GRUNDBEGRIFFE DER MATHEMATIK

Vorlesungsskript, Wintersemester 2010/2011

Christian Clason

Stand vom 29. Januar 2011

Institut für Mathematik und Wissenschaftliches Rechnen
Karl-Franzens-Universität Graz

INHALTSVERZEICHNIS

Teil I

MATHEMATISCHE GRUNDBEGRIFFE

ÜBERBLICK

Ein Grund für den Erfolg der modernen Mathematik ist ihre Abstraktheit – sie erlaubt es, gemeinsame Strukturen in den unterschiedlichsten Bereichen zu erkennen, und sinnvolle Aussagen über diese zu machen. Dabei darf dann natürlich nicht mehr auf die (zum Beispiel physikalischen) Hintergründe zurückgegriffen werden, da ansonsten die Anwendbarkeit in anderen Zusammenhängen (etwa der Wirtschaft) nicht mehr gewährleistet ist. Um also Erkenntnisse und Begründungen einsichtig und nachprüfbar zu formulieren, ist eine eigene, präzise, Fachsprache notwendig.

Wie Juristen klar zwischen „Besitz“ und „Eigentum“ unterscheiden, haben auch in der Mathematik scheinbar intuitiv verständliche umgangssprachliche Begriffe eine präzise technische Bedeutung. Und wie ein Schuldspruch ohne rigorose Urteilsbegründung inakzeptabel ist, sind auch an mathematische Begründungen strenge formale Anforderung gestellt. Dies stellt erfahrungsgemäß eine der größten Hürden für Einsteiger dar.

Wohlgemerkt: Es geht hier nicht um eine rigorose Fundierung (geschweige denn Philosophie) der Mathematik, die einen Formalismus erfordert, der den Rahmen dieser Veranstaltung sprengen würde (und eigentlich erst nach einigen Semestern Mathematikstudium sinnvoll ist, wenn man erkennen kann, *was* eigentlich begründet werden soll).[1] In diesen Kapiteln wollen wir vielmehr einen Überblick über Grammatik und Vokabular der Fachsprache Mathematik geben, damit Sie diese nicht gleichzeitig mit dem Stoff der ersten „richtigen“ Mathematik-Vorlesungen erlernen müssen.

Wie beim Erlernen jeder Sprache gilt auch hier: Erst Übung macht den Meister!

[1] Dies geschieht üblicherweise in einer Vorlesung *Grundlagen der Mathematik*. Für den interessierten Leser werden aber als Fußnoten Hinweise gegeben, wie solch eine Formalisierung aussehen kann.

ELEMENTARE LOGIK

1

Die Grammatik der Mathematik ist die *Logik*. Sie gibt die Regeln vor, wie aus als richtig erkannten Aussagen neue Aussagen abgeleitet werden können, deren Richtigkeit dadurch festgelegt ist. Ziel dieses Kapitels ist es, für die naive (das heißt informelle) Logik eine präzise Schreibweise zu formulieren, die für den abstrakten Rahmen der Mathematik tauglich ist. Insbesondere soll diese Logik nach Möglichkeit den intuitiven, alltäglichen Gebrauch wiedergeben.

1.1 AUSSAGEN

Das fundamentale Element der mathematischen Sprache ist die *Aussage*. Eine Aussage ist ein (umgangssprachlicher oder formelhafter) Satz, der entweder *wahr* oder *falsch* ist. Auch wenn wir nicht wissen, welches von beiden gilt, muss erkennbar sein, dass eine und nur eine der beiden Möglichkeiten zutreffen kann. Wenn die Aussage p wahr ist, sagt man auch: „p gilt“ oder „es gilt, dass p“.

Beispiel 1.1.

- „Graz ist eine Stadt.“ ist eine Aussage, deren *Wahrheitswert* „wahr“ ist.
- „Frankreich ist die Hauptstadt von Europa.“ ist eine Aussage mit dem Wahrheitswert „falsch“.
- „7+3“ ist keine Aussage.
- „7+3=4“ ist eine Aussage.
- „7+x=4“ ist dagegen keine Aussage, da der Wahrheitswert nicht feststehen kann, solange man nicht festlegt, was „x“ sein soll.[1]

[1] Um mit einem weit verbreiteten Missverständnis aufzuräumen: Mathematik besteht nicht aus Formeln, sondern aus Aussagen! (Welche gelegentlich Aussagen über Formeln sind.)

1.2 VERKNÜPFUNGEN VON AUSSAGEN

Neue Aussagen lassen sich bilden, indem zwei Aussagen verknüpft werden. Dazu legt man den Wahrheitswert der zusammengesetzten Aussage in Abhängigkeit vom Wahrheitswert der zu verknüpfenden Aussagen fest. Eine Möglichkeit ist die *Wahrheitstafel*: Wenn wir die beiden Aussagen p und q zu einer neuen Aussage r verknüpfen wollen, schreiben wir zum Beispiel (mit W für „wahr“ und F für „falsch“):

p	q	r
W	W	W
W	F	F
F	W	F
F	F	F

Damit ist für alle möglichen Belegungen von p und q der Wahrheitswert von r festgelegt: r ist wahr, wenn p und q beide wahr sind, und sonst falsch. Diese Verknüpfung nennt man *Konjunktion* oder *logisches Und*, und schreibt statt r üblicherweise $p \wedge q$, gesprochen „p und q“. Damit können wir ausdrücken, ob von zwei Aussagen beide wahr sind.

Beispiel 1.2. Wir betrachten die Aussagen p: „4 ist eine gerade Zahl“, deren Wahrheitswert wir als „wahr“ festlegen, und q: „4 ist durch 3 teilbar“, deren Wahrheitswert „falsch“ sein soll. Dann ist $p \wedge q$ die Aussage „4 ist eine gerade Zahl und durch 3 teilbar“, deren Wahrheitswert „falsch“ ist.

Wollen wir dagegen sagen, dass *mindestens eine* von zwei Aussagen p und q wahr sind, verwenden wir das *logische Oder* (auch *Disjunktion*) $p \vee q$, gesprochen „p oder q“:

p	q	$p \vee q$
W	W	W
W	F	W
F	W	W
F	F	F

Hier ist bereits der erste Unterschied zwischen der mathematischen Fachsprache und der Umgangssprache: das logische Oder ist nicht exklusiv, es drückt also *nicht* aus, dass von zwei Aussagen nur eine von beiden zutrifft!

Beispiel 1.3. Wir betrachten wieder die Aussagen p und q aus Beispiel 1.2. Dann ist $p \vee q$ die Aussage „4 ist gerade oder durch 3 teilbar“, deren Wahrheitswert „wahr“ ist. Ebenso gilt „6 ist gerade oder durch 3 teilbar“.

Aus einer Aussage p lässt sich auch eine neue Aussage bilden, indem man den Wahrheitswert von p umkehrt: Wir definieren die *Negation* $\neg p$ durch:

p	$\neg p$
W	F
F	*W*

Die Aussage $\neg p$ („nicht p“) ist also wahr, wenn p falsch ist, und umgekehrt.

Beispiel 1.4. Wir betrachten die Aussage q aus Beispiel 1.2. Dann ist $\neg q$ die Aussage „4 ist nicht durch 3 teilbar“, deren Wahrheitswert „wahr“ ist.

IMPLIKATION. Viele (insbesondere mathematische) Aussagen haben die Form „wenn p, dann q“. Wie können wir solche Aussagen mit unseren Mitteln ausdrücken? Dafür müssen wir uns erst darüber im klaren sein, was es heissen soll, dass die Aussage „wenn p, dann q“ wahr ist. Betrachten wir das Beispiel „Wenn es heute Abend regnet, gehe ich ins Kino“. Wenn wir sagen, dass diese Aussage wahr ist, behaupten wir weder, dass es heute Abend regnet, noch, dass der Sprecher ins Kino geht. Es ist lediglich gemeint, das *falls* ein Sachverhalt eintritt (es heute Abend regnet), ein anderer Sachverhalt eintritt (nämlich der Sprecher ins Kino geht). Das einzige, was ausgeschlossen wird, ist, dass es regnet und der Sprecher nicht ins Kino geht. Die Wahrheit von „Wenn es heute Abend regnet, gehe ich ins Kino“ sagt nichts aus über den Fall, wenn es nicht regnet – der Sprecher kann insbesondere trotzdem ins Kino gehen. In der umgangssprachlichen Verwendung können Sie diesen Fall offen lassen oder durch Vorwissen einschränken (wenn es nicht regnet, ist Biergartenwetter). In der Mathematik müssen wir aber für jede zusammengesetzte Aussage alleine anhand des Wahrheitswerts der Einzelaussagen (hier: „Es regnet heute Abend“ und „Ich gehe ins Kino“) entscheiden können, ob sie wahr oder falsch ist – sonst würde dadurch keine Aussage definiert.

Ausgehend von dem obigen Beispiel legen wir fest, dass die *Implikation* „wenn p, dann q“ (kurz $p \Rightarrow q$) nur dann falsch sein soll, wenn p wahr und q falsch ist. Die Wahrheitstafel lautet also:[2]

p	q	$p \Rightarrow q$
W	*W*	*W*
W	F	F
F	*W*	*W*
F	F	*W*

Beachten Sie: Die Wahrheit von $p \Rightarrow q$ erlaubt keine Aussagen über die Wahrheit von p, wenn q wahr ist! Will man dagegen ausdrücken, dass die zwei Aussagen p und q den selben

[2] Überlegen Sie sich anhand von Wahrheitstafeln, was passiert, wenn Sie den Wahrheitswert in Fall 3 und 4 anders festlegen würden!

Wahrheitswert haben, muss man die Aussage $(p \Rightarrow q) \wedge (q \Rightarrow p)$ verwenden:

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$(p \Rightarrow q) \wedge (q \Rightarrow p)$
W	W	W	W	W
W	F	F	W	F
F	W	W	F	F
F	F	W	W	W

Diese Aussage bezeichnet man als *Äquivalenz* und schreibt dafür auch kurz $p \Leftrightarrow q$ (gesprochen „genau dann q, wenn p“).

Zusammengesetzte Aussagen lassen sich natürlich wieder verknüpfen: zum Beispiel kann man das *exklusive Oder* von p und q ausdrücken durch $(p \vee q) \wedge (\neg(p \wedge q))$, wie man sich anhand der stufenweisen Betrachtung der Wahrheitstafeln überlegen kann. Dabei ist auf die Reihenfolge zu achten, die hier durch die Klammern angegeben wird; man geht immer von innen nach außen vor. Um Klammern zu sparen, vereinbaren wir die folgenden *Bindungsstärken* (analog zur Punkt-vor-Strich-Regel):

$$\neg \quad \text{vor} \quad \wedge, \vee \quad \text{vor} \quad \Rightarrow, \Leftrightarrow$$

Die Aussage $\neg p \wedge q \Rightarrow r$ ist also eindeutig festgelegt als äquivalent zu $((\neg p) \wedge q) \Rightarrow r$.

Beispiel 1.5. Es seien p, q und r Aussagen. Dann können wir zum Beispiel folgende zusammengesetzte Aussagen *formalisieren*, das heisst, mit logischen Verknüpfungen ausdrücken:

- „Weil p nicht zutrifft, kann auch q nicht gelten“: $\neg p \Rightarrow \neg q$
- „Es gilt weder p noch q“: $\neg p \wedge \neg q$.
- „Weil sowohl p als auch q nicht gelten, muss r zutreffen“: $\neg p \wedge \neg q \Rightarrow r$.

1.3 TAUTOLOGIEN UND KONTRADIKTIONEN

Eine zusammengesetzte Aussage, die unabhängig von den Wahrheitswerten der verknüpften Aussagen immer wahr ist, nennt man *Tautologie*. So hat $p \vee (\neg p)$ stets den Wahrheitswert „wahr“: entweder p ist wahr, dann ist die Oder-Verknüpfung auch wahr, oder p ist falsch und damit $\neg p$ wahr – und die Disjunktion ist wahr.

Eine zusammengesetzte Aussage, die unabhängig vom Wahrheitswert der verknüpften Aussagen immer falsch ist, nennt man *Kontradiktion*. Das klassische Beispiel ist $p \wedge (\neg p)$: keine Aussage kann gleichzeitig wahr und falsch sein.

Von besonderem Interesse sind Tautologien, die die Äquivalenz von zwei zusammengesetzten Aussagen ausdrücken. Mit ihrer Hilfe können zusammengesetzte Aussagen so umgeformt werden, dass ihr Wahrheitswert leichter entscheidbar ist. (Wären Wahrheitstafeln das einzige

Mittel, um zu entscheiden, ob zusammengesetzte Aussagen wahr oder falsch sind, wäre die Mathematik nämlich eine äußerst mühselige Angelegenheit.)

Ein einfaches Beispiel ist die Aussage $(\neg(\neg p)) \Leftrightarrow p$ (auch *doppelte Negation* genannt), die unabhängig vom Wahrheitswert von p wahr ist:

p	$\neg p$	$\neg(\neg p)$	$(\neg(\neg p)) \Leftrightarrow p$
W	F	*W*	*W*
F	*W*	F	*W*

Ersetzen wir in einer zusammengesetzten Aussage also $\neg(\neg p)$ durch p, ändert sich der Wahrheitswert nicht.

Ein weiteres Beispiel betrifft die Implikation: Anstatt sie als neue Verknüpfung einzuführen, können wir sie durch Negation und Disjunktion ausdrücken: $(p \Rightarrow q)$ ist äquivalent zu $(\neg p \vee q)$, wie man sich mit Hilfe der Wahrheitstafel versichert:[3]

p	q	$p \Rightarrow q$	$\neg p$	$\neg p \vee q$	$(p \Rightarrow q) \Leftrightarrow (\neg p \vee q)$
W	*W*	*W*	F	*W*	*W*
W	F	F	F	F	*W*
F	*W*	*W*	*W*	*W*	*W*
F	F	*W*	*W*	*W*	*W*

Mit Hilfe dieser Tautologie können wir auch die Negation der Implikation bilden:

$$\neg(p \Rightarrow q) \Leftrightarrow \neg(\neg p \vee q) \Leftrightarrow p \wedge \neg q$$

wobei wir die doppelte Negation verwendet haben. Wir erkennen wieder: Die Aussage „p impliziert q“ ist dann und nur dann falsch, wenn p wahr und q falsch ist.

Eine extrem nützliche Tautologie ist die *Kontraposition*

$$(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p).$$

Wir bemühen wieder die Wahrheitstafel:

p	q	$p \Rightarrow q$	$\neg q$	$\neg p$	$\neg q \Rightarrow \neg p$	$(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$
W	*W*	*W*	F	F	*W*	*W*
W	F	F	*W*	F	F	*W*
F	*W*	*W*	F	*W*	*W*	*W*
F	F	*W*	*W*	*W*	*W*	*W*

Gilt $p \Rightarrow q$, so nennt man p eine *hinreichende Bedingung* für q. Wir können nämlich aus der Wahrheit von p auf die Wahrheit von q schließen, aber q kann auch wahr sein, wenn p falsch ist. Aufgrund der Kontraposition erkennen wir: ist umgekehrt q falsch, so muss auch p falsch sein. Gilt daher $p \Rightarrow q$, sagt man, dass q eine *notwendige Bedingung* für p ist: nur wenn q wahr ist, kann p wahr sein. Gilt $p \Leftrightarrow q$, so ist p eine notwendige und hinreichende Bedingung für q, und umgekehrt.

[3] Oft wird daher die Implikation $p \Rightarrow q$ *definiert* als die Aussage $\neg p \vee q$.

Ein weiteres wichtiges Beispiel von Tautologien sind die *de Morganschen Regeln*:

- $(\neg(p \vee q)) \Leftrightarrow (\neg p \wedge \neg q)$, aufgrund der Wahrheitstafel

p	q	$p \vee q$	$\neg(p \vee q)$	$\neg p \wedge \neg q$	$(\neg(p \vee q)) \Leftrightarrow (\neg p \wedge \neg q)$
W	W	W	F	F	W
W	F	W	F	F	W
F	W	W	F	F	W
F	F	F	W	W	W

- $(\neg(p \wedge q)) \Leftrightarrow (\neg p \vee \neg q)$, wofür die Wahrheitstafel dem Leser überlassen bleibt.

Es gelten weiterhin die *Distributivgesetze*:

- $(p \wedge (q \vee r)) \Leftrightarrow ((p \wedge q) \vee (p \wedge r))$,
- $(p \vee (q \wedge r)) \Leftrightarrow ((p \vee q) \wedge (p \vee r))$.

und die *Assoziativgesetze*

- $((p \wedge q) \wedge r) \Leftrightarrow (p \wedge (q \wedge r))$,
- $((p \vee q) \vee r) \Leftrightarrow (p \vee (q \vee r))$.

Beispiel 1.6. Wir können mit Hilfe dieser Regeln die Aussagen aus Beispiel 1.5 vereinfachen (oder zumindest umformen):

- Die Aussage $\neg p \Rightarrow \neg q$ ist nach Kontraposition äquivalent zu $q \Rightarrow p$.
- Die Aussage $\neg p \wedge \neg q$ ist nach der de Morganschen Regel äquivalent zu $\neg(p \vee q)$.
- Die Aussage $\neg p \wedge \neg q \Rightarrow r$ ist wieder äquivalent zu $\neg(p \vee q) \Rightarrow r$, was wir auch umformen können zu $\neg\neg(p \vee q) \vee r$. Doppelte Negation und das Assoziativgesetz erlauben uns die einfache Schreibweise $p \vee q \vee r$.

1.4 QUANTOREN

Viele mathematische Aussagen gelten für bestimmte oder auch alle Objekte einer Klasse; etwa hat jedes Dreieck drei Seiten. Bislang können wir aber nur Aussagen über spezifische Objekte oder Situationen präzise formulieren. Wir müssen also die im letzten Abschnitt betrachtete Logik erweitern.[4]

Der erste Schritt ist, Aussagen von spezifischen Objekten zu lösen. Statt der Aussage p: „die Kreide ist weiß“ betrachten wir die *Aussageform* $P(x)$: „x ist weiß“; x nennt man dabei eine

[4] Die in diesem Abschnitt behandelte Erweiterung wird als *Prädikatenlogik* bezeichnet, im Gegensatz zu der davor betrachteten *Aussagenlogik*.

freie Variable[5]. Aussageformen haben keinen Wahrheitswert; erst wenn die freie Variable durch ein konkretes Objekt ersetzt wird, erhält man eine Aussage, die wahr oder falsch sein kann. Man spricht dabei von einer *Belegung* von x. So entsteht durch die Belegung von x durch 4 aus der Aussageform $P(x)$: „x ist eine gerade Zahl“ die Aussage $P(4)$: „4 ist eine gerade Zahl“.

Nun wäre nichts gewonnen, wenn wir trotzdem für jede Belegung von x die Aussage $P(x)$ separat prüfen müssten. Wir müssen also Aussagen über Aussageformen bilden: Eine neue Aussage entsteht, wenn freie Variable *gebunden* werden. Dies geschieht durch die *Quantoren*:

- Wenn man ausdrücken möchte, dass alle Objekte einer Klasse eine bestimmte Eigenschaft haben (etwa das alle Dreiecke drei Seiten haben), verwendet man den *Allquantor* ($\forall x$, „für alle x gilt“).

  Beispiel: $\forall x : P(x)$, „für alle Zahlen x gilt: x ist eine gerade Zahl“, ist eine (falsche) Aussage.

  Mögliche umgangssprachliche Aussagen, die sich mit dem Allquantor ausdrücken lassen, sind „Alle Pflanzen sind essbar“ (oder auch nur „Pflanzen sind essbar“) und „jede Pflanze ist essbar“. In der Mathematik ist auch die Formulierung „für beliebige x gilt“ üblich.

- Mit dem *Existenzquantor* ($\exists x$, „es gibt ein x, so dass gilt“) kann man ausdrücken, dass es ein Objekt gibt[6], das gewisse vorgegebene Eigenschaften erfüllt (etwa die Lösung einer quadratischen Gleichung ist).

  Beispiel: $\exists x : P(x)$, „es gibt eine Zahl x, so dass gilt: x ist eine gerade Zahl“, ist eine (richtige) Aussage.

  Mögliche umgangssprachliche Aussagen, die sich mit dem Existenzquantor ausdrücken lassen, sind „es gibt essbare Pflanzen“, „manche Pflanzen sind essbar“ und „es gibt (mindestens) eine essbare Pflanze“.

  Beachten Sie hier: Eine Aussage mit Existenzquantor wie $\exists x P(x)$ ist wahr, solange es *mindestens* ein Objekt a gibt, für das $P(a)$ wahr ist – die Frage, ob für noch ein weiteres Objekt b die Aussage $P(b)$ gilt, bleibt offen. Will man ausdrücken, dass *genau*

[5] Der Buchstabe x ist hier willkürlich; $P(x)$ und $P(y)$ (oder auch $P(m)$) sind die selbe Aussageform.

[6] Diese Aussage hat freilich keinen ontologischen Charakter, sondern drückt lediglich aus, dass für eine bestimmte Belegung a von x der Wahrheitswert der Aussage $P(a)$ mit Hilfe von Tautologien auf den Wahrheitswert der Axiome zurückgeführt werden kann. Alles weitere ist Aufgabe der Philosophie der Mathematik.

ein Objekt existiert, muss man dies explizit fordern, etwa mit der Konstruktion[7]

$$\exists x : (P(x) \wedge (\forall y : (P(y) \Rightarrow (y = x)))).$$

Dies wird oft mit dem Zeichen $\exists! x : P(x)$ oder $\exists_1 x : P(x)$ abgekürzt. Denken Sie aber stets daran, dass eine Aussage der Form „es existiert genau ein“ immer aus zwei Teilen besteht – der *Existenz* einer Belegung a von x so dass $P(a)$ wahr ist, und der *Eindeutigkeit* dieser Belegung (dass also $P(b)$ für alle anderen Belegungen b ungleich a falsch ist).

Offen gelassen ist hier noch, was für Objekte überhaupt in Frage kommen. Dies ist in der Mathematik eigentlich zwingend nötig, und lässt sich mit den im Kapitel 2 vorgestellten Formulierungen präzise angeben; wir vereinbaren hier vorläufig, dass alle x betrachtet werden, für die $P(x)$ eine sinnvolle Aussage ist (im Beispiel oben etwa Zahlen bzw. Pflanzen, aber keine Schuhe).

NEGATION VON QUANTORENAUSSAGEN. Auch aus Aussagen, die Quantoren enthalten, können durch Verknüpfung neue gebildet werden. Genauso können die bereits bekannten Tautologien angewendet werden, um solche Aussagen umzuformen. Zunächst legen wir fest, wie die Negation von Quantorenaussagen mit Hilfe der Negation von Aussageformen ausgedrückt werden soll.

Dazu betrachten wir die falsche Aussage p: „Jeder Mensch hat eine Schwester“. Mit Hilfe der Aussageform $P(x)$: „x hat eine Schwester“ können wir p auch formulieren als $\forall x : P(x)$. Wie lautet nun die Negation $\neg p$, die ja wahr sein muss? Die Negation der Aussageform $P(x)$ lautet $\neg P(x)$: „ x hat keine Schwester“. Die nahe liegende Möglichkeit $\forall x : \neg P(x)$ bedeutet: „Für jeden Menschen gilt: er hat keine Schwester“. Das ist natürlich auch falsch; wir wollten vielmehr ausdrücken, dass nicht jeder Mensch eine Schwester hat, dass es also auch Menschen gibt, die keine Schwester haben. Letzteres können wir aber mit dem Existenzquantor ausdrücken: $\exists x : \neg P(x)$.

Umgekehrt gilt: die Negation der Aussage „es gibt einen fliegenden Elefanten“ muss natürlich „es gibt keinen fliegenden Elefanten“ sein. Verwenden wir die Aussageform $P(x)$: „x fliegt“, so können wir die ursprüngliche Aussage formulieren als $\exists x : P(x)$. Die Negation muss dann lauten $\forall x : \neg P(x)$, „für jeden Elefanten gilt: er fliegt nicht“. Die Alternative, $\exists x : \neg P(x)$, „es gibt einen Elefanten, der nicht fliegt“, erlaubt immer noch, dass es einen fliegenden Elefanten gibt – und damit hätten wir eine Kontradiktion: p und $\neg p$ könnten beide wahr sein.

[7] Dabei muss streng genommen natürlich erst klar gestellt werden, was die Gleichheit $x = y$ bedeutet. Dies ist nicht Teil der hier skizzierten Logik, sondern muss separat definiert werden (man spricht manchmal von einer *Theorie der Identität*). Gottfried Leibniz schlug folgende Definition vor: $x = y$ genau dann, wenn x jede Eigenschaft hat, die y hat, und umgekehrt. Damit lässt sich der Identitätsbegriff sauber fundieren; wir berufen uns im Weiteren aber auf die naive Anschauung.

Um Kontradiktionen zu vermeiden, haben wir also nur eine Wahl, die Negation von Quantorenaussagen festzulegen.[8] Wir halten fest:

- $(\neg(\forall x : P(x))) \Leftrightarrow (\exists x : (\neg P(x)))$,
- $(\neg(\exists x : P(x))) \Leftrightarrow (\forall x : (\neg P(x)))$.

Für $\neg(\exists x : P(x))$ („es gibt kein x, so dass $P(x)$ gilt“) wird oft auch kurz $\nexists x : P(x)$ geschrieben.

Wir werden im Rahmen der Beweisstrategien noch einmal auf dieses Thema zurückkommen.

REIHENFOLGE VON QUANTOREN. Die größte Schwierigkeit beim Umgang mit Quantoren taucht erfahrungsgemäß auf, wenn eine Aussageform mehrere durch Quantoren gebundene Variablen enthält: dann spielt nämlich die Reihenfolge der Quantoren eine wichtige Rolle. Wir betrachten das an einem einfachen Beispiel:

Beispiel 1.7. $P(x, y)$ ist die Aussageform „x trinkt gerne y“. Dann gibt es folgende Möglichkeiten, x und y durch Quantoren zu binden:

1. $\forall x : (\forall y : (P(x, y)))$: Für alle Personen x gilt: für alle Getränke y gilt: x trinkt gerne y.

   Jede Person trinkt gerne jedes Getränk.

2. $\forall y : (\forall x : (P(x, y)))$: Für alle Getränke y gilt: für alle Personen x gilt: x trinkt gerne y.

   Jedes Getränk wird also von jeder Person gerne getrunken. Dies ist eine äquivalente Aussage zu Fall 1.

3. $\forall x : (\exists y : (P(x, y)))$: Für alle Personen x gilt: es gibt ein Getränk y, so dass gilt: x trinkt gerne y.

   Jede Person hat also ein (möglicherweise unterschiedliches) Getränk, das sie gerne trinkt.

4. $\exists y : (\forall x : (P(x, y)))$: Es gibt ein Getränk y, so dass gilt: für alle Personen x gilt: x trinkt gerne y.

   Es existiert also (mindestens) ein Getränk, das jeder gerne trinkt. Beachten Sie den Unterschied zu Fall 3: Obwohl auch hier jeder gerne ein Getränk trinkt, wird zusätzlich behauptet, dass alle Personen das gleiche Getränk mögen!

[8] In der Formalisierung der mathematischen Logik wird dies dadurch erreicht, dass der Existenzquantor über den Allquantor definiert wird: $\exists x P(x)$ ist dort nur eine Kurzschreibweise für $\neg(\forall x(\neg P(x)))$.

5. $\forall y : (\exists x : (P(x, y)))$: Für alle Getränke y gilt: es gibt eine Person x, so dass gilt: x trinkt gerne y.

   Für jedes Getränk gibt es also eine Person, die es gerne trinkt. Im Unterschied zu Fall 3 wird hier behauptet, dass selbst für das seltsamste Getränk mindestens eine Person existiert, die es gerne trinkt. Umgekehrt kann es Personen geben, die gar kein Getränk gerne trinken.

6. $\exists x : (\forall y : (P(x, y)))$: Es gibt eine Person x, so dass gilt: für alle Getränke y gilt: x trinkt gerne y.

   Es existiert also (mindestens) eine Person, die jedes Getränk gerne trinkt. Im Gegensatz zu Fall 5 ist es hier die selbe Person, die die verschiedenen Getränke mag. (Vergleichen Sie Fall 3 und 4: Die Reihenfolge $\exists\forall$ ergibt eine stärkere Behauptung als $\forall\exists$.)

7. $\exists x : (\exists y : (P(x, y)))$: Es gibt eine Person x, so dass gilt: es gibt ein Getränk y, so dass gilt: x trinkt gerne y.

   Mindestens eine Person mag mindestens ein Getränk (was sicher richtig ist).

8. $\exists y : (\exists x : (P(x, y)))$: Es gibt ein Getränk y, so dass gilt: es gibt eine Person x, so dass gilt: x trinkt gerne y.

   Mindestens ein Getränk wird also von mindestens einer Person gerne getrunken. Dies ist wieder eine äquivalente Aussage zu der in Fall 7: Es gibt eine Person, die (irgend)ein Getränk gerne trinkt.

Merke: *Nur gleiche Quantoren dürfen vertauscht werden!*

Beachten Sie auch, dass bei der Negation von Aussagen, die mehrere Quantoren enthalten, die Negation Schritt für Schritt von aussen nach innen angewandt wird:

$$(\neg(\forall x : (\exists y : P(x, y)))) \Leftrightarrow (\exists x : \neg(\exists y : P(x, y))) \Leftrightarrow (\exists x : (\forall y : (\neg P(x, y)))).$$

Neben der Negation und der Vertauschbarkeit gleicher Quantoren existieren noch weitere Tautologien mit Quantorenaussagen, wie etwa die *Distributivgesetze*

- $(\forall x : (P(x) \wedge Q(x))) \Leftrightarrow (\forall x : P(x)) \wedge (\forall x : Q(x))$,
- $(\exists x : (P(x) \vee Q(x))) \Leftrightarrow (\exists x : P(x)) \vee (\exists x : Q(x))$.

