Studienbücherei

G. Asser
Grundbegriffe der Mathematik

I. Mengen. Abbildungen. Natürliche Zahlen

Mathematik für Lehrer
Band 1

Herausgegeben von:
W. Engel, S. Brehmer, M. Schneider, H. Wussing

Unter Mitarbeit von:
G. Asser, J. Böhm, J. Flachsmeyer, G. Geise, T. Glocke, K. Härtig, G. Kasdorf, O. Krötenheerdt, H. Lugowski, P.-H. Müller, G. Porath

Studienbücherei

Grundbegriffe der Mathematik

G. Asser

I. Mengen. Abbildungen.
Natürliche Zahlen

Mit 9 Abbildungen

Zweite, berichtigte Auflage

VEB Deutscher Verlag
der Wissenschaften
Berlin 1975

Verlagslektor: Dipl.-Math. E. Arndt
Umschlaggestaltung: R. Wendt

Printed in the German Democratic Republic
Lizenz-Nr. 206 · 435/204/75
Satz: IV/2/14 VEB Druckerei »Gottfried Wilhelm Leibniz«,
445 Gräfenhainichen/DDR · 3893
Offsetdruck: Mühlhäuser Druckhaus
LSV 1014
Bestellnummer 570 028 9
EVP 9,80 Mark

Vorwort der Herausgeber

Auf der Grundlage allgemeiner Konzeptionen für die weitere Entwicklung der Ausbildung von Fachlehrern an den allgemeinbildenden polytechnischen Oberschulen der Deutschen Demokratischen Republik erarbeiteten die Mitglieder der Fachkommission Mathematik beim Ministerium für Volksbildung und beim Ministerium für Hoch- und Fachschulwesen ein Studienprogramm. In ihm wurden die Grundsätze für die Ausbildung im Fach, die Studienpläne, Prüfungen und der Inhalt der Lehrveranstaltungen für die nächsten Jahre festgelegt. Dieses 1969 bestätigte Programm gilt für die Ausbildung von Fachlehrern mit dem Haupt- oder Nebenfach Mathematik an Universitäten, Pädagogischen Hochschulen, Technischen Hochschulen und Pädagogischen Instituten der DDR. Die Fachlehrer werden im Unterricht der Klassen 5 bis 10 der allgemeinbildenden polytechnischen Oberschulen eingesetzt. Aber auch die Lehrer für die Klassen 11 und 12 der Erweiterten Oberschule werden aus den gemäß dem Studienprogramm ausgebildeten Fachlehrern ausgewählt.

Mit der Erarbeitung des Studienprogramms entstand der Wunsch und in gewisser Hinsicht auch die Notwendigkeit, ein eigenständiges Lehrwerk zu schaffen, da die bisher verwendete Literatur die speziellen Belange der Lehrerausbildung naturgemäß nur bedingt oder gar nicht berücksichtigt. Um so erfreulicher ist die bereitwillige Mitarbeit vieler Hochschullehrer als Autoren, Gutachter und Mitherausgeber bei der Aufgabe, ein solches Werk herauszubringen, das sowohl zum Gebrauch neben den Vorlesungen als auch im Fernstudium verwendet werden kann und darüber hinaus auch für die Lehrer geeignet ist, welche bereits im Schuldienst stehen.

Alle Beteiligten sind bemüht, im Rahmen der Studienbücherei des VEB Deutscher Verlag der Wissenschaften ein Werk zu schaffen, das in Anlehnung an das Studienprogramm im Stoffumfang beschränkt ist, aber den modernen Entwicklungstendenzen in der Mathematik Rechnung trägt und die Bedürfnisse bei deren Umsetzung in Ausbildung und Erziehung an den allgemeinbildenden Schulen der DDR berücksichtigt. Im Interesse der Benutzer erfolgte

eine weitgehende Abstimmung für die einzelnen Bände u. a. in bezug auf die verwendeten Begriffe, Termini und Symbole, ohne daß dadurch der individuelle wissenschaftliche Stil der einzelnen Autoren beeinflußt wurde. Darüber hinaus werden sich in den einzelnen Bänden neben Beispielen auch Übungsaufgaben und kurze historische Einflechtungen finden, die den Leser in die Lage versetzen sollen, die engen Wechselwirkungen zwischen allgemeingesellschaftlicher und speziell-wissenschaftlicher Entwicklung zu erkennen und daraus Einsichten in die innermathematische Vorwärtsbewegung zu gewinnen.

Herausgeber und Verlag sind nunmehr in der Lage, den ersten Band der Reihe „Mathematik für Lehrer" vorzustellen. Er wurde von Herrn Prof. Dr. G. Asser verfaßt und führt in den Stoff der ersten beiden Studienjahre ein. Bis 1975 werden die weiteren Bände folgen, die den hauptsächlichsten Stoff der ersten beiden Studienjahre umfassen.

Bd. 2 Doz. Dr. J. Wisliceny
Grundbegriffe der Mathematik
Teil II: Rationale Zahlen, reelle Zahlen und komplexe Zahlen

Bd. 3 Prof. Dr. J. Flachsmeyer, Doz. Dr. L. Prohaska
Algebra

Bd. 4 Prof. Dr. S. Brehmer, Prof. Dr. H. Apelt
Analysis
Teil I: Folgen, Reihen, Funktionen

Bd. 5 Prof. Dr. S. Brehmer, Prof. Dr. H. Apelt
Analysis
Teil II: Differential- und Integralrechnung

Bd. 6 Prof. Dr. J. Böhm, Dr. W. Börner, Dr. E. Hertel,
Prof. Dr. O. Krötenheerdt, Prof. Dr. W. Mögling,
Doz. Dr. L. Stammler
Geometrie
Teil I: Axiomatischer Aufbau der euklidischen Geometrie

Bd. 7 Prof. Dr. J. Böhm, Dr. W. Börner, Dr. E. Hertel,
Prof. Dr. O. Krötenheerdt, Prof. Dr. W. Mögling,
Doz. Dr. L. Stammler
Geometrie
Teil II: Analytische Darstellung der euklidischen Geometrie, Abbildungen als Ordnungsprinzip in der Geometrie, geometrische Konstruktionen

Bd. 8 Doz. Dr. E. SCHRÖDER
Darstellende Geometrie

Bd. 9 Doz. Dr. H. KAISER
Numerische Mathematik und Rechentechnik
Teil I

Weitere Bände, die zur Verwendung im dritten und vierten Studienjahr geeignet sind, werden vorbereitet.

Herausgeber und Autoren hoffen, daß diese Reihe einen Beitrag zur Verbesserung der Aus- und Weiterbildung von Mathematiklehrern liefert. Sie werden kritische Hinweise, die zur Verbesserung der Bücher führen, gern entgegennehmen.

Den Herausgebern ist es eine angenehme Pflicht, beiden genannten Ministerien für die den Autoren und Herausgebern gewährte Unterstützung zu danken. Weiter gilt unser Dank der Leitung des VEB Deutscher Verlag der Wissenschaften und besonders dem Lektorat Mathematik mit Herrn W. ARNOLD und Fräulein E. ARNDT für die gute Zusammenarbeit während der Vorbereitungszeit der Reihe „Mathematik für Lehrer".

W. ENGEL S. BREHMER M. SCHNEIDER H. WUSSING

Vorwort des Autors

Der vorliegende erste Band der Studienbücherei „Mathematik für Lehrer" enthält eine Einführung in die allgemeine Mengenlehre und Abbildungstheorie sowie die Anfangsgründe der Arithmetik der natürlichen Zahlen. Die Einführung in die Mengenlehre erfolgt auf der Grundlage einer typentheoretischen Auffassung der Mengen. Diese ist für die meisten Anwendungen der Mengenlehre in der Mathematik voll ausreichend und hat den Vorteil leichter Verständlichkeit. Da die Abbildungen, wie es heute in der Mathematik weitgehend üblich ist, als Mengen von geordneten Paaren aufgefaßt werden, erscheint die Abbildungstheorie als ein etwas weiter ausgeführtes Kapitel der allgemeinen Mengenlehre. Hier finden auch einige Ausführungen über Relationen und Operationen ihren Platz. In einem Abschnitt über endliche Mengen wird gezeigt, wie allgemeine Aussagen über endliche Mengen, die wohl jedem mehr oder minder einleuchten werden, auf der Grundlage einer exakten Endlichkeitsdefinition streng bewiesen werden können. Wesentlich ist dabei, daß diese Beweise nicht den Begriff der natürlichen Zahl benutzen. In dem Kapitel über natürliche Zahlen besteht das Hauptanliegen darin, die wichtigsten Eigenschaften der natürlichen Zahlen systematisch und lückenlos aus wenigen ihrer Grundeigenschaften, dem Peanoschen Axiomensystem für die natürlichen Zahlen, herzuleiten.

Die Definitionen und Sätze werden in diesem Band in halbformalisierter Form aufgeschrieben, die Beweise jedoch grundsätzlich in der Umgangssprache inhaltlich geführt. Wir möchten unsere Meinung hierzu kurz darlegen. Die Formalisierungstechnik ist im Zusammenhang mit der mathematischen Grundlagenforschung entstanden und dort zu größter Perfektion entwickelt worden, wobei neben einer Formalisierung der Aussagen auch eine Formalisierung der Beweise eine grundsätzliche Rolle spielt. In den Grundlagen der Mathematik ist diese Formalisierung unumgänglich notwendig, da durch sie

die meisten Grundlagenprobleme erst einen über allgemein-philosophische Erörterungen hinausgehenden bestimmten Sinn erhalten. Seit einigen Jahren setzt es sich nun immer mehr durch, daß gewisse Elemente der Formalisierungstechnik, gewissermaßen als „logische Stenografie", in Vorlesungen und Publikationen zu verschiedensten Gebieten der Mathematik Eingang finden. Und selbst wenn ein Dozent in seinen Vorlesungen bewußt hierauf verzichtet, wird er nicht selten feststellen, daß es seine Hörer – ob er es mag oder nicht – anders halten. Sofern die Abkürzungstechnik vernünftig und vor allen Dingen einwandfrei gehandhabt wird, kann man ihre Verwendung nicht schlechthin verwerfen; sie hat sogar den Vorteil, daß die logische Struktur der Definitionen und Sätze klar ersichtlich wird. Leider wird jedoch in Vorlesungen und Veröffentlichungen, insbesondere aber bei der „Privatstenografie" der Studenten noch oft gegen die genannten Forderungen verstoßen. Ich möchte mit der Art der Darstellung dem Leser zeigen, wie eine einwandfreie und – wie ich meine – vernünftige Verwendung logischer Zeichen bei der Darlegung mathematischer Sachverhalte etwa aussehen kann. Natürlich muß man die einwandfreie Handhabung der Abkürzungstechnik üben, und hierbei möchte der vorliegende Band ebenfalls unterstützen. Drei grundsätzliche Bemerkungen seien dem Leser noch mit auf den Weg gegeben: (1) Man hüte sich vor der Annahme, daß die Verwendung logischer Abkürzungen ein Zeichen besonderer mathematischer Bildung sei. (2) Man glaube nicht, daß die Verwendung logischer Abkürzungen automatisch logische Exaktheit zur Folge hat oder gar Voraussetzung hierfür sei; sie hilft jedoch, eventuelle logische oder begriffliche Unklarheiten aufzudecken und zu beseitigen. (3) Man vergesse über der Form nicht den Inhalt (und das betrifft keineswegs in erster Linie das Problem der Formalisierung). Vor allem bemühe man sich stets, und das ist für den angehenden Lehrer besonders wichtig, auch komplizierte mathematische Zusammenhänge sprachlich einwandfrei zu formulieren.

Die in diesem Band erklärten Begriffe und formulierten Sätze werden dem Leser zum größten Teil aus der Schule bekannt sein, allerdings vorwiegend als mehr oder minder empirisch gewonnene Einzelfakten. Demgegenüber werden sie hier in einen systematischen Zusammenhang gebracht und exakt begründet. Diese Seite der Mathematik bereitet dem Anfänger beim Mathematikstudium erfahrungsgemäß erhebliche Schwierigkeiten und ist die berüchtigte Barriere beim Übergang von der Schule zur Hochschule. Möge dieser Band vielen Studenten, die mit großen Erwartungen ein Mathematikstudium aufnehmen, um einmal als Lehrer ihr Wissen und Können an die Schuljugend weiterzugeben, die ersten Schritte auf diesem Wege erleichtern.

Ich möchte es nicht versäumen, den Herausgebern dieser Reihe und insbesondere ihrem Initiator, Herrn Prof. Dr. W. Engel, sowie vielen in der

Lehrerausbildung tätigen Fachkollegen, die das Manuskript lasen und mir wertvolle Hinweise gaben, herzlich zu danken. Mein Dank gilt ferner dem VEB Deutscher Verlag der Wissenschaften für die Herausgabe dieses Bandes und den Setzern des VEB Druckerei „G. W. Leibniz" für ihre sorgfältige Arbeit bei der Drucklegung.

Greifswald, im Februar 1973 GÜNTER ASSER

Inhalt

Überblick über die wichtigsten im vorliegenden Band eingeführten Zeichen

Das Zeichen „$:=$“ (gelesen: ... ist definitionsgemäß gleich ...) wird als Definitionszeichen für Terme verwendet; der links vom Zeichen stehende Term ist eine neu eingeführte (meistens kürzere) Bezeichnung für den rechts vom Zeichen stehenden Term.

Das Zeichen „$:\Leftrightarrow$“ (gelesen: ... gilt definitionsgemäß genau dann, wenn ...) wird analog als Definitionszeichen für Eigenschaften und Beziehungen benutzt.

1. Logische Zeichen

$p \wedge q :\Leftrightarrow p$ und q	(Konjunktion)
$p \vee q :\Leftrightarrow p$ oder q	(Alternative)
$p \Rightarrow q :\Leftrightarrow$ wenn p, so q	(Implikation)
$p \Leftrightarrow q :\Leftrightarrow p$ genau dann, wenn q	(Äquivalenz)
$\neg p :\Leftrightarrow$ nicht p	(Negation)
$\bigwedge_x H(x) :\Leftrightarrow$ für jedes x gilt $H(x)$	(Generalisierung)
$\bigwedge_{x \in M} H(x) :\Leftrightarrow \bigwedge_x (x \in M \Rightarrow H(x))$	
$\bigvee_x H(x) :\Leftrightarrow$ es gibt ein x mit $H(x)$	(Partikularisierung)
$\bigvee_{x \in M} H(x) :\Leftrightarrow \bigvee_x (x \in M \wedge H(x))$	

2. Mengentheoretische Zeichen

$x \in M :\Leftrightarrow x$ ist Element von M	
$x \notin M :\Leftrightarrow \neg x \in M$	
$M \subseteq N :\Leftrightarrow \bigwedge_x (x \in M \Rightarrow x \in N)$	(Inklusion)

$M \subset N :\Leftrightarrow M \subseteqq N \wedge M \neq N$ (echte Inklusion)
$\{x : H(x)\} :=$ Menge aller x mit der Eigenschaft $H(x)$
$M \cap N := \{x : x \in M \wedge x \in N\}$ (Durchschnitt)
$M \cup N := \{x : x \in M \vee x \in N\}$ (Vereinigung)
$M \backslash N := \{x : x \in M \wedge x \notin N\}$ (Differenzmenge)
$\bigcap \mathfrak{M} := \{x : \bigwedge_M (M \in \mathfrak{M} \Rightarrow x \in M)\}$
$\bigcup \mathfrak{M} := \{x : \bigvee_M (M \in \mathfrak{M} \wedge x \in M)\}$
$\mathfrak{P}(M) := \{X : X \subseteqq M\}$ (Potenzmenge)
$\emptyset :=$ leere Menge
$\{a\} := \{x : x = a\}$ (Einermenge)
$\{a, b\} := \{x : x = a \vee x = b\}$ (Zweiermenge)
$\mathsf{N} :=$ Menge aller natürlichen Zahlen
$\mathsf{N}^* :=$ Menge aller von Null verschiedenen natürlichen Zahlen
$\mathsf{Z} :=$ Menge aller ganzen Zahlen
$\mathsf{Q} :=$ Menge aller rationalen Zahlen
$\mathsf{R} :=$ Menge aller reellen Zahlen
$\mathsf{R}_+ := \{x : x \in \mathsf{R} \wedge x \geqq 0\}$
$\mathsf{R}_- := \{x : x \in \mathsf{R} \wedge x \leqq 0\}$
$\mathsf{R}^* := \{x : x \in \mathsf{R} \wedge x \neq 0\}$
$\mathsf{R}_+^* := \{x : x \in \mathsf{R} \wedge x > 0\}$ $(= \mathsf{R}^* \cap \mathsf{R}_+)$
$\mathsf{R}_-^* := \{x : x \in \mathsf{R} \wedge x < 0\}$ $(= \mathsf{R}^* \cap \mathsf{R}_-)$

3. Abbildungstheoretische Zeichen

$(x, y) :=$ geordnetes Paar aus x und y
$M \times N := \{(x, y) : x \in M \wedge y \in N\}$ (Produktmenge)
$B_F(x) := \{y : (x, y) \in F\}$ (volles Bild von x bei F)
$U_F(y) := \{x : (x, y) \in F\}$ (volles Urbild von y bei F)
$D(F) := \{x : B_F(x) \neq \emptyset\}$ (Definitionsbereich von F)
$W(F) := \{y : U_F(y) \neq \emptyset\}$ (Wertebereich von F)
$F^{-1} := \{(y, x) : (x, y) \in F\}$ (Umkehrkorrespondenz)
$G \circ F := \{(x, z) : \bigvee_y ((x, y) \in F \wedge (y, z) \in G)\}$ (Verkettung)
$F \mid X := \{(x, y) : x \in X \wedge (x, y) \in F\}$ (Einschränkung)
$F : M \to N :\Leftrightarrow F$ eindeutige Abbildung von M in N
$R/M :=$ Restsystem von R nach M
$F : \Sigma \leftrightarrows \Sigma' :\Leftrightarrow F$ Isomorphismus von Σ auf Σ'
$\Sigma \cong \Sigma' :\Leftrightarrow \bigvee_F (F : \Sigma \leftrightarrows \Sigma')$
$F : \Sigma \rightarrowtail \Sigma' :\Leftrightarrow F$ Homomorphismus von Σ in Σ'

1. Grundbegriffe der Mengenlehre

1.1. Einleitung

In allen Gebieten der Mathematik spielen heute der Mengen- und der Abbildungsbegriff sowie eine Reihe hiermit zusammenhängender allgemeiner Begriffsbildungen eine beherrschende Rolle. Ihr systematisches Studium bildet den Gegenstand einer eigenen mathematischen Disziplin, der sogenannten (allgemeinen) Mengenlehre oder Mengentheorie. Man kann behaupten, daß die mengentheoretische Betrachtungsweise einen bedeutsamen Einfluß auf die Entwicklung der Mathematik unseres Jahrhunderts ausgeübt hat und das Entstehen vieler wichtiger Teilgebiete der heutigen Mathematik ohne das Fundament der allgemeinen Mengenlehre nicht möglich gewesen wäre. Dabei hat die Mengenlehre wesentlich zur Präzisierung, Vereinfachung und Vereinheitlichung des Begriffssystems der Mathematik beigetragen. Daher wird auch der Mathematikunterricht an den Schulen in immer stärkerem Maße von mengentheoretischen Auffassungen durchdrungen. Es ist folglich nur natürlich, daß die Vermittlung von Grundkenntnissen der Mengenlehre einen entsprechenden Platz in der Mathematikausbildung der Lehrerstudenten an den Universitäten und Pädagogischen Hochschulen unserer Republik einnimmt.

Als Begründer der Mengenlehre ist der Hallenser Mathematikprofessor GEORG CANTOR (1845–1918) anzusehen. Natürlich hatte man schon lange vor dem Erscheinen der grundlegenden Arbeiten CANTORS zur Mengenlehre Gesamtheiten von mathematischen Objekten, wie Zahlen, Punkten usw., betrachtet, wenn man es vielleicht auch anders ausdrückte. Wesentlich neu bei CANTOR war, daß er die Mengen und gewisse Beziehungen zwischen ihnen zu selbständigen Gegenständen seiner mathematischen Untersuchungen machte. Dabei wurde CANTOR durch sehr konkrete mathematische Fragestellungen auf diese Untersuchungen geführt. Man kann sich heute kaum

noch vorstellen, welche große Energie CANTOR aufwenden mußte, um sich gegen zahlreiche Vorurteile seiner mathematischen Zeitgenossen durchzusetzen. Seine Arbeiten wurden von vielen seiner damaligen Kollegen als unklar und falsch abgelehnt, insbesondere war der einflußreiche und bedeutende Berliner Mathematiker LEOPOLD KRONECKER (1823–1891) einer seiner erbittertsten Widersacher. Man berief sich dabei unter anderem auf CARL FRIEDRICH GAUSS (1777–1855), der wohl größten mathematischen Autorität des vorigen Jahrhunderts, der in Anspielung auf den Grenzwertbegriff der Analysis einmal folgendes geäußert hatte: „So protestiere ich zuvörderst gegen den Gebrauch einer unendlichen Größe als einer Vollendeten, welches in der Mathematik niemals erlaubt ist. Das Unendliche ist nur eine façon de parler, indem man eigentlich von Grenzen spricht, denen gewisse Verhältnisse so nahe kommen als man will, während anderen ohne Einschränkung zu wachsen gestattet ist." Demgegenüber sah CANTOR in der Tat auch unendliche Mengen als etwas durchaus Vollendetes, Fertiges an, was insbesondere der Ausgangspunkt seiner Lehre von den transfiniten Zahlen war. Man pflegt heute jene beiden unterschiedlichen Auffassungen vom Unendlichen, das Unendlichwerden und das Unendlichsein, durch die Bezeichnungen *potentiell unendlich* und *aktual unendlich* zu unterscheiden. Es wurde übrigens CANTOR nicht nur der Vorwurf gemacht, er habe gegen Gesetze der Logik und Mathematik verstoßen, es wurden ihm sogar Verstöße gegen Grundsätze der Religion nachgesagt.

Die ersten großen Erfolge der Mengenlehre stellten sich zu Beginn unseres Jahrhunderts in einem Grenzgebiet zwischen Analysis und Geometrie, der Theorie der Punktmengen ein. Auf ihrer Grundlage entstand eine ganz neue mathematische Disziplin, die man zunächst „Analysis des Unendlichen" oder „Analysis situs" nannte und die heute als allgemeine oder mengentheoretische Topologie bezeichnet wird. Ihre Ideen spielen gegenwärtig in vielen Zweigen der Mathematik eine beherrschende Rolle. Die Theorie der Punktmengen entwickelte sich ihrerseits in engem Zusammenhang mit der Theorie der reellen Funktionen, vor allem der Maß- und Integrationstheorie. Es ist bemerkenswert, daß im Anfangsstadium die Theorie der Punktmengen noch als Teilgebiet der allgemeinen Mengenlehre angesehen wurde. Es spiegelt sich das noch deutlich im Inhalt der Arbeiten CANTORS und der ersten Lehrbücher zur Mengenlehre wider, so in dem im Jahre 1914 in Greifswald entstandenen klassischen Lehrbuch „Grundzüge der Mengenlehre" von FELIX HAUSDORFF (1868–1942; HAUSDORFF schied, um der Einweisung in ein Konzentrationslager zu entgehen, am 29. 1. 1942 freiwillig aus dem Leben), das von großer Bedeutung für die Durchsetzung der Cantorschen Ideen war. Dabei übten die genannten Gebiete einen starken Einfluß

auf die Entwicklung der Mengenlehre aus, wie man überhaupt sagen kann, daß die Mengenlehre stets maßgeblich von ihren Anwendungen in den verschiedenen Bereichen der Mathematik befruchtet wurde und auch heute noch befruchtet wird.

Allerdings ergaben sich auch eine Reihe von Schwierigkeiten. Es zeigte sich einerseits, daß bei unvorsichtigem Vorgehen Widersprüche auftreten, und andererseits deckte man einige Fakten auf, die man sich zunächst nur schwer erklären konnte. Dadurch erhielten die Diskussionen um die Mengenlehre neue Nahrung, die sich in den zwanziger Jahren zu heftigen Auseinandersetzungen über die logisch-philosophischen Grundlagen der Mathematik ausweiteten. Zugleich lösten sie indes auch fundierte Untersuchungen über die Grundlagen der Mathematik und speziell der Mengenlehre aus, und es kam zur forcierten Entwicklung solcher Gebiete wie der Mathematischen Logik und der Grundlagen der Mathematik (die man heute verbreitet als Metamathematik bezeichnet); insbesondere entstanden in dieser Zeit die verschiedenen axiomatischen Systeme der Mengenlehre.

Heute wird die Mengenlehre von der überwiegenden Mehrheit der Mathematiker als Fundament der Mathematik voll anerkannt. Dabei werden – ähnlich wie zu CANTORS Zeiten – vielfach bereits weit über den Rahmen der klassischen Mengenlehre hinausgehende mengentheoretische Bildungen benutzt, deren volle logische Rechtfertigung zum Teil noch aussteht. Es gibt allerdings auch heute noch einige wissenschaftliche Schulen in der Mathematik (ihre Anhänger nennen sich Intuitionisten, Konstruktivisten u. ä.), die ernste Bedenken gegen gewisse Grundprinzipien der Mengenlehre und auf ihnen beruhenden mathematischen Überlegungen haben, wobei sich diese Bedenken zugleich gegen gewisse von der Mehrheit der Mathematiker anerkannte logische Schlußweisen richten (es handelt sich hierbei insbesondere um gewisse Arten „indirekter Schlüsse"). Die Ergebnisse dieser Schulen sind von großer methodischer Bedeutung, die Verabsolutierung der Auffassungen dieser Schulen führt indes zu einer von der „klassischen Mathematik" erheblich abweichenden Mathematik. Auf die hiermit zusammenhängenden Probleme können wir an dieser Stelle nicht eingehen.

1.2. Das Mengenbildungsprinzip

In dem im Jahre 1895 erschienenen ersten Teil seiner Arbeit „Beiträge zur Begründung der transfiniten Mengenlehre" gibt CANTOR die folgende berühmte Definition einer Menge: *„Eine Menge ist eine Zusammenfassung bestimmter*

wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die Elemente der Menge genannt werden – zu einem Ganzen." Wir merken sofort an, daß es sich bei dieser Formulierung nicht um eine Definition in dem heute in der Mathematik einzig üblichen Sinne einer sogenannten expliziten Definition handeln kann, bei der verlangt wird, daß man den definierten Begriff (in unserem Fall „Menge") stets eliminieren kann, indem man ihn durch den definierenden Sachverhalt (in unserem Fall „Zusammenfassung bestimmter ... zu einem Ganzen") ersetzt. Denn im vorliegenden Fall würden dabei an die Stelle des Begriffs Menge lediglich eine Reihe von anderen undefinierten Begriffen, wie Zusammenfassung, Objekt usw., gesetzt werden. Die Cantorsche Formulierung ist daher zunächst nur als eine ungefähre Beschreibung dessen anzusehen, was vorliegen muß, damit von einer Menge gesprochen wird.

Bevor wir den eigentlichen Inhalt der Cantorschen „Definition" genauer herausarbeiten, betrachten wir einige Beispiele für Mengen, wie sie häufig in der Mathematik auftreten:

(1) Die Menge aller natürlichen Zahlen.

(2) Die Menge aller Primzahlen.

(3) Die Menge aller natürlichen Zahlen n, für die die Gleichung $x^n + y^n = z^n$ durch von Null verschiedene ganze Zahlen x, y, z lösbar ist.

(4) Die Menge aller reellen Zahlen x, die den Ungleichungen $5 \leqq x < 7$ genügen.

(5) Die Menge aller Punkte einer gegebenen Ebene, die von einem festen Punkt O dieser Ebene den Abstand 1 haben.

(6) Die Menge aller Punkte X einer gegebenen Ebene, für deren Abstände $|A_1X|$, $|A_2X|$ von zwei festen Punkten A_1, A_2 der Ebene die Beziehung $|A_1X| + |A_2X| = 2$ gilt.

In allen diesen Fällen sind die Elemente der jeweiligen Menge durch eine *Eigenschaft* bzw. *Aussage* $H(x)$ charakterisiert, wobei ein Objekt x_0 dann und nur dann der betreffenden Menge angehört, wenn $H(x_0)$ gilt. In Beispiel (1) ist es die Aussage (Eigenschaft) „x ist eine natürliche Zahl", die auf die Zahlen $0, 1, 2, \ldots$, nicht aber z. B. auf die Zahlen -5, $\frac{1}{2}$, π usw. zutrifft. In Beispiel (4) ist es die Aussage „x ist eine reelle Zahl mit $5 \leqq x < 7$", die z. B. auf die Zahlen 5, $5\frac{1}{3}$, 6, $6\frac{9}{10}$, 2π usw., nicht aber z. B. auf die Zahlen 4, 7, $7\frac{1}{2}$, π usw. zutrifft. In Beispiel (6) ist es die Aussage „X ist ein Punkt der gegebenen Ebene mit $|A_1X| + |A_2X| = 2$"; ist hierbei $|A_1A_2| > 2$, so gibt es keinen Punkt X der Ebene, der der verlangten Gleichung genügt, die betrachtete Menge ist „leer"; ist $|A_1A_2| = 2$, so erfüllen genau die Punkte der Verbin-

dungsstrecke der Punkte A_1, A_2 die verlangte Gleichung; ist schließlich $|A_1A_2| < 2$, so ist die betrachtete Menge Peripherie einer Ellipse mit den Brennpunkten A_1, A_2 und dem großen Durchmesser 2, die im Fall $A_1 = A_2$ (d. h. $|A_1A_2| = 0$) in einen Kreis vom Durchmesser 2 ausartet.

Wir stellen weiterhin fest, daß in einigen der betrachteten Beispiele genau fixiert ist, innerhalb welches Grundbereichs die Elemente der betreffenden Menge zu wählen sind. So haben wir in (4) ausdrücklich alle reellen Zahlen x mit der Eigenschaft $5 \leqq x < 7$ und nicht etwa nur die natürlichen (in diesem Fall würde die zugehörige Menge nur die beiden Zahlen 5 und 6 enthalten) oder die rationalen Zahlen mit dieser Eigenschaft betrachtet. In (5) und (6) betrachteten wir nur die Punkte einer gegebenen Ebene und nicht etwa alle Punkte des Raumes mit der jeweiligen Eigenschaft. Aber auch in den übrigen Beispielen läßt sich aus dem Zusammenhang, wenn auch nicht immer in eindeutiger Weise, ein Grundbereich finden, innerhalb dessen man sich die jeweilige Mengenbildung vollzogen denken kann. So kommt in (2) als Grundbereich nur der Bereich der natürlichen Zahlen in Frage, da sich die in der Mathematik übliche Definition der Primzahlen (vgl. 3.7. (48)) grundsätzlich auf natürliche Zahlen bezieht. In (1) kann man als Grundbereich z. B. den Bereich der natürlichen Zahlen (in diesem Fall sind sämtliche Objekte des Grundbereichs Elemente der betrachteten Menge), den der ganzen Zahlen, den der rationalen Zahlen oder den der reellen Zahlen nehmen, sicher aber nicht die Gesamtheit aller Punkte einer Ebene oder die reellen Zahlen zwischen 0 und 1.

Nach diesen Vorbemerkungen dürfte der folgende allgemeine Ansatz hinreichend motiviert sein: Vorgegeben sei ein bestimmter *Grundbereich* E von Objekten (man nennt sie in der Mengenlehre heute häufig *Urelemente*) und eine Eigenschaft oder Aussage $H(x)$, die für die Objekte x aus E definiert ist, wobei also für ein beliebiges Objekt x_0 aus E sinnvoll die Frage gestellt werden kann, ob $H(x)$ auf x_0 zutrifft oder nicht. Zur Vermeidung von Mißverständnissen weisen wir darauf hin, daß wir nicht voraussetzen, daß man wirklich in der Lage sein muß, für jedes konkrete x_0 die richtige Antwort auf diese Frage zu geben. So kann man z. B. (vgl. das obige Beispiel (3)) für jede natürliche Zahl n sinnvoll die Frage stellen, ob die Gleichung $x^n + y^n = z^n$ durch von Null verschiedene ganze Zahlen x, y, z lösbar ist oder nicht, die konkrete Antwort auf diese Frage ist indes bis heute erst für wenige Zahlen n bekannt, es wird vermutet, daß sie für alle natürlichen Zahlen $n \geqq 3$ negativ ausfällt (diese Vermutung geht bereits auf den französischen Juristen und Liebhabermathematiker Pierre de Fermat (1601–1665) zurück, dem viele wichtige mathematische Entdeckungen zu verdanken sind; Fermat behauptete übrigens, einen Beweis für diese Vermutung zu besitzen). Über die Natur der

Objekte des Grundbereichs E werden keine Einschränkungen gemacht: Es kann sich dabei um ganz reale Objekte handeln, wie z. B. die von einem Betrieb erzeugten Fertigprodukte, die Mitarbeiter eines Betriebes usw. („Objekte unserer Anschauung" in der Terminologie CANTORS), oder auch Objekte begrifflicher Natur, wie Zahlen, Punkte usw. („Objekte unseres Denkens").

Die Cantorsche Mengendefinition wollen wir nun dahingehend interpretieren, daß man für eine gegebene Aussage oder Eigenschaft $H(x)$ alle diejenigen Objekte x, auf die $H(x)$ zutrifft, zu einer Menge zusammenfassen kann, genauer, daß *es eine Menge gibt, die genau jene x als Elemente hat.* Bezeichnen wir diese Menge zur Abkürzung mit M, so besteht also für ein beliebiges Objekt x die folgende Beziehung:

(7) *x ist Element von M genau dann, wenn $H(x)$.*

Für x [*ist*] *Element von* M, x *gehört* [*als Element*] *zu* M, x *ist* [*als Element*] *in* M *enthalten* – alles Synonyma für denselben Sachverhalt – schreibt man heute allgemein $x \in M$ ($\in$ ist dabei eine stilisierte Form des kleinen griechischen Buchstaben Epsilon). Verwenden wir zur Abkürzung von „genau dann, wenn", der sogenannten logischen Äquivalenz (Gleichwertigkeit), das Zeichen $\Leftrightarrow$, so können wir die Beziehung (7) zwischen der Menge M und der sie definierenden Aussage (Eigenschaft) $H(x)$ in der Form

(7′) $\quad x \in M \Leftrightarrow H(x)$

schreiben. Daß diese Beziehung für alle x gilt, deuten wir kurz durch

$$\bigwedge_x (x \in M \Leftrightarrow H(x))$$

an, wobei also $\bigwedge_x \ldots$ als „für alle x (gilt) . . ." oder „für jedes x (gilt) . . ." zu lesen ist (Generalisierung). Die Existenz einer derartigen Menge geben wir schließlich durch

(8) $$\bigvee_M \bigwedge_x (x \in M \Leftrightarrow H(x))$$

wieder, wobei also $\bigvee_M \ldots$ Abkürzung für „es gibt ein(e Menge) M mit . . ." oder „es existiert ein M mit . . ." verwendet wird (Partikularisierung). In Zweifelsfällen, falls es also nicht aus dem Zusammenhang klar hervorgeht, ist anzugeben, innerhalb welches Grundbereichs E die Betrachtungen verlaufen.

Die in der angegebenen Weise interpretierte Cantorsche Mengendefinition wollen wir Mengenbildungsprinzip nennen. Es handelt sich hierbei faktisch um ein Axiom oder Postulat: Es wird ohne Beweis postuliert, daß es bei gegebener Aussage $H(x)$ über die Objekte eines gegebenen Grundbereichs

E stets eine Menge M geben soll, für die (7) gilt. Das Mengenbildungsprinzip gibt damit eine ganz bestimmte inhaltliche Vorstellung wieder, die mit dem Mengenbegriff verknüpft werden soll. In der mengentheoretischen Literatur sind für (8) (von für unsere Betrachtungen unwesentlichen Unterschieden abgesehen) die folgenden anderen Bezeichnungen üblich: Mengenbildungsaxiom, Komprehensionsaxiom (Komprehension = Zusammenfassung), Aussonderungsaxiom.

Zur Klärung der Begriffe merken wir an, daß es sich bei den durch Zusammenfassung von Objekten entstehenden Mengen um ganz neue abstrakte Objekte handeln soll. So ist z. B. eine Menge von Zahlen begrifflich etwas ganz anderes als eine Zahl, auch wenn vielleicht diese Menge (vgl. 1.5. (20)) nur eine einzige Zahl als Element enthält. Es ist daher nicht sinnvoll zu fragen, ob eine gewisse Menge M von Urelementen Element einer anderen Menge N von Urelementen ist oder nicht: Die Beziehung $x \in M$ ist zunächst grundsätzlich nur dann definiert, wenn x ein Urelement und M eine Menge von Urelementen ist.

Es soll allerdings ausdrücklich zugelassen werden, daß die Gesamtheit aller überhaupt bildbaren Mengen von Urelementen aus einem gegebenen Grundbereich E als neuer Grundbereich $\mathfrak{E}$ von Urelementen für die Bildung sogenannter *Mengensysteme* oder *Mengen zweiter Stufe* genommen werden kann, deren Elemente dann *Mengen erster Stufe*, d. h. Mengen von Objekten aus E sind. Es ist vielfach üblich, derartige Mengensysteme durch große deutsche Buchstaben (Frakturbuchstaben) zu bezeichnen. Die Bildung von Mengensystemen erfolgt mittels eines zu (8) analogen Mengenbildungsprinzips:

$$(9) \qquad \bigvee_{\mathfrak{M}} \bigwedge_{X} (X \in \mathfrak{M} \Leftrightarrow H(X))$$

(gelesen: Es gibt ein Mengensystem $\mathfrak{M}$, so daß für jede Menge X gilt: X ist Element von $\mathfrak{M}$ genau dann, wenn $H(X)$), wobei jetzt $H(X)$ eine gegebene sinnvolle Aussage (Eigenschaft) über Mengen erster Stufe ist.

Einige Beispiele mögen das näher erläutern. Als Grundbereich E nehmen wir die Gesamtheit aller Punkte einer Ebene. Spezielle (Punkt-)Mengen erster Stufe sind dann die Geraden, Strecken, Kreise, Dreiecke usw. Wir bemerken, daß die Wörter „Kreis", „Dreieck" usw. in der Geometrie in sehr unterschiedlicher Bedeutung verwendet werden und man vielfach erst aus dem Zusammenhang entnehmen kann, was gemeint ist; so bezeichnet das Wort „Kreis" manchmal die Kreislinie (d. h. die Menge aller derjenigen Punkte, deren Abstand vom Mittelpunkt gleich dem Radius r ist), manchmal die Kreisscheibe unter Einschluß der Kreislinie (d. h. die Menge aller Punkte, deren Abstand vom Mittelpunkt $\leqq r$ ist) und manchmal die Kreisscheibe unter Ausschluß der Kreislinie (d. h. die Menge aller Punkte, deren Abstand vom Mittelpunkt $< r$ ist). Mit Hilfe dieser Mengen können wir dann z. B. die folgenden Mengensysteme bilden:

(10) Das System aller derjenigen Geraden, die durch einen gegebenen Punkt gehen.

(11) Das System aller derjenigen Geraden, die zu einer gegebenen Geraden parallel sind.

(12) Das System aller derjenigen Geraden, die Tangenten an einem gegebenen Kreis sind.

(13) Das System aller derjenigen Kreise, die eine gegebene Gerade als Tangente haben.

(14) Das System aller derjenigen Kreise, die zu einem gegebenen Kreis konzentrisch sind.

Im Bedarfsfall können anschließend in analoger Weise Mengen dritter Stufe gebildet werden, deren Elemente Mengen zweiter Stufe, d. h. Mengensysteme sind usw. Da die Mengen einer beliebigen Stufe k als Mengen erster Stufe aufgefaßt werden können, die als Elemente Mengen der Stufe $k-1$ besitzen, können auf jeder Stufe im wesentlichen dieselben Begriffe eingeführt werden und gelten jeweils auch dieselben Sätze. Wir beschränken uns daher im folgenden auf Untersuchung der Verhältnisse in der jeweils niedrigsten Stufe.

Die Beschränkung der Mengenbildungen auf Objekte eines bestimmten Grundbereichs und die dadurch bedingte begriffliche Unterscheidung zwischen Urelementen (oder Mengen nullter Stufe), Mengen erster Stufe, Mengen zweiter Stufe (oder Mengensystemen) usw. haben vor allem den Zweck, die bei sogenannten „uferlosen" Mengenbildungen auftretenden Antinomien (Widersprüche) zu vermeiden. Am bekanntesten und einfachsten zu formulieren ist die im Jahre 1901 von dem englischen Mathematiker, Logiker, Philosophen und Sozialkritiker BERTRAND RUSSELL (1872–1970) entdeckte Antinomie der Menge aller Mengen, die sich nicht selbst als Element enthalten. Verzichtet man bei den Mengen auf eine Stufenunterscheidung (wie das z. B. in der Cantorschen „Definition" der Fall ist), so kann diese Menge – zumindest versuchsweise – gebildet werden. Bezeichnen wir die Russellsche Menge mit m, so gilt also für eine beliebige Menge x

$$x \in m \Leftrightarrow x \notin x, \tag{15}$$

wobei das Zeichen $\notin$, wie allgemein üblich, die Negation der Elementbeziehung bezeichnet. Setzen wir in (15) für x speziell die Menge m ein, so erhalten wir:

$$m \in m \Leftrightarrow m \notin m,$$

und das ist ein Widerspruch (denn es kann nicht ein gewisser Sachverhalt genau dann vorliegen, wenn er nicht vorliegt). Zur Vermeidung dieser und ähnlicher Antinomien entwickelte RUSSELL die hier dargelegten Grundgedanken eines Stufenaufbaus der Mengenlehre, der nach ihm auch Typentheorie genannt wird, und baute diesen zusammen mit A. N. WHITEHEAD (1861–1947) in dem in den Jahren 1910 bis 1913 erschienenen dreibändigen Werk der „Principia Mathematica" zur logischen Grundlage für die Mathematik aus.

Wir möchten bemerken, daß die Gesamtheit aller Mengen, die sich nicht selbst als Element enthalten (wie z. B. auch die Gesamtheit aller Mengen), nicht von

vornherein von antinomischer Natur ist. Grund für das Auftreten der Russellschen Antinomie ist in erster Linie, daß diese Gesamtheit wieder als Menge angesehen wird. In neuerer Zeit setzt sich in der Mathematik immer mehr ein stufenfreier Aufbau der Mengenlehre durch, bei dem der Mengenbegriff dem umfassenderen Begriff der Klasse untergeordnet wird. Dabei sind diese Klassen – anschaulich gesprochen – beliebige Zusammenfassungen von Mengen, und als Mengen werden auch nur solche Klassen angesehen, die als Elemente von Klassen auftreten können. Die Gesamtheit aller Mengen, die sich nicht selbst als Element enthalten, wie auch die Gesamtheit aller Mengen und alle sonst bekannten Gesamtheiten von Mengen, die bei Auffassung als Mengen zu Antinomien führen, erweisen sich bei einem derartigen Aufbau als Klassen, die keine Mengen sind, als sogenannte „Unmengen". Auf nähere Einzelheiten über stufenfreie Begründungen der Mengenlehre können wir hier jedoch nicht eingehen.

1.3. Das Extensionalitätsprinzip

Wir kommen nun zur Formulierung eines weiteren Grundprinzips der Mengenlehre, von dem man ebenfalls ursprünglich annahm, daß es lediglich eine Definition sei, das aber genau genommen ebenfalls den Charakter eines Axioms oder Postulats hat. Es betrifft die Frage, wann Mengen M und N aus Elementen eines gegebenen Grundbereichs E als identisch angesehen werden sollen. Die Antwort auf diese Frage gibt das folgende

Extensionalitätsprinzip (Extension = Umfang, Ausdehnung). *Mengen M und N sind genau dann gleich (identisch), wenn sie dieselben Elemente enthalten,* d. h., *wenn für jedes x gilt: x ist Element von M genau dann, wenn x Element von N ist;* in Zeichen:

$$(1) \qquad M = N \Leftrightarrow \bigwedge_x (x \in M \Leftrightarrow x \in N).$$

Durch (1) wird eine weitere wichtige inhaltliche Vorstellung festgelegt, die mit dem Mengenbegriff verknüpft sein soll. Es wird nämlich postuliert, daß jede Menge eindeutig durch die in ihr enthaltenen Elemente (ihren Umfang) bestimmt sein soll, unabhängig z. B. davon, durch welche Eigenschaft ihrer Elemente sie zunächst definiert wurde. Betrachten wir beispielsweise im Grundbereich aller Dreiecke einer gegebenen Ebene die Eigenschaften „x ist gleichseitig" und „x ist gleichwinklig", so legt nach dem Mengenbildungsprinzip jede von ihnen eine bestimmte Menge fest; nach dem Extensionalitätsprinzip handelt es sich indes in beiden Fällen um dieselbe Menge, denn bekanntlich ist jedes gleichseitige Dreieck gleichwinklig und jedes gleichwinklige Dreieck gleichseitig.

Als eine wichtige Folgerung aus dem Extensionalitätsprinzip erhalten wir, daß *es zu jeder Aussage $H(x)$ über die Objekte eines gegebenen Grundbereichs E nur eine einzige Menge M gibt, die alle und nur diejenigen Objekte x enthält, auf die $H(x)$ zutrifft.* Denn aus

$$\bigwedge_x (x \in M_1 \Leftrightarrow H(x)) \text{ und } \bigwedge_x (x \in M_2 \Leftrightarrow H(x))$$

folgt

$$\bigwedge_x (x \in M_1 \Leftrightarrow x \in M_2),$$

und nach (1) ist dann $M_1 = M_2$. Die durch die Aussage (Eigenschaft) $H(x)$ eindeutig bestimmte Menge M mit der Eigenschaft

$$\bigwedge_x (x \in M \Leftrightarrow H(x))$$

(ihre Existenz ist durch das Mengenbildungsprinzip gesichert) bezeichnet man mit $\{x : H(x)\}$ (gelesen: Menge aller x mit $H(x)$). Für ein beliebiges Objekt x_0 aus dem betrachteten Grundbereich gilt also

(2) $$x_0 \in \{x : H(x)\} \Leftrightarrow H(x_0).$$

Sind $H_1(x)$ und $H_2(x)$ Aussagen über die Objekte eines gegebenen Grundbereichs E, so ist nach (1)

$$\{x : H_1(x)\} = \{x : H_2(x)\} \text{ genau dann, wenn für jedes } x_0 \text{ gilt: } x_0 \in \{x : H_1(x)\} \Leftrightarrow x_0 \in \{x : H_2(x)\}.$$

Andererseits ist nach (2) $x_0 \in \{x : H_1(x)\}$ logisch äquivalent mit $H_1(x_0)$ und $x_0 \in \{x : H_2(x)\}$ logisch äquivalent mit $H_2(x_0)$. Folglich gilt

(3) $$\{x : H_1(x)\} = \{x : H_2(x)\} \Leftrightarrow \bigwedge_x (H_1(x) \Leftrightarrow H_2(x)).$$

Für Mengensysteme wird entsprechend das folgende Extensionalitätsprinzip gefordert:

(4) $$\mathfrak{M} = \mathfrak{N} \Leftrightarrow \bigwedge_X (X \in \mathfrak{M} \Leftrightarrow X \in \mathfrak{N})$$

(*Mengensysteme $\mathfrak{M}$, $\mathfrak{N}$ sind genau dann gleich, wenn sie dieselben Mengen X erster Stufe als Element enthalten*), und analog in höheren Stufen. Ist $H(X)$ eine Aussage über Mengen erster Stufe, so bezeichnet $\{X : H(X)\}$ das eindeutig bestimmte System $\mathfrak{M}$ aller derjenigen Mengen X erster Stufe, auf die $H(X)$ zutrifft, und analog in höheren Stufen. Entsprechend (2) und (3) gilt dann

(5) $$X_0 \in \{X : H(X)\} \Leftrightarrow H(X_0),$$

(6) $$\{X : H_1(X)\} = \{X : H_2(X)\} \Leftrightarrow \bigwedge_X (H_1(X) \Leftrightarrow H_2(X)).$$

Etwas genauer können wir den Inhalt des Extensionalitätsprinzips folgendermaßen beschreiben: Nennen wir im Sinne einer expliziten Definition Mengen M und N *umfangsgleich* (in Zeichen: $M \ominus N$), wenn sie dieselben Elemente enthalten, so gelten zunächst die folgenden Sätze:

(7) *Für jede Menge M gilt:* $M \ominus M$.

(8) *Wenn* $M_1 \ominus M_2$ *und* $M_2 \ominus M_3$, *so* $M_1 \ominus M_3$.

(9) *Wenn* $M_1 \ominus M_2$, *so* $M_2 \ominus M_1$.

Eine Beziehung (in unserem Fall zwischen Mengen), die die Eigenschaften (7), (8), (9) besitzen, nennt man allgemein eine *Äquivalenzrelation* (vgl. 2.5. (12)). Es zeigt sich nun, daß es sehr viele verschiedene Äquivalenzrelationen zwischen Mengen gibt, unter denen die Identität eine ausgezeichnete Rolle spielt (sie ist die in einem bestimmten Sinne kleinste Äquivalenzrelation). Durch das Extensionalitätsprinzip wird nun gerade postuliert, daß *die Umfangsgleichheit die Identität sein soll*, d. h., im Fall $M \ominus N$ die Buchstaben M und N dieselbe Menge bezeichnen. Hierin ist insbesondere enthalten, daß im Fall $M \ominus N$ jedes Mengensystem $\mathfrak{M}$, das die Menge M enthält, auch die Menge N enthält (und umgekehrt), und gerade diese Eigenschaft der Umfangsgleichheit läßt sich nicht (z. B. mittels des Mengenbildungsprinzips) aus ihrer Definition beweisen.

1.4. Mengenalgebra

Im vorliegenden Abschnitt wollen wir einige allgemeine Operationen für Mengen definieren und deren wichtigste Eigenschaften herleiten. Zu Ehren des englischen Logikers George Boole (1815–1869) werden diese Operationen heute vielfach Boolesche Operationen genannt. Bei allen im folgenden betrachteten Mengen, Mengensystemen usw. soll es sich um solche über demselben Grundbereich E handeln.

Es seien zunächst M_1, M_2 beliebige Mengen von Objekten aus E. Nehmen wir im Mengenbildungsprinzip als $H(x)$ die Aussage „$x \in M_1$ und $x \in M_2$", so erhalten wir, daß *es eine Menge M gibt, die alle und nur diejenigen Objekte aus E als Element enthält, die sowohl zu M_1 als auch zu M_2 gehören*, für die also bei beliebigem x gilt:

$$x \in M \Leftrightarrow x \in M_1 \wedge x \in M_2,$$

wobei wir die logische Konjunktion (das Wörtchen „und") durch das Zeichen $\wedge$ wiedergegeben haben. Diese auf Grund des Extensionalitätsprinzips durch M_1 und M_2 eindeutig bestimmte Menge M nennt man den *Durchschnitt* der Mengen M_1, M_2; man bezeichnet den Durchschnitt der Mengen M_1, M_2 heute allgemein mit $M_1 \frown M_2$, während in der älteren Literatur die Bezeichnung $M_1 \cdot M_2$ weit verbreitet war. Es gilt also

(1) $x \in M_1 \frown M_2 :\Leftrightarrow x \in M_1 \wedge x \in M_2$

bzw. bei Verwendung der in 1.3. eingeführten Bezeichnungsweise $\{x : H(x)\}$

(1') $\quad M_1 \cap M_2 :=: \{x : x \in M_1 \wedge x \in M_2\}$.

Der Doppelpunkt vor dem Äquivalenz- bzw. Gleichheitszeichen soll dabei andeuten, daß es sich bei der entsprechenden Zeile um eine Definition handelt (zu lesen: $x \in M_1 \cap M_2$ gilt definitionsgemäß genau dann, wenn $x \in M_1$ und $x \in M_2$, bzw. $M_1 \cap M_2$ ist definitionsgemäß gleich der Menge aller x mit $x \in M_1$ und $x \in M_2$).

Nehmen wir als $H(x)$ entsprechend die Aussage „$x \in M_1$ oder $x \in M_2$", so erhalten wir die heute allgemein mit $M_1 \cup M_2$ und in der älteren Literatur verbreitet mit $M_1 + M_2$ bezeichnete *Vereinigung* oder *Vereinigungsmenge* von M_1 und M_2. Die Vereinigung wird also definiert durch

(2) $\quad x \in M_1 \cup M_2 :\Leftrightarrow x \in M_1 \vee x \in M_2$

bzw.

(2') $\quad M_1 \cup M_2 :=: \{x : x \in M_1 \vee x \in M_2\}$,

wobei wir die logische Alternative (das Wörtchen „oder") durch das Zeichen $\vee$ wiedergegeben haben.

Zur Vermeidung von Mißverständnissen machen wir darauf aufmerksam, daß das Wörtchen „oder" im üblichen mathematischen Sprachgebrauch im nichtausschließenden Sinne (lateinisch: vel) verwendet wird. Die Menge $M_1 \cup M_2$ enthält also genau diejenigen x, die in wenigstens einer der Mengen M_1, M_2 enthalten sind, unter Einschluß aller derjenigen x, die zu beiden Mengen, d. h. zum Durchschnitt $M_1 \cap M_2$ gehören. Im Gegensatz dazu sollte man das ausschließende Oder (lateinisch: aut-aut) durch „entweder-oder" ausdrücken. Die zugehörige Menge bezeichnet man übrigens aus einem weiter unten ersichtlichen Grund als *symmetrische Differenz* $M_1 \triangle M_2$; sie wird definiert durch

(3) $\quad x \in M_1 \triangle M_2 :\Leftrightarrow x \in M_1 \dot{\vee} x \in M_2$

bzw.

(3') $\quad M_1 \triangle M_2 :=: \{x : x \in M_1 \dot{\vee} x \in M_2\}$,

wobei das Zeichen $\dot{\vee}$, das wir allerdings im folgenden nicht systematisch verwenden werden, die logische Antivalenz „entweder ... oder ..." bezeichnet. Ein Objekt x gehört genau dann zu $M_1 \triangle M_2$, wenn es in wenigstens einer der Mengen M_1, M_2 und nicht in beiden Mengen enthalten ist, wenn es also zu genau einer dieser beiden Mengen gehört.

Schließlich bezeichnen wir die Menge aller derjenigen x, die zu M_1, aber nicht zu M_2 gehören, mit $M_1 \backslash M_2$ und nennen sie die *Mengendifferenz* oder

Differenzmenge. Deuten wir die Negation von $x \in M$ durch $\neg x \in M$ oder kürzer $x \notin M$ an, das Zeichen $\neg$ ist also als Abkürzung für das Wörtchen „nicht" anzusehen, so können wir die Differenzmenge durch

(4) $\quad x \in M_1 \backslash M_2 : \Leftrightarrow x \in M_1 \wedge x \notin M_2$

bzw.

(4') $\quad M_1 \backslash M_2 := \{x : x \in M_1 \wedge x \notin M_2\}$

charakterisieren.

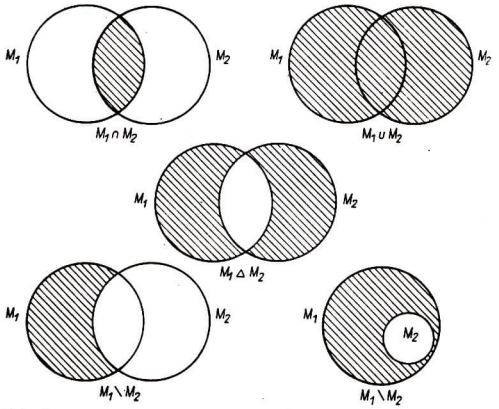


Abb. 1

Man kann sich die bisher eingeführten mengentheoretischen Operationen gut mittels sogenannter Eulerscher Kreise (Venn-Diagramme) veranschaulichen (vgl. Abb. 1). Die vollständige Charakterisierung der Mengen $M_1 \cap M_2$, $M_1 \cup M_2$, $M_1 \triangle M_2$, $M_1 \backslash M_2$ liefert die folgende Tabelle:

M_1	M_2	$M_1 \cap M_2$	$M_1 \cup M_2$	$M_1 \triangle M_2$	$M_1 \backslash M_2$
1	1	1	1	0	0
1	0	0	1	1	1
0	1	0	1	1	0
0	0	0	0	0	0

Die vier Zeilen entsprechen den unter M_1, M_2 angegebenen vier Möglichkeiten, daß nämlich ein gegebenes Objekt x entweder der betreffenden Menge an-

gehört (durch 1 angedeutet) oder nicht angehört (durch 0 angedeutet). Eine 1 oder 0 in der entsprechenden Spalte unter $M_1 \cap M_2, \ldots, M_1 \backslash M_2$ deutet an, ob im jeweiligen Fall x dieser Menge angehört oder nicht. So entspricht die dritte Zeile dem Fall $x \notin M_1, x \in M_2$, und die 1 unter $M_1 \triangle M_2$ gibt an, daß in diesem Fall x zu $M_1 \triangle M_2$ gehört. Zugleich beschreibt diese Tabelle den genauen Gebrauch von logischer Konjunktion („und"), Alternative („oder") und Antivalenz („entweder ... oder ...").

Wir kommen nun zu den wichtigsten algebraischen Rechengesetzen für die bisher eingeführten Mengenoperationen. Als erstes merken wir an, daß sowohl der Durchschnitt als auch die Vereinigung und die symmetrische Differenz dem sogenannten Kommutativgesetz genügen, d. h., *für beliebige Mengen* M_1, M_2 *gilt*

(5) $M_1 \cap M_2 = M_2 \cap M_1$;
(6) $M_1 \cup M_2 = M_2 \cup M_1$;
(7) $M_1 \triangle M_2 = M_2 \triangle M_1$.

Wir wollen am Beispiel von (5) zunächst grundsätzlich klarmachen, was hierfür eigentlich zu beweisen ist. Dazu erinnern wir daran, daß nach dem Extensionalitätsprinzip Mengen (und hierum handelt es sich bei $M_1 \cap M_2$ und $M_2 \cap M_1$) genau dann gleich sind, wenn sie dieselben Elemente enthalten. Daher ist zum Nachweis von (5) zu zeigen, daß bei beliebigem x folgendes gilt:

(5') $x \in M_1 \cap M_2 \Leftrightarrow x \in M_2 \cap M_1$.

Hierfür brauchen wir jedoch nur die Tabelle für $M_1 \cap M_2$ und $M_2 \cap M_1$ zu betrachten:

M_1	M_2	$M_1 \cap M_2$	$M_2 \cap M_1$
1	1	1	1
1	0	0	0
0	1	0	0
0	0	0	0

Da wir unter $M_1 \cap M_2$ dieselbe Verteilung von 1 und 0 haben wie unter $M_2 \cap M_1$, ist unsere Behauptung bewiesen. Analog beweist man (6) und (7). Natürlich bringen (5), (6), (7) lediglich zum Ausdruck, daß die logische Konjunktion, Alternative und Antivalenz kommutativ sind.

Als nächstes zeigt man, daß sowohl der Durchschnitt als auch die Vereinigung und die symmetrische Differenz dem sogenannten Assoziativgesetz genügen, d. h., *für beliebige Mengen* M_1, M_2, M_3 *gilt*

(8) $M_1 \cap (M_2 \cap M_3) = (M_1 \cap M_2) \cap M_3$;
(9) $M_1 \cup (M_2 \cup M_3) = (M_1 \cup M_2) \cup M_3$;
(10) $M_1 \triangle (M_2 \triangle M_3) = (M_1 \triangle M_2) \triangle M_3$.

Wir machen darauf aufmerksam, daß wie im folgenden Fall der Distributivgesetze in den jeweiligen Tabellen bereits acht Fälle zu unterscheiden sind. Abb. 2 veranschaulicht das Assoziativgesetz für die symmetrische Differenz mittels Eulerscher Kreise.

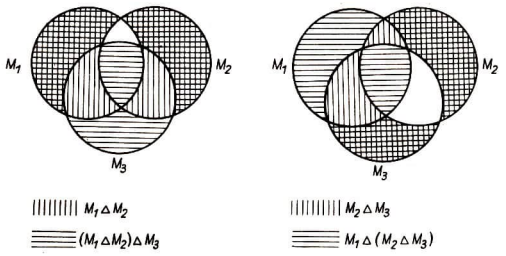


Abb. 2

Weiterhin gelten für Durchschnitt und Vereinigung die folgenden sogenannten Distributivgesetze:

(11) $(M_1 \cup M_2) \cap M_3 = (M_1 \cap M_3) \cup (M_2 \cap M_3)$;
(12) $(M_1 \cap M_2) \cup M_3 = (M_1 \cup M_3) \cap (M_2 \cup M_3)$.

Der Beweis für (11), der sogenannten *rechtsseitigen Distributivität* (= Verteilbarkeit) *des Durchschnitts bezüglich der Vereinigung*, wird durch die folgende Tabelle erbracht:

M_1	M_2	M_3	$M_1 \cup M_2$	$(M_1 \cup M_2) \cap M_3$	$M_1 \cap M_3$	$M_2 \cap M_3$	$(M_1 \cap M_3) \cup (M_2 \cap M_3)$
1	1	1	1	1	1	1	1
1	1	0	1	0	0	0	0
1	0	1	1	1	1	0	1
1	0	0	1	0	0	0	0
0	1	1	1	1	0	1	1
0	1	0	1	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0

Analog beweist man (12), die sogenannte *rechtsseitige Distributivität der Vereinigung bezüglich des Durchschnitts*. Wir empfehlen, sich (11) und (12) auch mittels Eulerscher Kreise zu veranschaulichen.

Die bisher genannten Sätze lassen eine gewisse Analogie zwischen den Rechengesetzen für Vereinigung und Durchschnitt und denen für Addition und Multiplikation von (z. B. reellen) Zahlen erkennen. Sie zeigen aber auch bereits einen wesentlichen Unterschied. Während nämlich für Vereinigung und

Durchschnitt beide Distributivgesetze gelten, ist bekanntlich beim Zahlenrechnen zwar die Multiplikation distributiv bezüglich der Addition (d. h., für beliebige reelle Zahlen a, b, c gilt $(a+b)\cdot c = a\cdot c + b\cdot c$), aber nicht die Addition bezüglich der Multiplikation. Noch deutlicher wird der Unterschied, wenn man beachtet, daß es sich bei Durchschnitt und Vereinigung um sogenannte idempotente Operationen handelt, d. h., *für beliebige Mengen M gilt*

(13) $M \cap M = M$;

(14) $M \cup M = M$.

Diese Gesetze haben beim Zahlenrechnen überhaupt kein Analogon mehr. Wir erwähnen weiterhin, daß für Durchschnitt und Vereinigung die folgenden merkwürdigen Verschmelzungssätze erfüllt sind: *Für beliebige Mengen M_1, M_2 gilt*

(15) $(M_1 \cup M_2) \cap M_1 = M_1$;

(16) $(M_1 \cap M_2) \cup M_1 = M_1$.

Als ein Kuriosum erwähnen wir schließlich, daß auch *der Durchschnitt distributiv bezüglich der symmetrischen Differenz ist:*

(17) $(M_1 \triangle M_2) \cap M_3 = (M_1 \cap M_3) \triangle (M_2 \cap M_3)$.

Die Assoziativgesetze (8), (9), (10) beinhalten, daß es bei einem Durchschnitt, einer Vereinigung und einer symmetrischen Differenz aus drei Mengen M_1, M_2, M_3 nicht darauf ankommt, in welcher Weise man sie durch Klammersetzung aus zweigliedrigen Durchschnitten, Vereinigungen bzw. symmetrischen Differenzen aufbaut. Daher kann man – wie auch beim Zahlenrechnen – ganz auf Klammersetzungen verzichten und einfach $M_1 \cap M_2 \cap M_3$, $M_1 \cup M_2 \cup M_3$ bzw. $M_1 \triangle M_2 \triangle M_3$ schreiben. Es liegt auf der Hand (systematisch kommen wir hierauf in 3.5. zurück), daß Analoges auch für Durchschnitte, Vereinigungen und symmetrische Differenzen aus vier und mehr Mengen gilt. Wir merken an, daß allgemein $M_1 \cap \cdots \cap M_k$ ($k \geqq 2$) die Menge aller derjenigen x ist, die sämtlichen Mengen $M_1, \ldots, M_k$ angehören. Entsprechend ist $M_1 \cup \cdots \cup M_k$ die Menge aller x, die in wenigstens einer der Mengen $M_1, \ldots, M_k$ enthalten sind, und $M_1 \triangle \cdots \triangle M_k$ die Menge aller x, die einer ungeraden Anzahl der Mengen $M_1, \ldots, M_k$ angehören. Ebenso lassen sich auch die Kommutativgesetze (5), (6), (7) leicht auf mehrgliedrige Durchschnitte, Vereinigungen und symmetrische Differenzen übertragen. Unter Verwendung der Kommutativgesetze (5), (6) gelangt man von den rechtsseitigen Distributivgesetzen (11), (12) sofort zu den folgenden linksseitigen Distributivgesetzen:

(11′) $M_3 \cap (M_1 \cup M_2) = (M_3 \cap M_1) \cup (M_3 \cap M_2)$;

(12′) $M_3 \cup (M_1 \cap M_2) = (M_3 \cup M_1) \cap (M_3 \cup M_2)$.

Schließlich kann man durch mehrfache Anwendung der einfachen Distributivgesetze auch ohne Schwierigkeiten die vom Zahlenrechnen bekannten Verallgemeinerungen der Distributivgesetze erhalten, wie z. B.

$$(M_1 \cap M_2) \cup (M_3 \cap M_4 \cap M_5)$$
$$= (M_1 \cup M_3) \cap (M_1 \cup M_4) \cap (M_1 \cup M_5) \cap (M_2 \cup M_3)$$
$$\cap (M_2 \cup M_4) \cap (M_2 \cup M_5).$$

Nehmen wir im Mengenbildungsprinzip als $H(x)$ die Aussage „$x = x$", die auf alle Objekte des betrachteten Grundbereichs E zutrifft, so erhalten wir eine Menge, die sämtliche Objekte aus E enthält und die wir folglich mit dem Grundbereich E identifizieren können. Nehmen wir die Aussage „$x \neq x$", die auf kein Objekt zutrifft, so erhalten wir eine Menge, die kein Objekt enthält; man nennt diese Menge die *leere Menge* und bezeichnet sie heute meistens mit $\emptyset$. Die leere Menge spielt für Durchschnitt und Vereinigung eine ähnliche Rolle wie die Zahl Null bei der Multiplikation und Addition. *Für jede Menge M gilt*

(18) $M \cap \emptyset = \emptyset$;

(19) $M \cup \emptyset = M$.

Die vom Zahlenrechnen bekannte Eigenschaft, daß ein Produkt auch nur dann Null ist, wenn wenigstens ein Faktor Null ist, hat dagegen kein Analogon. Vielmehr ist die Beziehung $M_1 \cap M_2 = \emptyset$ charakteristisch dafür, daß die Mengen M_1, M_2 kein Element gemeinsam haben. Man nennt solche Mengen M_1, M_2 *elementefremd* oder *disjunkt:*

(20) M_1 *disjunkt* (*elementefremd*) *zu* $M_2 : \Leftrightarrow M_1 \cap M_2 = \emptyset$.

Als Gesetze für die Differenz seien hier vor allem die folgenden genannt:

(21) $(M_1 \cap M_2) \backslash M_3 = (M_1 \backslash M_3) \cap (M_2 \backslash M_3)$;

(22) $(M_1 \cup M_2) \backslash M_3 = (M_1 \backslash M_3) \cup (M_2 \backslash M_3)$;

(23) $M_1 \backslash (M_2 \cap M_3) = (M_1 \backslash M_2) \cup (M_1 \backslash M_3)$;

(24) $M_1 \backslash (M_2 \cup M_3) = (M_1 \backslash M_2) \cap (M_1 \backslash M_3)$;

(25) $M_1 \backslash (M_2 \backslash M_3) = (M_1 \backslash M_2) \cup (M_1 \cap M_3)$;

(26) $M \backslash \emptyset = M$;

(27) $M \backslash M = \emptyset$;

(28) $M_1 \backslash (M_1 \backslash M_2) = M_1 \cap M_2$;

(29) $M_1 \triangle M_2 = (M_1 \cup M_2) \backslash (M_1 \cap M_2)$;

(30) $M_1 \triangle M_2 = (M_1 \backslash M_2) \cup (M_2 \backslash M_1)$.

Die Beweise aller dieser Sätze, die wir dem Leser als Übungsaufgaben überlassen, werden in bekannter Weise nach der Tabellenmethode geführt. Man kann sie sich auch leicht, was wir dem Leser dringend empfehlen, mittels

Eulerscher Kreise veranschaulichen. Bei den Gesetzen (23), (24), die man häufig de-Morgansche Regeln nennt, ist der Austausch von Vereinigung und Durchschnitt zu beachten. Die Sätze (29) und (30) zeigen, daß die symmetrische Differenz prinzipiell entbehrlich ist, da sie sich mittels Vereinigung, Durchschnitt und Differenz ausdrücken läßt. Durch (30) wird übrigens die Bezeichnung symmetrische Differenz erklärt. Aus (29) und (26) entnimmt man leicht, daß für disjunkte Mengen und nur für solche die symmetrische Differenz mit der Vereinigung übereinstimmt:

(31) M_1 *disjunkt zu* $M_2 \Leftrightarrow M_1 \triangle M_2 = M_1 \cup M_2$.

Alle im vorliegenden Abschnitt behandelten Begriffe lassen sich entsprechend auf Mengen höherer Stufe übertragen, wobei ganz analoge Sätze gelten. So wird man z. B. als Durchschnitt $\mathfrak{M}_1 \cap \mathfrak{M}_2$ zweier Mengensysteme $\mathfrak{M}_1$, $\mathfrak{M}_2$ das System aller derjenigen Mengen X erster Stufe definieren, die sowohl zu $\mathfrak{M}_1$ als auch zu $\mathfrak{M}_2$ gehören, d. h.

(32) $\mathfrak{M}_1 \cap \mathfrak{M}_2 := \{X : X \in \mathfrak{M}_1 \wedge X \in \mathfrak{M}_2\}$,

usw. Nimmt man im Mengenbildungsprinzip für Mengensysteme als $H(X)$ die auf keine Menge erster Stufe zutreffende Aussage „$X \neq X$", so erhält man die Existenz eines Mengensystems $\mathfrak{O}$, das keine Menge erster Stufe als Element enthält und das als *leeres System* bezeichnet wird. Das leere System $\mathfrak{O}$ ist bei unserem Ansatz zunächst begrifflich durchaus von der leeren Menge erster Stufe zu unterscheiden. Zum Beispiel kann die leere Menge erster Stufe Element eines gewissen Mengensystems $\mathfrak{M}$ sein (das dann vom leeren System verschieden ist!), niemals kann das aber bei unserer Stufenvereinbarung für das leere System der Fall sein. Es werden jedoch im folgenden keine Mißverständnisse auftreten, wenn wir das leere System und auch die leeren Mengen höherer Stufe sämtlich mit dem Symbol $\emptyset$ bezeichnen. Man muß sich nur in jedem Fall klar machen, als leere Menge welcher Stufe im betreffenden Zusammenhang dieses Symbol verstanden werden muß.

1.5. Die Inklusion

Eine wichtige Relation zwischen Mengen ist die Teilmengenbeziehung oder Inklusion. Man nennt eine Menge M_1 eine *Teil-* oder *Untermenge* der Menge M_2, wenn jedes Element der Menge M_1 auch Element der Menge M_2 ist. Ausführlicher sagt man dafür auch, daß M_1 *als Teilmenge in* M_2 *enthalten ist,*

und schreibt hierfür $M_1 \subseteq M_2$. Es gilt also

(1) $\quad M_1 \subseteq M_2 : \Leftrightarrow \bigwedge_x (x \in M_1 \Rightarrow x \in M_2)$,

wobei das Zeichen $\Rightarrow$ als Abkürzung für die logische Implikation „wenn ..., so ..." verwendet wird. Ist M_1 eine Teilmenge von M_2, so nennt man M_2 eine *Obermenge* von M_1 und schreibt $M_2 \supseteq M_1$, d. h.

(2) $\quad M_2 \supseteq M_1 : \Leftrightarrow M_1 \subseteq M_2$.

Man sagt hierfür auch, daß die Menge M_2 die Menge M_1 *umfaßt*.

Aus der angegebenen Definition erhält man mühelos die folgenden Grundeigenschaften der Inklusion:

(3) *Für jede Menge M gilt $M \subseteq M$* (Reflexivität);

(4) $\bigwedge_{M_1, M_2, M_3} (M_1 \subseteq M_2 \wedge M_2 \subseteq M_3 \Rightarrow M_1 \subseteq M_3)$ (Transitivität);

(5) $\bigwedge_{M_1, M_2} (M_1 \subseteq M_2 \wedge M_2 \subseteq M_1 \Rightarrow M_1 = M_2)$ (Antisymmetrie).

Eine Relation, die diese drei Eigenschaften besitzt, nennt man heute allgemein eine *teilweise Ordnung* (vgl. 2.5. (24)).

Die Inklusion hat also ähnliche Eigenschaften wie die $\leq$-Beziehung für reelle Zahlen. Allerdings müssen wir auch hier sofort wieder auf einen wesentlichen Unterschied hinweisen. Während die $\leq$-Beziehung für Zahlen *linear* ist, d. h., für beliebige reelle Zahlen a, b stets $a \leq b$ oder $b \leq a$ gilt, gibt es Mengen M_1, M_2, die *unvergleichbar* sind, für die weder $M_1 \subseteq M_2$ noch $M_2 \subseteq M_1$ gilt.

