n$, sodass*

$$p(z) = \prod_{i=1}^{n}(z-\alpha_i).$$

Das Polynom zerfällt also über $\mathbb{C}$ in genau n Linearfaktoren.

Beweis. Nach dem Fundamentalsatz der Algebra hat p eine Nullstelle, die man nach Theorem 6.5.10 abspalten kann. übrig bleibt ein Polynom q, dessen Grad um 1 kleiner ist als der von p. Auf q kann man wieder den Fundamentalsatz anwenden, usw. Das Korollar folgt mittels vollständiger Induktion. □

Das ist sehr praktisch, doch leider gibt es keine Möglichkeit, für allgemeine Polynome hohen Grades diese Linearfaktoren (d.h. die Nullstellen) zu bestimmen. Nils Henrik Abel (1802–1829) hat nämlich im Jahr 1824 den folgenden Satz bewiesen:

Theorem 6.5.12 (Abel)**.** *Für jedes $n \geq 5$ existiert ein Polynom p mit rationalen Koeffizienten vom Grad n, das eine reelle Nullstelle r besitzt mit der Eigenschaft, dass r nicht geschrieben werden kann als algebraischer Ausdruck, der rationale Zahlen, Additionen, Subtraktionen, Multiplikationen, Divisionen und k–te Wurzeln enthält. Anders ausgedrückt existiert keine Formel und damit kein endlicher algebraischer Algorithmus, der aus den Koeffizienten eines Polynoms vom Grad $n \geq 5$ die Nullstellen berechnet.*

Beweis. Der Beweis dieses Satzes gehört in die höhere Algebra und kann unter dem Kapitel Galoistheorie z.B. in [**Scheja, Storch 1988**] nachgelesen werden. □

6.6. Die Quaternionen $\mathbb{H}$

Eine letzte interessante Frage kann man noch über Zahlenmengen stellen, die sich direkt aus der Definition von $\mathbb{C}$ als Körperstruktur auf $\mathbb{R}\times\mathbb{R}$ ergibt. Kann man z.B. auf $\mathbb{R}\times\mathbb{R}\times\mathbb{R} = \mathbb{R}^3$ auch eine Körperstruktur einführen?

Die Suche nach der Antwort auf diese Frage hat auch den Mathematiker Sir William Rowan Hamilton (1805–1865), einen der bedeutendsten Wissenschafter seiner Epoche beschäftigt, und im Jahr 1843 präsentierte er schließlich die Arbeit „On a new Species of Imaginary Quantities connected with a theory of Quaternions“ bei einem Treffen der Royal Irish Academy.

Doch Hamilton hatte es nicht geschafft, auf $\mathbb{R}^3$ eine Körperstruktur einzuführen. Er hatte zwar keine Probleme gehabt, auf jedem $\mathbb{R}^n$ durch komponentenweise Definition eine Addition zu erklären, die eine abelsche Gruppe ergab, doch die Multiplikation hatte nicht gelingen wollen. Was er dann zusammengebracht hat, war eine algebraische Struktur im $\mathbb{C}^2 = \mathbb{R}^4$ zu definieren.

Sei $\mathbb{H} = \mathbb{C} \times \mathbb{C}$ gegeben. Wir definieren Verknüpfungen auf $\mathbb{H}$ durch

$$\begin{aligned}(z_0, z_1) + (w_0, w_1) &:= (z_0 + w_0, z_1 + w_1)\\ (z_0, z_1)(w_0, w_1) &:= (z_0w_0 - z_1\overline{w_1}, z_0w_1 + z_1\overline{w_0}).\end{aligned}$$

Wir können die algebraischen Eigenschaften von $(\mathbb{H}, +, \cdot)$ untersuchen. Wegen der komponentenweisen Definition der Addition folgt sofort, dass $(\mathbb{H}, +)$ eine abelsche Gruppe ist.

