Oliver Deiser

Grundbegriffe der wissenschaftlichen Mathematik

Sprache, Zahlen und erste Erkundungen

Springer

Springer-Lehrbuch

Oliver Deiser

Grundbegriffe der wissenschaftlichen Mathematik

Sprache, Zahlen und erste Erkundungen

Springer

PD Dr. Oliver Deiser
Fachbereich Mathematik
Freie Universität Berlin
Arnimallee 6
14195 Berlin
Deutschland
deiser@mi.fu-berlin.de

ISSN 0937-7433
ISBN 978-3-642-11488-5 e-ISBN 978-3-642-11489-2
DOI 10.1007/978-3-642-11489-2
Springer Heidelberg Dordrecht London New York

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Mathematics Subject Classification (2010): 00-01, 00A05, 00A07

Einbandentwurf: WMXDesign GmbH, Heidelberg

Gedruckt auf säurefreiem Papier

Springer ist Teil der Fachverlagsgruppe Springer Science+Business Media (www.springer.com)

für meine Eltern

Inhalt

Vorwort

Dieses Buch behandelt in kompakter und dichter Form die Sprache der modernen Mathematik und einige ihrer grundlegenden Objekte und Begriffe. Das entstehende Abbild der wissenschaftlichen Mathematik ist zweifach unvollständig: Zum einen werden die logisch-mengentheoretischen Grundlagen der Mathematik informal und damit nicht in der heute möglichen Exaktheit präsentiert. Zum anderen bilden die vorgestellten Themen eine persönlich gefärbte Auswahl, die noch vielfältig ergänzt werden könnte, und sie decken auch immer nur erste Schritte ab, sodaß der eigentliche Charakter der mathematischen Gebiete und ihre wechselseitigen Beziehungen nur bedingt ans Licht kommen. Um diese beiden Unvollständigkeiten zu beseitigen, benötigt man viele Bücher und viele Jahre Studium. Dieser Text ist dabei nur einführend, ergänzend und begleitend. Er ist keineswegs trivial, aber letztendlich doch einfach, auch wenn es für den Anfänger einige Zeit brauchen wird, seine Einfachheit zu sehen. Er steht im Regal der mathematischen Grundausbildung, aufgrund seiner individuellen Auswahl und Perspektive aber nicht dort, wo ein Kanon etabliert werden soll. Er will ein im Vergleich zur Allgemeinbildung deutlich vertieftes Verständnis mathematischer Grundbegriffe und Theorien vermitteln und dem Leser eine konzentrierte Begegnung mit der Mathematik ermöglichen. Ziel ist die Bildung oder Stärkung einer sprachlichen und inhaltlichen Grundlage, die sich in der weiter- und tiefer-reichenden Auseinandersetzung mit mathematischen Ergebnissen, Fragen und Denkweisen bewährt. Das Buch will also nicht nur Wissen vermitteln, sondern auch zu etwas befähigen.

Die ins Auge gefaßte Lesergruppe umfaßt damit alle, die an der Mathematik bereits interessiert sind und im Idealfall sogar schon ein Bedürfnis nach erhöhter Genauigkeit und Systematik mitbringen. Unter diese Gruppe fallen Schüler um die Abiturzeit, die ein Mathematikstudium in Erwägung ziehen, Teilnehmer eines mathematischen Vor- oder Einführungskurses an der Universität, Studierende der Mathematik in den ersten Semestern und nicht zuletzt auch von der Strenge und Schönheit der Mathematik faszinierte Laien. Auch für Lehrer und Dozenten kann das Buch, von einer anderen Warte aus betrachtet, interessant sein.

Neben Interesse wird lediglich vorausgesetzt, daß der Leser mit der Schulmathematik in rudimentärer Form vertraut ist. Dabei müssen keine Wissenslücken befürchtet werden, denn alles, was auf die Bühne kommt, wird durch eine Definition vorgestellt. Die Darstellung ist direkt, und den für diesmal höheren Zielen der Dichte und Kompaktheit folgend verzichtet sie, nicht leichten Herzens, auf eine ideengeschichtlich geleitete Anordnung und historische Einbettung ihres Inhalts.

Die Gestaltung des Textes folgt einer recht strengen Symmetrie, und der Leser findet einige Bemerkungen hierzu im Anschluß an dieses Vorwort. Dort gehen wir auch auf die Möglichkeit der nichtlinearen Lektüre ein: Es ist nicht notwendig, die ersten beiden Abschnitte „Die Sprache der Mathematik“ und „Zahlen“ gelesen zu haben, um die „Ersten Erkundungen“ mit Gewinn lesen zu können. Nach den Ausführungen zum Aufbau stellen wir die Themen der drei Abschnitte des Textes im Überblick vor. Der Orientierung dient darüber hinaus die Einführung, das strukturierte Inhaltsverzeichnis, die Tafel der Notationen und der Index. Am Ende des Buches findet sich zudem eine kommentierte Zusammenstellung ergänzender und weiterführender Literatur.

Mein Dank gilt der Deutschen Telekom Stiftung, die die Entstehung des Buches im Rahmen des „M-Bridge“-Projekts entscheidend gefördert hat. Weiter danke ich Jan-David Hardtke, Kristine Kaiser, Ilja Klebanov, Sophie Knell, Simon Lücking, Ute Skambraks und Kerstin Weller, die das Manuskript kritisch gelesen und in Tutorien an der FU Berlin erprobt haben. Heinz Jaskolla hat den Text durch zahllose Anregungen und Vorschläge ganz wesentlich verbessert.

Berlin, im November 2009

Oliver Deiser

Aufbau und Themen des Buches

Das Buch zerfällt in die drei Abschnitte „Die Sprache der Mathematik“, „Zahlen“ und „Erste Erkundungen“. Die „Sprache der Mathematik“ versammelt, was in der Mathematik überall vorkommt und verwendet wird, so etwa Mengen und Funktionen samt ihrem elementaren begrifflichen Umfeld. Die „Zahlen“ stellen die Grundobjekte vor, mit denen die Sprache operiert und weitere Objekte bildet. Die „Ersten Erkundungen“ zeigen, was man entdecken kann, wenn man mathematische Fragestellungen genauer verfolgt. Sie sind der Beginn von großen Theorien, die sich oft gegenseitig befruchten und beleuchten.

Die beiden ersten Abschnitte wiegen zusammen genommen in etwa so viel wie der dritte Abschnitt. Sie weisen jeweils drei Kapitel auf, während der dritte Abschnitt aus sechs Kapiteln besteht. Jedes Kapitel hat sechs Unterkapitel, an die sich eine Übungssektion mit je zweimal zwölf Übungsaufgaben anschließt. Die Übungen sind den einzelnen Unterkapiteln zugeordnet und geben dem Leser die Gelegenheit, sich die vorgestellten Begriffe zu eigen zu machen. Es ergeben sich so $3 + 3 + 6 = 12$ Kapitel mit $12 \cdot 6 = 72$ Unterkapiteln und $12 \cdot 24 = 288$ Übungsaufgaben. Hierbei wird keine Zahlenmystik angestrebt, sondern dieser Aufbau des Buches dient dazu, den Inhalten des Buches Form, Struktur und Symmetrie zu geben. Ein Ausufern einzelner Themen wird so von vorneherein unmöglich, und der Autor steht unter dem freiwilligen Zwang der Auswahl und Beschränkung. Ein „13. Kapitel“ stellt der Exkurs über Mächtigkeiten am Ende des ersten Abschnitts dar, der mit seinen irritierenden Ergebnissen und unbefriedigend beantworteten Fragen bewußt die Symmetrie stört. Der Exkurs nimmt räumlich die Hälfte eines der anderen Kapitel ein und besitzt drei Unterkapitel mit zwölf zugehörigen Übungen.

Die klassische lineare Lektüre des Buches ist denen zu empfehlen, die Systematik, logischen Aufbau, Genauigkeit und Gründlichkeit im Sinne einer Fundamentbildung schätzen und die dafür bereit sind, die Begegnung mit mathematischen Theorien etwas zurückzustellen. Aber auch eine nichtlineare Lektüre ist möglich. Die sechs Kapitel im dritten Abschnitt sind in ihrer Anordnung nicht zufällig, aber doch auch weitgehend unabhängig voneinander. Zudem genügt dort vielfach ein Grundverständnis der mathematischen Sprache und ein Grundwissen über die Zahlen, sodaß der dritte Abschnitt letztendlich nur bildlich auf den beiden ersten Abschnitten so ruht wie ein Dach auf Säulen und Säulen auf ei-

nem Fundament. Wer gerne möchte, kann mit dem zahlentheoretischen Kapitel über Teiler beginnen, selbst der Jargon der Mengen wird dort sparsam verwendet. Ergänzend kann dann bei Bedarf das erste Kapitel des zweiten Abschnitts herangezogen werden, das die Struktur der natürlichen Zahlen und insbesondere die Induktion diskutiert. In ähnlicher Weise verhält es sich mit dem Kapitel über Grenzwerte. Wir stellen dort die fundamentalen Struktureigenschaften der reellen Zahlen kurz vor, eine ausführliche Behandlung und Motivation findet der Leser im zweiten und dritten Kapitel des Abschnitts über Zahlen. Das Kapitel über Matrizen kann ohne Kenntnis des Körperbegriffs gelesen werden, erscheint mit entsprechendem Vorwissen aber in größerer Allgemeinheit. Am Ende des Kapitels über Matrizen führt eine Brücke zu den Graphen, über die man auch erst dann gehen kann, sobald man sich für Fragen der Graphentheorie interessiert. Das Kapitel über Graphen selbst ist wie das Kapitel über Teiler elementar zugänglich und sei insbesondere denen empfohlen, die den spielerischen Charakter der Mathematik schätzen. Das Kapitel über Gruppen wirkt sicherlich weniger „ad hoc“, wenn der Leser mit dem Aufbau des Zahlsystems und mit den Struktureigenschaften von bijektiven Funktionen vertraut ist, aber wer abstrakte Strukturen schätzt und als solche kennenlernen möchte, braucht hier kein Vorspiel auf dem Theater. Für das Kapitel über Wahrscheinlichkeiten schließlich ist die Kenntnis des Grenzwertbegriffs unentbehrlich.

Im Fließtext werden oft Behauptungen aufgestellt, die in den Übungsabschnitten wieder auftauchen und bewiesen werden sollen. Der Text bleibt dort lückenhaft, wo er durch eigenständiges Denken ergänzt werden kann. Er läßt sich grob dem zuordnen, was in didaktischen Kreisen als „modifizierte Methode von Robert Moore“ bekannt ist. Moore hat fortgeschrittenen Studenten lediglich einige mathematische Begriffe und Sätze angeschrieben, die sie dann selbständig untersuchen und beweisen sollten. In angepaßter Form wurde seine Methode dann nicht nur für den wissenschaftlichen Nachwuchs, sondern auch für alle Studierenden der Mathematik angewendet. Der Lernende soll die Dinge hinterfragen und aktiv nach kreativen Antworten auf aufgeworfene Fragen suchen. Der Ansatz wurzelt in langen Traditionen der Aufklärung: „sapere aude“ steht bei Horaz, und Kant hat dies gewichtig mit „Habe Mut, dich deines eigenen Verstandes zu bedienen“ übersetzt. Für die Mathematik gilt die Interpretation: „Habe Mut, Beweise selbst zu führen.“

Der Gedanke der eigenständigen Arbeit führt dazu, daß im ersten Abschnitt kein einziger Beweis einer elementaren Eigenschaft der behandelten Begriffe vorgeführt wird. Beweise finden sich erst dort, wo die Beweislast aufgrund ihres Gewichts nicht dem Leser überlassen werden kann und das Nachvollziehen von Argumenten den Verstand schon genug fordert. Damit die Lektüre des Buches aber auch außerhalb von begleiteten Kursen erleichtert wird, werden zu einigen Übungen Lösungshinweise gegeben, und ein Anhang stellt ausführliche Lösungsvorschläge für $3 \cdot 8 = 24$ Übungen des ersten Abschnitts bereit, also für ein Drittel der dortigen Aufgaben. Diese sind in den Übungssektionen mit einem „L“ für „Lösung“ gekennzeichnet.

Nach diesen Bemerkungen zum Aufbau des Buches wollen wir nun noch die behandelten Themen vorstellen.

Erster Abschnitt: Die Sprache der Mathematik

Unter „Sprache der Mathematik“ verstehen wir hier nicht eine formale Beschreibung der Syntax mathematischer Aussagen, sondern einen relativ weit gefaßten Satz an Begriffsbildungen, Sprechweisen und Konventionen, die in der Mathematik überall vorkommen, keine Merkmale einer Theoriebildung aufweisen und die jeder Mathematiker kennt und „fließend“ beherrscht. Naturgemäß spielen dabei grundlegende logische und mengentheoretische Konzepte eine wichtige Rolle.

Wir beginnen im Kapitel „Mathematisches Argumentieren“ mit einer ausführlichen Diskussion der mathematischen Junktoren. Wir betrachten ihre Verwendung und Semantik unter verschiedenen Gesichtspunkten, und wir isolieren einige Beweismuster, zu denen die Aussagenlogik Anlaß gibt. Schließlich behandeln wir die Quantoren im Zusammenspiel mit Funktionen, Relationen und Konstanten, wobei wir hier ein naives Grundverständnis des Funktions- und Relationsbegriffs annehmen. Das Kapitel endet mit einer tabellarischen Übersicht aussagenlogischer Tautologien und Quantorenregeln.

Das zweite Kapitel widmet sich dann dem Mengenbegriff und damit dem Grundbegriff der modernen mathematischen Sprache. Wir zeigen die logischen Probleme der uneingeschränkten Zusammenfassung von Objekten zu Mengen auf und stellen dann die wichtigsten „harmlosen“ Mengenbildungen und Operationen mit Mengen zusammen. Dabei lernen wir in der Definition des geordneten Paares auch ein einfaches Beispiel für die universelle interpretative Kraft des Mengenbegriffs kennen. Schließlich führen wir Potenzmengen und allgemeinere Mengensysteme ein und diskutieren die häufig in der Mathematik anzutreffende Stufung in „Punkt, Menge von Punkten, Mengensystem von Punktmengen“.

Das dritte Kapitel behandelt Relationen und Funktionen. Wir definieren Relationen als Mengen von geordneten Paaren und besprechen die wichtigsten Struktureigenschaften von Relationen und ihre Kombinationen, unter denen die Äquivalenzrelationen und die partiellen Ordnungen herausragen. Anschließend führen wir Funktionen als spezielle Relationen ein und versammeln die zugehörigen Begriffsbildungen, die in der Mathematik häufig verwendet werden. Ein Abschnitt über Wohldefiniertheit und Kongruenzrelationen diskutiert Funktionen im Zusammenspiel mit Relationen. Ein universelles funktionales Thema in der Mathematik ist schließlich das der Isomorphie mathematischer Strukturen, und wir geben eine elementare, aber zugleich auch sehr allgemeine Einführung in dieses Thema am Ende des Kapitels.

Die Darstellung der mathematischen Sprache mündet in einen anspruchsvolleren Exkurs über Mächtigkeiten, den man auch als ein erstes Beispiel für eine mathematische Theoriebildung ansehen kann. Wir vergleichen die Größe von Mengen mit Hilfe von Injektionen und Bijektionen, beweisen den Satz von Cantor-Bernstein, definieren den Begriff der Unendlichkeit und zeigen, daß Größenunterschiede im Unendlichen existieren. Weiter formulieren wir die Kontinuumshypothese und deuten die Problematik an, die dem Linearkontinuum im Rahmen der üblichen Fundierung der Mathematik innewohnt.

Der gesamte erste Abschnitt einschließlich des Exkurses kommt ohne Zahlen aus. In den Übungen gestatten wir uns aber gelegentliche Anleihen bei den Zahlen (einschließlich der Induktion), um Beispiele zur Verfügung zu haben, mit denen der Leser vertraut ist. Auch wird gegen Ende des Exkurses das Fehlen der Zahlen deutlich, sodaß auch von theoretischer Seite der Beginn des zweiten Abschnitts herbeigeführt wird.

Zweiter Abschnitt: Zahlen

Die Zahlen als Grundobjekte der Mathematik bilden das Thema des zweiten Abschnitts. Speziell hier gilt, daß der Leser schon viel mitbringt, und wir gestatten uns deswegen einen recht detaillierten und genauen Blick auf einen untadeligen Aufbau des Zahlsystems.

Das erste Kapitel beginnt mit einer informalen Beschreibung des Zählens. Diese Beschreibung führt uns zur Formulierung eines Induktionsprinzips und weiter zur Definition einer Dedekind-Struktur, die das Wesen des Zählens einfängt. Für interessierte Leser beweisen wir einen Isomorphiesatz und skizzieren die Möglichkeit der Konstruktion einer derartigen Struktur. Irgendeine Dedekind-Struktur dient uns dann im folgenden als Menge der natürlichen Zahlen $\mathbb{N}$. Aus der genuinen Nachfolgerbildung gewinnen wir mit Hilfe von rekursiven Definitionen die Addition, die Multiplikation und die Exponentiation. Danach definieren wir die Ordnung auf den natürlichen Zahlen, und wir formulieren und beweisen eine Verstärkung der vollständigen Induktion und das zugehörige Prinzip des kleinsten Elements. In einem Ausblick betrachten wir dann die rekursiv definierten Funktionen auf den natürlichen Zahlen noch genauer und gelangen so zu einer Definition der algorithmischen Berechenbarkeit einer Funktion.

Im zweiten Kapitel führen wir mit Hilfe der natürlichen Zahlen die ganzen Zahlen $\mathbb{Z}$ und dann mit Hilfe der ganzen Zahlen die rationalen Zahlen $\mathbb{Q}$ ein. Dabei verwenden wir die eleganten Methoden der Algebra, ohne uns dabei in einem allzu theoretischen und abstrakten Umfeld zu verlieren. Am Ende der Darstellung fassen wir die „üblichen Rechenregeln“ zum Begriff eines Körpers zusammen und weiter die „üblichen Ordnungseigenschaften“ zum Begriff eines angeordneten Körpers.

Die Frage, ob wir mit den rationalen Zahlen ein „Kontinuum“ konstruiert haben, beantworten wir im dritten Kapitel negativ, indem wir zeigen, daß die Ordnung von $\mathbb{Q}$ eher durch Lücken als durch Vollständigkeit glänzt. Im Zentrum stehen hier die Begriffe des Supremums und des Infimums, mit deren Hilfe wir die Lücken von $\mathbb{Q}$ identifizieren können. Mit Hilfe von Dedekindschen Schnitten konstruieren wir dann die reellen Zahlen $\mathbb{R}$. Wir zeigen die archimedische Ordnung unserer Konstruktion, die die Existenz von infinitesimalen Größen ausschließt, und wir skizzieren den Beweis einer algebraischen Charakterisierung des angeordneten Körpers der reellen Zahlen. Anschließend betrachten wir die natürliche Frage, ob wir auf $\mathbb{R}^2$, $\mathbb{R}^3$, … eine Multiplikation einführen können, die zusammen mit der punktweisen Addition eine Körperstruktur erzeugt. Diese Frage führt uns zu den komplexen Zahlen $\mathbb{C}$ und weiter zu einer knappen Darstellung der Hamiltonschen Quaternionen $\mathbb{H}$.

Dritter Abschnitt: Erste Erkundungen

Der letzte Abschnitt des Buches besteht aus sechs Kapiteln, die jeweils eine elementare, aber aus der Perspektive des Anfängers durchaus auch anspruchsvolle Einführung in einen mathematischen Themenbereich geben. Die Auswahl der sechs Gebiete und ihre inhaltliche Ausgestaltung ist dabei subjektiv, und sie erhebt keinerlei Anspruch auf eine vollständige Abdeckung des Wichtigsten.

Am Anfang steht eine Untersuchung des Teilbarkeitsbegriffs der natürlichen Zahlen. Wir beweisen die elementaren Eigenschaften dieses Begriffs und untersuchen den größten gemeinsamen Teiler und das kleinste gemeinsame Vielfache zweier Zahlen. (Das ist alles viel spannender, als man es von der Schule vielleicht in Erinnerung hat.) Die Frage nach einer effektiven Berechnung des größten gemeinsamen Teilers führt uns zum Euklidischen Algorithmus, einem einzigartig schönen und fruchtbaren Juwel der Mathematik. Er liefert uns neue Einsichten über den größten gemeinsamen Teiler zweier Zahlen, indem er diesen in effektiver Weise als Linearkombination der Zahlen darzustellen gestattet. Danach führt uns ein „Filmschnitt“ zu den Atomen der Teilbarkeitsrelation, den Primzahlen. Wir formulieren eine Reihe von Fragen, die sich bei der Betrachtung des Primzahl-Begriffs ergeben, und von denen einige bis heute ungelöst geblieben sind. Nachdem wir die Unendlichkeit der Primzahlen gezeigt haben, geben wir einen klassischen und einen modernen Beweis für den Hauptsatz der elementaren Zahlentheorie, die Eindeutigkeit der Primfaktorzerlegung. Als Anwendung diskutieren wir das klassische Argument für die Existenz von irrationalen Zahlen, das die Darstellung der Lücken von $\mathbb{Q}$ im zweiten Abschnitt ergänzt.

Von der Zahlentheorie springen wir zu den Grundlagen der Analysis und studieren im zweiten Kapitel den Grenzwertbegriff. Wir betrachten konvergente Folgen und Reihen sowie Häufungspunkte von Mengen. Anschließend diskutieren wir verschiedene äquivalente Fassungen des Stetigkeitsbegriffs und beweisen die wichtigsten elementaren Sätze über stetige Funktionen. Als Ausblick führen wir den Begriff der offenen Menge ein und formulieren mit seiner Hilfe die Stetigkeit einer Funktion. Schließlich spannen wir einen Bogen zum zweiten Abschnitt, indem wir den metrischen Vollständigkeitsbegriff in Beziehung zur Existenz von Suprema und Infima bringen.

Kapitel drei beschäftigt sich mit Matrizen, rechteckigen Gebilden aus Zahlen. Wir motivieren sie als kompakte Notationen für lineare Gleichungssysteme. Nach einem Blick auf den Lösungsraum eines derartigen Systems stellen wir den Gauß-Jordanschen Algorithmus vor, der unsere Systeme in effektiver und befriedigender Weise mit Hilfe von elementaren Matrizenumformungen löst. Anschließend betrachten wir den Zusammenhang zwischen Matrizen und linearen Abbildungen und gelangen so zur Definition einer Matrizenmultiplikation. Zu einer ganz anderen Anwendung der Matrizen führt uns dann die Frage, wie wir den transitiven Abschluß einer endlichen Relation effektiv berechnen können. Hierzu führen wir eine neue Matrizenmultiplikation ein und diskutieren den Algorithmus von Warshall. Insgesamt ergeben sich drei Aspekte, die die fundamentale Bedeutung des Konzepts illustrieren: Matrizen als Notationen für Gleichungssysteme, Matrizen als lineare Abbildungen, Matrizen als Darstellung von Relationen.

Die Gruppen sind das Thema der vierten Erkundung. Wir motivieren den Begriff sehr knapp mit Hilfe von endlichen Permutationen. (Für den Leser, der den zweiten Abschnitt studiert hat, sind die Struktur-Eigenschaften einer Gruppe wahrlich nichts Neues.) Wir stellen einige Beispiele für Gruppen zusammen und untersuchen Folgerungen aus den Gruppenaxiomen. Danach behandeln wir Untergruppen, Nebenklassen und Faktorgruppen, und wir beweisen den Satz von Lagrange.

Das fünfte Kapitel gehört dem weiten Reich der endlichen Kombinatorik an und bietet eine Einführung in die Graphentheorie. Wir führen zunächst die wichtigsten Begriffe wie „Kantenzug, Weg, Kreis, Zusammenhang“ ein, um über Graphen angemessen reden zu können. Anschließend fragen wir nach der Existenz von Eulerzügen, die es erlauben, einen Graphen in einem Zug zu zeichnen. Mit dem Algorithmus von Hierholzer finden wir ein einfaches Verfahren, das es uns erlaubt, im Falle der Existenz einen Eulerzug zu konstruieren. Nach einer Diskussion der Frage, wie man ein Labyrinth algorithmisch erkundet, wenden wir uns dem viel schwierigeren Problem der Existenz eines Kreises zu, der jede Ecke eines Graphen genau einmal besucht. Wir beweisen den Satz von Dirac, der ein hinreichendes Kriterium für die Existenz eines derartigen sogenannten Hamiltonkreises etabliert.

Das Buch endet mit einer Einführung in den Begriff eines Wahrscheinlichkeitsmaßes. Wir definieren abzählbare Wahrscheinlichkeitsräume und deuten die Probleme an, die der überabzählbare Fall aufwirft. Wir verzichten vollständig auf die Einführung von Dichtefunktionen und führen statt dessen eine Integralschreibweise für diskrete Summen ein, die nicht nur elegant ist, sondern auch auf die Notationen und wichtigsten Integrationssätze der allgemeinen Wahrscheinlichkeitstheorie vorbereitet. (Erfahrungsgemäß ist das allgemeine Integral $\int f \, d\mu$ auch für Fortgeschrittene nicht leicht zu verdauen, sodaß eine Vorbereitung anhand des einfachsten Falls sicher nicht schadet.) Wir verwenden unsere μ-Integrale dann zur Definition des Erwartungswerts und der Varianz einer Zufallsvariablen. Das Kapitel schließt mit einem Beweis des schwachen Gesetzes der großen Zahl und einer elementaren Formulierung seiner starken Version.

Einführung

Die systematisch definierende und beweisende Mathematik beginnt bei den alten Griechen. Im Zentrum stehen ideelle geometrische und arithmetische Objekte. In der Neuzeit setzte sich die Beschreibung der Geometrie mit Hilfe der Zahlen durch, und in der Moderne fand die Mathematik im scheinbar harmlosen Begriff der Menge ihren axiomatischen Grundbegriff, auf den im Verbund mit der mathematischen Logik jedes mathematische Objekt zurückgeführt werden konnte. Auf dieser Grundlage wurden die faszinierenden Landschaften der heutigen Mathematik errichtet. Das historische Erbe wurde vollständig integriert und in atemberaubender Form erweitert und neu interpretiert.

Sich in diesem Land zurechtzufinden ist nicht leicht, denn man spricht dort eine eigene Sprache, die aufgrund ihrer einzigartigen Informationsdichte und Genauigkeit bewundert wird, aber auch als trocken und formal empfunden werden kann. Wenn man darüber nachdenkt, warum die moderne Mathematik so spricht wie sie spricht, lassen sich zwei Leitmotive ausmachen, die man mit „Konsistenz“ und „Weite“ bezeichnen kann.

Konsistenz

Die Mathematik duldet keine inneren Widersprüche. Das war schon immer so, aber der Anspruch an Genauigkeit, Klarheit und Präzision wuchs mit der Mathematik, und in der Moderne sprang dieser Anspruch auf eine neue Stufe. Die Mathematik operiert heute sehr frei mit unendlichen Konstrukten, die wir in der Natur nicht wiederfinden, die sich aber zu ihrer Beschreibung oft überraschend gut eignen. Daß das ideelle Feuer des Unendlichen beherrscht werden muß, zeigen die um 1900 zu datierenden Brandstellen der mengentheoretischen Paradoxien. Die Zähmung des Unendlichen gelang dann in der ersten Hälfte des 20. Jahrhunderts. Wir wissen heute sehr viel über die Grundlagen der Mathematik, über Beweise, Mengen, Relationen, über algorithmisch berechenbare und allgemeine Funktionen und über die beiden Grundstrukturen der natürlichen und der reellen Zahlen. Der Erwerb und die Nutzung dieses Wissens führte zu einer tiefgreifenden Veränderung der bis ins 19. Jahrhundert üblichen mathematischen Sprache. Vergleicht man Lehrbücher und Forschungsarbeiten, die vor und nach diesem Übergang geschrieben wurden, so ist ein enormer Anstieg an Dichte, Formalismus, Komplexität, Symbolik und Systematik festzustellen. Wer anstrebt, Mathematiker zu werden, muß sich dieser Herausforderung stellen. Der Anfänger wird sich, wie ja auch viele Experten, nicht unbedingt mit Grundlagenfragen auseinandersetzen, aber er wird die Sprache der modernen Mathematik lernen wollen.

Weite

Das Motiv der Weite läßt sich auch durch „Anspruch“, „Ziele“, „Niveau“ beschreiben, durch das, wie weit man kommen will. Die Sprache der modernen Mathematik ermöglicht viele Konstruktionen und Ergebnisse, die in früheren Zeiten undenkbar gewesen wären. Wir möchten diesen Gedanken durch drei weit über diesen Text hinausgehende Beispiele illustrieren. Viele andere könnten ebenso gut hier stehen.

Als erstes Beispiel betrachten wir das klassische Problem der Bestimmung von Flächeninhalten. Bereits Archimedes konnte durch Ausschöpfung die Fläche von einfachen geometrischen Figuren berechnen, zu denen auch der Kreis gehörte. Eine allgemeine und flexible Methode stellte dann erst die Differential- und Integralrechnung von Leibniz und Newton zur Verfügung. Die Flächen- und Volumenberechnungen wurden, hatte man die neue Theorie erst einmal verinnerlicht, viel einfacher und auch in komplizierteren Fällen möglich. Weiter zeigte sich, daß das Integral in der Mathematik und in der mathematischen Naturbeschreibung an vielen Stellen eine zentrale Rolle spielt. In Folge des verallgemeinerten Funktionsbegriffs und der Untersuchung von Funktionen durch Fourierreihen wurde schließlich eine präzise Definition und genauere Untersuchung des Integrals notwendig. Dies führte Mitte des 19. Jahrhunderts zum Begriff der Riemann-Integrierbarkeit einer Funktion, und wenig später wurden auch die ersten detaillierten Konstruktionen der reellen Zahlen gegeben. Bereits zu Beginn des 20. Jahrhunderts fand dann Henri Lebesgue ein allgemeineres Integral, das durch seine guten „im höheren mathematischen Alltag“ benötigten Eigenschaften zu überzeugen wußte. Das Lebesgue-Integral wiederum bildete den Ausgangspunkt für das allgemeine maßtheoretische Integral, das heute insbesondere in der Wahrscheinlichkeitstheorie und Funktionalanalysis überall verwendet wird. Der Weg dorthin ist ohne die Beherrschung der mengentheoretischen Sprache und der Zahlen nicht zu gehen, und ohne ein hohes Maß an formaler Exaktheit und Sorgfalt hätte er nicht gegangen werden können.

Ein zweites Beispiel bildet die 1995 von Andrew Wiles bewiesene Fermatsche Vermutung: Für alle natürlichen Zahlen $n \geq 3$ ist die Gleichung $a^n + b^n = c^n$ nicht in natürlichen Zahlen $a, b, c \geq 1$ lösbar. Während etwa $3^2 + 4^2 = 5^2$ und $5^2 + 12^2 = 13^2$ gilt, gibt es also keine $a, b, c \geq 1$ mit $a^3 + b^3 = c^3$, und das gleiche gilt für die Exponenten 4, 5, 6, … Das Problem ist für einige kleine Exponenten noch relativ einfach zu lösen, aber zur Lösung des allgemeinen Falls mußten Methoden der algebraischen Geometrie verwendet und Teilresultate mehrerer Mathematiker zusammengetragen und erweitert werden. Dies ist kein Einzelfall: Die Mathematik setzt heute sehr oft fortgeschrittene analytische oder algebraische Methoden ein, um Licht auf einfache Fragen über die natürlichen Zahlen zu werfen. Auch hier ist ein hohes Maß an Genauigkeit und komplexer Begriffsbildung unerläßlich, denn nur dadurch werden die Beweise überschau- und überprüfbar.

Zuletzt sei mit dem zweiten Unvollständigkeitssatz von Kurt Gödel aus dem Jahr 1931 ein auch aus erkenntnistheoretischer Sicht tiefes mathematisches Ergebnis genannt. Grob formuliert besagt dieser Satz, daß eine hinreichend starke axiomatische Theorie ihre Widerspruchsfreiheit nicht beweisen kann (es sei

denn, sie ist widerspruchsvoll, denn dann kann sie alles beweisen). Insbesondere können wir die Widerspruchsfreiheit der Zahlentheorie und damit auch der viel stärkeren mengentheoretisch fundierten Mathematik nicht mit endlichen kombinatorischen Methoden beweisen – wozu David Hilbert um 1920 im „Hilbertschen Programm" aufgerufen hatte. Um einen derartigen Satz überhaupt erst formulieren und weiter dann beweisen zu können, ist eine mathematische Definition von „axiomatische Theorie", „Beweis" und „Widerspruchsfreiheit" unerläßlich, d. h. es muß möglich sein, die Mathematik als formales System aufzufassen, das mit ganz im Endlichen verbleibenden Argumenten untersucht werden kann. Das Ergebnis von Gödel versieht dann das Motiv der Konsistenz mit einem bemerkenswerten Zusatz: Die Mathematik duldet keine inneren Widersprüche, aber sie kann nicht beweisen, daß keine Widersprüche vorhanden sind.

✳

Diesem „Ausblick vorab" sei noch hinzugefügt, daß der Eingangsbereich der wissenschaftlichen Mathematik nicht nur die Grundlage für alles Weitere ist, sondern auch selbst bereits viel Interessantes zu bieten hat!

1. Abschnitt

Die Sprache der Mathematik

1. Mathematisches Argumentieren

In der Mathematik wird eine modifizierte Form der Umgangssprache verwendet. Diese Sprache ist einerseits sehr karg, andererseits wird sie ständig durch neue und oft sehr komplexe Begriffe angereichert, deren Bedeutung man erst erlernen muß. Wir betrachten im folgenden das Grundgerüst der mathematischen Sprache, das dann mit hierarchisch aufeinander aufbauenden mathematischen Begriffsbildungen zum Leben erweckt wird.

Zunächst werfen wir einen Blick auf logische Verbindungen von Aussagen und diskutieren Wahrheitstafeln und einige Grundmuster der mathematischen Beweisführung. Danach betrachten wir Quantoren sowie die inhaltlichen Bestandteile mathematischer Aussagen und gelangen so zu einer relativ genauen Beschreibung der heutigen mathematischen Sprache. Schließlich stellen wir noch einige Tautologien und Quantorenregeln tabellarisch zusammen.

Aussagen und Junktoren

Die Mathematik betrachtet Aussagen wie etwa

„Die Zahl 5 ist eine Primzahl.“,

„Wenn n kleiner als m ist, so ist stets auch n + 1 kleiner als m + 1.“,

„Die Sinus-Funktion ist periodisch und beschränkt.“,

„Nicht jede stetige Funktion auf den reellen Zahlen ist differenzierbar.“.

Auch „Die Zahl 4 ist eine Primzahl“ ist eine mathematische Aussage, nicht aber „Die Zahl 4 ist männlich“ oder „Diese Banane ist noch nicht reif.“ Eine genaue Definition dessen, was eine sinnvolle mathematische Aussage ist und was nicht, ist möglich, liegt aber außerhalb der Intentionen dieser Einführung in das mathematische Argumentieren. Ein naives Grundverständnis, welches z. B. natürliche Zahlen deutlich von Bananen trennt, genügt, um anzufangen.

In obigen Beispielen ist von einer *Konstanten* wie der Zahl 5, einer *Funktion* wie dem Sinus und einer *Relation* wie „kleiner als“ die Rede. Bevor wir aber diese inhaltlichen Elemente mathematischer Aussagen genauer betrachten, wenden wir uns dem einfacheren Thema der logischen Junktoren zu, die beliebige Aussagen A, B, C, D, … mit einander verknüpfen. Die für die Mathematik wichtigsten Junktoren sind:

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_1, © Springer-Verlag Berlin Heidelberg 2010

(i) die Negation *nicht/non*, in Zeichen: $\neg$,

(ii) die Konjunktion *und*, in Zeichen: $\wedge$,

(iii) die Disjunktion *oder*, in Zeichen: $\vee$,

(iv) die Implikation *folgt/impliziert*, in Zeichen: $\rightarrow$,

(v) die Äquivalenz *genau dann, wenn*, in Zeichen: $\leftrightarrow$.

Viele andere Verknüpfungen sind denkbar, zum Beispiel ein „weder noch". Überwiegend werden aber die obigen Junktoren eingesetzt und andere Junktoren werden mit ihrer Hilfe ausgedrückt. Zeichen für andere Verknüpfungen haben keinen allgemeinen Bekanntheitsgrad erreicht.

Verknüpfungen, die wir hier nicht betrachten, sind etwa „A kommt vor B" oder „A ist möglich", „A ist notwendig". Man kann solche Verknüpfungen innerhalb der sogenannten nichtklassischen Aussagenlogik studieren. In der allgemeinen Mathematik kommen sie nicht vor.

Mit Hilfe der logischen Verknüpfungen werden aus Aussagen neue Aussagen gebildet. Für jede Aussage A ist die Negation von A, also $\neg$ A, eine neue Aussage. Die Negation ist, wie man sagt, ein einstelliger Junktor. Die anderen Junktoren der obigen Tabelle sind zweistellig: Für alle Aussagen A und B sind die Konjunktion $A \wedge B$, die Disjunktion $A \vee B$, die Implikation $A \rightarrow B$, und die Äquivalenz $A \leftrightarrow B$ neue Aussagen.

Da man die Junktoren iteriert anwenden kann, ist es erforderlich, Klammern zu setzen. Dadurch wird zum Beispiel die Aussage $(A \wedge B) \rightarrow C$ von der Aussage $A \wedge (B \rightarrow C)$ unterscheidbar.

Um Klammern zu sparen, wird folgende Bindungsstärke der Junktoren vereinbart, von stark bindend zu schwach bindend:

$\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$. *(Bindungsstärke der Junktoren)*

Damit ist zum Beispiel

$\neg A \wedge B \rightarrow C$ die Aussage $((\neg A) \wedge B) \rightarrow C$,

und nicht etwa die Aussage $\neg (A \wedge (B \rightarrow C))$. Die vereinbarte Bindungsstärke ist nichts weiter als eine Konvention, die sich in der Praxis bewährt hat.

Zur Verbesserung der Lesbarkeit kann man die Struktur eines Ausdrucks auch durch graphische Mittel wie Abstände und Kursivstellungen verdeutlichen. So schreiben wir zum Beispiel:

A oder B *impliziert* C.

Das längliche *genau dann, wenn* kürzen wir gerne durch *gdw* ab. (Im Englischen hat sich hier das Kunstwort *iff* für *if and only if* durchgesetzt.)

Statt „A impliziert B" sagen wir gleichwertig auch „(aus) A folgt B", „A zieht B nach sich", „wenn A, so auch B", „A ist hinreichend für B" oder „B ist notwendig für A". Hinter den Sprechweisen „hinreichend" und „notwendig" verbirgt also nichts weiter als eine Implikation, die in zwei verschiedenen Richtungen gelesen wird.

Semantik der Junktoren

In den folgenden Abschnitten diskutieren wir die Bedeutung (Semantik) und Verwendung der Junktoren unter verschiedenen Gesichtspunkten:

(a) Junktoren in der Umgangssprache und in der Mathematik.

(b) Hierarchischer Aufbau der Junktoren.

(c) Junktoren als Operatoren auf Wahrheitswerten.

(d) Junktoren in mathematischen Beweisen.

Junktoren in der Umgangssprache und in der Mathematik

Feinsinnige umgangssprachliche Schattierungen der Junktoren fallen in der Mathematik weg. Ein Beispiel ist: „Ich werde krank und gehe ins Krankenhaus" (um gesund zu werden), im Gegensatz zu „Ich gehe ins Krankenhaus und werde krank" (weil ich mich angesteckt habe). In der Mathematik ist „A und B" stets gleichwertig zu „B und A".

Ebenso ist „A oder B" mathematisch immer gleichwertig zu „B oder A". Umgangssprachlich ist dagegen der Satz „Verlassen Sie mein Haus oder ich rufe die Polizei." vollkommen natürlich, während „Ich rufe die Polizei oder Sie verlassen mein Haus." befremdlich wirkt, und allenfalls dazu verwendet werden kann, um dem Dieb mitzuteilen, daß er bei einem Mathematiker eingebrochen ist.

Ein logisch korrektes „oder" kann in der Umgangssprache unangemessen oder unhöflich sein: „Kommst Du oder Deine Frau zu meiner Feier?" fragt man nicht, wenn man beide eingeladen hat. Ebenso antwortet man auf die Frage „Willst du hierbleiben oder nach Hause gehen?" nicht mit einem logisch korrekten „Ja", sondern man sagt, was man möchte.

Das mathematische „oder" wird durchgehend in einem nicht ausschließlichen Sinne verwendet, ist also von einem „entweder … oder" zu unterscheiden. Das umgangssprachliche „oder" wird dagegen manchmal ausschließlich verwendet und manchmal auch nicht. Wenn die Mutter ihr Kind fragt: „Willst du nun den gelben oder den roten Ball?", so ist „Beide!" als Antwort zwar alles andere als unwahrscheinlich, entspricht aber nicht der Intention des Fragenden. Dagegen läßt die Frage „War Alkohol oder Übermüdung der Grund für den Unfall?" die Antwort „Wohl beides!" zu, die vom Sprecher vielleicht sogar erwartet wird.

Auf der anderen Seite gibt es auch Wendungen der Umgangssprache, die den mathematischen Gebrauch einer logischen Verknüpfung illustrieren. Der Satz: „Wenn Du ein guter Koch bist, bin ich der Kaiser von China." ist eine witzige Variante von „Du bist ein miserabler Koch." Eine solche Behauptung kann Diskussionen auslösen, dagegen wird man in der Mathematik keinen Widerspruch ernten, wenn man behauptet: „Wenn 4 ungerade ist, so ist 0 = 1."

Hierarchischer Aufbau der Junktoren

Die Aufgabe, die Semantik der mathematischen Junktoren zu präzisieren, können wir durch einen hierarchischen Aufbau reduzieren, bei dem lediglich die Negation und die Konjunktion als Grundverknüpfungen angesehen werden. Die anderen Junktoren können dann nämlich wie folgt eingeführt werden:

(a) $A \vee B$ *wird definiert als* $\neg(\neg A \wedge \neg B)$,

(b) $A \to B$ *wird definiert als* $\neg A \vee B$,

(c) $A \leftrightarrow B$ *wird definiert als* $(A \to B) \wedge (B \to A)$.

Es bleibt nun nur noch die Aufgabe, die Basisjunktoren „und" und „non" zu präzisieren. Die Konjunktion bereitet kaum Probleme: „A und B" bedeutet „sowohl A, als auch B". Zur Semantik der Negation bemerken wir, daß wir sie im „klassischen" Sinne verstehen, d.h. eine doppelte Negation „non non A" ist stets gleichwertig zu A. Als Regel: Wir dürfen doppelte Negationen streichen. Viel mehr kann man an dieser Stelle nicht sagen. Weitere Einsichten in die Semantik der Konjunktion und der Negation werden wir aber unten bei der Diskussion der Wahrheitstafeln und der Verwendung der Junktoren in mathematischen Beweisen gewinnen.

Die Definition der Junktoren „oder", „impliziert" und „genau dann, wenn" ist ein Beispiel für zwei tragende Prinzipien wissenschaftlicher Vorgehensweise:

Reduziere Probleme so weit wie möglich.

Definiere Neues unter Verwendung des Alten.

Die erste Maxime befreit uns von unnötigen Komplexitäten und erlaubt uns in vielen Fällen, die Dinge klarer zu sehen. Die zweite Maxime führt darüber hinaus zu einem Begriffsgebäude, das auf möglichst wenigen Grundbegriffen ruht. In höchster Strenge und Vollkommenheit ist die Befolgung der zweiten Maxime nur in der Mathematik möglich – und nötig.

Wie in vielen anderen Situationen gibt es auch im vorliegenden Fall verschiedene Möglichkeiten der Problemreduktion. Wir hätten auch „oder" und „non" als Basisjunktoren betrachten können (und wir hätten dann $A \wedge B$ definiert als $\neg(\neg A \vee \neg B)$). Weiter gibt es einen interessanten hierarchischen Aufbau der Junktoren, der auf den Basisjunktoren „und" und „impliziert" beruht und also die Negation als einen definierten Junktor ansieht. Hierzu führen wir ein spezielles Aussagensymbol $\bot$ ein, das sog. *Falsum*. Semantisch steht das Falsum für eine Aussage wie $0 = 1$. Nun definieren wir:

$\neg A$ *als* $A \to \bot$,

und $A \vee B$ und $A \leftrightarrow B$ wie oben. Damit sind dann alle Junktoren auf die Junktoren $\wedge$, $\to$ und das Aussagensymbol $\bot$ zurückgeführt. Wir kommen auf diesen Ansatz, der die für das mathematische Beweisen so zentrale Implikation an die Spitze stellt, noch zurück.

Junktoren und Wahrheitswerte

Dem Leser wird aufgefallen sein, daß wir bislang nicht von der „Wahrheit“ oder „Falschheit“ einer Aussage gesprochen haben, und auch nicht von ihrer „Gültigkeit“ oder „Ungültigkeit“. Er wird vielleicht semantische Erklärungen der Form

„$A \wedge B$ ist genau dann wahr, wenn sowohl A als auch B wahr ist.“,

„$A \to B$ ist genau dann falsch, wenn A wahr und B falsch ist.“

erwartet haben. Der Grund für das Fehlen solcher Wendungen ist, daß „wahr“ und „falsch“ selbst eine sehr schwierige Semantik haben und damit keine Patentlösung sind, die mathematischen Junktoren zu erläutern. Sicherlich ist gegen eine Behauptung wie „Die Aussage ‚5 ist eine gerade Zahl.‘ ist falsch.“ nicht viel einzuwenden, denn man kann diese Behauptung einfach als die einfachere Behauptung lesen, daß 5 eine ungerade Zahl ist. Problematisch ist nun aber die Sicht, daß jede mathematische Aussage wahr oder falsch ist. Denn man darf fragen: Ist mathematische Wahrheit mehr als Beweisbarkeit? Wer hier mit „ja“ antwortet, muß auf die Nachfrage, was dieses Mehr genau ist, eine Antwort bereithalten und hier gelangen wir dann sehr schnell von der Mathematik zur Philosophie der Mathematik. Wer aber mit „nein“ antwortet, muß die Ansicht, daß jede mathematische Aussage wahr oder falsch ist, aufgeben, denn es gibt mathematische Aussagen, die sich weder beweisen noch widerlegen lassen. (Das ist das Thema der berühmten Gödelschen Unvollständigkeitssätze.)

Wir wollen hier die Wahrheitsdiskussion gar nicht erst eröffnen. Das ist auch gar nicht notwendig, und zudem fehlt uns an dieser Stelle ohnehin die für eine derartige Diskussion notwendige mathematische Erfahrung und ein Einblick in das mathematische Beweisen. Statt den Boden der Mathematik zu verlassen, wollen wir versuchen, die Mathematik zu finden, die in den vertrauten Sprechweisen „wahr“ und „falsch“ steckt. Wir betrachten also die Junktoren selbst als gewisse Gegenstände der Mathematik und nicht als Bestandteile ihrer Sprache. Anschließend können wir dann fragen, was sich von unseren Untersuchungen auf die mathematische Sprache übertragen lässt.

Die „Mathematisierung“ der Junktoren verläuft wie folgt: Wir arbeiten mit zwei prinzipiell beliebigen sog. *Wahrheitswerten* „w“ und „f“. Liegt nun eine mit Hilfe von Junktorenzeichen $\neg$, $\wedge$, …, Aussagensymbolen $A, B, C, \ldots$ und Klammern gebildete Aussage vor (eine bestimmte Zeichenkette wie $\neg A \wedge (B \to C)$), so weisen wir den darin enthaltenen Aussagensymbolen Wahrheitswerte „w“ oder „f“ zu, z. B. A = „w“, B = „f“, C = „f“, usw., und errechnen dann eindeutig den Wahrheitswert „w“ oder „f“ der zusammengesetzten Aussage. Um diese Berechnung durchführen zu können, müssen wir nur angeben, wie die Junktoren auf den Wahrheitswerten „w“ und „f“ operieren. Dies geschieht durch Angabe von sog. *Wahrheitstafeln* für die Junktoren. Wir wählen $\wedge$ und $\neg$ als Basisjunktoren und legen folgende Wahrheitstafeln für diese Junktoren fest:

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

$\neg$	A
f	w
w	f

Die dritte Zeile der $\wedge$-Tafel besagt z. B.: Hat A den Wert „f“ und B den Wert „w“, so hat die Aussage $A \wedge B$ den Wahrheitswert „f“. Die Tafel für die Negation kann man so zusammenfassen: Der Wahrheitswert von $\neg A$ ist immer der entgegengesetzte Wahrheitswert von A. Folglich ist z. B. der Wahrheitswert von $\neg\neg A$ stets gleich dem Wahrheitswert von A. Das Streichen der doppelten Negation steckt hier also in der Wahrheitstafel der Negation.

Aus der Definition von $A \vee B$ als $\neg(\neg A \wedge \neg B)$ errechnet sich dann die folgende Wahrheitstafel für die Disjunktion:

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

$\neg$	$(\neg$	A	$\wedge$	$\neg$	B)
w	f	w	f	f	w
w	f	w	f	w	f
w	w	f	f	f	w
f	w	f	w	w	f
4	1		3	2	

Die Ziffern unter den Spalten der rechten Tafel geben die Reihenfolge an, in der die Spalten gefüllt werden. Sie ergeben sich aus dem Aufbau der Aussage. Die letzte gefüllte Spalte ist die *Ergebnisspalte* der Wahrheitstafel.

Ebenso erhalten wir die folgende Wahrheitstafel für die Implikation:

A	$\rightarrow$	B
w	w	w
w	f	f
f	w	w
f	w	f

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

Der Leser wird analog die vier Zeilen www, wff, ffw, fwf für die Wahrheitstafel der Äquivalenz $A \leftrightarrow B$ finden.

Als ein etwas komplexeres Beispiel berechnen wir noch die Wahrheitstafel des sog. *Kontrapositionsgesetzes*, also der Aussage $(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$:

A	→	B	↔	¬	B	→	¬	A
w	w	w	w	f	w	w	f	w
w	f	f	w	w	f	f	f	w
f	w	w	w	f	w	w	w	f
f	w	f	w	w	f	w	w	f
	1		5	2		4	3	

Ist die Ergebnisspalte einer Wahrheitstafel durchgehend mit „w“ gefüllt, so nennt man die betrachtete Aussage *allgemeingültig* oder eine *Tautologie*. Das Kontrapositionsgesetz ist also ein Beispiel für eine Tautologie. Einfachere Beispiele sind die Tautologien $A \to A$, $A \vee \neg A$, $A \leftrightarrow A$, $A \leftrightarrow \neg\neg A$.

Wahrheitstafeln für Aussagen zu erstellen kann schnell sehr aufwendig werden. Ist eine Aussage aus $A_1, \ldots, A_n$ aufgebaut, so besteht die zugehörige Wahrheitstafel aus 2^n Zeilen, da alle möglichen w-f-Kombinationen für die Aussagensymbole $A_1, \ldots, A_n$ berücksichtigt werden müssen.

Zur Klärung der Bedeutung des Wahrheitstafelverfahrens betrachten wir die logischen Verknüpfungen unter einem weiteren Gesichtspunkt:

Junktoren in mathematischen Beweisen

In mathematischen Beweisen werden bestimmte Schlußregeln verwendet, und diese Schlußregeln werfen vielleicht das beste Licht auf die Junktoren. Wir betrachten hierzu obigen Ansatz mit $\wedge$ und $\to$ als Basisjunktoren sowie einem speziellen Aussagensymbol, dem Falsum $\bot$, das insbesondere der Definition von $\neg A$ als $A \to \bot$ dient. Beweise entstehen, indem wir bestimmte Aussagen als Annahmen ansehen und dann wiederholt Schlußregeln auf diese Annahmen anwenden. Eine Annahme A gilt als Beweis von A unter der Annahme A. Was darüber hinaus als Beweis gilt, wird durch folgende Schlußregeln beschrieben:

Definition *(aussagenlogische Schlußregeln, Gentzen-Kalkül)*

(S1) Hat man $A \wedge B$ bewiesen, so hat man auch A bewiesen.

(S2) Hat man $A \wedge B$ bewiesen, so hat man auch B bewiesen.

(S3) Hat man in einem Beweis A bewiesen und in einem zweiten Beweis B, so ergeben beide Beweise zusammen einen Beweis von $A \wedge B$.

(S4) Hat man in einem Beweis $A \to B$ bewiesen und in einem zweiten Beweis A, so ergeben beide Beweise zusammen einen Beweis von B.

(S5) Hat man A bewiesen, so hat man $B \to A$ bewiesen für jede beliebige Aussage B, und der Beweis von $B \to A$ hängt nicht von der Annahme von B ab.

(S6) Hat man $\bot$ bewiesen, so hat man jede beliebige Aussage bewiesen.

(S7) Hat man $\neg\neg A$ bewiesen, so hat man A bewiesen.

Die Regel (S4) ist scholastisch auch als *modus ponens* bekannt, (S6) als *ex falso quodlibet* und (S7) als *duplex negatio affirmat*. Weitere Schlußregeln muß man nicht zulassen, alle anderen mathematischen Argumente wie z. B. die Fallunterscheidung lassen sich aus diesen Regeln gewinnen.

In unseren Schlußregeln ist von „wahr“ und „falsch“ nicht einmal mehr symbolisch die Rede. Es geht nur noch um Beweisbarkeit.

Während einer Beweisführung darf man jederzeit beliebige Annahmen machen. Einzig die Regel (S5) erlaubt es dann, sich von einer Annahme auch wieder zu befreien. Die Regel (S5) wird deswegen auch als *Abbinden* von Annahmen bezeichnet. Bei allen anderen Schlußregeln bleibt die Abhängigkeit der bewiesenen Aussage von den bislang gemachten Annahmen erhalten.

Wir beweisen zur Illustration der Schlußregeln die Implikation von „links nach rechts“ (die sog. *Hin-Richtung*) im Kontrapositionsgesetz, also die Aussage $(A \to B) \to (\neg B \to \neg A)$. Hierbei sind A und B beliebige Aussagen.

Beweis von $(A \to B) \to (\neg B \to \neg A)$ mit Hilfe der Schlußregeln

Wir nehmen $A \to B$ an. Unser Ziel ist $\neg B \to \neg A$, also $(B \to \bot) \to (A \to \bot)$. Hierzu nehmen wir $B \to \bot$ an und wollen nun $A \to \bot$ zeigen. Hierzu wiederum nehmen wir A an und streben nun einen Beweis von $\bot$ an. Zusammenfassend haben wir also die drei Aussagen

$A \to B,\ B \to \bot,\ A$

angenommen und unser Beweisziel ist $\bot$. Aus $A \to B$ und A erhalten wir B mit modus ponens. Die Annahme $B \to \bot$ und erneute Anwendung von modus ponens liefert $\bot$. Nun wenden wir dreimal hintereinander die Abbindungsregel (S5) an und erhalten der Reihe nach

(1) $A \to \bot$, also $\neg A$, *(Abbinden der Annahme A)*,

(2) $\neg B \to \neg A$, *(Abbinden der Annahme $B \to \bot = \neg B$)*,

(3) $(A \to B) \to (\neg B \to \neg A)$. *(Abbinden der Annahme $A \to B$)*.

In (1) haben wir $\neg A$ bewiesen unter Annahme von $A \to B$ und $B \to \bot$. In (2) haben wir $\neg B \to \neg A$ bewiesen unter Annahme von $A \to B$. In (3) haben wir schließlich wie gewünscht die Aussage $(A \to B) \to (\neg B \to \neg A)$ bewiesen, und der Beweis hängt von keiner Annahme mehr ab.

Dieses Argument können wir kompakt als *Beweisbaum* notieren, in dem Anwendungen von Schlußregeln durch waagrechte Striche dargestellt werden:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \to B\ ^{①} \qquad A\ ^{②}}{B} \qquad B \to \bot\ ^{③}}{\bot}}{A \to \bot}}{(B \to \bot) \to (A \to \bot)}}{(A \to B) \to ((B \to \bot) \to (A \to \bot))}$$

(S4)

(S4)

(S5: Abbinden von Annahme ②)

(S5: Abbinden von Annahme ③)

(S5: Abbinden von Annahme ①)

Natürlich werden Beweise normalerweise nicht in dieser Ausführlichkeit geführt. Häufig auftauchende Argumentationsmuster werden in kompakter Form verwendet, und bereits bewiesene Resultate können als Hilfssätze verwendet werden. Prinzipiell haben wir für die Aussagenlogik aber nicht mehr zur Verfügung als obige elementare Schlußregeln.

Nach diesen Ausführungen können wir nun genau angeben, was das Wahrheitstafelverfahren leistet. Es gilt, wie man zeigen kann, der folgende Satz:

Satz *(Argumentative Beweisbarkeit der Tautologien)*
Für jede Aussage sind gleichwertig:

(i) Das Wahrheitstafelverfahren zeigt, daß die Aussage eine Tautologie ist.

(ii) Es gibt einen (von keinen Annahmen abhängigen) Beweis der Aussage mit Hilfe der Schlußregeln (S1) – (S7).

Ein argumentativer Beweis einer Aussage führt in vielen Fällen zu einem besseren Verständnis als das mechanische Wahrheitstafelverfahren. So zeigt zum Beispiel obige Argumentation, daß die *ex falso quodlibet* und die *duplex negatio affirmat* Regel in einem Beweis der Hin-Richtung nicht verwendet werden müssen. Für die Rück-Richtung $(\neg B \to \neg A) \to (A \to B)$ wird dann aber die Regel (S7) gebraucht. Damit bringt diese Analyse eine Asymmetrie im Kontrapositionsgesetz ans Licht: Die Rück-Richtung ist logisch komplizierter als die Hin-Richtung.

Die Implikation an die Spitze zu stellen und mit Hilfe von Schlußregeln zu beschreiben entspricht nicht zuletzt auch der mathematischen Erfahrung. Die Tautologie $A \to B \leftrightarrow (\neg A \vee B)$ ist zwar vielfach nützlich, jedoch wird die Implikation $A \to B$ von vielen Mathematikern nicht als ein statisches $\neg A \vee B$ empfunden, sondern als eine dynamische Beziehung zwischen A und B. Bei einem Beweis von $A \to B$ werden oft bestimmte innere Vorstellungen über den Gehalt von A in solche über den Gehalt von B übergeführt, und dieser Vorgang ist dynamisch und verläuft in der Regel über mehrere Zwischenstufen. Zu klären, welche Implikationsverhältnisse zwischen verschiedenen Begriffen bestehen, ist ein Grundmuster der mathematischen Tätigkeit und führt oft zu neuen Begriffsbildungen und Fragen. Die bewiesenen Implikationen dienen weiter dann dazu, in komplexeren Beweisen schnell voranzukommen: Hat man A bewiesen und weiß man aus einem früheren Beweis, daß $A \to B$ gilt, so hat man nach der modus ponens Schlußregel B bewiesen. Die Implikation darf damit als die Königin unter den mathematischen Junktoren gelten. Zugleich ist die Schlußregel des modus ponens die wichtigste Regel, die uns von bereits bewiesenen Aussagen zu neuen bewiesenen Aussagen führt.

Aussagenlogische Beweismuster

Viele häufig zu findende Beweisstrukturen der Mathematik beruhen auf aussagenlogischen Äquivalenzen. Wir wollen einige von ihnen zusammenstellen.

Zunächst betrachten wir Beweismuster, die die Implikation betreffen. Dabei verwenden wir die Gleichwertigkeit von $\neg A$ und $A \to \bot$. Bei unserem $\wedge$, $\to$, $\bot$-Ansatz gilt sie per Definition, und bei der Wahl von $\wedge$ und $\neg$ als Basisjunktoren ist $\neg A \leftrightarrow A \to \bot$ eine Tautologie, wobei $\bot$ definiert wird als $A_0 \wedge \neg A_0$ für eine beliebige Aussage A_0.

Die Äquivalenz von $\neg A$ und $A \to \bot$ wird in der Mathematik für die berühmtberüchtigten *Widerspruchsbeweise* verwendet, die in gelehrten Kreisen auch als Beweise des Typs *reductio ad absurdum* bekannt sind. Sollen wir eine Aussage B beweisen, so können wir wie folgt vorgehen. Wir nehmen $\neg B$ an und argumentieren dann solange, bis wir eine offenbar falsche Aussage abgeleitet haben. Damit haben wir dann $\neg B \to \bot$, also $\neg\neg B$ gezeigt. Streichen der doppelten Negation liefert nun wie gewünscht B. Wir halten fest:

Struktur eines Widerspruchsbeweises einer Aussage A

Annahme, es gilt non A. … … … Also gilt $\bot$, *Widerspruch!*

Die Pünktchen stehen hier und im folgenden für das eigentliche Argument.

Ist die Aussage, die wir beweisen wollen, von der Form $\neg A$, so nehmen wir natürlich nicht $\neg\neg A$ an, sondern gleich A selbst. Wir zeigen dann $\bot$, und haben also $A \to \bot$ und damit wie gewünscht $\neg A$ gezeigt. Hier kann man sich also das Streichen der doppelten Negation sparen.

Eng mit den Widerspruchsbeweisen verwandt sind die sog. *indirekten Beweise*, die auf dem Kontrapositionsgesetz ruhen. Das Prinzip des indirekten Beweisens besteht darin, anstelle von $A \to B$ zu zeigen, daß $\neg B \to \neg A$ gilt:

Struktur eines indirekten Beweises einer Implikation „aus A folgt B“

Es gelte non B. … … … Also gilt non A.

Indirekte Beweise werden in der Mathematik sehr häufig verwendet. Ob einem die Beweisstruktur $A \to B$ oder aber $\neg B \to \neg A$ angemessener, klarer, oder einfacher erscheint, muß man von Fall zu Fall entscheiden, und oft ist es auch eine Frage des Geschmacks. Speziell gilt dies für Implikationen der Form $A \to \neg B$, da dann die Kontraposition gleichwertig ist zu $B \to \neg A$.

Ein weiteres wichtiges Beweismuster ist die *Fallunterscheidung*. Zu zeigen ist eine Aussage A. Dies kann dadurch erreicht werden, daß man eine geeignete Aussage B findet, für die man sowohl $B \to A$ als auch $\neg B \to A$ beweisen kann. Dann hat man A bewiesen, denn für alle B ist $A \leftrightarrow (B \to A) \wedge (\neg B \to A)$ eine Tautologie. Wir halten wieder fest:

Struktur eines Beweises einer Aussage A durch Fallunterscheidung

1. Fall: Es gelte B. … … … Also gilt A.

2. Fall: Es gelte non B. … … … Also gilt A.

Auch hier hat man sich also eine zusätzliche Voraussetzung verschafft, allerdings auf Kosten eines zweigeteilten Arguments. Das Finden einer geeigneten

Aussage B ist eine Frage der mathematischen Erfahrung und Kreativität. Auch Varianten können hilfreich sein, zum Beispiel eine Aufspaltung des zweiten Falls in zwei Unterfälle.

Schließlich gibt es noch das Muster des ökonomischen Beweises einer Äquivalenzenkette. Zu zeigen ist, daß Aussagen $A_1, \ldots, A_n$ paarweise äquivalent sind. D. h. zu zeigen ist, daß $A_i \leftrightarrow A_j$ für alle $i \neq j$ gilt. Hierzu kann man wie folgt zyklisch vorgehen:

Struktur eines zyklischen Beweises der Äquivalenz von $A_1, \ldots, A_n$

A_1 folgt A_2: Es gelte A_1. … … … Also gilt A_2.
A_2 folgt A_3: Es gelte A_2. … … … Also gilt A_3.
…
A_n folgt A_1: Es gelte A_n. … … … Also gilt A_1.

Hinter diesem Beweismuster steckt die sog. *Transitivität* der Implikation, die auch als *Kettenschluß* bezeichnet wird: Gilt $A \to B$ und $B \to C$, so gilt auch $A \to C$.

Es kann bei diesem Vorgehen sehr nützlich sein, die Aussagen $A_1, \ldots, A_n$ in einer geeigneten Reihenfolge so anzuordnen, daß die zu zeigenden Implikationen sich natürlich auseinander ergeben und dadurch der Beweis möglichst einfach und kurz wird. Wie man eine solche gute Anordnung findet, ist eine Frage, die sich oft nur experimentell beantworten lässt.

Die paarweise Äquivalenz von drei Aussagen A_1, A_2, A_3 ist gleichwertig zu $(A_1 \leftrightarrow A_2) \wedge (A_2 \leftrightarrow A_3)$, nicht aber zu $(A_1 \leftrightarrow A_2) \leftrightarrow A_3$ oder zu $A_1 \leftrightarrow (A_2 \leftrightarrow A_3)$. Üblicherweise wird aber vereinbart, daß die klammerfreie Äquivalenzenkette die paarweise Äquivalenz der beteiligten Aussagen bedeuten soll, d. h.

$$A_1 \leftrightarrow A_2 \leftrightarrow A_3 \quad \textit{ist definiert als} \quad (A_1 \leftrightarrow A_2) \ \wedge \ (A_2 \leftrightarrow A_3).$$

Analoges gilt für $A_1 \leftrightarrow A_2 \leftrightarrow \ldots \leftrightarrow A_n$. Dagegen wird $A_1 \to A_2 \to A_3$ oft nicht als $(A_1 \to A_2) \wedge (A_2 \to A_3)$ gelesen, sondern als $A_1 \to (A_2 \to A_3)$. Der Grund ist, daß $A_1 \to (A_2 \to A_3)$ äquivalent ist zu $A_1 \wedge A_2 \to A_3$. Unter iterierter Rechtsklammerung gilt dann allgemeiner

$$A_1 \to A_2 \to \ldots \to A_n \to A_{n+1} \quad \textit{gdw} \quad A_1 \wedge \ldots \wedge A_n \to A_{n+1}.$$

Damit haben von rechts her geklammerte Implikationsketten eine überraschend einfache und sympathische Bedeutung.

Welche Bedeutung eine Kette von Folgerungen hat, ist meistens aus dem Kontext heraus klar. Z. B. bedeutet

$A \to C$ *impliziert*
$A \wedge B \to C$ *impliziert*
$A \wedge B \to C \vee D$,

dass $(A \to C) \to (A \wedge B \to C)$, und dass $(A \wedge B \to C) \to (A \wedge B \to C \vee D)$. Man kann hier auch stilisierte Pfeile oder Doppelpfeile statt „impliziert“ verwenden. Erfahrungsgemäß führen die speziell vom Anfänger gerne im Übermaß verwendeten

Doppelpfeile aber zu schlecht lesbaren Argumenten, insbesondere dann, wenn das nächste Glied der Implikationskette nicht unmittelbar aus dem vorhergehenden folgt, sondern weitere Voraussetzungen oder Zwischenresultate in den Schluß eingehen. In der Regel ist dann ein Ausschreiben des Arguments in ganzen Sätzen vorzuziehen.

Die Sprache der Mathematik

Junktoren und Grundaussagen A_0, A_1, A_2, … reichen für die Bedürfnisse der Mathematik nicht aus. Um zu sehen, wie weit der aussagenlogische Ansatz reicht, setzen wir A_n = „n ist eine Primzahl“. Dann bedeutet zum Beispiel $A_n \wedge A_{n+2}$, daß n ein Primzahlzwilling ist. Aber schon die Formulierung einer so einfachen Behauptung wie „Der Nachfolger jeder Primzahl ungleich 2 ist keine Primzahl.“ bereitet Schwierigkeiten: Die Aussage $A_n \to \neg A_{n+1}$ berücksichtigt die Zahl 2 nicht, und zudem haben wir nicht zum Ausdruck gebracht, daß wir über alle Primzahlen ungleich 2 etwas behaupten. Symbolisch können wir schreiben:

$$(A_3 \to \neg A_4) \wedge (A_4 \to \neg A_5) \wedge (A_5 \to \neg A_6) \wedge \ldots ,$$

aber dieser Ausdruck ist unendlich lang und wirft dadurch ganz neue Fragen und Probleme auf. Ebenso benötigen wir einen unendlich langen Ausdruck, um zu formulieren, daß es beliebig große Primzahlen gibt. Auch eine andere Wahl von Grundaussagen A_n löst diese Probleme nicht. Wir müssen also unsere Sprache substantiell erweitern.

Eine Betrachtung von allgemeinen mathematischen Aussagen zeigt, daß die Mathematik neben den aussagenlogischen Junktoren noch die folgenden Sprachelemente benötigt:

(a) Quantoren wie „für alle“, „es gibt (mindestens) ein“, „es gibt genau ein“, …

(b) Variablen wie a, b, c, … x, y, z, A, B, C, …

(c) Funktionen wie $+$, $\cdot$, $\circ$, $\sqrt{\ }$, sin, …

(d) Relationen wie $=$, $<$, $\leq$, „kongruent“, „prim“, „Element von“, …

(e) Konstanten wie 0, 1, 2, e, π, …

Mit Hilfe von Quantoren, Variablen, Funktionen, Relationen und Konstanten kann, wie die Erfahrung zeigt, im Zusammenspiel mit den Junktoren jede mathematische Aussage ausgedrückt und dadurch präzisiert werden. Wir betrachten hierzu einige Beispiele für typische mathematische Aussagen:

(i) Die reelle Funktion f ist periodisch.

(ii) n ist eine Primzahl.

(iii) Es gibt unendlich viele Primzahlen.

(iv) Die Funktion $\circ$ ist nicht kommutativ.

Diese Aussagen können wir in der Quantorensprache schreiben als:

(i) Es gibt ein $a \neq 0$, sodaß für alle x gilt: $f(x) = f(x + a)$.

(ii) n ist größer als 1 und für alle a, b gilt: Ist $a \cdot b = n$, so ist $a = 1$ oder $b = 1$.

(iii) Für alle m gibt es ein n mit: n ist größer als m und n ist eine Primzahl.

(iv) Es gibt x, y mit: $x \circ y \neq y \circ x$.

Schreiben wir symbolisch $\forall$ für „für alle" und $\exists$ für „es gibt (mindestens) ein", so können wir diese Aussagen formal schreiben als

(i) $\exists a\, (a \neq 0 \wedge \forall x\, (f(x) = f(x + a)))$.

(ii) $n > 1 \wedge \forall a, b\, (a \cdot b = n \rightarrow a = 1 \vee b = 1)$.

(iii) $\forall m\, \exists n\, (n > m \wedge n$ ist eine Primzahl$)$.

(iv) $\exists x, y\, (x \circ y \neq y \circ x)$.

Diese Beispiele ließen sich beliebig fortsetzen. Man kann sogar umgekehrt fordern: Eine Aussage gilt nur dann als mathematisch präzise, wenn sie sich in der Quantorensprache formulieren läßt.

Die genaue Definition der Quantorensprache ist Aufgabe der mathematischen Logik. Wir begnügen uns hier mit beschreibenden Ausführungen, die vor allem den Umgang mit Quantoren illustrieren sollen. Funktionen und Relationen werden wir später noch eingehend betrachten, für jetzt genügt ein naives Verständnis dieser Dinge. Auch auf Fragen nach der Gültigkeit oder Beweisbarkeit von Aussagen können wir hier nicht weiter eingehen. Betonen möchten wir aber, daß ein Analogon zum Wahrheitstafelverfahren für die Quantorensprache nicht mehr zur Verfügung steht. Aussagen dieser Sprache lassen sich nicht mehr mechanisch beweisen! Dagegen kann das argumentative Beweisen durch eine Erweiterung der Schlußregeln (S1) – (S7) eingefangen und präzisiert werden. Diese Schlußregeln spiegeln dann genau das Argumentieren wider, das bei der Beweisführung mathematischer Sätze allerorts verwendet wird.

Die beiden oben schon verwendeten Quantoren „für alle" und „es gibt ein" stehen im Zentrum des Interesses. Sie beziehen sich auf einen kontextabhängigen Bereich, etwa die natürlichen oder die reellen Zahlen. Quantifiziert wird dabei mit Hilfe von Variablen:

(i) „für alle x gilt die Eigenschaft A(x)", in Zeichen: $\forall x\, A(x)$,

(ii) „es gibt ein y mit der Eigenschaft A(y), in Zeichen: $\exists y\, A(y)$.

Mehrere aufeinanderfolgende Allquantoren können wir beliebig vertauschen, denn $\forall x\, \forall y\, A(x, y)$ ist äquivalent zu $\forall y\, \forall x\, A(x, y)$. Wir verwenden in einer solchen Situation auch die Schreibweise $\forall x, y\, A(x, y)$. Analoges gilt für den Existenzquantor. Dagegen dürfen gemischte All- und Existenzquantoren in der Regel nicht vertauscht werden. Wir betrachten hierzu für eine Eigenschaft A(x, y) die beiden Aussagen:

$\forall x\, \exists y\, A(x, y)$ und $\exists y\, \forall x\, A(x, y)$.

Die zweite Aussage impliziert die erste, aber die Umkehrung ist i. a. falsch. Die erste Aussage behauptet, daß es für jedes x ein „gutes" y gibt. Dieses y hängt i. a.

von x ab. Die zweite Aussage behauptet viel stärker, daß es ein y gibt, welches für alle x „gut“ ist.

Zwischen den beiden Quantoren bestehen die folgenden Zusammenhänge:

$\neg\,\exists x\, A(x)$ *ist gleichwertig zu* $\forall x \neg A(x)$,

$\neg\,\forall x\, A(x)$ *ist gleichwertig zu* $\exists x \neg A(x)$. *(Verneinungsregeln für Quantoren)*

Damit kann man, im Sinne der Reduktion von Problemen, den Existenzquantor durch den Allquantor definieren oder umgekehrt.

Wiederholte Anwendung der Verneinungsregeln liefert z. B.:

$\neg\,\forall x\,\exists y\,\forall z\, A(x, y, z)$ *ist gleichwertig zu* $\exists x\,\forall y\,\exists z \neg A(x, y, z)$.

Allgemein kann man eine Negation in eine komplexe Aussage hineinziehen, wenn man dabei All- und Existenzquantoren vertauscht.

Andere Quantoren wie zum Beispiel „es gibt genau ein x“ lassen sich mit Hilfe des All- und Existenzquantors ausdrücken (siehe Übungen). Als Zeichen für „es gibt genau ein“ ist zuweilen $\exists!$ zu finden. So bringt zum Beispiel $\exists!x\ (f(x) = 0)$ zum Ausdruck, daß die Funktion f eine eindeutige Nullstelle besitzt. Für weitere Quantoren wie „es gibt genau zwei“ haben sich keine Zeichen durchgesetzt, man schreibt derartige Quantoren bei Bedarf umgangssprachlich aus.

Prinzipiell können beliebige Symbole als Variable verwendet werden. Jedoch suggeriert der Kontext oft eine Zeichenwahl. So werden z. B. n, m bevorzugt für die natürlichen Zahlen eingesetzt, x, y, z bevorzugt für die reellen Zahlen, i, j bevorzugt als Indexvariable, usw.

Ebenso sind die Funktionen, Relationen und Konstanten kontextabhängig. Die Addition + bedeutet im Kontext der natürlichen Zahlen etwas anderes als im Kontext eines Vektorraumes. Eine Ausnahme bildet die Gleichheit =, die in jedem Kontext die Identität bedeutet. Die Verschiedenheit $x \neq y$ zweier Objekte ist definiert als $\neg(x = y)$.

Bei der Untersuchung eines Objektbereichs werden ständig neue Funktionen, Relationen und Konstanten der Sprache hinzugefügt, unter Befolgung der Maxime „Definiere Neues aus Altem“. Die Zahlentheorie kann zum Beispiel mit der Addition und der Multiplikation als Funktionen, sowie Konstanten 0 und 1 auf ihrem Objektbereich beginnen. Nun können die Relationen „n ist kleiner als m“, in Zeichen $n < m$, sowie „m ist ein Teiler von n“, in Zeichen $m \mid n$, eingeführt werden durch

$n < m$ *wird definiert als* $\exists k\,(k \neq 0 \wedge n + k = m)$,

$m \mid n$ *wird definiert als* $\exists k\,(m \cdot k = n)$.

Die Bedeutung, Gültigkeit und Beweisbarkeit von Aussagen ist kontextabhängig. Im Kontext der rationalen Zahlen bedeutet zum Beispiel die Aussage

$$\forall x, y\,(x < y \;\rightarrow\; \exists z\,(x < z \wedge z < y)),$$

daß zwischen je zwei rationalen Zahlen immer noch eine weitere rationale Zahl liegt. Diese Aussage ist für die rationalen Zahlen beweisbar, im Kontext der ganzen Zahlen ist sie dagegen falsch, da hier z. B. $\neg\,\exists z\,(0 < z \wedge z < 1)$ gilt.

Wir haben damit das Grundgerüst der heutigen mathematischen Sprache vollständig zusammengetragen. Diese Sprache ist eine Quantorensprache, die die reine Aussagenlogik substantiell erweitert. Kontextabhängig werden Relationen, Funktionen und Konstanten eingeführt und untersucht. Das Einführen neuer Begriffe hat einen hierarchischen Charakter, indem die Definition eines neuen Objekts auf bereits definierte Objekte zurückgreift. Was man nun als nicht weiter definierte, nur mit Hilfe von Axiomen beschriebene Grundbegriffe ansieht, ist eine Frage der Ziele, die man verfolgt, und auch eine Frage, welchen Grad an Präzision man letztendlich erreichen möchte. Die Mathematik hat über Jahrhunderte mit Zahlen und geometrischen Figuren gearbeitet, ohne sie genau zu definieren. Das funktionierte und funktioniert auch heute noch sehr gut, da man in Beweisen immer nur bestimmte Eigenschaften der Grundobjekte verwendet, auf die man sich einigen kann, ohne die Grundobjekte genau zu definieren. Dennoch strebt die Mathematik nach höchster Genauigkeit, und speziell bei ihrem heutigen selbstbewußten Umgang mit unendlichen Objekten ist eine Klärung der Fundamente unumgänglich, um ein gewisses Vertrauen in die Widerspruchsfreiheit des gesamten Systems etablieren zu können. Hier spielt nun eine ausgezeichnete Grundrelation die herausragende Rolle, nämlich die Elementbeziehung „x ist ein Element der Menge y“. Die Sprache der modernen Mathematik ist mengentheoretisch geprägt, und man kann sogar die gesamte Mathematik aus der Mengenlehre heraus aufbauen. Den wichtigsten mengentheoretischen Begriffen, die heute in der wissenschaftlichen Mathematik überall verwendet werden, widmet sich das nächste Kapitel.

Aussagenlogische Tautologien

Wir versammeln in tabellarischer Form einige häufig verwendete und daneben auch interessante Zusammenhänge für die Junktoren. Alle folgenden Aussagen sind Tautologien, für beliebige Aussagen A, B, C.

(T1) $\neg\neg A \leftrightarrow A$, *(Stabilität, duplex negatio affirmat)*

$A \vee \neg A$, *(Gesetz vom ausgeschlossenen Dritten, tertium non datur)*

$A \to A$,

$A \to (B \to A)$,

(T2) $\bot \leftrightarrow A \wedge \neg A$,

$\bot \to A$, *(ex falso quodlibet)*

$\neg A \leftrightarrow A \to \bot$,

(T3) $A \to B \leftrightarrow \neg B \to \neg A$, *(Kontrapositionsgesetz)*

(T4) $A \wedge (A \to B) \to B$, *(modus ponens, tautologische Form)*

$\neg B \wedge (A \to B) \to \neg A$, *(modus tollens, tautologische Form)*

(T5) $(A \leftrightarrow B) \quad \leftrightarrow \quad (\neg A \leftrightarrow \neg B)$,
$\neg (A \leftrightarrow B) \quad \leftrightarrow \quad (A \vee B) \wedge \neg (A \wedge B)$,

(T6) $\neg (A \to B) \quad \leftrightarrow \quad A \wedge \neg B$,
$\neg A \to B \quad \leftrightarrow \quad A \vee B$,

(T7) $A \to (B \to C) \quad \leftrightarrow \quad A \wedge B \to C$, *(Auflösung von Implikationsketten)*
$A \to (B \to C) \quad \leftrightarrow \quad B \to (A \to C)$,

(T8) $A \leftrightarrow (B \to A) \wedge (\neg B \to A)$, *(Fallunterscheidung)*

(T9) $(A \to B) \wedge (B \to C) \quad \to \quad (A \to C)$, *(Kettenschluß)*
$(A \wedge B \to C) \wedge (A \to B) \to (A \to C)$, *(Fregescher Kettenschluß)*

(T10) $\neg (A \wedge B) \quad \leftrightarrow \quad \neg A \vee \neg B$,
$\neg (A \vee B) \quad \leftrightarrow \quad \neg A \wedge \neg B$, *(de Morgansche Regeln)*

(T11) $A \vee (B \wedge C) \quad \leftrightarrow \quad (A \vee B) \wedge (A \vee C)$,
$A \wedge (B \vee C) \quad \leftrightarrow \quad (A \wedge B) \vee (A \wedge C)$, *(Distributivgesetze)*

(T12) $(A \to C) \wedge (B \to C) \quad \leftrightarrow \quad (A \vee B) \to C$,
$(A \to C) \vee (B \to C) \quad \leftrightarrow \quad (A \wedge B) \to C$,
$(A \to B) \wedge (A \to C) \quad \leftrightarrow \quad A \to B \wedge C$,
$(A \to B) \vee (A \to C) \quad \leftrightarrow \quad A \to B \vee C$,

(T13) $((A \to B) \to A) \to A$, *(Peirce-Formel)*

(T14) $A \wedge B \to C \vee D \quad \leftrightarrow \quad (A \to C) \vee (B \to D)$.

Daß diese Aussagen Tautologien sind, kann mit Hilfe von Wahrheitstafeln, durch inhaltliches Argumentieren oder mit Hilfe der Schlußregeln gezeigt werden. Oft folgen neue Tautologien auch aus der geschickten Kombination von bereits bekannten Tautologien. Für die zweite Aussage in (T12) gilt zum Beispiel die folgende schrittweise Umformung in paarweise äquivalente Aussagen:

$(A \to C) \vee (B \to C) \quad \leftrightarrow \quad (A \wedge B) \to C \quad gdw_{\text{nach (T5)}}$

$\neg((A \to C) \vee (B \to C)) \quad \leftrightarrow \quad \neg((A \wedge B) \to C) \quad gdw_{\text{nach (T10)}}$

$\neg(A \to C) \wedge \neg(B \to C) \quad \leftrightarrow \quad \neg((A \wedge B) \to C) \quad gdw_{\text{nach (T6)}}$

$A \wedge \neg C \wedge B \wedge \neg C \quad \leftrightarrow \quad A \wedge B \wedge \neg C$.

womit wir bei einer offenbar richtigen Aussage angelangt sind.

Der Leser möge eine Wahrheitstafel für die Tautologie

$(A \to C) \vee (B \to C) \quad \leftrightarrow \quad (A \wedge B) \to C$

erstellen. Dabei zeigt sich, daß die linke wie die rechte Seite nur für die Wahrheitswerte w, w, f von A, B, C falsch ist. Dies kann man sich auch leicht argumentativ klarmachen (und damit die Tautologie beweisen).

Quantorenregeln

Die folgenden Äquivalenzen und Implikationen gelten für alle mathematischen Eigenschaften A(x), B(x), A(x, y).

$$
\begin{array}{lllll}
(Q1) & \neg\,\forall x\,A(x) & \leftrightarrow & \exists x \neg\, A(x), \\
 & \neg\,\exists x\,A(x) & \leftrightarrow & \forall x \neg\, A(x), \\
(Q2) & \forall x\,\forall y\,A(x,y) & \leftrightarrow & \forall y\,\forall x\,A(x,y), \\
 & \exists x\,\exists y\,A(x,y) & \leftrightarrow & \exists y\,\exists x\,A(x,y), \\
(Q3) & \forall x\,(A(x) \wedge B(x)) & \leftrightarrow & \forall x\,A(x) \;\wedge\; \forall x\,B(x), \\
 & \exists x\,(A(x) \vee B(x)) & \leftrightarrow & \exists x\,A(x) \;\vee\; \exists x\,B(x), \\
(Q4) & \forall x\,A(x) \;\vee\; \forall x\,B(x) & \rightarrow & \forall x\,(A(x) \vee B(x)), \\
 & \exists x\,(A(x) \wedge B(x)) & \rightarrow & \exists x\,A(x) \;\wedge\; \exists x\,B(x), \\
(Q5) & \forall x\,(A(x) \rightarrow B(x)) & \rightarrow & \forall x\,A(x) \;\rightarrow\; \forall x\,B(x), \\
 & \exists x\,(A(x) \rightarrow B(x)) & \leftrightarrow & \forall x\,A(x) \;\rightarrow\; \exists x\,B(x).
\end{array}
$$

Wir zeigen zur Illustration die erste Implikation in (Q4).

Beweis von: $\forall x\,A(x) \;\vee\; \forall x\,B(x) \;\rightarrow\; \forall x\,(A(x) \vee B(x))$

Es gelte also $\forall x\,A(x) \vee \forall x\,B(x)$. Dann gilt $\forall x\,A(x)$ oder es gilt $\forall x\,B(x)$.
Gilt $\forall x\,A(x)$, so gilt offenbar die Abschwächung $\forall x\,(A(x) \vee B(x))$.
– Gilt andernfalls $\forall x\,B(x)$, so gilt ebenfalls $\forall x\,(A(x) \vee B(x))$.

Den zweiten Teil von (Q4) zeigt man ähnlich. Er ergibt sich aber auch aus dem ersten Teil durch Kontraposition und Anwendung der Verneinungsregel. (Wir wenden hier die bewiesene Aussage auf $\neg\,A(x)$ und $\neg\,B(x)$ an, und streichen am Ende doppelte Negationen.)

Es ist instruktiv, sich Gegenbeispiele für die fehlenden Implikationen von rechts nach links in (Q4) und (Q5) klarzumachen.

Übungen

Übung 1 *(Aussagen und Junktoren, I) (L)*
Sie stehen vor zwei verschlossenen Türen. Hinter der einen Türe ist ein Schatz, hinter der anderen eine Ziege. Zwischen den beiden Türen sitzt ein Zwerg, der weiß, wo der Schatz ist. Der Zwerg hat die Eigenart, stets zu lügen oder stets die Wahrheit zu sagen. Er ist allwissend, und insbesondere weiß er, ob er ein Lügner ist oder nicht.

Sie dürfen dem Zwerg genau eine aussagenlogische Frage stellen, die er Ihnen mit „ja" oder mit „nein" beantworten wird. (Z. B. „Ist der Schatz links, wenn er rechts ist?") Welche Frage stellen Sie, um zu erfahren, hinter welcher Türe der Schatz liegt?

Übung 2 *(Aussagen und Junktoren, II)*
Der Zwerg aus obiger Aufgabe sitzt nun an einer Weggabelung. Der eine Weg führt zum Dorf der lügenden und der andere Weg zum Dorf der wahrheitssagenden Zwerge. Welche ja-nein-Frage, die diesmal keine Junktoren enthalten darf, können Sie stellen, um zu erfahren, welcher Weg zu welchem Dorf führt?

Übung 3 *(Aussagen und Junktoren, III)*
Sie stehen drei Personen A, B und C gegenüber. Jede Person ist entweder ein Lügner oder ein Wahrheitssager. Sie wissen, daß eine der drei Personen ein Wahrheitssager ist und heute Geburtstag hat.

A sagt zu Ihnen: „B ist ein Wahrheitssager."

C sagt zu Ihnen: „Einer von uns ist ein Lügner."

Welche Person hat Geburtstag?

Übung 4 *(Aussagen und Junktoren, IV)*
Sie stehen drei Personen mit den Schildern A, B und C gegenüber. Sie wissen, daß eine Person immer die Wahrheit sagt (der Wahrheitssager), eine Person immer lügt (der Lügner), und eine Person zufällig die Wahrheit sagt oder lügt (der Stochastiker). Die Personen A, B, C wissen jeweils, wer der Wahrheitssager, wer der Lügner und wer der Stochastiker ist.

Sie dürfen drei aussagenlogische Fragen stellen, um die drei Personen zu identifizieren. Welche Fragen stellen Sie?

[Beispiel für eine Frage an Person B: „Ist A der Wahrheitssager, wenn C der Stochastiker ist?]

Übung 5 *(Semantik der Junktoren: Junktoren in der Umgangssprache, I) (L)*
In der Umgangssprache werden die beiden folgenden Sprechweisen gleichwertig gebraucht:

(a) Hunde und Katzen dürfen nicht an Bord.

(b) Hunde oder Katzen dürfen nicht an Bord.

Welche Tautologie wird hier verwendet?

Übung 6 *(Semantik der Junktoren: Junktoren in der Umgangssprache, II)*
„Wenn der Hahn kräht auf dem Mist, ändert sich das Wetter oder es bleibt wie es ist."

Auf welche Tautologie wird hier angespielt?

Übung 7 *(Semantik der Junktoren: Hierarchischer Aufbau der Junktoren, I)*
Welche Bedeutung hat die verneinte Äquivalenz $\neg(A \leftrightarrow B)$?

Übung 8 *(Semantik der Junktoren: Hierarchischer Aufbau der Junktoren, II)*
Wir betrachten die beiden folgenden zweistelligen Junktoren:

„nicht beide zugleich", in Zeichen: $\uparrow$,

„weder noch", in Zeichen: $\downarrow$.

Der Junktor $\uparrow$ wird auch „nand", „Sheffer-Strich" oder „Unverträglichkeit" genannt, der Junktor $\downarrow$ auch „nor" oder „Nihilition".

Die Wahrheitstafeln dieser Junktoren sind:

A	$\uparrow$	B
w	f	w
w	w	f
f	w	w
f	w	f

A	$\downarrow$	B
w	f	w
w	f	f
f	f	w
f	w	f

(a) Definieren Sie $\uparrow$ und $\downarrow$ durch $\neg$ und $\wedge$.

(b) Definieren Sie $\neg$ und $\wedge$ jeweils durch $\uparrow$ bzw. $\downarrow$.

Übung 9 *(Semantik der Junktoren: Hierarchischer Aufbau der Junktoren, III) (L)*
Läßt sich die Negation mit Hilfe der Junktoren $\rightarrow$, $\wedge$, $\vee$ (ohne Falsum) definieren?

Übung 10 *(Semantik der Junktoren: Hierarchischer Aufbau der Junktoren, IV) (L)*
Geben Sie die achtzeiligen Wahrheitstafeln für die dreistelligen Junktoren

$*_{\geq 2}$ = „mindestens zwei der drei", $*_{0,3}$ = „wenn eines, dann alle drei"

an, und definieren Sie $*_{\geq 2}(A, B, C)$ und $*_{0,3}(A, B, C)$ mit Hilfe der üblichen Junktoren.

Übung 11 *(Semantik der Junktoren: Hierarchischer Aufbau der Junktoren, V)*
Schreiben Sie eine „zufällige" Wahrheitstafel für A, B, C an, also eine Tafel mit acht Zeilen und vier Spalten $*$, A, B, C, wobei die $*$-Spalte z. B. durch das Werfen einer Münze mit w oder f gefüllt wird. Die Tafel können Sie dann als einen dreistelligen Junktor $*(A, B, C)$ ansehen. Definieren Sie nun den Junktor $*$ mit Hilfe der üblichen Junktoren.

Zeigen Sie nun allgemein, daß jeder n-stellige Junktor $*(A_1, \ldots A_n)$, $n \geq 1$, mit Hilfe der üblichen Junktoren definiert werden kann.

[Betrachten Sie eine Konjunktion von 2^n Aussagen der Form $L_1 \wedge \ldots \wedge L_n \to L$, wobei jedes L_i entweder A_i oder $\neg A_i$ und L entweder $\bot$ oder $\neg\bot$ ist. Alternativ läßt sich die Behauptung auch durch Induktion nach n zeigen.]

Übung 12 *(Semantik der Junktoren: Junktoren und Wahrheitswerte, I) (L)*
Eine Aussage A heißt *erfüllbar*, wenn $\neg$ A keine Tautologie ist, d. h. wenn die Ergebnisspalte der Wahrheitstafel von A mindestens einmal den Wert „w" enthält. Zeigen oder widerlegen Sie, daß für alle Aussagen A und B gilt:

(i) Sind A und $A \to B$ erfüllbar, so ist B erfüllbar.

(ii) Ist A erfüllbar und ist $A \to B$ eine Tautologie, so ist B erfüllbar.

(iii) Ist A nicht erfüllbar und ist $A \vee B$ eine Tautologie, so ist B eine Tautologie.

Übung 13 *(Semantik der Junktoren: Junktoren und Wahrheitswerte, II)*
Geben Sie Aussagen A, B, C an derart, daß gilt:

(i) $A \wedge B \wedge C$ ist nicht erfüllbar.

(ii) $A \wedge B$, $A \wedge C$, $B \wedge C$ sind erfüllbar.

Übung 14 *(Semantik der Junktoren: Junktoren und Wahrheitswerte, III)*
Welche Beziehung besteht zwischen der Erfüllbarkeit/Allgemeingültigkeit der Aussagen A, B, C, D und der Erfüllbarkeit/Allgemeingültigkeit der Aussage $A \vee B \vee C \vee D$ bzw. der Aussage $A \wedge B \wedge C \wedge D$?

Übung 15 *(Semantik der Junktoren: Junktoren und Wahrheitswerte, IV)*
Mit $\wedge$ und $\neg$ als Basisjunktoren sei $\bot$ definiert als $A_0 \wedge \neg A_0$ für eine beliebige Aussage A_0. Begründen Sie, daß die für die Widerspruchsbeweise verantwortliche Tautologie $A \to \bot \leftrightarrow \neg A$ als Spezialfall des Kontrapositionsgesetzes aufgefaßt werden kann.

Übung 16 *(Semantik der Junktoren: Junktoren in mathematischen Beweisen, I) (L)*
Beweisen Sie folgende Aussagen mit Hilfe der Schlußregeln (S1) – (S5):

(i) $(A \to (B \to C)) \to (A \wedge B \to C)$,

(ii) $A \to \neg\neg A$,

(iii) $\neg\neg\neg A \to \neg A$.

Übung 17 *(Semantik der Junktoren: Junktoren in mathematischen Beweisen, II)*
Beweisen Sie folgende Aussagen mit Hilfe der Schlußregeln (S1) – (S7):

(i) $A \vee \neg A$,

(ii) $(\neg A \to \neg B) \to (B \to A)$,

(iii) $((A \to B) \to A) \to A$.

Übung 18 *(Die Sprache der Mathematik, I)*
Definieren Sie den Quantor „es gibt genau ein x mit A(x)", in Zeichen $\exists!\, x\, A(x)$, mit Hilfe des Existenz- und Allquantors und der Gleichheit.

Übung 19 *(Die Sprache der Mathematik, II) (L)*
Formulieren Sie nur mit Hilfe der Quantoren, der Junktoren und der Gleichheit die Aussagen:

(a) Es gibt mindestens drei verschiedene Objekte.

(b) Es gibt genau drei verschiedene Objekte.