Ist jedes Element der Menge M_1 auch Element von M_2 und gibt es ein Element $x_0 \in M_2$ mit $x_0 \notin M_1$, d. h., ist $M_1 \subseteq M_2$ und $M_1 \neq M_2$, so heißt M_1 eine *echte Teilmenge* von M_2 und M_2 eine *echte Obermenge* von M_1; in Zeichen drückt man das durch $M_1 \subset M_2$ und $M_2 \supset M_1$ aus:

(6) $\quad M_1 \subset M_2 : \Leftrightarrow M_1 \subseteq M_2 \wedge M_1 \neq M_2$;

(7) $\quad M_2 \supset M_1 : \Leftrightarrow M_1 \subset M_2$.

Wir machen darauf aufmerksam, daß in der Literatur die Inklusion vielfach durch $M_1 \subset M_2$ und dann die echte Inklusion durch $M_1 \subsetneq M_2$ bezeichnet wird. Im Hinblick auf die Analogien zur $\leq$- und $<$-Beziehung für Zahlen halten wir dies jedoch für wenig zweckmäßig.

Aus der Definition des Durchschnitts folgt zunächst unmittelbar, daß *der Durchschnitt eine gemeinsame Teilmenge von M_1 und M_2 ist*, d. h.

(8) $\quad M_1 \cap M_2 \subseteq M_1, \quad M_1 \cap M_2 \subseteq M_2$.

Wir wollen nun zeigen, daß *der Durchschnitt die bezüglich der Inklusion größte Menge mit dieser Eigenschaft ist*, d. h., ist Z eine beliebige gemeinsame Teil-

menge von M_1 und M_2, so ist $Z \subseteq M_1 \cap M_2$. In logischer Abkürzung drückt sich das so aus:

(9) $\bigwedge_Z (Z \subseteq M_1 \wedge Z \subseteq M_2 \Rightarrow Z \subseteq M_1 \cap M_2)$.

Es sei also $Z \subseteq M_1$ und $Z \subseteq M_2$ und x ein beliebiges Element der Menge Z. Wegen $Z \subseteq M_1$ ist dann $x \in M_1$, und wegen $Z \subseteq M_2$ ist $x \in M_2$. Also ist $x \in M_1 \cap M_2$. Da das für jedes $x \in Z$ gilt, ist $Z \subseteq M_1 \cap M_2$, was zu zeigen war.

Wir merken an, daß *der Durchschnitt $M_1 \cap M_2$ auch die einzige Menge ist, die die Bedingungen* (8) *und* (9) *erfüllt*. Zum Beweis nehmen wir an, es sei D eine beliebige Menge mit diesen Eigenschaften, d. h., es gelte

(8′) $D \subseteq M_1, \quad D \subseteq M_2$;

(9′) $\bigwedge_Z (Z \subseteq M_1 \wedge Z \subseteq M_2 \Rightarrow Z \subseteq D)$.

Die Bedingung (8′) besagt dann gerade, daß D eine Menge Z ist, die die Voraussetzungen von (9) erfüllt, so daß wegen (9) $D \subseteq M_1 \cap M_2$ gilt. Andererseits besagt (8), daß $M_1 \cap M_2$ eine Menge Z ist, die die Voraussetzungen von (9′) erfüllt, so daß wegen (9′) $M_1 \cap M_2 \subseteq D$ gilt. Aus $D \subseteq M_1 \cap M_2$ und $M_1 \cap M_2 \subseteq D$ folgt aber nach (5) $D = M_1 \cap M_2$. Also ist in der Tat $M_1 \cap M_2$ die einzige Menge D, für die (8′) und (9′) gelten.

Ganz analog beweist man (Übungsaufgabe), daß *die Vereinigung $M_1 \cup M_2$ die bezüglich der Inklusion kleinste gemeinsame Obermenge von M_1 und M_2 ist*, d. h.

(10) $M_1 \subseteq M_1 \cup M_2, \quad M_2 \subseteq M_1 \cup M_2$;

(11) $\bigwedge_Z (M_1 \subseteq Z \wedge M_2 \subseteq Z \Rightarrow M_1 \cup M_2 \subseteq Z)$,

und ebenso wie beim Durchschnitt erkennt man, daß *die Vereinigung eindeutig durch diese beiden Eigenschaften charakterisiert ist.*

Ferner beweist man leicht, daß *Durchschnitt und Vereinigung bezüglich der Inklusion monoton sind*, d. h.

(12) $M_1 \subseteq M_2 \Rightarrow M_1 \cap M_3 \subseteq M_2 \cap M_3$;

(13) $M_1 \subseteq M_2 \Rightarrow M_1 \cup M_3 \subseteq M_2 \cup M_3$.

Für die Differenz gelten dabei die folgenden Gesetze:

(14) $M_1 \subseteq M_2 \Rightarrow M_1 \backslash M_3 \subseteq M_2 \backslash M_3$;

(15) $M_1 \subseteq M_2 \Rightarrow M_3 \backslash M_1 \supseteq M_3 \backslash M_2$.

Als Folgerung aus bereits Bewiesenem erhalten wir, daß *sich die Inklusion auch sehr einfach mittels Durchschnitt oder Vereinigung ausdrücken läßt*. Es gilt

nämlich

(16) $M_1 \subseteq M_2 \Leftrightarrow M_1 \cap M_2 = M_1$;
(17) $M_1 \subseteq M_2 \Leftrightarrow M_1 \cup M_2 = M_2$.

Wir beweisen als Beispiel die Behauptung (17); der Beweis für (16) kann analog erbracht werden (Übungsaufgabe). Zunächst nehmen wir an, es sei $M_1 \subseteq M_2$. Dann ist nach (13) (mit M_2 als M_3) $M_1 \cup M_2 \subseteq M_2 \cup M_2 = M_2$ (vgl. 1.4.(14)), und da die umgekehrte Inklusion $M_2 \subseteq M_1 \cup M_2$ nach (10) ebenfalls gilt, ist nach (5) $M_1 \cup M_2 = M_2$. Ist umgekehrt $M_1 \cup M_2 = M_2$, so ist wegen $M_1 \subseteq M_1 \cup M_2$ auch $M_1 \subseteq M_2$.

Als nächstes wollen wir eine Eigenschaft der Inklusion behandeln, die dem Anfänger meistens zunächst etwas Kopfzerbrechen bereitet. Wir behaupten nämlich, daß *die leere Menge* $\emptyset$ *Teilmenge jeder beliebigen Menge* M *ist*:

(18) $\emptyset \subseteq M$.

Zum Beweis haben wir folgendes zu zeigen:

(18') $\bigwedge_x (x \in \emptyset \Rightarrow x \in M)$.

Warum ist das der Fall? Hierzu ist einfach folgendes zu sagen: Beim üblichen mathematischen Sprachgebrauch ist eine Implikation „wenn p, so q", deren *Voraussetzung* oder *Prämisse* p falsch ist, grundsätzlich wahr. Mithin gilt (18') einfach deshalb, weil die Voraussetzung $x \in \emptyset$ für jedes x falsch ist, da ja die leere Menge kein Element enthält. Man könnte das im vorliegenden Fall vielleicht noch durch folgendes Argument ergänzen: Wäre nicht jedes Element der leeren Menge auch Element von M, so müßte man in der leeren Menge ein Element finden können, das nicht zu M gehört; das geht aber deshalb nicht, weil die leere Menge ja kein Element enthält.

Wir betrachten nun bei gegebener Menge M die Eigenschaft „$X \subseteq M$", die auf genau diejenigen Mengen (!) zutrifft, die Teilmengen von M sind. Das Mengensystem (!), das aus allen diesen Mengen besteht, heißt die *Potenzmenge* von M und wird mit $\mathfrak{P}(M)$ bezeichnet. Es gilt also

(19) $\mathfrak{P}(M) := \{X : X \subseteq M\}$.

Da nach (18) bzw. (3) die leere Menge und die Menge M Teilmengen von M sind, ist stets $\emptyset \in \mathfrak{P}(M)$ und $M \in \mathfrak{P}(M)$. Falls die Menge M selbst leer ist, reduziert sich die Potenzmenge $\mathfrak{P}(M)$ auf dasjenige Mengensystem, dessen einziges Element die leere Menge ist (dieses Mengensystem ist nicht leer!).

Es sei nun a ein fest gewähltes Objekt aus dem gegebenen Grundbereich E. Wir betrachten die Aussage „$x = a$". Man sieht sofort, daß diese Aussage einzig und allein auf das Objekt a zutrifft. Folglich enthält die durch diese

Aussage definierte Menge das Objekt a und nur dieses als Element. Man nennt sie daher die *Einermenge* aus (dem Objekt) a; für sie ist die Bezeichnung $\{a\}$ üblich:

(20) $\{a\} := \{x : x = a\}$.

Aus der Definition der Einermenge ergibt sich sofort

(21) $a \in M \Leftrightarrow \{a\} \subseteq M$.

Wegen $a \in \{a\}$ ist die Einermenge $\{a\}$ sicher von der leeren Menge verschieden, also $\{a\} \supset \emptyset$. Andererseits sieht man sofort, daß es keine Menge N mit $\emptyset \subset N \subset \{a\}$ geben kann. Für beides zusammen sagt man auch, daß die Menge $\{a\}$ bezüglich der Inklusion ein *oberer Nachbar* der leeren Menge ist. Umgekehrt ist unmittelbar klar, daß auch jeder obere Nachbar der leeren Menge eine Einermenge ist. Allgemein kann man zeigen (Übungsaufgabe), daß eine Menge M_2 genau dann oberer Nachbar einer Menge M_1 ist, wenn $M_2 = M_1 \cup \{a\}$ gilt, wobei a ein nicht zu M_1 gehöriges Objekt ist (im Fall $a \in M_1$ ist natürlich $M_1 \cup \{a\} = M_1$).

Sind a, b beliebige Objekte aus E, so wird

(21) $\{a, b\} := \{a\} \cup \{b\}$

gesetzt. Man erkennt sofort, daß folgendes gilt:

(22) $\{a, b\} = \{x : x = a \vee x = b\}$,

so daß die Menge $\{a, b\}$ die Objekte a, b und nur diese als Elemente enthält. Aus dem Kommutativgesetz für die Vereinigung (vgl. 1.4.(6)) folgt

(23) $\{a, b\} = \{b, a\}$,

was natürlich andererseits auch ein Spezialfall des Extensionalitätsprinzips ist (die Mengen $\{a, b\}$ und $\{b, a\}$ enthalten dieselben Elemente und stimmen daher überein). Ist $a = b$, so ist natürlich $\{a, b\} = \{a\} = \{b\}$, während im Fall $a \neq b$ die Menge $\{a, b\}$ ein gemeinsamer oberer Nachbar der Einermengen $\{a\}$ und $\{b\}$ ist. Man nennt im Fall $a \neq b$ die Menge $\{a, b\}$ eine *Zweiermenge*. Die Zweiermengen sind dann gerade die oberen Nachbarn von Einermengen.

Analog wird für Objekte a, b, c aus E

(24) $\{a, b, c\} := \{a\} \cup \{b\} \cup \{c\}$

gesetzt. Hieraus folgt

(25) $\{a, b, c\} = \{x : x = a \vee x = b \vee x = c\}$,

so daß die Menge $\{a, b, c\}$ die Objekte a, b, c und nur diese als Elemente enthält. Aus dem Kommutativgesetz für die Vereinigung folgt, daß es bei der Menge $\{a, b, c\}$ nicht darauf ankommt, in welcher Reihenfolge man die Objekte

a, b, c zwischen den geschweiften Klammern aufzählt, d. h.

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

Ferner ist

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

Sind die Objekte a, b, c paarweise verschieden, in diesem Fall nennt man $\{a, b, c\}$ eine *Dreiermenge*, so ist also $\{a, b, c\}$ ein gemeinsamer oberer Nachbar der Zweiermengen $\{a, b\}$, $\{a, c\}$, $\{b, c\}$. Sind gewisse der Objekte a, b, c gleich, so reduziert sich natürlich $\{a, b, c\}$ auf eine Zweier- oder sogar auf eine Einermenge.

Ist allgemein n eine beliebige natürliche Zahl und sind $a_1, \ldots, a_n$ vorgegebene Objekte aus E, so wird

(28) $\{a_1, \ldots, a_n\} := \{a_1\} \cup \cdots \cup \{a_n\}$

gesetzt. Dann ist

(29) $\{a_1, \ldots, a_n\} = \{x : x = a_1 \vee \cdots \vee x = a_n\}$,

so daß die Menge $\{a_1, \ldots, a_n\}$ die Objekte $a_1, \ldots, a_n$ und nur diese als Elemente enthält. Auch bei der Menge $\{a_1, \ldots, a_n\}$ kommt es nicht darauf an, in welcher Reihenfolge man die Objekte $a_1, \ldots, a_n$ zwischen den geschweiften Klammern aufzählt. Sind die Objekte $a_1, \ldots, a_n$ paarweise verschieden, so enthält die Menge $\{a_1, \ldots, a_n\}$ genau n Elemente, während sie sich andernfalls auf eine bestimmte Menge $\{a_{i_1}, \ldots, a_{i_k}\}$ $(1 \leqq i_1 < \cdots < i_k \leqq n)$ mit $k < n$ paarweise verschiedenen Elementen reduziert. Die Mengen aus n Elementen sind dabei genau die (bezüglich der Inklusion) oberen Nachbarn von Mengen aus $n - 1$ Elementen.

Wir bemerken abschließend, daß sich auch alle im vorliegenden Abschnitt eingeführten Begriffsbildungen unmittelbar auf Mengen höherer Stufe übertragen lassen, wobei ganz analoge Sätze gelten. Es sei darauf hingewiesen, daß die Potenzmenge einer Menge k-ter Stufe natürlich allgemein eine Menge $(k + 1)$-ter Stufe ist.

1.6. Durchschnitt und Vereinigung eines Mengensystems

In 1.4. hatten wir den Durchschnitt $M_1 \cap M_2$ der Mengen M_1, M_2 als Menge aller Objekte x definiert, die sowohl zu M_1 als auch zu M_2 gehören. Unter Verwendung der in 1.5.(23) eingeführten Bezeichnungsweise können wir dafür auch sagen, daß $M_1 \cap M_2$ die Menge aller derjenigen x ist, die in allen Mengen

des Mengensystems $\{M_1, M_2\}$ enthalten sind, denn dieses Mengensystem besteht ja gerade aus den Mengen M_1, M_2. Setzen wir an die Stelle des Mengensystems $\{M_1, M_2\}$ ein beliebiges Mengensystem $\mathfrak{M}$, so gelangen wir zum allgemeinen Begriff des *Durchschnitts eines Mengensystems* $\mathfrak{M}$ (genauer: der Mengen des Mengensystems $\mathfrak{M}$). Dieser Durchschnitt, wir wollen ihn mit $\bigcap \mathfrak{M}$ bezeichnen, ist also die Menge aller derjenigen Objekte x des zugrundeliegenden Bereichs E von Urelementen, die in sämtlichen Mengen X des Systems $\mathfrak{M}$ als Element enthalten sind:

$$(1) \qquad x \in \bigcap \mathfrak{M} :\Leftrightarrow \bigwedge_X (X \in \mathfrak{M} \Rightarrow x \in X)$$

bzw.

$$(1') \qquad \bigcap \mathfrak{M} := \{x : \bigwedge_X (X \in \mathfrak{M} \Rightarrow x \in X)\}.$$

Ist speziell $\mathfrak{M} = \{X : H(X)\}$, wobei $H(X)$ eine Aussage über Mengen erster Stufe ist, so schreibt man statt $\bigcap \{X : H(X)\}$ auch $\bigcap_{H(X)} X$ (gelesen: Durchschnitt aller Mengen X mit der Eigenschaft $H(X)$).

Entsprechend wird die *Vereinigung eines Mengensystems* $\mathfrak{M}$, wir wollen sie mit $\bigcup \mathfrak{M}$ bezeichnen, als Menge aller derjenigen Objekte x definiert, die in wenigstens einer Menge X des Systems $\mathfrak{M}$ als Element enthalten sind:

$$(2) \qquad x \in \bigcup \mathfrak{M} :\Leftrightarrow \bigvee_X (X \in \mathfrak{M} \wedge x \in X)$$

bzw.

$$(2') \qquad \bigcup \mathfrak{M} := \{x : \bigvee_X (X \in \mathfrak{M} \wedge x \in X)\}.$$

Im Fall $\mathfrak{M} = \{X : H(X)\}$ schreibt man analog wie beim Durchschnitt statt $\bigcup \{X : H(X)\}$ auch $\bigcup_{H(X)} X$.

Wir machen ausdrücklich darauf aufmerksam, daß der Durchschnitt und die Vereinigung, eines Mengensystems, d. h. einer Menge zweiter Stufe, Mengen erster Stufe sind. Bei sinngemäßer Verallgemeinerung der hier eingeführten Begriffe auf Mengen höherer Stufe erhält man allgemein als Durchschnitt und Vereinigung einer Menge k-ter Stufe ($k \geqq 2$) Mengen $(k-1)$-ter Stufe.

Von den Eigenschaften des allgemeinen Durchschnitts und der allgemeinen Vereinigung seien hier zunächst die folgenden genannt:

$$(3) \qquad \bigcap \{M_1, M_2\} = M_1 \cap M_2;$$
$$(4) \qquad \bigcup \{M_1, M_2\} = M_1 \cup M_2$$

und allgemeiner

$$(3') \qquad \bigcap \{M_1, \ldots, M_n\} = M_1 \cap \cdots \cap M_n;$$
$$(4') \qquad \bigcup \{M_1, \ldots, M_n\} = M_1 \cup \cdots \cup M_n.$$

Ferner gilt

(5) $\bigcap(\mathfrak{M}_1 \cup \mathfrak{M}_2) = \bigcap \mathfrak{M}_1 \frown \bigcap \mathfrak{M}_2$;
(6) $\bigcup(\mathfrak{M}_1 \cup \mathfrak{M}_2) = \bigcup \mathfrak{M}_1 \cup \bigcup \mathfrak{M}_2$;
(7) $\mathfrak{M}_1 \subseteq \mathfrak{M}_2 \Rightarrow \bigcap \mathfrak{M}_1 \supseteq \bigcap \mathfrak{M}_2$;
(8) $\mathfrak{M}_1 \subseteq \mathfrak{M}_2 \Rightarrow \bigcup \mathfrak{M}_1 \subseteq \bigcup \mathfrak{M}_2$.

Die Behauptungen (3) (vgl. den ersten Absatz dieses Abschnitts), (3'), (4), (4') sind unmittelbar Folgerungen aus den Definitionen.

Zum Beweis von (5) sei zunächst x ein beliebiges Element der Menge $\bigcap(\mathfrak{M}_1 \cup \mathfrak{M}_2)$. Dann ist x Element jeder Menge X des Mengensystems $\mathfrak{M}_1 \cup \mathfrak{M}_2$, also wegen $\mathfrak{M}_1 \subseteq \mathfrak{M}_1 \cup \mathfrak{M}_2$ insbesondere jeder Menge des Systems $\mathfrak{M}_1$ und wegen $\mathfrak{M}_2 \subseteq \mathfrak{M}_1 \cup \mathfrak{M}_2$ auch jeder Menge des Systems $\mathfrak{M}_2$. Folglich ist x sowohl Element von $\bigcap \mathfrak{M}_1$ als auch von $\bigcap \mathfrak{M}_2$ und mithin von $\bigcap \mathfrak{M}_1 \frown \bigcap \mathfrak{M}_2$. Da das für jedes $x \in \bigcap(\mathfrak{M}_1 \cup \mathfrak{M}_2)$ der Fall ist, gilt

(5a) $\bigcap(\mathfrak{M}_1 \cup \mathfrak{M}_2) \subseteq \bigcap \mathfrak{M}_1 \frown \bigcap \mathfrak{M}_2$.

Es sei nun umgekehrt x ein beliebiges Element der Menge $\bigcap \mathfrak{M}_1 \frown \bigcap \mathfrak{M}_2$. Dann ist $x \in \bigcap \mathfrak{M}_1$ und $x \in \bigcap \mathfrak{M}_2$, und folglich ist x Element sowohl jeder Menge X des Systems $\mathfrak{M}_1$ als auch jeder Menge X des Systems $\mathfrak{M}_2$. Ist nun X_0 eine beliebige Menge des Systems $\mathfrak{M}_1 \cup \mathfrak{M}_2$, so ist $X_0 \in \mathfrak{M}_1$ oder $X_0 \in \mathfrak{M}_2$. In jedem dieser beiden Fälle ist aber, wie gesagt, $x \in X_0$. Da das für alle Mengen $X_0 \in \mathfrak{M}_1 \cup \mathfrak{M}_2$ gilt, ist $x \in \bigcap(\mathfrak{M}_1 \cup \mathfrak{M}_2)$. Mithin ist jedes Element der Menge $\bigcap \mathfrak{M}_1 \frown \bigcap \mathfrak{M}_2$ auch Element von $\bigcap(\mathfrak{M}_1 \cup \mathfrak{M}_2)$, d. h.

(5b) $\bigcap \mathfrak{M}_1 \frown \bigcap \mathfrak{M}_2 \subseteq \bigcap(\mathfrak{M}_1 \cup \mathfrak{M}_2)$.

Aus (5a) und (5b) folgt aber auf Grund der Antisymmetrie der Inklusion (1.5.(5)) sofort die Gleichung (5).

Analog wird (6) bewiesen (Übungsaufgabe).

Zum Beweis von (7) seien $\mathfrak{M}_1$, $\mathfrak{M}_2$ beliebige Mengensysteme, die der Voraussetzung $\mathfrak{M}_1 \subseteq \mathfrak{M}_2$ von (7) genügen, und es sei x ein beliebiges Element der Menge $\bigcap \mathfrak{M}_2$. Dann gilt $x \in X$ für alle $X \in \mathfrak{M}_2$ und wegen $\mathfrak{M}_1 \subseteq \mathfrak{M}_2$ damit erst recht für alle $X \in \mathfrak{M}_1$, und folglich ist $x \in \bigcap \mathfrak{M}_1$. Wenn aber jedes Element x der Menge $\bigcap \mathfrak{M}_2$ auch Element von $\bigcap \mathfrak{M}_1$ ist, ist $\bigcap \mathfrak{M}_2$ eine Teilmenge von $\bigcap \mathfrak{M}_1$ und $\bigcap \mathfrak{M}_1$ eine Obermenge von $\bigcap \mathfrak{M}_2$, wie in (7) behauptet wird.

Analog wird (8) bewiesen (Übungsaufgabe).

In Verallgemeinerung von 1.5.(8) und 1.5.(9) kann man behaupten, daß *die Menge* $\bigcap \mathfrak{M}$ *die bezüglich der Inklusion größte Menge ist, die Teilmenge aller*

Mengen X des Systems $\mathfrak{M}$ ist, d. h.

$$(9) \qquad \bigwedge_X (X \in \mathfrak{M} \Rightarrow \bigcap \mathfrak{M} \subseteq X),$$

$$(10) \qquad \bigwedge_Z \left(\bigwedge_X (X \in \mathfrak{M} \Rightarrow Z \subseteq X) \Rightarrow Z \subseteq \bigcap \mathfrak{M}\right),$$

wobei $\bigcap \mathfrak{M}$ auch die einzige Menge ist, die diesen Bedingungen genügt. Analog ist die Menge $\bigcup \mathfrak{M}$ die bezüglich der Inklusion kleinste Menge, die alle Mengen X des Systems $\mathfrak{M}$ umfaßt, d. h.

$$(11) \qquad \bigwedge_X (X \in \mathfrak{M} \Rightarrow X \subseteq \bigcup \mathfrak{M}),$$

$$(12) \qquad \bigwedge_Z \left(\bigwedge_X (X \in \mathfrak{M} \Rightarrow X \subseteq Z) \Rightarrow \bigcup \mathfrak{M} \subseteq Z\right),$$

wobei auch hier $\bigcup \mathfrak{M}$ die einzige Menge ist, die diesen Bedingungen genügt.

Wir beweisen als Beispiel die Behauptungen für $\bigcup \mathfrak{M}$. Zum Beweis von (11) sei X eine beliebige Menge des Systems $\mathfrak{M}$ und x ein beliebiges Element aus X. Dann ist x in wenigstens einer Menge des Systems $\mathfrak{M}$ (nämlich in X) als Element enthalten, und folglich gilt $x \in \bigcup \mathfrak{M}$. Also ist in der Tat $X \subseteq \bigcup \mathfrak{M}$, wie in (11) behauptet wird. Zum Beweis von (12) sei Z eine beliebige Menge, die die Voraussetzung von (12) erfüllt, die also sämtliche Mengen X des Systems $\mathfrak{M}$ umfaßt, und es sei x ein beliebiges Element der Menge $\bigcup \mathfrak{M}$. Dann gehört x wenigstens einer Menge X_0 des Systems $\mathfrak{M}$ an, so daß wegen $X_0 \subseteq Z$ auch $x \in Z$ gilt. Da das für jedes $x \in \bigcup \mathfrak{M}$ der Fall ist, gilt $\bigcup \mathfrak{M} \subseteq Z$, wie in (12) behauptet wurde. Wir nehmen schließlich an, V sei eine beliebige Menge, die den Bedingungen

$$(11') \qquad \bigwedge_X (X \in \mathfrak{M} \Rightarrow X \subseteq V),$$

$$(12') \qquad \bigwedge_Z \left(\bigwedge_X (X \in \mathfrak{M} \Rightarrow X \subseteq Z) \Rightarrow V \subseteq Z\right)$$

genügt. Nach (11') ist V eine Menge Z, die der Voraussetzung von (12) genügt, und daher $\bigcup \mathfrak{M} \subseteq V$. Umgekehrt ist wegen (11) die Menge $\bigcup \mathfrak{M}$ eine Menge Z, die der Voraussetzung von (12') genügt, und daher $V \subseteq \bigcup \mathfrak{M}$. Aus beidem zusammen folgt $\bigcup \mathfrak{M} = V$, d. h., $\bigcup \mathfrak{M}$ ist die einzige Menge V, die den Bedingungen (11') und (12') genügt.

Als ein Kuriosum sei erwähnt, daß man formal als Vereinigung des leeren Systems die leere Menge erhält, während sich als Durchschnitt des leeren Systems die Menge E aller Urelemente ergibt (denn jedes Objekt ist in sämtlichen Mengen des leeren Systems enthalten, da das leere System keine Mengen enthält).

In Verallgemeinerung von 1.4.(20) nennt man schließlich ein Mengensystem $\mathfrak{M}$ *disjunkt*, wenn keine zwei verschiedenen Mengen aus $\mathfrak{M}$ ein Element

gemeinsam haben:

(13) $\mathfrak{M}$ *disjunkt* $: \Leftrightarrow \bigwedge_{X_1, X_2} (X_1 \in \mathfrak{M} \wedge X_2 \in \mathfrak{M} \wedge X_1 \neq X_2 \Rightarrow X_1 \cap X_2 = \emptyset)$.

(13) ist in der Tat eine Verallgemeinerung von 1.4.(20), denn es gilt

(14) $M_1 \neq M_2 \Rightarrow (M_1$ *disjunkt zu* $M_2 \Leftrightarrow \{M_1, M_2\}$ *disjunkt*$)$.

Die Voraussetzung $M_1 \neq M_2$ ist hierbei wesentlich: Nach (13) ist nämlich (Implikation mit falscher Prämisse!) jedes Einersystem $\{M\}$ disjunkt, während nach 1.4.(20) eine Menge M nur im Fall $M = \emptyset$ zu sich selbst disjunkt ist. Nach (13) ist auch das leere System disjunkt.

2. Grundbegriffe der Abbildungstheorie

2.1. Einleitung

Ein weiterer, eng mit dem Mengenbegriff verknüpfter Grundbegriff der Mathematik ist der Begriff der Funktion oder Abbildung. Die Bezeichnung „Funktion" geht auf GOTTFRIED WILHELM LEIBNIZ (1646—1717) zurück und fand durch den bedeutenden Schweizer Mathematiker JOHANN BERNOULLI (1667—1748) weite Verbreitung. Bei BERNOULLI und seinem berühmten Schüler LEONHARD EULER (1707—1783) finden wir die Auffassung der Funktion als einer durch einen analytischen Ausdruck gegebenen gesetzmäßigen Abhängigkeit einer veränderlichen Größe y von einer anderen veränderlichen Größe x bzw. einer „curvae quaecumque libero manus ductu descripta" (Kurve, die sich mit der freien Hand zeichnen läßt). Insbesondere die Ergebnisse des französischen Mathematikers JOSEPH FOURIER (1768—1830) über die nach ihm benannten trigonometrischen Reihen, zu denen er im Zusammenhang mit Untersuchungen zur Wärmelehre gelangte, führten zu der Frage, ob die Auffassung von BERNOULLI und EULER nicht zu eng ist. Es ist bemerkenswert, daß auch die Untersuchungen FOURIERS, gerade wegen der großen Allgemeinheit der durch trigonometrische Reihen erzeugbaren Funktionen, auf Bedenken seiner Zeitgenossen stießen, ähnlich wie 75 Jahre später die Untersuchungen CANTORS, die übrigens ebenfalls durch Ergebnisse über trigonometrische Reihen ausgelöst wurden. Im Jahre 1837 legte der in Berlin und später als Nachfolger von GAUSS in Göttingen wirkende P. G. LEJEUNE-DIRICHLET (1805—1859) die Grundlage für unseren heutigen Funktionsbegriff, indem er die Forderung der analytischen Darstellbarkeit rigoros fallen ließ und an eine Funktion nur noch die Forderung stellte, daß durch sie jedem Argumentwert x nach einer gegebenen Vorschrift ein bestimmter Funktionswert y zugeordnet wird. Über die Natur der Vorschrift wurden dabei keinerlei Einschränkungen mehr gemacht: Sie konnte wie bei BERNOULLI und EULER durch einen analytischen Ausdruck gegeben

sein, sie konnte in irgendeinem anderen mathematischen Verfahren bestehen, z. B. wie bei FOURIER in der Berechnung des Grenzwertes einer trigonometrischen Reihe an einer Stelle x ihres Konvergenzbereichs, sie konnte verbal beschrieben sein, wie bei der sogenannten Dirichletschen Funktion, die jeder rationalen Zahl x den Wert 1 und jeder irrationalen Zahl x den Wert 0 zuordnet, sie konnte aber z. B. auch in einer physikalischen Meßvorschrift oder dergleichen bestehen. Dieselbe Idee entwickelte übrigens schon im Jahre 1834 der berühmte russische Mathematiker N. I. LOBATSCHEWSKI (1792—1856), dessen Hauptverdienst die Entdeckung der nichteuklidischen Geometrie ist. In unseren Tagen hat sich gezeigt, daß für manche Zwecke der höheren Analysis und ihrer Anwendungen in der Physik eine nochmalige Verallgemeinerung des Funktionsbegriffs erforderlich ist. Diese verallgemeinerten Funktionen werden heute meist *Distributionen* genannt.

Das besondere Merkmal einer Funktion besteht nach der heutigen Auffassung darin, daß durch sie jedem Element x aus einer gegebenen Menge M, dem Definitionsbereich der Funktion, ein durch x eindeutig bestimmtes Element y aus einer evtl. anderen Menge N zugeordnet wird. Daneben benötigt man in der Mathematik aber auch Zuordnungen, bei denen evtl. manchen Elementen aus M kein und anderen Elementen aus M mehrere Elemente aus N entsprechen. Bis vor kurzem war für solche evtl. mehrdeutigen Zuordnungen im Anschluß an HAUSDORFF die Bezeichnung Abbildung üblich. In der letzten Zeit hat es sich jedoch immermehr eingebürgert, auch eine Abbildung grundsätzlich als eindeutig anzusehen, d. h. das Wort „Abbildung" synonym mit dem Wort „Funktion" zu verwenden. Wir wollen daher im folgenden eine nicht notwendig eindeutige Zuordnung eine Korrespondenz nennen; in der Literatur wird hierfür vielfach auch die Bezeichnung Relation verwendet. Die Abbildungen oder Funktionen sind dann spezielle Korrespondenzen, nämlich sogenannte eindeutige Korrespondenzen.

2.2. Geordnetes Paar und Produktmenge

Zur Präzisierung der in der Einleitung genannten Begriffe benötigen wir zunächst den Begriff des *geordneten Paares* (a, b) aus gegebenen Objekten a und b. Hierunter wollen wir ein durch die Objekte a, b festgelegtes neues (abstraktes) Objekt verstehen, wobei wir lediglich verlangen, daß *geordnete Paare* (a_1, b_1) *und* (a_2, b_2) *genau dann gleich sind, wenn sowohl* $a_1 \doteq a_2$ *als auch* $b_1 = b_2$ *ist:*

$$(1) \qquad (a_1, b_1) \doteq (a_2, b_2) \Leftrightarrow a_1 = a_2 \wedge b_1 = b_2.$$

Hiernach ist es erforderlich, das geordnete Paar (a, b) sorgfältig von der Zweiermenge $\{a, b\}$ zu unterscheiden; denn es ist stets $\{a, b\} = \{b, a\}$, aber natürlich im allgemeinen $(a, b) \neq (b, a)$ (nach (1) ist $(a, b) = (b, a)$ genau dann, wenn $a = b$). In der Literatur finden sich eine ganze Reihe von verschiedenen Definitionen für das geordnete Paar, die allerdings weitgehend technischer Natur sind. Es handelt sich dabei jeweils um die mengentheoretische Konstruktion eines Objektes (a, b) mit der in (1) geforderten Eigenschaft. So kann man z. B. nach dem polnischen Mathematiker C. KURATOWSKI zeigen, daß für Objekte a, b aus demselben Grundbereich E die Zweiermenge $\{\{a, b\}, \{a\}\}$ zweiter Stufe (aber natürlich genauso die Zweiermenge $\{\{a, b\}, \{b\}\}$) die Bedingung (1) erfüllt. Für die Zwecke der Mathematik ist die Festlegung auf eine bestimmte Definition des geordneten Paares ohne Belang; wir benötigen nur das folgende

Prinzip der Paarbildung. *Zu beliebigen Objekten a, b kann das geordnete Paar (a, b) gebildet werden, wobei für diese Bildung die Bedingung* (1) *erfüllt ist.*

Das Objekt a heißt dabei die *erste Komponente* oder das *erste Glied* und das Objekt b die *zweite Komponente* oder das *zweite Glied* des geordneten Paares (a, b). Bei der Paarbildung wollen wir ausdrücklich zulassen, daß die Komponenten eines Paares verschiedenen Grundbereichen angehören, insbesondere kann also z. B. die erste Komponente ein Objekt a eines bestimmten Grundbereichs E und die zweite Komponente eine Menge M von Objekten aus E sein (bei der Definition von KURATOWSKI ist das nicht ohne weiteres möglich).

Sind M_1, M_2 beliebige Mengen (über evtl. unterschiedlichen Grundbereichen E_1, E_2), so definiert man als *Produktmenge* oder *kartesisches Produkt* $M_1 \times M_2$ der Mengen M_1, M_2 die Menge aller geordneten Paare $p = (x_1, x_2)$ mit $x_1 \in M_1$ und $x_2 \in M_2$:

$$(2) \qquad M_1 \times M_2 := \{p : \bigvee_{x_1, x_2} (x_1 \in M_1 \wedge x_2 \in M_2 \wedge p = (x_1, x_2))\},$$

wobei wir für die rechte Seite dieser Definitionsgleichung auch kurz

$$\{(x_1, x_2) : x_1 \in M_1 \wedge x_2 \in M_2\}$$

schreiben wollen.

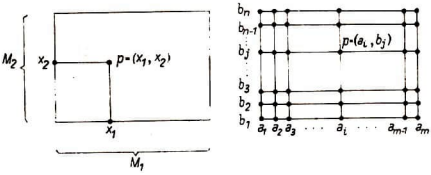


Abb. 3

Man kann sich das Produkt $M_1 \times M_2$ am einfachsten in der aus der analytischen Geometrie bekannten Weise durch ein Rechteck veranschaulichen, dessen Seiten die Mengen M_1, M_2 symbolisieren. Das geordnete Paar $p = (x_1, x_2)$ wird dabei durch den Punkt des Rechtecks dargestellt, dessen erste Koordinate das Element $x_1 \in M_1$ und dessen zweite Koordinate das Element $x_2 \in M_2$ ist. Ist $M_1 = \{a_1, \ldots, a_m\}$ eine Menge aus m Elementen und $M_2 = \{b_1, \ldots, b_n\}$ eine Menge aus n Elementen, so entartet natürlich das Rechteck in ein Punktgitter aus $m \cdot n$ Punkten (Abb. 3).

Wir merken zunächst an, daß im allgemeinen $M_1 \times M_2$ von $M_2 \times M_1$ verschieden, die Bildung der Produktmenge also nicht kommutativ ist. Ebensowenig gilt das Assoziativgesetz. Dagegen gelten bei beliebigem M_1, M_2, N die folgenden Beziehungen:

$$(3) \qquad \begin{aligned} (M_1 \cap M_2) \times N &= (M_1 \times N) \cap (M_2 \times N), \\ N \times (M_1 \cap M_2) &= (N \times M_1) \cap (N \times M_2) \end{aligned}$$

und

$$(4) \qquad \begin{aligned} (M_1 \cup M_2) \times N &= (M_1 \times N) \cup (M_2 \times N), \\ N \times (M_1 \cup M_2) &= (N \times M_1) \cup (N \times M_2), \end{aligned}$$

d. h. *das Mengenprodukt ist rechts- und linksseitig distributiv bzgl. Durchschnitt und Vereinigung* (vgl. 1.4.(11)). Der Beweis z. B. der ersten Regel (4) ergibt sich aus der folgenden für jedes Paar (x, y) geltenden Kette von logischen Äquivalenzen:

$$\begin{aligned} (x, y) \in (M_1 \cup M_2) \times N &\Leftrightarrow x \in M_1 \cup M_2 \wedge y \in N \\ &\Leftrightarrow (x \in M_1 \vee x \in M_2) \wedge y \in N \\ &\Leftrightarrow (x \in M_1 \wedge y \in N) \vee (x \in M_2 \wedge y \in N) \\ &\Leftrightarrow (x, y) \in M_1 \times N \vee (x, y) \in M_2 \times N \\ &\Leftrightarrow (x, y) \in (M_1 \times N) \cup (M_2 \times N). \end{aligned}$$

Man kann sich (3) und (4) leicht am oben erwähnten Rechteckmodell veranschaulichen.

Gemischte Anwendung von (3) liefert

$$(3') \qquad \begin{aligned} &(M_1 \cap M_2) \times (N_1 \cap N_2) \\ &= (M_1 \times N_1) \cap (M_1 \times N_2) \cap (M_2 \times N_1) \cap (M_2 \times N_2). \end{aligned}$$

Hierbei ist nun

$$(M_1 \times N_1) \cap (M_2 \times N_2) = (M_1 \times N_2) \cap (M_2 \times N_1);$$

denn für jedes Paar (x, y) gilt

$$\begin{aligned} &(x, y) \in (M_1 \times N_1) \cap (M_2 \times N_2) \\ &\Leftrightarrow (x, y) \in M_1 \times N_1 \wedge (x, y) \in M_2 \times N_2 \\ &\Leftrightarrow x \in M_1 \wedge y \in N_1 \wedge x \in M_2 \wedge y \in N_2 \\ &\Leftrightarrow (x, y) \in M_1 \times N_2 \wedge (x, y) \in M_2 \times N_1 \\ &\Leftrightarrow (x, y) \in (M_1 \times N_2) \cap (M_2 \times N_1). \end{aligned}$$

Folglich ist nach 1.4.(13)

$$(M_1 \times N_1) \cap (M_1 \times N_2) \cap (M_2 \times N_1) \cap (M_2 \times N_2) = (M_1 \times N_1) \cap (M_2 \times N_2),$$

und hieraus und aus (3') folgt

(5) $$(M_1 \cap M_2) \times (N_1 \cap N_2) = (M_1 \times N_1) \cap (M_2 \times N_2).$$

Setzt man in (5) $N_1 = N_2 = N$, so gelangt man natürlich sofort zu (3) zurück. Auch (5) kann man sich leicht am Rechteckmodell veranschaulichen, mittels dessen man zugleich leicht erkennen kann, daß eine analoge Verschärfung von (4) nicht möglich ist (Abb. 4).

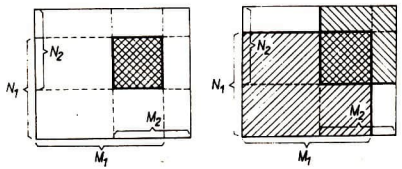


Abb. 4

Es seien noch folgende Rechengesetze genannt, deren Beweise keine Schwierigkeiten bieten, und die wir daher dem Leser als Übungsaufgabe überlassen wollen:

(6) $$(M_1 \setminus M_2) \times N = (M_1 \times N) \setminus (M_2 \times N),$$
$$N \times (M_1 \setminus M_2) = (N \times M_1) \setminus (N \times M_2);$$

(7) $$M_1 \subseteq M_2 \Rightarrow M_1 \times N \subseteq M_2 \times N,$$
$$M_1 \subseteq M_2 \Rightarrow N \times M_1 \subseteq N \times M_2;$$

(8) $$M_1 \times M_2 = \emptyset \Leftrightarrow M_1 = \emptyset \vee M_2 = \emptyset;$$

(9) $$M_1 \times N \subseteq M_2 \times N \wedge N \neq \emptyset \Rightarrow M_1 \subseteq M_2,$$
$$N \times M_1 \subseteq N \times M_2 \wedge N \neq \emptyset \Rightarrow M_1 \subseteq M_2.$$

Durch zweimalige Anwendung von (9) erhalten wir schließlich noch die folgenden wichtigen Kürzungsregeln:

(10) $$M_1 \times N = M_2 \times N \wedge N \neq \emptyset \Rightarrow M_1 = M_2,$$
$$N \times M_1 = N \times M_2 \wedge N \neq \emptyset \Rightarrow M_1 = M_2.$$

Wir sehen also, daß für das Mengenprodukt viele vom Zahlenrechnen bekannte Rechengesetze gelten, aber ebenso eine Reihe wichtiger Rechengesetze, wie insbesondere das Kommutativ- und Assoziativgesetz verletzt sind.

Mit Hilfe des geordneten Paares kann man das (*geordnete*) ***Tripel*** aus gegebenen Objekten a, b, c durch

(11) $\quad (a, b, c) := ((a, b), c)$

definieren. Auf Grund von (1) gilt dann

$$\begin{aligned}(a_1, b_1, c_1) = (a_2, b_2, c_2) &\Leftrightarrow ((a_1, b_1), c_1) = ((a_2, b_2), c_2)\\ &\Leftrightarrow (a_1, b_1) = (a_2, b_2) \wedge c_1 = c_2\\ &\Leftrightarrow a_1 = a_2 \wedge b_1 = b_2 \wedge c_1 = c_2,\end{aligned}$$

d. h.

(12) $\quad (a_1, b_1, c_1) = (a_2, b_2, c_2) \Leftrightarrow a_1 = a_2 \wedge b_1 = b_2 \wedge c_1 = c_2.$

Bezeichnen wir a als die *erste*, b als die *zweite* und c als die *dritte Komponente* des Tripels (a, b, c), so besagt (12), daß *Tripel* (a_1, b_1, c_1) *und* (a_2, b_2, c_2) *genau dann gleich sind, wenn sie komponentenweise übereinstimmen*. Die Menge aller Tripel (x_1, x_2, x_3) mit $x_1 \in M_1, x_2 \in M_2, x_3 \in M_3$ bezeichnet man mit $M_1 \times M_2 \times M_3$:

(13) $\quad M_1 \times M_2 \times M_3 := \{(x_1, x_2, x_3) : x_1 \in M_1 \wedge x_2 \in M_2 \wedge x_3 \in M_3\}.$

Aus (11) folgt unmittelbar

(13′) $\quad M_1 \times M_2 \times M_3 = (M_1 \times M_2) \times M_3.$

Wir merken an, daß es beim Tripel — analog wie beim geordneten Paar — wiederum nur darauf ankommt, daß die Bedingung (12) erfüllt ist. Daher hätten wir das Tripel (x, y, z) ebensogut durch $(x, (y, z))$ definieren können; bei dieser Definition wäre natürlich $M_1 \times M_2 \times M_3 = M_1 \times (M_2 \times M_3)$.

Allgemein definieren wir für gegebene Objekte $a_1, \ldots, a_n$ (n natürliche Zahl $\geqq 2$) das *n-Tupel* $(a_1, \ldots, a_n)$ induktiv durch

(14) $\quad (a_1, \ldots, a_n) := ((a_1, \ldots, a_{n-1}), a_n),$

wobei als 1-Tupel (a_1) einfach das Objekt a_1 selbst zu nehmen ist. Dann gilt bei beliebigem n

(15) $\quad (a_1, \ldots, a_n) = (b_1, \ldots, b_n) \Leftrightarrow a_1 = b_1 \wedge \cdots \wedge a_n = b_n.$

Bezeichnen wir $a_1, \ldots, a_n$ als die *Komponenten* oder *Glieder* des n-Tupels $(a_1, \ldots, a_n)$ (genauer a_i $(1 \leqq i \leqq n)$ als die *i-te Komponente* von $(a_1, \ldots, a_n)$), so sind also allgemein *n-Tupel* $(a_1, \ldots, a_n)$ *und* $(b_1, \ldots, b_n)$ *genau dann gleich, wenn sie komponentenweise (gliedweise) übereinstimmen*, und das ist die wirklich wesentliche Eigenschaft der n-Tupel. Die Menge aller n-Tupel $(x_1, \ldots, x_n)$ mit $x_1 \in M_1, \ldots, x_n \in M_n$ wird mit $M_1 \times \cdots \times M_n$ bezeichnet. Im Fall $M_1 = \cdots = M_n = M$ wird für $M_1 \times \cdots \times M_n$ auch kurz M^n geschrieben. In diesem Sinne ist also z. B. M^2 die Menge aller geordneten Paare (x, y) mit $x \in M$ und $y \in M$.

2.3. Korrespondenzen

Unter einer *Korrespondenz aus einer Menge M in eine Menge N* wollen wir eine beliebige Teilmenge F der Produktmenge $M \times N$ verstehen:

(1) $\quad F$ *Korrespondenz aus M in N* $: \Leftrightarrow F \subseteq M \times N$.

Eine Korrespondenz aus M in N ist also eine beliebige Menge von geordneten Paaren (x, y), deren erste Komponenten sämtlich zu M und deren zweite Komponenten sämtlich zu N gehören. Korrespondenzen wollen wir bei den folgenden allgemeinen Betrachtungen vorwiegend mit F, G usw. bezeichnen. Ist F eine Korrespondenz aus M in N und $(x, y) \in F$, so nennen wir y ein *Bild* von x und x ein *Urbild* von y bei F und sagen auch, daß *durch F dem Element x das Element y zugeordnet wird*. Zur Abkürzung verwenden wir hierfür die Schreibweise xFy:

(2) $\quad xFy :\Leftrightarrow (x, y) \in F$.

Die Menge aller Bilder y eines gegebenen Elements $x \in M$ heißt das *volle Bild* von x bei der Korrespondenz F und werde im folgenden mit $B_F(x)$ bezeichnet:

(3) $\quad B_F(x) := \{y : xFy\}$.

Entsprechend heißt die Menge aller Urbilder x eines gegebenen Elements $y \in N$ das *volle Urbild* von y bei F und werde mit $U_F(y)$ bezeichnet:

(4) $\quad U_F(y) := \{x : xFy\}$.

Da es in einer gegebenen Teilmenge F der Produktmenge $M \times N$ keineswegs zu jedem $x \in M$ ein Paar (x, y) zu geben braucht, das x als erste Komponente hat, kann das volle Bild $B_F(x)$ eines Elements $x \in M$ durchaus leer sein. In diesem Fall wird dem Element $x \in M$ durch die Korrespondenz F kein Bild zugeordnet. Die Menge derjenigen $x \in M$, für die $B_F(x) \neq \emptyset$ gilt, nennt man den *Definitionsbereich* der Korrespondenz F, manchmal auch den *Urbildbereich*, den *Vorbereich* oder den *Argumentbereich*. Der Definitionsbereich einer Korrespondenz F soll im folgenden allgemein mit $D(F)$ bezeichnet werden:

(5) $\quad D(F) := \{x : B_F(x) \neq \emptyset\}\ (= \{x : \bigvee_y xFy\})$.

Entsprechend nennt man die Menge aller derjenigen $y \in N$, deren volles Urbild nicht leer ist, die also Bild wenigstens eines Elements $x \in M$ sind, den *Wertebereich* der Korrespondenz F, manchmal auch den *Wertevorrat*, den *Bildbereich*, den *Nachbereich* oder den *Gegenbereich*. Der Wertebereich einer Korrespondenz F soll im folgenden allgemein mit $W(F)$ bezeichnet werden:

(6) $\quad W(F) := \{y : U_F(y) \neq \emptyset\}\ (= \{y : \bigvee_x xFy\})$.

In der englischsprachigen Literatur sind für den Definitions- bzw. Wertebereich die Bezeichnungen *domain* bzw. *image* verbreitet, auf Grund dessen man für diese Mengen auch international vielfach die Abkürzungen dom (F) bzw. im (F) verwendet.

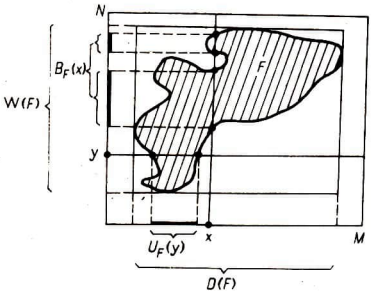


Abb. 5

Veranschaulicht man sich die Produktmenge $M \times N$ in der oben angegebenen Weise durch ein Rechteck mit den Seiten M, N, so erscheint jede Korrespondenz F aus M in N als eine gewisse Menge von Punkten dieses Rechtecks. Die Mengen $B_F(x)$, $U_F(y)$, $D(F)$ und $W(F)$ haben dann die in Abb. 5 dargestellte einfache Bedeutung. Ist $M = \{a_1, \ldots, a_m\}$ eine Menge aus m Elementen und $N = \{b_1, \ldots, b_n\}$ eine solche aus n Elementen, so kann man z. B. die Punkte des Gitters $M \times N$, die zu einer Korrespondenz F gehören, durch eine 1 und die übrigen durch eine 0 kennzeichnen. Man erhält auf diese Weise eine Charakterisierung von F durch eine *Tabelle* oder *Matrix*, wobei man allerdings meistens die Elemente aus M den Zeilen und denen von N die Spalten zuordnet:

F	b_1	$\ldots$	b_j	$\ldots$	b_n
a_1	α_{11}	$\ldots$	α_{1j}	$\ldots$	α_{1n}
$\cdot$	$\cdot$		$\cdot$		$\cdot$
$\cdot$	$\cdot$		$\cdot$		$\cdot$
$\cdot$	$\cdot$		$\cdot$		$\cdot$
a_i	α_{i1}	$\ldots$	α_{ij}	$\ldots$	α_{in}
$\cdot$	$\cdot$		$\cdot$		$\cdot$
$\cdot$	$\cdot$		$\cdot$		$\cdot$
$\cdot$	$\cdot$		$\cdot$		$\cdot$
a_m	α_{m1}	$\ldots$	α_{mj}	$\ldots$	α_{mn}

Dabei gilt

$$\alpha_{ij} = \begin{cases} 1, \text{ falls } (a_i, b_j) \in F, \\ 0, \text{ falls } (a_i, b_j) \notin F. \end{cases}$$

Für jede Korrespondenz F aus M in N ist offenbar $D(F) \subseteq M$ und $W(F) \subseteq N$. Ist speziell $D(F) = M$, so heißt F eine Korrespondenz *von M in N*; ist $W(F) = N$, so heißt F eine Korrespondenz *aus M auf N*; ist schließlich sowohl $D(F) = M$ als auch $W(F) = N$, so heißt F eine Korrespondenz *von M auf N*:

M \ N	in	auf
aus	$D(F) \subseteq M$ $W(F) \subseteq N$	$D(F) \subseteq M$ $W(F) = N$
von	$D(F) = M$ $W(F) \subseteq N$	$D(F) = M$ $W(F) = N$

Wegen $F \subseteq D(F) \times W(F)$ (vgl. Abb. 5) ist jede Korrespondenz aus M in N eine Korrespondenz von $D(F)$ auf $W(F)$.

Als spezielle Mengen *sind Korrespondenzen F, G aus M in N genau dann gleich, wenn sie dieselben geordneten Paare als Element enthalten:*

$$(7) \qquad F = G \Leftrightarrow \bigwedge_{x,y} ((x, y) \in F \Leftrightarrow (x, y) \in G).$$

Da nach (3) $(x, y) \in F$ logisch äquivalent ist mit $y \in B_F(x)$ und analog $(x, y) \in G$ mit $y \in B_G(x)$ und da die Beziehung $\bigwedge_y (y \in B_F(x) \Leftrightarrow y \in B_G(x))$ gerade besagt, daß die Mengen $B_F(x)$ und $B_G(x)$ gleich sind, folgt aus (7)

$$(8) \qquad F = G \Leftrightarrow \bigwedge_x B_F(x) = B_G(x),$$

d. h., *Korrespondenzen sind genau dann gleich, wenn bei beliebigem x die Bildmengen $B_F(x)$ und $B_G(x)$ übereinstimmen.* Analog erhält man

$$(8') \qquad F = G \Leftrightarrow \bigwedge_y U_F(y) = U_G(y).$$

Entsprechend gilt nach Definition der Inklusion

$$(9) \qquad F \subseteq G \Leftrightarrow \bigwedge_{x,y} ((x, y) \in F \Rightarrow (x, y) \in G).$$

Man nennt in diesem Fall die Korrespondenz F eine *Einschränkung* der Korrespondenz G und G eine *Erweiterung* oder *Fortsetzung* von F. In Analogie zu (8) bzw. (8') gilt

$$(10) \qquad F \subseteq G \Leftrightarrow \bigwedge_x B_F(x) \subseteq B_G(x);$$

$$(10') \qquad F \subseteq G \Leftrightarrow \bigwedge_y U_F(y) \subseteq U_G(y).$$

Für eine beliebige Korrespondenz F versteht man unter der zu F *inversen Korrespondenz* F^{-1} die Menge aller und nur der geordneten Paare (y, x), für die $(x, y) \in F$ gilt:

$$(11) \qquad F^{-1} := \{(y, x) : (x, y) \in F\},$$

wobei $\{(y, x) : (x, y) \in F\}$ eine Abkürzung für

$$\{p : \bigvee_{x,y} ((x, y) \in F \wedge p = (y, x))\}$$

ist. Die Korrespondenz F^{-1} wird auch als *Umkehrkorrespondenz* zu F bezeichnet. Ist F eine Korrespondenz aus M in N, so ist F^{-1} eine Korrespondenz aus N in M, und zwar ordnet die Korrespondenz F^{-1} einem Element $y \in N$ gerade diejenigen $x \in M$ als Bild zu, denen bei der Korrespondenz F das Element y als Bild zugeordnet ist, die also Urbild von y bei der Korrespondenz F sind:

$$(12) \qquad yF^{-1}x \Leftrightarrow xFy.$$

Aus (12) folgt, daß bei beliebigem x gilt:

$$x \in B_{F^{-1}}(y) \Leftrightarrow x \in U_F(y),$$

d. h.

$$(13) \qquad B_{F^{-1}}(y) = U_F(y).$$

Analog erhält man

$$(13') \qquad U_{F^{-1}}(x) = B_F(x),$$

und aus (13) bzw. (13') folgt

$$(14) \qquad D(F^{-1}) = W(F), \; W(F^{-1}) = D(F).$$

Also ist F^{-1} genau dann eine Korrespondenz von N in M, wenn F eine Korrespondenz aus M auf N ist, usw. Aus der Definition der inversen Korrespondenz folgt schließlich auf Grund von (7) unmittelbar, daß *die zu* F^{-1} *inverse Korrespondenz gleich der ursprünglichen Korrespondenz* F *ist:*

$$(15) \qquad (F^{-1})^{-1} = F.$$

Es sei schließlich F eine Korrespondenz aus M in N und G eine Korrespondenz aus N in P. Unter der *Verkettung*, dem *Produkt* oder der *Hintereinanderausführung* $G \circ F$ (gelesen etwa: G nach F) versteht man die Korrespondenz aus M in P, die einem Element $x \in M$ alle diejenigen $z \in P$ zuordnet, die Bild wenigstens eines Elements y aus $B_F(x)$ bei der Korrespondenz G sind:

$$(16) \qquad (x, z) \in G \circ F :\Leftrightarrow \bigvee_{y} ((x, y) \in F \wedge (y, z) \in G),$$

d. h.

$$(16') \qquad G \circ F := \{(x, z) : \bigvee_{y} ((x, y) \in F \wedge (y, z) \in G)\}.$$

Wir zeigen, daß *die Verkettung von Korrespondenzen assoziativ ist:*

(17) $F_1 \circ (F_2 \circ F_3) = (F_1 \circ F_2) \circ F_3$,

und ferner das folgende merkwürdige Inversionsgesetz gilt:

(18) $(G \circ F)^{-1} = F^{-1} \circ G^{-1}$.

Ist nämlich $(x, z) \in F_1 \circ (F_2 \circ F_3)$, so existiert nach (16) ein y_1 mit $(x, y_1) \in F_2 \circ F_3$ und $(y_1, z) \in F_1$ und wiederum nach (16) ein y_2 mit $(x, y_2) \in F_3$ und $(y_2, y_1) \in F_2$. Dann ist aber $(y_2, z) \in F_1 \circ F_2$ und $(x, z) \in (F_1 \circ F_2) \circ F_3$. Also ist $F_1 \circ (F_2 \circ F_3) \subseteq (F_1 \circ F_2) \circ F_3$. Entsprechend zeigt man, daß $(F_1 \circ F_2) \circ F_3 \subseteq F_1 \circ (F_2 \circ F_3)$ ist, und damit ist (17) bewiesen. Ist $(z, x) \in (G \circ F)^{-1}$, so ist nach (11) $(x, z) \in G \circ F$, und folglich gibt es nach (16) ein y mit $(x, y) \in F$ und $(y, z) \in G$. Dann ist aber $(z, y) \in G^{-1}$ und $(y, x) \in F^{-1}$ und folglich $(z, x) \in F^{-1} \circ G^{-1}$. Mithin ist $(G \circ F)^{-1} \subseteq F^{-1} \circ G^{-1}$. Entsprechend zeigt man, daß $F^{-1} \circ G^{-1} \subseteq (G \circ F)^{-1}$, und damit ist (18) bewiesen.

Ein besonders wichtiges Verfahren zur Charakterisierung einer Korrespondenz besteht darin, daß man eine für die Objekte x eines Grundbereichs E_1 und die Objekte y eines (evtl. anderen) Grundbereichs E_2 definierte Eigenschaft oder Aussage $H(x, y)$ betrachtet und im Grundbereich $E_1 \times E_2$ mittels des Mengenbildungsprinzips die Menge aller derjenigen Paare $p = (x, y)$ bildet, für die die Eigenschaft $H(x, y)$ erfüllt ist bzw. auf die die Aussage $H(x, y)$ zutrifft. Die auf Grund des Extensionalitätsprinzips eindeutig bestimmte Menge aller dieser Paare bezeichnet man naturgemäß durch

$$\{p : \bigvee_{x,y} (p = (x, y) \wedge H(x, y))\} \quad \text{oder kurz} \quad \{(x, y) : H(x, y)\}.$$

Diese Menge F kann aufgefaßt werden als Korrespondenz aus E_1 in E_2 oder allgemeiner aus M in N, sofern $D(F) \subseteq M \subseteq E_1$ und $W(F) \subseteq N \subseteq E_2$ ist, wobei $D(F) = \{x : \bigvee_y H(x, y)\}$ und $W(F) = \{y : \bigvee_x H(x, y)\}$ gilt. In Analogie zu 1.3. (2) gilt:

(19) $(x_0, y_0) \in \{(x, y) : H(x, y)\} \Leftrightarrow H(x_0, y_0)$,

und in Analogie zu 1.3. (3) erhalten wir:

(20) $\{(x, y) : H_1(x, y)\} = \{(x, y) : H_2(x, y)\} \Leftrightarrow \bigwedge_{x,y} (H_1(x, y) \Leftrightarrow H_2(x, y))$.

Diese Art der Charakterisierung einer Korrespondenz dürfte wohl gemeint sein, wenn man davon spricht, daß eine Korrespondenz durch eine Zuordnungsvorschrift definiert ist. Dabei ist dann aber zu beachten, daß verschiedene Zuordnungsvorschriften $H_1(x, y)$ und $H_2(x, y)$ dieselbe Korrespon-

denz definieren können, wenn sie nämlich umfangsgleich sind, d. h., wenn $\bigwedge_{x,y} (H_1(x, y) \Leftrightarrow H_2(x, y))$ gilt.

Wir wollen die in diesem Abschnitt eingeführten allgemeinen Begriffsbildungen noch an einem Beispiel aus dem täglichen Leben illustrieren. Dabei sei $M = \{a_1, \ldots, a_m\}$ eine Menge von m Betrieben und $N = \{b_1, \ldots, b_n\}$ eine Menge von n Arten von Produkten. Die Aussage „x produziert y" definiert dann eine bestimmte Korrespondenz F aus M in N, bei der jedem Betrieb a_μ aus M diejenigen Produkte b_ν aus N als Bild zugeordnet sind, die vom Betrieb a_μ produziert werden. So beschreibt z. B. für $m = 4$, $n = 5$ die Matrix

F	b_1	b_2	b_3	b_4	b_5
a_1	1	0	0	1	0
a_2	1	0	1	1	0
a_3	0	0	1	0	1
a_4	1	0	1	1	1

den Fall, daß der Betrieb a_1 die Produkte b_1 und b_4, nicht aber die Produkte b_2, b_3, b_5 produziert, daß a_2 die Produkte b_1, b_3, b_4, nicht aber b_2, b_5 produziert usw. Für einen gegebenen Betrieb a_μ aus M ist die Bildmenge $B_F(a_\mu)$ die Menge aller derjenigen Produkte aus N, die durch a_μ produziert werden, während für ein gegebenes Produkt b_ν aus N die Urbildmenge $U_F(b_\nu)$ die Menge aller derjenigen Betriebe aus M ist, die b_ν produzieren. Im betrachteten Spezialfall ist beispielsweise $B_F(a_3) = \{b_3, b_5\}$, $U_F(b_1) = \{a_1, a_2, a_4\}$, $U_F(b_2) = \emptyset$. Der Definitionsbereich $D(F)$ ist die Menge aller derjenigen Betriebe aus M, die wenigstens eines der Produkte aus N produzieren, während der Wertebereich $W(F)$ die Menge aller derjenigen Produkte aus N ist, die von wenigstens einem Betrieb aus M produziert werden. Im betrachteten Spezialfall ist $D(F) = M$ und $W(F) = \{b_1, b_3, b_4, b_5\} \subset N$, es handelt sich also um eine Korrespondenz von M in N (von M auf $\{b_1, b_3, b_4, b_5\}$). Deuten wir genauer F als die z. B. im Planjahr 1972 durch „x produziert y" definierte Korrespondenz und F' als die durch dieselbe Aussage im Planjahr 1973 definierte Korrespondenz, so besagt die Gleichung $F = F'$ (vgl. (8)), daß im Jahre 1973 jeder Betrieb aus M dieselben Produkte aus N produziert wie im Jahre 1972, oder auch (vgl. (9)), daß 1973 jedes Produkt aus N von denselben Betrieben aus M produziert wird wie 1972. Die Inklusion $F \subseteq F'$ besagt demgegenüber, daß 1973 kein Betrieb aus M weniger Produkte aus N produziert als 1972. Die zu F inverse Korrespondenz F^{-1} ordnet jedem Produkt b_ν aus N diejenigen Betriebe a_μ aus M als Bild zu, die das Produkt b_ν produzieren. Im betrachteten Spezialfall wird also F^{-1} durch die Matrix

F^{-1}	a_1	a_2	a_3	a_4
b_1	1	1	0	1
b_2	0	0	0	0
b_3	0	1	1	1
b_4	1	1	0	1
b_5	0	0	1	1

beschrieben. Zur Illustration der Verkettung sei schließlich $P = \{c_1, \ldots, c_p\}$ eine Menge von p Sorten von Rohstoffen, und es bedeute G die durch die Aussage „die Produktion von y erfordert z" definierte Korrespondenz aus N in P; sie ordnet einem beliebigen Produkt b_ν aus N gerade diejenigen Rohstoffe c_π aus P als Bild zu, die zur Produktion von b_ν erforderlich sind. Mit $p = 4$ sei beispielsweise

G	c_1	c_2	c_3	c_4
b_1	1	0	0	1
b_2	0	1	0	1
b_3	1	0	1	0
b_4	1	0	0	0
b_5	0	0	0	0

Die Korrespondenz $G \circ F$ ordnet dann, wie man leicht nachprüft, jedem Betrieb a_μ aus M diejenigen Rohstoffe c_π aus P als Bild zu, die er zur Produktion der Produkte aus N braucht. Im betrachteten Beispiel wird $G \circ F$ durch die folgende Matrix beschrieben:

$G \circ F$	c_1	c_2	c_3	c_4
a_1	1	0	0	1
a_2	1	0	1	1
a_3	1	0	1	0
a_4	1	0	1	1

Wir empfehlen, sich am betrachteten Beispiel die Beziehung (18) zu veranschaulichen.

2.4. Abbildungen und Funktionen

Wichtigster Spezialfall der Korrespondenzen sind die sogenannten eindeutigen Korrespondenzen. Dabei heißt eine Korrespondenz F aus M in N *eindeutig*, wenn sie jedem $x \in M$ höchstens ein Element $y \in N$ als Bild zuordnet:

(1) $$F \text{ eindeutig} :\Leftrightarrow \bigwedge_{x,y_1,y_2} (xFy_1 \wedge xFy_2 \Rightarrow y_1 = y_2).$$

Da eine Korrespondenz F genau den Elementen $x \in D(F)$ wenigstens ein Element als Bild zuordnet, sind also die eindeutigen Korrespondenzen dadurch charakterisiert, daß durch sie jedem Element $x \in D(F)$ genau ein Element y als Bild zugeordnet ist. Dieses durch x eindeutig bestimmte y heißt das *Bild* oder der *Wert* von x bei der Korrespondenz F und wird üblicherweise mit $F(x)$ bezeichnet.

Veranschaulicht man sich die Korrespondenzen aus M in N in der in 2.3 beschriebenen Weise als Teilmengen eines Rechtecks mit den Seiten M, N,

so erscheinen die eindeutigen Korrespondenzen als diejenigen Teilmengen, die schlicht über der Menge M liegen, d. h., bei denen keine zwei verschiedenen Punkte aus F denselben Punkt $x \in M$ als Abszisse haben (Abb. 6). Für einen gegebenen Punkt $x \in D(F)$ ist $F(x)$ derjenige Punkt y der Ordinate N, für den $(x, y) \in F$ gilt. Charakterisiert man im Fall $M = \{a_1, \ldots, a_m\}$, $N = \{b_1, \ldots, b_n\}$ die Korrespondenzen aus M in N in der in 2.3. beschriebenen Weise durch 0, 1-Matrizen mit m Zeilen und n Spalten, so spiegelt sich die Eindeutigkeit einer Korrespondenz aus M in N darin wider, daß in der zugehörigen Matrix in jeder Zeile höchstens einmal die Ziffer 1 auftritt (vorausgesetzt natürlich, daß die Elemente $b_1, \ldots, b_n$ paarweise verschieden sind).

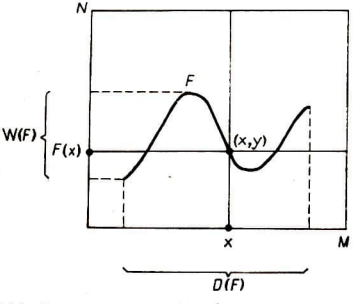


Abb. 6

Wir weisen darauf hin, daß das Bild $F(x)$ sorgfältig von der Bildmenge $B_F(x)$ zu unterscheiden ist. Es gilt aber

(2) $\quad F$ *eindeutig* $\wedge\, x \in D(F) \Rightarrow B_F(x) = \{F(x)\}$.

Aus der Definition des Wertes $F(x)$ folgt ferner

(3) $\quad F$ *eindeutig* $\Rightarrow F = \{(x, y) : x \in D(F) \wedge y = F(x)\}$

oder kurz

(3′) $\quad F$ *eindeutig* $\Rightarrow F = \{(x, F(x)) : x \in D(F)\}$.

Außerdem erhält man leicht

(4) $\quad F$ *eindeutig* $\Rightarrow W(F) = \{y : \bigvee_x (x \in D(F) \wedge y = F(x))\}$

oder kurz

(4′) $\quad F$ *eindeutig* $\Rightarrow W(F) = \{F(x) : x \in D(F)\}$.

Eine Korrespondenz F, die nicht eindeutig ist, bei der also für wenigstens ein $x \in D(F)$ die Bildmenge $B_F(x)$ keine Einermenge ist, heißt *mehrdeutig*.

Die eindeutigen Korrespondenzen aus M in N werden auch (*eindeutige*) *Abbildungen* oder *Funktionen* aus M in N genannt:

(5) $\quad f$ *Abbildung* (*Funktion*) *aus* M *in* $N : \Leftrightarrow f \subseteq M \times N \wedge f$ *eindeutig*

(zur typographischen Unterscheidung von beliebigen Korrespondenzen bezeichnen wir bei den folgenden allgemeinen Betrachtungen Abbildungen vorwiegend mit f, g usw.). Eine Abbildung (Funktion) *von* M *in* N ist dann eine Abbildung aus M in N, deren Definitionsbereich die Menge M ist. Für „f ist Abbildung von M in N" ist heute die Bezeichnung $f : M \to N$ oder auch $M \xrightarrow{f} N$ weit verbreitet:

$$f : M \to N : \Leftrightarrow f \subseteq M \times N \wedge f \text{ eindeutig} \wedge D(f) = M.$$

Die Menge aller Abbildungen von einer Menge M in eine Menge N wird häufig mit N^M bezeichnet:

(6) $\quad N^M := \{f : f : M \to N\}.$

Eine Abbildung *aus* M *auf* N ist entsprechend wie bei Korrespondenzen eine Abbildung aus M in N, deren Wertebereich die Menge N ist, und eine Abbildung *von* M *auf* N ist eine Abbildung aus M in N, deren Definitionsbereich gleich M und deren Wertebereich gleich N ist. Um auszudrücken, daß der Wertebereich einer Abbildung f aus M in N gleich N, also f eine Abbildung aus M auf N ist, sagt man heute vielfach auch, die Abbildung f sei *surjektiv* oder eine *Surjektion* (die Vorsilbe sur (= auf) kommt aus dem Französischen).

Als spezielle Korrespondenzen sind nach 2.3. (8) Abbildungen f, g aus M in N genau dann gleich, wenn bei beliebigem x die Bildmengen $B_f(x)$ und $B_g(x)$ übereinstimmen. Bei einer Abbildung f ist nun aber die Bildmenge $B_f(x)$ entweder leer (wenn nämlich $x \notin D(f)$) oder gleich der Einermenge $\{f(x)\}$ (wenn $x \in D(f)$). Folglich gilt für beliebige Abbildungen (Funktionen) f, g

(7) $\quad f = g \Leftrightarrow D(f) = D(g) \wedge \bigwedge_x (x \in D(f) \Rightarrow f(x) = g(x));$

denn die Gleichung $\{f(x)\} = \{g(x)\}$ ist ja äquivalent mit $f(x) = g(x)$. Sind speziell f, g Abbildungen von M in N, so ist $D(f) = D(g) = M$ und (7) vereinfacht sich zu

(7′) $\quad f = g \Leftrightarrow \bigwedge_x (x \in M \Rightarrow f(x) = g(x)).$

Im allgemeinen Fall ist die Bedingung $D(f) = D(g)$ dagegen wesentlich, was leider manchmal übersehen wird. Entsprechend erhält man aus 2.3. (9), daß

eine Abbildung g genau dann eine Fortsetzung einer Abbildung f (f Einschränkung von g) ist, wenn der Definitionsbereich von g den Definitionsbereich von f umfaßt und für alle x aus dem gemeinsamen Definitionsbereich $D(f)$ die Werte $f(x)$ und $g(x)$ übereinstimmen:

$$(8) \qquad f \subseteq g \Leftrightarrow D(f) \subseteq D(g) \wedge \bigwedge_x (x \in D(f) \Rightarrow f(x) = g(x)).$$

Hieraus folgt leicht, daß es für eine gegebene Abbildung g von M in N und für eine gegebene Teilmenge X von M eine und nur eine Abbildung f von X in N mit $f \subseteq g$ gibt, nämlich $f = \{(x, g(x)) : x \in X\}$. Diese Abbildung f nennt man die *Einschränkung* (manchmal auch *Beschränkung*) *von g auf X*; sie wird heute meistens mit $g \mid X$ (gelesen: g auf X) bezeichnet:

$$(9) \qquad g \mid X := \{(x, g(x)) : x \in X\}.$$

Es ist unmittelbar klar, daß die Umkehrkorrespondenz f^{-1} einer Abbildung f mehrdeutig sein kann, also keine (eindeutige) Abbildung zu sein braucht. Andererseits gibt es durchaus mehrdeutige Korrespondenzen F, deren Umkehrkorrespondenz F^{-1} eindeutig, also eine Abbildung ist; nach 2.3. (15) sind das genau diejenigen mehrdeutigen Korrespondenzen, die Umkehrkorrespondenzen von Abbildungen sind. Allgemein heißt eine Korrespondenz F aus M in N, deren inverse Korrespondenz F^{-1} eindeutig ist, eine *eindeutig umkehrbare* Korrespondenz (die in der Literatur meistens anzutreffende Bezeichnung „umkehrbar eindeutig" ist als sprachlich falsch zu verwerfen):

$$(10) \qquad F \textit{ eindeutig umkehrbar} :\Leftrightarrow F^{-1} \textit{ eindeutig}.$$

Insbesondere ist eine *eindeutig umkehrbare Abbildung* f aus M in N eine solche Abbildung aus M in N, deren Umkehrkorrespondenz eine Abbildung aus N in M ist. In diesem Fall heißt f^{-1} auch die zu f *inverse Abbildung* oder die *Umkehrabbildung* zu f. Die eindeutig umkehrbaren Abbildungen werden allgemein auch als *eineindeutige Abbildungen* (kurz: 1-1-Abbildung) bezeichnet:

$$(11) \qquad f \textit{ 1-1-Abbildung aus } M \textit{ in } N$$
$$:\Leftrightarrow f \subseteq M \times N \wedge f \textit{ eindeutig} \wedge f^{-1} \textit{ eindeutig}.$$

Auf Grund von 2.3. (15) ist klar, daß die Umkehrabbildung einer 1-1-Abbildung aus M in N eine 1-1-Abbildung aus N in M ist:

$$(12) \qquad f \textit{ 1-1-Abbildung aus } M \textit{ in } N \Rightarrow f^{-1} \textit{ 1-1-Abbildung aus } N \textit{ in } M.$$

Außerdem gilt

$$(13) \qquad f \textit{ 1-1-Abbildung aus } M \textit{ in } N$$
$$\Rightarrow \bigwedge_x (x \in D(f) \Rightarrow f^{-1}(f(x)) = x) \wedge \bigwedge_y (y \in W(f) \Rightarrow f(f^{-1}(y)) = y).$$

Für „f ist eine eindeutige Abbildung von M in N" sagt man heute vielfach, daß f *injektiv* oder eine *Injektion* ist. Eine Abbildung, die sowohl injektiv als auch surjektiv, die also eine 1-1-Abbildung von M auf N ist, nennt man dann auch *bijektiv* oder eine *Bijektion*.

