Vorlesungsnotizen Mathematik 1 für den Lehramtsstudiengang Sekundarstufe, Wintersemester 2021/22

Birgit Richter, Version vom 26. Januar 2022

Fachbereich Mathematik der Universität Hamburg, Bundesstrasse 55, 20146 Hamburg

Inhaltsverzeichnis

KAPITEL I

Grundlagen aus Logik und Mengenlehre

I.1. Vorwort

Dieses Skript baut auf dem Skript von Prof. Dr. Thomas Schmidt aus dem Wintersemester 2020/21 auf, ist aber an einigen Stellen abgeändert.

Der Vorlesungszyklus Mathematik 1 bis 4 für das Lehramt der Sekundarstufe deckt vor allem Themen aus den Bereichen Analysis und lineare Algebra ab. Diese Bereiche sind grundlegend für die Mathematik und werden auch im Studium des Kernfachs Mathematik als Erstes unterrichtet. Die lineare Algebra ist dabei eine allgemeine Lehre linearer Strukturen, zu der unter anderem der Umgang mit linearen Gleichungssystemen, linearen Abbildungen, Vektoren und Matrizen zählt und und zu der die sogenannte analytische Geometrie mit Punkten, Geraden und Ebenen gehört. Die Analysis ist die Lehre vom mathematischen Umgang mit dem Unendlichen, der letztlich immer mittels Grenzübergängen und Grenzwerten erfolgt. Zentrale Themen der Analysis sind reelle und komplexe Zahlen, Grenzwerte, Ableitungen und Integrale. In der Vorlesung Mathematik 1 liegt der Schwerpunkt zunächst auf allgemeinen Grundlagen der Mathematik, die nicht unbedingt einem der genannten Bereiche zuzuordnen sind, und dann auf ersten Themen der Analysis.

Das Ziel der Vorlesungen Mathematik 1 bis 4 ist, eine fachwissenschaftliche Einführung in das mathematische Arbeiten mit präzise definierten Begriffen, Lehrsätzen und vollständigen Beweisen zu geben. Im Schulunterricht der Sekundarstufe relevante Themen werden dabei in größerer Ausführlichkeit behandelt. Daneben sollen ein deutlich über den Schulstoff hinausgehendes mathematisches Hintergrundwissen, die Fähigkeit zur eigenständigen Einarbeitung in mathematische Themen und Konzepte, ein gewisser Überblick über mathematische Gebiete sowie ein Gefühl für die Mathematik als lebendige Wissenschaft vermittelt werden.

In diesem initialen Kapitel geht es um Aspekte der mathematischen Logik, der mathematischen Arbeitsweise und der Mengenlehre, die als Grundlage der gesamten modernen Mathematik betrachtet werden und den streng mathematische Aufbau aller weiteren Theorie erst ermöglichen. Die Beschäftigung mit den Grundlagen ist dabei keineswegs einfach, da man auch sehr einleuchtende Sachverhalte weiter hinterfragen und begründen muss.

Hier werden wir sowohl die Logik als auch die Mengenlehre hauptsächlich von einem naiven Standpunkt betrachten, der für ein Mathematik-Studium und das wissenschaftliche Arbeiten in den allermeisten mathematischen Gebieten ausreicht. Es sind fundiertere Zugänge zu Logik und Mengenlehre bekannt. Solche Aspekte können wir hier aber bestenfalls andeuten. Bevor man sie im Detail verstehen kann, braucht man erst einmal mehr Erfahrung mit dem mathematischen Denken und Arbeiten.

I.2. Aussagenlogik und Wahrheitstafeln

Grundlegend für alle Gebiete der modernen Mathematik ist der Umgang mit Aussagen, die entweder als wahr **(w)** oder als falsch **(f)** zu betrachten sind und denen einer dieser beiden sogenannten Wahrheitswerte zugeordnet wird. Insbesondere bewegen sich mathematische Aussagen nicht in Grauzonen. Ein Satz wie „Die Zahl 10 ist groß.“ erfüllt also nicht den Anspruch einer mathematischen Aussage, denn ab wann man von „groß“ spricht wird nicht ersichtlich. Wenn man „groß“ natürlich im Vorfeld definiert hat, dann ist es etwas Anderes. Dagegen sind

„Die Zahl 10 ist größer als die Zahl 15.“

und

„Die Zahl 10 ist mindestens so groß wie die Zahl 10.“

Beispiele für sinnvolle mathematische Aussagen, von denen die erste natürlich falsch, die zweite wahr ist. Dass eine Aussage gilt, ist eine alternative mathematische Sprechweise dafür, dass die Aussage wahr ist.

Aussagen können im Rahmen der Aussagenlogik durch logische Grundoperationen verneint und auf verschiedene Weisen zusammengesetzt werden. Formal notiert man verneinte und zusammengesetzte Aussagen mit Hilfe von Junktoren. Dies sind logische Symbolen, die vor oder zwischen die Aussage(n) geschrieben werden. Die fünf wichtigsten Grundoperationen und die Anwendung der zugehörigen Junktoren mit Platzhaltern A und B für Aussagen sind in folgender Tabelle zusammengefasst, wobei die Operationen, Schreib- und Sprechweisen innerhalb eines Feldes Alternativen gleicher Bedeutung sind:

Operation	**Junktoren**	**Bedeutung, Sprechweisen**
Verneinung Negation	$\neg A$	nicht A
logisches Und Konjunktion	$A \wedge B$	A und B
logisches Oder Disjunktion	$A \vee B$	A oder B
Folgerung Implikation	$A \Longrightarrow B$ $B \Longleftarrow A$	Aus A folgt B B folgt aus A A impliziert B
Äquivalenz	$A \Longleftrightarrow B$	A ist äquivalent zu B A gleichbedeutend mit B

Die Wahrheitswerte der verneinten und zusammengesetzten Aussagen ergeben sich dabei einzig und allein aus den Wahrheitswerten von A und B und werden durch folgende *Wahrheitstafeln* festgelegt:

A	$\neg A$
w	f
f	w

A	B	$A \wedge B$	$A \vee B$	$A \Longrightarrow B$	$A \Longleftrightarrow B$
w	w	w	w	w	w
w	f	f	w	f	f
f	w	f	w	w	f
f	f	f	f	w	w

Hervorzuheben ist hierbei insbesondere, dass das logische Oder nicht exklusiv ist, das heißt, $A \vee B$ ist *auch* dann wahr, wenn A, B beide wahr sind, zusätzlich zu den Fällen, in denen eine der Aussagen A, B wahr, eine falsch ist.

Daneben mag zunächst überraschen, dass die Implikation $A \Longrightarrow B$ bei falscher Prämisse A (dritte/vierte Zeile Tabellenkörper) als wahr festgelegt wird, also „aus Falschem alles folgt“.

Beispiele dazu sind

$$\begin{aligned} 3 \text{ ist ungerade} &\Longrightarrow 2 \cdot 4 = 8 \ \text{(wahr)} \\ 3 \text{ ist ungerade} &\Longrightarrow 2 \cdot 4 = 7 \ \text{(falsch)} \\ 3 \text{ ist gerade} &\Longrightarrow 2 \cdot 4 = 8 \ \text{(wahr)} \\ 3 \text{ ist gerade} &\Longrightarrow 2 \cdot 4 = 7 \ \text{(wahr)} \end{aligned}$$

Um dies zu verstehen, kann man auch an das Sprichwort „Wer A sagt, der muss auch B sagen.“ denken, das als Implikation aus den Teilaussagen „A wird gesagt.“ und „B wird gesagt.“ zusammengesetzt ist. Wenn man A und B sagt (erste Zeile), ist dies im Sinn des Sprichworts „richtig“. Wenn man A nicht sagt, ergibt sich keine Verpflichtung. Man kann dann B sagen (dritte Zeile) oder auch nicht (vierte Zeile); beides wäre „richtig“. Die einzige Möglichkeit, gemäß Sprichwort „falsch“ zu handeln, ist in der Tat die, dass zwar A, aber nicht B gesagt wird (zweite Zeile).

Natürlich kann man mehrere Junktoren kombinieren, wobei die Reihenfolge der Auswertung genau wie bei den Grundrechenarten im Allgemeinen durch Klammern anzuzeigen ist, und kann für jede logische Formel aus endlich vielen Aussagen und endlich vielen Junktoren Wahrheitstafeln ableiten.

Es kommt vor, dass verschiedene logische Formeln in allen Fällen mit dem gleichen Wahrheitswert belegt sind. In diesem Fall sind sie logisch äquivalent und wir können sie mit „ $\Longleftrightarrow$ “markieren.

Beispiele für logisch äquivalente Formeln, gebildet aus Aussagen A, B, C, sind die Kommutativ- und Assoziativgesetzte:

$$A \wedge B \iff B \wedge A, \qquad A \vee B \iff B \vee A,$$
$$(A \wedge B) \wedge C \iff A \wedge (B \wedge C), \qquad (A \vee B) \vee C \iff A \vee (B \vee C).$$

Dementsprechend kann man bei Formeln dieser Gestalt die Reihenfolge der Aussagen vertauschen und auf Klammerung verzichten. Weitere Beispiele für äquivalente Formeln sind:

$$(A \vee B) \wedge C \iff (A \wedge C) \vee (B \wedge C), \qquad (A \wedge B) \vee C \iff (A \vee C) \wedge (B \vee C),$$
$$A \Longrightarrow B \iff (\neg A) \vee B, \qquad A \Longleftrightarrow B \iff (A \wedge B) \vee (\neg(A \vee B)).$$

Die beiden Regeln der oberen Zeile kann man sich als Distributivgesetze für $\wedge$ und $\vee$ merken. Mit der unteren Zeile können die Pfeil-Symbole $\Longrightarrow$ und $\Longleftrightarrow$ durch $\neg$, $\wedge$ und $\vee$ ausgedrückt werden und erweisen sich damit im Prinzip als redundant. In der Tat ist die Pfeil-Notation aber sehr intuitiv und wird in der Praxis daher oft benutzt.

Es kann sich beim Ausfüllen einer Wahrheitstafel auch herausstellen, dass eine logische Formel für alle möglichen Kombinationen von Wahrheitswerten der Einzelaussagen stets wahr ist. Solche Formeln heißen *Tautologien*. Die Tautologien der Form $(\ldots) \Longleftrightarrow (\ldots)$ bedeuten dabei gemäß der Wahrheitstafel für die Äquivalenz, dass die linke und die rechte Teilformel stets denselben Wahrheitswert haben, somit logisch äquivalent sind und beliebig durch einander ausgetauscht werden können.

Bekannte *Beispiele für Tautologien*, gebildet aus Aussagen A, B, C, sind:

$(\neg(\neg A)) \Longleftrightarrow A$	(Gesetz der doppelten Negation),
$A \vee (\neg A)$	(Satz vom ausgeschlossenen Dritten),
$(A \wedge B) \Longrightarrow (A \vee C)$,	
$(\neg(A \wedge B)) \Longleftrightarrow ((\neg A) \vee (\neg B))$ $(\neg(A \vee B)) \Longleftrightarrow ((\neg A) \wedge (\neg B))$	(*De Morgansche Gesetze* zur Negation der Konjunktion und Disjunktion),
$(\neg(A \Longrightarrow B)) \Longleftrightarrow (A \wedge (\neg B))$	(Negation der Implikation).

Weitere Beispiele für Tautologien, an die sich später diskutierte Schluß- und Beweistechniken anlehnen, sind:

$(A \wedge (A \Longrightarrow B)) \Longrightarrow B$	(Modus ponens),
$((A \Longrightarrow B) \wedge (B \Longrightarrow C)) \Longrightarrow (A \Longrightarrow C)$	(Transitivität der Implikation),
$(A \Longrightarrow B) \Longleftrightarrow ((\neg B) \Longrightarrow (\neg A))$	(Kontrapositions-Prinzip),
$(A \Longleftrightarrow B) \Longleftrightarrow ((A \Longrightarrow B) \wedge (B \Longrightarrow A))$	(Charakterisierung der Äquivalenz durch gegenseitige Implikation).

I.3. Prädikatenlogik und Quantoren

In der mathematischen Praxis kommt man mit der Aussagenlogik allein nicht weit, sondern benötigt schnell die allgemeinere Prädikatenlogik. Prädikate sind dabei Aussagen mit freien Variablen. Ein Beispiel ist „x ist größer als die Zahl 15.“ mit einer freien Variable x. Freie Variablen sind Platzhalter für noch unbestimmte Objekte, und solang diese nicht bestimmt sind, man etwa im Beispiel den Wert von x nicht kennt, solange kann man über wahr oder falsch nicht sinnvoll entscheiden. Somit kann und soll Prädikaten erst einmal kein Wahrheitswert zugeordnet werden. Erst wenn man für die freien Variablen sinnvoll konkrete Objekte oder Zahlen einsetzt, zum Beispiel die Zahl 10 für x, erst dann ergeben sich wieder individuelle Aussagen, die als wahr oder falsch zu betrachten sind.

Man kann mit einem Prädikat auch auf andere Art Aussagen ohne freie Variablen bilden. Beispielsweise lässt sich mit dem Prädikat „x ist größer als die Zahl 15.“ einerseits die Aussage

„Alle natürlichen Zahlen sind größer als die Zahl 15.“ (falsch)

bilden und andererseits die Aussage

„Es gibt eine natürliche Zahl, die größer als die Zahl 15 ist.“ (wahr).

Dies sind tatsächlich Beispiele für das Hinzufügen von *Quantoren*, den entscheidenden logischen Operatoren der Prädikatenlogik, die bei einem Prädikat $P(x)$ mit einer freien Variable x die Variable binden und zu einer Aussage ohne freie Variable führen.

Die Kombination aus Prädikat und Quantor liefert eine Aussage mit Wahrheitswert. In der Praxis gibt man in Kombination mit dem Quantor auch eine Grundmenge[1] M von Elementen an, die man für die Variable x einzusetzen erlaubt. Im eben betrachteten Beispiel war diese Grundmenge M die Menge der natürlichen Zahlen.

Zu Schreib- und Sprechweisen für die beiden maßgeblichen und im Beispiel schon betrachteten Quantoren halten wir fest:

Quantor	**Schreibweisen**	**Bedeutung, Sprechweisen**
All-Quantor	$\forall x{\in}M\colon P(x)$	Für alle x aus M gilt $P(x)$. Sei x aus M beliebig. Dann gilt $P(x)$.
Existenz-Quantor	$\exists x{\in}M\colon P(x)$	Es gibt ein x aus M mit $P(x)$. Für ein x aus M gilt $P(x)$.

Hierbei wird die Aussage $\forall x{\in}M\colon P(x)$ als wahr betrachtet, wenn $P(x)$ für jedes Objekt aus der Grundmenge M, das an Stelle von x eingesetzt wird, wahr ist. In allen anderen Fällen gilt $\forall x{\in}M\colon P(x)$ als falsch. Analog betrachtet man $\exists x{\in}M\colon P(x)$ als wahr, wenn es ein Objekt in der Grundmenge M gibt, dessen Einsetzen zu einer wahren Aussage $P(x)$ führt. In allen anderen Fällen gilt $\exists x{\in}M\colon P(x)$ als falsch. Betont sei dabei, dass mit einem x immer *mindestens ein* x gemeint ist; diese Interpretation ist in der Mathematik allgemein üblich.

Die Symbole $\forall$ als umgedrehtes A und $\exists$ als gespiegeltes E erinnern dabei an die Namen der Quantoren. Gelegentlich wird auch die Notation $\exists!x{\in}M\colon P(x)$ mit einem Existenz- und Eindeutigkeitsquantor $\exists!$ verwendet, um auszudrücken, dass $P(x)$ für genau ein Objekt – also für ein einziges, aber kein weiteres – aus M wahr ist.

Seien Sie bitte vorsichtig, wenn Sie Aussagen verneinen, die Quantoren enthalten. Die Verneinung der Aussage „Alle Primzahlen sind ungerade“ ist *nicht* „Alle Primzahlen sind nicht ungerade“ (also „Alle Primzahlen sind gerade“) sondern „Es gibt eine Primzahl, die nicht ungerade ist“. Diese Aussage ist wahr: Die Zahl 2 ist gerade aber trotzdem eine Primzahl.

Für die Negation von Quantoren gelten die folgenden Gesetzmäßigkeiten. Dies sind Versionen der de Morganschen Gesetze aus Abschnitt I.2:

$$(\neg(\forall x{\in}M\colon P(x))) \iff (\exists x{\in}M\colon (\neg P(x))),$$
$$(\neg(\exists x{\in}M\colon P(x))) \iff (\forall x{\in}M\colon (\neg P(x))).$$

Die erste Regel besagt , dass „$P(x)$ gilt nicht für alle x aus M.“ gleichbedeutend ist mit „Für ein x aus M gilt $P(x)$ nicht.“ Die zweite Regel drückt aus, dass „Für kein x aus M gilt $P(x)$.“ gleichbedeutend mit „Für alle x aus M gilt $P(x)$ nicht.“ ist.

Natürlich treten in der Mathematik auch Prädikate mit mehreren freien Variablen auf, wobei die Variablen dann durch mehrere Quantoren gebunden werden können. Beispielsweise kann man aus einem Prädikat

[1] Tatsächlich werden hier Quantoren unter Verwendung von Mengen und im nächsten Abschnitt Mengen unter Verwendung von Quantoren erklärt. Eine konsistente Einführung der Begriffe erfordert ein vorsichtigeres Vorgehen: Man würde dazu an dieser Stelle nur die „Mengen-freien“ Aussagen $\forall x\colon P(x)$ und $\exists x\colon P(x)$ einführen, wobei im Hintergrund ein sogenanntes Diskursuniversum von zulässigen Objekten steht, die für die Variable x eingesetzt werden können. Im nächsten Schritt könnte man dann axiomatisch Mengen und die Elementbeziehung $\in$ einführen — unter Verwendung solcher Quantoren über das Diskursuniversum der Mengen. Schließlich würde man $\forall x \in M\colon P(x)$ und $\exists x \in M\colon P(x)$ als abkürzende Schreibweisen für $\forall x\colon (x \in M \implies P(x))$ und $\exists x\colon (x \in M \implies P(x))$ erklären. Sind diese Grundlagen einmal geklärt, so laufen alle fortan relevanten Quantoren aber tatsächlich über Grundmengen M.

$P(x,y)$ mit zwei freien Variablen x und y und aus Quantoren über Grundmengen M und N unter anderem die Aussagen $\forall x \in M : \forall y \in N : P(x,y)$ und $\exists y \in N : \forall x \in M : P(x,y)$ bilden.

Es ist dabei wichtig zu wissen, dass zwei Quantoren gleichen Typs (also $\forall$ mit $\forall$ und $\exists$ mit $\exists$) in der Reihenfolge vertauscht werden dürfen, ohne dass sich die Bedeutung der Aussage ändert. Beispielsweise ist $\forall x{\in}M : \forall y{\in}N : P(x,y)$ äquivalent zu $\forall y{\in}N\colon \forall x{\in}M\colon P(x,y)$.

Die Vertauschung eines All-Quantors mit einem Existenz-Quantor ist dagegen nicht ohne Unterschied in der Bedeutung möglich: Tatsächlich bedeutet $\forall x \in M : \exists y \in N : P(x,y)$, dass zu jedem $x{\in}M$ ein von x abhängiges $y{\in}N$ existiert, so dass $P(x,y)$ gilt. Dagegen symbolisiert $\exists y{\in}N : \forall x{\in}M : P(x,y)$ die Existenz eines $y{\in}N$ (jetzt wirklich nur ein y, das nicht von einem x abhängen darf), so dass für alle $x{\in}M$ Gültigkeit von $P(x,y)$ besteht. Man kann sich den Unterschied zwischen diesen beiden Aussagen klarmachen, indem man für $P(x,y)$ das einfache Prädikat „y ist größer als x“ einsetzt und als Mengen M und N die Menge der natürlichen Zahlen: Die Aussage $\forall x{\in}M : \exists y{\in}N : P(x,y)$ bedeutet dann

„Für alle natürlichen Zahlen x gibt es eine natürliche Zahl y, die größer ist als die zuerst gegebene Zahl x.“

und ist richtig. Die Aussage $\exists y{\in}N : \forall x{\in}M : P(x,y)$ dagegen bedeutet

„Es gibt eine natürliche Zahl y, die größer ist als alle natürlichen Zahlen x.“

und ist falsch.

Wie sich die Negationsregeln auf Formeln mit mehreren Quantoren auswirken, kann man schließlich anhand des Beispiels

$$(\neg(\forall x{\in}M : \exists y{\in}N : P(x,y))) \iff (\exists x{\in}M : \forall y{\in}N : (\neg P(x,y)))$$

verstehen: Beim Hereinziehen der Negation wird jeder der Quantoren durch den anderen ausgetauscht.

I.4. Zum Aufbau der Mathematik und logischen Schlüssen

Der Grundanspruch der Mathematik als Wissenschaft ist der, nur von einem begrenzten System von Grundannahmen, sogenannten Axiomen, auszugehen und alle Aussagen und Lehrsätze der Theorie durch lückenlose Argumentation mittels logischer Schlüsse aus den Axiomen abzuleiten. Die für die Herleitung einer Aussage benötigte Argumentation nennt man einen Beweis und spricht davon, die Aussage zu zeigen, zu beweisen, nachzuweisen oder zu verifizieren.

Hierzu können

- Definitionen (allgemein vereinbarte Abkürzungen oder Festlegungen)

getroffen und bereits bewiesene Aussagen verwendet werden. Für Gleichheiten beziehungsweise Äquivalenzen, die per Definition festgelegt werden, verwendet man dabei die Symbole $:=$, $=:$ beziehungsweise $:\iff$, $\iff:$, wobei der Doppelpunkt auf der Seite der neu eingeführten Bildung steht. Es kommt gelegentlich vor, dass bei einer Definition nicht direkt ersichtlich ist, warum ein eingeführtes Objekt in allen zulässigen Fällen sinnvoll erklärt oder von der richtigen Bauart ist. In solchen Situationen ist zusammen mit der Definition die sogenannte *Wohldefiniertheit* zu zeigen, das heißt, es ist im Zusammenhang mit der Definition auch die Sinnhaftigkeit des eingeführten Objekts zu beweisen.

Einmal bewiesene Aussagen bezeichnet man je nach Bedeutung, deren Einschätzung etwas subjektiv sein mag, als

- Sätze (wesentliche Erkenntnisse),
- Hauptsätze/Theoreme (Sätze von besonders weitreichender Bedeutung),
- Lemmata/Hilfssätze (kleinere/größere Hilfsresultate von eher spezieller Natur),
- Propositionen (Resultate von etwas allgemeinerem Nutzen),
- Korollare (vergleichsweise direkte Folgerungen aus anderen Resultaten).

Das wohl wichtigste Grundprinzip des Beweisens und des logischen Schließens, das sich an den Modus ponens aus Abschnitt I.2 anlehnt, besteht darin, bei zwei gegebenen Aussagen A und B aus der Wahrheit von einerseits A und andererseits $A \implies B$ auf die Wahrheit von B zu schließen. Dies erklärt zu einem gewissen Grad auch die Notation $A \implies B$ mit dem Pfeilsymbol: Gilt die Implikation, so kann man von A auf B schließen. Nützlich ist ein solcher logischer Schluss beispielsweise dann, wenn $A \implies B$ gemäß einem bereits bewiesenen mathematischen Satz gilt und die Wahrheit von A deutlich leichter zu prüfen ist als die von B.

Als konkretes Beispiel könnte der Satz die aus der Schule bekannte Regel zur Teilbarkeit durch 3 sein, die man in der Form

$$\forall x \in \mathbb{N} : ((3 \text{ teilt die Quersumme von } x) \iff (3 \text{ teilt } x))$$

mit der Menge $\mathbb{N}$ der natürlichen Zahlen schreiben kann und die ein generell anwendbares, notwendiges und hinreichendes Kriterium für die Teilbarkeit durch 3 darstellt.

Möchte man hiermit 873 auf Teilbarkeit durch 3 prüfen, so lässt sich dies logisch wie folgt aufdröseln: Zunächst setzt man 873 für x ein und erhält nach Berechnung der Quersumme 18 von 873, dass 3 genau dann 18 teilt, wenn es 873 teilt. Somit sind die Implikationen

$$\text{„}(3 \text{ teilt } 18) \implies (3 \text{ teilt } 873)\text{“}$$
$$\text{und}$$
$$\text{„}(3 \text{ teilt } 18 \text{ nicht}) \implies (3 \text{ teilt } 873 \text{ nicht})\text{“}$$

beide wahr (vergleiche mit Abschnitt I.2). Nun folgt der eigentliche logische Schluss im obigen Sinn: Da „3 teilt 18“ wahr ist (kleines Einmaleins) und die erste der beiden genannten Implikationen wahr ist, lässt sich schließen, dass „3 teilt 873“ wahr ist. Damit ist die Teilbarkeitsfrage in diesem Fall geklärt. Stellt man für eine andere Zahl anstelle von 873 fest, dass 3 die Quersumme nicht teilt, so läuft der Schluss natürlich analog, dann aber über die zweite Implikation, die das „nicht“ enthält.

Bei zukünftigen Schlüssen und Beweisen werden wir viel weniger ins Detail gehen und den begrifflichen und formalen Aufwand dadurch deutlich verringern. Trotzdem kann es immer mal wieder nützen, den gerade am Beispiel erläuterten Ablauf eines typischen logischen Schlusses zu durchdringen und im Hinterkopf zu behalten.

I.5. Grundlagen der Mengenlehre

In diesem Abschnitt werden grundlegende Definitionen und Notationen der Mengenlehre eingeführt. Wie schon zu Kapitelanfang angedeutet, beschränken wir uns dabei größtenteils auf eine Betrachtung vom Standpunkt der naiven Mengenlehre. Weiterführendes zur axiomatischen Fundierung der Mengenlehre und zu Gründen, warum es einer solchen bedarf, wird am Ende dieses Abschnitts kurz angerissen.

Grundlegend für die Mengenlehre ist in jedem Fall die Vorstellung, dass eine Menge unterscheidbare mathematische Objekte als Elemente enthält, wobei die Elementbeziehung durch das Symbol $\in$ zum Ausdruck gebracht wird. Dies ist eine Grundvorstellung, die nicht weiter formalisiert wird. Die folgende Definition verbindet diese Vorstellung mit Grundnotationen und der Festlegung, dass die Gleichheit von Mengen an ihren Elementen hängt und somit Mengen durch ihre Elemente vollständig charakterisiert sind:

Definition I.5.1. Als eine *Menge* M betrachtet man jede Ansammlung endlich oder unendlich vieler wohlunterscheidbarer mathematischer Objekte und erklärt die Aussage $x \in M$ (lies: „x ist ein Element von M“) für wahr, wenn das Objekt x in M enthalten ist. Darauf aufbauend definiert man für Mengen M und N:

- Die *Mengen-Inklusion* $M \subset N$ (lies: „M ist *Teilmenge* von N“ oder „M ist in N enthalten“ beziehungsweise äquivalent $N \supset M$ (lies: „N *Obermenge* von M“ oder „N enthält M“) bedeutet, dass jedes Element von M auch Element von N ist, also $\forall x \in M : x \in N$ gilt.
- Die *Mengen-Gleichheit* $M = N$ bedeutet, dass $M \subset N$ und $N \subset M$ gelten, also M und N dieselben Elemente enthalten.
- Die *echte* Mengen-Inklusion $M \subsetneqq N$ beziehungsweise äquivalent $N \supsetneqq M$ bedeutet, dass $M \subset N$, aber nicht $M = N$ gilt.

Für die logischen Negationen von $x \in M$, $M \subset N$, $N \supset M$, $M = N$ schreibt man naheliegenderweise $x \notin M$, $M \not\subset N$, $N \not\supset M$, $M \neq N$.

Bei der Mengen-Inklusion ist auch Gleichheit erlaubt, es gilt also $M \subset M$ für jede Menge M, was natürlich auch als $M \supset M$ geschrieben werden kann. Analog zu den Ungleichheitszeichen $\leqslant$ und $\geqslant$ bei Zahlen werden deshalb in mancher Literatur die alternativen Symbole $\subseteq$, $\supseteq$, $\subseteqq$, $\supseteqq$ für die Mengen-Inklusion verwendet. Für die strikte Mengen-Inklusion sind auch $\subsetneq$, $\supsetneq$ gebräuchlich — und seltener auch nur $\subset$, $\supset$. In dieser Vorlesung halten wir uns aber an die Konventionen der obigen Definition.

Insbesondere ist es nach der Definition sinnlos zu fragen, ob eine Menge dasselbe Objekt, beispielsweise dieselbe Zahl, mehrfach enthält, denn das Konzept der Menge sieht nur vor, dass ein Objekt entweder als

Element enthalten ist oder eben nicht. Ein Objekt ist in einer Menge nicht zweimal oder dreimal enthalten, jedenfalls nicht in dem Sinn, dass es sich um eine andere Menge handeln würde als bei nur einmaligem Enthaltensein. Die Menge $\{1, 1, 1, 1\}$ ist also zum Beispiel gleich der Menge $\{1\}$.

Um Mengen konkret durch Angabe ihrer Elemente zu beschreiben, listet man die Elemente wie folgt zwischen geschweiften Klammern auf:

Notation I.5.2. Man schreibt

- $\emptyset$ für die Menge, die kein Element enthält, die sogenannte *leere Menge*,
- $\{x\}$ für die Menge, die das Objekt x als einziges Element enthält,
- $\{x, y\}$ für die Menge, die genau die zwei Objekte x und y als Elemente enthält,
- allgemeiner $\{x_1, x_2, x_3, \ldots, x_{n-1}, x_n\}$ mit irgendeiner natürlichen Zahl n für die Menge, die genau die n Objekte x_1, x_2, x_3, …, x_{n-1}, x_n als Elemente enthält,
- $\{x_1, x_2, x_3, \ldots, \}$ für die Menge, die genau die unendlich vielen Objekte x_1, x_2, x_3, … als Elemente enthält,

Die Formulierung, dass *genau* die genannten Elemente enthalten sind, bedeutet dabei, dass außer den Objekten in der – eventuell mit … angedeuteten – Liste *keine weiteren* Objekte Element der Menge sind.

Wird hierbei dasselbe Objekt mehrfach gelistet, so hat dies gemäß den oben schon gemachten Bemerkungen keine andere Auswirkung als eine nur einfache Nennung.

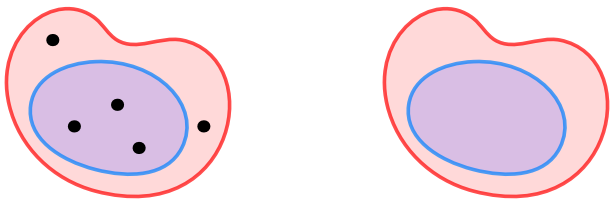
Beispiele I.5.3. Beispiele für Mengen mit konkret angegebenen Elementen sind:

- $\{5, 13\}$ (Menge mit den zwei Elementen 5 und 13),
- $\{\text{rot}, \text{grün}, \text{blau}\}$ (Menge der RGB-Grundfarben; 3 Elemente),
- $\{+, -, \cdot, :\}$ (Menge der Grundrechenarten; 4 Elemente),
- $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, \text{A}, \text{B}, \text{C}, \text{D}, \text{E}, \text{F}\}$ (Menge der Hexadezimal-Ziffern; 16 Elemente),
- $\{-7, -6, -5, \ldots, 17, 18\}$, die Menge der ganzen Zahlen von -7 bis 18; 26 Elemente,
- $\{1, 2, 3, 4, \ldots\}$ (Menge der natürlichen Zahlen; unendlich viele Elemente),
- $\{\{1, 2, 3, 4, \ldots\}, \{+, -, \cdot, :\}, \{5, 13\}, \{2, 5, -\}\}$, ein Mengensystem mit 4 Mengen als Elementen,
- $\mathbb{P} = \{2, 3, 5, 7, 11, 13, 17, 19, 23, \ldots\}$, die Menge der Primzahlen; unendlich viele Elemente,
- $\{\ , !, ", \#, \$, \%, \&, ', (,), 0, 1, 2, 3, \ldots, \text{x}, \text{y}, \text{z}, \{, |, \}, \tilde{}\}$, die Menge der ASCII-Zeichen ohne Steuerzeichen; 95 Elemente.

Notationen mit Pünktchen können problematisch sein, wenn die eigentliche Bildungsregel nicht angegeben wird und nicht jede*r unbedingt zur gleichen Interpretation kommt. Dennoch sind Pünktchen oft praktisch, dass wir sie mit etwas Vorsicht weiterhin verwenden.

Die Veranschaulichung von Mengen und Mengen-Inklusionen kann man *in geeigneten Fällen* mit Bereichen oder Umrissen in der Zeichenebene angehen: Besonders in Fällen mit endlich vielen Elementen ist es üblich, die Elemente in einem Bereich/Umriss von schematischer Bedeutung einzutragen oder zumindest anzudeuten.

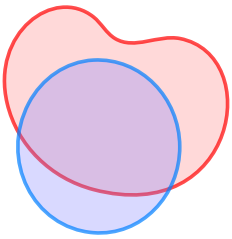
Beispiele zur Darstellung der Mengen-Inklusion zeigt Abbildung I.5.1; bei der ersten mit angedeuteten Elementen.



(I.5.1)

Vorsicht! Die obigen Zeichnungen können je nach Kontext auch Teilmengen in der Ebene mit unendlich vielen Punkten veranschaulichen. Wichtig ist also immer die Situation, die man betrachtet. Solche Zeichnungen ersetzen niemals einen Beweis!

Mengen können sich auch überlappen, ohne ineinander zu liegen:

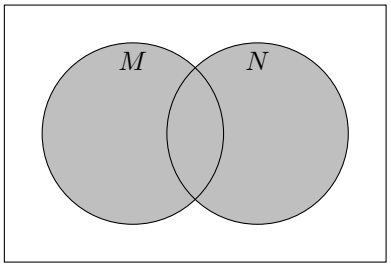


(I.5.2)

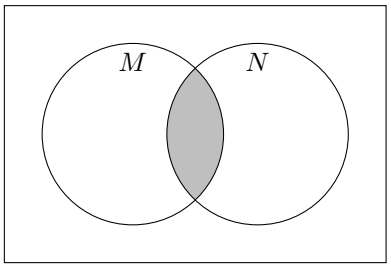
Mit den im Folgenden definierten Grundoperationen ergeben sich aber automatisch auch andere Möglichkeiten zur Angabe von Mengen:

Eine schematische Veranschaulichung von Mengenrelationen ist die durch sogenannte Venn-Diagramme (nach John Venn, 1834–1923, UK) wie in den folgenden vier Bildern.

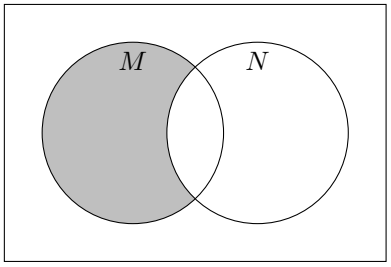
Definition I.5.4. Die *Grundoperationen* mit Mengen M, N und einem Mengensystem $\mathcal{S}$, also einer Menge $\mathcal{S}$ von Mengen, sind:
Aussonderungen: Ist $P(x)$ ein Prädikat, für dessen Variable x die Elemente von M sinnvoll eingesetzt werden können, so enthält die Aussonderungsmenge $\{x \in M \mid P(x)\}$ genau die Elemente x von M, für die $P(x)$ gilt.
Vereinigungen: Die Vereinigungsmenge $M \cup N$ von M und N enthält alle Elemente von M und alle Elemente von N und sonst keine weiteren Elemente. Für jedes x ist also $x \in M \cup N$ gleichbedeutend mit $(x \in M) \vee (x \in N)$.



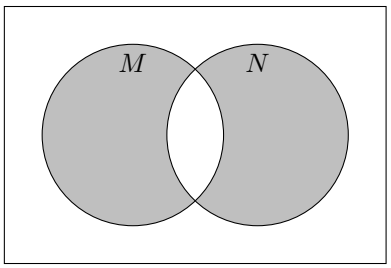
Allgemeiner hat die Vereinigungsmenge die drei gleichbedeutende Schreibweisen $\bigcup \mathcal{S} = \bigcup_{M \in \mathcal{S}} M = \bigcup \{M \mid M \in \mathcal{S}\}$ und ist die Menge, deren Elemente genau die Elemente der Elemente von $\mathcal{S}$ sind. Für jedes x ist also $x \in \bigcup_{M \in \mathcal{S}} M$ gleichbedeutend mit $\exists M \in \mathcal{S} \colon x \in M$.
Durchschnitte: Die Schnittmenge $M \cap N$ von M und N ist die Menge, die alle gemeinsamen Elemente von M und N und sonst keine weiteren Elemente enthält. Für jedes x ist also $x \in M \cap N$ gleichbedeutend mit $(x \in M) \wedge (x \in N)$.



Allgemeiner ist die Schnittmenge $\bigcap \mathcal{S} = \bigcap_{M\in\mathcal{S}} M = \bigcap\{M \mid M \in \mathcal{S}\}$ die Menge, die genau die Elemente enthält, die in allen Elementen von $\mathcal{S}$ enthalten sind. Für jedes x ist also $x \in \bigcap_{M\in\mathcal{S}} M$ gleichbedeutend mit $\forall M \in \mathcal{S}\colon x \in M$.
Mengen-Differenzen: Die Differenzmenge $M \setminus N := \{x \in M \mid x \notin N\}$ ist die Menge genau der Elemente von M, die keine Elemente von N sind.



Die *symmetrische Differenzmenge* ist $M\Delta N := (M \setminus N) \cup (N \setminus M) = (M \cup N) \setminus (M \cap N)$.



Aus Regeln für die logischen Operatoren $\vee$ und $\wedge$ ergeben sich entsprechende Regeln für den Umgang mit den Mengen-Operatoren $\cup$ und $\cap$. Für $\cup$ beispielsweise gelten die Kommutativ- und Assoziativgesetzte

$$M \cup N = N \cup M \text{ und } (M_1 \cup M_2) \cup M_3 = M_1 \cup (M_2 \cup M_3) = \bigcup_{M\in\{M_1,M_2,M_3\}} M,$$

für $\cap$ gelten diese analog, für das Zusammenspiel gelten die „Distributivgesetze“

$$(M_1 \cup M_2) \cap N = (M_1 \cap N) \cup (M_2 \cap N) \text{ und } (M_1 \cap M_2) \cup N = (M_1 \cup N) \cap (M_2 \cup N).$$

Insbesondere kann man beim Zusammentreffen von ausschließlich Vereinigungen oder ausschließlich Schnitten (aber nicht beim Zusammenspiel der beiden) auf Klammern verzichten und gebräuchliche Schreibweisen wie

$$\bigcup_{i=m}^{n} M_i := M_m \cup M_{m+1} \cup \ldots \cup M_n = \bigcup_{M\in\{M_m,M_{m+1},\ldots,M_n\}} M,$$
$$\bigcap_{i=m}^{n} M_i := M_m \cap M_{m+1} \cap \ldots \cap M_n = \bigcap_{M\in\{M_m,M_{m+1},\ldots,M_n\}} M$$

für beliebige ganze Zahlen m, n mit $m \leq n$ erklären. Wir benutzen die sinnvolle Konvention $\bigcup_{i=m}^{n} M_i := \emptyset$ im Fall $m > n$.

Ist allgemeiner für jedes Element i einer Menge I, genannt Indexmenge, eine Menge M_i gegeben[2] und ist $\mathcal{S}$ ein Mengensystem, das als Elemente genau all diese M_i enthält, so verwenden wir

$$\bigcup_{i\in I} M_i := \bigcup_{M\in\mathcal{S}} M \text{ und } \bigcap_{i\in I} M_i := \bigcap_{M\in\mathcal{S}} M$$

[2]Später werden wir in dieser Situation von einer Familie $(M_i)_{i\in I}$ von Mengen oder äquivalent von einer Abbildung $I \to \mathcal{S}$, $i \mapsto M_i$ sprechen.

auch für beliebige Indexmengen I. Auch hier benutzen wir $\bigcup_{i\in\emptyset} M_i := \emptyset$.

Unter Verwendung der Schnittmenge definieren wir außerdem:

Definition I.5.5. Zwei Mengen M und N heißen (zueinander) *disjunkt*, wenn

$$M \cap N = \emptyset$$

gilt. Allgemeiner heißen endliche viele Mengen $M_1, M_2, \ldots, M_n$ mit natürlicher Zahl n beziehungsweise unendliche viele Mengen $M_1, M_2, M_3, \ldots$ *paarweise disjunkt*, falls $M_i \cap M_j = \emptyset$ für alle $i, j \in \{1, 2, \ldots, n\}$ mit $i \neq j$ gilt. Analog heißen die in einem Mengensystem $\mathcal{S}$ enthaltenen Mengen *paarweise disjunkt*, wenn $M \cap N = \emptyset$ für alle $M, N \in \mathcal{S}$ mit $M \neq N$ gilt.

Beachten Sie, dass ein Schnitt von Mengen leer sein kann, *ohne* dass die Mengen paarweise disjunkt sind. So haben zum Beispiel die Mengen $M_1 = \{1, 3\}$, $M_2 = \{3, 5\}$ und $M_3 = \{1, 5\}$ leeren Schnitt, aber je zwei dieser Mengen haben einen nichtleeren Schnitt.

Für *disjunkte* Mengen M und N bezeichnet man $M \cup N$ als *disjunkte Vereinigung* von M und N und schreibt für $M \cup N$ auch

$$M \sqcup N.$$

Der Unterschied zwischen $\cup$ und $\sqcup$ besteht also einzig darin, dass $\sqcup$ nur zwischen disjunkten Mengen verwendet werden darf und durch seine Verwendung die Disjunktheit dann mit anzeigt. Analog verwendet man die Notationen $\bigsqcup_{i=m}^{n} M_i$ und $\bigsqcup_{M\in\mathcal{S}} M$. Gelegentlich wird das Symbol $\sqcup$ allerdings auch bei (möglicherweise) nicht disjunkten Mengen verwendet, um anzuzeigen, dass die Mengen vor der Vereinigung durch Umbenennung von Elementen disjunkt gemacht werden.

Auch für Mengen-Differenzen und ihr Zusammenspiel mit Vereinigungen und Schnitten lassen sich verschiedene allgemeingültige Regeln angeben, die teils in den Lernwerkstätten und Übungen behandelt werden. Hier halten wir vor allem eine etwas andere Sichtweise auf Differenzmengen fest:

Definition I.5.6. Für eine fixierte Menge $\mathcal{X}$, genannt die Grundmenge, und eine Teilmenge M von $\mathcal{X}$ nennt man $M^c := \mathcal{X} \setminus M$ das *Komplement* von M in $\mathcal{X}$.

Da bei der Komplement-Notation nur noch M und nicht mehr $\mathcal{X}$ auftritt, ist die Schreibweise M^c mit Bedacht zu verwenden und nur dann sinnvoll, wenn die zugrundeliegende Menge $\mathcal{X}$ klar ist und nicht variiert. Ein Vorteil der Komplement-Notation ist aber, dass mengentheoretische Folgerungen aus den Verneinungsregeln der Abschnitte I.2 und I.3 sehr prägnant angegeben werden können: Tatsächlich gelten für Teilmengen M, N, M_i einer fixierten Grundmenge $\mathcal{X}$ die Regel

$$(M^c)^c = M$$

und die mengentheoretischen *de Morganschen Gesetze*

$$(M \cup N)^c = M^c \cap N^c, \qquad (M \cap N)^c = M^c \cup N^c,$$
$$\left(\bigcup_{i\in I} M_i\right)^c = \bigcap_{i\in I} M_i^c, \qquad \left(\bigcap_{i\in I} M_i\right)^c = \bigcup_{i\in I} M_i^c$$

mit beliebiger Indexmenge I und ergänzender Konvention $\bigcap_{i\in\emptyset} M_i := \mathcal{X}$. Mit anderen Worten werden Vereinigungen unter Komplement-Bildung zu Schnitten und umgekehrt.

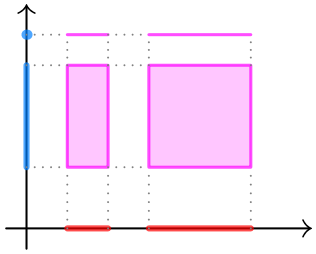
Als Nächstes wird die Liste der Grundoperationen mit Mengen noch etwas erweitert:

Definition I.5.7. Es seien M und N Mengen.

- Die Menge
$$M \times N = \{\{\{m\}, \{m, n\}\}, m \in M, n \in N\}$$
heißt das *kartesische Produkt der Mengen M und N*.
- Wir schreiben abkürzend (m, n) für ein Element $\{\{m\}, \{m, n\}\}$ von $M \times N$ und nennen (m, n) das *geordnete Paar mit erstem Eintrag m und zweitem Eintrag n*.

Beispiel I.5.8. Es ist $\{2, 4, 7\}\times\{4, 3\} = \{(2, 4), (2, 3), (4, 4), (4, 3), (7, 4), (7, 3)\}$.

Im Zusammenhang mit geordneten Paaren und kartesischen Produkten sei zunächst betont, dass (x, y) *nicht* dasselbe ist wie (y, x) (außer natürlich im Fall $x = y$). Dementsprechend ist auch $M \times N$ *nicht* dasselbe wie $N \times M$ (außer wenn $M = N$ oder $M = \emptyset$ oder $N = \emptyset$). Es kommt also auf die Reihenfolge entscheidend an; dies ist gerade das Wesen des *geordneten* Paars und des kartesischen Produkts.



(I.5.3)

Man kann das kartesische Produkt im Fall von Mengen M und N von Zahlen auf der Zahlengeraden veranschaulichen, indem man die in $M \times N$ enthaltenen Paare genau wie in der Schulmathematik als Koordinatenpunkte in der Zeichenebene interpretiert. Trägt man M auf der ersten und N auf der zweiten Achse des ebenen Koordinatensystems auf (womit man tatsächlich $M \times \{0\}$ und $\{0\} \times N$ zeichnet), so ergibt sich das kartesische Produkt $M \times N$ in der in Abbildung I.5.3 veranschaulichten Weise.

Neben geordneten Paaren (x_1, x_2) braucht man oft auch *Tripel* (x_1, x_2, x_3), Quadrupel (x_1, x_2, x_3, x_4) und allgemein für eine beliebige natürliche Zahl n sogenannte *n-Tupel* $(x_1, x_2, \ldots, x_n)$. Diese betrachtet man als Elemente des kartesischen Produkts mehrerer Mengen:

Definition I.5.9. Es seien n eine natürliche Zahl und M_1, M_2, …, M_n Mengen. Das *n-fache kartesische Produkt*

$$M_1 \times M_2 \times \ldots \times M_n,$$

oft als $\prod_{i=1}^{n} M_i$ notiert, ist die Menge der n-Tupel $(x_1, x_2, \ldots, x_n)$ mit $x_i \in M_i$ für alle $i \in \{1, 2, \ldots, n\}$. Dabei wird die Gleichheit $(x_1, x_2, \ldots, x_n) = (\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n)$ von n-Tupeln als gleichbedeutend mit $\forall i \in \{1, 2, \ldots, n\}\colon x_i = \widetilde{x}_i$ definiert.

Im oben formal eingeschlossenen Fall $n = 1$ trifft man die sehr naheliegende Konvention, dass ein 1-Tupel (x_1) nichts anderes als das Element x_1 und ein 1-faches kartesische Produkt M_1 nichts anderes als die Menge M_1 ist.

Besonders häufig braucht man das n-fache kartesische Produkt

$$M^n := \underbrace{M \times M \times \ldots \times M}_{n \text{ Faktoren}}$$

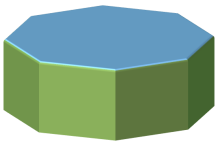
einer Menge M mit sich selbst. Es ist üblich, $((x_1, x_2), x_3) = (x_1, (x_2, x_3)) = (x_1, x_2, x_3)$ zu setzen, also derartige iterierte Paare mit Tripeln zu identifizieren. Dementsprechend gilt dann

$$(M_1 \times M_2) \times M_3 = M_1 \times (M_2 \times M_3) = M_1 \times M_2 \times M_3,$$

und man kann auf Klammerung verzichten.

Vor diesem Hintergrund können wir das kartesische Produkt $M \times N$ auch im Fall einer Menge M von Koordinatenpunkten in der Ebene und einer Menge N von Zahlen auf der Zahlengeraden veranschaulichen. In diesem Fall enthält $M \times N$ Tripel $(x, y, z) = ((x, y), z)$ aus ebenen Koordinaten (x, y) und einer weiteren Zahl z, wobei die Tripel insgesamt als räumliche Koordinaten interpretiert werden können. In einer 3D-Skizze ist somit M (beziehungsweise $M \times \{0\}$) in der Ebene der ersten beiden Achsen und N (beziehungsweise $\{(0, 0)\} \times N$) auf der dritten Achse zu zeichnen. Das Produkt $M \times N$ ergibt sich dann, wie in Abbildung I.5.4 dargestellt, als Menge von Koordinatenpunkten im 3-dimensionalen Raum.

(I.5.4)



Quelle: Wiki Commons

Auf zwei letzte wichtige Grundoperationen mit Mengen wird hier nur kurz eingegangen:

Definition I.5.10. Sei M eine Menge. Die Potenzmenge $\mathcal{P}(M)$ von M ist das Mengensystem, das genau die Teilmengen von M inklusive $\emptyset$ und M selbst als Elemente enthält.

Die wohl einfachste Potenzmenge ist $\mathcal{P}(\emptyset) = \{\emptyset\}$. Beachten Sie aber aber $\{\emptyset\} \neq \emptyset$. Ein konkreteres Beispiel ist $\mathcal{P}(\{6, 28, 496\}) = \{\emptyset, \{6\}, \{28\}, \{496\}, \{6, 28\}, \{6, 496\}, \{28, 496\}, \{6, 28, 496\}\}$. Allgemein kann mit etwas Kombinatorik gezeigt werden: Hat M genau n verschiedene Elemente für eine nicht-negative ganze Zahl n, so hat $\mathcal{P}(M)$ genau 2^n verschiedene Elemente. Wir kommen darauf in Beispiel II.2.5 zurück.

Axiom I.5.11 (Auswahlaxiom)**.** Sei $\mathcal{S}$ ein System disjunkter nicht-leerer Mengen. Dann gibt es eine Teilmenge A von $\bigcup_{M\in\mathcal{S}} M$, so dass A mit jeder in $\mathcal{S}$ enthaltenen Menge M genau ein Element gemeinsam hat.

Bei der Auswahloperation handelt es sich tatsächlich um eine Ausformulierung des sogenannten Auswahlaxioms der Mengenlehre, also um eine Grundannahme der Mathematik, die wir im Folgenden als gesetzt annehmen. Die Annahme betrifft dabei nur die pure Existenz von Auswahlmengen A mit der beschriebenen Eigenschaft. Wie die Auswahl eventuell zustande kommt und ob sie konkret angegeben werden kann, ist unerheblich. Etwas mehr hierzu wird unten im Anhang gesagt. Tatsächlich ist ein genaueres Verständnis des Auswahlaxioms an dieser Stelle aber nicht erforderlich.

Schließlich soll auch auf *Grenzen der naiven Mengenlehre* mit der Grundvorstellung einer Menge als Ansammlung von Objekten hingewiesen werden: Tatsächlich würde man im naiven Rahmen zunächst davon ausgehen, dass man auch das Mengensystem $\mathcal{S}$ *aller* Mengen, manchmal Allmenge genannt, bilden kann. Dies ist aber hochproblematisch, denn durch Aussonderung könnte man auch die Menge $\mathcal{R} := \{M \in \mathcal{S} \mid M \notin M\}$ bilden, und die Frage, ob $\mathcal{R}$ sich selbst enthält, ergäbe den Widerspruch $\mathcal{R} \in \mathcal{R} \iff \mathcal{R} \notin \mathcal{R}$. Dieser als *Russellsches Paradoxon* bekannte Widerspruch hat maßgeblich zur Entwicklung eines modernen Mengenbegriffs beigetragen und das Verständnis dafür geprägt, dass die naive Mengenlehre nicht ausreicht und man tatsächlich ein *präzises Axiomensystem als Grundlage* der Mengenlehre braucht, in dessen Rahmen die Bildung einer *Menge aller Mengen* nicht *zulässig* ist.

Als Grundlage der Mengenlehre und damit der Mathematik hat sich letztlich das *Zermelo-Fraenkel-Axiomensystem der Mengenlehre*, benannt nach den Mathematikern E. Zermelo (1871–1953) und A. Fraenkel (1891–1965), bewährt. Dieses Axiomensystem ist heutzutage sehr weitgehend akzeptiert und ähnelt an vielen Stellen oben schon diskutierten Bildungen. Eine detaillierte Behandlung der Axiome geht über den Vorlesungsstoff hinaus. Es soll allerdings auch nicht fälschlicherweise der Eindruck entstehen, dass die Axiome unzugänglich wären oder nicht ohne Weiteres hingeschrieben werden könnten. Daher seien die Axiome für Interessierte zumindest im Anhang festgehalten (siehe Abschnitt V.1)

Die historische Entwicklung von der naiven Mengenlehre des späten 19. Jahrhunderts zu dem in der obigen Form um das Jahr 1930 komplettierten Axiomensystem war übrigens ein langer Prozess mit vielen Beteiligten. Einen ersten Hinweis, dass die naive Mengenlehre an ihre Grenzen stößt, mag man tatsächlich in einer als *prägnante Anekdote* überlieferten Konversation zwischen den Mathematikern und Gründervätern der Mengenlehre G. Cantor (1845–1918) und R. Dedekind (1831–1916) sehen: „Dedekind äußerte hinsichtlich des Begriffs der Menge, er stelle sich eine Menge vor wie einen geschlossenen Sack, der ganz bestimmte Dinge enthalte, die man aber nicht sähe, und von denen man nichts wisse, außer daß sie vorhanden und bestimmt seien. Einige Zeit später gab Cantor seine Vorstellung einer Menge zu erkennen: Er richtete seine kolossale Figur hoch auf, beschrieb mit erhobenem Arm eine großartige Geste und sagte mit einem ins Unbestimmte gerichteten Blick: ‚Eine Menge stelle ich mir vor wie einen Abgrund.'" (Überlieferung nach F. Bernstein [**1**]).

KAPITEL II

Abbildungen, Relationen, Zahlen

In diesem und den nächsten Kapiteln beschäftigen wir uns viel mit den grundlegenden, auch aus der Schule bekannten *Zahlbereichen*

$\mathbb{N} = \{1, 2, 3, \ldots\}$ (*natürliche Zahlen*), $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$,
$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ (*ganze Zahlen*),
$\mathbb{Q}$ (*rationale Zahlen*; Brüche mit Zähler aus $\mathbb{Z}$, Nenner aus $\mathbb{N}$),
$\mathbb{R}$ (*reelle Zahlen*; Zahlen der Zahlengeraden)

und den *Rechengrundgesetzen* auf diesen Bereichen. Für die Zahlbereiche gelten die Inklusionen

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Das Ziel dieses Kapitels ist zweigeteilt: Zum einen werden allgemeine, für die Mathematik fundamentale Konzepte wie Abbildungen/Funktionen und Relationen eingeführt. Zum anderen werden mit Hilfe dieser Konzepte präzise Konstruktionen der Zahlbereiche $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$ nur auf Grundlage der (Axiome der) Mengenlehre gegeben und damit die Existenz dieser Zahlbereiche bewiesen werden. Die Konstruktion von $\mathbb{R}$ (und $\mathbb{C}$), die etwas mehr Analysis braucht, wird auf ein späteres Kapitel vertagt.

II.1. Abbildungen

Die gleichbedeutenden Begriffe Abbildung und Funktion sind in der Mathematik sehr grundlegend. Dieser Begriff lässt sich wie folgt fassen:

Definition II.1.1. Es seien $\mathcal{X}$ und $\mathcal{Y}$ beliebige Mengen. Eine *Funktion* oder *Abbildung* f von $\mathcal{X}$ nach $\mathcal{Y}$ ist eine Zuordnungsvorschrift, die jedem Element x von $\mathcal{X}$ ein eindeutiges Element y von $\mathcal{Y}$ zuordnet. Man sagt, f bilde das Element x auf das zugehörige Element y ab, dieses y sei das Bild von x unter f oder dieses y sei der *Funktionswert* von f an der Stelle x oder zum *Argument* x. Man nennt $\mathcal{X}$ die Definitionsmenge oder den *Definitionsbereich* und $\mathcal{Y}$ das Ziel oder den *Zielbereich* von f. Im Fall $\mathcal{X} = \mathcal{Y}$ spricht man von einer *Selbstabbildung*.

Besonders betont sei bei dieser Definition das *zentrale Existenz- und Eindeutigkeitsprinzip*, gemäß dem zu einem beliebigen Element x von $\mathcal{X}$ *genau ein* zugehöriges Element y von $\mathcal{Y}$ existiert. Sowohl die Nicht-Existenz des Elements y als auch die Existenz mehr als eines Elements y (zum gleichen x) werden ausgeschlossen.

Notationen II.1.2. Es seien $\mathcal{X}$, $\mathcal{Y}$ Mengen und f eine Abbildung von $\mathcal{X}$ nach $\mathcal{Y}$.

(a) Dass f Definitionsbereich $\mathcal{X}$ und Ziel $\mathcal{Y}$ hat, also eine Abbildung von $\mathcal{X}$ nach $\mathcal{Y}$ ist, drückt man durch die Notation $f\colon \mathcal{X} \to \mathcal{Y}$ aus.
(b) Ordnet f einem Element $x \in \mathcal{X}$ den Funktionswert $y \in \mathcal{Y}$ zu, so schreibt man $f(x)$ für y oder notiert kurz $x \mapsto y$.

Dabei ist es Konvention, den Pfeil $\to$ bei Angabe von Definitionsbereich und Ziel, aber den etwas anderen Pfeil $\mapsto$ für Zuordnungen der Elemente zu nutzen.

Beispiele II.1.3. In der Praxis gibt man eine Abbildung f von $\mathcal{X}$ nach $\mathcal{Y}$ konkret an, indem man neben $\mathcal{X}$, $\mathcal{Y}$ den Funktionsterm $f(x)$ oder die Zuordnung $x \mapsto f(x)$ beziehungsweise $x \mapsto y$ für alle $x \in \mathcal{X}$ eindeutig spezifiziert. Wir betrachten hierzu folgende Beispiele:

- Eine Abbildung
$$f\colon \{1,2,4,5\} \to \{-3,-2,0,4,10\}$$
ist durch die Zuordnungen
$$1 \mapsto -2, \quad 2 \mapsto -2, \quad 4 \mapsto 4, \quad 5 \mapsto 10$$
gegeben

  Die gleiche Zuordnungsvorschrift kann auch anders angegeben werden, z.B. durch eine *Wertetabelle*
$$\begin{array}{c|c|c|c|c} x & 1 & 2 & 4 & 5 \\ \hline f(x) & -2 & -2 & 4 & 10 \end{array},$$
durch $f(x) := x^2-3x$ oder durch $f(x) := \begin{cases} -2 & \text{falls } x < 3 \\ 6x-20 & \text{falls } x > 3 \end{cases}$ für alle $x \in \{1,2,4,5\}$. Wie zuletzt eine *geschweifte Klammer für Fallunterscheidungen* zu verwenden, ist allgemein üblich und bedeutet an dieser Stelle einerseits $f(x) := -2$ im Fall $x < 3$ und andererseits $f(x) := 6x-20$ im Fall $x > 3$.
- Eine andere Abbildung
$$f\colon \{\emptyset, 5, \mathbb{N}\} \to \mathbb{Q}$$
erhält man durch die Festlegungen
$$f(\emptyset) := \frac{3}{2}, \quad f(5) := 4, \quad f(\mathbb{N}) := -\frac{1}{12}.$$
- Eine weitere Abbildung
$$f\colon \mathbb{N} \to \mathbb{N}$$
wird durch
$$f(n) := n^2 \quad \text{für alle } n \in \mathbb{N}$$
definiert. Diese Abbildung ordnet zum Beispiel $1 \mapsto 1$, $2 \mapsto 4$, $3 \mapsto 9$, $4 \mapsto 16$ zu. Da es unendlich viele Zahlen in $\mathbb{N}$ gibt, reichen diese vier Beispiele (oder jede andere endliche Anzahl) aber nicht aus, um die Abbildung in Gänze zu beschreiben. Zum Beispiel haben 1, 2, 3, 4 dieselben Funktionswerte wie unter f auch unter $g\colon \mathbb{N} \to \{0,1,4,9,16\}$ mit
$$g(n) := \begin{cases} n^2, & \text{falls } n \leq 4 \\ 0, & \text{falls } n \geq 5 \end{cases}$$
für alle $n \in \mathbb{N}$ und unter $h\colon \mathbb{N} \to \mathbb{N}$ mit $h(n) := n^2 + (n-1)(n-2)(n-3)(n-4)$ für alle $n \in \mathbb{N}$, doch ab 5 unterscheiden sich die Werte von f, g und h.
- Man kann die Addition natürlicher Zahlen als Abbildung
$$f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$$
auffassen, die durch
$$f((m,n)) := m+n \quad \text{für alle } (m,n) \in \mathbb{N} \times \mathbb{N}$$
festgelegt ist. Dies ist ein erstes Beispiel einer *Abbildung von zwei Variablen*, die beispielsweise $(1,2) \mapsto 3$ und $(4,2) \mapsto 6$ zuordnet und von beiden Einträgen m und n der einzusetzenden Paare (m,n) abhängt.

  Zur Vereinfachung der Notation schreiben wir bei solchen Abbildungen zukünftig $f(m,n)$ statt $f((m,n))$. Übrigens ist es durchaus üblich, die hier betrachtete Abbildung + statt f zu nennen, und analog dazu auch bei anderen Abbildungen, die sich direkt aus einer Rechenoperation ergeben, das Rechensymbol als Name der Abbildung zu verwenden.
- Eine Funktion
$$f\colon \mathbb{R} \to \mathbb{R}$$
erhält man durch die Festlegung
$$f(x) := x^3-x \quad \text{für alle } x \in \mathbb{R}.$$
In alternativer verbreiteter Schreibweise kann diese Funktion auch als $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3-x$ angegeben werden.

- Eine weitere Funktion

$$f\colon \mathcal{X} \to \mathcal{Y} \quad \text{von } \mathcal{X} := \left\{x \in \mathbb{R} \mid 1 \leq x < \frac{11}{2}\right\} \text{ nach } \mathcal{Y} := \left\{y \in \mathbb{R} \mid -\frac{1}{2} < y < \frac{13}{2}\right\}$$

  wird durch

$$f(x) := -x^2{+}6x{-}3 \quad \text{für alle } x \in \mathcal{X}$$

  definiert. Um die *Wohldefiniertheit* dieser Funktion f sicherzustellen, muss man allerdings noch zeigen, dass $-\frac{1}{2} < -x^2{+}6x{-}3 < \frac{13}{2}$ für alle $x \in \mathcal{X}$ gilt: Tatsächlich erhält man durch Umschreiben $-x^2{+}6x{-}3 = 6{-}(x{-}3)^2$ und Rechnen mit Ungleichungen aber sogar $-\frac{1}{4} = 6{-}\left(\frac{5}{2}\right)^2 < -x^2{+}6x{-}3 \leq 6$ für alle $x \in \mathcal{X}$, womit die Wohldefiniertheit von f gesichert ist.

  Nebenbei sehen wir, dass $g(x) := -x^2{+}6x{-}3$ für $x \in \mathcal{X}$ auch eine Funktion $g\colon \mathcal{X} \to \mathcal{Z}$ in den kleineren Zielbereich $\mathcal{Z} := \left\{z \in \mathbb{R} \mid -\frac{1}{4} < z \leq 6\right\} \subsetneq \mathcal{Y}$ ergibt.
- Die Versuche, $f\colon \mathbb{N} \to \mathbb{N}$ durch $f(n) := n^2{-}n$ für alle $n \in \mathbb{N}$ oder $g\colon \mathbb{N}_0 \to \mathbb{N}_0$ durch $g(0) := n$ und $g(n) := 0$ für alle $n \in \mathbb{N}$ zu definieren, geben allerdings *keine* wohldefinierten Abbildungen, da $f(1)$ nicht im Zielbereich $\mathbb{N}$ liegt und $g(0)$ nicht eindeutig definiert ist. (Auch $g(0) := \mathbb{N}$ geht nicht, da zwar $\mathbb{N} \subset \mathbb{N}_0$, aber eben nicht $\mathbb{N} \in \mathbb{N}_0$ gilt).
- Weiterhin ist auch $h\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto \frac{1}{x}$ *keine* wohldefinierte Funktion, da $\frac{1}{0}$ nicht erklärt und deshalb der Funktionswert $h(0)$ nicht definiert ist. Die Variante $\widetilde{h}\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $x \mapsto \frac{1}{x}$ vermeidet dieses Problem natürlich und ist wohldefiniert.

Einfache Möglichkeiten zur Beschreibung einer Abbildung bestehen, falls der Definitionsbereich nur endlich viele Elemente hat. Dann kann die Abbildung durch Wertetabellen oder Zuordnungspfeile vollständig dargestellt werden. Die entscheidende Abbildungseigenschaft verlangt dabei, dass bei jedem Element des Definitionsbereichs genau ein Zuordnungspfeil beginnt und zum zugehörigen Element im Ziel weist. Dagegen bestehen für ein Element des Ziels prinzipiell alle Möglichkeiten: Dort können kein Pfeil, ein Pfeil, mehrere Pfeile oder im Extremfall einer konstanten Abbildung auch alle Pfeile enden.

Dagegen kann man für Definitionsbereiche mit unendlich vielen Elementen, nur einige Fälle, aber nie die gesamte Zuordnungsvorschrift auf diese Weise darstellen.
Natürlich müssen Definitionsbereich und Ziel einer Abbildung nicht unbedingt verschieden oder als Mengen disjunkt sein, sondern können gemeinsame Elemente enthalten (wie es mit Ausnahme der Additions-Abbildung tatsächlich bei allen obigen Beispielen und Bildern der Fall ist). Normalerweise stellt man diese beiden Mengen dennoch nicht überlappend dar und trägt gemeinsame Elemente zweimal separat ein. Sogar wenn Definitionsbereich und Ziel übereinstimmen, zieht man es der Übersichtlichkeit halber meist vor, zwei Kopien derselben Menge einzuzeichnen. Gelegentlich weicht man hiervon aber auch ab und verwendet alternative Darstellungen.