Übung 20 *(Die Sprache der Mathematik, III) (L)*
Lesen Sie „$\forall x$" als „für alle x", „M(x)" als „x ist ein Mensch", „Z(x)" als „x ist ein Zeitpunkt" und „L(x, y)" als „x lügt zum Zeitpunkt y".
Drücken Sie nun die folgenden Aussagen formal aus:

(i) Manche Menschen lügen immer.

(ii) Alle Menschen lügen manchmal.

(iii) Manche Menschen sagen manchmal die Wahrheit.

(iv) Es gibt genau einen Menschen, der immer die Wahrheit sagt.

(v) Von je zwei verschiedenen Menschen sagt zu einem gewissen Zeitpunkt der eine die Wahrheit und der andere nicht.

Übung 21 *(Die Sprache der Mathematik, IV)*
Lesen Sie „$\forall x$" als „für alle Menschen x", „b" als „der Butler", „g" als „der Gärtner", „l" als „der Lord", „K(x, y)" als „x hat y ermordet", „A(x, y)" als „x hat Angst vor y", „H(x, y)" als „x haßt y", und formalisieren Sie die folgenden Aussagen:

(a) Der Butler oder der Gärtner hat den Lord umgebracht, oder der Lord hat Selbstmord begangen.

(b) Man mordet nur die, die man haßt und vor denen man Angst hat.

(c) Diejenigen, die der Lord haßt, mag der Gärtner.

(d) Diejenigen, die der Lord haßt, haßt auch der Butler.

(e) Der Lord haßt sich selbst, und er haßt den Gärtner.

(f) Der Butler haßt alle, die Angst vor dem Lord haben.

(g) Jeder mag den Lord, den Butler oder den Gärtner.

Nehmen Sie nun diese Informationen als gegeben an, und ermitteln Sie durch eine möglichst detaillierte logische Argumentation den Mörder des Lords.

Übung 22 (Aussagenlogische Tautologien)
Beweisen Sie einige der aussagenlogischen Tautologien der obigen Tabelle mit Hilfe von Wahrheitstafeln, semantisch (d. h. durch inhaltliche Argumentation), oder mit Hilfe der Schlußregeln.

Übung 23 (Quantorenregeln, I)
Beweisen Sie einige der Quantorenregeln der obigen Tabelle und geben Sie Gegenbeispiele für die fehlenden Implikationen an.

Übung 24 (Quantorenregeln, II)
Betrachten Sie ein Zweipersonenspiel, und lesen Sie x_i als „der i-te Zug von Spieler I“ und y_i als „der i-te Zug von Spieler II“. Die Anzahl der Züge wird auf ein festes n (z. B. 500) begrenzt, und für jede mögliche Partie $x_1, y_1, x_2, y_2, \ldots, \ldots, x_n, y_n$ steht fest, ob Spieler I oder Spieler II diese Partie gewonnen hat (es gibt kein Unentschieden).

(a) Formulieren Sie mit Hilfe der Quantoren $\forall$ und $\exists$ die Aussagen „Spieler I besitzt eine Gewinnstrategie“ und „Spieler II besitzt eine Gewinnstrategie“.

(b) Zeigen Sie mit Hilfe der Verneinungsregeln für die Quantoren, daß genau einer der beiden Spieler eine Gewinnstrategie besitzt.

(c) Was läßt sich sagen, wenn bestimmte Partien mit „Remis“ enden?

(d) Wie lassen sich die Ergebnisse auf das Schachspiel anwenden?

[Zu (c): Definieren Sie zwei Hilfsspiele, bei denen Spieler I bzw Spieler II genau die Remispartien und seine Gewinnpartien des Originalspiels gewinnt.]

2. Mengen

Georg Cantor, der Begründer der Mengenlehre – und damit einer der Mitbegründer der modernen Mathematik – hat Ende des 19. Jahrhunderts folgende intuitive Beschreibung des Mengenbegriffs gegeben:

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

Ist M eine Menge und ist x ein Element von M, so schreiben wir

$x \in M$. *(Elementrelation)*

Das Zeichen $\in$ ist ein stilisiertes griechisches ε (Epsilon). Wir lesen den Ausdruck „$x \in M$" als „x ist ein Element von M", „x epsilon M", „x ist in M als Element enthalten" oder kurz als „x in M".

Was genau eine Menge ist, läßt sich immer nur umschreiben, aber nicht im üblichen Sinne definieren. Der Mengenbegriff ließe sich nur dann exakt definieren, wenn man auf andere wiederum undefinierte Grundbegriffe zurückgreifen wollte. Irgendwo muß man anfangen, die reine Logik genügt für die Mathematik nicht. Speziell der Mengenbegriff hat sich als undefinierter Grundbegriff sehr gut bewährt. Alles läßt sich auf ihn zurückführen, weitere undefinierte Grundbegriffe müssen nicht verwendet werden. Anders: Wir dürfen annehmen, daß jedes mathematische Objekt eine Menge ist. Wir werden im weiteren Verlauf an einigen Stellen andeuten, wie eine derartige mengentheoretische Interpretation von mathematischen Objekten wie Relationen, Funktionen, Zahlen durchgeführt werden kann. Zum Verständnis der Grundlagen der Mengenlehre und zur sicheren Beherrschung der mengentheoretischen Sprechweisen sind genauere Einblicke in diesen universellen Aspekt des Mengenbegriffs aber nicht notwendig.

Wir beschränken uns hier auf Mengen, die aus mathematischen Objekten gebildet sind und engen also Cantors allgemeineren Mengenbegriff etwas ein, der beliebige Objekte „unserer Anschauung oder unseres Denkens" zuläßt. Wir bilden also z. B. keine Mengen aus Äpfeln oder Birnen.

Innerhalb einer axiomatischen Mengenlehre wird die Existenz von Mengen durch Axiome genauer geregelt. Für uns genügt Cantors Beschreibung, und wir bilden Mengen in einer sehr freien Art und Weise. Auf die logischen Probleme einer beliebigen „Zusammenfassung von Objekten zu einem Ganzen" werden wir unten aber noch zu sprechen kommen.

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_2,

Extensionalität

Ein grundlegendes Prinzip der Mengenlehre ist das sog. *Extensionalitätsprinzip*, das der Leser wahrscheinlich intuitiv bereits angewendet hat. Dieses Prinzip bringt zum Ausdruck, daß eine Menge vollständig durch ihre Elemente bestimmt ist, und nicht noch etwa eine „Farbe" besitzt. Es lautet genau:

Haben zwei Mengen dieselben Elemente, so sind sie gleich.
(Extensionalitätsprinzip)

In unserer Quantorensprache können wir dieses Prinzip so formulieren:

$\forall x\, \forall y\, \forall z\, ((z \in x \leftrightarrow z \in y) \rightarrow x = y)$
(formale Version des Extensionalitätsprinzips)

In dieser formalen Version nehmen wir an, daß alle Objekte unserer Theorie Mengen sind. Wenn wir dieser Konvention nicht folgen möchten, so müssen wir ein Prädikat M(x) für „x ist Menge" in unsere Sprache mit aufnehmen. Die formale Version des Extensionalitätsprinzips lautet dann:

$\forall x\, \forall y\, (M(x) \wedge M(y) \wedge \forall z\, (z \in x \leftrightarrow z \in y) \rightarrow x = y)$.

Im folgenden nehmen wir aber an, daß die mathematischen Objekte genau die Mengen sind. Dies vereinfacht viele Formulierungen. Die Quantoren laufen über alle Mengen: „$\forall x$" ist gleichwertig zu „für alle Objekte x" und zu „für alle Mengen x".

Die Umkehrung der Implikation im Extensionalitätsprinzip gilt aus rein logischen Gründen: Sind zwei Objekte gleich, so stimmen sie in allen Eigenschaften überein, und insbesondere also in ihren Elementen. Formal:

$\forall x\, \forall y\, (x = y \rightarrow \forall z\, (z \in x \leftrightarrow z \in y))$.

Zusammengenommen gilt also:

Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente haben.

Das Extensionalitätsprinzip hat zur Folge, daß Reihenfolge und Wiederholungen bei der Mengenbildung keine Rolle spielen. Die Zusammenfassung der Objekte a, b, c, d liefert zum Beispiel dieselbe Menge wie die Zusammenfassung von a, c, b, d oder von a, b, d, c, a, b. Die Elemente einer Menge haben keine Ordnung wie bei einer Liste, und sie tauchen auch nicht mehrfach auf wie bei einer sog. Multimenge. Der Mengenbegriff ruht damit auf relativ komplizierten Abstraktionen, da in konkreter Rede und Schrift eine Reihenfolge der betrachteten Einzelobjekte immer vorhanden ist.

Eine wichtige Relation zwischen Mengen im Umfeld des Extensionalitätsprinzips ist die Teilmengenrelation:

Definition *(Teilmenge, Obermenge)*
Für alle Mengen A, B definieren wir:

$A \subseteq B$, *falls* $\forall a \in A\ a \in B$, *(Teilmenge)*

$A \subset B$, *falls* $A \subseteq B \wedge A \neq B$, *(echte Teilmenge)*

$A \supseteq B$, *falls* $B \subseteq A$, *(Obermenge)*

$A \supset B$, *falls* $B \subset A$ *(echte Obermenge)*

Gilt $A \subseteq B$ und $B \subseteq A$, so gilt $A = B$ nach dem Extensionalitätsprinzip. Dies führt zu einem häufig verwendeten mengentheoretischen Beweisprinzip: Zu zeigen ist, daß zwei Mengen A und B gleich sind. Wir zeigen hierzu in zwei Schritten, daß $A \subseteq B$ und $B \subseteq A$ gilt. Dann haben wir $A = B$ bewiesen. Diese Zerlegung der Beweislast in zwei Teile bringt oft eine erhebliche Erleichterung mit sich.

Wichtige Eigenschaften, die für alle Mengen A, B, C gelten, sind:

(i) $A \subseteq A$, *(Reflexivität)*

(ii) $A \subseteq B \wedge B \subseteq A \rightarrow A = B$, *(Antisymmetrie)*

(iii) $A \subseteq B \wedge B \subseteq C \rightarrow A \subseteq C$. *(Transitivität)*

Komprehensionen mit Hilfe von Eigenschaften

Definition *(Mengenkomprehension)*
Sei $\mathscr{E}(x)$ eine Eigenschaft. Dann sei, im Falle der Existenz,

$\{ x \mid \mathscr{E}(x) \}$

die Menge aller Objekte x mit der Eigenschaft $\mathscr{E}(x)$. Die Menge $\{ x \mid \mathscr{E}(x) \}$ heißt dann die zu $\mathscr{E}(x)$ gehörige *Mengenkomprehension*.

Ist $M = \{ x \mid \mathscr{E}(x) \}$, so gilt für alle x nach Definition:

(+) $x \in M$ *gdw* $\mathscr{E}(x)$.

Der Leser wird sich vielleicht über den Zusatz „im Falle der Existenz" gewundert haben. Der Grund für diesen Zusatz ist, daß die Zusammenfassung von Objekten zu einer Menge nicht in jedem Falle möglich ist. Cantor war dies bewußt, und die Betonung „zu einem Ganzen" in seiner Beschreibung des Mengenbegriffs deutet möglicherweise darauf hin. Einen ersten Hinweis auf die Problematik gibt die Zusammenfassung aller Objekte, also die Komprehension

$V = \{ x \mid x = x \}$. *(mathematisches Universum)*

Nachdem sicher $V = V$ gilt, haben wir nach der Äquivalenz (+):

$V \in V$.

Liefert also jede Zusammenfassung von Objekten eine Menge – wovon man naiv ausgehen könnte –, so gibt es Mengen, die sich selbst enthalten. Diese Merkwür-

digkeit führt aber noch nicht direkt zu einem Widerspruch. Einen Widerspruch erhalten wir aber, wenn wir die Eigenschaft, sich selbst als Element zu enthalten, negieren. Wir setzen:

$R = \{ x \mid x \notin x \}$. *(Russell-Zermelo-Komprehension)*

Nach (+) gilt dann für alle x:

$x \in R$ *gdw* $x \notin x$.

Setzen wir hier für x nun R selbst ein, so erhalten wir:

$R \in R$ *gdw* $R \notin R$.

Dieser unerreicht prägnante Widerspruch ist als *Russell-Zermelo-Paradoxon* bekannt. Das Argument zeigt, daß die Zusammenfassung von Objekten einer Theorie nicht in jedem Falle ein Objekt der Theorie liefern kann. Es werden keine mathematischen Konstruktionen verwendet, das Argument ist logischer Natur.

Eine umgangssprachliche Variante der Paradoxie ist Russells Dorfbarbier. Dieser Barbier behauptet: „Ich schneide genau den Dorfbewohnern die Haare, die sich selbst nicht die Haare schneiden." Diese Aussage erweist sich nun aber als unhaltbar: Wer sollte dem Dorfbarbier die Haare schneiden? Er muß sich die Haare schneiden, wenn er sie sich nicht selber schneidet. Und schneidet er sie sich selbst, so verletzt er ebenfalls seine Aussage.

Über die Russell-Zermelo-Paradoxie ist sehr viel nachgedacht worden, und wir können hier auf die tiefsinnigen Fragen der mathematischen Grundlagenforschung und auf das über ihnen errichtete logisch-axiomatische Gebäude nicht weiter eingehen. Das heute übliche Vorgehen, das alle aufgetretenen Paradoxien vermeidet, ist aber sehr einfach zu beschreiben. An die Stelle der beliebigen Zusammenfassung von Objekten tritt die *Aussonderung* von Objekten mit einer bestimmten Eigenschaft aus einer gegebenen Menge. Für jede Menge M und jede Eigenschaft $\mathscr{E}(x)$ existiert, per Axiom, die Menge

$N = \{ x \in M \mid \mathscr{E}(x) \} = \{ x \mid x \in M \wedge \mathscr{E}(x) \}$. *(Aussonderungsprinzip)*

Dieses Prinzip wird noch durch weitere Prinzipien (Axiome) ergänzt, die die Existenz von Mengen mit bestimmten Eigenschaften garantieren. Viele von diesen Prinzipien fordern direkt, daß bestimmte Komprehensionen $\{ x \mid \mathscr{E}(x) \}$ Mengen sind. So ist etwa, per Axiom, die Komprehension $\{ x \mid x = a \vee x = b \}$ für alle mathematischen Objekte a, b eine Menge.

Zusammengenommen erzeugen die Axiome der Mengenlehre ein genügend reichhaltiges und bislang äußerst stabiles mathematisches Universum. Das Aussonderungsprinzip genügt, um der Mathematik in unkomplizierter Weise alle Zusammenfassungen zur Verfügung zu stellen, die sie benötigt. Denn für jede „natürliche" in der Mathematik auftauchende Eigenschaft $\mathscr{E}(x)$, für die eine Komprehension gewünscht wird, läßt sich immer eine Menge M finden, sodaß $\mathscr{E}(x)$ nur für Elemente von M zutrifft, d. h. es gilt

$\{ x \mid \mathscr{E}(x) \} = \{ x \in M \mid \mathscr{E}(x) \}$.

In diesem Fall ist also die Zusammenfassung aller Objekte mit der Eigenschaft $\mathscr{E}(x)$ eine Menge. Anders formuliert: Das Phänomen der inkonsistenten Zusammenfassungen fällt in der Mathematik üblicherweise gar nicht auf. Man kann ohne große Bedenken die Komprehension $\{ x \mid \mathscr{E}(x) \}$ durchführen, und man darf mit der so gebildeten Menge frei operieren und mit ihrer Hilfe wieder andere Mengen bilden.

Die prinzipiellen Limitationen bleiben. Die Zusammenfassungen

$$V = \{ x \mid x = x \} \quad \text{und} \quad R = \{ x \mid x \notin x \}$$

sind in der axiomatischen Mengenlehre keine Mengen. Für R haben wir dies bereits gezeigt: Wäre R eine Menge, so würde $R \in R \leftrightarrow R \notin R$ gelten, was nicht sein kann. Und das Universum V kann unter dem Aussonderungsprinzip keine Menge sein, da sonst $R = \{ x \in V \mid x \notin x \}$ ebenfalls eine Menge wäre.

Statt von Zusammenfassungen oder Komprehensionen spricht man auch von *Klassen*, und Zusammenfassungen, die keine Mengen sind, nennt man *echte Klassen*. Die Klassen V und R sind Beispiele für echte Klassen. Die Intuition ist: Echte Klassen entstehen durch zu große, uferlose Zusammenfassungen. Ist eine Zusammenfassung beschränkt, so hinterläßt sie ein „fertiges Ganzes", eine Menge. Mengen sind in diesem Sinne „kleine Klassen". Wir können eine beliebige Zusammenfassung betrachten und ihr ein Zeichen wie V oder einen Namen wie „Russell-Zermelo-Klasse" geben, aber wir können i. a. nicht mehr so frei und sorglos mit ihnen operieren wie mit einer Menge.

Einfache Mengenbildungen

Nach diesen Vorbereitungen können wir uns der Bildung neuer Begriffe zuwenden. Die hier nicht weiter verfolgte Axiomatik der Mengenlehre garantiert, daß die folgenden Zusammenfassungen durchweg Mengen sind.

Definition *(elementare Mengenbildungen)*
Wir definieren:

$$\emptyset = \{ x \mid x \neq x \}, \qquad \textit{(leere Menge)}$$

$$\{ a \} = \{ x \mid x = a \}, \qquad \textit{(Einermenge)}$$

$$\{ a, b \} = \{ x \mid x = a \vee x = b \}, \qquad \textit{(Paarmenge)}$$

$$\{ a_1, \ldots, a_n \} = \{ x \mid x = a_1 \vee \ldots \vee x = a_n \}. \qquad \textit{(Angabe der Elemente)}$$

Statt $\emptyset$ schreiben wir oft auch 0 oder $\{\,\}$.

Es gilt $\{ a, a \} = \{ a \}$. Weiter ist $\{ a, b \} = \{ b, a \}$. Oft werden aber geordnete Paare (a, b) in der Mathematik gebraucht, bei denen die Reihenfolge eine Rolle spielt. Geordnete Paare kann man durch folgende geistreiche Konstruktion über die Paarmengenbildung einführen.

Definition *(geordnetes Paar)*
Für alle a, b setzen wir $(a, b) = \{\{a\}, \{a, b\}\}$. Die Menge (a, b) heißt das *geordnete Paar* oder *Kuratowski-Paar* von a und b.

Diese Definition des geordneten Paares ist das Paradebeispiel für das oben schon angesprochene erstaunliche Phänomen, daß sich alle mathematischen Objekte als Mengen interpretieren lassen. Interpretieren heißt hier, daß wir nicht behaupten, das Paar (a, b) ist, in einem ontologischen oder intuitiven Sinne, die Menge $\{\{a\}, \{a, b\}\}$. Wir behaupten lediglich, daß die Definition von (a, b) als $\{\{a\}, \{a, b\}\}$ uns alle erwünschten Eigenschaften liefert. Die einzige Eigenschaft geordneter Paare, die in der Mathematik wirklich gebraucht wird, ist, daß für alle Objekte a, b, c, d gilt:

$(a, b) = (c, d)$ *gdw* $a = c$ und $b = d$. *(Korrektheit der Paardefinition)*

Diese Eigenschaft gilt für unsere Paardefinition, wie man leicht nachprüft.

Wir definieren nun nach dem Prinzip „Neues aus Altem“ auch noch Tripel, Quadrupel, usw., nämlich durch

$(a, b, c) = ((a, b), c)$, *(Tripel)*

$(a, b, c, d) = ((a, b, c), d)$, *(Quadrupel)*

$(a, b, c, d, e) = ((a, b, c, d), e)$, … *(Quintupel, usw.)*

Man prüft wieder die Korrektheit dieser Definitionen. Für Tripel gilt zum Beispiel $(a, b, c) = (d, e, f)$ genau dann, wenn $a = d$, $b = e$ und $c = f$ gilt.

Eine Auflösung der Tripeldefinition liefert ein schon schwer zu durchschauendes Strukturungetüm:

$$(a, b, c) = ((a, b), c) = \{\{(a, b)\}, \{(a, b), c\}\} = \{\{\{\{a\}, \{a, b\}\}\}, \{\{\{a\}, \{a, b\}\}, c\}\}.$$

Operationen mit Mengen

Wir definieren eine Reihe von häufig verwendeten elementaren Operationen mit Mengen.

Definition *(elementare Mengenoperationen)*
Für alle Mengen A, B setzen wir:

$A \cap B = \{a \mid a \in A \wedge a \in B\}$, *(Durchschnitt)*

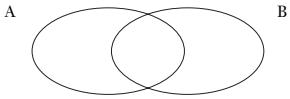
$A \cup B = \{a \mid a \in A \vee a \in B\}$, *(Vereinigung)*

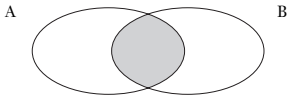
$A - B = A \setminus B = \{a \mid a \in A \wedge a \notin B\}$, *(Differenz)*

$A^c = \{a \in M \mid a \notin A\}$ für eine Grundmenge M, *(Komplementbildung)*

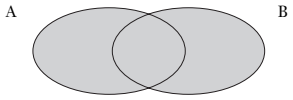
$A \Delta B = (A - B) \cup (B - A)$. *(symmetrische Differenz)*

Die elementaren Mengenoperationen können wir mit Hilfe von Diagrammen leicht visualisieren. Die hervorgehobenen Bereiche deuten dabei das Ergebnis der durchgeführten Operationen an.

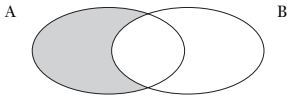




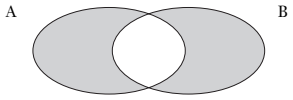
$A \cap B$



$A \cup B$

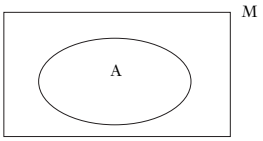


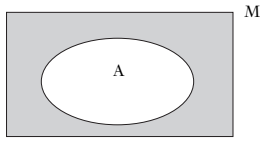
$A - B$



$A \Delta B$

Die Komplementbildung bezüglich einer Grundmenge M stellt man oft in der folgenden Form dar:

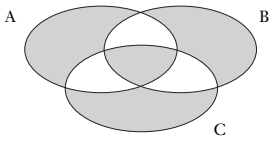




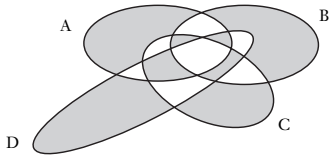
$A^c = M - A$

Mengendiagramme wurden seit Leibniz und Euler benutzt. Visualisierungen des obigen Typs wurden dann vor allem von John Venn verwendet und werden in der Literatur heute auch oft als *Venn-Diagramme* bezeichnet.

Die beiden folgenden Diagramme visualisieren schließlich die symmetrischen Differenzen $(A \Delta B) \Delta C$ mit 7 Schnitt-Teilen sowie $((A \Delta B) \Delta C) \Delta D$ mit 15 Schnitt-Teilen. (Vgl. hierzu auch die Übungen 9 und 10.)



$(A \Delta B) \Delta C$



$((A \Delta B) \Delta C) \Delta D$

Einige Rechengesetze der Operationen $\cap$, $\cup$, $-$, … behandeln wir in den Übungen.

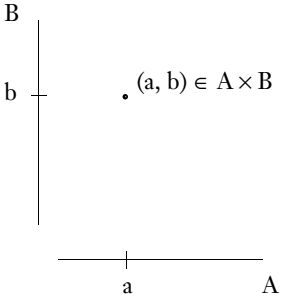
Mit Hilfe der geordneten Paare können wir nun auch Produkte definieren:

Definition *(Kreuzprodukt)*

Für alle Mengen A, B setzen wir:

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

Die Menge $A \times B$ heißt das *Kreuzprodukt* oder *kartesische Produkt* von A und B.



Hierbei ist die operationale Komprehension

$\{ (a, b) \mid a \in A \wedge b \in B \}$

eine Kurzschreibweise für die Komprehension

$\{ z \mid \exists a \in A \; \exists b \in B \;\; z = (a, b) \}$,

die von der Form $\{ z \mid \mathscr{E}(z) \}$ ist. Derartige Varianten der Komprehension verwenden wir von nun an ohne weiteren Kommentar.

Wir definieren $A \times B \times C = (A \times B) \times C$, usw. Es gilt dann nach obiger Definition des Tripels (a, b, c) als ((a, b), c) wie gewünscht:

$$A \times B \times C = \{ (a, b, c) \mid a \in A, \; b \in B, \; c \in C \}.$$

Wir schreiben oft auch A^2 anstelle von $A \times A$. Weiter sei $A^3 = A^2 \times A$, usw.

Potenzmengen

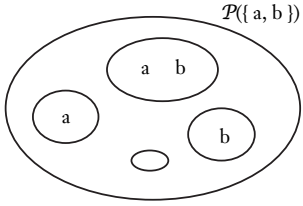
Wir fassen nun alle Teilmengen einer Menge zu einem neuen Objekt zusammen:

Definition *(Potenzmenge)*

Wir definieren für jede Menge A:

$$\mathcal{P}(A) = \{ B \mid B \subseteq A \}.$$

Die Menge $\mathcal{P}(A)$ heißt die *Potenzmenge von A*.



Beispielsweise gilt:

$\mathcal{P}(\varnothing) = \{ \varnothing \}$,

$\mathcal{P}(\{ \varnothing \}) = \{ \varnothing, \{ \varnothing \} \}$,

$\mathcal{P}(\{ a, b \}) = \{ \varnothing, \{ a \}, \{ b \}, \{ a, b \} \}$.

$\mathcal{P}(\{ a, b, c \}) = \{ \varnothing, \{ a \}, \{ b \}, \{ c \}, \{ a, b \}, \{ a, c \}, \{ b, c \}, \{ a, b, c \} \}$.

Die Existenz der Potenzmenge wird in einem axiomatischen Aufbau durch ein eigenes Axiom garantiert, das kühn und direkt fordert, daß für alle A die Zusammenfassung $\{ B \mid B \subseteq A \}$ eine Menge ist. Formal:

$\forall A \, \exists \mathscr{A} \, \forall B \, (B \in \mathscr{A} \leftrightarrow B \subseteq A)$. *(Potenzmengenaxiom)*

Mengensysteme

Von Interesse sind oft „gute" Teilmengen der Potenzmenge. Wir definieren hierzu:

Definition *(Mengensystem auf einer Menge)*
Sei A eine Menge, und sei $\mathscr{A} \subseteq \mathcal{P}(A)$. Dann heißt $\mathscr{A}$ ein *Mengensystem auf* A.

Alle Elemente eines Mengensystems $\mathscr{A}$ auf A sind also Teilmengen der Menge A. Die einfachsten Beispiele für Mengensysteme auf A sind $\varnothing$ und $\mathcal{P}(A)$. Sind $\mathscr{A}$ und $\mathscr{B}$ Mengensysteme auf A, so auch $\mathscr{A} \cup \mathscr{B}$ und $\mathscr{A} \cap \mathscr{B}$.

Für Mengensysteme können wir eine Schnitt- und Vereinigungsoperation einführen:

Definition *(Durchschnitt und Vereinigung eines Mengensystems)*
Für ein Mengensystem $\mathscr{A}$ setzen wir:

$\bigcap \mathscr{A} = \{ a \mid \forall B \in \mathscr{A} \; a \in B \}$, *falls* $\mathscr{A} \neq \varnothing$, *(großer Durchschnitt)*

$\bigcup \mathscr{A} = \{ a \mid \exists B \in \mathscr{A} \; a \in B \}$. *(große Vereinigung)*

Für $\mathscr{A} = \{ \{ a, b, c \}, \{ b, c \}, \{ b, d \} \}$ ist z. B. $\bigcup \mathscr{A} = \{ a, b, c, d \}$ und $\bigcap \mathscr{A} = \{ b \}$.

In dieser Definition verwenden wir die sogenannten *beschränkten Quantoren* „für alle $x \in X$" und „es gibt ein $x \in X$". Diese lassen sich wie folgt auf die üblichen Quantoren zurückführen:

$\forall x \in X \; A(x, X)$ *wird definiert als* $\forall x \, (x \in X \rightarrow A(x, X))$,

$\exists x \in X \; A(x, X)$ *wird definiert als* $\exists x \, (x \in X \wedge A(x, X))$.

Wir schreiben zudem „$\forall x, y \in X$" für „$\forall x \in X \, \forall y \in X$". Analoge Notationen gelten für den Existenzquantor.

Ein Mengensystem $\mathscr{A}$ auf A heißt *abgeschlossen* unter Vereinigungen, falls für alle $B, C \in \mathscr{A}$ gilt, daß $B \cup C \in \mathscr{A}$. Analog ist die Abgeschlossenheit unter Durchschnitten und Komplementbildungen definiert.

Definition *(Mengenverband, Mengenalgebra)*
Sei $\mathscr{A}$ ein Mengensystem auf einer Menge A. Ist $\mathscr{A}$ abgeschlossen unter Vereinigungen und Durchschnitten und gilt $\varnothing \in \mathscr{A}$, so heißt $\mathscr{A}$ ein *Mengenverband* auf A. Ist $\mathscr{A}$ zudem abgeschlossen unter Komplementbildung in A, so heißt $\mathscr{A}$ eine *Mengenalgebra* auf A.

In der allgemeinen Mathematik treten Mengen in vielen Fällen in der folgenden Form auf: Studiert wird ein Bereich X von Objekten, etwa gewisse Zahlen oder gewisse Funktionen. Dieser Bereich X ist eine Menge und die Elemente von X werden meist mit passenden kleinen Buchstaben bezeichnet, etwa x, y, z für reelle Zahlen, n, m, k für natürliche Zahlen, f, g, h für Funktionen. Für Teilmengen von X verwendet man nun große Buchstaben wie A, B, C. Schließlich werden dann Zeichen wie $\mathcal{A}, \mathcal{B}, \mathcal{C}$ für Mengensysteme auf X verwendet. Ist X die Menge der reellen Zahlen, so bezeichnen also

a, b, c, … reelle Zahlen,

A, B, C, … Mengen von reellen Zahlen (z. B. Intervalle), und

$\mathcal{A}, \mathcal{B}, \mathcal{C}$ Mengen von Mengen reeller Zahlen,

etwa $\mathcal{A} = \{ I \mid I \text{ ist ein offenes reelles Intervall} \}$. Diese Komplexitätsstufung in „Punkt, Menge von Punkten, Mengensystem" und zugehörigen Zeichen a, A, $\mathcal{A}$ erleichtert in vielen Fällen die Lesbarkeit, sie ist aber keineswegs zwingend und wird auch nicht konsequent durchgeführt. In der Mengenlehre ist sie unzweckmäßig, denn streng genommen ist ja bereits $(a, b) = \{\{a\}, \{a, b\}\}$ ein Mengensystem. Hier sind dann Schreibweisen wie $\bigcap x$ durchaus üblich; das „kleine x" kann dabei ein sehr kompliziertes Mengensystem sein.

Übungen

Übung 1 *(Komprehensionen mit Hilfe von Eigenschaften, I) (L)*
Wir definieren eine Zusammenfassung von Objekten S durch

$S = \{ x \mid \text{es gibt kein y mit } x \in y \text{ und } y \in x \}$.

Zeigen Sie, daß S keine Menge sein kann.

Übung 2 *(Komprehensionen mit Hilfe von Eigenschaften, II)*
Sei M eine Menge. Zeigen Sie mit Hilfe des Aussonderungsschemas, daß es eine Menge N gibt mit:

(i) Für alle $x \in N$ gilt $x \in M$.

(ii) $N \notin M$.

Übung 3 *(Einfache Mengenbildungen, I) (L)*
Zeigen Sie, daß für alle a, b, c, d gilt:

$(a, b) = (c, d)$ *gdw* $a = c$ und $b = d$.

Übung 4 *(Einfache Mengenbildungen, II)*
Wir setzen $0 = \emptyset$ und $1 = \{\emptyset\}$. Wir definieren für alle a, b:

$(a, b)^* = \{\{a, 0\}, \{b, 1\}\}$.

Zeigen Sie, daß auch diese alternative Paardefinition korrekt ist, d. h. für alle a, b, c, d gilt: $(a, b)^* = (c, d)^*$ *gdw* $a = c$ und $b = d$.

Übung 5 *(Operationen mit Mengen, I) (L)*
Zeigen Sie, daß für alle Mengen A, B, C gilt:

(i) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

(ii) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.

Vergleichen Sie (i) mit der Tautologie $(A \vee B) \wedge C \leftrightarrow (A \wedge C) \vee (B \wedge C)$. Diskutieren Sie allgemein den Zusammenhang zwischen Tautologien und Rechenregeln für Mengen. Welche Regel entspricht z. B. der Tautologie $\neg (A \vee B) \leftrightarrow \neg A \wedge \neg B$?

Übung 6 *(Operationen mit Mengen, II)*
Seien A_0, A_1, A_2 Mengen. Finden Sie Mengen $B_0 \subseteq A_0$, $B_1 \subseteq A_1$, $B_2 \subseteq A_2$ mit den Eigenschaften:

$A_0 \cup A_1 \cup A_2 = B_0 \cup B_1 \cup B_2$, $B_0 \cap B_1 = B_1 \cap B_2 = B_0 \cap B_2 = \emptyset$.

Übung 7 *(Operationen mit Mengen, III) (L)*
Sei A eine Menge. Bestimmen Sie jeweils alle Mengen B mit:

(i) $A \cap B = A$,

(ii) $A \cup B = A$.

(iii) $B - (B - A) = A$.

Übung 8 *(Operationen mit Mengen, IV)*
Zeigen Sie, daß für alle Mengen A, B gilt:

$A - B = A - (A \cap B)$.

Übung 9 *(Operationen mit Mengen, V)*
Zeigen oder widerlegen Sie, daß für alle Mengen A, B, C gilt:

(i) $A \Delta B = B \Delta A = (A \cup B) - (A \cap B)$,

(ii) $A \Delta (B \Delta C) = (A \Delta B) \Delta C$,

(iii) $(A \Delta B) \cap C = (A \cap C) \Delta (B \cap C)$.

(iv) $(A \Delta B) \cup C = (A \cup C) \Delta (B \cup C)$.

Übung 10 *(Operationen mit Mengen, VI) (L)*
Seien $A_1, \ldots, A_n$, $n \geq 2$, Mengen. Zeigen Sie:

$A_1 \Delta A_2 \Delta \ldots \Delta A_n = \{ a \mid$ die Anzahl aller i mit $a \in A_i$ ist ungerade $\}$.

Übung 11 *(Operationen mit Mengen, VII)*
Seien $A_1, \ldots, A_n$ Mengen mit $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n$. Dann gilt:

(i) $A_1 - \ldots - A_n = (A_1 - A_2) \cup \ldots \cup (A_{n-1} - A_n)$, *falls* n gerade,

(ii) $A_1 - \ldots - A_n = (A_1 - A_2) \cup \ldots \cup (A_{n-2} - A_{n-1}) \cup A_n$, *falls* n ungerade.

Hierbei ist die wiederholte Differenzenbildung mit Rechtsklammerung zu lesen, d. h. wir vereinbaren $A - B - C = A - (B - C)$ für alle A, B, C.

Übung 12 *(Potenzmengen, I)*
Zeigen Sie, daß für alle Mengen A, B gilt:

(i) $\emptyset, A \in \mathcal{P}(A)$,

(ii) $A \subseteq B$ *genau dann, wenn* $\mathcal{P}(A) \subseteq \mathcal{P}(B)$,

(iii) $\mathcal{P}(A) \in \mathcal{P}(\mathcal{P}(A))$.

Übung 13 *(Potenzmengen, II)*
Geben Sie alle Elemente von $\mathcal{P}(\mathcal{P}(\{ a \}))$ an.

Übung 14 *(Potenzmengen, III) (L)*
Zeigen oder widerlegen Sie, daß für alle Mengen A, B gilt:

(i) $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$,

(ii) $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$.

Übung 15 *(Potenzmengen, IV)*
Sei A eine Menge und sei $\mathcal{A} = \mathcal{P}(A)$. Zeigen Sie, daß für alle $X, Y \in \mathcal{A}$ ein eindeutiges $Z \in \mathcal{A}$ existiert mit den Eigenschaften:

(i) $X, Y \subseteq Z$,

(ii) $\forall Z' \in \mathcal{A}\ (X, Y \subseteq Z' \rightarrow Z \subseteq Z')$.

Übung 16 *(Mengensysteme, I)*
Zeigen Sie, daß für alle Mengen A, B gilt:

$\bigcap \{ A, B \} = A \cap B, \quad \bigcup \{ A, B \} = A \cup B.$

Übung 17 *(Mengensysteme, II)*
Warum wird $\mathcal{A} \neq \emptyset$ in der Definition von $\bigcap \mathcal{A}$ gefordert?
Welche Zusammenfassung wäre $\bigcap \emptyset$ nach der Definition von $\bigcap \mathcal{A}$?

Übung 18 *(Mengensysteme, III) (L)*
Zeigen oder widerlegen Sie, daß für alle Mengensysteme $\mathcal{A}, \mathcal{B}$ gilt:

(i) $\bigcup(\mathcal{A} \cup \mathcal{B}) = \bigcup \mathcal{A} \cup \bigcup \mathcal{B}$,

(ii) $\bigcap(\mathcal{A} \cap \mathcal{B}) = \bigcap \mathcal{A} \cap \bigcap \mathcal{B}$, *falls* $\mathcal{A} \cap \mathcal{B} \neq \emptyset$.

(iii) $\bigcup(\mathcal{A} \cap \mathcal{B}) = \bigcup \mathcal{A} \cap \bigcup \mathcal{B}$,

(iv) $\bigcap(\mathcal{A} \cup \mathcal{B}) = \bigcap \mathcal{A} \cup \bigcap \mathcal{B}$, *falls* $\mathcal{A}, \mathcal{B} \neq \emptyset$.

Übung 19 *(Mengensysteme, IV)*
Zeigen Sie, daß für alle Mengensystem $\mathcal{A}$ und $\mathcal{B}$ gilt:

(i) $\bigcup \mathcal{A} \cap \bigcup \mathcal{B} = \bigcup \{ A \cap B \mid A \in \mathcal{A}, B \in \mathcal{B} \}$,

(ii) $\bigcap \mathcal{A} \cup \bigcap \mathcal{B} = \bigcap \{ A \cup B \mid A \in \mathcal{A}, B \in \mathcal{B} \}$, *falls* $\mathcal{A}, \mathcal{B} \neq \emptyset$.

Übung 20 *(Mengensysteme, V)*
Sei $\mathcal{A} \neq \emptyset$ ein Mengensystem. Schreiben Sie $\bigcap \mathcal{A}$ als Aussonderung, d. h. finden Sie eine Menge M und eine Eigenschaft $\mathcal{E}(x)$ sodaß gilt

$\bigcap \mathcal{A} = \{ x \in M \mid \mathcal{E}(x) \}$.

Übung 21 *(Mengensysteme, VI)*
Sei A eine Menge, und seien $B, C, D \subseteq A$.

(i) Bestimmen Sie die kleinste Mengenalgebra $\mathcal{A}$ auf A mit $B \in \mathcal{A}$.

(ii) Bestimmen Sie den kleinsten Mengenverband $\mathcal{V}$ auf A mit $B, C, D \in \mathcal{V}$.

Übung 22 *(Mengensysteme, VII)*
Sei $A = \{ 1, 2, 3, 4 \}$, und seien $B = \{ 1 \}$, $C = \{ 2, 3 \}$.
Geben Sie die kleinste Mengenalgebra $\mathcal{A}$ auf A an mit $B, C \in \mathcal{A}$.

Übung 23 *(Mengensysteme, VIII) (L)*
Sei A eine Menge. Kann es eine Mengenalgebra $\mathcal{A}$ auf A geben, die genau fünf Elemente besitzt?

Übung 24 *(Mengensysteme, IX)*
Zeigen Sie, daß für alle Mengen A, B gilt:

$A \times B \in \mathcal{P}(\mathcal{P}(\mathcal{P}(A \cup B)))$.

3. Relationen und Funktionen

Die letzte der Sprachhürden, die wir zu nehmen haben, ist das Erlernen der Begriffe und Notationen im Umfeld der Relationen und der spezielleren Funktionen. Die elementaren Mengenbildungen zusammen mit dem Jargon der Relationen und Funktionen bilden den Grundstock des wissenschaftlichen mathematischen Sprechens. All diese Dinge werden in der Mathematik an allen Ecken und Enden verwendet.

Wir führen hier die Relationen und damit auch die Funktionen auf den Mengenbegriff zurück, und geben damit weitere Beispiele für eine mengentheoretische Interpretation allgemeiner mathematischer Konzepte. Dieses Vorgehen begeistert die einen durch seine Einfachheit und Klarheit („Endlich wird einem gesagt, was eine Funktion ist!"), die anderen stößt es eher ab („Das sind doch keine dynamischen Funktionen!"). Den letzteren sei noch einmal gesagt, daß wir keine fruchtbaren inneren Anschauungen verändern oder umbiegen wollen. Es geht um strukturelles, ökonomisches, präzises Definieren, wie es in der Mathematik ab einer gewissen Stufe ohnehin unumgänglich ist. Und warum sollte man diese Stufe dann nicht relativ weit unten ansetzen, wenn es einfach möglich ist? Die Notationen müßten wir ohnehin einführen.

Der Autor gehört sicher zur ersten Gruppe. Der einfache und klare mengentheoretische Relations- und Funktionsbegriff ist für ihn etwas Wunderbares. Die innere Anschauung bleibt davon unberührt, ja sie kann vor dem Hintergrund der formalen Definition erst richtig gewürdigt werden.

Relationen und ihre Struktureigenschaften

Definition *(Relation)*
Eine Menge R heißt eine *Relation*, falls jedes Element von R ein geordnetes Paar ist. R heißt *Relation auf* einer Menge A, falls $R \subseteq A \times A$.

Gilt $(a, b) \in R$, so sagen wir auch, daß a in der Relation R zu b steht. Anstelle von $(a, b) \in R$ schreiben wir auch a R b oder R(a, b). Gilt a R b und b R c, so schreiben wir auch a R b R c.

Für eine Relation R definieren wir:

$\mathrm{dom}(R) = \{ a \mid \exists b\ (a, b) \in R \}$, *(Definitionsbereich, „domain")*

$\mathrm{rng}(R) = \{ b \mid \exists a\ (a, b) \in R \}$, *(Wertebereich, „range")*

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_3, © Springer-Verlag Berlin Heidelberg 2010

Die Menge $\mathrm{field}(R) = \mathrm{dom}(R) \cup \mathrm{rng}(R)$ heißt das *Feld* von R. Das Feld von R ist die kleinste Menge A derart, daß R eine Relation auf A ist.

Weiter können wir für jede Relation R eine Umkehrrelation definieren:

$$R^{-1} = \{ (a, b) \mid (b, a) \in R \}. \qquad \textit{(Umkehrrelation)}$$

Der Übergang von R zu R^{-1} entspricht dem Übergang von „kleiner" zu „größer", von „höher" zu „niedriger", von „ist ein Kind von" zu „ist ein Elternteil von". Offenbar gilt

$$\mathrm{dom}(R^{-1}) = \mathrm{rng}(R), \quad \mathrm{rng}(R^{-1}) = \mathrm{dom}(R), \quad \mathrm{field}(R^{-1}) = \mathrm{field}(R).$$

Für Relationen gibt es eine Handvoll grundlegender Struktureigenschaften. So schließt zum Beispiel „a ist größer als b" aus, daß „b ist größer als a" gilt, und umgekehrt folgt aus „a ist ähnlich zu b", daß auch „b ist ähnlich zu a" gilt – andernfalls würde eine Relation den Namen „größer" bzw. „ähnlich" nicht verdient haben.

Die fünf wichtigsten Struktureigenschaften einer Relation sind nun die folgenden.

Definition *(Struktureigenschaften von Relationen)*
Eine Relation R auf A heißt:

(i) *reflexiv*, *falls* $\forall a \in A\ (a, a) \in R$,

(ii) *irreflexiv*, *falls* $\forall a \in A\ (a, a) \notin R$,

(iii) *symmetrisch*, *falls* $\forall a, b \in A\ (a\ R\ b \to b\ R\ a)$,

(iv) *antisymmetrisch*, *falls* $\forall a, b \in A\ (a\ R\ b \wedge b\ R\ a \to a = b)$,

(v) *transitiv*, *falls* $\forall a, b, c \in A\ (a\ R\ b \wedge b\ R\ c \to a\ R\ c)$.

Wir können uns eine Relation R auf A als eine „Punktwolke" im kartesischen Produkt $A \times A$ vorstellen: Wir tragen ein $(a, b) \in A \times A$ genau dann als Punkt ein, wenn a R b gilt. Dann bedeutet die Reflexivität von R, daß die gesamte Diagonale $\{ (a, a) \mid a \in A \}$ zur Punktwolke gehört. Die Symmetrie von R bedeutet, daß die Punktwolke invariant gegenüber der Spiegelung an der Diagonalen ist. Die Transitivität hat dagegen bei dieser Darstellung von R keine besonders anschauliche Bedeutung.

Eine ganz andere Möglichkeit, sich eine Relation R auf A zu visualisieren, ist die folgende: Wir fassen A als Menge von Punkten auf und verbinden diejenigen Punkte a, b mit einem Pfeil, für die a R b gilt. Nun hat die Transitivität von R eine sehr anschauliche Bedeutung, denn sie besagt: Gibt es einen Pfeil von a nach b und weiter einen Pfeil von b nach c in unserem Diagramm, so gibt es immer auch einen Pfeil von a nach c. Die Symmetrie besagt, daß es zu jedem Pfeil von a nach b auch immer den umgekehrten Pfeil von b nach a gibt. Wir können im symmetrischen Fall also die Pfeile ganz weglassen und einfach alle a, b, für die a R b gilt, mit einer Linie verbinden. Die Irreflexivität von R besagt schießlich, daß unser Diagramm keine „Schlingen" besitzt, die von einem Punkt a zu a selbst führen. Diese Form der Visualisierung ist vor allem für Relationen auf endlichen Men-

gen sehr nützlich und wir werden im Kapitel über Graphen darauf zurückkommen.

Der Leser kann bereits an dieser Stelle die Tragweite des Konzepts sehen: Relationen decken abstrakte mathematische Dinge wie die Element-Beziehung ebenso ab wie die Modellierung von konkreten Verkehrsnetzen. Sobald wir die Beziehungen betrachten, die zwischen Objekten eines Bereichs bestehen, werden wir Relationen definieren und untersuchen. Damit treten Relationen überall auf, wo es um mehr geht als die reine Anzahl von Dingen.

Wir führen noch einige Sprechweisen und Notationen für mehrstellige Relationen ein. Eine Menge R heißt eine *dreistellige Relation auf A*, falls $R \subseteq A^3$. Statt $(a_1, a_2, a_3) \in R$ schreiben wir auch $R(a_1, a_2, a_3)$. Analog sind 4-, 5-, … stellige Relationen definiert. Konsequent nennen wir ein $R \subseteq A$ dann auch eine *einstellige Relation* auf A und schreiben R(a) für $a \in R$. Hier wird die Wortbedeutung „Relation" („Beziehung") mißbraucht und man spricht auch deswegen oft von einem *(einstelligen) Prädikat* auf A. Allgemein werden die Begriffe *Prädikat* und *Relation* in der Mathematik synonym verwendet.

Es gibt drei Haupttypen von Relationen: Äquivalenzrelationen, Ordnungen und Funktionen. Diese drei Typen spielen eine universelle Rolle in der Mathematik, und wir wollen sie der Reihe nach vorstellen.

Äquivalenzrelationen

In der Mathematik tauchen häufig Situationen auf, in denen Objekte, die sich nur in als unwesentlich betrachteten Eigenschaften unterscheiden, einander gleichgestellt werden. Kommt es uns zum Beispiel nur auf den ganzzahligen Rest der Division einer natürlichen Zahl durch 7 an, so betrachten wir die Zahlen 4, 11, 18, … als äquivalent, oder, wie man in diesem Fall sagt, als kongruent modulo 7. Derartige Gleichstellungen sind mathematisch von großer Bedeutung, denn sie sind das Ergebnis von Abstraktionsvorgängen. Wenn wir abstrahieren, sehen wir von bestimmten Merkmalen der betrachteten Objekte ab (das lateinische „abstrahere" bedeutet „absehen von, wegnehmen"). Nach diesem Abstreifen von gewissen Eigenschaften sind zwei Objekte a und b, die sich nur in diesen Eigenschaften unterscheiden, gleichwertig und austauschbar geworden – vorausgesetzt, unsere weiteren Betrachtungen kehren nicht zu den vergessenen Merkmalen zurück.

Auch außerhalb der Mathematik ist diese Form der Gleichstellung nicht unbekannt. Unsere zyklische Einteilung der Wochentage entspricht dem oben erwähnten Rechnen „modulo 7". Weiter sagen wir zum Beispiel: „Sie fahren das gleiche Auto wie ich". Dabei meinen wir den Typ des Autos, auf das Nummernschild kommt es uns dabei nicht an. Eine Antwort könnte sein: „Ja, wenn man von der Farbe absieht". Der Angesprochene äußert damit ein etwas feineres Verständnis von „gleiches Auto", das vom Merkmal der Farbe nicht absehen will. Allgemein kennt unsere Umgangssprache den feinen Unterschied zwischen „das Gleiche" und „dasselbe". Man kann im Restaurant das gleiche Essen bestellen

wie der Nachbar, nicht aber dasselbe. (In der Mathematik wird diese Unterscheidung vermieden, „gleich“ und „identisch“ werden hier synonym verwendet und mit dem Zeichen „=“ ausgedrückt.)

Auch Klassifikationen kennen wir aus dem Alltag: Wir teilen auf in Männer und Frauen, in Kinder, Berufstätige und Rentner, in Selbständige und Arbeitnehmer, in Notenstufen von „sehr gut“ bis „ungenügend“. Jeden Klassifizierungsvorgang können wir auch als Abstraktionsvorgang auffassen. Wenn wir zum Beispiel von der individuellen Person absehen und uns nur für ihr Geschlecht interessieren, so erhalten wir eine Klasseneinteilung einer betrachteten Gruppe in Männer und Frauen. Umgekehrt führt jede Klassifikation zu einer Abstraktion, indem genau diejenigen Objekte einander gleichgestellt werden, die einer gemeinsamen Klasse angehören.

Mathematisch wird der Abstraktionsvorgang durch den Begriff der Äquivalenzrelation und der Begriff der Klassifikation durch den Begriff der Zerlegung einer Menge gefaßt. Beide Begriffe hängen, wie obige informale Diskussion schon andeutet, eng miteinander zusammen.

Wir beginnen mit dem Äquivalenzbegriff. Er versammelt genau die Eigenschaften, die wir von jedem Abstraktionsvorgang erwarten dürfen:

Definition *(Äquivalenzrelation)*
Eine Relation $\sim$ auf A heißt eine *Äquivalenzrelation*, falls $\sim$ reflexiv, symmetrisch und transitiv ist.

Ist $\sim$ eine Äquivalenzrelation auf A, so setzen wir für alle Elemente a von A:

$a/\sim \;=\; \{\, b \in A \mid a \sim b \,\}$. *(Äquivalenzklasse von a, „a modulo $\sim$“)*

Damit setzen wir dann:

$A/\sim \;=\; \{\, a/\sim \mid a \in A \,\}$. *(Faktorisierung, „A modulo $\sim$“)*

Eine Äquivalenzklasse $a/\sim$ ist also eine nichtleere Teilmenge von A, und die Faktorisierung $A/\sim$ ist ein Mengensystem auf A. Die wesentlichen Eigenschaften dieser Objekte sind:

(i) $a/\sim \;=\; b/\sim$ *gdw* $a \sim b$,

(ii) $a/\sim \,\cap\, b/\sim \;=\; \varnothing$ *gdw* $\text{non}(a \sim b)$,

(iii) $\bigcup A/\sim \;=\; A$.

Eine Äquivalenzrelation auf A teilt also die Menge A in gewisse nichtleere und paarweise disjunkte „Bereiche“ oder „Regionen“ ein (die Äquivalenzklassen der Relation). Die Faktorisierung ist die Menge all dieser Bereiche. Elemente, die demselben Bereich angehören, sehen wir als gleichwertig an. Eine Äquivalenzrelation ist damit wie gewünscht eine „unscharfe“ oder „vergröberte“ mathematische Gleichheit.

Wir diskutieren nun noch den engen Zusammenhang zwischen Gleichsetzungen und Klassifikationen. Hierzu definieren wir:

Definition *(Zerlegung)*
Ein Mengensystem $\mathfrak{Z}$ auf A heißt eine *Zerlegung* oder *Klasseneinteilung* von A, falls gilt:

(a) $B \neq \emptyset$ für alle $B \in \mathfrak{Z}$,

(b) $B \cap C = \emptyset$ für alle $B, C \in \mathfrak{Z}$ mit $B \neq C$,

(c) $\bigcup \mathfrak{Z} = A$.

Eine Zerlegung von A ist also ein System von nichtleeren und paarweise disjunkten Mengen derart, daß jedes Element von A in einem (und folglich genau einem) Element der Zerlegung vorkommt.

Nach den obigen Eigenschaften (i) – (iii) ist für jede Äquivalenzrelation $\sim$ auf A die Faktorisierung $A/\sim$ eine Zerlegung von A. Ist nun $\mathfrak{Z}$ eine beliebige Zerlegung von A, so setzen wir für alle $a, b \in A$:

$a \sim_{\mathfrak{Z}} b$, *falls* „es gibt ein $Z \in \mathfrak{Z}$ mit $a, b \in Z$".

Es ist leicht zu sehen, daß $\sim_{\mathfrak{Z}}$ eine Äquivalenzrelation auf A ist. Zudem liefert die Faktorisierung wieder die ursprüngliche Zerlegung, d. h. es gilt $A/\sim_{\mathfrak{Z}} = \mathfrak{Z}$.

In der Faktorisierung $A/\sim = \{ a/\sim \mid a \in A \}$ tauchen im allgemeinen die „Beiträge" $a/\sim$ der Mengenkomprehension mehrfach auf. Gilt $a \sim b$, so ist $a/\sim = b/\sim$ und wir können die Komprehension z. B. auch ohne b durchführen. Diese Beobachtung führt uns zur Frage der „minimalen Erzeugung" von $A/\sim$. Hierzu definieren wir:

Definition *(Repräsentanten)*
Sei $\sim$ eine Äquivalenzrelation auf A, und sei $a \in A$. Dann heißt jedes $b \in a/\sim$ ein *Repräsentant* der Äquivalenzklasse $a/\sim$.
Eine Teilmenge S von A heißt ein *vollständiges Repräsentantensystem*, falls S genau einen Repräsentanten aus jeder Äquivalenzklasse enthält.

Für ein vollständiges Repräsentantensystem S gilt also:

(a) $A/\sim = \{ a/\sim \mid a \in S \}$,

(b) $\text{non}(a \sim b)$ für alle $a, b \in S$ mit $a \neq b$.

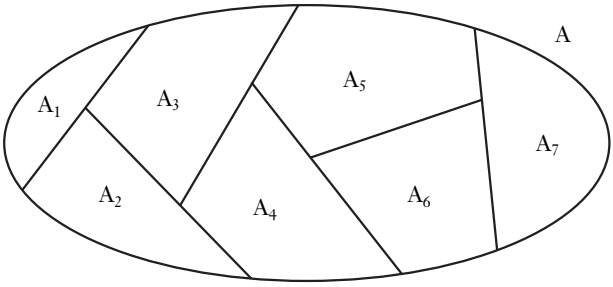
Die Menge S ist also ein minimaler Erzeuger der Faktorisierung $A/\sim$.

In vielen Fällen läßt sich ein vollständiges Repräsentantensystem einfach finden. So ist zum Beispiel 0, 1, …, 5, 6 ein derartiges System für das Rechnen modulo 7. Manchmal ist ein vollständiges Repräsentantensystem allerdings nicht zu sehen. Ein Beispiel ist die sog. *Vitali-Äquivalenzrelation* auf den reellen Zahlen, die definiert wird durch

$x \sim y$, *falls* „der Abstand zwischen x und y ist rational".

In der Tat ist der Satz, daß jede Äquivalenzrelation ein vollständiges Repräsentantensystem besitzt, ein Axiom der Mathematik. (Er ist elementar äquivalent zum sog. *Auswahlaxiom*.) Damit hat uns der Begriff der Äquivalenzrelation schon zu Fragen geführt, die die Grundlagen der Mathematik betreffen.

Wir betrachten zur Illustration und Visualisierung der Begriffe eine Zerlegung $\mathscr{Z}$ einer Menge A in sieben nichtleere Teilmengen:



Zerlegung einer Menge A in Mengen $A_1, \ldots, A_7$

Es gilt also $\mathscr{Z} = \{ A_1, \ldots, A_7 \}$, $\bigcup \mathscr{Z} = A_1 \cup \ldots \cup A_7 = A$, $A_i \neq \emptyset$ für alle i und $A_i \cap A_j = \emptyset$ für alle $i \neq j$.

Weiter sei $\sim$ die durch die Zerlegung $\mathscr{Z}$ definierte Äquivalenzrelation $\sim$ auf A, d. h. wir setzen für alle $a, b \in A$:

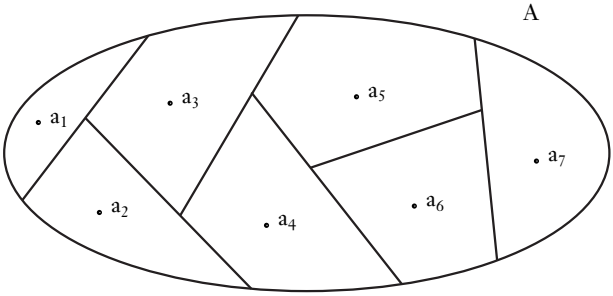
$a \sim b$, *falls* „es gibt ein $1 \leq i \leq 7$ mit $a, b \in A_i$“.

Dann gilt

$A/\sim \; = \mathscr{Z} = \{ A_1, \ldots, A_7 \}$,

d. h. die Faktorisierung $A/\sim$ ist gerade wieder die Zerlegung $\mathscr{Z}$.

Zur Bildung eines vollständigen Repräsentantensystems wählen wir aus jedem „Land“ der Zerlegung genau einen „Bewohner“ aus:



Bildung eines vollständigen Repräsentantensystems für $\sim$

Für jede solche Wahl $a_1 \in A_1, \ldots, a_7 \in A_7$ ist die Menge $\{ a_1, \ldots, a_7 \} \subseteq A$ ein vollständiges Repräsentantensystem unserer Äquivalenzrelation. Es gilt $a_i/\sim = A_i$ für alle i, also $A/\sim \; = \{ a_1/\sim, \; \ldots, \; a_7/\sim \}$.

Ordnungen

Neben den Äquivalenzrelationen gehören die Ordnungen zu den wichtigsten Relationen. Auch diese sind uns durch den Alltag sehr vertraut: Wir sprechen in den verschiedensten Situation von „besser“ und „schlechter“, von „schneller“ und „langsamer“, von „größer“ und „kleiner“. Wir unterliegen dabei leider oft einem Bedürfnis, je zwei Objekte miteinander vergleichen zu wollen. Natürlicher ist in vielen Fällen ein „partieller“ Ansatz, der zwei vorliegende Dinge nur manchmal miteinander vergleicht und andernfalls gar keine Aussage trifft. Diese Ordnungsidee können wir mathematisch wieder durch eine Kombination von Eigenschaften einer Relation einfangen:

Definition *(partielle Ordnung)*
Eine Relation $\leq$ auf A heißt eine *partielle Ordnung (vom nicht strikten Typ)* auf A, falls $\leq$ reflexiv, antisymmetrisch und transitiv ist.

Ausformuliert gilt also für alle $a, b, c \in A$:

$a \leq a$, $a \leq b \wedge b \leq a$ *impliziert* $a = b$, $a \leq b \leq c$ *impliziert* $a \leq c$.

Das bevorzugte Zeichen für partielle Ordnungen ist $\leq$, zusammen mit verwandten Symbolen wie $\preccurlyeq$, $\leq^*$, usw.

Stillschweigend wird für eine partielle Ordnung $\leq$ immer definiert:

$a < b$, *falls* $a \leq b$ und $a \neq b$ für alle $a, b \in A$.

Die Relation $<$ ist irreflexiv und transitiv auf A. Eine irreflexive und transitive Relation heißt eine *partielle Ordnung (vom strikten Typ)*. Ist $<$ eine strikte partielle Ordnung auf A, so definieren wir

$a \leq b$, *falls* $a < b$ oder $a = b$ für alle $a, b \in A$.

Dann ist die Relation $\leq$ eine partielle Ordnung auf A vom nicht strikten Typus.

Wir haben also immer partielle Ordnungen beider Typen vorliegen. Die Zeichen $\leq$, $\preccurlyeq$, $\leq^*$, … stehen immer für nichtstrikte Typen, die zugehörigen Zeichen $<$, $\prec$, $<^*$, … für die strikten Versionen.

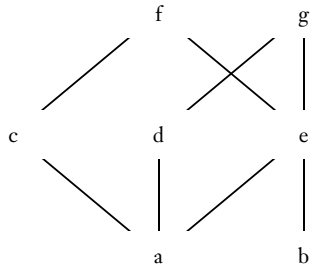
Nichtstrikte partielle Ordnungen unterscheiden sich von den Äquivalenzrelationen „nur“ durch den Austausch der Bedingungen „symmetrisch“ und „antisymmetrisch“. Dieser Austausch führt zu einer vollkommen anderen Welt. Partielle Ordnungen zerlegen eine Menge nicht in Teilgebiete, sondern sie ordnen die Elemente der Menge in einer netzartigen Struktur an. Diese Struktur kann man für endliche Mengen A mit Hilfe von Diagrammen sehr anschaulich machen: Wir zeichnen die Elemente von A als benannte Punkte und verbinden zwei Punkte $a < b$ genau dann mit einem Pfeil, wenn es kein c gibt mit $a < c < b$. Dieser Zusatz führt zu übersichtlichen Diagrammen, die nicht mit Transitivitätspfeilen überladen sind. Weiter kann man statt der Pfeile auch einfache Verbindungslinien einzeichnen, wenn man vereinbart, daß Punkte, die weiter oben stehen, größer sind als Elemente, die weiter unten stehen.

So gilt im Diagramm rechts:

$a < c < f$, $a < d < g$,

$a < e < f$, $a < e < g$,

$b < e < f$, $b < e < g$.



Darüber hinaus gelten keine weiteren Relationen, so gilt z. B. weder $a < b$ noch $d < f$.

Zur Beschreibung der Struktur einer partiellen Ordnung führen wir die folgenden Begriffe ein:

Definition *(vergleichbar, Kette, Antikette)*
Sei $\leq$ eine partielle Ordnung auf A.

(a) $a, b \in A$ heißen *vergleichbar*, falls $a \leq b$ oder $b \leq a$ gilt. Andernfalls heißen a und b *unvergleichbar*.

(b) Ein $B \subseteq A$ heißt eine *Kette*, falls je zwei Elemente von B vergleichbar sind.

(c) Ein $C \subseteq A$ heißt eine *Antikette*, falls je zwei verschiedene Elemente von C unvergleichbar sind.

(d) Eine Kette B heißt *maximal*, falls es keine Kette C gibt mit $B \subset C$. Analog sind maximale Antiketten definiert.

In der partiellen Ordnung des obigen Diagramms sind zum Beispiel a und g vergleichbar, a und b sind dagegen unvergleichbar. Die Menge { a, d, g } ist eine maximale Kette und die Menge { c, d, e } eine maximale Antikette der Ordnung.

Weiter zeichnen wir noch gewisse Elemente von partiellen Ordnungen aus:

Definition *(maximales, minimales, größtes, kleinstes Element)*
Sei $\leq$ eine partielle Ordnung auf A.

(a) Ein $a \in A$ heißt *maximal* (bzw. *minimal*), falls es kein b gibt mit $a < b$ (bzw. $b < a$).

(b) Ein $a \in A$ heißt das *größte* (bzw. *kleinste*) Element der Ordnung, falls $b \leq a$ (bzw. $a \leq b$) für alle $b \in A$.

Im obigen Diagramm sind a und b die minimalen Elemente und f und g die maximalen Elemente der Ordnung. Die Ordnung hat weder ein größtes noch ein kleinstes Element.

Die Inklusion liefert die wichtigsten Beispiele für partielle Ordnungen. Für jedes Mengensystem $\mathcal{A}$ auf einer Menge A ist die Teilmengenrelation $\subseteq$ eine partielle Ordnung auf $\mathcal{A}$, und die echte Inklusion $\subset$ ist die zugehörige strikte Version.

Einen schärferen Begriff als die partiellen Ordnungen bilden die sogenannten linearen Ordnungen.

Definition *(lineare Ordnung)*
Eine partielle Ordnung auf A heißt eine *lineare* oder *totale Ordnung* auf A, falls je zwei Elemente von A vergleichbar sind.

Für eine partielle Ordnung $\leq$ (oder $<$) auf A gilt also:

$\forall a, b \in A\ (a \leq b \ \vee \ b \leq a)$. *(Vergleichbarkeitsbedingung)*

Die Elemente von A werden in einer linearen Ordnung nicht mehr netzartig, sondern in einer abstrakten „Reihe“ oder „Linie“ angeordnet. In waagrechten Diagrammen vereinbaren wir dann, daß die Elemente von links nach rechts „wachsen“. So gilt $a < b < c < d$ in der wie folgt visualisierten Ordnung:

a —— b —— c —— d

Funktionen

Funktionen gehören zu den prominentesten Objekten der Mathematik überhaupt. Intuitiv ordnet eine Funktion f einem Objekt a ein neues Objekt b eindeutig zu. Wir schreiben dann $b = f(a)$. Ein präziser Funktionsbegriff, der nicht von Umrechnungen oder Zuordnungen sprechen muß, läßt sich nun aus dem Relationsbegriff relativ einfach gewinnen. Wir betrachten hierzu zwei weitere natürliche Struktureigenschaften einer Relation:

Definition *(rechts- und linkseindeutige Relationen)*
Eine Relation R heißt *rechtseindeutig*, falls es für alle $a \in \mathrm{dom}(R)$ genau ein b gibt mit $a\ R\ b$. Analog heißt R *linkseindeutig*, falls es für alle $b \in \mathrm{rng}(R)$ genau ein a gibt mit $a\ R\ b$.

In einem Punkt-Pfeil-Diagramm einer Relation bedeutet die Rechtseindeutigkeit, daß von jedem Punkt höchstens ein Pfeil wegführt und die Linkseindeutigkeit, daß zu jedem Punkt höchstens ein Pfeil hinführt. Bei dieser Visualisierung ist die folgende Definition des modernen mathematischen Funktionsbegriffs dann keine Überraschung mehr:

Definition *(Funktion)*
Eine Relation f heißt eine *Funktion* oder *Abbildung*, falls f rechtseindeutig ist. Eine Funktion f heißt eine Funktion *auf* einer Menge A, falls $\mathrm{dom}(f) = A$.

Damit haben wir Funktionen als gewisse Relationen, Relationen als Mengen von geordneten Paare und geordnete Paare als gewisse Mengen eingeführt. Unsere Anschauungen von Paarbildungen, Beziehungen und funktionalen Zuordnungen sind in diesen Definitionen vielfach nur als Schatten vorhanden. Das kann bei derart fundamentalen Begriffsbildungen wohl nicht anders sein, wichtig

ist aber, daß eine derartige Reduktion, die alle erforderlichen Eigenschaften zur Verfügung stellt, überhaupt möglich ist. Wir haben eine präzise, klare und letztendlich sehr einfache Definition gefunden, die nicht von „Zuordnungen", „Berechnungen", „Pfeilen" usw. sprechen muß.

Definition *(die Notation f(a) = b)*

Ist f eine Funktion, so schreiben wir

$f(a) = b$ *anstelle von* $(a, b) \in f$,

oder auch $a \overset{f}{\longmapsto} b$.

Gilt $f(a) = b$, so sagen wir, daß a durch f auf b *abgebildet* wird und nennen b den *Funktionswert* von f an der *Stelle* a oder für das *Argument* a.

Ist dom(f) von der Form A^2, so nennen wir f auch eine *zweistellige Funktion* auf A. Für alle $(a_1, a_2) \in A^2$ schreiben wir dann oft kurz $f(a_1, a_2)$ anstelle von $f((a_1, a_2))$. Analoges gilt für drei- und mehrstellige Funktionen.

Ist f eine Funktion und $A \subseteq \mathrm{dom}(f)$, so definieren wir die *Einschränkung* $f|A$ von f auf A durch

$(f|A)(a) = f(a)$ für alle $a \in A$.

In Mengenschreibweise gilt also einfach $f|A = \{(a, b) \in f \mid a \in A\}$. Die Funktion $f|A$ hat den Definitionsbereich A und nimmt auf A dieselben Funktionswerte an wie f.

Für jede Menge A definieren wir die *Identität* id_A auf A durch

$\mathrm{id}_A(a) = a$ für alle $a \in A$.

Schließlich können wir für jede Menge A und jedes c die *konstante Funktion* const_c^A auf A mit *Wert* c definieren durch

$\mathrm{const}_c^A(a) = c$ für alle $a \in A$.

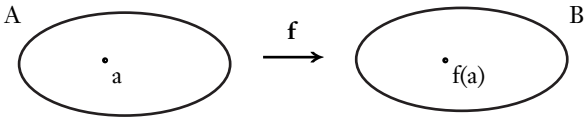
Die folgende Notation für Funktionen wird durchgehend verwendet:

Definition *(die Notation $f : A \to B$)*

Ist f eine Funktion, so schreiben wir

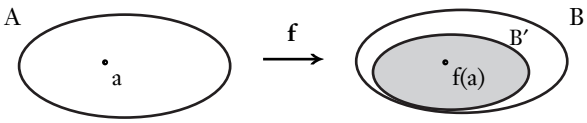
$f : A \to B$, *falls* $\mathrm{dom}(f) = A$ und $\mathrm{rng}(f) \subseteq B$.

Die Funktion f heißt dann eine Funktion *von* A *nach* B, und die Menge B ein *Wertevorrat* von f.



eine Funktion $f : A \to B$

Daß wir in der Notation $f : A \to B$ zwar $dom(f) = A$ verlangen, nicht aber $rng(f) = B$, ist eine bewährte Konvention.



eine Funktion $f : A \to B$ mit Wertebereich $B' = rng(f)$;
im allgemeinen ist B' eine echte Teilmenge des Wertevorrats B

Es gilt z. B. $id_A : A \to A$ und $const_c^A : A \to A$, falls $c \in A$. In jedem Falle gilt $const_c^A : A \to \{ c \}$.

Zur Notation $f : A \to B$ gehört weiter die folgende Mengenoperation:

Definition *(^AB)*

Für Mengen A und B setzen wir

$$^AB = \{ f \mid f : A \to B \}.$$

Weiter betrachten wir den Fall, bei dem Definitionsbereich und Wertevorrat übereinstimmen:

Definition *(Operation, abgeschlossen)*

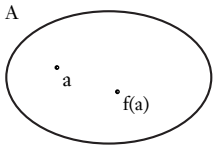
Eine Funktion $f : A \to A$ nennen wir auch eine *Operation* auf A.

Ebenso heißt $f : A^2 \to A$ eine *zweistellige Operation* auf A, usw.

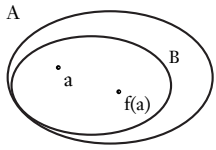
Ein $B \subseteq A$ heißt *abgeschlossen* unter einer n-stelligen Operation f, falls gilt:

$f(a_1, \ldots, a_n) \in B$ für alle $a_1, \ldots, a_n \in B$.

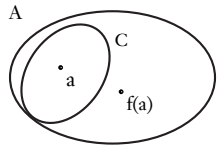
Ein $B \subseteq A$ ist genau dann abgeschlossen unter f, wenn wir $f \mid B^n : B^n \to B$ schreiben können. Die Funktion führt dann also aus der Menge B nicht heraus, wenn sie für Argumente in B ausgewertet wird.



eine Operation f auf A,
d. h. es gilt $f : A \to A$



unter f abgeschlossenes $B \subseteq A$,
d. h. $f(a) \in B$ für alle $a \in B$



unter f nicht abgeschlossenes $C \subseteq A$,
d. h. es gibt ein $a \in C$ mit $f(a) \notin C$

Für alle A ist ^AA die Menge aller einstelligen und $^{A \times A}A$ die Menge aller zweistelligen Operationen auf A.

Struktureigenschaften von Funktionen

Ist $f: A \to B$ eine Funktion, so ist das durch f gegebene Abbildungsverhältnis zwischen A und B von Interesse: Wird jeder Funktionswert nur einmal angenommen? Wird jeder Wert im Wertevorrat B als Funktionswert angenommen?

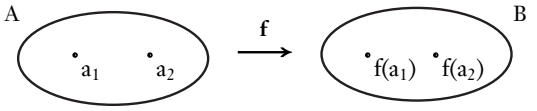
Definition *(injektiv, surjektiv, bijektiv)*
Eine Funktion $f: A \to B$ heißt:

(a) *injektiv*, *falls* „f ist linkseindeutig“,

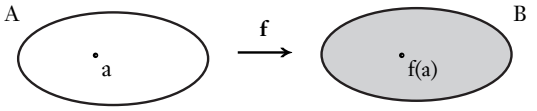
(b) *surjektiv*, *falls* $\mathrm{rng}(f) = B$,

(c) *bijektiv*, *falls* „f ist injektiv und surjektiv“.

Injektive Funktionen nehmen also keinen Wert zweimal an, surjektive Funktionen nehmen jeden Wert des ins Auge gefaßten Wertevorrats an, und Bijektionen sind vollständige Paarbildungen zwischen den Elementen ihres Definitionsbereichs und ihres Wertevorrats.



eine injektive Funktion $f: A \to B$, d.h es gilt $f(a_1) \neq f(a_2)$ für alle $a_1 \neq a_2$



eine surjektive Funktion $f: A \to B$, d.h. $B = rng(f)$

Die Injektivität von f können wir in Funktionsschreibweise auch so ausdrücken:

$\forall a, b \in A\,(f(a) = f(b) \to a = b)$.

Diese Bedingung ist unabhängig von der Angabe eines Wertevorrats. Dagegen benötigen die Eigenschaften „surjektiv“ und „bijektiv“ die Angabe eines Wertevorrats B.

Ist $f: A \to B$ injektiv, so ist $f: A \to \mathrm{rng}(f)$ bijektiv. Weiter können wir für injektive Funktionen eine Umkehrfunktion erklären:

Definition *(Umkehrfunktion)*
Ist f injektiv, so heißt die Relation f^{-1} die *Umkehrfunktion* von f.

Gilt dann $f(a) = b$, so gilt $f^{-1}(b) = a$. Weiter ist rng(f) der Definitionsbereich von f^{-1} und $f^{-1}: \mathrm{rng}(f) \to A$ ist bijektiv.

Verknüpfung von Funktionen

Zwei Funktionen können wir miteinander verknüpfen, wenn die zweite Funktion auf allen Funktionswerten der ersten definiert ist:

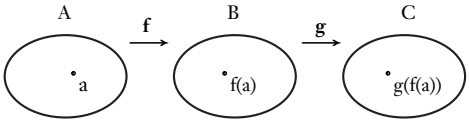
Definition *(Komposition)*
Seien $f : A \to B$ und $g : B \to C$ Funktionen.
Dann definieren wir die *Komposition*
oder *Verknüpfung*

$h = g \circ f$

durch

$h(a) = g(f(a))$ für alle $a \in A$.

Wir lesen $g \circ f$ als „g nach f".



Verknüpfung $g \circ f$ von $f : A \to B$ und $g : B \to C$

Für $h = g \circ f$ gilt $h : A \to C$. Sind f und g injektiv, so ist auch h injektiv. Gleiches gilt für „surjektiv" und „bijektiv".

Ist f eine injektive Funktion auf A, so gilt $f^{-1} \circ f = id_A$. Gilt umgekehrt $g \circ f = id_A$ für Funktionen $f : A \to B$ und $g : B \to A$, so ist die Funktion f injektiv und zudem gilt $g \mid rng(f) = f^{-1}$.

Mengen als Funktionswerte

Unser Funktionsbegriff ist durch eindeutige Zuordnungen bestimmt. In vielen Fällen sind punktweise gelesene Zuordnungen nicht eindeutig. So besitzen viele Menschen heute mehrere Telefonnummern und mehrere email-Adressen, reelle Funktionen können mehrere Nullstellen besitzen (oder auch gar keine), usw. Will man in einer solchen relationalen Situation einen funktionalen Kontext aufrecht erhalten, so kann man mit Mengen arbeiten und die Mehrdeutigkeiten zu einem Objekt zusammenfassen. Ist zum Beispiel $\mathcal{F}$ eine Menge von Funktionen von A nach B, so können wir für jedes $b \in B$ eine Funktion $G : \mathcal{F} \to \mathcal{P}(A)$ definieren durch

$G(f) = \{ a \in A \mid f(a) = b \}$ für alle $f \in \mathcal{F}$.

Diese Funktion ordnet also jeder Funktion f in $\mathcal{F}$ die Menge ihrer Argumente zu, die auf b abgebildet werden.

Allgemeiner können wir jeder Relation R mit $dom(R) \subseteq A$ und $rng(R) \subseteq B$ eine Funktion $G : A \to \mathcal{P}(B)$ zuordnen durch

$G(a) = \{ b \in B \mid a \, R \, b \}$ für alle $a \in A$.

Aus der Funktion G läßt sich dann die Relation R zurückgewinnen, denn es gilt $a \, R \, b$ genau dann, wenn $b \in G(a)$.

Bilder und Urbilder

Jede Funktion $f: A \to B$ gibt Anlaß zu zwei Operationen auf den Potenzmengen von A bzw. B. Gegeben eine Teilmenge X von A können wir die Menge der Funktionswerte $\{ f(x) \mid x \in X \}$ betrachten. Umgekehrt können wir für eine Teilmenge Y von B die Menge aller $x \in A$ betrachten, deren Funktionswerte in Y liegen:

Definition *(Bild, Urbild)*
Sei $f: A \to B$. Dann setzen wir für jedes $X \subseteq A$ und jedes $Y \subseteq B$:

$f[X] = \{ f(x) \mid x \in X \}$, *(Bild von X)*

$f^{-1}[Y] = \{ a \in A \mid f(a) \in Y \}$. *(Urbild von Y)*

Speziell ist also $f[A] = \mathrm{rng}(f)$ und $f^{-1}[B] = A$.

Das Urbild einer Menge ist immer definiert, unabhängig davon, ob f injektiv ist, d. h. unabhängig davon, ob die Umkehrfunktion f^{-1} existiert.

Die Folgennotation

Wir führen eine weitere Notation für Funktionen ein, die anschaulich dem Beschriften von Objekten entspricht. Haben wir 100 Umzugskartons vorliegen, so können wir sie z. B. mit den Zahlen 0, …, 99 beschriften. Wir sprechen dann von „Karton Nr. 75“ und haben so der Zahl 75 einen gewissen Karton k_{75} zugeordnet. Allgemeiner sind beliebige „Indizes“ zur Beschriftung möglich.

Diese Anschauung können wir formal nun so fassen:

Definition *(Folgennotation)*
Ein Ausdruck $\langle x_i \mid i \in I \rangle$ bedeutet die Funktion f mit den Eigenschaften:

(a) $\mathrm{dom}(f) = I$,

(b) $f(i) = x_i$ für alle $i \in I$.

Die Funktion $\langle x_i \mid i \in I \rangle$ heißt dann eine *Folge* oder *Familie* mit *Indexmenge* I.

Neben $\langle x_i \mid i \in I \rangle$ sind auch Varianten wie $(x_i)_{i \in I}$ gebräuchlich.

Jede Funktion f können wir als Folge schreiben, denn es gilt

$f = \langle f(i) \mid i \in \mathrm{dom}(f) \rangle$.

Obwohl Folgen und Funktionen identische Begriffe sind, bewährt sich die Notation in der Praxis. Ob ein Indizieren der Form x_i oder die klassische Funktionsnotation f(i) bevorzugt wird ist oft Geschmackssache. Die Folgenschreibweise wird insbesondere dann gern verwendet, wenn I die Menge der natürlichen Zahlen ist.

Die Indexschreibweise wird dann auch in suggestiver Weise an anderen Stellen eingesetzt. So setzen wir für eine Folge $\langle M_i \mid i \in I \rangle$ von Mengen:

$\bigcap_{i \in I} M_i = \bigcap \{ M_i \mid i \in I \}$, $\bigcup_{i \in I} M_i = \bigcup \{ M_i \mid i \in I \}$.

Weiter definieren wir:

Definition *(allgemeines Produkt)*
Sei $\langle B_i \mid i \in I\rangle$ eine Folge von Mengen. Dann setzen wir

$$\times\langle B_i \mid i \in I\rangle = \{ f \mid f \text{ ist Funktion auf } I,\ f(i) \in B_i \text{ für alle } i \in I \}.$$

Statt $\times\langle B_i \mid i \in I\rangle$ schreiben wir auch wieder $\times_{i\in I} B_i$. Ist $B_i = B$ für alle $i \in I$, so ist das Produkt $\times_{i\in I} B_i$ gleich der Funktionenmenge ${}^I B$.

Die Aussage, daß $\times_{i\in I} B_i$ nichtleer ist, falls alle B_i nichtleere Mengen sind, ist äquivalent zur Aussage, daß jede Äquivalenzrelation ein vollständiges Repräsentantensystem besitzt. Wir können das Nichtverschwinden des allgemeinen Kreuzprodukts für nichtleere Faktoren also wieder als Basis-Axiom der Mathematik ansehen.

Wohldefiniertheit und Kongruenzrelationen

Gegeben sei eine Menge A und eine Äquivalenzrelation $\sim$ auf A. Die Relation $\sim$ zerlegt die Menge A in die Äquivalenzklassen $a/\sim$, $a \in A$. Oft wird nun mit den Äquivalenzklassen als „Punkten" gerechnet und nicht mehr mit den ursprünglichen Elementen von A, d. h. es werden Funktionen eingeführt, die auf der Faktorisierung $A/\sim$ definiert wird. Nicht selten wird aber trotzdem der Funktionswert $f(a/\sim)$ mit Hilfe von a und nicht mit Hilfe der Äquivalenzklasse $a/\sim$ definiert. Es ist dann immer die *Wohldefiniertheit* von f oder die sog. *Unabhängigkeit von der Wahl der Repräsentanten* zu zeigen, d. h. man muß zeigen:

Für alle $a, b \in A$ mit $a \sim b$ gilt $f(a/\sim) = f(b/\sim)$. *(Wohldefiniertheit von f)*

Diese Voraussetzung spiegelt genau wider, daß die Funktion den Abstraktionsvorgang respektiert, der zur Bildung der Äquivalenzklassen geführt hat.

Ist umgekehrt $f : A \to A$ eine bereits definierte Operation auf A, so wird man fragen, ob die Äquivalenzrelation $\sim$ diese Operation respektiert, d. h. man fragt, ob durch die Vorschrift

$$f'(a/\sim) = f(a)/\sim \quad \text{für alle } a \in A$$

eine Funktion $f' : A/\sim \to A/\sim$ wohldefiniert ist. Dies ist offenbar genau dann der Fall, wenn für alle $a, b \in A$ gilt:

Ist $a \sim b$, so ist $f(a) \sim f(b)$. *(Kongruenzbedingung)*

In diesem Fall heißt $\sim$ eine *Kongruenzrelation* für die Operation f. Analoges gilt für mehrstellige Operationen. Für eine zweistellige Operation $f : A^2 \to A$ lautet die Kongruenzbedingung z. B., daß für alle $a, b, c, d \in A$ gilt:

Ist $a \sim b$ und $c \sim d$, so ist $f(a, c) \sim f(b, d)$. *(zweistellige Kongruenzbedingung)*

Dann ist $f' : A/\sim^2 \to A/\sim$ mit $f'(a/\sim, b/\sim) = f(a, b)/\sim$ wohldefiniert.

Schließlich sind in analoger Weise Kongruenzrelationen definiert, die mehrere ein- oder mehrstellige Operationen auf A respektieren. Alle respektierten Operationen lassen sich dann von A auf die Faktorisierung $A/\sim$ übertragen.

Damit haben wir ein starkes Begriffsvokabular entwickelt: Wir können Abstraktionen durchführen (eine Äquivalenzrelation $\sim$ auf einer Menge A einführen) und weiter dann diejenigen Operationen betrachten, die diese Abstraktionen respektieren. Beispiele diskutieren wir in den Übungen.

Isomorphismen

Mit Hilfe von Bijektionen können wir ausdrücken, daß zwei mathematische Objekte strukturell gleich sind. Wir betrachten hierzu zwei Mengen A und B, auf denen jeweils eine zweistellige Relation R bzw. S definiert ist. Dann heißen (A, R) und (B, S) *isomorph*, falls es eine Bijektion $F : A \to B$ gibt, sodaß für alle $a, b \in A$ gilt:

$a \mathrel{R} b$ *gdw* $F(a) \mathrel{S} F(b)$. *(Isomorphiebedingung für Relationen)*

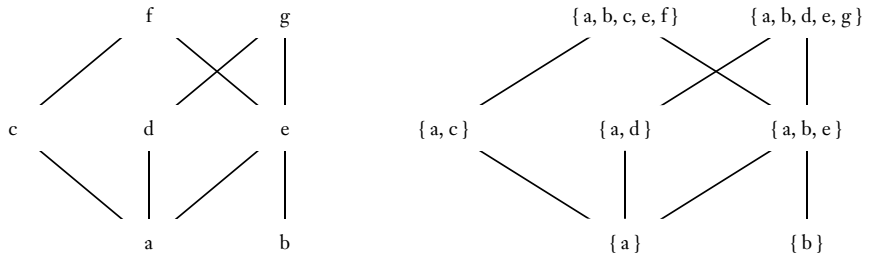
In diesem Fall heißt F ein *Isomorphismus* zwischen (A, R) und (B, S).

Eine anschauliche Vorstellung ist: Wir geben jedem $a \in A$ den neuen Namen F(a). Die Isomorphiebedingung besagt dann gerade, daß die Struktur (A, R) durch diese Umbenennung in die Struktur (B, S) übergeht.

Zur Illustration betrachten wir noch einmal die oben angegebene partielle Ordnung auf der Menge $A = \{ a, \dots, g \}$. Diese Ordnung ist isomorph zu der im Diagramm rechts unten gezeigten Inklusionsordnung, die auf einem Teilsystem B der Potenzmenge von A definiert ist. Ein Isomorphismus $F : A \to B$ ist definiert durch

$F(a) = \{ a \}$, $F(b) = \{ b \}$, $F(c) = \{ a, c \}$, $F(d) = \{ a, d \}$,

$F(e) = \{ a, b, e \}$, $F(f) = \{ a, b, c, e, f \}$, $F(g) = \{ a, b, d, e, g \}$.



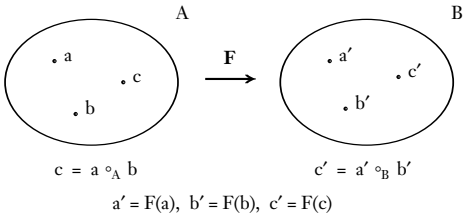
Dieser Übergang von einem Element y zur Menge $\{ x \mid x \leq y \}$ läßt sich für partielle Ordnungen immer durchführen. Er zeigt dann den ansprechenden Satz, daß jede partielle Ordnung $\leq$ auf einer Menge A isomorph zur Inklusionsordnung $\subseteq$ auf einem Mengensystem B auf A ist. Die Inklusionsordnung ist in diesem Sinne universell, es gibt keine Beispiele für partielle Ordnungen, die strukturell nicht von der Form $(B, \subseteq)$ wären.

Isomorphismen für Operationen

Da die Mathematik Strukturen untersucht und nicht die Namen von Objekten, ist es nicht verwunderlich, daß der Isomorphiebegriff überall in der Mathematik auftaucht. Im allgemeinen werden dabei zwei Mengen A und B nicht nur durch eine oder mehrere Relationen strukturiert, sondern auch durch Operationen. Ist auf A und B jeweils eine zweistellige Operation $\circ_A$ bzw. $\circ_B$ erklärt, so heißen $(A, \circ_A)$ und $(B, \circ_B)$ *isomorph*, wenn es eine Bijektion $F : A \to B$ gibt, sodaß für alle $a, b \in A$ gilt:

$F(a \circ_A b) = F(a) \circ_B F(b)$. *(Isomorphiebedingung für Funktionen)*

Die Funktion F heißt dann wieder ein *Isomorphismus* zwischen A und B.



Die Isomorphiebedingung $F(a \circ_A b) = F(a) \circ_B F(b)$ besagt, daß folgende Vorgehensweisen das gleiche Ergebnis c′ liefern: (1) Wir verknüpfen a und b in A und schicken das Ergebnis c mit F nach B. (2) Wir schicken a und b mit F nach B und verknüpfen dort a′ und b′.

Schließlich gibt es auch noch eine einfache Isomorphiebedingung für Konstanten, d. h. besonders ausgezeichnete Elemente der Strukturen (wie zum Beispiel der 0 und der 1 in Zahlstrukturen). Sind auf A und B besondere Elemente c_A und c_B ausgezeichnet, so fordern wir

$F(c_A) = c_B$. *(Isomorphiebedingung für Konstanten)*

Ist auf A und B jeweils ein Satz von Relationen, Operationen und Konstanten erklärt, so werden die Isomorphiebedingungen als Paket gefordert. So sind zum Beispiel zwei Strukturen $(A, R, \circ_A, c_A)$ und $(B, S, \circ_B, c_B)$ isomorph, wenn es eine Bijektion $F : A \to B$ gibt, die alle drei genannten Isomorphiebedingungen erfüllt. Analoges gilt für mehrere Relationen, Operationen und Konstanten. Dabei dürfen die Relationen und Funktionen auch eine beliebige Stellenzahl haben.

Der Isomorphiebegriff für Strukturen spielt eine große Rolle für Charakterisierungen und axiomatische Beschreibungen. Die Frage „Was sind die natürlichen Zahlen?“ beantwortet ein Mathematiker gerne wie folgt: „Die natürlichen Zahlen sind eine Menge N mit den und jenen Eigenschaften.“ Die Eigenschaften formuliert er mit Hilfe von Funktionen, Relationen und Konstanten. Er zeigt dann auf Wunsch die beiden folgenden Sätze:

(1) Es gibt eine Menge N mit den angegebenen Struktureigenschaften. *(Existenzsatz)*

(2) Je zwei Mengen N und M mit diesen Struktureigenschaften sind isomorph. *(Eindeutigkeitssatz)*

Mehr kann und will die Mathematik in der Regel nicht über das „Wesen“ der natürlichen Zahlen aussagen, und das gleiche gilt für die rationalen, reellen und komplexen Zahlen. Doch kann es sein, daß sich bei der Analyse eines Konstrukti-

onsproblems zeigt, daß es eine oder mehrere besonders natürliche, schöne, elegante, ausgezeichnete, einfache Methoden gibt, das gewünschte Objekt zu gewinnen. Der Mathematiker spricht dann in seiner Begeisterung etwas vage von einer „kanonischen" Konstruktion. Wir werden auf diese Dinge bei der Diskussion der natürlichen Zahlen im zweiten Abschnitt noch zurückkommen.

Übungen

Übung 1 *(Relationen und ihre Struktureigenschaften, I) (L)*
Für alle natürlichen Zahlen n, m setzen wir:

n R m, *falls* $|n - m|$ ist gerade,

n S m, *falls* $|n - m|$ ist ungerade.

(Hierbei ist $|a|$ der Betrag einer ganzen Zahl a, also $|a| = a$, falls $a \geq 0$, und $|a| = -a$, falls $a < 0$.)
Welche Struktureigenschaften haben diese Relationen?

Übung 2 *(Relationen und ihre Struktureigenschaften, II)*
Sei R eine Relation auf A. Welche Struktureigenschaften vererben sich von R auf R^{-1}? Welche Struktureigenschaften besitzt $R \cup R^{-1}$?

Übung 3 *(Relationen und ihre Struktureigenschaften, III) (L)*
Sei R eine Relation auf A. Zeigen Sie, daß es eine $\subseteq$-kleinste Relation R^* auf A gibt mit $R^* \supseteq R$ und R^* transitiv.

Übung 4 *(Relationen und ihre Struktureigenschaften, IV)*
Seien R und S Relationen auf A. Wir definieren eine Verknüpfung $\circ$ durch

$$R \circ S = \{(a, c) \mid \text{es gibt ein } b \text{ mit } a\ R\ b \text{ und } b\ S\ c\}.$$

(i) Untersuchen Sie exemplarisch die Struktureigenschaften von $R \circ S$ in Abhängigkeit von den Struktureigenschaften von R und S.

(ii) Wie läßt sich die Transitivität von R mit Hilfe der Verknüpfung $\circ$ ausdrücken?

(iii) Zeigen Sie, daß $\circ$ assoziativ ist, d. h. für alle Relationen R, S, T auf A gilt $(R \circ S) \circ T = R \circ (S \circ T)$.

Übung 5 *(Äquivalenzrelationen, I)*
Sei A eine Menge. Wir definieren für alle $a, b \in A$:

$a \sim_1 b$, *falls* $a = b$,

$a \sim_2 b$, *falls* $a = a$.

Zeigen Sie, daß $\sim_1$ und $\sim_2$ Äquivalenzrelationen sind und bestimmen Sie die Äquivalenzklassen.

Übung 6 *(Äquivalenzrelationen, II)*
Sei $\sim$ eine Äquivalenzrelation auf A. Zeigen Sie, daß $\mathfrak{Z} = A/\sim$ eine Zerlegung von A ist.

Übung 7 *(Äquivalenzrelationen, III)*
Sei $\mathfrak{Z}$ eine Zerlegung von A. Für $a, b \in A$ setzen wir:

$a \sim_{\mathfrak{Z}} b$, *falls* es gibt ein $Z \in \mathfrak{Z}$ mit $a, b \in Z$.

Zeigen Sie, daß $\sim_{\mathfrak{Z}}$ eine Äquivalenzrelation mit $A/\sim_{\mathfrak{Z}} = \mathfrak{Z}$ ist.

Übung 8 *(Äquivalenzrelationen, IV) (L)*
Seien $\sim_1$ und $\sim_2$ Äquivalenzrelationen auf A. Wir setzen für alle $a, b \in A$:

$a \sim b$, *falls* $a \sim_1 b$ und $a \sim_2 b$.

Zeigen Sie, daß $\sim$ eine Äquivalenzrelation auf A ist und visualisieren Sie die Äquivalenzklassen mit Hilfe von Zerlegungen.

Übung 9 *(Äquivalenzrelationen, V)*
Seien $\sim_1$ und $\sim_2$ Äquivalenzrelationen auf A. Wir setzen für alle $a, b \in A$:

$a \, R \, b$, *falls* $a \sim_1 b$ oder $a \sim_2 b$.

Welche Struktureigenschaften hat diese Relation R?