Ist f eine Abbildung von M in N und g eine Abbildung von N in P (auf diesen Fall wollen wir uns hier beschränken), *so ist $g \circ f$ eine Abbildung von M in P, und bei beliebigem $x \in M$ gilt $(g \circ f)(x) = g(f(x))$*

$$(14) \qquad f: M \to N \wedge g: N \to P \Rightarrow g \circ f: M \to P \wedge \bigwedge_x (x \in M \Rightarrow (g \circ f)(x) = g(f(x))).$$

Beweis. Nach Definition 2.3. (16) der Verkettung ist zunächst klar, daß $g \circ f$ eine Korrespondenz aus M in P ist, wobei für beliebiges $x \in M$ und $z \in P$

$$(*) \qquad (x, z) \in g \circ f \Leftrightarrow \bigvee_y ((x, y) \in f \wedge (y, z) \in g)$$

gilt. Ist nun x ein beliebiges Element aus M, so ist wegen $D(f) = M$ nach (3') $(x, f(x)) \in f$ und wegen $f(x) \in N = D(g)$ ebenfalls nach (3') $(f(x), g(f(x))) \in g$. Folglich ist nach (*) (mit $y = f(x)$) $(x, g(f(x))) \in g \circ f$. Mithin gibt es zu jedem $x \in M$ ein $z \in P$ (nämlich $g(f(x))$) mit $(x, z) \in g \circ f$, d. h., $g \circ f$ ist eine Korrespondenz von M in P. Unsere Behauptung (14) ist bewiesen, wenn wir zeigen können, daß bei beliebigem $x \in M$ das Element $g(f(x))$ das einzige $z \in P$ mit $(x, z) \in g \circ f$ ist. Es sei also z ein beliebiges Element mit $(x, z) \in g \circ f$. Wegen (*) existiert dann ein y, so daß $(x, y) \in f$ und $(y, z) \in g$. Auf Grund der Eindeutigkeit von f muß hierbei $y = f(x)$ und auf Grund der Eindeutigkeit von g dann $z = g(f(x))$ sein, was zu zeigen war.

Ist f eine Abbildung von M auf N und g eine Abbildung von N auf P, so wird (Beweis!) $g \circ f$ eine Abbildung von M auf P. Sind f und g beides 1-1-Abbildungen, so ist auch $g \circ f$ eine 1-1-Abbildung.

Ist M eine beliebige (nichtleere) Menge, so bezeichnen wir mit $\mathfrak{T}(M)$ die Menge aller 1-1-Abbildungen von M auf sich:

$$(15) \qquad \mathfrak{T}(M) := \{f : f \textit{ 1-1-Abbildung von } M \textit{ auf } M\}.$$

Dann gehört sicher die durch

$$(16) \qquad e_M := \{(x, x) : x \in M\}$$

definierte *identische Abbildung* der Menge M, durch die jedes $x \in M$ auf sich selbst abgebildet wird, zu $\mathfrak{T}(M)$. Ferner ist nach dem Gesagten mit f stets auch f^{-1} und mit f und g stets auch das Produkt $g \circ f$ in $\mathfrak{T}(M)$ enthalten. Dabei sind

die folgenden Rechengesetze erfüllt:

(17a) $f \circ e_M = e_M \circ f = f$ *für alle* $f \in \mathfrak{T}(M)$;
(17b) $f \circ f^{-1} = f^{-1} \circ f = e_M$ *für alle* $f \in \mathfrak{T}(M)$;
(17c) $f_1 \circ (f_2 \circ f_3) = (f_1 \circ f_2) \circ f_3$ *für alle* $f_1, f_2, f_3 \in \mathfrak{T}(M)$.

Man sagt hierfür, daß *die Menge* $\mathfrak{T}(M)$ *bezüglich der Verkettungsoperation eine Gruppe bildet* und nennt diese Gruppe auch die *Permutations-* oder *Transformationsgruppe* der Menge M, da man die eineindeutigen Abbildungen einer Menge M auf sich auch *Permutationen* oder *Transformationen* von M zu nennen pflegt. Der Gruppenbegriff spielt in fast allen Bereichen der höheren Mathematik eine wesentliche Rolle. Sein systematisches Studium ist Gegenstand eines Teilgebietes der Algebra, der sogenannten Gruppentheorie.

Aus (17c) folgt (vgl. 3.5. (13)), daß es auch bei vier- und mehrgliedrigen Produkten nicht auf die Klammersetzung (wohl aber im allgemeinen auf die Reihenfolge der Faktoren) ankommt, also z. B. bei $f_1, f_2, f_3, f_4 \in \mathfrak{T}(M)$

$$\begin{aligned}((f_1 \circ f_2) \circ f_3) \circ f_4 = (f_1 \circ f_2) \circ (f_3 \circ f_4) &= (f_1 \circ (f_2 \circ f_3)) \circ f_4 \\ &= f_1 \circ (f_2 \circ (f_3 \circ f_4))\end{aligned}$$

ist, wofür man auch kurz $f_1 \circ f_2 \circ f_3 \circ f_4$ schreibt, usw.

Ist $M = \{a_1, \ldots, a_n\}$ eine Menge aus n Elementen, so bezeichnet man die Permutation, die dem Element a_ν das Element a_{i_ν} $(\nu = 1, \ldots, n)$ zuordnet, häufig mit

$$(18) \qquad \begin{pmatrix} a_1 & \ldots & a_n \\ a_{i_1} & \ldots & a_{i_n} \end{pmatrix}$$

oder kurz, indem man nur die Indizes notiert, mit

$$\begin{pmatrix} 1 & \ldots & n \\ i_1 & \ldots & i_n \end{pmatrix}.$$

Man beachte, daß hierbei grundsätzlich $\{a_{i_1}, \ldots, a_{i_n}\} = \{a_1, \ldots, a_n\}$ gilt, d. h. in der unteren Zeile, abgesehen von der Reihenfolge, dieselben Elemente wie in der oberen Zeile erscheinen. Die identische Permutation e_M wird dabei durch

$$\begin{pmatrix} a_1 & \ldots & a_n \\ a_1 & \ldots & a_n \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} 1 & \ldots & n \\ 1 & \ldots & n \end{pmatrix}$$

wiedergegeben. Die Darstellung (18) für die zu einer Permutation f inverse Permutation f^{-1} erhält man, indem man jeweils unter a_ν dasjenige Element a_{j_ν} notiert, unterhalb dessen in der Darstellung (18) von f das Element a_ν erscheint $(\nu = 1, \ldots, n)$. Die Darstellung (18) für das Produkt (die Verkettung) $g \circ f$ zweier Permutationen f, g erhält man, indem man jeweils unter a_ν das-

jenige Element notiert, das in der Darstellung (18) von g unter demjenigen Element a_{i_ν} steht, das in der Darstellung (18) von f unter a_ν zu finden ist (man beachte die Reihenfolge g nach f!).

Ist $M = \{a_1\}$ eine Einermenge, so besteht offensichtlich $\mathfrak{T}(M)$ nur aus der identischen Permutation (als Kuriosum merken wir an, daß es auch im Fall $M = \emptyset$ genau eine 1-1-Abbildung von M auf sich gibt, nämlich die durch die leere Menge repräsentierte leere Abbildung von $\emptyset$ auf sich).

Ist $M = \{a_1, a_2\}$ eine Zweiermenge, so enthält $\mathfrak{T}(M)$ die folgenden beiden Permutationen:

$$f_1 = e_M = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \qquad f_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Dabei gilt $f_1 \circ f_1 = f_2 \circ f_2 = f_1, f_1 \circ f_2 = f_2 \circ f_1 = f_2$.

Ist $M = \{a_1, a_2, a_3\}$ eine Dreiermenge, so besteht $\mathfrak{T}(M)$ bereits aus sechs Permutationen, nämlich

$$f_1 = e_M = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad f_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad f_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

$$f_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad f_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad f_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Die verschiedenen Produkte dieser Permutationen sind systematisch in der folgenden *Gruppentafel* zusammengestellt, bei der allgemein im Schnittpunkt der i-ten Zeile mit der j-ten Spalte das Produkt $f_i \circ f_j$ aufgeführt ist:

	f_1	f_2	f_3	f_4	f_5	f_6
f_1	f_1	f_2	f_3	f_4	f_5	f_6
f_2	f_2	f_1	f_5	f_6	f_3	f_4
f_3	f_3	f_4	f_1	f_2	f_6	f_5
f_4	f_4	f_3	f_6	f_5	f_1	f_2
f_5	f_5	f_6	f_2	f_1	f_4	f_3
f_6	f_6	f_5	f_4	f_3	f_2	f_1

Daß in der Tabelle in der ersten Zeile dieselben Elemente wie in der Kopfzeile und in der ersten Spalte dieselben Elemente wie in der Kopfspalte auftreten, ist natürlich nur ein Ausdruck dafür, daß für $f_1 = e_M$ die Gleichungen (17a) gelten. Die Gleichung $f_2 \circ f_2 = e_M$ besagt (vgl. (17b)), daß $f_2^{-1} = f_2$ (und analog $f_3^{-1} = f_3$, $f_6^{-1} = f_6$) ist, während aus den Gleichungen $f_4 \circ f_5 = f_5 \circ f_4 = e_M$ folgt, daß $f_4^{-1} = f_5$ und $f_5^{-1} = f_4$ gilt – was man natürlich auch direkt bestätigen kann. Die Gleichungen $f_2 \circ f_3 = f_5$ und $f_3 \circ f_2 = f_4$ lassen erkennen, daß die Multiplikation von Permutationen im allgemeinen nicht kommutativ ist.

Im Fall einer Vierermenge $M = \{a_1, a_2, a_3, a_4\}$ besteht die Menge $\mathfrak{T}(M)$ bereits aus 24 Elementen (vgl. 3. 6. (16)), so daß hier die Aufstellung der Gruppentafel eine schon recht erhebliche Rechenarbeit erfordert.

In den nun folgenden Betrachtungen sei $M = \{a_1, \ldots, a_n\}$ eine beliebige Menge aus $n \geqq 2$ Elementen. Es seien ferner $a_{i_1}, \ldots, a_{i_k}$ paarweise verschiedene Elemente der Menge M, wobei wir annehmen wollen, daß $k \geqq 2$ ist. Unter dem *Zyklus* $(a_{i_1} a_{i_2} \ldots a_{i_k})$ oder kurz $(i_1\ i_2 \ldots i_k)$ versteht man diejenige Permutation der Menge M, die a_{i_1} in a_{i_2}, a_{i_2} in $a_{i_3}, \ldots,$ $a_{i_{k-1}}$ in a_{i_k} und schließlich a_{i_k} in a_{i_1} abbildet und alle übrigen Elemente der Menge M ungeändert läßt. Im Fall $n = 6$ ist also z. B.

$$(1\ 5\ 2\ 6) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 4 & 2 & 1 \end{pmatrix}.$$

Die Zahl $k \geqq 2$ heißt dabei die *Länge* des Zyklus. Auf die Betrachtung von Zyklen der Länge 1 (die naturgemäß als identische Permutation zu definieren wären), wollen wir hier grundsätzlich verzichten. Es gilt der folgende wichtige Satz:

(19) *Jede von der identischen Permutation verschiedene Permutation f der Menge M läßt sich als (evtl. in einen Faktor ausgeartetes) Produkt von paarweise elementefremden Zyklen darstellen.*

Zum Beweis sei f eine beliebige Permutation der Menge M mit $f \neq e_M$. Dann existiert ein Element $a_{i_1} \in M$ mit $f(a_{i_1}) \neq a_{i_1}$. Wir setzen $f(a_{i_1}) = a_{i_2}$, $f(a_{i_2}) = a_{i_3}, \ldots$. Offenbar gelangen wir nach höchstens n Schritten zu einem Element a_{i_k}, für das $f(a_{i_k}) = a_{i_\varkappa}$ mit $1 \leqq \varkappa \leqq k$ wird. Wir nehmen an, daß das beim Element a_{i_k} zum ersten Mal geschieht, also für alle $\varkappa < k$ die Beziehungen $f(a_{i_\varkappa}) \neq a_{i_\lambda}$ mit $1 \leqq \lambda \leqq \varkappa$ gelten. Dann muß $f(a_{i_k}) = a_{i_1}$ sein; wäre nämlich beispielsweise $f(a_{i_k}) = a_{i_\mu}$ mit $1 < \mu \leqq k$, so wäre wegen $a_{i_\mu} = f(a_{i_{\mu-1}})$ und $f(a_{i_k}) = f(a_{i_{\mu-1}})$ offenbar $a_{i_k} = f(a_{i_{k-1}}) = a_{i_{\mu-1}}$ im Widerspruch dazu, daß für alle $\varkappa$, λ mit $1 \leqq \lambda \leqq \varkappa < k$ die Beziehung $f(a_{i_\varkappa}) \neq a_{i_\lambda}$ gelten sollte. Es sind dann folgende Fälle möglich:

Fall 1. Für alle $b \in M \setminus \{a_{i_1}, \ldots, a_{i_k}\}$ ist $f(b) = b$. In diesem Fall ist $f = (a_{i_1} \ldots a_{i_k})$, und unser Satz ist bewiesen.

Fall 2. Es gibt ein $b_{j_1} \in M \setminus \{a_{i_1}, \ldots, a_{i_k}\}$ mit $f(b_{j_1}) \neq b_{j_1}$. In diesem Fall bauen wir in analoger Weise, beginnend mit b_{j_1}, einen zweiten Zyklus $(b_{j_1} \ldots b_{j_l})$ auf. Man sieht leicht ein, daß dieser Zyklus zum Zyklus $(a_{i_1} \ldots a_{i_k})$ elementefremd ist (Beweis!). Es sind dann folgende Fälle möglich:

Fall 1. Für alle $c \in M \setminus (\{a_{i_1}, \ldots, a_{i_k}\} \cup \{b_{j_1}, \ldots, b_{j_l}\})$ ist $f(c) = c$. In diesem Fall ist $f = (b_{j_1} \ldots b_{j_l}) \circ (a_{i_1} \ldots a_{i_k})$ und unser Satz bewiesen.

Fall 2. Es existiert ein $c \in M \setminus (\{a_{i_1}, \ldots, a_{i_k}\} \cup \{b_{j_1}, \ldots, b_{j_l}\})$ mit $f(c) \neq c$. In diesem Fall fahren wir in analoger Weise fort.

Das beschriebene Verfahren muß nach einer endlichen Anzahl von Schritten, spätestens nämlich, wenn die Menge M ausgeschöpft ist, mit dem entsprechenden Fall 1 abbrechen. Man erhält dann die gesuchte Darstellung von f als Produkt von elementefremden Zyklen, womit der behauptete Satz bewiesen ist.

Man erkennt leicht, daß elementefremde Zyklen f_1, f_2 bei der Multiplikation (Verkettung) von Permutationen vertauschbar sind, d. h., daß für sie $f_1 \circ f_2 = f_2 \circ f_1$

gilt. Weiterhin kann man zeigen, daß die Darstellung einer Permutation als Produkt von elementefremden Zyklen bis auf die Reihenfolge (und Klammerung) der Faktoren eindeutig ist. Den in (19) auftretenden Ausnahmefall der identischen Permutation beseitigt man gerne dadurch, daß man formal auch ein *leeres Produkt* zuläßt und dieses gleich der identischen Permutation setzt.

Die Zyklen der Länge 2 heißen *Transpositionen*. Bei der Transposition $(a_i\ a_j)$ werden also lediglich die beiden Elemente a_i und a_j miteinander vertauscht. Als teilweise Verschärfung von Satz (19) gilt:

(20) *Jede Permutation f der Menge M läßt sich als Produkt von Transpositionen darstellen.*

Für die identische Permutation ist das klar; denn sind a_i, a_j zwei beliebige Elemente aus M $(a_i \neq a_j)$, so wird $e_M = (a_i\ a_j) \circ (a_i\ a_j)$. Ist $f = (a_{i_1} \dots a_{i_k})$ ein Zyklus einer Länge $k \geqq 3$, so wird (Beweis!)

$$f = (a_{i_1}\ a_{i_k}) \circ (a_{i_1}\ a_{i_{k-1}}) \circ \cdots \circ (a_{i_1}\ a_{i_3}) \circ (a_{i_1}\ a_{i_2}).$$

Ist schließlich f weder das eine noch das andere, so stellt man f zunächst nach Satz (19) als Produkt von elementefremden Zyklen dar und zerlegt anschließend jeden evtl. vorhandenen Faktor einer Länge $k \geqq 3$ in der angegebenen Weise in Transpositionen.

Wir merken an, daß in der Darstellung einer beliebigen Permutation als Produkt von Transpositionen die Faktoren im allgemeinen nicht mehr elementefremd gewählt werden können. Daher ist diese Darstellung nicht mehr eindeutig (schon e_M ist auf unendlich viele Weisen als Produkt von Transpositionen darstellbar!). Aus demselben Grunde sind hierbei die Faktoren auch im allgemeinen nicht mehr vertauschbar (nur elementefremde Zyklen sind stets vertauschbar!). Es läßt sich allerdings zeigen (schwierige Übungsaufgabe), daß bei einer beliebigen Permutation f entweder alle sie darstellenden Produkte von Transpositionen eine gerade Anzahl von Faktoren oder alle eine ungerade Anzahl von Faktoren haben. Im ersten Fall heißt f eine *gerade Permutation*, im zweiten eine *ungerade Permutation*.

Mit Hilfe des Begriffs der eineindeutigen Abbildung kann man eine notwendige und hinreichende Bedingung dafür angeben, daß (endliche) Mengen M und N dieselbe Anzahl von Elementen besitzen. Will man z. B. feststellen, ob in einem Zimmer dieselbe Anzahl von Menschen und Stühlen vorhanden ist, so braucht man keineswegs die anwesenden Personen und die vorhandenen Stühle abzuzählen, sondern nur zu bitten, Platz zu nehmen; stellt sich dabei heraus, daß für jede Person ein Stuhl vorhanden ist (also keine Person stehen bleibt) und hinterher kein Stuhl frei geblieben ist, so sind jedenfalls gleich viele Personen und Stühle vorhanden. Wesentlich ist hierbei offensichtlich nur, daß durch das Niedersetzen eine 1-1-Abbildung zwischen den Personen und den Stühlen hergestellt wird. Man sagt allgemein, Mengen M und N seien *gleichmächtig* (in der älteren Literatur ist daneben die farblose Bezeichnung „äquivalent" üblich), und schreibt dafür $M \sim N$, wenn es eine 1-1-Abbildung von M auf N gibt:

(21) $M \sim N :\Leftrightarrow \bigvee_f$ (*f 1-1-Abbildung von M auf N*).

Auf Grund des zuvor Gesagten ist bei endlichen Mengen die Gleichmächtigkeit ein Ausdruck dafür, daß die betrachteten Mengen dieselbe Anzahl von Elementen besitzen, ohne daß wir allerdings bisher die Begriffe „endliche Menge" und „Anzahl" genauer präzisiert haben (vgl. 2.8.). Aus (21) ergeben sich leicht die folgenden Eigenschaften der Gleichmächtigkeit:

(22a) *Für jede Menge M gilt* $M \sim M$ (Reflexivität);

(22b) $\bigwedge_{M_1,M_2,M_3} (M_1 \sim M_2 \wedge M_2 \sim M_3 \Rightarrow M_1 \sim M_3)$ (Transitivität);

(22c) $\bigwedge_{M_1,M_2} (M_1 \sim M_2 \Rightarrow M_2 \sim M_1)$ (Symmetrie).

Eine Relation mit den Eigenschaften (22) nennt man allgemein eine *Äquivalenzrelation* (vgl. 2.5. (12)). Die Eigenschaft (22a) folgt daraus, daß bei beliebigem M die identische Abbildung e_M eine 1-1-Abbildung von M auf sich ist (im Fall $M = \emptyset$ ist $e_M = \emptyset$, und die leere Menge von geordneten Paaren erfüllt formal die Bedingungen einer 1-1-Abbildung von $\emptyset$ auf $\emptyset$). Zum Beweis von (22b) genügt es zu bemerken, daß die Verkettung $g \circ f$ einer 1-1-Abbildung f von M_1 auf M_2 mit einer 1-1-Abbildung g von M_2 auf M_3 eine 1-1-Abbildung von M_1 auf M_3 ist. Und (22c) folgt schließlich daraus, daß die Umkehrabbildung f^{-1} einer 1-1-Abbildung f von M_1 auf M_2 eine 1-1-Abbildung von M_2 auf M_1 ist.

Als einfache Folgerung aus den Definitionen erhält man noch

(23) $M_1 \sim M_2 \wedge N_1 \sim N_2 \wedge M_1 \cap N_1 = \emptyset \wedge M_2 \cap N_2 = \emptyset$
$\Rightarrow M_1 \cup N_1 \sim M_2 \cup N_2.$

Ist nämlich f eine 1-1-Abbildung von M_1 auf M_2, g eine 1-1-Abbildung von N_1 auf N_2, so ist (Beweis!) $f \cup g$ $(= \{(x, y) : (x, y) \in f \vee (x, y) \in g\})$ eine 1-1-Abbildung von $M_1 \cup N_1$ auf $M_2 \cup N_2$.

Ähnlich beweist man (Übungsaufgabe):

(24) $M_1 \sim M_2 \wedge N_1 \sim N_2 \Rightarrow M_1 \times N_1 \sim M_2 \times N_2.$

Wegen seiner (bereits auf CANTOR zurückgehenden) Beweisidee verdient der folgende Satz besonderes Interesse:

(25) *Für keine Menge M gilt* $M \sim \mathfrak{P}(M)$.

Zum Beweis betrachten wir eine beliebige Menge M und eine beliebige 1-1-Abbildung f von M in $\mathfrak{P}(M)$ (z. B. ist die Korrespondenz $\{(x, \{x\}) : x \in M\}$ eine solche). Wir zeigen, daß f keine Abbildung auf $\mathfrak{P}(M)$ sein kann, d. h., daß wenigstens eine Menge $X_0 \subseteq M$ existiert, die bezüglich f kein Urbild besitzt. Dazu sei X_0 die Menge aller derjenigen $x \in M$, die nicht Element der x zugeordneten Menge $f(x) \subseteq M$ sind:

(*) $X_0 := \{x : x \in M \wedge x \notin f(x)\}.$

Angenommen, es gibt ein $x_0 \in M$ mit $f(x_0) = X_0$. Nach (*) gilt dann

$$x_0 \in f(x_0) \Leftrightarrow x_0 \in X_0 \Leftrightarrow x_0 \notin f(x_0),$$

und das ist ein Widerspruch. Also ist unsere Annahme falsch, und es gilt (25). Zur Illustration dieses überraschenden Schlusses (vgl. Abb. 7) betrachten wir die Korrespondenz

$$F := \{(x, y) : x \in M \wedge y \in M \wedge y \in f(x)\},$$

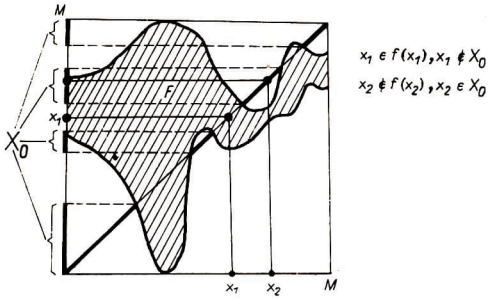


Abb. 7

bei der offenbar $B_F(x) = f(x)$ für alle $x \in M$ (da f eine 1-1-Abbildung von M in $\mathfrak{P}(M)$ sein sollte, sind die Mengen $B_F(x)$ paarweise verschieden). Ferner sei

$$G := \{(x, x) : x \in M \wedge x \notin f(x)\}.$$

Dann besteht G offenbar aus allen denjenigen Punkten der „Diagonalen" $e_M (= \{(x, x) : x \in M\})$, die nicht zu F gehören ($G = e_M \backslash F$), und wegen (*) gilt ferner $D(G) = W(G) = X_0$. Die Menge X_0 ist nun deshalb von allen Mengen $f(x)$ mit $x \in M$ verschieden, weil sich X_0 von einer beliebigen Menge $f(x_0)$ wenigstens im Element x_0 unterscheidet ($x_0 \in f(x_0) \triangle X_0$). Wegen dieser anschaulichen Deutung nennt man das verwendete Beweisverfahren häufig Cantorsches *Diagonalverfahren*.

Zum Abschluß wollen wir noch einige häufig in der Mathematik benutzte Bezeichnungen und Begriffe einführen, die eigentlich nur andere Benennungen oder Auffassungen von bereits behandelten Begriffen sind. Es sei dazu I eine beliebige Menge, die wir im vorliegenden Zusammenhang auch *Indexmenge* oder *Indexbereich* nennen wollen. Es sei ferner x eine Abbildung von I in eine gewisse Menge M ($x : I \to M$), die also jedem *Index* i aus I ein eindeutig bestimmtes Element $x(i)$ aus M als Bild zuordnet, wobei man statt $x(i)$ – im Sinne der Bezeichnung Index – auch x_i schreibt (hier ist also der Buchstabe x ein Funktionszeichen!). Die Abbildung x, also nach (3') die Menge

aller Paare (i, x_i) mit $i \in I$, wird dann eine *Familie* oder *(allgemeine) Folge* mit dem Indexbereich I genannt und mit $(x_i : i \in I)$ oder $(x_i)_{i \in I}$ bezeichnet. Das dem Index $i \in I$ durch x zugeordnete Element x_i heißt das zum Index i gehörige *Glied* dieser Familie oder Folge. Wir merken an, daß durchaus zugelassen ist, daß ein Element x aus M in einer Familie $(x_i)_{i \in I}$ mehrfach vorkommt (was bei Mengen unmöglich ist!), d. h., daß es mehrere Indizes $i \in I$ geben kann, für die $x_i = x$ gilt. Daher ist die Familie $(x_i)_{i \in I}$ wohl von der *Menge ihrer Glieder* zu unterscheiden, die naturgemäß nichts anderes als der Wertebereich der Abbildung x ist und nach (4′) durch $\{x_i : i \in I\}$ gegeben wird (wofür man manchmal auch $\{x_i\}_{i \in I}$ schreibt). Nach (7′) *sind Familien (Folgen)* $(x_i)_{i \in I}$ *und* $(y_i)_{i \in I}$ *mit dem Indexbereich* I *genau dann gleich, wenn sie gliedweise übereinstimmen:*

$$(26) \qquad (x_i)_{i \in I} = (y_i)_{i \in I} \Leftrightarrow \bigwedge_i (i \in I \Rightarrow x_i = y_i).$$

Eine Familie (Folge) $(y_j)_{j \in J}$ heißt *Teilfamilie (Teilfolge)* der Familie (Folge) $(x_i)_{i \in I}$, wenn $(y_j)_{j \in J} \subseteq (x_i)_{i \in I}$ im Sinne der Inklusion von Korrespondenzen (vgl. 2.3.(9)), so daß nach (8) gilt:

$$(27) \qquad (y_j)_{j \in J} \subseteq (x_i)_{i \in I} \Leftrightarrow J \subseteq I \wedge \bigwedge (j \in J \Rightarrow y_j = x_j).$$

Besonders wichtig ist der Spezialfall, daß der Indexbereich I die Menge $\mathbf{N}$ der natürlichen Zahlen oder die Menge $\mathbf{N}^*$ der positiven natürlichen Zahlen ist. Ist außerdem noch die Menge M, in die die Abbildung x erfolgt, die Menge $\mathbf{R}$ aller reellen Zahlen (also $W(x) \subseteq \mathbf{R}$), so nennt man $(x_i)_{i \in I}$ eine *reelle Zahlenfolge*. Ist z. B. für $i \in \mathbf{N}^*$ allgemein $x_i = \frac{1}{i}$, so erhält man die Zahlenfolge $\left(\frac{1}{i}\right)_{i \in \mathbf{N}^*}$ mit den Gliedern $1, \frac{1}{2}, \frac{1}{3}, \ldots$. Ist für $i \in \mathbf{N}$ allgemein $x_i = (-1)^i$, so erhält man die Zahlenfolge $((-1)^i)_{i \in \mathbf{N}}$ mit den Gliedern $1, -1, 1, -1, \ldots$; in diesem Fall ist $\{x_i\}_{i \in \mathbf{N}} = \{1, -1\}$.

Wichtig ist weiterhin der Fall, daß der Indexbereich I die Menge $\{1, \ldots, n\}$, d. h. die Menge aller natürlichen Zahlen i mit $1 \leqq i \leqq n$ ist, wobei n eine gegebene natürliche Zahl $\geqq 1$ ist. Auf Grund von (26) gilt

$$(26') \qquad (x_i)_{i \in \{1, \ldots, n\}} = (y_i)_{i \in \{1, \ldots, n\}} \Leftrightarrow x_1 = y_1 \wedge \cdots \wedge x_n = y_n,$$

d. h., die Folgen $(x_i)_{i \in \{1, \ldots, n\}}$ erfüllen die wesentliche Eigenschaft 2.2.(15) der n-Tupel, so daß in der Literatur auch vielfach die Menge M^n aller n-Tupel von Elementen der Menge M als Menge aller Abbildungen von $\{1, \ldots, n\}$ in M definiert wird. An Stelle der Menge $\{1, \ldots, n\}$ kann dabei natürlich auch irgendein anderer Indexbereich $I = \{i_1, \ldots, i_n\}$ aus n Elementen genommen werden.

Es ist durchaus auch zugelassen, daß die Glieder einer Familie oder Folge Mengen sind. Eine *Mengenfamilie* oder *Mengenfolge* $(M_i)_{i\in I}$ ist also nichts anderes als eine Abbildung M von I in ein gewisses Mengensystem $\mathfrak{M}$, durch die jedem Index $i \in I$ eine eindeutig bestimmte Menge $M(i)$ (oder M_i) aus $\mathfrak{M}$ zugeordnet wird (hier hat also der Buchstabe M die Rolle eines Funktionszeichens!).

Unter dem *Durchschnitt* bzw. der *Vereinigung* einer Mengenfamilie $(M_i)_{i\in I}$ werden dann die folgendermaßen definierten Mengen verstanden:

$$\bigcap_{i\in I} M_i := \{x : \bigwedge_i (i \in I \Rightarrow x \in M_i)\}, \tag{28}$$

$$\bigcup_{i\in I} M_i := \{x : \bigvee_i (i \in I \wedge x \in M_i)\}. \tag{29}$$

In Verallgemeinerung der Gesetze 1.4.(23) und 1.4.(24) gilt dann

$$N\setminus\bigcap_{i\in I} M_i = \bigcup_{i\in I} (N\setminus M_i), \quad N\setminus\bigcup_{i\in I} M_i = \bigcap_{i\in I} (N\setminus M_i), \tag{30}$$

wobei $(N\setminus M_i)_{i\in I}$ diejenige Mengenfamilie (Abbildung!) ist, bei der einem beliebigen Index $i \in I$ die Menge $N\setminus M_i$ zugeordnet ist. Als eine allgemeine Form der Distributivgesetze sind die folgenden Sätze anzusehen:

$$\begin{aligned} \bigcap_{i\in I} M_i \cup \bigcap_{j\in J} N_j &= \bigcap_{(i,j)\in I\times J} (M_i \cup N_j), \\ \bigcup_{i\in I} M_i \cap \bigcup_{j\in J} N_j &= \bigcup_{(i,j)\in I\times J} (M_i \cap N_j), \end{aligned} \tag{31}$$

wobei für gegebene Mengenfamilien $(M_i)_{i\in I}$, $(N_j)_{j\in J}$ beispielsweise unter $(M_i \cup N_j)_{(i,j)\in I\times J}$ die Mengenfamilie mit dem Indexbereich $I \times J$ verstanden wird, bei der das zum Index $(i, j) \in I \times J$ gehörige Glied die Menge $(M_i \cup N_j)$ ist. Ist der Indexbereich I Vereinigung der beiden disjunkten Indexbereiche I_1, I_2, so gilt

$$\begin{aligned} \bigcap_{i\in I} M_i &= \bigcap_{i\in I_1} M_i \cap \bigcap_{i\in I_2} M_i, \\ \bigcup_{i\in I} M_i &= \bigcup_{i\in I_1} M_i \cup \bigcup_{i\in I_2} M_i, \end{aligned} \tag{32}$$

wobei $(M_i)_{i\in I_1}$ bzw. $(M_i)_{i\in I_2}$ diejenigen Teilfamilien von $(M_i)_{i\in I}$ sind, die sich durch Einschränkung (vgl. (9)) der Abbildung M auf I_1 bzw. I_2 ergeben. (32) läßt sich sofort auf den Fall ausdehnen, daß I eine Vereinigung einer disjunkten Familie $(I_j)_{j\in J}$ von Indexbereichen ist, wobei die Familie $(I_j)_{j\in J}$ *disjunkt* heißt, wenn $I_{j_1} \cap I_{j_2} = \emptyset$ für alle $j_1, j_2 \in J$ mit $j_1 \neq j_2$:

$$\begin{aligned} \bigcap_{i\in\bigcup_{j\in J} I_j} M_i &= \bigcap_{j\in J} \Big(\bigcap_{i\in I_j} M_i\Big), \\ \bigcup_{i\in\bigcup_{j\in J} I_j} M_i &= \bigcup_{j\in J} \Big(\bigcup_{i\in I_j} M_i\Big) \end{aligned} \tag{32'}$$

(die Voraussetzung über die Disjunktheit der Indexbereiche ist sowohl bei (32) als auch bei (32′) unwesentlich). Die Beweise von (30), (31), (32), (32′) sind sämtlich sehr einfach und seien dem Leser als Übungsaufgabe überlassen.

2.5. Relationen

Es sei M eine beliebige Menge und k eine natürliche Zahl $\geqq 2$. Unter einer *k-stelligen Relation in M* versteht man eine beliebige Teilmenge der Menge M^k aller k-Tupel von Elementen aus M:

(1) $\quad R$ *k-stellige Relation in* $M : \Leftrightarrow R \subseteq M^k$.

Ist $(x_1, \ldots, x_k) \in R$, so sagt man, daß die Relation R auf das k-Tupel $(x_1, \ldots, x_k)$ *zutrifft* oder *die Elemente* $x_1, \ldots, x_k$ (in dieser Reihenfolge genommen) *in der Relation R stehen*.

Besonders wichtig sind die *zweistelligen* oder *binären* Relationen, mit denen wir uns im vorliegenden Abschnitt ausschließlich beschäftigen wollen und die daher einfach Relationen (ohne besondere Angabe der Stellenzahl 2) genannt werden sollen:

(2) $\quad R$ *Relation in* $M : \Leftrightarrow R \subseteq M^2$.

Offenbar ist danach eine Relation in M nichts anderes als eine Korrespondenz aus M in M, so daß wir gemäß 2.3.(2) statt $(x, y) \in R$ auch xRy schreiben. Nach 2.3.(11) bzw. 2.3.(16) sind mit R stets auch R^{-1} und mit R und S stets auch $S \circ R$ Relationen in M.

Als Beispiele für (zweistellige) Relationen seien hier nur folgende genannt: Die $\leqq$- und die $<$- Relation z. B. in der Menge R aller reellen Zahlen, die auf genau die Paare $(x, y) \in \mathsf{R} \times \mathsf{R}$ zutrifft, für die $x \leqq y$ bzw. $x < y$ gilt; die Inklusion bzw. echte Inklusion zwischen Mengen eines gegebenen Mengensystems $\mathfrak{M}$ (z. B. des Systems $\mathfrak{E}$ aller Mengen aus Elementen eines gegebenen Grundbereichs E), die auf genau die Paare $(X, Y) \in \mathfrak{M} \times \mathfrak{M}$ zutrifft, für die $X \subseteq Y$ bzw. $X \subset Y$ gilt; die Teilbarkeitsrelation in der Menge N aller natürlichen Zahlen, die auf genau die Paare $(x, y) \in \mathsf{N} \times \mathsf{N}$ zutrifft, bei denen x ein Teiler von y ist, usw.

Wir hatten bereits in 2.1. bemerkt, daß man vielfach allgemeiner jede Korrespondenz aus einer Menge M_1 in eine (evtl. andere) Menge M_2 als (binäre) Relation und dann jede Teilmenge einer Produktmenge $M_1 \times \cdots \times M_k$ als k-stellige Relation bezeichnet. Bei dieser allgemeinen Auffassung wird z. B. die durch die Eigenschaft „P liegt auf g“ für Punkte P und Geraden g beispiels-

weise einer euklidischen Ebene definierte Korrespondenz eine Relation zwischen Punkten und Geraden.

Wir betrachten nun einige häufig benötigte Eigenschaften von Relationen. Dabei setzen wir voraus, daß es sich stets um Relationen in einer festen nichtleeren Menge M handelt.

Eine Relation $R \subseteq M \times M$ heißt *reflexiv in M*, wenn jedes $x \in M$ zu sich selbst in der Relation R steht, d. h., R auf alle Paare (x, x) mit $x \in M$ zutrifft:

(3) $\quad R$ *reflexiv in* $M : \Leftrightarrow \bigwedge_x (x \in M \Rightarrow xRx)$.

Unter Verwendung der in 2.4.(16) eingeführten identischen Abbildung $e_M (:= \{(x, x) : x \in M\})$ können wir auch schreiben:

(3′) $\quad R$ *reflexiv in* $M \Leftrightarrow e_M \subseteq R$.

Danach ist eine Relation R genau dann nicht reflexiv in M, wenn ein Paar $(x, x) \in M \times M$ existiert, für das $\neg\, xRx$ gilt. Gilt $\neg\, xRx$ für alle Paare $(x, x) \in M \times M$, so nennt man die Relation R *irreflexiv in* M:

(4) $\quad R$ *irreflexiv in* $M : \Leftrightarrow \bigwedge_x (x \in M \Rightarrow \neg\, xRx)$,

wofür wir offenbar auch schreiben können:

(4′) $\quad R$ *irreflexiv in* $M \Leftrightarrow R \cap e_M = \emptyset$.

Eine Relation $R \subseteq M \times M$ heißt *transitiv*, wenn aus xRy und yRz stets xRz folgt (die Voraussetzungen $R \subseteq M \times M$ und xRy, yRz stellen bereits sicher, daß die Elemente x, y und z zu M gehören müssen):

(5) $\quad R$ *transitiv* $: \Leftrightarrow \bigwedge_{x,y,z} (xRy \wedge yRz \Rightarrow xRz)$.

Unter Verwendung der Verkettung läßt sich (5) auch schreiben als

(5′) $\quad R$ *transitiv* $\Leftrightarrow R \circ R \subseteq R$.

Eine Relation $R \subseteq M \times M$ heißt *symmetrisch*, wenn aus xRy stets yRx folgt:

(6) $\quad R$ *symmetrisch* $: \Leftrightarrow \bigwedge_{x,y} (xRy \Rightarrow yRx)$.

Nur eine andere Schreibweise für (6) ist

(6′) $\quad R$ *symmetrisch* $\Leftrightarrow R \subseteq R^{-1}$.

Wir merken an, daß man (6) bzw. (6′) sofort verschärfen kann zu

(7) $\quad R$ *symmetrisch* $\Leftrightarrow \bigwedge_{x,y} (xRy \Leftrightarrow yRx)$

bzw.

(7′) $\quad R$ *symmetrisch* $\Leftrightarrow R = R^{-1}$.

Eine Relation R ist also nicht symmetrisch, wenn es wenigstens ein Paar $(x, y) \in M \times M$ gibt, so daß xRy und $\neg yRx$. Gilt $\neg yRx$ für alle Paare (x, y), die in der Relation R stehen, so heißt R *asymmetrisch*:

(8) $\quad R$ *asymmetrisch* $: \Leftrightarrow \bigwedge_{x,y} (xRy \Rightarrow \neg yRx)$.

Eine Relation R ist also genau dann asymmetrisch, wenn für kein Paar $(x, y) \in M \times M$ sowohl xRy als auch yRx gilt:

$$R \text{ asymmetrisch} \Leftrightarrow \neg \bigvee_{x,y} (xRy \wedge yRx),$$

wofür wir offenbar auch

(8′) $\quad R$ *asymmetrisch* $\Leftrightarrow R \cap R^{-1} = \emptyset$

schreiben können.

Schließlich heißt eine Relation $R \subseteq M \times M$ *antisymmetrisch*, wenn aus xRy und yRx stets $x = y$ folgt:

(9) $\quad R$ *antisymmetrisch* $: \Leftrightarrow \bigwedge_{x,y} (xRy \wedge yRx \Rightarrow x = y)$

oder

(9′) $\quad R$ *antisymmetrisch* $\Leftrightarrow R \cap R^{-1} \subseteq e_M$.

Aus unseren Definitionen folgt z. B. sofort, daß *mit einer Relation R auch die zu ihr inverse Relation R^{-1} reflexiv, irreflexiv, transitiv, symmetrisch, asymmetrisch oder antisymmetrisch ist.* Ferner ist sofort zu sehen, daß *jede asymmetrische Relation irreflexiv ist:*

(10) $\quad R \subseteq M \times M \wedge R$ *asymmetrisch* $\Rightarrow R$ *irreflexiv in* M.

Umgekehrt *ist eine transitive Relation $R \subseteq M \times M$, wenn sie irreflexiv in M ist, auch asymmetrisch:*

(11) $\quad R \subseteq M \times M \wedge R$ *transitiv* $\wedge R$ *irreflexiv in* $M \Rightarrow R$ *asymmetrisch.*

Gäbe es nämlich ein Paar $(x, y) \in M \times M$, für das sowohl xRy als auch yRx gilt, so würde auf Grund der Transitivität auch xRx gelten, im Widerspruch zur Irreflexivität.

Eine wichtige Rolle spielen in der Mathematik die sogenannten *Äquivalenzrelationen*, die definiert sind durch

(12) $\quad R$ *Äquivalenzrelation in* M
$\quad : \Leftrightarrow R \subseteq M \times M \wedge R$ *reflexiv in* $M \wedge R$ *transitiv* $\wedge R$ *symmetrisch.*

Ist R eine Äquivalenzrelation in M und gilt xRy, so sagt man auch, daß y *bezüglich* (oder *nach* oder *modulo*) R *zu* x *äquivalent ist.* Die Menge $B_R(x)$ aller der-

jenigen $y \in M$, die zu x in der Relation R stehen, nennt man die *Äquivalenz-* oder *Restklasse von x nach* (oder *modulo*) R. Das System aller dieser Restklassen heißt das *Restsystem von M nach* (oder *modulo*) R und wird meist mit M/R bezeichnet:

$$(13) \qquad M/R := \{B_R(x) : x \in M\}.$$

Als erstes stellen wir fest, daß *bei beliebigem $x, y \in M$ die Restklassen $B_R(x)$ und $B_R(y)$ genau dann gleich sind, wenn xRy gilt:*

$$(14) \qquad B_R(x) = B_R(y) \Leftrightarrow xRy.$$

Ist nämlich $B_R(x) = B_R(y)$, so ist wegen $y \in B_R(y)$ (Reflexivität von R) auch $y \in B_R(x)$, und mithin gilt xRy. Gilt umgekehrt xRy und ist $z \in B_R(y)$, so gilt yRz und wegen der Transitivität von R auch xRz, d. h. $z \in B_R(x)$, so daß $B_R(y) \subseteq B_R(x)$; da ferner wegen der Symmetrie von R aus xRy auf yRx geschlossen werden kann, gilt entsprechend $B_R(x) \subseteq B_R(y)$, also folgt in der Tat aus xRy stets $B_R(x) = B_R(y)$.

Als nächtes behaupten wir, daß für das Restsystem M/R einer Äquivalenzrelation R in M stets die folgenden Eigenschaften erfüllt sind:

$$(15a) \qquad \bigwedge_X (X \in M/R \Rightarrow X \neq \emptyset),$$

$$(15b) \qquad \bigcup (M/R) = M,$$

(15c) M/R *ist disjunkt.*

Die Eigenschaften (15a) und (15b) folgen unmittelbar aus der Reflexivität von R. Auf Grund derer ist für jedes $x \in M$ nämlich $x \in B_R(x)$ und mithin $B_R(x) \neq \emptyset$ und $x \in \bigcup (M/R)$. Zum Beweis von (15c) zeigen wir, daß Restklassen $B_R(x)$, $B_R(y)$, die ein Element gemeinsam haben, übereinstimmen: Ist $z \in B_R(x)$, $z \in B_R(y)$, so gilt xRz und yRz, also nach (14) $B_R(x) = B_R(z)$ und $B_R(y) = B_R(z)$, so daß in der Tat $B_R(x) = B_R(y)$.

Ein Mengensystem $\mathfrak{Z}$ mit den unter (15a) bis (15c) angegebenen Eigenschaften des Systems M/R wird heute meist eine *Zerlegung* der Menge M genannt:

$$(16) \qquad \mathfrak{Z} \textit{ Zerlegung von } M :\Leftrightarrow \bigwedge_X (X \in \mathfrak{Z} \Rightarrow X \neq \emptyset) \wedge \bigcup \mathfrak{Z} = M \wedge \mathfrak{Z} \textit{ disjunkt.}$$

Eine Zerlegung von M ist also eine Einteilung von M in nichtleere Teilmengen, so daß jedes Element $x \in M$ genau einer dieser Teilmengen angehört (aus $\bigcup \mathfrak{Z} \subseteq M$ folgt nämlich, daß alle Mengen aus $\mathfrak{Z}$ Teilmengen von M sind, aus $M \subseteq \bigcup \mathfrak{Z}$ folgt, daß jedes Element aus M wenigstens einer Menge $X \in \mathfrak{Z}$ angehört, und die Disjunktheit von $\mathfrak{Z}$ besagt, daß kein Element aus M zu zwei

verschiedenen Mengen des Systems $\mathfrak{Z}$ gehört). Damit können wir (15a) bis (15c) zusammenfassen zu:

(15) R *Äquivalenzrelation in* $M \Rightarrow M/R$ *Zerlegung von* M.

Danach gehört also jedes Element x der Menge M einer und nur einer Restklasse des Systems M/R an, die dann aus allen und nur den Elementen y aus M besteht, die bezüglich R zu x äquivalent sind. Je zwei Elemente einer Restklasse sind äquivalent, während Elemente aus verschiedenen Restklassen nicht äquivalent sind.

Wir wollen nun zeigen, daß auch die Umkehrung von (15) gilt, genauer, daß man zu jeder Zerlegung $\mathfrak{Z}$ einer Menge M eine eindeutig bestimmte Äquivalenzrelation $R_{\mathfrak{Z}}$ in M finden kann, so daß $\mathfrak{Z}$ das Restsystem von $R_{\mathfrak{Z}}$ wird:

(17) $\mathfrak{Z}$ *Zerlegung von* $M \Rightarrow \bigvee\limits_{R_{\mathfrak{Z}}} (R_{\mathfrak{Z}}$ *Äquivalenzrelation in* $M \wedge \mathfrak{Z} = M/R_{\mathfrak{Z}})$.

Zum Beweis betrachten wir bei gegebener Menge M und gegebener Zerlegung $\mathfrak{Z}$ von M die durch

(18) $f_{\mathfrak{Z}} := \{(x, X) : x \in M \wedge X \in \mathfrak{Z} \wedge x \in X\}$

definierte Korrespondenz aus M in $\mathfrak{Z}$. Aus den Eigenschaften einer Zerlegung folgt mühelos, daß f eine eindeutige Abbildung von M auf $\mathfrak{Z}$ ist, und zwar ordnet $f_{\mathfrak{Z}}$ jedem Element $x \in M$ gerade diejenige Menge X der Zerlegung $\mathfrak{Z}$ zu, die x als Element enthält. Man nennt $f_{\mathfrak{Z}}$ die zur Zerlegung $\mathfrak{Z}$ gehörige *kanonische Abbildung*. Wenn es nun überhaupt eine Äquivalenzrelation R in M gibt, für die $\mathfrak{Z} = M/R$ wird, so muß wegen $x \in f_{\mathfrak{Z}}(x)$ bei beliebigem $x \in M$ jeweils $B_R(x) = f_{\mathfrak{Z}}(x)$ sein, also auf Grund von (14)

(*) $xRy \Leftrightarrow f_{\mathfrak{Z}}(x) = f_{\mathfrak{Z}}(y)$

gelten. Wir sehen nun umgekehrt (*) als Definition einer Relation $R_{\mathfrak{Z}}$ in M an, setzen also

(19) $R_{\mathfrak{Z}} := \{(x, y) : x \in M \wedge y \in M \wedge f_{\mathfrak{Z}}(x) = f_{\mathfrak{Z}}(y)\}$.

Man erkennt mühelos, daß die so definierte Relation $R_{\mathfrak{Z}}$ eine Äquivalenzrelation in M ist. Die Behauptung $\mathfrak{Z} = M/R_{\mathfrak{Z}}$ folgt unmittelbar daraus, daß bei beliebigem $x \in M$ die Gleichung $B_{R_{\mathfrak{Z}}}(x) = f_{\mathfrak{Z}}(x)$ gilt, denn für jedes $y \in M$ ist

$$y \in B_{R_{\mathfrak{Z}}}(x) \Leftrightarrow xR_{\mathfrak{Z}}y \Leftrightarrow f_{\mathfrak{Z}}(x) = f_{\mathfrak{Z}}(y) \Leftrightarrow y \in f_{\mathfrak{Z}}(x),$$

wobei im letzten Schritt benutzt wird, daß im Fall $y \in f_{\mathfrak{Z}}(x)$ die Mengen $f_{\mathfrak{Z}}(x), f_{\mathfrak{Z}}(y) \in \mathfrak{Z}$ das Element y gemeinsam haben, also wegen der Disjunktheit von $\mathfrak{Z}$ gleich sind.

Wir fassen die vorangehenden Resultate zusammen in dem folgenden

Hauptsatz über Äquivalenzrelationen. *Jede Äquivalenzrelation R in einer nichtleeren Menge M bewirkt eine eindeutig bestimmte Zerlegung M/R von M in Restklassen, wobei Elemente $x, y \in M$ genau dann derselben Restklasse angehören, wenn sie zueinander in der Relation R stehen. Umgekehrt existiert zu jeder Zerlegung $\mathfrak{Z}$ der Menge M eine eindeutig bestimmte Äquivalenzrelation $R_{\mathfrak{Z}}$, so daß $\mathfrak{Z}$ Restsystem $M/R_{\mathfrak{Z}}$ dieser Äquivalenzrelation ist.*

Wir wollen den Inhalt des Hauptsatzes an zwei Beispielen aus dem täglichen Leben erläutern. Dabei bedeute M die Menge aller Schüler, die zu einem bestimmten Zeitpunkt eine bestimmte Schule besuchen. Durch die Eigenschaft „y ist Klassenkamerad von x" wird (wenn wir jeden Schüler als Klassenkameraden von sich selbst ansehen) eine bestimmte Äquivalenzrelation in M definiert. Die Restklassen dieser Äquivalenzrelation sind dann gerade die verschiedenen Schulklassen. Fassen wir andererseits alle Schüler aus M, die einen bestimmten Vornamen haben, in jeweils einer Menge zusammen, so erhalten wir eine Zerlegung von M. Die zu dieser Zerlegung gehörige Äquivalenzrelation wird etwa durch die Eigenschaft „y hat denselben Vornamen wie x" charakterisiert.

Auf Grund von (12) ist klar, daß sowohl die identische Relation e_M als auch die Relation $M \times M$ Äquivalenzrelationen in M sind. Dabei ist offenbar e_M die bezüglich der Inklusion kleinste und $M \times M$ die bezüglich der Inklusion größte Äquivalenzrelation in M. Für e_M besteht das Restsystem M/e_M aus allen Einermengen $\{x\}$ mit $x \in M$, während für $M \times M$ das Restsystem $M/(M \times M)$ die Einermenge $\{M\}$ ist.

Es sei noch ein wichtiges Verfahren zur Erzeugung von Äquivalenzrelationen behandelt, das wiederum in dem Sinne universell ist, daß mit seiner Hilfe jede Äquivalenzrelation erzeugt werden kann. Dazu sei f eine beliebige Abbildung von der nichtleeren Menge M auf (bzw. in) eine gewisse Menge N. Wir setzen (vgl. (19)):

$$(20) \qquad R_f := \{(x, y) : x \in M \wedge y \in M \wedge f(x) = f(y)\}.$$

Man prüft mühelos nach, daß R_f eine Äquivalenzrelation in M ist, die man auch die durch f *induzierte* (erzeugte) Äquivalenzrelation nennt. Offenbar besteht das Restsystem M/R_f gerade aus allen vollen Urbildern $U_f(y)$ mit $y \in W(f)$:

$$(20') \qquad M/R_f = \{U_f(y) : y \in W(f)\}.$$

Die zur Zerlegung M/R_f gehörige kanonische Abbildung von M auf M/R_f, sie sei mit $\tilde{f}$ bezeichnet, ordnet dann einem beliebigen $x \in M$ die Menge aller

derjenigen $y \in M$ zu, die bei der Abbildung f auf $f(x)$ abgebildet werden:

(21) $\quad \check{f}(x) := \{y : y \in M \wedge f(y) = f(x)\} \quad (x \in M)$.

Daher ist die Korrespondenz

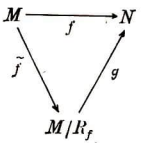
(22) $\quad g := \{(\xi, z) : \xi \in M/R_f \wedge z \in N \wedge \bigvee_x (x \in \xi \wedge z = f(x))\}$,

die einer beliebigen Klasse $\xi \in M/R_f$ das allen $x \in \xi$ gemeinsame Bild $z = f(x)$ aus N zuordnet, eine 1-1-Abbildung von M/R_f auf (bzw. in) N, und dabei gilt:

(23) $\quad f = g \circ \check{f}$,

was man sich gern durch ein Diagramm der Form

(23′)



veranschaulicht. *Jede eindeutige Abbildung f von einer Menge M auf (in) eine Menge N ist also Verkettung der kanonischen Abbildung $\check{f}$ von M auf das Restsystem M/R einer geeigneten Äquivalenzrelation R in M* (nämlich der durch f induzierten Äquivalenzrelation (20)) *und einer 1-1-Abbildung g von M/R auf (in) N.* Dabei ist offenbar die Abbildung f genau dann eineindeutig, wenn $\check{f}(x) = \{x\}$ $(x \in M)$. *Umgekehrt ist natürlich auch jede Verkettung $g \circ f_{\mathfrak{Z}}$ der zu einer Zerlegung $\mathfrak{Z}$ von M* (dem Restsystem M/R einer Äquivalenzrelation R in M) *gehörigen kanonischen Abbildung $f_{\mathfrak{Z}}$ mit einer 1-1-Abbildung g von $\mathfrak{Z}$ auf (in) eine beliebige Menge N eine eindeutige Abbildung von M auf (in) N.* Die von der kanonischen Abbildung $f_{\mathfrak{Z}}$ induzierte Äquivalenzrelation $R_{f_{\mathfrak{Z}}}$ stimmt natürlich mit der durch (19) gegebenen Äquivalenzrelation $R_{\mathfrak{Z}}$ überein.

Für (23) sagt man auch, daß das Diagramm (23′) kommutativ ist.

Die im obigen Beispiel betrachtete Relation „y hat denselben Vornamen wie x" wird z. B. durch die Abbildung f induziert, die jedem Schüler $x \in M$ seinen Vornamen zuordnet. Die Abbildung f von der Menge M der Schüler in die Menge N aller Lehrer der betrachteten Schule, die jedem Schüler seinen Klassenlehrer zuordnet, induziert die durch die Eigenschaft „y ist Klassenkamerad von x" charakterisierte Äquivalenzrelation (vorausgesetzt, daß kein Lehrer der Schule Klassenlehrer zweier verschiedener Klassen ist).

Wir kommen nun zu den wichtigsten Arten von Ordnungsrelationen, die ebenfalls in vielen Gebieten der Mathematik von Bedeutung sind. Es sei

wieder M eine beliebige (nichtleere) Menge. Eine Relation $R \subseteq M \times M$ heißt eine *reflexive teilweise Ordnung in* (oder *von*) M, wenn R reflexiv, transitiv und antisymmetrisch ist.

(24) *R reflexive teilweise Ordnung in M*
$:\Leftrightarrow R \subseteq M \times M \wedge R$ *reflexiv in* $M \wedge R$ *transitiv* $\wedge R$ *antisymmetrisch.*

Eine Relation $S \subseteq M \times M$ heißt eine *irreflexive teilweise Ordnung in* M, wenn S irreflexiv und transitiv ist:

(25) *S irreflexive teilweise Ordnung in M*
$:\Leftrightarrow S \subseteq M \times M \wedge S$ *irreflexiv in* $M \wedge S$ *transitiv.*

Als Musterbeispiele für reflexive teilweise Ordnungen können die $\leqq$-Relation für (reelle) Zahlen, die Inklusion für Mengen, aber z. B. auch die Teilbarkeitsrelation im Bereich der natürlichen Zahlen (vgl. 3.7.) dienen. Musterbeispiele für irreflexive teilweise Ordnungen sind die $<$-Relation für Zahlen und die echte Inklusion für Mengen. Der Zusatz „teilweise" bezieht sich darauf, daß es bei einer solchen Ordnung – wie z. B. bei der Inklusion – *unvergleichbare Elemente* geben kann, d. h. Elemente $x, y \in M$ mit $x \neq y$, für die weder xRy (bzw. xSy) noch yRx (bzw. ySx) gilt. Ein besonders extremes Beispiel ist die identische Relation e_M, von der man leicht nachprüft, daß sie eine reflexive teilweise Ordnung ist, und bei der je zwei verschiedene Elemente aus M unvergleichbar sind (daher nennt man e_M manchmal auch *totale Unordnung*). Statt teilweise Ordnung sagt man vielfach auch *Halbordnung* oder *partielle Ordnung*; in der neueren Literatur werden die teilweisen Ordnungen häufig auch einfach *Ordnungen* genannt.

Wir merken als erstes an, daß *mit einer Relation R (S) stets auch die zu ihr inverse Relation R^{-1} (S^{-1}) eine reflexive (irreflexive) teilweise Ordnung ist.* Die zur $\leqq$-Relation ($<$-Relation) für Zahlen inverse Ordnung ist die $\geqq$-Relation ($>$-Relation), zur $\subseteq$-Relation ($\subset$-Relation) für Mengen ist die $\supseteq$-Relation ($\supset$-Relation) invers (vgl. 1.5.(2) und 1.5.(7)).

Bei Betrachtung der Definitionen (24) und (25) mag vielleicht auffallen, daß die in (24) geforderte Antisymmetrie in (25) kein Entsprechen hat. Das liegt daran, daß die hier sinngemäß zu fordernde Asymmetrie auf Grund von (11) bereits aus der Irreflexivität und Transitivität von S folgt, während Analoges – wie das Beispiel der Äquivalenzrelationen (mit $R \cap R^{-1} = R$ statt $R \cap R^{-1} = e_M$) zeigt – für die Antisymmetrie nicht gilt. Allgemein kann man jedoch leicht zeigen (Übungsaufgabe), daß für eine reflexive und transitive Relation R (eine solche nennt man meistens eine *Quasiordnung*) durch

$$xTy :\Leftrightarrow xRy \wedge yRx \quad (\text{d. h. } T := R \cap R^{-1})$$

eine Äquivalenzrelation in M definiert wird, und die zusätzliche Forderung der Antisymmetrie stellt gerade sicher, daß T die identische Relation e_M ist.

Wir wollen nun zeigen, daß zwischen den reflexiven und den irreflexiven teilweisen Ordnungen ein sehr einfacher Zusammenhang besteht, daß *man nämlich in kanonischer Weise*, so wie man das von der $\leqq$- und der $<$-Relation für Zahlen kennt, *aus einer reflexiven teilweisen Ordnung eine irreflexive teilweise Ordnung erhalten kann und umgekehrt.*

Es sei dazu zunächst R eine reflexive teilweise Ordnung in M. Wir definieren mit ihrer Hilfe die Relation R_i durch

(26) $\quad xR_iy :\Leftrightarrow xRy \wedge x \neq y$

($x < y :\Leftrightarrow x \leqq y \wedge x \neq y$), d. h., wir setzen

(26′) $\quad R_i := R \setminus e_M$.

Unsere Behauptung ist, daß *die Relation R_i eine irreflexive teilweise Ordnung in M ist:*

(27) $\quad$ *R reflexive teilweise Ordnung in M*
$\Rightarrow$ R_i irreflexive teilweise Ordnung in M.

Zunächst folgt aus (26′) sofort $R_i \cap e_M = \emptyset$, d. h. (vgl. (4′)), R_i ist irreflexiv in M. Zum Nachweis der Transitivität nehmen wir an, es gelte xR_iy und yR_iz. Dann gilt wegen (26) xRy, $x \neq y$, yRz und $y \neq z$. Aus xRy und yRz folgt wegen der Transitivität von R sofort xRz. Es bleibt also zum Nachweis von xR_iz zu zeigen, daß auch $x \neq z$ gilt. Wir nehmen an, es wäre $x = z$. Dann würde sowohl xRy als auch (wegen yRz) yRx gelten, und wegen der Antisymmetrie von R wäre $x = y$, was ja nicht der Fall sein sollte.

Es sei nun S eine beliebige irreflexive teilweise Ordnung in M. Wir definieren mit ihrer Hilfe eine Relation S_r durch

(28) $\quad xS_ry :\Leftrightarrow xSy \vee x = y$

($x \leqq y :\Leftrightarrow x < y \vee x = y$), d. h., wir setzen

(28′) $\quad S_r := S \cup e_M$.

Unsere Behauptung ist, daß *die Relation S_r eine reflexive teilweise Ordnung in M ist:*

(29) $\quad$ *S irreflexive teilweise Ordnung in M*
$\Rightarrow$ S_r reflexive teilweise Ordnung in M.

Zunächst folgt aus (28′) sofort $e_M \subseteq S_r$, d. h. (vgl. (3′)), die Relation S_r ist reflexiv in M. Zum Nachweis der Transitivität von S_r nehmen wir an, es gelte

$xS_r y$ und $yS_r z$. Nach (28) muß dann einer der folgenden vier Fälle vorliegen:

(i) $xSy \wedge ySz$,
(ii) $xSy \wedge y = z$,
(iii) $x = y \wedge ySz$,
(iv) $x = y \wedge y = z$.

Im ersten Fall gilt xSz wegen der Transitivität von S und damit erst recht $xS_r z$, im zweiten und im dritten Fall gilt trivialerweise xSz und damit $xS_r z$, im vierten Fall gilt $x = z$ und damit ebenfalls $xS_r z$. Zum Nachweis der Antisymmetrie von S_r merken wir zunächst an, daß *für beliebige Relationen* R_1, R_2 *in* M *stets*

(30) $\quad (R_1 \cup R_2)^{-1} = (R_1^{-1} \cup R_2^{-1})$

gilt. Denn für beliebiges $(x, y) \in M \times M$ ist

$$\begin{aligned}(x, y) \in (R_1 \cup R_2)^{-1} &\Leftrightarrow (y, x) \in R_1 \cup R_2\\ &\Leftrightarrow (y, x) \in R_1 \vee (y, x) \in R_2\\ &\Leftrightarrow (x, y) \in R_1^{-1} \vee (x, y) \in R_2^{-1}\\ &\Leftrightarrow (x, y) \in R_1^{-1} \cup R_2^{-1}.\end{aligned}$$

Folglich wird

$$\begin{aligned}S_r \cap S_r^{-1} &= (S \cup e_M) \cap (S \cup e_M)^{-1} = (S \cup e_M) \cap (S^{-1} \cup e_M)\\ &= (S \cap S^{-1}) \cup e_M = e_M;\end{aligned}$$

denn es ist $S \cap S^{-1} = \emptyset$ wegen der Asymmetrie von S (vgl. (11)). Die Gleichung $S_r \cap S_r^{-1} = e_M$ besagt aber gerade (vgl. (9')) die Antisymmetrie von S_r. Wir merken an, daß auch die Transitivität von S_r durch einen analogen Schluß bewiesen werden kann, wenn man sich zuvor die folgenden Distributivgesetze verschafft (Übungsaufgabe):

$$\begin{aligned}&R_1 \circ (R_2 \cup R_3) = R_1 \circ R_2 \cup R_1 \circ R_3,\\ &(R_2 \cup R_3) \circ R_1 = R_2 \circ R_1 \cup R_3 \circ R_1.\end{aligned} \tag{31}$$

Beachtet man, daß (in Verallgemeinerung von 2.4.(17)) für jede Relation R in M

(32) $\quad R \circ e_M = e_M \circ R = R$

gilt, so wird

$$\begin{aligned}S_r \circ S_r &= (S \cup e_M) \circ (S \cup e_M)\\ &= S \circ S \cup S \circ e_M \cup e_M \circ S \cup e_M \circ e_M\\ &\subseteq S \cup e_M = S_r,\end{aligned}$$

da wegen der Transitivität von S (vgl. (5')) ja $S \circ S \subseteq S$ ist, und die Gleichung $S_r \circ S_r \subseteq S_r$ besagt gerade die Transitivität von S_r.

Aus (26') und (28') erhält man noch leicht, daß *für jede reflexive teilweise Ordnung R in M*

(33) $\quad (R_i)_r = R$

gilt. Denn

$$(R_i)_r = R_i \cup e_M = (R \setminus e_M) \cup e_M = R \cup e_M = R,$$

da wegen der Reflexivität von R offenbar $R \cup e_M = R$ ist.

Analog *gilt für jede irreflexive teilweise Ordnung S in M*

(34) $\quad (S_r)_i = S.$

Eine Relation $R \subseteq M \times M$ heißt *linear in M*, wenn bei beliebigem $x, y \in M$ stets wenigstens einer der Fälle xRy oder yRx eintritt, wenn also beliebige Elemente $x, y \in M$ durch R *vergleichbar* sind:

(35) $\quad R$ *linear in* $M : \Leftrightarrow \bigwedge_{x,y} (x \in M \wedge y \in M \Rightarrow xRy \vee yRx),$

wofür man auch

(35') $\quad R$ *linear in* $M \Leftrightarrow R \cup R^{-1} = M \times M$

schreiben kann. Man erkennt leicht, daß *jede lineare Relation reflexiv ist:*

(36) $\quad R$ *linear in* $M \Rightarrow R$ *reflexiv in* $M.$

Eine Relation $S \subseteq M \times M$ heißt *konnex in M*, wenn bei beliebigem $x, y \in M$ stets wenigstens einer der Fälle xSy oder ySx oder $x = y$ eintritt, anders ausgedrückt, wenn je zwei verschiedene Elemente aus M durch S vergleichbar sind:

(37) $\quad S$ *konnex in* $M : \Leftrightarrow \bigwedge_{x,y} (x \in M \wedge y \in M \Rightarrow xSy \vee ySx \vee x = y),$

wofür man auch

(37') $\quad S$ *konnex in* $M \Leftrightarrow S \cup S^{-1} \cup e_M = M \times M$

schreiben kann. Aus (35') bzw. (37') folgt sofort, daß *mit einer Relation auch die zu ihr inverse Relation linear bzw. konnex ist.* Mittels (26') bzw. (28') erhält man weiterhin

(38) $\quad R$ *linear in* $M \Rightarrow R_i$ *konnex in* $M,$

(39) $\quad S$ *konnex in* $M \Rightarrow S_r$ *linear in* $M.$

Eine reflexive (irreflexive) teilweise Ordnung in M, die außerdem linear (konnex) ist, heißt eine *reflexive (irreflexive) totale Ordnung*. In der älteren Literatur werden die totalen Ordnungen auch kurz *Ordnungen* genannt.

(40) *R reflexive totale Ordnung in M*
$:\Leftrightarrow$ *R reflexive teilweise Ordnung in $M \wedge R$ linear in M,*

(41) *S irreflexive totale Ordnung in M*
$:\Leftrightarrow$ *S irreflexive teilweise Ordnung in $M \wedge S$ konnex in M.*

Mittels (27) und (38) bzw. (29) und (39) erhält man

(42) *R reflexive totale Ordnung in M*
$\Rightarrow$ *R_i irreflexive totale Ordnung in M,*

(43) *S irreflexive totale Ordnung in M*
$\Rightarrow$ *S_r reflexive totale Ordnung in M.*

Bei einer irreflexiven totalen Ordnung S läßt sich die Konnexität leicht zur sogenannten *Trichotomie* verschärfen, die besagt, daß für beliebige Elemente $x, y \in M$ stets genau einer der Fälle xSy oder ySx oder $x = y$ eintritt, also stets wenigstens einer dieser Fälle und niemals zwei von ihnen eintreten:

(44) *S trichotom in M*
$$:\Leftrightarrow \bigwedge_{x,y} (x \in M \wedge y \in M \Rightarrow (xSy \vee ySx \vee x = y) \wedge \neg (xSy \wedge ySx) \wedge \neg (xSy \wedge x = y) \wedge \neg (ySx \wedge x = y)).$$

Der behauptete Satz besagt dann

(45) *S irreflexive totale Ordnung in $M \Rightarrow S$ trichotom in M.*

Zum Beweis genügt es zu bemerken, daß sich auf Grund der Asymmetrie von S (vgl. (11)) die Fälle xSy und ySx ausschließen, und auf Grund der Irreflexivität von S ebenso die Fälle xSy und $x = y$ wie auch ySx und $x = y$.

2.6. Operationen

Es sei M eine beliebige Menge und k eine natürliche Zahl ≥ 1. Unter einer *k-stelligen Operation in M* verstehen wir eine Abbildung o von der Menge M^k aller k-Tupel $(x_1, \ldots, x_k)$ von Elementen aus M in die Menge M:

(1) *o k-stellige Operation in M* $:\Leftrightarrow o : M^k \to M$.

Das dem k-Tupel $(x_1, \ldots, x_k) \in M^k$ durch o zugeordnete Element $o(x_1, \ldots, x_k)$ (es wäre genau genommen natürlich mit $o((x_1, \ldots, x_k))$ zu bezeichnen) heißt das *Resultat* der Operation o für $(x_1, \ldots, x_k)$. Vielfach spricht man schon dann von einer k-stelligen Operation in M, wenn o eine Abbildung *aus* M^k *in* M ist, und nennt die Operationen o, für die $D(o) = M^k$ ist, die also jedem k-Tupel ein Resultat zuordnen, *unbeschränkt ausführbare Operationen*. Wir wollen hier Abbildungen aus M^k in M, wenn sie gelegentlich eine Rolle spielen, *beschränkt ausführbare* oder *partielle Operationen* nennen (bei ihnen ist also die Ausführbarkeit, d. h. die Existenz des Resultats, auf $D(o)$ beschränkt). Die „unbeschränkt ausführbaren" Operationen werden dabei als spezielle beschränkt ausführbare Operationen angesehen. Nicht selten spricht man auch schon von einer k-stelligen Operation, wenn eine Abbildung o von (oder aus) einer Produktmenge $M_1 \times \cdots \times M_k$ in eine Menge N vorliegt.

Besonders wichtig sind wieder die *zweistelligen* oder *binären* Operationen, mit denen wir uns im vorliegenden Abschnitt vorwiegend beschäftigen wollen und die daher einfach Operationen (ohne besondere Angabe der Stellenzahl 2) genannt werden sollen:

(2) $\quad o$ *Operation in* $M :\Leftrightarrow o : M^2 \to M$.

Das einem Paar $(x, y) \in M \times M$ durch eine binäre Operation o zugeordnete Resultat $o(x, y)$ wird im folgenden — wie man das in der Mathematik meistens tut — mit xoy bezeichnet:

(3) $\quad xoy := o(x, y)$.

Wir merken an, daß eine einstellige Operation in M nichts anderes als eine Abbildung von M in M ist.

Durch

(4) $\quad R_o := \{(x, y, z) : x \in M \wedge y \in M \wedge z = xoy\}$

wird jeder binären Operation o eine bestimmte dreistellige Relation R_o zugeordnet (bei der Definition 2.2.(11) der Tripel ist sogar $R_o = o$), die folgende Bedingungen erfüllt:

(5) $\quad \bigwedge_{x,y} (x \in M \wedge y \in M \Rightarrow \bigvee_{z} (z \in M \wedge (x, y, z) \in R))$,

(6) $\quad (x, y, z_1) \in R \wedge (x, y, z_2) \in R \Rightarrow z_1 = z_2$.

Umgekehrt kann jede dreistellige Relation R in M, die diese Bedingungen erfüllt, als eine binäre Operation in M aufgefaßt werden (läßt man die Bedingung (5) fort, so erhält man gerade die beschränkt ausführbaren Operationen).