NAIVE MENGENLEHRE

2

Da es in der Mathematik darum geht, Gemeinsamkeiten herauszuarbeiten, hat sich die Mengenlehre als ausgesprochen fruchtbare „lingua franca“ der Mathematik herausgestellt: Die Aussagen und Begründungen so verschiedener Teilgebiete wie Analysis, Algebra, Geometrie, Topologie und Wahrscheinlichkeitsrechnung können alle mit Hilfe der Begriffe der Mengenlehre dargestellt werden.

2.1 MENGEN UND IHRE ELEMENTE

Unter einer *Menge* verstehen wir eine Zusammenfassung wohlunterschiedener Objekte unseres Denkens oder unserer Anschauung zu einem Ganzen[1]. Dabei ist der fundamentale Begriff der der Zugehörigkeit eines Objekts zu einer Menge: Ist das Objekt x in der Menge M enthalten, so sagen wir, x ist *Element* von M und schreiben $x \in M$.[2] Ist umgekehrt x nicht in M enthalten, schreiben wir $x \notin M$. Damit $x \in M$ eine Aussage(nform) definiert, die wir mit den Mitteln der Logik behandeln können, muss also für jedes Objekt x und jede Menge M eine und nur eine der beiden Möglichkeiten zutreffen. Für unsere Zwecke genügt es daher, Mengen dadurch zu definieren, dass jedes gegebene Objekt entweder Element der Menge ist oder nicht.

Eine fundamentale Eigenschaft von Mengen ist, dass sie durch ihre Elemente eindeutig festgelegt werden; dies bezeichnet man als *Extensionalitätsprinzip*. Um eine bestimmte Menge anzugeben, haben wir dabei die folgenden Möglichkeiten:

- Durch komplette Aufzählung aller Elemente. Beispiele sind die Mengen $\{1, 2, 3\}$ und $\{\text{Clason, Lettl, Müller, Propst, Ring, Tomaschek}\}$.
- Elemente können auch durch Auslassungszeichen ersetzt werden, wenn eindeutig erkennbar ist, welche Elemente ausgelassen werden.

  Beispiel: $\{1, 2, 3, \dots, 10\}$, $\{1, 2, 3, \dots\}$, $\{2, 4, 6, 8, \dots\}$.

[1] Diese anschauliche – *naive* – Definition geht auf Georg Cantor zurück.

[2] Auch in der *axiomatischen* Mengenlehre ist $\in$ der fundamentale Begriff, aus dem alles weitere abgeleitet wird.

- Durch Angabe einer Eigenschaft, die alle Elemente der Menge (und nur diese) erfüllen; ein Beispiel ist die oben genannte Menge aller Dozenten dieser Veranstaltung. Um solche Mengen präzise anzugeben, verwenden wir Aussageformen:

  $$\{x : P(x)\}$$

  ist die Menge aller Belegungen a von x, für die die Aussage $P(a)$ wahr ist. Insbesondere gilt für alle a die Tautologie $(a \in \{x : P(x)\}) \Leftrightarrow P(a)$. Man spricht von einer *prädikativen Definition*. Dies ist bei weitem die wichtigste Möglichkeit, da sie auch sehr abstrakte Mengen zulässt. Ausserdem kann man jede Aufzählung in dieser Form angeben: Die Menge $\{1, 2\}$ lässt sich zum Beispiel schreiben als $\{x : (x = 1) \vee (x = 2)\}$.

  Beispiel: $\{x : x \text{ gerade Zahl}\}$, $\left\{x : x^2 = 2\right\}$.

Beachten Sie, dass das Extensionalitätsprinzip verlangt, dass Mengen aus *wohlunterschiedenen* Objekten bestehen sollen. So beschreiben $\{1, 2, 3, 3\}$ und $\{1, 2, 3\}$ die selbe Menge, da beide Aufzählungen genau die Elemente 1, 2 und 3 enthalten. Ebenso spielt die Reihenfolge der Aufzählung keine Rolle: $\{3, 2, 1\}$ und $\{1, 2, 3\}$ sind die selbe Menge.

Mit Hilfe von Mengen können wir nun auch explizit festlegen, welche Belegungen für Aussageformen in Frage kommen: Nämlich nur solche, die Elemente einer vorgegebenen Menge M sind. Wir schreiben dann:

- $\forall x \in M : P(x)$ beziehungsweise
- $\exists x \in M : P(x)$.

Beispiel: $\forall x \in \{1, 2, 3\} : x > 0$ oder $\exists x \in \{1, 2, 3\} : x < 3$.

Umgekehrt verlangen wir, dass in einer prädikativen Definition immer eine Grundmenge angegeben wird: $\{x \in X : P(x)\}$. (Nicht zuletzt, weil zum Beispiel $\left\{x : x^2 = 4\right\}$ keine Menge definiert, da nicht feststeht, ob -2 Element ist oder nicht, solange wir nicht $\left\{x \in \mathbb{R} : x^2 = 4\right\}$ oder $\left\{x \in \mathbb{N} : x^2 = 4\right\}$ spezifizieren.)

TEILMENGEN. Eine neue Menge entsteht, wenn wir Elemente aus einer vorgegebenen Menge auswählen. So können wir aus $\{1, 2, 3\}$ die neue Menge $\{2, 3\}$ bilden, und erhalten eine *Teilmenge*. Wir präzisieren dies:

Wir betrachten zwei Mengen M und N. Dann sagen wir, M ist *Teilmenge* von N genau dann, wenn

$$\forall x \in M : x \in N$$

gilt. Wir schreiben dafür auch $M \subseteq N$. Umgekehrt nennen wir M *Obermenge* von N genau dann, wenn

$$N \subseteq M$$

gilt, und schreiben dafür $M \supseteq N$.

Damit können wir auch das Extensionalitätsprinzip formalisieren: M ist *gleich* N genau dann, wenn

$$(M \subseteq N) \wedge (N \subseteq M)$$

gilt. Jedes Element von M ist auch in N enthalten und umgekehrt, und daher bezeichnen M und N die selbe Menge; wir schreiben $M = N$. Sind M und N nicht gleich, schreiben wir $M \neq N$. Insbesondere bedeutet das Extensionalitätsprinzip für prädikativ definierte Mengen, dass $\{x \in X : P(x)\} = \{x \in X : Q(x)\}$ ist genau dann, wenn $P(x) \Leftrightarrow Q(x)$ für alle $x \in X$ gilt.

Die Gleichheit gibt uns ein weiteres nützliches Mittel in die Hand, um Mengen anzugeben: Wir behaupten einfach, dass die neue Menge N gleich ist mit einer bereits bekannten (oder explizit angegebenen) Menge M. Wir schreiben in diesem Fall $N := M$; der Doppelpunkt soll klarstellen, dass dies eine Definition, und keine Aussage, ist. So können wir die Menge $\mathbb{N} := \{1, 2, 3, \dots\}$ der natürlichen Zahlen definieren.

Schliesslich nennen wir M *echte Teilmenge* von N genau dann, wenn

$$(\forall x \in M : x \in N) \wedge (M \neq N)$$

gilt. Wir schreiben dafür auch $M \subsetneq N$, wenn diese Eigenschaft von Bedeutung ist (jede echte Teilmenge ist natürlich a fortiori eine Teilmenge).[3]

2.2 DIE LEERE MENGE

Eine Menge, die keine Elemente enthält, nennen wir *leere Menge*; solch eine Menge wird mit $\emptyset$ (oder, früher häufiger, $\{\}$) bezeichnet.[4]

Die leere Menge hat einige kontraintuitive Eigenschaften, die aber logisch zwingend sind. Wir legen zuerst fest, wie Quantoren über die leere Menge zu verstehen sind:

- Existenzaussagen sind immer falsch: Für eine beliebige Aussageform $P(x)$ hat $\exists x \in \emptyset : P(x)$ immer den Wahrheitswert „falsch“. Dies ist unmittelbar einleuchtend: Da die leere Menge keine Elemente enthält, kann sie insbesondere kein Element x enthalten, für das $P(x)$ wahr ist.

[3] Oft findet man auch die Schreibweise $M \subset N$. Leider wird diese in der Literatur sehr uneinheitlich gehandhabt: Je nach Autor wird sie für allgemeine Teilmengen oder für echte Teilmengen verwendet.

[4] Für unsere Zwecke ist diese naive Definition ausreichend. Eine formale Definition wäre $\emptyset := \{x \in X : \neg(x = x)\}$ für eine beliebige Menge X.

- Allaussagen sind dagegen immer wahr: Für eine beliebige Aussageform $P(x)$ hat $\forall x \in \emptyset : P(x)$ immer den Wahrheitswert „wahr“. Dies folgt zwingend aus der obigen Festlegung und der Negation von Quantoren: $\neg(\exists x \in \emptyset : \neg P(x))$ ist wahr, und nach Festlegung gleichbedeutend mit $\forall x \in \emptyset : \neg(\neg P(x))$ was (aufgrund der doppelten Negation) den gleichen Wahrheitswert wie $\forall x \in \emptyset : P(x)$ hat.

Daraus folgt sofort, dass eine leere Menge Teilmenge jeder Menge ist: Sei M eine beliebige Menge, dann gilt (mit $P(x)$: „$x \in M$“)

$$\forall x \in \emptyset : x \in M,$$

und damit nach Definition $\emptyset \subseteq M$.

Weiter folgern wir, dass es nur eine leere Menge geben kann: Sind $\emptyset$ und $\{\}$ beide leere Mengen (d.h. enthalten keine Elemente), so gilt einerseits $\emptyset \subseteq \{\}$ (eine leere Menge ist Teilmenge jeder Menge, wozu auch leere Mengen zählen) und umgekehrt aus dem gleichen Grund $\{\} \subseteq \emptyset$. Dies ist aber genau die Definition der Gleichheit; daher gilt $\emptyset = \{\}$.

2.3 MENGENOPERATIONEN

Nun wollen wir aus gegebenen Mengen neue Mengen bilden.

VEREINIGUNG, DURCHSCHNITT, KOMPLEMENT. Neue Mengen lassen sich auch bilden, indem wir zwei Mengen zu einer zusammenfassen. So können wir aus den Mengen $\{1, 2, 3\}$ und $\{2, 3, 4\}$ die Menge $\{1, 2, 3, 4\}$ bilden. Für beliebige Mengen A und B soll die *Vereinigung* beider Mengen, die wir mit $A \cup B$ bezeichnen, alle Elemente enthalten, die in einer der beiden Mengen enthalten sind. Seien im folgenden stets $A = \{x : P(x)\}$ und $B = \{x : Q(x)\}$ einer gegebenen Menge X. (Da wir jede Menge M in prädikativer Form schreiben können als $\{x : x \in M\}$, stellt dies keine Einschränkung dar). Dann definieren wir:

$$A \cup B := \{x \in X : (x \in A) \vee (x \in B)\} = \{x \in X : P(x) \vee Q(x)\}.$$

Umgekehrt können wir alle Elemente auswählen, die in beiden Mengen enthalten sind, etwa aus $\{1, 2, 3\}$ und $\{2, 3, 4\}$ die neue Menge $\{2, 3\}$ bilden. Für zwei beliebige Mengen A und B besteht also der *Durchschnitt* $A \cap B$ aus allen Elementen, die in beiden Mengen enthalten sind:

$$A \cap B := \{x \in X : (x \in A) \wedge (x \in B)\} = \{x \in X : P(x) \wedge Q(x)\}.$$

Zwei Mengen nennen wir *disjunkt*, wenn ihre Schnittmenge leer ist: $A \cap B = \emptyset$.

Eine weitere Möglichkeit ist die *(Mengen-)Differenz* $A \setminus B$, die alle Elemente von A enthält, die nicht in B sind (im letzten Beispiel also $\{1\}$). Für beliebige Mengen definieren wir:

$$A \setminus B := \{x \in X : (x \in A) \wedge (x \notin B)\} = \{x \in X : P(x) \wedge \neg Q(x)\}.$$

Beachten Sie, dass B keine Teilmenge von A sein muss, und dass $A \setminus B \neq B \setminus A$ ist. Die *symmetrische Differenz* $A \triangle B$ ist definiert als:

$$A \triangle B := (A \setminus B) \cup (B \setminus A),$$

Beispiel: $A := \{1, 2, 3\}$, $B := \{2, 3, 4\}$, $A \setminus B = \{1\}$, $B \setminus A = \{4\}$, $A \triangle B = \{1, 4\}$.

Häufig ist die Menge A dabei eine „Grundmenge“ (etwa die Menge aller natürlichen Zahlen $\mathbb{N}$) und B eine Teilmenge (etwa die Menge aller geraden Zahlen). Dann bezeichnet man $A \setminus B$ auch als *Komplement von* B *in* A und schreibt auch B^c (manchmal auch $\overline{B}$ oder $\complement B$). Dies lässt sich auch ausdrücken als

$$B^c := \{x \in A : x \notin B\} = \{x \in A : \neg Q(x)\}.$$

Die Operationen kann man sich mit Hilfe von Venn-Diagrammen verdeutlichen.

Mit Hilfe dieser Definitionen und den Tautologien aus Abschnitt 1.3 können wir nun einige Rechenregeln für Mengenoperationen aufstellen. Es seien X eine Menge und A, B, C Teilmengen von X. Ferner bezeichne A^c das Komplement von A in X. Dann gilt:

- $(A^c)^c = A$,
- $(A \cup B)^c = A^c \cap B^c$,
- $(A \cap B)^c = A^c \cup B^c$,
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Überlegen wir uns, warum die erste Gleichheit gelten muss. Ein Venn-Diagramm reicht nicht, um alle möglichen Zweifel auszuschließen – die wirklich harten Fälle, an denen die Aussage scheitern könnte, sind eventuell so verwinkelt, dass wir sie gar nicht zeichnen können. Wir müssen also ausschließlich auf Basis der Definitionen und Tautologien argumentieren. Nach Definition ist $A^c = \{x \in X : x \notin A\}$. Fassen wir $x \notin A$ als die Aussageform $\neg(x \in A)$ auf, können wir folgern:

$$(A^c)^c = \big(\{x \in X : \neg(x \in A)\}\big)^c = \{x \in X : \neg(\neg(x \in A))\}.$$

Aufgrund der doppelten Negation gilt aber:

$$\forall x \in X : \big((x \in A) \Leftrightarrow \neg(\neg(x \in A))\big).$$

Nach dem Extensionalitätsprinzip muss daher $(A^c)^c = A$ gelten.

Auf ähnliche Weise können auch die restlichen Identitäten gezeigt werden – was Ihnen als Übung überlassen bleibt.

KARTESISCHES PRODUKT. Wir können aus zwei Mengen A und B auch eine neue Menge bilden, indem wir Paare von Elementen bilden: für $a \in A$ und $b \in B$ betrachten wird das *geordnete Paar* (a, b). Wie der Name andeutet, spielt die Reihenfolge eine wichtige Rolle: $(a, b) \neq (b, a)$. Wir legen fest:[5]

$$(a, b) = (c, d) :\Leftrightarrow (a = c \wedge b = d)$$

Die Menge aller solcher Paare bezeichnet man als das *kartesische Produkt* von A und B:

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

Beispiel: $A = \{a, b, c\}$, $B = \{1, 2\}$, $A \times B = \{(a, 1), (a, 2), (b, 1), (b, 2), (c, 1), (c, 2)\}$.

Häufig trifft der Fall ein, dass mehrmals die gleiche Menge verwendet wird:

$$A \times A = \{(a_1, a_2) \,:\, a_1 \in A \wedge a_2 \in A\},$$

wofür man auch kurz A^2 schreibt.

Beispiel: $\{0, 1\}^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

Diese Konstruktion kann verallgemeinert werden: Aus den Elementen der drei Mengen A, B, C kann man geordnete Tripel (a, b, c) bilden, und daraus das kartesische Produkt $A \times B \times C$ (und schreibt wieder A^3 für $A \times A \times A$).[6]

2.4 POTENZMENGEN

Durch unsere Definition von Mengen ist ausdrücklich zugelassen, dass die Elemente einer Menge selber Mengen sind. Zum Beispiel ist $\{\{1, 2\}, \{2, 3\}\}$ eine Menge; auch „gemischte" Mengen sind möglich, etwa die Menge $\{1, 2, 3, \{1, 2, 3\}\}$. Ganz wichtig ist hier jedoch, zwischen Teilmengen und Elementen einer Menge zu unterscheiden! Insbesondere sollten Sie den Unterschied zwischen x und der Menge $\{x\}$, die x enthält, beachten.

Beispiel 2.1. Sei $A := \{1, 2, \{1\}, \{2, 3\}, \{3\}\}$. Dann gilt:

- $1 \in A$ (klar), $\{1\} \subseteq A$ (weil $1 \in A$), $\{1\} \in A$ (weil $\{1\}$ in der Auflistung vorkommt).
- $2 \in A$, $\{2\} \subseteq A$, $\{2\} \notin A$.
- $3 \notin A$ (denn $3 \neq \{3\}$), $\{3\} \not\subseteq A$ (weil $3 \notin A$), $\{3\} \in A$.
- $\{2, 3\} \not\subseteq A$ (da $3 \notin A$), aber $\{2, 3\} \in A$.

[5] Wir hätten geordnete Paare (a, b) auch rigoros als spezielle Mengen der Form $\{a, \{a, b\}\}$ einführen, und dann die hier festgelegte Äquivalenz aus den Tautologien der Mengenlehre ableiten können.

[6] Auch Tripel, Quadrupel, etc., können wir auf den Mengenbegriff zurückführen – wobei wir dabei streng genommen zeigen müssen, dass $A \times (B \times C)$ und $(A \times B) \times C$ auf die selbe Menge von Tripeln führen.

Eine häufig auftretende Menge von Mengen ist die *Potenzmenge* einer Menge M, die alle Mengen N enthält, die Teilmenge von M sind. Formal definieren wir die Potenzmenge $\mathcal{P}(M)$, indem wir festlegen:

$$N \in \mathcal{P}(M) \Leftrightarrow N \subseteq M$$

Beispiel 2.2. Für $M := \{1, 2, 3\}$ ist

$$\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}.$$

Bei der Definition der Potenzmenge haben wir die Menge nicht explizit (durch Liste der Elemente oder durch Aussageformen) angegeben, sondern implizit darüber definiert, wie man nachprüfen kann, ob ein Objekt Element der Menge ist. Wollten wir die Potenzmenge über eine Aussageform definieren, müssten wir schreiben $\mathcal{P}(M) := \{N \in X : N \subseteq M\}$, wobei X die „Menge aller Mengen" ist. Solch ein Objekt würde aber auf alle möglichen Widersprüche führen, wovon der bekannteste sicher die *Russellsche Antinomie* ist: Wir betrachten die Menge aller Mengen, die sich nicht selbst als Element enthalten (Mengen dürfen ja selber Mengen enthalten); formal definiert als $R := \{M \in X : M \notin M\}$. Gilt dann $R \in R$?

- Falls $R \in R$ gilt, ist nach Definition $R \notin R$ – eine Kontradiktion.
- Ist dagegen $R \notin R$, so muss gelten $R \in R$: nach Definition war ja R die Menge *aller* Mengen M, für die $M \notin M$ gilt – und da sollte R gerade nicht dabei sein. Wieder erhalten wir eine Kontradiktion.

Um solche inakzeptablen Widersprüche zu vermeiden, müssen wir also verbieten, prädikative Mengen über solche „universellen" Aussageformen zu definieren. Dies ist ein weiterer Grund, warum wir für prädikativ definierte Mengen zwingend eine Grundmenge angeben müssen: Für eine *explizit gegebene* Menge X ist $S := \{M \in X : M \notin M\}$ ohne Gefahr, da hier der Fall $S \notin S$ die Möglichkeit offen lässt, dass $S \notin X$ ist – womit der Widerspruch vermieden wäre.[7]

[7] So garantiert das *Aussonderungsaxiom* in der axiomatischen Mengenlehre nach Zermelo und Fränkel (welche die Grundlage für die moderne Mathematik darstellt) lediglich, dass für eine beliebige gegebene Menge solche prädikativ definierten *Teil*mengen existieren. Zusätzlich wird das *Fundierungsaxiom* eingeführt, um die lästigen selbst-enthaltenden Mengen explizit auszuschliessen.

FUNKTIONEN

Eine Funktion ist das abstrakte Abbild der Abhängigkeit einer Größe von einer anderen, und daher ein zentraler Begriff in der modernen Mathematik.

3

3.1 DEFINITION VON FUNKTIONEN

Seien X und Y nichtleere Mengen. Unter einer *Funktion* $f : X \to Y$ verstehen wir eine Zuordnungsvorschrift, die *jedem* Element $x \in X$ *genau ein* Element $y \in Y$ zuordnet. Wir schreiben dafür $y = f(x)$, und nennen y den *Wert von* f *an der Stelle* x oder *das Bild von* x *unter* f. Umgekehrt heisst x ein *Urbild von* y *unter* f.

Beachten Sie: Der *Definitionsbereich* X und der *Wertebereich* Y sind ein fester Bestandteil der Funktionsdefinition: $f : X \to Y$ und $g : V \to W$ sind dann und nur dann die gleiche Funktion, wenn $X = V$, $Y = W$ und $f(x) = g(x)$ für alle $x \in X$ gilt. Eine Funktion wird also üblicherweise in der folgenden Form angegeben:

$$f : X \to Y, \quad x \mapsto f(x),$$

wobei $x \mapsto f(x)$ die Zuordnungsvorschrift ist.

Beispiel 3.1. Auf der Menge X aller Menschen wird eine Funktion $f : X \to X$ definiert, indem man jeder Person ihre biologische Mutter zuordnet.

Die folgenden Beispiele sind unterschiedliche Funktionen:

- $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ($\mathbb{R}$ ist die Menge der reellen Zahlen),
- $f : \mathbb{N} \to \mathbb{N}, x \mapsto x^2$,
- $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}, x \mapsto \frac{1}{x}$,
- $f : \mathbb{R} \to \mathbb{R}, x \mapsto \begin{cases} 1, & x \in \mathbb{N}, \\ 0, & x \in \mathbb{R} \setminus \mathbb{N}. \end{cases}$

Die folgenden Beispiele sind *keine* Funktionen:

- $f : \mathbb{R} \to \mathbb{R}, x \mapsto \frac{1}{x}$, (diese Zuordnungsvorschrift ist für $x = 0$ nicht anwendbar),
- $f : \mathbb{R} \to \mathbb{R}, x \mapsto \begin{cases} 1, & x \in \mathbb{N}, \\ 0, & x \in \mathbb{R}, \end{cases}$ (für $x \in \mathbb{N}$ ist die Zuordnung nicht eindeutig).

Die Zuordnung von Stelle x und Funktionswert $f(x)$ können wir auch als Paarbildung $(x, f(x))$ auffassen. Dazu definieren wir den *Graph* einer Funktion $f : X \to Y$ als eine Teilmenge des kartesischen Produkts von X und Y:

$$G_f := \{(x, y) \in X \times Y : x \in X \text{ und } y = f(x)\}.$$

Es gilt also $f(x) = y$ genau dann, wenn $(x, y) \in G_f$ ist.[1] Die Darstellung einer Funktion über ihren Graphen ist eine Abstraktion der Zusammenstellung von Funktionswerten in einer Tabelle oder der üblichen Veranschaulichung reeller Funktionen als „Kurve“.

Seien X und Y Mengen, und $A \subseteq X$ eine Teilmenge von X. Dann können wir folgende nützliche Funktionen definieren:

- Die *Identität* $\mathrm{id}_X : X \to X, x \mapsto x$, bildet jedes Element von X auf sich selbst ab.
- Die *Einbettung* $j : A \to X, x \mapsto x$, bildet jedes Element von A auf sich selbst ab (das aber als Element von X aufgefasst wird).
- Die *Projektionen* in $X \times Y$ sind die Funktionen

$$p_1 : X \times Y \to X, \quad (x, y) \mapsto x,$$
$$p_2 : X \times Y \to Y, \quad (x, y) \mapsto y.$$

- Sei $f : X \to Y, x \mapsto f(x)$, eine Funktion. Die *Restriktion von* $f : X \to Y$ *auf* A, geschrieben $f|_A : A \to Y, x \mapsto f(x)$, erfüllt $f|_A(x) = f(x)$ für alle $x \in A$.
- Sei $f : A \to Y, x \mapsto f(x)$, eine Funktion. Eine *Erweiterung von* f *auf* X ist eine Funktion $g : X \to Y$,

$$x \mapsto \begin{cases} f(x), & x \in A, \\ g(x), & x \in X \setminus A, \end{cases}$$

  wobei jeweils $g(x) \in Y$ beliebig ist. Es gilt also $g(x) = f(x)$ für alle $x \in A$. (Erweiterungen sind im allgemeinen nicht eindeutig.)

[1] Umgekehrt kann man eine Funktion $f : X \to Y$ rigoros definieren als Tripel (X, Y, G_f), wenn G_f eine Teilmenge von $X \times Y$ ist, die für alle $x \in X$ genau ein Paar (x, y) enthält.

3.2 BILD UND URBILD

Manchmal ist es nützlich anzugeben, wie eine Funktion auf eine ganze Menge von Elementen wirkt. Dies kann man wie folgt angeben: Sei $f : X \to Y$ eine Funktion (die genaue Zuordnungsvorschrift ist hier nicht von Interesse, und wir nehmen ab sofort immer an, dass X und Y Mengen sind).

- Für $A \subseteq X$ ist das *Bild von* A *unter* f definiert als die Menge

$$f(A) := \{y \in Y : \exists x \in A : f(x) = y\} \subseteq Y.$$

- Für $B \subseteq Y$ ist das *Urbild von* B *unter* f definiert als die Menge

$$f^{-1}(B) := \{x \in X : \exists y \in B : f(x) = y\} \subseteq X.$$

Beispiel 3.2. Sei $X = \{a, b, c, \ldots, z\}$, $Y = \{2, 3, \ldots, 9\}$, und $f : X \to Y$ die „SMS“-Funktion, die jeden Buchstaben auf die entsprechende Telefontaste abbildet (das heisst $f(a) = 2$, $f(z) = 9, \ldots$).

- Für $A = \{s, e, r, v, u, s\}$ ist das Bild $f(A) = \{f(s), f(e), f(r), f(v), f(u), f(s)\} = \{7, 3, 7, 8, 8, 7\} = \{3, 7, 8\}$.
- Für $B = \{7, 8\}$ ist das Urbild $f^{-1}(B) = \{p, q, r, s, t, u, v\}$, da $f(p) = f(q) = f(r) = f(s) = 7$ und $f(t) = f(u) = f(v) = 8$ (und diese Liste vollständig ist).

Beispiel 3.3. Wir betrachten die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$.

- Sei $A = \{x \in \mathbb{R} : 0 < x < 2\}$. Dann ist $f(A) = \{x \in \mathbb{R} : 0 < x < 4\}$ und $f^{-1}(A) = \left\{x \in \mathbb{R} : -\sqrt{2} < x < \sqrt{2}\right\} \setminus \{0\}$.
- Sei $B = \{x \in \mathbb{R} : -2 < x < 0\}$. Dann ist $f(B) = f(A) = \{x \in \mathbb{R} : 0 < x < 4\}$, aber $f^{-1}(B) = \emptyset$.

Beachten Sie: $f(A)$ ist nur eine (leider etwas irreführende, aber weit verbreitete) Kurzschreibweise für eine Menge, und nicht der Wert der Funktion f, angewendet auf eine Menge. Ebenso darf $f^{-1}(B)$ nicht mit der Umkehrfunktion (siehe Abschnitt 3.3) von f verwechselt werden. (Insbesondere kann man immer das Urbild – zur Not als leere Menge – angeben, auch wenn keine Umkehrfunktion existiert.)

Oft wird umgekehrt eine Menge als das Bild einer Funktion charakterisiert; man definiert diese dann meist kurz in der Form $\{f(a) : a \in A\}$ $(= f(A))$. So kann man etwa $\left\{n^2 : n \in \mathbb{N}\right\}$ für die Menge aller Quadratzahlen oder $G_f = \{(x, f(x)) : x \in X\}$ für den Graphen der Funktion $f : X \to Y$ schreiben.

3.3 VERKNÜPFUNGEN UND UMKEHRFUNKTION

Ein wichtiges Konzept ist die Hintereinanderausführung von zwei Funktionen. Seien $f : X \to Y$, $x \mapsto f(x)$, und $g : Y \to Z$, $y \mapsto g(y)$ Funktionen. Dann bezeichnen wir die Funktion

$$g \circ f : X \to Z, \quad x \mapsto g(f(x))$$

(gesprochen „g nach f“) als *Verknüpfung* (oder *Komposition*) *von* f *und* g.

Beachten Sie:

- $g \circ f$ bezeichnet *eine* Funktion.
- Es ist hier wichtig, dass der Wertebereich von f eine Teilmenge des Definitionsbereichs von g ist. Für $f : A \to B$ und $g : C \to D$ mit $B \not\subseteq C$ ist $g \circ f$ eine sinnlose Bezeichnung.
- Selbst wenn $f \circ g$ und $g \circ f$ beide definiert sind, sind sie im allgemeinen nicht gleich.

Beispiel 3.4. Wir betrachten $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ und $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto x + 3$. Dann ist

- $g \circ f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 + 3$,
- $f \circ g : \mathbb{R} \to \mathbb{R}$, $x \mapsto (x + 3)^2 = x^2 + 6x + 9$.

Eine wichtige Frage ist, ob wir eine Funktion f per Verknüpfung „rückgängig machen“ können. Da die Reihenfolge der Komposition eine Rolle spielt, müssen wir deshalb unterscheiden, ob wir f zuerst oder zuletzt anwenden.

Seien $f : X \to Y$ und $g : Y \to X$ Funktionen. Dann ist

- g eine *Links-Inverse zu* f, falls $g \circ f = \mathrm{id}_X$ gilt (falls also $g(f(x)) = x$ für alle $x \in X$ gilt), und
- g eine *Rechts-Inverse zu* f, falls $f \circ g = \mathrm{id}_Y$ gilt (falls also $f(g(y)) = y$ für alle $y \in Y$ gilt),
- g eine *Inverse zu* f, falls g sowohl Links-Inverse als auch Rechts-Inverse zu f ist.