Übung 10 *(Äquivalenzrelationen, VI)*
Sei $f : A \to B$ eine Funktion, und sei $\sim$ eine Äquivalenzrelation auf B. Für alle $a, b \in A$ setzen wir $a \equiv b$, falls $f(a) \sim f(b)$. Zeigen Sie, daß $\equiv$ eine Äquivalenzrelation ist und bestimmen Sie die Äquivalenzklassen.

Übung 11 *(Äquivalenzrelationen, VII)*
Sei R reflexiv und transitiv auf A. Wir definieren für alle $a, b \in A$:

$a \sim b$, *falls* $a \, R \, b$ und $b \, R \, a$.

Zeigen Sie, daß $\sim$ eine Äquivalenzrelation auf A ist.

Übung 12 *(Ordnungen, I) (L)*
Sei $\mathfrak{A}$ ein Mengensystem auf A. Zeigen Sie, daß die Inklusion eine partielle Ordnung auf $\mathfrak{A}$ ist.
Visualisieren Sie sich diese Ordnung für $A = \{ 1, 2, 3 \}$ und $\mathfrak{A} = \mathcal{P}(A)$.

Übung 13 *(Ordnungen, II)*
Sei $\leq$ eine partielle Ordnung auf A. Wir setzen für alle $a, b \in A$:

$a \leq^* b$, *falls* $b \leq a$.

Zeigen Sie, daß $\leq^*$ eine partielle Ordnung ist, und daß $\leq^*$ linear ist, falls dies für $\leq$ gilt. Sind $\leq$ und $\leq^*$ für lineare Ordnungen immer isomorph?

Übung 14 *(Ordnungen, III)*
Seien $(A, \leq)$ und $(B, \leq)$ linear geordnet, und es gelte $A \cap B = \emptyset$.
Sei $C = A \cup B$. Wir definieren für alle $a, b \in C$:

$a \leq b$, *falls* $(a, b \in A \wedge a \leq b) \vee (a, b \in B \text{ und } a \leq b) \vee (a \in A \wedge b \in B)$.

Zeigen Sie, daß $\leq$ eine lineare Ordnung auf C ist und visualisieren Sie diese Ordnung.

Übung 15 *(Ordnungen, IV)*
Seien $(A, \leq)$ und $(B, \leq)$ linear geordnet. Sei $C = A \times B$. Wir definieren für alle $(a_1, b_1), (a_2, b_2) \in C$:

$(a_1, b_1) \leq (a_2, b_2)$, *falls* $b_1 \leq b_2 \vee (b_1 = b_2 \wedge a_1 \leq a_2)$.

Zeigen Sie, daß $\leq$ eine lineare Ordnung auf C ist und visualisieren Sie diese Ordnung.

Übung 16 *(Funktionen, I) (L)*
Seien $f : A \to B$, $g : B \to C$, $h : C \to D$ Funktionen. Zeigen Sie:

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

Übung 17 *(Funktionen, II)*
Zeigen Sie:

(i) Ist $f : A \to B$ injektiv, so ist $f : A \to \text{rng}(f)$ bijektiv.

(ii) Ist $f : A \to B$ bijektiv, so gibt es ein $g : B \to A$ bijektiv mit $g \circ f = \text{id}_A$.

(iii) Ist $f : A \to B$ surjektiv, so existiert ein injektives $g : B \to A$.

(iv) Sind $f : A \to B$, $g : B \to C$ bijektiv, so ist $g \circ f : A \to C$ bijektiv.

Übung 18 *(Funktionen, III)*
Seien $f : A \to B$ und $g : B \to C$ bijektiv. Zeigen Sie:

$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Übung 19 *(Funktionen, IV) (L)*
Sei $f : A \to B$ mit einer nichtleeren Menge A. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(i) $f : A \to B$ ist injektiv.

(ii) Es gibt ein $g : B \to A$ mit $g \circ f = \text{id}_A$.

Formulieren und beweisen Sie eine ähnliche Äquivalenz für Surjektionen.

Übung 20 *(Funktionen, V)*
Sei $f : A \to B$. Zeigen Sie:

(i) $f[f^{-1}[X]] \subseteq X$ für alle $X \subseteq B$,

(ii) $f^{-1}[f[X]] \supseteq X$ für alle $X \subseteq A$.

(iii) $f[X - Y] \supseteq f[X] - f[Y]$ für alle $X, Y \subseteq A$,

(iv) $f^{-1}[X - Y] = f^{-1}[X] - f^{-1}[Y]$ für alle $X, Y \subseteq B$,

(v) $f[\bigcap \mathscr{X}] \subseteq \bigcap \{f[X] \mid X \in \mathscr{X}\}$ für alle $\mathscr{X} \subseteq \mathcal{P}(A)$, $\mathscr{X} \neq \emptyset$,

(vi) $f[\bigcup \mathscr{X}] = \bigcup \{f[X] \mid X \in \mathscr{X}\}$ für alle $\mathscr{X} \subseteq \mathcal{P}(A)$,

(vii) $f^{-1}[\bigcap \mathscr{X}] = \bigcap \{f^{-1}[X] \mid X \in \mathscr{X}\}$ für alle $\mathscr{X} \subseteq \mathcal{P}(B)$, $\mathscr{X} \neq \emptyset$,

(viii) $f^{-1}[\bigcup \mathscr{X}] = \bigcup \{f^{-1}[X] \mid X \in \mathscr{X}\}$ für alle $\mathscr{X} \subseteq \mathcal{P}(B)$.

(Zur Vereinfachung können Sie die Aussagen (v) – (viii) zunächst für $X \cap Y$ und $X \cup Y$ statt allgemeiner $\bigcap \mathscr{X}$ und $\bigcup \mathscr{X}$ formulieren und beweisen.)

Zeigen Sie, daß die Gleichheit in (i), (ii), (iii) und (v) im allgemeinen nicht gilt. Unter welcher Voraussetzung an X gilt Gleichheit in (i)? Unter welcher Voraussetzung an f gilt Gleichheit in (ii), (iii) und (v)?

Übung 21 *(Wohldefiniertheit und Kongruenzrelationen, I) (L)*
Sei R eine reflexive und transitive Relation auf A, und sei $\sim$ die wie oben definierte Äquivalenzrelation, d. h. für alle $a, b \in A$ gilt $a \sim b$ genau dann, wenn a R b und b R a. Wir definieren nun für alle $a/\sim, b/\sim \in A/\sim$:

$a/\sim \leq b/\sim$, *falls* a R b.

Zeigen Sie, daß $\leq$ eine wohldefinierte partielle Ordnung auf $A/\sim$ ist.

Übung 22 *(Wohldefiniertheit und Kongruenzrelationen, II)*
Sei $m \geq 1$ eine natürliche Zahl. Für ganze Zahlen a, b definieren wir:

$a \equiv b$, *falls* $|a - b|$ ist ohne Rest durch m teilbar.

Zeigen Sie, daß $\equiv$ eine Kongruenzrelation für die Addition und die Multiplikation auf den ganzen Zahlen ist.

Übung 23 *(Isomorphismen, I) (L)*
Sei $\leq$ eine partielle Ordnung auf A. Wir definieren $F : A \to \mathcal{P}(A)$ durch

$F(a) = \{ b \in A \mid b \leq a \}$ für alle $a \in A$.

Sei $\mathscr{A} = \mathrm{rng}(F)$. Zeigen Sie, daß F ein Isomorphismus zwischen $(A, \leq)$ und $(\mathscr{A}, \subseteq)$ ist.

Übung 24 *(Isomorphismen, II)*
Sei $(A, R, \circ_A, c_A)$ isomorph zu $(B, S, \circ_B, c_B)$, und sei $(B, S, \circ_B, c_B)$ isomorph zu $(C, T, \circ_C, c_C)$. Zeigen Sie, daß $(A, R, \circ_A, c_A)$ isomorph zu $(C, T, \circ_C, c_C)$ ist.

Exkurs: Mächtigkeiten

Die Entwicklung der mathematischen Sprache führte uns zu den Strukturbegriffen „injektiv, surjektiv, bijektiv“ für Funktionen. Auf der Grundlage dieser Begriffe läßt sich nun, zunächst ganz ohne die Verwendung von Zahlen, eine mathematische Theorie der Mächtigkeiten von Mengen entwickeln, und wir wollen hier die Grundzüge dieser Theorie vorstellen. Das dabei auftretende Phänomen der Größenunterschiede im Unendlichen hat eine große mathematische und philosophische Strahlkraft, und es ist darüber hinaus für jede Form der kontinuierlichen Mathematik, die auf der Menge der reellen Zahlen basiert, von großer Bedeutung.

Mächtigkeitsvergleiche

Eine Herde von weißen und schwarzen Schafen können wir vergleichen, indem wir die Schafe paarweise durch ein Tor schicken und ermitteln, ob am Ende Schafe einer Farbe übrigbleiben. Mathematisch betrachtet konstruieren wir hier injektive Funktionen. In der Tat legen die Begriffe „injektiv“ und „bijektiv“ die folgenden Größenvergleiche für beliebige Mengen nahe:

Definition *(Mächtigkeitsvergleiche)*
Seien M und N Mengen. Wir schreiben:

$|M| \leq |N|$, *falls* „es gibt ein injektives $f : M \to N$“.

$|M| = |N|$, *falls* „es gibt ein bijektives $f : M \to N$“.

$|M| < |N|$, *falls* $|M| \leq |N|$, aber non$(|M| = |N|)$.

Gilt $|M| \leq |N|$, so sagen wir, daß die *Mächtigkeit* von M kleinergleich der Mächtigkeit von N ist. Analoge Sprechweisen verwenden wir für $|M| = |N|$ *(gleiche Mächtigkeit)* und $|M| < |N|$ *(kleinere Mächtigkeit)*.

Einige Eigenschaften der Mächtigkeitsvergleiche sind leicht zu zeigen. So ist die Gleichmächtigkeit zum Beispiel reflexiv, symmetrisch und transitiv auf der Klasse aller Mengen. Für jede Menge A definiert damit

$X \equiv Y$, *falls* $|X| = |Y|$ für alle $X, Y \subseteq A$

eine Äquivalenzrelation auf $\mathcal{P}(A)$.

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_4, © Springer-Verlag Berlin Heidelberg 2010

Der Satz von Cantor-Bernstein

Der Mächtigkeitsvergleich ist offenbar reflexiv und transitiv. Die Schreibweise suggeriert zudem die Antisymmetrie, d. h. die Gültigkeit der folgenden Implikation:

$|M| \leq |N|$ und $|N| \leq |M|$ *impliziert* $|M| = |N|$. *(Satz von Cantor-Bernstein)*

Der Beweis dieser Aussage ist eine nichttriviale Angelegenheit. Wir müssen zwei gegebene Injektionen

$f_1 : M \to N$ und $f_2 : N \to M$

zu einer Bijektion $f : M \to N$ verschmelzen. Daß dies letztendlich doch in einer relativ einfachen und dazu auch konstruktiven Art und Weise möglich ist, gehört zu den Juwelen der Mathematik. Als Korollar erhalten wir, daß der Mächtigkeitsvergleich $|M| \leq |N|$ die Strukturmerkmale einer partiellen Ordnung besitzt.

Cantor hat den Satz lange vermutet, aber keinen Beweis gesehen. Beweise wurden dann von Dedekind, Bernstein, Zermelo und anderen gefunden. Der folgende Beweis stammt von Dedekind. Es genügt für das folgende, wenn der Leser die Problemstellung des Satzes von Cantor-Bernstein zur Kenntnis nimmt. Der Beweis ist für diejenigen gedacht, die an dieser Stelle ein anspruchsvolleres Ergebnis studieren möchten.

Wir beweisen zunächst ein verwandtes Resultat, aus welchem wir dann den eigentlichen Satz leicht ableiten können.

Satz *(Satz von Cantor-Bernstein, Inklusionsform)*

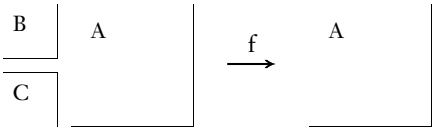
Seien A, B, C paarweise disjunkte Mengen, und sei

$f : A \cup B \cup C \to A$ bijektiv.

Dann gibt es Bijektionen

$b : A \cup B \cup C \to A \cup B$,

$b' : A \cup B \to A$.



Anders formuliert: Alle Mengen, die bzgl. der Inklusion zwischen zwei gleichmächtigen Mengen liegen, sind gleichmächtig mit diesen Mengen.

Beweis

Wir setzen:

$$Z = \bigcap \{ D \subseteq A \cup C \mid C \subseteq D,\ f[D] \subseteq D \}.$$

Die Menge des Schnitts ist nichtleer, da die Menge $D = A \cup C$ die geforderten Eigenschaften hat. Es gilt zudem $C \subseteq Z$, denn C ist eine Teilmenge jeder Menge D der Schnittbildung.

Weiter gilt:

(+) $f[Z] = Z - C$.

Beweis von (+)

Es gilt $f[Z] \subseteq Z$, denn ist $x \in Z$, so ist $f(x) \in D$ für alle D wie in der Schnittbildung, also $f(x) \in Z$. Wegen $rng(f) \subseteq A$ ist also $f[Z] \subseteq Z - C$.

Sei nun $x \in Z - C$ beliebig. Dann ist $Z - \{x\}$ eine echte Teilmenge von Z und eine Obermenge von C. Nach Definition von Z ist also das Bild $f[Z - \{x\}]$ keine Teilmenge von $Z - \{x\}$. Wegen

$$f[Z - \{x\}] \subseteq f[Z] \subseteq Z$$

gibt es also ein $y \in Z - \{x\}$ mit $f(y) = x$. Also ist $x \in f[Z]$.

Nach (+) ist $f|Z : Z \to Z - C$ bijektiv, und damit ist

$$b = f|Z \cup id_{(A \cup B) - Z}$$

eine Bijektion von $A \cup B \cup C$ nach $A \cup B$.

Schließlich ist dann $b' = f \circ b^{-1}$ eine Bijektion von $A \cup B$ nach A.

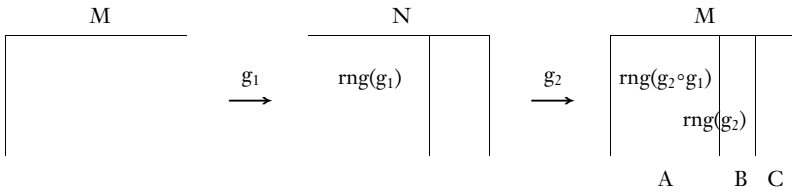
Wir erhalten hieraus ohne weitere Schwierigkeiten:

Korollar *(Satz von Cantor-Bernstein)*
Seien $g_1 : M \to N$ und $g_2 : N \to M$ injektiv.
Dann existiert eine Bijektion $g : M \to N$.

Beweis

Wir definieren:

$$A = rng(g_2 \circ g_1), \quad B = rng(g_2) - A, \quad C = M - rng(g_2).$$



Dann sind A, B, C paarweise disjunkt, und es gilt

$$M = A \cup B \cup C.$$

Weiter ist $g_2 \circ g_1 : A \cup B \cup C \to A$ bijektiv. Nach dem Satz gibt es ein bijektives $b : M \to A \cup B$. Aber es gilt $A \cup B = rng(g_2)$, und damit ist

$$g_2^{-1} \circ b : M \to N$$

bijektiv.

Als nächstes zeigen wir den nach dem Satz von Cantor-Bernstein zweiten Hauptsatz der elementaren Mächtigkeitstheorie: Die Potenzmenge einer Menge M ist immer von größerer Mächtigkeit als die Mächtigkeit von M selbst. Dieser starke Satz hat einen überraschend kurzen und einfachen Beweis, der als *Cantorsches Diagonalverfahren* bekannt ist.

Satz *(Satz von Cantor)*
Sei M eine Menge, und sei $f : M \to \mathcal{P}(M)$ eine Funktion.
Dann ist f nicht surjektiv. Genauer gilt: Sei

$D = \{ x \in M \mid x \notin f(x) \}$.

Dann gilt $D \notin \mathrm{rng}(f)$.

Beweis
Annahme, $D \in \mathrm{rng}(f)$. Sei dann $x^* \in M$ mit $f(x^*) = D$. Für alle $x \in M$ gilt nach Definition von D:

$x \in D$ *gdw* $x \notin f(x)$.

Speziell gilt also für x^*:

$x^* \in D$ *gdw* $x^* \notin f(x^*)$.

Wegen $f(x^*) = D$ haben wir also:

$x^* \in D$ *gdw* $x^* \notin D$,

– *Widerspruch*.

Korollar
Für alle Mengen M gilt $|M| < |\mathcal{P}(M)|$.

Beweis
Nach dem Satz von Cantor gilt $|M| \neq |\mathcal{P}(M)|$. Die injektive Funktion
– $f : M \to \mathcal{P}(M)$ mit $f(x) = \{ x \}$ für alle $x \in M$ zeigt, daß $|M| \leq |\mathcal{P}(M)|$.

Der Leser ist beim Studium des Beweises vielleicht an die Russell-Zermelo-Antinomie erinnert worden. In der Tat war Cantors allgemeiner Satz die Quelle für Russells Konstruktion der Klasse $R = \{ x \mid x \notin x \}$. Genaueres diskutieren wir in den Übungen.

Es gibt auch noch einen dritten Hauptsatz der elementaren Mächtigkeitstheorie: Für alle Mengen M und N gilt $|M| \leq |N|$ oder $|N| \leq |M|$. Je zwei Mengen sind also in ihrer Mächtigkeit vergleichbar. Diese Aussage ist richtig, läßt sich aber nur mit weitergehenden Hilfsmitteln beweisen, und wir können auf diesen Vergleichbarkeitssatz für Mächtigkeiten hier nicht weiter eingehen.

Unendlichkeiten

Der Begriff der Injektion dominiert die elementare Mengenlehre. Er ermöglicht uns den Vergleich der „Größe“ von beliebigen Mengen. Weiter läßt er sich auch zur Definition der Unendlichkeit verwenden:

Definition *(unendlich, endlich)*
Eine Menge M heißt *(Dedekind-) unendlich*, falls gilt:

Es gibt ein injektives $f : M \to M$ mit $rng(f) \neq M$.

Weiter heißt eine Menge N *(Dedekind-) endlich*, falls N nicht unendlich ist.

Gleichwertig zur Dedekind-Unendlichkeit ist: Es existiert eine echte Teilmenge N von M und eine bijektive Funktion $f : M \to N$.

Die endlichen Mengen sind also dadurch gekennzeichnet, daß jede injektive Operation auf ihnen automatisch bijektiv ist.

Ist M unendlich, so gilt

$$|M| < |\mathcal{P}(M)| < |\mathcal{P}(\mathcal{P}(M))| < \ldots$$

nach dem Satz von Cantor. Wir lassen uns auf ungeahnte Weiten ein, wenn wir unendliche Mengen und Potenzmengen unendlicher Mengen zulassen: Es gibt Größenunterschiede im Unendlichen!

Wir versuchen nun, die Struktur des Unendlichen etwas genauer zu beschreiben, indem wir Stufen der Unendlichkeit einführen. Hierzu definieren wir:

Definition *(abzählbar, überabzählbar)*
Eine Menge M heißt *abzählbar unendlich* oder von *von der ersten unendlichen Mächtigkeit*, falls gilt:

(a) M ist unendlich.

(b) Jede Teilmenge von M ist endlich oder gleichmächtig zu M.

Wir schreiben dann symbolisch auch $|M| = \aleph_0$ [Aleph-0].
Weiter heißt eine Menge A *abzählbar*, falls A endlich oder abzählbar unendlich ist, und *überabzählbar*, falls A nicht abzählbar ist.

Eine analoge Definition isoliert eine weitere Stufe des Unendlichen:

Definition *(von der zweiten unendlichen Mächtigkeit)*
Eine Menge M heißt *von der zweiten unendlichen Mächtigkeit*, falls gilt:

(i) M ist überabzählbar.

(ii) Jede Teilmenge von M ist abzählbar oder gleichmächtig zu M.

Wir schreiben dann symbolisch auch $|M| = \aleph_1$ [Aleph-1].

Ist M unendlich, so ist $\mathcal{P}(M)$ überabzählbar nach dem Satz von Cantor. Es stellt sich dann die Frage, wie groß der Sprung von M zu $\mathcal{P}(M)$ ist. Wir formulieren hierzu:

Kontinuumshypothese
Sei M abzählbar unendlich. Dann ist $\mathcal{P}(M)$ von der zweiten unendlichen Mächtigkeit.

Diese Aussage ist als Hypothese formuliert. Es ist nicht gelungen, sie zu beweisen oder zu widerlegen. Dagegen ist es aber gelungen zu zeigen, daß sie im Rahmen der üblichen, als widerspruchsfrei vorausgesetzten mengentheoretischen Axiomatik tatsächlich weder beweisbar noch widerlegbar ist. Die Kontinuumshypothese ist, wie man sagt, unabhängig von dieser Axiomatik.

Hat man die natürlichen Zahlen $\mathbb{N}$ und die reellen Zahlen $\mathbb{R}$ zur Verfügung, so lassen sich die Begriffe „endlich", „abzählbar" und „Kontinuumshypothese" weiter erläutern und motivieren. Denn eine Menge M ist, wie man zeigen kann, genau dann Dedekind-endlich, wenn ein $n \in \mathbb{N}$ existiert mit $|M| = |\{0, \ldots, n-1\}|$. Weiter läßt sich zeigen, daß eine Menge M genau dann abzählbar unendlich ist, wenn $|M| = |\mathbb{N}|$ gilt. Die Abzählbarkeit einer unendlichen Menge B bedeutet also, daß wir die Elemente von B vollständig in der Form $b_0, b_1, \ldots, b_n, \ldots$, auflisten können, mit Indizes n aus den natürlichen Zahlen. Für die reellen Zahlen $\mathbb{R}$ gilt weiter die fundamentale Mächtigkeitsbeziehung

(#) $\quad |\mathbb{R}| = |\mathcal{P}(\mathbb{N})|,$

die wir in den Übungen ausgehend von einem Grundverständnis der reellen Zahlen diskutieren. Damit folgt aus dem Satz von Cantor, daß die reellen Zahlen überabzählbar sind. Anders formuliert: Ist $x_0, x_1, \ldots, x_n, \ldots$ eine Folge reeller Zahlen, so gibt es eine reelle Zahl x mit $x \neq x_n$ für alle n. Die Kontinuumshypothese besagt dann, daß die reellen Zahlen von der zweiten unendlichen Mächtigkeit sind, oder, anders formuliert, daß gilt:

„Ist P eine unendliche Menge reeller Zahlen, so existiert entweder eine Bijektion zwischen P und $\mathbb{N}$ oder eine Bijektion zwischen P und $\mathbb{R}$."

Diese Aussage ist, im verschärften Sinne der nachgewiesenen Unabhängigkeit, offen.

Wir haben die Sprache der Mathematik nun recht genau kennengelernt, und dabei im Umfeld von Mengen und Bijektionen sogar eine unerwartet reichhaltige und subtile Theorie entdeckt. Als nächstes wollen wir nun aber die Zahlen nicht mehr nur naiv in Übungsaufgaben und Beispielen verwenden, sondern im Rahmen unserer Sprache einführen und untersuchen.

Übungen

Übung 1 *(Mächtigkeitsvergleiche, I)*
Sei A eine Menge. Für $X, Y \subseteq A$ setzen wir:

$X \equiv Y$, *falls* „es gibt ein bijektives $f : X \to Y$“.

Zeigen Sie, daß $\equiv$ eine Äquivalenzrelation auf $\mathcal{P}(A)$ ist.

Übung 2 *(Mächtigkeitsvergleiche, II)*
Für eine Menge M sei ${}^{M}\{0, 1\} = \{ f \mid f : M \to \{0, 1\} \}$.
Zeigen Sie, daß $|\mathcal{P}(M)| = |{}^{M}\{0, 1\}|$.

Übung 3 *(Mächtigkeitsvergleiche, III)*
Seien A, B Mengen mit $|A| = |B|$. Zeigen Sie, daß $|\mathcal{P}(A)| = |\mathcal{P}(B)|$.

Übung 4 *(Der Satz von Cantor-Bernstein, I)*
Sei A eine Menge, und sei $\equiv$ die Äquivalenzrelation der Gleichmächtigkeit auf $\mathcal{P}(A)$. Für $X, Y \subseteq A$ setzen wir:

$X/\equiv \ \leq \ Y/\equiv$, *falls* „es gibt ein injektives $f : X \to Y$“.

Zeigen Sie, daß $\leq$ eine wohldefinierte partielle Ordnung auf $\mathcal{P}(A)/\equiv$ ist.

Übung 5 *(Der Satz von Cantor-Bernstein, II)*
Seien $f : M \to N$ und $g : N \to M$ injektiv. Zeigen Sie, daß es Mengen $S \subseteq M$ und $T \subseteq \mathrm{rng}(g)$ mit $S \cap T = \emptyset$ und $S \cup T = M$ gibt, sodaß die Funktion

$h = f|S \cup g^{-1}|T$

eine Bijektion von M nach N ist.

Übung 6 *(Der Satz von Cantor-Bernstein, III)*
Leiten Sie die Inklusionsform des Satzes von Cantor-Bernstein aus dem Satz von Cantor-Bernstein ab.

Übung 7 *(Der Satz von Cantor-Bernstein, IV)*
Sei $\mathbb{N}$ die Menge der natürlichen und $\mathbb{R}$ die Menge der reellen Zahlen. Zeigen Sie mit Hilfe des Satzes von Cantor-Bernstein:

(a) $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$,

(b) $|\mathbb{N}| = |\mathbb{N}^2|$,

(c) $|\mathbb{R}| = |\mathbb{R}^2|$.

Übung 8 *(Der Satz von Cantor-Bernstein, V)*
Zeigen Sie:

(i) $|\mathbb{N}| = |(\mathbb{N} \times \{0\}) \cup (\mathbb{N} \times \{1\})|$.

(ii) Sind A, B disjunkt, so ist $|\mathcal{P}(A \cup B)| = |\mathcal{P}(A) \times \mathcal{P}(B)|$.

(iii) Geben Sie mit Hilfe von (i) und (ii) einen neuen Beweis für die Existenz einer Bijektion zwischen $\mathbb{R}$ und $\mathbb{R}^2$.

[Argumentieren Sie für (iii) etwa wie folgt:
$|\mathbb{R}| = |\mathcal{P}(\mathbb{N})| = |\mathcal{P}(\mathbb{N} \times \{0\} \cup \mathbb{N} \times \{1\})| = |\mathcal{P}(\mathbb{N} \times \{0\}) \times \mathcal{P}(\mathbb{N} \times \{1\})| = |\mathbb{R} \times \mathbb{R}|$.]

Übung 9 *(Der Satz von Cantor-Bernstein, VI)*
Wir definieren Mengen A, B, C durch

$A = \{0, 2, 4, \ldots\} = \{n \in \mathbb{N} \mid n \text{ ist gerade}\}, \quad C = \{1\},$

$B = \{3, 4, 5, \ldots\} = \{n \in \mathbb{N} \mid n \text{ ist ungerade, } n \geq 3\}.$

Sei $f : \mathbb{N} \to A$ die Bijektion mit $f(n) = 2n$ für alle $n \in \mathbb{N}$.
Bestimmen Sie die Bijektionen $b : \mathbb{N} \to A \cup B$ und $b' : A \cup B \to A$, die im Beweis der Inklusionsform des Satzes von Cantor-Bernstein für den vorliegenden Fall konstruiert werden. Bestimmen Sie hierzu die Menge Z und überzeugen Sie sich noch einmal von der Aussage (+).

Übung 10 *(Unendlichkeiten, I)*
Sei $f : \{0, 1, 2, 3, 4, 5\} \to \mathcal{P}(\{0, 1, 2, 3, 4, 5\})$ die folgende Funktion:

$f(0) = \{0, 1, 2\}$, $f(3) = \{3\}$,

$f(1) = \{2, 4\}$, $f(4) = \emptyset$,

$f(2) = \{0, 1, 3, 4, 5\}$, $f(5) = \{0, 2, 3, 4, 5\}$.

Bestimmen Sie die Diagonalmenge D dieser Funktion wie im Beweis des Satzes von Cantor. Visualisieren Sie die Funktion f durch ein Quadrat mit 36 Feldern und D als eine Diagonale.

Übung 11 *(Unendlichkeiten, II)*
Wie kann man die Antinomie von Russell-Zermelo aus dem Satz von Cantor gewinnen?

[Betrachten Sie als Funktion die Identität auf $V = \{x \mid x \text{ ist Menge}\}$.]

Übung 12 *(Unendlichkeiten, III)*
Zeigen Sie direkt, daß $|\mathbb{N}| < |\mathbb{R}|$ gilt, indem Sie die Diagonalmethode des Beweises des Satzes von Cantor auf reelle Zahlen in Dezimaldarstellung anwenden.

2. Abschnitt

Zahlen

1. Natürliche Zahlen

Wir untersuchen in diesem Kapitel die Struktur des Zählens. Im Vordergrund steht hier die Nachfolgerbildung und das zugehörige Induktionsprinzip. Durch diese Strukturmerkmale lassen sich die natürlichen Zahlen charakterisieren. Mit Hilfe rekursiver Definitionen läßt sich weiter die gesamte Arithmetik und die Ordnung der natürlichen Zahlen aus der Nachfolgerbildung gewinnen.

Wir beginnen mit einer informalen Diskussion, die zu einer präzisen Definition einer Zählstruktur führt.

Nachfolger und Induktion

Das Zählen ist bestimmt durch die Nachfolgerbildung, die einer Zahl n ihren direkten Nachfolger S(n) zuordnet:

0, 1, 2, 3, …, n, S(n), …, …, …, m, S(m), …

Die Elemente dieser Reihe – wie auch immer wir sie konstruieren und benennen wollen – heißen *natürliche Zahlen*. Die Menge aller natürlichen Zahlen bezeichnen wir mit $\mathbb{N}$.

Die Zählreihe beginnt mit einer ausgezeichneten Zahl, dem Anfangselement der Zählreihe. Wir nennen dieses Anfangselement die *Null* der Zählreihe. Bei der Einführung der Arithmetik wird die Null auch die übliche Rolle einer Null übernehmen. Hier wollen wir nur das Zählen selbst betrachten, der Name für das Anfangselement ist prinzipiell beliebig.

Die Null ist kein Nachfolger einer Zahl, jede von Null verschiedene Zahl ist dagegen der Nachfolger genau einer anderen Zahl. Es werden immer neue Zahlen gebildet, die Nachfolgerfunktion ist injektiv: Haben zwei Zahlen denselben Nachfolger, so sind sie gleich. Weiter wird die Zählreihe durch die bei Null beginnende Nachfolgerbildung vollkommen bestimmt und ausgeschöpft. Diese komplexe und keineswegs unproblematische Intuition versuchen wir durch ein *Induktionsschema* auszudrücken:

> Gilt eine Eigenschaft $\mathscr{E}$ für die Null, und gilt $\mathscr{E}$ mit n stets auch für S(n), so gilt $\mathscr{E}$ für alle n.
>
> *(Induktionsschema für Eigenschaften)*

Formal:

$$\mathscr{E}(0) \wedge \forall n(\mathscr{E}(n) \to \mathscr{E}(S(n))) \to \forall n\, \mathscr{E}(n).$$

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_5, © Springer-Verlag Berlin Heidelberg 2010

Statt von einer Eigenschaft $\mathscr{E}(n)$ können wir auch von der Menge aller natürlichen Zahlen reden, auf die die Eigenschaft $\mathscr{E}$ zutrifft. Wir setzen dann

$$X = \{ n \in \mathbb{N} \mid \mathscr{E}(n) \}.$$

Ist umgekehrt X eine Teilmenge von $\mathbb{N}$, so ist „$n \in X$" eine Eigenschaft für natürliche Zahlen. Das Induktionsschema lautet bei dieser Sicht der Dinge dann wie folgt:

> Ist X eine Teilmenge von $\mathbb{N}$, die die Null enthält und die unter der Nachfolgerfunktion abgeschlossen ist, so ist $X = \mathbb{N}$.
>
> *(mengentheoretisches Induktionsaxiom)*

Formal:

$$\forall X \subseteq \mathbb{N}\, (0 \in X \wedge \forall n\, (n \in X \to S(n) \in X) \;\to\; X = \mathbb{N}).$$

Im Gegensatz zum Induktionsschema für Eigenschaften, wo wir ein Induktions-Axiom pro Eigenschaft und damit unendlich viele Axiome fordern, können wir also in einer mengentheoretischen Umgebung das Induktionsprinzip durch ein einziges Axiom ausdrücken.

Diese intuitive Beschreibung des Zählens übersetzen wir nun in eine mathematische Begriffsbildung.

Dedekind-Strukturen

Definition *(Dedekind-Struktur)*

Sei D eine Menge, d ein Element von D, und sei $S : D \to D$ eine injektive Funktion mit $d \notin \mathrm{rng}(S)$. Dann heißt (D, S, d) eine *Dedekind-Struktur* oder eine *Zählreihe*, falls für alle $X \subseteq D$ das folgende Induktionsaxiom gilt:

(Ind_X) Ist $d \in X$ und ist mit jedem $n \in X$ auch $S(n) \in X$, so ist $X = D$.

Die Funktion S heißt dann die *Nachfolgerfunktion* und d die *Null* der Dedekind-Struktur (D, S, d).

Die wesentlichen Struktureigenschaften sind also:

(a) $d \notin \mathrm{rng}(S)$.

(b) S ist injektiv.

(c) Für alle $X \subseteq D$ gilt (Ind_X).

Sei X eine Teilmenge von D, etwa $X = \{ n \in D \mid \mathscr{E}(n) \}$. Wir wollen zeigen, daß $X = D$ gilt, d. h. jedes $n \in D$ erfüllt die Eigenschaft $\mathscr{E}(n)$. Wir zeigen hierzu:

(1) Es gilt $d \in X$. *(Induktionsanfang)*

(2) Sei $n \in X$. Dann gilt $S(n) \in X$. *(Induktionsschritt)*

Haben wir (1) und (2) bewiesen, so ist die Menge X gleich der Menge D nach dem Induktionsaxiom (Ind_X).

Ein Beweis von X = D nach dem Schema (1) und (2) heißt ein *Beweis durch Induktion*. Im Induktionsschritt (2) darf man „$n \in X$“ verwenden, um „$S(n) \in X$“ zu zeigen. Diese oft sehr hilfreiche zusätzliche Voraussetzung unterscheidet einen induktiven Beweis von einem Beweis, der direkt zeigt, daß jedes $n \in D$ ein Element von X ist. Die Aussage „$n \in X$“ nennt man auch die *Induktionsvoraussetzung*. Wir kürzen sie zuweilen mit *I. V.* ab.

Wir erläutern den Einsatz der Induktion in Dedekind-Strukturen an einem einfachen Beispiel.

Satz *(die Nachfolgerfunktion hat keine Fixpunkte)*
Sei (D, S, d) eine Dedekind-Struktur. Dann gilt $S(n) \neq n$ für alle $n \in D$.

Beweis
Sei $X = \{ n \in D \mid S(n) \neq n \}$. Wir zeigen, daß X die Voraussetzungen des Induktionsaxioms erfüllt. Nach (Ind_X) gilt dann X = D.

Induktionsanfang:

Es gilt $S(d) \neq d$, da sonst d im Wertebereich von S liegen würde.
Also ist $d \in X$.

Induktionsschritt von n nach S(n):

Es gelte also $n \in X$ (die Induktionsvoraussetzung). Dann gilt $S(n) \neq n$.
Da S injektiv ist, haben also S(n) und n verschiedene Nachfolger.
Also gilt $S(S(n)) \neq S(n)$, d. h. $S(n) \in X$. ∎

Mit Induktion zeigt man z. B. auch, daß $\text{rng}(S) = D - \{ d \}$ gilt.

Dem induktiven Beweisen in einer Dedekind-Struktur (D, S, d) steht das rekursive Definieren zur Seite: Wir können eine Funktion f mit Definitionsbereich D definieren, indem wir zunächst f(d) festlegen, und dann weiter für alle $n \in D$ festlegen, wie sich f(S(n)) aus f(n) ergibt, d. h. wir dürfen zur Definition von f(S(n)) den Funktionswert f(n) verwenden. Die Funktionswerte können dabei Elemente von D sein, oder auch beliebige Objekte wie z. B. endliche Folgen in D.

Ein Blick auf das Induktionsaxiom macht es glaubhaft, daß dadurch in eindeutiger Weise eine Funktion definiert wird, deren Definitionsbereich die Menge D ist. Eine strenge Rechtfertigung des rekursiven Definierens ist möglich, wir wollen hier aber darauf verzichten.

Schematisch verläuft eine rekursive Definition also wie folgt:

(1) Wir definieren f(d). *(Rekursionsanfang)*

(2) Wir definieren f(S(n)) unter Verwendung von f(n). *(Rekursionsschritt)*

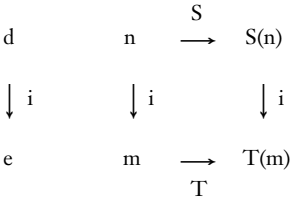
Dadurch wird eine eindeutige Funktion f mit dom(f) = D definiert.

Die Rekursion wird verwendet, um den folgenden fundamentalen Satz zu beweisen, der besagt, daß es bis auf Isomorphie nur eine Dedekind-Struktur gibt. Der Leser kann beim ersten Lesen den Beweis überschlagen.

Satz *(Eindeutigkeitssatz für Dedekind-Strukturen)*
Je zwei Dedekind-Strukturen (D, S, d) und (E, T, e) sind isomorph, d. h. es gibt eine Bijektion $i : D \to E$ mit den Eigenschaften:

(i) $i(d) = e$,

(ii) $i(S(n)) = T(i(n))$ für alle $n \in D$.



Beweis
Wir definieren eine Funktion $i : D \to E$ mit Hilfe von Rekursion über die Dedekind-Struktur (D, S, d):

$i(d) = e$, *(Rekursionsanfang)*

$i(S(n)) = T(i(n))$ für alle $n \in D$. *(Rekursionsschritt)*

Dann gelten die Eigenschaften (i) und (ii) nach Konstruktion. Es ist also nur noch zu zeigen, daß $i : D \to E$ bijektiv ist. Wir beginnen mit:

(+) $i : D \to E$ ist surjektiv.

Beweis von (+)
Sei $Y = rng(i)$. Dann gilt $e \in Y$. Ist $y \in Y$, so gibt es ein $n \in D$ mit $i(n) = y$. Dann ist aber $T(y) = i(S(n)) \in Y$. Nach dem Induktionsaxiom (Ind_Y) für die Dedekind-Struktur (E, T, e) gilt also $Y = E$.

Weiter zeigen wir nun:

(++) i ist injektiv.

Beweis von (++)
Wir zeigen durch Induktion nach $n \in D$:

(#) Für alle $m \in D$ gilt: Ist $m \neq n$, so ist $i(m) \neq i(n)$.

Induktionsanfang n = d
Für alle $m \neq d$ gilt $i(m) \neq i(d) = e$, denn ist $m = S(m')$, so ist $i(m) = i(S(m')) = T(i(m')) \neq e$, da $e \notin rng(T)$.

Induktionsschritt von n nach S(n)
Wir müssen zeigen, daß für alle $m \in D$ gilt:

Ist $m \neq S(n)$, so ist $i(m) \neq i(S(n))$, wobei $i(S(n)) = T(i(n))$.

Dies ist klar für $m = d$, da dann wieder $i(m) = e \notin rng(T)$ gilt.
Sei also $m = S(m')$ und es gelte $m \neq S(n)$. Dann gilt $S(m') \neq S(n)$, und damit $m' \neq n$. Also ist $i(m') \neq i(n)$ nach Induktionsvoraussetzung. Dann ist aber $T(i(m')) \neq T(i(n))$ nach Injektivität von T, also wie gewünscht

$i(m) = i(S(m')) = T(i(m')) \neq T(i(n)) = i(S(n))$. –

Damit haben wir also den richtigen Begriff gefunden – vorausgesetzt, wir können auch einen Existenzsatz nachreichen, der sicherstellt, daß es überhaupt eine Dedekind-Struktur gibt. Die Konstruktion einer solchen Struktur ist eine Aufgabe der Mengenlehre. Eine Möglichkeit, die alles aus der leeren Menge gewinnt, ist die folgende:

Wir definieren die Menge D als die kleinste Menge, die die leere Menge als Element enthält, und die mit jedem n auch die Menge $n \cup \{ n \}$ als Element enthält. Wir setzen dann $d = \emptyset$ und $S(n) = n \cup \{ n \}$ für alle $n \in D$. Dann ist, wie man zeigen kann, (D, S, d) eine Dedekind-Struktur. Die Existenz der Menge D wird dabei durch das sog. Unendlichkeitsaxiom der Mengenlehre sichergestellt, das im wesentlichen direkt fordert, daß D existiert. Bei diesem Vorgehen gilt dann:

$0 = \emptyset$,

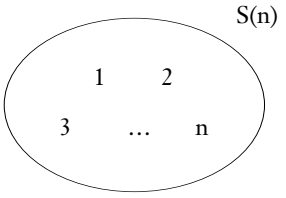
$1 = S(0) = 0 \cup \{ 0 \} = \{ 0 \}$,

$2 = S(1) = 1 \cup \{ 1 \} = \{ 0, 1 \}$,

$3 = S(2) = 2 \cup \{ 2 \} = \{ 0, 1, 2 \}$,

…

$S(n) = n \cup \{ n \} = \{ 0, \ldots, n \}$.



Jede natürliche Zahl ist hier also identisch mit der Menge der natürlichen Zahlen, die durch den Zählprozeß bis hin zu n entstanden sind. Insbesondere hat jede natürliche Zahl n genau n verschiedene Elemente, und die $\in$-Relation übernimmt die Rolle der Ordnung $<$ auf den natürlichen Zahlen.

Viele andere Konstruktionen sind möglich. Als Nachfolgerbildung können wir z. B. auch $S(n) = \{ n \}$ wählen. Dann ist $0 = \emptyset$, $1 = \{ 0 \}$, $2 = \{ 1 \} = \{ \{ \emptyset \} \}$, usw. Weiter ist auch eine Konstruktion der Form $0 = \emptyset$, $1 = (0)$, $2 = (0, 0)$, $3 = (0, 0, 0)$, usw. denkbar, die vielleicht am ehesten der Vorstellung $|$, $||$, $|||$, … entspricht.

In allen Fällen erhalten wir eine Dedekind-Struktur, und nach dem Isomorphiesatz sind alle Zugänge prinzipiell gleichwertig. Die spezielle Nachfolgerbildung $S(n) = n \cup \{ n \}$ besticht allerdings durch ihre guten Eigenschaften und gilt heute als der kanonische Ansatz zur Definition der natürlichen Zahlen innerhalb einer mengentheoretischen Umgebung.

Die natürlichen Zahlen

Wir nehmen im folgenden irgendeine feste Dedekind-Struktur $(\mathbb{N}, S, d)$ als gegeben an und nennen die Elemente von $\mathbb{N}$ *natürliche Zahlen*. Wir verwenden die üblichen Zahlzeichen:

$0 = d$, $1 = S(0)$, $2 = S(1) = S(S(d))$, $3 = S(2)$, usw.

Bei der Einführung der Arithmetik und der Ordnung auf $\mathbb{N}$ greifen wir nur auf die Struktur-Eigenschaften von Dedekind-Strukturen zurück und nicht auf die Eigenheiten einer speziellen Konstruktion.

Die Arithmetik der natürlichen Zahlen

Mit Hilfe von Rekursion können wir die gesamte Arithmetik auf $\mathbb{N}$ einführen:

Definition *(Addition auf $\mathbb{N}$)*
Wir definieren für alle $m \in \mathbb{N}$ durch Rekursion über $n \in \mathbb{N}$:

$m + 0 = m$,

$m + S(n) = S(m + n)$ für alle $n \in \mathbb{N}$.

Die Funktion $+ : \mathbb{N}^2 \to \mathbb{N}$ heißt die *Addition* auf $\mathbb{N}$. Für alle $m, n \in \mathbb{N}$ heißt $m + n$ die *Summe* von m und n, und m und n heißen die *Summanden* des Terms $m + n$.

Speziell gilt dann also:

$m + 1 = m + S(0) = S(m + 0) = S(m)$ für alle $m \in \mathbb{N}$.

Mit Hilfe der Addition können wir nun auch die Multiplikation rekursiv definieren:

Definition *(Multiplikation auf $\mathbb{N}$)*
Wir definieren für alle $m \in \mathbb{N}$ durch Rekursion über $n \in \mathbb{N}$:

$m \cdot 0 = 0$,

$m \cdot S(n) = (m \cdot n) + m$ für alle $n \in \mathbb{N}$.

Die Funktion $\cdot : \mathbb{N}^2 \to \mathbb{N}$ heißt die *Multiplikation* auf $\mathbb{N}$. Für alle $m, n \in \mathbb{N}$ heißt $m \cdot n$ das *Produkt* von m und n, und m und n heißen die *Faktoren* des Terms $m \cdot n$.

Wir schreiben wie üblich oft auch mn anstelle von $m \cdot n$. Weiter vereinbaren wir, daß die Multiplikation stärker bindet als die Addition, d. h. $mn + k$ ist die Summe $(mn) + k$ und nicht etwa das Produkt $m(n + k)$.

Mit Hilfe der Multiplikation wird schließlich auch noch die Exponentiation auf $\mathbb{N}$ rekursiv eingeführt:

Definition *(Exponentiation auf $\mathbb{N}$)*
Für alle $m \in \mathbb{N}$ definieren wir durch Rekursion über $n \in \mathbb{N}$:

$m^0 = 1$,

$m^{S(n)} = m^n \cdot m$ für alle $n \in \mathbb{N}$.

Die Funktion $\cdot^{\cdot} : \mathbb{N}^2 \to \mathbb{N}$ heißt die *Exponentiation* auf $\mathbb{N}$. Für alle $m \in \mathbb{N}$ heißt weiter die Funktion $m^{\cdot} : \mathbb{N} \to \mathbb{N}$ die Exponentiation zur *Basis* m. Für alle $m, n \in \mathbb{N}$ heißt m^n die *Potenz* von m und n, und m heißt die *Basis* und n der *Exponent* des Terms m^n.

Aus diesen Definitionen kann man alle vertrauten Eigenschaften der arithmetischen Operationen auf den natürlichen Zahlen ableiten. So gelten universell, d. h. für alle natürlichen Zahlen n, m, k, … die Gesetze:

$n + (m + k) = (n + m) + k$, $n \cdot (m \cdot k) = (n \cdot m) \cdot k$, *(Assoziativität)*

$n + m = m + n$, $n \cdot m = m \cdot n$, *(Kommutativität)*

$n \cdot (m + k) = n \cdot m + n \cdot k$, *(Distributivität)*

$n^m \cdot n^k = n^{m+k}$, $(n^m)^k = n^{m \cdot k}$, $n^k \cdot m^k = (n \cdot m)^k$. *(Exponentiationsregeln)*

Aus der Assoziativität und Kommutativität der Addition folgen viele weitere einfache Dinge, wie zum Beispiel

$n + S(m) = n + (m + 1) = n + (1 + m) = (n + 1) + m = S(n) + m$.

Die Ordnung der natürlichen Zahlen

Mit Hilfe der Addition läßt sich nun die Ordnung auf den natürlichen Zahlen definieren. Wir setzen für $n, m \in \mathbb{N}$:

$n \leq m$, *falls* „es gibt ein $k \in \mathbb{N}$ mit $n + k = m$“.

Wir schreiben wie üblich auch $n < m$, falls $n \leq m$ und $n \neq m$.

Die Relation $\leq$ erweist sich als eine lineare Ordnung, d. h. für alle natürlichen Zahlen n, m, k gilt:

$n \leq n$, *(Reflexivität)*

$n \leq m$ und $m \leq n$ *impliziert* $n = m$, *(Antisymmetrie)*

$n \leq m$ und $m \leq k$ *impliziert* $n \leq k$, *(Transitivität)*

$n \leq m$ *oder* $m \leq n$. *(Vergleichbarkeit, Linearität)*

Daß $\leq$ eine partielle Ordnung ist, ist eine gute Übungsaufgabe. Wir zeigen hier die Linearität. Dabei verwenden wir einfache Eigenschaften wie zum Beispiel „$0 + n = n$ für alle $n \in \mathbb{N}$“, die man leicht nachweisen kann (in diesem Fall z. B. durch Induktion oder durch Kommutativität $0 + n = n + 0 = n$).

Satz *(Linearität der Ordnung $\leq$)*
Für alle n, m gilt $n \leq m$ oder $m \leq n$.

Beweis
Sei $m \in \mathbb{N}$ beliebig. Wir setzen:

$L = \{ n \in \mathbb{N} \mid n \leq m \}$, $R = \{ n \in \mathbb{N} \mid m \leq n \}$, sowie

$X = L \cup R$.

Wir zeigen durch Induktion, daß $X = \mathbb{N}$.

Induktionsanfang:

Es gilt $0 + n = n$, also ist $0 \leq n$ und damit $0 \in L \subseteq X$.

Induktionsschritt von n nach S(n):

Nach Induktionsvoraussetzung ist $n \in X = L \cup R$.
Wir unterscheiden zwei Fälle (und im ersten Fall zwei Unterfälle).

1. Fall: $n \in L$.

Sei $k \in \mathbb{N}$ mit $n + k = m$.
Ist $k = 0$, so ist $n = m$ und $m + 1 = S(n)$, also $S(n) \in R \subseteq X$.
Andernfalls ist $k = S(k')$ für ein k' (da $\mathrm{rng}(S) = \mathbb{N} - \{ 0 \}$). Dann gilt

$$S(n) + k' = n + S(k') = n + k = m.$$

Also ist $S(n) \in L \subseteq X$.

2. Fall: $n \in R$.

Sei $k \in \mathbb{N}$ mit $m + k = n$. Dann ist $m + S(k) = S(m + k) = S(n)$,
also gilt $S(n) \in R \subseteq X$.

Starke Induktion und Prinzip des kleinsten Elements

Die Ordnung $\leq$ auf den natürlichen Zahlen erlaubt es uns, ein starkes neues Induktionsprinzip zu formulieren und aus dem alten abzuleiten. Bei diesem Prinzip haben wir die Induktionsvoraussetzung nicht nur für eine Zahl zur Verfügung, sondern für alle „bereits durchlaufenen" Zahlen. Wir definieren hierzu:

Definition *(die Anfangsstücke W(n))*
Für alle $n \in \mathbb{N}$ setzen wir

$$W(n) = \{ m \in \mathbb{N} \mid m < n \} \quad \text{für alle } n \in \mathbb{N}.$$

und nennen $W(n)$ das durch n gegebene *Anfangsstück* von $\mathbb{N}$.

Wir zeigen nun:

Satz *(starke Induktion)*
Sei $X \subseteq \mathbb{N}$ und für alle $n \in \mathbb{N}$ gelte:

(+) $W(n) \subseteq X$ *impliziert* $n \in X$.

Dann gilt $X = \mathbb{N}$.

Beweis
Wir zeigen durch Induktion nach n, daß $W(n) \subseteq X$ für alle $n \in \mathbb{N}$ gilt.

Induktionsanfang:

Trivialerweise ist $W(0) = \varnothing \subseteq X$.

Induktionsschritt von n nach S(n):

Nach I. V. gilt $W(n) \subseteq X$. Nach (+) ist also auch $n \in X$. Folglich ist $W(S(n)) = W(n) \cup \{ n \} \subseteq X$.

Damit können wir wie folgt zeigen, daß eine Eigenschaft $\mathscr{E}(n)$ für alle $n \in \mathbb{N}$ gilt: Wir zeigen für ein beliebiges n, daß $\mathscr{E}(n)$ gilt, und dürfen dabei annehmen, daß $\mathscr{E}(m)$ für alle $m < n$ bereits gezeigt ist.

Analog können wir bei der rekursiven Definition einer Funktion f mit Definitionsbereich $\mathbb{N}$ alle Werte f(m), $m < n$, zur Definition von f(n) verwenden.

Wir folgern aus der starken Induktion noch eine fundamentale Eigenschaft der natürlichen Zahlen.

Korollar *(Prinzip des kleinsten Elements)*

Sei $X \subseteq \mathbb{N}$ nichtleer. Dann besitzt X ein kleinstes Element (bzgl. $\leq$), d. h. es gibt ein $n \in X$ derart, daß $n \leq m$ für alle $m \in X$ gilt.

Beweis

Annahme nicht. Wir zeigen durch starke Induktion, daß $n \notin X$ für alle $n \in \mathbb{N}$ gilt, *im Widerspruch* zu X nichtleer.

Induktionsschritt n:

Nach I. V gilt $m \notin X$ für alle $m < n$, d. h. es gilt $W(n) \cap X = \emptyset$. Dann gilt aber auch $n \notin X$, da sonst n das kleinste Element von X wäre.

De facto sind die starke Induktion und das Prinzip des kleinsten Elements rein logisch äquivalent. Wir diskutieren dies in den Übungen.

Allgemein heißt eine lineare Ordnung $\leq$ auf einer Menge M eine *Wohlordnung*, falls jedes nichtleere $X \subseteq M$ ein kleinstes Element besitzt. Die Ordnung auf den natürlichen Zahlen ist also eine Wohlordnung.

Das Prinzip des kleinsten Elements wird oft wie folgt angewendet. Wir wollen wieder zeigen, daß $X = \mathbb{N}$ für eine Teilmenge X von $\mathbb{N}$ gilt. *Annahme*, dies ist nicht der Fall. Dann ist die Komplementmenge $Y = \mathbb{N} - X$ nichtleer und besitzt also ein kleinstes Element n*. (Ist hier $X = \{ n \in \mathbb{N} \mid \mathscr{E}(n) \}$, so ist n* also das kleinste „Gegenbeispiel" für die Eigenschaft $\mathscr{E}(n)$.) Wir argumentieren nun solange, bis wir ein $n' < n^*$ gefunden haben, das ebenfalls ein Element von Y ist. Damit ist dann der erwünschte *Widerspruch* erreicht, denn n′ widerspricht der minimalen Wahl von n*.

Rekursive Funktionen und algorithmische Berechenbarkeit

Wir hatten die arithmetischen Operationen auf den natürlichen Zahlen rekursiv eingeführt. Als Ausblick betrachten wir nun das Konzept einer rekursiv definierten Funktion noch genauer und gelangen in zwei Stufen zu einer einfachen Definition von „algorithmisch berechenbar".

Für die erste Stufe benötigen wir Basisfunktionen, Kompositionen und die sogenannte primitive Rekursion.

Als *Basisfunktionen* bezeichnen wir die folgenden Funktionen:

$Z : \mathbb{N} \to \mathbb{N}$, $\quad Z(n) = 0$ für alle n, *(Nullfunktion)*

$S : \mathbb{N} \to \mathbb{N}$, $\quad S(n) = n + 1$ für alle n, *(Nachfolgerfunktion)*

$I_i^k : \mathbb{N}^k \to \mathbb{N}$, $\quad I_i^k(n_1, \ldots, n_k) = n_i$ für alle $n_1, \ldots, n_k$, *(Projektionen)*

wobei die Projektionen für alle $k \geq 1$ und alle $1 \leq i \leq k$ definiert sind.

Sind $h : \mathbb{N}^k \to \mathbb{N}$ und $g_i : \mathbb{N}^{k'} \to \mathbb{N}$, $1 \leq i \leq k$, Funktionen, so ist die *Komposition* dieser Funktionen, in Zeichen $h(g_1, \ldots, g_k)$, die Funktion $f : \mathbb{N}^{k'} \to \mathbb{N}$ mit

$f(n_1, \ldots, n_{k'}) = h(g_1(n_1, \ldots, n_{k'}), \ldots, g_k(n_1, \ldots, n_{k'}))$ für alle $n_1, \ldots, n_{k'}$.

Seien $h : \mathbb{N}^{k+1} \to \mathbb{N}$ und $g : \mathbb{N}^{k-1} \to \mathbb{N}$ Funktionen für ein $k \geq 1$. (Ist $k = 1$, so fassen wir g als ein Element von $\mathbb{N}$ auf.) Dann ist die *primitive Rekursion* von g und h, in Zeichen $\mathrm{rec}(g, h)$, die Funktion $f : \mathbb{N}^k \to \mathbb{N}$ mit:

$f(0, p_1, \ldots, p_{k-1}) = g(p_1, \ldots, p_{k-1})$, *(Rekursionsanfang)*

$f(n+1, p_1, \ldots, p_{k-1}) = h(n, f(n, p_1, \ldots, p_{k-1}), p_1, \ldots, p_{k-1})$ *(Rekursionsschritt)*

für alle n und alle „Parameter" $p_1, \ldots, p_{k-1} \in \mathbb{N}$.

Damit können wir nun definieren:

Definition *(primitiv rekursive Funktionen)*

Die *primitiv rekursiven Funktionen* sind wie folgt definiert:

(i) Jede Basisfunktion ist primitiv rekursiv.

(ii) Eine Komposition $h(g_1, \ldots, g_k)$ primitiv rekursiver Funktionen $h, g_1, \ldots, g_k$ ist primitiv rekursiv.

(iii) Eine primitive Rekursion $\mathrm{rec}(g, h)$ primitiv rekursiver Funktionen g und h ist primitiv rekursiv.

Für ein $A \subseteq \mathbb{N}^k$ definieren wir die *Indikatorfunktion* $\mathrm{ind}_A : \mathbb{N}^k \to \{0, 1\}$ von A durch $\mathrm{ind}_A(n_1, \ldots, n_k) = 1$, falls $(n_1, \ldots, n_k) \in A$, und $\mathrm{ind}_A(n_1, \ldots, n_k) = 0$, sonst. Damit können wir unseren Rekursionsbegriff auf Relationen übertragen:

Definition *(primitiv rekursive Relationen)*

Eine Relation $A \subseteq \mathbb{N}^k$, $k \geq 1$, heißt *primitiv rekursiv*, falls $\mathrm{ind}_A : \mathbb{N}^k \to \mathbb{N}$ primitiv rekursiv ist.

Es ist leicht zu sehen, daß die Addition, Multiplikation und Exponentiation auf $\mathbb{N}$ primitiv rekursiv sind. Allgemeiner kann eine Fülle von Funktionen und Relationen als primitiv rekursiv nachgewiesen werden. Alle primitiv rekursiven Funktionen sind im intuitiven Sinne algorithmisch berechenbar, und konkret können wir die vermöge der Schemata (i) – (iii) vorliegende Definition einer primitiv rekursiven Funktion $f : \mathbb{N}^k \to \mathbb{N}$ einem Computer beibringen und die Werte $f(n_1, \ldots, n_k)$ für beliebige $n_1, \ldots, n_k$ ausrechnen, wenn wir physikalische

Limitationen vernachlässigen. Umgekehrt stellt sich die Frage, ob jede durch einen Computer berechenbare Funktion primitiv rekursiv ist. Die Antwort ist nein. Es zeigt sich, daß geschachtelte Rekursionen durch die primitive Rekursion nicht in allen Fällen abgedeckt sind. Ein Beispiel ist die *Ackermann-Funktion* $ac : \mathbb{N}^2 \to \mathbb{N}$, die definiert wird durch

$$ac(m, n) = \begin{cases} 2n, & \textit{falls } m = 0, \\ 1, & \textit{falls } m > 0,\ n = 0, \\ ac(m-1, ac(m, n-1)) & \textit{sonst}. \end{cases}$$

Der Leser möge sich das enorme Wachstum dieser Funktion vor Augen führen, indem er einige Werte ac(m, n) für kleine m und n berechnet. In der Tat kann man zeigen, daß die Funktion $h : \mathbb{N} \to \mathbb{N}$ mit h(n) = ac(n, n) für alle n schneller wächst als jede primitiv rekursive Funktion. Die Funktionen h und ac sind damit nicht primitiv rekursiv. Andererseits sind diese Funktionen mit Hilfe eines Computers berechenbar.

5	10	32	$2^{2^{16}}$			
4	8	16	2^{16}	...		
3	6	8	16	2^{16}	...	
2	4	4	4	4	4	4
1	2	2	2	2	2	2
0	0	1	1	1	1	1
	0	**1**	**2**	**3**	**4**	**5**

Ein anderes Beispiel für eine berechenbare, aber nicht primitiv rekursive Funktion erhalten wir durch Diagonalisierung. Wir können alle einstelligen primitiv rekursiven Funktionen auflisten als $f_0, f_1, f_2, \ldots, f_n, \ldots$ Nun definieren wir eine Funktion $g : \mathbb{N} \to \mathbb{N}$ „diagonal" durch $g(n) = f_n(n) + 1$ für alle n. Dann ist $g \neq f_n$ für alle n. Die Funktion g läßt sich, mit etwas Programmieraufwand, tatsächlich mit Hilfe eines Computers berechnen.

Wir müssen also unsere Definition der primitiv rekursiven Funktion erweitern, und im Hinblick auf das sehr allgemeine Diagonalargument muß diese Erweiterung substantiell sein und eine neue Idee ins Spiel bringen. Erstaunlicherweise sind wir aber nur einen Schritt von einer umfassenden Definition entfernt: In Berechnungen können wir eine „unbeschränkte Suche" starten, die entweder ein Ergebnis liefert oder aber divergiert, d. h. unendlich lange läuft. Diese unbeschränkte Suche fügen wir nun zu den primitiv rekursiven Funktionen hinzu.

Ist $g : \mathbb{N}^{k+1} \to \mathbb{N}$, $k \geq 1$, eine Funktion, so ist die μ-*Rekursion* von g, in Zeichen μg, die Funktion $f : A \to \mathbb{N}$ mit

$A = \{ (n_1, \ldots, n_k) \mid$ es gibt ein n mit $g(n_1, \ldots, n_k, n) = 0 \} \subseteq \mathbb{N}^k$,

$f(n_1, \ldots, n_k) =$ „das kleinste n mit $g(n_1, \ldots, n_k, n) = 0$" für alle $(n_1, \ldots, n_k) \in A$.

Damit definieren wir nun:

Definition *(partiell rekursiv)*

Eine Funktion $f : A \to \mathbb{N}$, $A \subseteq \mathbb{N}^k$, $k \geq 1$, heißt *partiell rekursiv*, falls es primitiv rekursive Funktionen $g : \mathbb{N}^{k+1} \to \mathbb{N}$ und $u : \mathbb{N} \to \mathbb{N}$ gibt mit $f = u \circ \mu g$.

Eine partiell rekursive Funktion entsteht also aus einer primitiv rekursiven Funktion durch eine unbeschränkte Nullstellensuche, gefolgt von einer abschließenden primitiv rekursiven Umrechnung der Suchergebnisse (beschrieben durch u).

Der Definitionsbereich einer k-stelligen partiell rekursiven Funktion ist im allgemeinen eine echte Teilmenge von $\mathbb{N}^k$. Dieses Merkmal verhindert es, die partiell rekursiven Funktionen erneut durch Diagonalisierung zu transzendieren. Wir können diese Funktionen wieder als $f_0, f_1, \ldots, f_n, \ldots$ auflisten, und wir können $g : A \to \mathbb{N}$, $A \subseteq \mathbb{N}$, definieren durch $g(n) = f_n(n) + 1$, falls $n \in \mathrm{dom}(f_n)$. Aber wir können nicht mehr „$g \neq f_n$ für alle n“ folgern, da im allgemeinen $n \notin \mathrm{dom}(g)$ gilt.

Gilt $\mathrm{dom}(f) = \mathbb{N}^k$ für eine k-stellige partiell rekursive Funktion f, so heißt f auch μ-*rekursiv* oder kurz *rekursiv*.

Man kann beweisen, daß die partiell rekursiven Funktionen genau mit den Funktionen $f : A \to \mathbb{N}$, $A \subseteq \mathbb{N}^k$, $k \geq 1$, zusammenfallen, die durch eine beliebig gewählte höhere Programmiersprache im theoretischen Sinne berechnet werden können. Insbesondere sind die Ackermann-Funktion und die Diagonalisierung der primitiv rekursiven Funktionen μ-rekursiv, und die Diagonalisierung der partiell rekursiven Funktionen ist partiell rekursiv.

Insgesamt ist eine Vielzahl von unterschiedlich motivierten Definitionen des Begriffs „algorithmisch berechenbar“ gefunden worden, die allesamt genau die partiell rekursiven Funktionen liefern. Dies ist ein starkes Argument dafür, daß man einen wichtigen und natürlichen Begriff gefunden hat. Daß alle intuitiv berechenbaren Funktionen unter diesen Begriff fallen, läßt sich aufgrund der Unschärfe der intuitiven Berechenbarkeit naturgemäß nicht beweisen. Diese Hypothese ist als Churchsche These bekannt.

Die μ-rekursiven Funktionen bilden nur einen kleinen Teil aller Funktionen auf den natürlichen Zahlen. Denn ihre Menge ist abzählbar, während die Menge aller Funktionen von $\mathbb{N}$ nach $\mathbb{N}$ überabzählbar ist. Diese Überlegung zeigt die Weite des allgemeinen mathematischen Funktionsbegriffs.

Übungen

Übung 1 *(Nachfolger und Induktion, I)*

Wir betrachten folgenden Beweisversuch der Aussage „Jede Schafherde ist einfarbig.“:

„Wir zeigen die Aussage durch Induktion über die Anzahl $n \geq 1$ der Elemente einer Schafherde H. Für $n = 1$ ist die Aussage klar. Im Induktionsschritt von n nach $n + 1$ seien $s_1, \ldots, s_{n+1}$ paarweise verschiedene Schafe. Nach Induktionsvoraussetzung sind die Herden $H_1 = \{ s_1, \ldots, s_n \}$ und $H_2 = \{ s_2, \ldots, s_{n+1} \}$ einfarbig, und deswegen ist auch $H_1 \cup H_2 = \{ s_1, \ldots, s_{n+1} \}$ einfarbig.“

Wo steckt der Fehler bei dieser Argumentation?

Übung 2 *(Nachfolger und Induktion, II)*
Wie bewerten Sie die Aussage:

„Wer arm ist und einen Cent findet, ist immer noch arm."

Übung 3 *(Dedekind-Strukturen, I)*
Sei (D, S, d) eine Dedekind-Struktur. Zeigen Sie:

$\mathrm{rng}(S) = D - \{ d \}$.

Übung 4 *(Dedekind-Strukturen, II)*
Geben Sie eine Menge D, eine Funktion $S : D \to D$ und ein $d \in D$ an mit:

(i) $S : D \to D - \{ d \}$ ist bijektiv.

(ii) (D, S, d) ist keine Dedekind-Struktur.

Übung 5 *(Dedekind-Strukturen, III)*
Geben Sie eine Menge D, eine Funktion $S : D \to D$ und ein $d \in D$ an mit:

(i) $S : D \to D$ ist bijektiv und es gilt $S(x) \neq x$ für alle $x \in D$.

(ii) Für alle $X \subseteq D$ gilt das Induktionsaxiom (Ind_X).

(iii) (D, S, d) ist keine Dedekind-Struktur.

Übung 6 *(Dedekind-Strukturen, IV)*
Eine Menge M heißt *induktiv*, falls $\emptyset \in M$ und $n \cup \{ n \} \in M$ für alle $n \in M$. Wir nehmen an, daß eine induktive Menge M_0 existiert und setzen $\mathbb{N} = \bigcap \{ M \subseteq M_0 \mid M \text{ ist induktiv} \}$. Zeigen Sie, daß $\mathbb{N}$ induktiv ist und daß $(\mathbb{N}, S, \emptyset)$ mit $S(n) = n \cup \{ n \}$ eine Dedekind-Struktur ist.

Übung 7 *(Dedekind-Strukturen, V)*
Sei $\mathbb{N}$ die über $d = \emptyset$ und die Nachfolgerbildung $S(n) = n \cup \{ n \}$ eingeführte Menge der natürlichen Zahlen. Zeigen Sie, daß für alle $n \in \mathbb{N}$ gilt:

(i) Für alle $m \in n$ ist $m \subseteq n$. (ii) $\bigcup S(n) = n$.

Übung 8 *(Die Arithmetik der natürlichen Zahlen, I)*
Zeigen Sie, daß für alle $n, m \in \mathbb{N}$ gilt:

$n + m = m + n$.

Übung 9 *(Die Arithmetik der natürlichen Zahlen, II)*
Zeigen Sie, daß für alle $n, m, k \in \mathbb{N}$ gilt:

$n + (m + k) = (n + m) + k$.

Übung 10 *(Die Arithmetik der natürlichen Zahlen, III)*
Zeigen Sie, daß für alle $n, m \in \mathbb{N}$ gilt:

$n \cdot m = m \cdot n$.

Übung 11 *(Die Arithmetik der natürlichen Zahlen, IV)*
Zeigen Sie, daß für alle $n, m, k \in \mathbb{N}$ gilt:

$$n \cdot (m + k) = n \cdot m + n \cdot k.$$

Übung 12 *(Die Arithmetik der natürlichen Zahlen, V)*
Zeigen Sie, daß es für alle $n, m \in \mathbb{N}$ mit $m \neq 0$ Zahlen $q, r \in \mathbb{N}$ gibt mit

$$n = m \cdot q + r, \quad 0 \leq r < m.$$

Übung 13 *(Die Arithmetik der natürlichen Zahlen, VI)*
Zeigen Sie, daß für alle $n, m, k \in \mathbb{N}$ gilt:

$$n^m \cdot n^k = n^{m+k}, \quad (n^m)^k = n^{m \cdot k}, \quad n^k \cdot m^k = (n \cdot m)^k.$$

Übung 14 *(Die Arithmetik der natürlichen Zahlen, VII)*
Zeigen Sie, daß für alle $n \in \mathbb{N}$ gilt:

$$2 \cdot \Sigma_{m \leq n}\, m = n(n + 1).$$

Hierbei ist die *Summe* $\Sigma_{m<n}\, a_m$ natürlicher Zahlen a_m rekursiv definiert durch $\Sigma_{m \leq 0}\, a_m = 0$ und $\Sigma_{m \leq n+1}\, a_m = (\Sigma_{m \leq n}\, a_m) + a_{n+1}$.

Übung 15 *(Die Arithmetik der natürlichen Zahlen, VIII)*
Zeigen Sie, daß für alle $n \in \mathbb{N}$ gilt:

$$\Sigma_{m<n}\, (2m + 1) = n^2.$$

Übung 16 *(Die Ordnung der natürlichen Zahlen, I)*
Zeigen Sie, daß die Relation $\leq$ auf $\mathbb{N}$ eine partielle Ordnung ist, d. h. $\leq$ ist reflexiv, antisymmetrisch und transitiv auf $\mathbb{N}$.

Übung 17 *(Die Ordnung der natürlichen Zahlen, II)*
Sei $n \in \mathbb{N}$, und sei $f : \{0, \ldots, n\} \to \{0, \ldots, n\}$. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) f ist bijektiv. (b) f ist injektiv. (c) f ist surjektiv.

Übung 18 *(Starke Induktion und Prinzip des kleinsten Elements, I)*
Sei $X \subseteq \mathbb{N}$ mit den Eigenschaften:

(i) $0, 1 \in X$. (ii) Für alle $n \in X$ ist $n + 2 \in X$.

Zeigen Sie, daß $X = \mathbb{N}$ gilt.

Übung 19 *(Starke Induktion und Prinzip des kleinsten Elements, II)*
Sei

$$A = \{a\,3 + b\,4 \mid a, b \in \mathbb{N}\}.$$

Zeigen Sie, daß $n \in A$ für alle $n \geq 6$ gilt.

Übung 20 *(Starke Induktion und Prinzip des kleinsten Elements, III)*
Sei $X \subseteq \mathbb{N}$ mit den Eigenschaften:

(i) Gibt es keine $m, k < n$ mit $m\,k = n$, so ist $n \in X$.

(ii) Für alle $n, m \in X$ ist $n \cdot m \in X$.

Zeigen Sie, daß $X = \mathbb{N}$ gilt.

Übung 21 *(Starke Induktion und Prinzip des kleinsten Elements, IV)*
Sei $A \subseteq \mathbb{N}$. Sei $X \subseteq A$, und für alle $n \in A$ gelte:

$W(n) \cap A \subseteq X$ *impliziert* $n \in X$.

Zeigen Sie, daß $X = A$ gilt.

Übung 22 *(Starke Induktion und Prinzip des kleinsten Elements, V)*
Sei $X \subseteq \mathbb{N}$ und für alle $n \in \mathbb{N}$ gelte:

(a) $\{ 2k \mid k < n \} \subseteq X$ *impliziert* $2n \in X$,

(b) $\{ 2k \mid k \in \mathbb{N} \} \cup \{ 2k + 1 \mid k < n \} \subseteq X$ *impliziert* $2n + 1 \in X$.

Zeigen Sie, daß $X = \mathbb{N}$, und interpretieren Sie dieses Ergebnis als eine starke Induktion über alle geraden Zahlen gefolgt von einer starken Induktion über alle ungeraden Zahlen.

Übung 23 *(Starke Induktion und Prinzip des kleinsten Elements, VI)*
Formulieren Sie die starke Induktion für ein $X \subseteq \mathbb{N}$ und das Prinzip des kleinsten Elements für ein $Y \subseteq \mathbb{N}$ als formale quantifizierte Aussagen. Wie hängen diese beiden Prinzipien logisch miteinander zusammen?

[Verwenden Sie das Kontrapositionsgesetz.]

Übung 24 *(Starke Induktion und Prinzip des kleinsten Elements, VII)*
Sei $<$ eine lineare Ordnung auf M. Zeigen Sie, daß äquivalent sind:

(a) Jede nichtleere Teilmenge von M besitzt ein kleinstes Element.

(b) Jede absteigende Folge in M ist schließlich konstant, d. h. ist $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Folge in M mit $x_{n+1} \leq x_n$ für alle n, so gibt es ein n_0 derart, daß $x_n = x_{n_0}$ für alle $n \geq n_0$ gilt.

2. Ganze und rationale Zahlen

Aus den natürlichen Zahlen können wir durch „kanonische“ algebraische Konstruktionen zuerst die ganzen und weiter dann die rationalen Zahlen gewinnen. Die Arithmetik und die Ordnung überträgt sich dabei von $\mathbb{N}$ auf $\mathbb{Z}$ und weiter von $\mathbb{Z}$ auf $\mathbb{Q}$. Wir erreichen insgesamt einen Zahlbereich $\mathbb{Q}$, in welchem wir so frei wie möglich addieren, subtrahieren, multiplizieren und dividieren können.

Wir wollen nun diese Erweiterung des Zahlbereichs vorstellen, ohne dabei den algebraischen Jargon in größter Allgemeinheit einzuführen.

Konstruktion der ganzen Zahlen

Wir können zwei natürliche Zahlen a und b immer addieren, aber nicht immer subtrahieren. Diese Beschränkung ist insbesondere bei algebraischen Umformungen hinderlich. Wir erweitern deswegen die natürlichen Zahlen um die sog. negativen Zahlen. Für jede natürliche Zahl n wird eine neue Zahl $-n$ eingeführt, und auf dem so erweiterten Zahlbereich wird eine Addition erklärt derart, daß $n + (-n) = 0$ für alle natürlichen Zahlen gilt. Technisch können wir diese Erweiterung mit Hilfe eines prinzipiell beliebigen Zeichens „–“ durchführen und die ganzen Zahlen $\mathbb{Z}$ als Menge $\mathbb{N} \cup \{ -n \mid n \in \mathbb{N} \}$ definieren, wobei $-n = (-, n)$. Auf dieser Menge läßt sich dann leicht eine Addition, Multiplikation und Ordnung erklären.

Eleganter ist die folgende algebraische Konstruktion der ganzen Zahlen, die auch dem Prinzip „Neues aus Altem“ besser gerecht wird. Die Idee ist hier, ein Paar (n, m) von natürlichen Zahlen als die Subtraktion $n - m$ zu lesen. Dann ist $(n, 0) = n - 0 = n$, aber $(0, n) = 0 - n = -n$. Ebenso ist $(n, m) = (n + k, m + k)$ für alle k, denn es gilt $(n + k) - (m + k) = n - m$. Gewisse Paare werden also miteinander gleichgesetzt, was wir mit Hilfe einer Äquivalenzrelation durchführen können. Dabei muß eine Subtraktion nicht vorausgesetzt werden, denn $n - m = n' - m'$ ist gleichwertig zu $n + m' = n' + m$.

Definition *(Äquivalenzrelation zur Konstruktion der ganzen Zahlen)*
Wir definieren für alle $(n, m), (n', m') \in \mathbb{N} \times \mathbb{N}$:

$(n, m) \sim (n', m')$, *falls* $n + m' = n' + m$.

Man zeigt, daß $\sim$ eine Äquivalenzrelation auf $\mathbb{N}^2$ ist. Damit definieren wir nun:

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_6, © Springer-Verlag Berlin Heidelberg 2010

Definition *(ganze Zahlen)*
Wir setzen $\mathbb{Z} = \mathbb{N}^2/\sim$. Die Elemente von $\mathbb{Z}$ heißen *ganze Zahlen.*

Die Elemente von $\mathbb{Z}$ schreiben wir zur Vereinfachung der Notation in der Form $[n, m]$ anstelle von $(n, m)/\sim$.
Nun definieren wir eine Addition und eine Multiplikation auf $\mathbb{Z}$:

Definition *(Arithmetik auf $\mathbb{Z}$)*
Für alle $[n, m], [n', m'] \in \mathbb{Z}$ setzen wir:

$$[n, m] + [n', m'] = [n + n', m + m'],$$

$$[n, m] \cdot [n', m'] = [n\,n' + m\,m', n\,m' + m\,n'].$$

Man zeigt leicht, daß diese Operationen wohldefiniert sind. Sie lassen sich leicht motivieren, wenn wir wie oben erwähnt $[n, m]$ als $n - m$ lesen. Wenn die vertrauten Rechengesetze gelten sollen, so ist

$$(n - m) \cdot (n' - m') = (n\,n' + m\,m') - (n\,m' + m\,n').$$

Damit können wir die Multiplikation gar nicht anders definieren. Analoge Überlegungen gelten für die Addition.
Wie üblich vereinbaren wir, daß die Multiplikation stärker bindet als die Addition, und daß wir in symbolischen Rechnungen Malpunkte weglassen können. Damit ist dann z. B. $ab + c = (a \cdot b) + c$ für alle ganzen Zahlen a, b, c. Weiter setzen wir $a^0 = [1, 0]$ für alle $a \in \mathbb{Z}$ und definieren rekursiv $a^{n+1} = a^n \cdot a$ für alle $n \in \mathbb{N}$.

Definition *(additiv Inverses, Differenz)*
Für alle $[n, m] \in \mathbb{Z}$ setzen wir

$$-[n, m] = [m, n].$$

Die ganze Zahl $-[n, m]$ heißt das *additiv Inverse* von $[n, m]$. Weiter sei:

$$[n, m] - [n', m'] = [n, m] + (-[n', m']) \quad \text{für alle } [n, m], [n', m'] \in \mathbb{Z}.$$

Die ganze Zahl $[n, m] - [n', m']$ heißt auch die *Differenz* von $[n, m]$ und $[n', m']$.

Damit ist eine Subtraktionsoperation auf den ganzen Zahlen eingeführt.
Die Bezeichnung von $-[n, m]$ als additiv Inverses ist in der Tat gerechtfertigt, denn es gilt:

$$[n, m] - [n, m] = [n, m] + [m, n] = [n + m, m + n] = [n + m, n + m] = [0, 0].$$

Zur weiteren Vereinfachung der Notation schreiben wir

$$n \quad \text{für} \quad [n, 0] \quad \text{für alle } n \in \mathbb{N}.$$

In Übereinstimmung mit dieser Notation sehen wir auch $\mathbb{N}$ als eine Teilmenge von $\mathbb{Z}$ an, indem wir $n \in \mathbb{N}$ und $[n, 0]$ miteinander identifizieren. Diese Identifikation respektiert die arithmetischen Operationen auf den beiden Zahlbereichen, denn es gilt für alle $n, m \in \mathbb{N}$:

$n + m = [n, 0] + [m, 0] = [n + m, 0]$, $n \cdot m = [n, 0] \cdot [m, 0] = [nm, 0]$.

Weiter gilt mit dieser Notation für alle $[n, m] \in \mathbb{Z}$:

$[n, m] = [n, 0] + [0, m] = [n, 0] - [m, 0] = n - m$,

d. h. alle ganzen Zahlen lassen sich als Differenz zweier Elemente des Bereichs $\mathbb{N}$ schreiben, der unseren Ausgangspunkt bildete. Damit ist unser Vorhaben, eine Subtraktion zu ermöglichen, in minimaler Weise durchgeführt.

Rechengesetze und Ordnung der ganzen Zahlen

Wir stellen die wichtigsten Struktureigenschaften der Arithmetik auf den ganzen Zahlen zusammen. Sie dominieren die ganzen Zahlen derart, daß auf die konkrete Definition einer ganzen Zahl a als einer Äquivalenzklasse $[n, m]$ nicht mehr zurückgegriffen werden muß, sobald diese Gesezte etabliert sind.

Satz *(Rechengesetze für $\mathbb{Z}$)*
Für alle $a, b, c \in \mathbb{Z}$ gilt:

(i) $a + (b + c) = (a + b) + c$. *(Assoziativgesetz für die Addition)*

(ii) $a + 0 = a$. *(Neutralität der Null für die Addition)*

(iii) $a - a = 0$. *(Existenz von additiven Inversen)*

(iv) $a + b = b + a$. *(Kommutativgesetz für die Addition)*

(v) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. *(Assoziativgesetz für die Multiplikation)*

(vi) $a \cdot 1 = a$. *(Neutralität der 1 für die Multiplikation)*

(vii) $1 \cdot 1 = 1$, $(-1) \cdot (-1) = 1$. *(Existenz von multiplikativen Inversen für 1, –1)*

(viii) $a \cdot b = b \cdot a$. *(Kommutativgesetz für die Multiplikation)*

(ix) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$. *(Distributivgesetz)*

Der Beweis sei dem Leser zur Übung überlassen.

Schließlich definieren wir eine Ordnung auf den ganzen Zahlen mit Hilfe der Ordnung auf den natürlichen Zahlen.

Definition *(Ordnung auf $\mathbb{Z}$)*
Für alle $[n, m], [n', m'] \in \mathbb{Z}$ setzen wir:

$[n, m] \leq [n', m']$, *falls* $n + m' \leq n' + m$.

Man zeigt, daß $\leq$ eine wohldefinierte lineare Ordnung auf $\mathbb{Z}$ ist. Weiter ist diese Ordnung eine Fortsetzung der Ordnung auf $\mathbb{N}$, d. h. gilt $n \leq m$ in $\mathbb{N}$, so gilt auch $[n, 0] \leq [m, 0]$ in $\mathbb{Z}$. Im Umfeld der Ordnung definieren wir:

Definition *(negativ, positiv, nicht negativ, Betrag)*
Ein $a \in \mathbb{Z}$ heißt *negativ*, falls $a < 0$ gilt, *positiv*, falls $0 < a$ gilt, und *nicht negativ*, falls $a \geq 0$ gilt. Weiter ist der *Betrag* $|a|$ von a definiert durch

$|a| = -a$, falls $a < 0$, und $|a| = a$ sonst.

Für alle $a, b \in \mathbb{Z}$ gilt die *Dreiecksungleichung* $|a + b| \leq |a| + |b|$, sowie die Produktregel $|a\,b| = |a|\,|b|$.

Die Ordnung auf $\mathbb{Z}$ können wir algebraisch charakterisieren, ohne dabei auf die Äquivalenzklassen zurückzugreifen. Denn für alle $a, b \in \mathbb{Z}$ gilt:

$a \leq b$ *gdw* es gibt ein $n \in \mathbb{N}$ mit $a + n = b$.

Diese Definition verwendet die natürlichen Zahlen $\mathbb{N}$. Für Freunde der Zahlentheorie sei erwähnt, daß es sogar eine Möglichkeit gibt, die natürlichen Zahlen rein arithmetisch innerhalb von $\mathbb{Z}$ zu definieren. Denn es gilt der (nichttriviale) zahlentheoretische Satz, daß jede natürliche Zahl eine Summe von vier Quadraten ist. Damit gilt dann für alle $a, b \in \mathbb{Z}$:

$a \leq b$ *gdw* es gibt $c_1, c_2, c_3, c_4 \in \mathbb{Z}$ mit $a + c_1^2 + c_2^2 + c_3^2 + c_4^2 = b$.

Die Ordnung auf den ganzen Zahlen respektiert die Arithmetik, denn für alle ganzen Zahlen a, b, c gilt:

$a \leq b$ *gdw* $a + c \leq b + c$,

$a \leq b$ und $0 \leq c$ *impliziert* $a \cdot c \leq b \cdot c$.

Weiter gilt $a \leq b$ genau dann, wenn $-b \leq -a$.

Rein ordnungstheoretische Struktureigenschaften der Ordnung auf den ganzen Zahlen sind:

Satz *(Struktur der Ordnung auf $\mathbb{Z}$)*
(a) Jedes $a \in \mathbb{Z}$ besitzt einen direkten Vorgänger und einen direkten Nachfolger. *(Existenz von Vorgängern und Nachfolgern)*

(b) Für alle $b \in \mathbb{Z}$ existieren $a, c \in \mathbb{Z}$ mit $a < b < c$. *(Unbeschränktheit)*

Fassen wir $\mathbb{N}$ wie im letzten Kapitel als Zählreihe 0, 1, 2, …, n, S(n), … auf, so ist $\mathbb{Z}$ die nach links fortgesetzte „Zählreihe“

…, P(m), m, …, –2, –1, 0, 1, 2, …, n, S(n), …,

bei der jedes Element m einen eindeutigen Vorgänger P(m) besitzt. Bei dieser Sicht steht nicht so sehr der Wunsch nach algebraischer Abgeschlossenheit im Vordergrund, sondern das einfachere Symmetriebedürfnis, die Sonderrolle der Null aufzuheben, die in $\mathbb{N}$ keinen Vorgänger besitzt. Dieser Ansatz kann analog zur Entwicklung von $\mathbb{N}$ wie im letzten Kapitel durchgeführt werden und liefert einen zu unserer algebraischen Konstruktion isomorphen Bereich der ganzen Zahlen samt Arithmetik und Ordnung.

Konstruktion der rationalen Zahlen

Bei der Betrachtung der Rechengesetze für die ganzen Zahlen fällt auf, daß inverse Elemente für die Addition immer existieren, während 1 und – 1 die einzigen ganzen Zahlen sind, die ein multiplikatives Inverses besitzen: Gilt $a \cdot b = 1$ für ganze Zahlen a, b, so ist $a = 1$ oder $a = -1$. In $\mathbb{Z}$ können wir nicht frei dividieren, so wie wir in $\mathbb{N}$ nicht frei subtrahieren konnten. Dies führt uns zur Konstruktion der rationalen Zahlen. Die Idee ist hier, ein Paar (a, b) ganzer Zahlen a, b mit $b \neq 0$ als Bruch a/b zu lesen. Die Durchführung verläuft analog zur Konstruktion von $\mathbb{Z}$.

Definition *(Äquivalenzrelation zur Konstruktion der rationalen Zahlen)*
Sei $\mathbb{Z}^* = \mathbb{Z} - \{ 0 \}$. Wir definieren für alle $(a, b), (c, d) \in \mathbb{Z} \times \mathbb{Z}^*$:

$(a, b) \sim (c, d)$, *falls* $a \cdot d = c \cdot b$.

Dann ist $\sim$ eine Äquivalenzrelation auf $\mathbb{Z} \times \mathbb{Z}^*$ und wir definieren:

Definition *(rationale Zahlen)*
Wir setzen $\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}^*)/\sim$. Die Elemente von $\mathbb{Q}$ heißen *rationale Zahlen.*

Wir schreiben rationale Zahlen zur Vereinfachung der Notation in der *Bruch-Form* a/b anstelle von $(a, b)/\sim$.
Aus der Definition von $\sim$ folgt die Kürzungsregel

$(ac)/(bc) = a/b$ für alle $a \in \mathbb{Z}$ und $b, c \in \mathbb{Z}^*$.

Ein Bruch a/b mit $a \neq 0$ heißt *gekürzt*, wenn es keine $a' \in \mathbb{Z}$ und $b', c \in \mathbb{Z}^*$, $c \neq 1$, gibt derart, daß $a/b = (a'c)/(b'c)$.
Wir definieren nun eine Addition und eine Multiplikation auf $\mathbb{Q}$:

Definition *(Arithmetik auf $\mathbb{Q}$)*
Für alle $a/b, c/d \in \mathbb{Q}$ setzen wir:

$a/b + c/d = (ad + bc)/(bd)$,

$a/b \cdot c/d = (ac)/(bd)$.

Wieder sind diese Operationen wohldefiniert.

Definition *(multiplikativ Inverses, Division)*
Wir definieren für alle $a/b \in \mathbb{Q}$ mit $a \neq 0$:

$(a/b)^{-1} = b/a$.

Die Zahl $(a/b)^{-1}$ heißt das *multiplikative Inverse* von a/b. Wir setzen:

$(a/b) / (c/d) = (a/b) \cdot (c/d)^{-1}$ für alle $a/b, c/d \in \mathbb{Q}$ mit $c \neq 0$.

Die Zahl $(a/b)/(c/d)$ heißt der *Quotient* der rationalen Zahlen a/b und c/d.

Damit ist eine Divisionsoperation auf den rationalen Zahlen eingeführt. Nach Definition der Multiplikation und der Inversenbildung gilt

$$(a/b)/(c/d) = (a/b) \cdot (c/d)^{-1} = (a/b) \cdot (d/c) = (ad)/(bc).$$

Zur Rechtfertigung der Bezeichnung von $(a/b)^{-1}$ als multiplikatives Inverses rechnen wir:

$$(a/b) / (a/b) = a/b \cdot b/a = (ab)/(ab) = 1/1.$$

Wir können wieder $\mathbb{Z}$ als Teilmenge von $\mathbb{Q}$ auffassen, indem wir jedes $a \in \mathbb{Z}$ mit $a/1 \in \mathbb{Q}$ identifizieren. Die Arithmetik auf $\mathbb{Z}$ wird dadurch respektiert, und für alle $a \in \mathbb{Z}$ und $b \in \mathbb{Z}^*$ gilt

$$a/b = a/1 \cdot 1/b = a/1 \cdot (b/1)^{-1} = a \cdot b^{-1},$$

d. h. alle rationalen Zahlen lassen sich als Quotient zweier ganzer Zahlen schreiben. Damit ist $\mathbb{Q}$ eine minimale Erweiterung von $\mathbb{Z}$, in der Divisionen durchgeführt werden können.

Für alle $q \neq 0$ ist $q^{-1} = 1/q$, denn ist $q = a/b$, so ist $q^{-1} = b/a$ nach Definition des Inversen, und ebenso gilt $1/q = (1/1)/(a/b) = b/a$ nach Definition der Division.

Rechengesetze und Ordnung der rationalen Zahlen

Für die rationalen Zahlen gelten (aufgrund der Analogie der Konstruktion) alle Rechengesetze aus der obigen Tabelle für $\mathbb{Z}$. Zusätzlich gilt nun aber wegen $q^{-1} = 1/q$ für alle $q \neq 0$ wie gewünscht:

Satz *(Existenz von multiplikativen Inversen für rationale Zahlen ungleich 0)*
Für alle $q \in \mathbb{Q} - \{0\}$ gilt $q \cdot q^{-1} = 1$.

Daß die Null eine Sonderrolle bei der Multiplikation bei Vorhandensein einer Subtraktion spielen muß, kann man zum Beispiel so einsehen: Soll $0 + 0 = 0$, $0 \neq 1$ und das Distributivgesetz $(a + b)\, c = ac + bc$ gelten, so kann die Null kein multiplikatives Inverses haben, d. h. es kann kein a geben mit $0 \cdot a = 1$. Denn wegen

$$0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$$

gilt $0 \cdot a = 0$ (durch Subtraktion von $0 \cdot a$ auf beiden Seiten). Eine gute Addition zusammen mit dem Distributivgesetz schließt also die Existenz eines multiplikativen Inversen der Null aus. Um so erfreulicher ist, daß multiplikative Inverse in den rationalen Zahlen für alle $q \neq 0$ existieren.

Die Ordnung auf $\mathbb{Q}$ kann mit Hilfe der Ordnung auf $\mathbb{Z}$ wie folgt definiert werden:

Definition *(Ordnung auf $\mathbb{Q}$)*
Wir setzen für alle $a, b, c, d \in \mathbb{Z}$ mit $b, d > 0$:

$a/b \leq c/d$, *falls* $a \cdot d \leq c \cdot b$.

Dann ist $\leq$ eine wohldefinierte lineare Ordnung auf $\mathbb{Q}$. Wir erweitern die früheren Sprechweisen (positiv, negativ, Betrag, …). Die Ordnung respektiert erneut die Arithmetik auf $\mathbb{Q}$.

Wesentliche Struktureigenschaften der rationalen Ordnung sind:

Satz *(Struktur der Ordnung auf $\mathbb{Q}$)*

(a) Für alle $q < p$ in $\mathbb{Q}$ existiert ein $r \in \mathbb{Q}$ mit $q < r < p$. *(Dichtheit)*

(b) Für alle $q \in \mathbb{Q}$ existieren $p, r \in \mathbb{Q}$ mit $p < q < r$. *(Unbeschränktheit)*

Man kann zeigen, daß diese beiden Bedingungen zusammen mit der Abzählbarkeit von $\mathbb{Q}$ die Ordnung bis auf Isomorphie festlegen, d. h. jede abzählbare, dichte und unbeschränkte lineare Ordnung $(M, \leq)$ ist isomorph zur Ordnung auf den rationalen Zahlen. Wir diskutieren diesen Satz in den Übungen mit Lösungshinweisen.

Körper

Damit wir nicht immer von den „üblichen Rechengesetzen" sprechen müssen, fassen wir die gewonnenen arithmetischen Struktureigenschaften in einem Begriff zusammen.

Definition *(Körper)*

Sei K eine Menge, und seien + und · zweistellige Operationen auf K. Weiter seien $0, 1 \in K$. Dann heißt die Struktur $(K, +, \cdot, 0, 1)$ ein *Körper* mit Addition +, Multiplikation ·, additiv neutralem Element 0 und multiplikativ neutralem Element 1, falls für alle $a, b, c \in K$ gilt:

(i) $a + (b + c) = (a + b) + c$, *(Assoziativgesetz für die Addition)*

(ii) $a + 0 = a$, *(Neutralität der Null)*

(iii) es gibt ein $a' \in K$ mit $a + a' = 0$, *(Existenz von additiven Inversen)*

(iv) $a + b = b + a$, *(Kommutativgesetz für die Addition)*

(v) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. *(Assoziativgesetz für die Multiplikation)*

(vi) $a \cdot 1 = a$, *(Neutralität der Eins)*

(vii) $a \neq 0$ *impliziert* es gibt ein $a' \in K$ mit $a \cdot a' = 1$. *(Existenz von multiplikativen Inversen)*

(viii) $a \cdot b = b \cdot a$. *(Kommutativgesetz für die Multiplikation)*

(ix) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$. *(Distributivgesetz)*

(x) $0 \neq 1$. *(Verschiedenheit der neutralen Elemente)*

Oft schreiben wir kurz K statt $(K, +, \cdot, 0, 1)$.