Als Beispiele für (zweistellige) Operationen seien genannt: die Addition und Multiplikation z. B. im Bereich R der reellen Zahlen, die Bildung des Durch-

schnitts, der Vereinigung, der symmetrischen Differenz und der Differenzmenge im System $\mathfrak{E}$ aller Mengen über einem gegebenen Grundbereich E, die Bildung des Relationenprodukts $S \circ R$ in der Menge $\mathfrak{P}(M \times M)$ aller binären Relationen in M usw. Die Abbildung, die einer binären Relation R in M ihre inverse Relation R^{-1} zuordnet, ist eine einstellige Operation (in $\mathfrak{P}(M \times M)$), ebenso die Abbildung, die einer Zahl $a \in \mathsf{R}$ die Zahl $-a$ zuordnet usw. Die Bildung der Produktmenge $X \times Y$ ist bei unserer Auffassung keine Operation z. B. im System $\mathfrak{E}$ aller Mengen über einem Grundbereich E, da das Resultat nicht wieder zu $\mathfrak{E}$ gehört. Um eben auch derartige Fälle mit zu erfassen, faßt man manchmal den Begriff der Operation allgemeiner und sieht schon jede Abbildung aus (oder von) einer Menge $M_1 \times M_2$ in eine (evtl. andere) Menge N als Operation an. Weitere interessante Beispiele für solche allgemeinen Operationen liefert die Vektorrechnung, wo man die Bildung des skalaren Vielfachen eines Vektors als Operation auffassen kann, die jedem Paar aus Skalar und Vektor als Resultat einen Vektor zuordnet, während das sogenannte Skalarprodukt (innere Produkt) eine Operation ist, die jedem Paar von Vektoren als Resultat ein Skalar zuordnet. Als Beispiele für beschränkt ausführbare Operationen erwähnen wir die Bildung des Quotienten $\frac{a}{b}$ von Zahlen aus R (beschränkt hier auf die Paare $(a, b) \in \mathsf{R} \times \mathsf{R}$, für die $b \neq 0$ ist) oder die Bildung der Differenz $a - b$ für Zahlen aus N (beschränkt hier auf die Paare $(a, b) \in \mathsf{N} \times \mathsf{N}$, für die $a \geqq b$ ist) usw.

Wir wollen nun eine Reihe von in der Mathematik häufig auftretenden Eigenschaften von Operationen systematisch zusammenstellen, denen wir auch im vorangehenden größtenteils schon begegnet sind. Auf allgemeine Zusammenhänge zwischen diesen Eigenschaften kann dabei nur in einigen einfachen Fällen eingegangen werden. Wir verwenden bei den grundlegenden Definitionen die nun bereits hinlänglich geübte Abkürzungstechnik, ohne den jeweiligen Sachverhalt noch immer breit zu erörtern. Bei den im folgenden auftretenden Operationen o, o_1, o_2 soll es sich stets um Operationen in einer fest gegebenen Menge M handeln, so daß sich die Quantifizierungen der Form „für jedes $x \ldots$“ und „es gibt ein $x \ldots$“ grundsätzlich auf Elemente aus M beziehen, was wir durch $\bigwedge_{x \in M} \ldots$ (als Abkürzung für $\bigwedge_{x} (x \in M \Rightarrow \ldots)$) bzw. $\bigvee_{x \in M} \ldots$ (als Abkürzung für $\bigvee_{x} (x \in M \wedge \ldots)$) andeuten wollen.

(7) o *kommutativ* $: \Leftrightarrow \bigwedge_{x, y \in M} xoy = yox,$

(8) o *assoziativ* $: \Leftrightarrow \bigwedge_{x, y, z \in M} xo(yoz) = (xoy)oz,$

(9_l) o_1 *linksseitig distributiv bzgl.* $o_2 :\Leftrightarrow \bigwedge_{x,y,z\in M} xo_1(yo_2z) = (xo_1y)o_2(xo_1z)$,

(9_r) o_1 *rechtsseitig distributiv bzgl.* $o_2 :\Leftrightarrow \bigwedge_{x,y,z\in M} (yo_2z)o_1x = (yo_1x)o_2(zo_1x)$,

(9) o_1 *(beidseitig) distributiv bzgl.* o_2
$:\Leftrightarrow o_1$ *linksseitig distributiv bzgl.* $o_2 \wedge o_1$ *rechtsseitig distributiv bzgl.* o_2.

Wir merken an, daß eine kommutative Operation o_1, die bzgl. einer Operation o_2 linksseitig distributiv ist, auch rechtsseitig distributiv bzgl. o_2 ist, und umgekehrt, d. h., bei einer kommutativen Operation folgt aus der einseitigen (rechts- oder linksseitigen) Distributivität bereits die beidseitige Distributivität.

(10) o *idempotent* $:\Leftrightarrow \bigwedge_{x\in M} xox = x$.

In den folgenden Definitionen seien e_l, e_r, e Elemente aus M.

(11_l) e_l *linksseitig neutrales Element für* $o :\Leftrightarrow \bigwedge_{x\in M} e_lox = x$,

(11_r) e_r *rechtsseitig neutrales Element für* $o :\Leftrightarrow \bigwedge_{x\in M} xoe_r = x$,

(11) e *(beidseitig) neutrales Element für* o
$:\Leftrightarrow e$ *linksseitig neutrales Element für* o
$\wedge\ e$ *rechtsseitig neutrales Element für* o.

Wir merken an, daß es bei gegebener Operation o durchaus möglich ist, daß es für o weder ein links- noch ein rechtsseitig neutrales Element gibt. Ein einfaches Beispiel hierfür ist die in der Zweiermenge $\{a, b\}$ $(a \neq b)$ durch die folgende Operationstabelle gegebene Operation:

	a	b
a	a	a
b	a	a

Es kann auch mehrere linksseitig oder mehrere rechtsseitig neutrale Elemente geben. Als Beispiele nennen wir die durch die Tabellen

	a	b
a	a	b
b	a	b

bzw.

	a	b
a	a	a
b	b	b

gegebenen Operationen (im ersten Fall sind a und b beide linksseitig neutral, im zweiten Fall sind a und b beide rechtsseitig neutral). *Wenn allerdings für*

eine Operation o sowohl ein linksseitig neutrales Element e_l als auch ein rechtsseitig neutrales Element e_r vorhanden ist, dann gibt es jeweils nur eines, wobei überdies $e_l = e_r$ gilt, d. h., dieses ist dann zugleich beidseitig neutrales Element. Insbesondere kann es also für eine Operation o höchstens ein beidseitig neutrales Element geben. Zum Beweis dieser Behauptung nehmen wir an, es sei e_l ein gewisses linksseitig neutrales Element für o und e_r ein rechtsseitig neutrales Element. Ist dann e_l' ein beliebiges linksseitig neutrales Element, so gilt

(i) $e_l o e_r = e_r$, da e_l linksseitig neutral ist;
(ii) $e_l' o e_r = e_r$, da e_l' linksseitig neutral ist;
(iii) $e_l o e_r = e_l$, da e_r rechtsseitig neutral ist;
(iv) $e_l' o e_r = e_l'$, da e_r rechtsseitig neutral ist.

Aus (i) und (iii) folgt $e_r = e_l$, und aus (ii) und (iv) folgt $e_r = e_l'$, so daß $e_l = e_l'$. Damit ist gezeigt, daß es nur ein einziges linksseitig neutrales Element gibt, das wir mit e_l^* bezeichnen wollen. Da aber in den vorangehenden Überlegungen e_r ein ganz beliebiges rechtsseitig neutrales Element war, ist damit zugleich gezeigt, daß jedes rechtsseitig neutrale Element mit e_l^* übereinstimmt, also auch nur ein rechtsseitig neutrales Element e_r^* existiert, das zudem gleich e_l^* ist. Wir haben diesen Beweis deshalb so ausführlich wiedergegeben, weil er wegen seiner Abstraktheit dem Anfänger meistens erhebliche Schwierigkeiten bereitet.

Bei einer kommutativen Operation ist natürlich jedes linksseitig neutrale Element auch rechtsseitig neutral und umgekehrt, d. h., jedes einseitig neutrale Element ist beidseitig neutral. Daher besitzt eine kommutative Operation o entweder kein oder genau ein neutrales Element.

In den folgenden Definitionen bezeichnet R eine binäre Relation in M.

(12_l) *o linksseitig monoton bzgl. R* $: \Leftrightarrow \bigwedge_{x,y,z \in M} (xRy \Rightarrow (xoz)R(yoz))$,

(12_r) *o rechtsseitig monoton bzgl. R* $: \Leftrightarrow \bigwedge_{x,y,z \in M} (xRy \Rightarrow (zox)R(zoy))$,

(12) *o (beidseitig) monoton bzgl. R*
$: \Leftrightarrow$ *o linksseitig monoton bzgl. $R \wedge o$ rechtsseitig monoton bzgl. R.*

Besonders wichtig ist der Spezialfall, daß hierbei R eine reflexive oder irreflexive teilweise oder totale Ordnung ist. Für den Fall einer irreflexiven totalen Ordnung sei hier das folgende häufig benötigte Resultat genannt:

(13) *S irreflexive totale Ordnung in M* $\wedge$ *o linksseitig monoton bzgl. S*
$\Rightarrow \bigwedge_{x,y,z \in M} ((xoz)S(yoz) \Rightarrow xSy) \wedge \bigwedge_{x,y,z \in M} (xoz = yoz \Rightarrow x = y)$.

Die erste Behauptung besagt, daß man im Fall einer irreflexiven totalen Ordnung in (12_l) die Implikation „$\Rightarrow$“ umkehren kann, während die zweite Behauptung beinhaltet, daß man bei einer Operation o, die in bezug auf eine irreflexive totale Ordnung linksseitig monoton ist, eine Gleichung der Form $xoz = yoz$ stets durch z kürzen kann, daß o rechtsseitig kürzbar ist (siehe unten). Zum Beweis der ersten Behauptung seien x, y, z Elemente aus M mit $(xoz)S(yoz)$. Auf Grund der Konnexität von S tritt dann für die Elemente x, y wenigstens einer der Fälle xSy oder ySx oder $x = y$ ein. Wir müssen zeigen, daß die letzten beiden Fälle nicht möglich sind: Würde ySx gelten, dann wäre wegen der linksseitigen Monotonie von o bzgl. S auch $(yoz)S(xoz)$ erfüllt, was wegen der Asymmetrie von S im Widerspruch zu $(xoz)S(yoz)$ steht; wäre $x = y$, so wäre $xoz = yoz$, was wegen der Irreflexivität von S ebenfalls im Widerspruch zu $(xoz)S(yoz)$ steht. Analog beweist man die zweite Behauptung. Ein entsprechendes Resultat gilt natürlich auch für Operationen, die bzgl. einer irreflexiven totalen Ordnung rechtsseitig monoton sind.

Als Beispiel sei genannt, daß bekanntlich die Addition im Bereich der reellen Zahlen (wie z. B. auch die Multiplikation im Bereich der positiven reellen Zahlen) bzgl. der $<$-Relation monoton ist. Allein aus dieser Tatsache folgt nach (19) bereits, daß man aus $a + c < b + c$ stets auf $a < b$ und aus $a + c = b + c$ stets auf $a = b$ schließen kann (ohne also z. B. von der Subtraktion Gebrauch zu machen).

Bei einer reflexiven Ordnung R spricht man vielfach von *echter Monotonie*, wenn die betrachtete Operation bzgl. der zu R gehörigen irreflexiven Ordnung R_i monoton ist (echte Monotonie der Addition von reellen Zahlen bzgl. der $\leqq$-Relation bedeutet also Monotonie der Addition bzgl. der $<$-Relation).

Kehrt sich beim Resultat xoz die Relation um, so spricht man meistens von *Antimonotonie*, also z. B.

(14_l) *o linksseitig antimonoton bzgl.* $R \Leftrightarrow: \bigwedge_{x,y,z \in M} (xRy \Rightarrow (yoz)R(xoz))$.

Beispiel. Die Subtraktion von reellen Zahlen $(xoy := x - y)$ ist bezüglich der $<$-Relation linksseitig monoton und rechtsseitig antimonoton, die Multi-

plikation von reellen Zahlen ist, wenn man sie auf $\mathsf{R} \times \mathsf{R}^*_-$ (R^*_- Menge aller negativen reellen Zahlen) einschränkt, d. h. als partielle Operation in R mit dem Definitionsbereich $\mathsf{R} \times \mathsf{R}^*_-$ betrachtet, bezüglich der $<$-Relation linksseitig antimonoton ($x < y \wedge z < 0 \Rightarrow y \cdot z < x \cdot z$). Für bezüglich einer irreflexiven totalen Ordnung antimonotone Operationen gilt ein zu (13) analoges Resultat.

(15_l) *o linksseitig kürzbar* $: \Leftrightarrow \bigwedge_{z_1, z_2, y \in M} (yoz_1 = yoz_2 \Rightarrow z_1 = z_2)$,

(15_r) *o rechtsseitig kürzbar* $: \Leftrightarrow \bigwedge_{z_1, z_2, y \in M} (z_1oy = z_2oy \Rightarrow z_1 = z_2)$,

(15) *o beidseitig kürzbar*
$: \Leftrightarrow$ *o linksseitig kürzbar* $\wedge$ *o rechtsseitig kürzbar*.

Häufig tritt der Fall ein, daß z. B. (15_l) nicht für alle $z_1, z_2, y \in M$, sondern nur für $z_1, z_2 \in N_1 (\subseteq M)$ und $y \in N_2 (\subseteq M)$ erfüllt ist. In diesem Fall heißt o linksseitig kürzbar in $N_1 \times N_2$:

$(15'_l)$ *o linksseitig kürzbar in* $N_1 \times N_2$
$: \Leftrightarrow \bigwedge_{z_1, z_2 \in N_1} \bigwedge_{y \in N_2} (yoz_1 = yoz_2 \Rightarrow z_1 = z_2)$
(und analog bei (15_r) und (15)).

Bei einer z. B. rechtsseitig kürzbaren Operation o besitzt eine Gleichung der Form $zoy = x$ bei gegebenem $x, y \in M$ höchstens eine Lösung $z \in M$. Die Menge aller derjenigen geordneten Paare $(x, y) \in M \times M$, für die diese Gleichung lösbar ist, für die also ein $z \in M$ mit $zoy = x$ existiert, bezeichnen wir mit $D_l(o)$:

(16_l) $D_l(o) := \{(x, y) : \bigvee_{z \in M} zoy = x\}$.

Dann erhalten wir vermöge

$17_l)$ $xo_l y = z : \Leftrightarrow (x, y) \in D_l(o) \wedge z \in M \wedge zoy = x$.

eine im allgemeinen partielle Operation o_l in M $(D(o_l) = D_l(o))$, die wir die *linksseitige Umkehrung von o* nennen wollen. Entsprechend wird bei einer linksseitig kürzbaren Operation o in

(16_r) $D_r(o) := \{(x, y) : \bigvee_{z \in M} yoz = x\}$

durch

(17_r) $xo_r y = z : \Leftrightarrow (x, y) \in D_r(o) \wedge z \in M \wedge yoz = x$

eine im allgemeinen partielle Operation o_r in M $(D(o_r) = D_r(o))$ definiert, die die *rechtsseitige Umkehrung von o* genannt wird.

Bei einer kommutativen Operation folgt natürlich aus der linksseitigen Kürzbarkeit die rechtsseitige Kürzbarkeit und umgekehrt, so daß eine einseitig kürzbare kommutative Operation stets beidseitig kürzbar ist. Überdies ist bei einer (beidseitig) kürzbaren Operation o stets $o_r = o_l$; denn es ist $D(o_r) = D(o_l)$, und bei beliebigem $(x, y) \in D(o_r)$ gilt

$$xo_r y = z \Leftrightarrow yoz = x \Leftrightarrow zoy = x \Leftrightarrow xo_l y = z.$$

Wir wollen die letzten Definitionen an einigen Beispielen erläutern: Zunächst sei o die Addition in $\mathbf{R}$, d. h. $xoy := x + y$ $(x, y \in \mathbf{R})$. Diese Operation ist beidseitig kürzbar, wobei $o_r = o_l$ die Subtraktionsoperation ist, d. h. $xo_r y = xo_l y = x - y$. Die Operation $xoy := x \cdot y$ $(x, y \in \mathbf{R})$ ist beidseitig kürzbar in $\mathbf{R}^* \times \mathbf{R}^*$ $(\mathbf{R}^* = \mathbf{R} \setminus \{0\})$ (genauer: linksseitig kürzbar in $\mathbf{R}^* \times \mathbf{R}$ und rechtsseitig kürzbar in $\mathbf{R} \times \mathbf{R}^*$), wobei $o_r = o_l$ die Divisionsoperation ist, d. h. $xo_r y = xo_l y = x : y$ $(x \in \mathbf{R},\ y \in \mathbf{R}^*)$. Schließlich betrachten wir noch die durch $xoy := x^y\ (= e^{y \cdot \ln x})$ $(D(o) = \mathbf{R}_+^* \times \mathbf{R}^*,\ \mathbf{R}_+^* := \{x : x \in \mathbf{R} \wedge x > 0\})$ definierte (nicht kommutative) Operation. Diese Operation ist rechtsseitig kürzbar in $\mathbf{R}_+^* \times \mathbf{R}^*$ und linksseitig kürzbar in $(\mathbf{R}_+^* \setminus \{1\}) \times \mathbf{R}^*$; die Umkehroperation o_l wird gegeben durch $xo_l y = \sqrt[y]{x}$, die Umkehroperation o_r dagegen durch $xo_r y = \dfrac{\ln x}{\ln y}\ (= \log_y x)$.

2.7. Mathematische Strukturen

Es ist von Mathematikern und Philosophen immer wieder die Frage aufgeworfen worden, was denn eigentlich Mathematik sei, was Gegenstand der als Mathematik bezeichneten Wissenschaft ist. Wir glauben nicht, daß man auf diese Frage eine für die gesamte gegenwärtige Mathematik voll verbindliche Antwort geben kann. Grundsätzlich kann man aber wohl feststellen, daß sich jeder zusammenhängende Komplex von mathematischen Untersuchungen auf eine mehr oder minder fest umrissene sogenannte mathematische Struktur bzw. eine Klasse derartiger Strukturen bezieht. Dabei ist eine *mathematische Struktur* oder *allgemeine Algebra* durch eine Menge M, die *Trägermenge* der Struktur, gewisse *ausgezeichnete Elemente* $a_1, \ldots, a_k$ aus M und gewisse *Grundrelationen* $R_1, \ldots, R_m$ und *Grundoperationen* $o_1, \ldots, o_n$ in M jeweils einer bestimmten Stellenzahl $i_1, \ldots, i_m$ bzw. $j_1, \ldots, j_n$ festgelegt (in komplizierteren Fällen — vgl. den Anfang dieses Abschnitts — können es auch mehrere Trägermengen und Relationen und Operationen zwischen Elementen dieser unter-

schiedlichen Trägermengen sein). Es ist heute allgemein üblich, als *Struktur* oder *Algebra* ein $(k + m + n + 1)$-Tupel $(M, a_1, \ldots, a_k, R_1, \ldots, R_m, o_1, \ldots, o_n)$ zu bezeichnen, dessen Komponenten die angegebene Bedeutung haben (wobei auch die Fälle $k = 0$ und (oder) $m = 0$ und (oder) $n = 0$ zugelassen sind). Das besondere Charakteristikum mathematischer Untersuchungen besteht darin, daß man aus gewissen aus der Erfahrung (durch Abstraktion aus realen Verhältnissen) gewonnenen und häufig in Axiomen fixierten Eigenschaften der Elemente von M, die sich mittels $a_1, \ldots, a_k$, $R_1, \ldots, R_m$, $o_1, \ldots, o_n$ ausdrücken lassen, durch logische Schlüsse weitere derartige Eigenschaften ableitet. Etwa in diesem Sinne ist die heute häufig zu findende Formulierung zu interpretieren, daß die moderne Mathematik Strukturwissenschaft ist. Allerdings trägt die genannte Beschreibung noch in keiner Weise der Tatsache Rechnung, daß nicht die Untersuchung schlechthin jeder Struktur, die man sich irgendwie ausdenken mag, bedeutungsvoll ist (die mathematischen Begriffe sind keineswegs freie Schöpfungen des menschlichen Geistes, wie vom Idealismus behauptet wird, sondern abstrakter Ausdruck bestimmter Verhältnisse der materiellen Welt). Das Kriterium für die Bedeutung mathematischer Untersuchungen ist allein die Praxis, wobei man sich jedoch hüten muß, diese zu eng zu fassen und nur einseitig in direkten Anwendungen der Mathematik in Technik, Natur- und Gesellschaftswissenschaften zu sehen. Viele Untersuchungen der sogenannten reinen Mathematik (in der letzten Zeit hat sich dafür auch die nicht sehr geistvolle Bezeichnung „Theoretische Mathematik" eingebürgert) haben zunächst nur den Zweck, den Bestand an gesicherten mathematischen Erkenntnissen zu vergrößern. Sie schaffen indes — und hierin äußert sich ihre Bedeutung — zugleich einen Vorlauf, indem sie Resultate und vor allem Begriffe und Methoden für mögliche Anwendungen bereitstellen. Die Geschichte der Mathematik insbesondere der letzten 150 Jahre zeigt, daß zahllose mathematische Resultate und ganze mathematische Theorien, die zunächst durch rein theoretische Erwägung konzipiert wurden, unversehens größte Bedeutung für die Anwendung hatten. So setzte die Relativitätstheorie von Albert Einstein (1879—1955) unabdingbar die großartigen geometrischen Erkenntnisse von C. F. Gauss, Bernhard Riemann (1826 bis 1866) und vielen anderen voraus, die moderne Wahrscheinlichkeitsrechnung mit ihren zahlreichen Anwendungen in Natur- und Gesellschaftswissenschaften wäre ohne das Fundament z. B. der Maßtheorie von Henri Lebesgue (1875—1941) u. a. nicht denkbar, und weder die Konstruktion noch der Betrieb programmgesteuerter Rechenanlagen wäre möglich, hätten sich nicht in einer 100 Jahre währenden Entwicklung die hierfür notwendigen theoretischen Prinzipien herausgebildet. Aus diesem Grunde ist eine Einteilung der Mathematik in reine und angewandte Mathematik gegenwärtig

nicht mehr sinnvoll; beide sind heute eng miteinander verzahnt und bilden eine dialektische Einheit.

Im Zusammenhang mit der genannten Strukturauffassung gewinnen eine Reihe von allgemeinen Begriffen Bedeutung, die die gesamte heutige Mathematik durchziehen. Wir betrachten hier nur die drei wichtigsten, wobei wir uns der Einfachheit halber auf Strukturen beschränken, deren Grundrelationen und -operationen sämtlich die Stellenzahl 2 haben.

Eine Struktur $\mathcal{Z}' = (M', a_1', \ldots, a_k', R_1', \ldots, R_m', o_1', \ldots, o_n')$ heißt *Unter-* oder *Teilstruktur* der Struktur $\mathcal{Z} = (M, a_1, \ldots, a_k, R_1, \ldots, R_m, o_1, \ldots, o_n)$ (im allgemeinen Fall müssen $\mathcal{Z}$ und $\mathcal{Z}'$ gleiche *Signatur* haben, d. h. die einander entsprechenden Relationen bzw. Operationen gleiche Stellenzahl besitzen), wenn $M' \subseteq M$ ist und dabei folgendes gilt:

(1a) $a_\varkappa' = a_\varkappa \quad (\varkappa = 1, \ldots, k)$,

(1b) $\bigwedge_{x,y \in M'} (xR_\mu' y \Leftrightarrow xR_\mu y) \qquad (\mu = 1, \ldots, m)$,

(1c) $\bigwedge_{x,y \in M'} o_\nu'(x, y) = o_\nu(x, y) \qquad (\nu = 1, \ldots, n)$.

Die Bedingung (1b) kann offenbar auch als

(1b') $R_\mu' = R_\mu \cap (M' \times M') \qquad (\mu = 1, \ldots, m)$

geschrieben werden. Man nennt in diesem Fall R_μ' die *Einschränkung* von R_μ auf M' ($\subseteq M$). Aus (1c) folgt, daß für $\nu = 1, \ldots, n$ mit x, y stets auch $o_\nu(x, y)$ zu M' gehört (denn o_ν' ist nach Voraussetzung Operation in M'), wofür man auch sagt, daß M' bezüglich $o_1, \ldots, o_n$ *abgeschlossen* ist. Folglich ist jede Operation o_ν' die Einschränkung (vgl. 2.4.(9)) der entsprechenden Operation o_ν auf $M' \times M'$:

(1c') $o_\nu' = o_\nu \mid (M' \times M') \qquad (\nu = 1, \ldots, n)$.

Zum Beispiel ist im Fall

$M = \mathsf{R}$,	$M' = \mathsf{N}$,
$R_1 = \leqq$-Relation in R,	$R_1' = \leqq$-Relation in N,
$o_1 =$ Addition in R,	$o_1' =$ Addition in N

$(M', 0, R_1', o_1')$ Unterstruktur von $(M, 0, R_1, o_1)$.

Die Struktur $\mathcal{Z}' = (M', a_1', \ldots, a_k', R_1', \ldots, R_m', o_1', \ldots, o_n')$ heißt *isomorph* zur Struktur $\mathcal{Z} = (M, a_1, \ldots, a_k, R_1, \ldots, R_m, o_1, \ldots, o_n)$ (im allgemeinen Fall ist wieder vorauszusetzen, daß die einander entsprechenden Relationen und Operationen in $\mathcal{Z}$ und $\mathcal{Z}'$ dieselbe Stellenzahl haben) — man schreibt $\mathcal{Z} \cong \mathcal{Z}'$ — wenn es eine 1-1-Abbildung f von M auf M' gibt, die man dann

einen *Isomorphismus* oder eine *isomorphe Abbildung* nennt (man schreibt hierfür $f: \Sigma \overset{\sim}{\to} \Sigma'$), so daß folgendes gilt:

$$(2a) \qquad f(a_\varkappa) = a'_\varkappa \qquad (\varkappa = 1, \ldots, k),$$

$$(2b) \qquad \bigwedge_{x,y \in M} (f(x) R'_\mu f(y) \Leftrightarrow x R_\mu y) \qquad (\mu = 1, \ldots, m),$$

$$(2c) \qquad \bigwedge_{x,y,z \in M} (o'_\nu(f(x), f(y)) = f(z) \Leftrightarrow o_\nu(x, y) = z) \qquad (\nu = 1, \ldots, n),$$

wobei man statt (2c) offenbar auch

$$(2c') \qquad \bigwedge_{x,y \in M} o'_\nu(f(x), f(y)) = f(o_\nu(x, y)) \qquad (\nu = 1, \ldots, n)$$

schreiben kann. Für (2) sagt man auch, daß sich bei der Abbildung f die Grundelemente, -relationen und -operationen *übertragen*. Man zeigt leicht, daß *die Isomorphie von Strukturen eine Äquivalenzrelation ist:*

$$(3a) \qquad \Sigma \cong \Sigma,$$

$$(3b) \qquad \Sigma_1 \cong \Sigma_2 \wedge \Sigma_2 \cong \Sigma_3 \Rightarrow \Sigma_1 \cong \Sigma_3,$$

$$(3c) \qquad \Sigma \cong \Sigma' \Rightarrow \Sigma' \cong \Sigma.$$

Zunächst ist nämlich die identische Abbildung e_M ein Isomorphismus jeder Struktur Σ mit der Trägermenge M auf sich. Ist f_1 ein Isomorphismus von Σ_1 auf Σ_2, f_2 ein Isomorphismus von Σ_2 auf Σ_3, so ist $f_2 \circ f_1$ ein Isomorphismus von Σ_1 auf Σ_3. Ist schließlich f ein Isomorphismus von Σ auf Σ', so ist f^{-1} ein Isomorphismus von Σ' auf Σ.

Die Isomorphie zweier Strukturen Σ, Σ' beinhaltet, daß man (mittels des Isomorphismus f) jede Eigenschaft der Elemente der Trägermenge M von Σ bezüglich der Grundelemente, -relationen und -operationen von Σ in eine analoge Eigenschaft der Elemente der Trägermenge M' von Σ' bezüglich der jeweils entsprechenden Grundelemente, -relationen und -operationen von Σ' übersetzen kann, wobei auch umgekehrt (mittels f^{-1}) jeder Eigenschaft der Elemente von M' eine Eigenschaft der Elemente von M entspricht. Ist z. B. die Grundoperation o_1 von Σ rechtsseitig monoton bzgl. der Grundrelation R_1 von Σ, so ist die o_1 entsprechende Grundoperation o'_1 von Σ' rechtsseitig monoton bzgl. der R_1 entsprechenden Relation R'_1, und umgekehrt. Gilt nämlich für o_1 und R_1 2.6.(12_r) und sind x', y', z' beliebige Elemente aus M' mit $x' R'_1 y'$, so existieren wegen der Eineindeutigkeit von f eindeutig bestimmte Elemente $x, y, z \in M$, so daß $x' = f(x)$, $y' = f(y)$, $z' = f(z)$. Wegen (2b) und $x' R'_1 y'$ gilt dann $x R_1 y$, und wegen 2.6.(12_r) folgt hieraus $o_1(z, x) R_1 o_1(z, y)$. Wiederum wegen (2b) folgt hieraus $f(o_1(z, x)) R'_1 f(o_2(z, y))$, und wegen (2c') ist

$$f(o_1(z, x)) = o'_1(f(z), f(x)) = o'_1(z', x')$$

und analog $f(o_1(z,y)) = o_1'(z',y')$. Also gilt $o_1'(z',x')\,R_1'o_1'(z',y')$ für alle $x',y',z' \in M'$ mit $x'R_1'y'$, und das besagt ja gerade, daß o_1' bzgl. R_1' rechtsseitig monoton ist. Ist das Grundelement a_1 von Σ linksseitig neutrales Element für die Grundoperation o_1 von Σ, so ist a_1' linksseitig neutrales Element für o_1', und umgekehrt usw.

Eine Abschwächung des Isomorphiebegriffs ist der Begriff der Homomorphie. Dabei heißt eine Struktur Σ' *homomorphes Bild* der Struktur Σ, wenn es eine eindeutige (also nicht notwendig eineindeutige) Abbildung f der Trägermenge M von Σ auf die Trägermenge M' von Σ' gibt — sie heißt dann ein *Homomorphismus* oder eine *homomorphe Abbildung* von Σ auf Σ' (in Zeichen $f: \Sigma \asymp \Sigma'$) —, so daß die Bedingungen (2) erfüllt sind. Wir bemerken, daß *die Homomorphie von Strukturen reflexiv und transitiv*, aber natürlich im allgemeinen nicht mehr symmetrisch ist.

Zur Einübung der zuletzt eingeführten abstrakten Begriffsbildungen wollen wir noch ein häufig benötigtes allgemeines Strukturtheorem beweisen, das allgemeines Homomorphietheorem genannt wird. Dazu seien $\Sigma = (M, a_1, \ldots, a_k, R_1, \ldots, R_m, o_1, \ldots, o_n)$ und $\Sigma' = (M', a_1', \ldots, a_k', R_1', \ldots, R_m', o_1', \ldots, o_n')$ Strukturen und f ein Homomorphismus von Σ auf Σ'. Es bezeichne R_f die von f in M induzierte Äquivalenzrelation (vgl. 2.5.(20)), für die also bei beliebigem $x, y \in M$ gilt:

(4) $\qquad xR_fy \Leftrightarrow f(x) = f(y),$

und es sei M/R_f das Restsystem dieser Äquivalenzrelation. In M/R_f erklären wir nun auf folgende Weise Relationen $\bar{R}_\mu (\mu = 1, \ldots, m)$. Es seien ξ, η beliebige Elemente aus M/R_f, d. h. Klassen von untereinander bzgl. R_f äquivalenten Elementen aus M. Wir wählen dann aus jeder dieser Klassen einen beliebigen „Repräsentanten" x^*, y^* aus ($x^* \in \xi$, $y^* \in \eta$) und setzen fest, daß gelten soll:

(5) $\qquad \xi \bar{R}_\mu \eta :\Leftrightarrow x^* R_\mu y^* \qquad (\mu = 1, \ldots, m).$

Wir müssen uns zunächst überlegen, daß *diese Definition unabhängig ist von der speziellen Auswahl der Repräsentanten aus den Klassen* ξ, η, d. h., daß wir zu keinem anderen Resultat gelangen, wenn wir statt x^*, y^* andere Repräsentanten x, y nehmen. Es seien also x, y evtl. andere Elemente aus den Klassen ξ, η. Nach Definition der Klassen folgt dann: xR_fx^* und yR_fy^*, also nach (4) $f(x) = f(x^*)$, $f(y) = f(y^*)$. Gilt nun $x^*R_\mu y^*$, so gilt nach der Homomorphiebedingung (2b) auch $f(x^*)\,R_\mu' f(y^*)$ und folglich $f(x)R_\mu' f(y)$, also wiederum nach (2b) $xR_\mu y$, wobei alle diese Schlüsse umkehrbar sind. Also liefert in der Tat die Verwendung von x, y an Stelle von x^*, y^* bei (5) nichts anderes. Auf ähnliche Weise erklären wir in M/R_f Operationen $\bar{o}_\nu$ ($\nu = 1, \ldots, n$):

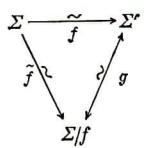
(6) $$\bar{o}_\nu(\xi, \eta) = \zeta :\Leftrightarrow \bigvee_{x^*, y^*, z^*} (x^* \in \xi \wedge y^* \in \eta \wedge z^* \in \zeta \wedge o_\nu(x^*, y^*) = z^*) \qquad (\nu = 1, \ldots, n).$$

Wie bei (5) zeigt man mittels (2c), daß auch diese Definition unabhängig von der speziellen Auswahl der Repräsentanten x^*, y^*, z^* aus den Klassen ξ, η, ζ ist. Schließlich setzen wir:

(7) $\qquad \bar{a}_\varkappa := \{x : x \in M \wedge xR_fa_\varkappa\} \qquad (\varkappa = 1, \ldots, k).$

Die Struktur $\bar{\Sigma} = (M/R_f, \bar{a}_1, \ldots, \bar{a}_k, \bar{R}_1, \ldots, \bar{R}_m, \bar{o}_1, \ldots, \bar{o}_n)$ nennt man die durch den Homomorphismus f erzeugte *Restklassen-* oder *Faktorstruktur* und schreibt $\bar{\Sigma} = \Sigma/R_f$ der kurz $\bar{\Sigma} = \Sigma/f$. Diese Faktorstruktur hat nun die bemerkenswerte Eigenschaft, und das ist gerade die Aussage des allgemeinen Homomorphietheorems, daß *die kanonische Abbildung* $\bar{f}$ (vgl. 2.5.(21)), *die einem beliebigen* $x \in M$ *die* x *enthaltende Restklasse aus* M/R_f *zuordnet*, wegen (5), (6), (7) *ein Homomorphismus von* Σ *auf* $\bar{\Sigma} = \Sigma/f$ *ist, während die Abbildung* g (vgl. 2.5.(22)), *die einer beliebigen Restklasse aus* M/R_f *das allen Elementen dieser Restklasse gemeinsame Bild bei* f *zuordnet, sich leicht* (Beweis!) *als Isomorphismus von* $\bar{\Sigma} = \Sigma/f$ *auf* Σ' *erweist.* Jeder Homomorphismus f von Σ auf Σ' ist also Verkettung $g \circ \bar{f}$ des kanonischen Homomorphismus $\bar{f}$ von Σ auf eine Faktorstruktur $\bar{\Sigma}$ und eines Isomorphismus von $\bar{\Sigma}$ auf Σ', was man sich gern (vgl. 2.5.(23')) auch durch das Diagramm

(8)



veranschaulicht.

Die vorangehenden Überlegungen sind im gewissen Sinne umkehrbar. Dazu merken wir zunächst an, daß wir bei der Definition von Σ/R_f neben der Tatsache, daß R_f eine Äquivalenzrelation ist, nur die Unabhängigkeit der Definitionen (5) und (6) von der Wahl der Repräsentanten x^*, y^* (z^*) aus den Restklassen ξ, η (ζ) benötigten. Analoges ist nun bei einer beliebigen Äquivalenzrelation R in M offenbar genau dann der Fall, wenn sie die folgenden *Verträglichkeitsbedingungen* erfüllt:

$$(9) \qquad \bigwedge_{x,y,x',y' \in M} (xRx' \wedge yRy' \Rightarrow (xR_\mu y \Leftrightarrow x'R_\mu y')) \qquad (\mu = 1, \ldots, m),$$

$$(10) \qquad \bigwedge_{x,y,x',y' \in M} (xRx' \wedge yRy' \Rightarrow o_\nu(x,y) R o_\nu(x',y')) \qquad (\nu = 1, \ldots, n).$$

Eine Äquivalenzrelation R in M, für die (9) und (10) gilt, heißt eine *Kongruenzrelation* der Struktur $\Sigma = (M, a_1, \ldots, a_k, R_1, \ldots, R_m, o_1, \ldots, o_n)$, und die vorangehenden Überlegungen lehren, daß man zu jeder Kongruenzrelation R einer Struktur Σ die Restklassenstruktur Σ/R bilden kann. Man zeigt leicht (Beweis!), daß *in jedem solchen Fall die kanonische Abbildung* $f_{M/R}$ (vgl. 2.5.(18)) *ein Homomorphismus von* Σ *auf* Σ/R *ist und jede Verkettung dieses kanonischen Homomorphismus mit einem Isomorphismus von* Σ/R *auf eine beliebige Struktur* Σ' *einen Homomorphismus von* Σ *auf* Σ' *ergibt.*

Die in 2.5. für „reine Mengen" durchgeführten Überlegungen können gewissermaßen als der Spezialfall $k = m = n = 0$ der vorangehenden Ausführungen angesehen werden.

2.8. Das Auswahlprinzip

Alle bisher betrachteten Mengenbildungen erfolgten mit Hilfe des in 1.2. behandelten Mengenbildungsprinzips mittels einer die Elemente der jeweiligen Menge charakterisierenden Eigenschaft $H(x)$. Daneben hat man in der Mathematik häufig Mengen zu bilden, für deren Elemente sich nicht ohne weiteres eine gemeinsame Eigenschaft angeben läßt. Diese Mengenbildungen benutzen ein erstmalig im Jahre 1904 durch den in Göttingen (später in Zürich und Freiburg i. Br.) wirkenden ERNST ZERMELO (1871—1953) klar formuliertes Prinzip, das man nach ihm als (Zermelosches) Auswahlprinzip oder Auswahlaxiom bezeichnet. In seiner einfachsten Form hat es den folgenden einfachen Sachverhalt zum Inhalt: *Gegeben sei ein nichtleeres System $\mathfrak{M}$ aus paarweise disjunkten nichtleeren Mengen; dann gibt es stets (wenigstens) eine Menge A* (eine sogenannte *Auswahlmenge* für $\mathfrak{M}$), *die aus jeder Menge M des Systems $\mathfrak{M}$ genau ein Element a_M in dem Sinne auswählt, daß A mit M genau das Element a_M gemeinsam hat.* Die Voraussetzung $\mathfrak{M} \neq \emptyset$ schließt den Trivialfall $\mathfrak{M} = \emptyset$ aus, in welchem natürlich $A = \emptyset$ eine Auswahlmenge ist. Die Voraussetzung, daß die leere Menge nicht in $\mathfrak{M}$ enthalten sein soll, ist bei der angegebenen Formulierung des Auswahlaxioms wesentlich, da natürlich keine Menge A aus der leeren Menge ein Element auswählen kann. Ebenso ist die Voraussetzung, daß $\mathfrak{M}$ disjunkt ist, wesentlich. Ist nämlich $\mathfrak{M} = \{M_1, M_2, M_3\}$ mit

$$M_1 \neq \emptyset,\ M_2 \neq \emptyset,\ M_1 \cap M_2 = \emptyset,\ M_1 \subseteq M_3,\ M_2 \subseteq M_3$$

(beispielsweise $M_3 = M_1 \cup M_2$), so hat jede Menge A, die mit M_1 und M_2 je genau ein Element a_{M_1} bzw. a_{M_2} gemeinsam hat, mit M_3 die beiden verschiedenen ($M_1 \cap M_2 = \emptyset$!) Elemente a_{M_1} und a_{M_2} gemeinsam. In Abkürzungstechnik drückt sich das Auswahlprinzip folgendermaßen aus:

$$(1) \qquad \mathfrak{M} \neq \emptyset \wedge \bigwedge_M (M \in \mathfrak{M} \Rightarrow M \neq \emptyset) \wedge \mathfrak{M} \textit{ disjunkt}$$

$$\Rightarrow \bigvee_A \bigwedge_M (M \in \mathfrak{M} \Rightarrow \bigvee_a M \cap A = \{a\}).$$

Wir merken an, daß man stets erreichen kann, daß die Auswahlmenge A nur Elemente enthält, die in wenigstens einer Menge des Systems $\mathfrak{M}$ vorkommen, d. h. $A \subseteq \bigcup \mathfrak{M}$ gilt (was in (1) nicht unbedingt verlangt ist); ist nämlich A' eine nach (1) existierende Auswahlmenge, für die das nicht der Fall ist, so wird $A = A' \cap \bigcup \mathfrak{M}$ eine Auswahlmenge, die diese zusätzliche Bedingung erfüllt. Wir merken weiterhin an, daß es zu einem Mengensystem $\mathfrak{M}$, das die Voraussetzungen des Auswahlprinzips erfüllt, im allgemeinen (wenn nämlich $\mathfrak{M}$ nicht

nur aus Einermengen besteht, in diesem Fall ist übrigens $\bigcup \mathfrak{M}$ eine Auswahlmenge für $\mathfrak{M}$) mehrere, unter Umständen sehr viele Auswahlmengen gibt. Im Gegensatz zur Bildung von Mengen mit Hilfe des Mengenbildungsprinzips ist also die Mengenbildung mittels Auswahlprinzip nicht eindeutig.

Es sind heute zahlreiche äquivalente Formulierungen zum Auswahlprinzip bekannt. Wir wollen uns hier auf den Beweis einer häufig benötigten anderen Fassung des Auswahlprinzips beschränken, daß *es nämlich zu jeder mehrdeutigen Korrespondenz F von einer Menge M in eine Menge N eine eindeutige Korrespondenz (also Funktion) f von M in N mit* $f \subseteq F$ *gibt* (Abb. 8):

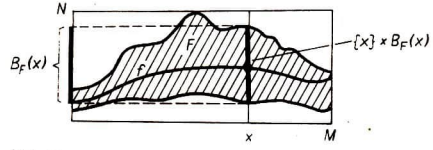


Abb. 8

(2) F *Korrespondenz von* M *in* N
$\Rightarrow \bigvee_f$ (f *eindeutige Korrespondenz von* M *in* $N \wedge f \subseteq F$).

Eine Funktion f mit der angegebenen Eigenschaft heißt eine *Auswahlfunktion* für F. Zum Beweis (wir nehmen o. B. d. A. an, daß $M \neq \emptyset$ ist), betrachten wir das folgende Mengensystem $\mathfrak{M}_F$:

$$\mathfrak{M}_F := \{\{x\} \times B_F(x) : x \in M\}. \tag{3}$$

Es gilt: (i) $\mathfrak{M}_F \neq \emptyset$ (da $M \neq \emptyset$). (ii) $\emptyset \notin \mathfrak{M}_F$; ist nämlich $\{x\} \times B_F(x) \in \mathfrak{M}_F$, so ist $x \in M$ und daher (F sollte Korrespondenz von M in N sein!) $B_F(x) \neq \emptyset$, also auch $\{x\} \times B_F(x) \neq \emptyset$. (iii) $\mathfrak{M}_F$ ist disjunkt; sind nämlich $\{x_1\} \times B_F(x_1)$, $\{x_2\} \times B_F(x_2)$ zwei verschiedene Elemente des Systems $\mathfrak{M}_F$, so ist $x_1 \neq x_2$, und dann ist $(\{x_1\} \times B_F(x_1)) \cap (\{x_2\} \times B_F(x_2)) = \emptyset$. Folglich gibt es nach (1) für $\mathfrak{M}_F$ eine Auswahlmenge A. Diese Menge A (sie ist eine Menge von Paaren (x, y) mit $x \in M$ und $y \in N$) ist zugleich eine Korrespondenz f mit der in (2) geforderten Eigenschaft (Beweis!).

Obwohl die im Auswahlprinizp (1) (bzw. in (2)) fixierte Forderung sicher außerordentlich einleuchtend ist, wird es gegenwärtig keineswegs von allen Mathematikern voll anerkannt. Es zeigt sich jedoch, daß man viele grundlegende Resultate der Mathematik ohne seine Hilfe nicht erhalten kann. Andererseits gelangt man gerade mit Hilfe des Auswahlaxioms zu einer Reihe von Ergebnissen, die recht paradox anmuten (indes keineswegs echte Widersprüche darstellen). Auf genauere Zusammenhänge können wir hier nicht eingehen.

2.9. Endliche Mengen

Nachdem in den vorangehenden Betrachtungen bereits gelegentlich der Begriff der endlichen Menge auftauchte, den wir dort stets im naiven, anschaulichen Sinne verwendeten, wollen wir uns jetzt einer etwas genaueren Untersuchung dieses Begriffs zuwenden. Es werden im folgenden drei einander äquivalente Definitionen der endlichen Menge betrachtet. Wir beginnen mit einer Definition, die inhaltlich wohl am nächstliegenden ist, aber den Nachteil hat, daß sie den Begriff der natürlichen Zahl benutzt, den wir systematisch erst im nächsten Kapitel behandeln werden. Die zweite Definition, die im wesentlichen auf BERTRAND RUSSELL zurückgeht und die nicht den Begriff der natürlichen Zahl verwendet, ist zwar etwas abstrakter, aber auch noch recht einleuchtend. Die dritte Definition benutzt eine merkwürdige, von dem Braunschweiger Mathematiker RICHARD DEDEKIND (1831—1916) entdeckte einfache charakteristische Eigenschaft der endlichen Mengen. Daneben gibt es in der Literatur noch zahlreiche weitere Endlichkeitsdefinitionen, auf die wir hier jedoch nicht eingehen können.

Mit N bezeichnen wir wie bisher die Menge aller natürlichen Zahlen einschließlich Null, d. h. der Zahlen $0, 1, 2, \ldots$. Ist n eine beliebige natürliche Zahl, so bezeichnet man die Menge aller Zahlen $m \in \mathsf{N}$, die kleiner als n sind, für die also $m < n$ gilt, als den durch die Zahl n bestimmten *Abschnitt* der Menge der natürlichen Zahlen; wir wollen ihn im folgenden mit $\mathcal{A}(n)$ bezeichnen.

$$(1) \qquad \mathcal{A}(n) := \{m : m \in \mathsf{N} \wedge m < n\}.$$

Der Abschnitt $\mathcal{A}(n)$ besteht also aus den Zahlen $0, 1, 2, \ldots, n-1$ und nur diesen. Auf Grund unserer Definition ist klar, daß auch die leere Menge ein Abschnitt der Menge der natürlichen Zahlen ist, nämlich der durch die Zahl 0 bestimmte Abschnitt $\mathcal{A}(0)$ (es gibt keine Zahl $m \in \mathsf{N}$, die kleiner als 0 ist).

Nun zur exakten Definition der endlichen Mengen: Im anschaulichen Sinne ist eine nichtleere Menge M genau dann endlich, wenn es eine natürliche Zahl n gibt, so daß man die Elemente der Menge M mit Hilfe der Zahlen $1, 2, \ldots, n$ oder — was für unsere Zwecke meistens bequemer ist — mit Hilfe der Zahlen $0, 1, \ldots, n-1$ durchnumerieren kann. Daneben gilt natürlich auch die leere Menge als endlich. Dabei bedeutet *durchnumerieren*, daß man eine 1-1-Abbildung zwischen den Elementen der Menge M und den Zahlen $0, 1, \ldots, n-1$ herstellen kann, so daß jedes Element aus M eine eindeutig bestimmte Nummer $0, 1, \ldots, n-1$ besitzt, und auch umgekehrt zu jeder Nummer ein Element gehört, das diese Nummer besitzt. Dieser Sachverhalt wird aber andererseits gerade durch $M \sim \mathcal{A}(n)$ wiedergegeben (vgl. 2.4.(21)). Damit gelangen wir zu der folgenden

Definition. *Eine Menge M heißt endlich, wenn es eine natürliche Zahl n gibt, so daß die Menge M dem Abschnitt $\mathcal{A}(n)$ gleichmächtig ist:*

$$(2) \qquad M \text{ endlich} :\Leftrightarrow \bigvee_n (n \in \mathsf{N} \wedge M \sim \mathcal{A}(n)).$$

Die leere Menge braucht hierbei nicht mehr besonders berücksichtigt zu werden; denn wegen $\mathcal{A}(0) = \emptyset$ ist auf Grund der Reflexivität der Gleichmächtigkeit nach (2) auch die leere Menge endlich (hierbei ist natürlich wesentlich, daß wir 0 als natürliche Zahl ansehen).

Wir weisen darauf hin, daß wir in (2) nur verlangt haben, daß die Menge M wenigstens einem Abschnitt $\mathcal{A}(n)$ gleichmächtig ist. Man zeigt leicht, daß *jede endliche Menge auch nur höchstens einem Abschnitt gleichmächtig sein kann.* Also gilt: *Jede endliche Menge ist genau einem Abschnitt $\mathcal{A}(n)$ gleichmächtig.* Zum Beweis dieser Behauptung nehmen wir an, es sei $M \sim \mathcal{A}(n_1)$ und $M \sim \mathcal{A}(n_2)$, und zeigen, daß dann $n_1 = n_2$ gilt. Hierzu genügt es wegen der Transitivität und Symmetrie der Gleichmächtigkeit (vgl. 2.4.(22)) zu zeigen, daß aus der Gleichmächtigkeit von Abschnitten $\mathcal{A}(n_1)$, $\mathcal{A}(n_2)$ die Gleichheit der sie bestimmenden Zahlen n_1, n_2 folgt:

$$(3) \qquad \mathcal{A}(n_1) \sim \mathcal{A}(n_2) \Rightarrow n_1 = n_2.$$

Dem Beweis von (3) schicken wir den folgenden einfachen abbildungstheoretischen Hilfssatz voraus:

$$(4) \qquad \begin{aligned} &f \text{ 1-1-Abbildung von } M \text{ auf } N \wedge x \in M \wedge y \in N \\ &\Rightarrow \bigvee_g (g \text{ 1-1-Abbildung von } M \text{ auf } N \wedge g(x) = y). \end{aligned}$$

Ist nämlich f 1-1-Abbildung von M auf N mit $f(x) = y'$, $f(x') = y$ ($x, x' \in M$; $y, y' \in N$), so wird

$$g := (f \setminus \{(x, y'), (x', y)\}) \smile \{(x', y'), (x, y)\}$$

eine 1-1-Abbildung von M auf N mit $g(x) = y$ (im Fall $f(x) = y$, in dem bereits f die gewünschte Eigenschaft hat, wird offenbar $g = f$).

Beim Beweis von (3) verwenden wir die aus der Schule bekannte Beweismethode der vollständigen Induktion, deren systematische Begründung wir erst im folgenden Abschnitt geben werden, und zwar führen wir den Beweis durch vollständige Induktion z. B. über n_1. Der Anfangsschritt $n_1 = 0$ ist trivial, da $\mathcal{A}(0)$ gleich der leeren Menge und diese nur sich selbst gleichmächtig ist, so daß $\mathcal{A}(n_2) = \emptyset$ und daher auch $n_2 = 0$ sein muß (denn $\mathcal{A}(0)$ ist der einzige leere Abschnitt!). Wir nehmen nun an (Induktionsvoraussetzung), daß (3) für $n_1 = n$ bei beliebigem n_2 schon bewiesen ist, und zeigen (Induktionsbehauptung), daß dann (3) auch für $n_1 = n + 1$ bei beliebigem n_2 gilt. Es sei also $\mathcal{A}(n + 1) \sim \mathcal{A}(n_2)$, und es sei f eine 1-1-Abbildung von $\mathcal{A}(n + 1)$ auf $\mathcal{A}(n_2)$.

Wegen $n \in \mathcal{A}(n+1)$ ist $\mathcal{A}(n_2) \neq \emptyset$, also $n_2 \neq 0$ und folglich $n_2 - 1 \in \mathcal{A}(n_2)$. Nach (4) gibt es eine 1-1-Abbildung g von $\mathcal{A}(n+1)$ auf $\mathcal{A}(n_2)$ mit $g(n) = n_2 - 1$, und dann ist $g \setminus \{(n, n_2 - 1)\}$ eine 1-1-Abbildung von $\mathcal{A}(n)$ auf $\mathcal{A}(n_2 - 1)$. Folglich muß nach Induktionsvoraussetzung $n = n_2 - 1$ und damit $n + 1 = n_2$ sein, womit die Induktionsbehauptung bewiesen ist.

Die eindeutig bestimmte natürliche Zahl n, für die für eine gegebene endliche Menge M die Bedingung $M \sim \mathcal{A}(n)$ erfüllt ist, heißt die *Elementeanzahl* oder *Kardinalzahl* von M und wird mit $|M|$ bezeichnet; in der Literatur sind auch die Bezeichnungen $\mathrm{card}(M)$ und (nach CANTOR) $\overline{\overline{M}}$ üblich. Es gilt also:

(5) $\quad |M| = n :\Leftrightarrow M \sim \mathcal{A}(n).$

In diesem Sinne besitzt die leere Menge die Elementeanzahl 0, alle Einermengen $\{a\}$ haben die Elementeanzahl 1, alle Zweiermengen $\{a, b\}$ mit $a \neq b$ haben die Elementeanzahl 2 usw. Allgemein gilt

(6) $\quad M$ *endlich* $\wedge\, x \notin M \Rightarrow M \cup \{x\}$ *endlich* $\wedge\, |M \cup \{x\}| = |M| + 1.$

Ist nämlich $|M| = n$ und f eine 1-1-Abbildung von M auf $\mathcal{A}(n)$, so wird im Fall $x \notin M$ (im Fall $x \in M$ ist $M \cup \{x\} = M$) offenbar $f \cup \{(x, n)\}$ eine 1-1-Abbildung von $M \cup \{x\}$ auf $\mathcal{A}(n+1)$, und damit gilt in der Tat $|M \cup \{x\}| = n + 1 = |M| + 1$.

Mittels (4) erhält man leicht das folgende Gegenstück zu (6):

(7) $\quad M$ *endlich* $\wedge\, x \in M \Rightarrow M \setminus \{x\}$ *endlich* $\wedge\, |M \setminus \{x\}| = |M| - 1.$

Ist nämlich $|M| = n$ und f eine 1-1-Abbildung von M auf $\mathcal{A}(n)$, so muß wegen $x \in M$ offenbar $\mathcal{A}(n) \neq \emptyset$, also $n \neq 0$ sein. Mithin ist $n - 1 \in \mathcal{A}(n)$, und es gibt nach (4) eine 1-1-Abbildung g von M auf $\mathcal{A}(n)$ mit $g(x) = n - 1$. Dann ist aber $g \setminus \{(x, n-1)\}$ eine 1-1-Abbildung von $M \setminus \{x\}$ auf $\mathcal{A}(n-1)$, so daß in der Tat $|M \setminus \{x\}| = n - 1 = |M| - 1$ wird.

Durch vollständige Induktion über die Elementeanzahl $|M|$ von M kann man (7) zu dem folgenden naheliegenden Satz verallgemeinern:

(7′) $\quad M$ *endlich* $\wedge\, N \subseteq M \Rightarrow N$ *endlich* $\wedge\, |N| \leqq |M|.$

Wegen der Reflexivität der Gleichmächtigkeit sind alle Abschnitte $\mathcal{A}(n)$ endlich, und es gilt $|\mathcal{A}(n)| = n$. Ferner sieht man auf Grund der Transitivität der Gleichmächtigkeit sofort, daß *jede zu einer endlichen Menge M gleichmächtige Menge N endlich ist und dieselbe Elementeanzahl wie jene hat:*

(8) $\quad M$ *endlich* $\wedge\, N \sim M \Rightarrow N$ *endlich* $\wedge\, |N| = |M|.$

Weitere wichtige Anzahlformeln werden wir in 3.6. behandeln.

Wir kommen nun zur Russellschen Definition der endlichen Mengen. Dazu sei E ein beliebiger Grundbereich und es bezeichne $\mathfrak{H}(E)$ das System aller im Sinne

von (2) endlichen Mengen aus Elementen des Bereichs E:

$$\mathfrak{F}(E) := \{X : X \subseteq E \wedge X \text{ endlich}\}.$$

Dann gilt

$$(9) \qquad \mathfrak{F}(E) \subseteq \mathfrak{P}(E) \wedge \emptyset \in \mathfrak{F}(E) \wedge \bigwedge_X \bigwedge_x (X \in \mathfrak{F}(E) \wedge x \in E \Rightarrow X \cup \{x\} \in \mathfrak{F}(E)).$$

Ein System $\mathfrak{M}$ von Mengen aus Elementen des Bereichs E mit den Eigenschaften (9) wollen wir kurz ein *induktives Mengensystem über E* nennen:

$$(9') \qquad \mathfrak{M} \textit{ induktiv über } E \Leftrightarrow \mathfrak{M} \subseteq \mathfrak{P}(E) \wedge \emptyset \in \mathfrak{M} \wedge \bigwedge_X \bigwedge_x (X \in \mathfrak{M} \wedge x \in E \Rightarrow X \cup \{x\} \in \mathfrak{M}).$$

Nach (9) ist $\mathfrak{F}(E)$ ein induktives Mengensystem über E; ebenso ist natürlich auch das System $\mathfrak{P}(E)$ aller Mengen aus Elementen des Bereichs E induktiv über E. Wir zeigen, daß $\mathfrak{F}(E)$ *das bzgl. Inklusion kleinste induktive Mengensystem über E ist* (während natürlich $\mathfrak{P}(E)$ das größte derartige System ist):

$$(10) \qquad \mathfrak{M} \textit{ induktiv über } E \Rightarrow \mathfrak{F}(E) \subseteq \mathfrak{M},$$

d. h.

$$(10') \qquad \mathfrak{M} \textit{ induktiv über } E \wedge M \textit{ endlich} \wedge M \subseteq E \Rightarrow M \in \mathfrak{M}.$$

Den Beweis für (10′) führen wir durch vollständige Induktion über $|M|$. Ist $|M| = 0$, also $M = \emptyset$, so gilt die Behauptung auf Grund der Forderung, daß das induktive Mengensystem $\mathfrak{M}$ die leere Menge als Element enthält. Wir nehmen nun an, (10′) sei schon für alle Teilmengen von E mit der Elementeanzahl n bewiesen, und es sei M eine Teilmenge von E mit $|M| = n + 1$ und f eine 1-1-Abbildung von $\mathcal{A}(n + 1)$ auf M. Dann ist offenbar $f\{(n, f(n))\}$ eine 1-1-Abbildung von $\mathcal{A}(n)$ auf $M\backslash\{f(n)\}$, also $M\backslash\{f(n)\}$ eine Teilmenge von E mit $|M\backslash\{f(n)\}| = n$. Nach Induktionsvoraussetzung ist daher $M\backslash\{f(n)\} \in \mathfrak{M}$, und da das induktive Mengensystem $\mathfrak{M}$ mit einer Menge X bei beliebigem $x \in E$ stets auch $X \cup \{x\}$ enthält, ist folglich $M = (M\backslash\{f(n)\}) \cup \{f(n)\}$ Element von $\mathfrak{M}$, was zu zeigen war.

Nach (10) ist also eine beliebige endliche Menge M von Elementen aus E in jedem induktiven Mengensystem über E als Element enthalten. Umgekehrt gehört natürlich eine Menge M, die in jedem induktiven Mengensystem enthalten ist, speziell dem als induktiv erkannten System $\mathfrak{F}(E)$ aller endlichen Mengen über E an und ist mithin endlich. Also gilt

$$(11) \qquad M \textit{ endlich} \Leftrightarrow \bigwedge_{\mathfrak{M}} (\mathfrak{M} \textit{ induktiv} \Rightarrow M \in \mathfrak{M}).$$

An dieser Äquivalenz ist nun bemerkenswert, daß auf der rechten Seite nur Begriffsbildungen der allgemeinen Mengenlehre auftreten, es sich also um eine rein mengentheoretische Charakterisierung der endlichen Mengen handelt, die den Begriff der natürlichen Zahl nicht mehr verwendet. Definiert man

$$(11') \qquad M \textit{ endlich } (R) :\Leftrightarrow \bigwedge_{\mathfrak{M}} (\mathfrak{M} \textit{ induktiv} \Rightarrow M \in \mathfrak{M})$$

(wobei sich diese Definition – wie in 1.2. generell verabredet – auf Mengen und Mengensysteme über einem fixierten Grundbereich E bezieht), so wird dadurch ein gewisser abstrakter Begriff der allgemeinen Mengenlehre erklärt, von dem sich

allerdings zeigt – und das ist der Inhalt des Satzes (!) (11) –, daß der umfangsgleich dem durch (2) mittels natürlicher Zahlen definierten Begriff der endlichen Menge ist:

(11) M *endlich* $(R) \Leftrightarrow M$ *endlich*.

Es ist nun nicht schwer, auf der Grundlage der Definition (11′) die anschaulich mehr oder minder evidenten Eigenschaften endlicher Mengen als Sätze der reinen Mengenlehre zu formulieren und zu beweisen. Wir wollen das an einigen Beispielen zeigen:

(12) M_1 *endlich* $(R) \wedge M_2$ *endlich* $(R) \Rightarrow M_1 \cup M_2$ *endlich* (R).

Wir merken an, daß hierbei M_1, M_2 und alle im folgenden Beweis auftretenden Mengen und Mengensysteme solche über demselben Grundbereich E sein sollen. Zum Beweis von (12) betrachten wir bei fester im Russellschen Sinne endlichen Menge M_1 das System $\mathfrak{M}^*$ aller derjenigen Mengen X, für die $M_1 \cup X$ im Russellschen Sinne endlich ist:

$$\mathfrak{M}^* := \{X : M_1 \cup X \text{ endlich } (R)\}.$$

Wir zeigen

(12′) $\mathfrak{M}^*$ *ist induktiv* (*über* E).

Damit ist offenbar (12) bewiesen; ist nämlich M_2 eine beliebige im Russellschen Sinne endliche Menge, so ist M_2 nach (11′) Element jedes induktiven Mengensystems $\mathfrak{M}$, damit wegen (12′) Element von $\mathfrak{M}^*$, und das besagt ja nach Definition von $\mathfrak{M}^*$ gerade, daß $M_1 \cup M_2$ im Russellschen Sinne endlich ist. Zum Beweis von (12′) ist folgendes zu zeigen:

(i) $\emptyset \in \mathfrak{M}^*$,

(ii) $X \in \mathfrak{M}^* (\wedge\, x \in E) \Rightarrow X \cup \{x\} \in \mathfrak{M}^*$.

Im vorliegenden Fall bedeutet (i), daß $M_1 \cup \emptyset = M_1$ im Russellschen Sinne endlich ist, was wir ja gerade vorausgesetzt haben. Zum Beweis von (ii) ist zu zeigen:

(iii) $M_1 \cup X$ *endlich* $(R) \Rightarrow M_1 \cup (X \cup \{x\})$ *endlich* (R).

Dazu sei $\mathfrak{M}_0$ ein beliebiges induktives Mengensystem. Nach Voraussetzung von (iii) ist dann $M_1 \cup X \in \mathfrak{M}_0$ und folglich nach (9′) auch $M_1 \cup (X \cup \{x\}) = (M_1 \cup X) \cup \{x\} \in \mathfrak{M}_0$. Also ist $M_1 \cup (X \cup \{x\})$ Element jedes induktiven Mengensystems, und das ist ja nach (11′) die Behauptung von (iii).

Analog erhält man durch Betrachtung des Systems

$$\mathfrak{M}^* := \{X : \bigwedge_N (N \subseteq X \Rightarrow N \text{ endlich } (R))\}$$

den folgenden Satz (Übungsaufgabe):

(13) M *endlich* $(R) \wedge N \subseteq M \Rightarrow N$ *endlich* (R).

Als nächstes wollen wir zeigen, daß folgendes gilt (vgl. (8)):

(14) M *endlich* $(R) \wedge N \sim M \Rightarrow N$ *endlich* (R).

Aus Gründen der Allgemeinheit wollen wir hierbei zulassen, daß M und N Mengen über unterschiedlichen Grundbereichen E_1 bzw. E_2 sind. Zum Beweis von (14)

betrachten wir das System $\mathfrak{M}^*$ aller derjenigen Mengen X von Elementen aus E_1, für die gilt: Alle zu X gleichmächtigen Mengen Y von Elementen aus E_2 sind im Russellschen Sinne endlich:

$$\mathfrak{M}^* := \{X : X \subseteq E_1 \wedge \bigwedge_Y (Y \subseteq E_2 \wedge Y \sim X \Rightarrow Y \textit{ endlich } (R))\}$$

Wir werden zeigen:

(14′) $\mathfrak{M}^*$ *ist induktiv über* E_1.

Damit ist offenbar (14) bewiesen; ist nämlich M eine beliebige im Russellschen Sinne endliche Menge von Elementen aus E_1, so ist M nach (11′) Element jedes induktiven Mengensystems $\mathfrak{M}$ über E_1, damit wegen (14′) speziell Element von $\mathfrak{M}^*$, und das besagt ja nach Definition von $\mathfrak{M}^*$ gerade, daß jede zu M gleichmächtige Menge Y von Elementen aus E_2, insbesondere also N, im Russellschen Sinne endlich ist. Zum Beweis von (14′) ist nach (9′) folgendes zu zeigen:

(i) $\quad \emptyset \in \mathfrak{M}^*$,

(ii) $\quad X \in \mathfrak{M}^* \wedge x \in E_1 \Rightarrow X \cup \{x\} \in \mathfrak{M}^*$.

Die Behauptung (i) folgt unmittelbar aus der Tatsache, daß jede zur leeren Menge gleichmächtige Menge Y leer ist, und da die leere Menge trivialerweise im Russellschen Sinne endlich ist (sie ist ja nach (9′) Element jedes induktiven Mengensystems!), ist jede zur leeren Menge gleichmächtige Menge im Russellschen Sinne endlich und folglich $\emptyset \in \mathfrak{M}^*$. Zum Beweis von (ii) sei X eine beliebige Menge aus $\mathfrak{M}^*$, d. h., es gelte

(iii) $\quad Y \subseteq E_2 \wedge Y \sim X \Rightarrow Y \textit{ endlich } (R)$.

Es ist zu zeigen, daß dann jede zu $X \cup \{x\}$ gleichmächtige Teilmenge Z von E_2 im Russellschen Sinne endlich ist. Der Fall $x \in X$ ist trivial, weil in diesem Fall $X \cup \{x\} = X$ und mithin unsere Behauptung wegen (iii) richtig ist. Es sei also $x \notin X$, $Z \sim X \cup \{x\}$, f 1-1-Abbildung von Z auf $X \cup \{x\}$ und $\mathfrak{M}_0$ ein beliebiges induktives Mengensystem über E_2. Dann ist $Z \setminus \{f^{-1}(x)\} \sim X$, denn $f \mid (Z \setminus \{f^{-1}(x)\})$ ist 1-1-Abbildung von $Z \setminus \{f^{-1}(x)\}$ auf X, und es gilt nach (iii): $Z \setminus \{f^{-1}(x)\}$ endlich (R). Das heißt, $Z \setminus \{f^{-1}(x)\}$ ist Element jedes induktiven Mengensystems über E_2, speziell also von $\mathfrak{M}_0$. Dann ist nach (9′) auch $Z = (Z \setminus \{f^{-1}(x)\}) \cup \{f^{-1}(x)\}$ Element von $\mathfrak{M}_0$. Da hierbei $\mathfrak{M}_0$ ein ganz beliebiges induktives Mengensystem über E_2 war, ist also Z Element jedes induktiven Mengensystems über E_2 und daher nach (11′) im Russellschen Sinne endlich. Also ist jede zu $X \cup \{x\}$ gleichmächtige Menge Z im Russellschen Sinne endlich und damit nach der Definition von $\mathfrak{M}^*$ in der Tat $X \cup \{x\} \in \mathfrak{M}^*$.

Mit Hilfe von (14) können wir nun leicht den folgenden Satz beweisen:

(15) $\quad M_1 \textit{ endlich } (R) \wedge M_2 \textit{ endlich } (R) \Rightarrow M_1 \times M_2 \textit{ endlich } (R)$.

Ist hierbei M_1 eine Menge über dem Grundbereich E_1, M_2 eine Menge über dem Grundbereich E_2, so wird $M_1 \times M_2$ eine Menge über dem Grundbereich $E_1 \times E_2$. Zum Beweis von (15) betrachtet man bei gegebener im Russellschen Sinne endlichen Menge $M_1 \subseteq E_1$ das folgende Mengensystem $\mathfrak{M}^*$ über E_2:

$$\mathfrak{M}^* := \{X : X \subseteq E_2 \wedge M_1 \times X \textit{ endlich } (R)\}.$$

Wir zeigen, daß folgendes gilt:

(15′) $\mathfrak{M}^*$ *ist induktiv über* E_2.

Man erhält (15) aus (15′) in derselben Weise, wie wir oben (12) aus (12′) gewonnen haben. Zum Beweis von (15′) ist folgendes zu beweisen:

(i) $\emptyset \in \mathfrak{M}^*$;
(ii) $X \in \mathfrak{M}^* \wedge x \in E_2 \Rightarrow X \cup \{x\} \in \mathfrak{M}^*$.

Die Behauptung (i) gilt wegen $M_1 \times \emptyset = \emptyset$ trivial. Mithin bleibt folgendes zu beweisen:

(iii) $M_1 \times X$ *endlich* $(R) \Rightarrow M_1 \times (X \cup \{x\})$ *endlich* (R).

Hierzu beachten wir, daß $M_1 \times (X \cup \{x\}) = (M_1 \times X) \cup (M_1 \times \{x\})$ ist. Also genügt es wegen (12) zu zeigen, daß die Menge $M_1 \times \{x\}$ im Russellschen Sinne endlich ist. Das folgt aber unmittelbar aus (14); denn offenbar ist $\{((y, x), y) : y \in M_1\}$ eine 1-1-Abbildung von $M_1 \times \{x\}$ auf M_1, d. h. $M_1 \times \{x\} \sim M_1$, und M_1 war ja als im Russellschen Sinne endlich vorausgesetzt.

Analog beweist man (Übungsaufgabe), daß folgendes gilt:

(16) M *endlich* $(R) \Rightarrow \mathfrak{P}(M)$ *endlich* (R);

in Worten: *Jede endliche Menge hat nur endlich viele Teilmengen.*

Wir wollen schließlich noch den folgenden merkwürdigen Satz beweisen:

(17) M *endlich* $(R) \Rightarrow \neg \bigvee_N (N \subset M \wedge N \sim M)$;

in Worten: *Keine endliche Menge ist einer ihrer echten Teilmengen gleichmächtig.* Zum Beweis von (17) sei

$$\mathfrak{M}^* := \{X : \neg \bigvee_N (N \subset X \wedge N \sim X)\}.$$

Wir zeigen

(17′) $\mathfrak{M}^*$ *ist induktiv.*

Aus (17′) folgt (17) in analoger Weise wie (12) aus (12′). Zum Beweis von (17′) ist folgendes zu zeigen:

(i) $\emptyset \in \mathfrak{M}^*$;
(ii) $X \in \mathfrak{M}^* \Rightarrow X \cup \{x\} \in \mathfrak{M}^*$.

Die Behauptung (i) gilt trivialerweise, da die leere Menge keine echte Teilmenge besitzt und folglich auch keiner ihrer echten Teilmengen gleichmächtig sein kann. Den Beweis von (ii) führen wir indirekt, setzen also voraus, daß $X \in \mathfrak{M}^*$, und nehmen an, daß $X \cup \{x\} \notin \mathfrak{M}^*$. Dann muß es eine Menge Y mit $Y \subset X \cup \{x\}$ und $Y \sim X \cup \{x\}$ geben. Es sei dann f 1-1-Abbildung von $X \cup \{x\}$ auf Y. Ist hierbei $x \notin Y$, so wird $f \setminus \{(x, f(x))\}$ eine 1-1-Abbildung von X auf $Y \setminus \{f(x)\}$ $(\subset Y \subseteqq X)$, im Widerspruch zu $X \in \mathfrak{M}^*$. Ist dagegen $x \in Y$, so existiert nach (4) eine 1-1-Abbildung g von $X \cup \{x\}$ auf Y mit $g(x) = x$, und dann wird $g \setminus \{(x, x)\}$ eine 1-1-Abbildung von X auf $Y \setminus \{x\}$ $(\subset Y \subseteqq X)$, ebenfalls im Widerspruch zu $X \in \mathfrak{M}^*$. Also ist unsere Annahme falsch, und es gilt (ii).

Aus (17) folgt unmittelbar, daß *jede Menge* M, *die sich eineindeutig auf eine ihrer echten Teilmengen abbilden läßt, unendlich ist* (wobei wir unter einer *unendlichen*

Menge natürlich eine Menge verstehen, die nicht endlich ist). Damit ergibt sich sofort, daß z. B. *die Menge* N *aller natürlichen Zahlen unendlich ist*, denn die Korrespondenz σ, die einer beliebigen Zahl $n \in \mathsf{N}$ die Zahl $n + 1$ zuordnet, ist eine 1-1-Abbildung von N auf $\mathsf{N} \setminus \{0\}$.