Speziell für eine Funktion $f\colon \mathcal{X} \to \mathcal{Y}$ von $\mathcal{X} \subset \mathbb{R}$ nach $\mathcal{Y} \subset \mathbb{R}$ kann man oft auch den *(Funktions-)Graphen*

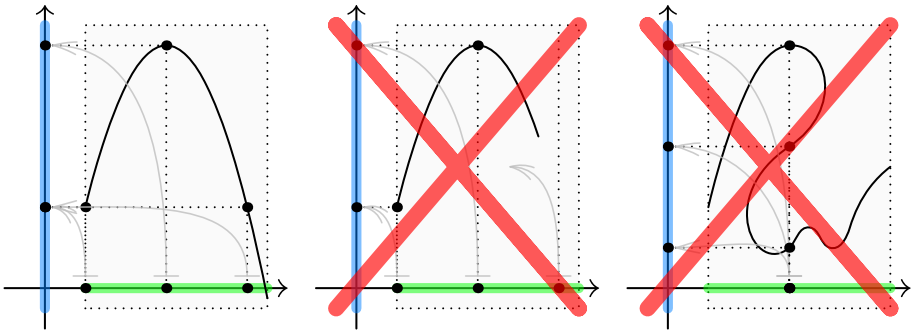
$$\mathrm{G}_f := \{(x, y) \in \mathcal{X}{\times}\mathcal{Y} \mid f(x) = y\} \subset \mathbb{R}^2$$

als Menge von Koordinatenpunkten der Ebene darstellen. Man trägt dazu den Definitionsbereich $\mathcal{X}$ auf der ersten, den Zielbereich $\mathcal{Y}$ auf der zweiten Achse des ebenen Koordinatensystems ein und bekommt als Darstellung von G_f in gutartigen Fällen eine Kurve durch alle Punkte der Form $(x, f(x))$ mit $x \in \mathcal{X}$.

Was gutartig heißt, sehen wir später genauer. Als Beispiel einer Abbildung, die keinen wirklich aussagekräftigen Graphen hat, kann man die Funktion $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

betrachten.



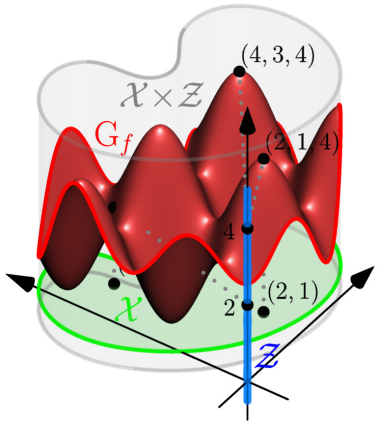
(II.1.1)

Die definierende Eigenschaft der Funktion verlangt dabei, dass sich wie im ersten Bild der Abbildung II.1.1 über jedem Punkt $x \in \mathcal{X}$ (den wir jetzt immer mit $(x,0) \in \mathcal{X} \times \{0\}$ identifizieren) genau ein Punkt von G_f befindet. Dass sich wie in den beiden anderen Bildern über einem Punkt von $\mathcal{X}$ kein Punkt oder mehrere Punkte der Kurve befinden, dass die Kurve also „aufhört“ oder unter/über sich selbst „zurück läuft“ ist für einen Funktionsgraphen als Kurve somit nicht möglich. Der entscheidende Vorteil der Graphendarstellung ist der, dass man an G_f oft *alle* zu f gehörigen Zuordnungen ablesen kann und nicht nur die Zuordnung einiger Beispiel-Elemente. Daher wird mit G_f (jedenfalls im Rahmen der Zeichengenauigkeit) in guten Fällen die *gesamte Zuordnungsvorschrift dargestellt.*

Auch für eine Funktion $f\colon \mathcal{X} \to \mathcal{Z}$ von einem 2-dimensionalen Definitionsbereich $\mathcal{X} \subset \mathbb{R}^2$ nach $\mathcal{Z} \subset \mathbb{R}$, also eine *Funktion von zwei Variablen*, kann man die Zuordnung einzelner Elemente mittels Pfeilen verdeutlichen oder in gutartigen Fällen die Funktion insgesamt durch ihren (Funktions-)Graphen

$$\mathrm{G}_f := \{(x,y,z) \in \mathcal{X}\times\mathcal{Z} \mid f(x,y) = z\} \subset \mathbb{R}^3$$

darstellen, wobei wir wie in Abschnitt I.5 wieder $((x,y),z) = (x,y,z)$ und $\mathbb{R}^2\times\mathbb{R} = \mathbb{R}^3$ identifizieren. Beide Möglichkeiten werden für eine Beispiel-Funktion[1] in Abbildung II.1.2 gezeigt. Dabei bildet der Graph jetzt eine Fläche im 3-dimensionalen Raum durch alle Punkte der Form $(x,y,f(x,y))$ mit $(x,y) \in \mathcal{X}$, es befindet sich aber nach wie vor über jedem Punkt $(x,y) \in \mathcal{X}$ (natürlich identifiziert mit $(x,y,0) \in \mathcal{X}\times\{0\}$) genau ein Punkt von G_f.



(II.1.2)

Als letzten Fall betrachten wir Darstellungen einer Funktion $f\colon \mathcal{X} \to \mathcal{Y}$ von $\mathcal{X} \subset \mathbb{R}$ in einen 2-dimensionalen Zielbereich $\mathcal{Y} \subset \mathbb{R}^2$. Im gutartigen Fall spricht man bei einer solchen Funktion von einer *Kurve* f in $\mathbb{R}^2$ und stellt oft das Bild

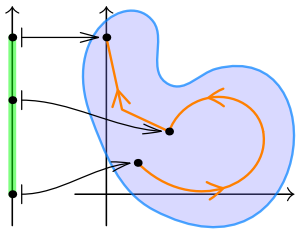
$$f(\mathcal{X}) := \{(y,z) \in \mathcal{Y} \mid \exists x \in \mathcal{X}\colon f(x) = (y,z)\} \subset \mathbb{R}^2$$

[1] Tatsächlich ist die in Abbildung II.1.2 geplottete Funktion auf dem dargestellten Definitionsbereich $\mathcal{X} \subset \mathbb{R}^2$ durch $f(x,y) := 2{-}2\sin((x{+}y)\frac{\pi}{2})\sin((x{-}y)\frac{3\pi}{8})$ für $(x,y) \in \mathcal{X}$ gegeben. Die hier verwendete Sinus-Funktion sin und die Kreiszahl π werden in der Vorlesung aber erst deutlich später eingeführt.

von f dar, obwohl es für eine Gesamtdarstellung der Funktion auch hier den Graphen

$$\mathrm{G}_f := \{(x,y,z) \in \mathcal{X}\times\mathcal{Y} \mid f(x) = (y,z)\} \subset \mathbb{R}^3,$$

braucht. Ein Beispiel für die erste Darstellungsweise wird in Abbildung II.1.3 gezeigt. Tatsächlich enthält in diesem Fall das Bild $f(\mathcal{X})$ schon relativ viel Information über die Funktion und kann zudem in der Ebene $\mathbb{R}^2$ leichter gezeichnet, interpretiert und verstanden werden als der Graph G_f, der im Raum $\mathbb{R}^3$ darzustellen ist. Das Bild muss sich übrigens nicht unbedingt so gutartig wie in Abbildung II.1.3 verhalten, sondern kann sich auch in sogenannten Selbstschnitten überkreuzen und/oder auf sich selbst zurück laufen. Der Graph einer solchen Funktion hat dagegen die Eigenschaft, dass sich in jedem der parallelen Ebenenstücke $\{x\}\times\mathcal{Y}$ mit $x \in \mathcal{X}$ genau ein Punkt von G_f befindet.



(II.1.3)

Der Graph G_f einer Funktion $f\colon \mathcal{X} \to \mathcal{Y}$ ist immer auch das Bild der Graphenabbildung $\mathcal{X} \to \mathcal{X}\times\mathcal{Y}$, $x \mapsto (x, f(x))$. Insofern haben wir mit den Graphen G_f der Abbildungen II.1.3 beziehungsweise II.1.2 auch schon Beispiele für *Bilder* von Funktionen von einem 1- bzw. 2-dimensionalen Definitionsbereich $\mathcal{X}$ in den 3-*dimensionalen Zielbereich* $\mathcal{X}\times\mathcal{Y} \subset \mathbb{R}^3$ gesehen.

Als nächstes führen wir Konzepte im Umfeld der gerade schon diskutierten Begriffe Bild und Graph präzise und allgemein ein:

Definition II.1.4. Es sei $f\colon \mathcal{X} \to \mathcal{Y}$ eine Abbildung von einer Menge $\mathcal{X}$ in eine Menge $\mathcal{Y}$.

(I) Das *Bild einer Teilmenge* $A \subset \mathcal{X}$ unter f erklärt man als

$$f(A) := \{y \in \mathcal{Y} \mid \exists x \in A\colon f(x) = y\} \subset \mathcal{Y}.$$

Oft wird dies auch etwas informeller als $f(A) = \{f(x) \mid x \in A\}$ geschrieben. Für einzelne $x \in \mathcal{X}$ ist das Bild $f(\{x\}) = \{f(x)\}$. Speziell heißt $\mathrm{Bild}(f) := f(\mathcal{X}) \subset \mathcal{Y}$ das *Bild der Abbildung* f.

(II) Das *Urbild einer Teilmenge* $B \subset \mathcal{Y}$ unter f erklärt man als

$$f^{-1}(B) := \{x \in \mathcal{X} \mid f(x) \in B\} \subset \mathcal{X}.$$

Für einzelne $y \in \mathcal{Y}$ gilt damit[2] $f^{-1}(\{y\}) = \{x \in \mathcal{X} \mid f(x) = y\}$. Die definierende Eigenschaft der Abbildung erzwingt generell $f^{-1}(\mathcal{Y}) = \mathcal{X}$ und auch $f^{-1}(\mathrm{Bild}(f)) = \mathcal{X}$.

(III) Den *(Funktions-)Graphen* G_f von f erklärt man als

$$\mathrm{G}_f := \{(x,y) \in \mathcal{X}\times\mathcal{Y} \mid f(x) = y\} \subset \mathcal{X}\times\mathcal{Y}.$$

Oft wird dies auch – etwas informeller – als $\mathrm{G}_f = \{(x, f(x)) \mid x \in \mathcal{X}\}$ geschrieben, und anstelle von G_f notiert man gleichbedeutend auch $\mathrm{Graph}(f)$.

Bemerkungen II.1.5.

- Das Bild einer Abbildung $f\colon \mathcal{X} \to \mathcal{Y}$ ist die Menge B mit den wenigsten möglichen Elementen, so dass noch $\mathrm{G}_f \subset \mathcal{X}\times B$ gilt. Das Bild *kann* eine *echte* Teilmenge des Ziels sein.
- Für eine Abbildung $f\colon \mathcal{X} \to \mathcal{Y}$ und $A \subset \mathcal{X}$, $B \subset \mathcal{Y}$ sagt man, f bilde (Elemente) von A *nach* B ab, wenn $f(A) \subset B$ gilt. Stärker sagt man, f bilde A *auf* B ab, wenn $f(A) = B$ gilt (wobei der Unterschied sich in der Präposition „auf“ anstelle von „nach“ niederschlägt). Insbesondere bildet f immer von $\mathcal{X}$ *nach* $\mathcal{Y}$ und $\mathcal{X}$ *auf* $f(\mathcal{X}) = \mathrm{Bild}(f)$ ab.
- Manchmal spricht man bei einer Abbildung auch vom Wertebereich/Bildbereich und meint entweder Bild oder Ziel. Wir vermeiden diese Begriffe aufgrund der Doppeldeutigkeit vorerst.

[2] In mancher Literatur wird für das Urbild $f^{-1}(\{y\})$ auch $f^{-1}(y)$ geschrieben. Wir vermeiden dies aber fürs Erste, um Verwechslungen mit der demnächst eingeführten Umkehrfunktion f^{-1} vorzubeugen.

- Vorsicht! Obwohl das Urbild mit dem Symbol $f^{-1}(\ldots)$ notiert wird, ist das Urbild auch in Fällen definiert, in denen die demnächst eingeführte Umkehrfunktion f^{-1} *nicht* existiert. Betrachten wir zum Beispiel die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(x) = 0$ für alle $x \in \mathbb{R}$, so ist $f^{-1}\{0\} = \mathbb{R}$.

Erste Regeln für (Ur-)Bilder lesen wir direkt aus den Definitionen ab: Für $A_1 \subset A_2 \subset \mathcal{X}$ gilt stets $f(A_1) \subset f(A_2)$ und für $B_1 \subset B_2 \subset \mathcal{Y}$ stets $f^{-1}(B_1) \subset f^{-1}(B_2)$. Des Weiteren gilt:

Satz II.1.6. *Es sei $f\colon \mathcal{X} \to \mathcal{Y}$ eine Abbildung von einer Menge $\mathcal{X}$ in eine Menge $\mathcal{Y}$. Für Teilmengen A_1, A_2 von $\mathcal{X}$ und B_1, B_2 von $\mathcal{Y}$ gelten dann stets:*

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2), \qquad f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2),$$
$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2), \qquad f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2).$$

Dass beim Bild des Schnitts tatsächlich nur „$\subset$“ und nicht „$=$“ gelten kann, sieht man schon an einer Abbildung $f\colon \{a_1, a_2\} \to \{y\}$ mit $a_1 \neq a_2$, denn dann ist $f(\{a_1\} \cap \{a_2\}) = f(\emptyset) = \emptyset$, aber $f(\{a_1\}) \cap f(\{a_2\}) = \{f(a_1)\} \cap \{f(a_2)\} = \{y\} \cap \{y\} = \{y\}$.

Zum Beweis des Satzes nutzen wir typische Techniken, um Mengen-Inklusionen und Mengen-Gleichheiten nachzuweisen:

Beweis. Wir behandeln erst das Bild der Vereinigung: Für $i \in \{1, 2\}$ gilt $A_1 \cup A_2 \supset A_i$, nach Vorbemerkung zum Satz dann auch $f(A_1 \cup A_2) \supset f(A_i)$ und daher $f(A_1 \cup A_2) \supset f(A_1) \cup f(A_2)$. Als Nächstes zeigen wir die umgekehrte Inklusion $f(A_1 \cup A_2) \subset f(A_1) \cup f(A_2)$ gemäß Definition der Inklusion: Sei $y \in f(A_1 \cup A_2)$. Nach Definition des Bilds gibt es dann $x \in A_1 \cup A_2$ mit $f(x) = y$. Dabei ist $x \in A_1$ oder $x \in A_2$. Im ersten Fall folgt $y = f(x) \in f(A_1)$, im zweiten Fall $y = f(x) \in f(A_2)$. In beiden Fällen gilt also $y \in f(A_1) \cup f(A_2)$. Damit ist die Inklusion $f(A_1 \cup A_2) \subset f(A_1) \cup f(A_2)$ und insgesamt auch $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$ gezeigt.

Beim Bild des Schnitts erhält man aus $A_1 \cap A_2 \subset A_i$ für $i \in \{1, 2\}$ mit der Vorbemerkung $f(A_1 \cap A_2) \subset f(A_i)$ und insgesamt $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$.

Die Regel für das Urbild der Vereinigung ergibt sich durch schrittweise Äquivalenzumformung[3] mit der Definition von Urbild und Vereinigung durch

$$\begin{aligned} x \in f^{-1}(B_1 \cup B_2) &\iff f(x) \in B_1 \cup B_2 \iff (f(x) \in B_1) \vee (f(x) \in B_2) \\ &\iff (x \in f^{-1}(B_1)) \vee (x \in f^{-1}(B_2)) \iff x \in f^{-1}(B_1) \cup \in f^{-1}(B_2) \end{aligned}$$

für $x \in \mathcal{X}$. Beim Urbild des Schnitts geht man analog vor (mit $\cap$ statt $\cup$ und $\wedge$ statt $\vee$). □

Sehr wichtige Eigenschaften von Abbildungen, die ab jetzt immer wieder auftreten, sind:

Definition II.1.7. Es sei $f\colon \mathcal{X} \to \mathcal{Y}$ eine Abbildung von einer Menge $\mathcal{X}$ in eine Menge $\mathcal{Y}$.

(I) Man nennt f *injektiv* oder eine *Injektion*, wenn zu jedem $y \in \mathcal{Y}$ höchstens ein $x \in \mathcal{X}$ mit $f(x) = y$ existiert, mit anderen Worten also, wenn für alle $x, \widetilde{x} \in \mathcal{X}$ die Implikation

$$f(\widetilde{x}) = f(x) \implies \widetilde{x} = x$$

gilt. Man nennt injektive Abbildungen auch Einbettungen und zeigt Injektivität von f gelegentlich durch die Notation $f\colon \mathcal{X} \hookrightarrow \mathcal{Y}$ an.

(II) Man nennt f *surjektiv* oder eine *Surjektion*, wenn zu jedem $y \in \mathcal{Y}$ mindestens ein $x \in \mathcal{X}$ mit $f(x) = y$ existiert, mit anderen Worten also, wenn $\mathrm{Bild}(f) = \mathcal{Y}$ gilt. Gelegentlich zeigt man Surjektivität von f durch die Notation $f\colon \mathcal{X} \twoheadrightarrow \mathcal{Y}$ an.

(III) Man nennt f *bijektiv* oder eine *Bijektion*, wenn zu jedem $y \in \mathcal{Y}$ genau ein $x \in \mathcal{X}$ mit $f(x) = y$ existiert, mit anderen Worten also, wenn f sowohl injektiv als auch surjektiv ist.

In *graphischen Darstellungen mit Zuordnungspfeilen* bedeuten Injektivität, Surjektivität bzw. Bijektivität, dass bei jedem Element des Ziels $\mathcal{Y}$ höchstens ein, mindestens ein beziehungsweise genau ein Pfeil endet.

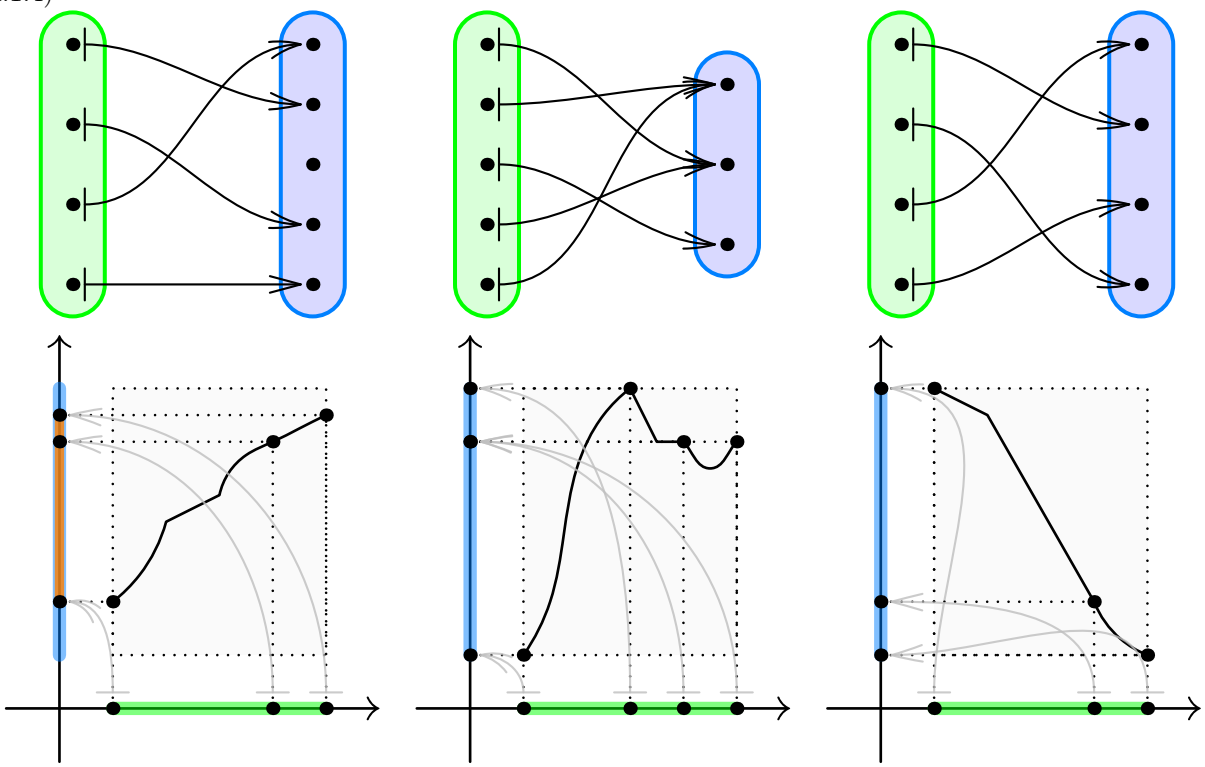
---

[3] Aneinandergereihte Äquivalenzen $S_1 \iff S_2 \iff S_3 \iff \ldots \iff S_{n-1} \iff S_n$ für Aussagen $S_1, S_2, \ldots, S_n$ verwenden wir ab jetzt immer als abkürzende Schreibweise mit der auch bei Äquivalenzumformungen verbreiteten Bedeutung, dass die auftretenden Aussagen wechselseitig äquivalent sind. In der ursprünglichen Notation der Aussagenlogik müssten wir hierfür streng genommen $(S_1 \iff S_2) \wedge (S_2 \iff S_3) \wedge \ldots \wedge (S_{n-1} \iff S_n)$ schreiben.

Für Definitions- und Zielbereiche $\mathcal{Y}$ mit endlich vielen Elementen wird dies in den oberen Bildern der Abbildung II.1.4 dargestellt. Wenn nicht alle (eventuell unendlich vielen) Elemente und/oder Zuordnungspfeile graphisch dargestellt werden können, gilt das Gesagte sinngemäß für alle prinzipiell in Frage kommenden Elemente und Pfeile.

Auch am Graphen G_f einer Funktion $f\colon \mathcal{X} \to \mathcal{Y}$ mit $\mathcal{X} \subset \mathbb{R}$ und $\mathcal{Y} \subset \mathbb{R}$ kann man die Abbildungseigenschaften ablesen, was beispielhaft in den unteren Bildern der Abbildung II.1.4 gezeigt wird: Injektivität bedeutet, dass sich auf jeder horizontalen Geraden in einer Höhe $y \in \mathcal{Y}$ höchstens ein Punkt von G_f befindet. Dies kann sich in gutartigen Fällen nur so manifestieren, dass die Graphenkurve G_f von links nach rechts durchgehend steigt oder fällt. Surjektivität liegt vor, wenn es auch immer mindestens einen solchen Punkt gibt, also jede horizontale Gerade einer Höhe $y \in \mathcal{Y}$ den Graph G_f tatsächlich trifft. Bijektivität erfordert beides zusammen.

(II.1.4)



Injektionen (links), Surjektionen (mittig) und Bijektionen (rechts) $f\colon \mathcal{X} \to \mathcal{Y}$ für endliche Mengen $\mathcal{X}$ und $\mathcal{Y}$ (oben) und in Graphendarstellung für $\mathcal{X} = \{x \in \mathbb{R} \mid 1 \leq x \leq 5\}$ und $\mathcal{Y} = \{y \in \mathbb{R} \mid 1 \leq y \leq 6\}$ (unten), links unten mit Bild $f(\mathcal{X}) = \{y \in \mathbb{R} \mid 2 \leq y \leq \frac{11}{2}\} \subsetneq \mathcal{Y}$

Bemerkungen II.1.8.
Ob eine Abbildung injektiv/surjektiv/bijektiv ist, hängt auch bei (wie in der Schule) durch Funktionsterme gegebenen Funktionen entscheidend davon ab, mit welchem Definitionsbereich und Ziel diese betrachtet werden.

Für eine *Injektion, Surjektion beziehungsweise Bijektion* $\mathcal{X} \to \mathcal{Y}$ zwischen Mengen $\mathcal{X}$ und $\mathcal{Y}$ mit endlich vielen Elementen ist zwingend erforderlich, dass $\mathcal{Y}$ *mindestens so viele, höchstens so viele beziehungsweise genau so viele Elemente* enthält wie $\mathcal{X}$.

Wie später noch Thema sein wird, kann man auch die „Größe“ von Mengen mit unendlich vielen Elementen zu einem gewissen Grad über Injektionen, Surjektionen, Bijektionen vergleichen.

Aus einer beliebigen Abbildung $f\colon \mathcal{X} \to \mathcal{Y}$ kann man durch Verkleinerung des Ziels stets die Surjektion $\widetilde{f}\colon \mathcal{X} \to \mathrm{Bild}(f)$ mit $\widetilde{f}(x) := f(x)$ für alle $x \in \mathcal{X}$ gewinnen. Ist f surjektiv, so ist dabei $\widetilde{f} = f$. Ist f injektiv, so ist $\widetilde{f}$ sogar bijektiv.

Bei einer *Bijektion* $f\colon \mathcal{X} \to \mathcal{Y}$ erfolgt eine *Eins-zu-eins-Zuordnung* (der Elemente) von $\mathcal{X}$ und $\mathcal{Y}$, bei einer Injektion entsprechend eine Eins-zu-eins-Zuordnung (der Elemente) von $\mathcal{X}$ und Bild(f). Manchmal nutzt man dies, um die einander zugeordneten Mengen und Elemente vollständig miteinander zu identifizieren.

Beispiele II.1.9.

- Die Abbildungen $f\colon \mathbb{N} \to \mathbb{N}$, $f(n) = n + 1$ ist injektiv, aber nicht surjektiv, weil die 1 nicht im Bild liegt.
- Die Abbildung $g\colon \{0, 1\} \to \{0\}$ mit $g(0) = g(1) = 0$ ist surjektiv, aber nicht injektiv.
- Die Abbildung $\sigma\colon \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}$ mit der Wertetabelle

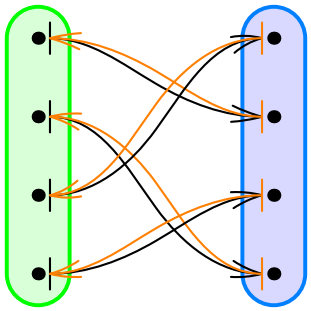
x	1	2	3	4
$\sigma(x)$	3	4	1	2

  ist eine Bijektion.

Die Abbildungseigenschaften werden von nun immer wieder auftreten und häufig eine wichtige Rolle spielen. Eng verbunden mit Bijektivität ist folgender Begriff:

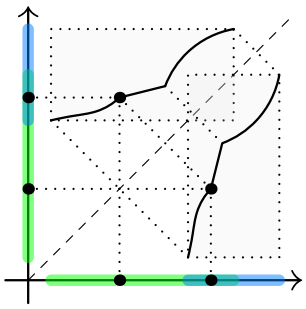
Definition II.1.10. Es seien $\mathcal{X}$ und $\mathcal{Y}$ Mengen sowie $f\colon \mathcal{X} \to \mathcal{Y}$ eine Abbildung. Eine Abbildung $g\colon \mathcal{Y} \to \mathcal{X}$ heißt *Umkehrfunktion* oder *Umkehrabbildung* von f, wenn folgende Äquivalenz für alle $x \in \mathcal{X}$ und $y \in \mathcal{Y}$ gilt:

$$f(x) = y \quad \Longleftrightarrow \quad g(y) = x.$$



(II.1.5)

Prinzipiell bedeutet dies nichts anderes, als dass die Umkehrfunktion aus der Funktion durch Umkehrung aller Zuordnungspfeile entsteht. Diese für das Konzept entscheidende Anschauung wird in einem einfachen Fall durch die Abbildung II.1.5 verdeutlicht. Bei einer Funktion $f\colon \mathcal{X} \to \mathcal{Y}$ von $\mathcal{X} \subset \mathbb{R}$ nach $\mathcal{Y} \subset \mathbb{R}$ geht der Graph $\mathrm{G}_g = \{(y, x) \in \mathbb{R}^2 \mid (x, y) \in \mathrm{G}_f\}$ der Umkehrfunktion $g\colon \mathcal{Y} \to \mathcal{X}$ wie in Abbildung II.1.6 aus dem Graph G_f von f durch Spiegelung an der ersten Winkelhalbierenden des Koordinatensystems hervor.



(II.1.6)

Was ist die Umkehrfunktion der Abbildung σ aus den Beispielen II.1.9?

Lemma II.1.11. *Es seien $\mathcal{X}$ und $\mathcal{Y}$ Mengen. Falls überhaupt eine Umkehrfunktion zu $f\colon \mathcal{X} \to \mathcal{Y}$ existiert, ist diese eindeutig bestimmt.*

BEWEIS. Für zwei Umkehrfunktionen $g\colon \mathcal{Y} \to \mathcal{X}$ und $\widetilde{g}\colon \mathcal{Y} \to \mathcal{X}$ zu f ergibt die Definition

$$g(y) = x \iff f(x) = y \iff \widetilde{g}(y) = x$$

für alle $x \in \mathcal{X}$, $y \in \mathcal{Y}$. Für $y \in \mathcal{Y}$ führt die Wahl $x := g(y) \in \mathcal{X}$ zur Äquivalenz $g(y) = g(y) \iff \widetilde{g}(y) = g(y)$. Da die linke Seite trivial gilt, gilt auch $\widetilde{g}(y) = g(y)$ für alle $y \in \mathcal{Y}$, und damit ist $\widetilde{g} = g$. □

Bemerkung II.1.12. Wegen der Eindeutigkeit bezeichnen wir die Umkehrfunktion zu f mit $f^{-1}\colon \mathcal{Y} \to \mathcal{X}$.

Die Notationen für Umkehrfunktion und Urbild sind insoweit konsistent, dass bei Existenz der Umkehrfunktion f^{-1} die Gleichheit $f^{-1}(\{y\}) = \{f^{-1}(y)\}$ (mit Urbild links und Umkehrfunktion rechts) für alle $y \in \mathcal{Y}$ gilt.

Satz II.1.13. *Es seien $\mathcal{X}$ und $\mathcal{Y}$ Mengen und $f\colon \mathcal{X} \to \mathcal{Y}$ eine Abbildung. Dann gilt:*

$$f \text{ ist bijektiv.} \qquad \iff \qquad \text{Es gibt eine Umkehrabbildung zu } f.$$

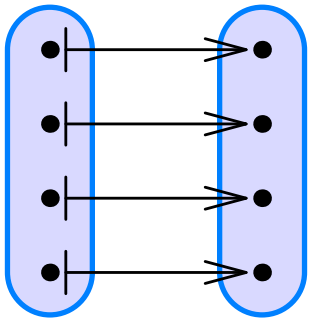
BEWEIS. „$\Longrightarrow$“: Sei f bijektiv und sei $y \in \mathcal{Y}$ beliebig. Zu y gibt es wegen der Bijektivität von f ein eindeutiges $x \in \mathcal{X}$ mit $f(x) = y$. Wir definieren eine Abbildung $g\colon \mathcal{Y} \to \mathcal{X}$ durch die Festlegung $g(y) := x$ für beliebiges $y \in \mathcal{Y}$ und das erwähnte eindeutig zugehörige x mit $f(x) = y$. Damit gilt offensichtlich die Äquivalenz $f(x) = y \iff g(y) = x$ für alle $x \in \mathcal{X}$ und $y \in \mathcal{Y}$, also ist g eine Umkehrabbildung zu f.

„$\Longleftarrow$“: Sei $g\colon \mathcal{Y} \to \mathcal{X}$ eine Umkehrabbildung zu f und sei $y \in \mathcal{Y}$ beliebig. Dann ist $g(y) =: x \in \mathcal{X}$ und das ist äquivalent zu $f(x) = y$. Damit ist f surjektiv. Nehmen wir an, dass $x_1 \neq x_2 \in \mathcal{X}$ gilt, aber $f(x_1) = y = f(x_2)$ ist. Dann wäre $g(y)$ sowohl gleich x_1 als auch gleich x_2 und somit wäre g keine Abbildung. □

Einige spezielle Abbildungen existieren für beliebige Mengen und können auch als weitere Beispiele von Abbildungen angesehen werden:

Beispiele II.1.14.

(a) Für jede Menge $\mathcal{X}$ ist die *Identität* oder *identische Abbildung* $\mathrm{id}_{\mathcal{X}}\colon \mathcal{X} \to \mathcal{X}$ von $\mathcal{X}$ durch $\mathrm{id}_{\mathcal{X}}(x) := x$ für alle $x \in \mathcal{X}$ gegeben. Diese Abbildung nimmt die Zuordnung $x \mapsto x$ für alle $x \in \mathcal{X}$ vor, ordnet also — wie in Abbildung II.1.7 — jedem Element von $\mathcal{X}$ wieder das Element selbst zu. Die Identität ist stets bijektiv.



(II.1.7)

(b) Für Mengen $\mathcal{X}$, $\mathcal{Y}$ und ein fixiertes Element $y \in \mathcal{Y}$ heißt $\mathcal{X} \to \mathcal{Y}$, $x \mapsto y$ die *konstante Abbildung* von $\mathcal{X}$ nach $\mathcal{Y}$ mit konstantem (Funktions-)Wert y.

Ist $f\colon \mathcal{X} \to \mathcal{Y}$ die konstante Abbildung mit Wert y, so schreiben wir das als $f \equiv y$ („f ist konstant gleich y“). Das normale Gleichheitszeichen wird an dieser Stelle bewusst vermieden, da die Abbildung f und das Element y des Zielbereichs verschiedene Objekte sind. Alternativ kann man $f(x) = y$ für alle $x \in \mathcal{X}$ ausschreiben und auf die Verwendung von $\equiv$ verzichten.

(c) Für eine *fixierte* Teilmenge $A \subset \mathcal{X}$ einer Menge $\mathcal{X}$ wird die *charakteristische Funktion* oder *Indikatorfunktion* $1_A\colon \mathcal{X} \to \{0, 1\}$ der Menge A durch

$$1_A(x) := \begin{cases} 1, & \text{falls } x \in A, \\ 0, & \text{falls } x \notin A \end{cases}$$

für alle $x \in \mathcal{X}$ definiert.

Etwas anders betrachtet kann man auch für ein *fixiertes* Element $x \in \mathcal{X}$ einer Menge $\mathcal{X}$ eine Abbildung $\delta_x \colon \mathcal{P}(\mathcal{X}) \to \{0,1\}$ durch $\delta_x(A) := 1_A(x)$ für alle $A \in \mathcal{P}(\mathcal{X})$ erhalten.

Wir kommen nun zu weiteren Grundoperationen mit Abbildungen:

Definition II.1.15. Es seien $\mathcal{X}$, $\mathcal{Y}$ und $\mathcal{Z}$ Mengen sowie $f \colon \mathcal{X} \to \mathcal{Y}$ und $g \colon \mathcal{Y} \to \mathcal{Z}$ Abbildungen. Dann ist die *Komposition* (auch *Verkettung* oder *Hintereinanderausführung* genannt) von f und g die Abbildung

$$g \circ f \colon \mathcal{X} \to \mathcal{Z}$$

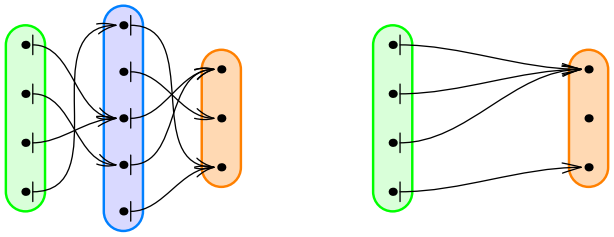
mit

$$(g \circ f)(x) := g(f(x)) \text{ für alle } x \in \mathcal{X}.$$

Beispiel II.1.16. Für $f \colon \mathbb{N} \to \mathbb{Z}$ mit $f(n) := 7-3n$ für $n \in \mathbb{N}$ und $g \colon \mathbb{Z} \to \mathbb{Q}$ mit $g(z) := z^2-3z+4^z$ für $z \in \mathbb{Z}$ ist $g \circ f \colon \mathbb{N} \to \mathbb{Q}$ durch $(g \circ f)(n) = (7-3n)^2 - 3(7-3n) + 4^{7-3n}$ für $n \in \mathbb{N}$ gegeben.

Anschaulich bedeutet die Komposition, dass man erst den Zuordnungen von f, dann denen von g folgt und auf diese Weise die Zuordnungen von $g \circ f$ erhält; dazu vergleiche Abbildung II.1.8.

(II.1.8)



Sowohl anschaulich einleuchtend als auch problemlos zu beweisen ist dann:

Satz II.1.17. *Für Mengen $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ sowie Abbildungen $f \colon \mathcal{W} \to \mathcal{X}$, $g \colon \mathcal{X} \to \mathcal{Y}$ und $h \colon \mathcal{Y} \to \mathcal{Z}$ gilt*

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Beweis. Für $w \in \mathcal{W}$ ergibt sich mit der Definition der Komposition

$$((h \circ g) \circ f)(w) = (h \circ g)(f(w)) = h(g(f(w))) = h((g \circ f)(w)) = (h \circ (g \circ f))(w). \qquad \square$$

Bemerkung II.1.18. Die Frage, ob die Komposition kommutativ ist, ob also $g \circ f$ gleich $f \circ g$ ist, macht nur für Selbstabbildungen f, g einer Menge $\mathcal{X}$ Sinn. Die Antwort ist aber auch in diesem Fall im Allgemeinen „Nein!“. Zum Beispiel gilt für die konstanten Abbildungen $f \colon \{0,1\} \to \{0,1\}$ mit $f \equiv 0$ und $g \colon \{0,1\} \to \{0,1\}$ mit $g \equiv 1$ offensichtlich $0 \equiv f \circ g \neq g \circ f \equiv 1$.

Übrigens liegt selbst für bijektive $f, g \colon \mathcal{X} \to \mathcal{X}$ im Allgemeinen keine Kommutativität vor: Zum Beispiel bei $f, g \colon \{1,2,3\} \to \{1,2,3\}$ mit $f(1) := 2$, $f(2) := 1$, $f(3) := 3$ und $g(1) := 1$, $g(2) := 3$, $g(3) := 2$ ist $3 = (g \circ f)(1) \neq (f \circ g)(1) = 2$.

Satz II.1.19. *Es seien $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ Mengen sowie $f \colon \mathcal{X} \to \mathcal{Y}$ und $g \colon \mathcal{Y} \to \mathcal{Z}$ Abbildungen. Dann gilt:*

Sind f, g beide injektiv $\implies$ $g \circ f$ ist injektiv.	*Ist $g \circ f$ injektiv $\implies$ f ist injektiv.*
Sind f, g beide surjektiv $\implies$ $g \circ f$ ist surjektiv.	*Ist $g \circ f$ surjektiv $\implies$ g ist surjektiv.*
Sind f, g beide bijektiv $\implies$ $g \circ f$ ist bijektiv.	*Ist $g \circ f$ bijektiv $\implies$ f ist injektiv und g ist surjektiv.*

Beweis. Die ersten beiden Zeilen werden in den Lernwerkstätten und Übungen behandelt. Die dritte folgt direkt daraus. $\square$

Definition II.1.20. Es seien $\mathcal{X}, \mathcal{Y}$ Mengen sowie $f \colon \mathcal{X} \to \mathcal{Y}$ und $g \colon \mathcal{Y} \to \mathcal{X}$ Abbildungen. Dann heißen f und g *zueinander invers* und g heißt *inverse Abbildung*, *inverse Funktion* oder kurz das *Inverse* von f, wenn gilt:

$$g \circ f = \mathrm{id}_{\mathcal{X}} \text{ und } f \circ g = \mathrm{id}_{\mathcal{Y}}.$$

Lemma II.1.21. *Es seien $\mathcal{X}, \mathcal{Y}$ Mengen. Falls überhaupt ein Inverses zu $f \colon \mathcal{X} \to \mathcal{Y}$ existiert, ist dieses eindeutig bestimmt.*

BEWEIS. Sind $g\colon \mathcal{Y} \to \mathcal{X}$ und $\widetilde{g}\colon \mathcal{Y} \to \mathcal{X}$ zwei Inverse zu f, so folgt

$$g = g \circ \mathrm{id}_{\mathcal{Y}} = g \circ (f \circ \widetilde{g}) = (g \circ f) \circ \widetilde{g} = \mathrm{id}_{\mathcal{X}} \circ \widetilde{g} = \widetilde{g}. \qquad \square$$

Tatsächlich ist das Inverse nichts anderes als die Umkehrfunktion:

Satz II.1.22. *Es seien $\mathcal{X}, \mathcal{Y}$ Mengen. Genau dann ist $g\colon \mathcal{Y} \to \mathcal{X}$ die Umkehrfunktion von $f\colon \mathcal{X} \to \mathcal{Y}$, wenn g das Inverse zu f ist.*

BEWEIS. „$\Longrightarrow$“: Ist g die Umkehrabbildung zu f, so ergibt Einsetzen von $y = f(x)$ beziehungsweise $x = g(y)$ in der Definition, dass $g(f(x)) = x$ für alle $x \in \mathcal{X}$ und $f(g(y)) = y$ für alle $y \in \mathcal{Y}$ gelten. Damit ist g auch das Inverse zu f.

„$\Longleftarrow$“: Ist g das Inverse zu f, so sind $g \circ f = \mathrm{id}_{\mathcal{X}}$ und $f \circ g = \mathrm{id}_{Y}$ beide bijektiv. Nach dem vorigen Satz ist dann f injektiv und surjektiv, also auch bijektiv. Damit existiert die Umkehrfunktion f^{-1} von f. Nach „$\Longrightarrow$“ ist f^{-1} invers zu f. Mit der Eindeutigkeit des Inversen folgt $g = f^{-1}$, also ist g die Umkehrfunktion zu f. $\square$

Insbesondere sind *Bijektivität, Umkehrbarkeit* (d.h. die Umkehrfunktion existiert) *und Invertierbarkeit* (d.h. das Inverse existiert) einer Abbildung alle drei *exakt gleichbedeutend.*

Zum Abschluss des Abschnitts sammeln wir noch einige weitere Grunddefinitionen bei Abbildungen ein. Zum Beispiel kann man den Definitionsbereich immer (künstlich) verkleinern und bei Werten in einem kartesischen Produkt in die sogenannten Komponenten aufspalten:

Definition II.1.23. Es seien $\mathcal{X}, \mathcal{Y}$ Mengen. Die *Einschränkung* einer Abbildung $f\colon \mathcal{X} \to \mathcal{Y}$ auf eine Teilmenge A des Definitionsbereichs $\mathcal{X}$ ist die Abbildung

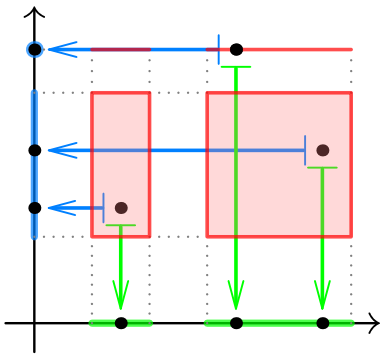
$$f|_A\colon A \to \mathcal{Y}$$

mit dem kleineren Definitionsbereich A, aber der auf diesem unveränderten Zuordnungsvorschrift $f|_A(x) := f(x)$ für alle $x \in A$.

Definition II.1.24. Es sei $n \in \mathbb{N}$, und $\mathcal{X}_1, \ldots, \mathcal{X}_n$ seien Mengen. Für $i \in \{1, 2, \ldots, n\}$ wird die *i-te Projektion*

$$\mathrm{p}_i\colon \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n \to \mathcal{X}_i$$

des n-fachen kartesischen Produkts $\mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n$ durch die Festlegung $\mathrm{p}_i(x_1, x_2, \ldots, x_n) := x_i$ für alle $(x_1, x_2, \ldots, x_n) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n$ definiert.



(II.1.9)

Auch allgemein nutzt man x_i mit $i \in \{1, 2, \ldots, n\}$ als Standard-Bezeichnung für den Eintrag $\mathrm{p}_i(x)$ des Tupels $x \in \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n$.

Bemerkungen II.1.25.

(a) Eine Funktion

$$f\colon \mathcal{X} \to \mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n$$

mit Werten im n-fachen kartesischen Produkt $\mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n$ entspricht eins-zu-eins einzelnen Funktionen $f_1\colon \mathcal{X} \to \mathcal{Y}_1$, $f_2\colon \mathcal{X} \to \mathcal{Y}_2$, …, $f_n\colon \mathcal{X} \to \mathcal{Y}_n$ mit

$$f(x) = (f_1(x), f_2(x), \ldots, f_n(x)) \text{ für alle } x \in \mathcal{X}.$$

Man nennt $f_1, f_2, \ldots, f_n$ die *Komponenten(-funktionen) von* f und kann jedes f_i mit $i \in \{1, 2, \ldots, n\}$ als $f_i = \mathrm{p}_i \circ f$ mit der i-ten Projektion p_i von $\mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n$ schreiben, woraus die Existenz und Eindeutigkeit von $f_1, f_2, \ldots, f_n$ zu gegebenem f klar wird.

In Analogie zu Elementen der Zielmenge schreibt man die Funktion mit Komponenten f_1,f_2, …, f_n als $(f_1, f_2, \ldots, f_n)$ und verwendet $f_1, f_2, \ldots, f_n$ als Standard-Bezeichnungen für die Komponenten einer Funktion f mit Werten in $\mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n$.

Besonders häufig arbeitet man im Fall $\mathcal{Y}_1 = \mathcal{Y}_2 = \ldots = \mathcal{Y}_n = \mathcal{Y}$, also für eine Funktion

$$f\colon \mathcal{X} \to \mathcal{Y}^n,$$

mit den n Komponentenfunktionen $f_1, f_2, \ldots, f_n\colon \mathcal{X} \to \mathcal{Y}$ der Funktion f.

(b) Die *Diagonalabbildung* $\mathcal{X} \to \mathcal{X}^n$, $x \mapsto (x, x, \ldots, x)$ kann auch als die Abbildung

$$(\mathrm{id}_\mathcal{X}, \mathrm{id}_\mathcal{X}, \ldots, \mathrm{id}_\mathcal{X})\colon \mathcal{X} \to \mathcal{X}^n$$

geschrieben werden, deren Komponenten alle $\mathrm{id}_\mathcal{X}$ sind. Sie ist stets injektiv.

(c) Das *kartesische Produkt von Abbildungen* $f_1\colon \mathcal{X}_1 \to \mathcal{Y}_1$, $f_2\colon \mathcal{X}_2 \to \mathcal{Y}_2$ ist die Abbildung

$$f_1 \times f_2\colon \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}_1 \times \mathcal{Y}_2,$$

die durch $(f_1 \times f_2)(x_1, x_2) := (f_1(x_1), f_2(x_2))$ für alle $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$ definiert wird. Die Komponenten von $f_1 \times f_2$ sind $\mathrm{p}_{\mathcal{Y}_1} \circ (f_1 \times f_2) = f_1 \circ \mathrm{p}_{\mathcal{X}_1}\colon \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}_1$ und $\mathrm{p}_{\mathcal{Y}_2} \circ (f_1 \times f_2) = f_2 \circ \mathrm{p}_{\mathcal{X}_2}\colon \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{Y}_2$ mit den Projektionen $\mathrm{p}_{\mathcal{X}_1}, \mathrm{p}_{\mathcal{X}_2}$ und $\mathrm{p}_{\mathcal{Y}_1}, \mathrm{p}_{\mathcal{Y}_2}$ der kartesischen Produkte $\mathcal{X}_1 \times \mathcal{X}_2$ und $\mathcal{Y}_1 \times \mathcal{Y}_2$. Die Produkt-Bildung $\times$ bei Abbildungen ist assoziativ und ist analog für eine beliebige endliche Zahl von Funktionen sinnvoll.

Als letztes besprechen wir noch kurz Mengen von Abbildungen:

Definition II.1.26. Sind $\mathcal{X}, \mathcal{Y}$ Mengen, so schreiben wir $\mathrm{Abb}(\mathcal{X}, \mathcal{Y})$ oder $\mathcal{Y}^\mathcal{X}$ für die *Menge der Abbildungen von* $\mathcal{X}$ *nach* $\mathcal{Y}$.

Bemerkungen II.1.27. Es sei $n \in \mathbb{N}$, und $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Y}_1$, $\mathcal{Y}_2$, …, $\mathcal{Y}_n$ seien Mengen.

- Für jeden Definitionsbereich $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ mit genau n verschiedenen Elementen x_1, x_2, …, x_n ist
$$\mathrm{Abb}(\mathcal{X}, \mathcal{Y}) \to \mathcal{Y}^n, \quad f \mapsto (f(x_1), f(x_2), \ldots, f(x_n))$$
eine Bijektion. Daher kann man eine Abbildung $f \in \mathrm{Abb}(\mathcal{X}, \mathcal{Y})$ auf dem n-elementigen Definitionsbereich $\mathcal{X}$ auch als n-Tupel $(f(x_1), f(x_2), \ldots, f(x_n)) \in \mathcal{Y}^n$ ihrer Werte betrachten und bekommt in diesem Fall eine naheliegende Identifikation von $\mathrm{Abb}(\mathcal{X}, \mathcal{Y})$ mit $\mathcal{Y}^n$. In Anlehnung hieran kann man sich eine Abbildung $f \in \mathrm{Abb}(\mathbb{N}, \mathcal{Y})$ mit Definitionsbereich $\mathbb{N}$ als unendliches Tupel $(f(x_1), f(x_2), f(x_3), \ldots)$ ihrer Werte vorstellen – ein Objekt, das wir so noch nicht definiert haben, das aber für $\mathrm{Abb}(\mathbb{N}, \mathcal{Y})$ die Schreibweise $\mathcal{Y}^\mathbb{N}$ nahelegt. Hierdurch wird für $\mathrm{Abb}(\mathcal{X}, \mathcal{Y})$ auch bei allgemeinen Definitionsbereich $\mathcal{X}$ die in der vorigen Definition eingeführte und an eine Potenz erinnernde Schreibweise $\mathcal{Y}^\mathcal{X}$ motiviert.
- Der oben erwähnten Eins-zu-Eins-Entsprechung zwischen Funktionen mit Werten im Produkt der Mengen Y_i, $\mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n$, und dem Tupel ihrer Komponentenfunktionen liegt tatsächlich eine Bijektion
$$\mathrm{Abb}(\mathcal{X}, \mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n) \to \mathrm{Abb}(\mathcal{X}, \mathcal{Y}_1) \times \mathrm{Abb}(\mathcal{X}, \mathcal{Y}_2) \times \ldots \times \mathrm{Abb}(\mathcal{X}, \mathcal{Y}_n)$$
oder mit anderen Worten (und vielleicht intuitiver)
$$(\mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n)^\mathcal{X} \to \mathcal{Y}_1^\mathcal{X} \times \mathcal{Y}_2^\mathcal{X} \times \ldots \times \mathcal{Y}_n^\mathcal{X}$$
zugrunde, die die Zuordnung $f \mapsto (f_1, f_2, \ldots, f_n)$ vornimmt (mit den Komponenten f_i von f). Dies ermöglicht eine häufig verwendete Identifikation, die wir teils in die Notation für die Komponenten eingebaut haben. Deutlicher noch sieht man dies an der zugehörigen Umkehrabbildung, die $(f_1, f_2, \ldots, f_n) \mapsto (f_1, f_2, \ldots, f_n)$ zuordnet, wobei das Tupel der Einzelfunktionen $f_1, f_2, \ldots, f_n$ (links) und die Funktion mit Komponenten $f_1, f_2, \ldots, f_n$ (rechts) in unserer Notation nicht mehr unterscheidbar sind.

II.2. Natürliche und ganze Zahlen, Induktion und Rekursion

Natürlich wurde die Menge $\mathbb{N} = \{1, 2, 3, \ldots\}$ der natürlichen Zahlen bereits verwendet und ist Ihnen geläufig. Dennoch möchte man $\mathbb{N}$ in der Mathematik präziser einführen, und prinzipiell kann man den Aufbau der Mathematik dann auch so gestalten, dass $\mathbb{N}$ vor seiner Einführung nicht vorkommt. Die Einhaltung einer derartigen konsequenten Reihenfolge wäre aber an vielen Stellen so umständlich, dass dies für eine Vorlesung nicht in Frage kommt. Die präzise Einführung von $\mathbb{N}$ mag zunächst penibel und überflüssig scheinen, wird aber den richtigen Rahmen für das *wichtige Beweisverfahren der vollständigen Induktion*, für *rekursive Definitionen* und die Eingrenzung der entscheidenden Eigenschaften von $\mathbb{N}$ bieten. Tatsächlich fordern wir zur Einführung von $\mathbb{N}$ folgendes Axiom:

Axiom II.2.1 (Peano-Axiome der natürlichen Zahlen)**.** Es gibt eine Menge $\mathbb{N}$, eine injektive Abbildung $S\colon \mathbb{N} \to \mathbb{N}$ und ein Element $1 \in \mathbb{N} \setminus S(\mathbb{N})$, so dass für alle Teilmengen M von $\mathbb{N}$ gilt: Ist $1 \in M$ und $S(M) \subset M$, so folgt $M = \mathbb{N}$.

Es handelt sich hierbei um eine kurze und prägnante Zusammenfassung eines fünf separate Axiome umfassenden Axiomensystems, das auf G. Peano (1858–1932) zurückgeht. Wir diskutieren die Original-Axiome und erläutern damit zugleich das Vorausgehende:

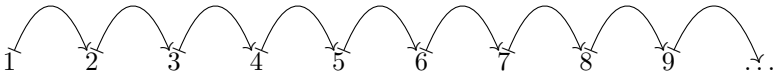
- Die erste Forderung $1 \in \mathbb{N}$ sichert überhaupt die *Existenz einer natürlichen Zahl* 1.
- Das zweite Axiom postuliert für jede natürliche Zahl $n \in \mathbb{N}$ die *Existenz und Eindeutigkeit eines Nachfolgers* $S(n) \in \mathbb{N}$, den wir normalerweise als $n{+}1$ bezeichnen. Dieses Axiom ist oben in der Wohldefiniertheit der Abbildung $S\colon \mathbb{N} \to \mathbb{N}$ enthalten.
- Das dritte Axiom $1 \notin S(\mathbb{N})$ besagt, dass 1 nicht Nachfolger einer natürlichen Zahl ist und begründet die herausgehobene Rolle der 1 als erste natürliche Zahl.
- Peanos viertes Axiom fordert $S(m) = S(n) \Longrightarrow m = n$ für alle $m, n \in \mathbb{N}$, entspricht oben der Injektivität von S und sichert die Eindeutigkeit des Vorgängers jeder natürlichen Zahl (sofern es einen Vorgänger gibt).
- Das letzte Axiom schließlich enthält die oben gestellte Forderung für Teilmengen M von $\mathbb{N}$ und kann in Worten so formuliert werden, dass aus $1 \in M$ und der Gültigkeit der Implikation $n \in M \Longrightarrow S(n) \in M$ für alle $n \in \mathbb{N}$ schon $M = \mathbb{N}$ folgt. Dieses Postulat ist als *Induktionsaxiom* bekannt und wird in Folge noch sehr genau diskutiert.

Natürlich beschreiben diese Axiome nichts anderes als die gewohnte Struktur der natürlichen Zahlen, wobei die *Sukzessions- oder Nachfolge-Abbildung* S, die in Abbildung II.2.1 veranschaulicht wird, einfach dem „Weiterzählen“ von einer natürlichen Zahl zur nächsten entspricht und die formale Einführung der nachfolgenden Zahlen durch

$$2 := S(1), \quad 3 := S(2), \quad 4 := S(3), \quad 5 := S(4), \quad 6 := S(5), \quad 7 := S(6), \quad 8 := S(7), \quad 9 := S(8), \ldots$$

ermöglicht.

(II.2.1)



Auch wenn all dies mehr oder weniger vertraut scheinen mag, sei an dieser Stelle trotzdem die Warnung angebracht, dass bei $\mathbb{N}$ *als Menge mit unendlich vielen Elementen* ganz anderes passieren kann als bei Mengen mit nur endlich vielen Elementen. So erhält man aus der Injektion $S\colon \mathbb{N} \to \mathbb{N}$ zum einen die Bijektion $\widetilde{S}\colon \mathbb{N} \to S(\mathbb{N})$, durch die man die Elemente von $\mathbb{N}$ und $S(\mathbb{N})$ eins-zu-eins identifizieren kann. Zum anderen ist aber doch $S(\mathbb{N}) \subsetneqq \mathbb{N}$, denn $\mathbb{N}$ enthält das Element 1, aber $1 \notin S(\mathbb{N})$.

Genau diese paradox scheinende Eigenschaft des Unendlichen illustriert das berühmte Gedankenspiel in *Hilberts Hotel*: In diesem Hotel gibt es unendlich viele Zimmer, für jede natürliche Zahl als Raumnummer eines. Sind alle Zimmer belegt, so ist dieses besondere Hotel aber dennoch nicht voll. Kommt nämlich in dieser Situation ein neuer Gast an, so bittet man alle bereits eingezogenen Gäste, aus ihrem derzeitigem Zimmer n in das Nachfolge-Zimmer $S(n)$ zu ziehen, sozusagen entlang der Umzugswege in Abbildung II.2.1. Damit wird Zimmer 1 frei und kann durch den neuen Gast bezogen werden.

Weitere Eigenschaften von $\mathbb{N}$ (und S) können wir aus den Axiomen folgern. Zum Beispiel sind bei $\{n \in \mathbb{N} \mid S(n) \neq n\}$ gemäß dem dritten und vierten Axiom die Voraussetzungen des Induktionsaxioms

erfüllt. Letzteres zeigt dann, dass die angegebene Menge ganz $\mathbb{N}$ ist und sich daher jedes $n \in \mathbb{N}$ von seinem Nachfolger unterscheidet. In ähnlicher Weise gibt die Anwendung des Induktionsaxioms auf $S(\mathbb{N}) \cup \{1\}$ für jede Zahl $n \in \mathbb{N}$ außer 1 die Existenz eines Vorgängers $p \in \mathbb{N}$ mit $S(p) = n$, der gemäß dem vierten Axiom außerdem eindeutig ist.

Schon hieran kann man die Wichtigkeit des Induktionsaxioms erkennen, auf das wir demnächst noch genau eingehen und ohne das man solche einleuchtenden Folgerungen tatsächlich nicht[4] zur Verfügung hätte.

Ist $\mathbb{N}$ einmal eingeführt, so ist der Aufwand zur Definition von $\mathbb{N}_0$ und $\mathbb{Z}$ eher gering:

Definition II.2.2. Wir setzen

$$\mathbb{N}_0 := \mathbb{N} \sqcup \{0\}$$

mit einer fixierten Zahl $0 \notin \mathbb{N}$ und

$$\mathbb{Z} := \mathbb{N}_0 \sqcup (-\mathbb{N})$$

mit einer Kopie [5] $-\mathbb{N}$ von $\mathbb{N}$, so dass die Abbildung $\mathbb{N} \to -\mathbb{N}$, die jedem $n \in \mathbb{N}$ ein neues Element $-n \in (-\mathbb{N}) \setminus \mathbb{N}_0$ zuordnet, bijektiv ist. Wir nennen die Zahlen aus $\mathbb{N}$ auch die positiven ganzen Zahlen, die aus $-\mathbb{N}$ die negativen ganzen Zahlen.

Das Negative $-z \in \mathbb{Z}$ einer ganzen Zahl $z \in \mathbb{Z}$ definieren wir ergänzend, nachdem es für positive z schon eingeführt ist, durch $-0 := 0$ und $-(-n) := n$ für $n \in \mathbb{N}$. Die Nachfolge-Abbildung setzen wir durch die Festlegungen $S(0) := 1$, $S(-1) := 0$ und $S(-S(n)) := -n$ für $n \in \mathbb{N}$, natürlich unter Beibehaltung von $S(n)$ für $n \in \mathbb{N}$, zu einer bijektiven Abbildung $S\colon \mathbb{Z} \to \mathbb{Z}$ fort.

Wir halten fest, dass für jede Zahl $z \in \mathbb{Z}$ ein eindeutiger Nachfolger $S(z) \in \mathbb{Z}$ und ein eindeutiger Vorgänger $S^{-1}(z) \in \mathbb{Z}$ (mit der Umkehrabbildung $S^{-1}\colon \mathbb{Z} \to \mathbb{Z}$ der Bijektion $S\colon \mathbb{Z} \to \mathbb{Z}$) existieren und $S^{-1}(z) \neq z \neq S(z)$ erfüllen. (Letzteres folgt leicht aus $S(n) \neq n$ für alle $n \in \mathbb{N}$).

Bemerkung II.2.3. Übrigens erfüllt anstelle von $(\mathbb{N}, 1)$ auch $(\mathbb{N}_0, 0)$ die Peano-Axiome mit $S\colon \mathbb{N}_0 \to \mathbb{N}_0$ erweitert durch $S(0) := 1$. Insofern gibt es zwischen $\mathbb{N}$ und $\mathbb{N}_0$ bisher, wo wir nur die Nachfolge-Abbildung S und ein Anfangselement 1 bzw. 0 betrachten, keinen strukturellen Unterschied, und wir hätten alternativ auch $\mathbb{N}_0$ axiomatisch einführen und $\mathbb{N} := \mathbb{N}_0 \setminus \{0\}$ setzen können. Sobald wir die Grundrechenarten angehen, entstehen aber selbstverständlich strukturelle Unterschiede zwischen 1 und 0 und somit zwischen $\mathbb{N}$ und $\mathbb{N}_0$.

Auf dem Induktionsaxiom fußt das *fundamentale Beweisprinzip der vollständigen Induktion (VI) zum Nachweis einer von $n \in \mathbb{N}$ abhängigen Aussage*[6] $A(n)$ für alle $n \in \mathbb{N}$, bei dem man wie folgt vorgeht:

(VI)
$$\left.\begin{array}{l} \text{Induktionsanfang: Zeige } A(1). \\ \text{Induktionsschritt: Zeige die Implikation } A(n) \Longrightarrow A(n{+}1) \text{ für alle } n \in \mathbb{N}. \end{array}\right\}$$

Dabei benutzen wir $n{+}1 := S(n)$ als die allgemein übliche Bezeichnung für die auf n folgende natürliche Zahl.

Bemerkungen II.2.4.

- Der Induktionsanfang wird manchmal auch Induktionsverankerung genannt. Er lässt sich oft vergleichsweise einfach erledigen, darf aber natürlich nicht vergessen werden.
- Für den Induktionsschritt fixiert man ein beliebiges $n \in \mathbb{N}$. Man betrachtet die *Induktionsannahme* oder Induktionsvoraussetzung $A(n)$ als gegeben und leitet dann unter Verwendung dieser Annahme die *Induktionsbehauptung* $A(n{+}1)$ her.

---

[4] Alle Peano-Axiome außer dem Induktionsaxiom sind auch für $\mathbb{N}$ mit $S(n) := n{+}2$ für alle ungeraden $n \in \mathbb{N}$ und $S(n) := n$ für alle geraden $n \in \mathbb{N}$ erfüllt. In diesem Modell wären aber anders als üblich alle geraden Zahlen ihr eigener Nachfolger. Zudem sind die Peano-Axiome außer dem Induktionsaxiom auch für die positiven reellen Zahlen $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x > 0\}$ anstelle von $\mathbb{N}$ mit $S(x) := x{+}1$ für alle $x \in \mathbb{R}_+$ erfüllt und ebenso für $\mathbb{N}$ mit $S(n) := n{+}2$ für alle $n \in \mathbb{N}$. Dabei gäbe es neben 1 aber weitere Zahlen, die keinen Vorgänger besitzen, zum Beispiel $\frac{1}{2}$ im ersten und 2 im zweiten Fall.

[5] Die etwas vage Einführung von $-\mathbb{N}$ als Kopie kann mengentheoretisch untermauert werden, indem man nur von einem neuen Vorzeichenobjekt $\ominus$ ausgeht, $-\mathbb{N} := \{\ominus\}\times\mathbb{N}$ setzt und dann für $n \in \mathbb{N}$ die Notation $-n := (\ominus, n) \in -\mathbb{N}$ festlegt.

[6] Genauer ist $A(n)$ ein Prädikat mit freier Variable n, für die natürliche Zahlen eingesetzt werden können, und das Ziel ist der Nachweis der Aussage $\forall n{\in}\mathbb{N}\colon A(n)$.

- Dass das Beweisverfahren funktioniert und tatsächlich die Gültigkeit von $A(n)$ für alle $n \in \mathbb{N}$ sicherstellt, ist durch das Induktionsaxiom gerechtfertigt: Sind Induktionsanfang und Induktionsschritt gemacht, so wissen wir für die Menge $M := \{n \in \mathbb{N} \mid A(n) \text{ gilt}\}$ einerseits $1 \in M$ und andererseits $n \in M \implies S(n) \in M$. Wir können daher per Induktionsaxiom schließen, dass $M = \mathbb{N}$ ist. Nach Wahl von M bedeutet dies, dass $A(n)$ für alle $n \in \mathbb{N}$ gilt.

  (Diese Argumentation braucht jetzt, nachdem wir sie einmal gemacht haben, natürlich *nicht* bei jedem Induktionsbeweis wiederholt zu werden.)
- Induktionsbeweise sind nicht konstruktiv, da man schon vor Beginn des Beweises wissen muss, wie die zu zeigende Aussage $A(n)$ für alle $n \in \mathbb{N}$ aussieht. Um ein Gefühl zu bekommen und eine sinnvolle Vermutung $A(n)$ aufstellen zu können, experimentiert man typischerweise mit den ersten paar natürlichen Zahlen. Ein erfolgreicher Induktionsbeweis zementiert dann, dass eine Aussage $A(n)$, die man für die ersten paar natürlichen Zahlen n verifiziert hat, tatsächlich und unwiderruflich für *alle* natürlichen Zahlen n gilt. Mit etwas mehr Erfahrung kann man manchmal auch mit möglichen Induktionsschritten experimentieren, um sinnvoll eine Vermutung $A(n)$ aufstellen zu können. Wie man zur Vermutung findet, ist tatsächlich nicht vorgeschrieben und muss auch nicht unbedingt dokumentiert werden.
- Man muss *nicht jede* von $n \in \mathbb{N}$ abhängige Aussage $A(n)$ mit vollständiger Induktion beweisen. Manchmal kann man solche Aussagen $A(n)$ auch für jedes $n \in \mathbb{N}$ ohne Verwendung der vorigen Aussagen $A(1), A(2), \dots, A(n-1)$ zeigen und braucht dann keinen Induktionsbeweis anzusetzen. Insbesondere ist dies der Fall, wenn man beim Versuch eines Induktionsschritts feststellt, dass man die Induktionsbehauptung ohne Verwendung der Induktionsannahme zeigen kann. Auch wenn man formal nichts falsch gemacht hat, empfiehlt es sich in dieser Situation trotzdem, das Argument noch einmal neu und induktionsfrei aufzuschreiben.