Oft sieht man auch den Begriff *Umkehrfunktion* für eine Inverse, und schreibt f^{-1}. (Beachten Sie den Unterschied zwischen dem Funktionswert der Umkehrfunktion $f^{-1}(x)$ und dem Kehrwert des Funktionswerts $f(x)^{-1}$.) Hat eine Funktion eine Umkehrfunktion, so nennt man sie *invertierbar*.

Beispiel 3.5.

- Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto 2x + 3$ hat die Inverse $g : \mathbb{R} \to \mathbb{R}$, $x \mapsto \frac{1}{2}(x - 3)$.

- Die *Nullfunktion* $f : \mathbb{R} \to \mathbb{R}, x \mapsto 0$, hat weder Links- noch Rechts-Inverse: Für jede Funktion $g : \mathbb{R} \to \mathbb{R}$ gilt $(g \circ f)(x) = g(0)$ für alle $x \in \mathbb{R}$. Und da 0 durch g nur auf ein einziges Element y abgebildet werden kann, ist $(g \circ f)(x) \neq x$ für $x \neq y$. Damit kann keine Funktion g eine Links-Inverse sein. Umgekehrt ist $(f \circ g)(x) = 0$ für alle $x \in \mathbb{R}$, also gibt es keine Funktion g, für die zum Beispiel $(f \circ g)(1) = 1$ gilt. Wir können also auch keine Rechts-Inverse finden.

3.4 INJEKTIV, SURJEKTIV, BIJEKTIV

Wir führen nun bequeme Kriterien ein, um zu entscheiden, ob eine Funktion Links- oder Rechts-Inverse besitzt. Dafür betrachten wir noch einmal das letzte Beispiel, und überlegen uns, woran das Finden einer Links- und einer Rechts-Inversen gescheitert ist.

Im ersten Fall hatten wir das Problem, dass alle Elemente x durch f auf das selbe Element abgebildet wurden – diese Zuordnung lässt sich aber nicht rückgängig machen: Wenn wir nur den Wert $y = f(x)$ gegeben haben, können wir nicht entscheiden, *welches* Element x nun ursprünglich vorlag. Im zweiten Fall liegt bei genauer Betrachtung ein anderes Problem vor: Es gibt Elemente $y \in \mathbb{R}$, die wir durch f gar nicht erreichen können; für die also $(f \circ g)(y) = y$ unmöglich ist.

Dies motiviert die folgende Definition: Sei $f : X \to Y$ eine Funktion. Dann nennen wir f

- *injektiv*, falls für alle $x, y \in X$ gilt: $f(x) = f(y) \Rightarrow x = y$ (oder, äquivalent durch Kontraposition, $x \neq y \Rightarrow f(x) \neq f(y)$),
- *surjektiv*, falls für alle $y \in Y$ gilt: es gibt ein $x \in X$, so dass $f(x) = y$ ist,
- *bijektiv*, falls f injektiv und surjektiv ist.

Beispiel 3.6. Sei $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geqslant 0\}$ die Menge der nicht-negativen reellen Zahlen. In diesem Beispiel verwenden wir die Tatsache, dass für $x \in \mathbb{R}_+$ nach Definition $\sqrt{x}$ diejenige Zahl größer oder gleich Null ist, deren Quadrat x ist, und dass für alle $x \in \mathbb{R}_+$ genau eine solche Zahl existiert.

- $f_1 : \mathbb{R}_+ \to \mathbb{R}_+, x \mapsto x^2$ ist injektiv und surjektiv (und daher bijektiv): Sei $b \geqslant 0$ beliebig, dann ist $a := \sqrt{b} \geqslant 0$, und $f_1(a) = \sqrt{b}^2 = b$. Also ist f_1 surjektiv.

  Seien nun $a, b \geqslant 0$ mit $f_1(a) = f_1(b)$, also $a^2 = b^2$. Dann ist auch $a = \sqrt{a^2} = \sqrt{b^2} = b$ (da $a, b \geqslant 0$), und damit ist f_1 injektiv.
- $f_2 : \mathbb{R} \to \mathbb{R}_+, x \mapsto x^2$ ist surjektiv, aber nicht injektiv: $f_2(-2) = 4 = f_2(2)$. Die Surjektivität folgt aus der gleichen Argumentation wie für f_1.
- $f_3 : \mathbb{R}_+ \to \mathbb{R}, x \mapsto x^2$ ist injektiv, aber nicht surjektiv: Zum Beispiel existiert für $b = -1$ kein $a \in \mathbb{R}$, so dass $a^2 = b$ gilt. Die Injektivität folgt aus der gleichen Argumentation wie für f_1.

- $f_4 : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ ist weder injektiv (aus dem gleichen Grund wie f_2) noch surjektiv (aus dem gleichen Grund wie f_3).

Hier wird noch einmal deutlich, warum Definitionsbereich und Wertebereich zu der Funktionsdefinition dazugehören müssen.

Es gilt (wie wir später beweisen werden):

- f hat eine Links-Inverse, genau dann, wenn f injektiv ist,
- f hat eine Rechts-Inverse, genau dann, wenn f surjektiv ist,
- f hat eine Inverse, genau dann, wenn f bijektiv ist.

Manchmal wird auch für eine lediglich injektive Funktion $f : X \to Y$ eine Umkehrfunktion gebraucht. In diesem Fall definiert man f^{-1} als die Inverse der Funktion $\tilde{f} : X \to f(X)$, $x \mapsto f(x)$ (die dann nach Konstruktion surjektiv ist.)

RELATIONEN

4

Mit Hilfe von Funktionen lassen sich Beziehungen zwischen Objekten ausdrücken. Allerdings hat der Funktionsbegriff wichtige Einschränkungen: eine Funktion $f : X \to Y$ verknüpft jedes Objekt x aus dem Definitionsbereich X immer nur mit jeweils einem Objekt $f(x)$ aus der Zielmenge Y. Ausserdem muss für jedes x aus X zwingend eine Zuordnung festgelegt werden. Will man allgemeinere Beziehungen ausdrücken, braucht man auch einen allgemeineren Begriff: den der Relation.

4.1 DEFINITION, BEISPIELE, EIGENSCHAFTEN

Seien M und N Mengen. Eine *Relation von* M *auf* N ist eine Teilmenge $R \subseteq M \times N$. Falls für zwei Elemente $a \in M$ und $b \in N$ gilt, dass $(a, b) \in R$ ist, sagen wir „a steht in Relation R zu b“ und schreiben kurz $a\,R\,b$.[1] In den meisten Fällen ist $M = N$; man nennt dann R kurz eine *Relation auf M*.

Beispiel 4.1. Die folgenden Definitionen von $R \subseteq M \times M$ sind jeweils eine Relation:

1. Sei M die Menge aller Studierenden der Grundbegriffe der Mathematik. Für $a, b \in M$ soll $a\,R\,b$ gelten, falls die Person a die Person b namentlich kennt.
2. Sei $M = \mathbb{N}$ und $(a, b) \in R$, falls a und b den selben Rest bei Division durch 5 lassen. (Zum Beispiel steht 5 in Relation zu 10, und auch 7 zu 12.) Dies definiert eine Relation auf $\mathbb{N}$, nämlich eine *Kongruenzrelation (modulo* 5). Man schreibt für $a\,R\,b$ in diesem Fall häufig $a \equiv b \mod 5$.
3. Sei $M = \mathbb{R}$ und $(a, b) \in R$ falls $a \leqslant b$ gilt (und analog für $a \geqslant b$).
4. Sei $M = \mathcal{P}(X)$ für eine beliebige Menge X und $(A, B) \in R$, falls $A \subseteq B$ gilt (und analog für $A \supseteq B$).
5. Sei $M = \mathbb{R}$ und $(a, b) \in R$ falls $a < b$ gilt (und analog für $a > b$).

[1] Funktionen können also als Spezialfall von Relationen definiert werden.

Der Begriff der Relation ist sehr allgemein. Anhand der folgenden Eigenschaften kann man aber wichtige Typen von Relationen unterscheiden. Sei M eine Menge und R eine Relation auf M. Dann nennen wir R:

- *reflexiv*, falls für alle $x \in M$ gilt: $x \mathrel{R} x$,
- *transitiv*, falls für alle $x, y, z \in M$ gilt: aus $x \mathrel{R} y$ und $y \mathrel{R} z$ folgt $x \mathrel{R} z$,
- *symmetrisch*, falls für alle $x, y \in M$ gilt: aus $x \mathrel{R} y$ folgt $y \mathrel{R} x$
- *antisymmetrisch*, falls für alle $x, y \in M$ gilt: aus $x \mathrel{R} y$ und $y \mathrel{R} x$ folgt $x = y$.

Beispiel 1 ist reflexiv, aber weder transitiv noch symmetrisch oder antisymmetrisch. Beispiel 2, 3 und 4 sind reflexiv und transitiv; Beispiel 2 ist symmetrisch, 3 und 4 sind antisymmetrisch. Beispiel 5 ist transitiv und antisymmetrisch, aber nicht reflexiv.

4.2 ORDNUNGSRELATIONEN

Eine Relation R auf einer Menge M, die reflexiv, transitiv und antisymmetrisch ist, nennt man *Ordnungsrelation* (oder *Halbordnung*). Für $(a, b) \in R$ schreibt man in diesem Fall häufig auch $a \preceq b$ oder, äquivalent, $b \succeq a$. Zwei Elemente $a, b \in M$ heissen *vergleichbar* bezüglich $\preceq$, wenn $a \preceq b$ oder $b \preceq a$ gilt. Sind alle Elemente in M vergleichbar (gilt also für alle $a, b \in M$, dass a und b vergleichbar sind), so nennt man $\preceq$ *Totalordnung* (manchmal auch einfach *Ordnung*).

Beispiel 4.2.

1. $\leqslant$ ist eine Totalordnung auf $\mathbb{R}$.
2. Für eine gegebene Menge M ist $\subseteq$ eine Ordnungsrelation auf $\mathcal{P}(M)$, aber im allgemeinen keine Totalordnung: Für $M = \{1, 2, 3\}$ und $\{1\}, \{2\} \in \mathcal{P}(M)$ gilt weder $\{1\} \subseteq \{2\}$ noch $\{2\} \subseteq \{1\}$.
3. Auf der Menge der Wörter ist die lexikographische Ordnung eine Totalordnung.
4. $<$ ist *keine* Ordnungsrelation auf $\mathbb{R}$, da $<$ nicht reflexiv ist.
5. Auf $\mathbb{N} \times \mathbb{N}$ sei $(n_1, n_2) \preceq (m_1, m_2)$ genau dann, wenn $n_1 \leqslant m_1$ und $n_2 \leqslant m_2$ gilt. Dann definiert dies eine Ordnungsrelation, die aber keine Totalordnung ist.

Hat man eine Ordnung, kann man nach dem größten (und kleinsten Element) bezüglich dieser Ordnung fragen. Sei $\preceq$ eine Ordnungsrelation auf der Menge M.

- Ein $x \in M$, für das für alle $y \in M$ gilt $x \preceq y$, heisst *kleinstes Element* oder *Minimum* von M.
- Ein $x \in M$, für das für alle $y \in M$ gilt $x \succeq y$, heisst *größtes Element* oder *Maximum* von M.

- Ein $x \in M$, für das für alle $y \in M$ aus $y \preceq x$ folgt, dass $y = x$ ist, heisst *minimales Element* von M.
- Ein $x \in M$, für das für alle $y \in M$ aus $y \succeq x$ folgt, dass $y = x$ ist, heisst *maximales Element* von M.

Beachten Sie den Unterschied zwischen Minimum und minimalem Element: Die Menge $\{(2,2),(3,3),(1,5)\} \subseteq \mathbb{N} \times \mathbb{N}$ hat bezüglich der Ordnung im Beispiel 4.2.5 die zwei minimalen Elemente $(1,5)$ und $(2,2)$, aber kein Minimum. Ein Minimum muss kleiner als alle anderen Elemente sein, während es für ein minimales Element reicht, dass keine kleineren Elemente existieren. Wir werden später zeigen, dass wenn eine Menge ein Minimum hat, dieses eindeutig sein muss (ebenso ein Maximum).

Eine Menge muss also nicht unbedingt ein Maximum oder ein Minimum enthalten. Wir können uns aber fragen, ob ein geeignetes Element nicht in einer grösseren Menge gefunden werden kann. Das wollen wir präzisieren:

Sei $\preceq$ eine Ordnung auf der Menge M und $A \subseteq M$ eine Teilmenge. Dann heisst $x \in M$

- eine *obere Schranke* von A, wenn für alle $y \in A$ gilt $x \succeq y$,
- eine *untere Schranke* von A, wenn für alle $y \in A$ gilt $x \preceq y$.

Besitzt A eine obere *und* eine untere Schranke, so nennt man A *beschränkt*, ansonsten *unbeschränkt*.

Beispiel 4.3. Wir betrachten die übliche Ordnung $\leqslant$ auf $\mathbb{R}$. Dann ist

- $\{x \in \mathbb{R} : 0 < x < 1\} \subseteq \mathbb{R}$ beschränkt, da 0 eine untere Schranke und 1 eine obere Schranke ist. Ebenso sind auch $2, 3, 4, \ldots$ obere und $-1, -2, -3, \ldots$ untere Schranken.
- $\{x \in \mathbb{R} : 0 < x\} \subseteq \mathbb{R}$ unbeschränkt, da zwar 0 eine untere Schranke ist, aber keine obere Schranke in $\mathbb{R}$ existiert.

Schranken sind also in der Regel nicht eindeutig. Allerdings können wir die Menge aller oberen (oder unteren) Schranken bilden, und wiederum fragen, ob diese ein kleinstes (bzw. größtes) Element besitzt: Sei wieder $\preceq$ eine Ordnung auf der Menge M, und $A \subseteq M$ eine Teilmenge.

- Ist $y \in M$ das Minimum der Menge $\{x \in M : x \text{ ist obere Schranke von } A\}$, so heisst y das *Supremum* von A.
- Ist $y \in M$ das Maximum der Menge $\{x \in M : x \text{ ist untere Schranke von } A\}$, so heisst y das *Infimum* von A.

Im Gegensatz zu Minima und Maxima müssen Infima und Suprema also nicht in A liegen. Die Menge $M = \{x \in \mathbb{R} : 0 < x < 1\}$ hat zum Beispiel in $\mathbb{R}$ das Infimum 0 und das Supremum 1. Ebenso hat die Menge $N = \{x \in \mathbb{R} : 0 \leqslant x < 1\}$ in $\mathbb{R}$ Supremum 1 und Infimum 0, aber hier ist $0 \in N$, also gleichzeitig ein Minimum. Hingegen besitzt N kein Maximum.

4.3 ÄQUIVALENZRELATIONEN

Eine Relation R auf einer Menge M, die reflexiv, transitiv und symmetrisch ist, nennt man *Äquivalenzrelation*. Für $(a, b) \in R$ schreibt man in diesem Fall häufig auch $a \sim b$ (gesprochen „a ist äquivalent zu b“).

Beispiel 4.4. Die folgenden Relationen $\sim$ sind Äquivalenzrelationen auf M:

1. Sei M eine beliebige Menge und $a \sim b$ genau dann, wenn $a = b$ ist.
2. Für $M = \mathbb{N}$ und $c \in \mathbb{N} \setminus \{0\}$ ist die Kongruenzrelation modulo c eine Äquivalenzrelation.
3. Sei $M = \mathbb{R}$ und $x \sim y$ genau dann, wenn ein $n \in \mathbb{Z}$ existiert, so dass $x = y + n2\pi$ gilt.
4. Sei $M = \mathbb{N} \times \mathbb{N}$ und $(n, m) \sim (n', m')$ genau dann, wenn $n + m' = n' + m$ gilt.

Mit Hilfe von Äquivalenzrelationen kann man also ausdrücken, dass zwei Elemente sich in gewisser (aber nicht unbedingt jeder) Hinsicht identisch verhalten. Es liegt daher nahe, solche Elemente zusammenzufassen (man spricht auch von *identifizieren*): Sei $\sim$ eine Äquivalenzrelation auf der Menge M. Für $a \in M$ heisst die Menge

$$[a]_\sim := \{x \in M : x \sim a\}$$

die *Äquivalenzklasse* von a unter $\sim$. Jedes $x \in [a]_\sim$ heisst *Repräsentant* von $[a]_\sim$. Die Menge aller Äquivalenzklassen

$$M/_\sim := \{[a]_\sim : a \in M\}$$

nennt man die *Quotientenmenge* von M unter $\sim$.

Beispiel 4.5.

1. Sei M eine beliebige Menge und $a \sim b$ genau dann, wenn $a = b$ ist. Dann ist die Äquivalenzklasse von $x \in M$ die Menge $[x]_\sim = \{x\}$, und die Quotientenmenge von M ist die Menge aller einelementigen Teilmengen: $M/_\sim = \{\{x\} : x \in M\}$.
2. Sei $M = \mathbb{N} \cup \{0\}$ und $\sim$ die Kongruenzrelation modulo 3. Dann sind die Äquivalenzklassen $[1]_\sim = \{1, 4, 7, 10, \dots\}$, $[2]_\sim = \{2, 5, 8, 11, \dots\}$, $[3]_\sim = \{0, 3, 6, 9, \dots\}$ und $[4]_\sim = \{1, 4, 7, 10, \dots\} = [1]_\sim$. Die Quotientenmenge ist also $M/_\sim = \{[1]_\sim, [2]_\sim, [3]_\sim\}$. (Im Zusammenhang mit Kongruenzrelationen spricht man auch oft von *Restklassen* und *Restklassenmengen*.)

3. Wir können ebene Winkel (gemessen im Bogenmaß) auffassen als Äquivalenzklassen unter der Relation in Beispiel 4.4.3. Dann wäre der rechte Winkel definiert als die Äquivalenzklasse $[\frac{\pi}{2}]_\sim$, und $\mathbb{R}/_\sim$ wäre die Menge der Winkel.
4. Mit Hilfe der Relation in Beispiel 4.4.4 kann man die ganzen Zahlen als Äquivalenzklassen unter $\sim$ definieren (was wir in einem späteren Kapitel behandeln werden).

Teil II

EINFÜHRUNG IN DAS MATHEMATISCHE ARBEITEN

ÜBERBLICK

Wir haben jetzt die wichtigsten Grundbegriffe kennengelernt, um mathematische Sachverhalte ausdrücken zu können: mit Hilfe logischer Verknüpfungen und der Mengenlehre sind prinzipiell alle mathematischen Aussagen formulierbar. Im folgenden Teil werden wir nun lernen, solche Aussagen zu *beweisen*, das heisst, ihren Wahrheitsgehalt zu entscheiden, und diese Entscheidung so darzulegen, dass sie von keinem hinreichend mathematisch Gebildeten angezweifelt werden kann. (Einige Beweise haben wir – quasi als Vorgeschmack – in dieser Veranstaltung bereits geführt.)

Insbesondere sollen Sie lernen, für typische mathematische Aussagen Beweise

- zu finden und
- aufzuschreiben.

Dies sind zwei sehr unterschiedliche Tätigkeiten, die beide gleichermaßen Übung verlangen.

LOGISCHE BAUSTEINE VON BEWEISEN

5

Die Mathematik ist eine *deduktive* Wissenschaft: Ausgehend von einmal definierten Grundbegriffen (zum Beispiel der Menge mitsamt der Elementbeziehung) und festgelegten Eigenschaften oder Beziehungen (die wir ohne weitere Begründung als gültig annehmen[1]) werden durch Anwendung logischer Schlussregeln weitere Behauptungen abgeleitet oder *bewiesen*. Dies geschieht dadurch, dass die zu beweisende Aussage als letztes Glied einer Kette von Aussagen steht, von denen jede entweder bereits als wahr erkannt (durch Festlegung oder Beweis) oder ihr Wahrheitswert mit Hilfe von Tautologien auf den Wahrheitswert einer früheren Aussage in der Kette zurückgeführt werden kann. Solche Tautologien werden in diesem Kontext *Schlussregeln* genannt. Die gesamte Kette ist ein *Beweis* für die letzte Aussage, die dann auch *Satz* oder *Theorem* genannt wird.[2]

Wir möchten nun einige wichtige Schlussregeln anführen und beispielhaft anwenden (der Nachweis, etwa per Wahrheitstafel, dass es sich um Tautologien handelt, bleibt dem Leser überlassen). Natürlich müssen Sie diese Regeln (und insbesondere ihre Namen) nicht auswendig lernen; ihre Verwendung ist für Sie sicher bereits selbstverständlich (oder wird es Ihnen rasch sein). Auch werden einzelne Beweisschritte in der Regel nie in dieser Ausführlichkeit angegeben. Trotzdem ist es wichtig, sich einmal explizit klarzumachen, aus welchen Einzelschritten ein mathematischer Beweis aufgebaut ist (und insbesondere, welche sprachlichen Formulierungen für welche Schlussregeln stehen).

- *Einsetzen von Definitionen*: Jede mathematische Definition ist eine logische Äquivalenz; der Wahrheitswert einer Aussage ändert sich nicht, wenn ein Begriff durch seine Definition ersetzt wird.

[1] Aussagen, die durch Konvention als wahr festgelegt werden, nennt man *Axiome*. *Welche* Aussagen als wahr festgelegt werden, ist dabei erstmal willkürlich. Die allgemein akzeptierten Axiome beruhen auf dem Wunsch, mit möglichst wenig Axiomen möglichst viele „sinnvolle“ Aussagen abzuleiten, ohne auf Widersprüche zu kommen, und haben sich in der Praxis (innerhalb der Mathematik und in ihrer Anwendung) als ausgesprochen nützlich erwiesen.

[2] Es ist ein wesentlicher Erfolg der Mathematik des letzten Jahrhunderts, dass *bewiesen* wurde, dass die so ableitbaren Aussagen genau die wahren Aussagen (wie etwa die durch Wahrheitstafeln festgestellten) sind. Dies gelang Kurt Gödel in seiner Doktorarbeit, und wird als *Vollständigkeitssatz* bezeichnet.

Beispiel: Eine gegebene Funktion $f : X \to Y$ ist surjektiv. Also existiert für alle $y \in Y$ ein $x \in X$ mit $f(x) = y$.

- *modus ponens*: $(p \wedge (p \Rightarrow q)) \Rightarrow q$.

  Beispiel: Eine gegebene Funktion f ist differenzierbar, und wenn f differenzierbar ist, ist f stetig. Also ist f stetig.

- *modus tollens*: $((\neg q) \wedge (p \Rightarrow q)) \Rightarrow (\neg p)$.

  Beispiel: Eine gegebene Funktion f ist nicht stetig, und wenn f differenzierbar ist, ist f stetig. Also ist f nicht differenzierbar.

- *Verkettung*: $((p \Rightarrow q) \wedge (q \Rightarrow r)) \Rightarrow (p \Rightarrow r)$.

  Beispiel: Wenn eine Funktion zweimal differenzierbar ist, dann ist sie differenzierbar, und wenn sie differenzierbar ist, dann ist sie stetig. Also ist eine Funktion stetig, wenn sie zweimal differenzierbar ist.

- *Modus tollendo ponens*: $(\neg p \wedge (q \vee p)) \Rightarrow q$.

  Beispiel: Es gilt $xy = 0$ (also $x = 0$ oder $y = 0$), und $x = 2$. Also ist $y = 0$.

- *Fallunterscheidung*: $((p \Rightarrow q) \wedge (\neg p \Rightarrow q)) \Rightarrow q$.

  Beispiel: Wenn n gerade ist, dann ist $n^2 + n$ gerade, und wenn n ungerade ist, dann ist $n^2 + n$ gerade. Also ist $n^2 + n$ gerade.

- *Reductio ad absurdum*: $((p \Rightarrow q) \wedge (p \Rightarrow \neg q)) \Rightarrow \neg p$.

  Diese Schlussregel ist der Kern des *Beweis durch Widerspruch*, der im nächsten Abschnitt ausführlich besprochen wird.

- Natürlich sind auch die in Abschnitt 1.3 besprochenen Tautologien (doppelte Negation, de Morganschen Gesetze, Distributivgesetze) gültige Schlussregeln.

Für den Umgang mit Quantorenaussagen gelten zusätzliche Schlussregeln. Hier geht es weniger darum, *dass* diese Regeln gelten (dies steckt eigentlich bereits in der Definition der Quantoren), sondern darum, *wie* sie in einem Beweis verwendet werden. Wir betrachten eine Aussageform $P(x)$ mit freier Variable x. Die ersten beiden Schlussregeln betreffen die *Verwendung* von gültigen Quantorenaussagen in Beweisen.

- *Universelle Spezialisierung*: Wenn $a \in X$ ein beliebiges Element von X ist, und wir wissen, dass $(\forall x \in X : P(x))$ gilt, so dürfen wir verwenden, dass $P(a)$ gilt.

  Beispiel: Wir betrachten eine differenzierbare Funktion f, und wissen, dass alle differenzierbaren Funktionen stetig sind. Also können wir folgern, dass f stetig ist.

- *Existenzielle Spezialisierung*: Wenn wir wissen, dass $(\exists x \in X : P(x))$ gilt, dürfen wir ein $b \in X$ einführen, für das $P(b)$ gilt. (Der Buchstabe b darf vorher nicht bereits anderweitig definiert worden sein.)

  Diese Schlussregel wird verwendet, um ein Element mit der durch $P(x)$ gegebenen Eigenschaft aus X auszuwählen und im weiteren Verlauf des Beweises zu verwenden.

  Beispiel: Wir wissen, dass das Polynom $p(x)$ eine Nullstelle hat. Diese bezeichnen wir mit z, und untersuchen dann, ob z reell ist, und falls ja, ob z positiv oder negativ ist.

Die nächsten beiden Schlussregeln legen fest, wann man Quantorenaussagen als bewiesen ansehen darf. Wir werden in Kapitel 6.5 ausführlicher auf ihre Anwendung eingehen.

- *Universelle Generalisierung*: Wenn wir wissen, dass für ein *beliebiges* Element $a \in X$ gilt, dass $P(a)$ wahr ist, so können wir folgern, dass $(\forall x \in X : P(x))$ gilt.

  Wichtig ist hier, dass $a \in X$ beliebig ist in dem Sinne, dass wir statt a jedes andere Element aus X einsetzen könnten, ohne dass sich an der Gültigkeit der Begründung, warum $P(a)$ gilt, etwas ändern würde.

  Beispiel: Sei n eine beliebige natürliche Zahl. Dann ist (wie wir später zeigen werden) $n^2 + n$ gerade. Also gilt für alle $n \in \mathbb{N} : n^2 + n$ ist gerade.

- *Existenzielle Generalisierung*: Wenn wir wissen, dass für ein *bestimmtes* $b \in X$ gilt, dass $P(b)$ wahr ist, so können wir folgern, dass $(\exists x \in X : P(x))$ gilt.

  Unter einem „bestimmten“ Element ist hier ein (mehr oder weniger) konkret angegebes Element gemeint.

  Beispiel: Für $2 \in \mathbb{N}$ ist $2^2 = 4$, also existiert ein $n \in \mathbb{N}$, so dass $n^2 = 4$ gilt.

In einem mathematischen Beweis werden Sie also immer wieder auf diese drei Bausteine stossen:

- Einsetzen von Definitionen,
- Anwenden bereits bewiesener Resultate, und deren
- Verknüpfung mit Hilfe logischer Schlussregeln.

Die Formulierung eines Beweises geschieht dabei üblicherweise nicht in der formal-logischen Schreibweise, sondern in natürlich-sprachigen Sätzen, wie es schon mehrmals in den zurückliegenden Abschnitten gezeigt wurde. Wichtig ist nur, dass dem Leser jederzeit klar ist, dass und vor allem wie die umgangssprachliche Argumentation formalisiert werden *kann*.[3]

[3] Dies dient vor allem der Verständlichkeit, aber auch der Durchführbarkeit (wenn auch auf Kosten der leichten Erkennbarkeit von Fehlern). Der letzte Versuch, Mathematik von Grund auf und streng formal darzustellen, war die *Principia Mathematica* von Russell und Whitehead (1910–1913). Der Beweis der „gelegentlich nützlichen Tatsache“, dass $1 + 1 = 2$ gilt, wird – nach einiger Vorarbeit – auf Seite 86 des zweiten Bandes abgeschlossen. Das Projekt wurde nach dem dritten Band abgebrochen.

Beispielhaft schauen wir uns noch einmal den Beweis an, dass zwei leere Mengen $\emptyset$ und $\{\}$ die gleiche Menge sind, und zerlegen ihn in seine logische Bausteine. Seien M und N Mengen und $L(N)$ die Aussageform „N ist eine leere Menge“. Die bereits bekannte Tatsache, dass eine leere Menge Teilmenge jeder Menge ist, können wir dann formalisieren als $\forall N : \forall M : (L(N) \Rightarrow N \subseteq M)$.[4]

1. *Sind $\emptyset$ und $\{\}$ beide leere Mengen (d.h. enthalten keine Elemente),*

   Als Anfangspunkt des Argumentes werden die Aussagen $L(\emptyset)$ und $L(\{\})$ als wahr angenommen (wir führen also einen direkten Beweis, siehe Kapitel 6.1).

2. *so gilt einerseits $\emptyset \subseteq \{\}$ (eine leere Menge ist Teilmenge jeder Menge, wozu auch leere Mengen zählen)*

   Universelle Spezialisierung: In der Aussage $\forall N : \forall M : (L(N) \Rightarrow N \subseteq M)$ (bereits bewiesene Tatsache) setzen wir für die freie Variable N das spezielle Element $\emptyset$ ein. Also gilt: $\forall M : (L(\emptyset) \Rightarrow \emptyset \subseteq M)$.