Wichtig ist nun, daß auch die Umkehrung der Implikation (17) gilt: *Jede Menge M, die sich nicht eineindeutig auf eine echte Teilmenge von sich selbst abbilden läßt, ist endlich*; oder anders ausgedrückt: *Jede unendliche Menge läßt sich eineindeutig auf eine ihrer echten Teilmengen abbilden.* Der Beweis hierfür ist recht schwierig und benutzt entscheidend das Auswahlaxiom. Wir werden daher diese Umkehrung erst später (vgl. 3.5. (35)) beweisen. Unter Benutzung hiervon wird die folgende Dedekindsche Endlichkeitsdefinition sinnvoll:

$$(18) \qquad M \text{ endlich } (D) :\Leftrightarrow \neg \bigvee_{N} (N \subset M \wedge N \sim M).$$

Auch diese Definition zeichnet sich dadurch aus, daß auf der rechten Seite nur Begriffsbildungen der allgemeinen Mengenlehre auftreten, insbesondere also der Begriff der natürlichen Zahl nicht verwendet wird. Eine Reihe der im vorangehenden bewiesenen Sätze über endliche Mengen (wie z. B. (13) und (14)) lassen sich auch mühelos unter Verwendung der Dedekindschen Definition beweisen, während bei anderen Sätzen ganz erhebliche Schwierigkeiten auftreten. Wir empfehlen dem Leser, diese Dinge selbst zu durchdenken.

3. Das System der natürlichen Zahlen

3.1. Einleitung

Der Zahlbegriff ist das Resultat eines komplizierten und langwierigen historischen Entwicklungsprozesses. Ein genaueres Studium dieses Prozesses zeigt, daß der Zahlbegriff auf sehr frühen Stufen der menschlichen Gesellschaft aus unmittelbaren Bedürfnissen der Praxis entstand und sich bei der weiteren gesellschaftlichen Entwicklung selbst entwickelte und vervollkommnete, daß also der Zahlbegriff keine unveränderliche Kategorie ist, die unserem Verstande a priori eigen, d. h. vor jeder Erfahrung dem Menschen schon bei seiner Geburt gegeben ist. Ebensowenig sind der Zahlbegriff und die arithmetischen Operationen freie Schöpfungen des menschlichen Geistes, sondern ein von einer Reihe spezieller konkreter Merkmale befreiter, abstrakter Ausdruck realer Beziehungen der materiellen Welt. Auf diese außerordentlich wichtigen philosophischen Fragen, die für das richtige Verständnis des Wesens der Mathematik von grundlegender Bedeutung sind, kann hier nicht näher eingegangen werden.

Im vorliegenden Abschnitt werden wir den Begriff der natürlichen Zahl und die wichtigsten elementaren Operationen und Relationen im Bereich der natürlichen Zahlen einer genauen logischen Analyse unterwerfen. Die dabei gewonnenen Resultate werden dem Leser zum größten Teil — allerdings weitgehend empirisch — aus der Schule bekannt sein. Demgegenüber besteht das Hauptanliegen der folgenden Ausführungen darin, dem Leser und zukünftigen Lehrer systematisch die logischen Zusammenhänge aufzudecken und ihn mit grundlegenden Beweisgedanken vertraut zu machen.

Unter natürlichen Zahlen verstehen wir im folgenden die Zahlen $0, 1, 2, \ldots$, sehen also insbesondere die Zahl Null als natürliche Zahl an. Die Frage, ob Null eine natürliche Zahl ist oder nicht, ist ausschließlich eine Frage der Konvention (und nicht etwa der Weltanschauung oder dergleichen). Die der Zahl Null lange Zeit zugeschriebene Sonderrolle besteht eigentlich nur darin, daß

sie bei der historischen Entwicklung des Zahlbegriffs erst sehr spät entstanden ist. Die Menge der natürlichen Zahlen wird nach wie vor mit N bezeichnet.

Unserer Kenntnis der natürlichen Zahlen entnehmen wir, daß die natürlichen Zahlen mindestens zwei wesentliche Aufgaben erfüllen: Einmal benutzt man sie zur Angabe der Anzahl der Elemente von endlichen Mengen, d. h. als Kardinalzahlen, zum anderen (hier allerdings meistens unter Ausschluß der Zahl Null) zum Durchnumerieren der Elemente einer endlichen Menge, d. h. als Ordinalzahlen. Der Leser mache sich sorgfältig klar, daß zwischen Kardinal- und Ordinalzahlen ein grundlegender begrifflicher Unterschied besteht; es ist z. B. etwas anderes, ob ich 50 Seiten oder die 50. Seite eines Buches zu studieren habe.

Zu einer Präzisierung des Kardinalzahlbegriffs gelangt man, wenn man die Kardinalzahlen als Äquivalenzklassen (vgl. 2.5.(13)) der Gleichmächtigkeit (vgl. 2.4.(21)) z. B. im System $\mathfrak{E}$ aller Mengen über einem gegebenen Grundbereich E definiert. Diese Auffassung geht im wesentlichen bereits auf CANTOR und DEDEKIND zurück und spiegelt in abstrakter Form den Inhalt des historisch entstandenen Anzahlbegriffs wider. Im Sinne dieser Definition ist z. B. die Zahl 5 der Inbegriff, das System aller derjenigen Mengen (über E), die sich eineindeutig auf die Finger meiner rechten Hand abbilden lassen. Die Addition von Kardinalzahlen wird durch die Vereinigung disjunkter Mengen (vgl. 2.4.(23) und 3.6.(7)), die Multiplikation durch das kartesische Produkt (vgl. 2.4.(24) und 3.6.(12)) und die $\leqq$-Beziehung durch die Inklusion (vgl. 2.9.(7') und 3.6.(4)) repräsentiert. In ähnlicher Weise kann man den Ordinalzahlbegriff präzisieren, worauf wir hier jedoch nicht eingehen können.

Wir werden uns im folgenden auf keine bestimmte Definition der natürlichen Zahlen stützen. Vielmehr werden wir einige Grundeigenschaften der natürlichen Zahlen als Axiome an die Spitze stellen, aus denen wir durch mathematische Schlüsse alle weiteren uns interessierenden Eigenschaften ableiten werden. Das von uns verwendete Axiomensystem wurde in nur unwesentlich anderer Form im Jahre 1891 von dem italienischen Logiker und Mathematiker GIUSEPPE PEANO (1858—1932) aufgestellt und wird daher heute allgemein *Peanosches Axiomensystem* genannt, obwohl die grundlegenden Ideen bereits von DEDEKIND stammen. Die mengentheoretische Definition der natürlichen Zahlen z. B. als Kardinalzahlen endlicher Mengen hat vor allem den Sinn, sich durch mengentheoretische Konstruktionen mathematische Objekte zu verschaffen, die diesen Axiomen genügen — die ein *Modell* für das Peanosche Axiomensystem bilden – und für die damit auch alle Folgerungen aus dem Axiomensystem gelten. Ihnen kommt insofern grundsätzliche Bedeutung zu, als sie zeigen, daß für eine logische Begründung auch des Zahlbegriffs die

Grundprinzipien der Mengenlehre ausreichen. Sie dürfen allerdings auch nicht (insbesondere philosophisch) überschätzt werden, spiegeln sie doch nur einige Aspekte des historisch gewachsenen Zahlbegriffs wider.

3.2. Das Peanosche Axiomensystem für die natürlichen Zahlen

Als Peanosches Axiomensystem bezeichnet man das folgende System von Aussagen über natürliche Zahlen:

(1) *Die Zahl Null ist eine natürliche Zahl.*
(2) *Jede natürliche Zahl besitzt eine eindeutig bestimmte natürliche Zahl als unmittelbaren Nachfolger.*
(3) *Jede natürliche Zahl ist unmittelbarer Nachfolger höchstens einer natürlichen Zahl.*
(4) *Die Zahl Null ist kein unmittelbarer Nachfolger einer natürlichen Zahl.*
(5) *Die Menge aller natürlichen Zahlen ist die bzgl. Inklusion kleinste Menge, die die Zahl Null und mit einer natürlichen Zahl auch deren unmittelbaren Nachfolger enthält.*

Das Peanosche Axiomensystem charakterisiert also die natürlichen Zahlen als Elemente der Trägermenge einer Struktur $(\mathbf{N}, 0, \sigma)$ (vgl. 2.7.) mit einem ausgezeichneten Element 0 (der Zahl Null) – Axiom (1) – und einer einstelligen Operation σ in $\mathbf{N}$ (die einer beliebigen Zahl $n \in \mathbf{N}$ ihren eindeutig bestimmten *unmittelbaren Nachfolger* – die Zahl $n + 1$ – zuordnet) – Axiom (2) –, in der folgende Eigenschaften erfüllt sind:

(3′) *Die Operation σ ist eindeutig umkehrbar*, d. h.

$$\bigwedge_{m,n \in \mathbf{N}} (\sigma(m) = \sigma(n) \Rightarrow m = n).$$

(4′) $$\neg \bigvee_{n \in \mathbf{N}} \sigma(n) = 0, \quad \text{d. h.} \quad 0 \notin W(\sigma).$$

(5′) $$\bigwedge_{M} (0 \in M \wedge \bigwedge_{n \in \mathbf{N}} (n \in M \Rightarrow \sigma(n) \in M) \Rightarrow \mathbf{N} \subseteq M).$$

Eine derartige Struktur wird heute vielfach *Peano-Struktur* (oder *Peano-Algebra*) genannt. Für die Tatsache, daß in $(\mathbf{N}, 0, \sigma)$ die Axiome (3′) bis (5′) gelten, sagt man auch, daß *diese Struktur ein Modell für diese Axiome ist.*

Bezeichnen wir für eine gegebene natürliche Zahl $n \in \mathbf{N}$ eine Zahl $m \in \mathbf{N}$ mit $\sigma(m) = n$ als *unmittelbaren Vorgänger* von n, so besagt Axiom (3), daß *jede natürliche Zahl höchstens einen unmittelbaren Vorgänger hat*, während in Axiom (4) festgestellt wird, daß *die Zahl Null* (in $\mathbf{N}$) *keinen unmittelbaren*

Vorgänger besitzt. Durch Axiom (5) – das auch Induktionsaxiom genannt wird – wird in präziser Form zum Ausdruck gebracht, daß *man ausgehend von der Zahl Null durch fortgesetzte Nachfolgerbildung schließlich alle natürlichen Zahlen erhält.*

Wir betonen nochmals, daß wir hier nicht die Absicht haben, die Peanoschen Axiome näher zu begründen, d. h. auf irgendwelche anderen Gesetzmäßigkeiten, z. B. auf die in 2.9. genannten Sätze über im Russellschen oder Dedekindschen Sinne endliche Mengen zurückzuführen. Wir sehen sie vielmehr als Grundeigenschaften der natürlichen Zahlen an, aus denen wir alle weiteren uns interessierenden Eigenschaften ableiten. Die im vorliegenden Kapitel bewiesenen Sätze können damit als Sätze aufgefaßt werden, die in jeder Peano-Struktur gelten.

Wir kommen nun zu einigen ersten einfachen Folgerungen aus dem Peanoschen Axiomensystem.

Als erstes wollen wir zeigen, daß *die Zahl Null die einzige natürliche Zahl ist, die keinen unmittelbaren Vorgänger hat,* daß also jede natürliche Zahl $n \neq 0$ wenigstens einen und mithin wegen Axiom (3) genau einen unmittelbaren Vorgänger besitzt:

(6) $\bigwedge_{n \in \mathbb{N}} (n \neq 0 \Rightarrow \bigvee_{m \in \mathbb{N}} \sigma(m) = n).$

Diese Behauptung ist offenbar gleichwertig mit

(6′) $W(\sigma) = \mathbb{N} \setminus \{0\}.$

Zum Beweis dieser Tatsache betrachten wir die Menge M, die die Zahl Null und alle diejenigen natürlichen Zahlen enthält, die einen unmittelbaren Vorgänger besitzen:

$$M := \{0\} \cup \{n : n \in \mathbb{N} \wedge \bigvee_{m \in \mathbb{N}} \sigma(m) = n\}.$$

Offensichtlich ist $0 \in M$. Ferner ist mit einer natürlichen Zahl n stets auch ihr unmittelbarer Nachfolger $\sigma(n)$ in M enthalten; denn $\sigma(n)$ besitzt einen unmittelbaren Vorgänger, nämlich die Zahl n. Folglich enthält nach Axiom (5) M alle natürlichen Zahlen, d. h., jede natürliche Zahl $n \neq 0$ hat wenigstens einen unmittelbaren Vorgänger.

Nur eine andere Formulierung von (3′) ist, daß *verschiedene Zahlen auch verschiedene Nachfolger haben:*

(3″) $m \neq n \Rightarrow \sigma(m) \neq \sigma(n).$

Hiermit erhalten wir leicht, daß *jede natürliche Zahl n von ihrem unmittelbaren Nachfolger $\sigma(n)$ verschieden ist:*

(7) $\bigwedge_{n} (n \in \mathbb{N} \Rightarrow n \neq \sigma(n)).$

Zum Beweis dieser Behauptung betrachten wir die Menge M aller derjenigen natürlichen Zahlen n, die von ihrem unmittelbaren Nachfolger verschieden sind:

$$M := \{n : n \in \mathsf{N} \wedge n \neq \sigma(n)\}.$$

Nach Axiom (4) ist $0 \in M$; denn die Zahl Null ist kein Nachfolger, also sicher von $\sigma(0)$ verschieden. Ist ferner $n \in M$, so ist nach Definition der Menge M dann $n \neq \sigma(n)$, also nach (3'') auch $\sigma(n) \neq \sigma(\sigma(n))$ und mithin $\sigma(n) \in M$. Damit folgt aber nach Axiom (5), daß M alle natürlichen Zahlen enthält, also in der Tat jede natürliche Zahl n von ihrem unmittelbaren Nachfolger $\sigma(n)$ verschieden ist.

Im vorangehenden haben wir zwei Beweise durch „vollständige Induktion" gegeben, die auf einer unmittelbaren Anwendung des Induktionsaxioms (5) beruhten. Meistens pflegt man jedoch Beweise durch vollständige Induktion so zu führen, daß man eine gegebene Aussage über natürliche Zahlen zunächst für die Zahl Null beweist, sodann zeigt, daß aus der Gültigkeit der betreffenden Aussage für eine beliebige natürliche Zahl n ihre Gültigkeit für die Zahl $n + 1$ folgt, und dann behauptet, daß die betrachtete Aussage für alle natürlichen Zahlen richtig ist. Dieser sogenannte *Schluß von n auf n + 1* erhält seine Rechtfertigung durch den folgenden Satz, in dem wir allerdings den unmittelbaren Nachfolger der Zahl n zunächst noch mit $\sigma(n)$ bezeichnen (im folgenden Abschnitt werden wir nach Definition der Addition zeigen können, daß für jede natürliche Zahl n die Beziehung $\sigma(n) = n + 1$ gilt):

Rechtfertigungssatz für Beweise durch vollständige Induktion. *Es sei $H(x)$ eine beliebige Aussage über natürliche Zahlen. Gilt diese Aussage für die Zahl Null und folgt für eine beliebige natürliche Zahl n aus ihrer Gültigkeit für die Zahl n ihre Gültigkeit für die Zahl $\sigma(n)$, so gilt die Aussage $H(x)$ für alle natürlichen Zahlen:*

$$(8) \qquad H(0) \wedge \bigwedge_{n \in \mathsf{N}} (H(n) \Rightarrow H(\sigma(n))) \Rightarrow \bigwedge_{n \in \mathsf{N}} H(n).$$

Den Nachweis der Gültigkeit von $H(0)$ nennt man den *Anfangsschritt* des Induktionsbeweises, den Beweis von $\bigwedge_{n \in \mathsf{N}} (H(n) \Rightarrow H(\sigma(n)))$ den *Induktionsschritt*. Die Voraussetzung $H(n)$ im Beweis des Induktionsschritts nennt man die *Induktionsvoraussetzung*, die daraus zu beweisende Behauptung $H(\sigma(n))$ heißt die *Induktionsbehauptung*. Wir machen darauf aufmerksam, daß man auf Grund des Rechtfertigungssatzes Sätze, die für alle natürlichen Zahlen gelten, durch vollständige Induktion beweisen kann, aber nirgends behauptet ist, daß man sie durch vollständige Induktion beweisen muß.

Zum Beweis von (8) sei $H(x)$ eine Aussage über natürliche Zahlen, für die die Voraussetzungen von (8) erfüllt sind. Wir bezeichnen mit M die Menge aller derjenigen natürlichen Zahlen x, für die $H(x)$ gilt:

$$M := \{x : x \in \mathbf{N} \wedge H(x)\}.$$

Auf Grund der Voraussetzung von (8) ist dann $0 \in M$, und es gilt

$$\bigwedge_{n \in \mathbf{N}} (n \in M \Rightarrow \sigma(n) \in M).$$

Also ist nach (5') $\mathbf{N} \subseteq M$, und das ist ja nach Definition von M gerade die Behauptung von (8).

Wir merken abschließend an, daß man aus der Gültigkeit von (8) (für beliebige Aussagen $H(x)$) auch leicht die Gültigkeit von (5) folgern kann, d. h., (8) und (5) sind logisch äquivalent, so daß man vielfach auch (8) als Induktionsaxiom verwendet. Zum Beweis sei M eine beliebige Menge, die die Voraussetzung von (5) erfüllt, die also die Zahl Null und mit einer natürlichen Zahl n stets auch die Zahl $\sigma(n)$ enthält. Wir betrachten dann die folgende Aussage $H(x)$ über natürliche Zahlen: „$x \in M \cap \mathbf{N}$" („x ist eine natürliche Zahl, die zu M gehört"). Offenbar sind für $H(x)$ die Voraussetzungen von (8) erfüllt, so daß im Fall der Gültigkeit von (8) die Aussage $H(x)$ auf alle natürlichen Zahlen zutrifft, und das besagt ja gerade, daß alle natürlichen Zahlen zu M gehören, was zu zeigen war.

3.3. Die Addition und Multiplikation natürlicher Zahlen

In den Peanoschen Axiomen ist zunächst nur von einer Operation in $\mathbf{N}$, der *Nachfolgeoperation* σ die Rede. Im vorliegenden Abschnitt wollen wir die Addition und Multiplikation von natürlichen Zahlen definieren und die wichtigsten Eigenschaften der natürlichen Zahlen bezüglich dieser Operationen behandeln.

Meistens pflegt man die Addition (und dann analog die Multiplikation) von natürlichen Zahlen induktiv zu definieren, und zwar als zweistellige Operation im Bereich der natürlichen Zahlen, die für beliebiges $m, n \in \mathbf{N}$ den sogenannten *Rekursionsgleichungen*

$$(1) \qquad m + 0 = m, \quad m + \sigma(n) = \sigma(m + n)$$

genügt (wobei wir — wie üblich — das Zeichen $+$ als Operationszeichen für die Addition verwendet haben und im Sinne von 2.6.(3) $m + n$ als Abkürzung für $+(m, n)$ gilt). Diese Definition bedarf jedoch einer grundsätzlichen Rechtfertigung. Es ist nämlich zunächst in keiner Weise gesichert, ob es überhaupt eine Operation gibt, die den Rekursionsgleichungen (1) genügt, und wenn das der Fall ist, ob es nur eine derartige Operation gibt (sonst müßte genauer

gesagt werden, welche der verschiedenen Operationen, die (1) erfüllen, die Addition sein soll).

Die Einzigkeit der durch die angegebenen Rekursionsgleichungen festgelegten Operation läßt sich leicht nachweisen. Dazu nehmen wir an, es seien $+_1$ und $+_2$ zweistellige Operationen in $\mathbf{N}$, welche (1) erfüllen. Wir zeigen, daß dann für beliebige natürliche Zahlen m, n

$$(2) \qquad m +_1 n = m +_2 n$$

gilt, so daß nach 2.4.(7′) in der Tat $+_1 = +_2$ ist (man beachte, daß nach Voraussetzung $+_1$ und $+_2$ Abbildungen von $\mathbf{N} \times \mathbf{N}$ in $\mathbf{N}$ sind). Den Beweis von (2) führen wir bei beliebigem, aber festem $m \in \mathbf{N}$ durch vollständige Induktion über n, wenden also den Rechtfertigungssatz 3.2.(8) auf die Aussage „$m +_1 x = m +_2 x$“ an. Da $+_1$ und $+_2$ beide die erste der Gleichungen (1) erfüllen sollen, ist $m +_1 0 = m$ und $m +_2 0 = m$, also $m +_1 0 = m +_2 0$, d. h., die betrachtete Aussage gilt für die Zahl Null. Wir nehmen nun an, unsere Aussage sei für die Zahl n schon bewiesen (Induktionsvoraussetzung), und zeigen, daß sie dann auch für die Zahl $\sigma(n)$ gilt (Induktionsbehauptung). Nach der zweiten Gleichung aus (1), der ja sowohl $+_1$ als auch $+_2$ genügen sollen, ist $m +_1 \sigma(n) = \sigma(m +_1 n)$ und $m +_2 \sigma(n) = \sigma(m +_2 n)$, wobei nach Induktionsvoraussetzung $m +_1 n = m +_2 n$ gilt, so daß in der Tat

$$m +_1 \sigma(n) = m +_2 \sigma(n)$$

ist, was zu zeigen war.

Wie steht es nun mit der Existenz einer derartigen Operation? Der Beweis hierfür ist etwas schwieriger. Wir benutzen dabei einen Kunstgriff, der auf den ungarischen Mathematiker Laszlo Kalmár zurückgeht. Zunächst zeigen wir nämlich, daß es zu jeder natürlichen Zahl m eine (und übrigens auch nur eine) Abbildung f_m von $\mathbf{N}$ in $\mathbf{N}$ gibt, die die Rekursionsgleichungen

$$(3_m) \qquad f_m(0) = m, \quad f_m(\sigma(n)) = \sigma(f_m(n))$$

erfüllt. Im Fall $m = 0$ leistet offenbar die Funktion

$$f_0 := \{(n, n) : n \in \mathbf{N}\} \; (= e_{\mathbf{N}})$$

das Verlangte; denn es ist $f_0(0) = 0$ und $f_0(\sigma(n)) = \sigma(n) = \sigma(f_0(n))$. Wir nehmen nun an, daß für die Zahl m bereits eine Funktion f_m gefunden ist, für die (3_m) gilt, und konstruieren mit ihrer Hilfe eine Funktion $f_{\sigma(m)}$, die den Gleichungen $(3_{\sigma(m)})$ genügt. Dazu setzen wir

$$f_{\sigma(m)} := \{(n, \sigma(f_m(n))) : n \in \mathbf{N}\}.$$

Dann gilt

$$f_{\sigma(m)}(0) = \sigma(f_m(0)) = \sigma(m)$$

und

$$f_{\sigma(m)}(\sigma(n)) = \sigma(f_m(\sigma(n))) = \sigma(\sigma(f_m(n))) = \sigma(f_{\sigma(m)}(n)),$$

so daß in der Tat $f_{\sigma(m)}$ die Gleichungen $(3_{\sigma(m)})$ erfüllt. (Die vorangehenden Ausführungen sind natürlich eine Kurzfassung eines Beweises durch vollständige Induktion über m für den Satz: *Zu jeder natürlichen Zahl m existiert eine Funktion f_m, die den Rekursionsgleichungen (3_m) genügt*). Wir setzen nun $m + n := f_m(n)$, wobei also $f_m(n)$ der Wert einer bestimmten (der eindeutig bestimmten) Funktion f_m mit (3_m) an der Stelle n ist. Dadurch wird offenbar jedem geordneten Paar (m, n) von natürlichen Zahlen genau eine natürliche Zahl $m + n$ zugeordnet, d. h. eine zweistellige Operation in N definiert. Diese Operation erfüllt die Rekursionsgleichungen (1), denn es gilt

$$m + 0 = f_m(0) = m, \quad m + \sigma(n) = f_m(\sigma(n)) = \sigma(f_m(n)) = \sigma(m + n).$$

Also gibt es genau eine zweistellige Operation $+$ in der Menge N, die den Rekursionsgleichungen (1) genügt, und diese Operation heißt die *Addition von natürlichen Zahlen*. Das Resultat $m + n$ der Anwendung der Additionsoperation auf ein Paar (m, n) von natürlichen Zahlen wird die *Summe* der Zahlen m, n genannt, die dann ihrerseits die *Summanden* heißen.

Wir wollen nun zeigen, daß *die Additionsoperation assoziativ* (vgl. 2.6.(8)) *und kommutativ* (vgl. 2.6.(7)) *ist*, d. h., *für beliebige natürliche Zahlen n_1, n_2, n_3 gilt:*

(4) $\quad n_1 + (n_2 + n_3) = (n_1 + n_2) + n_3,$

(5) $\quad n_1 + n_2 = n_2 + n_1.$

Den Beweis von (4) führen wir (bei festem $n_1, n_2 \in \mathbb{N}$) durch vollständige Induktion über n_3. Auf Grund von (1) gilt

$$n_1 + (n_2 + 0) = n_1 + n_2 = (n_1 + n_2) + 0,$$

d. h., (4) gilt für $n_3 = 0$. Wir nehmen nun an, (4) sei für $n_3 = n$ schon bewiesen, und zeigen, daß (4) dann auch für $n_3 = \sigma(n)$ gilt: Nach (1) ist

$$n_1 + (n_2 + \sigma(n)) = n_1 + \sigma(n_2 + n) = \sigma(n_1 + (n_2 + n)),$$

nach Induktionsvoraussetzung gilt $n_1 + (n_2 + n) = (n_1 + n_2) + n$ und folglich

$$\sigma(n_1 + (n_2 + n)) = \sigma((n_1 + n_2) + n),$$

und wiederum nach (1) ist

$$\sigma((n_1 + n_2) + n) = (n_1 + n_2) + \sigma(n);$$

also gilt in der Tat

$$n_1 + (n_2 + \sigma(n)) = (n_1 + n_2) + \sigma(n),$$

was zu zeigen war.

Zum Beweis von (5) zeigen wir zunächst, daß bei beliebigem $n_1 \in \mathbf{N}$

(6) $\quad 0 + n_1 = n_1$

gilt. Das ergibt sich mühelos durch vollständige Induktion über n_1; denn im Fall $n_1 = 0$ gilt (6) auf Grund von (1) trivial, und wenn (6) für $n_1 = n$ schon bewiesen ist, wird $0 + \sigma(n) = \sigma(0 + n) = \sigma(n)$, d. h., so gilt (6) auch für $n_1 = \sigma(n)$. Ferner zeigen wir durch vollständige Induktion über n_2, daß bei beliebigem $n_1, n_2 \in \mathbf{N}$

(7) $\quad n_1 + \sigma(n_2) = \sigma(n_1) + n_2$

gilt. Im Fall $n_2 = 0$ gilt (7); denn nach (1) ist

$$n_1 + \sigma(0) = \sigma(n_1 + 0) = \sigma(n_1) = \sigma(n_1) + 0.$$

Wir nehmen nun an, (7) sei für $n_2 = n$ schon bewiesen; dann wird

$$n_1 + \sigma(\sigma(n)) = \sigma(n_1 + \sigma(n)) = \sigma(\sigma(n_1) + n) = \sigma(n_1) + \sigma(n),$$

d. h., es gilt (7) auch für $n_2 = \sigma(n)$.

Mittels (6) und (7) kann schließlich (5) durch vollständige Induktion über n_2 bewiesen werden. Nach (1) ist zunächst $n_1 + 0 = n_1$, und nach (6) ist $n_1 = 0 + n_1$, also ist $n_1 + 0 = 0 + n_1$, d. h., (5) gilt für $n_2 = 0$. Wir nehmen nun an, (5) sei für $n_2 = n$ schon bewiesen. Dann wird

$$n_1 + \sigma(n) = \sigma(n_1 + n) = \sigma(n + n_1) = n + \sigma(n_1),$$

wobei nach (7) $n + \sigma(n_1) = \sigma(n) + n_1$ ist, so daß $n_1 + \sigma(n) = \sigma(n) + n_1$ wird, d. h., (5) gilt für $n_2 = \sigma(n)$.

Die erste Gleichung (1) besagt offenbar, daß *die Zahl Null rechtsseitig und* wegen (5) (bzw. (6)) dann natürlich auch *linksseitig neutrales Element für die Addition ist* (vgl. 2.6.(11)), wobei die allgemeinen Betrachtungen aus 2.6. lehren, daß 0 auch die einzige natürliche Zahl x ist, für die $m + x = m$ bei beliebigem $m \in \mathbf{N}$ gilt.

Bezeichnen wir den unmittelbaren Nachfolger der Zahl Null mit 1

(8) $\quad 1 := \sigma(0),$

so wird bei beliebigem $n \in \mathbf{N}$ nach (1) $\sigma(n) = \sigma(n + 0) = n + \sigma(0) = n + 1$, d. h., *für jedes* $n \in \mathbf{N}$ *gilt*

(9) $\quad \sigma(n) = n + 1,$

so daß wir von nun an statt $\sigma(n)$ — wie üblich — auch $n + 1$ schreiben können (was wir systematisch allerdings erst ab 3.5. tun wollen).

Es gelten für die Addition ferner noch die folgenden beiden Sätze, die wir in folgenden Abschnitt benötigen werden:

(10) $m + n = 0 \Leftrightarrow m = 0 \wedge n = 0;$

(11) $m + n = m \Leftrightarrow n = 0.$

Da bei (10) und (11) die Implikationen von rechts nach links trivial gelten, genügt es, die Implikationen von links nach rechts zu beweisen. Zum Beweis von (10) sei also $m + n = 0$ und wir nehmen an, es wäre beispielsweise $n \neq 0$. Dann gäbe es nach 3.2.(6) eine Zahl $k \in \mathsf{N}$, so daß $n = \sigma(k)$, und es wäre

$$m + n = m + \sigma(k) = \sigma(m + k) = 0,$$

im Widerspruch zu Axiom 3.2.(4). Der Beweis von (11) ergibt sich leicht durch vollständige Induktion über m und sei dem Leser als Übungsaufgabe überlassen.

In Analogie zu (1) definieren wir die Multiplikation natürlicher Zahlen induktiv durch die folgenden Rekursionsgleichungen:

(12) $m \cdot 0 = 0, \quad m \cdot \sigma(n) = (m \cdot n) + m.$

In die Rekursionsgleichungen für die Multiplikation geht also bereits die zuvor durch (1) induktiv definierte Addition ein. Entsprechend wie bei der Addition hat man sich natürlich davon zu überzeugen, daß es genau eine binäre Operation $\cdot$ in N gibt, die den Rekursionsgleichungen (12) genügt. Dies sei dem Leser als Übungsaufgabe überlassen. Diese eindeutig bestimmte Operation wird die *Multiplikation von natürlichen Zahlen* genannt. Das Resultat $m \cdot n$ der Anwendung der Multiplikationsoperation auf ein gegebenes Paar (m, n) von natürlichen Zahlen heißt das *Produkt* dieser Zahlen, die dann ihrerseits *Faktoren* bzw. *Multiplikator* und *Multiplikand* genannt werden.

Durch vollständige Induktion über m zeigt man leicht (Übungsaufgabe), daß *für alle* $m \in \mathsf{N}$

(13) $0 \cdot m = 0$

gilt. Ferner ist nach (12) $m \cdot 1 = m \cdot \sigma(0) = m \cdot 0 + m = 0 + m = m$, d. h., *für jedes* $m \in \mathsf{N}$ *ist*

(14) $m \cdot 1 = m,$

die Zahl 1 *ist rechtsseitig neutrales Element für die Multiplikation* (vgl. 2.6.(11_r)). Durch vollständige Induktion über m kann man leicht zeigen, daß *die Zahl* 1 *auch linksseitig neutrales Element für die Multiplikation ist*, d. h., *für jedes* $m \in \mathsf{N}$ *gilt*

(15) $1 \cdot m = m.$

Durch vollständige Induktion beweist man, daß *die Multiplikation assoziativ, kommutativ sowie rechts- und linksseitig distributiv bzgl. der Addition ist*, d. h. *für beliebige natürliche Zahlen* n_1, n_2, n_3 *gilt*

(16) $(n_1 \cdot n_2) \cdot n_3 = n_1 \cdot (n_2 \cdot n_3)$;

(17) $n_1 \cdot n_2 = n_2 \cdot n_1$;

(18_r) $(n_1 + n_2) \cdot n_3 = n_1 \cdot n_3 + n_2 \cdot n_3$;

(18_l) $n_3 \cdot (n_1 + n_2) = n_3 \cdot n_1 + n_3 \cdot n_2$

(wobei wir bei (18) bereits von der üblichen Konvention Gebrauch gemacht haben, daß das Pluszeichen stärker als das Malzeichen trennt, d. h.,

$$n_1 \cdot n_3 + n_2 \cdot n_3$$

ist eine abkürzende Schreibweise für $(n_1 \cdot n_3) + (n_2 \cdot n_3)$). Die durchweg einfachen Beweise seien dem Leser als Übungsaufgaben überlassen, wobei wir nur den Hinweis geben möchten, diese Sätze in folgender Reihenfolge zu beweisen: (18_r), (17), (18_l), (16).

Schließlich *gilt bei beliebigem* $m, n \in \mathbb{N}$

(19) $m \cdot n = 0 \Leftrightarrow m = 0 \vee n = 0$,

(20) $m \cdot n = 1 \Leftrightarrow m = 1 \wedge n = 1$.

Wir können uns offenbar sowohl bei (19) als auch bei (20) auf den Nachweis der Implikation von links nach rechts beschränken. Es sei also $m \cdot n = 0$, und wir nehmen an, es wäre sowohl $m \neq 0$ als auch $n \neq 0$. Dann gäbe es nach 3.2.(6) natürliche Zahlen m_1 und n_1, so daß $m = \sigma(m_1)$, $n = \sigma(n_1)$, und es wäre

$$m \cdot n = \sigma(m_1) \cdot \sigma(n_1) = \sigma(m_1) \cdot n_1 + \sigma(m_1) = \sigma(\sigma(m_1) \cdot n_1 + m_1) = 0,$$

im Widerspruch zu Axiom 3.2.(4). Ist $m \cdot n = 1$, so muß nach (19) sowohl $m \neq 0$ als auch $n \neq 0$ sein. Folglich gibt es natürliche Zahlen m_1, n_1 mit $m = \sigma(m_1)$, $n = \sigma(n_1)$, und dann wird

$$m \cdot n = \sigma(m_1) \cdot \sigma(n_1) = \sigma(m_1) \cdot n_1 + \sigma(m_1) = \sigma(\sigma(m_1) \cdot n_1 + m_1),$$

so daß wegen $m \cdot n = 1$ $(= \sigma(0))$ nach Axiom 3.2.(3)

$$\sigma(m_1) \cdot n_1 + m_1 = 0,$$

also nach (10) $\sigma(m_1) \cdot n_1 = 0$ und $m_1 = 0$ und damit $n_1 = 0$ sein muß, woraus sofort $m = \sigma(m_1) = 1$ und $n = \sigma(n_1) = 1$ folgt.

3.4. Die Ordnung der natürlichen Zahlen

Mit Hilfe der Addition können wir nun die übliche $\leqq$-Relation und die $<$-Relation für natürliche Zahlen definieren:

(1) $m \leqq n :\Leftrightarrow \bigvee_{k \in \mathbb{N}} m + k = n.$

Wir zeigen als erstes, daß *die so definierte* $\leqq$*-Relation eine reflexive totale Ordnung in* $\mathbf{N}$ *ist* (vgl. 2.5.(40)), d. h., *für beliebige natürliche Zahlen* n, n_1, n_2, n_3 *gilt*

(2) $\quad n \leqq n,$

(3) $\quad n_1 \leqq n_2 \wedge n_2 \leqq n_3 \Rightarrow n_1 \leqq n_3,$

(4) $\quad n_1 \leqq n_2 \wedge n_2 \leqq n_1 \Rightarrow n_1 = n_2,$

(5) $\quad n_1 \leqq n_2 \vee n_2 \leqq n_1.$

Zum Beweis von (2) genügt es zu bemerken, daß wegen $n + 0 = n$ eine natürliche Zahl k existiert (nämlich $k = 0$), für die $n + k = n$ ist, und mithin in der Tat $n \leqq n$ gilt. Zum Beweis von (3) sei $n_1 \leqq n_2$ und $n_2 \leqq n_3$. Dann existieren wegen (1) natürliche Zahlen k_1, k_2 mit $n_1 + k_1 = n_2$ und $n_2 + k_2 = n_3$. Dann wird aber nach 3.3.(4) $n_1 + (k_1 + k_2) = (n_1 + k_1) + k_2 = n_2 + k_2 = n_3$, d. h., für die natürliche Zahl $k = k_1 + k_2$ gilt $n_1 + k = n_3$, und es ist in der Tat $n_1 \leqq n_3$. Zum Beweis von (4) sei $n_1 \leqq n_2$ und $n_2 \leqq n_1$, und zwar sei $n_1 + k_1 = n_2$ und $n_2 + k_2 = n_1$. Dann wird $n_1 + (k_1 + k_2) = (n_1 + k_1) + k_2 = n_2 + k_2 = n_1$, und auf Grund von 3.3.(11) muß daher $k_1 + k_2 = 0$ sein, woraus mittels 3.3.(10) $k_1 = k_2 = 0$ folgt. Dann ist aber wegen $n_1 + k_1 = n_2$ in der Tat $n_1 = n_2$, was zu zeigen war.

Dem Beweis von (5), d. h. der Linearität der $\leqq$-Relation, schicken wir zunächst zwei Bemerkungen voraus: Wegen $0 + n = n$ gilt für jedes $n \in \mathbf{N}$ die Beziehung

(6) $\quad 0 \leqq n,$

die Zahl 0 *ist* also *bzgl. der* $\leqq$*-Relation kleinste Zahl in* $\mathbf{N}$. Entsprechend folgt aus $n + 1 = \sigma(n)$ (vgl. 3.3.(9)), daß für jedes $n \in \mathbf{N}$

(7) $\quad n \leqq \sigma(n)$

gilt. Damit können wir nun (5) leicht durch vollständige Induktion z. B. über n_1 beweisen. Wegen $0 \leqq n_2$ gilt erst recht $0 \leqq n_2 \vee n_2 \leqq 0$, d. h., (5) ist für $n_1 = 0$ richtig. Wir nehmen nun an, (5) sei für $n_1 = n$ schon bewiesen, es gelte also $n \leqq n_2 \vee n_2 \leqq n$, und zeigen, daß dann (5) auch für $n_1 = \sigma(n)$ gilt. Ist $n_2 \leqq n$, so ist wegen $n \leqq \sigma(n)$ auf Grund von (3) $n_2 \leqq \sigma(n)$, und es gilt in der Tat $\sigma(n) \leqq n_2 \vee n_2 \leqq \sigma(n)$. Ist dagegen $n \leqq n_2$, so existiert eine Zahl $k \in \mathbf{N}$ mit $n + k = n_2$; ist hierbei $k = 0$, so ist $n = n_2$ und folglich wegen (7) $n_2 \leqq \sigma(n)$, also gilt erst recht $\sigma(n) \leqq n_2 \vee n_2 \leqq \sigma(n)$; ist dagegen $k \neq 0$, so existiert nach 3.2.(6) eine natürliche Zahl k_1 mit $k = \sigma(k_1)$, und es wird $n_2 = n + k = n + \sigma(k_1) = \sigma(n) + k_1$ (vgl. 3.2.(7)), also $\sigma(n) \leqq n_2$, und damit

gilt ebenfalls erst recht $\sigma(n) \leqq n_2 \vee n_2 \leqq \sigma(n)$. Also ist (5), wenn es für $n_1 = n$ gilt, auch für $n_1 = \sigma(n)$ richtig, was noch zu zeigen war.

Mittels 2.5.(26) und 2.5.(42) erhalten wir sofort, daß *durch*

(8) $\quad m < n : \Leftrightarrow m \leqq n \wedge m \neq n$

eine irreflexive totale Ordnung in N *definiert wird*, d. h., für die durch (8) definierte $<$-Relation in N gelten die folgenden Sätze:

(9) $\quad \neg \bigvee_{n \in \mathbf{N}} n < n,$

(10) $\quad \bigwedge_{n_1, n_2, n_3 \in \mathbf{N}} (n_1 < n_2 \wedge n_2 < n_3 \Rightarrow n_1 < n_3),$

(11) $\quad \bigwedge_{n_1, n_2 \in \mathbf{N}} (n_1 < n_2 \vee n_1 = n_2 \vee n_2 < n_1).$

Aus 2.5.(11) folgt, daß *die $<$-Relation auch asymmetrisch ist:*

(12) $\quad \bigwedge_{n_1, n_2 \in \mathbf{N}} (n_1 < n_2 \Rightarrow \neg\, n_2 < n_1),$

und Satz 2.5.(45) lehrt, daß *die $<$-Relation trichotom ist:*

(13) $\quad \bigwedge_{n_1, n_2 \in \mathbf{N}} ((n_1 < n_2 \vee n_1 = n_2 \vee n_2 < n_1)$
$\wedge \neg (n_1 < n_2 \wedge n_1 = n_2) \wedge \neg (n_1 < n_2 \wedge n_2 < n_1)$
$\wedge \neg (n_1 = n_2 \wedge n_2 < n_1)),$

d. h., von den drei Fällen $n_1 < n_2$, $n_1 = n_2$, $n_2 < n_1$ tritt bei beliebigem $n_1, n_2 \in \mathbf{N}$ stets genau einer ein. Schließlich folgt aus 2.5.(28) und 2.5.(33), daß *bei beliebigem* $m, n \in \mathbf{N}$

(14) $\quad m \leqq n \Leftrightarrow m < n \vee m = n$

gilt, wodurch insbesondere die Lesart „m ist kleiner oder gleich n" für $m \leqq n$ gerechtfertigt wird (da man $m < n$ als „m ist kleiner als n" liest).

Die $<$-Relation läßt sich auch leicht direkt mit Hilfe der Addition ausdrücken: *Bei beliebigem* $m, n \in \mathbf{N}$ *gilt*

(15) $\quad m < n \Leftrightarrow \bigvee_{k \in \mathbf{N}} (k \neq 0 \wedge m + k = n).$

Zum Beweis nehmen wir zunächst an, es sei $m < n$. Dann ist nach (8) $m \leqq n$ und $m \neq n$. Wegen (1) gibt es folglich eine natürliche Zahl k mit $m + k = n$, und hierbei muß $k \neq 0$ sein, da andernfalls $m = n$ wäre. Gibt es umgekehrt eine Zahl $k \neq 0$ mit $m + k = n$, so ist wegen (1) $m \leqq n$; es muß aber auch $m \neq n$ sein, da andernfalls $m + k = m$ wäre und dann nach 3.3.(11) $k = 0$ sein müßte.

Als nächstes zeigen wir, daß *die Addition bezüglich der $\leqq$-Relation monoton ist* (vgl. 2.6.(12)), wobei wir uns wegen der Kommutativität der Addition

natürlich auf den Nachweis z. B. der rechtsseitigen Monotonie beschränken können:

$$(16) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (n_1 \leqq n_2 \Rightarrow m + n_1 \leqq m + n_2).$$

Es sei also $n_1 \leqq n_2$. Dann exisitiert eine natürliche Zahl k mit $n_1 + k = n_2$. Folglich wird $(m + n_1) + k = m + (n_1 + k) = m + n_2$, und mithin ist $m + n_1 \leqq m + n_2$. Entsprechend zeigt man unter Benutzung von 3.3.(18_l), daß auch *die Multiplikation* (rechtsseitig) *monoton bzgl. der* $\leqq$*-Relation ist:*

$$(17) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (n_1 \leqq n_2 \Rightarrow m \cdot n_1 \leqq m \cdot n_2).$$

Analog beweist man mittels 3.3.(18), daß *die Addition und die Multiplikation auch monoton bzgl. der* $<$*-Relation sind, wobei allerdings bei der Multiplikation zusätzlich* $m \neq 0$ *vorausgesetzt werden muß:*

$$(18) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (n_1 < n_2 \Rightarrow m + n_1 < m + n_2),$$

$$(19) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (n_1 < n_2 \wedge m \neq 0 \Rightarrow m \cdot n_1 < m \cdot n_2).$$

Beim Beweis von (19) wird wesentlich von 3.3.(19) Gebrauch gemacht $(m \neq 0 \wedge k \neq 0 \Rightarrow m \cdot k \neq 0)$.

Aus (18) bzw. (19) folgt mittels 2.6.(13), daß auch *die Umkehrungen von* (18) *bzw.* (19) *gelten und die Addition und Multiplikation kürzbare Operationen* (vgl. 2.6.(15)) *sind:*

$$(20) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (m + n_1 < m + n_2 \Rightarrow n_1 < n_2),$$

$$(21) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (m \cdot n_1 < m \cdot n_2 \Rightarrow n_1 < n_2),$$

$$(22) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (m + n_1 = m + n_2 \Rightarrow n_1 = n_2),$$

$$(23) \qquad \bigwedge_{m,n_1,n_2 \in \mathbb{N}} (m \cdot n_1 = m \cdot n_2 \wedge m \neq 0 \Rightarrow n_1 = n_2).$$

Aus (22) bzw. (23) folgt mittels 2.6.(17), daß *die Addition und die Multiplikation beschränkt ausführbare Umkehroperationen besitzen*, die wir *Subtraktion* bzw. *Division* nennen und für die wir die Operationszeichen — bzw. : verwenden. Sie sind definiert durch

$$(24) \qquad n - m = k :\Leftrightarrow m + k = n,$$

bzw.

$$(25) \qquad n : m = k :\Leftrightarrow m \cdot k = n,$$

wobei wir statt $n : m$, wie üblich auch $\frac{n}{m}$ schreiben. Der Definitionsbereich für die Subtraktion bzw. Division wird dabei gegeben durch (vgl. 2.6.(16))

$$(26) \qquad D(-) = \{(n, m) : \bigvee_{k \in \mathbf{N}} m + k = n\} = \{(n, m) : m \leqq n\},$$

$$(27) \qquad D(:) = \{(n, m) : m \neq 0 \wedge \bigvee_{k \in \mathbf{N}} m \cdot k = n\} = \{(n, m) : m \neq 0 \wedge m \mid n\},$$

wobei $\mid$ die Teilbarkeitsrelation bezeichnet, die wir systematisch in 3.7. behandeln werden.

Wir merken an, daß wegen 3.2.(7) die Behauptung (7) unmittelbar zu

$$(28) \qquad \bigwedge_{n \in \mathbf{N}} n < \sigma(n)$$

verschärft werden kann. Ferner gilt bei beliebigem $n \in \mathbf{N}$

$$(29) \qquad \neg \bigvee_{m \in \mathbf{N}} (n < m \wedge m < \sigma(n)).$$

Offenbar ist (29) äquivalent mit

$$\bigwedge_{m \in \mathbf{N}} (n < m \Rightarrow \neg\, m < \sigma(n)),$$

wobei $\neg\, m < \sigma(n)$ wegen (11) und (14) seinerseits mit $\sigma(n) \leqq m$ äquivalent ist. Also sind

$$(29') \qquad \bigwedge_{m \in \mathbf{N}} (n < m \Rightarrow \sigma(n) \leqq m)$$

und analog auch (Beweis!)

$$(29'') \qquad \bigwedge_{m \in \mathbf{N}} (m < \sigma(n) \Rightarrow m \leqq n)$$

äquivalente Formulierungen für (29). Durch (28) und (29) wird übrigens nachträglich die für $\sigma(n)$ benutzte Bezeichnungsweise *unmittelbarer Nachfolger* von n (nämlich in der $<$-Relation) gerechtfertigt.

Als Hilfssatz zum Beweis von (29') zeigen wir zunächst:

$$(30) \qquad \bigwedge_{k \in \mathbf{N}} (k \neq 0 \Rightarrow k \geqq 1)$$

(wobei $k \geqq l$ allgemein nur eine andere Schreibweise für $l \leqq k$ sein soll). Ist nämlich $k \in \mathbf{N}$ und $k \neq 0$, so existiert nach 3.2.(6) eine natürliche Zahl k_1 mit $k = \sigma(k_1) = k_1 + 1$. Hierbei ist nach (6) $k_1 \geqq 0$, und Anwendung von (16) (genauer der linksseitigen Monotonie der Addition bezüglich der $\leqq$-Relation) liefert: $k = k_1 + 1 \geqq 0 + 1 = 1$, was zu beweisen war.

Die Behauptung (29') kann nun folgendermaßen bewiesen werden: Es sei $n < m$. Nach (15) gibt es dann eine natürliche Zahl $k \neq 0$ mit $n + k = m$.

Wegen (30) ist hierbei $k \geqq 1$, und (16) ergibt $\sigma(n) = n + 1 \leqq n + k = m$, wie in (29′) behauptet wurde.

Es sei darauf hingewiesen, daß der unmittelbare Nachfolger $\sigma(n)$ durch die Bedingungen (28) und (29) eindeutig charakterisiert ist, d. h., *erfüllt eine natürliche Zahl n' die Bedingungen*

(28*) $\quad n < n'$,

(29*) $\quad \neg \bigvee_{m \in \mathbb{N}} (n < m \wedge m < n')$,

so ist $n' = \sigma(n)$. Denn aus (28*) und (29′) folgt $\sigma(n) \leqq n'$, und aus (28) und der zu (29*) äquivalenten Bedingung

(29*′) $\quad \bigwedge_{m \in \mathbb{N}} (n < m \Rightarrow n' \leqq m)$

folgt $n' \leqq \sigma(n)$, so daß wegen (4) in der Tat $n' = \sigma(n)$ ist.

Sehr häufig verwendet man in der Mathematik das folgende

Prinzip der kleinsten Zahl. *In jeder nichtleeren Menge von natürlichen Zahlen gibt es eine (eindeutig bestimmte) kleinste Zahl*, d. h. eine Zahl, die kleiner als alle anderen Zahlen der betrachteten Menge ist:

(31) $\quad M \subseteqq \mathbb{N} \wedge M \neq \emptyset \Rightarrow \bigvee_{m_0} (m_0 \in M \wedge \bigwedge_{m} (m \in M \Rightarrow m_0 \leqq m))$.

Die kleinste Zahl in einer gegebenen nichtleeren Menge M von natürlichen Zahlen wird üblicherweise mit $\min M$ (gelesen: Minimum von M) bezeichnet. Insbesondere ist also $\min\{n_1, \ldots, n_k\}$ die kleinste der Zahlen $n_1, \ldots, n_k$. Zum Beweis von (31) betrachten wir die Menge N aller natürlichen Zahlen n, die kleiner oder gleich allen Zahlen m der Menge M sind:

$$N := \{n : n \in \mathbb{N} \wedge \bigwedge_{m} (m \in M \Rightarrow n \leqq m)\}.$$

Wegen (6) ist offenbar $0 \in N$. Andererseits enthält die Menge N sicher nicht alle natürlichen Zahlen; denn ist $m \in M$, so ist wegen $m < \sigma(m)$ die Zahl $\sigma(m)$ nicht Element von N. Daher gibt es eine Zahl m_0, so daß $m_0 \in N$ und $\sigma(m_0) \notin N$; denn andernfalls enthielte N nach dem Induktionsaxiom 3.2.(5) sämtliche natürlichen Zahlen. Wegen $m_0 \in N$ gilt nach Definition von N:

$$\bigwedge_{m} (m \in M \Rightarrow m_0 \leqq m).$$

Außerdem muß m_0 Element von M sein; denn andernfalls wäre $m_0 < m$ für alle $m \in M$, und nach (29′) wäre $\sigma(m_0) \leqq m$ für alle $m \in M$, also $\sigma(m_0) \in N$ – was ja nicht der Fall sein sollte. Also enthält in der Tat jede nichtleere Menge von natürlichen Zahlen eine kleinste Zahl. Daß diese kleinste Zahl eindeutig bestimmt ist, ist auch leicht einzusehen: Sind nämlich m_0 und m_{00} kleinste

Zahlen in M, gilt also

(i) $m_0 \in M$, (ii) $m \in M \Rightarrow m_0 \leqq m$,
(i') $m_{00} \in M$, (ii') $m \in M \Rightarrow m_{00} \leqq m$,

so ist wegen (i) und (ii') $m_{00} \leqq m_0$ und wegen (i') und (ii) $m_0 \leqq m_{00}$, also nach (4) in der Tat $m_0 = m_{00}$.

Mit Hilfe des Prinzips der kleinsten Zahl ergibt sich leicht der folgende

Rechtfertigungssatz für Beweise durch ordnungstheoretische Induktion. *Es sei $H(x)$ eine beliebige Aussage über natürliche Zahlen. Gilt diese Aussage für die Zahl Null und folgt bei beliebigem $n \in \mathsf{N}$ aus der Gültigkeit von $H(x)$ für alle Zahlen $m < n$ die Gültigkeit von $H(x)$ für die Zahl n, so gilt die Aussage $H(x)$ für alle natürlichen Zahlen:*

$$(32) \qquad H(0) \wedge \bigwedge_{n \in \mathsf{N}} \Big(\bigwedge_{m \in \mathsf{N}} (m < n \Rightarrow H(m)) \Rightarrow H(n) \Big) \Rightarrow \bigwedge_{n \in \mathsf{N}} H(n).$$

Im Gegensatz zur vollständigen Induktion 3.2.(8) ist also bei der ordnungstheoretischen Induktion im Induktionsschritt aus der *Induktionsvoraussetzung*, daß die Aussage $H(x)$ für alle Zahlen $m < n$ gilt, die *Induktionsbehauptung*, daß $H(x)$ dann auch für die Zahl n gilt, zu erschließen. Bei der gegebenen Formulierung ist übrigens rein formal der *Anfangsschritt* $H(0)$ im Induktionsschritt enthalten, brauchte also nicht gesondert gefordert zu werden. Bei praktischen Anwendungen ist allerdings der Fall $n = 0$ meistens gesondert zu behandeln, so daß wir ihn extra aufgeführt haben.

Zum Beweis von (32) sei $H(x)$ eine Aussage über natürliche Zahlen, für die die Voraussetzungen von (32) erfüllt sind. Wir nehmen an, die Behauptung von (32) wäre falsch, d. h., es gäbe eine natürliche Zahl n_0, für die die Aussage $H(x)$ nicht gilt. Dann wäre die Menge M aller derjenigen natürlichen Zahlen n, für die $H(x)$ falsch ist, nicht leer. Also enthielte die Menge M eine kleinste Zahl m_0. Nach Definition der Menge M müßte dann (wegen $m_0 \in M$) die Aussage $H(x)$ für die Zahl m_0 falsch sein, während sie für alle Zahlen $m < m_0$ gültig ist. Das ist aber ein Widerspruch dazu, daß bei beliebigem n (speziell also für $n = m_0$) aus der Gültigkeit von $H(x)$ für alle Zahlen $m < n$ die Gültigkeit von $H(x)$ für die Zahl n folgen sollte. Also ist unsere Annahme falsch und $H(x)$ gilt für alle natürlichen Zahlen.

Für den Nachweis, daß eine Aussage $H(x)$ über natürliche Zahlen für alle Zahlen $n \geqq n_0$ gilt, wobei n_0 eine bestimmte natürliche Zahl bedeutet, kann man die folgende modifizierte Form der Beweise durch vollständige

Induktion benutzen:

$$(33)\qquad H(n_0) \wedge \bigwedge_{n \in \mathbf{N}} (n \geqq n_0 \wedge H(n) \Rightarrow H(\sigma(n))) \Rightarrow \bigwedge_{n \in \mathbf{N}} (n \geqq n_0 \Rightarrow H(n)).$$

Zum Beweis von (33) sei $H(x)$ eine Aussage über natürliche Zahlen, für die die Voraussetzungen von (33) erfüllt sind. Es sei

$$M := \{x : H(x)\} \cup \{x : x \in \mathbf{N} \wedge x < n_0\}.$$

Man zeigt leicht, daß die Menge M die Voraussetzungen des Induktionsaxioms 3.2.(5) erfüllt, also nach diesem alle natürlichen Zahlen enthält. Dann müssen aber (da die Menge $\{x : x \in \mathbf{N} \wedge x < n_0\}$ nur die Zahlen enthält, die kleiner als n_0 sind) alle Zahlen $n \geqq n_0$ in der Menge $\{x : H(x)\}$ enthalten sein, und das ist ja gerade die Behauptung von (33).

Man kann ohne Schwierigkeit eine entsprechende Modifikation von (32) formulieren und beweisen, was wir dem Leser als Übungsaufgabe überlassen wollen.

Zum Abschluß sei noch das folgende Prinzip der größten Zahl bewiesen: *In jeder nichtleeren nach oben beschränkten Menge M von natürlichen Zahlen gibt es eine (eindeutig bestimmte) größte Zahl;* dabei heißt eine Menge M von natürlichen Zahlen *nach oben beschränkt*, wenn es eine natürliche Zahl n_0 gibt, so daß alle Zahlen m aus M kleiner oder gleich n_0 sind, d. h., wenn keine Zahl der Menge M größer als n_0 ist (eine solche Zahl n_0 wird dann auch eine *obere Schranke* für die Menge M genannt):

$$(34)\qquad M \subseteq \mathbf{N} \wedge M \neq \emptyset \wedge \bigvee_{n_0 \in \mathbf{N}} \bigwedge_{m} (m \in M \Rightarrow m \leqq n_0)$$
$$\Rightarrow \bigvee_{m_0} (m_0 \in M \wedge \bigwedge_{m} (m \in M \Rightarrow m \leqq m_0)).$$

Die größte Zahl in einer nach oben beschränkten nichtleeren Menge M von natürlichen Zahlen wird üblicherweise mit $\max M$ (gelesen: Maximum von M) bezeichnet. Insbesondere ist also $\max\{n_1, \ldots, n_k\}$ die größte der Zahlen $n_1, \ldots, n_k$ (wobei zu beachten ist, daß *jede nichtleere endliche Menge von natürlichen Zahlen nach oben beschränkt ist* – Beweis!). Es sei M eine beliebige Menge, die die Voraussetzungen des Satzes (34) erfüllt. Mit N bezeichnen wir die Menge aller derjenigen natürlichen Zahlen n, die größer oder gleich allen Zahlen m aus M sind (die also obere Schranken für M sind):

$$N := \{n : n \in \mathbf{N} \wedge \bigwedge_{m} (m \in M \Rightarrow m \leqq n)\}.$$

Diese Menge N ist offenbar nicht leer (da M nach oben beschränkt ist). Folglich enthält N nach (31) eine kleinste Zahl m_0. Wir behaupten, daß m_0 größte Zahl in M ist. Wegen $m_0 \in N$ ist sicher $m \leqq m_0$ für alle $m \in M$. Es bleibt also zu zeigen, daß m_0 zu M gehört. Dazu nehmen wir an, das wäre nicht der Fall;

dann gälte $m < m_0$ für alle $m \in M$. Da M nicht leer ist (hier wird diese Voraussetzung von (34) entscheidend benutzt), wäre $m_0 \neq 0$. Es gäbe also ein $m_1 \in \mathbf{N}$ mit $m_0 = \sigma(m_1)$, und dann gälte nach (29'') $m \leqq m_1$ für alle $m \in M$. Mithin wäre m_1 Element von N, im Widerspruch dazu, daß m_0 kleinste Zahl aus N sein sollte, aber nach (28) $m_1 < m_0$ ist.

3.5. Induktive Definitionen

Bei der Definition der Addition und der Multiplikation von natürlichen Zahlen in 3.3. haben wir bereits zwei Beispiele für sogenannte **induktive Definitionen** kennengelernt. In beiden Fällen war zunächst zu beweisen, daß es jeweils **genau eine** zweistellige Operation in N gibt, die den formulierten Rekursionsgleichungen 3.3.(1) bzw. 3.3.(12) genügt. Nachdem das geschehen war, konnten wir die entsprechenden induktiven Definitionen durch einwandfreie **explizite Definitionen** ergänzen, indem wir nämlich z. B. die Addition von natürlichen Zahlen als die eindeutig bestimmte Operation in N definierten, die die Rekursionsgleichungen 3.3.(1) erfüllt. In diesem Sinne spricht man von einer **Rechtfertigung** der induktiven Definition 3.3.(1). Gemeint ist damit also, daß die Rekursionsgleichungen 3.3.(1) eine eindeutig bestimmte Operation charakterisieren und in dieser Hinsicht als Definition dieser Operation angesehen werden können.

Es zeigt sich nun, daß man in analoger Weise sehr allgemeine Arten von *induktiven* oder – wie man auch sagt – *rekursiven Definitionen* rechtfertigen kann. Wir wollen hier nur ein besonders wichtiges Beispiel für einen solchen allgemeinen **Rechtfertigungssatz** etwas eingehender diskutieren. Dazu sei M eine beliebige Menge von irgendwelchen mathematischen Objekten, x_0 ein fest vorgegebenes Element aus M und F eine bestimmte Abbildung von $\mathbf{N} \times M$ in M. Wir betrachten dann die folgenden *Rekursionsgleichungen:*

$$(1) \qquad f(0) = x_0, \quad f(n+1) = F(n, f(n))$$

(wobei wir von nun an statt $\sigma(n)$ wie üblich $n+1$ schreiben). Die Frage ist, ob es unter den angegebenen Voraussetzungen stets genau eine Abbildung f von N in M gibt, die bei beliebigem $n \in \mathbf{N}$ den Gleichungen (1) genügt. Diese Frage wird nahegelegt durch die Tatsache, daß ja durch die Gleichungen (1) die Funktionswerte $f(0)$, $f(1)$, $f(2)$, ... der Reihe nach eindeutig festgelegt sind, wobei auf Grund des Induktionsaxioms 3.2.(5) der zu einem beliebigen Argumentwert $n \in \mathbf{N}$ gehörige Funktionswert $f(n)$ auch erreicht wird. Diese

Bemerkung wird auch heute noch vielfach als ausreichende Begründung für die Rechtmäßigkeit induktiver Definitionen angesehen. Sie ist indes keineswegs ein strenger Beweis des Satzes, daß *es genau eine Abbildung f von* $\mathbf{N}$ *in* M *gibt, die den Rekursionsgleichungen* (1) *genügt*, als der sie eigentlich wohl gedacht ist. Hierauf hat bereits im Jahre 1888 RICHARD DEDEKIND mit allem Nachdruck hingewiesen, von dem auch der erste strenge Beweis des genannten Rechtfertigungssatzes stammt. Der nachfolgende Beweis des Rechtfertigungssatzes, der vorwiegend technischer Natur ist, kann bei einem ersten Studium übergangen werden.

Wir zeigen als erstes, daß es höchstens eine Abbildung f von $\mathbf{N}$ in M gibt, die den Rekursionsgleichungen (1) genügt. Sind nämlich f_1, f_2 Abbildungen von $\mathbf{N}$ in M, die beide die Rekursionsgleichungen (1) erfüllen, so gilt zunächst $f_1(0) = x_0$, $f_2(0) = x_0$, d. h. $f_1(0) = f_2(0)$. Wir nehmen nun an, daß für eine gewisse Zahl n die Gleichung $f_1(n) = f_2(n)$ bereits gilt. Dann gilt aber auch

$$f_1(n+1) = F(n, f_1(n)) = F(n, f_2(n)) = f_2(n+1).$$

Daraus folgt, daß die Gleichung $f_1(n) = f_2(n)$ für alle natürlichen Zahlen n gilt, und das besagt ja gerade, daß die Funktionen f_1 und f_2 übereinstimmen.

Für den Existenzbeweis betrachten wir das System $\mathfrak{M}$ aller derjenigen Mengen $N \subseteq \mathbf{N} \times M$, die folgende Bedingungen erfüllen:

(i) $(0, x_0) \in N$,

(ii) $(n, x) \in N \Rightarrow (n+1, F(n, x)) \in N$,

und bezeichnen mit N^* den Durchschnitt dieses Mengensystems (wir bemerken, daß z. B. die Menge $\mathbf{N} \times M$ zu $\mathfrak{M}$ gehört). Man zeigt leicht, daß auch die Menge N^* die Bedingungen (i) und (ii) erfüllt, d. h. $\mathbf{N}^* \in \mathfrak{M}$; denn $(0, x_0)$ gehört zu allen Mengen des Systems $\mathfrak{M}$ und damit zu dessen Durchschnitt N^*, und liegt das Paar (n, x) in N^*, so gehört (n, x) und mithin nach (ii) auch $(n+1, F(n, x))$ zu jeder Menge des Systems $\mathfrak{M}$, und folglich liegt $(n+1, F(n, x))$ in N^*. Wir werden nun zeigen, daß folgendes gilt:

(iii) *Zu jeder natürlichen Zahl n gibt es genau ein Element x aus M, so daß* $(n, x) \in N^*$.

Offenbar besagt (iii) gerade, daß N^* eine (eindeutige) Abbildung von $\mathbf{N}$ in M ist, und die Bedingungen (i) bzw. (ii) für N^* besagen gerade, daß folgendes gilt:

(i*) $N^*(0) = x_0$,

(ii*) $N^*(n) = x \Rightarrow N^*(n+1) = F(n, x)$,

wobei wir statt (ii*) offenbar auch kurz $N^*(n+1) = F(n, N^*(n))$ schreiben können. Das heißt aber, daß die Abbildung N^* die Rekursionsgleichungen (1) erfüllt, womit der Existenzbeweis erbracht ist. Es bleibt also (iii) zu beweisen. Das tun wir durch vollständige Induktion über n.

Anfangsschritt: Da $(0, x_0) \in N^*$ ist, genügt es zu zeigen, daß kein Paar $(0, x)$ mit $x \neq x_0$ zu N^* gehört. Angenommen, das wäre doch der Fall. Dann würden wir die Menge $N_1 = N^* \setminus \{(0, x)\}$ betrachten. Man erkennt leicht, daß sie ebenfalls die Bedingungen (i) und (ii) erfüllen würde; denn das in N^* enthaltene Paar $(0, x_0)$

ist (wegen $x \neq x_0$) in N_1 verblieben, und mit (n, x) gehört stets $(n+1, F(n, x))$ zu N_1, da wir ja kein Paar entfernt haben, dessen erste Komponente gleich $n+1$ ist. Die Beziehung $N_1 \in \mathfrak{M}$ hätte aber (vgl. 1.6.(9)) $\bigcap \mathfrak{M} = N^* \subseteq N_1$ zur Folge, was im Widerspruch zu $N_1 \subset N^*$ steht.

Induktionsschritt: Wir nehmen an, daß bereits gezeigt ist, daß es genau ein $x \in M$ mit $(n, x) \in N^*$ gibt, und zeigen, daß es dann auch genau ein $y \in M$ mit $(n+1, y) \in N^*$ gibt. Die Existenz eines solchen y ist klar, denn wegen Eigenschaft (ii) von N^* leistet $F(n, x)$ das Verlangte. Es bleibt also zu zeigen, daß es daneben kein weiteres derartiges y geben kann. Der Beweis hierfür verläuft analog wie im Anfangsschritt, indem man zeigt, daß andernfalls die Menge

$$N^* \setminus \{(n+1, y)\} \subset N^*$$

die Bedingungen (i) und (ii) erfüllen würde, was wie dort zum Widerspruch führt (beim Beweis von (ii) ist dabei lediglich zu beachten, daß das herausgenommene Paar $(n+1, y)$ im Fall $y \neq F(n, x)$ nicht von der Form $(n+1, F(n, z))$ mit $(n, z) \in N^*$ sein kann).

In vielen Anwendungsfällen, so z. B. bei der Addition (vgl. 3.3.(3)), der Multiplikation und der anschließend zu behandelnden Potenzierung hat man es faktisch mit dem Spezialfall zu tun, daß die Werte der Funktion F nicht vom ersten Argument abhängen, d. h. eine Abbildung G von M in M existiert, so daß bei beliebigem $(n, x) \in \mathbb{N} \times M$

$$F(n, x) = G(x)$$

ist. In diesem Fall reduzieren sich die Rekursionsgleichungen (1) auf

$$(1') \qquad f(0) = x_0, \quad f(n+1) = G(f(n)),$$

wobei also x_0 ein festes Element aus M und G eine gegebene Abbildung von M in M ist.

Wir kommen nun zu einer Reihe von weiteren Beispielen von induktiven Definitionen. Zunächst wählen wir in (1') für x_0 die Zahl 1 und für G bei festem $m \in \mathbb{N}$ diejenige Abbildung G_m von $\mathbb{N}$ in sich, die der Zahl $x \in \mathbb{N}$ die Zahl $x \cdot m$ zuordnet. Mittels des Rechtfertigungssatzes erhalten wir, daß *es genau eine Abbildung p_m von $\mathbb{N}$ in sich gibt, die den Rekursionsgleichungen*

$$(2) \qquad p_m(0) = 1, \quad p_m(n+1) = p_m(n) \cdot m$$

genügt. Setzen wir noch

$$m \uparrow n := p_m(n)$$

(gelesen: m hoch n), so nehmen die Gleichungen (2) die Form

$$(2') \qquad m \uparrow 0 = 1, \quad m \uparrow (n+1) = (m \uparrow n) \cdot m$$

an, durch die eine eindeutig bestimmte zweistellige Operation in $\mathbb{N}$ definiert wird, die man als *Potenzierung* bezeichnet. Es ist allgemein üblich, das Resultat der Anwendung der Potenzierungsoperation auf ein Paar (m, n) von

natürlichen Zahlen statt durch $m \uparrow n$ durch m^n zu bezeichnen. In diesem Fall heißt m die *Basis* und n der *Exponent*. Wir haben zunächst die andere Schreibweise vorgestellt, um ganz deutlich zu machen, daß es sich auch bei der Potenzierung – wie bei der Addition und der Multiplikation – um eine Operation handelt und die übliche Exponentenschreibweise nur eine gewisse Konvention über die Schreibung der Werte dieser Operation ist. In Exponentenschreibweise erhalten die Rekursionsgleichungen (2') die bekanntere Form

(2'') $\quad m^0 = 1, \quad m^{n+1} = m^n \cdot m.$

Mittels der Rekursionsgleichungen (2'') erhält man mühelos die folgenden Potenzgesetze:

(3) $\quad m^{n_1+n_2} = m^{n_1} \cdot m^{n_2},$

(4) $\quad m^{n_1 \cdot n_2} = (m^{n_1})^{n_2},$

(5) $\quad (m_1 \cdot m_2)^n = m_1^n \cdot m_2^n.$

Die durch vollständige Induktion über n_2 bzw. n zu führenden Beweise seien dem Leser als Übungsaufgaben überlassen. Bei Verwendung des Operationszeichens $\uparrow$ für die Potenzierung nimmt das Gesetz (3) die Form $m \uparrow (n_1 + n_2) = (m \uparrow n_1) \cdot (m \uparrow n_2)$ an, die man als eine Art abgewandelter linksseitiger Distributivität der Potenzierung bzgl. der Addition ansehen kann, abgewandelt insofern, als bei der Verteilung aus der Addition die Multiplikation wird. Das Gesetz (4) nimmt entsprechend die Form $m \uparrow (n_1 \cdot n_2) = (m \uparrow n_1) \uparrow n_2$ an, ist also eine Art Ersatz für die (nicht geltende) Assoziativität der Potenzierungsoperation. Das Gesetz (5) nimmt schließlich die Form

$$(m_1 \cdot m_2) \uparrow n = (m_1 \uparrow n) \cdot (m_2 \uparrow n)$$

an und beinhaltet in dieser Form die rechtsseitige Distributivität der Potenzierung bzgl. der Multiplikation.

Für die Potenzierung gelten ferner die folgenden Monotoniegesetze, deren Beweise ebenfalls dem Leser überlassen bleiben:

(6) $\quad m_1 \leqq m_2 \Rightarrow m_1^n \leqq m_2^n,$

(7) $\quad m_1 < m_2 \wedge n \neq 0 \Rightarrow m_1^n < m_2^n,$

(8) $\quad n_1 \leqq n_2 \wedge m \neq 0 \Rightarrow m^{n_1} \leqq m^{n_2},$

(9) $\quad n_1 < n_2 \wedge m > 1 \Rightarrow m^{n_1} < m^{n_2}.$

Als nächstes Beispiel behandeln wir die induktive Definition der allgemeinen Summe und des allgemeinen Produkts. Dazu sei $(a_\nu)_{\nu \in \mathbb{N}}$ eine mittels der Zahlen aus $\mathbb{N}$ indizierte Familie (Folge) von Elementen einer gegebenen Menge M (vgl. S. 64), d. h. eine Abbildung, die jeder natürlichen

Zahl ν ein bestimmtes Element a_ν der Menge M zuordnet. Weiter setzen wir voraus, daß für die Elemente der Menge M eine als Addition „+" bzw. als Multiplikation „·" bezeichnete zweistellige Operation erklärt sei (es kann also z. B. M die Menge der natürlichen Zahlen mit der in 3.3. erklärten Addition bzw. Multiplikation sein, es kann aber auch M die Menge R der reellen Zahlen mit der dort üblichen Addition bzw. Multiplikation sein oder dergleichen). Dann wird durch

$$(10) \qquad \Sigma(0) = a_0, \quad \Sigma(n+1) = \Sigma(n) + a_{n+1}$$

bzw.

$$(11) \qquad \Pi(0) = a_0, \quad \Pi(n+1) = \Pi(n) \cdot a_{n+1}$$

eine bestimmte Abbildung Σ bzw. Π von N in M definiert, wobei man statt $\Sigma(n)$ bzw. $\Pi(n)$ auch ausführlicher $\sum_{\nu=0}^{n} a_\nu$ (gelesen: Summe ν von 0 bis n über a_ν) bzw. $\prod_{\nu=0}^{n} a_\nu$ (gelesen: Produkt ν von 0 bis n über a_ν) schreibt. Damit nehmen die Rekursionsgleichungen (10) und (11) folgende Form an:

$$(10') \qquad \sum_{\nu=0}^{0} a_\nu = a_0, \quad \sum_{\nu=0}^{n+1} a_\nu = \sum_{\nu=0}^{n} a_\nu + a_{n+1},$$

$$(11') \qquad \prod_{\nu=0}^{0} a_\nu = a_0, \quad \prod_{\nu=0}^{n+1} a_\nu = \prod_{\nu=0}^{n} a_\nu \cdot a_{n+1}.$$

Es ist also z. B.

$$\sum_{\nu=0}^{0} a_\nu = a_0, \quad \sum_{\nu=0}^{1} a_\nu = a_0 + a_1, \quad \sum_{\nu=0}^{2} a_\nu = (a_0 + a_1) + a_2,$$

$$\sum_{\nu=0}^{3} a_\nu = ((a_0 + a_1) + a_2) + a_3, \ \ldots$$

(man beachte die Klammerung!). Bei der vereinbarten Schreibweise soll es allerdings auf die Bezeichnung des sogenannten *Summationsindex* ν nicht ankommen, also statt $\sum_{\nu=0}^{n} a_\nu$ ebensogut $\sum_{\mu=0}^{n} a_\mu$, $\sum_{i=0}^{n} a_i$ usw. geschrieben werden dürfen. Wir merken an, daß zur Berechnung des Wertes $\sum_{\nu=0}^{n} a_\nu$ natürlich nur die Glieder $a_0, \ldots, a_n$ der gegebenen unendlichen Folge $(a_\nu)_{\nu \in \mathbb{N}}$ benötigt werden. Die Betrachtung unendlicher Folgen dient lediglich der formalen Vereinfachung der induktiven Definition.