Beispiele II.2.5. Unter leichtem Vorgriff auf noch einzuführende Konzepte und Rechenregeln geben wir Beispiele der Anwendung von (VI):

- Wir wollen zeigen:

  $$\text{Für alle } n \in \mathbb{N} \text{ ist } 3^{2n}-2^n \text{ durch 7 teilbar.}$$

  Wir argumentieren durch vollständige Induktion nach $n \in \mathbb{N}$.

  Induktionsanfang für $n = 1$: Die Behauptung reduziert sich wegen $3^{2\cdot 1}-2^1 = 7$ zu „7 ist durch 7 teilbar“ und ist offensichtlich richtig.

  Induktionsschritt von n auf $n+1$ für $n \in \mathbb{N}$: Wir möchten von der Induktionsannahme „$3^{2n}-2^n$ ist durch 7 teilbar“ auf die Induktionsbehauptung „$3^{2n+2}-2^{n+1}$ ist durch 7 teilbar“ schließen. Dazu schreiben wir

  $$3^{2n+2} - 2^{n+1} = 9\cdot 3^{2n} - 9\cdot 2^n + 9\cdot 2^n - 2\cdot 2^n = 9\cdot(3^{2n}-2^n) + 7\cdot 2^n,$$

  wobei einerseits nach Induktionsannahme $3^{2n}-2^n$ und damit auch $9\cdot(3^{2n}-2^n)$ durch 7 teilbar ist und andererseits $7\cdot 2^n$ wegen des Faktors 7 durch 7 teilbar ist. Mit den beiden Summanden ist auch die Summe $3^{2n+2}-2^{n+1}$ durch 7 teilbar und die Induktionsbehauptung gezeigt.
- Wir wollen zeigen, dass die sukzessive Addition der ungeraden Zahlen die Quadratzahlen ergibt, genauer

  $$1+3+5+\dots+(2n-3)+(2n-1) = n^2 \quad \text{für alle } n \in \mathbb{N}.$$

  Wir argumentieren durch vollständige Induktion nach $n \in \mathbb{N}$.

  Induktionsanfang für $n = 1$: Die Behauptung reduziert sich zu $1 = 1^2$ und gilt offensichtlich.

  Induktionsschritt von n auf $n+1$ für $n \in \mathbb{N}$: Wir möchten von der Induktionsannahme

  $$1+3+5+\dots+(2n-3)+(2n-1) = n^2$$

  auf die Induktionsbehauptung

  $$1+3+5+\dots+(2n-1)+(2n+1) = (n+1)^2$$

  schließen. Dazu rechnen wir unter Verwendung der Induktionsannahme im zweiten Schritt

  $$\begin{aligned} 1+3+5+\dots+(2n-1)+(2n+1) &= (1+3+5+\dots+(2n-3)+(2n-1)) + (2n+1) \\ &= n^2 + (2n+1) = n^2+2n+1 = (n+1)^2 \end{aligned}$$

und erhalten die Induktionsbehauptung.

In der praktischen Anwendung treten auch etliche *Varianten des Induktionsprinzips* auf: Zum Beispiel kann man eine Aussage $A(z)$ für alle $z \in \{z_0, z_0{+}1, z_0{+}2, \ldots\}$ mit fixiertem $z_0 \in \mathbb{Z}$ (häufig auch $z_0 = 0$) nachweisen, indem man den Induktionsanfang bei $A(z_0)$ und Induktionsschritt für alle $z \in \{z_0, z_0{+}1, z_0{+}2, \ldots\}$ von $A(z)$ zu $A(z{+}1)$ durchführt. Man kann die Variable beim Induktionsschritt selbstverständlich anders benennen und den Schritt etwa von $A(z{-}1)$ zu $A(z)$ für alle $z \in \{z_0{+}1, z_0{+}2, z_0{+}3, \ldots\}$ durchführen. Man kann mehrere Einzelfälle als Induktionsanfang behandeln oder beim Induktionsschritt in Zweierschritten von $A(z)$ zu $A(z{+}2)$ übergehen (nützlich etwa dann, wenn für gerade und ungerade z unterschiedliches Vorgehen geboten ist). Explizit erwähnen wir folgende *Variante*, die ebenfalls den Nachweis von $A(n)$ für alle $n \in \mathbb{N}$ erlaubt:

$$(\widetilde{\text{VI}}) \qquad \left.\begin{array}{l} \textit{Induktionsanfang:} \text{ Zeige } A(1). \\ \textit{Induktionsschritt:} \text{ Zeige } A(1) \wedge A(2) \wedge \ldots \wedge A(n) \Longrightarrow A(n{+}1) \text{ für alle } n \in \mathbb{N}. \end{array}\right\}$$

Da man hier mehr Annahmen zur Verfügung hat, bietet ($\widetilde{\text{VI}}$) *beim Induktionsschritt mehr Flexibilität* als (VI). Tatsächlich sind die beiden Prinzipien aber gleichwertig, denn ($\widetilde{\text{VI}}$) kann durch Anwendung des Original-Prinzips (VI) auf $\widetilde{A}(n) := A(1) \wedge A(2) \wedge \ldots \wedge A(n)$ aus diesem abgeleitet werden.

Entscheidend ist *bei allen Varianten* des Induktionsprinzip aber vor allem, dass *durch Induktionsanfang und -schritt ein Dominoeffekt entsteht*, dass etwa durch $A(z_0)$, den Schritt von $A(z_0)$ zu $A(z_0{+}1)$, den Schritt von $A(z_0{+}1)$ zu $A(z_0{+}2)$, den Schritt von $A(z_0{+}2)$ zu $A(z_0{+}3)$, ... *alle* gewünschten $A(z)$ abgedeckt werden. Statt verschiedene Varianten auswendig zu lernen, sollte man tatsächlich besser dieses Zusammenspiel von Induktionsanfang und Induktionsschritt im Hinterkopf behalten und sich überlegen, dass es funktioniert.

Beispiel II.2.6. Wir möchten zeigen:

Für alle $n \in \mathbb{N}_0$ und alle Mengen M mit genau n Elementen hat $\mathcal{P}(M)$ genau 2^n Elemente.

Dazu argumentieren wir durch vollständige Induktion nach $n \in \mathbb{N}_0$.

Induktionsanfang für $n = 0$: Die einzige Menge mit 0 Elementen ist die leere Menge $\emptyset$. Für diese hat $\mathcal{P}(\emptyset) = \{\emptyset\}$ genau $2^0 = 1$ Elemente (nämlich das eine Element $\emptyset$).

Induktionsschritt von $n{-}1$ auf n für $n \in \mathbb{N}$: Wir haben als Induktionsannahme, dass $\mathcal{P}(M')$ für jede Menge M' mit genau $(n{-}1)$ Elementen genau 2^{n-1} Elemente hat. Wir möchten daraus für eine gegebene Menge M mit genau n Elementen schließen, dass $\mathcal{P}(M)$ genau 2^n Elemente hat. Sei dazu x irgendein ab jetzt fixiertes Element von M. Dann hat $M' := M \setminus \{x\}$ genau $(n{-}1)$ Elemente, und nach Induktionsannahme hat $\mathcal{P}(M')$ genau 2^{n-1} Elemente. Wir können nun die Elemente A von $\mathcal{P}(M)$, also die Teilmengen A von M, darauf betrachten, ob $x \notin A$ oder $x \in A$ gilt. Im Fall $x \notin A$ ist $A \subset M'$, also $A \in \mathcal{P}(M')$. Im Fall $x \in A$ ist $A = A' \sqcup \{x\}$ für die Menge $A' := A \setminus \{x\} \subset M'$, also für ein $A' \in \mathcal{P}(M')$. Insgesamt ist somit

$$\mathcal{P}(M) = \mathcal{P}(M') \sqcup \{A \in \mathcal{P}(M) \mid A = A' \sqcup \{x\} \text{ für ein } A' \in \mathcal{P}(M')\}$$

die disjunkte Vereinigung von $\mathcal{P}(M')$ und einer weiteren Menge, die für jedes Element A' von $\mathcal{P}(M')$ genau ein Element A enthält. (Beachte dafür $A' \neq B' \Longrightarrow A' \sqcup \{x\} \neq B' \sqcup \{x\}$ für $A', B' \in \mathcal{P}(M')$!) Damit hat $\mathcal{P}(M)$ genau doppelt so viele Elemente wie $\mathcal{P}(M')$, also wie behauptet $2{\cdot}2^{n-1} = 2^n$ Elemente.

Beispiel II.2.7. Wir möchten die *Existenz der Primfaktorzerlegung* zeigen:

Für alle $n \in \mathbb{N} \setminus \{1\}$ gibt es $k \in \mathbb{N}$ und Primzahlen $p_1, p_2, \ldots, p_k \in \mathbb{N}$ mit $n = p_1 p_2 {\cdot} \ldots {\cdot} p_k$.

Wir argumentieren durch vollständige Induktion nach $n \in \mathbb{N} \setminus \{1\}$.

Induktionsanfang für $n = 2$: Für $n = 2$ gilt die Behauptung mit $k = 1$ und der Primzahl 2.

Induktionsschritt von $2, 3, \ldots, n{-}1$ auf n für $n \in \mathbb{N} \setminus \{1, 2\}$: Wir haben als Induktionsannahme, dass die behauptete Darstellung für $1, 2, \ldots, n{-}1$ jeweils existiert. Wir möchten daraus die Existenz der entsprechenden Darstellung von n ableiten. Im Fall, dass n eine Primzahl ist, ist die Darstellung trivial (mit $k = 1$, $p_1 = n$). Im Fall, dass n keine Primzahl ist, können wir $n = ab$ mit $a, b \in \{2, 3, \ldots, n{-}1\}$ schreiben. Gemäß Induktionsannahme gibt es daher einerseits $k \in \mathbb{N}$ und Primzahlen $p_1, p_2, \ldots, p_k \in \mathbb{N}$ mit $a = p_1 p_2 {\cdot} \ldots {\cdot} p_k$, andererseits $\ell \in \mathbb{N}$ und Primzahlen $q_1, q_2, \ldots, q_\ell \in \mathbb{N}$ mit $b = q_1 q_2 {\cdot} \ldots {\cdot} q_\ell$. Setzen wir $p_{k+i} := q_i$ für alle $i \in \{1, 2, \ldots, \ell\}$, so ergibt sich mit

$$n = ab = (p_1 p_2 {\cdot} \ldots {\cdot} p_k)(q_1 q_2 {\cdot} \ldots {\cdot} q_\ell) = p_1 p_2 {\cdot} \ldots {\cdot} p_{k+\ell}$$

die behauptete Darstellung von n (mit $k+\ell \in \mathbb{N}$ anstelle von k). Wir erhalten also die Induktionsbehauptung.

Eine Version des Induktionsprinzips zur Erklärung eines von $n \in \mathbb{N}$ abhängigen Objekts X_n für alle $n \in \mathbb{N}$ ist das *Prinzip der rekursiven Definition*, das wir schematisch wie folgt festhalten:

(RD)
$$\left.\begin{array}{l} \textit{Rekursionsanfang:}\ \text{Definiere } X_1. \\ \textit{Rekursionsschritt:}\ \text{Für alle } n \in \mathbb{N} \text{ definiere } X_{n+1} \text{ unter Rückgriff auf } X_n. \end{array}\right\}$$

Hierbei ist (nur jetzt einmal, nicht bei jeder Anwendung) zu überlegen, dass das Objekt X_n durch diese Festlegungen tatsächlich für alle $n \in \mathbb{N}$ definiert wird. Genau dies ergibt aber die Anwendung des Induktionsprinzips (VI) auf die n-abhängige Aussage „X_n ist (wohl)definiert“ (jedenfalls sofern beim Rekursionsschritt X_{n+1} *für gegebenes* X_n immer wohldefiniert ist). Analog zum Induktionsprinzip besitzt auch (RD) Varianten mit Rekursionsanfang bei beliebigem $z_0 \in \mathbb{Z}$, mehreren Einzelfällen als Rekursionsanfang, Rekursionsschritt mit Rückgriff auf mehrere Vorgänger-Objekte, et cetera. Damit wird es möglich, viele naheliegende Definitionen präziser als mit Pünktchen (oder auch überhaupt erst) hinzuschreiben:

Beispiele II.2.8.

- Ausgehend von der bijektiven Nachfolge-Abbildung $S\colon \mathbb{Z} \to \mathbb{Z}$ kann man die *Summe* $z+n \in \mathbb{Z}$ und die *Differenz* $z-n \in \mathbb{Z}$ von $z \in \mathbb{Z}$ und $n \in \mathbb{N}_0$ erklären, dies aber auf verschiedene Weisen hinschreiben: Mit Pünktchen lauten die Definitionen
$$z+n := \underbrace{S(S(S(\ldots(S(z))\ldots)))}_{n \text{ Anwendungen von } S} \quad \text{und} \quad z-n := \underbrace{S^{-1}(S^{-1}(S^{-1}(\ldots(S^{-1}(z))\ldots)))}_{n \text{ Anwendungen von } S^{-1}}.$$
Mit dem Rekursionsprinzip können wir dieselben Definitionen ganz präzise und frei von Andeutungen durch Pünktchen für alle $z \in \mathbb{Z}$ durch
$$z \pm 0 := z, \quad z + S(n) := S(z+n) \text{ für } n \in \mathbb{N}_0, \quad z - S(n) := S^{-1}(z-n) \text{ für } n \in \mathbb{N}_0$$
treffen. Wenn im Vorfeld $z+1 := S(z)$, $z-1 := S^{-1}(z)$ vereinbart wird, können beide Varianten etwas vertrauter hingeschrieben werden, nämlich als
$$z+n := (\ldots(((z\underbrace{+1)+1)+1)+\ldots)+1}_{n \text{ Summanden } 1} \text{ und } z-n := (\ldots(((z\underbrace{-1)-1)-1)-\ldots)-1}_{n \text{ Subtrahenden } 1}$$
und mit Rekursion als
$$z \pm 0 := z, \quad z + (n+1) := (z+n) + 1 \text{ für } n \in \mathbb{N}_0, \quad z - (n+1) := (z-n) - 1 \text{ für } n \in \mathbb{N}_0.$$
Um $z \pm \widetilde{z} \in \mathbb{Z}$ sogar für beliebige $z, \widetilde{z} \in \mathbb{Z}$ zu erklären, vereinbart man im Nachhinein noch
$$z + (-n) := z - n \text{ für } n \in \mathbb{N} \quad \text{und} \quad z - (-n) := z + n \text{ für } n \in \mathbb{N}.$$
- Ausgehend von der Summe kann das *Produkt*
$$n \cdot z := nz := \underbrace{(\ldots((z+z)+z)+\ldots)+z}_{n \text{ Summanden } z} \in \mathbb{Z}$$
von $n \in \mathbb{N}_0$ und $z \in \mathbb{Z}$ erklärt werden. Die Präzisierung dieser Pünktchen-Definition durch Rekursion für alle $z \in \mathbb{Z}$ ist
$$0z := 0 \quad \text{und} \quad (n+1)z := (nz)+z \text{ für } n \in \mathbb{N}_0.$$
Um $z\widetilde{z} \in \mathbb{Z}$ sogar für beliebige $z, \widetilde{z} \in \mathbb{Z}$ zu erklären, ergänzt man die Festlegung
$$(-n)z := n(-z) \text{ für } n \in \mathbb{N}.$$
- Mit Hilfe des Produkts können *Potenzen*
$$z^n := \underbrace{(\ldots((z \cdot z) \cdot z) \cdot \ldots) \cdot z}_{n \text{ Faktoren } z} \in \mathbb{Z}$$
mit Basis $z \in \mathbb{Z}$ und Exponent $n \in \mathbb{N}$ definiert werden. Die präzisere rekursive Definition für jedes $z \in \mathbb{Z}$ lautet
$$z^1 := z \quad \text{und} \quad z^{n+1} := z^n \cdot z \text{ für } n \in \mathbb{N}.$$

Ergänzend definiert man

$$z^0 := 1 \text{ (zumindest) für } z \in \mathbb{Z} \setminus \{0\}.$$

Der Ausdruck 0^0 *bleibt generell undefiniert* (begründet zum Beispiel dadurch, dass man die Regeln $z^0 = 1$ für alle $z \in \mathbb{Z} \setminus \{0\}$ und $0^n = 0$ *nicht beide* konsistent fortsetzen kann). Später wird es sich in manchen Zusammenhängen als sinnvoll erweisen, 0^0 als 1 festzulegen.

Auf Basis der Definitionen kann man die Kommutativität und Assoziativität der Addition und Multiplikation, die Distributivgesetze und weitere bekannte Rechenregeln ($z-z = 0$, Klammer-Regeln, binomische Formel, et cetera) für den (symbolischen) Umgang mit ganzen Zahlen herleiten. Zu einem großen Teil kann dies mit Induktionsbeweisen nachgewiesen werden, was hier (abgesehen von einem Beispiel in den Übungen) aber ausgespart werden soll. Wir treffen ab jetzt außerdem die üblichen Konventionen zur Klammereinsparung: Es gilt Punkt- vor Strich-Rechnung, und die dadurch und durch Assoziativität überflüssigen Klammern werden weggelassen. *Quotienten* $z : n = z/n = \frac{z}{n} \in \mathbb{Z}$ von $z \in \mathbb{Z}$ und $n \in \mathbb{Z} \setminus \{0\}$ können im Rahmen der ganzen Zahlen natürlich nur dann definiert werden, wenn z durch n teilbar ist. Formal würde man den Quotienten $\frac{z}{n}$ an dieser Stelle daher als die eindeutige Zahl $q \in \mathbb{Z}$, sofern eine solche denn existiert, mit $nq = z$ festlegen.

Beispiele II.2.9.

- Die *Fakultät*

$$n! := n(n-1)\cdot \ldots \cdot 3\cdot 2\cdot 1 \in \mathbb{N}$$

von $n \in \mathbb{N}_0$ wird rekursiv definiert durch

$$0! := 1, \quad (n+1)! := (n+1)(n!) \text{ für } n \in \mathbb{N}_0.$$

Beispiele sind $0! = 1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $6! = 720$, $7! = 5040$, $8! = 40320$.

- Die *Fibonacci-Zahlen* $F_n \in \mathbb{N}_0$ mit $n \in \mathbb{N}_0$ sind rekursiv durch

$$F_0 := 0, \quad F_1 := 1 \quad \text{und} \quad F_{n+1} := F_{n-1}+F_n \text{ für alle } n \in \mathbb{N}$$

definiert. Die ersten Fibonacci-Zahlen sind $F_0 = 0$, $F_1 = 1$, $F_2 = 1$, $F_3 = 2$, $F_4 = 3$, $F_5 = 5$, $F_6 = 8$, $F_7 = 13$, $F_8 = 21$, $F_9 = 34$, $F_{10} = 55$. Überraschenderweise kann man (z.B. mit vollständiger Induktion) auch die für alle $n \in \mathbb{N}$ gültige geschlossene *Formel von Binet*

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

beweisen. Etwas mehr zu Beweis und Hintergrund dieser Formel folgt in den Übungen.

- Für $k \leq \ell$ in $\mathbb{Z}$ wird die Summe von $\ell-k+1$ Zahlen $a_k, a_{k+1}, \ldots, a_{\ell-1}, a_\ell \in \mathbb{Z}$ mit dem *Summenzeichen* $\sum$ (nach dem Sigma genannten griechischen Buchstaben Σ) als

$$\sum_{i=k}^{\ell} a_i := a_k+a_{k+1}+\ldots+a_{\ell-1}+a_\ell \in \mathbb{Z}$$

abgekürzt. Rekursiv kann man für gegebene $a_i \in \mathbb{Z}$ erst

$$\sum_{i=0}^{0} a_i := a_0 \quad \text{und} \quad \sum_{i=0}^{n} a_i := \left(\sum_{i=0}^{n-1} a_i\right) + a_n \text{ für } n \in \mathbb{N},$$

darauf aufbauend dann

$$\sum_{i=k}^{\ell} a_i := \sum_{i=0}^{\ell-k} a_{k+i} \text{ für } k \leq \ell \text{ in } \mathbb{Z}$$

definieren. Spezialfälle sind Summen von Einsen $\sum_{i=k}^{\ell} 1 = \ell-k+1$ und Produkte $nz = \sum_{i=1}^{n} z$.

- Analog wird für $k \leq \ell$ in $\mathbb{Z}$ das Produkt von $a_k, a_{k+1}, \ldots, a_{\ell-1}, a_\ell \in \mathbb{Z}$ mit dem *Produktzeichen* $\prod$ (nach dem Pi genannten griechischen Buchstaben Π) als

$$\prod_{i=k}^{\ell} a_i := a_k a_{k+1}\cdot \ldots \cdot a_{\ell-1}a_\ell \in \mathbb{Z}$$

abgekürzt. Die rekursive Definition kann genau wie beim Summenzeichen ausgeschrieben werden. Spezialfälle sind Potenzen $z^n = \prod_{i=1}^{n} z$ und Fakultäten $n! = \prod_{i=1}^{n} i$.

In Ergänzung zu (II.2.9) und (II.2.9) notiert man auch $\sum_{i\in I} a_i := \sum_{j=1}^{n} a_{i_j}$ und $\prod_{i\in I} a_i := \prod_{j=1}^{n} a_{i_j}$ für jede endliche Indexmenge $I = \{i_1, i_2, \ldots, i_n\}$ mit $n \in \mathbb{N}$ Elementen.

Ergänzend nutzt man für „leere“ Summen und Produkte die Konventionen $\sum_{i=k}^{\ell} a_i := 0$, $\prod_{i=k}^{\ell} a_i := 1$ im Fall $k > \ell$ sowie $\sum_{i\in\emptyset} a_i := 0$, $\prod_{i\in\emptyset} a_i := 1$.

Wichtig für das Rechnen mit Summen- und Produktzeichen sind die Regeln für *Indexverschiebung* (mit $k, \ell \in \mathbb{Z}$, $v, a_i \in \mathbb{Z}$; entspricht Substitution $j = i+v$)

$$\sum_{i=k}^{\ell} a_i = \sum_{j=k+v}^{\ell+v} a_{j-v}, \qquad \prod_{i=k}^{\ell} a_i = \prod_{j=k+v}^{\ell+v} a_{j-v},$$

Indexlaufumkehr (mit $k, \ell \in \mathbb{Z}$, $a_i \in \mathbb{Z}$; entspricht Substitution $j = k+\ell-i$)

$$\sum_{i=k}^{\ell} a_i = \sum_{j=k}^{\ell} a_{k+\ell-j}, \qquad \prod_{i=k}^{\ell} a_i = \prod_{j=k}^{\ell} a_{k+\ell-j}$$

und *Verhalten bei Summen und Produkten* (mit Indexmenge I, $a_i, b_i, z \in \mathbb{Z}$, $n \in \mathbb{N}$)

$$\sum_{i\in I}(a_i+b_i) = \left(\sum_{i\in I} a_i\right) + \left(\sum_{i\in I} b_i\right), \qquad \prod_{i\in I}(a_i \cdot b_i) = \left(\prod_{i\in I} a_i\right)\left(\prod_{i\in I} b_i\right),$$

$$\sum_{i\in I}(za_i) = z\sum_{i\in I} a_i, \qquad \prod_{i\in I}(a_i^n) = \left(\prod_{i\in I} a_i\right)^n.$$

Wir diskutieren als Nächstes die Ihnen allen bekannte Darstellung ganzer Zahlen, bei denen Ziffern an unterschiedlichen Positionen unterschiedliches Gewicht besitzen, also die Darstellung in sogenannten Stellenwertsystemen:

Bemerkungen II.2.10.

- Im *Dezimalsystem* (auch dekadisches System oder Zehnersystem genannt) werden die bereits erklärten 1-stelligen Zahlen $0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \in \mathbb{N}_0$ als Ziffern verwendet. Allgemeiner erhält man eine $(n+1)$-stellige Zahl mit $n \in \mathbb{N}_0$ durch Hintereinanderschreiben solcher Ziffern $z_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ und unter Verwendung der *Basis* $10 := S(9)$ als

  $$z_n z_{n-1} \ldots z_2 z_1 z_0 := \sum_{i=0}^{n} z_i \cdot 10^i \in \mathbb{N}_0,$$

  wobei das Hintereinanderschreiben von Ziffern ohne dazwischen gestelltes Symbol hier ausnahmsweise *nicht* für das Produkt steht. Man kann dies als rekursive Definition betrachten, bei der der Rekursionsanfang mit der Definition der 1-stelligen Zahlen bereits erfolgt ist und im Rekursionsschritt für jedes $n \in \mathbb{N}$ die $(n+1)$-stellige Zahl $zm := z\cdot 10^n + m$ mit Ziffer $z \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ und n-stelliger natürlicher Zahl m definiert wird.

  Werden Ziffern ohne Verknüpfungszeichen und Erläuterung verwendet oder hintereinander geschrieben, so ist standardmäßig immer dieses System gemeint.
- In *Stellenwertsystemen mit beliebiger Basis* $b \in \mathbb{N} \setminus \{1\}$ arbeitet man mit b Ziffern. Man verwendet (jedenfalls heutzutage und in unserem Kulturkreis) im Fall $b \leq 10$ normalerweise die b ersten Ziffern $0, 1, 2, \ldots, b-2, b-1 \in \mathbb{Z}$ des Dezimalsystems in üblicher Bedeutung und muss für $b > 10$ Ziffersymbole für die Dezimalzahlen $10, 11, 12, \ldots, b-2, b-1 \in \mathbb{Z}$ hinzufügen. Für jedes $n \in \mathbb{N}_0$ werden $(n+1)$-stellige Zahlen mit (im jeweiligen System zulässigen) Ziffern z_i in Analogie zum Dezimalsystem durch

  $$(z_n z_{n-1} \ldots z_2 z_1 z_0)_b := \sum_{i=0}^{n} z_i \cdot b^i \in \mathbb{N}_0$$

erklärt (wobei die als Subskript anzugebende Basis b benannt wird). An verschiedenen Stellen übliche Basen sind $b = 2$ (Binärsystem/Dualsystem/Zweiersystem), $b = 3$ (Ternärsystem), $b = 8$ (Oktalsystem/Achtersystem), $b = 10$ (Dezimalsystem aus (II.2.10)), $b = 12$ (Duodezimalsystem/Zwölfersystem), $b = 16$ (Hexadezimalsystem/Sechzehnersystem). Beim letztgenannten System sind als Zusatz-Ziffern A := 10, B := 11, C := 12, D := 13, E := 14, F := 15 weitgehend üblich.

Beispiele II.2.11. Nachrechnen zeigt $(11111100101)_2 = (2202212)_3 = (3745)_8 = 2021 = (1205)_{12} = (7E5)_{16}$.

Wie für die Basis 10 ist auch für beliebige Basen $b \in \mathbb{N} \setminus \{1\}$ richtig, dass jede natürliche Zahl eine Zifferndarstellung in dieser Basis hat und verschiedene Zifferndarstellungen ohne führende Null-Ziffern zu verschiedenen natürlichen Zahlen gehören. Dies müsste man natürlich beweisen, worauf wir hier aber verzichten.

Zum Abschluss dieses Abschnitts sei angemerkt, dass *man die natürlichen Zahlen auch allein aus den Axiomen der Mengenlehre* aus Abschnitt I.5 *konstruieren kann*, womit sich die Peano-Axiome tatsächlich von zusätzlichen Postulaten zu herleitbaren Eigenschaften wandeln. Mit anderen Worten können die Existenz der natürlichen Zahlen und ihre Grundeigenschaften rein mengentheoretisch unterbaut werden, was im Detail aber über den Vorlesungsstoff hinausgeht und nur im Anhang im Abschnitt V.2 ausgeführt wird.

II.3. Beweisstrategien

Vorausgeschickt sei die Warnung, dass es kein Patentrezept zum Finden eines mathematischen Beweises gibt. Einige Beweise kann man nur über Erfahrung, Intuition, Umformulieren des Problems, Betrachtung aus verschiedenen Blickwinkeln und das Ausprobieren verschiedener Ansätze angehen. Und am Ende scheitert man eventuell trotzdem...

Dennoch seien hier einige Hinweise zu generellen Strategien und häufigen Vorgehensweisen zusammengetragen. Zu einem großen Teil lehnen sich die Hinweise aber eng an Definitionen an und kamen auf die ein oder andere Weise schon vor.

Strategien für Beweise. Generell hängt das Vorgehen bei einem Beweis stark von der behaupteten Aussage ab. Es folgen Hinweise sowohl für spezielle Fälle als auch allgemeiner Natur.

(a) Beweise von *Aussagen mit Teilaussagen* A, B:
- Implikation $A \Longrightarrow B$: Dies ist der Prototyp eines logischen Schlusses. Es gibt hierfür drei prinzipielle Möglichkeiten:
  Beim *direkten Beweis* nimmt man A als wahr an und argumentiert, dass B dann ebenfalls wahr sein muss; vergleiche mit Abschnitt I.4.
  Der *Beweis durch Kontraposition* nutzt das Kontrapositions-Prinzip aus Abschnitt I.2: Man zeigt $(\neg B) \Longrightarrow (\neg A)$, geht also von $\neg B$ aus und schließt auf $\neg A$.
  Beim *indirekten Beweis* oder *Widerspruchsbeweis*, auch *reductio ad absurdum* genannt, nimmt man A und zudem die Widerspruchsannahme $\neg B$ als wahr an und zeigt, dass hieraus ein Widerspruch entsteht. Dieses Vorgehen ist durch die Definition der Implikation selbst gerechtfertigt: Man führt $A \wedge (\neg B)$ als den einzigen Fall, in dem $A \Longrightarrow B$ falsch ist, zum Widerspruch und schließt diesen somit aus.
  Die logischen Verneinungen bei Kontraposition und indirektem Beweis betreffen oft Aussagen mit Quantoren. In solchen Fällen denke man an die zugehörigen Regeln aus Abschnitt I.3, auch dann, wenn die Quantoren in Worten und nicht explizit als Formelzeichen auftreten.
- Äquivalenz $A \Longleftrightarrow B$: Meist zeigt man die Richtungen „$\Longrightarrow$“ und „$\Longleftarrow$“ separat. Dafür kann jede der gerade besprochenen Strategien zum Einsatz kommen.
  Seltener, aber nicht völlig ungewöhnlich ist, eine Äquivalenz durch Aneinandersetzen bekannter oder einfacher Äquivalenzen (zum Beispiel Äquivalenzumformungen) $A \Longleftrightarrow H_1$, $H_1 \Longleftrightarrow H_2$, $H_2 \Longleftrightarrow H_3$, …, $H_{n-1} \Longleftrightarrow H_n$, $H_n \Longleftrightarrow B$ (mit Hilfsaussagen $H_1, H_2, \ldots, H_n$) zu zeigen.

(b) In fast jedem Zusammenhang kommen *Fallunterscheidungen* in Betracht, bei denen im Beweis Fälle mit verschiedenen Annahme einzeln abgearbeitet werden. Entscheidend ist natürlich, dass die Fälle die Gesamtheit aller Möglichkeiten abdecken müssen.

(c) *Existenz- und Eindeutigkeitsbeweise*:
- Zu Existenzbeweisen lässt sich wenig Allgemeines sagen.
  Nur um auf eine Existenzaussage der Form $\exists x \in \mathcal{X} : \forall y \in \mathcal{Y} : P(x, y)$ zu schließen, bietet sich manchmal ein indirekter Beweis an, da die Widerspruchsannahme für jedes $x \in \mathcal{X}$ ein $y_x \in \mathcal{Y}$ mit $\neg P(x, y_x)$ gibt und man mit den „Gegenbeispielen“ y_x eventuell gut argumentieren kann.
- Bei Eindeutigkeitsbeweisen ist für zwei Objekte x, y mit gewissen behaupteten Eigenschaften die Gleichheit $x = y$ zu zeigen. Dies kann direkt oder indirekt geschehen. Bei letzterem nimmt man an, dass x, y mit $x \neq y$ die behaupteten Eigenschaften haben und erzeugt einen Widerspruch.

(d) Beweise von *Aussagen über Mengen* M, N:
- Mengen-Inklusion $M \subset N$. Die Implikation $x \in M \Longrightarrow x \in N$ kann direkt, durch Kontraposition (entspricht dem Nachweis $N^c \subset M^c$, wenn $M, N \subset \mathcal{X}$ für eine Grundmenge $\mathcal{X}$) oder durch Widerspruch (entspricht dem Nachweis $M \setminus N \subset \emptyset$) geschehen.
  Oft kann man auch abstrakt ohne Betrachtung einzelner Elemente argumentieren, zum Beispiel durch Zusammensetzen schon bekannter Inklusionen.
- Mengen-Gleichheit $M = N$: Meist zeigt man die Inklusionen „$\subset$“ und „$\supset$“ separat.
  Seltener kann man direkt die Äquivalenz $x \in M \Longleftrightarrow x \in N$ nachweisen oder abstrakter argumentieren.
- Das Widerlegen solcher Aussagen ist viel einfacher. Um $M \not\subset N$ beziehungsweise $M \neq N$ zu zeigen, müssen Sie nur ein Element $x \in M \setminus N$ beziehungsweise $x \in M \Delta N$, also ein Gegenbeispiel angeben.

(e) Zu Beweisen von *Aussagen über Zahlen* x, y lässt sich nur weniger aussagekräftig sagen:
- Gelegentlich weist man eine Ungleichung $x \leq y$ für $x, y \in \mathbb{R}$ nach, indem man $x \leq y + \varepsilon$ für alle $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ (oder nur für alle $\varepsilon \in \mathbb{Q}$ mit $\varepsilon > 0$) zeigt. Genaueres dazu später noch!
- Gelegentlich weist man eine Gleichheit $x = y$ nach, indem man „$\leq$“ und „$\geq$“ separat zeigt.

Anders als bei Mengen sind die beschriebenen Vorgehensweisen hier aber nicht kanonisch.

(f) Beweise von *Aussagen über n-Tupel* x, y (z.B. Paare oder Tripel):

- Die Gleichheit $x = y$ von n-Tupeln zeigt man oft, indem man $x_i = y_i$ für alle $i \in \{1, 2, \ldots, n\}$ nachweist.

(g) Beweise von *Aussagen über Abbildungen* $f, g \colon \mathcal{X} \to \mathcal{Y}$:

- Gleichheit $f = g$ von Abbildungen: Vorab ist zu prüfen, dass f und g den gleichen Definitionsbereich $\mathcal{X}$ und je nach genauer Auffassungsweise (siehe Abschnitt II.4) gleichen Zielbereich $\mathcal{Y}$ haben. Man zeigt dann oft $f(x) = g(x)$ für alle $x \in \mathcal{X}$.
- Injektivität von f: Die Implikation $f(x) = f(\widetilde{x}) \Longrightarrow x = \widetilde{x}$ für alle $x, \widetilde{x} \in \mathcal{X}$ können Sie direkt, per Kontraposition ($x \neq \widetilde{x} \Longrightarrow f(x) \neq f(\widetilde{x})$) oder indirekt ($f(x) = f(\widetilde{x})$ für $x \neq \widetilde{x}$ führt zum Widerspruch) nachweisen. Dies ist ein spezieller Fall eines Eindeutigkeitsbeweises.
- Surjektivität von f: Oft gibt man sich ein beliebiges $y \in \mathcal{Y}$ vor und zeigt die Existenz eines $x \in \mathcal{X}$ mit $f(x) = y$. Dies ist ein spezieller Fall eines Existenzbeweises.
- Bijektivität von f: Oft zeigt man Injektivität von f und Surjektivität von f separat.
- Das Widerlegen ist meist wieder einfacher: Für $f \neq g$ muss man nur *ein* $x \in \mathcal{X}$ mit $f(x) \neq g(x)$ angeben, für Nicht-Injektivität von f *zwei* $x, \widetilde{x} \in \mathcal{X}$ mit $x \neq \widetilde{x}$, $f(x) = f(\widetilde{x})$, für Nicht-Surjektivität von f *ein* $y \notin \mathrm{Bild}(f)$ (was allerdings $f(x) \neq y$ für *alle* $x \in \mathcal{X}$ bedeutet).

Oft kann man auch abstrakter ohne Betrachtung einzelner Elemente argumentieren, zum Beispiel über die Komposition und bereits bekannte Eigenschaften gewisser Abbildungen.

(h) Beweise von *Aussagen* $A(n)$*, die von* $n \in \mathbb{N}$ *(oder von* $z \in \{z_0, z_0{+}1, z_0{+}2, \ldots\} \subset \mathbb{N}$*)* abhängen:

- Das *Prinzip der vollständigen Induktion* wurde in Abschnitt II.2 besprochen.
- Das *Prinzip des kleinsten Gegenbeispiels* ist eine *indirekte Variante des Induktionsprinzips*, die aber relativ selten benötigt wird. Man zeigt dabei den Induktionsanfang $A(1)$ und macht die Widerspruchsannahme, dass $A(n)$ *nicht* für alle $n \in \mathbb{N}$ gilt. Diese Annahme erzwingt, dass einer

der Induktionsschritte beim Induktionsprinzip der Form ($\widetilde{\text{VI}}$) scheitern muss. Es gibt also ein $n \in \mathbb{N}$, so dass $A(1)$, $A(2)$, $\dots A(n)$ alle wahr sind, aber $A(n{+}1)$ falsch (und somit $n{+}1$ das hypothetische „kleinste Gegenbeispiel“) ist. Kann man auf dieser Grundlage einen Widerspruch herleiten, so hat man $A(n)$ für alle $n \in \mathbb{N}$ gezeigt.

(i) Ab und zu kann man bei Beweisen *Ausdehnungsprozeduren* einsetzen, also eine Aussage im ersten Schritt für spezielle Objekte beziehungsweise eine Variable in einer Teilmenge zeigen, im zweiten Schritt unter Rückgriff auf den ersten für etwas weniger spezielle Objekte bzw. eine größere Teilmenge und erst im n-ten Schritt irgendwann allgemein bzw. für die ganze Menge. Bei Aussagen über Zahlen kann sich dies zum Beispiel so gestalten, dass man eine Aussage erst für natürliche Zahlen, dann für ganze Zahlen, dann für rationale Zahlen und schließlich für reelle Zahlen nachweist.

(j) In einem Beweis können die bisher genannten Techniken beliebig kombiniert oder auch dieselbe Technik mehrfach angewandt werden. *Allgemeine Tipps*, um einen Beweis zu finden oder zu verifizieren sind:

- Das Ziel des Beweises bewusst aufschreiben, zum Beispiel als „Zu zeigen: …“ oder „Behauptung: …“. Dies ist am Anfang der Lösung einer Beweisaufgabe immer gern gesehen.
- Eine formale Prüfung der Aussage vornehmen, etwa, ob gleichgesetzte Objekte überhaupt vom gleichen Typ (Zahl, Paar, n-Tupel, Menge, Abbildung) sind, Argumente im Definitionsbereich einer Abbildung liegen, bei der Komposition von Abbildungen das Ziel der inneren Abbildung gleich dem Definitionsbereich der äußeren Abbildung ist! Man kann dabei auch Grenzfälle abklopfen, zum Beispiel den, dass eine Menge leer oder die ganze Grundmenge ist, eine Zahl den größten oder kleinsten möglichen Wert (zum Beispiel Null) annimmt, et cetera. Auch wenn die Aussagen der Vorlesung und der Übungen in der Regel (formal) korrekt sind, so hilft die Prüfung doch oft beim besseren Verständnis der Ausgangssituation.
- Sich beim Beweis einer Implikation $A \Longrightarrow B$ von Anfang *und* Ende annähern, also sowohl überlegen, was mit A gezeigt werden kann, als auch, was denn reichen würde, um damit B zu zeigen.
- Wenn der allgemeine Fall nicht in Reichweite scheint, dann zuerst einen einfachen Fall oder wichtigen Modellfall betrachten und erst danach dessen Lösung verallgemeinern
- Teilstücke der Argumentation immer wieder durchgehen, kritisch hinterfragen und auf Korrektheit prüfen, auch anhand von speziellen Grenz- und Modellfällen.

Je schwieriger und umfangreicher der Beweis und sein Kontext sich gestalten, desto wichtiger werden gerade die letztgenannten Tipps.

II.4. Relationen

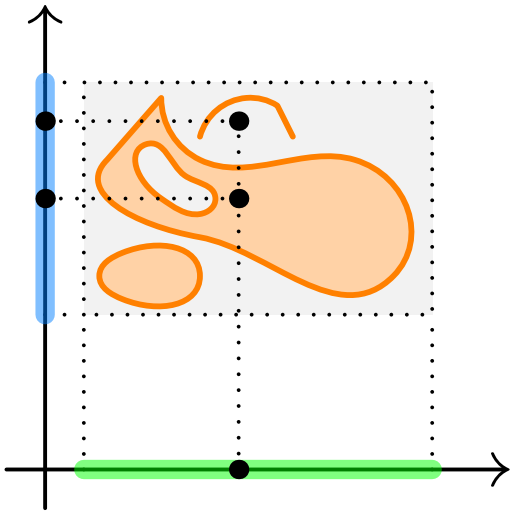
Relationen treten in der Mathematik in vielfältiger Gestalt auf. Wir beginnen hier mit dem abstrakten Konzept und wenden uns danach den wichtigen Spezialfällen zu.

Definition II.4.1. Es seien $\mathcal{X}$ und $\mathcal{Y}$ beliebige Mengen.

(a) Eine (zweistellige) *Relation* R zwischen (den Elementen von) $\mathcal{X}$ und $\mathcal{Y}$ ist ein Tripel $(\mathcal{X}, \mathcal{Y}, R)$ mit einer Teilmenge R des kartesischen Produkts $\mathcal{X}\times\mathcal{Y}$. Eine Relation zwischen $\mathcal{X}$ und $\mathcal{X}$ bezeichnen wir als Relation zwischen den Elementen von $\mathcal{X}$ oder Relation auf $\mathcal{X}$.

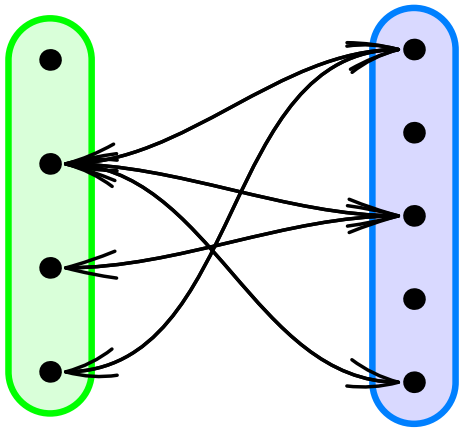
(b) Die *Menge aller Relationen zwischen* $\mathcal{X}$ *und* $\mathcal{Y}$ (die formal gleich $\{\mathcal{X}\}\times\{\mathcal{Y}\}\times\mathcal{P}(\mathcal{X}\times\mathcal{Y})$ ist) bezeichnen wir mit $\mathrm{Rel}(\mathcal{X}, \mathcal{Y})$. Wir kürzen $\mathrm{Rel}(\mathcal{X}, \mathcal{X})$ durch $\mathrm{Rel}(\mathcal{X})$ ab.

Ein $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ mit $\mathcal{X}, \mathcal{Y} \subset \mathbb{R}$:



(II.4.1)

Darstellung einer Relation R zwischen Mengen $\mathcal{X}$ und $\mathcal{Y}$ mit endlich vielen Elementen:



(II.4.2)

Bemerkung II.4.2. Für beliebige $n \in \mathbb{N}$ kann man eine *n-stellige Relation* zwischen Mengen $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ als Tupel $(\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n, T)$ mit einer Teilmenge T des n-fachen kartesischen Produkts $\mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n$ erklären. Wir behandeln nur den zuvor betrachteten und mit Abstand wichtigsten Fall $n = 2$.

Das wesentliche Objekt in der Definition einer Relation ist der letzte Eintrag des Tupels, also bei $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ die Teilmenge $R \subset \mathcal{X} \times \mathcal{Y}$. Die wahre Bedeutung von Relationen erschließt sich aber tatsächlich weniger aus der Definition und mehr aus folgenden Sprech-, Schreib- und Betrachtungsweisen:

Notation II.4.3. Es seien $\mathcal{X}$, $\mathcal{Y}$ Mengen, $x \in \mathcal{X}$, $y \in \mathcal{Y}$ und $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$. Wir sagen und schreiben im Fall $(x, y) \in R$, dass x mit y bezüglich R in Relation steht, oder, dass die Aussage xRy gilt. Sehr oft interpretieren wir eine Relation nicht als Tupel oder Teilmenge, sondern als eine Beziehung zwischen Elementen x und y; vergleiche mit Abbildung II.4.2. Die Beziehung kommt zum Ausdruck, indem wir R in der gerade eingeführten Notation xRy zwischen x und y schreiben. In Zukunft werden wir anstelle von R nicht nur andere Buchstaben, sondern häufig auch anders geartete Symbole verwenden.

Beispiele II.4.4.

(a) Für beliebige Mengen $\mathcal{X}, \mathcal{Y}$ enthält $\mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ die Leer-Relation $R = (\mathcal{X}, \mathcal{Y}, \emptyset)$ (für die xRy nie gilt) und die All-Relation $R = (\mathcal{X}, \mathcal{Y}, \mathcal{X} \times \mathcal{Y})$ (für die xRy immer gilt).

(b) Durch

$$y \sim_g z :\Longleftrightarrow z - y \text{ ist gerade}, \qquad y \sim_u z :\Longleftrightarrow z - y \text{ ist ungerade}$$

für $y, z \in \mathbb{Z}$ werden zwei Relationen $\sim_g$, $\sim_u$ in $\mathrm{Rel}(\mathbb{Z})$ erklärt, von denen für jedes $(y, z) \in \mathbb{Z}^2$ eine gilt und eine nicht. Zum Beispiel gelten $2 \sim_g 4$, $(-2) \sim_u 5$, $(-7) \sim_g (-7)$ und $9 \sim_u 0$. Die zugehörigen Teilmengen sind die Mengen $R_g = \{(y, z) \in \mathbb{Z}^2 \mid z-y \text{ ist gerade}\}$ und $R_u = \{(y, z) \in \mathbb{Z}^2 \mid z-y \text{ ist ungerade}\}$.

(c) Für jede Menge M ist die *Gleichheit „=“ von Elementen* eine Relation $(M, M, R_=)$ auf M mit der Teilmenge $R_= = \Delta_M := \{(x, y) \in M^2 \mid x = y\}$.

(d) Für jede Grundmenge $\mathcal{X}$ geben die *Gleichheit, Ungleichheit und (strikte) Inklusion von Mengen*, also alle Symbole $\square \in \{=, \neq, \subset, \supset, \subsetneq, \supsetneq\}$, Relationen zwischen Teilmengen von $\mathcal{X}$ oder mit anderen Worten Relationen auf $\mathcal{P}(\mathcal{X})$.

(e) Für jede Grundmenge $\mathcal{X}$ und jedes Mengensystem $\mathcal{S}$ (zum Beispiel $\mathcal{S} = \mathcal{P}(\mathcal{X})$) ist *die Element-Beziehung* „$\in$“ eine Relation auf $\mathcal{X} \times \mathcal{S}$ mit zugehöriger Teilmenge $R_\in = \{(x, M) \in \mathcal{X} \times \mathcal{S} \mid x \in M\}$. Analog kann man „$\ni$“ als Relation auf $\mathcal{S} \times \mathcal{X}$ verstehen.

(f) Die *Gleichheit, Ungleichheit und kleiner, größer für Zahlen*, also alle Symbole $\square \in \{=, \neq, <, >, \leq, \geq\}$, geben Relationen auf jedem Zahlbereich $\mathbb{B} \in \{\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$. Dabei sind Gleichheit und Ungleichheit generell definiert. Auch die anderen Symbole sind Ihnen prinzipiell vertraut. Präzise können „$<$“ und „$>$“ durch

$$z > y :\iff y < z :\iff z-y \in \mathbb{N} \quad \text{für } y, z \in \mathbb{Z}$$

erklärt und später auf $y, z \in \mathbb{Q}$ und $y, z \in \mathbb{R}$ verallgemeinert werden. Darauf aufbauend bedeuten $y \leq z$ und $z \geq y$ natürlich nichts anderes als $(y < z) \vee (y = z)$.

(g) Die *Gleichheit und kleiner und größer Beziehungen zwischen Abbildungen*, also alle Symbole $\square \in \{=, <, >, \leq, \geq\}$, geben Relationen auf Mengen $\mathrm{Abb}(\mathcal{X}, \mathbb{B})$ von Abbildungen (mit beliebiger Menge $\mathcal{X}$ und $\mathbb{B} \in \{\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$), wobei man für $f, g \in \mathrm{Abb}(\mathcal{X}, \mathbb{B})$ die meisten genannten Symbole $\square$ standardmäßig im punktweisen Sinn

$$f \square g :\iff f \square g \text{ auf } \mathcal{X} :\iff \forall x \in \mathcal{X} \colon f(x) \square g(x)$$

versteht. Die Ungleichheit $f \neq g$ verstehen wir als logisches Gegenteil $\exists x \in \mathcal{X} \colon f(x) \neq g(x)$ der Gleichheit $f = g$.

(h) Die *Teilbarkeitsrelation* „$|$“ ist auf jedem ganzzahligen Zahlbereich $\mathbb{B} \in \{\mathbb{N}_0, \mathbb{N}, \mathbb{Z}\}$ sinnvoll. Dabei bedeutet $t|z$ mit $t, z \in \mathbb{B}$, dass t ein Teiler von z ist, oder als Formel

$$t \mid z :\iff \exists d \in \mathbb{B} \colon td = z \quad \text{für } t, z, \in \mathbb{B}.$$

(i) Für jede feste Grundmenge $\mathcal{X}$ ist *Disjunktheit von Teilmengen* eine Relation auf $\mathcal{P}(\mathcal{X})$.

(j) Jede Funktion $f \colon \mathcal{X} \to \mathcal{Y}$ zwischen Mengen $\mathcal{X}, \mathcal{Y}$ induziert eine Relation $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ mit $R = \mathrm{G}_f$, wobei $\mathrm{G}_f = \{(x, y) \in \mathcal{X} \times \mathcal{Y} \mid f(x) = y\}$ definiert war, oder äquivalent mit

$$xRy :\iff f(x) = y \quad \text{für alle } x \in \mathcal{X},\ y \in \mathcal{Y}.$$

Die Relation R erbt von der Funktion f die folgende entscheidende Eigenschaft:

(II.4.3) Für alle $x \in \mathcal{X}$ gibt es genau ein $y \in \mathcal{Y}$ mit xRy.

Tatsächlich kommt jede Relation $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ mit der Eigenschaft (II.4.3) des letzten Beispiels durch eine eindeutig bestimmte Abbildung $f \colon \mathcal{X} \to \mathcal{Y}$ zustande, denn man kann den Funktionswert $f(x)$ als das eindeutige, zu $x \in \mathcal{X}$ gehörige $y \in \mathcal{Y}$ mit xRy festsetzen. Wir erhalten eine 1-zu-1-Korrespondenz zwischen Abbildungen $\mathcal{X} \to \mathcal{Y}$ und Relationen in $\mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ mit der Eigenschaft (II.4.3) und können Abbildungen fortan als spezielle Relationen auffassen. Dies können wir auch benutzen, um den Begriff der Abbildung aus Abschnitt II.1 – wo, wir erinnern uns, der nicht formal definierte Begriff „Zuordnungsvorschrift“ einging – völlig präzise auf den Punkt zu bringen und mengentheoretisch zu unterfüttern:

Definition II.4.5. Es seien $\mathcal{X}$ und $\mathcal{Y}$ Mengen. Eine Abbildung $f \colon \mathcal{X} \to \mathcal{Y}$ von $\mathcal{X}$ nach $\mathcal{Y}$ ist eine Relation $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$, bei der zu jedem $x \in \mathcal{X}$ *genau* ein $y \in \mathcal{Y}$ mit xRy existiert. Für jedes $x \in \mathcal{X}$ wird das eindeutige y mit xRy als $f(x)$ bezeichnet.

Bemerkungen II.4.6.

- Beim praktischen Umgang mit Abbildungen ist die Präzisierung des Begriffs selten relevant. Sie zeigt aber, dass auch der Abbildungsbegriff mengentheoretisch unterfüttert und allein auf die Axiome der Mengenlehre gegründet werden kann. Manchmal hilft die Präzisierung auch beim Umgang mit Grenzfällen. Sie klärt etwa, dass von leerem Definitionsbereich $\emptyset$ in beliebiges Ziel $\mathcal{Y}$ genau eine Abbildung existiert, die leere Abbildung $\emptyset \to \mathcal{Y}$, die der Leer-Relation (die in diesem Fall zugleich die All-Relation ist) entspricht.
- Durch die Präzisierung des Abbildungsbegriffs wird die *Gleichheit $f_1 = f_2$ von Abbildungen $f_1 \in \mathrm{Abb}(\mathcal{X}_1, \mathcal{Y}_1)$ und $f_2 \in \mathrm{Abb}(\mathcal{X}_2, \mathcal{Y}_2)$* auf die Gleichheit von Tripeln und (Teil-)Mengen zurückgeführt und stellt sich als *gleichbedeutend mit $\mathcal{X}_1 = \mathcal{X}_2$, $\mathcal{Y}_1 = \mathcal{Y}_2$ und $f_1(x) = f_2(x)$ für alle $x \in \mathcal{X}_1$* heraus. Dies haben wir in Abschnitt II.1, dem Einschub zu Beweisstrategien und obigem Beispiel (7) teils schon benutzt, hatten dort aber eher nur Gleichheit von Abbildungen mit *a priori* gleichem Definitionsbereich und Ziel angesprochen.
- Man nennt eine Relation R zwischen Mengen $\mathcal{X}$ und $\mathcal{Y}$ linkstotal, wenn jedes $x \in \mathcal{X}$ mit *mindestens* einem $y \in \mathcal{Y}$ bezüglich R in Relation steht. Man nennt sie rechtseindeutig, wenn jedes $x \in \mathcal{X}$ mit *höchstens* einem $y \in \mathcal{Y}$ bezüglich R in Relation steht. Beide Bedingungen zusammen ergeben die obige Forderung, dass jedes x mit *genau* einem y in Relation steht, wir können also festhalten: *Eine Funktion $\mathcal{X} \to \mathcal{Y}$ ist nichts anderes als eine linkstotale und rechtseindeutige Relation zwischen $\mathcal{X}$ und $\mathcal{Y}$.*
- Für jede linkstotale Relation R zwischen Mengen $\mathcal{X}$ und $\mathcal{Y}$ gibt es mindestens eine Abbildung $f\colon \mathcal{X} \to \mathcal{Y}$ mit $\mathrm{G}_f \subset R$.

  Dies ergibt sich durch Anwendung des Auswahlaxioms aus Abschnitt I.5 auf das System der disjunkten nicht-leeren Mengen $R \cap (\{x\}\times\mathcal{Y})$ mit $x \in \mathcal{X}$ und Festlegung von $f(x)$ mit $x \in \mathcal{X}$ als y-Eintrag des ausgewählten Paars (x, y) aus $R \cap (\{x\}\times\mathcal{Y})$. Tatsächlich ist obige Aussage sogar äquivalent zum Auswahlaxiom, denn für ein System $\mathcal{S}$ disjunkter nicht-leerer Mengen kann man sie auf die linkstotale Relation $R \in \mathrm{Rel}(\mathcal{S}, \bigcup_{M\in\mathcal{S}} M)$ mit $R := \bigcup_{M\in\mathcal{S}}(\{M\}\times M)$ anwenden und erhält erst eine Abbildung $f\colon \mathcal{S} \to \bigcup_{M\in\mathcal{S}} M$ mit $\mathrm{G}_f \subset R$ und daraus dann die Auswahlmenge $A := \mathrm{Bild}(f) \subset \bigcup_{M\in\mathcal{S}} M$ mit $A \cap M = \{f(M)\}$ für jedes $M \in \mathcal{S}$. □

Als Nächstes führen wir Grundoperationen mit Relationen ein, die teils schon bekannte Operationen mit Abbildungen verallgemeinern:

Definition II.4.7. Es seien $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ beliebige Mengen.

- Die *Komposition von Relationen* $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ und $S \in \mathrm{Rel}(\mathcal{Y}, \mathcal{Z})$ ist die Relation $S \circ R := RS \in \mathrm{Rel}(\mathcal{X}, \mathcal{Z})$ definiert durch
$$x(RS)z :\iff \exists y \in \mathcal{Y} : (xRy \wedge ySz) \text{ für } x \in \mathcal{X}, z \in \mathcal{Z}.$$
- Die *Umkehrrelation* zu $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ ist die Relation $R^{-1} \in \mathrm{Rel}(\mathcal{Y}, \mathcal{X})$ mit
$$yR^{-1}x :\iff xRy \text{ für alle } x \in \mathcal{X}, y \in \mathcal{Y}.$$
- Die *komplementäre Relation* zu $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ ist die Relation $R^c \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ mit
$$xR^cy :\iff \neg(xRy) \text{ für alle } x \in \mathcal{X}, y \in \mathcal{Y}.$$
  Dies bedeutet, dass die Teilmenge $R^c \subset \mathcal{X} \times \mathcal{Y}$ das Komplement von R in $\mathcal{X}\times\mathcal{Y}$ ist.

Bemerkungen II.4.8. Es seien $\mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ Mengen.

- Die Komposition von Relationen verallgemeinert die Komposition von Abbildungen.
- Die Umkehrrelation verallgemeinert die Umkehrfunktion und existiert immer. Im Gegensatz zu Funktionen ist also für Relationen die Umkehrbarkeit stets gegeben.
- Die Komplement-Bildung hat bei Abbildungen kein Analogon.
- Die Komposition ist assoziativ: Für $R \in \mathrm{Rel}(\mathcal{W}, \mathcal{X})$, $S \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$, $T \in \mathrm{Rel}(\mathcal{Y}, \mathcal{Z})$ gilt
$$(RS)T = R(ST).$$
- Die Umkehrung und die Komplement-Bildung sind involutorisch: Für $R \in \mathrm{Rel}(\mathcal{X}, \mathcal{Y})$ gilt
$$(R^{-1})^{-1} = R = (R^c)^c.$$

Beispiele II.4.9.

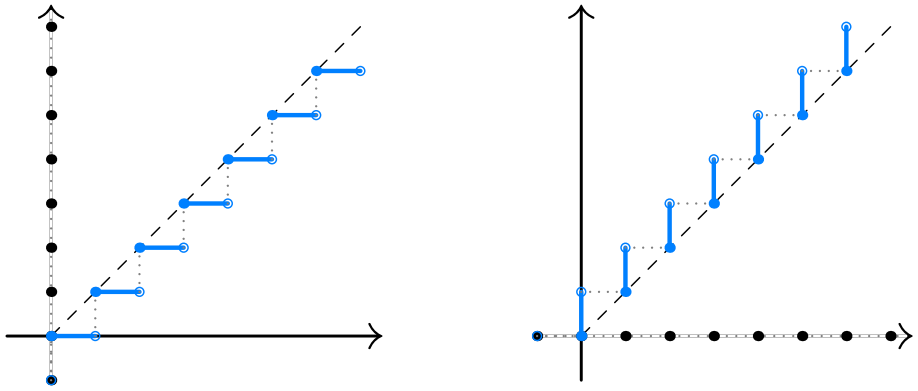
(a) Für die Relationen des früheren Beispiels aus II.4.4 gelten $(\sim_g)^{-1} = \sim_g$, $(\sim_u)^{-1} = \sim_u$, $(\sim_g)^c = \sim_u$ und $(\sim_u)^c = \sim_g$.

(b) Für die Relationen $\subset, \subsetneq, \in, <, \leq$ sind $\supset, \supsetneq, \ni, >, \geq$ (sofern definiert zwischen Elementen, Mengen, Zahlen, Abbildungen) die Umkehrungen. Komplementär zu $\subset, \in, <, \leq$ sind $\not\subset, \notin, \geq, >$ (sofern definiert bei Elementen, Mengen, Zahlen).

(c) Die *nächstkleinere ganze Zahl* zu einer reellen Zahl $x \in \mathbb{R}$, also die eindeutige Zahl $z \in \mathbb{Z}$ mit $z \leq x < z{+}1$, schreibt man mit der sogenannten *Gauß-Klammer* als $\lfloor x \rfloor \in \mathbb{Z}$. Die *Abrunden-Funktion* $A\colon \mathbb{R} \to \mathbb{Z}$, $x \mapsto \lfloor x \rfloor$ ist surjektiv, aber nicht injektiv und daher nicht bijektiv. Als Funktion ist sie daher *nicht* umkehrbar.

Die Umkehr*relation* existiert aber immer und ist in diesem Fall die Relation $A^{-1} \in \mathrm{Rel}(\mathbb{Z}, \mathbb{R})$ mit $A^{-1} = \{(z, r) \in \mathbb{Z}{\times}\mathbb{R} \mid z \leq r < z{+}1\}$ (siehe Abbildung II.4.4).

(II.4.4)



Die beiden wirklich wichtigen Klassen von Relationen werden nun in Kürze über die Gültigkeit (einiger) der folgenden Eigenschaften definiert.

Definition II.4.10. Eine Relation R auf einer Menge $\mathcal{X}$ heißt

- *reflexiv*, wenn xRx für alle $x \in \mathcal{X}$ gilt,
- *symmetrisch*, wenn für alle $x, y \in \mathcal{X}$ gilt:
$$xRy \implies yRx,$$
- *asymmetrisch*, wenn es *kein* Paar $(x, y) \in \mathcal{X}^2$ mit xRy und yRx gibt,
- *antisymmetrisch*, wenn für alle $x, y \in \mathcal{X}$ gilt:
$$(xRy \text{ und } yRx) \implies x = y,$$
- *transitiv*, wenn für alle $x, y, z \in \mathcal{X}$ gilt:
$$(xRy \text{ und } yRz) \implies xRz.$$

Bemerkungen II.4.11. Für $R \in \mathrm{Rel}(\mathcal{X})$ gilt:

- Ist R symmetrisch, so gilt für alle $x, y \in \mathcal{X}$ automatisch auch: $xRy \iff yRx$. Deshalb ist R genau dann symmetrisch, wenn $R^{-1} = R$ gilt.
- Asymmetrie und Symmetrie von R schließen einander aus (außer wenn $R = \emptyset$).
- Asymmetrie und Reflexivität von R schließen einander aus (außer wenn $\mathcal{X} = \emptyset$). Antisymmetrie kann man als schwächere Form von Asymmetrie sehen, die Reflexivität noch erlaubt.
- Alle fünf gerade definierten Eigenschaften übertragen sich von R auf R^{-1} (und wegen $\left(R^{-1}\right)^{-1} = R$ natürlich auch von R^{-1} auf R).

II.4.1. Ordnungsrelationen. Wir können nun eine wichtige Klasse von Relationen definieren und diskutieren:

Definition II.4.12.

(I) Eine *Ordnungsrelation*, *partielle Ordnung* oder *Halbordnung* auf einer Menge $\mathcal{X}$ ist eine *reflexive*, *antisymmetrische* und *transitive* Relation auf $\mathcal{X}$.

(II) Eine *strikte Ordnungsrelation*, *strikte partielle Ordnung* oder *strikte Halbordnung* auf einer Menge $\mathcal{X}$ ist eine *asymmetrische* und *transitive* Relation auf $\mathcal{X}$.

Bemerkungen II.4.13.

- Vorsicht! Eine strikte Ordnungsrelation ist nicht etwa eine Ordnungsrelation mit Zusatzeigenschaft. Vielmehr kann $R \in \mathrm{Rel}(\mathcal{X})$ auf $\mathcal{X} \neq \emptyset$ *nie* zugleich Ordnungsrelation und strikte Ordnungsrelation sein, weil Reflexivität und Asymmetrie einander ausschließen.
- Es besteht aber eine 1-zu-1-Korrespondenz zwischen Ordnungsrelationen und strikten Ordnungsrelationen durch folgende zueinander inverse Operationen: Zu jeder Ordnungsrelation $\trianglelefteq$ auf $\mathcal{X}$ gehört eine strikte Ordnungsrelation $\triangleleft$ auf $\mathcal{X}$, indem wir setzen: $x \triangleleft y :\iff (x \trianglelefteq y \wedge x \neq y)$ für $x, y \in \mathcal{X}$. Umgekehrt gehört zu jeder strikten Ordnungsrelation $\triangleleft$ auf $\mathcal{X}$ eine Ordnungsrelation $\trianglelefteq$ auf $\mathcal{X}$ mit $x \trianglelefteq y :\iff (x \triangleleft y \vee x = y)$ für $x, y \in \mathcal{X}$.
- Da sich alle relevanten Eigenschaften übertragen, ist die Umkehrrelation einer (strikten) Ordnungsrelation wieder eine (strikte) Ordnungsrelation.

Beispiele II.4.14. In den früheren Beispielen gilt:

(a) Die Relationen $=, \subset, \supset, \leq, \geq$ sind Ordnungsrelationen, aber (außer auf leerer Grundmenge) keine strikten Ordnungsrelationen.
(b) Die Relationen $\subsetneq, \supsetneq, <, >$ sind strikte Ordnungsrelationen, aber (außer auf leerer Grundmenge) keine Ordnungsrelationen.
(c) Die Ungleichheitsrelation auf Mengen $\neq$ ist symmetrisch, ist aber auf einer Grundmenge mit mindestens zwei Elementen weder reflexiv noch asymmetrisch noch antisymmetrisch noch transitiv und damit weder Ordnungsrelation noch strikte Ordnungsrelation.
(d) Die Teilbarkeitsrelationen „$|$“ auf $\mathbb{N}$ und $\mathbb{N}_0$ sind Ordnungsrelationen, aber keine strikten Ordnungsrelationen. Die Teilbarkeitsrelation „$|$“ auf $\mathbb{Z}$ dagegen ist zwar reflexiv und transitiv, aber weder asymmetrisch noch antisymmetrisch (wie man zum Beispiel an $(-1)|1$ und $1|(-1)$ sieht) und damit weder Ordnungsrelation noch strikte Ordnungsrelation.

Definition II.4.15. Es sei $\mathcal{X}$ eine Menge.