   Universelle Spezialisierung: In der Aussage $\forall M : (L(\emptyset) \Rightarrow \emptyset \subseteq M)$ (bereits bewiesene Tatsache) setzen wir für die freie Variable M das spezielle Element $\{\}$ ein. Also gilt: $L(\emptyset) \Rightarrow \emptyset \subseteq \{\}$.

   Modus ponens: $L(\emptyset)$ (erstes Element in der Kette) und $L(\emptyset) \Rightarrow \emptyset \subseteq \{\}$ (vorheriges Element der Kette) impliziert $\emptyset \subseteq \{\}$.

3. *und umgekehrt aus dem gleichen Grund $\{\} \subseteq \emptyset$.*

   Universelle Spezialisierung: In der Aussage $\forall N : \forall M : (L(N) \Rightarrow N \subseteq M)$ (bereits bewiesene Tatsache) setzen wir für die freie Variable N das spezielle Element $\{\}$ ein. Also gilt: $\forall M : (L(\emptyset) \Rightarrow \emptyset \subseteq M)$.

   Universelle Spezialisierung: In der Aussage $\forall M : (L(\{\}) \Rightarrow \{\} \subseteq M)$ (bereits bewiesene Tatsache) setzen wir für die freie Variable M das spezielle Element $\emptyset$ ein. Also gilt: $L(\{\}) \Rightarrow \{\} \subseteq \emptyset$.

   Modus ponens: $L(\{\})$ (erstes Element in der Kette) und $L(\{\}) \Rightarrow \{\} \subseteq \emptyset$ (vorheriges Element der Kette) impliziert $\{\} \subseteq \emptyset$.

4. *Dies ist aber genau die Definition der Gleichheit; daher gilt $\emptyset = \{\}$.*

   Einsetzen der Definition: Die Aussage $\emptyset \subseteq \{\} \wedge \{\} \subseteq \emptyset$ ist logisch äquivalent zu $\emptyset = \{\}$. Dies war der zu zeigende Satz.

Zum Abschluss soll noch auf einige falsche, aber leider in Lösungsversuchen immer wieder anzutreffende „Schlussregeln“ hingewiesen werden.

[4] Um die Übersichtlichkeit nicht noch mehr zu behindern, geben wir die Menge X, in der M und N liegen muss und über die quantifiziert wird, nicht explizit an.

- *Verwechslung von Implikation und Äquivalenz*: q und $p \Rightarrow q$ impliziert *nicht* p!

  Falsches Beispiel: Wenn eine differenzierbare Funktion f an der Stelle x ein Minimum hat, so ist dort die Ableitung $f'(x) = 0$. Es gilt $f'(x) = 0$, also hat f in x ein Minimum. (Tatsächlich kann f in x auch ein Maximum haben.)

- *Falsche Kontraposition*: $\neg p$ und $p \Rightarrow q$ impliziert *nicht* $\neg q$!

  Falsches Beispiel: Eine bijektive Funktion ist surjektiv, und f ist nicht bijektiv. Also ist f nicht surjektiv. (Es kann auch sein, dass f surjektiv, aber nicht injektiv ist.)

- *Fehlende Prämisse*: $p \Rightarrow q$ impliziert *nicht* q!

  Falsches Beispiel: Eine bijektive Funktion ist surjektiv, und f ist injektiv. Also ist f surjektiv. (Die Injektivität von f tut hier nichts zur Sache; ohne die *Bijektivität* (d.h. Wahrheit von p) können wir aus $p \Rightarrow q$ nichts folgern.)

Lassen Sie sich nicht von der Offensichtlichkeit dieser Fehlschlüsse in diesen (bewusst einfachen und konstruierten) Beispielen täuschen: Auch nur ein solcher Fehlschluss in einer langen und komplizierten Argumentation hat schon viele Beweise vereitelt – insbesondere, da sie in solchem Umfeld deutlich schwerer zu erkennen sind. Gerade deshalb ist es wichtig, die logische Struktur hinter der mathematischen Argumentation in Beweisen zu erkennen.

MATHEMATISCHE BEWEISSTRATEGIEN

6

Im letzten Kapitel haben wir uns damit beschäftigt, aus welchen Bausteinen ein Beweis aufgebaut ist. Nun wenden wir uns der Frage zu, wie man einen Beweis im Ganzen konstruiert. Von besonderem Interesse wird dabei sein, wie man Aussagen der Form $p \Rightarrow q$ beweist, da die meisten mathematischen Sätze, die bewiesen werden müssen, in dieser Form vorliegen. Diesem Thema sind die ersten vier Abschnitte gewidmet. Der letzte Abschnitt behandelt die zweite häufige Form, nämlich Existenz- und Allaussagen.

Eines zu Beginn: Es gibt keine Patentrezepte, die man auswendig lernen und je nach Fall abarbeiten kann. Mathematik ist (auch) eine kreative Betätigung; für viele macht genau dieser schöpferische Umgang mit den mathematischen Objekten, die doch strikten Regeln gehorchen, den eigentlichen Reiz der Mathematik aus. Sinn dieses Kapitels ist vielmehr, darzustellen, wie die Struktur der zu beweisenden Aussage die Struktur ihres Beweises beeinflusst.

6.1 DIREKTER BEWEIS

Die meisten mathematischen Theoreme haben die logische Struktur $p \Rightarrow q$: explizit in der Form „wenn p gilt, dann q“, oder etwas versteckter als „Sei p. Dann q.“

Wie beweist man eine solche Implikation, ohne sie in komplizierte Tautologien zu verpacken? In einem *direkten Beweis* nimmt man an, dass p gilt, und leitet daraus (und aus bereits als wahr erkannten Aussagen) mit Hilfe logischer Schlussregeln die Aussage q ab. Um einzusehen, dass solch eine logische Argumentation genügt, betrachten wir wieder die Wahrheitstafel der Implikation: Ist p falsch, ist $p \Rightarrow q$ immer wahr, unabhängig von der Wahrheit von q. Der einzige interessante Fall ist, wenn p wahr ist: dann muss auch q wahr sein, damit die Implikation gilt. Es genügt also nachweisen, dass aus der Wahrheit von p zwingend die Wahrheit von q folgt: Steht als erstes Glied in einer korrekten Beweiskette p, und als letztes Glied q, dann können wir $p \Rightarrow q$ als bewiesen akzeptieren.

Ein direkter Beweis von $p \Rightarrow q$ fängt also immer (sinngemäß) mit den Worten „Sei p“ (d.h. „als wahr gegeben“) und endet mit dem Worten „Also gilt q.“ Als ersten Schritt sollten Sie sich dabei überlegen, welche Aussagen als Voraussetzungen gegeben sind. Ein Beispiel:

Satz 6.1. *Für* $a, b \in \mathbb{R}$ *mit* $0 < a < b$ *gilt* $a^2 < b^2$.

Die Voraussetzungen sind, dass $a > 0$, $b > 0$ und $b > a$ gelten. (Und ausserdem, dass a und b reelle Zahlen sind, also die üblichen Rechenregeln gelten.) Daraus sollen wir mit Hilfe logischer Schlussregeln ableiten, dass $a^2 < b^2$ gilt.

Beweis. Es seien $a, b \in \mathbb{R}$ und es gelte $0 < a < b$. Da $a > 0$ gilt, können wir $a < b$ auf beiden Seiten mit a multiplizieren und erhalten die Ungleichung $a^2 < ab$. Genauso folgt aus $a < b$ und $b > 0$, dass $ab < b^2$ gilt. Aus der Transitivität der „kleiner"-Relation folgt nun $a^2 < b^2$, was zu zeigen war. □

Beachten Sie, dass jede der Voraussetzungen im Beweis auch verwendet wurde. (Tatsächlich ist die Aussage falsch, wenn jeweils eine davon weggelassen wird.) In Übungs- und Klausuraufgaben können Sie davon ausgehen, dass Sie die gegebenen Voraussetzungen verwenden müssen, um einen richtigen Beweis zu finden.

Wir betrachten noch ein Beispiel. Dafür geben wir zuerst eine präzise Definition der Teilbarkeit:

Definition 6.2. Seien $a, b \in \mathbb{N}$. Wir sagen a *teilt* b (kurz: $a \mid b$), wenn ein $q \in \mathbb{N}$ existiert, so dass $aq = b$ gilt.

Satz 6.3. *Seien* $a, b, c \in \mathbb{N}$. *Wenn gilt:* a *teilt* b *und* b *teilt* c*, dann gilt auch:* a *teilt* c.

Hier ist also zu zeigen, dass die Implikation „$(a \mid b \wedge b \mid c) \Rightarrow a \mid c$" gilt.[1] Wir dürfen also $a \mid b$ und $b \mid c$ als wahr ansehen, und müssen daraus ableiten, dass $a \mid c$ wahr ist.

Beweis. Angenommen, dass $a \mid b$ und $b \mid c$ gilt. Dann existiert nach Definition ein $q \in \mathbb{N}$, so dass $aq = b$, sowie ein $r \in \mathbb{N}$, so dass $br = c$ ist. Daraus folgt

$$c = br = (aq)r = (qr)a.$$

Daher gibt es ein $s = (qr) \in \mathbb{N}$, so dass $as = c$ gilt. Also gilt $a \mid c$, wieder nach Definition. □

Ein wesentlicher Punkt dieses Beweises (neben der existenziellen Spezialisierung und Generalisierung) war, dass wir uns in der Argumentation auf die mathematische Definition von „a teilt b" berufen haben. Dies sollte (besonders in Übungsaufgaben während der ersten Semester) immer der erste Schritt bei der Suche nach einem Beweis sein. Oft wird dadurch bereits klar, wie Sie argumentieren müssen.

Wir betrachten noch ein Beispiel.

[1] Beachten Sie, dass der Satz nicht in dieser logischen Notation formuliert war. Auch mathematische Sätze sind deutsche Sätze, und müssen als solche lesbar sein.

Definition 6.4. Eine Zahl $n \in \mathbb{Z}$ heisst *gerade*, wenn ein $k \in \mathbb{Z}$ existiert mit $n = 2k$ (also $2 \mid n$ gilt). Gibt es ein $j \in \mathbb{Z}$ mit $n = 2j + 1$, so nennen wir n *ungerade*.[2]

Lemma 6.5. *Ist $n \in \mathbb{Z}$ gerade, dann ist auch n^2 gerade.*

Ein *Lemma* ist eine mathematische Aussage, die bewiesen wird, um sie später im Rahmen eines Beweises einer anderen Aussage anzuwenden. Diese Namensgebung ist keineswegs zwingend; wenn Ihnen aber in einer Vorlesung oder einem Lehrbuch ein Lemma begegnet, können Sie davon ausgehen, dass es im weiteren Verlauf benötigt wird.

Beweis. Sei $n \in \mathbb{Z}$ gerade. Nach Definition existiert also ein $k \in \mathbb{Z}$ mit $n = 2k$. Daher gilt $n^2 = (2k)^2 = 2(2k^2)$. Wegen $2k^2 \in \mathbb{Z}$ ist also auch n^2 gerade, was zu zeigen war. □

6.2 INDIREKTER BEWEIS

Aufgrund der Kontraposition ist die Aussage $p \Rightarrow q$ logisch äquivalent mit der Aussage $(\neg q) \Rightarrow (\neg p)$. Können wir also $(\neg q) \Rightarrow (\neg p)$ beweisen, haben wir damit gleichzeitig auch gezeigt, dass $p \Rightarrow q$ gilt. In einem *indirekten Beweis* nimmt man daher an, dass $\neg q$ gilt, und leitet daraus $\neg p$ ab (führt also einen direkten Beweis der Kontraposition der Aussage).

Ob für eine zu zeigende Aussage ein direkter oder ein indirekter Beweis günstiger ist, lässt sich üblicherweise nicht im Voraus erkennen (auch wenn man nach einiger Zeit ein gewisses Gefühl entwickelt, in welcher Situation welches Vorgehen eher zum Ziel führt). Kommt man mit einer Strategie nicht weiter, so versucht man eben die andere.

Lemma 6.6. *Sei $n \in \mathbb{Z}$. Ist n^2 gerade, dann ist auch n gerade.*

Wollten wir einen direkten Beweis der Aussage versuchen, würden wir annehmen, dass n^2 gerade ist, und daraus folgern, dass ein $k \in \mathbb{Z}$ existiert, so dass $n^2 = 2k$ ist. Jetzt würden wir gerne dem Beweis von Lemma 6.5 folgen, aber ohne weitere Informationen über k (etwa: ist k Quadratzahl?) kommen wir nicht weiter. Die Kontraposition des Lemmas hingegen ist: Wenn n ungerade ist, ist n^2 auch ungerade. Dies sieht Lemma 6.5 hingegen sehr ähnlich, so dass wir hoffen können, dass der Beweis auch sehr ähnlich lauten wird. In der Tat können wir Lemma 6.6 so beweisen:

Beweis. Wir führen einen indirekten Beweis: Sei also n ungerade. Dann existiert ein $k \in \mathbb{Z}$ mit $n = 2k + 1$. Daher gilt $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$. Wegen $2k^2 + 2k \in \mathbb{Z}$ ist also auch n^2 ungerade. Nach Kontraposition folgt, dass wenn n^2 gerade ist, auch n gerade ist. □

[2] Der Beweis, dass eine natürliche Zahl entweder gerade oder ungerade, aber nie beides ist, basiert auf der Eindeutigkeit der Division mit Rest, und sei deshalb erstmal ausgelassen.

Es ist (vor allem in längeren Beweisen) hilfreich für den Leser, wenn man die verwendete Beweistechnik explizit angibt.

Wir betrachten noch ein Beispiel mit einer etwas komplizierteren Aussage:

Satz 6.7. *Seien* X, Y *Mengen und* $f : X \to Y$ *eine Abbildung. Wenn für alle* $h_1, h_2 : X \to X$ *aus* $f \circ h_1 = f \circ h_2$ *folgt, dass* $h_1 = h_2$ *ist, dann ist* f *injektiv.*

Die Voraussetzung enthält jetzt selber eine Implikation, nämlich $(f \circ h_1 = f \circ h_2) \Rightarrow (h_1 = h_2)$. Aus ihrer Gültigkeit für alle h_1, h_2 müssen wir schliessen, dass f injektiv ist. In einem direkten Beweis müssten wir dazu annehmen, dass die Implikation gilt. Dies ist aber ein ungünstiger Start für einen Beweis – wir dürfen ja zum Beispiel nicht davon ausgehen, dass $h_1 = h_2$ gilt, sondern nur, dass dies unter gewissen Voraussetzungen (die wir nicht unter Kontrolle haben) gilt.

Hingegen ist ein indirekter Beweis deutlich einfacher zu konstruieren: Wir nehmen an, dass f nicht injektiv ist, und zeigen, dass die Negation der Allaussage gilt, dass es also *eine* Situation gibt, in der die Implikation nicht gilt. Da die Implikation $p \Rightarrow q$ nur dann falsch ist, wenn p wahr und q falsch ist, reicht es ein Paar h_1, h_2 anzugeben, für das $f \circ h_1 = f \circ h_2$, aber $h_1 \neq h_2$ ist.

Beweis. Wir führen einen indirekten Beweis: Sei also f nicht injektiv, dann existieren $x_1, x_2 \in X$ mit $x_1 \neq x_2$ und $f(x_1) = f(x_2)$. Wir betrachten nun die beiden konstanten Funktionen

$$\begin{aligned} h_1 &: X \to X, \quad x \mapsto x_1 \\ h_2 &: X \to X, \quad x \mapsto x_2 \end{aligned}$$

Da $x_1 \neq x_2$ vorausgesetzt war, ist auch $h_1 \neq h_2$. Andererseits gilt für alle $x \in X$

$$(f \circ h_1)(x) = f(x_1) = f(x_2) = (f \circ h_2)(x),$$

also $f \circ h_1 = f \circ h_2$. Also existieren $h_1, h_2 : X \to X$ mit $h_1 \neq h_2$ und $f \circ h_1 = f \circ h_2$. Nach Kontraposition ist damit die zu zeigende Aussage bewiesen. □

6.3 BEWEIS DURCH WIDERSPRUCH

Noch „indirekter“ geht man in einem *Beweis durch Widerspruch* vor. Das Kernargument hier ist die in Kapitel 5 aufgeführte Schlussregel der *reductio ad absurdum*, meistens in der Form $((\neg p \Rightarrow q) \wedge (\neg p \Rightarrow \neg q)) \Rightarrow p$. Bevor wir sie für den Beweis von Implikationen anwenden, wollen wir ihren Gebrauch für einfache Aussagen untersuchen.

Der bekannteste Widerspruchsbeweis ist sicher der folgende:

Satz 6.8. *Die Wurzel von* 2 *ist irrational.*

Beweis. Wir führen einen Beweis durch Widerspruch: Angenommen, $\sqrt{2}$ ist eine rationale Zahl (also als Bruch darstellbar). Dann gibt es ganze Zahlen $k \in \mathbb{Z}$ und $l \in \mathbb{Z} \setminus 0$, so dass $\sqrt{2} = \frac{k}{l}$ und l und k teilerfremd sind. (Wir kürzen den Bruch also, soweit es geht – bis eben l und k keinen gemeinsamen Teiler ausser 1 mehr haben). Durch Quadrieren erhalten wir $2 = \frac{k^2}{l^2}$, und damit gilt

$$k^2 = 2l^2. \tag{6.1}$$

Also ist k^2 gerade, und wir folgern mit Lemma 6.6, dass auch k gerade ist.

Nach Definition existiert daher ein $j \in \mathbb{Z}$ mit $k = 2j$. Einsetzen in (6.1) ergibt $4j^2 = 2l^2$, und nach Kürzen $l^2 = 2j^2$. Also ist l^2 gerade, und daher (wieder nach Lemma 6.6) auch l. Wir haben also gezeigt, dass k und l beide den Teiler 2 besitzen. Das ist aber ein Widerspruch dazu, dass k und l teilerfremd sind. Die Annahme, dass $\sqrt{2}$ rational ist, kann also nicht wahr sein, womit die zu zeigende Aussage bewiesen ist. □

Welche Aussage q jeweils zum Widerspruch zu führen (d.h. q und $\neg q$ abzuleiten) ist, stellt dabei jeweils die Kernfrage dar. Auch hier gibt es keine Kochrezepte: Es ist gerade die Freiheit dieser Wahl, die den Beweis durch Widerspruch zu so einem mächtigen Werkzeug (und Mathematik zu einer kreativen Tätigkeit) macht.

Wir betrachten noch ein wichtiges Beispiel:

Satz 6.9. *Es gibt keine surjektive Funktion von einer Menge M in ihre Potenzmenge $\mathcal{P}(M)$.*

In einem Beweis durch Widerspruch nehmen wir nun an, dass es eine Funktion $f : M \to \mathcal{P}(M)$ gibt, die surjektiv ist, und leiten daraus eine neue Aussage und deren Negation ab. Die Verbindung von Mengen mit Mengen von Mengen erinnert uns an die Russellsche Antinomie: Wir versuchen daher, einen ähnlichen Widerspruch zu konstruieren.

Beweis. Sei M eine beliebige Menge und $f : M \to \mathcal{P}(M)$ eine surjektive Funktion. Wir betrachten nun die Menge $A = \{x \in M : x \notin f(x)\}$, die offensichtlich Element der Potenzmenge von M ist (und auch die leere Menge oder ganz M sein kann). Weil f surjektiv ist, existiert ein $a \in M$ mit $A = f(a)$. Ist jetzt $a \in A$? Wir machen eine Fallunterscheidung:

1. Angenommen, $a \in A$. Dann gilt nach Definition von A, dass $a \notin f(a) = A$, was ein Widerspruch ist.
2. Ist aber $a \notin A$, dann gilt (wegen $f(a) = A$), dass $a \in A$ ist, und damit erhalten wir den selben Widerspruch.

Die Annahme, dass es eine surjektive Funktion $f : M \to \mathcal{P}(M)$ gibt, hat also auf einen Widerspruch geführt, und kann daher nicht wahr sein. □

In einem Beweis durch Widerspruch ist es nicht zwingend, dass man sowohl q als auch $\neg q$ direkt aus der Annahme ableitet. Wenn man weiss, dass q wahr ist (etwa, da q ein Axiom oder eine bereits bewiesene Aussage ist), reicht es, $\neg q$ abzuleiten:

Satz 6.10. *Die Menge der natürlichen Zahlen hat kein größtes Element.*

Beweis. Wir führen einen Beweis durch Widerspruch: Angenommen, es gibt eine größte natürliche Zahl, d.h. es existiert ein $n \in \mathbb{N}$ so dass $n \geqslant m$ für alle $m \in \mathbb{N}$ gilt. Wählen wir $m = n + 1$, so bedeutet das, dass $n \geqslant n + 1$ gilt. Daraus folgt aber, dass $0 \geqslant 1$ ist – ein Widerspruch, da wir (nach der Definition der Ordnung auf $\mathbb{N}$) wissen, dass $0 < 1$ ist. Also kann keine größte natürliche Zahl existieren. □

Setzt man in der obigen Form der *reductio ad absurdum* für p eine Implikation $a \Rightarrow b$ ein, erhält man (durch die Definition der Implikation als $\neg a \vee b$) die Tautologie

$$((a \wedge \neg b) \Rightarrow (q \wedge \neg q)) \Rightarrow (a \Rightarrow b).$$

Der Beweis durch Widerspruch geht also wie folgt: Man nimmt an, dass a wahr und b falsch ist, und leitet daraus eine Aussage q und ihre Negation $\neg q$ her. (Wieder ist der zentrale Punkt, q geeignet zu wählen.) Vergleichen Sie dieses Vorgehen mit dem im direkten und im indirekten Beweis: Bei einem Beweis durch Widerspruch dürfen wir a *und* $\neg b$ als gegeben voraussetzen. Die Tatsache, dass wir eine zusätzliche Aussage zur Verfügung haben, macht die Stärke dieser Beweisstrategie aus.

Ein Beispiel soll das Vorgehen demonstrieren:

Satz 6.11. *Seien* X, Y *Mengen und* $f : X \to Y$ *eine Funktion. Wenn* f *eine Links-Inverse besitzt, dann ist* f *injektiv.*

Beweis. Wir führen einen Beweis durch Widerspruch: Wir nehmen an, f besitzt eine Links-Inverse g, ist aber nicht injektiv. Da g eine Links-Inverse zu f ist, gilt $g(f(x)) = x$ für alle $x \in X$. Aus der Nicht-Injektivität von f folgt ausserdem, dass $x_1, x_2 \in X$ existieren, so dass $f(x_1) = f(x_2)$ und $x_1 \neq x_2$ gilt. Dann ist aber

$$x_1 = g(f(x_1)) = g(f(x_2)) = x_2,$$

was ein Widerspruch zu $x_1 \neq x_2$ ist.[3] □

[3] Diesen Beweis hätte man auch – sogar kürzer – als indirekten (und als direkten) Beweis formulieren können, da wir eine der Annahmen in einen Widerspruch verwickelt haben, ohne sie im Beweis explizit verwendet zu haben. (In einem „echten" Widerspruchsbeweis führt man den Widerspruch mit einer dritten Aussage – weder Prämisse noch Folgerung – her.) Der hier gegebene Beweis bleibt natürlich richtig.

6.4 BEWEISE MIT FALLUNTERSCHEIDUNG

Wir haben bereits gesehen (etwa im Beweis von Satz 6.9), dass es manchmal hilfreich ist, den Beweis aufzuspalten: Zuerst nimmt man an, dass eine gewisse Eigenschaft gilt ($a \in A$), die man verwenden kann, um den Satz zu beweisen. Danach weist man nach, dass auch im Fall, dass die Eigenschaft nicht gilt ($a \notin A$), die zu beweisende Aussage gültig ist. Ähnlich wie im Beweis durch Widerspruch besteht der Nutzen dieser Technik darin, quasi „umsonst“ eine neue Aussage zur Verfügung zu haben, die man im Beweis einsetzen kann (in einem Fall $a \in A$, im anderen $a \notin A$).

Satz 6.12. *Sei* $n \in \mathbb{N}$. *Dann ist* $n^2 + n$ *gerade.*

Beweis. Sei $n \in \mathbb{N}$. Wir machen eine Fallunterscheidung:

1. n ist gerade. Dann existiert ein $k \in \mathbb{N}$ mit $n = 2k$, und es gilt

$$n^2 + n = 4k^2 + 2k = 2(2k^2 + k).$$

   Also ist $n^2 + n$ gerade.

2. n ist ungerade. Dann existiert ein $k \in \mathbb{N}$ mit $n = 2k + 1$, und es gilt

$$n^2 + n = (2k + 1)^2 + 2k + 1 = (4k^2 + 4k + 1) + 2k + 1 = 2(2k^2 + 3k + 1).$$

   Also ist $n^2 + n$ gerade.

Da jede natürliche Zahl n entweder gerade oder ungerade ist, ist $n^2 + n$ immer gerade. □

In diesem Fall konnte jeweils nur ein Fall zutreffen; dies ist aber nicht unbedingt notwendig. Kritisch ist dabei nur, dass am Ende wirklich *alle möglichen* Fälle abgedeckt sind, insbesondere, wenn die einzelnen Fälle wieder in Unter-Fälle aufgeteilt werden müssen.

Satz 6.13. *Seien* A, B, C *Mengen. Wenn* $A \subseteq C$ *und* $B \subseteq C$ *ist, dann ist* $(A \cup B) \subseteq C$.

Beweis. Es gelte $A \subseteq C$ und $B \subseteq C$, und sei x ein beliebiges Element in $A \cup B$. Dann ist nach Definition $x \in A$ oder $x \in B$.

1. $x \in A$. Dann ist wegen $A \subseteq C$ auch $x \in C$.
2. $x \in B$. Dann ist wegen $B \subseteq C$ auch $x \in C$.

Da immer (mindestens) einer der beiden Fälle zutrifft, ist $x \in C$. Da $x \in A \cup B$ beliebig war, gilt $(A \cup B) \subseteq C$. □

Fallunterscheidungen sind oft notwendig, wenn der zu beweisende Satz die Form $p \Rightarrow (q \vee r)$ hat. Eine mögliche Strategie ist dann die folgenden: Im ersten Fall nehmen wir (zusätzlich zu p) eine Aussage s an und leiten q her. Im anderen Fall leiten wir r aus $\neg s$ ab. Dabei ist es möglich, dass die jeweiligen Fälle mit ganz unterschiedlichen Strategien bewiesen werden:

Satz 6.14. *Sei $x \in \mathbb{R}$. Wenn $x^2 \geqslant x$ ist, dann ist entweder $x \leqslant 0$ oder $x \geqslant 1$.*

Beweis. Sei $x^2 \geqslant x$. Wenn $x \leqslant 0$ ist, folgt direkt die Aussage. Ist andererseits $x > 0$, dürfen wir auf beiden Seiten von $x^2 \geqslant x$ durch x dividieren, und erhalten $x \geqslant 1$. □

Eine Aufspaltung des Beweises ist oft auch notwendig, wenn die zu beweisende Aussage eine Konjunktion enthält: Um $p \Rightarrow (q \wedge r)$ zu beweisen, führen wir zwei separate Beweise: einmal $p \Rightarrow q$, und danach $p \Rightarrow r$. Eine sehr häufige Aussage dieser Form ist die Gleichheit von zwei Mengen A und B: $A = B$ gilt nach Definition genau dann, wenn $(A \subseteq B) \wedge (B \subseteq A)$ gilt.

Auch Äquivalenzen der Form $p \Leftrightarrow q$ sind Konjunktionen $(p \Rightarrow q) \wedge (q \Rightarrow p)$, und müssen deshalb in der Regel separat bewiesen werden.

Satz 6.15. *Seien A und B Mengen. Dann gilt $B \subseteq (B \setminus A)$ genau dann, wenn $A \setminus B = A$ ist.*

Beweis. Wir beweisen zuerst die Implikation $(B \subseteq (B \setminus A)) \Rightarrow (A \setminus B = A)$. Sei also $B \subseteq (B \setminus A)$. Wir zeigen jetzt, dass $(A \setminus B) = A$ gilt, indem wir die beiden Inklusionen nachweisen:

1. $(A \setminus B) \subseteq A$: Sei $x \in A \setminus B$ beliebig. Dann ist nach Definition $x \in A$.
2. $A \subseteq (A \setminus B)$: Sei $x \in A$ beliebig. Wir müssen zeigen, dass x nicht in B liegt, und führen einen Beweis durch Widerspruch. Nehmen wir also an, dass $x \in B$ ist. Aus der Voraussetzung, dass $B \subseteq (B \setminus A)$ gilt, folgt daraus, dass $x \in B \setminus A$ liegt, und damit $x \notin A$. Dies ist ein Widerspruch, weshalb $x \notin B$ gelten muss. Also ist $x \in A \setminus B$.

Nun beweisen wir die andere Implikation: Sei $(A \setminus B) = A$. Wir zeigen nun, dass $B \subseteq (B \setminus A)$ gilt. Sei dafür $x \in B$ beliebig. Dann kann x auf keinen Fall in $A \setminus B$ liegen, und aus der Voraussetzung $(A \setminus B) = A$ folgt sofort $x \notin A$. Also ist $x \in B$ und $x \notin A$, und daher gilt nach Definition $x \in B \setminus A$. Da x beliebig war, haben wir $B \subseteq (B \setminus A)$ bewiesen. □

Ist die Äquivalenz von mehr als zwei Aussagen zu beweisen, etwa $(p \Leftrightarrow q) \wedge (q \Leftrightarrow r)$ (auch kurz $p \Leftrightarrow q \Leftrightarrow r$ geschrieben), so kann man sich eines *Ringschlusses* bedienen: Statt alle Implikationen zu zeigen (in diesem Fall vier: $p \Rightarrow q$, $q \Rightarrow p$, $q \Rightarrow r$, $r \Rightarrow q$), beweist man nur drei geeignete, und beruft sich für die Gültigkeit der restlichen Implikationen auf die Verkettung. So kann man aus den drei Implikationen $p \Rightarrow q$, $q \Rightarrow r$ und $r \Rightarrow p$ alle übrigen herleiten. Die Reihenfolge der drei Aussagen in der „Kette“ ist dabei beliebig (kann aber für die Schwierigkeit der Beweisführung große Unterschiede machen).