Man nennt die Aussagen (i) – (x) auch die *Körperaxiome*. Das Wort „Körper" geht auf Dedekind zurück und ist eine heute veraltete Bezeichnung für eine strukturierte Ansammlung von Personen oder Dingen. Es ist heute noch in „Körperschaft" gebräuchlich.

Wir haben gezeigt, daß die rationalen Zahlen $(\mathbb{Q}, +, \cdot, 0, 1)$ einen Körper bilden. Drei weitere wichtige Körper werden wir im folgenden Kapitel kennenlernen.

Gelten für eine Struktur $(K, +, \cdot, 0, 1)$ die Eigenschaften (i) – (vi) und (viii) – (x), so nennt man die Struktur einen *kommutativen Ring (mit 1)*. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot, 0, 1)$ bilden, wie wir gesehen haben, einen derartigen Ring.

Wie oben schreiben wir $-a$ für das (eindeutig bestimmte) a' mit $a + a' = 0$ und a^{-1} für das (eindeutig bestimmte) a' mit $a \cdot a' = 1$, falls $a \neq 0$. Statt $a + (-b)$ schreiben wir auch $a - b$, und statt $a \cdot b^{-1}$ schreiben wir auch a/b.

In einem Körper gelten alle vertrauten Rechenregeln, z. B.:

Satz *(elementare arithmetische Eigenschaften von Körpern)*
Sei K ein Körper. Dann gilt für alle $x, y \in K$:

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

(ii) $x\,y = 0$ *gdw* $x = 0$ oder $y = 0$, *(Nullteilerfreiheit)*

(iii) $(-1)\,x = -x$,

(iv) $(-x)(-y) = x\,y$.

Beweis
Die Aussagen (i) – (iii) seien dem Leser zur Übung überlassen. Wir beweisen hier (iv) unter Verwendung von (iii). Es gilt:

$(-x)(-y) =_{(iii)} (-1)\,x\,(-y) = (-1)(-y)\,x =_{(iii)} (-(-y))\,x = y\,x = x\,y.$ ∎

Weiter gelten in jedem Körper die bekannten Rechenregeln für das Bruchrechnen, z. B.

$a/b + c/d = (ad + bc)/bd$ für alle $a, b, c, d \in K$ mit $b, d \neq 0$.

Angeordnete Körper

Wir wollen nun noch den Ordnungsaspekt in unsere Betrachtungen integrieren.

Definition *(angeordneter Körper)*
Sei K ein Körper, und sei $\leq$ eine lineare Ordnung auf K. Dann heißt $(K, +, \cdot, 0, 1, \leq)$ (oder kurz K) ein *angeordneter Körper*, falls für alle $a, b, c \in K$ gilt:

(i) $a \leq b$ *impliziert* $a + c \leq b + c$,

(ii) $0 \leq a, b$ *impliziert* $0 \leq a \cdot b$.

Die rationalen Zahlen bilden, wie wir gezeigt haben, einen angeordneten Körper. (Analog bilden die ganzen Zahlen einen sog. angeordneten Ring.)

Definition *(Betrag, positiv, negativ)*
Sei K ein angeordneter Körper. Für jedes $x \in K$ ist der *Betrag* von x, in Zeichen $|x|$, definiert durch $|x| = x$, falls $x \geq 0$, und $|x| = -x$, falls $x \leq 0$. Ist $x \in K$ mit $x > 0$, so heißt x ein *positives* Element von K, und gilt $x < 0$, so heißt x ein *negatives Element* von K. Wir setzen

$$K^+ = \{ x \in K \mid 0 < x \}.$$

Die Menge K^+ der positiven Elemente ist abgeschlossen unter der Addition und Multiplikation. Allgemein gelten in angeordneten Körpern alle vertrauten Eigenschaften für die Addition und Negation von positiven und negativen Zahlen. Eine Zusammenstellung dieser Eigenschaften findet der Leser in den Übungen.

Ein angeordneter Körper K hat die *Charakteristik 0*, d.h. es gilt für alle natürlichen Zahlen $n \geq 1$, daß $n1 \neq 0$, wobei $n1 = 1 + \ldots + 1$ das n-fache der 1 von K ist, d.h. wir definieren rekursiv

$$0\,1 = 0, \quad (n+1)1 = n1 + 1 \quad \text{für alle } n \in \mathbb{N}.$$

Die Nullteilerfreiheit eines angeordneten Körpers führt nun dazu, daß jeder angeordnete Körper die rationalen Zahlen umfaßt:

Für alle $n, m \in \mathbb{N}$, $m \neq 0$, identifizieren wir das Körperelement $\pm (n1)\,(m1)^{-1}$ mit $\pm\, n/m \in \mathbb{Q}$. Die Arithmetik von $\mathbb{Q}$ wird, wie man unschwer einsieht, unter dieser Identifikation respektiert. Wir können also ohne Einschränkung annehmen, daß $\mathbb{Q}$ eine Teilmenge jedes angeordneten Körpers ist. Die rationalen Zahlen sind in diesem Sinne der kleinste angeordnete Körper.

Übungen

Übung 1 *(Konstruktion der ganzen Zahlen, I)*
Für alle $(n, m), (n', m') \in \mathbb{N} \times \mathbb{N}$ setzen wir:

$$(n, m) \sim (n', m'), \quad \textit{falls} \quad n + m' = n' + m.$$

Zeigen Sie, daß $\sim$ eine Äquivalenzrelation auf $\mathbb{N}^2$ ist. Welche Eigenschaft der Addition auf $\mathbb{N}$ verwenden Sie zum Beweis der Transitivität?

Übung 2 *(Konstruktion der ganzen Zahlen, II)*
Zeigen Sie, daß die Addition und die Multiplikation auf $\mathbb{Z}$ wohldefiniert sind.

Übung 3 *(Rechengesetze und Ordnung der ganzen Zahlen, I)*
Beweisen Sie möglichst viele der Rechengesetze für die ganzen Zahlen.

Übung 4 *(Rechengesetze und Ordnung der ganzen Zahlen, II)*
Beweisen Sie mit Hilfe des Distributivgesetzes und der Rechenregeln für die Addition, daß $(-1)\cdot(-1) = 1$ gilt (ohne auf die Definition der Addition und der Multiplikation auf $\mathbb{Z}$ zurückzugreifen).

[Betrachten Sie die Summe von $(-1)(-1)$ und (-1).]

Übung 5 *(Rechengesetze und Ordnung der ganzen Zahlen, III)*
Für alle $[n,m], [n',m'] \in \mathbb{Z}$ setzen wir:

$[n,m] \leq [n',m']$, *falls* $n + m' \leq n' + m$.

Zeigen Sie, daß $\leq$ eine wohldefinierte lineare Ordnung auf $\mathbb{Z}$ ist.

Übung 6 *(Rechengesetze und Ordnung der ganzen Zahlen, IV)*
Zeigen Sie, daß für alle ganzen Zahlen a, b gilt:

$a \leq b$ *gdw* „es gibt ein $n \in \mathbb{N}$ mit $a + n = b$“.

Übung 7 *(Konstruktion der rationalen Zahlen, I)*
Für alle $(a,b), (c,d) \in \mathbb{Z} \times \mathbb{Z}^*$ setzen wir:

$(a,b) \sim (c,d)$, *falls* $a \cdot d = c \cdot b$.

Zeigen Sie, daß $\sim$ eine Äquivalenzrelation auf $\mathbb{Z} \times \mathbb{Z}^*$ ist.

Übung 8 *(Konstruktion der rationalen Zahlen, II)*
Zeigen Sie, daß die rationalen Zahlen abzählbar sind, d. h. es gibt eine Folge $q_0, q_1, \ldots, q_n, \ldots, n \in \mathbb{N}$, in der alle rationalen Zahlen vorkommen.

[Zählen Sie gekürzte Brüche a/b ab nach der Summe $|a| + |b|$:
$0,\ 1/1,\ -1/1,\ 2/1,\ -2/1,\ 1/2,\ -1/2,\ 3/1,\ -3/1,\ 1/3,\ -1/3,\ \ldots$]

Übung 9 *(Rechengesetze und Ordnung der rationalen Zahlen, I)*
Wir setzen für alle $a, b, c, d \in \mathbb{Z}$ mit $b, d > 0$:

$a/b \leq c/d$, *falls* $a \cdot d \leq c \cdot b$.

Zeigen Sie, daß $\leq$ eine wohldefinierte lineare Ordnung auf $\mathbb{Q}$ ist.

Übung 10 *(Rechengesetze und Ordnung der rationalen Zahlen, II)*
Zeigen Sie, daß für alle $a, b, c, d \in \mathbb{Z}$ mit $b, d \neq 0$ gilt:

$a/b \leq c/d$ *gdw* $a \cdot b \cdot d^2 \leq c \cdot d \cdot b^2$.

Übung 11 *(Rechengesetze und Ordnung der rationalen Zahlen, III)*
Zeigen Sie, daß die rationalen Zahlen dicht geordnet sind:

Für alle $q < p$ in $\mathbb{Q}$ gibt es ein $r \in \mathbb{Q}$ mit $q < r < p$.

Übung 12 *(Rechengesetze und Ordnung der rationalen Zahlen, IV)*
Sei M abzählbar, und sei ≤ eine dichte und unbeschränkte lineare Ordnung auf M. Zeigen Sie, daß (M, ≤) isomorph zur Ordnung (ℚ, ≤) der rationalen Zahlen ist.

[Zählen Sie M und ℚ auf als $x_0, x_1, \ldots, x_n, \ldots$ bzw. $q_0, q_1, \ldots, q_n, \ldots$, und konstruieren Sie rekursiv Werte $f(x_n)$ so, daß die Elemente $x_0, \ldots, x_n$ in M genauso liegen wie die Elemente $f(x_0), \ldots, f(x_n)$ in ℚ, d. h. es gilt $x_i < x_j$ genau dann, wenn $f(x_i) < f(x_j)$. Definieren Sie ein geeignetes $f(x_n) = q_k$ mit möglichst kleinem Index k, damit alle rationalen Zahlen irgendwann als Wert von f erscheinen. Dann ist die Funktion $f : M \to \mathbb{Q}$ ein Isomorphismus.]

Übung 13 *(Körper, I)*
Sei K ein Körper. Zeigen Sie, daß für alle $x, y \in K$ gilt:

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

(ii) $x\,y = 0$ *gdw* $x = 0$ oder $y = 0$, *(Nullteilerfreiheit)*

(iii) $(-1)\,x = -x$.

Übung 14 *(Körper, II)*
Sei K ein Körper. Zeigen Sie, daß für alle $a, b, c, d \in K$ mit $b, d \neq 0$ gilt:

$a/b + c/d = (ad + bc)/bd$,

wobei x/y definiert ist als xy^{-1} für alle $x, y \in K$, $y \neq 0$.
Beweisen Sie weitere Rechengesetze des Bruchrechnens.

Übung 15 *(Körper, III)*
Führen Sie eine Addition und Multiplikation auf { 0, 1 } ein, sodaß ein Körper auf { 0, 1 } entsteht. Läßt sich dieser Körper anordnen?

Übung 16 *(Körper, IV)*
Konstruieren Sie einen Körper auf { 0, 1, 2, 3 }.

Übung 17 *(Körper, V)*
Zeigen Sie, daß es keinen Körper mit genau sechs Elementen geben kann.

[Betrachten Sie a = 1 + 1 und b = 1 + 1 + 1 und zeigen Sie 1 + 1 + 1 + 1 + 1 + 1 = 0.]

Übung 18 *(Körper, VI)*
Zeigen Sie, daß wir für endliche K in der Körperdefinition die Existenz von multiplikativ inversen Elementen durch die Nullteilerfreiheit ersetzen können, d. h. wir fordern stattdessen, daß $x\,y \neq 0$ für alle $x, y \in K - \{ 0 \}$ gilt.

[Für endliche Mengen K ist jede Injektion $f : K \to K$ eine Bijektion.]

Übung 19 *(Angeordnete Körper, I)*
Sei K ein angeordneter Körper. Zeigen Sie, daß für alle $x, y, x', y' \in K$ gilt:

(i) $x < y$ und $x' < y'$ *impliziert* $x + x' < y + y'$,

(ii) $x < y$ und $0 < z$ *impliziert* $zx < zy$,

(iii) $x, y < 0$ *impliziert* $0 < xy$,

(iv) $0 < x^2$ für alle $x \neq 0$, insbesondere also $0 < 1$,

(v) $x < 0 < y$ *impliziert* $xy < 0$.

Übung 20 *(Angeordnete Körper, II)*
Sei K ein angeordneter Körper. Zeigen Sie, daß für alle $x, y \in K$ gilt:

(i) $|x + y| \leq |x| + |y|$, *(Dreiecksungleichung)*

(ii) $|xy| = |x|\,|y|$. *(Produktregel)*

Übung 21 *(Angeordnete Körper, III)*
Sei K ein angeordneter Körper. Zeigen Sie, daß K die Charakteristik 0 besitzt, d.h. es gilt $n1 \neq 0$ für alle $n \geq 1$, wobei wieder $n1 \in K$ rekursiv definiert wird durch $01 = 0$, $(n+1)1 = n1 + 1$ für alle $n \in \mathbb{N}$.

Übung 22 *(Angeordnete Körper, IV)*
Sei K ein angeordneter Körper. Zeigen Sie, daß die Identifikation von $\pm (n1)\,(m1)^{-1} \in K$ mit $\pm n/m \in \mathbb{Q}$ für alle $n, m \in \mathbb{N}$, $m \neq 0$, die Arithmetik von $\mathbb{Q}$ respektiert.

Übung 23 *(Angeordnete Körper, V)*
Sei K ein Körper. Sei $M \subseteq K$ mit:

(a) K ist die disjunkte Vereinigung von M, $\{0\}$, $-M = \{-x \mid x \in M\}$.

(b) M ist abgeschlossen unter Addition und Multiplikation.

Wir setzen dann für alle $x, y \in K$:

$x \leq y$ *falls* „es gibt ein $z \in M \cup \{0\}$ mit $x + z = y$“.

Zeigen Sie, daß dadurch K zu eine angeordneten Körper mit $K^+ = M$ wird.

Übung 24 *(Angeordnete Körper, VI)*
Sei K ein angeordneter Körper. Dann ist $(K, \leq)$ dicht geordnet, d.h. für alle $x < y$ in K gibt es ein z in K mit $x < z < y$.

3. Reelle und komplexe Zahlen

Wir lernen in diesem Kapitel zwei weitere Zahlbereiche kennen, nämlich die reellen Zahlen und die komplexen Zahlen. Beide Zahlbereiche bilden einen Körper.

Obere Schranken und Suprema

Der Körper $\mathbb{Q}$ der rationalen Zahlen sieht auf den ersten Blick wie ein hervorragendes Modell eines arithmetischen mathematischen Kontinuums aus. Dieser Eindruck entsteht vor allem aufgrund der Eigenschaft der Dichtheit und der Existenz beliebig kleiner Größen: Für alle $q < p$ existiert ein r mit $q < r < p$, und für alle n gilt $1/n \cdot n = 1$. Wir können also das Intervall von 0 bis 1 in beliebig feine Teile zerlegen. Welche Punkte eines Kontinuums sollen hier noch fehlen?

Der Schein trügt, und daß er trügt, ist eine tiefe Erkenntnis der griechischen Mathematik: Es gibt irrationale Zahlen. Die rationalen Zahlen reichen nicht aus, um ein mathematisches Kontinuum zu modellieren. Wir werden im Kapitel über Zahlentheorie das klassische Argument kennenlernen. Hier verfolgen wir einen modernen Gedanken, der die Schwächen der rationalen Zahlen aufdeckt und der de facto viel stärker zeigt, daß den rationalen Zahlen noch „fast alle" Punkte fehlen, um etwas zu bilden, was wir als Kontinuum ansehen.

Der Begriff eines Kontinuums läßt sich in der Sprache der linearen Ordnungen beschreiben. Hierzu führen wir noch einige weitere Begriffe über lineare Ordnungen ein.

Definition *(beschränkt, obere und untere Schranke)*
Sei $(M, \leq)$ eine lineare Ordnung. Ein $X \subseteq M$ heißt *nach oben beschränkt*, falls ein $s \in M$ existiert, sodaß $x \leq s$ für alle $x \in X$ gilt. Jedes derartige s heißt dann eine *obere Schranke* von X in M.
Analog sind die Begriffe *nach unten beschränkt* und *untere Schranke* von X definiert. Ein X heißt *beschränkt (schlechthin)*, falls X sowohl nach oben als auch nach unten beschränkt ist.

Ist X eine beschränkte Menge von Punkten eines anschaulichen Linearkontinuums (eine stetige Linie), so können wir eine beliebige obere Schranke s von X so weit zur Menge hinschieben, bis sie die Menge X berührt. Von einem Kontinuum erwarten wir, daß es diesen Berührpunkt tatsächlich gibt. Analoges gilt für untere Schranken. Wir können diese Berührpunkte leicht formal definieren:

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_7,

Definition *(Supremum, Infimum)*
Sei $(M, \leq)$ eine lineare Ordnung, und sei $X \subseteq M$. Existiert eine kleinste obere Schranke s^* von X, so heißt diese das *Supremum* von X, in Zeichen

$s^* = \sup(X)$.

Existiert eine größte untere Schranke s' von X, so heißt diese das *Infimum* von X, in Zeichen $s' = \inf(X)$.

Es gilt also $s^* = \sup(X)$ genau dann, wenn gilt:

(a) s^* ist eine obere Schranke von X.

(b) Ist s eine obere Schranke von X, so ist $s^* \leq s$.

Lineare Vollständigkeit

Damit können wir nun die entscheidende Eigenschaft definieren, die wir von einem Kontinuum erwarten:

Definition *(vollständig)*
Eine lineare Ordnung $(M, \leq)$ heißt *vollständig*, falls jede nichtleere nach oben beschränkte Teilmenge X von M ein Supremum in M besitzt.

Automatisch besitzt dann jede nach unten beschränkte Teilmenge auch ein Infimum.

Die Vollständigkeit zusammen mit guten arithmetischen Eigenschaften würde uns für ein Modell eines Kontinuums genügen. Wir definieren also:

Definition *(arithmetisches Kontinuum)*
Ein angeordneter Körper $(K, +, \cdot, 0, 1, \leq)$ heißt ein *arithmetisches Kontinuum*, falls die lineare Ordnung $(K, \leq)$ vollständig ist.

Wir zeigen nun, daß die rationalen Zahlen dem Anspruch der Vollständigkeit nicht genügen. Stärker zeigen wir:

Satz *(Überabzählbarkeit einer vollständigen dichten Ordnung)*
Sei $(K, \leq)$ eine dichte und vollständige lineare Ordnung mit mehr als einem Element. Dann ist K überabzählbar. Mit anderen Worten:

Ist $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Folge in K, so gibt es ein $x^* \in K$ mit $x^* \neq x_n$ für alle n.

Beweis
Seien $x_0, x_1, \ldots, x_n, \ldots$ paarweise verschiedene Elemente von K.
Ohne Einschränkung sei $x_0 < x_1$ (sonst vertauschen wir x_0 und x_1).
Wir definieren $g(0) = 0$, $g(1) = 1$, und setzen dann rekursiv für $n \geq 2$:

$g(n)$ = „das kleinste k mit $x_{g(n-2)} < x_k < x_{g(n-1)}$ oder $x_{g(n-1)} < x_k < x_{g(n-2)}$“.

Existiert $g(n)$ nicht, so ist $x^* = (x_{g(n-2)} + x_{g(n-1)})/2$ wie gewünscht.
Ist aber $g(n)$ für alle n definiert, so gilt nach Konstruktion:

$$x_{g(0)} < x_{g(2)} < x_{g(4)} < \ \dots \ < x_{g(5)} < x_{g(3)} < x_{g(1)}.$$

Sei $X = \{ x_{g(2n)} \mid n \in \mathbb{N} \}$. Dann ist X nach oben beschränkt in K, also existiert $x^* = \sup(X)$. Nach Konstruktion von g kann dann aber x^* kein Element der Folge der x_n sein (sonst wäre $x^* = x_{g(n)}$ für ein geeignetes n).

Da ein arithmetisches Kontinuum $(K, +, \cdot, \leq)$ eine vollständige und dichte Ordnung besitzt, kann also ein solches Kontinuum nicht abzählbar sein. Es ist bemerkenswert, daß diese Notwendigkeit nur auf den Ordnungseigenschaften, nicht aber auf dem Vorhandensein arithmetischer Operationen ruht.

Speziell zeigt der Satz:

Korollar *(Unvollständigkeit der rationalen Zahlen)*
Die Ordnung der rationalen Zahlen ist unvollständig.

Beweis
Es gibt eine Folge $\langle q_n \mid n \in \mathbb{N} \rangle$ rationaler Zahlen mit $\{ q_n \mid n \in \mathbb{N} \} = \mathbb{Q}$, etwa die Aufzählung der gekürzten Brüche $q = \pm n/m$, $n \in \mathbb{N}$, $m \in \mathbb{N} - \{ 0 \}$ nach steigender Summe $n + m$:

$$0,\ 1/1,\ -1/1,\ 1/2,\ -1/2,\ 2/1,\ -2/1,\ 1/3,\ -1/3,\ 3/1,\ -3/1,$$
$$1/4,\ -1/4,\ 2/3,\ -2/3,\ 3/2,\ -3/2,\ 4/1,\ -4/1,\ \dots$$

Wir wollen an dieser Stelle das klassische Argument der alten Griechen wenigstens notieren. Sie zeigten, daß die Quadratwurzel aus 2 keine rationale Zahl ist. Anders: Die Länge der Diagonalen eines Einheitsquadrats ist nicht von der Form n/m für alle $n, m \in \mathbb{N}$. Ohne von Wurzeln, irrationalen Zahlen und Diagonalen zu reden, können wir dieses Ergebnis in unserem Kontext so formulieren:

Die Menge $\{ q \in \mathbb{Q} \mid q \leq 0 \text{ oder } q^2 < 2 \}$ besitzt kein Supremum in $\mathbb{Q}$.

Oder noch einmal anders formuliert:

Die Funktion $f: \mathbb{Q} \to \mathbb{Q}$ mit $f(q) = q^2 - 2$ für alle $q \in \mathbb{Q}$ besitzt keine Nullstelle.

Wir zeigen dieses Ergebnis im Kapitel über „Teiler“ im dritten Abschnitt.

Obiger Beweis der Unvollständigkeit von $\mathbb{Q}$ erlaubt zwar nicht unmittelbar, gewisse Zahlgrößen wie die Quadratwurzel aus 2 als irrational zu erkennen, aber er zeigt andererseits auch viel mehr: Wir müssen notwendig überabzählbar viele Punkte zu $\mathbb{Q}$ hinzufügen, um eine vollständige Erweiterung der rationalen Zahlen zu erzeugen. Unser scheinbar so harmloser Anspruch der Existenz von Suprema zwingt uns, nach Objekten zu suchen, die nicht nur das vertraute Reich des Endlichen, sondern auch das sich anschließende Reich des abzählbar Unendlichen verlassen. Damit ist folgendes Unternehmen gewagter als es aussieht:

Konstruktion der reellen Zahlen

Unser Ziel ist die Konstruktion einer vollständigen Erweiterung von $\mathbb{Q}$. Hierzu betrachten wir die Unvollständigkeitsstellen von $\mathbb{Q}$ als Objekte, und wir verwenden diese Objekte dann dazu, die Lücken von $\mathbb{Q}$ zu stopfen. Wir definieren:

Definition *(Schnitt, Lücke)*
Ein Paar (L, R) nichtleerer Teilmengen L, R von $\mathbb{Q}$ heißt ein *(Dedekindscher) Schnitt* in $\mathbb{Q}$, falls gilt:

(i) $L \cap R = \emptyset$, $L \cup R = \mathbb{Q}$.

(ii) $q < r$ für alle $q \in L$ und alle $r \in R$,

(iii) $\sup(L) \in L$, falls $\sup(L)$ existiert.

Ein Schnitt (L, R) heißt eine *Lücke* von $\mathbb{Q}$, falls $\sup(L)$ nicht existiert.

Ein Schnitt (L, R) zerlegt $\mathbb{Q}$ also in einen linken Anteil L und einen rechten Anteil R. Daß wir Suprema im Falle der Existenz stets dem linken Teil zurechnen, ist eine reine Konvention, die oft nützlich ist, zuweilen aber auch hinderlich sein kann. Wir vereinbaren deswegen: Wird ein Paar (L, R) mit (i) und (ii) als Schnitt bezeichnet, obwohl $q^* = \sup(L) \in R$ gilt, so ist damit stillschweigend der wirkliche Schnitt $(L \cup \{ q^* \}, R - \{ q^* \})$ gemeint.

Die Lücken von $\mathbb{Q}$ „markieren" genau die Stellen, an denen $\mathbb{Q}$ nicht vollständig ist. Wir fassen nun diese Lücken und alle anderen Schnitte als „Punkte" auf:

Definition *(Definition von $\mathbb{R}$)*
Wir setzen $\mathbb{R} = \{ (L, R) \mid (L, R) \text{ ist ein Schnitt in } \mathbb{Q} \}$.
Die Elemente von $\mathbb{R}$ heißen *reelle Zahlen* oder *Punkte des Linearkontinuums.*

Schnitte sind durch bestimmte Teilmengen von $\mathbb{Q}$ gegeben. Die Definition von $\mathbb{R}$ basiert damit letztendlich auf der Existenz von $\mathcal{P}(\mathbb{Q})$ oder gleichwertig der Existenz von $\mathcal{P}(\mathbb{N})$.

Wir sehen im folgenden $\mathbb{Q}$ als Teilmenge von $\mathbb{R}$ an, indem wir jedes $q \in \mathbb{Q}$ mit dem Schnitt (L_q, R_q) identifizieren, wobei $L_q = \{ r \in \mathbb{Q} \mid r \leq q \}$ und $R_q = \mathbb{Q} - L_q$.

Auf $\mathbb{R}$ können wir eine natürliche Ordnung einführen:

Definition *(Ordnung der reellen Zahlen)*
Wir definieren für alle Schnitte (L_1, R_1), (L_2, R_2):

$(L_1, R_1) \leq (L_2, R_2)$, *falls* $L_1 \subseteq L_2$.

Es ist leicht zu sehen, daß $\leq$ eine lineare Ordnung auf der Menge der reellen Zahlen ist, die zudem die Ordnung auf $\mathbb{Q}$ respektiert. Weiter gilt nun wie gewünscht:

Satz *(Vollständigkeit der reellen Ordnung)*
$(\mathbb{R}, \leq)$ ist vollständig.

Beweis
Sei $\mathcal{S}$ eine nichtleere nach unten beschränkte Menge von Schnitten in $\mathbb{Q}$.
Wir zeigen, daß $\mathcal{S}$ ein Infimum besitzt. Hierzu setzen wir:

$L^* = \bigcap_{(L,R)\in\mathcal{S}} L, \quad R^* = \mathbb{Q} - L^*.$

Dann ist (L^*, R^*) ein Schnitt. Weiter gilt für alle Schnitte (L', R'):

(L', R') ist eine untere Schranke von $\mathcal{S}$ *gdw*
$L' \subseteq L$ für alle $(L, R) \in \mathcal{S}$ *gdw* $L' \subseteq L^*$.

Damit ist also $(L^*, R^*) = \inf(\mathcal{S})$. ∎

Ebenso definiert $L^* = \bigcup_{(L,R)\in\mathcal{S}} L, \ R^* = \mathbb{Q} - L^*$ das Supremum einer nach oben beschränkten Menge $\mathcal{S}$ von Schnitten, wobei wir hier auf die oben angesprochene Liberalisierung des Schnittbegriffs zurückgreifen, denn hier gilt $\sup(L^*) \in R^*$, falls das Supremum von $\mathcal{S}$ eine rationale Zahl ist, die nicht der Menge $\mathcal{S}$ angehört.

Es bleibt nun übrig, die Arithmetik von $\mathbb{Q}$ nach $\mathbb{R}$ zu liften. Zur Vereinfachung der Notation identifizieren wir einen Schnitt (L, R) im folgenden mit seiner rechten Hälfte R. Dies führt zu keinem Informationsverlust, da $L = \mathbb{Q} - R$ gilt. Damit ist dann zum Beispiel ein Ausdruck „$R < 0$“ definiert, der gleichbedeutend ist mit „$(\mathbb{Q} - R, R) < (L_0, R_0)$“, wobei wieder $L_0 = \{ q \in \mathbb{Q} \mid q \leq 0 \}$.

Definition *(Addition auf $\mathbb{R}$)*
Wir definieren für alle Schnitte R_1, R_2:

$R_1 + R_2 = \{ q + r \mid q \in R_1, r \in R_2 \}.$

Man zeigt, daß $R_1 + R_2$ wieder ein Schnitt ist, und daß die Addition die Körperaxiome (i) – (iv) erfüllt. Zudem setzt sie die Addition auf $\mathbb{Q}$ fort.

Die Einführung der Multiplikation ist ebenso einfach, wird aber durch technische Vorzeichenprobleme etwas behindert. Für jeden Schnitt R wird $|R|$ wie üblich definiert als das additive Inverse $-R$ von R, falls $R < 0$, und als R sonst. Weiter definieren wir das Vorzeichen oder *Signum* $\mathrm{sg}(R)$ eines Schnitts R als 1, falls $R > 0$, als 0, falls $R = 0$, und als -1, falls $R < 0$. Damit können wir nun definieren:

Definition *(Multiplikation auf $\mathbb{R}$)*
Wir definieren für alle Schnitte $R_1, R_2 > 0$:

$R_1 \cdot R_2 = \{ q \cdot r \mid q \in R_1, r \in R_2 \}.$

Weiter sei $0 \cdot R_2 = R_1 \cdot 0 = 0$ für alle Schnitte R_1, R_2.
Damit setzen wir nun für alle Schnitte R_1, R_2:

$R_1 \cdot R_2 = (\mathrm{sg}(R_1)\,\mathrm{sg}(R_2))\,(|R_1| \cdot |R_2|),$

unter Verwendung der üblichen Multiplikation für Vorzeichen.

Die Multiplikation auf $\mathbb{R}$ setzt wieder die Multiplikation auf $\mathbb{Q}$ fort, und ein genaues und etwas mühsames Nachrechnen zeigt dann:

Satz *(über $\mathbb{R}$)*
$(\mathbb{R}, +, \cdot, \leq)$ ist ein angeordneter Körper.

Der Leser möge so viele Körperaxiome nachweisen, bis er das Gefühl hat, daß es einmal auch wieder gut sein muß (siehe Übung 9). Dabei ist eine Beschränkung auf positive Elemente keine große Einschränkung. Die additiven und multiplikativen Inversen sollten aber identifiziert werden.

Aufgrund der Vollständigkeit der Ordnung haben wir unser Ziel also erreicht:

Korollar *($\mathbb{R}$ ist ein Kontinuum)*
Der Körper der reellen Zahlen ist ein arithmetisches Kontinuum.

Das archimedische Axiom

In der Geschichte der Differential- und Integralrechnung spielt die Diskussion um unendlich kleine Größen eine große Rolle, die sich zum Beispiel in der Leibnizschen „dx"-Notation widerspiegelt. Die im 19. Jahrhundert erfolgte Fundierung der Analysis ging einen Weg, der infinitesimale Größen explizit vermied und zeigte, daß man die Analysis ohne diese Größen aufbauen kann. Wir wollen hier noch zeigen, daß und wie obige Konstruktion der reellen Zahlen über Dedekindsche Schnitte infinitesimale Größen ausschließt. Wir betrachten hierzu den folgenden Ordnungsbegriff:

Definition *(archimedisch angeordnete Körper)*
Ein angeordneter Körper K heißt *archimedisch angeordnet*, falls für alle $x, y \in K$ mit $0 < x < y$ gilt:

Es gibt ein $n \in \mathbb{N}$ mit $y \leq nx$. *(archimedisches Axiom)*

Hierbei definieren wir wieder $0x = 0$ und rekursiv $(n + 1)x = nx + x$ für alle $n \in \mathbb{N}$.

In einem archimedisch angeordneten Körper kann man also durch endliche Vervielfachung eines beliebig kleinen positiven Elements jedes andere positive Element übertreffen. Jedes positive Element des Körpers eignet sich als Maßstab, mit dessen Hilfe beliebige andere positive Elemente ausgemessen werden können. Damit ist also die Existenz von unendlich kleinen Körperelementen x (d. h. $0 < x < 1/(n\,1)$ für alle $n \geq 1$) und unendlich großen Körperelementen y (d. h. $y > (n\,1)$ für alle $n \in \mathbb{N}$) ausgeschlossen.

Das archimedische Axiom ist, wie man leicht einsehen kann, jeweils äquivalent zu den folgenden Versionen, für die wir wieder $\mathbb{Q} \subseteq K$ annehmen (vermöge der Identifizierung von $\pm\, n/m \in \mathbb{Q}$ mit $\pm\, (n\,1)/(m\,1) \in K$):

(a) Für alle $y > 0$ gibt es ein $n \in \mathbb{N}$ mit $n > y$.

(b) $0 = \inf(\{ 1/n \mid n \in \mathbb{N}, n \neq 0 \})$.

(c) Es gibt ein $z > 0$ derart, daß $\{ nz \mid n \in \mathbb{N} \}$ nicht nach oben beschränkt ist.

(d) $\mathbb{Q}$ ist dicht in K, d.h. für alle $x < y$ in K gibt es ein $q \in \mathbb{Q}$ mit $x < q < y$.

Bemerkenswerterweise schließt nun die lineare Vollständigkeit infinitesimale Größen aus:

Satz *(lineare Vollständigkeit impliziert archimedische Anordnung)*
Ein arithmetisches Kontinuum $(K, +, \cdot, \leq)$ ist archimedisch angeordnet.

Beweis
Sei $x \in K$, $x > 0$, und sei $X = \{ nx \mid n \in \mathbb{N} \}$. Wir zeigen, daß X nach oben unbeschränkt ist. *Andernfalls* existiert $x^* = \sup(X)$. Dann ist $x^* - x < x^*$, also ist $x^* - x$ keine obere Schranke von X. Sei also $n \in \mathbb{N}$ mit $x^* - x < nx$. Dann ist aber

$$x^* < nx + x = (n + 1)x,$$

also ist x^* keine obere Schranke von X, *Widerspruch*. ∎

Eine Version der Vollständigkeit, die prinzipiell mit infinitesimalen Größen verträglich ist, werden wir im dritten Abschnitt im Kapitel über Grenzwerte kennenlernen.

Charakterisierung der reellen Zahlen

Wir haben mit den reellen Zahlen ein arithmetisches Kontinuum konstruiert. Der folgende Eindeutigkeitssatz besagt nun, daß wir den angeordneten Körper $\mathbb{R}$ als „das“ arithmetische Kontinuum bezeichnen können:

Satz *(Isomorphiesatz für die reellen Zahlen)*
Sei K ein arithmetisches Kontinuum. Dann ist K isomorph zum angeordneten Körper der reellen Zahlen, d.h. es gibt ein bijektives $f : K \to \mathbb{R}$, sodaß für alle $x, y \in K$ gilt:

(i) $f(x + y) = f(x) + f(y)$,

(ii) $f(x \cdot y) = f(x) \cdot f(y)$,

(iii) $x \leq y$ *gdw* $f(x) \leq f(y)$.

Beweis *(Skizze)*
Wir nehmen wieder ohne Einschränkung an, daß $\mathbb{Q} \subseteq K$ gilt.
Wir konstruieren nun einen Isomorphismus $f : K \to \mathbb{R}$ wie folgt.
Auf $\mathbb{Q}$ sei f die Identität, d.h. wir setzen $f(q) = q$ für alle $q \in \mathbb{Q}$.

Für alle anderen $x \in K$ sei

$f(x)$ = „das in $\mathbb{R}$ bestimmte Supremum aller $q \in \mathbb{Q}$, die in K kleiner als x sind".

Dann ist f eine wohldefinierte Bijektion, und wie man mit nicht allzuviel Mühe zeigen kann, ein Isomorphismus zwischen K und $\mathbb{R}$. Hier wird entscheidend benutzt, daß $\mathbb{Q}$ dicht in K und $\mathbb{R}$ ist.

Es gibt also eine überraschend einfache algebraisch-ordnungstheoretische Charakterisierung des klassischen mathematischen Linearkontinuums. Die Vollständigkeit der Ordnung legt einen angeordneten Körper bis auf Isomorphie fest.

Komplexe Zahlen und Quaternionen

Für jede natürliche Zahl $n \geq 1$ lassen sich Elemente des n-dimensionalen Kontinuums $\mathbb{R}^n = \{ (x_1, \ldots, x_n) \mid x_i \in \mathbb{R} \text{ für alle } 1 \leq i \leq n \}$, auch *Vektoren* genannt, einfach addieren, indem man setzt:

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

Diese punktweise Addition von Vektoren hat die mittlerweile vertrauten Struktur-Eigenschaften: Die Operation ist assoziativ und kommutativ, es gibt ein neutrales Element $0 = (0, \ldots, 0)$, und für jeden Vektor $(x_1, \ldots, x_n)$ gibt es ein additiv Inverses, nämlich $(-x_1, \ldots, -x_n)$.

Eine natürliche Frage ist nun: Können wir auf dem $\mathbb{R}^n$ auch eine Multiplikation einführen, sodaß zusammen mit der Addition ein Körper entsteht?

Der naive Ansatz der punktweisen Multiplikation $\cdot_p$ scheitert schon in der Ebene. Diese Operation ist zwar assoziativ und kommutativ, und zudem ist (1, 1) multiplikativ neutral, da $(x, y) \cdot_p (1, 1) = (x \cdot 1, y \cdot 1) = (x, y)$, jedoch gibt es vom Nullvektor $0 = (0, 0)$ verschiedene Vektoren, die kein multiplikatives Inverses besitzen. So ist zum Beispiel die Gleichung $(1, 0) \cdot_p (x, y) = (1, 1)$ unlösbar, da ja $0 \cdot y = 0$ für alle y gilt. Also hat (1, 0) kein multiplikativ inverses Element für die punktweise Multiplikation $\cdot_p$.

Die Antwort auf obige Frage lautet: Es gibt eine Multiplikation auf dem $\mathbb{R}^2$, durch die $\mathbb{R}^2$ zum sog. Körper der komplexen Zahlen wird. Dagegen gibt es keine Körper-Multiplikation auf dem $\mathbb{R}^n$ für alle $n \geq 3$. Weiter gibt es noch zwei „gute" multiplikative Operationen auf dem $\mathbb{R}^4$ und dem $\mathbb{R}^8$. Unter diesen Operationen wird der $\mathbb{R}^4$ zum Zahlbereich der sog. Quaternionen und der $\mathbb{R}^8$ zum Zahlbereich der sog. Oktaven. Die Multiplikation der Quaternionen ist aber nicht mehr kommutativ, und die Multiplikation der Oktaven ist weder kommutativ noch assoziativ. Wir wollen uns hier auf die komplexen Zahlen konzentrieren, die in der mathematischen Analysis eine überragende Rolle spielen. Am Ende des Kapitels geben wir aber die Definition der Quaternionen noch an.

Definition *(die komplexen Zahlen $\mathbb{C}$)*
Wir setzen $\mathbb{C} = \mathbb{R}^2$. Die Elemente von $\mathbb{C}$ heißen *komplexe Zahlen.*
Für alle $(x_1, y_1), (x_2, y_2) \in \mathbb{C}$ setzen wir:

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2,\ y_1 + y_2),$$

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2,\ x_1 y_2 + y_1 x_2),$$

unter Verwendung der reellen Arithmetik auf der rechten Seite.

Wir reichen unten sowohl eine Motivation als auch eine anschauliche Interpretation der Multiplikation nach.

Die beiden Operationen sind, wie man leicht sieht, assoziativ und kommutativ. Weiter ist die komplexe Zahl $0 = (0, 0)$ additiv neutral. Die Zahl $(1, 0)$ ist multiplikativ neutral, denn für alle $(x, y) \in \mathbb{C}$ gilt:

$$(x, y) \cdot (1, 0) = (x \cdot 1 - y \cdot 0,\ y \cdot 1 + x \cdot 0) = (x, y).$$

Für alle $(x, y) \in \mathbb{C}$ ist $(-x, -y)$ additiv invers zu (x, y). Weiter ist für alle $(x, y) \in \mathbb{C}$, $(x, y) \neq 0$, die komplexe Zahl $(x/w, -y/w)$ mit $w = x^2 + y^2$ multiplikativ invers zu (x, y), denn

$$(x, y) \cdot (x/w, -y/w) = ((x^2 + y^2)/w,\ (xy - yx)/w) = (w/w, 0/w) = (1, 0).$$

Schließlich gilt auch das Distributivgesetz, wie man leicht nachrechnet. Damit haben wir gezeigt:

Satz *($\mathbb{C}$ ist ein Körper)*
Die komplexen Zahlen $\mathbb{C}$ bilden einen Körper mit additiv neutralem Element $0 = (0, 0)$ und multiplikativ neutralem Element $1 = (1, 0)$.

Wir können wieder $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auffassen, indem wir jedes $x \in \mathbb{R}$ mit $(x, 0) \in \mathbb{C}$ identifizieren. Diese Identifikation respektiert die reelle Arithmetik, und sie ist auch mit der Notation $1 = (1, 0)$ für das neutrale Element der Multiplikation konsistent. Zudem gilt nun

$$x_0 \cdot (x, y) = (x_0, 0) \cdot (x, y) = (x_0 x,\ x_0 y)$$

für alle $x_0, x, y \in \mathbb{R}$. Damit setzt die komplexe Multiplikation die übliche skalare Multiplikation (Streckung eines Vektors) fort.

Neben $1 = (1, 0)$ ist $(0, 1)$ der zweite kanonische Vektor der Ebene. Er ist eine prominente Figur in der Theorie der komplexen Zahlen:

Definition *(imaginäre Einheit)*
Wir setzen $i = (0, 1)$ und nennen $i \in \mathbb{C}$ die *imaginäre Einheit*.

Für die imaginäre Einheit gilt:

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

d.h. i ist in $\mathbb{C}$ eine Quadratwurzel von -1, und weiter gilt dies dann auch für $-i$.

Das Polynom $z^2 + 1$ hat also die beiden Nullstellen i und $-i$ in $\mathbb{C}$. Allgemeiner gilt der folgende nichttriviale Satz, der die komplexen Zahlen von den reellen Zahlen hervorhebt, und den wir hier ohne Beweis angeben:

Satz *(Fundamentalsatz der Algebra)*
Jedes Polynom $P(z) = \alpha_n z^n + \ldots + \alpha_1 z + \alpha_0$ mit $n \geq 1$ und Koeffizienten $\alpha_i \in \mathbb{C}$, $\alpha_n \neq 0$, besitzt eine Nullstelle in $\mathbb{C}$.

Aus der Gleichung $i^2 = -1$ folgt, daß sich der Körper der komplexen Zahlen nicht anordnen läßt, denn in einem angeordneten Körper ist das Quadrat einer von 0 verschiedenen Zahl immer positiv. Mit dem Zahlbegriff verbinden wir seit der Kindheit ein „größer" und „kleiner", ein „mehr" und „weniger". Beim Übergang von $\mathbb{R}$ nach $\mathbb{C}$ müssen wir diese Intuition lockern, oder uns weigern, die Elemente von $\mathbb{C}$ Zahlen zu nennen. Rechnen können wir mit ihnen hervorragend. Die Zahl i erlaubt zudem eine klammerfreie Darstellung von komplexen Zahlen, die den Rechenaspekt erhöht und den Zahlcharakter dieser Gebilde unterstützt. Für alle $(x, y) \in \mathbb{C}$ gilt nämlich

$$(x, y) = x + i\,y.$$

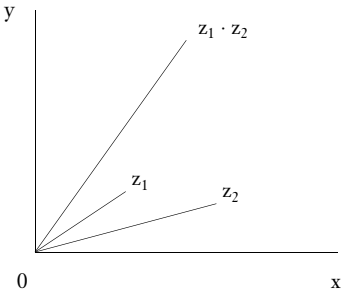
Damit kann eine Rechnung mit komplexen Zahlen als Rechnung mit reellen Zahlen aufgefaßt werden, in der zusätzlich das Objekt i erscheint, das aufgrund der Multiplikations-Eigenschaft $i^2 = -1$ in Rechnungen zuweilen auch wieder verschwindet. Damit kann, ausgehend von reellen Zahlen, ein imaginärer Umweg entstehen, der am Ende der Rechnung wieder in $\mathbb{R}$ landet. In dieser Form wurden die komplexen Zahlen lange verwendet, bevor sie sich dann schließlich als mathematisch reale Objekte ohne imaginäre Anmutung durchsetzten.

Wir haben oben die Multiplikation ad hoc präsentiert. Sie läßt sich aus den Forderungen ableiten, daß $1 = (1, 0)$ multiplikativ neutral und $i^2 = (0, 1)^2 = -1$ sein soll. Damit geht $\mathbb{C}$ in eindeutiger Weise aus der Forderung hervor, daß $(1, 0)$ die Rolle der 1 übernimmt und $(0, 1)$ eine Wurzel von -1 ist.

Die Multiplikation in $\mathbb{C}$ läßt eine sehr sympathische geometrische Interpretation zu. Für jede komplexe Zahl (x, y) definieren wir den *Betrag* oder die *Länge* $|(x, y)|$ von (x, y) durch

$$|(x, y)| = \sqrt{x^2 + y^2}.$$

Der Betrag von $z = x + i\,y$ ist also die Euklidische Länge des Vektors (x, y). Weiter definieren wir das *Argument* von $(x, y) \neq 0$ als den im Gegenuhrzeigersinn gemessenen Winkel des Vektors (x, y) zur positiven x-Achse. Das Produkt zweier komplexer Zahlen erhält man nun wie folgt:



„Multipliziere die Längen der beiden Vektoren und addiere ihre Argumente."
(geometrische Multiplikationsregel)

Die Längenbehauptung dieser Regel läßt sich durch einfaches Nachrechnen beweisen, für die Winkeladdition kann man die trigonometrischen Funktionen heranziehen. Der Beweis fällt bei einer umfassenderen Untersuchung der komplexen Zahlen einfach ab, sobald die Exponentialfunktion eingeführt ist. Man kann aber die Regel ohne jeden analytischen Aufwand beweisen, wenn man elementares geometrisches Argumentieren zuläßt:

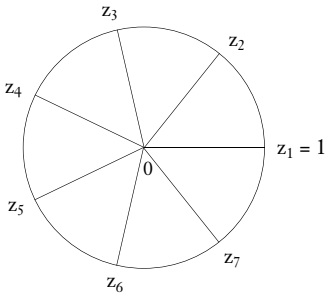
Beweis der Multiplikationsregel durch geometrische Argumentation

Wir lesen die Regel als Definition einer Multiplikation $*$ für Vektoren im $\mathbb{R}^2$. Elementare geometrische Überlegungen zeigen nun, daß $\mathbb{R}^2$ unter dieser Multiplikation zu einem Körper wird. Das Inverse von $(x, y) \neq 0$ ergibt sich z. B. aufgrund der geometrischen Multiplikationsregel durch Spiegelung von (x, y) an der x-Achse und Skalierung auf die inverse Länge. Weiter gilt unter der Multiplikationsregel

$$i * i = -1,$$

denn $i = (0, 1)$ hat Länge 1 und einen rechten Winkel mit der x-Achse, dessen Verdopplung auf die negative Achse führt. Zudem ist $(1, 0)$ multiplikativ neutral, denn $(1, 0)$ hat Länge 1 und Argument 0. Damit haben wir einen Körper $(\mathbb{R}^2, +, *)$ mit neutralem Element $1 = (1, 0)$ und den Eigenschaften $i * i = -1$ für $i = (0, 1)$ vorliegen. Nach obiger Bemerkung ist die Multiplikation $*$ identisch mit der Multiplikation auf $\mathbb{C}$. ∎

Der Leser kann sich mit der geometrischen Multiplikationsregel nun sofort die n verschiedenen Lösungen der komplexen Gleichung $z^n = 1$ visualisieren, die sog. *komplexen n-ten Einheitswurzeln*. Sie bilden ein regelmäßiges n-Eck aus Punkten auf dem Einheitskreis der Ebene. Die Zahl $1 = (1, 0)$ gehört diesem n-Eck stets an.



die Lösungen $z_1, \ldots, z_7$ von $z^7 = 1$

Wir wollen die komplexen Zahlen nun noch etwas genauer untersuchen. Hierzu führen wir noch einige nützliche Begriffe ein.

Definition *(Realteil, Imaginärteil)*

Für alle $z = (x, y) \in \mathbb{C}$ heißt x der *Realteil* und y der *Imaginärteil* von z, in Zeichen $x = \mathrm{Re}(z)$ und $y = \mathrm{Im}(z)$.

Für alle komplexen Zahlen z gilt also

$$z = \mathrm{Re}(z) + i\,\mathrm{Im}(z).$$

Zwei komplexe Zahlen sind genau dann gleich, wenn sie in Real- und Imaginärteil übereinstimmen. Zudem gilt für alle komplexen Zahlen z, z':

$$\mathrm{Re}(z + z') = \mathrm{Re}(z) + \mathrm{Re}(z'), \quad \mathrm{Im}(z + z') = \mathrm{Im}(z) + \mathrm{Im}(z').$$

Definition *(komplexe Konjugation)*
Für alle $z \in \mathbb{C}$ setzen wir $\bar{z} = \mathrm{Re}(z) - i\,\mathrm{Im}(z)$ und nennen $\bar{z}$ die *komplexe Konjugation* von z.

Anschaulich ist die komplexe Konjugation einfach die Spiegelung des Vektors z der Ebene $\mathbb{R}^2$ an der x-Achse. Für alle $z, u, w \in \mathbb{C}$ gelten die folgenden Eigenschaften, deren Beweis wir dem Leser zur Übung überlassen:

(a) $\mathrm{Re}(z) = (z + \bar{z})/2, \quad \mathrm{Im}(z) = (z - \bar{z})/(2i).$

(b) $(\bar{z})^- = z.$

(c) $(u + w)^- = \bar{u} + \bar{w}, \quad (uw)^- = \bar{u}\,\bar{w}.$

(d) $|z|^2 = z\bar{z}.$

Quaternionen

Wir wollen nun noch schildern, wie sich der $\mathbb{R}^4$ mit einer „guten" Multiplikation ausstatten läßt. Hierzu seien

$$1 = e_1 = (1,0,0,0), \quad i = e_2 = (0,1,0,0),$$

$$j = e_3 = (0,0,1,0), \quad k = e_4 = (0,0,0,1).$$

Wir setzen nun

$$1\,1 = 1, \quad 1\,i = i\,1 = i, \quad 1\,j = j\,1 = j, \quad 1\,k = k\,1 = k,$$

$$i^2 = j^2 = k^2 = -1,$$

$$ij = k, \quad ji = -k, \quad ik = -j, \quad ki = j, \quad jk = i, \quad kj = -i.$$

Diese Regeln für die Multiplikation kann man sich leicht merken, wenn man i, j, k zyklisch anordnet. Die Multiplikation eines benachbarten Paares der Reihe i, j, k, i, j springt „vorwärts" auf das nächste Element und „rückwärts" auf das vorangehende Element mit einem Minuszeichen.

Wir definieren nun eine Multiplikation auf dem $\mathbb{R}^4$ durch

$$(x_1, \ldots, x_4) \cdot (y_1, \ldots, y_4) = \Sigma_{1 \le n, m \le 4} (x_n\, y_m)\, (e_n\, e_m) \in \mathbb{R}^4,$$

wobei $x_n\, y_m$ das reelle Produkt von x_n und y_m bezeichnet und in der Summe wie üblich skalar multipliziert und punktweise addiert wird.

Ausrechnen unter Verwendung obiger Werte für $e_n\, e_m$ liefert dann die folgende Produktregel

$$\begin{aligned}(x_1, \ldots, x_4) \cdot (y_1, \ldots, y_4) = {} & (x_1\, y_1 - x_2\, y_2 - x_3\, y_3 - x_4\, y_4)\ 1 + \\ & (x_1\, y_2 + x_2\, y_1 + x_3\, y_4 - x_4\, y_3)\ i + \\ & (x_1\, y_3 - x_2\, y_4 + x_3\, y_1 + x_4\, y_2)\ j + \\ & (x_1\, y_4 + x_2\, y_3 - x_3\, y_2 + x_4\, y_1)\ k.\end{aligned}$$

Berechnen wir $(y_1, \ldots, y_4) \cdot (x_1, \ldots, x_4)$, so ändern sich in der Produktregel genau die Vorzeichen der sechs Terme $x_3\, y_4$, $x_4\, y_3$, $x_4\, y_2$, $x_2\, y_4$, $x_2\, y_3$ und $x_3\, y_2$.

Wir definieren:

Definition *(die Quaternionen $\mathbb{H}$)*
Wir setzen $\mathbb{H} = \mathbb{R}^4$ und versehen $\mathbb{H}$ mit der punktweisen Addition und obiger Multiplikation. Die Elemente von $\mathbb{H}$ heißen *(Hamiltonsche) Quaternionen*.

Für alle $\alpha \in \mathbb{R}$ und $p = (x_1, \ldots, x_4) \in \mathbb{H}$ gilt $(\alpha, 0, 0, 0) \cdot p = (\alpha x_1, \ldots, \alpha x_4) = \alpha p$. Wir können also wieder $\alpha \in \mathbb{R}$ mit $(\alpha, 0, 0, 0) \in \mathbb{H}$ identifizieren. Für alle Quaternionen $p = (x_1, x_2, x_3, x_4)$ gilt dann $p = x_1 + x_2 i + x_3 j + x_4 k$. Wir setzen

$$\bar{p} = x_1 - x_2 i - x_3 j - x_4 k,$$

und nennen wieder $\bar{p}$ die *Konjugierte* von p. Weiter sei $w = x_1^2 + x_2^2 + x_3^2 + x_4^2$. Man rechnet nun leicht nach, daß $p \cdot \bar{p} = \bar{p} \cdot p = w$ gilt. Damit ist also für alle Quaternionen $p \neq 0$ die Quaternion $\bar{p}/w$ multiplikativ invers zu p. Weiter zeigt man nun leicht, daß die Quaternionen einen Schiefkörper bilden, d. h. es gelten alle Körperaxiome mit Ausnahme der Kommutativität der Multiplikation. Zusätzlich gilt das zweite Distributivgesetz $(p_1 + p_2) p_3 = p_1 p_3 + p_2 p_3$, das sich nun ja mangels Kommutativität nicht mehr aus dem ersten ableiten läßt.

Oft ist es nützlich, Quaternionen als Elemente von $\mathbb{C}^2$ aufzufassen. Wir identifizieren also jedes $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ mit $((x_1, x_2), (x_3, x_4)) \in \mathbb{C}^2$. Auf $\mathbb{C}^2$ definieren wir die Addition punktweise und die Multiplikation durch

$$(u, v) \cdot (w, z) = (u w - v \bar{z},\ u z + v \bar{w}) \quad \text{für alle } (u, v), (w, z) \in \mathbb{C}^2,$$

unter Verwendung der komplexen Multiplikation und Konjugation. Abgesehen von den beiden komplexen Konjugationen ist diese Multiplikation strukturell identisch mit der Multiplikation auf $\mathbb{C}$!

Für alle $u = (x_1, x_2)$, $v = (x_3, x_4)$, $w = (y_1, y_2)$, $z = (y_3, y_4) \in \mathbb{C}$ gilt nun:

$$\begin{aligned}
&(u, v) \cdot (w, z) = \\
&((x_1, x_2)(y_1, y_2) - (x_3, x_4)(y_3, -y_4),\ (x_1, x_2)(y_3, y_4) + (x_3, x_4)(y_1, -y_2)) = \\
&((x_1 y_1 - x_2 y_2, x_1 y_2 + x_2 y_1) - (x_3 y_3 + x_4 y_4, -x_3 y_4 + x_4 y_3), \\
&\qquad (x_1 y_3 - x_2 y_4, x_1 y_4 + x_2 y_3) + (x_3 y_1 + x_4 y_2, -x_3 y_2 + x_4 y_1)) = \\
&((x_1 y_1 - x_2 y_2 - x_3 y_3 - x_4 y_4,\ x_1 y_2 + x_2 y_1 + x_3 y_4 - x_4 y_3), \\
&\qquad (x_1 y_3 - x_2 y_4 + x_3 y_1 + x_4 y_2,\ x_1 y_4 + x_2 y_3 - x_3 y_2 + x_4 y_1)),
\end{aligned}$$

und damit haben wir obige Definition wiedergefunden. Die beiden Zugänge zu den Quaternionen über $\mathbb{R}^4$ und $\mathbb{C}^2$ sind also äquivalent. Die Eleganz der $\mathbb{C}^2$-Darstellung zeigt sich z. B. in der Identifikation der multiplikativen Inversen. Für alle $(u, v) \in \mathbb{C}^2$ sei $w = u \bar{u} + v \bar{v} \in \mathbb{R}$. Ist $(u, v) \neq 0$, so ist $w \neq 0$ und es gilt

$$(u, v)(\bar{u}/w, -v/w) = ((u \bar{u} + v \bar{v})/w, (-u v + v u)/w) = (1, 0) = 1.$$

Ebenso gilt auch $(\bar{u}/w, -v/w)(u, v) = 1$ und damit ist $(\bar{u}/w, -v/w)$ multiplikativ invers zu (u, v).

Übungen

Übung 1 *(Obere Schranken und Suprema, I)*
Sei $(M, \leq)$ eine lineare Ordnung. Welche Elemente der Ordnung bezeichnen im Falle der Existenz $\sup(\emptyset)$, $\sup(M)$, $\inf(\emptyset)$, $\inf(M)$?

Übung 2 *(Obere Schranken und Suprema, II)*
Sei $(M, \leq)$ eine lineare Ordnung, und sei $X \subseteq M$. Weiter sei Y die Menge der oberen Schranken von X. Zeigen Sie, daß $\sup(X)$ genau dann existiert, wenn $\inf(Y)$ existiert, und daß in diesem Falle $\sup(X) = \inf(Y)$ gilt.

Übung 3 *(Obere Schranken und Suprema, III)*
Sei $(M, \leq)$ eine lineare Ordnung, und seien $X, Y \subseteq M$.
Seien $x^* = \sup(X)$ und $y^* = \sup(Y)$. Zeigen Sie:

$\max(x^*, y^*) = \sup(X \cup Y)$,

wobei $\max(x^*, y^*) = x^*$, falls $y^* \leq x^*$, und $\max(x^*, y^*) = y^*$ sonst.

Übung 4 *(Obere Schranken und Suprema, IV)*
Sei $(M, \leq)$ eine lineare Ordnung, und sei $\mathcal{S} \subseteq \mathcal{P}(M)$ nichtleer.
Für alle $X \in \mathcal{S}$ existiere das Supremum von X. Zeigen Sie:

$\sup(\bigcup \mathcal{S}) = \sup(\{ \sup(X) \mid X \in \mathcal{S} \})$.

Wie hängt diese Aussage mit der vorhergehenden Übung zusammen? Formulieren Sie zudem eine analoge Aussage für Infima.

Übung 5 *(Lineare Vollständigkeit, I)*
Sei $(M, \leq)$ eine vollständige lineare Ordnung. Zeigen Sie, daß jede nichtleere nach unten beschränkte Teilmenge von M ein Infimum besitzt.

Übung 6 *(Lineare Vollständigkeit, II)*
Sei $(M, \leq)$ eine lineare Ordnung. Für $L, R \subseteq M$ schreiben wir $L \leq R$, falls $x \leq y$ für alle $x \in L$ und $y \in R$ gilt. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(i) $(M, \leq)$ ist vollständig.

(ii) Für alle nichtleeren $L, R \subseteq M$ mit $L \leq R$ gibt es ein z mit $L \leq z \leq R$, d. h. es gilt $x \leq z \leq y$ für alle $x \in L$ und $y \in R$.

Übung 7 *(Lineare Vollständigkeit, III)*
Geben Sie eine genaue Begründung für den Zusatz

„sonst wäre $x^* = x_{g(n)}$ für ein geeignetes n“

am Ende des Beweises der Überabzählbarkeit einer dichten vollständigen Ordnung.

Übung 8 *(Konstruktion der reellen Zahlen, I)*
Wir definieren nun für alle Schnitte (L_1, R_1), (L_2, R_2) in $\mathbb{Q}$:

$(L_1, R_1) \leq (L_2, R_2)$, *falls* $L_1 \subseteq L_2$.

Zeigen Sie, daß $\leq$ eine lineare Ordnung ist.

Übung 9 *(Konstruktion der reellen Zahlen, II)*
Zeigen Sie (durch exemplarischen Nachweis von Körperaxiomen), daß $(\mathbb{R}, +, \cdot, \leq)$ ein angeordneter Körper ist. Identifizieren Sie insbesondere das additive Inverse $-(L, R)$ für alle Schnitte (L, R) sowie das multiplikative Inverse $(L, R)^{-1}$ für alle Schnitte $(L , R) \neq 0$, wobei $0 = (L_0, R_0)$.

Übung 10 *(Das archimedische Axiom, I)*
Sei $(K, +, \cdot, \leq)$ ein angeordneter Körper. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) K ist archimedisch angeordnet, d. h. für alle $0 < x < y$ gibt es ein $n \in \mathbb{N}$ mit $y \leq nx$.

(b) Für alle $y > 0$ gibt es ein $n \in \mathbb{N}$ mit $n > y$.

(c) $0 = \inf(\{ 1/n \mid n \in \mathbb{N}, n \neq 0 \})$.

(d) Es gibt ein $z > 0$, sodaß $\{ nz \mid n \in \mathbb{N} \}$ nicht nach oben beschränkt ist.

(e) $\mathbb{Q}$ ist dicht in K, d. h. für alle $x < y$ in K gibt es ein $q \in \mathbb{Q}$ mit $x < q < y$.

Übung 11 *(Das archimedische Axiom, II)*
Sei $(K, +, \cdot, \leq)$ ein archimedisch angeordneter Körper. Zeigen Sie:

(a) Für alle $x > 1$ und alle y in K gibt es ein $n \in \mathbb{N}$ mit $x^n \geq y$.

(b) Für alle $0 < z < 1$ und alle $0 < y$ in K gibt es ein $n \in \mathbb{N}$ mit $z^n < y$.

Übung 12 *(Das archimedische Axiom, III)*
Sei b eine natürliche Zahl mit $b \geq 2$. Für alle $m \in \mathbb{N}$ und jede Folge $\langle a_n \mid n \geq 1 \rangle$ mit $a_n \in \{ 0, \ldots, b - 1 \}$ für alle $n \geq 1$ sei

$$m, a_1 a_2 \ldots a_n \ldots = m + \sup(\{ \Sigma_{1 \leq i \leq n} a_i / b^i \mid n \geq 1 \}).$$

Ist $x \in \mathbb{R}$ und $x = \pm m, a_1 a_2 \ldots a_n \ldots$, so heißt $\pm m, a_1 a_2 \ldots a_n \ldots$ eine *b-adische Darstellung* von x oder eine *b-adische Bruchentwicklung* von x. Ist $b = 2$, so heißt die Darstellung *dyadisch*, und ist $b = 10$, so heißt die Darstellung eine *Dezimaldarstellung*.

(a) Zeigen Sie, daß das Supremum in der Definition von $m, a_1 a_2 \ldots a_n \ldots$ existiert.

(b) Zeigen Sie, daß $1 = 0,a\,a\,a \ldots a \ldots$ für $a = b - 1$ gilt.

(c) Zeigen Sie, daß jede reelle Zahl mindestens eine und höchstens zwei b-adische Darstellungen besitzt.

Übung 13 *(Charakterisierung der reellen Zahlen)*
Führen Sie den Beweis des Charakterisierungssatzes im Detail aus.

Übung 14 *(Komplexe Zahlen und Quaternionen, I)*
Zeigen Sie, daß das Distributivgesetz für die komplexen Zahlen gilt.

Übung 15 *(Komplexe Zahlen und Quaternionen, II)*
Leiten Sie die Definition der komplexen Multiplikation aus den folgenden Forderungen ab:

(i) $(1, 0)$ ist multiplikativ neutral.

(ii) Es gilt $i^2 = (-1, 0)$ für $i = (0, 1)$.

Nehmen Sie dabei an, daß alle üblichen Rechenregeln gelten.

Übung 16 *(Komplexe Zahlen und Quaternionen, III)*
Zeigen Sie durch elementare geometrische Argumentation, daß $\mathbb{R}^2$ zusammen mit der Vektoraddition und der mit Hilfe der geometrischen Multiplikationsregel definierten Operation $*$ einen Körper bildet.

Übung 17 *(Komplexe Zahlen und Quaternionen, IV)*
Sei $z \in \mathbb{C}$ und sei $n \in \mathbb{N} - \{ 0 \}$. Bestimmen Sie mit Hilfe der geometrischen Multiplikationsregel alle n-ten komplexen Wurzeln von z, d. h. alle $w \in \mathbb{C}$ mit $w^n = z$.

Übung 18 *(Komplexe Zahlen und Quaternionen, V)*
Seien $0 < a < b$ reelle Zahlen, und sei $R = \{ z \in \mathbb{C} \mid a < |z| < b \}$.
Bestimmen Sie die Menge $\{ 1/z \mid z \in R \}$.

Übung 19 *(Komplexe Zahlen und Quaternionen, VI)*
Schreiben Sie die folgenden komplexen Zahlen z und ihre Kehrwerte $1/z$ in der Form $x + iy$ mit $x, y \in \mathbb{R}$:

(a) $(3 - 2i)/(5i - 1)$, (b) i^7, (c) $17i/(i - 3)^2$.

Übung 20 *(Komplexe Zahlen und Quaternionen, VII)*
Zeigen Sie sowohl rechnerisch als auch mit Hilfe der geometrischen Multiplikationsregel, daß für alle $z, u, w \in \mathbb{C}$ gilt:

(a) $\mathrm{Re}(z) = (z + \bar{z})/2$, $\mathrm{Im}(z) = (z - \bar{z})/(2i)$.

(b) $(\bar{z})^- = z$.

(c) $(u + w)^- = \bar{u} + \bar{w}$, $(uw)^- = \bar{u}\,\bar{w}$.

(d) $|z|^2 = z\bar{z}$.

Übung 21 (Komplexe Zahlen und Quaternionen, VIII)
Seien $z, w \in \mathbb{C}$, und es gelte $|z| = 1$ oder $|w| = 1$. Zeigen Sie, daß $|z - w| = |1 - \bar{z}w|$.

Übung 22 (Komplexe Zahlen und Quaternionen, IX)
Wir definieren für alle $u, w \in \mathbb{C}$:

$u * w = (uw)^-$.

Welche Körperaxiome gelten für $(\mathbb{R}^2, +, *)$, und welche sind verletzt?

Übung 23 (Komplexe Zahlen und Quaternionen, X)
Zeigen Sie, daß es auf dem $\mathbb{R}^3$ keine assoziative, kommutative und mit der punktweisen Addition distributive Multiplikation $\cdot : \mathbb{R}^3 \to \mathbb{R}^3$ gibt mit den Eigenschaften:

(i) $(0, 1, 0)(0, 1, 0) = (-1, 0, 0)$,

(ii) $(\alpha, 0, 0)(x, y, z) = (\alpha x, \alpha y, \alpha z)$ für alle $\alpha \in \mathbb{R}$ und $(x, y, z) \in \mathbb{R}^3$.

[Sei $1 = (1, 0, 0)$, $i = (0, 1, 0)$, $j = (0, 0, 1) \in \mathbb{R}^3$. Dann gilt $i^2 = -1$ und

$(\alpha, \beta, \gamma) = \alpha + \beta i + \gamma j$ für alle $\alpha, \beta, \gamma \in \mathbb{R}$,

wobei wir α für $(\alpha, 0, 0)$ schreiben. Seien nun $\alpha, \beta, \gamma \in \mathbb{R}$ definiert durch

$i \cdot j = \alpha + \beta i + \gamma j$.

Dann gilt $-j = (-\beta + \gamma\alpha) + (\alpha + \gamma\beta) i + \gamma^2 j$, also $\gamma^2 = -1$, was unmöglich ist.]

Übung 24 (Komplexe Zahlen und Quaternionen, XI)
Eine Quaternion u heißt *rein imaginär*, falls ihre erste Komponente gleich 0 ist, d.h. u ist von der Form $(0, \alpha, \beta, \gamma) \in \mathbb{R}^4$. Zeigen Sie, daß für rein imaginäre Quaternionen u, v gilt:

(i) Es gibt ein $\alpha \in \mathbb{R}$, $\alpha \leq 0$, mit $u^2 = \alpha$.

(ii) Es gibt ein $\beta \in \mathbb{R}$ mit $uv + vu = \beta$.

(iii) $uv = -(u \circ v) + u \times v$, wobei

$$(0, \alpha, \beta, \gamma) \circ (0, \alpha', \beta', \gamma') = \alpha\alpha' + \beta\beta' + \gamma\gamma',$$

$$(0, \alpha, \beta, \gamma) \times (0, \alpha', \beta', \gamma') = (0, \beta\gamma' - \gamma\beta', -\alpha\gamma' + \alpha'\gamma, \alpha\beta' - \beta\alpha').$$

3. Abschnitt

Erste Erkundungen

1. Teiler

Wir arbeiten im folgenden mit den ganzen Zahlen

$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

Wir interessieren uns in erster Linie für natürliche Zahlen, aber in $\mathbb{Z}$ können wir freier rechnen, weil wir im Gegensatz zu $\mathbb{N}$ beliebig subtrahieren können.

Im folgenden sind a, b, c, n, m, k, … stets Elemente von $\mathbb{Z}$.

Teilbarkeit

Die wichtigste Relation der elementaren Zahlentheorie ist die Teilbarkeit ohne Rest:

Definition *(teilbar, Teiler, Vielfaches, $d \mid a$)*

Ein a heißt *teilbar (ohne Rest)* durch ein d, in Zeichen $d \mid a$, falls gilt:

Es gibt ein k mit $kd = a$.

Die Zahl d heißt dann auch ein *Teiler* von a, und a heißt ein *Vielfaches* von d.

Wir stellen die elementaren Eigenschaften der Teilbarkeitsrelation in einem Satz zusammen.

Satz *(Eigenschaften der Teilbarkeitsrelation)*

Für alle a, b, c, d, …, n, m, … gilt:

(T1) $a \mid a$, $1 \mid a$, $a \mid 0$, $0 \mid a$ *gdw* $a = 0$, $d \mid a$ *gdw* $|d| \mid |a|$,

(T2) $d \mid a$ und $a \mid d$ *gdw* $|d| = |a|$,

(T3) $d \mid a$ und $a \neq 0$ *impliziert* $|d| \leq |a|$,

(T4) $e \mid d$ und $d \mid a$ *impliziert* $e \mid a$,

(T5) $d \mid a$ *impliziert* $d \mid (ab)$ und $(nd) \mid (na)$,

(T6) $(ca) \mid (cb)$ und $c \neq 0$ *impliziert* $a \mid b$,

(T7) $d \mid a$ und $d \mid b$ *impliziert* $d \mid (na + mb)$.

Der Beweis dieser Eigenschaften sei dem Leser zur Übung überlassen.

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_8, © Springer-Verlag Berlin Heidelberg 2010

Die Eigenschaft (T7) wird häufig verwendet, und wir betrachten sie deswegen noch etwas genauer. Hierzu definieren wir:

Definition *(Linearkombination)*
Ein c heißt eine *Linearkombination* von a und b, falls n und m existieren mit

$$c = na + mb.$$

So sind zum Beispiel

$$13 = 2\cdot 4 + 1\cdot 5, \quad 8 = 2\cdot 4 + 0\cdot 5, \quad 0 = 0\cdot 4 + 0\cdot 5, \quad -1 = 1\cdot 4 + (-1)\cdot 5$$

Linearkombinationen der Zahlen 4 und 5. Die Eigenschaft (T7) besagt nun: Ist d ein Teiler von a und von b, so ist d auch ein Teiler jeder Linearkombination von a und b. Eine genaue Beschreibung der möglichen Linearkombinationen zweier Zahlen a und b wird sich im Laufe unserer Untersuchungen ergeben.

Wir betrachten nun die sog. Division mit Rest, die den Versuch beschreibt, eine Zahl a durch eine Zahl d zu teilen. Für d = 3 gilt z. B.

$$14 = 4\cdot d + 2, \quad 18 = 6\cdot d + 0, \quad -5 = -2\cdot d + 1, \quad -9 = -3\cdot d + 0.$$

Für die Reste r = 2, 0, 1, 0 dieser Beispiele gilt $0 \le r < d$. Wir zeigen allgemein:

Satz *(Division mit Rest)*
Für alle a und alle $d \ge 1$ gibt es eindeutig bestimmte q und r mit:

$$a = qd + r, \quad 0 \le r < d.$$

Das „q" steht hierbei für „Quotient" und das „r" für „Rest". Die Darstellung $a = qd + r$ mit $0 \le r < d$ nennen wir auch die *Division* von a durch d mit *Rest* r.

Beweis
Wir nehmen zunächst an, daß $a \ge 0$ gilt. Sei dann $q \ge 0$ maximal mit $qd \le a$. Sei $r = a - qd$. Dann gilt $a = qd + r$, und nach Wahl von q ist $0 \le r < d$.
Sei nun $a < 0$. Nach dem bereits Bewiesenen gibt es q', r' mit

$$-a = q'd + r', \quad 0 \le r' < d.$$

Ist $r' = 0$, so ist $a = -q'd + 0$, und wir sind fertig. Andernfalls sei $q = -q' - 1$ und $r = -r' + d$. Dann ist

$$a = (-q')d - r' = qd + d + r - d = qd + r, \quad 0 \le r < d.$$

Damit ist die Existenzaussage bewiesen. Zum Beweis der Eindeutigkeit sei

$$a = q_1 d + r_1 = q_2 d + r_2 \text{ mit } 0 \le r_1 \le r_2 < d.$$

Dann gilt $r_2 - r_1 = (q_1 - q_2)d$, also $d \mid (r_2 - r_1)$. Dann ist aber $r_2 = r_1$, denn andernfalls wäre $r_2 - r_1 \ne 0$ und damit ist $d \le r_2 - r_1 < d$, was nicht sein kann. Wegen $r_1 = r_2$ ist dann aber $q_1 d = q_2 d$, und wegen $d \ne 0$ also $q_1 = q_2$. Damit ist auch die Eindeutigkeitsaussage des Satzes gezeigt. ∎

Sind $a = qd + r$ und $b = q'd + r'$ die Divisionen von a bzw. b durch d, so gilt

$$a - b = (q - q')d + r - r'.$$

Wegen $-d < r - r' < d$ gilt also $r = r'$ genau dann, wenn d ein Teiler von $a - b$ ist. Zwei Zahlen a und b haben also bei Division mit d genau dann den gleichen Rest, wenn ihre Differenz durch d teilbar ist. Wir definieren entsprechend:

Definition *(kongruent modulo d, $a \equiv b \bmod(d)$)*
Sei $d \geq 1$. Zwei Zahlen a und b heißen *kongruent modulo* d, in Zeichen $a \equiv b \bmod(d)$, falls $d \mid (a - b)$.

Diesem Kongruenzbegriff liegt folgende „geometrische" Vorstellung der Kongruenz zugrunde: Wir können den Punkt a mit dem Punkt b zur Deckung bringen, indem wir a um ganzzahlige Vielfache von d entlang der Zahlengeraden verschieben.

Wir schreiben auch $a \equiv b \equiv c \equiv \ldots \bmod(d)$, falls die Zahlen a, b, c paarweise kongruent modulo d zueinander sind. Es gilt zum Beispiel:

$$1 \equiv 5 \equiv 9 \equiv \ldots \bmod(4), \quad 1 \equiv -3 \equiv -7 \equiv \ldots \bmod(4).$$

Die Kongruenzrelation wird uns im weiteren immer wieder begegnen und viele Beispiele für algebraische Strukturen zur Verfügung stellen.

Größter gemeinsamer Teiler

Von fundamentaler Bedeutung für alles weitere ist folgende Begriffsbildung:

Definition *(größter gemeinsamer Teiler, ggT(a, b))*
Seien a, b nicht beide Null. Dann heißt die größte Zahl $d \geq 1$, die sowohl ein Teiler von a als auch ein Teiler von b ist, der *größte gemeinsame Teiler* von a und b, in Zeichen $d = \mathrm{ggT}(a, b)$. Wir setzen weiter $\mathrm{ggT}(0, 0) = 0$.

Sind a und b nicht beide gleich Null, so existiert ein $d \geq 1$ wie in der Definition in der Tat. Denn 1 ist ein gemeinsamer Teiler von a und b, und jeder weitere positive gemeinsame Teiler d von a und b ist kleinergleich $|a|$ oder kleinergleich $|b|$. Es gibt also nur endlich viele gemeinsame Teiler $d \geq 1$ von a und b, und einer von ihnen ist der größte. Wir werden unten ein Verfahren zur effektiven Bestimmung von ggT(a, b) kennenlernen.

Da jede Zahl die Null teilt, existiert kein wörtlicher größter gemeinsamer Teiler für die Wahl $a = b = 0$. Die Setzung von $\mathrm{ggT}(0, 0) = 0$ ist aber notationell bequem und sinnvoll. So gilt z. B. $\mathrm{ggT}(a, a) = |a|$ für alle a.

Es gilt zum Beispiel

$$\mathrm{ggT}(2, 5) = 1, \quad \mathrm{ggT}(10, 12) = 2, \quad \mathrm{ggT}(-6, 9) = 3, \quad \mathrm{ggT}(-4, -8) = 4.$$

Allgemein gilt:

Satz *(Eigenschaften des größten gemeinsamen Teilers)*
Für alle a, b, c, n, m gilt:

(G1) $ggT(a, a) = ggT(0, a) = |a|$,

(G2) $ggT(a, b) = ggT(b, a) = ggT(|a|, |b|)$,

(G3) $a \mid b$ *gdw* $ggT(a, b) = |a|$,

(G4) $ggT(a, b) \leq ggT(a, na + mb)$,

(G5) $ggT(a, b) = ggT(a, na + b)$.

Der Beweis dieser Eigenschaften kann wieder dem Leser überlassen bleiben. Wir bemerken aber noch, daß die Ungleichung in (G4) im allgemeinen keine Gleichung ist. Denn es gilt

$ggT(2, 3) = 1$, aber $ggT(2, 1 \cdot 2 + 2 \cdot 3) = ggT(2, 8) = 2$.

In Analogie zur Definition des größten gemeinsamen Teilers definieren wir:

Definition *(kleinstes gemeinsames Vielfaches, kgV(a, b))*
Seien $a, b \neq 0$. Dann heißt die kleinste Zahl $v \geq 1$, die sowohl ein Vielfaches von a als auch ein Vielfaches von b ist, das *kleinste (positive) gemeinsame Vielfache* von a und b, in Zeichen $v = kgV(a, b)$. Weiter setzen wir $kgV(0, a) = kgV(a, 0) = 0$ für alle a.

Es gilt $kgV(a, b) \leq |ab|$. Als Anwendung der Division mit Rest zeigen wir:

Satz *(Teilereigenschaft des kgV)*
$a \mid w$ und $b \mid w$ *impliziert* $kgV(a, b) \mid w$.

Beweis
Die Aussage ist klar für $w = 0$, da dann $c \mid w$ für alle c gilt. Sei also $w \neq 0$. Sei $v = kgV(a, b)$. Wegen $w \neq 0$ ist $v \geq 1$. Sei also $w = qv + r$ mit $0 \leq r < v$. Dann gilt $r = w - qv$, und damit $a \mid r$ und $b \mid r$ nach (T7). Also ist r ein gemeinsames Vielfaches von a und b. Wegen $r < v = kgV(a, b)$ ist dann aber $r = 0$. Folglich gilt $w = qv$ und damit $v \mid w$. ∎

Damit können wir nun weitere Eigenschaften der ggT-Funktion beweisen:

Satz *(weitere Eigenschaften des größten gemeinsamen Teilers)*
Für alle $a, b, c, \ldots$ gilt:

(G6) $e \mid a$ und $e \mid b$ *impliziert* $e \mid ggT(a, b)$,

(G7) $ggT(ca, cb) = |c| \, ggT(a, b)$,

(G8) $d \mid ab$ und $ggT(d, a) = 1$ *impliziert* $d \mid b$.

(G9) $a \mid c$, $b \mid c$ und $ggT(a, b) = 1$ *impliziert* $ab \mid c$.

(G10) $ggT(a, c) = 1$ und $ggT(b, c) = 1$ *impliziert* $ggT(ab, c) = 1$.

Beweis

zu (G6): Sei $d = ggT(a, b)$. Die Aussage ist klar für $d = 0$ oder $e = 0$. Seien also $d, e \neq 0$, und sei $v = kgV(e, d) \geq 1$. Dann ist a ein Vielfaches von e und von d, also gilt $v \mid a$. Ebenso gilt $v \mid b$. Damit ist also v ein gemeinsamer Teiler von a und b, also $d \geq v = kgV(e, d) \geq d$. Folglich ist $v = d$, und damit $e \mid d$.

zu (G7): Seien $d = ggT(a, b)$ und $d' = ggT(ca, cb)$. O.E. seien $a, b \neq 0$ und $c > 0$. Wegen $cd \mid ca$ und $cd \mid cb$ gilt $cd \leq d'$. Wegen $c \mid ca$ und $c \mid cb$ gilt $c \mid d'$ nach (G6). Also gibt es ein e mit $d' = ce$. Dann gilt $ce \mid ca$ und $ce \mid cb$, also $e \mid a$ und $e \mid b$. Dann ist aber $e \leq d$ und damit $d' = ce \leq cd$.

zu (G8): Es gilt $d \mid bd$ und $d \mid ba$. Nach (G6) gilt also $d \mid ggT(bd, ba)$. Aber $ggT(bd, ba) = |b| \, ggT(d, a) = |b|$. Also $d \mid b$.

zu (G9): Sei $v = kgV(a, b) \leq |ab|$, und sei $v = eb$. Dann gilt $a \mid eb$ und $ggT(a, b) = 1$, also $a \mid e$ nach (G8). Also ist $v \geq |ab|$, und damit $v = |ab|$. Da c ein Vielfaches von a und b ist, gilt $v \mid c$ und damit $ab \mid c$.

zu (G10): Sei $d \geq 1$ ein Teiler von ab und c. Wir zeigen, daß $d = 1$ gilt. Wegen $ggT(a, c) = 1$ und $d \mid c$ gilt $ggT(d, a) = 1$. Wegen $d \mid ab$ gilt dann aber $d \mid b$ nach (G8). Wegen $ggT(b, c) = 1$ und $d \mid c$ ist dann aber $d = 1$. –

Analog zu (G7) gilt

(#) $kgV(ca, cb) = |c| \, kgV(a, b)$ für alle a, b, c.

Denn $|c| \, kgV(a, b)$ ist ein Vielfaches von ca und cb, und deswegen gilt „$\leq$“. Zum Beweis von „$\geq$“ sei $kgV(ca, cb) = cd$. O.E. ist $cd \neq 0$. Dann gilt $ca \mid cd$ und $cb \mid cd$, also $a \mid d$ und $b \mid d$. Folglich gilt $kgV(a, b) \mid d$ und damit $c \cdot kgV(a, b) \mid kgV(ca, cb)$.

Damit können wir auch folgenden ansprechenden Zusammenhang beweisen:

Satz *(Produktsatz für ggT und kgV)*
Für alle a, b gilt: $|a \cdot b| = ggT(a, b) \cdot kgV(a, b)$.

Beweis

Sei $d = ggT(a, b)$. Die Aussage ist klar für $d = 0$. Ist $d = 1$, so ist $kgV(a, b) = |ab|$ wie im Beweis von (G9) oben. Sei also $d \geq 1$, und seien $a = da'$ und $b = db'$. Dann ist $ggT(a', b') = 1$ und damit gilt mit (#):

$$|ab| = d^2 \, |a'b'| = d \cdot d \cdot kgV(a', b') = ggT(a, b) \cdot kgV(da', db') =$$

$$ggT(a, b) \cdot kgV(a, b). \quad –$$

Unsere Untersuchungen der Teilbarkeit sind nun so weit gediehen, daß der Leser an dieser Stelle zu den Abschnitten über Primzahlen und der Eindeutigkeit der Primfaktorzerlegung springen kann, wenn er möchte. Die natürliche Fortsetzung scheint aber an dieser Stelle die Diskussion eines Verfahrens zu sein, das den größten gemeinsamen Teiler zweier Zahlen effektiv ermittelt.

Der Euklidische Algorithmus

Der Euklidische Algorithmus ist ein effektives Verfahren zur Bestimmung des größten gemeinsamen Teilers d* zweier Zahlen a und b. Er gehört zu den ältesten, schönsten und fruchtbarsten Algorithmen der Mathematik. Das Verfahren wird uns nicht nur den größten gemeinsamen Teiler d* zweier Zahlen a und b liefern, sondern auch Koeffizienten n und m derart, daß

$$d^* = n\,a + m\,b.$$

So ist zum Beispiel $3 = ggT(129, 33)$ und es gilt $3 = -1 \cdot 129 + 4 \cdot 33$.

Wegen $ggT(a, b) = ggT(|a|, |b|) = ggT(|b|, |a|)$ genügt es, das Verfahren für Paare a, b mit $a > b > 0$ anzugeben. Es ist bestimmt durch die sog. *Wechselwegnahme:* In jedem Schritt des Algorithmus liegen zwei Zahlen vor (zu Beginn sind dies a und b). Wir ziehen nun die kleinere der beiden Zahlen von der größeren ab und erhalten so ein neues Zahlenpaar. Dies wiederholen wir so lange, bis beide Zahlen gleich sind. Wir werden zeigen, daß der so errechnete gemeinsame Wert des letzten Zahlenpaares nichts anderes ist als der größte gemeinsame Teiler des Ausgangspaares.

Wir verwenden folgende Fassung, bei der mehrere Schritte der gerade geschilderten Wechselwegnahme zu einem zusammengefaßt werden.

Algorithmus von Euklid

Sei $a > b > 0$. Wir setzen $a_0 = a$, $a_1 = b$ und berechnen nun schrittweise $a_2, \ldots, a_i, \ldots$ durch Division mit Rest wie folgt:

$$a_0 = q_0 \cdot a_1 + a_2 \qquad \text{mit } 0 < a_2 < a_1,$$

$$a_1 = q_1 \cdot a_2 + a_3 \qquad \text{mit } 0 < a_3 < a_2,$$

$$\ldots$$

$$a_i = q_i \cdot a_{i+1} + a_{i+2} \qquad \text{mit } 0 < a_{i+2} < a_{i+1},$$

$$\ldots$$

Wir beenden das Verfahren mit a_{i^*+1} als Ergebnis, sobald wir einen Index i^* gefunden haben mit $a_{i^*} = q_{i^*} \cdot a_{i^*+1}$. Wir setzen dann noch $a_{i^*+2} = 0$.

Das Verfahren muß nach endlich vielen Schritten abbrechen und ein Ergebnis a_{i^*+1} liefern, da nach Konstruktion $a_0 > a_1 > \ldots > a_i > \ldots > 0$ gilt.

Der Algorithmus ist ein Beispiel für eine (nach endlich vielen Schritten abbrechende) rekursive Definition von Zahlen a_i. Im Rekursionsanfang setzen wir $a_0 = a$ und $a_1 = b$. Im Rekursionschritt wird a_{i+2} mit Rückgriff auf a_i und a_{i+1} definiert durch die eindeutige Darstellung

$$a_i = q_i \cdot a_{i+1} + a_{i+2},\ 0 \le a_{i+2} < a_{i+1}.$$

Sobald $a_{i+2} = 0$ ist, wird die Rekursion abgebrochen.

Der Algorithmus verläuft für das obige Paar $a = a_0 = 129$ und $b = a_1 = 33$ wie folgt:

$$\begin{aligned} 129 &= 3 \cdot 33 + 30, & q_0 &= 3, \quad a_2 = 30, \\ 33 &= 1 \cdot 30 + 3, & q_1 &= 1, \quad a_3 = 3, \\ 30 &= 10 \cdot 3, & q_2 &= 10. \end{aligned}$$

Das Ergebnis der Berechnung ist also 3. Zurückrechnen liefert nun:

$$3 = 33 - 1 \cdot 30 = 33 - 1 \cdot (129 - 3 \cdot 33) = -1 \cdot 129 + 4 \cdot 33.$$

Damit haben wir obige Darstellung von $3 = ggT(129, 33)$ gefunden.

Daß der Algorithmus in der Tat das leistet, was wir von ihm behaupten, wollen wir nun beweisen.

Satz *(Korrektheit des Euklidischen Algorithmus)*
Sei d^* das Ergebnis des Euklidischen Algorithmus für $a > b > 0$.
Dann gilt $d^* = ggT(a, b)$.

Beweis
Seien $a_0 > a_1 > a_2 > \ldots > a_{i^*+1} = d^*$ und $q_0, \ldots, q_{i^*}$ die Zahlen, die der Euklidische Algorithmus für $a = a_0$ und $b = a_1$ liefert. Sei $d = ggT(a, b)$.
Wir zeigen durch Induktion nach i:

(+) $d = ggT(a_i, a_{i+1})$ für alle i mit $0 \le i \le i^*$.

Induktionsanfang i = 0

Es gilt $ggT(a_0, a_1) = ggT(a, b) = d$.

Induktionsschritt von i nach i + 1

Es gelte also $d = ggT(a_i, a_{i+1})$ (Induktionsvoraussetzung).
Dann gilt

$$ggT(a_{i+1}, a_{i+2}) = ggT(a_{i+1}, a_i - q_i a_{i+1}) =_{(G5)} ggT(a_{i+1}, a_i) =_{(G2)} d.$$

Damit haben wir aber:

$$d =_{(+)\,\text{für}\, i = i^*} ggT(a_{i^*}, a_{i^*+1}) = ggT(q_{i^*} a_{i^*+1}, a_{i^*+1}) = ggT(q_{i^*} d^*, d^*) = d^*.$$

Der Leser sieht, daß die Korrektheit des Euklidischen Algorithmus letztendlich auf der einfachen Eigenschaft (G5) ruht. In der oben geschilderten Fassung der Wechselwegnahme wird dies besonders deutlich: Ist a', b' das aktuell vorliegende Paar mit $a' \neq b'$, so ist $a' - b'$, b' oder $a', b' - a'$ das nächste Paar, je nachdem, ob $b' < a'$ oder $a' < b'$ gilt. Da aber

$$ggT(a', b') = ggT(a' - b', b') = ggT(a', b' - a')$$

gilt, bewahrt die Wechselwegnahme in jedem Schritt den größten gemeinsamen Teiler des Ausgangspaares a, b.

Linearkombinationen

Wir beweisen nun auch noch das oben schon angekündigte Ergebnis über Linearkombinationen:

Satz *(Linearkombinationen und größter gemeinsamer Teiler)*
Sei d = ggT(a, b). Dann können wir mit Hilfe des Euklidischen Algorithmus Zahlen n und m finden derart, daß $d = n\,a + m\,b$.

Beweis
Die Behauptungen sind leicht einzusehen, falls a = b gilt oder falls eine der beiden Zahlen gleich 0 ist. Weiter genügt es, die Aussage für den Fall $a > b > 0$ zu beweisen. Hierzu seien wieder $a_0 > a_1 > a_2 > \ldots > a_{i^*+1} = d$ und $q_0, \ldots, q_{i^*}$ die Zahlen, die der Algorithmus für $a = a_0$ und $b = a_1$ liefert. Wir zeigen durch starke Induktion nach i:

(+) a_i ist eine Linearkombination von a und b für alle $0 \le i \le i^* + 1$.

Induktionsschritt i
Sei also $a_{i'} = n_{i'}\,a + m_{i'}\,b$ für alle $i' < i$ (Induktionsvoraussetzung). Wegen $a_0 = a$ und $a_1 = b$ gilt die Behauptung für i = 0 und i = 1. Sei also $i \ge 2$. Dann gilt

$$a_i = a_{i-2} - q_{i-2}\,a_{i-1} =$$
$$n_{i-2}\,a + m_{i-2}\,b - q_{i-2}\,(n_{i-1}\,a + m_{i-1}\,b) =$$
$$(n_{i-2} - q_{i-2}\,n_{i-1})\,a + (m_{i-2} - q_{i-2}\,m_{i-1})\,b.$$

∎ Nach (+) ist insbesondere $d = a_{i^*+1}$ eine Linearkombination von a und b.

Nach dem Beweis von (+) können wir für $a > b > 0$ also d = ggT(a, b) wie folgt effektiv als Linearkombination von a und b darstellen: Wir berechnen zuerst die Zahlen a_i und q_i für $0 < i \le i^*$ des Euklidischen Algorithmus für a und b. Nun definieren wir rekursiv:

$$n_0 = 1,\quad m_0 = 0,\quad n_1 = 0,\quad m_1 = 1,$$
$$n_i = n_{i-2} - q_{i-2}\,n_{i-1},\quad m_i = m_{i-2} - q_{i-2}\,m_{i-1} \qquad \text{für alle } i \text{ mit } 2 \le i \le i^* + 1.$$

Dann gilt $a_i = n_i\,a + m_i\,b$ für alle i mit $0 \le i \le i^* + 1$, und damit ist

$$d = n_{i^*+1}\,a + m_{i^*+1}\,b.$$

Die Darstellung des größten gemeinsamen Teilers ist keineswegs eindeutig. So ist etwa

$$1 = 1 \cdot 3 - 1 \cdot 2 = 3 \cdot 3 - 4 \cdot 2.$$

Wir halten weiter fest:

Korollar *(Identifizierung der Linearkombinationen von a und b)*
Die Linearkombinationen von a und b sind genau die Zahlen der Form kd, wobei $d = ggT(a, b)$.
Insbesondere gilt: Sind a, b nicht beide gleich Null, so ist d die kleinste positive Linearkombination von a und b.

Beweis
Sei $d = n\,a + m\,b$. Dann ist $kd = (kn)\,a + (km)\,b$. Also sind alle Zahlen der Form kd Linearkombinationen von a und b.
Ist umgekehrt $c = n'\,a + m'\,b$ eine Linearkombination von a und b, so ist d nach (T7) ein Teiler von c. Also existiert ein k mit $c = kd$.

Unsere Ergebnisse über Linearkombinationen lassen sich auch für einfache Beweise der Eigenschaften der ggT-Funktion nutzen. Wir betrachten noch einmal die wichtigen Eigenschaften:

(G6) $e \mid a$ und $e \mid b$ *impliziert* $e \mid ggT(a, b)$,

(G7) $ggT(ca, cb) = |c|\, ggT(a, b)$,

die wir bei unserer Analyse des Euklidischen Algorithmus und der Linearkombinationen nicht verwendet haben.

Zum Beweis von (G6) sei $d = ggT(a, b)$, und es sei

$$d = n\,a + m\,b.$$

Wegen $e \mid a$ und $e \mid b$ gilt dann $e \mid d$ auch nach (T7).