Die Multiplikation ist assoziativ:

$$\begin{aligned}\big((z_0, z_1)(w_0, w_1)\big)(t_0, t_1) &= (z_0w_0 - z_1\overline{w_1}, z_0w_1 + z_1\overline{w_0})(t_0, t_1) =\\ &= (z_0w_0t_0 - z_1\overline{w_1}t_0 - z_0w_1\overline{t_1} - z_1\overline{w_0}\overline{t_1}, z_0w_0t_1 - z_1\overline{w_1}t_1 + z_0w_1\overline{t_0} + z_1\overline{w_0}\overline{t_0}) =\\ &= (z_0, z_1)(w_0t_0 - w_1\overline{t_1}, w_0t_1 + w_1\overline{t_0}) =\\ &= (z_0, z_1)\big((w_0, w_1)(t_0, t_1)\big),\end{aligned}$$

das Element $(1, 0)$ ist das Einselement bezüglich der Multiplikation (das ist leicht), und jedes Element verschieden von $0 = (0, 0)$ besitzt ein Inverses:

$$(z_0, z_1)^{-1} = \left(\frac{z_0}{|z_0|^2 + |z_1|^2}, \frac{-z_1}{|z_0|^2 + |z_1|^2}\right).$$

Beidseitig gelten die Distributivgesetze, doch das Kommutativgesetz bezüglich der Multiplikation ist **nicht** erfüllt. Eine algebraische Struktur dieser Art nennt man **Schiefkörper**.

Die Quaternionen der Form $(z, 0)$ bilden einen Körper, der isomorph zu $\mathbb{C}$ ist, und daher werden wir diese Elemente in Zukunft auch mit den komplexen Zahlen identifizieren und wieder z schreiben.

Wenn wir spezielle Elemente betrachten, erhalten wir erstaunliche Ergebnisse:

$$\begin{aligned}(0, 1)(0, 1) &= (-1, 0)\\ (0, i)(0, i) &= (-1, 0).\end{aligned}$$

Die Quaternionen enthalten also noch zwei „Wurzeln“ von -1. Wir schreiben $j := (0, 1)$ und $k := (0, i)$ und erhalten so die Rechenregeln

$$\begin{array}{ccc} i^2 = -1, & j^2 = -1, & k^2 = -1,\\ ij = k, & jk = i, & ki = j,\\ ji = -k, & kj = -i, & ik = -j. \end{array}$$

Aus der Definition der Quaternionen lässt sich leicht zeigen, dass man jedes $q \in \mathbb{H}$ eindeutig schreiben kann als $z_0 + z_1j$ (Achtung auf die Reihenfolge!) mit komplexen Koeffizienten z_0 und z_1 oder als $q = x_0 + a_1i + a_2j + a_3k$ mit reellen Koeffizienten a_i.

Der Betrag einer Quaternione ist

$$|(z_0, z_1)| = \sqrt{|z_0|^2 + |z_1|^2},$$

und die konjugierte Quaternione ist

$$\overline{(z_0, z_1)} = (z_0, -z_1).$$

Es gilt analog zu den komplexen Zahlen $|q|^2 = q\overline{q}$.

Interessant ist vielleicht noch eine weitere Darstellung der Quaternionen als Paar (a, A) mit eine reellen Zahl a und einem Vektor $A \in \mathbb{R}^3$. In diesem Fall lassen sich die Operationen hinschreiben als

$$\begin{aligned}(a, A) + (b, B) &= (a + b, A + B)\\ (a, A)(b, B) &= (ab - AB, aB + Ab + A \times B),\end{aligned}$$

also fast so wie die Operationen in $\mathbb{C}$, bis auf den Term $A \times B$ in der Definition der Multiplikation. An dieser Definition kann man auch schon erahnen, dass Quaternionen etwas mit Drehungen zu tun haben.

Die Frage, die sich die Mathematiker jetzt stellten war, ob niemand klug genug war, die richtige Definition einer Multiplikation zu finden, oder ob die Schwierigkeiten einen mathematischen Grund haben.