Satz 6.16. *Seien* X, Y *Mengen und* $f : X \to Y$ *eine Funktion. Dann sind folgende Aussagen äquivalent:*

a) f *ist injektiv.*

b) f *hat eine Links-Inverse.*

c) Für alle Funktionen $h_1, h_2 : X \to X$ *folgt aus* $f \circ h_1 = f \circ h_2$*, dass* $h_1 = h_2$ *ist.*

Wir machen den Ringschluss $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$. Da wir die Implikation $(c) \Rightarrow (a)$ bereits in Satz 6.7 bewiesen haben, sind nur noch die ersten beiden Implikationen offen:

1. Sei f injektiv. Wir zeigen, dass f eine Links-Inverse g besitzt, indem wir nachweisen, dass eine *Funktion* mit den gewünschten Eigenschaften konstruierbar ist. Sei dazu z ein beliebiges Element in X. Wir wählen $g : Y \to X$ wie folgt:

$$g : y \mapsto \begin{cases} x, & \text{falls } y = f(x) \in f(X), \\ z, & \text{falls } y \notin f(X). \end{cases}$$

   Da f injektiv ist, ist die Wahl von x im ersten Fall eindeutig möglich; durch den zweiten Fall wird sicher gestellt, dass g auf ganz Y definiert ist. Wir haben also dadurch eine gültige Funktion $g : Y \to X$ definiert. Weiterhin gilt für $x \in X$, dass $g(f(x)) = x$ ist. Damit ist g eine Links-Inverse zu f.

2. Angenommen, f besitzt eine Links-Inverse g. Wir zeigen, dass aus $f \circ h_1 = f \circ h_2$ folgt, dass $h_1 = h_2$ ist. Seien also $h_1, h_2 : X \to X$ mit $f \circ h_1 = f \circ h_2$. Dann ist für alle $x \in X$

$$h_1(x) = g(f(h_1(x))) = g(f(h_2(x))) = h_2(x),$$

   da nach Voraussetzung $g \circ f = \mathrm{id}_X$ gilt. Also ist $h_1 = h_2$.

6.5 BEWEISE VON QUANTORENAUSSAGEN

6.5.1 AUSSAGEN MIT ALLQUANTOR

Um Aussagen der Form $\forall x \in X : P(x)$ zu beweisen, wendet man die Schlussregel der universellen Generalisierung an: Man nimmt an, dass a ein beliebiges Element von X ist, und leitet ab, dass $P(a)$ gilt. Da dies der gleiche Ansatz wie in einem Beweis der Implikation $(a \in X) \Rightarrow P(a)$ ist, sind die bereits behandelten Strategien wie indirekter Beweis und Beweis durch Widerspruch auch hier anwendbar. Wir wollen deshalb nicht weiter darauf eingehen.

6.5.2 AUSSAGEN MIT EXISTENZQUANTOR

Der Beweis einer Existenzaussage $\exists x \in X : P(x)$ ist auf den ersten Blick denkbar einfach: Man gibt ein Objekt, das die geforderte Eigenschaft hat, an; entweder direkt (wie den Teiler im Beweis von Satz 6.3) oder in Form einer garantiert durchführbaren Konstruktionsvorschrift (wie für die Links-Inverse im Beweis von Satz 6.16).[4] Allerdings muss dabei begründet werden, dass das präsentierte Objekt tatsächlich die gewünschte Eigenschaft hat, und (dies wird oft vergessen) ein Element in X ist.

Für das nächste Beispiel definieren wir $\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$.

Satz 6.17. *Es gibt eine bijektive Funktion* $f : \mathbb{N}_0 \to \mathbb{Z}$.

Beweis. Wir definieren eine Funktion $f : \mathbb{N}_0 \to \mathbb{Z}$ durch die Vorschrift

$$f(n) = \begin{cases} k, & \text{falls ein } k \in \mathbb{N}_0 \text{ existiert mit } n = 2k - 1, \\ -k, & \text{falls ein } k \in \mathbb{N}_0 \text{ existiert mit } n = 2k. \end{cases}$$

Da $k \in \mathbb{N}_0$ ist, ist k und $-k$ in $\mathbb{Z}$. Und da jede Zahl $n \in \mathbb{N}_0$ entweder gerade oder ungerade ist, wird durch diese Zuordnungsvorschrift jedem $n \in \mathbb{N}_0$ ein eindeutiges $m \in \mathbb{Z}$ zugewiesen; also ist dadurch tatsächlich eine Funktion $f : \mathbb{N}_0 \to \mathbb{Z}$ definiert.

Wir müssen nun nachweisen, dass diese Funktion bijektiv ist. Wir zeigen zuerst die Surjektivität: Sei $m \in \mathbb{Z}$ beliebig. Falls $m > 0$ ist, so gilt für $n = 2m - 1 \in \mathbb{N}$, dass $f(n) = m$ ist. Gilt dagegen $m \leqslant 0$, so wählen wir $n = -2m \in \mathbb{N}_0$ und erhalten $f(n) = m$. Für jedes $m \in \mathbb{Z}$ existiert also ein $n \in \mathbb{N}_0$ mit $f(n) = m$.

Nun zur Injektivität: Seien $n, m \in \mathbb{N}_0$ mit $f(n) = f(m) = k$. Wir machen wieder eine Fallunterscheidung:

1. Angenommen, $k > 0$: Dann ist $n = 2k - 1$ und $m = 2k - 1$, also $n = m$.
2. Gilt dagegen $k \leqslant 0$, dann ist $n = -2k$ und $m = -2k$, also wieder $n = m$.

Damit ist f injektiv, und auch bijektiv. □

Beachten Sie, dass wir nicht angeben mussten, *wie* die Funktion f gefunden wurde – dies ist für die logische Struktur des Beweises völlig nebensächlich (wenn auch natürlich didaktisch wertvoll). Im Gegenteil ist dies die häufigste Fehlerursache in einem wichtigen Typ von Existenzaussagen: Der Lösung von Gleichungen.

[4] Eine dritte Variante ist der Beweis durch Widerspruch: Man nimmt an, dass alle $x \in X$ die Eigenschaft $P(x)$ *nicht* haben, und leitet daraus einen Widerspruch ab. Solch ein Existenzbeweis wird *nichtkonstruktiv* genannt, da er nicht erlaubt, in konkreten Situationen ein Objekt mit der gewünschten Eigenschaft zu gewinnen. Die *konstruktivistische Mathematik* (nach L. E. J. Brouwer) lehnt solche Beweise ab; für die meisten Mathematiker sind die mit dieser Methode gewonnenen Resultate aber zu nützlich, um auf sie zu verzichten. Es wird trotzdem versucht, konstruktive Beweise für bisher nichtkonstruktiv bewiesene Resultate zu finden.

Satz 6.18. *Es gibt ein* $x \in \mathbb{R}$*, so dass* $\sqrt{x^2-5} = \sqrt{x+1}$ *gilt.*

Eine oft gesehene „Lösung“ ist die folgende (meistens völlig unkommentierte) Kette von Gleichungen:

$$\begin{aligned} \sqrt{x^2-5} &= \sqrt{x+1} \\ x^2-5 &= x+1 \\ x^2-x-6 &= 0 \\ (x-3)(x+2) &= 0 \\ x = 3 \text{ oder } x &= -2 \end{aligned}$$

Dabei wird impliziert, dass wir die Wahl zwischen diesen beiden Werten haben. Dies ist aber nicht korrekt; $x = -2$ dürfen wir noch nicht mal in die Wurzel einsetzen. Der Fehler hier ist, dass diese Rechnung für unsere Zwecke „rückwärts“ läuft: Aus der Tatsache, dass $\sqrt{x^2-5} = \sqrt{x+1}$ gilt, wird abgeleitet, dass x einen bestimmten Wert hat. Wir müssen aber genau die andere Richtung zeigen: Wenn x einen bestimmten Wert hat (hier: $x = 3$), dann gilt $\sqrt{x^2-5} = \sqrt{x+1}$. Den ersten Schritt, eine Gleichung auf beiden Seiten zu quadrieren, können wir aber nicht immer rückgängig machen ($1 = 1$ impliziert $1^2 = 1^2$, aber $(-1)^2 = 1^2$ impliziert nicht $-1 = 1$).[5] Ein mathematischer Beweis sollte deshalb immer „vorwärts“ aufgeschrieben werden: Beginnend mit als wahr gegebenen Aussagen, und endend mit der zu beweisenden Aussage. Korrekt wäre der folgende Beweis.

Beweis. Wähle $x = 3 \in \mathbb{R}$. Dann ist $\sqrt{3^2-5} = \sqrt{4} = \sqrt{3+1}$. □

Natürlich ist die „rückwärts“-Rechnung ein probates Mittel, um den Kandidaten $x = 3$ zu finden.

Ein Spezialfall sind *Existenz- und Eindeutigkeitsbeweise*. Dabei handelt es sich um Beweise von Aussagen der Form „Es gibt genau ein …“. Wie in Kapitel 1.4 bereits betont, handelt es sich dabei um zwei Aussagen, die auch separat beweisen werden müssen: eine Existenzaussage („es gibt ein …“) und eine Eindeutigkeitsaussage („es kann kein zweites … geben“). Für die Existenzaussage geht man wie oben geschildert vor, um ein Objekt x mit den gewünschten Eigenschaften zu erhalten. Die Eindeutigkeitsaussage wird bewiesen, indem man annimmt, dass ein ansonsten beliebiges Objekt y mit der geforderten Eigenschaft existiert, und daraus ableitet, dass $y = x$ gilt[6] (es sich also nur um verschiedene Namen für das selbe Objekt handelt).

Satz 6.19. *Sei* M *eine Menge und* $A \subseteq M$*. Dann existiert genau eine Menge* $B \subseteq M$ *mit* $A \cap B = \emptyset$ *und* $A \cup B = M$*.*

[5] Ein weiterer Grund, in mathematischen Argumenten immer die Implikationen explizit anzugeben.
[6] Dies ist ein direkter Beweis von $\forall y \in X : (P(y) \Rightarrow x = y)$ für ein gegebenes $x \in X$.

Beweis. Wir beweisen zuerst die Existenz. Wähle $B = A^c = M \setminus A \subseteq M$. Dann gilt sowohl $A \cap A^c = \emptyset$ als auch $A \cup A^c = M$ (siehe Übungsaufgabe).

Nun zeigen wir die Eindeutigkeit. Sei $C \subseteq M$ eine beliebige Menge mit $A \cap C = \emptyset$ und $A \cup C = M$. Nun gilt wegen $A \cap C = \emptyset$, dass $C \subseteq A^c$ ist: Für alle $x \in M$ folgt aus $x \in C$ dass $x \notin A$ ist. Umgekehrt folgt aus $A \cup C = M$, dass $A^c \subseteq C$ gilt: Für alle $x \in M$ folgt aus $x \notin A$ dass $x \in C$ sein muss. Also gilt $A^c \subseteq C$, und damit auch $C = A^c$. □

Teil III

UNENDLICHE MENGEN

ÜBERBLICK

Eine der großen Leistungen während der Fundierung der modernen Mathematik in der Zeit um 1900 war die Präzisierung und rigorose Untersuchung des Begriffs der „Unendlichkeit", frei von jedem philosophischen Beiwerk. Als ein wesentliches Resultat stellte sich dabei heraus, dass man unendliche Mengen bezüglich Ihrer „Größe" weiter klassifizieren kann: in *abzählbar* unendliche Mengen (darunter vor allem die Menge der natürlichen Zahlen $\mathbb{N}$) und *überabzählbar* unendliche Mengen (wie etwa die Menge der reellen Zahlen $\mathbb{R}$). Ein fundamentaler Unterschied zwischen beiden Klassen ist, dass Allaussagen über abzählbare Mengen mit Hilfe des Prinzips der vollständigen Induktion bewiesen werden können. Bevor wir die obigen Begriffe näher untersuchen, wollen wir zuerst dieses wichtige Beweisprinzip ausführlich darstellen.

VOLLSTÄNDIGE INDUKTION

7

Es ist eine der charakterisierenden Eigenschaften der natürlichen Zahlen, dass jede Zahl n stets einen eindeutigen Nachfolger $n+1$ besitzt, und dass es eine „erste“ Zahl gibt, die selber kein Nachfolger ist – um etwas über alle natürlichen Zahlen auszusagen, kann man sie also „der Reihenfolge nach“ untersuchen. (Genau diese Idee steckt hinter dem Begriff „abzählbar“.) Dies können wir uns für Beweise und Definitionen zu Nutze machen. Im ersten Fall erhalten wir das Prinzip der *vollständigen Induktion*, im zweiten das Prinzip der *rekursiven Definition*.

7.1 DAS PRINZIP DER VOLLSTÄNDIGEN INDUKTION

Das *Prinzip der vollständigen Induktion* ist eine besondere Schlussregel für Allaussagen über die natürlichen Zahlen. Geschrieben in der in Kapitel 5 gebrauchten Form lautet es

$$(P(1) \wedge (\forall n \in \mathbb{N} : P(n) \Rightarrow P(n+1))) \Rightarrow (\forall n \in \mathbb{N} : P(n)).$$

Bei einem Induktionsbeweis der Aussage $\forall n \in \mathbb{N} : P(n)$ geht man also wie folgt vor:

1. *Induktionsbeginn*: Man zeigt, dass $P(1)$ gilt.
2. *Induktionsschritt*: Man nimmt an, dass für beliebiges $n \in \mathbb{N}$ die Aussage $P(n)$ gilt, und leitet daraus ab, dass $P(n+1)$ gilt. (Die Aussage $P(n)$ nennt man auch *Induktionsannahme*.)

Mit dem Induktionsschritt haben wir also eine Implikationskette konstruiert, die – wie eine Dominoreihe – über alle natürlichen Zahlen läuft. Der Induktionsbeginn entspricht dann dem Umwerfen des ersten Steins, der dann den nächsten umwirft: $P(1)$ ist wahr, woraus folgt, dass $P(2)$ wahr ist, woraus folgt, dass $P(3)$ wahr ist, woraus folgt …[1]

[1] Dass wir damit wirklich jede der (unendlich vielen) natürlichen Zahlen „erwischen“, ist nicht selbstverständlich. Tatsächlich ist dies eine der wesentlichen Eigenschaften, die durch die mathematische Definition der natürlichen Zahlen gewährleistet wird.

Ähnlich wie beim Beweis durch Widerspruch liegt der Vorteil bei dieser Beweistechnik gegenüber dem direkten Beweis von $n \in \mathbb{N} \Rightarrow P(n)$ darin, dass wir im Beweis eine zusätzliche Aussage (nämlich $P(n-1)$) verwenden können.

Wir verdeutlichen das Prinzip an einem klassischen Beispiel. Hierfür verwenden wir die in der Mathematik übliche Schreibweise für die Summe der Zahlen $a_1, a_2, \ldots, a_n$:[2]

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n.$$

Satz 7.1. *Für alle $n \in \mathbb{N}$ gilt*

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

Beweis. Wir führen den Beweis durch vollständige Induktion. Für $n \in \mathbb{N}$ ist die Aussage $P(n)$ die Gültigkeit der Gleichung $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

Induktionsbeginn: Wir müssen zeigen, dass $P(1)$ gilt, dass also $\sum_{k=1}^{1} k = \frac{1(1+1)}{2}$ gilt. Auf der linken Seite der Gleichung steht nur der Summand 1, und die rechte Seite kann vereinfacht werden zu $\frac{2}{2} = 1$. Also gilt

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

Induktionsschritt: Angenommen, $P(n)$ gilt. Wir müssen jetzt daraus folgern, dass $P(n+1)$ gilt. In anderen Worten, um zu zeigen, dass $\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2}$ ist, dürfen wir verwenden, dass $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ gilt. Dazu formen wir die Summe so um, dass wir die Induktionsannahme $P(n)$ anwenden können (was wir in Rechnungen immer mit „IA“ kennzeichnen wollen):

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

Andererseits erhalten wir durch Ausmultiplizieren

$$\frac{(n+1)(n+2)}{2} = \frac{n^2 + 3n + 2}{2}.$$

Zusammen gilt also:

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

was zu zeigen war. □

[2] Analog definiert man das Produkt $\prod_{k=1}^{n} a_k = a_1 \cdot a_2 \cdots a_n$. Eine mathematisch saubere Definition ohne „Pünktchen“ erhält man, indem man die Rekursion aus Kapitel 7.2 anwendet.

Beachten Sie, dass wir die Umformungen am Schluss wieder „vorwärts“ aufgeschrieben haben, und nicht in der Form, in der wir die Rechnungen (etwa auf einem Schmierzettel) durchprobiert haben. Ebenso gilt, dass wir die verwendete Beweisstrategie explizit angegeben haben. Hilfreich ist auch oft, die Annahme $P(n)$ und die daraus zu folgernde Aussage $P(n+1)$ möglichst explizit hinzuschreiben.

Wir müssen eine Induktion nicht bei $n = 1$ beginnen; genauso gut können wir bei $n = 0$ anfangen, wenn auch die Aussage $P(0)$ gilt. Umgekehrt können wir auch Aussagen für alle natürlichen Zahlen ab einer gewissen Zahl n_0 mit Induktion beweisen. Allgemein beweist man (analog zu oben) für gegebenes $n_0 \in \mathbb{N}_0$ eine Aussage der Form $\forall n \in \{m \in \mathbb{N}_0 \,:\, m \geqslant n_0\} : P(n)$, indem man $P(n_0)$ nachweist, und dann für alle $n \geqslant n_0$ zeigt, dass aus $P(n)$ die Aussage $P(n+1)$ folgt.

Satz 7.2. *Für alle $n \in \mathbb{N}$ mit $n \geqslant 5$ gilt $2^n > n^2$.*

Beweis. Wir verwenden vollständige Induktion. Da $2^5 = 32 > 25 = 5^2$ ist, ist der Induktionsbeginn gezeigt. Sei nun $n \geqslant 5$ und $2^n > n^2$. Wir müssen zeigen, dass $2^{n+1} > (n+1)^2$ ist. Nach Induktionsannahme gilt:

$$2^{n+1} = 2 \cdot 2^n \overset{\text{IA}}{>} 2n^2.$$

Andererseits ist $(n+1)^2 = n^2 + 2n + 1$. Können wir also beweisen, dass $n^2 > 2n + 1$ ist, sind wir fertig. Hier können wir verwenden, dass $n \geqslant 5 > 1$ (und $3 > 1$) ist:

$$n^2 \geqslant 5n = 2n + 3n > 2n + 1.$$

Also gilt $2^{n+1} > 2n^2 > n^2 + 2n + 1 = (n+1)^2$, was zu zeigen war. □

Nicht nur elementare arithmetische Aussagen können mit vollständiger Induktion bewiesen werden. Das folgende Lemma werden wir im nächsten Kapitel benötigen:

Lemma 7.3. *Die Funktion $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$,*

$$(a, b) \mapsto \frac{(a+b-2)(a+b-1)}{2} + b$$

ist surjektiv.

Beweis. Wir überlegen uns zunächst, dass die obige Vorschrift wirklich eine Abbildung von $\mathbb{N} \times \mathbb{N}$ nach $\mathbb{N}$ definiert, dass also

$$\frac{(a+b-2)(a+b-1)}{2} + b \in \mathbb{N}$$

gilt für alle $(a, b) \in \mathbb{N} \times \mathbb{N}$. Ist $(a, b) = (1, 1)$ so ist

$$\frac{(a+b-2)(a+b-1)}{2} + b = 1 \in \mathbb{N}.$$

Sei nun $(a, b) \in \mathbb{N} \times \mathbb{N} \setminus \{(1, 1)\}$. Dann ist $m := a + b - 2 \in \mathbb{N}$. Nach Satz 6.12 ist $m^2 + m = m(m + 1) = (a + b - 2)(a + b - 1)$ gerade, woraus

$$\frac{(a + b - 2)(a + b - 1)}{2} + b \in \mathbb{N}$$

folgt.

Wir beweisen jetzt, dass f surjektiv ist. Wir zeigen dazu mit Induktion, dass es für jedes $n \in \mathbb{N}$ ein $(a, b) \in \mathbb{N} \times \mathbb{N}$ mit $n = f(a, b)$ gibt. (Der Lesbarkeit halber wollen wir auch im Folgenden stets $f(a, b)$ für $f((a, b))$ schreiben.)

Induktionsbeginn: Wir haben bereits gesehen, dass $f(1, 1) = 1$ ist. Also existiert für $n = 1$ ein Paar $(a, b) = (1, 1) \in \mathbb{N} \times \mathbb{N}$ mit $f(a, b) = n$.

Induktionsschritt: Sei nun $n \in \mathbb{N}$ und es existiere $(a, b) \in \mathbb{N} \times \mathbb{N}$ mit $n = f(a, b)$ (Induktionsannahme). Wir zeigen, dass $n + 1 = f(c, d)$ für ein geeignetes $(c, d) \in \mathbb{N} \times \mathbb{N}$ gilt. Dazu machen wir eine Fallunterscheidung.

1. Es ist $a = 1$. Dann gilt

   $$n+1 = f(a, b)+1 = \frac{(b - 1)b}{2}+b+1 = \frac{b^2 + b}{2}+1 = \frac{b(b + 1)}{2}+1 = f(b+1, 1).$$

   Also gilt die Aussage für $(c, d) = (b + 1, 1)$.

2. Es ist $a > 1$. Dann ist $(c, d) := (a - 1, b + 1) \in \mathbb{N} \times \mathbb{N}$ und es gilt

   $$\begin{aligned} n + 1 = f(a, b) + 1 &= \frac{(a + b - 2)(a + b - 1)}{2} + b + 1 = \\ &= \frac{\Big((a - 1) + (b + 1) - 2\Big)\Big((a - 1) + (b + 1) - 1\Big)}{2} + b + 1 \\ &= f(a - 1, b + 1). \end{aligned}$$

Damit ist in jedem Fall $(c, d) \in \mathbb{N} \times \mathbb{N}$ gefunden, und die Surjektivität gezeigt. □

Eine oft sehr nützliche Variante des Induktionsbeweises verwendet die *starke Induktion:*

$$[\forall n \in \mathbb{N} : (\forall k < n : P(k)) \Rightarrow P(n))] \Rightarrow \forall n \in \mathbb{N} : P(n).$$

Hier haben wir der Übersichtlichkeit halber einfach $k < n$ statt $k \in \{m \in \mathbb{N} : m < n\}$ geschrieben. Der Name „starke Induktion“ rührt daher, dass wir bei dem Beweis von $P(n)$ nicht nur $P(n - 1)$ verwenden dürfen, sondern alle $P(k)$ für $k < n$. Wir können uns zusätzlich den Induktionsbeginn sparen. Haben wir nämlich den starken Induktionsschritt $\forall n \in \mathbb{N} : (\forall k < n : P(k)) \Rightarrow P(n))$ gezeigt, können wir daraus folgern, dass $P(1)$ gilt: Die Aussage „$\forall k < 1 : P(k)$“ ist immer wahr, da die Menge $\{m \in \mathbb{N} : m < 1\}$ die leere Menge

ist, für die jede Allaussage gilt. Aus $P(1)$ folgt dann, dass $P(2)$ gilt, und aus $P(1)$ und $P(2)$ folgt $P(3)$ …(Mit (vollständiger) Induktion kann man rigoros beweisen, dass daraus $P(n)$ für alle $n \in \mathbb{N}$ folgt.)

Bei der starken Induktion gehen wir also wie folgt vor: Wir nehmen an, dass für beliebiges $n \in \mathbb{N}$ die Aussage $P(k)$ für alle $k < n$ gilt, und leiten daraus ab, dass $P(n)$ wahr ist.

Wir können damit eine wichtige Eigenschaft der natürlichen Zahlen beweisen, und gleichzeitig ein sehr elegantes Beispiel für einen Induktionsbeweis geben.

Satz 7.4 (Wohlordnungsprinzip)**.** *Jede nichtleere Teilmenge der natürlichen Zahlen hat ein kleinstes Element.*

Beweis. Wir führen einen indirekten Beweis: Wir nehmen an, $N \subseteq \mathbb{N}$ hat kein kleinstes Element, und zeigen, dass $N = \emptyset$ ist. Da dies äquivalent ist mit der Aussage $\forall n \in \mathbb{N} : n \notin N$, reicht es, diese zu zeigen, und zwar mit starker Induktion. Sei $n \in \mathbb{N}$ beliebig, und gelte $k \notin N$ für alle $k < n$. Dann kann aber auch n nicht in N liegen: Wäre nämlich $n \in N$, so wäre n gerade das kleinste Element von N (da ja nach Induktionsannahme $k \notin N$ für alle $k < n$ ist), dass es nach Voraussetzung nicht geben kann. Das ist ein Widerspruch, also ist $n \notin N$ und damit der Induktionsschritt gezeigt. □

Es ist also nicht immer offensichtlich, wie man eine Allaussage durch vollständige Induktion beweisen kann. Wir betrachten ein weiteres Beispiel:

Satz 7.5. *Sei M eine Menge und R eine Ordnungsrelation auf M. Dann hat jede nichtleere endliche Teilmenge $A \subseteq M$ ein minimales Element bezüglich R.*

Auf den ersten Blick hat diese Aussage nichts mit natürlichen Zahlen zu tun, geschweige denn mit vollständiger Induktion. Bei genauerem Überlegen finden wir aber einen Anhaltspunkt bei der Bedingung, dass A endlich ist: Eine endliche Menge hat eine bestimmte Anzahl von Elementen, und diese Anzahl muss eine natürliche Zahl sein (wir werden sehen, dass diese intuitive Vorstellung auch eine mathematisch korrekte Definition ergibt). Wir führen also die Induktion nach der Anzahl der Elemente von A durch.

Beweis. Wir beweisen durch Induktion nach n, dass jede Teilmenge $A \subseteq M$ mit genau n Elementen ein minimales Element a enthält, dass also ein $a \in A$ existiert, so dass für alle $x \in A$ gilt: aus $x \preceq a$ folgt $x = a$.

Induktionsbeginn: Angenommen, A hat nur ein Element, es ist also $A = \{a\}$ für ein $a \in M$. Für alle $x \in A$ gilt dann $x = a$, und damit insbesondere die behauptete Implikation.

Induktionsschritt: Angenommen, jede Teilmenge von M mit genau n Elementen hat ein minimales Element. Wir müssen zeigen, dass dann auch jede Teilmenge von M mit genau $n + 1$ Elementen ein minimales Element hat. Sei also A eine beliebige Teilmenge von M mit $n + 1$ Elementen. Wähle nun ein beliebiges Element $a \in A$, und betrachte $A' = A \setminus \{a\}$.

Dann hat A' genau n Elemente, und daher nach Induktionsannahme ein minimales Element $b \in A'$. Wir zeigen nun, dass entweder a oder b ein minimales Element von A ist. Dafür machen wir eine Fallunterscheidung:

1. Es gilt $a \preceq b$: Dann ist a das gewünschte minimale Element. Sei nämlich $x \in A$ mit $x \preceq a$. Angenommen, es gilt $x \in A'$. Aufgrund der Transitivität von R folgt aus $x \preceq a$, dass $x \preceq b$ gilt, und da b minimales Element von A' ist, erhalten wir $x = b$. Also ist $x \preceq a$ und $a \preceq b = x$, und aus der Antisymmetrie folgt $x = a \notin A'$. Das ist ein Widerspruch, also muss $x \in A \setminus A'$ sein. Dann gilt aber gerade $x = a$, wie gefordert.
2. Es gilt nicht $a \preceq b$: Dann ist b das gewünschte minimale Element. Sei $x \in A$ beliebig. Wenn $x \in A'$ ist, so gilt die Implikation $x \preceq b \Rightarrow x = b$, da b minimales Element von A' ist. Ist $x \notin A'$, so muss $x = a$ gelten, und da $a \preceq b$ nicht gilt, ist die Implikation $x \preceq b \Rightarrow x = b$ auch wahr.

□

7.2 REKURSIVE DEFINITION

Umgekehrt können wir die Grundidee der Induktion auch verwenden, um für eine gegebene Menge X eine Funktion $f : \mathbb{N} \to X$ zu definieren[3]. Statt jeder natürlichen Zahl explizit ein Element zuzuordnen (zum Beispiel $f(n) = n$ für $f : \mathbb{N} \to \mathbb{N}$), geben wir eine explizite Zuordnung nur für $n = 1$ an, und definieren für $n > 1$ den Wert $f(n)$ in Abhängigkeit von $f(n-1)$ (zum Beispiel $f(n) = f(n-1) + 1$). Dass dieses Vorgehen zu einer wohldefinierten Funktion führt, kann rigoros bewiesen werden.

Satz 7.6 (Rekursive Definition)**.** *Sei* X *eine Menge und* $a \in X$ *sowie die Funktion* $h : \mathbb{N} \times X \to X$ *gegeben. Dann existiert eine eindeutig bestimmte Funktion* $f : \mathbb{N} \to X$*, so dass gilt*

$$\begin{aligned} f(1) &= a, \\ f(n+1) &= h(n, f(n)) \quad \textit{für alle } n \in \mathbb{N}. \end{aligned}$$

Auf diese Weise können die Potenzen a^n für gegebenes $a \in \mathbb{R}$ und beliebiges $n \in \mathbb{N}$ mathematisch sauber definiert werden: Durch die Wahl $h : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$, $h(n, x) = x \cdot a$ werden die Potenzen durch

$$\begin{aligned} a^1 &= a, \\ a^{n+1} &= a^n \cdot a \quad \text{für alle } n \in \mathbb{N}, \end{aligned}$$

eindeutig festgelegt (die letzte Zeile wird als *Rekursionsvorschrift* bezeichnet). Diese Definition kann man nun verwenden, um zum Beispiel durch Induktion die Gültigkeit der Rechenregeln zu beweisen.