Zum Beweis von (G7) sei wieder $d = ggT(a, b)$ und $d' = ggT(ca, cb)$, und ohne Einschränkung sei $c \geq 0$. Wegen $cd \mid ca$ und $cd \mid cb$ gilt $cd \leq d'$. Zum Beweis der anderen Ungleichung sei $d = na + mb$. Dann ist $dc = nac + mbc$ eine Linearkombination von ac und bc, und wegen $dc \geq 0$ also $dc \geq d'$ nach dem Korollar.

Eine direktere Möglichkeit, die Eigenschaft (G6) mit Hilfe des Euklidischen Algorithmus zu beweisen, ist die folgende. Es seien wieder $a_0, \ldots, a_{i^*+1}$ die Zahlen, die der Euklidische Algorithmus für $a > b > 0$ erzeugt. Dann gilt:

(+) $e \mid a_i$ für alle i mit $0 \leq i \leq i^* + 1$.

Damit gilt $e \mid a_{i^*+1}$ (und $a_{i^*+1} = d$).
Wir können aber (+) leicht durch starke Induktion nach i zeigen:

Induktionsschritt i
Sei e ein Teiler von $a_{i'}$ für alle $0 \leq i' < i$ (Induktionsvoraussetzung).
Wegen $e \mid a$, $e \mid b$, $a = a_0$ und $b = a_1$ gilt die Behauptung für $i = 0$ und $i = 1$.
Ist $i \geq 2$, so gilt $a_i = a_{i-2} - q_{i-2}\, a_{i-1}$, und damit gilt $e \mid a_i$ nach Induktionsvoraussetzung und (T7).

Damit können wir nun gut gerüstet den wohl faszinierendsten Objekten der Zahlentheorie begegnen, nämlich den Primzahlen.

Primzahlen

Definition *(Primzahl, zusammengesetzte Zahl)*
Ein $p \geq 2$ heißt eine *Primzahl* oder eine *unzerlegbare Zahl*, wenn es keinen Teiler d von p gibt mit $1 < d < p$. Andernfalls heißt p eine *zusammengesetzte* oder *zerlegbare* Zahl.

Daß wir der 1 keinen Status als Primzahl einräumen liegt unter anderem an der angestrebten Eindeutigkeit der Primfaktorzerlegung jeder Zahl $n \geq 2$, die wir unten zeigen werden. So ist zum Beispiel $20 = 2^2 \cdot 5$. Wäre die 1 eine Primzahl, so hätten wir verschiedene Darstellungen, nämlich $20 = 1 \cdot 2^2 \cdot 5 = 1^2 \cdot 2^2 \cdot 5 = \ldots$.

Die Reihe der Primzahlen kann man mit einem *Siebverfahren* ermitteln: Wir schreiben die natürlichen Zahlen $n \geq 2$ der Reihe nach auf:

2, 3, 4, 5, 6, 7, 8, …,

und streichen, die 2 stehenlassend, alle Vielfachen der 2. Anschließend betrachten wir die erste Zahl 3, die diesen Prozeß überlebt hat und streichen alle Vielfachen der 3, wobei die 3 selbst wieder unangetastet bleibt. So fortfahrend erhalten wir eine Liste der Primzahlen:

2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, …,

2, 3, , 5, , 7, , 9, , 11, , 13, , 15, , 17, , 19, , 21, , 23, , 25, …,

2, **3**, , 5, , 7, , , , 11, , 13, , , , 17, , 19, , , , 23, , 25, …,

2, 3, , **5**, , 7, , , , 11, , 13, , , , 17, , 19, , , , 23, , , …,

usw. *(Sieb des Eratosthenes)*

Eine Durchführung dieser Methode per Hand oder mit Hilfe eines Computers ergibt folgende Liste der 168 Primzahlen, die kleiner als 1000 sind:

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71,
73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 127, 131, 137, 139, 149, 151, 157, 163, 167, 173,
179, 181, 191, 193, 197, 199, 211, 223, 227, 229, 233, 239, 241, 251, 257, 263, 269, 271, 277, 281,
283, 293, 307, 311, 313, 317, 331, 337, 347, 349, 353, 359, 367, 373, 379, 383, 389, 397, 401, 409,
419, 421, 431, 433, 439, 443, 449, 457, 461, 463, 467, 479, 487, 491, 499, 503, 509, 521, 523, 541,
547, 557, 563, 569, 571, 577, 587, 593, 599, 601, 607, 613, 617, 619, 631, 641, 643, 647, 653, 659,
661, 673, 677, 683, 691, 701, 709, 719, 727, 733, 739, 743, 751, 757, 761, 769, 773, 787, 797, 809,
811, 821, 823, 827, 829, 839, 853, 857, 859, 863, 877, 881, 883, 887, 907, 911, 919, 929, 937, 941,
947, 953, 967, 971, 977, 983, 991, 997.

Im Intervall von $n = 10^7 = 10000000$ bis $n + 1000$ befinden sich dagegen nur noch 61 Primzahlen, nämlich

n + 19, n + 79, n + 103, n + 121, n + 139, n + 141, n + 169, n + 189, n + 223, n + 229, n + 247, n + 253, n + 261, n + 271, n + 303, n + 339, n + 349, n + 357, n + 363, n + 379, n + 439, n + 451, n + 453, n + 457, n + 481, n + 511, n + 537, n + 583, n + 591, n + 609, n + 643, n + 651, n + 657, n + 667, n + 687, n + 691, n + 721, n + 723, n + 733, n + 741, n + 747, n + 759, n + 763, n + 769, n + 789, n + 799, n + 813, n + 819, n + 831, n + 849, n + 867, n + 871, n + 873, n + 877, n + 891, n + 931, n + 943, n + 961, n + 967, n + 987, n + 993

Viele Fragen aus der zauberhaften Welt der Primzahlen konnten bereits von den alten Griechen mit bestechenden Argumenten beantwortet werden, andere sind bis heute offen geblieben. Alles scheint greifbar zu sein, und doch haben wir stets den Eindruck, daß uns das rechte Verständnis der Dinge noch abgeht. Diese Liebesheirat von Einfachheit und Tiefsinn fasziniert Menschen mit mathematischen Neigungen seit Jahrtausenden und spricht sowohl den Forschergeist an als auch das Streben nach Schönheit und Harmonie.

Wir wollen einige Fragen etwas näher betrachten. Zunächst führt der doch recht radikale Ausdünnungsprozeß der Siebmethode zu der Frage:

1) Gibt es beliebig große Primzahlen?

Die Antwort ist „ja", und ein entsprechender Beweis findet sich schon bei Euklid. Ebenso ist aber, wie wir zeigen werden, auch die folgende umgekehrte Frage zu bejahen:

2) Gibt es für jedes $k \geq 1$ ein $n \geq 1$ derart, daß alle Zahlen n, $n + 1$, $n + 2$, …, $n + k$ zusammengesetzt sind?

Die gepaart auftretenden Primzahlen 3 und 5, 5 und 7, 11 und 13, 17 und 19, …, 857 und 859, 881 und 883, …, führen zu der Frage:

3) Gibt es beliebig große Primzahlzwillinge, d. h. Primzahlen p derart, daß auch p + 2 eine Primzahl ist?

Dieses Problem ist bis zum heutigen Tage offen.

Eine geistreiche Beobachtung ist die folgende:

$4 = 2 + 2$, $6 = 3 + 3$, $8 = 3 + 5$, $10 = 3 + 7 = 5 + 5$, $12 = 5 + 7$,

$14 = 3 + 11 = 7 + 7$, …

Wir fragen also allgemein:

4) Ist jede gerade Zahl $n > 2$ die Summe zweier Primzahlen?

Diese sog. *Goldbachsche Vermutung* ist ebenfalls noch unbewiesen. Sie wurde mit Hilfe von Computern für sehr viele Zahlen bestätigt.

Schließlich blicken wir auf den Ausdünnungsprozeß als Ganzes und fragen:

5) Was lässt sich über die Verteilung der Primzahlen sagen? Wieviele Primzahlen kleinergleich n gibt es bei vorgegebenem n ungefähr?

Die Verteilung der Primzahlen beschreibt ein bemerkenswertes Resultat, das Gauß 1793 anhand des Studiums von Primzahltabellen vermutet hat, das aber erst 1896 von Hadamard und Vallée Poussin bewiesen werden konnte. Sei hierzu $\pi(n)$ die Anzahl der Primzahlen, die kleinergleich n sind. Dann gilt, mit dem Logarithmus ln zur Basis e:

$$\pi(n) \sim n/\ln(n) \qquad \textit{(Primzahlsatz)}$$

Das Symbol $\sim$ wird hier als „asymptotisch gleich" gelesen. Die genaue Bedeutung eines Ausdrucks „$A(n) \sim B(n)$" ist, daß der Wert von $A(n)/B(n)$ mit wachsendem n gegen 1 strebt. Nach dem Primzahlsatz würden wir also abschätzen, daß es in etwa $\pi'(n) = n/\ln(n)$ viele Primzahlen kleinergleich n gibt. Eine Tabelle zeigt, wie gut die Abschätzung für die ersten Zehnerpotenzen ist:

n	$\pi(n)$	$n/\ln(n)$	$\pi(n)\ln(n)/n$
10	4	4,34	0,9210
10^2	25	21,71	1,1513
10^3	168	144,76	1,1605
10^4	1229	1085,74	1,1320
10^5	9592	8685,89	1,1043
10^6	78498	72382,41	1,0845
10^7	664579	620420,69	1,0712
10^8	5761455	5428681,02	1,0613

Nach dem Primzahlsatz konvergieren die Werte in der rechten Spalte gegen 1.

Der Primzahlsatz ist ein Beispiel dafür, daß zur Untersuchung der natürlichen Zahlen weitergehende Zahlbereiche und zugehörige Funktionen eingesetzt werden können, wie hier die reelle Logarithmusfunktion. Erstaunlicherweise sind es sogar die komplexen Zahlen, die besonders gut geeignet sind, Licht auf die natürlichen Zahlen zu werfen.

Wir wenden uns nun der ersten oben aufgeworfenen Frage der Unendlichkeit der Primzahlen zu. Es ist keineswegs klar, daß nicht nach endlich vielen Schritten alle natürlichen Zahlen, die größer als ein gewisses p sind, durch das Sieb des Eratosthenes fallen. Dann gäbe es eine größte Primzahl. Daß dies nicht der Fall ist, besagt der folgende klassische Satz.

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

Beweis
Sei $m \geq 1$, und seien $p_1, \ldots, p_m$ beliebige Primzahlen. Wir setzen:

$$n = p_1 \cdot \ldots \cdot p_m + 1.$$

Für alle $1 \leq i \leq m$ gilt $p_i \mid n - 1$, also ist p_i kein Teiler von n. Sei nun p^* prim mit $p^* \mid n$ (etwa p^* = „der kleinste Teiler von n größergleich 2"). Dann gilt $p^* \neq p_i$ für alle $1 \leq i \leq n$. Also gibt es keine endliche Liste aller Primzahlen.

Die Zahl n des Beweises selbst ist nicht notwendig eine Primzahl. Ist die 2 nicht unter den Zahlen p_i, so ist das Produkt aller p_i ungerade und damit n selbst gerade, also durch 2 teilbar. Ein anderes Beispiel ist

$$n = 2 \cdot 5 \cdot 11 + 1 = 111 = 3 \cdot 37.$$

Interessanter sind hier die Produkte der ersten n Primzahlen:

Definition *(Euklidische Zahlen)*
Sei $n \geq 1$, und seien $p_1 < p_2 < \ldots < p_n$ die ersten n Primzahlen. Wir setzen

$$en(n) = p_1 \cdot \ldots \cdot p_n + 1,$$

und nennen en(n) die n-te *Euklidische Zahl.*

Hier steht „en" für *Euclidean number*. Es gilt also

$$en(1) = 2 + 1 = 3, \quad en(2) = 2 \cdot 3 + 1 = 7, \quad en(3) = 2 \cdot 3 \cdot 5 + 1 = 31,$$
$$en(4) = 2 \cdot 3 \cdot 5 \cdot 7 + 1 = 211, \quad en(5) = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 + 1 = 2311.$$

Die ersten 5 Euklidischen Zahlen sind also Primzahlen. Doch die sich an dieser Stelle aufdrängende Vermutung, daß alle Euklidischen Zahlen prim sind, ist falsch. Computerberechnungen haben ergeben, daß en(11) und en(75) Primzahlen sind, während alle Euklidischen Zahlen en(k) mit $6 \leq k \leq 74$, $k \neq 11$, zusammengesetzt sind. Speziell sieht en(6) = 30031 auf den ersten Blick vielleicht wie eine Primzahl aus, es gilt aber $30031 = 59 \cdot 509$.

Es ist zudem ein offenes Problem, ob unendlich viele Euklidische Zahlen prim sind, und der Leser sieht, daß selbst ein Jahrtausende altes und äußerst kurzes Argument Fragen aufwerfen kann, die die Mathematiker mit all ihren ausgereiften Methoden bislang nicht beantworten konnten. Gerade aus der Sicht des Anfängers sieht der Primzahlsatz wie eine uneinnehmbare Burg aus, während die Frage nach den Euklidischen Zahlen einfach wirkt und nach einer einfachen Lösung ruft. Ein geistreiches kurzes Argument würde nicht überraschen. Es scheint aber kein solches Argument zu geben, und wir gelangen zu der Erkenntnis, daß der Schwierigkeitsgrad mathematischer Fragen schwer einzuschätzen und zu messen ist.

Diese Betrachtungen zeigen auch, daß Vermutungen, die durch eine Reihe von Beispielen belegt werden, keinen Beweis ersetzen können. Es ist „Zufall", daß die ersten fünf Euklidischen Zahlen prim sind. Ebenso könnte es „Zufall" sein, daß sich alle bislang untersuchten geraden Zahlen größer als 2 als Summe zweier Primzahlen schreiben ließen. Die zugehörige Goldbachsche Vermutung bleibt offen. Ob wahr oder falsch: In jedem Falle können experimentelle Überlegungen und Berechnungen zu interessanten mathematischen Vermutungen und neuen Einsichten führen und abgesehen von ihrem spielerischen Wert scheint dies gerade ihre Funktion in der Mathematik zu sein.

Eindeutigkeit der Primfaktorzerlegung

Eine *Primfaktorzerlegung* einer Zahl $n \geq 2$ ist eine Darstellung von n als Produkt von Primzahlen. Die Primzahlen erscheinen bei dieser Betrachtung als die „multiplikativen Bausteine" der natürlichen Zahlen. So ist zum Beispiel:

$$4 = 2^2,\quad 6 = 2 \cdot 3,\quad 9 = 3^2,\quad 15 = 3 \cdot 5,\quad 16 = 2^4,\quad 18 = 2 \cdot 3^2,\quad \ldots$$

Um eine Primfaktorzerlegung von n zu finden, argumentieren wir induktiv. Für $n = 2$ ist die Aussage klar. Für ein $n > 2$ überprüfen wir für alle d mit $2 \leq d < n$ der Reihe nach, ob $d \mid n$ gilt oder nicht. Ist kein d ein Teiler von n, so ist n eine Primzahl (und eine Primfaktorzerlegung von sich selbst). Finden wir dagegen ein erstes d mit $d \mid n$, so ist d eine Primzahl und $d \cdot Z$ eine Primfaktorzerlegung von n, wobei Z eine nach Induktionsvoraussetzung existierende Primfaktorzerlegung von $n/d < n$ ist.

Stärker gilt: Ist p ein Primteiler von n, so existiert eine Primfaktorzerlegung von n, in der p vorkommt (nämlich $p \cdot Z$ mit Z Primfaktorzerlegung von n/p).

Wie eine Faktorisierung von n effektiv durchgeführt werden kann, ist eine interessante Frage, die aber auf einem anderen Blatt steht.

Es ist keineswegs selbstverständlich, daß eine Primfaktorzerlegung bis auf die Reihenfolge der Faktoren eindeutig ist. Der klassische Beweis ruht auf der folgenden Beobachtung:

Satz *(Teilbarkeitssatz von Euklid)*
Sei p prim, und seien a, b derart, daß $p \mid ab$. Dann gilt $p \mid a$ oder $p \mid b$.

Beweis
Gilt $p \mid a$, so ist nichts weiter zu zeigen. Sei also p kein Teiler von a.
Dann ist $ggT(p, a) = 1$. Nach (G8) gilt also $p \mid b$. ∎

Wir geben für diesen wichtigen Satz noch einen etwas direkteren Beweis mit Hilfe von Linearkombinationen.

Direkter Beweis des Teilbarkeitssatzes
Ist p kein Teiler von a, so gilt $ggT(p, a) = 1$. Also existieren n, m mit

$$np + ma = 1, \quad \text{und damit gilt} \quad npb + mab = b.$$

Wegen $p \mid ab$ gibt es ein k mit $ab = kp$. Also gilt $p(nb + mk) = b$,
und folglich ist p ein Teiler von b. ∎

Problemlos ergibt sich folgende Verallgemeinerung:

Satz *(allgemeiner Teilbarkeitssatz)*
Sei p eine Primzahl, und seien $a_1, \ldots, a_n$ mit $n \geq 1$ derart, daß $p \mid a_1 \cdot \ldots \cdot a_n$.
Dann gibt es ein i mit $p \mid a_i$.

Beweis

Wir zeigen die Aussage durch Induktion nach n. Für $n = 1$ ist nichts zu zeigen, und für $n = 2$ haben wir die Aussage bereits bewiesen. Gilt nun $p \mid (a_1 \dots a_{n+1})$ mit $n \geq 2$, so setzen wir $a = a_1 \dots a_n$. Dann gilt $p \mid a\, a_{n+1}$ und damit $p \mid a$ oder $p \mid a_{n+1}$. Im zweiten Fall sind wir fertig, und im ersten Fall gibt es wegen $a = a_1 \dots a_n$ nach Induktionsvoraussetzung ein i mit $p \mid a_i$. ∎

Mit anderen Worten: Teilt eine Primzahl ein Produkt, so teilt sie mindestens einen Faktor des Produkts. Daraus folgt sofort:

Korollar

Ist p eine Primzahl und gilt $p \mid n$, so kommt p in jeder Primfaktorzerlegung von n vor.

Damit können wir nun den „Fundamentalsatz der Zahlentheorie" zeigen:

Satz *(Eindeutigkeit der Primfaktorzerlegung)*

Jedes $n \geq 2$ läßt sich eindeutig schreiben in der Form

$$n = p_1^{e_1} \cdot \dots \cdot p_k^{e_k},$$

mit Primzahlen $p_1 < \dots < p_k$, $k \geq 1$, und Exponenten $e_i \geq 1$.

Beweis

Wir zeigen die Behauptung durch starke Induktion nach $n \geq 2$.

Induktionsschritt n

Sei p der kleinste Primteiler von n. Ist $n = p$, so ist die Aussage trivial. Sei also $n > p$, und seien Z_1 und Z_2 Primfaktorzerlegungen von n wie im Satz. Nach dem Korollar kommt p in beiden Zerlegungen vor. Nach Induktionsvoraussetzung stimmen dann aber die Zerlegungen Z_1' und Z_2' von n/p überein, die aus Z_1 und Z_2 durch Verminderung des p-Exponenten um 1 hervorgehen. Dann stimmen aber offenbar auch Z_1 und Z_2 überein. ∎

Definition *(kanonische Primfaktorzerlegung)*

Sei $n \geq 2$. Dann heißt die Darstellung $n = p_1^{e_1} \cdot \dots \cdot p_k^{e_k}$ mit $p_1 < \dots < p_k$ und Exponenten $e_i \geq 1$ die *kanonische Primfaktorzerlegung* von n.

Anhand der kanonischen Primfaktorzerlegung von n lassen sich alle Teiler von n leicht ablesen. Ebenso können wir den größten gemeinsamen Teiler und das kleinste gemeinsame Vielfache von n und m sofort ermitteln, wenn die Primfaktorzerlegungen von n und m bekannt sind. Im allgemeinen ist aber die Berechnung der Primfaktorzerlegung einer Zahl sehr aufwendig, und der Euklidische Algorithmus bleibt die erste Wahl zur Berechnung des größten gemeinsamen Teilers. Die Gleichung $nm = ggT(n, m) \cdot kgV(n, m)$ liefert dann auch das kleinste gemeinsame Vielfache.

Wir geben noch einen weiteren Beweis der Eindeutigkeit der Primfaktorzerlegung, der den Teilbarkeitssatz nicht heranzieht. Dieser Beweis ist erst Anfang des 20. Jahrhunderts von Ernst Zermelo gefunden worden. Er ist ein Paradebeispiel für eine geistreiche starke Induktion.

Zweiter Beweis der Eindeutigkeit der Primfaktorzerlegung

Wir zeigen die Behauptung durch starke Induktion nach $n \geq 2$.

Induktionsschritt n

Annahme, es gibt zwei verschiedene kanonische Primfaktorzerlegungen

$$p_1^{e_1} \cdot \ldots \cdot p_k^{e_k} = n = q_1^{d_1} \cdot \ldots \cdot q_{k'}^{d_{k'}}.$$

Dann ist jedes p_i verschieden von jedem q_j, da wir sonst zwei unterschiedliche kanonische Primfaktorzerlegungen von $n/p_i = n/q_j < n$ gewinnen könnten.
Ohne Einschränkung sei $p_1 < q_1$. Dann gilt $p_1 q_1 < q_1^2 \leq n$ (denn es gilt $d_1 \geq 2$ oder $k' \geq 2$, da n keine Primzahl sein kann). Also ist

$$m = n - p_1 q_1 > 0.$$

Wegen $p_1 \mid n$ und $q_1 \mid n$ sind p_1 und q_1 Teiler von m. Dann kommen aber sowohl p_1 als auch q_1 in der nach Induktionsvoraussetzung eindeutigen kanonischen Primfaktorzerlegung von $m < n$ vor. (Wegen $p_1 \mid m$ und $q_1 \mid m$ gibt es Primfaktorzerlegungen von m, in denen p_1 bzw. q_1 auftauchen; aufgrund der Eindeutigkeit folgt die Behauptung.)
Also gilt $p_1 q_1 \mid m$, und damit $p_1 q_1 \mid n$. Also ist q_1 ein Teiler von n/p_1, und damit existiert eine Primfaktorzerlegung von n/p_1, in der q_1 vorkommt. Aber es gilt

$$n/p_1 = p_1^{e_1 - 1} \cdot \ldots \cdot p_k^{e_k}.$$

Wegen $n/p_1 < n$ ist nach Induktionsvoraussetzung also q_1 eine der Primzahlen p_i, *Widerspruch*. ∎

Irrationale Verhältnisse

Als eine Anwendung der Eindeutigkeit der Primfaktorzerlegung zeigen wir die Existenz irrationaler Zahlen: Es gibt eine reelle Zahl x, die sich nicht als Bruch a/b mit ganzen Zahlen a, b, $b \neq 0$, schreiben lässt. Konkret zeigen wir, daß $x = \sqrt{2}$, die positive Quadratwurzel aus 2, eine irrationale Zahl ist. Diese Aussage können wir formulieren, ohne den Bereich der ganzen Zahlen zu verlassen. Denn die Irrationalität dieser Wurzel können wir ausdrücken als:

Für alle natürlichen Zahlen $a, b \geq 1$ ist $(a/b)^2 \neq 2$.

Gleichwertig hierzu ist wiederum die Aussage des folgenden Satzes:

Satz *(Irrationalität von $\sqrt{2}$)*

Es gibt keine ganzen Zahlen $a, b \geq 1$ mit $a^2 = 2 \cdot b^2$.

Beweis

Annahme doch. Seien $a = 2^{e_1} \cdot u_1$ und $b = 2^{e_2} \cdot u_2$, mit ungeraden Zahlen u_1, u_2 und $e_1, e_2 \geq 0$. Wegen $a^2 = 2 \cdot b^2$ gilt dann:

$$(+) \quad a^2 = 2^{2e_1} \cdot u_1^2 = 2^{2 \cdot e_2 + 1} \cdot u_2^2.$$

Dann ist $2e_1$ gerade, aber $2e_2 + 1$ ungerade. Also ist $2e_1 \neq 2 \cdot e_2 + 1$. Aber die Zahlen u_1^2 und u_2^2 sind ungerade, und damit liefert (+) zwei verschiedene kanonische Primfaktorzerlegungen von a^2, nämlich

$$a^2 = 2^{2e_1} \cdot Z_1 = 2^{2 \cdot e_2 + 1} \cdot Z_2,$$

mit kanonischen Primfaktorzerlegungen Z_1 von u_1^2 und Z_2 von u_2^2, in denen der Faktor 2 nicht vorkommt. *Widerspruch.*

Kurz: $a^2 = 2 \cdot b^2$ ist für ganze Zahlen unmöglich, denn der 2-Exponent der eindeutigen Primfaktorzerlegung von a^2 ist gerade (möglicherweise 0), der von $2 \cdot b^2$ aber ungerade. Oder noch einmal anders formuliert: a^2 können wir geradzahlig oft halbieren, $2b^2$ ungeradzahlig oft.

Die Irrationalität der Quadratwurzel aus 2 taucht häufig ganz vorne in Listen der wichtigsten oder schönsten mathematischen Resultate aller Zeiten auf. Jeder Mathematiker sollte davon wissen, und fast möchte man sagen, daß jeder davon wissen sollte, denn nicht zuletzt ist das Ergebnis auch kulturgeschichtlich von großer Bedeutung. Es widerlegte die Pythagoreische Doktrin „Alles ist Zahl" im Sinne von „Jede Größe ist ein Verhältnis positiver natürlicher Zahlen". Modern formuliert lautet die Doktrin, daß die rationalen Zahlen genügen, um ein mathematisches Kontinuum zu bilden. Die Irrationalität der Quadratwurzel aus 2 zeigt, daß ein mathematisches Kontinuum mehr Punkte umfassen muß als die rationalen Zahlen.

Übungen

Übung 1 *(Teilbarkeit, I)*

Zeigen Sie, daß für alle a, b, c, d, …, n, m, … gilt:

(T1) $a \mid a$, $1 \mid a$, $a \mid 0$, $0 \mid a$ *gdw* $a = 0$, $d \mid a$ *gdw* $|d| \mid |a|$,

(T2) $d \mid a$ und $a \mid d$ *gdw* $|d| = |a|$,

(T3) $d \mid a$ und $a \neq 0$ *impliziert* $|d| \leq |a|$,

(T4) $e \mid d$ und $d \mid a$ *impliziert* $e \mid a$,

(T5) $d \mid a$ *impliziert* $d \mid (ab)$ und $(nd) \mid (na)$,

(T6) $(ca) \mid (cb)$ und $c \neq 0$ *impliziert* $a \mid b$,

(T7) $d \mid a$ und $d \mid b$ *impliziert* $d \mid (na + mb)$.

Übung 2 (Teilbarkeit, II)
Zeigen Sie, daß für alle n gilt:

(i) $6 \mid (n\,(n+1)\,(n+2))$,

(ii) $2 \mid (n^2 - n)$, $6 \mid (n^3 - n)$, $30 \mid (n^5 - n)$,

(iii) $8 \mid (n^2 - 1)$, falls n ungerade.

Übung 3 (Teilbarkeit, III)
Zeigen Sie, daß die Kongruenz modulo m eine Äquivalenzrelation auf $\mathbb{Z}$ ist. Zeigen Sie weiter, daß aus $a \equiv b \bmod(m)$ und $c \equiv d \bmod(m)$ stets folgt:

(i) $a + c \equiv b + d \bmod(m)$,

(ii) $ka \equiv kb \bmod(m)$ für alle k,

(iii) $a \cdot c \equiv b \cdot d \bmod(m)$,

(iv) $a^n \equiv b^n \bmod(m)$ für alle $n \geq 0$.

(v) $a \equiv b \bmod(d)$ für alle Teiler d von m mit $d \geq 1$.

Insbesondere ist die Kongruenz modulo m also eine Kongruenzrelation für die Addition und die Multiplikation auf $\mathbb{Z}$.

Übung 4 (Teilbarkeit, IV)
Bestimmen Sie für $m = 13$ und $m = 15$ alle Paare a, b mit $0 \leq a, b < m$ mit $ab \equiv 1 \bmod(m)$.

Sehen Sie eine Besonderheit für $m = 13$? Erstellen Sie weitere Tabellen, um Ihre Hypothesen zu überprüfen.

Übung 5 (Größter gemeinsamer Teiler, I)
Zeigen Sie, daß für alle a, b, c, n, m gilt:

(G1) $ggT(a, a) = ggT(0, a) = |a|$,

(G2) $ggT(a, b) = ggT(b, a) = ggT(|a|, |b|)$,

(G3) $a \mid b$ *gdw* $ggT(a, b) = |a|$,

(G4) $ggT(a, b) \leq ggT(a, na + mb)$,

(G5) $ggT(a, b) = ggT(a, na + b)$.

Übung 6 (Größter gemeinsamer Teiler, II)
Sei $v = kgV(m, n)$. Zeigen Sie, daß für alle a, b gilt:

$a \equiv b \bmod(m)$ und $a \equiv b \bmod(n)$ *gdw* $a \equiv b \bmod(v)$.

Übung 7 *(Größter gemeinsamer Teiler, III)*
Sei $n \geq 1$, und seien $a_1, \ldots, a_n$ ganze Zahlen, die nicht alle gleich 0 sind. Dann ist der *größte gemeinsame Teiler* von $a_1, \ldots, a_n$, in Zeichen $ggT(a_1, \ldots, a_n)$, definiert als das größte $m \geq 1$ mit $m \mid a_i$ für alle $1 \leq i \leq n$. Wir setzen zudem $ggT(0, \ldots, 0) = 0$. Zeigen Sie, daß für alle $n \geq 2$ und alle $a_1, \ldots, a_{n+1}$ gilt:

$$ggT(a_1, \ldots, a_{n+1}) = ggT(ggT(a_1, \ldots, a_n), a_{n+1}).$$

Übung 8 *(Größter gemeinsamer Teiler, IV)*
Sei $A \subseteq \mathbb{N}$ unendlich, und sei $d \geq 1$ die größte Zahl, die alle $a \in A$ teilt. Zeigen Sie: Es gibt $a_1, \ldots, a_n \in A$ mit $d = ggT(a_1, \ldots, a_n)$.

Übung 9 *(Der Euklidische Algorithmus, I)*
Führen Sie den Euklidischen Algorithmus für die Zahlen 84 und 132 durch, und stellen Sie $ggT(84, 132)$ in der Form $a \cdot 84 + b \cdot 132$ dar.

Übung 10 *(Der Euklidische Algorithmus, II)*
Zeigen Sie, daß sich die Reste, die der Euklidische Algorithmus erzeugt, nach jedem zweiten Schritt mindestens halbiert haben.

Übung 11 *(Linearkombinationen, I)*
Zeigen Sie, daß für alle $n \geq 2$ und alle $a_1, \ldots, a_n$ gilt:

$$\{ m_1 a_1 + \ldots + m_n a_n \mid m_i \in \mathbb{Z} \} = \{ k \cdot ggT(a_1, \ldots, a_n) \mid k \in \mathbb{Z} \}.$$

Übung 12 *(Linearkombinationen, II)*
(a) Zeigen Sie, daß für alle a, b gilt:

$$|a \cdot b| = ggT(a, b) \cdot kgV(a, b).$$

(b) Sei $d = ggT(k, m)$. Zeigen Sie, daß für alle a, b gilt:

$$ka \equiv kb \mod(m) \quad gdw \quad a \equiv b \mod(m/d).$$

Ist also $ggT(k, m) = 1$, so gilt:

$$ka \equiv kb \mod(m) \quad gdw \quad a \equiv b \mod(m).$$

[Zeigen Sie (a) mit Hilfe von (G8) zuerst für den Fall $ggT(a, b) = 1$. Im allgemeinen Fall ist dann weiter Übung 6 nützlich.]

Übung 13 *(Linearkombinationen, III)*
Ein nichtleeres $I \subseteq \mathbb{Z}$ heißt ein *Ideal*, falls I abgeschlossen unter Linearkombinationen ist, d. h. für alle $a, b \in I$ und alle $n, m \in \mathbb{Z}$ ist $na + mb \in I$. Zeigen Sie, daß für jedes Ideal I ein eindeutiges $d \in \mathbb{N}$ existiert mit

$$I = \{ kd \mid k \in \mathbb{Z} \}.$$

Übung 14 *(Linearkombinationen, IV)*
Sei $A \subseteq \mathbb{N}$ nichtleer und abgeschlossen unter Addition, d. h. für alle $a, b \in A$ ist $a + b \in A$. Weiter sei 1 der größte gemeinsame Teiler aller Zahlen in A. Zeigen Sie: Es gibt ein $k_0 \in \mathbb{N}$ derart, daß $k \in A$ für alle $k \geq k_0$ gilt.

[Seien $a_1, \ldots, a_n \in A$ mit $ggT(a_1, \ldots, a_k) = 1$. Seien $n_i \in \mathbb{Z}$ mit $n_1 a_1 + \ldots + n_k a_k = 1$. Gruppierung in positive und negative Summanden liefert $1 = c - d$ mit $c, d \in A$. Dann ist $k_0 = (d + 1)(d - 1)$ wie gewünscht: Wir schreiben hierzu ein $k \geq k_0$ als $k = a\,d + r$, mit $0 \leq r < d$. Dann ist $a \geq d - 1$, und $k = a\,d + r = a\,d + r\,(c - d) \in A$.]

Übung 15 *(Primzahlen, I)*
Sei p eine Primzahl mit $p \equiv 1 \bmod(3)$. Zeigen Sie, daß $p \equiv 1 \bmod(6)$.

Übung 16 *(Primzahlen, II)*
Sei $p = 2^n - 1$ eine Primzahl. Zeigen Sie, daß n eine Primzahl ist.

[Es gilt $(a^n - 1) = (a^{n-1} + a^{n-2} + \ldots + a^1 + 1)(a - 1)$ für alle a.]

Übung 17 *(Primzahlen, III)*
Sei $k \geq 1$ beliebig. Zeigen Sie, daß es ein n gibt, so daß alle Zahlen $n,\ n + 1,\ n + 2,\ \ldots,\ n + k$ zusammengesetzt sind.

Übung 18 *(Primzahlen, IV)*
Zeigen Sie, daß es unendlich viele Primzahlen p der Form $p = 4k + 3$ gibt.

[Betrachten Sie $n = 2 \cdot 2 \cdot p_1 \cdot \ldots \cdot p_n - 1$.]

Übung 19 *(Primzahlen, V)*
Zeigen Sie, daß es unendlich viele Primzahlen p der Form $p = 6k + 5$ gibt.

Übung 20 *(Primzahlen, VI)*
Für $n \geq 1$ sei $F_n = 2^{2^n} - 1$ die n-te *Fermatsche Zahl*. Zeigen Sie:

$ggT(F_n, F_m) = 1$ für alle $n < m$.

Folgern Sie hieraus, daß es unendlich viele Primzahlen gibt.

[Zeigen Sie, daß F_n ein Teiler von $F_m - 2$ ist. Dann ist jeder gemeinsame Teiler von F_n und F_m ein Teiler von $F_m - 2$ und von F_m und daher gleich 1.
Zum Beweis von $F_n \mid (F_m - 2)$ schreiben Sie $(F_m - 2)/F_n$ in der Form $(a^k - 1)/(a + 1)$, $a \geq 1$, $k \geq 2$ gerade, und beweisen und verwenden, daß für diese a, k gilt:
$(a^k - 1)/(a + 1) = a^{k-1} - a^{k-2} + \ldots - a^0$.]

Übung 21 *(Eindeutigkeit der Primfaktorzerlegung, I)*
Seien $n, m \geq 2$. Wie lassen sich alle Teiler von n anhand der kanonischen Primfaktorzerlegung von n beschreiben? Wie ermittelt man den größten gemeinsamen Teiler ggT(n, m) und das kleinste gemeinsame Vielfache kgV(n, m) von n und m anhand der kanonischen Primfaktorzerlegungen von n und m? Zeigen Sie mit Hilfe dieser Ergebnisse noch einmal, daß $n \cdot m = ggT(n, m) \cdot kgV(n, m)$.

Übung 22 (Eindeutigkeit der Primfaktorzerlegung, II)
Die Zahlen 1, 4, 9, 16, …, n^2, …, heißen *Quadratzahlen*.
Seien $a, b \geq 1$ ungerade. Zeigen Sie, daß $a^2 + b^2$ keine Quadratzahl ist.

Übung 23 (Eindeutigkeit der Primfaktorzerlegung, III)
Zeigen Sie, daß die positive Quadratwurzel aus 3 irrational ist.

Übung 24 (Eindeutigkeit der Primfaktorzerlegung, IV)
Seien $a_0, \ldots, a_{k-1}$ ganze Zahlen, und sei $x \in \mathbb{R} - \mathbb{Z}$ mit

$$x^k + a_{k-1} x^{k-1} + \ldots + a_0 = 0.$$

Zeigen Sie, daß x irrational ist.

[Wir nehmen an, daß $x = n/m$ für $n \in \mathbb{Z}$ und $m \in \mathbb{N}^+$ gilt, wobei der Bruch n/m gekürzt sei. Dann gilt $n^k + a_{k-1} n^{k-1} m^1 + \ldots + a_0 m^k = 0$, im Widerspruch zu $ggT(n, m) = 1$.]

2. Grenzwerte

Wir präzisieren in diesem Kapitel die anschaulichen Vorstellungen des Grenzwerts einer Folge reeller Zahlen und der Stetigkeit einer reellwertigen Funktion. Zudem beweisen wir einige grundlegende Sätze im Umfeld dieser Begriffe.

Das Kapitel kann auch ohne eine genauere Kenntnis der im zweiten Abschnitt durchgeführten Konstruktion und Untersuchung von $\mathbb{R}$ gelesen werden. Es genügt, wenn der Leser ein Grundverständnis der reellen Zahlen mitbringt und mit den Begriffen „Infimum“ und „Supremum“ und den zugehörigen Sprechweisen und Notationen vertraut ist. Wir stellen diese Dinge noch einmal zusammen.

Für ein $X \subseteq \mathbb{R}$ und ein $s \in \mathbb{R}$ schreiben wir $X \leq s$, falls $x \leq s$ für alle $x \in X$ gilt. Analog ist $s \leq X$ definiert. Gilt $X \leq s$ ($s \leq X$), so heißt s eine *obere (untere) Schranke von X*. Ein $X \subseteq \mathbb{R}$ heißt nach *oben (unten) beschränkt*, falls ein s existiert mit $X \leq s$ ($s \leq X$). Ein $X \subseteq \mathbb{R}$ heißt *beschränkt*, falls X nach oben und nach unten beschränkt ist.

Ein $s^* \in \mathbb{R}$ heißt das *Supremum* von $X \subseteq \mathbb{R}$, in Zeichen $s^* = \sup(X)$, falls s^* die kleinste obere Schranke von X ist, d. h. es gilt:

$X \leq s^*$ und für alle $s \in \mathbb{R}$ mit $X \leq s$ ist $s^* \leq s$.

Analog heißt ein s^* das *Infimum* von X, in Zeichen $s^* = \inf(X)$, falls s^* die größte untere Schranke von X ist.

Wir verwenden im folgenden, daß jede nichtleere beschränkte Teilmenge X von $\mathbb{R}$ ein Infimum $x_0 = \inf(X)$ und ein Supremum $x_1 = \sup(X)$ in $\mathbb{R}$ besitzt. Diese fundamentale Eigenschaft der reellen Zahlen nennen wir auch die *(lineare) Vollständigkeit* von $\mathbb{R}$.

Konvergente Folgen

In der mathematischen Analysis spielt folgendes anschauliche Prinzip eine tragende Rolle:

Ist $x_0, x_1, \ldots, x_n, \ldots, n \in \mathbb{N}$, eine Folge reeller Zahlen, deren Glieder immer dichter beieinander liegen, so nähert sich diese Folge einer eindeutig bestimmten reellen Zahl an, dem sog. Grenzwert der Folge.

Diese Anschauung besitzt zwei vage Bestandteile: Zum einen die „Verdichtung“ einer Folge und zum anderen ihre „Annäherung“ an eine Zahl. Die Verdichtungseigenschaft präzisieren wir wie folgt, ohne dabei einen Grenzwertbegriff zu verwenden:

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_9, © Springer-Verlag Berlin Heidelberg 2010

Definition *(Cauchyfolge)*
Eine Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ in $\mathbb{R}$ heißt eine *Cauchyfolge*, falls gilt:

$$\forall \varepsilon > 0 \; \exists n_0 \; \forall n, m \geq n_0 \; |x_n - x_m| < \varepsilon. \qquad \textit{(Cauchybedingung)}$$

Die Annäherung einer Folge an eine reelle Zahl können wir durch eine ganz ähnliche Bedingung zum Ausdruck bringen:

Definition *(konvergente Folge, Grenzwert)*
Eine Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ in $\mathbb{R}$ *konvergiert*, falls ein $x \in \mathbb{R}$ existiert, sodaß gilt:

$$\forall \varepsilon > 0 \; \exists n_0 \; \forall n \geq n_0 \; |x_n - x| < \varepsilon. \qquad \textit{(Konvergenzbedingung)}$$

In diesem Fall heißt dann x ein *Grenzwert* der Folge.

Man kann leicht sehen, daß eine konvergente Folge eine Cauchyfolge ist, und daß ein Grenzwert im Falle der Existenz eindeutig bestimmt ist. Im Falle der Existenz schreiben wir dann

$$\lim(\langle x_n \mid n \in \mathbb{N} \rangle), \quad \lim_{n \to \infty} x_n, \quad \lim_{n \in \mathbb{N}} x_n \text{ oder } \lim_n x_n$$

für den eindeutig bestimmten Grenzwert x der Folge $\langle x_n \mid n \in \mathbb{N} \rangle$.
Damit können wir nun obiges Prinzip präzise formulieren und beweisen:

Satz *(Konvergenz von Cauchyfolgen in $\mathbb{R}$)*
Jede Cauchyfolge in $\mathbb{R}$ konvergiert.

Beweis
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Cauchyfolge in $\mathbb{R}$. Wir finden den Grenzwert dieser Folge mit Hilfe von Suprema und Infima. Die Menge $\{ x_n \mid n \in \mathbb{N} \}$ ist beschränkt (!), und damit können wir für alle $n \in \mathbb{N}$ definieren:

$$y_n = \sup(\{ x_m \mid m \geq n \}).$$

Dann gilt $y_0 \geq y_1 \geq \ldots \geq y_n \geq \ldots$ und $\{ y_n \mid n \in \mathbb{N} \}$ ist beschränkt, denn für alle n gilt $x_n \leq y_n$, und damit ist jede untere Schranke von $\{ x_n \mid n \in \mathbb{N} \}$ auch eine untere Schranke von $\{ y_n \mid n \in \mathbb{N} \}$. Also existiert

$$x^* = \inf(\{ y_m \mid m \in \mathbb{N} \}).$$

Wir zeigen, daß $\lim_{n \in \mathbb{N}} x_n = x^*$. Sei hierzu $\varepsilon > 0$ beliebig. Sei n_0 derart, daß $|x_n - x_m| < \varepsilon/2$ für alle $n, m \geq n_0$. Weiter sei $n_1 \geq n_0$ derart, daß $y_{n_1} - x^* < \varepsilon/2$. Nach Definition von y_{n_1} gibt es ein $n_2 \geq n_1$ mit $y_{n_1} - x_{n_2} \leq \varepsilon/2$. Dann gilt:

$$(+) \quad |x^* - x_{n_2}| \leq \varepsilon/2.$$

Denn es gilt $x^* \leq x_{n_2} \leq y_{n_1}$ oder $x_{n_2} \leq x^* \leq y_{n_1}$. Im ersten Fall folgt (+) aus $y_{n_1} - x^* \leq \varepsilon/2$, im zweiten aus $y_{n_1} - x_{n_2} \leq \varepsilon/2$. Für alle $n \geq n_2$ gilt dann wegen (+) und $n, n_2 \geq n_0$:

$$|x^* - x_n| = |x^* - x_{n_2} + x_{n_2} - x_n| \leq |x^* - x_{n_2}| + |x_{n_2} - x_n| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

∎

Der Beweis motiviert die folgenden Begriffe:

Definition *(Limes Inferior, Limes Superior)*
Sei $\langle x_n \mid n \in \mathbb{N}\rangle$ eine beschränkte Folge in $\mathbb{R}$ (d.h. $\{ x_n \mid n \in \mathbb{N} \}$ ist beschränkt). Dann definieren wir:

$\liminf(\langle x_n \mid n \in \mathbb{N}\rangle) = \sup(\{ \inf(\{ x_m \mid m \geq n \}) \mid n \in \mathbb{N} \})$,

$\limsup(\langle x_n \mid n \in \mathbb{N}\rangle) = \inf(\{ \sup(\{ x_m \mid m \geq n \}) \mid n \in \mathbb{N} \})$.

Die reelle Zahl $\liminf(\langle x_n \mid n \in \mathbb{N}\rangle)$ heißt der *Limes Inferior* der Folge und $\limsup(\langle x_n \mid n \in \mathbb{N}\rangle)$ ihr *Limes Superior*.

Wir verwenden auch die suggestive Notation

$\liminf_{n \in \mathbb{N}} x_n = \sup_{n \in \mathbb{N}} \inf_{m \geq n} x_m$, $\limsup_{n \in \mathbb{N}} x_n = \inf_{n \in \mathbb{N}} \sup_{m \geq n} x_m$.

Diese Werte haben eine sehr anschauliche Bedeutung: Wir bilden zunächst das Infimum der Folge. Nun streichen wir das erste Folgenglied und korrigieren unser Infimum evtl. nach oben. Danach streichen wir auch das zweite Folgenglied und korrigieren unser Infimum, usw. Der Grenzwert der so korrigierten Infima ist der Limes Inferior der Folge. Analog erhalten wir den Limes Superior, wenn wir das Supremum der Folge iteriert durch Streichen von Elementen der Folge absenken.

Offenbar gilt für alle beschränkten Folgen $\langle x_n \mid n \in \mathbb{N}\rangle$:

$\liminf_{n \in \mathbb{N}} x_n \leq \limsup_{n \in \mathbb{N}} x_n$.

Der Fall der Gleichheit der beiden Werte verdient eine Formulierung als Satz:

Satz *(Limes als Limes Inferior und als Limes Superior)*
Sei $\langle x_n \mid n \in \mathbb{N}\rangle$ eine beschränkte Folge in $\mathbb{R}$. Dann sind äquivalent:

(a) $\langle x_n \mid n \in \mathbb{N}\rangle$ konvergiert.

(b) $\liminf_{n \in \mathbb{N}} x_n = \limsup_{n \in \mathbb{N}} x_n$.

In diesem Fall gilt dann $\lim_{n \in \mathbb{N}} x_n = \liminf_{n \in \mathbb{N}} x_n = \limsup_{n \in \mathbb{N}} x_n$.

Der Beweis sei dem Leser zur Übung überlassen.

In Spezialfällen kann der Limes einer Folge einfacher ausgedrückt werden. Eine Folge $\langle x_n \mid n \in \mathbb{N}\rangle$ heißt *monoton wachsend*, falls $x_n \leq x_{n+1}$ für alle $n \in \mathbb{N}$ gilt, und *streng monoton wachsend*, falls $x_n < x_{n+1}$ für alle $n \in \mathbb{N}$ gilt. Analog sind monoton fallende und streng monoton fallende Folgen definiert. Ist nun $\langle x_n \mid n \in \mathbb{N}\rangle$ monoton wachsend und beschränkt, so gilt

$\lim_{n \in \mathbb{N}} x_n = \sup(\{ x_n \mid n \in \mathbb{N} \})$.

Analog ist $\inf(\{ x_n \mid n \in \mathbb{N} \})$ der Grenzwert einer monoton fallenden beschränkten Folge $\langle x_n \mid n \in \mathbb{N}\rangle$.

Häufungspunkte

Nachdem wir definiert haben, wann ein Folge sich einem Punkt annähert, wollen wir nun noch präzisieren, was es bedeutet, daß sich eine Menge von reellen Zahlen um einen Punkt herum verdichtet. Hierzu brauchen wir einige Notationen und Begriffe. Für alle $x, y \in \mathbb{R}$ definieren wir:

$$]x, y[= \{ z \in \mathbb{R} \mid x < z < y \}, \qquad \textit{(offenes Intervall)}$$

$$[x, y] = \{ z \in \mathbb{R} \mid x \leq z \leq y \}. \qquad \textit{(abgeschlossenes Intervall)}$$

Analog sind die halboffenen Intervalle $[x, y[$ und $]y, x]$ definiert, und weiter definieren wir auch noch die unbeschränkten Intervalle $]\infty, y[= \{ z \in \mathbb{R} \mid z < y \}$, $[x, \infty[= \{ z \in \mathbb{R} \mid x \leq z \}$, $]-\infty, \infty[= \mathbb{R}$, usw.

Ebenso einfach wie wichtig sind die folgenden Umgebungen eines Punktes:

Definition *(ε-Umgebung)*
Für alle $x \in \mathbb{R}$ und alle $\varepsilon > 0$ heißt die Menge $U_\varepsilon(x) =]x - \varepsilon, x + \varepsilon[$ die *(offene)* ε-*Umgebung* von x.

Damit können wir nun definieren:

Definition *(Häufungspunkt)*
Sei $X \subseteq \mathbb{R}$. Ein $x \in \mathbb{R}$ heißt ein *Häufungspunkt* von X, falls gilt:

Für alle $\varepsilon > 0$ ist $(U_\varepsilon(x) - \{ x \}) \cap X \neq \emptyset$.

Ist x ein Häufungspunkt von X, so ist $U_\varepsilon(x) \cap X$ unendlich für alle $\varepsilon > 0$ (!).

Ein Häufungspunkt x von X kann der Menge X angehören oder nicht. Eine abzählbare Menge kann überabzählbar viele Häufungspunkte haben. So ist z. B. jede reelle Zahl ein Häufungspunkt von $\mathbb{Q}$.

Die Vollständigkeit von $\mathbb{R}$ führt nun zu folgendem fundamentalen Existenzsatz über Häufungspunkte:

Satz *(Satz von Bolzano-Weierstraß)*
Sei $X \subseteq \mathbb{R}$ unendlich und beschränkt. Dann existiert ein Häufungspunkt von X.

Beweis
Wir setzen

$$Y = \{ x \in \mathbb{R} \mid X \cap]-\infty, x] \text{ ist endlich} \}.$$

Dann gilt $\inf(X) \in Y$ (da X beschränkt) und $\sup(X) \notin Y$ (da X unendlich). Also ist $Y \neq \emptyset$ und nach oben beschränkt, und damit existiert $x^* = \sup(Y)$. Weiter gilt $Y =]-\infty, x^*[$ oder $Y =]-\infty, x^*]$ nach Definition von Y.

1. Fall: Es gilt $x^* \notin Y$.

Dann ist $]-\infty, x^*] \cap X = (]-\infty, x^* - \varepsilon] \cap X) \cup (]x^* - \varepsilon, x^*] \cap X)$ unendlich für alle $\varepsilon > 0$. Wegen $x^* - \varepsilon \in Y$ ist dann aber $]x^* - \varepsilon, x^*] \cap X$ unendlich für alle $\varepsilon > 0$. Also ist x^*ein Häufungspunkt von X.

2. Fall: Es gilt $x^* \in Y$.

Dann ist $]x^*, x^* + \varepsilon] \cap X$ unendlich für alle $\varepsilon > 0$, da sonst $x^* + \varepsilon \in Y$ für ein $\varepsilon > 0$ wäre, im Widerspruch zu $x^* = \sup(Y)$. Also ist auch in diesem Fall x^* ein Häufungspunkt von X. ∎

Aus diesem Satz folgt, daß wir aus einer beschränkten Folge in $\mathbb{R}$ eine konvergente Teilfolge „extrahieren" oder „ausheben" können. Dabei nennen wir eine Folge $\langle y_n \mid n \in \mathbb{N} \rangle$ eine *Teilfolge* von $\langle x_n \mid n \in \mathbb{N} \rangle$, falls eine streng monoton wachsende Funktion $g : \mathbb{N} \to \mathbb{N}$ existiert mit $y_n = x_{g(n)}$. Die Folge $\langle y_n \mid n \in \mathbb{N} \rangle$ heißt dann die durch g bestimmte Teilfolge von $\langle x_n \mid n \in \mathbb{N} \rangle$. So bestimmt zum Beispiel die Funktion g mit $g(n) = 2n$ die Teilfolge $\langle x_{2n} \mid n \in \mathbb{N} \rangle$.

Der Satz von Bolzano-Weierstraß liefert nun:

Korollar *(Existenz von konvergenten Teilfolgen)*
Jede beschränkte Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ reeller Zahlen besitzt eine konvergente Teilfolge.

Beweis
Ist $X = \{ x_n \mid n \in \mathbb{N} \}$ endlich, so gibt es ein $c \in \mathbb{R}$ mit $x_n = c$ für unendlich viele $n \in \mathbb{N}$. Offenbar ist dann die konstante Folge $\langle c \mid n \in \mathbb{N} \rangle$ eine konvergente Teilfolge von $\langle x_n \mid n \in \mathbb{N} \rangle$.
Sei also X unendlich. Nach Voraussetzung ist X beschränkt, und nach dem Satz von Bolzano-Weierstraß existiert also ein Häufungspunkt x^* von X. Wir definieren nun rekursiv

$g(0)$ = „das kleinste k mit $|x^* - x_k| < 1$",

$g(n)$ = „das kleinste $k > g(n-1)$ mit $|x^* - x_k| < 1/2^n$" für alle $n \geq 1$.

Ein solches k existiert, da $U_{1/2^n}(x^*) \cap X$ unendlich ist für alle $n \in \mathbb{N}$.
Nach Konstruktion ist $g : \mathbb{N} \to \mathbb{N}$ streng monoton steigend und es gilt

$\lim_{n \in \mathbb{N}} x_{g(n)} = x^*$,

denn ist $\varepsilon > 0$ und $1/2^{n_0} \leq \varepsilon$, so ist $|x^* - x_{g(n)}| < 1/2^n \leq \varepsilon$ für alle $n \geq n_0$. ∎

Intervallschachtelungen

Folgendes Prinzip ist vielfach nützlich:

Satz *(Intervallschachtelung)*
Seien $I_n = [a_n, b_n]$, $a_n \leq b_n$, beschränkte Intervalle mit $I_{n+1} \subseteq I_n$ für alle $n \in \mathbb{N}$. Dann gilt $\bigcap I_n \neq \emptyset$.

Beweis

Wegen $a_n \leq b_0$ und $b_n \geq a_0$ für alle $n \in \mathbb{N}$ existieren

$a = \sup(\{ a_n \mid n \in \mathbb{N} \})$ und $b = \inf(\{ b_n \mid n \in \mathbb{N} \})$.

Dann gilt $a \leq b$ und $\bigcap_{n \in \mathbb{N}} I_n = [a, b] \neq \emptyset$.

Zur Illustration zeigen wir den Satz von Bolzano-Weierstraß mit Hilfe einer Intervallschachtelung.

Zweiter Beweis des Satzes von Bolzano-Weierstraß

Sei also X eine unendliche Teilmenge des beschränkten Intervalls $[a, b]$. Durch iterierte Intervallhalbierung definieren wir rekursiv Intervalle $I_0 = [a, b] \supseteq I_1 \supseteq I_2 \supseteq \ldots \supseteq I_n \supseteq \ldots$, sodaß gilt:

$I_n \cap X$ ist unendlich für alle n.

(Da $I_n \cap X$ nach I. V. unendlich ist, muß die linke oder die rechte Hälfte von I_n unendlich viele Elemente von X enthalten.)

Sei x^* das eindeutige Element von $\bigcap_{n \in \mathbb{N}} I_n$. Dann ist x^* ein Häufungspunkt von X, denn für alle $\varepsilon > 0$ ist $I_n \subseteq U_\varepsilon(x^*)$ für alle n mit $(b - a)/2^n < \varepsilon$.

Reihen

Mit Hilfe des Grenzwertbegriffs können wir die unendliche Summe

$$x_0 + x_1 + \ldots + x_n + \ldots$$

von gewissen Folgen $x_0, x_1, \ldots, x_n, \ldots$ reeller Zahlen definieren.

Definition *(unendliche Reihe, Partialsumme, Summe)*

Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Folge in $\mathbb{R}$. Für alle $n \in \mathbb{N}$ setzen wir

$$y_n = \Sigma_{i \leq n} x_i.$$

Dann heißt y_n die n-te *Partialsumme* der Folge $\langle x_n \mid n \in \mathbb{N} \rangle$. Weiter heißt $\langle y_n \mid n \in \mathbb{N} \rangle$ die durch $\langle x_n \mid n \in \mathbb{N} \rangle$ gegebene *unendliche Reihe*. Im Falle der Konvergenz von $\langle y_n \mid n \in \mathbb{N} \rangle$ setzen wir

$$\Sigma_{n \in \mathbb{N}} x_n = \lim_{n \in \mathbb{N}} y_n,$$

und nennen die reelle Zahl $\Sigma_{n \in \mathbb{N}} x_n$ die *Summe* von $\langle x_n \mid n \in \mathbb{N} \rangle$.

Es ist üblich, die Notation $\Sigma_{n \in \mathbb{N}} x_n$ nicht nur für den Limes der Partialsummen y_n zu verwenden, sondern zugleich auch für die betrachtete Reihe $\langle y_n \mid n \in \mathbb{N} \rangle$ selbst. So kann man dann sagen: „Sei $\Sigma_{n \in \mathbb{N}} x_n$ eine unendliche Reihe.“, was einfacher und suggestiver ist als „Sei $\langle \Sigma_{i \leq n} x_i \mid n \in \mathbb{N} \rangle$ eine unendliche Reihe“. Der Ausdruck „$\Sigma_{n \in \mathbb{N}} x_n$“ bedeutet also immer die Folge der Partialsummen der Zahlen x_n und im Fall der Konvergenz dieser Folge auch ihren Grenzwert.

Die Reihen $\Sigma_{n\in\mathbb{N}}(-1)^n = 1 - 1 + 1 - 1 + \ldots$ und $\Sigma_{n\in\mathbb{N}} 1 = 1 + 1 + 1 + 1 + \ldots$ sind beide divergent. Die Partialsummen der ersten Reihe pendeln zwischen 1 und 0 hin und her. Dagegen übertreffen die Partialsummen der zweiten Reihe schließlich jede reelle Zahl, und eine Sprechweise der „Konvergenz gegen unendlich" drängt sich hier auf. Wir führen sie allgemein für Folgen ein:

Definition *(uneigentliche Konvergenz oder bestimmte Divergenz von Folgen)*
Eine Folge $\langle x_n \mid n \in \mathbb{N}\rangle$ in $\mathbb{R}$ heißt *uneigentlich konvergent* oder *bestimmt divergent* gegen ∞, in Zeichen $\lim_{n\in\mathbb{N}} x_n = \infty$, falls gilt:

$\forall k \in \mathbb{N}\, \exists n_0 \in \mathbb{N}\, \forall n \geq n_0\ x_n \geq k.$

Analog ist $\lim_{n\in\mathbb{N}} x_n = -\infty$ definiert.

Gilt $x_n \geq 0$ für alle Summanden einer Reihe $\Sigma_{n\in\mathbb{N}} x_n$, so wachsen die Partialsummen monoton und die Reihe konvergiert genau dann, wenn die Partialsummen nach oben beschränkt sind. In diesem Fall gilt $\Sigma_{n\in\mathbb{N}} x_n = \sup_{n\in\mathbb{N}} \Sigma_{i\leq n} x_i$, andernfalls ist $\Sigma_{n\in\mathbb{N}} x_n = \infty$. Ein Beispiel hierzu ist die *harmonische Reihe*

$$\Sigma_{n\in\mathbb{N}}\, 1/(n+1) \;=\; 1 + 1/2 + 1/3 + \ldots + 1/n + \ldots$$

Ein zeitloses Argument der Mathematik zeigt, daß diese Reihe bestimmt divergiert, d.h. es gilt $\Sigma_{n\in\mathbb{N}}\, 1/(n+1) = \infty$. Wir diskutieren dies in den Übungen.

Die Berechnung des Grenzwerts einer unendliche Reihe ist nur selten in einfacher Art und Weise möglich. Einen solchen Fall bilden die Reihen der Form $\Sigma_{n\in\mathbb{N}} x^n$, die sogenannten (unendlichen) *geometrischen Reihen.* Ist $|x| \geq 1$, so ist $\Sigma_{n\in\mathbb{N}} x^n$ divergent. Ist dagegen $|x| < 1$, so gilt

$$\Sigma_{n\in\mathbb{N}}\, x^n \;=\; 1/(1-x). \qquad \textit{(Konvergenz der geometrischen Reihe für } |x| < 1\textit{)}$$

In den Übungen findet der Leser einen Hinweis auf den überraschend einfachen Beweis dieser Konvergenzaussage. Es gilt also zum Beispiel:

$$\Sigma_{n\in\mathbb{N}}\,(1/2)^n \;=\; 1 + 1/2 + 1/4 + 1/8 + \ldots \;=\; 1/(1-1/2) \;=\; 2,$$

$$\Sigma_{n\in\mathbb{N}}\,(-1/3)^n \;=\; 1 - 1/3 + 1/9 - 1/27 + \ldots \;=\; 1/(1+1/3) \;=\; 3/4.$$

Die erste Reihe kann man sich durch eine Verteilung zweier Kuchen an unendlich viele Kunden durch wiederholte Halbierung anschaulich machen.

Wir geben schließlich noch einige Beispiele für konvergente unendliche Reihen an, die man mit weitergehenden Methoden berechnen kann. Durch die ihnen innewohnende Schönheit und Magie dürfen sie für sich stehen:

$$\Sigma_{n\in\mathbb{N}}\,(-1)^n/(n+1) \;=\; 1 - 1/2 + 1/3 - 1/4 + \ldots \;=\; \ln(2),$$

$$\Sigma_{n\in\mathbb{N}}\,(-1)^n/(2n+1) \;=\; 1 - 1/3 + 1/5 - 1/7 + \ldots \;=\; \pi/4,$$

$$\Sigma_{n\geq 1}\, 1/n^2 \;=\; 1 + 1/4 + 1/9 + 1/16 + \ldots \;=\; \pi^2/6.$$

Die Berechnung der letzten Reihe gelang bereits Leonhard Euler. Ebenso kann man die Summen $\Sigma_{n\geq 1}\, 1/n^k$ für gerade k bestimmen. Dagegen entziehen sich die Summen für ungerade k einer konkreten Berechnung. Man weiß durch ein Ergebnis von Apéry (1979), daß $\Sigma_{n\geq 1}\, 1/n^3 = 1{,}202\ldots$ irrational ist.

Stetige Funktionen

Anschaulich bedeutet die Stetigkeit einer Funktion, daß die Funktion keine Sprünge macht, oder, etwas genauer, daß sich jeder Funktionswert f(a) nur wenig verändert, wenn sich das Argument a nur hinreichend wenig verändert.

Mit Hilfe von Grenzwerten können wir den Stetigkeitsbegriff präzise fassen. Im folgenden betrachten wir Funktionen $f : A \to \mathbb{R}$ mit einem prinzipiell beliebigen Definitionsbereich $A \subseteq \mathbb{R}$. Der Leser denke in erster Linie an Definitionsbereiche, die die Form eines Intervalls haben oder die die Vereinigung von endlich vielen Intervallen sind.

Wir definieren:

Definition *(Stetigkeit)*

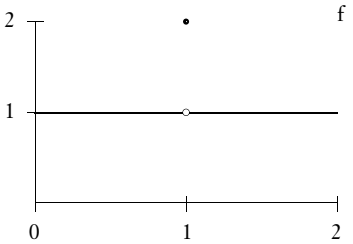
Eine Funktion $f : A \to \mathbb{R}$ heißt *stetig in* a für ein $a \in A$, falls für alle Folgen $\langle x_n \mid n \in \mathbb{N} \rangle$ in A gilt:

$\lim_{n \in \mathbb{N}} x_n = a$ *impliziert* $\lim_{n \in \mathbb{N}} f(x_n) = f(a)$.

Weiter heißt f *stetig (schlechthin)*, falls f stetig in allen $a \in A$ ist.

Im Stetigkeitsfall gilt also $f(\lim_{n \in \mathbb{N}} x_n) = \lim_{n \in \mathbb{N}} f(x_n)$ für alle konvergenten Folgen $\langle x_n \mid n \in \mathbb{N} \rangle$ in A, deren Grenzwert in A liegt.

Wir betrachten einige Beispiele für stetige und unstetige Funktionen. Die vielleicht einfachste Unstetigkeitsstelle ist die „Punktierung" einer konstanten Funktion. Sei hierzu $f : [0, 2] \to [0, 2]$ mit $f(x) = 1$, falls $x \neq 1$, und $f(1) = 2$. Dann ist f unstetig im Punkt 1 und stetig in allen anderen Punkten des Intervalls $[0, 2]$. Zum Beweis der Unstetigkeit von f im Punkt 1 definieren wir eine Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ in $[0, 2]$ durch $x_n = 1 + 1/2^n$ für alle n. Dann gilt $\lim_{n \in \mathbb{N}} x_n = 1$, aber $\lim_{n \in \mathbb{N}} f(x_n) = \lim_{n \in \mathbb{N}} 1 = 1 \neq f(1)$.

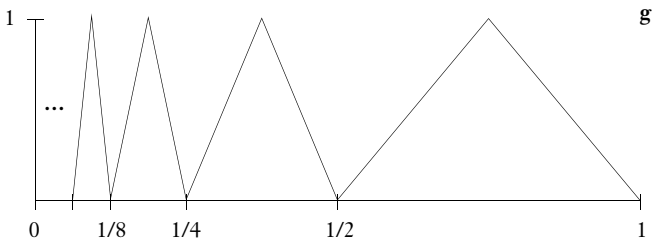


Also ist f unstetig im Punkt 1. Ist aber $x \in [0, 2]$, $x \neq 1$, und ist $\langle x_n \mid n \in \mathbb{N} \rangle$ eine gegen x konvergente Folge in $[0, 2]$, so gibt es ein $\varepsilon > 0$ mit $1 \notin [x - \varepsilon, x + \varepsilon]$ und ein $n_0 \in \mathbb{N}$ mit $x_n \in [x - \varepsilon, x + \varepsilon]$ für alle $n \geq n_0$. Dann gilt $f(x_n) = 1$ für alle $n \geq n_0$. Also ist $\lim_{n \in \mathbb{N}} f(x_n) = 1 = f(x)$. Dies zeigt, daß die Funktion f in allen von 1 verschiedenen Punkten ihres Definitionsbereichs stetig ist.

Wir können auch auf einer dichten Menge punktieren. Sei $\mathrm{ind}_{\mathbb{Q}} : \mathbb{R} \to [0, 1]$ die Indikatorfunktion von $\mathbb{Q}$, d. h. $\mathrm{ind}_{\mathbb{Q}}(x) = 1$, falls $x \in \mathbb{Q}$ und $\mathrm{ind}_{\mathbb{Q}}(x) = 0$ sonst. Der Leser zeigt leicht, daß diese Funktion in jedem $x \in \mathbb{R}$ unstetig ist.

Als nächstes betrachten wir die im folgenden Diagramm visualisierte Funktion $g : [0, 1] \to \mathbb{R}$, die sich in immer schmäler werdenden Zacken der Höhe 1 der y-Achse annähert. Für jede Definition von $g(0) \in \mathbb{R}$ ist diese Funktion unstetig

im Punkt 0. Denn für jedes $y \in [0, 1]$ gibt es eine gegen 0 konvergente Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ in $[0, 1]$ mit $g(x_n) = y$ für alle n, sodaß also $\lim_{n \in \mathbb{N}} g(x_n) = y$.



Jede konstante Funktion ist stetig. Weiter ist die Identität $id_A : A \to \mathbb{R}$ stetig für alle $A \subseteq \mathbb{R}$. Sind $f : A \to B$ stetig in a und $g : B \to \mathbb{R}$ stetig in f(a), so ist auch die Verknüpfung $h = g \circ f$ stetig in a. Eine Untersuchung der Definition der Addition und der Multiplikation auf $\mathbb{R}$ zeigt, daß die punktweise Addition und Multiplikation zweier stetiger Funktionen wieder stetig ist. Folglich ist jedes reelle Polynom $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \quad \text{für alle } x \in \mathbb{R},$$

eine stetige Funktion. Dagegen sind punktweise Limiten von Polynomen im allgemeinen nicht mehr stetig:

Wir betrachten die Funktionen $h_n : [0, 1] \to [0, 1]$, $n \in \mathbb{N}$, mit $h_n(x) = x^{n+1}$ für alle $x \in [0, 1]$. Diese stetigen Funktionen verbinden die Punkte (0, 0) und (1, 1) durch immer flacher werdende Kurven. Für jedes $x \in [0, 1[$ gilt $\lim_{n \in \mathbb{N}} h_n(x) = 0$, und damit konvergieren die Funktionen h_n punktweise gegen die Funktion h mit $h(x) = 0$ für alle $x \in [0, 1[$ und $h(1) = 1$. Die Stetigkeit von Funktionen kann bei einem punktweisen Grenzübergang also verloren gehen.

Wir diskutieren nun noch einige äquivalente Formulierungen der Stetigkeit. Die folgende sog. ε-δ-Stetigkeit ist vielleicht die direkteste Präzisierung der „kleinen Änderung von f(a) bei kleiner Änderung von a":

Satz *(ε-δ-Formulierung der Stetigkeit)*
Sei $f : A \to \mathbb{R}$, und sei $a \in A$. Dann sind äquivalent:

(a) f ist stetig in a.

(b) $\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in A \; (|x - a| < \delta \to |f(x) - f(a)| < \varepsilon)$.

Beweis

(a) ⇝ (b):

Wir führen den Beweis indirekt. Sei also $\varepsilon > 0$ derart, daß für alle $\delta > 0$ ein $x \in A$ existiert mit $|x - a| < \delta$ und $|f(x) - f(a)| \geq \varepsilon$.
Dann existiert für alle $n \in \mathbb{N}$ ein $x_n \in A$ mit

$$|x_n - a| < 1/2^n \text{ und } |f(x) - f(a)| \geq \varepsilon.$$

Dann konvergiert aber $\langle x_n \mid n \in \mathbb{N} \rangle$ gegen a, während $\langle f(x_n) \mid n \in \mathbb{N} \rangle$ nicht gegen f(a) konvergiert. Also ist f nicht stetig in a.

(b) ⌢ (a):
Sei $\langle x_n \mid n \in \mathbb{N}\rangle$ eine Folge mit $\lim_n x_n = a$. Wir zeigen $\lim_n f(x_n) = f(a)$.
Sei hierzu $\varepsilon > 0$ beliebig. Für dieses ε sei nun $\delta > 0$ wie in (b).
Wegen $\lim_n x_n = a$ gibt es ein n_0 derart, daß $|x_n - a| < \delta$ für alle $n \geq n_0$.
Nach Wahl von δ ist dann aber $|f(x_n) - f(a)| < \varepsilon$ für alle $n \geq n_0$.

Der Wertebereich stetiger Funktionen

Eine Funktion, die auf einem Intervall $[a, b]$ definiert ist und sich dort stetig ändert, kann anschaulich keine Werte auslassen: Um von einem $f(x)$ zu einem $f(y)$ zu gelangen, muß jeder Wert z als Funktionswert angenommen werden, der zwischen $f(x)$ und $f(y)$ liegt. Diese Anschauung wollen wir nun präzisieren und beweisen.

Im folgenden sei stets $[a, b]$ ein beschränktes reelles Intervall. Derartige Intervalle heißen auch *kompakt*. Wir nehmen zudem immer $a < b$ an.

Zunächst zeigen wir:

Satz *(Nullstellensatz)*
Sei $f : [a, b] \to \mathbb{R}$ stetig, und $f(a)$ und $f(b)$ haben verschiedene Vorzeichen.
Dann gibt es ein $c \in [a, b]$ mit $f(c) = 0$.

Beweis
Wir nehmen der Einfachheit halber an, daß $f(a) < 0$ und $f(b) > 0$ gilt.
Aus dem ε-δ-Kriterium folgt, daß für alle $y \in]a, b[$ mit $f(y) > 0$ gilt:

(+) Es gibt ein $\delta > 0$ mit: $x \in [a, b]$ und $f(x) > 0$ für alle $x \in]y - \delta, y + \delta[$.

Wir setzen $X = \{ x \in [a, b] \mid f(x) > 0 \}$. Dann ist X nichtleer und beschränkt, also existiert $x^* = \inf(X) \in [a, b]$. Nach (+) und Definition von X ist $f(x^*) \leq 0$.
Wegen $f(x^*) \leq 0$ ist $x^* \notin X$ und damit als Infimum von X ein Häufungspunkt von X. Also gibt es eine Folge $\langle x_n \mid n \in \mathbb{N}\rangle$ in X mit $\lim_n x_n = x^*$. Dann gilt $f(x^*) = \lim_n f(x_n) \geq 0$, da $f(x_n) \geq 0$ für alle n. Insgesamt also $f(x^*) = 0$.

Aus diesem Satz folgt, daß jedes reelle Polynom $f(x) = \alpha_n x^n + \ldots + \alpha_1 x + \alpha_0$ mit $\alpha_i \in \mathbb{R}$, $\alpha_n \neq 0$, n ungerade, eine Nullstelle besitzt. Denn ist $b > 0$ genügend groß, so haben $f(-b)$ und $f(b)$ verschiedene Vorzeichen.

Aus dem Nullstellensatz folgt stärker:

Korollar *(Zwischenwertsatz)*
Sei $f : [a, b] \to \mathbb{R}$ stetig, und es gelte $c, d \in \mathrm{rng}(f)$ mit $c < d$.
Dann ist $[c, d] \subseteq \mathrm{rng}(f)$.

Beweis
Sei $e \in]c, d[$. Sei $c = f(a')$, $d = f(b')$. Sei I das durch a' und b' gegebene nichtleere kompakte Intervall. Wir setzen

$g(x) = f(x) - e$ für alle $x \in I$.

Dann ist $g : I \to \mathbb{R}$ stetig und $g(a')$ und $g(b')$ haben verschiedene Vorzeichen. Nach dem Nullstellensatz gibt es ein x mit $g(x) = 0$. Also ist

$f(x) = g(x) + e = e$. ▁

Weiter existieren nun aber für stetige Funktionen auf kompakten Intervallen nicht nur Zwischen-, sondern auch Extremwerte:

Satz *(Annahme von Maximum und Minimum)*
Sei $f : [a, b] \to \mathbb{R}$ stetig. Dann gibt es ein $c \in [a, b]$ mit

$f(x) \leq f(c)$ für alle $x \in [a, b]$.

Analog existiert ein $d \in [a, b]$ mit $f(d) \leq f(x)$ für alle $x \in [a, b]$.

Beweis
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Folge in $[a, b]$ mit:

(i) $\lim_n f(x_n) = \sup(\mathrm{rng}(f))$, *falls* $\mathrm{rng}(f)$ nach oben beschränkt ist,

(ii) $f(x_n) \geq n$ für alle $n \in \mathbb{N}$, *sonst.*

Sei $\langle y_n \mid n \in \mathbb{N} \rangle$ eine gegen $c \in [a, b]$ konvergierende Teilfolge von $\langle x_n \mid n \in \mathbb{N} \rangle$. Dann konvergiert $\langle f(y_n) \mid n \in \mathbb{N} \rangle$ gegen $f(c)$ und damit ist (ii) ausgeschlossen.
Weiter ist $f(c) = \lim_n f(y_n) = \lim_n f(x_n) = \sup(\mathrm{rng}(f))$, also ist c wie gewünscht. ▁

Insgesamt haben wir gezeigt:

Korollar *(Wertebereich stetiger Funktionen auf kompakten Intervallen)*
Sei $f : [a, b] \to \mathbb{R}$ stetig. Dann existieren $c \leq d$ mit $\mathrm{rng}(f) = [c, d]$.

Differentialquotienten

Zur Illustration des Stetigkeitsbegriffs definieren wir noch das grundlegende Konzept der Differentialrechnung, ohne es genauer zu untersuchen.

Ist $g : \mathbb{R} \to \mathbb{R}$ eine Gerade mit Steigung c, d. h. gilt $g(x) = cx + d$ für ein $d \in \mathbb{R}$ und alle $x \in \mathbb{R}$, so können wir die Steigung c dieser Funktion berechnen, sobald wir zwei Funktionswerte $g(a)$ und $g(b)$ mit $a \neq b$ kennen. Denn es gilt

$(g(b) - g(a))/(b - a) = (cb + d - ca - d)/(b - a) = c.$

Wir versuchen nun, die Steigung einer beliebigen Funktion f in einem Punkt a - über den Quotienten $(f(b) - f(a))/(b - a)$ zu erklären. Dies gelingt, wenn dieser Quotient gute Stetigkeitseigenschaften besitzt:

Definition *(Differentialquotient)*
Sei $f : A \to \mathbb{R}$ eine Funktion, und sei $a \in A$ ein Häufungspunkt von A. Dann heißt f *differenzierbar* in a mit *Differentialquotient* oder *Ableitung* $c \in \mathbb{R}$, falls für jede gegen a konvergente Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ in $A - \{ a \}$ gilt:

$$c = \lim_{n \in \mathbb{N}} \frac{f(x_n) - f(a)}{x_n - a}.$$

Definieren wir $h(x) = (f(x) - f(a))/(x - a)$ für $x \in A - \{ a \}$ und $h(a) = c$, so besagt die Differenzierbarkeit von f in a mit Differentialquotient c gerade, daß die Funktion $h : A \to \mathbb{R}$ stetig in a ist. Damit ist die Differenzierbarkeit von f eine Stetigkeitsforderung an eine mit Hilfe von f berechnete Funktion, und es ist nicht überraschend, daß diese Forderung die Stetigkeit der Funktion f im Punkt a impliziert.

Offene Mengen und Umgebungen

Wir wollen nun noch eine weitere Charakterisierung der Stetigkeit etablieren, der in der Mathematik eine große Bedeutung zukommt. Hierzu definieren wir:

Definition *(offene Menge)*
Ein Menge U reeller Zahlen heißt *offen*, wenn für alle $x \in U$ ein $\varepsilon > 0$ existiert mit $U_\varepsilon(x) \subseteq U$.

Anschaulich besagt diese Bedingung, daß kein Punkt von U am Rand von U liegt. Ein $U \subseteq \mathbb{R}$ ist offenbar genau dann offen, wenn ein System $\mathcal{S}$ offener Intervalle existiert mit $U = \bigcup \mathcal{S}$. Die offenen Intervalle erzeugen in diesem Sinne die offenen Mengen. Die leere Menge ist nach obiger Definition offen. Das Mengensystem $\mathcal{U}$ der offenen Mengen ist weiter abgeschlossen unter endlichen Durchschnitten und beliebigen Vereinigungen, d. h. für alle $U_1, \ldots, U_n \in \mathcal{U}$ ist $U_1 \cap \ldots \cap U_n \in \mathcal{U}$ und für alle $\mathcal{V} \subseteq \mathcal{U}$ ist $\bigcup \mathcal{V} \in \mathcal{U}$.
Wir definieren weiter:

Definition *(Umgebung)*
Sei $x \in \mathbb{R}$. Ein $P \subseteq \mathbb{R}$ heißt eine *Umgebung* von x, falls ein offenes $U \subseteq P$ existiert mit $x \in U$.

Als Umgebung von x gilt also jede Menge P, die ein offenes Intervall enthält, dem x angehört. Die Menge P selbst muß nicht offen sein. Ist P aber offen, so nennen wir P eine *offene Umgebung* von x.
Wir wollen nun mit Hilfe dieser Begriffe eine Charakterisierung der Stetigkeit einer Funktion $f : A \to \mathbb{R}$ geben. Da unsere Funktionen aber einen beliebigen Definitionsbereich A haben können, brauchen wir noch relative Begriffe:

Definition *(offen in, Umgebung in)*
Sei $A \subseteq \mathbb{R}$. Ein $V \subseteq A$ heißt *offen in* A, falls ein offenes $U \subseteq \mathbb{R}$ existiert mit

$$V = U \cap A.$$

Weiter heißt für ein $x \in A$ ein $P \subseteq A$ eine *Umgebung von* x *in* A, falls ein in A offenes V existiert mit $x \in V$ und $V \subseteq P$.

Ist etwa $A = [a, b]$, so ist $[a, c[$ offen in A für alle $a \leq c \leq b$, denn es gilt

$$[a, c[\; = \;]a - 1, c[\; \cap \; A$$

für die offene Menge $]a - 1, c[$. Wegen $B \cap \mathbb{R} = B$ ist weiter jedes $B \subseteq \mathbb{R}$ offen in sich selbst.

Ist A offen, so ist P genau dann offen in A, wenn P offen ist. Der Leser, dem die relativen Begriffe zunächst unsympathisch erscheinen, kann also für das folgende zur Vereinfachung annehmen, daß der Definitionsbereich A einer Funktion offen ist. In diesem Fall kann überall der relative Zusatz „in A“ gestrichen werden.

Der folgende Satz ist nun nichts weiter als eine Umformulierung des Satzes über die ε-δ-Stetigkeit:

Satz *(Umgebungs-Formulierung der Stetigkeit)*
Sei $f : A \to \mathbb{R}$, und sei $a \in A$. Dann sind äquivalent:

(a) f ist stetig in a.

(b) Für jede offene Umgebung U von f(a) gibt es eine in A offene Umgebung V von a mit $f[V] \subseteq U$.

(c) Für jede Umgebung P von f(a) ist $f^{-1}[P]$ eine Umgebung von a in A.

Die Stetigkeit von f in einem Punkt a besagt also, daß die Urbilder von Umgebungen von f(a) Umgebungen von a sind. Der Umgebungsbegriff ist bei dieser punktweisen Urbild-Charakterisierung der Stetigkeit der angemessene. Ist zum Beispiel $f : \mathbb{R} \to \mathbb{R}$ die Funktion mit $f(x) = 1$ für $0 \leq x \leq 1$ und $f(x) = 0$ sonst, so ist f stetig im Punkt 1/2. Dagegen ist aber das Intervall $[0, 1]$ das Urbild von $U_\varepsilon(1)$ für alle $0 < \varepsilon < 1$. Die Urbilder offener Umgebungen von $f(1/2) = 1$ sind hier also nicht notwendig offen. Für eine Charakterisierung der Stetigkeit in jedem Punkt genügen allerdings die offenen Mengen:

Satz *(Urbild-Formulierung der Stetigkeit)*
Sei $f : A \to \mathbb{R}$. Dann sind äquivalent:

(a) f ist stetig.

(b) Für alle offenen $U \subseteq \mathbb{R}$ ist $f^{-1}[U]$ offen in A.

Beweis
(b) $\curvearrowright$ (a):
Aus (b) folgt, daß für alle $a \in A$ die Urbilder von Umgebungen von f(a) Umgebungen von a in A sind. Damit ist f stetig nach obigem Satz.

(a) $\curvearrowright$ (b):
Sei $U \subseteq \mathbb{R}$ offen, und sei $V = f^{-1}[U]$. Wir zeigen, daß V offen in A ist. Sei hierzu $a \in V$ beliebig. Wir zeigen, daß ein $\varepsilon > 0$ existiert mit

$$U_\varepsilon(a) \cap A \subseteq V.$$

Annahme nicht. Dann existiert für alle $n \in \mathbb{N}$ ein $x_n \in U_{1/2^n}(a) \cap A$ mit $x_n \notin V$. Dann konvergiert aber $\langle x_n \mid n \in \mathbb{N} \rangle$ gegen a, und aufgrund der Stetigkeit von f in a konvergiert dann $\langle f(x_n) \mid n \in \mathbb{N} \rangle$ gegen f(a). Wegen $f(a) \in U$ und U offen gibt es dann aber ein n mit $f(x_n) \in U$. Dann ist aber $x_n \in f^{-1}[U] = V$, *Widerspruch*.

Hat eine Funktion einen offenen Definitionsbereich, so haben wir also folgende bestechend einfache Formulierung der Stetigkeit einer Funktion in jedem Punkt:

Die Urbilder von offenen Mengen sind offen.

In dieser Form wird die Stetigkeit in der Mathematik heute in jedem abstrakten geometrischen Kontext definiert, in dem gewisse Mengen als „offen" ausgezeichnet sind:

Ist X eine Menge und $\mathcal{U}$ ein System von Teilmengen von X mit $\emptyset, X \in \mathcal{U}$, so heißt $(X, \mathcal{U})$ ein *topologischer Raum* und $\mathcal{U}$ eine *Topologie* auf X, falls $\mathcal{U}$ abgeschlossen unter endlichen Durchschnitten und beliebigen Vereinigungen ist. Die Elemente von $\mathcal{U}$ heißen dann die *offenen Mengen* des Raumes, und ein $P \subseteq X$ heißt eine *Umgebung* eines Punktes $x \in X$, falls es ein offenes U gibt mit $x \in U$ und $U \subseteq P$. Sind $(X, \mathcal{U})$ und $(Y, \mathcal{V})$ topologische Räume, so heißt ein $f : X \to Y$ *stetig*, falls $f^{-1}[V] \in \mathcal{U}$ gilt für alle $V \in \mathcal{V}$. Für lokale Betrachtungen der Stetigkeit in einem Punkt muß aber, wie wir gesehen haben, „offen" durch „Umgebung" ersetzt werden: Ein $f : X \to Y$ heißt *stetig in* x, falls das Urbild jeder Umgebung von f(x) eine Umgebung von x ist.

Auch die relativen Begriffe finden wir hier wieder: Ist $(X, \mathcal{U})$ ein topologischer Raum und $A \subseteq X$, so ist $(A, \mathcal{U}_A)$ ein topologischer Raum, wobei die sog. Relativtopologie $\mathcal{U}_A$ definiert ist durch $\mathcal{U}_A = \{ U \cap A \mid U \in \mathcal{U} \}$. Die Menge A erbt in diesem Sinne eine Topologie von X. Speziell erbt jedes $A \subseteq \mathbb{R}$ die Topologie der offenen Mengen, und damit haben wir dann obige Charakterisierung der Stetigkeit einer Funktion $f : A \to \mathbb{R}$ wiedergefunden: f ist genau dann stetig im Sinne der Folgenstetigkeit, wenn $f : A \to \mathbb{R}$ eine stetige Abbildung zwischen den topologischen Räumen $(A, \mathcal{U}_A)$ und $(\mathbb{R}, \mathcal{U})$ ist.

Von der Folgenstetigkeit gelangten wir zur ε-δ-Stetigkeit. Eine auf den ersten Blick etwas aufgeblasene Umformulierung führte dann zur Formulierung der Stetigkeit mit Hilfe von offenen Teilmengen von $\mathbb{R}$ und Umgebungen von Punkten. Die Struktureigenschaften der offenen Mengen führten schließlich zum Begriff des topologischen Raumes und einem allgemeinen Stetigkeitsbegriff, der stetige Abbildungen als gewisse strukturerhaltende Abbildungen kennzeichnet. Das ist ein langer Weg der Abstraktion und Verallgemeinerung, und dieser Zwischenabschnitt ist eher als Ausblick gedacht denn als Einführung. Der Leser möge sich zudem vor Augen halten, daß die präzise Formulierung des Stetigkeitsbegriffs erst im 19. Jahrhundert gegeben worden ist, während man seit

Newton und Leibniz Differential- und Integralrechnung betrieb. Und danach vergingen noch einmal einige Jahrzehnte, bis im frühen 20. Jahrhundert die abstrakten topologischen Begriffe eingeführt wurden.

Metrische Vollständigkeit

Für Leser, die die Konstruktion und Charakterisierung von $\mathbb{R}$ als linear vollständig angeordneten Körper im zweiten Abschnitt verfolgt haben, werfen wir nun noch einmal einen Blick auf die Konvergenz von Folgen in beliebigen angeordneten Körpern. Unsere Verdichtungs- und Konvergenzbegriffe können wir nämlich allgemein definieren:

Definition *(Cauchyfolge, konvergente Folge, Grenzwert)*
Sei K ein angeordneter Körper, und sei $\langle x_n \mid n \in \mathbb{N}\rangle$ eine Folge in K.

(a) $\langle x_n \mid n \in \mathbb{N}\rangle$ heißt eine *Cauchyfolge* in K, falls gilt:

$$\forall \varepsilon \in K^+ \, \exists n_0 \, \forall n, m \geq n_0 \; |x_n - x_m| < \varepsilon. \qquad \textit{(Cauchybedingung)}$$

(b) $\langle x_n \mid n \in \mathbb{N}\rangle$ *konvergiert* in K, falls ein $x \in K$ existiert, sodaß gilt:

$$\forall \varepsilon \in K^+ \, \exists n_0 \, \forall n \geq n_0 \; |x_n - x| < \varepsilon. \qquad \textit{(Konvergenzbedingung)}$$

In diesem Fall heißt dann x ein *Grenzwert* der Folge.

Wir schreiben wieder $\lim_{n \to \infty} x_n$ oder $\lim_{n \in \mathbb{N}} x_n$ für den eindeutig bestimmten Grenzwert einer konvergenten Folge.
Nun definieren wir:

Definition *(metrische Vollständigkeit)*
Ein angeordneter Körper K heißt *metrisch vollständig*, falls gilt:

Jede Cauchyfolge in K konvergiert in K. *(metrische Vollständigkeit)*

Die Bezeichnung als metrische Vollständigkeit rührt daher, daß wir in den Definitionen einer Cauchyfolge und eines Grenzwertes den Abstand $|x_n - x_m|$ von x_n und x_m bzw. den Abstand $|x - x_n|$ von x und x_n betrachten.
Wir haben oben gezeigt, daß der angeordnete Körper $\mathbb{R}$ metrisch vollständig ist. Der Beweis benutzt außer der linearen Vollständigkeit keine speziellen Eigenschaften von $\mathbb{R}$. Er zeigt:

Satz *(lineare Vollständigkeit impliziert metrische Vollständigkeit)*
Jedes arithmetische Kontinuum ist metrisch vollständig.

Beim Versuch, die Umkehrung dieses Satzes zu beweisen, bleiben wir stecken. Es zeigt sich aber, daß der Beweis durchgeführt werden kann, wenn wir die Anordnung als archimedisch voraussetzen:

Satz *(Vollständigkeit bei archimedischer Anordnung)*
Sei $(K, +, \cdot, \leq)$ ein metrisch vollständiger und archimedisch angeordneter Körper. Dann ist $(K, +, \cdot, \leq)$ ein arithmetisches Kontinuum.

Beweis
Sei X eine nichtleere und nach oben beschränkte Teilmenge von K.
Sei $x_0 \in X$ beliebig. Für alle $n \geq 1$ existiert nach dem archimedischen Axiom

k_n = „das kleinste k mit: $x_0 + k \cdot 1/n$ ist eine obere Schranke von X“.

Sei $x_n = x_0 + k_n/n$ für alle $n \geq 1$. Dann ist $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Cauchyfolge. Denn sei $\varepsilon > 0$. Nach dem archimedischen Axiom gibt es ein $n_0 \geq 1$ mit $1/n_0 < \varepsilon$. Dann gilt für alle $n \geq n_0$, daß $|x_n - x_{n_0}| \leq 1/n_0 < \varepsilon$. Also existiert

$x^* = \lim_{n \in \mathbb{N}} x_n$.

– Dann gilt $x^* = \sup(X)$, wie man leicht nachprüft.

Für einen angeordneten Körper ist also „(linear) vollständig“ äquivalent zu „metrisch vollständig und archimedisch“. Die reellen Zahlen sind damit bis auf Isomorphie der einzige angeordnete Körper, in dem jede Cauchyfolge konvergiert und in dem für alle $x > 0$ die Folge $x, 2x, 3x, \ldots, nx, \ldots$ jedes $y > 0$ übertrifft. Wir können hier auf die archimedische Anordnung nicht verzichten. Die Konstruktion von metrisch vollständigen, aber nicht archimedisch angeordneten Körpern ist jedoch eine nichttriviale Angelegenheit.

Übungen

Übung 1 *(Konvergente Folgen, I)*
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Cauchyfolge in $\mathbb{R}$. Zeigen Sie, daß $X = \{ x_n \mid n \in \mathbb{N} \}$ beschränkt ist.

Übung 2 *(Konvergente Folgen, II)*
Zeigen Sie, daß jede konvergente Folge in $\mathbb{R}$ eine Cauchyfolge ist.

Übung 3 *(Konvergente Folgen, III)*
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine konvergente Folge in $\mathbb{R}$. Zeigen Sie, daß der Grenzwert dieser Folge eindeutig bestimmt ist.

Übung 4 *(Konvergente Folgen, IV)*
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine Folge in $\mathbb{R}$. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) $\langle x_n \mid n \in \mathbb{N} \rangle$ ist eine Cauchyfolge.

(b) $\forall \varepsilon > 0 \; \exists n_0 \; \forall n \geq n_0 \; |x_n - x_{n_0}| < \varepsilon$.

Übung 5 (Konvergente Folgen, V)
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine monoton wachsende beschränkte Folge in $\mathbb{R}$.
Zeigen Sie, daß $\lim_{n \in \mathbb{N}} x_n = \sup(\{ x_n \mid n \in \mathbb{N} \})$.

Übung 6 (Konvergente Folgen, VI)
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine beschränkte Folge in $\mathbb{R}$. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) $\langle x_n \mid n \in \mathbb{N} \rangle$ konvergiert.

(b) $\liminf_{n \in \mathbb{N}} x_n = \limsup_{n \in \mathbb{N}} x_n$.

Zeigen Sie weiter, daß in diesem Fall gilt:

$\liminf_{n \in \mathbb{N}} x_n = \limsup_{n \in \mathbb{N}} x_n = \lim_{n \in \mathbb{N}} x_n$.

Übung 7 (Konvergente Folgen, VII)
Seien $\langle x_n \mid n \in \mathbb{N} \rangle$ und $\langle y_n \mid n \in \mathbb{N} \rangle$ konvergente Folgen in $\mathbb{R}$, und seien $x = \lim_n x_n$ und $y = \lim_n y_n$. Zeigen Sie, daß $\langle x_n + y_n \mid n \in \mathbb{N} \rangle$ gegen $x + y$ und $\langle x_n \cdot y_n \mid n \in \mathbb{N} \rangle$ gegen $x \cdot y$ konvergiert.
(Allgemeiner gilt dies für Folgen in einem angeordneten Körper.)

Übung 8 (Häufungspunkte, I)
Für jedes $X \subseteq \mathbb{R}$ sei X' die Menge der Häufungspunkte von X.
Zeigen oder widerlegen Sie, daß für alle $X, Y \subseteq \mathbb{R}$ gilt:

(i) $(X \cap Y)' = X' \cap Y'$, (ii) $(X \cup Y)' = X' \cup Y'$.

Übung 9 (Häufungspunkte, II)
Zeigen Sie, daß die folgenden Aussagen für alle $X \subseteq \mathbb{R}$ und alle $x \in \mathbb{R}$ äquivalent sind:

(a) x ist ein Häufungspunkt von X.