(I) Eine Ordnungsrelation $\trianglelefteq$ auf $\mathcal{X}$ heißt eine *Totalordnung*, *totale Ordnung* oder *lineare Ordnung* auf einer Teilmenge T von $\mathcal{X}$, wenn $(x \trianglelefteq y) \vee (y \trianglelefteq x)$ für alle $x, y \in T$ gilt. Wir nennen T dann eine *total* oder *linear geordnete Teilmenge* von $\mathcal{X}$ oder eine *Kette* in $\mathcal{X}$.
(II) Wir verwenden dieselben Begriffe für eine strikte Ordnungsrelation $\triangleleft$ auf $\mathcal{X}$, wenn sie für die „nichtstrikte“ Ordnungsrelation $\trianglelefteq$ auf $\mathcal{X}$ mit $x \trianglelefteq y :\iff ((x \triangleleft y) \vee (x = y))$ für alle $x, y \in \mathcal{X}$ erfüllt sind. Speziell ist $\triangleleft$ genau dann eine *strikte Totalordnung*, wenn für alle $x, y \in T$ eine (und dann automatisch genau eine) der drei Aussagen $x \triangleleft y$, $y \triangleleft x$, $x = y$ gilt.

Bemerkungen II.4.16. Es Sei $\mathcal{X}$ eine Menge.

- Bei einer Totalordnung $\trianglelefteq$ auf $\mathcal{X}$ können zwei Elemente $x, y \in \mathcal{X}$ stets auf irgendeine Weise verglichen werden: Es gilt stets $x \trianglelefteq y$ oder $y \trianglelefteq x$. Bei einer allgemeinen Ordnungsrelation dagegen kann es passieren (vergleiche die folgenden Beispiele), dass für zwei Elemente $x, y \in \mathcal{X}$ weder $x \trianglelefteq y$ noch $y \trianglelefteq x$ gilt, also überhaupt kein Vergleich zwischen x und y gezogen werden kann. Diese Möglichkeit, dass man in manchen Fällen eben gar nicht vergleichen kann, ist der Grund, warum man auch von *nur partiellen* Ordnungen oder *Halb*ordnungen spricht. Analog verhält es sich natürlich bei strikten Ordnungsrelationen.
- Die Umkehrrelation einer (strikten) Totalordnung ist eine (strikte) Totalordnung. Dies folgt nach vorigen Bemerkungen direkt aus der Totalordnungseigenschaft.
- In den Übungen zeigen Sie für Komplemente: $\trianglelefteq \in \mathrm{Rel}(\mathcal{X})$ ist genau dann eine Totalordnung auf $\mathcal{X}$, wenn $\trianglelefteq^c$ eine strikte Totalordnung auf $\mathcal{X}$ ist. Umgekehrt damit auch: $\triangleleft \in \mathrm{Rel}(\mathcal{X})$ ist genau dann eine strikte Totalordnung auf $\mathcal{X}$, wenn $\triangleleft^c$ eine Totalordnung auf $\mathcal{X}$ ist.

Beispiele II.4.17.

(a) Die Relationen $\leq, \geq$ sind Totalordnungen und $<, >$ strikte Totalordnungen auf den reellen Zahlen $\mathbb{R}$, insbesondere auch auf $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$ und jeder Teilmenge $T \subset \mathbb{R}$.
(b) Zwischen Paaren, Tripeln oder Tupeln von Zahlen oder mit anderen Worten auf $\mathbb{B}^n$ mit $\mathbb{B} \in \{\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ und $n \in \mathbb{N} \setminus \{1\}$ gibt es mehrere Möglichkeiten, Ordnungsrelationen zu definieren:

Für $n = 2$ kann man zum Beispiel Ordnungsrelationen $\leq_{\text{komp}}$ und $\leqq_{\text{komp}}$ auf $\mathbb{B}^2$ komponentenweise durch

$$x \leq_{\text{komp}} y :\Longleftrightarrow (x_1 \leq y_1, x_2 \leq y_2) \qquad \text{für } x, y \in \mathbb{B}^2,$$
$$x \leqq_{\text{komp}} y :\Longleftrightarrow ((x_1 < y_1, x_2 < y_2) \vee (x_1 = y_1, x_2 = y_2) \qquad \text{für } x, y \in \mathbb{B}^2$$

erklären, aber $\leq_{\text{komp}}$ und $\leqq_{\text{komp}}$ sind *keine* Totalordnungen auf $\mathbb{B}^2$. Trotzdem gibt es für diese Relationen aber Ketten in $\mathbb{B}^2$ wie $\{(3,2),(6,2),(7,4),(8,9)\}$ (nur für $\leq_{\text{komp}}$) und $\{x \in \mathbb{B}^2 \mid x_1 = x_2\}$ (für $\leq_{\text{komp}}$ und $\leqq_{\text{komp}}$), bei denen eine (strikte) Ordnung der Elemente bezüglich beider Einträge zugleich vorliegt oder vorgenommen werden kann.

Alternativ erhält man durch die sogenannte *lexikographische Ordnung* (die sich an die übliche Sortierung von Worten zunächst nach dem ersten Buchstaben, dann nach dem zweiten, dem dritten und den folgenden anlehnt)

$$x \leq_{\text{lex}} y :\Longleftrightarrow ((x_1 < y_1) \vee (x_1 = y_1, x_2 \leq y_2)) \text{ für } x, y \in \mathbb{B}^2$$

tatsächlich eine Totalordnung $\leq_{\text{lex}}$ auf $\mathbb{B}^2$.

Selbstverständlich kann man bei diesen Bildungen auch die Umkehrrelationen, die zugehörigen strikten Ordnungsrelationen und Verallgemeinerungen auf $\mathbb{B}^n$ mit beliebigem $n \in \mathbb{N}\backslash\{1\}$ betrachten.

(c) Die Relationen $\leq$, $\geq$, $<$, $>$ auf $\text{Abb}(\mathcal{X}, T)$ mit $T \subset \mathbb{R}$, also zwischen Abbildungen, sind dagegen keine *(strikten) Totalordnungen*, jedenfalls sofern $\mathcal{X}$ und T mindestens zwei Elemente haben: Sind nämlich $f, g \in \text{Abb}(\mathcal{X}, T)$ mit $f(x) < g(x)$ für ein $x \in \mathcal{X}$ und $g(\widetilde{x}) < f(\widetilde{x})$ für ein anderes $\widetilde{x} \in \mathcal{X}$, so gilt weder $f \leq g$ noch $g \leq f$ (bzw. weder $f < g$ noch $g < f$ noch $f = g$). Trotzdem gibt es für diese Relationen aber Ketten in $\text{Abb}(\mathcal{X}, T)$, etwa die Teilmenge aller konstanten Abbildungen $\mathcal{X} \to T$.

(d) Die Relationen $\subset$, $\supset$, $\subsetneqq$, $\supsetneqq$ sind *keine* (strikten) Totalordnungen auf $\mathcal{P}(\mathcal{X})$, sofern die Grundmenge $\mathcal{X}$ zwei verschiedene Elemente $x \neq \widetilde{x}$ enthält, denn für $\{x\}, \{\widetilde{x}\} \in \mathcal{P}(\mathcal{X})$ gilt dann weder $\{x\} \subset \{\widetilde{x}\}$ noch $\{\widetilde{x}\} \subset \{x\}$ (beziehungsweise weder $\{x\} \subsetneqq \{\widetilde{x}\}$ noch $\{\widetilde{x}\} \subsetneqq \{x\}$ noch $\{x\} = \{\widetilde{x}\}$. Es gibt aber für diese Relationen (viele) Ketten in $\mathcal{P}(\mathcal{X})$, z.B. ist für $\mathcal{X} = \mathbb{N}$ das Mengensystem

$$\{\emptyset, \{-3\}, \{-3,5\}, \{-3,5,0\}, \{-3,5,0,8,-2\}, \{-3,5,0,8,-2,4\}, \{-3,5,0,8,-2,4,1,2,3\}\}$$

ein Beispiel einer Kette in $\mathcal{P}(\mathbb{N})$. Es gibt in $\mathcal{P}(\mathbb{N})$ neben solchen endlichen auch unendliche Ketten von analoger Natur.

Definition II.4.18. Es sei $\trianglelefteq$ eine Ordnungsrelation auf einer Menge $\mathcal{X}$ und T sei eine Teilmenge von $\mathcal{X}$.

(I) Wir nennen $s \in \mathcal{X}$ eine *obere Schranke* (bzw. *untere Schranke*) für T in $\mathcal{X}$, wenn $x \trianglelefteq s$ (bzw. $s \trianglelefteq x$) für alle $x \in T$ gilt. Ist eine obere Schranke (bzw. untere Schranke) für T selbst Element von T, so heißt sie ein *größtes Element* (bzw. ein *kleinstes Element*) von T.

(II) Wir nennen $m \in T$ ein *maximales Element* (bzw. *minimales Element*) von T, wenn für jedes $x \in T$ mit $m \trianglelefteq x$ (bzw. $x \trianglelefteq m$) schon $x = m$ gilt.

(III) Ein $s_* \in \mathcal{X}$ heißt *kleinste obere* (*größte untere*) Schranke für T, wenn s_* selbst obere (untere) Schranke für T ist und $s_* \trianglelefteq s$ (bzw. $s \trianglelefteq s_*$) für alle oberen (unteren) Schranken $s \in \mathcal{X}$ für T erfüllt).

Lemma II.4.19. *Es sei $\trianglelefteq$ eine Ordnungsrelation auf $\mathcal{X}$ und $T \subset \mathcal{X}$.*

Falls ein größtes (oder kleinstes) Element von T existiert, ist dieses immer eindeutig und ist auch das eindeutige maximale (oder minimale) Element von T und die kleinste obere (oder größte untere Schranke) für T

Beweis. Es existiere ein größtes Element $g \in T$ von T. Wir begründen, dass

- dieses größte Element eindeutig ist: Ist auch $\widetilde{g} \in T$ ein größtes Element von T, so gilt sowohl $g \trianglelefteq \widetilde{g}$ als auch $\widetilde{g} \trianglelefteq g$, und die Antisymmetrie von $\trianglelefteq$ gibt $\widetilde{g} = g$.
- g die kleinste obere Schranke für T ist: Nach Definition ist g obere Schranke für T und $g \trianglelefteq s$ für jede obere Schranke $s \in \mathcal{X}$ für T. Also ist g die kleinste obere Schranke für T.
- g das eindeutige maximale Element von T ist: Da jedes $x \in T$ mit $g \trianglelefteq x$ zusätzlich $x \trianglelefteq g$ und dann per Antisymmetrie auch $x = g$ erfüllt, ist g ein maximales Element von T. Ist auch $m \in T$ ein maximales Element, so gilt mit $m \trianglelefteq g$ sofort auch $m = g$. Daher ist g tatsächlich das eindeutige maximale Element von T.

Existiert stattdessen ein kleinstes Element von T, so kann man analog argumentieren oder durch Übergang zu $\trianglelefteq^{-1}$ auf das Vorige reduzieren. □

Lemma II.4.20. *Für total geordnetes T sind maximale (minimale) Elemente von T dasselbe wie größte (kleinste) Elemente von T. Insbesondere greift Lemma II.4.19 dann auch für ein maximales (minimales) Element.*

Beweis. Gemäß Lemma II.4.19 ist ein größtes Element von T stets auch ein maximales Element von T. Wir zeigen, dass umgekehrt ein maximales Element $m \in T$ von T stets auch ein größtes Element von T ist: Sei dazu $x \in T$ beliebig. Da T total geordnet ist, gilt entweder $x \trianglelefteq m$ oder $m \trianglelefteq x$. Im zweiten Fall folgt $x = m$ per Maximalität von m und damit $x \trianglelefteq m$ gemäß der Reflexivität von $\trianglelefteq$. Also gilt tatsächlich $x \trianglelefteq m$ für jedes $x \in T$, und m ist ein größtes Element von T.

Analog oder durch Übergang zu $\trianglelefteq^{-1}$ behandelt man kleinste und minimale Elemente. □

Beispiele II.4.21.

- Bezüglich der Totalordnung $\leq$ auf $\mathbb{R}$ ist das größte/kleinste Element einer Teilmenge im üblichen Sinn zu verstehen, zum Beispiel ist 5 das größte Element von $\{x \in \mathbb{R} \mid x<0\} \sqcup \{5\} \sqcup \{2\}$. Es gibt Teilmengen ohne größtes Element, aber mit oberer Schranke, etwa $\{x \in \mathbb{R} \mid x<0\}$, und auch Teilmengen ohne obere Schranke, zum Beispiel $\mathbb{N}$.
- Eine größtes/kleinstes Element bezüglich der Totalordnung $\geq$ ist ein kleinstes/größtes Element im herkömmlichen Sinn. Um Verwirrung zu vermeiden, wendet man die obigen Begriffe in diesem Wortlaut daher nur auf $\leq$ und ähnliche, „nach oben gerichtete“ Relationen an.
- Wir betrachten $T := \{(1,1),(1,2),(1,3),(1,4),(2,2)\} \subset \mathbb{N}^2$ bezüglich der Ordnungsrelationen $\leq_{\text{komp}}$, $\leqq_{\text{komp}}$, $\leq_{\text{lex}}$ (siehe Beispiele II.4.17). Dieses T hat bezüglich $\leq_{\text{komp}}$ genau $(1,4)$ und $(2,2)$ als maximale Elemente und $(2,4)$ als kleinste obere Schranke, hat bezüglich $\leqq_{\text{komp}}$ genau $(1,2)$, $(1,3)$, $(1,4)$ und $(2,2)$ als maximale Elemente und $(3,5)$ als kleinste obere Schranke und hat bezüglich der Totalordnung $\leq_{\text{lex}}$ das größte Element $(2,2)$.
- Bezüglich der Ordnungsrelation $\subset$ hat die Teilmenge $T := \{\{1\},\{1,2\},\{1,2,3\},\{2,3\},\{2,4\}\}$ von $\mathcal{P}(\{1,2,3,4,5\})$ genau $\{1,2,3\}$ und $\{2,4\}$ als maximale Elemente. Obere Schranken für T sind genau $\{1,2,3,4\}$ und $\{1,2,3,4,5\}$. Ein größtes Element von T gibt es nicht.

Es folgt ein allgemeines Resultat über Ordnungsrelationen, dessen Bedeutung sich nicht unbedingt auf den ersten Blick erschließt, das aber später wichtige Anwendungen hat:

Lemma II.4.22 (Zornsches Lemma)**.** *Es sei $\trianglelefteq$ eine Ordnungsrelation auf einer Menge $\mathcal{X}$. Wenn für jede Kette in $\mathcal{X}$ eine obere Schranke in $\mathcal{X}$ existiert, dann gibt es ein maximales Element von $\mathcal{X}$.*

Das Zornsche Lemma ist äquivalent zum Auswahlaxiom der Mengenlehre und sein Beweis kann entweder mit Hilfe sogenannter Ordinalzahlen oder elementar geführt werden. Beides geht über den Vorlesungsstoff hinaus, weshalb wir die elementare Argumentation nur im Anhang als angeben (siehe Abschnitt V.3).

Als Nächstes kommen wir zu Ordnungseigenschaften der natürlichen und ganzen Zahlen bezüglich der Totalordnung $\leq$ und der zugehörigen strikten Totalordnung $<$. Hier benutzen wir die übliche Ordnungsrelation, dass $n < S(n)$ für alle $n \in \mathbb{N}_0$ ist und dass $S^{-1}(n) < n$ ist für alle $n \in \mathbb{Z}$. Verwenden wir „kleinste Zahl“ für „kleinstes Element“, so gilt:

Lemma II.4.23. *In $\mathbb{N}$ ist 1 die kleinste Zahl und in $\mathbb{N}_0$ ist 0 die kleinste Zahle. Eine größte Zahl gibt es weder in $\mathbb{N}$ noch in $\mathbb{N}_0$. In $\mathbb{Z}$ gibt es weder eine kleinste noch eine größte Zahl.*

Diese einleuchtenden Aussagen lassen sich auch auf Grundlage der bisherigen Definitionen verifizieren: Zum Beispiel übersetzt man die Aussage, dass 1 die kleinste Zahl in $\mathbb{N}$ ist, mit den Definitionen des kleinsten Elements und der Relation $\leq$ auf $\mathbb{Z}$ in die zu ihr äquivalenten Aussagen $\forall n \in \mathbb{N}\colon 1 \leq n$ und $\forall n \in \mathbb{N}\colon n-1 \in \mathbb{N}_0$. Letztere erkennt man dann aufgrund des früher zu Nachfolgern und Vorgängern ganzer Zahlen Gesagten als richtig. Dass es keine größte Zahl in $\mathbb{N}$ gibt, liegt natürlich einfach an $n+1 > n$ für alle $n \in \mathbb{N}$. Mit denselben Argumenten bestätigt man die anderen obigen Aussagen sowie für jedes $z \in \mathbb{Z}$, dass

(∗∗) $z+1$ die kleinste Zahl in $\{y \in \mathbb{Z} \mid z < y\}$ und $z-1$ die größte Zahl in $\{y \in \mathbb{Z} \mid y < z\}$

ist.

Definition II.4.24. Eine Ordnungsrelation auf einer Menge $\mathcal{X}$ heißt eine *Wohlordnung* auf $\mathcal{X}$, wenn sie eine totale Ordnung auf ganz $\mathcal{X}$ ist und bezüglich ihr jede nicht-leere Teilmenge von $\mathcal{X}$ ein kleinstes Element besitzt.

Proposition II.4.25. In jeder nicht-leeren Teilmenge von $\mathbb{Z}$, die eine obere Schranke in $\mathbb{Z}$ besitzt, existiert eine größte Zahl, und in jeder nicht-leeren Teilmenge von $\mathbb{Z}$, die eine untere Schranke in $\mathbb{Z}$ besitzt, existiert eine kleinste Zahl. Insbesondere haben $\mathbb{N}$ und $\mathbb{N}_0$ die Wohlordnungseigenschaft, gemäß der in jeder ihrer nicht-leeren Teilmengen eine kleinste Zahl existiert.

Beweis. Wir zeigen erst die Existenz kleinster Zahlen in T für $\emptyset \neq T \subset \mathbb{Z}$ mit unterer Schranke z_0 für T in $\mathbb{Z}$. Dazu argumentieren wir indirekt: Angenommen, es gibt *keine* kleinste Zahl in T. Dann zeigen wir durch Induktion, dass jedes $z \in \{z_0, z_0{+}1, z_0{+}2, \ldots\}$ untere Schranke für T ist. Der Induktionsanfang für $z = z_0$ ist per Voraussetzung gegeben. Für den Induktionsschritt sei $z \in \mathbb{Z}$ untere Schranke für T, also $T \subset \{y \in \mathbb{N} \mid z \leq y\}$. Zudem ist aber $z \notin T$ (denn sonst wäre z kleinste Zahl in T) und damit sogar $T \subset \{y \in \mathbb{N} \mid z < y\} = \{y \in \mathbb{N} \mid z{+}1 \leq y\}$, wobei die Gleichheit aus $(**)$ resultiert. Dies bedeutet, dass $z{+}1$ untere Schranke für T und der Induktionsschritt komplett ist. Insgesamt sind dann aber alle unteren Schranken $z_0, z_0{+}1, z_0{+}2, \ldots \notin T$, da man sonst eine kleinste Zahl in T bekäme. Dies steht im Widerspruch zu $T \neq \emptyset$ und beweist insgesamt die Existenz der kleinsten Zahl in T.

Die Existenz der größten Zahl in T für $\emptyset \neq T \subset \mathbb{Z}$ mit oberer Schranke ergibt sich analog oder durch Anwendung des Vorigen auf $\{-z \mid z \in \mathbb{Z}\}$.

Die Wohlordnungseigenschaft von $\mathbb{N}$ und $\mathbb{N}_0$ folgt, da dort gemäß Lemma II.4.23 1 beziehungsweise 0 kleinstes Element und damit untere Schranke für jede Teilmenge ist. □

Gemäß der vorigen Proposition ist die Standard-Ordnung „$\leq$“ eine Wohlordnung auf $\mathbb{N}$ und $\mathbb{N}_0$. Auf $\mathbb{Z}$ ist „$\leq$“ zwar keine Wohlordnung, man kann auf $\mathbb{Z}$ und jeder Menge, die in Bijektion zu $\mathbb{N}$ steht, aber ohne Probleme eine Wohlordnung erzeugen. Auf $\mathbb{Z}$ sieht eine mögliche Wohlordnung $\trianglelefteq$ zum Beispiel so aus, dass

$$0 \trianglelefteq 1 \trianglelefteq -1 \trianglelefteq 2 \trianglelefteq -2 \trianglelefteq 3 \trianglelefteq -3 \trianglelefteq 4 \trianglelefteq -4 \trianglelefteq \ldots$$

gilt. Auf $\mathbb{R}$ und anderen „großen“ Mengen dagegen kann man eine Wohlordnung nicht konstruktiv erhalten. Ihre pure Existenz ist dennoch sichergestellt durch:

Satz II.4.26. *Für jede Menge $\mathcal{X}$ gibt es eine Wohlordnung auf $\mathcal{X}$.*

Den Beweis finden Sie in Abschnitt V.4.

II.4.2. Äquivalenzrelationen. Die zweite wichtige Klasse spezieller Relationen ist folgende:

Definition II.4.27. Eine *Äquivalenzrelation* auf einer Menge $\mathcal{X}$ ist eine *reflexive*, *symmetrische* und *transitive* Relation auf $\mathcal{X}$.

Bemerkung II.4.28. Für die Umkehrrelation und Selbst-Komposition einer Äquivalenzrelation $\sim$ gelten stets $\sim^{-1} = \sim$ (folgt aus Symmetrie) und $\sim\sim = \sim$ (folgt aus Reflexivität und Transitivität).

Wir werden in Kürze konkrete Beispiele von Äquivalenzrelationen diskutieren, beschäftigen uns aber zuvor mit der entscheidenden abstrakten Eigenschaft, dass sogenannte Äquivalenzklassen gebildet werden können und die nützlichen Eigenschaften des nächsten Satzes aufweisen:

Definition II.4.29. Es sei $\sim$ eine Äquivalenzrelation auf einer Menge $\mathcal{X}$.

(I) Die *Äquivalenzklasse* von $x \in \mathcal{X}$ bezüglich $\sim$ ist

$$[x]_\sim := \{y \in \mathcal{X} \mid y \sim x\} \subset \mathcal{X}.$$

Ist die betrachtete Äquivalenzrelation im Kontext klar, so notieren wir auch $[x]$ für $[x]_\sim$.

(II) Die *Quotientenmenge* $\mathcal{X}/\sim$ (lies: $\mathcal{X}$ modulo $\sim$) von $\mathcal{X}$ bezüglich der Äquivalenzrelation $\sim$ ist die *Menge aller Äquivalenzklassen* von $\sim$, also

$$\mathcal{X}/\sim := \{[x]_\sim \mid x \in \mathcal{X}\} \subset \mathcal{P}(\mathcal{X}).$$

(III) Die *kanonische Projektion* oder *Quotientenabbildung* von $\mathcal{X}$ nach $\mathcal{X}/\!\sim$ ist die stets surjektive Abbildung

$$\pi_\sim \colon \mathcal{X} \to \mathcal{X}/\!\sim, \quad x \mapsto [x]_\sim.$$

Ergibt sich die Äquivalenzrelation aus dem Kontext, so schreiben wir auch π für $\pi_\sim$.

Satz II.4.30. *Sei $\mathcal{X}$ eine Menge mit $x, y \in \mathcal{X}$. Für eine Äquivalenzrelation $\sim$ auf $\mathcal{X}$ gilt*

$$y \sim x \iff y \in [x]_\sim \iff [y]_\sim = [x]_\sim \iff [y]_\sim \cap [x]_\sim \neq \emptyset$$

und für die komplementäre Relation $\not\sim := \sim^c$ dementsprechend

$$y \not\sim x \iff y \notin [x]_\sim \iff [y]_\sim \neq [x]_\sim \iff [y]_\sim \cap [x]_\sim = \emptyset.$$

Beweis. Wir zeigen nur die obere Zeile von Äquivalenzen, weil sich die untere durch Negation daraus ergibt. Da $y \sim x \iff y \in [x]_\sim$ per Definition der Äquivalenzklasse $[x]_\sim$ gilt, werden wir tatsächlich nur als Ringschluss die drei Implikationen in

$$y \sim x \implies [y]_\sim = [x]_\sim \implies [y]_\sim \cap [x]_\sim \neq \emptyset \implies y \sim x$$

nachweisen:

- Erste Implikation: Es gelte $y \sim x$. Wir verifizieren $[y]_\sim \subset [x]_\sim$ und $[y]_\sim \supset [x]_\sim$ separat. Für „$\subset$“ sei $z \in [y]_\sim$, also $z \sim y$. Zusammen mit $y \sim x$ und Transitivität von $\sim$ folgt $z \sim x$, also wie benötigt $z \in [x]_\sim$. Für „$\supset$“ bemerken wir, dass per Symmetrie von $\sim$ auch $x \sim y$ gilt, und greifen dann auf „$\subset$“ mit vertauschten Rollen von x und y zurück.
- Zweite Implikation: Hierfür ist nur $[x]_\sim \neq \emptyset$ sicherzustellen. Dies ist aber gegeben, weil Reflexivität $x \sim x$ und damit $x \in [x]_\sim$ garantiert.
- Dritte Implikation: Es sei $z \in [y]_\sim \cap [x]_\sim$, also $z \sim x$ und $z \sim y$. Per Symmetrie gilt auch $y \sim z$. Mit Transitivität folgt aus $y \sim z$ und $z \sim x$ dann $y \sim x$. □

Bemerkungen II.4.31. Insbesondere ergibt sich aus dem Satz:

- Jede Äquivalenzklasse $[x]_\sim$ kann auch als $[y]_\sim$ mit beliebigem $y \in [x]_\sim$ geschrieben werden. Man hat bei der Darstellung einer Äquivalenzklasse durch ein Element und die eckigen Klammern in der Regel also eine Wahl, welches Element man nennt. Man sagt daher häufig, dass ein $y \in [x]_\sim$ (wie auch x selbst) ein *Repräsentant der Äquivalenzklasse* $[x]_\sim$ ist.
- Verschiedene Äquivalenzklassen sind stets disjunkt, also ist die Quotientenmenge $\mathcal{X}/\!\sim$ ein System disjunkter Mengen. Da jedes Element $x \in \mathcal{X}$ in einer Äquivalenzklasse liegt (nämlich in $x \in [x]_\sim$), bedeutet dies insgesamt

$$\mathcal{X} = \bigsqcup(\mathcal{X}/\!\sim) = \bigsqcup\{[x]_\sim \mid x \in \mathcal{X}\}.$$

Mit anderen Worten bringt eine Äquivalenzrelation stets eine *Partition*, also eine disjunkte Zerlegung der Grundmenge in Äquivalenzklassen mit sich.

Auch umgekehrt erhält man aus einer Partition $\mathcal{X} = \bigsqcup\{M \mid M \in \mathcal{S}\}$ einer Menge $\mathcal{X}$ mittels eines Systems $\mathcal{S}$ disjunkter Mengen, durch die Festlegung

$$x \sim_{\mathcal{S}} y :\iff \exists M \in \mathcal{S} \colon (x \in M \wedge y \in M) \text{ für } x, y \in \mathcal{X}$$

eine Äquivalenzrelation $\sim_{\mathcal{S}} \in \mathrm{Rel}(\mathcal{X})$.

Dabei ergeben sich als Äquivalenzklassen von $\sim_{\mathcal{S}}$ gerade die in $\mathcal{S}$ enthaltenen Mengen, und aus einer Partition in Äquivalenzklassen erhält man die Äquivalenzrelation zurück, es gelten also stets $\mathcal{X}/\!\sim_{\mathcal{S}} = \mathcal{S}$ und $\sim_{\mathcal{X}/\sim} = \sim$. Somit sind die beschriebenen Übergänge von Äquivalenzrelation zu Partition und von Partition zu Äquivalenzrelation zueinander invers, und man kann Äquivalenzrelationen auf $\mathcal{X}$ tatsächlich 1-zu-1 mit Partitionen von $\mathcal{X}$ identifizieren.

Beispiele II.4.32.

(a) Die All-Relation auf $\mathcal{X}$ ist eine (triviale) Äquivalenzrelation. Bei dieser ist $[x] = \mathcal{X}$ für alle $x \in \mathcal{X}$, ganz $\mathcal{X}$ ist also die einzige Äquivalenzklasse.

(b) Die Gleichheitsrelation „=“ ist eine Äquivalenzrelation auf M und kann als stark vereinfachter Prototyp einer solchen Relation angesehen werden. Bei $=$ haben alle Äquivalenzklassen $[x]_= = \{x\}$ mit $x \in M$ genau ein Element. Die Quotientenabbildung $M \to M/=$ ist bijektiv und erlaubt die kanonische Identifikation von $M/=$ mit M.

(c) Sehr zentrale Beispiele von Äquivalenzrelationen sind die *modulo-n-Relationen* $\sim_n$ auf $\mathbb{Z}$, die für jede feste Zahl $n \in \mathbb{Z}$ durch

$$y \sim_n x :\iff n \mid (y-x) \iff \exists z \in \mathbb{Z} : y-x = nz \text{ für } x, y \in \mathbb{Z}$$

erklärt werden (und deren Reflexivität, Symmetrie und Transitivität man leicht prüft). Typischerweise betrachtet man nur $n \in \mathbb{N} \setminus \{1\}$, da $\sim_{-n}=\sim_n$ gilt, $\sim_0$ die Gleichheitsrelation und $\sim_1$ die All-Relation ist. Die Relation $\sim_2$ ist die Relation $\sim_g$ des früheren Beispiels (siehe Beispiele II.4.4) (wohingegen die Relation $\sim_u$ von früher *keine* Äquivalenzrelation und hier irrelevant ist). Man verwendet bei den modulo-Relationen die allgemeinen Schreibweisen

$$y = x \mod n \text{ für } y \sim_n x$$

(gelesen „y ist gleich x modulo n“ oder „y ist kongruent zu x modulo n“) und

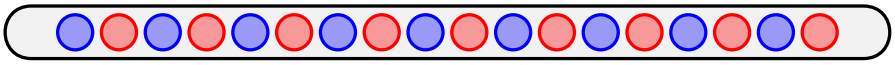
$$\mathbb{Z}/n\mathbb{Z} \text{ für } \mathbb{Z}/\sim_n .$$

Konkret gelten beispielsweise $25 = 7 \mod 9$ und $3 \cdot 6385 + 4 = -11 \mod 3$.

Die *Äquivalenzklassen modulo* 2 sind

$$\begin{aligned} [0]_{\mathbb{Z}/2\mathbb{Z}} &= \{\ldots, -6, -4, -2, 0, 2, 4, 6, \ldots\} &&= [2]_{\mathbb{Z}/2\mathbb{Z}} = [-2]_{\mathbb{Z}/2\mathbb{Z}}\,, \\ [1]_{\mathbb{Z}/2\mathbb{Z}} &= \{\ldots, -5, -3, -1, 1, 3, 5, 7, \ldots\} &&= [3]_{\mathbb{Z}/2\mathbb{Z}} = [-1]_{\mathbb{Z}/2\mathbb{Z}} \end{aligned}$$

(II.4.5)



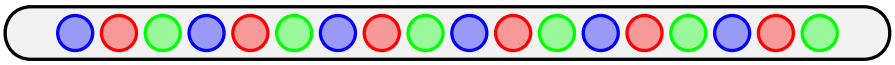
und geben die in Abbildung II.4.5 gezeigte Zerlegung $\mathbb{Z} = [0]_{\mathbb{Z}/2\mathbb{Z}} \sqcup [1]_{\mathbb{Z}/2\mathbb{Z}}$ von $\mathbb{Z}$ in die *Mengen der geraden und ungeraden Zahlen.*

Die *Äquivalenzklassen modulo* 3 sind

$$\begin{aligned} [0]_{\mathbb{Z}/3\mathbb{Z}} &= \{\ldots, -9, -6, -3, 0, 3, 6, 9, \ldots\} &&= [3]_{\mathbb{Z}/3\mathbb{Z}} = [-3]_{\mathbb{Z}/3\mathbb{Z}}\,, \\ [1]_{\mathbb{Z}/3\mathbb{Z}} &= \{\ldots, -8, -5, -2, 1, 4, 7, 10, \ldots\} &&= [4]_{\mathbb{Z}/3\mathbb{Z}} = [-2]_{\mathbb{Z}/3\mathbb{Z}}\,, \\ [2]_{\mathbb{Z}/3\mathbb{Z}} &= \{\ldots, -7, -4, -1, 2, 5, 8, 11, \ldots\} &&= [5]_{\mathbb{Z}/3\mathbb{Z}} = [-1]_{\mathbb{Z}/3\mathbb{Z}} \end{aligned}$$

und geben die in Abbildung II.4.6 gezeigte Zerlegung $\mathbb{Z} = [0]_{\mathbb{Z}/3\mathbb{Z}} \sqcup [1]_{\mathbb{Z}/3\mathbb{Z}} \sqcup [2]_{\mathbb{Z}/3\mathbb{Z}}$ in drei Äquivalenzklassen mit je unendlich vielen Elementen, die bei Division durch 3 Rest 0, 1 bzw. 2 ergeben.

(II.4.6)



Analog ist $\mathbb{Z}$ in Äquivalenzklassen modulo $4, 5, 6, 7, 8, 9, \ldots$ zerlegt. Ein Beispiel für eine Äquivalenzklasse modulo 9 ist

$$[7]_{\mathbb{Z}/9\mathbb{Z}} = \{\ldots, -38, -29, -20, -11, -2, 7, 16, 25, 34, 43, 52, 61, 70, 79, \ldots\}.$$

(d) Ein Beispiel einer Äquivalenzrelation $\sim_T$ auf $\mathbb{N}$ erhält man durch

$$m \sim_T n :\iff m \text{ und } n \text{ besitzen gleich viele Teiler in } \mathbb{N} \text{ für } m, n \in \mathbb{N}.$$

Beispiele von Äquivalenzklassen bezüglich $\sim_T$ sind

$\{1\}$	(genau 1 Teiler),
$\{2,3,5,7,11,13,17,19,\ldots\}$	(genau 2 Teiler; Primzahlen $p \in \mathbb{P}$),
$\{4,9,25,49,121,169,\ldots\}$	(genau 3 Teiler; Zahlen p^2 mit $p \in \mathbb{P}$),
$\{6,8,10,14,15,21,22,\ldots\}$	(genau 4 Teiler; Zahlen p^3 oder pq mit $p \neq q$ in $\mathbb{P}$),
$\{16,81,625,2401,14641,\ldots\}$	(genau 5 Teiler; Zahlen p^4 mit $p \in \mathbb{P}$),
$\{12,18,20,28,32,44,45,\ldots\}$	(genau 6 Teiler; Zahlen p^5 oder p^2q mit $p \neq q$ in $\mathbb{P}$).

Insgesamt zerlegt die Äquivalenzrelation $\sim_T$ die natürlichen Zahlen in die Äquivalenzklasse $\{1\}$ mit einem Element und unendlich viele weitere Äquivalenzklassen mit jeweils unendlich vielen Elementen.

(e) Für jede Menge $\mathcal{X}$ wird durch

$$f \propto g :\Longleftrightarrow \exists \gamma \in \mathbb{R}\backslash\{0\}\colon \forall x \in \mathcal{X}\colon g(x) = \gamma f(x) \text{ für } f,g \in \mathrm{Abb}(\mathcal{X},\mathbb{R})$$

eine Äquivalenzrelation $\propto$ auf $\mathrm{Abb}(\mathcal{X},\mathbb{R})$ definiert. Im Fall $\mathcal{X} = \mathbb{R}$ enthält die Äquivalenzklasse von $\mathrm{id}_{\mathbb{R}}$ bezüglich $\propto$ alle linearen Funktionen $\mathbb{R} \to \mathbb{R}$, $x \mapsto \gamma x$ mit Steigungsparameter $\gamma \in \mathbb{R}\backslash\{0\}$, und die Äquivalenzklasse von $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ bezüglich $\propto$ enthält alle quadratischen Funktionen $\mathbb{R} \to \mathbb{R}$, $x \mapsto \gamma x$ mit (Öffnungs-)Parameter $\gamma \in \mathbb{R}\backslash\{0\}$.

(f) Bezüglich der Äquivalenzrelationen $\sim^{123}$ auf $\mathcal{P}(\mathbb{N})$ mit

$$M \sim^{123} N :\Longleftrightarrow M \cap \{1,2,3\} = N \cap \{1,2,3\} \text{ für } M,N \subset \mathbb{N}$$

enthält die Äquivalenzklasse der Menge der ungeraden Zahlen

$$[\{1,3,5,7,9,\ldots\}] = \{M \in \mathcal{P}(\mathbb{N}) \mid 1 \in M, 2 \notin M, 3 \in M\}$$

sowohl endliche Mengen wie $\{1,3\}$, $\{1,3,8\}$, $\{1,3,7,19\}$, $\{1,3,4,5,6,7,8,9,10,11,12,13\}$ als auch unendliche Mengen wie $\mathbb{N}\backslash\{2\}$, $\mathbb{P}\Delta\{1,2\}$ und $\{1,3,5,7,9,\ldots\}$ selbst. Insgesamt besteht $\mathcal{P}(\mathbb{N})/\sim^{123}$ aus 8 Äquivalenzklassen mit jeweils unendlich vielen Elementen.

(g) Bezüglich der Äquivalenzrelationen $\sim^{\infty}$ auf $\mathcal{P}(\mathbb{N})$ mit

$$M \sim^{\infty} N :\Longleftrightarrow M\Delta N \text{ hat nur endlich viele Elemente für } M,N \subset \mathbb{N}$$

enthält die Äquivalenzklasse der Menge $\mathbb{P} \subset \mathbb{N}$ der Primzahlen nur unendliche Mengen wie zum Beispiel $\{19,23,29,31,37,41\ldots\}$ (Primzahlen ab 19) und $\{1,3,5,7,\ldots,101,103,107,109,113,127,\ldots\}$ (ungerade Zahlen bis 103, Primzahlen ab 107). Hier enthält $\mathcal{P}(\mathbb{N})/\sim^{\infty}$ unendlich viele Äquivalenzklassen mit je unendlich vielen Elementen.

(h) Jede Funktion $f\colon \mathcal{X} \to \mathcal{Y}$ zwischen Mengen $\mathcal{X}$ und $\mathcal{Y}$ induziert eine Äquivalenzrelation $\sim^f$ auf $\mathcal{X}$ durch

$$x \sim^f y :\Longleftrightarrow f(x) = f(y) \text{ für alle } x,y \in \mathcal{X}. \tag{II.4.7}$$

(Reflexivität, Symmetrie, Transitivität besagen hier $f(x) = f(x)$, $f(x) = f(y) \Longrightarrow f(y) = f(x)$ und $(f(x) = f(y)) \wedge (f(y) = f(z)) \Longrightarrow f(x) = f(z)$ für alle $x,y,z \in \mathcal{X}$ und gelten offensichtlich.)

Spezialfälle dieser Bildung haben wir in den Beispielen (d) und (f) gesehen: Die Relation $\sim_T$ ergibt sich für die Funktion $f\colon \mathbb{N} \to \mathbb{N}$, die $n \in \mathbb{N}$ auf die Anzahl der Teiler von n abbildet. Die Relation $\sim^{123}$ ergibt sich für $f\colon \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\{1,2,3\}), M \mapsto M \cap \{1,2,3\}$.

Etwas allgemeiner kann man übrigens auch für $f\colon \mathcal{X} \to \mathcal{Y}$ und *für eine beliebige Äquivalenzrelation* $\cong$ *auf dem Ziel* $\mathcal{Y}$ anstelle der Gleichheitsrelation durch

$$x \simeq^f y :\Longleftrightarrow f(x) \cong f(y) \text{ für alle } x,y \in \mathcal{X}$$

eine Äquivalenzrelation $\simeq^f$ auf $\mathcal{X}$ erhalten.

Mit Äquivalenzrelationen bringt man in der Mathematik ganz allgemein und formal korrekt zum Ausdruck, dass man die in Relation stehenden Elemente miteinander und mit ihrer Äquivalenzklasse im Hinblick

auf gewisse, an der entsprechenden Stelle relevante Eigenschaften identifizieren kann. Ist speziell für Elemente $x \in \mathcal{X}$ nur der Funktionswert $f(x) \in \mathcal{Y}$ unter einer fixierten Abbildung $f\colon \mathcal{X} \to \mathcal{Y}$ relevant, so wird genau dies im nächsten Satz formalisiert:

Satz II.4.33. *Es seien $\mathcal{X}$ und $\mathcal{Y}$ Mengen, $\sim$ eine Äquivalenzrelation auf $\mathcal{X}$ und $f\colon \mathcal{X} \to \mathcal{Y}$ eine Abbildung mit der Eigenschaft*

$$y \sim x \Longrightarrow f(y) = f(x) \text{ für alle } x, y \in \mathcal{X}.$$

Dann gibt es genau eine Abbildung $f_*\colon \mathcal{X}/\!\sim \to \mathcal{Y}$ mit

$$f_*([x]) = f(x) \text{ für alle } x \in \mathcal{X}.$$

Beweis. Wir zeigen, dass $f_*\colon \mathcal{X}/\!\sim \to \mathcal{Y}$ durch $f_*([x]) := f(x)$ für $x \in \mathcal{X}$ wohldefiniert wird, dass also für $x, y \in \mathcal{X}$ aus $[y] = [x]$ stets $f(y) = f(x)$ folgt. Da $[y] = [x]$ gemäß dem vorigen Satz $y \sim x$ bedeutet, ist genau dies aber durch die vorausgesetzte Eigenschaft gesichert. Mit der Wohldefiniertheit ist dann auch die Existenz und Eindeutigkeit von f_* klar. □

Beispiel II.4.34. Wir betrachten die Abbildung $r\colon \mathbb{Z} \to \{0, 1, 2\}$, die ganze Zahlen auf ihren Rest bei Division durch 3 abbildet, also $r(x) = x \bmod 3$ mit $r(x) \in \{0, 1, 2\}$ (oder mit anderen Worten $x = 3\lfloor x/3 \rfloor + r(x)$ für alle $x \in \mathbb{Z}$) erfüllt. Die zugehörige Relation $\sim^r$ (im Sinne von aus (II.4.7)) ist die modulo-3-Relation.

Da $(y = x \bmod 6) \Longrightarrow (y = x \bmod 3) \Longrightarrow r(y) = r(x)$ gilt, greift der Satz sowohl für die Modulo-6- als auch die Modulo-3-Relation und gibt eine eindeutige Abbildung $r_*\colon \mathbb{Z}/6\mathbb{Z} \to \{0, 1, 2\}$ mit $r_*([x]_{\mathbb{Z}/6\mathbb{Z}}) = r(x)$ beziehungsweise $r_*\colon \mathbb{Z}/3\mathbb{Z} \to \{0, 1, 2\}$ mit $r_*([x]_{\mathbb{Z}/3\mathbb{Z}}) = r(x)$ für alle $x \in \mathbb{Z}$. Konkret sieht man an

$$\begin{array}{ccccccc} r(-2) = 1, & r(-1) = 2, & r(0) = 0, & r(1) = 1, & r(2) = 2, & r(3) = 0, & r(4) = 1, \\ r_*([-2]) = 1, & r_*([-1]) = 2, & r_*([0]) = 0, & r_*([1]) = 1, & r_*([2]) = 2, & r_*([3]) = 0, & r_*([4]) = 1 \end{array}$$

(für Äquivalenzklassen entweder modulo 6 oder modulo 3), dass der Übergang von r zu r_* naheliegend und eher eine Formalität ist. Entscheidend ist aber, dass zum Beispiel mit $4 = -2 \bmod 6$ auch $[4] = [-2]$ ist und dies $r(4) = r(-2)$ erzwingt: Wäre nämlich $r(4) \neq r(-2)$, so könnte man den Wert von r_* auf $[4] = [-2]$ nicht wie benötigt festlegen. Dass auf diese Weise tatsächlich keine Probleme entstehen (weil die Abbildung nämlich auf allen Elementen einer Äquivalenzklasse denselben Wert hat), das wird in diesem Beispiel durch die erwähnten Implikationen $(y = x \bmod 6) \Longrightarrow (y = x \bmod 3) \Longrightarrow r(y) = r(x)$ sichergestellt – und genauso im allgemeinen Fall durch die Voraussetzung $y \sim x \Longrightarrow f(y) = f(x)$.

Tatsächlich kann die hier betrachtete Abbildung r nach allen modulo-n-Relationen mit durch 3 teilbarem n faktorisieren, aber nicht nach irgendwelchen anderen modulo-Relationen: Die Faktorisierung modulo 2 scheitert zum Beispiel daran, dass dann tatsächlich $[0] = [2]$, aber $r(0) = 0 \neq r(2) = 2$ ist und kein sinnvoller Wert von r_* auf $[0] = [2]$ festgelegt werden kann.

Bemerkungen II.4.35.

- Die Voraussetzung $y \sim x \Longrightarrow f(y) = f(x)$ für alle $x, y \in \mathcal{X}$ bedeutet, dass f auf Äquivalenzklassen von $\sim$ konstant ist. Mit $\sim^f$ aus (II.4.7) kann diese Voraussetzung äquivalent als $x \sim y \Longrightarrow x \sim^f y$ für alle $x, y \in \mathcal{X}$ geschrieben werden. Insbesondere ist $\sim^f$ die Relation mit den größten Äquivalenzklassen, für die der Satz anwendbar ist.
- Im Satz ist
  - Die Abbildung f_* genau dann injektiv, wenn

    $$y \sim x \Longleftrightarrow f(y) = f(x)$$

    für alle $x, y \in \mathcal{X}$) gilt:
    Ist f_* injektiv, so erhalten wir $f(y) = f(x) \Longrightarrow y \sim x$ für alle $x, y \in \mathcal{X}$ aus den Schlüssen

    $$f(y) = f(x) \implies f_*([y]) = f_*([x]) \implies [y] = [x] \implies y \sim x.$$

    Da die Umkehr-Implikation im Satz vorausgesetzt wird, ist damit $y \sim x \Longleftrightarrow f(y) = f(x)$ für alle $x, y \in \mathcal{X}$ gezeigt.
    Gilt $y \sim x \Longleftrightarrow f(y) = f(x)$ für alle $x, y \in \mathcal{X}$, so erhalten wir

    $$f_*([y]) = f_*([x]) \Longrightarrow [y] = [x]$$

für alle $x, y \in \mathcal{X}$ durch die Schlüsse

$$f_*([y]) = f_*([x]) \implies f(y) = f(x) \implies y \sim x \implies [y] = [x].$$

Damit ist f_* injektiv.

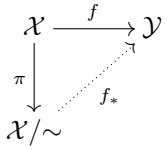
– $\mathrm{Bild}(f_*) = \mathrm{Bild}(f)$ und insbesondere f_* genau dann surjektiv, wenn f surjektiv ist: Die Gleichheit $\mathrm{Bild}(f_*) = \mathrm{Bild}(f)$ liest man aus $f_*([x]) = f(x)$ für alle $x \in \mathcal{X}$ ab und erhält dann auch die Aussage zur Surjektivität.

- Es sei $\pi\colon \mathcal{X} \to \mathcal{X}/\!\sim$ die Quotientenabbildung. Die Schlussfolgerung des Satzes kann abstrakter so formuliert werden, dass genau eine Abbildung $f_*\colon \mathcal{X}/\!\sim \to \mathcal{Y}$ mit

$$f_* \circ \pi = f$$

existiert. Damit wird die Abbildung f in die Faktoren f_* und π zerlegt, was die Verwendung des Begriffs „Faktorisierung“ erklärt. Insbesondere können wir durch die Wahl $\sim\, = \sim^f$ und gemäß der vorigen Bemerkung jede Abbildung f als Hintereinanderausführung der Surjektion π und der Injektion f_* schreiben. Schränkt man f zusätzlich auf das Bild ein, betrachtet man also $f\colon \mathcal{X} \to f(\mathcal{X})$, so ist f_* bijektiv.

Man kann sich die Situation des Satzes anhand des rechts gezeigten Diagramms verdeutlichen und merken. Die Gleichheit $f = f_* \circ \pi$ bringt man dabei auch so zum Ausdruck, dass man von einem *kommutativen Diagramm* spricht, also einem Diagramm, in dem der direkte Weg von $\mathcal{X}$ nach $\mathcal{Y}$ mit f derselben Abbildung entspricht wie der Weg von $\mathcal{X}$ über $\mathcal{X}/\!\sim$ nach $\mathcal{Y}$ mit π und f_*.



II.5. Rationale Zahlen

In diesem Abschnitt diskutieren wir eine Möglichkeit zur *formalen Einführung der rationalen Zahlen*, also der Brüche mit ganzzahligem Zähler und Nenner, letzterer ungleich Null. Die Darstellung solcher Brüche ist bekanntlich nicht eindeutig, zum Beispiel ist $\frac{3}{2} = \frac{6}{4} = \frac{-18}{-12}$, es können also die drei Zähler-Nenner-Paare $(3,2)$, $(6,4)$, $(-18,-12)$ (und viele weitere) zur Darstellung desselben Bruchs herangezogen werden. Der Zusammenhang, dass zwei Paare denselben Bruch darstellen, gibt tatsächlich eine Äquivalenzrelation auf Zähler-Nenner-Paaren, und verschiedene Zähler-Nenner-Paare können als verschiedene Repräsentanten desselben Bruchs betrachtet werden. Deshalb liegt es nahe, mit einer geeigneten Äquivalenzrelation zu arbeiten:

Proposition II.5.1. Durch die Festlegung

$$(z,n) \sim_{\mathbb{Q}} (y,m) :\iff mz = ny \text{ für alle } (z,n),(y,m) \in \mathbb{Z}\times(\mathbb{Z}\backslash\{0\})$$

ist eine Äquivalenzrelation auf $\mathbb{Z}\times(\mathbb{Z}\backslash\{0\})$ gegeben.

Beweis. Wir zeigen die definierenden Eigenschaften (wobei $(z,n),(y,m),(x,\ell) \in \mathbb{Z}\times(\mathbb{Z}\backslash\{0\})$):

- Reflexivität: $(z,n) \sim_{\mathbb{Q}} (z,n)$ bedeutet $nz = nz$ und das gilt für alle $(z,n) \in \mathbb{Z}\times(\mathbb{Z}\backslash\{0\})$.
- Symmetrie: Es gilt sogar $(z,n) \sim_{\mathbb{Q}} (y,m) \iff (y,m) \sim_{\mathbb{Q}} (z,n)$, denn die ausgeschriebene linke Seite $mz = ny$ und die ausgeschriebene rechte Seite $ny = mz$ besagen dasselbe.
- Transitivität: Wir zeigen $((z,n) \sim_{\mathbb{Q}} (y,m) \wedge (y,m) \sim_{\mathbb{Q}} (\ell,x)) \implies (z,n) \sim_{\mathbb{Q}} (\ell,x)$: Die Prämisse bedeutet $mz = ny$ und $\ell y = mx$, und daraus folgt $m\ell z = \ell mz = \ell ny = n\ell y = nmx = mnx$. Wegen $m \neq 0$ erhalten wir $\ell z = nx$, also wie erforderlich $(z,n) \sim_{\mathbb{Q}} (x,\ell)$. □

Definition II.5.2. Wir definieren die Menge der rationalen Zahlen

$$\mathbb{Q} := (\mathbb{Z}\times(\mathbb{Z}\backslash\{0\}))/\sim_{\mathbb{Q}}$$

als die Quotientenmenge der Äquivalenzrelation $\sim_{\mathbb{Q}}$ aus der vorausgehenden Proposition und vereinbaren die Schreibweisen

$$z : n := z/n := \frac{z}{n} := [(z,n)] \text{ mit } z \in \mathbb{Z}, n \in \mathbb{Z}\backslash\{0\}$$

für die Äquivalenzklassen dieser Relation. Wir identifizieren außerdem jede ganze Zahl $z \in \mathbb{Z}$ mit der Äquivalenzklasse $\frac{z}{1} \in \mathbb{Q}$ und fassen in dieser Weise $\mathbb{Z}$ als Teilmenge von $\mathbb{Q}$ auf.

Dass diese Definition sinnvoll ist, liegt vor allem daran, dass sich aus ihr sofort die grundlegende Kürzungs- beziehungsweise Erweiterungsregel

$$\frac{z}{n} = \frac{mz}{mn} \text{ für } z \in \mathbb{Z}, m, n \in \mathbb{Z}\backslash\{0\}$$

ergibt, denn diese Regel bedeutet nach Definition der Äquivalenzrelation nichts anderes als die offensichtliche Gleichheit $mnz = nmz$.

Das Rechnen mit rationalen Zahlen kann nun wie folgt eingeführt werden:

Definition II.5.3. Addition, Subtraktion, Multiplikation und Division rationaler Zahlen werden unter Verwendung der Addition, Subtraktion und Multiplikation ganzer Zahlen durch

$$\begin{aligned} \frac{z}{n} \pm \frac{y}{m} &:= \frac{mz \pm ny}{nm}, \\ \frac{z}{n} \cdot \frac{y}{m} &:= \frac{zy}{nm} \text{ und} \\ \frac{z}{n} \Big/ \frac{y}{m} &:= \frac{zm}{ny} \end{aligned}$$

für $y, z \in \mathbb{Z}$, $m, n \in \mathbb{Z}\backslash\{0\}$ (bei der Division auch $y \in \mathbb{Z}\backslash\{0\}$) definiert.

Dass diese Festlegungen tatsächlich nur von den Äquivalenzklassen $\frac{z}{n}$, $\frac{y}{m}$ und nicht den (Einträgen der) verwendeten Repräsentanten (z, n), (y, m) abhängen, ist zunächst überhaupt nicht klar. Daher ist hier – wie es für das Arbeiten mit Äquivalenzklassen sehr typisch ist – die *Wohldefiniertheit zu zeigen*:

Satz II.5.4. *Die obige Definition der Addition, Subtraktion, Multiplikation und Division rationaler Zahlen hängt nicht von der Vertreterwahl ab.*

Beweis. Für die Wohldefiniertheit der Addition und Subtraktion zeigen wir, dass sich aus $\frac{z}{n} = \frac{z'}{n'}$ und $\frac{y}{m} = \frac{y'}{m'}$ stets $\frac{mz \pm ny}{nm} = \frac{m'z' \pm n'y'}{n'm'}$ ergibt. Dazu schreiben wir die Prämisse mit der Definition der Äquivalenzrelation $\sim_{\mathbb{Q}}$ aus zu $n'z = nz'$ und $m'y = my'$ und bekommen dann

$$n'm'(mz \pm ny) = m'mn'z \pm n'nm'y = mnm'z' \pm nmn'y' = mn(m'z' \pm n'y'),$$

was wie erforderlich $\frac{mz \pm ny}{nm} = \frac{m'z' \pm n'y'}{n'm'}$ bedeutet. Für die Wohldefiniertheit der Multiplikation zeigen wir, dass aus $\frac{z}{n} = \frac{z'}{n'}$ und $\frac{y}{m} = \frac{y'}{m'}$ stets $\frac{zy}{nm} = \frac{z'y'}{n'm'}$ folgt. Dazu nutzen wir wieder $n'z = nz'$ und $m'y = my'$, rechnen $n'm'zy = n'zm'y = nz'my' = nmz'y'$ und bekommen wie gewünscht $\frac{zy}{nm} = \frac{z'y'}{n'm'}$. Die Wohldefiniertheit der Division prüft man ganz ähnlich. □

Insbesondere gelten nach diesen Definitionen für $y, z \in \mathbb{Z}$, $n \in \mathbb{Z}\backslash\{0\}$ stets

$$\begin{aligned} \frac{z}{1} \pm \frac{y}{1} &= \frac{z \pm y}{1}, \\ \frac{z}{1} \cdot \frac{y}{1} &= \frac{zy}{1}, \\ \frac{z}{1} \Big/ \frac{n}{1} &= \frac{z}{n}. \end{aligned}$$

Die ersten beiden Gleichungen besagen dabei, dass die Addition, Subtraktion und Multiplikation auf $\mathbb{Q}$ und $\mathbb{Z}$ *konsistent mit* der (in der Definition vereinbarten) Auffassung von $\mathbb{Z}$ als Teilmenge von $\mathbb{Q}$ zusammenpassen. Die dritte Gleichung besagt, dass die Division von $z \in \mathbb{Z}$ durch $n \in \mathbb{Z}\backslash\{0\}$ in $\mathbb{Q}$ die Äquivalenzklasse $\frac{z}{n}$ gibt, oder anders herum, dass die $\mathbb{Q}$ ausmachenden Äquivalenzklassen $\frac{z}{n}$ die Quotienten z/n der gerade eingeführten Division sind. Dies macht die *Unterscheidung zwischen Äquivalenzklassen und Quotienten fortan unnötig* und erklärt und rechtfertigt, warum wir das Symbol „/“ in beiden Zusammenhängen verwendet haben. (Übrigens besteht auch mit der früher am Rande und nur für den Fall $z = nq$ erwähnten Division in $\mathbb{Z}$ natürlich Konsistenz, da in diesem Fall $\frac{z}{1} \Big/ \frac{n}{1} = \frac{q}{1}$ gilt.)

Aus den Definitionen lassen sich auch auf $\mathbb{Q}$ die *Kommutativität und Assoziativität der Addition und Multiplikation*, die *Distributivgesetze* und weitere bekannte Rechenregeln ableiten. Davon, solche Rechenregeln im Detail nachzuweisen, sehen wir hier ab.

Schließlich können auch die *Standard-Ordnungsrelationen* „$<$“, „$>$“, „$\leq$“, „$\geq$“ *auf rationale Zahlen erweitert* werden: Wir erklären dazu zuerst die strikten Totalordnungen $<$ und $>$ auf $\mathbb{Q}$ durch

$$\frac{z}{n} > \frac{y}{m} :\Longleftrightarrow \frac{y}{m} < \frac{z}{n} :\Longleftrightarrow ny < mz \text{ für } y, z \in \mathbb{Z},\ m, n \in \mathbb{N}.$$

Beachten Sie, dass es genügt, nur *Nenner* aus $\mathbb{N}$ zu betrachten, weil $\frac{z}{n}$ mit $z \in \mathbb{Z}$, $n \in -\mathbb{N}$ auch als $\frac{-z}{-n}$ mit $-z \in \mathbb{Z}$, $-n \in \mathbb{N}$ geschrieben werden kann. Die zugehörigen (nicht-strikten) Totalordnungen $\leq$ und $\geq$ auf $\mathbb{Q}$ ergeben sich durch

$$q \geq p :\Longleftrightarrow p \leq q :\Longleftrightarrow (p < q \vee p = q)$$

für $p, q \in \mathbb{Q}$. Dass diese Festlegungen wohldefiniert sind und man in der Tat totale (strikte) Ordnungsrelationen erhält, prüft man ausgehend von den entsprechenden Eigenschaften auf $\mathbb{Z}$ recht problemlos. Außerdem besteht auch hier insofern Konsistenz, dass $\frac{y}{1} < \frac{z}{1} \Longleftrightarrow y < z$ für alle $y, z \in \mathbb{Z}$ gilt. Genauer gehen wir auf den Umgang und das Rechnen mit Ungleichungen im Kontext der reellen Zahlen ein.

Tatsächlich ist diese Konstruktion der rationalen Zahlen $\mathbb{Q}$ ein illustratives Beispiel für den Umgang mit Äquivalenzrelationen, und wir werden auch im nächsten Abschnitt noch einmal darauf zurückkommen. Soweit es den praktischen Umgang mit rationalen Zahlen und das Rechnen mit Brüchen angeht, unterfüttert die beschriebene Konstruktion aber vor allem die bekannten und üblichen Rechenregeln, während man auf die Konstruktion selbst im Weiteren nicht mehr zurückgreifen muss.

II.6. Mächtigkeit von Mengen

Die Mächtigkeit oder Kardinalität einer Menge verallgemeinert die Anzahl der Elemente der Menge, bleibt aber auch für Mengen mit unendliche vielen Elementen sinnvoll und erlaubt einen Vergleich verschiedener Grade von Unendlichkeit. Um dies präzise fassen zu können, erweist es sich als sinnvoll, zuerst den Vergleich von Mächtigkeiten einzuführen:

Definition II.6.1. Es seien M und N Mengen.

(I) Man nennt M und N *gleichmächtig* oder *von gleicher Kardinalität*, notiert $|M| = |N|$, wenn es eine *Bijektion* von M nach N gibt.

(II) Man nennt N *mindestens gleichmächtig* zu M, notiert $|N| \geq |M|$, und M *höchstens gleichmächtig* zu N, notiert $|M| \leq |N|$, wenn es eine *Injektion* von M nach N gibt.

(III) Man nennt N *echt mächtiger* als M, notiert $|N| > |M|$, und M *echt weniger mächtig* als N, notiert $|M| < |N|$, wenn $|M| \leq |N|$, aber nicht $|M| = |N|$ gilt.

Bemerkungen II.6.2.

- Für *endliche Mengen* M und N bedeutet N gleichmächtig/mindestens gleichmächtig zu/echt mächtiger als M, dass N *genau so viele/mindestens so viele/echt mehr Elemente* als M hat. Dies unterstreicht, dass man bei der Notation $|M|$ die Anzahl der Elemente von M im Hinterkopf haben sollte (auch wenn wir die Mächtigkeit $|M|$ noch nicht definiert haben, sondern nur die *Gleich*mächtigkeit $|M| = |N|$).
- Für die Gleichmächtigkeit beliebiger Mengen M, N und L zeigt man leicht

$$\begin{aligned} |M| &= |M|, \\ |M| = |N| &\Longleftrightarrow |N| = |M|, \\ \big(|L| = |M| \wedge |M| = |N|\big) &\Longrightarrow |L| = |N|. \end{aligned}$$

  Damit ist Gleichmächtigkeit eine Äquivalenzrelation auf der Potenzmenge $\mathcal{P}(\mathcal{X})$ jeder fixierten Menge $\mathcal{X}$. Ohne eine Grundmenge $\mathcal{X}$ zu fixieren, können wir dagegen nicht formal sauber von einer Äquivalenzrelation sprechen, da wir die Menge aller Mengen als Grundmenge bräuchten und diese nicht existiert; vergleiche mit Abschnitt I.5.

- Für den Vergleich von Mächtigkeiten beliebiger Mengen M, N und L gelten generell die an (strikte) Ordnungsrelationen erinnernden Regeln
  - $|M| \leq |M|$,
  - $\big((|M| \leq |N|) \wedge (|N| \leq |M|)\big) \Longrightarrow |M| = |N|$,
  - $\neg((|M| < |N|) \wedge (|N| < |M|))$,
  - $\big((|L| \leq |M|) \wedge (|M| \leq |N|)\big) \Longrightarrow |L| \leq |N|$,
  - $\big((|L| < |M|) \wedge (|M| < |N|)\big) \Longrightarrow |L| < |N|$.

  Während die meisten dieser Regeln naheliegend sind, handelt es sich bei der zweiten Regel um einen bekannten *Satz von Cantor-Schröder-Bernstein*, dessen Beweis bei Interesse im Anhang (siehe Abschnitt V.6) nachgelesen werden kann und eine etwas trickreiche Konstruktion einer Bijektion $M \to N$ aus einer Injektion $M \to N$ und einer Injektion $N \to M$ erfordert.

  Alles in allem können damit für jede fixierte Menge $\mathcal{X}$ die Mindestens-Gleichmächtigkeit und Höchstens-Gleichmächtigkeit als Ordnungsrelationen auf $\mathcal{P}(\mathcal{X})$ modulo der Gleichmächtigkeit angesehen werden und die echt größere und echt geringe Mächtigkeit als strikte Ordnungsrelationen auf $\mathcal{P}(\mathcal{X})$.
- Eine etwas fortgeschrittenere Argumentation mit dem Wohlordnungssatz zeigt auch die Totalordnungseigenschaft, dass stets eine der drei Möglichkeiten $|\mathcal{X}| \leq |\mathcal{Y}|$, $|\mathcal{Y}| \leq |\mathcal{X}|$, $|\mathcal{X}| = |\mathcal{Y}|$ eintritt.

Alles in allem gelten damit für den Vergleich von Mächtigkeiten dieselben Regeln wie für den Vergleich von Zahlen, so dass die eingeführte Notation mit den Symbolen $=$, $\leq$, $\geq$, $<$, $>$ sich als sehr sinnvoll und suggestiv erweist.

Aufbauend auf (dem Vergleich von) Mächtigkeiten können wir einige weitere Begriffe spezifizieren:

Definition II.6.3.
- Mit dem Konzept der Gleichmächtigkeit können wir präzisieren, dass eine Menge M
  - genau $n \in \mathbb{N}_0$ Elemente hat, notiert $|M| = n$, wenn $|M| = |\{1, 2, \ldots, n\}|$ gilt. Im Fall $n = 0$ verstehen wir $\{1, 2, \ldots, n\} = \emptyset$, so dass $|M| = 0$ genau $M = \emptyset$ bedeutet.
  - *endlich* ist, notiert $|M| < \infty$, wenn ein $n \in \mathbb{N}_0$ mit $|M| = n$ existiert,
  - *unendlich* ist, wenn sie nicht endlich ist.

  Gemäß dem Auswahlaxiom (oder einer Folgerung daraus) kann man in jeder unendlichen Menge M rekursiv $x_1 \in M$, $x_2 \in M \setminus \{x_1\}$, $x_3 \in M \setminus \{x_1, x_2\}$, $x_4 \in M \setminus \{x_1, x_2, x_3\}$, $\ldots$ wählen, daraus eine Injektion $\mathbb{N} \to M$, $n \mapsto x_n$ erhalten und so $|\mathbb{N}| \leq |M|$ einsehen. In diesem Sinne *sind die natürlichen Zahlen die kleinste unendliche Menge.*
- Eine alternative, äquivalente Definition (un)endlicher Mengen geht auf Dedekind zurück und kommt ohne expliziten Rückgriff auf die (Struktur der) natürlichen Zahlen aus. Dabei erklärt man eine Menge M als unendlich, wenn eine echte Teilmenge $T \subsetneq M$ mit $|T| = |M|$ existiert, und andernfalls als endlich. Inspiriert wird diese Definition durch die schon in Abschnitt II.2 erwähnte Eigenschaft, dass die Nachfolge-Abbildung eine Bijektion von $\mathbb{N}$ auf $\mathbb{N}\setminus\{1\}$ ist und damit $|\mathbb{N}\setminus\{1\}| = |\mathbb{N}|$ für die echte Teilmenge $\mathbb{N}\setminus\{1\} \subsetneq \mathbb{N}$ gilt.