[3] Eine Funktion $f : \mathbb{N} \to X$ wird auch *Folge* (in X) genannt, das Element $f_n := f(n) \in f(\mathbb{N})$ bezeichnet man als *Folgeglied*.

Satz 7.7. *Für alle* $n, m \in \mathbb{N}$ *und* $a \in \mathbb{R}$ *gilt* $a^{m+n} = a^m \cdot a^n$.

Beweis. Seien $a \in \mathbb{R}$ und $m \in \mathbb{N}$ beliebig. Wir führen einen Beweis durch Induktion nach n.

Induktionsbeginn: Sei $n = 1$. Dann ist $a^{m+1} = a^m \cdot a = a^m \cdot a^1$.

Induktionsschritt: Sei $n \in \mathbb{N}$ beliebig, und gelte $a^{m+n} = a^m \cdot a^n$. Dann ist wegen der Assoziativität von Addition in $\mathbb{N}$ und Multiplikation in $\mathbb{R}$:

$$a^{m+(n+1)} = a^{(m+n)+1} = a^{m+n} \cdot a \overset{\text{IA}}{=} (a^m \cdot a^n) \cdot a = a^m(a^n \cdot a) = a^m \cdot a^{n+1},$$

wobei wir im zweiten und letzten Schritt die rekursive Definition der Potenz verwendet haben. □

Auf dieselbe Weise kann auch die Summe $\sum_{k=1}^{n} a_k$ und das Produkt $\prod_{k=1}^{n} a_k$ über $a_k \in \mathbb{R}$ mit $1 \leqslant k \leqslant n$ sauber definiert werden.

Als ein weiteres Beispiel betrachten wir die *Binomialkoeffizienten* $\binom{n}{k}$ für $n, k \in \mathbb{N}_0$ mit $n \geqslant k \geqslant 0$, die zum Beispiel die Anzahl der k-elementigen Teilmengen einer Menge mit n Elementen bezeichnen. Eine Möglichkeit, diese anzugeben, benutzt das *Pascalsche Dreieck*: Wir schreiben an die Spitze eines Dreiecks eine 1, und füllen die darunterliegenden Zeilen, indem wir für jeden Eintrag jeweils die Zahlen links und rechts darüber addieren (wobei Zahlen ausserhalb des Dreiecks als 0 genommen werden). Der $k$te Eintrag in der $n$ten Spalte (jeweils beginnend mit 0 gezählt) ist dann $\binom{n}{k}$ (siehe Abbildung 7.1).

Wir können diese Konstruktionsvorschrift als rekursive Definition schreiben. Dabei nutzen wir aus, dass auf dem Rand des Dreiecks immer 1 steht. (Dies kann man durch Induktion aus der oben angegebenen Bildungsvorschrift herleiten.)

Definition 7.8. Für $n, k \in \mathbb{N}_0$ mit $n \geqslant k \geqslant 0$ ist der Binomialkoeffizient $\binom{n}{k}$ definiert als

$$\binom{n}{k} := \begin{cases} 1 & \text{falls } k = 0 \text{ oder } k = n, \\ \binom{n-1}{k-1} + \binom{n-1}{k} & \text{falls } 0 < k < n. \end{cases}$$

Häufig sucht man *geschlossene* Darstellungen von rekursiv definierten Objekten, das heisst eine Darstellung für $f(n)$, die nicht auf vorherige Elemente $f(n-1)$ zurückgreift. Durch kombinatorische Überlegungen kommt man auf folgende Formel:[4]

Satz 7.9. *Für* $n, k \in \mathbb{N}_0$ *mit* $n \geqslant k \geqslant 0$ *gilt*

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

[4] Ein allgemeines Verfahren, um geschlossene Darstellungen für rekursive Funktionen zu erhalten, beruht auf der Theorie der *Erzeugendenfunktionen*.

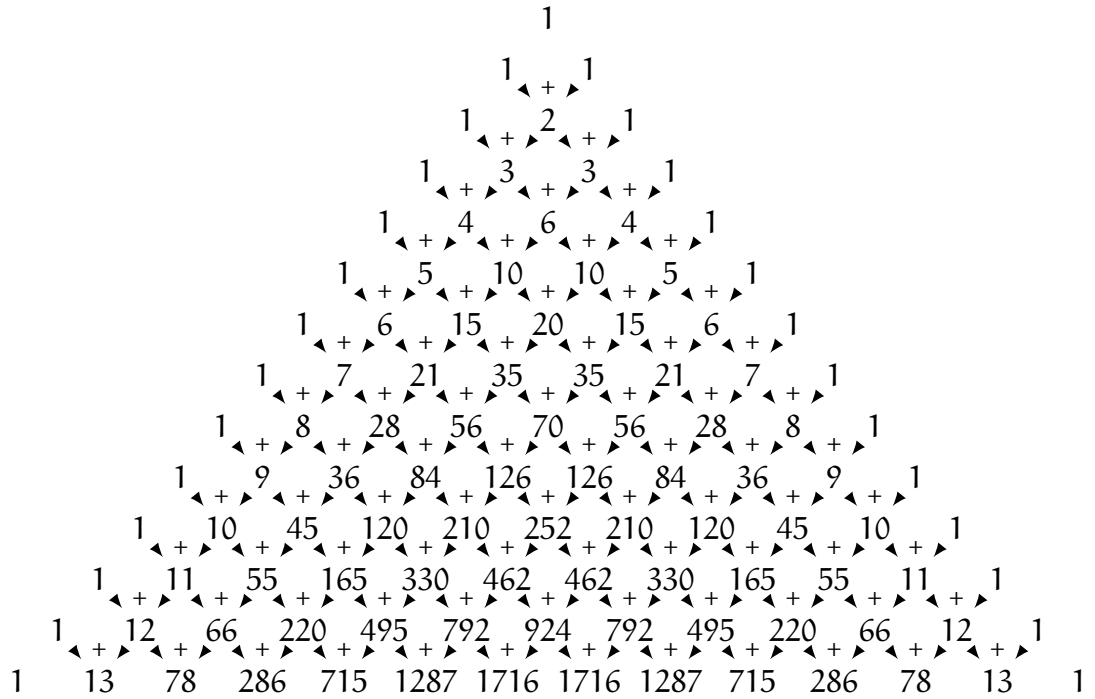


Abbildung 7.1: Das Pascalsche Dreieck. In der $n$ten Zeile stehen die Binomialkoeffizienten $\binom{n}{0}, \ldots, \binom{n}{n}$ (dabei wird die Spitze als nullte Zeile gezählt).

Hier ist $k! = 1 \cdot 2 \cdot \cdots \cdot k$ das Produkt der ersten k natürlichen Zahlen, wobei $0! = 1$ definiert wird (hier versteckt sich wieder eine rekursive Definition!).

Beweis. Sei $n \geqslant 0$ beliebig, und definiere $c(n,k) = \frac{n!}{k!(n-k)!}$. Wir müssen beweisen, dass $c(n,k)$ genau die in Definition 7.8 geforderten Eigenschaften erfüllt.

Wir zeigen zuerst, dass die Bedingungen für $k = 0$ und $k = n$ erfüllt sind. Für $k = 0$ haben wir

$$c(n,0) = \frac{n!}{0!(n-0)!} = \frac{n!}{1 \cdot n!} = 1,$$

und für $k = n$

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

Nun beweisen wir, dass für alle übrigen Fälle, also $0 < k < n$, die Rekursionsvorschrift erfüllt

ist. Durch Erweitern erhalten wir

$$\begin{aligned} c(n-1,k)+c(n-1,k-1) &= \frac{(n-1)!}{k!(n-1-k)!}+\frac{(n-1)!}{(k-1)!(n-1-(k-1))!} \\ &= \frac{(n-1)!(n-k)}{k!(n-k)!}+\frac{(n-1)!k}{k!(n-k)!} = \frac{(n-1)!(n-k+k)}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} = c(n,k). \end{aligned}$$

Also erfüllt $c(n,k) = \frac{n!}{k!(n-k)!}$ genau die Definition von $\binom{n}{k}$, und da rekursive Definitionen eindeutig sind, muss die gewünschte Identität gelten. □

Analog zur starken Induktion kann eine rekursive Definition von $f(n+1)$ auch auf mehrere „Vorgänger“ $f(k)$ für $k \leqslant n$ zurückgreifen. Dabei muss sicher gestellt werden, dass genug Startwerte festgelegt sind, um die Rekursionsvorschrift anzuwenden. Das bekannteste Beispiel sind die *Fibonacci-Zahlen*, definiert durch

$$F(1) = F(2) = 1, \qquad F(n+1) = F(n) + F(n-1) \text{ für } n \geqslant 2.$$

Diese Definition liefert die Zahlenfolge

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

Auch diese Rekursion hat eine explizite Darstellung:

$$F(n) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

(Die Zahl $\frac{1+\sqrt{5}}{2}$ heisst *goldener Schnitt*.) Dass die rechte Seite tatsächlich für alle $n \in \mathbb{N}$ eine natürliche Zahl liefert, ist alles andere als offensichtlich. Man kann aber wieder (wenn auch mit etwas mehr Aufwand) zeigen, dass diese Darstellung die Rekursionsrelation erfüllt.

UNENDLICHE MENGEN

8

Wieviele natürliche Zahlen gibt es? Dies ist eine unsinnige Frage, da als Antwort auf „wieviele“ eine Zahl erwartet wird, aber keine Zahl existiert, so dass alle natürlichen Zahlen kleiner oder gleich sind. („Unendlich“ ist keine Zahl!) Für einen Mathematiker ist das aber unbefriedigend (um David Hilbert zu zitieren: „In der Mathematik gibt es kein Ignorabimus[1].“) Wir müssen also sorgfältiger vorgehen. Statt nach der expliziten Anzahl von Elementen zu fragen, wollen wir erstmal nur entscheiden, ob zwei gegebene Mengen in einem gewissen (noch zu definierenden) Sinn „gleich groß“ sind (wobei dies für endliche Mengen der üblichen Anzahl von Elementen entsprechen soll). Dafür können wir aber ohne Zahlen auskommen: Es genügt, die Elemente beider Mengen paarweise zusammenzufassen; geht die Paarbildung auf, sind die Mengen gleich groß. In diesem Falle können wir also *jedem* Element der ersten Menge ein *eindeutiges* Element der zweiten Menge zuordnen – aus Sicht der Mathematik beschreibt das eine bijektive Funktion zwischen den beiden Mengen. Dies motiviert die folgende Definition:

Definition 8.1. Zwei Mengen A und B heissen *gleichmächtig*, falls es eine bijektive Funktion $f : A \to B$ gibt. In diesem Fall schreiben wir $A \sim B$.

Die Gleichmächtigkeit erfüllt die Eigenschaften einer Äquivalenzrelation:[2]

Satz 8.2. *Seien* A, B, C *Mengen. Dann gilt:*

1. $A \sim A$,
2. $A \sim B$ *und* $B \sim C$ *impliziert* $A \sim C$,
3. $A \sim B$ *genau dann, wenn* $B \sim A$ *gilt.*

[1] lat.: „wir werden nicht wissen“

[2] Gleichmächtigkeit ist aber keine Äquivalenzrelation, weil wir dafür eine Menge angeben müssten, auf welcher sie eine Relation ist. Das wäre dann jedoch die „Menge aller Mengen“, die es aber, wie wir gesehen haben, nicht geben kann.

Beweis. Für jede Menge A ist id_A eine bijektive Abbildung von A nach A. Die Hintereinanderausführung $g \circ f$ eine bijektive Abbildung von A nach C, falls $f : A \to B$ und $g : B \to C$ beide bijektiv sind. Und wenn $f : A \to B$ eine bijektive Abbildung ist, dann existiert die ebenfalls bijektive Umkehrabbildung $f^{-1} : B \to A$. □

Diese Definition kann nun ohne weiteres auch auf unendliche Mengen, und insbesondere auf $\mathbb{N}$, angewendet werden:

Definition 8.3. Sei A eine Menge. Dann heisst A

- *endlich*, falls A leer ist oder ein $n \in \mathbb{N}$ existiert, so dass A gleichmächtig ist zu der Menge $\{1, \ldots, n\}$,
- *unendlich*, falls A nicht endlich ist,
- *abzählbar unendlich*, falls es eine bijektive Funktion $f : \mathbb{N} \to A$ gibt,
- *überabzählbar unendlich*, falls A unendlich ist und es keine bijektive Funktion $f : \mathbb{N} \to A$ gibt.

Wir fassen endliche und abzählbar unendliche Mengen unter dem Begriff *abzählbar* zusammen. Offenbar ist $\mathbb{N}$ wegen Satz 8.2.1 abzählbar unendlich. Dass es wirklich überabzählbar unendliche Mengen gibt, folgt aus Satz 6.9: Es gibt keine Bijektion von $\mathbb{N}$ in die Potenzmenge von $\mathbb{N}$. Da für jedes $n \in \mathbb{N}$ gilt, dass $\{n\} \in \mathcal{P}(\mathbb{N})$ ist, kann $\mathcal{P}(\mathbb{N})$ aber auch nicht endlich sein. Also ist die Potenzmenge von $\mathbb{N}$ überabzählbar unendlich.

Bemerkung. Mit der Gleichmächtigkeit können wir die *relative* Größe von beliebigen Mengen angeben. Es ist nun möglich, auf Basis dieser Definition auch die *absolute* Größe (genannt *Mächtigkeit*) von unendlichen Mengen zu definieren (die für endliche Mengen genau der Anzahl der Elemente entspricht). Man erhält dadurch die *Kardinalzahlen*, für die sich eine äusserst reichhaltige Theorie entwickelt hat.

Wir betrachten nun die drei unendlichen Mengen $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$. Wir beginnen mit $\mathbb{Z}$, und beweisen als Vorbereitung folgendes:

Lemma 8.4. *Die Mengen $\mathbb{N}$ und $\mathbb{N}_0$ sind gleichmächtig.*

Beweis. Wir definieren die Abbildung $f : \mathbb{N} \to \mathbb{N}_0$, $n \mapsto n - 1$. Dann ist f surjektiv, denn für alle $m \in \mathbb{N}_0$ existiert genau ein $n = m + 1 \in \mathbb{N}$ mit $f(n) = m$. Mit der Injektivität von f folgt, dass f bijektiv ist. □

Bei unendlichen Mengen ändert also das Hinzufügen von endlich vielen Elementen nichts an der Abzählbarkeit. Der nächste Schritt ist:

Satz 8.5. *Die Menge $\mathbb{Z}$ ist abzählbar unendlich.*

Beweis. Es kann $\mathbb{Z}$ nicht endlich sein, denn $\mathbb{N} \subseteq \mathbb{Z}$ ist bereits unendlich. Da die Gleichmächtigkeit transitiv ist, und $\mathbb{N}$ und $\mathbb{N}_0$ nach Lemma 8.4 gleichmächtig sind, reicht es, eine Bijektion von $\mathbb{N}_0$ nach $\mathbb{Z}$ anzugeben. Wir definieren $f : \mathbb{N}_0 \to \mathbb{Z}$ durch die Vorschrift

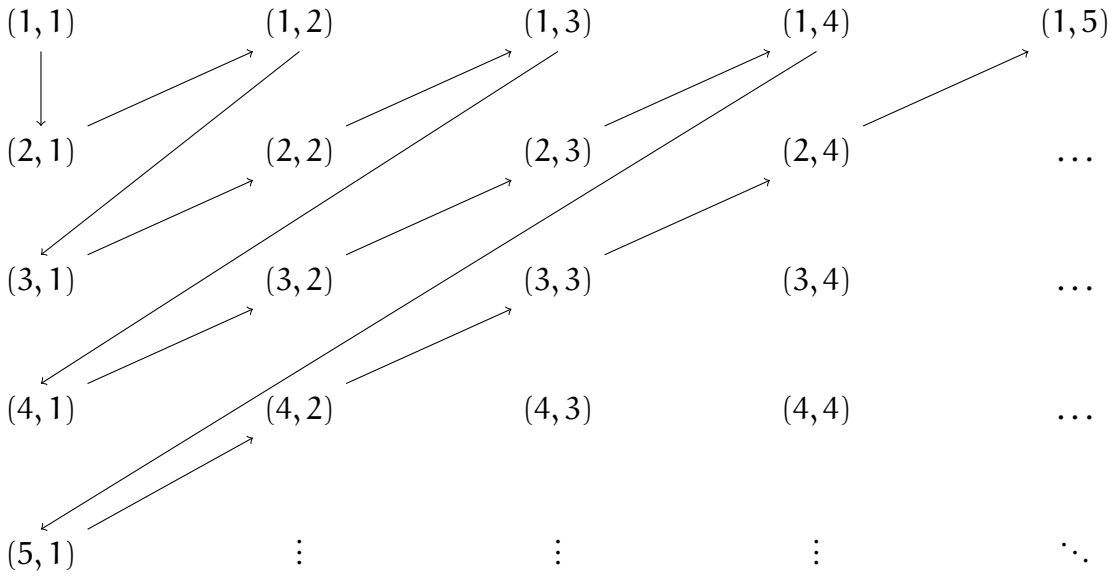
$$f(n) = \begin{cases} k, & \text{falls ein } k \in \mathbb{N}_0 \text{ existiert mit } n = 2k - 1, \\ -k, & \text{falls ein } k \in \mathbb{N}_0 \text{ existiert mit } n = 2k. \end{cases}$$

Dann ist f bijektiv (siehe Satz 6.17). □

Nun zu $\mathbb{Q}$. Da wir jeden Bruch $\frac{n}{m} \in \mathbb{Q}$ als Paar $(n, m) \in \mathbb{Z} \times \mathbb{N}$ auffassen können, und da $\mathbb{Z}$ und $\mathbb{N}$ gleichmächtig sind, betrachten wir zuerst die Menge $\mathbb{N} \times \mathbb{N}$.

Satz 8.6. *Die Menge $\mathbb{N} \times \mathbb{N}$ ist abzählbar unendlich.*

Der Beweis dieser Aussage ist als *Cantors erstes Diagonalargument* bekannt: Wir stellen uns die Paare $(m, n) \in \mathbb{N} \times \mathbb{N}$ in einer (unendlichen) Matrix aufgeschrieben vor, und zählen sie entlang der Diagonalen ab, für die die Summe der beiden Komponenten konstant ist:



Diese Abzählung (hier angedeutet durch die Pfeile) kann man explizit angeben: Es handelt sich um *Cantors Paarungsfunktion*, die wir schon in Lemma 7.3 kennengelernt haben.

Beweis. Wir betrachten wieder die Funktion $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ definiert durch

$$f(a, b) = \frac{(a + b - 2)(a + b - 1)}{2} + b,$$

die nach Lemma 7.3 surjektiv ist. Es bleibt zu zeigen, dass sie injektiv ist. Seien also (a, b), $(c, d) \in \mathbb{N} \times \mathbb{N}$ mit $f(a, b) = f(c, d)$. Wir zeigen $(a, b) = (c, d)$, d.h. $a = c$ und $b = d$.

Wir setzen dazu $n = a + b - 2 \in \mathbb{N}_0$ und $m = c + d - 2 \in \mathbb{N}_0$. Wegen $f(a, b) = f(c, d)$ gilt (nach etwas Umformen)

$$\frac{n(n+1)}{2} - \frac{m(m+1)}{2} = d - b. \tag{8.1}$$

Wir zeigen zunächst $n = m$. Angenommen es ist $n \neq m$. Wir können dann ohne Einschränkung annehmen, dass $n > m$ gilt (sonst vertauschen wir im Folgenden n und m). Wegen $c - 2 \geqslant -1$ ist $m \geqslant d - 1$ und damit $d - b \leqslant d - 1 \leqslant m$. Aus (8.1) erhalten wir

$$n(n+1) - m(m+1) = 2(d-b) \leqslant 2m = m + m < m + n. \tag{8.2}$$

Nun gilt

$$\begin{aligned} n(n+1) - m(m+1) &= n^2 - m^2 + n - m = (n-m)(n+m) + n - m \\ &= (n-m)(n+m+1) \geqslant n + m + 1 \end{aligned}$$

(beachten Sie, dass $n - m \geqslant 1$ gilt, da $n > m$ ist). Aus (8.2) folgt nun der Widerspruch $n + m + 1 < n + m$.

Aus (8.1) und $n = m$ folgt $b = d$. Aus $a + b - 2 = n = m = c + d - 2$ und $b = d$ folgt auch $a = c$. □

Ähnlich geht man nun vor, um zu zeigen, dass $\mathbb{Q}$ abzählbar ist. Dabei muss man beachten, dass $\frac{2}{4} = \frac{1}{2}$ gilt, aber $(2, 4) \neq (1, 2)$ ist – die Paarungsfunktion wäre über $\mathbb{Q}$ also nicht mehr wohldefiniert. Wir müssen die kürzbaren Brüche daher „überspringen“. Dadurch erhalten wir eine Funktion, die wir nicht mehr explizit angeben, aber mit Hilfe des Wohlordnungsprinzips rekursiv definieren können.[3] Sind die positiven Brüche abzählbar, gehen wir für die negativen Brüche (und die Null) dann analog zum Beweis von Satz 8.5 vor. Dies lässt sich mit etwas Aufwand mathematisch sauber beweisen, was wir hier aber nicht tun wollen. Wir halten fest:

Satz 8.7. *Die Menge der rationalen Zahlen $\mathbb{Q}$ ist abzählbar unendlich.*

Mit den gleichen Argumenten kann man allgemein beweisen:

Satz 8.8. *Es gilt:*

1. *Die Produktmenge zweier abzählbar unendlicher Mengen ist abzählbar.*
2. *Die Vereinigung von abzählbar vielen abzählbar unendlichen Mengen ist abzählbar.*

[3] Ein anderer, 2000 von Calkin und Wilf publizierter, eleganter Beweis verwendet eine Darstellung der rationalen Zahlen mit Hilfe eines binären Baums, in dem nur die bereits gekürzten Brüche vorkommen.

Beweisskizze:

1. Seien A und B abzählbar unendliche Mengen. Also existieren bijektive Funktionen $f_1 : \mathbb{N} \to A$ und $f_2 : \mathbb{N} \to B$. Dann ist die Funktion $f : \mathbb{N} \times \mathbb{N} \to A \times B$, $(n_1, n_2) \mapsto (f_1(n_1), f_2(n_2))$ bijektiv. Da $\mathbb{N} \times \mathbb{N}$ abzählbar ist, folgt daraus auch die Abzählbarkeit von $A \times B$.

2. Es sind abzählbar unendlich viele Mengen gegeben, also können wir sie in der Form A_n, $n \in \mathbb{N}$ bezeichnen. Wir wenden wieder Cantors erstes Diagonalargument an: Die Mengen A_n sind alle abzählbar, also kann man sie in der Form $A_n = \{a_{n,1}, a_{n,2}, \dots, \}$ angeben. Die Vereinigungsmenge schreiben wir jetzt als (unendliche) Matrix auf:

$$\begin{array}{ccccc} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & \dots \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & \dots \\ a_{3,1} & a_{1,2} & a_{3,3} & a_{3,4} & \dots \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

   Überspringen wir nun die Elemente $a_{j,k}$, die bereits in einer Menge A_i für $i < j$ enthalten waren, definiert das analog zur Abzählbarkeit von $\mathbb{Q}$ wieder die gewünschte Bijektion.

Ist $\mathbb{R}$ auch abzählbar? Nein, denn selbst das (beschränkte) Intervall $(0, 1]$ kann schon nicht mehr abgezählt werden.

Satz 8.9. *Das Intervall* $(0, 1] = \{x \in \mathbb{R} : 0 < x \leqslant 1\}$ *ist überabzählbar.*

Der Beweis beruht auf der selben Konstruktion wie im Beweis von Satz 6.9, die auch als *Cantors zweites Diagonalargument* bezeichnet wird. Die Fassung, die wir hier angeben, verwendet die eindeutige Dezimaldarstellung der reellen Zahlen: für jede reelle Zahl $x \in (0, 1]$ existiert genau eine Folge $a_1, \dots$ mit $a_n \in \{0, 1, \dots, 9\}$ für alle $n \in \mathbb{N}$, so dass man x schreiben kann als

$$x = 0.\, a_1\, a_2\, a_3\, \dots,$$

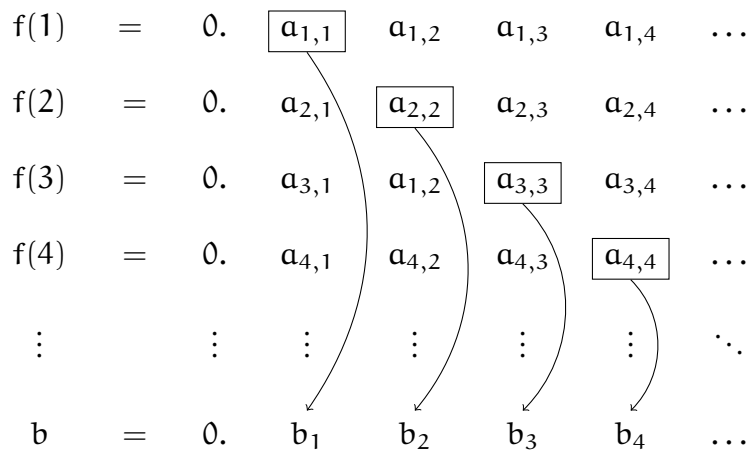
wobei es kein $n \in \mathbb{N}$ gibt mit $a_k = 0$ für alle $k > n$. Die letzte Bedingung gewährleistet die Eindeutigkeit, da ja bekanntlich jede abbrechende Dezimalzahl äquivalent mit einer Neunerperiode geschrieben werden kann. (Wir schreiben also zum Beispiel $0.4999\dots$ statt 0.5.)

Beweis. Da für jedes $n \in \mathbb{N}$ gilt, dass $\frac{1}{n} \in (0, 1]$ ist, kann $(0, 1]$ nicht endlich sein. Wir müssen nun zeigen, dass $(0, 1]$ nicht abzählbar unendlich ist, dass es also keine bijektive Abbildung von $\mathbb{N}$ nach $(0, 1]$ geben kann. Sei nun $f : \mathbb{N} \to (0, 1]$ gegeben, wobei wir für $n \in \mathbb{N}$

wieder $f(n) = 0.\,a_{n,1}\,a_{n,2}\,a_{n,3}\,\ldots$ in Dezimaldarstellung schreiben. Wir konstruieren ein $b \in (0,1]$, das nicht im Bild von f liegt. Definiere b_n für alle $n \in \mathbb{N}$ als

$$b_n = \begin{cases} 2 & \text{falls } a_{n,n} = 1 \\ 1 & \text{falls } a_{n,n} \neq 1 \end{cases}$$

und setze $b = 0.\,b_1\,b_2\,b_3\,\ldots$. Denken wir uns die $f(n)$ in Dezimaldarstellung der Reihe nach untereinander geschrieben, wird das Diagonalargument deutlicher:



Da b eine nicht abbrechende Dezimalzahl ist, gilt $b \in \mathbb{R}$ und insbesondere $b \in (0,1]$. Andererseits existiert kein $n \in \mathbb{N}$, so dass $f(n) = b$ gilt, denn für alle $n \in \mathbb{N}$ ist nach Konstruktion $b_n \neq a_{n,n}$, also unterscheiden sich b und $f(n)$ in der $n$ten Dezimalstelle, und die Dezimaldarstellung ist eindeutig. Die Funktion f ist deshalb nicht surjektiv, und damit auch nicht bijektiv. □

Also ist insbesondere $\mathbb{R}$ überabzählbar. Als weiteren Beleg, dass man Gleichmächtigkeit nicht einfach mit einer naiven Anschauung, die auf der Anzahl der Elemente von endlichen Mengen oder Dimensionsbegriffen basiert, gleichsetzen kann, sei folgende Tatsache angeführt: $\mathbb{R}^2$ und $\mathbb{R}$ sind gleichmächtig. Wieder reicht es, eine bijektive Abbildung zwischen dem Quadrat $(0,1]^2$ und dem Intervall $(0,1]$ anzugeben: Sei $(x,y) \in (0,1]^2$ mit der nicht-abbrechenden Dezimaldarstellung

$$x = 0.\,x_1\,x_2\,x_3\,\ldots,$$
$$y = 0.\,y_1\,y_2\,y_3\,\ldots,$$

gegeben. Um die Abbildungsvorschrift nach $(0,1]$ zu definieren, gruppieren wir die Dezimalstellen von x (und analog von y) in Blöcke $X_j = x_{j1}\,x_{j2}\,x_{j3}\,\ldots\,x_{jk_j}$ mit $k_j \geqslant 1$ und $x_{j1} = x_{j2} = \cdots = x_{j,k_j-1} = 0, x_{jk_j} \neq 0$ (wir gehen also immer bis einschließlich zur ersten

von Null verschiedenen Dezimalstelle). Die gesuchte Abbildung erhalten wir dann durch die Zuordnung

$$f(x, y) = 0.\,X_1\,Y_1\,X_2\,Y_2\,X_3\,Y_3\ldots\,.$$

(Würden wir die Dezimalstellen einfach abwechseln, hätte zum Beispiel $0.101010\ldots$ kein Urbild.) Das ist wieder eine eindeutige reelle Zahl in $(0, 1]$, und für gegebenes $z \in (0, 1]$ kann man aus der Dezimaldarstellung sofort die des Urbilds $f^{-1}(z) \in (0, 1]^2$ ablesen. So ist beispielsweise das Urbild $(x, y) = f^{-1}(z)$ von

$$z = 0.3\,009\,01\,2\,2\,05\,007\,1\,08\,0008\ldots$$

gegeben durch

$$\begin{aligned} x &= 0.\quad 3\quad 01\quad 2\quad 007\quad 08\quad \ldots, \\ y &= 0.\quad 009\quad 2\quad 05\quad 1\quad 0008\quad \ldots\,. \end{aligned}$$

Um aus $\mathbb{R}$ eine „mächtigere“ Menge zu konstruieren, müssen wir daher wieder die Potenzmenge bemühen – mit der sich die Konstruktion wiederholen lässt …

Teil IV

ZAHLBEGRIFFE

ÜBERBLICK

Das Zählen gehört zu den elementaren Erfordernissen der Interaktion mit Mitmenschen und Umwelt. Man sagt: das sind fünf Früchte, oder: es gibt acht Planeten. Wird die Anzahl von den zu zählenden Objekten abstrahiert – Was haben fünf Früchte und fünf Planeten gemein? – führt diese Abstraktion zu den „natürlichen Zahlen“ $1, 2, 3, \ldots$.