Ist φ eine eindeutige Abbildung von N in sich, so wird unter $\sum_{\nu=0}^{n} a_{\varphi(\nu)}$ in naheliegender Weise der Wert $\Sigma_\varphi(n)$ der durch

$$\Sigma_\varphi(0) = a_{\varphi(0)}, \quad \Sigma_\varphi(n+1) = \Sigma_\varphi(n) + a_{\varphi(n+1)}$$

definierten Abbildung von N in M verstanden. In diesem Sinne ist also z. B.

$$\sum_{\nu=0}^{3} a_{\nu+2} = ((a_2 + a_3) + a_4) + a_5, \quad \sum_{\nu=0}^{3} a_{2\nu} = ((a_0 + a_2) + a_4) + a_6$$

und analog beim Produkt. Schließlich definiert man noch

$$(10'') \qquad \sum_{\nu=n_0}^{n_0+n} a_\nu := \sum_{\nu=0}^{n} a_{\nu+n_0}, \quad \sum_{\nu=n_0}^{n_0+n} a_{\varphi(\nu)} := \sum_{\nu=0}^{n} a_{\varphi(\nu+n_0)},$$

so daß also z. B.

$$\sum_{\nu=5}^{8} a_\nu = ((a_5 + a_6) + a_7) + a_8, \quad \sum_{\nu=5}^{8} a_{2\nu} = ((a_{10} + a_{12}) + a_{14}) + a_{16}$$

wird.

Wir kommen nun zu den wichtigsten Rechengesetzen für die allgemeine Summe und das allgemeine Produkt. Bei den folgenden Gesetzen wird zunächst nur vorausgesetzt, daß *es sich bei der betrachteten Addition „+" bzw. Multiplikation „·" um eine assoziative Operation in M handelt*. Dann gilt für beliebiges m, n mit $n \neq 0$

$$(12) \qquad \sum_{\nu=0}^{m+n} a_\nu = \sum_{\nu=0}^{m} a_\nu + \sum_{\nu=m+1}^{m+n} a_\nu, \quad \prod_{\nu=0}^{m+n} a_\nu = \prod_{\nu=0}^{m} a_\nu \cdot \prod_{\nu=m+1}^{m+n} a_\nu.$$

Wir merken an, daß in (12) das Assoziativgesetz für „+" bzw. „·" als Spezialfall enthalten ist, also (12) eine gewisse Verallgemeinerung des Assoziativgesetzes für „+" bzw. „·" darstellt. In der Tat wird für $m = 0$, $n = 2$

$$\sum_{\nu=0}^{m+n} a_\nu = \sum_{\nu=0}^{2} a_\nu = (a_0 + a_1) + a_2$$

und

$$\sum_{\nu=0}^{m} a_\nu + \sum_{\nu=m+1}^{m+n} a_\nu = a_0 + (a_1 + a_2)\,.$$

Der Beweis von (12) erfolgt durch vollständige Induktion über n, wobei der Anfangsschritt $n = 1$ (wir hatten $n \neq 0$ vorausgesetzt!) trivial ist. Wir nehmen daher an, daß (12) für $n = n_0 \geqq 1$ schon bewiesen ist, und zeigen, daß dann (12) auch für $n = n_0 + 1$ gilt. Das liefert aber unmittelbar die folgende Gleichungskette:

$$\begin{aligned}
\sum_{\nu=0}^{m+n_0+1} a_\nu &= \sum_{\nu=0}^{m+n_0} a_\nu + a_{m+n_0+1} \\
&= \left(\sum_{\nu=0}^{m} a_\nu + \sum_{\nu=m+1}^{m+n_0} a_\nu\right) + a_{m+n_0+1} \\
&= \sum_{\nu=0}^{m} a_\nu + \left(\sum_{\nu=m+1}^{m+n_0} a_\nu + a_{m+n_0+1}\right) \\
&= \sum_{\nu=0}^{m} a_\nu + \sum_{\nu=m+1}^{m+n_0+1} a_\nu\,.
\end{aligned}$$

Unter Benutzung von (12) können wir nun das folgende allgemeine Assoziativgesetz beweisen: *Für jede beliebig geklammerte Summe $s(a_1, \ldots, a_n)$ bzw. jedes beliebig geklammerte Produkt $p(a_1, \ldots, a_n)$ der Elemente $a_1, \ldots, a_n$ aus M (in dieser Reihenfolge!) gilt*

$$(13) \qquad s(a_1, \ldots, a_n) = \sum_{\nu=1}^{n} a_\nu, \quad p(a_1, \ldots, a_n) = \prod_{\nu=1}^{n} a_\nu,$$

wobei also $\sum_{\nu=1}^{n} a_\nu \left(\prod_{\nu=1}^{n} a_\nu\right)$ die (das) induktiv definierte *kanonisch geklammerte* Summe (Produkt) bedeutet. Der Beweis von (13) erfolgt durch ordnungstheoretische Induktion (vgl. 3.4.(32)) über die Anzahl n der Summanden (bzw. Faktoren). Der Anfangsschritt $n = 1$ wie auch der Fall $n = 2$ gelten trivial; im Fall $n = 3$ reduziert sich (13) im wesentlichen auf das übliche Assoziativgesetz, da es hier neben der kanonischen Klammerung $(a_1 + a_2) + a_3$ nur noch die Klammerung $a_1 + (a_2 + a_3)$ gibt (die Außenklammern haben wir, wie üblich, fortgelassen). Wir nehmen nun an, (13) sei schon für alle Summen mit $k < n$ Summanden bewiesen, und es sei $s(a_1, \ldots, a_n)$ eine beliebig geklammerte Summe mit den n Summanden $a_1, \ldots, a_n$. Dann existiert offenbar eine Zahl $k < n$, so daß

$$s(a_1, \ldots, a_n) = s_1(a_1, \ldots, a_k) + s_2(a_{k+1}, \ldots, a_n),$$

wobei $s_1(a_1, \ldots, a_k)$ bzw. $s_2(a_{k+1}, \ldots, a_n)$ in geeigneter Weise geklammerte Summen mit den Summanden $a_1, \ldots, a_k$ bzw. $a_{k+1}, \ldots, a_n$ sind. Ist z. B.

$$s(a_1, a_2, a_3, a_4, a_5, a_6, a_7) = ((a_1 + a_2) + a_3) + ((a_4 + a_5) + (a_6 + a_7)),$$

so wird $k = 3$ und

$$s_1(a_1, a_2, a_3) = (a_1 + a_2) + a_3, \; s_2(a_4, a_5, a_6, a_7) = (a_4 + a_5) + (a_6 + a_7).$$

Da nun s_1 und s_2 weniger als n Summanden haben, gilt nach Induktionsvoraussetzung

$$s_1(a_1, \ldots, a_k) = \sum_{\nu=1}^{k} a_\nu, \quad s_2(a_{k+1}, \ldots, a_n) = \sum_{\nu=k+1}^{n} a_\nu,$$

und nach (12) wird

$$s(a_1, \ldots, a_n) = \sum_{\nu=1}^{k} a_\nu + \sum_{\nu=k+1}^{n} a_\nu = \sum_{\nu=1}^{n} a_\nu,$$

was zu zeigen war.

Bei den nun folgenden Überlegungen setzen wir voraus, daß *die betrachtete Addition „+" bzw. Multiplikation „·" eine Operation in der gegebenen Menge M ist, die sowohl assoziativ als auch kommutativ ist.* Dann gilt das folgende allgemeine Assoziativ-Kommutativgesetz: *Für beliebige Elemente*

$a_1, \ldots, a_n$ aus M und jede Permutation π der Indizes $1, \ldots, n$ gilt

$$(14) \qquad \sum_{\nu=1}^{n} a_{\pi(\nu)} = \sum_{\nu=1}^{n} a_\nu, \quad \prod_{\nu=1}^{n} a_{\pi(\nu)} = \prod_{\nu=1}^{n} a_\nu.$$

Der Spezialfall $n = 2$ und $\pi = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ liefert das gewöhnliche Kommutativgesetz. Der Spezialfall $n = 3$, $\pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ liefert

$$(a_1 + a_2) + a_3 = (a_2 + a_3) + a_1,$$

woraus unter Anwendung des (als Spezialfall in (14) enthaltenen) gewöhnlichen Kommutativgesetzes sofort

$$(a_1 + a_2) + a_3 = a_1 + (a_2 + a_3),$$

d. h. das gewöhnliche Assoziativgesetz, folgt. (14) ist also in der Tat eine Verallgemeinerung sowohl des Kommutativ- als auch des Assoziativgesetzes (wir werden auch beim Beweis wesentlich das Assoziativgesetz benötigen).

Beim Beweis von (14) können wir uns auf Grund des Satzes 2.4.(20) (der Leser, der 2.4.(20) nicht studiert hat, muß (14) ohne Beweis zur Kenntnis nehmen) auf den Fall beschränken, daß die betrachtete Permutation π eine Transposition $(i\ j)$ mit $1 \leqq i < j \leqq n$ ist, und wegen

$$(i\ j) = (j\ n) \circ (i\ n) \circ (j\ n)$$

genügt es, den Fall $\pi = (k\ n)$ mit $1 \leqq k < n$ zu behandeln. Hierzu merken wir zunächst an, daß auf Grund des Assoziativ- und des Kommutativgesetzes bei beliebigem $a, b, c, d \in M$ die Gleichung

$$(a + b) + (c + d) = (a + d) + (c + b)$$

gilt (Beweis!). Folglich wird

$$\begin{aligned} \sum_{\nu=1}^{n} a_\nu &= \left(\sum_{\nu=1}^{k-1} a_\nu + a_k\right) + \left(\sum_{\nu=k+1}^{n-1} a_\nu + a_n\right) \\ &= \left(\sum_{\nu=1}^{k-1} a_\nu + a_n\right) + \left(\sum_{\nu=k+1}^{n-1} a_\nu + a_k\right) \\ &= \sum_{\nu=1}^{k} a_{\pi(\nu)} + \sum_{\nu=k+1}^{n} a_{\pi(\nu)} = \sum_{\nu=1}^{n} a_{\pi(\nu)} \end{aligned}$$

(im Fall $k = 1$ fehlt der Anteil $\sum_{\nu=1}^{k-1} a_\nu$, im Fall $k = n - 1$ der Anteil $\sum_{\nu=k+1}^{n-1} a_\nu$, so daß sich der Beweis leicht vereinfacht).

Schließlich beweist man durch vollständige Induktion über n, daß *für jede assoziative und kommutative Addition „+" bzw. Multiplikation „·" bei beliebigem*

$a_1, \ldots, a_n, b_1, \ldots, b_n$ aus M die folgenden Gleichungen gelten:

$$(15) \qquad \sum_{\nu=1}^{n}(a_\nu + b_\nu) = \sum_{\nu=1}^{n} a_\nu + \sum_{\nu=1}^{n} b_\nu, \quad \prod_{\nu=1}^{n}(a_\nu \cdot b_\nu) = \prod_{\nu=1}^{n} a_\nu \cdot \prod_{\nu=1}^{n} b_\nu.$$

Als nächstes nehmen wir an, daß *die Multiplikation „·" linksseitig bzw. rechtsseitig distributiv bzgl. der Addition „+" ist. Dann gilt bei beliebigem $a, a_1, \ldots, a_n \in M$* (Beweis!)

$$(16) \qquad a \cdot \left(\sum_{\nu=1}^{n} a_\nu\right) = \sum_{\nu=1}^{n}(a \cdot a_\nu) \text{ bzw. } \left(\sum_{\nu=1}^{n} a_\nu\right) \cdot a = \sum_{\nu=1}^{n}(a_\nu \cdot a).$$

Ist „·" beidseitig distributiv bzgl. „+", so erhalten wir durch Anwendung beider Gleichungen (16), wenn wir hier für a eine Summe $\sum_{\mu=1}^{m} b_\mu$ mit $b_1, \ldots, b_m \in M$ einsetzen:

$$(17) \qquad \sum_{\nu=1}^{n} a_\nu \cdot \sum_{\mu=1}^{m} b_\mu = \sum_{\nu=1}^{n}\left(\sum_{\mu=1}^{m}(a_\nu \cdot b_\mu)\right).$$

Die rechte Seite von (17) kann aufgefaßt werden als eine bestimmte nicht kanonisch geklammerte Summe, deren Summanden die in einer bestimmten Reihenfolge genommenen sämtlichen Produkte $a_\nu \cdot b_\mu$ $(\nu = 1, \ldots, n;$ $\mu = 1, \ldots, m)$ sind. *Ist die Addition „+" noch assoziativ und kommutativ, so kann man diese Summe beliebig umklammern und ihre Summanden in eine beliebige andere Reihenfolge bringen, ohne daß sich ihr Wert ändert.* Dafür schreibt man dann auch kurz

$$\sum_{\nu=1}^{n}\sum_{\mu=1}^{m}(a_\nu \cdot b_\mu) \quad \text{oder} \quad \sum_{\substack{\nu=1,\ldots,n \\ \mu=1,\ldots,m}}(a_\nu \cdot b_\mu)$$

oder dergleichen.

Von nun an bezeichne $+$ bzw. $\cdot$ wieder grundsätzlich die in 3.3. definierte Addition und Multiplikation von natürlichen Zahlen. Durch vollständige Induktion zeigt man leicht, daß *man* in bekannter Weise *die Multiplikation als iterierte Addition und* analog *die Potenzierung als iterierte Multiplikation auffassen kann*, d. h., für beliebiges $n \neq 0$ gilt

$$(18) \qquad m_1 = \cdots = m_n = m \Rightarrow \sum_{\nu=1}^{n} m_\nu = m \cdot n \wedge \prod_{\nu=1}^{n} m_\nu = m^n.$$

Damit (18) auch im Fall $n = 0$ gilt, führt man formal eine *leere Summe* und ein *leeres Produkt* ein, wobei man vereinbart, daß die leere Summe den Wert 0 und das leere Produkt den Wert 1 haben soll.

Als nächstes betrachten wir die durch die Rekursionsgleichungen

$$(19) \qquad \varphi(0) = 1, \quad \varphi(n+1) = \varphi(n) \cdot (n+1)$$

definierte Abbildung von $\mathbf{N}$ in $\mathbf{N}$. Den Funktionswert $\varphi(n)$ bezeichnet man allgemein mit $n!$ (gelesen: n Fakultät), womit wir (19) auch als

$$(19') \qquad 0! = 1, \quad (n+1)! = n! \cdot (n+1)$$

schreiben können. Offenbar gilt also:

$$0! = 1, \quad 1! = 1, \quad 2! = 1 \cdot 2, \quad 3! = 1 \cdot 2 \cdot 3 = 6,$$
$$4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24, \quad 5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120, \ldots,$$

allgemein

$$(20) \qquad n! = \prod_{\nu=1}^{n} \nu,$$

wobei (20) gemäß unserer Vereinbarung über das leere Produkt auch im Fall $n = 0$ richtig ist.

Als Beispiel für eine kompliziertere Art von induktiver Definition (sogenannter *mehrfacher Rekursion*) betrachten wir die folgenden Rekursionsgleichungen:

$$(21) \qquad \begin{aligned} &\psi(n, 0) = 1, \quad \psi(0, k+1) = 0, \\ &\psi(n+1, k+1) = \psi(n, k) + \psi(n, k+1). \end{aligned}$$

Man kann zeigen, daß es genau eine Abbildung ψ von $\mathbf{N} \times \mathbf{N}$ in $\mathbf{N}$ gibt, die bei beliebigem $n, k \in \mathbf{N}$ den Gleichungen (21) genügt. Die sukzessive Berechnung der Funktionswerte $\psi(n, k)$, die man üblicherweise mit $\binom{n}{k}$ (gelesen: n über k) bezeichnet, erfolgt nach folgendem Schema:

(22)

$n \backslash k$	0	1	2	3	4	5	.	.	.
0	1	0	0	0	0	0	.	.	.
1	1	1	0	0	0	0	.	.	.
2	1	2	1	0	0	0	.	.	.
3	1	3	3	1	0	0	.	.	.
4	1	4	6	4	1	0	.	.	.
5	1	5	10	10	5	1	.	.	.
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Zunächst wird entsprechend der ersten Gleichung die erste Spalte mit Einsen ausgefüllt und entsprechend der zweiten Gleichung die erste Zeile von der zweiten Spalte an mit Nullen. Dann kann man mittels der dritten Gleichung für $n = 0$, d. h. $\psi(1, k+1) = \psi(0, k) + \psi(0, k+1)$, die zweite Zeile ausfüllen, indem man nämlich in eine beliebige Spalte die Summe aus der in dieser Spalte und der links daneben stehenden Spalte der ersten Zeile stehenden Zahlen schreibt. Mittels der dritten Gleichung für $n = 1$, d. h. $\psi(2, k+1)$

$= \psi(1, k) + \psi(1, k+1)$ kann man anschließend die dritte Zeile (neben $n = 2$) ausfüllen usw. Sieht man von den oberhalb der Diagonalen stehenden Nullen ab, so kann man sich die Erzeugung der Funktionswerte $\psi(n, k)$ an dem folgenden einfacheren Schema merken:

(22′)

					1					
				1		1				
			1		2		1			
		1		3		3		1		
	1		4		6		4		1	
1		5		10		10		5		1
.		.		.		.		.		.

das in der Literatur als *Pascalsches Dreieck* bekannt ist (nach dem französischen Mathematiker Blaise Pascal (1623–1662)). Die Zahlen $\binom{n}{k}$, die man wegen ihres Auftretens im gleich zu behandelnden binomischen Satz auch als *Binomialkoeffizienten* bezeichnet, spielen in vielen Gebieten der Mathematik eine wesentliche Rolle.

Durch vollständige Induktion über n bestätigt man leicht die folgenden (nach (22) zu vermutenden) Beziehungen:

(23) $\quad n < k \Rightarrow \binom{n}{k} = 0,$

(24) $\quad \binom{n}{n} = 1 \quad$ *für alle* $n \in \mathbb{N}$,

(25) $\quad n \geqq k \Rightarrow \binom{n}{k} = \binom{n}{n-k}.$

Die Behauptung (23) ist für $n = 0$ offensichtlich richtig, da im Fall $k > 0$ eine natürliche Zahl l mit $k = l + 1$ existiert und nach der zweiten Gleichung (21) $\binom{0}{l+1} = 0$ sein soll. Wir nehmen an, daß (23) für $n = n_0$ bereits gilt, d. h. für alle $k > n_0$ die Beziehung $\binom{n_0}{k} = 0$ erfüllt ist, und zeigen, daß dann auch für alle Zahlen $k > n_0 + 1$ die Beziehung $\binom{n_0+1}{k} = 0$ gilt. Ist jedoch $k > n_0 + 1$, so existiert eine Zahl l mit $k = l + 1$, wobei $l > n_0$ und damit erst recht $l + 1 > n_0$ ist, so daß nach Induktionsvoraussetzung sowohl $\binom{n_0}{l} = 0$ als auch $\binom{n_0}{l+1} = 0$ wird. Dann wird aber

$$\binom{n_0+1}{k} = \binom{n_0+1}{l+1} = \binom{n_0}{l} + \binom{n_0}{l+1} = 0 + 0 = 0,$$

was zu zeigen war. Den ((23) benutzenden) Induktionsbeweis für (24) überlassen wir dem Leser als Übungsaufgabe. Beim Beweis von (25) ist der Anfangsschritt $n = 0$ (der sich wegen der Voraussetzung $k \leqq n$ auf den Fall $n = k = 0$ reduziert) trivial. Wir nehmen daher an, (25) sei für $n = n_0$ schon bewiesen, und zeigen, daß (25) dann auch für $n = n_0 + 1$ gilt. Da sich die Fälle $k = 0$ und $k = n_0 + 1$ auf die nach (21) und (24) gültige Gleichung

$$\binom{n_0+1}{0} = \binom{n_0+1}{n_0+1} = 1$$

reduzieren, können wir voraussetzen, daß die Zahl k den Ungleichungen $0 < k < n_0 + 1$ genügt. Dann wird $k = l + 1$ mit $0 \leqq l < n_0$, und es gilt nach Induktionsvoraussetzung:

$$\begin{aligned}\binom{n_0+1}{k} &= \binom{n_0+1}{l+1} = \binom{n_0}{l} + \binom{n_0}{l+1} = \binom{n_0}{n_0-l} + \binom{n_0}{n_0-(l+1)} \\ &= \binom{n_0}{n_0-l} + \binom{n_0}{(n_0-l)-1} = \binom{n_0+1}{n_0-l} = \binom{n_0+1}{(n_0+1)-k},\end{aligned}$$

was zu beweisen war.

Binomischer Satz. *Für beliebige (natürliche) Zahlen a, b gilt für jeden natürlichen Exponenten $n \geqq 1$:*

$$(26) \qquad (a+b)^n = \sum_{\nu=0}^{n} \binom{n}{\nu} a^{n-\nu} b^{\nu}.$$

Wir merken an, daß (26) *auch für beliebige reelle oder komplexe Zahlen a, b richtig ist* (wir werden nämlich beim Beweis nur solche Umformungen verwenden, die auch für beliebige reelle oder komplexe Zahlen gelten). Die Bezeichnung binomischer Satz rührt daher, daß man in der älteren Literatur Terme der Form $a + b$ *Binome* nannte. Rein formal gilt übrigens (26) auch für $n = 0$. Indes ist der Fall $n = 0$ nicht als Anfangsschritt der Induktion ausreichend, da wir im Induktionsschritt maßgeblich (Frage: wo?) die Voraussetzung $n \geqq 1$ (d. h. $n + 1 \geqq 2$) ausnutzen werden. Bei der Niederschrift von (26) haben wir übrigens von der üblichen Konvention Gebrauch gemacht, das Multiplikationszeichen fortzulassen (genau genommen müßte auf der rechten Seite $\binom{n}{\nu} \cdot a^{n-\nu} \cdot b^{\nu}$ stehen).

Der Beweis von (26) erfolgt durch vollständige Induktion über n. Im Anfangsschritt $n = 1$ steht in (26) auf der linken Seite $a + b$ und auf der rechten Seite $\binom{1}{0} a^1 b^0 + \binom{1}{1} a^0 b^1$, was offensichtlich gleich $a + b$ ist. Wir nehmen nun an, daß (26) für den Exponenten $n \geqq 1$ schon bewiesen ist, und zeigen, daß

(26) dann auch für den Exponenten $n+1$ gilt. Nach (2''), der Induktionsvoraussetzung, der linksseitigen Distributivität der Multiplikation bzgl. der Addition und (16) gilt zunächst

$$\begin{aligned}(a+b)^{n+1} &= (a+b)^n (a+b)\\ &= \left(\sum_{\nu=0}^{n} \binom{n}{\nu} a^{n-\nu} b^\nu\right)(a+b)\\ &= \left(\sum_{\nu=0}^{n} \binom{n}{\nu} a^{n-\nu} b^\nu\right) a + \left(\sum_{\nu=0}^{n} \binom{n}{\nu} a^{n-\nu} b^\nu\right) b\\ &= \sum_{\nu=0}^{n} \binom{n}{\nu} a^{n+1-\nu} b^\nu + \sum_{\nu=0}^{n} \binom{n}{\nu} a^{n-\nu} b^{\nu+1}.\end{aligned}$$

Nach (12) $\left(\text{mit } m = 0,\ a_\nu = \binom{n}{\nu} a^{n+1-\nu} b^\nu\right)$ und (10'') wird

$$\begin{aligned}(*)\qquad \sum_{\nu=0}^{n} \binom{n}{\nu} a^{n+1-\nu} b^\nu &= \binom{n}{0} a^{n+1} b^0 + \sum_{\nu=1}^{n} \binom{n}{\nu} a^{n+1-\nu} b^\nu\\ &= a^{n+1} + \sum_{\nu=1}^{n} \binom{n}{\nu} a^{n+1-\nu} b^\nu\\ &= a^{n+1} + \sum_{\nu=0}^{n-1} \binom{n}{\nu+1} a^{n-\nu} b^{\nu+1},\end{aligned}$$

und entsprechend wird nach (12) $\left(\text{mit } m = n-1,\ a_\nu = \binom{n}{\nu} a^{n-\nu} b^{\nu+1}\right)$ und (24)

$$\begin{aligned}\sum_{\nu=0}^{n} \binom{n}{\nu} a^{n-\nu} b^{\nu+1} &= \sum_{\nu=0}^{n-1} \binom{n}{\nu} a^{n-\nu} b^{\nu+1} + \binom{n}{n} a^0 b^{n+1}\\ &= \sum_{\nu=0}^{n-1} \binom{n}{\nu} a^{n-\nu} b^{\nu+1} + b^{n+1}.\end{aligned}$$

Also gilt nach (15) und (21)

$$\begin{aligned}(a+b)^{n+1} &= a^{n+1} + \sum_{\nu=0}^{n-1} \left(\binom{n}{\nu+1} + \binom{n}{\nu}\right) a^{n-\nu} b^{\nu+1} + b^{n+1}\\ &= a^{n+1} + \sum_{\nu=0}^{n-1} \binom{n+1}{\nu+1} a^{n-\nu} b^{\nu+1} + b^{n+1}.\end{aligned}$$

Nochmalige Anwendung der Umformung (*) $\left(\text{mit } \binom{n+1}{\nu} \text{ anstelle von } \binom{n}{\nu}\right)$ ergibt

$$(a+b)^{n+1} = \sum_{\nu=0}^{n} \binom{n+1}{\nu+1} a^{n+1-\nu} b^\nu + b^{n+1} = \sum_{\nu=0}^{n+1} \binom{n+1}{\nu} a^{n+1-\nu} b^\nu,$$

was zu zeigen war.

Setzen wir in (26) speziell $a = b = 1$, so erhalten wir

$$(27) \qquad \sum_{\nu=0}^{n} \binom{n}{\nu} = 2^n.$$

Durch vollständige Induktion über n (Übungsaufgabe) beweist man leicht die folgende Gleichung:

$$(28) \qquad \sum_{\nu=0}^{n} \binom{m+\nu}{\nu} = \binom{m+n+1}{n},$$

woraus für $m = 0$ speziell

$$(29) \qquad \binom{n+1}{n} = n+1$$

und dann mittels (25)

$$(29') \qquad \binom{n+1}{1} = n+1$$

folgt. Wenden wir auf die in (29) links und rechts auftretenden Binomialkoeffizienten die Beziehung (25) an, so erhalten wir

$$(30) \qquad \sum_{\nu=0}^{n} \binom{m+\nu}{m} = \binom{m+n+1}{m+1}.$$

Setzen wir noch $m + n = p$, so kann (30) auch geschrieben werden als

$$(30') \qquad \sum_{\nu=m}^{p} \binom{\nu}{m} = \binom{p+1}{m+1} \quad (p \geqq m).$$

Es sei ferner das folgende wichtige Additionstheorem für die Binomialkoeffizienten genannt, dessen Beweis wir dem Leser als Übungsaufgabe überlassen:

$$(31) \qquad \binom{n_1+n_2}{k} = \sum_{\varkappa=0}^{k} \binom{n_1}{\varkappa}\binom{n_2}{k-\varkappa}.$$

Als nächstes wollen wir die folgende wichtige Beziehung beweisen:

$$(32) \qquad 0 < k \leqq n \Rightarrow \binom{n}{k} \cdot k! = n \cdot (n-1) \cdots (n-k+1).$$

Der Beweis von (32) erfolgt durch vollständige Induktion über n, wobei der Anfangsschritt $n = 0$ trivial ist, da es keine natürliche Zahl k mit $0 < k \leqq 0$ gibt (Implikation mit falscher Prämisse!). Wir nehmen nun an, daß (32) für $n = n_0$ bei beliebigem k bereits gilt, und zeigen, daß dann (32) auch für $n = n_0 + 1$ bei beliebigem k richtig ist. Es sei also $0 < k \leqq n_0 + 1$. Dann existiert eine Zahl l mit $0 \leqq l \leqq n_0$, so daß $k = l + 1$ wird. Ist hierbei $l = 0$,

so wird $k = 1$ und nach (29')

$$\binom{n_0+1}{k} \cdot k! = \binom{n_0+1}{1} \cdot 1 = n_0 + 1,$$

und das ist gerade die Behauptung von (32) im Fall $n = n_0 + 1$, $k = 1$. Ist $l = n_0$, so wird $k = n_0 + 1$, und nach (24) ist

$$\binom{n_0+1}{k} \cdot k! = \binom{n_0+1}{n_0+1} \cdot (n_0+1)! = (n_0+1) \cdot n_0 \cdots 1,$$

und das ist die Behauptung von (32) im betrachteten Fall. Ist schließlich $0 < l < n_0$, so wird nach Induktionsvoraussetzung

$$\begin{aligned}
\binom{n_0+1}{k} \cdot k! &= \binom{n_0+1}{l+1} \cdot (l+1)! \\
&= \binom{n_0}{l} (l+1)! + \binom{n_0}{l+1} \cdot (l+1)! \\
&= n_0 (n_0-1) \cdots (n_0-l+1)(l+1) \\
&\quad + n_0(n_0-1) \cdots (n_0-l+1)(n_0-l) \\
&= n_0 (n_0-1) \cdots (n_0-l+1)(l+1+n_0-l) \\
&= (n_0+1) n_0 \cdots (n_0-l+1) \\
&= (n_0+1) n_0 \cdots (n_0+1-k+1),
\end{aligned}$$

und das ist die Behauptung von (32) im Fall $n = n_0 + 1, 0 < l < n_0$. Wir bemerken, daß die gesonderte Behandlung der Fälle $l = 0$ und $l = n_0$ deshalb erforderlich ist, weil in diesen Fällen die Induktionsvoraussetzung nicht anwendbar ist.

Da $\binom{n}{k}$ eine natürliche Zahl ist, folgt aus (32) speziell, daß im Fall $0 < k \leqq n$ die Zahl $k!$ ein Teiler von $n \cdot (n-1) \cdots (n-k+1)$ ist, so daß wir nach 3.4.(25) den Satz (32) auch in der Form

$$(32') \qquad 0 < k \leqq n \Rightarrow \binom{n}{k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{1 \cdot 2 \cdots k}$$

schreiben können, was durch Erweitern mit $(n-k)!$ noch auf die symmetrische Form

$$(32'') \qquad k \leqq n \Rightarrow \binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

gebracht werden kann.

Als weitere Anwendung des Rechtfertigungssatzes (1') wollen wir zeigen, daß die natürlichen Zahlen durch die Peanoschen Axiome (vgl. 3.2.) bis auf Isomorphie eindeutig charakterisiert sind, genauer:

(33) *Jede Peano-Struktur* $(\tilde{\mathbb{N}}, \tilde{0}, \tilde{\sigma})$ *ist der Peano-Struktur* $(\mathbb{N}, 0, \sigma)$ *isomorph,*

wobei unter einer *Peano-Struktur* (vgl. 3.2.) ein beliebiges Modell des Peanoschen Axiomensystems verstanden wird (zum Isomorphiebegriff vgl. 2.7.(2)). Hieraus folgt natürlich nach 2.7.(3) sofort, daß *je zwei Peano-Strukturen zueinander isomorph sind*. Zum Beweis von (33) bemerken wir zunächst, daß es nach (1′) (genau) eine Abbildung f von N in $\tilde{\mathsf{N}}$ gibt, die folgende Eigenschaften besitzt:

$$(34) \qquad f(0) = \tilde{0}, \quad f(\sigma(n)) = \tilde{\sigma}(f(n)) \quad (n \in \mathsf{N}).$$

Wenn wir zeigen können, daß f eine 1-1-Abbildung von N auf $\tilde{\mathsf{N}}$ ist, haben wir (33) bewiesen; denn dann ist nach (34) f ein Isomorphismus von $(\mathsf{N}, 0, \sigma)$ auf $(\tilde{\mathsf{N}}, \tilde{0}, \tilde{\sigma})$. Hierzu brauchen wir offenbar nur noch zu beweisen, daß es zu jedem $x \in \tilde{\mathsf{N}}$ genau eine Zahl $n \in \mathsf{N}$ mit $f(n) = x$ gibt. Zum Beweis dieser Behauptung sei M die Menge aller derjenigen $x \in \tilde{\mathsf{N}}$, für die das der Fall ist. Dann ist zunächst $\tilde{0} \in M$. Denn einerseits ist nach (34) die Zahl 0 ein Urbild von $\tilde{0}$ bei der Abbildung f. Es ist aber 0 auch die einzige Zahl mit dieser Eigenschaft; gäbe es nämlich eine natürliche Zahl $n \neq 0$ mit $f(n) = \tilde{0}$, dann wäre nach 3.2.(6) $n = \sigma(n_1)$ mit $n_1 \in \mathsf{N}$, und nach (34) wäre

$$\tilde{0} = f(\sigma(n_1)) = \tilde{\sigma}(f(n_1)),$$

im Widerspruch dazu, daß es nach Axiom 3.2.(4′) für $(\tilde{\mathsf{N}}, \tilde{0}, \tilde{\sigma})$ kein $x \in \tilde{\mathsf{N}}$ mit $\tilde{0} = \tilde{\sigma}(x)$ gibt. Es sei nun x_0 ein beliebiges Element der Menge M, d. h. ein Element aus $\tilde{\mathsf{N}}$, das genau ein Urbild n_0 bei der Abbildung f besitzt. Wir behaupten, daß dann auch $\tilde{\sigma}(x_0)$ zu M gehört. Wegen $f(n_0) = x_0$ ist nach (34) zunächst

$$f(\sigma(n_0)) = \tilde{\sigma}(f(n_0)) = \tilde{\sigma}(x_0),$$

d. h., die Zahl $\sigma(n_0)$ ist Urbild von $\tilde{\sigma}(x_0)$ bei der Abbildung f. Ist andererseits n ein beliebiges Urbild von $\tilde{\sigma}(x_0)$ bei f, so muß zunächst $n \neq 0$ sein; denn sonst wäre $\tilde{0} = f(0) = \tilde{\sigma}(x_0)$, im Widerspruch zu Axiom 3.2.(4′) für $(\tilde{\mathsf{N}}, \tilde{0}, \tilde{\sigma})$. Also gibt es eine Zahl $n_1 \in \mathsf{N}$ mit $n = \sigma(n_1)$. Dann wird aber $\tilde{\sigma}(x_0) = f(\sigma(n_1)) = \tilde{\sigma}(f(n_1))$, woraus nach Axiom 3.2.(3′) für $(\tilde{\mathsf{N}}, \tilde{0}, \tilde{\sigma})$ sofort $x_0 = f(n_1)$ folgt, so daß nach Voraussetzung (n_0 sollte das einzige Urbild von x_0 sein) $n_1 = n_0$ und mithin $n = \sigma(n_1) = \sigma(n_0)$ sein muß. Also ist $\sigma(n_0)$ das einzige Urbild von $\tilde{\sigma}(x_0)$, d. h. $\tilde{\sigma}(x_0) \in M$. Damit besitzt die Menge $M (\subseteq \tilde{\mathsf{N}})$ die folgenden Eigenschaften:

$$\tilde{0} \in M, \quad \bigwedge_{x \in \tilde{\mathsf{N}}} (x \in M \Rightarrow \tilde{\sigma}(x) \in M).$$

Folglich enthält nach Axiom 3.2.(5′) für $(\tilde{\mathsf{N}}, \tilde{0}, \tilde{\sigma})$ die Menge M sämtliche Elemente aus $\tilde{\mathsf{N}}$, und das besagt ja nach Definition von M gerade, daß zu jedem Element $x \in \tilde{\mathsf{N}}$ genau eine Zahl $n \in \mathsf{N}$ mit $f(n) = x$ existiert.

Ein Axiomensystem, das nur untereinander isomorphe Modelle besitzt, heißt *kategorisch* oder *monomorph*. Insbesondere ist also nach (34) das Peanosche Axiomensystem kategorisch, wobei man aus dem Beweis von (34) noch leicht entnehmen kann, daß es zu je zwei Modellen dieses Axiomensystems jeweils nur eine einzige isomorphe Abbildung des einen auf das andere Modell gibt. Man sagt hierfür auch, daß das Peanosche Axiomensystem *strikt kategorisch* ist. Wir bemerken, daß man in jeder Peano-Struktur, so wie wir das in den vorangehenden Abschnitten am Beispiel der Struktur $(\mathsf{N}, 0, \sigma)$ vorgeführt haben, eine Addition $\tilde{+}$, eine Multiplikation $\tilde{\cdot}$, eine reflexive totale Ordnung $\tilde{\leqq}$ usw. definieren kann, wobei dann der

konstruierte Isomorphismus f von $(\mathsf{N}, 0, \sigma)$ auf $(\tilde{\mathsf{N}}, \tilde{0}, \tilde{\sigma})$ zugleich auch ein Isomorphismus von $(\mathsf{N}, 0, \sigma, +, \cdot, \leqq, \ldots)$ auf $(\tilde{\mathsf{N}}, \tilde{0}, \tilde{\sigma}, \dot{+}, \tilde{\cdot}, \tilde{\leqq}, \ldots)$ wird.

Wir wollen schließlich noch einen Beweis für die in 2.9.(18) bereits erwähnte Behauptung skizzieren, daß auch die Umkehrung von Satz 2.9.(17) gilt, und zwar wollen wir zeigen, daß *jede unendliche Menge M einer ihrer echten Teilmengen gleichmächtig ist:*

$$(35) \qquad M \text{ unendlich} \Rightarrow \bigvee_N (N \subset M \wedge N \sim M)$$

(wegen der in 2.9.(11) bewiesenen Äquivalenz von Endlichkeit im Sinne von 2.9.(2) und Russellscher Endlichkeit können wir die Voraussetzung von (35) interpretieren als: M ist keinem Abschnitt $\mathcal{A}(n)$ ($n \in \mathsf{N}$) gleichmächtig). Dazu merken wir zunächst an, daß die Funktion σ, die einer beliebigen natürlichen Zahl n ihren unmittelbaren Nachfolger $n+1$ zuordnet, eine 1-1-Abbildung von N auf $\mathsf{N} \setminus \{0\}$ ist, d. h., die Menge N ist ihrer echten Teilmenge $\mathsf{N} \setminus \{0\}$ gleichmächtig. Wir werden nun folgendes zeigen:

$$(36) \qquad M \text{ unendlich} \Rightarrow \bigvee_{N_0} (N_0 \subseteqq M \wedge N_0 \sim \mathsf{N}).$$

Daraus ist leicht (35) zu erhalten. Ist nämlich g 1-1-Abbildung von N_0 auf N mit $g(x_0) = 0$, so wird (Beweis!, vgl. Abb. 9) $g^{-1} \circ \sigma \circ g$ 1-1-Abbildung von N_0 auf $N_0 \setminus \{x_0\}$ und $(g^{-1} \circ \sigma \circ g) \cup e_{M \setminus N_0}$ 1-1-Abbildung von M auf $M \setminus \{x_0\}$ ($\subset M$), womit (35) bewiesen ist.

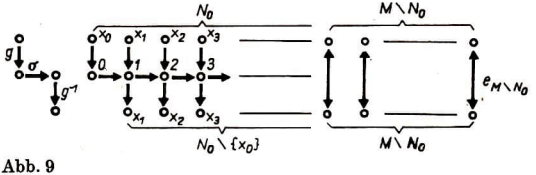


Abb. 9

Zum Beweis von (36) betrachten wir bei gegebenem $x_0 \in M$ zunächst die folgende Menge F:

$$F := \{(X, x) : X \subseteqq M \wedge X \neq \emptyset \wedge x \in X\} \cup \{(\emptyset, x_0)\}.$$

Offenbar ist F eine Korrespondenz von $\mathfrak{P}(M)$ in M. Es sei dann gemäß 2.8.(2) f eine Auswahlfunktion für F, d. h. eine (eindeutige) Abbildung von $\mathfrak{P}(M)$ in M mit $f \subseteqq F$. Die Funktion f ordnet der leeren Menge das Element x_0 und jeder nichtleeren Teilmenge X von M ein eindeutig bestimmtes Element der Menge X zu. Es sei φ die durch die folgenden Rekursionsgleichungen definierte Abbildung von N in $\mathfrak{P}(M)$:

$$\varphi(0) = \emptyset, \quad \varphi(n+1) = \varphi(n) \cup \{f(M \setminus \varphi(n))\}.$$

Durch vollständige Induktion über n zeigt man leicht, daß $|\varphi(n)| = n$ bei beliebigem n ist (hierbei wird wesentlich benutzt, daß M unendlich ist), wobei außerdem $\varphi(0) \subset \varphi(1) \subset \varphi(2) \subset \cdots$ gilt und jeweils $\varphi(n+1) \setminus \varphi(n)$ eine Einermenge ist. Setzen wir

$$N_0 := \bigcup_{n \in \mathsf{N}} \varphi(n)$$

und bezeichnen mit ψ diejenige Abbildung von N in M, die einer gegebenen natürlichen Zahl n das einzige Element aus $\varphi(n+1) \setminus \varphi(n)$ zuordnet, so wird ψ eine 1-1-Abbildung von N auf N_0, womit (36) bewiesen ist.

3.6. Kombinatorische Anzahlbestimmungen

Im vorliegenden Abschnitt wollen wir die bereits in 2.8. begonnenen Anzahlbestimmungen endlicher Mengen systematisch weiterführen. Derartige Anzahlbestimmungen bilden den wesentlichen Gegenstand der sogenannten Kombinatorik, deren einfachste Grundbegriffe wir hier ebenfalls behandeln werden. In 2.9. hatten wir als Elementeanzahl oder Kardinalzahl $|M|$ einer endlichen Menge M die eindeutig bestimmte natürliche Zahl n definiert, für die $M \sim \mathcal{A}(n)$ gilt, wobei $\mathcal{A}(n)$ die Menge aller natürlichen Zahlen i mit $0 \leqq i < n$ bezeichnete. Dort wurden bereits eine Reihe von einfachen Anzahlsätzen vermerkt, die wir noch einmal kurz zusammenstellen wollen:

(1) $\quad |M| = 0 \Leftrightarrow M = \emptyset;$

(2) $\quad |M| = n \wedge x \notin M \Rightarrow |M \cup \{x\}| = n + 1;$

(3) $\quad |M| = n \wedge x \in M \Rightarrow (n > 0 \wedge) \, |M \setminus \{x\}| = n - 1;$

(4) $\quad |M| = n \wedge N \subseteq M \Rightarrow (N \textit{ endlich} \wedge) \, |N| \leqq n;$

(5) $\quad |M| = n \wedge N \sim M \Rightarrow (N \textit{ endlich} \wedge) \, |N| = n.$

Als unmittelbare Folgerung aus (4) erhalten wir, daß *der Durchschnitt $M \cap N$ einer endlichen Menge M mit einer beliebigen Menge N stets endlich ist und $|M \cap N| \leqq |M|$ gilt*:

(6) $\quad M \textit{ endlich} \Rightarrow M \cap N \textit{ endlich} \wedge |M \cap N| \leqq |M|.$

Als erstes wollen wir nun den folgenden naheliegenden Satz beweisen:

(7) $\quad |M_1| = n_1 \wedge |M_2| = n_2 \wedge M_1 \cap M_2 = \emptyset \Rightarrow |M_1 \cup M_2| = n_1 + n_2.$

Den Beweis von (7) führen wir durch vollständige Induktion über n_2. Im Fall $n_2 = 0$ (Anfangsschritt) ist $M_2 = \emptyset$, und die Behauptung von (7) gilt trivial. Wir nehmen nun an (Induktionsvoraussetzung), (7) sei für $n_2 = n$ schon bewiesen, und zeigen (Induktionsbehauptung), daß (7) dann auch für $n_2 = n + 1$ richtig ist. Es sei also $|M_2| = n + 1$. Dann ist nach (1) M_2 nicht leer. Also gibt es ein x mit $x \in M_2$. Nach (3) ist dann $|M_2 \setminus \{x\}| = n$. Wegen $M_1 \cap M_2 = \emptyset$ ist auch $M_1 \cap (M_2 \setminus \{x\}) = \emptyset$, und es ist $x \notin M_1 \cup (M_2 \setminus \{x\})$. Folglich wird nach Induktionsvoraussetzung $|M_1 \cup (M_2 \setminus \{x\})| = n_1 + n$, und (2) liefert schließlich

$$|M_1 \cup M_2| = |M_1 \cup (M_2 \setminus \{x\}) \cup \{x\}| = n_1 + n + 1,$$

womit die Induktionsbehauptung bewiesen ist.

Durch vollständige Induktion über k kann man (7) sofort verallgemeinern zu

(8) $$|M_1| = n_1 \wedge \cdots \wedge |M_k| = n_k \wedge \bigwedge_{i,j} (1 \leqq i < j \leqq k \Rightarrow M_i \cap M_j = \emptyset)$$
$$\Rightarrow |M_1 \cup \cdots \cup M_k| = n_1 + \cdots + n_k.$$

Als unmittelbare Folgerung aus (7) erhalten wir

(9) $$|M_1| = n_1 \wedge |M_2| = n_2 \wedge M_2 \subseteq M_1 \Rightarrow (n_2 \leqq n_1 \wedge) |M_1 \setminus M_2| = n_1 - n_2.$$

Denn offenbar ist im Fall $M_2 \subseteq M_1$ stets $M_1 = M_2 \cup (M_1 \setminus M_2)$ mit
$$M_2 \cap (M_1 \setminus M_2) = \emptyset,$$
also nach (7) $|M_1| = |M_2| + |M_1 \setminus M_2|$, d. h.
$$n_1 = n_2 + |M_1 \setminus M_2|,$$
woraus mittels 3.4.(24) unmittelbar die Behauptung von (9) folgt (daß mit M_1 auch $M_1 \setminus M_2$ endlich, also $|M_1 \setminus M_2|$ eine natürliche Zahl ist, folgt aus (4)).

Sind M_1, M_2 endliche Mengen mit $|M_1| = n_1$, $|M_2| = n_2$, so können wir allgemein nur behaupten, daß $|M_1 \cup M_2| \leqq n_1 + n_2$ ist:

(10) $$|M_1| = n_1 \wedge |M_2| = n_2 \Rightarrow |M_1 \cup M_2| \leqq n_1 + n_2.$$

Unter Verwendung von (9) können wir indes (10) sofort zu der folgenden merkwürdigen Gleichung verschärfen:

(11) $$|M_1| = n_1 \wedge |M_2| = n_2 \Rightarrow |M_1 \cup M_2| + |M_1 \cap M_2| = n_1 + n_2.$$

Wir merken an, daß (7) als Spezialfall in (11) enthalten ist. Zum Beweis von (11) beachten wir, daß
$$M_1 \cup M_2 = M_1 \cup (M_2 \setminus (M_1 \cap M_2))$$
mit $M_1 \cap (M_2 \setminus (M_1 \cap M_2)) = \emptyset$. Also wird nach (7)
$$|M_1 \cup M_2| = |M_1| + |M_2 \setminus (M_1 \cap M_2)|,$$
wobei wegen $M_1 \cap M_2 \subseteq M_2$ nach (9) $|M_2 \setminus (M_1 \cap M_2)| = n_2 - |M_1 \cap M_2|$ ist. Mithin wird $|M_1 \cup M_2| = n_1 + n_2 - |M_1 \cap M_2|$, was nach 3.4.(24) nur eine andere Schreibweise der Behauptung von (11) ist.

Als nächstes wollen wir zeigen, daß folgendes gilt:

(12) $$|M_1| = n_1 \wedge |M_2| = n_2 \Rightarrow |M_1 \times M_2| = n_1 \cdot n_2.$$

Der Beweis von (12) erfolgt durch vollständige Induktion über n_2. Der Anfangsschritt $n_2 = 0$ ist wegen $M_1 \times \emptyset = \emptyset$ trivial. Wir nehmen daher an, (12) sei für $n_2 = n$ schon bewiesen, und zeigen, daß dann (12) auch für $n_2 = n + 1$ richtig ist. Im Fall $n_2 = n + 1$ können wir offenbar M_2 in der Form $M_2 = M_2' \cup \{x\}$ mit $x \notin M_2'$ und $|M_2'| = n$ darstellen, so daß nach Induktionsvoraussetzung $|M_1 \times M_2'| = n_1 \cdot n$ ist. Andererseits ist offenbar
$$M_1 \times M_2 = M_1 \times (M_2' \cup \{x\}) = (M_1 \times M_2') \cup (M_1 \times \{x\})$$

mit $(M_1 \times M_2') \cap (M_1 \times \{x\}) = \emptyset$. Schließlich ist $\{((y, x), y) : y \in M_1\}$ eine 1-1-Abbildung von $M_1 \times \{x\}$ auf M_1, so daß $M_1 \times \{x\} \sim M_1$ und daher nach (5) $|M_1 \times \{x\}| = |M_1| = n_1$ wird. Damit erhalten wir nach (7)

$$|M_1 \times M_2| = |M_1 \times M_2'| + |M_1 \times \{x\}| = n_1 \cdot n + n_1 = n_1 \cdot (n+1),$$

womit die Induktionsbehauptung bewiesen ist.

Durch vollständige Induktion über k kann (12) sofort verallgemeinert werden zu

(13) $\quad |M_1| = n_1 \wedge \cdots \wedge |M_k| = n_k \Rightarrow |M_1 \times \cdots \times M_k| = n_1 \cdots n_k,$

woraus sich im Spezialfall $n_1 = \cdots = n_k = n$ nach 3.5.(18) sofort

(13') $\quad |M| = n \wedge k \geqq 1 \Rightarrow |M^k| = n^k$

ergibt. In 2.4.(26') hatten wir nun bereits bemerkt, daß man die Menge M^k aller k-Tupel von Elementen einer beliebigen Menge M auch auffassen kann als Menge aller Abbildungen von einem beliebigen Indexbereich $I = \{i_1, \ldots, i_k\}$ aus k Elementen (d. h. mit $|I| = k$) in die Menge M. Genauer gilt

(14) $\quad |I| = k \Rightarrow M^I \sim M^k,$

wobei (vgl. 2.4.(6)) M^I die Menge aller Abbildungen von I in M bezeichnet. Der exakte Beweis von (14) sei dem Leser als Übungsaufgabe überlassen. Damit erhalten wir auf Grund von (13') und (5)

(14') $\quad |M| = n \wedge |I| = k \Rightarrow |M^I| = n^k;$

in Worten: *Die Anzahl aller Abbildungen f von einer Menge I mit k Elementen in eine Menge M mit n Elementen ist gleich n^k.*

Unter Verwendung von (14') können wir nun leicht den folgenden Satz beweisen:

(15) $\quad |M| = n \Rightarrow |\mathfrak{P}(M)| = 2^n;$

in Worten: *Eine Menge aus n Elementen besitzt insgesamt 2^n Teilmengen.* Zum Beweis von (15) zeigen wir, daß für jede Menge M folgendes gilt:

(15') $\quad \mathfrak{P}(M) \sim \{0, 1\}^M.$

Hierzu ordnen wir jeder Menge $X \subseteq M$ ihre durch

$$\chi_X^M(x) := \begin{cases} 1, & \text{falls } x \in X, \\ 0, & \text{falls } x \in M \backslash X, \end{cases}$$

definierte *charakteristische Funktion* zu. Man zeigt leicht, daß die durch $\{(X, \chi_X^M) : X \subseteq M\}$ definierte Korrespondenz eine 1-1-Abbildung von $\mathfrak{P}(M)$ auf die Menge aller Abbildungen von M in $\{0, 1\}$, d. h. auf $\{0, 1\}^M$ ist, so daß in der Tat (15') und damit auf Grund von (14') auch (15) gilt (dabei ist lediglich noch $|\{0, 1\}| = 2$ zu beachten).

Wir kommen nun zur kombinatorischen, d. h. mengentheoretischen Deutung der in 3.5.(19′) induktiv definierten Funktion $n!$. Wir behaupten, daß folgendes gilt:

(16) $\quad |M| = n \Rightarrow |\mathfrak{T}(M)| = n!;$

in Worten: *Die Anzahl aller 1-1-Abbildungen einer Menge aus n Elementen auf sich (aller Permutationen einer Menge aus n Elementen) ist gleich n!*. Den Beweis von (16) führen wir durch vollständige Induktion über n. Der Anfangsschritt $n = 0$ ist trivial, da in diesem Fall $M = \emptyset$ und rein formal die leere Abbildung (leere Menge von geordneten Paaren) die einzige 1-1-Abbildung von $\emptyset$ auf $\emptyset$ ist. Wem dies zu gekünstelt oder abstrakt erscheint, mag in (16) den Fall $n = 0$ ausschließen und den Induktionsbeweis mit $n = 1$ beginnen: Hier ist M eine Einermenge und in der Tat e_M die einzige 1-1-Abbildung von M auf sich (vgl. S. 60). Wir nehmen nun an, daß (16) für die Zahl n bereits gilt, und zeigen, daß die Behauptung von (16) dann auch für jede Menge M mit $n + 1$ Elementen richtig ist. Es sei $M = \{a_1, \ldots, a_{n+1}\}$ eine solche. Mit $F_i (i = 1, \ldots, n + 1)$ bezeichnen wir die Menge aller 1-1-Abbildungen f von M auf sich, für die $f(a_{n+1}) = a_i$ gilt. Offenbar sind die Mengen F_i paarweise disjunkt, und es gilt

$$\mathfrak{T}(M) = F_1 \cup \cdots \cup F_{n+1}.$$

Ferner ist $F_{n+1} \sim \mathfrak{T}(M \setminus \{a_{n+1}\})$; denn die Korrespondenz Φ, die einer beliebigen Permutation

$$\begin{pmatrix} a_1 & \ldots & a_n & a_{n+1} \\ a_{i_1} & \ldots & a_{i_n} & a_{n+1} \end{pmatrix}$$

aus F_{n+1} die Permutation

$$\begin{pmatrix} a_1 & \ldots & a_n \\ a_{i_1} & \ldots & a_{i_n} \end{pmatrix}$$

der Menge $M \setminus \{a_{n+1}\}$ $(= \{a_1, \ldots, a_n\})$ zuordnet, ist eine 1-1-Abbildung von F_{n+1} auf $\mathfrak{T}(M \setminus \{a_{n+1}\})$. Also ist nach Induktionsvoraussetzung (nach (3) ist $|M \setminus \{a_{n+1}\}| = n$) und (5) $|F_{n+1}| = n!$. Schließlich ist bei beliebigem $i = 1, \ldots, n$ auch $F_i \sim F_{n+1}$, also ebenfalls nach (5) $|F_i| = n!\,(i = 1, \ldots, n)$; denn die Korrespondenz Φ_i, die einer beliebigen Permutation f aus F_i die Permutation $(a_i\, a_{n+1}) \circ f$ zuordnet (bzgl. der Definition der Transposition $(a_i\, a_{n+1})$ vgl. 2.4.(20)), ist eine 1-1-Abbildung von F_i auf F_{n+1}: Ist nämlich

$$f = \begin{pmatrix} a_1 & \ldots & a_j & \ldots & a_{n+1} \\ a_{i_1} & \ldots & a_{n+1} & \ldots & a_i \end{pmatrix}$$

eine beliebige Permutation aus F_i, so wird

$$\Phi_i(f) = (a_i\, a_{n+1}) \circ f = \begin{pmatrix} a_1 & \ldots & a_j & \ldots & a_{n+1} \\ a_{i_1} & \ldots & a_i & \ldots & a_{n+1} \end{pmatrix},$$

d. h. $\Phi_i(f) \in F_{n+1}$, Φ_i ist also eine Abbildung von F_i in F_{n+1}; ist g eine beliebige Permutation aus F_{n+1}, so ist analog $(a_i\, a_{n+1}) \circ g \in F_i$, wobei wegen $(a_i\, a_{n+1}) \circ (a_i\, a_{n+1}) = e_M$

$$\Phi_i((a_i\, a_{n+1}) \circ g) = (a_i\, a_{n+1}) \circ (a_i\, a_{n+1}) \circ g = g$$

gilt, d. h., Φ_i ist Abbildung von F_i auf F_{n+1}; sind schließlich f_1, f_2 Permutationen aus F_i mit $\Phi_i(f_1) = \Phi_i(f_2)$, so ist $(a_i\, a_{n+1}) \circ f_1 = (a_i\, a_{n+1}) \circ f_2$, und folglich (Multiplikation beider Seiten der letzten Gleichung mit $(a_i\, a_{n+1})$) auch $f_1 = f_2$, d. h., Φ_i ist eindeutig umkehrbar. Aus dem Bewiesenen folgt mittels (8) und 3.5.(19)

$$|\mathfrak{P}(M)| = |F_1| + \cdots + |F_{n+1}| = n! \cdot (n+1) = (n+1)!,$$

womit die Induktionsbehauptung bewiesen ist.

Zur kombinatorischen Deutung der durch 3.5.(21) induktiv definierten Binomialkoeffizienten $\binom{n}{k}$ bezeichnen wir für eine gegebene Menge M und eine gegebene natürliche Zahl k mit $\mathfrak{K}_k(M)$ das System aller derjenigen Teilmengen von M, die die Elementeanzahl k besitzen:

$$(17) \qquad \mathfrak{K}_k(M) := \{X : X \subseteq M \wedge |X| = k\}.$$

Die Elemente des Systems $\mathfrak{K}_k(M)$, d. h. die k-elementigen Teilmengen von M, heißen auch *Kombinationen ohne Wiederholungen von Elementen aus M zur Klasse k.*

Wir behaupten, daß folgendes gilt:

$$(18) \qquad |M| = n \Rightarrow |\mathfrak{K}_k(M)| = \binom{n}{k},$$

d. h., $\binom{n}{k}$ *ist die Anzahl aller der Teilmengen einer Menge M von n Elementen, die genau k Elemente enthalten.*

Nach (18) bestimmt sich z. B. die Anzahl der Tipmöglichkeiten beim Zahlenlotto („5 aus 90") zu

$$\binom{90}{5} = \frac{90 \cdot 89 \cdot 88 \cdot 87 \cdot 86}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 43\,949\,268$$

und beim Sportfesttoto („6 aus 49") zu

$$\binom{49}{6} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 13\,983\,816.$$

Den Beweis von (18) führen wir durch vollständige Induktion über n. Der Fall $n = 0$ ist nach 3.5.(21) trivial, da $\mathfrak{K}_0(\emptyset) = \{\emptyset\}$ und $\mathfrak{K}_{k+1}(\emptyset) = \emptyset$ (es gibt genau eine Teilmenge X der leeren Menge mit $|X| = 0$, nämlich $X = \emptyset$, und

keine Teilmenge X der leeren Mengen mit $|X| = k + 1 > 0$). Wir nehmen nun an, (18) sei für $n = n_0$ bereits bewiesen, und zeigen, daß (18) dann auch für $n = n_0 + 1$ gilt. Es sei also $M = \{a_1, \ldots, a_{n_0+1}\}$ eine beliebige Menge mit $n_0 + 1$ Elementen und k eine beliebige natürliche Zahl. Ist $k = 0$, so wird $\mathfrak{K}_k(M) = \{\emptyset\}$ und in der Tat $|\mathfrak{K}_k(M)| = \binom{n_0+1}{k}$. Ist dagegen $k > 0$, so wird

$$\mathfrak{K}_k(M) := \mathfrak{K}_k(M \setminus \{a_{n_0+1}\}) \cup \mathfrak{K}_k^*,$$

wenn wir mit $\mathfrak{K}_k^*$ das System aller $X \in \mathfrak{K}_k(M)$ mit $a_{n_0+1} \in X$ bezeichnen, wobei überdies $\mathfrak{K}_k(M \setminus \{a_{n_0+1}\}) \cap \mathfrak{K}_k^* = \emptyset$ ist. Wegen $|M \setminus \{a_{n_0+1}\}| = n_0$ ist nach Induktionsvoraussetzung

$$|\mathfrak{K}_k(M \setminus \{a_{n_0+1}\})| = \binom{n_0}{k}.$$

Ferner ist $\mathfrak{K}_k^* \sim \mathfrak{K}_{k-1}(M \setminus \{a_{n_0+1}\})$ (hier brauchen wir die Voraussetzung $k > 0$), da sich jede Menge $X \in \mathfrak{K}_k^*$ in eindeutiger Weise in der Form $X_0 \cup \{a_{n_0+1}\}$ mit $X_0 \in \mathfrak{K}_{k-1}(M \setminus \{a_{n_0+1}\})$ darstellen läßt. Folglich ist nach (5)

$$|\mathfrak{K}_k^*| = |\mathfrak{K}_{k-1}(M \setminus \{a_{n_0+1}\})|,$$

also nach Induktionsvoraussetzung $|\mathfrak{K}_k^*| = \binom{n_0}{k-1}$. Dann wird aber nach (7) und 3.5.(21)

$$|\mathfrak{K}_k(M)| = |\mathfrak{K}_k(M \setminus \{a_{n_0+1}\})| + |\mathfrak{K}_k^*| = \binom{n_0}{k} + \binom{n_0}{k-1} = \binom{n_0+1}{k},$$

was zu zeigen war.

Da für eine beliebige Menge M mit $|M| = n$ offenbar

$$\mathfrak{P}(M) = \mathfrak{K}_0(M) \cup \mathfrak{K}_1(M) \cup \cdots \cup \mathfrak{K}_n(M)$$

ist, wobei die Mengensysteme $\mathfrak{K}_\nu(M)$ $(\nu = 0, \ldots, n)$ überdies paarweise disjunkt sind, wird nach (8)

$$|\mathfrak{P}(M)| = \sum_{\nu=0}^{n} |\mathfrak{K}_\nu(M)|,$$

was wegen (15) und (18) auch als

$$2^n = \sum_{\nu=0}^{n} \binom{n}{\nu}$$

geschrieben werden kann. Damit sind wir zu einem neuen mengentheoretischen Beweis der Formel 3.5.(27) gelangt. Da im Fall $k \leq n$ die Korrespondenz, die einer beliebigen Menge X aus $\mathfrak{K}_k(M)$ die Menge $M \setminus X$ zuordnet, vor allem nach (9) eine 1-1-Abbildung von $\mathfrak{K}_k(M)$ auf $\mathfrak{K}_{n-k}(M)$ ist, wird

$$|\mathfrak{K}_k(M)| = |\mathfrak{K}_{n-k}(M)|,$$

und das ist der mengentheoretische Inhalt der Formel 3.5.(25).

In analoger Weise lassen auch alle weiteren Formeln in den Binomialkoeffizienten eine mehr oder minder einfache kombinatorische Deutung zu (welche kombinatorische Bedeutung hat z. B. das Auftreten der Binomialkoeffizienten im binomischen Satz?).

Für zahlreiche Anwendungen ist die folgende Verallgemeinerung des Begriffs der Kombination zur Klasse k von Bedeutung. Es sei M eine endliche Menge mit $|M| = n$ und $(n_1, \ldots, n_r)$ ein r-Tupel von natürlichen Zahlen, das der Bedingung $\sum_{\varrho=1}^{r} n_\varrho = n$ genügt. Unter einer $(n_1, \ldots, n_r)$-*Kombination* (ohne Wiederholungen) *der Elemente aus* M verstehen wir ein beliebiges r-Tupel $(X_1, \ldots, X_r)$ von Teilmengen von M mit folgenden Eigenschaften:

$$X_1 \cup \cdots \cup X_r = M, \quad X_{\varrho_1} \cap X_{\varrho_2} = \emptyset \ (\varrho_1 \neq \varrho_2), \ |X_\varrho| = n_\varrho.$$

Mit $\mathfrak{K}_{(n_1,\ldots,n_r)}(M)$ bezeichnen wir das System aller dieser $(n_1, \ldots, n_r)$-Kombinationen:

$$(19) \qquad \mathfrak{K}_{(n_1,\ldots,n_r)}(M) := \{(X_1, \ldots, X_r) : (X_1, \ldots, X_r) \textit{ ist } (n_1, \ldots, n_r)\textit{-Kombination von } M\}.$$

Dann gilt

$$(20) \qquad |M| = n \wedge \sum_{\varrho=1}^{r} n_\varrho = n \Rightarrow |\mathfrak{K}_{(n_1,\ldots,n_r)}(M)| = \frac{n!}{n_1! \cdots n_r!}.$$

In (20) ist speziell enthalten, daß im Fall $n = \sum_{\varrho=1}^{r} n_\varrho$ die Zahl $n_1! \cdots n_r!$ stets ein Teiler von $n!$ ist (denn $|\mathfrak{K}_{(n_1,\ldots,n_r)}(M)|$ ist seiner Natur nach eine natürliche Zahl!). Im Fall $r = 2$, $n_1 = k$ $(\leqq n)$, $n_2 = n - k$ nimmt die Behauptung von (20) die Gestalt

$$|\mathfrak{K}_{(k,n-k)}(M)| = \frac{n!}{k!\,(n-k)!} \left(= \binom{n}{k}\right)$$

an. Diese Gleichung ist natürlich nur eine andere Fassung der Behauptung von (18); denn im Fall $k \leqq n$ ist offenbar $X \subseteq M$ genau dann eine Kombination der Elemente von M zur Klasse k, wenn das Paar $(X, M \backslash X)$ eine $(k, n-k)$-Kombination in M ist.

Mit Hilfe von (20) lassen sich Fragen folgenden Typs beantworten: Auf wie viele Weisen lassen sich die 32 verschiedenen Karten eines Kartenspiels an drei Spieler verteilen, so daß jeder Spieler zehn Karten erhält und die beiden restlichen Karten im Talon verbleiben? Antwort:

$$\frac{32!}{10! \cdot 10! \cdot 10! \cdot 2!}.$$

Dem Beweis von (20) schicken wir den folgenden fast trivialen Hilfssatz voraus:

Es sei M_2 eine endliche Menge, f eine Abbildung von M_1 auf M_2 und p eine natürliche Zahl, so daß das volle Urbild $U_f(y)$ (vgl.2.3.(4)) *eines beliebigen Elements $y \in M_2$ jeweils aus genau p Elementen besteht. Dann ist auch die Menge M_1 endlich, und es gilt $|M_1| = p \cdot |M_2|$.*

Ist nämlich $|M_2| = m$ und dabei $M_2 = \{a_1, \ldots, a_m\}$, so wird

$$M_1 = U_f(a_1) \cup \cdots \cup U_f(a_m),$$

wobei die Mengen $U_f(a_\mu)$ $(\mu = 1, \ldots, m)$ paarweise disjunkt sind (vgl. 2.5.(20')) und bei beliebigem $\mu = 1, \ldots, m$ nach Voraussetzung $|U_f(a_\mu)| = p$ ist. Dann ist aber nach (8) $|M_1| = p \cdot m = p \cdot |M_2|$, wie behauptet wurde.

Zum Beweis von (20) sei zunächst $(A_1, \ldots, A_r)$ eine feste $(n_1, \ldots, n_r)$-Kombination der Menge M (eine solche existiert wegen $\sum_{\varrho=1}^{r} n_\varrho = n$). Eine Permutation π der Menge M soll eine $(A_1, \ldots, A_r)$-Permutation genannt werden, wenn π jede der Mengen $A_1, \ldots, A_r$ auf sich abbildet, d. h., wenn $\{\pi(x) : x \in A_\varrho\} = A_\varrho$ für $\varrho = 1, \ldots, r$. Die Menge aller $(A_1, \ldots, A_r)$-Permutationen von M werde mit $\mathfrak{T}(A_1, \ldots, A_r)$ bezeichnet. Für beliebiges $\pi \in \mathfrak{T}(A_1, \ldots, A_r)$ ist offenbar $\pi_\varrho := \pi \mid A_\varrho$ (vgl. 2.4.(9)) eine Permutation der Menge A_ϱ, und es gilt: $\pi = \pi_1 \cup \cdots \cup \pi_r$. Sind umgekehrt $\pi_1, \ldots, \pi_r$ beliebige Permutationen der Mengen $A_1, \ldots, A_r$, so wird $\pi = \pi_1 \cup \cdots \cup \pi_r$ eine $(A_1, \ldots, A_r)$-Permutation. Die $(A_1, \ldots, A_r)$-Permutationen entsprechen also eineindeutig den r-Tupeln $(\pi_1, \ldots, \pi_r) \in \mathfrak{T}(A_1) \times \cdots \times \mathfrak{T}(A_r)$, d. h.

$$\mathfrak{T}(A_1, \ldots, A_r) \sim \mathfrak{T}(A_1) \times \cdots \times \mathfrak{T}(A_r),$$

und mithin gilt nach (5), (13) und (17)

$$|\mathfrak{T}(A_1, \ldots, A_r)| = |\mathfrak{T}(A_1)| \cdots |\mathfrak{T}(A_r)| = n_1! \cdots n_r!.$$

Ist nun π eine beliebige Permutation der Menge M, so sei $f(\pi) := (X_1, \ldots, X_r)$ die durch $X_\varrho := \{\pi(x) : x \in A_\varrho\}$ $(\varrho = 1, \ldots, r)$ definierte $(n_1, \ldots, n_r)$-Kombination. Man sieht leicht, daß f eine Abbildung von $\mathfrak{T}(M)$ auf $\mathfrak{K}_{(n_1, \ldots, n_r)}(M)$ ist. Für eine beliebige $(n_1, \ldots, n_r)$-Kombination $(X_1, \ldots, X_r)$ und eine beliebige Permutation $\pi \in \mathfrak{T}(M)$ mit $f(\pi) = (X_1, \ldots, X_r)$ wird

$$U_f((X_1, \ldots, X_r)) = \{\pi \circ \pi^* : \pi^* \in \mathfrak{T}(A_1, \ldots, A_r)\}$$

(Beweis!). Mithin ist bei beliebigem $(X_1, \ldots, X_r) \in \mathfrak{K}_{(n_1, \ldots, n_r)}(M)$

$$U_f((X_1, \ldots, X_r)) \sim \mathfrak{T}(A_1, \ldots, A_r),$$

also $|U_f((X_1, \ldots, X_r))| = n_1! \cdots n_r!$. Damit können wir (mit $M_1 = \mathfrak{T}(M)$, $M_2 = \mathfrak{K}_{(n_1, \ldots, n_r)}(M)$, $p = n_1! \cdots n_r!$) den zuvor bewiesenen Hilfssatz anwenden und erhalten mittels (17):

$$n! = n_1! \cdots n_r! \cdot |\mathfrak{K}_{(n_1, \ldots, n_r)}(M)|,$$

was zu zeigen war.

Die Zahlen $\frac{n!}{n_1! \cdots n_r!}$ mit $\sum_{\varrho=1}^{r} n_\varrho = n$ nennt man wegen ihres Auftretens im folgenden sogenannten polynomischen Satz auch *Polynomialkoeffizienten*.

Polynomischer Satz. *Für beliebige (natürliche) Zahlen* $a_1, \ldots, a_r$ $(r \geqq 2)$ *gilt für jeden natürlichen Exponenten* $n \geqq 1$:

$$(21) \qquad (a_1 + \cdots + a_r)^n = \sum_{n_1 + \cdots + n_r = n}{}' \frac{n!}{n_1! \cdots n_r!} a_1^{n_1} \cdots a_r^{n_r},$$

wobei $\sum_{n_1 + \cdots n_r = n}{}'$ die (in einer beliebigen Reihenfolge genommene) Summe über alle möglichen Indexkombinationen $(n_1, \ldots, n_r)$ mit $0 \leqq n_\varrho \leqq n$ $(\varrho = 1, \ldots, r)$ und $\sum_{\varrho=1}^{r} n_\varrho = n$ bezeichnet. Im Fall $r = 2$ geht offenbar (21) wegen 3.5. (32″) in den binomischen Satz über.

Den Beweis von (21) führt man am bequemsten bei festem n durch vollständige Induktion über die Anzahl r der Summanden. Im Fall $r = 2$ (Anfangsschritt) reduziert sich – wie bereits bemerkt – (21) auf den binomischen Satz, ist also die Behauptung richtig. Wir nehmen nun an, (21) wäre bereits für alle Summen mit r Summanden bewiesen, und zeigen, daß (21) dann auch für alle Summen mit $r + 1$ Summanden gilt:

$$\begin{aligned}
&(a_1 + \cdots + a_{r+1})^n = (a_1 + \cdots + a_{r-1} + (a_r + a_{r+1}))^n \\
&= \sum_{n_1 + \cdots + n_{r-1} + m = n}{}' \frac{n!}{n_1! \cdots n_{r-1}! \cdot m!} a_1^{n_1} \cdots a_{r-1}^{n_{r-1}} (a_r + a_{r+1})^m \\
&= \sum_{n_1 + \cdots + n_{r-1} + m = n}{}' \frac{n!}{n_1! \cdots n_{r-1}!\, m!} a_1^{n_1} \cdots a_{r-1}^{n_{r-1}} \\
&\qquad \sum_{n_r + n_{r+1} = m}{}' \frac{m!}{n_r!\, n_{r+1}!} a_r^{n_r} \cdot a_{r+1}^{n_{r+1}} \\
&= \sum_{n_1 + \cdots + n_{r+1} = n}{}' \frac{n!}{n_1! \cdots n_{r+1}!} a_1^{n_1} \cdots a_{r+1}^{n_{r+1}}.
\end{aligned}$$

Dabei hat man sich bei der letzten Umformung nur zu überlegen, daß (nach Kürzen durch $m!$) beide Summen aus denselben Summanden aufgebaut sind.