(b) Es gibt eine Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ in X mit $x_n \neq x_m$ für alle $n \neq m$ und $\lim_n x_n = x$.

Übung 10 (Häufungspunkte, III)
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine beschränkte Folge in $\mathbb{R}$ derart, daß $x_n \neq x_m$ für alle $n \neq m$ gilt. Zeigen Sie, daß der Limes Inferior der Folge der kleinste Häufungspunkt von $X = \{ x_n \mid n \in \mathbb{N} \}$ ist.
Zeigen Sie weiter, daß diese Aussage nicht immer gilt, wenn wir nur fordern, daß X unendlich ist.

Übung 11 (Häufungspunkte, IV)
Sei $\langle x_n \mid n \in \mathbb{N} \rangle$ eine beschränkte Folge reeller Zahlen derart, daß $x_n \neq x_m$ für alle $n \neq m$ gilt. Zeigen Sie, daß für alle $x \in \mathbb{R}$ äquivalent sind:

(a) $\lim_{n \to \infty} x_n = x$.

(b) x ist der einzige Häufungspunkt von $\{ x_n \mid n \in \mathbb{N} \}$.

Übung 12 *(Reihen, I)*
Seien $\Sigma_{n\in\mathbb{N}}\, x_n$ und $\Sigma_{n\in\mathbb{N}}\, y_n$ konvergente Reihen. Zeigen Sie:

$\Sigma_{n\in\mathbb{N}}\, x_n + y_n \;=\; (\Sigma_{n\in\mathbb{N}}\, x_n) \;+\; (\Sigma_{n\in\mathbb{N}}\, y_n)$.

Übung 13 *(Reihen, II)*
Zeigen Sie, daß $\Sigma_{n\in\mathbb{N}}\, 1/(n+1) = \infty$. *(Divergenz der harmonischen Reihe)*

[Fassen Sie Summanden der Reihe in genügend lange aufeinanderfolgende Blöcke zusammen, sodaß die Summe jedes Blocks größergleich 1/2 ist.]

Übung 14 *(Reihen, III)*
Sei $x \in \mathbb{R}$ mit $|x| < 1$. Zeigen Sie, daß

$\Sigma_{n\in\mathbb{N}}\, x^n \;=\; 1/(1-x)$. *(Berechnung der geometrischen Reihe für $|x| < 1$)*

[Zeigen Sie, daß für alle $n \in \mathbb{N}$ gilt: $\Sigma_{0\le i\le n}\, x^i = (1 - x^{n+1})/(1-x)$.]

Übung 15 *(Reihen, IV)*
Zeigen Sie, daß $\Sigma_{n\ge 1}\, 1/(n(n+1)) = 1$.

[Zeigen Sie, daß $\Sigma_{1\le i\le n}\, 1/(i(i+1)) = n/(n+1)$ für alle $n \ge 1$.]

Übung 16 *(Reihen, V)*
Sei $\Sigma_{n\in\mathbb{N}}\, x_n$ eine konvergente Reihe. Zeigen Sie, daß das arithmetische Mittel der Partialsummen gegen die Summe der Reihe konvergiert, d.h. es gilt

$\lim_{n\in\mathbb{N}}\, (s_0 + \ldots + s_n)/(n+1) \;=\; s$,

wobei $s_n = \Sigma_{i\le n}\, x_i$ für alle $n \ge 1$ und $s = \Sigma_{n\in\mathbb{N}}\, x_n$.

Übung 17 *(Stetige Funktionen, I)*
Zeigen Sie:

(i) Ist $f : A \to \mathbb{R}$ konstant, so ist f stetig.

(ii) Für alle $A \subseteq \mathbb{R}$ ist die Identität $\mathrm{id}_A : A \to \mathbb{R}$ stetig.

(iii) Ist $f : A \to B$ stetig in a und $g : B \to \mathbb{R}$ stetig in f(a), so ist $h = g \circ f$ stetig in a.

Übung 18 *(Stetige Funktionen, II)*
Sei $f : [0,1] \to [0,1]$ stetig. Zeigen Sie:

Es gibt ein $x \in [0,1]$ mit $f(x) = x$.

Übung 19 *(Stetige Funktionen, III)*
Sei $f : \mathbb{R} \to \mathbb{R}$ eine stetige Funktion und für alle $\varepsilon > 0$ gebe es $x, y \in\,]-\varepsilon, \varepsilon[$ mit $f(x) > 0$ und $f(y) < 0$. Zeigen Sie, daß $f(0) = 0$.

Übung 20 (Offene Mengen und Umgebungen, I)
Sei $U \subseteq \mathbb{R}$. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) U ist offen.

(b) Es gibt ein System $\mathcal{S}$ von offenen Intervallen mit $U = \bigcup \mathcal{S}$.

Übung 21 (Offene Mengen und Umgebungen, II)
Sei $\mathcal{U}$ ein System offener Teilmengen in $\mathbb{R}$. Zeigen Sie:

(i) $\bigcup \mathcal{U}$ ist offen.

(ii) Ist $\mathcal{U}$ endlich, so ist $\bigcap \mathcal{U}$ offen (mit $\bigcap \varnothing = \mathbb{R}$).

Übung 22 (Offene Mengen und Umgebungen, III)
Ein $C \subseteq \mathbb{R}$ heißt *abgeschlossen*, falls $\mathbb{R} - C$ offen ist. Weiter heißt ein $D \subseteq A$ abgeschlossen *in* A, falls es ein abgeschlossenes C gibt mit $D = A \cap C$. Zeigen Sie, daß für eine Funktion $f : A \to \mathbb{R}$ die folgenden Aussagen äquivalent sind:

(a) f ist stetig.

(b) Für alle abgeschlossenen $C \subseteq \mathbb{R}$ ist $f^{-1}[C]$ abgeschlossen in A.

Übung 23 (Metrische Vollständigkeit, I)
Sei $(K, +, \cdot, \leq)$ ein angeordneter Körper. Dann sind äquivalent:

(a) $(K, +, \cdot, \leq)$ ist ein arithmetisches Kontinuum, d. h. $(K, \leq)$ ist eine vollständige lineare Ordnung.

(b) Jede monoton wachsende beschränkte Folge $\langle x_n \mid n \in \mathbb{N} \rangle$ in K konvergiert in K.

Übung 24 (Metrische Vollständigkeit, II)
Geben Sie einen detaillierten Beweis für die Aussage „$x^* = \sup(X)$“ im Beweis des Satzes, daß die metrische Vollständigkeit zusammen mit dem archimedischen Axiom die lineare Vollständigkeit impliziert.

3. Matrizen

Matrizen sind rechteckige Gebilde aus Zahlen. Wir notieren eine Matrix A mit m Zeilen und n Spalten in der Form $\langle a_{i,j} \mid 1 \le i \le m, 1 \le j \le n \rangle$ oder kurz als $A = (a_{i,j})_{i,j}$ oder $A = (a_{ij})_{ij}$, wenn n und m fest gewählt sind. Die übliche graphische Darstellung einer derartigen $(m \times n)$-Matrix A ist:

$$\begin{array}{cccc} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{array}$$

Matrizen treten in vielen verschiedenen Kontexten auf und man kann sie aus ganz unterschiedlichen Blickwinkeln betrachten: Sie vereinfachen ganz ohne jeden theoretischen Hintergrund die Notationen beim Lösen von linearen Gleichungssystemen, sie repräsentieren lineare Abbildungen zwischen Vektorräumen und sie bilden einen ansprechenden Rahmen zur Untersuchung von beliebigen endlichen Relationen. Diese Aspekte wollen wir in diesem Kapitel vorstellen.

Vektoren

Im folgenden sei $K = \mathbb{Q}$ oder $K = \mathbb{R}$. Der Leser, der mit dem Körperbegriff vertraut ist, kann allgemeiner für K einen beliebigen endlichen oder unendlichen Körper einsetzen, unsere Überlegungen bleiben dann unverändert gültig.

Weiter sei $n \ge 1$ eine festgewählte natürliche Zahl. Wir betrachten die Menge

$$K^n = \{ (x_1, \dots, x_n) \mid x_i \in K \text{ für alle } 1 \le i \le n \}$$

aller n-Tupel mit Einträgen in K. Die Elemente von K^n nennen wir auch *Vektoren* und die Elemente von K auch *Skalare*. Ist $x \in K^n$ ein Vektor, so sei x(i) seine i-te *Komponente* für alle $1 \le i \le n$. Ist also $x = (x_1, \dots, x_n)$, so ist $x(i) = x_i$ für alle i.

Der *Nullvektor* des K^n ist der Vektor x mit $x(i) = 0$ für alle i. Wir bezeichnen ihn oft wie die Null des Körpers K mit 0. Diese Mehrfachbedeutung des Nullzeichens ist in der Regel unproblematisch.

Für alle $1 \le i \le n$ ist der i-te *Einheitsvektor* e_i des K^n definiert durch

$$e_i(j) = \begin{cases} 1, & \textit{falls } i = j, \\ 0 & \textit{sonst.} \end{cases}$$

Für $n = 3$ gilt also z. B. $e_2 = (0, 1, 0)$ und für $n = 4$ ist $e_2 = (0, 1, 0, 0)$.

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_10, © Springer-Verlag Berlin Heidelberg 2010

Für Vektoren $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in K^n$ und Skalare $\alpha \in K$ führen wir die folgenden arithmetischen Operationen ein:

$x + y = (x_1 + y_1, \ldots, x_n + y_n)$, *(Vektoraddition)*

$\alpha x = (\alpha x_1, \ldots, \alpha x_n)$. *(Skalarmultiplikation)*

Der K^n bildet mit diesen Operationen einen sog. Vektorraum, auf dessen allgemeine Definition wir hier verzichten können.

Für jeden Vektor $x \in K^n$ gilt

$x = \Sigma_{1 \le i \le n} x(i)\, e_i$.

So ist z. B. $(2, 3, -2) = 2 \cdot (1, 0, 0) + 3 \cdot (0, 1, 0) - 2 \cdot (0, 0, 1)$.

Lineare Gleichungssysteme

Wir betrachten ein lineares Gleichungssystem mit m Gleichungen und n Variablen, „Unbestimmten" oder „Unbekannten" $x_1, \ldots, x_n$:

$a_{1,1} x_1 + a_{1,2} x_2 + \ldots + a_{1,n} x_n = b_1$, *(Gleichung 1)*

$a_{2,1} x_1 + a_{2,2} x_2 + \ldots + a_{2,n} x_n = b_2$, *(Gleichung 2)*

$\ldots$ $\ldots$

$a_{m,1} x_1 + a_{m,2} x_2 + \ldots + a_{m,n} x_n = b_m$. *(Gleichung m)*

Hierbei sind alle $a_{i,j}$ und alle b_i gegebene Elemente von K. Die Skalare $a_{i,j}$ heißen die *Koeffizienten* des Gleichungssystems, und $b = (b_1, \ldots, b_m)$ heißt der *rechte Vektor* des Systems. Das Gleichungssystem heißt *homogen*, falls der rechte Vektor der Nullvektor ist.

Das Gleichungssystem heißt *lösbar*, falls $x_1, \ldots, x_n \in K$ existieren, sodaß alle Gleichungen des Systems simultan erfüllt sind. Der Vektor $(x_1, \ldots, x_n) \in K^n$ heißt dann *eine Lösung* des Systems. Die Menge

$L = \{ (x_1, \ldots, x_n) \in K^n \mid (x_1, \ldots, x_n)$ ist eine Lösung des Systems $\}$

heißt die *Lösungsmenge* oder der *Lösungsraum* des Systems.

Wir notieren die Koeffizienten des Systems als $(m \times n)$-Matrix $A = (a_{ij})_{ij}$. Formal ist eine derartige Matrix eine Abbildung $A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to K$, und wir schreiben dann a_{ij} für die Funktionswerte A(i, j).

Oft ist es auch nützlich, den rechten Vektor in die Koeffizientenmatrix zu integrieren, indem wir der Koeffizientenmatrix eine weitere aus $b_1, \ldots, b_m$ gebildete Spalte hinzufügen. Wir bezeichnen diese Matrix mit (A, b) und nennen sie die *erweiterte Koeffizientenmatrix*. Sie hat m Zeilen und n + 1 Spalten. Die i-te Zeile dieser Matrix ist $(a_{i,1}, \ldots, a_{i,n}, b_i)$.

Ein Gleichungssystem mit erweiterter Koeffizientenmatrix (A, b) und Unbekannten $x_1, \ldots, x_n$ notieren wir oft auch in der kompakten Form

$A x = b$. *(Notation für lineare Gleichungssysteme)*

Als nächstes wollen wir einer Koeffizientenmatrix A eine Funktion zuordnen, die die Berechnungen durchführt, die entstehen, wenn wir in einem Gleichungssystem A x = b für die Unbestimmten Skalare von K einsetzen.

Definition *(von einer Matrix induzierte Abbildung)*
Sei $A = (a_{ij})_{ij}$ eine $(m \times n)$-Matrix in K. Dann definieren wir eine Funktion $f_A : K^n \to K^m$ durch:

$$f_A(x_1, \ldots, x_n)(i) = \Sigma_{1 \le j \le n}\, a_{i,j}\, x_j = a_{i,1}\, x_1 + a_{i,2}\, x_2 + \ldots + a_{i,n}\, x_n$$

für alle $(x_1, \ldots, x_n) \in K^n$ und alle $1 \le i \le m$.

Die Funktion f_A heißt die von der Matrix A *induzierte* oder die der Matrix A *zugeordnete* Abbildung.

Die Abbildung f_A wertet für $x_1, \ldots, x_n \in K$ alle m Terme auf der linken Seite eines Gleichungssystems A x = b aus. Die Ergebnisse dieser Auswertungen fassen wir zu einem Vektor $f_A(x)$ der Länge m zusammen. Die Frage der Lösbarkeit des Gleichungssystems A x = b lautet nun einfach:

Gilt $b \in rng(f_A)$?

Weiter gilt:

$$L = \{ x \in K^n \mid f_A(x) = b \},$$

d. h. der Lösungsraum ist das Urbild des Vektors b unter der Abbildung f_A.

Wir wollen nun die geometrische Struktur des Lösungsraumes L von A x = b noch genauer ergründen. Hierzu betrachten wir das zugeordnete homogene Gleichungssystem Ax = 0 und setzen

$$L_0 = \{ x \in K^n \mid A\,x = 0 \}.$$

Für die Lösungsmenge L_0 gilt:

Satz *(Struktur des Lösungsraumes eines homogenen Systems)*

(i) $0 \in L_0$,

(ii) $x + y \in L_0$ für alle $x, y \in L_0$,

(iii) $\alpha\, x \in L_0$ für alle $x \in L_0$ und alle $\alpha \in K$.

Beweis
Sei $f = f_A$ die zugeordnete Abbildung. Wir rechnen:

$$f(0) = f(0 \cdot 0) = 0 \cdot f(0) = 0,$$

$$f(x+y) = f(x) + f(y) = 0 + 0 = 0 \quad \text{für alle } x, y \in L_0,$$

$$f(\alpha\, x) = \alpha\, f(x) = \alpha\, 0 = 0 \quad \text{für alle } x \in L_0 \text{ und alle } \alpha \in K.$$

– Aus $L_0 = \{ x \in K^n \mid f(x) = 0 \}$ folgen die Behauptungen.

Die aufgetauchten Eigenschaften von f werden wir im Abschnitt über „Lineare Abbildungen“ studieren. Hier halten wir die Strukturaussagen (i) – (iii) begrifflich fest:

Definition *(Unterraum)*
Ein $U \subseteq K^n$ heißt ein *Unterraum* oder *Untervektorraum* des K^n, falls $0 \in U$ gilt und U abgeschlossen unter Vektoraddition und Skalarmultiplikation ist, d. h. für alle $x, y \in U$ und alle $\alpha \in K$ gilt $x + y \in U$ und $\alpha x \in U$.

Der Satz besagt also, daß die Lösungsmenge eines homogenen linearen Gleichungssystems ein Unterraum ist. Umgekehrt zeigt man in der linearen Algebra, daß auch jeder Unterraum als Lösungsraum eines homogenen Gleichungssystems auftaucht.

Für die Lösungsmenge L von $A x = b$ gilt die folgende Darstellung, deren Beweis dem Leser zur Übung überlassen bleiben kann:

Satz *(Struktur des Lösungsraumes eines inhomogenen Systems)*
Sei $y = (y_1, \ldots, y_n) \in L$. Dann gilt

$$L = L_0 + y, \quad \text{wobei } L_0 + y = \{ x + y \mid x \in L_0 \}.$$

Kennen wir also die Lösung L_0 des zugeordneten homogenen Systems, so liefert uns jede spezielle Lösung von $A x = b$ alle Lösungen von $A x = b$.

Wir definieren:

Definition *(affiner Unterraum)*
Ein $X \subseteq K^n$ heißt ein *affiner Unterraum* des K^n, falls ein Unterraum U des K^n und ein $y \in K^n$ existieren mit

$$X = U + y = \{ x + y \mid x \in U \}.$$

Damit ist die Lösungsmenge eines linearen Gleichungssystems also ein affiner Unterraum des K^n.

Affine Unterräume sind i. a. nicht mehr abgeschlossen unter der Addition von Vektoren. Es gilt aber eine bemerkenswerte andere Abschlußeigenschaft.

Satz *(Abschlußeigenschaft eines affinen Unterraums)*
Sei $X = U + y$ ein affiner Unterraum, und seien $x_1, \ldots, x_k \in X$.
Zudem seien $\alpha_1, \ldots, \alpha_k \in K$ mit $\Sigma_{1 \le i \le k} \alpha_i = 1$. Dann gilt $\Sigma_{1 \le i \le k} \alpha_i x_i \in X$.

Beweis
Sei $x_i = u_i + y$ für alle $1 \le i \le k$, mit u_i in U. Dann gilt:

$$\Sigma_{1 \le i \le k} \alpha_i x_i = \Sigma_{1 \le i \le k} \alpha_i (u_i + y) = (\Sigma_{1 \le i \le k} \alpha_i u_i) + 1\, y \in U + y = X.$$

Zwei Gleichungen und zwei Unbekannte

Bevor wir uns einem allgemeinen Lösungsverfahren zuwenden, betrachten wir lineare Gleichungssysteme mit zwei Gleichungen und zwei Unbekannten noch etwas genauer. Wir notieren ein derartiges System in der Form

$$a\,x + b\,y = e, \qquad \textit{(Gleichung 1)}$$

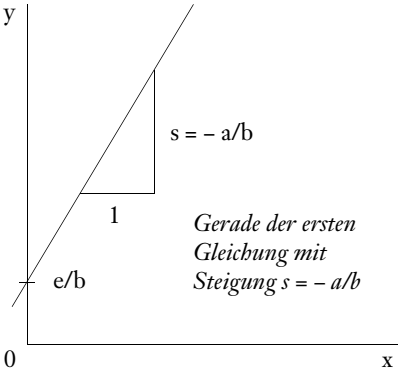
$$c\,x + d\,y = f. \qquad \textit{(Gleichung 2)}$$

mit Unbekannten x und y und vorgegebenen reellen Zahlen a, …, f. Wir nehmen der Einfachheit halber an, daß $b, d \neq 0$ gilt. Dann können wir die beiden Gleichungen schreiben als

$$y = -(a/b)\,x + e/b, \qquad \textit{(Geradengleichung 1)}$$

$$y = -(c/d)\,x + f/d. \qquad \textit{(Geradengleichung 2)}$$

Die beiden Gleichungen beschreiben also Geraden in der Ebene, und die Schnittpunkte dieser Geraden bilden die Lösungsmenge des Systems. Sind die Geraden nicht parallel, so existiert genau eine Lösung. Dies ist genau dann der Fall, wenn die Steigungen $-a/b$ und $-c/d$ der Geraden verschieden sind, d. h. falls $ad - bc \neq 0$. (Dieser Wert ist unabhängig von e und f.) Sind die Geraden parallel aber verschieden, so existiert keine Lösung. Andernfalls fallen die beiden Geraden zusammen und es gibt unendlich viele Lösungen.



Liegen die Steigungen der Geraden nahe beieinander, d. h. ist $|ad - bc|$ sehr klein, so treten bemerkenswerte Effekte auf. Wir betrachten hierzu folgendes Beispiel (aus einer Arbeit von William Kahan):

$$0{,}2161\,x + 0{,}1441\,y = 0{,}1440,$$

$$1{,}2969\,x + 0{,}8648\,y = 0{,}8642.$$

Die eindeutige Lösung dieses Gleichungssystems ist $(x, y) = (2, -2)$. Für $\bar{x} = 0{,}9911$ und $\bar{y} = -0{,}4870$, ergibt sich

$$0{,}2161\,\bar{x} + 0{,}1441\,\bar{y} = 0{,}1440 + 0{,}00000001,$$

$$1{,}2969\,\bar{x} + 0{,}8648\,\bar{y} = 0{,}8642 - 0{,}00000001,$$

d. h. der von $(2, -2)$ deutlich verschiedene Vektor $(\bar{x}, \bar{y})$ könnte als Lösung des Systems gelten, wenn wir mit weniger als acht Nachkommastellen Genauigkeit rechnen. Umgekehrt betrachtet: Ändern wir die rechte Seite (0,1440, 0,8642) minimal um $(10^{-8}, -10^{-8})$ ab, so springt die exakte Lösung von $(2, -2)$ zu $(\bar{x}, \bar{y})$. Man sagt in der Numerik, daß das Gleichungssystem *schlecht konditioniert* ist.

Das Gauß-Jordansche Eliminationsverfahren

Wir stellen nun ein Lösungsverfahren für unsere Gleichungssysteme vor. Der Verfahrenstyp wird oft als *Gauß-Jordansches Eliminationsverfahren* bezeichnet, seine Ursprünge reichen aber bis in die vorchristliche chinesische Mathematik zurück. Den Ausgangspunkt bildet die Beobachtung, daß sich der Lösungsraum eines Gleichungssystems

$$a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n = b_1, \qquad \textit{(Gleichung 1)}$$
$$a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,n}x_n = b_2, \qquad \textit{(Gleichung 2)}$$
$$\ldots \qquad\qquad \ldots$$
$$a_{m,1}x_1 + a_{m,2}x_2 + \ldots + a_{m,n}x_n = b_m. \qquad \textit{(Gleichung m)}$$

durch die folgenden vier elementaren Operationen nicht ändert:

$(E_i' = \alpha \cdot E_i)$	Multiplikation der i-ten Gleichung mit $\alpha \in K$, $\alpha \neq 0$.
$(E_i' = E_i + \alpha E_j)$	Addition des α-fachen der j-ten zur i-ten Gleichung, $i \neq j$.
$(\sigma_{i,j})$	Vertauschung der i-ten und der j-ten Gleichung.
$(\pi_{i,j})$	Vertauschung der i-ten und der j-ten Spalte des Systems (einschließlich der Variablen x_i und x_j).

Die ersten drei dieser Operationen können wir analog auch für Matrizen ausführen, indem wir „n-te Gleichung“ durch „n-te Zeile“ ersetzen. Bei der Spaltenvertauschung ist dagegen etwas Umsicht geboten, da in Matrizenschreibweise keine benannten Variablen mehr zur Verfügung stehen. Wir können aber Spaltenvertauschungen auch für Matrizen zulassen, wenn wir nebenbei mitschreiben, welche Vertauschungen wir durchgeführt haben. Dann geht keinerlei Information verloren und wir können Matrizen weiterhin als bequeme Notation für Gleichungssysteme ansehen.

Wir streben durch wiederholte Ausführung dieser Operationen folgende Form an:

Definition *(D_r-Form)*

Eine $(m \times n)$-Matrix A ist in *D_r-Form* für ein $0 \leq r \leq \min(m, n)$, falls gilt:

(i) $a_{i,i} = 1$ für alle $1 \leq i \leq r$,

(ii) $a_{i,j} = 0$ für alle $1 \leq i, j \leq r$ mit $i \neq j$,

(iii) $a_{i,j} = 0$ für alle i, j mit $r < i \leq m$.

In einer derartigen Matrix steht also links oben die $(r \times r)$-Diagonalmatrix D_r, deren Diagonale mit Einsen bestückt ist und die außerhalb ihrer Diagonalen nur Null-Einträge aufweist. Weiter sind alle Einträge in den Zeilen $r + 1, \ldots, m$ der Matrix gleich 0.

Der Lösungsraum eines Gleichungssystems in D_r-Form läßt sich nun aus der erweiterten Koeffizientenmatrix direkt ablesen:

Satz *(Lösungsraum für D_r-Formen)*
Sei $A\,x = b$ ein Gleichungssystem mit einer $(m \times n)$-Matrix $A = (a_{ij})_{ij}$ in D_r-Form. Sei L der Lösungsraum des Systems. Dann gilt:

Ist $b_i \neq 0$ für ein i mit $r < i \leq m$, so ist das System unlösbar.
Andernfalls ist der K^n-Vektor $b^* = (b_1, \ldots, b_r, 0, \ldots, 0) \in L$, und es gilt:

(a) Ist $r = n$, so ist $L = \{ b^* \}$.

(b) Ist $r < n$, so ist $L = L_0 + b^*$ mit dem folgenden Lösungsraum L_0 des homogenen Systems $A\,x = 0$:

$$L_0 = \{ \alpha_{r+1}\, y_{r+1} + \ldots + \alpha_n\, y_n \mid \alpha_i \in K \text{ für alle } r < i \leq n \},$$

wobei $y_i = (-a_{1i}, \ldots, -a_{ri}, 0, \ldots, 0) + e_i \in K^n$ für alle $r < i \leq n$.

Beweis
Wir zeigen, daß in (b) jede Lösung von $A\,x = 0$ ein Element von L_0 ist.
(Die anderen Behauptungen des Satzes sind unschwer einzusehen).
Sei also $(\alpha_1, \ldots, \alpha_n)$ eine beliebige Lösung von $A\,x = 0$. Dann gilt

$$\alpha_i + \Sigma_{r<j\leq n}\, a_{ij}\, \alpha_j = 0 \quad \text{für alle } 1 \leq i \leq r,$$

und damit ist

$$(\alpha_1, \ldots, \alpha_n) = \alpha_{r+1}\, y_{r+1} + \ldots + \alpha_n\, y_n \in L_0.$$

Wir zeigen nun, daß wir durch unsere vier Operationen eine D_r-Form erreichen können:

Satz *(Eliminationsverfahren)*
Sei $A\,x = b$ ein Gleichungssystem einer $(m \times n)$-Matrix $A = (a_{ij})_{ij}$.
Dann läßt sich (A, b) durch die vier elementaren Operationen in die Form (A', b') überführen mit A' in D_r-Form für ein $r \leq \min(m, n)$.
Ist $\pi = \pi_k \circ \ldots \circ \pi_1 : \{ 1, \ldots, n \} \rightarrow \{ 1, \ldots, n \}$ die Bijektion der durchgeführten Spaltenvertauschungen $\pi_1, \ldots, \pi_k$, so gilt

$$L = \{ \pi^{-1}(x) \mid x \in L' \},$$

für die Lösungsräume L von $A\,x = b$ und L' von $A'\,x = b'$.

Beweis
Wir konstruieren rekursiv (A^i, b^i) für $i = 0, 1, \ldots, r$, $r \leq \min(n, m)$.
Dabei ist die j-te Spalte von A^i der Vektor $e_j \in K^m$ für alle $1 \leq j \leq i$.

Zu Beginn sei $(A^0, b^0) = (A, b)$. Sei nun (A^i, b^i) konstruiert für ein $i \geq 0$.
Ist A^i in D_i-Form, so ist $(A', b') = (A^i, b^i)$ mit $r = i$ wie gewünscht.
Andernfalls existieren Indizes $i^*, j^* > i$ mit $A^i(i^*, j^*) \neq 0$.

Nach einer evtl. Zeilenvertauschung gefolgt von einer evtl. Spaltenvertauschung können wir also annehmen, daß $A^i(i+1, i+1) \neq 0$.
Durch Zeilenmultiplikation und -addition können wir nun die $(i+1)$-te Spalte der Matrix (A^i, b^i) in den Spaltenvektor e_{i+1} überführen. Mit der so entstehenden Matrix (A^{i+1}, b^{i+1}) wiederholen wir das Verfahren.

Das Verfahren produziert eine Matrix $(A', b') = (A^r, b^r)$ mit A' in D_r-Form. Eine Induktion nach $i \leq r$ zeigt, daß

$$L = \{ (\pi_{k(i)} \circ \ldots \circ \pi_1)^{-1}(x) \mid x \in K^n, A^i x = b^i \},$$

wobei $\pi_1, \ldots, \pi_{k(i)}$ die Spaltenvertauschungen sind, die bis zum i-ten Schritt durchgeführt wurden. Dies zeigt die Behauptung über den Lösungsraum. ∎

Wir können das im Beweis enthaltene Verfahren leicht deterministisch machen und so einen „Algorithmus von Gauß-Jordan“ erhalten, indem wir im Rekursionsschritt i* und j* geeignet minimal wählen. Gilt $A^i(i+1, i+1) \neq 0$, so sind gar keine Vertauschungen notwendig. Ist $A^i(i^*, i+1) \neq 0$ für ein $i^* > i$, so genügt eine Zeilenvertauschung. Wir können damit die Anzahl der Spaltenvertauschungen so gering wie möglich halten.
Das Verfahren zeigt für quadratische Matrizen:

Korollar *(Überführung in Diagonalform)*
Sei A eine $(n \times n)$-Matrix. Dann sind äquivalent:

(a) $f_A : K^n \to K^n$ ist bijektiv.

(b) A läßt sich mit Hilfe der ersten drei elementaren Operationen in die Diagonalmatrix D_n überführen.

Beweis
Die Aussage ist klar, falls wir alle vier Operationen zulassen. Wird aber eine Vertauschung der $(i+1)$-ten Spalte im Rekursionsschritt notwendig, so bleibt diese Spalte eine Spalte von

$$A^{i+1}, \ldots, A^r,$$

da alle ihre Komponenten ab der Stelle $i+1$ gleich 0 sind.
Folglich ist $A'(n, n) = 0$, d. h. es gilt $r < n$. Damit ist aber $A' x = b'$ unlösbar für alle b', deren n-te Komponente von 0 verschieden ist. Also ist $A x = b$ nicht für alle b lösbar, und damit ist f_A nicht bijektiv. ∎

Wir führen zur Illustration das Eliminationsverfahren nun noch für das folgende Gleichungssystem mit $K = \mathbb{Q}$ durch. Die erweiterte Koeffizientenmatrix (A^0, b^0) ist rechts notiert.

$$\begin{array}{rcrcrcrcl|c}
x_1 & - & x_2 & + & x_3 & & & = & -1 \\
x_1 & - & x_2 & & & - & x_4 & = & 1 \\
 & & & - & x_3 & - & x_4 & = & 2 \\
-2x_1 & + & 2x_2 & + & x_3 & & & = & -1
\end{array}
\qquad
\begin{array}{cccc|c}
1 & -1 & 1 & 0 & -1 \\
1 & -1 & 0 & -1 & 1 \\
0 & 0 & -1 & -1 & 2 \\
-2 & 2 & 1 & 0 & -1
\end{array}$$

Zuerst berechnen wir (A^1, b^1) durch „Ausräumen der ersten Spalte“:

$$\begin{array}{cccc|c}
1 & -1 & 1 & 0 & -1 \\
0 & 0 & -1 & -1 & 2 \\
0 & 0 & -1 & -1 & 2 \\
0 & 0 & 3 & 0 & -3
\end{array}
\qquad
\begin{array}{cccc|c}
1 & 0 & 1 & -1 & -1 \\
0 & -1 & -1 & 0 & 2 \\
0 & -1 & -1 & 0 & 2 \\
0 & 0 & 3 & 0 & -3
\end{array}$$

(A^1, b^1) Spaltentausch $\pi_{2,4}$

Wir notieren nun den Spaltentausch $\pi_{2,4}$ und berechnen dann (A^2, b^2):

$$\begin{array}{cccc|c}
1 & 0 & 1 & -1 & -1 \\
0 & 1 & 1 & 0 & -2 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3 & 0 & -3
\end{array}
\qquad
\begin{array}{cccc|c}
1 & 0 & 1 & -1 & -1 \\
0 & 1 & 1 & 0 & -2 \\
0 & 0 & 3 & 0 & -3 \\
0 & 0 & 0 & 0 & 0
\end{array}$$

(A^2, b^2) Zeilentausch $\sigma_{3,4}$

Wir tauschen die dritte und vierte Zeile und erhalten nun (A^3, b^3) in D_3-Form, indem wir die Koordinate (3, 3) auf 1 normieren und dann oberhalb dieser Koordinate „ausräumen“:

$$\begin{array}{cccc|c}
1 & 0 & 0 & -1 & 0 \\
0 & 1 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0
\end{array}$$

$(A^3, b^3) = (A', b')$

$L' = \{\alpha(1, 0, 0, 1) \mid \alpha \in \mathbb{Q}\} + (0, -1, -1, 0).$

$L = \{\alpha(1, 1, 0, 0) \mid \alpha \in \mathbb{Q}\} + (0, 0, -1, -1).$

Den Lösungsraum L' von $A^3 x = b^3$ lesen wir ab, und Vertauschung der Koordinaten 2 und 4 liefert den Lösungsraum L des ursprünglichen Systems $A x = b$.

Lineare Abbildungen

Wir wollen nun die Struktureigenschaften der einer $(m \times n)$-Matrix A zugeordneten Abbildung f_A noch genauer beschreiben.

Definition *(lineare Abbildung)*
Eine Funktion $f : K^n \to K^m$ heißt eine *lineare Abbildung*, falls für alle Vektoren $x, y \in K^n$ und alle Skalare $\alpha, \beta \in K$ gilt:

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y). \qquad \textit{(Linearitätsbedingung)}$$

Lineare Abbildungen respektieren also die Skalarmultiplikation und die Vektoraddition. Es zeigt sich nun, daß diese Abbildungen genau die Abbildungen der Form f_A sind:

Satz *(Hauptsatz über lineare Abbildungen)*
(a) Für jede $(m \times n)$-Matrix A ist $f_A : K^n \to K^m$ eine lineare Abbildung.
(b) Für jede lineare Abbildung $f : K^n \to K^m$ existiert eine eindeutig bestimmte $(m \times n)$-Matrix A mit $f = f_A$.

Beweis

zu (a):
Die Behauptung folgt leicht durch direktes Nachrechnen.

zu (b):
Wir definieren eine $(m \times n)$-Matrix $A = (a_{ij})_{ij}$ durch

$$a_{ij} = f(e_j)(i) \quad \text{für alle } 1 \le i \le m,\ 1 \le j \le n.$$

Die Spalten von A werden also mit den Bildern der Einheitsvektoren $e_1, \ldots, e_n$ unter der linearen Abbildung f gefüllt.

Für alle $x = (x_1, \ldots, x_n)$ gilt $x = \Sigma_{1 \le j \le n}\, x_j\, e_j$. Nach Linearität der Abbildung f gilt also:

$$f(x) = \Sigma_{1 \le j \le n}\, x_j\, f(e_j).$$

Also gilt für alle $1 \le i \le m$:

$$f(x)(i) = \Sigma_{1 \le j \le n}\, x_j\, f(e_j)(i) = \Sigma_{1 \le j \le n}\, x_j\, a_{ij} = f_A(x)(i).$$

Also gilt $f = f_A$. Dies zeigt die Existenzbehauptung.

Zum Beweis der Eindeutigkeit sei B eine weitere $(m \times n)$-Matrix mit $f = f_B$. Dann gilt $f_B(e_j) = f(e_j)$ für alle $1 \le j \le n$. Dann gilt aber

$$b_{ij} =_{\text{Definition von } f_B} f_B(e_j)(i) =_{f = f_B} f(e_j)(i) =_{\text{Definition von } a_{ij}} a_{ij} \quad \text{für alle } i, j.$$

Also ist $A = B$. ∎

Damit haben wir gezeigt, daß das Feld der durch die Linearitätsbedingung $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ bestimmten Abbildungen kleiner ist als man denken könnte: Jede lineare Abbildung ist die Berechnungsfunktion der linken Seite eines linearen Gleichungssystems. Umgekehrt sind alle diese Berechnungsfunktionen auch lineare Abbildungen.

Wir definieren:

Definition *(darstellende Matrix einer linearen Abbildung)*
Ist $f : K^n \to K^m$ eine lineare Abbildung. Dann heißt die eindeutige $(m \times n)$-Matrix A mit $f = f_A$ die *darstellende Matrix* von f.

Der Begriff der linearen Abbildung führt uns nun zu einer geistreichen Multiplikation von Matrizen.

Matrizenmultiplikation

Sind die Abbildungen $f : K^n \to K^m$ und $g : K^m \to K^r$ linear, so ist auch die Komposition $h = g \circ f$ linear von K^n nach K^r, wie man leicht nachrechnet. Seien A, B, C die darstellenden Matrizen von g, f und h, d. h. es gelte

$$g = f_A, \quad f = f_B, \quad h = f_C.$$

Es stellt sich nun die Frage, wie wir die $(r \times n)$-Matrix C aus der $(r \times m)$-Matrix A und der $(m \times n)$-Matrix B berechnen können. Eine Analyse dieser Fragestellung führt zur folgenden Definition der Multiplikation $A \cdot B$:

Definition *(Matrizenmultiplikation)*
Seien $A = (a_{ij})_{ij}$ eine $(r \times m)$-Matrix und $B = (b_{ij})_{ij}$ eine $(m \times n)$-Matrix.
Dann ist das *Produkt* $C = A \cdot B$, $C = (c_{ij})_{ij}$ definiert durch

$$c_{ij} = \Sigma_{1 \le k \le m} a_{ik} b_{kj} \quad \text{für alle } 1 \le i \le r,\ 1 \le j \le n.$$

Der Leser führe diese Multiplikation anhand einiger einfacher Matrizen in $\mathbb{Q}$ durch. Er wird dann sehen, warum die Matrizenmultiplikation mit der Merkregel „Zeile mal Spalte“ versehen ist.

Es gilt nun wie erwünscht:

Satz *(Matrizenmultiplikation und Komposition von linearen Abbildungen)*
Sind $f_A : K^m \to K^r$ und $f_B : K^n \to K^m$ linear, so ist $f_A \circ f_B = f_{A \cdot B}$.

Beweis
Für alle $1 \le j \le n$ gilt:

$$(f_A \circ f_B)(e_j) = f_A(f_B(e_j)) = f_A(b_{1j}, \ldots, b_{mj}) = f_A(\Sigma_{1 \le k \le m} b_{kj} e_k) =$$

$$\Sigma_{1 \le k \le m} b_{kj} f_A(e_k) = \Sigma_{1 \le k \le m} b_{kj} (a_{1k}, \ldots, a_{rk}).$$

Ist $C = (c_{ij})_{ij}$ die darstellende Matrix von $f_A \circ f_B$, so gilt also für alle i, j:

$c_{ij} = (f_A \circ f_B)(e_j)(i) = \Sigma_{1 \le k \le m} b_{kj} (a_{1k}, \ldots, a_{rk})(i) = \Sigma_{1 \le k \le m} b_{kj} a_{ik}$.

Also ist $C = A \cdot B$.

Als Korollar erhalten wir:

Korollar *(Assoziativität der Matrizenmultiplikation)*
Für alle geeigneten Matrizen A, B, C in K gilt $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Beweis
Es gilt $(f_A \circ f_B) \circ f_C = f_A \circ (f_B \circ f_C)$ nach Assoziativität der Komposition. Nach obigem Satz ist also $f_{(A \cdot B) \cdot C} = f_{A \cdot (B \cdot C)}$. Aufgrund der Eindeutigkeit der darstellenden Matrix ist dann aber $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Die Matrizenmultiplikation untermauert unsere Schreibweise $A x = b$ für ein Gleichungssystem. Denn wir können $x \in K^n$ als Spaltenvektor mit n Einträgen auffassen, d. h. als $(n \times 1)$-Matrix. Dann ist $A \cdot x$ das Produkt einer $(m \times n)$-Matrix und einer $(n \times 1)$-Matrix, und für alle x gilt $A \cdot x = b$ genau dann, wenn x eine Lösung des Gleichungssystems mit erweiterter Koeffizientenmatrix (A, b) ist.

Relationen und Matrizen

Sei R eine Relation auf der endlichen Menge $M = \{1, \ldots, n\}$. Visualisieren wir die Beziehungen $a R b$ für alle $a, b \in M$ durch einen Pfeil von a nach b, so ist folgende Frage natürlich:

Gibt es für einen Startpunkt $s \in M$ und ein Ziel $z \in M$ einen R-Zug von s nach z?

Dabei ist ein *R-Zug* der Länge $k \ge 1$ von s nach z eine Folge $a_0, \ldots, a_k$ in M mit $a_0 = a$, $a_k = z$ und $a_i R a_{i+1}$ für alle $i < k$.

Zur Untersuchung der Frage definieren wir für zwei Relationen S_1 und S_2 auf M ihre *Verknüpfung* $S_1 \circ S_2$ durch:

$S_1 \circ S_2 = \{ (a, c) \mid$ es gibt ein $b \in M$ mit $a S_1 b$ und $b S_2 c \}$.

Nun definieren wir rekursiv für alle $k \ge 1$:

$R^1 = R$, $R^{k+1} = R^k \circ R$ für alle k.

Eine Induktion nach $k \ge 1$ zeigt, daß für alle $a, b \in M$ und alle $k \ge 1$ gilt:

$a R^k b$ *gdw* „es gibt einen R-Zug der Länge k von a nach b".

Gibt es einen R-Zug von a nach b, so gibt es auch einen R-Zug von a nach b, dessen Länge kleinergleich der Anzahl n der Elemente von M ist (Herausschneiden von Schleifen). Damit lautet unsere Frage:

Gibt es ein $k \le n$ mit $s R^k z$?

Ist R reflexiv, so gilt, wie man leicht einsieht, $R = R^1 \subseteq \ldots \subseteq R^n$, und dann vereinfacht sich unsere Frage zu:

Gilt s R^n z ?

Die *transitive Hülle* R* von R ist definiert als die kleinste transitive Relation, die R als Teilmenge enthält. Nach unseren Überlegungen gilt also $R^* = R^n$, falls R reflexiv ist, und $R^* = \bigcup_{1 \le k \le n} R^k$ in jedem Fall (siehe Übungen).

Für eine effektive Berechnung der transitiven Hülle von R sind 0-1-Matrizen nützlich.

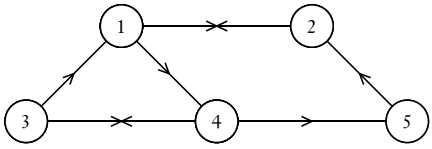
Definition *(darstellende Matrix von R)*

Die *darstellende Matrix* $A_R = (a_{ij})_{ij}$ von R ist definiert durch

$$a_{ij} = \begin{cases} 1, & \textit{falls } i \; R \; j, \\ 0, & \textit{sonst.} \end{cases}$$

So stellt zum Beispiel die Matrix auf der linken Seite die auf der rechten Seite visualisierte Relation R auf { 1, …, 5 } dar:

$$\begin{array}{ccccc} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{array}$$



Wir führen nun eine neue „logische" Multiplikation für 0-1-Matrizen ein, die die Verknüpfung $S_1 \circ S_2$ von Relationen beschreibt. Hierzu definieren wir eine Addition auf { 0, 1 } durch:

$0 + 0 = 0, \quad 0 + 1 = 1 + 0 = 1 + 1 = 1.$ *(logische 0-1-Addition)*

Diese Addition entspricht der Wahrheitswertfunktion des Junktors „oder". Die übliche Multiplikation auf { 0, 1 } entspricht bereits der Wahrheitswertfunktion des Junktors „und". Im folgenden bedeutet + immer die logische Addition. So ist z. B. (0, 1, 1, 0) + (1, 1, 0, 0) = (1, 1, 1, 0), usw.

Mit diesen Operationen definieren wir nun:

Definition *(logische Matrizenmultiplikation)*

Seien A, B (n × n)-Matrizen mit 0-1-Einträgen. Dann ist das *logische Produkt* A ∗ B von A und B die (n × n)-Matrix $C = (c_{ij})_{ij}$ mit

$$c_{ij} = \Sigma_{1 \le k \le n} a_{ik} b_{kj} = a_{i1} \cdot b_{1j} + \ldots + a_{in} \cdot b_{nj} \quad \text{für alle } 1 \le i, j \le n.$$

Diese Matrizenmultiplikation folgt also wieder der Regel „Zeile mal Spalte", lediglich die Arithmetik der Addition ist modifiziert.

Es gilt wieder $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ für alle 0-1-Matrizen A, B, C, wie man leicht nachrechnet. Die Assoziativität folgt aber auch aus dem folgenden Satz, der den Zusammenhang mit der Verknüpfung von Relationen herstellt.

Satz *(Verknüpfung von Relationen und logische Matrizenmultiplikation)*
Seien S_1, S_2 Relationen auf $M = \{ 1, \ldots, n \}$. Dann gilt

$A_{S_1 \circ S_2} = A_{S_1} * A_{S_2}$.

Beweis
Seien $A_{S_1} = (a_{ij})_{ij}$, $A_{S_2} = (b_{ij})_{ij}$, $A_{S_1} * A_{S_2} = (c_{ij})_{ij}$.
Dann gilt für alle i, j:

$c_{ij} = 1$ *gdw*

$\sum_{1 \le k \le n} a_{ik} b_{kj} = 1$ *gdw*

es gibt ein $1 \le k \le n$ mit $a_{ik} b_{kj} = 1$ *gdw*

es gibt ein $1 \le k \le n$ mit $a_{ik} = 1$ und $b_{kj} = 1$ *gdw*

es gibt ein $1 \le k \le n$ mit $(i, k) \in S_1$ und $(k, j) \in S_2$ *gdw*

– $(i, j) \in S_1 \circ S_2$.

Korollar
Es gilt $A_{R^k} = (A_R)^k$ für alle $k \ge 1$.

Wir können also die transitive Hülle R* von R berechnen, indem wir für $A = A_R$ der Reihe nach die Produkte

$A * A = A^2$, $A^2 * A = A^3$, …, A^n

berechnen. Ist R reflexiv, so ist A^n die darstellende Matrix A_{R^*} von R*. Allgemein ist A_{R^*} die punktweise logische Addition der Matrizen A^1, …, A^n. Diese Berechnung von R* benötigt aber unnötig viele Rechenschritte, und wir wollen deswegen noch einen effizienteren Algorithmus vorstellen.

Der Warshall-Algorithmus

Sei wieder R eine Relation auf $M = \{ 1, \ldots, n \}$. Wir definieren für alle $k \le n$ eine Relation $R^{(k)}$ auf M durch:

$i\ R^{(k)}\ j$ *falls* „es gibt einen R-Zug $i = a_0, \ldots, a_m = j$ mit $1 \le a_1, \ldots, a_{m-1} \le k$“

für alle $i, j \in M$. Offenbar gilt:

$R = R^{(0)} \subseteq R^{(1)} \subseteq \ldots \subseteq R^{(n)}$ = „die transitive Hülle von R“.

Bei diesen Relationen steht der obere Index also nicht für die Länge eines R-Zuges, sondern für die Menge $\{ 1, \ldots, k \}$ der Elemente von M, die für einen R-Zug verwendet werden dürfen.

Der sog. Algorithmus von Warshall berechnet nun in effektiver Weise die darstellenden Matrizen der Relationen $R^{(k)}$. Wir formulieren das Verfahren vorab und beweisen dann, was es leistet.

Algorithmus von Warshall

Sei A eine $(n \times n)$-Matrix mit 0-1-Einträgen. Wir definieren rekursiv $(n \times n)$-Matrizen $A^{(k)}$ mit 0-1-Einträgen für alle $k \leq n$. Das Ergebnis der Berechnung ist die Matrix $A^{(n)}$.

Zunächst setzen wir $A^{(0)} = A$. Sei nun $A^{(k)}$ definiert für ein $k < n$. Für alle $1 \leq i \leq n$ mit $A^{(k)}(i, k+1) = 1$ sei die i-te Zeile von $A^{(k+1)}$ die punktweise logische Summe der i-ten und $(k+1)$-ten Zeile von $A^{(k)}$. Die anderen Zeilen von $A^{(k+1)}$ übernehmen wir unverändert aus der Matrix $A^{(k)}$.

Ist die Matrix $A^{(k)}$ berechnet, so ist die $(k+1)$-te Spalte die „aktive" Spalte und die $(k+1)$-te Zeile die „aktive" Zeile. Die Einsen der aktiven Spalte markieren genau diejenigen Zeilen, auf die wir die aktive Zeile logisch addieren. Die aktive Spalte und Zeile bleiben dabei, wie man leicht sieht, unverändert.

Wir zeigen nun:

Satz *(Korrektheit des Warshall-Algorithmus)*

Sei $A = A_R$ die darstellende Matrix von R, und seien $A^{(0)}, \ldots, A^{(n)}$ die durch den Warshall-Algorithmus berechneten Matrizen. Dann gilt für alle $k \leq n$:

$A^{(k)}$ ist die darstellende Matrix $A_{R^{(k)}}$ der Relation $R^{(k)}$.

Beweis

Wir zeigen die Behauptung durch Induktion nach $k \leq n$.

Induktionsanfang $k = 0$:

Es gilt $A^{(0)} = A = A_R = A_{R^{(0)}}$.

Induktionsschritt von k nach $k+1$ für $k < n$:

Für alle $1 \leq i, j \leq n$ gilt:

$A^{(k+1)}(i,j) = 1$ *gdw*

$A^{(k)}(i,j) = 1$ *oder* $A^{(k)}(i,k+1) = 1$ und $A^{(k)}(k+1,j) = 1$ *gdw* I.V.

$(i,j) \in R^{(k)}$ *oder* $(i,k+1) \in R^{(k)}$ und $(k+1,j) \in R^{(k)}$ *gdw*

$(i,j) \in R^{(k)}$ *oder* es gibt $1 \leq a_1, \ldots, a_m, b_1, \ldots, b_{m'} \leq k$ mit $i, a_1, \ldots, a_m, k+1, b_1, \ldots, b_{m'}, j$ ist ein R-Zug *gdw*

$(i,j) \in R^{(k+1)}$,

wobei wir für die letzte Rückrichtung die Existenz von R-Zügen verwenden, die kein Element von M mehrfach durchlaufen. ∎

Damit ist also das Ergebnis $A^{(n)}$ des Warshall-Algorithmus die darstellende Matrix der transitiven Hülle von R.

Wir beobachten noch, daß wir uns die Matrizen $A^{(i)}$ während der Berechnung nicht merken müssen. Es genügt somit ein „Speicherplatz" für eine 0-1-Matrix.

Übungen

Übung 1 (Vektoren)
Sei $\bar{e}_1 = (0, 1, 1)$, $\bar{e}_2 = (1, 0, 1)$, $\bar{e}_3 = (1, 1, 0)$. Finden Sie für jeden Vektor $v = (b_1, b_2, b_3) \in K^3$ Skalare $\alpha_1, \alpha_2, \alpha_3 \in K$ mit

$$v = \alpha_1 \bar{e}_1 + \alpha_2 \bar{e}_2 + \alpha_3 \bar{e}_3.$$

Zeigen Sie weiter, daß $\alpha_1, \alpha_2, \alpha_3$ eindeutig bestimmt sind.

Übung 2 (Lineare Gleichungssysteme, I)
Seien U, V Unterräume des K^n. Zeigen Sie, daß $U \cap V$ ein Unterraum des K^n ist. Unter welcher Bedingung ist $U \cup V$ ein Unterraum des K^n? Welche Aussagen gelten für affine Unterräume?

Übung 3 (Lineare Gleichungssysteme, II)
Seien $u_1, \ldots, u_k, v_1, \ldots, v_m \in K^n$, und seien

$$U = \{ \alpha_1 u_1 + \ldots + \alpha_k u_k \mid \alpha_i \in K \}, \quad V = \{ \beta_1 v_1 + \ldots + \beta_m v_m \mid \beta_i \in K \}.$$

Zeigen Sie, daß U und V Unterräume des K^n sind, und beschreiben Sie den kleinsten Unterraum W des K^n mit $W \supseteq U \cup V$.

Übung 4 (Das Gauß-Jordansche Eliminationsverfahren, I)
Sei A eine $(n \times n)$-Matrix mit $a_{ii} = 1$ für alle i und $a_{ij} = 0$ für alle $i > j$. Beschreiben Sie, wie Sie eine Lösung von $Ax = b$ ohne Ausräumen des rechten oberen Teils der Matrix A berechnen können. Vergleichen Sie die Anzahl der dazu nötigen Rechenschritte mit der Zahl der Rechenschritte, die ein Ausräumen der Matrix benötigt.

Übung 5 (Das Gauß-Jordansche Eliminationsverfahren, II)
In welche so weit wie möglich ausgeräumte Form können wir eine $(m \times n)$-Matrix A bringen, wenn wir nur die ersten drei elementaren Operationen verwenden dürfen, also keine Spaltenvertauschungen zugelassen sind? Beschreiben Sie, wie Sie anhand der Überführung von A in diese Form die Lösungen eines Gleichungssystems $Ax = b$ bestimmen können.

Übung 6 (Lineare Abbildungen, I)
Sei $n \geq 1$ und sei $A \subseteq \{ 1, \ldots, n \}$. Wir definieren $f, g, h : K^n \to K$ durch

$$f(x_1, \ldots, x_n) = \Sigma_{i \in A} x_i,$$

$$g(x_1, \ldots, x_n) = \Sigma_{i \in A} x_i + 1,$$

$$h(x_1, \ldots, x_n) = \Sigma_{i \in A} 2 \cdot x_i \quad \text{für alle } (x_1, \ldots, x_n) \in K.$$

Zeigen Sie oder widerlegen Sie, daß f, g, h lineare Abbildungen sind, und geben Sie im positiven Fall die darstellenden Matrizen der Abbildungen an.

Übung 7 *(Lineare Abbildungen, II)*
Sei A eine $(n \times n)$-Matrix über K. Zeigen Sie, daß $f_A : K^n \to K^n$ eine lineare Abbildung ist.

Übung 8 *(Lineare Abbildungen, III)*
Seien $f, g : K^n \to K^m$ lineare Abbildungen. Weiter seien $\alpha, \beta \in K$.
Sei $h = \alpha f + \beta g$, d. h.

$h(x) = \alpha f(x) + \beta g(x)$ für alle $x \in K^n$.

Zeigen Sie, daß $h : K^n \to K^m$ eine lineare Abbildung ist.

Übung 9 *(Lineare Abbildungen, IV)*
Seien $f, g : K^n \to K^n$ lineare Abbildungen. Weiter sei $h = f \circ g$.
Zeigen Sie, daß $h = f \circ g$ linear ist.

Übung 10 *(Lineare Abbildungen, V)*
Wir definieren $f, g : \mathbb{Q}^3 \to \mathbb{Q}^3$ durch:

$f(x_1, x_2, x_3) = (x_1 + x_3,\ x_3,\ x_1 - x_2 + x_3)$,

$g(x_1, x_2, x_3) = (x_3,\ 2x_1,\ x_2 - x_1)$ für alle $x_1, x_2, x_3 \in \mathbb{Q}$.

Zeigen Sie, daß f und g lineare Abbildungen sind und geben Sie die darstellenden Matrizen von f, g, $f \circ f$, $g \circ g$, $f \circ g$ und $g \circ f$ an.

Übung 11 *(Lineare Abbildungen, VI)*
Sei $f : K^n \to K^m$ eine lineare Abbildung. Zeigen Sie, daß äquivalent sind:

(a) f ist injektiv.

(b) $\{ x \in K^n \mid f(x) = 0 \} = \{ 0 \}$.

Übung 12 *(Lineare Abbildungen, VII)*
Sei $f : K^n \to K^m$ eine lineare Abbildung. Zeigen Sie:

(a) $f^{-1}[\{ 0 \}]$ ist ein Unterraum des K^n.

(b) rng(f) ist ein Unterraum des K^m.

Übung 13 *(Matrizenmultiplikation, I)*
Beweisen Sie durch direktes Nachrechnen, daß die Matrizenmultiplikation assoziativ ist.

Übung 14 *(Matrizenmultiplikation, II)*
Sei $n \geq 2$. Geben Sie $(n \times n)$-Matrizen A, B an mit $A \cdot B \neq B \cdot A$.

Übung 15 *(Matrizenmultiplikation, III)*
Für eine quadratische Matrix A sei spur(A) die Summe der Einträge auf der Diagonalen von A. Zeigen Sie, daß für alle $(n \times m)$-Matrizen A und $(m \times n)$-Matrizen B gilt, daß $\text{spur}(A \cdot B) = \text{spur}(B \cdot A)$.

Übung 16 *(Matrizenmultiplikation, IV)*
Sei $A = (a_{ij})_{ij}$ die $(n \times n)$-Matrix mit $a_{i,i+1} = 1$ für alle $1 \leq i < n$ und $a_{ij} = 0$ sonst. Bestimmen Sie die Matrizen $A^2, A^3, \ldots, A^n$.

Übung 17 *(Matrizenmultiplikation, V)*
Sei $A = (a_{ij})_{ij}$ eine $(n \times n)$-Matrix mit $a_{ij} = 0$ für alle $i \geq j$. Zeigen Sie, daß A^n die Nullmatrix ist.

Übung 18 *(Matrizenmultiplikation, VI)*
Sei A eine $(m \times n)$-Matrix über K und sei $\alpha \in K$. Geben Sie $(m \times m)$-Matrizen $C_{\alpha,i}$, $C_{i,\alpha j}$, $C_{\sigma(i,j)}$ und $(n \times n)$-Matrizen $C_{\pi(i,j)}$ an mit:

(I) $C_{\alpha,i} A$ = „die Multiplikation der i-ten Zeile von A mit α“,

(II) $C_{i,\alpha j} A$ = „die Addition des α-fachen der j-ten zur i-ten Zeile von A“,

(III) $C_{\sigma(i,j)} A$ = „die Vertauschung der i-ten und der j-ten Zeile von A“,

(IV) $A\, C_{\pi(i,j)}$ = „die Vertauschung der i-ten und der j-ten Spalte von A“.

Übung 19 *(Matrizenmultiplikation, VII)*
Sei A eine $(n \times n)$-Matrix mit $f_A : K^n \to K^n$ bijektiv. Zeigen Sie, daß es ein Produkt $C = C_m \ldots C_1$ von Matrizen des Typs (I) – (III) der vorherigen Aufgabe gibt mit $C \cdot A = D_n$. Formulieren Sie einen Algorithmus zur Konstruktion von C. Zeigen Sie weiter, daß auch $A \cdot C = D_n$ gilt.

Übung 20 *(Relationen und Matrizen, I)*
Seien R, S, T Relationen auf M, und sei $\Delta_M = \{ (x, x) \mid x \in M \}$. Zeigen Sie:

(i) $(R^{-1})^{-1} = R$, $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$, $R \circ (S \circ T) = (R \circ S) \circ T$,

(ii) R ist reflexiv *gdw* $\Delta_M \subseteq R$, R ist irreflexiv *gdw* $\Delta_M \cap R = \emptyset$,

(iii) R ist symmetrisch *gdw* $R = R^{-1}$,

(iv) R ist antisymmetrisch *gdw* $R \cap R^{-1} \subseteq \Delta_M$,

(v) R ist transitiv *gdw* $R \circ R \subseteq R$.

Übung 21 *(Relationen und Matrizen, II)*
Für eine Relation R auf $M = \{ 1, \ldots, n \}$ sei $R^* = \bigcup_{1 \leq k \leq n} R^k$. Zeigen Sie:

(a) R^* ist die transitive Hülle von R, d. h. die kleinste transitive Relation auf M, die R als Teilmenge enthält.

(b) $(R \cup \Delta_M)^*$ mit $\Delta_M = \{ (a, a) \mid a \in M \}$ ist die transitiv-reflexive Hülle von R, d. h. die kleinste transitive und reflexive Relation auf M, die R als Teilmenge enthält.

(c) Ist R reflexiv, so ist $R^k \subseteq R^m$ für alle $k \leq m$.

(d) Geben Sie eine unendliche Relation R auf $\mathbb{N}$ an mit:
$R^k \cap R^m = \emptyset$ für alle $k \neq m$.

Übung 22 (Relationen und Matrizen, III)
Bildet die logische Addition zusammen mit der Multiplikation einen Körper auf { 0, 1 } ?

Übung 23 (Relationen und Matrizen, IV)
Führen Sie den Algorithmus von Warshall durch für M = { 1, …, 6 } und R = { (1, 1), (1, 2), (2, 4), (4, 6), (2, 5), (3, 6), (4, 5), (5, 3), (6, 2) }.

Übung 24 (Relationen und Matrizen, V)
Sei R eine Relation auf M = { 1, …, n } mit darstellender Matrix A. Bestimmen Sie die Anzahl der Rechenschritte, die zur Berechnung der transitiven Hülle von R verwendet werden, wenn

(a) die logischen Produkte A, …, A^n berechnet werden,

(b) der Algorithmus von Warshall verwendet wird.

4. Gruppen

Wir führen in diesem Kapitel den Begriff einer Gruppe ein, der zu den wichtigsten algebraischen Strukturbegriffen der Mathematik gehört. Er eignet sich insbesondere zur Beschreibung und Untersuchung von Symmetrien. Gruppen tauchen nahezu überall auf, der Bogen reicht von der Zahlentheorie und der endlichen Kombinatorik bis hin zur Quantenmechanik und Elementarteilchenphysik.

Nach einer Definition des Begriffs stellen wir einige Beispiele für Gruppen zusammen und untersuchen dann die Gruppenaxiome auf ihre elementaren Folgerungen. Schließlich besprechen wir Untergruppen und zeigen den Satz von Lagrange.

Der Begriff der Gruppe

Für eine natürliche Zahl $n \geq 1$ betrachten wir die Menge

$$S_n = \{ f \mid f : \{ 1, \ldots, n \} \to \{ 1, \ldots, n \} \text{ bijektiv} \}$$

aller *Permutationen* der Zahlen 1, …, n. Eine Permutation ist also eine Umordnung der Zahlen 1, …, n. Wir können ein $f \in S_n$ in verschiedenen Weisen notieren. Offiziell als

$$f = \{ (1, f(1)), \ldots, (n, f(n)) \}, \quad \text{suggestiv als}$$

$$f = 1 \mapsto f(1), \ \ldots, \ n \mapsto f(n), \quad \text{oder auch kurz als n-Tupel}$$

$$f = (f(1), \ldots, f(n)).$$

Wir können zwei Permutationen hintereinander ausführen: Sind $f, g \in S_n$, so ist $g \circ f \in S_n$. Weitere Eigenschaften der Permutationen sind:

(a) Für alle $f, g, h \in S_n$ gilt

$$f \circ (g \circ h) = (f(g(h(1)), \ldots, f(g(h(n))) = (f \circ g) \circ h.$$

(b) Es gibt eine triviale Permutation, nämlich $e = (1, 2, \ldots, n)$.
Für alle $f \in S_n$ gilt:

$$f \circ e = e \circ f = f.$$

(c) Wir können eine Permutation wieder rückgängig machen.
Für alle $f \in S_n$ ist $f^{-1} = \{ (f(1), 1), \ldots, (f(n), n) \} \in S_n$. Für alle $f \in S_n$ gilt:

$$f \circ f^{-1} = f^{-1} \circ f = e.$$

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_11, © Springer-Verlag Berlin Heidelberg 2010

Die gleichen Eigenschaften gelten allgemeiner für die Menge aller Bijektionen auf einer beliebigen Menge M. Sie gelten aber auch für viele Strukturen, die aus Zahlen und arithmetischen Operationen gebildet sind. Ein Beispiel ist die Addition auf den ganzen Zahlen. Hier gilt für alle $a, b, c \in \mathbb{Z}$:

(a) $a + (b + c) = (a + b) + c$,

(b) $a + 0 = 0 + a = a$.

(c) $a - a = -a + a = 0$.

Für die Addition gilt sogar zusätzlich $a + b = b + a$ für alle $a, b \in \mathbb{Z}$. Bereits für die Menge S_3 ist diese Vertauschbarkeit aber nicht mehr allgemein gültig. Denn seien

$f = (1, 3, 2), \quad g = (2, 1, 3).$

Dann gilt

$g \circ f = (2, 3, 1), \quad f \circ g = (3, 1, 2).$

Also ist die Kommutativität eine sicher wichtige, aber nicht stets anzutreffende Eigenschaft innerhalb eines umfassenden strukturellen Kontextes. Diesem Kontext geben wir den Namen „Gruppe“:

Definition *(Gruppe)*
Sei G eine Menge, und sei $\cdot : G \times G \to G$ eine Operation auf G.
Dann heißt das Paar $(G, \cdot)$ eine *Gruppe*, falls gilt:

(a) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in G$. *(Assoziativgesetz)*

(b) Es existiert ein $e \in G$ mit: Für alle $a \in G$ ist $a \cdot e = e \cdot a = a$.
(Existenz eines neutralen Elements)

(c) Für alle $a \in G$ existiert ein $b \in G$ mit $a \cdot b = b \cdot a = e$.
(Existenz eines inversen Elements)

Jedes e wie in (b) heißt ein *neutrales Element* von G. Jedes b wie in (c) heißt ein zu a *inverses Element* von G (bzgl. e). In (c) ist e irgendein fest gewähltes neutrales Element. Wir werden gleich zeigen, daß ein neutrales Element und weiter auch ein zu a inverses Element stets eindeutig bestimmt ist.
Die Aussagen (a) – (c) werden auch als die *Gruppenaxiome* bezeichnet.

Definition *(abelsche Gruppen)*
Eine Gruppe $(G, \cdot)$ heißt *abelsch* oder *kommutativ*, falls gilt:

$a \cdot b = b \cdot a$ für alle $a, b \in G$. *(Kommutativgesetz)*

Ist $(G, \cdot)$ eine Gruppe, so heißt die Funktion $\cdot$ die *Gruppenoperation* auf G. Häufig verwendete Zeichen für die Gruppenoperation sind $\cdot$, $\circ$, $*$, $+$. Das Pluszeichen wird dabei nur für abelsche Gruppen verwendet. Die anderen Zeichen können für abelsche oder nichtabelsche Gruppen verwendet werden.

Statt $a \cdot b$ schreiben wir auch kurz ab. Aufgrund des Assoziativgesetzes können wir Klammern weglassen. Es gilt z. B:

$abcd = (ab)(cd) = ((ab)c)d = a(b(cd)))$, usw.

Einen strengen Beweis für die Freiheit der Klammerung (die das Weglassen von Klammern ermöglicht), führt man durch Induktion über die Anzahl der Faktoren eines Ausdrucks.

Beispiele für Gruppen

Wir stellen einige Beispiele für Gruppen vor.

Triviale Gruppen

Jede Einermenge $G = \{ a \}$ ist eine Gruppe unter der Operation $\cdot : G^2 \to G$ mit $a \cdot a = a$.

Zahlen unter der Addition

Die Strukturen $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, und $(\mathbb{H}, +)$ mit der üblichen Addition sind abelsche Gruppen.

Zahlen unter der Multiplikation

Seien $\mathbb{Q}^*$, $\mathbb{R}^*$, $\mathbb{C}^*$, $\mathbb{H}^*$ die Zahlenmengen $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ ohne die Null. Dann sind $\mathbb{Q}^*$, $\mathbb{R}^*$, $\mathbb{C}^*$ und $\mathbb{H}^*$ unter der üblichen Multiplikation Gruppen. Mit Ausnahme von $\mathbb{H}^*$ sind diese Gruppen zudem abelsch.

Addition von Vektoren

Für alle $n \geq 1$ bildet $\mathbb{R}^n$ mit der punktweisen Addition $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ eine abelsche Gruppe. Der Nullvektor $0 = (0, \ldots, 0)$ ist neutral und $(-x_1, \ldots, -x_n)$ ist invers zu $(x_1, \ldots, x_n)$.

Restklassengruppen

Für alle $m \geq 1$ bildet $\mathbb{Z}_m$ unter der Addition von Restklassen eine abelsche Gruppe. Weiter bildet $\mathbb{Z}_m^* = \mathbb{Z}_m - \{ 0 \}$ unter der Multiplikation von Restklassen genau dann eine Gruppe, wenn m eine Primzahl ist.

Kleinsche Vierergruppe

Seien 1, a, b, c paarweise verschieden. Wir definieren eine Multiplikation $*$ auf $V = \{ 1, a, b, c \}$ durch die rechtsstehende Multiplikationstafel. Dann ist $(V, *)$ eine abelsche Gruppe.

$*$	1	a	b	c
1	1	a	b	c
a	a	1	c	b
b	b	c	1	a
c	c	b	a	1

Permutationsgruppen

Für jede Menge M bildet $G = \{ f \mid f : M \to M \text{ ist bijektiv} \}$ zusammen mit der Verknüpfung $\circ$ von Funktionen eine Gruppe. Das neutrale Element ist id_M, die Identität auf M. Für alle $f \in G$ ist die Umkehrfunktion f^{-1} invers zu f. Die Gruppe G heißt die *Permutationsgruppe* von M.

Umkehrung der Gruppenoperation
Ist $(G, \cdot)$ eine Gruppe, so ist auch $(G, *)$ eine Gruppe, wobei $a * b = b \cdot a$ für alle $a, b \in G$ gesetzt wird.

Gruppe der invertierbaren Elemente
Sei $\cdot : M^2 \to M$ eine assoziative Operation auf einer Menge M, und es existiere ein neutrales Element e für diese Operation. Sei G die Menge der invertierbaren Elemente von M, d. h. es gilt

$$G = \{ a \in M \mid \text{es gibt ein } b \in G \text{ mit } ab = ba = e \}.$$

Dann ist $(G, \cdot \mid G^2)$ eine Gruppe.

Gruppen und Körper
Ist K ein Körper, so sind $(K, +)$ und $(K^*, \cdot)$ abelsche Gruppen, wobei wieder $K^* = K - \{ 0 \}$ gesetzt wird. Umgekehrt bildet für zwei kommutative Gruppen $(K, +)$ und $(K^*, \cdot)$ das Tripel $(K, +, \cdot)$ genau dann einen Körper, wenn das Distributivgesetz $a(b + c) = ab + ac$ und zudem $0 \cdot a = 0$ für alle $a \in K$ gilt.

Folgerungen aus den Gruppenaxiomen

Wir wollen nun einige elementare Folgerungen aus den Gruppenaxiomen ableiten. Den Anfang bildet eine vertraute Regel:

Satz *(Kürzungsregel)*
Sei $(G, \cdot)$ eine Gruppe. Dann gilt für alle $a, b, c \in G$:

(i) $a\,b = a\,c$ *impliziert* $b = c$,

(ii) $b\,a = c\,a$ *impliziert* $b = c$.

Beweis
Sei a' invers zu a. Dann gilt:

$$b = e\,b = a'\,a\,b = a'\,a\,c = e\,c = c.$$

– Die zweite Regel wird analog bewiesen.

Als nächstes zeigen wir, daß wir anstelle von *einem* neutralen oder inversen Element von *dem* neutralen oder inversen Element reden können:

Satz *(Eindeutigkeit des neutralen und des inversen Elements)*
Sei $(G, \cdot)$ eine Gruppe. Dann gilt:

(i) Seien e, e' neutrale Elemente von G. Dann gilt $e = e'$.

(ii) Sei $a \in G$, und seien b, c invers zu a. Dann gilt $b = c$.

Beweis

zu (i): Es gilt $e = e \cdot e' = e'$.

zu (ii): Es gilt $ab = e = ac$. Also ist $b = c$ nach der Kürzungsregel. ∎

Das bevorzugte Zeichen für das neutrale Element einer mit Hilfe des Pluszeichens notierten Gruppe ist 0. Im Kontext von $\cdot, \circ, *$ wird neben e vor allem auch das Zeichen 1 verwendet.

Das Inverse eines Gruppenelements a bezeichnen wir mit a^{-1}. Für additiv notierte Gruppen $(G, +)$ verwenden wir jedoch die Notation $-a$ anstelle von a^{-1}. Wir schreiben dann weiter $a - b$ anstelle von $a + (-b)$.

In einer Gruppe $(G, \cdot)$ gelten also neben dem Assoziativgesetz die Gesetze:

$$a\,e = e\,a = a, \quad a\,a^{-1} = a^{-1}\,a = e \quad \text{für alle } a \in G.$$

Für additiv notierte Gruppen $(G, +)$ lesen sich diese Aussagen als:

$$a + 0 = 0 + a = a, \quad a - a = -a + a = 0 \quad \text{für alle } a \in G.$$

Für die Inversenbildung gelten die beiden folgenden Regeln, die häufig verwendet werden:

Satz *(Regeln für die Inversenbildung)*

Sei $(G, \cdot)$ eine Gruppe. Dann gilt für alle $a, b \in G$:

(i) $(a^{-1})^{-1} = a$.

(ii) $(a\,b)^{-1} = b^{-1}\,a^{-1}$.

Beweis

zu (i):

Es gilt $a\,a^{-1} = e = (a^{-1})^{-1}\,a^{-1}$. Nach der Kürzungsregel ist dann aber $a = (a^{-1})^{-1}$.

zu (ii):

Es gilt $(a\,b)(b^{-1}\,a^{-1}) = a\,(b\,b^{-1})\,a^{-1} = a\,e\,a^{-1} = a\,a^{-1} = e$.
Analog ist $(b^{-1}\,a^{-1})(a\,b) = e$. Wegen der Eindeutigkeit des Inversen ist also $(ab)^{-1} = b^{-1}\,a^{-1}$. ∎

Exponentiation und Vervielfachung

Für jede Gruppe $(G, \cdot)$ können wir eine Exponentiation einführen. Für $a \in G$ definieren wir hierzu rekursiv:

$$a^0 = e, \quad a^{n+1} = a^n \cdot a \quad \text{für alle } n \in \mathbb{N}.$$

Wir schreiben weiter a^{-n} für $(a^n)^{-1}$. Damit ist a^n für alle $n \in \mathbb{Z}$ definiert. Es gelten die üblichen Gesetze für die Exponentiation, wie man durch Induktion und Inversenbildung beweisen kann:

$a^n a^m = a^{n+m}$ für alle $a \in G$ und alle $n, m \in \mathbb{Z}$,

$(a^n)^m = a^{n \cdot m}$ für alle $a \in G$ und alle $n, m \in \mathbb{Z}$,

$a^n b^n = (ab)^n$ für alle $a, b \in G$ mit $ab = ba$ und alle $n \in \mathbb{Z}$.

Die doppelte Verwendung des Malzeichens für die Gruppenoperation und für die Multiplikation auf den ganzen Zahlen ist in der Regel unproblematisch.

Für additiv notierte (also abelsche) Gruppen schreiben wir na anstelle von a^n und nennen na ein *Vielfaches* von a. Es gilt dann für alle $a, b \in G$ und alle ganzen Zahlen n, m aufgrund obiger Regeln:

$na + ma = (n + m)a$,

$n(ma) = (nm)a$,

$na + nb = n(a + b)$.

Schließlich definieren wir für alle $a_1, \ldots, a_n \in G$ das Produkt der a_i durch:

$\Pi_{1 \le i \le n} a_i = a_1 \cdot \ldots \cdot a_n$.

In Erweiterung der Konvention $a^0 = e$ vereinbaren wir, daß das Produkt über die leere Folge gleich e sein soll. Dann gilt für alle $n \ge 0$, daß $\Pi_{1 \le i \le n} a = a^n$.

Für abelsche Gruppen schreiben wir Σ anstelle von Π, also

$\Sigma_{1 \le i \le n} a_i = a_1 + \ldots + a_n$.

Hier können wir aufgrund der Kommutativität sogar eine Summe für beliebige endliche Teilmengen E von G einführen. Wir setzen:

$\Sigma E = \Sigma_{a \in E} a = a_1 + \ldots + a_n$,

wobei $E = \{ a_1, \ldots, a_n \}$ mit $a_i \ne a_j$ für alle $1 \le i < j \le n$. Das Gruppenelement ΣE heißt die *Summe* der Menge E in der Gruppe. Für $E = \varnothing$ gilt $\Sigma E = 0$.

Untergruppen

Viele Gruppen besitzen eine Reihe von Teilmengen, die selbst eine Gruppenstruktur aufweisen. Sie spielen bei der Untersuchung der „Mutter-Gruppe“ eine wesentliche Rolle. Wir definieren:

Definition *(Untergruppe)*
Sei $(G, \cdot)$ eine Gruppe, und sei $H \subseteq G$. H heißt *Untergruppe* von G, falls H zusammen mit der von G ererbten Multiplikation eine Gruppe bildet, d. h. falls $(H, \cdot | H^2)$ eine Gruppe ist.

Implizit ist in dieser Definition enthalten, daß die Menge H abgeschlossen unter der Gruppenoperation ist, d. h. es gilt $a \cdot b \in H$ für alle $a, b \in H$.

Für jede Gruppe $(G, \cdot)$ mit neutralem Element e sind $\{ e \}$ und G Untergruppen von G, die sog. *trivialen Untergruppen* von G.

Wir geben einige Beispiele für Untergruppen von konkreten Gruppen. Unter der Addition ist $\mathbb{Z}$ eine Untergruppe von $\mathbb{Q}$, und $\mathbb{Q}$ selbst ist eine Untergruppe von $\mathbb{R}$. Jede Gerade durch den Nullpunkt ist eine Untergruppe von $(\mathbb{R}^2, +)$. Die Untergruppen von $(\mathbb{Z}, +)$ sind genau die Teilmengen von $\mathbb{Z}$ der Form

$$\mathbb{Z}d = \{ kd \mid k \in \mathbb{Z} \}.$$

Denn jede Menge $\mathbb{Z}d$ ist eine Untergruppe von $\mathbb{Z}$, wie leicht zu sehen ist. Ist umgekehrt H eine Untergruppe von $\mathbb{Z}$, so ist H abgeschlossen unter Addition $a + b$ und weiter dann unter der Vervielfachung na, d. h. H ist abgeschlossen unter Linearkombinationen $na + mb$ mit $a, b \in H$ und $n, m \in \mathbb{Z}$. Nach den Ergebnissen des vorletzten Kapitels ist dann also $H = \{ 0 \}$ oder $H = \mathbb{Z}d$, wobei d das kleinste positive Element von H ist.

Als nächstes beweisen wir einen nützlichen Satz, der ein hinreichendes und notwendiges Kriterium dafür aufstellt, wann eine Teilmenge einer Gruppe eine Untergruppe der Gruppe bildet.

Satz *(Untergruppenkriterium)*

Sei $(G, \cdot)$ eine Gruppe. Dann sind die folgenden Aussagen äquivalent:

(i) H ist eine Untergruppe von G.

(ii) $H \neq \emptyset$, und für alle $a, b \in H$ gilt $ab \in H$, $a^{-1} \in H$.

(iii) $H \neq \emptyset$, und für alle $a, b \in H$ gilt $ab^{-1} \in H$.

Beweis

(i) ⇝ (ii):

Da $(H, \cdot \mid H^2)$ eine Gruppe ist, ist $H \neq \emptyset$ und $ab \in H$ für alle $a, b \in H$. Sei d das neutrale Element von H. Dann gilt – wie in jeder Gruppe –, daß $d^2 = d$. Also ist $d\,d = d\,e$ und damit $d = e$ nach der Kürzungsregel. Das neutrale Element von H ist also das neutrale Element von G. Ebenso sind die Inversen im Sinne von H die Inversen im Sinne von G: Denn sei $a \in H$. Dann existiert ein $b \in H$ mit $ab = d = e$. Multiplikation in G von links mit a^{-1} zeigt, daß $b = a^{-1}$. Also ist $a^{-1} \in H$.

(ii) ⇝ (iii):

Seien $a, b \in H$. Dann ist $c = b^{-1} \in H$ und weiter dann $ab^{-1} = ac \in H$ nach Voraussetzung.

(iii) ⇝ (i):

Wir betrachten die Menge H zusammen mit der von G ererbten Multiplikation. Dann gilt das Assoziativgesetz für alle Elemente von H, da es für alle Elemente von G gilt.

Sei nun $a \in H$ beliebig. Dann ist $aa^{-1} = e \in H$. Für alle $a \in H$ ist dann weiter $a^{-1} = ea^{-1} \in H$. Da die Multiplikation auf H für Elemente von H mit der Multiplikation auf G übereinstimmt, gilt

$$ae = ea = a \text{ und } aa^{-1} = a^{-1}a = e \quad \text{für alle } a \in H.$$

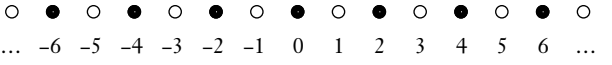
Schließlich ist $ab \in H$ für alle $a, b \in H$, denn es gilt $b^{-1} \in H$ nach dem bereits Gezeigten, und damit ist $ab = a(b^{-1})^{-1} \in H$ nach Voraussetzung. Also ist $(H, \cdot | H^2)$ eine Gruppe. ∎

Wir wollen nun noch eine abstrakte Konstruktionsmethode für Untergruppen einer beliebigen Gruppe vorstellen.

Definition *(von einem Element erzeugte Untergruppe)*
Sei $(G, \cdot)$ eine Gruppe, und sei $a \in G$ beliebig. Dann heißt

$$H = \{ a^n \mid n \in \mathbb{Z} \}$$

die von a *erzeugte Untergruppe* von G.



von a = 2 erzeugte Untergruppe H = { n 2 | n ∈ ℤ } der Gruppe (ℤ, +)

In der Tat ist H eine Untergruppe von G, denn das neutrale Element $e = a^0$ ist in H und für alle $a^n, a^m \in H$ ist $a^n \cdot (a^m)^{-1} = a^{n-m} \in H$ nach den Rechenregeln für die Exponentiation. Die Gruppe H ist die kleinste Untergruppe von G, die a als Element enthält. Sie ist immer abelsch, unabhängig davon, ob G abelsch ist oder nicht.

Definition *(zyklische Gruppen)*
Eine Gruppe $(G, \cdot)$ heißt *zyklisch*, wenn es ein $a \in G$ gibt derart, daß die von a erzeugte Untergruppe ganz G ist. Jedes derartige a heißt dann ein *Generator* oder *Erzeuger* von G.

Eine zyklische Gruppe ist also notwendig abelsch. Nach obiger Überlegung ist jede Untergruppe von $(\mathbb{Z}, +)$ zyklisch. Bereits die Kleinsche Vierergruppe ist dagegen nicht zyklisch.
Wir definieren weiter:

Definition *(Ordnung eines Gruppenelements)*
Sei $(G, \cdot)$ eine Gruppe. Ein $a \in G$ heißt *von endlicher Ordnung*, falls ein $n \geq 1$ existiert mit $a^n = e$. Wir setzen dann

$$\mathrm{ord}_G(a) = \text{„das kleinste } n \geq 1 \text{ mit } a^n = e\text{“},$$

und nennen $\mathrm{ord}_G(a)$ die *Ordnung* von a in G.

Hat $a \in G$ endliche Ordnung, so ist $\mathrm{ord}_G(a)$ gleich der Anzahl der Elemente der von a erzeugten Untergruppe.
Für jede abelsche Gruppe (G, +) bildet

$$H = \{ a \in G \mid a \text{ hat endliche Ordnung} \}$$

eine Untergruppe von G, die sog. *Torsions-Untergruppe* von G.

Nebenklassen und Faktorgruppen

Jede Untergruppe H einer Gruppe G induziert eine natürliche Äquivalenzrelation auf der Gruppe: Die Identität $x = y$ ist gleichwertig mit $x\,y^{-1} = e$. Statt „gleich e“ fordern wir nun schwächer „$\in H$“:

Definition *(von einer Untergruppe induzierte Äquivalenzrelationen)*
Sei $(G, \cdot)$ eine Gruppe, und sei H eine Untergruppe von G. Dann setzen wir für alle $x, y \in G$:

$$x \sim_H y, \ \textit{falls}\ \ x\,y^{-1} \in H, \quad x \sim^H y, \ \textit{falls}\ \ x^{-1}\,y \in H.$$

Diese abgeschwächten Gleichheiten sind, wie wir gleich zeigen werden, Äquivalenzrelationen auf G. Wir definieren hierzu für alle $a \in G$ und alle $B, C \subseteq G$:

$$aB = \{ab \mid b \in B\}, \quad Ba = \{ba \mid b \in B\}, \quad BC = \{bc \mid b \in B, c \in C\}.$$

Nun zeigen wir:

Satz *(über $\sim_H$ und $\sim^H$)*
Sei $(G, \cdot)$ eine Gruppe, und sei H eine Untergruppe von G.
Dann sind $\sim_H$ und $\sim^H$ Äquivalenzrelationen auf G. Weiter gilt:

$$a/\sim_H = Ha, \quad a/\sim^H = aH \quad \text{für alle } a \in G.$$

Beweis
Wir zeigen die Aussagen über $\sim = \sim_H$. Der Beweis für $\sim^H$ ist analog.

$\sim$ *ist reflexiv:* Für alle $x \in G$ ist $x\,x^{-1} = e \in H$. Also ist $x \sim x$.

$\sim$ *ist symmetrisch:* Sei $x \sim y$. Dann ist $z = x\,y^{-1} \in H$. Also gilt

$$y\,x^{-1} = (x\,y^{-1})^{-1} = z^{-1} \in H.$$

Also gilt $y \sim x$.

$\sim$ *ist transitiv:* Seien $x \sim y$ und $y \sim z$. Dann sind $h_1 = x\,y^{-1}$ und $h_2 = y\,z^{-1}$ Elemente von H und damit ist

$$x\,z^{-1} = x\,y^{-1}\,y\,z^{-1} = h_1\,h_2 \in H.$$

Also ist $x \sim z$.

Also ist $\sim$ eine Äquivalenzrelation auf G. Für alle $a, x \in G$ gilt:

$x \sim a$ *gdw* $xa^{-1} \in H$ *gdw*

es gibt ein $h \in H$ mit $xa^{-1} = h$ *gdw*

es gibt ein $h \in H$ mit $x = ha$ *gdw* $x \in Ha$.

Also gilt $a/\sim = Ha$.

Wir definieren:

Definition *(Nebenklassen)*
Die Äquivalenzklassen Ha der Relation $\sim_H$ heißen die *Rechtsnebenklassen* von H in G. Analog heißen die Äquivalenzklassen aH der Relation $\sim^H$ die *Linksnebenklassen* von H in G.

Im allgemeinen ist $aH \neq Ha$. Die Beziehung $aH = Ha$ ist aber eine wünschenswerte Eigenschaft, denn sie führt dazu, daß wir auf den Nebenklassen eine Gruppenoperation einführen können:

Definition *(Normalteiler)*
Sei $(G, \cdot)$ eine Gruppe. Eine Untergruppe H von G heißt ein *Normalteiler* von G, falls $aH = Ha$ für alle $a \in G$ gilt. Wir setzen dann:

$G/H = \{ aH \mid a \in G \}$,

$aH \cdot bH = (ab)H$ für alle $a, b \in G$,

und nennen $(G/H, \cdot)$ die *Faktorgruppe* von G bzgl. H.

In der Tat ist $(G/H, \cdot)$ eine wohldefinierte Gruppe.

Ist G abelsch, so ist jede Untergruppe ein Normalteiler. Dagegen existieren Untergruppen der Permutationsgruppe S_3, die keine Normalteiler sind.

Für beliebige Gruppen liefern Normalteiler folgende Abschwächung der Kommutativität, die in der Gruppentheorie eine wichtige Rolle spielt:

Definition *(auflösbare Gruppen)*
Eine Gruppe $(G, \cdot)$ heißt *auflösbar*, falls Untergruppen $G_0, \ldots, G_n$ von G existieren mit den Eigenschaften:

(a) $G = G_0 \supseteq \ldots \supseteq G_n = \{ e \}$,

(b) G_{i+1} ist ein Nomalteiler von G_i für alle $i < n$,

(c) G_i/G_{i+1} ist abelsch für alle $i < n$.

Die Folge $G_0, \ldots, G_n$ heißt dann eine *Normalreihe* von G.

Die Permutationsgruppen S_2, S_3 und S_4 sind auflösbar, wie folgende Ketten von Normalteilern mit abelschen Faktoren zeigen:

$S_2 \supset \{ (1, 2) \}$,

$S_3 \supset S_3^+ \supset \{ (1, 2, 3) \}$,

$S_4 \supset S_4^+ \supset \{ (1, 2, 3, 4), (2, 1, 4, 3), (3, 4, 1, 2), (4, 3, 2, 1) \} \supset \{ (1, 2, 3, 4) \}$,

wobei $S_n^+ = \{ f \in S_n \mid \operatorname{sgn}(f) = 1 \}$ mit

$\operatorname{sgn}(f) = \Pi_{1 \leq i < j \leq n} (f(i) - f(j))/(i - j) \in \{ -1, 1 \}$,

dem *Signum* der Permutation f. Konkret ist

$S_3^+ = \{ (1, 2, 3),\ (2, 3, 1),\ (3, 1, 2) \},$

$S_4^+ = \{ (1, 2, 3, 4),\ (1, 4, 2, 3),\ (1, 3, 4, 2),\ (2, 1, 4, 3),\ (2, 3, 1, 4),\ (2, 4, 3, 1),$
$(3, 1, 2, 4),\ (3, 4, 1, 2),\ (3, 2, 4, 1),\ (4, 1, 3, 2),\ (4, 2, 1, 3),\ (4, 3, 2, 1) \}.$

Dagegen sind, wie man zeigen kann, die Gruppen S_n für $n \geq 5$ nicht mehr auflösbar. Die Nichtauflösbarkeit dieser Gruppen verwendet man in der Galoistheorie der Algebra, um zu zeigen, daß es keine durch Wurzelausdrücke gebildeten Lösungsformeln für Polynomgleichungen fünften und höheren Grades gibt.

Der Satz von Lagrange

Für eine endliche Menge M sei wieder $|M|$ die Anzahl der Elemente von M, die *Kardinalität* von M. Für eine endliche Gruppe G heißt $|G|$ auch die *Ordnung* von G. Wir zeigen nun:

Satz *(Satz von Lagrange)*
Sei $(G, \cdot)$ eine endliche Gruppe, und sei H eine Untergruppe von G.
Dann ist $|H|$ ein Teiler von $|G|$. Genauer gilt

$$|G| = (H : G) \cdot |H|,$$

wobei $(H : G)$ die Anzahl der Rechtsnebenklassen von H in G ist.

Beweis
Für $a \in G$ definieren wir $f_a : H \to Ha$ durch $f_a(x) = xa$ für alle $x \in H$.
Dann ist $f_a : H \to Ha$ bijektiv. Also gilt:

(+) $|H| = |Ha|$ für alle $a \in G$.

Nach obigem Satz ist $Z = \{ Ha \mid a \in G \}$ eine Zerlegung von G. Also gilt

$$|G| = |Z| \cdot |H|,$$

denn G ist die disjunkte Vereinigung von $|Z|$-vielen Mengen, die nach (+) alle jeweils $|H|$-viele Elemente besitzen. ∎

Statt „Anzahl der Rechtsnebenklassen" können wir auch „Anzahl der Linksnebenklassen" schreiben, denn die Funktion f mit $f(aH) = Ha$ für alle $a \in G$ ist eine wohldefinierte Bijektion zwischen den Links- und Rechtsnebenklassen.
Aus dem Satz folgt unmittelbar:

Korollar *(Untergruppen von Gruppen mit Primzahlordnung)*
Sei G eine Gruppe, und $|G|$ sei eine Primzahl.
Dann sind $\{ e \}$ und G die einzigen Untergruppen von G.

Damit wird jede Gruppe mit Primzahlordnung von jedem ihrer Elemente ungleich e erzeugt:

Korollar *(Gruppen mit Primzahlordnung)*
Sei G eine Gruppe mit Primzahlordnung. Dann gilt für alle $a \in G$ mit $a \neq e$, daß $G = \{ a^n \mid n \in \mathbb{Z} \}$.
Insbesondere ist G zyklisch und damit abelsch.

Beweis
$H = \{ a^n \mid n \in \mathbb{Z} \}$ ist eine von $\{ e \}$ verschiedene Untergruppe von G.
Also gilt $H = G$ nach dem vorangehenden Korollar. ∎

Übungen

Übung 1 *(Der Begriff der Gruppe, I)*
Für alle $n \geq 2$ heißt ein $g \in S_n$ eine *Transposition*, falls es $1 \leq i < j \leq n$ gibt mit

$g(i) = j, \quad g(j) = i, \quad g(k) = k$ für alle $1 \leq k \leq n$ mit $k \neq i, j$.

(a) Sei $f = (2, 4, 1, 3) \in S_4$. Finden Sie Transpositionen $g_1, ..., g_k$ mit $f = g_1 \circ ... \circ g_k$.

(b) Zeigen Sie allgemein, daß jedes $f \in S_n$, $n \geq 2$, eine Komposition von Transpositionen ist.

Übung 2 *(Der Begriff der Gruppe, II)*
Sei $n \geq 1$. Für alle $f \in S_n$ sei $\operatorname{sgn}(f) = \Pi_{1 \leq i < j \leq n} (f(i) - f(j))/(i - j)$ das *Signum* von f. Zeigen Sie, daß für alle $f, g \in S_n$ gilt:

(i) $\operatorname{sgn}(f) \in \{ -1, 1 \}$,

(ii) $\operatorname{sgn}(f \circ g) = \operatorname{sgn}(f) \cdot \operatorname{sgn}(g)$,

(iii) $\operatorname{sgn}(f) = 1$ *gdw* f ist eine Komposition einer geraden Anzahl von Transpositionen.

Übung 3 *(Der Begriff der Gruppe, III)*
Stellen Sie einen Zusammenhang her zwischen der Gruppe S_3 und allen Spiegelungen und Drehungen eines gleichseitigen Dreieckes.

Übung 4 *(Der Begriff der Gruppe, IV)*
Sei G eine Menge, und sei $\cdot : G \times G \rightarrow G$ eine Operation auf G mit:

(a) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in G$. *(Assoziativgesetz)*

(b) Es existiert ein $e \in G$ mit: Für alle $x \in G$ ist $x \cdot e = x$.
(Existenz eines rechtsneutralen Elements)

(c) Für alle $x \in G$ existiert ein $y \in G$ mit $x \cdot y = e$.
(Existenz eines rechtsinversen Elements)

Zeigen Sie, daß $(G, \cdot)$ eine Gruppe ist.

Übung 5 (Der Begriff der Gruppe, V)
Sei G eine nichtleere Menge, und sei $\cdot : G \times G \to G$ eine assoziative Operation auf G. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) $(G, \cdot)$ ist eine Gruppe.

(b) Für alle $a, b \in G$ sind die Gleichungen „$x \cdot a = b$“ und „$a \cdot x = b$“ lösbar in der Variablen x in G.

Übung 6 (Der Begriff der Gruppe, VI)
Seien G_1, G_2 Gruppen in multiplikativer Schreibweise. Wir definieren eine Multiplikation auf $G_1 \times G_2$ durch:

$(x_1, x_2) \cdot (y_1, y_2) = (x_1 y_1, x_2 y_2)$ für alle $(x_1, x_2), (y_1, y_2) \in G_1 \times G_2$.