Vor allem bietet das Konzept der Gleichmächtigkeit aber eine Möglichkeit, auch unendliche Menge nach Größenvergleich mit $\mathbb{N}$ zu unterscheiden:

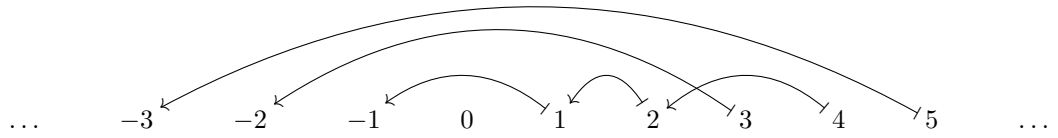
Definition II.6.4. Eine Menge M heißt
- *höchstens abzählbar*, wenn $|M| \leq |\mathbb{N}|$ gilt,
- *abzählbar (unendlich)*, wenn $|M| = |\mathbb{N}|$ gilt,
- *überabzählbar*, wenn $|M| > |\mathbb{N}|$ gilt.

Bemerkungen II.6.5.
- *Abzählbarkeit* einer Menge M *bedeutet, dass man die Elemente* von M *mit natürlichen Zahlen nummerieren und* in der Form $M = \{x_1, x_2, \ldots, x_{n-1}, x_n\}$ (endliche Liste im Fall $|M| < |\mathbb{N}|$ mit $n = |M| \in \mathbb{N}_0$) oder $M = \{x_1, x_2, x_3, \ldots\}$ (unendliche Liste im Fall $|M| = |\mathbb{N}|$) *nacheinander aufzählen kann.*
- Neben $\mathbb{N}$ und $\mathbb{N}_0$ selbst ist die Menge $\mathbb{Z}$ der ganzen Zahlen abzählbar unendlich.

  Zum Beispiel ist die Abbildung

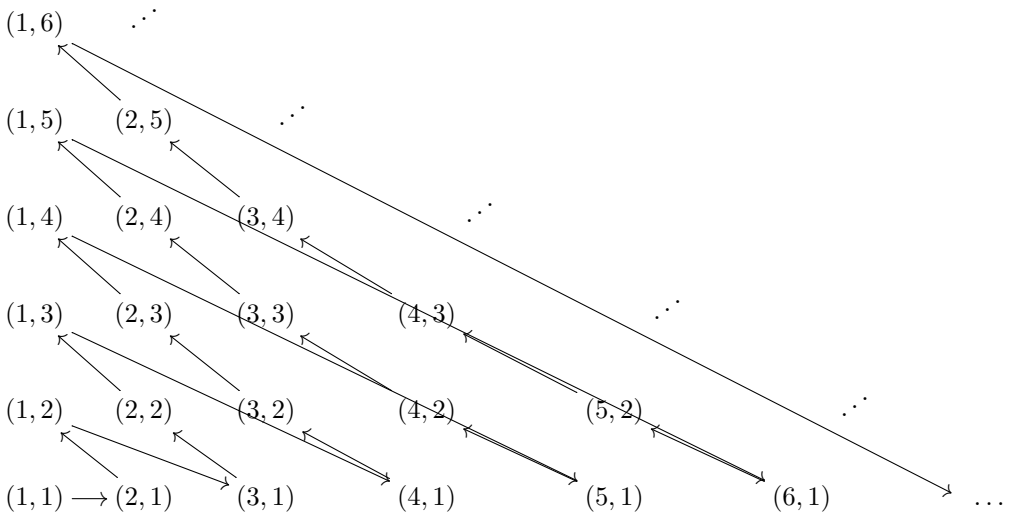
$$f\colon \mathbb{N}_0 \to \mathbb{Z}, \quad f(n) = \begin{cases} \frac{n}{2}, & \text{falls } n \text{ gerade ist,} \\ -\frac{n+1}{2}, & \text{falls } n \text{ ungerade ist,} \end{cases}$$

eine Bijektion.



- Das *erste Cantorsche Diagonalverfahren* zeigt, dass die kartesischen Produkte $\mathbb{N}^2$, $\mathbb{Z}^2$ und die Menge $\mathbb{Q}$ der rationalen Zahlen abzählbar unendlich sind. Dazu konstruiert man erst eine Bijektion $\mathbb{N} \to \mathbb{N}^2$, hinter der die Grundidee steht, die Paare in $\mathbb{N}^2$ gemäß des in Abbildung II.6.1 gezeigten Diagonalschemas zu nummerieren:

(II.6.1)



Ist damit $|\mathbb{N}^2| = |\mathbb{N}|$ gezeigt, so kann dies mit $|\mathbb{Z}| = |\mathbb{N}|$ und folglich $|\mathbb{Z}^2| = |\mathbb{N}^2|$ zu $|\mathbb{Z}^2| = |\mathbb{N}|$ zusammengesetzt werden. Weiter folgt nun $|\mathbb{Q}| = |\mathbb{N}|$, denn zu einem gilt mit $\mathbb{N} \subset \mathbb{Q}$ natürlich $|\mathbb{N}| \leq |\mathbb{Q}|$, zum anderen ergibt sich mit der Konstruktion $\mathbb{Q} = (\mathbb{Z}\times(\mathbb{Z}\backslash\{0\}))/\sim_{\mathbb{Q}}$ des Abschnitts II.5 und dem Vorausgehenden, dass $|\mathbb{Q}| \leq |\mathbb{Z}^2| = |\mathbb{N}|$ gilt.

- Beim *zweiten Cantorschen Diagonalverfahren* handelt es sich um ein Widerspruchsargument zum Nachweis, dass die Menge $\mathbb{R}$ der reellen Zahlen und ihre Teilmenge $\mathbb{I} := \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$ tatsächlich *überabzählbar* sind. Man verwendet dazu (worauf wir später noch genauer eingehen), dass jedes $x \in \mathbb{I}$ eine Darstellung $0, z_1 z_2 z_3 \ldots$ mit unendlich vielen Nachkomma-Ziffern $z_1, z_2, z_3, \ldots \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ hat. Angenommen, man könnte in dieser Darstellung nun alle $x \in \mathbb{I}$ in einer Reihenfolge

$$\begin{aligned} x_1 &= 0, \mathbf{z_{11}} z_{12} z_{13} z_{14} \ldots \\ x_2 &= 0, z_{21} \mathbf{z_{22}} z_{23} z_{24} \ldots \\ x_3 &= 0, z_{31} z_{32} \mathbf{z_{33}} z_{34} \ldots \\ x_4 &= 0, z_{41} z_{42} z_{43} \mathbf{z_{44}} \ldots \\ &\vdots \quad \vdots \end{aligned}$$

auflisten. Dann könnte man für jeden Index $i \in \mathbb{N}$ eine von der fett gedruckten Ziffer z_{ii} auf der Diagonale verschiedene Ziffer $z_i^* \in \{1, 2, 3, 4, 5, 6, 7, 8\} \setminus \{z_{ii}\}$ wählen und erhielte $x^* := 0, z_1^* z_2^* z_3^* z_4^* \ldots \in \mathbb{I}$. Die Zahl x^* unterscheidet sich wegen $z_i^* \neq z_{ii}$ an der i-ten Nachkommastelle von der i-ten Zahl

auf der Liste, unterscheidet sich also[7] von *jeder* Zahl auf der Liste und kommt auf der Liste nicht vor. *Widerspruch*! Also können die $x \in \mathbb{I}$ und erst recht alle $x \in \mathbb{R}$ nicht nummeriert werden, $\mathbb{I}$ und $\mathbb{R}$ sind in der Tat überabzählbar.

- Tatsächlich sind $\mathbb{R}$ und $\mathbb{I}$ gleichmächtig zur Potenzmenge $\mathcal{P}(\mathbb{N})$ (siehe [**4**]).
- Einige weiterführende Bemerkungen zu Mächtigkeiten von Mengen finden sich im Anhang im Abschnitt V.5.

[7] An dieser Stelle ist etwas Vorsicht geboten, da bei der unendlichen Darstellung von Zahlen mit eigentlich nur endlich vielen Nachkomma-Stellen die Doppeldeutigkeit $0, z_1 z_2 \ldots z_{k-1} z_k 0000 \ldots = 0, z_1 z_2 \ldots z_{k-1}(z_k - 1)9999 \ldots$ besteht. Um trotzdem sicher sein zu können, dass x^* nicht auf der Liste vorkommt, wurden bei der Wahl der z_i^* die Ziffern 0 und 9 ausgeschlossen.

KAPITEL III

Algebraische Grundstrukturen

In diesem Kapitel werden wir das Konzept eines Zahlbereichs samt seiner Rechenarten darauf sehr weitgehend verallgemeinern. Dies ist nützlich, um gemeinsame Eigenschaften der Zahlbereiche und der Rechenarten sowie weiterer, ähnlich gearteter Operationen einordnen, beschreiben und abstrahieren zu können.

III.1. Verknüpfungen, Halbgruppen und Gruppen

Wir beginnen mit der Einführung von Verknüpfungen, die Rechenarten verallgemeinern:

Definition III.1.1.

(I) Unter einer *Verknüpfung* auf einer Menge G verstehen wir eine Abbildung $*\colon G \times G \to G$. Wir verwenden die Notation $g * h$ statt $*(g,h)$ für das Bild von $(g,h) \in G \times G$ unter der Verknüpfung $*$.

(II) Wir nennen eine Verknüpfung $*$ auf einer Menge G *assoziativ*, wenn für alle $g,h,k \in G$ gilt:

$$(g * h) * k = g * (h * k).$$

(III) Wir nennen eine Verknüpfung $*$ auf einer Menge G *kommutativ*, wenn für alle $g,h \in G$ gilt:

$$g * h = h * g.$$

Beispiele III.1.2.

(a) Verknüpfungen auf endlichen Mengen können durch *Verknüpfungstafeln* genannte Tabellen angegeben werden, indem man man $g*h$ in das Feld in der zu g gehörigen Zeile und zu h gehörigen Spalte einträgt. Zum Beispiel sind eine Verknüpfung τ auf $\{2,3,5,7\}$ und eine Verknüpfung $\rtimes$ auf einer beliebigen 3-elementigen Menge $\{e,\alpha,\beta\}$ wie folgt gegeben:

τ	2	3	5	7
2	5	2	7	3
3	2	3	5	7
5	7	5	7	2
7	3	7	2	5

$\rtimes$	e	α	β
e	e	α	β
α	α	α	α
β	β	β	β

Dabei ist τ nicht assoziativ, weil zum Beispiel

$$(2\,\tau\,2)\tau 5 = 7 \neq 2\tau(2\,\tau\,5) = 3,$$

aber kommutativ. Das erkennt man bei gleicher Reihenfolge in Eingangsspalte und Kopfzeile an Spiegelsymmetrie bezüglich der Diagonalen.

Dagegen ist $\rtimes$ assoziativ, denn die Verknüpfung mehrerer Elemente gibt unabhängig von der Klammerung das am weitesten links stehende Element $\neq e$, falls ein solches existiert, und e sonst. Aber sie ist nicht kommutativ, weil zum Beispiel gilt

$$\alpha \rtimes \beta = \alpha \neq \beta \rtimes \alpha = \beta.$$

Die Darstellung einer Verknüpfung durch solche Tafeln wird aber schon bei relativ wenigen Elementen unübersichtlich.

(b) Die Addition $+$ und die Multiplikation $\cdot$ sind assoziative und kommutative Verknüpfungen auf jedem der Zahlbereiche $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

(c) Die Subtraktion $-$ ist keine Verknüpfung auf $\mathbb{N}$ oder $\mathbb{N}_0$ (zum Beispiel ist $1-2 \notin \mathbb{N}_0$). Sie ist eine Verknüpfung auf $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$, ist dort aber weder assoziativ (zum Beispiel $(0-0)-1 \neq 0-(0-1)$) noch kommutativ ($0-1 \neq 1-0$).

(d) Die Division : (die wir normalerweise mit dem Schrägstrich oder Bruchstrich notieren) ist keine Verknüpfung auf irgendeinem Zahlbereich, der 0 enthält, (Division durch 0 undefiniert) und auch nicht auf $\mathbb{N}$ oder $\mathbb{Z}\backslash\{0\}$ ($\frac{1}{2} \notin \mathbb{Z}$). Sie ist eine Verknüpfung auf $\mathbb{Q}\backslash\{0\}$ und $\mathbb{R}\backslash\{0\}$, ist dort aber weder assoziativ ($\frac{1/1}{2} \neq \frac{1}{1/2}$) noch kommutativ ($\frac{1}{2} \neq \frac{2}{1}$).

(e) Das Potenzieren $(m,n) \mapsto m^n$ ist eine Verknüpfung auf $\mathbb{N}$ und, sobald man irgendeine Konvention für $0^0 \in \mathbb{N}_0$ festlegt, auch auf $\mathbb{N}_0$, ist aber weder assoziativ ($2^{(1^2)} \neq (2^1)^2$) noch kommutativ ($2^1 \neq 1^2$).

(f) Für jede Menge $\mathcal{X}$ ist die Komposition $\circ$ von Selbstabbildungen eine Verknüpfung auf $\mathrm{Abb}(\mathcal{X})$. Diese Verknüpfung ist assoziativ, aber für $|\mathcal{X}| \geq 2$ nicht kommutativ (siehe Satz und Bemerkung in Abschnitt II.1). Die Komposition ist auch eine assoziative Verknüpfung auf gewissen Teilmengen von $\mathrm{Abb}(\mathcal{X})$, zum Beispiel auf der Menge der Injektionen $\mathcal{X} \to \mathcal{X}$, der Menge der Surjektionen $\mathcal{X} \to \mathcal{X}$ und der Menge der Bijektionen $\mathcal{X} \to \mathcal{X}$.

(g) Für jede Menge $\mathcal{X}$ sind die Mengen-Operationen Vereinigung ($\cup$), Durchschnitt ($\cap$), Differenz ($\setminus$) und symmetrische Differenz (Δ) Verknüpfungen auf der Potenzmenge $\mathcal{P}(\mathcal{X})$. Dabei sind $\cup$, $\cap$ und Δ kommutativ und assoziativ (vergleiche mit Abschnitt I.5), während $\setminus$ für $|\mathcal{X}| \geqslant 1$ weder assoziativ ($(\{x\} \setminus \emptyset) \setminus \{x\} \neq \{x\} \setminus (\emptyset \setminus \{x\})$) noch kommutativ ($\{x\} \setminus \emptyset \neq \emptyset \setminus \{x\}$) ist.

(h) Etwas ungewöhnliche Beispiele für Verknüpfungen sind der größte gemeinsame Teiler und das kleinste gemeinsame Vielfache zweier natürlicher Zahlen. Man kann mittels der Primfaktorzerlegung zeigen, dass es sich um assoziative und kommutative Verknüpfungen auf $\mathbb{N}$ handelt.

Oftmals nützlich im Zusammenhang mit verschiedenen Verknüpfungen ist:

Notation III.1.3. Sei $*$ eine Verknüpfung auf einer Menge G. Für $g \in G$ und $A, B \subset G$ verwenden wir häufig die auf Mengen erweiterten Notationen

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

für die Bilder der kartesischen Produkte $\{g\}\times A$, $A\times\{g\}$ und $A\times B$ unter $*$.

Bemerkungen III.1.4.

- Tatsächlich wird mit der Definition von $A * B$ eine Verknüpfung auf $\mathcal{P}(G)$ erklärt, die genau dann assoziativ beziehungsweise kommutativ ist, wenn die Verknüpfung $*$ auf G dies ist.
- Speziell ist die sich aus der Addition $+$ auf einem Zahlbereich $\mathbb{B} \in \{\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ ergebende Verknüpfung $+$ auf $\mathcal{P}(\mathbb{B})$ mit $A+B = \{a+b \,|\, (a,b) \in A\times B\}$ als *Minkowski-Addition* von Mengen bekannt.
- Konkret ergibt sich beispielsweise aus $A = \{0,4,7\}$ und $B = \{1,-5\}$ durch Multiplikation und Subtraktion $3A-B = \{-1,5,11,17,20,26\}$.
- Die eingeführte Notation ist sehr nützlich und erlaubt es beispielsweise, die Mengen der geraden beziehungsweise ungeraden Zahlen in $\mathbb{N}$ und $\mathbb{Z}$ prägnant als $2\mathbb{N}$ und $2\mathbb{Z}$ beziehungsweise $2\mathbb{N}-1$ und $2\mathbb{Z}+1 = 2\mathbb{Z}-1$ zu schreiben.
- Dass bei Verwendung solcher Notationen *etwas Vorsicht* geboten ist, erkennt man aber daran, dass schon für $A \subset \mathbb{N}$ meistens $A+A \neq 2A$ und $A-A \neq \{0\}$ gelten: Beispielsweise für $A = \{1,2\}$ ergibt sich $A+A = \{2,3,4\}$, aber $2A = \{2,4\}$. Weiterhin gilt $\mathbb{N}-\mathbb{N} = \mathbb{Z}$.

Notation III.1.5. Es sei $*$ eine Verknüpfung auf einer Menge G.

- Für jedes $n \in \mathbb{N}$ definieren wir die *komponentenweise* Anwendung der Verknüpfung $*$ *auf Tupeln* durch
$$x * y := (x_1 * y_1, x_2 * y_2, \ldots, x_n * y_n) \in G^n \text{ für } x, y \in G^n.$$
- Für jede Menge $\mathcal{X}$ definieren wir die *punktweise* Anwendung der Verknüpfung $*$ *auf Abbildungen* $f_1, f_2 \colon \mathcal{X} \to G$ durch
$$(f_1 * f_2)(x) := f_1(x) * f_2(x) \in G \text{ für alle } x \in \mathcal{X}$$
und legen damit eine Abbildung $f_1 * f_2 \colon \mathcal{X} \to G$ fest.

Bemerkungen III.1.6.

- Mit dieser Definition von $x * y$ und von $f_1 * f_2$ erhalten wir eine Verknüpfung auf dem kartesischen Produkt G^n und eine Verknüpfung auf der Menge $\mathrm{Abb}(\mathcal{X}, G)$ der Abbildungen $\mathcal{X} \to G$. Beide diese Verknüpfungen sind (wenn $\mathcal{X} \neq \emptyset$) genau dann assoziativ beziehungsweise kommutativ, wenn $*$ auf G dies ist.
- Insbesondere sind hiermit die komponentenweise Summe/Differenz $x \pm y$, das komponentenweise Produkt $x \cdot y$, der komponentenweise Quotient $\frac{x}{y}$ von Tupeln $x, y \in \mathbb{B}^n$ und die komponentenweise Summe/Differenz $f_1 \pm f_2$, das komponentenweise Produkt $f_1 \cdot f_2$, der komponentenweise Quotient $\frac{f_1}{f_2}$ von Abbildungen $f_1, f_2 \colon \mathcal{X} \to \mathbb{B}$ erklärt, jedenfalls soweit die Rechenoperationen Verknüpfungen auf den Zahlbereichen $\mathbb{B} \in \{\mathbb{N}, \mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{Z}\backslash\{0\}, \mathbb{Q}\backslash\{0\}, \mathbb{R}\backslash\{0\}\}$ sind.

Definition III.1.7. Es sei $*$ eine Verknüpfung auf einer Menge G.

(I) Wir nennen $e \in G$ ein *neutrales Element* bezüglich $*$, wenn $e * g = g = g * e$ für alle $g \in G$ gilt.

(II) Falls ein neutrales Element $e \in G$ für $*$ existiert, so nennen wir $h \in G$ ein zu $g \in G$ bezüglich $*$ *inverses Element* oder *Inverses* zu $g \in G$, wenn $h * g = e = g * h$ gilt. Existiert ein Inverses zu $g \in G$, so heißt g (in G) *invertierbar*.

Bemerkungen III.1.8. Es sei $*$ eine Verknüpfung auf einer Menge G.

(a) Existiert ein *neutrales Element* $e \in G$ für $*$, so ist dieses stets *eindeutig*. Ist nämlich $e' \in G$ ein weiteres neutrales Element für $*$, so folgt aus der Definition $e' = e' * e = e$. Daher ist es beim Umgang mit inversen Elementen nicht nötig, das neutrale Element explizit zu benennen, solange es existiert.

(b) Existiert ein *inverses Element* $h \in G$ zu $g \in G$ und ist $*$ zumindest assoziativ, so ist das inverse Element h zu g *eindeutig*. Ist nämlich $h' \in G$ ein weiteres inverses Element zu g, so folgt nach Definition

$$h' = h' * e = h' * (g * h) = (h' * g) * h = e * h = h.$$

Dasselbe Argument haben wir für die Komposition von Abbildungen schon in Abschnitt II.1 benutzt.

(c) Bei assoziativer Verknüpfung $*$ wird die Notation $g^{-1} \in G$ für das inverse Element zu $g \in G$ sinnvoll. Wir halten fest, dass g^{-1} nach Definition $g^{-1} * g = e = g * g^{-1}$ erfüllt und folgende allgemeine Regeln für Inverse gelten:

$$\begin{aligned} e^{-1} &= e && \text{für das neutrale Element } e \in G, \\ \left(g^{-1}\right)^{-1} &= g && \text{für invertierbares } g \in G, \\ \left(g * h\right)^{-1} &= h^{-1} * g^{-1} && \text{für invertierbare } g, h \in G\,. \end{aligned}$$

Für die letzte Regel rechnet man dabei mit den Definitionen und Assoziativität nach, dass $h^{-1} * g^{-1} * g * h = h^{-1} * e * h = h^{-1} * h = e = g * g^{-1} = g * e * g^{-1} = g * h * h^{-1} * g^{-1}$ gilt. Diese Regel ist auch als Socken-Schuhe-Regel bekannt: Wenn Sie sich anziehen, ziehen Sie erst die Socken an, dann die Schuhe. Beim Ausziehen ist das (hoffentlich) andersherum!

(d) Manchmal werden auch *linksneutrale Elemente* $\ell \in G$ mit $\ell * g = g$ für alle $g \in G$ und *rechtsneutrale Elemente* $r \in G$ mit $g = g * r$ für alle $g \in G$ betrachtet. Solche Elemente sind im Allgemeinen nicht eindeutig bestimmt. Sobald aber ein linksneutrales Element ℓ *und* ein rechtsneutrales Element r *existieren*, sind beide eindeutig, stimmen überein und sind damit auch das eindeutige neutrale Element. Das ergibt sich aus $\ell = \ell * r = r$.

(e) Ebenso werden (Existenz des neutralen Elements e vorausgesetzt) manchmal *Linksinverse* $h \in G$ zu $g \in G$ mit $h * g = e$ und *Rechtsinverse* $h \in G$ zu $g \in G$ mit $g * h = e$ betrachtet. Linksinverse und Rechtsinverse zu g sind im Allgemeinen nicht eindeutig bestimmt. Sobald aber ein Linksinverses h' *und* ein Rechtsinverses h zu einem g *existieren* und $*$ assoziativ ist, sind beide eindeutig, stimmen überein und sind damit auch das eindeutige Inverse zu g. Das zeigt die Rechnung aus Bemerkung (2).

Definition III.1.9.

(I) Eine *Halbgruppe* ist ein Paar $(G, *)$ aus einer Menge G und einer *assoziativen* Verknüpfung $*$ auf G. Gibt es für $*$ ein neutrales Element, so spricht man von einem *Monoid* (oder einer *Halbgruppe mit neutralem Element*).

(II) Eine *Gruppe* $(G,*)$ ist ein Monoid, in dem es zu jedem Element $g \in G$ ein inverses Element $g^{-1} \in G$ gibt.

(III) Eine Halbgruppe oder Gruppe $(G,*)$ heißt *kommutativ* oder *abelsch*, wenn $*$ eine kommutative Verknüpfung ist.

Bemerkung III.1.10. Später verzichtet man sehr häufig auf die explizite Angabe der Verknüpfung $*$, die sich oft aus dem Kontext ergibt, und spricht nur davon, dass die zugrundeliegende Menge G eine (Halb-)Gruppe ist. Ist speziell die Verknüpfung eine Addition oder Multiplikation (gewisser Objekte), so spricht man von einer *additiven (Halb-)Gruppe* G oder *multiplikativen (Halb-)Gruppe* G.

Beispiele III.1.11.

- Mit den Verknüpfungstafeln (die erste gegenüber der für τ im früheren Beispiel (1) nur bei der Verknüpfung von 5 mit sich modifiziert, die zweite unverändert)

$\circledast$	2	3	5	7
2	5	2	7	3
3	2	3	5	7
5	7	5	3	2
7	3	7	2	5

und

$\rtimes$	e	α	β
e	e	α	β
α	α	α	α
β	β	β	β

wird $(\{2,3,5,7\},\circledast)$ zu einer abelschen Gruppe mit neutralem Element 3 und den Inversen 2^{-1}=7, 7^{-1}=2, 5^{-1}=5 sowie $(\{e,\alpha,\beta\},\rtimes)$ zu einem nicht-kommutativen Monoid (mit neutralem Element e) und nicht-invertierbaren Elementen α, β.

Die Eigenschaft des neutralen Elements erkennt man in der Verknüpfungstafel daran, dass die zugehörige Spalte und Zeile Kopien der Eingangsspalte und Kopfzeile sind. Für die Verknüpfungstafel einer Gruppe ist neben der Existenz des neutralen Elements erforderlich (aber noch nicht ausreichend[1]), dass jedes Element in jeder Zeile und jeder Spalte genau einmal auftritt.

- Für die üblichen Zahlbereiche $\mathbb{B}$ mit der Addition $+$ und der Multiplikation $\cdot$ sind $(\mathbb{B},+)$ und $(\mathbb{B},\cdot)$, wie in folgender Tabelle angegeben, kommutative Halbgruppen (kHG), kommutative Monoide (kM) oder sogar kommutative Gruppen (kG), wobei jeweils die stärkste Eigenschaft angegeben wird:

	$\mathbb{N}$	$\mathbb{N}_0$	$\mathbb{Z}$	$\mathbb{Q}$	$\mathbb{R}$	$\mathbb{Z}\backslash\{0\}$	$\mathbb{Q}\backslash\{0\}$	$\mathbb{R}\backslash\{0\}$
$+$	kHG	kM	kG			—		
$\cdot$	kM					kM	kG	

Das *neutrale Element der Addition* ist immer 0, das *neutrale Element der Multiplikation* ist immer 1. Das *additive Inverse* zu $x \in \mathbb{B}$ ist $-x$, das *multiplikative Inverse* zu $x \in \mathbb{B}\backslash\{0\}$ ist $\frac{1}{x}$.

All dies ergibt sich sofort aus den üblichen Rechenregeln in den Zahlbereichen.

- *Produkt-(Halb-)Gruppen:* Für jedes $n \in \mathbb{N}$ und jede (Halb-)Gruppe $(G,*)$ ist auch $(G^n,*)$ mit der komponentenweisen Verknüpfung $*$ eine (Halb-)Gruppe, auf die sich auch Kommutativität und/oder Existenz des neutralen Elements überträgt.
  Allgemeiner ergibt sich auch als Produkt von $n \in \mathbb{N}$ (Halb-)Gruppen $(G_1,*)$, $(G_2,*)$, …, $(G_n,*)$ eine (Halb-)Gruppe $(G_1\times G_2\times\ldots\times G_n,*)$, deren Verknüpfung durch
  $$(g_1,g_2,\ldots,g_n)*(g_1',g_2',\ldots,g_n') := (g_1{*}g_1',g_2{*}g_2',\ldots,g_n{*}g_n')$$
  für $(g_1,g_2,\ldots,g_n),(g_1',g_2',\ldots,g_n') \in G_1\times G_2\times\ldots\times G_n$ komponentenweise erklärt ist.
- *Abbildungs-(Halb-)Gruppen:* Für jede Menge $\mathcal{X}$ und jede (Halb-)Gruppe $(G,*)$ ist auch $(\mathrm{Abb}(\mathcal{X},G),*)$ mit der punktweisen Verknüpfung $*$ eine (Halb-)Gruppe, auf die sich auch Kommutativität und/oder

[1] In der Tat definiert nebenstehende Verknüpfungstafel eine kommutative Verknüpfung $\star$ auf $\{e,\alpha,\beta,\gamma,\delta,\varepsilon\}$ mit neutralem Element e, wobei jedes Element in jeder Zeile und Spalte genau einmal auftritt, wegen $(\delta\star\gamma)\star\beta = \delta \neq \delta\star(\gamma\star\beta) = \beta$ aber dennoch keine Assoziativität und keine Gruppe vorliegt:

$\star$	e	α	β	γ	δ	ε
e	e	α	β	γ	δ	ε
α	α	e	δ	β	ε	γ
β	β	δ	e	ε	γ	α
γ	γ	β	ε	e	α	δ
δ	δ	ε	γ	α	e	β
ε	ε	γ	α	δ	β	e

Existenz des neutralen Elements überträgt. Speziell sind $(\mathrm{Abb}(\mathcal{X},\mathbb{B}),+)$ und $(\mathrm{Abb}(\mathcal{X},\mathbb{B}),\cdot)$ insoweit (Halb-)Gruppen wie $(\mathbb{B},+)$ und $(\mathbb{B},\cdot)$.

(Halb-)Gruppen von Selbstabbildungen: Für jede Menge $\mathcal{X}$ ist $(\mathrm{Abb}(\mathcal{X}),\circ)$, die Menge der Selbstabbildungen von $\mathcal{X}$ mit der Komposition $\circ$, ein (für $|\mathcal{X}|\geq 2$ nicht kommutatives) Monoid. Das neutrale Element ist $\mathrm{id}_{\mathcal{X}}$. Die Menge der Bijektionen $\mathcal{X}\to\mathcal{X}$ ist mit $\circ$ sogar eine (für $|\mathcal{X}|\geq 3$ nicht kommutative) Gruppe, in der das Inverse einer Bijektion deren Umkehrabbildung ist.

Die Menge der Injektionen $\mathcal{X}\to\mathcal{X}$ und die Menge der Surjektionen $\mathcal{X}\to\mathcal{X}$ stimmen übrigens für endliches $\mathcal{X}$ mit der Menge der Bijektionen $\mathcal{X}\to\mathcal{X}$ überein, werden für unendliches $\mathcal{X}$ aber durch $\circ$ zu nicht-kommutativen Halbgruppen mit neutralem Element, in denen für jedes Element ein Linksinverses beziehungsweise Rechtsinverses existiert, dieses für nicht-invertierbare Elemente aber nicht eindeutig ist.

- *(Halb-)Gruppen von Mengen:* Für jede Menge $\mathcal{X}$ sind $(\mathcal{P}(\mathcal{X}),\cup)$ und $(\mathcal{P}(\mathcal{X}),\cap)$ kommutative Halbgruppen mit neutralem Element $\emptyset$ beziehungsweise $\mathcal{X}$. Weiterhin ist $(\mathcal{P}(\mathcal{X}),\Delta)$ sogar eine kommutative Gruppe mit neutralem Element $\emptyset$ und der (bei dieser Betrachtungsweise ungewohnten Eigenschaft), dass mit $M\Delta M=\emptyset$ für alle $M\subset\mathcal{X}$ jedes Element zu sich selbst invers ist.
- Mit dem größten gemeinsamen Teiler zweier natürlicher Zahlen wird $\mathbb{N}$ zu einer kommutativen Halbgruppe, mit dem kleinsten gemeinsamen Vielfachen zweier natürlicher Zahlen sogar zu einer kommutativen Halbgruppe mit neutralem Element 1.

Definition III.1.12. Wir bezeichnen die Menge der Modulo-n-Äquivalenzklassen künftig auch als den Modulo-n-*Restklassenring*

$$\mathbb{Z}/n\mathbb{Z}=\mathbb{Z}/\sim_n .$$

Nach Definition von Modulo-Relationen und Äquivalenzklassen gilt

$$[x]_{\mathbb{Z}/n\mathbb{Z}}=\{y\in\mathbb{Z}\mid y\sim_n x\}=\{y\in\mathbb{Z}\mid y-x=nz \text{ für ein } z\in\mathbb{Z}\}=\{x+nz\mid z\in\mathbb{Z}\}=x+n\mathbb{Z}.$$

Daher werden wir zukünftig auch die Schreibweise $[x]_{\mathbb{Z}/n\mathbb{Z}}=x+n\mathbb{Z}$ benutzen.

Satz III.1.13. *Es sei $n\in\mathbb{N}\setminus\{1\}$.*

(1) *Die Addition und die Multiplikation geben durch*

$$[x]_{\mathbb{Z}/n\mathbb{Z}}+[y]_{\mathbb{Z}/n\mathbb{Z}}:=[x+y]_{\mathbb{Z}/n\mathbb{Z}} \text{ und } [x]_{\mathbb{Z}/n\mathbb{Z}}\cdot[y]_{\mathbb{Z}/n\mathbb{Z}}:=[xy]_{\mathbb{Z}/n\mathbb{Z}} \text{ für } x,y\in\mathbb{Z}$$

wohldefinierte Verknüpfungen auf $\mathbb{Z}/n\mathbb{Z}$. Im Fall der Addition kann $[x]_{\mathbb{Z}/n\mathbb{Z}}+[y]_{\mathbb{Z}/n\mathbb{Z}}$ äquivalent als Minkowski-Addition von Teilmengen von $\mathbb{Z}$ interpretiert werden.

(2) *Mit diesen Verknüpfungen ist $(\mathbb{Z}/n\mathbb{Z},+)$ eine abelsche Gruppe mit neutralem Element $[0]_{\mathbb{Z}/n\mathbb{Z}}$ und Inversem $[-x]_{\mathbb{Z}/n\mathbb{Z}}$ zu $[x]_{\mathbb{Z}/n\mathbb{Z}}$ sowie $(\mathbb{Z}/n\mathbb{Z},\cdot)$ ein kommutatives Monoid mit neutralem Element $[1]_{\mathbb{Z}/n\mathbb{Z}}$. Kürzen wir $\mathbb{Z}/p\mathbb{Z}^\times:=\mathbb{Z}/p\mathbb{Z}\setminus\{[0]_{\mathbb{Z}/p\mathbb{Z}}\}$ ab, so ist darüber hinaus $(\mathbb{Z}/p\mathbb{Z}^\times,\cdot)$ genau für die Primzahlen $p\in\mathbb{N}$ eine abelsche Gruppe.*

Beweis von Teil (1) des Satzes. Für die Wohldefiniertheit der Addition ist $[x']=[x]\wedge[y']=[y]\implies[x'+y']=[x+y]$ für die Äquivalenzklassen bezüglich der Modulo-n-Relation zu zeigen. Sind $x'-x$ und $y'-y$ durch n teilbar, so sind, stets auch $(x'+y')-(x+y)=(x'-x)+(y'-y)$ durch n teilbar. Bei Interpretation als Minkowski-Addition ergibt sich die identische Vorschrift $[x]_{\mathbb{Z}/n\mathbb{Z}}+[y]_{\mathbb{Z}/n\mathbb{Z}}=(x+n\mathbb{Z})+(y+n\mathbb{Z})=x+y+n\mathbb{Z}+n\mathbb{Z}=x+y+n\mathbb{Z}=[x+y]_{\mathbb{Z}/n\mathbb{Z}}$.

Für die Wohldefiniertheit der Multiplikation ist analog $[x']=[x]\wedge[y']=[y]\implies[x'y']=[xy]$ zu zeigen. Dies ist aber ebenfalls gegeben, weil Teilbarkeit von $x'-x$ und $y'-y$ durch n die Teilbarkeit von $x'y'-xy=(x'-x)y'+x(y'-y)$ durch n nach sich zieht. □

Den Beweis von Teil (III.1.13) des Satzes gehen wir erst nach einigen Beispielen hierzu an:

Beispiele III.1.14.

- Für Restklassen in $\mathbb{Z}/2\mathbb{Z}$ und $\mathbb{Z}/7\mathbb{Z}$ gelten

$$[3]+[-5]=[-2] \text{ in } \mathbb{Z}/2\mathbb{Z},$$
$$[1]+[1]=[0] \text{ in } \mathbb{Z}/2\mathbb{Z},$$
$$[3]\cdot[5]=[1] \text{ in } \mathbb{Z}/7\mathbb{Z}.$$

Dies bedeutet letztlich dasselbe wie

$$3-5=-2 \mod 2,\ 1+1=0 \mod 2,\ 3\cdot 5=1 \mod 7,$$

was wir auch in Abschnitt II.4.2 schon hinschreiben konnten. Jetzt verstehen wir aber genauer, wie weit diese Art des Rechnens trägt.

- Da $\mathbb{Z}/n\mathbb{Z}$ aus den n Restklassen $[0]=n\mathbb{Z}$, $[1]=1+n\mathbb{Z}$, $[2]=2+n\mathbb{Z}$, …, $[n-2]=(n-2)+n\mathbb{Z}$, $[n-1]=(n-1)+n\mathbb{Z}$ besteht (was man formal durch Division mit Rest einsieht), lassen sich die Addition und die Multiplikation auf $\mathbb{Z}/n\mathbb{Z}$ für festes $n\in\mathbb{N}$ durch Verknüpfungstafeln vollständig angegeben. Wir zeigen hier die Verknüpfungstafeln

  – für $\mathbb{Z}/2\mathbb{Z}$:

$+$	$[0]$	$[1]$
$[0]$	$[0]$	$[1]$
$[1]$	$[1]$	$[0]$

$\cdot$	$[0]$	$[1]$
$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$

  – für $\mathbb{Z}/3\mathbb{Z}$:

$+$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[1]$	$[2]$
$[1]$	$[1]$	$[2]$	$[0]$
$[2]$	$[2]$	$[0]$	$[1]$

$\cdot$	$[0]$	$[1]$	$[2]$
$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$
$[2]$	$[0]$	$[2]$	$[1]$

  – für $\mathbb{Z}/4\mathbb{Z}$:

$+$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[1]$	$[2]$	$[3]$
$[1]$	$[1]$	$[2]$	$[3]$	$[0]$
$[2]$	$[2]$	$[3]$	$[0]$	$[1]$
$[3]$	$[3]$	$[0]$	$[1]$	$[2]$

$\cdot$	$[0]$	$[1]$	$[2]$	$[3]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[1]$	$[0]$	$[1]$	$[2]$	$[3]$
$[2]$	$[0]$	$[2]$	$[0]$	$[2]$
$[3]$	$[0]$	$[3]$	$[2]$	$[1]$

  Man erkennt insbesondere, dass die Multiplikation zwar auf $\mathbb{Z}/2\mathbb{Z}^\times$ und $\mathbb{Z}/3\mathbb{Z}^\times$ eine Verknüpfung ist, aber wegen $[2]\cdot[2]=[0]$ nicht auf $\mathbb{Z}/4\mathbb{Z}\setminus\{[0]\}$.

Beweis von Teil (2) des Satzes. Da die Wohldefiniertheit der Verknüpfungen schon gesichert ist, ergeben sich die für alle $n\in\mathbb{N}$ gültigen (Halb-)Gruppeneigenschaften von $(\mathbb{Z}/n\mathbb{Z},+)$ und $(\mathbb{Z}/n\mathbb{Z},\cdot)$ sofort aus denen von $(\mathbb{Z},+)$ und $(\mathbb{Z},\cdot)$.

Ist $p\in\mathbb{N}$ *keine* Primzahl, so ist zunächst $\mathbb{Z}/1\mathbb{Z}\setminus\{[0]\}=\emptyset$ im Fall $p=1$ wegen Nicht-Existenz des neutralen Elements keine Gruppe. In den anderen Nicht-Primzahl-Fällen können wir $n=xy$ mit $x,y\in\{2,3,\dots,n-1\}$ schreiben. Da für $[x],[y]\in\mathbb{Z}/n\mathbb{Z}\setminus\{[0]\}$ dann $[x]\cdot[y]=[n]=[0]\notin\mathbb{Z}/n\mathbb{Z}\setminus\{[0]\}$ eintritt, ist $\cdot$ nur auf $\mathbb{Z}/n\mathbb{Z}$, aber nicht auf $\mathbb{Z}/n\mathbb{Z}\setminus\{[0]\}$ wohldefinierte Verknüpfung, und $(\mathbb{Z}/n\mathbb{Z}\setminus\{[0]\},\cdot)$ ist keine Gruppe.

Es bleibt also die Gruppeneigenschaft von $(\mathbb{Z}/p\mathbb{Z}^\times,\cdot)$ im Fall einer Primzahl $p\in\mathbb{P}$ zu verifizieren. Dazu zeigen wir für jede Restklasse $[y]\in\mathbb{Z}/p\mathbb{Z}^\times$ mit $y\in\mathbb{Z}$ zuerst die Existenz eines multiplikativ Inversen in $\mathbb{Z}/p\mathbb{Z}$. Wir schreiben dazu $[y]=[x]$ mit Hilfe eines Repräsentanten $x\in\{1,2,\dots,p-1\}$, der sich als Rest bei der Division von y durch p ergibt (wobei der Rest 0 ausgeschlossen ist, weil andernfalls $[y]=[0]$ wäre). Wir zeigen nun indirekt, dass $[0]$, $[x]$, $[2x]$, …, $[(p-1)x]$, also mit anderen Worten die Restklassen $[nx]$ mit $n\in\{0,1,2,\dots,p-1\}$, in $\mathbb{Z}/p\mathbb{Z}$ alle verschieden sind. Andernfalls müsste nämlich $[kx]=[\ell x]$ für $k,\ell\in\{0,1,2,\dots,p-2,p-1\}$ mit $k<\ell$ gelten, und für $m:=\ell-k\in\{1,2,3,\dots,p-2,p-1\}$ bekämen wir $[mx]=[0]$, mit anderen Worten also $p|(mx)$. Mit der Eindeutigkeit der Primfaktorzerlegung (deren Beweis bisher noch nicht vorgestellt wurde, aber vielleicht später noch nachgetragen wird) ergibt sich, dass die Primzahl p ein Primfaktor von m oder einer von x sein müsste, also $p|m$ oder $p|x$ gälte. Beides ist aber ausgeschlossen, weil m und x in $\{0,1,2,\dots,p-1\}$ sind. Damit sind $[nx]$ mit $n\in\{0,1,2,\dots,p-1\}$ in der Tat p verschiedene Elemente von $\mathbb{Z}/p\mathbb{Z}$. Da $\mathbb{Z}/p\mathbb{Z}$ genau die p Elemente $[0]$, $[1]$, $[2]$, …, $[p-1]$ enthält, muss somit $[nx]=[1]$ für ein $n\in\{1,2,\dots,p-1\}$ gelten (wobei $n=0$ wegen $[0x]=[0]\neq[1]$ ausgeschlossen ist). Nach Definition der Multiplikation gilt somit auch $[n]\cdot[x]=[1]=[x]\cdot[n]$ für $[n]\in\mathbb{Z}/p\mathbb{Z}^\times$, also ist $[n]$ das gesuchte Inverse zu $[y]=[x]$. An dieser Stelle können wir nun mit einem kurzen allgemeinen Argument folgern, dass für beliebige $[y],[z]\in\mathbb{Z}/p\mathbb{Z}^\times$ auch $[y]\cdot[z]\neq[0]$ ist, weshalb die Multiplikation überhaupt eine Verknüpfung auf $\mathbb{Z}/p\mathbb{Z}^\times$ ist: Wäre $[y]\cdot[z]=[0]$, so ergäbe sich mit dem Inversen $[n]$ zu $[y]$ der Widerspruch

$[z] = [n]\cdot[y]\cdot[z] = [n]\cdot[0] = [0]$. Mit der gerade begründeten Verknüpfungseigenschaft und der Existenz der Inversen ist dann klar, dass $(\mathbb{Z}/p\mathbb{Z}^\times, \cdot)$ eine abelsche Gruppe ist (denn die restlichen benötigten Eigenschaften vererben sich von $(\mathbb{Z}/p\mathbb{Z}, \cdot)$). □

Anhand dieser Beispiele diskutieren wir kurz einige allgemeine Begriffe bei Gruppen:

Definition III.1.15. Es sei $(G, *)$ ein Monoid mit neutralem Element e und $g \in G$.
(I) Wir erklären *Potenzen* g^n von g mit Exponent $n \in \mathbb{N}$ rekursiv durch $g^0 := e$ und $g^{n+1} := g * g^n$ für $n \in \mathbb{N}_0$.
(II) Gilt $g^2 = g$, so heißt g *idempotent*.

Man beachte, dass diese Definition prinzipiell auch auf additive Gruppen wie $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{Z}/n\mathbb{Z}, +)$ angewandt werden kann. Dort entsprechen die Potenzen allerdings Vielfachen, die man in der üblichen Schreibweise als nx beziehungsweise $[nx]$ notiert.

Bemerkungen III.1.16.
- Ist g idempotent, so folgt mit vollständiger Induktion $g^n = g$ für alle $n \in \mathbb{N}$.
- Ein neutrales Element ist immer idempotent und weitere Idempotente, falls existent, sind nicht invertierbar (denn für jedes invertierbare Idempotente $g \in G$ gilt $g = g^{-1} * g^2 = g^{-1} * g = e$). In einer Gruppe ist daher das neutrale Element immer das einzige idempotente Element.
- In den Monoiden $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$ ist neben dem neutralen Element 1 einzig 0 idempotent.
- In $(\mathbb{Z}/n\mathbb{Z}, \cdot)$ sind $[1]$ und $[0]$ immer idempotent, für $n = 6$ ist $[3]$ mit $[3]\cdot[3] = [9] = [3]$ in $(\mathbb{Z}/6\mathbb{Z}, \cdot)$ ein weiteres Beispiel eines Idempotenten. Finden Sie weitere nicht-triviale Idempotente für andere Werte von n?

Wir führen eine grundlegende Klasse von Gruppen mit besonders gutartiger und einfacher Struktur ein:

Definition III.1.17. Eine Gruppe $(G, *)$ mit neutralem Element e heißt *zyklisch der Ordnung* $n \in \mathbb{N}$, wenn $G = \{e, g, \ldots, g^{n-1}\}$ für ein $g \in G$ mit $g^n = e \neq g^k$ für alle $k \in \{1, 2, \ldots, n-1\}$ gilt. Ein solches Element g heißt ein *Erzeuger* der Gruppe $(G, *)$.

Bemerkungen III.1.18.
- Für einen Erzeuger g in einer zyklischen Gruppe $(G, *)$ der Ordnung n sind e, g, g^2, g^3, …, g^{n-1} alle verschieden (weil aus $g^\ell = g^k$ für $0 \leq k < \ell < n$ schon $g^{\ell-k} = e$ folgt). Insbesondere ist daher $|G| = n$.
- Es kann in einer zyklischen Gruppe mehrere Erzeuger geben. Ein konkretes Beispiel folgt unten.
- Zyklische Gruppen sind stets abelsch, denn in der Darstellung mit einem Erzeuger g erhält man die Kommutativität $g^k g^\ell = g^{k+\ell} = g^\ell g^k$ für alle $k, \ell \in \{0, 1, 2, \ldots, n-1\}$.

Beispiele III.1.19.
- Die Gruppe $(\mathbb{Z}/7\mathbb{Z}^\times, \cdot)$ ist zyklisch der Ordnung 6 mit genau $[3]$ und $[5] = [3]^{-1}$ als Erzeugern, denn $[3]^1 = [3]$, $[3]^2 = [2]$, $[3]^3 = [6]$, $[3]^4 = [4]$, $[3]^5 = [5]$, $[3]^6 = [1]$ und $[5]^1 = [5]$, $[5]^2 = [4]$, $[5]^3 = [6]$, $[5]^4 = [2]$, $[5]^5 = [3]$, $[5]^6 = [1]$ geben jeweils alle Elemente von $\mathbb{Z}/7\mathbb{Z}^\times$, während für die anderen vier Elemente $[1]^1 = [2]^3 = [4]^3 = [6]^2 = [1]$ eintritt.
- Die Gruppe $(\mathbb{Z}/4\mathbb{Z}, +)$ ist zyklisch der Ordnung 4 mit genau $[1]$ und $[3]$ als Erzeugern, denn $[1]$, $[1] + [1] = [2]$, $[1] + [2] = [3]$, $[3] + [1] = [0]$ und $[3]$, $[3] + [3] = [2]$, $[3] + [2] = [1]$, $[3] + [1] = [0]$ sind jeweils alle Elemente von $\mathbb{Z}/4\mathbb{Z}$, während $[0]$ die triviale Gruppe erzeugt und $[2] + [2] = [0]$ gilt.
- Allgemein gilt:
  - Für jedes $n \in \mathbb{N}$ ist die *additive Gruppe* $(\mathbb{Z}/n\mathbb{Z}, +)$ des Restklassenrings $\mathbb{Z}/n\mathbb{Z}$ zyklisch der Ordnung n mit Erzeuger $[1]$.
    (Begründung: Mit $[1\cdot1] = [1]$, $[2\cdot1] = [2]$, …, $[(n-1)\cdot1] = [n-1]$, $[n\cdot1] = [0]$ erhalten wir alle Elemente von $\mathbb{Z}/n\mathbb{Z}$.)
  - In $(\mathbb{Z}/p\mathbb{Z}, +)$ mit einer Primzahl $p \in \mathbb{P}$ sind sogar alle Elemente außer $[0]$ Erzeuger. (Begründung: Dies ergibt sich aus dem Beweis des letzten Satzes, in dem wir gesehen hatten, dass für jedes $x \in \mathbb{Z}/p\mathbb{Z}^\times = \mathbb{Z}/p\mathbb{Z}\backslash\{[0]\}$ die p Restklassen $[x]$, $[2x]$, $[3x]\ldots$, $[(p-1)x]$, $[px] = [0]$ alle verschieden sind und daher die p Elemente von $\mathbb{Z}/p\mathbb{Z}$ sein müssen.)

– Für jede Primzahl $p \in \mathbb{P}$ ist die *multiplikative Gruppe* $(\mathbb{Z}/p\mathbb{Z}^\times, \cdot)$ des Restklassenrings $\mathbb{Z}/p\mathbb{Z}$ zyklisch der Ordnung $p-1$. Das ergibt sich aus dem Kleinen Satz von Fermat, den wir später zeigen werden.

Als nächstes Thema behandeln wir *Permutationen*. Mit solchen formalisiert man in der Mathematik oft Veränderungen einer Reihenfolge. Wir erwähnen Permutationen an dieser Stelle aber vor allem wegen Permutationsgruppen wie den *symmetrischen Gruppen* Σ_n und den alternierenden Gruppen A_n.

Definition III.1.20.

(I) Eine *Permutation* einer endlichen Menge $\mathcal{X}$ ist eine bijektive Selbstabbildung von $\mathcal{X}$.

(II) Für $n \in \mathbb{N}$ schreiben wir Σ_n für die Menge der Permutationen von $\{1, 2, \ldots, n\}$ und nennen $(\Sigma_n, \circ)$ die *symmetrische Gruppe* auf n Elementen n.

(III) Permutationen $\pi \in \Sigma_n$ mit $n \in \mathbb{N}$ notieren wir manchmal durch ihre Wertetabelle

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n-1 & n \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(n-1) & \pi(n) \end{pmatrix} \tag{III.1.1}$$

Bemerkungen III.1.21.

- Beachten Sie, dass die Permutation in (III.1.1) mit der Permutation übereinstimmt, die durch
$$\begin{pmatrix} \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n-1) & \sigma(n) \\ \pi(\sigma(1)) & \pi(\sigma(2)) & \pi(\sigma(3)) & \ldots & \pi(\sigma(n-1)) & \pi(\sigma(n)) \end{pmatrix}$$
beschrieben wird mit einer beliebigen Permutation $\sigma \in \Sigma_n$.
- Man kann die Betrachtung von Permutationen immer auf den Modellfall von Σ_n über der Grundmenge $\{1, 2, \ldots, n\}$ reduzieren, da jede endliche Menge $\mathcal{X}$ durch eine Bijektion mit $\{1, 2, \ldots, n\}$ für $n := |\mathcal{X}|$ identifiziert werden kann.
- Dass $(\Sigma_n, \circ)$ (und allgemein die Menge der Bijektionen $\mathcal{X} \to \mathcal{X}$ mit der Komposition $\circ$) tatsächlich eine Gruppe ist, haben wir früher schon beobachtet. Diese Gruppe ist *nur* für $n \in \{1, 2\}$ (beziehungsweise $|\mathcal{X}| \in \{1, 2\}$) abelsch.
- Mit vollständiger Induktion kann man zeigen, dass es genau $n!$ Permutationen einer n-elementigen Menge gibt. Es gilt also $|\Sigma_n| = n!$.

Beispiel III.1.22. Die Permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 1 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 5 & 1 & 3 \\ 1 & 4 & 5 & 2 & 3 \end{pmatrix} \in \Sigma_5$$

ist die Bijektion $\pi\colon \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4, 5\}$ mit $1 \mapsto 2$, $2 \mapsto 4$, $3 \mapsto 3$, $4 \mapsto 1$, $5 \mapsto 5$.

Definition III.1.23. Es sei π eine Permutation einer endlichen Menge $\mathcal{X}$.

(I) Wir nennen π die *Transposition* von $a, b \in \mathcal{X}$ mit $a \neq b$ in $\mathcal{X}$, wenn $\pi(a) = b$, $\pi(b) = a$ und $\pi(x) = x$ für alle $x \in \mathcal{X} \setminus \{a, b\}$ gilt.

(II) Wir nennen π einen *Zykel* der Länge $\ell \in \mathbb{N}$ oder einen ℓ-Zykel, wenn es ein $a \in \mathcal{X}$ mit $\pi^\ell(a) = a \neq \pi^k(a)$ für alle $k \in \{1, 2, \ldots, \ell-1\}$ und $\pi(x) = x$ für alle $x \in \mathcal{X} \setminus \{a, \pi(a), \pi^2(a), \ldots, \pi^{\ell-1}(a)\}$ gibt. Für solche Zykel wird die Zykelschreibweise $\pi = (a, \pi(a), \pi^2(a), \ldots, \pi^{\ell-1}(a))$ benutzt. Manchmal läßt man die Kommata weg.

Bemerkungen III.1.24.

- Ein Zykel der Länge 1 ist die Identität, ein Zykel der Länge 2 ist eine Transposition.
- Eine Transposition π ist selbstinvers. Für einen Zykel π der Länge ℓ in $\mathcal{X}$ gilt $\pi^\ell = \mathrm{id}_\mathcal{X}$.

Beispiel III.1.25. Was ist die Zykelschreibweise der Permutation aus Beispiel III.1.22?

Beispiele III.1.26. Wir betrachten die gut überschaubaren Beispiele Σ_2 und Σ_3. Beachten Sie, dass $\Sigma_1 = \{\mathrm{id}\}$ gilt und dass Σ_4 schon die Mächtigkeit $|\Sigma_4| = 4! = 24$ hat.

- Die $2! = 2$ Elemente der symmetrischen Gruppe $(\Sigma_2, \circ)$ sind die Identität id und eine Transposition τ mit den Darstellungen
$$\mathrm{id} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \text{ und } \tau = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$
Offensichtlich ist τ selbstinvers und $(\Sigma_2, \circ)$ ist zyklisch der Ordnung 2. In Zykelschreibweise kann man $\tau = (1,2) = (2,1)$ ausdrücken.
- Die $3! = 6$ Elemente der symmetrischen Gruppe $(\Sigma_3, \circ)$ sind die Identität id, drei Transpositionen τ_1, τ_2, τ_3 und zwei 3-Zykel σ_1, σ_2 mit den Darstellungen
$$\mathrm{id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad \tau_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$$
$$\tau_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$
Bezüglich der Komposition, der Verknüpfung der Gruppe, hängen die Elemente zum Beispiel durch
$$\sigma_1 = \tau_2 \circ \tau_1 = \tau_3 \circ \tau_2 = \tau_1 \circ \tau_3 = \sigma_2^2 = \sigma_2^{-1}$$
zusammen. In Zykelschreibweise kann man $\tau_1 = (1,2) = (2,1)$, $\tau_2 = (1,3) = (3,1)$ und $\sigma_1 = (1,2,3) = (2,3,1) = (3,1,2)$ ausdrücken. Wie sehen die restlichen Zykelbeschreibungen aus?

Zur Einführung einer weiteren Gruppe von Permutationen und für wichtige Anwendungen später in der linearen Algebra brauchen wir:

Satz III.1.27. *Sei π eine Permutation einer endlichen Menge $\mathcal{X}$.*

(a) *Dann kann π als Komposition $\pi = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_{m-1} \circ \tau_m$ einer endlichen Zahl $m \in \mathbb{N}_0$ von Transpositionen $\tau_1, \tau_2, \ldots, \tau_{m-1}, \tau_m$ in $\mathcal{X}$ geschrieben werden.*

(b) *Die Parität (gerade oder ungerade) der Zahl m in (a) oder mit anderen Worten die Restklasse $[m] \in \mathbb{Z}/2\mathbb{Z}$ ist durch π eindeutig bestimmt.*

Definition III.1.28.

(I) Wir nennen eine Permutation π einer endlichen Menge $\mathcal{X}$ eine *gerade beziehungsweise ungerade Permutation*, wenn die Parität von m in Teil (b) des Satzes gerade beziehungsweise ungerade ist. Das *Vorzeichen* $\mathrm{sign}(\pi) \in \{1, -1\}$ einer Permutation π erklären wir für gerade Permutationen π zu 1, für ungerade Permutationen π zu -1.

(II) Für $n \in \mathbb{N}$ setzen wir
$$A_n := \{\pi \in \Sigma_n \mid \pi \text{ ist gerade}\} = \{\pi \in \Sigma_n \mid \mathrm{sign}(\pi) = 1\},$$
und nennen $(A_n, \circ)$ die *alternierende Gruppe* auf n Elementen.

Bemerkungen III.1.29.

(a) Für Permutationen π und σ von $\mathcal{X}$ gilt $\mathrm{sign}(\pi^{-1}) = \mathrm{sign}(\pi)$ und $\mathrm{sign}(\sigma \circ \pi) = \mathrm{sign}(\sigma)\mathrm{sign}(\pi)$.
(Begründung für letzteres: Sind σ bzw. π Kompositionen von ℓ bzw. m Transpositionen, so ist $\sigma \circ \pi$ Komposition von $\ell+m$ Transpositionen. Im Fall $\mathrm{sign}(\sigma) = \mathrm{sign}(\pi)$ haben ℓ, m gleiche Parität, $\ell+m$ ist gerade, und es gilt $\mathrm{sign}(\sigma \circ \pi) = 1 = \mathrm{sign}(\sigma)\mathrm{sign}(\pi)$. Im Fall $\mathrm{sign}(\sigma) = -\mathrm{sign}(\pi)$ dagegen haben ℓ, m verschiedene Parität, $\ell+m$ ist ungerade, und es gilt $\mathrm{sign}(\sigma \circ \pi) = -1 = \mathrm{sign}(\sigma)\mathrm{sign}(\pi)$.)

(b) Insbesondere folgt aus Bemerkung (a), dass für $\pi, \sigma \in A_n$ auch $\pi^{-1} \in A_n$ und $\sigma \circ \pi \in A_n$ gelten. Erst dies (zusammen mit früheren Beobachtungen) stellt sicher, dass $(A_n, \circ)$ in der Tat eine Gruppe ist. Die Gruppe $(A_n, \circ)$ ist *nur* für $n \in \{1, 2, 3\}$ abelsch.

(c) Für $n \in \mathbb{N}\backslash\{1\}$ enthält die Teilmenge A_n von Σ_n die Hälfte der Permutationen aus Σ_n, es gilt also $|A_n| = \frac{1}{2}|\Sigma_n| = \frac{n!}{2}$.
(Begründung: Sei $\tau \in \Sigma_n$ eine beliebige Transposition, die als solche insbesondere selbstinvers mit $\mathrm{sign}(\tau) = -1$ ist. Mit der vorausgehenden Bemerkung (1) und den Gruppeneigenschaften von $(\Sigma_n, \circ)$ folgt, dass
$$A_n \to \Sigma_n\backslash A_n,\ \pi \mapsto \tau \circ \pi$$
eine wohldefinierte Bijektion von A_n nach $\Sigma_n\backslash A_n$ ist. Damit gilt $|A_n| = |\Sigma_n\backslash A_n| = |\Sigma_n| - |A_n|$, und durch Auflösen ergibt sich $|A_n| = \frac{1}{2}|\Sigma_n|$.)

Beispiele III.1.30. Wir betrachten die Beispiele A_2, A_3 und A_4. Beachten Sie, dass $A_1 = \{\mathrm{id}\}$ und $A_2 = \{\mathrm{id}\}$ und dass A_5 schon Mächtigkeit $|A_5| = \frac{5!}{2} = 60$ hat.

- Die $\frac{3!}{2} = 3$ Elemente der alternierenden Gruppe $(A_3, \circ)$ sind die Identität id und die beiden als Elemente von $(\Sigma_3, \circ)$ schon betrachteten 3-Zykel σ_1 und σ_2. Es gelten $\sigma_1^2 = \sigma_2$, $\sigma_2^2 = \sigma_1$ und $\sigma_1 \circ \sigma_2 = \mathrm{id} = \sigma_2 \circ \sigma_1$. Insbesondere ist $(A_3, \circ)$ zyklisch der Ordnung 3.
- Die $\frac{4!}{2} = 12$ Elemente der alternierenden Gruppe $(A_4, \circ)$ sind die Identität id, acht 3-Zykel η_1, η_2, η_3, η_4, η_5, η_6, η_7, η_8 und drei Kompositionen ϑ_1, ϑ_2, ϑ_3 von je zwei Transpositionen mit den Darstellungen

$$\mathrm{id} = \begin{pmatrix}1&2&3&4\\1&2&3&4\end{pmatrix}, \quad \eta_1 = \begin{pmatrix}1&2&3&4\\2&3&1&4\end{pmatrix}, \quad \eta_2 = \begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix}, \quad \eta_3 = \begin{pmatrix}1&2&3&4\\2&4&3&1\end{pmatrix},$$
$$\eta_4 = \begin{pmatrix}1&2&3&4\\4&1&3&2\end{pmatrix}, \quad \eta_5 = \begin{pmatrix}1&2&3&4\\3&2&4&1\end{pmatrix}, \quad \eta_6 = \begin{pmatrix}1&2&3&4\\4&2&1&3\end{pmatrix}, \quad \eta_7 = \begin{pmatrix}1&2&3&4\\1&3&4&2\end{pmatrix},$$
$$\eta_8 = \begin{pmatrix}1&2&3&4\\1&4&2&3\end{pmatrix}, \quad \vartheta_1 = \begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}, \quad \vartheta_2 = \begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix}, \quad \vartheta_3 = \begin{pmatrix}1&2&3&4\\4&3&2&1\end{pmatrix}.$$

  Dass $(A_4, \circ)$ nicht abelsch ist, sieht man beispielsweise an $\eta_1 \circ \vartheta_1 = \eta_5$ und $\vartheta_1 \circ \eta_1 = \eta_8$.

Es verbleibt, den Beweis des letzten Satzes durchzuführen:

Beweis von Satz III.1.27. Für den ersten Teil argumentieren wir per vollständiger Induktion nach $|\mathcal{X}| \in \mathbb{N}_0$:

Den Induktionsanfang für $|\mathcal{X}| = 0$ erledigt die Beobachtung, dass dann $\mathcal{X} = \emptyset$ ist, dass die (formal) bijektive leere Abbildung die einzige Abbildung $\emptyset \to \emptyset$ und die einzige Permutation von $\emptyset$ ist, und, dass die leere Abbildung (formal) die Komposition von 0 Transpositionen ist.

Für den Induktionsschluss von $|\mathcal{X}|-1$ zu $|\mathcal{X}|$ sei $|\mathcal{X}| \in \mathbb{N}$ und π Permutation von $\mathcal{X}$. Wir wählen $x_1 \in \mathcal{X}$. Im Fall $\pi(x_1) = x_1$ sei $\tau_1 := \mathrm{id}_{\mathcal{X}}$, im Fall $\pi(x_1) \neq x_1$ sei $\tau_1 := (x_1, \pi(x_1))$ die Transposition von $\mathcal{X}$ mit $\tau_1(x_1) = \pi(x_1)$ und $\tau_1(\pi(x_1)) = x_1$. In beiden Fällen folgt $(\tau_1 \circ \pi)(x_1) = x_1$, und wegen Bijektivität erhalten wir $(\tau_1 \circ \pi)(x) \in \mathcal{X} \setminus \{x_1\}$ für alle $x \in \mathcal{X} \setminus \{x_1\}$. Wir können daher die Permutation $\widetilde{\pi}$ von $\mathcal{X} \setminus \{x_1\}$ mit $\widetilde{\pi}(x) = (\tau_1 \circ \pi)(x)$ für alle $x \in \mathcal{X} \setminus \{x_1\}$ bilden und $\widetilde{\pi}$ nach Induktionsannahme für ein $m \in \mathbb{N}$ als Komposition von $m-1$ Transpositionen schreiben. Wir erweitern diese zu $m-1$ Transpositionen $\tau_2, \tau_3, \ldots, \tau_m$ auf $\mathcal{X}$ mit $\tau_2(x_1) = \tau_3(x_1) = \ldots \tau_m(x_1) = x_1$ und bekommen $\tau_1 \circ \pi = \tau_2 \circ \tau_3 \circ \ldots \circ \tau_{m-1} \circ \tau_m$. Da τ_1 selbstinvers ist (Eigenschaft von id_X und jeder Transposition), erhalten wir

$$\pi = \tau_1^{-1} \circ \tau_2 \circ \tau_3 \circ \ldots \circ \tau_{m-1} \circ \tau_m = \tau_1 \circ \tau_2 \circ \tau_3 \circ \ldots \circ \tau_{m-1} \circ \tau_m.$$

Damit ist π die Komposition der Transpositionen $\tau_1, \tau_2, \ldots, \tau_{m-1}, \tau_m$ (im Fall $\pi(x_1) = x_1$ nach Weglassen von $\tau_1 = \mathrm{id}_{\mathcal{X}}$), und die Induktionsbehauptung ist gezeigt.

Für den zweiten Teil nehmen wir der Einfachheit halber $\mathcal{X} = \{1, 2, \ldots, n\}$ an. (Wir könnten für allgemeines $\mathcal{X}$ eine beliebige Totalordnung auf $\mathcal{X}$ einführen und damit analog argumentieren). Wir betrachten die Menge

$$\mathrm{FS}(\pi) := \{(i,j) \in \{1, 2, \ldots, n\}^2 \mid i < j, \pi(i) > \pi(j)\}$$

der sogenannten Fehlstände einer Permutation $\pi \in \Sigma_n$ und argumentieren mit der Zahl der Fehlstände $|\mathrm{FS}(\pi)| \in \mathbb{N}_0$. Für diesen Beweis entscheidend ist nun:

> *Behauptung:* Für eine beliebige Permutation $\pi \in \Sigma_n$ und eine Transposition $\tau \in \Sigma_n$ haben die Zahlen der Fehlstände $|\mathrm{FS}(\tau \circ \pi)|$ und $|\mathrm{FS}(\pi)|$ unterschiedliche Parität.