Es sei nun wieder M eine beliebige Menge mit $|M| = n$ und k eine natürliche Zahl mit $1 \leqq k \leqq n$. Unter einer *Variation ohne Wiederholungen von Elementen aus M zur Klasse k* versteht man ein beliebiges k-Tupel $(x_1, \ldots, x_k)$ aus paarweise verschiedenen Elementen der Menge M. Die Menge aller dieser k-Tupel wollen wir mit $\mathfrak{V}_k(M)$ bezeichnen:

$$(22) \qquad \mathfrak{V}_k(M) := \{(x_1, \ldots, x_k) : (x_1, \ldots, x_k) \in M^k \wedge x_1, \ldots, x_k \text{ } \textit{paarweise verschieden}\}.$$

Während es also bei den Kombinationen zur Klasse k nicht darauf ankommt, in welcher Reihenfolge ihre Elemente genommen werden (die Kombinationen sind Mengen), spielt bei einer Variation zur Klasse k diese Reihenfolge (wie eben bei k-Tupeln) eine wesentliche Rolle. Wir wollen zeigen, daß folgendes gilt:

(23) $$|M| = n \wedge 1 \leqq k \leqq n \Rightarrow |\mathfrak{V}_k(M)| = n \cdot (n-1) \cdot \cdots \cdot (n-k+1).$$

Diesen Anzahlsatz benötigt man bei der Lösung von Aufgaben folgenden Typs: Bei wie vielen der Zahlen zwischen 1000 und 9999 besteht die übliche Dezimaldarstellung aus paarweise verschiedenen Ziffern (wie z. B. bei 1023 usw.)? Zur Beantwortung dieser Frage betrachten wir zunächst alle Zahlen $x_1 \cdot 10^3 + x_2 \cdot 10^2 + x_3 \cdot 10 + x_4$ mit $(x_1, x_2, x_3, x_4) \in \mathfrak{V}_4(\{0, \ldots, 9\})$. Ihre Anzahl ist nach (23) gleich

$$10 \cdot 9 \cdot 8 \cdot 7 = 5040.$$

Unter den betrachteten Zahlen kommen nun aber neben den uns interessierenden noch alle Zahlen der Form $x_2 \cdot 10^2 + x_3 \cdot 10 + x_4$ mit

$$(x_2, x_3, x_4) \in \mathfrak{V}_3(\{1, \ldots, 9\})$$

vor. Deren Anzahl ist, wiederum nach (23), gleich

$$9 \cdot 8 \cdot 7 = 504,$$

womit sich die zu bestimmende Anzahl zu

$$5040 - 504 = 4536$$

ergibt.

Ein Beweis von (23) wird z. B. mittels des oben bewiesenen Hilfssatzes folgendermaßen erhalten: Die Korrespondenz f, die einer beliebigen Variation $(x_1, \ldots, x_k) \in \mathfrak{V}_k(M)$ die Kombination $\{x_1, \ldots, x_k\} \in \mathfrak{K}_k(M)$ zuordnet, ist eine Abbildung von $\mathfrak{V}_k(M)$ auf $\mathfrak{K}_k(M)$, für die offenbar bei beliebigem $\{x_1, \ldots, x_k\} \in \mathfrak{K}_k(M)$ die Beziehung $|U_f(\{x_1, \ldots, x_k\})| = k!$ gilt. Mithin wird nach (18) und 3.5.(32′)

$$|\mathfrak{V}_k(M)| = |\mathfrak{K}_k(M)| \cdot k! = \binom{n}{k} \cdot k! = n(n-1) \cdots (n-k+1),$$

was zu zeigen war.

Verzichtet man auf die in (22) gestellte Forderung, daß die Elemente $x_1, \ldots, x_k$ paarweise verschieden sind, so gelangt man zu den sogenannten *Variationen mit Wiederholungen von Elementen aus M zur Klasse k*. Bezeichnen wir die Menge aller dieser mit $\mathfrak{V}_k^w(M)$, so wird offenbar

(24) $$\mathfrak{V}_k^w(M) := M^k,$$

und daher gilt nach (13′)

(25) $$|M| = n \Rightarrow |\mathfrak{V}_k^w(M)| = n^k.$$

Mit Hilfe dieses Anzahlsatzes bestimmt sich z. B. die Anzahl der Tipmöglichkeiten beim Fußballtoto (mit Zusatzzahl) zu $3^{13} = 1\,594\,323$.

Die abschließend zu behandelnden Kombinationen mit Wiederholungen sind insofern etwas problematisch, als sie der Sache nach zwar ebenfalls recht einfach sind, ihre exakte Definition aber einige Schwierigkeiten bereitet. Der Grund hierfür ist, daß es in unserem Begriffssystem keine Objekte gibt, die den Charakter von nicht geordneten Gesamtheiten mit mehrfach auftretenden Elementen haben. Am einfachsten können wir eine Kombination mit Wiederholungen dadurch beschreiben, daß wir mittels einer Abbildung α von M in N für jedes Element $x \in M$ angeben, wie oft es in der betreffenden Kombination gezählt werden soll. Eine solche Abbildung α nennen wir eine *Kombination mit Wiederholungen von Elementen aus M zur Klasse k*, wenn $\sum_{x \in M} \alpha(x) = k$ gilt. Die Menge aller dieser Kombinationen wollen wir mit $\mathfrak{K}_k^w(M)$ bezeichnen:

$$(26) \qquad \mathfrak{K}_k^w(M) := \{\alpha : \alpha : M \to \mathsf{N} \wedge \sum_{x \in M} \alpha(x) = k\}.$$

Wir behaupten, daß folgendes gilt:

$$(27) \qquad |M| = n (\geqq 1) \Rightarrow |\mathfrak{K}_k^w(M)| = \binom{n+k-1}{k}.$$

Mit Hilfe dieses Anzahlsatzes lassen sich Aufgaben folgenden Typs lösen: Wie viele verschiedene Würfe sind mit fünf gleichen (= ununterscheidbaren) Würfeln möglich? Im vorliegenden Fall nehmen wir als Menge M die Menge $\{1, \ldots, 6\}$ der auf jedem Würfel vorhandenen Augenzahlen. Die Abbildung α ordnet jeder Augenzahl $x \in M$ die Anzahl der Würfel zu, die bei einem bestimmten Wurf die Augenzahl x zeigen. Sie charakterisiert vollständig die bei einem bestimmten Wurf hinsichtlich der Augenzahlen zu beobachtende Situation. Durch die Bedingung $\sum_{x \in M} \alpha(x) = 5$ wird festgelegt, daß wir es mit fünf Würfeln zu tun haben. Damit ergibt sich nach (27) die gesuchte Anzahl zu

$$|\mathfrak{K}_5^w(M)| = \binom{10}{5} = 252.$$

Ferner folgt aus (27), daß die Anzahl der Summanden im polynomischen Satz (21) gleich $\binom{n+r-1}{n}$ ist.

Den Beweis von (27) führen wir bei festem k durch vollständige Induktion über n. Im Fall $n = 1$, d. h. $M = \{a_1\}$, ist nur der Fall $\alpha(a_1) = k$ möglich, so daß $|\mathfrak{K}_k^w(M)| = 1$ wird, also wegen $\binom{k}{k} = 1$ die Behauptung von (27) erfüllt ist. Wir nehmen nun an, die Behauptung von (27) sei für alle Mengen M mit

n Elementen bereits bewiesen, und zeigen, daß die Behauptung von (27) dann auch für alle Mengen M mit $|M| = n + 1$ gilt. Es sei also $M = \{a_1, \ldots, a_{n+1}\}$ eine beliebige Menge mit $n + 1$ Elementen. Dann ist offenbar

$$\mathfrak{R}_k^w(M) = \mathfrak{R}_0^* \cup \cdots \cup \mathfrak{R}_k^*,$$

wobei

$$\mathfrak{R}_\varkappa^* = \{\alpha : \alpha : M \to \mathsf{N} \wedge \sum_{x \in M} \alpha(x) = k \wedge \alpha(a_{n+1}) = \varkappa\}$$

($\varkappa = 0, \ldots, k$). Die Menge $\mathfrak{R}_\varkappa^*$ besteht also aus jeweils denjenigen Kombinationen aus $\mathfrak{R}_k^w(M)$, in denen das Element a_{n+1} mit genau der Vielfachheit $\varkappa$ auftritt. Die Mengen $\mathfrak{R}_0^*, \ldots, \mathfrak{R}_k^*$ sind dabei paarweise disjunkt. Bei beliebigem $\alpha \in \mathfrak{R}_\varkappa^*$ ist nun $f_\varkappa(\alpha) := \alpha \setminus \{(a_{n+1}, \varkappa)\}$ eine Abbildung von $M \setminus \{a_{n+1}\}$ in N mit

$$\sum_{x \in M \setminus \{a_{n+1}\}} (f_\varkappa(\alpha))(x) = k - \varkappa, \quad \text{d. h.} \quad f_\varkappa(\alpha) \in \mathfrak{R}_{k-\varkappa}^w(M \setminus \{a_{n+1}\}).$$

Man zeigt nun leicht (Beweis!), daß $f_\varkappa$ eine 1-1-Abbildung von $\mathfrak{R}_\varkappa^*$ auf $\mathfrak{R}_{k-\varkappa}^w(M \setminus \{a_{n+1}\})$ ist, also $\mathfrak{R}_\varkappa^* \sim \mathfrak{R}_{k-\varkappa}^w(M \setminus \{a_{n+1}\})$ gilt. Mithin wird wegen $|M \setminus \{a_{n+1}\}| = n$ nach Induktionsvoraussetzung

$$|\mathfrak{R}_\varkappa^*| = \binom{n + k - \varkappa + 1}{k - \varkappa},$$

also nach (8) und 3.5. (28)

$$\begin{aligned} |\mathfrak{R}_k^w(M)| &= |\mathfrak{R}_0^*| + |\mathfrak{R}_1^*| + \cdots + |\mathfrak{R}_k^*| \\ &= \binom{n+k-1}{k} + \binom{n+k-2}{k-1} + \cdots + \binom{n-1}{0} \\ &= \sum_{\varkappa=0}^{k} \binom{n-1+\varkappa}{\varkappa} = \binom{n+k}{k} \end{aligned}$$

was zu zeigen war.

Durch geeignete Kombination der gewonnenen Resultate lassen sich wesentlich kompliziertere Anzahlbestimmungen durchführen, worauf wir hier jedoch nicht näher eingehen können.

3.7. Elemente der Teilbarkeitstheorie

Wir kommen nun zur Behandlung einer weiteren wichtigen Relation im Bereich der natürlichen Zahlen, der sogenannten *Teilbarkeitsrelation*. Bekanntlich heißt die natürliche Zahl m ein *Teiler* der natürlichen Zahl n und n ein *Viel-*

faches von m, in Zeichen: $m \mid n$ (gelesen: m [ist] Teiler von n oder m teilt n), wenn es eine natürliche Zahl q gibt, so daß $m \cdot q = n$:

$$(1) \qquad m \mid n :\Leftrightarrow \bigvee_{q \in \mathbb{N}} m \cdot q = n.$$

Ein Vergleich mit der Definition 3.4.(1) zeigt, daß die Teilbarkeitsrelation das multiplikative Analogon zur $\leqq$-Relation ist.

Wir zeigen als erstes, daß *die Teilbarkeitsrelation eine reflexive teilweise Ordnung in* $\mathbb{N}$ *ist* (vgl. 2.5.(24)), d. h., *für beliebige natürliche Zahlen* n, n_1, n_2, n_3 *gilt*

$$(2) \qquad n \mid n,$$

$$(3) \qquad n_1 \mid n_2 \wedge n_2 \mid n_3 \Rightarrow n_1 \mid n_3,$$

$$(4) \qquad n_1 \mid n_2 \wedge n_2 \mid n_1 \Rightarrow n_1 = n_2.$$

Zum Beweis von (2) genügt es zu bemerken, daß wegen $n \cdot 1 = n$ eine natürliche Zahl q existiert, für die $n \cdot q = n$ ist, und mithin in der Tat $n \mid n$ gilt. Zum Beweis von (3) sei $n_1 \mid n_2$ und $n_2 \mid n_3$. Dann existieren wegen (1) natürliche Zahlen q_1, q_2 mit $n_1 \cdot q_1 = n_2$ und $n_2 \cdot q_2 = n_3$. Dann wird aber $n_1 \cdot (q_1 \cdot q_2) = (n_1 \cdot q_1) \cdot q_2 = n_2 \cdot q_2 = n_3$, d. h., für die Zahl $q = q_1 \cdot q_2$ gilt $n_1 \cdot q = n_3$, und es ist in der Tat $n_1 \mid n_3$. Zum Beweis von (4) sei $n_1 \mid n_2$ und $n_2 \mid n_1$, und zwar sei $n_1 \cdot q_1 = n_2$ und $n_2 \cdot q_2 = n_1$. Dann wird $n_1 \cdot (q_1 \cdot q_2) = (n_1 \cdot q_1) \cdot q_2 = n_2 \cdot q_2 = n_1 \cdot 1$, woraus im Fall $n_1 \neq 0$ nach 3.4.(23) $q_1 \cdot q_2 = 1$ und dann nach 3.3.(20) $q_1 = q_2 = 1$ folgt, so daß in der Tat $n_1 = n_2$ ist, während im Fall $n_1 = 0$ wegen $n_1 \cdot q_1 = n_2$ auch $n_2 = 0$ ist.

Im Gegensatz zur $\leqq$-Relation ist natürlich die Teilbarkeitsrelation keine totale Ordnung, d. h., es gibt unvergleichbare Elemente (z. B. gilt weder $2 \mid 3$ noch $3 \mid 2$). Wir werden jedoch sehen, daß starke Analogien zwischen der Teilbarkeitsrelation und der Inklusion bestehen.

Zunächst ergibt sich aus $m \cdot 0 = 0$ sofort, daß *die Zahl* 0 *Vielfaches jeder natürlichen Zahl* m *ist:*

$$(5) \qquad \bigwedge_{m \in \mathbb{N}} m \mid 0.$$

Die Zahl 0 ist also bezüglich der Teilbarkeitsrelation *größtes Element* in $\mathbb{N}$ (bezüglich der $\leqq$-Relation ist in $\mathbb{N}$ kein größtes Element vorhanden), und wegen (4) ist die Zahl 0 auch die einzige Zahl mit dieser Eigenschaft; denn ist $m \mid n_0$ für alle $m \in \mathbb{N}$, so ist speziell $0 \mid n_0$, und umgekehrt ist nach (5) $n_0 \mid 0$, also gilt nach (4) $n_0 = 0$.

Wegen $1 \cdot n = n$ *ist die Zahl* 1 *Teiler jeder natürlichen Zahl* n*:*

$$(6) \qquad \bigwedge_{n \in \mathbb{N}} 1 \mid n.$$

d. h., die Zahl 1 ist bezüglich der Teilbarkeitsrelation *kleinstes Element* in N (wie die Zahl 0 bzgl. der $\leqq$-Relation und die leere Menge bzgl. der Inklusion), und wegen (4) ist die Zahl 1 auch die einzige Zahl mit dieser Eigenschaft (Beweis!).

Ferner erkennt man leicht, daß zwischen der Teilbarkeitsrelation und der $\leqq$-Relation folgende Beziehung besteht:

(7) $\quad m \mid n \wedge n \neq 0 \Rightarrow m \leqq n,$

d. h., vom Fall $n = 0$ abgesehen ist die Teilbarkeitsrelation eine Teilrelation der $\leqq$-Relation. Ist nämlich $m \mid n$, so existiert eine Zahl $q \in \mathbb{N}$ mit $m \cdot q = n$, wobei im Fall $n \neq 0$ die Zahl q von Null verschieden, also $q \geqq 1$ ist. Dann ist aber (vgl. 3.3.(17)) $m = m \cdot 1 \leqq m \cdot q = n$, was zu zeigen war.

Als nächstes zeigt man, daß *bei beliebigem* $m, n_1, n_2 \in \mathbb{N}$ *folgendes gilt*:

(8) $\quad m \mid n_1 \wedge m \mid n_2 \Rightarrow m \mid n_1 + n_2,$

(9) $\quad m \mid n_1 \wedge m \mid n_2 \wedge n_1 \geqq n_2 \Rightarrow m \mid n_1 - n_2.$

Der Beweis von (8) sei dem Leser als Übungsaufgabe überlassen. Da (9) im Fall $m = 0$ trivial gilt (Beweis!), können wir im folgenden Beweis für (9) $m \neq 0$ voraussetzen. Ist dann $n_1 = q_1 m$, $n_2 = q_2 m$, so ist mit $n_1 \geqq n_2$ nach 3.4.(21) und 3.4.(22) auch $q_1 \geqq q_2$, so daß eine Zahl $q \in \mathbb{N}$ mit $q_2 + q = q_1$ existiert, und folglich wird $q_2 m + qm = q_1 m$, also $n_2 + qm = n_1$. Hieraus folgt nach 3.4.(24) $n_1 - n_2 = qm$, d. h. $m \mid n_1 - n_2$, was zu zeigen war.

Ohne Schwierigkeit erhält man noch, daß *bei beliebigem* $m, n, k \in \mathbb{N}$ *folgendes gilt*:

(10) $\quad m \mid n \Rightarrow m \cdot k \mid n \cdot k,$

d. h., *die Multiplikation ist monoton bzgl. der Teilbarkeitsrelation* (vgl. 2.6.(12)). Natürlich läßt sich (10) sofort zu

(10′) $\quad m_1 \mid m_2 \wedge n_1 \mid n_2 \Rightarrow m_1 \cdot n_1 \mid m_2 \cdot n_2$

verschärfen, woraus mittels (6) noch

(10″) $\quad n_1 \mid n_2 \Rightarrow n_1 \mid k \cdot n_2$

folgt.

Als Hilfssatz für die weiteren Überlegungen benötigen wir den folgenden

Satz über die Division mit Rest. *Es sei* (n, m) *ein beliebiges Paar von natürlichen Zahlen mit* $m \neq 0$. *Dann gibt es genau ein Paar* (q, r) *von natürlichen Zahlen, so daß folgendes gilt:*

(11) $\quad n = q \cdot m + r \quad$ *und* $\quad 0 \leqq r < m.$

Die durch n und m eindeutig bestimmten Zahlen q und r heißen der *Quotient* bzw. der *Rest* bei Division von n durch m und sollen im folgenden mit $\boldsymbol{q}(n, m)$ bzw. $\boldsymbol{r}(n, m)$ bezeichnet werden.

Wir zeigen als erstes, daß es höchstens ein derartiges Paar (q, r) von natürlichen Zahlen geben kann. Dazu nehmen wir an, es gelte

(i) $\quad n = q_1 \cdot m + r_1 \quad$ mit $\quad 0 \leqq r_1 < m,$

(ii) $\quad n = q_2 \cdot m + r_2 \quad$ mit $\quad 0 \leqq r_2 < m.$

Wäre hierbei $r_1 \neq r_2$, so müßte nach 3.4.(13) entweder $r_1 < r_2$ oder $r_2 < r_1$ gelten. Ohne Beschränkung der Allgemeinheit können wir uns auf die Betrachtung des Falles $r_2 < r_1$ beschränken. In diesem Fall gäbe es nach 3.4.(15) eine natürliche Zahl $k \neq 0$ mit $r_2 + k = r_1$, und nach (i), (ii) und 3.4.(22) wäre $q_2 \cdot m = q_1 \cdot m + k$, also $q_1 \cdot m < q_2 \cdot m$ und mithin nach 3.4.(21) $q_1 < q_2$, also nach 3.4.(29′) $q_1 + 1 \leqq q_2$. Dann müßte aber wegen $r_1 < m$ nach (i), 3.4.(18), 3.4.(17) und (ii)

$$\begin{aligned} n &= q_1 \cdot m + r_1 < q_1 \cdot m + m \\ &= (q_1 + 1) \cdot m \leqq q_2 \cdot m \leqq q_2 \cdot m + r_2 = n, \end{aligned}$$

d. h. $n < n$ gelten, was offenbar ein Widerspruch ist. Also ist unsere Annahme $r_1 \neq r_2$ falsch, und es gilt $r_1 = r_2$. Dann wird aber nach (i) und (ii) auf Grund von 3.4.(22) $q_1 \cdot m = q_2 \cdot m$, also wegen $m \neq 0$ nach 3.4.(23) $q_1 = q_2$, womit die Einzigkeit eines Paares (q, r) mit (11) bewiesen ist.

Zum Existenzbeweis betrachten wir die Menge M aller natürlichen Zahlen x, für die die Ungleichung $x \cdot m \leqq n$ erfüllt ist. Die Menge M ist sicher nicht leer, da die Zahl 0 zu M gehört. Die Menge M ist ferner nach oben beschränkt, da wegen $m \neq 0$ alle Zahlen $x \in M$ der Ungleichung $x \leqq n$ genügen. Folglich gibt es nach 3.4.(34) in M eine größte Zahl q. Für diese gilt

(12) $\quad q \cdot m \leqq n < (q + 1) \cdot m.$

Wegen $q \cdot m \leqq n$ existiert eine natürliche Zahl r, so daß $q \cdot m + r = n$, wobei wegen $q \cdot m + r < q \cdot m + m$ nach 3.4.(20) $r < m$ ist. Also erfüllen die konstruierten Zahlen q, r die Bedingungen (11), was noch zu zeigen war.

Zugleich haben wir gezeigt, daß *es bei beliebigem* $m \neq 0$ *und* n *genau eine Zahl* $q \in \mathbb{N}$ *gibt, für die* (12) *gilt.*

Wir merken an, daß *im Fall* $m \neq 0$ *die Zahl* m *genau dann ein Teiler der Zahl* n *ist, wenn* $\boldsymbol{r}(n, m) = 0$ *ist:*

(13) $\quad m \neq 0 \Rightarrow (m \mid n \Leftrightarrow \boldsymbol{r}(n, m) = 0).$

Als nächstes wollen wir zeigen, daß *es zu beliebigen natürlichen Zahlen* m, n *genau eine natürliche Zahl* d *gibt, die den folgenden Bedingungen*

genügt:

(14a) $d \mid m, \quad d \mid n,$

(14b) $\bigwedge_{t \in \mathbf{N}} (t \mid m \wedge t \mid n \Rightarrow t \mid d).$

Eine Zahl d mit den Eigenschaften (14a) nennt man einen *gemeinsamen Teiler* der Zahlen m, n, und die durch m und n eindeutig bestimmte Zahl d mit den Eigenschaften (14a) und (14b) heißt der *größte gemeinsame Teiler* der Zahlen m und n und werde im folgenden mit $m \sqcap n$ bezeichnet (in der Literatur ist auch die Bezeichnung (m, n) üblich, die wir jedoch konsequent für das geordnete Paar benutzen; vielfach wird auch die Bezeichnung $\mathrm{ggT}(m, n)$ verwendet). Wir machen mit allem Nachdruck darauf aufmerksam, daß entsprechend (14b) der größte gemeinsame Teiler *die bezüglich der teilweisen Ordnung | größte Zahl ist, die gemeinsamer Teiler der Zahlen m und n ist, und nicht etwa bezüglich der totalen Ordnung* $\leqq$, wie vielfach fälschlich definiert wird. Im Bereich der natürlichen Zahlen ist das wegen (7) zwar im wesentlichen dasselbe, jedoch ist die Bedingung (14b) und nicht die Bedingung

(14b') $\bigwedge_{t \in \mathbf{N}} (t \mid m \wedge t \mid n \Rightarrow t \leqq d)$

die entscheidende Eigenschaft des größten gemeinsamen Teilers. Dabei ist u. a. zu beachten, daß man es mit Teilbarkeit auch in anderen Bereichen als den natürlichen Zahlen zu tun hat (z. B. bei Polynomen), in denen es kein Analogon zur $\leqq$-Relation gibt. Die Teilbarkeitsrelation und der durch (14a) und (14b) charakterisierte größte gemeinsame Teiler sind ihrer Natur nach zunächst allein mittels der Multiplikation definiert, während die $\leqq$-Relation (vgl. 3.4.(1)) auf der Addition basiert. Wir machen ferner darauf aufmerksam, daß die Eigenschaften (14a) und (14b) das formale Analogon zu den Eigenschaften 1.5.(8) und 1.5.(9) des Durchschnitts bezüglich der Inklusion sind.

Zum Nachweis der Einzigkeit des größten gemeinsamen Teilers nehmen wir an, daß die Zahlen d_1 und d_2 die Eigenschaften (14) besitzen. Aus (14a) für d_1 und (14b) für d_2 folgt dann $d_1 \mid d_2$ und aus (14b) für d_1 und (14a) für d_2 analog $d_2 \mid d_1$, so daß nach (4) in der Tat $d_1 = d_2$ ist.

Da im Fall $m = 0$ die Zahl $d = n$ und im Fall $n = 0$ die Zahl $d = m$ die Eigenschaften (14) besitzt (Beweis!), beschränken wir uns im folgenden Existenzbeweis auf den Fall $m \neq 0$, $n \neq 0$. Wir geben zunächst einen Beweis, der auf C. F. Gauss zurückgeht und auf einer wichtigen Darstellungsmöglichkeit für $m \sqcap n$ beruht. Dazu betrachten wir die Menge $D(m, n)$ aller natürlichen Zahlen $x \neq 0$, die sich in der Form $am - bn$ mit $a, b \in \mathbf{N}$ darstellen lassen,

$$D(m, n) := \{x : x \in \mathbf{N} \wedge x \neq 0 \wedge \bigvee_{a,b \in \mathbf{N}} x = am - bn\},$$

sowie analog die Menge $D(n, m)$ aller natürlichen Zahlen $x \neq 0$, die eine Darstellung der Form $\bar{a}n - \bar{b}m$ mit $\bar{a}, \bar{b} \in \mathbf{N}$ besitzen.

Wir beweisen als erstes, daß *im Fall* $m \neq 0$, $n \neq 0$ *die Mengen* $D(m, n)$ *und* $D(n, m)$ *gleich sind*. Dazu sei $x = am - bn$ eine beliebige Zahl der Menge $D(m, n)$. Dann ist wegen $x \neq 0$ offenbar $am \geqq bn \geqq b$, d. h., es existiert eine Zahl $\bar{a} \in \mathbf{N}$ mit $b + \bar{a} = am$. Ebenso ist $an \geqq a$, und es existiert eine Zahl $\bar{b} \in \mathbf{N}$ mit $\bar{b} + a = an$. Dann ist $bn + \bar{a}n = \bar{b}m + am$, und es gilt $am + \bar{a}n = x + bn + \bar{a}n = x + \bar{b}m + am$ und mithin $\bar{a}n = x + \bar{b}m$, also $x = \bar{a}n - \bar{b}m$, so daß $x \in D(n. m)$. Folglich ist $D(m, n) \subseteqq D(n, m)$, und aus Symmetriegründen gilt dann natürlich auch die umgekehrte Inklusion.

Wegen $m \in D(m, n)$ $(a = 1, b = 0)$ ist die Menge $D(m, n)$ nicht leer. Folglich gibt es nach 3.4.(31) in $D(m, n)$ eine (bezüglich $\leqq$) kleinste Zahl d. Wir behaupten, daß diese Zahl d die Eigenschaften (14) besitzt. Zum Beweis von (14b) gehen wir davon aus, daß wegen $d \in D(m, n)$ natürliche Zahlen a_0, b_0 mit $d = a_0 m - b_0 n$ existieren (wobei wegen $d \in D(m, n)$ natürlich $a_0 m > b_0 n$ ist). Ist nun t eine beliebige natürliche Zahl mit $t \mid m$ und $t \mid n$, so ist nach (10″) $t \mid a_0 m$ und $t \mid b_0 n$, also nach (9) $t \mid a_0 m - b_0 n$, d. h. $t \mid d$, wie für (14b) zu zeigen war. Zum Beweis von (14a) wenden wir auf m und d (wegen $d \in D(m, n)$ ist $d \neq 0$) den Satz (11) über die Division mit Rest an:

$$m = qd + r \quad \text{mit} \quad r < d.$$

Nun läßt sich d (wegen $D(m, n) = D(n, m)$) auch in der Form $d = \bar{a}_0 n - \bar{b}_0 m$ mit $\bar{a}_0, \bar{b}_0 \in \mathbf{N}$ darstellen. Folglich wird

$$m = qd + r = q(\bar{a}_0 n - \bar{b}_0 m) + r = q\bar{a}_0 n - q\bar{b}_0 m + r,$$

also $(q\bar{b}_0 + 1)\, m = q\bar{a}_0 n + r$, und mithin

$$r = (q\bar{b}_0 + 1)\, m - (q\bar{a}_0)\, n,$$

woraus im Fall $r \neq 0$ unmittelbar $r \in D(m, n)$ folgen würde, was wegen $r < d$ im Widerspruch zur Minimalität von d in $D(m, n)$ steht. Also ist $r = r(m, d) = 0$ und wegen (13) d ein Teiler von m. Aus Symmetriegründen ist dann natürlich auch $d \mid n$.

Damit haben wir zugleich den folgenden wichtigen Satz bewiesen:

(15) $\quad m \neq 0 \wedge n \neq 0$

$$\Rightarrow m \sqcap n = \min\{x : x \in \mathbf{N} \wedge x \neq 0 \wedge \bigvee_{a,b \in \mathbf{N}} x = am - bn\}$$

$$= \min\{x : x \in \mathbf{N} \wedge x \neq 0 \wedge \bigvee_{\bar{a},\bar{b} \in \mathbf{N}} x = \bar{a}n - \bar{b}m\}.$$

Man kann übrigens noch leicht zeigen, daß *die im Beweis von* (15) *betrachtete Menge* $D(m, n)$ *gerade aus sämtlichen positiven Vielfachen der Zahl* $m \sqcap n$ *besteht* (Übungsaufgabe).

Aus der Definition (14) ergeben sich mühelos die folgenden Eigenschaften des größten gemeinsamen Teilers (vgl. dazu die Sätze 1.4.(5), 1.4.(8), 1.4.(13) und 1.5.(16) über den Durchschnitt):

(16) $m \sqcap n = n \sqcap m$,

(17) $n_1 \sqcap (n_2 \sqcap n_3) = (n_1 \sqcap n_2) \sqcap n_3$,

(18) $n \sqcap n = n$,

(19) $m \mid n \Leftrightarrow m \sqcap n = m$.

Wir beweisen als Beispiel das Assoziativgesetz (17) und überlassen die analogen Beweise der anderen Behauptungen dem Leser als Übungsaufgaben. Zur Abkürzung werde $n_1 \sqcap (n_2 \sqcap n_3)$ gleich d gesetzt. Dann gilt zunächst nach (14a) $d \mid n_1$ und $d \mid n_2 \sqcap n_3$ und wegen $n_2 \sqcap n_3 \mid n_2$ und $n_2 \sqcap n_3 \mid n_3$ nach (3) auch $d \mid n_2$ und $d \mid n_3$. Folglich ist nach (14b) $d \mid n_1 \sqcap n_2$ und nochmals nach (14b) $d \mid (n_1 \sqcap n_2) \sqcap n_3$. Also gilt $n_1 \sqcap (n_2 \sqcap n_3) \mid (n_1 \sqcap n_2) \sqcap n_3$, und analog zeigt man, daß auch umgekehrt $(n_1 \sqcap n_2) \sqcap n_3 \mid n_1 \sqcap (n_2 \sqcap n_3)$, so daß nach (4) in der Tat (17) gilt.

Auf Grund von (17) können wir also, ohne Mißverständnisse befürchten zu müssen, in $n_1 \sqcap (n_2 \sqcap n_3)$ und $(n_1 \sqcap n_2) \sqcap n_3$ die Klammern fortlassen und kurz $n_1 \sqcap n_2 \sqcap n_3$ schreiben, wobei nach 3.5.(13) analoges bei vier- und mehrgliedrigen $\sqcap$-Termen möglich ist. In diesem Sinne bezeichnet also

$$n_1 \sqcap n_2 \sqcap \cdots \sqcap n_k$$

einen beliebig geklammerten $\sqcap$-Term aus den Zahlen $n_1, \ldots, n_k$ (in dieser Reihenfolge) z. B. den kanonisch geklammerten Term

$$(\cdots ((n_1 \sqcap n_2) \sqcap n_3) \sqcap \cdots) \sqcap n_k.$$

Man zeigt leicht, daß *bei beliebigem k die Zahl $d = n_1 \sqcap \cdots \sqcap n_k$ durch die folgenden Eigenschaften charakterisiert ist* (für $k = 3$ ist das im obigen Beweis für (17) enthalten):

(20a) $d \mid n_1, \ldots, d \mid n_k$,

(20b) $\bigwedge_{t \in \mathbb{N}} (t \mid n_1 \wedge \cdots \wedge t \mid n_k \Rightarrow t \mid d)$.

Aus diesem Grunde nennt man $n_1 \sqcap \cdots \sqcap n_k$ den *größten gemeinsamen Teiler der Zahlen $n_1, \ldots, n_k$*.

Auf Grund von (16) und (17) kommt es nach 3.5.(14) in $n_1 \sqcap \cdots \sqcap n_k$ auch nicht auf die Reihenfolge der Glieder an, d. h., *für jede Permutation π der Indizes $1, \ldots, k$ gilt*

(16′) $n_{\pi(1)} \sqcap \cdots \sqcap n_{\pi(k)} = n_1 \sqcap \cdots \sqcap n_k$,

was man übrigens auch unmittelbar (20) entnehmen kann.

Auch der Satz (15) läßt sich leicht auf mehrgliedrige $\sqcap$-Terme verallgemeinern. Dazu ist es allerdings zweckmäßig, von ganzen Zahlen Gebrauch zu machen. Offenbar können wir nach (15) sagen, daß *im Fall* $m \neq 0$, $n \neq 0$ *der größte gemeinsame Teiler* $m \sqcap n$ *die kleinste positive Zahl ist, die sich in der Form* $\alpha m + \beta n$ *mit* $\alpha, \beta \in \mathbf{Z}$ (Z Menge aller ganzen Zahlen) *darstellen läßt.* In Verallgemeinerung hiervon gilt:

(15′) *Sind die Zahlen* $n_1, \ldots, n_k \in \mathbf{N}$ *nicht sämtlich gleich Null, so ist* $n_1 \sqcap \cdots \sqcap n_k$ *die kleinste positive Zahl, die sich in der Form*
$\alpha_1 n_1 + \cdots + \alpha_k n_k$
mit $\alpha_1, \ldots, \alpha_k \in \mathbf{Z}$ *darstellen läßt.*

Natürlich läßt sich (15′) unter Verwendung der Differenz auch allein in natürlichen Zahlen ausdrücken, wird dann aber entsprechend der Vielzahl von Möglichkeiten recht kompliziert (z. B. ist also $n_1 \sqcap n_2 \sqcap n_3$ die kleinste von Null verschiedene Zahl in der Menge aller der natürlichen Zahlen, die sich in einer der Formen $\pm a_1 n_1 \pm a_2 n_2 \pm a_3 n_3$ mit $a_1, a_2, a_3 \in \mathbf{N}$ darstellen lassen).

Aus (19) und (5) folgt noch leicht, daß *bei beliebigem* $n \in \mathbf{N}$

(21) $\quad n \sqcap 0 = 0 \sqcap n = n$

ist (wie schon im Beweis von (15) bemerkt), d. h., *die Zahl* 0 *ist beidseitig neutrales Element der Operation* $\sqcap$ (vgl. 2.6.(11)). Entsprechend folgt aus (19) und (6), daß *bei beliebigem* $n \in \mathbf{N}$ (in Analogie zu 1.4.(18))

(22) $\quad n \sqcap 1 = 1 \sqcap n = 1.$

Schließlich merken wir an, daß *die Multiplikation beidseitig distributiv bzgl. der Operation* $\sqcap$ *ist* (vgl. 2.6.(9)), d. h., *bei beliebigem* $m, n_1, n_2 \in \mathbf{N}$ *gilt*

(23) $$\begin{aligned} m \cdot (n_1 \sqcap n_2) &= (m \cdot n_1) \sqcap (m \cdot n_2), \\ (n_1 \sqcap n_2) \cdot m &= (n_1 \cdot m) \sqcap (n_2 \cdot m). \end{aligned}$$

Wegen der Kommutativität der Multiplikation können wir uns natürlich auf den Nachweis z. B. der linksseitigen Distributivität beschränken, und da (23) in den Fällen $n_1 = 0$ und $n_2 = 0$ nach (21) trivial gilt, können wir beim nachfolgenden Beweis $n_1 \neq 0$ und $n_2 \neq 0$ voraussetzen. Wir weisen nach, daß die Zahl $d = m \cdot (n_1 \sqcap n_2)$ die charakteristischen Eigenschaften des größten gemeinsamen Teilers von $m \cdot n_1$ und $m \cdot n_2$ besitzt. Wegen $n_1 \sqcap n_2 \mid n_1$ und $n_1 \sqcap n_2 \mid n_2$ gilt nach (10) $m \cdot (n_1 \sqcap n_2) \mid m \cdot n_1$ und $m \cdot (n_1 \sqcap n_2) \mid m \cdot n_2$, d. h., d ist ein gemeinsamer Teiler von $m \cdot n_1$ und $m \cdot n_2$. Es bleibt also zu zeigen, daß jeder gemeinsame Teiler t der Zahlen $m \cdot n_1$ und $m \cdot n_2$ auch ein Teiler von d ist. Dazu beachten wir, daß nach (15) natürliche Zahlen a, b mit

$$n_1 \sqcap n_2 = a n_1 - b n_2$$

existieren. Dann wird aber

$$d = m \cdot (n_1 \sqcap n_2) = amn_1 - bmn_2,$$

woraus nach (10″) und (9) die genannte Behauptung folgt.

Der auf einer geschickten Anwendung des Prinzips der kleinsten Zahl beruhende Gaußsche Beweis für die Existenz des größten gemeinsamen Teilers hat den Nachteil, daß er zunächst lediglich ein sogenannter reiner Existenzbeweis ist, d. h. kein Berechnungsverfahren für $m \sqcap n$ liefert; denn es ist in keiner Weise zu erkennen, wie man bei vorgegebenem $m, n \in \mathbf{N}$ die Koeffizienten a, b bzw. $\bar{a}, \bar{b}$ in der Darstellung (15) von $m \sqcap n$ ermitteln kann. Aus dem Beweis von (15) ist lediglich zu entnehmen, wie man aus den Koeffizienten a, b die Koeffizienten $\bar{a}, \bar{b}$ (und analog aus $\bar{a}, \bar{b}$ auch a, b) berechnen kann:

$$(24) \qquad \bar{a} = am - b, \quad \bar{b} = an - a; \qquad a = \bar{a}n - \bar{b}, \quad b = \bar{a}m - \bar{a}.$$

Ein wichtiges Berechnungsverfahren für den größten gemeinsamen Teiler findet sich bereits im siebenten Buch der „Elemente" des Euklid (etwa 365 bis 300 v. u. Z.), dem großartigen Sammelwerk der mathematischen Kenntnisse der griechischen Antike (der Name „Elementarmathematik" bedeutete übrigens ursprünglich „Mathematik gemäß den Elementen des Euklid", die in überarbeiteter Form noch bis in das vergangene Jahrhundert die Grundlage der Schulmathematik bildeten). Es wird demgemäß heute *euklidischer Algorithmus* genannt. Dieser Algorithmus besteht in einer merkwürdigen sukzessiven Anwendung der in (11) behandelten Division mit Rest. Es seien dabei m, n beliebige von Null verschiedene natürliche Zahlen, wobei wir o. B. d. A. wegen (16) $m \geqq n$ voraussetzen können (was aber nur der Abkürzung des Verfahrens dient). Der Systematik halber sei $m = r_0$ und $n = r_1$ gesetzt. Im ersten Schritt dividieren wir $r_0 (= m)$ durch $r_1 (= n)$ mit Rest, wobei wir $q(r_0, r_1) = q_1$ und $r(r_0, r_1) = r_2$ setzen:

$$(25_0) \qquad r_0 = q_1 r_1 + r_2 \quad \text{mit} \quad r_2 < r_1.$$

Ist hierbei $r_2 = 0$, so ist das Verfahren beendet (und $r_0 \sqcap r_1 = r_1$). Ist dagegen $r_2 \neq 0$, so wird in analoger Weise r_1 durch r_2 mit Rest dividiert, so daß bei $q(r_1, r_2) = q_2$, $r(r_1, r_2) = r_3$

$$(25_1) \qquad r_1 = q_2 r_2 + r_3 \quad \text{mit} \quad r_3 < r_2$$

wird. Ist hierbei $r_3 = 0$, so ist das Verfahren beendet. Ist dagegen auch noch $r_3 \neq 0$, so wird

$$(25_2) \qquad r_2 = q_3 r_3 + r_4 \quad \text{mit} \quad r_4 < r_3$$

gesetzt und allgemein das Verfahren mit

$$(25_i) \qquad r_i = q_{i+1} r_{i+1} + r_{i+2} \quad \text{mit} \quad r_{i+2} < r_{i+1}$$

solange weitergeführt, bis erstmalig $r_{k+2} = 0$ wird:

$$
\begin{aligned}
r_0 &= q_1 r_1 + r_2 \quad \text{mit} \quad 0 < r_2 < r_1, \\
r_1 &= q_2 r_2 + r_3 \quad \text{mit} \quad 0 < r_3 < r_2, \\
&\cdot \\
&\cdot \\
&\cdot \\
r_{k-1} &= q_k r_k + r_{k+1} \text{ mit } 0 < r_{k+1} < r_k, \\
r_k &= q_{k+1} r_{k+1}.
\end{aligned} \tag{25}
$$

Da die Reste $r_2, r_3, \ldots$ eine echt monoton fallende Folge von natürlichen Zahlen bilden, erfolgt nach endlich vielen Schritten ein Abbruch. Wir behaupten nun, daß *der letzte nicht verschwindende Rest r_{k+1} der größte gemeinsame Teiler der Zahlen r_0 (= m) und r_1 (= n) ist.* Das ist eine unmittelbare Folgerung aus:

$$n < m \wedge m = qn + r \Rightarrow m \sqcap n = n \sqcap r. \tag{26}$$

Mittels (26) erhalten wir nämlich aus (25):

$$
\begin{aligned}
r_0 \sqcap r_1 = r_1 \sqcap r_2 = r_2 \sqcap r_3 = \cdots = r_{k-1} \sqcap r_k \\
= r_k \sqcap r_{k+1} = r_{k+1} \sqcap 0 = r_{k+1}.
\end{aligned}
$$

Zum Beweis von (26) genügt es jedoch, folgendes zu bemerken:

$$t \mid m \wedge t \mid n \Rightarrow t \mid n \wedge t \mid r \Rightarrow t \mid n \sqcap r,$$

so daß wegen $m \sqcap n \mid m$ und $m \sqcap n \mid n$ speziell $m \sqcap n \mid n \sqcap r$; umgekehrt gilt

$$t \mid n \wedge t \mid r \Rightarrow t \mid m \wedge t \mid n \Rightarrow t \mid m \sqcap n,$$

so daß wegen $n \sqcap r \mid n$ und $n \sqcap r \mid r$ auch $n \sqcap r \mid m \sqcap n$. Also ist

$$m \sqcap n = n \sqcap r,$$

wie behauptet.

Mittels (25) läßt sich auch leicht die in (15) festgestellte Darstellung des größten gemeinsamen Teilers berechnen. Nach (25_0) wird nämlich zunächst

$$r_2 = m - q_1 n.$$

Setzen wir dies in (25_1) ein, so erhalten wir

$$r_3 = n - q_2(m - q_1 n) = (q_1 q_2 + 1) n - q_2 m,$$

und (25_2) liefert

$$r_4 = (q_2 q_3 + 1) m - (q_1 q_2 q_3 + q_1 + q_3) n$$

usw., bis schließlich nach (25_{k-1}) eine Darstellung von r_{k+1} (= $m \sqcap n$) in der Form $am - bn$ bzw. $\bar{a}n - \bar{b}m$ gewonnen wird.

Das folgende Beispiel ($m = 133$, $n = 91$) möge das Verfahren erläutern:

$$\begin{aligned} 133 &= 1 \cdot 91 + 42, \\ 91 &= 2 \cdot 42 + 7, \\ 42 &= 6 \cdot 7. \end{aligned} \tag{*}$$

Also gilt $133 \sqcap 91 = 7$. Ferner ergibt sich aus (*)

$$42 = 133 - 91,$$

$$7 = 91 - 2 \cdot 42 = 91 - 2 \cdot (133 - 91) = 3 \cdot 91 - 2 \cdot 133,$$

und (24) liefert

$$7 = (3 \cdot 91 - 2) \cdot 133 - (3 \cdot 133 - 3) \cdot 91 = 271 \cdot 133 - 396 \cdot 91.$$

Dieses Beispiel lehrt bereits, daß die Koeffizienten $a, b, \bar{a}, \bar{b}$ in der Darstellung (15) sehr groß werden können.

Natürliche Zahlen m, n heißen *teilerfremd* oder *relativ prim*, wenn ihr größter gemeinsamer Teiler die Zahl 1 ist:

(27) m *teilerfremd zu* $n :\Leftrightarrow m \sqcap n = 1$.

Wir merken an, daß die Teilerfremdheit das formale Analogon zur Disjunktheit von Mengen (vgl. 1.4.(20)) ist, da ja die Zahl 1 hinsichtlich der Teilbarkeitsrelation dieselbe Rolle wie die leere Menge bzgl. der Inklusion spielt. Nach Satz (15) *sind von Null verschiedene Zahlen* m, n *genau dann teilerfremd, wenn es natürliche Zahlen* a, b ($\bar{a}, \bar{b}$) *gibt, so daß* $am - bn = 1$ ($\bar{a}n - \bar{b}m = 1$):

$$(28) \qquad m \neq 0 \wedge n \neq 0 \Rightarrow (m \sqcap n = 1 \Leftrightarrow \bigvee_{a,b \in \mathbb{N}} am - bn = 1).$$

Damit erhalten wir leicht den folgenden Satz:

$$(29) \qquad \begin{aligned} &m \neq 0 \wedge n \neq 0 \wedge m = (m \sqcap n) \cdot m_1 \wedge n = (m \sqcap n) \cdot n_1 \\ &\Rightarrow m_1 \text{ teilerfremd zu } n_1. \end{aligned}$$

Denn nach (15) gibt es natürliche Zahlen a, b mit $m \sqcap n = am - bn = (m \sqcap n)\, am_1 - (m \sqcap n)\, bn_1 = (m \sqcap n)(am_1 - bn_1)$, und hieraus folgt nach 3.4.(23) $am_1 - bn_1 = 1$, also sind nach (28) die Zahlen m_1, n_1 teilerfremd, was zu zeigen war.

Mittels (28) können wir ferner den folgenden wichtigen Satz beweisen:

$$(30) \qquad k \mid m \cdot n \wedge k \sqcap m = 1 \Rightarrow k \mid n;$$

in Worten: *Teilt eine Zahl* k *ein Produkt* $m \cdot n$ *und ist sie zu einem der Faktoren (beispielsweise zu* m*) teilerfremd, so teilt sie den anderen Faktor.* Ist eine der Zahlen k oder m gleich 0, so gilt (30) trivial, so daß wir beim folgenden Beweis $k \neq 0$, $m \neq 0$ voraussetzen können. Dann gibt es nach (28) natürliche Zahlen a, b mit $ak - bm = 1$, und Multiplikation beider Seiten dieser Gleichung mit n

ergibt $akn - bmn = n$. Hierbei ist nun $k \mid akn$, und wegen $k \mid mn$ gilt nach (10″) $k \mid bmn$, also nach (9) $k \mid akn - bmn$, d. h. $k \mid n$, was zu zeigen war.

Mit Hilfe von (30) können wir schließlich den folgenden Satz beweisen:

(31) $m_1 \mid n \wedge m_2 \mid n \wedge m_1 \sqcap m_2 = 1 \Rightarrow m_1 \cdot m_2 \mid n$.

Ist nämlich $n = m_1 q_1$, so gilt nach (30) wegen $m_2 \mid q_1 m_1$ und $m_1 \sqcap m_2 = 1$ offenbar $m_2 \mid q_1$, d. h., es gibt eine Zahl $q \in \mathsf{N}$ mit $m_2 q = q_1$, und dann wird $n = m_1 q_1 = m_1 m_2 q$, und das besagt ja gerade, daß $m_1 \cdot m_2$ ein Teiler von n ist.

Das duale Gegenstück zum größten gemeinsamen Teiler und formale Analogon zur Vereinigungsmenge ist das sogenannte *kleinste gemeinsame Vielfache* zweier natürlicher Zahlen m, n, das durch die Bedingungen

(32a) $m \mid v, \quad n \mid v,$

(32b) $\bigwedge_{s \in \mathsf{N}} (m \mid s \wedge n \mid s \Rightarrow v \mid s)$

charakterisiert wird (vgl. 1.5.(10), (11)). Zunächst ist natürlich vor allem wieder zu zeigen, daß *es bei beliebigem $m, n \in \mathsf{N}$ genau eine natürliche Zahl v gibt, die die Bedingungen* (32) *erfüllt*. Diese durch m und n eindeutig bestimmte Zahl v wird dann das kleinste gemeinsame Vielfache von m und n genannt und im folgenden mit $m \sqcup n$ bezeichnet (in der Literatur sind hierfür auch die Bezeichnungen $[m, n]$ und $\mathrm{kgV}(m, n)$ üblich).

Der Nachweis für die Einzigkeit von v erfolgt analog dem Beweis der Einzigkeit einer Zahl d mit den Eigenschaften (14) und sei dem Leser als Übungsaufgabe überlassen. Zum Existenzbeweis merken wir zunächst an, daß im Fall $m = 0$ und im Fall $n = 0$ die Zahl $v = 0$ die Eigenschaften (32) besitzt, so daß wir uns auf die Betrachtung des Falles $m \neq 0$, $n \neq 0$ beschränken können. Wir wollen zeigen, daß in diesem Fall die Zahl

$$v = \frac{m \cdot n}{m \sqcap n}$$

das Verlangte leistet (wegen $m \sqcap n \mid m \cdot n$ und $m \sqcap n \neq 0$ ist v eine natürliche Zahl). Zum Beweis dieser Behauptung sei $m = (m \sqcap n)\, m_1$ und $n = (m \sqcap n)\, n_1$. Dann wird offenbar

$$v = (m \sqcap n)\, m_1 n_1,$$

woraus zunächst unmittelbar (32a) abgelesen werden kann. Zum Nachweis von (32b) sei s ein beliebiges gemeinsames Vielfaches von m und n, etwa $s = mq_1 = nq_2$. Dann ist s insbesondere ein Vielfaches von $m \sqcap n$, etwa $s = (m \sqcap n)\, q$, und es gilt

$$(m \sqcap n)\, q = mq_1 = (m \sqcap n)\, m_1 q_1,$$
$$(m \sqcap n)\, q = nq_2 = (m \sqcap n)\, n_1 q_2,$$

woraus nach 3.4.(22) (wegen $m \sqcap n \neq 0$) $q = m_1 q_1 = n_1 q_2$ folgt. Also ist $m_1 \mid q$ und $n_1 \mid q$, wobei nach (29) $m_1 \sqcap n_1 = 1$ ist. Mithin gilt nach (31) auch $m_1 n_1 \mid q$, d. h., es existiert eine Zahl $q' \in \mathbf{N}$ mit $m_1 n_1 q' = q$. Dann wird aber $s = (m \sqcap n) q = (m \sqcap n) m_1 n_1 q' = v q'$, so daß in der Tat $v \mid s$, was noch zu zeigen war.

Damit haben wir zugleich den folgenden wichtigen Satz bewiesen:

$$(33) \qquad m \neq 0 \wedge n \neq 0 \Rightarrow m \sqcup n = \frac{m \cdot n}{m \sqcap n}.$$

Hieraus folgt unmittelbar, daß *bei teilerfremden Zahlen m, n das kleinste gemeinsame Vielfache gleich dem Produkt ist:*

$$(34) \qquad m \sqcap n = 1 \Rightarrow m \sqcup n = m \cdot n,$$

wobei im Fall $m \neq 0$, $n \neq 0$ offenbar auch die Umkehrung gilt.

Aus der Definition (32) ergeben sich mühelos die folgenden Eigenschaften des kleinsten gemeinsamen Vielfachen (vgl. dazu die Sätze 1.4.(6), 1.4.(9), 1.4.(14) und 1.5.(17) über die Vereinigung):

$$(35) \qquad m \sqcup n = n \sqcup m,$$

$$(36) \qquad n_1 \sqcup (n_2 \sqcup n_3) = (n_1 \sqcup n_2) \sqcup n_3,$$

$$(37) \qquad n \sqcup n = n,$$

$$(38) \qquad m \mid n \Leftrightarrow m \sqcup n = n.$$

Die Beweise hierfür verlaufen analog den Beweisen der Sätze (16) bis (19) und seien dem Leser als Übungsaufgaben überlassen. Auf Grund von (36) können wir, ohne Mißverständnisse befürchten zu müssen, in drei- und mehrgliedrigen $\sqcup$-Termen die Klammern fortlassen, also $n_1 \sqcup \cdots \sqcup n_k$ für einen beliebig geklammerten $\sqcup$-Term aus den Zahlen $n_1, \ldots, n_k$ (in dieser Reihenfolge), z. B. den kanonisch geklammerten Term $(\cdots((n_1 \sqcup n_2) \sqcup n_3) \sqcup \cdots) \sqcup n_k$ schreiben. *Die Zahl $v = n_1 \sqcup \cdots \sqcup n_k$ wird* dann *bei beliebigem k durch die Eigenschaften*

$$(39a) \qquad n_1 \mid v, \ldots, n_k \mid v,$$

$$(39b) \qquad \bigwedge_{s \in \mathbf{N}} (n_1 \mid s \wedge \cdots \wedge n_k \mid s \Rightarrow v \mid s)$$

charakterisiert und daher das *kleinste gemeinsame Vielfache der Zahlen $n_1, \ldots, n_k$* genannt. Man zeigt leicht, daß in Verallgemeinerung von (33) folgendes gilt:

$$(40) \qquad n_1 \neq 0 \wedge \cdots \wedge n_k \neq 0 \Rightarrow n_1 \sqcup \cdots \sqcup n_k = \frac{n_1 \cdots n_k}{m_1 \sqcap \cdots \sqcap m_k},$$

wobei

$$m_i = n_1 \cdots n_{i-1} \cdot n_{i+1} \cdots n_k \quad (i = 1, \ldots, k).$$

Auf Grund von (35) und (36) kommt es in $n_1 \sqcup \cdots \sqcup n_k$ ebenfalls nicht auf die Reihenfolge der Glieder an, d. h., *für jede Permutation π der Indizes* $1, \ldots, k$ *gilt*

(35′) $\quad n_{\pi(1)} \sqcup \cdots \sqcup n_{\pi(k)} = n_1 \sqcup \cdots \sqcup n_k.$

Aus (38) und (5) entnehmen wir noch leicht (wie bereits im Beweis von (33) bemerkt), daß *bei beliebigem* $n \in \mathbb{N}$

(41) $\quad n \sqcup 0 = 0 \sqcup n = 0$

ist, und entsprechend folgt aus (38) und (6), daß *bei beliebigem* $n \in \mathbb{N}$ (in Analogie zu 1.4.(19))

(42) $\quad n \sqcup 1 = 1 \sqcup n = n.$

In Analogie zu (23) gilt ferner

(43) $$\begin{aligned} m \cdot (n_1 \sqcup n_2) &= (m \cdot n_1) \sqcup (m \cdot n_2), \\ (n_1 \sqcup n_2) \cdot m &= (n_1 \cdot m) \sqcup (n_2 \cdot m)\,. \end{aligned}$$

Es entsteht nun naturgemäß die Frage, ob in Analogie zu Durchschnitt und Vereinigung (vgl. 1.4.(11) und 1.4.(12)) die Distributivgesetze

(44) $\quad (n_1 \sqcup n_2) \sqcap n_3 = (n_1 \sqcap n_3) \sqcup (n_2 \sqcap n_3),$

(45) $\quad (n_1 \sqcap n_2) \sqcup n_3 = (n_1 \sqcup n_3) \sqcap (n_2 \sqcup n_3)$

und die Verschmelzungssätze (vgl. 1.4.(15) und 1.4.(16))

(46) $\quad (n_1 \sqcup n_2) \sqcap n_1 = n_1,$

(47) $\quad (n_1 \sqcap n_2) \sqcup n_1 = n_1$

gelten. Wir merken ohne Beweis (Übungsaufgaben!) an, daß *das tatsächlich der Fall ist*.

Ein weiterer wichtiger Begriff der Teilbarkeitstheorie ist der Begriff der *Primzahl*. Hierunter wird bekanntlich eine natürliche Zahl $p > 1$ verstanden, die nur durch 1 und sich selbst teilbar ist:

(48) $\quad p$ *Primzahl* $: \Leftrightarrow p \in \mathbb{N} \wedge p > 1 \wedge \bigwedge_{t \in \mathbb{N}} (t \mid p \Rightarrow t = 1 \vee t \doteq p).$

Wir machen ausdrücklich darauf aufmerksam, daß die Zahl 1, die ja auch nur durch 1 und sich selbst teilbar ist, nicht zu den Primzahlen gerechnet wird. Im Prinzip ist das natürlich eine Konvention, die allerdings einen bestimmten Grund hat, auf den wir etwas später eingehen werden. Die Primzahlen sind mithin dadurch charakterisiert, daß *sie genau zwei Teiler besitzen* (die Zahl 1 besitzt nur einen Teiler). Bezüglich der teilweisen Ordnung $\mid$ sind die Primzahlen gerade die oberen Nachbarn der Zahl 1, die ja nach (6) das bezüglich

der Teilbarkeitsrelation kleinste Element ist. Sie spielen also in dieser Hinsicht (vgl. S. 36) dieselbe Rolle wie die Einermengen bei der Inklusion.

Als erstes wollen wir zeigen, daß *sich jede natürliche Zahl $n > 1$ auf wenigstens eine Weise als Produkt von Primzahlen darstellen läßt*, wobei dieses Produkt allerdings auch in einen einzelnen Faktor ausarten kann. Lassen wir auch das *leere Produkt* zu und schreiben wir diesem den Wert 1 bei, so gilt der behauptete Satz auch im Fall $n = 1$:

$$(49) \qquad n \geqq 1 \Rightarrow \bigvee_{k,p_1,\ldots,p_k} (k \in \mathbf{N} \wedge p_1 \text{ Primzahl} \wedge \cdots \wedge p_k \text{ Primzahl} \wedge n = \prod_{\varkappa=1}^{k} p_\varkappa).$$

Wir beweisen (49) indirekt, nehmen also an, es gäbe eine natürliche Zahl $n \geqq 1$, die sich nicht als Produkt von Primzahlen darstellen läßt. Dann ist die Menge aller der Zahlen $n \geqq 1$, die sich nicht als Produkt von Primzahlen darstellen lassen, nicht leer, enthält also nach 3.4.(31) eine kleinste Zahl n_0. Dabei muß $n_0 > 1$ sein, da verabredungsgemäß die Zahl 1 durch das leere Produkt dargestellt wird. Es kann auch n_0 keine Primzahl sein, da sich in diesem Fall n_0 als Produkt aus einem Faktor darstellen läßt. Folglich besitzt n_0 einen Teiler t, der von 1 und n_0 verschieden ist, d. h. $n_0 = t \cdot q$, wobei nach (7) offenbar $1 < t < n_0$ und $1 < q < n_0$ gilt. Also lassen sich t und q als Produkte von Primzahlen darstellen (denn n_0 sollte die kleinste Zahl $\geqq 1$ sein, die keine derartige Darstellung besitzt), etwa $t = \prod_{\varkappa=1}^{k} p_\varkappa$, $q = \prod_{\lambda=1}^{l} p'_\lambda$, und dann wäre

$$n_0 = t \cdot q = \prod_{\varkappa=1}^{k} p_\varkappa \cdot \prod_{\lambda=1}^{l} p'_\lambda$$

eine Darstellung von n_0 als Produkt von Primzahlen, die es ja nach Annahme nicht geben sollte. Folglich ist unsere Annahme falsch, und es gilt (49).

Wir wollen im folgenden unter einer *Primzahlzerlegung der Zahl n* ein k-Tupel $(k \geqq 0)$ $(p_1, \ldots, p_k)$ von Primzahlen $p_1, \ldots, p_k$ verstehen, für das $p_1 \leqq p_2 \leqq \cdots \leqq p_k$ und $\prod_{\varkappa=1}^{k} p_\varkappa = n$ gilt, wobei wir die Zahl k auch die *Länge* dieser Primzahlzerlegung nennen wollen. Nach dem allgemeinen Assoziativ-Kommutativgesetz 3.5. (14) ist klar, daß *jede natürliche Zahl n, die sich überhaupt als Produkt von Primzahlen darstellen läßt* – und nach (49) gilt das für jede natürliche Zahl $n \geqq 1$ – *auch eine Primzahlzerlegung* im hier betrachteten Sinne *besitzt*.

Wir wollen nun zeigen, daß schärfer folgendes gilt:

(50) *Jede natürliche Zahl $n \geqq 1$ besitzt genau eine Primzahlzerlegung.*

Man nennt (50) Hauptsatz über die eindeutige Primzahlzerlegung. Unsere Einschränkung auf Primzahlprodukte $\prod_{\varkappa=1}^{k} p_\varkappa$ mit $p_1 \leqq p_2 \leqq \cdots \leqq p_k$ stellt sicher, daß es im strengen Sinne genau eine Zerlegung gibt, während andernfalls die Zerlegung nur bis auf die Reihenfolge der Faktoren eindeutig ist. Für die Gültigkeit von (50) ist offenbar wesentlich, daß wir die Zahl 1 nicht zu den Primzahlen gerechnet haben, denn andernfalls wäre noch nicht einmal die Länge der Primzahlzerlegung eindeutig bestimmt.

Die Existenz einer Primzahlzerlegung ist – wie bereits bemerkt – durch (4) bewiesen. Für die Einzigkeit wollen wir aus einem später ersichtlichen Grund zwei verschiedene Beweise geben. Der erste Beweis stützt sich auf den unmittelbar aus (30) folgenden und schon bei EUKLID vorhandenen Hilfssatz

(51) $\quad p$ *Primzahl* $\wedge\, p \mid m \cdot n \Rightarrow p \mid m \vee p \mid n$;

in Worten: *Teilt eine Primzahl p ein Produkt $m \cdot n$, so teilt sie wenigstens einen der Faktoren m oder n.* Zum Beweis nehmen wir an, daß die Primzahl p das Produkt $m \cdot n$, aber nicht z. B. den Faktor m teilt, und zeigen, daß dann p notwendig den anderen Faktor n teilen muß. Ist nämlich p kein Teiler von m, so ist offenbar nach (19) $m \sqcap p = 1$ und mithin nach (30) p ein Teiler von n, was zu zeigen war.

Durch vollständige Induktion über k (Übungsaufgabe!) kann man (51) leicht verallgemeinern zu

(51′) $\quad p$ *Primzahl* $\wedge\, p \mid \prod_{\varkappa=1}^{k} n_\varkappa \Rightarrow p \mid n_1 \vee \cdots \vee p \mid n_k$.

Hier zunächst noch die genaue Formulierung des Einzigkeitssatzes:

(50′) $\quad n \geqq 1 \wedge (p_1, \ldots, p_k)$ *Primzahlzerlegung für n*

$\qquad \wedge\, (p'_1, \ldots, p'_l)$ *Primzahlzerlegung für n*

$\qquad \Rightarrow k = l \wedge p_1 = p'_1 \wedge \cdots \wedge p_k = p'_k$.

Der erste Beweis für (50′) erfolgt durch ordnungstheoretische Induktion über n (vgl. 3.4.(32)). Der Anfangsschritt $n = 1$ ist trivial: Da jede nichtleere Primzahlzerlegung $\prod_{\varkappa=1}^{k} p_\varkappa$ $(k \geqq 1)$ größer als 1 ist, besitzt die Zahl 1 nur die leere Zerlegung. Wir nehmen nun an, daß für alle Zahlen m mit $1 \leqq m < n$ die Einzigkeit der Primzahlzerlegung schon bewiesen ist, und zeigen, daß sie dann auch für die Zahl n gilt. Dazu seien $(p_1, \ldots, p_k)$ und $(p'_1, \ldots, p'_l)$ beliebige

Primzahlzerlegungen für die Zahl n (> 1), wobei wir o. B. d. A. annehmen können, daß $p_1 \leqq p_1'$ ist (andernfalls würden wir lediglich die Rollen von $(p_1, \ldots, p_k)$ und $(p_1', \ldots, p_l')$ vertauschen). Wegen $\prod_{\varkappa=1}^{k} p_\varkappa = \prod_{\lambda=1}^{l} p_\lambda'$ ist dann $p_1 \mid \prod_{\lambda=1}^{l} p_\lambda'$, so daß nach (51′) die Zahl p_1 wenigstens einen der Faktoren $p_1', \ldots, p_l'$ teilen und mithin (da $p_1', \ldots, p_l'$ Primzahlen sind) mit diesen übereinstimmen muß. Wegen $p_1 \leqq p_1'$ muß das der Faktor p_1' sein; denn wäre $p_1 \neq p_1'$, so wäre $p_1 < p_1'$, und wegen $p_\lambda' \geqq p_1'$ ($\lambda = 2, \ldots, l$) könnte dann p_1 mit keinem der Faktoren $p_1', \ldots, p_l'$ übereinstimmen. Folglich ist nach 3.4.(23) $\prod_{\varkappa=2}^{k} p_\varkappa = \prod_{\lambda=2}^{l} p_\lambda'$, d. h., die Zahl $m = \prod_{\varkappa=2}^{k} p_\varkappa < n$ hat die Primzahlzerlegungen $(p_2, \ldots, p_k)$ und $(p_2', \ldots, p_l')$, und nach Induktionsvoraussetzung muß dann $k = l$ und $p_2 = p_2', \ldots, p_k = p_k'$ sein, was zusammen mit der schon bewiesenen Gleichung $p_1 = p_1'$ die Behauptung von (50′) für die Zahl n ergibt.

Obwohl bereits bei EUKLID der Satz (51) mit allen dazu notwendigen Vorbereitungen zu finden ist, wurde von ihm der relativ kleine Schritt zum Hauptsatz (50) nicht mehr vollzogen. Wir wissen nicht, ob den griechischen Mathematikern dieser Satz bekannt war, und wenn, warum ihn EUKLID in seinem sonst so systematischen Werk nicht formuliert hat. Sicher ist jedoch, daß EUKLID ihn nicht stillschweigend als eine evidente Tatsache angesehen und benutzt hat.

Der nachfolgende zweite Beweis von (50′) ist sehr neuen Datums und geht auf E. ZERMELO zurück. Er benutzt nicht den Satz (51) und damit auch nicht die zum Beweis von (51) erforderlichen Grundtatsachen über den größten gemeinsamen Teiler. Er ermöglicht dadurch einen wesentlich anderen Aufbau der Anfangsgründe der Teilbarkeitstheorie, wie er heute auch in der Schule vorgenommen wird, obwohl dabei manche Zusammenhänge unerkannt bleiben. Der Zermelosche Beweis von (50′) erfolgt analog wie der Beweis des Satzes (49) indirekt. Es wird also angenommen, daß es eine Zahl $n \geqq 1$ gibt, die zwei verschiedene Primzahlzerlegungen besitzt. Dann gibt es auch eine kleinste derartige Zahl, und diese sei mit n bezeichnet, wobei $n = \prod_{\varkappa=1}^{k} p_\varkappa$ und $n = \prod_{\lambda=1}^{l} p_\lambda'$ zwei verschiedene Primzahlzerlegungen der Zahl n seien. Dann muß offenbar $p_1 \neq p_1'$ sein; denn andernfalls wäre $\prod_{\varkappa=2}^{k} p_\varkappa = \prod_{\lambda=2}^{l} p_\lambda'$, und wegen $\prod_{\varkappa=2}^{k} p_\varkappa < n$ müßte $k = l$ und $p_2 = p_2', \ldots, p_k = p_k'$ sein (da ja n die kleinste

Zahl mit mehreren Primzahlzerlegungen sein sollte), was zusammen mit $p_1 = p_1'$ ein Widerspruch zu der Voraussetzung ist, daß $(p_1, \ldots, p_k)$ und $(p_1', \ldots, p_l')$ verschiedene Zerlegungen der Zahl n sind. Ohne Beschränkung der Allgemeinheit können wir annehmen, daß $p_1 < p_1'$ ist. Wir betrachten dann die Zahl

$$m = n - p_1 \cdot p_2' \cdots p_l'.$$

Wegen $p_1 < p_1'$ ist $p_1 \cdot p_2' \cdots p_l' < p_1' \cdot p_2' \cdots p_l'$ und mithin $m \in \mathsf{N}$ und außerdem $m < n$. Offenbar ist

(i) $$m = p_1(p_2 \cdots p_k - p_2' \cdots p_l')$$

mit $p_2 \cdots p_k - p_2' \cdots p_l' < m < n$, so daß $p_2 \cdots p_k - p_2' \cdots p_l'$ genau eine Primzahlzerlegung $(p_1'', \ldots, p_r'')$ besitzt, aus der nach (i) die eindeutig bestimmte Primzahlzerlegung $(p_1^*, \ldots, p_{r+1}^*)$ von m dadurch gewonnen wird, daß man die Primzahl p_1 an der richtigen Stelle einordnet. Andererseits ist

(ii) $$m = (p_1' - p_1)\, p_2' \cdots p_l'$$

mit $p_1' - p_1 \leqq m < n$, so daß auch $p_1' - p_1$ genau eine Primzahlzerlegung $(p_1''', \ldots, p_s''')$ besitzt, aus der sich nach (ii) die eindeutig bestimmte Primzahlzerlegung $(p_1^*, \ldots, p_{r+1}^*)$ von m dadurch gewinnen läßt, daß man die Primzahlen $p_2', \ldots, p_l'$ an den richtigen Stellen einordnet. Da nun p_1 unter den Primzahlen $p_1^*, \ldots, p_{r+1}^*$ vorkommt, muß sie auch unter den Primzahlen $p_1''', \ldots, p_s''', p_2', \ldots, p_l'$ vorhanden sein. Nun ist aber wegen

$$p_1 < p_1' \leqq p_2' \leqq \cdots \leqq p_l'$$

die Primzahl p_1 von allen Primzahlen $p_2', \ldots, p_l'$ verschieden. Also muß p_1 unter den Primzahlen $p_1''', \ldots, p_s'''$ vorkommen, und hieraus folgt wegen $p_1' - p_1 = \prod_{\sigma=1}^{s} p_\sigma'''$, daß p_1 ein Teiler von $p_1' - p_1$ ist. Dann ist aber nach (8) p_1 auch Teiler von $(p_1' - p_1) + p_1$, d. h. Teiler von p_1', so daß $p_1 = p_1'$ sein muß, was wir bereits als unmöglich erkannt haben. Also ist unsere Annahme, daß es eine natürliche Zahl $n \geqq 1$ mit zwei verschiedenen Primzahlzerlegungen gibt, falsch und (50′) bewiesen.

Man wird zugeben müssen, daß der Zermelosche Beweis zwar recht originell, aber keineswegs ganz einfach ist.

Faßt man die in der eindeutigen Primzahlzerlegung einer natürlichen Zahl $n \geqq 1$ evtl. mehrfach auftretenden Primzahlen im zugehörigen Produkt $\prod_{\varkappa=1}^{k} p_\varkappa$ zu Potenzen zusammen, so erhält man: *Jede natürliche Zahl* $n \geqq 1$

läßt sich auf genau eine Weise in der Form

$$(52) \qquad n = \prod_{\lambda=1}^{l} p_\lambda^{\nu_\lambda}$$

mit paarweise verschiedenen Primzahlen $p_1 < p_2 < \cdots < p_l$ *und von Null verschiedenen natürlichen Exponenten* $\nu_1, \ldots, \nu_l$ *darstellen.* Die Darstellung (52) wollen wir auch die *Primzahlpotenzdarstellung von n* nennen.

Bevor wir diesen Satz noch etwas weiter ausbauen, wollen wir einen anderen wichtigen Satz beweisen, der gleichfalls auf EUKLID zurückgeht:

(53) *Die Menge P aller Primzahlen ist unendlich.*

Nach 2.9.(2) haben wir hierfür zu zeigen, daß die Menge P keinem Abschnitt $\mathcal{A}(n)$ ($n \in \mathbb{N}$) gleichmächtig ist. Das beweisen wir wiederum indirekt. Wir nehmen also an, es gäbe eine natürliche Zahl n mit $P \sim \mathcal{A}(n)$, und es sei f eine 1-1-Abbildung von $\mathcal{A}(n)$ auf P, so daß also P aus den genau n Primzahlen

$$p_0 = f(0), \quad \ldots, \quad p_{n-1} = f(n-1)$$

besteht. Wir betrachten dann die folgende natürliche Zahl m

$$m = p_0 \cdots p_{n-1} + 1.$$

Da offenbar $m \geqq 2$ ist, läßt sich m nach (49) als nichtleeres Produkt von Primzahlen darstellen. Ist dann p ein beliebiger Faktor aus diesem Produkt, so ist $p \mid m$. Daraus folgt, daß die Primzahl p von den Primzahlen $p_0, \ldots, p_{n-1}$ verschieden ist; denn der Rest $r(m, p_\nu)$ bei der Division von m durch p_ν ist gleich 1, und mithin kann nach (13) p_ν kein Teiler von m sein. Also ist $p \notin \{p_0, \ldots, p_{n-1}\}$, entgegen unserer Annahme, daß $\{p_0, \ldots, p_{n-1}\}$ sämtliche Primzahlen enthält. Dieser Widerspruch widerlegt unsere Annahme und beweist (53).