Zeigen Sie, daß $G_1 \times G_2$ mit dieser Multiplikation eine Gruppe ist.

Übung 7 (Der Begriff der Gruppe, VII)
Sei $(G, \cdot)$ eine Gruppe. Wir definieren wieder für beliebige $A, B \subseteq G$:

$AB = \{ xy \mid x \in A, y \in B \}$.

Ist $\mathcal{P}(G)$ mit dieser Multiplikation eine Gruppe? Welche Gruppenaxiome gelten?

Übung 8 (Der Begriff der Gruppe, VIII)
Finden Sie drei verschiedene Elemente $a, b, c \neq e$ der Gruppe S_4, die zusammen mit $e \in S_4$ isomorph zur Kleinschen Vierergruppe sind, d. h. es gilt $a^2 = b^2 = c^3 = e$ und $xy = z$ für alle x, y, z mit $\{ x, y, z \} = \{ a, b, c \}$.

Übung 9 (Folgerungen aus den Gruppenaxiomen, I)
Sei $(G, \cdot)$ eine Gruppe. Zeigen oder widerlegen Sie:

(a) Für alle $x \in G$ gilt $x^2 = x$ genau dann, wenn $x = e$.

(b) Für alle $x, y \in G$ gilt: Ist $xy = e$, so ist $x = e$ oder $y = e$.

(c) Für alle $x, y \in G$ gilt $(xy)^2 = x^2 y^2$.

Übung 10 (Folgerungen aus den Gruppenaxiomen, II)
Sei $(G, \cdot)$ eine Gruppe, und sei $a \in G$. Wir definieren $f : G \to G$ durch

$f(x) = ax$ für alle $x \in G$.

Zeigen Sie, daß $f : G \to G$ bijektiv ist.

Übung 11 (Folgerungen aus den Gruppenaxiomen, III)
Sei $(G, \cdot)$ eine Gruppe mit neutralem Element e, und sei $a \in G$.
Weiter sei $n \geq 1$ derart, daß $a^n = e$ und $a^i \neq e$ für alle $1 \leq i < n$. Zeigen Sie:

(i) Für alle i, j mit $0 \leq i < j < n$ ist $a^i \neq a^j$.

(ii) Für alle $m \in \mathbb{Z}$ existiert ein $0 \leq i < n$ mit $a^m = a^i$.

Übung 12 *(Folgerungen aus den Gruppenaxiomen, IV)*
Sei $(G, \cdot)$ eine Gruppe mit $a^2 = e$ für alle $a \in G$. Zeigen Sie, daß die Gruppe abelsch ist.

Übung 13 *(Folgerungen aus den Gruppenaxiomen, V)*
Sei $(G, \cdot)$ eine endliche abelsche Gruppe, und sei $G = \{ a_1, \ldots, a_n \}$ mit paarweise verschiedenen a_i. Zeigen Sie, daß $a_1^2 \cdot \ldots \cdot a_n^2 = e$.

Übung 14 *(Exponentiation und Vervielfachung)*
Sei $(G, \cdot)$ eine Gruppe. Zeigen Sie, daß für alle $a, b \in G$ und $n, m \in \mathbb{Z}$ gilt:

(i) $a^n a^m = a^{n+m}$,

(ii) $(a^n)^m = a^{n \cdot m}$,

(iii) $a^n b^n = (ab)^n$, falls $ab = ba$.

Geben Sie weiter eine Gruppe $(H, \cdot)$ und $a, b \in H$ an mit $a^2 b^2 \neq (a b)^2$.

Übung 15 *(Untergruppen, I)*
Sei $(G, \cdot)$ eine Gruppe, und seien H_1 und H_2 Untergruppen von G. Zeigen Sie, daß $H_1 \cap H_2$ eine Untergruppe von G ist.

Übung 16 *(Untergruppen, II)*
Sei $(G, \cdot)$ eine Gruppe, und seien H_1 und H_2 Untergruppen von G. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(i) $H_1 \cup H_2$ ist eine Untergruppe von G.

(ii) Es gilt $H_1 \subseteq H_2$ oder $H_2 \subseteq H_1$.

Übung 17 *(Untergruppen, III)*
Sei $(G, \cdot)$ eine Gruppe, und sei $a \in G$. Zeigen Sie, daß $H = \{ a^n \mid n \in \mathbb{Z} \}$ eine Untergruppe von G ist.

Übung 18 *(Untergruppen, IV)*
Sei $(G, \cdot)$ eine Gruppe, und sei $H = \{ a \in G \mid ab = ba \text{ für alle } b \in G \}$. Zeigen Sie, daß H eine abelsche Untergruppe von G ist.

Übung 19 *(Untergruppen, V)*
Sei $(G, \cdot)$ eine Gruppe, und seien H_1 und H_2 Untergruppen von G. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(i) $H_1 H_2$ ist eine Untergruppe von G.

(ii) $H_1 H_2 \subseteq H_2 H_1$.

(iii) $H_2 H_1 \subseteq H_1 H_2$.

(iv) $H_1 H_2 = H_2 H_1$.

Hierbei sei $AB = \{ ab \mid a \in A, b \in B \}$ für alle $A, B \subseteq G$.

Übung 20 (Nebenklassen und Faktorgruppen, I)
Geben Sie eine Untergruppe H der Permutationsgruppe S_3 auf $\{1, 2, 3\}$ an, die kein Normalteiler von S_3 ist. Bestimmen Sie die Links- und Rechtsnebenklassen gH und Hg für alle $g \in S_3$. Geben Sie weiter $g_1, \ldots, g_n$ an, sodaß $\{g_1, \ldots, g_n\}$ ein vollständiges Repräsentantensystem sowohl für $S_3/\sim_H$ als auch für $S_3/\sim^H$ ist.

Übung 21 (Nebenklassen und Faktorgruppen, II)
Sei H ein Normalteiler von G, und sei

$$G/H = \{aH \mid a \in G\} \quad (= \{Ha \mid a \in G\}),$$

$$aH \cdot bH = (ab)H \quad \text{für alle } a, b \in G.$$

Zeigen Sie, daß diese Multiplikation wohldefiniert ist, d. h. aus $aH = a'H$ und $bH = b'H$ folgt stets $(ab)H = (a'b')H$, und daß G/H unter dieser Multiplikation eine Gruppe bildet. Wo wird gebraucht, daß H ein Normalteiler ist?

Übung 22 (Nebenklassen und Faktorgruppen, III)
Sei $(G, \cdot)$ eine Gruppe. Für alle $a, b \in G$ sei

$$[a, b] = aba^{-1}b^{-1}$$

der *Kommutator* von a und b, und für alle $A, B \subseteq G$ sei

$$[A, B] = \{[a, b] \mid a \in A, b \in B\}$$

der *Kommutator* der Mengen A und B. Zeigen Sie:

(a) $[G, G]$ ist ein Normalteiler und $G/[G, G]$ ist abelsch.

(b) Ist N ein Normalteiler von G mit G/N abelsch, so ist $[G, G] \subseteq N$.

(c) G ist genau dann auflösbar, wenn die Kette $G^0 \supseteq G^1 \supseteq G^2 \supseteq \ldots$ in der Untergruppe $\{e\}$ endet, wobei $G^0 = G$ und $G^{i+1} = [G^i, G^i]$ für alle i.

Übung 23 (Nebenklassen und Faktorgruppen, IV)
Zeigen Sie, daß

$$S_2 \supset \{(1, 2)\}, \quad S_3 \supset S_3^+ \supset \{(1, 2, 3)\}, \quad \text{sowie}$$

$$S_4 \supset S_4^+ \supset \{(1, 2, 3, 4), (2, 1, 4, 3), (3, 4, 1, 2), (4, 3, 2, 1)\} \supset \{(1, 2, 3, 4)\}.$$

Ketten von Normalteilern mit kommutativen Faktoren sind.
Zeigen Sie weiter, daß $[S_5, S_5] = S_5^+$ und $[S_5^+, S_5^+] = S_5^+$ gilt und folgern Sie mit Hilfe der vorangehenden Übung, daß S_5 nicht auflösbar ist.

Übung 24 (Der Satz von Lagrange)
Sei $(G, \cdot)$ eine endliche Gruppe mit genau n Elementen. Zeigen Sie:

$$a^n = e \quad \text{für alle } a \in G.$$

5. Graphen

Wir stellen in diesem Kapitel die Anfänge der Graphentheorie vor. Ein Graph besteht aus „Ecken“, die durch gewisse „Kanten“ verbunden sind. Die Kanten können gerichtet sein oder ungerichtet, und sie können weiter auch noch mit Zahlen beschriftet sein, die ihre „Länge“ angeben. Wir fragen dann zum Beispiel nach Konstruktionen für kürzeste Wege von einer Ecke zu einer anderen, oder nach der Existenz eines Kantenzuges, der alle Ecken oder alle Kanten des Graphen genau einmal besucht, usw.

Die Anfänge der Graphentheorie liegen im 18. und 19. Jahrhundert, als kombinatorische Fragen mathematisch untersucht wurden, etwa die Frage, wie man aus einem Labyrinth herausfindet oder die Frage, wie viele Farben man braucht, um eine Landkarte zu färben. In jüngerer Zeit ist das Interesse an effektiven algorithmischen Lösungen für graphentheoretische Fragestellungen durch komplexe technische Anwendungen stark gestiegen. Die Berechnung von optimalen Routen mit verschiedenen Nebenbedingungen ist hier ein typisches Beispiel.

Endliche Graphen

Wir konzentrieren uns hier auf den einfachsten Graphen-Typ:

Definition *(Graph)*
Ein *(endlicher, ungerichteter, einfacher) Graph* ist ein Paar $G = (E, K)$ mit:

(a) E ist eine endliche nichtleere Menge,

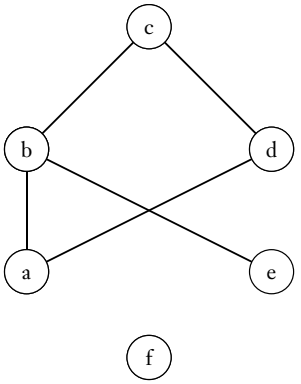
(b) $K \subseteq \{\{a, b\} \mid a, b \in E,\ a \neq b\}$.

Die Elemente von E heißen die *Ecken* von G und die Elemente von K die *Kanten* von G. Gilt $\{a, b\} \in K$, so sagen wir, daß die Ecken a und b durch eine Kante *verbunden* sind. Die Ecke b heißt dann ein *Nachbar* von a in G. Wir schreiben für Kanten oft auch kurz ab anstelle von $\{a, b\}$.
Die Anzahl $|E|$ der Ecken eines Graphen G heißt die *Ordnung* von G und die Anzahl $|K|$ der Kanten von G heißt die *Größe* von G.

Graphen können wir visualisieren, indem wir die Ecken E als benannte Punkte zeichnen, und dann genau die Punktpaare ab mit einer Linie verbinden, die eine Kante des Graphen bilden. Unsere Graphen sind nicht gerichtet, d. h. die Verbindungen sind keine Pfeile. Weiter gibt es keine Schlingen der Form $aa \in K$, und wir erlauben auch keine mehrfachen Verbindungen zwischen zwei Ecken.

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_12,

Ein Beispiel für einen visualisierten Graphen zeigt das folgende Diagramm:



Ecken: a, b, c, d, e, f
Kanten: ab, ad, bc, be, cd
Nachbarn von a: b, d
Nachbarn von b: a, c, e
…
Nachbarn von f: keine

Einige wichtige spezielle Graphen sind die folgenden:

Definition *(die Graphen K_n, $K_{n,m}$, C_n)*
Wir definieren für alle $n, m \geq 1$:

$K_n = (\{1, \ldots, n\}, \{ij \mid 1 \leq i < j \leq n\})$,

$K_{n,m} = (\{1, \ldots, n+m\}, \{ij \mid 1 \leq i \leq n,\ n+1 \leq j \leq n+m\})$,

$C_n = (\{1, \ldots, n\}, \{i(i+1) \mid 1 \leq i < n\} \cup \{n1\})$, falls $n \geq 3$.

Der Leser visualisiere sich diese Graphen anhand von Diagrammen. Zu den Graphen K_n und $K_{n,m}$ gehören die beiden folgenden Begriffsbildungen.

Definition *(vollständig)*
Ein Graph heißt *vollständig*, wenn je zwei seiner Ecken durch eine Kante verbunden sind.

Die Graphen K_n sind vollständig. Der Graph K_n wird oft auch als der *kanonische vollständige Graph* mit n Ecken bezeichnet.

Definition *(bipartit, vollständig bipartit)*
Ein Graph $G = (E, K)$ heißt *bipartit*, wenn sich E derart in zwei Mengen E_1 und E_2 zerlegen läßt, daß jede Kante von G eine Ecke aus E_1 mit einer Ecke aus E_2 verbindet. (Die Mengen E_1, E_2 können leer sein.)
G heißt *vollständig bipartit*, falls eine derartige Zerlegung existiert, sodaß jede Ecke aus E_1 mit jeder Ecke aus E_2 verbunden ist.

Die Graphen $K_{n,m}$ sind vollständig bipartit für die Zerlegung $E_1 = \{1, \ldots, n\}$ und $E_2 = \{n+1, \ldots, n+m\}$ der Eckenmenge $\{1, \ldots, n+m\}$.
Wir definieren:

Definition *(Grad)*
Sei $G = (E, K)$ ein Graph. Dann setzen wir für alle $a \in E$

$N(a) = \{ b \in E \mid ab \in K \}$,

$d(a) =$ „die Anzahl der Elemente von $N(a)$“,

und nennen $d(a)$ den *Grad* der Ecke a in G.

Hierbei steht „d“ für „degree“. Es gilt folgende Summenregel:

Satz *(Gradsumme)*
Sei $G = (E, K)$ ein Graph der Größe k. Dann gilt $\Sigma_{a \in E}\, d(a) = 2 \cdot k$.

Der Satz ist anschaulich klar, denn jede Kante trägt zwei „Einheiten“ zur linken Summe bei. Ein ausführlicher Beweis verläuft wie folgt:

Beweis
Für alle $a \in E$ sei $S(a) = \{ (a, b) \mid ab \in K \}$. Dann ist $d(a)$ die Anzahl der Elemente von $S(a)$, und die Mengen $S(a)$, $a \in E$, sind paarweise disjunkt, da wir hier geordnete Paare verwenden. Damit haben wir

$$\left| \bigcup_{a \in E} S(a) \right| = \Sigma_{a \in E} \, |S(a)| = \Sigma_{a \in E}\, d(a).$$

Andererseits ist

$$\left| \bigcup_{a \in E} S(a) \right| = |\{ (a, b) \mid ab \in K \}| = 2 \cdot |K| = 2 \cdot k.$$

Isomorphe Graphen

Die Namen der Ecken sind für graphentheoretische Fragen unerheblich. Ein Beispiel ist: „Besitzt dieser Graph eine Ecke vom Grad 3?“ Diese Unerheblichkeit wird in folgendem Isomorphiebegriff ausgedrückt:

Definition *(isomorph, Isomorphismus)*
Zwei Graphen $G_1 = (E_1, K_1)$ und $G_2 = (E_2, K_2)$ heißen *isomorph*, falls es eine Bijektion $f : E_1 \to E_2$ gibt, sodaß für alle $a, b \in E_1$ gilt:

$ab \in K_1$ *gdw* $f(a)f(b) \in K_2$.

Jede solche Bijektion f heißt dann ein *Isomorphismus* zwischen G_1 und G_2.

Zwei isomorphe Graphen haben die gleiche Ordnung und Größe, und sie stimmen weiter in allen Eigenschaften überein, die sich mit Hilfe von Ecken und Kanten formulieren lassen.

Ist $G = (E, K)$ ein Graph der Ordnung n mit Eckenmenge $E = \{ a_1, \ldots, a_n \}$, so setzen wir $E' = \{ 1, \ldots, n \}$ und $K' = \{ ij \mid a_i a_j \in K \}$. Dann ist $G' = (E', K')$ isomorph zu G, und die Funktion $f : E \to E'$ mit $f(a_i) = i$ für alle $1 \le i \le n$ ist ein Isomorphismus. Diese Überlegung zeigt, daß es prinzipiell genügt, Graphen zu betrachten, deren Eckenmenge von der Form $\{ 1, \ldots, n \}$ ist.

Kantenzüge, Wege und Kreise

Einen Graphen erkunden wir entlang seiner Kanten. Wir definieren hierzu:

Definition *(Kantenzug, geschlossen, offen, besuchte Ecken und Kanten)*
Sei G = (E, K) ein Graph. Eine Folge $a_0, \ldots, a_n$ in E heißt ein *Kantenzug* der *Länge* n in G *von* a_0 *nach* a_n, falls $a_i a_{i+1} \in K$ für alle $i < n$.
Gilt $a_0 = a_n$, so heißt der Kantenzug *geschlossen*. Andernfalls heißt er *offen*.
Die Ecken $a_0, \ldots, a_n$ heißen die *besuchten Ecken* und die Kanten $a_i a_{i+1}$ die *besuchten Kanten* des Kantenzuges.

Eine einzelne Ecke a gilt immer als Kantenzug der Länge 0.

Im oben dargestellten Graphen ist z. B. a, d, a, b, a ein geschlossener Kantenzug der Länge 4. Er besucht die Ecken a, b, d und die Kanten a d und a b.

Ecken und Kanten können in einem Kantenzug mehrfach besucht werden. Für einfache Besuche führen wir eigene Begriffe ein:

Definition *(Weg, einfacher Kantenzug)*
Ein Kantenzug $a_0, \ldots, a_n$ in einem Graphen heißt ein *Weg*, falls er keine Ecke zweimal besucht, d. h. es gilt $a_i \neq a_j$ für alle $i < j \leq n$. Er heißt ein *einfacher Kantenzug*, falls keine Kante zweimal besucht wird, d. h. es gilt $a_i a_{i+1} \neq a_j a_{j+1}$ für alle $i < j < n$.

Jeder Weg ist offenbar ein einfacher Kantenzug. Die Umkehrung ist i. a. falsch. Im Graphen oben ist zum Beispiel b, c, d, a, b, e ein einfacher Kantenzug, aber kein Weg.

Schließlich definieren wir noch Kreise:

Definition *(Kreis)*
Ein geschlossener Kantenzug $a_0, \ldots, a_n, a_0$ mit $n \geq 2$ in einem Graphen G heißt ein *Kreis* in G mit (n + 1)-vielen Ecken, falls $a_0, \ldots, a_n$ ein Weg ist.
Ein Graph G heißt *kreisfrei*, falls es keinen Kreis in G gibt.
Weiter heißt ein Graph G ein *Kreis*, wenn es einen Kreis $a_0, \ldots, a_n, a_0$ in G gibt, der alle Kanten besucht.

Die kleinsten Kreise eines Graphen sind also bei dieser Definition „Dreiecke" der Form a, b, c, a. Ein Kantenzug a, b, a gilt nicht als Kreis.

Obiger Graph enthält einen Kreis mit 4 Ecken, etwa a, b, c, d, a. Dagegen gibt es keinen Kreis, der die Ecke e enthält.

Die Graphen C_n sind Kreise. Der Graph C_n wird auch als der *kanonische Kreis* mit n Ecken bezeichnet. In C_n ist der Kantenzug 1, …, n, 1 ein Kreis mit n Ecken.

Erreichbarkeit und Zusammenhang

Die wichtigste Relation in einem Graphen ist die Erreichbarkeitsbeziehung:

Definition *(erreichbar)*
Ist G = (E, K) ein Graph, so heißt eine Ecke b *erreichbar* von einer Ecke a, falls es einen Kantenzug $a_0, \ldots, a_n$ in G gibt mit $a_0 = a$ und $a_n = b$.

Die Erreichbarkeit ist, wie man leicht zeigt, eine Äquivalenzrelation auf E. Wir können also definieren:

Definition *(Zusammenhangskomponenten, zusammenhängend)*
Die Äquivalenzklassen der Erreichbarkeitsrelation in einem Graphen G heißen die *Zusammenhangskomponenten* von G. Ein Graph G heißt *zusammenhängend*, falls jede Ecke von jeder anderen aus erreichbar ist.

Ein Graph G = (E, K) ist also genau dann zusammenhängend, wenn die Eckenmenge E die einzige Zusammenhangskomponente von G ist.
Im Umfeld der Erreichbarkeitsrelation definieren wir weiter:

Definition *(Brücke)*
Eine Kante k eines Graphen G heißt eine *Brücke*, wenn das Streichen von k aus G die Anzahl der Zusammenhangskomponenten erhöht.

Leicht zu sehen ist: Eine Kante ab ist genau dann eine Brücke, wenn der Kantenzug a, b der einzige Weg in G von a nach b ist.
In der Sprache der Relationen ist die Erreichbarkeitsrelation nichts anderes als die transitive Hülle der Relation $R = \{ (a, b) \mid ab \in K \} \cup \{ (a, a) \mid a \in E \}$. Damit ist der im Kapitel über Matrizen vorgestellte Algorithmus von Warshall geeignet, die Erreichbarkeitsrelation effektiv zu berechnen, und er liefert insbesondere auch eine Antwort auf die Frage, ob ein gegebener Graph zusammenhängend ist oder nicht. Wir verweisen den interessierten Leser also auf das dritte Kapitel, für das folgende ist die Kenntnis der dortigen Ergebnisse aber nicht notwendig.

Definition *(Abstand)*
Sei G = (E, K) ein zusammenhängender Graph. Für $a, b \in E$ setzen wir

d(a, b) = „die Länge eines kürzesten Kantenzuges von a nach b in G“,

und nennen d(a, b) den *Abstand* von a und b in G.

Der Buchstabe „d“ steht hier für „distance“. Die Funktion $d : G^2 \to \mathbb{N}$ hat die Eigenschaften einer *Metrik*, d. h. für alle $a, b, c \in E$ gilt $d(a, a) = 0$, $d(a, b) = d(b, a)$, sowie die *Dreiecksungleichung* $d(a, c) \leq d(a, b) + d(b, c)$.

Ist G nicht zusammenhängend, so kann man $d(a, b) = \infty$ vereinbaren, falls die Ecken a und b in verschiedenen Zusammenhangskomponenten des Graphen liegen. Unter den üblichen Rechenregeln für ∞ gelten dann immer noch die metrischen Eigenschaften für die Funktion d. Die Zusammenhangskomponente einer Ecke a ist dann gegeben durch $\{ b \in E \mid d(a, b) < \infty \}$.

Eulerzüge

Wir wenden uns nun der Frage zu, ob wir einen Graphen „in einem Zug" zeichnen können. Zusätzlich sollen Start- und Endpunkt übereinstimmen:

Definition *(Eulerzug)*
Ein *Eulerzug* in einem Graphen G = (E, K) ist ein geschlossener Kantenzug $a_0, \ldots, a_n = a_0$ in G derart, daß jede Kante von G genau einmal besucht wird. Ein Graph heißt *Eulersch*, wenn ein Eulerzug existiert.

Ist $a_0, \ldots, a_n = a_0$ ein Eulerzug in G, so ist n die Größe von G und für die Kantenmenge K gilt $K = \{ a_i a_{i+1} \mid i < n \}$. Jeder Kreis C_n, $n \geq 3$, ist offenbar Eulersch. Weiter sind z. B. die vollständigen Graphen K_3 und K_5 sowie die vollständigen bipartiten Graphen $K_{2,2}$ und $K_{2,4}$ Eulersch, nicht aber die Graphen K_2, K_4 und $K_{1,3}$. Hiervon kann man sich leicht überzeugen.

Erstaunlicherweise gibt es ein einfaches Kriterium für die Existenz eines Eulerzuges. Zudem existiert ein schneller und durchsichtiger Algorithmus, mit dessen Hilfe wir Eulerzüge finden können.

Wir beginnen mit folgender Beobachtung:

Satz *(notwendiges Kriterium für die Existenz von Eulerzügen)*
Sei G = (E, K) Eulersch. Dann haben alle Ecken einen geraden Grad.

Beweis
Sei $a_0, a_1, \ldots, a_n = a_0$ ein Eulerzug. Sei e eine von a_0 verschiedene Ecke. Besuchen wir die Ecke e insgesamt n-mal auf dem Eulerzug, so werden dabei genau 2n verschiedene Kanten mit der Ecke e besucht, denn wir laufen in die Ecke bei jedem Besuch hinein und wieder hinaus, und keine Kante wird mehrfach besucht. Da alle Kanten des Graphen besucht werden, gilt also $d(e) = 2n$. Da nun aber auch $a_1, \ldots, a_n, a_0, a_1$ ein Eulerzug ist, zeigt das Argument auch, daß $d(a_0)$ gerade ist. ∎

Existiert ein Eulerzug in G, so ist G zusammenhängend, wobei wir Ecken vom Grad 0 stillschweigend streichen. Damit haben wir zwei notwendige Bedingungen für die Existenz eines Eulerzuges gefunden. Wir wollen nun zeigen, daß diese beiden Bedingungen auch hinreichend sind. Hierzu ist folgendes Lemma hilfreich:

Satz *(Rückkehrlemma)*
Sei G = (E, K) ein Graph, und jede Ecke in G habe einen geraden Grad. Sei a ∈ E mit d(a) > 0. Dann existiert ein einfacher geschlossener Kantenzug positiver Länge in G, der die Ecke a besucht.

Beweis
Wir starten bei a und besuchen solange eine bislang nicht besuchte Kante, bis wir wieder bei der Ecke a ankommen. (Nach dem „rein-raus"-Argument des obigen Satzes können wir in der Tat immer eine neue Kante finden, wenn wir noch nicht bei a angekommen sind.) Wir erhalten einen nach Konstruktion einfachen und geschlossenen Kantenzug positiver Länge, der a besucht. ∎

Damit zeigen wir nun:

Satz *(Kriterium für die Existenz von Eulerzügen)*
Sei G = (E, K) ein zusammenhängender Graph, und jede Ecke habe einen geraden Grad. Dann existiert ein Eulerzug in G.

Beweis
Sei Z ein einfacher geschlossener Kantenzug in G der Länge k. Ist k die Größe von G, so ist Z ein Eulerzug. Andernfalls zeigen wir, daß wir Z zu einem einfachen geschlossenen Kantenzug Z^+ erweitern können, der mehr als k Kanten besucht. Nach endlich vielen Iterationen haben wir dann einen Eulerzug gefunden.

Sei G′ der Graph, der aus G durch Streichen der auf Z besuchten Kanten hervorgeht. Dann hat jede Ecke immer noch einen geraden Grad in G′. Da G zusammenhängend ist, gibt es eine Ecke a von Z und eine Kante ab, die nicht auf Z besucht wird (!). Nach dem Rückkehrlemma existiert ein einfacher und geschlossener Kantenzug

W = a, b, …, a

in G′. Wir fügen nun W in Z an irgendeiner Stelle ein, an der die Ecke a in Z besucht wird. Der so erhaltene Kantenzug Z^+ in G ist einfach und geschlossen, und er besucht mehr als k Kanten. ∎

Der reduzierte Graph G′ ist i. a. nicht mehr zusammenhängend. Das Rückkehrlemma setzt aber keinen Zusammenhang voraus.

Aus dem Beweis gewinnen wir den sog. *Algorithmus von Hierholzer*. Gegeben sei ein zusammenhängender Eulerscher Graph. Wir dürfen annehmen, daß die Eckenmenge des Graphen von der Form E = { 1, …, n } ist. Der Algorithmus konstruiert nun eine Folge von einfachen und geschlossenen Kantenzügen Z_0, …, Z_m in G derart, daß Z_m ein Eulerzug ist.

Algorithmus von Hierholzer

Wir beginnen mit $Z_0 = 1$.

Ist Z_i konstruiert, aber noch kein Eulerzug, so sei a die erste Ecke auf Z_i, von der eine noch unbesuchte Kante wegführt. Wir konstruieren nun einen einfachen, geschlossenen und in a beginnenden Kantenzug W, indem wir immer die kleinste Ecke wählen, zu der eine bislang unbesuchte Kante hinführt. Finden wir keine solche Ecke mehr, so ist W konstruiert. Wir fügen nun W in Z_i an der ersten Stelle des Besuchs der Ecke a ein und erhalten so Z_{i+1}.

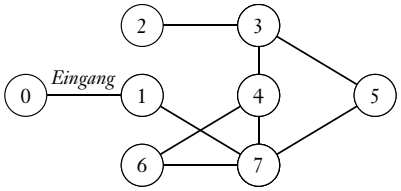
Dieses Vorgehen wird solange iteriert, bis ein Eulerzug gefunden ist.

Obiger Beweis zeigt, daß der Algorithmus tatsächlich einen Eulerzug konstruiert: Der bei a beginnende Kantenzug W im Rekursionsschritt endet automatisch wieder bei a.

Erkundung eines Labyrinths

In einem Labyrinth sehen wir von jedem Raum nur die ablaufenden Gänge, der Bauplan des Labyrinths liegt uns nicht vor. Die Räume eines Labyrinths bilden die Ecken und seine Gänge die Kanten eines zusammenhängenden Graphen. Es stellt sich also die Frage, wie wir einen zusammenhängenden Graphen möglichst schnell vollständig erkunden, wenn unsere Sicht auf den aktuellen Standort beschränkt ist. Wir werden dabei gewisse Kanten zweimal besuchen müssen, denn ist eine Ecke eine Sackgasse (d.h. ihr Grad ist gleich 1), so müssen wir in die Ecke hinein- und auch wieder zurücklaufen.

Es gibt nun ein überraschend einfaches Verfahren, das ein Labyrinth ausgehend von einer Startecke vollständig erkundet und dabei jede Kante genau einmal in jeder Richtung durchläuft und schließlich auch wieder aus dem Labyrinth herausführt. Dabei darf man an den besuchten Kanten gewisse Markierungspfeile anbringen. Es ist aber nicht nötig, während der Erkundung eine Karte des Labyrinths anzufertigen.



Wir legen einen zusammenhängenden Graphen G = (E, K) mit Eckenmenge $E = \{0, 1, \ldots, n\}$ zugrunde, der das Labyrinth modelliert. Zudem sei 0 die Startecke und 1 ihr eindeutiger Nachbar, d.h. die Kante 0 1 ist der Eingang in das Labyrinth. Wir konstruieren nun rekursiv einen geschlossenen Kantenzug $a_0, \ldots, a_m$ mit $a_0 = a_m = 0$ und $m = 2 \cdot |K|$ in G. Dabei werden die Kanten von G mit roten oder gelben Pfeilen markiert. Hierbei steht „rot" für „bereits durchlaufen (in Pfeilrichtung)" und zugleich für „fortan verboten", während die gelben Pfeile eine Sonderrolle spielen, die man mit „nicht durchlaufen, solange noch andere Optionen zur Verfügung stehen" umschreiben kann. Genaueres geht aus dem Algorithmus hervor.

Sagen wir im folgenden, daß a b markiert wird, so zeigt der Markierungspfeil immer von a nach b.

Algorithmus zur Erkundung eines Labyrinths

Wir setzen $a_0 = 0$ und $a_1 = 1$ und markieren 0 1 rot.
Sei nun $a_0, \ldots, a_i$ konstruiert für ein $i \geq 1$. Ist $a_i = 0$, so stoppen wir.
Wird a_i zum ersten Mal besucht, so markieren wir $a_i\, a_{i-1}$ gelb. (Für $i = 1$ wird also insbesondere 10 gelb markiert.) Im Falle der Existenz sei a_{i+1} eine Ecke derart, daß $a_i\, a_{i+1}$ weder gelb noch rot markiert ist; andernfalls sei a_{i+1} die eindeutige Ecke, für die $a_i\, a_{i+1}$ gelb markiert ist. In beiden Fällen markieren wir $a_i\, a_{i+1}$ rot und wiederholen das Verfahren.

Jede durchlaufene Kante wird also sofort in Durchlaufrichtung mit einem roten Pfeil markiert. Wird eine Ecke zum ersten Mal besucht, so wird die Besuchskante in entgegengesetzter Richtung mit einem gelben Pfeil markiert. Rote Kanten dürfen in der Folge unter keinen Umständen in Pfeilrichtung durchlaufen werden, und Kanten mit einem gelben Pfeil werden so lange wie möglich gemieden. Trägt eine Kante einen gelben Pfeil, so trägt sie auch einen roten Pfeil in umgekehrter Richtung. Wir zeigen nun:

Satz *(Korrektheit des Erkundungs-Algorithmus)*

Ist $Z = a_0, \ldots, a_m$ der durch den Algorithmus konstruierte Kantenzug in G, so gilt $m = 2\,|K|$, $a_m = 0$, und für jede Kante $a\,b \in K$ existieren i und j mit $a_i = a$, $a_{i+1} = b$ und $a_j = b$, $a_{j+1} = a$.

Beweis

Sei $Z = a_0, \ldots, a_m$. Wir zeigen zunächst, daß der Kantenzug Z geschlossen ist:

(+) Es gilt $a_m = 0$.

Beweis von (+)

Für $a \neq 0$ gibt es zu Beginn d(a)-viele Möglichkeiten, die Ecke a zu betreten und d(a)-viele Möglichkeiten, sie wieder zu verlassen. Durch jeden Besuch von a werden beide Möglichkeiten jeweils um 1 reduziert. Damit können wir jede besuchte Ecke $a \neq 0$ von Z auch wieder verlassen. Nach Konstruktion stoppt das Verfahren also mit $a_m = 0$.

Wir zeigen nun durch starke Induktion nach $1 \leq i < m$:

(++) a_i wird genau $d(a_i)$-mal auf Z besucht.

Beweis von (++)

Da die gelb markierte Kante 10 nach (+) durchlaufen wird, gilt die Aussage für $a_1 = 1$. Sei nun $i \geq 2$ und die Aussage gelte für alle a_j, $j < i$. Wurde a_i in $a_1, \ldots, a_{i-1}$ besucht, so ist nichts zu zeigen. Andernfalls wird aber $a_i\, a_{i-1}$ gelb markiert. *Annahme*, a_i wird nicht $d(a_i)$-mal besucht. Dann wird die gelb markierte Kante $a_i\, a_{i-1}$ nach den Regeln des Algorithmus nicht durchlaufen, und damit wird auch a_{i-1} nicht $d(a_{i-1})$-mal besucht, *Widerspruch*.

Nach Konstruktion von Z genügt es zu zeigen:

(+++) Jede Ecke a von G wird von Z besucht.

Beweis von (+++)

Wegen G zusammenhängend gibt es einen Weg $b_0, \ldots, b_k$ von 1 nach a. Wir zeigen durch Induktion nach $j \leq k$, daß b_j von Z besucht wird. Dies ist klar für $b_0 = 1$. Wird nun aber b_j für ein $j < k$ von Z besucht, so wird b_j nach (++) sogar $d(b_j)$-oft von Z besucht, und alle diese Besuche führen auf verschiedenen Kanten von b_j weg. Also ist $b_j b_{j+1}$ eine besuchte Kante von Z, und damit wird also auch b_{j+1} von Z besucht.

– Aus (++) und (+++) folgt $m = (\Sigma_{a \in E - \{0\}} d(a)) + 1 = \Sigma_{a \in E} d(a) = 2 |K|$.

Der Beweis von (+) benötigt die gelben Markierungspfeile nicht. Jede Erkundung eines Labyrinths, die es vermeidet, eine Kante in derselben Richtung zweimal zu durchlaufen, führt aus dem Labyrinth wieder heraus. Die gelben Markierungspfeile stellen sicher, daß die gesamte Zusammenhangskomponente der Startecke 0 erkundet wird. Kommt es uns nur darauf an, einen Schatz im Labyrinth zu finden, so können wir die gelben Pfeile zurücklaufen, sobald der Schatz gefunden ist.

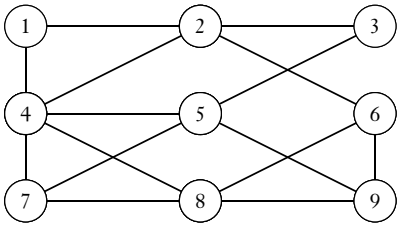
Unser Verfahren zur Erkundung eines Labyrinths läßt sich auch für „Eulersche Fragestellungen“ verwenden: Ein Postbote wird ausgehend vom Postamt gerne jede Straße je einmal in jeder Richtung durchlaufen, da er dann die Post für beide Straßenseiten leichter verteilen kann. Weiter will er am Ende wieder am Postamt ankommen. Obiger Satz zeigt, daß diese Ansprüche für jeden beliebigen zusammenhängenden Graphen erfüllt werden können, ganz ohne eine Gradbedingung.

Hamiltonkreise

Wir stellen nun die zur Existenz von Eulerzügen analoge Frage, wann und wie man geschlossene Kantenzüge findet, die jede Ecke des Graphen genau einmal besuchen:

Definition *(Hamiltonkreis)*

Sei G ein Graph der Ordnung n. Ein Kreis mit n Ecken in G heißt ein *Hamiltonkreis*. G heißt *Hamiltonsch*, falls ein Hamiltonkreis in G existiert.



In diesem Graphen ist
1, 2, 3, 5, 9, 6, 8, 7, 4, 1
ein Hamiltonkreis.

War das Problem der Existenz von Eulerzügen überraschend einfach, so ist nun das Problem der Existenz von Hamiltonkreisen überraschend schwierig. Es scheint keinen schnellen Algorithmus zu geben, der einen gegebenen Graphen daraufhin überprüft, ob er Hamiltonsch ist oder nicht. Ebenso scheint es kein einfaches notwendiges und hinreichendes Kriterium dafür zu geben, wann ein Graph Hamiltonsch ist. Hinreichend ist die folgende Reichhaltigkeitsbedingung:

Satz *(Satz von Gabriel Dirac)*
Sei $G = (E, K)$ ein Graph der Ordnung $n^* \geq 3$, und es gelte $d(a) \geq n^*/2$ für alle $a \in E$. Dann ist G Hamiltonsch.

Beweis
Sei $W = a_0, \ldots, a_n$ ein maximaler Weg in G, d. h. ein Weg in G, den wir nach links und nach rechts nicht mehr fortsetzen können. Dann werden alle Nachbarecken von a_0 und von a_n auf W besucht, da wir sonst den Weg verlängern könnten. Insbesondere ist $n \geq d(a_0) \geq n^*/2 \geq 3/2$. Andererseits ist $n^* \geq n + 1$, da W ein Weg ist. Weiter gilt:

(+) Es gibt ein $i < n$, sodaß $C = a_0, \ldots, a_i, a_n, a_{n-1}, \ldots, a_{i+1}, a_0$ ein Kreis ist.

(Im Falle $i = n - 1$ ist $C = a_0, \ldots, a_n, a_0$ ein Kreis.)

Beweis von (+)
Gesucht ist ein $i < n$, sodaß a_i mit a_n und a_{i+1} mit a_0 in G verbunden ist. Wir setzen hierzu

$$A = \{ i < n \mid a_i a_n \in K \}, \quad B = \{ i < n \mid a_{i+1} a_0 \in K \}.$$

Dann gilt $|A|, |B| \geq n^*/2 > n/2$, da alle Nachbarecken von a_n und a_0 besucht werden. Wegen $A \cup B \subseteq \{ 0, \ldots, n - 1 \}$ gilt dann aber

$$|A \cap B| = |A| + |B| - |A \cup B| > n/2 + n/2 - n = 0,$$

also ist $A \cap B$ nichtleer.

Ist $n + 1 = n^*$, so haben wir mit C einen Hamiltonkreis gefunden. Andernfalls gilt:

(++) Ist a eine auf C nicht besuchte Ecke, so gibt es ein $j \leq n$ mit $aa_j \in K$.

Beweis von (++)
Andernfalls gilt $\{ a_0, \ldots, a_n \} \cap N(a) = \emptyset$, und wegen $n \geq n^*/2$ gilt dann

$$n^* = |E| \geq (n + 1) + d(a) \geq (n^*/2 + 1) + n^*/2 = n^* + 1,$$

Widerspruch.

Aus (++) erhalten wir durch Aufschneiden des Kreises C an der Ecke a_j und Anfügen von a einen Weg in G, der länger ist als W. Diesen Weg können wir zu einem maximalen Weg fortsetzen, und dann liefert (+) wieder einen entsprechend langen Kreis. Nach endlicher Iteration dieses Verfahrens haben wir dann einen Hamiltonkreis von G gefunden. ∎

Der Beweis liefert einen effizienten „Algorithmus von Dirac“, der einen Hamiltonkreis in Graphen findet, die der Reichhaltigkeitsbedingung des Satzes genügen. Der Leser ist aufgefordert, diesen Algorithmus explizit zu notieren.

Der Satz von Gabriel Dirac (1952) kann noch zum Satz von Øystein Ore (1960) verbessert werden: Ein Graph $G = (E, K)$ der Ordnung $n^* \geq 3$ ist Hamiltonsch, falls $d(a) + d(b) \geq n^*$ für alle $a, b \in E$ mit $a \neq b$ und $ab \notin K$ gilt. Noch stärker ist der Satz von Bondy-Chvátal (1972): Sei $G = (E, K)$ ein Graph der Ordnung $n^* \geq 3$, und seien $a, b \in E$ mit $a \neq b$, $ab \notin K$ und $d(a) + d(b) \geq n^*$. Dann ist G genau dann Hamiltonsch, wenn $(E, K \cup \{ ab \})$ Hamiltonsch ist.

Übungen

Übung 1 *(Endliche Graphen, I)*

Sei $G = (E, K)$ ein Graph. Sei $U = \{ a \in E \mid d(a) \text{ ist ungerade} \}$.
Zeigen Sie, daß $|U|$ eine gerade Zahl ist.

Übung 2 *(Endliche Graphen, II)*

Ein Graph $G = (E, K)$ heißt *planar*, wenn er in der Ebene so gezeichnet werden kann, daß sich keine Kanten überschneiden. Für einen planaren Graphen sei e die Anzahl der Ecken, k die Anzahl der Kanten und f die Anzahl der Flächen, wobei die äußere Fläche mitzählt. (Z. B. hat ein Graph in der Form einer Acht drei Flächen.)
Zeigen Sie, daß für jeden planaren Graphen gilt:

$$e - k + f = 2. \qquad \textit{(Eulersche Polyederformel)}$$

[Wir beweisen die Aussage zuerst für Graphen mit $f = 1$. Die allgemeine Aussage beweisen wir dann durch Induktion nach f.]

Folgern Sie, daß es nur 5 regelmäßige konvexe Polyeder im $\mathbb{R}^3$ gibt (die *Platonischen Körper*): ein Tetraeder mit 4, einen Kubus mit 6, ein Oktaeder mit 8, ein Dodekaeder mit 12, und ein Ikosaeder mit 20 Flächen. (Dabei heißt ein $P \subseteq \mathbb{R}^3$ *konvex*, falls für alle $a, b \in P$ die Strecke von a nach b eine Teilmenge von P ist.)

[Seien e, k, f die Anzahl der Ecken, Kanten, Flächen eines Polyeders. Für ein regelmäßiges Polyeder sei n die Zahl der Kanten einer Fläche und d der Grad der Ecken. Begründen Sie, daß die Eulersche Polyederformel gilt, und folgern Sie für regelmäßige Polyeder die Beziehung

$$(+) \quad 1/n + 1/d = 1/2 + 1/k,$$

die sich aus der Eulerschen Polyederformel und $nf = 2k$, $de = 2k$ ergibt. Bestimmen Sie die möglichen Lösungen von (+) mit $n \geq 3$ und $d \geq 3$, nämlich

$$(n, d, k) = (3, 3, 6),\ (3, 4, 12),\ (3, 5, 30),\ (4, 3, 12),\ (5, 3, 30).$$

Beobachten Sie hierzu, daß $n \leq 3$ oder $d \leq 3$ gelten muß, falls (+) mit positiven Zahlen (n, d, k) erfüllt wird.].

Übung 3 *(Endliche Graphen, III)*
Zeigen Sie:

(a) Ist $f : G_1 \to G_2$ ein Isomorphismus zwischen zwei Graphen, so gilt $d_{G_1}(a) = d_{G_2}(f(a))$ für alle $a \in E_1$.

(b) Skizzieren Sie alle Graphen für die Eckenmenge $E = \{ 1, 2, 3 \}$. Welche dieser Graphen sind isomorph? Welche Graphen sind aus welchen Gründen nicht isomorph?

(c) Geben Sie zwei nichtisomorphe Graphen mit den Ecken $\{ 1, \ldots, 5 \}$ an, sodaß die Grade beider Graphen die Zahlen 2, 2, 2, 3, 3 sind.

Übung 4 *(Endliche Graphen, IV)*
Ein Graph $G = (E, K)$ heißt *selbstkomplementär*, falls G isomorph zu seinem komplementären Graphen $G^c = (E, \{ ab \mid a, b \in E, a \neq b, ab \notin K \})$ ist.

(a) Geben Sie je einen selbstkomplementären Graphen an mit den Ecken $E_1 = \{ 1, 2, 3, 4 \}$ und $E_2 = \{ 1, 2, 3, 4, 5 \}$.

(b) Zeigen Sie: Ist G selbstkomplementär, so ist die Anzahl seiner Ecken durch 4 ohne Rest oder durch 4 mit Rest 1 teilbar.

Übung 5 *(Kantenzüge, Wege und Kreise, I)*
Sei $a_0, \ldots, a_n$ ein einfacher geschlossener Kantenzug in einem Graphen G. Zeigen Sie, daß $i < j \leq n$ existieren, sodaß $a_i, a_{i+1}, \ldots, a_j$ ein Kreis ist. Zeigen Sie weiter, daß für jedes a_i, $i \leq n$, ein Kreis in G durch a_i existiert. Kann hier auf die Voraussetzung „einfach" verzichtet werden?

Übung 6 *(Kantenzüge, Wege und Kreise, II)*
Sei $G = (E, K)$ ein Graph, und seien $k_1, k_2, k_3 \in K$ paarweise verschieden. Es gebe einen Kreis, der k_1 und k_2 besucht, und einen Kreis, der k_2 und k_3 besucht. Zeigen Sie, daß es einen Kreis gibt, der k_1 und k_3 besucht.

Übung 7 *(Kantenzüge, Wege und Kreise, III)*
Sei $G = (E, K)$ ein Graph. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) G ist bipartit. (b) Jeder Kreis in G hat eine gerade Länge.

Inbesondere ist also jeder kreisfreie Graph bipartit.

Übung 8 *(Erreichbarkeit und Zusammenhang, I)*
Sei $G = (E, K)$ ein Graph, und seien $a, b \in E$. Zeigen Sie:
Ist b erreichbar von a, so existiert ein Weg von a nach b.

Übung 9 *(Erreichbarkeit und Zusammenhang, II)*
Sei $G = (E, K)$ ein Graph. Für $a, b \in E$ sei $a \equiv b$, falls b erreichbar von a ist. Zeigen Sie, daß $\equiv$ eine Äquivalenzrelation auf E ist. Genauer ist $\equiv$ die $\subseteq$-kleinste Äquivalenzrelation R auf E mit aRb für alle $ab \in K$.

Übung 10 *(Erreichbarkeit und Zusammenhang, III)*
Sei G = (E, K) ein zusammenhängender Graph. Zeigen Sie, daß die Abstandsfunktion $d : G^2 \to \mathbb{N}$ eine Metrik ist, d. h. es gilt für alle $a, b, c \in E$:

(a) $d(a, a) = 0$,

(b) $d(a, b) = d(b, a)$,

(c) $d(a, c) \leq d(a, b) + d(b, c)$.

Übung 11 *(Erreichbarkeit und Zusammenhang, IV)*
Sei G = (E, K) ein Graph. Zeigen Sie, daß für alle $ab \in K$ die folgenden Aussagen äquivalent sind:

(a) ab ist eine Brücke von G.

(b) a, b ist der einzige Weg von a nach b.

Übung 12 *(Erreichbarkeit und Zusammenhang, V)*
Sei G ein Graph, und jede Ecke von G habe einen geraden Grad.
Zeigen Sie, daß G keine Brücken besitzt.

Übung 13 *(Erreichbarkeit und Zusammenhang, VI)*
Für einen Graphen G = (E, K) sei

$$G^c = (E, \{ ab \mid a, b \in E, a \neq b, ab \notin K \}),$$

der zu G *komplementäre Graph*.
Zeigen Sie, daß der Graph G^c eines unzusammenhängenden Graphen G zusammenhängend ist.

Übung 14 *(Erreichbarkeit und Zusammenhang, VII)*
Sei G ein zusammenhängender Graph der Ordnung n derart, daß jede Ecke von G den Grad 2 besitzt. Zeigen Sie:

G ist isomorph zum Kreis C_n.

Übung 15 *(Eulerzüge, I)*
Sei G ein Graph, und seien a, b verschiedene Ecken von G. Ein *offener Eulerzug* von a nach b in G ist ein einfacher in a beginnender und in b endender Kantenzug in G, der alle Kanten durchläuft.

Zeigen Sie, daß die folgenden Aussagen für jeden zusammenhängenden Graphen G = (E, K) und alle $a, b \in E$, $a \neq b$, äquivalent sind:

(a) Es gibt einen offenen Eulerzug von a nach b.

(b) Die Grade d(a), d(b) sind ungerade, und für alle $c \in E - \{ a, b \}$ ist der Grad d(c) gerade.

Übung 16 *(Eulerzüge, II)*
Zeigen Sie, daß die folgenden Aussagen für jeden zusammenhängenden Graphen G = (E, K) äquivalent sind:

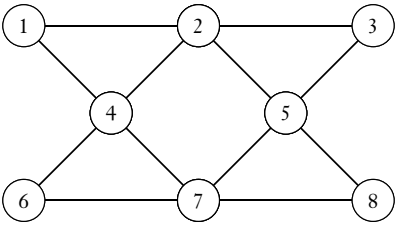
(a) G ist Eulersch.

(b) G ist die Vereinigung von Kreisen $K_1, \ldots, K_n$ in G mit paarweise disjunkten Kanten.

Übung 17 *(Eulerzüge, III)*
Sei G = (E, K) ein zusammenhängender Graph. Zeigen Sie, daß es einen geschlossenen Kantenzug gibt, der jede Kante genau zweimal durchläuft.

Übung 18 *(Eulerzüge, IV)*
Betrachten Sie ein Dominospiel mit Zahlenpaaren von 0, …, 6 auf den Steinen, ohne die Steine mit gleichen Zahlen. (Das Spiel hat dann also 21 Steine.) Das Dominoproblem lautet: Kann man die Steine so in einer geschlossenen Kette anordnen, daß, wie beim Domino üblich, aneinanderliegende Steine dort, wo sie sich berühren, stets die gleiche Zahl aufweisen? Klären Sie dieses Problem in der Sprache der Graphentheorie. Geben Sie im Falle der Existenz eine konkrete Lösung an. Für welche n gibt es eine Lösung, wenn auf den Dominosteinen die Zahlen 0, 1, …, n erscheinen?

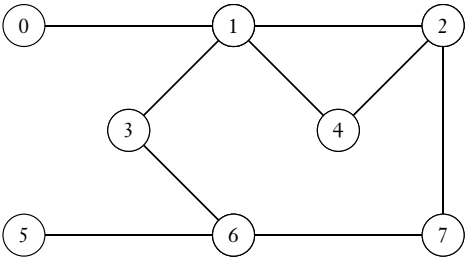
Übung 19 *(Eulerzüge, V)*
Bestimmen Sie mit dem Algorithmus von Hierholzer einen Eulerzug für den folgenden Graphen:



Übung 20 *(Eulerzüge, VI)*
Was findet der Algorithmus von Hierholzer, wenn er auf einen nicht notwendig zusammenhängenden Graphen angewendet wird, dessen Ecken alle einen geraden Grad haben? Was kann passieren, wenn der Algorithmus auf einen beliebigen Graphen angewendet wird?

Übung 21 *(Erkundung eines Labyrinths, I)*
Wie verläuft der Erkundungsalgorithmus, wenn die Startkante 0 1 in einen Kreis C_n hineinführt?

Übung 22 (Erkundung eines Labyrinths, II)
Führen Sie den Erkundungsalgorithmus für den folgenden Graphen durch. Wählen Sie dabei im Falle mehrerer möglicher nächster Ecken immer die kleinste mögliche Ecke.



Übung 23 (Hamiltonkreise, I)
Zeigen Sie, daß die Graphen der fünf regelmäßigen Platonischen Körper aus Übung 2 Hamiltonsch sind.

Übung 24 (Hamiltonkreise, II)
Sei G = (E, K) ein Graph der Ordnung n* mit $d(a) \geq n^*/2$ für alle $a \in E$. Zeigen Sie ohne Verwendung des Satzes von Dirac, daß G zusammenhängend ist. Kann die Schranke $n^*/2$ noch verbessert werden?

6. Wahrscheinlichkeiten

Die mathematische Zähmung des intuitiven Wahrscheinlichkeitsbegriffs ist das Thema dieses Kapitels. Wir sagen beim Würfeln: „Die Wahrscheinlichkeit, mit einem Würfel eine 1 zu würfeln, ist 1/6." oder „Die Wahrscheinlichkeit, mit einem Paar Würfel eine 8 in der Summe zu würfeln, ist 5/36." In der Physik reden wir von Aufenthaltswahrscheinlichkeiten eines Elektrons in einem Orbital. Und ein Blick auf die Wettervorhersage informiert uns: „Die Niederschlagswahrscheinlichkeit beträgt morgen in Hamburg 90 Prozent".

Denken wir über die Verwendung des Begriffs der Wahrscheinlichkeit in diesen Beispielen nach, so tritt schnell die Frage auf, ob und wann so etwas wie eine „echte" Wahrscheinlichkeit vorliegt oder ob unsere Wahrscheinlichkeitsaussagen nur aufgrund unserer Unkenntnis der Situation entstehen – „Gott würfelt nicht". Diese Fragen können und wollen wir hier nicht berühren. Unsere mathematischen Wahrscheinlichkeitsmaße sind einfach bestimmte Funktionen, deren Eigenschaften unserer Intuition über den Begriff entstammen. Die Mathematik würfelt nicht, aber sie kann den Wurf eines Würfels modellieren.

Abzählbare Wahrscheinlichkeitsräume

Einfache Zufallsexperimente können wir durch eine Funktion $\nu : A \to [0, 1]$ modellieren, die den Elementen a einer gewissen abzählbaren Menge A von „Stichproben" oder „Ergebnissen" einen Wert $\nu(a)$ so zuweist, daß sich alle Werte zu 1 aufsummieren. Der Wert $\nu(a)$ heißt dann die Wahrscheinlichkeit von a bei ν. Allgemeiner können wir dann auch von der Wahrscheinlichkeit $\mu(B)$ eines beliebigen „Ereignisses" $B \subseteq A$ reden, indem wir alle $\nu(a)$, $a \in B$, aufsummieren. Der Wurf eines fairen Würfels kann in dieser Weise durch $A = \{1, \ldots, 6\}$ und $\nu(a) = 1/6$ für alle $a \in A$ modelliert werden. Dann gilt $\mu(\text{„die Augenzahl ist gerade"}) = \mu(\{2, 4, 6\}) = \nu(2) + \nu(4) + \nu(6) = 1/2$, usw.

Wir definieren:

Definition *(abzählbare Verteilungen der Eins)*
Sei A eine abzählbare Menge. Eine Funktion $\nu : A \to [0, 1] \subseteq \mathbb{R}$ heißt eine *Verteilung der Eins* auf A, falls $\Sigma_{a \in A}\, \nu(a) = 1$.

Hier und im folgenden verwenden wir, daß wir abzählbar unendlich viele nichtnegative reelle Zahlen in beliebiger Reihenfolge anordnen können, ohne

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2_13, © Springer-Verlag Berlin Heidelberg 2010

daß dies das Konvergenzverhalten und den Grenzwert der zugehörigen unendlichen Reihe verändern würde. Ist nämlich $\Sigma_{n\in\mathbb{N}}\, x_n$ eine konvergente Reihe reeller Zahlen mit $x_n \geq 0$ für alle $n \in \mathbb{N}$, so gilt für jede Bijektion $\pi : \mathbb{N} \to \mathbb{N}$, daß

$$\Sigma_{n\in\mathbb{N}}\, x_n = \sup(\{\, \Sigma_{n\in E}\, x_n \mid E \subseteq \mathbb{N} \text{ endlich} \,\}) = \sup(\{\, \Sigma_{n\in F}\, x_{\pi(n)} \mid F \subseteq \mathbb{N} \text{ endlich} \,\}) = \Sigma_{n\in\mathbb{N}}\, x_{\pi(n)}.$$

Ist also X eine abzählbar unendliche Menge nichtnegativer Zahlen, so können wir $\Sigma_{x\in X}\, x$ definieren als $\Sigma_{n\in\mathbb{N}}\, f(n)$ für eine beliebige Bijektion $f : \mathbb{N} \to X$, vorausgesetzt, die Reihe $\Sigma_{n\in\mathbb{N}}\, f(n)$ konvergiert. Ist g eine reellwertige Funktion auf einer beliebigen abzählbar unendlichen Menge A mit $g(x) \geq 0$ für alle $x \in A$, so ist analog $\Sigma_{a\in A}\, g(a)$ definiert als $\Sigma_{n\in\mathbb{N}}\, g(\pi(n))$ mit einer beliebigen Bijektion $\pi : \mathbb{N} \to A$.

Eine Verteilung der Eins auf einer Menge induziert nun das folgende Gewicht von beliebigen Teilmengen der Menge:

Definition *(abzählbares Wahrscheinlichkeitsmaß)*
Sei ν eine Verteilung der Eins auf der abzählbaren Menge A.
Für alle $B \subseteq A$ setzen wir:

$$\mu(B) = \Sigma_{b\in B}\, \nu(b).$$

Dann heißt $\mu : \mathcal{P}(A) \to [0, 1]$ das *durch* ν *induzierte (Wahrscheinlichkeits-) Maß* auf A und $(A, \mathcal{P}(A), \mu)$ ein *(abzählbarer) Wahrscheinlichkeitsraum*. Die Menge A heißt *Grundmenge* oder *Ergebnisraum* und ihre Elemente nennen wir *Elementarereignisse* oder *Stichproben*. Die Menge $\mathcal{P}(A)$ heißt *Ereignisraum* und für jedes $B \in \mathcal{P}(A)$ heißt $\mu(B)$ die μ-*Wahrscheinlichkeit* des Ereignisses B.

Statt $(A, \mathcal{P}(A), \mu)$ schreiben wir oft einfach auch (A, μ). Bei dieser Schreibweise ist A immer als abzählbar vorausgesetzt. Zur Definition von (A, μ) genügt es, eine abzählbare Grundmenge A und eine Verteilung ν der Eins auf A anzugeben.

Einige Beispiele und Konstruktionsmethoden für Wahrscheinlichkeitsräume diskutieren wir in den folgenden Zwischenabschnitten sowie in den Übungen.

Einfache Modellbildungen

Einen fairen Münzwurf können wir durch $A = \{ 0, 1 \}$ und $\nu(0) = \nu(1) = 1/2$ modellieren, wobei 0 für „Kopf“ und „1“ für „Zahl“ steht. Ist μ das von ν induzierte Wahrscheinlichkeitsmaß auf $\mathcal{P}(A)$, so gilt

$$\mu(\varnothing) = 0,\ \mu(\{ 0 \}) = \mu(\{ 1 \}) = 1/2,\ \mu(\{ 0, 1 \}) = 1.$$

Ist $A = \{ 1, \ldots, 6 \}$, so induziert die Verteilung $\nu : A \to [0, 1]$ mit $\nu(a) = 1/6$ für alle $a \in A$ einen Wahrscheinlichkeitsraum (A, μ), den wir als ein geeignetes mathematisches Modell für den Wurf eines ungezinkten Würfels ansehen. Es gilt dann $\mu(\{ 1, 2 \}) = 1/3$, $\mu(A - \{ 1 \}) = 5/6$, usw. Allgemeiner eignen sich alle fünf platonischen Körper für „würfelnde“ Zufallsexperimente, deren mathematische Verteilungen durch $\nu_4(a) = 1/4$, $\nu_6(a) = 1/6$, $\nu_8(a) = 1/8$, $\nu_{12}(a) = 1/12$ und $\nu_{20}(a) = 1/20$ für alle Elementarereignisse a bestimmt sind.

Wirft man eine Münze und einen Würfel, so ist $A = \{0, 1\} \times \{1, \ldots, 6\}$ ein geeigneter Ergebnisraum. Sind dabei Münze und Würfel fair, so induziert die Verteilung „$\nu(a) = 1/12$ für alle $a \in A$“ den passenden Wahrscheinlichkeitsraum. Der Wurf eines Dodekaeders läßt sich also durch dieses Experiment simulieren.

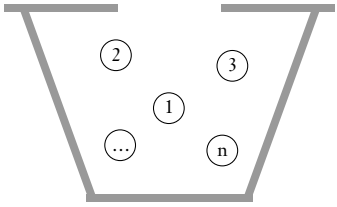
Werfen wir zwei faire Würfel und betrachten die Summe der Augen, so ist, wie man sich leicht überlegt, die Ergebnismenge $S = \{2, \ldots, 12\}$ und die folgende Verteilung ν zur Modellierung geeignet:

$\nu(2) = \nu(12) = 1/36$, $\nu(3) = \nu(11) = 2/36$, $\nu(4) = \nu(10) = 3/36$,

$\nu(5) = \nu(9) = 4/36$, $\nu(6) = \nu(8) = 5/36$, $\nu(7) = 6/36$.

Der Ansatz hinter diesen Berechnungen ist, den zweifachen Würfelwurf zunächst durch die Ergebnismenge $A = \{1, \ldots, 6\}^2$ und die Verteilung $\nu(a_1, a_2) = 1/36$ für alle $(a_1, a_2) \in A$ zu beschreiben und dann das „Zufallsverhalten“ der Funktion $f : A \to \mathbb{R}$ mit $f(a_1, a_2) = a_1 + a_2$ für alle $(a_1, a_2) \in A$ zu untersuchen. Diese Überlegung führt zu den sogenannten Zufallsvariablen, die wir unten genauer betrachten.

Urnenmodelle



Eine ganze Familie von Zufallsexperimenten liefert das Ziehen aus einer Urne. Wir ziehen nacheinander k Kugeln aus einer Urne mit n Kugeln, die mit den Zahlen 1, 2, …, n beschriftet sind. Hier ist zu unterscheiden, ob die Kugeln nach dem Ziehen zurückgelegt werden und ob die Reihenfolge für das Ergebnis der Ziehung eine Rolle spielt. Wir diskutieren die vier sich ergebenden Varianten in den Übungen. Dabei spielen die folgenden Werte eine wichtige Rolle:

$|A|$ = „die Anzahl der Elemente von A“, *(Betrag oder Mächtigkeit von A)*

$n!$ = $1 \cdot \ldots \cdot n$, mit $0! = 1$, *(n-Fakultät)*

$\binom{n}{k}$ = $n!/(k! \cdot (n-k)!)$, *(Binomialkoeffizienten, „n über k“, „k aus n“)*

$\binom{n}{k_1, \ldots, k_r}$ = $n!/(k_1! \cdot \ldots \cdot k_r!)$, *(Multinomialkoeffizienten, „n über $k_1, \ldots, k_r$“)*

die für alle endlichen Mengen A und alle $n, k, k_1, \ldots, k_r \in \mathbb{N}$ mit $0 \le k \le n$ und $k_1 + \ldots + k_r = n$ definiert sind.

Die Binomialkoeffizienten geben an, auf wieviele Arten es möglich ist, k Elemente aus einer Menge mit n Elementen auszuwählen, d.h. $\binom{n}{k}$ ist die Anzahl der k-elementigen Teilmengen von $\{1, \ldots, n\}$. Da es genau 2^n Teilmengen der Menge $\{1, \ldots, n\}$ gibt, gilt $\Sigma_{0 \le k \le n} \binom{n}{k} = 2^n$ für alle n.

Analog ist $\binom{n}{k_1, \ldots, k_r}$ die Anzahl der Zerlegungen der Menge $\{1, \ldots, n\}$ in Mengen $A_1, \ldots, A_r$, die genau $k_1, \ldots, k_r$ Elemente enthalten, d.h. es gilt

$$\binom{n}{k_1, \ldots, k_r} = |\{(A_1, \ldots, A_r) \mid A_i \cap A_j = \varnothing \text{ für } i \ne j,\ \bigcup_{1 \le i \le r} A_i = \{1, \ldots, n\},\ |A_i| = k_i \text{ für alle } i\}|.$$

Damit gibt es also $\binom{49}{6}$ Ergebnisse einer Lottoziehung und $\binom{32}{10,10,2}$ Skatspiele, also Möglichkeiten, 32 Karten auf drei Spieler I, II, III mit je 10 Karten und einen Skat mit zwei Karten zu verteilen.

Die Multinomialkoeffizienten verallgemeinern die Binomialkoeffizienten, denn es gilt $\binom{n}{k} = \binom{n}{k,(n-k)}$ für alle $0 \le k \le n$. In der Tat entsprechen die k-elementigen Teilmengen von $\{ 1, \ldots, n \}$ den Zerlegungen dieser Menge in zwei Teile mit k und n – k Elementen.

Das *Pascalsche Dreieck* liefert eine Möglichkeit der rekursiven Berechnung der Binomialkoeffizienten. Die äußeren Werte jeder Zeile sind 1, und jeder andere Wert einer Zeile ist die Summe der beiden über ihm stehenden Werte. Die (n + 1)-te Zeile listet dann die Binomialkoeffizienten $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{n}$ auf.

```
                1
              1   1
            1   2   1
          1   3   3   1
        1   4   6   4   1
      1   5  10  10   5   1
```

Die Bezeichnung „Binomial-“ und „Multinomialkoeffizienten“ ist motiviert durch

$$(x + y)^n = \Sigma_{0 \le k \le n} \binom{n}{k} x^k y^{n-k}, \qquad \textit{(Binomialsatz)}$$

$$(x_1 + \ldots + x_r)^n = \Sigma_{0 \le k_i \le n,\, k_1 + \ldots + k_r = n} \binom{n}{k_1, \ldots, k_r} x_1^{k_1} \cdot \ldots \cdot x_r^{k_r}. \qquad \textit{(Multinomialsatz).}$$

Binomial- und Multinomialverteilungen

Obige Anzahlaussagen können wir auch so formulieren: $\binom{n}{k}$ ist die Anzahl der 0-1-Tupel $(a_1, \ldots, a_n)$, die genau k Einsen aufweisen. Denn hat $A \subseteq \{ 1, \ldots, n \}$ genau k Elemente, so formen wir das 0-1-Tupel $(a_1, \ldots, a_n)$ mit $a_i = 1$ für $i \in A$ und $a_i = 0$ sonst. Dadurch entsteht eine Bijektion zwischen den k-elementigen Teilmengen von $\{ 1, \ldots, n \}$ und den 0-1-Tupeln $(a_1, \ldots, a_n)$ mit genau k Einsen. Analog ist $\binom{n}{k_1, \ldots, k_r}$ die Anzahl der Tupel $(a_1, \ldots, a_n)$ mit Einträgen in $\{ 1, \ldots, r \}$, die für alle $1 \le i \le r$ genau k_i-oft die Zahl i aufweisen.

Definieren wir also für ein $0 \le p \le 1$ und ein $n \ge 1$

$$b_n^p(k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{für alle } 0 \le k \le n,$$

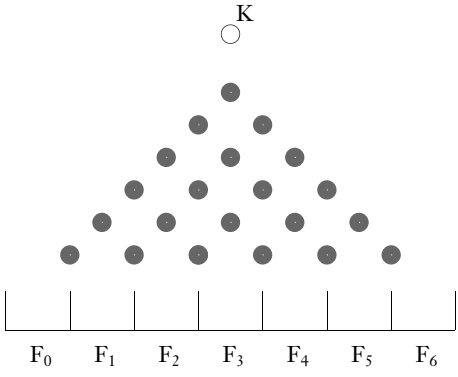
so ist $b_n^p(k)$ die Wahrscheinlichkeit, in einem n-fach wiederholten Zufallsexperiment genau k Erfolge zu erzielen, wenn p die Erfolgswahrscheinlichkeit ist. Die Verteilung b_n^p auf $\{ 0, \ldots, n \}$ heißt die *Binomialverteilung* für n und p.

Interessieren wir uns für r verschiedene Ergebnisse $a_1, \ldots, a_r$ eines Experiments, und hat jedes Ergebnis a_i die Wahrscheinlichkeit p_i, so ist analog

$$b_n^{p_1, \ldots, p_r}(k_1, \ldots, k_r) = \binom{n}{k_1, \ldots, k_r} p_1^{k_1} \ldots p_r^{k_r} \quad \text{für alle } 0 \le k_i \le n \text{ mit } k_1 + \ldots + k_r = n$$

die Wahrscheinlichkeit, bei einer n-fachen Wiederholung des Experiments genau k_i-oft das Ergebnis a_i zu erhalten, für alle $1 \le i \le r$. Die Verteilung $b_n^{p_1, \ldots, p_r}$ heißt die *Multinomialverteilung* für n und $p_1, \ldots, p_r$.

Sind also die Tageswahrscheinlichkeiten für „Regen, bewölkt, Sonne“ gleich 1/10, 1/5 bzw. 7/10, so ist die Wahrscheinlichkeit für eine Woche mit einem Regen-, zwei bewölkten und vier Sonnentagen gleich $\binom{7}{1,2,4} \cdot 1/10 \cdot (2/10)^2 \cdot (7/10)^4 = (105 \cdot 4 \cdot 7^4)/10^7 = 0{,}1008421$, also etwa 10%.



Die Binomialverteilung (und das Pascalsche Dreieck) taucht auch im folgenden Zufallsexperiment auf, dem sog. *Galton-Brett*: Eine Kugel K fällt in einem wie im Diagramm rechts angeordneten Nagelbrett stufenweise nach unten, und zwar jeweils mit Wahrscheinlichkeit p nach rechts und mit Wahrscheinlichkeit q = (1 – p) nach links. Dann wird die Wahrscheinlichkeit, daß die Kugel in einem der am Ende des Brettes angebrachten Fächer $F_0, \ldots, F_n$ landet, durch die Binomialverteilung b_n^p beschrieben. In unserem Diagramm ist n = 6 und die Wahrscheinlichkeit, bei p = 1/2 im Fach F_3 zu landen, berechnet sich zu $\binom{6}{3} \cdot 1/2^3 \cdot 1/2^3 = 20/64$. In der Tat führen genau 20 Pfade nach F_3. Sie lassen sich durch 0-1-Tupel $(a_1, \ldots, a_6)$ mit genau drei 1-Einträgen beschreiben, wobei „0" für „links" und „1" für „rechts" steht.

Normierung von Reihen und geometrische Verteilung

Aus jeder nichttrivialen konvergenten Summe nichtnegativer Zahlen können wir eine Verteilung der Eins durch Normierung erhalten: Ist $\Sigma_{n \in \mathbb{N}} a_n = b > 0$ für reelle Zahlen $a_n \geq 0$, so definiert $\nu(n) = a_n/b$ für alle $n \in \mathbb{N}$ eine Verteilung der Eins auf $\mathbb{N}$. Ein Beispiel liefert die geometrische Reihe $\Sigma_{n \in \mathbb{N}} q^n = (1 - q)^{-1}$ für ein q mit $0 \leq q < 1$. Wiederholen wir nämlich ein Zufallsexperiment mit einer Erfolgswahrscheinlichkeit p = 1 – q, so ist

$$\nu(n) = q^n p$$

für alle $n \in \mathbb{N}$ die Wahrscheinlichkeit, in den ersten n Versuchen einen Mißerfolg zu erzielen und danach einen Erfolg. Die Funktion ν heißt die *geometrische Verteilung* zum Parameter q. In der Tat ist ν die Normierung der geometrischen Reihe.

Gleichverteilungen und Dirac-Maße

Wir definieren nun allgemein:

Definition *(Gleichverteilung auf einer endlichen Menge)*
Sei A eine nichtleere Menge, und sei $\nu(a) = 1/|A|$ für alle $a \in A$.
Dann heißt das durch ν induzierte Maß die *Gleichverteilung* auf A.

Jedes Element der Grundmenge A erhält hier das gleiche Gewicht. Für alle $B \subseteq A$ ist $\mu(B)$ die Anzahl der Elemente von B geteilt durch die Anzahl der Elemente von A. Andererseits können wir auch einem einzigen Punkt die gesamte Masse zuweisen:

Definition *(Dirac-Maß)*
Sei A eine beliebige nichtleere Menge, und sei $a \in A$. Weiter sei

$\nu(b) = 1$, falls $b = a$, und $\nu(b) = 0$, sonst.

Dann heißt das durch ν induzierte Maß das *Dirac-Maß* auf A im Punkt a und wird mit $\delta_{a,A}$ bezeichnet.

Für das Dirac-Maß $\delta_{a,A}$ gilt also $\delta_{a,A}(B) = 1$, falls $a \in B$, und $\delta_{a,A}(B) = 0$, sonst, für alle $B \subseteq A$.

Gewichtete Summen und Produkte

Sind $\nu_1, \nu_2 : A \to [0, 1]$ Verteilungen der Eins, so können wir die Verteilungen mitteln, indem wir $\nu(a) = \nu_1(a)/2 + \nu_2(a)/2$ für alle $a \in A$ setzen. Statt der Mittelung mit Faktor 1/2 ist eine Mittelung mit Faktoren 1/3 und 2/3 usw. möglich.

Allgemein sei A eine abzählbare Menge und seien $\nu_n : A \to [0, 1]$ Verteilungen der Eins für alle $n \in \mathbb{N}$. Weiter sei auch $\nu : \mathbb{N} \to [0, 1]$ eine Verteilung der Eins. Dann definieren wir

$$\nu^*(a) = \Sigma_{n \in \mathbb{N}} \nu(n) \cdot \nu_n(a) \quad \text{für alle } a \in A.$$

Daß ν^* wieder eine Verteilung der Eins auf A ist, folgt aus dem folgenden auch andernorts oft nützlichen Summationssatz, der das Kommutativ- und Assoziativgesetz für nichtnegative reelle Zahlen ins Unendliche ausdehnt:

Satz *(Summationssatz)*
Seien $x_{n,m}$ reelle Zahlen mit $x_{n,m} \geq 0$ für alle $n, m \in \mathbb{N}$. Weiter sei $\pi : \mathbb{N} \to \mathbb{N}^2$ bijektiv. Es existiere

$$s^* = \sup(\{ \Sigma_{(n,m) \in E} x_{n,m} \mid E \subseteq \mathbb{N}^2, E \text{ endlich} \}).$$

Dann gilt:

$$(+) \quad \Sigma_{n \in \mathbb{N}} \Sigma_{m \in \mathbb{N}} x_{n,m} = \Sigma_{m \in \mathbb{N}} \Sigma_{n \in \mathbb{N}} x_{n,m} = \Sigma_{k \in \mathbb{N}} x_{\pi(k)} = s^*.$$

Beweis
Für alle $n_0, m_0, k_0 \in \mathbb{N}$ seien

$$S_{n_0,m_0} = \Sigma_{n \leq n_0} \Sigma_{m \leq m_0} x_{n,m}, \quad S'_{n_0,m_0} = \Sigma_{m \leq m_0} \Sigma_{n \leq n_0} x_{n,m},$$

$$T_{k_0} = \Sigma_{k \leq k_0} x_{\pi(k)}.$$

Dann gilt $S_{n_0,m_0} = S'_{n_0,m_0}$ und $S_{n_0,m_0}, S'_{n_0,m_0}, T_{k_0} \leq s^*$ für alle $n_0, m_0, k_0 \in \mathbb{N}$. Hieraus folgt $\Sigma_{n \leq n_0} \Sigma_{m \in \mathbb{N}} x_{n,m} \leq s^*$ für alle n_0 und $\Sigma_{m \leq m_0} \Sigma_{n \in \mathbb{N}} x_{n,m} \leq s^*$ für alle m_0, und damit dann weiter, daß alle in (+) betrachteten Reihen konvergent und kleinergleich s^* sind.
Zum Beweis der anderen Ungleichungen sei $E \subseteq \mathbb{N}^2$ endlich. Wir wählen dann n^* so groß, daß $E \subseteq \{ (n, m) \in \mathbb{N}^2 \mid n, m \leq n^* \} \cap \{ \pi(k) \mid k \leq n^* \}$. Dann gilt $\Sigma_{(n,m) \in E} x_{n,m} \leq S_{n^*,n^*}, S'_{n^*,n^*}, T_{n^*}$. Folglich sind alle in (+) betrachteten unendlichen Reihen größergleich s^*.

Damit können wir definieren:

Definition *(gewichtete Summe von Wahrscheinlichkeitsmaßen)*
Sind in obiger Situation μ_n die von den Verteilungen ν_n induzierten Maße, so nennen wir das von ν^* induzierte Maß μ^* die durch ν *gewichtete Summe* der μ_n, in Zeichen $\mu^* = \Sigma_{n \in \mathbb{N}} \nu(n) \mu_n$.

Wir modellieren zur Illustration folgende Situation: Ein Spieler wirft eine Münze, entscheidet dann mit Wahrscheinlichkeit 1/2, ob er aufhört oder noch einmal eine Münze wirft, usw. Als Grundmenge A können wir hier die Menge der nichtleeren endlichen 0-1-Folgen wählen, und für das modellierende Wahrscheinlichkeitsmaß gilt

$$\mu(\{ 0 \}) = \mu(\{ 1 \}) = 1/4,\ \mu(\{ 01 \}) = \mu(\{ 00 \}) = \mu(\{ 10 \}) = \mu(\{ 11 \}) = 1/16, \text{ usw.},$$

denn das Ereignis 10 entsteht zum Beispiel durch folgenden Ablauf: Der Spieler wirft eine 1, entscheidet sich weiterzumachen, wirft eine 0, und entscheidet sich aufzuhören.

Das Maß μ können wir einfach als gewichtete Summe notieren. Ist τ_n die Gleichverteilung auf den endlichen 0-1-Folgen der Länge n, und $\tau_n(s) = 0$ für alle anderen Folgen s, so gilt $\mu = \Sigma_{n \geq 1} 1/2^n \tau_n$.

Als nächstes betrachten wir Produkte von Wahrscheinlichkeitsräumen. Seien hierzu $\nu_1 : A \to [0, 1]$ und $\nu_2 : B \to [0, 1]$ Verteilungen der Eins. Weiter sei $C = A \times B$. Wir setzen

$$\nu((a, b)) = \nu_1(a) \cdot \nu_2(b) \quad \text{für alle } a \in A \text{ und } b \in B.$$

Dann ist ν eine Verteilung der Eins auf C, und wir können definieren:

Definition *(Produkt von Wahrscheinlichkeitsmaßen)*
Sind in obiger Situation μ_1 und μ_2 die von ν_1 bzw. ν_2 induzierten Maße, so heißt das von ν induzierte Maß μ das *Produkt* von μ_1 und μ_2, in Zeichen

$$\mu = \mu_1 \times \mu_2.$$

Weiter nennen wir (C, μ) den *Produktraum* von (A, μ_1) und (B, μ_2).

Es gilt dann $\mu(C) = \Sigma_{(a, b) \in C} \mu_1(\{ a \}) \cdot \mu_2(\{ b \})$ für alle $C \subseteq A \times B$.

Rekursiv definieren wir $\mu_1 \times \ldots \times \mu_{n+1} = (\mu_1 \times \ldots \times \mu_n) \times \mu_{n+1}$ und haben damit beliebig lange endliche Produkte von Wahrscheinlichkeitsmaßen und -räumen zur Verfügung. Diese Produktbildung entspricht der unabhängigen Hintereinanderausführung von Zufallsexperimenten, die durch $\mu_1, \ldots, \mu_n$ modelliert werden. Speziell gilt der leicht zu zeigende Satz:

Satz *(Produkte von Gleichverteilungen)*
Seien $\mu_1, \ldots, \mu_n$ die Gleichverteilungen auf $A_1, \ldots, A_n$, und sei

$$\mu = \mu_1 \times \ldots \times \mu_n.$$

Dann ist μ die Gleichverteilung auf $A = A_1 \times \ldots \times A_n$.

Bildmaße

Wir können ein auf einer Menge A definiertes Wahrscheinlichkeitsmaß μ mit Hilfe von Funktionen auf andere Mengen übertragen:

Definition *(Bildmaß)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $T : A \to B$ eine Funktion. Dann setzen wir:

$$\mu_T(C) = \mu(\{ a \in A \mid T(a) \in C \}) \quad \text{für alle } C \subseteq B.$$

Die Funktion $\mu_T : \mathcal{P}(B) \to [0, 1]$ heißt das *Bildmaß* von μ bzgl. T und wird auch mit $T(\mu)$ oder $\mu \circ T^{-1}$ bezeichnet.

Es ist leicht zu sehen, daß (B, μ_T) wieder ein Wahrscheinlichkeitsraum ist. Das Maß μ_T wird zudem induziert von der Verteilung $\nu_T : B \to [0, 1]$ der Eins auf B mit $\nu_T(b) = \mu(T^{-1}(\{ b \}))$ für alle $b \in B$.

Ist (A, μ) ein Wahrscheinlichkeitsraum und sind $T : A \to B$ sowie $S : B \to C$ Funktionen, so gilt $\mu_{S \circ T} = (\mu_T)_S$, oder, in den beiden alternativen Notationen, $(S \circ T)(\mu) = S(T(\mu))$ bzw. $\mu \circ (S \circ T)^{-1} = (\mu \circ T^{-1}) \circ S^{-1}$.

Ist $(A \times B, \mu)$ der Produktraum von (A, μ_1) und (B, μ_2), so ist μ_1 das Bildmaß der Projektion $pr_1 : A \times B \to A$ mit $pr_1((a, b)) = a$ für alle $(a, b) \in A \times B$. Denn für alle $C \subseteq A \times B$ ist

$$\mu_{pr_1}(C) = \mu(\{ (a, b) \in A \times B \mid pr_1((a, b)) \in C \}) = \mu(\{ (a, b) \in A \times B \mid a \in C \}) =$$

$$\mu(C \times B) = \Sigma_{(a, b) \in C \times B}\, \mu_1(\{ a \})\, \mu_2(\{ b \}) = \mu_1(C) \cdot \mu_2(B) = \mu_1(C) \cdot 1 = \mu_1(C).$$

Analog ist μ_2 das Bildmaß der Projektion $pr_2 : A \times B \to B$ mit $pr_2((a, b)) = b$ für alle $(a, b) \in A \times B$.

Additivität und Stetigkeit von Wahrscheinlichkeitsmaßen

Die grundlegenden Eigenschaften eines Wahrscheinlichkeitsmaßes sind die folgenden:

Satz *(Elementare Eigenschaften eines Wahrscheinlichkeitsmaßes)*
Sei (A, μ) ein Wahrscheinlichkeitsraum. Dann gilt:

(W1) $\mu(\varnothing) = 0$, $\mu(A) = 1$,

(W2) $\mu(B \cup C) = \mu(B) + \mu(C)$ für alle disjunkten $B, C \subseteq A$, *(Additivität)*

(W3) $\mu(\bigcup_{n \in \mathbb{N}} B_n) = \sup_{n \in \mathbb{N}} \mu(B_n)$ für alle $B_0 \subseteq \ldots \subseteq B_n \subseteq \ldots \subseteq A$. *(Aufwärts-Stetigkeit)*

Der Beweis dieses Satzes sei dem Leser zur Übung empfohlen.

Ist umgekehrt A abzählbar und $\mu : \mathcal{P}(A) \to [0, 1]$ eine Funktion, die die Aussagen (W1) – (W3) erfüllt, so ist $\nu(a) = \mu(\{ a \})$ für alle $a \in A$ eine Verteilung der Eins, die μ induziert. Damit sind unsere Wahrscheinlichkeitsmaße genau die Funktionen μ mit den Eigenschaften (W1) – (W3).

Die Eigenschaften (W1) – (W3) werden in der allgemeinen Wahrscheinlichkeitstheorie zur Definition eines Wahrscheinlichkeitsraumes verwendet. Hier betrachtet man auch überabzählbare Grundmengen A. Zur Modellierung eines Bogenschusses auf eine Scheibe ist z. B. ein Kreis $K \subseteq \mathbb{R}^2$ mit normierter Fläche 1 eine natürliche Grundmenge, und K ist überabzählbar. Ebenso liefert der unendliche Münzwurf die überabzählbare Grundmenge A aller Folgen $\langle b_n \mid n \in \mathbb{N} \rangle$ mit $b_n \in \{ 0, 1 \}$ für alle $n \in \mathbb{N}$. Für überabzählbare Grundmengen ist obiger Ansatz über Verteilungen der Eins nun nicht mehr allgemein geeignet. Dieses Phänomen ist für den intuitiven Flächenbegriff auf K, der wahrscheinlichkeitstheoretisch der Gleichverteilung auf K und damit dem zufälligen Bogenschuß entspricht, klar: Jede einpunktige Menge $\{ a \}$ hat Fläche 0, und damit ist die Fläche von K also nicht die Summe der Flächen seiner Punkte. Ebenso hat jede unendliche Folge in A beim unendlichen Münzwurf die Wahrscheinlichkeit 0, während alle Folgen zusammengenommen die Wahrscheinlichkeit 1 haben. Das ist irritierend genug, aber man wird sich damit anfreunden können, ohne diskrete Verteilungen der Eins zu leben und statt dessen eine Funktion $\mu : \mathcal{P}(A) \to [0, 1]$ ein Wahrscheinlichkeitsmaß zu nennen, falls die Aussagen (W1), (W2) und (W3) gelten. Hier taucht dann aber ein neues Problem auf: Es ist ungemein schwierig bis unmöglich, Wahrscheinlichkeitsmaße zu konstruieren, die auf allen Teilmengen eines überabzählbaren Grundraumes A definiert sind und für die $\mu(\{ a \}) = 0$ für alle $a \in A$ gilt (vgl. die Übungen zu diesem Zwischenabschnitt). Man muß deswegen noch eine weitere und diesmal wirklich bittere Pille schlucken und den Definitionsbereich von Wahrscheinlichkeitsmaßen reduzieren. Hierzu eignen sich Mengensysteme, die unter abzählbaren Operationen abgeschlossen sind:

Definition *(σ-Algebra)*
Eine Mengenalgebra $\mathcal{A}$ auf einer Menge A heißt eine *σ-Algebra*, falls für alle Folgen $\langle A_n \mid n \in \mathbb{N} \rangle$ in $\mathcal{A}$ gilt, daß $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Damit wird nun definiert:

Definition *(allgemeiner Wahrscheinlichkeitsraum)*
Sei $\mathcal{A}$ eine σ-Algebra auf A, und sei $\mu : \mathcal{A} \to [0, 1] \subseteq \mathbb{R}$. Dann heißt μ ein *Wahrscheinlichkeitsmaß* auf $\mathcal{A}$, falls gilt:

(W1) $\mu(\emptyset) = 0, \quad \mu(A) = 1,$

(W2) $\mu(B \cup C) = \mu(B) + \mu(C)$ für alle disjunkten $B, C \in \mathcal{A}$, *(Additivität)*

(W3) $\mu(\bigcup_{n \in \mathbb{N}} B_n) = \sup_{n \in \mathbb{N}} \mu(B_n)$ für alle $B_n \in \mathcal{A}$ mit $B_0 \subseteq \ldots \subseteq B_n \subseteq \ldots$ *(Aufwärts-Stetigkeit)*

In diesem Fall nennen wir $(A, \mathcal{A}, \mu)$ einen *Wahrscheinlichkeitsraum* mit *Grundmenge* A und *Ereignisraum* $\mathcal{A}$.

Wir wollten dem Leser diese Definition nicht vorenthalten, aber wir wollen nun auch gleich wieder zu den abzählbaren Wahrscheinlichkeitsräumen $(A, \mu) = (A, \mathcal{P}(A), \mu)$ zurückkehren. Die allgemeine Theorie verdient ein umfangreiches Vorspiel über Mengensysteme und die Konstruktion von Maßen auf σ-Algebren, für das hier nicht der Ort ist. Jedoch wollen wir der allgemeinen Theorie Tribut zollen, indem wir einige ihrer geistreichen Notationen importieren:

Summen als Integrale

Wir führen eine Integralschreibweise für Summen ein, die für sich genommen nützlich ist und zudem auf die Notationen und Sätze der allgemeinen Wahrscheinlichkeitstheorie vorbereitet. Da wir im folgenden die Konvergenz von Reihen oft stillschweigend behaupten, halten wir vorab fest:

Satz *(Majoranten- und Umordnungskriterium für Reihen)*
Sei $\Sigma_{n\in\mathbb{N}}\, a_n$ eine konvergente Reihe mit $a_n \geq 0$ für alle $n \in \mathbb{N}$.
Weiter seien b_n, $n \in \mathbb{N}$, reelle Zahlen mit $|b_n| \leq a_n$ für alle $n \in \mathbb{N}$.
Dann konvergiert die Reihe $\Sigma_{n\in\mathbb{N}}\, b_n$ gegen ein b mit $|b| \leq \Sigma_{n\in\mathbb{N}}\, a_n$.
Weiter ist die Reihe $\Sigma_{n\in\mathbb{N}}\, b_n$ invariant unter Umordnungen, d. h. für alle Bijektionen $\pi : \mathbb{N} \to \mathbb{N}$ gilt $\Sigma_{n\in\mathbb{N}}\, b_{\pi(n)} = \Sigma_{n\in\mathbb{N}}\, b_n$.

Beweis
Für alle $n < m$ ist $|\Sigma_{n<i\leq m}\, b_i| \leq \Sigma_{n<i\leq m}\, a_i$. Also ist die Folge der Partialsummen von $\Sigma_{n\in\mathbb{N}}\, b_n$ eine Cauchyfolge und daher konvergiert $\Sigma_{n\in\mathbb{N}}\, b_n$ gegen ein b. Aus $|\Sigma_{i<n}\, b_i| \leq \Sigma_{i<n}\, a_i$ für alle n folgt zudem $|b| \leq \Sigma_{n\in\mathbb{N}}\, a_n$.
Ist $\pi : \mathbb{N} \to \mathbb{N}$ bijektiv, so konvergiert $\Sigma_{n\in\mathbb{N}}\, b_{\pi(n)}$ gegen ein c, denn es gilt $|b_{\pi(n)}| \leq a_{\pi(n)}$ für alle n und die Reihe $\Sigma_{n\in\mathbb{N}}\, a_{\pi(n)}$ konvergiert.
Es bleibt zu zeigen, daß $c = b$ ist. Sei hierzu $\varepsilon > 0$ beliebig, und sei $n_0 \in \mathbb{N}$ derart, daß $\Sigma_{k\geq n_0}\, a_k < \varepsilon/2$. Dann gilt auch $|b - \Sigma_{i<n_0}\, b_i| \leq \Sigma_{k\geq n_0}\, a_k < \varepsilon/2$.
Sei nun n_1 derart, daß jedes $i < n_0$ ein Element von $\{\pi(0), \ldots, \pi(n_1)\}$ ist.
Dann gilt für alle $n \geq n_1$:

$$|b - \Sigma_{i<n}\, b_{\pi(i)}| \leq |b - \Sigma_{i<n_0}\, b_i| + |\Sigma_{i<n_0}\, b_i - \Sigma_{i<n}\, b_{\pi(i)}| \leq$$

$$\varepsilon/2 + |\Sigma_{i<n,\,\pi(i)\geq n_0}\, b_{\pi(i)}| \leq \varepsilon/2 + \Sigma_{k\geq n_0}\, a_k \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

– Da $\varepsilon > 0$ beliebig ist, folgt $b = c$.

Man kann zeigen, daß die alternierende harmonische Reihe $\Sigma_{n\in\mathbb{N}}\, b_n$, $b_n = (-1)^n/(n+1)$ für alle n, konvergiert (gegen $\ln(2)$), während es für jedes $x \in \mathbb{R}$ eine Bijektion $\pi : \mathbb{N} \to \mathbb{N}$ gibt mit $\Sigma_{n\in\mathbb{N}}\, b_{\pi(n)} = x$. Für diese konvergente Reihe gilt also kein Umordnungssatz. Das Majorantenkriterium ist verletzt, da die harmonische Reihe $\Sigma_{n\in\mathbb{N}}\, |b_n|$ nicht konvergiert.

Nach diesen Vorbereitungen können wir nun definieren:

Definition *(Integralschreibweise für abzählbare Summen)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $f : A \to \mathbb{R}$ eine Funktion. Die Funktion f heißt μ-*integrierbar*, falls $\Sigma_{a \in A} |f(a)| \, \mu(\{a\})$ konvergiert. Wir setzen dann:

$$\int f \, d\mu = \int f(x) \, \mu(dx) = \Sigma_{a \in A} f(a) \, \mu(\{a\}).$$

Wir nennen die reelle Zahl $\int f \, d\mu$ das μ-*Integral* von f.
Weiter setzen wir für alle $B \subseteq A$:

$$\int_B f \, d\mu = \int f(x) \cdot \mathrm{ind}_B(x) \, \mu(dx),$$

wobei $\mathrm{ind}_B : A \to \{0, 1\}$ die Indikatorfunktion auf B ist, d. h. es gilt $\mathrm{ind}_B(b) = 1$, falls $b \in B$, und $\mathrm{ind}_B(b) = 0$, sonst.

In der Tat führt die Konvergenz der Reihe $\Sigma_{a \in A} |f(a)| \, \mu(\{a\})$ nach dem Majorantenkriterium dazu, daß auch die Reihe $\Sigma_{a \in A} f(a) \, \mu(\{a\})$ konvergiert, und zwar gegen einen von einer Aufzählung von A unabhängigen Grenzwert.

Ist $A = \{a_n \mid n \in \mathbb{N}\}$ und bilden wir für alle $n \in \mathbb{N}$ im $\mathbb{R}^2$ das Rechteck mit den gegenüberliegenden Ecken $(n, 0)$ und $(n + \mu(\{a_n\}), f(a_n))$, so ist das μ-Integral von f die Fläche aller Rechtecke, wobei negative Funktionswerte zu negativen Flächen führen.

Die folgenden Eigenschaften sind leicht einzusehen:

Satz *(Eigenschaften des Integrals)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien f, g μ-integrierbar. Dann ist $\alpha f + \beta g$ μ-integrierbar für alle $\alpha, \beta \in \mathbb{R}$ und es gilt:

(i) $\int \alpha f + \beta g \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu$ für alle $\alpha, \beta \in \mathbb{R}$. *(Linearität)*

(ii) $\int f \, d\mu \leq \int g \, d\mu$, falls $f \leq g$. *(Monotonie)*

Hierbei lesen wir alle Operationen und Relationen punktweise, d. h. $\alpha f + \beta g$ ist die Funktion h mit $h(a) = \alpha f(a) + \beta g(a)$ für alle $a \in A$, und ebenso bedeutet $f \leq g$, daß $f(a) \leq g(a)$ für alle $a \in A$.

Für eine Funktion $f : A \to \mathbb{R}$ definieren wir den *Positivteil* $f^+ : A \to \mathbb{R}$ und den *Negativteil* $f^- : A \to \mathbb{R}$ von f durch

$f^+(a) = f(a)$, falls $f(a) \geq 0$, $f^+(a) = 0$, falls $f(a) < 0$,

$f^-(a) = 0$, falls $f(a) \geq 0$, $f^-(a) = -f(a)$, falls $f(a) < 0$.