Zum Nachweis dieser Behauptung benutzen wir zunächst, dass die Transposition τ nach Definition die Form $\tau = (k, \ell)$ mit $k, \ell \in \{1, 2, \ldots, n\}$, $k \neq \ell$, hat. Da $(k, \ell) = (\ell, k)$ ist, behandeln wir ohne Einschränkung den Fall $k < \ell$. Für ein Paar $(i, j) \in \{1, 2, \ldots, n\}^2$ mit $i < j$ und $\{\pi(i), \pi(j)\} =: \{x, y\}$ mit $x < y$ unterscheiden wir folgende Fälle:

(1) Fall $x, y \notin \{k, \ell\}$: Dann ist $\tau(x) = x$, $\tau(y) = y$, also $\tau(\pi(i)) = \pi(i)$, $\tau(\pi(j)) = \pi(j)$, und es folgt

$$(i,j) \in \mathrm{FS}(\tau \circ \pi) \iff (i,j) \in \mathrm{FS}(\pi).$$

(2) Fall $x < k$, $y \in \{k, \ell\}$ sowie Fall $x \in \{k, \ell\}$, $\ell < y$: Neben $x < y$ gilt dann $\tau(x) < \tau(y)$. Es folgt

$$\tau(\pi(i)) > \tau(\pi(j)) \Longleftrightarrow \pi(i) > \pi(j)$$

und

$$(i,j) \in \mathrm{FS}(\tau \circ \pi) \Longleftrightarrow (i,j) \in \mathrm{FS}(\pi).$$

(3) Fall $x = k < y < \ell$ sowie Fall $k < x < \ell = y$: Neben $x < y$ gilt dann $\tau(x) > \tau(y)$ und es folgt

$$\tau(\pi(i)) > \tau(\pi(j)) \Longleftrightarrow \pi(i) < \pi(j)$$

und

$$(i,j) \in \mathrm{FS}(\tau \circ \pi) \Longleftrightarrow (i,j) \notin \mathrm{FS}(\pi).$$

(4) Fall $x = k$, $y = \ell$: Dann ist $\tau(x) = y$, $\tau(y) = x$, also $\tau(\pi(i)) = \pi(j)$, $\tau(\pi(j)) = \pi(i)$, und es folgt $(i,j) \in \mathrm{FS}(\tau \circ \pi) \Longleftrightarrow (i,j) \notin \mathrm{FS}(\pi)$.

In den Fällen (1) und (2) entnehmen wir, dass Fehlstände (i,j) von π bei $\tau \circ \pi$ weiterbestehen. In den Fällen (3) und (4) treten bei $\tau \circ \pi$ gegenüber π Fehlstände hinzu oder fallen weg. Die Situation (3) tritt dabei (da es $\ell - k - 1$ natürliche Zahlen zwischen k und ℓ gibt) für genau $2(\ell - k - 1)$ Paare $(x,y) \in \{1, 2, \ldots, n\}^2$ mit $x < y$ und dementsprechend auch für $2(\ell - k - 1)$ Paare $(i,j) \in \{1, 2, \ldots, n\}^2$ mit $i < j$ ein. Die Situation (4) liegt für genau ein Paar (i,j) mit $i < j$ vor (nämlich das mit $\{\pi(i), \pi(j)\} = \{k, \ell\}$). Insgesamt ist damit $|\mathrm{FS}(\tau \circ \pi) \,\Delta\, \mathrm{FS}(\pi)| = 2(\ell - k - 1) + 1$ ungerade, wegen $|A| + |B| = |A \Delta B| + 2|A \cap B|$ ist auch $|\mathrm{FS}(\tau \circ \pi)| + |\mathrm{FS}(\pi)|$ ungerade, also muss von $|\mathrm{FS}(\tau \circ \pi)|$ und $|\mathrm{FS}(\pi)|$ eins gerade, eins ungerade sein. Dies zeigt die obige Behauptung.

Da $|\mathrm{FS}(\mathrm{id})| = 0$ für die Identität $\mathrm{id} \in \Sigma_n$ gerade ist und für Transpositionen $|\mathrm{FS}(\tau_1)|$ ungerade ist, kann man mit vollständiger Induktion zeigen, dass die Komposition einer geraden beziehungsweise ungeraden Anzahl von Transpositionen stets gerades beziehungsweise ungerades Vorzeichen hat. Ist für $\pi \in \Sigma_n$ also $|\mathrm{FS}(\pi)|$ gerade, so kann $\pi = \tau_1 \circ \tau_2 \circ \ldots \circ \tau_{m-1} \circ \tau_m$ nur für gerades $m \in \mathbb{N}_0$ gelten. Ist $|\mathrm{FS}(\pi)|$ ungerade, so kann selbiges nur für ungerades $m \in \mathbb{N}_0$ gelten. Dies zeigt die behauptete Eindeutigkeit der Parität von m. □

III.2. Ringe und Körper

Nachdem wir Gruppen als algebraische Strukturen mit *einer* Verknüpfung kennengelernt haben, kommen wir nun zu algebraischen Strukturen mit *zwei* Verknüpfungen, die in vielen Beispielen durch Addition und Multiplikation gegeben sind. Sie kennen zum Beispiel den Ring der ganzen Zahlen $(\mathbb{Z}, +, .)$, bei dem die Addition und die Multiplikation verträglich sind in dem Sinne, dass Distributivgesetze gelten:

Definition III.2.1.

(a) Ein *Ring* ist ein Tripel $(R, +, \cdot)$ bestehend aus einer Menge R und Verknüpfungen $+$ und $\cdot$ auf R, so dass
- $(R, +)$ eine abelsche Gruppe ist,
- $(R, \cdot)$ ein Monoid ist
- und folgende *Distributivgesetze* gelten:

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z),$$
$$(x + y) \cdot z = (x \cdot z) + (y \cdot z), \qquad \text{für alle } x, y, z \in R.$$

(b) Bezüglich der *Addition* $+$ beziehungsweise in der *additiven Gruppe* $(R, +)$ eines Rings $(R, +, \cdot)$ bezeichnet man das neutrale Element als die *Null* 0 oder das *Nullelement* 0_R von $(R, +, \cdot)$ und das Inverse zu $x \in R$ als das *additiv Inverse* $-x \in R$ zu x.

(c) Bezüglich der *Multiplikation* $\cdot$ eines Rings $(R, +, \cdot)$ bezeichnet man das neutrale Element als die *Eins* 1 oder das *Einselement* 1_R von $(R, +, \cdot)$, ein invertierbares Element $x \in R$ als eine *Einheit* von R und sein Inverses als das *multiplikativ Inverse* $x^{-1} \in R$ zu x.

(d) Ein Ring heißt *kommutativ*, wenn neben seiner additiven Gruppe $(R, +)$ auch seine multiplikative Halbgruppe $(R, \cdot)$ kommutativ ist.

Bemerkungen III.2.2. Es sei $(R, +, \cdot)$ ein Ring.

- Wir verwenden beim Rechnen mit Elementen von R dieselben Konventionen zur Notationsvereinfachung wie bei Zahlen. Konkret gehören dazu das Weglassen des Multiplikationspunkts ($xy := x \cdot y$ für $x, y \in R$), die Einführung der *Subtraktion* ($x-y := x+(-y)$ für $x, y \in R$), die Konvention *Punkt-vor Strich-Rechnung*, derzufolge etwa die Klammern auf den rechten Seiten der Distributivgesetze entfallen können, und die üblichen *Konventionen zur Einsparung von* aufgrund Assoziativität unnötigen *Klammern*. Zudem verwendet man für $x \in R$, $n \in \mathbb{N}_0$ die Notation x^n, für invertierbares x auch $x^{-n} := (x^{-1})^n$, für *Potenzen* der multiplikativen Halbgruppe.
- Ist $(R, +, \cdot)$ ein Ring, so gelten für alle $x, y \in R$ die Regeln

$$\begin{aligned} 0+x = x = x+0, \\ 1x = x = x1, \\ 0x = 0 = x0, \\ (-x)y = -(xy) = x(-y). \end{aligned}$$

  Wir können also in Zukunft ohne Mehrdeutigkeit $-xy$ notieren.

  (Begründungen: Die ersten beiden Regeln gelten per Definition des Null- und des Einselements. Die dritte Regel ergibt sich durch $0x = 0x+x-x = (0+1)x-x = 1x-x = x-x = 0$ und eine analoge Rechnung. Zum Nachweis der vierten Regel reichen wegen der Kommutativität der Addition die Rechnungen $xy+(-x)y = (x-x)y = 0y = 0$ und $xy+x(-y) = x(y-y) = x0 = 0$.)

Beispiele III.2.3. Aus den in Abschnitt III.1 betrachteten Beispielen von additiven und multiplikativen (Halb-)Gruppen ergeben sich nach Verifikation der Distributivgesetze Beispiele von Ringen:

(a) Der *Nullring* ist der (bis auf Umbenennung des Elements eindeutige) Ring $(\{0\}, +, \cdot)$ mit nur einem Element 1=0. Dies ist der einzige Ring mit $1 = 0$: Aus $1 = 0$ und der vorausgehenden Folgerung ergibt sich $x = 1x = 0x = 0$ für jedes Ringelement x.

(b) Die *ganzen Zahlen* $(\mathbb{Z}, +, \cdot)$, die *rationalen Zahlen* $(\mathbb{Q}, +, \cdot)$, die *reellen Zahlen* $(\mathbb{R}, +, \cdot)$ und die *Restklassenringe* $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ mit $n \in \mathbb{N}$ sind *kommutative Ringe*.

(c) *Produkt-Ringe:* Für jedes $n \in \mathbb{N}$ und jeden (kommutativen) Ring $(R, +, \cdot)$ ist auch $(R^n, +, \cdot)$ mit den komponentenweisen Verknüpfungen + und $\cdot$ ein (kommutativer) Ring mit Nullelement $0_{R^n} = (0_R, 0_R, \ldots, 0_R)$ und Einselement $1_{R^n} = (1_R, 1_R, \ldots, 1_R)$.

  Allgemeiner ist für jedes $n \in \mathbb{N}$ und (kommutative) Ringe $(R_1, +, \cdot)$, $(R_2, +, \cdot)$, …, $(R_n, +, \cdot)$ auch $(R_1 \times R_2 \times \ldots \times R_n, +, \cdot)$ mit den komponentenweisen Verknüpfungen ein (kommutativer) Ring.

(d) *Abbildungs-Ringe*: Für jede Menge $\mathcal{X}$ und jeden (kommutativen) Ring $(R, +, \cdot)$ ist $(\mathrm{Abb}(\mathcal{X}, R), +, \cdot)$ mit den punktweisen Verknüpfungen + und $\cdot$ ein (kommutativer) Ring.

  Dagegen ergibt $(\mathrm{Abb}(R), +, \circ)$ mit der Komposition $\circ$ für $R \neq \{0_R\}$ *keinen* Ring: Zwar ist $(\mathrm{Abb}(R), +)$ abelsche Gruppe, $(\mathrm{Abb}(R), \circ)$ ist Halbgruppe mit neutralem Element, und das „rechte" Distributivgesetz $(f+g) \circ h = f \circ h + g \circ h$ gilt für alle $f, g, h \in \mathrm{Abb}(R)$. Das „linke" Distributivgesetz $f \circ (g+h) = f \circ g + f \circ h$ gilt aber zum Beispiel für $f \equiv 1_R$, nicht, da dann $f \circ (g+h) \equiv 1_R$ verschieden von $f \circ g + f \circ h \equiv 1_R + 1_R$ ist (denn $R \neq \{0_R\}$ bedeutet $1_R \neq 1_R + 1_R$).

(e) *Mengen-Ringe:* Für jede Menge $\mathcal{X}$ ist $(\mathcal{P}(\mathcal{X}), \Delta, \cap)$ ein kommutativer Ring (mit Nullelement $\emptyset$, Einselement $\mathcal{X}$ und $M \Delta M = \emptyset$ für alle $M \in \mathcal{P}(\mathcal{X})$).

(f) Als weitere wichtige Beispiele von Ringen lernen wir später in diesem Abschnitt *Polynom-Ringe* kennen.

Bemerkungen III.2.4.

- In einem Ring $(R, +, \cdot)$ kann für eine endliche Indexmenge I und $a_i \in R$ das Summenzeichen $\sum_{i \in I} a_i$ und im kommutativen Fall auch das Produktzeichen $\prod_{i \in I} a_i$ wie in Abschnitt II.2 sinnvoll erklärt werden.
- In einem Ring $(R, +, \cdot)$ können wir Vielfache von 1 definieren, und zwar als $2_R := 1_R + 1_R \in R$, $3_R := 2_R + 1_R \in R$ und allgemein als $n_R := \sum_{i=1}^{n} 1_R \in R$, $(-n)_R := -(n_R) \in R$ für $n \in \mathbb{N}_0$.

  Sind die Elemente z_R mit $z \in \mathbb{Z}$ alle voneinander verschieden, so kann man $z \in \mathbb{Z}$ mit $z_R \in R$ identifizieren und so $\mathbb{Z}$ als Teilmenge von R auffassen. Man spricht in diesem Fall von einem *Ring*

der Charakteristik 0. Zum Beispiel haben $(\mathbb{B},+,\cdot)$, $(\mathbb{B}^n,+,\cdot)$, $(\mathrm{Abb}(\mathcal{X},\mathbb{B}),+,\cdot)$ mit $\mathbb{B}\in\{\mathbb{Z},\mathbb{Q},\mathbb{R}\}$ Charakteristik 0.

Sind andernfalls die z_R mit $z\in\mathbb{Z}$ *nicht* alle verschieden, so gibt es ein kleinstes $n\in\mathbb{N}$ mit $n_R=0_R$, man kann $[z]\in\mathbb{Z}/n\mathbb{Z}$ mit $z_R\in R$ identifizieren und so $\mathbb{Z}/n\mathbb{Z}$ als Teilmenge von R auffassen. In diesem Fall spricht man von einem *Ring der (endlichen) Charakteristik* $n\in\mathbb{N}$. Zum Beispiel haben $(\mathbb{Z}/n\mathbb{Z},+,\cdot)$, $((\mathbb{Z}/n\mathbb{Z})^k,+,\cdot)$, $(\mathrm{Abb}(\mathcal{X},\mathbb{Z}/n\mathbb{Z}),+,\cdot)$ Charakteristik n und $(\mathcal{P}(\mathcal{X}),\Delta,\cap)$ mit $\mathcal{X}\neq\emptyset$ Charakteristik 2. (Charakteristik 1 hat übrigens nur der Nullring.)

- Manche Autoren fordern bei Definition eines Rings nicht allgemein, dass ein neutrales Element 1 der multiplikativen Halbgruppe existieren muss, und bezeichnen einen Ring, bei dem dies doch der Fall ist, explizit als Ring mit Eins oder unitären Ring. Wir bleiben aber bei der obigen Konvention, gemäß der jeder Ring ein Ring mit Eins ist.

Definition III.2.5. Es sei $(R,+,\cdot)$ ein *kommutativer* Ring.

(a) Wir nennen $y\in R$ einen *Teiler* von $z\in R$ (in R) und notieren $y\mid z$ (in R), wenn ein $x\in R$ mit $xy=z$ existiert.

(b) Wir nennen $y\in R\setminus\{0\}$ einen *Nullteiler* (in R), wenn ein $x\in R\setminus\{0\}$ mit $xy=0$ existiert. Gibt es in $R\setminus\{0\}$ keinen Nullteiler, so nennen wir $(R,+,\cdot)$ *nullteilerfrei*.

(c) Ist $(R,+,\cdot)$ nullteilerfrei mit $R\neq\{0\}$, so sprechen wir von einem *Integritätsbereich*.

Bemerkung III.2.6. In einem kommutativen Ring $(R,+,\cdot)$ ist wegen $0y=0$ jedes Element $y\in R$ ein Teiler von 0. Dies macht ein $y\in R$ mit $y\neq 0$ aber noch nicht unbedingt zum Nullteiler, denn dafür ist $xy=0$ eben auch mit $x\neq 0$ erforderlich.

Beispiele III.2.7.

- Die ganzen Zahlen $(\mathbb{Z},+,\cdot)$, die rationalen Zahlen $(\mathbb{Q},+,\cdot)$ und die reellen Zahlen $(\mathbb{R},+,\cdot)$ sind Integritätsbereiche.

  Dagegen ist $(\mathbb{B}^2,+,\cdot)$ mit $\mathbb{B}\in\{\mathbb{Z},\mathbb{Q},\mathbb{R}\}$ und den komponentenweisen Verknüpfungen *kein* Integritätsbereich, denn dort gilt $(1,0)\cdot(0,1)=(0,0)=0_{\mathbb{B}^2}$. Aus dem gleichen Grund sind auch $(\mathbb{B}^n,+,\cdot)$ mit $n\geq 2$ und $(\mathrm{Abb}(\mathcal{X},\mathbb{B}),+,\cdot)$ mit $|\mathcal{X}|\geq 2$ *keine* Integritätsbereiche.
- Der Restklassenring $(\mathbb{Z}/n\mathbb{Z},+,\cdot)$ mit $n\in\mathbb{N}$ ist genau dann ein Integritätsbereich, wenn n eine Primzahl ist. Die folgt aus den Resultaten des Abschnitts III.1 zur multiplikativen Gruppe $(\mathbb{Z}/p\mathbb{Z}^\times,\cdot)$. Ist $n\in\mathbb{N}\setminus\{1\}$ *keine* Primzahl, also $n=ab$ mit $a,b\in\{2,3,\ldots,n-1\}$, so sieht man dies auch sofort an $[a][b]=[n]=[0]$, womit $[a]$ und $[b]$ Nullteiler in $\mathbb{Z}/n\mathbb{Z}$ sind.

Bemerkungen III.2.8.

(a) Dass ein kommutativer Ring $(R,+,\cdot)$ nullteilerfrei ist, bedeutet, dass die Implikation

$$xy=0 \implies x=0\vee y=0$$

beziehungsweise, äquivalent dazu

$$x\neq 0\wedge y\neq 0 \implies xy\neq 0$$

für alle $x,y\in R$ gilt.

(b) Als vielleicht wichtigste Konsequenz gilt in nullteilerfreien Ringen und Integritätsbereichen $(R,+,\cdot)$ die *Kürzungsregel*

$$xz=yz \implies x=y \text{ für alle } x,y\in R, z\in R\setminus\{0\}$$

(auch dann, wenn z nicht invertierbar ist). Um die Kürzungsregel einzusehen, schreibt man $xz=yz$ äquivalent als $(x-y)z=0$ und benutzt dann die vorige Bemerkung (a).

(c) Da man für einen Integritätsbereich $(R,+,\cdot)$ (wie für jeden Ring) immer $\mathbb{Z}$ oder $\mathbb{Z}/n\mathbb{Z}$ als Teilmenge von R auffassen kann und $(\mathbb{Z}/n\mathbb{Z},+,\cdot)$ aber nur für Primzahlen n Integritätsbereich ist, stellt sich heraus, dass die Charakteristik eines Integritätsbereichs stets null oder eine Primzahl ist.

Definition III.2.9.

(a) Ein *Schiefkörper* oder *Divisionsring* ist ein Ring $(K,+,\cdot)$ mit $K\neq\{0\}$, in dem jedes Element in $K\setminus\{0\}$ multiplikativ invertierbar ist, also ein Ring $(K,+,\cdot)$, für den $(K\setminus\{0\},\cdot)$ eine Gruppe ist.

(b) Ein *Körper* ist ein kommutativer Schiefkörper, also ein Ring $(K,+,\cdot)$, für den $(K\setminus\{0\},\cdot)$ eine abelsche Gruppe ist.

Beispiele III.2.10.

- Die rationalen Zahlen $(\mathbb{Q},+,\cdot)$ und die reellen Zahlen $(\mathbb{R},+,\cdot)$ sind Körper.
- Aus dem gleichen Grund wie bei Integritätsbereichen sind aber $(\mathbb{Q}^n,+,\cdot)$, $(\mathbb{R}^n,+,\cdot)$ mit $n\geq 2$ und $(\mathrm{Abb}(\mathcal{X},\mathbb{Q}),+,\cdot)$, $(\mathrm{Abb}(\mathcal{X},\mathbb{R}),+,\cdot)$ mit $|\mathcal{X}|\geq 2$ *keine* Körper.
- Der kommutative Restklassenring $(\mathbb{Z}/n\mathbb{Z},+,\cdot)$ mit $n\in\mathbb{N}$ ist genau dann ein Körper, wenn n eine Primzahl ist. Den Beweis hierfür haben wir schon in Abschnitt III.1 gesehen, als dort gezeigt wurde, dass $(\mathbb{Z}/n\mathbb{Z}\setminus\{0\},\cdot)$ genau für Primzahlen $n\in\mathbb{N}$ eine abelsche Gruppe ist. Man erhält also für jede Primzahl p einen endlichen Körper $(\mathbb{Z}/p\mathbb{Z},+,\cdot)$ mit genau p Elementen und Charakteristik p und schreibt in Anlehnung an den englischsprachigen Fachbegriff „field“ für Körper auch $\mathbb{F}_p$ für $\mathbb{Z}/p\mathbb{Z}$.
- Der vielleicht bekannteste Schiefkörper, der kein Körper ist, ist der $\mathbb{R}$ (und $\mathbb{C}$) erweiternde Zahlbereich der sogenannten Quaternionen, auf den wir an dieser Stelle aber nicht näher eingehen. Eine gute Einführung findet sich in **[16]**.

Bemerkungen III.2.11.

(a) Ein Körper $(K,+,\cdot)$ ist die algebraische Struktur mit den besten Eigenschaften der beiden Verknüpfungen. Wenn wir rekapitulieren, bedeutet dies insgesamt, dass
- $(K,+)$ eine abelsche Gruppe ist,
- $(K\setminus\{0\},\cdot)$ eine abelsche Gruppe ist, und dass
- die Distributivgesetze für alle Element von K gelten.

(b) Man kann daher in einem Körper $(K,+,\cdot)$ sehr weitgehend wie in den Bereichen $\mathbb{Q}$ und $\mathbb{R}$ der rationalen und reellen Zahlen rechnen. Speziell lässt sich über das zu Ringen Gesagte hinaus die Division $\frac{x}{y}:=x/y:=xy^{-1}$ für alle $x,y\in K$ mit $y\neq 0$ erklären.

(c) Jeder Körper ist insbesondere ein Integritätsbereich (denn für $y\in K\setminus\{0\}$ mit $xy=0$ für $x\in K$ folgt $x=xyy^{-1}=0y^{-1}=0$, so dass y kein Nullteiler sein kann).

(d) Die Charakteristik eines Körpers oder auch eines Schiefkörpes ist stets null oder eine Primzahl.

Ein weiteres zentrales Beispiel für eine Ringstruktur ergibt das Rechnen mit Polynomen:

Bemerkung III.2.12. Wir möchten *Polynome* (in einer Unbestimmten X) wie beispielsweise

$$\begin{aligned} p &:= 2X^2-3X+4 &&= 0X^3 &&+2X^2 &&+(-3)X &&+4\,,\\ q &:= 3X^3+6X^2+X &&= 3X^3 &&+6X^2 &&+1X &&+0 \end{aligned}$$

als symbolische Ausdrücke addieren wie in

$$p+q=(0{+}3)X^3+(2{+}6)X^2+(-3{+}1)X+(4{+}0)=3X^3+8X^2-2X+4$$

und multiplizieren wie in

$$\begin{aligned} pq &= (0{\cdot}3)X^6+(2{\cdot}3{+}0{\cdot}6)X^5+(-3{\cdot}3{+}2{\cdot}6{+}0{\cdot}1)X^4+(4{\cdot}3{-}3{\cdot}6{+}2{\cdot}1{+}0{\cdot}0)X^3\\ &\qquad +(4{\cdot}6{-}3{\cdot}1{+}2{\cdot}0)X^2+(4{\cdot}1{-}3{\cdot}0)X+(4{\cdot}0)\\ &= 6X^5+3X^4-4X^3+21X^2+4X. \end{aligned}$$

Um diese Rechenoperationen allgemein und formal einführen zu können, identifizieren wir p und q mit ihren jeweiligen *Koeffizientenfolgen* $a\colon\mathbb{N}_0\to\mathbb{Z}$, $i\mapsto a_i$ und $b\colon\mathbb{N}_0\to\mathbb{Z}$, $i\mapsto b_i$, die wir uns als unendliche Tupel $(a_0,a_1,a_2,a_3,a_4,a_5,a_6,\ldots)=(4,-3,2,0,0,0,0,\ldots)\in\mathbb{Z}^{(\mathbb{N}_0)}$ und $(b_0,b_1,b_2,b_3,b_4,b_5,b_6,\ldots)=(0,1,6,3,0,0,0,\ldots)\in\mathbb{Z}^{(\mathbb{N}_0)}$ mit nur endlich vielen Nicht-Null-Einträgen vorstellen. (Zur Notation $\mathbb{Z}^{(\mathbb{N}_0)}$ siehe dabei die folgende Definition.)

Formal lassen sich diese Überlegungen zum symbolischen Rechnen mit Polynomen in folgende algebraische Konstruktion umsetzen:

Definition III.2.13. Sei $(R,+,\cdot)$ ein Ring.

(a) Wir setzen

$$R[X] := \{a \in \mathrm{Abb}(\mathbb{N}_0, R) \mid a_i \neq 0 \text{ nur für endliche viele } i \in \mathbb{N}_0\},$$

wobei wir $a_i \in R$ für den Funktionswert von $i \in \mathbb{N}_0$ unter a schreiben und $a \in R[X]$ auch durch Aufzählung der Funktionswerte in der Form $a = (a_0, a_1, a_2, a_3, \ldots)$ angeben.

(b) Wir nennen $a = (a_0, a_1, a_2, a_3, \ldots) \in R[X]$ ein *Polynom* über dem Grundring $(R, +, \cdot)$ in der Unbestimmten X, den Eintrag $a_i \in R$ den *Koeffizient* i-ter Ordnung von a und $(a_0, a_1, a_2, a_3, \ldots) \in R^{(\mathbb{N}_0)}$ die *Koeffizientenfolge* von a.

(c) Wir erklären die *Summe* $a+b \in R[X]$ und das *Produkt* $ab \in R[X]$ *von Polynomen* $a, b \in R[X]$ durch

$$(a+b)_k := a_k + b_k,$$

$$(ab)_k := \sum_{i=0}^{k} a_i b_{k-i} \text{ für } k \in \mathbb{N}_0.$$

Dies ist wohldefiniert, da $(a+b)_k \neq 0$ und $(ab)_k \neq 0$ nur für endliche viele $k \in \mathbb{N}_0$ gelten kann.

(d) Wir nennen $(R[X], +, \cdot)$ mit den gerade definierten Verknüpfungen den *Polynomring* über dem Grundring $(R, +, \cdot)$ in der *Unbestimmten* X.

Proposition III.2.14. Für jeden kommutativen Ring $(R, +, \cdot)$ ist der Polynomring $(R[X], +, \cdot)$ über $(R, +, \cdot)$ ein kommutativer Ring mit $0_{R[X]} = (0_R, 0_R, 0_R, 0_R, \ldots)$, $1_{R[X]} = (1_R, 0_R, 0_R, 0_R, \ldots)$ und additiv Inversen $-(a_0, a_1, a_2, a_3, \ldots) = (-a_0, -a_1, -a_2, -a_3, \ldots)$ zu $(a_0, a_1, a_2, \ldots) \in R[X]$.

Beweis. Dass $(R[X], +)$ (mit $0_{R[X]}$ und Inversen wie angegeben) eine abelsche Gruppe ist, folgt problemlos daraus, dass $(\mathrm{Abb}(\mathbb{N}_0, R), +)$ diese Eigenschaft hat, die Null in $R[X]$ liegt und Inverse zu Elementen von $R[X]$ in $R[X]$ bleiben. Es gilt daher nur, die multiplikativen Eigenschaften und Distributivgesetze des Polynomrings zu verifizieren: Die benötigte Eigenschaft von $1_{R[X]}$ entnehmen wir aus der Definition der Multiplikation (wo bei Multiplikation mit $1_{R[X]}$ von links oder rechts einzig der Summand für $i = 0$ bzw. $i = k$ ungleich Null ist und bleibt). Für multiplikative Assoziativität machen wir die Rechnung (für $a, b, c \in R[X]$, $k \in \mathbb{N}_0$)

$$\begin{aligned}((ab)c)_k &= \sum_{j=0}^{k} \sum_{i=0}^{j} a_i b_{j-i} c_{k-j} = \sum_{0 \leq i \leq j \leq k} a_i b_{j-i} c_{k-j} = \sum_{i=0}^{j} \sum_{j=i}^{k} a_i b_{j-i} c_{k-j} \\ &= \sum_{i=0}^{j} \sum_{j=0}^{k-i} a_i b_j c_{k-i-j} = (a(bc))_k\end{aligned}$$

mit Umsortierung der Summationsindizes und Indexverschiebung, wobei $\sum_{0 \leq i \leq j \leq k}$ für $\sum_{(i,j) \in I_k}$ mit der Indexmenge

$$I_k := \{(\bar{i}, \bar{j}) \in \mathbb{N}_0^2 \mid \bar{i} \leq \bar{j} \leq k\}$$

steht. Die Distributivgesetze des Polynomrings ergeben sich auf naheliegende Weise durch Ausmultiplizieren von Summen in $(R, +, \cdot)$. Ist schließlich $(R, +, \cdot)$ kommutativ, so bekommen wir multiplikative Kommutativität im Polynomring durch die Rechnung (für $a, b \in R[X]$, $k \in \mathbb{N}_0$)

$$(ab)_k = \sum_{i=0}^{k} a_i b_{k-i} = \sum_{i=0}^{k} b_{k-i} a_i = \sum_{i=0}^{k} b_i a_{k-i} = (ba)_k$$

mit Indexlaufumkehr. Damit sind alle benötigten Eigenschaften gezeigt. □

Bemerkungen III.2.15. Es sei $(R, +, \cdot)$ ein kommutativer Ring.

(a) Durch Identifikation der Elemente $r \in R$ des Grundrings mit $(r, 0, 0, 0, \ldots) \in R[X]$ verstehen wir

$$R \subset R[X].$$

Wir bezeichnen die Elemente von R in diesem Zusammenhang als *Konstanten* oder *konstante Polynome* und bekommen $r(a_0, a_1, a_2, a_3, \ldots) = (ra_0, ra_1, ra_2, ra_3, \ldots)$ sowie $(a_0, a_1, a_2, a_3, \ldots)r = (a_0 r, a_1 r, a_2 r, a_3 r, \ldots)$ für $r \in R$ und $(a_0, a_1, a_2, a_3, \ldots) \in R[X]$.

(b) Wir verstehen die Unbestimmte X selbst als Polynom

$$X := (0, 1, 0, 0, 0, \ldots) \in R[X]$$

mit $pX = Xp$ für alle $p \in R[X]$ und erhalten induktiv, dass das Monom rX^k der Ordnung $k \in \mathbb{N}_0$ mit Koeffizient $r \in R$ die Koeffizientenfolge $(0, 0, \ldots, 0, 0, r, 0, 0, 0, \ldots) \in R[X]$ mit genau k Nullen vor einem r als dem einzigen nicht-trivialen Koeffizienten besitzt.

(c) Die Potenzen von X^i sind wie in einem allgemeinen Ring zu verstehen. Insbesondere ist $X^0 = 1 \in R \subset R[X]$.

(d) Damit können wir jedes Polynom $p = (a_0, a_1, a_2, a_3, \ldots) \in R[X] \setminus \{0\}$ in Standard-Form

$$p = a_\ell X^\ell + a_{\ell-1} X^{\ell-1} + \ldots + a_2 X^2 + a_1 X + a_0 = \sum_{i=0}^{\ell} a_i X^i$$

schreiben, wobei ℓ die größte Zahl in $\mathbb{N}_0$ mit $a_\ell \neq 0$ sei. Diese existiert, weil mindestens ein Koeffizient, aber insgesamt nur endliche viele Koeffizienten ungleich Null sind. Wir nennen hierbei a_ℓ den *Höchstkoeffizienten* und $\mathrm{Grad}(p) := \ell \in \mathbb{N}_0$ den *Grad* des Polynoms p. Für das Nullpolynom $0_{R[X]}$ treffen wir die Konvention, dass sein Grad $-\infty$ sei. Ein Polynom mit Höchstkoeffizient 1 heißt auch normiertes Polynom.

(e) Direkt aus der Definition als Koeffizientenfolgen ergibt sich für Polynome die Möglichkeit des *Koeffizientenvergleichs*: Aus der Gleichheit $\sum_{i=0}^{\ell} a_i X^i = \sum_{i=0}^{m} b_i X^i$ von Polynomen in $R[X]$ mit Höchstkoeffizienten $a_\ell \neq 0 \neq b_m$ folgen die Übereinstimmungen $\ell = m$ der Grade und $a_i = b_i$ der Koeffizienten für alle $i \in \{1, 2, \ldots, \ell\}$.

(f) An dieser Stelle erhalten wir für

$$p = \sum_{i=0}^{\ell} a_i X^i \in R[X] \text{ und } q = \sum_{i=0}^{m} b_i X^i \in R[X]$$

ganz allgemein die Rechenregeln für Summe und Produkt von Polynomen

$$p + q = \sum_{k=0}^{\max\{\ell,m\}} (a_k + b_k) X^k \text{ und } pq = \sum_{k=0}^{\ell+m} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) X^k = \sum_{k=0}^{\ell+m} \left(\sum_{i=\max\{0,k-m\}}^{\min\{k,\ell\}} a_i b_{k-i} \right) X^k,$$

wobei $\max\{x, y\}$ bzw. $\min\{x, y\}$ die größere bzw. kleinere von zwei Zahlen $x, y \in \mathbb{R}$ bezeichnet und wir natürlich $a_{\ell+1} = a_{\ell+2} = a_{\ell+3} = \ldots = 0$ sowie $b_{m+1} = b_{m+2} = b_{m+3} = \ldots = 0$ verstehen.

Wir halten noch fest:

Bemerkungen III.2.16.

- Für einen Integritätsbereich $(R, +, \cdot)$ ist auch der Polynomring $(R[X], +, \cdot)$ ein Integritätsbereich. (Begründung: Für Polynome $p = \sum_{i=0}^{\ell} a_i X^i \in R[X] \setminus \{0\}$ und $q = \sum_{i=0}^{m} b_i X^i \in R[X] \setminus \{0\}$ mit Höchstkoeffizienten $a_\ell \neq 0 \neq b_m$ ergibt sich als Höchstkoeffizient $(\ell+m)$-ter Ordnung von $p+q$ gemäß Obigem $\sum_{i=\max\{0,\ell\}}^{\min\{\ell+m,\ell\}} a_i b_{\ell+m-i} = a_\ell b_m \neq 0$. Damit ist $pq \neq 0$ in $R[X]$.)
- Selbst für einen Körper K ist der Polynomring $K[X]$ *nie* ein Körper, denn das Polynom X ist (wie auch jedes andere Polynom vom Grad ≥ 1) nicht invertierbar.

In der Schulmathematik werden Polynome eher als Funktionen oder Funktionsterme betrachtet. Auch in unserer Betrachtungsweise sind Polynome mit einer zugehörigen Polynomfunktion verbunden.

Definition III.2.17. Sei $(R, +, \cdot)$ ein kommutativer Ring. Die zu einem Polynom $p = \sum_{i=0}^{\ell} a_i X^i \in R[X]$ gehörige *Polynomfunktion* $\widehat{p}\colon R \to R$ ist durch

$$\widehat{p}(r) := \sum_{i=0}^{\ell} a_i r^i \in R \text{ für alle } r \in R$$

gegeben. Wir nennen $r \in R$ eine *Nullstelle* von $p \in R[X]$ und auch von $\widehat{p} \in \mathrm{Abb}(R)$, wenn $\widehat{p}(r) = 0_R$ gilt.

Es gibt trotzdem gute Gründe, die Konzepte der Polynome und der Polynomfunktionen sauber auseinanderzuhalten:

Beispiel III.2.18. Wir betrachten die Primzahl 2 und $R = \mathbb{Z}/2\mathbb{Z}$. Das Polynom $p = X^2 + X$ ist natürlich nicht das Nullpolynom in $\mathbb{Z}/2\mathbb{Z}[X]$, aber die zugehörige Polynomfunktion $\widehat{p}$ ist die Nullabbildung, weil beide Funktionswerte

$$[0]^2 + [0] = [0] \text{ und } [1]^2 + [1] = [1] + [1] = [2] = [0]$$

trivial sind.

Finden Sie weitere Beispiele für andere Primzahlen? Können Sie für jede beliebige Primzahl ein Beispiel angeben für ein nicht-triviales Polynom, welches als Polynomfunktion konstant null ist?

Bemerkungen III.2.19. Es sei $(R, +, \cdot)$ ein kommutativer Ring.

- Das *Polynom* $p = \sum_{i=0}^{\ell} a_i X^i \in R[X]$ ist formal als Koeffizientenfolge definiert und als Objekt symbolischen Rechnens mit einer Unbestimmten X zu verstehen. Die *Polynomfunktion* $\widehat{p} \in \mathrm{Abb}(R)$ ist *als Abbildung* $R \to R$ definiert (die übrigens ganz anders abbildet als die Koeffizientenfolge $\mathbb{N}_0 \to R$ von p und mit dieser nicht unmittelbar zusammenhängt).

  Noch etwas anderes ist der Funktions*wert* $\widehat{p}(r) = \sum_{i=0}^{\ell} a_i r^i$ der Polynomfunktion $\widehat{p}$, der für jedes einzelne $r \in R$ selbst Element von R ist. Zwar besteht die Funktion $\widehat{p}$ im Wesentlichen in der Zuordnungsregel $r \mapsto \sum_{i=0}^{\ell} a_i r^i$ oder mit anderen Worten $\widehat{p}(r) = \sum_{i=0}^{\ell} a_i r^i$ für alle $r \in R$, die Gleichsetzung $\widehat{p} = \sum_{i=0}^{\ell} a_i r^i$ ohne „(r)“ links ist aber formal *nicht* korrekt.
- Die Addition und Multiplikation im Polynomring $R[X]$ entsprechen der punktweisen Addition und Multiplikation von Polynomfunktionen in dem Sinn, dass

  $$\widehat{p+q} = \widehat{p} + \widehat{q} \text{ und } \widehat{pq} = \widehat{p}\widehat{q} \text{ in } \mathrm{Abb}(R)$$

  für alle $p, q \in R[X]$ gelten. Diese Übereinstimmung bedeutet letztlich, dass man auch mit (formal eingeführten) Polynomen mit denselben Rechenregeln wie im Ring $(R, +, \cdot)$ rechnen kann – und speziell bei Polynomen über Zahlbereichen einfach mit den üblichen Rechenregeln.

  (Begründung: Die Gleichheit $\widehat{p+q} = \widehat{p}+\widehat{q}$ ist klar, da auch $p+q$ komponentenweise definiert wurde. Die Gleichheit $\widehat{pq} = \widehat{p}\widehat{q}$ zeigen wir, indem wir $p = \sum_{i=0}^{\ell} a_i X^i$, $q = \sum_{j=0}^{m} b_j X^j$ schreiben und mit der Kommutativität von $(R, +, \cdot)$ nachrechnen, dass

  $$\widehat{p}(r)\,\widehat{q}(r) = \Big(\sum_{i=0}^{\ell} a_i r^i\Big)\Big(\sum_{j=0}^{m} b_j r^j\Big) = \sum_{i=0}^{\ell}\sum_{j=0}^{m} a_i b_j r^{i+j}$$

  $$= \sum_{i=0}^{\ell}\sum_{k=i}^{i+m} a_i b_{k-i} r^k = \sum_{0\le i\le k\le i+m\le \ell+m} a_i b_{k-i} r^k = \sum_{k=0}^{\ell+m}\sum_{i=\max\{0,k-m\}}^{\min\{k,\ell\}} a_i b_{k-i} r^k = \widehat{pq}(r)$$

  für alle $r \in R$ gilt.)
- Als Konsequenz der vorigen Bemerkung bildet die Menge der Polynomfunktionen $R \to R$ mit der punktweisen Addition und Multiplikation ebenfalls einen kommutativen Ring.

Wie bei ganzen Zahlen besteht auch bei Polynomen die Möglichkeit zur Division mit Rest und bringt einige nützliche Konsequenzen:

Satz III.2.20. *Es sei $(K, +, \cdot)$ ein Körper.*

(a) Polynomdivision: *Für Polynome $p, q \in K[X]$ mit $q \neq 0$ gibt es eindeutig bestimmte Polynome $r, s \in K[X]$ mit*[2] $\mathrm{Grad}(r) < \mathrm{Grad}(q)$*, so dass $p = s \cdot q + r$ gilt.*

(b) Abspalten von Linearfaktoren: *Ist $x_0 \in K$ eine Nullstelle von $p \in K[X]$, so kann p als $p = s\cdot(X - x_0)$ mit $s \in K[X]$ geschrieben werden.*

(c) *Ein Polynom p über K vom Grad $n \in \mathbb{N}_0$ hat höchstens n verschiedene Nullstellen in K.*

(d) *Hat K unendlich viele Elemente, so ist ein Polynom über K durch die zugehörige Polynomfunktion eindeutig bestimmt.*

Beweis. Beweis von Teil (a) des Satzes: Wir zeigen zuerst Existenz von r und s: Im Fall $\mathrm{Grad}(q) > \mathrm{Grad}(p)$ können wir $r := p$ und $s := 0$ wählen und erhalten trivial $p = s \cdot q + r$ mit $\mathrm{Grad}(r) = \mathrm{Grad}(p) < \mathrm{Grad}(q)$. Für $\mathrm{Grad}(q) \le \mathrm{Grad}(p)$ zeigen wir die Existenz durch Induktion nach $\mathrm{Grad}(p) \in \mathbb{N}_0$. Beim Induktionsanfang für $\mathrm{Grad}(p) = 0$ ist auch $\mathrm{Grad}(q) = 0$ und damit $q \in K \setminus \{0\}$, so dass wir $p = s \cdot q + r$ für

[2] Wir erlauben $r = 0$ mit $\mathrm{Grad}(r) = -\infty < \mathrm{Grad}(q) \in \mathbb{N}_0$.

$r := 0$ und $s := pq^{-1}$ erhalten. Für den Induktionsschritt habe $p \in K[X]$ Grad $\ell \in \mathbb{N}$ und Höchstkoeffizient a_ℓ sowie $q \in K[X]$ Grad $m \in \mathbb{N}_0$ und Höchstkoeffizient b_m, insbesondere $b_m \neq 0$, und es sei $m \leq \ell$. Dann verschwindet bei $\widetilde{p} := p - a_\ell b_m^{-1} X^{\ell-m} q \in K[X]$ der Koeffizient ℓ-ter Ordnung, es gilt also $\mathrm{Grad}(\widetilde{p}) < \ell$. Per Induktionsannahme gibt es daher $r, \widetilde{s} \in K[X]$ mit $\mathrm{Grad}(r) < \mathrm{Grad}(q)$, so dass $\widetilde{p} = \widetilde{s} \cdot q + r$ gilt. Durch Umformen erhalten wir daraus $p = (\widetilde{s} + a_\ell b_m^{-1} X^{\ell-m}) \cdot q + r$, also die Induktionsbehauptung für das bereits eingeführte r und $s := \widetilde{s} + a_\ell b_m^{-1} X^{\ell-m} \in K[X]$.

Um die Eindeutigkeit von r und s nachzuweisen, ist für $r, \widetilde{r}, s, \widetilde{s} \in K[X]$ mit $\mathrm{Grad}(r) < \mathrm{Grad}(q)$, $\mathrm{Grad}(\widetilde{r}) < \mathrm{Grad}(q)$ und $\widetilde{s} \cdot q + \widetilde{r} = s \cdot q + r$ zu zeigen, dass $\widetilde{r} = r$, $\widetilde{s} = s$ sein muss. Dazu schreiben wir die Gleichung als $(\widetilde{s} - s) \cdot q = r - \widetilde{r}$ und bemerken im Fall $\widetilde{s} \neq s$, dass $\mathrm{Grad}((\widetilde{s} - s) \cdot q) \geq \mathrm{Grad}(q)$ gilt, weil der Grad des Produkts die Summe der Grade ist. Dies benutzt die Nullteilerfreiheit. Andererseits gilt aber $\mathrm{Grad}(r - \widetilde{r}) < \mathrm{Grad}(q)$, und wir erreichen einen Widerspruch. Also muss $\widetilde{s} = s$ sein, und dann folgt sofort auch $\widetilde{r} = r$.

Beweis von Teil (b) des Satzes: Aus Teil (a) mit $q := X - x_0 \in K[X]$ lesen wir $p = s \cdot (X - x_0) + r$ für $r, s \in K[X]$ mit $\mathrm{Grad}(r) < \mathrm{Grad}(q) = 1$, also $r \in K$ ab. Durch Einsetzen von x_0 erhalten wir $0 = \widehat{p}\,(x_0) = \widehat{s}\,(x_0) \cdot 0 + r = r$, also gilt wie behauptet $p = s \cdot (X - x_0)$.

Beweis von Teil (c) des Satzes:
Wir argumentieren indirekt: Sei $p \in K[X]$ ein Polynom mit $\mathrm{Grad}(p) = n \in \mathbb{N}_0$ und $(n+1)$ verschiedenen Nullstellen $x_1, x_2, \ldots, x_n, x_{n+1} \in K$. Dann erreichen wir durch $(n+1)$-malige Anwendung von Teil (b) des Satzes die Form $p = s \prod_{i=1}^{n+1} (X - x_i)$ mit $s \in K[X]$. Im Fall $s \neq 0$ folgt $\mathrm{Grad}(p) \geq n+1$, im Fall $s = 0$ folgt $\mathrm{Grad}(p) = -\infty$. Wir erhalten also in jedem Fall einen Widerspruch zu $\mathrm{Grad}(p) = n \in \mathbb{N}_0$, und die Behauptung ist bewiesen.

Beweis von Teil (d) des Satzes: Wären $p, q \in K[X]$ *verschiedene* Polynome mit $p(x) = q(x)$ für alle $x \in K$, so hätte $p - q \in K[X] \setminus \{0\}$ mit $\mathrm{Grad}(p - q) \in \mathbb{N}_0$ unendlich viele Nullstellen in K und dies stünde im Widerspruch zu Teil (c) des Satzes. □

Bemerkung III.2.21. Die Teile (c) und (d) Satzes gelten nicht nur über Körpern, sondern allgemeiner für normierte Polynome über Integritätsbereichen. Der Beweis von Teil (c) muss dort modifiziert werden.

III.2.1. Polynomdivision. Mit Teil (a) des Satzes geht das *Rechenverfahren der Polynomdivision* einher, mit dem man für gegebene Polynome $p, q \in K[X]$, $q \neq 0$, über einem Körper K den Multiplikator s und den Rest r mit $p = s \cdot q + r$ und $\mathrm{Grad}(r) < \mathrm{Grad}(q)$ bestimmt. Man baut dabei exakt auf der Idee des vorgestellten Induktionsbeweises auf und bestimmt durch Division des Leitmonoms $a_\ell X^\ell$ von p durch das Leitmonom $b_m X^m$ von q zunächst das Leitmonom $a_\ell b_m^{-1} X^{\ell-m}$ von s. Nun betrachtet man den Korrekturterm $\widetilde{p} := p - a_\ell b_m^{-1} X^{\ell-m} \cdot q$ und dividiert im nächsten Schritt das Leitmonom von $\widetilde{p}$ durch $b_m X^m$, um den nächsten Term von s zu erhalten. Danach folgt die nächste Korrektur, und so weiter.

Als konkretes Beispiel führen wir dieses Verfahren hier für $p := X^4 - 4X^3 + 4X^2 - 2$ und $q := 2X^2 - 4X - 6$ in $\mathbb{Q}[X]$ wie folgt durch:

$$
\begin{array}{l}
(\mathrm{X}^4 - 4\mathrm{X}^3 \qquad + 4\mathrm{X} - 2) \,:\, (2\mathrm{X}^2 - 4\mathrm{X} - 6) = \tfrac{1}{2}\mathrm{X}^2 - \mathrm{X} - \tfrac{1}{2} + \dfrac{-4\mathrm{X} - 5}{2\mathrm{X}^2 - 4\mathrm{X} - 6}\,. \\
\underline{-(\mathrm{X}^4 - 2\mathrm{X}^3 - 3\mathrm{X}^2)} \\
\qquad -2\mathrm{X}^3 + 3\mathrm{X}^2 \\
\qquad \underline{-(-2\mathrm{X}^3 + 4\mathrm{X}^2 + 6\mathrm{X})} \\
\qquad\qquad -\ \mathrm{X}^2 - 2\mathrm{X} \\
\qquad\qquad \underline{-(-\ \mathrm{X}^2 + 2\mathrm{X} + 3)} \\
\qquad\qquad\qquad -4\mathrm{X} - 5
\end{array}
$$

Dabei wird im jedem Schritt ein farbiges Leitmonom der linken Seite durch das Leitmonom $2X^2$ von q (das in jedem Schritt gleich bleibt) dividiert, um das farblich entsprechende Monom auf der rechten Seite zu erhalten. Dieses Monom wird dann mit ganz $q = 2X^2 - 4X - 6$ multipliziert, um das immer noch farblich entsprechende Korrekturpolynom links zu erhalten. Nach Subtraktion der Korrektur beginnt der nächste Schritt. Da sich der Grad des Polynoms links in jedem Schritt verringert, ist dieser Grad nach endlich vielen Schritten $< \mathrm{Grad}(q)$, womit das Verfahren endet und das verbleibende, hier violett gefärbte Polynom den

Rest r bildet. Insgesamt haben wir im Beispielfall $s = \frac{1}{2}X^2-X-\frac{1}{2}$ und $r = -4X-5$ gefunden und mit anderen Worten

$$X^4-4X^3+4X^2-2 = \big(\frac{1}{2}X^2-X-\frac{1}{2}\big)\cdot(2X^2-4X-6)+(-4X-5)$$

eingesehen.

Zum Abschluss dieses Abschnitts halten wir fest, dass die Konstruktion des Polynomrings über einem Grundring iteriert werden kann (was aber tatsächlich nur deshalb funktioniert, weil wir die Bildung allgemein über Ringen und nicht nur über Körpern vorgenommen haben):

Bemerkung III.2.22. Ein *Polynom* über einem Ring $(R,+,\cdot)$ *in zwei Unbestimmten* X und Y hat die Form

$$\sum_{\substack{i\in\{0,1,2,\ldots,\ell\}\\ j\in\{0,1,2,\ldots,m\}}} a_{ij}X^iY^j := \sum_{j=0}^{m}\left(\sum_{i=0}^{\ell} a_{ij}X^i\right)Y^j \in \big(R[X]\big)[Y]$$

mit Koeffizienten $a_{ij}\in R$. Man schreibt daher

$$R[X,Y] := (R[X])[Y]$$

und nennt $R[X,Y]$ den Polynomring über R in zwei Unbestimmten X und Y (für die übrigens auch bei nichtkommutativem Grundring stets $XY=YX$ und allgemeiner $pX=Xp$, $pY=Yp$ für alle $p\in R[X,Y]$ gelten). Iteration dieser Vorgehensweise ergibt den Polynomring $R[X_1,X_2,\ldots,X_n] := \big(\ldots\big(\big(R[X_1]\big)[X_2]\big)\ldots\big)[X_n]$ über R in $n\in\mathbb{N}$ Unbestimmten X_1, X_2, …, X_n. Genauer gehen wir auf das Rechnen mit Polynomen mehrerer Variablen an dieser Stelle aber nicht ein.

III.3. Homomorphismen, Unter- und Faktorstrukturen

Definition III.3.1.

(a) Ein *(Gruppen-))Homomorphismus* von einer Gruppe $(G,*)$ in eine Gruppe $(H,\circledast)$ ist eine Abbildung $f\colon G\to H$ mit $f(g_1{*}g_2)=f(g_1)\circledast f(g_2)$ für alle $g_1,g_2\in G$.

(b) Ein *(Ring-)Homomorphismus* von einem Ring $(R,+,\cdot)$ in einen Ring $(S,\oplus,\odot)$ ist eine Abbildung $f\colon R\to S$ mit $f(x{+}y)=f(x)\oplus f(y)$ und $f(x{\cdot}y)=f(x)\odot f(y)$ für alle $x,y\in R$, sowie $f(1_R)=1_S$. Sind $(R,+,\cdot)$ und $(S,\oplus,\odot)$ sogar Körper, so spricht man von einem *Körperhomomorphismus*.

(c) Wir vereinbaren für Gruppen, Ringe, Körper gleichermaßen: Ein *Monomorphismus*, *Epimorphismus* bzw. *Isomorphismus* ist ein injektiver, surjektiver bzw. bijektiver Homomorphismus. Gibt es zwischen zwei Gruppen/Ringen/Körpern einen Isomorphismus, so heißen diese zueinander *isomorph*. Ein *Endomorphismus* ist ein Homomorphismus mit gleichem Definitionsbereich und Ziel bezüglich der gleichen Verknüpfungen darauf. Ein *Automorphismus* ist ein bijektiver Endomorphismus. Die Mengen aller Homo- und aller Isomorphismen $\mathcal{X}\to\mathcal{Y}$ notieren wir als $\mathrm{Hom}(\mathcal{X},\mathcal{Y})$ und $\mathrm{Iso}(\mathcal{X},\mathcal{Y})$, für Endo- und Automorphismen vereinbaren wir $\mathrm{End}(\mathcal{X}) := \mathrm{Hom}(\mathcal{X},\mathcal{X})$ und $\mathrm{Aut}(\mathcal{X}) := \mathrm{Iso}(\mathcal{X},\mathcal{X})$.

Bemerkungen III.3.2.

(a) Homomorphismen sind strukturerhaltende Abbildungen. Im Fall eines Isomorphismus kann man die Elemente in Definitionsbereich und Ziel 1-zu-1 identifizieren. Daher erhalten Isomorphismen alle algebraischen Eigenschaften, und zueinander isomorphe Gruppen/Ringe/Körper verhalten sich in algebraischer Hinsicht völlig gleich.

(b) Ein Gruppenhomomorphismus $f\colon G\to H$ erfüllt automatisch $f(e_G)=e_H$ für die neutralen Elemente $e_G\in G$ und $e_H\in H$ sowie $f(g)^{-1}=f(g^{-1})$ für alle $g\in G$.

(Nachweis: Mit $f(e_G)=f(e_G)\circledast f(e_G)\circledast f(e_G)^{-1}=f(e_G{*}e_G)\circledast f(e_G)^{-1}=f(e_G)\circledast f(e_G)^{-1}=e_H$ erhalten wie die Behauptung über die neutralen Elemente. Für die Regel zu den Inversen rechnen wir dann $f(g)\circledast f(g^{-1})=f(g{*}g^{-1})=f(e_G)=e_H$ und $f(g^{-1})\circledast f(g)=f(g^{-1}{*}g)=f(e_G)=e_H$.)

(c) Ein Ring- beziehungsweise Körperhomomorphismus $f\colon R\to S$ ist insbesondere ein Gruppenhomomorphismus von $(R,+)$ in $(S,\oplus)$ und erfüllt neben $f(1_R)=1_S$ automatisch $f(0_R)=0_S$, $-f(x)=f(-x)$ für $x\in R$ sowie $f(x)^{-1}=f(x^{-1})$ für invertierbare $x\in R$. Im Körperfall ist f auch Gruppenhomomorphismus von $(R\setminus\{0\},\cdot)$ in $(S\setminus\{0\},\odot)$ mit $f(x)^{-1}=f(x^{-1})$ für alle $x\in R\setminus\{0\}$. Dies folgt aus der vorigen Bemerkung beziehungsweise analog zu dieser.

(Die Forderung $f(1_R) = 1_S$ in der Definition kann man aber nicht weglassen. Dass diese nicht aus den anderen Bedingungen folgt, erkennt man am Beispiel der Nullabbildung $f\colon R \to S$, $x \mapsto 0_S$.)

Beispiele III.3.3.

- Die identische Abbildung ist für jede Gruppe/jeden Ring/jeden Körper ein Automorphismus.
- Für $\mathbb{B} \in \{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$ ist die Abbildung $\mathbb{B} \to \mathbb{B}$, $x \mapsto -3x$ ein Gruppenendomorphismus von $(\mathbb{B}, +)$, denn es gilt $-3(x+y) = -3x+(-3y)$ für $x, y \in \mathbb{B}$. Für $\mathbb{B} \in \{\mathbb{Q}, \mathbb{R}\}$ handelt es sich sogar um einen Gruppenautomorphismus
- Für $\mathbb{B} \in \{\mathbb{Q}, \mathbb{R}\}$ ist die Abbildung $\mathbb{B}\backslash\{0\} \to \mathbb{B}\backslash\{0\}$, $x \mapsto x^2$ ein Gruppenendomorphismus von $(\mathbb{B}\backslash\{0\}, \cdot)$, denn es gilt $(xy)^2 = x^2y^2$ für $x, y \in \mathbb{B}\backslash\{0\}$.
- Die Abbildung $\mathbb{Z} \to \mathbb{Q} \setminus \{0\}$, $x \mapsto 2^x$ ist ein Gruppenmonomorphismus von der additiven Gruppe $(\mathbb{Z}, +)$ in die multiplikative Gruppe $(\mathbb{Q} \setminus \{0\}, \cdot)$, denn es gilt $2^{x+y} = 2^x 2^y$ für $x, y \in \mathbb{Z}$.
- Für jedes $1 \neq n \in \mathbb{N}$ ist die Quotientenabbildung $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $x \mapsto [x]_{\mathbb{Z}/n\mathbb{Z}}$ des Restklassenrings $\mathbb{Z}/n\mathbb{Z}$ ein Ringepimorphismus, denn es gelten $[x+y]_{\mathbb{Z}/n\mathbb{Z}} = [x]_{\mathbb{Z}/n\mathbb{Z}}+[y]_{\mathbb{Z}/n\mathbb{Z}}$ und $[xy]_{\mathbb{Z}/n\mathbb{Z}} = [x]_{\mathbb{Z}/n\mathbb{Z}}[y]_{\mathbb{Z}/n\mathbb{Z}}$ für alle $x, y \in \mathbb{Z}$ sowie $[1]_{\mathbb{Z}/n\mathbb{Z}} = 1_{\mathbb{Z}/n\mathbb{Z}}$.
- Die Einbettung $\mathbb{Q} \to \mathbb{R}$, $x \mapsto x$ von $\mathbb{Q}$ in $\mathbb{R}$ ist ein Körpermonomorphismus.
- Für jeden Ring $(R, +, \cdot)$ ist die Vertauschung der Einträge $f\colon R^2 \to R^2$, $(x_1, x_2) \mapsto (x_2, x_1)$ ein Ringautomorphismus. Allgemeiner ist für $n \in \mathbb{N}$ und $\pi \in \Sigma_n$ die *Koordinatenpermutation* $f_\pi\colon R^n \to R^n$, $(x_1, x_2, \ldots, x_n) \mapsto (x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)})$ ein Ringautomorphismus.
- Für jeden kommutativen Ring $(R, +, \cdot)$ ist die Abbildung $R[X] \to \mathrm{Abb}(R, R)$, $p \mapsto \widehat{p}$ ein Ringhomomorphismus und ebenso für jedes feste $r \in R$ die Abbildung $R[X] \to R$, $p \mapsto \widehat{p}(r)$. Bei beiden Bildungen spricht man auch vom Einsetzungshomomorphismus.

  (Für einen Integritätsbereich R mit unendlich vielen Elementen ist die erste Variante sogar Ring*mono*morphismus. Das folgt aus dem letzten Satz in Abschnitt III.2 und der darauf folgenden Bemerkung.)

Bemerkungen III.3.4.

(a) Für jede Gruppe $(G, *)$ wird $\mathrm{Aut}(G)$ mit der Komposition von Abbildungen zu einer Gruppe, der *Automorphismengruppe* von $(G, *)$:

Für $\phi, \psi \in \mathrm{Aut}(G)$ rechnen wir nach, dass $\psi\circ\phi \in \mathrm{Aut}(G)$ wiederum ein Homomorphismus ist: Es gilt

$$\psi(\phi(g_1{*}g_2)) = \psi(\phi(g_1){*}\phi(g_2)) = \psi(\phi(g_1)){*}\psi(\phi(g_2))$$

für alle $g_1, g_2 \in G$.

Zudem ergibt sich $\phi^{-1} \in \mathrm{Aut}(G)$ für die Umkehrfunktion ϕ^{-1}

$$\phi^{-1}(g_1{*}g_2) = \phi^{-1}(\phi(\phi^{-1}(g_1)){*}\phi(\phi^{-1}(g_1))) = \phi^{-1}(\phi(\phi^{-1}(g_1){*}\phi^{-1}(g_2))) = \phi^{-1}(g_1){*}\phi^{-1}(g_2)$$

für alle $g_1, g_2 \in G$, so dass diese wiederum strukturerhaltend ist.

(b) Für eine Gruppe $(G, *)$ und eine abelsche Gruppe $(H, \circledast)$ werden $\mathrm{Hom}(G, H)$ und $\mathrm{End}(H)$ *mit punktweiser Verknüpfung* $\circledast$ auch abelsche Gruppen. Man spricht von *Homo- bzw. Endomorphismengruppen*.

(Nachweis: Für $\phi, \psi \in \mathrm{Hom}(G, H)$ zeigen wir $\phi\circledast\psi \in \mathrm{Hom}(G, H)$ durch die auf Kommutativität von $\circledast$ gegründete Rechnung

$$\begin{aligned}(\phi\circledast\psi)(g_1{*}g_2) &= \phi(g_1{*}g_2)\circledast\psi(g_1{*}g_2)\\ &= \phi(g_1)\circledast\phi(g_2)\circledast\psi(g_1)\circledast\psi(g_2)\\ &= \phi(g_1)\circledast\psi(g_1)\circledast\phi(g_2)\circledast\psi(g_2)\\ &= (\phi\circledast\psi)(g_1)\circledast(\phi\circledast\psi)(g_2)\end{aligned}$$

für $g_1, g_2 \in G$. Dass die konstante Abbildung $G \to H$ auf das neutrale Element in H das neutrale Element in $\mathrm{Hom}(G, H)$ ist, ist klar. Schließlich können wir Inverse zu $\phi \in \mathrm{Hom}(G, H)$ als punktweise Inverse ϕ^{-1} erhalten, denn die Rechnung

$$\begin{aligned}\phi^{-1}(g_1{*}g_2) &= \phi(g_1{*}g_2)^{-1}\\ &= (\phi(g_1)\circledast\phi(g_2))^{-1}\\ &= \phi(g_2)^{-1}\circledast\phi(g_1)^{-1}\\ &= \phi(g_1)^{-1}\circledast\phi(g_2)^{-1}\end{aligned}$$

unter erneuter Verwendung der Kommutativität von $\circledast$ ergibt $\phi^{-1} \in \mathrm{Hom}(G,H)$.)

(c) Für eine (meist additiv notierte) abelsche Gruppe $(G,+)$ wird $(\mathrm{End}(G),+,\circ)$ mit punktweiser Addition und Komposition sogar ein Ring, der *Endomorphismenring* von $(G,+)$. Anders als bei $(\mathrm{Abb}(G),+,\circ)$ (was gemäß Abschnitt III.2 *kein* Ring ist) ergibt sich für Endomorphismen $\phi,\psi,\chi \in \mathrm{End}(G)$ nämlich das „linke" Distributivgesetz $\phi\circ(\psi+\chi) = \phi\circ\psi+\phi\circ\chi$ durch die Rechnung

$$(\phi\circ(\psi+\chi))(g) = \phi(\psi(g)+\chi(g)) = \phi(\psi(g))+\phi(\chi(g)) = (\phi\circ\psi+\phi\circ\chi)(g)$$

für alle $g \in G$ mit der Homomorphismus-Eigenschaft von ϕ. Die anderen benötigten Eigenschaften folgen aus Bemerkung (b) oder sind leicht zu prüfen.

Definition III.3.5.

(a) Der *Kern eines Gruppenhomomorphismus* $f\colon G \to H$ zwischen Gruppen $(G,*)$ und $(H,\circledast)$ ist

$$\mathrm{Kern}(f) := f^{-1}(\{e_H\}) = \{g \in G \mid f(g) = e_H\} \subset G$$

mit dem neutralen Element e_H von $(H,\circledast)$.

(b) Der *Kern eines Ringhomomorphismus* $f\colon R \to S$ zwischen Ringen $(R,+,\cdot)$ und $(S,\oplus,\odot)$ ist

$$\mathrm{Kern}(f) := f^{-1}(\{0_S\}) = \{x \in R \mid f(x) = 0_S\} \subset R.$$

Bemerkungen III.3.6. Wegen $f(e_G) = e_H$ für das neutrale Element e_G von $(G,*)$ gilt in (a) stets $e_G \in \mathrm{Kern}(f)$ und wegen $f(0_R) = 0_S$ in (b) stets $0_R \in \mathrm{Kern}(f)$.

Am Kern lässt sich erkennen, ob ein Homomorphismus ein Monomorphismus ist (und so wird diese Eigenschaft in Zukunft dann auch oft gezeigt):

Satz III.3.7.

(a) *Für einen Homomorphismus $f\colon G \to H$ von Gruppen $(G,*)$ und $(H,\circledast)$ gilt:*

$$f \text{ ist injektiv} \iff \mathrm{Kern}\, f = \{e_G\}.$$

(b) *Für einen Homomorphismus $f\colon R \to S$ von Ringen $(R,+,\cdot)$ und $(S,\oplus,\odot)$ gilt:*

$$f \text{ ist injektiv} \iff \mathrm{Kern}\, f = \{0_R\}.$$

(c) *Ein Körperhomomorphismus $f\colon K \to L$ von Körpern $(K,+,\cdot)$ und $(L,\oplus,\odot)$ ist immer injektiv.*

Beweis. Teil (a) wird in den Übungen gezeigt. Teil (b) folgt durch Anwendung von Teil (a) auf die additiven Gruppen. Für Teil (c) reicht es, $\mathrm{Kern}(f) \subset \{0_K\}$ für jeden Körperhomomorphismus $f\colon K \to L$ zu zeigen (denn nach der Bemerkung ist dies gleichbedeutend mit $\mathrm{Kern}(f) = \{0_K\}$ und gibt gemäß Teil (b) Injektivität). Dazu sei $x \in \mathrm{Kern}(f)$, also $x \in K$ mit $f(x) = 0_L$. Wäre $x \neq 0_K$, so bekämen wir mit $1_L = f(1_K) = f(xx^{-1}) = f(x)\odot f(x^{-1}) = 0_L \odot f(x^{-1}) = 0_L$ einen Widerspruch. Also ist $x = 0_K$, und $\mathrm{Kern}(f) \subset \{0_K\}$ ist gezeigt. □

Als nächsten erklären wir Unterstrukturen von Gruppen, Ringen und Körpern, mit denen wir ohne eine solche explizite Benennung tatsächlich schon oft umgegangen sind:

Definition III.3.8.