Oft ist man aber nicht an den Zahlen an sich, sondern an dem Vergleich zweier Zahlen interessiert: Wieviel mehr Früchte hast Du als ich? Dabei fand man die Nützlichkeit nicht-positiver Zahlen (inklusive 0). Auf ähnliche Weise führte die Untersuchung von Verhältnissen von Strecken auf Brüche von ganzen Zahlen. Dann entdeckte man, dass nicht alle Streckenlängen sich so darstellen lassen, und stiess auf weitere mathematisch oder anwendungsbedingte Notwendigkeiten, das System der Zahlen zu ergänzen – wieder zunächst intuitiv, als (fiktive) Hilfsgrößen in Berechnungen.

Als die Mathematik mit der axiomatischen Mengenlehre neu begründet wurde, mussten diese Zahlbegriffe, und vor allem die Existenz der Menge der natürlichen Zahlen, gesichert werden. Ausgehend von den natürlichen Zahlen kann man, wie in diesem Kapitel geschildert, sukzessive eine Konstruktion der ganzen, rationalen, reellen und komplexen Zahlen durchführen, und zwar mit Hilfe von Äquivalenzklassen. Dabei wird jeweils auf die gewünschten algebraischen Eigenschaften der Addition und der Multiplikation geachtet sowie im Fall der reellen Zahlen auf die Vollständigkeit (in einer für die moderne Analysis wichtigen Form). Auf die Konstruktion der natürlichen Zahlen mit Hilfe der axiomatischen Mengenlehre wird im Anhang kurz eingegangen. Auch die Darstellungen in den folgenden Kapiteln ist nur skizziert; die (zumeist einfachen) Beweise bleiben dem Leser als Übung überlassen.

Im mathematischen Alltag geht man allerdings selten auf diese Konstruktionen zurück, sondern beruft sich auf eine axiomatische Definition der jeweiligen Zahlenmengen anhand ihrer charakterisierenden Eigenschaften.

DIE NATÜRLICHEN ZAHLEN

9

Eine wesentliche Aufgabe der Grundlegung der Mathematik besteht darin, die Existenz der Menge der natürlichen Zahlen aus den Axiomen der Mengenlehre abzuleiten. Im Rahmen dieser Vorlesung gehen wir aber der Einfachheit halber von der Existenz der natürlichen Zahlen aus[1] und schreiben wie bisher $\mathbb{N}$ für die Menge $\{1, 2, 3, \dots\}$ der natürlichen Zahlen.

Die natürlichen Zahlen sollen folgende Eigenschaften erfüllen:

- Es gibt eine *Addition* $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ und eine *Multiplikation* $\cdot : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Beide Operationen sind für sich allein *assoziativ* und *kommutativ*, und gemeinsam erfüllen sie das *Distributivgesetz*

  $$\forall n, m, k \in \mathbb{N} : n(m + k) = (n \cdot m) + (n \cdot k)$$

  Ausserdem darf man bezüglich Addition und Multiplikation kürzen, das heisst aus $m + k = n + k$ oder $m \cdot k = n \cdot k$ folgt $m = n$.

- Es ist eine Ordnungsrelation $\leqslant$ gegeben, nämlich durch

  $$\begin{aligned} n < m &:\Leftrightarrow \exists k \in \mathbb{N} : n + k = m, \\ n \leqslant m &:\Leftrightarrow (n < m) \vee (n = m). \end{aligned} \tag{9.1}$$

  Die Addition und Multiplikation sind monoton bezüglich der Relationen $<$ und $\leqslant$, das heisst für alle $n, m, k \in \mathbb{N}$ gilt

  $$n < m \Leftrightarrow n + k < m + k \Leftrightarrow n \cdot k < m \cdot k, \tag{9.2}$$

  und analog für $\leqslant$.

- Schließlich gilt in $\mathbb{N}$ das in Abschnitt 7 besprochene Prinzip der vollständigen Induktion.

[1] Frei nach dem Motto: „Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk." (Leopold Kronecker). Die Konstruktion der natürlichen Zahlen wird in Anhang A skizziert.

DIE GANZEN ZAHLEN

10

Eine wünschenswerte Operation, nämlich die Subtraktion, kann auf $\mathbb{N}$ nur bedingt definiert werden: Falls $n \leqslant m$ gilt, ist $n-m$ keine natürliche Zahl. Statt nun die Null und negative ganze Zahlen und dazu weitere Axiome über die Subtraktion hinzuzufügen, ist es eleganter und ohne zusätzliche Annahmen möglich, mit Paaren natürlicher Zahlen zu arbeiten. Jede positive oder negative ganze Zahl lässt sich nämlich als Differenz zweier natürlicher Zahlen darstellen, doch sind diese nicht eindeutig bestimmt. Deshalb muss man mit Äquivalenzklassen arbeiten.

10.1 KONSTRUKTION VON $\mathbb{Z}$

Wir möchten also diejenigen Zahlenpaare identifizieren, die die gleiche Differenz haben – ohne die Subtraktion zu verwenden, die wir ja erst einführen möchten. Die Identität zweier Differenzen lässt sich aber so umformen, dass wir nur die Addition von natürlichen Zahlen benötigen. So erhalten wir die folgende, bereits in Beispiel 4.4.4 eingeführte, Relation auf $\mathbb{N} \times \mathbb{N}$: Für $n, m, k, j \in \mathbb{N}$ sei

$$(n, m) \sim (k, j) :\Leftrightarrow n + j = m + k. \tag{10.1}$$

Die *ganzen Zahlen* werden nun definiert als Äquivalenzklassen der Relation $\sim$,

$$[(n, m)]_\sim = \{(k, j) \in \mathbb{N} \times \mathbb{N} : (n, m) \sim (k, j)\}.$$

Der Übersichtlichkeit halber schreiben wir im Folgenden einfach $[(n, m)]$. Die Menge der ganzen Zahlen $\mathbb{Z}$ ist dann die Quotientenmenge von $\mathbb{N} \times \mathbb{N}$ unter $\sim$, also die Menge aller Äquivalenzklassen bezüglich der Relation $\sim$:

$$\mathbb{Z} := (\mathbb{N} \times \mathbb{N})/_\sim = \{[(n, m)] : n, m \in \mathbb{N}\}$$

In Abbildung 10.1 sind die Paare in $\mathbb{N} \times \mathbb{N}$ als kleine Kreise eingezeichnet, die in einer Äquivalenzklasse liegenden Paare sind mit durchgezogenen Geraden verbunden. Die Gerade

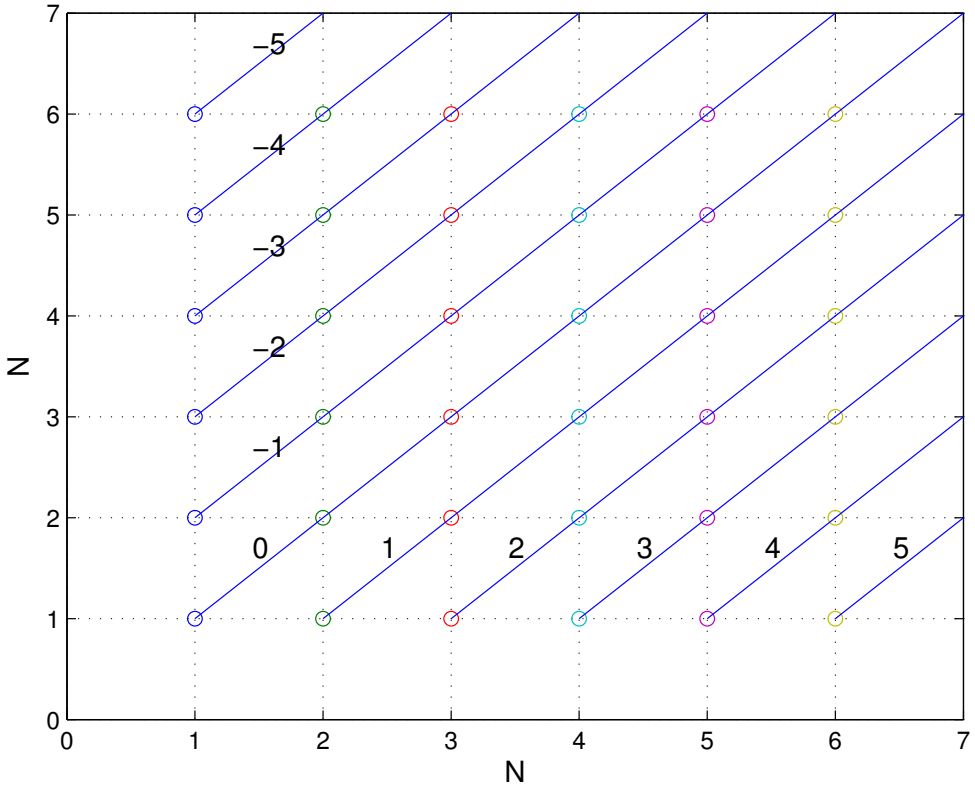


Abbildung 10.1: Graphische Darstellung der ganzen Zahlen als Äquivalenzklassen auf $\mathbb{N} \times \mathbb{N}$.

zur ganzen Zahl 0 geht durch $(1, 1)$, unterhalb davon liegen die Klassen der positiven Zahlen, oberhalb die der negativen ganzen Zahlen.

Zwischen den Äquivalenzklassen wird nun eine Addition, eine Multiplikation und eine Relation „kleiner als“ eingeführt. Wir verwenden dafür der Einfachheit halber wieder die Zeichen $+, \cdot, <$, müssen uns aber bewusst sein, dass damit andere Verknüpfungen und eine andere Relation gemeint sind. Insbesondere kann man etwa nicht eine natürliche Zahl und eine Äquivalenzklasse, die aus Paaren natürlicher Zahlen besteht, addieren. Auf die Einbettung von $\mathbb{N}$ in die Menge der ganzen Zahlen kommen wir am Ende des Kapitels zurück.

10.2 ARITHMETISCHE OPERATIONEN AUF $\mathbb{Z}$

Im Folgenden seien immer $[(n, m)], [(k, j)] \in \mathbb{Z}$. Die *Addition* auf $\mathbb{Z}$ ist die komponentenweise Addition von Paaren:

$$[(n, m)] + [(k, j)] := [(n + k, m + j)]$$

Die *Multiplikation* wird so definiert, dass das Distributivgesetz gilt. Insbesondere muss es für das Produkt von Differenzen, $(n - m) \cdot (k - j) = (nk + mj) - (mk + nj)$, gelten:

$$[(n, m)] \cdot [(k, j)] := [(nk + mj), (mk + nj)]$$

Aus Assoziativität, Kommutativität und Distributivität der Operationen auf $\mathbb{N}$ folgert man leicht die entsprechenden Rechengesetze auf $\mathbb{Z}$.

Für beide Operationen gibt es ein *neutrales Element* in $\mathbb{Z}$, dass bei Addition beziehungsweise Multiplikation keine Wirkung hat:

$$\widehat{0} := [(1, 1)] \quad \text{(neutrales Element der Addition)}$$
$$\widehat{1} := [(2, 1)] \quad \text{(neutrales Element der Multiplikation)}$$

Die *additiv inverse* Zahl zu $[(n, m)]$ ist dasjenige Element in $\mathbb{Z}$, das zu $[(n, m)]$ addiert $\widehat{0}$ ergibt (anschaulich: umgekehrtes Vorzeichen hat, vergleiche Abbildung 10.1):

$$-[(n, m)] := [(m, n)]$$

Damit ist klar, dass für alle Elemente in $\mathbb{Z}$ eine additiv inverse Zahl existiert. Die *Subtraktion* ist jetzt als Addition mit der inversen Zahl definiert:

$$[(n, m)] - [(k, j)] := [(n, m)] + (-[(k, j)])$$

Schliesslich definiert man die *Ordnung* so, dass $n - m$ kleiner ist als $k - j$, falls $n + j$ kleiner als $m + k$:

$$[(n, m)] < [(k, j)] :\Leftrightarrow n + j < m + k$$

Um die Sinnhaftigkeit dieser Definitionen für $+, \cdot, <$ auf $\mathbb{Z}$ nachzuweisen, muss gezeigt werden, dass die rechte Seite jeweils unabhängig von der speziellen Wahl von Repräsentanten der Äquivalenzklassen sind. (Diese Bedingung nennt man *Wohldefiniertheit*.) Wir beweisen dies beispielhaft für die Addition. Seien etwa

$$(n, m) \sim (\overline{n}, \overline{m}) \quad \text{und} \quad (k, j) \sim (\overline{k}, \overline{j}) \in \mathbb{N} \times \mathbb{N}.$$

je zwei Repräsentanten zweier Äquivalenzklassen. Dann gilt also

$$n + \overline{m} = m + \overline{n} \quad \text{und} \quad k + \overline{j} = j + \overline{k}.$$

Addieren wir beide Gleichungen, erhalten wir

$$(n + k) + (\overline{m} + \overline{j}) = (m + j) + (\overline{n} + \overline{k}),$$

und daher nach Definition von $\sim$, dass

$$(n + k, m + j) \sim (\overline{n} + \overline{k}, \overline{m} + \overline{j}).$$

Also ist $[(n + k, m + j)] = [(\overline{n} + \overline{k}, \overline{m} + \overline{j})]$, was zu zeigen war.

Wir betrachten nun die Identifikation der natürlichen Zahlen mit ganzen Zahlen. Die injektive Funktion $j : \mathbb{N} \to \mathbb{Z}$,

$$j(n) = [(n + 1, 1)]$$

ordnet jeder natürlichen n Zahl eine positive (das heisst $j(n) > \widehat{0}$) ganze Zahl zu und ist *verknüpfungstreu* (das heisst $j(n + m) = j(n) + j(m)$ und $j(n \cdot m) = j(n) \cdot j(m)$). Spricht man daher von der Summe einer natürlichen Zahl n und einer ganzen Zahl m, meint man damit immer die (ganze) Zahl $j(n) + m$.

Bemerkung. Die Verknüpfung $+$ ordnet je zwei Elementen a, b aus $\mathbb{Z}$ genau ein Element $a + b$ aus $\mathbb{Z}$ zu und ist assoziativ. Es gibt ein neutrales Element $\widehat{0}$ mit $a + \widehat{0} = a$ für alle $a \in \mathbb{Z}$. Zu jedem $a \in \mathbb{Z}$ gibt es ein inverses Element, nämlich $-a$, so dass $a + (-a) = \widehat{0}$. Eine algebraische Struktur $(\mathbb{Z}, +)$ mit diesen Eigenschaften nennt man eine *Gruppe*. Man nennt die Gruppe *kommutativ*, wenn die Verknüpfung kommutativ ist. Ferner ist auch die Verknüpfung $\cdot$ eine innere Verknüpfung auf $\mathbb{Z}$, die assoziativ und kommutativ ist, und sie erfüllt zusammen mit $+$ das Distributivgesetz. Daher nennt man die Struktur $(\mathbb{Z}, +, \cdot)$ einen *kommutativen Ring*.

Allerdings gibt es in $\mathbb{Z}$ nicht für alle Zahlen *multiplikative* Inverse (und damit die Möglichkeit der Division), und dies motiviert die Einführung von Brüchen und damit die Erweiterung zu den rationalen Zahlen.

DIE RATIONALEN ZAHLEN

11

Die Division als Umkehrung der Multiplikation ist in $\mathbb{Z}$ nicht ausführbar. Ähnlich wie bei der Erweiterung der natürlichen Zahlen zu den ganzen Zahlen werden zur Ermöglichung der Division Äquivalenzklassen von Paaren ganzer Zahlen gebildet.

11.1 BRÜCHE UND RATIONALE ZAHLEN

Es liegt nahe, rationale Zahlen als Paar von Zähler und Nenner darzustellen. Allerdings müssen wir nun beachten, dass zwei verschiedene Brüche die selbe rationale Zahl darstellen können (nämlich wenn beide durch Kürzen auf den selben Bruch zurückgeführt werden können). Wir wollen solche Paare also identifizieren, und definieren daher die folgende Relation.

Definition 11.1. Für $a, c \in \mathbb{Z}$ und $b, d \in \mathbb{Z} \setminus \{\widehat{0}\}$ sei

$$(a, b) \equiv (c, d) :\Leftrightarrow ad = bc.$$

Beachten Sie, dass diese Definition und die Definition von $\sim$ in (10.1) formal sehr ähnlich sind: es ist nur $+$ durch $\cdot$ ersetzt.

Durch $\equiv$ wird eine Äquivalenzrelation auf $\mathbb{Z} \times (\mathbb{Z} \setminus \{\widehat{0}\})$ definiert. Deren Äquivalenzklassen werden *rationale Zahlen* (manchmal auch *Bruchzahlen*) genannt (siehe Abbildung 11.1): Für $(a, b) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{\widehat{0}\})$ definieren wir den *Bruch* $\frac{a}{b}$ als die Äquivalenzklasse von (a, b):[1]

$$\frac{a}{b} := [(a, b)]_{\equiv} = \left\{(c, d) \in \mathbb{Z} \times (\mathbb{Z} \setminus \{\widehat{0}\}) \,:\, (a, b) \equiv (c, d)\right\}.$$

[1] Streng genommen muss man immer zwischen der *rationalen Zahl* $\frac{a}{b}$ (also der Äquivalenzklasse $[(a, b)]$) und dem *Bruch* mit Zähler a und Nenner b (also dem speziellen Repräsentanten (a, b)) unterscheiden. Dies geschieht aber meistens nicht explizit, wenn aus dem Kontext klar ist, ob Äquivalenzklasse oder Repräsentant gemeint ist.

Analog zu $\mathbb{Z}$ definieren wir die Menge der rationalen Zahlen

$$\mathbb{Q} := \left(\mathbb{Z} \times (\mathbb{Z} \setminus \{\hat{0}\})\right) /_{\equiv} = \left\{ \frac{a}{b} \,:\, a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{\hat{0}\} \right\}.$$

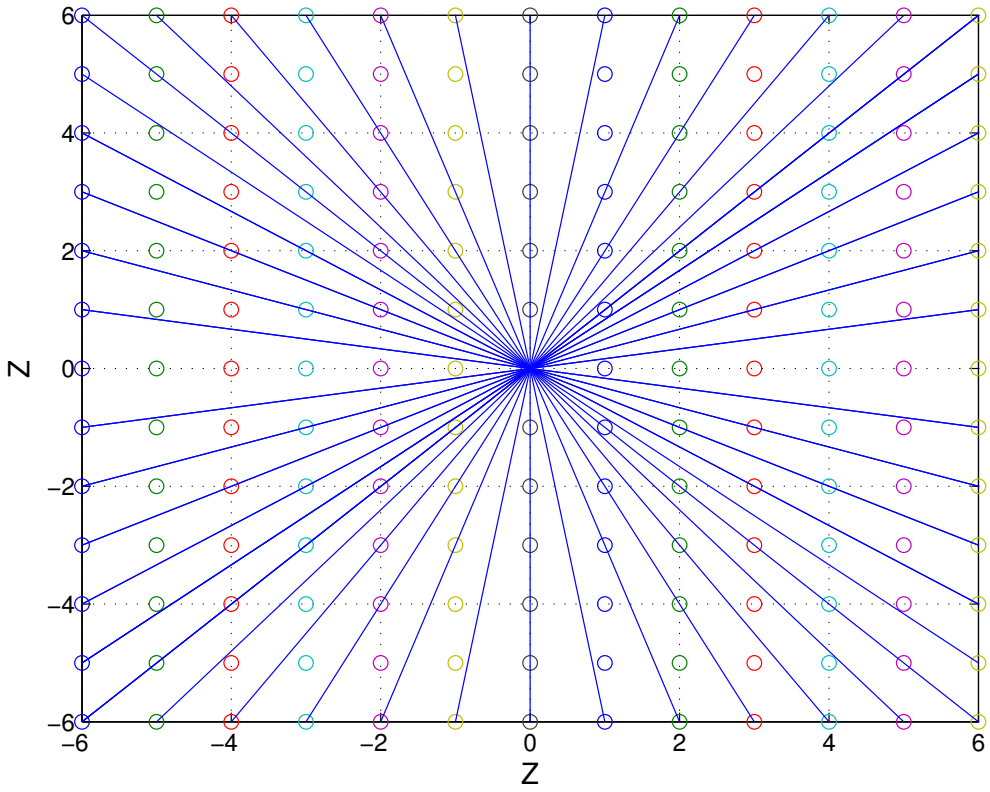


Abbildung 11.1: Graphische Darstellung der rationalen Zahlen als Äquivalenzklassen auf $\mathbb{Z} \times \mathbb{Z} \setminus \{\hat{0}\}$. (Der Übersichtlichkeit halber sind einige Äquivalenzklassen ausgelassen.)

Wir führen als nächstes auf $\mathbb{Q}$ eine Addition, eine Multiplikation und eine Relation „kleiner als“ ein und verwenden dafür wieder die Zeichen $+, \cdot, <$.

11.2 RECHNEN MIT BRÜCHEN

Im Folgenden seien $a, c \in \mathbb{Z}$ und $b, d \in \mathbb{Z} \setminus \{\hat{0}\}$. Die *Addition*, *additive Inverse*, *Subtraktion* und *Multiplikation* definieren wir über die entsprechende Operation in $\mathbb{Z}$:

$$\frac{a}{b} + \frac{c}{d} := \frac{ad + bc}{bd}, \quad -\frac{a}{b} := \frac{-a}{b}, \quad \frac{a}{b} - \frac{c}{d} := \frac{a}{b} + \left(-\frac{c}{d}\right), \quad \frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd}.$$

Das *neutrale Element* der Addition beziehungsweise Multiplikation ist definiert als

$$\overline{0} := \frac{\hat{0}}{\hat{1}}, \quad \overline{1} := \frac{\hat{1}}{\hat{1}}.$$

Ist $a \neq \hat{0}$, können wir auch ein *multiplikativ inverses Element* definieren als

$$\left(\frac{a}{b}\right)^{-1} := \frac{b}{a},$$

und damit die *Division* durch

$$\frac{c}{d} \Big/ \frac{a}{b} := \frac{c}{d} \cdot \left(\frac{a}{b}\right)^{-1}.$$

Die *Ordnung* auf $\mathbb{Q}$ ist gegeben durch

$$\frac{a}{b} < \frac{c}{d} :\Leftrightarrow \begin{cases} ad < bc, \text{ falls } \hat{0} < b, d \text{ oder } b, d < \hat{0}, \\ bc < ad, \text{ sonst,} \end{cases}$$

und analog zu (9.1) wird daraus die Ordnungsrelation $\leqslant$ definiert. Auch in $\mathbb{Q}$ sind Addition und Multiplikation monoton bezüglich dieser Ordnung. Mit dieser Ordnung definieren wir auch den *Betrag* einer rationalen Zahl q als

$$|q| = \begin{cases} q & \text{falls } q \geqslant 0, \\ -q & \text{falls } q < 0. \end{cases}$$

Wieder müssen wir nachweisen, dass die rechten Seiten unabhängig sind von der speziellen Wahl der Repräsentanten von $\frac{a}{b}$ und $\frac{c}{d}$. Um die Wohldefiniertheit der Addition in $\mathbb{Q}$ zu zeigen, seien etwa $(a, b) \equiv (\overline{a}, \overline{b})$ und $(c, d) \equiv (\overline{c}, \overline{d})$. Distributivität und Kommutativität der Operationen auf $\mathbb{Z}$ sowie Anwendung von $a\overline{b} = b\overline{a}$ und $c\overline{d} = d\overline{c}$ liefern

$$\begin{aligned}(ad + bc)\overline{bd} &= ad\overline{bd} + bc\overline{bd} = a\overline{b}d\overline{d} + b\overline{b}c\overline{d} = b\overline{a}d\overline{d} + b\overline{b}d\overline{c} = bd\overline{a}\overline{d} + bd\overline{b}\overline{c} \\ &= bd(\overline{a}\overline{d} + \overline{b}\overline{c}),\end{aligned}$$

und daher nach Definition der Äquivalenzklassen

$$\frac{ad + bc}{bd} = \frac{\overline{a}\overline{d} + \overline{b}\overline{c}}{\overline{bd}}.$$

Da die arithmetischen Operationen in $\mathbb{Q}$ über die Operationen in $\mathbb{Z}$ definiert sind, kann man Assoziativität, Kommutativität und Distributivität der Operationen in $\mathbb{Q}$ unter der Verwendung der entsprechenden Eigenschaften in $\mathbb{Z}$ zeigen. Ausserdem gibt es ähnlich wie beim Übergang von $\mathbb{N}$ nach $\mathbb{Z}$ eine Einbettung von $\mathbb{Z}$ in $\mathbb{Q}$, vermittelt durch die injektive, verknüpfungstreue Funktion $j : \mathbb{Z} \to \mathbb{Q}$, $a \mapsto \frac{a}{\hat{1}}$.

DIE REELLEN ZAHLEN

12

Wie lang ist die Diagonale eines Quadrates, dessen Seitenlänge 1 ist? Nach dem Satz von Pythagoras ist die quadrierte Diagonallänge 2. Satz 6.8 besagt aber, dass es keine rationale Zahl gibt, deren Quadrat 2 ist. Also gibt es in $\mathbb{Q}$ gewissermaßen Lücken, man sagt, $\mathbb{Q}$ ist nicht vollständig. Wir ergänzen daher die Menge $\mathbb{Q}$ zu den reellen Zahlen $\mathbb{R}$. Unterschiedliche Charakterisierungen dieser Lücken führen zu unterschiedlichen, aber beweisbar äquivalenten Konstruktionen von $\mathbb{R}$. Wir skizzieren hier einen (von Cantor 1883 vorgestellten) Ansatz, der auch die Grundlage für die Vervollständigung sehr viel allgemeinerer Strukturen darstellt.[1] Dabei werden die reellen Zahlen als Äquivalenzklassen von (rationalen) Cauchy-Folgen aufgefasst. Wir setzen hier Grundkenntnisse über (reelle) Folgen und ihren Kalkül voraus.

12.1 FUNDAMENTALFOLGEN

Unter einer *rationalen Folge* verstehen wir eine Funktion $r : \mathbb{N} \to \mathbb{Q}$, wobei wir für die Funktionswerte (oder *Folgenglieder*) kurz $r_n := r(n)$ und für die Folge selber $(r_n)_{n\in\mathbb{N}}$ oder auch nur (r_n) schreiben wollen. Rationale Folgen, die Cauchy-Folgen im Sinne der Analysis sind, bezeichnen wir als Fundamentalfolgen:

Definition 12.1. Eine rationale Folge (r_n) heisst *Fundamentalfolge*, falls für alle rationalen $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $|r_m - r_n| < \varepsilon$ für alle $m, n \geqslant n_0$ gilt.

Die Glieder einer Fundamentalfolge liegen also mit steigendem Index immer dichter beieinander. Mit solchen Folgen können wir den „Lücken“ in $\mathbb{Q}$ beliebig nahe kommen, sie aber nie erreichen. Beispielsweise bilden die abgebrochenen Dezimaldarstellungen

$$r_1 = 1, \quad r_2 = 1.4, \quad r_3 = 1.41, \quad r_4 = 1.414, \quad r_5 = 1.4142, \ldots$$

eine Fundamentalfolge, die sich $\sqrt{2} \notin \mathbb{Q}$ annähert. Die Idee ist nun, die Zahl $\sqrt{2}$ durch diese Folge darzustellen (den Grenzwert haben wir ja nicht zur Verfügung, da er als reelle Zahl

[1] Alternativen sind Intervallschachtelungen oder die Dedekindschen Schnitte.

gerade erst konstruiert werden soll). Dabei müssen wir berücksichtigen, dass wir uns auf viele Arten dieser Lücke annähern können (etwa auch durch die rekursiv definierte Folge $a_1 = 1$, $a_{n+1} = a_n/2 + 1/a_n$). Wir müssen also wieder Äquivalenzklassen betrachten.

Dafür führen wir zuerst die Menge $\mathcal{F}$ der Fundamentalfolgen ein. Auf $\mathcal{F}$ kann man nun Addition und Multiplikation definieren: Für zwei Fundamentalfolgen (r_n) und (s_n) sei

$$(r_n) + (s_n) := (r_n + s_n), \qquad (r_n) \cdot (s_n) := (r_n \cdot s_n).$$

Durch einfache Abschätzung vergewissert man sich, dass dies innere Verknüpfungen sind. Auch eine Ordnungsrelation können wir definieren: Für $(r_n), (s_n) \in \mathcal{F}$ sei $(r_n) < (s_n)$, falls es ein rationales $\varepsilon > 0$ und ein $n_0 \in \mathbb{N}$ gibt, so dass $r_n + \varepsilon < s_n$ ist für alle $n \geqslant n_0$. (Anschaulich bedeutet dies, dass wir die Folgenglieder von (r_n) und (s_n) – bis auf endlich viele – durch eine rationale Zahl voneinander trennen können.) Die Ordnungsrelation $\leqslant$ wird dann wieder wie in (9.1) festgelegt.

12.2 DIE REELLEN ZAHLEN ALS ÄQUIVALENZKLASSEN

Wir möchten nun diejenigen Folgen identifizieren, die dieselbe Lücke annähern – wobei wir wieder nicht auf den erst zu konstruierenden Grenzwert zurückgreifen können. Wenn zwei Folgen aber der selben Lücke immer näher kommen, müssen sie sich dabei auch einander näher kommen. Das motiviert die folgende Definition:

Definition 12.2. Eine rationale Folge (r_n) heisst *Nullfolge*, falls für alle rationalen $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $|r_n| < \varepsilon$ für alle $n \geqslant n_0$ gilt.