Dann gilt $f = f^+ - f^-$ und $|f| = f^+ + f^-$. Zudem ist f genau dann μ-integrierbar, wenn sowohl f^+ als auch f^- μ-integrierbar sind, und in diesem Fall gilt:

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu. \qquad \textit{(Zerlegung in Positiv- und Negativteil)}$$

Ist $T : A \to B$, so ist eine Funktion $h : B \to \mathbb{R}$ genau dann T_μ-integrierbar, wenn die Funktion $h \circ T : A \to \mathbb{R}$ μ-integrierbar ist, und in diesem Fall gilt

$$\int h \, dT_\mu = \int h \circ T \, d\mu. \qquad \textit{(Transformationsformel für Bildmaße)}$$

Unabhängigkeit

Würfeln wir zweimal hintereinander, so ist das Eintreten einer 6 im zweiten Wurf unabhängig vom Ergebnis des ersten Wurfs, und speziell ist die Wahrscheinlichkeit, im zweiten Wurf eine 6 zu würfeln, immer noch gleich 1/6, auch wenn bereits im ersten Wurf eine 6 gewürfelt wurde. Allgemein ist die Wahrscheinlichkeit, im ersten Wurf a_1 und dann im zweiten Wurf a_2 zu würfeln, gleich dem Produkt der Einzelwahrscheinlichkeiten, also gleich $1/6 \cdot 1/6 = 1/36$. Wir definieren nun „Unabhängigkeit" über diese Produktregel:

Definition *(unabhängige Ereignisse)*
Sei (A, μ) ein Wahrscheinlichkeitsraum. Zwei Ereignisse $B, C \subseteq A$ heißen *unabhängig* (*unter* oder *bei* μ), falls gilt:

$$\mu(B \cap C) = \mu(B) \cdot \mu(C).$$

Sei z. B. $n \geq 2$ und μ die Gleichverteilung auf $A = \{1, \ldots, 6\}^n$, d. h. wir modellieren den n-fachen Wurf eines Würfels. Dann sind alle Ereignisse B und C, die nur mit Hilfe zweier verschiedener Indizes i und j aus $\{1, \ldots, n\}$ definiert sind, unabhängig voneinander. Denn seien

$$B = \{(a_1, \ldots, a_n) \in A \mid \mathscr{E}(a_i)\},$$

$$C = \{(a_1, \ldots, a_n) \in A \mid \mathscr{F}(a_j)\},$$

für gewisse Eigenschaften $\mathscr{E}$ und $\mathscr{F}$ und $1 \leq i, j \leq n$ mit $i \neq j$. Wir setzen:

k_0 = „die Anzahl aller $1 \leq w \leq 6$ mit $\mathscr{E}(w)$",

k_1 = „die Anzahl aller $1 \leq w \leq 6$ mit $\mathscr{F}(w)$".

Dann hat die Menge B genau $k_0 \cdot 6^{n-1}$ viele, die Menge C genau $k_1 \cdot 6^{n-1}$ und die Menge $B \cap C = \{(a_1, \ldots, a_n) \in A \mid \mathscr{E}(a_i)$ und $\mathscr{F}(a_j)\}$ wegen $i \neq j$ genau $k_0 \cdot k_1 \cdot 6^{n-2}$ viele Elemente. Da μ die Gleichverteilung auf $\{1, \ldots, 6\}^n$ ist, gilt also

$$\mu(B \cap C) = (k_0\, k_1)/6^2 = k_0/6 \cdot k_1/6 = \mu(B) \cdot \mu(C).$$

Dieses Ergebnis entspricht unserer Intuition, daß sich der i-te und der j-te Wurf des Würfels in keiner Weise gegenseitig beeinflussen.

Wir betrachten die Unabhängigkeit zweier Ereignisse noch unter einem etwas anderen Blickwinkel. Modellieren wir ein Zufallsexperiment durch einen Wahrscheinlichkeitsraum (A, μ) und erhalten wir die Information, daß ein Ereignis B eingetreten ist, so verändert diese Information unser Maß μ. Ist beim Wurf eines Würfels $B = \{2, 4, 6\}$, so arbeiten wir fortan mit der Verteilung $\nu(1) = \nu(3) = \nu(5) = 0$, $\nu(2) = \nu(4) = \nu(6) = 1/3$ der Eins, und nicht mehr mit der Gleichverteilung. (Bei dieser Betrachtung ist die Interpretation von Wahrscheinlichkeiten als Maß unserer Ignoranz besonders bestechend.) Allgemein definieren wir:

Definition *(bedingte Wahrscheinlichkeit)*
Sei (A, μ) ein Wahrscheinlichkeitsraum. Dann heißt für alle $B, C \subseteq A$ mit $\mu(B) \neq 0$ der Wert

$$\mu_B(C) = \mu(B \cap C)/\mu(B)$$

die *bedingte Wahrscheinlichkeit* von C *gegeben* B.

Zwei Ereignisse B und C mit $\mu(B) \neq 0$ sind damit genau dann unabhängig, wenn die Information, daß das Ereignis B eingetreten ist, an der Wahrscheinlichkeit von C nichts ändert, d. h. falls die bedingte Wahrscheinlichkeit von C gegeben B gleich $\mu(C)$ ist.

Allgemeiner definieren wir nun noch die Unabhängigkeit einer Folge von Ereignissen:

Definition *(unabhängige Folgen)*
Sei (A, μ) ein Wahrscheinlichkeitsraum. Eine Folge $\langle A_i \mid i \in I \rangle$ in $\mathcal{P}(A)$ heißt *unabhängig* (*unter* μ), falls für alle endlichen $J \subseteq I$ mit $J \neq \emptyset$ gilt:

$$\mu(\bigcap_{j \in J} A_j) = \prod_{j \in J} \mu(A_j).$$

Es genügt hier nicht zu fordern, daß $\mu(A_i \cap A_j) = \mu(A_i) \cdot \mu(A_j)$ für alle $i < j$ gilt. Sei hierzu μ die Gleichverteilung auf $A = \{0, 1\} \times \{0, 1\}$. Wir betrachten

$$B = \{(0, 0), (0, 1)\}, \quad C = \{(0, 0), (1, 0)\}, \quad D = \{(0, 1), (1, 0)\}.$$

Dann sind je zwei Elemente von $\mathcal{A} = \{B, C, D\}$ unabhängig, während das System $\mathcal{A}$ selbst nicht unabhängig ist:

$$\mu(B \cap C) = \mu(B \cap D) = \mu(C \cap D) = 1/4 = \mu(B)\,\mu(C) = \mu(C)\,\mu(D) = \mu(B)\,\mu(D),$$

$$\mu(B \cap C \cap D) = \mu(\emptyset) = 0 \neq 1/8 = \mu(B)\,\mu(C)\,\mu(D).$$

Zufallsvariable

In vielen Fällen interessiert uns nicht nur das Eintreten eines Elementarereignisses $a \in A$, sondern auch ein von a abhängiger reeller Wert $X = X(a)$, etwa die Summe der Augen a_i für $a = (a_1, \ldots, a_{10}) \in \{1, \ldots, 6\}^{10}$ beim 10-fachen Wurf eines Würfels. Obwohl dieser Wert einfach durch eine Funktion $X : A \to \mathbb{R}$ gegeben wird, erhält er eine Zufälligkeit durch die Zufälligkeit von $a \in A$. So können wir in unserem Beispiel nach der Wahrscheinlichkeit fragen, daß X zwischen 10 und 20 liegt. Diese Betrachtungen motivieren die folgende Sprechweise:

Definition *(Zufallsvariable, Verteilung einer Zufallsvariable)*
Sei (A, μ) ein Wahrscheinlichkeitsraum. Dann heißt jede Funktion $X : A \to \mathbb{R}$ eine *(reellwertige) Zufallsvariable* auf A.
Das Bildmaß μ_X auf $\mathrm{rng}(X)$ heißt auch die *Verteilung* der Zufallsvariable X.

Ist $X : A \to \mathbb{R}$ eine Zufallsvariable und $\mathscr{E}$ eine Eigenschaft reeller Zahlen, so schreiben wir suggestiv $\mu(\mathscr{E}(X))$ statt $\mu(\{ a \in A \mid \mathscr{E}(X(a)) \})$. Zum Beispiel ist

$$\mu(10 \le X \le 20) = \mu(\{ a \in A \mid 10 \le X(a) \le 20 \}).$$

Ähnlich ist $\{ X \in B \}$ eine Kurzform von $\{ a \in A \mid X(a) \in B \}$ $(= X^{-1}[B])$, usw.

Eine wichtige Rolle bei der wiederholten Auswertung eines Zufallsexperiments spielt der folgende Mittelwert einer Zufallsvariablen:

Definition *(Erwartungswert)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $X : A \to \mathbb{R}$ eine μ-integrierbare Zufallsvariable. Dann setzen wir

$$E(X) = \int X \, d\mu,$$

und nennen $E(X)$ den *Erwartungswert* von X.

Ist μ die Gleichverteilung auf einer Menge $A = \{ a_1, \ldots, a_n \}$ mit genau n Elementen, so gilt $E(X) = (X(a_1) + X(a_2) + \ldots + X(a_n))/n$. In diesem Fall ist also $E(X)$ das arithmetische Mittel der Werte $X(a_i)$.

Existiert $E(X^2)$ für eine Zufallsvariable X, so existiert auch $E(X)$, denn es gilt $|X| \le X^2 + 1$. Mit $E(X)$ und $E(X^2)$ existiert dann weiter auch $E((X - E(X))^2)$. Damit können wir definieren:

Definition *(Varianz)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $X : A \to \mathbb{R}$ eine Zufallsvariable auf A derart, daß $E(X^2)$ existiert. Dann definieren wir

$$V(X) = E((X - E(X))^2) = \int (X - E(X))^2 \, d\mu,$$

und nennen $V(X)$ die *Varianz* von X. Weiter heißt $\sigma(X) = \sqrt{V(X)}$ die *Standardabweichung* von X.

Sind $X, Y : A \to \mathbb{R}$ Zufallsvariablen, deren Varianzen existieren, so existiert auch $E(XY)$, denn es gilt $|XY| \le X^2 + Y^2$. Diese Eigenschaft benutzen wir in der folgenden Definition der Kovarianz zweier Zufallsvariablen:

Definition *(Kovarianz, Korrelationskoeffizient, unkorreliert)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $X, Y : A \to \mathbb{R}$ Zufallsvariablen, deren Varianzen existieren. Dann definieren wir

$$\mathrm{Cov}(X, Y) = E((X - E(X))(Y - E(Y))),$$

$$\rho(X, Y) = \mathrm{Cov}(X, Y)/(\sigma(X)\,\sigma(Y)), \quad \text{falls } \sigma(X), \sigma(Y) > 0$$

und nennen $\mathrm{Cov}(X, Y)$ die *Kovarianz* und $\rho(X, Y)$ den *Korrelationskoeffizienten* von X und Y. X und Y heißen *unkorreliert*, wenn $\mathrm{Cov}(X, Y) = 0$.

Grundlegende Eigenschaften des Erwartungswerts und der Varianz versammelt der folgende Satz, dessen Beweis dem Leser überlassen bleiben kann:

Satz *(elementare Eigenschaften von E(X), V(X), Cov(X, Y))*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $X, Y, X_1, \ldots, X_n : A \to \mathbb{R}$ Zufallsvariablen, deren Varianzen existieren. Dann gilt:

(a) $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$ für alle $\alpha, \beta \in \mathbb{R}$.

(b) $Cov(X, Y) = Cov(Y, X) = E(XY) - E(X)E(Y)$,
$V(X) = Cov(X, X) = E(X^2) - E(X)^2$.

(c) $Cov(\alpha X + \beta, \gamma Y + \delta) = \alpha\gamma\, Cov(X, Y)$ für alle $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

(d) $(\Sigma_{1 \le i \le n} X_i)^2$ ist μ-integrierbar und es gilt
$V(\Sigma_{1 \le i \le n} X_i) = \Sigma_{1 \le i \le n} V(X_i) + 2 \cdot \Sigma_{1 \le i < j \le n} Cov(X_i, X_j)$.

Nach (b) und (c) gilt also speziell

$Cov(X, Y) = Cov(X - E(X), Y - E(Y))$,

$V(X) = V(X - E(X))$.

Wir dürfen also beim Rechnen mit Varianzen und Kovarianzen annehmen, daß unsere Zufallsvariablen den Erwartungswert Null haben (denn $E(X - E(X)) = 0$).

Wir zeigen nun noch:

Satz *(Ungleichung von Cauchy-Schwarz)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $X, Y : A \to \mathbb{R}$ Zufallsvariablen, deren Varianzen existieren. Dann existiert der Erwartungswert von $X \cdot Y$ und es gilt

$E(XY)^2 \le E(X^2) \cdot E(Y^2)$. *(Cauchy-Schwarz-Ungleichung)*

Speziell gilt $|Cov(X, Y)| \le \sigma(X)\,\sigma(Y)$ und $|\rho(X, Y)| \le 1$, falls $\sigma(X), \sigma(Y) > 0$.

Beweis
Ist $E(X^2) = 0$, so gilt $\Sigma_{a \in A} X^2(a)\,\mu(\{a\}) = 0$, also gilt für alle $a \in A$, daß $X(a) = 0$ oder $\mu(\{a\}) = 0$. Also ist $0 = \Sigma_{a \in A} X(a)Y(a)\mu(\{a\}) = E(XY)$ und damit $E(XY)^2 = 0 \le E(X^2) \cdot E(Y^2)$. Ebenso gilt dies, falls $E(Y^2) = 0$.

Seien also $E(X^2), E(Y^2) > 0$. Wir setzen:

$X' = X/\sqrt{E(X^2)}$,

$Y' = Y/\sqrt{E(Y^2)}$.

Die Cauchy-Schwarz-Ungleichung folgt nun aus folgender Abschätzung:

$$\int X'Y'\,d\mu \le \int (X'^2 + Y'^2)/2\,d\mu = 1/2 + 1/2 = 1,$$

für die wir verwenden, daß $0 \le (\alpha - \beta)^2$ und also $\alpha\beta \le (\alpha^2 + \beta^2)/2$ für alle $\alpha, \beta \in \mathbb{R}$ gilt. Der Zusatz folgt aus der Cauchy-Schwarz-Ungleichung für die Zufallsvariablen $X' = X - E(X)$ und $Y' = Y - E(Y)$.

Schließlich dehnen wir den Begriff der Unabhängigkeit auf Zufallsvariable aus:

Definition *(Unabhängigkeit von Zufallsvariablen)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $\mathcal{X} = \langle X_i \mid i \in I \rangle$ eine Folge von Zufallsvariablen auf A. Dann heißt $\mathcal{X}$ *unabhängig* (*unter* oder *bei* μ), wenn für alle Folgen $\langle B_i \mid i \in I \rangle$ in $\mathcal{P}(\mathbb{R})$ gilt, daß $\langle \{ X_i \in B_i \} \mid i \in I \rangle$ unabhängig ist.

Speziell sind also zwei Zufallsvariablen X und Y auf A unabhängig, wenn für alle $B, C \subseteq \mathbb{R}$ gilt, daß $\mu(X \in B \text{ und } Y \in C) = \mu(X \in B) \cdot \mu(X \in C)$, d.h. falls für alle $B, C \subseteq \mathbb{R}$ die Ereignisse $X^{-1}[B]$ und $Y^{-1}[C]$ unabhängig sind.

Für alle Zufallsvariablen X, Y gilt:

(+) Sind X, Y unabhängig, so sind X, Y unkorreliert.

Die Umkehrung ist im allgemeinen nicht gültig. Der Beweis dieser Aussagen sei wieder dem Leser überlassen.

Nach Aussage (d) des obigen Satzes gilt also für paarweise unabhängige Zufallsvariablen $X_1, \ldots, X_n$ die folgende sympathische Summenregel für Varianzen:

$$V(X_1 + \ldots + X_n) = V(X_1) + \ldots + V(X_n).$$

Ist (A, μ), $A \subseteq \mathbb{R}$, ein Wahrscheinlichkeitsraum, so hätte man gerne einen Wahrscheinlichkeitsraum (B, ν) und unabhängige Zufallsvariablen $\langle X_n \mid n \in \mathbb{N} \rangle$ auf der Grundmenge B, deren Verteilung gleich μ ist, d.h. für alle $a \in A$ und alle $n \in \mathbb{N}$ gilt $\nu(X_n = a) = \mu(\{ a \})$. Dann modellieren die Zufallsvariablen X_n ein unendlich oft wiederholtes Zufallsexperiment mit Verteilung μ. Die Konstruktion einer derartigen Folge von unabhängigen Zufallsvariablen ist eine nichttriviale Angelegenheit und eine Aufgabe der allgemeinen Wahrscheinlichkeitstheorie, da die Menge B überabzählbar ist, wenn A mehr als ein Element besitzt. Man verallgemeinert hierzu die Produktbildung $(A, \mu) \times \ldots \times (A, \mu)$ ins Unendliche und erhält einen Wahrscheinlichkeitsraum $(B, \mathcal{B}, \nu)$ mit einer σ-Algebra $\mathcal{B}$ auf der Menge B aller Funktionen $f : \mathbb{N} \to A$. Dann sind die Projektionen $X_n : B \to A$ mit $X_n(f) = f(n)$ für alle $f \in B$ und alle $n \in \mathbb{N}$ unabhängige Zufallsvariable mit Verteilung μ wie gewünscht.

Das Gesetz der großen Zahl

Das Augenzählen beim Würfeln hat den Erwartungswert $(1 + \ldots + 6)/6 = 3{,}5$. Würfeln wir 100 mal, so erwarten wir ungefähr 350 als Summe der Augen. Allgemein können wir ein Zufallsexperiment wiederholt in unabhängiger Weise n-mal hintereinander durchführen und dabei die Ergebnisse $X(1), \ldots, X(n)$ für eine Zufallsvariable X von Interesse notieren. Dann sollte für große Zahlen n das arithmetische Mittel $(X(1) + \ldots + X(n))/n$ der Ergebnisse „in der Regel“ nahe beim Erwartungswert E(X) der ausgewerteten Zufallsvariablen X liegen. Diese Anschauung wollen wir nun präzisieren und beweisen.

Im folgenden sei (A, μ) ein Wahrscheinlichkeitsraum, und es sei $X : A \to \mathbb{R}$ eine Zufallsvariable, deren Varianz existiert. Für alle $n \geq 1$ sei

$(A^n, \mu_n) = (A, \mu) \times \ldots \times (A, \mu)$

das n-fache Produkt von (A, μ), und für alle $1 \leq i \leq n$ sei

$X_i^n(a_1, \ldots, a_n) = X(a_i)$ für alle $(a_1, \ldots, a_n) \in A^n$.

Schließlich sei für alle $n \geq 1$

$Y^n = (\Sigma_{1 \leq i \leq n} X_i^n)/n$

das arithmetische Mittel der Zufallsvariablen X_i^n, $1 \leq i \leq n$.

Wir wollen zeigen, daß Y^n für große n mit hoher Wahrscheinlichkeit kaum von $E(X)$ abweicht. Eine mögliche Präzisierung dieser Aussage ist:

Für alle $\varepsilon > 0$ gilt $\lim_{n \geq 1} \mu_n(|Y^n - E(X)| < \varepsilon) = 1$.

Hierzu beweisen wir vorab eine allgemeine Ungleichung:

Satz *(Ungleichung von Bienaymé-Chebyshev)*
Für alle reellen Zahlen $t > 0$ gilt $\mu(|X| \geq t) \leq E(X^2)/t^2$.

Beweis

$$E(X^2) = \int X^2 \, d\mu \geq \int_{\{|X| \geq t\}} X^2 \, d\mu \geq \int_{\{|X| \geq t\}} t^2 \, d\mu = t^2 \, \mu(|X| \geq t).$$

–

Hiermit erhalten wir nun schnell:

Satz *(schwaches Gesetz der großen Zahl)*
Für alle $n \geq 1$ und alle $\varepsilon > 0$ gilt

$\mu_n(|Y^n - E(X)| \geq \varepsilon) \leq V(X)/(n\varepsilon^2)$.

Insbesondere gilt also $\lim_{n \geq 1} \mu_n(|Y^n - E(X)| < \varepsilon) = 1$.

Beweis
Sei $n \geq 1$ und $\varepsilon > 0$. Die Zufallsvariablen X_i^n, $1 \leq i \leq n$, sind unabhängig und es gilt $E(X_i^n) = E(X)$ und $V(X_i^n) = V(X)$ für alle $1 \leq i \leq n$. Wir rechnen:

$\mu_n(|Y^n - E(X)| \geq \varepsilon) = \mu_n(|(\Sigma_{1 \leq i \leq n} X_i^n)/n - E(X)| \geq \varepsilon) =$

$\mu_n(|(\Sigma_{1 \leq i \leq n} X_i^n) - nE(X)| \geq n\varepsilon) =$

$\mu_n(|\Sigma_{1 \leq i \leq n} (X_i^n - E(X_i^n))| \geq n\varepsilon) \quad \leq_{\text{Bienaymé-Chebychev}}$

$E((\Sigma_{1 \leq i \leq n} (X_i^n - E(X_i^n)))^2)/(n^2\varepsilon^2) \quad =_{X_i^n,\, 1 \leq i \leq n,\, \text{paarweise unkorreliert}}$

$E(\Sigma_{1 \leq i \leq n} (X_i^n - E(X_i^n))^2)/(n^2\varepsilon^2) =$

$(\Sigma_{1 \leq i \leq n} V(X_i^n))/(n^2\varepsilon^2) = n\,V(X)/(n^2\varepsilon^2) = V(X)/(n\varepsilon^2)$.

–

Wir wollen nun noch eine stärkere Version des Gesetzes der großen Zahl formulieren, die besagt, daß alle unendlichen Folgen in A, deren arithmetisches Mittel nicht gegen E(X) konvergiert, eine vernachlässigbare Menge „vom Maß Null" bilden. Diese Aussage läßt sich mit Hilfe der oben angesprochenen unendlichen Produktmaße präzisieren, die ein unendlich oft wiederholtes Zufallsexperiment modellieren. Wir können das Ergebnis aber auch elementar in einer äquivalenten Form formulieren. Hierzu sei

$\mathcal{F} = \{ \langle a_n \mid n \geq 1 \rangle \mid a_n \in A \text{ für alle } n \geq 1 \}$,

die Menge aller unendlichen Folgen (ab n = 1) in der Menge A. Für alle endlichen Folgen $s = \langle b_1, \ldots, b_m \rangle$, $m \geq 1$, in A sei weiter

$N_s = \{ \langle a_n \mid n \geq 1 \rangle \in \mathcal{F} \mid a_i = b_i \text{ für alle } 1 \leq i \leq m \}$

die Menge aller Folgen in $\mathcal{F}$ mit Anfangsstück s. Dann gilt:

Satz *(starkes Gesetz der großen Zahl)*
Sei N die Menge aller Folgen $\langle a_n \mid n \geq 1 \rangle \in \mathcal{F}$ mit:

$(\Sigma_{1 \leq i \leq n} a_i)/n$ konvergiert nicht gegen E(X).

Sei $\varepsilon > 0$. Dann existieren endliche Folgen $s_n = \langle a_i^n \mid 1 \leq i \leq k_n \rangle$ in A, $n \in \mathbb{N}$, mit den Eigenschaften:

(i) $N \subseteq \bigcup_{n \in \mathbb{N}} N_{s_n}$,

(ii) $\mu(\{ a_1^n \}) \cdot \ldots \cdot \mu(\{ a_{k_n}^n \}) < \varepsilon/2^n$ für alle $n \in \mathbb{N}$.

Die Eigenschaft (i) besagt, daß alle „schlechten" unendlichen Folgen von denjenigen Folgen überdeckt werden, die mit einem s_n beginnen. Weiter besagt dann (ii), daß $\mu_{k_n}(\{ s_n \}) < \varepsilon/2^n$ im k_n-fachen Produkt von μ gilt, wobei k_n die Länge von s_n ist. Damit hat N in einem unendlichen Produktraum, der alle endlichen Produkträume fortsetzt, eine Wahrscheinlichkeit, die kleiner als $\Sigma_{n \geq 1} \varepsilon/2^n = \varepsilon$ ist. Da dies für alle $\varepsilon > 0$ gilt, ist N eine Menge vom Maß Null in diesem Produktraum. Obige Formulierung kommt aber ganz ohne unendliche (und sogar ohne endliche) Produkträume aus.

Übungen

Übung 1 *(Abzählbare Wahrscheinlichkeitsräume, I)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $A_1, \ldots, A_n \subseteq A$.
Zeigen Sie:

(i) $\mu(\bigcup_{1 \le i \le n} A_i) \le \Sigma_{1 \le i \le n} \mu(A_i)$,

(ii) $\mu(\bigcup_{1 \le i \le n} A_i) \ge \Sigma_{1 \le i \le n} \mu(A_i) - \Sigma_{1 \le i < j \le n} \mu(A_i \cap A_j)$.

(iii) $\mu(\bigcup_{1 \le i \le n} A_i) \le \Sigma_{1 \le i \le n} \mu(A_i) - \Sigma_{1 \le i < j \le n} \mu(A_i \cap A_j) + \Sigma_{1 \le i < j < k \le n} \mu(A_i \cap A_j \cap A_k)$.

Übung 2 *(Abzählbare Wahrscheinlichkeitsräume, II)*
Wir betrachten das blinde Ziehen von k Kugeln aus einer Urne mit n numerierten Kugeln $1, \ldots, n$ in den folgenden vier Varianten:

(1) mit Reihenfolge, mit Zurücklegen,

(2) mit Reihenfolge, ohne Zurücklegen,

(3) ohne Reihenfolge, ohne Zurücklegen,

(4) ohne Reihenfolge, mit Zurücklegen.

Wir setzen hierzu:

$A_1 = \{1, \ldots, n\}^k = \{(a_1, \ldots, a_k) \mid 1 \le a_i \le n \text{ für alle } i\}$,

$A_2 = \{(a_1, \ldots, a_k) \in A_1 \mid a_i \ne a_j \text{ für alle } i \ne j\}$,

$A_3 = \{(a_1, \ldots, a_k) \in A_1 \mid a_1 < \ldots < a_k\}$,

$A_4 = \{(a_1, \ldots, a_k) \in A_1 \mid a_1 \le \ldots \le a_k\}$.

Begründen Sie, daß die folgenden Anzahlen $|A_i|$ gelten und daß die Wahrscheinlichkeitsmaße $\mu_1, \ldots, \mu_4$ die vier Urnenziehungen modellieren:

$|A_1| = n^k$, $\quad |A_2| = n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$,

$|A_3| = |A_2|/k! = \binom{n}{k}$, $\quad |A_4| = \binom{n+k-1}{k}$,

$\mu_i(\{(a_1, \ldots, a_k)\}) = 1/|A_i|$ für alle $i = 1, 2, 3$ und alle $(a_1, \ldots, a_k) \in A_i$,

$\mu_4(\text{„genau } k_1 \text{ mal } 1, \ldots, \text{genau } k_n \text{ mal } n\text{“}) = \binom{k}{k_1, \ldots, k_n}/n^k$,
für alle $0 \le k_1, \ldots, k_n \le k$ mit $\Sigma_{1 \le i \le n} k_i = k$,

wobei die Binomialkoeffizienten $\binom{n}{k}$ und die Multinomialkoeffizienten $\binom{n}{k_1, \ldots, k_r}$ definiert sind durch

$\binom{n}{k} = n!/(k! \cdot (n-k)!)$, $\quad \binom{n}{k_1, \ldots, k_r} = n!/(k_1! \cdot \ldots \cdot k_r!)$.

[Zur Berechnung von $|A_4|$ betrachten wir die Abbildung g auf A_4 mit $g(a_1, \ldots, a_k) = (b_1, \ldots, b_k)$, $b_i = a_i + i - 1$ für alle i, wobei $a_1 \le \ldots \le a_k$.]

Übung 3 *(Abzählbare Wahrscheinlichkeitsräume, III)*
Sei $A = \{ 1, \ldots, k \}$ und sei (A^n, μ) die Gleichverteilung auf A^n für ein $n \geq 1$. Bestimmen Sie:

(i) $\mu(\{ (a_1, \ldots, a_n) \in A^n \mid \text{es gibt ein } i \text{ mit } a_i = 1 \})$,

(ii) $\mu(\{ (a_1, \ldots, a_n) \in A^n \mid \text{es gibt } i, j \text{ mit } a_i = 1 \text{ und } a_j = 2 \})$,

(iii) $\mu(\{ (a_1, \ldots, a_n) \in A^n \mid \text{es gibt ein } i < n \text{ mit } a_i = 1 \text{ und } a_{i+1} = 2 \})$.

Übung 4 *(Abzählbare Wahrscheinlichkeitsräume, IV)*
Zeigen Sie, daß sich jedes Wahrscheinlichkeitsmaß als gewichtete Summe von Dirac-Maßen schreiben läßt.

Übung 5 *(Abzählbare Wahrscheinlichkeitsräume, V)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei (A^n, ν) sein n-faches Produkt für ein $n \geq 1$. Weiter seien $B, C \subseteq A$ und $i, j \in \{ 1, \ldots, n \}$. Bestimmen Sie die μ-Wahrscheinlichkeiten von

(i) $\nu(\{ (a_1, \ldots, a_n) \in A^n \mid a_i \in B \text{ und } a_j \in C \})$ und

(ii) $\nu(\{ (a_1, \ldots, a_n) \in A^n \mid a_i \in B \text{ oder } a_j \in C \})$

mit Hilfe von $\mu(B)$ und $\mu(C)$.

Übung 6 *(Additivität und Stetigkeit von Wahrscheinlichkeitsmaßen, I)*
Sei A eine abzählbare Menge, und sei $\mu : \mathcal{P}(A) \to [0, 1]$. Zeigen Sie, daß (A, μ) genau dann ein Wahrscheinlichkeitsraum ist, wenn die beiden folgenden Bedingungen gelten:

(W*1) $\mu(A) = 1$,

(W*2) $\mu(\bigcup_{n \in \mathbb{N}} A_n) = \Sigma_{n \in \mathbb{N}} \mu(A_n)$ für alle paarweise disjunkten $A_n \subseteq A$. *(Sigma-Additivität)*

Formulieren und beweisen Sie zudem eine zur Aufwärts-Stetigkeit (W3) äquivalente Abwärts-Stetigkeit.

Übung 7 *(Additivität und Stetigkeit von Wahrscheinlichkeitsmaßen, II)*
Sei (A, μ) ein Wahrscheinlichkeitsraum. Sei $\langle A_n \mid n \in \mathbb{N} \rangle$ eine Folge von paarweise disjunkten Teilmengen von A. Zeigen Sie:

$\mu(\bigcup_{n \in \mathbb{N}} A_n) = \Sigma_{n \in \mathbb{N}} \mu(A_n)$.

Übung 8 *(Additivität und Stetigkeit von Wahrscheinlichkeitsmaßen, III)*
Sei A eine Menge, und sei $\mathcal{S}$ eine Menge von σ-Algebren auf A. Zeigen Sie, daß $\bigcap \mathcal{S}$ eine σ-Algebra auf A ist, und folgern Sie, daß es für alle Mengensysteme $\mathcal{B}$ auf A eine kleinste σ-Algebra $\sigma(\mathcal{B})$ gibt mit $\mathcal{B} \subseteq \sigma(\mathcal{B})$.

Übung 9 *(Additivität und Stetigkeit von Wahrscheinlichkeitsmaßen, IV)*
Sei $K = \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = (2\pi)^{-2} \}$ die Kreislinie mit Umfang 1. Für alle $v, w \in K$ sei $\alpha(v, w) \in [0, 2\pi[$ der Winkel zwischen v und w im Bogenmaß, gemessen von v nach w gegen den Uhrzeigersinn. Für alle $v \in K$ und $\alpha \in \mathbb{R}$ sei $\mathrm{rot}_\alpha(v) \in K$ die Drehung von v um den Winkel α gegen den Uhrzeigersinn, und für $A \subseteq K$ sei $\mathrm{rot}_\alpha(A) = \{ \mathrm{rot}_\alpha(v) \mid v \in A \}$ die Drehung der Menge A um den Winkel α. Wir definieren nun die *Vitali-Äquivalenzrelation* auf K durch:

$v \sim w$, *falls* „es gibt ein $q \in \mathbb{Q} \cap [0, 1[$ mit $\alpha(v, w) = q\,2\pi$“ für alle $v, w \in K$.

Sei V ein vollständiges Repräsentantensystem für $\sim$. Benutzen Sie die Menge V um zu zeigen, daß es keine Funktion $\mu : \mathcal{P}(K) \to [0, 1]$ gibt mit den Eigenschaften (W1) – (W3) sowie:

(i) $\mu(\{ v \}) = 0$ für alle $v \in K$.

(ii) $\mu(\mathrm{rot}_\alpha(A)) = \mu(A)$ für alle $A \subseteq K$ und alle $\alpha \in [0, 2\pi[$.

[Argumentieren Sie, daß sowohl $\mu(V) = 0$ als auch $\mu(V) > 1/n$ für ein $n \geq 1$ den geforderten Eigenschaften von μ widerspricht.]

Jede Modellierung des „Glücksrads“ muß sich also auf eine σ-Algebra $\mathcal{A}$ auf K beschränken, die nicht die volle Potenzmenge von A ist. Erweitern Sie dieses Ergebnis auch auf die Modellierung eines zufälligen Pfeilwurfs auf eine Kreisscheibe und auf die zufällige Wahl eines Punktes in $[0, 1[$.

Übung 10 *(Additivität und Stetigkeit von Wahrscheinlichkeitsmaßen, V)*
Zeigen Sie mit Hilfe einer zur vorherigen Übung verwandten Äquivalenzrelation $\sim$ auf der Menge $\mathcal{F}$ aller 0-1-Folgen $f : \mathbb{N} \to \{ 0, 1 \}$, daß jede Modellierung eines unendlichen Münzwurfs lediglich auf einer σ-Algebra $\mathcal{A}$ auf $\mathcal{F}$ definiert werden kann, die eine echte Teilmenge von $\mathcal{P}(\mathcal{F})$ ist.

Übung 11 *(Summen als Integrale, I)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien f, g μ-integrierbar. Zeigen Sie:

(i) $\int \alpha f + \beta g \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu$ für alle $\alpha, \beta \in \mathbb{R}$. *(Linearität)*

(ii) $\int f \, d\mu \leq \int g \, d\mu$, falls $f \leq g$. *(Monotonie)*

(iii) $\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$. *(Zerlegung in Positiv- und Negativteil)*

(iv) Ist $T : A \to B$, so ist eine Funktion $h : B \to \mathbb{R}$ genau dann μ_T-integrierbar, wenn die Funktion $h \circ T : A \to \mathbb{R}$ μ-integrierbar ist, und in diesem Fall gilt

$$\int h \, d\mu_T = \int h \circ T \, d\mu.$$ *(Transformationsformel für Bildmaße)*

Übung 12 *(Summen als Integrale, II)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $f, f_n : A \to [0, \infty[$. Alle f_n seien μ-integrierbar und das Supremum der Integrale existiere in $\mathbb{R}$. Weiter gelte $f_n \uparrow f$, d. h. $\langle f_n(a) \mid n \in \mathbb{N} \rangle$ konvergiert monoton wachsend gegen $f(a)$ für alle $a \in A$. Zeigen Sie, daß f μ-integrierbar ist und daß gilt:

$$\int f \, d\mu \;=\; \sup_{n \in \mathbb{N}} \int f_n \, d\mu.$$

Übung 13 *(Summen als Integrale, III)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $f, g, f_n : A \to \mathbb{R}$. Alle f_n sowie g seien μ-integrierbar und es gelte $|f_n(a)| \leq g(a)$ für alle $n \in \mathbb{N}$ und alle $a \in A$. Weiter gelte $\lim_{n \in \mathbb{N}} f_n(a) = f(a)$ für alle $a \in A$. Zeigen Sie, daß f μ-integrierbar ist und daß gilt:

$$\int f \, d\mu \;=\; \lim_{n \in \mathbb{N}} \int f_n \, d\mu.$$

Übung 14 *(Unabhängigkeit, I)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $B \subseteq A$. Zeigen Sie, daß für jede Zerlegung $\mathfrak{Z}$ von A in Mengen mit positivem Maß gilt:

$$\mu(B) \;=\; \Sigma_{Z \in \mathfrak{Z}} \, \mu(Z) \, \mu_Z(B). \qquad \textit{(Formel von der totalen Wahrscheinlichkeit)}$$

Interpretieren Sie diese Formel weiter als eine gewichtete Summe von Wahrscheinlichkeitsmaßen.

Übung 15 *(Unabhängigkeit, II)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $B \subseteq A$ mit $\mu(B) > 0$. Weiter sei $Z \subseteq A$ mit $\mu(Z), \mu(Z^c) > 0$. Zeigen Sie:

$$\mu_B(Z) \;=\; \mu(Z) \, \mu_Z(B) / (\mu(Z) \, \mu_Z(B) + \mu(Z^c) \, \mu_{Z^c}(B)). \qquad \textit{(Formel von Bayes)}$$

Übung 16 *(Unabhängigkeit, III)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $A_1, \ldots, A_n \subseteq A$. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) Die Mengen $A_1, \ldots, A_n$ sind unabhängig.

(b) Für alle $B_1, \ldots, B_n$ mit $B_i = A_i$ oder $B_i = A_i^c$ für alle i gilt
$\mu(\bigcap_{1 \leq i \leq n} B_i) \;=\; \mu(B_1) \cdot \ldots \cdot \mu(B_n)$.

Übung 17 *(Zufallsvariable, I)*
Sei μ die Gleichverteilung auf $A = \{ 1, \ldots, n \}$. Sei $X(i) = i$ für alle $i \in A$. Bestimmen Sie $E(X)$ und $V(X)$.

Übung 18 *(Zufallsvariable, II)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und seien $X, Y, X_1, \ldots, X_n : A \to \mathbb{R}$ Zufallsvariable, deren Varianzen existieren. Zeigen Sie:

(a) $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$ für alle $\alpha, \beta \in \mathbb{R}$.

(b) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X) = E(XY) - E(X)E(Y)$,
$V(X) = \mathrm{Cov}(X, X) = E(X^2) - E(X)^2$.

(c) $\mathrm{Cov}(\alpha X + \beta, \gamma Y + \delta) = \alpha\gamma \,\mathrm{Cov}(X, Y)$ für alle $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

(d) $V(\Sigma_{1 \le i \le n} X_i) = \Sigma_{1 \le i \le n} V(X_i) + 2 \cdot \Sigma_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j)$.

(e) Sind X, Y unabhängig, so sind X, Y unkorreliert.

Übung 19 *(Zufallsvariable, III)*
Sei (A, μ) ein Wahrscheinlichkeitsraum. Begründen Sie folgende Interpretation der Kovarianz $\mathrm{Cov}(X, Y)$ zweier Zufallsvariablen $X, Y : A \to \mathbb{R}$:

„Eine große positive Kovarianz von X und Y bedeutet, daß X oft einen großen Wert annimmt, wenn Y einen großen Wert annimmt, und daß dabei die Vorzeichen von X und Y übereinstimmen."

Übung 20 *(Zufallsvariable, IV)*
Geben Sie einen Wahrscheinlichkeitsraum (A, μ) und Zufallsvariable $X, Y : A \to \mathbb{R}$ an, sodaß die beiden folgenden Eigenschaften gelten:

(i) X und Y sind unkorreliert,

(ii) X und Y sind abhängig.

Übung 21 *(Zufallsvariable, V)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $X : A \to [0, \infty[$ eine Zufallsvariable, deren Erwartungswert existiert. Zeigen Sie:

$E(X) \le \Sigma_{n \in \mathbb{N}} \mu(X > n) \le E(X) + 1$.

Übung 22 *(Zufallsvariable, VI)*
Seien (A, μ) und (A^n, ν) Wahrscheinlichkeitsräume (für ein festes $n \ge 1$). Für $1 \le i \le n$ sei $X_i((a_1, \ldots, a_n)) = a_i$. Alle X_i haben die Verteilung μ, d. h. für alle $1 \le i \le n$ und alle $a \in A$ gilt $\nu(X_i = a) = \mu(\{ a \})$. Zeigen Sie, daß die folgenden Aussagen äquivalent sind:

(a) Die Zufallsvariablen $X_1, \ldots, X_n$ sind unabhängig.

(b) ν ist das n-fache Produktmaß von μ.

Übung 23 *(Das Gesetz der großen Zahl, I)*
Sei (A, μ) ein Wahrscheinlichkeitsraum, und sei $X : A \to \mathbb{R}$ eine Zufallsvariable. Weiter sei $f : [0, \infty[\to [0, \infty[$ monoton wachsend, und es existiere $E(f \circ |X|)$. Zeigen Sie, daß für alle $\varepsilon \geq 0$ mit $f(\varepsilon) > 0$ gilt:

$$\mu(|X| \geq \varepsilon) \leq E(f \circ |X|)/f(\varepsilon).$$

Übung 24 *(Das Gesetz der großen Zahl, II)*
Bestimmen Sie jeweils für den Münzwurf und für das Würfeln mit Hilfe des Gesetzes der großen Zahl das kleinste $n \geq 1$ mit

$$\mu_n(|Y^n - E(X)| \geq 1/10) \leq 1/100,$$

wobei $X(i) = i$ für alle i gilt.

Lösungsvorschläge

Wir stellen in diesem Anhang Lösungen für einige Übungsaufgaben des ersten Abschnitts vor. Sie dienen in erster Linie dem Leser, der das Buch im Selbststudium liest.

1.1 Mathematisches Argumentieren

Übung 1:

Wir stellen dem Zwerg die Frage:

„Lügst du genau dann, wenn hinter der linken Tür der Schatz versteckt ist?"

Wir unterscheiden vier Fälle:

1. Der Zwerg lügt und der Schatz ist hinter der linken Tür.
2. Der Zwerg sagt die Wahrheit und der Schatz ist hinter der linken Tür.
3. Der Zwerg lügt und der Schatz ist nicht hinter der linken Tür.
4. Der Zwerg sagt die Wahrheit und der Schatz ist nicht hinter der linken Tür.

Die richtigen Antworten auf die Fragen sind:

1. Fall: „ja", 2. Fall: „nein", 3. Fall: „nein", 4. Fall: „ja".

Als Antwort des Zwergs erhalten wir also:

1. Fall: „nein", 2. Fall: „nein", 3. Fall: „ja", 4. Fall: „ja".

Damit ist der Schatz hinter der linken Tür, wenn wir „nein" als Antwort erhalten, und hinter der rechten Tür, wenn wir „ja" als Antwort erhalten.

Bemerkung

Auch „Bist du entweder ein Lügner oder der Schatz ist links?" führt zur Entdeckung des Schatzes, mit einem ausschließlichen „entweder … oder". (Vgl. hierzu auch Übung 7.)

Übung 5:

Die erste Aussage ist eine Kurzform von

„Hunde dürfen nicht an Bord und Katzen dürfen nicht an Bord".

Die verwendete Tautologie ist $(A \to C) \wedge (B \to C) \leftrightarrow (A \vee B) \to C$, siehe (T12) in der Tabelle aussagenlogischer Tautologien.

O. Deiser, *Grundbegriffe der wissenschaftlichen Mathematik*, Springer-Lehrbuch,
DOI 10.1007/978-3-642-11489-2, © Springer-Verlag Berlin Heidelberg 2010

Übung 9:

Die Negation läßt sich nicht mit Hilfe der Junktoren $\rightarrow, \wedge, \vee$ definieren. Denn für alle Aussagen B und C haben $B \rightarrow C$, $B \wedge C$ und $B \vee C$ den Wahrheitswert w, falls B und C den Wahrheitswert w haben. Ist also D eine aus A mit Hilfe von $\rightarrow, \wedge, \vee$ aufgebaute Aussage (z. B. $D = A \rightarrow (A \rightarrow A) \wedge A$), so hat D den Wahrheitswert w, falls w der Wahrheitswert von A ist. Dagegen besitzt $\neg A$ den Wahrheitswert f, falls A den Wahrheitswert w besitzt.

Übung 10:

Die Wahrheitstafeln der dreistelligen Junktoren $*_{\geq 2}$ und $*_{0,3}$ lauten:

$*_{\geq 2}$	A	B	C
w	w	w	w
w	w	w	f
w	w	f	w
w	f	w	w
f	w	f	f
f	f	w	f
f	f	f	w
f	f	f	f

$*_{0,3}$	A	B	C
w	w	w	w
f	w	w	f
f	w	f	w
f	f	w	w
f	w	f	f
f	f	w	f
f	f	f	w
w	f	f	f

Wir können diese Junktoren mit Hilfe der üblichen Junktoren z. B. definieren durch

$*_{\geq 2}(A, B, C) = (A \wedge B) \vee (B \wedge C) \vee (A \wedge C)$,

$*_{0,3}(A, B, C) = A \vee B \vee C \rightarrow A \wedge B \wedge C$, oder auch

$*_{0,3}(A, B, C) = (A \rightarrow B \wedge C) \wedge (B \rightarrow A \wedge C) \wedge (C \rightarrow A \wedge B)$.

Übung 12:

zu (i): Diese Aussage ist widerlegbar. Sei nämlich A eine erfüllbare, aber nicht allgemeingültige Aussage, und sei $B = A \wedge \neg A$. Dann ist $A \rightarrow B$ erfüllbar, denn jede Zeile von $A \rightarrow B$, die für A den Wert f aufweist, besitzt den Wert w in der Ergebnisspalte. Dagegen ist B nicht erfüllbar.

zu (ii): Diese Aussage ist beweisbar. Sei nämlich Z eine Zeile der Wahrheitstafel von A, die den Wert w in der Ergebnisspalte aufweist, und sei Z′ eine Zeile der Wahrheitstafel von $A \rightarrow B$, die dieselbe w-f-Belegung von A wie die Zeile Z besitzt. Da $A \rightarrow B$ nach Voraussetzung eine Tautologie ist und A den Wert w in Z′ erhält, muß notwendig auch B den Wert w in Z′ erhalten. Damit ist B erfüllbar (bei der durch Z′ gegebenen w-f-Belegung von B).

zu (iii): Diese Aussage ist beweisbar. Denn nach Voraussetzung hat jede Zeile der Wahrheitstafel von A den Wert f und jede Zeile der Wahrheitstafel von $A \vee B$ den Wert w in der Ergebnisspalte. Ist aber f der Wahrheitswert von A und w der Wahrheitswert von $A \vee B$, so ist notwendig w der Wahrheitswert von B. Damit weist die Ergebnisspalte von B nur die Werte w auf.

Übung 16, (iii):

Wir beweisen $\neg\neg\neg A \to \neg A$ mit Hilfe von (S1) – (S5) (ohne (S7)). Wir nehmen $\neg A = A \to \bot$ und A an. Modus ponens liefert $\bot$ (unter Annahme von $\neg A$ und A). Nach Abbinden von $A \to \bot$ haben wir $\neg\neg A = (A \to \bot) \to \bot$ bewiesen (unter Annahme von A). Nun nehmen wir $\neg\neg\neg A = \neg\neg A \to \bot$ an und erhalten $\bot$ mit modus ponens (unter Annahme von A und $\neg\neg\neg A$). Abbinden von A liefert $\neg A = A \to \bot$ (unter Annahme von $\neg\neg\neg A$), und Abbinden von $\neg\neg\neg A$ liefert schließlich $\neg\neg\neg A \to \neg A$ ohne Abhängigkeit von Annahmen.

Bemerkung

Wir können also $\neg\neg B \to B$ ohne (S7) beweisen, wenn B von der Form $\neg A$ ist. Im allgemeinen Fall kann aber, wie man mit weitergehenden Methoden zeigen kann, auf die Regel (S7) nicht verzichtet werden.

Als Beweisbaum notiert lautet obiges Argument:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \to \bot \;{}^{①} \qquad A \;{}^{②}}{\bot}\ (S4)}{(A \to \bot) \to \bot}\ (S5\text{: Abbinden von Annahme ①}) \qquad \dfrac{((A \to \bot) \to \bot) \to \bot \;{}^{③}}{\bot}\ (S4)}{A \to \bot}\ (S5\text{: Abbinden von Annahme ②})}{(((A \to \bot) \to \bot) \to \bot) \to (A \to \bot)}\ (S5\text{: Abbinden von Annahme ③})$$

Bemerkung

Das Argument enthält auch einen Beweis von $A \to \neg\neg A$, also von Teil (ii) der Übung: Nach dem Abbinden von Annahme ① ist $\neg\neg A$ unter Annahme von A bewiesen. Abbinden von A liefert dann $A \to \neg\neg A$ ohne Abhängigkeit von Annahmen.

Übung 19:

zu (a): $\exists x_1, x_2, x_3\, (x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_2 \neq x_3)$.

zu (b): Wir verwenden die Aussage φ aus Teil (a) und formulieren:

$$\varphi \wedge \forall x_1, x_2, x_3, x_4\, (x_1 = x_2 \vee x_1 = x_3 \vee x_1 = x_4 \vee x_2 = x_3 \vee x_2 = x_4 \vee x_3 = x_4).$$

Übung 20:

zu (i): $\exists x\,(M(x) \wedge \forall y\,(Z(y) \rightarrow L(x, y)))$,

zu (ii): $\forall x\,(M(x) \rightarrow \exists y\,(Z(y) \wedge L(x, y)))$,

zu (iii): $\exists x, y\,(M(x) \wedge Z(y) \wedge \neg L(x, y))$,

zu (iv): $\exists x_1\,(M(x_1) \wedge \forall x_2\,(M(x_2) \wedge \forall y\,(Z(y) \rightarrow \neg L(x_2, y)) \leftrightarrow x_1 = x_2))$,

zu (v): $\forall x_1, x_2\,(M(x_1) \wedge M(x_2) \wedge x_1 \neq x_2 \rightarrow \exists y\,(Z(y) \wedge \neg L(x_1, y) \wedge L(x_2, y)))$.

1.2 Mengen

Übung 1:

Sei $S = \{ x \mid \neg \exists y\,(x \in y \wedge y \in x) \}$. Wir unterscheiden zwei Fälle.

1. Fall: $S \in S$. Dann gibt es ein y mit $S \in y$ und $y \in S$, nämlich $y = S$. Also ist $S \notin S$ nach Definition von S, *Widerspruch*.

2. Fall: $S \notin S$. Nach Definition von S gibt es dann ein y mit $S \in y$ und $y \in S$. Dann gilt aber $y \in S$ und $S \in y$. Also ist $y \notin S$ nach Definition von S, *im Widerspruch* zu $y \in S$.

In beiden Fällen ergibt sich also ein Widerspruch. S kann also, wie die Russell-Zermelo-Komprehension R, keine Menge sein.

Übung 3:

$(a, b) = (c, d)$ *impliziert* $a = c$ und $b = d$:

Sei also $(a, b) = (c, d)$. Dann gilt $\{ \{ a \}, \{ a, b \} \} = \{ \{ c \}, \{ c, d \} \}$.

Ist $a = b$, so ist $\{ \{ a \}, \{ a, b \} \} = \{ \{ a \} \} = \{ \{ c \}, \{ c, d \} \}$, also $\{ c \} = \{ a \}$ und $\{ c, d \} = \{ a \}$. Also ist $c = d = a = b$ in diesem Fall.

Andernfalls ist $a \neq b$. Dann ist $\{ a, b \} \neq \{ c \}$, also $\{ a, b \} = \{ c, d \}$. Dann ist aber $c \neq d$ und damit $\{ a \} \neq \{ c, d \}$. Also ist $\{ a \} = \{ c \}$ und somit $a = c$. Aus $\{ a, b \} = \{ c, d \} = \{ a, d \}$ und $a \neq b$ folgt dann aber auch $b = d$.

$a = c$ und $b = d$ *impliziert* $(a, b) = (c, d)$:

Ist $a = c$ und $b = d$, so ist $(a, b) = \{ \{ a \}, \{ a, b \} \} = \{ \{ c \}, \{ c, d \} \} = (c, d)$.

Übung 5:

Wir beweisen die Aussage (i). Die Aussage (ii) wird analog bewiesen.

$(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$:

Sei $x \in (A \cup B) \cap C$. Dann ist $x \in A$ oder $x \in B$. Zudem ist $x \in C$.
Ist $x \in A$, so ist $x \in A \cap C$ und damit auch $x \in (A \cap C) \cup (B \cap C)$.
Ist $x \in B$, so ist $x \in B \cap C$ und damit ebenfalls $x \in (A \cap C) \cup (B \cap C)$.

$(A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$:

Sei $x \in (A \cap C) \cup (B \cap C)$. Dann ist $x \in A \cap C$ oder $x \in B \cap C$.
Ist $x \in A \cap C$, so ist $x \in A \cup B$ und $x \in C$ und damit $x \in (A \cup B) \cap C$.
Ist $x \in B \cap C$, so ist $x \in A \cup B$ und $x \in C$, also ebenfalls $x \in (A \cup B) \cap C$.

Damit ist die Aussage (i) bewiesen. Ersetzen wir in (i) formal

$\wedge$ für $\cap$, $\vee$ für $\cup$, $\leftrightarrow$ für $=$,

und lesen wir den so entstehenden Ausdruck aussagenlogisch, so ergibt sich die Tautologie $(A \vee B) \wedge C \leftrightarrow (A \wedge C) \vee (B \wedge C)$.

Sei allgemein „$M_1 \subseteq M_2$“ oder „$M_1 = M_2$“ eine Rechenregel für Mengen, wobei M_1 und M_2 Mengenterme sind, die mit Hilfe des Durchschnitts $\cap$, der Vereinigung $\cup$, der relativen Komplementbildung $(\,)^c$ bezüglich einer Grundmenge M, der leeren Menge $\varnothing$ und der Grundmenge M gebildet sind. (Im obigen Fall wäre also $M_1 = (A \cup B) \cap C$, $M_2 = (A \cap C) \cup (B \cap C)$ und unsere Regel ist von der Form „$M_1 = M_2$“; eine Grundmenge muß hier nicht angegeben werden, wir können aber $M = A \cup B \cup C$ setzen.)
Wir führen dann die folgenden Ersetzungen durch:

$\wedge$ für $\cap$, $\vee$ für $\cup$, $\neg$ für $(\,)^c$, $\perp$ für $\varnothing$, $\neg \perp$ für M, $\rightarrow$ für $\subseteq$, $\leftrightarrow$ für $=$.

Die Tautologie $\neg (A \vee B) \leftrightarrow \neg A \wedge \neg B$ geht in dieser Weise aus der Rechenregel $(A \cup B)^c = A^c \cap B^c$ hervor. Die Regel $A \cap B \subseteq A$ entspricht der Tautologie $A \wedge B \rightarrow A$, die Regel $A = A$ der Tautologie $A \leftrightarrow A$, die Regel $(A^c)^c = A$ der Tautologie $\neg \neg A \leftrightarrow A$, die Regel $\varnothing \subseteq A$ der Tautologie $\perp \rightarrow A$, die Regel $A \cup A^c = M$ der Tautologie $A \vee \neg A \leftrightarrow \neg \perp$, die sich zu $A \vee \neg A$ verkürzen läßt (denn für alle Aussagen C ist C genau dann eine Tautologie, wenn $C \leftrightarrow \neg \perp$ eine Tautologie ist).

Es gilt nun allgemein, daß „$M_1 \subseteq M_2$“ bzw. „$M_1 = M_2$“ genau dann eine allgemeingültige Rechenregel für Mengen ist, wenn der Ersetzungsvorgang eine aussagenlogische Tautologie liefert. Um dies einzusehen, ordnen wir einem Ausdruck „$x \in C$“ den Wert w zu, falls x ein Element von C ist, und den Wert f sonst. Diese Zuordnung entspricht für $A \cap B$, $A \cup B$, A^c, $\varnothing$ und M genau den Wahrheitstafeln für $\wedge$, $\vee$, $\neg$, $\perp$ und $\neg \perp$; so erhält zum Beispiel „$x \in (A \cap B)$“ genau dann den Wert w, wenn „$x \in A$“ und „$x \in B$“ den Wert w erhalten; „$x \in \varnothing$“ erhält, wie $\perp$, den Wert f, und „$x \in M$“ erhält, wie $\neg \perp$, den Wert w, usw. Vermöge dieser Entsprechung erhält „$x \in M_1$“ den Wert der Zeile der M_1 zugeordneten Wahrheitstafel, die durch die Werte von „$x \in A_1$“, …, „$x \in A_n$“ gegeben ist, wobei A_1, …, A_n die Mengen sind, aus denen M_1 gebildet ist. Gleiches gilt für „$x \in M_2$“. Damit gilt „$x \in M_1$ gdw $x \in M_2$“ für alle x, d. h. $M_1 = M_2$, wenn die „$M_1 = M_2$“ zugeordnete Aussage eine Tautologie ist. Ist umgekehrt diese Aussage keine Tautologie, so können wir für ein beliebiges x Mengen A_1, …, A_n der Form $A_i = \{ x \}$ oder $A_i = \varnothing$ definieren, so daß die Werte von „$x \in A_i$“ einer f-Zeile der Wahrheitstafel entsprechen. Dann sind die Werte von „$x \in M_1$“ und „$x \in M_2$“ für die durch die A_i definierten Mengen M_1 und M_2 verschieden, und damit gilt $M_1 \neq M_2$. Ähnliche Überlegungen gelten für „$M_1 \subseteq M_2$“.

Übung 7:

zu (i): Es gilt $A \cap B = A$ genau dann, wenn $A \subseteq B$: Denn ist $A \subseteq B$, so ist offenbar $A \cap B = A$. Ist umgekehrt $A \cap B = A$, so ist jedes Element von A ein Element von $A \cap B$ und damit insbesondere ein Element von B, d. h. es gilt $A \subseteq B$.

zu (ii): Es gilt $A \cup B = A$ genau dann, wenn $B \subseteq A$: Denn ist $B \subseteq A$, so ist offenbar $A \cup B = A$. Ist umgekehrt $A \cup B = A$, so ist jedes Element von $A \cup B$ ein Element von A und damit ist insbesondere jedes Element von B ein Element von A, d. h. es gilt $B \subseteq A$.

zu (iii): Es gilt

$$B - (B - A) = \{ x \in B \mid x \notin (B - A) \} = \{ x \in B \mid \text{non}(x \in B \wedge x \notin A) \} =$$

$$\{ x \in B \mid x \notin B \vee x \in A \} = \{ x \in B \mid x \in A \} = A \cap B.$$

Also gilt $B - (B - A) = A$ genau dann, wenn $A \cap B = A$. Nach (i) ist dies genau dann der Fall, wenn $A \subseteq B$.

Übung 10:

Wir geben zwei Beweise für die Aussage:

(+) $A_1 \Delta \ldots \Delta A_n = \{ a \mid$ die Anzahl aller i mit $a \in A_i$ ist ungerade $\}$.

Erster Beweis

Sei a beliebig, und sei k die Anzahl aller i mit $a \in A_i$.

Wir nehmen zunächst an, daß k gerade ist. Sei also $k = 2m$.
Wir können die Mengen A_i beliebig umordnen und klammern (vgl. Übung 9). Es gilt also

$$A_1 \Delta \ldots \Delta A_n = (A'_1 \Delta A'_2) \Delta \ldots \Delta (A'_{2m-1} \Delta A'_{2m}) \Delta A'_{2m+1} \Delta \ldots \Delta A'_n,$$

wobei die Mengen A'_i eine Umordnung der Mengen A_i sind mit $a \in A'_i$ für alle $1 \leq i \leq 2m$ und $a \notin A'_i$ für alle $2m < i \leq n$. Wir setzen nun $B_i = A'_{2i-1} \Delta A'_{2i}$ für $1 \leq i \leq m$. Dann gilt

$$A_1 \Delta \ldots \Delta A_n = B_1 \Delta \ldots \Delta B_m \Delta A'_{2m+1} \Delta \ldots \Delta A'_n,$$

und a ist kein Element irgendeiner der (n – m)-vielen Mengen der rechten Seite der Gleichung. Damit ist $a \notin A_1 \Delta \ldots \Delta A_n$ (denn für alle Mengen C_i ist $C_1 \Delta \ldots \Delta C_n \subseteq C_1 \cup \ldots \cup C_n$).

Ist dagegen k ungerade, also $k = 2m + 1$ für ein gewisses m, so erhalten wir in ähnlicher Weise eine Darstellung

$$A_1 \Delta \ldots \Delta A_n = B_1 \Delta \ldots \Delta B_m \Delta A'_{2m+1} \Delta \ldots \Delta A'_n$$

in der a ein Element von A'_n und sonst keiner Menge der rechten Seite ist.
Wir setzen nun $B = B_1 \Delta \ldots \Delta B_m \Delta A'_{2m+1} \Delta \ldots \Delta A'_{n-1}$.
Dann ist $A_1 \Delta \ldots \Delta A_n = B \Delta A'_n$, und es gilt $a \notin B$ und $a \in A'_n$.
Also ist $a \in B \Delta A'_n$ und damit $a \in A_1 \Delta \ldots \Delta A_n$.

Zweiter Beweis

Wir zeigen die Aussage durch Induktion nach $n \geq 2$.

Induktionsanfang n = 2:

Es gilt $A_1 \Delta A_2 = (A_1 - A_2) \cup (A_2 - A_1) =$

$\{ a \mid$ die Anzahl aller i mit $a \in A_i$ ist gleich 1 $\} =$

$\{ a \mid$ die Anzahl aller i mit $a \in A_i$ ist ungerade $\}$.

Induktionsschritt von n nach n + 1, $n \geq 2$:

Es gelte (+) für $A_1 \Delta \ldots \Delta A_n$ (Induktionsvoraussetzung). Dann gilt:

$A_1 \Delta \ldots \Delta A_{n+1} = (A_1 \Delta \ldots \Delta A_n) \Delta A_{n+1} =$

$\{ a \mid$ die Anzahl aller $1 \leq i \leq n$ mit $a \in A_i$ ist ungerade $\} \Delta A_{n+1} =$

$(\{ a \mid$ die Anzahl aller $1 \leq i \leq n$ mit $a \in A_i$ ist ungerade $\} - A_{n+1}) \cup$

$(A_{n+1} - \{ a \mid$ die Anzahl aller $1 \leq i \leq n$ mit $a \in A_i$ ist ungerade $\}) =$

$\{ a \mid$ die Anzahl aller $1 \leq i \leq n$ mit $a \in A_i$ ist ungerade und $a \notin A_{n+1} \} \cup$

$\{ a \mid a \in A_{n+1}$ und die Anzahl aller $1 \leq i \leq n$ mit $a \in A_i$ ist gerade $\} =$

$\{ a \mid$ die Anzahl aller $1 \leq i \leq n + 1$ mit $a \in A_i$ ist ungerade $\}$.

Also gilt (+) für $A_1 \Delta \ldots \Delta A_{n+1}$.

Übung 14:

Die Aussage (i) ist widerlegbar. Sei nämlich $A = \{ 0 \}$ und $B = \{ 1 \}$. Dann ist

$\mathcal{P}(A \cup B) = \mathcal{P}(\{ 0, 1 \}) = \{ \emptyset, \{ 0 \}, \{ 1 \}, \{ 0, 1 \} \}$, aber

$\mathcal{P}(A) \cup \mathcal{P}(B) = \{ \emptyset, \{ 0 \} \} \cup \{ \emptyset, \{ 1 \} \} = \{ \emptyset, \{ 0 \}, \{ 1 \} \}$.

Dagegen ist (ii) beweisbar. Denn für alle Mengen A, B, C gilt:

$C \in \mathcal{P}(A \cap B)$ *gdw* $C \subseteq A \cap B$ *gdw* $C \subseteq A$ und $C \subseteq B$ *gdw*

$C \in \mathcal{P}(A)$ und $C \in \mathcal{P}(B)$ *gdw* $C \in \mathcal{P}(A) \cap \mathcal{P}(B)$.

Übung 18:

Wir beweisen (i), also $\bigcup (\mathcal{A} \cup \mathcal{B}) = \bigcup \mathcal{A} \cup \bigcup \mathcal{B}$.

$\bigcup (\mathcal{A} \cup \mathcal{B}) \subseteq \bigcup \mathcal{A} \cup \bigcup \mathcal{B}$:

Sei $x \in \bigcup (\mathcal{A} \cup \mathcal{B})$. Dann existiert ein $C \in \mathcal{A} \cup \mathcal{B}$ mit $x \in C$.
Ist $C \in \mathcal{A}$, so ist $x \in \bigcup \mathcal{A}$ und damit $x \in \bigcup \mathcal{A} \cup \bigcup \mathcal{B}$.
Ist $C \in \mathcal{B}$, so ist $x \in \bigcup \mathcal{B}$ und damit ebenfalls $x \in \bigcup \mathcal{A} \cup \bigcup \mathcal{B}$.

$\bigcup \mathcal{A} \cup \bigcup \mathcal{B} \subseteq \bigcup (\mathcal{A} \cup \mathcal{B})$:

Sei $x \in \bigcup \mathcal{A} \cup \bigcup \mathcal{B}$. Ist $x \in \bigcup \mathcal{A}$, so existiert ein $A \in \mathcal{A}$ mit $x \in A$.
Dann ist aber $A \in \mathcal{A} \cup \mathcal{B}$ und damit $x \in \bigcup (\mathcal{A} \cup \mathcal{B})$.
Der Fall $x \in \bigcup \mathcal{B}$ wird analog bewiesen.

Dagegen sind die Aussagen (ii), (iii) und (iv) widerlegbar. Wir betrachten hierzu die Mengensysteme

$$\mathcal{A}_1 = \{\{0\},\{1\}\}, \quad \mathcal{B}_1 = \{\{0\},\{2\}\},$$

$$\mathcal{A}_2 = \{\{0,1\}\}, \quad \mathcal{B}_2 = \{\{0,2\}\},$$

$$\mathcal{A}_3 = \{\{0\}\}, \quad \mathcal{B}_3 = \{\{1\}\}.$$

Dann ist $\mathcal{A}_1 \cap \mathcal{B}_1 = \{\{0\}\} \neq \emptyset$ und $\mathcal{A}_3, \mathcal{B}_3 \neq \emptyset$. Zudem gilt:

$$\bigcap(\mathcal{A}_1 \cap \mathcal{B}_1) = \bigcap\{\{0\}\} = \{0\} \neq \emptyset = \emptyset \cap \emptyset = \bigcap \mathcal{A}_1 \cap \bigcap \mathcal{B}_1,$$

$$\bigcup(\mathcal{A}_2 \cap \mathcal{B}_2) = \bigcup \emptyset = \emptyset \neq \{0\} = \{0,1\} \cap \{0,2\} = \bigcup \mathcal{A}_2 \cap \bigcup \mathcal{B}_2,$$

$$\bigcap(\mathcal{A}_3 \cup \mathcal{B}_3) = \bigcap\{\{0\},\{1\}\} = \emptyset \neq \{0\} \cup \{1\} = \bigcap \mathcal{A}_3 \cup \bigcap \mathcal{B}_3.$$

Damit bildet $\mathcal{A}_1, \mathcal{B}_1$ ein Gegenbeispiel zu (ii), $\mathcal{A}_2, \mathcal{B}_2$ ein Gegenbeispiel zu (iii) und $\mathcal{A}_3, \mathcal{B}_3$ ein Gegenbeispiel zu (iv).

Bemerkung

Leicht nachzuweisen sind dagegen die Inklusionen

$$\bigcap(\mathcal{A} \cap \mathcal{B}) \supseteq \bigcap \mathcal{A} \cap \bigcap \mathcal{B}, \quad \text{falls } \mathcal{A} \cap \mathcal{B} \neq \emptyset,$$

$$\bigcup(\mathcal{A} \cap \mathcal{B}) \subseteq \bigcup \mathcal{A} \cap \bigcup \mathcal{B},$$

$$\bigcap(\mathcal{A} \cup \mathcal{B}) \subseteq \bigcap \mathcal{A} \cup \bigcap \mathcal{B}, \quad \text{falls } \mathcal{A}, \mathcal{B} \neq \emptyset.$$

Übung 23:

Es kann keine Mengenalgebra $\mathcal{A}$ auf A geben, die genau 5 Elemente besitzt. *Annahme*,

$$\mathcal{A} = \{\emptyset, A, B, C, D\}$$

ist eine Mengenalgebra auf A mit genau 5 Elementen. Dann gilt aber $B^c = A - B \in \mathcal{A}$, und da B von $\emptyset$ und A verschieden ist, ist B^c von $\emptyset$, A und B verschieden. Also gilt $B^c = C$ oder $B^c = D$. Ohne Einschränkung sei $B^c = C$. Analog ist aber auch $D^c = A - D \in \mathcal{A}$ und verschieden von $\emptyset$, A und D. Also gilt $D^c = B$ oder $D^c = C$. Im ersten Fall ist $D = B^c = C$, im zweiten Fall $D = C^c = B$, *im Widerspruch* zu $D \neq C$ und $D \neq B$.

1.3 Relationen und Funktionen

Übung 1:

R ist reflexiv: Für alle n gilt n R n, da $|n - n| = 0$ gerade ist.

R ist nicht irreflexiv: Es gilt 0 R 0.

R ist symmetrisch: Gilt n R m, so ist $|n - m|$ gerade. Wegen $|n - m| = |m - n|$ gilt dann aber auch m R n.

R ist nicht antisymmetrisch: Es gilt 0 R 2 und 2 R 0, aber $0 \neq 2$.

R ist transitiv: Gilt n R m und m R k, so sind $|n - m|$ und $|m - k|$ gerade. Also gibt es ganze Zahlen a, b mit $n - m = 2a$ und $m - k = 2b$. Dann gilt $n - k = n - m + m - k = 2a + 2b = 2(a + b)$. Also ist $|n - k|$ gerade, und damit gilt n R k.

S ist nicht reflexiv: Es gilt non(0 S 0), da $|0 - 0| = 0$ gerade ist.

S ist irreflexiv: Für alle n ist $|n - n| = 0$ gerade, d. h. es gilt non(n S n).

S ist symmetrisch: Gilt n S m, so ist $|n - m|$ ungerade. Wegen $|n - m| = |m - n|$ gilt dann aber auch m S n.

S ist nicht antisymmetrisch: Es gilt 0 R 1 und 1 R 0, aber $0 \neq 1$.

S ist nicht transitiv: Es gilt 0 S 1 und 1 S 2, aber non(0 S 2).

Übung 3:

Wir definieren ein Mengensystem $\mathcal{A}$ durch

$$\mathcal{A} = \{ S \subseteq A \times A \mid S \text{ ist transitiv}, S \supseteq R \}.$$

Wegen $A \times A \in \mathcal{A}$ ist $\mathcal{A} \neq \emptyset$. Wir können also definieren:

$$R^* = \bigcap \mathcal{A}.$$

Wir zeigen, daß R* die gewünschten Eigenschaften besitzt. Zunächst gilt $R^* \subseteq A \times A$, also ist R* eine Relation auf A. Es gilt $R \subseteq R^*$, da $R \subseteq S$ für alle $S \in \mathcal{A}$. Weiter ist R* transitiv, denn sind $a, b, c \in A$ mit a R* b und b R* c, so gilt a S b und b S c für alle $S \in \mathcal{A}$, und damit auch a S c für alle $S \in \mathcal{A}$ (da alle $S \in \mathcal{A}$ transitiv sind). Dann ist aber a R* c nach Definition von R*. Ist schließlich $R' \supseteq R$ eine transitive Relation auf A, so ist $R' \in \mathcal{A}$ und damit $R' \supseteq \bigcap \mathcal{A} = R^*$. Damit haben wir bewiesen, daß R* die $\subseteq$-kleinste transitive Relation auf A ist mit $R^* \supseteq R$.

Bemerkung

Eine konstruktivere Darstellung von R* kann man mit Hilfe von Rekursion und der Verknüpfung $\circ$ für Relationen aus Aufgabe 4 erhalten.
Wir definieren hierzu rekursiv $R^1 = R$, $R^{n+1} = R^n \circ R$ für alle $n \geq 1$.
Man kann nun zeigen, daß $R^* = \bigcup_{n \geq 1} R^n$ gilt, d. h. es gilt a R* b genau dann, wenn ein $n \geq 1$ existiert mit a R^n b. Wir kommen auf diese Darstellung in Kapitel 3.3 über Matrizen zurück.

Übung 8:

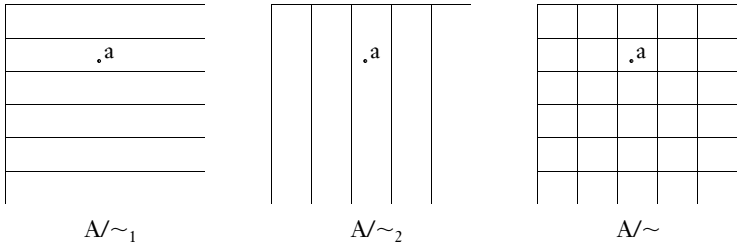
$\sim$ *ist reflexiv:* Sei $a \in A$. Dann gilt $a \sim_1 a$ und $a \sim_2 a$ aufgrund der Reflexivität von $\sim_1$ und $\sim_2$. Nach Definition von $\sim$ gilt dann aber $a \sim a$.

$\sim$ *ist symmetrisch:* Seien $a, b \in A$ mit $a \sim b$. Dann gilt $a \sim_1 b$ und $a \sim_2 b$. Aufgrund der Symmetrie von $\sim_1$ und $\sim_2$ gilt dann auch $b \sim_1 a$ und $b \sim_2 a$. Also gilt $b \sim a$ nach Definition von $\sim$.

$\sim$ *ist transitiv:* Seien $a, b, c \in A$ mit $a \sim b$ und $b \sim c$. Dann gilt $a \sim_i b$ und $b \sim_i c$ für $i = 1, 2$. Aufgrund der Transitivität von $\sim_i$ gilt dann also $a \sim_i c$ für $i = 1, 2$, also $a \sim c$ nach Definition von $\sim$.

Zur Visualisierung von $\sim$ betrachten wir die durch $\sim_1$ und $\sim_2$ definierten Zerlegungen von A, also die Faktorisierungen $A/\sim_1$ und $A/\sim_2$. Dann ist die Faktorisierung $A/\sim$ die gröbste gemeinsame Verfeinerung der Zerlegungen $A/\sim_1$ und $A/\sim_2$: die Äquivalenzklassen von $\sim$ sind genau die Schnitte der Äquivalenzklassen von $\sim_1$ und $\sim_2$. Denn für alle $a \in A$ gilt

$$a/\sim \;=\; \{\, b \in A \mid a \sim_1 b \text{ und } a \sim_2 b \,\} = \{\, b \in A \mid a \sim_1 b \,\} \cap \{\, b \in A \mid a \sim_2 b \,\} \;=\; a/\sim_1 \cap\, a/\sim_2.$$



Bemerkung

Sind die Faktorisierungen endlich und hat $A/\sim_1$ genau n und $A/\sim_2$ genau m Elemente, so hat $A/\sim$ mindestens $k = \max(n, m)$ und höchstens $n \cdot m$ Elemente. Im Fall $\sim_1 = \sim_2$ gilt $\sim = \sim_1$ und dann hat $A/\sim$ genau n Elemente. Diese Überlegung zeigt die Grenzen der Genauigkeit des obigen Diagramms, das $n \cdot m$ Äquivalenzklassen suggeriert.

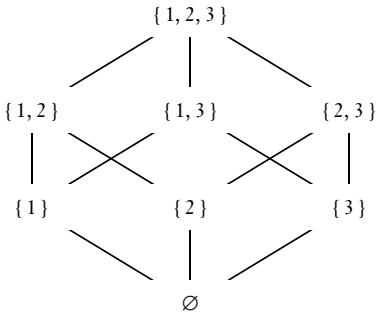
Übung 12:

$\subseteq$ *ist reflexiv auf* $\mathcal{A}$*:* Sei $B \in \mathcal{A}$. Dann gilt offenbar $B \subseteq B$.

$\subseteq$ *ist antisymmetrisch auf* $\mathcal{A}$*:* Seien $B, C \in \mathcal{A}$ mit $B \subseteq C$ und $C \subseteq B$. Dann gilt $B = C$ nach dem Extensionalitätsprinzip.

$\subseteq$ *ist transitiv auf* $\mathcal{A}$*:* Seien $B, C, D \in \mathcal{A}$ mit $B \subseteq C$ und $C \subseteq D$. Sei $x \in B$ beliebig. Dann ist $x \in C$ wegen $B \subseteq C$. Weiter ist dann $x \in D$ wegen $C \subseteq D$. Dies zeigt, daß $B \subseteq D$.

Wir können die Inklusionsordnung für A = { 1, 2, 3 } und $\mathcal{A} = \mathcal{P}(A)$ durch das rechtsstehende Diagramm visualisieren (vgl. hierzu die Bemerkungen im Abschnitt über Ordnungen).



Übung 16:

Es gilt $\mathrm{dom}(h \circ (g \circ f)) = \mathrm{dom}(g \circ f) = \mathrm{dom}(f) = \mathrm{dom}((h \circ g) \circ f)$.
Für alle $a \in A$ gilt zudem:

$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a)))$,

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

Damit sind die Funktionen $h \circ (g \circ f)$ und $(h \circ g) \circ f$ gleich.

Übung 19:

(i) $\curvearrowright$ *(ii):*

Sei also $f : A \to B$ injektiv, und sei $a_0 \in A$ beliebig. Wir definieren $g : B \to A$ durch $g(b) =$ „das eindeutige $a \in A$ mit $f(a) = b$“, falls $b \in \mathrm{rng}(f)$, und $g(b) = a_0$ andernfalls. Dann gilt $(g \circ f)(a) = g(f(a)) = a$ für alle $a \in A$, d. h. $g \circ f = \mathrm{id}_A$.

(ii) $\curvearrowright$ *(i):*

Sei $g : B \to A$ mit $g \circ f = \mathrm{id}_A$. Seien $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$. Dann gilt $a_1 = g(f(a_1)) = g(f(a_2)) = a_2$. Also ist f injektiv.

Für Surjektionen gilt folgende Äquivalenz (hier auch für $A = \emptyset$):

(a) $f : A \to B$ ist surjektiv.

(b) Es gibt ein $g : B \to A$ mit $f \circ g = \mathrm{id}_B$.

Zum Beweis:

(a) $\curvearrowright$ *(b):* Sei also $f : A \to B$ surjektiv. Wir definieren $g : B \to A$ durch:

$g(b) =$ „ein $a \in A$ mit $f(a) = b$“ für alle $b \in B$.

Wegen $\mathrm{rng}(f) = B$ existiert für alle $b \in B$ in der Tat ein $a \in A$ mit $f(a) = b$. Nach Konstruktion gilt dann $(f \circ g)(b) = f(g(b)) = b$ für alle $b \in B$, d. h. es gilt $f \circ g = \mathrm{id}_B$.

(b) $\curvearrowright$ *(a):* Sei $g : B \to A$ mit $f \circ g = \mathrm{id}_B$. Sei $b \in B$ beliebig. Dann gilt $f(g(b)) = b$, also ist $b \in \mathrm{rng}(f)$. Damit ist f surjektiv.

Übung 21:

$\leq$ *ist wohldefiniert:* Seien $a, b \in A$ mit $a \mathrel{R} b$. Seien $a \sim a'$ und $b \sim b'$. Wir zeigen, daß $a' \mathrel{R} b'$. Wegen $a \sim a'$ gilt $a' \mathrel{R} a$, und wegen $b \sim b'$ gilt $b \mathrel{R} b'$. Also haben wir $a' \mathrel{R} a$, $a \mathrel{R} b$ und $b \mathrel{R} b'$. Aufgrund der Transitivität von R gilt dann aber wie gewünscht $a' \mathrel{R} b'$.

$\leq$ *ist reflexiv:* Sei $a/\sim \in A/\sim$. Wegen R reflexiv gilt $a \mathrel{R} a$, also $a/\sim \leq a/\sim$.

$\leq$ *ist antisymmetrisch:* Seien $a/\sim, b/\sim \in A/\sim$ mit $a/\sim \leq b/\sim$ und $b/\sim \leq a/\sim$. Dann gilt $a \mathrel{R} b$ und $b \mathrel{R} a$, also $a \sim b$ und damit $a/\sim = b/\sim$.

$\leq$ *ist transitiv:* Seien $a/\sim, b/\sim, c/\sim \in A/\sim$ mit $a/\sim \leq b/\sim$ und $b/\sim \leq c/\sim$. Dann gilt $a \mathrel{R} b$ und $b \mathrel{R} c$. Wegen R transitiv gilt dann $a \mathrel{R} c$ und damit $a/\sim \leq c/\sim$.

Übung 23:

F ist injektiv:

Seien $a, c \in A$ mit $F(a) = F(c)$. Dann gilt $\{ b \in A \mid b \leq a \} = \{ d \in A \mid d \leq c \}$. Insbesondere ist $a \in \{ d \in A \mid d \leq c \}$ und $c \in \{ b \in A \mid b \leq a \}$. Dann ist aber $a \leq c$ und $c \leq a$, und damit $a = c$ nach Antisymmetrie von $\leq$.

Nach Definition von $\mathcal{A} = \mathrm{rng}(F)$ ist also $F: A \to \mathcal{A}$ bijektiv. Es bleibt zu zeigen, daß die Funktion F die partiellen Ordnungen $\leq$ auf A und $\subseteq$ auf $\mathcal{A}$ respektiert, d. h. daß für alle $a, c \in A$ gilt:

(+) $a \leq c$ *gdw* $F(a) \subseteq F(c)$.

Aber für alle $a, c \in A$ gilt:

$a \leq c$ *gdw* „für alle $b \in A$ mit $b \leq a$ gilt $b \leq c$“ *gdw*

$\{ b \in A \mid b \leq a \} \subseteq \{ d \in A \mid d \leq c \}$ *gdw* $F(a) \subseteq F(c)$.

Literatur

Wir stellen ein knappes kommentiertes Literaturverzeichnis zu den Themen dieses Buches zusammen. Neben neueren ein- und weiterführenden Texten findet der Leser hier auch einige klassische Lehrbücher. Die Auswahl ist dabei persönlich gefärbt.

1. Abschnitt: Die Sprache der Mathematik

Viele Einführungen in die Mathematik enthalten einen Abschnitt über Wahrheitstafeln, Junktoren und Quantoren, und das gleiche gilt für Mengen, Funktionen und Relationen. Eine genauere Behandlung der mathematischen Sprache und des mathematischen Beweisens fällt dann in das Gebiet der mathematischen Logik. Ein lesenswerter Klassiker hierzu ist [Tarski 1971]. Das Buch [Halmos 1960] ist eine knappe und elementare Einführung in die Welt der Mengen. In [Deiser 2009] findet der Leser eine weitergehende Darstellung, die auch die historische Entwicklung einbezieht. Der Klassiker unter den Lehrbüchern zur Mengenlehre ist sicher [Hausdorff 1914].

2. Abschnitt: Zahlen

Der Sammelband [Ebbinghaus 1988] enthält Aufsätze unterschiedlichen Schwierigkeitsgrads zum Thema „Zahlen". Das Buch [Rautenberg 2008] bietet eine elementar gehaltene Einführung in das Zahlsystem. In [Deiser 2008] werden neben Konstruktionen und Charakterisierungen der reellen Zahlen auch topologische und maßtheoretische Themen behandelt. Der Klassiker unter den Darstellungen des Zahlsystems ist das Buch [Landau 1930], in dem die Konstruktionen wohl zum ersten Mal im Detail durchgeführt wurden. Zur Geschichte des Zahlbegriffs verweisen wir auf [Gericke 1980].

3. Abschnitt: Erste Erkundungen

Teiler: Viele elementare Ergebnisse der Teilbarkeitstheorie finden sich bereits in den ca. 300 vor Chr. verfaßten „Elementen" des Euklid. Gauß hat dann 1801 durch seine „Disquisitiones Arithmeticae" die moderne Zahlentheorie begründet. Ein erster Klassiker unter den Lehrbüchern ist dann [Hardy / Wright 1979]. Neuere Einführungen sind etwa [Bundschuh 2008] und [Scheid / Frommer 2006]. In [Forster 1996] werden algorithmische Aspekte betont.

Grenzwerte: Unter den vielen Lehrbüchern zur Analysis seien stellvertretend [Behrends 2008], [Forster 2008], [Rudin 2008] und [Walter 2007] genannt. Eine historisch geleitete Darstellung bietet [Hairer / Wanner 2008].

Matrizen: Mehr über Matrizen, lineare Gleichungssysteme und die allgemeine Theorie der Vektorräume findet der Leser etwa in den Lehrbüchern [Bosch 2008], [Fischer 2008] und [Koecher 2002]. Für numerische Aspekte verweisen wir auf [Deuflhard / Hohmann 2008] und [Trefethen / Bau 1997].

Gruppen: Die Grundbegriffe der Gruppentheorie kommen in vielen Lehrbüchern zur linearen Algebra zur Sprache. Eine einführende Einzeldarstellung ist [Alexandroff 2007]. Wer mehr über die Auflösbarkeit von Gleichungen erfahren möchte, findet in [Edwards 1993] eine Einführung in die Galois-Theorie, die sich an Originalarbeiten orientiert. Weiter verweisen wir hierzu auf [Lang 1979].

Graphen: Die Theorie der Graphen wird in eigenen Lehrbüchern behandelt, und sie bildet zumeist auch ein Hauptstück in den Einführungen in die diskrete Mathematik. Wir nennen hier [Aigner 2006] und [Diestel 2006]. Der Klassiker zur Graphentheorie ist [König 1936].

Wahrscheinlichkeiten: Einführungen in die mathematische Modellierung des Zufalls und die Grundbegriffe der Statistik findet der Leser in [Georgii 2009] und [Krengel 2005]. Weiter führt [Dudley 2002].

Literaturverzeichnis

Aigner, Martin 2006 Diskrete Mathematik; *6. Auflage. Vieweg+Teubner, Braunschweig.*

Alexandroff, Pavel 2007 Einführung in die Gruppentheorie; *11. Auflage. Harri Deutsch, Frankfurt.*

Behrends, Ehrhard 2008 Analysis I; *4. Auflage. Vieweg+Teubner, Braunschweig.*

Bosch, Siegfried 2008 Lineare Algebra; *3. Auflage. Springer, Berlin.*

Bundschuh, Peter 2008 Einführung in die Zahlentheorie; *6. Auflage. Springer, Berlin.*

Deiser, Oliver 2008 Reelle Zahlen; *2. Auflage. Springer, Berlin.*

– 2009 Einführung in die Mengenlehre; *3. Auflage. Springer, Berlin.*

Deuflhard, Peter / Hohmann, Andreas 2008 Numerische Mathematik 1; *4. Auflage. De Gruyter, Berlin.*

Diestel, Reinhard 2006 Graphentheorie; *3. Auflage. Springer, Berlin.*

Dudley, R. M. 2002 Real Analysis and Probability; *2. Auflage. Cambridge University Press, Cambridge.*

Ebbinghaus, Heinz-Dieter et al. 1992 Zahlen; *3. Auflage. Springer, Berlin.*

Edwards, Harold 1993 Galois Theory; *2. Auflage. Springer, Berlin.*

Euklid um 300 v. Chr. Elemente; *Deutsche Übersetzung: „Die Elemente. Bücher I – XIII", 4. Auflage 2003, Harri Deutsch, Frankfurt.*

Gauß, Carl Friedrich 1801 Disquisitiones Arithmeticae; *Deutsche Übersetzung: „Untersuchungen über höhere Arithmetik", Springer, Berlin 1889.*

Georgii, Hans-Otto 2009 Stochastik; *4. Auflage. Walter de Gruyter, Berlin.*

Fischer, Gerd 2008 Lineare Algebra; *16. Auflage. Vieweg+Teubner, Braunschweig.*

Forster, Otto 1996 Algorithmische Zahlentheorie; *Vieweg, Braunschweig.*

– 2008 Analysis I; *9. Auflage. Vieweg + Teubner, Braunschweig.*

Gericke, Helmuth 1970 Geschichte des Zahlbegriffs; *Bibliographisches Institut, Mannheim.*

Halmos, Paul Richard 1960 Naive Set Theory; *Van Nostrand, Princeton.* Deutsche Übersetzung: „Naive Mengenlehre", 4. Auflage 1976, Vandenhoeck & Ruprecht, Göttingen.

Hardy, G. H. / Wright, E. M. 1979 An Introduction to the Theory of Numbers; *5. Auflage. Clarendon Press, Oxford.*

Hausdorff, Felix 1914 Grundzüge der Mengenlehre; *Veit & Comp., Leipzig.* Kommentierter und mit Essays versehener Nachdruck 2002 bei Springer, Berlin (Band II der Hausdorff-Werkausgabe).

Hairer, Ernst / Wanner, Gerhard 2008 Analysis by Its History; *2. Auflage. Springer, Berlin.*

Koecher, Max 2002 Lineare Algebra und analytische Geometrie; *4. Auflage. Springer, Berlin.*

Krengel, Ulrich 2005 Einführung in die Wahrscheinlichkeitstheorie und Statistik; *8. Auflage. Vieweg + Teubner, Braunschweig.*

Landau, Edmund 1930 Grundlagen der Analysis; *Leipzig.*

Lang, Serge 1979 Algebraische Strukturen; *Vandenhoeck & Ruprecht, Göttingen.*

Rautenberg, Wolfgang 2007 Messen und Zählen; *Heldermann, Berlin.*

Rudin, Walter 2008 Analysis; *4. Auflage. Oldenbourg Verlag, München.*

Scheid, Harald / Frommer, Andreas 2006 Zahlentheorie; *4. Auflage. Spektrum, Heidelberg.*

Tarski, Alfred 1971 Einführung in die mathematische Logik; *5. Auflage. Vandenhoeck & Ruprecht, Göttingen.*

Trefethen, Lloyd / Bau, David 1997 Numerical Linear Algebra; *Society for Industrial and Applied Mathematics (SIAM), Philadelphia.*

Walter, Wolfgang 2007 Analysis 1; 7. *Auflage. Springer, Berlin.*

Notationen

Mathematisches Argumentieren

$\neg$, 20, 22
$\wedge$, 20
$\vee$, 20, 22
$\rightarrow$, 20, 22
$\leftrightarrow$, 20, 22
$A_1 \leftrightarrow A_2 \leftrightarrow A_3$, 29
$A_1 \rightarrow A_2 \rightarrow A_3$, 29
$\forall$, 31
$\exists$, 31

Mengen

$\in$, 41
$A \subseteq B$, $A \subset B$, 43
$A \supseteq B$, $A \supset B$, 43
$\{ x \mid \mathscr{E}(x) \}$, 43
$\{ x \in M \mid \mathscr{E}(x) \}$, 44
$\varnothing$, 45
$\{ a \}$, 45
$\{ a, b \}$, 45
$\{ a_1, \ldots, a_n \}$, 45
(a, b), 46
(a, b, c), 46
$A \cap B$, 46
$A \cup B$, 46
$A - B$, 46
A^c, 46
$A \Delta B$, 46
$A \times B$, 48
$\mathcal{P}(A)$, 48
$\bigcap \mathscr{A}$, $\bigcup \mathscr{A}$, 49
$\forall x \in X$, $\exists x \in X$, 49
$\forall x, y \in X$, 49

Relationen und Funktionen

$a \mathrel{R} b$, 55
$R(a, b)$, 55
$a \mathrel{R} b \mathrel{R} c$, 55
$\mathrm{dom}(R)$, 55
$\mathrm{rng}(R)$, 55
$\mathrm{field}(R)$, 56
R^{-1}, 56
$R(a_1, \ldots, a_n)$, 57
$a/\sim$, 58
$A/\sim$, 58
$f(a) = b$, 64
$f(a_1, \ldots, a_n)$, 64
$f \restriction A$, 64
id_A, 64
const_c^A, 64
$f : A \rightarrow B$, 64
${}^A B$, 65
f^{-1}, 66
$g \circ f$, 67
$f[X]$, 68
$f^{-1}[Y]$, 68
$\langle x_i \mid i \in I \rangle$, 68
$\bigcap_{i \in I}$, 68
$\bigcup_{i \in I}$, 68
$\bigtimes \langle B_i \mid i \in I \rangle$, 69
$\bigtimes_{i \in I}$, 69
$R \circ S$, 72

Mächtigkeiten

$|M| \leq |N|$, 77
$|M| = |N|$, 77
$|M| < |N|$, 77
$\aleph_0$, $\aleph_1$, 81

Zahlen

S, 87
$\mathbb{N}$, 87, 91
0, 1, 2, 3, 91
S(n), 91
$h(g_1, \ldots, g_k)$, 96
rec(g, h), 96
ind_A, 96
μg, 97
$\mathbb{Z}$, 104
$|a|$, 106
$\mathbb{Q}$, 107
a/b, 107
$|x|$, 111
K^+, 111
sup(X), 116
inf(X), 116
$\mathbb{R}$, 118
$\mathbb{C}$, 123
i, 123
$|(x, y)|$, 124
Re(z), 125
Im(z), 125
$\bar{z}$, 126
1, i, j, k, 126
$\mathbb{H}$, 127

Teiler

$d \mid a$, 135
ggT(a, b), 137
kgV(a, b), 138
$\sim$, 146
en(n), 147

Grenzwerte

sup(X), 157
inf(X), 157
lim, 158
liminf, 159
limsup, 159
$]x, y[$, 160
$[x, y]$, 160
$U_\varepsilon(x)$, 160
$\Sigma_{n \in \mathbb{N}} x_n$, 162
$\lim_{n \in \mathbb{N}} x_n = \infty$, 163

Matrizen

f_A, 179
$E_i' = \alpha \cdot E_i$, 182
$E_i' = E_i + \alpha E_j$, 182
$\sigma_{i,j}$, 182
$\pi_{i,j}$, 182
D_r, 182
$S_1 \circ S_2$, 188
A_R, 189
$R^{(k)}$, 190

Gruppen

0, 201
e, 201
1, 201
a^{-1}, 201
– a, 201
a^n, 201
$\Pi_{1 \le i \le n} a_i$, 202
$\Sigma_{1 \le i \le n} a_i$, 202
$\mathrm{ord}_G(a)$, 204
$x \sim_H y$, 205
$x \sim^H y$, 205
G/H, 206
(H : G), 207
sgn(f), 208

Graphen

a b, 213
K_n, 214
$K_{n,m}$, 214
C_n, 214
N(a), 215
d(a), 215
d(a, b), 217
G^c, 226

Wahrscheinlichkeiten

$|A|$, 231
n!, 231
$\binom{n}{k}$, 231
$\binom{n}{k_1, \ldots, k_r}$, 231
$\delta_{a,A}$, 234
$\Sigma_{n \in \mathbb{N}} \nu(n)\, \mu_n$, 235
$\mu_1 \times \mu_2$, 235
$T(\mu)$, 236
$\mu \circ T^{-1}$, 236
$\int f \, d\mu$, 239
ind_B, 239
E(X), 242
V(X), 242
$\sigma(X)$, 242
Cov(X, Y), 242
$\rho(X, Y)$, 242

Index