(a) Eine *Untergruppe* U einer Gruppe $(G,*)$ ist eine Teilmenge U von G mit $e_G \in U$ (für das neutrale Element e_G von $(G,*)$) sowie $g * h \in U$ und $g^{-1} \in U$ für alle $g,h \in U$.

(b) Ein *Unterring* U eines Rings $(R,+,\cdot)$ ist eine Untergruppe U von $(R,+)$ mit $1 \in U$ und $xy \in U$ für alle $x,y \in U$.

(c) Ein *Teilkörper* U eines Körpers $(K,+,\cdot)$ ist ein Unterring U von $(K,+,\cdot)$ mit $x^{-1} \in U$ für alle $x \in U\backslash\{0\}$.

Beispiele III.3.9.

- Für jedes $n \in \mathbb{N}$ ist $n\mathbb{Z}$ Untergruppe von $(\mathbb{Z},+)$, aber für $n \in \mathbb{N}\backslash\{1\}$ wegen $1 \notin n\mathbb{Z}$ kein Unterring von $(\mathbb{Z},+,\cdot)$.
- $\mathbb{Z}$ ist Unterring von $(\mathbb{Q},+,\cdot)$, und $\mathbb{Q}$ ist Teilkörper von $(\mathbb{R},+,\cdot)$.
- Für jedes $n \in \mathbb{N}$ ist A_n Untergruppe von $(\Sigma_n,\circ)$.

- Für eine Gruppe $(G,*)$ und eine abelsche Gruppe $(H,\circledast)$ sind $\mathrm{Aut}(G)$ Untergruppe von $(\{f \in \mathrm{Abb}(G) \mid f \text{ bijektiv}\},\circ)$ und $(\mathrm{Hom}(G,H),*)$ Untergruppe von $(\mathrm{Abb}(G,H),*)$.
- Für hier stets additiv betrachtete Restklassengruppen kann $\mathbb{Z}/2\mathbb{Z}$ wegen $\mathbb{Z}/2\mathbb{Z} \not\subset \mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \not\subset \mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \not\subset (\mathbb{Z}/2\mathbb{Z})^2$ schon formal keine Untergruppe von $\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$ oder $(\mathbb{Z}/2\mathbb{Z})^2$ sein.

  Interessanter ist aber, ob die algebraische Struktur von $\mathbb{Z}/2\mathbb{Z}$ in $\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$ oder $(\mathbb{Z}/2\mathbb{Z})^2$ enthalten ist, ob $\mathbb{Z}/2\mathbb{Z}$ also zumindest isomorph zu einer Untergruppe ist.

  Die Antwort lautet, dass $\mathbb{Z}/2\mathbb{Z}$ *nicht* isomorph zu irgendeiner Untergruppe von $\mathbb{Z}/3\mathbb{Z}$ ist, aber isomorph zur Untergruppe $\{[0]_{\mathbb{Z}/4\mathbb{Z}},[2]_{\mathbb{Z}/4\mathbb{Z}}\}$ von $\mathbb{Z}/4\mathbb{Z}$ und auch zu jeder der drei Untergruppen
  $\{([0]_{\mathbb{Z}/2\mathbb{Z}},[0]_{\mathbb{Z}/2\mathbb{Z}}),([0]_{\mathbb{Z}/2\mathbb{Z}},[1]_{\mathbb{Z}/2\mathbb{Z}})\}$,
  $\{([0]_{\mathbb{Z}/2\mathbb{Z}},[0]_{\mathbb{Z}/2\mathbb{Z}}),([1]_{\mathbb{Z}/2\mathbb{Z}},[0]_{\mathbb{Z}/2\mathbb{Z}})\}$,
  $\{([0]_{\mathbb{Z}/2\mathbb{Z}},[0]_{\mathbb{Z}/2\mathbb{Z}}),([1]_{\mathbb{Z}/2\mathbb{Z}},[1]_{\mathbb{Z}/2\mathbb{Z}})\}$ von $(\mathbb{Z}/2\mathbb{Z})^2$ ist.

Bemerkungen III.3.10.
- In jeder Gruppe $(G,*)$ ist die triviale Untergruppe $\{e_G\}$ die kleinste und ganz G die größte Untergruppe.

  In jedem Ring $(R,+,\cdot)$ ist der *Grundring* $R_0 := \{z_R \mid z \in \mathbb{Z}\}$ der kleinste und ganz R der größte Unterring. Dabei ist R_0 stets isomorph zu $\mathbb{Z}$, wenn $(R,+,\cdot)$ Charakteristik 0 hat, oder zu $\mathbb{Z}/n\mathbb{Z}$, wenn $(R,+,\cdot)$ Charakteristik $n \in \mathbb{N}$ hat.

  In jedem Körper $(K,+,\cdot)$ ist der *Primkörper* $K_0 := \{z_K \cdot n_K^{-1} \mid n,z \in \mathbb{Z}, n_K \neq 0 \text{ in } K\}$ der kleinste und ganz K der größte Teilkörper. Dabei ist K_0 stets isomorph zu $\mathbb{Q}$, wenn $(K,+,\cdot)$ Charakteristik 0 hat, oder zu $\mathbb{F}_p$, wenn $(K,+,\cdot)$ Charakteristik $p \in \mathbb{P}$ hat.
- Dass $U \subset G$ Untergruppe einer Gruppe $(G,*)$ ist, ist auch durch $U \neq \emptyset$ und $g^{-1} * h \in U$ für alle $g,h \in U$ charakterisiert.

  Dass $U \subset K$ Teilkörper eines Körpers $(K,+,\cdot)$ ist, ist charakterisiert durch $U\backslash\{0\} \neq \emptyset$ und $x-y \in U$ für alle $x,y \in U$ sowie $x^{-1}y \in U$ für alle $x,y \in U\backslash\{0\}$.
- Die Terminologie ist sinnvoll, weil mit ihr für eine Untergruppe U von $(G,*)$ auch $(U,*)$ eine Gruppe, für einen Unterring U von $(R,+,\cdot)$ auch $(U,+,\cdot)$ ein Ring und für einen Teilkörper U von $(K,+,\cdot)$ auch $(U,+,\cdot)$ ein Körper ist.
- Man schreibt kurz, dass $U \subset G$ Untergruppe, $U \subset R$ Unterring, $U \subset K$ Teilkörper ist.
- Die definierenden Eigenschaften einer Untergruppe U von $(G,*)$ werden auch so ausgedrückt, dass $\emptyset \neq U$ unter der Verknüpfung $*$ und unter Inversenbildung abgeschlossen ist. Analog ist ein Unterring unter Addition, additiver Inversenbildung und Multiplikation abgeschlossen, ein Teilkörper zusätzlich unter multiplikativer Inversenbildung.

Satz III.3.11. *Für jede Teilmenge $A \subset G$ in einer Gruppe $(G,*)$ gibt es eine bezüglich Mengen-Inklusion kleinste Untergruppe $U \subset G$ mit $A \subset U$. Dieses U heißt die von A* erzeugte Untergruppe *und wird mit $\langle A\rangle$ bezeichnet. Für $n \in \mathbb{N}$ und $g_1,g_2,\ldots,g_n \in G$ wird $\langle g_1,g_2,\ldots,g_n\rangle := \langle\{g_1,g_2,\ldots,g_n\}\rangle$ vereinbart. Analog versteht man* erzeugte Unterringe und Teilkörper.

Beweis. Es ist Existenz von $U = \langle A\rangle$ zu beweisen. Da es mit G selbst zumindest eine Untergruppe V von $(G,*)$ mit $A \subset V$ gibt, ist $\mathcal{S}_A := \{V \subset G \mid V \text{ Untergruppe von } (G,*), A \subset V\}$ nicht leer, und wir können $U := \bigcap \mathcal{S}_A \subset G$ setzen. Da $A \cup \{e_G\} \subset U$ gilt und die Abgeschlossenheitseigenschaften von Untergruppen bei beliebigem Durchschnitt erhalten bleiben, ist $U \in \mathcal{S}_A$ das gesuchte kleinste Element von $\mathcal{S}_A$. Für Unterringe und Teilkörper argumentiert man analog. □

Bemerkungen III.3.12.
- Die Konstruktion im Beweis benutzt man häufig, um die Existenz einer kleinsten (oft von einer Teilmenge erzeugten) Menge mit gewissen Durchschnitts-stabilen Eigenschaften zu begründen. Versionen dieses Arguments werden Sie im Lauf des Studiums an vielen Stellen wiedersehen.
- Ist die von einem einzelnen $g \in G$ erzeugte Untergruppe $\langle g\rangle$ in einer Gruppe $(G,*)$ endlich, so ist $(\langle g\rangle,*)$ zyklisch von Ordnung $|\langle g\rangle|$ und hat g (im für zyklische Gruppen erklärten Sinn) als Erzeuger. Tatsächlich kann eine zyklische Gruppe äquivalent als eine Gruppe $(G,*)$ definiert werden, für die $|G| < \infty$ gilt und für die ein einzelnes $g \in G$ mit $\langle g\rangle = G$ existiert.

Beispiel III.3.13. In der additiven Gruppe $(\mathbb{Z},+)$ ist $\langle x\rangle = x\mathbb{Z}$ für jedes $x \in \mathbb{Z}$.

Ein zu Unterstrukturen duales Konzept sind sogenannte Faktorstrukturen, an deren Definition wir uns nun unter Rückgriff auf Quotienten einer Äquivalenzrelation annähern:

Satz III.3.14. *Für jede Untergruppe U einer Gruppe $(G,*)$ wird durch*

$$x \sim_U y :\Longleftrightarrow x^{-1} * y \in U \text{ für } x, y \in G$$

eine Äquivalenzrelation $\sim_U$ auf G definiert, bei der $\left|[y]_{\sim_U}\right| = \left|[x]_{\sim_U}\right|$ für alle $x, y \in G$ gilt, also alle Äquivalenzklassen

$$[x]_{\sim_U} = x * U$$

mit $x \in G$ gleiche Kardinalität haben.

Wir nennen die Äquivalenzklassen bezüglich $\sim_U$ auch Linksnebenklassen und vereinbaren die Bezeichnung

$$G/U := G/\sim_U = \{[x]_{\sim_U} \mid x \in G\}$$

für die Quotientenmenge. Insbesondere greift dies bei einem Ring oder Körper $(R, +, \cdot)$ für jede Untergruppe U von $(R, +)$ mit der durch $x \sim_U y \iff x - y \in U$ gegebenen Äquivalenzrelation.

Beweis. Wir verifizieren zunächst die definierenden Eigenschaften der Äquivalenzrelation $\sim_U$: Reflexivität liegt vor, weil $x^{-1} * x = e_G \in U$ für alle $x \in G$ gilt. Symmetrie ist erfüllt, weil für $x, y \in G$ aus $x^{-1} * y \in U$ schon $y^{-1} * x = (x^{-1} * y)^{-1} \in U$ folgt. Transitivität ergibt sich, weil für $x, y, z \in G$ aus $x^{-1} * y \in U$ und $y^{-1} * z \in U$ auch $x^{-1} * z = (x^{-1} * y) * (y^{-1} * z) \in U$ folgt.

Die Gleichheit $[x]_{\sim_U} = x * U$ ergibt sich aus $x \sim_U y \iff x^{-1} * y \in U \iff y \in x * U$ für $y \in G$.

Weiter erhalten wir für feste $x, y \in G$ durch $\lambda(z) := y * x^{-1} * z$ für $z \in G$ eine Abbildung $\lambda\colon [x]_{\sim_U} \to [y]_{\sim_U}$, denn $z \in [x]_{\sim_U} = x * U$ impliziert

$$y * x^{-1} * z \in y * x^{-1} * x * U = y * U = [y]_{\sim_U}.$$

Analog gibt $\mu(z) := x * y^{-1} * z$ eine Abbildung $\mu\colon [y]_{\sim_U} \to [x]_{\sim_U}$, die die Umkehrabbildung zu λ ist. Somit ist λ eine Bijektion und es gilt $\left|[y]_{\sim_U}\right| = \left|[x]_{\sim_U}\right|$. □

Bemerkung III.3.15. Es sei U eine Untergruppe einer Gruppe $(G,*)$. Eine sehr ähnliche Äquivalenzrelation auf G erhält man durch die Festlegung $x \,{}_U\!\sim y :\Longleftrightarrow y * x^{-1} \in U$ für $x, y \in G$. Die Äquivalenzklassen $[x]_{U\sim} = U * x$ heißen Rechtsnebenklassen. Sie verhalten sich weitgehend analog zu den Linksnebenklassen, sind aber im nicht-abelschen Fall im Allgemeinen verschieden von diesen.

Als ein naheliegendes Beispiel für den Nutzen von Nebenklassen beweisen wir den Satz von Lagrange über grundlegende Kennzahlen (die wir vorher noch kurz definieren) von Gruppen, Untergruppen und ihren Elementen:

Definition III.3.16.

(a) Eine Gruppe $(G,*)$ heißt *endlich*, wenn $|G| < \infty$ ist, und $|G| \in \mathbb{N}$ heißt dann die *Ordnung der Gruppe* $(G,*)$.

(b) Die *Ordnung eines Elements* $g \in G$ in einer Gruppe $(G,*)$ ist – wenn existent – eine Zahl $m \in \mathbb{N}$ mit $g^m = e_G$ und $g^k \neq e_G$ für alle $k \in \{1, 2, \ldots, m-1\}$, wobei e_G das neutrale Element von $(G,*)$ bezeichnet. Existiert kein solches m, so heißt g von unendlicher Ordnung.

Satz III.3.17 (Satz von Lagrange)**.** *Für eine endliche Gruppe $(G,*)$ sind die Ordnungen der Untergruppen von G und die Ordnungen der Elemente von G alle Teiler der Ordnung $|G|$ der Gruppe.*

Beweis. Sei U eine Untergruppe von $(G,*)$. Nach dem vorigen Satz enthält jede Nebenklasse $[x]_{\sim_U} \in G/U$ genau so viele Elemente wie $U = [e_G]_{\sim_U} \in G/U$, es gilt also $|N| = |U|$ für alle Nebenklassen $N \in G/U$. Da G die disjunkte Vereinigung der $|G/U|$ Nebenklassen in $|G/U|$ ist, ergibt sich mit $|G| = |G/U| \cdot |U|$, dass $|U|$ ein Teiler von $|G|$ ist. Dies zeigt die Behauptung über die Ordnungen der Untergruppen.

Ein Element $g \in G$ hat wegen $|G| < \infty$ zunächst eine Ordnung $m \in \mathbb{N}$. Andernfalls wäre $g^k \neq e_G$ für alle $k \in \mathbb{N}$ und damit wären g, g^2, g^3, g^4, g^5, … unendliche viele verschiedene Elemente von G. Nun muss die von g erzeugte Untergruppe $\langle g \rangle$ gleich $\{e_G, g, g^2, g^3, \ldots, g^{m-1}\}$ sein und $|\langle g \rangle| = m$ erfüllen. Nach dem schon Gezeigten ist daher m ein Teiler von $|G|$ und auch die Behauptung über die Ordnungen der Elemente verifiziert. □

Beispiele III.3.18. Wir klassifizieren alle Elemente und Untergruppen nach ihrer Ordnung:

- In der additiven *Restklassengruppe* $(\mathbb{Z}/12\mathbb{Z}, +)$ mit $|\mathbb{Z}/12\mathbb{Z}| = 12$ gilt

Ordnung	1	2	3	4	6	12
Elemente	$[0]$	$[6]$	$[4], [8]$	$[3], [9]$	$[2], [10]$	$[1], [5], [7], [11]$
Untergruppen	$\{[0]\}$	$\{[0], [6]\}$	$\{[0], [4], [8]\}$	$\langle 3\rangle = \langle 9\rangle$	$\langle 2\rangle = \langle 10\rangle$	$\langle 1\rangle = \mathbb{Z}/12\mathbb{Z}$

  Tatsächlich sind in diesem Beispiel und allgemein in $(\mathbb{Z}/n\mathbb{Z}, +)$ mit $n \in \mathbb{N}$ alle Untergruppen zyklisch und jede aufgrund des Satzes von Lagrange zulässige Ordnung wird in $(\mathbb{Z}/n\mathbb{Z}, +)$ durch genau eine Untergruppe realisiert.
- In der symmetrischen Gruppe $(\Sigma_3, \circ)$ mit $|\Sigma_3| = 6$ (Bezeichnungen aus Abschnitt III.1) erhalten wir:

Ordnung	1	2	3	6
Elemente	id	$\tau_1 = (1,2)$, $\tau_2 = (1,3)$, $\tau_3 = (2,3)$	$\sigma_1 = (1,2,3)$, $\sigma_2 = (1,3,2)$	—
Untergruppen	$\{\mathrm{id}\}$	$\langle\tau_1\rangle = \{\mathrm{id}, \tau_1\}, \langle\tau_2\rangle, \langle\tau_3\rangle$	$\langle\sigma_1\rangle = \langle\sigma_2\rangle = \{\mathrm{id}, \sigma_1, \sigma_2\}$	Σ_3

  In diesem Beispiel sind alle echten Untergruppen (also alle außer Σ_3 selbst) zyklisch.
- In der alternierenden Gruppe $(A_4, \circ)$ mit $|A_4| = 12$ (Bezeichnungen aus Abschnitt III.1):

Ordnung	1	2	3	4	6	12
Elemente	id	$\vartheta_1, \vartheta_2, \vartheta_3$	$\eta_1,\eta_2,\eta_3,\eta_4,\eta_5,\eta_6,\eta_7,\eta_8$	———————	–	—
Untergruppen	$\{\mathrm{id}\}$	$\langle\vartheta_1\rangle,\langle\vartheta_2\rangle,\langle\vartheta_3\rangle$	$\langle\eta_1\rangle{=}\langle\eta_2\rangle,\langle\eta_3\rangle,\langle\eta_5\rangle,\langle\eta_7\rangle$	$\{\mathrm{id}, \vartheta_1, \vartheta_2, \vartheta_3\}$	–	A_4

  In diesem Beispiel gibt es keine Untergruppe der aufgrund des Satzes von Lagrange zulässigen Ordnung 6, und neben A_4 ist auch die echte Untergruppe $\{\mathrm{id}, \vartheta_1, \vartheta_2, \vartheta_3\}$ nicht zyklisch. Diese Untergruppe ist die berühmte Kleinsche Vierergruppe.

Als interessante Folgerung ergibt sich ein berühmter Grundsatz der Zahlentheorie:

Korollar III.3.19 (Kleiner Satz von Fermat)**.** *Es sei $p \in \mathbb{P}$ eine Primzahl. Dann gelten*

$$x^p = x \mod p \text{ für alle } x \in \mathbb{Z}$$

und

$$x^{p-1} = 1 \mod p \text{ für alle } x \in \mathbb{Z} \setminus p\mathbb{Z}.$$

Bemerkung III.3.20. Die Behauptung des Satzes kann ganz elementar formuliert werden: Ist $p \in \mathbb{P}$ eine Primzahl, so ist für jedes $x \in \mathbb{Z}$ entweder x oder $x^{p-1}-1$ durch p teilbar.

Der Beweis lässt sich mit den inzwischen erarbeiteten Techniken kurz und elegant führen:

Beweis des Kleinen Satzes von Fermat. Für $x \in \mathbb{Z} \setminus p\mathbb{Z}$ ist $[x]_{\mathbb{Z}/p\mathbb{Z}} \neq [0]_{\mathbb{Z}/p\mathbb{Z}}$ in $\mathbb{Z}/p\mathbb{Z}$, womit $[x]_{\mathbb{Z}/p\mathbb{Z}}$ ein Element der multiplikativen Gruppe $(\mathbb{Z}/p\mathbb{Z}^\times, \cdot)$ der Ordnung $|\mathbb{Z}/p\mathbb{Z}^\times| = p-1$ ist. Dass dies tatsächlich eine Gruppe ist, wurde in Abschnitt III.1 gezeigt und war nicht ganz einfach. Davon profitieren wir nun. Nach dem Satz von Lagrange ist die Ordnung $m \in \mathbb{N}$ von $[x]_{\mathbb{Z}/p\mathbb{Z}}$ in $(\mathbb{Z}/p\mathbb{Z}^\times, \cdot)$ ein Teiler von $p-1$, also $p-1 = \ell m$ für ein $\ell \in \mathbb{N}$. Es folgt

$$\left[x^{p-1}\right]_{\mathbb{Z}/p\mathbb{Z}} = [x]^{p-1}_{\mathbb{Z}/p\mathbb{Z}} = \left([x]^m_{\mathbb{Z}/p\mathbb{Z}}\right)^\ell = [1]^\ell_{\mathbb{Z}/p\mathbb{Z}} = [1]_{\mathbb{Z}/p\mathbb{Z}},$$

was $x^{p-1} = 1 \mod p$ bedeutet. Dies zeigt die zweite Behauptung. Die erste Behauptung folgt für $x \in \mathbb{Z} \setminus p\mathbb{Z}$ durch Multiplikation mit x. Sie gilt für $x \in p\mathbb{Z}$ mit $x^p = 0 = x \mod p$ aber ebenfalls. □

Schließlich möchten wir die Quotientenmenge G/U beziehungsweise R/U (die wir für jede (additive) Untergruppe U bilden können) wieder mit Verknüpfungen versehen und selbst zu einer Gruppe beziehungsweise einem Ring machen. Dafür benötigen wir für U tatsächlich noch eine etwas andere Struktur:

Definition III.3.21.

(a) Ein *Normalteiler* N einer Gruppe $(G, *)$ ist eine Untergruppe N von $(G, *)$ mit $g * N = N * g$ für alle $g \in G$.

(b) Ein *(beidseitiges) Ideal* I eines Rings $(R,+,\cdot)$ ist eine Untergruppe I von $(R,+)$ mit $rI \subset I$ und $Ir \subset I$ für alle $r \in R$.

Bemerkungen III.3.22.

(a) Ein Normalteiler ist also eine Untergruppe, für die die Links- und Rechtsnebenklassen übereinstimmen, und kann alternativ als Untergruppe N mit $g{*}N{*}g^{-1} \subset N$ für alle $g \in G$ charakterisiert werden. In einer *abelschen* Gruppe ist dies immer der Fall, so dass dort die Normalteiler nichts anderes als Untergruppen sind.

(b) Ein Ideal muss 1 nicht enthalten und daher kein Unterring sein, beispielsweise ist $x\mathbb{Z}$ für jedes $x \in \mathbb{Z}$ ein Ideal im Ring $(\mathbb{Z},+,\cdot)$, aber nur in den trivialen Fällen $x \in \{-1,1\}$ mit $x\mathbb{Z} = \mathbb{Z}$ ein Unterring in $(\mathbb{Z},+,\cdot)$. Andererseits muss ein Unterring kein Ideal sein, beispielsweise ist $\mathbb{Z}$ ein Unterring von $(\mathbb{Q},+,\cdot)$, aber kein Ideal in $(\mathbb{Q},+,\cdot)$ (denn beispielsweise ist $\frac{1}{2}\mathbb{Z} \not\subset \mathbb{Z}$).

(c) Analog zu erzeugten Untergruppen, Unterringen oder Teilkörpern lassen sich der von einer Teilmenge A *erzeugte Normalteiler* in einer Gruppe und das von einer Teilmenge A *erzeugte Ideal* in einem Ring definieren. Auch für diese schreibt man manchmal $\langle A \rangle$.

(d) Der Kern eines Gruppenhomomorphismus $f\colon G \to H$ zwischen Gruppen $(G,*)$ und $(H,\circledast)$ ist stets ein Normalteiler in $(G,*)$. Der Kern eines Ringhomomorphismus $\psi\colon R \to S$ zwischen Ringen $(R,+,\cdot)$ und $(S,\oplus,\odot)$ ist immer ein Ideal in $(R,+,\cdot)$:

Dass die Kerne (additive) Untergruppen sind, sieht man problemlos. Für $g \in G$, $n \in \operatorname{Kern}(f)$ ist zudem $f(g{*}n{*}g^{-1}) = f(g)\circledast e_H \circledast f(g)^{-1} = e_H$ und damit $g{*}\operatorname{Kern}(f){*}g^{-1} \subset \operatorname{Kern}(f)$. Somit ist also $\operatorname{Kern}(f)$ ein Normalteiler. Für $r \in R$ und $n \in \operatorname{Kern}(\psi)$ ist $\psi(rn) = \psi(r)0 = 0$ und analog $\psi(nr) = 0$. Dies bedeutet $r\operatorname{Kern}(\psi) \subset \operatorname{Kern}(\psi)$ und $\operatorname{Kern}(\psi)r \subset \operatorname{Kern}(\psi)$, so dass $\operatorname{Kern}(\psi)$ ein Ideal ist.

Satz III.3.23.

(a) *Ist N ein Normalteiler in einer Gruppe $(G,*)$, so erhalten wir durch*

$$[g]_{\sim_N} * [h]_{\sim_N} := [g{*}h]_{\sim_N} \text{ für } g,h \in G$$

beziehungsweise äquivalent durch

$$(g{*}N) * (h{*}N) := (g{*}h){*}N \text{ für } g,h \in G$$

eine wohldefinierte Verknüpfung auf G/N, mit der G/N zu einer Gruppe wird. Diese ist abelsch, falls $(G,*)$ abelsch ist. Wir nennen dann $(G/N,*)$ die Faktorgruppe *oder* Quotientengruppe *von G nach N.*

(b) *Ist I ein Ideal in einem Ring $(R,+,\cdot)$, so erhalten wir durch*

$$[r]_{\sim_I} + [s]_{\sim_I} := [r{+}s]_{\sim I} \text{ und } [r]_{\sim_I} \cdot [s]_{\sim_I} := [rs]_{\sim I} \text{ für } r,s \in R$$

beziehungsweise äquivalent durch

$$(r{+}I) + (s{+}I) := (r{+}s){+}I \text{ und } (r{+}I)\cdot(s+I) := (rs){+}I \text{ für } r,s \in R$$

wohldefinierte Verknüpfungen auf R/I, mit denen R/I ein Ring wird. Dieser ist kommutativ, falls $(R,+,\cdot)$ kommutativ ist. Wir nennen dann $(R/I,+,\cdot)$ den Restklassenring *von R nach I.*

Beispiel III.3.24. Ein wichtigstes Beispiel von Faktorgruppen und Restklassenringen sind die Restklassenringe $\mathbb{Z}/n\mathbb{Z}$, die wir bereits ausführlich kennengelernt haben. An dieser Stelle können wir sie aber als Spezialfälle einer übergeordneten Theorie verstehen.

Bemerkung III.3.25. Für einen Körper $(K,+,\cdot)$ sind $\{0\}$ und ganz K die einzigen Ideale von $(K,+,\cdot)$. Für das eine ist $K/\{0\}$ nur eine neue Version von K, für das andere hat K/K nur ein Element und wird zum Nullring, aber nicht zu einem Körper.

Zum Beweis des Satzes. Entscheidend ist der Beweis der Wohldefiniertheit der Verknüpfungen:

Bezüglich Teil (a) verifizieren wir für alternative Repräsentanten $g' \in g*N$, $h' \in h*N$ der Nebenklassen dazu $g'*h' \in (g*h)*N$ durch die Rechnung

$$g'*h' \in g*N*h*N = g*h*N*N = g*h*N,$$

bei der entscheidend eingeht, dass N Normalteiler ist.

Bezüglich Teil (b) bemerken wir zunächst, dass die Addition durch Teil (a) abgedeckt ist, weil I als Untergruppe in der *abelschen* Gruppe $(R,+)$ automatisch Normalteiler ist. Für die Multiplikation betrachten wir $r' \in r+I$, $s' \in s+I$ und bekommen durch die Rechnung

$$r's' \in (r+I)\cdot(s+I) \subset rs + rI + Is + I\cdot I \subset rs + I$$

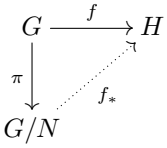
unter Verwendung der Ideal-Eigenschaft, dass $r's' \in [rs]_{\sim_I} = rs + I$ gilt.

Alles Weitere (Assoziativität, Kommutativität, neutrale und inverse Elemente) erhält man aus den entsprechenden Eigenschaften von G beziehungsweise R. □

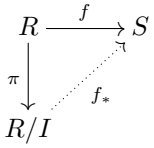
Zum Abschluss des Kapitels erwähnen wir die Faktorisierungssätze für Gruppen und für Ringe, bei denen es sich um weitgehende Analoga des Faktorisierungssatzes für Äquivalenzrelationen handelt:

Satz III.3.26.

(a) *Es seien $(G,*)$ und $(H,\circledast)$ Gruppen, N ein Normalteiler von $(G,*)$ mit der Quotientenabbildung $\pi\colon G \to G/N$ und $f\colon G \to H$ ein Gruppenhomomorphismus mit $N \subset \mathrm{Kern}(f)$. Dann gibt es genau einen Gruppenhomomorphismus $f_*\colon G/N \to H$, der $f_* \circ \pi = f$ erfüllt, also das folgende Diagramm kommutativ macht.*



(b) *Es seien $(R,+,\cdot)$ und $(S,\oplus,\odot)$ Ringe, I ein Ideal von $(R,+,\cdot)$ mit der Quotientenabbildung $\pi\colon R \to R/I$ und $f\colon R \to S$ ein Ringhomomorphismus mit $I \subset \mathrm{Kern}(f)$. Dann gibt es genau einen Ringhomomorphismus $f_*\colon R/I \to S$, der $f_* \circ \pi = f$ erfüllt, also das folgende Diagramm kommutativ macht.*



Bemerkung III.3.27. Gemäß der Bemerkung zu Normalteilern und Idealen ist der Satz stets mit $N = \mathrm{Kern}(f)$ beziehungsweise $I = \mathrm{Kern}(f)$ anwendbar, und genau in diesem Fall ist f_* injektiv. Außerdem gilt stets $\mathrm{Bild}(f_*) = \mathrm{Bild}(f)$, und insbesondere ist f_* genau dann surjektiv, wenn f surjektiv ist.

Beweis der Faktorisierungssätze. Die Existenz und Eindeutigkeit einer Abbildung f_* mit $f_* \circ \pi = f$ ergibt sich aus dem Satz des Abschnitts II.4.2 über die Faktorisierung nach einer Äquivalenzrelation, wenn man bedenkt, dass im Gruppenfall

$$x \sim_N y \iff y^{-1} * x \in N \implies y^{-1} * x \in \mathrm{Kern}(f)$$

und dies ist genau dann der Fall, wenn

$$f(y^{-1} * x) = e_G \iff f(x) = f(y)$$

für $x, y \in N$ gilt. Im Ringfall

$$x \sim_I y \iff x - y \in I \implies x - y \in \mathrm{Kern}(f)$$

und dies ist genau dann der Fall, wenn $f(x-y) = 0$ gilt, also $f(x) = f(y)$ für $x, y \in R$ gilt. Gemäß dem vorigen Satz kann zudem $(G/N,*)$ als Gruppe beziehungsweise $(R/I,+,\cdot)$ als Ring aufgefasst werden, und es bleibt nur die Homomorphismus-Eigenschaft von f_* nachzurechnen. Im Gruppenfall gelingt dies mit der Rechnung

$$f_*([g_1]_{\sim_N} * [g_2]_{\sim_N}) = f_*([g_1 * g_2]_{\sim_N}) = f(g_1 * g_2) = f(g_1) \circledast f(g_2) = f_*([g_1]_{\sim_N}) \circledast f_*([g_2]_{\sim_N})$$

für $[g_1]_{\sim_N}, [g_2]_{\sim_N} \in G/N$. Im Ringfall kann man analog vorgehen. □

KAPITEL IV

Reelle und komplexe Zahlen

In diesem Kapitel schließen wir die Diskussion der Zahlbereiche ab, indem wir nach $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ in Kapitel II auch die für die Analysis grundlegenden Bereiche der *reellen Zahlen* $\mathbb{R}$ und der *komplexen Zahlen* $\mathbb{C}$ präzise einführen.

IV.1. Reelle Zahlen

Die Menge $\mathbb{R}$ der reellen Zahlen kann man sich als die Menge aller *Dezimalzahlen (mit endlich oder unendlich vielen Nachkommastellen)* oder als die Menge aller *Punkte der Zahlengerade* vorstellen. Neben den rationalen Zahlen aus $\mathbb{Q}$ gehören zu $\mathbb{R}$ auch sogenannte *irrationale Zahlen* aus $\mathbb{R}\setminus\mathbb{Q}$ wie zum Beispiel Wurzeln aus ganzen Zahlen, die keine Quadratzahlen sind, Dezimalzahlen mit unendlich vielen nicht-periodischen Nachkommastellen, die Eulersche Zahl e, die Kreiszahl π, viele aus solchen gebildete algebraische Ausdrücke und überabzählbar viele weitere Zahlen.

Zur präzisen Einführung der reellen Zahlen gibt es verschiedene Möglichkeiten. Hier geben wir zunächst ein Axiomensystem an:

Axiom IV.1.1 (Axiome der reellen Zahlen)**.** Es gibt eine Menge $\mathbb{R}$, zwei Verknüpfungen $+$ und $\cdot$ auf $\mathbb{R}$ sowie eine Relation $<$ auf $\mathbb{R}$ mit folgenden Eigenschaften:

- *Körperaxiome:* Mit den beiden Verknüpfungen wird $(\mathbb{R}, +, \cdot)$ ein Körper.
- *Anordnungsaxiome:* Die Relation $<$ ist eine strikte Totalordnung auf $\mathbb{R}$ und ist kompatibel mit der Addition $+$ und der Multiplikation $\cdot$ in dem Sinne, dass für alle $r, x, y \in \mathbb{R}$ gilt:
$$x < y \implies r+x < r+y \text{ und } 0 < r, x < y \implies r\cdot x < r\cdot y.$$
- *Metrische Vollständigkeit:* Jede Intervallschachtelung in $\mathbb{R}$ besitzt einen Kern in $\mathbb{R}$.
- *Archimedisches Axiom:* Für jedes $x \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$ mit $x < n$.

Zu den Axiomen sind noch einige Erläuterungen zu geben und insbesondere die Begriffe der Intervallschachtelung und des Kerns einer solchen überhaupt einmal zu definieren. Wir gehen dies teils direkt, teils aber auch später in diesem Abschnitt an.

Bemerkungen IV.1.2.

- Wie üblich bezeichnen 0 und 1 das Null- und das Einselement des Körpers $\mathbb{R}$, und wir setzen wie vorher $n = \sum_{i=1}^{n} 1 \in \mathbb{R}$ für $n \in \mathbb{N}$. Außerdem finden in $\mathbb{R}$ alle allgemein für Körper besprochenen Regeln, Definitionen und Konventionen Anwendung.
- In der Kurzzusammenfassung besagen die Axiome:
$$\mathbb{R} \text{ ist ein } \textit{vollständiger, Archimedisch angeordneter Körper}.$$
- Für die rationalen Zahlen $\mathbb{Q}$ sind alle Axiome außer der Vollständigkeit erfüllt. In der Vollständigkeit besteht also der entscheidende Unterschied zwischen $\mathbb{Q}$ und $\mathbb{R}$.

  Die komplexen Zahlen $\mathbb{C}$, die wir später einführen, bilden einen (in geeignetem Sinn) vollständigen Körper, erlauben aber *keine* mit Addition und Multiplikation kompatible Anordnung.

Definition IV.1.3.

- Wir nennen eine Zahl $x \in \mathbb{R}$ *positiv*, wenn $0 < x$ gilt, und andernfalls *nichtpositiv*. Analog nennen wir $x \in \mathbb{R}$ *negativ*, wenn $x < 0$ gilt, und andernfalls *nichtnegativ*.
- Ausgehend von $<$ können wir wie üblich die nicht-strikte Totalordnung $\leq$ durch
$$x \leq y :\iff (x < y \vee x = y) \text{ für } x, y \in \mathbb{R}$$

einführen sowie $>$ und $\geq$ als die Umkehrrelationen zu $<$ und $\leq$ erklären.

Bemerkungen IV.1.4.

- Aus den Anordnungsaxiomen folgen die Regeln für $n \in \mathbb{N}$, $r, s, x, y, x_i, y_i \in \mathbb{R}$:

$$n > 0,$$
$$x \text{ ist positiv } \iff -x \text{ ist negativ},$$
$$x^2 \geq 0 \text{ mit „=“ nur falls } x = 0,$$
$$\sum_{i=1}^{n} x_i^2 \geq 0 \text{ mit „=“ nur falls alle } x_i = 0,$$
$$x_i \leq y_i \text{ für } i \in \{1, 2, \ldots, n\} \implies \sum_{i=1}^{n} x_i \leq \sum_{i=1}^{n} y_i \text{ mit „=“ nur falls alle } x_i = y_i,$$
$$0 < x_i \leq y_i \text{ für } i \in \{1, 2, \ldots, n\} \implies \prod_{i=1}^{n} x_i \leq \prod_{i=1}^{n} y_i \text{ mit „=“ nur falls alle } x_i = y_i,$$
$$\text{Ist } x < y, \text{dann gilt: } r < 0 \iff r \cdot x > r \cdot y,$$
$$0 < r < s \implies \frac{1}{r} > \frac{1}{s} > 0.$$

  Insbesondere ist $n \neq 0$ für alle $n \in \mathbb{N}$, weshalb der Körper $(\mathbb{R}, +, \cdot)$ der reellen Zahlen Charakteristik 0 hat und sein Primkörper mit dem Körper $(\mathbb{Q}, +, \cdot)$ der rationalen Zahlen identifiziert werden kann. Wir verstehen daher immer

$$\mathbb{Q} \subset \mathbb{R}.$$

  Weiterhin muss man sich unbedingt merken, dass die Multiplikation mit einem negativen Faktor und die Reziprokenbildung $\frac{1}{(\cdot)}$ bei positiven Zahlen das Ungleichheitszeichen umkehren. Konkrete Beispiel sind:

$$4 < 7. \text{ aber } -4 > -7 \text{ und } \frac{1}{4} > \frac{1}{7}.$$

  Wir beweisen nicht alle hier genannten Regeln, zeigen aber zwei ganz grundlegende Argumente:

  Zur Begründung für $n > 0$: Wäre $1 < 0$, so folgt durch Addition von -1 auch $0 < -1$, durch Multiplikation mit -1 dagegen $-1 < 0$. Widerspruch! Da somit weder $1 < 0$ noch $1 = 0$ gilt, ist zwingend $1 > 0$. Durch Addition von $n-1$ folgt $n > n-1$ für $n \in \mathbb{N}$ und mit Transitivität $n > 0$ für $n \in \mathbb{N}$.

  Zur Begründung für $x^2 \geq 0$: Für $x > 0$ gilt $x^2 = x \cdot x > 0 \cdot x = 0$ und daher insgesamt $x^2 > 0$. Für $x < 0$ gilt $x^2 = (-x) \cdot (-x) > 0 \cdot (-x) = 0$ und damit wieder $x^2 > 0$. Für $x = 0$ ist natürlich $x^2 = 0^2 = 0$.

Bevor wir nun die restlichen Axiome von $\mathbb{R}$ genauer besprechen, geben wir erst einige grundlegende Definitionen:

Definition IV.1.5. Die *erweiterten reellen Zahlen* sind

$$\overline{\mathbb{R}} := \mathbb{R} \sqcup \{-\infty, \infty\}$$

mit zusätzlichen Symbolen ∞ (Unendlich) und $-\infty$ (minus Unendlich). Wir setzen die Relation $<$ durch die Festlegung $-\infty < x < \infty$ für alle $x \in \mathbb{R}$ zu einer strikten Totalordnung auf $\overline{\mathbb{R}}$ fort.

Bemerkung IV.1.6. Wie wir in Mathematik 2 sehen, kann man mit ∞ und $-\infty$ zu einem gewissen Grad rechnen, aber zum Beispiel $\infty + (-\infty)$ *nicht* sinnvoll erklären. Daher wird $\overline{\mathbb{R}}$ *nicht* zu einem Körper oder überhaupt einer algebraischen Struktur im Sinn des Kapitels III.

Definition IV.1.7.

(a) Eine Zahl $x \in \overline{\mathbb{R}}$ *liegt zwischen* $r \in \overline{\mathbb{R}}$ und $s \in \overline{\mathbb{R}}$, wenn $r \leq x \leq s$ oder $s \leq x \leq r$ gilt, und sie *liegt* echt *zwischen* r und s, wenn $r < x < s$ oder $s < x < r$ gilt.

(b) Eine Teilmenge $I \subset \overline{\mathbb{R}}$ heißt ein *Intervall*, wenn I mit zwei Zahlen stets auch alle zwischen diesen liegenden enthält.

(c) Spezielle Intervalle sind

$$(a,b) :=]a,b[:= \{x \in \overline{\mathbb{R}} \mid a < x < b\} \text{ } \textit{offen},$$
$$(a,b] :=]a,b] := \{x \in \overline{\mathbb{R}} \mid a < x \leq b\} \text{ links } \textit{halboffen},$$
$$[a,b) := [a,b[:= \{x \in \overline{\mathbb{R}} \mid a \leq x < b\} \text{ rechts } \textit{halboffen},$$
$$[a,b] := \{x \in \overline{\mathbb{R}} \mid a \leq x \leq b\} \text{ } \textit{abgeschlossen}$$

mit *Randpunkten* $a, b \in \overline{\mathbb{R}}$, wobei oft nur $a < b$ bzw. $a \leq b$ betrachtet wird, weil man andernfalls $\emptyset$ erhält. Intervalle des Typs $[a,b]$ mit $a, b \in \mathbb{R}$ nennt man *kompakt*.

(d) Wir vereinbaren als naheliegenden Abkürzungen $\mathbb{R}_{>a} := (a,\infty)$, $\mathbb{R}_{<b} := (-\infty,b)$, $\mathbb{R}_{\geq a} := [a,\infty)$, $\mathbb{R}_{\leq b} := (-\infty,b]$ und bemerken, dass $\mathbb{R} = (-\infty,\infty)$, $\overline{\mathbb{R}} = [-\infty,\infty]$.

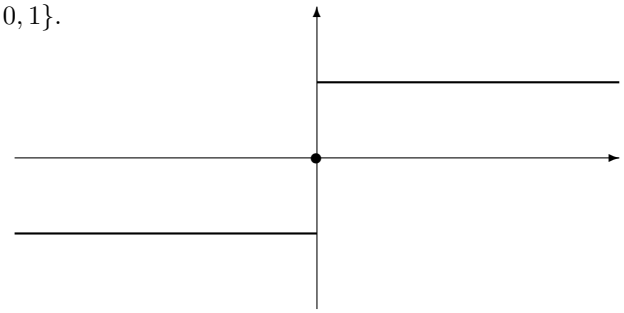
Bemerkung IV.1.8. Dass die speziellen Intervalle aus Teil (d) der Definition tatsächlich Intervalle sind, verifiziert man problemlos. Demnächst begründen wir außerdem, dass tatsächlich *alle* Intervalle in $\overline{\mathbb{R}}$ von dieser Form sind.

Definition IV.1.9.

(a) Die reelle *Vorzeichenfunktion* oder reelle *Signumfunktion* ist für $x \in \mathbb{R}$ definiert als:

$$\operatorname{sign}\colon \mathbb{R} \to \mathbb{R}, \quad \operatorname{sign}(x) := \begin{cases} 1, & \text{falls } x > 0, \\ 0, & \text{falls } x = 0, \\ -1, & \text{falls } x < 0 \end{cases}$$

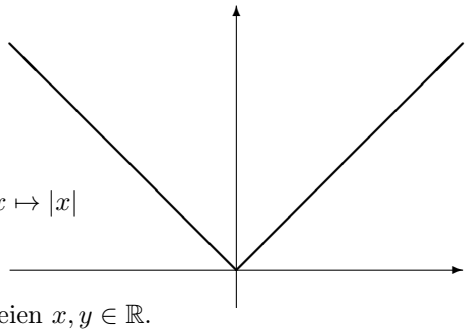
mit $\operatorname{Bild}(\operatorname{sign}) = \{-1, 0, 1\}$.



(b) Die reelle *Betragsfunktion* ist für $x \in \mathbb{R}$ definiert als:

$$|\cdot|\colon \mathbb{R} \to \mathbb{R}, \quad |x| := \begin{cases} x, & \text{falls } x \geq 0, \\ -x, & \text{falls } x \leq 0 \end{cases}$$

mit $\operatorname{Bild}(|\cdot|) = \mathbb{R}_{\geq 0}$.



Graph von $x \mapsto |x|$

Bemerkungen IV.1.10. Es seien $x, y \in \mathbb{R}$.

- Man kann den Betrag $|x|$ von x als den Abstand von x zu 0 auf der Zahlengerade und $|y-x|$ als den Abstand von y zu x auf den Zahlengerade interpretieren.

- Aus den Definitionen ergeben sich die Rechenregeln

$$\begin{aligned} x &= \operatorname{sign}(x)|x|, \\ \operatorname{sign}(xy) &= \operatorname{sign}(x)\operatorname{sign}(y), \\ x^2 &= |x|^2, \\ |x| = |y| &\iff (x = y \vee x = -y), \\ |x| \le y &\iff x \in [-y, y]. \end{aligned}$$

  Ist $y \neq 0$, so gilt auch

$$\operatorname{sign}(xy) = \operatorname{sign}(x)\operatorname{sign}(y) = \operatorname{sign}\Big(\frac{x}{y}\Big).$$

Beispiel IV.1.11. Das Auflösen von (Un-)Gleichungen mit Beträgen gelingt oft mit Fallunterscheidungen. Ein konkretes Beispiel ist die Ungleichung

$$\frac{1}{|x|} \ge \frac{1}{1-x} \text{ für } x \in \mathbb{R} \setminus \{0, 1\},$$

bei der die Unterscheidung folgender 3 Fälle sinnvoll ist:

- Fall $x < 0$: Wir multiplizieren mit $-x = |x| > 0$ und $1-x > 0$ und erhalten als äquivalente Ungleichung erst $1-x \ge -x$, dann $1 \ge 0$, was generell erfüllt ist.
- Fall $0 < x < 1$: Wir multiplizieren mit $x = |x| > 0$ und $1-x > 0$ und erhalten als äquivalente Ungleichung erst $1-x \ge x$, dann $x \le \frac{1}{2}$.
- Fall $x > 1$: Wir multiplizieren mit $x = |x| > 0$ und $1-x < 0$ und erhalten als äquivalente Ungleichung erst $1-x \le x$, dann $x \ge \frac{1}{2}$.

Zusammenfassend ist damit $(-\infty, 0) \sqcup (0, \frac{1}{2}] \sqcup (1, \infty)$ die Lösungsmenge der obigen Ungleichung.

Nun kommen wir zurück zu den noch nicht genauer diskutierten Axiomen von $\mathbb{R}$, dem Archimedischen Axiom und dem Vollständigkeitsaxiom:

Bemerkungen IV.1.12.

(a) Das Archimedische Axiom sichert unter anderem die Möglichkeit der *Division mit Rest*: Für $x \in \mathbb{R}$, $q \in \mathbb{R}_{>0}$ existiert stets eine eindeutige Darstellung $x = sq+r$ mit $s \in \mathbb{Z}$, $r \in [0, q)$. Speziell für $q = 1$ ist $s \in \mathbb{Z}$ mit $x \in [s, s+1)$ der in Abschnitt II.4 erwähnte ganzzahlige Anteil, der mit der Gauß-Klammer als $\lfloor x \rfloor \in \mathbb{Z}$ notiert wird.

(Begründung für Existenz von r, s: Gemäß des Archimedischen Axioms ist $T_{x,q} := \{z \in \mathbb{Z} \mid z > q^{-1}x\}$ nicht leer und hat eine untere Schranke in $\mathbb{Z}$. Nach Abschnitt II.4.1 enthält dann $T_{x,q} \subset \mathbb{Z}$ eine kleinste Zahl t und für $s := t-1 \in \mathbb{Z}$ gilt $sq \le x < (s+1)q$, und für $r := x-sq \in [0, q)$ erhalten wir $x = sq+r$.

Begründung für Eindeutigkeit von r, s: Es sei $\widetilde{s}q+\widetilde{r} = sq+r$ mit $s, \widetilde{s} \in \mathbb{Z}$, $r, \widetilde{r} \in [0, q)$. Im Fall $\widetilde{s} > s$ ist $\widetilde{s} \ge s+1$, und es ergibt sich mit $q \le (\widetilde{s}-s)q = r-\widetilde{r} \le r < q$ ein Widerspruch. Im Fall $\widetilde{s} < s$ folgt dieser analog. Also muss $\widetilde{s} = s$ und dann auch $\widetilde{r} = r$ sein.)

(b) Aus dem Archimedischen Axiom folgt, dass zu jedem $\varepsilon \in \mathbb{R}_{>0}$ – egal, wie klein ε auch immer sein mag, solange es nur positiv ist – ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$ existiert.

(Begründung: Nach dem Axiom gibt es $n \in \mathbb{N}$ mit $n > \frac{1}{\varepsilon} > 0$. Durch Reziprokenbildung folgt $\frac{1}{n} < \varepsilon$.)

(c) Als weitere Folgerung ergibt sich, dass $\mathbb{R}$ dicht geordnet ist, das heißt, echt zwischen zwei verschiedenen reellen Zahlen liegen unendlich viele weitere reelle Zahlen. Tatsächlich finden im Zwischenbereich immer sowohl unendlich viele rationale Zahlen aus $\mathbb{Q}$ als auch unendlich viele irrationale Zahlen aus $\mathbb{R} \setminus \mathbb{Q}$, weshalb auch $\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ dicht geordnet sind. Mit etwas anderen Worten enthalten für $a, b \in \mathbb{R}$ mit $a < b$ sowohl $(a, b) \cap \mathbb{Q}$ als auch $(a, b) \setminus \mathbb{Q}$ stets unendlich viele Elemente.

(Begründung: Für beliebiges $k \in \mathbb{N}$ gibt es nach Bemerkung (b) stets ein $n \in \mathbb{N}$ mit $\frac{k}{n} < b-a$. Nach Bemerkung (a) können wir $a = \frac{s}{n}+r$ mit $s \in \mathbb{Z}$ und $r \in \big[0, \frac{1}{n}\big)$ schreiben. Damit liegen die k rationalen Zahlen $\frac{s+1}{n}, \frac{s+2}{n}, \ldots, \frac{s+k}{n}$ alle in $(a, b) \cap \mathbb{Q}$ (denn $\frac{s+1}{n} = \frac{s}{n}+\frac{1}{n} > \frac{s}{n}+r = a$ und $\frac{s+k}{n} = \frac{s}{n}+\frac{k}{n} < \frac{s}{n}+r+b-a = b$). Da k beliebig war, enthält $(a, b) \cap \mathbb{Q}$ also unendlich viele Zahlen.

Um dasselbe für $(a, b) \setminus \mathbb{Q}$ zu zeigen, setzen wir voraus, dass es überhaupt eine irrationale Zahl $t \in \mathbb{R}$ gibt (Beispiele folgen in Mathematik 2), und wir können dann auch $t > 0$ annehmen. Wir argumentieren

nun wie zuvor, wählen zu $k \in \mathbb{N}$ ein $n \in \mathbb{N}$ mit $\frac{k}{n}t < b-a$, schreiben $a = \frac{s}{n}t+r$ mit $s \in \mathbb{Z}$, $r \in \left[0, \frac{1}{n}t\right)$ und erhalten die k irrationalen Zahlen $\frac{s+1}{n}t$, $\frac{s+2}{n}t$, …, $\frac{s+k}{n}t$ in $(a,b) \setminus \mathbb{Q}$.)

Definition IV.1.13. Eine *Intervallschachtelung* in $\mathbb{R}$ ist eine unendliche Folge nicht-leerer kompakter Intervalle

$$[a_1, b_1] \supset [a_2, b_2] \supset [a_3, b_3] \supset [a_4, b_4] \supset \dots$$

mit Randpunkten $-\infty < a_1 \leq a_2 \leq a_3 \leq a_4 \leq \dots \leq b_4 \leq b_3 \leq b_2 \leq b_1 < \infty$, so dass zu jedem $\varepsilon \in \mathbb{R}_{>0}$ ein $n \in \mathbb{N}$ mit $b_n - a_n < \varepsilon$ existiert Die letzte Bedingung kann man in Worten so ausdrücken, dass die Intervall-Längen $b_n - a_n$ mit wachsendem $n \in \mathbb{N}$ beliebig klein werden.

Als *Kern einer* solchen *Intervallschachtelung* bezeichnet man eine Zahl $c \in \mathbb{R}$ mit $c \in [a_n, b_n]$ für alle $n \in \mathbb{N}$, also $c \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$.

Lemma IV.1.14. *Der Kern einer Intervallschachtelung ist eindeutig.*

Beweis. Sind $c, \widetilde{c} \in \mathbb{R}$ zwei Kerne, so gilt mit $c, \widetilde{c} \in [a_n, b_n]$ auch $|\widetilde{c}-c| \leq b_n - a_n$ für alle $n \in \mathbb{N}$. Wäre $|\widetilde{c}-c| > 0$, so gäbe es ein $n \in \mathbb{N}$ mit $b_n - a_n < |\widetilde{c}-c|$, und wir erhielten einen Widerspruch. Also muss $|\widetilde{c}-c| = 0$ gelten, was $\widetilde{c} = c$ bedeutet. □

Bemerkungen IV.1.15.

- Erst mit der gerade gegebenen Definition wird das Axiom der metrischen Vollständigkeit

  „Jede Intervallschachtelung in $\mathbb{R}$ besitzt einen Kern in $\mathbb{R}$."

  sinnvoll. Die anschauliche Interpretation des Axioms ist, dass die reellen Zahlen $\mathbb{R}$, anders als die rationalen Zahlen $\mathbb{Q}$, die gesamte Zahlengerade füllen und keine „Lücken" mehr lassen.

  Konkret ist zum Beispiel $\sqrt{2} = 1,4142135623\dots$ in $\mathbb{R}$ der Kern der Intervallschachtelung

  $$\left[\frac{14}{10}, \frac{15}{10}\right] \supset \left[\frac{141}{100}, \frac{142}{100}\right] \supset \left[\frac{1414}{1000}, \frac{1415}{1000}\right] \supset \left[\frac{14142}{10000}, \frac{14143}{10000}\right] \supset \dots,$$

  wobei die Randpunkte allgemein als

  $$a_n := 10^{-n}\lfloor 10^n\sqrt{2}\rfloor \in \mathbb{Q} \text{ und } b_n := a_n + 10^{-n} \in \mathbb{Q}$$

  geschrieben werden können. Da aber $\sqrt{2} \notin \mathbb{Q}$ irrational ist (dazu in Mathematik 2 nochmal genauer), kommt $\sqrt{2}$ als Kern in $\mathbb{Q}$ nicht in Frage, und in $\mathbb{Q}$ besitzt diese Intervallschachtelung eben keinen Kern.
- Mit einem erst später eingeführten Konzept lässt sich metrische Vollständigkeit von $\mathbb{R}$ so charakterisieren: Jede Cauchy-Folge in $\mathbb{R}$ konvergiert in $\mathbb{R}$. Für $\mathbb{C}$ geht dies auch.

Weitere Begriffe, die mit der Anordnung von $\mathbb{R}$ zusammenhängen, sind:

Definition IV.1.16. Es sei A eine Teilmenge von $\overline{\mathbb{R}}$.

(a) Wir nennen A *von oben beschränkt* beziehungsweise *von unten beschränkt*, wenn A eine obere Schranke beziehungsweise untere Schranke in $\mathbb{R}$ besitzt. Ist A von oben und unten beschränkt, so heißt A *beschränkt*.

(b) Ein größtes Element bzw. kleinstes Element[1] von A nennt man auch *Maximum* bzw. *Minimum* von A und schreibt für dieses $\max A = \max_{x\in A} x$ bzw. $\min A = \min_{x\in A} x$.

(c) Eine *kleinste obere Schranke* bzw. *größte untere Schranke* für A in $\overline{\mathbb{R}}$ bezeichnet man als *Supremum* bzw. *Infimum* von A und notiert diese als $\sup A = \sup_{x\in A} x$ bzw. als $\inf A = \inf_{x\in A} x$.

Damit können wir eine etwas andere Vollständigkeitseigenschaft von $\mathbb{R}$ formulieren:

Satz IV.1.17. *Jede nicht-leere, von oben beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum in $\mathbb{R}$.*

Bemerkungen IV.1.18.

(a) Als Folgerung besitzt *sogar jede* Teilmenge $A \subset \overline{\mathbb{R}}$ ein Supremum in $\overline{\mathbb{R}}$, wobei $\sup A = \infty$ genau für von oben unbeschränktes A und $\sup A = -\infty$ genau für $A = \emptyset$ und für $A = \{-\infty\}$ eintritt.

[1] Da $\mathbb{R}$ total geordnet ist, fallen größte bzw. kleinste Elemente hier mit maximalen bzw. minimalen Elementen zusammen; vergleiche mit Abschnitt II.4.2.

(b) Für $A \subset \overline{\mathbb{R}}$ und $M \in \overline{\mathbb{R}}$ ergibt sich folgende *Charakterisierung des Supremums*:

$$M = \sup A \iff \Big(\underbrace{\forall x \in A\colon x \le M}_{M \text{ obere Schranke für } A} \quad \text{und } \forall L \in (-\infty, M)\colon \quad \underbrace{\exists y \in A\colon y > L}_{L \textit{ keine} \text{ obere Schranke für } A} \Big)$$

Diese Charakterisierung wird in Anwendungen des Öfteren benutzt.

(c) Für $A \subset \overline{\mathbb{R}}$ existiert $\max A$ genau dann, wenn $\sup A \in A$ gilt, und in diesem Fall ist $\sup A = \max A \in A$. Auch $\sup A \notin A$ kommt aber natürlich vor.

(d) Analoges gilt für das Infimum (und das Minimum).

(e) Aus dem Axiom folgt die Behauptung, dass die speziellen Intervalle aus Teil (d) der Intervalldefinition *alle* Intervalle I in $\overline{\mathbb{R}}$ sind.

Um dies einzusehen, nimmt man $I \neq \emptyset$ an, erklärt $a := \inf I \in [-\infty, \infty)$, $b := \sup I \in (-\infty, \infty]$ und überlegt sich dann, dass $I \in \{(a,b), (a,b], [a,b), [a,b]\}$ gilt.

(f) Mit erst später eingeführten Konzepten lässt sich Ordnungs-Vollständigkeit von $\mathbb{R}$ so charakterisieren: Jede beschränkte, monotone Folge in $\mathbb{R}$ konvergiert in $\mathbb{R}$.

Beweis des Satzes. Sei A nicht-leere, von oben beschränkte Teilmenge von $\mathbb{R}$. Für fixiertes $n \in \mathbb{N}$ betrachten wir obere Schranken für A der Form[2] $\frac{z}{2^n}$ mit $z \in \mathbb{Z}$. Wegen der Beschränktheitsvoraussetzung besitzt A eine obere Schranke, gemäß des Archimedischen Axioms dann auch eine obere Schranke der gegebenen Form und wegen der Wohlordnungseigenschaften von $\mathbb{Z}$ schließlich eine kleinste Schranke b_n der gegebenen Form. Setzen wir $a_n := b_n - \frac{1}{2^n}$, so ist insgesamt durch die Intervalle $[a_n, b_n]$ mit $n \in \mathbb{N}$ eine Intervallschachtelung gegeben, wobei $b_n - a_n = \frac{1}{2^n} \le \frac{1}{n}$ auch wieder wegen des Archimedischen Axioms für ein $n \in \mathbb{N}$ kleiner als jedes gegebene $\varepsilon \in \mathbb{R}_{>0}$ wird. Nach dem Axiom der metrischen Vollständigkeit besitzt die Intervallschachtelung einen Kern $M \in \mathbb{R}$. Jedes $x \in A$ erfüllt nun $x \le b_n \le M + \frac{1}{n}$ für alle $n \in \mathbb{N}$, also gilt auch $x \le M$, und M ist eine obere Schranke für A. Außerdem gibt es zu jedem $n \in \mathbb{N}$ ein $y \in A$ mit $y > a_n$ und folglich $y > M - \frac{1}{n}$, so dass M das Supremum von A sein muss. □

Bemerkung IV.1.19. Tatsächlich ist Ordnungs-Vollständigkeit von $\mathbb{R}$ unter Voraussetzung der Körper- und Anordnungsaxiome sogar äquivalent zur metrischen Vollständigkeit von $\mathbb{R}$ zusammen mit dem Archimedischen Axiom. Daher besteht ein äquivalentes Axiomensystem für $\mathbb{R}$ aus den Körperaxiomen, den Anordnungsaxiomen und der Ordnungs-Vollständigkeit.

Dass metrische Vollständigkeit und das Archimedische Axiome Ordnungs-Vollständigkeit implizieren, zeigt der vorausgehende Beweis.

Für den Umkehrschluss sei Ordnungs-Vollständigkeit von $\mathbb{R}$ vorausgesetzt. Ist $([a_n, b_n])_{n\in\mathbb{N}}$ eine Intervallschachtelung in $\mathbb{R}$, so existiert $c := \sup\{a_n \mid n \in \mathbb{N}\} \in \mathbb{R}$, und für alle $n \in \mathbb{N}$ gilt $a_n \le c$ und $c \le b_n$, weil b_n eine obere Schranke und c die *kleinste* obere Schranke für $\{a_n \mid n \in \mathbb{N}\}$ ist. Damit ist c der Kern der Intervallschachtelung, und die metrische Vollständigkeit von $\mathbb{R}$ ist nachgewiesen. Das Archimedische Axiome erhält man durch ein Widerspruchsargument. Wäre das Axiom nicht erfüllt, so gäbe es ein $x \in \mathbb{R}$ mit $n \le x$ für alle $n \in \mathbb{N}$. Damit wäre $\mathbb{N}$ von oben beschränkt und wir setzen $M := \sup \mathbb{N} \in \mathbb{R}$. Dies wäre die kleinste obere Schranke und $M-1$ wäre *keine* obere Schranke für $\mathbb{N}$. Es gäbe ein $n_0 \in \mathbb{N}$ mit $n_0 > M-1$, und $\mathbb{N} \ni n_0+1 > M$ stünde im Widerspruch zur Wahl von M als obere Schranke für $\mathbb{N}$. Es folgt die Gültigkeit des Archimedischen Axioms.

Um die Mathematik — wie in Abschnitt I.5 angekündigt — einzig auf das Zermelo-Fraenkel-Axiomensystem der Mengenlehre zu gründen, bleibt aber dennoch eine Konstruktion des Zahlbereichs $\mathbb{R}$ auf Grundlage bereits eingeführter Bildungen anzugeben. Hierzu gibt es mehrere Möglichkeiten. Eine recht elementare Vorgehensweise geht vom bereits eingeführten Zahlbereich $\mathbb{Q}$ aus und basiert auf der Verwendung sogenannter *Dedekindscher Schnitte*. Die Grundidee ist dabei, eine reelle Zahl $x \in \mathbb{R}$ mit der Teilmenge $\mathbb{Q}_{<x} := \{a \in \mathbb{Q} \mid a < x\}$ von $\mathbb{Q}$ zu identifizieren. Mit der Teilmenge verbindet man gedanklich die Zerlegung $\mathbb{Q} = \mathbb{Q}_{<x} \sqcup \mathbb{Q}_{\ge x}$, an der man insbesondere die „Schnittstelle“ x ablesen kann. Als formale Konstruktion führt man $\mathbb{R}$ daher als Menge von Teilmengen von $\mathbb{Q}$ ein, die die charakteristischen Eigenschaften der gerade besprochenen Mengen

[2] Mit Schranken der einfacheren Form $\frac{z}{n}$ können wir hier nicht (ohne Weiteres) arbeiten, da sich nicht immer eine Intervallschachtelung ergäbe. Zum Beispiel für $A = \{\frac{2}{5}\}$ bekämen wir nämlich $[a_2, b_2] = [0, \frac{1}{2}] \not\supset [a_3, b_3] = [\frac{1}{3}, \frac{2}{3}]$.

$\mathbb{Q}_{<x}$ aufweisen: Tatsächlich setzt man

$$\mathbb{R} := \left\{ A \in \mathcal{P}(\mathbb{Q}) \;\middle|\; \begin{array}{c} \emptyset \neq A \neq \mathbb{Q} \\ \forall a \in A : \forall b \in \mathbb{Q} \setminus A : a < b \\ \text{Es gibt keine größte Zahl in } A. \end{array} \right\}$$

und versteht dann $\mathbb{Q} \subset \mathbb{R}$, indem man $q \in \mathbb{Q}$ mit $\mathbb{Q}_{<q} \in \mathbb{R}$ identifiziert. Auf dem so definierten Bereich der reellen Zahlen erhält man nun die Kleiner-Relation $<$ als die strikte Mengeninklusion $\subsetneq$ und die Addition $+$ als die Minkowski-Addition $+$. Die Multiplikation $\cdot$ von $A, B \in \mathbb{R}$ kann man im Fall $A, B \geq 0$ durch $A \cdot B := (A_{>0} \cdot B_{>0}) \sqcup \mathbb{Q}_{\leq 0}$ erklären, wobei $\cdot$ auf der rechten Seite für die auf Mengen erweiterte Multiplikation (vgl. Abschnitt III.1) steht und wir $A_{>0} := \{a \in A \mid a > 0\}$ abgekürzt haben. Die Multiplikation mit negativen reellen Zahlen lässt sich (im Fortgang der Argumentation mit Hilfe des additiv Inversen) darauf zurückführen. Ausgehend von diesen Definitionen gilt es nun, die zuvor angegebenen Axiome von $\mathbb{R}$ zu verifizieren, was etwas Aufwand erfordert und hier nicht weiter besprochen wird. Neben der Zurückführung auf die Mengenlehre erreicht man mit dieser Konstruktion übrigens auch den Nachweis, dass das für $\mathbb{R}$ angegebene Axiomensystem widerspruchsfrei ist – jedenfalls, sofern das Zermelo-Fraenkel-Axiomensystem dies ist.