Arthur Cayley (1821–1895) hat versucht, die Methode noch einmal anzuwenden und auf $\mathbb{H} \times \mathbb{H} \cong \mathbb{R}^8$ eine Multiplikation einzuführen. Es gelang ihm, die Cayley-Zahlen oder Oktaven oder Okternionen $\mathbb{O}$ zu definieren, doch deren algebraische Eigenschaften lassen doch deutlich mehr zu wünschen übrig als die der Quaternionen. Okternionen sind nicht einmal mehr ein Schiefkörper. Es besitzt zwar jedes Element ein eindeutiges Inverses, doch die Multiplikation ist nicht assoziativ! Solch eine algebraische Struktur, in der über einer abelschen Gruppe eine Multiplikation definiert wird, die die Distributivität erfüllt und wo Einselement und Inverse existieren, heißt (nichtassoziative) **Divisionsalgebra**.

Ein tiefer Satz aus der Differentialgeometrie besagt nun, dass Divisionsalgebren über $\mathbb{R}^n$ nur in den Dimensionen 1, 2, 4 und 8 existieren, und in jeder dieser Dimensionen genau eine, nämlich $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ und $\mathbb{O}$. Es war also nicht Unfähigkeit, die die Mathematiker des 19. Jahrhunderts daran gehindert hat, über allen $\mathbb{R}^n$ eine Körperstruktur zu finden, sondern sie haben nach nicht Existentem gestrebt.

Damit beenden wir unseren Ausflug in die Welt der Zahlen. Beginnend von $\mathbb{N}$, der Klasse von Zahlen, deren Geschichte ihren Ursprung bereits in grauer Vorzeit hat, haben wir sie basierend auf exakten mathematischen Grundlagen neu entdeckt. Weiter ging es über die ganzen zu den rationalen Zahlen und darauf aufbauend zu den reellen Zahlen, der Grundlage der Analysis. Auf der Suche nach den Nullstellen der Polynome sind wir zu den komplexen Zahlen und Gauss' fundamentalem Theorem gelangt. Selbst die Frage, ob wir damit schon alle Zahlensysteme gefunden haben, die $\mathbb{R}$ als Unterkörper haben und in denen man dividieren kann, haben wir untersucht. Wir haben alle dieser Strukturen gefunden, und daher ist es jetzt an der Zeit, Neuland zu entdecken und zu erkunden, was Generationen von Mathematikern aus den hier präsentierten Prinzipien geschaffen haben.

Wir hoffen, dass Ihnen dieser Streifzug durch die Grundlagen der modernen Mathematik ein wenig Freude bereitet hat und wünschen allen LeserInnen viel Vergnügen mit all den Theorien, Strukturen und Anwendungen, die im Verlauf ihres Studiums noch kommen mögen.

Literaturverzeichnis

[Behrends 2003] Behrends, E., *Analysis, Band 1*, Ein Lehrbuch für den sanften Wechsel von der Schule zur Uni, Vieweg, Braunschweig/Wiesbaden, 2003.

[Beutelspacher 1999] Beutelspacher, A., *Das ist o.B.d.A. trivial*, Tips und Tricks zur Formulierung mathematischer Gedanken, Vieweg, Braunschweig/Wiesbaden, 1999.

[Bishop 1967] Bishop, E., *Foundations of constructive analysis*, McGraw–Hill, New York, 1967.

[Bronstein et al. 1989] Bronstein, I.N.; Semendjajew, K.A., *Taschenbuch der Mathematik*, Verlag Harri Deutsch, Thun, 1989.

[Cigler, Reichel 1987] Cigler, J.; Reichel, H.C., *Topologie*, B.I. Hochschultaschenbücher, Mannheim/Wien/Zürich, 1987.

[Heuser 1986] Heuser, H., *Lehrbuch der Analysis, Teil 1*, B.G. Teubner, Stuttgart, 1986.

[O'Connor, Robertson 1996] O'Connor, J.J; Robertson, E.F., *A history of set theory*, `http://www-groups.dcs.st-and.ac.uk/history/HistTopics/Beginnings_of_set_theory.html`, 1996.

[Remmert, Schumacher 2001] Remmert, R.; Schumacher, *Funktionentheorie 1*, Springer Verlag, 2001.

[Scheja, Storch 1988] Scheja, G.; Storch, U., *Lehrbuch der Algebra*, Teubner Verlag, Stuttgart, 1988.