Es sei nun wieder n eine natürliche Zahl $\geqq 1$ und p eine beliebige Primzahl. Mit $\exp_p(n)$ (gelesen: Exponent von p in n) bezeichnen wir den Exponenten ν, mit dem p in die eindeutige Primzahlpotenzdarstellung (52) von n eingeht, wobei wir $\exp_p(n) = 0$ setzen, wenn die Primzahl p in dieser Darstellung nicht als Faktor auftritt. Man erkennt leicht (Beweis!), daß $\exp_p(n)$ *die größte natürliche Zahl* ν^* *ist, so daß* p^{ν^*} *Teiler von n ist.* Eine solche größte Zahl existiert nach 3.4.(34), da die Menge aller Zahlen ν mit $p^\nu \mid n$ nach oben beschränkt ist: Ist nämlich $p^\nu \mid n$, so ist $\nu < n$; denn wegen $p \geqq 2$ ist nach 3.5.(6) $p^n \geqq 2^n$, und es ist andererseits – wie man leicht durch vollständige Induktion über n bestätigt (Beweis!) – stets $2^n > n$, und aus $p^n > n$ folgt nach 3.5.(9), daß erst recht für alle $m > n$ die Ungleichung $p^m > n$ gilt; folglich ist im Fall $n \neq 0$

nach (7) keine Zahl p^m mit $m > n$ ein Teiler von n. Es ist also

(54) $\quad \exp_p(n) := \max\{\nu : \nu \in \mathsf{N} \wedge p^\nu \mid n\}.$

Insbesondere gilt also

(55) $\quad \exp_p(n) \neq 0 \Leftrightarrow p \mid n.$

Ferner gilt nach (52)

(56) $\quad n \geqq 1 \Rightarrow n = \prod_{p \mid n} p^{\exp_p(n)},$

wobei das Produkt auf der rechten Seite entsprechend der „Laufanweisung" unter dem Produktzeichen über alle (etwa nach wachsender Größe geordneten) endlich vielen (!) Primzahlen zu erstrecken ist, die Teiler der Zahl n sind. Das ist nur eine andere Schreibweise für (52). Dabei schadet es übrigens nichts, wenn wir in das Produkt auf der rechten Seite von (56) noch einige Primzahlen aufnehmen, die keine Teiler von n sind; denn für diese ist nach (55) $\exp_p(n) = 0$ und $p^{\exp_p(n)} = 1$. Für unsere folgenden Betrachtungen ist es sogar zweckmäßig, das *formal unendliche Produkt* $\prod_{p \in P} p^{\exp_p(n)}$ zu betrachten und (56) in der Form

(56′) $\quad n \geqq 1 \Rightarrow n = \prod_{p \in P} p^{\exp_p(n)}$

zu schreiben (wobei also dieses unendliche Produkt als das Produkt der nur endlich vielen von 1 verschiedenen Faktoren der Form $p^{\exp_p(n)}$ $(p \in P)$ definiert ist).

Man zeigt leicht (Übungsaufgaben!), daß *im Fall* $m \neq 0$, $n \neq 0$ *die folgenden* merkwürdigen *Beziehungen gelten:*

(57) $\quad \bigwedge_{p \in P} \exp_p(m \cdot n) = \exp_p(m) + \exp_p(n),$

(58) $\quad m \mid n \Leftrightarrow \bigwedge_{p \in P} \exp_p(m) \leqq \exp_p(n),$

(59) $\quad \bigwedge_{p \in P} \exp_p(m \sqcap n) = \min\{\exp_p(m), \exp_p(n)\},$

(60) $\quad \bigwedge_{p \in P} \exp_p(m \sqcup n) = \max\{\exp_p(m), \exp_p(n)\}.$

Hierdurch werden die Multiplikation und die mit ihrer Hilfe definierte Teilbarkeitsrelation sowie die auf der Teilbarkeitsrelation beruhenden Operationen $\sqcap$ und $\sqcup$ auf die Addition, die mit ihrer Hilfe definierte $\leqq$-Relation sowie die auf der $\leqq$-Relation beruhenden Operationen min und max zurückgeführt.

Offenbar kann man die Beziehungen (59) und (60) auch in der Form

(59′) $\quad m \sqcap n = \prod_{p \in P} p^{\min\{\exp_p(m), \exp_p(n)\}},$

(60′) $\quad m \sqcup n = \prod_{p \in P} p^{\max\{\exp_p(m), \exp_p(n)\}}$

schreiben, und hierdurch wird in der Schule meistens der größte gemeinsame Teiler und das kleinste gemeinsame Vielfache definiert. Diese Definitionen benutzen indes maßgeblich — und das ist wohl zu beachten — den Hauptsatz (52). Setzt man diesen z. B. mit dem Zermeloschen Verfahren als bewiesen voraus, so kann man auf der Grundlage dieser Definitionen und der Sätze (57) und (58) die charakteristischen Eigenschaften (14) und (32) des größten gemeinsamen Teilers bzw. kleinsten gemeinsamen Vielfachen sowie die anderen in diesem Abschnitt behandelten Sätze der Teilbarkeitstheorie beweisen.

Als Beispiel betrachten wir den wichtigen Satz (30): Es sei $k \mid m \cdot n$ und $k \sqcap m = 1$. Dann ist bei beliebigem $p \in P$ nach (58) und (57)

$$\exp_p(k) \leqq \exp_p(m \cdot n) = \exp_p(m) + \exp_p(n)$$

und wegen $k \sqcap m = 1$ nach (59) und (55) $\min\{\exp_p(k), \exp_p(m)\} = 0$. Hieraus folgt, daß $\exp_p(k) \leqq \exp_p(n)$ bei beliebigem $p \in P$ ist; denn im Fall $\exp_p(k) = 0$ gilt dies trivial, während im Fall $\exp_p(k) \neq 0$ wegen $\min\{\exp_p(k), \exp_p(m)\} = 0$ offenbar $\exp_p(m) = 0$ sein muß und dann die behauptete Ungleichung aus $\exp_p(k) \leqq \exp_p(m) + \exp_p(n)$ folgt. Wenn aber bei beliebigem $p \in P$ die Ungleichung $\exp_p(k) \leqq \exp_p(n)$ gilt, ist nach (55) k ein Teiler von n, wie in (30) behauptet wurde. Man beachte, daß dieser Beweis von (30) sich grundlegend von dem auf S. 157 gegebenen Beweis unterscheidet, der maßgeblich (28), also die in (15) gegebene Charakterisierung des größten gemeinsamen Teilers benutzt. Wir empfehlen dem Leser, auch die anderen wesentlichen Grundgesetze der Teilbarkeitstheorie mittels (57) bis (60) zu beweisen.

Die vorangehenden Ausführungen legen die Frage nahe, wie man für eine gegebene natürliche Zahl $n \geqq 1$ die Primzahlpotenzdarstellung (52) effektiv herstellen kann. Das ist nun — um es gleich ganz deutlich zu sagen — nur durch systematisches Probieren möglich. Hierin liegt der große Nachteil z. B. der Bestimmung des größten gemeinsamen Teilers nach (59) gegenüber seiner Berechnung mittels des euklidischen Algorithmus (25). Wir empfehlen dem Leser, den größten gemeinsamen Teiler zweier sehr großer Zahlen, z. B. 10 436 877 und 128 412, einmal nach dem einen und dann nach dem anderen Verfahren zu bestimmen.

Zur Ermittlung der Primzahlpotenzdarstellung einer Zahl n probiert man die Primzahlen der Reihe nach durch, ob sie Teiler von n sind oder nicht und bestimmt auf diese Weise zunächst die kleinste Primzahl p_1, für die $p_1 \mid n$ gilt. Sodann bildet man solange Potenzen $p_1, p_1^2, \ldots$ dieser Primzahl, bis man erstmals zu einer Primzahlpotenz $p_1^{\nu_1+1}$ gelangt, die kein Teiler von n mehr ist, und hat dann in der Zahl ν_1 des Exponenten $\exp_{p_1}(n)$ gefunden. Anschließend

geht man zu der Zahl $n_1 = \dfrac{n}{p_1^{r_1}}$ über und setzt, beginnend mit der auf p_1 folgenden Primzahl und der Zahl n_1 das Verfahren in analoger Weise fort. Es bricht ab, wenn $n_k = \dfrac{n_{k-1}}{p_k^{r_k}}$ gleich 1 wird. Zur Abkürzung des Verfahrens kann man den folgenden Satz benutzen, dessen Beweis dem Leser als Übungsaufgabe überlassen sei:

$$(61) \qquad n > 1 \wedge \neg (n \textit{ Primzahl}) \Rightarrow \bigvee_p (p \textit{ Primzahl} \wedge p \mid n \wedge p^2 \leqq n).$$

Wir wollen dieses Verfahren an einem Beispiel erläutern: Es sei $n = 138875$. Man stellt leicht fest, daß n nicht durch 2 und nicht durch 3 teilbar ist, d. h. $\exp_2(n) = \exp_3(n) = 0$. Dagegen ist $5 \mid n$, und die Betrachtung der Potenzen 5, 5^2, 5^3, 5^4 führt zu $\exp_5(n) = 3$. Sodann wird das Verfahren mit $p = 7$ und $n_1 = \dfrac{n}{5^3} = 1111$ fortgesetzt. Man stellt fest, daß 7 kein Teiler von n_1 ist, während $11 \mid 1111$, jedoch 11^2 kein Teiler von 1111 ist. Folglich ist $\exp_7(n) = 0$, $\exp_{11}(n) = 1$. Schließlich setzen wir $n_2 = \dfrac{1111}{11} = 101$. Hier können wir das Verfahren bereits abbrechen; auf Grund der bisherigen Konstruktion ist nämlich keine Primzahl $p \leqq 11$ ein Teiler von 101, so daß wegen $11^2 = 121 > 101$ nach (61) 101 eine Primzahl sein muß. Damit erhalten wir die folgende Primzahlpotenzdarstellung von n:

$$138875 = 5^3 \cdot 11 \cdot 101.$$

Das geschilderte Verfahren setzt wesentlich voraus, daß man über die Menge der Primzahlen in ihrer natürlichen Anordnung verfügt oder zumindest in der Lage ist, diese beliebig weit fortzusetzen. Letzteres gelingt nun mit Hilfe einer auf den hellenistischen Mathematiker, Geographen und Astronomen Eratosthenes von Kyrene (etwa 275—194 v. u. Z.) zurückgehenden Methode, die man in recht anschaulicher Weise auch *Sieb des Eratosthenes* nennt. Dazu bezeichne n_0 eine beliebig große natürliche Zahl (im folgenden Beispiel ist $n_0 = 120$). Wir denken uns die Zahlen $2, 3, \ldots, n_0$ in ihrer natürlichen Anordnung der Reihe nach aufgeschrieben. Aus dieser Folge streichen wir zunächst alle diejenigen Zahlen n, die echte Vielfache der Zahl 2 sind, d. h. sich als $n = 2 \cdot q$ mit $q > 1$ darstellen lassen. Sodann gehen wir in der verbleibenden Folge zur nächsten Zahl (d. h. zur Zahl 3) über und streichen aus der verbliebenen Folge alle echten Vielfachen dieser Zahl. Im nächsten Schritt gehen wir in der nunmehr verbliebenen Folge wiederum zur nächsten Zahl (d. h. zur Zahl 5) über und streichen alle deren echte Vielfache usw. Im Endergebnis

bleibt die Menge aller Primzahlen p mit $p \leqq n_0$ in ihrer natürlichen Reihenfolge stehen. Wir merken an, daß nach (61) das Verfahren bereits abgebrochen werden kann, wenn man erstmalig zu einer nicht gestrichenen Zahl p mit $p^2 > n_0$ gelangt (also im Fall $n_0 = 120$ schon nach vier Schritten, da man am Beginn des fünften Schrittes die Zahl $p = 11$ findet und $11^2 = 121 > 120$ ist):

~~2~~	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~	11	~~12~~	13	~~14~~	~~15~~
~~16~~	17	~~18~~	19	~~20~~	~~21~~	~~22~~	23	~~24~~	~~25~~	~~26~~	~~27~~	~~28~~	29	~~30~~
31	~~32~~	~~33~~	~~34~~	~~35~~	~~36~~	37	~~38~~	~~39~~	~~40~~	41	~~42~~	43	~~44~~	~~45~~
~~46~~	47	~~48~~	~~49~~	~~50~~	~~51~~	~~52~~	53	~~54~~	~~55~~	~~56~~	~~57~~	~~58~~	59	~~60~~
61	~~62~~	~~63~~	~~64~~	~~65~~	~~66~~	67	~~68~~	~~69~~	~~70~~	71	~~72~~	73	~~74~~	~~75~~
~~76~~	~~77~~	~~78~~	79	~~80~~	~~81~~	~~82~~	83	~~84~~	~~85~~	~~86~~	~~87~~	~~88~~	89	~~90~~
~~91~~	~~92~~	~~93~~	~~94~~	~~95~~	~~96~~	97	~~98~~	~~99~~	~~100~~	101	~~102~~	103	~~104~~	~~105~~
~~106~~	107	~~108~~	109	~~110~~	~~111~~	~~112~~	113	~~114~~	~~115~~	~~116~~	~~117~~	~~118~~	~~119~~	~~120~~

Es sind also 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 sämtliche Primzahlen unterhalb 120.

Man hat nach dieser Methode sehr weitreichende Primzahltafeln aufgestellt (bis 10^7 reicht eine veröffentlichte Tafel von LEHMER, und bei der Accademia d'Italia gibt es eine Tafel bis $3 \cdot 10^9$). Diese Tafeln zeigen, daß die Primzahlen in der Menge der natürlichen Zahlen sehr unregelmäßig verteilt sind. So findet man einerseits außerordentlich lange Intervalle, die frei von Primzahlen sind, und andererseits stellt sich immer wieder einmal der Fall ein, daß zwei unmittelbar aufeinander folgende ungerade Zahlen (wie 41, 43, 101, 103 usw.) Primzahlen sind. Während sich die erste Beobachtung durch den nachfolgenden, überraschend einfach zu beweisenden Satz in sehr allgemeiner Form bestätigen läßt, führt die zweite Beobachtung auf ein bislang ungelöstes berühmtes mathematisches Problem, nämlich auf die Frage, ob die Menge aller derartigen *Primzahlzwillinge* endlich oder unendlich ist.

$$(62) \qquad \bigwedge_{n \in \mathbb{N}} (n \geqq 2 \Rightarrow \bigvee_{x \in \mathbb{N}} \{x+1, \ldots, x+n\} \cap P = \emptyset);$$

in Worten: *Zu jeder natürlichen Zahl $n \geqq 2$ existiert eine natürliche Zahl x, so daß unter den n aufeinander folgenden Zahlen $x+1, \ldots, x+n$ keine Primzahl vorkommt.* Setzen wir nämlich $x = (n+1)! + 1$, so ist — wie man unmittelbar sieht — die Zahl $x + 1$ $(= (n+1)! + 2)$ durch 2, die Zahl $x + 2$ $(= (n+1)! + 3)$ durch 3, ... und schließlich die Zahl

$$x + n \; (= (n+1)! + (n+1))$$

durch $n + 1$ teilbar, also in der Tat keine der Zahlen $x + 1, \ldots, x + n$ eine Primzahl.

In diesem Zusammenhang ist die Frage nach Abschätzungen der Anzahl $\pi(n)$ aller Primzahlen p mit $p \leqq n$ von Interesse. Schon EULER bewies, daß *bei unbeschränkt wachsendem n der Quotient* $\frac{\pi(n)}{n}$, d. h. das Verhältnis dieser Anzahl zur Anzahl aller natürlichen Zahlen k mit $1 \leqq k \leqq n$, *gegen Null strebt*, also die Primzahlen mit wachsendem n im Mittel spärlicher werden, in jedem hinreichend großen Abschnitt der Folge der natürlichen Zahlen die überwiegende Anzahl der Zahlen zusammengesetzt ist. Er konnte jedoch andererseits auch zeigen, daß *mit wachsendem n die Folge der Zahlen* $x_n = \sum_{p \leqq n} \frac{1}{p}$ $\left(\text{ebenso wie die Folge } y_n = \sum_{k=1}^{n} \frac{1}{k}\right)$ *unbeschränkt wächst*, während — wie in der Analysis gezeigt wird — z. B. die Folge $z_n = \sum_{k=1}^{n} \frac{1}{k^2}$ $\left(\text{gegen } \frac{\pi^2}{6}\right)$ konvergiert, also insbesondere beschränkt ist. Hieraus folgt, daß die Primzahlen wesentlich dichter als z. B. die Quadratzahlen liegen. Diese relativ elementar beweisbaren Tatsachen ergeben jedoch erst ein recht ungenaues Bild von der Verteilung der Primzahlen in der Folge der natürlichen Zahlen. Durch eingehendes Studium von großen Primzahltabellen gelangte man bereits zu Beginn des vorigen Jahrhunderts zu der merkwürdigen Feststellung, daß bei sehr großem n der Wert $\pi(n)$ in recht guter Näherung gleich $\frac{n}{\ln n}$ ist, aber viele berühmte Mathematiker dieser Zeit, wie GAUSS, LEGENDRE u. a. bemühten sich vergeblich, diese empirisch gefundene Tatsache allgemein zu beweisen. Erst dem bedeutenden russischen Mathematiker P. L. TSCHEBYSCHEFF (1821—1894) gelangen hierbei die ersten bedeutsamen Erfolge. Er zeigte u. a., daß *für alle hinreichend großen natürlichen Zahlen n die Ungleichung*

$$\ln 2 < \frac{\pi(n) \cdot \ln n}{n} < 2 \cdot \ln 2$$

gilt. Im Jahre 1894 konnten die französischen Mathematiker J. HADAMARD (1865 bis 1963) und C. DE LA VALLÉE-POUSSIN (1866—1962) mit tiefliegenden Hilfsmitteln der Analysis den ersten vollständigen Beweis des sogenannten Primzahlsatzes erbringen, der sich am einfachsten in der Grenzwertbeziehung

$$\lim_{n \to \infty} \frac{\pi(n) \cdot \ln n}{n} = 1$$

ausdrücken läßt. Im Jahre 1948 konnten P. ERDÖS und A. SELBERG den Beweisgang derart vereinfachen, daß in ihm nur noch einfache Tatsachen aus der reellen Analysis benötigt werden. Die weiteren Untersuchungen in dieser Richtung beschäftigen sich bis in unsere Tage mit dem Problem, die Differenz $1 - \frac{\pi(n) \cdot \ln n}{n}$ genauer abzuschätzen.

Ein weiteres wichtiges Kapitel arithmetischer Forschung war und ist mit der Frage nach der Anzahl der Primzahlen in gewissen anderen Teilmengen der Menge der natürlichen Zahlen verknüpft. Das wohl berühmteste Resultat in dieser Richtung ist der im Jahre 1840 von DIRICHLET ebenfalls mit analytischen Hilfsmitteln bewiesene Satz, daß *es in jeder „arithmetischen Progression" kn + l* $(n = 0, 1, 2, \ldots)$

mit teilerfremdem k und l, wie z. B. in der Folge 9, 14, 19, 24, ... ($k = 5$, $l = 9$), *unendlich viele Primzahlen gibt*. Die naheliegende Verallgemeinerung dieser Frage auf Zahlenmengen z. B. der Form $\{k_1 n^2 + k_2 n + k_3 : n = 0, 1, 2, \ldots\}$ führt sofort auf zahllose ungelöste Probleme; so ist bis heute unbekannt, ob die Menge

$$\{n^2 + 1 : n = 0, 1, 2, \ldots\},$$

d. h. die Folge 1, 2, 5, 10, 17, ..., endlich oder unendlich viele Primzahlen enthält.

In einem Brief an EULER aus dem Jahre 1742 warf GOLDBACH (1690—1764) das Problem auf, ob man jede natürliche Zahl $n \geqq 6$ als Summe von drei Primzahlen darstellen kann, wie er gleichfalls rein empirisch bemerkt hatte. Diese Behauptung ist offenbar bewiesen, wenn man zeigen kann, daß sich jede gerade natürliche Zahl $n \geqq 4$ als Summe von zwei Primzahlen darstellen läßt, eine Behauptung, für die bislang ebenfalls kein Gegenbeispiel bekannt ist und die man heute vielfach (nicht ganz zu Recht) Goldbachsche Vermutung nennt ($4 = 2 + 2$, $6 = 3 + 3$, $8 = 5 + 3$, ...). Die ersten Erfolge in Richtung eines Beweises dieser Vermutung konnten jedoch erst im Jahre 1930 von dem damals ganz jungen sowjetischen Mathematiker I. G. SCHNIRELMANN (1905—1938) erreicht werden, der mit relativ elementaren Hilfsmitteln zeigte, daß *es eine Zahl k gibt, so daß jede natürliche Zahl $n \geqq 1$ als Summe von höchstens k Primzahlen dargestellt werden kann*. Diese Zahl k erwies sich zunächst als außerordentlich groß und konnte in den folgenden Jahren bis auf 67 herabgedrückt werden. Im Jahre 1936 bewies der sowjetische Mathematiker I. M. WINOGRADOW mit neuartigen analytischen Hilfsmitteln, daß *sich jede hinreichend große ungerade Zahl als Summe von drei Primzahlen und damit jede hinreichend große gerade Zahl als Summe von vier Primzahlen darstellen läßt*. Die ursprüngliche und die verschärfte Goldbachsche Vermutung für gerade Zahlen harren dagegen noch immer ihrer Lösung.

Zum Abschluß wollen wir noch kurz eine wichtige auf GAUSS zurückgehende arithmetische Methode behandeln, die beim Beweis zahlreicher elementarer Sätze der Zahlentheorie wertvolle Dienste leistet. Wir knüpfen dazu an den Satz (11) über die Division mit Rest an, nach dem man bei gegebener natürlicher Zahl $m (\geqq 2)$ jede natürliche Zahl n in eindeutiger Weise in der Form

$$n = q\,m + r \quad \text{mit} \quad 0 \leqq r < m$$

darstellen kann, wobei wir den Rest r auch ausführlich mit $\mathfrak{r}(n, m)$ bezeichnen. Sind nun n_1, n_2 beliebige natürliche Zahlen, für die $\mathfrak{r}(n_1, m) = \mathfrak{r}(n_2, m)$ gilt, so sagt man, daß die Zahlen n_1, n_2 *kongruent modulo m* sind, und schreibt dafür $n_1 \equiv n_2 \bmod m$ (vielfach auch kurz $n_1 \equiv n_2\,(m)$):

(63) $$n_1 \equiv n_2 \bmod m : \Leftrightarrow \mathfrak{r}(n_1, m) = \mathfrak{r}(n_2, m).$$

Bei gegebenem $m (\geqq 2)$ wird durch (63) eine zweistellige Relation R_m in $\mathbf{N}$ definiert, von der man sofort feststellt, daß sie eine Äquivalenzrelation in $\mathbf{N}$ ist (vgl. 2.5.(12)). Setzen wir o. B. d. A. voraus, daß $n_1 \geqq n_2$ ist, so gilt (Beweis!)

(64) $$n_1 \equiv n_2 \bmod m \Leftrightarrow m \mid n_1 - n_2.$$

Insbesondere folgt aus (64)

(65) $n \equiv 0 \bmod m \Leftrightarrow m \mid n$.

Ferner erhält man mittels (64)

(66) $n_1 \equiv n_2 \bmod m \Rightarrow n_1 + k \equiv n_2 + k \bmod m$,

und zweimalige Anwendung von (66) liefert

(66′) $n_1 \equiv n_2 \bmod m \wedge k_1 \equiv k_2 \bmod m \Rightarrow n_1 + k_1 \equiv n_2 + k_2 \bmod m$.

Analog ergibt sich

(67) $n_1 \equiv n_2 \bmod m \Rightarrow n_1 \cdot k \equiv n_2 \cdot k \bmod m$

und

(67′) $n_1 \equiv n_2 \bmod m \wedge k_1 \equiv k_2 \bmod m \Rightarrow n_1 \cdot k_1 \equiv n_2 \cdot k_2 \bmod m$

sowie

(67″) $n_1 \equiv n_2 \bmod m \Rightarrow n_1^k \equiv n_2^k \bmod n$.

Da für $\boldsymbol{r}(n, m)$ genau die Werte $0, 1, \ldots, m-1$ möglich sind, folgt aus (63) sofort, daß *jede natürliche Zahl n genau einer dieser Zahlen kongruent modulo m ist:*

(68) $$\bigwedge_{n \in \mathbf{N}} \Big(\bigvee_{i \in \mathbf{N}} (0 \leqq i \leqq m-1 \wedge n \equiv i \bmod m) \wedge \neg \bigvee_{i,j \in \mathbf{N}} (0 \leqq i < j \leqq m-1 \wedge n \equiv i \bmod m \wedge n \equiv j \bmod m) \Big).$$

Wir wollen die Wirksamkeit der bisher bewiesenen Sätze an zwei einfachen Beispielen erläutern. Als erstes wollen wir zeigen, daß keine Zahl z. B. der Form $7k + 3$ eine Quadratzahl sein kann. Dazu merken wir zunächst an, daß jede Zahl n der Form $7k + 3$ der Bedingung $n \equiv 3 \bmod 7$ genügt. Wäre nun n eine Quadratzahl, etwa $n = q^2$ mit $q \in \mathbf{N}$, so müßte q nach (68) genau eine der Bedingungen

$$q \equiv 0 \bmod 7, \quad q \equiv 1 \bmod 7, \quad q \equiv 2 \bmod 7, \quad q \equiv 3 \bmod 7,$$
$$q \equiv 4 \bmod 7, \quad q \equiv 5 \bmod 7, \quad q \equiv 6 \bmod 7$$

erfüllen. Dann wäre aber nach (67″) bzw. (63)

$$q^2 \equiv 0 \bmod 7, \quad q^2 \equiv 1 \bmod 7, \quad q^2 \equiv 4 \bmod 7, \quad q^2 \equiv 2 \bmod 7,$$
$$q^2 \equiv 2 \bmod 7, \quad q^2 \equiv 4 \bmod 7, \quad q^2 \equiv 1 \bmod 7,$$

im Widerspruch zu $n \equiv 3 \bmod 7$.

Zur Erläuterung des nächsten Beispiels machen wir zunächst eine historische Vorbemerkung: Im Jahre 1796 bewies Gauss im Alter von 19 Jahren, *daß sich das regelmäßige n-Eck genau dann mit Zirkel und Lineal konstruieren läßt, wenn*

die Zahl n die Form

$$n = 2^m \cdot p_1 \cdots p_k$$

hat, wobei m eine beliebige natürliche Zahl ist und $p_1, \ldots, p_k$ paarweise verschiedene Primzahlen der Form $2^r + 1$ mit $r \geqq 1$ sind.

Es ergibt sich daher die naheliegende Frage, für welche Werte von r die Zahl $2^r + 1$ eine Primzahl ist. Zunächst sieht man leicht, *daß $2^r + 1$ mit $r \geqq 1$ nur dann eine Primzahl sein kann, wenn r eine Potenz von 2 ist.* Denn im Fall $r = k \cdot u$, u ungerade, wird

$$2^r + 1 = (2^k + 1)\,(2^{k(u-1)} - 2^{k(u-2)} + \cdots + 2^{2k} - 2^k + 1).$$

Also reduziert sich unsere Frage auf das Problem, für welche Werte von s die Zahl $2^{2^s} + 1$ eine Primzahl ist. Für $s = 0, 1, 2, 3, 4$ erhält man der Reihe nach die Primzahlen 3, 5, 17, 257, 65537. Dies sind auch die einzigen bis heute bekannten Primzahlen der Form $2^{2^s} + 1$, die auch Gaußsche oder Fermatsche Primzahlen genannt werden. Man weiß darüber hinaus lediglich noch, daß sich für $s = 5, 6, 7, 8, 9, 11, 12, 18, 23, 36, 38, 73$ mit Sicherheit keine Primzahlen ergeben. Für $s = 5$ (und ähnlich übrigens für die anderen genannten Werte) ergibt sich das folgendermaßen: Wegen $5 \cdot 2^7 = 640$ ist $5 \cdot 2^7 \equiv 641 - 1 \bmod 641$ und mithin nach (67″) $5^2 \cdot 2^{14} \equiv (641-1)^2 \bmod 641$. Andererseits ist

$$(641-1)^2 = 641^2 - 2 \cdot 641 + 1 \equiv 1 \bmod 641,$$

also $5^2 \cdot 2^{14} \equiv 1 \bmod 641$ und damit $5^4 \cdot 2^{28} \equiv 1 \bmod 641$. Folglich ist

$$2^{32} + 1 \equiv 2^{32} + 5^4 \cdot 2^{28} \bmod 641$$

und damit $2^{32} + 1 \equiv 2^{28}\,(2^4 + 5^4) \bmod 641$. Nun ist aber $2^4 + 5^4 = 641$, d. h. $2^4 + 5^4 \equiv 0 \bmod 641$, und hieraus folgt $2^{32} + 1 \equiv 0 \bmod 641$, d. h. $641 \mid 2^{32} + 1$, so daß in der Tat $2^{32} + 1$ keine Primzahl ist.

Nach dem oben genannten Resultat von Gauss lassen sich von den regelmäßigen Vielecken mit einer Eckenzahl unter 100 genau die mit folgenden Eckenzahlen mit Zirkel und Lineal konstruieren:

3, 4, 5, 6, 8, 10, 12, 15, 16, 17, 20, 24, 30, 32, 34, 40, 48, 51, 60, 64, 68, 80, 85, 96.

Von Gauss stammt übrigens auch die erste Konstruktion des regelmäßigen 17-Ecks mit Zirkel und Lineal.

Ist i eine der Zahlen $0, 1, \ldots, m-1$, so bezeichnen wir mit $K_i^{(m)}$ die Menge aller natürlichen Zahlen n, die bei Division durch m den Rest i lassen, d. h., die modulo m zur Zahl i kongruent sind:

$$(69) \qquad K_i^{(m)} := \{n : n \equiv i \bmod m\}\ (= \{n : \mathfrak{r}(n, m) = i\})\ (i = 0, \ldots, m-1).$$

Es ist unmittelbar klar, daß die Mengen $K_i^{(m)}$ gerade die Restklassen der Äquivalenzrelation R_m sind (vgl. 2.5.(12)), d. h.

(70) $\quad \mathsf{N}/R_m = \{K_0^{(m)}, K_1^{(m)}, \ldots, K_{m-1}^{(m)}\}.$

Daraus folgt (vgl. 2.5.(15)), daß das System $\{K_0^{(m)}, \ldots, K_{m-1}^{(m)}\}$ eine Zerlegung der Menge N bildet, was natürlich nur eine andere Formulierung für (68) ist.

Wir merken an, daß die Sätze (66′) und (67′) gerade besagen, daß die Relation R_m sogar eine Kongruenzrelation (vgl. 2.7.(10)) der Struktur $\mathcal{Z} = (\mathsf{N}, +, \cdot)$ ist. Daraus folgt, daß man die Restklassenstruktur $\tilde{\mathcal{Z}} = \mathcal{Z}/R_m = (\mathsf{N}/R_m, \tilde{+}, \tilde{\cdot})$ bilden kann, wobei $\tilde{+}$ und $\tilde{\cdot}$ durch repräsentantenweise Addition bzw. Multiplikation definiert sind. Im vorliegenden Fall wird allgemein

$$K_{i_1}^{(m)} \tilde{+} K_{i_2}^{(m)} = K_j^{(m)} \Leftrightarrow r(i_1 + i_2, m) = j \quad (0 \leqq i_1, i_2 \leqq m-1),$$
$$K_{i_1}^{(m)} \tilde{\cdot} K_{i_2}^{(m)} = K_j^{(m)} \Leftrightarrow r(i_1 \cdot i_2, m) = j \quad (0 \leqq i_1, i_2 \leqq m-1).$$

Die Abbildung $\tilde{f}$, die einer beliebigen natürlichen Zahl n diejenige Restklasse $K_i^{(m)}$ zuordnet, der n angehört, ist nach 2.7.(8) ein Homomorphismus von $\mathcal{Z}$ auf $\mathcal{Z}/R_m$. Die Struktur $\mathcal{Z}/R_m$ heißt der *Restklassenring modulo m*. Im Fall $m = 6$ werden z. B. die Addition $\tilde{+}$ und die Multiplikation $\tilde{\cdot}$ durch folgende Tabellen gegeben, wobei wir statt $K_i^{(6)}$ kurz i geschrieben haben:

$\tilde{+}$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$\tilde{\cdot}$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

Mittels (64) erhält man sofort, daß die Implikation (66) umkehrbar ist, d. h., daß *man aus* $n_1 + k \equiv n_2 + k \bmod m$ *stets auf* $n_1 \equiv n_2 \bmod m$ *schließen kann:*

(71) $\quad n_1 + k \equiv n_2 + k \bmod m \Rightarrow n_1 \equiv n_2 \bmod m\,.$

Wir fragen nun, wann Analoges bei (67) möglich ist. Unsere Behauptung ist, daß *man aus* $n_1 \cdot k \equiv n_2 \cdot k \bmod m$ *genau dann bei beliebigem* n_1, n_2 *auf* $n_1 \equiv n_2 \bmod m$ *schließen kann, wenn die Zahl k zum Modul m teilerfremd ist:*

(72) $\quad \bigwedge_{n_1, n_2} (n_1 k \equiv n_2 k \bmod m \Rightarrow n_1 \equiv n_2 \bmod m) \Leftrightarrow k \sqcap m = 1.$

Wir nehmen zunächst an, es sei $k \sqcap m = 1$ und es gelte $n_1 k \equiv n_2 k \bmod m$ mit o. B. d. A. $n_1 k \geqq n_2 k$. Dann ist nach (64) $m \mid k(n_1 - n_2)$, also nach (30) $m \mid n_1 - n_2$ und mithin $n_1 \equiv n_2 \bmod m$. Ist dagegen $k \sqcap m = d > 1$ und $k = d\,q$, so wählen wir $n_1, n_2 \in \mathsf{N}$ mit $n_1 > n_2$ und $(n_1 - n_2)\,d = m$. Dann ist $0 < n_1 - n_2 < m$, also m kein Teiler von $n_1 - n_2$, d. h. $n_1 \not\equiv n_2 \bmod m$. Andererseits ist

$$n_1 k - n_2 k = (n_1 - n_2)\,k = (n_1 - n_2)\,dq = m\,q, \text{ d. h. } m \mid n_1 k - n_2 k$$

und mithin $n_1 k \equiv n_2 k \bmod m$.

Aus (72) folgt speziell:

(72′) $\quad p$ *Primzahl* $\wedge\ k \not\equiv 0 \bmod p \wedge kn_1 \equiv kn_2 \bmod p \Rightarrow n_1 \equiv n_2 \bmod p.$

Ist der Modul m eine Primzahl p, so kann man also aus einer Kongruenz $kn_1 \equiv kn_2 \bmod p$ generell den Faktor k kürzen, wenn er „von Null verschieden ist" (d. h. $k \not\equiv 0 \bmod p$ gilt).

Eine Menge $\{n_0, \ldots, n_{m-1}\}$ aus m Zahlen, die aus jeder Restklasse $K_i^{(m)}$ (genau) einen Repräsentanten enthält, heißt ein *vollständiges Restsystem modulo m* (v. R. mod m). Zum Beispiel ist also $\{0, 4, 8, 15, 25, 29\}$ ein v. R. mod 6. Für manche Anwendungen ist der folgende Satz von Bedeutung:

(73) $\{n_0, \ldots, n_{m-1}\}$ *v. R. mod* $m \wedge k \sqcap m = 1 \wedge n \in \mathbf{N}$
$\Rightarrow \{k\,n_0 + n, \ldots, k\,n_{m-1} + n\}$ *v. R. mod* m.

Zum Beweis von (73) genügt es offenbar zu zeigen (warum?), daß die Zahlen $kn_i + n$ $(i = 0, \ldots, m-1)$ paarweise modulo m inkongruent sind. Dazu nehmen wir an, es sei $kn_{i_1} + n \equiv kn_{i_2} + n \bmod m$. Dann ist nach (71) $kn_{i_1} \equiv kn_{i_2} \bmod m$ und nach (72) $n_{i_1} \equiv n_{i_2} \bmod m$, also $n_{i_1} = n_{i_2}$, da $\{n_0, \ldots, n_{m-1}\}$ ein v. R. mod m ist. Da $\{0, 4, 8, 15, 25, 29\}$ ein v. R. mod 6 ist, ist nach (73) (mit $k = 7$, $n = 3$) auch $\{3, 31, 59, 108, 170, 206\}$ ein v. R. mod 6.

Eine einfache Folgerung aus Satz (26) ist

(74) $n_0 \in K_i^{(m)} \wedge n_0 \sqcap m = d \Rightarrow \bigwedge_n (n \in K_i^{(m)} \Rightarrow n \sqcap m = d)$,

also speziell

(74′) $n_0 \in K_i^{(m)} \wedge n_0 \sqcap m = 1 \Rightarrow \bigwedge_n (n \in K_i^{(m)} \Rightarrow n \sqcap m = 1)$;

in Worten: *Ist wenigstens ein Repräsentant n_0 einer Restklasse $K_i^{(m)}$ zum Modul m teilerfremd, so gilt das für alle Zahlen aus $K_i^{(m)}$.* Eine Restklasse $K_i^{(m)}$, deren sämtliche Elemente zum Modul m teilerfremd sind, heißt eine *prime Restklasse modulo m*. Nach (74) gilt dann

(75) $K_i^{(m)}$ *prime Restklasse modulo* $m \Leftrightarrow m \sqcap i = 1$.

Eine Menge $\{n_1, \ldots, n_s\}$ von natürlichen Zahlen, die aus jeder primen Restklasse modulo m genau ein Element enthält, wird ein *primes Restsystem modulo m* (p. R. mod m) genannt. Die Anzahl s stimmt offenbar mit der Anzahl der natürlichen Zahlen $x \leqq m$ überein, die zu m teilerfremd sind. Diese Anzahl wird üblicherweise mit $\varphi(m)$ bezeichnet, und die Abbildung φ heißt die *Eulersche Funktion*:

(76) $\varphi(m) := |\{x : x \in \mathbf{N} \wedge x \leqq m \wedge x \sqcap m = 1\}|$.

Zunächst ist unmittelbar klar, daß folgendes gilt:

(77) p *Primzahl* $\Rightarrow \varphi(p) = p - 1$

und allgemeiner (Beweis!)

(77′) p *Primzahl* $\Rightarrow \varphi(p^k) = p^k - p^{k-1}$.

Wir wollen als nächstes den folgenden Satz von Euler beweisen:

(78) $k \sqcap m = 1 \Rightarrow k^{\varphi(m)} \equiv 1 \bmod m$.

Ist speziell m eine Primzahl, so geht (78) wegen (77) in den folgenden Satz von Fermat über:

(78′) p *Primzahl* $\wedge \neg p \mid k \Rightarrow k^{p-1} \equiv 1 \bmod p$.

Zum Beweis von (78) zeigen wir zunächst, daß in Analogie zu (73) folgendes gilt:

(79) $\{n_1, \ldots, n_{\varphi(m)}\}$ p. R. *mod* $m \wedge k \sqcap m = 1$
$\Rightarrow \{kn_1, \ldots, kn_{\varphi(m)}\}$ p. R. *mod* m.

Da die Zahlen $kn_1, \ldots, kn_{\varphi(m)}$ wegen (72) paarweise inkongruent modulo m sind, brauchen wir nur zu zeigen, daß sie sämtlich zu m teilerfremd sind. Das ist jedoch unmittelbar klar.

Eine einfache Folgerung aus (79) ist, daß jede der Zahlen $kn_1, \ldots, kn_{\varphi(m)}$ des primen Restsystems $\{kn_1, \ldots, kn_{\varphi(m)}\}$ genau einer Zahl $n_{i_1}, \ldots, n_{i_{\varphi(m)}}$ des primen Restsystems $\{n_1, \ldots, n_{\varphi(m)}\}$ modulo m kongruent ist:

(*) $kn_1 \equiv n_{i_1} \bmod m, \ldots, kn_{\varphi(m)} \equiv n_{i_{\varphi(m)}} \bmod m,$

wobei $\{n_{i_1}, \ldots, n_{i_{\varphi(m)}}\} = \{n_1, \ldots, n_{\varphi(m)}\}$ ist. Aus (*) folgt nach (67′)

$$k^{\varphi(m)} \cdot n_1 \cdots n_{\varphi(m)} \equiv n_1 \cdots n_{\varphi(m)} \bmod m.$$

Hierbei sind nun $n_1, \ldots, n_{\varphi(m)}$ sämtlich zu m teilerfremd, und damit folgt aus (72) die Behauptung von (78).

Als letztes wollen wir zeigen, daß *die Eulersche Funktion* φ im folgenden Sinne *multiplikativ ist*;

(80) $m_1 \sqcap m_2 = 1 \Rightarrow \varphi(m_1 \cdot m_2) = \varphi(m_1) \cdot \varphi(m_2)$.

Hiernach ergibt sich bei beliebigem $m \in \mathbf{N}$ mit $m \geqq 2$ der Wert $\varphi(m)$, d. h. die Anzahl der zu m teilerfremden Zahlen $x \leqq m$ aus der Primzahlpotenzdarstellung $m = \prod_{p \mid m} p^{\exp_p(m)}$ von m, zu

$$\varphi(m) = \prod_{p \mid m} \varphi(p^{\exp_p(m)}) = \prod_{p \mid m} (p^{\exp_p(m)} - p^{\exp_p(m)-1})$$

$$= \prod_{p \mid m} p^{\exp_p(m)} \left(1 - \frac{1}{p}\right) = m \cdot \prod_{p \mid m} \left(1 - \frac{1}{p}\right),$$

so daß z. B.

$$\varphi(120) = 120 \cdot \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right) = 32$$

wird. Hiermit folgt aus dem Eulerschen Satz, daß für jede zu 120 teilerfremde Zahl k (z. B. $k = 100793$) die Beziehung $k^{32} \equiv 1 \bmod 120$ gilt.

Zum Beweis von (80) seien m_1 und m_2 beliebige teilerfremde natürliche Zahlen. Wir betrachten die Anordnung der Zahlen von 1 bis $m_1 \cdot m_2$ in folgender Tabelle:

(81)

1	2	...	k	...	m_2
$m_2 + 1$	$m_2 + 2$	...	$m_2 + k$	...	$2m_2$
⋮					
$(m_1 - 1)\, m_2 + 1$	$(m_1 - 1)\, m_2 + 2$	...	$(m_1 - 1)\, m_2 + k$	...	$m_1 \cdot m_2$

Offenbar ist $\varphi(m_1 \cdot m_2)$ die Anzahl aller derjenigen Zahlen n der Tabelle, die zu $m_1 \cdot m_2$ teilerfremd sind. Wegen $m_1 \sqcap m_2 = 1$ ist $n \sqcap (m_1 \cdot m_2) = 1$ genau dann, wenn $n \sqcap m_1 = 1$ und $n \sqcap m_2 = 1$ ist. Wir bestimmen nun zunächst diejenigen Zahlen aus (81), die zu m_2 teilerfremd sind. Dazu beachten wir, daß folgendes gilt:

a) In jeder Zeile von (81) steht nach (73) ein v. R. mod m_2,

b) Alle Zahlen einer Spalte von (81) liegen in derselben Restklasse modulo m_2.

c) In jeder Spalte von (81) steht nach (73) ein v. R. mod m_1.

Aus a) und b) folgt nach (74'), daß genau $\varphi(m_2) \cdot m_1$ Zahlen der Tabelle (81) zu m_2 teilerfremd sind, die auf $\varphi(m_2)$ volle Spalten von (81) verteilt sind. Unter diesen sind nun diejenigen auszuwählen, die zusätzlich zu m_1 teilerfremd sind. Dazu brauchen wir jedoch nur zu beachten, daß es wegen c) in jeder Spalte genau $\varphi(m_1)$ Zahlen gibt, die zu m_1 teilerfremd sind. Also enthält die Tabelle (81) in der Tat $\varphi(m_1) \cdot \varphi(m_2)$ Zahlen, die sowohl zu m_1 als auch m_2 teilerfremd sind, was zu zeigen war.

3.8. Die systematische Darstellung der natürlichen Zahlen

Die historische Entwicklung des Zahlbegriffs war eng mit der Entwicklung der Zahlbezeichnungen, und zwar sowohl der sprachlichen Benennungen als auch der Zahlnotierungen verbunden, die uns zugleich die wichtigsten Rückschlüsse auf den jeweiligen Entwicklungsstand des Zahlbegriffs gestatten, denn in der Überlieferung sind uns vor allem die Zahlbezeichnungen überkommen. An ein Bezeichnungssystem für die natürlichen Zahlen wird man vor allem die Forderung stellen, daß sich in ihm jede natürliche Zahl n in eindeutiger Weise bezeichnen läßt. Das einfachste Verfahren besteht offenbar darin, daß man mit Hilfe eines einzigen Grundzeichens | (Strich) jede natürliche Zahl $n \geqq 1$ als Folge von n Strichen schreibt. Diese Art der Bezeichnung hat natürlich den großen Nachteil, daß die Niederschrift einigermaßen großer Zahlen sehr lang und unübersichtlich wird, und sie ist daher für die Praxis kaum brauchbar.

Die für das praktische Rechnen vollkommenste Schreibweise der Zahlen beruht auf einem Prinzip, das u. a. auch unserer üblichen dezimalen Zifferndarstellung der natürlichen Zahlen zugrunde liegt. Bei der dezimalen oder dekadischen Zifferndarstellung werden zunächst die Zahlen von 0 bis 9 durch individuelle Zahlzeichen (Ziffern) **0, 1, ..., 9** bezeichnet, wobei wir im vorliegenden Abschnitt zur deutlichen begrifflichen Unterscheidung die *Ziffern* (= *Zahlzeichen*) in Fettdruck wiedergeben wollen. In diesem Sinne meinen wir also mit 8 die Zahl Acht, mit **8** aber die diese Zahl bezeichnende Ziffer; ist allgemein a eine der Zahlen 0, ..., 9 so bedeutet $\boldsymbol{a}$ die zu dieser Zahl gehörige Ziffer. Während die Zahlen Objekte begrifflicher Natur sind, sind die zugehörigen Zahlzeichen oder Ziffern letztlich geometrische Figuren einer

bestimmten Gestalt. Von der Zahl (aber nicht der Ziffer) 8 können wir z. B. behaupten, daß sie kleiner als die Zahl 9 ist, von der Ziffer (aber nicht der Zahl) 8, daß sie auf dieser Seite dieses Buches fünfzehnmal gedruckt ist (und zwar achtmal mager und siebenmal fett).

Jede natürliche Zahl n wird anschließend durch eine eindeutig bestimmte endliche Folge aus Ziffern $0, 1, \ldots, 9$ bezeichnet, wobei man sich eines sogenannten *Stellenwert-* oder *Positionsprinzips* bedient. Grundlage hierfür ist der anschließend (in allgemeinerer Form) zu beweisende Satz, daß *man jede natürliche Zahl* $n \geqq 1$ *in der Form*

$$(1) \qquad n = a_k \cdot 10^k + a_{k-1} \cdot 10^{k-1} + \cdots + a_1 \cdot 10 + a_0$$

darstellen kann, wobei der Exponent k *und die Koeffizienten* $a_\varkappa$ $(\varkappa = 0, 1, \ldots, k)$ *durch* n *eindeutig bestimmte natürliche Zahlen sind, die folgenden Bedingungen genügen:* $k \geqq 0$, $0 \leqq a_\varkappa \leqq 9$ $(\varkappa = 0, \ldots, k)$, $a_k \neq 0$. Umgekehrt liefert natürlich jede Summe (1) mit den genannten Bedingungen eine bestimmte natürliche Zahl $n \geqq 1$ (die Zahl 0 kann man dabei formal durch die leere Summe darstellen). Die Darstellung einer Zahl n in der Form (1) wollen wir auch als *Dezimaldarstellung* von n bezeichnen. Wird nun die Zahl n durch die Dezimaldarstellung (1) gegeben, so verwendet man zur Bezeichnung der Zahl n die Ziffernfolge (das „Wort") $a_k\, a_{k-1} \ldots a_1\, a_0$, was wir auch durch

$$(2) \qquad n \mathrel{\hat{=}}_{10} a_k\, a_{k-1} \ldots a_1\, a_0$$

(gelesen etwa: die Zahl n wird dezimal durch die Ziffernfolge $a_k\, a_{k-1} \ldots a_1\, a_0$ bezeichnet) wiedergeben wollen. Die Ziffernfolge $a_k\, a_{k-1} \ldots a_1\, a_0$ wollen wir dabei die *dezimale Zifferndarstellung von* n nennen. Die bislang ausgeschlossene Zahl 0 soll durch die Ziffer **0** bezeichnet werden. Wir merken an, daß die Zifferndarstellung jeder natürlichen Zahl $n \geqq 1$ mit einer von der Ziffer **0** verschiedenen Ziffer beginnt.

Die Bedeutung einer Ziffer in der (dezimalen) Zifferndarstellung einer natürlichen Zahl n hängt also außer von ihrer Gestalt maßgeblich von der Stelle (Position) ab, an der sie in der Zifferndarstellung auftritt. So hat in der dezimalen Zifferndarstellung 888 der Zahl 888 die erste (am weitesten links stehende) Ziffer 8 die Bedeutung von 8 Hunderter, die mittlere Ziffer 8 die Bedeutung von 8 Zehner und die dritte Ziffer 8 die Bedeutung von 8 Einer. So einfach uns auch heute diese Schreibweise erscheinen mag, ist sie doch das Ergebnis einer langen historischen Entwicklung. Das dezimale Positionssystem ist wahrscheinlich indischen Ursprungs, von wo es im frühen Mittelalter über den vorderen Orient, Nordafrika und Spanien nach Mitteleuropa gelangte. Aber auch hier dauerte es noch einige Jahrhunderte, bis es sich im 16. und 17. Jahrhundert gegen das bis dahin übliche römische Bezeichnungssystem durch-

setzte. Insbesondere war das Positionsprinzip den altgriechischen Mathematikern unbekannt, die ein alphabetisches Bezeichnungsprinzip verwendeten.

Ebenso wie die Zahl 10 kann auch jede andere natürliche Zahl $g \geqq 2$ als *Grundzahl* (*Basis*) für ein Positionssystem genommen werden. Die besonders weite Verbreitung des Dezimalsystems hat wahrscheinlich ihren historischen Ursprung in der Anzahl unserer Finger. Rein sachlich besitzt es keine nennenswerten Vorzüge, die es vor Positionssystemen mit anderer Basis auszeichnen. Die Darstellung der Zahlen mit Hilfe einer beliebigen Basis $g \geqq 2$ wurde zuerst 1654 von B. PASCAL untersucht. Sie beruht auf der zu (1) analogen Tatsache, daß *man bei gegebenem* $g \geqq 2$ ($g \in \mathbb{N}$) *jede natürliche Zahl* $n \geqq 1$ *in der Form*

$$(3) \qquad n = a_k \cdot g^k + a_{k-1} \cdot g^{k-1} + \cdots + a_1 \cdot g + a_0$$

darstellen kann, wobei der Exponent k *und die Koeffizienten* $a_\varkappa$ ($\varkappa = 0, \ldots, k$) *durch* n (*und* g) *eindeutig bestimmte natürliche Zahlen sind, die jetzt den folgenden Bedingungen genügen:* $k \geqq 0$, $0 \leqq a_\varkappa \leqq g - 1$ ($\varkappa = 0, \ldots, k$) *und* $a_k \neq 0$. Wir wollen allgemein (3) die *g-adische Darstellung von n* nennen. In Verallgemeinerung der dezimalen Zifferndarstellung benötigt man zur *g-adischen Zifferndarstellung* der natürlichen Zahlen g Ziffern (Zahlzeichen). Deuten wir die einer natürlichen Zahl a mit $0 \leqq a \leqq g - 1$ entsprechende Ziffer wieder allgemein durch $\boldsymbol{a}$ an, so wird im g-adischen Positionssystem die Zahl n mit der g-adischen Darstellung (3) wieder durch die Ziffernfolge $\boldsymbol{a_k a_{k-1} \ldots a_1 a_0}$ nun aber aus Ziffern $\boldsymbol{a_\varkappa}$ mit $0 \leqq a_\varkappa \leqq g - 1$ bezeichnet, was wir jetzt durch

$$(4) \qquad n \triangleq_g \boldsymbol{a_k a_{k-1} \ldots a_1 a_0}$$

(gelesen etwa: die Zahl n wird g-adisch durch die Ziffernfolge $\boldsymbol{a_k a_{k-1} \ldots a_1 a_0}$ bezeichnet) andeuten wollen. Die Zahl 0 wird wieder durch die ihr zugeordnete Ziffer bezeichnet.

Beim *Dual-* oder *Binärsystem* ($g = 2$) werden also nur zwei Ziffern, z. B. **0** und **1** benötigt (wobei vielfach statt **0** und **1** die Zeichen o und l verwendet werden). Das *Ternärsystem* ($g = 3$) benötigt drei Ziffern **0**, **1** und **2**. Beim *Duodezimalsystem* ($g = 12$) sind 12 Ziffern, etwa **0**, **1**, ..., **9**, **A**, **B** (mit **A** als Ziffer für 10, **B** als Ziffer für 11) erforderlich, usw. Man rechnet leicht nach, daß z. B. folgendes gilt:

$$3179 \triangleq_{10} \mathbf{3179}, \qquad 3179 \triangleq_2 \mathbf{110001101011}$$
$$3179 \triangleq_3 \mathbf{11100202}, \qquad 3179 \triangleq_{12} \mathbf{1A0B}.$$

Ein Positionssystem mit kleiner Grundzahl g hat offenbar den Nachteil, daß die Niederschrift größerer Zahlen sehr lang wird, es benötigt demgegenüber nur wenige Ziffern und hat demgemäß ein kleines Einmaleins. In einem Positionssystem mit großer Grundzahl (im alten Babylonien wurde ein solches mit

der Grundzahl 60 verwendet!) werden die Niederschriften der Zahlen kurz, es benötigt aber viele Ziffern und das für das Rechnen erforderliche Einmaleins wird sehr umfangreich.

Wir wollen nun den Satz (3) beweisen, der ja die Grundlage für die g-adische Zifferndarstellung der natürlichen Zahlen ist. Der Satz (1) ist natürlich nur der Spezialfall $g = 10$ des Satzes (3). Dazu nehmen wir zunächst an, daß uns eine Darstellung der Zahl $n \geqq 1$ in der Form (3) schon bekannt ist. Dann ist offenbar die Zahl a_0 der Rest bei der Division der betrachteten Zahl n durch die Grundzahl $g \geqq 2$ (vgl. 3.7.(11)):

(5_0) $\quad n = q_1 g + a_0 \quad$ mit $\quad 0 \leqq a_0 \leqq g - 1;$

die Zahl a_1 ist sodann der Rest bei der Division des Quotienten $q_1 = \boldsymbol{q}(n, g)$ durch g:

(5_1) $\quad q_1 = q_2 g + a_1 \quad$ mit $\quad 0 \leqq a_1 \leqq g - 1,$

die Zahl a_2 ist der Rest bei der Division des Quotienten $q_2 = \boldsymbol{q}(q_1, g)$ durch g:

(5_2) $\quad q_2 = q_3 g + a_2 \quad$ mit $\quad 0 \leqq a_2 \leqq g - 1$

usw. Dieses Verfahren wird nun solange fortgesetzt, bis erstmalig der Quotient $q_{k+1} = \boldsymbol{q}(q_k, g)$ gleich Null wird (wegen $g \geqq 2$ bilden die Quotienten eine echt monoton fallende Folge von natürlichen Zahlen, so daß dieser Fall nach endlich vielen Schritten eintritt), und der zugehörige Rest $\boldsymbol{r}(q_k, g)$ ist dann der in (3) auftretende Koeffizient a_k:

(5_k) $\quad q_k = 0 \cdot g + a_k \quad$ mit $\quad 0 < a_k \leqq g - 1$

wegen $q_k \neq 0$ ist $a_k \neq 0$, wie in (3) gefordert). Wenn sich die Zahl $n \geqq 1$ also überhaupt in der Form (3) darstellen läßt, müssen sich die Koeffizienten $a_0, a_1, \ldots, a_{k-1}, a_k$ nach folgendem Algorithmus ergeben:

$$
\begin{aligned}
n = q_0 &= q_1 g + a_0 && \text{mit} \quad q_1 \neq 0, \quad 0 \leqq a_0 \leqq g - 1, \\
q_1 &= q_2 g + a_1 && \text{mit} \quad q_2 \neq 0, \quad 0 \leqq a_1 \leqq g - 1, \\
&\cdot \\
(5) \qquad &\cdot \\
&\cdot \\
q_{k-1} &= q_k g + a_{k-1} && \text{mit} \quad q_k \neq 0, \quad 0 \leqq a_{k-1} \leqq g - 1, \\
q_k &= 0 \cdot g + a_k && \text{mit} \quad 0 < a_k \leqq g - 1
\end{aligned}
$$

(man vergleiche dies mit dem Euklidischen Algorithmus 3.7.(25)). Man sieht nun aber sofort, daß mit den nach (5) ermittelten Koeffizienten $a_\varkappa$ $(\varkappa = 0, \ldots, k)$ die Gleichung (3) gilt. Setzt man nämlich (5_k) in

(5_{k-1}) $\quad q_{k-1} = q_k g + a_{k-1}$

ein, so erhält man

$$q_{k-1} = a_k g + a_{k-1};$$

dies in

$$(5_{k-2}) \quad q_{k-2} = q_{k-1} g + a_{k-2}$$

eingesetzt, ergibt

$$q_{k-2} = a_k g^2 + a_{k-1} g + a_{k-2},$$

usw.; und schließlich folgt aus (5_1) und (5_0)

$$q_1 = a_k g^{k-1} + \cdots + a_2 g + a_1$$

sowie

$$n = q_0 = a_k g^k + \cdots + a_1 g + a_0.$$

Wendet man (5) mit der Grundzahl $g = 3$ bzw. $g = 12$ auf die Zahl $n = 3179$ an (vgl. obiges Beispiel), so erhält man:

$$\begin{aligned} 3179 &= 1059 \cdot 3 + 2, \\ 1059 &= 353 \cdot 3 + 0, \\ 352 &= 117 \cdot 3 + 2, \\ 117 &= 39 \cdot 3 + 0, \\ 39 &= 13 \cdot 3 + 0, \\ 13 &= 4 \cdot 3 + 1, \\ 4 &= 1 \cdot 3 + 1, \\ 1 &= 0 \cdot 3 + 1, \end{aligned} \qquad \begin{aligned} 3179 &= 264 \cdot 12 + 11, \\ 264 &= 22 \cdot 12 + 0, \\ 22 &= 1 \cdot 12 + 10, \\ 1 &= 0 \cdot 12 + 1, \end{aligned}$$

so daß in der Tat $3179 \triangleq_3 \mathbf{11100202}$ und $3179 \triangleq_{12} \mathbf{1A0B}$ gilt.

Man zeigt noch leicht (Beweis!), daß *der Exponent k in* (3) *durch die Ungleichungen*

$$(6) \qquad g^k \leqq n < g^{k+1}$$

charakterisiert ist (man sollte sich dabei auch überlegen, daß es bei gegebenem $g \geqq 2$ und $n \geqq 1$ genau eine natürliche Zahl k gibt, für die (6) gilt).

Wir kommen nun zur theoretischen Begründung des Additionsverfahrens für in g-adischer Zifferndarstellung gegebene natürliche Zahlen, das im Prinzip genauso wie das bekannte Additionsverfahren für in dezimaler Zifferndarstellung gegebene Zahlen verläuft. Es seien dazu n_1, n_2 von Null verschiedene natürliche Zahlen mit

$$n_1 \triangleq_g \boldsymbol{a}_{k_1} \ldots \boldsymbol{a}_0, \quad n_2 \triangleq_g \boldsymbol{b}_{k_2} \ldots \boldsymbol{b}_0,$$

d. h., es gelte

$$(7) \qquad n_1 = \sum_{\varkappa=0}^{k_1} a_\varkappa g^\varkappa, \quad n_2 = \sum_{\varkappa=0}^{k_2} b_\varkappa g^\varkappa$$

mit $0 \leqq a_\varkappa \leqq g-1$ $(\varkappa = 0, \ldots, k_1)$, $a_{k_1} \neq 0$, $0 \leqq b_\varkappa \leqq g-1$ $(\varkappa = 0, \ldots, k_2)$, $b_{k_2} \neq 0$. Setzen wir $k = \max\{k_1, k_2\}$ und im Fall $k_1 < k$ bzw. $k_2 < k$ noch $a_{k_1+1} = \cdots = a_k = 0$ bzw. $b_{k_2+1} = \cdots = b_k = 0$, so wird

$$(7') \qquad n_1 = \sum_{\varkappa=0}^{k} a_\varkappa g^\varkappa, \quad n_2 = \sum_{\varkappa=0}^{k} b_\varkappa g^\varkappa$$

mit $0 \leqq a_\varkappa, b_\varkappa \leqq g-1$ $(\varkappa = 0, \ldots, k)$ und $a_k + b_k \neq 0$. Wegen (vgl. (6))

$$g^{k_1} \leqq n_1 < g^{k_1+1}, \quad g^{k_2} \leqq n_2 < g^{k_2+1}$$

wird — wie man leicht zeigt —

$$(8) \qquad g^k \leqq n_1 + n_2 < g^{k+2},$$

so daß der Exponent k^* in der g-adischen Darstellung von $n_1 + n_2$, es sei dies

$$(9) \qquad n_1 + n_2 = \sum_{\varkappa=0}^{k^*} s_\varkappa g^\varkappa$$

$(0 \leqq s_\varkappa \leqq g-1$ $(\varkappa = 0, \ldots, k^*)$, $s_{k^*} \neq 0)$, entweder gleich k oder gleich $k+1$ ist. Die Aufgabe besteht nun darin, aus den Koeffizienten $a_\varkappa$ und $b_\varkappa$ $(\varkappa = 0, \ldots, k)$ die Koeffizienten $s_0, s_1, \ldots, s_k, s_{k+1}$ zu ermitteln, wobei im Fall $k^* = k$ natürlich $s_{k+1} = 0$ gesetzt ist.

Wir behaupten, daß sich die Koeffizienten $s_0, \ldots, s_{k+1}$ aus den Koeffizienten $a_0, \ldots, a_k, b_0, \ldots, b_k$ gemäß

$$(10) \qquad s_\varkappa = \boldsymbol{r}(a_\varkappa + b_\varkappa + \ddot{u}_\varkappa, g) \quad (\varkappa = 0, \ldots, k), \quad s_{k+1} = \ddot{u}_{k+1}$$

berechnen, wobei die *Überträge* $\ddot{u}_0, \ldots, \ddot{u}_{k+1}$ nach den folgenden *Rekursionsgleichungen* gewonnen werden:

$$(11) \qquad \ddot{u}_0 = 0, \quad \ddot{u}_{\varkappa+1} = \boldsymbol{q}(a_\varkappa + b_\varkappa + \ddot{u}_\varkappa, g) \quad (\varkappa = 0, \ldots, k).$$

Man erkennt leicht (Beweis!), daß für $\ddot{u}_\varkappa$ $(\varkappa = 0, \ldots, k+1)$ nur die Werte 0 und 1 in Frage kommen, so daß sich im Fall $g = 2$ die Werte $s_\varkappa$ und $\ddot{u}_{\varkappa+1}$ aus folgender Tabelle ablesen lassen:

$a_\varkappa$	$b_\varkappa$	$\ddot{u}_\varkappa$	$s_\varkappa$	$\ddot{u}_{\varkappa+1}$
0	0	0	0	0
0	1	0	1	0
1	0	0	1	0
1	1	0	0	1
0	0	1	1	0
0	1	1	0	1
1	0	1	0	1
1	1	1	1	1

Die analoge Tabelle für den Fall $g = 3$ hat bereits 18 Zeilen (wir empfehlen dem Leser, sie sich ausführlich aufzuschreiben), die für $g = 10$ hat 200 Zeilen, die für $g = 12$ hat 288 Zeilen, und bei beliebigem g hat sie $2\,g^2$ Zeilen.

Zum Beweis von (10) beachten wir, daß die Koeffizienten $a_0, \ldots, a_k$, $b_0, \ldots, b_k$, $s_0, \ldots, s_{k+1}$ nach (5) durch

$$
\begin{array}{lll}
q_\varkappa = q_{\varkappa+1}\,g + a_\varkappa \ (\varkappa = 0, \ldots, k) & \text{mit } q_0 = n_1, & q_{k+1} = 0, \\
q'_\varkappa = q'_{\varkappa+1}\,g + b_\varkappa \ (\varkappa = 0, \ldots, k) & \text{mit } q'_0 = n_2, & q'_{k+1} = 0, \\
q''_\varkappa = q''_{\varkappa+1}\,g + s_\varkappa \ (\varkappa = 0, \ldots, k+1) & \text{mit } q''_0 = n_1 + n_2, & q''_{k+2} = 0
\end{array} \tag{12}
$$

gegeben werden, wobei $q_{\varkappa+1} = \boldsymbol{q}\,(q_\varkappa, g)$, $a_\varkappa = \boldsymbol{r}\,(q_\varkappa, g)$ usw. gilt. Man zeigt nun leicht, daß bei beliebigem $\varkappa = 0, \ldots, k+1$

$$q''_\varkappa = q_\varkappa + q'_\varkappa + \ddot{u}_\varkappa \tag{13}$$

ist. Im Fall $\varkappa = 0$ gilt (13) trivial. Wenn aber (13) für den Index $\varkappa \leqq k$ schon bewiesen ist, wird nach (12)

$$q''_\varkappa = q_\varkappa + q'_\varkappa + \ddot{u}_\varkappa = (q_{\varkappa+1} + q'_{\varkappa+1})\,g + a_\varkappa + b_\varkappa + \ddot{u}_\varkappa$$

und mithin nach (11) und (12)

$$\begin{aligned} q''_{\varkappa+1} &= \boldsymbol{q}\,(q''_\varkappa, g) = q_{\varkappa+1} + q'_{\varkappa+1} + \boldsymbol{q}\,(a_\varkappa + b_\varkappa + \ddot{u}_\varkappa, g) \\ &= q_{\varkappa+1} + q'_{\varkappa+1} + \ddot{u}_{\varkappa+1}. \end{aligned}$$

Also gilt (13) für alle Indizes $\varkappa = 0, \ldots, k+1$. Aus (13) ist nun leicht (10) zu erhalten. Denn aus (12) und (13) folgt für $\varkappa = 0, \ldots, k$

$$q''_\varkappa = (q_{\varkappa+1} + q'_{\varkappa+1})\,g + a_\varkappa + b_\varkappa + \ddot{u}_\varkappa,$$

so daß in der Tat

$$s_\varkappa = \boldsymbol{r}\,(q''_\varkappa, g) = \boldsymbol{r}\,(a_\varkappa + b_\varkappa + \ddot{u}_\varkappa, g)$$

ist, und für $\varkappa = k+1$ liefert (13)

$$s_{k+1} = q''_{k+1} = q_{k+1} + q'_{k+1} + \ddot{u}_{k+1} = \ddot{u}_{k+1}.$$

Die vorangehenden Überlegungen lassen sich leicht auf Summen aus mehr als zwei Summanden verallgemeinern. Dabei kann natürlich bei hinreichend vielen Summanden der Übertrag $\ddot{u}_\varkappa\,(\varkappa = 1, \ldots, k+1)$ beliebige Werte $0, \ldots, g-1$ annehmen, und bei mehr als g Summanden treten neben dem Übertrag auf die links folgende Stelle auch Überträge auf höhere Stellen auf.

Die Multiplikation von in g-adischer Zifferndarstellung gegebenen Zahlen erfolgt ebenfalls nach einem dem bekannten Multiplikationsverfahren für Zahlen in dezimaler Zifferndarstellung analogen Verfahren. Zur theoretischen Begründung dieses Verfahrens merken wir zunächst an, daß bei beliebigem

$\lambda \geqq 0$ folgendes gilt:

$$(14) \qquad n \triangleq_g a_k \dots a_0 \Rightarrow n^t \cdot g^\lambda \triangleq_g a_k \dots a_0 \underbrace{0 \dots 0}_{\lambda\text{-mal}}.$$

Damit reduziert sich wegen

$$\left(\sum_{\varkappa=0}^{k} a_\varkappa g^\varkappa\right) \cdot \left(\sum_{\lambda=0}^{l} b_\lambda g^\lambda\right) = \sum_{\lambda=0}^{l} \left(b_\lambda \cdot \sum_{\varkappa=0}^{k} a_\varkappa g^\varkappa\right) g^\lambda$$

die Ermittlung der Ziffern der g-adischen Zifferndarstellung von

$$n_1 \cdot n_2 \left(= \left(\sum_{\lambda=0}^{k} a_\varkappa g^\varkappa\right) \cdot \left(\sum_{\lambda=0}^{l} b_\lambda g^\lambda\right)\right)$$

im wesentlichen auf die Aufgabe, bei gegebenem $\gamma = 0, \dots, g-1$ aus der g-adischen Zifferndarstellung von n die von $n \cdot \gamma$ zu ermitteln (wobei die Fälle $\gamma = 0$ und $\gamma = 1$ natürlich trivial sind). Es sei dazu

$$n \triangleq_g a_k \dots a_0, \quad n \cdot \gamma \triangleq_g p_{k^*} \dots p_0.$$

Dann ist im Fall $\gamma \neq 0$ zunächst nach (6) entweder $k^* = k$ oder $k^* = k+1$, und analog zu (10) und (11) gilt (Beweis!):

$$(15) \qquad p_\varkappa = r(a_\varkappa \gamma + \ddot{u}_\varkappa, g) \quad (\varkappa = 0, \dots, k), \quad p_{k+1} = \ddot{u}_{k+1}$$

mit

$$(16) \qquad \ddot{u}_0 = 0, \quad \ddot{u}_{\varkappa+1} = q(a_\varkappa \gamma + \ddot{u}_\varkappa, g) \quad (\varkappa = 0, \dots, k).$$

Für die Überträge $\ddot{u}_\varkappa$ $(\varkappa = 0, \dots, k+1)$ sind dabei nach (16) die Werte $0, \dots, g-2$ möglich (Beweis!). Im Fall $g = 3$ ergeben sich die Werte für $p_\varkappa$ und $\ddot{u}_\varkappa$ nach folgender Tabelle (bei der wir auf die Angabe der Trivialfälle $\gamma = 0$ und $\gamma = 1$ verzichten):

$a_\varkappa$	γ	$\ddot{u}_\varkappa$	$p_\varkappa$	$\ddot{u}_{\varkappa+1}$
0	2	0	0	0
1	2	0	2	0
2	2	0	1	1
0	2	1	1	0
1	2	1	0	1
2	2	1	2	1

Im Fall $g = 10$ hat die analoge Tabelle (ohne $\gamma = 0$ und $\gamma = 1$) 640 Zeilen, und bei beliebigem g sind es $g \cdot (g-2)^2$ Zeilen (der Fall $g = 2$ benötigt keine Tabelle, das Einmaleins reduziert sich auf $1 \cdot 1 = 1$).

Hier ein Beispiel für eine Multiplikation im Ternärsystem:

$$
\begin{array}{r}
2122 \cdot 1212 \\
\hline
2122 \\
12021 \\
\quad 2122 \\
\quad 12021 \\
\ddot{u}\colon\ 112211 \\
\hline
11212111
\end{array}
$$

Auch die Division (mit Rest) verläuft im g-adischen Positionssystem analog wie im Dezimalsystem. Wir verzichten auf eine genaue theoretische Begründung und geben nur ein Beispiel für eine Division im Ternärsystem:

$$
\begin{array}{l}
211211021 : 101 = 2021121 \\
202 \\
\hline
\quad 221 \\
\quad 202 \\
\quad \hline
\quad\ 121 \\
\quad\ 101 \\
\quad \hline
\quad\ 200 \\
\quad\ 101 \\
\quad \hline
\quad\ 222 \\
\quad\ 202 \\
\quad \hline
\quad\ 201 \\
\quad\ 101 \\
\quad \hline
\quad\ 100
\end{array}
$$

d. h., für $n_1 \triangleq_3 211211021$ und $n_2 \triangleq_3 101$ wird $q(n_1, n_2) \triangleq_3 2021121$ und $r(n_1, n_2) \triangleq_3 100$.

Wir merken an, daß sich die Division mit Rest im Dualsystem besonders einfach gestaltet.

Zum Abschluß wollen wir noch das Grundprinzip der bekannten Dreier- und Neunerprobe herausarbeiten. Dazu sei zunächst n eine beliebige natürliche Zahl mit $n \triangleq_{10} a_k \ldots a_0$, d. h. $n = \sum_{\varkappa=0}^{k} a_\varkappa \cdot 10^\varkappa$. Die *dezimale Quersumme* von n wird dann definiert durch

$$
(17) \qquad Q_{10}(n) := \sum_{\varkappa=0}^{k} a_\varkappa .
$$

Aus $10 \equiv 1 \bmod 3$ ($10 \equiv 1 \bmod 9$) folgt nun mittels 3.7.(67'') sofort, daß bei beliebigem $\varkappa$ auch $10^\varkappa \equiv 1 \bmod 3$ ($10^\varkappa \equiv 1 \bmod 9$) ist, und damit erhält man mittels 3.7.(66') und 3.7.(67)

(18) $\quad Q_{10}(n) \equiv n \bmod 3 \quad (Q_{10}(n) \equiv n \bmod 9).$

Insbesondere gilt also auf Grund 3.7.(65)

(19) $\quad 3 \mid n \Leftrightarrow 3 \mid Q_{10}(n), \quad 9 \mid n \Leftrightarrow 9 \mid Q_{10}(n).$

Ist g eine beliebige Grundzahl, $n \triangleq_g a_k \ldots a_0$, so nennen wir analog

(17') $\quad Q_g(n) := \sum_{\varkappa=0}^{k} a_\varkappa$

die *g-adische Quersumme* von n. In Verallgemeinerung von (18) bzw. (19) gilt dann

(18') $\quad g \equiv 1 \bmod k \Rightarrow Q_g(n) \equiv n \bmod k,$

(19') $\quad g \equiv 1 \bmod k \Rightarrow (k \mid n \Leftrightarrow k \mid Q_g(n)).$

Im Fall $g = 11$ tritt also z. B. an die Stelle der Dreier- und Neunerprobe eine Zweier- und Fünferprobe.

Namen- und Sachverzeichnis