Als weiterführende Zusatz-Information sei noch erwähnt, dass sich in Anlehnung an die Konstruktion mittels Dedekindscher Schnitte auch folgende Eindeutigkeitseigenschaft der reellen Zahlen nachweisen lässt:

Satz IV.1.20. *Erfüllen zwei Mengen $\mathbb{R}$ (mit $+, \cdot, <$) und $\widetilde{\mathbb{R}}$ (mit $\widetilde{+}, \widetilde{\cdot}, \widetilde{<}$) alle Axiome der reellen Zahlen, so gibt es einen Körperisomorphismus $\psi\colon \mathbb{R} \to \widetilde{\mathbb{R}}$, der zudem im Sinn von*

$$x < y \implies \psi(x) \widetilde{<} \psi(y)$$

für alle $x, y \in \mathbb{R}$ ordnungerhaltend ist.

Beweisskizze. Man betrachtet zunächst die Primkörper $\mathbb{Q}$ und $\widetilde{\mathbb{Q}}$ von $\mathbb{R}$ und $\widetilde{\mathbb{R}}$, die beide den rationalen Zahlen entsprechen und somit durch einen Körperisomorphismus $f\colon \mathbb{Q} \to \widetilde{\mathbb{Q}}$ identifiziert werden können. Mit den Anordnungsaxiomen kann gezeigt werden, dass f (wie auch f^{-1}) ordnungerhaltend ist. Insbesondere ist daher für jedes $x \in \mathbb{R}$ das Bild $f(\mathbb{Q}_{<x}) \subset \widetilde{\mathbb{R}}$ der nicht-leeren, von oben beschränkten Menge $\mathbb{Q}_{<x} \subset \mathbb{R}$ ebenfalls nicht-leer und von oben beschränkt. Gemäß der Ordnungs-Vollständigkeit von $\widetilde{\mathbb{R}}$ (deren Äquivalenz zur metrischen Vollständigkeit und dem Archimedischen Axiom ja schon diskutiert wurde) existiert dann für alle $x \in \mathbb{R}$ das Supremum $\widetilde{\sup} f(\mathbb{Q}_{<x})$ in $\widetilde{\mathbb{R}}$, und es lässt sich eine Abbildung $\psi\colon \mathbb{R} \to \widetilde{\mathbb{R}}$ durch $\psi(x) := \widetilde{\sup} f(\mathbb{Q}_{<x})$ für $x \in \mathbb{R}$ definieren.

Die benötigten Eigenschaften von ψ verifiziert man in mehreren Schritten.

Um zunächst einzusehen, dass ψ ein Körperhomomorphismus ist, kann man für $x, y \in \mathbb{R}$ von $\mathbb{Q}_{<x+y} = \mathbb{Q}_{<x} + \mathbb{Q}_{<y}$ ausgehen und erhält

$$\psi(x+y) = \widetilde{\sup} f(\mathbb{Q}_{<x} + \mathbb{Q}_{<y}) = \widetilde{\sup}[f(\mathbb{Q}_{<x}) \widetilde{+} f(\mathbb{Q}_{<y})] = \widetilde{\sup} f(\mathbb{Q}_{<x}) \,\widetilde{+}\, \widetilde{\sup} f(\mathbb{Q}_{<y}) = \psi(x) \widetilde{+} \psi(y).$$

Für die Multiplikation lässt sich ähnlich (aber mit Fallunterscheidung nach den Vorzeichen von x und y) argumentieren.

Dass ψ ordnungerhaltend und damit insbesondere injektiv ist, erhält man wie folgt: Für $x, y \in \mathbb{R}$ mit $x < y$ gibt es wegen der dichten Anordnung von $\mathbb{Q}$ in $\mathbb{R}$ rationale Zwischenstellen $p, q \in \mathbb{Q}$ mit $x < p < q < y$. Aus der Definition des Supremums und der Ordnungerhaltung unter f ergibt sich dann

$$\widetilde{\sup} f(\mathbb{Q}_{<x}) \widetilde{\leq} f(p) \widetilde{<} f(q) \widetilde{\leq} \widetilde{\sup} f(\mathbb{Q}_{<y}),$$

also wie benötigt $\psi(y) \widetilde{<} \psi(y)$.

Um den Beweis der Surjektivität von ψ vorzubereiten, weist man zunächst $\psi(q) = f(q)$ für $q \in \mathbb{Q}$ nach, mit anderen Worten also $f(q) = \widetilde{\sup} f(\mathbb{Q}_{<q})$. Dass es sich bei $f(q)$ um eine obere Schranke für $f(\mathbb{Q}_{<q})$ handelt, folgt, weil f ordnungerhaltend ist. Gäbe es noch eine kleinere obere Schranke für $f(\mathbb{Q}_{<q})$, so könnte diese wegen der dichten Anordnung als $\eta \in \widetilde{\mathbb{Q}}$ mit $\eta \widetilde{<} f(q)$ gewählt werden. Es ergäbe sich der Widerspruch, dass einerseits $f^{-1}(\eta) < q$, aber andererseits $f^{-1}(\eta)$ obere Schranke für $\mathbb{Q}_{<q}$ wäre. Also ist $f(q) = \widetilde{\sup} f(\mathbb{Q}_{<q})$.

Als weitere Vorbereitung zeigt man mit ähnlicher Argumentation $\psi(\sup A) = \widetilde{\sup} f(A)$ für nicht-leere, von oben beschränkte $A \subset \mathbb{Q}$. Direkt sieht man wieder, dass $\psi(\sup A)$ obere Schranke für $f(A) = \psi(A)$ ist. Gäbe es eine kleinere obere Schranke, so könnte diese als $\eta \in \widetilde{\mathbb{Q}}$ mit $\eta \widetilde{<} \psi(\sup A)$ gewählt werden. Nun wäre

$f^{-1}(\eta) < \sup A$ (denn mit $f^{-1}(\eta) \geq \sup A$ müsste auch $\eta = f(f^{-1}(\eta)) = \psi(f^{-1}(\eta)) \geq \psi(\sup A)$ gelten) und $f^{-1}(\eta)$ wäre eine kleinere obere Schranke für A als $\sup A$. In Anbetracht dieses Widerspruchs gilt wie behauptet $\psi(\sup A) = \widetilde{\sup} f(A)$.

Schließlich lässt sich die Surjektivität von ψ beweisen. Sei dazu $\alpha \in \widetilde{\mathbb{R}}$ beliebig. Die schon bei der Definition von ψ angestellten Überlegungen liefern dann die Existenz von $x := \sup f^{-1}(\widetilde{\mathbb{Q}}_{<\alpha}) \in \mathbb{R}$, und mit den zuletzt nachgewiesenen Eigenschaften folgt $\psi(x) = \psi(\sup f^{-1}(\widetilde{\mathbb{Q}}_{<\alpha})) = \sup \psi(f^{-1}(\widetilde{\mathbb{Q}}_{<\alpha})) = \sup f(f^{-1}(\widetilde{\mathbb{Q}}_{<\alpha})) = \sup \widetilde{\mathbb{Q}}_{<\alpha} = \alpha$.

Insgesamt ist $\psi\colon \mathbb{R} \to \widetilde{\mathbb{R}}$ ordnungerhaltender Körperisomorphismus. □

IV.2. Der Körper der komplexen Zahlen

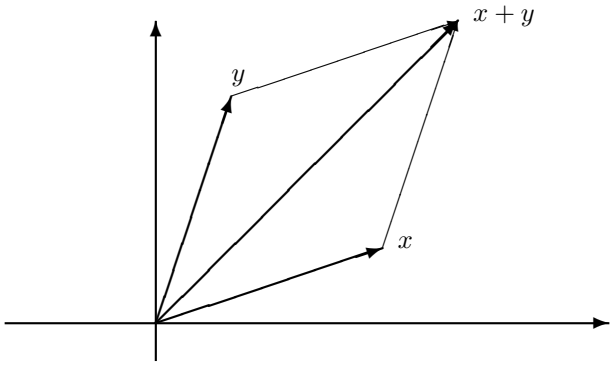
Elemente in $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ schreiben wir oft als Vektoren, also

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, x_1, x_2 \in \mathbb{R} \right\}.$$

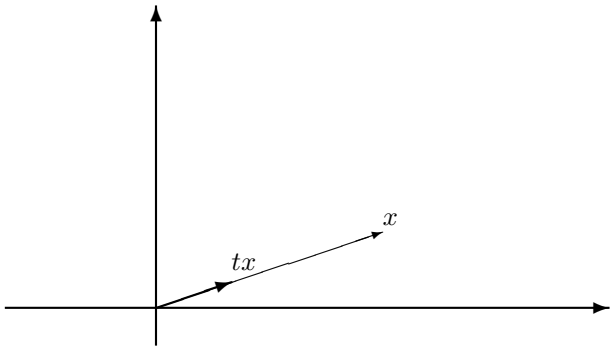
Definition IV.2.1.

(a) Die *Summe* von $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ist

$$x + y := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}.$$



(b) Die *Streckung von* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ *um den Wert* $t \in \mathbb{R}$ ist $tx := \begin{pmatrix} tx_1 \\ tx_2 \end{pmatrix}$.



(c) Die *Länge von* x oder auch die *euklidische Norm von* x ist

$$||x|| := \sqrt{x_1^2 + x_2^2}.$$

Bemerkung IV.2.2. Für alle $x, y, z \in \mathbb{R}^2$ und alle $s, t \in \mathbb{R}$ gilt:

(a) $(x + y) + z = x + (y + z)$.

(b) Mit $0 := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ist $0 + x = x = x + 0$.

(c) Für alle $x \in \mathbb{R}^2$ gibt es ein $-x \in \mathbb{R}^2$ mit $x + (-x) = 0 = (-x) + x$, nämlich

$$-x = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}.$$

(d) $x + y = y + x$.

(e) $s(tx) = (st)x$.

(f) $1x = x$.

(g) $t(x + y) = tx + ty$.

(h) $(s + t)x = sx + tx$.

Für den $\mathbb{R}^2$ sind diese Rechenregeln einfach nachzurechnen. Tun Sie das bitte. Analoge Aussagen gelten für den $\mathbb{R}^n$ für alle $n \geqslant 1$. Wir werden später genau diese Rechenregeln benutzen, um allgemein zu sagen, was ein Vektorraum sein soll. Insbesondere ist der $\mathbb{R}^2$ mit der Addition eine abelsche Gruppe.

Wir versehen die reelle Ebene mit einer Körperstruktur.

Wir definieren eine Multiplikation auf $\mathbb{R}^2$: Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ definieren wir

$$x \cdot y := \begin{pmatrix} x_1 y_1 - x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix}. \tag{IV.2.1}$$

Damit ist $(\mathbb{R}^2 \setminus \{0\}, \cdot)$ eine abelsche Gruppe:

(a) Der Nachweis der Assoziativität ist eine mühselige Rechnung, die Sie selbst durchführen. Kommutativität ist einfach zu sehen.

(b) Das Element $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist das neutrale Element:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \cdot y_1 - 0 \cdot y_2 \\ 1 \cdot y_2 + 0 \cdot y_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

(c) Ist $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, so ist

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{-1} = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} = \frac{1}{||x||^2} \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

das multiplikative Inverse:

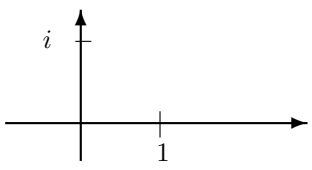
$$\frac{1}{x_1^2 + x_2^2} \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} x_1^2 - ((-x_2)x_2) \\ x_1 x_2 - x_1 x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(d) Dass die Distributivgesetze gelten, folgt mit einer Rechnung, die wir hier unterschlagen.

Wir schreiben $\mathbb{C}$ für $\mathbb{R}^2$ mit dieser Körperstruktur; $\mathbb{C}$ heißt auch *komplexe Zahlenebene*.

Definition IV.2.3.

(a) Das Element $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{C}$ heißt *imaginäre Einheit* und wird mit i bezeichnet.



(b) Ist $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}$, so heißt x_1 der *Realteil von* x und x_2 der *Imaginärteil von* x. Wir kürzen dies ab mit $\mathsf{Re}(x) = x_1$ und $\mathsf{Im}(x) = x_2$.

Bemerkung IV.2.4.

- Es gilt $i^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, also erfüllt i die Gleichung
$$i^2 = -1.$$
Wir können also i als eine Quadratwurzel aus -1 auffassen, aber Vorsicht: Was ist falsch an der Gleichungskette
$$-1 = i^2 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)(-1)} = \sqrt{1} = 1?$$
- Wir können jedes $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}$ schreiben als
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1 \cdot 1 + x_2 \cdot i$$
oder kurz $x_1 + x_2 i$.
- Mit dieser Umformulierung sieht die Multiplikation in $\mathbb{C}$ schon freundlicher aus: Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ist

$x \cdot y = (x_1 + x_2 i)(y_1 + y_2 i) = x_1 y_1 + x_1 y_2 i + x_2 y_1 i + x_2 y_2 i^2 = (x_1 y_1 - x_2 y_2) + (x_1 y_2 + x_2 y_1) i.$

Wir brauchen also nur die Rechenregeln in $\mathbb{R}$, $i^2 = -1$ und die Distributivgesetze.

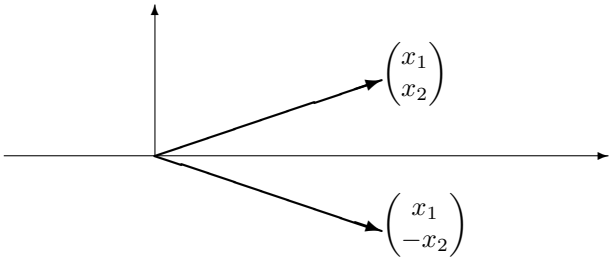
Definition IV.2.5.

(a) Die Abbildung
$$^{-}\colon \mathbb{C} \to \mathbb{C}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \overline{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$
heißt *komplexe Konjugation*. Wir schreiben auch $\overline{x_1 + x_2 i} = x_1 - x_2 i$.

(b) Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}$ heißt
$$|x| = \sqrt{x_1^2 + x_2^2}$$
der Absolutbetrag von x. Dies ist nichts anderes als die euklidische Norm des Vektors $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.

Bemerkung IV.2.6.

(a) Geometrisch ist die komplexe Konjugation die Spiegelung an der x_1-Achse



(b) Traditionell werden Elemente in $\mathbb{C}$ oft mit $z = x + iy$ notiert. Wir werden diese Schreibweise ab jetzt benutzen.

Satz IV.2.7. *Für alle $z, w \in \mathbb{C}$ gilt:*

(a) $\overline{\overline{z}} = z$,
(b) $\overline{z+w} = \overline{z} + \overline{w}$,
(c) $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$,
(d) $\overline{0} = 0, \overline{1} = 1, \overline{i} = -i$,
(e) $|z+w| \leqslant |z| + |w|$,
(f) $|z \cdot w| = |z||w|$,
(g) $|z|^2 = z \cdot \overline{z}$. *Insbesondere gilt für alle* $z \neq 0$, *dass*

$$z^{-1} = \frac{\overline{z}}{|z|^2}.$$

Beweis. Die Beweise der obigen Behauptungen bestehen aus elementaren Rechnungen. Wir führen exemplarisch (b) und (g) aus. Eine anschauliche Begründung für (e) ist, dass $|z+w|$ der Länge der Diagonale des Parallelograms entspricht, welches von z und w aufgespannt wird. Diese Länge ist kleiner gleich der Summe der Längen der aufspannenden Seiten.

Zu (b): Wir schreiben $z = x + iy$ und $w = u + iv$ und erhalten

$$\begin{aligned}\overline{z+w} =& \overline{x + iy + u + iv} \\ =& \overline{(x+u) + i(y+v)} \\ =& (x+u) - i(y+v) \\ =& x - iy + u - iv \\ =& \overline{z} + \overline{w}.\end{aligned}$$

Zu (g): Auch hier rechnen wir explizit nach:

$$z \cdot \overline{z} = (x+iy) \cdot (x-iy) = x^2 + iyx - ixy - i^2y^2 = x^2 + y^2 = |z|^2.$$

□

Wir benutzen komplexe Zahlen, um eine allgemeiner Lösung quadratischer Gleichungen mit reellen Koeffizienten anzugeben. Gesucht sind die Nullstellen von

$$ax^2 + bx + c \tag{IV.2.2}$$

also Elemente x_0 mit

$$ax_0^2 + bx_0 + c = 0.$$

Hierbei seien $a, b, c \in \mathbb{R}$. Ist $a = 0$, so suchen wir x_0 mit $bx_0 + c = 0$. Ist b auch null und verschwindet c nicht, so gibt es keine Lösung. Für $b \neq 0$ erhalten wir eine Lösung und zwar

$$x_0 = \frac{-c}{b}.$$

Für $a \neq 0$ sind die Nullstellen von (IV.2.2) gleich denen der Gleichung

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0,$$

so dass wir im Folgenden ohne Beschränkung der Allgemeinheit annehmen, dass der Höchstkoeffizient 1 ist. Gesucht sind also Nullstellen von

$$x^2 + \alpha x + \beta$$

mit $\alpha, \beta \in \mathbb{R}$. Wir formen um:

$$
\begin{aligned}
& x^2 + \alpha x + \beta = 0 \\
\Leftrightarrow\, & x^2 + \alpha x = -\beta \\
\Leftrightarrow\, & x^2 + 2 \cdot \frac{\alpha}{2} x + \frac{\alpha^2}{4} = \frac{\alpha^2}{4} - \beta \\
\Leftrightarrow\, & \left(x - \frac{\alpha}{2}\right)^2 = \frac{\alpha^2}{4} - \beta.
\end{aligned}
$$

Ist $\frac{\alpha^2}{4} - \beta > 0$, so können wir die Quadratwurzel ziehen und erhalten die beiden reellen Lösungen

$$x_{1/2} = -\frac{\alpha}{2} \pm \sqrt{\frac{\alpha^2}{4} - \beta}.$$

Ist $\frac{\alpha^2}{4} - \beta = 0$, so ist $-\frac{\alpha}{2}$ eine doppelte Nullstelle.

Für $\frac{\alpha^2}{4} - \beta < 0$ schreiben wir dies um zu $\frac{\alpha^2}{4} - \beta = -d$ mit $d > 0$ und erhalten zwei komplexe Lösungen

$$z_{1/2} = -\frac{\alpha}{2} \pm \sqrt{d} \cdot i.$$

Beachten Sie, dass die Lösungen zueinander komplex konjugiert sind: $\overline{z_1} = z_2$.

Zum Beispiel erhalten wir für $x^2 - 4x + 5$ die beiden Lösungen

$$z_{1/2} = 2 \pm \sqrt{4 - 5} = 2 \pm i.$$

KAPITEL V

Anhang

Die Themen im Anhang sind nicht prüfungsrelevant, aber vielleicht finden Sie diese ja trotzdem interessant.

V.1. Zermelo-Fraenkel Axiome

Die Zermelo-Fraenkel-Axiome der Mengenlehre fundieren eine Mathematik, in der alle mathematischen Objekte Mengen sind und insbesondere auch als Elemente von Mengen nur Mengen in Frage kommen. Die Axiome basieren auf einer Prädikatenlogik erster Stufe mit Gleichheitsrelation $=$, wobei alle Quantoren über das sogenannte Diskursuniversum aus allen Mengen laufen. Daneben wird nur auf die undefinierte Elementrelation $\in$ aufgebaut, auf deren Grundlage $\notin$ und $\subset$ wie früher in diesem Abschnitt definiert werden. Die genauen Axiome lauten dann wie folgt, wobei die zweiten eingeklammerten Absätze nicht zum Axiomensystem gehören, sondern Zusatzerläuterungen geben:

- *Extensionalitätsaxiom:* Zwei Mengen M und N sind genau dann gleich, wenn sie dieselben Elemente enthalten. Mit anderen Worten ist die Gleichheit $M = N$ gleichbedeutend mit $\forall x\colon (x \in M \iff x \in N)$.

  (Wir hatten dies zuvor als Definition der Mengen-Gleichheit betrachtet. Da die Gleichheit aber schon in der zugrunde gelegten Prädikatenlogik definiert ist, kann man sich diese Charakterisierung der Gleichheit genau genommen nur als Axiom und nicht als Definition wünschen.)
- *Leermengenaxiom:* Es existiert eine Menge $\emptyset$ ohne Elemente, die also $\forall x\colon x \notin \emptyset$ erfüllt.

  (Wegen des Extensionalitätsaxioms ist $\emptyset$ dabei eindeutig bestimmt. Das Leermengenaxiom ist das einzige Axiom, das die Existenz einer konkreten Menge fordert. Alle anderen Mengen können — was zunächst erstaunlich scheinen mag — mit Hilfe der anderen Axiome aus der leeren Menge gebildet werden.)
- *Paarmengenaxiom:* Für alle Mengen a und b existiert eine Menge $\{a, b\}$, die genau a und b als Elemente enthält, für die also $\forall x\colon (x \in \{a, b\} \iff (x = a) \vee (x = b))$ gilt.

  (Man kann die Menge $\{a, b\}$ als ein *ungeordnetes* Paar von a und b auffassen, denn die Gleichheit $\{a, b\} = \{\tilde{a}, \tilde{b}\}$ solcher Mengen bedeutet $(a = \tilde{a}) \wedge (b = \tilde{b})$ $(a = \tilde{b}) \wedge (b = \tilde{a})$. Hierauf aufbauend kann man das *geordnete* Paar (a, b) und das kartesische Produkt mengentheoretisch wie oben durch $(a, b) := \{\{a\}, \{a, b\}\}$ definieren und weist anhand dieser Definition nach, dass die Gleichheit geordneter Paare $(a, b) = (\tilde{a}, \tilde{b})$ tatsächlich wie gewünscht $(a = \tilde{a}) \wedge (b = \tilde{b})$ bedeutet.)
- *Vereinigungsaxiom:* Für jede Menge $\mathcal{S}$ (von Mengen) existiert eine mit $\bigcup \mathcal{S}$ bezeichnete Menge, deren Elemente genau die Elemente der Elemente von $\mathcal{S}$ sind, für die mit anderen Worten also $\forall x\colon (x \in \bigcup \mathcal{S} \iff \exists M \in \mathcal{S}\colon x \in M)$ gilt.
- *Unendlichkeitsaxiom:* Es existiert eine sogenannte induktive Menge, die zum einen die leere Menge $\emptyset$ als Element enthält und zum anderen für jedes ihrer Elemente x auch $x \cup \{x\}$ als Element enthält (wobei $\{x\} := \{x, x\}$ und $x \cup \{x\}$ gemäß den vorigen Axiomen existieren).

  (Dieses Axiom sichert die Existenz unendlicher Mengen und spielt vor allem bei der Konstruktion der natürlichen Zahlen $\mathbb{N}$, der wohl grundlegendsten unendlichen Menge, eine entscheidende Rolle.)
- *Potenzmengenaxiom:* Zu jeder Menge M existiert die sogenannte Potenzmenge $\mathcal{P}(M)$, die genau die Teilmengen von M als Elemente enthält, also $\forall T\colon (T \subset M \iff T \in \mathcal{P}(M))$ erfüllt.

- *Fundierungsaxiom/Regularitätsaxiom:* Jede Menge M außer der leeren Menge enthält ein Element N mit $M \cap N = \emptyset$ (wobei man vor Einführung der Schnittmenge statt $M \cap N = \emptyset$ eigentlich $\forall x\colon (x \notin M \vee x \notin N)$ schreiben muss).

  (Dieses Axiom sichert, dass keine unendliche Kette von ineinander enthalten Mengen des Typs $M_1 \ni M_2 \ni M_3 \ni \dots$ existieren kann (denn dann würde die unendliche Menge $M = \{M_1, M_2, M_3, \dots\}$ dem Axiom widersprechen). Das Axiom führt auch dazu, dass $N \notin N$ für jede Menge N gilt (denn bei Existenz von N mit $N \in N$ würde $M = \{N\}$ dem Axiom widersprechen).)
- *Aussonderungsaxiom:* Für jede Menge M und jedes Prädikat $P(y)$ mit einer freien Variable y gibt es eine Menge $\{y \in M \mid P(y)\}$, deren Elemente genau die Elemente x von M sind, für die $P(x)$ gilt. Diese Menge $\{y \in M \mid P(y)\}$ erfüllt also $\forall x\colon (x \in \{y \in M \mid P(y)\} \iff x \in M \wedge P(x)$ ist wahr).

  (Insbesondere können als Konsequenz des Axioms auch Schnittmengen und — wie vorher schon gesehen — Mengen-Differenzen definiert werden.)
- *Ersetzungsaxiom:* Ist M eine Menge, so kann jedes Element x von M durch eine beliebige von x abhängige Menge N_x ersetzt und auf diese Weise eine neue Menge $\{N_x \mid x \in M\}$ gebildet werden.
- *Auswahlaxiom:* Für jede Menge $\mathcal{S}$ von paarweise disjunkten nicht-leeren Mengen (das heißt $\forall M \in \mathcal{S} : M \neq \emptyset$ und $\forall M, N \in \mathcal{S} : M \neq N \Longrightarrow M \cap N = \emptyset$) kann eine Menge A gebildet werden, die genau ein Element aus jedem Element von $\mathcal{S}$ enthält (und sonst keine weiteren Elemente).

  (Dieses Axiom ermöglicht es, aus jeder in $\mathcal{S}$ enthaltenen Menge je ein Element auszuwählen und aus diesen Elementen eine neue Menge zu bilden. Der entscheidende Punkt ist dabei, dass solch eine Auswahl in sehr großer Allgemeinheit ermöglicht wird. Falls eine konkrete Regel für die Auswahl angegeben werden kann oder falls $\mathcal{S}$ nur endlich viele Mengen als Elemente hat, wird das Auswahlaxiom nicht benötigt und die Auswahl kann auch mit den anderen Axiomen bewerkstelligt werden. Das Auswahlaxiom greift aber selbst dann, wenn $\mathcal{S}$ unendlich viele Elemente hat und keine Regel für die Auswahl konstruktiv angegeben werden kann. Bei früheren Kontroversen um die Konstruktivität mathematischer Beweise stand daher auch die Sinnhaftigkeit des Auswahlaxioms zur Debatte. Heutzutage wird das Axiom aber von einer überwiegenden Mehrheit der Mathematiker*innen akzeptiert, und in vielen Bereichen der modernen Mathematik kann nicht darauf verzichtet werden.)

V.2. Konstruktion der natürlichen Zahlen

Satz V.2.1. *Es existieren eine Menge* $\mathbb{N}$*, ein Abbildung* S *und ein Element* 1*, die die Peano-Axiome erfüllen.*

Beweis. Gemäß des Unendlichkeitsaxioms gibt es eine Menge U mit $\emptyset \in U$ und $\forall x \in U\colon x \cup \{x\} \in U$. Wir erklären $\mathbb{N}$ als Durchschnitt aller Teilmengen von U, die diese Eigenschaften von U teilen, definieren $S\colon \mathbb{N} \to \mathbb{N}$ durch $S(n) := n \cup \{n\}$ für alle $n \in \mathbb{N}$, und setzen $1 := \emptyset \in \mathbb{N}$. Wegen $S(n) = n \cup \{n\} \neq \emptyset = 1$ für alle $n \in \mathbb{N}$ ist dann $1 \notin \S(\mathbb{N})$.

Zum Nachweis der Injektivität von S seien $m, n \in \mathbb{N}$ mit $m \cup \{m\} = n \cup \{n\}$. Wäre $m \neq n$, so müssten $m \in n$ und $n \in m$ gelten, und die Paarmenge $\{m, n\}$ widerspräche dem Fundierungsaxiom. Also muss $m = n$ sein, und S ist injektiv.

Ist schließlich $M \subset \mathbb{N}$ mit $1 \in M$ und $S(M) \subset M$, so ist auch $M \subset U$ mit $\emptyset \in M$ und $\forall x \in M\colon x \cup \{x\} \in M$. Damit ist M eine der Mengen, als deren Durchschnitt $\mathbb{N}$ erklärt wurde. Es folgt also $\mathbb{N} \subset M$ und insgesamt $M = \mathbb{N}$.

Damit haben $\mathbb{N}$, S und 1 alle benötigten Eigenschaften. □

Etwas weniger formell bedeutet die Konstruktion dieses Beweises, dass wir die natürlichen Zahlen als

$$1 := \emptyset, \qquad 2 := 1 \cup \{1\}, \qquad 3 := 2 \cup \{2\}, \qquad 4 := 3 \cup \{3\}, \qquad 5 := 4 \cup \{4\}, \dots$$

und konkreter als

$$1 := \emptyset, \quad 2 := \{\emptyset\}, \quad 3 := \{\emptyset, \{\emptyset\}\}, \quad 4 := \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \quad 5 := \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}, \dots$$

erhalten haben.

Des Weiteren lässt sich auch eine Eindeutigkeitseigenschaft der natürlichen Zahlen formal herleiten:

Satz V.2.2. *Sind die Peano-Axiome einerseits für* $(\mathbb{N}, S, 1)$*, andererseits für* $(\widetilde{\mathbb{N}}, \widetilde{S}, \widetilde{1})$ *erfüllt, so können* $\mathbb{N}$ *und* $\widetilde{\mathbb{N}}$ *durch eine Bijektion* $f\colon \mathbb{N} \to \widetilde{\mathbb{N}}$ *identifiziert werden, die* $f(1) = \widetilde{1}$ *und* $f \circ S = \widetilde{S} \circ f$ *erfüllt.*

Da wir alle weiteren Operationen mit natürlichen Zahlen wie Addition, Multiplikation, et cetera nur aus der Nachfolge-Abbildung und der Eins konstruiert haben, sind $\mathbb{N}$ und $\widetilde{\mathbb{N}}$ damit auch im Hinblick auf solche Operationen *strukturell vollständig äquivalent* und unterscheiden sich höchstens durch die Benennungen konkreter Zahlen.

Beweisskizze. Wir definieren $f(n) \in \widetilde{\mathbb{N}}$ für alle $n \in \mathbb{N}$, indem wir $f(1) := \widetilde{1}$ und rekursiv $f(S(n)) := \widetilde{S}(f(n))$ für alle $n \in \mathbb{N}$ festsetzen. Über das Induktionsaxiom von $(\mathbb{N}, S, 1)$ ist damit die Wohldefiniertheit einer Abbildung $f\colon \mathbb{N} \to \widetilde{\mathbb{N}}$ gesichert, die per Definition $f(1) = \widetilde{1}$ und $f \circ S = \widetilde{S} \circ f$ erfüllt.

Um Surjektivität von f nachzuweisen, beobachten wir einerseits $\widetilde{1} \in f(\mathbb{N})$ und andererseits

$$\widetilde{n} \in f(\mathbb{N}) \Longrightarrow \widetilde{S}(\widetilde{n}) \in f(\mathbb{N}),$$

wobei letzteres gilt, weil $\widetilde{n} \in f(\mathbb{N})$ ja $\widetilde{n} = f(n)$ für ein $n \in \mathbb{N}$ und damit auch $\widetilde{S}(\widetilde{n}) = \widetilde{S}(f(n)) = f(S(n))$ bedeutet. Aufgrund dieser Beobachtungen gibt das Induktionsaxiom von $(\widetilde{\mathbb{N}}, \widetilde{S}, \widetilde{1})$ schon $f(\mathbb{N}) = \widetilde{\mathbb{N}}$, und damit ist f surjektiv.

Der Nachweis der Injektivität von f braucht Vorbereitungen. Wir benutzen, dass Addition und Subtraktion auf $\mathbb{N}$ und $\mathbb{Z}$ wie zuvor besprochen rekursiv eingeführt werden können und Standard-Rechenregeln genügen. Da wir für $\widetilde{\mathbb{N}}$ die gleichen Axiome wie für $\mathbb{N}$ haben, können wir auch auf $\widetilde{\mathbb{N}}$ eine Addition $\widetilde{+}$ rekursiv einführen und erhalten analoge Eigenschaften. Jetzt ergibt sich aus der Vertauschbarkeit $f \circ S = \widetilde{S} \circ f$ von f mit $S, \widetilde{S}$ die Vertauschbarkeit von f mit $+, \widetilde{+}$ in der Form $f(m{+}n) = f(m)\widetilde{+}f(n)$ für alle $m, n \in \mathbb{N}$. Um dies formal einzusehen, führt man bei festem $m \in \mathbb{N}$ aufbauend auf der rekursiven Definition von $+, \widetilde{+}$ und f vollständige Induktion nach $n \in \mathbb{N}$ durch. (Wir gehen dazu nicht ins Detail.)

Zum eigentlichen Nachweis der Injektivität von f sei $f(m) = f(n)$ für gewisse $m, n \in \mathbb{N}$. Ist $m \neq n$, so gilt entweder $n{-}m \in \mathbb{N}$ oder $m{-}n \in \mathbb{N}$. Wir betrachten im Fall $n{-}m \in \mathbb{N}$ erst den Subfall $m{=}1$. In diesem ist $n{\neq}1$, und mit $\widetilde{1} = f(1) = f(n) = f(n{-}1)\widetilde{+}\widetilde{1}$ (Rechenregeln und Vertauschbarkeit von f mit Addition verwendet) erhalten wir den Widerspruch, dass $\widetilde{1}$ ein Nachfolger ist. Im Subfall $m \neq 1$ bekommen wir auf ähnliche Weise $\widetilde{1}\widetilde{+}f(m{-}1) = f(m) = f(n) = f(n{-}m)\widetilde{+}\widetilde{1}\widetilde{+}f(m{-}1)$. Elimination des letzten Summanden $f(m{-}1)$ auf beiden Seiten der resultierenden Gleichung führt erneut auf den Widerspruch, dass $\widetilde{1}$ ein Nachfolger ist. Damit ist der Fall $n{-}m \in \mathbb{N}$ ausgeschlossen. Dieselbe Argumentation mit vertauschten Rollen von m und n schließt aber auch $m{-}n \in \mathbb{N}$ aus. Daher muss $m = n$ gelten, und damit ist f injektiv.

Insgesamt haben wir gezeigt, dass f eine Bijektion mit den gewünschten Eigenschaften ist. □

V.3. Beweis des Zornschen Lemmas

Sei $\mathcal{K}$ das Mengensystem aller Ketten in $\mathcal{X}$, auf dem wir die Mengen-Inklusion „$\subset$“ als Ordnungsrelation betrachten.

Wir zeigen zunächst, dass die Behauptung des Lemmas folgt, sobald die Existenz eines maximalen Elements von $\mathcal{K}$ nachgewiesen ist. Sei also $M \in \mathcal{K}$ ein maximales Element von $\mathcal{K}$. Dann existiert für die Kette M in $\mathcal{X}$ nach Voraussetzung des Lemmas eine obere Schranke s in $\mathcal{X}$. Um zu zeigen, dass s auch ein maximales Element von $\mathcal{X}$ ist, sei weiter $x \in \mathcal{X}$ mit $s \trianglelefteq x$. Dann ist auch $M \cup \{x\}$ eine Kette in $\mathcal{X}$ (denn Reflexivität gibt $x \trianglelefteq x$, die Schrankeneigenschaft von s gibt $m \trianglelefteq s$ für alle $m \in M$ und mit Transitivität folgt $m \trianglelefteq x$ für alle $m \in M$). Also ist $M \cup \{x\} \in \mathcal{K}$, und wegen der Maximalität von M folgt $M = M \cup \{x\}$, also $x \in M$. Wegen der Schrankeneigenschaft von s bedeutet dies $x \trianglelefteq s$ und wegen Antisymmetrie von $\trianglelefteq$ dann $x = s$. Damit ist s das gewünschte maximale Element von $\mathcal{X}$.

Im Hauptteil des Beweises zeigen wir nun die Existenz eines maximalen Elements M von $\mathcal{K}$. Dazu verwenden wir erst das Auswahlaxiom, um auf die Existenz einer Auswahl-Abbildung $f\colon \mathcal{P}(\mathcal{X}) \setminus \{\emptyset\} \to \mathcal{X}$ mit $f(T) \in T$ für alle nicht-leeren $T \subset \mathcal{X}$ zu schließen. (Genauer kann die Existenz von f dadurch begründet werden, dass Bemerkung (II.4.6) zum Abbildungsbegriff aus dem aktuellen Abschnitt II.4 auf die umgekehrte Element-Relation „$\ni$“ zwischen $\mathcal{P}(\mathcal{X}) \setminus \{\emptyset\}$ und $\mathcal{X}$ angewandt wird.) Als Nächstes vereinbaren wir für Ketten $T \in \mathcal{K}$ die Notation $\widehat{T} := \{x \in \mathcal{X} \mid T \cup \{x\} \in \mathcal{K}\}$ und definieren dann eine Abbildung $g\colon \mathcal{K} \to \mathcal{K}$, die eine Kette wenn möglich um ein mit Hilfe von f ausgewähltes Element erweitert, durch

$$g(T) := \begin{cases} T, & \text{falls } T \text{ maximales Element von } \mathcal{K} \\ T \sqcup \left\{ f\left(\widehat{T} \setminus T\right) \right\}, & \text{andernfalls} \end{cases}$$

für alle $T \in \mathcal{K}$. Formal ist die Abbildung g wohldefiniert, weil es für nicht-maximales T eine Kette $\widetilde{T} \in \mathcal{K}$ mit $T \subsetneqq \widetilde{T}$ gibt und mit $T \subsetneqq \widetilde{T} \subset \widehat{T}$ dann $\widehat{T} \setminus T \neq \emptyset$, Wohldefiniertheit von $f\left(\widehat{T} \setminus T\right) \in \widehat{T} \setminus T$ und $T \sqcup \left\{ f\left(\widehat{T} \setminus T\right) \right\} \in \mathcal{K}$ sichergestellt sind.

Für den weiteren Beweis nennen wir ein System von Ketten $\mathcal{T} \subset \mathcal{K}$ einen *Turm*, wenn es folgende Eigenschaften hat:

(a) Es gilt $\emptyset \in \mathcal{T}$.
(b) Für jedes $T \in \mathcal{T}$ ist $g(T) \in \mathcal{T}$.
(c) Für jede Kette (von Ketten) $\mathcal{S} \subset \mathcal{T}$ ist $\bigcup \mathcal{S} \in \mathcal{T}$.

Nun argumentieren wir in aufeinander aufbauenden Schritten:

(a) *Behauptung.* Das Mengensystem $\mathcal{K}$ aller Ketten ist ein Turm.
Die Eigenschaften (1) und (2) sind für $\mathcal{K}$ klar. Für (3) ist zu zeigen, dass für eine Kette von Ketten $\mathcal{S} \subset \mathcal{K}$ auch $\bigcup \mathcal{S}$ eine Kette ist, also $\bigcup \mathcal{S} \in \mathcal{K}$ gilt. Es seien dazu $x, \widetilde{x} \in \bigcup \mathcal{S}$. Es gibt dann $T, \widetilde{T} \in \mathcal{S}$ mit $x \in T$ und $\widetilde{x} \in \widetilde{T}$, und wegen der Ketteneigenschaft von $\mathcal{S}$ gilt $T \subset \widetilde{T}$ oder $\widetilde{T} \subset T$. Somit gilt $x, \widetilde{x} \in \widetilde{T}$ oder $x, \widetilde{x} \in T$. Als Elemente *einer* Kette ($\widetilde{T}$ oder T) erfüllen $x, \widetilde{x}$ dann $x \trianglelefteq \widetilde{x}$ oder $\widetilde{x} \trianglelefteq x$. Damit ist $\bigcup \mathcal{S}$ eine Kette.
(b) *Behauptung.* Ein beliebiger Durchschnitt von Türmen ist wieder ein Turm.
Dies ist klar, das sich die drei Eigenschaften (1), (2), (3) problemlos auf den Durchschnitt übertragen.
(c) In Anbetracht der Schritte 1) und 2) können wir einen „kleinsten“ Turm $\mathcal{T}_0$ als Durchschnitt *aller* Türme $\mathcal{T} \subset \mathcal{K}$ erhalten. Damit definieren wir

$$\mathcal{V} := \{V \in \mathcal{T}_0 \mid \forall T \in \mathcal{T}_0 \colon ((T \subset V) \vee (V \subset T))\} \text{ und } \mathcal{T}_V := \{T \in \mathcal{T}_0 \mid (T \subset V) \vee (g(V) \subset T)\} \text{ für } V \in \mathcal{V}.$$

(d) *Behauptung.* Für jedes $V \in \mathcal{V}$ ist $\mathcal{T}_V$ ein Turm.
Die Eigenschaft (1) ist klar. Für (2) betrachten wir $T \in \mathcal{T}_V \subset \mathcal{T}_0$ und bemerken $g(T) \in \mathcal{T}_0$ (da $\mathcal{T}_0$ ein Turm ist). Es tritt nun einer der drei Fälle $T \subset V \subset g(T)$, $T \not\subset V$, $V \not\subset g(T)$ ein. Im Fall $T \subset V \subset g(T)$ gilt, da sich $g(T)$ von T um höchstens ein Element unterscheidet, $(V = g(T)) \vee (T = V)$ und damit insbesondere $(g(T) \subset V) \vee (g(V) \subset g(T))$, was $g(T) \in \mathcal{T}_V$ bedeutet. Im Fall $T \not\subset V$ gilt wegen $T \in \mathcal{T}_V$ notwendig $g(V) \subset T \subset g(T)$ und $g(T) \in \mathcal{T}_V$. Im Fall $V \not\subset g(T)$ gilt wegen $V \in \mathcal{V}$ notwendig $g(T) \subset V$ und damit $g(T) \in \mathcal{T}_V$. Also gilt $g(T) \in \mathcal{T}_V$ in allen Fällen, und die Eigenschaft (2) ist für $\mathcal{T}_V$ gezeigt. Für die Eigenschaft (3) betrachten wir eine Kette von Ketten $\mathcal{S} \subset \mathcal{T}_V \subset \mathcal{T}_0$ und bemerken $\bigcup \mathcal{S} \in \mathcal{T}_0$ (da $\mathcal{T}_0$ ein Turm ist). Nach Definition von $\mathcal{T}_V$ gilt entweder $\exists T \in \mathcal{S} \colon g(V) \subset T$, somit $g(V) \subset \bigcup \mathcal{S}$ und $\bigcup \mathcal{S} \in \mathcal{T}_V$, oder es gilt $\forall T \in \mathcal{S} \colon T \subset V$, somit $\bigcup \mathcal{S} \subset V$ und erneut $\bigcup \mathcal{S} \in \mathcal{T}_V$. Damit ist (3) für $\mathcal{T}_V$ gezeigt.
(e) *Behauptung.* Für alle $V \in \mathcal{V}$ gilt $\mathcal{T}_V = \mathcal{T}_0$.
Dies folgt, da einerseits $\mathcal{T}_V$ gemäß 4) ein Turm mit $\mathcal{T}_V \subset \mathcal{T}_0 \subset \mathcal{K}$ und andererseits $\mathcal{T}_0$ der Schnitt aller Türme $\subset \mathcal{K}$ ist.
(f) *Behauptung.* Auch $\mathcal{V}$ ist ein Turm.
Die Eigenschaft (1) ist klar. Für (2) betrachten wir $V \in \mathcal{V} \subset \mathcal{T}_0$ und bemerken wieder $g(V) \in \mathcal{T}_0$. Für jedes $T \in \mathcal{T}_0$ gilt dann $T \subset V$ oder $V \subsetneqq T$. Im ersten Fall folgt trivial $T \subset g(V)$. Im zweiten Fall benutzen wir $T \in \mathcal{T}_0 \overset{5)}{=} \mathcal{T}_V$ und erhalten $g(V) \subset T$. Insgesamt gilt für alle $T \in \mathcal{T}_0$ also $(T \subset g(V)) \vee (g(V) \subset T)$, wir erhalten $g(V) \in \mathcal{V}$, und die Eigenschaft (2) ist für $\mathcal{V}$ gezeigt. Für die Eigenschaft (3) betrachten wir eine Kette (von Ketten) $\mathcal{S} \subset \mathcal{V} \subset \mathcal{T}_0$ und bemerken wieder $\bigcup \mathcal{S} \in \mathcal{T}_0$. Nach Definition von $\mathcal{V}$ gilt für jedes $T \in \mathcal{T}_0$ entweder $\exists V \in \mathcal{S} \colon T \subset V$ und somit $T \subset \bigcup \mathcal{S}$, oder es gilt $\forall V \in \mathcal{S} \colon V \subset T$ und somit $\bigcup \mathcal{S} \subset T$. Insgesamt gilt $\forall T \in \mathcal{T}_0 \colon ((T \subset \bigcup \mathcal{S}) \vee (\bigcup \mathcal{S} \subset T))$, womit $\bigcup \mathcal{S} \in \mathcal{V}$ und (3) für $\mathcal{V}$ gezeigt sind.
(g) *Behauptung.* Es gilt $\mathcal{V} = \mathcal{T}_0$.

Dies folgt, da einerseits $\mathcal{V}$ gemäß 6) ein Turm mit $\mathcal{V} \subset \mathcal{T}_0 \subset \mathcal{K}$ und andererseits $\mathcal{T}_0$ der Schnitt aller Türme $\subset \mathcal{K}$ ist.

(h) *Behauptung.* $\mathcal{T}_0$ ist eine Kette (von Ketten).

Für $T, V \in \mathcal{T}_0 \overset{7)}{=} \mathcal{V}$ gilt $(T \subset V) \vee (V \subset T)$ nach Definition von $\mathcal{V}$.

(i) *Behauptung.* Es gibt ein maximales Element M von $\mathcal{K}$.

Da $\mathcal{T}_0$ nach 8) eine Kette von Ketten und per Definition ein Turm ist, folgt $M := \bigcup_{T \in \mathcal{T}_0} T \in \mathcal{T}_0$ gemäß Eigenschaft (3) des Turms $\mathcal{T}_0$. Weiter gilt $g(M) \in \mathcal{T}_0$ gemäß Eigenschaft (2) des Turms $\mathcal{T}_0$, und gemäß Konstruktion von M ergibt sich $g(M) \subset M$. Nach Konstruktion von g gilt aber andererseits $T \subsetneq g(T)$, wann immer $T \in \mathcal{K}$ nicht-maximales Element von $\mathcal{K}$ ist. Somit verbleibt für $M \in \mathcal{T}_0 \subset \mathcal{K}$ nur die Möglichkeit, dass M maximales Element von $\mathcal{K}$ ist.

Damit ist der Beweis komplett. □

V.4. Beweisskizze des Wohlordnungssatzes

Wir skizzieren einen Beweis für die Behauptung, dass jede Menge wohlgeordnet werden kann. Der Beweis basiert auf dem Zornschen Lemma und wendet dieses Lemma in typischer Manier an:

Beweisskizze. Wir betrachten die Menge von Definitionsbereichen und Graphen von Wohlordnungen

$$\mathcal{W} := \{(D, G) \in \mathcal{P}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}^2) \mid (D, D, G) \text{ ist eine Wohlordnung auf } D\}$$

(wobei das Tripel (D, D, G) wie ursprünglich definiert für die Relation auf D mit Graph $G \subset D^2$ steht) und erklären eine Ordnungsrelation $\oplus \in \mathrm{Rel}(\mathcal{W})$ durch

$$(D, G) \oplus (\widetilde{D}, \widetilde{G}) :\iff ((D \subset \widetilde{D}) \wedge (G \subset \widetilde{G}) \wedge (\forall x \in D\colon \forall y \in \widetilde{D} \setminus D\colon (x, y) \in \widetilde{G}))$$

für alle $(D, G), (\widetilde{D}, \widetilde{G}) \in \mathcal{W}$. Grob gesagt bedeutet $(D, G) \oplus (\widetilde{D}, \widetilde{G})$ damit, dass die Relation $\widetilde{R} = (\widetilde{D}, \widetilde{D}, \widetilde{G})$ die Relation $R = (D, D, G)$ so fortsetzt, dass die Elemente von $\widetilde{D} \setminus D$ bezüglich $\widetilde{R}$ „größer oder gleich“ den Elementen von D sind.

Wir zeigen nun, dass die Voraussetzung des Zornschen Lemmas für $\mathcal{W}$ mit der Relation $\oplus$ erfüllt ist, dass also für jede Kette $\mathcal{K} \subset \mathcal{W}$ eine obere Schranke in $\mathcal{W}$ existiert. Dazu setzen wir für eine solche Kette $D_\mathcal{K} := \bigcup\{D \mid (D, G) \in \mathcal{K}\} \subset \mathcal{P}(\mathcal{X})$ und $G_\mathcal{K} := \bigcup\{G \mid (D, G) \in \mathcal{K}\} \subset \mathcal{P}(\mathcal{X}^2)$. Damit ist $G_\mathcal{K} \subset (D_\mathcal{K})^2$ der Graph einer Relation $R_\mathcal{K} \in \mathrm{Rel}(D_\mathcal{K})$, und es ist nicht schwer zu sehen, dass $R_\mathcal{K}$ eine Totalordnung auf $D_\mathcal{K}$ ist (denn zum Nachweis von Reflexivität, Antisymmetrie, Transitivität und Totalordnungs-Eigenschaft operiert man mit höchstens drei Elementen von $D_\mathcal{K}$ und kann sich mit der Ketteneigenschaft immer darauf zurückziehen, dass diese alle im Definitionsbereich D *nur eines* $(D, G) \in \mathcal{K}$ liegen). Etwas schwieriger ist der folgende Nachweis, dass $R_\mathcal{K}$ sogar eine Wohlordnung ist: Sei $\emptyset \neq T \subset D_\mathcal{K}$. Wir wählen $t \in T$. Dann gilt $t \in D$ für ein $(D, G) \in \mathcal{K}$, und $T \cap D \neq \emptyset$ besitzt bezüglich der Wohlordnung (D, D, G) ein kleinstes Element $x \in T \cap D$. Wir zeigen, dass dieses x schon das kleinste Element von T bezüglich $R_\mathcal{K}$ ist. Sei dazu $y \in T \subset D_\mathcal{K}$. Dann gilt $y \in \widetilde{D}$ für ein $(\widetilde{D}, \widetilde{G}) \in \mathcal{K}$. Wir unterscheiden nun die Fälle $y \in D$ und $y \notin D$. Im Fall $y \in D$ ist $y \in T \cap D$ und gemäß Wahl von x somit $(x, y) \in G \subset G_\mathcal{K}$. Im Fall $y \notin D$ erinnern wir uns, dass $\mathcal{K}$ eine Kette ist. Da $y \in \widetilde{D} \setminus D$ ja $\widetilde{D} \subset D$ und damit $(\widetilde{D}, \widetilde{G}) \oplus (D, G)$ ausschließt, muss $(D, G) \oplus (\widetilde{D}, \widetilde{G})$ gelten. Die letzte Bedingung aus der Definition von $\oplus$ liefert für $x \in D$ und $y \in \widetilde{D} \setminus D$ dann $(x, y) \in \widetilde{G} \subset G_\mathcal{K}$. Somit ist $(x, y) \in G_\mathcal{K}$ oder mit anderen Worten $x R_\mathcal{K} y$ in allen Fällen gezeigt, x ist also bezüglich $R_\mathcal{K}$ kleinstes Element von T. Insgesamt erhalten wir, dass $R_\mathcal{K}$ eine Wohlordnung auf $D_\mathcal{K}$, also $(D_\mathcal{K}, G_\mathcal{K}) \in \mathcal{W}$ ist. Man prüft nun problemlos, dass $(D, G) \oplus (D_\mathcal{K}, G_\mathcal{K})$ für alle $(D, G) \in \mathcal{K}$ gilt und somit $(D_\mathcal{K}, G_\mathcal{K})$ eine obere Schranke für $\mathcal{K}$ ist.

Insgesamt ist die Voraussetzung des Zornschen Lemmas erfüllt, und dieses liefert nun die Existenz eines maximalen Elements (D, G) von $\mathcal{W}$, für das $R = (D, D, G)$ eine Wohlordnung auf $D \subset \mathcal{X}$ ist. Angenommen, es ist $D \subsetneq \mathcal{X}$. Dann könnten wir ein $x_0 \in \mathcal{X} \setminus D$ wählen, dieses x_0 durch die Festlegungen $D_0 := D \sqcup \{x_0\}$, $G_0 := G \sqcup (D_0 \times \{x\})$ als neues größtes Element hinzufügen und erhielten $(D_0, G_0) \in \mathcal{W}$ mit $(D, G) \oplus (D_0, G_0)$, aber $(D_0, G_0) \neq (D, G)$. Da dies im Widerspruch zur Maximalität von (D, G) stünde, muss tatsächlich $D = \mathcal{X}$ gelten. Dies bedeutet aber, dass R tatsächlich eine Wohlordnung auf ganz $\mathcal{X}$ ist. □

Tatsächlich stellen sich das *Auswahlaxiom*, das *Zornsche Lemma* und der *Wohlordnungssatz* sogar als *zueinander äquivalent* heraus. Da wir mit den vorausgehenden Beweisen schon gesehen haben, dass das Auswahlaxiom das Zornsche Lemma und das Zornsche Lemma den Wohlordnungssatz implizieren, ist für die Äquivalenz nur noch zu zeigen, dass der Wohlordnungssatz das Auswahlaxiom impliziert. Da mittels Wohlordnung sehr kanonisch (kleinste) Elemente ausgewählt werden können, ist letzteres tatsächlich vergleichsweise einfach: Für ein beliebiges System $\mathcal{S}$ disjunkter nicht-leerer Mengen liefert der Wohlordnungssatz die Existenz einer Wohlordnung auf $\bigcup \mathcal{S}$, und bezüglich dieser existiert in jeder Menge $M \in \mathcal{S}$, die ja nicht-leere Teilmenge von $\bigcup \mathcal{S}$ ist, ein kleinstes Element x_M (das durch seine Eigenschaft zu einem gewissen Grad konstruktiv charakterisiert ist). Gemäß dem Ersetzungsaxiom kann nun die Menge $\{x_M | M \in \mathcal{S}\}$ gebildet werden. Diese hat dann die Auswahleigenschaft, dass sie mit jeder der in $\mathcal{S}$ enthaltenen Mengen genau ein Element gemeinsam hat.

V.5. Ausblick: Kardinalzahlen

- Im Prinzip kann die Mächtigkeit einer Menge nicht nur vergleichen, sondern auch als eigenständige Eigenschaft einer einzelnen Menge definiert und betrachtet werden. Dazu ordnet man jeder Äquivalenzklasse[1] von gleichmächtigen Mengen M als Kenngröße eine gewisse Kardinalzahl zu, die die Anzahl der Elemente verallgemeinert. Für die kleinste unendliche Kardinalität, die Kardinalität von $\mathbb{N}$ und allen abzählbar unendlichen Mengen ist die Bezeichnung $\aleph_0$ üblich (mit dem Aleph $\aleph$, dem ersten Buchstaben des hebräischen Alphabets). Die nächstgrößere[2] Kardinalität, also die kleinste überabzählbare Kardinalität, nennt man $\aleph_1$.
- Die berühmte *Kontinuumshypothese* fragt nun, ob tatsächlich $|\mathbb{R}| = \aleph_1$ gilt oder nicht, ob also die nächstgrößere Kardinalität nach $|\mathbb{N}|$ schon die Kardinalität $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$ des sogenannten Kontinuums $\mathbb{R}$ ist oder es noch Zwischenstufen gibt. Dieses 1878 von Cantor formulierte Problem war über viele Jahrzehnte eine fundamentale offene Frage der Mathematik und wurde erst durch aufsehenerregende (Meta-)Sätze der berühmten Mathematiker K. Gödel (1906–1978) und P. Cohen (1934–2007) aus den Jahren 1938 und 1963 gelöst: Das erstaunliche Fazit lautet, dass die Frage der Kontinuumshypothese *im Rahmen des Zermelo-Fraenkel-Axiomensystems der Mengenlehre unentscheidbar* ist und auf Basis dieses Systems nicht beantwortet werden kann. Genauer können sowohl die Gültigkeit als auch die Nicht-Gültigkeit der Kontinuumshypothese in konsistenter Weise als Zusatz-Annahmen zum Axiomensystem hinzugefügt werden. Das Auftreten eines solchen prinzipiell unentscheidbaren Problems ist überraschend und zu einem gewissen Grad schockierend. Dennoch handelt es sich um einen Fakt der mathematischen Theorie, mit dem man (fortan) leben muss. *In der Praxis treten unentscheidbare Probleme glücklicherweise sehr selten auf*, und in den allermeisten mathematischen Disziplinen sind die Gültigkeit oder Nicht-Gültigkeit der Kontinuumshypothese und anderer unentscheidbarer Probleme kaum von Belang.
- Zum Abschluss dieses Abschnitts behandeln wir einen bekannten Satz von Cantor, demzufolge man durch Bildung der Potenzmenge immer noch größere Mächtigkeiten erhält und es daher bei der Mächtigkeit von Mengen kein Limit gibt:

  Satz V.5.1 (Satz von Cantor)**.** *Für jede Menge M gilt $|\mathcal{P}(M)| > |M|$.*

  Ausgehend von $|M| = n \Longrightarrow |\mathcal{P}(M)| = 2^n$ für endliche Mengen M (was prinzipiell schon bei der ursprünglichen Einführung der Potenzmenge erwähnt wurde) schreibt man für die Mächtigkeit $|\mathcal{P}(M)|$ der Potenzmenge auch allgemein $2^{|M|}$. Damit lautet der Satz von Cantor $2^{|M|} > |M|$, und die Kontinuumshypothese fragt nach $2^{|\mathbb{N}|} = \aleph_1$.

  Der elegante Beweis des Satzes lehnt sich an die Grundidee des Russellschen Paradoxons an:

[1] Die Äquivalenzklassen können, weil eben die Menge aller Mengen nicht existiert, nicht als Mengen gebildet werden. Abstrakt kann man sich eine „Ansammlung“ aller zu einer gegebenen Menge gleichmächtigen Mengen aber trotzdem vorstellen, und im verallgemeinerten Sinn sogenannter Klassen existiert diese Ansammlung auch als formales Objekt.

[2] Tatsächlich ergibt sich mit dem Wohlordnungssatz, dass die Kardinalzahlen nicht nur die Totalordnungs-, sondern auch die Wohlordnungseigenschaft haben. Nur deshalb kann man von einer nächstgrößeren Kardinalität überhaupt sprechen.

Da $M \to \mathcal{P}(M)$, $x \mapsto \{x\}$ eine Injektion ist, gilt $|M| \leq |\mathcal{P}(M)|$. Um $|M| < |\mathcal{P}(M)|$ zu zeigen, bleibt also $|M| = |\mathcal{P}(M)|$ (also die Existenz einer Bijektion $f\colon M \to \mathcal{P}(M)$) auszuschließen. *Angenommen*, es gäbe solch eine Bijektion f. Dann ließe sich die Teilmenge $T := \{x \in M \mid x \notin f(x)\} \in \mathcal{P}(M)$ bilden, und wegen der Surjektivität von f gäbe es ein $a \in M$ mit $f(a) = T$. Nun erhielte man einerseits im Fall $a \in T$, dass $a \notin f(a)$, also $a \notin T$ gelten müsste, andererseits im Fall $a \notin T$, dass $a \in f(a)$, also $a \in T$ gelten müsste. Damit ist in jedem Fall ein *Widerspruch* erreicht. Also existiert keine Bijektion $f\colon M \to \mathcal{P}(M)$, womit $|M| < |\mathcal{P}(M)|$ gezeigt ist. □

Unter anderem gibt der Satz von Cantor auch die Existenz einer Menge G mit $|G| > |\mathcal{P}^k(\mathbb{N})|$ für alle $k \in \mathbb{N}_0$, wobei sich $\mathcal{P}^k(\mathbb{N})$ durch k-fache Bildung der Potenzmenge in der Form $\mathcal{P}^k(\mathbb{N}) := \mathcal{P}(\mathcal{P}(\ldots\mathcal{P}(\mathbb{N})\ldots))$ ergibt. Man erhält dieses G einfach als $G := \mathcal{P}\big(\bigcup_{k\in\mathbb{N}_0} \mathcal{P}^k(\mathbb{N})\big)$.

Tatsächlich gibt es der Kardinalitäten insgesamt sogar „zu viele“, um diese in einer Menge zusammenfassen zu können. Genauer gilt, dass eine Menge $\mathcal{K}$ aller Kardinalitäten nicht existiert. Gäbe es nämlich eine solche Menge $\mathcal{K}$, so ließe sich auch eine Vereinigungsmenge $G := \bigcup_{\aleph\in\mathcal{K}} M_\aleph$ bilden, die Mengen $M_\aleph$ *jeder* Kardinalität $|M_\aleph| = \aleph$ als Teilmengen enthält. Insbesondere enthielte G eine Teilmenge der Kardinalität $|\mathcal{P}(G)|$, was $|\mathcal{P}(G)| \leq |G|$ bedeutete und damit im Widerspruch zum Satz von Cantor stände.

V.6. Beweis des Cantor-Schröder-Bernstein Theorems

Wir beweisen die Regel $(|M| \leq |N| \wedge |N| \leq |M|) \Longrightarrow |M| = |N|$, also des Satzes von Cantor-Schröder-Bernstein: Es sei sowohl $|M| \leq |N|$ als auch $|N| \leq |M|$. Per Definition gibt es dann Injektionen $f\colon M \to N$ und $g\colon N \to M$, und aus letzterer erhalten wir durch Verkleinerung des Zielbereichs eine Bijektion $\widetilde{g}\colon N \to \mathrm{Bild}(g)$ mit Umkehrbijektion $\widetilde{g}^{-1}\colon \mathrm{Bild}(g) \to N$. Wir definieren $A_n \subset M$ für alle $n \in \mathbb{N}_0$ durch den Rekursionsanfang $A_0 := M \setminus \mathrm{Bild}(g)$ und den Rekursionsschritt $A_{n+1} := g(f(A_n))$ für alle $n \in \mathbb{N}_0$. Damit können wir $A_* := \bigcup_{n\in\mathbb{N}_0} A_n$ setzen und $h\colon M \to N$ durch

$$h(x) := \begin{cases} f(x), & \text{für } x \in A_* \\ \widetilde{g}^{-1}(x), & \text{für } x \notin A_* \end{cases}$$

für alle $x \in M$ definieren, denn im Fall $x \notin A_*$ liegt x insbesondere in $M \setminus A_0 = \mathrm{Bild}(g)$, wo $\widetilde{g}^{-1}$ definiert ist.

Um Injektivität von h zu zeigen, betrachten wir $x, y \in M$ mit $h(y) = h(x)$ und unterscheiden Fälle: Im Fall $x, y \in A_*$ gilt $f(y) = f(x)$, und $y = x$ folgt per Injektivität von f. Im Fall $x, y \notin A_*$ gilt $\widetilde{g}^{-1}(y) = \widetilde{g}^{-1}(x)$, und $y = x$ folgt aus der Bijektivität von $\widetilde{g}^{-1}$. Im Fall $x \in A_*$, $y \notin A_*$ gilt $\widetilde{g}^{-1}(y) = f(x)$. Es gibt dann ein $n \in \mathbb{N}_0$ mit $x \in A_n$ und folglich $y = \widetilde{g}(f(x)) = g(f(x)) \in A_{n+1} \subset A_*$. Damit ist ein Widerspruch erreicht und das Auftreten dieses Falls tatsächlich ausgeschlossen. Analog sieht man, dass auch der Fall $x \notin A_*$, $y \in A_*$ nicht eintreten kann. Damit ist wie benötigt $y = x$ in allen möglichen Fällen gezeigt.

Um Surjektivität von h zu zeigen, sei $y \in N$. Da $\widetilde{g}^{-1}$ bijektiv ist, können wir $y = \widetilde{g}^{-1}(x)$ mit $x \in \mathrm{Bild}(g) = M \setminus A_0$ schreiben. Wir unterscheiden wieder Fälle: Im Fall $x \notin A_*$ erhalten wir direkt $y = h(x) \in \mathrm{Bild}(h)$. Im Fall $x \in A_*$ muss $x \in A_n = g(f(A_{n-1}))$ für ein $n \in \mathbb{N}$ gelten (denn wir hatten $A_* = \bigcup_{n\in\mathbb{N}_0} A_n$ gewählt, und $x \notin A_0$ ergab sich bei der Wahl von x). Es gibt also ein $x' \in A_{n-1}$ mit $x = g(f(x'))$, und wir erhalten $y = \widetilde{g}^{-1}(x) = \widetilde{g}^{-1}(g(f(x'))) = \widetilde{g}^{-1}(\widetilde{g}(f(x'))) = f(x')$ mit $x' \in A_*$. Dies bedeutet auch in diesem Fall $y = h(x') \in \mathrm{Bild}(h)$. Damit ist, wie für Surjektivität benötigt, $y \in \mathrm{Bild}(h)$ in allen Fällen gezeigt.

Insgesamt ist $h\colon M \to N$ injektiv und surjektiv, also auch bijektiv, und es gilt $|M| = |N|$. □

Literaturverzeichnis

[1] Zitat F. Bernstein aus: *R. Dedekind, Gesammelte mathematische Werke, Dritter Band*, herausgegeben von R. Fricke, E. Noether, Ö. Ore. S. 449, Vieweg, 1932.

[2] S. Bosch, *Lineare Algebra.* Springer, 2014.

[3] T. Bröcker, *Lineare Algebra und Analytische Geometrie.* Birkhäuser, 2004.

[4] O. Deiser, *Einführung in die Mengenlehre.* Springer, 2010.

[5] G. Fischer, *Lernbuch Lineare Algebra und Analytische Geometrie.* Springer, 2019.

[6] G. Fischer, B. Springborn, *Lineare Algebra.* Springer, 2020.

[7] O. Forster, *Analysis 1.* Springer, 2015.

[8] G. Greefrath, R. Oldenburg, H.-S. Siller, V. Ulm, H.-G. Weigand, *Didaktik der Analysis.* Springer, 2016.

[9] H.-W. Henn, A. Filler, *Didaktik der Analytischen Geometrie und Linearen Algebra.* Springer, 2015.

[10] P.R. Halmos, *Naive Set Theory.* Springer, 1974.

[11] H. Heuser, *Lehrbuch der Analysis. Teil 1.* Vieweg+Teubner, 2009.

[12] S. Hildebrandt, *Analysis 1.* Springer, 2006.

[13] K. Jänich, *Lineare Algebra.* Springer, 2008.

[14] K. Königsberger, *Analysis 1.* Springer, 2004.

[15] W. Walter, *Analysis 1.* Springer, 2009.

[16] H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Lamotke, K. Mainzer, J. Neukirch, A. Prestel, R. Remmert, *Zahlen.* Springer, 1992.