Die Menge aller Nullfolgen bezeichnen wir mit $\mathcal{N}$. Jede Nullfolge ist eine Fundamentalfolge, und Summe und Produkt von Nullfolgen sind auch Nullfolgen. Wir betrachten nun solche Fundamentalfolgen als äquivalent, die sich nur um eine Nullfolge unterscheiden: Für $(r_n), (s_n) \in \mathcal{F}$ wird durch

$$(r_n) \triangleq (s_n) :\Leftrightarrow \exists (t_n) \in \mathcal{N} : (r_n) = (s_n) + (t_n)$$

eine Äquivalenzrelation auf $\mathcal{F}$ definiert. Die reellen Zahlen $\mathbb{R}$ sind dann die Äquivalenzklassen bezüglich dieser Relation:

$$\mathbb{R} := \mathcal{F}/_{\triangleq} = \left\{ [(r_n)]_{\triangleq} \,:\, (r_n) \in \mathcal{F} \right\}.$$

Die Addition, Multiplikation und Ordnung auf $\mathbb{R}$ definieren wir wieder über die entsprechenden Begriffe in $\mathcal{F}$:

$$\begin{aligned} [(r_n)]_{\triangleq} + [(s_n)]_{\triangleq} &:= [(r_n) + (s_n)]_{\triangleq}, \\ [(r_n)]_{\triangleq} \cdot [(s_n)]_{\triangleq} &:= [(r_n) \cdot (s_n)]_{\triangleq}, \\ [(r_n)]_{\triangleq} < [(s_n)]_{\triangleq} &:\Leftrightarrow (r_n) < (s_n), \end{aligned}$$

wobei man auch hier nachweisen muss, dass die rechte Seite unabhängig von der Wahl der Repräsentaten (r_n) und (s_n) ist.

Die injektive, verknüpfungstreue Funktion $j : \mathbb{Q} \to \mathbb{R}$, $q \mapsto [(q)_{n\in\mathbb{N}}]_{\triangleq}$, vermittelt wieder die Einbettung von $\mathbb{Q}$ in $\mathbb{R}$. (Hier bezeichnet $(q)_{n\in\mathbb{N}}$ die konstante Folge $(q, q, q, \dots)$.) Das neutrale Element der Addition ist die Äquivalenzklasse der konstanten Folge $(\overline{0})_{n\in\mathbb{N}}$ – und damit genau die Menge $\mathcal{N}$ der Nullfolgen, – das der Multiplikation $[(\overline{1})_{n\in\mathbb{N}}]_{\triangleq}$. Analog zu $\mathbb{Q}$ können wir dann den Betrag einer reellen Zahl mit Hilfe der Ordnung und einer Fallunterscheidung einführen.

Die additive Inverse führt man wieder auf diejenige in $\mathbb{Q}$ zurück: $-[(r_n)]_{\triangleq} := [(-r_n)]_{\triangleq}$. Für die multiplikative Inverse müssen wir beachten, dass auch eine Fundamentalfolge, die keine Nullfolge ist, Folgenglieder gleich Null enthalten kann. Allerdings können das nur endlich viele sein; diese kann man (zum Beispiel durch 1) ersetzen, ohne dass sich die Äquivalenzklasse ändert. Für $[(r_n)]_{\triangleq} \neq [(\overline{0})_{n\in\mathbb{N}}]_{\triangleq}$ (und damit $(r_n) \notin \mathcal{N}$) definiert man also

$$([(r_n)]_{\triangleq})^{-1} := [(\tilde{r}_n^{-1})]_{\triangleq} \quad \text{mit} \quad \tilde{r}_n := \begin{cases} r_n & \text{falls } r_n \neq 0 \\ 1 & \text{falls } r_n = 1 \end{cases} \quad \text{für alle } n \in \mathbb{N},$$

wobei man noch beweisen muss, dass $[(\tilde{r}_n^{-1})]$ wieder eine Fundamentalfolge ist.

12.3 VOLLSTÄNDIGKEIT DER REELLEN ZAHLEN

Die so konstruierten reellen Zahlen haben nun keine Lücken mehr: Ein Element, dass durch reelle Zahlen beliebig genau angenähert werden kann, muss selbst eine reelle Zahl sein. Analytisch ausgedrückt: Jede Cauchy-Folge reeller Zahlen konvergiert, und der Grenzwert ist eine reelle Zahl.

Satz 12.3. *Ist (x_n) eine Cauchy-Folge reeller Zahlen, dann existiert ein $x \in \mathbb{R}$, so dass für alle $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $|x_n - x| < \varepsilon$ für alle $n \geqslant n_0$ gilt.*

Beweisskizze: Sei $(x_k)_{k\in\mathbb{N}} = ([(x_{n,k})_{n\in\mathbb{N}})]_{\triangleq})_{k\in\mathbb{N}}$ eine reelle Cauchy-Folge. Aufgrund der Definition von $\mathbb{R}$ gibt es zu jedem Folgenglied x_k ein $r_k \in \mathbb{Q}$ mit $|x_k - r_k| < 1/k$ (jeweils als reelle Zahl aufgefasst). Dann ist die rationale Folge[2] $(r_k)_{k\in\mathbb{N}}$ eine Fundamentalfolge. Diese Folge fassen wir nun mittels Einbettung als reelle Folge $([(r_k)_{n\in\mathbb{N}}]_{\triangleq})_{k\in\mathbb{N}}$ auf, welche dann in $\mathbb{R}$ gegen $x := [(r_n)_{n\in\mathbb{N}}]_{\triangleq}$ konvergiert. Damit konvergiert auch (x_k) gegen x. □

Eine Wiederholung dieser Konstruktion liefert also nichts Neues.

Bei der Arbeit mit reellen Zahlen stellt man sie sich natürlich nicht als Äquivalenzklassen von Folgen vor, sondern verwendet sie anhand der Eigenschaften, die man durch diese Konstruktion gewährleistet hat:

[2] Hier versteckt sich wieder – wie bei Cantor nicht anders zu erwarten – ein Diagonalargument.

(R1) $(\mathbb{R}, +, \cdot)$ ist ein Körper[3] mit dem neutralen Element der Addition 0 und dem der Multiplikation 1.[4]

(R2) Die Relation $\leqslant$ ist eine Totalordnung auf $\mathbb{R}$, bezüglich der die Addition reeller Zahlen sowie die Multiplikation mit nichtnegativen reellen Zahlen monoton ist (siehe (9.2)).

(R3) Die Ordnung ist archimedisch,[5] und jede Cauchy-Folge in $\mathbb{R}$ konvergiert.

Viele Lehrbücher der Analysis führen daher die Menge $\mathbb{R}$ der reellen Zahlen über die Postulate (R1), (R2) und – in einer äquivalenten Form – (R3) axiomatisch ein, und definieren $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ als spezielle Teilmengen. Deren Eigenschaften können dann mit Hilfe von (R1) und (R2) bewiesen werden.

[3] ein kommutativer Ring, für den auch $(\mathbb{R} \setminus \{0\}, \cdot)$ eine kommutative Gruppe ist

[4] Man schreibt natürlich letztlich wieder 0 und 1 für die reellen Zahlen $[(\overline{0})]_{\triangleq}$, $[(\overline{1})]_{\triangleq}$.

[5] Das heisst, für alle $x, y \in \mathbb{R}$ mit $x, y > 0$ existiert ein $n \in \mathbb{N}$, so dass $nx > y$ ist (wieder mit Hilfe der Einbettung in $\mathbb{R}$ aufgefasst). Dies kann man aus der entsprechenden Eigenschaft von $\mathbb{Q}$ folgern, die man mit Induktion beweist.

DIE KOMPLEXEN ZAHLEN

13

Zwar haben die reellen Zahlen keine Lücken mehr, aber trotzdem gibt es immer noch Bereiche, in denen man nicht mit ihnen auskommt. So besitzt die Gleichung $x^2 + 1 = 0$ keine Lösung in $\mathbb{R}$, da $x^2 = x \cdot x \geqslant 0$ für alle $x \in \mathbb{R}$ gilt. (Dies folgt direkt aus der Monotonie der Multiplikation mit nichtnegativen Zahlen und der Fallunterscheidung $x = 0$, $x > 0$ oder $-x > 0$.) Die komplexen Zahlen sind eine Erweiterung des Zahlenbereichs um die Lösung dieser (und ähnlicher) Gleichung(en).

Die Idee hinter der Konstruktion ist, die Lösung von $x^2 + 1 = 0$ als neue, unabhängige „Koordinate“ an die reellen Zahlen anzubringen – ähnlich, wie man aus den reellen Zahlen die euklidische Ebene $\mathbb{R}^2$ gewinnt. Wir betrachten also wieder Paare von (reellen) Zahlen, wobei wir für diese Paare eine Multiplikation definieren, die auch negative Quadrate liefern kann.

Definition 13.1. Die Menge der *komplexen Zahlen* $\mathbb{C}$ ist die Menge aller geordneten Paare reeller Zahlen $\mathbb{R} \times \mathbb{R}$ versehen mit Addition $+$ und Multiplikation $\cdot$, definiert durch

$$\begin{aligned}(a, b) + (c, d) &:= (a + c, b + d),\\ (a, b) \cdot (c, d) &:= (ac - bd, ad + bc).\end{aligned}$$

Die neutralen Elemente der Addition und der Multiplikation sind

$$\tilde{0} := (0, 0) \text{ beziehungsweise } \tilde{1} := (1, 0).$$

Die zu $(a, b) \in \mathbb{C}$ additiv inverse komplexe Zahl ist

$$-(a, b) := (-a, -b)$$

und die zu $(a, b) \in \mathbb{C} \setminus \{(0, 0)\}$ multiplikativ inverse komplexe Zahl ist gegeben durch

$$(a, b)^{-1} := (a \cdot (a^2 + b^2)^{-1}, -b \cdot (a^2 + b^2)^{-1}),$$

Mit den so definierten Verknüpfungen ist $(\mathbb{C}, +, \cdot)$ ein Körper. Die Multiplikation ist jetzt genau so gewählt, dass

$$(0,1)^2 = (0,1) \cdot (0,1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1,0) = -(1,0),$$

also $(0,1)^2 + (1,0) = (0,0)$, gilt.[1]

Die Funktion $j : \mathbb{R} \to \mathbb{C}, x \mapsto (x,0)$, ist injektiv und verknüpfungstreu, und erlaubt es, die Menge der reellen Zahlen mit der Teilmenge $\{(x,0) \in \mathbb{C} : x \in \mathbb{R}\}$ zu identifizieren. Es ist daher üblich, die komplexe Zahl $(a,0)$ einfach wieder mit a zu bezeichnen und für die „imaginäre Einheit“ $(0,1)$ die Abkürzung i zu verwenden, $i := (0,1)$. Da für alle $(a,b) \in \mathbb{C}$ mit $a, b \in \mathbb{R}$ die Gleichung

$$(a,b) = (a,0) + (b,0) \cdot (0,1)$$

gilt, lässt sich jede komplexe Zahl $z = (a,b) \in \mathbb{C}$ in der Form $z = a + b\,i$ schreiben. Damit können wir nun sagen, dass die Gleichung $x^2 + 1 = 0$ eine Lösung in $\mathbb{C}$ hat, nämlich $x = i$. Es gilt sogar noch mehr: In $\mathbb{C}$ hat jede Gleichung der Form $a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$ mit $n \in \mathbb{N}$ und $a_n, \ldots, a_0 \in \mathbb{C}$ eine Lösung; dies ist der *Fundamentalsatz der Algebra*.

Im Gegensatz zu den Konstruktionen der vorigen Kapitel mussten wir hier auf etwas anderes verzichten, um die gewünschte Eigenschaft zu garantieren: Erst der Verzicht auf die Monotonie der Multiplikation erlaubt ja, dass $z^2 < 0$ gelten kann.[2]

Satz 13.2. *Es gibt in $\mathbb{C}$ keine Relation $>$, die folgende zwei Eigenschaften hat:*

1) *Für jedes $z \in \mathbb{C}$ gilt genau eine der Beziehungen $z > \tilde{0}$, $z = \tilde{0}$ oder $-z > \tilde{0}$.*
2) *Aus $w > \tilde{0}$ und $z > \tilde{0}$ folgt $w + z > \tilde{0}$ und $wz > \tilde{0}$.*

Beweis. Angenommen man hätte eine Relation $>$ auf $\mathbb{C}$ mit den geforderten Eigenschaften. Dann müsste (wie im Reellen) für jedes $z \neq \tilde{0}$ gelten, dass $z^2 > \tilde{0}$ ist. Insbesondere wäre dann $\tilde{1}^2 > \tilde{0}$ und $i^2 > \tilde{0}$, folglich auch $\tilde{0} = i^2 + 1 > \tilde{0}$, und dies ist ein Widerspruch. □

Beachten Sie, dass man auf $\mathbb{C}$ durchaus eine Ordnung definieren kann, wenn wir eine der beiden Eigenschaften aufgeben. Verzichten wir auf die Totalordnungseigenschaft (1), so ist zum Beispiel $(a,0) < (b,0) :\Leftrightarrow a < b$ (wobei Letzteres die übliche Ordnung auf $\mathbb{R}$ bezeichnet) eine Ordnung, für die Eigenschaft (2) gilt – allerdings sind nur komplexe Zahlen mit $\operatorname{Im} z = 0$ vergleichbar. Umgekehrt ist die lexikographische Ordnung auf $\mathbb{R}^2$ eine Totalordnung, die aber nicht mit der Multiplikation verträglich ist ($(0,-1) < (0,1)$, aber $(0,-1)^2 = (1,0) > (-1,0) = (0,1)^2$).

[1] Tatsächlich wird durch diese Forderung und die Definition von $(1,0)$ als neutralem Element die Multiplikation in $\mathbb{C}$ eindeutig festgelegt. Sie hat auch eine geometrische Interpretation als Rotation mit Skalierung von Vektoren in der Ebene.

[2] Ein Beispiel sind die *Quaternionen*, die eingeführt wurden, um die geometrische Interpretationen der komplexen Zahlen als Rotation in der Ebene in den dreidimensionalen Raum zu übertragen. Allerdings muss man auch bei den folgenden Erweiterungen notwendigerweise auf weitere Eigenschaften verzichten; so sind die Quaternionen etwa nicht mehr kommutativ.

ANHANG

PEANO-AXIOME UND DIE KONSTRUKTION DER NATÜRLICHEN ZAHLEN

A

In diesem Kapitel möchten wir zeigen, wie die natürlichen Zahlen, die in Kapitel 9 anhand ihrer geforderten Eigenschaften eingeführt wurden, mathematisch fundieren werden können. Dabei gehen wir in zwei Schritten vor: Zuerst führen wir ein minimales Axiomensystem ein, dass die gewünschten Eigenschaften der natürlichen Zahlen gewährleistet. Dann skizzieren wir, wie aus der axiomatischen Mengenlehre eine Menge konstruiert werden kann, die die vorher axiomatisch geforderten Eigenschaften (nun beweisbar) erfüllt. Üblicherweise wird dabei die Menge der natürlichen Zahlen $\mathbb{N}_0$ als *inklusive* der Null konstruiert. Die Menge $\mathbb{N}$ erhält man dann zum Beispiel als Bild von $\mathbb{N}_0$ unter der Funktion $n \mapsto n + 1$.

A.1 PEANO-ARITHMETIK

Wir verlangen, dass die natürlichen Zahlen $\mathbb{N}_0$ zusammen mit einem Element 0 und einer *Nachfolgervorschrift* S die *Peano-Axiome*[1] erfüllen:

(P1) $0 \in \mathbb{N}_0$.

(P2) $S : \mathbb{N}_0 \to \mathbb{N}_0$ ist eine Funktion, das heisst für alle $n \in \mathbb{N}_0$ gilt $S(n) \in \mathbb{N}_0$.

(P3) Für alle $n \in \mathbb{N}_0$ gilt $S(n) \neq 0$.

(P4) Für alle $n, m \in \mathbb{N}_0$ gilt: Aus $S(n) = S(m)$ folgt $n = m$.

(P5) Für alle $M \subseteq \mathbb{N}_0$ gilt: Wenn $0 \in M$ gilt und für alle $n \in \mathbb{N}_0$ aus $n \in M$ auch $S(n) \in M$ folgt, so ist $M = \mathbb{N}_0$.

Die Axiome (P1) und (P3) gewährleisten also, dass die natürlichen Zahlen einen festgelegten Anfang (hier: 0) haben. Mit (P2) bekommen wir für jede natürliche Zahl genau einen Nachfolger. Wegen (P4) ist die Nachfolgervorschrift injektiv – dies garantiert zusammen mit (P1), dass $\mathbb{N}_0$ unendlich ist. Damit ist im wesentlichen die naive Vorstellung des Zählens

[1] Diese Axiome wurden 1889 von Guiseppe Peano aufgestellt.

axiomatisch festgehalten. Axiom (P5) ist nichts anderes als das Prinzip der vollständigen Induktion (in mengentheoretischer Schreibweise). Wir können also die einzelnen natürlichen Zahlen definieren als

$$0, \quad 1 := S(0), \quad 2 := S(1), \quad 3 := S(2), \quad \dots.$$

Addition und Multiplikation können nun rekursiv über die Nachfolgervorschrift S definiert werden. Die Addition $+ : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ ist die (wegen des Rekursionssatzes 7.6) eindeutige Abbildung, die für alle $n \in \mathbb{N}_0$

$$\begin{aligned} n + 0 &= n, \\ n + S(m) &= S(n + m) \end{aligned}$$

erfüllt. Die Multiplikation $\cdot : \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$ erhalten wir analog für alle $n \in \mathbb{N}_0$ durch

$$\begin{aligned} n \cdot 0 &= 0, \\ n \cdot S(m) &= (n \cdot m) + n. \end{aligned}$$

Die Ordnung wird wie in Kapitel 9 über die Addition definiert. Mit diesen Definitionen und großzügiger Anwendung der vollständigen Induktion kann man nun die kompletten arithmetischen und Ordnungseigenschaften von $\mathbb{N}_0$ beweisen, die in Kapitel 9 aufgezählt sind.

Was noch offen bleibt, ist die Frage, ob überhaupt solch eine Menge (im mathematischen Sinne) existiert, oder ob diese Forderungen gar unvereinbar mit anderen, genauso fundamentalen Objekten der Mathematik sind. Außerdem können wir nicht sicher sein, dass wir überhaupt von *den* natürlichen Zahlen sprechen können, oder ob nicht noch weitere Mengen existieren, die die Peano-Axiome erfüllen (und womöglich zusätzliche, unerwünschte Eigenschaften haben). Um diese Fragen zu beantworten, muss man die natürlichen Zahlen auf Basis der Mengenlehre konstruieren.

A.2 MENGENTHEORETISCHE KONSTRUKTION VON $\mathbb{N}_0$

Die grundlegende Idee ist, jede natürliche Zahl n mit der Menge ihrer Vorgänger $\{0, \dots, n-1\}$ zu identifizieren. Wir erhalten damit

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

und allgemein

$$n + 1 := n \cup \{n\}.$$

Diese Konstruktion, die auf John von Neumann zurückgeht, wollen wir nun im Rahmen des Zermelo-Fraenkel-Axiomensystems (siehe Anhang B) präzisieren. Wir nennen eine Menge M *induktiv*, wenn gilt: $\emptyset \in M$, und für alle $m \in M$ ist auch $m \cup \{m\} \in M$. Das Unendlichkeitsaxiom (*Unend*) garantiert direkt die Existenz einer solchen Menge. Wir definieren nun die natürlichen Zahlen $\mathbb{N}_0$ als Durchschnitt aller induktiven Mengen:

$$n \in \mathbb{N}_0 :\Leftrightarrow \forall M : (M \text{ induktiv } \Rightarrow n \in M).$$

Dass dadurch die Menge $\mathbb{N}_0$ eindeutig charakterisiert ist, folgt aus dem Aussonderungsschema (Aus_P), angewandt auf $P(x) :\Leftrightarrow \forall M : (M \text{ induktiv } \Rightarrow x \in M)$, und dem Extensionalitätsaxiom (*Ext*).

Definiert man nun die Nachfolgervorschrift $S(n) := n \cup \{n\}$, kann man leicht nachweisen, dass $(\mathbb{N}_0, \emptyset, S)$ die Peano-Axiome erfüllt. Insbesondere gilt (P5): wenn M eine induktive Teilmenge von $\mathbb{N}_0$ ist, muss $M = \mathbb{N}_0$ sein, da $\mathbb{N}_0$ nach Konstruktion die kleinste induktive Menge ist.

Ein schöner Nebeneffekt dieser Konstruktion ist, dass wir die Ordnung auf $\mathbb{N}_0$ besonders leicht definieren können: Es gelte $m < n$ genau dann, wenn $m \in n$ ist. (Addition und Multiplikation definiert man weiterhin über die Nachfolgervorschrift wie in Abschnitt A.1 beschrieben.)

DAS ZFC-AXIOMENSYSTEM

B

Der Vollständigkeit halber geben wir hier das Axiomensystem für die heute akzeptierte axiomatische Mengenlehre nach Zermelo und Fraenkel an. Abgesehen von der Prädikatenlogik erster Stufe und der Identität, wie sie in Kapitel 1 beschrieben sind, wird als elementarer Begriff lediglich die Element-von-Relation $\in$ eingeführt. Der Übersichtlichkeit halber schreiben wir wieder $x \notin y$ für $\neg(x \in y)$ und $x \neq y$ für $\neg(x = y)$. Alle Quantoren laufen über das sogenannte *Universum* aller Mengen.[1]

Die einzelnen *ZFC-Axiome* sind:

(*Ext*) $$\forall x : \forall y : (\forall z : (z \in x \Leftrightarrow z \in y)) \Rightarrow x = y$$

Das *Extensionalitätsaxiom* besagt, dass zwei Mengen gleich sind, wenn sie die selben Elemente haben. Die andere Implikationsrichtung folgt direkt aus der Definition der Identität. (Umgekehrt kann dies auch als Definition der Gleichheit verwendet werden, wenn die Implikation durch die Äquivalenz ersetzt wird.)

(*Leer*) $$\exists x : \forall y : y \notin x$$

Das *Leermengenaxiom* fordert die Existenz der leeren Menge, die mit $\emptyset$ bezeichnet wird. Dies garantiert, dass überhaupt Mengen existieren, über die wir etwas aussagen können.[2] Die Eindeutigkeit der leeren Menge folgt sofort aus dem Extensionalitätsaxiom.

Die nächsten vier Axiome erlauben die Bildung neuer Mengen.

(Aus_P) $$\forall z : \exists x : \forall y : (y \in x \Leftrightarrow (y \in z \wedge P(y)))$$

Das *Aussonderungsschema* erzeugt für jede Aussageform P ein Axiom, dass die Existenz einer Teilmenge einer gegebenen Menge z liefert, welche genau die Elemente y enthält, für die

[1] Diese Sammlung aller Mengen kann natürlich selbst keine Menge sein, wie die Russellsche Antinomie zeigt. Da die ZFC-Axiome die Mengenlehre „von außen“ beschreiben, ist dies zulässig; umgekehrt können aus den ZFC-Axiomen keine Aussagen über das Universum als Ganzes abgeleitet werden.

[2] Fordern wir umgekehrt die Existenz einer Menge per Axiom (etwa $\exists x : x = x$), können wir die Existenz der leeren Menge daraus und aus dem Separationsaxiom beweisen.

$P(y)$ wahr ist. Das Extensionalitätsaxiom garantiert wiederum die Eindeutigkeit von x. Dies ist die Grundlage der prädikativen Definition von Mengen, wie wir sie in Kapitel 2 kennen gelernt haben. (Die Bedingung $y \in z$ verhindert dabei die Konstruktion der Russellschen Antinomie.)

(*Paar*) $$\forall x, y : \exists z : (x \in z \wedge y \in z)$$

Das *Paarmengenaxiom* erlaubt uns für gegebenes x und y die Konstruktion der Menge $\{x, y\} := \{w \in z : w = x \vee w = y\}$ (die wegen (*Ext*) wieder eindeutig bestimmt ist). Damit können wir nun für gegebenes x die Menge $\{x\} := \{x, x\}$ bilden, die wir insbesondere für die Konstruktion der natürlichen Zahlen benötigen.

(*Ver*) $$\forall z : \exists x : \forall y : (y \in x \Leftrightarrow \exists w : (w \in z \wedge y \in w))$$

Das *Vereinigungsaxiom* liefert die Bildung von Vereinigungsmengen: Für jede Menge z existiert eine (wieder: eindeutige) Menge x, die genau die Elemente enthält, die in einem Element w von z enthalten sind. Für a, b können wir deshalb $a \cup b := \{w \in x : w \in a \vee w \in b\}$ definieren, indem wir (*Ver*) auf die Menge $z = \{a, b\}$ anwenden.

(*Pot*) $$\forall z : \exists x : \forall y : (y \in x \Leftrightarrow \forall w : (w \in y \Rightarrow w \in z))$$

Das *Potenzmengenaxiom* garantiert für jedes z die (eindeutige) Existenz der Potenzmenge $\mathcal{P}(z) := x$: ein Element y liegt in x dann und nur dann, wenn y Teilmenge von z ist.

(Ers_P) $$\forall z : (\forall w \in z : \exists! y : P(w, y)) \Rightarrow (\exists x : \forall y : (y \in x \Leftrightarrow \exists w : (w \in z \wedge P(w, y))))$$

Das *Ersetzungsschema* erzeugt für jede gegebene Funktion f ein Axiom, das für gegebene Menge z die Bildmenge $f(z) := x$ liefert. Dabei ist f charakterisiert durch den Definitionsbereich z und Zuordnungsvorschrift $w \mapsto y :\Leftrightarrow P(w, y)$ (die Bedingung $\forall w \exists! y P(w, y)$ gewährleistet dabei, dass es sich wirklich um eine Funktion handelt). Dies rechtfertigt auch die Schreibweise $\{f(x) : x \in X\}$ für entsprechende Mengen.

Mit den bisher eingeführten Axiomen kann man im Grunde schon die übliche Mengenlehre betreiben – wenn man sich auf endliche Mengen beschränkt. Die Existenz von unendlichen Mengen muss explizit gefordert werden; sie können nicht alleine aus endlichen Mengen konstruiert werden:

(*Unend*) $$\exists x : (\emptyset \in x \wedge \forall z : (z \in x \Rightarrow (z \cup \{z\} \in x)))$$

Das *Unendlichkeitsaxiom* garantiert also die Existenz einer speziellen unendlichen Menge mit den Elementen $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots$, und ist die Grundlage für die von Neumann-Konstruktion der natürlichen Zahlen.

(*Fund*) $$\forall z : (\exists y : y \in z \Rightarrow \exists y (y \in z \wedge \forall x : (x \notin z) \vee x \notin y)))$$

Das *Fundierungsaxiom* besagt, dass jede nichtleere Menge z ein Element y enthält, so dass z und y disjunkt ist. Damit werden bestimmte bizarre Mengen ausgeschlossen, wie etwa solche

Mengen, die sich selbst enthalten.[3] Dieses Axiom wird im wesentlichen nur für bestimmte Wohlordnungsaussagen benötigt.

Die obigen Axiome ergeben zusammen das *ZF-System* (für Zermelo-Fraenkel). Durch Hinzunehmen des *Auswahlaxioms* entsteht daraus das *ZFC-System* (C für „choice“).

$$(C)\quad \forall x : [\forall y : (y \in x \Rightarrow \exists z : z \in y) \wedge \forall y, z : y \in x \wedge z \in x \wedge y \neq z \Rightarrow \forall u : u \in y \Rightarrow u \notin z] \Rightarrow \exists y : \forall z : z \in x \Rightarrow \exists! u : u \in y \wedge u \in z$$

Dieses Axiom besagt, dass für jede nichtleere Menge x, die die nichtleeren und disjunkten Mengen y enthält, eine Menge u existiert, die genau ein Element aus jeder dieser Mengen y enthält. Dieses Axiom hat immer wieder zu Kontroversen geführt, da es die Definition sehr abstrakter (und unerwünschter) Mengen gestattet, ohne eine Konstruktionsmöglichkeit kennen zu müssen.[4] So existieren in ZFC Teilmengen der reellen Zahlen, die nicht messbar sind (d.h., die in einem gewissen Sinn keine sinnvolle „Länge“ haben). Umgekehrt kann man ohne Auswahlaxiom nicht beweisen, dass jeder (unendlich-dimensionale) Vektorraum eine Basis hat. Da dieses und ähnliche (nur mit (*C*) beweisbare) Resultate in der Praxis sehr fruchtbar sind, akzeptieren die meisten Mathematiker dieses Axiom. Trotzdem suchen einige Mathematiker weiterhin nach neuen Beweisen, die das Auswahlaxiom nicht verwenden, oder nach einem Ersatz für (*C*), das die gewünschten Resultate, aber nicht die Konstruktion von „Monstern“ erlaubt.

Tatsächlich kann diese Liste noch weiter reduziert werden, da die Gültigkeit einiger der obigen Forderungen bereits aus den übrigen bewiesen werden kann:

- (*Leer*) folgt aus (*Unend*) per Aussonderung,
- (*Paar*) folgt aus (Ers_P), (*Leer*) und (*Pot*),
- (Aus_P) ist ein Spezialfall von (Ers_P).

Die erste Folgerung geht dabei auf Fraenkel, die letzten beiden auf Zermelo zurück.

[3] So verletzt zum Beispiel die Menge $\{\{\{\dots\}\}\}$, die nur ein Element enthält – und zwar sich selbst! – das Fundierungsaxiom.

[4] Solche Definitionen werden in der intuitionistischen Mathematik kategorisch abgelehnt.